PSY 319
Statistical Methods in Psychology

Ibadan Distance Learning Centre Series

PSY 319
Statistical Methods in Psychology

By
A.M. Lawal
Department of Psychology
University of Ibadan

Published by
Distance Learning Centre
University of Ibadan

First Published 2009

ISBN 978-021-391-0

General Editor: Prof. F. O. Egbokhare
Series Editor: Olubunmi. I. Adeyemo

Typesetted @ Distance Learning Centre, University of Ibadan

Table of Contents

Vice-Chancellor’s Message

I congratulate you on being part of the historic evolution of our Centre for External Studies into a Distance Learning Centre. The reinvigorated Centre, is building on a solid tradition of nearly twenty years of service to the Nigerian community in providing higher education to those who had hitherto been unable to benefit from it.

Distance Learning requires an environment in which learners themselves actively participate in constructing their own knowledge. They need to be able to access and interpret existing knowledge and in the process, become autonomous learners.

Consequently, our major goal is to provide full multi media mode of teaching/learning in which you will use not only print but also video, audio and electronic learning materials.

To this end, we have run two intensive workshops to produce a fresh batch of course materials in order to increase substantially the number of texts available to you. The authors made great efforts to include the latest information, knowledge and skills in the different disciplines and ensure that the materials are user-friendly. It is our hope that you will put them to the best use.

Professor Olufemi A. Bamiro, FNSE
Vice-Chancellor

Foreword

The University of Ibadan Distance Learning Programme has a vision of providing lifelong education for Nigerian citizens who for a variety of reasons have opted for the Distance Learning mode. In this way, it aims at democratizing education by ensuring access and equity.

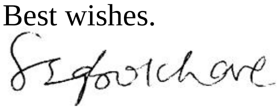
The U.I. experience in Distance Learning dates back to 1988 when the Centre for External Studies was established to cater mainly for upgrading the knowledge and skills of NCE teachers to a Bachelors degree in Education. Since then, it has gathered considerable experience in preparing and producing course materials for its programmes. The recent expansion of the programme to cover Agriculture and the need to review the existing materials have necessitated an accelerated process of course materials production. To this end, one major workshop was held in December 2006 which have resulted in a substantial increase in the number of course materials. The writing of the courses by a team of experts and rigorous peer review have ensured the maintenance of the University's high standards. The approach is not only to emphasize cognitive knowledge but also skills and humane values which are at the core of education, even in an ICT age.

The materials have had the input of experienced editors and illustrators who have ensured that they are accurate, current and learner friendly. They are specially written with distance learners in mind, since such people can often feel isolated from the community of learners. Adequate supplementary reading materials as well as other information sources are suggested in the course materials.

The Distance Learning Centre also envisages that regular students of tertiary institutions in Nigeria who are faced with a dearth of high quality textbooks will find these books very useful. We are therefore delighted to present these new titles to both our Distance Learning students and the University's regular students. We are confident that the books will be an invaluable resource to them.

We would like to thank all our authors, reviewers and production staff for the high quality of work.

Best wishes.



Professor Francis O. Egbokhare
Director

General Introduction and Course Objectives

This course discusses basic descriptive and inferential statistics including scales of measurement, hypothesis testing for both correlational and experimental techniques applicable to the behavioural, social, medical sciences and education. The course further cuts across statistical techniques such as Pearson, Spearman rank order correlations, chi-square, t-test for independent and repeated measures and analysis of variance. Specifically, the course provides students with the following abilities at the end:

1. the ability to use common statistical techniques to examine and describe data so as to reach objective, conclusions based on the obtained data;
2. the ability to understand statistical terms and research reports as found in psychology, other social sciences disciplines, education and the medical sciences;
3. the ability to select the appropriate statistical test for a given situation;
4. the ability to perform hypothesis tests, i.e., formulate a hypothesis, perform a statistical test, and report conclusions, and
5. the ability to read, understand and interpret statistical computer print-out or result with the use of Statistical Package for Social Sciences (SPSS) into behavioural implications.

Therefore, every lecture in this course must be effectively handled with total commitment and dedication as a student who is ready to excel in the area of statistics.

LECTURE ONE

Basic Descriptive Statistics

Introduction

In this lecture, I intend to introduce to you basic descriptive statistics. An example is regular frequency distributions used to describe data in a study to give it a more statistical meaning.

Objectives

At the end of this lecture, you should be able to:

1. explain descriptive statistics; and
2. discuss the frequency distribution is calculated.

Pre-Test

Define descriptive statistics

CONTENT

Descriptive Statistics

Descriptive statistics are statistical procedures used to organize and summarize the main characteristics of sample data. They describe and summarize information collected from a sample, usually reducing large amounts of numerical data in a more manageable form. When researchers collect data in a study, the data are typically a vast number of numbers, page after page of entries, or often today, a large computer spread sheet of rows and columns of numbers. Even to the trained eye of the researchers who collected them, the data can seem incomprehensible. The usual techniques used in descriptive statistics include the frequency distribution (e.g. how many male and female students attended all the PSY 319 classes?), the mean (e.g. on average, how many days did students come for PSY 319 classes?), the media (e.g. in computing the ages of students attending PSY 319 classes, in ascending/descending form, what age (s) falls in the middle?) and the mode (e.g. what age is most common in the ages of students attending PSY 319 classes?). Finally, for this course, another usual statistical technique for descriptive analysis is standard deviation.

Frequency Distribution

Once the data from a research are collected and coded (i.e. converted to numbers for analysis) they must be organized before they can be examined or presented to others. The best procedure for beginning the process of organizing and summarizing these numbers is to construct a frequency distribution. A frequency distribution is a table presenting the number of participant responses (e.g. scores, values) within the numerical categories of some scale of measurement. Thus, frequency distributions display the type of response made in a piece of research, as well as how many participants actually made each response (i.e. frequency).

A frequency distribution makes evident what cannot be seen in raw data. It shows clearly the scores of individuals whether high or low, whether it is concentrated in one area or spread out across the whole scale, thereby giving us an organized picture of the data. Frequency distribution can be depicted in tables or graphs, showing the scale of measurement and the number or frequency of individuals in each category.

Example 1:

Below are the scores of 30 students who sat for PSY 319 Test:

10	2	3	8	7	3	1	4	7	2
2	2	9	1	10	10	5	3	3	4
3	9	2	4	3	2	6	2	9	2

Supposing you want to explain the trend of the students' performance to the head of department, how will you summarize the data? One of the ways you can do this is to prepare a frequency distribution table that will show the tabular arrangement of the data together with the corresponding class frequencies. The frequency distribution will consist of the scores and the frequency associated with each score.

Table 1.1 Frequency distribution of students' performance on PSY 319 test

Score (x)	F
10	3
9	3
8	1
7	2
6	1
5	1
4	3
3	6
2	8
1	2
	$\sum$ **f = 30**

From Table 1, you can observe the result of the PSY 319 test. The result shows that 3 students had the highest score of 10. Similarly, 3 students had 9 score. However, majority of the students, especially 8 of them had lower score of 2. What is important to note here, is that the sum of the frequency (i.e. $\sum$ **f**) will give you the total number of students who participated in the PSY 319 test?

Summary

In the lecture, descriptive statistics was defined as the statistical procedure used to organize and summarize the main characteristics of sample data in a more manageable form. One of the descriptive statistical procedures commonly used is the regular frequency distribution. A frequency distribution is a table presenting the number of participant responses within the numerical categories of some scale of measurement.

Post-Test

Explain what a regular frequency distribution is

Reference

Dunn, D. S. (2001). *Statistics and Data Analysis for the Behavioural Sciences*. McGraw Hill, New York, NY 1020.

LECTURE TWO

Group Frequency Distribution

Introduction

In this lecture, you will be introduced to another type of frequency distribution. It is used for grouped data that presents a set of data with large scores in a statistically meaningful way.

Objectives

At the end of this lecture, you should be able to:

1. explain how the group frequency distribution is calculated and the table is constructed; and
2. identify the lower and upper limits of a class interval

Pre-Test

1. Define the range.
2. What is upper limit in grouped frequency distribution?

CONTENT

Group Frequency Distribution

Apart from the regular frequency distribution, there is what is called grouped frequency distribution for grouped data. When data covers a large number of values, listing all the scores in a simple frequency distribution table will be unwise. Rather, a grouped frequency distribution table will be appropriate for a large number of values. A grouped frequency distribution places raw scores into preset intervals of values. Instead of examining one score or frequency at a time, a range of scores or frequencies falling inside an interval can be considered. These intervals represent groups of scores; hence the name grouped frequency distribution.

