M.Com_Applied_Economics_IIYear_IVSem_IIIUnit

Dr. Tarannum Siddiqui

Measurement and scaling techniques

A) Measurement: Measurement is the process of observing and recording the observations that are collected as part of research. The recording of the observations may be in terms of numbers or other symbols to characteristics of objects according to certain prescribed rules. The respondent's, characteristics are feelings, attitudes, opinions etc. The most important aspect of measurement is the specification of rules for assigning numbers to characteristics. The rules for assigning numbers should be standardized and applied uniformly. This must not change over time or objects.

B) Scaling: Scaling is the assignment of objects to numbers or semantics according to a rule. In scaling, the objects are text statements, usually statements of attitude, opinion, or feeling.

Issues in attitude measurement-

When a researcher is interested in measuring the attitudes, feelings or opinions of respondents he/she should be clear about the following:

a) What is to be measured?

b) Who is to be measured?

c) The choices available in data collection techniques

Level of Measurement Scales-

The level of measurement refers to the relationship among the values that are assigned to the attributes, feelings or opinions for a variable. Typically, there are four levels of measurement scales or methods of assigning numbers:

a) Nominal scale,

(b) Ordinal scale,

(c) Interval scale, and

(d) Ratio scale.

(a) Nominal Scale is the crudest among all measurement scales but it is also the simplest scale. In this scale the different scores on a measurement simply indicate different categories. The nominal scale does not express any values or relationships between variables.

The nominal scale is often referred to as a categorical scale. The assigned numbers have no arithmetic properties and act only as labels. The only statistical operation that can be performed on nominal scales is a frequency count. We cannot determine an average except mode.

For example: labeling men as '1' and women as '2' which is the most common way of labeling gender for data recording purpose does not mean women are 'twice something or other' than men. Nor it suggests that men are somehow 'better' than women.

(b) Ordinal Scale involves the ranking of items along the continuum of the characteristic being scaled. In this scale, the items are classified according to whether they have more or less of a characteristic.

The main characteristic of the ordinal scale is that the categories have a logical or ordered relationship. This type of scale permits the measurement of degrees of difference, (i.e. 'more' or 'less') but not the specific amount of differences (i.e. how much 'more' or 'less'). This scale is very common in marketing, satisfaction and attitudinal research. Using ordinal scale data, we can perform statistical analysis like Median and Mode, but not the Mean.

For example, a fast food home delivery shop may wish to ask its customers: How would you rate the service of our staff? (1) Excellent • (2) Very Good • (3) Good • (4) Poor • (5) Worst •

(c) Interval Scale is a scale in which the numbers are used to rank attributes such that numerically equal distances on the scale represent equal distance in the characteristic being measured. An interval scale contains all the information of an ordinal scale, but it also one allows to compare the difference/distance between attributes. Interval scales may be either in numeric or semantic formats. The interval scales allow the calculation of averages like Mean, Median and Mode and dispersion like Range and Standard Deviation.

For example, the difference between '1' and '2' is equal to the difference between '3' and '4'. Further, the difference between '2' and '4' is twice the difference between '1' and '2'. Measuring temperature is an example of interval scale. But, we cannot say 40°C is twice as hot as 20°C.

(i) Example of interval scale in numeric format

Food supplied is:						
Fresh	1	2	3	4	5	
Tastes good	1	2	3	4	5	Indicate your score on the concerned black and circle the appropriate number on each line.
Value for money	1	2	3	4	5	
Attractive packaging	1	2	3	4	5	
Prompt time delivery	1	2	3	4	5	

(ii) Example of Interval Scale in Sematic Format

Please indicate your views on the food supplied by XXX fast shop by scoring them on a five points scale from 1 to 5 (that is, 1 = excellent, 2 = very good, 3 = good, 4 = poor, 5 = worst). Indicate your views by ticking the appropriate responses below:

Food Supplied is:	Excellent	Very good	Good	Poor	Worst
Fresh					
Tastes good					
Value for money					
Attractive packaging					
Prompt time delivery					

(d) Ratio Scale is the highest level of measurement scales. This has the properties of an interval scale together with a fixed (absolute) zero point. The absolute zero point allows us to construct a meaningful ratio.

Ratio scales permit the researcher to compare both differences in scores and relative magnitude of scores. Examples of ratio scales include weights, lengths and times.

For example, the number of customers of a bank's ATM in the last three months is a ratio scale. This is because you can compare this with previous three months.

