



The contents of this report are not to be used for advertising, publication, or promotional purposes. Citation of trade names does not constitute an official endorsement or approval of the use of such commercial products.

The findings of this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

The D-series of reports includes publications of the  
Environmental Effects of Dredging Programs:

Dredging Operations Technical Support

Long-Term Effects of Dredging Operations

Interagency Field Verification of Methodologies for  
Evaluating Dredged Material Disposal Alternatives  
(Field Verification Program)



PRINTED ON RECYCLED PAPER

USACEWES



*11/17/96*  
**Dredging Operations Technical  
Support Program**

*11/17  
W39m  
11-96*  
Miscellaneous Paper D-96-2  
July 1996

# **Applications Guide for Statistical Analyses in Dredged Sediment Evaluations**

by Joan U. Clarke, Dennis L. Brandon

U.S. Army Corps of Engineers  
Waterways Experiment Station  
3909 Halls Ferry Road  
Vicksburg, MS 39180-6199

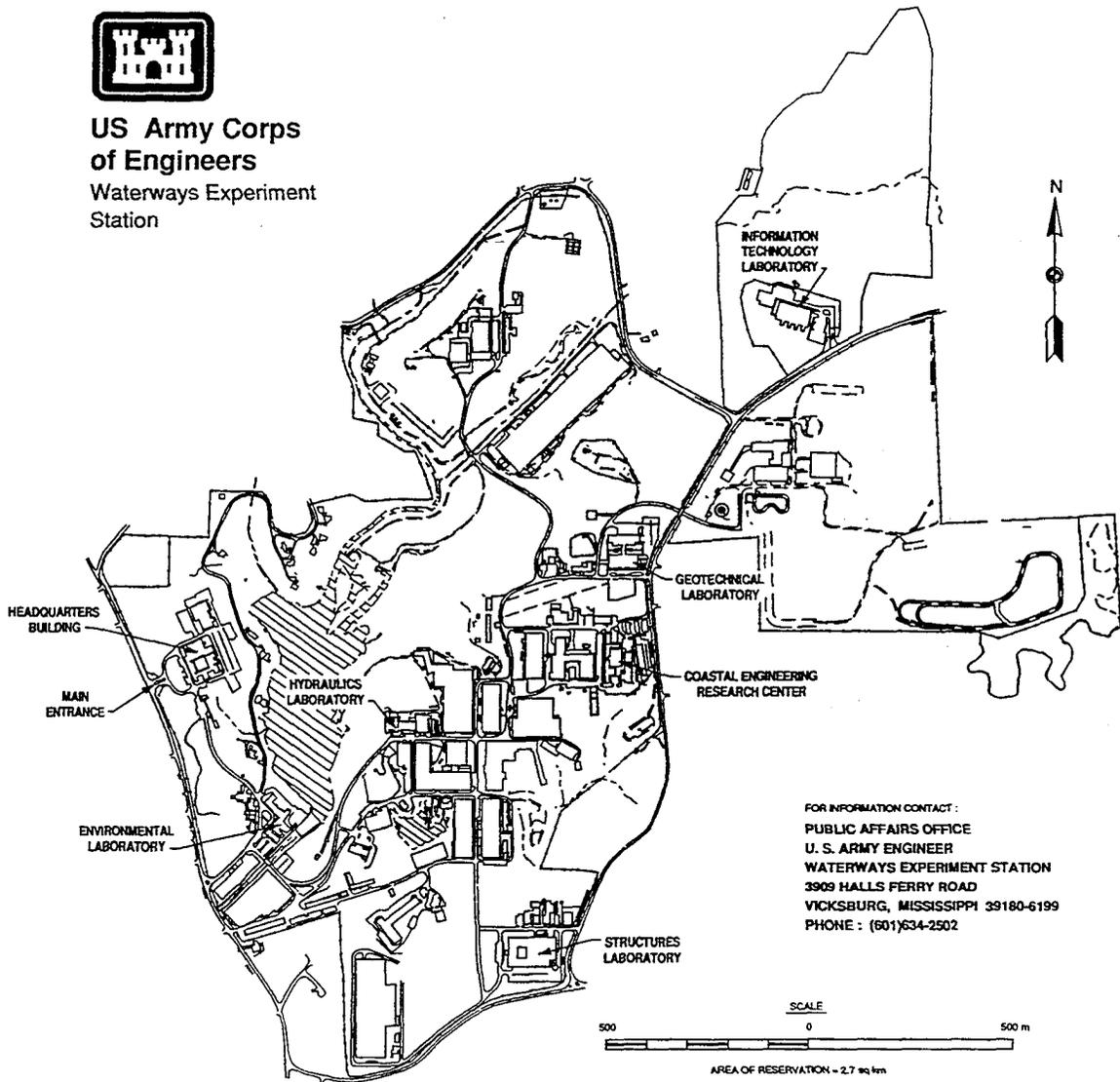
Final report

Approved for public release; distribution is unlimited

Prepared for U.S. Army Corps of Engineers  
Washington, DC 20314-1000



**US Army Corps  
of Engineers**  
Waterways Experiment  
Station



### Waterways Experiment Station Cataloging-in-Publication Data

Clarke, Joan U.

Applications guide for statistical analyses in dredged sediment evaluations / by Joan U. Clarke, Dennis L. Brandon ; prepared for U.S. Army Corps of Engineers.

241 p. : ill. ; 28 cm. -- (Miscellaneous paper ; D-96-2)

Includes bibliographic references.

1. Marine sediments -- Analysis -- Statistics. 2. Dredging (Biology) 3. Dredging spoil -- Analysis -- Statistics. I. Brandon, Dennis L. II. United States. Army. Corps of Engineers. III. U.S. Army Engineer Waterways Experiment Station. IV. Environmental Laboratory (U.S. Army Engineer Waterways Experiment Station) V. Dredging Operations Technical Support Program (U.S. Army Engineer Waterways Experiment Station) VI. Title. VII. Series: Miscellaneous paper (U.S. Army Engineer Waterways Experiment Station) ; D-96-2.

TA7 W34m no.D-96-2



US Army Corps  
of Engineers  
Waterways Experiment  
Station

# Environmental Effects of Dredging Programs



## Dredging Operations Technical Support Report Summary

---

