DETECTION OF TRENDS
IN
WATER QUALITY PARAMETERS

By

Robert Hale Montgomery

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Resource Development

1981

ABSTRACT

DETECTION OF TRENDS
IN
WATER QUALITY PARAMETERS

By

Robert Hale Montgomery

With the advent of standards and criteria for water quality parameters, there has been an increasing concern about the changes of these parameters over time. Thus, sound statistical methods dealing with the detection of trends in water quality parameters are needed.

The method presented provides: 1) formulation of a problem (hypothesis), 2) selection of water quality parameter(s) and data, 3) data analysis techniques, and 4) statistical tests for detection of trends. A review of water quality parameters and certain topics in statistics is also provided. The techniques are explained in a non-statistical manner to allow usage for those not well versed in statistical theory.

ACKNOWLEDGMENTS

I would like to deeply express my thanks to the members of my committee, Dr. Kenneth Reckhow, Dr. Stanley Zarnoch, and Dr. Eckhart Dersch, for their assistance, advice, and criticism. Dr. Reckhow, my major professor, served as the cornerstone of my graduate career. His knowledge and professional attributes set the standards that I will attempt to uphold. Dr. Zarnoch's knowledge and friendship played a vital role in allowing me to grasp the numerous statistical concepts. Dr. Dersch provided the impetus to show that research must be applied to management needs in order to be valuable.

Special thanks go to Janine Jurack for the long and tedious job of typing this document and to Paul Schneider for the preparation of figures. Also, I would like to thank my graduate student cohorts V. David Lee, Micheal Beaulac, Jonathon Simpson, and Ralph Ancil for their interactions we had at Michigan State University.

Finally, to my present and future family, I dedicate this thesis to, for without them none of this would have been possible.

This research was supported by a grant from the National Oceanic and Atmospheric Administration, grant number 03-78-B01-109.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter I

INTRODUCTION

Study Objectives

The earth is endowed with numerous natural resources available for man's use. However, long-term maintenance of the environment is a prerequisite to the continuing usage of natural resources. Thus, society is responsible for the protection and propagation of the environment for future generations. As society becomes more technologically oriented, the use of qualitative knowledge in decision planning can be balanced and supported by quantitative information. The use of quantitative techniques in water quality planning, for example, provides the necessary "hard" evidence that planners and managers have lacked in recent years. An especially weak area of water quality management is quantitative methods dealing with trend detection in water quality parameters. Trend detection techniques are vital in the development of planning, policy and management of water resource systems.

The objective of this study is to provide sound statistical methods dealing with the detection of trends in water quality parameters. The trend detection method is based on strong statistical techniques, but attempts to explain the techniques in a non-statistical manner. The intention of the study is to make the trend detection method available for use by persons not well versed in statistical theory.

The remaining sections of Chapter 1 describe water quality management, the application and usefulness of trend detection techniques to water quality management, and provide a general review of trend theory.

Chapter 2 discusses water quality parameters. The emphasis is on data availability and on selecting the appropriate parameter for a given management decision or concern. Chapter 3 provides a general review of the field of statistics that apply to the application and use of trend detection techniques, particularly those dealing with time series data and analysis. The trend detection method is explained in Chapter 4 and all phases necessary for its use are discussed (i.e., hypothesis testing, data preparation, data analysis, statistical tests, and time series modeling). Two applications of the trend detection method are at the end of Chapter 4. The conclusions and recommendations on the trend detection method are discussed in Chapter 5. There are also two Appendices: A) containing water quality parameter data sets, and B) containing statistical tables needed for the statistical methods.

Water Quality Management

As the demand for water increased over the years, there has been a profound expansion of the goals among water quality management agencies. First, it was protection of the public health. Later, as both the amount and variety of waste discharges grew, concern for the protection of other beneficial uses was expressed. Recently, esthetic or social goals of water quality have been added to the growing list of objectives for water quality management. The result is an emphasis on resource oriented management instead of management oriented to use. Thus, the concern has shifted from pollutants to the water resource itself. In this context, resource protection becomes the object of public policy, and the various beneficial uses appear as subsystems which must be

managed in a coordinated and integrated fashion to achieve resource policy objectives (McGauhey, 1968).

Water quality management programs, or agencies, have two broad objectives: prevention and abatement. The prevention objective is related to maintaining the existing "good" water quality, while abatement refers to reducing or moderating existing pollution conditions. These two objectives can be subdivided into seven basic activities, which are (Ward, 1973):

1. Planning
2. Research
3. Aid Programs
4. Technical Assistance
5. Regulation
6. Legal Enforcement
7. Data Collection, Processing, and Dissemination

Planning, research and aid programs are generally classified as prevention, while technical assistance, regulation and legal enforcement are classified as abatement. Data collection, processing and dissemination provide support to the first six activities. Figure 1 illustrates this type of water quality management structure.

The planning activity in water quality management has received increased emphasis in recent years, at both the state and federal level. The state level emphasis has been on program planning, while at the federal level, emphasis has primarily been on project planning. Project and program planning are similar in that a plan for future action is being developed, but the plans differ in the amount of detail required. Program planning deals with broad definitions in program and

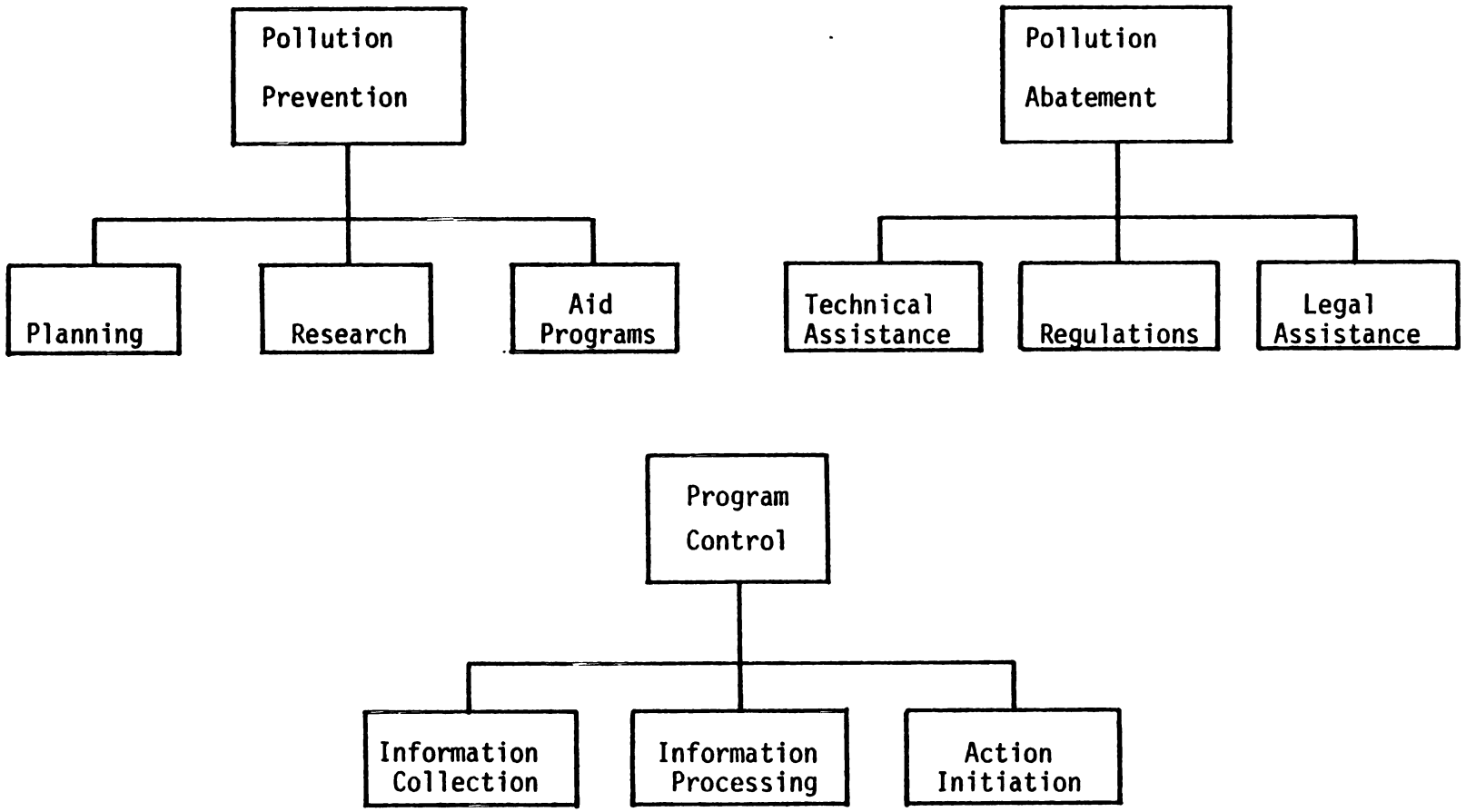


Figure 1. Water quality management structure as suggested by Ward (1973).

work plans. Consequently, general overall trends in water quality best serve the program planning function at the state level. Project planning entails specific planning related to detailed evaluation of future water quality effects as a result of specific causes (e.g., waste treatment plant location, number of pollution sources). Thus, the change in a specific parameter over time is the concern in project planning. The detection of trends in water quality parameters is extremely important for both program and project planning.

The general objective of research is to conduct studies and investigations on issues relating to water quality. Included in the research objective is the implementation of research findings. All too often, research results are not put to an effective use. It is in the research phase that the detection of trends in water quality parameters is carried out.

The activity denoted under aid programs is: 1) accepting and supervising of loans and grants, 2) processing applications and administration of loans or grants, and 3) certification of the need for money for various projects. The information used to verify a need for a certain project can often deal with how a parameter(s) has or will change when the project is implemented. Thus, past trends can provide the impetus to invoke a project and future changes can be used to evaluate the effectiveness of the project.

The objective of technical assistance is to advise, consult, and cooperate on technical matters concerning water quality. Technical assistance may be given to other agencies (federal, state, local) or to private enterprises. In order for technical assistance to be given, valid information about water quality parameters is needed. A major

form of information disseminated concerns the changes in water quality parameters over time and methods of determining (quantifying) the change.

The objective of water quality regulations is the maintenance of water quality. This objective has traditionally been the most important to public agencies given the money and manpower devoted to its accomplishment. Regulation entails the development of water quality standards and the routine surveillance of water quality to insure compliance with the established standards. Regulation also involves developing pollution abatement requirements and establishing procedures to assure that future actions do not cause violations of standards. Important uses of trend detection for regulation are: 1) to evaluate the progress of abatement programs toward meeting standards, and 2) to identify emerging water quality problem areas so that prevention can be affected, before abatement programs are required.

The object of legal enforcement is to enforce water quality standards when standards cannot be maintained by persuasion. The need for quantitative factual information on water quality parameters for legal evidence is paramount. Until recently, environmental lawyers had only qualitative information to support their position. Subjective evaluations, however, may be easily contradicted by the opposing side's experts. With the use of statistical techniques to provide quantitative information on water quality trends the courts should have a clear path to follow (i.e., what the data dictate).

Successful accomplishment of the six objectives outlined depends upon good data collection, processing and dissemination. The first step toward good water quality data is to select a representative

sample. Following proper sample preparation, handling, and laboratory procedures, with quality control, the results of the analysis are processed. Data processing involves screening, verification, interpretation, indexing, storage and retrieval of the data. Data dissemination is accomplished through the generation of reports. The idea of proper sampling design cannot be overstressed. For a detailed review of sampling design, consult Cochran (1977), Beckers et al. (1973), Sanders et al. (1976), and Ward (1973).

All phases of planning, development and operation of a water quality management system require data to quantify existing states of the system, to forecast future changes and trends, and to predict the response of the system to intervention. To evaluate alternative ways of combating perceived conditions of water quality deterioration requires objective measurements that: 1) describe the observed conditions, and 2) can be related to appropriate measures of water resource management. A number of parameters have been observed which might provide objective criteria to study trends in water systems. An important concern with water quality parameter selection for trend detection analysis is whether the parameter provides the best measure of the perceived condition or qualities of concern. The kind of data required and their spatial and temporal resolution depend upon the nature of the problem and type of intervention under consideration. A partial list of needs and uses of water quality data is shown in Figure 2 (Sherwani and Moreau, 1975).

Relatively few studies have been conducted on trends in the quality of the nation's waters. The emphasis of past water quality trend analysis has been primarily with river water quality (Wolman, 1971; EPA, 1974;

Public Interest:	Public Health
	Aesthetics
	Nuisance
	Ecological balance
	Conservation
	Natural state preservation
	Recreation
Planning:	Water and related land use planning
	Economic planning
	Urban planning
Regulation and Control:	Identification of sources
	Fate of pollutants
	Description of present state of quality
	Prediction of water quality
	Evaluation of trends
	Available control strategies and tactics
	Measurement of progress in pollution abatement
	Episodic effects
	Non-degradation policy
	Research
	Legislation
	Public hearings
	User-oriented reports

Figure 2. Water quality data uses (Sherwani and Moreau, 1975)

Lettenmaier, 1977), and only a few with lake trend analysis (Rockwell et al., 1979; Chapra, 1980; Dobson, 1980). The reason for the lack of water quality trend studies is due to the number of disabilities that interfere with a truly adequate statistical analysis of time series data. The major disabilities that affect the use of proper statistical procedures are (Wolman, 1971):

1. Short records of water quality data.
2. Techniques of observation and analysis have changed over the years.
3. Changing location or frequency of observations.
4. Correlations that relate specific variables to limnologic and hydrologic behavior are rarely available.
5. Natural background often hides water quality trends.
6. Explanation of trends requires a knowledge of the economy and land use in the area.

Most measures of national growth suggest that demands on water resources are increasing. It is these increasing demands for water resources that usually result in deterioration of water quality. Thus, even without proof, one might assume that water quality conditions are getting worse. The problem is to precisely determine the relationships between the pressures posed by society and the responses of water quality parameters. The main interest for the management of a water resource is not its present state, but how the state of the system has changed and will change in the future. Thus, the use of trend detection techniques to quantify the change in water quality parameters over time is critical.

By providing information on water quality trends, one can determine whether a given water resource is improving, deteriorating or stationary under the current conditions. Therefore, management and policy decisions can be developed, with a degree of certainty, based on the changing system. For example, if a lake with excellent water quality was exhibiting a declining trend, emphasis could be placed on an effort to control the situation. Conversely, a lake exhibiting poor water quality, but with an improving trend is evidence of good recent management. If planning is to be based on sound information and the effectiveness of management alternatives are to be judged, then observational tools (e.g., trend detection methods) must be designed and implemented in water quality management.

Trend Theory

Lake water quality parameters are subject to continual change over time. The inputs to the lake, outputs from the lake, and the lake itself are variable. Thus, lake water quality data arise from a nonstationary process. To provide a representation of the changing conditions over a period of time, a time sequence of measurements on water quality parameters is needed. The data exist in a time series, which preserves the order of occurrence. The measurements may be taken at approximately regular intervals or values may be obtained as averages over fixed periods. The measurements should be taken over a sufficiently long period of time using similar methods. Usually most existing water quality data are: 1) not measured over sufficient periods of time, 2) measured at different times and time intervals, and 3) variable,

due to changes in sampling methods and laboratory techniques. Thus, inherent problems exist in trend detection of water quality parameters.

In order for a trend to appear, the data from a time series must display a non-random pattern. In a non-random series the observations cannot be explained by purely random variation at a given level of significance. Non-randomness can arise from: 1) presence of a trend, 2) cyclicality, and 3) serial correlation. A series exhibits a trend if the values of its members show a tendency to increase or decrease. A cyclic, or periodic, series occurs when the values rise or fall in a regular fashion. Serial correlation occurs when a value is dependent upon previous values.

A time series can be considered to be made up of two parts: 1) a systematic or deterministic component changing over time in a regular and predictable way, and 2) a random component superimposed on the regular part (Sherwani and Moreau, 1975). The changes in the deterministic part are those attributable to fundamental changes in the nature of the process itself. The random component represents short-term fluctuations due to transitory factors. Often, a trend may be obscured by superimposed random variation. Also, it is difficult to detect small trends in a short series.

A general time series equation is

$$x_t = \mu_t + e_t \qquad (1)$$

where: x_t = time series of data

μ_t = the trend

e_t = random variation (noise term)

Two of the most common trends are step and linear trends (Lettenmaier, 1977). A step trend is an instantaneous jump in the mean level at some point in time, i.e.,

$$x_t = \mu_1 + [\mu_2 - \mu_1]_T + e_t \tag{2}$$

where: μ_1 = true mean of first part of record

μ_2 = true mean of second part of record

$[L]_T$ = function with value zero for $t \leq T$ and L for $t > T$, where t is start of step trend

A linear trend is simply a uniform increase in the mean level:

$$x_t = \mu + t/T\Delta\mu + e_t \tag{3}$$

where: t = time length of each time unit

T = total time length of time series

For large sample sizes, or when the variance of e_t is known, the power of the classical two-tailed t-test (Breiman, 1973) against step and linear trends, where the noise terms e_t are statistically independent and normally distributed with mean zero and variance σ_e^2, is

$$1 - \beta = F(N_T - W_{1-\alpha/2}) \tag{4}$$

where: F = cumulative distribution function of a standard normal probability distribution

$W_{1-\alpha/2}$ = standard normal quartile at probability level $1 - \alpha/2$

$$N_T = \frac{T_r \; n}{2\sigma} \quad \text{(for a step trend)} \tag{5}$$

$$N_T = \frac{T_r \; n(n+1)\;(n-1)}{n\sqrt{12\sigma_e}} \approx \frac{T_r \; n}{12\sigma} \quad \text{(for a linear trend)} \tag{6}$$

$$T_r = \text{trend magnitude} \tag{7}$$

$|\mu_1 - \mu_2|$ for step trend

$\Delta\mu$ for linear trend

Trend equations can be expanded to include seasonal movement or modified to a multiplicative system (log or polynomial) instead of the additive system, as in the step and linear trend (consult Lettenmaier (1977) and Sherwani and Moreau (1975) for more information).

A step trend is a permanent increase or decrease in the value of the water quality parameter. This may, for example, result from a permanent change in land use or the construction of a waste treatment plant. A linear trend is a steady upward (or downward) movement in a water quality parameter. This may, for example, result from a change in agricultural or urban runoff (i.e., nutrient loads).

The classical hypothesis testing framework may be used to determine the existence or nonexistence of trends in water quality parameters. The null hypothesis, H_0, is that no trend exists, while the alternative hypothesis, H_1, is that a trend does exist. The choice between H_0 and H_1 is made on the basis of a test statistic, computed from the data. The test statistic is compared to a probability outcome distribution, and H_0 is either accepted or rejected at a given level of confidence.

CHAPTER II

WATER QUALITY

The quality of water is affected by many factors, some of which are 1) nutrients, 2) organic material, 3) toxic chemicals, 4) dissolved and suspended solids, 5) dissolved oxygen, and 6) pH. The number of factors required to specify water quality is limitless, changing both the physical, chemical and biological environment and the socio-economic activities within the water system. Any consideration of water quality should also include water quantity, because quality and quantity of water are interrelated. A knowledge of ambient water quality is required to: 1) ascertain reliable and accurate information on the current status of water quality required for planning of water resources, 2) provide reliable data to assess long-term trends and overall changes in water quality, 3) determine the degree to which water quality is improved as a result of pollution abatement measures, 4) indicate problem areas requiring corrective actions, and 5) determine the extent of compliance and non-compliance with water quality standards (Sherwani and Moreau, 1975).

Water Quality Parameters

Water quality management seeks to insure levels of pollutants in surface waters which will not interfere with desired uses. Thus, water quality depends directly on the use of the water. There are numerous physical, chemical, and biological parameters which are significant in

determining water quality. Figure 3 lists some parameters and the role they play for various water uses.

Water quality parameters may be general or specific purpose parameters. The general purpose parameters provide basic information about water quality. Some general purpose parameters are: 1) dissolved oxygen, 2) pH, 3) temperature, and 4) conductivity. Special purpose parameters relate to definite processes, activities, or pollution sources affecting water quality. Examples of specific purpose parameters are: 1) BOD, biological oxygen demand, 2) fecal coliforms, and 3) toxic chemicals.

There are numerous parameters that affect water quality and discussion of all of them is not possible here. Therefore, only the major parameters are discussed below. For more information on these parameters and others, consult: 1) EPA, (1976), 2) Wetzel, (1975), 3) Hutchinson (1957), 4) McNeely (1979).

Alkalinity refers to the quantity and types of compounds, mainly inorganic carbon, which collectively shift the pH to the alkaline side of neutrality. The forms of inorganic carbon in fresh waters are: 1) free carbon dioxide (CO_2), 2) carbonic acid (H_2CO_3), 3) calcium bicarbonate (HCO_3^-), and 4) carbonate ($CO_3^=$). These four species of inorganic carbon form an equilibrium between each other based primarily on pH. As the pH increases, the dominant species moves from free carbon dioxide to carbonate ions. Alkalinity is a measure of the buffering capacity of water. Since pH has a direct effect on organisms and an indirect effect on the toxicity of certain pollutants in water, the buffering capacity of water is very important to water quality. Long term trends in alkalinity are extremely important in those areas

	Parameter	Use
1.	Temperature	Aquatic life, Industrial use, assimilative capacity, recreation
2.	Turbidity	Drinking water, recreation, industrial use
3.	pH	Industrial use, aquatic life, recreation
4.	Dissolved Oxygen	Aquatic life, aesthetics, industrial use, assimilative capacity
5.	BOD	Food and beverage industries, recreation, assimilative capacity
6.	Suspended Solids	Aesthetics, photosynthesis, reservoir capacity depletion, hydroelectric power generation, navigation
7.	Total Dissolved Solids	Irrigation, water supply, industrial use
8.	Coliform	Direct-contact water-based recreation, water supply, food and beverage industries, irrigation
9.	Nutrients	Eutrophication, aesthetic degradation, secondary effects on aquatic life
10.	Organics	Water supply, industrial use, aquatic life
11.	Heavy Metals	Water supply, aquatic life
12.	Radioactivity	Water supply
13.	Oil	Recreation, industrial use
14.	Color	Aesthetics, water supply, recreation, industrial use
15.	Conductivity	General parameter of water quality
16.	Chlorophyll	Biological activity

Figure 3. Important water quality parameters and related uses (modified from Sherwani and Moreau (1975))

receiving acid rain. Acid rain tends to lower the natural pH of waters with low buffering capacity. Thus, monitoring of alkalinity changes may serve as an indicator of deteriorating water quality due to acid rain. Another important parameter for trend analysis in water quality is the hypolimnetic inorganic carbon accumulation. The accumulation of inorganic carbon in the hypolimnion of a lake can be used to estimate indirectly the organic production in a lake's epilimnion and metalimnion (for more information consult Wetzel, 1975).

The use of biological organisms for evaluating water quality is an old concept that only recently has received attention. The response of sensitive biological indicator organisms to an environmental stress provides an early indicator of changing water quality. Biological organisms are: 1) uniquely sensitive to multiple environmental stresses operating consecutively or simultaneously, and 2) integrate the effects of environmental stresses over time. Thus, the ability to selectively accumulate, biomagnify, and show the synergistic effects to exposure from environmental stresses gives bioindicators useful properties for determining trends in water quality. Figure 4 lists the major communities of biological organisms and the parameters used to describe them for water quality evaluation. One problem with the use of biological organisms is the difficulty in obtaining quantitative numbers and levels on the organisms. The use of biological organisms for detection of trends in water quality has tremendous possibilities, especially when dealing with water systems affected by numerous stresses. For more information on biological organisms and monitoring consult 1) Worf, (1980), and 2) Cairns, et al. (1977).

Organic carbon is of two forms, particulate or dissolved, and

Community	Parameter
Plankton	Counts and identification Chlorophyll a Biomass as ashfree weight
Periphyton	Counts and identification Chlorophyll a Biomass as ashfree weight
Macrophyton	Areal coverage Identification Biomass as ashfree weight
Macroinvertebrate	Counts and identification Biomass as ashfree weight Flesh tainting Toxic substances in tissue
Fish	Toxic substances in tissue Counts and identification Biomass as wet weight Condition factor Flesh tainting Age and growth

Figure 4. Major communities of biological organisms used in biological monitoring (Weber, 1980)

results from outside of the lake (allochthonous) or from within the lake (autochthonous). The dissolved to particulate organic carbon ratio approximates 6:1 to 10:1 in most natural water bodies (Wetzel and Rich, 1973). The biochemical trnasformations of particulate and dissolved organic matter by microbial metabolism are fundamental to the dynamics of nutrient cycling and energy flux within aquatic ecosystems. Thus, trends in organic carbon may provide valuable information on possible changes in water quality.

Iron is an essential trace element required by both plants and animals. In some marl lakes where iron is precipitated by the highly alkaline conditions, iron may be the limiting factor for algal growth. Iron is also a vital oxygen transport element in the blood of all vertebrate and some invertebrate animals. Iron exists in solution in water as either the ferrous (Fe^{++}) or ferric (Fe^{+++}) state. Amounts of iron in solution in natural water, and rate of oxidation of Fe^{++} to Fe^{+++}, as occurs in oxygenated waters, are dependent primarily on: 1) pH, 2) E_h (redox potential), and 3) temperature. As the hypolimnion of a lake goes anaerobic, ferric iron is reduced to ferrous iron. This reduction in the state of iron causes a release of iron-bound phosphorus from the sediments, because ferrous iron is quite soluble. Thus, the changing concentration of both forms of iron over time are very important in lakes, consult: 1) Hutchinson (1957), and 2) Stumm and Morgan, (1970).

Microbiological organisms (bacteria, viruses, etc.) have been used to determine the safety of water for drinking, swimming and shellfish harvesting, for it is well known that water may serve as a medium for the transfer of disease (EPA, 1976). The exact relationship between numbers of specific disease-causing organisms in water and the potential

for transmission is unknown. However, the numbers and biomass of bacteria tend to increase with increasing concentrations of inorganic and organic compounds in lakes (i.e., change from oligotrophic to eutrophic lake). The seasonal distribution of bacterial populations is highly variable between lakes and within a lake between years. In some cases, bacterial populations are correlated to numbers of phytoplanktonic algae (Wetzel, 1975). Bacteria of the coliform group are usually considered the primary indicators of fecal contamination from warmblooded animals.

Nitrogen occurs in fresh water in numerous forms: 1) ammonium (NH_4^+), 2) nitrite (NO_2^-), 3) nitrate (NO_3^-), 4) dissolved molecular (N_2), and 5) a large number of particulate and dissolved organic compounds (Wetzel, 1975). Ammonia is a pungent, gaseous, alkaline compound of nitrogen and hydrogen and is highly soluble in water. Ammonia is generated primarily from decomposition of organic matter by heterotrophic bacteria and as an excretory product of animals. It is present primarily as NH_4^+ and as undissociated NH_4OH (NH_3 is sometimes used), the latter being highly toxic (Trussell, 1972). The proportions of NH_3 to NH_4OH are dependent on 1) pH, 2) temperature, and 3) ionic strength (salinity). Ammonia usually increases as lake productivity increases or when the hypolimnion is anaerobic. Nitrite (NO_2^-) levels in natural lake waters are usually very low, while nitrate (NO_3^-) is much more common. Nitrogen may often be the limiting factor in algal productivity, especially when large amounts of phosphorus have been added to the lake by man's activities. Vollenweider (1968) found a direct correlation, with some exceptions, between productivity of algae and average concentrations of nitrogen. Since nitrogen is a dominant

factor in lake systems, the analysis of trends in nitrogen concentrations are extremely useful in the examination of water quality.

Oxygen is a fundamental parameter in lakes and is a major parameter for evaluating water quality. Dissolved oxygen has a direct effect on the maintenance of aquatic life. Insufficient dissolved oxygen in the water causes: 1) decrease in numbers and kinds of aquatic life, 2) decomposition of organic materials, and 3) release of nutrients from the sediments and formation of anaerobic gases (e.g., hydrogen sulfide and methane). Numerous criteria for dissolved oxygen have been established for most forms of aquatic life (EPA, 1976). While mean oxygen concentrations are mainly used for water quality, oxygen deficits may also provide valuable information. The oxygen deficit, in lakes, is the difference in the amount of oxygen present at the beginning and at the end of stratification below a given depth. The oxygen deficit reflects the amount of organic matter synthesized by measuring the rate of oxygen utilization, thus provides an indirect estimate of lake productivity. The use of oxygen as a parameter for water quality trend analysis is extremely effective.

pH is a measure of the hydrogen ion activity in water and results from the dissociation of water to H^+ and OH^- ions. The pH usually is defined as the logarithm of the reciprocal of the concentration of free hydrogen ions (i.e., $pH = \log \frac{1}{H^+}$). In truth, pH measures the activity of the hydrogen ion and not the concentration. The pH of natural waters is governed mainly by the carbonate system (discussed in inorganic carbon). pH affects the dissociation of weak acids and bases of many toxic compounds. The solubility of metal compounds in bottom sediments or suspended material also is affected by pH. As with

alkalinity and inorganic carbon, pH may provide valuable information on water quality in areas receiving acid rain.

Phosphorus is one of the major nutrients required for algal and macrophyte nutrition and is often the limiting factor for productivity in lakes. When phosphorus is limiting, increased supplies of phosphorus have led to a condition of accelerated eutrophication or aging of waters. The majority of phosphorus in lake water is bound organically in organic phosphates and cellular constituents (Wetzel, 1975). The only significant form of inorganic phosphorus in natural waters is orthophosphate ($PO_4^{=}$, $H_2PO_4^{-}$). Phosphorus is often stored, consolidated in lake sediments (i.e., phosphorus sink), with some being released under anaerobic conditions. The amount of phosphorus, and other nutrients, retained by a lake is a function of: 1) the phosphorus loading to the lake, 2) the volume of the euphotic zone (zone that receives light), 3) the extent of biological activities, 4) the lake detention time, and 5) the level of discharge (outflow) from the lake (EPA, 1976). Phosphorus is probably the most studied of all lake parameters, and only a small fraction is presented here. For more information on phosphorus in natural waters consult: 1) Hutchinson, (1957), 2) Wetzel, (1975), and 3) Hynes, (1970).

Physical parameters describe the physical characteristics of lakes. Some physical parameters are: 1) temperature, 2) inflow and outflow, 3) lake volume, 4) lake and shoreline area, 5) mean depth, 6) hydraulic detention time, and 7) sediment characteristics. Changes in the physical state of a lake may have definite effects on lake water quality. For example, reduction in lake volume may decrease the hypolimnion of the lake. Thus, an oxygen depletion may occur due to the small volume of water (and oxygen) in the hypolimnion. The usefulness of physical

parameters is increased when dealing with reservoirs, where physical characteristics are easily controlled. For more information on physical parameters consult: 1) Hutchinson, (1957), and 2) Wetzel, (1975).

The salinity of inland waters is made up primarily of the major anions, bicarbonate (HCO_3^-), carbonate ($CO_3^=$), sulfate ($SO_4^=$), and clorides (CL^-) and major cations, calcium (Ca^{++}), magnesium (Mg^{++}), sodium (Na^+) and potassium (K^+). The proportions of major ions in natural waters tend towards Ca > Mg $\geq$ K and CO_3 > SO_4 > Cl, with Na and CL having larger concentrations in soft waters (Wetzel, 1975). Magnesium, sodium, potassium, and chloride are relatively conservative ions and undergo minor spatial and temporal changes within a lake. Calcium, inorganic carbon, and sulfate are non-conservative (dynamic) ions and their concentrations are strongly influenced by biotic activities. Soft and hard waters refer to conditions of low and high salinity levels, respectively in natural waters. The ratios of monovalent:divalent cations and the proportions of cations influence the metabolism of many organisms. Therefore, salinity can indirectly affect seasonal population succession and productivity of certain algae and macrophytes.

Silica (SiO_2) is a major component in algal production, especially for diatom algae which utilize silica for cell wall formation. Seasonal population dynamics of diatoms can greatly influence silica concentrations in natural waters (Lund, 1949; 1950). In a lake dominated by diatoms, sedimenting diatom frustules can accumulate within the sediments and be lost permanently to the system. Thus, silica concentrations will exhibit decreasing trends and affect diatom populations. The concentration of silica is very important for water quality trends in lakes where diatom algae are present (e.g., Lake Michigan).

Other water quality parameters, and references for information on each of the parameters, are: 1) toxic compounds (EPA, 1976), 2) micro-nutrients, (boron, calcium, cobalt, molybdenum, zinc, etc.) (Wetzel, 1975), 3) dissolved and suspended solids (Hutchinson, 1957), and 4) organic matter (Hutchinson, 1957). The water quality parameters discussed above are far from all inclusive, but are usually the major parameters that are used for evaluating water quality and for which sufficient data has been collected for to allow good trend analysis.

Data Selection

The process of selecting data for use in water quality trend analysis consists of four phases:

1. Determine desired output
2. Determine availability of data
3. Choose data to fit needs
4. Place data in desired form

The first three are discussed in this section, while the last, data preparation, is discussed in Chapter 4.

