UNIT - II

RESEARCH PROBLEM

What is a research problem?

- The term ‘problem’ means a question or issue to be examined.

- Research Problem refers to some difficulty /need which a researcher experiences in the context of either theoretical or practical situation and wants to obtain a solution for the same.

HOW DO WE KNOW WE HAVE A RESEARCH PROBLEM?

- Customer complaints
- Conversation with company employees
- Observation of inappropriate behaviour or conditions in the firm
- Deviation from the business plan
- Success of the firm's competitor's
- Relevant reading of published material (trends, regulations)
- Company records and reports.

The first step in the research process – definition of the problem involves two activities:

☆Identification / Selection of the Problem

☆Formulation of the Problem

PROBLEM IDENTIFICATION

☆ This step involves identification of a few problems and selection of one out of them, after evaluating the alternatives against certain selection criteria.

SOURCES OF PROBLEMS

- Reading
- Academic Experience
- Daily Experience
- Exposure to Field Situations
- Consultations
- Brainstorming
- Research
- Intuition

CRITERIA OF SELECTION

The selection of one appropriate researchable problem out of the identified problems requires evaluation of those alternatives against certain criteria. They are:

- Internal / Personal criteria – Researcher's Interest, Researcher's Competence, Researcher's own Resource: finance and time.

- External Criteria or Factors – Researchability of the problem, Importance and Urgency, Novelty of the Problem, Feasibility, Facilities, Usefulness and Social Relevance, Research Personnel.

DEFINITION / FORMULATION OF THE RESEARCH PROBLEM

- Formulation is the process of refining the research ideas into research questions and objectives.

- Formulation means translating and transforming the selected research problem/topic/idea into a scientifically researchable question. It is concerned with specifying exactly what the research problem is.

- Problem definition or Problem statement is a clear, precise and succinct statement of the question or issue that is to be investigated with the goal of finding an answer or solution.

- There are two ways of stating a problem:

    1) Posting question / questions
    2) Making declarative statement / statements

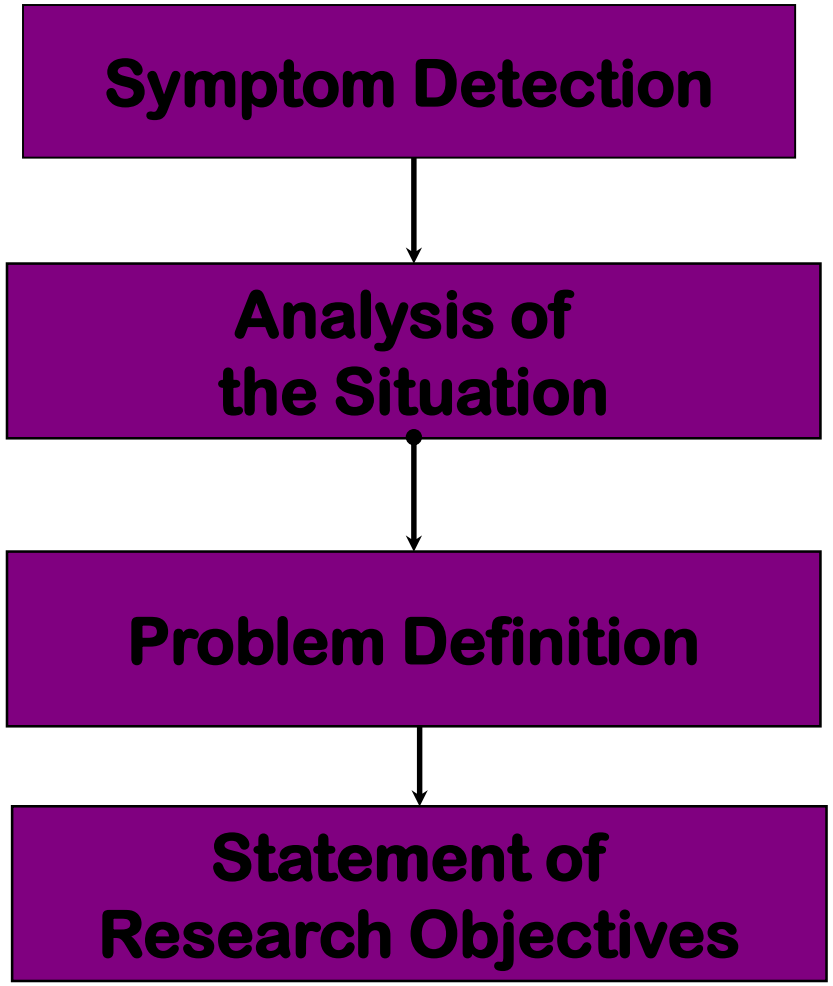
STEPS IN DEFINING THE PROBLEM

- **STATEMENT OF THE PROBLEM IN A GENERAL WAY.**
- **UNDERSTANDING THE NATURE OF PROBLEM**
- **SURVEYING THE AVAILABLE LITERATURE**
- **DEVELOPING IDEAS THROUGH DISCUSSIONS**
- **REPHRASING THE RESEARCH PROBLEM**

CRITERIA OF A GOOD RESEARCH PROBLEM

- Clear and Unambiguous
- Empirical
- Verifiable
- Interesting
- Novel and Original
- Availability of Guidance

Defining Problem, Results in Clear Cut Research Objectives..



