1987

Two Educational Comparisons of Linear and Circular Statistics

Robert E. Watson
Loyola University Chicago

TWO EDUCATIONAL COMPARISONS OF
LINEAR AND CIRCULAR STATISTICS

by

Robert E. Watson

A Dissertation Submitted to the Faculty of the Graduate School
of Loyola University of Chicago in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

May
1987

ACKNOWLEDGEMENTS

I would like to express my appreciation to Dr. Jack Kavanagh, the Director of my Dissertation Committee, for his support, encouragement and guidance in this research study. I would also like to thank Drs. Todd Hoover, Steven Miller, and Ronald Morgan for their advice and help in preparing this manuscript.

I would like to thank Drs. John Wozniak and Kay Smith for their support. I also would like to thank the many family and friends who have helped me throughout this course of study. I am grateful to Pat O'Brien for her able assistance in typing and editing the text. Thanks for Annie, a special friend.

Finally, I would like to thank Hazel and Bob Watson, my parents, for their love and guidance.

VITA

The author, Robert Edward Watson, is the son of Robert Watson and Hazel (Nicholson) Watson. He was born June 14, 1937 in Chicago, Illinois.

His elementary education was obtained in the Catholic schools of Chicago, Illinois and Grand Junction, Colorado. His secondary education was obtained at St. Philip High School in Chicago, Illinois.

In September 1956, Mr. Watson entered De Paul University, receiving the degree of bachelor of philosophy in February 1968. He was accepted for graduate study in education at Loyola University in September 1968. He received the degree of master of education in February 1973. In January 1975, he was accepted by the Graduate School of Loyola University for a master's program in the area of curriculum instruction with an emphasis on reading. In May 1979, he was accepted for a doctorate degree in research methods by the Graduate School of Loyola University. In 1987 he was elected a member of Alpha Sigma Mu.

During the past twenty years, Mr. Watson has worked as a teacher and administrator in the Catholic elementary and secondary schools of Lake County, Illinois. He has also taught at Xavier College in Chicago and at Loyola University of Chicago.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CONTENTS OF APPENDICES

Page

CHAPTER I

INTRODUCTION

Statistical analysis is an important consideration in the overall process of problem solving. Statistics can be classified into two general types: descriptive and inferential.

Minium (1978) defines the function of descriptive statistics as providing meaningful and convenient techniques for describing features of data that are of interest. These include the mean, median, standard deviation and the percentile. In each instance, the statistic describes a feature of the data set.

Inferential statistics can be classified into two types: those that utilize hypothesis testing and those that apply estimation techniques. Both are concerned with the population from which the sample is taken. Hypothesis testing asks the question, "Is the value we put on the population reasonable in light of the evidence from the sample?" Estimation techniques ask the question, "What is the population value?"

Inferential tests are of two types: parametric and nonparametric. The distinction between parametric and nonparametric tests lies in the assumptions upon which the tests are constructed. Parametric tests generally have limitations placed upon the distribution of the sample. It is assumed that the data fulfill the conditions of the interval scale. Nonparametric tests require ordinal data (Hays, 1981). Examples of parametric tests are the two sample z-test, analysis of variance, two sample t-test, and significance test and confidence intervals for regression and correlation. Examples of nonparametric tests are the chi-squared test, Wilcoxon, Kruskal-Wallis, Tukey's quick test and Kendall's coefficient of concordance, W.

The present study was designed to examine an alternate analytical method (circular statistics) which is used to analyze data in the form of angular measurement. The data base used in the study was acquired in the field of education. Circular descriptive procedures including measures of central tendency, measures of dispersion and a description of the distribution of the data set were compared to the corresponding linear procedures. Inferential circular procedures including one-sample tests, two-sample tests and circular correlation were compared with the corresponding linear procedures.

The educational data base used in the present study exhibited the mathematical properties associated with a mathematical function. Hays (1973) defines the mathematical relationship as a function if each member of the domain is paired with one and only one member of the range. An example of a mathematical function is $f(x) = y = x + 1$. The conditions stated in the definition of a function are met. For each real number, x, there corresponds one and only one f(x). The domain is represented by x; the range by f(x). When 2 is substituted for x, $f(2) = 3$. The domain is 2 and the range is 3. Now consider the relation, $f(x) = y^2 = x$. If $x = 4$, then y can equal either 2 or -2. In this example the domain, x, is paired with two values in the range, 2 and -2. Therefore, $f(x) = y^2 = x$ does not fulfill the conditions of a function.

Circular statistical procedures are designed to be applied to data which are represented by a periodic function. The functional relationship established between the domain and the range may be described in a number of different ways. The example, $f(x) = y = x + 1$ will graph a line and is called a linear function. A quadratic equation such as $f(x) = y = x^2$ will graph a parabola; and $f(x) = \sin(x)$ establishes a periodic function. Another periodic function is the wrapping function described as a line wrapped about a circle by Crosswhite, Hawkinson and Sachs (1983, p. 127). This is accomplished by mapping the points of the line onto the points of the circle. Each point on the circle may then be described by a central angle and the unit radius. The circular function established by this mapping fulfills the definition of a periodic function.

If statistical data have a functional relationship which is periodic, they reportedly can be analyzed best with circular methods. Batschelet (1965) defines a period function as $f(x) = f(x + a) = f(x + 2a)$... where x represents an element of the domain, f(x) is an element of the range and a is greater than zero. Examples of periodic functions are the trigonometric functions, sine and cosine. The length of the period is 2 pi.

The time study is another form of the period function. Periods of time such as the hour, day, month, or year may be mapped onto a circle and analyzed with circular methods. Hill (1955) describes an example of this type of study in the field of medicine. His study of the hourly distribution of births in a hospital exemplified the use of periodic data based upon a time interval.

Directional studies such as vanishing points in ornithology and wind direction in meteorology demonstrate other applications of the periodic function. The circular shape of the earth also lends itself to this form of description. For example, Schmidt-Koenig (1963) used circular methods to analyze an ornithology study involving the homing direction of pigeons.

The final example of a periodic function considered here is Guttman's quasicircumplex. Guttman's method for scaling a psychological test using a quasicircumplex provides another example where circular methods are employed in the analysis of data. Guttman (1954, p. 260) defined a circumplex as "A system of variables which has a circular law of order ..." and a quasicircumplex as a perfect circumplex including deviations. A quasicircumplex is a method by which data can be circularly scaled. McCormick and Kavanagh (1981) demonstrated that the Interpersonal Checklist fulfills the requirements of a quasicircumplex which thus allows the checklist to be circularly scaled.

Studies using trigonometric functions, time periods, direction and those fulfilling the criteria of Guttman's quasicircumplex demonstrate the properties associated with a periodic function and as such may employ circular analysis. Mardia (1972) lists examples of disciplines in which circular statistics are used: geology, meteorology, biology, geography, economics, psychology, medicine, and astronomy.

Gumble (1954) in his work on the circular normal distribution, von Mises distribution, used geophysical data which included amount of rainfall per year and temperature changes. He also included egg production figures and automobile production and sales figures. Guttman's (1954) development of the circumplex and quasicircumplex is based on Thurstone's mental abilities tests. Guttman's landmark paper provided social science with another way to scale data. If Guttman's conditions are met, circular scaled data must be evaluated using circular methods.

Data from a number of disciplines have been evaluated using circular statistical methods. Biology and related fields have used circular statistics rather extensively.

The use of circular methods began to appear in the social science literature after 1950. Guttman's (1954) work provided psychologists with a method to circularly scale tests. The use of the quasicircumplex in scaling psychological tests preceded the development of many significance tests needed to evaluate the circular data. Watson and Williams (1956) unified the inference problems and renewed interest in the theoretical mathematics associated with circular statistics. Since 1950 social scientists, especially in psychology, have used circular statistical methods to study interest inventories, personality inventories and ability tests. Circular correlation was first developed in the 1970s.

It is postulated here that educational research can benefit from the application of circular statistical analysis. The school year, the class day and the class period are natural time periods. Budget considerations are based upon fiscal periods; attendance is based also upon a time period. Statistical analysis of educational research problems relating to these areas may benefit from use of circular statistical methodologies. Questions concerning enrollment patterns, busing and student distribution may be expressed as directional data. Many tests used in education, such as intelligence tests and interest inventories, meet the criteria of Guttman's quasicircumplex. These examples are viewed as establising the need for circular analysis in the field of education.

The present investigation was designed to demonstrate that educational data may be represented, described and analyzed as circular data. In the following study, the comparisons of linear and circular statistical methods are presented. Computer programs were developed for the circular tests used in these comparisons. A comparative presentation of the linear and circular methods used in the present investigation is presented in Table 1.1.

Table 1.1

A comparative survey of the linear and circular methods used in the investigation

Methods	Linear Methods	Circular Methods
Descriptive Methods		
Central Tendency	Mean Median Mode	Mean Angle Length of Mean Vector Angular Range
Dispersion	Variance Standard Deviation Range Skewness Kurtosis	Angular Variance Angular Deviation Angular Range Angular Skewness Angular Kurtosis
Inferential Methods	Analysis of Variance Chi-squared Pearson Correlation	Rayleigh Test V-Test Watson U^2 Test Watson-Williams Test Circular Correlation Chi-squared Test

One set of data was obtained from the Unit District R. In 1972 Unit District R went to a "45-15" schedule. This type of time schedule provides real data fitting a quadrimodal model. The district had three levels of instruction (grade school, junior high school, and high school) offering excellent opportunities for a correlational study.

The second set of data was provided by Carmel High School for Boys, Mundelein, IL. Carmel's student population consists of male high school students from Lake County, IL. The data afford the opportunity to evaluate enrollment patterns using several methods. Bivariate studies are presented using 1980 census data to evaluate attendance and the socioeconomic status of the students' communities. Correlation studies regarding student attendance is also presented.

Comparisons of linear and circular statistical analytic methods are possible in a number of instances. There are, however, traditional linear methods of analysis which do not have corresponding methods in circular statistics, such as multiple regression and multivariate analysis. The study of a multidimensional model in circular analysis continues to focus on the three-dimensional sphere. Higher dimensions have yet to be considered.

Statistical packages were made available to the researcher with the advent of the computer. Much of the growth in statistical analysis parallels the development of computer technology. Neither of the two major main frame statistical packages, SPSS nor SAS, support circular methods. Computer programs for circular statistics must be developed for the main frame and microcomputers to facilitate the use of circular analysis.

The microcomputer programs developed and used for the circular analysis are presented in Appendix A. These programs were written in standard Pascal. They are menu-formatted and include those circular tests that were employed in this comparative study. Table 1.2 lists the circular tests that have been developed and programmed for the microcomputer.

Table 1.2

Programmed Circular Tests

1. Measures of Central Tendency a. mean vector, r b. mean angle c. median angle	2. Measures of Dispersion a. angular variance b. mean angular deviation c. range d. skewness and kurtosis
3. Tests for Goodness of Fit a. chi-squared test b. Watson U^2 test	4. Tests for Randomness a. Rayleigh test b. The V test
5. Two Sample Tests a. Watson-Williams test b. chi-squared test c. Watson's U^2 test d. the run test	6. Multisample Tests a. Watson-Williams test b. chi-squared test
7. Bivariate General a. means b. standard deviations c. variance d. covariance e. correlation	8. Bivariate Tests a. Hotelling's one sample b. Hotelling's two sample 9. Circular Correlation

In summary, circular data are found in the field of education. Examples of directional studies (Schmidt-Koenig, 1963; Mardia, 1972), time studies (Hill, 1955; Mardia, 1972) and psychological scaling methods (Guttman, 1954; McCormick and Kavangh, 1981) utilizing circular methods were discussed. This present comparative investigation was designed to evaluate data systematically using both linear statistics and circular statistics. In what follows, a comparison of the results of each method of analysis is made in light of an educational problem solving model. To facilitate this comparison, the circular statistical tests that were used were programmed in standard Pascal for the microcomputer.

CHAPTER II

A HISTORICAL REVIEW OF THE LITERATURE RELATED TO LINEAR AND CIRCULAR STATISTICS

The literature of circular statistics parallels that of traditional statistics. The foundations of circular statistics lie in the 17th and 18th centuries in the work of leading mathematicians such as Gauss, Bessel, and Bernoulli. Theoretical statisticians who developed traditional methods were also instrumental in the development of circular statistics. Pearson, Rayleigh, Hotelling, Hodges, Rao, Watson, and Williams all contributed to the literature of both methodologies.

This section traces circular statistics from its origins to the present. A historical review of the development of circular statistics will be divided into three time periods: early period--prior to 1900, middle period--1900 to 1950, and current period--1950 to present.

The second part of this review traces the use of the computer and microcomputer in statistical research and addresses the availability of statistical programs.

Early Period--Prior to 1900

The mathematical origins of circular statistics are found in the late 17th century. Historically significant mathematicians associated with the foundations of circular statistics include Bernoulli, Bessel, Gauss, and Rayleigh. The earliest applications of circular statistics were in astronomy, specifically, mapping the orbits of the planets. The focus of those studies involved ellipses, circles, and spheres.

Daniel Bernoulli (1734), questioned if the similarity of orbital planes of the planets could have arisen by chance. His study was based on the distribution of the normals of each of the orbital planes. Bernoulli's question was finally answered by Watson (1970) who applied the Rayleigh Test (a circular statistic) to demonstrate that the similarity of planetary orbits could not have happened by chance alone.

Gauss developed his theory of errors in 1809 while working on his second great work "Theoria motus corporum coelestium in sectionibus conicis solem ambientium." The circular data would have supported a directional theory of errors. Because the observation errors were small, Gauss developed a linear theory instead of a circular theory.

Friedrich Bessel, a contemporary and friend of Gauss, developed a family of functions. Von Mises (1918) later used these functions in defining the circular normal distribution used in circular statistics.

Lord Rayleigh (1880), in his work on acoustics was the first to study the significance of the mean vector. The results of this study did not reach the general public until Lord Rayleigh (1905) published a response to Karl Pearson's (1905) letter in Nature which generalized the significance of the mean vector.

Statistical analysis owes much to the development of mathematics. The Renaissance produced the calculus attributed to Newton and

Leibniz. The Age of Enlightenment produced such men of genius as Gauss and Euler. Modern statistical analysis was first proposed and used by Sir Francis Galton (1889) in studies on heredity. Circular statistics was first developed to answer the questions of astronomy, but after 1900, interest in this form of mathematical analysis expanded to include distribution theory and tests of significance.

Middle Period--1900 to 1950

During this fifty year period the constructs upon which modern statistical theory are based were developed. Significant and creative thought characterized the era. As probability theory advanced, statistical breakthroughs occurred in distribution theory. The study of uniform distributions led to the development of the circular normal distribution, the wrapped Cauchy distribution, the wrapped Poisson distribution, and the Cardioid distribution among others. Circular statistics' tests for Goodness of Fit and Randomness were first developed during this period. Two-sample tests and Hotelling's bivariate methods were also introduced.

The study of the mean vector that Rayleigh first began in 1880 was generalized by Karl Pearson in 1905. Rayleigh (1905) responded with his estimated value of the mean vector based upon his previous research. The exact sampling distribution of the mean vector was developed in 1905 by J.C. Kluyver (1906). The Rayleigh Test for uniformity was formalized by Watson (1956) after Greenwood and Durand (1955) developed a table of critical values. Since that time Stephens

(1969), Zar (1974), and Papakonstantinou (1979) have refined the table of critical values for the mean vector.

A wrapped normal distribution is a normal distribution wrapped about a circle. The distribution was first studied by de Hass-Lorentz in 1911. Tables were developed by Schuler and Gebelein (1955). The supporting mathematical works were done by Stephens (1963) and Bingham (1971).

Von Mises (1918) introduced a distribution analogous to the normal distribution of linear statistics to measure the difference in atomic weights from integral values. Grumble, Greenwood, and Durand (1953) produced a table of values including a table of probability distributions. Downs (1966) demonstrated that under certain conditions the bivariate distribution is in fact a von Mises distribution.

Hotelling (1931) developed a bivariate confidence ellipse used in linear statistics. In addition he developed parametric one-sample and two-sample tests. Batschelet (1981) incorporated Hotelling's method in a discussion of circular statistics when considering linear variates.

Levy first introduced the Cauchy distribution in 1939. Levy had worked on the Wrapped Poisson Distribution and The Wrapped Normal Distribution as well. Additional theoretical mathematics were presented by Winter (1947).

The cosine function was developed by Jeffreys (1948). The function is expressed in terms of the polar distance and polar angle. The resulting function is heart-shaped, hence, the name Cardioid Distribution.

The fifty year period saw many changes in circular statistics. Lord Rayleigh's initial interest with a uniform distribution led to the theoretical development of several circular distributions. The corresponding advancements in mathematics provided the theoretical framework. The time from 1900 to 1950 was a time of steady progress in the development of statistical methodology. The groundwork was laid for the future development of circular methods.

Current Period--1950 to Present

This period marks a breakthrough in circular statistics. Mardia (1972) attributes much of this renewed interest to a paper written by Watson and Williams (1956) which unified the inference problems attributed to the von Mises and Fisher distributions. It is also a time which was right for growth. Watson, Stephens, Rao, and Guttman all contributed significantly to the literature during this period.

Computer technology developed rapidly and made statistical procedures such as multiple regression and multivariate analysis practical. Along with the technical advancements came breakthroughs in circular statistics. Advancements in computer science provided the impetus for new use of circular statistics. Biology, Psychology, and Economics are examples of disciplines suggestive of creative analytic approaches. All three may be classified by Torgenson's (1958) criteria of "less well- developed sciences" which required new techniques to establish their constructs. The literature review for the current period will focus upon the historical development of those tests that will be used to analyze the educational data.

Guttman (1954) developed the circumplex and quasicircumplex method of scaling psychological tests. Guttman (1957) applied this method of scaling to mental ability tests. Examples of the use of the circumplex in psychological scaling include the Minnesota Multiphasic Personality Inventory by Slater (1962), Strong, Kuder, Holland and ACT inventories by Cole (1973), and the Interpersonal Checklist by McCormick and Kavanagh (1981).

A two-sample significance test was developed by Watson and Williams (1956). It was designed as a nonparametric test of the von Mises distributions. Stephens (1972) revised the test and also proposed the use of a test statistic Z instead of F. Rao and Sengupta (1972) developed a multisample test for large samples.

Greenwood and Durand (1955) first proposed the V-test as an alternate two-sample test to the Watson-Williams test. A better mathematical approximation was developed in 1957 by the same authors which led to a revision by Durand and Greenwood (1958) of the V-test. The V-test provides a significance test for circular data where a predetermined angle already exists.

During the early 1960s Watson continued to publish extensively in the field of circular statistics. He authored several papers on significance tests for the circle and the sphere. They included the Watson U^2 (1961, 1962) test--in 1961 for Goodness of Fit and in 1962 he demonstrated the Watson U^2 version for use with two samples. The latter is an extension of the Goodness of Fit test where U^2 is asymptotically distributed under the null hypothesis. Stephens (1964, 1965) studied the U^2 distribution throughout this period. Stephens (1965) developed the first four moments for the U^2 distribution.

Following Watson's early work with parametric tests, an interest in nonparametric tests was generated. The run test established by Wald and Wolfowitz for linear statistics was modified for circular statistics by Barton and David (1958). The uniform-score test was proposed by Watson and Wheeler (1964) and Mardia (1967). The uniform-score test as presented by Mardia (1967) originated as a special bivariate problem similar to Hodges-Ajne's test. Mardia, in addition to the uniform-score test, developed a table of critical values for the test statistic, β. This interest in the development of parametric and nonparametric tests continued into the latter part of the 1960's.

The progress in circular statistics during the latter part of the 1960s can be attributed to several theoretical statisticians: Beran, Mardia, Rao, Schach, Stephens, and Watson continue to appear. Beran's (1969) discussion of the uniformity of a circular distribution for two-sample nonparametric tests did for nonparametric procedures what the Watson-Williams paper of 1956 did for circular distribution theory. Schach (1969, 1970) contributed to the uniform-scores test independently of Watson-Wheeler. The development of this test can be followed in four papers appearing from 1967 through 1969, Mardia

(1967, 1968, 1969a, 1969b). Stephens (1969) did a Monte Carlo study comparing the power of the Kuiper's V_n, Watson's U^2, and Ajne's A_n. Rao (1966), working with the von Mises distribution as an alternate hypothesis, compared the Rayleigh, Watson's U^2, Ajne's A_n, and Kuiper's V_n. Using Monte Carlo methods Rao established that the power of the Rayleigh is sufficient for large values of K.

This period provided circular statistics with the parametric and nonparametric tests necessary for inferential statistical analysis. With significance tests established, the interest shifted to different methods of analysis in the 1970s. Batschelet (1981) considered bivariate methods, periodic regression, and circular correlation. Although discussed prior to the 1970s the literature now begins to consider these topics in greater depth.

Bivariate methods applicable to circular statistics were first introduced by Hotelling in 1931. He established a confidence ellipse in addition to his one- and two-sample parametric tests. Several nonparametric tests were developed for use with bivariate data. Mardia (1967) developed a bivariate two-sample test based upon his uniform-scores test. The Wilcoxon's sign-rank test represents another linear statistic which was extended to circular analysis by Moore (1980).

Batschelet (1981) discussed periodic regression as the fitting of the cosine function to a bivariate distribution. The independent variable is circular and the dependent variable is linear. Periodic regression differs significantly from linear regression. Batschelet identifies circular regression as primarily a descriptive tool. If the distributions are independent and normal they may be treated in a parametric manner allowing for confidence intervals and the use of analysis of variance (Bliss, 1970). Batschelet considers the case where independence is not assumed. Ware and Bowden (1977) showed that independence is unnecessary when observations were made upon the same individual.

Circular correlation is the last circular method to be considered. Its literature begins in the 1970s continuing into the 1980s. Circular correlation is used in statistical inference. Mardia (1975) addressed this problem of circular correlation, suggesting that the statistic used be the mean vector. Stephens (1979) developed a table of critical values for the correlation coefficient, r. A parametric test was developed by Jupp and Mardia (1980) for the general case involving two circular variates. A nonparametric rank test based upon the correlation coefficient was developed by Mardia (1975). Significance tests considering data where one variate was linear and the other circular were also developed. A parametric test for circular-linear data is attributed to Mardia (1976) and Johnson and Wehrly (1977). This test was the forerunner of Mardia's nonparametric rank test--the general case. An alternative nonparametric test was proposed by Batschelet et al (1973).

In summary, the 1970s saw the development of circular methods expanded to include linear bivariate methods, periodic regression and circular correlation. The application of circular methods to

Biological Science is well established with the possible exception of circular correlation (Batschelet, 1965; 1971). The use of the Guttman model to scale psychological test data has also been established. The use of circular statistics for analysis of educational data has not been well established. An ERIC Search did not identify any such studies.

Computers as a tool to conduct educational research are widely accepted. The October 1983 issue of Phi Delta Kappan (1983) discussed the use of computers in school settings. Although the school use differs from research applications, interest in computers including microcomputers is well established. Mainframe statistical packages such as SPSS and SAS Freeman (1984) are used extensively. However, neither SPSS nor SAS support circular statistics. The leading statistical packages designed for the microcomputer StatPac, (1984), and SYSTAT (March 1985) also do not support circular methods. StatPac is similar in design to SPSS. The use of the microcomputer in the social sciences at the university level had been restricted to word processing Freeman (1984). This was due primarily to the size of the data bases. SPSS has issued a microcomputer version of the mainframe program. The hardware required to run the SPSS program for the microcomputer, however, will require a hard disk and a math coprocessor 8087 chip. Currently, it is designed for the family of PC microcomputers consisting of the IBM-PCXT and the IBM-PCAT. As quality software develops the use of the microcomputers in social science research will increase. Programs for circular procedures are not available for the microcomputer. The development of such programs is essential to access the benefits of the application of circular statistics to educational data.

CHAPTER III

METHOD

The statistical methods applied to the analysis of selected data sets are described in this chapter. The specific computer programs used with the circular methods are discussed.

Comparisons are made between selected statistics from both linear (measures of central tendency including mean, median and mode; measures of dispersion including variance, standard deviation; and inferential methods including analysis of variance, chi-squared and correlation) and circular applications (measures of central tendency including mean vector, mean angle; measures of dispersion including angular deviation, angular variance and angular range; and inferential methods including one-sample tests, two-sample tests, and circular correlation). The selected descriptive and inferential methods include the bivariate methods of correlation and regression. Each statistical test is discussed in terms of its purpose, assumptions, test statistic, decision rule, and interpretation, where applicable.

The computer programs developed for the circular methods used in these comparisons are included in Appendix A. A discussion of each circular statistical program used accompanies each comparison. The discussion includes program use, restrictions, documentation, and verification. The section presents a description of the programming process. In the verification section, the test data used to confirm program accuracy is described. The programming test data used in the verification process are taken from examples cited by Batschelet (1981) and Mardia (1972).

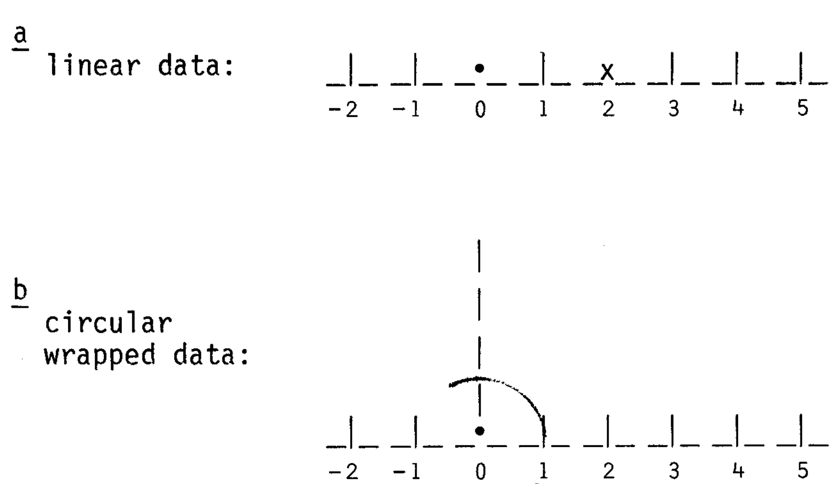
The two data sets used in this analysis are provided by Carmel High School for Boys in Mundelein, Illinois and Unit District R. The Carmel data represents the number of students attending Carmel High School for Boys and the communities in which they live. The Unit District R data provide the daily absentee rates for High School X, Intermediate School Y, and Elementary School Z.

Linear and Circular Statistics

The major distinction between linear and circular statistics is the difference in presentation of data. Figure 3.1 below illustrates this difference.

Figure 3.1

Illustration of Linear and Circular Data



The description of a property as a linear measurement is described by Torgerson (1958) as "...the assignment of numbers to a system to represent that property. In order to represent the property, an isomorphism (a one-to-one relationship), must obtain certain characteristics of the number system involved and the relations between various quantities (instances) of the property to be measured" (p. 15). He identifies these characteristics as order, distance, and origin. Any unidimensional scale must possess order and may exhibit either one or both of the remaining characteristics. In the example, the number two is assigned as a measurement of the characteristic.

The characteristics of measurement for linear data are also assumed by circular data. In addition angular data requires that the function be periodic. A periodic function may be thought of as a function that repeats itself. This periodic repetition is illustrated by the circle in Figure 3.1. The linear measurement is two units. When it is wrapped about the circle the corresponding point on the circle represents the data in circular form. The circular data is expressed by the polar coordinates (r, θ). The radius is assigned the value, one, and the polar angle, θ, is used to describe the characteristic. The circular value corresponding to the linear measurement of two would have an arc length of two and a circular value of 114.65 degrees. Angles that are congruent to each other, such as 630°, -90°, and 270°, are assigned the same position on the unit circle. In the illustration, the circumference of the circle is

6.28. The arc length represented by 8.28 is assigned the same position on the circle as an arc of length of two. The angular measure describing both distances is the same. The calculations are as follows.

The circumference of the circle is:

$$C = 2 \times \pi \times r$$

$$C = 2 \times 3.14 \times 1$$

$$C = 6.28$$

The measure of the angle in degrees is:

$$A = \text{length of arc} \times (180 / \text{pi})$$

$$A = 2 \times (180/3.14)$$

$$A = 114.65^\circ$$

$$A = 2.001 \text{ radians}$$

In summary, linear data requires at least a unidimensional scale exhibiting the property of order. The real number system may be used to describe the data. The geometric representation of linear data requires that the scale fulfills the three properties of order, distance, and origin. Circular data may assign many values to the same point on the circle. In the example, 114.65 degrees and 474.65 degrees are represented by the same point when a circular function is applied. On a linear scale they are represented by two distinct points. Comparisons that follow (arithmetic mean, mean angle; standard deviation, angular deviation; linear correlation, circular correlation) consider the type of data being analyzed and how the analysis and description of that data differs using linear and circular methods.

Measures of Central Tendency

Descriptive methods include measurements of dispersion and measurements of central tendency. Hays (1981) describes measures of central tendency by saying, "Indices of central tendency are ways of describing the "typical" or the 'average' value in the distribution" (p. 142). The measures of central tendency to be compared are: 1) the linear and circular mean; and 2) the linear and circular median. The mode is not considered here. Hays (1981) describes measures of dispersions by saying, "Indices of variability, on the other hand, describe the 'spread' or the extent of difference among the observations making up the distribution" (p. 142). In the present investigation, variance, standard deviation, and range are the measures of dispersion that are compared using linear and circular statistical methods. Each statistic is discussed in terms of its purpose, assumptions, test statistic, decision rule, and interpretation, where applicable.

The Arithmetic Mean and the Mean Angle

The purpose of the arithmetic mean is to describe the "average" value. The mean is used as a descriptive statistic because it is responsive to the exact position of each value. Minium (1978) states, "The mean is amenable to arithmetic and algebraic manipulation in a way that the others are not." This quality allows it to be used in inferential methods. Hays (1981) in discussing the "best guess" measure of central tendency states, "If both the size of the errors and their sign are considered, and we want zero error in the long run, then the mean serves as a best guess."

The statistic is provided by the sum of the scores divided by the total number of scores. The arithmetic mean is defined as:

$$\bar{X} = (X_1 + X_2 + \ldots X_n) / n$$

where $\bar{X}$ represents the arithmetic mean of a sample, X_i represents the ith score, and n represents the total number of scores.

The circular statistic comparable to the mean is the mean angle. The purpose of the mean angle is to find an index that is representative of all the scores in a group. Because circular data is periodic, an angular measure is required to describe it. The angular measurement provides that each point on the circle may be expressed by an angle measuring from 0 to 360 degrees. The process used to describe the arithmetic mean does not adequately describe the mean angle. An example illustrates this difference. Angular measures of 0, 330, and 30 degrees have an arithmetic mean of 120. A plot of these points on the unit circle indicates a value of 0 degrees better describes the "average angle." The mean angle is determined by the use of vector algebra. Each of the data points are defined as unit vectors. Through the application of vector algebra a mean vector is found and consequently a mean angle. As stated in the linear case, the mean angle is the statistic used when computation is required in inferential methods. The deviation from the mean angle is $\theta - \bar{\theta}$. In linear statistics the sum of the deviations are equal to zero. In circular statistics the

$$\sin(\theta - \bar{\theta}) = 0$$

provides a similar relationship.

In summary the linear arithmetic mean and the circular mean angle both describe the "average," however, each must employ a different mathematical method due to the unique way that each method describes its data.

Circular statistics uses the mean angle rather than the arithmetic mean to describe circular data. The mean angle describes the data as a periodic or wrapped function. The circular data is expressed as unit vectors. The coordinates describing each circular measurement may be expressed in trigonometric form:

$$x = \cos\theta$$

$$y = \sin\theta$$

when the radius of the circle is one. The mean vector is then calculated. The arithmetic means for the cosine and sine values are calculated:

$$\bar{x} = (\cos\theta_1 + \cos\theta_2 + \ldots + \cos\theta_n) / n$$

$$\bar{y} = (\sin\theta_1 + \sin\theta_2 + \ldots + \sin\theta_n) / n$$

The mean vector is defined as:

$$r = (\bar{x}^2 + \bar{y}^2)^{\frac{1}{2}}$$

The mean angle is defined as:

$$\text{If } \bar{x} > \theta \quad \text{then } \theta = \arctan\ (\bar{y}\ /\ \bar{x})$$

$$\text{If } \bar{x} < \theta \quad \text{then } \theta = \arctan\ (\bar{y}\ /\ \bar{x}) + 180$$

Special conditions exist where the arctan is undefined. These conditions exist when $x = 0$.

If $x = 0$ and y 0 then = 90°

If $x = 0$ and y 0 then = 270°

In summary the mean angle is the analogous statistic to the arithmetic mean. They exhibit similar statistical properties.

<u>The Median and the Mean Direction</u>

The median used in linear statistics is responsive to the number of scores above and below it, but not to their exact position in the distribution. This quality of the median allows it to be less affected by extreme scores than the arithmetic mean. Because of this characteristic the median is a better choice than the mean when a distribution is extremely skewed.

Many of the qualities exhibited by the linear median are found in the circular median direction. The median direction is described by Batschelet (1981) in this manner, "For this purpose we divide the circular sample by a diameter in such a way that half of the sample points lie on one side and half on the other side of the diameter" (p. 18). The median in both statistical systems divides the points into two equal groups. The position of the diameter in circular statistics must reflect the distribution of the data, otherwise misleading information will be obtained. The linear median is defined as the score below which fifty percent of the scores in the distribution occur.

Mardia (1972) defines the median direction as "a population whose median direction is any solution of

$$\int_{\varepsilon_0}^{\varepsilon_0 + \pi} f(\theta)\, d\theta = \int_{\varepsilon_0 + \pi}^{\varepsilon_0 + 2\pi} f(\theta)\, d\theta = \tfrac{1}{2}$$

with the parent density $f(\theta)$ satisfying

$$f(\varepsilon_0) > f(\varepsilon_0 + \pi)\text{"} \qquad \text{(p. 28)}$$

This may be interpreted as dividing the circle in such a way that half the sample points fall on one side of the diameter and half fall on the other side. This solution is unique when the distribution is unimodal with an odd N. If N is even, then the mean direction passes between the two points.

The median direction is similar in interpretation to the linear median. Mardia (1972) stated, "Hence, the mean deviation is a minimum when measured from the median direction" (p. 31). This effect is the same in linear and circular statistics.

<u>Discussion of the Programming Process</u>

The program called "MEANDIR" provides the user with mean vector, r, the coordinates of the mean vector, the mean angle, and the median angle. The user is required to enter the angular measure in degrees.

The angles are sorted using a bubble sort and provide the user with the sine and cosine values of each measure entered.

The only restriction placed upon the user of this program is with the median. The program requires an input of at least three angular measures. The user must input the data as the program requires.

The program was written in standard Pascal on the IBM Personal Computer. An error situation exists if only two scores are entered and the larger score is entered first. This is due to the fact that a buffer is constructed against which the entry is compared. The only time this error affects the program is in the situation mentioned where two scores are considered, which constitutes a trivial situation.

Verification of the program's accuracy was based upon data taken from Batschelet (1981, p. 14, table 1.3.1). The data represented was taken from a study by K. Schmidt-Koenig (1963) on the homing ability of pigeons. Further verification was provided by Mardia (1972) through the use of Example 2.1 (p. 22). Both tests verified the accuracy of the program. The use of Batschelet's data required the 5 degrees measure be expressed as 365 degrees for verification of the median values. All tabled values are accurate to the thousandth's place. Some rounding did occur in the last decimal place--the ten thousandth's position. The results obtained from using the program "MEANDIR" were verified using the results shown in the Batschelet and Mardia texts. The presentation of data using a table format allows the user to compare the values of the sine and cosine functions for each entry, providing the user with a visual check of the programming procedure.

Measures of Dispersion

In this subsection, three measures of dispersion are discussed for linear and circular statistics. The linear measures of dispersions are variance, standard deviation, and range. Corresponding circular measures of dispersions include angular variance, mean angular deviation, and range. Also included in this section is a comparison of linear and circular skewness and kurtosis. A discussion of the statistical methods includes the purpose, assumptions, test statistic, decision rule, and interpretation, where applicable.

Linear and Angular Variance

The purpose of linear variance is to describe quantitatively the extent scores cluster about a point on a line. The point in linear statistics that variance considers is the arithmetic mean. Hays (1981) says, "The variance summarizes how different the various cases are from each other, just as it reflects how different each case is from the mean" (p. 163). In circular statistics the purpose of angular variance is similar to that of linear statistics. The point about which the scores are clustered is the mean angle. The degree of dispersion is based upon the mean vector, r. For $r = 1$ no dispersion is evident, if $r = 0$, the dispersion is maximized.

Linear variance is the mean of the squared deviations. A squared deviation score expresses the square of the distance a point lies from the mean. The linear variance is defined mathematically by Minium (1978 p. 85) as:

$$s^2 = \frac{\sum (X - \overline{X})^2}{n}$$

The analogous circular statistic is defined by Batschelet (1981) as:

$$\frac{1}{n} \Sigma (\phi - \bar{\phi})^2 = 2 (1 - r)$$

Mathematically there may be several formuli that are asymptotically equivalent to the linear variance. In addition to Batschelet's formula for angular variance is another developed by Mardia (1972, p. 24). Mardia's formula is appropriate when the values cluster about the mean angle. When the divergence is large, Mardia's value tends to infinity whereas Batschelet's remains finite. For this reason Batschelet's procedure is used in this work. Mardia (1972) states his formula in these terms, "We shall show...that an appropriate transformation of s to the range $(0,\infty)$ is given by:

$$s = [-2\log (1 - S_0)]^{\frac{1}{2}} \quad \text{(p.24).}$$

The relationship between the mean vector and sample size is noted by Batschelet (1981), "Consequently, all measures derived from r also depend on the sample size" (p. 35). Using the definition of variance:

$$s^2 = 2 (1 - r)$$

Batschelet (1981) expresses the relationship between linear and angular variance as, "This quantity is asymptotically equivalent to the variance (1.4.6) in linear statistics" (p. 34).

Standard and Mean Angular Deviation

The mathematical definition for the standard deviation is related a sample value and thus uses n - 1 rather than n for the divisor. Glass (1984) defines the standard deviation as "...the positive square root of the variance" (p. 51). Mathematically this is shown as:

$$s = \sqrt{\frac{\Sigma(X - \bar{X})^2}{n - 1}}$$

Batschelet (1981) defines the mean angular deviation as "taking the square root, we obtain a measure of dispersion which is equivalent to the standard deviation in linear statistics. We call:

$$s = [2\,(1 - r)]^{\frac{1}{2}}$$

the mean angular deviation or, ..." (p. 34).

Both angular and standard deviation are expressed in the same units of measurement as the original measurements. Consequently they are used in both descriptive and inferential statistics. Due to their mathematical definitions the standard deviation reflects the properties of the arithmetic mean and angular deviation reflects the properties of the mean vector.

Linear and Circular Range

The purpose of linear and circular range is to provide a description of variability. In linear statistics the range is the difference between the largest and smallest score. It, therefore, is simple to calculate, and its meaning is easily understood.

The circular range is defined by Minium (1981) as "of practical value can also be the range. This is the length of the smallest arc that contains all sample points" (p. 41). Any angular measure greater than two pi may be expressed as an angle between zero and two pi. This feature of circular data makes the circular range an attractive descriptive statistic. The difference between all adjacent arcs must be calculated to determine the largest. This arc length is then subtracted from two pi to determine the range. Although the circular range is primarily a descriptive tool a number of nonparametric tests make use of it.

Linear and Circular Kurtosis

Another dimension of descriptive statistics is the shape of the distribution. The "flatness" and "peakness" that the distribution exhibits about the mode is known as kurtosis. Linear and circular statistics both exhibit the same descriptive qualities. When the statistic or parameter is less than zero, the distribution is platykurtic or flat. A mesokurtic dispersion is reflected when the distribution is normal. If the statistic or parameter is greater than zero a leptkurtic, or slender, dispersion is exhibited.

Mathematically the two statistics used to describe kurtosis differ. The linear case is defined to be the fourth moment of a score. If it is taken about the origin, $E(X^4)$ demonstrates kurtosis. Kurtosis about the mean is defined as $E[X - E(X)]^4$. The linear statistic is defined by Glass (1984) as "..., the mean of the frequency distribution of Z-scores (i.e., Z's raised to the fourth power) minus the constant 3 (which is the mean Z^4-value for the normal curve)" (p. 70). Mathematically, this may be stated as

$$\gamma_2 = \frac{\sum_{i=1}^{n} Z_i^4}{N} - 3$$

Kurtosis for a circular distribution is based upon the second trigonometric moment with rectangular components of:

$$a = \int_0^{2\pi} \cos 2\phi \; f(\phi) \, d\phi$$

$$b = \int_0^{2\pi} \sin 2\phi \; f(\phi) \, d\phi$$

The second moment uses the angular measure 2ϕ. The mean angle, m_2°, and the vector length, $\overline{R}_2$, represent values based upon angular data which has been multiplied by two. The value, $\bar{x}_0$, represents the mean angle and S_0 represents the variance of the first moment. Mardia (1972, section 3.7.3) provides the mathematical development supporting his definition of kurtosis:

$$g^{\circ}_2 = [\ \bar{R}_2 \ \cos(m^{\circ}_2 - 2\bar{x}_0\ ) - (1 - S_0\ )^4\ ] \ /\ S_0{}^2.$$

as stated in section 2.7.2. The denominator serves to offset the effects of dispersion. Batschelet (1965) was the first to consider kurtosis for a circular distribution.

Linear kurtosis is used as a descriptive method when the statistical series is long enough to warrant grouping. For small samples, measures of kurtosis tend to be unreliable. The primary purpose is to quantitatively indicate a sample's departure from normality.

Circular kurtosis provides the same type of interpretive information as does linear kurtosis. Circular grouped data require a statistical correction to compensate for the biasness associated with the mean vector. A table of Sheppard-type corrections may be found in Mardia (1972, p. 38). Kurtosis and skewness are only meaningful if the distribution is unimodal. This results because they are defined in terms of the second trigonometric moment.

Linear and Circular Skewness

The final descriptive method to be considered here is skewness. The concept of skewness is relevant when the scatter is greater on one side of the point of central tendency. If the mean is greater than the median, then the distribution is positively skewed. When the mean is less than the median, the distribution is negatively skewed. Skewness, like kurtosis, can be used to describe a sample's departure from the normal distribution. It is a rough estimate and should not take the place of a goodness-of-fit test.

The linear statistic used to describe skewness was first introduced by Pearson. It is defined to be:

$$ck = \frac{u - \text{Mode}}{\sigma}$$

An alternate form was proposed by Pearson using the third moment:

$$\gamma_1 = \frac{\sum_{i=1}^{N} Z_i^3}{N}$$

Hays (1981) states, "For example, the third moment about the mean will be used in certain measures of skewness: the third moment will be zero for a symmetric distribution, negative for a skewness to the left, positive to the right" (p. 167).

The circular statistic used to describe skewness is attributed to the second trigonometric moment. Its derivation is based upon the rectangular coordinates:

$$a_2 = \int_0^{2\pi} \cos 2\phi \; f(\phi) \; d(\phi)$$

$$b_2 = \int_0^{2\pi} \sin 2\phi \; f(\phi) \; d(\phi)$$

Batschelet (1981) defines the measures of skewness as

$$g_1 = \frac{r_2 \sin(\bar{\phi} - 2\phi)}{s^3}$$

where r_2 is the vector length and ϕ_2 the mean angle. The subscript 2 indicates that each of the angles were doubled in accordance with the second trigonometric moment.

The statements regarding sample size made for kurtosis hold for skewness as well. When using circular grouped data it is necessary to adjust the bias mean vectors r_1 and r_2. The adjustment procedure for the mean vector, when the data is grouped is described in Batschelet (1981, p. 37). Since the variance is defined as a function of r_1 it also must be adjusted. This adjustment is only required for circular grouped data.

In this subsection, we have considered ten measures of dispersions: linear variance, mean angular variance, standard deviation, mean angular deviation, linear and circular range, linear and circular kurtosis, and linear and circular skewness. The corresponding linear and circular statistics generally provide the researcher with similar information concerning the data under question. Linear and circular methods do differ in their mathematical formulation. The differences result from the fundamental difference between linear and circular data. Circular grouped data requires a correction due to the biasness of the mean vector. Since Batschelet (1981) defines variance in terms of the mean vector a correction is also required for angular variance and mean angular deviation.