Group frequency distributions are not difficult to construct but there are a few guidelines that you need to typically follow. In this lecture, these guidelines will be reviewed and demonstrated using scores on self efficacy of 30 students who were just transferred to a new campus. The scale has 8 items; where high score on it indicates higher self efficacy, while low score suggests lower efficacy. Respondents rate their agreement with each statement on a 5-point Likert format (i.e. I =strongly disagree to 5 = strongly agree). The scores range from 8-40. I think, though the following guidelines will prove to be helpful to you in calculating grouped frequency distribution:

a. Begin by calculating the difference between the highest and the lowest scores in a set of data i.e. H-L+1, where H= highest score, and L-lowest score.
b. You now determine the number of class interval. Intervals are symbols defining a class e.g. 45-49. The end numbers 45 and 49 are called class limits (45 is lower limit and 49 is upper limit). Ideally, when constructing a grouped frequency distribution, the number of intervals should not be more than 10 class intervals. Also, the class width/size of each interval should be a simple number e.g. 2, 4, 5, or 10 (Alarape, 2005); this will enable the reader to quickly see how you have divided your range. The lowest score in each class interval should be the multiple of the width and the intervals should be of equal width; covering all the range of scores without leaving any unaccounted for.
c. The third step entails arranging intervals in a table by starting from the lowest one and working up to the highest point. The lowest score in the distribution is the ***lower limit*** of the first interval and the ***upper limit*** of the interval can be determined using the following formula: *upper limit minus lower limit plus (class size minus -1).*
d. Finally, enter all the raw scores into their appropriate scoring intervals and under a column labeled “frequency” or “f”

Example

Let us now consider the example on self efficacy of the 30 transferred students aforementioned to show the steps involved in constructing grouped frequency distribution. The 30 students' scores on self- efficacy is given below, where N=30

10	10	22	35	30	33	9	31	27	40
33	30	21	29	14	22	17	34	20	32
22	32	22	36	9	22	10	19	10	12

You need to first determine the range of the scores. The *range* is determined by R = H – L, where H denotes the highest score, L denotes the lowest score. In the example above, the range is R = 40-9 = 31. This suggests that you need 31 rows to show the score of each student in a simple frequency distribution, which may be unwise because of the large scores. In this situation, grouped frequency distribution will be appropriate. The next thing you need to do to compute grouped frequency distribution for the data stated above is to compute the number of class interval following this formula.

CI = $\frac{H-L}{i}+1$; where H – highest score, L – lowest score, i = class size/width and CI= class interval. In this example, using a class size of 5 C.I = $\frac{40-9}{5}+1$ = $\frac{31}{5}+1$

= 6.2 + 1 = 7.2 i. e. 7 approximately. Having determined the number of the class interval for the data which is 7, you have to start from the lowest score falling in the first class interval starting from the bottom upward.

You now take each score from the data given and locate its appropriate class interval putting a tally mark in front of the row of that particular interval. The tally mark is made by drawing a slash representing each score that falls within that interval. After the fourth slash, the fifth slash is used to cross out the earlier four slashes indicating block of fives. This process makes counting easy after the tallying. After this, the number of cases that fall within a class interval is counted to obtain the frequency (f). The above data are used as an example for computing a grouped frequency distribution.

Table 2.1: A grouped Frequency Distribution Table of Student's Self Efficacy

Class Interval	Tally Mark	Frequency
39 – 43	/	1
34 – 38	|||	3
29 – 33	卌 |||	8
24 – 28	/	1
19 – 23	卌 |||	8
14 – 18	||	2
9 – 13	卌 ||	7
		$\sum$ f= 30 = N

From the above table you can see the frequency of occurrence within each of the class of the class interval. For instance, there are seven students whose scores fall within a class interval of 9 – 13 with frequency of 7. It worth noting that if there are errors in the tally marks, the sum of the frequency ($\sum$ f) will not be equal to the total number (N) of the transferred students whose self efficacy was measured. As you can see, a grouped frequency distribution like this one provides information quickly than a simple or regular frequency distribution

.

Summary

In the lecture, grouped frequency distribution has been well differentiated from the simple or regular frequency distribution based on it procedures in computing the table. In the grouped frequency distribution, there are class intervals because it is appropriate for data with large values or scores.

Post-Test

1. What is lower limit?
2. What is a range?

Reference

Alarape, A. I. (2005). Statistical Methods and Computer Application, in B. Udegbe, S. Balogun, H. Osinowo and G. Sunmola (eds). *Psychology: Perspectives in Human Behaviour*. Ibadan, Department of Psychology, University of Ibadan (Chapter 3), 59-88.

LECTURE THREE

Measures of Central Tendency

Introduction

In this lecture, I intend to introduce to you another descriptive statistics referred to as measures of central tendency. The measures of central tendency include the mean, the median and the mode. Also, in this lecture, standard deviation will be calculated.

Objectives

At the end of this lecture, you should be able to:

1. discuss the three measures of central tendency; and
2. differentiate all the three measures of central tendency.

Pre-Test

1. Discuss the three measures of central tendency.
2. What are the differences among the mean, the median and the mode?

CONTENT

The Measures of Central Tendency

Measures of central tendency are descriptive statistics that identify the central location of a sample of data. The central tendency of a set of data is the best single indicator describing the representative value (s) of any sample. Measures of central tendency are very useful in discerning differences between groups of individuals. The most common ways of measuring central tendency are *the mean, the median and the mode*. Each of these measures is reviewed in this lecture with comprehensive examples. All these measures of central tendency have different characteristics and their computation differ.

The Mean

The mean is the arithmetic mean or average of a set of scores. It is calculated by summing the set of scores and dividing them by the N (number) of scores. That is, $\overline{X} = \frac{\sum x}{N}$

where $\overline{X}$ = is the mean, $\sum$ = is the summation sign, x = is the score and N = is the number of scores. The mean is far the single most useful measure of central tendency available; it can be used to analyze data from interval or ratio scales of measurement.

Example 1

If we have an array of scores from a sample that were randomly drawn, let's say 12, 15, 10, 14 and 8, their arithmetic mean is: $\overline{X}$ $= \frac{12+15+10+14+8}{5}$, $= \frac{59}{5}$, $\overline{X}$ = 11.8.

While the arithmetic mean is used for ungrouped data however, there is a particular formula by which to calculate the mean for a set of grouped data.

The Mean of Grouped Frequency Distribution

In determining the mean of a frequency distribution for a grouped data, simply multiply each score by its frequency and then add up the result. Thereafter, divide your result by the number of scores (N), which can be found by summing the frequencies. The formula for mean of grouped data is $\overline{X} = \frac{\sum fx_c}{N}$, where, x_c = is midpoint of a class interval, f = is the number of scores that fall within each interval and $\overline{X}$, $\sum$ and N remain the same as earlier stated. To compute the mean for grouped frequency distribution following the earlier example.

Table 3.1: Showing the Grouped Frequency Distribution

Scores	**Midpoint (x)**	**F**	**Fx**
39 – 43	41	1	41
34 – 38	36	3	108
29 – 33	31	8	248
24 – 28	26	1	26
19 – 23	21	8	168
14 – 18	16	2	32
9 – 13	11	7	77
Sum		**N = 30**	$\sum \boldsymbol{fx_c} = \mathbf{700}$

In Table 3.1 the mid-point is calculated by adding the lower and upper limits of each of the class interval and divided it by 2 e. g. 9 +13 = 22 /2 = 11. To determine the mean for the grouped data, you sum the *fx* and divide it by N, that is, Mean = $\frac{\sum fx_c}{N} = \frac{700}{30} = 23.33$

The Median

The median is the number of scores that precisely divides a distribution of data in half. The median can be used when calculating ordinal, interval or ratio scale data. When you are interested in the exact mid-point of a distribution, the median is appropriate statistic. Fifty percent of a distribution's observations will fall above the median and fifty percent will fall below it.

To calculate the median in a distribution for ungrouped data as in case of these 10 score data.