REVIEW OF LITERATURE

REVIEW OF LITERATURE

- Literature Review is the documentation of a comprehensive review of the published and unpublished work from secondary sources of data in the areas of specific interest to the researcher.
- The main aim is to find out problems that are already investigated and those that need further investigation.
- It is an extensive survey of all available past studies relevant to the field of investigation.
- It gives us knowledge about what others have found out in the related field of study and how they have done so.

PURPOSE OF REVIEW

- To gain a background knowledge of the research topic.
- To identify the concepts relating to it, potential relationships between them and to formulate researchable hypothesis.
- To identify appropriate methodology, research design, methods of measuring concepts and techniques of analysis.
- To identify data sources used by other researchers.
- To learn how others structured their reports.

How to conduct the Literature Survey?

- Identify the relevant sources.
- Extract and Record relevant information.
- Write-up the Literature Review.

SOURCES OF LITERATURE

- Books and Journals
- Electronic Databases
  - Bibliographic Databases
  - Abstract Databases
  - Full-Text Databases
- Govt. and Industry Reports
- Internet
- Research Dissertations / Thesis

RECORDING THE LITERATURE

- The most suitable method of recording notes is the card system.

- The recording system involves use of two sets of cards:
  - Source cards (3”x 5”) – used for noting bibliographic information.
  - Note cards (5”x 8”) – used for actual note taking.

SOURCE CARDS

- Source Cards serve two purposes:

  a) Provide documentary information for foot notes.

  b) It is used for compiling bibliography to be given at the end of the report.

SOURCE CARDS

* Source Cards can be coded by a simple system inorder to relate them to the corresponding note cards.

1) Marking a combination of letters and a number on the right hand top corner that begins with ‘C’. For example; C1, C2 etc.

OR

2) Marking the letter ‘B’ or ‘J’ or ‘R’ (B=Books, J=Journal, R=Report) on the left hand top corner.

SOURCE CARDS

- **The recording of bibliographic information should be made in proper bibliographic format.**

- The format for citing a book is:
  Author's name, (year), Title of the book, Place of publication, Publisher's name.
- For Example; Koontz Harold (1980), Management, New Delhi, McGraw-Hill International.

- The format for citing a journal article is:
  Author's name, (year), Title of the article, Journal name, Volume (number), pages.
- For Example; Sheth J.N (1973), A Model of Industrial Buying Behaviour, Journal of Marketing, 37(4), 50-56.

NOTE CARDS

- Detailed Information extracted from a printed source is recorded on the note cards.

- It is desirable to note a single fact or idea on each card, on one side only.

How to write the review?

* There are several ways of presenting the ideas of others within the body of the paper.

* For Example; If you are referring the major influencing factors in the Sheth's model of Industrial Buying Behaviour, it can be written as,

1) Sheth (1973, p-50) has suggested that, there are a number of influencing factors ……..
2) According to Sheth (1973) model of industrial buying behaviour, there are a number of influencing factors…….

How to write the review?

3) In some models of industrial buying behaviour, there are a number of influencing factors (Sheth, 1973).

4) In some models of industrial buying behaviour, there are a number of influencing factors[1].

1. Sheth J.N (1973), A Model of Industrial Buying Behaviour, Journal of Marketing, 37(4), 50-56.

Points to be kept in mind while reviewing literature..

- Read relevant literature.
- Refer original works.
- Read with comprehension.
- Read in time.
- Index the literature.

FORMULATION OF HYPOTHESIS

HYPOTHESIS - MEANING

- A hypothesis is an assumption about relations between variables.

- Hypothesis can be defined as a logically conjectured relationship between two or more variables expressed in the form of a testable statement.

  Relationships are conjectured on the basis of the network of associations established in the theoretical framework formulated for the research study.

VARIABLES

- Anything that can vary can be considered as a variable.

- A variable is anything that can take on differing or varying values.

For example; Age, Production units, Absenteeism, Sex, Motivation, Income, Height, Weight etc.

Note: The values can differ at various times for the same object or person (or) at the same time for different objects or persons.

Variable / Attribute

- A variable is a characteristic that takes on two or more values whereas, an attribute is a specific value on a variable (qualitative).

For example;

- The variable SEX/GENDER has 2 attributes - Male and Female.
- The variable AGREEMENT has 5 attributes – Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree.

Types of Variables

- **Explanatory vs Extraneous Variable**

  The variables selected for analysis are called explanatory variables and all other variables that are not related to the purpose of the study but may affect the dependant variable are extraneous.

- **Dependant vs Independent Variable**

  The variable that changes in relationship to changes in another variable(s) is called dependant variable.

  The variable whose change results in the change in another variable is called an independent variable.

  OR

  An independent variable is the one that influences the dependant variable in either a positive or negative way.

HYPOTHESIS

- Research Hypothesis is a predictive statement that relates an independent variable to a dependant variable.
- Hypothesis must contain atleast one independent variable and one dependant variable.
- Hypothesis are tentative, intelligent guesses as to the solution of the problem.
- Hypothesis is a specific statement of prediction. It describes in concrete terms what you expect to happen in the study.
- Hypothesis is an assumption about the population of the study.
- It delimits the area of research and keeps the researcher on the right track.

PROBLEM (VS) HYPOTHESIS

- Hypothesis is an assumption, that can be tested and can be proved to be right or wrong.

- A problem is a broad question which cannot be directly tested. A problem can be scientifically investigated after converting it into a form of hypothesis.