Discussion of the Programming Process

The circular statistics used to describe the measures of dispersion are presented in a menu format. The program which provides the

Angular Variance, Mean Angular Deviation, Range, Skewness, and Kurtosis is entitled "MenuA." The user is required to enter the measure of each angle in degrees. A maximum of 100 entries is allowed per sample. If a larger sample is used the universal constant, j, which adjusts the size of the arrays, must be adjusted. This adjustment is necessary throughout the other menus for sample sizes larger than 100. The functions r2 and meanang2 multiply each of the data entries by two. These functions are used in determining kurtosis and skewness. The same procedure is used to create a unimodal distribution from bimodal distributions which are 180 degrees apart.

The major difficulty in developing a menu format was the actual construction of the menu itself. Pascal is considered by most to be an excellent programming language. However, its inability to format character data made the layout of the menu a tedious, time-consuming process. A need to re-enter data for each menu is certainly a drawback. These programs are intended for small sample analysis. Large data sets would require a data file. This is beyond the scope of the present study. The range presented a need to identify the largest arc between any two adjacent data points. An array was used to store the arc length between the adjacent points on the circle. A bubble sort was then used to identify the largest arc length. The use of a menu format facilitates the uses of the statistical procedures.

Verification of the angular variance and mean angular deviation were based upon the data taken from K. Schmidt-Koenig (1963) as stated in Batschelet (1981). The variance and mean angular deviation provided by the programs were identical to the values stated by Batschelet (1981) of 0.1998 and 0.447 radians (p. 36).

To verify the programs for kurtosis and skewness the universal constant, j, was set to 1000. The adjustment requires in excess of 256K of memory for the compiler to work. The data used in the verification is presented in Mardia (1972, table 2.6). The measure for skewness was 0.3217 and the measure for kurtosis was 1.4882. The program reflected a skewness measure of 0.3215 and a kurtosis measure of 1.4887.

Circular Tests of Uniformity

Tests of uniformity are generally classified as tests of goodness-of-fit in linear statistics. Circular statistics, however, provides several tests which are specifically designed for use in determining uniformity or randomness. Batschelet (1981) identifies the Rayleigh test and the V-test as the two most important. These tests are being considered in a separate category because the test statistics used are unique to circular statistics. The Rayleigh test is a function of the length of the mean vector, r. The V-test uses the test statistic, u, which is composed of the length of the mean vector, the difference between the mean angle and the projected direction, and the sample size. Neither test has a direct counterpart in linear statistics. Both tests require that a random circular sample be obtained. Batschelet (1981) states that a random sample must meet the following requirements: "..., we assume that each observation is drawn from the same population and that all observations are mutually independent" (p. 52). The Rayleigh and V-tests are the most used goodness-of-fit tests in circular statistics.

The Rayleigh Test

The Rayleigh test was first proposed by Lord Rayleigh and later generalized by Karl Pearson. Its purpose is to distinguish if a given set of circular data reflects a uniform or random distribution. Batschelet (1981) states the null hypothesis as, "H_0: The parent population is uniformly distributed (randomness)" (p. 54). The alternate hypothesis is usually expressed in terms of a directed distribution such as the von Mises distribution. Two assumptions are placed upon the data when using the Rayleigh test. The first assumes that if grouped angular data is being evaluated then the mean vector length, r, has been adjusted for grouping. The second assumes that each angular measure has been adjusted for axial data.

The test statistic is the mean vector length, r. The sum of the unit vectors will ideally be zero if the points are unformly distributed about the unit circle. This relationship was first explored by Lord Rayleigh (1880). Mardia (1972, p. 95) defines the distribution of r as:

$$H_n(R) = R \int_0^{\infty} J_0^n(uR)\, J_1(u)\, du \;.$$

J_n is a family of functions known as the Bessel functions. Tables of critical values of r have been developed by Stephens (1969), Schmidt-Koenig (1963), and Papakonstantinou (1979). For large samples, greater than 30,

$$z = nr^2$$

may be used as a test statistic. Significances occurs when $P < \alpha$ or when $z < z(\alpha)$. In these cases the length of the mean vector differs significantly from zero.

The Rayleigh test is the most frequently used goodness-of-fit test. The length of the mean vector is used as a statistic to determine if a sample exhibits directedness. A unimodal parent population found to be significant using the Rayleigh test provides proof that the directions are concentrated about the mean angle.

The V-test

The V-test is another test of randomness. It requires a known angle, θ_0 , and tests if the angular measures describing the data cluster about the given angle. The angular data must be a random sample of a parent population. The form used here was developed by Greenwood and Durnad (1958). The assumptions concerning grouped and axial data made previously in the Rayleigh test also hold for the V-test. The null hypothesis is stated by Batschelet (1981) as, "H_0: The parent population is uniformly distributed (randomness)" (p. 59). Like the Rayleigh test the alternate hypothesis is usually a directional distribution such as the von Mises distribution.

The test statistic used by the V-test is defined by Batschelet (1981) as, "

$$u = (2n)^{\frac{1}{2}} v$$

where v is given by

$$v = r \cos (\bar{\phi} - \theta_0)\text{" (p. 59).}$$

The parameter v is a distance on the vector having the polar coordinates (r, θ_0). The distance will reach a maximum value when $\bar{\phi}$ and θ_0 are equal. A minimum value will be reached when the difference between $\bar{\phi}$ and θ_0 is pi or 180 degrees. If the circle is a unit circle the values will range from -1 to +1.

If u is greater than or equal to the critical value, then the null hypothesis is rejected. A table of critical values is given by Batschelet (1981). These table values are a modification of the Rayleigh tables.

The V-test should only be used for randomness. It is not applicable where directedness is already indicated and it is to be determined if the mean direction differs significantly from the theoretical direction. If significance is found about a known direction, it is more powerful in rejecting randomness than is the Rayleigh test. The known angle must be selected before the experiment for the test to have meaning. The V-test makes use of the expected direction. A statistical test which utilizes as much information as possible is usually preferred.

Discussion of the Programming Process

The program used for the Rayleigh and V-tests is "MenuB." The input requires the user to enter each angular measure in degrees. For grouped or axial data additional information is required. The main menu consists of the Rayleigh and V-test. The user then is advised of the test's purpose, assumptions, and the null hypothesis. A choice is required based upon the stated assumptions. The test statistic is calculated and the decision rule is shown. A table of critical values is required for the interpretation of the Rayleigh and V-tests.

The Rayleigh test provides four choices based upon the given assumptions:

1. If the data is unimodal and the sample size is 30 or less.
2. If the data is unimodal and the sample size is 30 or more.
3. If the data is multimodal and the sample size is 30 or less.
4. If the data is multimodal and the sample size is 30 or more.

Batschelet (1981) uses a sample size of 30 as a suggested break point. A sample size of 30 may be used with either method. All four selections return the parameters, r, the length of the mean vector, and n, the number of angles. Using these parameters the critical value, P, may be obtained from the previously mentioned table of critical values. Selections one and two require only the angular measures as input. Selections three and four require the number of axis. For axial data to be used the arc between each axis must be of equal length. Selections one and three are based upon the distribution of the length of the mean vector, r. Selections two and four use the z statistic

$$z = nr^2$$

to determine the test statistic.

The V-test provides the user with two choices:

1. If the data is unimodal.
2. If the data is axial.

The test statistic used is u. Axial data will require additional input. The user is provided with the parameters, u and n. The program does not consider grouped data.

The verification of the Rayleigh program was based upon two sets of data. The first set of data was taken from a study by Schmidt-Koenig (1963). This is presented by Mardia (1972) in example 6.3. The second set of data is found in example 1.6.4 of Batschelet (1981). The data comes from a study by Wagner (1952). Using program one from the Rayleigh test the parameters $r = 0.2229$ and $n = 10$ were the same as those stated in the Schmidt-Koenig study. The parameters obtained when program three was applied to the Wagner data were found to be identical to those shown in the study. These parameters for the axial data were $n = 13$ and $r = .7348$. Programs two and four required that the parameter u be calculated and that the sample size be thirty or more. The samples used to verify these programs each had fewer than thirty measurements. Sample size determines which statistic, r or u, is to be used. The algorithm, therefore, may be verified using a sample size smaller than thirty. Program two produced a z-score equal to 0.6045. The calculated value was based upon the parameters r and n from the Wagner study and produced a calculated z-score of 0.6043. A similar comparison was done with the Schmidt-Koenig data, here the program value was 0.4967 and the calculated value was 0.4968. Program four was verified using the Wagner data. The program value was 7.0189 and the calculated value was 7.0191. Based upon these limited verification studies, it appears that the four programs which constitute the Rayleigh provide accurate parameters.

Verification of the V-test was done using data from example 4.3.1 of Batschelet (1981). Both conditions were tested:

1. If the data is unimodal.
2. If the data is axial.

The first program based upon unimodal data from the Keeton (1969) study returned a u value of 2.3937 based on a sample size, n, of 11. These values were identical to those stated in example 4.3.1. The Keeton data was also used as a test of axial data. The data had to be modified by dividing each angular measurement by two. The algorithm could then address the data as if it were bimodal. Theoretically the program should produce the same result as the unimodal case. A u = 2.3937 was obtained verifying the program's computational accuracy.

Goodness-of-fit Tests

The goodness-of-fit tests to be considered here are the linear and circular chi-squared tests and the Watson U^2 test for circular data. Hays (1981) discussing goodness-of-fit tests states, "Tests ...

based on a single sample distribution are called 'goodness-of-fit' tests" (p. 540). The goodness-of-fit test appearing in the literature most often is the chi-squared test. For this reason the linear and circular chi-squared tests are compared in the subsection presented below. The other three tests (the Rayleigh, V, and the Watson U^2 tests) are used to demonstrate the unique statistical characteristics associated with circular analysis.

<u>The Linear and Circular Chi-squared Tests</u>

The purposes of the chi-squared test for linear and circular statistics are similar. This similarity may be seen in comparing Hays statement of purpose for the linear case with that of Batschelet for the circular. Hays (1981) states, "Chi-squared tests of goodness-of-fit may be carried out for any hypothetical population distribution we might specify, provided that the population distribution is discrete, or is thought of as grouped into some relatively small set of class intervals" (p. 540). Batschelet (1981) states the purpose of the circular case as, "To test whether the given distribution fits the sample" (p. 71).

Data for the linear case must be randomly selected from independent observations. If grouped data is used then the categories must be mutually exclusive and exhaustive. The angular data must be randomly selected and a circular distribution must be compared. In addition the expected frequency in each group interval is recommended to be at least four.

The linear and circular null hypotheses are essentially the same. Daniel (1978) states the linear null hypothesis as, "The sample has been drawn from a population that follows a specified distribution" (p. 256). Batschelet (1981) states the null hypothesis for the circular distribution as, "The parent distribution coincides with the given circular distribution" (p. 72).

The mathematical formula developed by Karl Pearson for chi-squared is the same for both statistical methods. The only difference is in the presentation of the data, one is linear the other is circular. Minium (1978) defines chi-squared as,

$$\chi^2 = \sum \left[\frac{(f_o - f_e)^2}{f_e} \right]$$

where: f_e is the expected frequency and f_o is the obtained frequency and summation is over the number of discrepancies..." (p. 428). Batschelet (1981) defines the circular form as "Let k be the number of groups, n_i the observed and e_i the expected frequency in the ith group. Then the test statistic is:

$$X = \sum_{n=1}^{k} (n_i - e_i)^2 / e_i \quad ." \quad (p. 72)$$

The terminology is slightly different between the linear and circular formuli, however, conceptually they are identical.

The null hypothesis is rejected if the critical value, P, is smaller than the preassigned level, α. The linear statistic requires a "large" sample size, Daniel (1978) indicates a sample size of thirty or more is usually adequate. He also cautions that for each population parameter estimated, one degree of freedom is subtracted. Batschelet (1981) notes that the chi-squared statistic is an estimation. The error associated with the estimation becomes large when the sample size is small.

The Watson U-squared Test

The Watson U^2 test uses the mean squared deviation as a test statistic. G. S. Watson (1961) developed this circular test based upon Cramer and von Mises's work. Mardia (1972) demonstrated that the test statistic, U^2, was distribution-free. This characteristic allows the test to be classified as nonparametric.

The purpose of the test is to determine whether a certain circular sample fits a given circular distribution. The test assumes that the data is a random sample composed of angular measures. Group data is not allowed. It is assumed that the data has been rearranged in ascending order. Batschelet (1981) states the null hypothesis as, "The parent distribution coincides with the given theoretical distribution" (p. 80).

The circular test statistic, U^2, is based upon the linear W^2 statistic developed by Cramer-von Mises. Mardia (1972) defines the Cramer-von Mises statistic as: "

$$W^2 = \sum_{i=1}^{n} [U_i - \{ (2i - 1) / (2n) \}]^2 + 1 / (12n) ,$$

where the U_i are the uniform order statistics..." (p. 181). Batschelet (1981) defines the circular test statistic as, "The mean square deviation is essentially an integral" (p. 80). By working out and rearranging the terms, one obtains an expression which is suitable for numerical calculations. Let $F(\phi)$ be the distribution function of the theoretical distribution. Let

$$v_i = F(\phi_i) \qquad (i = 1,2,\ldots,n)$$

$$v = 1 / n \, \Sigma \, v_i$$

$$c = 2i - 1$$

that is, the numbers 1,3,5,...,2n-1. Then the test statistic is

$$U^2{}_n = v^2{}_i - (c_i \, v_i / n) + n[1/3 - (v_i - \tfrac{1}{2})^2]$$

(p. 80). If the calculated value of U^2 exceeds the critical value the null hypothesis is rejected. A table of critical values was developed by Stevens (1964).

The U^2 statistic is a form of variance distributed about the origin. The statistic demonstrates several desirable characteristics. It is invariant under rotation. It may be used for unimodal and multimodal distributions. Small sample sizes do not limit its use.

<u>Discussion of the Programming Process</u>

The program, MenuC, includes two tests for goodness-of-fit. They are the chi-squared test and the Watson U^2 test. MenuC has a menu format. The main menu provides the user with a choice between the chi-squared and the Watson U^2 tests. Each test provides a second menu advising the user of its purpose, assumptions, and the null hypothesis. The user is required to choose one of the following conditions based upon the distribution under consideration:

1. If the given distribution is uniform.
2. For distributions other than choice a.

The data is then entered. A third menu provides the calculated value of the test statistic and the decision rule. To determine significance, a table of critical values must be supplied by the user.

The data entry for the chi-squared test requires different inputs based upon the aforementioned conditions. Condition one for a uniform distribution requires that three separate inputs be made. The first input asks for the number of equal arcs into which the circle will be subdivided. The second entry requires the total number of angles. The third input asks that each angle be entered in degrees. Condition two requires that the user make two entries for each set of grouped data. The input consists of the calculated frequency followed by the expected frequency.

The procedure for the Watson U^2 test requires two separate entries determined by the distribution. The two conditions are the same as those used in the chi-squared test. If the distribution is uniform, condition one is selected. The two required entries consist of the total number of angles and the measure of each angle expressed in degree. Distributions other than the uniform distribution require the use of the second condition. A normal distribution may also be considered under condition two. The inputs consists of: the total number of angles, the measure of each angle, and the probability of each angular measurement. The probability is expressed as a uniform order statistic. The example illustrates the computation. Under condition two, assume that the distribution is normal and the angular measure is twenty degrees. What should be entered as the probability of twenty degrees?

$$F(\phi) = 20/360 = 0.0556$$

The value 0.0556 is entered as the probability.

The program was written in standard Pascal and compiled on the IBM Personal Computer. The Pascal compiler used DOS 2.0. If DOS 2.1 is used to run the MenuC.EXE, an error message will result. To use later editions of DOS 2.x or DOS 3.x, the program must be compiled using that edition of DOS. The program used the SORT function previously mentioned. Individual procedures were developed for the chi-squared and the Watson U^2 tests. The procedure written for the individual tests included the data input. Previously the data input was treated either in the main body of the program or as a separate function.

Verification of the chi-squared program was based on data provided by Goodman (1970). The angular measures do not appear in the publication. Batschelet (1981, p. 74) provides a diagram from which the fifty angles were estimated. The estimated values were:

002	005	011	017	020	026	035
039	039	045	045	048	055	070
075	075	079	086	090		
108	116	122	129	129	133	
146	149	155	167	171	175	
187	215	212	212	208		
232	242					
268	268	286				
300	314	323				
335	339	339	350	365	369	

The chi-squared program for a uniform distribution provided a chi-squared value of 5.6000 with 9 degrees of freedom. These are identical to values stated by Batschelet (1981, p. 73). The second condition considering various distributions was verified using data from Mardia (1972, p. 136, Table 6.2). The distribution considered was a von Mises distribution. The program calculated a chi-squared value of 9.6478 with 11 degrees of freedom. Mardia (1972, p. 135) shows a chi-squared value of 9.6 with 11 degrees of freedom. Mardia shows all chi-squared values to one decimal place. Confirmation was also established using the data from Batschelet (1981, p. 75, Table 4.8.1). This is the same data used in the uniform distribution verification. The uniform probability was determined to be 0.1 for each of the groups. Chi-squared was found to be 5.6000 with 9 degrees of freedom, verified by Batschelet.

The Watson U^2 test for a uniform distribution used two sets of sample data in the verification process. The data from Batschelet (1981, p. 81, Table 4.10.1) was entered in the program. A calculated U^2 value of 0.1361 was obtained. This value is identical to the figure determined by Batschelet. Mardia (1972, p. 182) provided another set of data. The program produced a calculated U^2 value of 0.1158. Mardia indicates a U^2 value of 0.1157. Condition two of the Watson U^2 test is used with distributions other than the uniform distribution. Condition two, however, may be used to test the uniform distribution. A uniform distribution was applied to the data from Batschelet (1981, p.118, Table 4.10.1) using condition two of the

Watson U^2 test. The calculated value of U^2 provided by the program was 0.1361. This figure is identical to the value shown in Batschelet.

Two-sample and Multisample Tests

The following discussion of inferential tests considers two types. First, multisample tests that may be used with two or more samples are discussed. The multisample comparisons will be made between the t-test and the Watson-Williams test, and the circular and linear chi-squared test. Second, tests that are specifically designed for use with only two samples are discussed. They are the run test and the Watson U^2 test. The Wald-Wolfowitz test is compared to the run test in circular statistics. The Watson U^2 test does not have a comparable linear statistic.

The T-test and the Watson-Williams Test

The t-test is used in linear statistics to compare the arithmetic means of two samples. If the sample size is small, a t-test should be used provided Fisher's assumption of homogeneity of variance holds. The mathematical derivation of the t-test makes three assumptions: first, the raw scores for both sampling distributions are normally distributed; second, the individual observations are independent; third, the variances of the samples are equal or homogeneous. The null hypothesis and the alternate hypothesis are stated as:

$$H_0:\ u_x - u_y = 0$$

$$H_A:\ u_x - u_y \neq 0$$

The distribution proposed by Gosset and known as Student's distribution or the t-distribution, pools the variances of the two samples to estimate the population variance. Minium (1978) defines the estimate of the pooled variance as

$$s^2_{\bar{X}-\bar{y}} = \frac{x^2 - y^2}{(n_x - 1)\ (n_y - 1)} \quad (1/n + 1/n)$$

where $x = X - \bar{X}$ and $y = Y - \bar{Y}$.

The calculated t-value uses the estimated standard error of the difference between two independent means. The three assumptions stated previously must apply. The formula for calculating t is:

$$t = \frac{(\bar{X} - \bar{Y}) - (u_{\bar{X}} - u_{\bar{Y}})}{s_{\bar{X}-\bar{Y}}}$$

A t-test is available for dependent means. An adjustment is made to the standard error as follows:

$$s^2_{\bar{X}-\bar{Y}} = s^2_{\bar{X}} + s^2_{\bar{Y}} - 2rs_{\bar{X}}s_{\bar{Y}}$$

A table of critical values of t are found in most statistical books. The degrees of freedom equal

$$df = n - 2$$

where n is the total number of scores. When using dependent scores, n represents the number of paired scores. The decision rule states if

$$t_{cal} > + t_{crit} \quad \text{or} \quad t_{cal} < -t_{crit}$$

then the null hypothesis is rejected.

The Watson-Williams test was first developed in 1956. Stephens (1972) proposed that a g > 1 be used to reduce the error when approximating the F distribution. The data used must be angular and represent independent random samples. The two samples must be taken from a von Mises distribution having equal concentration parameters. It is assumed that $K \geqq 2$. These statements are analogous to the three assumptions stated for the t-test. Batschelet (1981) states the null hypothesis as "Let θ_1 and θ_2 be the (unknown) mean angles of the populations. Then the null hypothesis states:

$$H_o: \theta_1 = \theta_2." \text{ (p. 98)}$$

The multisample form is expressed as:

$$H_o: \theta_1 = \theta_2 = \ldots = \theta_n.$$

The test statistic that Batschelet (1981) provides is an approximation to the F distribution. The error associated with the approximation is small, provided the parameter, K, is greater than two. The test statistic for the two sample form of the test is:

$$F = g(n - 2) \frac{R_1 + R_2 - R}{n - (R_1 + R_2)} .$$

Where R_1 and R_2 are the lengths of the resultant vectors. The two resultant vectors are pooled to form R. This is accomplished by combining the two samples into one sample and computing the resultant vector, R. The variable g is defined as

$$g = 1 + \frac{3}{8K} .$$

Note, as K approaches infinity, g approaches 1. The formula for the multisample group is:

$$F = g \frac{(n - k) \ (R - R)}{(k - 1) \ (n - R)}$$

where k is the number of samples. The F statistic has k-1 and n-k degrees of freedom. Significance is obtained if the calculated value of F exceeds the critical value of F. Tables of F values are available in most statistics books.

The Watson-Williams test is a parametric test. Consequently, it is limited by restrictive assumption. The assumption $K > 2$ is necessary to correct for the error associated with the approximation of the F value. It is a powerful test when these assumptions are met. The multisample form may be used for a test of homogeneity.

Linear and Circular Chi-squared Two-Sample and Multisample Tests

The purpose of the linear and circular chi-squared test is to determine if the samples differ significantly from each other. Multisample tests compare two or more samples. The data are composed of independent random samples. The form of the data was discussed when the goodness-of-fit tests were considered previously. The circular assumption made in the goodness-of-fit test recommends that all expected frequencies be greater than or equal to 4, and is amended for the multisample test to five or more.

The test statistic is the same as that used in the goodness-of-fit discussion. For both the linear and circular two-sample test, a 2 x c contingency table is used. The linear and circular multisample tests employ a r x c contingency table. Batschelet (1981) defines the expected probability as:

$$e_{ij} = M_i N_j / N \quad (i = 1,2,\ldots,r) \; (j = 1,2,\ldots,c)$$

where e is the expected frequency of the ith, jth term. M is the sum of the frequencies in the ith row, c is the sum of the frequencies in the jth column and N is the total.

The degrees of freedom are determined by

$$(r - 1)\ (c - 1).$$

The null hypothesis is rejected if the calculated value of chi-squared is greater than the critical value. Tables of chi-squared values are readily available.

The Linear and Circular Runs Tests

The linear test was proposed by Wald and Wolfowitz (1940). Barton and David (1958) modified the linear test for use in circular analysis. Both the linear and circular tests make the same assumptions. The data are composed of two independent random samples. The populations from which they are drawn are continuous. A comparison of the statement of the null hypothesis of each test indicates that the purposes of the two tests are identical. Daniel (1978) indicates: "The Wald-Wolfowitz runs tests (T-106) uses the number of runs present in the data from two samples to test the null hypothesis that the samples come from identical populations against the alternative that the populations differ in any respect whatsoever-..." (p. 103). Batschelet (1981) states the null hypothesis as: "The two samples are drawn from the same population" (p. 121).

The test statistic used is the total number of runs. Daniel (1978) defines a run as: "...a sequence of events, items, or symbols that is preceded and followed by an event, item, or symbol of a different type, or by none at all" (p. 53). Daniel defines the total number of runs to be r and Batschelet refers to them as h.

Table values were developed for the linear test by Swed-Eisenhart (1943). A table of critical values for the circular test was published by Asano (1965).

Ties between the two samples present a serious problem. This limits the uses of the test for large samples. Daniel (1978) provides an asymptotic normal statistic that can be used for large samples. This, however, does not alleviate the problem of ties. If significance is found, the cause is not identified for either statistic. Further inspection of the data is required.

The Watson U-squared Test

The test was developed by G. S. Watson in 1962 and is based upon the sum of squares methodology proposed by Cramer and von Mises. The use of the deviations between distribution functions as a statistical measure was proposed by Kolomogorov (1933). The purpose of the test is to determine if the two samples differ significantly from each other. The null hypothesis may be stated as: The two samples are taken from the same population. The circular data consists of two independent random samples. The populations from which the samples are taken must be continuous.

The test statistic is defined by Batschelet (1981) as: "Let $n + m = N$ and $d_1, d_2, d_3, \ldots, d_N$ be the difference between the sample distribution functions at the N sample points. If d denotes the arithmetic mean of the d_k's, we consider the quantity

$$U^2 = nm/N^2 [\, d_k^2 - 1/n\, (d_k)^2 \,]." \text{ (p. 115).}$$

The number of elements in sample one is represented by n and m represents the elements in sample two.

If the calculated value of U^2 exceeds the critical value, the null hypothesis is rejected. A table of critical values was published by Zar (1974). If significance is indicated, the test does not identify the statistic with which it is associated.

Because the Watson U^2 test is a nonparametric test, tied scores present a problem. The test is most useful with small samples where ties are less likely to occur. Batschelet (1981) recommends that if grouping is used, it be limited to not more than five degrees.

Discussion of the Programming Process

Two menus were developed for the two-sample and multisample tests. MenuD contains the two-sample Watson-Williams test, the chi-squared test, and the Watson U^2 test. MenuE includes a two-sample test, the runs test, and the multisample forms of the Watson-Williams and the chi-squared test. Both programs are menu driven. The main menu requires a choice of statistical tests. Each test provides a second menu stating the purpose, assumptions, and the null hypothesis upon which the test is based. A third menu provides either the test statistic or the parameters used to determine the test statistic and the decision rule.

Data entry form MenuD is incorporated into the main program. A function entitled, input, is called by the main program. A two-sample test requires that a minimum of five entries be made. From the main menu, the user is required to choose the test statistic. The program asks for the following four inputs:

Enter the data for the first set.

1. Enter the number of angles you have.
2. Enter the measure of each angle.

Enter the data for the second set.

3. Enter the number of angles you have.
4. Enter the measure of each angle.

The second menu of the Watson-Williams test provides the user with the parameters, r and n, to determine the von Mises parameter, K. A table of K values based on the parameter, r, is needed. Entry of the K value provides the user with an F value and the appropriate degrees of freedom. The second menu of the chi-squared test requires the entry of the number of groups or equal arcs into which the circle will be subdivided.

The data input for MenuE takes two forms. The runs test is a two-sample test and requires the five inputs stated previously. The multisample test requires entry of the number of samples to be compared. Each sample has two data inputs: the number of angles for the sample and the measure of each of the angles. The Watson-Williams test requires the K value estimated from the parameters be supplied. Chi-squared requires the number of equal arcs into which the circle will be subdivided.

The runs test was included as part of MenuE because of the Pascal compiler. The IBM PC used to program the circular statistics was equipped with 180K drives. The Pascal compiler exceeded the capacity of the disk drives when compiling the four two-sample tests. The overflow error was obtained when the PASIBF.SYS and PASIBF.BIN files were constructed. Each of the statistical tests used were programmed as individual procedures. A function named, SORT, was programmed to sort the data used in the runs test. In addition, functions rl and input were used. Input has been discussed previously. The length of the mean vector is determined by the function rl.

Verification of the Watson-Williams two-sample and multisample tests was based on data from Duelli and Wehner (1973). The program for the Watson-Williams two-sample test provided an F value of 8.71

with 1 and 18 degrees of freedom. The parameter, r, equal to 0.9905 is used to estimate the parameter, K. Batschelet (1981, p. 102) indicates an F value equivalent to 8.6 with 1 and 18 degrees of freedom. The difference between the F value provided by the program and the F value stated by Batschelet is attributed to Batschelet's assumption that

$$g = 1 + 3/8K = 1$$

The program calculates the actual value of g when the estimated parameter, K, is entered instead of assuming that the value approximates one. Verification of the multisample form of the Watson-Williams test employed the data from Batschelet (1981, p. 100, example 6.2.1). The program for the multisample form of the Watson-Williams test provided an F value of 8.707 with 1 and 18 degrees of freedom. These values are identical to the values obtained from the two-sample form.

The verification for the chi-squared two-sample and multisample tests was based on a distribution of angles representing the data from Batschelet (1981, p. 111, Table 6.4.1). Both the two-sample and the multisample chi-squared programs produced a chi-squared value of 77.0271. Batschelet (1981, p. 110) indicates a chi-squared value of 77.3. Batschelet's figures when run on a Hewlett Packard 41CV calculator with the H.P. statistical module produced a chi-squared value of 77.0.

The Watson two-sample U^2 test was verified using two sets of data. The first set was taken from Mardia (1972, p. 202, example 7.11). The program produced a U^2 value of 0.3205. Mardia shows a U^2

value of 0.3204. To verify the Mardia figure it is necessary to break the ties in the same manner as he suggests. The second set of data was taken from Batschelet (1981, p. 118, Table 6.6.2). The program produced a rounded value of 0.261, identical to the Batschelet figure.

Verification of the runs test was conducted using data from Mardia (1971, p. 204, example 7.12). The program defined four runs with nine elements in the first set and ten in the second set. Mardia's values are identical to those obtained from the computer program.

Bivariate Methods

The bivariate methods as addressed by Batschelet are expressed in terms of linear variables. Corresponding circular methods, therefore, are not applicable here. The discussion that follows is limited to areas where linear data may present a particular problem in interpretation. Computer programs were written to describe the bivariate statistics used. A description of the computer programs for Hotelling's one and two-sample tests is provided.

Each point in a Cartesian space may be described as an ordered pair. These ordered pairs in turn may be described as a set of vectors having various lengths. Consequently a mean vector may be provided through the use of vector algebra. This is not the same as the mean vector, r, associated with angular data. The ordered pairs (x,y) consist of two dependent variables. Each ordered pair may be described as a vector with length:

$$r = (x^2 + y^2)^{\frac{1}{2}} .$$

Circular data requires that the radius of each data point be one.

A confidence ellipse developed by Hotelling is defined by Batschelet as "...a region that covers the population centre with a given probability" (p. 141). It is analogous to the confidence interval associated with univariate methods.

The statistical tests associated with the bivariate methods are identical to those found in linear statistics, and are well known. That is to say, the correlation procedures are based upon Pearson's correlation coefficient, r. Hotelling's one-sample and two-sample tests are identical to those used in linear statistics. Point estimation techniques used in linear statistics are based upon an elliptical model.

Discussion of the Programming Process

MenuF consists of three linear bivariate procedures, General Information, Hotelling's one-sample test and Hotelling's two-sample test. The three procedures are from the field of linear statistics. Batschelet (1981) includes a discussion of bivariate linear methods in his work on circular statistics. Since MenuF considers linear methods, the discussion of the programming process will be abbreviated.

The procedure entitled, General Information, provides the user with the means, variances, and the standard deviations of the two variables. The covariance and correlation are depicted along with a table showing the deviations used in their calculations. Two inputs are required. The user is requested to enter the number of ordered pairs and then the coordinates of each ordered pair. The verification of the program was based on data from Batschelet (1981, p. 133). The eight statistics produced by the program procedure, General Information, were identical to the values obtained by Batschelet.

Hotelling's one-sample test is used to test if $(\bar{x},\bar{y})$ deviate significantly from the origin. The input consists in the number of ordered pairs and the coordinates of each ordered pair. The program provides the user with the degrees of freedom and requests an F value as input. The calculated T and critical T values are provided along with the decision rule. The procedure was verified using data from Batschelet (1981, p. 148). An F value of 10.92 was entered based on 2 and 6 degrees of freedom and an alpha level of 0.01. The program provided a calculated T^2 value of 33.15 and a critical T^2 value of 25.48, compared to Batschelet's calculated value of 33.1 and critical value of 25.5.

Hotelling's two-sample test considers if the centers of the two samples differ significantly from each other. The input is similar to the one-sample test except, two sample sets are entered. The program provides the user with the degrees of freedom necessary to select an F value. The F value is entered and the program supplies the calculated T^2 value and the critical T^2 value. The verification of the procedure used data from Batschelet (1981, p. 153). An F value of 7.56 was used based on 2 and 10 degrees of freedom. The program produced a calculated T^2 value of 135.39 compared to Batschelet's 135.4. A critical T^2 value of 16.63 was obtained from both the program and Batschelet.

Linear and Circular Regression

Regression is a bivariate method. In linear statistics, regression and correlation belong to a unified theory. The main purpose may be either descriptive or inferential. Batschelet (1981) describes the purpose of circular regression as "...fitting regression lines and estimating parameters such as mean level, amplitude, and acrophase-...However, tests of significance are not always available. In short: Regression analysis is mainly a descriptive tool" (p. 178).

The data consists of ordered pairs, where the first variable represents the independent variable and the second variable represents the dependent variable. Linear statistics requires that both variables be represented as linear measurements. Batschelet (1981) considers circular-linear regression, where the independent variable is expressed as an angular measurement and the dependent variable is expressed as a linear measurement. The comparison of linear and circular regression will consist of linear-linear and circular-linear data.

Linear regression uses a least squares method to fit a straight line to the data. The equation of the straight line is defined as:

$$z'_y = r(z_x)$$

where:

z'_y is the predicted standard score of Y

r is the coefficient of correlation between X and Y

z_x is the standard score value from which z' is predicted

The assumptions made by linear regression are stated by Glass (1984) as "The regression procedures that have been discussed and illustrated make three assumptions:

1. The Y scores are normally distributed at all points along the regression line; that is, the residuals are normally distributed. (There is no assumption that the independent variables are normally distributed.)

2. There is a linear relationship between the Y's and the $\hat{Y}$'s at all points along the straight regression line, the residuals have a mean of zero.

3. The variance of the residuals is homogenous at all points along the regression line. This characteristic is known as homoscedasicity" (p. 141).

Circular-linear regression fits a sinusoidal curve to the data using the method of least squares. The general form of the equation is:

$$g(t) = M + A\cos(\omega t - \phi_1) + A\cos(2\omega t - \phi_2) + \ldots + A\cos(k\omega t - \phi_k).$$

Batschelet (1981) lists three criteria to be considered when fitting the trigonometric polynomial:

a) The model should fit the data points with sufficient accuracy

b) The model curve should be as smooth as possible

c) The number of parameters should be as small as possible since too many parameters are hardly meaningful.

To attain a model with sufficient accuracy requires the use of several parameters. This in turn conflicts with the need to limit the number of parameters. The comparisons made here restrict the trigonometric polynomial to the first harmonic (k=1). Batschelet (1981) defines the first harmonic as "A simple model of a periodic phenomenon is a sine curve given by the equation

$$Y = M + A \cos \omega (t - t_0).$$

Here, t is the independent variable, usually the time, subject to a certain period, T. The equation contains four parameters:

M = mean level or mesor,

A = amplitude ($A \geqq 0$),

ω = angular frequency

t_0 = peak phase or acrophase (Greek akrso = high)

The angular frequency, ω, is related to the period, T, by the formula

$$\omega = 360/T \text{ or } 2\pi/T." \quad \text{(p. 159)}$$

It is assumed that the period T be either 360 degrees or 24 hours which results in the angular frequency being one. Time series analysis must be used when the period is estimated. The three remaining parameters are estimated using the system of normal equations established by Gauss. A function g(t) consisting of the first and second harmonic would require that five parameters be estimated.

Hays (1981) refutes the linear assumption that the Y scores be normally distributed about the regression line in these words: "It is not necessary to make any assumptions at all about the form of the distribution, the variability of Y scores within X columns or 'arrays,' or the true level of measurement represented by the scores in order to employ linear regression..." (p. 460). When inferential methods are considered this assumption must not be disregarded. Batschelet (1981) issues the following caution regarding the use of the first harmonic, "Experience has shown that linear models are not always suitable even when there is only one peak and one trough within a cycle" (p. 162). Here, the "linear model" refers to the first harmonic. The use of the second and third harmonics may be necessary to accurately describe the data. The increase in the number of parameters impedes interpretation.

Linear and Circular Correlation

Linear correlation is usually considered a descriptive statistic. The measure of linear correlation considered here are Pearson's product-moment. Linear correlation establishes an association between two variables. The measure of association varies between positive one and negative one with a correlation of zero indicating that no relationship exists between the two variables.

Circular correlation is concerned primarily with statistical inference. The type of circular data used determine the form of circular correlation to be used. Correlation for five forms of circular data are considered: circular-circular data (when the angles are uniformly distributed); a parametric test for circular-circular data; a nonparametric test for circular-circular data; a parametric test for linear-circular data; and a nonparametric test for linear-circular data.

Pearson's Product-Moment Coefficient of Correlation

Pearson's product-moment correlation assumes that the line of best fit is a straight line. The bivariate distribution should exhibit the property of homoscedasticity. Both distributions composing the bivariate distribution are assumed to be continuous. If the stated assumptions are disregarded, the product-moment correlation lacks precision.

The product-moment correlation is defined by Minium (1978) as:

$$"r = \frac{\Sigma xy}{n \quad S_X S_Y}$$

where: $x = (X - \bar{X})$, $y = (Y - \bar{Y})$, xy is the sum of the products of the paired deviation scores, n = number of pairs of scores, and S_X,

S_Y are the standard deviations of the two distributions" (p. 146). The Σ xy/n product is known as the covariance.

Interpretation of the correlation coefficient requires that the assumptions stated previously are met. If the line of best fit is not a straight line, the correlation coefficient is underestimated. When the characteristic of homoscedasticity does not apply, the value of the association varies. The correlation coefficient may be overestimated at one point and underestimated at another. If the two variables are discontinuous, the correlation coefficient is over estimated. The degree of relationship cannot be interpreted in direct proportion to the magnitude of the correlation coefficient.

Circular Correlation

Circular correlation was developed in the 1970s and early 1980s. Unlike linear correlation which is primarily a descriptive tool, circular correlation is classified as an inferential statistic. The purpose of circular correlation is stated by Batschelet (1981): "We are interested in making statistical inference from the sample to the population (p. 178). Five forms of circular correlation are considered based upon the type of data and the restrictions placed upon that data. Each form of circular correlation is discussed individually. It is assumed that each pair of observations is independent of the next pair of observations.

Circular-circular correlation. We first considered the special case of paired samples where the variates are assumed to be uniformly distributed on the circle. The purpose of the test is to determine if the variates are correlated.

Circular correlation is based on the difference between the two variates, $(\omega_i - \phi_i)$. A high correlation exists if the differences are clustered about a mean angle. The test statistic will be the length of the mean vector:

$$r = [(\Sigma \operatorname{Cos} \delta_i)^2 + (\Sigma \operatorname{Sin} \delta_i)^2]^{\frac{1}{2}} / n$$

The length of the mean vector varies between 1 and 0. Negative correlation may be obtained by substituting $-\phi$, for ϕ . Mardia (1975) suggested that the larger of the correlations be used to represent the association. A distribution of r values was calculated by Stephens (1979) using $r\sqrt{n}$ as the test statistic. Batschelet (1981, p. 183, Table 9.2.2) provides a table of critical values.

Circular-circular data (nonparametric form). The second form of circular correlation assumes a random sample of bivariate measurements. The measurement for each of the variates must be expressed as angular data. A nonparametric test using rank correlation is proposed. The circle is divided into n equal arcs. Each angular measurement is assigned an angular measure consistent with its rank. For example, if there were 10 pairs of angles and the second ordered pair ranked (1,3), then the angular measures assigned would be (36°,108°). The mean vector is then calculated using the same method as stated previously,

$$r = \text{Max}\ (r_+ , r_-).$$

Batschelet (1981, p. 187) defined the test statistic as:

$$r^2 = -(n - 1)^{-1} \ln [1 - (1 - P)^{\frac{1}{2}}].$$

A table of critical values for r^2 was developed by Mardia (1975).

As a nonparametric test, ties are not permitted and must be broken. This requirement restricts the use of this form of correlation. Group data may not be used with rank correlation.

Circular-circular data (parametric form). The third form of circular correlation assumes a random sample of bivariate data. Both variates are expressed in terms of angular measures and are assumed to be independent. The parametric test was proposed by Jupp and Mardia (1980).

The mathematical concept upon which the test is based is described by Batschelet (1981) as: "Mathematically, what we do is to calculate the sum of the squared carried correlations between the vectors" (p. 190). He continues to define the correlation as:

$$\begin{aligned} r = {} & [r^2_{cc} + r^2_{cs} + r^2_{sc} + r^2_{ss} \\ & + 2(r_{cc}\ r_{ss} + r_{cs}\ r_{sc})\ r_1\ r_2 \\ & - 2(r_{cc}\ r_{cs} + r_{sc}\ r_{ss})\ r_2 \\ & - 2(r_{cc}\ r_{sc} + r_{cs}\ r_{ss})\ r_1] \\ & / [(1 - r^2_1)\ (1 - r^2_1)] \end{aligned}$$

When n is sufficiently large, the test statistic, nr^2, reflects a chi-squared distribution with four degrees of freedom. The chi-squared distribution holds only if the angular measures are independent. The well known restriction regarding parametric tests holds here as well.

Linear-circular data (nonparametric form). The fourth procedure for a correlation coefficient was proposed by Mardia (1976). The nonparametric test is based on the use of rank correlation. This procedure assumes that the ordered pair is composed of a linear and circular data. It is recommended that this test be used on data which exhibit a single peak and trough per period.

The procedure used to develop the test statistic U_n requires that both ordinates be ranked. The circular variate is first ranked and then expressed as:

$$\phi' = i \bullet \varepsilon (i = 1,2,3,\ldots,n) \quad \text{where } \varepsilon = 360/n$$

If an ordered pair had a circular rank of 6 and n = 8, then the modified value would equal 270 degrees. This method is similar to that used in the second procedure. The rank of the linear variate is assigned the number of unit masses the rank represents. If the linear variate had a rank of two, it would be assigned a value of two. Batschelet (1981) defines the mean vector as:

$$r = (C^2 + S^2)^{\frac{1}{2}}$$

where:

$$C = \Sigma\, r_i \operatorname{Cos} \phi_i', \qquad S = \Sigma\, r_i \operatorname{Sin} \phi_i'$$

The use of unit masses prevents the value of r to ever equal one. The following adjustment was proposed by Mardia (1976) along with a table of critical values.

$$U_n = a_n (C^2 + S^2).$$

Batschelet (1981, p. 195) defines the test statistic as "most suitable for a test statistic is the quantity U_n defined by:

$$U_n = c_n D_n$$

where:

$$c_n = 24 / [a_n n^2 (n + 1)]."$$

The value, c_n, equals:

$[n_1 + 5\cot^2 (\pi / n) + 4\cot^4 (\pi / n)]^{-1}$ (for even n)

$2 \sin4 (\pi / n) / [1 + \cos (\pi / n)]^3$ (for odd n)

As a nonparametric test, ties must be eliminated. The distribution may be skewed, however, the assumption that there exists one peak and one trough per period is required for meaningful results.

Linear-circular data (parametric form). The fifth and final correlation procedure considered is a parametric test where one variate is linear and the other, circular. Batschelet (1981, p. 193) states the restrictions placed upon the test:

"This procedure is parametric for two reasons:

(a) It is only applicable if the regression line is (approximately) sinusoidal.

(b) The joint distribution of x and y should be (approximately) normal."

If the linear and circular variates are independent, the test statistic $n * r^2$ approximates a chi-squared distribution with two degrees of freedom, when n is large.

The circular correlation used in the test statistic is defined as:

$$r^2 = (r^2_{yC} + r^2_{yS} - 2r_{yC}\, r_{yC}\, r_{CS}) \,/\, (1 - r^2_{CS})$$

where:

$$r_{yC} = \text{corr}\ (y, \cos\phi),$$
$$r_{yS} = \text{corr}\ (y, \sin\phi),$$
$$r_{CS} = \text{corr}\ (\cos\phi, \sin\phi).$$

The three correlations depicted are calculated as Pearson's product-moment correlations.

Discussion of the Programming Process

MenuG consists of five correlation procedures. The main menu requires the user to select one procedure. Each procedure provides a secondary menu stating the assumptions upon which the test is based and a reference indicating where a table of critical values may be found. The first procedure for circular-circular data where both distributions are uniform provides the user with both the positive and negative correlation. It identifies the larger value as the test statistic. The second procedure is a parametric test for circular-circular data. The user is provided with the correlation, the number of ordered pairs, the correlation squared and the test statistic. The third choice is a nonparametric test for circular-circular data. It provides the same information as choice "a", but uses r^2 as the test statistic. The fourth procedure, a parametric test for linear-circular data, produces output identical to procedure two. The fifth procedure is a nonparametric procedure for linear-circular data. The output is shown as the test statistic, U^2.