Example 2

Consider the following data and calculate the median score for the distribution.

26 32 21 12 15 11 27 16 18 21.

You need to first re-arrange the scores from the lowest to the highest as follows: 11 12 15 16 18 21 21 26 27 32.

When you have an even number of scores like this, you can use this formula:

$$\textit{Median score} = \frac{N+1}{2}$$

In this example, $\frac{10+1}{2} = \frac{11}{2} = 5.5$

This suggests that the median score can be found in 5.5th scores into the array, which places it between the scores of 18 and 21. That is, 18 + 21/2 = 19.5. On the other hand, when you have an odd number of scores as follows: 55 67 78 83 88 92 98. You arrange the scores from the lowest to the highest and use the same formula.

$$\textit{Median score} = \frac{7+1}{2} = \frac{8}{2} = 4.$$

Locate the 4th score in the original array, which is 83 as the median score. While this process is used for ungrouped data, however, there is a particular procedure for calculating the median for grouped data.

The Median for Grouped Frequency Distribution

To compute the median for grouped data you may follow these steps.

1. You first divide the number of cases (N) in the distribution by 2 that is, N/2. Since the median lies between half of the number of score above and below the other half of the distribution, you need to count from the lowest interval until you locate the interval containing the median.
2. You then ascertain the number of scores needed in this interval to make half the number of scores.

   Furthermore, you need to divide the value you obtained by the number of scores within the interval where the median is located.
2. Then, multiply the value you obtain by the size of class interval. Add the result obtained to the lower limit of the median class interval and check your result by adding from the top of the intervals.
3. Finally, take away the value you obtained from the upper limit of the interval where the median is located.

Considering the previous data, you can compute the median of grouped frequency distribution as follows.

Table 3.2: Showing the Median for Group Frequency Distribution

Scores	**F**	
39 – 43	1	
34 – 38	3	13
29 – 33	8	
24 – 28	1	
19 – 23	8	
14 – 18	2	9
9 – 13	7	
	N = 30	

In Table 4, 13 = is the number of scores above the median interval, while 9 = is the number of scores below the media interval. Therefore,

The Median = 18. 5 + 6/8 x 5 (5 = class size),

= 18. 5 + (0.75 x 5),

= 18. 5 + 3.75, = 22.25

The Mode

The mode is the third measure of central tendency. The mode is not calculated; rather, it is simply reported once the various frequencies within a distribution of data are known. In short, the mode is the most frequently occurring score in a given distribution of scores.

Example 3

In a given distribution of scores like the following:

3 2 2 4 6 7 7 7 6 9. The mode is the value with the highest frequency which is 7 for ungrouped data. That is,

Score	0	1	2	3	4	5	6	7	8	9
Frequency	0	0	2	1	1	0	2	3	0	1

Calculating mode of grouped data with intervals

When you have grouped data with class interval, you can either determine the crude or interpolated mode. Though, interpolated mode is more accurate than the crude mode. However, the two forms of mode are calculated in this lecture.

The Crude Mode is the mid-point value of the class interval having the highest frequency. Considering the earlier given data, the class intervals with the highest frequency are 29 – 33 and 19 - 23. In this case, we have two class intervals with highest frequency of 8 each; this is referred to as bimodal. However, the crude modes of the distribution are the mid-points, which are 31 and 21. If in the data, there is a single class interval with highest frequency, the mid-point of the class interval is the crude mode.

Interpolated Mode occurs when the use of crude mode is not feasible, the interpolated mode is favoured. This of course occurs when we have a large sample data and the distribution is skewed. In this kind of situation, you use interpolated mode to determine the mode of the distribution. The formula for calculating the interpolated mode is given as:

$M_0 = \text{LRL} + \frac{d1}{d_1 + d_2} xi$, where: LRL = Lower real limits, d_1 = difference between the frequency of the modal interval and frequency of the preceding interval, d_2 = difference between the frequency of the modal interval and the frequency of the next following interval, i = class interval size.

Example 3.3

Calculate an interpolated mode for the following data,

70-79	20
60-69	25 (modal interval)
50-59	21

Modal interval is the interval with the highest frequency that is 60-69, where the preceding interval is 50 – 59 and the following interval is 70 – 79. To calculate the interpolated mode for this data.

LRL = 59.5, d_1, = 25-21 = 4, d_2 = 25 – 20 = 5 and i = 10.

$M_0 = 59.5 + \frac{4}{4+5}$ x 10

= 59. 5 + 4/9 x 10

= 59.5 + 0.44 x 10

= 59.5 + 4.4

= $\underline{63.9}$

Therefore, the interpolated mode is 63.9. The interpolated mode within the modal class interval improves the estimate of the mode by allowing the adjoining frequencies to add their weight in arriving at a final estimate.

Standard deviation

The standard deviation is the average deviation between an observed score and the mean of a distribution. Alarape (2005) explains standard deviation to be a measure of the distance between each score and the mean of that particular distribution. The standard deviation (SD) is determined by taking the square root of the variance or

SD = $\sqrt{\frac{\sum x^2}{n}}$, where "*x*" is a deviation from the mean and "n" equals the sample size.

To calculate the standard deviation use the following steps;

1. First, find each deviation (*x*) from the mean x-$\overline{X}$, where x is the score and $\overline{X}$ is the mean.
2. Second, square each deviation, finding x^2
3. Third, sum the squared deviations, finding $\sum x^2$
4. Fourth, divide the sum by n, finding $\sum x^2/n$
5. Finally, extract the positive square root of the result of step 4.

Example 7

Let's say you are asked to calculate a standard deviation for the following scores: 6, 5, 3, 2.

Table 3.4: Showing Calculation of a Standard Deviation

X	$\overline{X}$	**X-** $\overline{X}$	**(X-** $\overline{X}$ **)^2**
6	4	2	4
5	4	1	1
3	4	-1	1
2	4	-2	4
$\sum$ =**16**			**10**

$$SD = \sqrt{\frac{10}{4}}$$

$= \sqrt{2.5}$

$= 1.58$

Summary

In the lecture, the three measures of central tendency which include the mean, the median and the mode were discussed and calculated using the formulae for ungrouped and grouped data. Standard deviation was finally explained and calculated.

Post-Test

What is the major difference between the crude mode and the interpolated mode?

Reference

Alarape, A. I. (2005). Statistical Methods and Computer Application, in B. Udegbe, S. Balogun, H. Osinowo and G. Sunmola (eds). *Psychology: Perspectives in Human Behaviour*. Ibadan, Department of Psychology, University of Ibadan (Chapter 3), 59-88.

LECTURE FOUR

Scales of Measurement

Introduction

This lecture introduces to you the scales of measurement that are vital to understanding statistical techniques and statistical interpretation. Efforts are made in this lecture to simplify each of the four scales of measurement for your better understanding. You are therefore implored to be attentive as this lecture will greatly help you to master the statistical techniques discussed in this course.

Objectives

At the end of the lecture, you should be able to:

1. discuss the four scales of measurement in research and statistics; and
2. differentiate characteristics of the four scales of measurement

Pre-Test

1. Mention and explain the four scales of measurement.
2. Differentiate characteristics of nominal scale from the ordinal scale of measurement.

CONTENT

Scales of Measurement

When data are collected in a piece of research, they are usually based on some forms of measurement. Measurement in this sense refers to the human tendency to attach meaning to phenomena that are yet to be understood. Whether you measure some aspect of nature in the physical world or people’s beliefs in the social world, you try to categorize this in measurement. Such measurement can involve qualitative (descriptive) or quantitative (numerical) observations, objects, places, events, ideas etc. in some systematic manner. About four types of scales of measurement are commonly used. These are *nominal, ordinal, interval and ratio scales*.

The Nominal Scale

A nominal scale uses number to identify qualitative differences among measurements. The measurements made by a nominal scale are names, labels or categories, and no quantitative distribution can be drawn among them. In other words, nominal data are more of categories, labeling, numbering or identification that cannot be said to be greater than the other. Examples of variables or questions on nominal include gender (male or female), hobbies (please check any of the below activities you do during your leisure time), religion (Are you a Muslim, Christian traditionalist's?) etc. Coding male as 1 and female as 2 does not mean that female is higher or greater than male, rather, these are just for categorization or labeling or identification

The Ordinal Scale

An ordinal scale ranks or orders observations based on whether they are greater than or lesser than one another. Ordinal scales do not provide information about how close or distant observations are from one another; rather they show superiority. For example, when you observed who graduated with first class in the department of psychology, it must be a student with highest Cumulative Grade Point Average (CGPA). Other variables on ordinal scale include job status, which include Senior Middle and Junior management staff.