CHARACTERISTICS OF HYPOTHESIS

- Conceptual Clarity - It should be clear and precise.
- Specificity - It should be specific and limited in scope.
- Consistency - It should be consistent with the objectives of research.
- Testability - It should be capable of being tested.
- Expectancy - It should state the expected relationships between variables.

CHARACTERISTICS OF HYPOTHESIS

- Simplicity - It should be stated as far as possible in simple terms.
- Objectivity - It should not include value judgments, relative terms or any moral preaching.
- Theoretical Relevance - It should be consistent with a substantial body of established or known facts or existing theory.
- Availability of Techniques – Statistical methods should be available for testing the proposed hypothesis.

SOURCES OF HYPOTHESIS

- Discussions with colleagues and experts about the problem, its origin and objectives in seeking a solution.
- Examination of data and records for possible trends, peculiarities.
- Review of similar studies.
- Exploratory personal investigation / Observation.
- Logical deduction from the existing theory.
- Continuity of research.
- Intuition and personal experience.

TYPES OF HYPOTHESIS

Descriptive Hypothesis

These are assumptions that describe the characteristics (such as size, form or distribution) of a variable. The variable may be an object, person, organisation, situation or event.

Examples:

- **“Public enterprises are more amenable for centralized planning”.**

Relational Hypothesis [Explanatory Hypothesis]

These are assumptions that describe the relationship between two variables. The relationship suggested may be positive, negative or causal relationship.

Examples:

- **“Families with higher incomes spend more for recreation”.**

Causal Hypothesis state that the existence of or change in one variable causes or leads to an effect on another variable. The first variable is called the independent variable and the latter is the dependant variable.

Null Hypothesis

When a hypothesis is stated negatively, it is called null hypothesis. It is a 'no difference', 'no relationship' hypothesis. ie., It states that, no difference exists between the parameter and statistic being compared to or no relationship exists between the variables being compared.

It is usually represented as H_O or H_0 .

Example:

- **H_0: There is no relationship between a family's income and expenditure on recreation.**

Alternate Hypothesis

It is the hypothesis that describes the researcher's prediction that, there exist a relationship between two variables or it is the opposite of null hypothesis. It is represented as H_A or H_1.

Example:

H_A: **There is a definite relationship between family's income and expenditure on recreation.**

FORMS OF RELATIONSHIPS

NON-DIRECTIONAL

- There IS a relationship between
- X & Y

- X….linked….Y

Vs DIRECTIONAL

- If X goes up, Y ….
- or
- As X increases, Y…
- X = Independent
- variable
- Y = Dependent variable

DIRECTIONAL HYPOTHESES-

"X" causes "Y" to change

- If X changes (increases/decreases) then Y will (increase/decrease)
- a causal link

DIRECTION OF RELATIONSHIP

- If X increases, Y increases

  A **POSITIVE** relationship

- If X increase, Y decreases

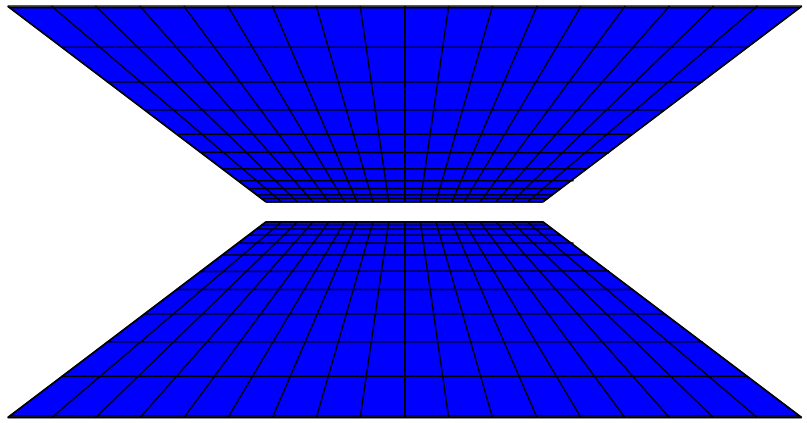
  A **NEGATIVE** or **INVERSE** relationship

- As X changes, Y does NOT change...>

  No Change...>**NO RELATIONSHIP**

NON-DIRECTIONAL HYPOTHESES
- the weakest form

- There Is
- a relationship
- between X & Y
  - non-causal
  - correlational statement
  - X…..Y



