5-1-1975

Introductory statistics for social workers

Greg Alan Bergman
University of Nebraska at Omaha

INTRODUCTORY STATISTICS
FOR SOCIAL WORKERS

by

Greg Alan Bergman

RESEARCH PROJECT

Presented to the Faculty of

The Graduate College in the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Master of Social Work

Graduate School of Social Work

Under the Supervision of Ronald H. Ozaki, D. S. W.

Omaha, Nebraska

May, 1975

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to Dr. Ronald Ozaki, whose suggestions have been invaluable throughout the preperation of this text.

I would also like to express my gratitude to Drs. Briggs and Jones, whose teachings in statistics provided me with the information needed to begin such a project.

A deep gratitude goes to my wife, Susan, for her suggestions and support throughout the work on this project. This text would not have been possible without her patience and indulgence.

Finally, my sincerest gratitude goes to my parents, Ivan and Marjorie Ann, to whom I would like to dedicate this text.

INTRODUCTION

An Increase in the Utilization of Research

In a 1969 article (Weinberger, R., & Tripodi, T., 1969, p. 439), it was concluded that between the years of 1956 and 1965, "the number and the percentages of research articles increased over time." This same conclusion can be drawn by looking at the quantity of citings of research articles in social work journals between the years of 1969 to 1974. In 1969, 32 citings were made, as compared with 41 citings in 1973, and 45 citings in 1974. In looking at Psychological Abstracts, 30 citings were made in 1965, as compared with 123 for 1973.

Not only are more research articles being written in recent years, but also more emphasis is being placed on the application of research for the social work practioner. According to a recent article (Bose, 1973), among the key prerequisites for effective planning of social welfare services is the availability of reliable and valid statistical data and information. This article also goes on to say that research is needed at every stage of the planning process: diagnosis of social situation, identification of needs, determination of planning objectives, development of policy, setting of priorities, formulation of programs and services, monitoring of implementation, and evaluation of progress.

Research is also important when using the casework approach as a social worker. In an article entitled "Implications of Research For the Goals of Casework" (Reid, 1970, p. 141), it was suggested that the effectiveness of casework may be blunted by a tendency of caseworkers to pursue treatment goals that are excessively global and overambitious. The author concluded that recent "research suggests that modes of casework practice directed at the definition and resolution of very specific target problems in brief courses of treatment may be more effective than typical approaches currently in use." (Reid, 1970, p. 141).

Northwood and Reed (1969, p. 165) state that research in race relations can be used to provide knowledge to help "implement the institutional values of the profession," and to establish feasible and realizeable change goals. Research has also been used to anlyze current welfare programs (Weiss & Rein, 1969; Williams & Evans, 1969; Weiss, 1970), and in the administration of social action programs (Tripodi, Epstein, & MacMurray, 1970).

The two conclusions that can be drawn from these facts are that recent articles are being utilized more by social workers; and research is being used by social workers in casework, groupwork, social planning, and administration.

The Need for Research in Social Work Curriculum

The core courses of all accredited social work schools consist of Practice, Human Behavior in the Social Environment (HBSE), Social Welfare Policy and Services (SWPS), Practicum, and Research. Research is a vital part of all of these courses. In Practice, research is constantly being undertaken to expand the knowledge of this area. (Flanagen, 1971; Segal, 1972; Reid, 1970). In HBSE, new theories are being developed daily in the areas of Personality and Human Development theories. (Maddi, 1972; Hall & Lindzey, 1970; Johnson & Medinnus, 1969). Many recent articles illustrate the use of research in the SWPS area. (Orlans, 1971; Weiss & Rein, 1969; Engstrom, 1970; Gil, 1968).

Research is also utilized in Practicum. As the student leaves the classroom environment and enters the working world, all of the previously tested theories are put to work. Even the Research course utilizes research as is illustrated in such articles as "The Evaluation of Social Work Teaching Competence: Issues, Pseudo-Issues, and Future Directions" (Feldman, 1970), and also in terms of the value of teaching research to students (Gil, 1968 and Lewis, 1969).

If social work is to "become a learned profession, social work education must have as its goal the scientifically oriented and committed professional practitioner." (Gil, 1968). Thus, the need for research in the student's curriculum goes unquestioned. In addition, the central concerns of research courses should be directed toward preparing a professional practitioner to be "appreciative of science, to develop in his professional self-attitude an understanding likely to promote a spirit of inquiry, and to undertake descriptive studies of limited scope, answering questions arising out of social agency practice." (Lewis, 1969).

The Council on Social Work Education's Standards for the Accreditation of Baccalaureate Degree Programs in Social Work (Effective July 1, 1974) states that "An educational program that prepares for beginning professional practice shall demonstrate that it: ...provides content in ... social research." (C.S.W.E., 1974, p. 13).

Not only is research mandated, but also recent articles suggest that the "student involvement in research should constitute a basic ingredient of all graduate study, and teaching should therefore include the guidance of students in developing constructive research undertakings related to specific curriculum areas." (Gil, 1968).

If a student is to "develop constructive research undertakings," (Gil, 1968), a basic knowledge is needed on the student's part in Statistics. It is the purpose of this text to provide the student with the knowledge needed in this area, as an introduction into social work research. The final goals of this author are to develop in the student the appreciation and understanding of the utilization of research in the social work profession, and to give the student the skills needed to undertake independent research.

Organization of the Text

Most statistic's textbooks (Cain, 1972; Armore, 1966; Hoel, 1971; Hope, 1967; Noether, 1971) dwell on methodology. Unless the student is a mathematic's major, and very few social work students have a business undergraduate major, about 7% in 1960 and 2% in 1966 (Golden, Pins, & Jones, 1972, p. 20), this does not prove very useful for the practitioner.

Therefore, this text will give the student both the methodology for finding a statistic, as well as the "uses" of each statistic given. These uses will represent why it would be used, when it would be used, and what it represents.

The purposes of this two-fold approach are: to provide the student with the skills to analyze data, and to provide information which will be utilized every time he reads a professional journal article that has any statistical analysis in it.

The text is divided as follows:

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I.

STATISTICS

What is Statistics?

The term statistics brings to the minds of most people visions of charts and graphs, endless columns of numbers, and other methods of showing how the government is spending money. "At one time the word referred exclusively to numerical information required by governments for the conduct of state." (Noether, 1971). But how does one go about defining statistics?

One way may be to use the definition espoused by mathematicians, who have on occasion defined mathematics as "being those professional activities engaged in by mathematicians." (Anderson & Bancroft, 1952). Thus, statistics comprises those professional activities engaged in by statisticians. This, of course, leaves one a bit bewildered.

Webster (1960) defines statistics as "facts or data of a numerical kind, assembled, classified, and tabulated so as to present significant information about a given subject."

For our purposes, statistics, when used in the plural, will mean a set of numerical data or numerical records. When used in the singular, the term statistic is sometimes properly used to denote a particular entity in a set of numbers or a particular item in a numerical record.

A statistic is also a computed measure, such as an average, a t-value, a correlation, or a Chi-Square value, based on a sample.

Why Study Statistics?

Statistical methods and concepts have much to offer social work students. It is generally recognized that research workers require the use of statistical methods to plan research projects and evaluate research findings. However, an understanding of statistical theory, methods, and concepts is of substantial importance to others in the profession of social work.

Social workers not directly involved in research need a good understanding of statistics to comprehend much of the material contained in professional journals. Social workers need not be highly trained statisticians in order to read and understand many of the published professional articles. They need not be skilled statisticians, either, if they aspire towards research work. However, knowledge of elementary statistical methods, concepts, and terminology is necessary when the social worker attempts an independent research project. He will then be able to better formulate his research problem. He will also be able to more fully profit from the statistical results, and will be better able to understand how to interpret the results within the framework of the research problem.

Applicability of Statistical Methods

Statistical methods have a wide application. For example, statistical methods are used in the fields of social work, "psychology, education, sociology, anthropology, engineering, medicine, business, economics, and accounting." (Armore, 1966).

Statistical methods are applicable to numerical data. There is no limit to the kind and quantity of numerical data to which statistical methods can be applied. When a problem needs to be solved, a decision needs to be made, or a question needs to be answered, it is not always necessary to employ statistical methods. However, when the need to use statistical methods arises, it is always because there is a problem to be solved, a decision to be made, or a question to be answered.

Generally, a most important part of the application of statistical methods is to carefully and precisely define the problem to be solved, the decision to be made, or the question to be answered. Once this is done, statistical methods may then be applied.

CHAPTER II.

TERMS, CONCEPTS, AND OPERATIONS

Populations and Samples

Statistical methods deal with two types of data: population data and sample data. Population can be defined as the complete set of data relating to an area of interest. For example, if a social worker wanted to know the age distribution of all welfare recipients at his agency, then the ages of all the recipients make up the population, a population of ages. If there are 1500 recipients at the agency, then there is a population of 1500 ages.

If these 1500 recipients consisted of 1000 females and 500 males, a new term can be introduced. For example, if the worker was interested in the age distribution of the recipients as well as the seperate age distributions for male and female recipients, then the 1500 ages make up the population, while the 500 ages of the males and the 1000 ages of the females make up subpopulations.

The illustration presented represents finite populations. These are populations which are in some way limited and contain a determinable number of items. Finite populations may be small or they may be large.

On the other hand, a population may be infinite. An infinite population contains an unlimited or indefinitely large number of items. For example, researchers may be interested in the results of a treatment method for alcoholics. The population for this study may be infinite because it includes all alcoholics in this country today and for years ahead. Some texts use the term universe instead of population. These terms may be used interchangeably.

If a set of data contains less than the full scope of information in which you are interested, it represents a sample. For example, if a community organization worker is interested in the effects of a new program for trailer court residents in the county, of which there are 800, but obtains information on only 95, then the 95 trailer court residents represents a sample of the population of the 800 trailer court residents. Hence, a sample represents a part of the population.

It should be noted that a collection of data may represent a population from one point of view and a sample from another. For example, the 1500 welfare recipients are a population in terms of the agency, but are only a sample in terms of the total number of welfare recipients in the United States.

Descriptive and Inductive Statistics

Descriptive statistics provide methods to organize, summarize, and describe sets of data which represent populations. In other words, descriptive statistics is the term used when statistical methods are applied to a set of data which comprise the entire population, when complete information (i.e. the total population) is available.

Usually social workers and other researchers deal with sample data rather than population data. This is quite understandable when it is realized that collection of data for an entire population is often expensive, time consuming, and, sometimes, impossible. Sample data are collected when it is desired to estimate one or more population characteristic. Use of sample information to estimate population characteristics is a form of inductive inference. That is, it involves the reasoning from the particular (sample) to the general (population). For example, in a sample of families from a community, 68% have one or more children in school. This finding is generalized to the population by estimating that 68% of all families in the community have one or more children in school. When statistical methods are applied to sample data in order to make generalizations to the population, the methods used are called inductive statistics.

In conclusion, all statistical methods are directed towards description of a population. Descriptive methods are used to describe a population based on population data. Inductive methods are used to describe a population based on sample data. In addition, when one wants to just describe sample data, descriptive methods may also be used.

Variables and Symbollic Notation

A variable is a quantity that varies; or, it is a quantity that takes on different values. Variables fall into two categories: discrete or continuous. A variable is discrete if it can take on only specified values. For example, a die can only take on numbers which are not less than one, nor greater than six. Furthermore, only the specified values: 1, 2, 3, 4, 5, or 6 are possible. Values in between are impossible. Hence discrete variables contain gaps between them.

A *continuous* variable is one which can take on any value between specified limits. For example, a person's height may be 62 inches, or we might measure it as 61.7 inches, or 61.68 inches, or 61.67917 inches. A continuous variable is a theoretical concept.

A quantity which can have only a single value is called a *constant*.

Each field of knowledge develops its own language. In the field of statistics, the student needs to learn not only new technology, but also a new way of communicating. He must learn to use *symbols*, such as Greek letters, to represent variables and operations.