The first step in data selection is to determine the desired output, or the information that is needed. When dealing with water quality trends, for example, the desired output may be the changes in the overall quality of water or changes in a specific parameter. Thus, the desired output may determine the exact parameter to use. For example, if the management concern is whether a lake is maintaining its ability to sustain trout populations, trends in temperature and dissolved oxygen, the two critical factors for trout, should be analyzed. Often, the

desired output may be achieved by analyzing one of a number of parameters. For example, when trends in overall water quality are the desired outputs, examination of phosphorus, nitrogen, oxygen or others may be used. Also, when overall water quality is desired, the analysis in trends in trophic status indicators may provide more information than does any single parameter. Some trophic state indicators are: 1) Carlson (1977), 2) Walker (1979), and 3) Reckhow (1980) who provides a review of trophic state indicators.

Once the desired output has been defined, the data that are available must be determined. The major concern with data availability is which parameters have been sampled. Water quality data has been collected by countless individuals for numerous reasons. Ideally, one would like to use their own data. However, when dealing with water quality trend analysis, especially annual trends, one will often need to use data from other sources. There are numerous water quality data banks, for example: 1) STORET (EPA), 2) WATSTORE (USGS), 3) NAWDEX (USGS), and 4) several state data systems (see Edwards (1980) for a review of water quality data systems). An important issue when dealing with different or multiple data sources is whether the data are mutually compatible. Similar sampling designs, sampling devices, laboratory techniques, and methods are a prerequisite to combine data. A major concern in time series analysis is the need for continuously spaced data over time, without missing values. Water quality data usually are unevenly spaced in time, with missing values. Thus, the combination of data from different sources may be necessary. Usually, the more data one has, the more information one can obtain. However, extreme care must be taken when combining data sets from various sources.

Based on the desired output and the availability of data, data should be chosen for trend detection analysis. Ideally, one has a definite parameter in mind based on the desired output and excellent data on these parameters collected over time. In reality, the choice of a parameter(s) and the choice of data on this parameter to use, are highly subjective. It may be necessary to use the parameter desired, with low sample size, or use another parameter (related to that desired, limnologically or by correlation) with a larger sample size. The choice may be remedied by analysis of a few parameters in order to use all of the results in the management and planning of water resources. In closing, the selection of an appropriate parameter(s) and associated data is extremely important in the application of trend detection analysis, for the information produced from the analysis can only be applied on the basis of the data used.

CHAPTER III

STATISTICS

Introduction

In order to apply trend detection techniques, a general knowledge of statistics is necessary. The key statistical topics relevant to the application of time series analysis are descriptive statistics, probability distributions, and hypothesis testing. The information in this chapter provides the basic theory of these statistical concepts in an informal manner. The first section presents the statistical aspects of descriptive statistics. The application (when and where to use descriptive statistics) is discussed in the beginning of Chapter 4. The second section on probability distributions is a theoretical presentation which is needed to understand hypothesis testing. Finally, hypotheses testing is discussed and provides the basis of statistical trend detection techniques.

This chapter is not meant to serve as a substitute for a statistical textbook, but is intended to present only the necessary areas of statistics needed to properly apply and analyze the trend detection techniques presented in Chapter 4. The following books are suggested for a more complete treatment on these subjects: 1) classical statistics (Sokal and Rolf, 1969; Bhattacharyya and Johnson, 1977; Neter and Wasserman, 1974), 2) nonparametric statistics (Conover, 1971; Siegel, 1956; Hollander and Wolfe, 1973), and 3) data analysis and regression

(Mosteller and Tukey, 1977; Reckhow and Chapra, 1980; Tukey, 1977; Reckhow, 1980; Chatterjee and Price, 1977).

Descriptive Statistics

In order to apply statistical methods, a concise description of the data is necessary. This can be achieved by performing arithmetic operations on the data to obtain values for one or more descriptive measures or statistics.

There are three catagories of statistics used to describe random variables in a particular population: 1) measures of central tendency (location of an ordinary value), 2) measures of dispersion (relative distance of extreme values from a central value), and 3) measures of relationship between variables (degree of similarity or dissimilarity in magnitude).

Measures of Central Tendency

In most sets of data there is a tendency for the observed values to group themselves about some central value. This central value is characteristic of the data and may be used to describe the central tendency of the data's distribution. The statistics that describe this phenomenon are measures of location or central tendency. Common measures of location include the arithmetic mean, median, and mode.

The average, or arithmetic mean, is the most frequently used of all statistical measures. The population arithmetic mean (or simply mean) of a particular random variable Y in a population is usually denoted

by the Greek letter μ or μ_X. The estimate of μ_X for samples of size n is:

$$\hat{\mu}_X = \bar{X} = \sum_{i=1}^{n} X_i/n \qquad (8)$$

where: $\hat{\mu}_X = \bar{X}$ = estimate of population mean (sample mean)

$X_i = i^{th}$ observation

n = sample size (i.e., number of X's).

The mean is calculated by summing all the individual observations (X_1, X_2, X_3, ..., X_n) of a sample and dividing this sum by the number of items (n) in the sample. If, for example, three total phosphorus concentration measurements of 1, 2, and 6 mg/l were taken from a lake, their sample mean concentration is:

$$\bar{X} = \frac{1 + 2 + 6}{3} = \frac{9}{3} = 3 \text{ mg/l}$$

The median is the value of a random variable that ranks midway between the largest and smallest values. It can also be defined by the value of the variable (in an ordered array) that has an equal number of items on either side of it, (i.e., divides the frequency distribution in half).

If n is an odd integer, the sample median is the $(n + 1)/2^{th}$ number in the ordered array. For example, consider this array of oxygen concentrations 1, 1, 2, 3, 7, 8, 11, 12, 14, 19, 20 mg/1. It has 11 observations and the (11 + 1)/2, or the sixth number, is equal to 8. Hence, the median oxygen concentration of this set of values is 8 mg/l. If n is even, the median is calculated as the midpoint between the $(n/2)^{th}$ and the $[(n/2) + 1]^{th}$ variate. Thus, from a sample of 6 oxygen

concentrations, 5, 8, 10, 11, 12, 13 mg/l, the median would be the mid-point between 6/2 = 3rd and [(6/2) + 1] = 4^{th} value. For this example, the median is 10.5 mg/l. The median is especially useful when the random variates exhibit skewed (asymmetric) distributions (Mosteller and Tukey, 1977), for example, when dealing with lake phosphorus concentration, in a eutrophic lake.

The mode refers to the value occupied by the greatest number of individuals in a frequency distribution. When applied to a frequency distribution it is the value of the variable where the probability density function peaks. Given a set of nitrate concentrations of 1, 2, 3, 3, 3, 4, 4, 5 mg/l, the mode would be equal to 3 mg/l.

The midrange is the average of the largest and smallest value. The midrange provides a quick and easy measure of location but is subject to extreme variation from sample to sample unless the samples are quite large.

The geometric mean is the antilog of the arithmetic mean of the logarithms of a set of values, and it is always as small or smaller than the arithmetic mean of the same set of values. The geometric mean is computed as:

$$G.M._X = \text{antilog} \frac{1}{n} \sum_{i=1}^{n} \log X \qquad (9)$$

The harmonic mean is the reciprocal of the arithmetic mean of the reciprocals of a set of values and is always as small or samller than the geometric mean of the same set of values. The harmonic mean is computed as:

$$H_X = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X}} \tag{10}$$

In making a decision as to which measure of location should be used with a given set of data, a primary consideration is the intended use once of the measure once selected. In addition, the advantages and disadvantages inherent in each of the measures of location should be known. If the distribution of the data is symmetrical and unimodel, the mean, median, and mode are identical, but as the distribution becomes skewed, differences among these measures will occur. This is illustrated in Figure 5.

In addition to the intuitive appeal for the use of the mean, the mean also has smaller variability from sample to sample, is easier to work with mathematically, and has more desirable properties in connection probability distributions than other measures. However, the mean is sensitive to extreme values, particularly when n is small.

In general, the mean is recommended when the distribution is normal or uniform. Robust statistics (e.g., median) are preferred when the data distribution is skewed or irregularly shaped or when insufficient information is available, i.e., n is small (Mosteller and Tukey, 1977; Tukey, 1977). If the distribution is questionable as to whether the distribution is normal or not, it is desirable to use a couple different measures of location to provide more information.

Measures of Dispersion

In most sets of statistical data the numerical values will not be identical, but will be scattered or dispersed to some degree. The

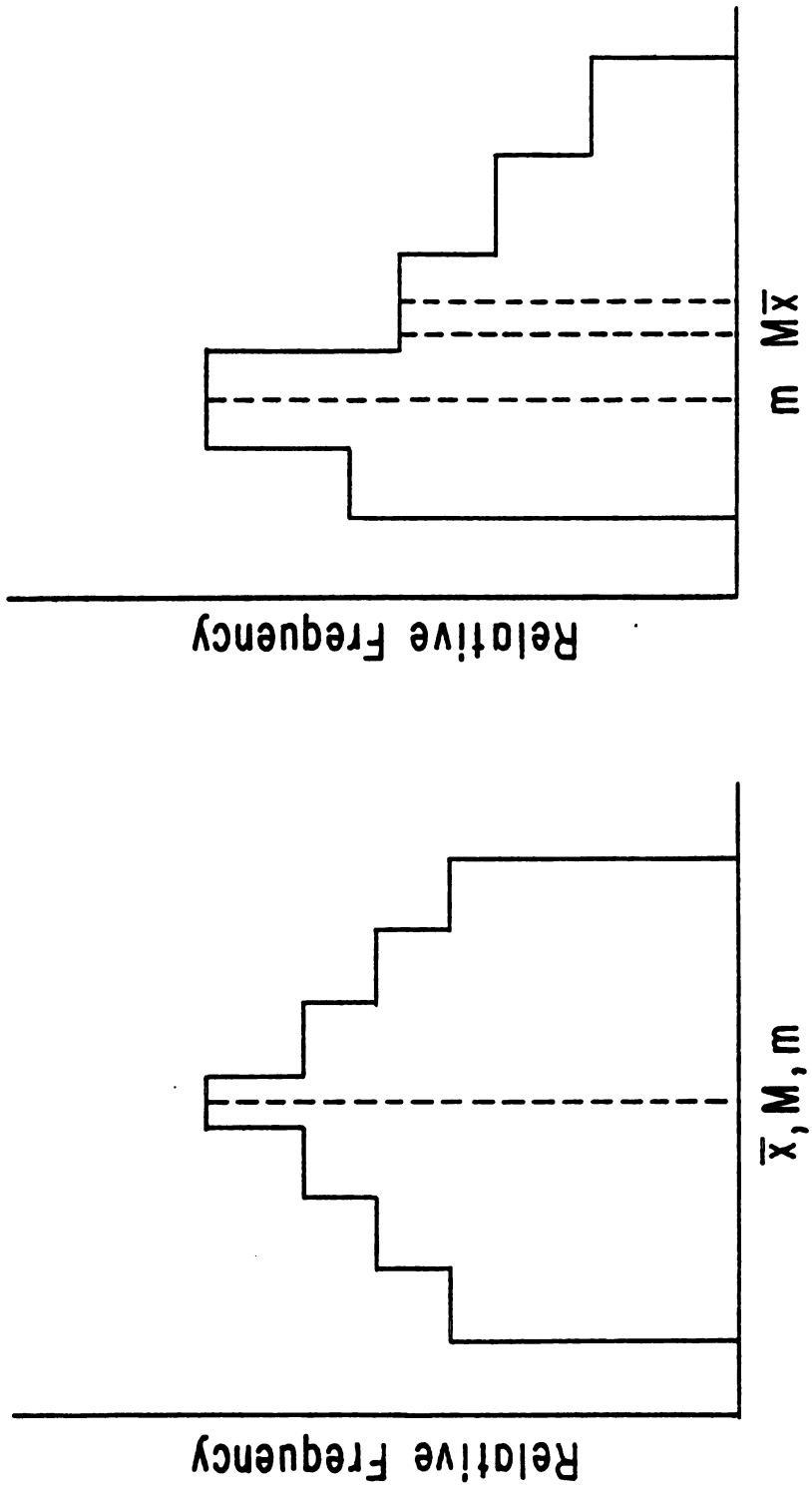


Figure 5. Effect of distribution shape on mean ($\bar{x}$), median (M), and mode (m).

statistics used to measure this characteristic of the data are measures of variation or dispersion. Several sets of data could have the same, or nearly the same, mean, median, or mode, but vary considerably in the level of dispersion around a central value. Thus, a more complete description of the data results when we evaluate one of the measures of variation in addition to one or more of the measures of location.

The range is the difference between the largest and smallest values in the data. Using data set A.1 (found in Appendix A), the range is 7.1. Although the range is easy to calculate and is commonly used as a rough-and-ready measure of variability, it is generally not a satisfactory measure of variation for three reasons. First, the calculation involves only two of the observations, regardless of sample size. Therefore, it utilizes only a fraction of the available information concerning variation in the data. Secondly, since the range tends to become larger as sample size increases, it is improper to compare ranges from two sets of data with different sample sizes. Finally, the range is very unstable except in small sample sizes. With repeated samples taken from the same source, the ranges will exhibit more variation from sample to sample than will other measures of variation. However, the use of the range differs from the other measures in that it provides a relatively good measure of variation for small numbers of observations.

Among measures of variation, the standard deviation and its square the variance, are almost universally accepted as the most useful dispersion statistics. The standard deviation for a particular random variable X in a population usually is denoted by the lower case Greek letter sigma (σ), and the variance by σ^2 or σ_X^2. The estimate of σ^2 from a particular sample size n is:

$$\hat{\sigma}^2 = s^2 = \left[\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2 \Big/ n\right] \Big/ (n - 1) \qquad (11)$$

Using data set A.1, an estimate of the variance and standard deviation of the data are as follows:

$s^2 = [(20.8^2 + \ldots + 15.9^2) - (20.8 + 20.9 + \ldots 15.9)^2/11] / 10$

$= [4727.37 - (226.9)^2/11] / 10$

$= (4727.37 - 4680.32) / 10$

$s^2 = 4.70$

$s = \sqrt{4.70}$

$s = 2.17$

Strictly speaking, only one parameter σ or σ^2 is needed to describe the dispersion in X. However, the squared form is much easier to work with mathematically (variances are ascribed to independent causal agents are additive) while the unsquared form has the advantage of being expressed in the same units as X, the variable measured. Thus, both are usually calculated and used.

When the data are presented in an ordered array, the interquartile range is the difference between the value at the 75 percent level and the value at the 25 percent level. The interquartile range provides a description of the dispersion in the central half of the distribution. Since the interquartile range, like the median, is based on order statistics, it is robust in situations with extreme data (i.e., outliers) and skewed distributions. These percent levels can be altered to accommodate more or less of the variable as desired. The interquartile range for data set A.1 is 3.55.

The mean of median absolute deviation is computed by:

$$A.D. = \frac{\sum_{i=1}^{n} |X_i - \bar{X}|}{n} \tag{12}$$

where: A.D. = mean or median absolute deviation

$\bar{X}$ = mean or median

The value of $\bar{X}$ is either the mean or median depending whether the mean or median absolute deviation is desired. The choice between these two is equivalent to the choice between the mean and median. The mean and median absolute deviation for data set A.1 is 1.42 and 1.38, respectively.

The coefficient of variation is a measure of relative, rather than absolute variation, since it is a unitless quantity. It is calculated by:

$$C.V. = \frac{\hat{\sigma}}{\hat{\mu}} = \frac{s}{\bar{X}} \tag{13}$$

and can be expressed as a ratio or a percentage. It's primary advantage is that it is independent of the unit of measurement and can therefore be used to compare the relative variations of two or more sets of data, regardless of the units involved. The coefficient of variation for data set A.1 is .105.

Standard deviations of various statistics are generally known as standard errors. The standard error of a statistic, for example the mean, is the standard deviation of a distribution of means for samples of a given sample size n. The standard error is used as a measure of the reliability of an estimate. The following are the estimates of standard error for the mean and median, respectively (for standard errors of other estimators, consult Sokal and Rohlf (1969)).

$$s_{\bar{X}} = \frac{s}{\sqrt{n}} \tag{14}$$

$$s_{med} = (1.2533)\, s_{\bar{X}} \tag{15}$$

where: $s_{\bar{X}}$ = standard error of mean

s_{med} = standard error of median

The estimates for $s_{\bar{X}}$ and s_{med} for data set A.1 are .65 and .82, respectively. It should be noted that $s_{\bar{X}}$ is valid for any population with finite variance and s_{med} for large samples from normal populations.

Measures of Relationships

Measures of relationship commonly used are correlation and regression. There has been much confusion on the subject matter of correlation and regression for several reasons. First, the mathematical relations between the two methods of analysis are similar. Second, earlier statistical texts did not make a sufficiently clear distinction between the two approaches. Finally, while the approach chosen by an investigator may be correct in terms of his intentions, the data available for analysis may be such as to make one or the other technique inappropriate.

Regression is intended to describe the dependence of a variable Y on an independent variable X. Regression equations lend support to hypotheses regarding the possible causation of changes in Y by changes in X; for purposes of prediction, of Y in terms of X; and for purposes of explaining some of the variation of Y by X, by using the latent variable as a statistical control.

Correlation, by contrast, is concerned largely with whether two

variables are interdependent of covary (i.e., vary together). One variable is not expressed as a function of the other, hence no distinction between dependent and independent variables is made. The intent is to estimate the degree to which these variables vary together.

Regression is the proper measure of relationship when a random variable Y is dependent on, or caused by, one or more controllable or fixed variates X_i, which are said to be independent. Ordinary methods of linear regression require a proposed model to be linear in the parameters but not necessarily in the relation of Y to X. The simple linear statistical model is:

$$Y = \beta_0 + \beta_1 X + e \qquad (16)$$

where: Y = dependent variable

β_o = origin (extrapolated value of Y when x is fixed at zero)

β_1 = slope (average change in Y per unit change in x)

X = independent or fixed variable

e = random error

The random error is composed of two basic parts: 1) failure of the linear form properly to describe the relation between Y and X (i.e., nonlinear bias) and 2) random contributions of latent variables to Y.

Regression is based on four assumptions (Sokal and Rohlf, 1969).

1. The independent variable X is measured without error (i.e., the X values are known). This means that only Y the dependent variable, is a random variable, and X does not vary at random.
2. The expected value for the variable Y for any given X is described by the linear function $\mu_Y = \beta_o + \beta_1 X$.

Another way of stating this assumption is that the parametric means μ_Y of the values of Y are a function of X and lie on a straight line described by this equation.

3. For any given value of X the Y's are independently and identically distributed. The distributions must be normally distributed when confidence intervals or hypothesis testing is needed. Figure 6 illustrates this assumption. By taking repeated measurements each year, a frequency distribution of nutrient concentrations (Y) to the independent variates (X = time (years)) is generated. Due to the inherent variability in lakes and in time, it is obvious that a frequency distribution of values of Y (nutrient concentration) around the expected value will result. This assumption states that these sample values must be independently and identically distributed.

4. The variance of Y given X, is equal for all X's. This means that variances of the samples along the regression line are homoscedastic. Thus, the variance around the regression line is constant and independent of the magnitude of X and Y.

Commonly, least squares estimation is used to derive estimators for the regression parameters β_o and β_1. The general idea is to minimize for samples of size n the squared random errors of Equation 16 with respect to β_o and β_1. This is so the estimators obtained specify an estimator of Y, given X, which has the smallest variance of any linear equation having unbiased estimators of the same parameters. An estimate for β_1 is:

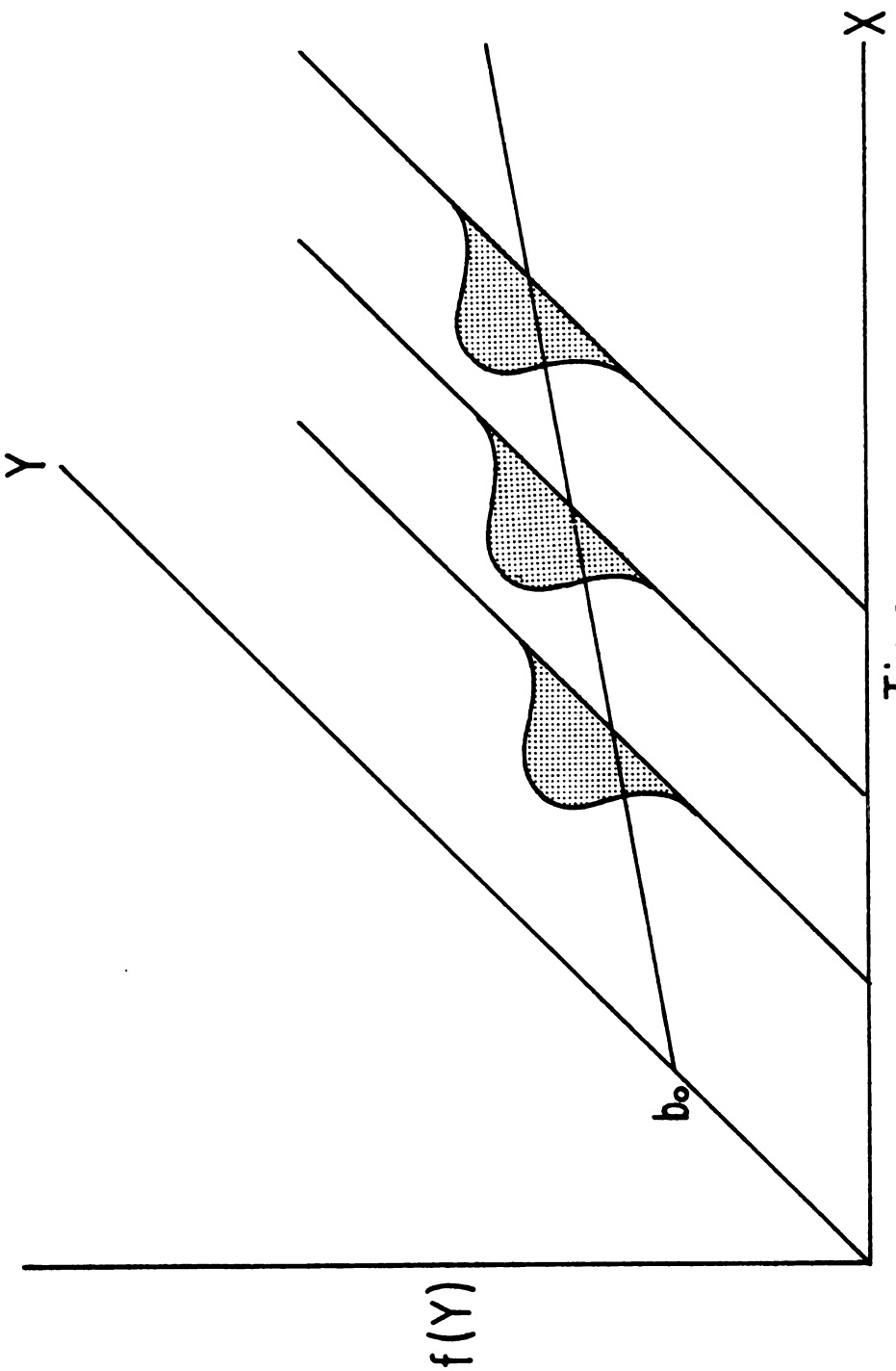


Figure 6. Normally distributed Y's given x (independent data assumption).

$$\hat{\beta}_1 = b_1 = \frac{\left[\sum_{i=1}^{n}\left(X_i\ Y_i\right) - \left(\sum_{i=1}^{n} X_i\right)\left(\sum_{i=1}^{n} Y_i\right)\Big/ n\right]}{\left[\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2\Big/ n\right]} = \frac{sp_{xy}}{ss_X} \qquad (17)$$

where: sp_{xy} = sum of cross products of X and Y

ss_X = sum of squares of X

β_o can then be estimated by:

$$\hat{\beta}_o = b_o = \bar{Y} - b_1 \bar{X} \qquad (18)$$

Given a linear model, an observation Y may be partitioned into three parts: 1) the mean, 2) the deviation of the regression line from the mean (regression effect), and 3) the deviation of the observation from the regression line (error). Therefore, the sum of squares of Y may be partioned into a sum of squares caused by regression and a residual (error) sum of squares, which measures the failure of the observed values to fall exactly on the regression line. The equation takes the form:

$$ss_Y = ss_R + ss_e \qquad (19)$$

where: ss_Y = sum of squares of Y

ss_R = regression sum of squares

ss_e = residual (error) sum of squares

When ss_e is divided by n - 2 degrees of freedom, the error sum of squares is an unbiased estimate of the variation of Y, given X, if the linear model is correct, i.e.,

$$(s_y^2|X) = ss_e/(n - 2) \qquad (20)$$

where: $(s_y^2|X)$ = variance of Y given X

Usually, the error sum of squares is obtained by subtracting the regression sum of squares from the total sum of squares for Y. The regression sum of squares is computed as:

$$ss_R = b_1 \, sp_{xy} \tag{21}$$

Also note that:

$$ss_R/ss_Y = (sp_{XY})^2/ss_x \, ss_Y = r^2_{XY} \tag{22}$$

where: r^2_{XY} = coefficient of determination (represents percentage of the total variation explained by the model)

This (r^2_{XY}) shows that the proportion of the total sum of squares of Y attributable to linear regression on X is the square of the correlation between X and Y.

In order to develop interval estimates or tests of hypotheses for the parameters and predictions, the validity of the assumptions of normality, homogeneous variance, and linearity must be checked. The procedures for satisfying this requirement are in the next chapter. Adjustments may be made to the data for: 1) non-normality by transformations (discussed in the next chapter), 2) heterogeneous variance by transformations or weighted regressions, and 3) nonlinearity by transformations and adding higher degrees of polynomials to the model (consult an advanced regression analysis text for more information concerning these problems).

Given satisfaction in the assumptions, the (1-α) 100 percent confidence interval estimates may be computed from:

$$b_o \pm t_{\alpha/2,\ n-2}\ (s_Y|X)\ \sqrt{(1/n) + \bar{X}^2/ss_X} \tag{23}$$

$$b_1 \pm t_{\alpha/2,\ n-2}\ (s_y|X)/\sqrt{ss_X} \tag{24}$$

where: $t_{\alpha/2,\ n-2}$ = value from t distribution

α = alpha level, i.e., desired level of significance

The expression to the right of $t_{\alpha/2,\ n-2}$ is the standard error of estimate for b_o and b_1.

Tests of hypotheses about parameters (e.g., H: $\beta_1 = \beta_1'$) are computed from:

$$t = (b_1 - \beta_1') \ /\ [(s_y|x)/\sqrt{ss_x}] \tag{25}$$

where: β_1' = the slope you are testing for

The appropriate critical levels for a two-tailed are $\pm t_{\alpha/2,\ n-2}$ and for a one-tailed $t_{\alpha,\ n-2}$. The theory of hypotheses testing will be discussed in a later section and present the information about those tests to those unfamiliar with the subject.

The following is an example of developing a linear regression model and estimation of parameters for data set A.2. In this example all assumptions are assumed valid, which in reality, is probably not the case.

When using time as the independent variable it is convenient to recode the time units into whole numbers from 1 to n. For this example, time (years) are 1968-1979, which are recoded from 1 to 12. The first step in developing a regression model is to calculate the slope (b_1) of the equation (Equation 17):

$$b_1 = \frac{\left[\sum_{i=1}^{n}\left(X_i\ Y_i\right) - \left(\sum_{i=1}^{n} X_i\right)\left(\sum_{i=1}^{n} Y_i\right)\Big/n\right]}{\left[\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2\Big/n\right]} = \frac{sp_{XY}}{ss_X}$$

$$b_1 = \frac{[(1\cdot 215. + 2\cdot 237. + \ldots + 12\cdot 335) - (1 + 2 + \ldots + 12)\cdot(215. + 237. + \ldots + 335.)/12]}{[(1^2 + 2^2 + \ldots + 12^2) - (1 + 2 + \ldots + 12)^2/12]}$$

$$= \frac{[22877 - (78\cdot 274)]}{[650 - 5071} = \frac{1505}{143} = \frac{sp_{XY}}{ss_X}$$

$$b_1 = 10.52$$

The intercept (b_o) can then be solved (Equation 18):

$$b_o = \bar{Y} - b_1 \bar{X}$$

$$b_o = 275. - (10.52 \cdot 6.5)$$

$$= 205.62$$

Therefore, the regression model is $Y = 205.62 + 10.52X + e$.

The quantity ss_Y is calculated the same as ss_X, with y's substituted for X's (Equation 17 (denominator)).

$$ss_Y = (215.^2 + 237.^2 + \ldots + 335^2) - [(215. + 237. + \ldots + 335)^2/12]$$

$$= 917452 - (3288^2/12)$$

$$= 917542 - 900912$$

$$ss_Y = 16540.$$

The regression sum of squares (SS_R) is (Equation 21):

$$ss_R = b_1\ SP_{xy}$$

$$ss_R = 10.52 \cdot 1505$$

$$= 15832.6$$

Then, the residual (error) sum of squares is calculated (Equation 19):

$$ss_e = ss_Y - ss_R$$

$$ss_e = 16540 - 15832.6$$

$$= 707.4$$

The variance of Y given X is (Equation 20):

$$(s_Y^2|X) = ss_e/(n - 2)$$

$$(s_Y^2|X) = 707.4/10$$

$$= 70.74$$

The coefficient of determination (r^2_{xy}), which represents the percentage of the total variation explained by the model, is calculated (Equation 22):

$$r^2_{XY} = \frac{ss_R}{ss_Y}$$

$$r^2_{XY} = \frac{15832.6}{16540.}$$

$$= .957$$

Finally, standard errors for estimates of b_o and b_1 are calculated (Equation 23 and 24):

$$s_{b_o} = (s_Y|X)\sqrt{(1/n) + \bar{X}^2/ss_X}$$

$$= 8.41\sqrt{(1/12) + 6.5^2/143}$$

$$= 1.32$$

$$s_{b_1} = (s_Y|X)\sqrt{ss_X}$$

$$= 8.41/\sqrt{143}$$

$$= .703$$

If the joint distribution of X and Y is a bivariate normal distribution, one of its parameters is the product moment correlation coefficient ρ (rho), or simply, the correlation coefficient.

The correlation coefficient ρ is a measure of the linear covariation of the variables, that is, it measures the degree of linear association between them. It may vary from -1 to +1, inclusive, and is a dimensionless quantity. As ρ increases in absolute value, so does the linear association. A positive correlation means as one variable increases, the other increases. A negative correlation means as one variable increases, the other decreases. Since ρ measures only linear relationship, the variable may be perfectly correlated in a curvilinear relationship, and ρ could be equal to zero.

The common estimate for ρ is:

$$\hat{\rho} = r = \frac{sp_{XY}}{\sqrt{ss_X \; ss_Y}} \tag{26}$$

The estimator of ρ is nearly unbiased for large sample sizes but slightly underestimates in small samples, especially in small magnitudes of correlation (Kendall and Stuart, 1967). For small samples ($4 < n < 15$) r may be adjusted (Olkin and Pratt, 1958) to provide an estimate that is nearly unbiased,

$$r^* = r[1 + (1 - r^2) / 2 (n - 4)] \tag{27}$$

where: r^* = adjusted correlation coefficient

To test if $\rho = 0$ a t-value is computed:

$$t = r/\sqrt{(1 - r^2)/(n - 2)} \tag{28}$$

and compared with critical levels of $\pm t_{\alpha/2,\, n-2}$ from the t distribution for a two-tailed test. Weir (1960) has shown a quick, approximate,

two-sided test (H_0: $\rho = 0$, given $n > 4$ and $\alpha = .05$) is achieved simply by noting if $r > 2/n$. Those interested on testing $\rho = \rho_0$ or construction of confidence intervals should consult Bhattacharyya and Johnson (1977).

Another method of estimating the relationship in random variables, where a linear relationship is not assumed, nor that the variables are normally distributed, is Spearman's rank correlation coefficient r_s. It is computed from:

$$r_s = 1 - \frac{6 \sum_{i=1}^{n} \left(d_i^2\right)}{n^3 - n} \qquad (29)$$

where: d_i = difference between the ranks of the ith pair of n pairs of observations.

For samples of $n > 100$, Spearman's coefficient may be tested with good approximation by using the procedures for product moment correlation. Otherwise, see Zar (1974) for tables of the distribution ($4 \leq n \leq 100$).

The following is an example of computing the product moment correlation (r and r*) and Spearman's rank correlation coefficient (r_s) using the same data (A.2) as in the regression analysis example. The estimate for ρ (i.e. r) is calculated using Equation 26:

$$r = \frac{1505}{\sqrt{143 \cdot 16540}}$$

$$= \frac{sp_{XY}}{\sqrt{ss_X \; ss_Y}}$$

$$= .978$$

The estimate is adjusted for small sample size in the following manner (Equation 27):

$$r^* = r\,[1 + (1 - r^2)/2\,(n - 4)]$$

$$= .978$$

The estimate is adjusted for small sample size in the following manner (Equation 27):

$$r^* = r\,[1 + (1 - r^2)/2\,(n - 4)]$$

$$= .978\,[1 + (1\ .978^2)/2\,(12 - 4)]$$

$$= .980$$

The Spearman's rank correlation coefficient (r_s) is computed using (Equation 29):

$$r_s = 1 - \frac{6 \sum_{i=1}^{n} \left(d_i^2\right)}{n^3 - n}$$

Ranks for the observations are found in data set 2 in Appendix A.

$$= 1 - \frac{6 \cdot [(1 - 1)^2 + (2 - 2)^2 + (3 - 3)^2 + (4-5)^2 + \ldots + (12 - 12)^2]}{12^3 - 12}$$

$$= 1 - \frac{6 \cdot 4}{12^3 - 12}$$

$$= .986$$

Probability Distributions

Given any continuous random variable Y, there is a corresponding mathematical expression or function f(Y) known as the frequency function of Y. For the theoretical or population distribution of Y the frequency function is the analog of the frequency distribution (histogram). Thus, f(Y) is a mathematical model that provides a basis for calculating theoretical frequencies or probabilities for any or all outcome classes of the variable. These functions are: 1) defined for all

values of the variable, 2) non-negative for all values of Y, and 3) such that the total area under the corresponding frequency curve and above the Y-axis is equal to one.