For example, the difference between 10 and 15 minutes is the same as the difference between 25 and 30 minutes and 30 minutes is twice as long as 15 minutes

Types of Scaling Technique-

(1) **Comparative Scales**

- (i) Paired Comparison,
- (ii) Rank Order,
- (iii) Constant Sum,
- (iv) Q-Sort and Other Procedures

(2) **Non-comparative Scales**

- (i) Continuous Rating Scales,
- (ii) Itemized Rating Scales- (A) Likert, (B) Semantic differential, (C) Stapel

(1) Comparative Scales-

In comparative scaling, the respondent is asked to compare one object with another. The comparative scales can further be divided into the following four types of scaling techniques: (i) Paired Comparison Scale, (ii) Rank Order Scale, (iii) Constant Sum Scale, and (iv) Q-sort Scale.

(i) Paired Comparison Scale: This is a comparative scaling technique in which a respondent is presented with two objects at a time and asked to select one object according to some criterion. The data obtained are ordinal in nature.

For example, there are four types of cold drinks Coke, Pepsi, Sprite, and Limca. The respondents can prefer Pepsi to Coke or Coke to Sprite, etc.

Brand	Coke	Pepsi	Sprite	Limca
Coke	-	1		
Pepsi		-		
Sprite	1	1	-	
Limca	1	1	1	-
No. of times preferred	2	3	1	0

(ii) Rank Order Scale: This is another type of comparative scaling technique in which respondents are presented with several items simultaneously and asked to rank them in the order of priority. This is an ordinal scale that describes the favoured and unfavoured objects, but does not reveal the distance between the objects. The resultant data in rank order is ordinal data. This yields better results when direct comparison are required between the given objects. The major disadvantage of this technique is that only ordinal data can be generated.

Brand	Rank
Coke	3
Pepsi	1
Limca	2
Sprite	4

(iii) Constant Sum Scale: In this scale, the respondents are asked to allocate a constant sum of units such as points, rupees, or chips among a set of stimulus objects with respect to some criterion. For example, you may wish to determine how important the attributes of price, fragrance, packaging, cleaning power, and lather of a detergent are to consumers. Respondents might be asked to divide a constant sum to indicate the relative importance of the attributes. The advantage of this

technique is saving time. However, main disadvantages are the respondents may allocate more or fewer points than those specified. The second problem is respondents might be confused.

Attribute	No. of Points
Price	50
Fragrance	05
Packaging	10
Cleaning Power	30
Lather	05
Total Points	100

(iv) Q-Sort Scale: This is a comparative scale that uses a rank order procedure to sort objects based on similarity with respect to some criterion. The important characteristic of this methodology is that it is more important to make comparisons among different responses of a respondent than the responses between different respondents. Therefore, it is a comparative method of scaling rather than an absolute rating scale. In this method the respondent is given statements in a large number for describing the characteristics of a product or a large number of brands of a product. **Such as- Prefer Most, Like, Neutral, Dislike, Prefer Least.**

(2) In non-comparative scaling-

In non-comparative scaling respondents need only evaluate a single object. Their evaluation is independent of the other object, which the researcher is studying. The non-comparative scaling techniques can be further divided into:

(i) Continuous Rating Scale, and

(ii) Itemized Rating Scale.

(i) Continuous Rating Scales: It is very simple and highly useful. In continuous rating scale, the respondent's rate the objects by placing a mark at the appropriate position on a continuous line that runs from one extreme of the criterion variable to the other. Example: Question: How would you rate the TV advertisement as a guide for buying?

Strong Agree	10	9	8	7	6	5	4	3	2	1	0	Strong Disagree

(ii) Itemized Rating Scales: Itemized rating scale is a scale having numbers or brief descriptions associated with each category. The categories are ordered in terms of scale position and the respondents are required to select one of the limited number of categories that best describes the product, brand, company, or product attribute being rated. Itemized rating scales are widely used in marketing research. Itemised rating scales is further divided into three parts, namely-

(a) Likert scale,

(b) Semantic Differential Scale, and

(c) Stapel Scale.