There are certain conventions which are generally followed in the use of symbols. Usually, the letters X, Y, Z, or the lower case letters x, y, z are used to represent variables. For example, if we measure the height (in inches) of three men as 70, 68, and 71, we may let X=height (which is the variable we are concerned with here). We can then let X_1, X_2, and X_3 stand for the height of each of the three men. The symbol, X_1, is read "X sub 1" or "X one." Then X_1=70 inches, X_2=68 inches, and X_3=71 inches.

As a convention, the early letters of the alphabet, usually lower case (a, b, c, or d) are used to denote constants. The letters k and g are also often used to denote constants. Sometimes the upper case letters A, B, C, D, K, or G are used to represent constants.

The foregoing discussion was meant to introduce the idea of symbollic notation. Other symbols, representing variables and operations, will be introduced as needed.

Measurement Scales

Measurement implies a scale. For example, to measure a person's height, we may use a scale, such as a yardstick, which tells how much height he has.

A scale may be thought of as a device for measuring: a foot rule is a scale, a speedometer is a scale, and, of course, a weighing machine is a scale. According to conventional views (Hammond & Householder, 1962), there are four types of scales used for measuring: the nominal scale, the ordinal scale, the interval scale, and the ratio scale.

Nominal Scales

The most elementary scale of measurement is a classificatory scale, called a nominal scale. The use of this scale involves setting up categories which are clearly defined and delineated, but in such a way that the only measure of a characteristic available from the scale is equivalence or equality. For example, we may classify college students according to their class level, using the nominal scale categories: Freshman, Sophomore, Junior, Senior, or Graduate. Students who fall into the same category are equivalent (equal).

That is, equivalent with respect to the characteristic measured -- class level. There is no other measurement information available from this scale other than equivalence (e.g. belonging to the same category).

The scale elements in a nominal scale may be denoted by names, numbers, letters, or any set of symbols. The only purpose served by the symbols is to uniquely identify the categories making up the scale. As an example, we could use the nominal scale when referring to the races of man as follows: Caucasoid, Negroid, and Mongoloid. Or we might set up the scale as: A, B, C, etc. In this case, we define: A=Caucasoid, B=Negroid, etc. Or we could use numbers: 1=Caucasoid, 2=Negroid, etc. We could use any other symbol, if we wished to. The use of symbols to denote scale elements in a nominal scale in no way implies more than merely category identification. Thus, measurement with the use of a nominal scale involves only the seperation of elements into mutually exclusive categories.

Ordinal Scales

The next highest scale of measurement is the ordinal scale. This is a scale of rank or relative importance. An example is the scale of titles assigned to faculty members in a university: Professor, Associate Professor, Assistant Professor, and Instructor. These titles make up an ordinal scale, since they indicate relative importance in faculty status.

An ordinal scale contains two degrees of measurement: equivalence and relative importance (greater than or less than).

In our example, faculty members with the same title, say Associate Professor, are equivalent (equal) in status. As we proceed from one end of the scale (Professor) to the other end of the scale (Instructor), there is a constant decline in relative status.

The scale elements in an ordinal scale may be expressed in terms of any set of symbols which provide the property of ranking. For example, if we wish to rank three students on their intelligence levels, we might set up a scale as follows: smart, smarter, smartest. Or we might use A, B, C, where A=smart, etc. Or we might use 1, 2, 3, where 1= smart, etc.

It should be noted that, while number or letter sequences may be used in the construction of an ordinal scale, as well as a nominal scale, the information supplied is different. The use of 1, 2, or 3 for a nominal scale provides only the information that there are three categories. In the ordinal scale, this information is provided, plus the information of relative importance. In the example of intelligence levels, 3 identifies an intelligence level category different from 1. Furthermore, 3 indicates more intelligence than is indicated by 1 or 2.

There are times when 3 will not always indicate more of something than 1 or 2. If, by definition, 3=smart, 2=smarter, and 1=smartest, then 1 would indicate more intelligence than is indicated by 2 or 3. In an ordinal scale, the number used to indicate the category is only a <u>symbol</u>, which is why the symbol 1 may indicate a higher relative status than the symbols 2 or 3.

Measuring with an ordinal scale rather than a nominal scale is usually preferable, because ordinal scale measurements provide data in a form which permits more, and better, analysis of data. The fact that ordinal scales do imply direction is, of course, an advantage. Nevertheless, although ordinal scales indicate greater than and less than, they provide no information as to "how much" of a difference exists. Where such information is requires or desireable, an interval scale must be employed.

Interval Scales

An interval scale is one for which it is possible to quantify the intervals, the distances, which lie between the numbers constituting the data. "Quantifying the intervals" means expressing their length or extent in terms of some fixed, well-understood unit of measurement. For example, suppose an exam in a human behavior course containing 25 questions of equal difficulty (assuming this is possible) is given to a class of students. If each correct answer scores one point, then test scores could range from 0 to 25. If this test measures knowledge in the human behavior area, then the scale of test scores (0, 1, 2, 3, ...25) represents an interval scale. An interval scale, then, contains three degrees of measurement: equivalence, relative importance, and a unique unit of measurement with a known distance between any two intervals.

Because of the existence of a unique unit of measurement, the interval scale provides a measure of the distance (interval) between scale values. In the example noted above, if two students obtained scores of 15 and 20, respectively, we could observe that the second score is 5 points higher than the first. Moreover, with an interval scale, ratios of scales intervals are meaningful. For example, consider test scores of 10, 15, and 25, which indicate intervals of 5 and 10 points. The interval ratio is 10/5=2. This indicates that the third score (25) exceeds the second (15) by twice as much as the second exceeds the first (10).

When available, interval scales are of enormous advantage over nominal scales and ordinal scales, because most of the ordinary procedures of arithmetic can be applied to the numerals which represent points on the scales. Interval scale numbers can be added, subtracted, multiplied, and divided, for example.

There are, however, several limitations of an interval scale. One limitation is the lack of an absolute zero point. For example, if a student did not answer any questions correctly on the human behavior exam, his score would be zero. However, this could hardly be taken to mean that he has no knowledge of the subject area. In an interval scale, the zero point is arbitrary. An additional limitation is the fact that, in an interval scale, the ratio between scale values is not meaningful. For example, if one student scored 10 on the exam, and another scored 20, it can not be concluded that the second student has twice as much knowledge as the first student. The scale values of an interval scale can only be expressed numerically.

Ratio Scales

The highest level of measurement is a ratio scale. A ratio scale has all of the properties of an interval scale, plus an absolute or true zero point. An interval scale can be converted into a ratio scale if it can be shown that it has an absolute zero or origin from which measurements are made. This zero must be absolute or real in the sense that there is only one possible point on the scale which may represent it. The scales on which we measure height and weight are examples of ratio scales. If an object is said to weigh zero pounds, we mean that this object has no weight.

In addition to the ratio between two scale intervals being meaningful, the ratio between two scale values are also meaningful in a ratio scale. This is due to the fact that the true zero point provides a measure of the interval (distance) of a given scale value from this true zero point. For example, a height of 20 yards means an interval of twenty yards from zero height. In the same sense, suppose two objects weigh 50 and 10 pounds, respectively. Then the ratio between values is 50/10=5. It can be concluded that the heavier object is five times as heavy as the lighter one. The scale values in a ratio scale can only be expressed numerically.

Please note Appendix A for a summary of measurement scales.

Rules for Summation

One of the key operations to the development and application of statistical procedures is the <u>summation operation</u>. Many of the statistical procedures which we will encounter in later chapters (e.g. Central Tendency, Variations, and Correlation) typically require the addition (summation) of sets of terms. For example, let X stand for "scores on an accreditation exam," and let X_i then represent a particular score. Let us consider four specific scores: X_1, X_2, X_3, X_4; then we may represent the summation of these four scores by the symbol: $\sum_{i=1}^{4} X_i$. The capital Greek letter $\sum$ (sigma) denotes the summation operation. The symbol: $\sum_{i=1}^{4} X_i$ is read as "the sum of X_i, i varying from 1 through 4." We may show the meaning of this summation symbol algebraically as follows:

$$\sum_{i=1}^{4} X_i = X_1 + X_2 + X_3 + X_4.$$

Suppose $X_1=10$, $X_2=16$, $X_3=14$, and $X_4=9$. Then we may write $\sum_{i=1}^{4} X_i = 10 + 16 + 14 + 9$ or $\sum_{i=1}^{4} X_i = 49$.

The letter i is called the <u>index</u> of the summation. The range of this index is specified by showing the lower limit (1) below the summation sign, $\sum$, and the upper limit (4) above it. In the above illustration, the <u>range of the summation index</u> is from 1 through 4.

The summation notation may be used to specify sums of any number of terms. In general, we may speak of N terms.

For example, we may speak of summing N X-scores. That is, when N=4, we mean 4 score; when N=500, we mean 500 scores, etc. Generally, then, we specify the summation of N X-scores as:

$$\sum_{i=1}^{N} X_i = X_1 + X_2 + X_3 + \dots + X_N.$$

There are three simple rules of summation which are helpful and frequently used in the development and application of statistical procedures:

Rule One states that the summation of a sum, or difference, of terms is equal to the sum, or difference, of the individual summations of each term.

RULE ONE $$\sum_{i=1}^{N} (X_i + Y_i - Z_i) = \sum_{i=1}^{N} X_i + \sum_{i=1}^{N} Y_i - \sum_{i=1}^{N} Z_i.$$

Rule Two states that the summation of a term involving a constant, K, times a variable, X_i, is equal to the constant times the summation of the variable.

RULE TWO $$\sum_{i=1}^{N} KX_i = K\sum_{i=1}^{N} X_i.$$

Finally, Rule Three states that the summation of a constant is equal to the constant times the number of times it is included in the summation.

RULE THREE $$\sum_{i=1}^{N} K = NK.$$

CHAPTER III.

FREQUENCY DISTRIBUTIONS

Organization of Raw Data

Very frequently, a researchist will find himself faced with a large mass of unorganized or raw data. For example, suppose that the researchist has measured the height of twenty-eight male college students, and recorded the heights, rounded to the nearest inch, as they were obtained. His record sheet would look like this: 68, 70, 64, 74, 71, 68, 66, 69, 70, 70, 67, 74, 67, 69, 71, 67, 69, 72, 68, 71, 69, 70, 67, 68, 68, 70, 69, and 70.

It is quite difficult to get a very adequate idea of the characteristics of this group of data from such a list of raw data. Since we know that height is a variable with characteristics of a ratio scale, it makes sense to put the scores in the order of their size: 64, 66, 67, 67, 67, 67, 68, 68, 68, 68, 68, 69, 69, 69, 69, 69, 70, 70, 70, 70, 70, 70, 71, 71, 71, 72, 74, and 74. This ranking does add to the meaningfulness of the set of heights. For example, it is now apparent that the heights range from a low of 64 inches to a high of 74 inches. It may also be observed that most of the heights occur more than once. It can be seen that a large number of the males were 67, 68, 69, or 70 inches tall, for example.

Ranking is a form of organization, which adds systematic arrangement to the formerly unorganized collection of raw data. However, in general, only limited insight into the data is possible by merely ranking. What is needed is summarization, as well as organization. Table 3.1 presents a possible form of data summarization, including systematic arrangement. The table is arranged in two columns. The first column lists the possible heights within the range of heights (X_i). The second column shows the frequency (f_i) with which each height occurs among the twenty-eight subjects.

Table 3.1 Simple Grouping of Heights of 28 Male College Students in Inches.

Score X_i	Frequency f_i
64	1
65	0
66	1
67	4
68	5
69	5
70	6
71	3
72	1
73	0
74	2

Another way of presenting the distribution in Table 3.1 is by means of a segmented line, or linear scale:

Frequency	1	0	1	4	5	5	6	3	1	0	2
Height	64	65	66	67	68	69	70	71	72	73	74

The summarization of the twenty-eight heights in Table 3.1 we will call simple grouping, because it is the simplest form of score grouping possible, and because no information is lost in the process. That is, the table leaves no doubt as to exactly which heights are included in the sample.