Let Y be a continuous random variable with a frequency function f(Y), shown graphically in Figure 7. The function is adjusted so that the total area under the curve is one and corresponds to probability one for the range of all possible values for Y. In addition, f(Y) has the property that, given any two numbers a and b, the probability that a randomly selected element of the population will have a Y value between a and b, inclusive, is equal to the area under the curve between the lines Y = a and Y = b as indicated by the shaded portion in Figure 7. Therefore, for any continuous distribution, the area under the curve between a and b is the probability $P(a \leq Y \leq b)$.

For every discrete random variable there is also a frequency function or probability distribution function which has essentially the same properties as continuous frequency functions. However, since the variable is discrete, the graph of the function does not result in a continuous curve, but in a bar diagram (Figure 8). The probability of each value that Y can take is represented by the height of the appropriate bar.

Among the more important properties of a theoretical distribution is a set of quantities known as the moments of the distribution. The moments characterize the distribution. In applied statistics, the first two moments are of most importance. The first moment about the origin is the mean μ of the theoretical distribution and is defined as the average value or expected value of the variable. The variance σ^2 of a random variable Y is defined as the second moment about the mean, the average

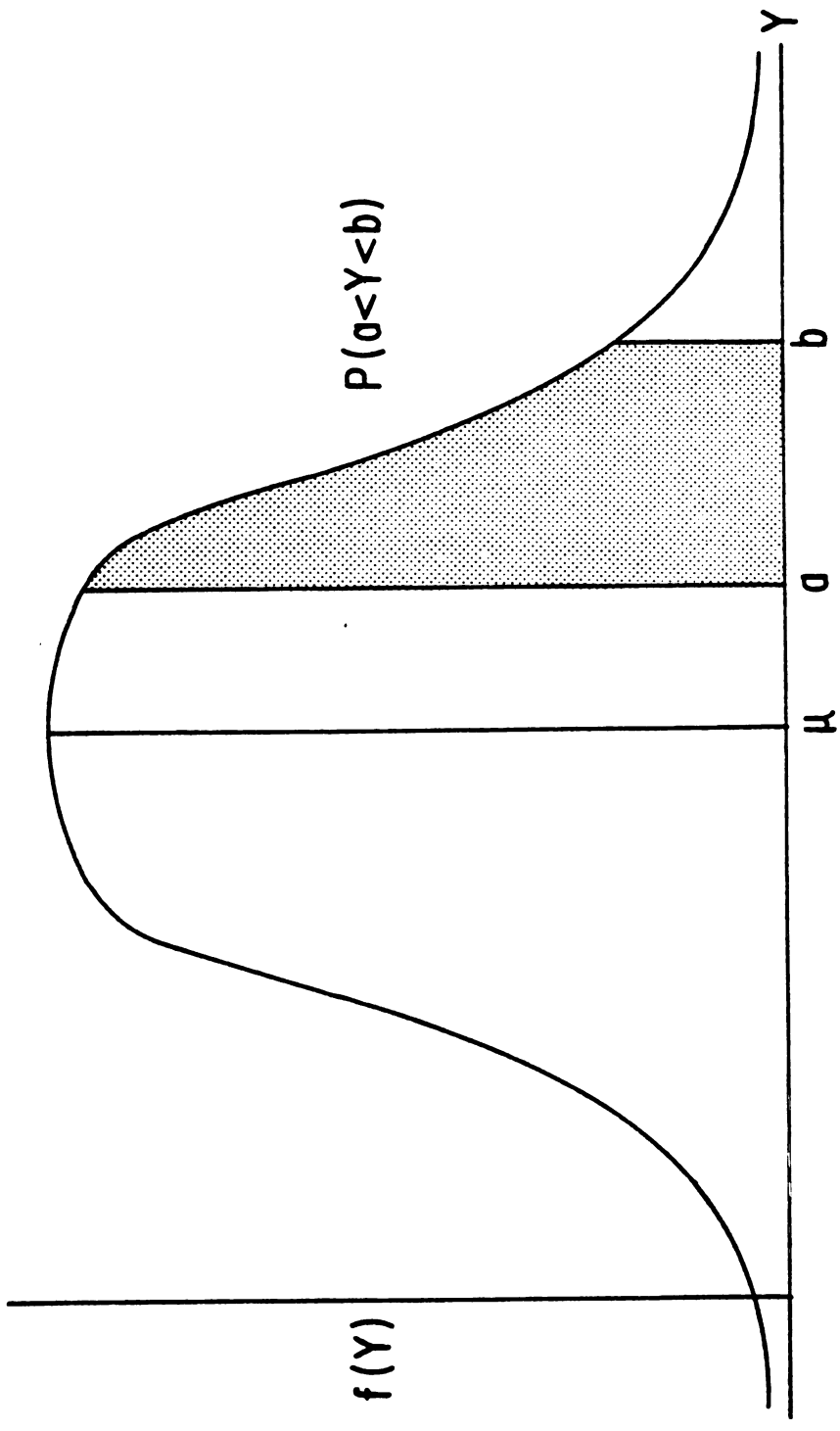


Figure 7. Frequency function for a continuous random variable (Y).

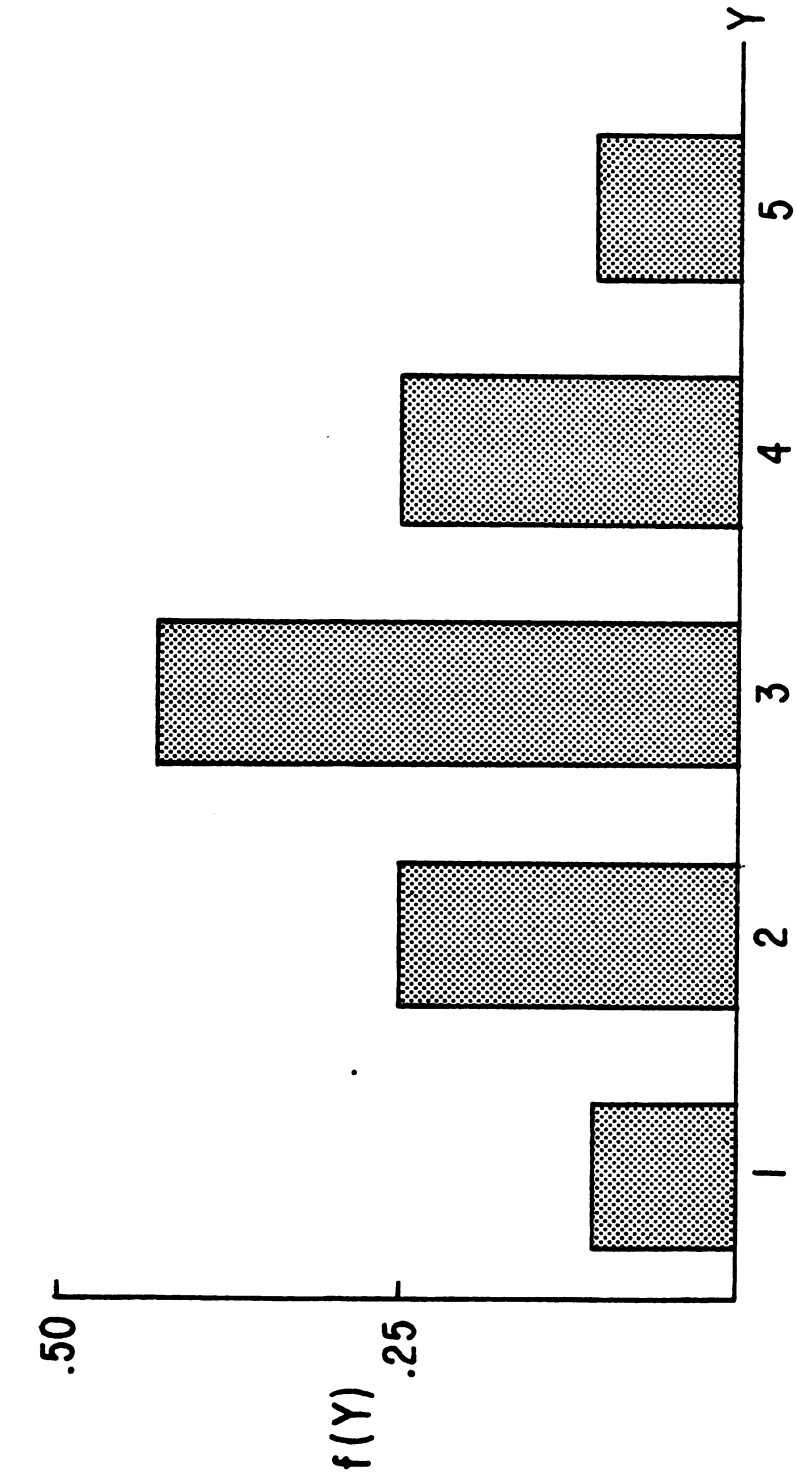


Figure 8. Frequency function for a discrete random variable (Y).

value of $(Y - \mu)^2$. The third and fourth moments, skewness and kurtosis, respectively, refer to distribution shape and will be discussed in the next chapter when dealing with distribution selection.

Probability or density distributions describe the relative frequency of occurrence of value taken on by random variables. The distributions used in practice are usually at best close approximations to the true distribution. The more widely used distributions are the binomial, multinomial, and poisson distributions for discrete variables and the normal or gaussian distribution for continuous variables.

While density distributions describe actual frequency (in populations) or probable occurrence (in samples) of different values of natural phenomena, sampling distributions describe the relative frequencies of different values of functions of random variables in samples of specified size. Three important sampling distributions associated with samples from normal distributions are the chi-square, student's t, and variance ratio (F) distributions.

Normal or Gaussian Distribution

The normal distribution provides a good approximation to many empirical frequency distributions found in biological sciences. Its strength results from the central limit theorem of Laplace (Mood and Graybill, 1963). The speed with which the distribution of means, drawn at random from distinctly non-normal distributions, converges to the normal distribution as sample size increases, provides the keystone to the support of most statistical procedures (Gill, 1978).

For a continuous random variable Y, the normal density function is:

$$f(Y) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(1/2\sigma^2)(Y-\mu)^2} \qquad (30)$$

The distribution of any normal variable depends only on the mean (μ) and variance (σ^2).

The normal density function is a symmetric, bell-shaped curve (Figure 9). The mean is a location parameter and the standard deviation, a scale parameter, is the distance from the mean to the inflection in the curve. An infinite number of normal curves exist, differing in location, scale or both.

Hypothesis Testing

The most frequent application of statistics in biological research is to test some scientific hypothesis. To decide whether or not any disagreement between the observed and the theoretical values is sufficient to warrant rejection of the theory, data are collected in a suitable manner and a statistical test on the hypothesis is used. A statistical hypothesis is an expression, in some manner, of the theory of concern. Statistical methods are important in biology because results of experiments are usually not clearcut and therefore need statistical tests to support decisions between alternative hypotheses. A statistical test examines a set of sample data and, on the basis of an expected distribution of the data, leads to a decision on whether to accept the hypothesis underlying the expected distribution, or whether to reject that hypothesis and accept an alternative one. The nature of the tests varies with the data and the hypothesis, but the same general philosophy of hypothesis testing is common to all tests.

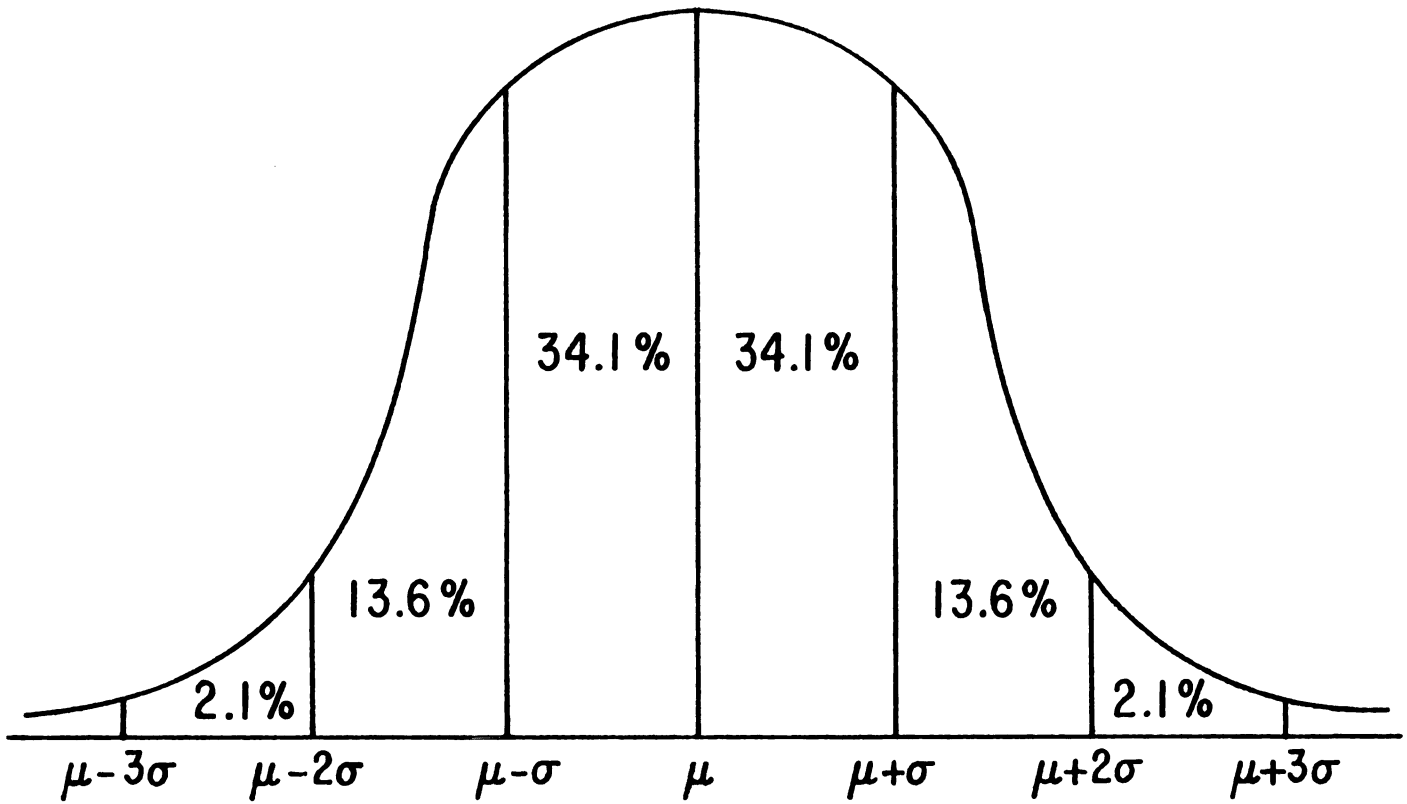


Figure 9 . The normal density function, with mean (μ) and standard deviation (σ).

The first step is to state the null hypothesis (H_0). The null hypothesis is a hypothesis of no differences (such as H_0: $\mu_1 = \mu_2$). It is formulated for the express purpose of being rejected. If it is rejected, the alternative hypothesis (H_1) is accepted. Usually, hypotheses tested by random samples cannot be absolutely disproved because the range of theoretical distributions of sampling variables extends to infinity. If one attempts to propose and test a direct hypothesis, such as H_0: $\mu_1 \neq \mu_2$ against the alternative H_1: $\mu_1 = \mu_2$, statistical assessment of probability of the evidence is not possible because the ability to obtain numerical values from sampling distributions, given H_0, depends on knowing the value of $\mu_1 - \mu_2$, the very value being questioned in the hypothesis. On the other hand, for an indirect hypothesis, such as H_0: $\mu_1 = \mu_2$, the value of $\mu_1 - \mu_2$, given H_0, is zero, and numerical values can be obtained from the sampling distribution.

Rejection of the indirect hypothesis proposed usually is a strong decision because the experimenter chooses a small degree of doubt (probability of being wrong). Therefore, when sufficiently conclusive results occur in a sample, one may be confident (but not certain) that the results represent the true status of the population. Acceptance of the indirect hypothesis proposed normally is a weak decision, because in that case the experimenter usually cannot completely control the probability of being wrong. Acceptance of a hypothesis, such as H_0: $\mu_1 = \mu_2$, should not be interpreted firmly, such as "no mean difference exists," but in a more qualified way. For example, "the experimental difference in means was not sufficient to provide high confidence that the true means differ." It is important to be cautious against over confidence in the results of any one isolated case by reminding us that "the one

chance in a million will undoubtedly occur sometime." However, one should not discard a significant result just because it leads to awkward conclusions or goes against one's personal bias.

Given that a null hypothesis is formed and data collected, the question arises "what is the probability of having obtained the observed outcome or an even more extreme outcome if there is actually no difference in populations?"[1] To answer this question, the relative frequency distribution of all possible outcomes, given the two populations are equivalent, must be examined. The specific observed outcome may then be compared with this distribution. Figure 10 shows a typical outcome distribution. The proportion of the probability in the tails is designated as α (alpha), the level of significance. The points on the outcome scale, such that no more than 100 $(\alpha/2)$% of the outcomes are beyond them in each tail are referred to as the critical values. The total area beyond these values is called the critical region or rejection region, and is shaded in Figure 10. This region contains the most rare 100 α% of the outcomes, given H_0 is true. The area inside the critical values is called the acceptance region (unshaded area). Given this distribution, the probability that an outcome will fall in the critical region, if H_0 is true, is less than or equal to α. In other words, if H_0 is true and a study is repeated numerous times, in about 100 α% of cases the observed outcome would fall into this region. The value chosen for α provides an arbitrary means of making a decision as to whether or not an observed outcome is rare or not rare under H_0.

---

[1]Here we consider the null hypothesis to be concerned with possible differences between two populations. This is the case with trend detection.

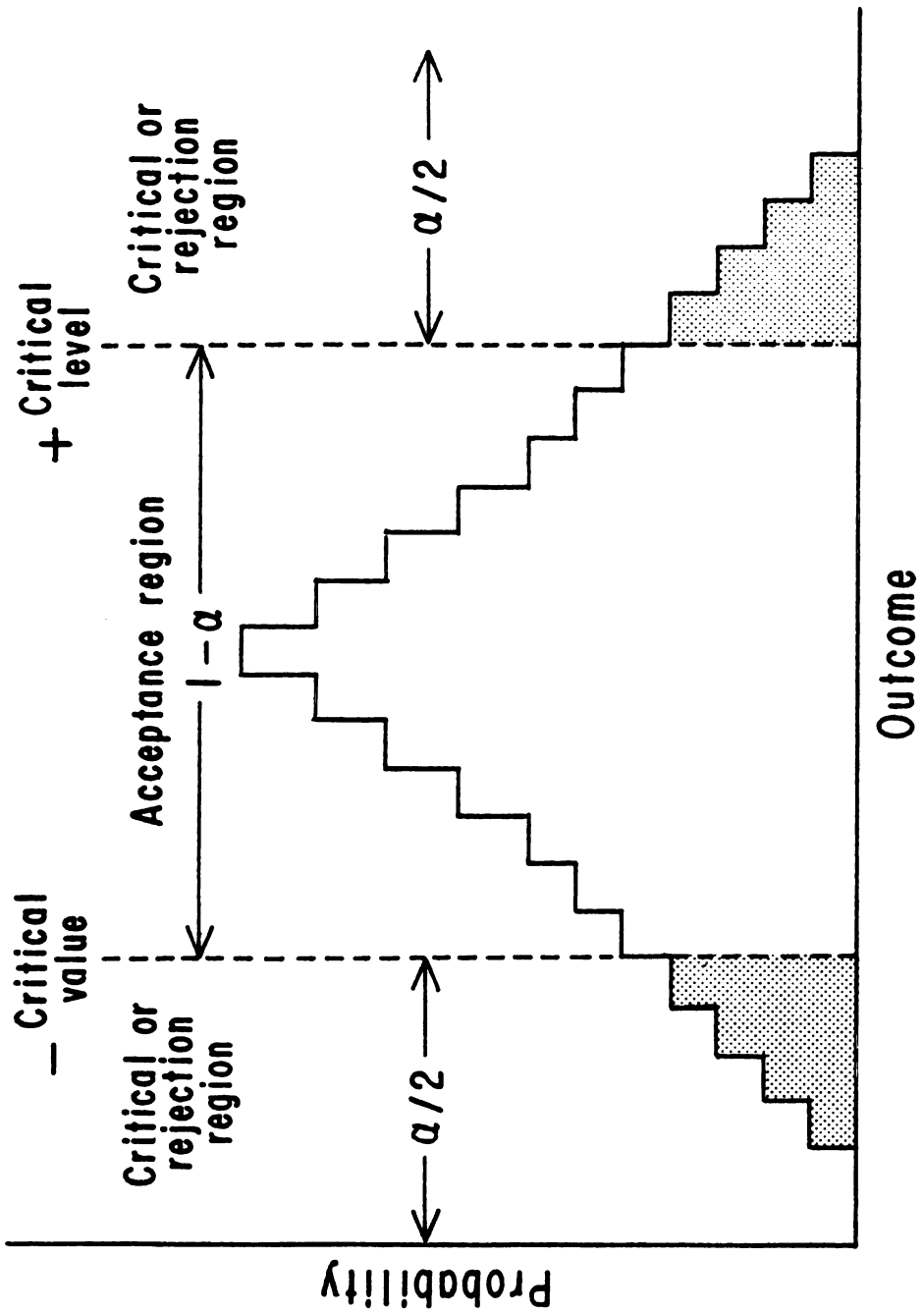


Figure 10. An outcome distribution.

In testing a null hypothesis, two possibilities arise. One is when the observed outcome falls in the critical region, then H_0 is rejected. The probability of such a result, or more extreme result, is considered to be too small ($\leq \alpha$) to be attributed solely to chance. Hence, H_1 (i.e., mean difference) would be statistically significant at the α or 100 α%) level. The second possibility is when the observed outcome falls in the acceptance region, then H_0 is accepted. The probability of such a result, or even more extreme, is greater than α. This, by the arbitrary definition established, is not a rare outcome if H_0 is true. Therefore, one would state that no statistically significant difference between the means had been demonstrated at the 100 α% level.

It is important to remember that a null hypothesis is not proved or disproved by a statistical significance test, but rather, precise probability statements may be made regarding the compatability of a set of observations with H_0. The true cause of observed differences in effects is almost always an inextricable combination of real differences and random processes. Although it would be useful to have the entire distribution of outcomes under H_0 available for every situation, the calculations are prohibitive. Therefore, tables have been prepared for most tests with critical levels for a few selected α levels. While the most commonly chosen values for α are 0.05 and 0.01, these values are arbitraty, and mainly for the sake of having concise tables, they have become standard. The choice of α will be discussed later in the section.

The outcome distribution discussion above is based on two-tailed tests (i.e., H_1: $\mu_1 \neq \mu_2$). However, in some cases, a set of restricted alternative hypotheses may be desired, those being H_1: $\mu_1 > \mu_2$ or H_1: $\mu_1 < \mu_2$. In these situations the entire 100 α% of the probability

is assigned to the appropriate single tail. The corresponding critical value is the basis for these one-tailed tests of H_0. When applying these one-sided tests, the critical value used is that for the entire α, as compared to two-tailed, where the α is divided in half (i.e., $\alpha/2$) and distributed to each tail.

When testing hypotheses in the manner described, two kinds of decision errors are possible. One, a true hypothesis may be rejected or, two, a false hypothesis may be accepted. These are referred to as Type I and Type II errors, respectively. These errors and the conditions under which they arise are shown in Figure 11. These errors occur with certain probabilities, since the decision to accept or reject H_0 is based upon the outcome of a random process. Type I error, or α, is equal to P (reject $H_0|H_0$ is true). Type II error, or β (not to be confused with β used in regression analysis), is equal to P (accept $H_0|H_0$ is false). The probability of Type I error is also referred to as the size of the critical region or as the level of significance of the test. Beta (β), the probability of Type II error, provides a more specific criterion for the assigning of critical regions for statistical significance tests. The general principle is: among all critical regions of the same size, use the one for which the probability of a Type II error is minimum. The probability of making a Type II error, β, depends upon the amount of real deviation from the null hypothesis. It is a function of the true value of the parameter being tested, and therefore usually cannot be evaluated. However, β can be lowered in three ways: 1) increase the sample size, 2) reduce experimental error, and 3) increase the permissible probability for Type I error (α). Decreasing β is equivalent to increasing 1 - β, the probability of rejecting an incorrect hypothesis.

Decision	True Situation in Population	
	Hypothesis (H_0) Correct	Hypothesis (H_0) Incorrect
Accept hypothesis (H_0)	no error probability = $1 - \alpha$	Type II error probability = β
Reject hypothesis (H_0)	Type I error probability = α	no error probability = $1 - \beta$

Figure 11. Errors in decisions based on tests of hypotheses.

The probability 1 - β is termed the power of the test to detect that a parameter differs from a specified value. Obviously, for any given test one would like to minimize β and maximize 1 - β. The curve described by 1 - β as a function of the magnitude to be detected (usually in σ units) is the power curve (Figure 12). This curve shows the probability of rejecting H_0 as a function of μ. It is interpreted in the following manner: as one moves away from μ (values under H_1) the probability of rejecting an incorrect hypothesis increases; as sample size increases, less of a change is necessary to develop the power. Hence, as stated before, increasing sample size will increase power, which decreases Type II error. This power curve is for a two-tailed test, they are usually expressed as one-tailed, where only half of the curve is presented.

Graphically, the relation between α and β is shown in Figure 13. The two distributions represent H_0 and H_1, with mean μ_o and μ_d (difference is mean that is being tested). The area under H_0 to the right of $\bar{Y}_c$ (critical level) is the rejection region. Values in this region are considered members of H_1 based on the arbitrary selection of an α level. However, there still is some probability (Type I error) that they belong to H_0. The area under H_1 to the left of $\bar{Y}_c$ is the probability of Type II error, β. While these values are not considered extreme enough to support H_1, there is some probability that they do belong to H_1. Thus, the two types of errors result from the overlap in the distributions of the two hypotheses, H_0 and H_1. If α was to be increased, then β would decrease (if μ_d did not change), which was stated earlier as a measure to decrease β. Also, if μ_d (i.e., H_1 is changed) is moved to the right, this would also cause a decrease in β. Both these situations are reversible and would cause increases in β.

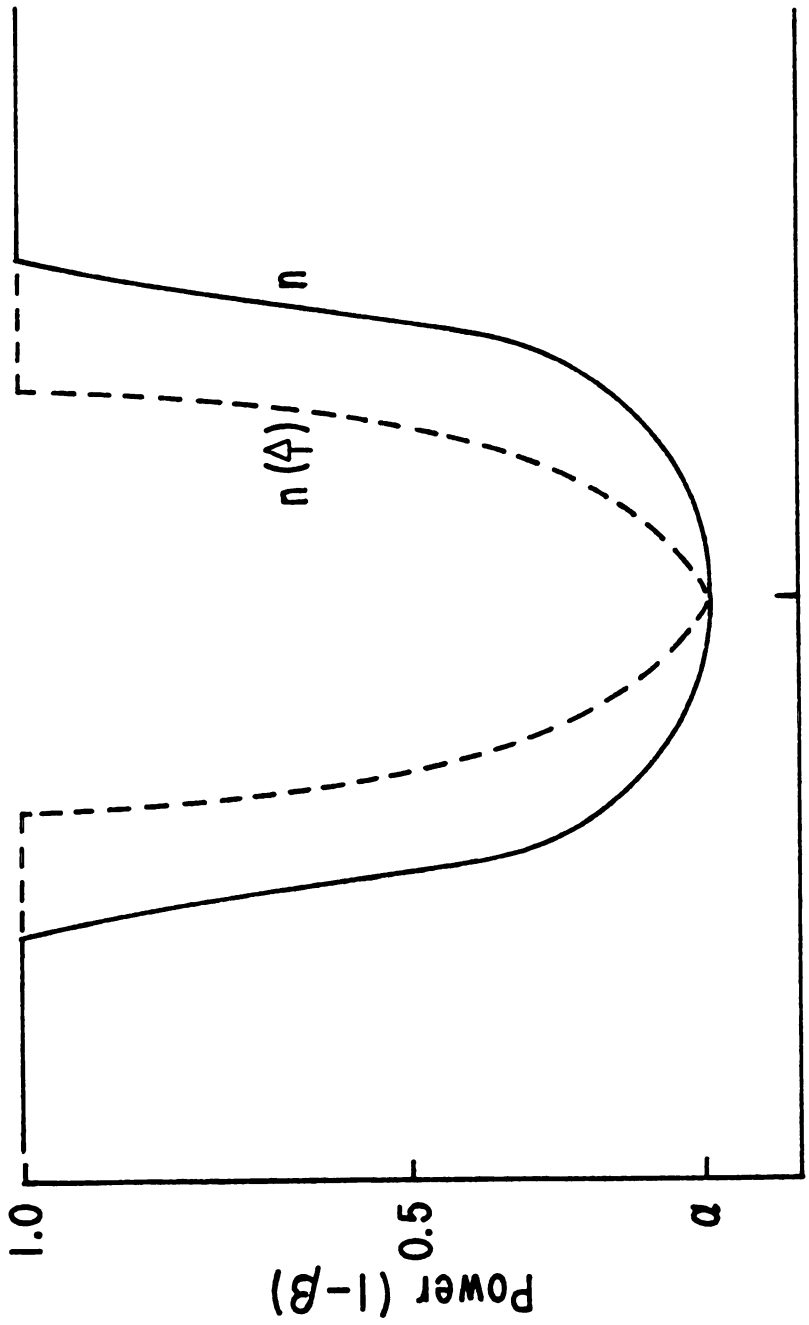


Figure 12. A typical power curve for n and for a larger n(↑).

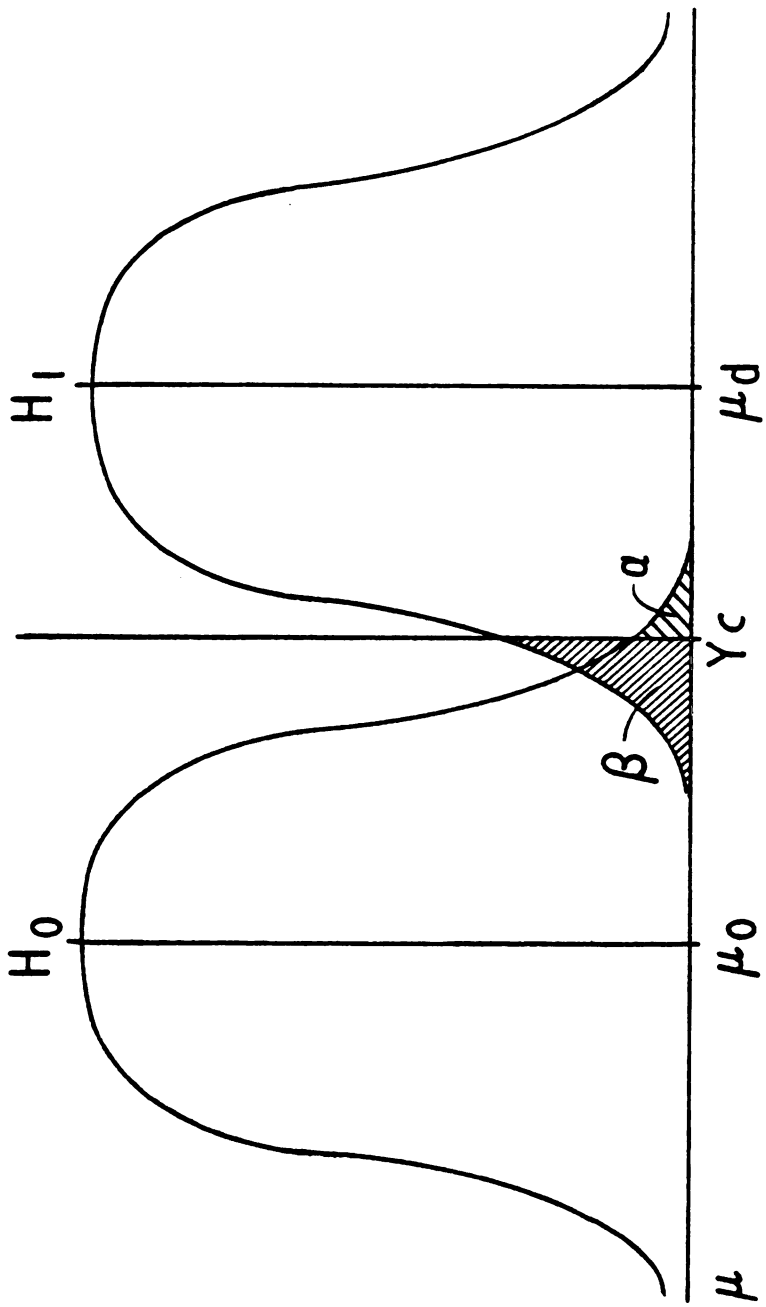


Figure 13. Relation between probability of errors, Type I error and Type II error

One would like to minimize both probabilities of making errors but usually only direct control of α, Type I error, is possible. In the past, it has been common only to report the statistical significance of results when $\alpha < 0.05$ or $\alpha < 0.01$ (the chance that a rejection of H_0 is less than 1 in 20 or 1 in 100, respectively). It is unfortunate that "nonsignificant" results, or results judged significant at values of α larger than 0.05, are often regarded as unpublishable. In many early experiments or fields of high variation and complexity, the probability of error $\alpha = 0.10$, 0.25 or even 0.50 may be acceptable.