The Interval Scale

An interval scale is quantitative; contains measurable equal distances between observations but lack a true zero value or point. Interval scales are used when the distance between observations is measurable, equal and ordinal, but a true zero point is unnecessary. The basic mathematical operations of addition, subtraction, multiplication and division can be performed on data collected from interval scales when a zero score appears on an interval scale, however, its placement does not mean that information stops at this or that the respondent does not have such behaviour at all. For instance, aggression scale is on interval scale of measurement. One scoring zero on the scale does not indicate that one is not aggressive at all. However, you can only score low; indicating that you are low in aggression. Interval scales are commonly used for behavioural measurement; such behavioural variables include intelligence, anxiety, depression, job satisfaction, organizational commitment etc.

The Ratio Scale

A ratio scale ranks observation contains equal and meaningful intervals, and has a true zero point or score. The ratio scale incorporates all of the properties found in the previous three scales of measurement as well as absolute zero point. Here, a zero score is meaningful because it indicates a true absence of information. For instance, the zero measurement on a ruler means that there is no object being measured, just as reading of 0 miles per hour (mph) on a speed meter indicates that the car is not in motion. Examples of variables on ratio scale of measurement include weight, height, length etc.

Summary

When you measure some aspects of behaviour or people's belief, you try to put them in category in form of measurement; which can be qualitative or quantitative in manner. There are four commonly used scales of measurement which include nominal, ordinal, interval and ratio scales. Nominal scale measures data such as names, label, categories which are qualitative in manner. Ordinal scale measures data such as observations ordered or ranked which are qualitative in nature. Interval scale measures data of equal intervals between observation with no true zero point including behavioural measurements. Finally, ratio scale measures data that have equal intervals between observation with true zero point and mostly on physical measurement.

Post-Test

1. Names variables that can be measured with ratio scale of measurement
2. What is the major difference between interval scale and ratio scale of measurements?

References

Alarape, I. A. (2005). Statistical Methods and Computer Applications, *Psychology Perspectives in Human Behaviour*, Revised and Enlarged Edition, (Chapter 3), 59 – 88/

Dunn, D. S. (2001). *Introduction to Statistics and Data Analysis as Tools for Researchers, Statistics and Data Analysis for the Behavioural Sciences*, (Chapter 1), 3-43

LECTURE FIVE

Hypotheses Testing

Introduction

In the present lecture, you shall be introduced to how you can formulate hypothesis in your research to meet your study objectives. The various types of hypothesis will be discussed and how you can accept or reject your stated hypothesis explained to you comprehensively.

Objectives

At the end of the lecture, you should be able to:

1. differentiate types of hypotheses stated in research; and
2. explain how to state a null and alternative hypothesis in research.

Pre-Test

1. Mention the two types of hypotheses
2. What is the major difference between the null hypothesis and the alternative hypothesis?

CONTENT

Hypotheses Testing

Any research is expected to be guided with statement of hypotheses that must have been derived from literature review. A hypothesis is a tentative statement of expectation of the research (Alarape 2005). In other words, a hypothesis proposes a relationship between two or more variables. However, hypothesis testing compares sample data and statistics to known or estimated population parameters. In hypothesis testing, data are collected from a sample of the population, which will either refute or support the hypothesis. In a nutshell, the researcher wants to be sure that he is not in error with his or her prediction. The researcher wants to prove that the null hypothesis (H_o) is false and that the alternate hypothesis is true. There are two main types of hypothesis statement mostly stated in the literature. These are Null Hypothesis and Alternative hypothesis.

The Null Hypothesis (H_0)

The null hypothesis traditionally indicates that all the population parameters in an experiment, which are represented by sample statistics, are equal. In other words, the null hypothesis predicts that a given independent variable will not cause a change or effect on the dependent variable. For example, a null hypothesis may state that there is no relationship between shoe size and sociability. There is no difference in academic performance of male and female students of University of Ibadan. All these aforementioned are examples of null hypotheses indicating no effect or relationship between variables.

The Alternative Hypothesis (H_1)

The alternative hypothesis specifies that a difference exists between the population parameters identified by the null hypothesis. In other words, the alternative hypothesis predicts that a given independent variable will cause a change or effect in the dependent variable. In alternative hypothesis, there are conviction; degree and direction of the relationship between or among two or more variables as predicted by the research. Examples of alternative hypothesis are as follows: male students will significantly record higher academic performance than female students in University of Ibadan. There will be a significant positive relationship between job satisfaction and job performance among employees of distance learning centre. Self esteem will have a significant effect on life satisfaction among retired Nigerian civil servants. However, alternative hypothesis that stated that there is a relationship between variables can either be classified into directional or non-directional hypothesis.

A *directional hypothesis* specifies the exact nature or direction of the relationship between variables. The example that there will be a significant job performance among employees of distance learning centre indicates the direction of the relationship as "positive". The hypothesis suggests that the higher the job satisfaction, the higher the job performance among the employees.

A Non-directional hypothesis anticipates that a difference will exist between variables but does not specify the nature, or direction of that difference. In other words, a non-directional hypothesis only specifies that there will be a difference between two variables or groups, but does not state which group will be higher on the dependent variable than the other. For example, a non-directional hypothesis is stated when you say that there will be a gender difference on job satisfaction. The stated hypothesis does not indicate if it is the male or the female that will be higher in job satisfaction.

Statistical Decision/Significance

A difference between means is described as being statistically significant; that is, one mean is larger in value than another. Significance testing entails using statistical tests and probabilities to determine whether sample data can be used to accept or reject a null hypothesis involving variables. The word significance in the context of statistics; refers to whether a result is statistically reliable, one that is sufficiently trustworthy so that an investigator can reasonably reject the null hypothesis. Thus, you are able to decide whether to accept or reject the null hypothesis. In social sciences, the minimum level at which you can reject H_0 (i.e. accept H_1) is at .05 level of significance. This suggests that the maximum measurement error should not exceed 5% so that one can be sure that the result is not due to chance. When H_1 is accepted, it means that the probability of being wrong in taking this decision is less than 5% (i.e. $P<.05$). On the other hand, when H_1 is rejected or H_0 is accepted, it means that the probability of being wrong in taking this decision is greater than 5% (i.e. $P>.05$).

Inferential Errors

By inferential errors, it means a situation where you accept H_0 as true when you should reject it; or one where you reject H_0 as false when the appropriate decision is to accept it. These errors are in two categories: Type I and Type II errors.

Type I Error involves rejecting the H_0 (i.e. a researcher believes that a significant difference exists) when, in fact, there is actually no difference.

In a nutshell, type I error is committed by rejecting H_0 when it is in fact true.

Type II Error involves accepting H_0 (i.e. a researcher believes that no significance difference does exists) when, in fact, there actually is a difference. In short, Type II error is committed by accepting H_0 when in fact it is false.

Summary

In the lecture, procedures for testing hypothesis were introduced. Two types of hypotheses were stated which are the null hypothesis and the alternative hypothesis. Furthermore, under the alternative hypothesis that states that there is a relationship between variables, there are directional and non-directional hypothesis. Finally, level at which a null hypothesis stated can be rejected or accepted was discussed to be .05 level of significance.

Post-Test

1. Define a directional hypotheses
2. Differentiate between a directional and a non directional hypothesis with examples.

Reference

Alarape, I. A. (2005). Statistical Methods and Computer Applications, *Psychology Perspectives in Human Behaviour,* Revised and Enlarged Edition, (Chapter 3), 59 – 88/

LECTURE SIX

Correlations

Introduction

The present lecture introduces you to some forms of correlational analyses. The correlation analyses discussed and worked in this lecture include Pearson r and Spearman rank order correlations. The two correlation analyses discussed in this lecture have peculiarities in their use. In other words, data where Pearson correlation is applicable, a spearman rank order correlation will not be appropriate.

Objectives

At the end of this lecture, you should be able to:

1. explain and calculate Pearson correlation; and
2. explain and calculate spearman rank order correlation

Pre-Test

What are the differences between Pearson and Spearman rank order correlations?