The data input process is the same for all five procedures. The program requires the user to enter the total number of ordered pairs followed by the measurement of each variate. A space must separate each variate. When linear-circular data is used, the linear measurement must be entered before the angular value.

Verification of the procedures for the first procedure was made using data from Batschelet (1981, p. 179, Table 9.2.1). The program showed a correlation of 0.8643, identical to Batschelet's value. A difference was noted between the negative correlations. Verification of the second procedure used the same data source as procedure one. The program provided a test statistic equal to 11.8397; when rounded it is equivalent to Batschelet's figure of 11.84. The program for the nonparametric form of a circular-circular correlation used the same Pascal functions as the first procedure. Identical results to those shown by Batschelet were obtained when rounded. This verified that the function used to compute the negative correlation was provided accurate output. The data used to verify procedure four was taken from Batschelet (1981, Table 9.4.1). The correlations used to compute the verifying test statistic were performed on the Hewlett Packard 41C statistical Module and similar results were obtained using EPISTST statistical programs. This statistic was compared to that provided by MenuG and found to be identical. The fifth procedure used the same data as procedure four. The test statistic, U_n, was found to be identical to the value shown by Batschelet, when rounded.

CHAPTER IV

RESULTS

4.1 Introduction

Comparisons of linear and circular statistical methods are based upon two educational hypotheses. Two forms of circular data are used in the comparisons: directional and periodic. The directional data consist of an angular measure depicting the community in which students of Carmel High School for Boys (hereafter, Carmel) resided. (See Appendix D for letter authorizing use of the Carmel data.) The estimated family income of a Carmel student is based upon the 1980 census report for the six-county metropolitan Chicago area. The periodic data express student attendance in Unit District R as an angular measure. The comparisons include attendance data from three instructional levels, elementary, junior high school, and senior high school. In addition, hypothetical data are used to demonstrate selected circular statistical procedures.

The research hypotheses are analyzed using the circular and linear statistical methods described in Chapter III. A conclusion based upon the statistical evidence is drawn for each hypothesis. The educational decisions based upon the circular and linear methods of analysis are compared.

4.2 Carmel High School for Boys' Study

Hypothesis #1: The Tuition at Carmel High School For Boys Should Be Increased

The first data set considers the feasibility of a tuition increase at a boys' secondary Catholic high school. Many factors affect such a decision. Here, the ability of a family to afford a tuition increase is considered. The median income of the community in which the family resides is used as an estimate of the family's income. The median income of the community is obtained from the 1980 census.

Description of the Data Set

The data set used in the present study was obtained from Carmel High School for Boys, located in Mundelein, Illinois, a northern suburb of Chicago. The data set provided by the school consists of the number of students residing within a community. The study requires that a community be included in the 1980 census report. Students residing in communities not identified by the 1980 census report are placed in the nearest census-reported community.

The estimated family income is the median income of the community in which the family resides. The income and population demographics used in the present study are reported by the Northeastern Illinois Planning Commission. The rank of the communities used in the present study are supplied by the Chicago Tribune (9/17/82, Section II, page 3) which published a ranking of 262 Chicagoland communities by median income as reported by the 1980 census.

Two sets of data are collected: one for use with linear methods of analysis and the other for use with circular methods of analysis. The linear data set consists of four variables: 1) the number of Carmel students residing within a community, 2) the rank of the community determined by the median income of the community, 3) the community's median income, and 4) the community's estimated total income based upon its median income and the number of Carmel students who reside within the community. Carmel provided the data set containing the number of students attending the school as of October 1, 1982 and the communities in which they resided. The estimated total income of students attending Carmel from a specified community is the product of the number of students residing within that community and the community's median income. Table 4.2.1 summarizes the linear data used.

The circular data set consists of the number of students attending Carmel by community, the angular measure showing the direction of the community, the median income of the community, and the total income of the community based on the number of students attending Carmel from that community. The number of students, median income, and total income are considered above. The angular measurement is obtained by locating each community on a map of northern Illinois. Carmel is located in Mundelein, Illinois and Mundelein, therefore, is assigned an angular measure of 0°. A protractor is used to measure the angular direction of each community in which Carmel students reside. Figure 4.2.1 depicts the angular distribution of the data set. Table 4.2.2 summarizes the circular data.

Table 4.2.1

Summary of Linear Data

Community	Number of Students	Rank	Income	Total Income
Barrington Hills	2	6	$75,078	$ 150,156
Lincolnshire	3	8	70,752	212,256
Riverwoods	2	11	64,235	128,470
Kildeer	9	14	61,064	549,576
Highland Park	4	17	57,612	230,448
Long Grove	11	24	52,492	577,412
Deerfield	2	29	49,076	98,152
Lake Bluff	1	30	49,013	49,013
Hawthorn Woods	12	38	44,430	533,160
Barrington	22	41	43,853	964,766
Libertyville	110	51	37,525	4,127,750
Palatine	1	53	36,614	36,614
Buffalo Grove	21	68	34,194	718,074
Wadsworth	5	71	33,807	169,035
Gurnee	15	91	31,473	472,095
Lake Zurich	27	96	31,000	837,000
Cary	5	105	30,313	151,565
Vernon Hills	11	108	30,225	332,475
Lindenhurst	14	111	30,108	421,512
Crystal Lake	1	114	29,908	29,908
Mundelein	98	121	29,200	2,861,600
Wheeling	4	149	27,476	109,904
Grayslake	35	153	27,263	954,205
McHenry	4	174	26,241	104,964
Island Lake	4	184	25,575	102,300
Wauconda	34	199	25,247	858,398
Lake Villa	13	211	24,648	320,424
Antioch	9	212	24,622	221,598
Waukegan	45	214	24,536	1,104,120
Fox River Grove	1	215	24,513	24,513
Round Lake Beach	13	217	24,324	316,212
Highwood	1	226	24,060	24,060
Zion	4	236	23,091	92,364
Round Lake Heights	1	237	23,045	23,045
Round Lake Park	3	244	22,557	67,671
Round Lake	11	247	22,400	246,400
Fox Lake	16	253	21,602	345,632
North Chicago	41	259	20,290	831,890
Total	615			$19,398,737

Figure 4.2.1

Graph of Frequency and Mean Angle

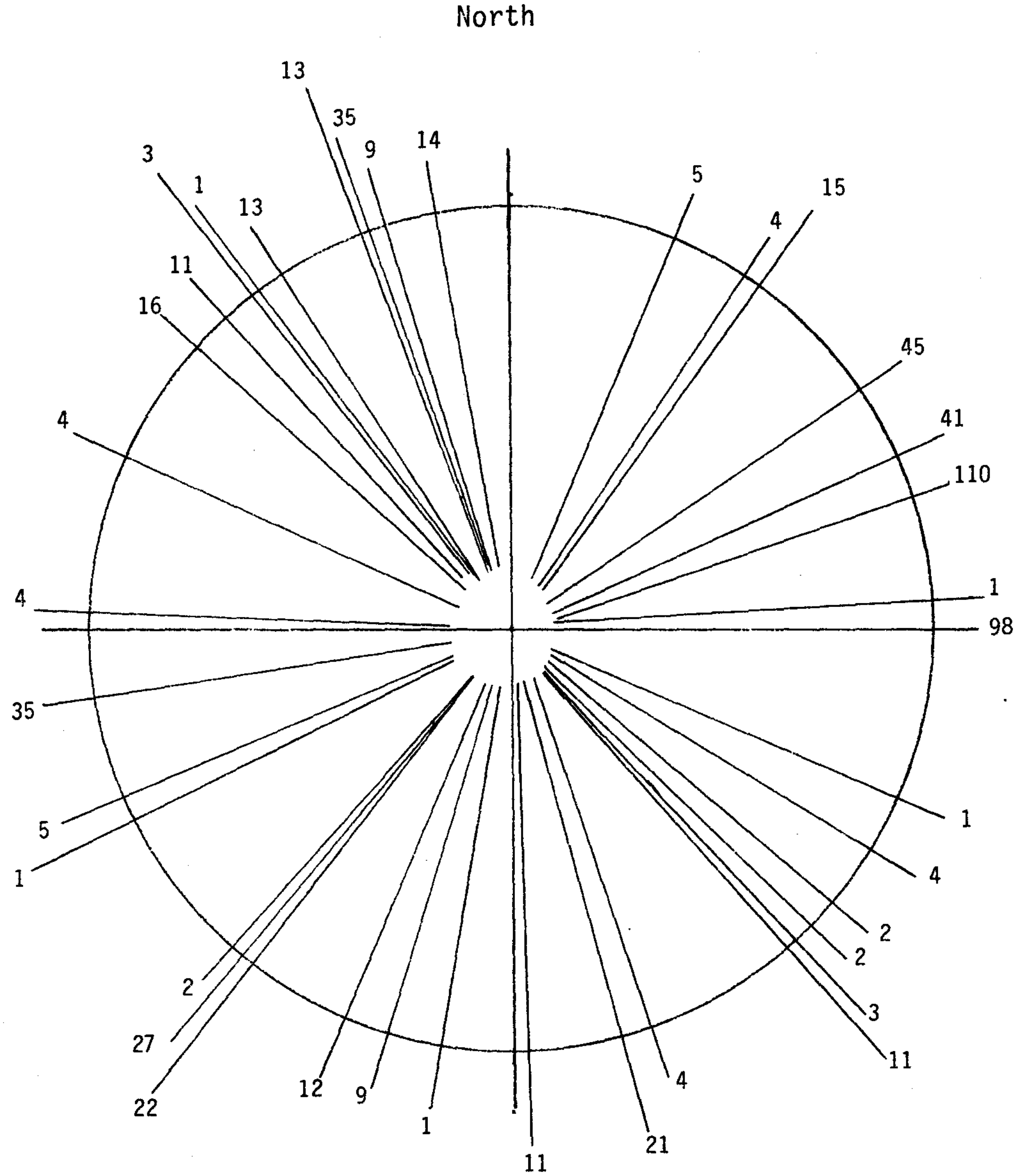


Table 4.2.2

Summary of Circular Data

Community	Number of Students	Angle	Rank	Income	Total Income
Barrington Hills	2	227	6	$75,078	$ 150,156
Lincolnshire	3	313	8	70,752	212,256
Riverwoods	2	316	11	64,235	128,470
Kildeer	9	252	14	61,064	549,576
Highland Park	4	329	17	57,612	230,448
Long Grove	11	272	24	52,492	577,412
Deerfield	2	320	29	49,076	98,152
Lake Bluff	1	4	30	49,013	49,013
Hawthorn Woods	12	246	38	44,430	533,160
Barrington	22	231	41	43,853	964,766
Libertyville	110	19	51	37,525	4,127,750
Palatine	1	260	53	36,614	36,614
Buffalo Grove	21	286	68	34,194	718,074
Wadsworth	5	68	71	33,807	169,035
Gurnee	15	55	91	31,473	472,095
Lake Zurich	27	229	96	31,000	837,000
Cary	5	202	105	30,313	151,565
Vernon Hills	11	312	108	30,225	332,475
Lindenhurst	14	100	111	30,108	421,512
Crystal Lake	1	187	114	29,908	29,908
Mundelein	98	0	121	29,200	2,861,600
Wheeling	4	289	149	27,476	109,904
Grayslake	35	109	153	27,263	954,205
McHenry	4	155	174	26,241	104,964
Island Lake	4	177	184	25,575	102,300
Wauconda	34	187	199	25,247	858,398
Lake Villa	13	110	211	24,648	320,424
Antioch	9	107	212	24,622	221,598
Waukegan	45	35	214	24,536	1,104,120
Fox River Grove	1	205	215	24,513	24,513
Round Lake Beach	13	122	217	24,324	316,212
Highwood	1	337	226	24,060	24,060
Zion	4	57	236	23,091	92,364
Round Lake Heights	1	126	237	23,045	23,045
Round Lake Park	3	127	244	22,557	67,671
Round Lake	11	132	247	22,400	246,400
Fox Lake	16	137	253	21,602	345,632
North Chicago	41	25	259	20,290	831,890
Total	615				$19,398,737

Discussion of Test Results

The data sets are analyzed using the descriptive and inferential methods discussed in Chapter III. The descriptive methods consider measures of central tendency and variability. Linear correlation is considered an inferential technique in the present study to permit a comparison between linear and circular correlation. Other inferential tests include the Rayleigh test, V-test, the Watson U^2 test, and the chi-squared test.

Measures of Central Tendency and Dispersion

Linear Methods of Central Tendency

The linear measures of central tendency used are the arithmetic mean and the median. The arithmetic mean, reflecting the average community income of the 615 students attending Carmel, is $31,542.66. The median income for the 615 students is $29,200.00. The student representing the median income lives in Mundelein. The four communities with the greatest student enrollment are: Libertyville (110), Mundelein (98). Waukegan (45), and North Chicago (41).

Linear Measures of Dispersion

The measures of dispersion used are variance, standard deviation and range. The variance depicts the spread or dispersion of the estimated community income in dollars squared. The variance based upon estimated student income by community is 83,836,914 squared dollars. The standard deviation ($9,156.25) reflects the dollar variability associated with the estimated family incomes. The range of the median income is $54,788.00. The most affluent community is Barrington Hills, ranked sixth in median income, $75,078. The least affluent community was North Chicago, ranked 259th with a median income of $20,290 (Chicago Tribune, 1982).

The arithmetic mean and standard deviation are calculated using the Hewlett-Packard statistical module, 00041-90030. These values are confirmed by the EPISTST version 3.1 statistical package.

Two measures of skewness and kurtosis are used in the present study. The first values are based upon the distribution of students. The measure of skewness (2.73) indicates a positively skewed distribution. Kurtosis for the distribution is 10.36 units. The second distribution considered is community income based upon the number of students attending Carmel from that community. The distribution is positively skewed (1.69 units), kurtosis (6.90 units).

Circular Measures of Central Tendency

The circular measures of central tendency include the mean angle, the mean vector, and the median angle. The computer program, MEANDIR, provides each of these measures. The mean angle based upon the angular measurement for each of the 615 students residing within a specific community is 26.0 degrees. The coordinates of the mean vector are

$$x = 0.2861$$

$$y = 0.1394.$$

The magnitude of the mean vector is

$$r = 0.3183.$$

The median angle is 55.0°. Table 4.2.3 depicts the angular values used in determing the median. The community corresponding to the median angle is Gurnee, Illinois.

Circular Measures of Dispersion

The circular measures of dispersion include the angular range, angular variance and the mean angular deviation. The values are obtained from the computer program MENUA. The angular range based upon the 615 angular measurements is 328.0° No adjustment in terms of Mod(360) is necessary. The angular variance is 78.1° squared using Batschelet's formula for variance. The angular deviation about the mean vector is 66.90°.

MENUA provides circular measures of skewness and kurtosis. A skewed value of -0.0397 is obtained for the distribution and kurtosis value of 0.2146. Both measures indicate minimal effect on the distribution.

Table 4.2.3

Data Used in Calculating Mean Angle and Magnitude

	Frequency	Angle	Cosine	Sine	Total Cosine	Total Sine
1.	98	0.0	1.0000	0.0000	98.0000	0.0000
2.	1	4.0	0.9976	0.0697	0.9976	0.0697
3.	110	19.0	0.9455	0.3256	104.0070	35.8126
4.	41	25.0	0.9063	0.4226	37.1586	17.3274
5.	45	35.0	0.8192	0.5736	36.8618	25.8110
6.	15	55.0	0.5736	0.8192	8.6036	12.2873
7.	4	57.0	0.5446	0.8387	2.1785	3.3547
8.	5	68.0	0.3746	0.9272	1.8730	4.6359
9.	14	100.0	-0.1737	0.9848	-2.4311	13.7873
10.	9	107.0	-0.2924	0.9563	-2.6314	8.6067
11.	35	109.0	-0.3256	-.9455	-11.3950	33.0931
12.	13	110.0	-0.3420	0.9397	-4.4463	12.2160
13.	13	122.0	-0.5299	0.8480	-6.8890	11.0246
14.	1	126.0	-0.5878	0.8090	-0.5878	0.8090
15.	3	127.0	-0.6018	0.7986	-1.8055	2.3959
16.	11	132.0	-0.6691	0.7431	-7.3605	8.1746
17.	16	137.0	-0.7314	0.6820	-11.7017	10.9119
18.	4	155.0	-0.9063	0.4226	-3.6252	1.6904
19.	4	177.0	-0.9986	0.0523	-3.9945	0.2093
20.	35	187.0	-0.9925	-0.1219	-34.7391	-4.2657
21.	5	202.0	-0.9272	-0.3746	-4.6359	-1.8731
22.	1	205.0	-0.9063	-0.4226	-0.9063	-0.4226
23.	2	227.0	-0.6820	-0.7314	-1.3640	-1.4627
24.	27	229.0	-0.6561	-0.7547	-17.7134	-20.3773
25.	22	231.0	-0.6293	-0.7772	-13.8449	-17.0974
26.	12	246.0	-0.4067	-0.9135	-4.8807	-10.9626
27.	9	252.0	-0.3090	-0.9511	-2.7811	-8.5595
28.	1	260.0	-0.1736	-0.9848	-0.1736	-0.9848
29.	11	272.0	-.0349	-0.9994	0.3840	-10.9933
30.	21	286.0	0.2756	-0.9613	5.7886	-20.1864
31.	4	289.0	0.3256	-0.9455	1.3023	-3.7821
32.	11	312.0	0.6691	-0.7431	7.3605	-8.1745
33.	3	313.0	0.6820	-0.7313	2.0460	-2.1940
34.	2	316.0	0.7193	-0.6946	1.4387	-1.3893
35.	2	320.0	0.7661	-0.6428	1.5321	-1.2856
36.	4	329.0	0.8572	-0.5150	3.4287	-2.0601
37.	1	337.0	0.9205	-0.3907	0.9205	-0.3907
	615				175.9746	85.7558

Inferential Methods

Linear Methods

The linear inferential methods used are the chi-squared test and linear correlation. The chi-squared test is used to determine if the students attending Carmel are uniformly distributed by community. The correlation study considers the relationship between the estimated median income of the community and the number of students attending Carmel from that community.

Tests of Randomness. A chi-squared value is calculated based on the number of students per community. The test requires the data to be grouped in quantities of 5 or more. Twelve communities shown on Table 4.2.1 had fewer than five students. These communities are combined with adjacent communities to meet the chi-squared requirement of five or more. The calculated chi-squared value, based on the 26 groups is 738.90. The critical chi-squared value at the .01 level of significance and 25 degrees of freedom is 44.31 (Batschelet, 1981, Table G, p. 542). The null hypothesis of uniformity is rejected.

Linear Correlation. Pearson's correlation coefficient comparing the number of students attending Carmel and the estimated income of the community in which students reside is -0.17. The correlation coefficient is not significant at the .05 level of significance.

The calculated chi-squared value is obtained by using the statistical module for the Hewlett Packard 41C calculator, number 00041-90030. The correlation information is provided by the EPISTAT statistical program for the I.B.M. personal computer.

Circular Methods

The inferential circular methods used in the present study are tests for randomness, tests for goodness-of-fit, and circular correlation. Tests for randomness include the Rayleigh test and the V-test. The chi-squared and Watson U^2 tests (goodness-of-fit tests) determine if the data are uniformly distributed. A parametric and nonparametric form of circular-linear correlation determines the relationship between the angular direction of the community and the number of students attending Carmel from that community. The computer programs comprising the statistical methods are MENUB, MENUC, and MENUG.

Tests of Randomness. The Rayleigh test and V-test are used to test for randomness. The data consist of angular measurements representing the communities in which students resided. The Rayleigh test uses the magnitude of the mean vector, r, as a test statistic. MENUB provides several options subject to the characteristics of the data. The options selected are a unimodal distribution and a sample size in excess of 30. For a large sample size, a z-score may be used as a test statistic. The calculated z-score provided by MENUB is 62.3104. The critical value is 6.91 for an infinite sample size and a .001 level of significance. Significance is established and the null hypothesis of uniformity is rejected.

The V-test determines if the angular measurements cluster about a given angle. Two hypothetical angles are considered: Mundelein (0°) and Libertyville (19°). The first hypothetical angle considered is 0°, representing the community in which the Carmel is located

(Mundelein). MENUB provides a u value of 10.0352 with n = 615. The critical value of u for n = 1000 and a significance level of .01 is 2.2360. Significance is established and the null hypothesis is rejected. The second hypothetical angle considered is 19°, representing the community from which the greatest number of students are enrolled (Libertyville). The computer program provides a u value of 11.0806 with 615 subjects. The critical value used to determine significance is the same as the value used in the Mundelein illustration. The null hypothesis of uniformity is rejected.

Goodness-of-Fit Tests. The goodness-of-fit tests applied to the data are the chi-squared test and the Watson U^2 test. They are used to determine if the data are distributed uniformly. MENUC provides the computer programs.

The circular chi-squared test uses the direction of the community in which the student resides as the input data. There are 615 such entries. Two forms of the test are available. The form used assumes a uniform distribution. The computer program requires that the data be grouped. Batschelet (1982) suggests that twelve or more groups be used, therefore, the distribution is divided into 12 equal groups. For groups of this size, a correction for grouping is not necessary. A chi-squared value of 950.0535 with 11 degrees of freedom is obtained. The chi-squared value is significant at the .01 level. The data is regrouped into 360 equal partitions, each partition representing a single degree. The chi-squared test is then applied and is significant at the .01 level.

The Watson U^2 test provides another test of uniformity. Option "w" of MENUC selects the computer program. A calculated U^2 value of 5.1080 is obtained. A critical value of .302 is obtained with n=1000 and a .005 level of significance. The null hypothesis stating that the distribution is uninformly distributed is rejected.

Circular Correlation

The form of circular correlation used in the present study requires that one variate be expressed as a circular measure and the other as a linear measure. A parametric and nonparametric test are used in the analysis. The computer program, MENUG, provides the parametric and nonparametric procedures.

The parametric test considers the relationship between the angular direction of the community and the number of students attending Carmel from that community. A correlation of .3489 is obtained using MENUG, option "d." Correlation in circular statistics is treated as an inferential method. When n is relatively large (38), the test statistic ($n * r^2$) approximates a chi-squared distribution with 2 degrees of freedom. The calculated test statistic ($n * r^2$) is 4.5043. A probability of .1085 is obtained for the calculated value. The correlation is not significant at the .05 level.

The nonparametric form of the circular-linear correlation is applied to the data set. Again, MENUG is employed using option "e." A test statistic U is obtained with a value of 1.4809. The critical values taken from Table X of Batschelet are 5.9 at the .05 level of significance and 8.85 at the .01 level of significance (Batschelet, 1981, p. 352). The correlation is not significant at the .05 level of significance.

A circular-linear correlation is conducted to determine the relationship between the angular direction of the community and the median income of that community. Both the parametric and nonparametric forms of circular-linear correlation are used.

The parametric test produces a correlation of .6262. The test statistic ($n * r^2$) is 14.9014. The value of the correlation is significant at the .01 level. The nonparametric test produces a value of 17.7633 for the test statistic, U, which is significant at the .01 level of significance.

Comparison of Linear and Circular Statistical Methods

Many factors affect the decision process of a family considering a private education for their children. The present study is designed to demonstrate a method of estimating the financial ability of a family to afford a tuition increase. The decision regarding a tuition increase is based upon the results of the linear and circular tests discussed previously.

Linear Methods of Analysis

The three areas of linear analysis used in this study are measures of central tendency, measures of dispersion and inferential methods. Table 4.2.4 summarizes the statistical results derived from the linear methods.

Table 4.2.4

Summary of Linear Test Results

Classification	Statistical Result	Significance
Central Tendency		
1. Mean (Income)	$31,542.66	Not Applicable
2. Median (Income)	$29,200.00	Not Applicable
3. Mode (Income	$37,525.00	Not Applicable
Dispersion		
1. Variance (Income)	$83,863,916.00	Not Applicable
2. Standard Deviation (Income)	$9,156.25	Not Applicable
3. Range (Income)	$54,788.00	Not Applicable
4. Skewness (Income)	1.69	Not Applicable
5. Kurtosis (Income)	6.90	Not Applicable
6. Skewness (Students)	2.73	Not Applicable
7. Kurtosis (Students)	10.36	Not Applicable
Inferential Methods		
1. Chi-squared (Distribution of Students)	738.90	Yes at .01 Level
2. Pearson Correlation (Income-Number of Students)	-.17	No at .05 Level

The measures of central tendency are based on the 1980 census. The present study considers a tuition increase for the 1983-1984 academic year. To project the effect of the economy on family income and expenditures, the consumer price index is used to estimate the amount of money a family had available for education in 1983 compared to 1980. The consumer price index for the year 1980 is 246.8. For 1983 the consumer price index is 298.4, an increase of 20.9% for the three-year period. Table 4.2.5 depicts the disposition of income in the census year 1980. The measures of central tendency used to estimate a family's income are the adjusted mean ($38,135), median ($35,302), and the mode ($45,368).

The standard deviation indicates the income dispersion about the mean. The standard deviation of $9,156 represents a narrow spread. Within one standard deviation ($40,699 and $22,387), 490 of the 615 estimated family incomes are located. The percentage of estimated family incomes within one standard deviation of the mean is 79.7%. The range of $54,788 exhibits greater variability than the standard deviation. The rank of the communities depicted in Table 4.2.1 indicates the diverse economic backgrounds of Carmel students. A spread of 253 communities lie between the two extremes of Barrington Hills (6th) and North Chicago (259th) and accounts for the dispersion associated with the range.

The distribution is positively skewed, 1.69. Of the 615 represented incomes, 410 or 66.7% of the incomes are below the arithmetic mean income of $31,542. The distribution exhibits a kurtosis value of 6.90. Due to the number of enrolled students from Mundelein (98) and Libertyville (110), a high kurtosis value is expected. The positively skewed value indicates that the school's enrollment must be further scrutinized.

Table 4.2.5

Disposition of Personal Income for the Year 1980

Category	Quarters I	II	III	IV	Average
Personal Income	2,088.2	2,114.5	2,182.1	2,256.2	2,160.2
Personal Tax and Nontax Payments	323.1	330.0	341.5	359.2	338.5
Disposable Personal Income	1,765.1	1,784.1	1,840.6	1,897.0	1,821.7
Personal Consumption Expenditures	1,631.0	1,626.8	1,682.2	1,751.0	1,672.8
Interest Paid by Consumers to Business	46.7	46.3	46.0	46.8	46.4
Personal Transfer Payments to Foreigners	1.0	1.0	1.0	1.6	1.2
Personal Savings	86.4	110.0	111.4	97.6	101.3
Total Expenditures	1,765.1	1,784.1	1,840.6	1,897.0	1,821.7

Note: All figures shown are in billions of dollars. Some discrepancies exist due to rounding. The data were taken from the Economic Report of the President (1982).

Table 4.2.6

Summary of Directional Data by Community Rank

Community	Angle	Rank	Cosine	Sine	Cumulative Cosine	Cumulative Sine
Barrington Hills	227	6	-.6820	-.7314	-4.092	-4.5684
Lincolnshire	313	8	.6820	-.7314	2.0460	-2.1942
Riverwoods	316	11	.7193	-.6947	1.4386	-1.3894
Kildeer	252	14	-.3090	-.9511	-2.7810	-8.5599
Highland Park	329	17	.8572	-.5150	3.4288	-2.0600
Long Grove	272	24	.0349	-.9994	.3839	-10.9934
Deerfield	320	29	.7660	-.6428	1.5320	-1.2856
Lake Bluff	4	30	.9976	.0698	.9976	.0698
Hawthorn Woods	246	38	-.4067	-.9136	-4.8804	-10.9632
Barrington	231	41	-.6293	-.7772	-13.8446	-17.0984
Libertyville	19	51	.9455	.3256	104.0050	35.8160
Palatine	260	53	-.1737	-.9838	-.1737	-.9838
Buffalo Grove	286	68	.2756	-.9613	5.7876	-20.1873
Wadsworth	68	71	.3746	.9272	1.8730	4.6360
Gurnee	55	91	.5736	.8192	8.6040	12.2880
Lake Zurich	229	96	-.6561	-.7547	-17.7147	-20.3769
Cary	202	105	-.9272	-.3746	-4.6360	-1.8730
Vernon Hills	312	108	.6691	-.7431	7.3601	-8.1741
Lindenhurst	100	111	-.1737	.9848	-2.4318	13.7872
Crystal Lake	187	114	-.9926	-.1219	-.9926	-.1219
Mundelein	0	121	1.0000	.0000	98.0000	0.0000
Wheeling	289	149	.3256	-.9455	1.3024	-3.7820
Grayslake	109	153	-.3256	.9455	-11.3960	33.0925
McHenry	155	174	-.9063	.4226	-3.6252	1.6904
Island Lake	177	184	-.9986	.0523	-3.9944	.2090
Wauconda	187	199	-.9926	-.1219	-33.7484	-4.1446
Lake Villa	110	211	-.3420	.9397	-4.4460	12.2161
Antioch	107	212	-.2924	.9564	-2.6316	8.6076
Waukegan	35	214	.8192	.5736	36.8640	25.8120
Fox River Grove	205	215	-.9063	-.4226	-.9063	-.4226
Round Lake Beach	122	217	-.5299	.8481	-6.8887	6.2933
Highwood	337	226	.9250	-.3907	.9250	-.3907
Zion	57	236	.5446	.8387	2.1784	3.3548
Round Lake Heights	126	237	-.5878	.8090	-.5878	.8090
Round Lake Park	127	244	-.6018	.7986	-1.8054	2.3958
Round Lake	132	247	-.6691	.7431	-7.3601	8.1741
Fox Lake	137	253	-.7314	.6820	-11.7024	10.9120
North Chicago	25	259	.9063	.4226	37.1586	17.3266

Two inferential tests (the chi-squared test and Pearson's correlation coefficient) are used to describe the data. The chi-squared test rejects the null hypothesis of a uniform distribution at the .01 level of significance. The correlation between the number of students attending Carmel and the estimated income of the community in which they reside is not significant at the .05 level.

Carmel's tuition and fees in 1980 were $1,095. The proposed increase for the 1983-84 school year would raise the tuition and fees to $1,570. This representd a 43.4% increase for the four year period. During the same period the consumer price index increased 20.9%. The cost of education at Carmel was increasing at twice the rate of other consumer goods and services. Tables 4.2.7 and 4.2.8 depict the dispersion of personal expenditures.

Enrollment during the period from 1980 to 1982 decreased from 659 to 605 students. The estimated enrollment for the 1983-84 school year was 618.

Table 4.2.7

Personal Consumption Expenditures for the Year 1980

Category	Expenditure	Percent
Average Personal Income	$2,160.2	100.0
Taxes	338.5	15.7
Interest Paid by Consumers to Business	46.4	2.1
Personal Transfer Payments to Foreigners	1.2	.1
Personal Savings	101.3	4.7
Personal Consumption Expenditures	1,672.8	77.4
Total	$2,160.2	100.0
Personal Consumption by Expenditure		
Motor Vehicles and Parts	$ 89.9	4.2
Furniture and Household Equipment	84.6	3.9
Other Durable Goods	37.3	1.7
Food	345.7	16.0
Clothing and Shores	104.8	4.8
Gasoline and Oil	89.0	4.1
Other Nondurable Goods	136.2	6.3
Housing	272.0	12.6
Electricity and Gas	55.7	2.6
Other Household Expenses	56.0	2.6
Transportation	64.1	3.0
Medical Care	143.6	6.7
Other Services	193.9	8.9
Total	$1,672.8	77.4

Note: The data used in this table are taken from Tables B-14 and B-23 of the 1982 Economic Report of the President. Some discrepancies arise due to rounding. The original source of the data is the Department of Commerce, Bureau of Economic Analysis.

Table 4.2.8

Adjusted Income for Residents of Mundelein

Based on Consumer Price Index for 1980

Items	Consumer Price Index 1980	Consumer Price Index 1983	Percent of Increase
All Items	246.8	298.4	20.9
Food and Beverages	248.0	284.4	14.7
Housing Total	263.3	323.1	22.7
- Shelter	281.7	344.8	22.4
- Fuel and Utilities	278.6	370.3	22.9
- Furnishings	205.4	238.5	16.1
Apparel and Makeup	178.4	196.5	10.2
Transportation	249.7	298.4	19.5
Medical Care	265.9	357.3	34.4
Entertainment	205.3	246.0	19.8
Other Goods and Services	214.5	288.3	34.4
Energy	361.1	419.3	16.2

Table 4.2.9 adjusts the dollar value for personal consumption items. The median income is adjusted to reflect the consumer price index for the period from 1980 to 1983. The median is used for two reasons. First, it is the statistical method used by the census to describe the "average" income of the communities. Second, it provides a better estimate of the school's student population considering the skewness of the distribution. Table 4.2.9 implies that the adjusted family expenditures for 1983 are proportionately more than in 1980.

The median income is $35,303 and the expenditures total $35,480. This reflects a $177 reduction in available funds for family expenditures. Additional funds would be needed to meet the proposed tuition increase of $238. Tuition had increased two-fold over the consumer price index. A family would appropriate the additional money from such areas as savings and other services (entertainment, education, etc.) to offset a tuition increase.

The Lake County Economic Development Commission, using data from the 1980 census, reports the median income of a Lake County family as $28,045. This is $1,155 below the 1980 Carmel median income of $29,200. Apparently, Carmel drew from a population whose income was above the mean for the area. The concern demonstrated for the positively skewed distribution appears unwarranted. A family needed $415 in additional funds to offset the $238 proposed tuition increase and the $177 increase in household expenditures due to the effect of inflation on the consumer price index. The $415 reflects a 1.2% increase which must be obtained from other areas of goods and services. The decision supported by the linear analysis is to increase the tuition to $1,570.

Table 4.2.9

Estimated Distribution of the Median Income for a Carmel Family as of September 1983

Item	1980 Value	Consumer Price Adjustment	1983 Value
Median Income	$29,200	120.9	$35,303
	======	=====	======
Expenditures			
Taxes	$ 4,584	-	$5,115
Interest	613	120.9	741
Transfer Payments	29	120.9	35
Savings	1,372	120.9	1,659
Motor Vehicles	1,226	134.4	1,648
Furniture, etc.	1,139	116.1	1,322
Other Durable Goods	496	134.4	667
Food	4,672	114.7	5,359
Clothing	1,402	110.2	1,545
Gasoline and Oil	1,197	116.2	1,391
Other Nondurable Goods	1,840	134.4	2,473
Housing	3,679	122.4	4,503
Utilities	760	122.9	934
Other Household Expenses	760	120.9	919
Transportation	876	119.5	1,047
Medical Care	1,956	134.4	2,629
Other Services	2,599	134.4	3,493
Total Expenditures	$29,200		$35,480
	======		======

Circular Methods of Analysis

The direction in which the community lies for each of the 615 students constitutes the angular data used in the circular analysis. The decision concerning the tuition increase is based upon the circular methods. Table 4.2.10 summarizes the circular methods used in the decision-making process.

The circular statistics used to describe the measures of central tendency are the mean angle, the median angle and the mean vector. The direction of the mean angle is 26.0°. Figure 4.2.1 depicts the circular distribution. The mean angle lies in a northeasterly direction of 26.0°. North Chicago is the community whose angular measure is associated most closely with the mean angle. The median income lies in the same direction as the community of Gurnee, 44°. The length of the mean vector (.3183) is used in the Rayleigh and V-test to determine uniformity.

The angular deviation is 66.9°. Fifty-four percent of the student body came from communities within one angular deviation, 319.1° (26° - 66.9°) and 92.9° (26° + 66.9°) of the mean angle. Included within this span are the communities of Waukegan (35°) and North Chicago (25°). From these two communities Carmel drew 86 students. The median family income, as indicated by the census for Waukegan was $24,536 and for North Chicago was $20,290. Both communities were below the median family income of Lake County, $28,045. A tuition increase has greater impact on communities such as Waukegan and North Chicago. The angular range added little additional information.

Table 4.2.10

Summary of Circular Test Results

Classification	Statistical Result	Significance
Central Tendency		
1. Mean Angle	26.0°	Not Applicable
2. Median Angle	55.0°	Not Applicable
3. Mean Vector, r	0.3183	Not Applicable
Dispersion		
1. Angular Range	328.0°	Not Applicable
2. Angular Variance	78.1°	Not Applicable
3. Mean Angular Deviation	66.9°	Not Applicable
4. Skewness	-0.0397	Not Applicable
5. Kurtosis	0.2146	Not Applicable
Inferential Methods		
Tests for Randomness		
1. Rayleigh Test	62.3104	Yes at the .01 level
2. V-test Mundelein	10.0352	Yes at the .01 level
V-test Libertyville	11.0806	Yes at the .01 level
Goodness-of-Fit Tests		
1. Chi-squared Test	950.0535	Yes at the .01 level
2. Watson U^2 Test	5.1080	Yes at the .01 level
Circular Correlation		
Direction and Number of Students		
1. Parametric	4.5043	Not at the .05 level
2. Nonparametric	1.4809	Not at the .05 level
Direction and Community Rank		
1. Parametric	17.8692	Yes at the .01 level

The circular distribution is minimally skewed (-0.0397). This indicates that the data are evenly distributed about the mean angle. The measure of kurtosis is 0.2146. Both circular kurtosis and skewness values differ from the linear results.

The inferential circular methods used in the present analysis are classified as follows: tests for randomness, goodness-of-fit tests, and circular correlation. The present analysis uses the goodness-of-fit tests as tests of uniformity.

The Rayleigh test is a test of uniformity. The null hypothesis of a uniform distribution is rejected at the .01 level of significance. The Rayleigh test provides assurance that the angular direction of the communities in which Carmel students reside is not randomly distributed.

Batschelet (1982, p. 58), states the purpose of the V-test is "to test whether the observed angles have a tendency to cluster about the given angle...." Two hypothetical angles are tested. The first is 0° representing Mundelein, the community in which Carmel is located. The second is 19° representing Libertyville, the community in which the greatest number of students reside. The V-tests are significant at the .01 level, indicating that the distribution clusters about the hypothetical angles.

The chi-squared and Watson U^2, goodness-of-fit tests, confirm that the circular distribution is not uniformly distributed. These goodness-of-fit tests are used to demonstrate the computer applications.

Circular correlation establishes a relationship between the direction of a community and the number of students attending Carmel from that community. Two forms of circular-linear correlation are used (parametric and nonparametric). Significance is not established by either method. The correlational relationship between the directions of the communities in which the students reside and the number of students attending Carmel cannot be established. However, a significant correlational relationship exists between the direction of the community and the rank of the community by median income. The test statistic used to approximate a X^2 distribution with 2 degrees of freedom is 17.8692 reflecting a probability level of less than .001. A review of the community ranks (Table 4.2.2), indicates that the communities in the southwest direction are most affluent and the communities in the northwest direction are the least affluent.

Based upon circular analysis, the suggested action is to seek funding through sources other than a tuition increase. The majority of the students who attend Carmel (52.4%) are from communities whose directions lie between 319.1° and 92.9°. Several of these communities [North Chicago (25°), Zion (57°), and Waukegan (35°)] are ranked among the lowest by median income. Families in such areas generally have less money available for education. A substantial tuition increase could prohibit student attendance from these communities. For these reasons, alternate methods of funding are suggested.

Comparison of Linear and Circular Decisions

Linear analysis suggests a tuition increase. The median income of $29,200, based on the 1980 census is above the median income for the Lake County area, $28,045. The majority of students (33.8%) came from the Mundelein, Libertyville area whose median incomes are ranked in the upper 50% of Chicagoland communities. It appears the monies are available if the education warrants the expenditure.

Circular analysis indicates that many students come from communities ranked among the lowest in the six-county metropolitan Chicago area. North Chicago, the community which best represents the mean angle (26.0°), is ranked 259th. Students from the North Chicago, Waukegan area represent 14.0% of the enrollment. The funds available for education for families residing in these communities may be limited. A tuition increase may affect their choice of a private education for their children.

A review of the linear data in Table 4.2.11 indicates the potential error in the linear analysis. Only 47 of the 615 students reside in communities ranked in the third quarter. However, 31.2% (192) of the 615 students reside in communities whose income is in the fourth quarter. Approximately two-thirds of the students reside in communities ranked in the upper half. The circular methods of analysis draw attention to this inequity.

Table 4.2.11

Summary of Linear Data by Quarter

Quarter	Number of Students	Percent of Students	Total Income	Average Income
First	200	32.5%	$ 8,375,847	$41,879
Second	176	28.6	5,275,190	29,973
Third	47	7.7	1,271,373	27,050
Fourth	192	31.2	4,476,327	23,314
Total	615	100.0%	$19,398,737	$31,543

Note: These averages are somewhat misleading due to the fact that a family may have had more than one student attending the school. However, since the school educated only boys, this number is well below the two-plus children per family found on the Census survey.

The decision implemented by Carmel is to increase tuition. It appears the majority of the families (66%) could afford the tuition increase. However, $20,000 was set aside for financial aid to assist families demonstrating financial need. Both linear and circular methods of analysis are used to arrive at a fair decision. This is one of several steps taken by Carmel to reverse the declining enrollment. Enrollment has increased since 1982, reaching a high of 716 students in Fall 1986. Financial aid in 1986-87 is $25,000.

4.3 Unit District R Comparison

Hypothesis #2a: The Four Tracts at Intermediate School Y Receive the Same Instruction Time and

Hypothesis #2b: The Absentee Rate Is Affected by the Instructional Level

Many areas of the school system, from funding to the number of hours of instruction a student receives, are affected by school attendance. Periodic data representing the number of students absent per day affords an opportunity to review the time period when absences occur.

Unit District R provided three levels of instruction: high school, intermediate and elementary. Due to high enrollment, the school system operated on a twelve-month schedule (45-15) from July 1, 1978 to June 30, 1979. The students in each of the District's schools are divided into four tracts: A, B, C, and D. Each academic quarter is divided into four fifteen-day sessions with each tract attending three of the four sessions.

The present study was designed to examine the attendance at the Intermediate School Y to determine if each of the four tracts receive equivalent instructional time. Factors influencing absenteeism such as, time of occurrence, weather and instructional level are considered. The investigation compares absenteeism at the High School X, Intermediate School Y and Elementary School Z to determine the effect the instructional level has upon daily attendance. A special focus of the study is identification of the "average" time period when absenteeism occurs. Identification of factors causing absenteeism and suggestions to minimize instruction time lost are made. Educational data, i.e., daily absenteeism, are analyzed to illustrate various uses of linear and circular statistical procedures.

Description of Data

The data provided by Unit District R represent the daily attendance for the year beginning July 1, 1978 and ending June 30, 1979. Attendance figures were provided by the district for two senior high schools, four intermediate schools and nine elementary schools. The present study considers only three of the district schools: High School X, Intermediate School Y, and Elementary School Z. Primary emphasis is placed on Intermediate School Y. The attendance for each of the four tracts (A, B, C, D) is compared by academic quarter. The quarterly attendance for High School X, Intermediate School Y, and Elementary School Z is compared to determine whether absentee rates differ between levels of instruction. Appendix B depicts the date the absences occur, the three tracts that are in attendance, and the number of absences for High School X, Intermediate School Y, and Elementary School Z.

The daily attendance by tract at Intermediate School Y is shown in Appendix C. Column 1 displays the attendance dates. The number of days each tract attended school differ. Tracts A, B, and D are affected by weather-related school closings (4), a teacher institute day (1), and a national holiday (1). An adjustment is made so that each of the four tracts are scheduled for a 177-day school year. The yearly adjustment is made by substituting the arithmetic mean as the absentee rate for each of the aforementioned school closings. The yearly mean absentee rates are used to minimize fluctuation.

	Tracts			
	A	B	C	D
High School X	150	147	144	144
Intermediate School Y	44	42	42	43
Elementary School Z	29	30	29	25

These changes affect the second and third quarters. Schedules A and D require a shift of one day from one quarter into the following quarter. To balance the number of days of attendance in the first and second quarters for Tract A, the attendance for 9/29/78 is moved from the first quarter and placed in the second quarter. Likewise, for Tract D, the attendance for 3/30/79 is moved from the third quarter to the fourth quarter. The adjusted school year consists of 177 days:

School Year by Quarter	
First Quarter	44
Second Quarter	46
Third Quarter	44
Fourth Quarter	43
Total	177

Appendix C reflects these changes.