The itemised rating scales can be in the form of : (a) graphic, (b) verbal, or (c) numeric as shown below :

Itemised Graphic Scale	Itemised Verbal Scale	Itemised Nemeric Scale
Favourable 	Completely Satisfied	-5
		-4
	Somewhat Satisfied	-3
		-2
Indifferent 		-1
	Neither Satisfied Nor dissatisfie	0
		+1
	Somewhat dissatisfied	+2
		+3
Unfavourable 		+4
	Completely dissatisfied	+5

(a) Likert Scale: Likert, is extremely popular for measuring attitudes, because, the method is simple to administer. With the Likert scale, the respondents indicate their own attitudes by checking how strongly they agree or disagree with carefully worded statements that range from very positive to very negative towards the attitudinal object. Respondents generally choose from five alternatives (say strongly agree, agree, neither agree nor disagree, disagree, strongly disagree). A Likert scale may include a number of items or statements. Disadvantage of Likert Scale is that it takes longer time to complete than other itemised rating scales because respondents have to read each statement. Despite the above disadvantages, this scale has several advantages. It is easy to construct, administer and use.

A Likert Scale for studying opinions on food products

Particular	Strongly Agree	Agree	Neither Agree nor disagree	Disagree	Strongly disagree
If the price of raw materials fall, firms too should reduce the price of the food products	1	2	3	4	5
There should be uniform price through out the country for food products	1	2	3	4	5
The food companies should concentrate more on keeping hygiene while manufacturing food products.	1	2	3	4	5
The expiry dates should be printed on the food products before the are delivered to consumers in the market.	1	2	3	4	5
There should be government regulations on the firms in keeping acceptable quality and on the prices.	1	2	3	4	5
Now-a-days most food companies are concerned only with profit making rather than taking care of quality.	1	2	3	4	5

(b) Semantic Differential Scale: This is a seven point rating scale with end points associated with bipolar labels (such as good and bad, complex and simple) that have semantic meaning. It can be used to find whether a respondent has a positive or negative attitude towards an object. It has been widely used in comparing brands, products and company images. It has also been used to develop advertising and promotion strategies and in a new product development study.

Examples of Semantic Differential Scale				
Modern	-	-	-	Old- fashioned
Good	-	-	-	Bad
Clean	-	-	-	Dirty
Important	-	-	-	Unimportant
Expensive	-	-	-	Inexpensive
Useful	-	-	-	Useless
Strong	-	-	-	Weak
Quick	-	-	-	Slow

(c) Staple Scale: The Stapel scale was originally developed to measure the direction and intensity of an attitude simultaneously. Modern versions of the Stapel scale place a single adjective as a substitute for the Semantic differential when it is difficult to create pairs of bipolar adjectives. The modified Stapel scale places a single adjective in the centre of an even number of numerical Values.

Selection of an appropriate scaling technique-

A number of issues decide the choice of scaling technique. Some significant issues are: 1) Problem Definition and Statistical Analysis,

2) The Choice between Comparative and Non-comparative Scales,

3) Type of Category Labels,

4) Number of Categories,

5) Balanced versus Unbalanced Scale, and

6) Forced versus Non-forced Categories

What is a Hypothesis? - Definition & Explanation

Ordinarily, when one talks about hypothesis, one simply means a mere assumption or some supposition to be proved or disproved. But for a researcher hypothesis is a formal question that he intends to resolve. Thus a hypothesis may be defined as a proposition or a set of proposition set forth as an explanation for the occurrence of some specified group of phenomena either asserted merely as a provisional conjecture to guide some investigation or accepted as highly probable in the light of established facts. Quite often a research hypothesis is a predictive statement, capable of being tested by scientific methods, that relates an independent variable to some dependent variable. For example, consider statements like the following ones:

"Students who receive counselling will show a greater increase in creativity than students not receiving counselling" Or "the automobile A is performing as well as automobile B."
These are hypotheses capable of being objectively verified and tested. Thus, we may conclude that a hypothesis states what we are looking for and it is a proposition, which can be put to a test to determine its validity.

Characteristics of hypothesis: Hypothesis must possess the following characteristics:

1. Hypothesis should be clear and precise. If the hypothesis is not clear and precise, the inferences drawn on its basis cannot be taken as reliable.
2. Hypothesis should be capable of being tested. In a swamp of untestable hypotheses, many a time the research programmes have bogged down. Some prior study may be done by researcher in order to make hypothesis a testable one. A hypothesis "is testable if other deductions can be made from it which, in turn, can be confirmed or disproved by observation."
3. Hypothesis should state relationship between variables, if it happens to be a relational hypothesis.
4. Hypothesis should be limited in scope and must be specific. A researcher must remember that narrower hypotheses are generally more testable and he should develop such hypotheses.
5. Hypothesis should be stated as far as possible in most simple terms so that the same is easily understandable by all concerned. But one must remember that simplicity of hypothesis has nothing to do with its significance.
6. Hypothesis should be consistent with most known facts i.e., it must be consistent with a substantial body of established facts. In other words, it should be one which judges accept as being the most likely.
7. Hypothesis should be amenable to testing within a reasonable time. One should not use even an excellent hypothesis, if the same cannot be tested in reasonable time for one cannot spend a life-time collecting data to test it.
8. Hypothesis must explain the facts that gave rise to the need for explanation. This means that by using the hypothesis plus other known and accepted generalizations, one should be able to deduce the original problem condition. Thus hypothesis must actually explain what it claims to explain; it should have empirical reference.