A simple grouping frequency table, such as Table 3.1, exhibits the <u>frequency distribution</u> of a collection of raw scores. The <u>frequency distribution</u> is a method for presenting a set of data in an economical fashion, a way that avoids the redundancy of listing identical scores. It will be observed that this summarization permits better understanding of how simple grouping provides an adequate understanding of the score distribution. Nearly always, further summarization is needed for a better understanding of the distribution of the data.

Further summarization of the data requires more condensed grouping of the scores. For example, Table 3.2 presents three <u>frequency tables</u> with a varying degree of summarization of the 28 heights in each.

Table 3.2 Frequency Tables of Heights of 28 Male College Students in Inches.

(A)		(B)		(C)	
Score Interval	Frequency (f_i)	Score Interval	Frequency (f_i)	Score Interval	Frequency (f_i)
64-65	1	64-66	2	64-67	6
66-67	5	67-69	14	68-71	19
68-69	10	70-72	10	72-75	3
70-71	9	73-75	2		
72-73	1				
74-75	2				

It will be observed that summarization beyond simple grouping is accompanied by some loss of information. However, the gain from adequate summarization in a well-constructed frequency table outweighs any loss of information. A frequency, with summarization beyond simple grouping, presents a more compact, more easily visualized representation of the frequency distribution.

In summarizing raw data, an important concern is how far to go. Table 3.3, generally based on the results of an article by H. A. Sturges (Sturges, 1926), summarizes the number of class intervals to use when constructing a frequency table.

Table 3.3 Approximate Number of Classes to Use When Constructing a Frequency Table.

Number of Items of Raw Data	Approximate Number of Classes to Use
15-29	5
30-59	6
60-99	7
100-199	8
200-499	9
500-999	10
1,000-1,999	11
2,000-3,999	12
4,000-7,999	13
8,000-14,999	14
15,000-34,999	15
35,000-69,999	16
70,000-149,999	17
150,000-299,999	18
300,000-499,999	19
500,000 and over	20

Construction of Frequency Distributions -- Frequency Tables

The primary rule to follow when constructing frequency distributions is to present the data in a form that makes sense and answers the questions most pertinent to the researcher. There are two commonly used means of representing frequency distributions: frequency tables and graphic representation.

Tally Tables and Simple Grouping

Construction of a frequency table begins with a clear understanding of the purposes the table is to serve. Generally, important uses of a frequency table are to study over-all distribution of the raw data, and to provide a tabular summary convenient for computation of various descriptive measures.

As a general rule, there is no need to organize the raw data into simple grouping tables. However, there may be instances when so little is known about the data, that simple grouping may be considered a desireable step. This may be accomplished through a simple procedure of setting up a tally table. In a tally table, scores are tallied in the appropriate cells, using vertical or slanted tally strokes.

To make a tally table, set up two cells marked "Scores (X_i)" and "Tally." Remember to list all scores from the lowest to the highest, even if the frequency is zero. Omission may lead to an erroneous impression concerning the score distribution, particularly if there are many scores with a zero frequency. Study of the tally table usually reveals various characteristics of the score distribution. As in simple grouping, the score range, and the scores with the highest frequencies will be quite evident. Table 3.4 is an example of a tally table using the data of the heights of the 28 male college students.

Table 3.4 Tally Table for a Simple Grouping: Heights of Male College Students.

Score (X_i)	Tally
64	/
65	
66	/
67	////
68	~~////~~
69	~~////~~
70	~~////~~ /
71	///
72	/
73	
74	//

Frequency Tables

Usually, a collection of raw data is summarized into a frequency table without the use of a simple grouping table. The first step in constructing a frequency table is to determine the number of items in the collection. We note that there are twenty-eight students whose heights have been obtained. Next, decide on the number of classes to use. Referring to Table 3.3, we observe that for 28 scores, five classes are suggested. Reviewing the raw data in Table 3.4, we note that the lowest score is 64 and the highest is 74. This indicates a range of 10, the difference between 74 and 64. Therefore, if we think in terms of summarizing the 28 scores into five classes, we must divide the full range of the scores, 10, into five equal parts (intervals). Thus, each class interval will be 2.

We are now ready to determine the discrete classes. This is accomplished by selecting the lowest discrete class limit.

In making this selection, it is best to select a round number equal to or lower than the lowest score and a multiple of the length of the class interval; which makes 64 a good selection, since it meets both of the guidelines. Then, by successively adding the length of the class interval to 64, the lower discrete class limits are found (e.g. 64, 66, 68, 70, 72, and 74). Table 3.5 shows how a frequency table is constructed. Note that the lower class limits are as we have just developed them. Then, the first discrete class becomes 64-65; the second, 66-67; etc.

Table 3.5 Frequency Table: Height of Male College Students.

Discrete Class Limits	Tally	Frequency (f_i)
64-65	/	1
66-67	~~////~~	5
68-69	~~////~~ ~~////~~	10
70-71	~~////~~ ////	9
72-73	/	1
74-75	//	2

It will be observed that there are six classes, even though five is the suggested number. There are several reasons for this: firstly, the length of the class interval decided upon and the score selected for the lowest discrete class limit are somewhat arbitrary; secondly, a sufficient number of classes must be established so that the highest discrete class can accomodate the highest score in the raw data. Choice of a convenient class length and selection of a suitable score for the lowest class limit may lead to a different number of classes than those suggested by Table 3.3, if the highest class is to accomodate the highest score.

We are now ready to determine the class frequencies. This is easily accomplished by tallying, as is shown in Table 3.5. Tallies are recorded in a manner similar to that explained for simple grouping. The only difference is that the tallies are recorded for any score, in our example, heights, falling between and including the discrete class limits. The total of these class frequencies should equal the total number of scores included in the raw data. This can then be used to check that all of the items were tallied. The two columns: "Discrete Class Limits" and "Frequency" provide all the basic data needed when constructing a frequency table.

Certain additional considerations need to be taken into account when constructing a frequency table. Overlapping discrete class limits must be avoided at all times. This occurs when the upper limit of one discrete class is the same as the lower limit of the next highest class. In addition, lower discrete class limits should be carefully selected since they stand out more than upper limits. Finally, frequency tables with classes of equal length are preferable during construction.

There are times when using classes of unequal length is considered more convenient. This may occur when the low, or high, end of the scale of values for the data stretches out with negligible frequencies. In this instance, it may be more convenient to summarize the string of values with very low frequencies into one or two classes. However, care is required when studying such a table, to take into account the variation in length of class intervals.

Sometimes, open-end class intervals are appropriate, especially when the raw data contains a few unusually low or high values. The following frequency table, calculated from a chart in Due and Friedlander (1973), shows an open-end class at each end of the table:

Table 3.6 Income Levels and the Number of Families in Each Category (1968).

Income	Families
Under $3,000	5,202,530
$3,000 to $4,999	6,111,710
$5,000 to $6,999	7,323,950
$7,000 to $9,999	11,819,340
$10,000 to $14,999	12,576,990
$15,000 and Over	7,475,480

The problem with open-end classes is that there is no way of knowing the range of the scores included. For example, the table indicates that there were 5 million families with an income under $3,000 in 1968, but it does not indicate either the frequencies for specific incomes under $3,000 nor the lowest income found.

The foregoing discussion for construction of a frequency table should be used with a great deal of thought. There is no such thing as a single, best-constructed frequency table. The ways to which the frequency table will be used must be carefully considered if the best table for a particular purpose is to be obtained.

Graphic Representation of Frequency Distributions

Graphic representations of a frequency distribution, as exhibited in a frequency table, assists in defining the shape of distributions. There are four common ways of representing frequency distributions graphically: bar diagrams, pie charts, histograms, and frequency polygons. The one used in a particular situation depends both on the kind of data to be presented, and on the needs for clarity of presentation.

In the case of categorical data, based on nominal scale data, bar diagrams and pie charts are usually used. The more common of the two, bar diagrams, are constructed from frequency distributions in the following way:

1. Frequency (f_i), proportions (p_i), and per cent (%) are placed along the vertical axis (ordinate or y-axis).
2. Score classes are placed on the horizontal axis (abscissa or x-axis).
3. Bars are constructed to the height of the class frequency, proportion, or per cent for each class.

Firstly, we ought to look at a hypothetical example to determine how proportions (p_i) and per cents (%) are found. If the total number of students enrolled in a social work program is 1,000 (N=1,000), and the frequency distribution is found to be Freshman=250, Sophomore=200, Junior=150, Senior=150, and Graduate=250, then this can be represented in Table 3.7.

Table 3.7 Class Levels, Frequencies, Proportions, and Per Cents of Students Enrolled in a Social Work Program.

Class	Frequency (f_i)	Proportion (p_i)	Per Cent ($\%_i$)
Freshman	250	.25	25.0
Sophomore	200	.20	20.0
Junior	150	.15	15.0
Senior	150	.15	15.0
Graduate	250	.25	25.0
	$\sum f_i=1{,}000=N$	$\sum p_i=1.00$	$\sum \%_i=100.0$

Once the frequencies in a frequency distribution are determined, proportions and per cents follow directly. The following formulas are used for conversion:

$$p_i = {}^{f}i/N \quad \text{and} \quad \% = 100p_i = 100\ {}^{f}i/N.$$

The proportion of observations in any class vary from 0 (f_i=0) to 1.00 (f_i=N). In some texts, proportions are called relative frequency. Both terms are interchangeable. The per cent, of course, vary from 0 to 100. Please note that f_i, p_i, and $\%_i$ are all discrete variables, since all of them are limited to a finite set of points for a given set of data. Frequency will always be a value between zero and N in discrete steps of one ($0 \leq f_i \leq N$). Similarly, a proportion will always be a value between zero and one in discrete steps of 1/N ($0 \leq p_i \leq 1$).

Figure 3.1 is a bar diagram which represents the data of Table 3.7. Note that the three scales are given in Figure 3.1 to emphasize that any one of the three could be used to present data in a bar diagram.

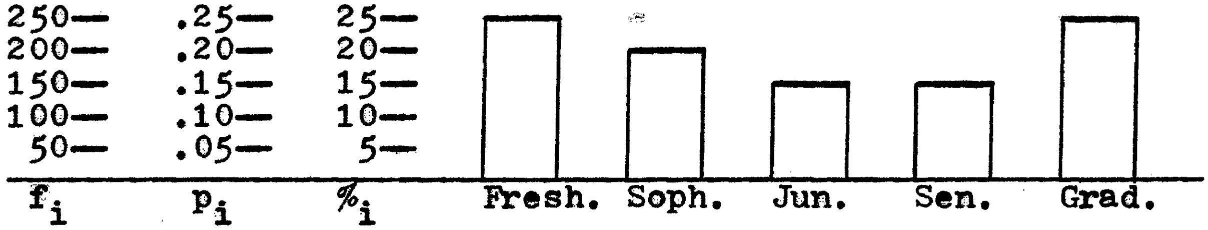


Figure 3.1 Bar Diagram of Class Levels of Students Enrolled in a Social Work Program.

The other way of presenting categorical data is by the use of a pie chart. This is illustrated in Figure 3.2. This method is especially useful in emphasizing the proportion of data falling into each category of an entire unit. The pie, of course, always equals 100 per cent.

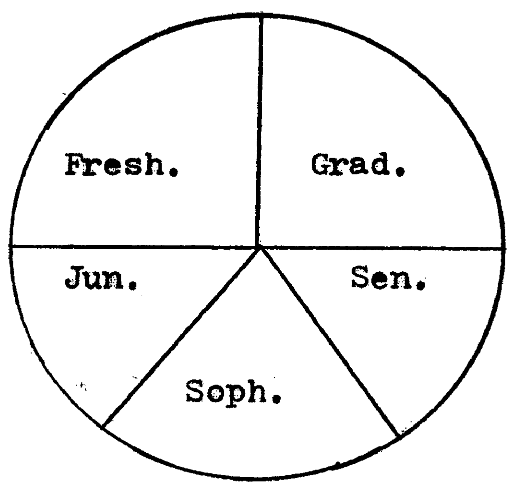


Figure 3.2 Pie Diagram of Class Levels of Students Enrolled in a Social Work Program.