CONTENT

Correlations

Correlation is a statistical tool or technique that is used to know whether there is a relationship or association between two variables (e.g. variables X and Y). It is calculated basically to know whether the nature of relationship between two variables is positive, negative or zero. Positive correlation suggests that as X increases, Y also increase or as X decreases, Y also decreases and this can be referred to as direct relationship.

Negative relationship occurs when X increases and Y variable decreases; and this is referred to as indirect or inverse relationship between x and y variables. The correlation coefficient of 1.00 indicates a perfect linear relationship between the two variables. Apart from the direction in correlation, there is what is called the degree of the relationship where coefficient ranging between .60 and 1.00 as "strong"; those ranging from .40 and .59 as “moderate” while those ranging from .01 and .39 as “weak” respectively. Finally, it is worth noting that in correlation, there is no cause and effect relationship; meaning X variable does not cause Y variable; rather, they could only be related. Within the scope of this course, you will be limited to Pearson r and Spearman rank order correlation (rh_0).

The Pearson Correlation

The Pearson correlation which is also called the Pearson product-moment correlation; specifically enables the researchers to assess the nature of the relationship between two variables, X and Y. Correlations are based in pair of variables and each pair is based on the responses of one person. In other words, scores on the two variables X and Y are gotten from same person (s). It is commonly denoted by “r”; and to calculate the Pearson correlation, there is computational formula for this

$$r = \frac{N\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{\left(N\Sigma X^2 - (\Sigma X)^2\right)\left(N\Sigma Y^2 - (\Sigma Y)^2\right)}}$$, the degree of freedom (df) = N-2

Example 6.1

The following data are given as a set of scores for 10 students who participated in a scale validation study (i.e. fraudulent intent and lie scales).

Table 6.1: Showing Scores on Fraudulent Intent scale and Lie Scale

X (Fraudulent)	20	5	18	6	19	3	4	3	17	18
Y(Lie scale)	8	2	10	3	8	4	3	2	7	9

First, you need the following to be able to calculate Pearson correlation based on the formula: the sum of X ($\sum X$), the sum of Y ($\sum Y$), the sum of the product (times) of X and Y ($\sum XY$), the sum of the square of X ($\sum X^2$) and the sum of the square of Y ($\sum Y^2$). Based on the example given,

$\sum X = 113$, $\sum Y = 56$, $\sum XY = 831$, $\sum X^2 = 1793$, $\sum Y^2 = 400$, $(\sum X)^2 = 12769$,

$(\sum Y)^2 = 3136$ and N = 10.

$$r = \frac{10\,x831-(113)(56)}{\sqrt{(10\,x1793-12769)(10\,x400-3136)}},= \frac{=8310-6328}{\sqrt{(17930-12769)(4000-3136)}},= \frac{=1982}{\sqrt{(5161)(864)}}$$

$$= \frac{1982}{\sqrt{4459104}},= \frac{1982}{2111.66}$$

r = 0.94 approximately to 2 decimal point

df = N-2, = 10 –2, = 8

Interpretation

At this stage, consult your statistical table to determine whether the calculated value is significant or not; using a (df) of 8 with critical value at .05 level of significance (one tail test) which is 0.55. This means that our calculated value (· 94) is greater than the critical value (. 55); suggesting that the H_0 is rejected. In other words, there is a positive relationship between fraudulent intent scale and lie scale ($r = .94$, $df=8$, $p<.05$). This is described, in that, the higher the fraudulent intent, the higher the tendency to tell lie among the students.

However, to determine the degree of the relationship between fraudulent intent and lay scales, there is a need to calculate coefficient of determination; that is denoted by r^2. It explains the variation in one variable that can be accounted for by variation in the other variable. In other words, one can say variable X predicts variable Y. When r^2 is multiplied by 100, it gives us percentage of the variance in Y that is associated with or accounted for by variance in X. Based on our r= .94, the coefficient of determination r^2= $(.94)^2$ =.88, when 0.88 is multiplied 100, it gives us 88.36% of variance in lie scale that is associated with or accounted for by variance in fraudulent intent scale.

The Spearman Correlation

Another type of correlation coefficient that is used to determine the nature of association or relationship between two ordinal data is called Spearman rank order correlation coefficient denoted by 'rho'. When actual values of the variables given are not available, the data may be ranked from 1 to N in order of size, importance, etc; and this statistical technique can be used to determine the coefficient. It is appropriate when the relationship between two variables is not linear but shows consistency of direction. In other words, small increase in X will lead to dramatic increase in Y. Here is the formula for the Spearman rank order correlation.

rho = 1 - $\frac{6\sum D^2}{N\left(N^2 - 1\right)}$, where D = the difference between ranks of the values of X and Y, N is the number of pairs of value (X,Y) in the data given.

Example 6.2

The following is given as the scores of 10 students in a debate. The scores are from lowest to highest.

Table 6.2: Showing the debate scores of the students

Debate Scores	Rank (R)
4	10
6	9
8	8
11	7
12	6
13	5
16	3.5 (average of position 3rd and 4th)
16	3.5 (average of position 3rd and 4th)
18	2
19	1

Table 6.2 shows how you can rank a set of scores.

Example 6.3

Use the following data to calculate the spearman rank order correlation, the procedure described above can be used for the exam and test scores of 10 PSY 319 distance learning students.

Table 6.3: Showing the Exams and Test scores for the students

Test scores	R_1	Exam score	R_2	$D=(R_1 - R_2)$	D^2
4	10	40	10	0	0
6	9	50	9	0	0
8	8	55	8	0	0
11	7	60	7	0	0
12	6	63	6	0	0
13	5	65	4	1	1
16	3.5	65	4	-0.5	0.25
16	3.5	65	4	-0.5	0.25
18	2	72	2	0	0
19	1	80	1	0	0
					ΣD^2 =0.50

$$\text{rho} = 1 - \frac{6\sum D^2}{N(N^2-1)}, \quad = 1 - \frac{6 \times 0.50}{10(10^2-1)}, = 1 - \frac{3}{10(100-1)}, = 1 - \frac{3}{10(99)}$$

$$= 1 - \frac{3}{990}, = 1 - 0.003, \qquad = 0.997$$

rho = 1.00

Interpretation

Check the statistical table to know whether the obtained calculated value is significant or not using the actual N as degree of freedom (df) at .05 level of significance (one – tail test). The critical value is 0.50. Since our calculated value (1.00) is greater than the critical value (.50) we reject H_0 and accept $H_{1.}$ This indicates that there is a significant relationship between the students' test and exam scores (rho=1.00;P<.05).

Summary

In the lecture, correlation analyses were discussed and calculated. Specifically, Pearson and Spearman rank order correlations were explained and calculated with their formulae. Examples were given and were calculated with interpretations.

Post-Test

What is the difference between Pearson and Spearman rank correlations?

Reference

Alarape, I. A. (2005). Statistical Methods and Computer Applications, *Psychology Perspectives in Human Behaviour*, Revised and Enlarged Edition, (Chapter 3), 59 – 88/

LECTURE SEVEN

The Chi-Square Test

Introduction

The present lecture introduces you to another statistical method commonly used in psychology. The chi-square test is a non-parametric statistics that can be used for nominal data. The two types of chi-square test discussed and calculated in this lecture are the chi-square test for goodness of fit and chi-square test for independence.

Objectives

At the end of this lecture, you should be able to:

1. discuss when to use chi-square test; and
2. differentiate the chi-square test for goodness of fit from test for independence.

Pre-Test

Mention the two commonly used chi-square test

CONTENT

The Chi Square Test

The chi-square (X^2) test is used when you have categorical or nominal data that can be changed to frequency. Given its non-parametric nature, the chi – square test needs not be applied to data that conform to any particular shape (i.e. normal distribution), though the observations must be nominal. It is useful in testing hypothesis about the form and shape of frequency distribution. One underlying assumption the chi – square does have, is that observations are randomly selected from some large population. Also, the number of expected observations within a given category should be reasonably large. For the scope of this course, explanation will be made on chi – square test for goodness of fit and chi – square test for independence. The steps involved for testing chi-square include the following:

a. verify that the available data are based on a nominal or an ordinal scale of measurement;
b. state null and alternative hypotheses; use a significance level of .05;
c. perform analysis with the use of the chi-square formula to determine whether to accept or reject the null hypothesis and compute any supporting statistics.
d. The general formula for chi-square is:

$$X^2 = \sum (fo - fe)^2 / fe,$$

df= c-1, where c is the number of category.