Basically, research is an organized method for finding out what will not work as well as what will work. Those who object to this idea, fail to appreciate that most positive knowledge is gained through acquisition of negative knowledge. It is important not to place undue emphasis on the exact level of significance achieved in a particular instance. Too many scientists use the significance test for support instead of illumination.

CHAPTER IV

TREND DETECTION METHODS

The statistical process of detecting trends in lake water quality data is a step-wise procedure and is illustrated in Figure 14. The process starts with the need to determine if some lake water quality parameter has or is changing over time (hypothesis formation). Data to solve this question are then placed in a usable form so that appropriate statistical techniques can be applied. The data are then plotted to yield general information about the change in the parameter over time. Exploratory data analysis is conducted to provide the necessary information as to what statistical technique should be applied. Finally, appropriate statistical tests are applied to the data, to test the hypothesis of concern. Thus, quantitative information is generated concerning trends in lake water quality parameters. The time series data may also be modeled. The model can be used to develop the pattern(s) of change and for forecasting future events. While these steps appear discrete on the diagram, in practice, information from one phase may often be used in another. In the following sections each of these phases will be discussed.

Hypothesis Formation

The problem of asserting the existence or nonexistence of trends in lake water quality data may be treated in the classical hypothesis testing framework (discussed in Chapter III). In this case, the null hypothesis, H_0, is that there is no change (no trend) in the underlying

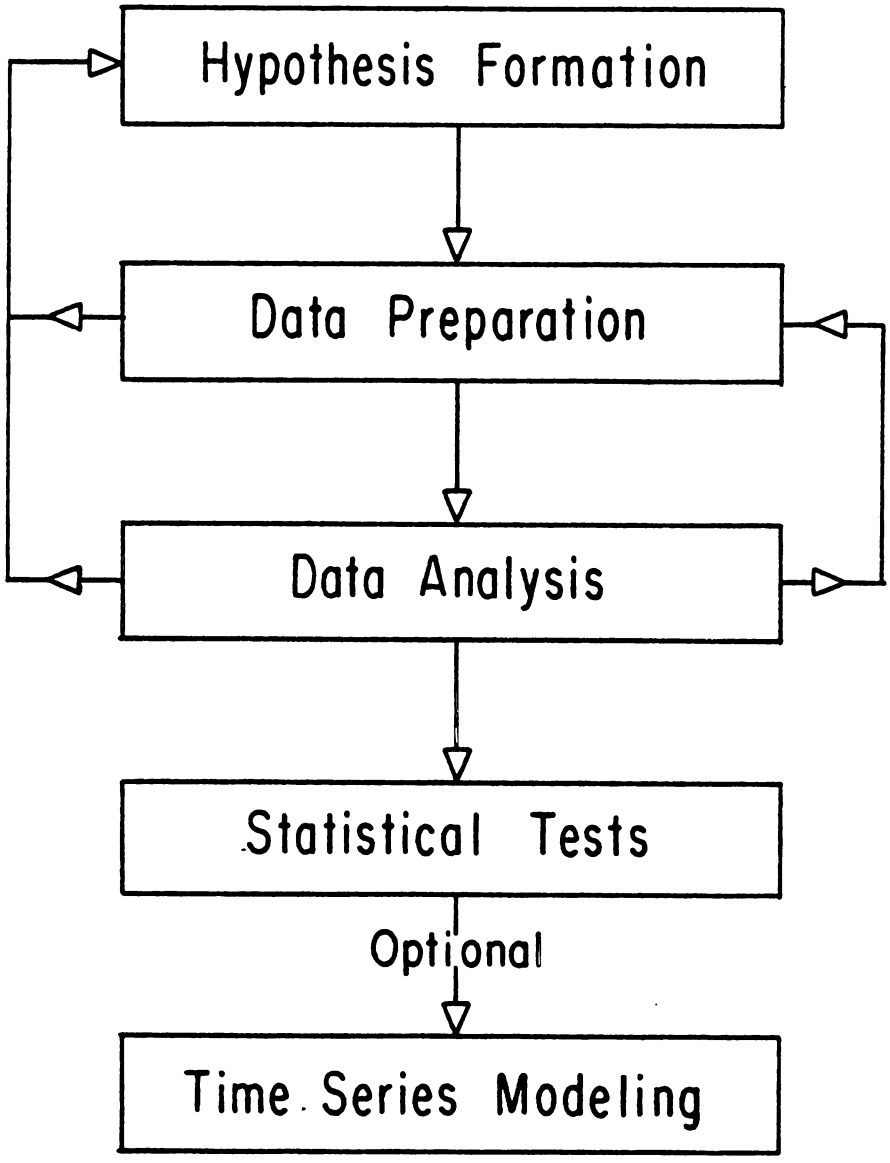


Figure 14 . Flow diagram for Trend Detection Methods.

population (lake) from which the data set (water quality parameter) was drawn. The alternative hypothesis, H_1, may be either a hypothesis that a trend does exist in the data (two-sided test) or that a positive (or a negative) trend exists in the data (one-sided test). If the concern is, "has there been a change in a given parameter," the power of a test is extremely important. The power (one of four possible outcomes, Figure 11) gives the probability, at a fixed confidence level, that the statistical test detects a trend of a specified magnitude, when in truth, there really is one. However, if the concern is of no change or trend, then the confidence interval $(1 - \alpha)$ is crucial, for it is the probability of concluding that H_0 is true, given H_0 is actually true in nature.

In developing a hypothesis there are three concerns: 1) selection of water quality parameter(s), 2) selection of a measure of central tendency, and 3) selection of one or two-tailed alternative hypothesis. The selection of a water quality parameter (discussed in Chapter II) will determine the limnological variable or process that information will be generated about. Hence, it is important to select the appropriate parameter that will provide the information about the question of concern. The selection of a measure of central tendency (discussed in Chapter III) is dependent upon: 1) what information is desired, and 2) the characterisitcs of the data and the distribution of the data. Ideally, one would like to have numerous observations in each time interval, so that the data's distribution can be analyzed and a measure of central tendency can be confidently chosen. However, in reality, one usually has limited data which were unevenly sampled through time using various methods. Hence, a subjective decision must be made, taking into consideration all the facets of each individual question, as to which

measure of central tendency should be used in a given situation. The third issue in selecting a null hypothesis is whether a one or two-tailed alternative is desired. If it is known that a parameter either increased or decreased, a one-tail H_1 should be used. A one-tailed test will maximize the probabilities of each outcome by placing all of the rejection region (α) at one tail of the outcome distribution. However, if the type of change is not known, a two-tailed should be used. It should be stressed, that when possible, the one-tailed alternative should be chosen. The construction of the two hypotheses (H_0 and H_1) may be simple or complex, depending on the question being asked and the data available. Hypothesis formation is an extremely important phase, for as with all statistical test, only the hypothesis stated is being tested. Hence, if this information is going to be used in planning and management one must be sure of what is being tested.

Data Preparation

Before any analysis can be conducted, the data must be prepared or placed in the desired form. By this, it is meant that the data be expressed as some measure of central tendency, such as the mean or the median. In order to apply trend detection techniques, there can be only one data point for each time unit. For example, when dealing with nitrogen changes over ten years, in yearly increments, only one data point for each year can be used. The data preparation problem arises when numerous observations are located in the same time unit, yet one value is needed to represent that discrete time unit. If the data are continuous throughout the interval and equally spaced, then any measure may be

directly calculated and used. However, when unequal sampling occurs over time, the data must be corrected. Two methods that can be used are regression analysis and polynomial models (regression is discussed in Chapter III, polynomial models are discussed later in this Chapter). For either, a model is constructed with the data for that time unit, and an estimate for the parameter is calculated at the middle of the time unit. Thus, one parameter value that represents that discrete time unit is generated. Data preparation, as in hypothesis formation, is highly variable and may be simple or complex depending on the situation. When dealing with numerous observations in a time unit the one value used for testing the null hypothesis should best represent the parameter in the fashion desired.

Data Analysis

Once a hypothesis is formed and the data are properly arranged (i.e., one data value per time unit) the data are ready to be explored and analyzed. The data analysis step will provide the necessary information to determine which method (statistical test) should be used to test the null hypothesis. This information may also be used to examine the hypothesis formation and data preparation steps. Thus, after examining the data, one may find it necessary to restate the null and alternative hypothesis and/or express the data in a given time unit in a different manner. As noted before, the conduction of the tasks in the first three sections of this chapter are interdependent and may cause reevaluation of one or two of the others. This same inter-relationship occurs within this section. While this step should follow the flow

diagram in Figure 14, it may be necessary, based on the outcome in a subsection, to change the flow. For example, if the data are non-normal when tested, one may wish to transform the data and retest for normality.

Graphical Techniques

The graphical techniques presented help provide a general knowledge about the parameter in the time series data. These plots provide: 1) a visual test for trends, 2) a check on validity of assumptions (normality, homogeneous variance), and 3) an examination of outliers. Although qualitative in nature, these graphs provide an important starting point for all data analysis in trend detection.

The first, and simplest, is the time series plot. The time series plot is a plot of the data (parameter) against time. These plots may be useful in visually detecting trends and should always be done. The plot may be expanded to include the standard deviation or quartile range around each value. Figure 15 shows a time series plot of the mean and standard deviation for mean annual ammonia concentrations in Lake Ontario from 1967-1977. From looking at the mean values only, one would expect a decreasing trend, but by including the standard deviation range, one may question the strength of the visual evidence. It may also be useful to plot both the mean and median (or other combination of measures of central tendency) to provide more information about the overall pattern and distribution of the parameter, in question, over time.

A slightly more sophisticated technique is the cumulative sum plot (Lettenmaier, 1977). The cumulatvie sum is simply the area under local mean to "j" minus a fraction (j/n) of the area under the global mean.

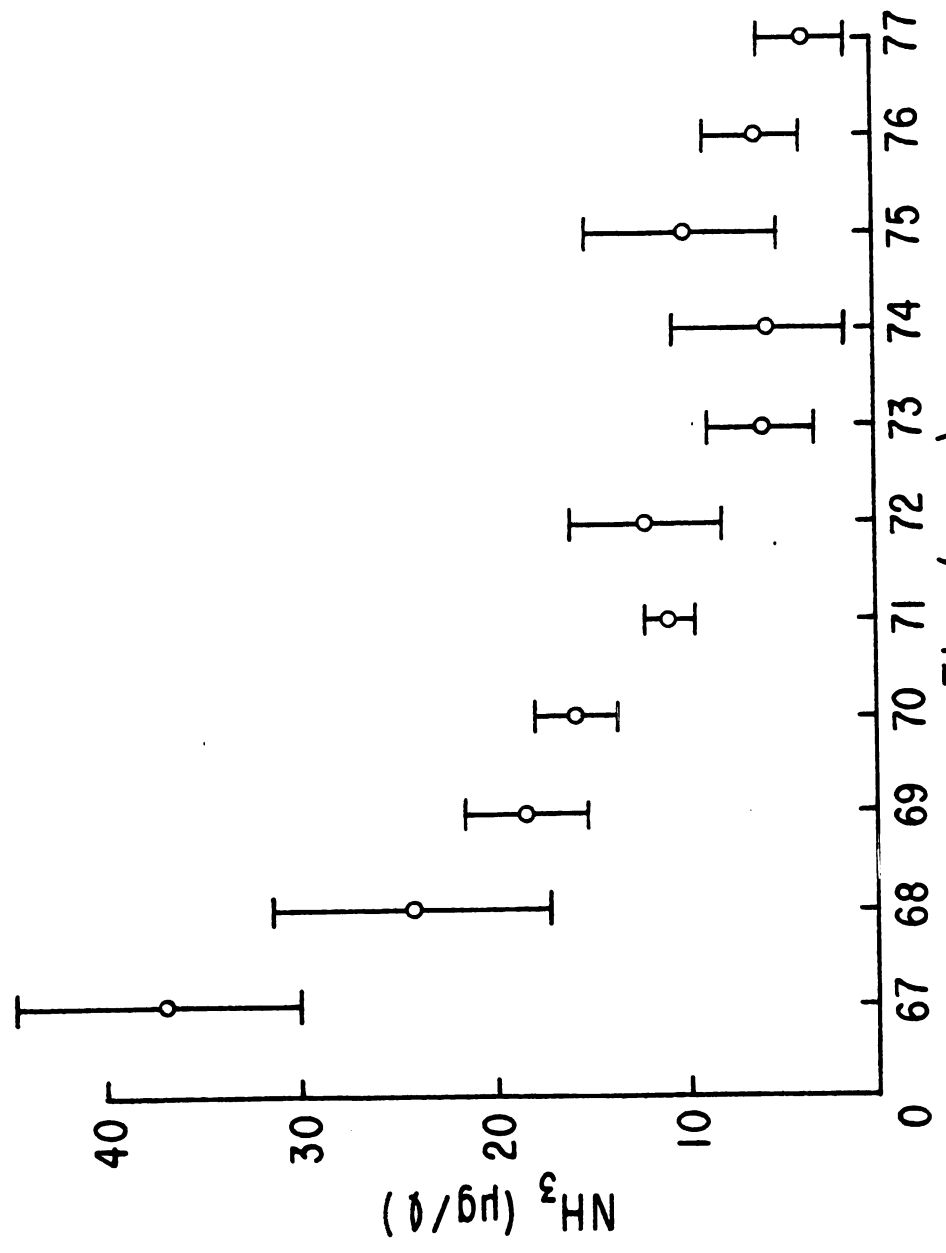


Figure 15. Time series plot of mean annual ammonia concentrations for Lake Ontario.

The cumulative sum is defined as

$$CU_j = \sum_{i=1}^{j} x_i - j/n \sum_{i=1}^{n} x_i \qquad (31)$$

where: $CU_j = j^{th}$ cusum point

$x_i = i^{th}$ data point

j = subinterval of n

n = sample size

The value for j is determined by the sample size (n) and the level of information desired. For example, if n = 100, one may choose j = 10 and thus 10 cusum points would be plotted. If more information (also more computational effort) is desired, a smaller value of j may be chosen. This provides more points and allows for a better representation of the trend in the data. There may be n - 1 cusum values calculated for any set of data. If the mean is constant (no trend), the cusum is zero; if there is a step change in the mean level, the cusum is linear; if a linear change exists, the cusum is quadratic. These three possibilities are shown in Figure 16, with t being the trend starting point.

The histogram (frequency plot) is a useful technique for examining the normality of the data's distribution. A normal distribution will have a bell-shaped curve. Also, the mean, median, and mode should all be identical. Two statistics that describe the shape of distributions are skewness and kurtosis, the third and fourth moments, respectively, Skewness, or asymmetry, means that one tail of the frequency curve is drawn out more than the other. Curves are skewed to the right or left, depending upon whether the right or left tails are drawn out. Kurtosis

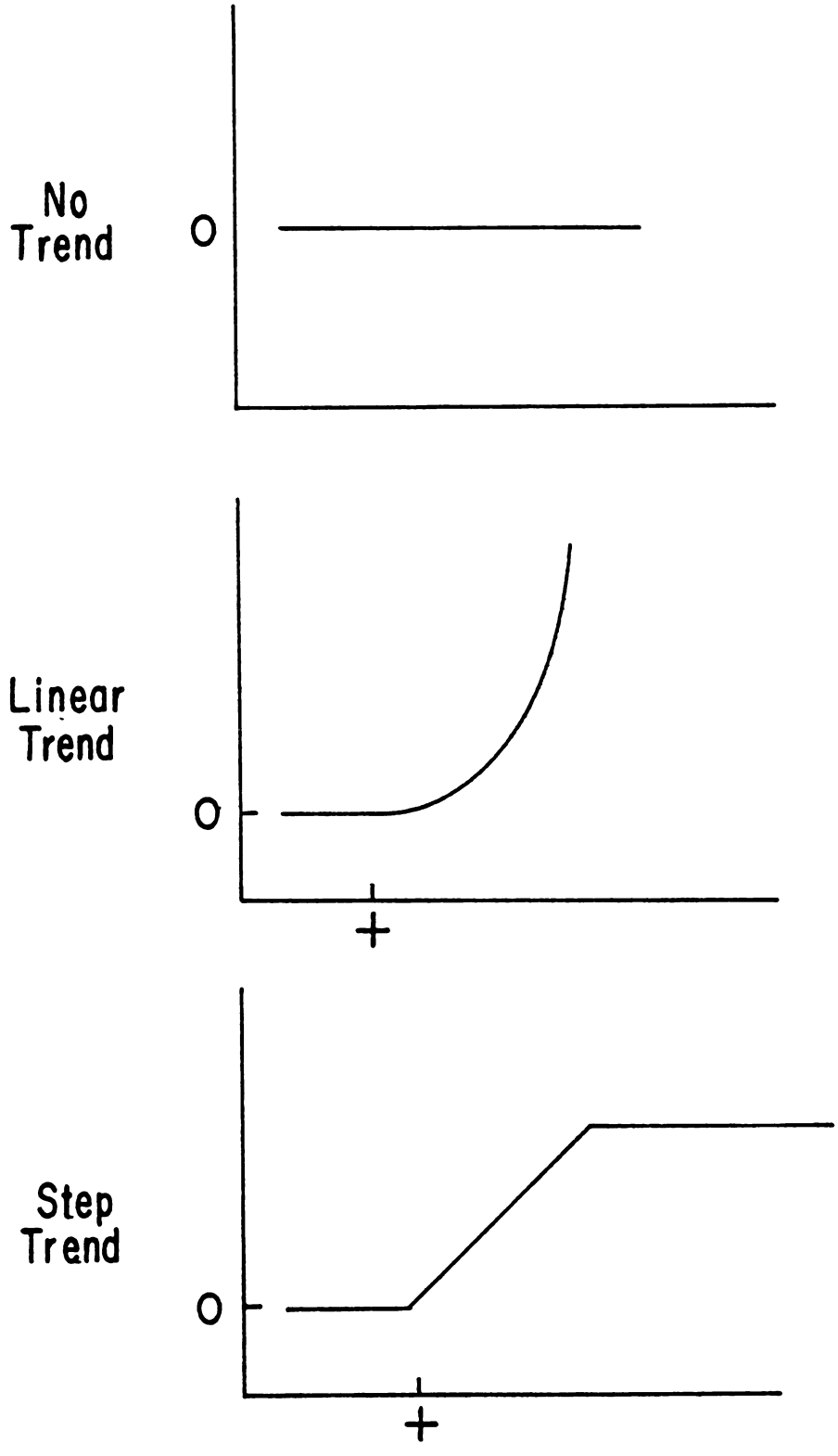


Figure 16 . CUSUM plot of step, linear, and no trend.

is a measure of peakedness of the distribution. A leptokurtic curve, higher peak than normal curve, results from having more data near the mean and at the tails and fewer data in the intermediate. Platykurtic, flattened curves, have more data in the intermediate area and less at the mean and in the tails. Figure 17 shows a normal curve along with skewed right, skewed left, leptokurtic, and platykurtic distributions. The calculation of skewness and kurtosis statistics is complicated and one should consult a statistical textbook for the procedure (e.g., Sokal and Rohlf, 1969).

The sample residuals (errors) may be plotted to check the validity of assumptions on homogeneous variance (Behnken and Draper, 1972). The sample residuals are estimates of the random errors and may be calculated from:

$$\hat{e}_i = X_i - \bar{X} \tag{32}$$

where: $X_i = i^{th}$ data point

$\bar{X}$ = mean of all X's

For use with regression analysis, the sample residuals are computed from:

$$\hat{e}_i = (Y_i - \bar{Y}) - b_1 (X_i - \bar{X}) \tag{33}$$

where: $Y_i = i^{th}$ Y observation or data point

$\bar{Y}$ = mean of all Y's

b_i = estimate regression line slope

A plot of e_i against X_i or e_i against $\hat{Y}_i$ (for regression) should produce a horizontal band of points. If the band of points increases or decreases in width, then heterogeneous variances usually are indicated.

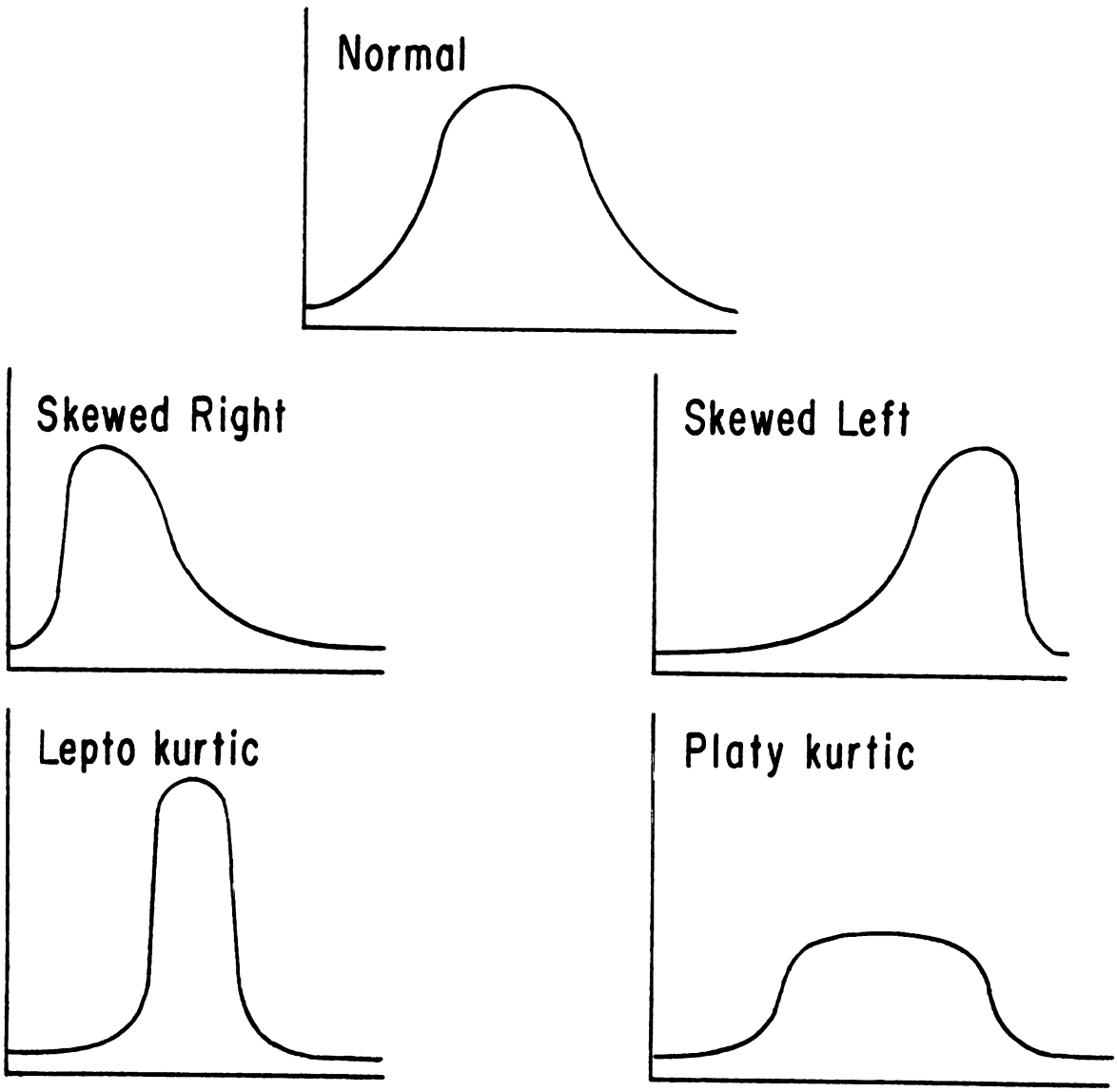


Figure 17. Normal, skewed, kurtotic distributions.

If the band of points is fairly uniform in width but is not horizontal, transformations of the addition of higher degrees of polynomials to the model may be necessary. If an unusually large e_i is observed ("outlier"), the corresponding X_i or Y_i should be identified to discover the possibility of a recording error or an unusual condition in the system. Techniques for examining "outliers" are discussed later in this section. While these techniques may lead one to identify a so-called outlier, the exclusion of data should occur only when concrete evidence for removal is present and not only because of test statistics or plotting techniques suggest it.

For sample sizes less than 50, these plotting techniques may not provide enough information to make proper decisions. Even with large sample sizes, the statistical tests which follow should be utilized to provide the maximum assurance that the assumptions required for the usage of the statistical tests are not violated.

Distribution Tests

An assumption central to most parametric statistical tests is that the data come from a population that exhibits a normal distribution. This normality assumption is relaxed by nonparametric statistics, which require only a continuous distribution. In the previous section, graphical techniques were used to provide qualitative information on the form of the data's distribution. This section provides a statistical test of the assumption on normality. Two important considerations of this test are: 1) the test determines only whether the data is nonnormal and not whether it is normal (i.e., proves the alternative hypothesis and not the

null hypothesis) and 2) when the sample size is small it is extremely difficult to reject the null hypothesis even if the data grossly deviates from normality or symmetry. Therefore, when dealing with samll sample sizes, one should use an appropriate statistical test and also intuitive knowledge of the parameter and the water system in determining whether to use parametric or nonparametric statistical tests for trend detection.

The W-test is used to evaluate the assumption that a sample has a normal distribution for samples $\leq$ 50 (Shapiro and Wilk, 1965). The W-test has been shown to be an effective technique for evaluating the assumption of normality against a wide spectrum of deviations from normality such as skewness and kurtosis of the distribution (Atkinson, 1967; Chen, 1971).

The W-test is used as follows:

Step 1. Rearrange the data to obtain the ordered array $X_1, X_2, \ldots, X_n$, where $X_1 \leq X_2 \leq \ldots \leq X_n$.

Step 2. Compute the variance (s^2) of the data (Equation 11 and from previous chapter).

Step 3. If n is even, set k = n/2; if n is odd, set k = (n - 1)/2. Then compute

$$b = \sum_{i=1}^{k} a_{n-1+1} (X_{n-i+1} - X_i) \tag{34}$$

where the values of a_{n-i+1} for i = 1, ..., k are given in Table 1 of Appendix B for n = 3, ..., 50.

Step 4. Compute

$$W = b^2 s^2 \tag{35}$$

If the test statistic W is less than the critical value $W_{\alpha, n}$ from Table 2 of Appendix B, one may reject the hypothesis of normality with probability of Type I error less than α.

For large sample sizes (n > 50) the Kolmogorov-Smirnov test is suggested (Siegel, 1956; Sokal and Rohlf, 1969).

A quick and easy measure that displays the general shape of the data's distribution is the mode to mean ratio (R) (Springer and Gifford, 1980). It is calculated using the following:

$$R = [1 + (C.V.)^2]^{-3/2} \qquad (36)$$

where: C.V. = coefficient of variation (Equation 13)

If R < 1, the distribution is skewed right. If R = 1, the distribution is normal or symmetric. If R > 1, the distribution is skewed left. This should not be substituted for a statistical test for normality, but may be used as a quick check technique.

Homogeneous Variance

In the biological sciences there is a common tendency toward positive correlation of the mean and variance over a wide range of a specific variable (i.e., groups with large means tend to have large variances and those with smaller means have smaller variances (Gill, 1978)). A check of the assumption of homogeneous variance is done by testing the hypothesis, H_0: $\sigma_1^2 = \sigma_2^2 = \ldots \sigma_t^2$, where t is the number of treatments (in this case, the number of time units). The first test described is a quick and easy method, while the second is more efficient

and capable of dealing with uneven numbers of data within the different time units.

Hartley's f_{max} test (Hartley, 1950) involves the ratio of the largest to the smallest of the variances within groups.

$$f_{max} = s^2_{max} \Big/ s^2_{min} \tag{36}$$

where: f_{max} = test statistic

s^2_{max} = maximum variance of all treatments (time unit)

s^2_{min} = minimum variance of all treatments (time unit)

If the test statistic exceeds the critical value, $f_{max,\ \alpha,\ t,\ v}$ (from Table 3 of Appendix B), for t groups with v = r - 1 df (r equals number of data points within each time step or unit), one may reject H_0. If replication within each group is only slightly unequal, the critical value should be corrected by using $v = (\Sigma r_i/t)-1$.

Bartlett's test (Bartlett, 1937) for homogenity of variances is used because it provides good efficiency and handles uneven replication within groups.

The following is the procedure for Bartlett's test:

Step 1. Convert the variances for X's to logarithms. It is convenient to add 1 to the data when zeros are present and multiply by 10, or a power of 10, to avoid technical problems with logarithms.

Step 2. Compute

$$X^2 = 2.3026 \left[\sum_{i=1}^{a} (n_i - 1) \log s^2 \right] - \left[\sum_{i=1}^{a} (n_i - 1) \log s_i^2 \right] \tag{37}$$

where: a = number of X's

s_i = i^{th} variance

$$s^2 = \frac{\sum_{i-1}^{a} (n_i - 1)\, s_i^2}{\sum_{i=1}^{a} (n_i - 1)}$$

The constant 2.3026 transforms the common logarithms to natural ones, which are needed for the formula. If this value x^2 exceeds $s^2_{\alpha,\ a-1}$ from Table 4 and Table 5 of Appendix B, one may reject H_0. If the computed s^2 is nearly equal to the critical value, a correction factor should be used (see Sokal and Rholf, 1969).

Examination of Outliers

If it can be determined that an extreme observation is a result of errors in recording data, parameter value determination, or foreign to the defined population the outlier may be excluded. Otherwise, the observation should be used in the computations, unless one has adopted some consistent procedure of testing outliers or of censoring or "Winzorizing" (replacing an extreme value by the next most extreme observation) all extreme data (Tukey, 1962). However, arbitrary exclusion of data that do not conform to one's perceived ideas can lead to serious bias in estimation as well as underestimation of experimental error.

The procedures of Grubbs and Beck (1972) may be used as an approximate test of whether the magnitude of the largest residual is so large, that the corresponding value may be excluded. The procedure is as follows:

Step 1. Compute residuals

$$e_i = X_i - \bar{X}$$

Step 2. Select largest residual (e_L), either

$$|e_{max}| \quad \text{or} \quad |e_{min}|$$

Step 3. Compute test statistic

$$A = e_L/s \tag{38}$$

where: A = test statistic

s = standard deviation

If the test statistic exceeds the critical value $A_{\alpha, n}$ from Table 6 of Appendix B, one may exclude the corresponding data point. One should chose a small Type I Error ($\alpha = 0.01$) for rejection of an outlier, because a Type I Error would result in the exclusion of an extreme but valid observation which would lead to a biased estimate of σ^2 (too small) and distortion of confidence statements. While a Type II Error would lead to an inflated estimate of σ^2 and conservative confidence statements about means.

Graphical methods, statistical tests, and intuitive knowledge must be used in considering whether to remove an outlier. When outliers are removed one must have: 1) sound evidence, 2) be capable of supporting such a removal, and 3) held responsible for the removal. One final point when dealing with outliers is that in some cases an extreme value or outlier may be more interesting than the rest of the data.

Time Dependency of Data

A very important issue when dealing with time series data in lakes, especially those with long detention times, is time correlated data (Yule, 1921; Bartlett, 1935). This time correlation arises from the fact that the value of a data point is dependent, to some degree, upon

data points previous to it in time. This time correlation of data is termed autocorrelation or serial correlation. In a lake system, an annual concentration will depend on the concentration in the lake in previous year, or years, to some degree. Thus, those lakes with long detention or residence times, (i.e., not much change in yearly water composition) will usually exhibit autocorrelated data. Autocorrelated data means that the random error terms are correlated over time. If autocorrelation is present, the assumption that the random error terms (e_i) are independent is violated.

Consider the time series of data X_1, X_2, ..., X_n, the most satisfactory estimate of the kth lag autocorrelation (ρ_k) is (Jenkins and Watts, 1968; Box and Jenkins, 1968)

$$r_k = \frac{C_k}{C_o} \tag{39}$$

where: r_k = estimate of ρ_k

C_k = estimate of Y_k (autocovariance)

C_o = estimate of Y_o

An estimate of Y_k, autocovariance of the kth lag is

$$C_k = \frac{1}{n} \sum_{i=1}^{n-k} (X_i - \bar{X})\,(X_{i+k} - \bar{X}) \tag{40}$$

An estimate of Y_o is

$$C_o = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2 \tag{41}$$

The Durbin-Watson test can be used to test for autocorrelation, (i.e., H_o: $\rho = 0$) (Durbin and Watson, 1951). The Durbin-Watson statistic d is defined in terms of the observed residuals by

$$d = \frac{\sum_{i=2}^{n} (e_i - e_{i-1})^2}{\sum_{i-1}^{n} e_i^2} \qquad (42)$$

where: e_i = i^{th} residual

The d statistic is compared to upper (d_u) and lower (d_u) bounds (Table 7 in Appendix B). If d falls below d_L, the hypothesis that the original residuals are uncorrelated is rejected, and if it falls above d_u, the hypothesis is accepted. When d falls between d_u and d_L, the test is inconclusive and further statistics are necessary (see Durbin and Watson, 1971). The formula above tests for positive autocorrelation, which is by far the most common in lakes. When testing for negative autocorrelation, the test statistic, 4 - d, is used. The remaining part of the test is conducted in the same manner as for positive autocorrelation, where if $4 - D < d_L$, one concludes negative autocorrelation ($\rho < 0$) exists. The two-sided test is done by employing both of the one-sided tests separately. Thus, the Type I Error for a two-tailed test is 2α. For more information on other methods of testing for autocorrelation, consult Kenkel (1975) and Sen (1978, 1979).