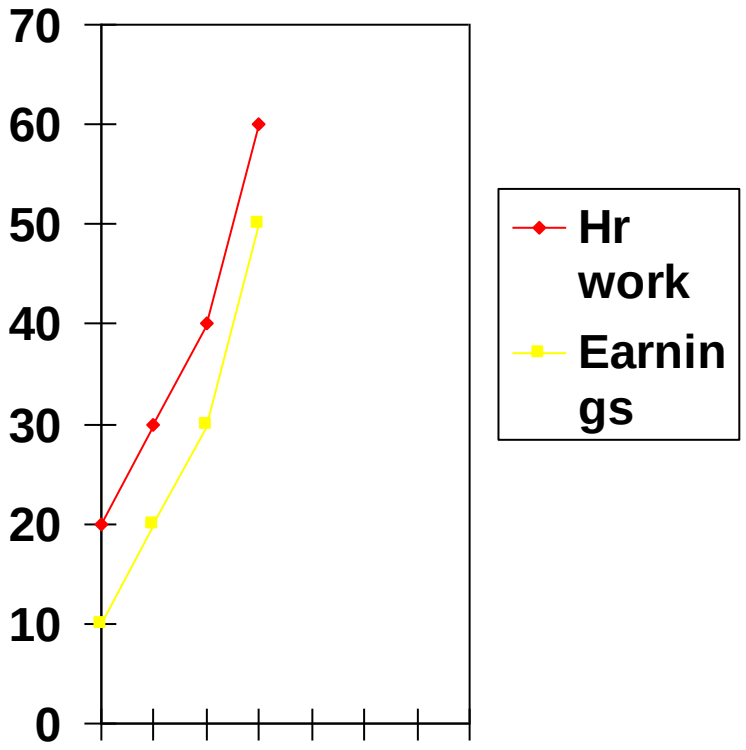
CORRELATIONAL RELATIONSHIP

Positive correlation

- When the values of TWO variables "go together"

or

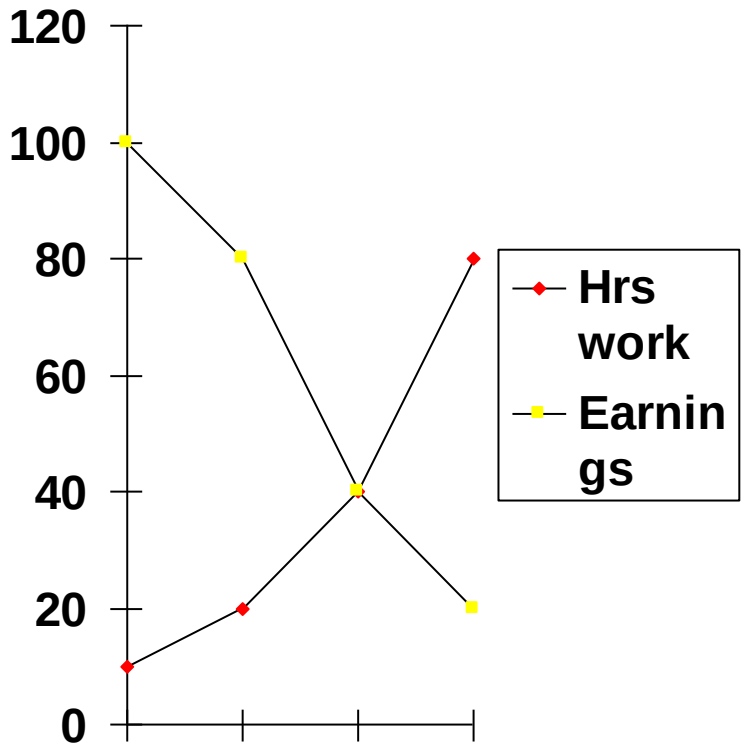
- Values on X & Y change in SAME DIRECTION



Negative Correlation

- When the values of two variables CO-VARY in Opposite direction

(as one goes up, the other goes down)

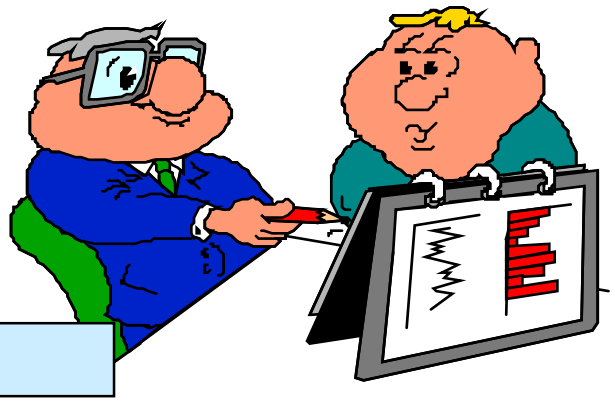


FUNCTIONS OR ROLE OF HYPOTHESIS

- It gives a definite point to the investigation and provides direction to the study.
- It determines the data needs.
- It specifies the sources of data.
- It suggests which type of research is likely to be more appropriate.
- It determines the most appropriate technique of analysis.
- It contributes to the development of theory.

Hypothesis Testing

Steps for Hypothesis Testing



Formulate H_0 and H_1

Select Appropriate Test

Choose Level of Significance

Calculate Test Statistic TS_{CAL}

Determine Prob Assoc with Test Stat

Compare with Level of Significance, α

Determine Critical Value of Test Stat TS_{CR}

Determine if TS_{CR} falls into (Non) Rejection Region

Reject/Do not Reject H_0

Draw Marketing Research Conclusion

Step 1: Formulate the Hypothesis

- A **null hypothesis** is a statement of the status quo, one of no difference or no effect. If the null hypothesis is not rejected, no changes will be made.
- An **alternative hypothesis** is one in which some difference or effect is expected.
- The null hypothesis refers to a specified value of the population parameter (e.g., μ, σ, π), not a sample statistic (e.g., $\overline{X}$).