Chi-Square Test for Goodness of Fit

The chi-square test for goodness of fit is a one-variable test; that tests whether obtained observations conform to (i.e. ‘fit’) or diverge from the population proportion specified by a null hypothesis. Let us consider a specific example to illustrate the chi – square test for goodness of fit.

Example 7.1

A lecturer asked 35 students in his class to complete a standard course evaluation by responding to “*statistic was my favorite class this semester*”. A student is expected to tick only one of the five rating options (i.e. strongly agree, agree undecided, disagree, strongly disagree). The following is the analysis of the students’ responses.

Observed Data

SA	**A**	**U**	**D**	**SD**
17	8	3	2	5

Please note that the lecturer elected to treat the responses as categorical data; perhaps he might have easily calculated a mean rating based on the 35 responses. Though, looking at the data, it seems to indicate that many students strongly agree and agree with the statement.

However, you need to statistically demonstrate this positive pattern of agreement deviates from what would be expected by chance. The expected data is the one specified by the null hypothesis. The chi-square test indicates whether there is a difference between some observed frequencies (the data drawn from a piece of research) and expected frequencies. In this example, the above scores are the observed frequencies (fo) of the students, you need to calculate the (fe) to arrive at the formula.

Table 7.1: Showing Chi-square for goodness of fit for the students responses.

	fo	**Fe**	**(fo-fe)**	**$(fo\text{-}fe)^2$**	**$(fo\text{-}fe)^2/fe$**
SA	17	7	10	100	14.29
A	8	7	1	1	0.14
U	3	7	-4	16	2.29
D	2	7	-5	25	3.57
SD	5	7	-2	4	0.57

$\sum (fo - fe)^2 / fe$ **=20.86**

To obtain the (fe) for the students, the arithmetic mean of (fo) is calculated. Then if (17+8+3+2+5) divided by 5 (number of score) will give you 7. Therefore,

$\sum (fo - fe)^2 / fe$ =14.29+0.14+2.29+3.57+0.57=20.86. The df=c-1, i.e. 5-1=4.

Interpretation

Now check your statistical table to know if to reject or accept Ho using .05 level of significance. The critical value is 9.488. You can observe that our calculated value (20.86) is greater than the critical value (9.49), we reject H_o and accept $H_{1.}$ (X^2 =20.86; df=4; P<.05). In other words, the result shows that the students significantly agreed with the statistics in that semester.

Chi-Square Test for Independence

The chi-square for independence indicates whether the frequencies associated with two variables (with two or more categories each) are statistically independent upon one another. This type of chi-square is used when you are interested in the relationship between two variables. Each observation is placed in one-and only one cell representing a joint relationship between one category from each variable. The null hypothesis usually predicts that there is no relationship between the two variables (i.e. the variables are independent of each other). In other words, the response on one variable is not related to the response on the second variable. However, if there is relationship between the two variables, it means that response of one variable depends on the other variable.

Example 7.2

A media specialist wants to find out how people learn about major new events generally. Do they read the newspaper or watch television? In turn, does their level of education influence the source they rely upon for the news? They specifically randomly sampled 206 adult residents from a community; asking them to indicate their educational status and their primary source of news. The data gathered are presented below:

Table 7.2: Showing educational status and sources of news

Educational Status	**Television**	**Newspapers**	**Total**
Secondary	47_a	62_b	109
Graduate	58_c	39_d	97
Total	**105**	**101**	**206**

Note that you can state a null hypothesis or alternative hypothesis based on the given example. For instance, H_0= educational status and news source are independent of each other. $H_{1=}$There is a relationship between educational status and news source. For the above data, you first label each of the cells with alphabetical letters, to identify each cell in order to avoid mix up when calculating them one by one e.g. a, b, c, & d respectively. The same computational formula for calculating the chi-square is applicable here but differs in arriving at the expected frequency (fe) and the degree of freedom (df) i.e. (r-1) (c-1). Fe = (row total) (column total)/n, where "n" is the sample size. In solving the example,

Fe_a=109x105/206=55.56

Fe_b=109x101/206=53.44

Fe_c=97x105/206=49.44
Fe_d=97x101/206=47.56
df= (2-1)(2-1)=1x1=1

Table 7.3: Showing Chi-square for Independence for the Participants

Fo	**Fe**	**(fo-fe)**	**$(fo\text{-}fe)^2$**	**$(fo\text{-}fe)^2/fe$**
47	55.56	-8.56	73.27	1.32
62	53.44	8.56	73.27	1.37
58	49.44	8.56	73.27	1.48
39	47.56	-8.56	73.27	1.54

$\sum (fo - fe)^2 / fe$ **=5.71**

Interpretation

Now check your statistical table to know if to reject or accept Ho. Our df is 1 under .05 level of significance. The critical value is 3.841. You can observe that our calculated value (5.71) is greater than the critical value (3.84), we reject H_o and accept $H_{1.}$ (X^2 =5.71; df=1; P<.05). The result shows that the two variables are not independent of one another. In other words, sources of news depend on level of education.

Summary

The chi-square (X^2) test is used when you have categorical or nominal data that can be changed to frequency. The mostly used types of chi-square tests are test for goodness of fit and test for independence. The chi-square formula for the two is the same but arriving at the fe and degree of freedom differs.

Post-Test

Does the calculation of fe for the two chi-square tests differs?

Reference

Dunn, D. S. (2001). *Introduction Statistics and Data Analysis as Tools for Researchers, Statistics and Data Analysis for the Behavioural Sciences*, (Chapter 1), 3-43

LECTURE EIGHT

T-Test

Introduction

In this lecture you will be introduced to another statistical method in psychology that is used for group comparison. Two types of t-test will be discussed and calculated in this lecture. The two types of t-test are t-test for independent samples and t-test for repeated measures.

Objectives

At the end of this lecture, you should be able to:

1. test a hypothesis using t-test for independent samples and repeated measures; and
2. differentiate the two t-tests and know when to use them.

Pre-Test

Why using t-test for independent samples?

CONTENT

T-Test

In the use of t-tests, you examine situations where two separate samples are drawn in order to compare the mean difference between the two samples. In other words, t-tests are used to compare two sample means, but not more than two. Because it is a parametric test, the t-test was designed for data analysis only when certain conditions are met which constitute the assumptions underlying the t-test.

Assumptions underlying the t–test

a. The sample data are drawn from a normally distributed population; not skewed distribution.

b. The data are either randomly sampled or individually sample from a larger population.

c. The dependent variable must be based on either interval or ratio scales

d. The samples are presumed to come from populations that have equal variances.

The two variations of the t-test presented in this lecture are t-test for independent samples and t-test for repeated measures.

T-Test for Independent Groups/Samples

The t-test for independent samples is specially designed to detect significant differences between two groups of subject in a study. An independent group design is one in which the researcher randomly selects subjects from the population of interest and randomly assigns them into two groups (experimental and control groups). It is basically used to compare the mean difference in two groups, thereby assessing whether the independent variable elicited an anticipated behavioural change in one group not the other. An example is given below to show the procedure in testing a hypothesis using the t-test for independent groups.

Example 8.1

A marriage counselor wants to investigate how regional background of a spouse affects marital success. He randomly assigned the participants into two groups where Group A constitute marriage partners from the same region of the country (i.e. either north or south) and Group B comprises of those whose marriage partners are from different regions of the country. Ratings of the success of the marriage (the higher the rating, the better the marriage) are obtained. Calculate this using t-test for independent samples.

Table 8.1: Showing Spouses' Marital Success Ratings

Group A	10,	9,	8,	8,	9,	10,	9,	8,	7,	8,
Group B	6,	7,	8	8	7	7	6	5	8	6

Hypothesis like

H$_0$: $\bar{X}_1 - \bar{X}_2 = 0$ (i.e spouse region does not affect marital success).

H$_1$: $\bar{X}_1 - \bar{X}_2 \neq 0$ (i.e spouse region affects marital success).