Frequency distributions of numerical data are usually represented as either a histogram or a frequency polygon. A histogram is merely a bar diagram where the classes are ordered and adjacent on a common scale; the bars are, therefore, also adjacent, and extend from the lower limit of a class to its upper limit.

Histograms are constructed from frequency distributions in the following way:

1. Frequency (f_i), proportions (p_i), and per cents ($\%_i$) are placed along the vertical axis (ordinate or y-axis).
2. Score classes are placed on the horizontal axis (abscissa or x-axis). The scale is marked off, however, at points corresponding to scores halfway between the whole-number points. A height of 71.0 inches, for example, falls in the center of the lower real limit, 70.5, and the upper real limit, 71.5. Thus, the score scale on the abscissa is marked off to identify the continuous class limits.
3. The frequency for each class is plotted as a horizontal line over the full length of the class interval at the proper height corresponding to the frequency, proportion, or per cent.
4. The ends of the horizontal line are connected by vertical lines to the corresponding continuous class limits marked off on the abscissa.

The histogram has the appearance of a bar diagram, with the bars of equal width (equal class intervals) closely positioned. There is no space between bars because continuous class limits are used, thereby eliminating gaps between the class intervals. In addition, construction of a histogram implies the assumption that the scores included in a class are evenly distributed over the class interval.

Figure 3.3 is a histogram which represents the data of Table 3.1. It is the <u>area</u> of each rectangular bar in a histogram which represents the class frequency. Of course, with equal class intervals, as in our illustration, the height of the bar also represents the class frequency. However, it is more useful to think in terms of area rather than heights of the bars. The use of areas under a curve, such as a histogram, to represent frequency and especially proportion (relative frequency), is a basic notion in statistical inference, as will become apparent in later chapters. Histograms can be used to show frequency (f_i), proportions (p_i), and per cent ($\%_i$) of data.

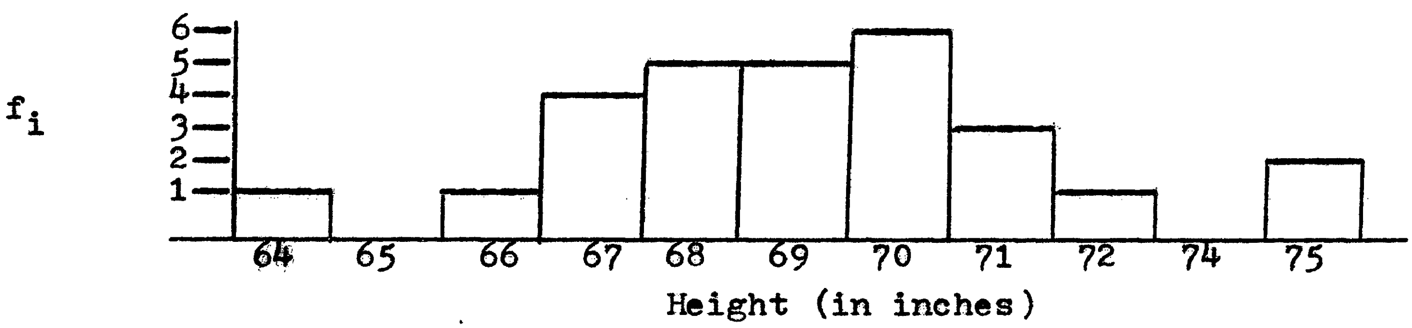


Figure 3.3 Histogram of 28 Heights.

<u>Frequency polygons</u> are also used to show frequency distributions of numerical data. In a frequency polygon, the horizontal axis (abscissa) contains the score scale, with the class midpoints specifically marked off on it. A frequency scale is marked off on the vertical axis (ordinate). The frequency for each class is plotted by placing a dot directly above the class midpoint, at the proper height corresponding to the frequency (or proportion or per cent), and the dots are connected by straight lines. It is customary to close the frequency polygon at both ends.

Figure 3.4 is a frequency polygon which represents the data of Table 3.1. Please note that plotting the total class frequency above the class midpoint means that the construction of the frequency polygon is based on the assumption that all scores included in a class are equal to the class midpoint. Finally, the <u>height</u> of the dot above the class midpoint represents the class frequency. This is in contrast to the histogram in which the area of each rectangular bar represents the class frequency (proportion or per cent).

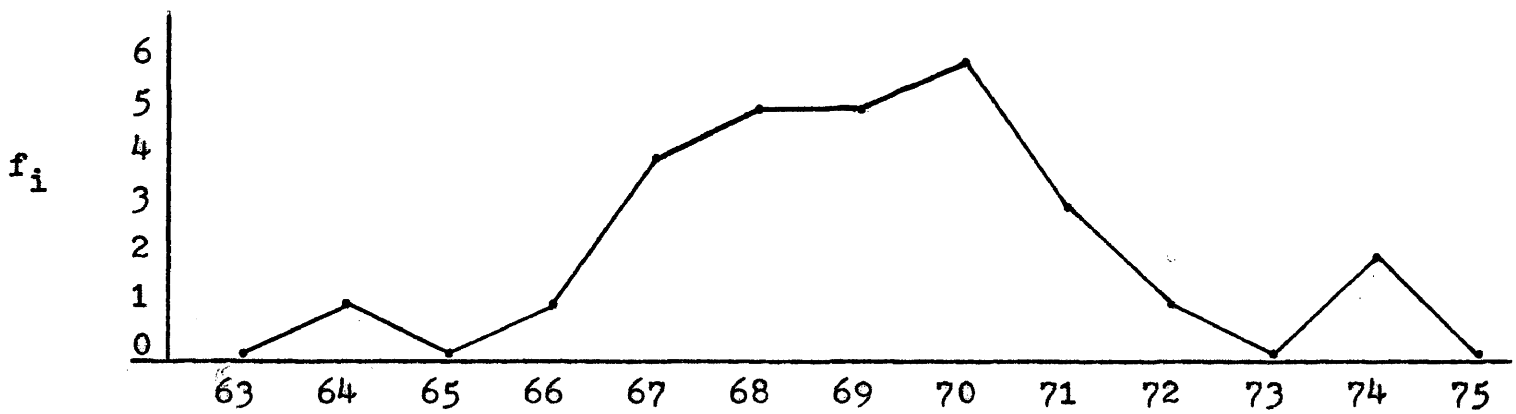


Figure 3.4 Frequency Polygon of 28 Heights.

Forms of Frequency Distributions

The shape of a frequency distribution may be expected to vary for different sets of data summarized into frequency tables. Figures 3.5 through 3.10 present histograms for various distribution types. Figure 3.5 presents a <u>rectangular distribution</u> in which the class frequencies are all equal. This is also called a <u>uniform distribution</u>. The histogram presents bars of equal height, indicating that the class frequencies are equal.

f_i

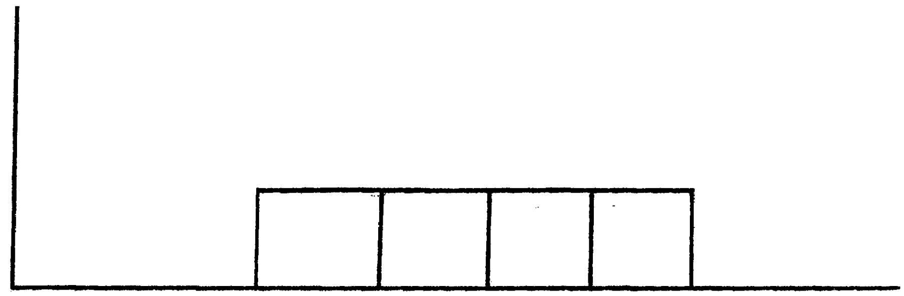


Figure 3.5 Histogram of a Rectangular Distribution.

Figure 3.6 presents a <u>symmetrical distribution with a central peak</u>. A distribution is symmetrical if one half of it can be folded over so that it is exactly superimposed over the other half. In other words, one side is a "mirror image" of the other. In the distribution of Figure 3.6, class frequencies continuously decline as one moves away from the middle class interval in either direction.

f_i

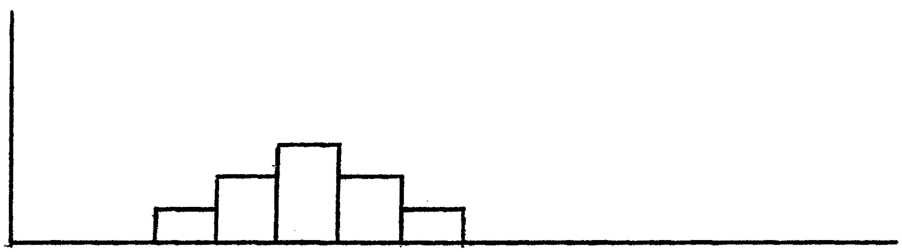


Figure 3.6 Histogram of a Symmetrical Distribution with a Central Peak.

Symmetrical distributions may also be bell-shaped, triangular, or any other shape, as long as one side is a mirror image of the other.

Figure 3.7 presents a <u>skewed distribution</u>. It will be observed from the histogram that the class containing the highest frequency is not centrally located, as in Figure 3.6, and that class frequencies continuously decline in either direction from this class. Moreover, the classes stretch out further in the downward direction, forming what may be called a "tail" of the distribution. Since this tail is in the left-hand portion of the histogram, this is called a <u>left-skewed distribution</u> or <u>negative-skewed distribution</u>.

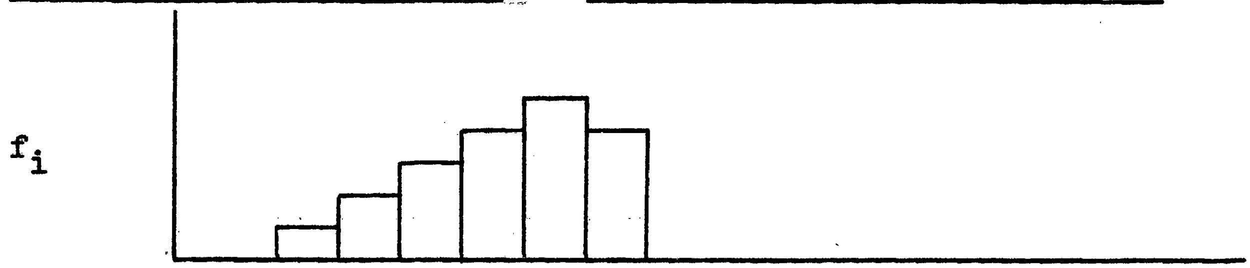


Figure 3.7 Histogram of a Left-Skewed Distribution.

Figure 3.8 presents a <u>right-skewed</u> or <u>positive-skewed</u> distribution. This distribution has the "tail" in the upper end of the histogram, and is a mirror image of a left-skewed distribution.

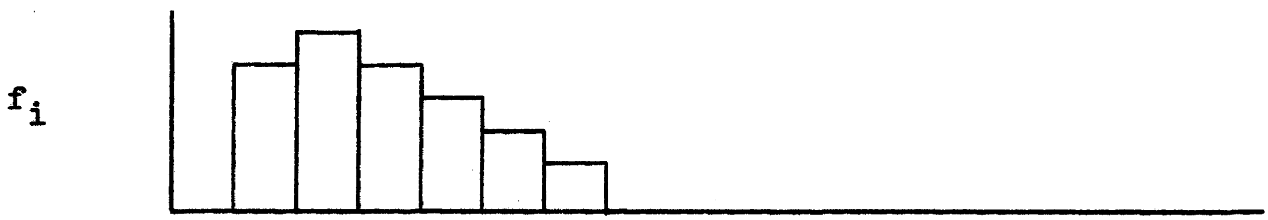


Figure 3.8 Histogram of a Right-Skewed Distribution.