Transformations

To correct for violations of the assumptions detected by analyzing the data, transformations of the data may be used. Transformations are valid because the linear scale has a relationship with other scales of measurement. In essence, the scale of measurement is arbitrary, therefore, one may transform data to a different scale of measurement so the transformed variates more closely satisfies the assumptions. A

fortunate fact of transformations is that often several departures from the assumptions are simultaneously cured by the same transformation to a new scale. For example, simply by correcting for nonnormality, homogeneous variance may also be obtained.

When a transformation is applied, tests of significance are performed on the transformed data. When the transformations are nonlinear (e.g., log, square root), confidence limits computed in the transformed scale changed to the original form are asymmetrical (Sokal and Rohlf, 1969). However, the use of the standard error in the original scale is misleading and should not be used with transformed data. The log and square root transofrmations are discussed below.

The most common of all transformations is the log transformation. The log transformation consists of converting all the data into logarithms, usually common logarithms. There are two situations where the log transformation is quite effective. First, is with skewed right distributions (very common in lake parameters). The log transformation of skewed data will usually result in a normal distribution. The second situation is when the mean is positively correlated with the variance (i.e., large means have large variances). The variance should become independent of the mean when the data undergoes a logarithmic transformation. When dealing with log transformation, two helpful hints are: 1) if the data includes zero, add 1 to all values to avoid the log of zero (i.e., negative infinity) and 2) if there are data between 0 and 1, multiply all values by 10 or some power of 10, to avoid the negative values that will result.

For stronger relationships of mean to variance (such as the mean directly proportional to the variance) or for skewed distributions, the

square root transformation may be appropriate for obtaining approximate normality as well as homogeneous variance. When the data include zero's, one-half should be added to remedy the problem of trying to take the square root of zero.

Other transformations are available, for example reciprocal, arcsine, and probit. For a further discussion of these and the subject of transformations, consult Tukey (1962), Sokal and Rohlf (1969), Neter and Wasserman (1974), and Mosteller and Tukey (1975).

A final form of data transformation is the process of removing seasonality in the data. This is only necessary when dealing with data in less than yearly time units. The commonest methods of removing seasonality are differencing and modeling. The differencing technique is used by forming a grand mean for each identical time unit (i.e., all the same month, all the same season) and form a new time series as the difference between the raw time series and the corresponding grand mean. Durbin and Murphy (1975) and Cleveland and Tiao (1976) consider the use of an additive-multiplicative model and ARIMA model, respectively. For more information on the seasonal adjustment of data, consult Wallis (1974).

Statistical Tests for Trend Detection

The final step in the trend detection method is to compute a test statistic using an appropriate statistical test. This test statistic is compared to a table of critical levels and the strength of evidence (i.e., level of alpha that rejects H_0) is determined. Classically, the alpha level has been chosen apriori (i.e., before the data are examined).

In the case of detecting trends in lake parameters the concern is with the degree of certainty that the data exhibit some type of trend. Thus, the alpha level that provides acceptance of H_1 (i.e., that a trend exists) should be used to provide the necessary quantitative information needed for planning, policy and management decisions. Three cases of data are possible, based on type of distribution and time dependency, and will determine which statistical test to be used. The three cases are:

1. Normal and Independent Data
2. Symmetric and Independent Data
3. Dependent Data

For each case, statistical tests are discussed to allow computation of the necessary statistics to test for the presence of trends.

Normal and Independent Data

All classical parametric statistics require that the data come from a normal population and are independent of each other (i.e., independent in time for trend analysis). However, when dealing with lake systems the validity of these assumptions is highly questionable. Many lake parameters will exhibit skewed distributions which may be converted to normality through transformations of the data. A more important issue is the small sample size of data usually available for trend detection. As discussed before, it is difficult to disprove normality on small data sets, even when the data grossly deviate from normality. The issue of autocorrelation in lake data is usually more of a problem than is normality. Since most lakes do not renew their total volume of water very

quickly, one data point collected in time will be dependent upon the previous data point in time (to some degree). Since the majority of lake parameters are autocorrelated over time, the assumption of independent data is violated. The violation of the independent data assumption has been almost completely ignored in past lake trend analyses. While it is unlikely that both of these assumptions would be valid, especially independence, it may be possible in some situations. Therefore, a statistical test is provided for this case. If doubt exists about these assumptions, one should use a test from one of the other cases that do not require the assumptions.

When testing for the difference between means of normal variables, four cases, for which different methods are needed, may be distinguished: 1) known variances or large sample size, 2) unknown but equal variances, 3) unequal variances but equal coefficients of variation, and 4) unequal variances and coefficients of variation (Gill, 1978). If the variances of the populations are known from prior research or the sample size is large ($n > 200$), case 1 should be used. Case 3 and 4 are less powerful than case 2 and are needed only when variances cannot be made equal through transformations. Since cases 1, 3, and 4 are rare, only the second case will be discussed here. For case 1, 3, and 4 a statistical textbook should be consulted (e.g., Gill, 1978; Sokal and Rohlf, 1969).

Case 2 is used when the variances are unknown but the hypothesis of equal variances has been accepted (the test for homogenious variance is in an earlier section). The statement that the variances are unknown, implies that the true population variance is not known, while a variance estimate, calculated from the sample data, is known, and used to test for equal variances. The variances discussed here are for the two

populations that are designated by the null hypothesis and not the individual variances for each time unit, as was tested for in the homogeneous variance section.

This case involves the use of the well known two-sample t-test. The test statistic is computed from

$$t = (\bar{X}_1 - \bar{X}_2) \quad [s \quad (1/n_1) + (1/n_2)] \tag{43}$$

where: $\bar{X}_1$ = mean of population 1

$\bar{X}_2$ = mean of population 2

s = standard deviation of population 1 and 2

n_1 = sample size of population 1

n_2 = sample size of population 2

The standard deviation (s) for the total population (μ_1 and μ_2) is

$$s = \frac{\left[\sum_{i=1}^{r_1}\left(X_{1i}^2\right) - \sum_{i=1}^{r_1}\left(X_{1i}^2\right)\Big/ r_i\right] + \left[\sum_{i=1}^{r_2}\left(X_{2i}^2\right) + \sum_{i=1}^{r_2}\left(X_{2i}\right)^2 r_2\right]}{(r_1 + r_2 - 2)} \tag{44}$$

The test statistic is compared to the critical values (Table 8 of Appendix B):

$$+ t_{\alpha,\ r_1 + r_2 - 2} \text{ for } H_1: \ \mu_1 > \mu_2$$

$$- t_{\alpha_1\ r\ \frac{1}{2}\ r\ -\ 2} \text{ for } H_1: \ \mu_1 > \mu_2$$

If the test statistic exceeds tha appropriate critical value, the null hypothesis is rejected and a trend is indicated with the probability of a Type I Error less than α.

Symmetric and Independent

By relaxing the assumption of normal distribution to symmetric distribution, while still assuming independent data, nonparametric tests may be used. Nonparametric tests should also be used if the data have a small sample size to bypass the assumption of normality. Again, it should be stressed that assuming independent data in lake systems is extremely questionable. The nonparametric tests suggested are the Mann-Whitney and Spearman Rho, and these should be applied to step and linear trends, respectively.

The Mann-Whitney test is one of the most powerful of the nonparametric tests, especially when dealing with step changes between the populations (Conover, 1971; Siegel, 1956). The data consist of two random samples. Let $X_1, X_2, \ldots, X_{n_1}$ denote the random sample of size n_1 from population 1, and let $Y_1, Y_1, \ldots, Y_{n_2}$ denote the random sample of size n_2 from population 2 (defined by null hypothesis). The test statistic is calculated in the following manner.

Step 1. Rank the combined sample of X's and Y's from smallest to largest. For ties, take average of ranks that would have been assigned to them had there been no ties.

Step 2. Sum the ranks for the X's

$$s = \sum_{i=1}^{n} R(X_i)$$

Step 3. Compute the test statistic

$$T = s - \frac{n_1(n_1 + 1)}{2} \qquad (45)$$

The lower critical values $W_{\alpha,\ n_1,\ n_2}$ are found in Table 9 of Appendix B. Substitute n_1 for n and n_2 for m in Table 9. The upper critical level is computed by:

$$W_{1-\alpha} = n_1 n_2 = W_{\alpha} \tag{46}$$

An alternative to using upper critical values, the statistic T', defined as

$$T' = n_1\ n_2 - T \tag{47}$$

may be used with the lower critical values whenever an upper tailed test is desired.

The following decisions rules should be used for their respective alternative hypothesis.

H_1: $\mu_1 < \mu_2$ Reject H_0 when T is less than W_{α}

H_1: $\mu_1 > \mu_2$ Reject H_0 when T is greater than $W_{1\ -\ \alpha}$ or if T' is less than W_{α}

When testing for continuous (i.e., linear) trends, the Spearman Rho test is suggested (Conover, 1971; Siegel, 1956). The data X_1, X_2, ..., X_n is used to compute a test statistic:

$$T = \sum_{i=1}^{n} (R\ (X_i) - i)^2 \tag{48}$$

where: $R(X_i)$ = rank of the i^{th} observation X_i in the sample size of n

The null hypothesis is rejected if T is less than $W_{\alpha,\ n}$ (H_0: $\mu_1 \geq \mu_2$) or if T is greater than 1 - W_{α} (H_0: $\mu_1 < \mu_2$), with W_{α} values found in Table 10 of Appendix B. The quantities of the Hotelling-Pabst Test statistics are used in Table 10.

Dependent Data

When dealing with lake parameters over time, time dependency is almost certain to occur. This autocorrelation leads to violation of the assumption of independent data. The sensitivity of statistical procedures to violations of the independence assumption is well known (Gastwirth and Rubin, 1971, 1975; Serfling, 1968). Thus, some procedure is needed if quantitative information is desired about trends in lake parameters over time. Two similar approaches are those of Lettenmaier (1976, 1977) and Sen (1963, 1965). The test statistics proposed by Sen were shown to be asymptotically normal for large sample sizes ($n > 1000$), but they were found unsuited for application to small and medium sample sizes (Lettenmaier, 1976). The method suggested here is that proposed by Lettenmaier (1976, 1977). In this method, the test critical levels are corrected based on the degree of dependence. The correction values were generated using Monte Carlo simulation. The data dependence is assumed to be of the lag one Markov type. Test rejection levels were generated based on ρ, n, α, and trend magnitude and compared to the independent case. For a more detailed discussion of the generation of the correction values consult Lettenmaier (1976, 1977).

When testing the trend hypothesis with dependent data, the following procedure should be used.

Step 1. Calculate the test statistic from data using the Mann-Whitney or the Spearman Rho test for step or linear trend hypothesis, respectively.

Step 2. The modified critical level is calculated as the upper or lower rejection level, plus a scaled difference:

$$W'_{\alpha_L} = L_1 + f(n, \alpha, \rho)\ (W_{\alpha_L} - L_1) \tag{49}$$

$$W'_{\alpha_u} = L_u + f(n, \alpha, \rho)\ (W_{\alpha_u} - L_u) \tag{50}$$

where: W'_{α_L} = modified lower critical level

W'_{α_u} = modified upper critical level

L_1 = lower bound of the test statistic of outcome distribution

L_u = upper bound of the test statistic of outcome distribution

$f(n, \alpha, \rho)$ = correction value from Table 11 or 12 from Appendix B for Mann-Whitney or Spearman Rho test, respectively

W_{α_L} = lower critical level for independent data

W_{α_u} = upper critical level for independent data

For both, the Mann-Whitney and Spearman Rho, the lower bound of the test statistic (L_1) is zero. The upper bound (L_u) is $[n(n^2 - 1)]/3$ for the Spearman Rho and $n_1\ n_2$ for the Mann-Whitney, where n_1 and n_2 are the number of data in the first and second partitions, respectively. The α levels used in $f(n, \alpha, \rho)$ to determine the correction value from Tables 11 and 12 of Appendix B are for a two-sided test. If a one-sided test is desired at significance level α, a value of 2α should be entered into $f(n, \alpha, \rho)$.

Step 3. The modified critical level is compared to the test statistic calculated in Step 1, with H_0 accepted or rejected under the same circumstances as for independent data.

The power associated with the data may be calculated for the Mann-Whitney and Spearman Rho by using an equivalent number of independent samples, based on the autocorrelation of the data, in the normalized power function. The procedure is as follows:

Step 1. Calculate the number of equivalent independent samples based on lag 1 Markov dependence (Matalas and Langbein, 1962).

$$\frac{1}{n^*} \quad \frac{1}{n} \; 1 + \frac{2(\rho^+ - \rho^{2+})}{(\rho^+ - 1)^2} \tag{51}$$

where: n^* = number of equivalent independent samples

n = sample size

ρ = lag 1 correlation coefficient

t = sampling interval ($t = 1$ for equal spaced samples)

Step 2. Calculate the trend number:

$$N_T = \frac{T_r (n^*)^{1/2}}{2\sigma} \tag{52}$$

where: N_T = trend number (dimensionless)

T_r = trend magnitude (for Step Trend $T_r = \hat{\mu}_1 - \hat{\mu}_2$)

σ = standard deviation of time series (s, an estimate of σ should be used)

Step 3. Calculate the power:

$$1 - \beta = F(N_T - t_{1 - \alpha/2, V}) \tag{53}$$

where: $1 - \beta$ = power of test

F = cumulative distribution function of a standard student's t distribution

$t_{1 - \alpha/2}$ = critical level of student's t distribution at probability level $1 - \alpha/2$ (two-tailed test)

V = degrees of freedom $(n - 2)$

Table 8 of Appendix B lists the values for the Student's t distribution.

The procedure for calculating the power of the Spearman Rho test, is the same as for the Mann-Whitney, except N_T' is substituted for N_T, where

$$N_T' = \frac{T_r \left[n^* (n^* + 1) (n^* - 1)\right]^{1/2}}{(12)^{1/2} n^* \sigma} \tag{54}$$

where: T_r = trend magnitude (for linear trend $T_r = \Delta\hat{\mu}$)

One problem with these procedures is that the standard deviation and lag 1 correlation coefficient are assumed to be known. When sample size is greater than 30 and data are independent, s may be substituted for σ with little sample error (Lettenmaier, 1977). However, sample error in the lag 1 correlation coefficient estimation is of more concern, since errors in ρ will effect the estimation in the equivalent independent sample size (n^*) and the variance (Bayley and Hammersley, 1946). These errors will be translated into uncertainty in the estimation of N_T or N_T'. Thus, the problem of uncertainty in the estimations of the parameters (n^*,s) should be considered when the results are being used for planning, policy and management of lake ecosystems.

Time Series Modeling

After completing the trend detection phase, one may wish to model the time series data to provide information on the pattern of change over time and to forecast future events. The subject of time series modeling has become increasing important and it's use is expanding. The attempt here is to provide basic information only and not an indepth study. The range of books, for those interested in a more complete review, include: 1) introductory texts (Kendall, 1973; Chatfield,1975), 2) spectral analysis (Jenkins and Watts, 1968; Bloomfield, 1976), 3) univariate forecasting (Box and Jenkins, 1970; Granger and Newbold, 1977), and 4) a general review article on time series analysis (Chatfield, 1977). Only the basic method of fitting polynomials and the use of linear regression analysis will be discussed here.

A general model for the deterministic (trend) component is the polynomial model. Polynomial curves are convenient approximations for nonlinear relationships. The polynomial function is of the general form:

$$\bar{Y} = a + bX + cX^2 + dX^3 + \ldots \qquad (55)$$

where: Y = dependent variable

X = independent variable

a, b, c, d = coefficients

The polynomial involving only the terms of X and X^2, will yield a parabola with one inflection point. As increasing powers of X are used, the curve becomes more and more complex and will fit data increasingly well. However, each added power to X, one degree of

freedom is lost and the test of significance is harder. For most work, a cubic polynomial is the upper limit of degree. The aim of a polynomial model is to obtain a better fitting regression to a set of points. This is done by adding a quadratic term (X^2) to the regression equation and observing if a significant portion of the residual sum of squares is removed. The same is done for cubic and possibly higher terms. The numeric procedures of calculating third or higher degree polynomials usually involves matrix inversion and the use of a computer. For a second degree polynomial, Steel and Torric (1960) or Sokal and Rohlf (1969) should be consulted for the necessary formulas.

The methods for developing a linear regression model was discussed in Chapter III. The major problem with the use of linear regression on water quality time series data is the violation of the assumptions for the use of linear regression. As has been discussed, most water quality parameters are autocorrelated. Thus, the assumption of independent data is violated. While the model itself may not be able to statistically test for a trend, the pattern or shape of the time series data can be illustrated. Linear regression can be used for forecasting future events if one assumes that conditions under which the data was collected will continue and be uneffected or changed. Other areas of regression analysis which may be of use are: 1) distribution-free regression analysis (Hollander and Wolfe, 1973), and 2) regression with autocorrelated residuals (Neter and Wasserman, 1974; Kendall, 1973).

Linear and polynomial regression are simple techniques that can be used for modeling the pattern of change in a parameter over time. However, their use in forecasting is limited. Forecasting should involve the use of more complicated models and mathematical techniques not

explained here. Gilchrist (1970) provides an excellent review of statistical forecasting. Also, the references cited at the beginning of this section can be used to provide more information on methods of forecasting with time series data.

Application

In order to allow for better understanding of the trend detection methods and their usage, two examples are presented, one for a linear trend and the other for a step trend. The linear trend uses real data, while the step trend is a hypothetical situation. The examples were chosen in an attempt to cover the majority of techniques. The format is both verbal and numerical to provide a clear explanation of the techniques. Some equations will not be repeated, but are referenced by the equation number or the step number for a given procedure.

Linear Trend Example

Two problems with Lake Ontario are the increased algal densities and nutrient loading to the lake. While phosphorus is often the limiting nutrient in clean (oligotrophic) lakes, nitrogen has an increasing effect as water quality declines. Thus, to provide information on these two problems, the mean nitrate concentration in the springtime for Lake Ontario is examined for possible trends in the years 1968-1979. This information can be used to compare algal population growth and the nitrogen loading over these same years.

The mean springtime nitrate concentrations (μg/l) from Lake Ontario during 1968-1979 are in data set 2 in Appendix A. The data are mean values of samples collected evenly through the spring. The variances are assumed equal between different years because of limnological knowledge. Based on the knowledge that nutrient loadings have increased in this time period, the null and alternative hypotheses are

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 < \mu_2$$

(i.e., a positive trend is hypothesized).

The analysis is started by graphically displaying the data in a time series plot (Figure 18) and a cumulative sum plot (Figure 19). The following shows the calculation for the cusum points. A value of j = 1 is chosen to provide the maximum number of cusum points (i.e., maximum information). The cusum points are calculated in the following manner (Equation 31):

$$cu_j = \sum_{i=1}^{h} X_i - j/n \sum_{i=1}^{n} X_i$$

$$cu_1 = 215. - 1/12 \cdot (215 + 237 + \ldots + 335)$$

$$= -59$$

$$cu_2 = (215 + 237) - 2/12 \cdot 3288$$

$$= -96$$

$$\vdots$$

$$cu_{11} = (215 + 237 + \ldots + 324) - 11/12 \cdot 3288$$

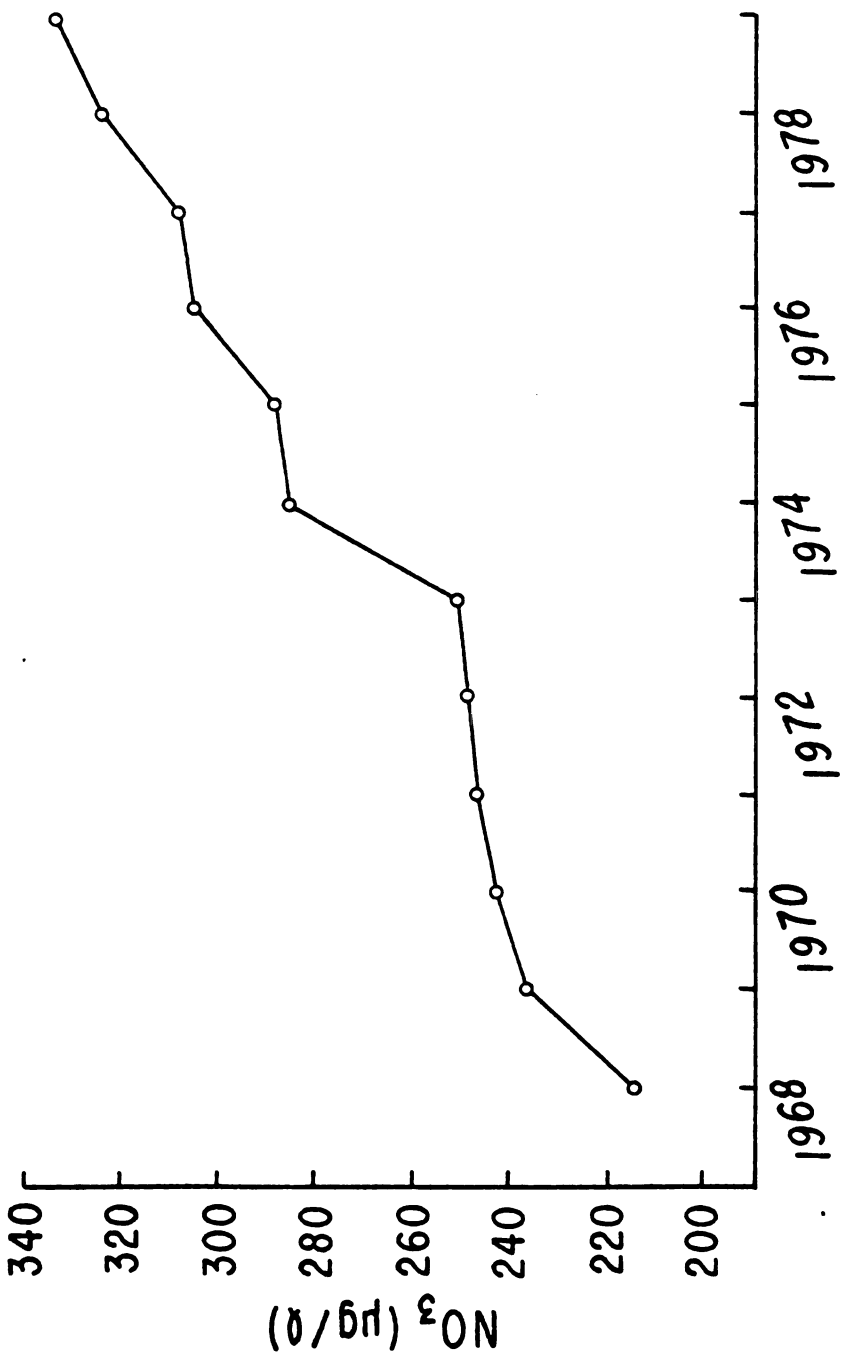


Figure 18. Time series plot of nitrate concentration (data set A.2).

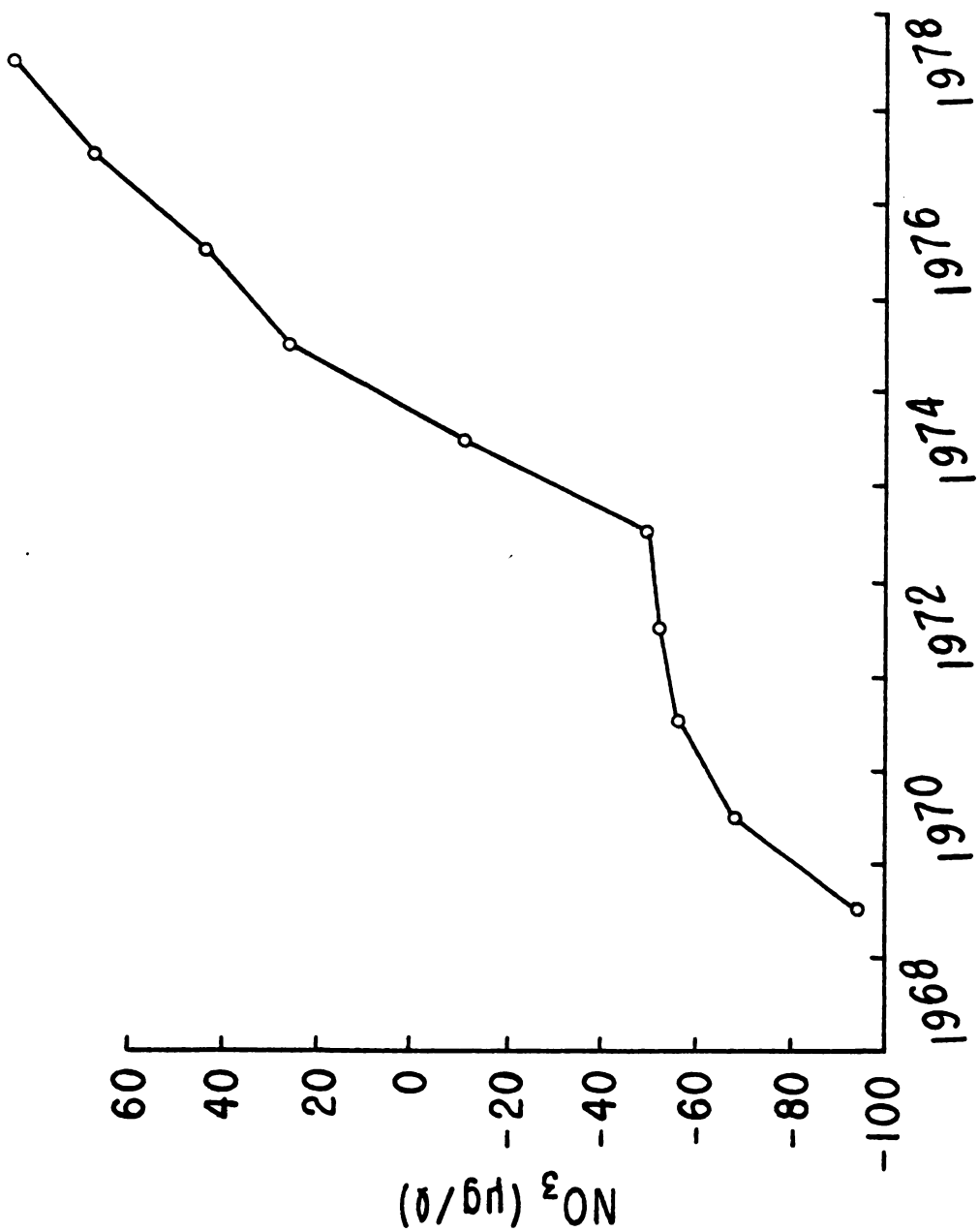


Figure 19. Cusum plot of nitrate concentrations (data set A.2).

Both the time series plot and cumulative sum plot seem to indicate a linear trend and further analysis should be conducted. The mode to mean ratio is used to provide a quick check of normality. The mode to mean ratio (Equation 36) is 0.97, which suggests the distribution is slightly skewed right. The skewness is shown in the relation between the mean and mode (Data set 2 in Appendix A). The W test is not used here due to the very small sample size and the inability of the test to determine non-normality under these conditions. Therefore, to be safe and not violate the normality assumption, the distribution is considered symetric, which is much easier to satisfy.

The next step is to estimate the autocorrelation coefficient and test to see if the autocorrelation is zero. To estimate ρ_1 (Equation 39) both c_k (k is equal to 1 in this case) and c_o need to be calculated in the following manner (Equation 40 and 41):

$$c_1 = \frac{1}{12}[(215 - 274)(237 - 274) + (237 - 274)(242 - 274) + \ldots + (324 - 274)(335 - 274)]$$

$$= 968.67$$

$$c_o = \frac{1}{12}[(215 - 274) + (327 - 274) + \ldots + (335 - 274)]$$

$$= 1378.3$$

$$r_1 = \frac{c_1}{c_o} = \frac{968.67}{1368.3} = .703$$

Thus, the estimate of the lag 1 autocorrelation coefficient is .703.

To test (H_0: $\rho = 0$) the Durbin-Watson test is used (Equation 42).

$$d = \frac{[-37 - (-59) + (-32) - (-37) + \ldots + (61 - 50)]}{(-59)^2 + (-37)^2 + \ldots + 61^2}$$

$$= .0073$$

Since the calculated d value is less than d_L (critical level) = 0.81, H_0 is rejected and the data are correlated. A technical note here is that it is useful to first calculate the residuals so as to make the calculations easier.

Based on dependent data and a linear trend hypothesis (from apriori information and examining time series and cusum plots), the Spearman Rho test is used, with adjusted critical levels, to test the hypothesis (H_1: $\mu_1 < \mu_2$).

Step 1. Using Equation 48, the test statistic is

$$T = (1 - 1)^2 + (2 - 2)^2 + (3 - 3)^2 + 5 - 4)^2 + 4 - 5)^2 + \ldots + (12 - 12)^2$$

$$= 2$$

Step 2. Calculate the modified critical level (Equation 49)

$$W'_{\alpha_L} = L_1 + f(n, \alpha, \rho)\ (W_{\alpha_L} - L_1)$$

where: $L_1 = 0$

$f(12, .01, .7) = .63$ from Table 12 Appendix B (generated)

$W_{\alpha_L} = W_{.005,\ 12} = 52$ from Table 10 Appendix B

$$W'_{\alpha_L} = 0 + (.11)\ (52 - 0)$$

$$= 5.72$$

Since the test statistic (T = 2) is less than the critical level ($W'_{\alpha_L} = 5.72$) one may reject the H_0 and conclude that a linear trend exists at a 99.5% confidence level.

The power associated with the Spearman Rho test for a linear trend in the mean springtime nitrate concentration is as follows:

Step 1. Calculate the number of equivalent independent samples Equation 51):

$$\frac{1}{n^*} \quad \frac{1}{n} \; 1 + \frac{2\,(\rho^t - \rho^{2t})}{(\rho^t - 1)^2}$$

$$\frac{1}{n^*} \quad \frac{1}{12} \; 1 + \frac{2\,(\rho^1 - \rho^{2.1})}{(\rho^1 - 1)^2}$$

$$n^* \quad 2$$

Step 2. Calculate the trend number (Equation 54):

$$N_T = \frac{T_r\,[n^*\,(n^* + 1)\,(n^* - 1)]^{1/2}}{12\,n^*\,\sigma}$$

where: $T_r = \Delta\mu$ for Linear Trends

$= 100$

$\sigma = 38.78$

$$N_T' = \frac{100\,[2.\,(2. + 1)\,(2. - 1)]^{1/2}}{12\,(2.12)\,(38.78)}$$

$= 0.86$

Step 3. Calculate the power (Equation 53):

$$1 - \beta = F(N_T - t_{1 - \alpha/2})$$

Substitute α for $\alpha/2$ because the test is one-tailed.

where: $t_{1 - .005,\ 10} = 3.169$

$$1 - \beta = F(.86 - 3.169)$$

$$= F(-2.309)$$

From the standard normal tables F(-2.31) is equal to .010, therefore the power is .01.

Based on the analysis, one may conclude that mean springtime nitrate levels have undergone a linear trend in Lake Ontario from 1968-1979 at the 99.5% confidence level. However, due to the small sample size and the large autocorrelation coefficient, only a power of .01 can be associated with the test.

Step Trend Example

The following hypothetical example illustrates the procedures for detection of step trends in lake parameters. For many years Lake A has had excellent water quality which allowed for tremendous recreation usage. These beautiful conditions prompted numerous housing developments on the lake and the surrounding area. Nutrient loads have increased significantly, to the point where the lake is experiencing accelerated eutrophication. The lake association consulted local officials and the state Department of Natural Resources for possible solutions. The lake exhibits a short detection time, therefore if nutrient loading is curtailed, water quality should respond fairly quick. A program was developed to control the limiting nutrient of the lake, phosphorus. The management program includes a sewage system, fertilizer control, and land use practices. The costs involved for the plan, especially for the construction of sewer pipes to connect to the local treatment plant, are high. During public meetings people expressed the concern of how sure can one be that these measures will "cure" the lake.

Another lake system, Lake B, nearby, underwent the exact same problem as Lake A. Since both lakes are limnologically very similar and the methods for restoration are similar, data from Lake B can be

used to provide the degree of quantitative evidence that the plan will be effective.

The hypothesis formed is that a step trend occurred starting at the time of implementation of the water quality plan. Mean annual total phosphorus is chosen as the lake parameter because of the availability of data and the significance it plays in lake eutrophication. Thus,

$$H_0: \mu_1 = \mu_2$$

$$H_0: \mu_1 > \mu_2$$

where: $\mu_1 = (X_1, \ldots, X_{15})$ and $\mu_2 = (X_{16}, \ldots, X_{30})$

The value which divides μ_1 and μ_2 is dependent upon when the management plan was initiated.