### *Applications Guide for Statistical Analyses in Dredged Sediment Evaluations (MP D-96-2)*

**ISSUE:** Dredged sediment evaluations often require statistical analysis of chemical or biological test results. However, the resulting data are frequently problematic for standard statistical procedures because of improper experimental design, insufficient replication, failure to meet statistical test assumptions, outliers, and missing or below detection limit observations. Such nonideal data can seriously affect the error rates of statistical tests. This in turn increases the likelihood of drawing false inferences concerning the potential of a dredged sediment for adverse biological effects.

**INVESTIGATIONS:** Simulations were conducted to investigate the impact of nonideal data on the performance of statistical tests recommended for dredged sediment evaluations. Statistical test error rates were assessed using data that violated the normality and equality of variances assumptions, as well as data that included outliers or nondetects.

**SUMMARY:** This report includes a brief introduction to statistical aspects of sediment sampling,

some basic experimental designs and problems that can arise, errors in statistical testing and the importance of power, testing the normality and equality of variances assumptions and implications of violations, the effect of outliers, methods for analyzing less-than detection limit data, and interpreting statistical test results. Program statements are provided for recommended statistical testing procedures using some popular statistical software packages. This report is intended as a companion to the statistics appendix of the Inland Testing Manual.

**AVAILABILITY OF REPORT:** The report is available on Interlibrary Loan Service from the U.S. Army Engineer Waterways Experiment Station (WES) Library, 3909 Halls Ferry Road, Vicksburg, MS 39180-6199; telephone (601) 634-2355.

To purchase a copy, call the National Technical Information Service (NTIS) at (703) 487-4780. For help in identifying a title for sale, call (703) 487-4780. NTIS report numbers may also be requested from the WES librarians.

**About the Authors:** Ms. Joan U. Clarke and Mr. Dennis L. Brandon are statisticians in the WES Environmental Laboratory. For further information about the Dredging Operations Technical Support Program, contact Mr. Thomas R. Patin, Program Manager, at (601) 634-3444.