The data in Appendix C depict the attendance for each of the four tracts. The number of students absent by tract are not identified. Each school day three tracts were in attendance: Example: Tracts A, B, D represent the three tracts attending school on July 11,

1978. Appendix C depicts the daily absentee rates by quarter. This rate includes the absentees for the three tracts in attendance. The second column in Appendix C identifies the tracts in attendance.

The school year consists of an adjusted 177 days per tract. An arc length of

$$360° / 177 = 2.0229°$$

represents each day of attendance. July 11, 1978 is assigned the angular measure of 0°. The last day of regular attendance is June 28, 1979 which is represented by an angular measure of 357.9664°. Column 3 of Appendix C provides the angular measure assigned to a time period or date. The daily absentee data for High School X, Intermediate School Y, and Elementary School Z by tract constitute the remaining data shown in Appendix C.

The data are analyzed using linear and circular methods. The circular analysis treats the data as a periodic measure of time. The period of time used in these comparisons is the academic year, July 11, 1978 to June 30, 1979, (177 days). The school system did close from July 1, 1978 to July 10, 1978.

Discussion of Statistical Tests

The data set is analyzed using the descriptive and inferential methods discussed in Chapter III. Emphasis is placed on those circular methods of analysis that are not used in the Carmel study.

The linear and circular statistical methods used in the present study are shown in Table 4.3.1. The descriptive tests listed are used in the Carmel illustration. The linear inferential methods include an analysis of variance using a completely random design (ANOVA-CRD) and an analysis of variance using a random block design (ANOVA-RBD). The circular inferential methods include tests for which computer programs are written and not used previously. They include the Watson-Williams two-sample and multi-sample tests. The present study includes a parametric and nonparametric test for circular-circular correlation where both variates are expressed as angular measures. The statistical tests demonstrate the different statistical interpretations associated with circular and linear analysis.

Table 4.3.1

Statistical Procedures Used in Unit District R Study

Classification	Linear	Circular
Central Tendency	Mean Median	Mean Angle Median Angle Mean Vector
Dispersion	Variance Standard Deviation Range	Angular Variance Mean Angular Deviation Angular Range
Inferential Methods		
a. One-Sample	Chi-squared	Rayleigh V-test
b. Two-Sample		Watson U^2 Watson-Williams
c. Multi-Sample	ANOVA - CRD ANOVA - RBD	Watson-Williams Chi-squared
d. Correlation	Pearson, r	Circular-Circular

Linear Measures of Central Tendency and Dispersion

Measures of Central Tendency

The measures of central tendency used are the arithmetic mean and the median. The EPISTST statistical program is used to calculate these values. The measures of central tendency for each school (High School X, Intermediate School Y, and Elementary School Z) are presented by quarter and by year in Tables 4.3.2, 4.3.3, and 4.3.4. Both the mean and median are based upon an adjusted 177-day school year.

The mean and median depicting the number of absences at Intermediate School Y are shown in Table 4.3.3. Both statistical measures are depicted by the quarter and by year. The designation of a tract refers to the absentee rate for all students scheduled when that tract is in session. The adjusted Tracts A, B, and D are modified to reflect the same number of attendance days (177) as Tract C.

Measures of Dispersion

Three linear measures of dispersion are used in the present study: the variance, standard deviation and range. The variance and standard deviation are calculated using the EPISTAT statistical procedures. The formulae used to calculate the variance and standard deviation are for sample data. The measures of dispersion for the three levels of instruction are shown in Tables 4.3.6, 4.3.7 and 4.3.8.

The three measures of dispersion are shown for the actual data as well as for the adjusted data. Since Tract C does not require an adjustment, it is not included in the adjusted absentee data. The ranges for the adjusted and actual data are identical. As a yearly measure of dispersion, the range is not the statistic of choice. Tracts A, B, and D experience a school closing of at least one day due to inclement weather. There are no reported absences on June 28, 1978, the last day of school. The range considers the difference between the largest and smallest values for each group. In the present study the range is not an appropriate measure of dispersion since the range encompasses all possible values (i.e., no students reported absent to all students reported absent).

Table 4.3.2

Measures of Central Tendency: High School X

	Actual Absentee Rates Quarter 1st	2nd	3rd	4th	Yearly
Group A					
Mean	136	146	188	118	150
Median	136	125	125	118	127
Group B					
Mean	144	135	189	124	147
Median	145	128	121	116	129
Group C					
Mean	142	151	153	128	144
Median	139	129	118	114	128
Group D					
Mean	136	156	163	122	144
Median	133	136	129	116	129

	Adjusted Absentee Rates Quarter 1st	2nd	3rd	4th	Yearly
Group A					
Mean	136	147	185	133	150
Median	136	126	131	118	128
Group B					
Mean	144	135	185	124	147
Median	145	129	125	116	130
Group D					
Mean	136	155	163	122	144
Median	133	136	134	116	130

Table 4.3.3

Measures of Central Tendency: Intermediate School Y

	Actual Absentee Rates Quarter				
	1st	2nd	3rd	4th	Yearly
Group A					
Mean	34	43	75	28	44
Median	33	32	38	27	32
Group B					
Mean	33	35	76	28	42
Median	32	32	41	27	31
Group C					
Mean	33	47	59	29	42
Median	32	36	42	29	33
Group D					
Mean	35	50	57	30	43
Median	34	39	42	29	37

	Adjusted Absentee Rates Quarter				
	1st	2nd	3rd	4th	Yearly
Group A					
Mean	34	43	72	28	44
Median	33	34	41	27	32
Group B					
Mean	33	36	72	28	42
Median	32	34	42	27	32
Group D					
Mean	35	50	56	30	43
Median	34	39	43	29	37

Table 4.3.4

Measures of Central Tendency: Elementary School Z

	Actual Absentee Rates Quarter				
	1st	2nd	3rd	4th	Yearly
Group A					
Mean	19	16	59	15	29
Median	19	22	24	15	19
Group B					
Mean	21	22	66	15	36
Median	20	20	22	15	19
Group C					
Mean	19	26	57	14	29
Median	18	23	22	14	19
Group D					
Mean	19	27	37	17	25
Median	18	23	26	16	20
	Adjusted Absentee Rates Quarter				
	1st	2nd	3rd	4th	Yearly
Group A					
Mean	19	26	56	15	29
Median	19	23	16	15	19
Group B					
Mean	21	22	63	15	30
Median	20	21	26	15	19
Group D					
Mean	18	27	37	17	25
Median	18	24	26	16	20

Table 4.3.5

Measures of Dispersion: High School X

	Actual Absentee Rates Quarter				
	1st	2nd	3rd	4th	Yearly
Group A					
Variance	1246.1	6052.8	47219.3	5550.3	14520.3
Standard Deviation	35.3	77.8	217.3	74.5	120.5
Range	201.0	404.0	1218.0	451.0	1301.0
Group B					
Variance	1274.5	1082.4	47306.3	3014.0	12791.6
Standard Deviation	35.7	32.9	217.5	54.9	113.1
Range	201.0	156.0	1213.0	379.0	1301.0
Group C					
Variance	506.3	5776.0	32220.3	5776.0	10941.2
Standard Deviation	22.5	76.0	179.5	76.0	104.6
Range	98.0	404.0	1218.0	451.0	1301.0
Group D					
Variance	1317.7	6130.9	14208.6	3271.8	6262.6
Standard Deviation	36.3	78.3	119.2	57.2	79.2
Range	201.0	404.0	654.0	423.0	737.0

	Adjusted Absentee Rates Quarter				
	1st	2nd	3rd	4th	Yearly
Group A					
Variance	1246.1	5791.2	42931.8	5550.3	14018.6
Standard Deviation	35.3	76.1	207.2	74.5	118.4
Group B					
Variance	1274.5	1043.3	43056.3	3014.0	12343.2
Standard Deviation	35.7	32.3	207.5	54.9	111.1
Group D					
Variance	1317.7	5867.6	13133.2	3226.2	6068.4
Standard Deviation	36.3	76.6	114.6	56.8	77.9

Table 4.3.6

Measures of Dispersion: Intermediate School Y

	Actual Absentee Rates Quarter				
	1st	2nd	3rd	4th	Yearly
Group A					
Variance	67.2	1376.4	12791.6	72.3	3636.1
Standard Deviation	8.2	37.1	113.1	8.5	60.3
Range	35.0	185.0	478.0	42.0	493.0
Group B					
Variance	64.0	169.0	12723.8	56.3	3340.8
Standard Deviation	8.0	13.0	112.8	7.5	57.8
Range	35.0	71.0	478.0	35.0	478.0
Group C					
Variance	53.3	1354.2	9273.7	74.0	2787.8
Standard Deviation	7.3	36.8	96.3	8.6	52.8
Range	33.0	185.0	478.0	50.0	493.0
Group D					
Variance	54.8	1317.7	3091.4	84.6	1218.0
Standard Deviation	7.4	36.3	55.6	9.2	34.9
Range	32.0	177.0	289.0	50.0	308.0

	Adjusted Absentee Rates Quarter				
	1st	2nd	3rd	4th	Yearly
Group A					
Variance	67.2	1310.4	11685.6	72.3	3504.6
Standard Deviation	8.2	36.2	108.1	8.5	59.2
Group B					
Variance	64.0	163.8	11642.4	56.3	3226.2
Standard Deviation	8.0	12.8	107.9	7.5	56.8
Group D					
Variance	54.8	1260.3	2883.7	86.5	1176.5
Standard Deviation	7.4	35.5	53.7	9.3	34.3

Table 4.3.7

Measures of Dispersion: Elementary School Z

	Actual Absentee Rates Quarter				
	1st	2nd	3rd	4th	Yearly
Group A					
Variance	34.8	169.0	12232.4	42.3	3158.4
Standard Deviation	5.9	13.0	110.6	6.5	56.2
Range	22.0	58.0	422.0	33.0	435.0
Group B					
Variance	32.5	36.0	12276.6	22.1	3249.0
Standard Deviation	5.7	6.0	110.8	4.7	57.0
Range	21.0	32.0	422.0	19.0	425.0
Group C					
Variance	36.0	166.4	11257.2	38.4	3091.4
Standard Deviation	6.0	12.9	106.1	6.2	55.6
Range	22.0	66.0	422.0	33.0	435.0
Group D					
Variance	33.6	163.8	894.0	31.4	331.2
Standard Deviation	5.8	12.8	29.9	5.6	18.2
Range	21.0	64.0	101.0	33.0	114.0

	Adjusted Absentee Rates Quarter				
	1st	2nd	3rd	4th	Yearly
Group A					
Variance	34.8	161.3	11172.5	42.3	3047.0
Standard Deviation	5.9	12.7	105.7	6.5	55.2
Group B					
Variance	31.4	38.4	11236.0	22.1	3136.0
Standard Deviation	5.6	6.2	106.0	4.7	56.0
Group D					
Variance	33.6	158.8	841.0	30.3	320.4
Standard Deviation	5.8	12.6	29.0	5.5	17.9

Linear Inferential Methods

Three linear inferential methods are used in the analysis (a chi-squared test, an analysis of variance, and the Pearson correlation coefficient). The chi-squared test determines if the daily absentee rate at Intermediate School Y is randomly distributed. An analysis of variance, ANOVA, using a completely random design (CRD) compares the yearly means of the four tracts for each school. An ANOVA using a random block design (RBD) compares a tract of Intermediate School Y by quarter. The Pearson's correlation coefficient considers if a relationship exists between the absentee rates of the three levels of instruction. The three inferential methods used in this analysis illustrate the power of linear statistics.

Chi-squared Test

A chi-squared test determines if the absentee rate for Intermediate School Y is randomly distributed. The data consists of the daily absentee rate for each day the school was in session (230 days). Each tract attended an adjusted 177-day school year. The statistical module, number 00041-90030, for the Hewlett Packard 41C calculator computes a chi-squared value of 14,652.95. The mean daily attendance, 42.83, is the expected probability. The tabled value of chi-squared at the .01 level of significance with 229 degrees of freedom is 281.71. The formula for calculating the tabled chi-squared is taken from Mininum (1978, p. 542). Significance is established at the .01 level and the null hypothesis of uniformity is rejected.

Analysis of Variance

An one-way analysis of variance using a completely random design is used to compare the means of the four tracks. Whenever a single tract is referred to, the absentee rate includes the three tracts in attendance. The EPISTAT statistical package is used to calculate the three ANOVAs. The variabilities of each are inspected to determine homogeneity of variance. The variances for each level are:

School	A	B	C	D
High School X	14,517.6	12,787.2	10,942.2	6,278.0
Intermediate School Y	3,633.3	3,341.9	2,782.7	1,215.5
Elementary School Z	3,156.3	3,245.4	3,089.0	332.0

Group D of Elementary School Z is noticeably different. The other three tracts are similar. The use of an ANOVA for the comparison is appropriate. Table 4.3.8 depicts the results of the three ANOVAs.

To facilitate presentation, the sum of squares, mean squares, and F values are rounded. Since the number of tracts and the total number of attendance days are equivalent for each school, the degrees of freedom are the same. Therefore, the tabled F values are identical. The tabled F value for an alpha level of .05 with three degrees of freedom in the "Between Group" and 686 degrees of freedom in the "Within Group" is 2.61 (Minium, 1978, Table H, p. 546). A comparison of calculated F values with the critical value indicates that at the .05 level of significance, the null hypothesis is not rejected.

Table 4.3.8

Analysis of Variance CRD

By Tracts by Year

School	Degrees of Freedom	Sum of Squares	Mean Squares	F
High School X				
Between Groups	3	4,801	1,600.3	.14
Within Groups	686	7,634,892	11,129.6	
Total	689	7,638,973		
Intermediate School Y				
Between Groups	3	601	200.3	.07
Within Groups	686	1,882,176	2,743.7	
Total	689	1,882,777		
Elementary School Z				
Between Groups	3	2,963	987.7	.40
Within Groups	686	1,688,385	2,461.2	
Total	689	1,691,248		

An ANOVA is used to compare each tract at Intermediate School Y by quarter. The EPISTAT program provides the variance to determine if the completely random design is appropriate. Following is a listing of the variances by tracts:

	Quarters			
Tract	First	Second	Third	Fourth
A	68.0	1,374.9	12,788.9	73.0
B	63.4	169.1	12,731.3	55.9
C	52.6	1,353.6	9,276.6	73.8
D	54.3	1,315.4	3,091.4	85.3

A review of the variability by tract and by quarter indicates that homogeneity of variance is not maintained. For example, variability in Tract A for the third quarter is 186 times greater than for the first quarter. An ANOVA using a random block design is employed to accommodate these differences in variability. The model requires that each quarter possess an equal number of days. This condition is satisfied by selecting the number of days per quarter based upon the quarter with the fewest attendance days. Unequal quarters are adjusted by dropping excess days. The Hewlett Packard statistical module is used to calculate the ANOVA. Table 4.3.9 summarizes these results.

Table 4.3.9

Analysis of Variance RBD

Intermediate School Y by Tract by Quarter

Tract	Degrees of Freedom	Sum of Squares	Mean Squares	F
A				
Between Groups	3	53,621	17,873.8	5.29
Within Blocks	39	139,250	3,570.5	1.06
Residuals	117	395,279	3,378.5	
Total	159	588,150		
B				
Between Groups	3	59,164	19,721.3	6.16
Within Blocks	39	131,292	3,366.5	1.05
Residual	117	374,512	3,201.0	
Total	159	564,968		
C				
Between Groups	3	23,921	7,973.7	2.87
Within Blocks	42	106,976	2,547.1	0.92
Residual	126	350,304	2,780.2	
Total	171	481,201		
D				
Between Groups	3	17,874	5,958.1	5.31
Within Blocks	40	43,479	1,087.0	0.97
Residual	120	134,711	1,122.6	
Total	163	196,064		

The critical F value for Tracts A and B with 3 and 117 degrees of freedom are 3.95 at the .01 level. Significance is established for Tracts A and B at the .01 level of significance. The tabled F value for Tract C with 3 and 126 degrees of freedom are 2.68 at the .05 level and 3.94 at the .01 level. Comparing the calculated F value of 2.87 with the tabled values establishes significance at the .05 level but not at the .01 level. The critical F value for Tract D with 3 and 120 degrees of freedom is 3.95 at the .01 level. The calculated F value of 5.31 is significant. The ANOVA indicates that the absentee rates for the same tract differ significantly at the .05 level. At the .01 level of significance, three of the four tracts (A, B, and D) indicate significance; C alone does not.

Pearson's Correlation Coefficient

Pearson's correlation coefficient establishes whether a relationship exists between absenteeism and the instructional level. The EPISTAT statistical package calculates the coefficient of correlation. The correlation considers the absentee rates for the 230-day school year for each level of instruction. Significance is established at the .01 level for all three levels. The three correlations considered are depicted in Table 4.3.10.

Table 4.3.10

Pearson's Correlation Coefficient

by Schools by Year

Schools	r	T	df	Significance
Intermediate School Y / High School X	.76	17.5	228	Yes .01 level
Intermediate School Y / Elementary School Z	.76	17.7	228	Yes .01 level
Elementary School Z / High School X	.52	9.2	228	Yes .01 level

Decisions Based on Linear Methods

The purpose of the present study is stated in Section 4.3. It may be expressed in terms of two questions. First, did each of the four tracts at Intermediate School Y receive equivalent instruction time? Second, does the level of instruction affect the absentee rate? This section addresses these questions using the results of the linear methods discussed previously.

Did each of the four tracts at Intermediate School Y receive equivalent instruction time? This question is examined from the following perspectives:

1. Examination of daily absentee rates
2. Description of data, daily absentee rates
3. Comparisons of attendance by tract and school.

The first procedure examines the daily absentee rates for the three schools. The raw data showing the absentee rates is shown in Appendix B. Appendix C depicts the daily attendance by tract and by school. Intermediate School Y is in session 230 days. Tracts A, B, and D are in attendance for 171 days. Six days of instruction are lost due to a holiday, a teacher institute day, and four days for emergency school closing due to inclement weather. Tract C is present for a period of 177 days. School closings are not experienced when Tract C was scheduled. The highest absentee rates are experienced during the period 1/15/79 to 1/25/79, while Tracts A, B, and D are in session. The data indicate that Tract C received 3.5% more instruction time than the other three tracts.

The second consideration is the statistical description of the data set. A data set is described by a measure for central tendency, a measure of dispersion, and a description of the distribution. The absentee rate for Intermediate School Y is discussed using the mean, standard deviation, and a chi-squared test for randomness.

Table 4.3.3 presents the means and medians for each tract at Intermediate School Y. The yearly means are similar [Tract A (43); Tract B (42); Tract C (42); Tract D (43)]. The mean quarterly rates for the third quarter depict the largest variability [Tract A (75); Tract B (76); Tract C (59); Tract D (57)]. This is expected due to the weather-related school closings during that time period.

The measures of dispersion are presented in Table 4.3.6 by quarter and by year for Intermediate School Y. The standard deviations for Tracts A (60.3); B (57.8) and C (52.8) are similar. Tract D (34.9) exhibits the least variability. Data in Appendix B indicate that on January 31, 1979 and February 1, 1979, the school is closed unofficially and all students are reported absent. Tracts A, B, and C are in attendance during this period. Thus, the difference in variability between Tract D and the other tracts is explained. A review of the quarterly values indicates that the variabilities observed in the third quarter are similar to the yearly figures. Standard deviations for the third quarter are Tracts A (113.1); B (112.8); C (96.3) and D (55.6). The majority of the variability associated with the absentee rate is attributed to the third quarter.

This is expected since the absentee rates for the period from 1/15/79 to 1/25/79 were at their highest levels.

A random distribution is selected as the distribution model to compare daily absenteeism. This model assumes that the number of absentees is the same for each school day. A chi-squared test determines uniformity. The mean absentee rate for the school (42.83) is used as the expected value. A chi-squared value of 14,652.95 is obtained. The value is significant at the .01 level. The assumption that the absentee rate is the same for each school day is rejected. This result is expected. The inclement weather results in more absences due to weather-related illnesses such as colds and flu. Inclement weather makes the transportation of students to and from school more difficult. Other considerations such as holidays and vacations affect daily attendance.

The third consideration is the comparison of attendance at Intermediate School Y by tracts. The comparison uses an one-way analysis of variance employing a random block design. Table 4.3.9 depicts the results of the ANOVA. The means of each tract differ significantly by quarter at the .05 level as discussed previously. Significance at the .01 level is established for Tracts A, B, and D. These findings affirm that the absentee rates by quarter are significantly different.

The question, "Does the level of instruction affect the absentee rate?" is considered. To answer this question a comparison of attendance at High School X, Intermediate School Y and Elementary

School Z is presented. Table 4.3.11 shows the correlation between the three schools. The three correlations are significant at the .01 level. The correlation between High School X and Intermediate School Y is the same as the correlation between Intermediate School Y and Elementary School Z (.76). This implies that 57.8% of the change in absenteeism at one school is associated with the change in absenteeism at the other school. The correlation between the elementary and high school is .52, indicating that only 27.0% of the change in absenteeism at one school is associated with the absentee rate at the other school. These results support the assumption that the instructional level affects the absentee rate.

The next comparison demonstrates whether the yearly mean absentee rates differ for the four tracts at each school. Table 4.3.8 depicts three one-way ANOVAs to compare the tracts by school. The null hypothesis that the means of each tract differ from the other three tracts is not rejected at the high school, intermediate or elementary levels. This implies that the absentee rates for the four tracts at each school are not significantly different. The same conclusion is reached by a review of the quarterly and yearly means. The results of the three ANOVAs provide statistical confirmation.

Conclusions and observations based upon the statistical results are presented. The first observation is based solely upon the number of days of instruction. Three of the four tracts, A, B, D, in attendance at the Intermediate School Y attended school 171 days. Tract C received 6 additional days of instruction or 3.5% more classroom time.

The discrepancy is primarily attributed to the closing of school for four days in January because of inclement weather. Tracts A, B, and C received two fewer days of instruction due to the unofficial school closings on January 31st and February 1st when all students are reported absent. The observation is verified by the ANOVA study comparing each tract by quarters. As observed, the means are significantly different. Linear procedures determine the quarter in which the absentee rates are statistically different. The time period when absences occur are not addressed by the statistical procedures. The chi-squared test confirms that the data are not uniformly distributed. The description of the distribution is not addressed in the present study.

The second observation concerns the tracking of students at Intermediate School Y. An ANOVA compares the means of the four tracts at Intermediate School Y. Significance is not established indicating that the absentee rates are not affected by the tracking process. The issue is answered clearly using linear methods. Pearson's correlation coefficient establishes that the absentee rates are influenced by the instructional level. The correlation between High School X and Intermediate School Y is the same at the correlation between Intermediate School Y and Elementary School Z ($r = .76$). However, the correlation between High School X and Elementary School Z is .52.

Yearly schedules offer unique educational opportunities for change. Proper scheduling can reduce energy costs, reduce absenteeism, and use existing facilities more efficiently. Space limitations necessitated implementation of a 12-month school schedule for Unit District R. However, if a facilities exist and education becomes a year-round activity each student would receive 33.3% more instruction time. The schools would become more cost efficient if operated over a 12-month period. During the period July 1, 1978 to June 30, 1979, Unit District R made maximum use of its instructional facilities. If the time period representing the average daily absentee rates could be identified, a schedule could be constructed to minimize the variability associated with each quarter. Circular statistics affords such a statistical tool. School closings for short periods during the hottest and coldest periods of the year would be cost efficient and would allow for maintenance of the buildings. The linear method affirms that the tracking system does work, when absentee rates are considered. The present study also indicates that one schedule may not be applicable for each level of instruction. These are several of the many factors that must be considered.

Circular Measures of Central Tendency and Dispersion

Measures of Central Tendency

The circular statistics comprising the measures of central tendency are the coordinates of the mean vector, the magnitude of the mean vector, and the mean angle. The measurements for Tracts A, B, and D are based upon on adjusted data reflecting a 177 day school year. The arithmetic mean is used to estimate daily absentee rates when classes are not held. The yearly attendance rates for the school are based upon a 230-day schedule. These data do not require adjustment and reflect actual daily attendance.

The absentee rates are considered as grouped data. As such, a correction of the mean vector, r, is necessary when the number of groups is small, 12. Batschelet (1981, p. 38) states, "If the number of groups exceeds 12, the correction has minimal effect and can often be dropped." The grouping process assigns the number of absences per day to a single group. When individual tracts are considered, 177 groups are used. The absentee rates for the school are based upon 230 groups. No correction is required with groups of this size. An example of time as a periodic measure is found in Mardia (1972, p. 11). In the present study the time period used is the school attendance year (230 days). The data are angular distributions of time expressed as periodic measures. The data are distributed about the circle. Each day is represented by an arc length of equal measure. The angular measures representing the individual school days are spaced equally about the circle. The number of students absent on a given day constitutes the weighting of the unit vector. The statistical objective determines the time period which represents the "average" daily attendance.

Tables 4.3.11, 4.3.12, 4.3.13 show the coordinates of the mean vector (m), the magnitude or the length of the mean vector (r), and the mean angle (Ø), for each level of instruction by tract and by school for the school year. The vectors for the elementary and intermediate levels lie in the third quadrant. At the high school level, the mean vectors for Tracts C, D, and combined lie in the second quadrant, adjacent to the third quadrant. Of the three levels, the mean vectors associated with Elementary School Z exhibit the largest magnitude. The closer the magnitude or length of the mean vector is to 1, the nearer the data points are to each other. The magnitudes of the mean vectors at the high school level are extremely small, indicating greater variability. The three mean angles are within 22° of each other.

Table 4.3.11

Coordinates of the Mean Vector

by Tract and by School

Tracts	Elementary School Z		Intermediate School Y		High School Z	
	x	y	x	y	x	y
A	-0.3202	-0.1498	-0.2487	-0.0785	-0.0924	-0.0257
B	-0.2760	-0.1940	-0.2204	-0.1045	-0.0847	-0.0128
C	-0.3523	-0.0625	-0.2337	-0.0050	-0.0588	0.0176
D	-0.2002	-0.0747	-0.2016	-0.0076	-0.0896	0.0170
Combined	-0.3063	-0.1201	-0.2363	-0.0479	-0.0848	0.0003

Table 4.3.12

Mean Vectors by Tract and by School

Tracts	Elementary School Z r	Intermediate School Y r	High School X r
A	0.3535	0.2608	0.0959
B	0.3374	0.2440	0.0857
C	0.3578	0.2337	0.0614
D	0.2137	0.2018	0.0912
Combined	0.3291	0.2411	0.0848

Table 4.3.13

Mean Angles by Tract and by School

Tracts	Elementary School Z Ø	Intermediate School Y Ø	High School X Ø
A	205.0802°	197.5173°	195.5327°
B	215.0991°	205.3613°	188.5972°
C	190.0681°	181.2318°	163.3511°
D	200.4509°	182.1518°	169.2646°
Combined	201.4134°	191.4494°	179.7970°

Measures of Dispersion

The measures of dispersion considered are angular variance, angular deviation, and the range. The range is not a viable measure of dispersion for this data set. By definition it is the smallest arc which contains all of the data points. Since the points are distributed about the continuum, the range is not applicable.

The angular variance and angular deviation are obtained by the computer programs discussed previously. Since the descriptive use of the angular variance is limited, it is expressed only in radian measure. The variability increases from elementary to high school at each level and for each tract. Tables 4.3.14 and 4.3.15 are shown below.

Table 4.3.14

Angular Variance by Tract and by School

Tracts	Elementary School Z radians	Intermediate School Y radians	High School X radians
A	1.2930	1.4784	1.8082
B	1.3252	1.5120	1.8286
C	1.2844	1.5326	1.8772
D	1.5726	1.5964	1.8176
Combined	1.3418	1.5178	1.8304

Table 4.3.15

Angular Deviation by Tract and by School

	Elementary School Z		Intermediate School Y		High School Z	
Tracts	degrees	radians	degrees	radians	degrees	radians
A	65.1511	1.1371	69.6656	1.2159	77.0452	1.3447
B	65.9574	1.1512	70.4528	1.2296	77.4786	1.3523
C	64.9341	1.1333	70.9312	1.2380	78.5015	1.3701
D	71.8508	1.2540	72.3925	1.2635	77.2453	1.3482
Combined	66.3692	1.1584	70.5878	1.2320	77.5168	1.3529

Test for Randomness

A description of a data set must include a discussion of its distribution. It is assumed that the rate of absenteeism is the same for each attendance day. The Rayleigh test is used to verify that the distribution is uniformly distributed. The null hypothesis for the Rayleigh test is stated: "The parent population is uniformly distributed." The test statistic is

$$z = nr^2$$

The z statistic is used when values of n are greater than 30. The test statistic was calculated using MenuB.

The Rayleigh test is used to verify uniformity at each level of instruction. The results of the Rayleigh test are shown in Table 4.3.16. At Elementary School Z and Intermediate School Y, the null hypothesis of uniformity is rejected for each of the four tracts and for the combined group. However, for High School X, the null hypothesis of uniformity is not rejected. The mean vectors associated with High School X are small. The Rayleigh test uses the mean vector as a measure of dispersion. Consequently, randomness can not be discounted as a distribution at the high school level.

Table 4.3.16

Rayleigh Test by Tract and by School

	Days Present	Mean Vector	Test Statistic	Critical Value	Significance level .01
High School X					
A	177	.0959	1.6278	4.5850	No
B	177	.0857	1.3000	4.5850	No
C	177	.0614	0.6673	4.5850	No
D	177	.0912	1.4722	4.5808	No
Combined	230	.0848	1.6539	4.5900	No
Intermediate School Y					
A	177	.2608	12.0389	4.5850	Yes
B	177	.2440	19.5379	4.5850	Yes
C	177	.2337	9.6670	4.5850	Yes
D	177	.2018	7.2080	4.5808	Yes
Combined	230	.2411	13.3697	4.5900	Yes
Elementary School Z					
A	177	.3535	22.1183	4.5850	Yes
B	177	.3374	20.1495	4.5850	Yes
C	177	.3578	22.6597	4.5850	Yes
D	177	.2137	8.0832	4.5808	Yes
Combined	230	.3291	24.9106	4.5900	Yes

Circular Inferential Methods

One-Sample Tests

Three one-sample tests (the V-test, the Watson U^2, and the Rayleigh test) are discussed in this section. Each of these tests is used to determine uniformity.

The V-test determines if observed angles have a tendency to cluster about a given or hypothetical angle. The V-test determines if the daily absentee rates for the three levels cluster about each other. Each mean angle depicting a level of instruction is used as a hypothetical angle. The V-test compares the hypothetical angle with the distribution of the other two levels of instruction.

An example shows how these comparisons are made. The mean angle representing the yearly absentee rate at Elementary School Z is used as the given or hypothetical angle. The mean angle of Intermediate School Y is compared to the given angle by the V-test. The test determines whether the mean angle of Intermediate School Y differs significantly from the given angle. When the null hypothesis, "The population is uniformly distributed," is rejected, significance is established. Six comparisons are made. The results are summarized in Table 4.3.17. The mean angle for each of the three schools is used as the hypothetical direction against which the other two levels of instruction are contrasted. The test statistic is

$$u = (2n)^{\frac{1}{2}} v.$$

The null hypothesis is evaluated at both the .05 and .01 levels of significance. The critical value at the .05 level is 1.6452 and at the .01 level is 2.3228 (Batschelet, 1981, Table I, p. 336). The null hypothesis is rejected for each comparison at the .05 level of significance. At the .01 level, significance is established at the elementary and intermediate levels, but not at the high school level.

Table 4.3.17

Summary of V-test Results with the

Given Angle as the Hypothetical Direction

Given Angle	School[a]	n	r	Mean Angle	u	Rejected .05	.01
H.S.X							
179.7970°	E.S.Z	230	.3291	201.4134°	6.5620	Yes	Yes
179.7970°	I.S.Y	230	.2411	191.4494°	5.0645	Yes	Yes
I.S.Y							
191.4494°	E.S.Z	230	.3291	201.4134°	6.9519	Yes	Yes
191.4494°	H.S.X	230	.0848	179.7970°	1.7813	Yes	No
E.S.Z							
201.4134°	I.S.Y.	230	.2411	191.4494°	5.0930	Yes	Yes
201.4134°	H.S.X.	230	.0840	179.7970°	1.6908	Yes	No

[a] H.S.X = High School X; I.S.Y = Intermediate School Y;
E.S.Z = Elementary School Z

The Watson U^2 test is used as a goodness-of-fit test or as a test of uniformity. It is not appropriate for large samples of grouped data (Batschelet, 1981, p. 115). To demonstrate the computer program, a data set is selected containing those dates when the absentee rate at Intermediate School Y is less than 20 students. Table 4.3.18 represents this data set. The number of absences for the nine-day period totals 157. Watson's U^2 statistic is calculated at 2.5742. The critical value with degrees of freedom (∞ , ∞) is .267. The null hypothesis of uniformity is rejected. Figure 4.3.1 shows the data clustered at the beginning of the fourth quarter. The mean vector for this data set is r = .4998 and the mean angle is 271.7459°.

Table 4.3.18

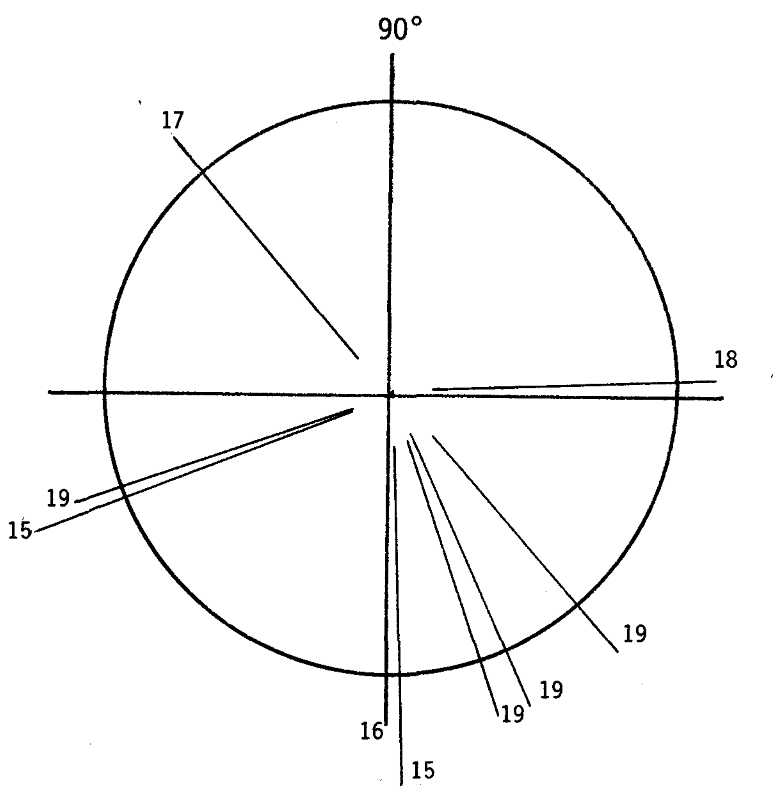
Absentee Rates Less than Twenty

at Intermediate School Y

Date	Measure	Frequency
07/13/78	3.1304°	18
11/08/78	131.4783°	17
01/23/79	198.7826°	19
01/29/79	201.9130°	15
04/02/79	270.7826°	16
04/03/79	272.3478°	15
04/26/79	289.2609°	19
05/01/79	294.2609°	19
05/16/79	311.4783°	19

Figure 4.3.1

Circular Distribution of the Data Set



The Rayleigh test is discussed previously under descriptive methods.

Two-Sample Test

The Watson U^2 two-sample test is not applicable for use with large data sets. This application is presented to demonstrate the use of the computer programs. The Watson U^2 test compares the absentee rates for Tracts A and C. The data set consists of the angular measures associated with absentee rates of less than 20 for Tracts A and C. The angular measures represent attendance as an ordinal number, i.e., the number of absences on the first day of school, on the second day of school, etc. Tract A is present each time the attendance rate is less than 20. Tract C is used in the comparison since it does not require a scheduling adjustment. The angular measures describing the period when the absences occur differ for each tract. Tract A associates 17 absences with an angular measure of 146.4408°, whereas the same 17 absences occur at 113.8984° for Tract C. This implies that Tract A attends more days of school than Tract C when this occurrence takes place. Table 4.3.19 depicts the angular measure and the number of absences associated with that measure for Tracts A and C.

The Watson U^2 test requires that ties be broken. Two ties occur, one at 274.5764° and the other at 296.9495°. A coin toss is used to break the ties. Both times, Tract C's frequencies are used. The test statistic used is Watson's U^2. A calculated U^2 value of .7314 is obtained. The critical U^2 value at the .01 level of significance with df (∞, ∞) is .268 (Mardia, 1972, p. 314). The null hypothesis of uniformity is rejected, implying that the time period when low attendance is encountered for Tract A is significantly different than for Tract C.

Discussion of Circular Methods

The use of circular statistics to study the time factors influencing absentee rates is stated in the section entitled, "Hypothesis 2a and 2b," on page 31. Circular methods are used to determine the effect of time upon daily absentee rates. The present analysis considers two questions based upon the results of these circular methods. The first question: "Did each tract at Intermediate School Y experience similar absentee rates during the same time periods?" The second question: "Does the level of instruction affect the time period when absenteeism occurs?" The following discussion considers the circular data from three perspectives.

1. Examination of angular measures of daily absentee rates.
2. Description of data.
3. Inferential comparisons of absentee rates by tract and by school.

The present circular study uses periodic data, i.e., an angular measure representing a specific time period. The circular data for the three levels of instruction are shown in Appendix C. Unit District R is selected for the present study for two reasons. First, Unit District R operates on a twelve-month schedule. Second, the number of absences provides a large data set on which to apply circular methods.

The circle divides the academic year into 4 quadrants:

0° to 90° representing the first quarter;

90° to 180° representing the second quarter;

180° to 270° representing the third quarter;

270° to 360° representing the fourth quarter.

If 5 or more attendance days appear in a specific month, the month is assigned to that attendance quarter. The angular measures (193° to 206°) represent the time period when the absentee rates are greatest at Elementary School Z. Intermediate School Y experiences its largest absentee rates between 161° and 207° and at High School X the largest number of absences occur between 161° and 205°. At each level of instruction the maximum number of absentees occur at the end of the second academic quarter (December) or at the beginning of the third academic quarter (January). The angular dispersion is greater at the intermediate and high school levels than at the elementary level, supporting the hypothesis that the attendance rate at the elementary level is more stable than at the intermediate and high school levels. A rationale can be made that at the lower level the school fills a two-fold function: education and child care.

The circular data are described using measures of central tendency, measures of dispersion and a description of the distribution. The measures of central tendency include the coordinates of the mean vector (m), the length or magnitude of the mean vector (r), and mean angle (Ø). Tables 4.3.11 through 4.3.16 show these results. The coordinates of the mean vectors (m) for Elementary School Z and Intermediate School Y indicate that the time period representing the average absentee rate occurs in the third quarter. At the high school level, the mean vectors lie in the third quadrant for Tracts A and B and in the second quadrant for Tracts C and D. The vector representing the combined absentee rate for High School X lies in the second quadrant. A vector has two components, a magnitude or distance and a direction. In circular statistics, the magnitude is represented as r, and the direction as the mean angle. The magnitudes by tract and for the combined groups are shown in Table 4.3.12. The table illustrates the difference in magnitudes associated with the instructional level implying that the instructional level affects the absentee rate. The length of the mean vector varies from 0 to 1. At 1, all absences occur on the same day. At 0, the absences are evenly distributed. The magnitudes decrease as the instructional level increases: elementary (.3291), the intermediate (.2411) and the high school level (.0848). The mean angles representing the angular direction of the mean vectors are 201° at the elementary level, 191° at the intermediate level and 180° at the high school level. These results suggest that weather conditions affect attendance at the lower levels more than at the upper levels. The lengths of the mean vectors indicate that the average absentee rate for the elementary, intermediate and high school differ. The magnitudes at the elementary and intermediate levels are sufficiently large, suggesting that the absentee rate is directional. However, the magnitude at the high school level (.0848) is very small, implying a more uniform distribution of absenteeism. The magnitudes and mean vectors suggest that students at the elementary level miss less school than their high school counterparts.

The dispersion of data about the mean angle is depicted in Tables 4.3.16 and 4.3.17. The angular deviation is the descriptive statistic of choice. Angular deviation is defined as a function of r:

$$s = [2\ (1-r)]^2\ .$$

Therefore, as r increases, the angular deviation decreases. Table 4.3.15 indicates that the level of instruction with the least amount of variability (66°) is the elementary level with the high school level exhibiting the most variability (77°). These results imply that as the level of instruction increases, the angular distribution becomes less directed. The maximum value that angular deviation can attain is 180°. The three angular deviations: High School X (78°), Intermediate School Y (71°), and Elementary School Z (66°) represent angular distributions which do not cluster about the mean angle. The differences in angular deviations denote that variability increases as the instructional level increases.

The comparison of the mean angular deviations for the four tracts at Intermediate School Y indicate a difference of less than 3°. This is expected since the magnitudes of the mean vectors deviate by less than .06. The 3° difference between the maximum and minimum angular deviations indicates that the absentee rates for the four tracts are similarly dispersed.

A discussion of the angular distributions is considered now. It is postulated that absences are uniformly distributed about the circle. The Rayleigh test is used to verify the assumption of uniformity. Table 4.3.16 depicts the test results. At Elementary School Z and Intermediate School Y, the null hypothesis of uniformity is rejected for each of the four tracts and for the combined enrollment. The Rayleigh test confirms that the time periods when absences occur are directed. However, at High School X, the null hypothesis of uniformity is not rejected. This result is interpreted as indicating that the circular data representing absentee rates at the high school level are not directed. These results are expected since the Rayleigh test uses the length of the mean vector as its test statistic.

The V-test is used to determine if the mean angles of the three instructional levels differ from one another. Table 4.3.17 shows the results of the three comparisons. Each mean angle for each level of instruction is used as the hypothetical direction and the other two levels are compared to it. In each of the six comparisons, significance is established at the .05 level, indicating that the mean angles do differ significantly. The result confirms the comment that the time period when absences occur is affected by the instructional level. At the .01 level of significance, the null hypothesis is not rejected at the high school level. This is expected since the test statistic used in the V-test is a function of r. The V-test generally uses a pre-determined hypothetical direction. In the present study the V-test uses a mean angle as the hypothetical angle. However, appropriate two-sample tests for samples of this size are lacking.

A two-sample or multi-sample test is needed to compare the circular distributions. A review of two-sample and multi-sample tests as presented by Batschelet (1981) and Mardia (1972) demonstrate a need for statistical tests employing large data sets. The Watson U^2 test is a small-sample test and thus not appropriate. The Watson-Williams two-sample test requires a von Mises distribution with a concentration parameter of k greater than 2. The circular distributions at Elementary School Z and Intermediate School Y possibly exhibit a von Mises distribution. However, the concentration parameter associated with magnitudes of .3291 and .2411 are far below the required value of 2. The Watson-Williams test, therefore, is not applicable. The same statement can be made for the multi-sample version of the Watson-Williams test. The multi-sample form of the circular chi-squared test is appropriate. However, the method of reporting absences used by Unit District R does not identify the number of absences by tract. A circular chi-squared test could be used to compare attendance by tract and by day of instruction provided the absentee rate for each tract for each attendance day is available.

A circular-circular correlation is used to establish if a relationship exists between the three levels of instruction. A parametric test statistic (r^2) is used. Three circular-circular correlations are calculated (Elementary School Z and Intermediate School Y; Elementary School Z and High School X; Intermediate School Y and High School X). The three correlations are significant with a probability of less than .001 level, establishing an angular relationship between the three levels of instruction. The correlations suggest that a similarity in the time periods when absences occur exists between the three levels of instruction. At first this appears to contradict the results of the V-test which imply that the mean angles differ significantly from each other. The circular-circular correlation is a test for independence. Consequently, the circular data sets may have a mean angle which differs significantly and a distribution which exhibits a dependent relationship. Two circular measures are discussed: direction (mean angle) and magnitude (measure of dispersion). The mean angles of the three instructional levels differ significantly and are based upon the direction of the mean vector. The dependent relationship established is based upon the length of three mean vectors.

It is demonstrated that the circular distributions for the four tracts at Intermediate School Y are directed. Weather conditions affect Tracts A and B to a greater extent than Tracts C and D. Observation of the data sets indicates that Tracts A, B, and D attended school 6 days less than Tract C. Because Unit Distrit R does not identify absentees by tract, the multi-sample chi-squared test can not be used. The Watson-Williams test is not appropriate due to the unique characteristics of the sample, consisting of a large sample size and small concentration parameter. Circular analysis does not answer the question: "Did the four tracts receive the same instruction time?" It does raise an interesting question: "Why do Tracts A and B experience absentee rates at a time different from Tracts C and D?" This may indicate that students in Tracts A and B are absent more often. Further analysis requires absentee rates reported by tract.