The mean annual total phosphorus concentrations are found in data set 3 of Appendix A. Figure 20 and 21 show the time-series plot and cusum plot, respectively. The sample residuals are shown in Figure 22 to check the validity of the assumption of homogeneous variance. The residuals are calculated by using

$$\hat{e}_i = X_i - \bar{X}$$

Therefore,

$$\hat{e}_1 = 77 - 77.7 = 7 - 0.7$$

$$\hat{e}_2 = 83 - 77.7 = 5.3$$

$$\vdots$$

$$\hat{e}_{30} = 64 - 77.7 = -13.7$$

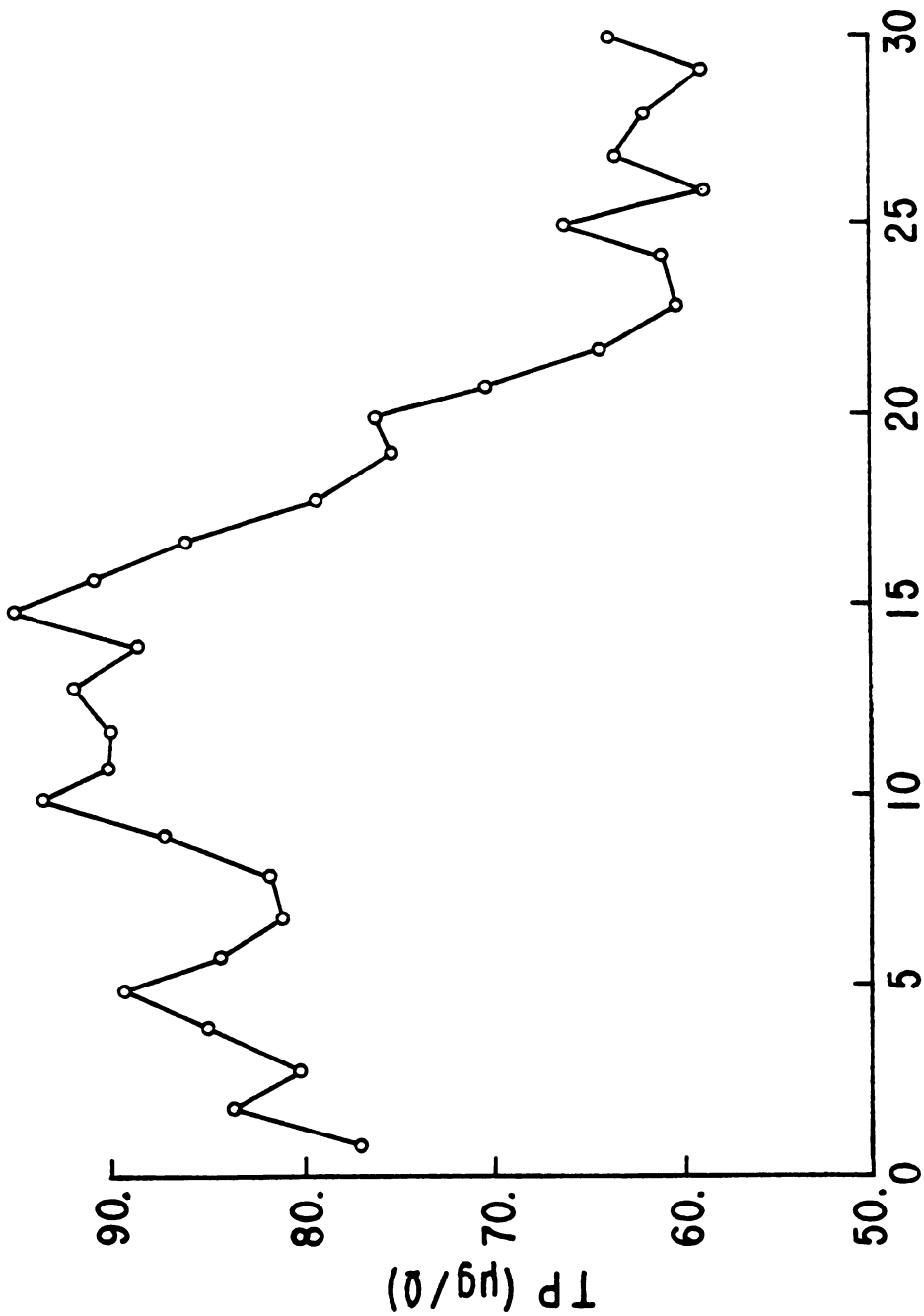


Figure 20. Time series plot of total phosphorus concentration (data set A.3).

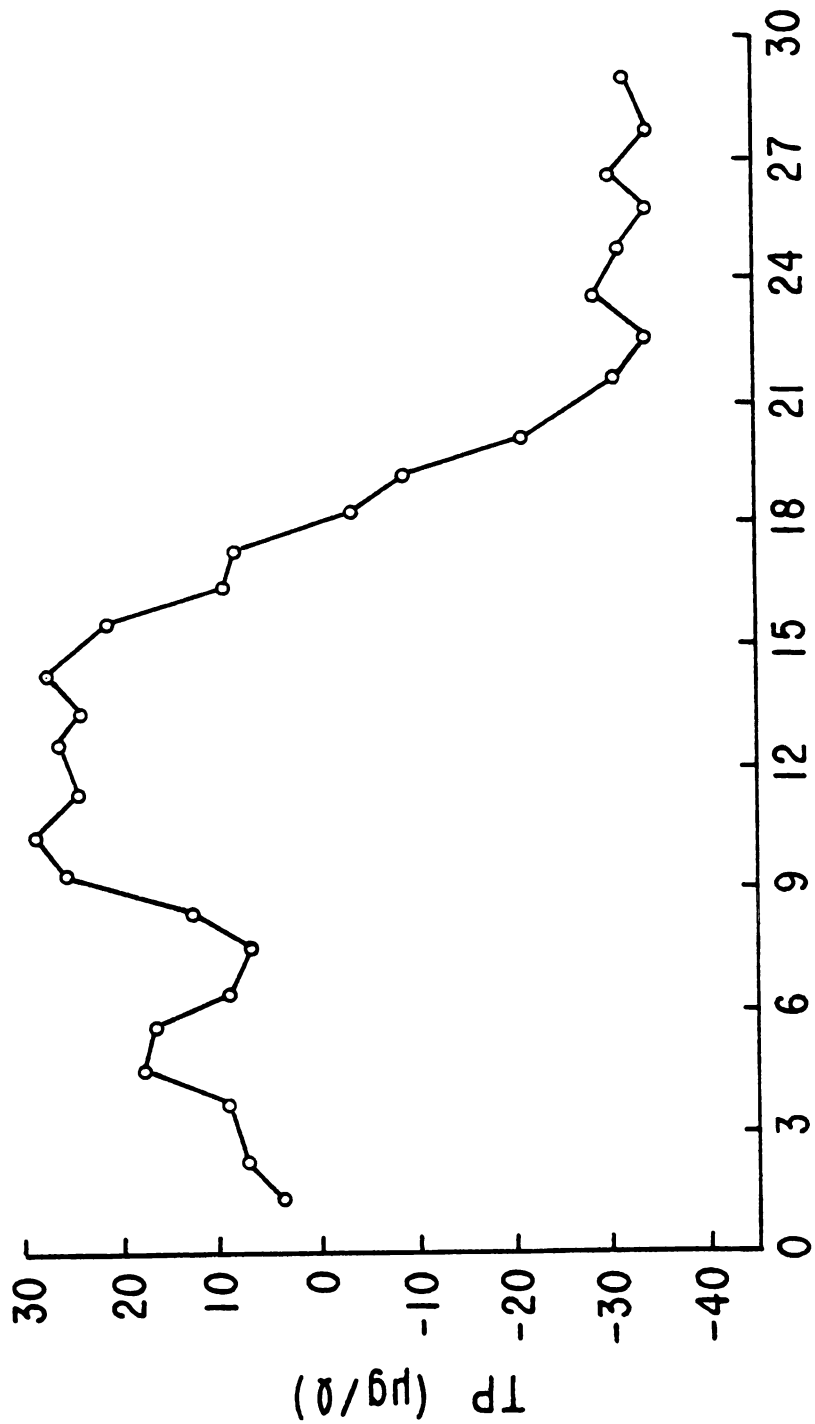


Figure 21. Cusum plot of total phosphorus concentrations (data set A.3).

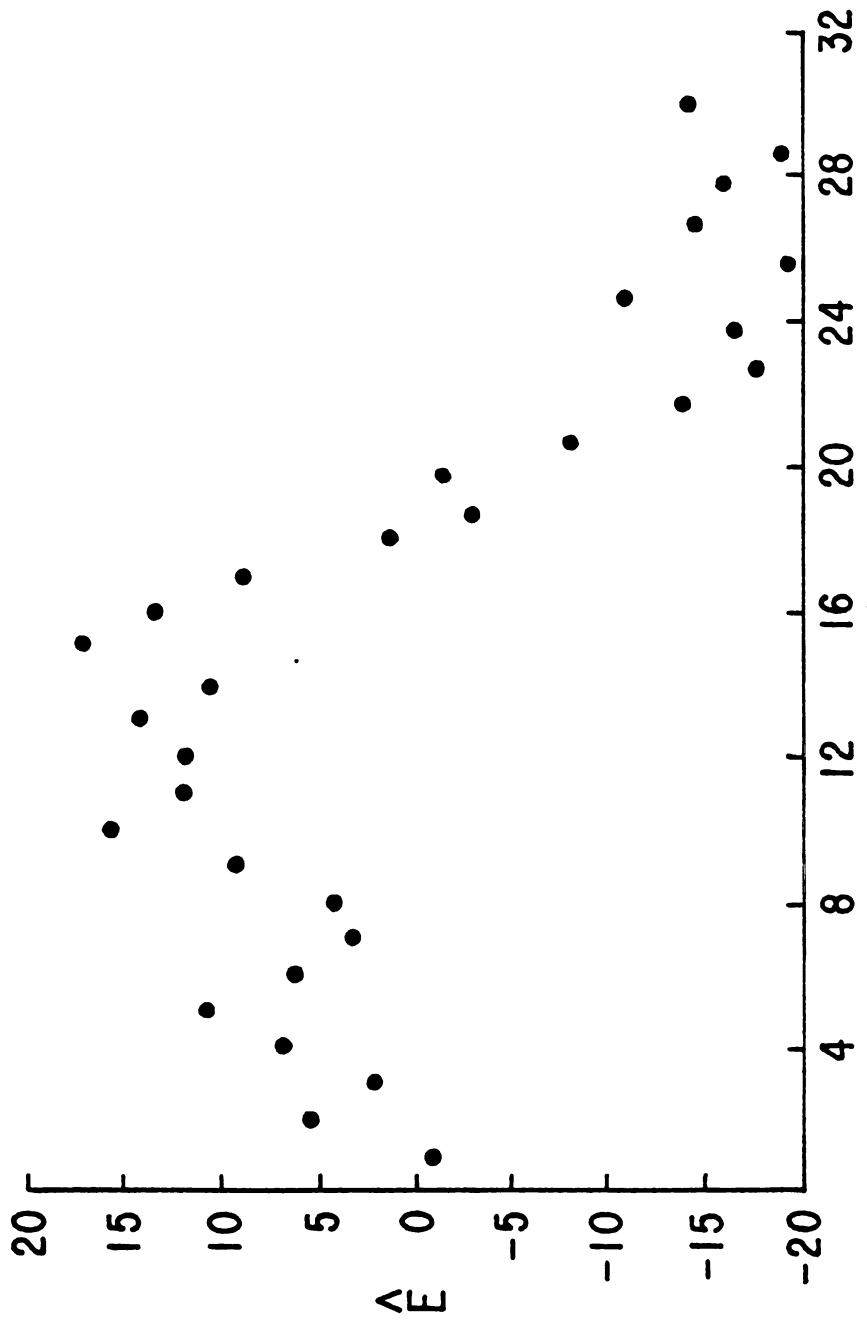


Figure 22. Plot of total phosphorus residuals (data set A.4).

The residual plot tends to be fairly horizontal and thus suggesting homogeneous variannce. Also, no extreme value is indicated in the plot.

The W-test is used to test for normality. The steps are:

Step 1. Order the array $X_1, X_2, \ldots, X_n$:

58.0, 59.0 ..., 94.0, 95.0

Step 2. Compute the variance (Equation 11):

$$s^2 = 143.80$$

Step 3. Since n is even,

$$K = n/2 = 30/2 = 15$$

Then, compute (Equation 34):

$$b = .0076\ (95.0 - 58.0) + .0227\ (94.0 - 59.0) + \ldots + .4254\ (81.0 - 80.0)$$

$$b = 25.66$$

Step 4. Compute test statistic (Equation 35):

$$W = 25.66^2/143.60$$

$$= 4.58$$

Since the test statistic (W = 4.58) exceeds the critical value ($W_{.09,\ 30} = .983$) one may not reject H_0, therefore one assumes normality. The mode to mean ratio is .99, which also suggests normality.

The lag 1 autocorrelation coefficient calculations are shown in the linear trend example, and thus repetition is not necessary. The value of r_1 (estimate of ρ_1) is 0.46. The Durbin-Watson was used to test whether ρ is significantly different from zero. The results (computation not shown) show ρ is significantly different from zero at the 99% confidence level.

The analysis of the data showed a possible step trend, normal and dependent data. Therefore, the Mann-Whitney test will be used to test the hypothesis of trend.

Step 1a. Rank all values from each population from smallest to largest (already done in W-test)

Step 1b. Sum the ranks for μ_1:

$$s = 13 + 18 + \ldots + 30$$

(Note, for values of 80.0 and 80.0, the ranks would have been 25 and 26, thus 25.5 is used for each.)

$$s = 325$$

Step 1c. Compute test statistic (Equation 45):

$$T = 325 - \frac{15\ (15 + 1)}{2}$$

$$= 205$$

Step 2. Modify the critical level (Equation 50):

$$W'_{\alpha_u} = L_u + f(n, \alpha, \rho)\ (W_{\alpha_u} - L_u)$$

where:

$$L_u = n_1 \cdot n_2$$

$$= 15 \cdot 15$$

$$= 225$$

$$f(.30, .01, .46) = .491 \text{ (interpolated from Table B.11)}$$

$$W_{.01,\ 15,\ 15} = 57$$

$$W_{\alpha_u} = W_{1\ -\ \alpha} = n_1 \cdot n_2 - W_{\alpha,\ n_1,\ n_2}$$

$$= (15 \cdot 15) - 57$$

$$= 168$$

$$W'_{\alpha_u} = 225 + (.491)\ (168 - 225)$$

$$= 197.01$$

Since the test statistic (T = 205) exceeds the modified critical level (W'_{α_u} = 197.01), H_0 is rejected and one may conclude a step trend at a 99.5% confidence level.

The associated power of the Mann-Whitney test is calculated in the following manner.

Step 1. Calculate number of equivalent independent samples (Equation 51):

$$\frac{1}{n^*} \approx \frac{1}{30}\ 1 + \frac{2\ (.46 - .46^2)}{(.46 - 1)^2}$$

$$n^* = 11$$

Step 2. Calculate the trend number (Equation 52):

$$N_T = \frac{T_r (n^*)^{1/2}}{2\sigma}$$

where: $T_r = |\mu_1 - \mu_2|$ for step trends

$$= |86.47 - 68.93|$$

$$= 17.54$$

$$N'_T = \frac{17.54\ (11)^{1/2}}{2 \cdot 11.99}$$

$$= 2.43$$

Step 3. Calculate the power (Equation 53):

$$1 - \beta = F(N'_T - t_{1 - \alpha,\ \nu})$$

where: $t_{1 - .005,\ 28} = 2.763$

$$= F(2.43 - 2.763)$$

$$= F(-.333)$$

From the standard normal tables F(-.33) is equal to .371, therefore the power of the test is .371.

Based on the test results one may conclude with 99.5% confidence and a power of .37 that a step trend is present in mean annual total phosphorus concentrations in Lake A.

CHAPTER V

CONCLUSIONS

Most phases of water quality management are concerned not only with the present condition, but with the spectrum from past to future conditions. Hence, information on the changes in water quality parameters over time is important for water quality management. The use of statistical techniques to detect trends in water quality parameters has the potential to provide the quantitative information needed for water resources management and planning decisions. The usefulness of the information generated by the detection of trends in water quality parameters should greatly increase in the future.

The method presented here, provides:

1. formulation of a problem (hypothesis)
2. selection of water quality parameter(s) and data
3. data analysis techniques
4. statistical tests for detection of trends

The development of any quantitative methodology usually has its roots in a need exhibited in the area of management or planning. Thus, the use of the trend detection technique arises from a need discerned in the management area for quantitative information on changes in water quality parameters over time. The starting point for the use of trend detection techniques is the determination of a desired output for which proper management, and planning decisions are to be based upon. This desired output determines the parameter(s) and data chosen, and the formulation of the null and alternative hypotheses. It is extremely important to properly consider the desired output in the selection of data and hypothesis formation, for decisions made based on the results of the

trend detection analysis, must be in lieu of the input information (i.e., parameter(s), data, and hypothesis).

The preliminary data analysis techniques are a necessary step for trend detection analysis. The data analysis step provides: 1) visual evidence of trends and validity of assumptions, 2) statistical tests to determine the data's distribution, 3) statistical tests for the verification of the assumptions of homogeneous variance and independent data, 4) methods of examining "outliers," and 5) types of data transformations. The data analysis techniques presented are the ones generally accepted by applied scientists. However, many other techniques are available and may be substituted, given the alternative technique is as powerful. The importance of preliminary data analysis cannot be overstressed. The information provided by the examining the data will: 1) determine which statistical tests to use for detecting trends in water quality parameters, and 2) provide a "good feel" for the data to allow for the utmost evaluation and application of the trend detection results to water quality management and planning decisions.

The autocorrelation present in most lake data, prohibits the use of most statistical techniques, including regression, based on the violation of the independency assumption. The violation of the independency assumption, as well as the normality assumption, is a major problem with the majority of lake analyses conducted in the past. The continued violation of these assumptions is not necessary, when there are techniques, such as the one presented here, that accommodate the dependent data situation. Thus, the statistical tests for dependent data are suggested for use, almost exclusively, when dealing with lake systems. The dependency of the data is used, along with sample size and alpha level, to correct the critical level of the statistical test.

As with most statistical techniques, there are some limitations to the trend detection method. The major problem, especially with trend detection analysis for water quality parameters, is the poor power associated with small sample sizes. The small sample size problem is compounded by increasing autocorrelation. When the sample size is small and a large autocorrelation of the data is present, the number of effective independent samples is exceedingly small. Since the power is a function of the effective independent samples, as effective samples decrease, so does the power. This is exemplified in the step trend example in the application section of Chapter 4. Based on a sample size of 12 and an autocorrelation of 0.7 the number of effective independent samples is 2, and this results in a power of 0.01.

Recalling Figure 11, the power is the probability of accepting H_1 when in reality H_1 is true. In trend detection, the power means the probability of detecting a trend when in truth a trend exists. From a management aspect, the power is important since the information usually desired is whether a parameter has changed over time (i.e., did phosphorus concentrations change?).

Another problem with the trend detection technique for dependent data, is the assumption that both the variance and autocorrelation are known. The assumption of known variance and autocorrelation is true when the estimates for the variance and autocorrelation are accurate (i.e., known with certainty). Accuracy in the estimate of variance occurs when: 1) the sample size exceeds 30, 2) the data are independent, and 3) the sampling design is good. Autocorrelation accuracy is very important, for the autocorrelation is used in calculating: 1) effective independent samples, and 2) variance estimate. Both the number of effective independent samples and the variance estimate are used to

calculate the test statistic (N_T or N_T') which determines the power of the test. The problem of not knowing the variance and autocorrelation may be accomodated by using upper confidence limits of estimates of the variance and autocorrelation to provide conservative results.

Water quality management requires an understanding of aquatic ecosystems. Basic to the understanding of aquatic systems is sound information on which an understanding can be derived from. Thus, for water quality management, information is needed on the changes in water quality parameters over time. The detection of trends in water quality parameters allows for the management of water resources for their full beneficial use. Thus, trend detection techniques for water quality parameters is an integral part of water resources management for protecting environmental quality.

LIST OF REFERENCES

LIST OF REFERENCES

Atkinson, J. D., 1967. An investigation of Goodness of Fit Tests for Normality. Masters Thesis. Iowa State Univ., Ames, Iowa.

Bartlett, M. S., 1935. Some Aspects of the Time-correlation Problem in Regard to Tests of Significance. J. R. Stat. Soc. 98:536-543.

Bartlett, M. S., 1937. Some Examples of Statistical Methods of Research in Agriculture and Applied Biology. J. R. Stat. Soc (suppl.) 4:137-147.

Bayley, G. V. and J. M. Hammersley, 1946. The Effective Number of Independent Observations in an Autocorrelated Time Series. J. R. Stat. Soc. (suppl.) 8(1):184-197.

Beckers, C. V., S. G. Chamberlain, and G. P. Grimsrud, 1972. Quantitative Methods for Preliminary Design of Water Quality Surveillance Systems. EPA-R5-72-001.

Behnken, D. W. and N. R. Draper, 1972. Residuals and their Variance Patterns. Technometrics 14:101-12.

Bhattacharyya, G. K. and R. A. Johnson, 1977. Statistical Concepts and Methods. J. Wiley & Sons, New York.

Bloomfield, P., 1976. Fourier Analysis of Time Series: An Introduction. J. Wiley & Sons, New York.

Box, G. P. and G. M. Jenkins, 1970. Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco, Calif.

Breiman, L., 1973. Statistics with a View Toward Applications. Houghton-Mifflin, Boston, Mass.

Cairns, J., Jr., K. L. Dickerson, and G. F. Westlake (editors), 1977. Biological Monitoring of Water and Effluent Quality, STP607, Amer. Soc. for Testing and Materials, Philadelphia, Pa.

Carlson, R. E., 1977. A Trophic State Index for Lakes. Limnol. Oceanogr. 22(2):361-369.

Chapra, S. C., 1980. Simulation of Recent and Projected Total Phosphorus Trends in Lake Ontario. JGLR 6(2):101-112.

Chatfield, C., 1975. The Analysis of Time-Series: Theory and Practice. Chapman and Hall, London.

Chatfield, C., 1977. Some Recent Developments in Time-Series Analysis. J. R. Stat. Soc. A 140(4):492-510.

Chatterjee, S. and B. Price, 1977. Regression Analysis by Example. J. Wiley & Sons, New York.

Chen, E. H., 1971. The Power of the Shapiro-Wilk W Test for Normality in Samples from Contaminated Normal Distributions. J. Am. Stat. Assoc. 66:760-762.

Clevland, W. P. and G. C. Tiao, 1976. Decomposition of Seasonal Time Series: A Model for the Census X-11 Program. J. Amer. Stat. Assoc. 71:581-587.

Cochran, W. G., 1963. Sampling Techniques. J. Wiley & Sons, New York.

Conover, C. J., 1971. Practical Nonparametric Statistics. J. Wiley & Sons, New York.

Dobson, F. H., 1980. Lake Ontario Phosphorus Management: Observations and Interpretations of Within-lake Trends 1965 to 1979. Paper presented at the Inter. Assoc. for Great Lakes Research Conference. Kingston, Ontario.

Durbin, J. and M. J. Murray, 1975. Seasonal Adjustment Based on a Mixed Additive-multiplicative Model. J. R. Stat. Soc. A 138:385-410.

Durbin, J. and G. S. Watson, 1951. Testing for Serial Correlation in Least-squares Regression. Biometrika 38:159-178.

Edwards, M. D., 1980. Water Data and Services Available from Participants in the National Water Data Exchange. Water Res. Bult. 16(1):1-14.

Environmental Protection Agency, 1974. National Water Quality Inventory. 1974, EPA-440/9-74-001.

Environmental Protection Agency, 1976. Quality Criteria for Water. U. S. Environ. Prot. Ag., Washington, D. C.

Gastwirth, J. L. and H. Rubin, 1971. Effect of Dependence on the Level of Some One-sample Tests. J. Amer. Stat. Assoc. 66:816-820.

Gastwirth, J. L. and H. Rubin, 1975. The Behavior of Robust Estimators on Dependent Data. Ann. Statist. 3:1070-1100.

Gilchrist, W., 1976. Statistical Forecasting. J. Wiley & Sons, New York.

Gill, J. L., 1978. Design and Analysis of Experiments in the Animal and Medical Sciences. Iowa State Univ. Press, Ames, Iowa.

Granger, C. W. J., and P. Newbold, 1977. Forecasting Economic Time Series. Academic Press, New York.

Grubbs, F. E. and G. Beck, 1972. Extension of Sample Sizes and Percentage Points for Significance Tests of Outlying Observations. Technometrics 14:847-854.

Hartley, H. O., 1950. The Maximum F-ratio as a Short-cut Test for Heterogeneity of Variance. Biometrika 37:308-312.

Hollander, M. and D. A. Wolfe, 1973. Nonparametric Statistical Methods. J. Wiley & Sons, New York.

Hutchinson, G. E., 1957. A Treatise on Limnology, Volume 1. J. Wiley & Sons, New York.

Hynes, H. B. N., 1970. The Ecology of Running Waters. Univ. of Toronto Press, Toronto, Canada.

Jenkins, G. M. and D. G. Watts, 1968. Spectral Analysis and Its Applications. Holden-Day, San Francisco, Calif.

Kendall, M. G., 1973. Time-Series. Hafner Press, New York.

Kendall, M. G. and A. Stuart, 1967. The Advanced Theory of Statistics, Vol. 2, Inference and Relationships. Griffin, London.

Kenkel, J. L., 1975. Small Sample Tests for Serial Correlation in Models Containing Lagged Dependent Variables. Review of Econ. and Stat. 57:383-386.

Lettenmaier, D. P., 1976. Detection of Trends in Water Quality Data from Records with Dependent Observations. Water Res. Bult. 12(5):1037-1046.

Lettenmaier, D. P., 1977. Detection of Trends in Stream Quality Monitoring Network Design and Data Analysis. Technical Report 51, C. W. Harris Hydraulic Lab, Dept. of Civil Eng., Univ of Washington.

Lund, J. W. G., 1949. Studies on Asterionella, I. The Origin and Nature of the Cells Producing Seasonal Maxima. J. Ecol. 37:389-419.

Lund, J. W. G., 1950. Studies on Asterionella formosa Hass, II. Nutrient Depletion and the Spring Maximum. J. Ecol. 38:1-35.

Matalas, N. C. and W. B. Langbein, 1962. Information Content of the Mean. J. Geophys. Res. 67 (9):3441-3448.

McNeely, R. N., V. P. Neimanis, and L. Dwyer, 1979. Water Quality Source Book: A Guide to Water Quality Parameters. Inland Waters Directorate, Ottawa, Canada.

Mosteller, F. and J. W. Tukey, 1977. Data Analysis and Regression. Addison-Wesley, Reading, Mass.

Mood, A. M. and F. A. Graybill, 1963. Introduction to the Theory of Statistics. McGraw-Hill, New York.

Neter, J. and W. Wasserman, 1974. Applied Linear Statistical Models. R. D. Irwin, Inc., Homewood, Ill.

Olkin, I. and J. W. Pratt, 1958. Unbiased Estimation of Certain Correlation Coefficients. Ann. Math. Stat. 29:201-211.

Reckhow, K. H., 1980. Lake Data Analysis and Nutrient Budget Modeling. Corvallis Environ. Res. Lab., Environmental Protection Agency (in press).

Reckhow, K. H. and S. C. Chapra, 1980. Engineering Approachs for Lake Management: Data Analysis and Modeling. Ann Arbor Science, Ann Arbor, Mich. (in press).

Rockwell, D. C, C. V. Marion, M. F. Palmer, D. S. Devault, and R. J. Bowden, 1979. Environmental Trends in Lake Michigan, In: Phosphorus Management Strategies for Lakes. Ed. R. C. Loehr. Ann Arbor Science, Ann Arbor, Mich.

Sanders, T. G., D. P. Adrian, and B. B. Berger, 1976. Designing a River Basin Sampling System, Report No. 62, Dept. of Civil Eng., Univ. of Mass.

Serfling, R. J., 1968. The Wilcoxon Two-sample Statistic on Strongly Mixing Processes. Ann. Math. Statist. 39:1202-1209.

Sen, P. N., 1963. On the Properties of U-Statistics When the Observations Are Not Independent. Calcutta Stat. Ass. Bult. 12(47):69-92.

Sen, P. N., 1965. Some Nonparametric Tests for M-Dependent Time Series. J. Amer. Stat. Ass. 60(1):134-147.

Sen, Z., 1978. Autorun Analysis of Hydrologic Time Series. J. Hydrology 36:75-85.

Sen, Z., 1979. Application of the Autorun Test to Hydrologic Data. J. Hydrology 42:1-7.

Sherwani, J. K. and D. H. Moreau, 1975. Strategies for Water Quality Monitoring. Environ. Sci. and Eng. Pub. No. 398. Univ. of North Carolina at Chapel Hill.

Shapiro, S. S. and M. B. Wilk, 1965. An Analysis of Variance Tests for Normality, J. Am. Stat. Ass. 67:215-216.

Siegel, S., 1956. Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill, New York.

Sokal, R. R. and F. J. Rohlf, 1969. Biometry. Freeman, San Francisco, Calif.

Springer, E. P. and G. F. Gifford, 1980. Spatial Variability of Rangeland Infiltration Rates, Water Res. Bult. 16(3):550-552.

Steel, R. G. D. and J. H. Torrie, 1960. Principles and Procedures in Statistics. McGraw-Hill, New York.

Stumm, W. and J. J. Morgan, 1970. Aquatic Chemistry, An Introduction Emphasizing Chemical Equilibria in Natural Waters. Wiley-Interscience, New York.

Trussell, R. P., 1972. The Percent Un-ionized Ammonia in Aqueous Ammonia Solutions at Different pH Levels and Temperatures. J. Fish. Res. Bd. Canada 29:1505-1507.

Tukey, J. W., 1962. The Future of Data Analysis. Ann. Math. Stat. 33:1-67.

Tukey, J. W., 1977. Exploratory Data Analysis. Addison-Wesley, Reading, Mass.

Vollenweider, R. A., 1968. The Scientific Basis of Lake and Stream Eutrophication, with Particular Reference to Phosphorus and Nitrogen as Eutrophication Factors. OECD Paris-Tech. Rep. DAS/DSI/68.

Walker, W. W., Jr., 1979. Use of Hypolimnetic Oxygen Depletion Rate as a Trophic State Index for Lakes. Water Resour. Res. 15(6):1463-1470.

Wallis, K. F., 1974. Seasonal Adjustment and Relations Between Variables. J. Amer. Stat. Ass. 69:18-31.

Ward, R. C., 1973. Data Acquisition Systems in Water Quality Management. EPA-R5-73-014.

Weber, C. I., 1980. Federal and State Biomonitoring Programs. In: Biological Monitoring for Environmental Effects. Ed. D. L. Worf. D. C. Heath, Lexington, Mass.

Weir, J. B., 1960. Standardized t. Nature (London) 185:558.

Wetzel, R. G., 1975. Limnology. W. B. Saunders, Philadelphia, Pa.

Wetzel, R. G. and P. H. Rich, 1973. Carbon in Freshwater Systems. In: Carbon and the Biosphere. Ed. G. M. Woodwell and E. V. Pecan. Tech. Info. Center, U. S. Atomic Energy Comm.

Wolman, G. M., 1971. The Nation's Rivers. Science 174(4012):905-918.

Worf, D. L. (Editor), 1980. Biological Monitoring for Environmental Effects. D. C. Heath, Lexington, Mass.

Yule, G. U., 1921. On the Time-correlation Problem with Especial Reference to the Variate-difference Correlation Method. J. Roy, Stat. Soc. 84:496-537.

Zar, J. H., 1974. Biostatistical Analysis. Prentice-Hall, Englewood Cliffs, N. J.

APPENDIX A

Data Set A.1: Mean total phosphorus concentrations (μg/l) in Lake Ontario's offshore waters in springtime from 1969-1979 (Dobson, 1980)

Year	Total Phosphorus (μg P/L)
1969	20.8
1970	20.9
1971	23.0
1972	21.5
1973	22.6
1974	21.8
1975	21.1
1976	21.6
1977	20.5
1978	17.2
1979	15.9

$\sum_{i=1}^{n} Y_i = 226.9$

$n = 11$

Mean	= 20.63
Median	= 21.1
Midrange	= 19.45
Geometric mean	= 20.51
Harmonic mean	= 20.39
Range	= 7.1
Variance	= 4.70
Standard deviation	= 2.17
Standard error of mean	= 0.65
Coefficient of variation	= 0.105

Data Set A.2. Mean nitrate concentrations (µg/1) for Lake Ontario in spring from 1968-1979 (from Great Lakes Environmental Research Lab).

Year	Rank	Nitrate (µg/1)	Rank		
1968	1	215.	1		
1969	2	237.	2	Mean	= 274.0
1970	3	242.	3	Median	= 267.5
1971	4	249.	5	Variance	= 1503.64
1972	5	247.	4	Standard deviation	= 38.78
1973	6	251.	6	Standard error of mean	= 11.19
1974	7	287.	8	Coefficient of variation	= 0.142
1975	8	286.	7	n	= 12
1976	9	306.	9		
1977	10	309.	10		
1978	11	324.	11		
1979	12	335.	12		

Data Set A.3. Hypothetical mean annual total phosphorus concentration for Lake B, over 30 years.