Example of a Hypothesis Test

For the data in Table 15.1, suppose we wanted to test the hypothesis that the mean familiarity rating exceeds 4.0, the neutral value on a 7 point scale. A significance level of $\alpha = 0.05$ is selected. The hypotheses may be formulated as:

$$H_0: \mu \leq 4.0$$

$$H_1: \mu > 4.0$$

$$t = (\bar{X} - \mu)/s_{\bar{X}}$$

$$s_{\bar{X}} = 0.293$$

$$t_{CAL} = (4.724\text{-}4.0)/0.293 = 2.471$$

One Sample : t Test

- The df for the *t* stat is *n* - 1. In this case, *n* - 1 = 28.
- The probability assoc with 2.471 is less than 0.05. So the null hypothesis is rejected

- Alternatively, the critical t_α value for a significance level of 0.05 is 1.7011
- Since, 1.7011 <2.471, the null hypothesis is rejected.

- The familiarity level does exceed 4.0.
- Note that if the population standard deviation was **known** to be 1.5, rather than estimated from the sample, a ***z* test** would be appropriate.

Step 1: Formulate the Hypothesis

- A null hypothesis may be rejected, but it can never be accepted based on a single test.
- In marketing research, the null hypothesis is formulated in such a way that its rejection leads to the acceptance of the desired conclusion.
- A new Internet Shopping Service will be introduced if more than 40% people use it:

$$H_0: \pi \leq 0.40$$

$$H_1: \pi > 0.40$$

Step 1: Formulate the Hypothesis

- In eg on previous slide, the null hyp is a **one-tailed test**, because the alternative hypothesis is expressed directionally.
- If not, then a **two-tailed test** would be required as foll:

$$H_0: \pi = 0.40$$

$$H_1: \pi \neq 0.40$$

Step 2: Select an Appropriate Test

- The **test statistic** measures how close the sample has come to the null hypothesis.
- The test statistic often follows a well-known distribution (eg, normal, *t*, or chi-square).
- In our example, the *z* statistic, which follows the standard normal distribution, would be appropriate.

$$z = \frac{p - \pi}{\sigma_p}$$

Where σ_p is standard deviation

Step 3: Choose Level of Significance

Type I Error

- Occurs if the null hypothesis is rejected when it is in fact true.
- The probability of type I error (α) is also called the **level of significance**.

Type II Error

- Occurs if the null hypothesis is not rejected when it is in fact false.
- The probability of type II error is denoted by β .
- Unlike α, which is specified by the researcher, the magnitude of β depends on the actual value of the population parameter (proportion).

It is necessary to balance the two types of errors.