In summary, the circular methods compare the time periods when absenteeism occurs at the three levels of instruction. The results appear paradoxical. The V-test is used to determine that the mean angles differ significantly at the .05 level. The results of the V-test imply that the time period when the "average" absentee rate occurs differ. The high school "average" is experienced first, then the intermediate level, and finally, the elementary level. These results imply that weather conditions affect the elementary level more than the high school level. Three circular-circular correlations are used to compare the three instructional levels. Each correlation is significant at the .001 level. The null hypothesis of independence is rejected. The paradox develops between the results of the V-test which indicates a significant difference between the mean angles of the three levels of instruction and the correlations which show a dependent relationship between the three levels. The paradox is resolved since one measure reflects angular measurement and the other measure reflects the magnitude of the mean vector. More students missed school in the third quarter at each level of instruction.

However, a difference exists regarding the influence of absentee rate during the other three quarters. At the high school level, absentee rates were higher throughout the year, hence, its mean angle was less than at the elementary level which has a lower absentee rate throughout the other three quarters. Circular analysis is a statistical approach which clarifies and explains the problem.

Discussion of Linear and Circular Decisions

One of the major differences between linear and circular analysis is the description of the data set. The linear data used in the present study consist of the daily absentee rates at the High School X, Intermediate School Y, and Elementary School Z. The circular data consist of an angular (periodic) measure representing a school day. Each absent student is assigned the angular measure representing the day absent. The application of linear and circular methods provides different perspectives in the problem-solving process.

Both linear and circular analysis considers the two questions: "Did each tract at Intermediate School Y receive equivalent instruction time? and "Did the level of instruction affect the absentee rate?" The data is discussed from three perspectives:

1. Examination of data set.
2. Descriptive methods.
3. Inferential methods.

Did each tract receive equivalent instruction time is considered first. The examination of data considers the number of days each tract attended school and the absentee rates for each of the four tracts. An examination of both the linear and circular data reveals that Tracts A, B, and D receive 6 fewer days of instruction than Tract D. The linear data identifies the six attendance days that are missed by Tracts A, B, and D as 10/09/78, 10/20/78, 01/15/79, 01/16/79, 01/24/79, and 01/25/79. From the period, January 15 through January 25, the school is closed four times. An examination of the attendance records reveals that these closings are weather-related. The circular data are based upon a 177 school day year. Each data point is treated as an ordinal number. The angular measurement representing the day that the school is closed may differ for each tract. Tract A experiences school closing at 101.6590°, 120.0001°, 187.1188°, 189,1527°, 207.4578°, and 209.4917°. The school is closed for Tract B at 99.6613°, 117.9664°, 193.2205°, 195.2544°, 207.4578°, and 209.4917°. Tract C is not scheduled when school closings occur. Tract D experiences school closings at 99.6611°, 117.9662°, 193.2205°, 195.2544°, 207.4578°, and 209.4917°. A review of the data indicates that Tract A had one more day of attendance than the other two tracts when the first school closing occurs. By the time that the January closings occur, Tracts A, B, and C have attended the same number of days. The circular distribution divides the school year into four quadrants. Each quadrant represents an academic quarter. Two closings occur at the beginning of the second quadrant, between 90° and 120°. The other four closings are at the beginning of the third quadrant (180° to 210°). Both methods provide information unique to the description of the data set used.

The description of a data set requires a measure of central tendency, a measure of dispersion and a discussion of the distribution. The primary linear and circular descriptive measures are compared.

Table 4.3.3 shows the linear measures of central tendency used in the present study. The yearly arithmetic mean for the four tracts differ by 2 students [Tract A (44), Tract B (42), Tract C (42), Tract D (43)]. The adjusted figure uses the mean to estimate attendance for the 6 days when the school is not in session. The adjusted means are identical to the actual means. A review of the quarterly means by tract indicates an increase in absenteeism in the third quarter.

The circular measure of central tendency used to describe the data set is the mean vector. The mean vector is defined in terms of length or magnitude and direction or mean angle. The magnitudes and mean angles representing the four tracts at Intermediate School Y are shown in Tables 4.3.12 and 4.3.13. The magnitudes for the four tracts differ by .06. Tract D accounts for 50 percent of the difference associated with the magnitudes. An inspection of the data indicates that the school is closed on 1/31/79 and 2/1/79 and the entire student body is reported absent. The tracts in attendance for this period are Tracts A, B, and C. Since Tract D is not in attendance at this period, its magnitude is smaller.

The mean angles indicate the time period when the "average" absentee rate occurs. The mean angles for the four tracts range from

81.1518° to 205.3613°, a span of 24.2095°. This span represents a large value since the absentee rate for a specific tract includes the absentee rates of the other two tracts in attendance.

Circular descriptive methods identify the mean angle or time period when the "average" attendance occurs. The magnitude provides a measure of dispersion about the mean angle. Due to the nature of the data set, linear and circular analysis provides different descriptive information.

The two measures of dispersion discussed are standard deviation (linear method) and angular deviation (circular method). Table 4.3.7 shows the adjusted standard deviations [Tract A (59.2), Tract B (56.8), Tract C (52.8), Tract D (34.3)]. Tract C does not require an adjustment. The standard deviation for Tract D is far less (1.4613) than the other three tracts. An inspection of the data indicates that Intermediate School Y reports the entire student population absent on 1/31/79 and 2/1/79. Tract D is not in attendance during that period, consequently, reducing the standard deviation.

The angular deviation describes the angular variability about the mean angle. Table 4.3.15 shows the measures of dispersions are 69.6656° for Tract A, 70.4528° for Tract B, 70.9312° for Tract C, and 72.3925° for Tract D. The difference in the angular deviation for the four tracts is 2.7269°. The difference in the angular deviation from Tract D and the other three tracts is 1.4613. This value represents 54 percent of the difference associated with the four angular deviations. The reason Tract D's angular deviation differs from the other three tracts is related to the two "unofficial" school closings on 1/31/79 and 2/1/79.

Two tests for randomness are used: the chi-squared test (linear method) and the Rayleigh test (circular method). The chi-squared test is based upon the attendance for the school year (230 days). It is significant at the .01 level. Consequently, the null hypothesis of uniformity is rejected. The Rayleigh test is used to determine randomness for each of the four tracts and for the school year (combined group). Significance at the .01 level is established for each tract and for the combined group. The null hypothesis of uniformity is rejected. Both linear and circular methods established that the data are not uniformly distributed.

Finally, linear and circular methods are used to determine if the data sets differ by tract at Intermediate School Y. The linear method used is an analysis of variance. Table 4.3.8 depicts the results of the analysis of variance for each level of instruction by tract. The F value for Intermediate School Y is .07. The null hypothesis stating that the means are equal is not rejected at the .05 level of significance.

Three circular methods are considered: the Watson-Williams two-sample and multi-sample, Watson's U^2 test and the chi-squared test. The Watson-Williams test is appropriate only for a von Mises distribution with a concentration parameter greater than 2. The data representing the four tracts do not meet this requirement. The Watson U^2 test is not suitable for a data set of this size. The chi-squared test is not used since the absentee rates reported by the district do not identify the number of absences by tract. If the absences are reported by tract, a chi-squared distribution would be applicable.

A review of the 9 remaining two-sample and multi-sample tests considered by Batschelet (1981) does not reveal a circular procedure appropriate for comparing the four tracts at Intermediate School Y. Mardia (1972, p. 153) acknowledges the lack of a circular test statistic comparable to Student's t-test in linear statistics. The lack of circular procedures to compare large sample sets restricts the use of circular analysis.

In summary, the results of linear and circular methods imply that the absentee rates for the 4 tracts at Intermediate School Y are similar. Linear analysis uses an ANOVA to compare the absentee rates of the four tracts. Both data sets indicate that Tracts A, B, and D receive six fewer days of instruction. The linear and circular measures of variability show that Tracts A, B, and C are in session when the school is closed unofficially and all students reported absent. The mean angles indicate that the angular distributions of Tracts A and B are similar, and Tracts C and D are similar.

The linear analysis compares the levels of instruction using Pearson's correlation coefficient. The circular analysis contrasts the three levels using the V-test and circular-circular correlation. The linear and circular data sets for High School X, Intermediate School Y, and Elementary School Z are compared. Each level attends school 230 days. Each school is closed officially on 6 days, 4 of which are weather-related. On January 31, 1979, all three levels are closed and the student body reported absent. On February 1, 1979, Intermediate School Y and Elementary School Z are closed and all students reported absent. Elementary School Z remains closed on February 2 with all students reported absent. Since the District had used its allotted number of snow days, the closings necessitated the recording of all students as absent.

The linear method used to establish the relationships between the absentee rates at the three schools is Pearson's correlation coefficient. Table 4.3.10 shows the three comparisons (Elementary School Z and Intermediate School Y, Elementary School Z and High School X, Intermediate School Y and High School Z). Each comparison is significant at the .01 level. The correlation is highest (.76) between Intermediate School Y and the other two levels. The correlation between Elementary School Z and High School X is .52. The correlation of .76 implies that 58% of the variability between the two sets is explained. A correlation of .52 implies that only 27% of the variability is shared. These findings suggest that the instructional level affects the attendance rates.

Two circular methods (the V-test and the circular-circular correlation) are used to determine if the instructional level affects the attendance rates. The V-test determines if the mean angles used as hypothetical angles differ significantly from the mean angles representing the other two instructional levels. Table 4.3.17 shows the results of the V-test. At the .05 level of significance, all three mean angles differ significantly. This implies that the time period when the "average" absentee rate occurs is different for each instructional level.

The three circular-circular correlations are significant at the .001 level, rejecting the null hypothesis of independence. This implies that a relationship exists between the levels of instruction. Although a circular relationship exists, based upon the correlational methods the variability of the distributions is sufficient to account for the difference in the mean angles observed in the V-test. An apparent paradox exists due to the nature of circular analysis. The V-test is based upon direction, whereas correlation is based upon magnitude. The paradox is thus resolved. These results imply that instructional level affects absentee rate. However, the effect may differ for each level of instruction.

Results of the present study show that linear and circular methods can be used to study a problem from two different perspectives. Linear analysis provides a mathematical perspective with which education is familiar. Circular methods provide an angluar view of the problem. The present study presents an educational concern, absenteeism, which demonstrates that a set of data may be analyzed using two statistical methodologies. Each statistical method has unique strengths. The circular tests used in the present study add to the statistical understanding of when absences occur, an important point when considering the amount of instruction time a student receives. Circular analysis has limitations. The lack of a circular method analogous to Student's t-test or an ANOVA is a serious limitation. Many questions in education are based on time. At what time during the instruction period is the student most receptive? What time period of the school represents the "average" attendance rate? How long can a child read independently? These questions illustrate several practical educational problems that can benefit from circular analysis. The present study suggests that the application of circular methods contributes to a more complete understanding of the problem.

4.4 Summary

The findings reported in Chapter IV indicate that circular methods of analysis have a place in education. Two data sets are considered. The first data set is used to analyze the feasibility of a tuition increase at a private high school. The second data set is used to determine the effect of tracting upon absentee rates and the influence of the instructional level upon absentee rates. Both questions appear to be basic to the educational process. The data is analyzed using linear and circular methodologies.

The decision reached by the private school to modify a tuition increase by developing a financial aid program reflects the need for circular analysis. The linear analysis indicates that the tuition increase is appropriate. Circular analysis, in the form of directional data, questions this decision. Upon detailed examination, the data reveal that approximately one-third of the student population reside in communities ranked in the lower socioeconomic quarter. A tuition increase without financial aid could have seriously affected enrollment. The results reported here indicate that the school's decision to raise tuition and to establish a financial aid fund has contributed to a 1986-87 enrollment at the largest level in the last ten years.

The second study uses periodic data to illustrate the use of circular statistics. A large data set of this type is difficult to analyze using circular methods. Circular analysis lacks a statistical procedure comparable to the linear analysis of variance. In the present study linear methods are superior when comparisons are required. An understanding of absenteeism requires a knowledge of when the absences occur. The question of time is of utmost importance. Circular analysis suggests that time period when absenteeism occurs is different for the elementary level, the intermediate level and the high school level. A similar distribution difference is observed when comparing Tracts A and B with Tracts C and D. Two slightly different distributions are suggested. The linear methods used to analyze the data do not identify these differences. The circular analysis provides questions for consideration. When did each tract experience its "average" absentee rate? How could a set of data be dependent and yet vary significantly? Linear methods of analysis do not address these questions. They are answered through the use of circular analysis.

Comparative strengths and weaknesses of circular analysis are demonstrated. Educators consider data daily which require either directional or period analysis. Since statistical analysis is an important element in the decision-making process, the use of circular statistics facilitates this process.

CHAPTER V

SUMMARY AND CONCLUSIONS

Introduction

This chapter summarizes material discussed in the foregoing chapters. The overall purpose of the present study is to demonstrate that circular statistics has a place in the field of education. The following discussion presents the statistical procedures associated with linear and circular data. The computer programs are discussed in light of the recent advances in the area of microcomputer hardware and software. The two educational applications (the Carmel High School for Boys' study and the Unit District R study) are summarized and the effect of the statistical procedures upon the decision-making process are detailed. The concluding section presents suggestions for further research.

Summary and Conclusions

Linear and Circular Data

The use of a statistical methodology is determined by the type of data (linear or circular). Linear data require a measurement expressed as real numbers. Circular data are defined by mapping a line onto a circle by a transcendental function (wrapping function). The manner in which data are represented dictates the statistical method. The circular data in the present study are of two forms: directional and periodic.

Linear and Circular Methods

The statistical procedures comprise two types: descriptive and inferential methods. Chapter III compares several descriptive and inferential, linear and circular procedures. These procedures include measures of central tendency, measures of dispersion, kurtosis, skewness, one-sample test, two-sample test, multi-sample tests, and correlation. Many linear tests are adapted for circular statistics, including the chi-squared test, correlation, the Watson-Williams test and the Watson U^2 test. Correlation in circular statistics is treated as an inferential method. In linear statistics, correlation is treated as either a descriptive or an inferential procedure. Circular tests such as the Rayleigh test and the V-test are based upon the mean vector and have no counterparts in linear analysis. Comparable circular procedures are lacking for such important linear inferential methods as Student's t-test and analysis of variance. The history of circular statistics (Chapter II) coincides with the development of linear statistics. Much of the theoretical research in circular analysis is attributed to the statisticians working in the field of linear analysis.

Computer Programs

The microcomputer programs for the circular procedures shown in Table 1.2 are written to facilitate the comparison of the two statistical methodologies. An IBM-PC computer is used to write the statistical routines. The machine is equipped with two 180K drives, and 256K of RAM. Pascal is the language of choice due to its structured format. The computer programs are compiled using an IBM Pascal Compiler. Eight computer programs are developed for the circular tests shown in Table 2.1. The Pascal programs are shown in Appendix A. A menu format is used wherever possible to facilitate use of the computer programs. Recent developments (the families of 286 and 386 microcomputers) in the field of microcomputers now allow the eight menus to be combined into one program. The computer programs are interactive in design, similar to a calculator, and are not applicable for analysis of large data sets. A data management procedure is needed before large data sets can be analyzed. The program routines are tested using data from biology and medicine (Batschelet, 1981 and Mardia, 1972). These computer programs are used in the analysis of both circular studies.

Application of Circular Statistics to Two Educational Studies

Two educational studies compare and contrast linear and circular statistical procedures. The Carmel High School for Boys' study illustrates a directional circular measure. The attendance study of High School X, Intermediate School Y and Elementary School Z of Unit District R represents a periodic circular measure. These studies compare the strengths and weaknesses associated with both linear and circular analyses.

Summary of Results of the Carmel High School for Boys' Study.

The Carmel High School for Boys' study is designed to address the question of the feasibility of a tuition increase. The linear data consist of the number of students residing in a specific community.

The student's family income is estimated as the median income of the community in which the family resides, as reported by the 1980 census. The median income for the student body is $29,200. Chicago and its environs are ranked by median income (Chicago Tribune, 1982). Two-thirds of the student body at Carmel High School for Boys reside in communities whose median income was ranked in the upper one-half of the Chicagoland communities. Linear analysis of the data supports the tuition increase.

The circular data set consists of the number of students residing in a community and the angular measure depicting the direction of the community. Descriptive circular statistics include the length of the mean vector (.3183) and measure of the mean angle (26°). The communities of Mundelein (0°), Libertyville (19°), North Chicago (25°) and Waukegan (35°) are clustered about the mean angle. Waukegan and North Chicago are ranked by median income as 214 and 259 respectively. The number of students attending Carmel from Waukegan and North Chicago total 86, constituting 14% of the student body. A tuition increase would have a greater negative financial impact on students from those areas. Circular analysis suggests that an increase in tuition would adversely affect enrollment. A review of the linear data by community rank indicates that 31% of the students reside in communities ranked in the lower quarter by median income.

The decision taken by the school was to increase tuition and to establish a fund for financial aid. Prior to this decision, Carmel High School for Boys had experienced a declining enrollment. Since the establishment of the financial aid fund ($20,000), the enrollment has increased each year. Many factors reportedly affect an increase in enrollment: an improving economy and increased building in the area are two examples. The number of students from communities ranked in the lower quarter by median income has increased. The economic growth experienced by the nation has not been experienced in Waukegan and North Chicago. These communities rely on industrial manufacturing for employment. The decision to provide financial help is an important factor reversing the enrollment pattern. Circular analysis provides the information that led to the establishment of the financial aid fund.

The action suggested by the two statistical analyses are summarized as follows:

Linear Methods:

1. Median income of Carmel families is $29,200.
2. Two-thirds of the student body reside in communities whose median income is ranked above the average for the Chicagoland area.
3. The recommendation is to increase tuition.

Circular Methods:

1. The mean angle indicates that students attending Carmel come from a northeasterly (26°).
2. The magnitude of the mean vector is .3183. This indicates that the majority of students attending Carmel reside in communities clustering about the mean angle.

3. The communities between 0° and 35° include Mundelein, Libertyville, Waukegan, and North Chicago. The students from these communities constitute 48 percent of the school's enrollment.
4. Waukegan and North Chicago are ranked 214th and 259th respectively by median income.
5. The recommendation is to increase tuition and to develop a financial aid program to help needy families.

<u>Summary of Results of the Unit District R Study</u>. This study is designed to analyze large data sets using circular and linear methods of analysis. Two questions are asked. First, did the tracting of students at Intermediate School Y affect student attendance? The two factors considered are the number of instruction days that each tract attends school and the daily absentee rates for the school. Second, does the instructional level affect the absentee rate?

Unit District R is chosen for the present study because it operated on a 45-15 schedule from 7/1/78 to 6/30/79 and the total number of students reported absent at High School X (33,652), Intermediate School Y (9,851) and Elementary School Z (6,515) represent three large data sets. The daily recorded absentee rates include the number of students absent for the three tracts in attendance. The number of students absent in a specific tract is not recorded. The linear data used in the present study contains the daily absences from each of the aforementioned schools, based on a 177-day adjusted school year. The circular data describing the yearly attendance consists of assigning an angular measure to each of the 230 days the school is in session. The circular data used to describe individual tracts consists of an angular measure representing the school day and the daily absentee rates when the tract is present.

The linear analysis shows that Tracts A, B, and D receive six fewer instruction days than does Tract C. The tracts are compared using an ANOVA with a CRD design. The null hypothesis of equality of means is not rejected. An ANOVA with an RBD design compares the means of the individual tracts by quarter. Significance is established indicating the means of the same tract differ significantly by quarter. The third quarter exhibits the highest number of absences. School is closed officially four times between 1/15/79 and 1/25/79 due to inclement weather.

The standard deviations indicate that Tract D does not exhibit the same variability as the other three tracts. A review of the data indicates the school is closed unofficially on January 31 and February 1 and the entire student population, consisting of Tracts A, B and C, are reported absent on those dates. The analysis of variance establishes that the absentee rates of the four tracts do not differ significantly. The actual number of school days Tracts A and B attend school is 169, representing 4.5% less instruction time.

The circular data used to compare the four tracts are adjusted to reflect a 177-day school year. Each day of the school year is assigned an angular measure ranging from 0° to 360° with angular increments of 2.0339°. The attendance days represent an ordinal value rather than a specific date. Thus, an angular measurement of 87.4577° represents the 44th day of school for Tract D. An examination of the data reveals that Tracts A, B, and D received 6 fewer days of instruction than Tract C. The lengths of the mean vectors for Tracts A, B, and C are similar. The magnitude of the mean vectors of Tract D deviates the most (.03) from the magnitudes of the other three tracts. The direction of the mean angles indicates that Tracts A and B differ from Tracts C and D. As discussed in the linear case, this occurs because Tracts A and B are scheduled each time the school is closed. Circular two-sample and multi-sample tests are unavailable to compare the four tracts. The lack of circular methods similar to the Student's t-test and an analysis of variance restricts the use of circular statistics.

The absentee rates for the three levels of instruction (high school, intermediate school, elementary school) are compared. Three Pearson's correlation coefficients are used in the linear analysis. Each correlation (high school/intermediate, high school/elementary, intermediate/elementary) is significant at the .01 level. Significance with a probability of less than .001 is obtained using a parametric form of circular correlation known as circular-circular correlation. The V-test is used to determine if the mean angles differ significantly from each other. Significance is established at the .05 level. This indicates that the distributions representing the three levels of instructions are dependent, however, the time period when the "average" absentee rate occurs differ. The measures of the mean angle decrease as the level of instruction increases. This indicates that the elementary school is affected by inclement weather more than the high school. It also indicates that the distribution of absences at the elementary level is more stable than at the other two levels. Both methods found a significant relationship between the levels of instruction. The circular methods identified an "average" time period providing a different perspective to the question. A comparative summary of the linear and circular statistical results and their implications are as follows:

Linear Methods

1. Absentee rates for Intermediate School Y reveal that Tracts A, B, and D attend school six fewer days than Tract C.
2. Tract D is not scheduled on 1/31/79 and 2/1/79 when the school is closed unofficially and all students are reported absent.
3. An analysis of variance using a CRD design indicates that the means (absentee rates) for the four tracts are not significantly different.
4. The conclusion reached is that students in the four tracts receive similar instructional time based on the number of days of attendance and the daily absentee rates.
5. Pearson's correlation coefficients comparing the absentee rates at the high school, intermediate and

elementary levels are significant.

6. The conclusion based upon Pearson's correlation coefficient is that a relationship exists between the absentee rates at each level of instruction.

Circular Methods

1. The absentee data at Intermediate School Y reveal six fewer days of instruction for Tracts A, B, and D than for Tract C.
2. The dispersion in the angular deviation suggests that the absentee rates for Tract D differ from Tracts A, B, and C. An inspection of the circular data reveals that Tract D is not scheduled on 1/31/79 and 2/1/79 when all students are reported absent.
3. None of the circular two-sample and multi-sample tests are appropriate for use with this data set.
4. The conclusion reached is that the four tracts receive similar instruction based upon the number of days absent and the absentee rates.
5. Three circular-circular correlations imply that a dependent relationship exists between the three levels of instruction and their absentee rates.
6. V-tests comparing the mean angles of the three instructional levels are significant, implying that the time the "average" absentee rate occurs is different for the three instructional levels.

7. The conclusion reached based upon the statistical methods suggests that absenteeism exhibits a dependent relationship at each level of instruction. However, each instructional level experiences different patterns of absenteeism.

Suggestions for Further Research

User-friendly computer programs must be developed before circular statistics can be used routinely by the educational community. The advent of the family of 386 machines will provide the hardware necessary to develop these programs. A data management system is necessary to handle data sets of more than 25 items. The future use of circular statistics in education requires the development of computer routines.

Many educational concerns are expressed as periodic measurement. What is the optimum length of a class period to maximize learning? At what period in the day is the student most responsive to instruction? Is class behavior affected by the time period when the class is scheduled? These examples of periodic measure are basic issues in education. Each should be examined carefully with circular procedures.

The development of yearly schedules to maximize the use of the physical plant and the school staff require a study of the time periods when absenteeism occurs. The use of circular statistics as shown in the Unit District R study help to identify such patterns of absenteeism.

The learning styles of individuals differ. A circular analysis of the different methods a student uses to learn over a specific time period would help identify the mode of instruction best suited for that individual.

Circular procedures have been used in the construction and analysis of ability tests and personality inventories. The use of Guttman's circumplex to scale the Minnesota Multiphasic Personality Inventory, the ACT inventories and the Interpersonal Checklists are three examples. These items can be evaluated using circular procedures to determine independence and homogeneity. The use of circular methods to analyze test items is demonstrated by several investigators (McCormick and Kavanagh, 1981; Humphrey and Benjamin, 1986; Guttman, 1954). The use of circular procedures in test construction is an interesting area for further research.

Mathematical research in the field of circular statistics has declined during the past decade. A renewed interest in the theoretical development of the mathematics is necessary. Circular procedures must be developed to identify specific causal relationships, (e.g., whether significance is attributed to the mean angle, angular deviation, magnitude, etc.). There is a need for new circular procedures, such as multi-sample tests. The development in circular analysis of a field analogous to experimental design in linear statistics is of interest.

Data represented as periodic or directional measure require circular analysis. The additional insights that this statistical process provides warrants its serious consideration.

BIBLIOGRAPHY

Asano, C. (1963). Runs test for circular distribution and a table of probabilities. Annals of the Institute for Statistical Mathematics. 17, 331-334.

Barton, D.E., & David, F.N. (1958). Runs in a ring. Biometrika. 45, 572-578.

Batschelet, E. (1965). Statistical methods for the analysis of problems in animal orientation and certain biological rhythms. Washington D.C.: American Institute of Biological Science.

Batschelet, E. (1971). Recent statistical methods for orientation data. Washington: American Institute of Biological Studies.

Batschelet, E. (1981). Circular Statistics in Biology. New York: Academic Press.

Batschelet, E., Hillman, D., Smolensky, M., & Halberg, F. (1973). Angular-linear correlation coefficient for rhythmometry and circannually changing human birth rates at different geographic latitudes. International Journal of Chronobiology. 1, 183-202.

Beran, R.J. (1969). The derivation of nonparametric two-sample tests from tests for uniformity of a circular distribution. Biometrika. 56, 561-570.

Bernoulli, D. (1734). Recherches physiques et astronomiques, sur le probleme propose pour la seconde fois par l'Academie Royale des Sciences des Paris. Recuiel des pieces qui ont remporte le prix de l'Academie Royale des Science, Tome III, 95-134.

Bingham, M.S. (1971). Stochastic processes with independent increments taking values in an abelian group. Proceedings of the London Mathematical Society. 22, 507-530.

Bliss, C.I. (1970). Statistics in Biology (Vol. 2), New York: McGraw-Hill.

Chicago, suburb ratings by median family income. (1982). Chicago Tribune. September 17, 1982, Section 2, Page 3.

Cole, N. (1973). On measuring the vocational interest of women. Journal of Counseling Psychology. 20, 105-112.

Crosswhite, F.J., Wawkinson, L.D., & Sachs, L. (1983). Calculus. Columbus, OH: Merrill.

Daniel, W.W. (1978). Applied nonparametric statistics. Boston, Houghton-Mifflin.

Downs, T.D. (1966). Some relationships among the von Mises distributions of different dimensions. Biometrika. 53, 269-272.

Durand, D., & Greenwood, J.A. (1958). Modification of the Rayleigh test for uniformity in analysis of two-dimensional orientation data. Journal of Geology. 66, 229-238.

Freeman, Joshua (1984). PC's and the social scientists: The chemistry is right. PC Magazine. 3(10), 192-197.

Galton, F. (1889). Natural inheritance. New York: MacMillan.

Glass, G.V., & Kenneth, H.D. (1984). Statistical methods in education and psychology. 2nd edition. Englewood Cliffs, NJ: Prentice-Hall.

Greenwood, J.A., & Durand, D. (1955). The distribution of length and components of the sum of n random unit vectors. Annals of Mathematical Statistics. 26, 233-246.

Gumbel, E.J. (1954). Application of the circular normal distribution. American Statistical Association Journal. 49, 267-297.

Gumbel, E.J., Greenwood, J.A., & Durand, D. (1953). The circular normal distribution: Theory and tables. Journal of the American Statistical Association. 48, 131-152.

Guttman, L. (1954). A new approach to factor analysis: the radex. In P. Lazarsfeld (Ed.). Mathematical thinking in the social sciences (pp. 258-349). Glencoe, IL: Free Press.

Guttmann, L. (1957). Empirical verification of the radex structure of mental abilities and personality traits. Educational and Psychological Measurement. 17, 391-407.

Hays, W.L. (1973). Statistics for the social sciences (2nd ed.). New York: Holt, Rinehart, Winston.

Hays, W.L. (1981). Statistics (3rd ed.). New York: Holt, Rinehart, Winston.

Hill, A.B. (1955). Principals of medical statistics (6th ed.). London: Lancet.

Hotelling, H. (1931). The generalization of Student's ratio. Annals of Mathematical Statistics. 2, 360-378.

Humphrey, L., & Benjamin, L. (1986). Using structural analysis of social behavior to assess critical but elusive family processes. American Psychologist. 41, 979-989.

Jeffreys, H. (1948). Theory of Probability (2nd ed.). Oxford, England: Oxford University Press.

Johnson, R.A., & Wehrly, T. (1977). Measures and models in angular correlation and angular-linear correlation. Journal of the Royal Statistical Society. B39, 222-229.

Jupp, P.E., & Mardia, K.V. (1980). A general correlation coefficient for directional data and related regression problems. Biometrika. 67, 163-173.

Kluyver, J.C. (1905). A local probability theorem. Ned. Akad. Wet. Proc.. 8, 341-351.

Mardia, K.V. (1967). A non-parametric test for the bivariate two-sample location problem. Journal of the Royal Statistical Society. B29, 320-342.

Mardia, K.V., (1968). Small sample power of a nonparametric test for the bivariate two-sample location problem in the normal case. Journal of the Royal Statistical Society. B30, 83-92.

Mardia, K.V., (1969a). On the Wheeler and Watson's two-sample test on a circle. Sankhya. A31, 177-190.

Mardia, K.V., (1969b). On the null distribution of a nonparametric test for the bivariate two-sample problem. Journal of the Royal Statistic Society, B31, 98-102

Mardia, K.V. (1972). Statistics of directional data. New York: Academic Press.

Mardia, K.V. (1975). Statistics of directional data. Journal of the Royal Statistical Society. B37, 349-393.

Mardia, K.V. (1976). Linear circular correlation coefficients and rhythmometry. Biometrika. 63, 403-405.

McCormick, C.C., & Kavanagh, J.A. (1981). Scaling interpersonal checklist items to a circular model. Applied Psychological Measurement. 5(4), 421-447.

Minium, E.W. (1978). Statistical reasoning in psychology and education (2nd ed.). New York: Wiley.

Moore, R.F. (1980). A modification of the Rayleigh test for vector data. Biometrika. 67, 175-180.

Pearson, J. (1905). The problem of the random walk. Nature. 72, 294, 342.

Papakonstantinou, V. (1979). Beitrage zur zirkularen Statistic. PhD. dissertation University of Zurich, Switzerland.

Rao, J.S., & Sengupta, S. (1972). Mathematical techniques for paleo-current analysis: Treatment of directional data. Journal of the International Association of Mathematical Geology. 4, 235-248.

Rao, J.S. (1969). Some contribution to the analysis of circular data. Unpublished doctoral dissertation. Indian Statistical Institute, Calcutta, India.

Rayleigh, Lord. (1880). On the resultant of a large number of vibrations of the same pitch and or arbitrary phase. Philosophic Magazine. 10, 73-80.

Rayleigh, Lord. (1905). The problem of the random walk. Nature. 72, 318.

Schach, S. (1969). The asynmptotic distribution of some nonlinear function of the two-sample rank vector. Annals of Mathematical Statistics. 40, 1011-1020.

Schach, S. (1970). The asynmptotic distribution of a class of nonparametric test statistics. Metrika. 15, 48-58.

Schmidt-Koenig, K. (1963). On the role of the loft, the distance and site of release in pigeon homing. Biological Bulletin. 125, 154-164.

Schuler, M., & Gebelin, H. (1955). Five place tables of elliptical functions based on Jacobi's parameter q. Berlin: Springer.

Slater, P.E. (1962). Parent behavior and the personality of the child. Journal of Genetic Psychology. 101,53-68.

Stephens, M.A. (1963). Random walk on a circle. Biometrika. 50, 385-390.

Stephens, M.A. (1964). The distribution of the goodness-of-fit statistic, U^2. II. Biometrika. 51, 393-397.

Stephens, M.A. (1965). Significance points for the two-sample statistic, U^2. Biometrika. 52, 661-663.

Stephens, M.A. (1969). A goodness-of-fit statistic for the circle, with some comparisons. Biometrika. 56, 161-168.

Stephens, M.A. (1969). Tests for randomness of direction against two circular alternatives. Journal of the American Statistical Association. 64, 280-289.

Stephens, M.A. (1972). Multisample tests for the von Mises distribution. Journal of the American Statistical Association. 67, 456-461.

Stephens, M.A. (1979). Vector correlation. Biometrika. 66, 41-48.

Swed, F.S. & Eisenhart, C. (1943). Tables for testing randomness of grouping in a sequence of alternatives. Annals of Mathematical Statistics. 14, 66-87.

Torgerson, W.S. (1958). Theory and methods of scaling. New York: Wiley.

von Mises, R. (1918). Uber die "Ganzzahligkeit" der Atomgewicht und verwandte Fragen. Physikal. Z. 19, 490-500.

Wald, A., & Wolfowitz, J. (1940). On a test whether two-samples are from the same population. Annals of Mathematical Statistics. 11, 147-162.

Ware, J.H., & Bowden, R.E. (1977). Circadian rhythm analysis when output is collected at intervals. Biometrics. 33, 566-571.

Watson, G.S. (1956). A test for randomness of direction. Monthly Notices of the Royal Astronomical Society. Geophysical Supplement. 7, 160-161.

Watson, G.S. (1961). Goodness-of-fit test on a circle. Biometrika. 48, 109-114.

Watson, G.S. (1962). Goodness-of-fit test on a circle II. Biometrika. 49, 57-63.

Watson, G.S. (1970). Orientation statistics in the earth science. Bulletin of the Geological Institute. University of Uppsala. 2, 73-89.

Watson, G.S., & Williams, E.J. (1956). On the construction of significance tests on the circle and the sphere. Biometrika. 43, 344-352.

Winter, A. (1947). On the shape of the angular case of the Cauchy's distribution curves. Annals of Mathematical Statistics. 18, 589-593.

Wheeler, S., & Watson, G.S. (1964). A distribution-free two-sample test on a circle. Biometrika. 51, 256-257.

Zar, J.H. (1974). Biostatistical Analysis. Englewood Cliffs, NJ: Prentice Hall.

APPENDIX A

Computer Programs for Circular Procedures

1. MEANDIR
2. MENU A
3. MENU B
4. MENU C
5. MENU D
6. MENU E
7. MENU F
8. MENU G

1. MEANDIR

```
PROGRAM MEAN (INPUT, OUTPUT);
   CONST
      j = 1000;
   TYPE
      datafile = array [1..j] of real;
   VAR
      n,i : integer;
      angle : real;
      data : datafile;
   PROCEDURE MEANXY (n : integer; data1 : datafile);
      Const
         pi = 3.1416;
      Var
         x,y,mean_angle,holder,sumcos,sumsin : real;
         i : integer;
      Begin
         sumcos := 0;
         sumsin := 0;
         write('                         ');
         Writeln('BELOW IS GIVEN THE RECTANGULAR');
         write('                         ');
         Writeln('COORDINATES OF THE MEAN VECTOR');
         writeln;
         write('                    ');
         Writeln('Angle     Cosine Values     Sine Values');
         writeln;
            for i := 1 to n do
                begin
                   holder := data1[i]*pi/180;
                   if holder <= 0 then
                      holder := 2*pi+holder;
                   Write('          ',i,'. ',data[i]:5:1);
                   Write('        ',Cos(holder):7:4);
                   Writeln('          ',Sin(holder):7:4);
                   sumcos := sumcos + Cos(holder);
                   sumsin := sumsin + Sin(holder);
                end;
         x := sumcos/n;
         y := sumsin/n;
         writeln;
         write('               TOTAL   ');
         Writeln('          ',sumcos:9:4,'          ',sumsin:9:4);
         writeln;
         Write('The coordinates of the Mean Vector');
         Write(' are x =  ',sumcos/n:5:4,' and y = ');
         Writeln(sumsin/n:5:4,'.');
         Write('The Mean Vector, r = ');
         Writeln(SQRT(SQR(sumcos/n) + SQR(sumsin/n)):5:4,'.');
         if x > 0 then
            begin
               mean_angle := (ARCTAN(y/x))*180/3.1416;
               while mean_angle < 0 do
               mean_angle := 360 + mean_angle;
```

```
      end
      else
      begin
         if x < 0 then
         mean_angle := (3.1416 + ARCTAN(y/x))*180/3.1416
         else
         begin
            if y > 0 then
            mean_angle := 180
            else
            begin
               if y < 0 then
               mean_angle := 270
               else
               Writeln('Mean Angle is undetermined');
            end;
         end;
      end;
   Writeln('The Mean Angle = ',mean_angle:3:1,' degrees.');
End;
PROCEDURE MEDIAN (n : integer; data1 : datafile);
   Var
      m : integer;
   Begin
      m := n DIV 2;
      if n MOD 2 <> 0
         then
         begin
            Write('The Median Angle = ',data1[m+1]:3:1);
            Writeln(' degrees.')
         end
         else
         begin
            Write('The Median Angle = ');
            Writeln((data1[m]+data1[m+1])/2:3:1,' degrees.');
         end;
   End;
FUNCTION SORT (n : integer; data1 : datafile) : datafile;
   Var
      pass,i : integer;
      flag : boolean;
      hold : real;
   Begin
      flag := true;
      pass := 1;
         while (pass <= n - 1) and (flag) do
            begin
               flag := false;
                  for i := 1 to n - pass do
                     if data1[i] > data1[i+1] then
                        begin
                           flag := true;
                           hold := data1[i];
```

```
                              data1[i] := data1[i+1];
                              data1[i+1] := hold;
                           end;
                     pass := pass + 1;
               end;
          SORT := data1;
       End;
    FUNCTION IN_PUT (n : integer) : datafile;
       Var
          data1 : datafile;
       Begin
          for i := 1 to n do
              begin
                 read(angle);
                 data1[i] := angle;
              end;
          IN_PUT := SORT(n,data1);
       End;
BEGIN
   Writeln('Input the number of angles you have.');
   Readln(n);
   writeln;
   Writeln('Input the measure of each angle in degrees.');
   data := IN_PUT(n);
   writeln;
   MEANXY(n,data);
   MEDIAN(n,data);
END.
```

2. MENU A

```
PROGRAM MENUA(Input,Output);
 CONST
    j = 1000;
  TYPE
    datafile = array [1..j] of real;
  VAR
    n,i : integer;
    v,angle : real;
    data : datafile;
    tests : char;
  PROCEDURE MenuA;
    Begin
      Write('--------------------------------');
      Writeln('--------------------------------');
      Writeln('                            MENU A');
      Writeln('       Measures of Dispersions, Skewness, and Kurtosis');
      Write('--------------------------------');
      Writeln('--------------------------------');
      Writeln;
      Writeln('1. Angular Variance                                   1. v');
      Writeln;
      Writeln('2. Mean Angular Deviation                             2. d');
      Writeln;
      Writeln('3. Angular Variance and                               3. c');
      Writeln('   Mean Angular Deviation');
      Writeln;
      Writeln('4. Range                                              4. r');
      Writeln;
      Writeln('5. Skewness                                           5. s');
      Writeln;
      Writeln('6. Kurtosis                                           6. k');
      for i := 1 to 3 do
        Writeln;
    End;
  PROCEDURE Variance(r:real);
    Const
      pi = 3.1416;
    Var
      s : real;
    Begin
      s := 180/pi*2*(1 - r);
      Writeln('The Angular Variance = ',s:3:1,' degrees.');
    End;
  Procedure Dviation(r:real);
    Const
      pi = 3.1416;
    Var
      s:real;
    Begin
      s := 180/pi*SQRT(2*(1 - r));
      Writeln('The Mean Angular Variance, s = ',s:3:1,' degrees.');
    End;
  FUNCTION r1(n:integer; data1:datafile) : real;
```

```
    Const
       pi = 3.1416;
    Var
       holder,sumcos,sumsin : real;
       i : integer;
       x,y : real;
    Begin
       sumcos := 0;
       sumsin := 0;
          for i := 1 to n do
              begin
                 holder := data1[i]*pi/180;
                 if holder < 0 then
                    holder := holder + 2*pi;
                 sumcos := sumcos + Cos(holder);
                 sumsin := sumsin + Sin(holder);
              end;
       x := sumcos/n;
       y := sumsin/n;
       r1 := (SQRT(SQR(sumcos/n) + SQR(sumsin/n)));
    End;
FUNCTION r2(n:integer; data1:datafile) : real;
    Const
       pi = 3.1416;
    Var
       holder,sumcos,sumsin : real;
       i : integer;
       x,y,z : real;
    Begin
       sumcos := 0;
       sumsin := 0;
          for i := 1 to n do
              begin
                 holder := data1[i]*pi/180;
                 if holder < 0 then
                    holder := holder + 2*pi;
                 sumcos := sumcos + Cos(2*holder);
                 sumsin := sumsin + Sin(2*holder);
              end;
       x := sumcos/n;
       y := sumsin/n;
       z := (SQRT(SQR(sumcos/n) + SQR(sumsin/n)));
       r2 := z;
    End;
 FUNCTION SORT(n:integer; data1:datafile) : datafile;
       Var
          pass,i : integer;
          flag : boolean;
          hold : real;
       Begin
          flag := true;
          pass := 1;
             while (pass <= n - 1) and (flag) do
```

```
         begin
            flag := false;
               for i := 1 to n - pass do
                  if data1[i] > data1[i+1] then
                     begin
                        flag := true;
                        hold := data1[i];
                        data1[i] := data1[i+1];
                        data1[i+1] := hold;
                     end;
               pass := pass + 1;
         end;
      SORT := data1;
   End;
PROCEDURE Range(n:integer; data1:datafile);
   Var
      diff : datafile;
   Begin
      for i := 1 to n-1 do
          diff[i] := data1[i+1] - data1[i];
      diff[n] := (360 + data1[1]) - data1[n];
      data1 := SORT(n,diff);
      Writeln('The Range = ',(360 - data1[n]):3:1,' degrees.');
   End;
FUNCTION meanangl(n:integer; data:datafile) : real;
   Const
      pi = 3.1416;
   Var
      holder,sumcos,sumsin : real;
      i : integer;
      x,y : real;
   Begin
      sumcos := 0;
      sumsin := 0;
         for i := 1 to n do
             begin
                holder := data[i]*pi/180;
                if holder < 0 then
                   holder := holder + 2*pi;
                sumcos := sumcos + Cos(holder);
                sumsin := sumsin + Sin(holder);
             end;
      x := sumcos/n;
      y := sumsin/n;
      if x > 0 then
         meanangl := ARCTAN(y/x)
         else
         begin
            if x < 0 then
            meanangl := pi + ARCTAN(y/x)
            else
            begin
               if y > 0 then
```

```
               meanang1 := pi
               else
               begin
                  if y < 0 then
                  meanang1 := 3/2*pi
                  else
                  Writeln('Mean Angle is undetermined.');
               end;
            end;
         end;
      End;
   FUNCTION meanang2(n:integer; data:datafile) : real;
      Const
         pi = 3.1416;
      Var
         holder,sumcos,sumsin : real;
         i : integer;
         x,y : real;
      Begin
         sumcos := 0;
         sumsin := 0;
            for i := 1 to n do
                begin
                   holder := data[i]*pi/180;
                   if holder < 0 then
                      holder := holder + 2*pi;
                   sumcos := sumcos + Cos(2*holder);
                   sumsin := sumsin + Sin(2*holder);
                end;
         x := sumcos/n;
         y := sumsin/n;
         if x > 0 then
            meanang2 := ARCTAN(y/x)
            else
            begin
               if x < 0 then
               meanang2 := pi + ARCTAN(y/x)
               else
               begin
                  if y > 0 then
                  meanang2 := pi
                  else
                  begin
                     if y < 0 then
                     meanang2 := 3/2*pi
                     else
                     Writeln('Mean Angle is undetermined.');
                  end;
               end;
            end;
      End;
PROCEDURE Skewness(n:integer; data1:datafile);
   Var
```

```
         rr,r,s,t,u : real;
      Begin
         r := r1(n,data1);
         s := SQRT(2*(1-r));
         rr := r2(n,data1);
         t := meanang1(n,data1);
         u := meanang2(n,data1);
         Write('The measure of Skewness = ');
         Writeln(rr*SIN(u - 2*t)/(s*s*s):5:4);
      End;
   PROCEDURE Kurtosis(n:integer; data1:datafile);
      Var
         r,rr,s,t,u : real;
      Begin
         r := r1(n,data1);
         s := SQRT(2*(1-r));
         rr := r2(n,data1);
         t := meanang1(n,data1);
         u := meanang2(n,data1);
         Write('The measure of Kurtosis = ');
         Writeln((rr*COS(u - 2*t) - SQR(SQR(1-s)))/SQR(s):5:4);
      End;
      FUNCTION IN_PUT(n : integer) : datafile;
         Var
            data1 : datafile;
         Begin
            Writeln('Input the measure of each angle.');
            Writeln;
            for i := 1 to n do
                 begin
                    read(angle);
                    data1[i] := angle;
                 end;
            IN_PUT := SORT(n,data1);
            readln;
         End;
BEGIN
MenuA;
Writeln('Input your choice of statistical tests.');
Readln(tests);
Writeln('Input the number of angles you have.');
Readln(n);
data := IN_PUT(n);
   Repeat
      Case tests of
      'v' : begin
             v := r1(n,data);
             Variance(v);
            end;
      'd' : begin
             v := r1(n,data);
             Dviation(v);
            end;
```

```
    'c' : begin
            v := r1(n,data);
            Variance(v);
            Dviation(v);
          end;
    'r' : Range(n,data);
    's' : Skewness(n,data);
    'k' : Kurtosis(n,data);
    End;
    Writeln;
    MenuA;
    Writeln('    Input your choice of statistic. If finished enter "f".');
    Readln(tests);
  Until tests = 'f';
END.
```