Year	Phosphorus (µg/l)	Rank	Year	Phosphorus (µg/l)	Rank
1	77.0	13	16	91.0	27
2	83.0	18	17	86.0	21
3	80.0	15	18	79.0	14
4	85.0	20	19	75.0	11
5	89.0	24	20	76.0	12
6	84.0	12	21	70.0	10
7	81.0	16	22	64.0	8
8	82.0	17	23	60.0	3
9	87.0	22	24	61.0	4
10	94.0	29	25	66.0	9
11	90.0	25.5	26	58.0	1
12	90.0	25.5	27	63.0	6
13	92.0	28	28	62.0	5
14	88.0	23	29	59.0	2
15	95.0	30	30	64.0	7

Overall Mean = 77.70 $Mean_1$ (1 - 15) = 86.47

Overall Variance = 143.80 $Mean_2$ (16 - 30) = 68.93

Overall Standard Deviation = 11.99

n = 30

APPENDIX B

TABLE B.1. SHAPIRO-WILK TEST FOR NORMALITY

Coefficients of Ordered Differences ($a_{i,n}$)

i	n=11	12	13	14	15	16	17	18	19	20
1	.5601	.5475	.5359	.5251	.5150	.5056	.4968	.4886	.4808	.4734
2	.3315	.3325	.3325	.3318	.3306	.3290	.3273	.3253	.3232	.3211
3	.2260	.2347	.2412	.2460	.2495	.2521	.2540	.2553	.2561	.2565
4	.1429	.1586	.1707	.1802	.1878	.1939	.1988	.2027	.2059	.2085
5	.0695	.0922	.1099	.1240	.1353	.1447	.1524	.1587	.1641	.1686
6	...	.0303	.0539	.0727	.0880	.1005	.1109	.1197	.1271	.1334
7	...	...	...	.0240	.0433	.0593	.0725	.0837	.0932	.1013
8	...	...	...	...	...	.0196	.0359	.0496	.0612	.0711
9	...	...	...	...	...	...	...	.0163	.0303	.0422
10	...	...	...	...	...	...	...	...	...	.0140

i	n=21	22	23	24	25	26	27	28	29	30
1	.4643	.4590	.4542	.4493	.4450	.4407	.4366	.4328	.4291	.4254
2	.3185	.3156	.3126	.3098	.3069	.3043	.3018	.2992	.2968	.2944
3	.2578	.2571	.2563	.2554	.2543	.2533	.2522	.2510	.2499	.2487
4	.2119	.2131	.2139	.2145	.2148	.2151	.2152	.2151	.2150	.2148
5	.1736	.1764	.1787	.1807	.1822	.1836	.1848	.1857	.1864	.1870
6	.1399	.1443	.1480	.1512	.1539	.1563	.1584	.1601	.1616	.1630
7	.1092	.1150	.1201	.1245	.1283	.1316	.1346	.1372	.1395	.1415
8	.0804	.0878	.0941	.0997	.1046	.1089	.1128	.1162	.1192	.1219
9	.0530	.0618	.0696	.0764	.0823	.0876	.0923	.0965	.1002	.1036
10	.0263	.0368	.0459	.0539	.0610	.0672	.0728	.0778	.0822	.0862
11	...	.0122	.0228	.0321	.0403	.0476	.0540	.0598	.0650	.0697
12	...	...	...	.0107	.0200	.0284	.0358	.0424	.0483	.0537
13	...	...	...	...	...	.0094	.0178	.0253	.0320	.0381
14	...	...	...	...	...	...	...	.0084	.0159	.0227
15	...	...	...	...	...	...	...	...	...	.0076

Table B.1. Coefficients of Ordered Differences ($a_{i,n}$) (cont.)

i	n=31	32	33	34	35	36	37	38	39	40
1	.4220	.4188	.4156	.4127	.4096	.4068	.4040	.4015	.3989	.3964
2	.2921	.2898	.2876	.2854	.2834	.2813	.2794	.2774	.2755	.2737
3	.2475	.2463	.2451	.2439	.2427	.2415	.2403	.2391	.2380	.2368
4	.2145	.2141	.2137	.2132	.2127	.2121	.2116	.2110	.2104	.2098
5	.1874	.1878	.1880	.1882	.1883	.1883	.1883	.1881	.1880	.1878
6	.1641	.1651	.1660	.1667	.1673	.1678	.1683	.1686	.1689	.1691
7	.1433	.1449	.1463	.1475	.1487	.1496	.1505	.1513	.1520	.1526
8	.1243	.1265	.1284	.1301	.1317	.1331	.1344	.1356	.1366	.1376
9	.1066	.1093	.1118	.1140	.1160	.1179	.1196	.1211	.1225	.1237
10	.0899	.0931	.0961	.0988	.1013	.1036	.1056	.1075	.1092	.1108
11	.0739	.0777	.0812	.0844	.0873	.0900	.0924	.0947	.0967	.0986
12	.0585	.0629	.0669	.0706	.0739	.0770	.0798	.0824	.0848	.0870
13	.0435	.0485	.0530	.0572	.0610	.0645	.0677	.0706	.0733	.0759
14	.0298	.0344	.0395	.0441	.0484	.0523	.0559	.0592	.0622	.0651
15	.0144	.0206	.0262	.0314	.0361	.0404	.0444	.0481	.0515	.0546
16	...	.0068	.0131	.0187	.0239	.0287	.0331	.0372	.0409	.0444
17	...	...	...	.0062	.0119	.0172	.0220	.0264	.0305	.0343
18	...	...	...	...	...	.0057	.0110	.0158	.0203	.0244
19	...	...	...	...	...	...	...	.0053	.0101	.0146
20	...	...	...	...	...	...	...	...	...	.0049

Table B.1. Coefficients of Ordered Differences ($a_{i,n}$) (cont.)

i	n=41	42	43	44	45	46	47	48	49	50
1	.3940	.3917	.3894	.3872	.3850	.3830	.3808	.3789	.3770	.3751
2	.2719	.2701	.2684	.2667	.2651	.2635	.2620	.2604	.2589	.2574
3	.2357	.2345	.2334	.2323	.2313	.2302	.2291	.2281	.2271	.2260
4	.2091	.2085	.2078	.2072	.2065	.2058	.2052	.2045	.2038	.2032
5	.1876	.1874	.1871	.1868	.1865	.1862	.1859	.1855	.1851	.1847
6	.1693	.1694	.1695	.1695	.1695	.1695	.1695	.1693	.1692	.1691
7	.1531	.1535	.1539	.1542	.1545	.1548	.1550	.1551	.1553	.1554
8	.1384	.1392	.1398	.1405	.1410	.1415	.1420	.1423	.1427	.1430
9	.1249	.1259	.1269	.1278	.1286	.1293	.1300	.1306	.1312	.1317
10	.1123	.1136	.1149	.1160	.1170	.1180	.1189	.1197	.1205	.1212
11	.1004	.1020	.1035	.1049	.1062	.1073	.1085	.1095	.1105	.1113
12	.0891	.0909	.0927	.0943	.0959	.0972	.0986	.0998	.1010	.1020
13	.0782	.0804	.0824	.0842	.0860	.0876	.0892	.0906	.0919	.0932
14	.0677	.0701	.0724	.0745	.0765	.0783	.0801	.0817	.0832	.0846
15	.0575	.0602	.0628	.0651	.0673	.0694	.0713	.0731	.0748	.0764
16	.0476	.0506	.0534	.0560	.0584	.0607	.0628	.0648	.0667	.0685
17	.0379	.0411	.0442	.0471	.0497	.0522	.0546	.0568	.0588	.0608
18	.0283	.0318	.0352	.0383	.0412	.0439	.0465	.0489	.0511	.0532
19	.0188	.0227	.0263	.0296	.0328	.0357	.0385	.0411	.0436	.0459
20	.0094	.0136	.0175	.0211	.0245	.0277	.0307	.0335	.0361	.0386
21	...	.0045	.0087	.0126	.0163	.0197	.0229	.0259	.0288	.0314
22	...	...	...	.0042	.0081	.0118	.0153	.0185	.0215	.0244
23	...	...	...	...	...	.0039	.0076	.0111	.0143	.0174
24	...	...	...	...	...	...	...	.0037	.0071	.0104
25	...	...	...	...	...	...	...	...	...	.0035

Source: Gill (1978).

Table B.2. Critical Values for W Statistic*

n	α=0.9	0.5	0.10	0.05	0.02	0.01
11	.973	.940	.876	.850	.817	.792
12	.973	.943	.883	.859	.828	.805
13	.974	.945	.889	.866	.837	.814
14	.975	.947	.895	.874	.846	.825
15	.975	.950	.901	.881	.855	.835
16	.976	.952	.906	.887	.863	.844
17	.977	.954	.910	.892	.869	.851
18	.978	.956	.914	.897	.874	.858
19	.978	.957	.917	.901	.879	.863
20	.979	.959	.920	.905	.884	.868
21	.980	.960	.923	.908	.888	.873
22	.980	.961	.926	.911	.892	.878
23	.981	.962	.928	.914	.895	.881
24	.981	.963	.930	.916	.898	.884
25	.981	.964	.931	.918	.901	.888
26	.982	.965	.933	.920	.904	.891
27	.982	.965	.935	.923	.906	.894
28	.982	.966	.936	.924	.908	.896
29	.982	.966	.937	.926	.910	.898
30	.983	.967	.939	.927	.912	.900
31	.983	.967	.940	.929	.914	.902
32	.983	.968	.941	.930	.915	.904
33	.983	.968	.942	.931	.917	.906
34	.983	.969	.943	.933	.919	.908
35	.984	.969	.944	.934	.920	.910
36	.984	.970	.945	.935	.922	.912
37	.984	.970	.946	.936	.924	.914
38	.984	.971	.947	.938	.925	.916
39	.984	.971	.948	.939	.927	.917
40	.985	.972	.949	.940	.928	.919
41	.985	.972	.950	.941	.929	.920
42	.985	.972	.951	.942	.930	.922
43	.985	.973	.951	.943	.932	.923
44	.985	.973	.952	.944	.933	.924
45	.985	.973	.953	.945	.934	.926
46	.985	.974	.953	.945	.935	.927
47	.985	.974	.954	.946	.936	.928
48	.985	.974	.954	.947	.937	.929
49	.985	.974	.955	.947	.937	.929
50	.985	.974	.955	.947	.938	.930

Source: **Gill (1978).**

*Nonnormality is indicated when the W statistic is *smaller* than the appropriate critical value.

TABLE B.3 UPPER PERCENTAGE POINTS OF F_{max} DISTRIBUTION (1-CDF)

Upper Percentage Points of F_{max} Distribution (1-CDF): Equal Replication

r	α	r=3	4	5	6	7	8	9	10	12
2	.25	15.4	23.1	33.3	46.8	58.3	70.8	83.3	96.7	121
	.10	42.5	60.1	98.2	129	162	196	231	267	342
	.05	87.5	112	202	266	333	403	475	550	704
	.01	448	729	1036	1362	1705	2063	2432	2813	3605
3	.25	8.15	11.7	15.1	18.5	21.9	25.2	28.4	31.6	37.8
	.10	16.8	21.0	30.9	37.7	44.4	50.9	57.4	63.7	76.1
	.05	27.8	39.2	50.7	62.0	72.9	83.5	93.9	104	124
	.01	85	120	151	184	216	249	281	310	361
4	.25	5.79	7.81	9.67	11.4	13.1	14.7	16.2	17.6	20.4
	.10	10.3	13.9	17.1	20.1	22.9	25.6	28.1	30.6	35.5
	.05	15.5	20.6	25.2	29.5	33.6	37.5	41.1	44.6	51.4
	.01	37	49	59	69	79	89	97	106	120
5	.25	4.66	6.05	7.28	8.41	9.46	10.4	11.4	12.2	13.9
	.10	7.68	9.86	11.8	13.5	15.2	16.7	18.1	19.4	21.9
	.05	10.8	13.7	16.5	18.7	20.8	22.9	24.7	26.5	29.9
	.01	22	28	33	38	42	46	50	54	60
6	.25	4.00	5.05	5.97	6.78	7.53	8.22	8.87	9.48	10.6
	.10	6.23	7.78	9.11	10.3	11.4	12.4	13.3	14.2	15.8
	.05	8.38	10.4	12.1	13.7	15.0	16.3	17.5	18.6	20.7
	.01	15.5	19.1	22	25	27	30	32	34	37
7	.25	3.57	4.41	5.14	5.77	6.35	6.88	7.37	7.82	8.66
	.10	5.32	6.52	7.52	8.41	9.20	9.93	10.6	11.2	12.4
	.05	6.94	8.44	9.70	10.8	11.8	12.7	13.5	14.3	15.8
	.01	12.1	14.5	16.5	18.4	20	22	23	24	27
8	.25	3.26	3.97	4.57	5.09	5.55	5.98	6.37	6.73	7.40
	.10	4.71	5.68	6.48	7.18	7.80	8.36	8.88	9.36	10.2
	.05	6.00	7.18	8.12	9.03	9.78	10.5	11.1	11.7	12.7
	.01	9.9	11.7	13.2	14.5	15.8	16.9	17.9	18.9	21
9	.25	3.03	3.64	4.15	4.59	4.98	5.34	5.66	5.96	6.51
	.10	4.26	5.07	5.74	6.32	6.82	7.28	7.70	8.09	8.78
	.05	5.34	6.31	7.11	7.80	8.41	8.95	9.45	9.91	10.7
	.01	8.5	9.9	11.1	12.1	13.1	13.9	14.7	15.3	16.6

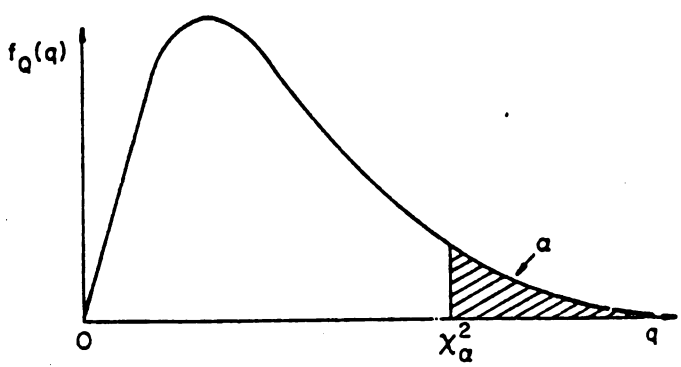
Table B.3. Upper Percentage Points of F_{max} Distribution (1-CDF): Equal Replication (cont.)

		t=3	4	5	6	7	8	9	10	12
10	.25	2.85	3.59	3.83	4.21	4.55	4.86	5.13	5.39	5.85
	.10	3.93	4.63	5.19	5.68	6.11	6.49	6.84	7.16	7.74
	.05	4.85	5.67	6.34	6.92	7.42	7.87	8.28	8.66	9.34
	.01	7.4	8.6	9.6	10.4	11.1	11.8	12.4	12.9	13.9
12	.25	2.58	3.02	3.38	3.68	3.95	4.18	4.40	4.60	4.95
	.10	3.45	4.00	4.44	4.81	5.13	5.42	5.68	5.92	6.35
	.05	4.16	4.79	5.30	5.72	6.09	6.42	6.72	7.00	7.48
	.01	6.1	6.9	7.6	8.2	8.7	9.1	9.5	9.9	10.6
15	.25	2.32	2.67	2.95	3.18	3.38	3.56	3.72	3.87	4.13
	.10	3.00	3.41	3.74	4.02	4.25	4.46	4.65	4.82	5.13
	.05	3.54	4.01	4.37	4.68	4.95	5.19	5.40	5.59	5.93
	.01	4.9	5.5	6.0	6.4	6.7	7.1	7.3	7.5	8.0
20	.25	2.07	2.33	2.53	2.70	2.85	2.98	3.09	3.20	3.38
	.10	2.57	2.87	3.10	3.29	3.46	3.60	3.73	3.85	4.06
	.05	2.95	3.29	3.54	3.76	3.94	4.10	4.24	4.37	4.59
	.01	3.8	4.3	4.6	4.9	5.1	5.3	5.5	5.6	5.9
30	.25	1.80	1.98	2.12	2.24	2.34	2.42	2.49	2.56	2.68
	.10	2.14	2.34	2.50	2.62	2.73	2.82	2.90	2.97	3.10
	.05	2.40	2.61	2.78	2.91	3.02	3.12	3.21	3.29	3.39
	.01	3.0	3.3	3.4	3.6	3.7	3.8	3.9	4.0	4.2
60	.25	1.51	1.62	1.70	1.76	1.81	1.86	1.90	1.93	2.00
	.10	1.71	1.82	1.90	1.96	2.02	2.07	2.11	2.14	2.21
	.05	1.85	1.96	2.04	2.11	2.17	2.22	2.26	2.30	2.36
	.01	2.2	2.3	2.4	2.4	2.5	2.5	2.6	2.6	2.7
∞		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Source: Gill (1978)

Table B. 4. Upper Percentage Points of Chi-Square Distribution

ν	α: 0.3	0.2	0.1	0.05	0.025	0.01	0.005	0.001
1	1.074	1.642	2.706	3.841	5.024	6.635	7.879	10.83
2	2.408	3.219	4.605	5.991	7.378	9.210	10.60	13.82
3	3.665	4.642	6.251	7.815	9.348	11.34	12.84	16.27
4	4.878	5.989	7.779	9.488	11.14	13.28	14.86	18.47
5	6.064	7.289	9.236	11.07	12.83	15.09	16.75	20.52
6	7.231	8.558	10.64	12.59	14.45	16.81	18.55	22.46
7	8.383	9.803	12.02	14.07	16.01	18.48	20.28	24.32
8	9.524	11.03	13.36	15.51	17.53	20.09	21.96	26.12
9	10.66	12.24	14.68	16.92	19.02	21.67	23.59	27.88
10	11.78	13.44	15.99	18.31	20.48	23.21	25.19	29.59
11	12.90	14.63	17.28	19.68	21.92	24.72	26.76	31.26
12	14.01	15.81	18.55	21.03	23.34	26.22	28.30	32.91
13	15.12	16.98	19.81	22.36	24.74	27.69	29.82	34.53
14	16.22	18.15	21.06	23.68	26.12	29.14	31.32	36.12
15	17.32	19.31	22.31	25.00	27.49	30.58	32.80	37.70
16	18.42	20.47	23.54	26.30	28.85	32.00	34.27	39.25
17	19.51	21.61	24.77	27.59	30.19	33.41	35.72	40.79
18	20.60	22.76	25.99	28.87	31.53	34.81	37.16	42.31
19	21.69	23.90	27.20	30.14	32.85	36.19	38.58	43.82
20	22.77	25.04	28.41	31.41	34.17	37.57	40.00	45.31
21	23.86	26.17	29.62	32.67	35.48	38.93	41.40	46.80
22	24.94	27.30	30.81	33.92	36.78	40.29	42.80	48.27
23	26.02	28.43	32.01	35.17	38.08	41.64	44.18	49.73
24	27.10	29.55	33.20	36.42	39.36	42.98	45.56	51.18
25	28.17	30.68	34.38	37.65	40.65	44.31	46.93	52.62
26	29.25	31.79	35.56	38.89	41.92	45.64	48.29	54.05
27	30.32	32.91	36.74	40.11	43.19	46.96	49.64	55.48
28	31.39	34.03	37.92	41.34	44.46	48.28	50.99	56.89
29	32.46	35.14	39.09	42.56	45.72	49.59	52.34	58.30
30	33.53	36.25	40.26	43.77	46.98	50.89	53.67	59.70



$P[Q > \chi^2_\alpha] = \alpha$. For two-tailed procedures, table should be entered at percentage corresponding to $\alpha/2$.

Table B. 4. Upper Percentage Points of Chi-Square Distribution (cont.)

ν	α: 0.3	0.2	0.1	0.05	0.025	0.01	0.005	0.001
31	34.60	37.36	41.42	44.99	48.23	52.19	55.00	61.10
32	35.66	38.47	42.58	46.19	49.48	53.49	56.33	62.49
33	36.73	39.57	43.75	47.40	50.73	54.78	57.65	63.87
34	37.80	40.68	44.90	48.60	51.97	56.06	58.96	65.25
35	38.86	41.78	46.06	49.80	53.20	57.34	60.27	66.62
36	39.92	42.88	47.21	51.00	54.44	58.62	61.58	67.99
37	40.98	43.98	48.36	52.19	55.67	59.89	62.88	69.35
38	42.05	45.08	49.51	53.38	56.90	61.16	64.18	70.70
39	43.11	46.17	50.66	54.57	58.12	62.43	65.48	72.05
40	44.16	47.27	51.81	55.76	59.34	63.69	66.77	73.40
41	45.22	48.36	52.95	56.94	60.56	64.95	68.05	74.74
42	46.28	49.46	54.09	58.12	61.78	66.21	69.34	76.08
43	47.34	50.55	55.23	59.30	62.99	67.46	70.62	77.42
44	48.40	51.64	56.37	60.48	64.20	68.71	71.89	78.75
45	49.45	52.73	57.51	61.66	65.41	69.96	73.17	80.08
46	50.51	53.82	58.64	62.83	66.62	71.20	74.44	81.40
47	51.56	54.91	59.77	64.00	67.82	72.44	75.70	82.72
48	52.62	55.99	60.91	65.17	69.02	73.68	76.97	84.04
49	53.67	57.08	62.04	66.34	70.22	74.92	78.23	85.35
50	54.72	58.16	63.17	67.50	71.42	76.15	79.49	86.66
51	55.78	59.25	64.30	68.67	72.62	77.39	80.75	87.97
52	56.83	60.33	65.42	69.83	73.81	78.62	82.00	89.27
53	57.88	61.41	66.55	70.99	75.00	79.84	83.25	90.57
54	58.93	62.50	67.67	72.15	76.19	81.07	84.50	91.87
55	59.98	63.58	68.80	73.31	77.38	82.29	85.75	93.17
56	61.03	64.66	69.92	74.47	78.57	83.51	86.99	94.46
57	62.08	65.74	71.04	75.62	79.75	84.73	88.24	95.75
58	63.13	66.82	72.16	76.78	80.94	85.95	89.48	97.04
59	64.18	67.89	73.28	77.93	82.12	87.17	90.72	98.32
60	65.23	68.97	74.40	79.08	83.30	88.38	91.95	99.61
61	66.27	70.05	75.51	80.23	84.48	89.59	93.19	100.9
62	67.32	71.13	76.63	81.38	85.65	90.80	94.42	102.2
63	68.37	72.20	77.75	82.53	86.83	92.01	95.65	103.4
64	69.42	73.28	78.86	83.68	88.00	93.22	96.88	104.7
65	70.46	74.35	79.97	84.82	89.18	94.42	98.11	106.0
66	71.51	75.42	81.09	85.96	90.35	95.63	99.33	107.3
67	72.55	76.50	82.20	87.11	91.52	96.83	100.6	108.5
68	73.60	77.57	83.31	88.25	92.69	98.03	101.8	109.8
69	74.64	78.64	84.42	89.39	93.86	99.23	103.0	111.1
70	75.69	79.71	85.53	90.53	95.02	100.4	104.2	112.3

Table B.4. Upper Percentage Points of Chi-Square Distribution (cont.)

ν	α: 0.3	0.2	0.1	0.05	0.025	0.01	0.005	0.001
71	76.73	80.79	86.64	91.67	96.19	101.6	105.4	113.6
72	77.78	81.86	87.74	92.81	97.35	102.8	106.6	114.8
73	78.82	82.93	88.85	93.95	98.52	104.0	107.9	116.1
74	79.86	83.00	89.96	95.08	99.68	105.2	109.1	117.3
75	80.91	85.07	91.06	96.22	100.8	106.4	110.3	118.6
76	81.95	86.13	92.17	97.35	102.0	107.6	111.5	119.8
77	82.99	87.20	93.27	98.48	103.2	108.8	112.7	121.1
78	84.04	88.27	94.37	99.62	104.3	100.0	113.9	122.3
79	85.08	89.34	95.48	100.7	105.5	111.1	115.1	123.6
80	86.12	90.41	96.58	101.9	106.6	112.3	116.3	124.8
81	87.16	91.47	97.68	103.0	107.8	113.5	117.5	126.1
82	88.20	92.54	98.78	104.1	108.9	114.7	118.7	127.3
83	89.24	93.60	99.88	105.3	110.1	115.9	119.9	128.6
84	90.28	94.67	101.0	106.4	111.2	117.1	121.1	129.8
85	91.32	95.73	102.1	107.5	112.4	118.2	122.3	131.0
86	92.36	96.80	103.2	108.6	113.5	119.4	123.5	132.3
87	93.40	97.86	104.3	109.8	114.7	120.6	124.7	133.5
88	94.44	98.93	105.4	110.9	115.8	121.8	125.9	134.7
89	95.48	99.99	106.5	112.0	117.0	122.9	127.1	136.0
90	96.52	101.1	107.6	113.1	118.1	124.1	128.3	137.2
91	97.56	102.1	108.7	114.3	119.3	125.3	129.5	138.4
92	98.60	103.2	109.8	115.4	120.4	126.5	130.7	139.7
93	99.64	104.2	110.8	116.5	121.6	127.6	131.9	140.9
94	100.7	105.3	111.9	117.6	122.7	128.8	133.1	142.1
95	101.7	106.4	113.0	118.8	123.9	130.0	134.2	143.3
96	102.8	107.4	114.1	119.9	125.0	131.1	135.4	144.6
97	103.8	108.5	115.2	121.0	126.1	132.3	136.6	145.8
98	104.8	109.5	116.3	122.1	127.3	133.5	137.8	147.0
99	105.9	110.6	117.4	123.2	128.4	134.6	139.0	148.2
100*	106.9	111.7	118.5	124.3	129.6	135.8	140.2	149.4

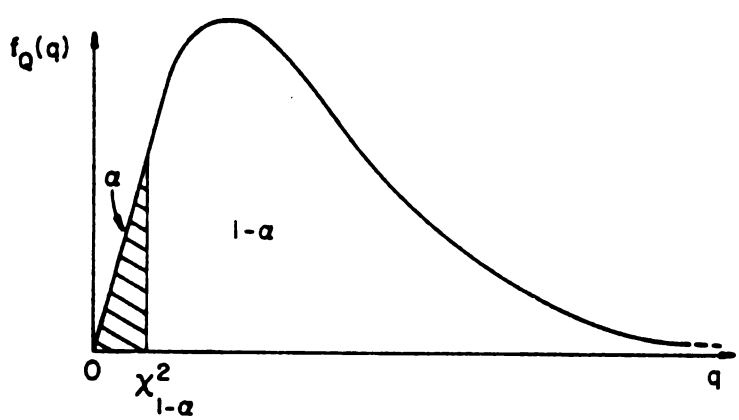
Source: Gill (1978)

*For $\nu > 100$, one may use the approximation $\chi^2_{\alpha,\nu} \approx (z_{1-\alpha}+\sqrt{2\nu-1})^2/2$, where $z_{1-\alpha}$ is an upper percentage point from the standard normal distribution (Table A.2).

Table B. 5. Lower Percentage Points of Chi-Square Distribution

ν \ $1-\alpha$:	0.999	0.995	0.99	0.975	0.95	0.9	0.8	0.7
1	0.002*	0.039*	0.157*	0.982*	0.004	0.016	0.064	0.148
2	0.002	0.010	0.020	0.051	0.103	0.211	0.446	0.713
3	0.024	0.072	0.115	0.216	0.352	0.584	1.005	1.424
4	0.091	0.207	0.297	0.484	0.710	1.064	1.649	2.195
5	0.210	0.412	0.554	0.831	1.145	1.610	2.343	3.000
6	0.381	0.676	0.872	1.237	1.635	2.204	3.070	3.828
7	0.598	0.989	1.239	1.690	2.167	2.833	3.822	4.671
8	0.857	1.344	1.646	2.180	2.733	3.490	4.594	5.527
9	1.152	1.735	2.088	2.700	3.325	4.168	5.380	6.393
10	1.479	2.156	2.558	3.247	3.940	4.865	6.179	7.267
11	1.834	2.603	3.053	3.816	4.575	5.578	6.989	8.148
12	2.214	3.074	3.571	4.404	5.226	6.304	7.807	9.034
13	2.617	3.565	4.107	5.009	5.892	7.042	8.634	9.926
14	3.041	4.075	4.660	5.629	6.571	7.790	9.467	10.82
15	3.483	4.601	5.229	6.262	7.261	8.547	10.31	11.72
16	3.942	5.142	5.812	6.908	7.962	9.312	11.15	12.62
17	4.416	5.697	6.408	7.564	8.672	10.09	12.00	13.53
18	4.905	6.265	7.015	8.231	9.390	10.86	12.86	14.44
19	5.407	6.844	7.633	8.907	10.12	11.65	13.72	15.35
20	5.921	7.434	8.260	9.591	10.85	12.44	14.58	16.27
21	6.447	8.034	8.897	10.28	11.59	13.24	15.44	17.18
22	6.983	8.643	9.542	10.98	12.34	14.04	16.31	18.10
23	7.529	9.260	10.20	11.69	13.09	14.85	17.19	19.02
24	8.085	9.886	10.86	12.40	13.85	15.66	18.06	19.94
25	8.649	10.52	11.52	13.12	14.61	16.47	18.94	20.87
26	9.222	11.16	12.20	13.84	15.38	17.29	19.82	21.79
27	9.803	11.81	12.88	14.57	16.15	18.11	20.70	22.72
28	10.39	12.46	13.56	15.31	16.93	18.94	21.59	23.65
29	10.99	13.12	14.26	16.05	17.71	19.77	22.48	24.58
30	11.59	13.79	14.95	16.79	18.49	20.60	23.36	25.51

*Divide these entries by 1000.



$P[Q > \chi^2_{1-\alpha}] = 1-\alpha$. For two-tailed procedures, table should be entered at percentage corresponding to $1-\alpha/2$.

Table B. 5. Lower Percentage Points of Chi-Square Distribution (cont.)

ν 1-α:	0.999	0.995	0.99	0.975	0.95	0.9	0.8	0.7
31	12.20	14.46	15.66	17.54	19.28	21.43	24.26	26.44
32	12.81	15.13	16.36	18.29	20.07	22.27	25.15	27.37
33	13.43	15.82	17.07	19.05	20.87	23.11	26.04	28.31
34	14.06	16.50	17.79	19.81	21.66	23.95	26.94	29.24
35	14.69	17.19	18.51	20.57	22.46	24.80	27.84	30.18
36	15.32	17.89	19.23	21.34	23.27	25.64	28.74	31.12
37	15.97	18.59	19.96	22.11	24.07	26.49	29.64	32.05
38	16.61	19.29	20.69	22.88	24.88	27.34	30.54	32.99
39	17.26	20.00	21.43	23.65	25.70	28.20	31.44	33.93
40	17.92	20.71	22.16	24.43	26.51	29.05	32.34	34.87
41	18.58	21.42	22.91	25.21	27.33	29.91	33.25	35.81
42	19.24	22.14	23.65	26.00	28.14	30.77	34.16	36.76
43	19.91	22.86	24.40	26.79	28.96	31.63	35.07	37.70
44	20.58	23.58	25.15	27.57	29.79	32.49	35.97	38.64
45	21.25	24.31	25.90	28.37	30.61	33.35	36.88	39.58
46	21.93	25.04	26.66	29.16	31.44	34.22	37.80	40.53
47	22.61	25.77	27.42	29.96	32.27	35.08	38.71	41.47
48	23.29	26.51	28.18	30.75	33.10	35.95	39.62	42.42
49	23.98	27.25	28.94	31.55	33.93	36.82	40.53	43.37
50	24.67	27.99	29.71	32.36	34.76	37.69	41.45	44.31
51	25.37	28.73	30.48	33.16	35.60	38.56	42.36	45.26
52	26.07	29.48	31.25	33.97	36.44	39.43	43.28	46.21
53	26.76	30.23	32.02	34.78	37.28	40.31	44.20	47.16
54	27.47	30.98	32.79	35.59	38.12	41.18	45.12	48.11
55	28.17	31.73	33.57	36.40	38.96	42.06	46.04	49.06
56	28.88	32.49	34.35	37.21	39.80	42.94	46.96	50.01
57	29.59	33.25	35.13	38.03	40.65	43.82	47.88	50.96
58	30.30	34.01	35.91	38.84	41.49	44.70	48.80	51.91
59	31.02	34.77	36.70	39.66	42.34	45.58	49.72	52.86
60	31.74	35.53	37.48	40.48	43.19	46.46	50.64	53.81
61	32.46	36.30	38.27	41.30	44.04	47.34	51.56	54.76
62	33.18	37.07	39.06	42.13	44.89	48.23	52.49	55.71
63	33.91	37.84	39.86	42.95	45.74	49.11	53.41	56.67
64	34.63	38.61	40.65	43.78	46.59	50.00	54.34	57.62
65	35.36	39.38	41.44	44.60	47.45	50.88	55.26	58.57
66	36.09	40.16	42.24	45.43	48.31	51.77	56.19	59.53
67	36.83	40.94	43.04	46.26	49.16	52.66	57.11	60.48
68	37.56	41.71	43.84	47.09	50.02	53.55	58.04	61.44
69	38.30	42.49	44.64	47.92	50.88	54.44	58.97	62.39
70	39.04	43.28	45.44	48.76	51.74	55.33	59.90	63.35

Table B. 5. Lower Percentage Points of Chi-Square Distribution (cont.)