Figure 3.9 presents a <u>U-distribution</u>. This is a symmetrical distribution, with class frequencies declining continuously as we move toward the middle of the distribution from either end. The histogram has the general shape of the letter "U".

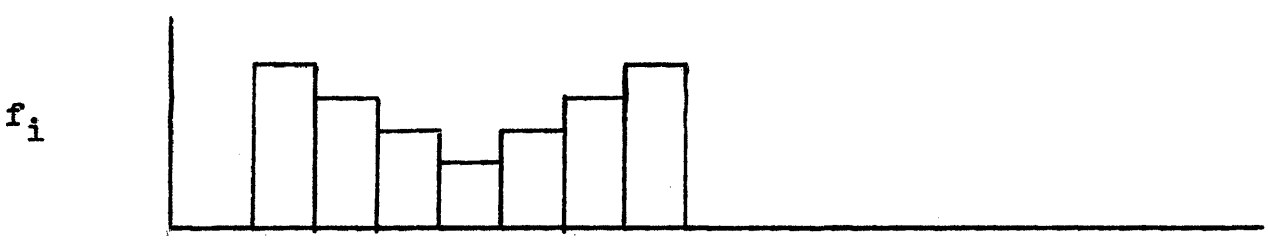


Figure 3.9 Histogram of a U-Distribution.

Figure 3.10 presents a <u>J-distribution</u>, so called because class frequencies rise continuously from the lowest class interval to the highest in such a way that the histogram resembles the letter "J".

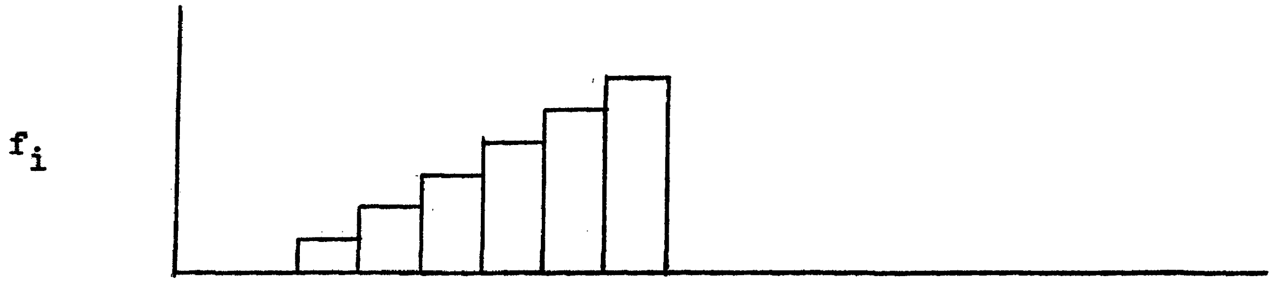


Figure 3.10 Histogram of a J-Distribution.

CHAPTER IV.

CENTRAL TENDENCY AND VARIABILITY

Introduction

We mentioned earlier that the two major types of statistics are descriptive and inferential. Before we can begin the study of inferential statistics, we must learn the techniques used to describe the characteristics of the data from which we wish to draw inferences.

Usually this is done in three ways: the form of the distribution (discussed in Chapter III), the central tendency of the distribution, and the variability of the distribution. The last two are typically reported by the use of numerical indices. In this chapter we will discuss the measures of central tendency and the measures of variability most commonly used in descriptive and inferential statistics.

While each measure of central tendency is defined in its own way, all of them may be considered different methods of finding that which is commonly called an average. Generally, an average is a single value which, in some sense, summarizes a set of data.

Being more specific, we can define measures of central tendency as an attempt to identify the most characteristic score in a group of scores.

Each of the averages we will study provides a measure of the central tendency of a set of numerical data. It should be noted that, although all averages may represent "typical values" of a set of data, this is not always the case. For example, consider the following four scores: 10, 15, 95, 99. It is possible, of course, to find a central value for these four scores somewhere between 15 and 95. However, it is clearly impossible to find any single value which is typical of all four scores.

There are three common measures of central tendency: the <u>mode</u>, the <u>median</u>, and the <u>mean</u>. Although each of these measures of central tendency is an average, each is defined in a different way, and each is used under different circumstances.

The Mode

The mode is the easiest of the three measures to obtain and it is also most subject to fluctuation when the values of a few scores are changed. For this reason, the mode is typically used only for a quick estimate of the central tendency of a distribution. The <u>mode</u> is defined as the most frequent score in the distribution, the score which appears most often. In Table 4.1, each mode is the score which appears more often than any other score in the distribution. Thus, in Sample A, the mode is 8, while the mode in Sample B is 6.

Table 4.1 Modes in Frequency Distributions.

Sample A.

Score (X_i)	Frequency (f_i)
9	6
8 (Mode)	7
7	3
6	6
5	6
4	6
3	5
2	4
1	6
0	1
	N=50

Sample B.

Score (X_i)	Frequency (f_i)
9	6
8	5
7	3
6 (Mode)	11
5	5
4	5
3	4
2	3
1	7
0	1
	N=50

Sometimes a set of data does not have a mode. For example, consider the following set of eight grades on a Social Welfare Policy exam: 73, 86, 92, 84, 86, 92, 73, 84. It will be observed that each grade appears twice. Hence, there is no mode since no grade appears more often than any or all the other grades. In such a case, the mode <u>is not zero</u>; it just does not exist. Also a set of scores may have more than one mode. To be a mode, a score need not have a frequency which is the maximum when all scores along the scale are considered. Actually, a modal score is one which has a maximum frequency along some portion of the score scale. A distribution which contains two or more modes is called a <u>multimodal distribution</u>.

The Median

The <u>median</u> is defined as that point on a scale below which 50 per cent of the cases fall. The median is not a score, but a point on a scale dividing the upper and lower half of the scores. That is, the median is a measure of <u>position</u>.

In a set of scores, the median is the score value centrally located so that the number of scores in the set with a higher value is equal to the number with a lower value. We shall use the symbol M to represent the median, whether it be determined from a set of sample data or from the full population.

There is a very simple procedure to find the median of a sample. When there is an even n, the procedure is as follows:

1. Arrange the scores in the order of their size.
2. Place the median between the $(n/2)^{th}$ score and the $[(n/2) + 1]^{th}$ score.
3. If there is a gap between these two scores, take their average, so that the median will be placed in the middle of the region.

For example, consider the set of data 4, 6, 7, 8, 9, 9. Since there are 6 scores, $n/2=3$ and $(n/2)+1=4$. The third score in order is 7 and the fourth is 8, so the median is 7.5. Here there was no gap between the two numbers, so Step 3 was not needed. But take the set of data 4, 6, 7, 9, 9, 10. Here n is again 6, so the third and fourth scores are again needed. In this case, however, 7 and 9 are not together, there is a gap of 2 between them. Adding the two scores and dividing by two gives a value of 8 for the median.

When a set of data has an odd number (n), the following procedure is followed:

1. Arrange the scores in order of size.

2. Place the median at the score that is the $[(n + 1)/2]^{th}$ score from the bottom of the distribution.

In the case of the following set of data: -8, -6, 0, 1, 5, the median of the five numbers is the $[(n + 1)/2]^{th}$ score, or the third score. Thus, M = 0.

The Arithmetic Mean

The arithmetic mean, usually referred to merely as the mean by statisticians, of a set of values is equal to the sum of the values divided by the number of values included in the summation. We will use the symbol $\overline{X}$ (read "X bar") to represent the mean of a set of sample data, and the symbol μ (the Greek letter Mu) to denote the mean of a set of population data.

For population data, we will use N to denote the number of items in the population. For sample data, we will use n to denote the number of items in the sample. That is, we will speak of a population of size N, and a sample of size n. The computation of the arithmetic mean of a set of scores (X_i), in symbollic notation is:

$$\text{Population: } \mu = \frac{X_1 + X_2 + X_3 + \ldots + X_N}{N} \quad \text{or } \mu = \frac{\sum_{i=1}^{N} X_i}{N}$$

$$\text{Sample: } \overline{X} = \frac{X_1 + X_2 + X_3 + \ldots + X_n}{n} \quad \text{or } \overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}.$$

Table 4.2 represents a sample of scores obtained on a hypothetical aptitude test. Following that is an example of computing the arithmetic mean for this sample.

Table 4.2 Scores on an Aptitude Test.

Scores (X_i)
115
124
130
118
127
106
124
121

$$\sum X_i = 965$$

$$n = 8 \quad \therefore \quad \bar{X} = 120.63$$

In this set of sample data, the sum of the values is equal to 965. Since there are 8 scores in the sample, the mean is calculated by dividing 965 by 8. Thus, $\bar{X} = 120.63$.

Although central tendencies are used in descriptive statistics, an important point to remember is that it has been proven by mathematical statisticians that the mean ($\bar{X}$) computed on the basis of sample data can be interpreted as an estimate of the mean (μ) in the population from which the sample was selected. This is very important when inferential statistics is being undertaken.

The Arithmetic Mean:
Frequency Distributions

Situations often arise when in which it is necessary to compute the arithmetic mean for data organized into frequency distributions. This may occur when the raw data is not available, or when it is inconvenient to work directly with a large volume of unorganized data. Computation of the mean of a frequency distribution can be done by using the following equation:

$$\mu = \frac{\sum_{i=1}^{c} X_i f_i}{\sum^{c} f_i} = \frac{\sum_{i=1}^{c} X_i f_i}{N} .$$

One assumption which must be made is that the scores in a class interval are equal to the class midpoint (X_m). By making this assumption and by using the above equation, where c is representative of the number of classes present, the mean for the data in Table 4.3 is μ=59.17.

Table 4.3 Frequency Table of Scores on a Personality Inventory Administered to 60 Subjects.

Discrete Class Limits	Midpoint (X_m)	f_i	$X_i f_i$
30-39	34.5	3	103.5
40-49	44.5	11	489.5
50-59	54.5	18	981.0
60-69	64.5	15	967.5
70-79	74.5	9	670.5
80-89	84.5	4	338.0

$$\mu = \frac{3550.0}{60}$$

$$\mu = 59.17$$

Variability

Typically, measurements of any kind exhibit variation. Reaction time to a stimulus varies from subject to subject, scores on a psychometric test vary from subject to subject, and the test grades for a class vary from pupil to pupil.

The extent of variation among the scores in a set is of primary importance in any application of statistical methods. For example, consider the two sets of scores: 150, 151, 153, 154 (Set A) and 70, 101, 204, 233 (Set B). In both sets, the mean is 152. However, Set A contains considerably less variation among the scores than Set B. Consequently, knowledge of the mean for Set A conveys more information concerning the individual scores in the set than is the case for Set B.

One of the simplest measures of variation is the range. The range is the distance between the highest and the lowest value in a set. In the above example, the highest score in Set A is 154, and the lowest is 150. Thus, the range is 4, the difference between these two extreme scores.

The score interval, 150-154, as well as the distance, 4, is referred to as the range. In Set B, the range, then, is 70-233 or 163. That is, either the distance or the interval is the range.

Fundamentally, measures of variation are of a different nature than measures of central tendency. That is, an average represents a point on the score scale, whereas a measure of variation represents a distance along the score scale. In essence, the most important measures of variation used in statistics represent, in some sense, the average distance of the scores in a set from a point of reference, usually the point of central tendency.

Standard Deviation

The standard deviation is the most important and most often used measure of variation in statistics. This measure of variation is based on the arithmetic mean as the point of reference. Therefore, the standard deviation is the appropriate measure of variation to use when the mean is used as the average.

Let us consider the set of four scores: 24, 28, 31, 37. Using the mean, 30, as the point of reference, we determine the deviation of each score from the mean as: 24-30=-6, 28-30=-2, 31-30=1, and 37-30=7. We may attempt to construct a measure of variation by computing the mean of these deviations from the mean. That is, add the four deviations and divide by 4. However, we note that the deviations add to zero.