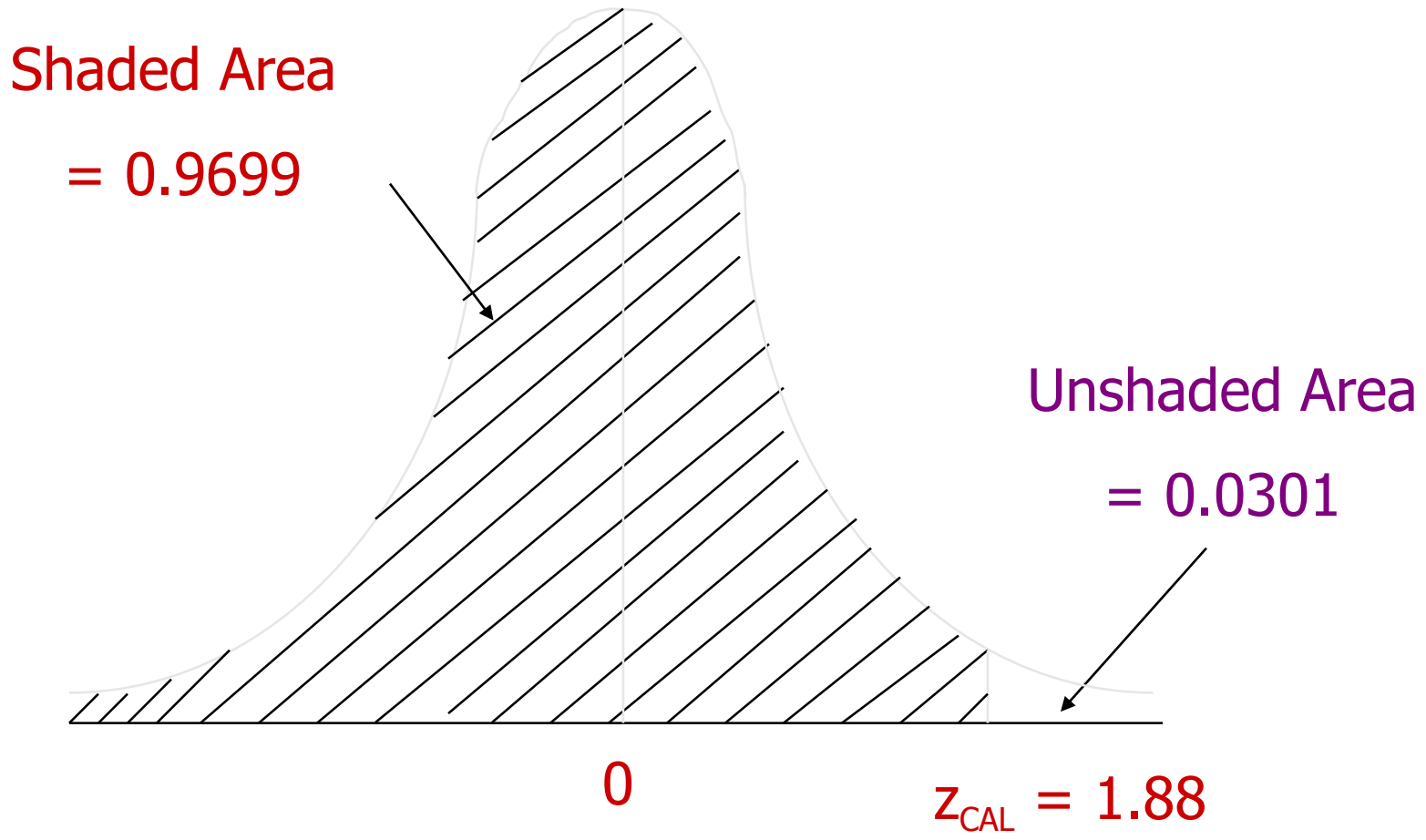
Step 3: Choose Level of Significance

Power of a Test

- The **power of a test** is the probability (1 - β) of rejecting the null hypothesis when it is false and should be rejected.
- Although β is unknown, it is related to α. An extremely low value of α (e.g., = 0.001) will result in intolerably high β errors.

Probability of z with a One-Tailed Test

Fig. 15.5



Step 4: Collect Data and Calculate Test Statistic

- The required data are collected and the value of the test statistic computed.
- In our example, 30 people were surveyed and 17 shopped on the internet. The value of the sample proportion is $\hat{p}$ 17/30 = 0.567.
- The value of σ_p is:

$$\sigma_p = 0.089$$

Step 4: Collect Data and Calculate Test Statistic

The test statistic z can be calculated as follows:

$$z_{CAL} = \frac{\hat{p} - \pi}{\sigma_p}$$

$$= \frac{0.567-0.40}{0.089}$$

$$= 1.88$$

Step 5: Determine Probability Value/ Critical Value

- Using standard normal tables (Table 2 of the Statistical Appendix), the area to the right of z_{CAL} is .0301 (z_{CAL} =1.88)
- The shaded area between 0 and 1.88 is 0.4699. Therefore, the area to the right of 1.88 is 0.5 - 0.4699 = 0.0301.
- Thus, the p-value is .0301

- Alternatively, the critical value of *z*, called z_{α}, which will give an area to the right side of the critical value of α=0.05, is between 1.64 and 1.65. Thus z_{α} =1.645.