The formula for t-test for independent samples is:

$$t = \frac{X - Y}{\sqrt{\left(\frac{\frac{\varepsilon x^2 - (\varepsilon x)^2}{n_1} + \frac{\varepsilon y^2 \, (\varepsilon y)^2}{n_2}}{(n_1 + n_2) - 2}\right)\left(\frac{1+1}{n_1 \; n_2}\right)}}$$

where

$\bar{X}$ = Mean rating of marital success for the first group

$\bar{Y}$ = Mean of rating of marital success for the second group

εx^2 = Sum of the square of the first group.

εy^2 = Sum of the square of the second group.

$(\varepsilon x)^2$ = The square of the sum of the scores in the first group.

$(\varepsilon y)^2$ = The square of the sum of the scores in the second group

n_1 = Number of scores in the first group

n_2 = Number of scores in the second group

Degree of freedom (df) = $n_1 + n_2 - 2$

$\Sigma X = 86$, $\Sigma X^2 = 748$, $\bar{X} = 8.6$, $\Sigma Y = 68$, $\Sigma Y^2 = 472$, $\bar{Y} = 6.8$, N=10

$$t = \frac{8.6 - 6.8}{\sqrt{\left(\frac{748 - \frac{(86)^2}{10} + 472\frac{(68)^2}{10}}{(10+10) - 2}\right)\left(\frac{1+1}{1010}\right)}}$$

t = 1.8/0.45

t= 4.0

df = $n_1 + n_2$ $^{-2}$ = 10 + 10 –2 =18

***Note*:** The + value may be negative, but it is not important; rather the absolute value of the t-value.

Interpretation

From the calculation, our t-value or calculated value is 4.0. This will be compared with the critical value (table value) in the statistical table using a (df) of 18 at one-tail test of 0.5 level of significance. The critical value is 1.734. Since our calculated value (4.00) is greater than the critical value (1.73), we will accept H_1 and reject H_0. This suggests that there is a significant difference between the two groups. t (18) = 4.0; P <.05. The interpretation is that spouses from same region ($\overline{X}$ = 8.6) are more successful in marriage than those from different region ($\overline{Y}$ = 6.8).

T-Test for Repeated Samples

In t-test for repeated measures, as the name applies, scores are being observed on same subject repeatedly. In other words, two measures obtained must have come from the same participant and that is why the number of subjects in the two groups must be the same (equal). This type of statistical method is appropriate for pretest-post-test design; and it can be used to control for extraneous variables. The formula for t-test for repeated measures is given below:

$$t = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{\Sigma D^2 - \frac{(\Sigma D)^2}{N}}{N(N-1)}}}$$

Where,

$\overline{X}$ = is the mean of the first group.

$\overline{Y}$ = is the mean of the second group.

D = is the difference between pair of X and Y scores.

N = is the number of pairs of scores.

Example 8.2

A researcher intends to investigate the influence of an anti-smoking campaign on attitudes of students toward cigarette. He randomly selects 20 students and randomly assigns them to experimental and control groups. An attitude scale is administered to both groups and their attitude toward cigarette is measured. The researcher then introduces an anti-smoking campaign to the students in the experimental group for a month of series of speaker, class discussion on the health implications of smoking. After this, their attitudes towards smoking were measured again. However, the control group was not exposed to any anti-smoking campaign. The following are the scores obtained in the two conditions.

Table 8.2: Showing Pre-Test and Post-Test Measures of Attitudes Toward Smoking

Pre-test (X)	**Post-test (Y)**	**D**	**D^2**
21	25	-4	16
29	23	6	36
22	20	2	4
36	38	-2	4
20	25	-5	25
18	22	-4	16
20	18	2	4
29	31	-2	4
28	26	2	4
30	33	-3	9

$\sum D^2 = 122$, $\sum D = -8$, $\overline{X} = 25.3$, $\overline{Y} = 26.1$, $N = 10$

$$t = \frac{25.3 - 26.1}{\sqrt{122 - \frac{(-8)^2}{10}}}, \quad = \frac{-0.8}{\sqrt{\frac{122 - \frac{64}{10}}{10 \ x \ 9}}}$$

$$\frac{-0.8}{\sqrt{\frac{122 - 6.4}{90}}}, \quad \frac{-0.8}{\sqrt{\frac{115.6}{90}}}, \quad = \frac{-0.8}{\sqrt{1.13}}$$

$$t = 0.71$$

df = N – 1 = 10 – 1 = 9

Interpretation

From the calculation, our t-value (calculated value) is 0.71. This will be compared with the table value using (df) = 9 at one-tail test of 0.5 level of significance. The critical value is 1.83. Since our calculated value (0.71) is less than the critical value (1.83) we accept Ho and reject Hi. This suggests that there is no significant difference in the two conditions, (t(9) = 0.71; P>0.5). The interpretation of this is that the anti-smoking campaign did not have significant effect on the students' attitudes towards smoking. In this case, the researcher needs to examine his intervention strategy to identify what could have gone wrong.

Summary

In the lecture, two types of t-test were discussed and calculated. The two t-test tests were t-test independent samples and t-test repeated measures.

Post-Test

Why do we use t-test for repeated samples?

LECTURE NINE

Analysis of Variance (ANOVA)

Introduction

The lecture introduces you to another statistical method called analysis of variance (ANOVA). ANOVA can be one-way or factorial. The one-way ANOVA involves one independent variable while the factorial ANOVA may be two or more independent variables having levels.

Objectives

At the end of this lecture, you should be able to:

1. explain when the use of ANOVA is appropriate; and
2. how to interpret ANOVA results.

Pre-Test

What is the difference between one-way ANOVA and factorial ANOVA?

CONTENT

Analysis of variance (ANOVA)

Analysis of variance is a parametric statistical method used for determining whether significant difference exists in an experiment containing three (3) or more sample means. As a statistical method, ANOVA divides variability, attributing portions of it to the effect of an independent variable (IV) on a dependent variable (DV). The assumptions underlying the use ANOVA include:

a. The population where the samples were drawn from must be normally distributed.

b. There should be random selection and random assignment of samples to experimental group.
c. There should be equal variance on the groups. In this course, discussion will be based on one-way ANOVA for independent groups and for repeated measures.

One-Way ANOVA for Independent Groups

A one-way ANOVA or one-variable ANOVA for independent groups is a statistical method for analyzing the variation found within the various levels of a single independent variable. In other words, it is used to identify differences among sample means of an independent variable. A one-way ANOVA is used for an independent variable occurring at least 3 levels; and subjects are randomly selected and assigned into different groups.

Example 9.1

A researcher wants to investigate the relative sociability scores of different major (Economics, Sociology and Psychology) in a university. The following scores were obtained on the sociability scale of 8 items from the students based on their major. The researcher then hypothesized: H_0: There will be significant effect of major on sociability; H_1: Psychology major will be more sociable than Economic and Sociology majors.

Table 9.1: Showing Students' Scores on Sociability Scale

Econs	0	1	3	5	0	3	5	1
Soc	3	5	6	6	3	6	5	6
Psy	5	7	9	9	7	9	5	9

Steps in Computation of a one-way ANOVA include:

a. Compute the sums of each group by adding the scores to obtain the grand total ($\sum X$) and the grand mean ($\overline{X}$) for each group;
b. Compute the sum of X^2 ($\sum X^2$) for each group
c. Add all the groups' sums to obtain the $\sum\sum X$. (iv)
d. Add all the sum of squares ($\sum X^2$) to obtain the $\sum\sum X^2$.