Thus, a slight modification is needed to find the standard deviation of a population or a sample from the deviations of a set of scores from their mean. Firstly, we will use the figure X_d to represent $X_i - \mu$, the deviation of a score (X_i) from the mean (μ) for the set. We will also use X_d to represent the deviation from the mean for sample data as well as population data. In other words, $X_d = X_i - \bar{X}$. Then we will modify this step by squaring the deviations from the mean and then obtaining the mean of the squared deviations. Using this approach, we obtain the following equations for the measures of variation:

$$\text{Population: } \sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N} = \frac{\sum_{i=1}^{N} X_d^2}{N}$$

$$\sigma = \sqrt{\sigma^2} \, .$$

The result σ^2 (σ is the lower case Greek letter Sigma) is the mean of the squared deviation from the mean. It is called the <u>variance</u>. The variance is an important measure of variation in statistics. However, σ^2 measures the variation of a set of terms of the <u>squares</u> of the deviations from the mean. It is general practice to compute, σ, the square root of the variance. This measure of variation, σ, is called the <u>standard deviation</u> of the population.

The <u>standard deviation</u> is a "measure of how, on the average, the scores in a set deviate from the mean." (Armore, 1966, p. 140). In other words, it is a measure of how far, on the average, the scores are dispersed, or scattered, about the mean. A larger value of σ means that the scores in a population are more scattered about the mean, on the average, than would be indicated by a smaller value of σ. In fact, the more closely the scores in a population are concentrated about the mean, the smaller is the value of σ. As an extreme, if $\sigma=0$, it means that all of the scores are at the mean; that is, there is no variation at all among the scores.

It should be noted that it is the <u>distance</u> of each score from the mean, not the <u>direction</u>, which is considered in the computation of the standard deviation, σ.

In our discussion of the arithmetic mean, a distinction was made between the mean computed from the population, and the mean computed from sample data. Along similar lines, a distinction is made between the standard deviation computed from the population and one computed from sample data. We will use the figure, s^2, to denote the variance computed from the sample data, and s to denote the standard deviation computed from sample data. The variance and standard deviation of sample data based on the sample mean, $\overline{X}$, is:

$$\text{Sample: } s^2 = \frac{\sum_{i=1}^{n} (X_i - s)^2}{n} = \frac{\sum_{i=1}^{n} X_d^2}{n}$$

$$s = \sqrt{s^2} \; .$$

This standard deviation, s, can be interpreted as an estimate of the standard deviation, σ, in the population from which the sample was selected.

Standard Deviations: Frequency Distributions

Computation of the standard deviation for data summarized into a frequency table starts off with an assumption. The assumption usually made is that the scores in a class interval are equal to the class midpoint. Table 4.4 presents a frequency table for a population of 50 scores on a matching test. Clearly, if we know all 50 individual scores, we could compute the standard deviation. This is not possible, however, so the class midpoint must be used.

Table 4.4 Frequency Table for a Population of 50 Scores on a Matching Test.

Discrete Class Limits	Midpoint X_i	f_i	$X_i f_i$	$X_i^2 f_i$	
110-119	114.5	10	1145.0	131,102.50	
120-129	124.5	11	1369.5	170,502.75	
130-139	134.5	13	1748.5	235,173.25	
140-149	144.5	9	1300.5	187,922.25	
150-159	154.5	4	618.0	95,481.00	$\sigma = 14.8$
160-169	164.5	2	329.0	54,120.50	
170-179	174.5	1	174.5	30,450.25	
Total:		50	6685.0	904,752.50	

To compute σ and s, with X_i representing the class midpoint, f_i representing the class frequency, and c equal to the number of classes, the following equations are used:

Population:

$$\sigma = \frac{1}{N} \sqrt{N \sum_{i=1}^{c} X_i^2 f_i - \left(\sum_{i=1}^{c} X_i f_i\right)^2}$$

Sample:

$$s = \sqrt{\frac{n \sum_{i=1}^{c} X_i^2 f_i - \left(\sum_{i=1}^{c} X_i f_i\right)^2}{n(n-1)}} .$$

Thus, using the above formula, the standard deviation for the frequency table in Table 4.4 is 14.8.

Summary

The foregoing definitions and terms are by no means exhaustive. There are other measures of central tendency (e.g. harmonic mean, geometric mean) and other measures of variation (e.g. quartile and decile variation). Nevertheless, for the purposes of this text, and most introductory statistics textbooks, the foregoing discussion is adequate.

CHAPTER V.

TESTING HYPOTHESES

Standard Scores

A score standing by itself is, usually, not meaningful. It must be evaluated with reference to the average level of scores, as well as to the variation among the scores. Thus, the concepts of central tendency, averages, and variation become important in making a score meaningful.

A widely used method for evaluating a score is to transform it to standard units. The original score is called a raw score, and the transformed score is called a standard score.

To transform a set of scores into standard scores (or standard units), we must first determine values for the mean and standard deviation of the set of data. Generally, the predetermined values are zero for the mean, and one for the standard deviation. This is accomplished by the following equations:

Population: $$z_i = \frac{X_i - \mu}{\sigma}$$

Sample: $$z_i = \frac{X_i - \bar{X}}{s} .$$

Standard scores, z_i, computed by these formulas have been shown (Armore, 1966, pp. 179-181) to have a mean of zero and a standard deviation of one.

For example, in a sample of $\overline{X}$=20 and s=5, then:

$$z_i = \frac{X_i - 20}{5} .$$

The raw scores from the sample can be transformed to standard scores, z_i, by substituting into this equation. For example, suppose 10 and 40 are two scores from this sample. Then:

$$z_{10} = \frac{10 - 20}{5} = -2 \quad \text{and} \quad z_{40} = \frac{40 - 20}{5} = 4.$$

Please note that z_i scores may be negative as well as positive. Raw scores below the mean result in z_i scores which are negative. Raw scores above the mean result in z_i scores which are positive. A raw score equal to the mean results in a z_i score equal to zero.

Standard scores, z_i, are in standard deviation units. For example, the raw score of 10 in the foregoing example is 10 points below the mean of 20. Also, 10 points amounts to two standard deviations, with s=5, away from the mean. Therefore, the corresponding z_i score of -2 indicates that the raw score of 10 is two standard deviations <u>below</u> the mean.

This indicates an important advantage of standard scores over raw scores. For example, a raw score of 50 is not meaningful in and of itself, but a standard score of 3 is meaningful, indicating a score which is three standard deviations <u>above</u> the mean. As mentioned previously, a negative standard score indicates a score <u>below</u> the mean, whereas, a positive standard score indicates a score <u>above</u> the mean, as the above example illustrates.

The Normal Distribution

Remembering our discussion of distributions in Chapter III, one of the most useful symmetrical distributions in all of statistics is a theoretical distribution known as the normal distribution. The normal distribution, or normal curve, represents a continuous distribution of an infinite population. That is, the variables of the distribution represent continuous variables. The variable, X, is a continuous variable. As defined in Chapter II, a continuous variable is one which can take on any value between specified limits. In the case of the normal distribution, the limits of the variable, X, are negative infinity and positive infinity. In other words, X may take on any value. The equation of this distribution is:

$$Y = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2} \quad \text{where}$$

Y = height of the curve at a specified value of X
σ = standard deviation of the normal population
μ = mean of the normal population
π = 3.1416 (pi)
e = 2.7183, a mathematically determined constant which is the base of the natural (Naperian) logarithms
X = a continuous variable.

Specific values must be assigned to μ and σ before a specific normal curve can be identified. In other words, the above equation defines a "family" of normal curves, and each time values are specified for μ and σ, a particular normal distribution is identified.

For example, let us specify μ=100 and σ=10, the particular normal curve identified is presented in Figure 5.1 on the left. Note that the normal curve is perfectly symmetrical, with a centrally located peak. Therefore, the general appearance is that of a hill or bell. Owing to the perfect symmetry of the distribution, and central peak, the <u>mean</u>, <u>median</u>, and <u>mode</u> are all equal for a normal distribution.

How is the normal curve changed if μ and σ are changed? Suppose, we specify μ=150 instead of 100, while σ remains the same. The normal curve identified by the values of μ=150 and σ=10 is presented in Figure 5.1 on the right. Observe that the only effect on the distribution is to shift it to a higher position on the X scale. In all other respects, the two normal distributions are identical.

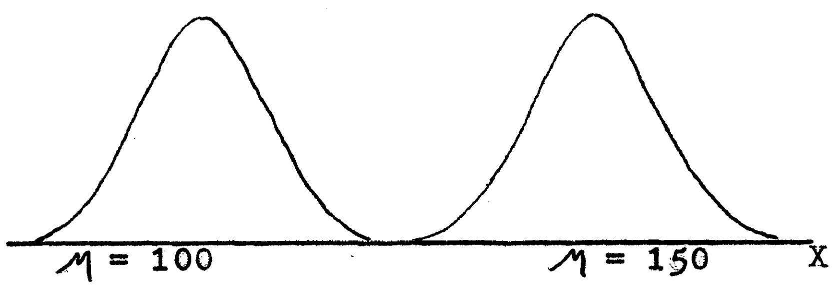


Figure 5.1 Normal Distributions with the Same Standard Deviation and Different Means.

Now, suppose the standard deviation changes. Figure 5.2 presents two normal distributions with the same mean, but with different standard deviations. Note that the distribution with the larger σ has a greater spread or dispersion about the mean than the one with a smaller σ.

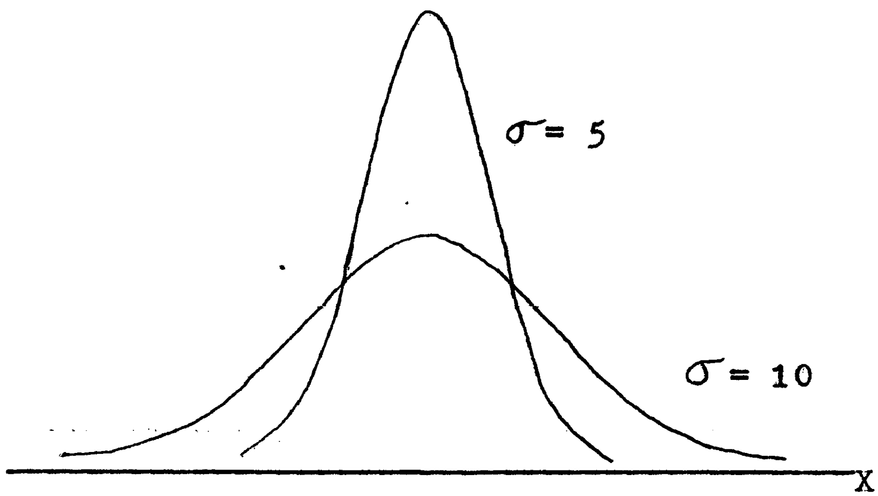


Figure 5.2 Normal Distributions with the Same Mean and Different Standard Deviations.

The most important use of the normal distribution in statistical use is related to the area under the curve. Note in Figure 5.3 that the distribution of the areas under the curve, when viewed on a comparable basis, are identical.