3. MENU B

```
PROGRAM MENUB(Input,Output);
   CONST
      j = 1000;
   TYPE
      datafile = array [1..j] of real;
   VAR
      m,n,i : integer;
      u,v,angle : real;
      data : datafile;
      tests : char;
   PROCEDURE Clear;
      Begin
         for i := 1 to 27 do
            Writeln;
      End;
   PROCEDURE MenuB;
      Begin
         Write('--------------------------------');
         Writeln('-----------------------------');
         writeln;
         Writeln('                              MENU B');
         Writeln('           Tests for Randomness and Goodness of Fit');
         writeln;
         Write('--------------------------------');
         Writeln('-----------------------------');
         writeln;
         writeln;
         Writeln('                      Tests of Randomness');
         writeln;
         Writeln('1. The Rayleigh Test                                   1. r');
         Writeln;
         Writeln('2. The V Test                                          2. v');
         for i := 1 to 6 do
           writeln;
         Writeln('     This menu requires the use of statistical tables.');
         Writeln('     The user will be asked for input from this source.');
         Writeln;
      End;
   FUNCTION r1(n:integer; data1:datafile) : real;
      Const
         pi = 3.1416;
      Var
         holder,sumcos,sumsin : real;
         i : integer;
         x,y : real;
      Begin
         sumcos := 0;
         sumsin := 0;
            for i := 1 to n do
                begin
                   holder := data1[i]*pi/180;
                   sumcos := sumcos + Cos(holder);
                   sumsin := sumsin + Sin(holder);
```

```
          end;
      x := sumcos/n;
      y := sumsin/n;
      r1 := (SQRT(SQR(sumcos/n) + SQR(sumsin/n)));
   End;
FUNCTION r2(n,m:integer; data1:datafile) : real;
   Const
      pi = 3.1416;
   Var
      holder,sumcos,sumsin : real;
      i : integer;
      x,y,z : real;
   Begin
      sumcos := 0;
      sumsin := 0;
         for i := 1 to n do
            begin
               holder := data1[i]*pi/180;
               sumcos := sumcos + Cos(m*holder);
               sumsin := sumsin + Sin(m*holder);
            end;
      x := sumcos/n;
      y := sumsin/n;
      z := (SQRT(SQR(sumcos/n) + SQR(sumsin/n)));
      r2 := z;
   End;
 FUNCTION SORT(n:integer; data1:datafile) : datafile;
      Var
         pass,i : integer;
         flag : boolean;
         hold : real;
      Begin
         flag := true;
         pass := 1;
            while (pass <= n - 1) and (flag) do
               begin
                  flag := false;
                     for i := 1 to n - pass do
                        if data1[i] > data1[i+1] then
                           begin
                              flag := true;
                              hold := data1[i];
                              data1[i] := data1[i+1];
                              data1[i+1] := hold;
                           end;
                     pass := pass + 1;
               end;
         SORT := data1;
      End;
   FUNCTION meanang1(n:integer; data:datafile) : real;
      Const
         pi = 3.1416;
      Var
```

```
      holder,sumcos,sumsin : real;
      i : integer;
      z,x,y : real;
   Begin
      sumcos := 0;
      sumsin := 0;
         for i := 1 to n do
             begin
                holder := data[i]*pi/180;
                sumcos := sumcos + Cos(holder);
                sumsin := sumsin + Sin(holder);
             end;
      x := sumcos/n;
      y := sumsin/n;
      if x > 0 then
         begin
         z := ARCTAN(y/x);
         meanangl := z;
         while z < 0 do
         z := z + 6.2832;
         meanangl := z;
         end
         else
         begin
            if x < 0 then
            meanangl := pi + ARCTAN(y/x)
            else
            begin
               if y > 0 then
               meanangl := pi
               else
               begin
                  if y < 0 then
                  meanangl := 3/2*pi
                  else
                  Writeln('Mean Angle is undetermined.');
               end;
            end;
         end;
   End;
FUNCTION meanang2(n,m:integer; data:datafile) : real;
   Const
      pi = 3.1416;
   Var
      holder,sumcos,sumsin : real;
      i : integer;
      x,y,z : real;
   Begin
      sumcos := 0;
      sumsin := 0;
         for i := 1 to n do
             begin
                holder := data[i]*pi/180;
```

```
                sumcos := sumcos + Cos(m*holder);
                sumsin := sumsin + Sin(m*holder);
             end;
      x := sumcos/n;
      y := sumsin/n;
      if x > 0 then
         begin
         z := ARCTAN(y/x);
         meanang2 := z;
         while z < 0 do
         z := z + 6.2832;
         meanang2 := z;
         end
         else
         begin
            if x < 0 then
            meanang2 := pi + ARCTAN(y/x)
            else
            begin
               if y > 0 then
               meanang2 := pi
               else
               begin
                  if y < 0 then
                  meanang2 := 3/2*pi
                  else
                  Writeln('Mean Angle is undetermined.');
               end;
            end;
         end;
   End;
FUNCTION IN_PUT(n : integer) : datafile;
   Var
      data1 : datafile;
   Begin
      Writeln('          Enter the measure of each angle.');
      for i := 1 to n do
           begin
              read(angle);
              data1[i] := angle;
           end;
      IN_PUT := SORT(n,data1);
      readln;
   End;
PROCEDURE RAYLEIGH(n:integer ; data1:datafile);
   Var
      z,x : real;
   Begin
      Clear;
      Write('--------------------------------');
      Writeln('-------------------------------------');
      Writeln('PURPOSE');
      Write('The purpose of the Rayleigh test ');
```

```
Writeln('is to test whether the population');
Write('from which the sample is drawn ');
Writeln('differs significantly from randomness.');
Writeln;
Writeln('ASSUMPTIONS');
Write('That for axial data the angles between');
Writeln(' modes are equal.');
Writeln;
Writeln('NULL HYPOTHESIS');
Writeln('That the population is uniformly distributed.');
Writeln;
Write('--------------------------------');
Writeln('------------------------------------');
writeln;
Write('1. If data is unimodal and sample size is 30 or less.');
Writeln('      1. a');
Write('2. If data is unimodal and sample size is 30 or more.');
Writeln('      2. b');
Write('3. If data is multimodal and sample size is 30 or less.');
Writeln('    3. c');
Write('4. If data is multimodal and sample size is 30 or more.');
Writeln('    4. d');
for i := 1 to 4 do
writeln;
Writeln('                     Enter your choice of test');
Readln(tests);
Case tests of
   'a' : begin
           x := r1(n,data1);
           Clear;
           Write('The parameters are the vector r = ',x:5:4);
           Writeln(' and the number of angles = ',n:3);
           writeln;
           Write('------------------------------------');
           Writeln('------------------------------------');
           writeln;
           Writeln('DECISION RULE');
           Write('If the critical level, P, is less than ');
           Writeln('the preassigned level, alpha, the');
           Write('null hypothesis is rejected. For P > alpha');
           Writeln(' randomness cannot be excluded.');
           for i := 1 to 14 do
               writeln;
           Writeln('                  To continue please enter "c".');
           Readln(tests);
        end;
   'b' : begin
           x := r1(n,data1);
           z := n*x*x;
           Clear;
           Write('The z score = ',z:5:4,'. Compare');
           Writeln(' this z score with the preassigned level');
           Writeln('using a table of critical z values.');
```

```
        writeln;
        Write('------------------------------------');
        Writeln('------------------------------------');
        writeln;
        Writeln('DECISION RULE');
        Write('If the value of z is > or = to the ');
        Writeln('preassigned table value, significance');
        Writeln('occurs.');
        for i := 1 to 14 do
            writeln;
        Writeln('                 To continue please enter "c".');
        Readln(tests);
      end;
  'c' : begin
        Writeln('              Enter the number of multimodal axes.');
        readln (m);
        x := r2(n,m,data1);
        Clear;
        Write('The parameters are the vector r = ',x:5:4);
        Writeln(' and the number of angles = ',n:3);
        writeln;
        Write('------------------------------------');
        Writeln('------------------------------------');
        writeln;
        Writeln('DECISION RULE');
        Write('If the critical level, P, is less than ');
        Writeln('the preassigned level, alpha, the');
        Write('null hypothesis is rejected. For P > alpha');
        Writeln(' randomness cannot be excluded.');
        for i := 1 to 14 do
            writeln;
        Writeln('                 To continue please enter "c".');
        Readln(tests);
      end;
  'd' : begin
        Writeln('              Enter the number of multimodal axes.');
        readln (m);
        x := r2(n,m,data1);
        z := n*x*x;
        Clear;
        Write('The z score = ',z:5:4,'.');
        Writeln(' Compare this z score with the preassigned level');
        Writeln('using a table of critical z values.');
        writeln;
        Write('------------------------------------');
        Writeln('------------------------------------');
        writeln;
        Writeln('DECISION RULE');
        Write('If the value of z is > or = to the preassigned');
        Writeln(' table value significance');
        Writeln('occurs.');
        for i := 1 to 14 do
            writeln;
```

```
                Writeln('                To continue please enter "c".');
                Readln(tests);
              end;
      end;
   End;
PROCEDURE VTEST(n:integer ; data1:datafile);
   Var
      d,r,u,v : real;
      m : integer;
   Begin
      Clear;
      Write('-----------------------------------');
      Writeln('-----------------------------------');
      Writeln('PURPOSE');
      Write('The purpose of the V TEST ');
      Writeln('is to determine whether the observed angles ');
      Write('have a tendency to cluster about the ');
      Writeln('given angle, thus whether the');
      Write('distribution differs significantly from ');
      Writeln('randomness.');
      Writeln;
      Writeln('ASSUMPTIONS');
      Write('That for axial data the angles between');
      Writeln(' modes are equal. It is important');
      Write('when entering axial data the hypothetical ');
      Writeln('angle not be multiplied. This');
      Writeln('is provided for in the program.');
      Writeln;
      Writeln('NULL HYPOTHESIS');
      Writeln('That the population is uniformly distributed.');
      Writeln;
      Write('-----------------------------------');
      Writeln('-----------------------------------');
      writeln;
      Write('1. If data is unimodal.                          ');
      Writeln('             1. a');
      Write('2. If data is axial.                             ');
      Writeln('             2. b');
      for i := 1 to 4 do
      writeln;
      Writeln('                   Enter your choice of test');
      Readln(tests);
      Case tests of
         'a' : begin
                Writeln('        Enter the measure of your hypothetical angle');
                Readln(angle);
                u := angle*3.1416/180;
                v := meanangl(n,data1);
                d := ABS(v - u);
                r := r1(n,data1);
                v := r*cos(d);
                u := SQRT(2*n)*v;
                Clear;
```

```
                    Write('The parameter u = ',u:5:4);
                    Writeln(' and the number of angles  n = ',n:3);
                    writeln;
                    Write('------------------------------------');
                    Writeln('-----------------------------------');
                    writeln;
                    Writeln('DECISION RULE');
                    Write('If u is greater than or equal to ');
                    Writeln('the preassigned level of alpha, the null ');
                    Write('hypothesis is rejected. For u less than the');
                    Writeln(' preassigned level, randomness');
                    Writeln('cannot be excluded.');
                    for i := 1 to 14 do
                        writeln;
                    Writeln('                   To continue please enter "c".');
                    Readln(tests);
                  end;
            'b' : begin
                    Writeln('               Enter the number of multimodal axes.');
                    Readln(m);
                    Writeln('         Enter the measure of your hypothetical angle');
                    Readln(angle);
                    u := angle*3.1416/180;
                    v := meanang2(n,m,data1);
                    d := ABS(v - u);
                    r := r2(n,m,data1);
                    v := r*cos(d);
                    u := SQRT(2*n)*v;
                    Clear;
                    Write('The parameter u = ',u:5:4);
                    Writeln(' and the number of angles  n = ',n:3);
                    writeln;
                    Write('------------------------------------');
                    Writeln('-----------------------------------');
                    writeln;
                    Writeln('DECISION RULE');
                    Write('If u is greater than or equal to ');
                    Writeln('the preassigned level of alpha, the null ');
                    Write('hypothesis is rejected. For u less than the');
                    Writeln(' preassigned level, randomness');
                    Writeln('cannot be excluded.');
                    for i := 1 to 14 do
                        writeln;
                    Writeln('                   To continue please enter "c".');
                    Readln(tests);
                  end;
          end;
        End;
BEGIN
Clear;
MenuB;
Writeln('          Enter your choice of statistical test. ');
Readln(tests);
```

```
Writeln('          Enter the number of angles you have. ');
Readln(n);
data := IN_PUT(n);
   Repeat
      Case tests of
      'r' : RAYLEIGH(n,data);
      'v' : VTEST(n,data);
      End;
      Writeln;
      Writeln;
      MenuB;
      Writeln('Enter your choice of statistical test. If finished enter "f".');
      Readln(tests);
   Until tests = 'f';
END.
```

4. MENU C

```
PROGRAM MENUC(Input,Output);
   CONST
      j = 1000;
   TYPE
      datafile = array [1..j] of real;
   VAR
      m,n,i : integer;
      u,v,angle : real;
      data : datafile;
      tests : char;
   PROCEDURE Clear;
      Begin
         for i := 1 to 27 do
            Writeln;
      End;
   PROCEDURE MenuC;
      Begin
         Write('--------------------------------');
         Writeln('-----------------------------');
         writeln;
         Writeln('                                MENU C');
         Writeln('                    Tests for Goodness of Fit');
         writeln;
         Write('--------------------------------');
         Writeln('-----------------------------');
         writeln;
         writeln;
         Writeln('1. The chi-squared Test                                     1. c');
         Writeln;
         Writeln('2. The Watson U-squared Test                                2. w');
         for i := 1 to 8 do
            writeln;
         Writeln('     This menu requires the use of statistical tables.');
         Writeln('     The user will be asked for input from this source.');
         Writeln;
      End;
   FUNCTION r1(n:integer; data1:datafile) : real;
      Const
         pi = 3.1416;
      Var
         holder,sumcos,sumsin : real;
         i : integer;
         x,y : real;
      Begin
         sumcos := 0;
         sumsin := 0;
            for i := 1 to n do
               begin
                  holder := data1[i]*pi/180;
                  if holder < 0 then
                     holder := holder + 2*pi;
                  sumcos := sumcos + Cos(holder);
                  sumsin := sumsin + Sin(holder);
```

```
          end;
      x := sumcos/n;
      y := sumsin/n;
      r1 := (SQRT(SQR(sumcos/n) + SQR(sumsin/n)));
   End;
FUNCTION cr1(n,k:integer; data1:datafile) : real;
   Const
      pi = 3.1416;
   Var
      i : integer;
      r,c : real;
   Begin
     r := r1(n,data1);
     c := (pi/k)/(sin(pi/k));
     cr1 := r*c;
   End;
FUNCTION SORT(n:integer; data1:datafile) : datafile;
      Var
         pass,i : integer;
         flag : boolean;
         hold : real;
      Begin
         flag := true;
         pass := 1;
            while (pass <= n - 1) and (flag) do
               begin
                  flag := false;
                     for i := 1 to n - pass do
                        if data1[i] > data1[i+1] then
                           begin
                              flag := true;
                              hold := data1[i];
                              data1[i] := data1[i+1];
                              data1[i+1] := hold;
                           end;
                     pass := pass + 1;
               end;
         SORT := data1;
      End;
   FUNCTION meanangl(n:integer; data1:datafile) : real;
      Const
         pi = 3.1416;
      Var
         holder,sumcos,sumsin : real;
         i : integer;
         x,y,z : real;
      Begin
         sumcos := 0;
         sumsin := 0;
         for i := 1 to n do
             begin
                holder := data1[i]*pi/180;
                if holder < 0 then
```

```
              holder := holder + 2*pi;
            sumcos := sumcos + Cos(holder);
            sumsin := sumsin + Sin(holder);
          end;
     x := sumcos/n;
     y := sumsin/n;
     if x > 0 then
       begin
       z := ARCTAN(y/x);
       meanangl := z;
       while z < 0 do
         z := z + 2*pi;
       meanangl := z;
       end
       else
       begin
         if x < 0 then
           meanangl := pi + ARCTAN(y/x)
           else
           begin
           if y > 0 then
             meanangl := pi
             else
             begin
             if y < 0  then
               meanangl := 3/2*pi
               else
               Writeln('Mean angle is undetermined.');
             end;
           end;
       end;
     End;
   FUNCTION IN_PUT (n : integer) : datafile;
     Var
       data1 : datafile;
     Begin
       Writeln('Enter the measure of each angle.');
       for i := 1 to n do
         begin
         read(angle);
         data1[i] := angle;
         end;
       IN_PUT := SORT(n,data1);
       readln;
     End;
PROCEDURE CHISQR;
   Var
     meanang,psi,sum,sume,chi,KK,II,v,x,y,z,e,r : real;
     i,j,k,c : integer;
     count : array[1..50] of integer;
     freq : array[1..50] of real;
     data1 : datafile;
   Begin
```

```
Clear;
Write('-----------------------------------');
Writeln('-----------------------------------');
writeln;
Writeln('PURPOSE');
Write('The purpose of the chi-squared test ');
Writeln('is to determine whether the given');
Write('distribution fits the sample. If the');
Writeln(' sample is uniformly distributed');
Writeln('it becomes a test for randomness.');
writeln;
Writeln('ASSUMPTIONS');
Write('That the data is suitably grouped, and');
Writeln(' each group contains an expected');
Write('frequency of four or more. Choice a ');
Writeln('below assumes equal sized groups,');
Write('this need not be the case with ');
Writeln('choice b. You will be asked for input from');
Write('statistical tables. One such source ');
Writeln('is "Circular Statistics in Biology" ');
Writeln('by Edward Batschelet.');
writeln;
Writeln('NULL HYPOTHESIS');
Write('The parent distribution coincides with');
Writeln(' the given circular distribution.');
writeln;
Write('-----------------------------------');
Writeln('-----------------------------------');
writeln;
Write('1. If the given distribution is uniform. ');
Writeln('                            1. a');
Write('2. For distributions other than choice a.');
Writeln('                            2. b');
writeln;
Writeln('                        Enter your choice of test.');
Readln(tests);
writeln;
Clear;
Write('Enter the number of arcs that the');
Writeln(' circle is to be subdivided into.');
Write('If the number of arcs exceeds 50 ');
Writeln('the program must be adjusted. For');
Write('arcs less than twelve a corrected r ');
Writeln('will be provided when needed.');
Readln(k);
Case tests of
   'a' : begin
           Writeln('Enter the number of angles.');
           Readln(n);
           data1 := IN_PUT(n);
           sum := 0;
           for j := 1 to k do
             begin
```

```
          c := 0;
          x := 360*j/k;
          y := 360/k;
          for  i := 1 to n do
            if (data1[i] < x) and (x-y <= data1[i]) then
              begin
              c := c + 1;
              count[j] := c;
              end;
          end;
        e := n/k;
        for j := 1 to k do
          begin
          chi := SQR(count[j] - e)/e;
          sum := sum + chi;
          end;
        writeln;
        Write('The parameters are chi-squared = ',sum:5:4);
        Writeln(' and the degrees of freedom = ',k-1:2,'.');
        Write('The degrees of freedom');
        Writeln(' equal the number of arcs less one.');
        writeln;
        Write('-----------------------------------------');
        Writeln('-------------------------------------');
        writeln;
        Writeln('DECISION RULE');
        Write('The null hypothesis is rejected if ');
        Writeln('the critical level, P, as determined');
        Write('from a table of chi-squared values, is');
        Writeln(' smaller than the preassigned level');
        Writeln('of significance.');
        for i := 1 to 9 do
          writeln;
        Writeln('                     To continue enter "c".');
        for i := 1 to 4 do
          writeln;
        Readln(tests);
      end;
'b' : begin
      writeln;
      Write('------------------------------');
      Writeln('-------------------------------');
      writeln;
      Write('This program requires two entries, ');
      Writeln('first the frequency and');
      Write('second the probability for that ');
      Writeln('frequency. Note that the');
      Write('arcs need not be equal. This program ');
      Writeln('computes the chi-squared');
      Write('value once the relevant data has ');
      Writeln('been entered.');
      writeln;
      Write('------------------------------');
```

```
          Writeln('------------------------------');
          for i := 1 to 4 do
            writeln;
          sum := 0;
          for j := 1 to k do
            begin
            Writeln('Enter the frequency for the ',j:2,' arc.');
            Readln(c);
            count[j] := c;
            Writeln('Enter the probability for the ',j:2,' arc.');
            Readln(e);
            chi := SQR(count[j] - e)/e;
            sum := sum + chi;
            end;
          Write('The parameters are chi-squared = ',sum:5:4);
          Writeln(' and the degrees of freedom = ',k-1:2,'.');
          Write('The degrees of freedom equal the ');
          Writeln('number of arcs less one.');
          Write('------------------------------------');
          Writeln('------------------------------------');
          writeln;
          Writeln('DECISION RULE');
          Write('The null hypothesis is rejected if ');
          Writeln('the critical level, P, as determined');
          Write('from a table of chi-squared values is ');
          Writeln('smaller than the preassigned level');
          Writeln('of signifance.');
          for i := 1 to 12 do
            writeln;
          Writeln('                    To continue enter "c".');
          for i := 1 to 3 do
            writeln;
          Readln(tests);
          end;
    end;
  End;
PROCEDURE WATSONU2;
  Var
    U,v2,v,c,sumv,sumv2,sumc : real;
    i : integer;
  Begin
    Clear;
    Write('--------------------------------------');
    Writeln('------------------------------------');
    writeln;
    Writeln('PURPOSE');
    Write('The purpose of the Watson U-squared test ');
    Writeln('is to determine whether the given');
    Write('distribution fits the sample. When the ');
    Writeln('sample is uniformly distributed, it');
    Writeln('becomes a test for randomness.');
    writeln;
    Writeln('ASSUMPTIONS');
```

```
Write('That the sample points, angles, are not ');
Writeln('grouped. The program will rearrange');
Write('them in ascending order. You may be asked');
Writeln(' for input from statistical tables.');
Write('One such source is "Circular Statistics');
Writeln(' in Biology" by Edward Batschelet.');
writeln;
Writeln('NULL HYPOTHESIS');
Write('The parent distribution coincides with');
Writeln(' the given theoretical distribution.');
writeln;
Write('-------------------------------------');
Writeln('-------------------------------------');
writeln;
Write('1. If the given distribution is uniform. ');
Writeln('                              1. a');
Write('2. For distributions other than choice a.');
Writeln('                              2. b');
writeln;
writeln;
Writeln('                       Enter your choice of test.');
Readln(tests);
writeln;
Case tests of
   'a' : begin
         Clear;
         Writeln('Enter the number of angles.');
         Readln(n);
         data := IN_PUT(n);
         sumv := 0;
         sumv2 := 0;
         sumc := 0;
         for i := 1 to n do
           begin
           v := data[i]/360;
           sumv := sumv + v;
           v2 := SQR(v);
           sumv2 := v2 + sumv2;
           c := (2*i - 1)*v/n;
           sumc := c + sumc;
           end;
         U := sumv2-sumc+n*(1/3-SQR(sumv/n-0.5));
         writeln;
         Writeln('The Watson U-squared value is ',U:5:4,'.');
         writeln;
         Write('-------------------------------------');
         Writeln('-------------------------------------');
         writeln;
         Writeln('DECISION RULE');
         Write('If the sample value of U-squared exceeds ');
         Writeln('the critical level, the table value,');
         Writeln('the null hypothesis is rejected.');
         for i := 1 to 15 do
```

```
              writeln;
             Writeln('                        To continue enter "c".');
             Readln(tests);
             end;
       'b' : begin
             Clear;
             sumv := 0;
             sumv2 := 0;
             sumc := 0;
             Writeln('Enter the number of angles.');
             Readln(n);
             data := IN_PUT(n);
             for i := 1 to n do
               begin
               Writeln('Enter the probability for ',data[i]:3:1,' degrees.');
               Readln(v);
               sumv := v + sumv;
               v2 := SQR(v);
               sumv2 := v2 + sumv2;
               c := (2*i - 1)*v/n;
               sumc := c +sumc;
               end;
             U := sumv2-sumc+n*(1/3-SQR(sumv/n-0.5));
             writeln;
             Writeln('The Watson U-squared value is ',U:5:4,'.');
             writeln;
             Write('----------------------------------------');
             Writeln('-------------------------------------');
             writeln;
             Writeln('DECISION RULE');
             Write('If the sample value of U-squared exceeds ');
             Writeln('the critical level, the table value,');
             Writeln('the null hypothesis is rejected.');
             for i := 1 to 15 do
               writeln;
             Writeln('                        To continue enter "c".');
             Readln(tests);
             end;
       end;
    End;
Begin
Clear;
MenuC;
Writeln('             Enter your choice of statistical test.');
Readln(tests);
  Repeat
    Case tests of
      'c' : CHISQR;
      'w' : WATSONU2;
      End;
      writeln;
      writeln;
      MenuC;
```

```
      Writeln('Enter your choice of statistical test. If finished enter "f".');
      Readln(tests);
      Clear;
  Until tests = 'f';
End.
```

5. MENU D

```
PROGRAM MENUD(Input,Output);
   CONST
      j = 1000;
   TYPE
      datafile = array [1..j] of real;
   VAR
      m,n1,n2,i : integer;
      u,v,angle : real;
      data1,data2 : datafile;
      tests : char;
   PROCEDURE Clear;
      Begin
         for i := 1 to 27 do
            Writeln;
      End;
   PROCEDURE MenuD;
     Begin
     Write('--------------------------------');
     Writeln('-----------------------------');
     writeln;
     Writeln('                              MENU D');
     Writeln('                     Two Sample Comparisons');
     writeln;
     Write('--------------------------------');
     Writeln('-----------------------------');
     writeln;
     Writeln('1. The Watson-William test                              1. w');
     writeln;
     Writeln('2. The chi-squared test                                 2. c');
     writeln;
     Writeln('3. The Watson''s U-squared test                          3. u');
     for i := 1 to 6 do
       writeln;
     Writeln('     This menu requires the use of statistical tables.');
     Writeln('     The user will be asked for input from this source.');
     Writeln;
     End;
   FUNCTION r1(n:integer; data1:datafile) : real;
      Const
         pi = 3.1416;
      Var
         holder,sumcos,sumsin : real;
         i : integer;
         x,y : real;
      Begin
         sumcos := 0;
         sumsin := 0;
            for i := 1 to n do
                begin
                   holder := data1[i]*pi/180;
                   if holder < 0 then
                      holder := holder + 2*pi;
                   sumcos := sumcos + Cos(holder);
```

```
            sumsin := sumsin + Sin(holder);
          end;
      x := sumcos/n;
      y := sumsin/n;
      r1 := (SQRT(SQR(sumcos/n) + SQR(sumsin/n)));
   End;
FUNCTION SORT(n:integer; data1:datafile) : datafile;
      Var
         pass,i : integer;
         flag : boolean;
         hold : real;
      Begin
         flag := true;
         pass := 1;
            while (pass <= n - 1) and (flag) do
               begin
                  flag := false;
                     for i := 1 to n - pass do
                        if data1[i] > data1[i+1] then
                           begin
                              flag := true;
                              hold := data1[i];
                              data1[i] := data1[i+1];
                              data1[i+1] := hold;
                           end;
                     pass := pass + 1;
               end;
         SORT := data1;
      End;
   FUNCTION IN_PUT(n : integer) : datafile;
      Var
         data1 : datafile;
      Begin
         Writeln('Enter the measure of each angle.');
         for i := 1 to n do
              begin
                 read(angle);
                 data1[i] := angle;
              end;
         IN_PUT := SORT(n,data1);
         readln;
      End;
   PROCEDURE WATSON_WILLIAMS(n1,n2:integer ; data1,data2:datafile);
     Var
       angle,sumsin,sumcos,mn,R,R11,R22,psi1,psi2,K,mr,F,g :real;
       n,i,j : integer;
     Begin
       Clear;
       Write('----------------------------------');
       Writeln('--------------------------------------');
       Writeln('PURPOSE');
       Write('To test whether the mean angles of');
       Writeln(' the two samples differ significantly');
```

```
Writeln('from each other.');
writeln;
Writeln('ASSUMPTIONS');
Write('First, the two samples must be taken ');
Writeln('from populations with a von Mises ');
Write('distribution. Second, the parameters');
Writeln(' of concentration, K, are equal for ');
Write('for each population. Third, that the ');
Writeln('parameter of concentration, K, is');
Writeln('large, K > 2.');
writeln;
Writeln('NULL HYPOTHESIS');
Writeln('The mean angles of the two samples are equal.');
Write('---------------------------------');
Writeln('-------------------------------------');
writeln;
Write('The von Mises distribution requires');
Writeln(' the use of statistical tables, such');
Write('as those found in "Circular Statistics ');
Writeln('in Biology" by Edward Batschelet.');
Write('Below are given the parameters needed');
Writeln(' to determine K, from table D1 of');
Writeln('the formentioned text.');
writeln;
R11 := r1(n1,data1)*n1;
R22 := r1(n2,data2)*n2;
n := n1 + n2;
mr := (R11 + R22)/n;
mn := (n1 + n2)/2;
Write('The mean vector r = ',mr:5:4,' and');
Writeln(' the average sample size = ',mn:5:2);
writeln;
Writeln('Enter the concentration parameter, K.');
Readln(k);
sumsin := 0;
sumcos := 0;
for i := 1 to n1 do
  begin
  angle := data1[i];
  if angle < 0 then
    angle := 360 + angle;
  sumsin := sumsin + Sin(3.1416*angle/180);
  sumcos := sumcos + Cos(3.1416*angle/180);
  end;
for i := 1 to n2 do
  begin
  angle := data2[i];
  if angle < 0 then
    angle := 360 + angle;
  sumsin := sumsin + Sin(3.1416*angle/180);
  sumcos := sumcos + Cos(3.1416*angle/180);
  end;
R := SQRT(SQR(sumcos) + SQR(sumsin));
```

```
    g := 1 + (3/(8*K));
    F := g*(n-2)*(R11+R22-R)/(n-(R11+R22));
    Clear;
    Write('F = ',F:6:4,' and the degrees of freedom are ');
    Writeln('1 and ',n-2:3,'. ');
    writeln;
    Write('----------------------------------');
    Writeln('--------------------------------------');
    writeln;
    Writeln('DECISION RULE');
    Write('If the statistic F as calculated is');
    Writeln(' greater than the critical F-value,');
    Write('as shown in a tables of values, the');
    Writeln(' null hypothesis is rejected.');
    for i := 1 to 14 do
      writeln;
    Writeln('                    To continue please enter "c".');
    Readln(tests);
End;
PROCEDURE CHISQR(n1,n2:integer ; data1,data2:datafile);
  TYPE
    count = Array[1..50] of integer;
    e = Array[1..50] of real;
  Var
    sum,i,j,k,c,n,r : integer;
    count1,count2,total : count;
    e1,e2 : e;
    chi,x,y : real;
  Begin
    Clear;
    Write('----------------------------------');
    Writeln('------------------------------------');
    Writeln('PURPOSE');
    Write('To test whether the two samples differ');
    Writeln(' significantly from each other.');
    Writeln('The type of difference is not specified.');
    writeln;
    Writeln('ASSUMPTIONS');
    Write('That all the expected frequencies, e, ');
    Writeln('are at least 5 and the samples');
    Writeln('are independent and random.');
    writeln;
    Writeln('NULL HYPOTHESIS');
    Write('The two samples are drawn from the');
    Writeln(' same population.');
    Write('----------------------------------');
    Writeln('------------------------------------');
    for i := 1 to 10 do
      writeln;
    Write('                ');
    Writeln('Enter the number of groups or equal arcs.');
    Readln(k);
    sum := 0;
```

```
chi := 0;
r := k - 1;
for i := 1 to 50 do
  begin
  count1[i] := 0;
  count2[i] := 0;
  end;
for j := 1 to k do
  begin
  c := 0;
  x := 360*j/k;
  y := 360/k;
  for i := 1 to n1 do
    if (data1[i] < x) and (x-y <= data1[i]) then
      begin
      c := c + 1;
      count1[j] := c;
      end;
  c := 0;
  for i := 1 to n2 do
    if (data2[i] < x) and (x-y <= data2[i]) then
      begin
      c := c + 1;
      count2[j] := c;
      end;
  total[j] := count1[j] + count2[j];
  e1[j] := total[j]*n1/(n1+n2);
  e2[j] := total[j]*n2/(n1+n2);
  if (e1[j] >= 4.5) and (e2[j] >= 4.5) then
    begin
    chi := (SQR(count1[j]-e1[j])/e1[j])+chi;
    chi := (SQR(count2[j]-e2[j])/e2[j])+chi;
    end
    else
    r := r - 1;
  end;
  Clear;
  Write('The calculated value of chi-squared equals ');
  Writeln(chi:5:4,' and the degrees of');
  Writeln('freedon equal ',r:2,'.');
  writeln;
  Write('------------------------------------');
  Writeln('------------------------------------');
  writeln;
  Writeln('DECISION RULE');
  Write('If the statistic chi-squared as calculated');
  Writeln(' is greater than the critical');
  Write('value as shown in a tables of values, the');
  Writeln(' null hypothesis is rejected.');
  for i := 1 to 13 do
    writeln;
  write('                        ');
  Writeln('To continue please enter "c".');
```

```
      Readln(tests);
    end;
  PROCEDURE WATSONU2(n1,n2:integer ; data1,data2:datafile);
    Var
      n11,n22,c,i,j,k,n : integer;
      d1,d2,U,sum1,sum2,e1,e2 : real;
      total : array[1..200] of real;
    Begin
      Clear;
      Write('----------------------------------');
      Writeln('-----------------------------------');
      Writeln('PURPOSE');
      Write('To test whether the two samples differ');
      Writeln(' significantly from each other.');
      Writeln('The type of difference is not specified.');
      writeln;
      Writeln('ASSUMPTIONS');
      Write('That the two samples are independent');
      Writeln(' and random samples of circular');
      Write('observations and that both samples ');
      Writeln('come from populations with a con-');
      Write('tinuous distribution. In case of ');
      Writeln('grouping the class interval should');
      Writeln('should not be greater than five.');
      writeln;
      Writeln('NULL HYPOTHESIS');
      Write('The two samples are drawn from the');
      Writeln(' same parent population.');
      Write('----------------------------------');
      Writeln('-----------------------------------');
      writeln;
      for k := 1 to 200 do
        total[k] := 0;
      j := 1;
      i := 1;
      e1 := 0;
      e2 := 0;
      sum1 := 0;
      sum2 := 0;
      n11 := n1;
      n22 := n2;
      n := n1 + n2;
      c := n;
      for k := 1 to n do
        begin
        if data1[i] = data2[j] then
          if (i+j) MOD 2 = 0 then
            begin
            total[k] := data1[i];
            i := i + 1;
            j := j + 1;
            n22 := n22 - 1;
            c := c - 1;
```

```
      end
      else
      begin
      total[k] := data2[j];
      i := i + 1;
      j := j + 1;
      n11 := n11 -1;
      c := c - 1;
      end
    else
    if (data1[i]>data2[j]) then
      begin
      total[k] := data2[j];
      if j <= n2 then
        j := j + 1
        else
        begin
        total[k] := data1[i];
        i := i + 1;
        end;
      end
      else
      if (data1[i]<data2[j]) then
        begin
        total[k] := data1[i];
        if i <= n1 then
          i := i + 1
          else
          begin
          total[k] := data2[j];
          j := j + 1;
          end;
        end;
end;
i := 1;
j := 1;
for k := 1 to c do
  begin
  if (total[k]=data1[i]) and (total[k]=data2[j]) then
    if (i+j) MOD 2 = 0 then
      begin
      e1 := 1/n11 + e1;
      d1 := e1 - e2;
      d2 := SQR(d1);
      sum1 := sum1 + d1;
      sum2 := sum2 + d2;
      j := j + 1;
      i := i + 1;
      end
      else
      begin
      e2:= 1/n22 + e2;
      d1 := e1 - e2;
```

```
            d2 := SQR(d1);
            sum1 := sum1 + d1;
            sum2 := sum2 + d2;
            j := j + 1;
            i := i + 1;
            end
          else
          if total[k] = data1[i] then
            begin
            e1 := 1/n11 + e1;
            d1 := e1 - e2;
            d2 := SQR(d1);
            sum1 := sum1 + d1;
            sum2 := sum2 + d2;
            i := i + 1;
            end
            else
            if total[k] = data2[j] then
              begin
              e2 := 1/n22 + e2;
              d1 := e1 - e2;
              d2 := SQR(d1);
              sum1 := sum1 + d1;
              sum2 := sum2 + d2;
              j := j + 1;
              end;
        end;
      n := n11 + n22;
      U := (n11*n22)/SQR(n)*(sum2-SQR(sum1)/n);
      Writeln('The calulated U-squared value is ',U:5:4,'.');
      Write('----------------------------------');
      Writeln('------------------------------------');
      writeln;
      Writeln('DECISION RULE');
      Write('If the calculated U-squared value ');
      Writeln('exceeds the table value we reject');
      Write('the null hypothesis and conclude ');
      Writeln('that the samples are significantly');
      Writeln('different.');
      writeln;
      Writeln('                 To continue please enter "c".');
      Readln(tests);
    End;
BEGIN
Clear;
MenuD;
Writeln('          Enter your choice of statistical test. ');
Readln(tests);
writeln;
Writeln('          Enter the data now for the first set.');
writeln;
Writeln('Enter the number of angles you have. ');
Readln(n1);
```

```
data1 := IN_PUT(n1);
writeln;
Writeln('          Enter the data now for the second set.');
writeln;
Writeln('Enter the number of angles you have. ');
Readln(n2);
data2 := IN_PUT(n2);
   Repeat
      Case tests of
      'w' : WATSON_WILLIAMS(n1,n2,data1,data2);
      'c' : CHISQR(n1,n2,data1,data2);
      'u' : WATSONU2(n1,n2,data1,data2);
      End;
      Writeln;
      Writeln;
      MenuD;
      Write('Enter your choice of statistical test.');
      Writeln(' If finished enter "f".');
      Readln(tests);
   Until tests = 'f';
END.
```