You will arrive at the following for each major:

	Econs(X_1)	Soc(X_2)	Psy(X_3)
$\overline{X}$	= 2.25	5	7.5
$\sum X$	= 18	40	60
$\sum X^2$	=70	212	472

$\sum\sum X = 118$

$\sum\sum X^2 = 754$

N = 8 + 8 + 8 = 24

You need to first calculate the correction term (CT) i.e.

$$CT = \frac{(\Sigma\Sigma X)^2}{N} = \frac{(118)^2}{24} = 580.17$$

$SS_t = \sum\sum X^2 - CT = 754 - 580.17 = 173.83$

$$SS_b = \frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \frac{(\Sigma x_3)^2}{n_3} - CT, \quad =$$

$$\frac{(18)^2}{8} + \frac{(40)^2}{8} + \frac{(60)^2}{8} - 580.17$$

$$= \frac{324}{8} + \frac{1600}{8} + \frac{3600}{8} - 580.17, \quad = 40.5 + 200 + 450 - 580.17$$

= 690.5 – 580.17

= <u>110.33</u>

$SS_w = SS_t - SS_b$

= 173.83 – 110.33

=<u>63.5</u>

Compute the degree of freedom (df) for each of the components (SS_t, SS_b, SS_w) df_t = total number of subjects minus 1, i.e. = N – 1 = 24 – 1 = <u>23</u>

df_b = number of levels/groups minus 1, i.e. = K – 1 = 3 – 1 = <u>2</u>

$df_w = df_t - df_b$, = 23 – 2 = <u>21</u>

Computes the mean square (MS) for each of the components

$MS_t = \frac{SS_t}{df_t}$ (Not needed for the analysis)

$MS_b = \frac{SS_b}{df_b} = \frac{110.33}{2} = \underline{55.17}$

$MS_w = \frac{SS_w}{df_w} = \frac{63.5}{21} = \underline{3.02}$

Compute the F-ratio:

$$F = \frac{MS_b}{MS_w} = \frac{55.17}{3.02} = \underline{18.27}$$

You then need to present your result in a table.

Table 9.2: Showing a One-way ANOVA of Effect of Majors on Sociability among Students

Source	SS	df	MS	F	P
Between Groups	110.33	2	55.17	18.27	<.05
Within Groups	63.5	21	3.02		
Total	173.83	23			

Interpretation of Result

From the calculation, our F-ratio is 18.27, you now consult the statistical table to determine either to reject or accept Ho. The dfs for the numerator and denominator are 2 and 21 respectively. Using the dfs at one tail-test of .05 level of significance; the critical value is 3.47. Since our calculated F-ratio (18.27) is greater than the table value (3.47), we accept H_1 and reject H_0. This suggests that major of students have influence on the sociability $F(2,21) = 18.27$; $P<.05$. The interpretation of this is that psychology major ($\overline{X}$ = 7.25) are more sociable than Economic ($\overline{X}$ = 2.25) and sociology ($\overline{X}$ = 5.0) majors. This confirms the stated hypothesis.

One-Way ANOVA for Repeated Measures

In one-way ANOVA for repeated measures, each of the same subjects is exposed to different treatment or experimental situations. This type of one-way ANOVA is advantageous because it requires few numbers of subjects. However, there may be problem of carry-over effect where the subject might be progressively more proficient at performing the task that might be due to learning and not the different experimental situation.

The statistical assumptions underlying the one-way repeated measures ANOVA are.

a. those dependent variables are based on either a ratio or an interval scale of measurement.
b. that the data collected within each level of the independent variable are independent of one another.
c. that the distribution of the population from which the measures in each level of the IV are drawn is assumed to be normal.
d. there is homogeneity (equal) of variance.

Example 9.2

An experiment was designed to test whether the difference in the performance of machines', each experienced operators work, on each of the machines for equal times. The number of units produced per machine was recorded for each operator. Test the hypothesis that there is a difference between the machines at significance levels of .05. The data collected for the experiment is as follows.

Table 9.2: Showing the Number of Units Produced by Five Machines

Subject	A	B	C	D	E
1	8	2	3	5	3
2	3	3	2	1	3
3	7	3	4	2	1
4	6	5	5	4	3
5	4	6	2	1	2

Steps in computation of One-way ANOVA for repeated measure:

a. Compute the sums for each treatment by adding the scores to obtain grand total $\sum X$ and the grand mean $\overline{X}$ for each treatment.
b. Obtain the sum of the squared values ($\sum X^2$) for each treatment.

c. Add all the treatments sums to obtain the $\sum\sum X$.

d. Add all the sum of squares ($\sum x^2$) to obtain $\sum\sum X^2$

	A	B	C	D	E
$\overline{X}$	= - 5.6	3.8	3.2	2.6	2.4
$\sum X$	= 28	19	16	13	12
$\sum X^2$	= 174	83	58	47	32

$\sum\sum X=88$

$\sum\sum X^2 = 394$

N = 5 + 5 + 5 + 5 = 25

Obtain the correction term (CT),

$$C.T = \frac{(\Sigma\Sigma x)^2}{N`}$$, where N is the total number of observations.

$$\frac{(88)^2}{25} = \frac{7744}{25} = 309.76$$

$SS_t = \sum\sum X^2 - C.T$

$= 394 - 309.76$

$= \underline{84.24}$

$$SS_{treat} = \frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \frac{(\Sigma x_3)^2}{n_3} + \frac{(\Sigma x_4)^2}{n_4} + \frac{(\Sigma x_5)^2}{n_5} \text{ - C.T}$$

$$= \frac{(28)}{5} + \frac{(19)^2}{5} + \frac{(16)^2}{5} + \frac{(13)^2}{5} + \frac{(12)^2}{5} - 309.76$$

$$= \frac{784}{5} + \frac{361}{5} + \frac{256}{5} + \frac{169}{5} + \frac{144}{5} - 309.76$$

$= 156.8 + 72.2 + 51.2 + 33.8 + 28.8 - 309.76$

$= 342.8 - 309.76$

$= \underline{33.04}$

Add the five scores recorded for each subject as follows: 21, 12, 17, 23, 15. Square each of the sums, add them together and divide the result by the number of scores as follows:

$$= \frac{(21) + (12)^2 + (17)^2 + (23)^2 + (15)^2}{5}$$

$$= \frac{1628}{5} = 325.6$$

To obtain sum of square for subjects (i.e. SS_{sub}), subtract the correction term from the value obtained in the immediate previous stage. This is, 325.6-309.76=15.84.

To obtain the sum of squares error (i.e. SS_{error}), subtract SS_{treat} and the SS_{sub} from the SS_t. That is,

$SS_{error}=SS_t-SS_{treat}-SS_{sub}=84.24-33.04-15.04=\underline{36.16}$

Compute the degrees of freedom (df) for each of he components (i.e. $SS_{t,}$ SS_{treat},$SS_{sub,}$ SS_{error})

Df_t= number of measures minus 1, i.e. 25-1=24

Df_{treat}=total number of treatment given minus 1, i.e.5-1=4

df_{sub}=total number of subjects used minus 1, i.e. 5-1=4

$df_{error}=df_t-df_{treat}-df_{sub}$ i.e. 24-4-4=16

Compute the mean squares for each of the components

$MS_t=SS_{t/}df_t$ (not needed for analysis)

$MS_{treat}=SS_{teat}/df_{treat}$, i.e. 33.04/4=8.26

$MS_{sub}=SS_{sub}/df_{sub}$, i.e. 15.84/4=3.96

$MS_{error}=SS_{error}/df_{error}$, i.e. 36.16/4=2.26

Compute the F-ratio

$F=MS_{treat}/MS_{error}$=8.26/2.26=3.65

Now, you need to present your result in a Table

Table 9.3: Showing a one-way ANOVA of Group Difference in the Machine Output

Source	**SS**	**Df**	**MS**	**F**	**P**
Subject	15.84	4	3.96	3.65	<.05
Treatment	33.04	4	8.26		
Error	36.16	16	2.26		
Total	**85.04**	**24**			

Interpretation

From the calculation, our F-ratio is 3.65, you now consult the statistical table to determine either to reject or accept Ho. The dfs for the numerator and denominator are 4 and 16 respectively. Using the dfs at one tail-test of .05 level of significance; the critical value is 3.01. Since our calculated F-ratio (3.65) is greater than the table value (3.01), we accept H_1 and reject H_0. This suggests that there is a difference in performance among machines $F(4,16) = 3.65$; $P<.05$. A significant F value does not mean that the means are significantly different from each other until carried out in multiple comparison. However, since there is a significant difference, there is a need to compute a statistical procedure called multiple comparisons to know the extent of significant difference in the means.

Summary

In the lecture, efforts were made to explain and work out two types of One-way ANOVA with comprehensive examples. A one-Way ANOVA is a one-variable statistical method but occurring at levels.

Post-Test

What is the difference between the two one-way ANOVA discussed in this lecture?

Reference

Alarape, I. A. (2005). Statistical Methods and Computer Applications, *Psychology Perspectives in Human Behaviour,* Revised and Enlarged Edition, (Chapter 3), 59 – 88/

Note: *All students are advised to further consult statistical textbooks for continuous practice. Do not base your knowledge of statistics on the course material alone.*