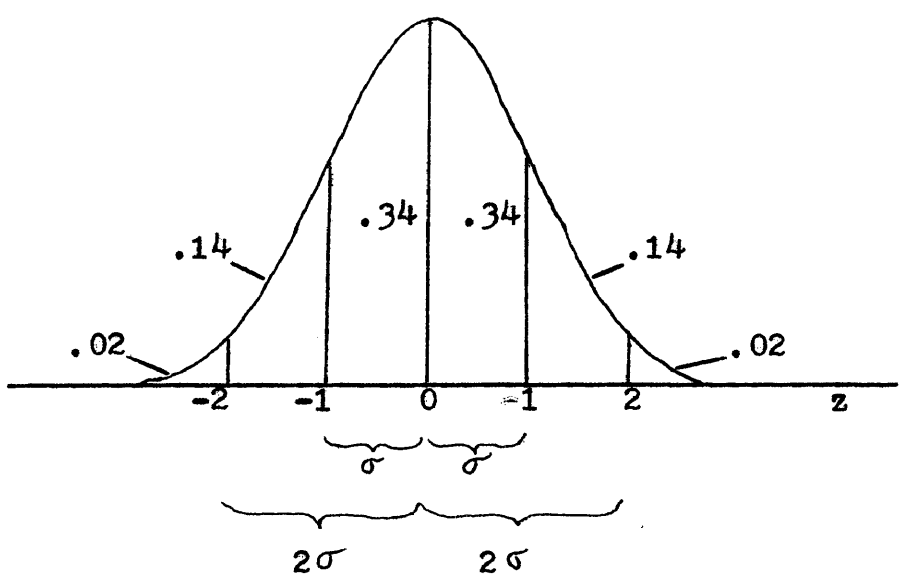


Figure 5.3 The Standard Normal Distribution ($\mu = 0$, $\sigma = 1$).

Generally, our interest in the normal distribution will be in the standard normal distribution expressed in z units. As can be observed in Figure 5.3, only 2% of the population is farther than 2z or two standard deviations from the mean.

Since the area corresponding to a specific interval along the horizontal scale indicates a percentage of the population, it may also be interpreted as a probability. This is the main use of the normal distribution in statistics. This is accomplished by using the Table of Areas Under the Normal Curve: Areas from the Mean to Specified Values of z. Please note Appendix B, Table I for this information.

An example will help illustrate. If the mean score on a vocational assessment test was 22 points with a standard deviation of 3 points, and these scores are distributed approximately like the normal curve, several facts can be determined.

What is the percentage of the sample of scores between 18 points and 23.2 points? We must first transform the scores 18 and 23.2 to standard units. With $\overline{X}=22$ and $s=3$, then: $z_{18}=(18-22)/3=-4/3=-1.33$ and $z_{23.2}=(23.2-22)/3=1.2/3=0.4$. From Table I in Appendix B, a z value of -1.33 is .4082 of the area, and a z value of 0.4 is .1554 of the area. In other words, .5636 or 56.36% of the sample falls between the scores of 18 and 23.2. It would then be stated that the probability is 56.36% that if a score is selected at random from the sample, it will have a value between 18 and 23.2.

Another, but very similar use of the normal distribution is to find the probability that, if a score is selected at random from the sample, it will be higher than a certain value. Using the example from the vocational assessment test above, what is the probability that a score will be above 28? Firstly, transforming 28 to standard units, we obtain $z_{28}=(28-22)/3=2$. By looking at Table I in Appendix B, we find that .4772 of the scores fall between the mean ($z=0$) and 28 ($z=2$). Since the curve of the normal distribution is symmetrical, half (.5000) of the total area under the curve lies on each side of the mean. Therefore, .5000-.4772=.0228 or about 2% of the scores, if selected at random, will be higher than 28 points.

Finally, it must be noted that we are making the assumption that the scores are measured on a continuous scale. Such an assumption is necessary if we are to represent a real sample or real population by the normal curve.

T-tests

Simply, t-tests are used for equality/inequality measures of a sample as compared to a population, when the sample size is less than 30. The t-test is also used in decision making questions, when the sample size is less than 30. If the sample size is more than 30, the procedures for z-tests are used, as described in the previous section.

In order to compare equality/inequality questions and decision making questions, a three-phase procedure must be followed.

This three-phase procedure (Armore, 1966, pp. 358-359) is as follows:

A. Formulate the Problem.

1. Specify the test parameter.
2. State the hypothesis to be tested (H_0).
3. Set up an alternative hypothesis (H_1), to be accepted in the event H_0 is rejected.
4. Specify α, the level of significance of the test.

B. Construct the Decision Model.

1. Specify the test statistic.
2. Determine the sampling distribution of the test statistic by specifying the theoretical sampling model which is appropriate.
3. Specify the test variable and the computational equation.
4. State the decision rule in terms of the test variable, defining the region of rejection on the one hand, and the region of acceptance or of no decision on the other.

C. Evaluate the Sample Data.

1. Compute the test statistic based upon the the sample data.
2. Compute the test variable.
3. Evaluate the test variable according to the decision rule.
4. Make the decision.

The equation for the t-test is:

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} \quad \text{with df} = n-1.$$

In this equation, $\bar{X}$ is the sample mean, μ is the population mean, s is the standard deviation of the sample, n is the number in the sample, and df is the degrees of freedom in the sample. The degrees of freedom of a sample is the number of decisions which can be made before the remaining decisions are predetermined. For example, if a sample has an n=7, then df=6. If we know that the total of the 7 scores is 70, we can decide six different scores, add these together, and subtract this sum from 70. This remainder is the predetermined score which is not part of the degrees of freedom.

An important point to keep in mind when answering equality/inequality questions and decision making questions is whether to use a one-tail or two-tail decision. That is, the types of questions which we will be answering will dictate the need for either a one-tail or two-tail decision. For example, if we are interested in making a decision concerning which students to send to represent the School of Social Work at a national convention, we might decide to send those students whose grade point averages are in the top 2.5% of the class. Then we would use a one-tail decision.

We know that when making z-tests, the top 2.5% is found to be in scores that are 1.96 standard scores above the mean. However, if the size of the sample is less than 30, we must use a t-test to decide how many student's grade point averages are in the top 2.5% of the class. If there are

25 students in the class, then the degrees of freedom would be 24. By looking at Table II in Appendix B, we see that a t value of 2.064 is the value above which 2.5% of the grade point averages would lie. Therefore, by using a one-tail decision, we can find out how many of the students are in the top 2.5% of their class.

Several examples may help illustrate this point of one-tail or two-tail decisions, as well as the three-phase procedure, and the major uses of the t-test.

In a study of Veteran's Administration Hospital residents, their mean I.Q. was 100.7 with a standard deviation of 3.6. Therefore, $\mu = 100.7$ and $\sigma = 3.6$. A sample from another Veteran's Administration Hospital had a sample mean, $\bar{X}$, equal to 103.4 and a standard deviation, s, equal to 2.7, with a sample size, n, equal to 25. Thus, in this sample, the degrees of freedom, df, is equal to 24.

The question is: can we infer from this sample the population parameters? Or, in other words, is this a sample from our population?

A. <u>Formulate the Problem</u>.

1. The test parameters are $\mu=100.7$ and $\sigma=3.6$.
2. H_0: $\mu=100.7$.
3. H_1: $\mu \neq 100.7$.
4. $\alpha = .05$, which means that 95 out of 100 times, the findings that $\bar{X}=\mu$ will <u>not</u> occur by chance.

B. Construct the Decision Model.

1. $\overline{X}$ is the sample mean.

2. The sampling distribution is the t-test, with n-1 degrees of freedom.

3. The test variable is $t = \frac{\overline{X} - \mu}{s/\sqrt{n}}$.

4. If $-2.064 < t < 2.064$ (from Table III in Appendix B), then the sample comes from the population.

C. Evaluate the Sample Data.

1. $\overline{X} = 103.4$.

2. Test variable, t, is equal to $\frac{103.4-100.7}{2.7/\sqrt{25}} = 5$.

3. t = 5 falls in the area of rejection.

4. So, the two measures are not equal. Therefore, this sample is not representative of the population.

Another example is one in which a principal in a junior high knows that the mean score on an aptitude test has been about 125 for students entering high school. He decides to look at the problem of whether or not the educational programs at his school prepare the students for high school. He selects a sample of 22 students, and finds a mean of 121.3 and a standard deviation of 12.7. What should the principal do?

A. Formulate the Problem.

1. The test parameter is μ, the population mean score on the aptitude test.

2. $H_0: \mu = 125$.

3. $H_1: \mu < 125$.

4. $\alpha = .05$.

B. <u>Construct the Decision Model.</u>

1. $\bar{X}$ is the sample mean.

2. The sampling distribution is the t-test, with n-1 degrees of freedom.

3. The test variable is $t = \dfrac{\bar{X} - \mu}{s/\sqrt{n}}$.

4. If $t < -1.721$ (from Table II in Appendix B), then reject H_0 and accept H_1.

C. <u>Evaluate the Sample Data.</u>

1. $\bar{X} = 121.3$.

2. Test variable, t, is equal to $\dfrac{121.3-125}{12.7/\sqrt{22}} = -1.37$.

3. $t = -1.37$ falls in the area of acceptance.

4. So, the principal will be satisfied with the quality of the educational programs in the junior high. Therefore, he will make no changes in the educational programs of his school, at this time.

In conclusion, the t-test is used when testing to see if a sample ($n < 30$) is representative of the population, or to make decisions. Questions such as at what level of significance and whether to make a one-tail or a two-tail decision must be answered. Then the equation:

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} \text{ with a df} = n-1,$$

may be used.

CHAPTER VI.

CORRELATION AND REGRESSION

Introduction

Many of the problems in statistical work involve two or more variables. Two of the techniques for dealing with data associated with two or more variables are correlation and regression. Although the emphasis is on two variables, the methods can be extended to deal with more than two.

In correlation studies, the several variables are studied simultaneously to see how they are interrelated. In regression studies, there is usually one particular variable of interest, and the remaining variables are studied for their relationship with the other one.

Correlation

Correlations are most often used to show relationships between variables or concepts. A measure of association or correlation is usually shown by a two digit figure between .00 and .99. This digit is called the "correlation coefficient."

Correlation coefficients also have directions. For example, two correlation coefficients of -.70 and .70 show an equal relationship between the variables. A minus sign is used to illustrate an inverse relationship, and no sign is used to illustrate a direct relationship.

For example, a correlation coefficient of .63 might be found between age and educational level. What this signifies is that there is a relationship of strength .63 that when a person's age increases, so does their educational level. On the other hand, a correlation of -.27 might be found between a male's age and the quantity of hair on his head. This illustrates that as a male person gets older, the amount of hair on his head becomes less.

There is no general agreement as to what is a strong correlation or a weak correlation. However, the following are guidelines as suggested in Golstein (1969, p. 268):

1. Strong or high correlations are in the range of .70 and above (or -.70 and below).
2. Moderately high or moderately strong correlations are in the range of .34 to .69 (or -.69 to -.34).
3. Weak or low correlations range between -.33 and .33.

Remember that the negative figure only represents that there is an inverse relationship between the variables, not that, for example, -.34 is less than .26. As a matter of fact, a correlation coefficient of -.34 shows a stronger relationship than one of .26.

There are six different types of correlations. Each is used with the different kind of scales: nominal, ordinal, interval, or ratio, as well as when the variables are discrete or continuous. These six are (Goldstein, 1969, p. pp. 269-270):

1. The contingency coefficient ("C") is used

when both variables are nominally scaled and are in more than two categories.

2. Rho, or the rank order coefficient, is used for continuous variables that are scaled on an ordinal scale, so that only the rank order of observations is available.
3. The bi-serial r ("r_{bis}") is used when one variable is continuous and the other is dichotomous, or divided into only two classes.
4. The tetrachoric r ("r_{tet}") is used when both variables are dichotomous, such as in a 2 by 2 table.
5. Phi coefficient ("Ø") is also used when both variables are dichotomous, such as in a 2 by 2 table.
6. The Pearsonian correlation ("r") is used to compute relationships when both variables are continuous and intervally, or ratio, scaled. It is the most accurate when there is a larger number of categories (10 or more) in each variable.

The formula of the correlation coefficient is:

$$\text{Correlation Coefficient: } r = \frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{(n-1)\, s_X s_Y} .$$

The interpretation of a correlation coefficient as a measure of the strength of the linear relationship between two variables is a purely mathematical interpretation. As such, it is completely lacking of any cause or effect implications. The fact that two variables tend to increase or decrease together does not imply that one has any direct or indirect effect on the other. Both may be influenced by other variables in such a manner as to give rise to a strong mathematical relationship. Which means, correlation does not show a cause-and-effect relationship between variables.