ν 1-α:	0.999	0.995	0.99	0.975	0.95	0.9	0.8	0.7
71	39.78	44.06	46.25	49.59	52.60	56.22	60.83	64.30
72	40.52	44.84	47.05	50.43	53.46	57.11	61.76	65.26
73	41.26	45.63	47.86	51.26	54.33	58.01	62.69	66.21
74	42.01	46.42	48.67	52.10	55.19	58.90	63.62	67.17
75	42.76	47.21	49.48	52.94	56.05	59.79	64.55	68.13
76	43.51	48.00	50.29	53.78	56.92	60.69	65.48	69.08
77	44.26	48.79	51.10	54.62	57.79	61.59	66.41	70.04
78	45.01	49.58	51.91	55.47	58.65	62.48	67.34	71.00
79	45.76	50.38	52.72	56.31	59.52	63.38	68.27	71.96
80	46.52	51.17	53.54	57.15	60.39	64.28	69.21	72.92
81	47.28	51.97	54.36	58.00	61.26	65.18	70.14	73.87
82	48.04	52.77	55.17	58.84	62.13	66.08	71.07	74.83
83	48.80	53.57	55.99	59.69	63.00	66.98	72.01	75.79
84	49.56	54.37	56.81	60.54	63.88	67.88	72.94	76.75
85	50.32	55.17	57.63	61.39	64.75	68.78	73.88	77.71
86	51.08	55.97	58.46	62.24	65.62	69.68	74.81	78.67
87	51.85	56.78	59.28	63.09	66.50	70.58	75.75	79.63
88	52.62	57.58	60.10	63.94	67.37	71.48	76.69	80.59
89	53.39	58.39	60.93	64.79	68.25	72.39	77.62	81.55
90	54.16	59.20	61.75	65.65	69.13	73.29	78.56	82.51
91	54.93	60.00	62.58	66.50	70.00	74.20	79.50	83.47
92	55.70	60.81	63.41	67.36	70.88	75.10	80.43	84.43
93	56.47	61.63	64.24	68.21	71.76	76.01	81.37	85.39
94	57.25	62.44	65.07	69.07	72.64	76.91	82.31	86.36
95	58.02	63.25	65.90	69.92	73.52	77.82	83.25	87.32
96	58.80	64.06	66.73	70.78	74.40	78.73	84.19	88.28
97	59.58	64.88	67.56	71.64	75.28	79.63	85.13	89.24
98	60.36	65.69	68.40	72.50	76.16	80.54	86.07	90.20
99	61.14	66.51	69.23	73.36	77.05	81.45	87.01	91.17
100*	61.92	67.33	70.06	74.22	77.93	82.36	87.95	92.13

Source: Gill (1978)

*For $\nu > 100$, one may use the approximation, $\chi^2_{1-\alpha,\nu} = (z_\alpha + \sqrt{2\nu-1})^2/2$, where z_α is a lower percentage point from the standard normal distribution (Table

TABLE B. 6. CRITICAL VALUES FOR TESTING ONE OUTLYING OBSERVATION

n	α=0.01	0.05	0.10	n	α=0.01	0.05	0.10
...	...	...	...	32	3.135	2.773	2.591
...	...	...	...	34	3.164	2.799	2.616
3	1.155	1.153	1.148	36	3.191	2.823	2.639
4	1.492	1.463	1.425	38	3.216	2.846	2.661
5	1.749	1.672	1.602	40	3.240	2.866	2.682
6	1.944	1.822	1.729	42	3.261	2.887	2.700
7	2.097	1.938	1.828	44	3.282	2.905	2.719
8	2.221	2.032	1.909	46	3.302	2.923	2.736
9	2.323	2.110	1.977	48	3.319	2.940	2.753
10	2.410	2.176	2.036	50	3.336	2.956	2.768
11	2.485	2.234	2.088	55	3.376	2.992	2.804
12	2.550	2.285	2.134	60	3.411	3.025	2.837
13	2.607	2.331	2.175	65	3.442	3.055	2.866
14	2.659	2.371	2.213	70	3.471	3.082	2.893
15	2.705	2.409	2.247	75	3.496	3.107	2.917
16	2.747	2.443	2.279	80	3.521	3.130	2.940
17	2.785	2.475	2.309	85	3.543	3.151	2.961
18	2.821	2.504	2.335	90	3.563	3.171	2.981
19	2.854	2.532	2.361	95	3.582	3.189	3.000
20	2.884	2.557	2.385	100	3.600	3.207	3.017
21	2.912	2.580	2.408	105	3.617	3.224	3.033
22	2.939	2.603	2.429	110	3.632	3.239	3.049
23	2.963	2.624	2.448	115	3.647	3.254	3.064
24	2.987	2.644	2.467	120	3.662	3.267	3.078
25	3.009	2.663	2.486	125	3.675	3.281	3.092
26	3.029	2.681	2.502	130	3.688	3.294	3.104
27	3.049	2.698	2.519	135	3.700	3.306	3.116
28	3.068	2.714	2.534	140	3.712	3.318	3.129
29	3.085	2.730	2.549	145	3.723	3.328	3.140
30	3.103	2.745	2.563				

Source: Gill (1978)

TABLE B. 7. *Signifance points of the Durbin-Watson statistics* d_L *and* d_U: $\alpha = .01$.

	$k' = 1$		$k' = 2$		$k' = 3$		$k' = 4$		$k' = 5$	
n	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0·81	1·07	0·70	1·25	0·59	1·46	0·49	1·70	0·39	1·96
16	0·84	1·09	0·74	1·25	0·63	1·44	0·53	1·66	0·44	1·90
17	0·87	1·10	0·77	1·25	0·67	1·43	0·57	1·63	0·48	1·85
18	0·90	1·12	0·80	1·26	0·71	1·42	0·61	1·60	0·52	1·80
19	0·93	1·13	0·83	1·26	0·74	1·41	0·65	1·58	0·56	1·77
20	0·95	1·15	0·86	1·27	0·77	1·41	0·68	1·57	0·60	1·74
21	0·97	1·16	0·89	1·27	0·80	1·41	0·72	1·55	0·63	1·71
22	1·00	1·17	0·91	1·28	0·83	1·40	0·75	1·54	0·66	1·69
23	1·02	1·19	0·94	1·29	0·86	1·40	0·77	1·53	0·70	1·67
24	1·04	1·20	0·96	1·30	0·88	1·41	0·80	1·53	0·72	1·66
25	1·05	1·21	0·98	1·30	0·90	1·41	0·83	1·52	0·75	1·65
26	1·07	1·22	1·00	1·31	0·93	1·41	0·85	1·52	0·78	1·64
27	1·09	1·23	1·02	1·32	0·95	1·41	0·88	1·51	0·81	1·63
28	1·10	1·24	1·04	1·32	0·97	1·41	0·90	1·51	0·83	1·62
29	1·12	1·25	1·05	1·33	0·99	1·42	0·92	1·51	0·85	1·61
30	1·13	1·26	1·07	1·34	1·01	1·42	0·94	1·51	0·88	1·61
31	1·15	1·27	1·08	1·34	1·02	1·42	0·96	1·51	0·90	1·60
32	1·16	1·28	1·10	1·35	1·04	1·43	0·98	1·51	0·92	1·60
33	1·17	1·29	1·11	1·36	1·05	1·43	1·00	1·51	0·94	1·59
34	1·18	1·30	1·13	1·36	1·07	1·43	1·01	1·51	0·95	1·59
35	1·19	1·31	1·14	1·37	1·08	1·44	1·03	1·51	0·97	1·59
36	1·21	1·32	1·15	1·38	1·10	1·44	1·04	1·51	0·99	1·59
37	1·22	1·32	1·16	1·38	1·11	1·45	1·06	1·51	1·00	1·59
38	1·23	1·33	1·18	1·39	1·12	1·45	1·07	1·52	1·02	1·58
39	1·24	1·34	1·19	1·39	1·14	1·45	1·09	1·52	1·03	1·58
40	1·25	1·34	1·20	1·40	1·15	1·46	1·10	1·52	1·05	1·58
45	1·29	1·38	1·24	1·42	1·20	1·48	1·16	1·53	1·11	1·58
50	1·32	1·40	1·28	1·45	1·24	1·49	1·20	1·54	1·16	1·59
55	1·36	1·43	1·32	1·47	1·28	1·51	1·25	1·55	1·21	1·59
60	1·38	1·45	1·35	1·48	1·32	1·52	1·28	1·56	1·25	1·60
65	1·41	1·47	1·38	1·50	1·35	1·53	1·31	1·57	1·28	1·61
70	1·43	1·49	1·40	1·52	1·37	1·55	1·34	1·58	1·31	1·61
75	1·45	1·50	1·42	1·53	1·39	1·56	1·37	1·59	1·34	1·62
80	1·47	1·52	1·44	1·54	1·42	1·57	1·39	1·60	1·36	1·62
85	1·48	1·53	1·46	1·55	1·43	1·58	1·41	1·60	1·39	1·63
90	1·50	1·54	1·47	1·56	1·45	1·59	1·43	1·61	1·41	1·64
95	1·51	1·55	1·49	1·57	1·47	1·60	1·45	1·62	1·42	1·64
100	1·52	1·56	1·50	1·58	1·48	1·60	1·46	1·63	1·44	1·65

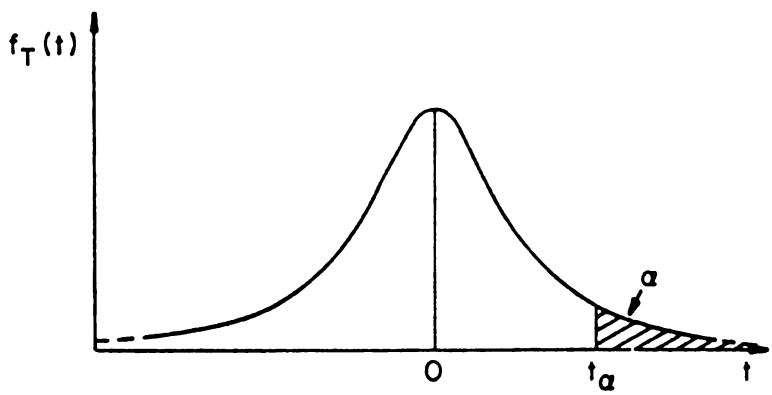
TABLE B.7. *Significance points of the Durbin-Watson statistics* d_L *and* d_U: $\alpha = .05$.

n	$k' = 1$		$k' = 2$		$k' = 3$		$k' = 4$		$k' = 5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1·08	1·36	0·95	1·54	0·82	1·75	0·69	1·97	0·56	2·21
16	1·10	1·37	0·98	1·54	0·86	1·73	0·74	1·93	0·62	2·15
17	1·13	1·38	1·02	1·54	0·90	1·71	0·78	1·90	0·67	2·10
18	1·16	1·39	1·05	1·53	0·93	1·69	0·82	1·87	0·71	2·06
19	1·18	1·40	1·08	1·53	0·97	1·68	0·86	1·85	0·75	2·02
20	1·20	1·41	1·10	1·54	1·00	1·68	0·90	1·83	0·79	1·99
21	1·22	1·42	1·13	1·54	1·03	1·67	0·93	1·81	0·83	1·96
22	1·24	1·43	1·15	1·54	1·05	1·66	0·96	1·80	0·86	1·94
23	1·26	1·44	1·17	1·54	1·08	1·66	0·99	1·79	0·90	1·92
24	1·27	1·45	1·19	1·55	1·10	1·66	1·01	1·78	0·93	1·90
25	1·29	1·45	1·21	1·55	1·12	1·66	1·04	1·77	0·95	1·89
26	1·30	1·46	1·22	1·55	1·14	1·65	1·06	1·76	0·98	1·88
27	1·32	1·47	1·24	1·56	1·16	1·65	1·08	1·76	1·01	1·86
28	1·33	1·48	1·26	1·56	1·18	1·65	1·10	1·75	1·03	1·85
29	1·34	1·48	1·27	1·56	1·20	1·65	1·12	1·74	1·05	1·84
30	1·35	1·49	1·28	1·57	1·21	1·65	1·14	1·74	1·07	1·83
31	1·36	1·50	1·30	1·57	1·23	1·65	1·16	1·74	1·09	1·83
32	1·37	1·50	1·31	1·57	1·24	1·65	1·18	1·73	1·11	1·82
33	1·38	1·51	1·32	1·58	1·26	1·65	1·19	1·73	1·13	1·81
34	1·39	1·51	1·33	1·58	1·27	1·65	1·21	1·73	1·15	1·81
35	1·40	1·52	1·34	1·58	1·28	1·65	1·22	1·73	1·16	1·80
36	1·41	1·52	1·35	1·59	1·29	1·65	1·24	1·73	1·18	1·80
37	1·42	1·53	1·36	1·59	1·31	1·66	1·25	1·72	1·19	1·80
38	1·43	1·54	1·37	1·59	1·32	1·66	1·26	1·72	1·21	1·79
39	1·43	1·54	1·38	1·60	1·33	1·66	1·27	1·72	1·22	1·79
40	1·44	1·54	1·39	1·60	1·34	1·66	1·29	1·72	1·23	1·79
45	1·48	1·57	1·43	1·62	1·38	1·67	1·34	1·72	1·29	1·78
50	1·50	1·59	1·46	1·63	1·42	1·67	1·38	1·72	1·34	1·77
55	1·53	1·60	1·49	1·64	1·45	1·68	1·41	1·72	1·38	1·77
60	1·55	1·62	1·51	1·65	1·48	1·69	1·44	1·73	1·41	1·77
65	1·57	1·63	1·54	1·66	1·50	1·70	1·47	1·73	1·44	1·77
70	1·58	1·64	1·55	1·67	1·52	1·70	1·49	1·74	1·46	1·77
75	1·60	1·65	1·57	1·68	1·54	1·71	1·51	1·74	1·49	1·77
80	1·61	1·66	1·59	1·69	1·56	1·72	1·53	1·74	1·51	1·77
85	1·62	1·67	1·60	1·70	1·57	1·72	1·55	1·75	1·52	1·77
90	1·63	1·68	1·61	1·70	1·59	1·73	1·57	1·75	1·54	1·78
95	1·64	1·69	1·62	1·71	1·60	1·73	1·58	1·75	1·56	1·78
100	1·65	1·69	1·63	1·72	1·61	1·74	1·59	1·76	1·57	1·78

Source: Kendal (1973)

TABLE B.8. UPPER PERCENTAGE POINTS OF STUDENT'S t DISTRIBUTION (1-CDF)

	α								
ν	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.0005
1	1.000	1.376	1.963	3.078	6.314	12.706	31.82	63.66	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.60
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.500	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	0.698	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	0.696	0.873	1.083	1.356	1.782	2.179	2.681	3.054	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	0.691	0.866	1.074	1.341	1.753	2.132	2.602	2.947	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.767
24	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.646



$P[T > t_\alpha] = \alpha$. For two-tailed procedures, table should be entered at column headed by desired value of $\alpha/2$. In all cases, ν = degrees of freedom.

TABLE B.8. UPPER PERCENTAGE POINTS OF STUDENT'S t DISTRIBUTION (1-CDF) (cont.)

	α								
ν	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.0005
31	0.682	0.854	1.054	1.310	1.696	2.040	2.453	2.744	3.634
32	0.682	0.853	1.054	1.309	1.694	2.037	2.449	2.738	3.622
33	0.682	0.853	1.053	1.308	1.692	2.034	2.445	2.733	3.611
34	0.682	0.852	1.053	1.307	1.691	2.032	2.441	2.728	3.601
35	0.682	0.852	1.052	1.306	1.690	2.030	2.438	2.724	3.592
36	0.681	0.852	1.052	1.306	1.688	2.028	2.434	2.720	3.582
37	0.681	0.852	1.051	1.305	1.687	2.026	2.431	2.716	3.574
38	0.681	0.851	1.051	1.304	1.686	2.024	2.428	2.712	3.566
39	0.681	0.851	1.050	1.304	1.685	2.023	2.426	2.708	3.559
40	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551
42	0.680	0.850	1.049	1.302	1.682	2.018	2.418	2.698	3.538
44	0.680	0.850	1.049	1.301	1.680	2.015	2.414	2.692	3.526
46	0.680	0.850	1.048	1.300	1.679	2.013	2.410	2.687	3.515
48	0.680	0.849	1.048	1.299	1.677	2.011	2.406	2.682	3.505
50	0.679	0.849	1.047	1.299	1.676	2.009	2.403	2.678	3.496
60	0.679	0.848	1.046	1.296	1.671	2.000	2.390	2.660	3.461
70	0.678	0.847	1.044	1.294	1.667	1.994	2.381	2.648	3.436
80	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.417
90	0.677	0.846	1.042	1.291	1.662	1.987	2.368	2.632	3.402
100	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.391
120	0.676	0.845	1.041	1.289	1.658	1.980	2.358	2.618	3.374
140	0.676	0.844	1.040	1.288	1.656	1.977	2.353	2.611	3.362
160	0.676	0.844	1.040	1.287	1.654	1.975	2.350	2.607	3.353
180	0.676	0.844	1.039	1.286	1.653	1.973	2.347	2.604	3.346
200	0.676	0.843	1.039	1.286	1.652	1.972	2.345	2.601	3.340
300	0.676	0.843	1.038	1.285	1.650	1.968	2.338*	2.592	3.323
400	0.676	0.843	1.038	1.284	1.649	1.966	2.335*	2.588	3.315
500	0.676	0.843	1.037	1.284	1.648	1.965	2.334*	2.586	3.310
1000	0.675	0.842	1.037	1.283	1.647	1.962	2.330*	2.581	3.301
∞	0.6745	0.8416	1.0364	1.2816	1.6448	1.9600	2.3263	2.5758	3.2905

Source: Gill (1978)

Table B.9. Quantiles of the Mann-Whitney Test Statistic[a]

n	p	$m = 2$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	.001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	.005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
2	.01	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	2	2
	.025	0	0	0	0	0	0	1	1	1	1	2	2	2	2	2	3	3	3	3
	.05	0	0	0	1	1	1	2	2	2	2	3	3	4	4	4	4	5	5	5
	.10	0	1	1	2	2	2	3	3	4	4	5	5	5	6	6	7	7	8	8
	.001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
	.005	0	0	0	0	0	0	0	1	1	1	2	2	2	3	3	3	3	4	4
3	.01	0	0	0	0	0	1	1	2	2	2	3	3	3	4	4	5	5	5	6
	.025	0	0	0	1	2	2	3	3	4	4	5	5	6	6	7	7	8	8	9
	.05	0	1	1	2	3	3	4	5	5	6	6	7	8	8	9	10	10	11	12
	.10	1	2	2	3	4	5	6	6	7	8	9	10	11	11	12	13	14	15	16
	.001	0	0	0	0	0	0	0	0	1	1	1	2	2	2	3	3	4	4	4
	.005	0	0	0	0	1	1	2	2	3	3	4	4	5	6	6	7	7	8	9
4	.01	0	0	0	1	2	2	3	4	4	5	6	6	7	9	8	9	10	10	11
	.025	0	0	1	2	3	4	5	5	6	7	8	9	10	11	12	12	13	14	15
	.05	0	1	2	3	4	5	6	7	8	9	10	11	12	13	15	16	17	18	19
	.10	1	2	4	5	6	7	8	10	11	12	13	14	16	17	18	19	21	22	23

SOURCE: **Conover (1971)**

[a] The entries in this table are the quantiles w_p of the Mann-Whitney test statistic T, given by Equation (5.3.2), for selected values of p. Note that $P(T < w_p) \leq p$. Upper quantiles may be found from the equation

$$w_{1-p} = mn - w_p$$

Critical regions correspond to values less than (or greater than) but not including the appropriate quantile.

Table B.9. (CONTINUED)

n	*p*	*m* = 2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	.001	0	0	0	0	0	0	1	2	2	3	3	4	4	5	6	6	7	8	8
	.005	0	0	0	1	2	2	3	4	5	6	7	8	8	9	10	11	12	13	14
5	.01	0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
	.025	0	1	2	3	4	6	7	8	9	10	12	13	14	15	16	18	19	20	21
	.05	1	2	3	5	6	7	9	10	12	13	14	16	17	19	20	21	23	24	26
	.10	2	3	5	6	8	9	11	13	14	16	18	19	21	23	24	26	28	29	31
	.001	0	0	0	0	0	0	2	3	4	5	5	6	7	8	9	10	11	12	13
	.005	0	0	1	2	3	4	5	6	7	8	10	11	12	13	14	16	17	18	19
6	.01	0	0	2	3	4	5	7	8	9	10	12	13	14	16	17	19	20	21	23
	.025	0	2	3	4	6	7	9	11	12	14	15	17	18	20	22	23	25	26	28
	.05	1	3	4	6	8	9	11	13	15	17	18	20	22	24	26	27	29	31	33
	.10	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	35	37	39
	.001	0	0	0	0	1	2	3	4	6	7	8	9	10	11	12	14	15	16	17
	.005	0	0	1	2	4	5	7	8	10	11	13	14	16	17	19	20	22	23	25
7	.01	0	1	2	4	5	7	8	10	12	13	15	17	18	20	22	24	25	27	29
	.025	0	2	4	6	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35
	.05	1	3	5	7	9	12	14	16	18	20	22	25	27	29	31	34	36	38	40
	.10	2	5	7	9	12	14	17	19	22	24	27	29	32	34	37	39	42	44	47
	.001	0	0	0	1	2	3	5	6	7	9	10	12	13	15	16	18	19	21	22
	.005	0	0	2	3	5	7	8	10	12	14	16	18	19	21	23	25	27	29	31
8	.01	0	1	3	5	7	8	10	12	14	16	18	21	23	25	27	29	31	33	35
	.025	1	3	5	7	9	11	14	16	18	20	23	25	27	30	32	35	37	39	42
	.05	2	4	6	9	11	14	16	19	21	24	27	29	32	34	37	40	42	45	48
	.10	3	6	8	11	14	17	20	23	25	28	31	34	37	40	43	46	49	52	55

Table B.9. (CONTINUED)

n	p	$m = 2$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	.001	0	0	0	2	3	4	6	8	9	11	13	15	16	18	20	22	24	26	27
	.005	0	1	2	4	6	8	10	12	14	17	19	21	23	25	28	30	32	34	37
9	.01	0	2	4	6	8	10	12	15	17	19	22	24	27	29	32	34	37	39	41
	.025	1	3	5	8	11	13	16	18	21	24	27	29	32	35	38	40	43	46	49
	.05	2	5	7	10	13	16	19	22	25	28	31	34	37	40	43	46	49	52	55
	.10	3	6	10	13	16	19	23	26	29	32	36	39	42	46	49	53	56	59	63
	.001	0	0	1	2	4	6	7	9	11	13	15	18	20	22	24	26	28	30	33
	.005	0	1	3	5	7	10	12	14	17	19	22	25	27	30	32	35	38	40	43
10	.01	0	2	4	7	9	12	14	17	20	23	25	28	31	34	37	39	42	45	48
	.025	1	4	6	9	12	15	18	21	24	27	30	34	37	40	43	46	49	53	56
	.05	2	5	8	12	15	18	21	25	28	32	35	38	42	45	49	52	56	59	63
	.10	4	7	11	14	18	22	25	29	33	37	40	44	48	52	55	59	63	67	71
	.001	0	0	1	3	5	7	9	11	13	16	18	21	23	25	28	30	33	35	38
	.005	0	1	3	6	8	11	14	17	19	22	25	28	31	34	37	40	43	46	49
11	.01	0	2	5	8	10	13	16	19	23	26	29	32	35	38	42	45	48	51	54
	.025	1	4	7	10	14	17	20	24	27	31	34	38	41	45	48	52	56	59	63
	.05	2	6	9	13	17	20	24	28	32	35	39	43	47	51	55	58	62	66	70
	.10	4	8	12	16	20	24	28	32	37	41	45	49	53	58	62	66	70	74	79
	.001	0	0	1	3	5	8	10	13	15	18	21	24	26	29	32	35	38	41	43
	.005	0	2	4	7	10	13	16	19	22	25	28	32	35	38	42	45	48	52	55
12	.01	0	3	6	9	12	15	18	22	25	29	32	36	39	43	47	50	54	57	61
	.025	2	5	8	12	15	19	23	27	30	34	38	42	46	50	54	58	62	66	70
	.05	3	6	10	14	18	22	27	31	35	39	43	48	52	56	61	65	69	73	78
	.10	5	9	13	18	22	27	31	36	40	45	50	54	59	64	68	73	78	82	87

Table B.9. (Continued)

n	p	$m = 2$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	.001	0	0	2	4	6	9	12	15	18	21	24	27	30	33	36	39	43	46	49
	.005	0	2	4	8	11	14	18	21	25	28	32	35	39	43	46	50	54	58	61
13	.01	1	3	6	10	13	17	21	24	28	32	36	40	44	48	52	56	60	64	68
	.025	2	5	9	13	17	21	25	29	34	38	42	46	51	55	60	64	68	73	77
	.05	3	7	11	16	20	25	29	34	38	43	48	52	57	62	66	71	76	81	85
	.10	5	10	14	19	24	29	34	39	44	49	54	59	64	69	75	80	85	90	95
	.001	0	0	2	4	7	10	13	16	20	23	26	30	33	37	40	44	47	51	55
	.005	0	2	5	8	12	16	19	23	27	31	35	39	43	47	51	55	59	64	68
14	.01	1	3	7	11	14	18	23	27	31	35	39	44	48	52	57	61	66	70	74
	.025	2	6	10	14	18	23	27	32	37	41	46	51	56	60	65	70	75	79	84
	.05	4	8	12	17	22	27	32	37	42	47	52	57	62	67	72	78	83	88	93
	.10	5	11	16	21	26	32	37	42	48	53	59	64	70	75	81	86	92	98	103
	.001	0	0	2	5	8	11	15	18	22	25	29	33	37	41	44	48	52	56	60
	.005	0	3	6	9	13	17	21	25	30	34	38	43	47	52	56	61	65	70	74
15	.01	1	4	8	12	16	20	25	29	34	38	43	48	52	57	62	67	71	76	81
	.025	2	6	11	15	20	25	30	35	40	45	50	55	60	65	71	76	81	86	91
	.05	4	8	13	19	24	29	34	40	45	51	56	62	67	73	78	84	89	95	101
	.10	6	11	17	23	28	34	40	46	52	58	64	69	75	81	87	93	99	105	111
	.001	0	0	3	6	9	12	16	20	24	28	32	36	40	44	49	53	57	61	66
	.005	0	3	6	10	14	19	23	28	32	37	42	46	51	56	61	66	71	75	80
16	.01	1	4	8	13	17	22	27	32	37	42	47	52	57	62	67	72	77	83	88
	.025	2	7	12	16	22	27	32	38	43	48	54	60	65	71	76	82	87	93	99
	.05	4	9	15	20	26	31	37	43	49	55	61	66	72	78	84	90	96	102	108
	.10	6	12	18	24	30	37	43	49	55	62	68	75	81	87	94	100	107	113	120

Table B.9. (CONTINUED)

n	p	$m = 2$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	.001	0	1	3	6	10	14	18	22	26	30	35	39	44	48	53	58	62	67	71
	.005	0	3	7	11	16	20	25	30	35	40	45	50	55	61	66	71	76	82	87
17	.01	1	5	9	14	19	24	29	34	39	45	50	56	61	67	72	78	83	89	94
	.025	3	7	12	18	23	29	35	40	46	52	58	64	70	76	82	88	94	100	106
	.05	4	10	16	21	27	34	40	46	52	58	65	71	78	84	90	97	103	110	116
	.10	7	13	19	26	32	39	46	53	59	66	73	80	86	93	100	107	114	121	128
	.001	0	1	4	7	11	15	19	24	28	33	38	43	47	52	57	62	67	72	77
	.005	0	3	7	12	17	22	27	32	38	43	48	54	59	65	71	76	82	88	93
18	.01	1	5	10	15	20	25	31	37	42	48	54	60	66	71	77	83	89	95	101
	.025	3	8	13	19	25	31	37	43	49	56	62	68	75	81	87	94	100	107	113
	.05	5	10	17	23	29	36	42	49	56	62	69	76	83	89	96	103	110	117	124
	.10	7	14	21	28	35	42	49	56	63	70	78	85	92	99	107	114	121	129	136
	.001	0	1	4	8	12	16	21	26	30	35	41	46	51	56	61	67	72	78	83
	.005	1	4	8	13	18	23	29	34	40	46	52	58	64	70	75	82	88	94	100
19	.01	2	5	10	16	21	27	33	39	45	51	57	64	70	76	83	89	95	102	108
	.025	3	8	14	20	26	33	39	46	53	59	66	73	79	86	93	100	107	114	120
	.05	5	11	18	24	31	38	45	52	59	66	73	81	88	95	102	110	117	124	131
	.10	8	15	22	29	37	44	52	59	67	74	82	90	98	105	113	121	129	136	144
	.001	0	1	4	8	13	17	22	27	33	38	43	49	55	60	66	71	77	83	89
	.005	1	4	9	14	19	25	31	37	43	49	55	61	68	74	80	87	93	100	106
20	.01	2	6	11	17	23	29	35	41	48	54	61	68	74	81	88	94	101	108	115
	.025	3	9	15	21	28	35	42	49	56	63	70	77	84	91	99	106	113	120	128
	.05	5	12	19	26	33	40	48	55	63	70	78	85	93	101	108	116	124	131	139
	.10	8	16	23	31	39	47	55	63	71	79	87	95	103	111	120	128	136	144	152

For n or m greater than 20, the pth quantile w_p of the Mann-Whitney test statistic may be approximated by

$$w_p = \frac{nm}{2} + x_p \frac{\sqrt{nm(n + m + 1)}}{12}$$

where x_p is the pth quantile of a standard normal random variable.

Table B.10. Quantiles of the Hotelling-Pabst Test Statistic[a]

n	$p = .001$	.005	.010	.025	.050	.100	$\frac{1}{3}n(n^2-1)$
4					2	2	20
5			2	2	4	6	40
6		2	4	6	8	14	70
7	2	6	8	14	18	26	112
8	6	12	16	24	32	42	168
9	12	22	28	38	50	64	240
10	22	36	44	60	74	92	330
11	36	56	66	86	104	128	440
12	52	78	94	120	144	172	572
13	76	110	130	162	190	226	728
14	106	148	172	212	246	290	910
15	142	194	224	270	312	364	1120
16	186	250	284	340	390	450	1360
17	238	314	356	420	480	550	1632
18	300	390	438	512	582	664	1938
19	372	476	532	618	696	790	2280
20	454	574	638	738	826	934	2660
21	546	686	758	870	972	1092	3080
22	652	810	892	1020	1134	1270	3542
23	772	950	1042	1184	1312	1464	4048
24	904	1104	1208	1366	1510	1678	4600
25	1050	1274	1390	1566	1726	1912	5200
26	1212	1462	1590	1786	1960	2168	5850
27	1390	1666	1808	2024	2216	2444	6552
28	1586	1890	2046	2284	2494	2744	7308
29	1800	2134	2306	2564	2796	3068	8120
30	2032	2398	2584	2868	3120	3416	8990

For n greater than 30, the quantiles of T may be approximated by

$$w_p \simeq \tfrac{1}{3}n(n^2-1) + x_p \cdot \frac{1}{6}\sqrt{\frac{n(n^2-1)}{n-1}}$$

where x_p is the pth quantile of a standard normal random variable

SOURCE. Conover (1971)

Table B.11. Estimated Correction Factor for the Mann Whitney Test.

N = 30

		$\alpha = .01$	$\alpha = .05$	$\alpha = .1$
	.1	.903	.038	.980
	.2	.715	.876	.822
	.3	.688	.769	.790
ρ	.4	.516	.609	.716
	.5	.474	.556	.632
	.6	.386	.459	.519
	.7	.283	.336	.405

N = 50

		$\alpha = .01$	$\alpha = .05$	$\alpha = .1$
	.1	.895	.910	.928
	.2	.781	.854	.857
	.3	.714	.787	.828
ρ	.4	.603	.706	.769
	.5	.478	.633	.694
	.6	.383	.510	.580
	.7	.192	.371	.465

Table B.11. Continued

N = 100

		$\alpha = .01$	$\alpha = .05$	$\alpha = .1$
	.1	.885	.945	.963
	.2	.851	.897	.923
	.3	.815	.863	.878
ρ	.4	.725	.743	.830
	.5	.612	.653	.756
	.6	.510	.597	.688
	.7	.374	.500	.584

Table B.12. Estimated Correction Factor for the Spearman Rho Test.

N = 30

		$\alpha = .01$	$\alpha = .05$	$\alpha = .1$
	.1	.923	.930	.939
	.2	.786	.854	.875
	.3	.785	.806	.830
ρ	.4	.682	.729	.762
	.5	.548	.639	.686
	.6	.463	.539	.588
	.7	.350	.418	.476

N = 50

		$\alpha = .01$	$\alpha = .05$	$\alpha = .1$
	.1	.925	.929	.939
	.2	.869	.895	.911
	.3	.777	.786	.857
ρ	.4	.698	.798	.814
	.5	.629	.711	.756
	.6	.546	.636	.676
	.7	.389	.507	.579

Table B.12. Continued

N = 100

		$\alpha = .01$	$\alpha = .05$	$\alpha = .1$
	.1	.944	.954	.970
	.2	.954	.941	.942
	.3	.874	.905	.923
ρ	.4	.817	.866	.887
	.5	.768	.815	.850
	.6	.686	.756	.786
	.7	.593	.669	.716