- Note, in determining the critical value of the test statistic, the area to the right of the critical value is either α or $\alpha/2$. It is α for a one-tail test and $\alpha/2$ for a two-tail test.

Steps 6 & 7: Compare Prob and Make the Decision

- If the prob associated with the calculated value of the test statistic (z_{CAL}) is less than the level of significance (α), the null hypothesis is rejected.
- In our case, the p-value is 0.0301.This is less than the level of significance of α =0.05. Hence, the **null hypothesis is rejected**.
- Alternatively, if the calculated value of the test statistic is greater than the critical value of the test statistic (z_{α}), the null hypothesis is rejected.

Steps 6 & 7: Compare Prob and Make the Decision

- The calculated value of the test statistic z_{CAL}= 1.88 lies in the rejection region, beyond the value of z_{α}=1.645. Again, the same conclusion to reject the null hypothesis is reached.

- Note that the two ways of testing the null hypothesis are equivalent but mathematically opposite in the direction of comparison.

- Writing Test-Statistic as TS:

  If the probability of TS_{CAL} < significance level (α) then reject H_0 but if $TS_{CAL} > TS_{CR}$ then reject H_0.

Step 8: Mkt Research Conclusion

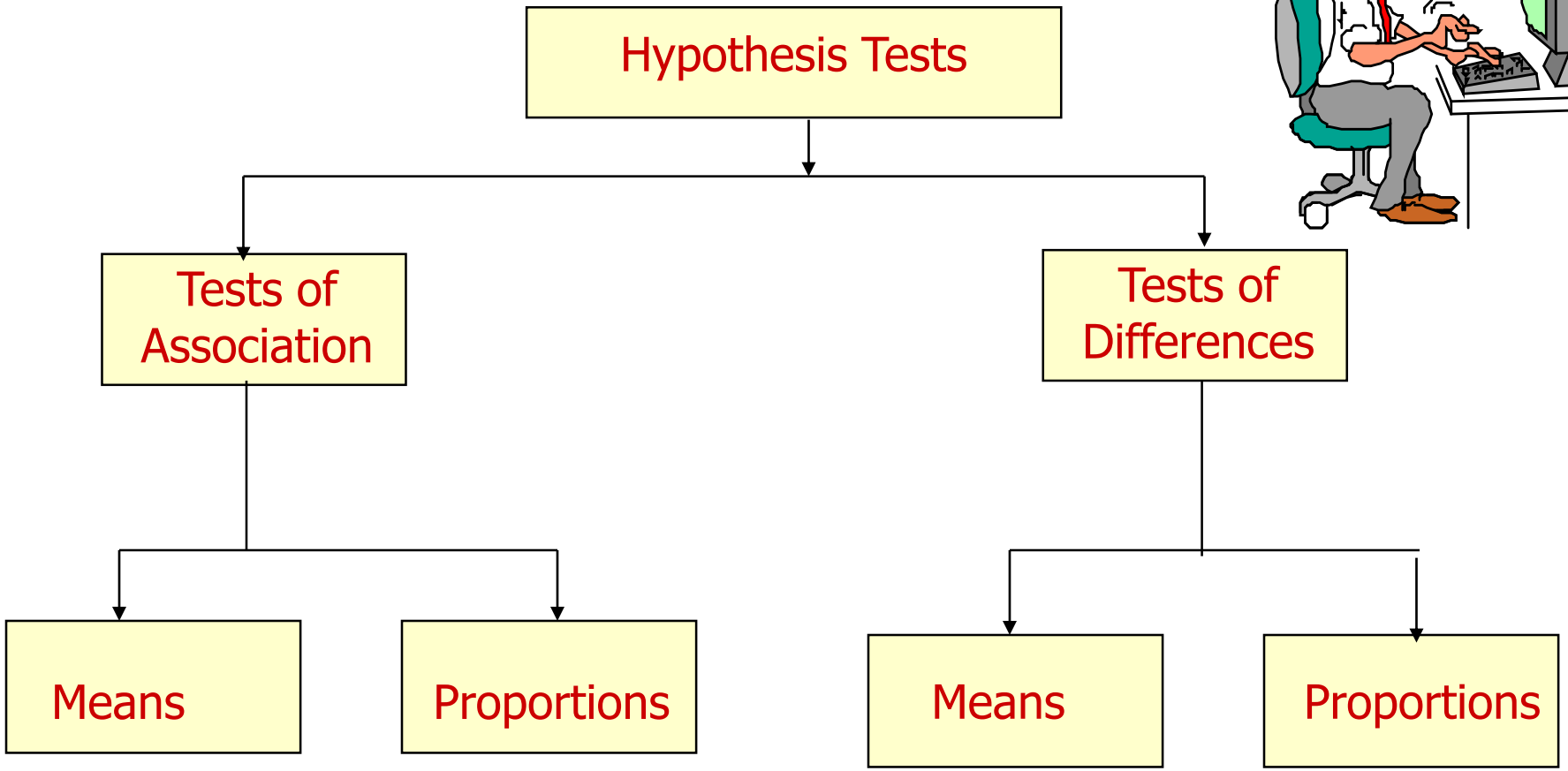
- The conclusion reached by hypothesis testing must be expressed in terms of the marketing research problem.