Regression

Usually we study the relationship between two or more variables in the hopes that any relationship that is found can be used to assist in making estimates or predictions of a particular one of the variables. The correlation coefficient is merely concerned with determining how strongly two such variables are linearly related, and it is not capable of solving prediction problems. Methods that have been designed to handle prediction problems are known as regression methods. Although there is linear regression, multiple linear regression, and nonlinear regression, the most common type of regression methods, and the one we will study in this text, is one known as the linear regression method.

One assumption which must be made is that there is a theoretical straight line that expresses the linear relationship between the theoretical mean value of Y ($\overline{Y}$), and the corresponding value of X ($\overline{X}$).

Thus, the problem of linear predictions reduces to the problem of fitting a straight line to a set of points. Now, the equation of a straight line is $Y = a + bX$ in which a and b are parameters determining the line. The parameter a determines where the line cuts the Y-axis. The parameter b determines the slope of the line.

Since the problem is to determine the values of the parameters a and b so that the line will fit a set of points well, the problem is essentially one of estimating the parameters a and b in some efficient manner.

It has been found that:

$$a = \overline{Y} \qquad \text{and} \qquad b = \frac{\sum_{i=1}^{n}(X_i - \overline{X})\, Y_i}{\sum_{i=1}^{n}(X_i - \overline{X})^2} .$$

Thus, the regression line of Y on X is:

$$\text{Regression Line:} \quad Y' = \overline{Y} + b(X-\overline{X}) \quad \text{where} \quad b = \frac{\sum_{i=1}^{n}(X_i - \overline{X})\, Y_i}{\sum_{i=1}^{n}(X_i - \overline{X})^2} .$$

The prime is placed on Y (Y') only to distinguish the straight line of Y from the observed value of Y, if we are given the value of X.

Summary

The basic difference between correlation and regression methods is that in the correlation problem, both the X and the Y variables are statistical variables whose values are determined only after the sample is obtained. In the regression problem, on the other hand, the X values are chosen in advance, so that the Y values are determined by the sample.

In terms of the purpose, a correlation coefficient is useful for describing how strongly two variables are linearly related. Regression is used to make predictions about a variable, when the other variable is known.

Correlation coefficients do not lend themselves readily to quantitative statements unless they are associated with regression. Thus, correlation is usually the first phase in the study of the relationship of two variables, whereas regression is the basic technique to use in such a study. From this two-step process, the researcher will gain knowledge of the relationship between the two variables, as well as being able to predict one variable with his knowledge of the other variable.

APPENDIX A

TABLE I.

PROPERTIES OF THE FOUR SCALES OF MEASUREMENT.

(Armore, 1966, p. 25)

Scale Properties	Nominal Scale	Ordinal Scale	Interval Scale	Ratio Scale
1. Equivalence. All items in a given category (or with a given scale value) are equal.	X	X	X	X
2. Relative importance. Indicates greater than or less than.		X	X	X
3. Provides a measure of the interval (distance) between scale values. Ratio of scale intervals are meaningful.			X	X
4. Provides a measure of the interval (distance) of a given scale value from the true zero point. Ratios of scale values are meaningful.				X

TABLE II.

FOUR LEVELS OF MEASUREMENT AND THE STATISTICS APPROPRIATE TO EACH LEVEL.

(Siegel, 1956, p. 30)

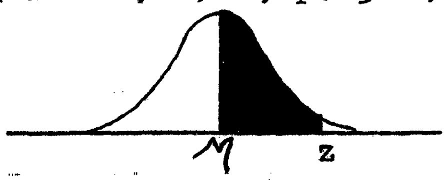
Scale	Defining Relations	Examples of Appropriate Statistics	Appropriate Statistical Tests
Nominal	1. Equivalence	Mode Frequency Contingency coefficient	Nonparametric tests
Ordinal	1. Equivalence 2. Greater than	Median Percentile Spearman r_s Kendall T Kendall W	Nonparametric tests
Interval	1. Equivalence 2. Greater than 3. Known ratio of any two intervals	Mean Standard deviation Pearson product-moment correlation	Nonparametric tests and parametic tests
Ratio	1. Equivalence 2. Greater than 3. Known ratio of any two intervals 4. Known ratio of any two scale values	Geometric mean Coefficient of variation	Nonparametric tests and parametric tests

APPENDIX B

TABLE I.

TABLE OF AREAS UNDER THE NORMAL CURVE.

(Armore, 1966, p. 500)



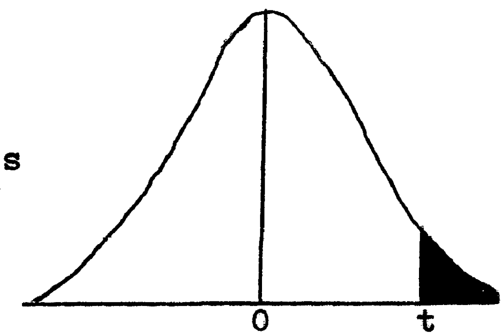
.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
.4774	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
.49865	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
.4999683									

TABLE II.

t DISTRIBUTION: ONE-TAIL.

(Hoel, 1971, p. 288)

The first column lists the number of degrees of freedom (df). The headings of the other columns give probabilities (P) for t to exceed the entry value. Use symmetry for negative t values.

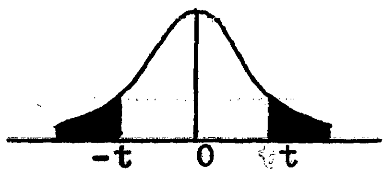


df \ P	.10	.05	.025	.01	.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
∞	1.282	1.645	1.960	2.326	2.576

TABLE III.

t DISTRIBUTION: TWO-TAIL.

(Armore, 1966, p. 501)



df	Probability (Total area in both tails) .5	.4	.3	.2	.1	.05	.02	.01	.001
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	636.619
2	.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.598
3	.765	.978	1.250	1.638	2.353	3.182	4.541	5.841	12.941
4	.741	.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	.727	.920	1.156	1.476	2.015	2.571	3.365	4.032	6.859
6	.718	.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.998	3.499	5.405
8	.706	.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.518	2.831	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.508	2.819	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.500	2.807	3.767
24	.685	.857	1.059	1.318	1.711	2.064	2.492	2.797	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.485	2.787	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.479	2.779	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659
∞	.674	.842	1.036	1.282	1.645	1.960	2.326	2.576	3.291

BIBLIOGRAPHY

Abstracts for Social Workers, 5(1969), 9(1973), and 10(1974). New York: National Association of Social Workers.

Anderson, Richard L. and T. A. Bancroft. Statistical Theory in Research. New York: McGraw-Hill Book Compnay, Inc., 1952.

Armore, Sidney J. Introduction to Statistical Analysis and Inference. New York: John Wiley & Sons, Inc., 1966.

Blank, Stanley S. Descriptive Statistics. New York: Appleton-Century-Crofts, 1968.

Bose, A. B. "Statistical Needs for Social Welfare Planning." Indian Journal of Social Work, 34(1): 65-74, 1973.

Cain, Rolene B. Elementary Statistical Concepts. Philadelphia: W. B. Saunders Company, 1972.

Chaiklin, H. "Research--As Related to Policy-Making." Catholic Charities Review, 56(1): 4-11, 1972.

Due, John F. and Ann F. Friedlaender. Government Finance Economics of the Public Sector. (5th ed.). Homewood, Illinois: Richard D. Irwin, Inc., 1973.

Engstrom, G. A. "Research Utilization: the Challenge of Applying SRS Research." Welfare in Review, 8(5): 1-7, 1970.

Feldman, R. A. "The Evaluation of Social Work Teaching Competence: Issues, Pseudo-Issues, and Future Directions." Iowa Journal of Social Work, 3(3): 63-77, 1970.

Flanagan, J. C. "Evaluation and Validation of Research Data in Primary Prevention." American Journal of Orthopsychiatry, 41(1): 117-123, 1971.

Games, Paul A. and George R. Klare. Elementary Statistics. New York: McGraw-Hill Book Company, 1967.

Gil, D. G. "Research, a Basic Ingrediant in the Study of Social Policy and Social Services." Journal of Education for Social Work, 4(1): 14-20, 1968.

Golden, Deborah, Arnulf M. Pins, and Wyatt Jones. Students in Schools of Social Work: A Study of Characteristics and Factors Affecting Career Choice and Practice Concentration. New York: Council on Social Work Education, 1972.

Goldstein, Harris K. Research Standards and Methods for Social Workers. (Rev. ed.). Northbrook, Illinois: Whitehall Company, 1969.

Hall, Calvin S. and Gardner Lindzey. Theories of Personality. (2nd ed.). New York: John Wiley & Sons, Inc., 1970.

Hammond, Kenneth R. and James E. Householder. Introduction to the Statistical Method. New York: Alfred A. Knopf, Inc., 1962.

Herstein, N. "The Latest Dimension of Social Work Research." Social Casework, 50(5): 269-275, 1969.

Hoel, Paul G. Elementary Statistics. (3rd ed.). New York: John Wiley & Sons, Inc., 1971.

Hope, K. Elementary Statistics: A Workbook. Oxford: Pergamon Press, 1967.

Johnson, Ronald C. and Gene R. Medinnus. Child Psychology. (2nd ed.). New York: John Wiley & Sons, Inc., 1969.

Lewis, H. "Perspectives on Social Work Research in a School's Curriculum." Journal of Social Work Process, 17: 81-96, 1969.

Maddi, Salvatore R. Personality Theories. (Rev. ed.). Illinois: The Dorsey Press, 1972.

Noether, Gottfried E. Introduction to Statistics. Boston: Houghton Mifflin Company, 1971.

Northwood, L. K. and R. L. Reed. "An Agenda for Research About Race and Social Work," in R. R. Miller, (ed.), Race, Research, and Reason: Social Work Perspective. Washington, D. C.: National Association of Social Workers, 1969.

Orlans, H. "The Political Uses of Social Research." Annals, 394(March): 28-35, 1971.

Psychological Abstracts, 39(1965) and 49 & 50(1973). Washington, D. C.: American Psychological Association.

Reid, W. J. "Implications of Research for the Goals of Casework." Smith College Studies in Social Work, 40(2): 140-154, 1970.

Reynolds, R. R., Jr. "Replication and Substantive Import: A Critique on the Use of Statistical Inference in Social Research." Sociology and Social Research, 53(3): 299-310, 1969.

Segal, S. P. "Research on the Outcome of Social Work Therapeutic Intervention: A Review of the Literature." Journal of Health and Social Behavior, 13(1): 3-17, 1972.

Siegel, Sidney. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill Book Company, 1956.

Standards for the Accreditation of Baccalaureate Degree Programs in Social Work. New York: Council on Social Work Education, 1974.

Sturges, H. A. "The Choice of a Class Interval." Journal of the American Statistical Association, 21(1926).

Tripodi, T., I. Epstein, and C. MacMurray. "Dilemmas in Evaluation: Implications for Administrators of Social Action Programs." American Journal of Orthopsychiatry, 40(5): 850-857, 1970.

Weinberger, R. and T. Tripodi. "Trends in Types of Research Reported in Selected Social Work Journals." Social Service Review, 43(4): 439-447, 1969.

Weiss, C. H. "The Politicization of Evaluation Research." Journal of Social Issues, 26(4): 57-68, 1970.

Weiss, R. and M. Rein. "The Evaluation of Broad-Aim Programs: A Cautionary Case and a Moral." Annals, 385(September): 133-142, 1969.

Williams, W. and J. W. Evans. "The Politics of Evaluation: The Case of Head Start." Annals, 385(September): 118-132, 1969.

Young, Robert K. and Donald J. Veldman. Introductory Statistics for the Behavioral Sciences. New York: Holt, Rinehart and Winston, 1965.