6. MENU E

```
PROGRAM MENUE(Input,Output);
   CONST
      j = 100;
   TYPE
      datafile = array [1..j] of real;
   VAR
      m,n1,n2,i : integer;
      u,v,angle : real;
      data1,data2 : datafile;
      tests : char;
   PROCEDURE Clear;
      Begin
         for i := 1 to 27 do
            Writeln;
      End;
   PROCEDURE MenuE;
     Begin
     Write('--------------------------------');
     Writeln('-----------------------------');
     writeln;
     Writeln('                            MENU E');
     Writeln('                Two Sample and Multisample Comparisons');
     writeln;
     Write('--------------------------------');
     Writeln('-----------------------------');
     writeln;
     Writeln('                    Two Sample Comparison Continued');
     writeln;
     Writeln('1. The run test                                          1. r');
     writeln;
     Writeln('                        Multisample Comparisons');
     writeln;
     Writeln('2. The Watson-Williams test                              2. w');
     writeln;
     Writeln('3. The chi-squared test                                  3. c');
     for i := 1 to 4 do
       writeln;
     Writeln('     This menu requires the use of statistical tables.');
     Writeln('     The user will be asked for input from this source.');
     Writeln;
     End;
   FUNCTION r1(n:integer; data1:datafile) : real;
      Const
         pi = 3.1416;
      Var
         holder,sumcos,sumsin : real;
         i : integer;
         x,y : real;
      Begin
         sumcos := 0;
         sumsin := 0;
            for i := 1 to n do
                begin
```

```
              holder := data1[i]*pi/180;
              if holder < 0 then
                 holder := holder + 2*pi;
              sumcos := sumcos + Cos(holder);
              sumsin := sumsin + Sin(holder);
            end;
      x := sumcos/n;
      y := sumsin/n;
      r1 := (SQRT(SQR(sumcos/n) + SQR(sumsin/n)));
   End;
FUNCTION SORT(n:integer; data1:datafile) : datafile;
      Var
         pass,i : integer;
         flag : boolean;
         hold : real;
      Begin
         flag := true;
         pass := 1;
            while (pass <= n - 1) and (flag) do
               begin
                  flag := false;
                     for i := 1 to n - pass do
                        if data1[i] > data1[i+1] then
                           begin
                              flag := true;
                              hold := data1[i];
                              data1[i] := data1[i+1];
                              data1[i+1] := hold;
                           end;
                     pass := pass + 1;
               end;
         SORT := data1;
      End;
   FUNCTION IN_PUT(n : integer) : datafile;
      Var
         data1 : datafile;
      Begin
         Writeln('Enter the measure of each angle.');
         for i := 1 to n do
              begin
                 read(angle);
                 data1[i] := angle;
              end;
         IN_PUT := SORT(n,data1);
         readln;
      End;
   PROCEDURE RUNTEST;
      Var
        n1,n2,c,i,j,k,n : integer;
        data1,data2,total : datafile;
      Begin
        Clear;
        Write('-------------------------------');
```

```
Writeln('------------------------------');
Writeln('PURPOSE');
Write('To test whether the two samples ');
Writeln('differ significantly fron each');
Write('other, particularly with respect ');
Writeln('to the mean angle.');
writeln;
Writeln('ASSUMPTIONS');
Write('The program assumes that neither ');
Writeln('sample will exceed 50 and that');
Write('the user will break all ties before');
Writeln(' entering the data. It is');
Write('also assumed that the samples ');
Writeln('are independent and randon, and');
Write('come from continuous circular ');
Writeln('distributions. Where grouping');
Write('occurs, the class interval');
Writeln(' should be small to avoid ties.');
writeln;
WriteLN('NULL HYPOTHESIS');
Write('The two samples are drawn fron the ');
Writeln('same circular population.');
Write('--------------------------------');
Writeln('------------------------------');
for i := 1 to 4 do
  writeln;
Writeln('          Enter the data now for the first set.');
writeln;
Writeln('Enter the number of angles you have. ');
Readln(n1);
data1 := IN_PUT(n1);
writeln;
Writeln('          Enter the data now for the second set.');
writeln;
Writeln('Enter the number of angles you have. ');
Readln(n2);
data2 := IN_PUT(n2);
Clear;
for i := 1 to 100 do
  total[i] := 400;
n := n1 + n2;
for i := 1 to n1 do
  total[i] := data1[i];
for j := 1 to n2 do
  total[n1+j] := data2[j];
total := SORT(n,total);
total[n+1] := total[1];
c := 0;
k := 1;
i := 1;
j := 1;
While (k <= n) do
  begin
```

```
       if data1[i] = total[k] then
         begin
         while (data1[i] = total[k]) do
           begin
           if i <> n1 then
             i := i + 1;
           k := k + 1;
           end;
         c := c + 1;
         end;
       if data2[j] = total[k] then
         begin
         while (data2[j] = total[k]) do
           begin
           if j <> n2 then
             j := j + 1;
           k := k + 1;
           end;
         c := c + 1;
         end;
       end;
     if ((total[n]=data1[n1]) and (total[n+1]=data1[1])) or
     ((total[n]=data2[n2]) and (total[n+1]=data2[1])) then
       c := c - 1;
     Write('The number of runs equals ',c:2,'. ');
     Writeln('The sample size for the first');
     Write('group is ',n1:2,' and the second ');
     Writeln('sample size is ',n2:2,'.');
     writeln;
     Write('--------------------------------');
     Writeln('------------------------------');
     writeln;
     Writeln('DECISION RULE');
     Write('Table values for the Run Test ');
     Writeln('are found in Edward Batschelet''s');
     Write('text, "Circular Statistics in Biology".');
     Writeln(' If P is smaller than');
     Write('the preassigned level, significance is ');
     Writeln('established.');
     for i := 1 to 12 do
       writeln;
     write('                    ');
     Writeln('To continue please enter "c".');
     Readln(tests);
   End;
PROCEDURE WATSON_WILLIAMS;
  Var
    F,K,g,R,mn,mr,angle,sumsin,sumcos,sumr : real;
    sumn,n,i,j : integer;
    size : array[1..100] of integer;
    rs,data : datafile;
  Begin
    Clear;
```

```
Write('----------------------------------');
Writeln('-----------------------------------');
Writeln('PURPOSE');
Write('To test whether the mean angles of');
Writeln(' the samples differ significantly');
Writeln('from each other.');
writeln;
Writeln('ASSUMPTIONS');
Write('First, the samples must be taken ');
Writeln('from populations with a von Mises ');
Write('distribution. Second, the parameters');
Writeln(' of concentration, K, are equal');
Write('for each population. Third, that the ');
Writeln('parameter of concentration, K,');
Writeln('is large, K > 2.');
writeln;
Writeln('NULL HYPOTHESIS');
Writeln('The mean angles of the samples are equal.');
Write('----------------------------------');
Writeln('-----------------------------------');
writeln;
for i := 1 to 4 do
  writeln;
Writeln('Enter the number of samples.');
Readln(n);
writeln;
sumsin := 0;
sumcos := 0;
sumn := 0;
sumr := 0;
for i := 1 to n do
  begin
  Write('Enter the number of angles you have for the ');
  Writeln(i:2,' sample.');
  Readln(size[i]);
  Write('Enter the measure of each angle for the ');
  Writeln(i:2,' sample.');
  for j := 1 to size[i] do
    begin
    read(angle);
    data[j] := angle;
    if angle < 0 then
      angle := 360 + angle;
    sumsin := sumsin + Sin(3.1416*angle/180);
    sumcos := sumcos + Cos(3.1416*angle/180);
    end;
  rs[i] := r1(size[i],data);
  readln;
  writeln;
  end;
for i := 1 to n do
  begin
  sumn := size[i] + sumn;
```

```
      sumr := (size[i]*rs[i]) + sumr;
      end;
    Clear;
    mn := sumn/n;
    mr := sumr/sumn;
    Write('The von Mises distribution requires');
    Writeln(' the use of statistical tables, such');
    Write('as those found in "Circular Statistics ');
    Writeln('in Biology" by Edward Batschelet.');
    Write('Below are given the parameters needed');
    Writeln(' to determine K, from table D1');
    Writeln('of the for mentioned text.');
    for i := 1 to 8 do
      writeln;
    Write('The mean vector r = ',mr:5:4,' and');
    Writeln(' the average sample size = ',mn:5:2);
    writeln;
    Writeln('Enter the concentration parameter, K.');
    Readln(k);
    R := SQRT(SQR(sumcos) + SQR(sumsin));
    g := 1 + (3/(8*K));
    F := (g*(sumn-n)*(sumr-R))/((n-1)*(sumn-sumr));
    Clear;
    Write('F = ',F:6:4,' and the degrees of freedom are ');
    Writeln(n-1:3,' and ',sumn-n:3,'. ');
    writeln;
    Write('-----------------------------------');
    Writeln('-------------------------------------');
    writeln;
    Writeln('DECISION RULE');
    Write('If the statistic F as calculated is');
    Writeln(' greater than the critical F-value,');
    Write('as shown in a tables of values, the');
    Writeln(' null hypothesis is rejected.');
    for i := 1 to 12 do
      writeln;
    Writeln('                  To continue please enter "c".');
    Readln(tests);
End;
PROCEDURE CHISQR;
  TYPE
    count = Array[1..50] of integer;
  Var
    p,m,s,sum,sumn,i,j,k,l,c,n,r : integer;
    counte,countr,countc : count;
    e,chi,chiminus,x,y : real;
    buffer : array[1..10,1..50] of integer;
    data : datafile;
  Begin
    Clear;
    Write('-----------------------------------');
    Writeln('------------------------------------');
    Writeln('PURPOSE');
```

```
Write('To test whether the samples differ');
Writeln(' significantly from each other. The');
Writeln('type of difference is not specified.');
writeln;
Writeln('ASSUMPTIONS');
Write('That all the expected frequencies, e, ');
Writeln('are at least 5 and the samples');
Write('are independent and random. This ');
Writeln('program will compare at most ten');
Writeln('samples at one time.');
writeln;
Writeln('NULL HYPOTHESIS');
Write('The samples are drawn from the');
Writeln(' same population.');
Write('----------------------------------');
Writeln('-----------------------------------');
for i := 1 to 8 do
  writeln;
Writeln('Enter the number of samples you have.');
Readln(s);
writeln;
Write('Enter the number of equal arcs that');
Writeln(' the samples are to be subdivided.');
Readln(k);
writeln;
for i := 1 to 100 do
  data[i] := 400;
for i := 1 to 50 do
  begin
  counte[i] := 0;
  countr[i] := 0;
  countc[i] := 0;
  end;
sum := 0;
chi := 0;
r := k - 1;
for i := 1 to s do
  begin
  Write('Enter the number of angles for the ');
  Writeln(i:2,' sample.');
  readln(n);
  Write('Enter the measure of each angle for ');
  Writeln('the ',i:2,' sample.');
  for j := 1 to n do
    Read(data[j]);
  readln;
  writeln;
  data := SORT(n,data);
  for j := 1 to k do
    begin
    c := 0;
    x := 360*j/k;
    y := 360/k;
```

```
    for l := 1 to n do
      if (data[l] < x) and (x-y <= data[l]) then
        c := c + 1;
    buffer[i,j] := c;
    end;
  countr[i] := n;
  end;
sumn := 0;
for j := 1 to k do
  begin
  for i := 1 to s do
    begin
    c := buffer[i,j];
    sum := sum + c;
    end;
  countc[j] := sum;
  sumn := sumn + countc[j];
  sum := 0;
  end;
c := 0;
l := 0;
chiminus := 0;
for j := 1 to k do
  begin
  l := l + 1;
  for i := 1 to s do
    begin
    e := (countr[i]*countc[j]/sumn);
    if e < 4.5 then
      counte[l] := j
      else
      chi := SQR(buffer[i,j]-e)/e + chi;
    end;
  end;
  c := 0;
  for p := 1 to 50 do
    if counte[p] > 0 then
      begin
      c := c + 1;
      m := counte[p];
      for i := 1 to s do
        begin
        e := (countr[i]*countc[m]/sumn);
        if e >= 4.5 then
          chiminus := (SQR(buffer[i,m]-e)/e) + chiminus;
        end;
      end;
  chi := chi - chiminus;
  Clear;
  Write('The calculated value of chi-squared equals ');
  Writeln(chi:5:4,' and the degrees of ');
  Write('freedon equal ',r-c:2,'.');
  if c <> 0 then
```

```
            begin
            Write(' The degrees of freedom ');
            Writeln('have been reduced because at');
            Write('least one element in the following ');
            Writeln('column/s do not meet the expected');
            Write('frequency, column/s');
            for i := 1 to 50 do
                if counte[i] > 0 then
                  begin
                  m := counte[i];
                  Write(' ',m:2);
                  end;
            Writeln('.');
            end;
          writeln;
          Write('-----------------------------------');
          Writeln('-----------------------------------');
          writeln;
          Writeln('DECISION RULE');
          Write('If the statistic chi-squared as calculated');
          Writeln(' is greater than the critical');
          Write('value as shown in a tables of values, the');
          Writeln(' null hypothesis is rejected.');
          for i := 1 to 12 do
            writeln;
          write('                    ');
          Writeln('To continue please enter "c".');
          Readln(tests);
      end;
Begin
MenuE;
Writeln('            Enter your choice of statistical test. ');
Readln(tests);
   Repeat
      Case tests of
      'r' : RUNTEST;
      'w' : WATSON_WILLIAMS;
      'c' : CHISQR;
      End;
      Writeln;
      Writeln;
      MenuE;
      Write('Enter your choice of statistical test.');
      Writeln(' If finished enter "f".');
      Readln(tests);
   Until tests = 'f';
End.
```

7. MENU F

```
PROGRAM MENUF(Input, Output);
  Const
    o = 100;
  Type
    datafile = Array[1..o] of real;
  Var
    tests : char;
    i,n,n2 : integer;
    x,y : real;
    datax,datay,datax2,datay2 : datafile;
  Procedure Clear;
    begin
    for i := 1 to 26 do
      writeln;
    end;
  Procedure MenuF;
    Var
      i : integer;
    begin
    Write('------------------------------');
    Writeln('------------------------------');
    writeln;
    Writeln('                      Bivariate Methods');
    writeln;
    Write('------------------------------');
    Writeln('------------------------------');
    writeln;
    Write('1. General Information              ');
    Writeln('                    1. g');
    Writeln('   a. means of the x and y coordinates');
    Writeln('   b. variances of the x and y coordinates');
    Writeln('   c. standard deviations of the x and y coordinates');
    Writeln('   d. covarience of x and y');
    Writeln('   e. coorelation between x and y');
    writeln;
    Write('2. Hotelling''s one-sample Test     ');
    Writeln('                    2. o');
    writeln;
    Write('3. Hotelling''s two-sample Test     ');
    Writeln('                    3. t');
    for i := 1 to 6 do
      writeln;
    end;
  Procedure General(n:integer ; datax,datay:datafile);
    Var
      sumx,sumy,sumx2,sumy2,sumxd,sumyd,sumxy,x,y : real;
      sx1,sx2,sy1,sy2,cov,r : real;
      i : integer;
    begin
    Clear;
    Write('--------------------------------');
    Writeln('------------------------------');
    Write('                 _       _ 2   ');
```

```
Writeln('        _      _ 2      _    _');
Write('  i       x    x-x    (x-x)    ');
Writeln('y    y-y    (y-y)    (x-x)(y-y)');
Write('--------------------------------');
Writeln('------------------------------');
x := 0;
y := 0;
for i := 1 to n do
  begin
  x := datax[i] + x;
  y := datay[i] + y;
  end;
sumx := 0;
sumy := 0;
sumx2 := 0;
sumy2 := 0;
sumxd := 0;
sumyd := 0;
sumxy := 0;
x := x/n;
y := y/n;
for i := 1 to n do
  begin
  sumx := datax[i] + sumx;
  sumxd := datax[i] - x + sumxd;
  sumx2 := SQR(datax[i] - x) + sumx2;
  Write(i:3,datax[i]:9:1,datax[i]-x:6:1,SQR(datax[i]-x):8:2);
  sumy := datay[i] + sumy;
  sumyd := datay[i] - y + sumyd;
  sumy2 := SQR(datay[i] - y) + sumy2;
  Write(datay[i]:7:1,datay[i]-y:6:1,SQR(datay[i]-y):8:2);
  sumxy := sumxy + (datax[i]-x)*(datay[i]-y);
  Writeln((datax[i]-x)*(datay[i]-y):12:2);
  end;
Write('--------------------------------');
Writeln('------------------------------');
Write('Total',sumx:7:1,sumxd:6:1,sumx2:8:2);
Write(sumy:7:1,sumyd:6:2,sumy2:8:2);
Writeln('    ',sumxy:8:2);
writeln;
sx2 := sumx2/(n-1);
sx1 := SQRT(sumx2/(n-1));
Writeln('The mean value of x equals ',x:5:2,'.');
Writeln('The variance of x equals ',sx2:5:2,'.');
Writeln('The standard deviation of x equals',sx1:5:2,'.');
writeln;
sy1 := SQRT(sumy2/(n-1));
sy2 := sumy2/(n-1);
Writeln('The mean value of y equals ',y:5:2,'.');
Writeln('The variance of y equals ',sy2:5:2,'.');
Writeln('The standard deviation of y equals',sy1:5:2,'.');
writeln;
Writeln('The Covariance of x and y equals ',sumxy/(n-1):5:2,'.');
```

```
  Write('The correlation coefficient of x and y equals ');
  Writeln((sumxy/((n-1)*sx1*sy1)):5:4,'.');
  writeln;
  Writeln('                    To continue enter "c".');
  Readln(tests);
  end;
Procedure Hotell(n:integer ; datax,datay:datafile);
  Var
    sumx,sumy,sumx2,sumy2,sumxd,sumyd,sumxy,x,y : real;
    F,T,sx1,sx2,sy1,sy2,cov,r : real;
    i : integer;
  begin
  Clear;
  Write('--------------------------------');
  Writeln('------------------------------');
  Writeln('PURPOSE                          _ _');
  Write('To test whether the sample center ');
  Writeln('(x,y) deviates significantly');
  Writeln('from the origin.');
  writeln;
  Writeln('ASSUMPTIONS');
  Write('That the population fron which ');
  Writeln('the sample is taken is bivariate');
  Writeln('and that the data is not grouped.');
  writeln;
  Writeln('NULL HYPOTHESIS');
  Write('That the population center, u = (u1,u2),');
  Writeln(' be at the origin, 0.');
  Write('--------------------------------');
  Writeln('------------------------------');
  x := 0;
  y := 0;
  for i := 1 to n do
    begin
    x := datax[i] + x;
    y := datay[i] + y;
    end;
  sumx := 0;
  sumy := 0;
  sumx2 := 0;
  sumy2 := 0;
  sumxd := 0;
  sumyd := 0;
  sumxy := 0;
  x := x/n;
  y := y/n;
  for i := 1 to n do
    begin
    sumx := datax[i] + sumx;
    sumxd := datax[i] - x + sumxd;
    sumx2 := SQR(datax[i] - x) + sumx2;
    sumy := datay[i] + sumy;
    sumyd := datay[i] - y + sumyd;
```

```
    sumy2 := SQR(datay[i] - y) + sumy2;
    sumxy := sumxy + (datax[i]-x)*(datay[i]-y);
    end;
  sx2 := sumx2/(n-1);
  sx1 := SQRT(sumx2/(n-1));
  sy1 := SQRT(sumy2/(n-1));
  sy2 := sumy2/(n-1);
  cov := sumxy/(n-1);
  r := cov/(sx1*sy1);
  T := n/(1-r*r)*(SQR(x/sx1)-((2*r*x*y)/(sx1*sy1))+SQR(y/sy1));
  for i := 1 to 3 do
    writeln;
  Write('It is necessary to calculate the ');
  Writeln('table value T. This statistic');
  Write('is calculated using a table of ');
  Writeln('F values. The degrees of freedom');
  Write('are 2 and ',n-2:3,'. The level of alpha is ');
  Writeln('determined by the user.');
  writeln;
  Writeln('                    Enter the table value of F.');
  Readln(F);
  F := 2*(n-1)*F/(n-2);
  Clear;
  Write('The calculated value of T-squared equals ');
  Writeln(T:5:2,'. The table value');
  Writeln('equals ',F:5:2);
  writeln;
  Write('-------------------------------');
  Writeln('-------------------------------');
  Writeln('DECISION RULE');
  Write('If the calculated T-squared value');
  Writeln(' is greater than the table');
  Writeln('value of T-squared the null hypothesis is rejected.');
  for i := 1 to 14 do
    writeln;
  Writeln('                    To continue enter "c".');
  Readln(tests);
  end;
Procedure Hotel2(n,n2:integer;datax,datay,datax2,datay2:datafile);
  Var
    sumx2,sumy2,sumxy,x,y : real;
    sumx22,sumy22,sumxy2,x2,y2 : real;
    xsum,ysum,r,t1,t2,F,T : real;
    i : integer;
  begin
  Clear;
  Write('---------------------------------');
  Writeln('---------------------------------');
  Writeln('PURPOSE');
  Write('To test whether the centers of the ');
  Writeln('two samples deviate significantly');
  Writeln('from each other.');
  writeln;
```

```
Writeln('ASSUMPTIONS');
Writeln('  1. The parent populations are bivariate normal.');
Write('  2. The parent populations have the');
Writeln(' same variance and covariance');
Writeln('  3. The two samples are independent of each other.');
Writeln('  4. The data is not grouped.');
writeln;
Writeln('NULL HYPOTHESIS');
Write('Let u1 and u2 be the populations');
Writeln(' centers. The statement u1 = u2');
Writeln('represents the null hypothesis.');
Write('--------------------------------');
Writeln('--------------------------------');
x := 0;
y := 0;
for i := 1 to n do
  begin
  x := datax[i] + x;
  y := datay[i] + y;
  end;
sumx2 := 0;
sumy2 := 0;
sumxy := 0;
x := x/n;
y := y/n;
for i := 1 to n do
  begin
  sumx2 := SQR(datax[i] - x) + sumx2;
  sumy2 := SQR(datay[i] - y) + sumy2;
  sumxy := sumxy + (datax[i]-x)*(datay[i]-y);
  end;
x2 := 0;
y2 := 0;
for i := 1 to n2 do
  begin
  x2 := datax2[i] + x2;
  y2 := datay2[i] + y2;
  end;
sumx22 := 0;
sumy22 := 0;
sumxy2 := 0;
x2 := x2/n2;
y2 := y2/n2;
for i := 1 to n2 do
  begin
  sumx22 := SQR(datax2[i] - x2) + sumx22;
  sumy22 := SQR(datay2[i] - y2) + sumy22;
  sumxy2 := sumxy2 + (datax2[i]-x2)*(datay2[i]-y2);
  end;
xsum := sumx2 + sumx22;
ysum := sumy2 + sumy22;
r := (sumxy + sumxy2)/SQRT(xsum*ysum);
t1 := (x-x2)/(SQRT(((1/n)+(1/n2))*(xsum/(n+n2-2))));
```

```
    t2 := (y-y2)/(SQRT(((1/n)+(1/n2))*(ysum/(n+n2-2))));
    T := (1/(1-r*r))*(SQR(t1)-(2*r*t1*t2)+SQR(t2));
    for i := 1 to 3 do
      writeln;
    Write('It is necessary to calculate the ');
    Writeln('table value T. This statistic');
    Write('is calculated using a table of ');
    Writeln('F values. The degrees of freedom');
    Write('are 2 and ',n+n2-3:3,'. The level of alpha is ');
    Writeln('determined by the user.');
    writeln;
    Writeln('                    Enter the table value of F.');
    Readln(F);
    F := (2*(n+n2-2)*F)/(n+n2-3);
    Clear;
    Write('The calculated value of T-squared equals ');
    Writeln(T:5:2,'. The table value');
    Writeln('equals ',F:5:2);
    writeln;
    Write('--------------------------------');
    Writeln('--------------------------------');
    Writeln('DECISION RULE');
    Write('If the calculated T-squared value');
    Writeln(' is greater than the table');
    Writeln('value of T-squared the null hypothesis is rejected.');
    for i := 1 to 14 do
      writeln;
    Writeln('                     To continue enter "c".');
    Readln(tests);
    end;
BEGIN
MenuF;
Writeln('          Enter the letter for the test of your choice.');
writeln;
Readln(tests);
writeln;
if (tests = 't') then
  Writeln('            Enter the data for the first sample.')
  else
  Writeln('             Enter the data now for the sample.');
writeln;
Writeln('Enter the number of sample points.');
Readln(n);
writeln;
Write('Enter the coordinates for each sample point.');
Writeln(' Enter the x ');
Writeln('ordinate first, a space, followed by the y ordinate.');
for i := 1 to n do
  begin
  Read(x);
  datax[i] := x;
  Read(y);
  datay[i] := y;
```

```
  end;
Readln;
While (tests <> 'f') do
  begin
  Case (tests) of
    'g' : General(n,datax,datay);
    'o' : Hotel1(n,datax,datay);
    't' : begin
          writeln;
          if (tests = 't') then
          Write('           ');
          Writeln('Enter the data now for the second sample.');
          writeln;
          Writeln('Enter the number of sample points.');
          Readln(n2);
          writeln;
          Write('Enter the coordinates for each sample point.');
          Writeln(' Enter the x ');
          Writeln('ordinate first, a space, followed by the y ordinate.');
          for i := 1 to n2 do
            begin
            Read(x);
            datax2[i] := x;
            Read(y);
            datay2[i] := y;
            end;
          Readln;
          Hotel2(n,n2,datax,datay,datax2,datay2);
          end;
    end;
  MenuF;
  Writeln('Enter the letter for the test of your choice. If finished');
  Writeln('enter "f".');
  Readln(tests);
  end;
END.
```

8. MENU G

```
PROGRAM MENUG(Input,Output);
   CONST
      j = 1000;
   TYPE
      datafile = array [1..j] of real;
   VAR
      n,i : integer;
      a,b,v,angle : real;
      rplus,rminus,data1,data2 : datafile;
      tests : char;
   PROCEDURE Clear;
      Begin
      for i := 1 to 26 do
        writeln;
      End;
   PROCEDURE Menu6;
      Begin
         writeln;
         Write('--------------------------------');
         Writeln('--------------------------------');
         Writeln('                              MENU 6');
         Writeln('                      Circular Correlation');
         Write('--------------------------------');
         Writeln('--------------------------------');
         writeln;
         writeln;
         Write('1. Paired Samples - Uniform Distribution');
         Writeln('                   1. a');
         Writeln;
         Write('2. Circular-circular Samples - Parametric');
         Writeln('                  2. b');
         Writeln;
         Write('3. Circular-circular Samples - Nonparametric');
         Writeln('               3. c');
         Writeln;
         Write('4. Circular-linear Samples - Parametric');
         Writeln('                    4. d');
         Writeln;
         Write('5. Circular-linear Samples - Nonparametric');
         Writeln('                 5. e');
         for i := 1 to 7 do
           Writeln;
      End;
   FUNCTION r1(n:integer; data1:datafile) : real;
      Const
         pi = 3.1416;
      Var
         holder,sumcos,sumsin : real;
         i : integer;
         x,y : real;
      Begin
         sumcos := 0;
         sumsin := 0;
```

```
        for i := 1 to n do
            begin
               holder := data1[i]*pi/180;
               if holder < 0 then
                  holder := holder + 2*pi;
               sumcos := sumcos + Cos(holder);
               sumsin := sumsin + Sin(holder);
            end;
      x := sumcos/n;
      y := sumsin/n;
      r1 := (SQRT(SQR(sumcos/n) + SQR(sumsin/n)));
   End;
FUNCTION r2(n:integer; data1:datafile) : real;
   Const
      pi = 3.1416;
   Var
      holder,sumcos,sumsin : real;
      i : integer;
      x,y,z : real;
   Begin
      sumcos := 0;
      sumsin := 0;
         for i := 1 to n do
            begin
               holder := data1[i]*pi/180;
               if holder < 0 then
                  holder := holder + 2*pi;
               sumcos := sumcos + Cos(2*holder);
               sumsin := sumsin + Sin(2*holder);
            end;
      x := sumcos/n;
      y := sumsin/n;
      z := (SQRT(SQR(sumcos/n) + SQR(sumsin/n)));
      r2 := z;
   End;
 FUNCTION SORT(n:integer; data1:datafile) : datafile;
      Var
         pass,i : integer;
         flag : boolean;
         hold : real;
      Begin
         flag := true;
         pass := 1;
            while (pass <= n - 1) and (flag) do
               begin
                  flag := false;
                     for i := 1 to n - pass do
                        if data1[i] > data1[i+1] then
                           begin
                              flag := true;
                              hold := data1[i];
                              data1[i] := data1[i+1];
                              data1[i+1] := hold;
```

```
                 end;
              pass := pass + 1;
           end;
        SORT := data1;
     End;
FUNCTION meanang1(n:integer; data:datafile) : real;
     Const
        pi = 3.1416;
     Var
        holder,sumcos,sumsin : real;
        i : integer;
        x,y : real;
     Begin
        sumcos := 0;
        sumsin := 0;
           for i := 1 to n do
               begin
                  holder := data[i]*pi/180;
                  if holder < 0 then
                     holder := holder + 2*pi;
                  sumcos := sumcos + Cos(holder);
                  sumsin := sumsin + Sin(holder);
               end;
        x := sumcos/n;
        y := sumsin/n;
        if x > 0 then
           meanang1 := ARCTAN(y/x)
           else
           begin
              if x < 0 then
              meanang1 := pi + ARCTAN(y/x)
              else
              begin
                 if y > 0 then
                 meanang1 := pi
                 else
                 begin
                    if y < 0 then
                    meanang1 := 3/2*pi
                    else
                    Writeln('Mean Angle is undetermined.');
                 end;
              end;
           end;
     End;
FUNCTION meanang2(n:integer; data:datafile) : real;
     Const
        pi = 3.1416;
     Var
        holder,sumcos,sumsin : real;
        i : integer;
        x,y : real;
     Begin
```

```
      sumcos := 0;
      sumsin := 0;
         for i := 1 to n do
             begin
                holder := data[i]*pi/180;
                if holder < 0 then
                   holder := holder + 2*pi;
                sumcos := sumcos + Cos(2*holder);
                sumsin := sumsin + Sin(2*holder);
             end;
      x := sumcos/n;
      y := sumsin/n;
      if x > 0 then
         meanang2 := ARCTAN(y/x)
         else
         begin
            if x < 0 then
            meanang2 := pi + ARCTAN(y/x)
            else
            begin
               if y > 0 then
               meanang2 := pi
               else
               begin
                  if y < 0 then
                  meanang2 := 3/2*pi
                  else
                  Writeln('Mean Angle is undetermined.');
               end;
            end;
         end;
   End;
Function Corr(n:integer;datax,datay:datafile):real;
  Var
    sumx,sumy,sumx2,sumy2,sumxd,sumyd,sumxy,x,y : real;
    sx1,sx2,sy1,sy2,cov,r : real;
    i : integer;
  begin
  x := 0;
  y := 0;
  for i := 1 to n do
    begin
    x := datax[i] + x;
    y := datay[i] + y;
    end;
  sumx := 0;
  sumy := 0;
  sumx2 := 0;
  sumy2 := 0;
  sumxd := 0;
  sumyd := 0;
  sumxy := 0;
  x := x/n;
```

```
  y := y/n;
  for i := 1 to n do
    begin
    sumx := datax[i] + sumx;
    sumxd := datax[i] - x + sumxd;
    sumx2 := SQR(datax[i] - x) + sumx2;
    sumy := datay[i] + sumy;
    sumyd := datay[i] - y + sumyd;
    sumy2 := SQR(datay[i] - y) + sumy2;
    sumxy := sumxy + (datax[i]-x)*(datay[i]-y);
    end;
  sx2 := sumx2/(n-1);
  sx1 := SQRT(sumx2/(n-1));
  sy1 := SQRT(sumy2/(n-1));
  sy2 := sumy2/(n-1);
  corr := sumxy/((n-1)*sx1*sy1);
  end;
PROCEDURE Uniform(n:integer;rplus,rminus:datafile);
  Var
    a,b,c : real;
  Begin
  Clear;
  Write('--------------------------------');
  Writeln('--------------------------------');
  Writeln('ASSUMPTIONS');
  Writeln('  1. That the two variates are dependent.');
  Writeln('  2. That the two variates are paired samples.');
  Write('  3. The two variates must be uniformly ');
  Writeln('distributed on the circle.');
  Write('  4. That the observed pairs are independent');
  Writeln(' of each other.');
  writeln;
  Writeln('TESTS of SIGNIFICANCE');
  Write('Any test for Randomness may be applied to the ');
  Writeln('correlation, r');
  Writeln('such as the Rayleigh Test or the V Test.');
  Write('--------------------------------');
  Writeln('--------------------------------');
  for i := 1 to 2 do
    writeln;
  a := r1(n,rplus);
  b := r1(n,rminus);
  if a > b then
     c := a
     else
     c := b;
  Write('The value of the positive correlation is ');
  Writeln(a:5:4,' and the value ');
  Write('of the negative correlation is ');
  Writeln(b:5:4,'. The correlation, r is ');
  Writeln('is the Max(r+,r-) and equals ',c:5:4);
  for i := 1 to 5 do
    writeln;
```

```
  Write('                        ');
  Writeln('To continue enter "c".');
  Readln(tests);
  End;
PROCEDURE Cir_Cir_par(n:integer;data1,data2:datafile);
  Var
    datax,datay : datafile;
    a,b,c,d,rcc,rsc,r1,rcs,rss,r2,rsqr,r : real;
  Begin
  Clear;
  Write('--------------------------------');
  Writeln('--------------------------------');
  Writeln('ASSUMPTIONS');
  Writeln('  1. That the two variates are circular.');
  Write('  2. The pairs are a randon sample of');
  Writeln(' bivariate measurements.');
  writeln;
  Writeln('TESTS of SIGNIFICANCE');
  Write('The test statistic approximates a chi-squared ');
  Writeln('distribution with');
  Write('4 df. A source is Table G, in');
  Writeln(' "Circular Statistics in Biology"');
  Writeln('by Edward Batschelet.');
  Write('--------------------------------');
  Writeln('--------------------------------');
  writeln;
  for i := 1 to n do
    begin
      while data1[i] < 0 do
        data1[i] := data1[i] + 360;
      while data2[i] < 0 do
        data2[i] := data2[i] + 360;
    end;
  for i := 1 to n do
    begin
    data1[i] := data1[i]*3.1416/180;
    data2[i] := data2[i]*3.1416/180;
    end;
  for i := 1 to n do
    begin
      datax[i] := Cos(data1[i]);
      datay[i] := Cos(data2[i]);
    end;
  rcc := Corr(n,datax,datay);
  for i := 1 to n do
    begin
      datax[i] := Sin(data1[i]);
      datay[i] := Cos(data2[i]);
    end;
  rsc := Corr(n,datax,datay);
  for i := 1 to n do
    begin
      datax[i] := Cos(data1[i]);
```

```
      datay[i] := Sin(data1[i]);
    end;
  r1 := Corr(n,datax,datay);
  for i := 1 to n do
    begin
      datax[i] := Cos(data1[i]);
      datay[i] := Sin(data2[i]);
    end;
  rcs := Corr(n,datax,datay);
  for i := 1 to n do
    begin
      datax[i] := Sin(data1[i]);
      datay[i] := Sin(data2[i]);
    end;
  rss := Corr(n,datax,datay);
  for i := 1 to n do
    begin
      datax[i] := Cos(data2[i]);
      datay[i] := Sin(data2[i]);
    end;
  r2 := Corr(n,datax,datay);
  a := rcc*rcc+rcs*rcs+rsc*rsc+rss*rss;
  b := (2*r1*r2)*(rcc*rss+rcs*rsc);
  c := (2*r2)*(rcc*rcs+rsc*rss);
  d := (2*r1)*(rcc*rsc+rcs*rss);
  rsqr := (a + b - c - d)/((1-r1*r1)*(1-r2*r2));
  r := SQRT(rsqr);
  Write('The correlation r = ',r:5:4,'.');
  Writeln('. The parameters are n = ',n:3);
  Write('and r-squared = ',r*r:5:4);
  Writeln(' The test statistic n*r*r = ',n*r*r:5:4,'.');
  for i := 1 to 5 do
    writeln;
  Write('                         ');
  Writeln('To continue enter "c".');
  Readln(tests);
  End;
PROCEDURE Cir_Cir_Nonpar(n:integer;data1,data2:datafile);
  Var
    c : real;
    j,k : integer;
    hold1,hold2,hold :datafile;
  Begin
  Clear;
  Write('--------------------------------');
  Writeln('--------------------------------');
  Writeln('ASSUMPTIONS');
  Writeln('  1. That the two variates are circular.');
  Write('  2. The pairs are a randon sample of');
  Writeln(' bivariate measurements.');
  writeln;
  Writeln('TESTS of SIGNIFICANCE');
  Write('The significance test for r-squared requires ');
```

```
  Writeln('a table of critical');
  Write('values. A source is Table W, in');
  Writeln(' "Circular Statistics in Biology"');
  Writeln('by Edward Batschelet.');
  Write('---------------------------------');
  Writeln('---------------------------------');
  for i := 1 to 4 do
    writeln;
  hold1 := SORT(n,data1);
  hold2 := SORT(n,data2);
  for i := 1 to n do
    begin
    for j := 1 to n do
      begin
      if data1[i] = hold1[j] then
        begin
        hold1[j] := j*360/n;
        for k := 1 to n do
          begin
          if data2[i] = hold2[k] then
            hold[j] := k*360/n;
          end;
        end;
      end;
    end;
    for i := 1 to n do
      begin
      rplus[i] := hold[i] - hold1[i];
      rminus[i] := hold1[i] + hold[i];
      end;
    a := r1(n,rplus);
    b := r1(n,rminus);
    if a > b then
       c := a
       else
       c := b;
    Write('The value of the positive correlation is ');
    Writeln(a:5:4,' and the value ');
    Write('of the negative correlation is ');
    Writeln(b:5:4,'. The correlation, r is ');
    Write('the Max(r+,r-) and equals ',c:5:4);
    Writeln('. The parameters are n = ',n:3);
    Writeln('and r-squared = ',c*c:5:4,'.' );
    for i := 1 to 5 do
    writeln;
    Write('                         ');
    Writeln('To continue enter "c".');
    Readln(tests);
  End;
PROCEDURE Cir_Lin_Par(n:integer;data1,data2:datafile);
  Var
    datax,datay : datafile;
    rcy,rsy,rcs,r : real;
```

```
Begin
Clear;
Write('--------------------------------');
Writeln('--------------------------------');
Writeln('ASSUMPTIONS');
Write('  1. That one variate is circular and');
Writeln(' other is linear.');
Write('  2. The data pairs are independent');
Writeln(' of each other.');
Writeln('  3. That the regression line is sinusoidal.');
Write('  4. The joint distribution of x and y should');
Writeln(' be normal.');
writeln;
Writeln('TESTS of SIGNIFICANCE');
Write('The test statistic approximates a chi-squared ');
Writeln('distribution with');
Write('2 degrees of freedom, when n is large.');
Writeln(' A reference is Table 6,');
Write('in "Circular Statistics in Biology"');
Writeln(' by Edward Batschelet.');
Write('--------------------------------');
Writeln('--------------------------------');
writeln;
for i := 1 to n do
  begin
    while data1[i] < 0 do
      data1[i] := data1[i] + 360;
  end;
for i := 1 to n do
  begin
  data1[i] := data1[i]*3.1416/180;
  end;
for i := 1 to n do
  begin
    datax[i] := Cos(data1[i]);
    datay[i] := data2[i];
  end;
rcy := Corr(n,datax,datay);
for i := 1 to n do
  begin
    datax[i] := Cos(data1[i]);
    datay[i] := Sin(data1[i]);
  end;
rcs := Corr(n,datax,datay);
for i := 1 to n do
  begin
    datax[i] := Sin(data1[i]);
    datay[i] := data2[i];
  end;
rsy := Corr(n,datax,datay);
r := (SQR(rcy)+SQR(rsy)-2*rcy*rsy*rcs)/(1-SQR(rcs));
Write('The correlation r = ',SQRT(r):5:4,'.');
Writeln(' The parameters are n = ',n:3);
```

```
  Write('and r-squared = ',r:5:4,'.');
  Writeln(' The test statistic n*r*r = ',n*r:5:4,'.');
  for i := 1 to 5 do
    writeln;
  Write('                         ');
  Writeln('To continue enter "c".');
  Readln(tests);
  End;
PROCEDURE Cir_Lin_Nonpar(n:integer;data1,data2:datafile);
  Const
    pi = 3.1416;
  Var
    datac,datal,datax,datay : datafile;
    x,e,a,psi,C,S,Dn,Cn,An,Un : real;
    i,j : integer;
  Begin
  Clear;
  Write('--------------------------------');
  Writeln('--------------------------------');
  Writeln('ASSUMPTIONS');
  Write('  1. That one variate is circular and');
  Writeln(' other is linear.');
  Write('  2. The ordered pairs of data must be ');
  Writeln('independent of each other.');
  Write('  3. For good results the regression ');
  Writeln('curve must exhibit a single');
  Writeln('     peak and a single trough per period.');
  writeln;
  Writeln('TESTS of SIGNIFICANCE');
  Write('The test statistic used in this nonparametric');
  Writeln(' test is U. A Table ');
  Write('of critical values, Table X, is found in');
  Writeln(' "Circular Statistics in');
  Writeln('Biology" by Edward Batschelet.');
  Write('--------------------------------');
  Writeln('--------------------------------');
  e := 2*pi/n;
  C := 0;
  S := 0;
  for i := 1 to n do
    begin
      datac[i] := 0;
      datal[i] := 0;
    end;
  for i := 1 to n do
    begin
      datax[i] := data1[i];
      datay[i] := data2[i];
    end;
  datax := Sort(n,datax);
  datay := Sort(n,datay);
  for i := 1 to n do
    begin
```

```
          for j := 1 to n do
            If (data1[i] = datax[j]) and (datac[i] <> j) then
               datac[i] := j;
        end;
      for i := 1 to n do
        data1[i] := datac[i]*e;
      for i := 1 to n do
        begin
          for j := 1 to n do
            If (data2[i] = datay[j]) and (data1[i] <> j) then
               data1[i] := j;
        end;
      for i := 1 to n do
        data2[i] := data1[i];
      for i := 1 to n do
        begin
          C := data2[i]*Cos(data1[i]) + C;
          S := data2[i]*Sin(data1[i]) + S;
        end;
      psi := SQR(Cos(e/2)/Sin(e/2));
      x := 1+Cos(e/2);
      if n MOD 2 = 0 then
         An :=  1/(1+5*psi+4*SQR(psi))
         else
         An := 2*SQR(SQR(Sin(e/2)))*x*x*x;
      Dn := An * (SQR(C) + SQR(S));
      Cn := 24/(An*n*n*(n+1));
      Un := Cn*Dn;
      Write('The calculated value of the test statistic, U,');
      Writeln(' is ',Un:5:4,'.');
      writeln;
      for i := 1 to 5 do
        writeln;
      Write('                         ');
      Writeln('To continue enter "c".');
      Readln(tests);
      End;
BEGIN
Menu6;
Writeln('            Input your choice of statistical tests.');
Readln(tests);
writeln;
Writeln('              Input the number of angles you have.');
Readln(n);
writeln;
Write('Enter the value of the first variate');
Writeln(' followed by a space and then');
Write('the value of the second variate. If');
Writeln(' one variate is linear enter');
Writeln('the circular variable first.');
for i := 1 to n do
  begin
  Read(a);
```

```
  Read(b);
  Data1[i] := a;
  Data2[i] := b;
  rplus[i] := a - b;
  rminus[i] := a + b;
  end;
Readln;
Writeln;
   Repeat
      Case tests of
        'a' : Uniform(n,rplus,rminus);
        'b' : Cir_Cir_Par(n,data1,data2);
        'c' : Cir_Cir_Nonpar(n,data1,data2);
        'd' : Cir_Lin_Par(n,data1,data2);
        'e' : Cir_Lin_Nonpar(n,data1,data2);
      end;
      Menu6;
      Writeln('    Input your choice of statistic. If finished enter "f".');
      Readln(tests);
   Until tests = 'f';
END.
```

APPENDIX B

Daily Absentee Rates of High School X, Intermediate School Y, Elementary School Z

Date	Schedule	Student Absences: High School Y	Intermediate X	Elementary Z
7/11/78	ABD	0	23	21
7/12/78	ABD	33	25	19
7/13/78	ABD	156	18	17
7/14/78	ABD	162	28	20
7/15/78	ABD	201	40	29
7/17/78	ABD	180	47	15
7/18/78	ABD	168	37	12
7/19/78	ABD	130	32	15
7/20/78	ABD	119	43	19
7/21/78	ABD	149	50	28
7/24/78	ABD	135	33	25
7/25/78	ABD	127	39	19
7/26/78	ABD	123	34	24
7/27/78	ABD	114	31	13
7/28/78	ABD	176	43	29
7/31/86	ABC	179	26	28
8/1/78	ABC	136	21	10
8/2/78	ABC	131	32	14
8/3/78	ABC	131	22	17
8/4/78	ABC	166	33	19
8/7/78	ABC	158	35	27
8/8/78	ABC	145	25	25
8/9/78	ABC	154	37	21
8/10/78	ABC	125	26	20
8/11/78	ABC	161	42	31
8/14/78	ABC	153	29	24
8/15/78	ABC	143	29	21
8/16/78	ABC	138	26	14
8/17/78	ABC	145	28	15
8/18/78	ABC	183	53	28
8/21/78	BCD	182	30	30
8/22/78	BCD	133	26	22
8/23/78	BCD	132	20	15
8/24/78	BCD	160	39	29
8/25/78	BCD	198	31	30
8/28/78	BCD	170	40	24
8/29/78	BCD	130	25	16
8/30/78	BCD	135	28	17
8/31/78	BCD	143	31	16
9/1/78	BCD	160	37	17
9/5/78	BCD	157	41	23
9/6/78	BCD	142	33	19
9/7/78	BCD	106	35	18
9/8/78	BCD	155	33	16
9/11/78	ACD	141	31	9
9/12/78	ACD	118	44	15
9/13/78	ACD	119	43	16

Date	Schedule	Student Absences: High School Y	Intermediate X	Elementary Z
9/14/78	ACD	140	35	17
9/15/78	ACD	132	41	14
9/18/78	ACD	180	42	12
9/19/78	ACD	105	42	14
9/20/78	ACD	116	40	11
9/21/78	ACD	112	40	11
9/22/78	ACD	121	42	12
9/25/78	ACD	128	30	22
9/26/78	ACD	125	25	18
9/27/78	ACD	133	31	17
9/28/78	ACD	100	28	25
9/29/78	ACD	132	32	28
10/2/78	ABD	135	41	26
10/3/78	ABD	121	40	18
10/4/78	ABD	115	35	23
10/5/78	ABD	120	37	23
10/6/78	ABD	144	38	20
10/10/78	ABD	148	30	18
10/11/78	ABD	109	26	26
10/12/78	ABD	103	31	16
10/13/78	ABD	250	35	25
10/16/78	ABD	152	39	16
10/17/78	ABD	104	25	17
10/18/78	ABD	110	29	17
10/19/78	ABD	136	32	19
10/23/78	ABC	123	31	17
10/24/78	ABC	145	28	14
10/25/78	ABC	107	27	12
10/26/78	ABC	108	28	15
10/27/78	ABC	131	35	17
10/30/78	ABC	124	24	17
10/31/78	ABC	128	26	14
11/1/78	ABC	121	27	18
11/2/78	ABC	123	25	25
11/3/78	ABC	124	31	30
11/6/78	ABC	143	24	29
11/7/78	ABC	105	24	22
11/8/78	ABC	105	17	19
11/9/78	ABC	97	21	27
11/10/78	ABC	126	28	30
11/13/78	BCD	142	35	26
11/14/78	BCD	136	29	19
11/15/78	BCD	141	26	14
11/16/78	BCD	188	41	24
11/17/78	BCD	178	55	25
11/20/78	BCD	144	42	25
11/21/78	BCD	153	50	20

Date	Schedule	Student Absences High School Y	Intermediate X	Elementary Z
11/22/78	BCD	188	70	26
11/27/78	BCD	169	59	28
11/28/78	BCD	134	39	17
11/29/78	BCD	108	30	20
11/30/78	BCD	127	36	16
12/1/78	BCD	234	88	44
12/4/78	BCD	129	41	30
12/5/78	BCD	94	39	22
12/6/78	BCD	96	39	21
12/7/78	ACD	489	202	70
12/8/78	ACD	157	67	39
12/11/78	ACD	120	41	21
12/12/78	ACD	85	35	21
12/13/78	ACD	108	31	35
12/14/78	ACD	110	26	24
12/15/78	ACD	128	35	32
12/18/78	ACD	130	40	35
12/19/78	ACD	104	39	21
12/20/78	ACD	125	44	22
12/21/78	ACD	223	48	21
1/2/79	ACD	429	179	78
1/3/79	ACD	266	130	48
1/4/79	ACD	143	64	31
1/5/79	ACD	138	56	32
1/8/79	ABD	161	40	29
1/9/79	ABD	138	30	31
1/10/79	ABD	108	28	26
1/11/79	ABD	127	39	28
1/12/79	ABD	144	43	27
1/17/79	ABD	737	308	111
1/18/79	ABD	346	179	74
1/19/79	ABD	519	234	109
1/22/79	ABD	123	26	14
1/23/79	ABD	113	19	13
1/26/79	ABD	265	99	31
1/29/79	ABC	122	15	20
1/30/79	ABC	94	29	16
1/31/79	ABC	1301	493	435
2/1/79	ABC	166	493	435
2/2/79	ABC	107	49	435
2/5/79	ABC	168	43	17
2/6/79	ABC	104	28	16
2/7/79	ABC	111	20	16
2/8/79	ABC	96	29	14
2/9/79	ABC	115	27	16
2/13/79	ABC	114	27	18
2/14/79	ABC	98	22	13