TESTING OF HYPOTHESIS IN RESEARCH METHODOLOGY -

As has been stated above that hypothesis testing determines the validity of the assumption (technically described as null hypothesis) with a view to choose between two conflicting

hypotheses about the value of a population parameter. Hypothesis testing helps to decide on the basis of a sample data, whether a hypothesis about the population is likely to be true or false. Statisticians have developed several tests of hypotheses (also known as the tests of significance) for the purpose of testing of hypotheses which can be classified as:

a. Parametric tests or standard tests of hypotheses; and

b. Non-parametric tests or distribution-free test of hypothesis.

Parametric tests usually assume certain properties of the parent population from which we draw samples. Assumptions like observations come from a normal population, sample size is large, assumptions about the population parameters like mean, variance, etc., must hold good before parametric tests can be used. But there are situations when the researcher cannot or does not want to make such assumptions. In such situations we use statistical methods for testing hypotheses, which are called non-parametric tests because such tests do not depend on any assumption about the parameters of the parent population. Besides, most non-parametric tests assume only nominal or ordinal data, whereas parametric tests require measurement equivalent to at least an interval scale. As a result, non-parametric tests need more observations than parametric tests to achieve the same size of Type I and Type II errors.

PROCEDURE FOR HYPOTHESIS TESTING IN RESEARCH METHODOLOGY

Procedure for 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? Procedure for hypothesis testing refers to all those steps that we undertake for making a choice between the two actions i.e., rejection and acceptance of a null hypothesis. The various steps involved in hypothesis testing are stated below:

1. **Making a formal statement:** The step consists in making a formal statement of the null hypothesis (H_0) and also of the alternative hypothesis (H_a). This means that hypotheses should be clearly stated, considering the nature of the research problem. For instance, Mr. Mohan of the Civil Engineering Department wants to test the load bearing capacity of an old bridge which must be more than 10 tons, in that case he can state his hypotheses.
Take another example. 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.
The formulation of hypotheses is an important step which must be accomplished with due care in accordance with the object and nature of the problem under consideration. It also indicates

whether we should use a one-tailed test or a two-tailed test. If Ha is of the type greater than (or of the type 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.

2. **Selecting a significance level:** The hypotheses are tested on a pre-determined level of significance and as such the same should be specified. Generally, in practice, either 5% level or 1% level is adopted for the purpose. The factors that affect the level of significance are: the magnitude of the difference between sample means, the size of the samples, the variability of measurements within samples; and whether the hypothesis is directional or non-directional (A directional hypothesis is one which predicts the direction of the difference between, say, means). In brief, the level of significance must be adequate in the context of the purpose and nature of enquiry.

3. **Deciding the distribution to use:** After deciding the level of significance, the next step in hypothesis testing is to determine the appropriate sampling distribution. The choice generally remains between normal distribution and the t-distribution. The rules for selecting the correct distribution are similar to those which we have stated earlier in the context of estimation.

4. **Selecting a random sample and computing an appropriate value:** Another step is to select a random sample(s) and compute an appropriate value from the sample data concerning the test statistic utilizing the relevant distribution. In other words, draw a sample to furnish empirical data.

5. **Calculation of the probability:** One has then to calculate the probability that the sample result would diverge as widely as it has from expectations, if the null hypothesis were in fact true.

6. **Comparing the probability:** Yet another step consists in comparing the probability thus calculated with the specified value for a , the significance level. If the calculated probability is equal to or smaller than the a value in case of one-tailed test (and a /2 in case of two-tailed test), then reject the null hypothesis (i.e., accept the alternative hypothesis), but if the calculated probability is greater, then accept the null hypothesis. In case we reject H0, we run a risk of (at most the level of significance) committing an error of Type I, but if we accept H0, then we run some risk (the size of which cannot be specified as long as the H0 happens to be vague rather than specific) of committing an error of Type II.