- In our example, we conclude that there is evidence that the proportion of Internet users who shop via the Internet is significantly greater than 0.40. Hence, the department store should introduce the new Internet shopping service.

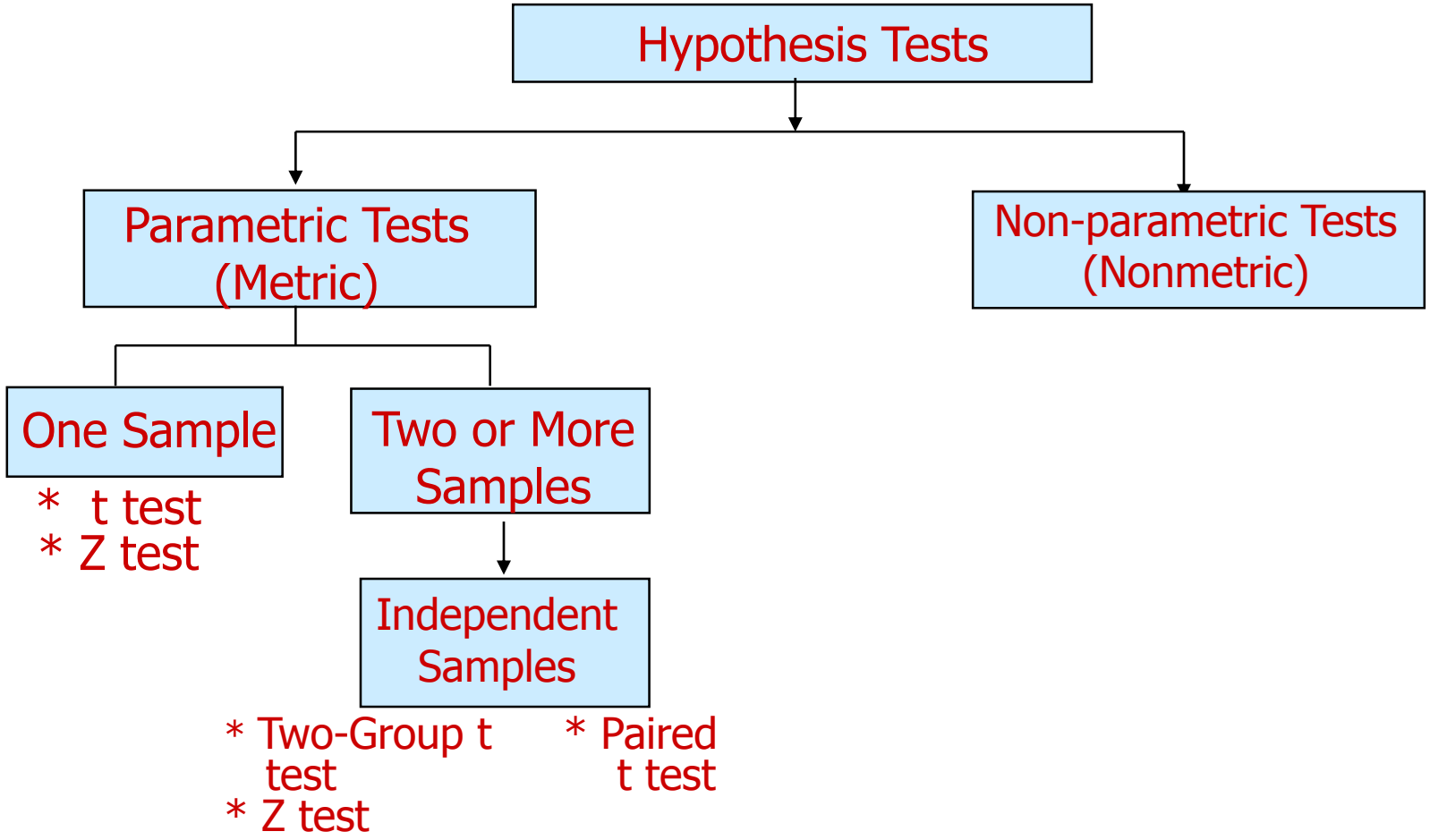
Using a t-Test

- Assume that the random variable X is normally dist, with unknown pop variance estimated by the sample variance s^2.
- Then a **t test** is appropriate.
- The t-statistic, $t = (\overline{X} - \mu)/s_{\overline{X}}$ is t distributed with n - 1 df.
- The **t dist** is similar to the normal distribution: bell-shaped and symmetric. As the number of df increases, the t dist approaches the normal dist.

Broad Classification of Hyp Tests



Hypothesis Testing for Differences



Two Independent Samples: Means

- In the case of means for two independent samples, the hypotheses take the following form.

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

- The two populations are sampled and the means and variances computed based on samples of sizes *n*1 and *n*2.
- The idea behind the test is similar to the test for a single mean, though the formula for standard error is different
- *Suppose we want to determine if internet usage is different for males than for females, using data in Table 15.1*

Internet Usage Data

Table 15.1

Respondent Number	Sex	Familiarity	Internet Usage	Attitude Toward Internet	Attitude Toward Technology	Usage of Internet Shopping	Usage of Internet Banking
1	1.00	7.00	14.00	7.00	6.00	1.00	1.00
2	2.00	2.00	2.00	3.00	3.00	2.00	2.00
3	2.00	3.00	3.00	4.00	3.00	1.00	2.00
4	2.00	3.00	3.00	7.00	5.00	1.00	2.00
5	1.00	7.00	13.00	7.00	7.00	1.00	1.00
6	2.00	4.00	6.00	5.00	4.00	1.00	2.00
7	2.00	2.00	2.00	4.00	5.00	2.00	2.00
8	2.00	3.00	6.00	5.00	4.00	2.00	2.00
9	2.00	3.00	6.00	6.00	4.00	1.00	2.00
10	1.00	9.00	15.00	7.00	5.00	1.00	2.00
11	2.00	4.00	3.00	4.00	3.00	2.00	2.00
12	2.00	5.00	4.00	6.00	4.00	2.00	2.00
13	1.00	6.00	8.00	8.00	5.00	2.00	1.00
14	1.00	6.00	8.00	3.00	2.00	2.00	2.00
15	1.00	6.00	5.00	5.00	4.00	1.00	2.00
16	2.00	4.00	3.00	4.00	3.00	2.00	2.00
17	1.00	6.00	9.00	5.00	3.00	1.00	1.00
18	1.00	4.00	4.00	5.00	4.00	1.00	2.00
19	1.00	7.00	14.00	6.00	6.00	1.00	1.00
20	2.00	6.00	6.00	6.00	4.00	2.00	2.00
21	1.00	6.00	9.00	4.00	2.00	2.00	2.00
22	1.00	5.00	5.00	5.00	4.00	2.00	1.00
23	2.00	3.00	2.00	4.00	2.00	2.00	2.00
24	1.00	7.00	15.00	6.00	6.00	1.00	1.00
25	2.00	6.00	6.00	5.00	3.00	1.00	2.00
26	1.00	6.00	13.00	6.00	6.00	1.00	1.00
27	2.00	5.00	4.00	5.00	5.00	1.00	1.00
28	2.00	4.00	2.00	3.00	2.00	2.00	2.00
29	1.00	4.00	4.00	5.00	3.00	1.00	2.00
30	1.00	3.00	3.00	7.00	5.00	1.00	2.00

Two Independent-Samples: *t* Tests

Table 15.14

Summary Statistics

	Number of Cases	Mean	Standard Deviation
Male	15	9.333	1.137
Female	15	3.867	0.435

***F* Test for Equality of Variances**

F value	2-tail probability
15.507	0.000

***t* Test**

Equal Variances Assumed			Equal Variances Not Assumed		
t value	Degrees of freedom	2-tail probability	*t* value	Degrees of freedom	2-tail probability
-4.492	28	0.000	-4.492	18.014	0.000

Two Independent Samples: Proportions

- Consider data of Table 15.1
- Is the proportion of respondents using the Internet for shopping the same for males and females?

The null and alternative hypotheses are:

$$H_0: \pi_1 = \pi_2$$
$$H_1: \pi_1 \neq \pi_2$$

- The test statistic is similar to the one for difference of means, with a different formula for standard error.

Summary of Hypothesis Tests for Differences

Sample	Application	Level of Scaling	Test/Comments
One Sample	Proportion	Metric	*Z* test
One Sample	Means	Metric	*t* test, if variance is unknown *z* test, if variance is known

Summary of Hypothesis Tests for Differences

Two Indep Samples	Application	Scaling	Test/Comments
Two indep samples	Means	Metric	Two-group *t* test *F* test for equality of variances
Two indep samples	Proportions	Metric Nonmetric	*z* test Chi -square test

THANK YOU