Date	Schedule	Student Absences: High School Y	Intermediate X	Elementary Z
2/15/79	ABC	103	26	18
2/16/79	ABC	108	22	22
2/19/79	BCD	181	49	26
2/20/79	BCD	129	41	17
2/21/79	BCD	209	44	20
2/22/79	BCD	88	42	17
2/23/79	BCD	111	42	16
2/26/79	BCD	112	45	21
2/27/79	BCD	118	48	13
2/28/79	BCD	112	48	17
3/1/79	BCD	126	39	15
3/2/79	BCD	139	47	48
3/5/79	BCD	188	54	114
3/6/79	BCD	119	49	109
3/7/79	BCD	93	33	89
3/8/79	BCD	106	36	62
3/9/79	BCD	140	35	75
3/12/79	ACD	128	31	32
3/13/79	ACD	111	36	30
3/14/79	ACD	178	67	32
3/15/79	ACD	160	54	38
3/16/79	ACD	118	45	29
3/19/79	ACD	150	32	34
3/20/79	ACD	134	45	26
3/21/79	ACD	134	42	22
3/22/79	ACD	134	37	22
3/23/79	ACD	165	43	22
3/26/79	ACD	137	46	25
3/27/79	ACD	118	43	16
3/28/79	ACD	105	33	16
3/29/79	ACD	83	33	13
3/30/79	ACD	91	42	19
4/2/79	ABD	134	16	21
4/3/79	ABD	112	15	19
4/4/79	ABD	94	21	14
4/5/79	ABD	105	28	17
4/6/79	ABD	118	27	16
4/9/79	ABD	149	41	21
4/10/79	ABD	137	23	19
4/11/79	ABD	123	41	23
4/12/79	ABD	158	37	22
4/23/79	ABD	122	29	15
4/24/79	ABD	113	25	18
4/25/79	ABD	117	27	18
4/26/79	ABD	116	19	12
4/27/79	ABD	133	28	17
4/30/79	ABC	110	29	11

Date	Schedule	Student Absences: High School Y	Intermediate X	Elementary Z
5/1/79	ABC	115	19	4
5/2/79	ABC	130	24	5
5/3/79	ABC	117	29	4
5/4/79	ABC	132	33	9
5/7/79	ABC	137	35	17
5/8/79	ABC	112	24	7
5/9/79	ABC	137	21	10
5/10/79	ABC	111	27	11
5/11/79	ABC	151	35	14
5/14/79	ABC	131	25	14
5/15/79	ABC	72	26	14
5/16/79	ABC	105	19	13
5/17/79	ABC	93	20	17
5/18/79	ABC	451	24	19
5/21/79	BCD	102	26	15
5/22/79	BCD	85	20	16
5/23/79	BCD	82	34	15
5/24/79	BCD	108	36	20
5/25/79	BCD	149	50	19
5/29/79	BCD	116	38	19
5/30/79	BCD	86	30	16
5/31/79	BCD	94	30	14
6/1/79	BCD	123	38	14
6/4/79	BCD	122	33	16
6/5/79	BCD	93	20	14
6/6/79	BCD	114	22	9
6/7/79	BCD	87	27	9
6/8/79	BCD	150	29	18
6/11/79	ACD	128	39	30
6/12/79	ACD	106	31	21
6/13/79	ACD	93	36	14
6/14/79	ACD	166	23	11
6/15/79	ACD	423	42	15
6/18/79	ACD	144	39	17
6/19/79	ACD	108	32	14
6/20/79	ACD	93	23	8
6/21/79	ACD	80	29	16
6/22/79	ACD	99	33	15
6/25/79	ACD	122	38	25
6/26/79	ACD	124	23	13
6/27/79	ACD	216	42	33
6/28/79	ACD	0	0	0

APPENDIX C

Daily Absentee Rates by Quarter

1. Tract A
2. Tract B
3. Tract C
4. Tract D

1. Tract A

STUDENT ATTENDANCE - Tract A
First Quarter

Date	Tracts	Angular Measure	Student Absences: High School X	Student Absences: Intermediate Y	Student Absences: Elementary Z
7/11/78	ABD	0.0000	0	23	21
7/12/78	ABD	2.0339	33	25	19
7/13/78	ABD	4.0678	156	18	17
7/14/78	ABD	6.1017	162	28	20
7/15/78	ABD	8.1356	201	40	29
7/17/78	ABD	10.1695	180	47	15
7/18/78	ABD	12.2034	168	37	12
7/19/78	ABD	14.2373	130	32	15
7/20/78	ABD	16.2712	119	43	19
7/21/78	ABD	18.3051	149	50	28
7/24/78	ABD	20.3390	135	33	25
7/25/78	ABD	22.3729	127	39	19
7/26/78	ABD	24.4068	123	34	24
7/27/78	ABD	26.4407	114	31	13
7/28/78	ABD	28.4746	176	43	29
7/31/78	ABC	30.5085	179	26	28
8/1/78	ABC	32.5424	136	21	10
8/2/78	ABC	34.5763	131	32	14
8/3/78	ABC	36.6102	131	22	17
8/4/78	ABC	38.6441	166	33	19
8/7/78	ABC	40.6780	158	35	27
8/8/78	ABC	42.7119	145	25	25
8/9/78	ABC	44.7458	154	37	21
8/10/78	ABC	46.7797	125	26	20
8/11/78	ABC	48.8136	161	42	31
8/14/78	ABC	50.8475	153	29	24
8/15/78	ABC	52.8814	143	29	21
8/16/78	ABC	54.9153	138	26	14
8/17/78	ABC	56.9492	145	28	15
8/18/78	ABC	58.9831	183	53	28
9/11/78	ACD	61.0170	141	31	9
9/12/78	ACD	63.0509	118	44	15
9/13/78	ACD	65.0848	119	43	16
9/14/78	ACD	67.1187	140	35	17
9/15/78	ACD	69.1526	132	41	14
9/18/78	ACD	71.1865	180	42	12
9/19/78	ACD	73.2204	105	42	14
9/20/78	ACD	75.2543	116	40	11
9/21/78	ACD	77.2882	112	40	11
9/22/78	ACD	79.3221	121	42	12
9/25/78	ACD	81.3560	128	30	22
9/26/78	ACD	83.3899	125	25	18
9/27/78	ACD	85.4238	133	31	17
9/28/78	ACD	87.4577	100	28	25

STUDENT ATTENDANCE - Tract A
Second Quarter

Date	Tracts	Angular Measure	Student Absences: High School X	Intermediate Y	Elementary Z
9/29/78	ACD	89.4916	132	32	28
10/2/78	ABD	91.5255	135	41	26
10/3/78	ABD	93.5594	121	40	18
10/4/78	ABD	95.5933	115	35	23
10/5/78	ABD	97.6272	120	37	23
10/6/78	ABD	99.6611	144	38	20
10/9/78	Free	101.6950	150	44	29
10/10/78	ABD	103.7289	148	30	18
10/11/78	ABD	105.7628	109	26	26
10/12/78	ABD	107.7967	103	31	16
10/13/78	ABD	109.8306	250	35	25
10/16/78	ABD	111.8645	152	39	16
10/17/78	ABD	113.8984	104	25	17
10/18/78	ABD	115.9323	110	29	17
10/19/78	ABD	117.9662	136	32	19
10/20/78	Free	120.0001	150	44	29
10/23/78	ABC	122.0340	123	31	17
10/24/78	ABC	124.0679	145	28	14
10/25/78	ABC	126.1018	107	27	12
10/26/78	ABC	128.1357	108	28	15
10/27/78	ABC	130.1696	131	35	17
10/30/78	ABC	132.2035	124	24	17
10/31/78	ABC	134.2374	128	26	14
11/1/78	ABC	136.2713	121	27	18
11/2/78	ABC	138.3052	123	25	25
11/3/78	ABC	140.3391	124	31	30
11/6/78	ABC	142.3730	143	24	29
11/7/78	ABC	144.4069	105	24	22
11/8/78	ABC	146.4408	105	17	19
11/9/78	ABC	148.4747	97	21	27
11/10/78	ABC	150.5086	126	28	30
12/7/78	ACD	152.5425	489	202	70
12/8/78	ACD	154.5764	157	67	39
12/11/78	ACD	156.6103	120	41	21
12/12/78	ACD	158.6442	85	35	21
12/13/78	ACD	160.6781	108	31	35
12/14/78	ACD	162.7120	110	26	24
12/15/78	ACD	164.7459	128	35	32
12/18/78	ACD	166.7798	130	40	35
12/19/78	ACD	168.8137	104	39	21
12/20/78	ACD	170.8476	125	44	22
12/21/78	ACD	172.8815	223	48	21
1/2/79	ACD	174.9154	429	179	78
1/3/79	ACD	176.9493	266	130	48
1/4/79	ACD	178.9832	143	64	31
1/5/79	ACD	181.0171	138	56	32

STUDENT ATTENDANCE - Tract A
Third Quarter

			Student Absences		
Date	Tracts	Angular Measure	High School X	Intermediate Y	Elementary Z
1/8/79	ABD	183.0510	161	40	29
1/9/79	ABD	185.0849	138	30	31
1/15/79	Closed	187.1188	150	44	29
1/16/79	Closed	189.1527	150	44	29
1/10/79	ABD	191.1866	108	28	26
1/11/79	ABD	193.2205	127	39	28
1/12/79	ABD	195.2544	144	43	27
1/17/79	ABD	197.2883	737	308	111
1/18/79	ABD	199.3222	346	179	74
1/19/79	ABD	201.3561	519	234	109
1/22/79	ABD	203.3900	123	26	14
1/23/79	ABD	205.4239	113	19	13
1/24/79	Closed	207.4578	150	44	29
1/25/79	Closed	209.4917	150	44	29
1/26/79	ABD	211.5256	265	99	31
1/29/79	ABC	213.5595	122	15	20
1/30/79	ABC	215.5934	94	29	16
1/31/79	ABC	217.6273	1301	493	435
2/1/79	ABC	219.6612	166	493	435
2/2/79	ABC	221.6951	107	49	435
2/5/79	ABC	223.7290	168	43	17
2/6/79	ABC	225.7629	104	28	16
2/7/79	ABC	227.7968	111	20	16
2/8/79	ABC	229.8307	96	29	14
2/9/79	ABC	231.8646	115	27	16
2/13/79	ABC	233.8985	114	27	18
2/14/79	ABC	235.9324	98	22	13
2/15/79	ABC	237.9663	103	26	18
2/16/79	ABC	240.0002	108	22	22
3/12/79	ACD	242.0341	128	31	32
3/13/79	ACD	244.0680	111	36	30
3/14/79	ACD	246.1019	178	67	32
3/15/79	ACD	248.1358	160	54	38
3/16/79	ACD	250.1697	118	45	29
3/19/79	ACD	252.2036	150	32	34
3/20/79	ACD	254.2375	134	45	26
3/21/79	ACD	256.2714	134	42	22
3/22/79	ACD	258.3053	134	37	22
3/23/79	ACD	260.3392	165	43	22
3/26/79	ACD	262.3731	137	46	25
3/27/79	ACD	264.4070	118	43	16
3/28/79	ACD	266.4409	105	33	16
3/29/79	ACD	268.4748	83	33	13
3/30/79	ACD	270.5087	91	42	19

STUDENT ATTENDANCE - Tract A
Fourth Quarter

Date	Tracts	Angular Measure	Student Absences: High School X	Intermediate Y	Elementary Z
4/2/79	ABD	272.5426	134	16	21
4/3/79	ABD	274.5765	112	15	19
4/4/79	ABD	276.6104	94	21	14
4/5/79	ABD	278.6443	105	28	17
4/6/79	ABD	280.6782	118	27	16
4/9/79	ABD	282.7121	149	41	21
4/10/79	ABD	284.7460	137	23	19
4/11/79	ABD	286.7799	123	41	23
4/12/79	ABD	288.8138	158	37	22
4/23/79	ABD	290.8477	122	29	15
4/24/79	ABD	292.8816	113	25	18
4/25/79	ABD	294.9155	117	27	18
4/26/79	ABD	296.9494	116	19	12
4/27/79	ABD	298.9833	133	28	17
4/30/79	ABC	301.0172	110	29	11
5/1/79	ABC	303.0511	115	19	4
5/2/79	ABC	305.0850	130	24	5
5/3/79	ABC	307.1189	117	29	4
5/4/79	ABC	309.1528	132	33	9
5/7/79	ABC	311.1867	137	35	17
5/8/79	ABC	313.2206	112	24	7
5/9/79	ABC	315.2545	137	21	10
5/10/79	ABC	317.2884	111	27	11
5/11/79	ABC	319.3223	151	35	14
5/14/79	ABC	321.3562	131	25	14
5/15/79	ABC	323.3901	72	26	14
5/16/79	ABC	325.4240	105	19	13
5/17/79	ABC	327.4579	93	20	17
5/18/79	ABC	329.4918	451	24	19
6/11/79	ACD	331.5257	128	39	30
6/12/79	ACD	333.5596	106	31	21
6/13/79	ACD	335.5935	93	36	14
6/14/79	ACD	337.6274	166	23	11
6/15/79	ACD	339.6613	423	42	15
6/18/79	ACD	341.6952	144	39	17
6/19/79	ACD	343.7291	108	32	14
6/20/79	ACD	345.7630	93	23	8
6/21/79	ACD	347.7969	80	29	16
6/22/79	ACD	349.8308	99	33	15
6/25/79	ACD	351.8647	122	38	25
6/26/79	ACD	353.8986	124	23	13
6/27/79	ACD	355.9325	216	42	33
6/28/79	ACD	357.9664	0	0	0

2. Tract B

STUDENT ATTENDANCE - Tract B
First Quarter

Date	Tracts	Angular Measure	Student Absences High School X	Intermediate Y	Elementary Z
7/11/78	ABD	0.0000	0	23	21
7/12/78	ABD	2.0339	33	25	19
7/13/78	ABD	4.0678	156	18	17
7/14/78	ABD	6.1017	162	28	20
7/15/78	ABD	8.1356	201	40	29
7/17/78	ABD	10.1695	180	47	15
7/18/78	ABD	12.2034	168	37	12
7/19/78	ABD	14.2373	130	32	15
7/20/78	ABD	16.2712	119	43	19
7/21/78	ABD	18.3051	149	50	28
7/24/78	ABD	20.3390	135	33	25
7/25/78	ABD	22.3729	127	39	19
7/26/78	ABD	24.4068	123	34	24
7/27/78	ABD	26.4407	114	31	13
7/28/78	ABD	28.4746	176	43	29
7/31/86	ABC	30.5085	179	26	28
8/1/78	ABC	32.5424	136	21	10
8/2/78	ABC	34.5763	131	32	14
8/3/78	ABC	36.6102	131	22	17
8/4/78	ABC	38.6441	166	33	19
8/7/78	ABC	40.6780	158	35	27
8/8/78	ABC	42.7119	145	25	25
8/9/78	ABC	44.7458	154	37	21
8/10/78	ABC	46.7797	125	26	20
8/11/78	ABC	48.8136	161	42	31
8/14/78	ABC	50.8475	153	29	24
8/15/78	ABC	52.8814	143	29	21
8/16/78	ABC	54.9153	138	26	14
8/17/78	ABC	56.9492	145	28	15
8/18/78	ABC	58.9831	183	53	28
8/21/78	BCD	61.0170	182	30	30
8/22/78	BCD	63.0509	133	26	22
8/23/78	BCD	65.0848	132	20	15
8/24/78	BCD	67.1187	160	39	29
8/25/78	BCD	69.1526	198	31	30
8/28/78	BCD	71.1865	170	40	24
8/29/78	BCD	73.2204	130	25	16
8/30/78	BCD	75.2543	135	28	17
8/31/78	BCD	77.2882	143	31	16
9/1/78	BCD	79.3221	160	37	17
9/5/78	BCD	81.3560	157	41	23
9/6/78	BCD	83.3899	142	33	19
9/7/78	BCD	85.4238	106	35	18
9/8/78	BCD	87.4577	155	33	16

STUDENT ATTENDANCE - Tract B
Second Quarter

Date	Tracts	Angular Measure	Student Absences: High School X	Intermediate Y	Elementary Z
10/2/78	ABD	89.4918	135	41	26
10/3/78	ABD	91.5257	121	40	18
10/4/78	ABD	93.5596	115	35	23
10/5/78	ABD	95.5935	120	37	23
10/6/78	ABD	97.6274	144	38	20
10/9/78	Free	99.6613	147	42	30
10/10/78	ABD	101.6952	148	30	18
10/11/78	ABD	103.7291	109	26	26
10/12/78	ABD	105.7630	103	31	16
10/13/78	ABD	107.7969	250	35	25
10/16/78	ABD	109.8308	152	39	16
10/17/78	ABD	111.8647	104	25	17
10/18/78	ABD	113.8986	110	29	17
10/19/78	ABD	115.9325	136	32	19
10/20/78	Free	117.9664	147	42	30
10/23/78	ABC	120.0003	123	31	17
10/24/78	ABC	122.0342	145	28	14
10/25/78	ABC	124.0681	107	27	12
10/26/78	ABC	126.1020	108	28	15
10/27/78	ABC	128.1359	131	35	17
10/30/78	ABC	130.1698	124	24	17
10/31/78	ABC	132.2037	128	26	14
11/1/78	ABC	134.2376	121	27	18
11/2/78	ABC	136.2715	123	25	25
11/3/78	ABC	138.3054	124	31	30
11/6/78	ABC	140.3393	143	24	29
11/7/78	ABC	142.3732	105	24	22
11/8/78	ABC	144.4071	105	17	19
11/9/78	ABC	146.4410	97	21	27
11/10/78	ABC	148.4749	126	28	30
11/13/78	BCD	150.5088	142	35	26
11/14/78	BCD	152.5427	136	29	19
11/15/78	BCD	154.5766	141	26	14
11/16/78	BCD	156.6105	188	41	24
11/17/78	BCD	158.6444	178	55	25
11/20/78	BCD	160.6783	144	42	25
11/21/78	BCD	162.7122	153	50	20
11/22/78	BCD	164.7461	188	70	26
11/27/78	BCD	166.7800	169	59	28
11/28/78	BCD	168.8139	134	39	17
11/29/78	BCD	170.8478	108	30	20
11/30/78	BCD	172.8817	127	36	16
12/1/78	BCD	174.9156	234	88	44
12/4/78	BCD	176.9495	129	41	30
12/5/78	BCD	178.9834	94	39	22
12/6/78	BCD	181.0173	96	39	21

STUDENT ATTENDANCE - Tract B
Third Quarter

			Student Absences		
		Angular	High School	Intermediate	Elementary
1/8/79	ABD	183.0510	161	40	29
1/9/79	ABD	185.0849	138	30	31
1/10/79	ABD	187.1188	108	28	26
1/11/79	ABD	189.1527	127	39	28
1/12/79	ABD	191.1866	144	43	27
1/15/79	Closed	193.2205	147	42	30
1/16/79	Closed	195.2544	147	42	30
1/17/79	ABD	197.2883	737	308	111
1/18/79	ABD	199.3222	346	179	74
1/19/79	ABD	201.3561	519	234	109
1/22/79	ABD	203.3900	123	26	14
1/23/79	ABD	205.4239	113	19	13
1/24/79	Closed	207.4578	147	42	30
1/25/79	Closed	209.4917	147	42	30
1/26/79	ABD	211.5256	265	99	31
1/29/79	ABC	213.5595	122	15	20
1/30/79	ABC	215.5934	94	29	16
1/31/79	ABC	217.6273	1301	493	435
2/1/79	ABC	219.6612	166	493	435
2/2/79	ABC	221.6951	107	49	435
2/5/79	ABC	223.7290	168	43	17
2/6/79	ABC	225.7629	104	28	16
2/7/79	ABC	227.7968	111	20	16
2/8/79	ABC	229.8307	96	29	14
2/9/79	ABC	231.8646	115	27	16
2/13/79	ABC	233.8985	114	27	18
2/14/79	ABC	235.9324	98	22	13
2/15/79	ABC	237.9663	103	26	18
2/16/79	ABC	240.0002	108	22	22
2/19/79	BCD	242.0341	181	49	26
2/20/79	BCD	244.0680	129	41	17
2/21/79	BCD	246.1019	209	44	20
2/22/79	BCD	248.1358	88	42	17
2/23/79	BCD	250.1697	111	42	16
2/26/79	BCD	252.2036	112	45	21
2/27/79	BCD	254.2375	118	48	13
2/28/79	BCD	256.2714	112	48	17
3/1/79	BCD	258.3053	126	39	15
3/2/79	BCD	260.3392	139	47	48
3/5/79	BCD	262.3731	188	54	114
3/6/79	BCD	264.4070	119	49	109
3/7/79	BCD	266.4409	93	33	89
3/8/79	BCD	268.4748	106	36	62
3/9/79	BCD	270.5087	140	35	75

STUDENT ATTENDANCE - Tract B
Fourth Quarter

			Student Absences		
4/2/79	ABD	272.5426	134	16	21
4/3/79	ABD	274.5765	112	15	19
4/4/79	ABD	276.6104	94	21	14
4/5/79	ABD	278.6443	105	28	17
4/6/79	ABD	280.6782	118	27	16
4/9/79	ABD	282.7121	149	41	21
4/10/79	ABD	284.7460	137	23	19
4/11/79	ABD	286.7799	123	41	23
4/12/79	ABD	288.8138	158	37	22
4/23/79	ABD	290.8477	122	29	15
4/24/79	ABD	292.8816	113	25	18
4/25/79	ABD	294.9155	117	27	18
4/26/79	ABD	296.9494	116	19	12
4/27/79	ABD	298.9833	133	28	17
4/30/79	ABC	301.0172	110	29	11
5/1/79	ABC	303.0511	115	19	4
5/2/79	ABC	305.0850	130	24	5
5/3/79	ABC	307.1189	117	29	4
5/4/79	ABC	309.1528	132	33	9
5/7/79	ABC	311.1867	137	35	17
5/8/79	ABC	313.2206	112	24	7
5/9/79	ABC	315.2545	137	21	10
5/10/79	ABC	317.2884	111	27	11
5/11/79	ABC	319.3223	151	35	14
5/14/79	ABC	321.3562	131	25	14
5/15/79	ABC	323.3901	72	26	14
5/16/79	ABC	325.4240	105	19	13
5/17/79	ABC	327.4579	93	20	17
5/18/79	ABC	329.4918	451	24	19
5/21/79	BCD	331.5257	102	26	15
5/22/79	BCD	333.5596	85	20	16
5/23/79	BCD	335.5935	82	34	15
5/24/79	BCD	337.6274	108	36	20
5/25/79	BCD	339.6613	149	50	19
5/29/79	BCD	341.6952	116	38	19
5/30/79	BCD	343.7291	86	30	16
5/31/79	BCD	345.7630	94	30	14
6/1/79	BCD	347.7969	123	38	14
6/4/79	BCD	349.8308	122	33	16
6/5/79	BCD	351.8647	93	20	14
6/6/79	BCD	353.8986	114	22	9
6/7/79	BCD	355.9325	87	27	9
6/8/79	BCD	357.9664	150	29	18

3. Tract C

STUDENT ATTENDANCE - Tract C
First Quarter

Date	Tracts	Angular Measure	Student Absences: High School X	Intermediate Y	Elementary Z
7/31/86	ABC	0.0000	179	26	28
8/1/78	ABC	2.0339	136	21	10
8/2/78	ABC	4.0678	131	32	14
8/3/78	ABC	6.1017	131	22	17
8/4/78	ABC	8.1356	166	33	19
8/7/78	ABC	10.1695	158	35	27
8/8/78	ABC	12.2034	145	25	25
8/9/78	ABC	14.2373	154	37	21
8/10/78	ABC	16.2712	125	26	20
8/11/78	ABC	18.3051	161	42	31
8/14/78	ABC	20.3390	153	29	24
8/15/78	ABC	22.3729	143	29	21
8/16/78	ABC	24.4068	138	26	14
8/17/78	ABC	26.4407	145	28	15
8/18/78	ABC	28.4746	183	53	28
8/21/78	BCD	30.5085	182	30	30
8/22/78	BCD	32.5424	133	26	22
8/23/78	BCD	34.5763	132	20	15
8/24/78	BCD	36.6102	160	39	29
8/25/78	BCD	38.6441	198	31	30
8/28/78	BCD	40.6780	170	40	24
8/29/78	BCD	42.7119	130	25	16
8/30/78	BCD	44.7458	135	28	17
8/31/78	BCD	46.7797	143	31	16
9/1/78	BCD	48.8136	160	37	17
9/5/78	BCD	50.8475	157	41	23
9/6/78	BCD	52.8814	142	33	19
9/7/78	BCD	54.9153	106	35	18
9/8/78	BCD	56.9492	155	33	16
9/11/78	ACD	58.9831	141	31	9
9/12/78	ACD	61.0170	118	44	15
9/13/78	ACD	63.0509	119	43	16
9/14/78	ACD	65.0848	140	35	17
9/15/78	ACD	67.1187	132	41	14
9/18/78	ACD	69.1526	180	42	12
9/19/78	ACD	71.1865	105	42	14
9/20/78	ACD	73.2204	116	40	11
9/21/78	ACD	75.2543	112	40	11
9/22/78	ACD	77.2882	121	42	12
9/25/78	ACD	79.3221	128	30	22
9/26/78	ACD	81.3560	125	25	18
9/27/78	ACD	83.3899	133	31	17
9/28/78	ACD	85.4238	100	28	25
9/29/78	ACD	87.4577	132	32	28

STUDENT ATTENDANCE - Tract C
Second Quarter

Date	Tracts	Angular Measure	Student Absences: High School X	Intermediate Y	Elementary Z
10/23/78	ABC	89.4916	123	31	17
10/24/78	ABC	91.5255	145	28	14
10/25/78	ABC	93.5594	107	27	12
10/26/78	ABC	95.5933	108	28	15
10/27/78	ABC	97.6272	131	35	17
10/30/78	ABC	99.6611	124	24	17
10/31/78	ABC	101.6950	128	26	14
11/1/78	ABC	103.7289	121	27	18
11/2/78	ABC	105.7628	123	25	25
11/3/78	ABC	107.7967	124	31	30
11/6/78	ABC	109.8306	143	24	29
11/7/78	ABC	111.8645	105	24	22
11/8/78	ABC	113.8984	105	17	19
11/9/78	ABC	115.9323	97	21	27
11/10/78	ABC	117.9662	126	28	30
11/13/78	BCD	120.0001	142	35	26
11/14/78	BCD	122.0340	136	29	19
11/15/78	BCD	124.0679	141	26	14
11/16/78	BCD	126.1018	188	41	24
11/17/78	BCD	128.1357	178	55	25
11/20/78	BCD	130.1696	144	42	25
11/21/78	BCD	132.2035	153	50	20
11/22/78	BCD	134.2374	188	70	26
11/27/78	BCD	136.2713	169	59	28
11/28/78	BCD	138.3052	134	39	17
11/29/78	BCD	140.3391	108	30	20
11/30/78	BCD	142.3730	127	36	16
12/1/78	BCD	144.4069	234	88	44
12/4/78	BCD	146.4408	129	41	30
12/5/78	BCD	148.4747	94	39	22
12/6/78	BCD	150.5086	96	39	21
12/7/78	ACD	152.5425	489	202	70
12/8/78	ACD	154.5764	157	67	39
12/11/78	ACD	156.6103	120	41	21
12/12/78	ACD	158.6442	85	35	21
12/13/78	ACD	160.6781	108	31	35
12/14/78	ACD	162.7120	110	26	24
12/15/78	ACD	164.7459	128	35	32
12/18/78	ACD	166.7798	130	40	35
12/19/78	ACD	168.8137	104	39	21
12/20/78	ACD	170.8476	125	44	22
12/21/78	ACD	172.8815	223	48	21
1/2/79	ACD	174.9154	429	179	78
1/3/79	ACD	176.9493	266	130	48
1/4/79	ACD	178.9832	143	64	31
1/5/79	ACD	181.0171	138	56	32

STUDENT ATTENDANCE - Tract C
Third Quarter

			Student Absences		
Date	Tracts	Angular Measure	High School X	Intermediate Y	Elementary Z
1/29/79	ABC	183.0510	122	15	20
1/30/79	ABC	185.0849	94	29	16
1/31/79	ABC	187.1188	1301	493	435
2/1/79	ABC	189.1527	166	493	435
2/2/79	ABC	191.1866	107	49	435
2/5/79	ABC	193.2205	168	43	17
2/6/79	ABC	195.2544	104	28	16
2/7/79	ABC	197.2883	111	20	16
2/8/79	ABC	199.3222	96	29	14
2/9/79	ABC	201.3561	115	27	16
2/13/79	ABC	203.3900	114	27	18
2/14/79	ABC	205.4239	98	22	13
2/15/79	ABC	207.4578	103	26	18
2/16/79	ABC	209.4917	108	22	22
2/19/79	BCD	211.5256	181	49	26
2/20/79	BCD	213.5595	129	41	17
2/21/79	BCD	215.5934	209	44	20
2/22/79	BCD	217.6273	88	42	17
2/23/79	BCD	219.6612	111	42	16
2/26/79	BCD	221.6951	112	45	21
2/27/79	BCD	223.7290	118	48	13
2/28/79	BCD	225.7629	112	48	17
3/1/79	BCD	227.7968	126	39	15
3/2/79	BCD	229.8307	139	47	48
3/5/79	BCD	231.8646	188	54	114
3/6/79	BCD	233.8985	119	49	109
3/7/79	BCD	235.9324	93	33	89
3/8/79	BCD	237.9663	106	36	62
3/9/79	BCD	240.0002	140	35	75
3/12/79	ACD	242.0341	128	31	32
3/13/79	ACD	244.0680	111	36	30
3/14/79	ACD	246.1019	178	67	32
3/15/79	ACD	248.1358	160	54	38
3/16/79	ACD	250.1697	118	45	29
3/19/79	ACD	252.2036	150	32	34
3/20/79	ACD	254.2375	134	45	26
3/21/79	ACD	256.2714	134	42	22
3/22/79	ACD	258.3053	134	37	22
3/23/79	ACD	260.3392	165	43	22
3/26/79	ACD	262.3731	137	46	25
3/27/79	ACD	264.4070	118	43	16
3/28/79	ACD	266.4409	105	33	16
3/29/79	ACD	268.4748	83	33	13
3/30/79	ACD	270.5087	91	42	19

STUDENT ATTENDANCE - Tract C
Fourth Quarter

Date	Tracts	Angular Measure	Student Absences: High School X	Intermediate Y	Elementary Z
4/30/79	ABC	272.5426	110	29	11
5/1/79	ABC	274.5765	115	19	4
5/2/79	ABC	276.6104	130	24	5
5/3/79	ABC	278.6443	117	29	4
5/4/79	ABC	280.6782	132	33	9
5/7/79	ABC	282.7121	137	35	17
5/8/79	ABC	284.7460	112	24	7
5/9/79	ABC	286.7799	137	21	10
5/10/79	ABC	288.8138	111	27	11
5/11/79	ABC	290.8477	151	35	14
5/14/79	ABC	292.8816	131	25	14
5/15/79	ABC	294.9155	72	26	14
5/16/79	ABC	296.9494	105	19	13
5/17/79	ABC	298.9833	93	20	17
5/18/79	ABC	301.0172	451	24	19
5/21/79	BCD	303.0511	102	26	15
5/22/79	BCD	305.0850	85	20	16
5/23/79	BCD	307.1189	82	34	15
5/24/79	BCD	309.1528	108	36	20
5/25/79	BCD	311.1867	149	50	19
5/29/79	BCD	313.2206	116	38	19
5/30/79	BCD	315.2545	86	30	16
5/31/79	BCD	317.2884	94	30	14
6/1/79	BCD	319.3223	123	38	14
6/4/79	BCD	321.3562	122	33	16
6/5/79	BCD	323.3901	93	20	14
6/6/79	BCD	325.4240	114	22	9
6/7/79	BCD	327.4579	87	27	9
6/8/79	BCD	329.4918	150	29	18
6/11/79	ACD	331.5257	128	39	30
6/12/79	ACD	333.5596	106	31	21
6/13/79	ACD	335.5935	93	36	14
6/14/79	ACD	337.6274	166	23	11
6/15/79	ACD	339.6613	423	42	15
6/18/79	ACD	341.6952	144	39	17
6/19/79	ACD	343.7291	108	32	14
6/20/79	ACD	345.7630	93	23	8
6/21/79	ACD	347.7969	80	29	16
6/22/79	ACD	349.8308	99	33	15
6/25/79	ACD	351.8647	122	38	25
6/26/79	ACD	353.8986	124	23	13
6/27/79	ACD	355.9325	216	42	33
6/28/79	ACD	357.9664	0	0	0

4. Tract D

STUDENT ATTENDANCE - Tract D
First Quarter

Date	Tracts	Angular Measure	Student Absences: High School X	Intermediate Y	Elementary Z
7/11/78	ABD	0.0000	0	23	21
7/12/78	ABD	2.0339	33	25	19
7/13/78	ABD	4.0678	156	18	17
7/14/78	ABD	6.1017	162	28	20
7/15/78	ABD	8.1356	201	40	29
7/17/78	ABD	10.1695	180	47	15
7/18/78	ABD	12.2034	168	37	12
7/19/78	ABD	14.2373	130	32	15
7/20/78	ABD	16.2712	119	43	19
7/21/78	ABD	18.3051	149	50	28
7/24/78	ABD	20.3390	135	33	25
7/25/78	ABD	22.3729	127	39	19
7/26/78	ABD	24.4068	123	34	24
7/27/78	ABD	26.4407	114	31	13
7/28/78	ABD	28.4746	176	43	29
8/21/78	BCD	30.5085	182	30	30
8/22/78	BCD	32.5424	133	26	22
8/23/78	BCD	34.5763	132	20	15
8/24/78	BCD	36.6102	160	39	29
8/25/78	BCD	38.6441	198	31	30
8/28/78	BCD	40.6780	170	40	24
8/29/78	BCD	42.7119	130	25	16
8/30/78	BCD	44.7458	135	28	17
8/31/78	BCD	46.7797	143	31	16
9/1/78	BCD	48.8136	160	37	17
9/5/78	BCD	50.8475	157	41	23
9/6/78	BCD	52.8814	142	33	19
9/7/78	BCD	54.9153	106	35	18
9/8/78	BCD	56.9492	155	33	16
9/11/78	ACD	58.9831	141	31	9
9/12/78	ACD	61.0170	118	44	15
9/13/78	ACD	63.0509	119	43	16
9/14/78	ACD	65.0848	140	35	17
9/15/78	ACD	67.1187	132	41	14
9/18/78	ACD	69.1526	180	42	12
9/19/78	ACD	71.1865	105	42	14
9/20/78	ACD	73.2204	116	40	11
9/21/78	ACD	75.2543	112	40	11
9/22/78	ACD	77.2882	121	42	12
9/25/78	ACD	79.3221	128	30	22
9/26/78	ACD	81.3560	125	25	18
9/27/78	ACD	83.3899	133	31	17
9/28/78	ACD	85.4238	100	28	25
9/29/78	ACD	87.4577	132	32	28

STUDENT ATTENDANCE - Tract D
Second Quarter

			Student Absences		
Date	Tracts	Angular Measure	High School X	Intermediate Y	Elementary Z
10/2/78	ABD	89.4916	135	41	26
10/3/78	ABD	91.5255	121	40	18
10/4/78	ABD	93.5594	115	35	23
10/5/78	ABD	95.5933	120	37	23
10/6/78	ABD	97.6272	144	38	20
10/9/78	Free	99.6611	144	43	25
10/10/78	ABD	101.6950	148	30	18
10/11/78	ABD	103.7289	109	26	26
10/12/78	ABD	105.7628	103	31	16
10/13/78	ABD	107.7967	250	35	25
10/16/78	ABD	109.8306	152	39	16
10/17/78	ABD	111.8645	104	25	17
10/18/78	ABD	113.8984	110	29	17
10/19/78	ABD	115.9323	136	32	19
10/20/78	Free	117.9662	144	43	25
11/13/78	BCD	120.0001	142	35	26
11/14/78	BCD	122.0340	136	29	19
11/15/78	BCD	124.0679	141	26	14
11/16/78	BCD	126.1018	188	41	24
11/17/78	BCD	128.1357	178	55	25
11/20/78	BCD	130.1696	144	42	25
11/21/78	BCD	132.2035	153	50	20
11/22/78	BCD	134.2374	188	70	26
11/27/78	BCD	136.2713	169	59	28
11/28/78	BCD	138.3052	134	39	17
11/29/78	BCD	140.3391	108	30	20
11/30/78	BCD	142.3730	127	36	16
12/1/78	BCD	144.4069	234	88	44
12/4/78	BCD	146.4408	129	41	30
12/5/78	BCD	148.4747	94	39	22
12/6/78	BCD	150.5086	96	39	21
12/7/78	ACD	152.5425	489	202	70
12/8/78	ACD	154.5764	157	67	39
12/11/78	ACD	156.6103	120	41	21
12/12/78	ACD	158.6442	85	35	21
12/13/78	ACD	160.6781	108	31	35
12/14/78	ACD	162.7120	110	26	24
12/15/78	ACD	164.7459	128	35	32
12/18/78	ACD	166.7798	130	40	35
12/19/78	ACD	168.8137	104	39	21
12/20/78	ACD	170.8476	125	44	22
12/21/78	ACD	172.8815	223	48	21
1/2/79	ACD	174.9154	429	179	78
1/3/79	ACD	176.9493	266	130	48
1/4/79	ACD	178.9832	143	64	31
1/5/79	ACD	181.0171	138	56	32

STUDENT ATTENDANCE - Tract D
Third Quarter

Date	Tracts	Angular Measure	Student Absences High School X	Intermediate Y	Elementary Z
1/8/79	ABD	183.0510	161	40	29
1/9/79	ABD	185.0849	138	30	31
1/10/79	ABD	187.1188	108	28	26
1/11/79	ABD	189.1527	127	39	28
1/12/79	ABD	191.1866	144	43	27
1/15/79	Closed	193.2205	144	43	25
1/16/79	Closed	195.2544	144	43	25
1/17/79	ABD	197.2883	737	308	111
1/18/79	ABD	199.3222	346	179	74
1/19/79	ABD	201.3561	519	234	109
1/22/79	ABD	203.3900	123	26	14
1/23/79	ABD	205.4239	113	19	13
1/24/79	Closed	207.4578	144	43	25
1/25/79	Closed	209.4917	144	43	25
1/26/79	ABD	211.5256	265	99	31
2/19/79	BCD	213.5595	181	49	26
2/20/79	BCD	215.5934	129	41	17
2/21/79	BCD	217.6273	209	44	20
2/22/79	BCD	219.6612	88	42	17
2/23/79	BCD	221.6951	111	42	16
2/26/79	BCD	223.7290	112	45	21
2/27/79	BCD	225.7629	118	48	13
2/28/79	BCD	227.7968	112	48	17
3/1/79	BCD	229.8307	126	39	15
3/2/79	BCD	231.8646	139	47	48
3/5/79	BCD	233.8985	188	54	114
3/6/79	BCD	235.9324	119	49	109
3/7/79	BCD	237.9663	93	33	89
3/8/79	BCD	240.0002	106	36	62
3/9/79	BCD	242.0341	140	35	75
3/12/79	ACD	244.0680	128	31	32
3/13/79	ACD	246.1019	111	36	30
3/14/79	ACD	248.1358	178	67	32
3/15/79	ACD	250.1697	160	54	38
3/16/79	ACD	252.2036	118	45	29
3/19/79	ACD	254.2375	150	32	34
3/20/79	ACD	256.2714	134	45	26
3/21/79	ACD	258.3053	134	42	22
3/22/79	ACD	260.3392	134	37	22
3/23/79	ACD	262.3731	165	43	22
3/26/79	ACD	264.4070	137	46	25
3/27/79	ACD	266.4409	118	43	16
3/28/79	ACD	268.4748	105	33	16
3/29/79	ACD	270.5087	83	33	13

STUDENT ATTENDANCE - Tract D
Fourth Quarter

Date	Tracts	Angular Measure	Student Absences: High School X	Intermediate Y	Elementary Z
3/30/79	ACD	272.5426	91	42	19
4/2/79	ABD	274.5765	134	16	21
4/3/79	ABD	276.6104	112	15	19
4/4/79	ABD	278.6443	94	21	14
4/5/79	ABD	280.6782	105	28	17
4/6/79	ABD	282.7121	118	27	16
4/9/79	ABD	284.7460	149	41	21
4/10/79	ABD	286.7799	137	23	19
4/11/79	ABD	288.8138	123	41	23
4/12/79	ABD	290.8477	158	37	22
4/23/79	ABD	292.8816	122	29	15
4/24/79	ABD	294.9155	113	25	18
4/25/79	ABD	296.9494	117	27	18
4/26/79	ABD	298.9833	116	19	12
4/27/79	ABD	301.0172	133	28	17
5/21/79	BCD	303.0511	102	26	15
5/22/79	BCD	305.0850	85	20	16
5/23/79	BCD	307.1189	82	34	15
5/24/79	BCD	309.1528	108	36	20
5/25/79	BCD	311.1867	149	50	19
5/29/79	BCD	313.2206	116	38	19
5/30/79	BCD	315.2545	86	30	16
5/31/79	BCD	317.2884	94	30	14
6/1/79	BCD	319.3223	123	38	14
6/4/79	BCD	321.3562	122	33	16
6/5/79	BCD	323.3901	93	20	14
6/6/79	BCD	325.4240	114	22	9
6/7/79	BCD	327.4579	87	27	9
6/8/79	BCD	329.4918	150	29	18
6/11/79	ACD	331.5257	128	39	30
6/12/79	ACD	333.5596	106	31	21
6/13/79	ACD	335.5935	93	36	14
6/14/79	ACD	337.6274	166	23	11
6/15/79	ACD	339.6613	423	42	15
6/18/79	ACD	341.6952	144	39	17
6/19/79	ACD	343.7291	108	32	14
6/20/79	ACD	345.7630	93	23	8
6/21/79	ACD	347.7969	80	29	16
6/22/79	ACD	349.8308	99	33	15
6/25/79	ACD	351.8647	122	38	25
6/26/79	ACD	353.8986	124	23	13
6/27/79	ACD	355.9325	216	42	33
6/28/79	ACD	357.9664	0	0	0

APPENDIX D

Letter of Authorization from

Carmel High School for Boys

ORDER OF CARMELITES - CATHOLIC SECONDARY SCHOOL EDUCATION OFFICE
St. Matthew's Parish, 1555 Glen Ellyn Road, Glendale Heights, Illinois 60139, 312/469-6300

23 March 1987

Mr. Robert E. Watson
131 School Street
Grayslake, Illinois 60030

Dear Mr. Watson:

During my tenure as principal of Carmel High School for Boys, in Mundelein, Illinois, I gave you verbal approval to use the name of Carmel High School for Boys in your dissertation for Loyola University of Chicago. It was my understanding that Carmel's name would be used in the 1982 Tuition Study for Carmel High School for boys.

With this letter I give you written approval to use the name of Carmel High School for Boys in your dissertation for Loyola University.

Sincerely,

Gregory L. Klein

Fr. Gregory L. Klein, O. Carm.
Superintendent

APPROVAL SHEET

The dissertation submitted by Robert E. Watson has been read and approved by the following committee:

Dr. Jack A. Kavanagh, Director
Professor, Education and
Chairman, Department of Foundations of Education, Loyola

Dr. Todd J. Hoover
Associate Professor, Education, Loyola

Dr. Steven I. Miller
Professor, Education, Loyola

Dr. Ronald R. Morgan
Associate Professor, Education, Loyola

The final copies have been examined by the director of the dissertation and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the dissertation is now given final approval by the Committee with reference to content and form.

The dissertation is therefore accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

4/14/87
Date

Jack A. Kavanagh
Director's Signature