# Contents

---

Preface . . . . .	x
1—Introduction . . . . .	1
2—Sediment Sampling . . . . .	3
3—Experimental Design . . . . .	5
Completely Randomized Design . . . . .	5
Randomized Blocks Design . . . . .	6
Unequal Replication, Inadequate Replication, No Replication . . . . .	7
Lack of Randomization . . . . .	8
Inadequate Interspersion . . . . .	9
Control Treatments . . . . .	9
Number of Experimental Organisms and Exposure Duration . . . . .	10
4—Statistical Errors, $\alpha$ , and Power . . . . .	11
Type I Error and Confidence . . . . .	11
The importance of $\alpha$ . . . . .	12
Effect of violations of assumptions on $\alpha$ . . . . .	14
Transformations and $\alpha$ . . . . .	14
Type II Error and Power . . . . .	17
Why power is important . . . . .	17
Sample size, effect size, $\alpha$ , and power . . . . .	17
Why $\alpha$ should not be restricted to 0.05 . . . . .	19
Increasing power by increasing the number of reference replicates . . . . .	20
Effect of frequency distribution and unequal variances on power . . . . .	22
Effect of transformation on power . . . . .	23
Calculating power . . . . .	24
Summary . . . . .	25
5—Tests of Assumptions: Normality . . . . .	27
Survival Data . . . . .	27
Testing for normality . . . . .	28
Example data . . . . .	29
Chemical Concentration Data in Environmental Samples . . . . .	31
Performance of Shapiro-Wilk's Test . . . . .	33
Consequences of failure to correctly identify sample distribution . . . . .	35

When not to assume normality . . . . .	39
Example data . . . . .	41
Summary . . . . .	44
6—Tests of Assumptions: Equality of Variances . . . . .	46
Variances and Power . . . . .	47
Equality of Variance Tests on Example Data . . . . .	48
Levene’s Test . . . . .	48
Bartlett’s Test . . . . .	49
Hartley’s $F_{\max}$ Test . . . . .	54
Cochran’s Test . . . . .	55
Performance of Levene’s Test . . . . .	55
Summary . . . . .	58
7—Outliers . . . . .	60
Effect of Outliers on Statistical Tests . . . . .	61
Effect on tests of assumptions . . . . .	61
Effect on <i>t</i> -tests . . . . .	63
Example data . . . . .	64
Detecting Outliers . . . . .	68
What To Do When Outliers Occur . . . . .	70
Summary . . . . .	71
8—Less-Than Detection Limit (Censored) Data . . . . .	73
Methods for Censored Data . . . . .	73
Applying Censored Data Methods . . . . .	75
One nondetect in one or more treatments . . . . .	75
Several nondetects in one or more treatments . . . . .	76
More than 50-percent censoring . . . . .	77
Severe censoring . . . . .	79
Complete censoring of a dredged sediment treatment . . . . .	79
Reference sediment data completely censored . . . . .	81
Complex problems . . . . .	83
Multiple Detection Limits . . . . .	86
Moderately censored data . . . . .	87
Highly censored data . . . . .	88
Nondetects and Estimation . . . . .	89
Evaluating censored data methods for estimation . . . . .	90
Comparisons with an action level . . . . .	94
Summary . . . . .	96
9—Interpreting Statistical Test Results . . . . .	99
Interpreting P-Values . . . . .	99
When Different Tests on the Same Data Have Different Results . . . . .	101
When LSD Is Significant and ANOVA Is Not or Vice Versa . . . . .	102
Interpreting LSD Output . . . . .	102
When Transformation Changes the Order of Treatment Means . . . . .	104
When Data for an Action Level Comparison Fail the Normality Test . . . . .	106

Analyzing a Blocked Design . . . . .	107
Summary . . . . .	111
References . . . . .	113
Appendix A: SAS Programs . . . . .	A1
Appendix B: SYSTAT Programs . . . . .	B1
Appendix C: SPSS Programs . . . . .	C1
SF 298	

## List of Figures

---

Figure 1. Completely randomized design . . . . .	5
Figure 2. Randomized complete blocks design . . . . .	6
Figure 3. Randomized incomplete blocks design . . . . .	6
Figure 4. Randomized design illustrating unequal, inadequate, and no replication . . . . .	7
Figure 5. Systematic design . . . . .	8
Figure 6. Randomized design with inadequate interspersion . . . . .	9
Figure 7. Mean $\alpha$ for LSD test using untransformed samples from normal and nonnormal distributions. Equal sample sizes . . .	14
Figure 8. Mean $\alpha$ for LSD test using untransformed samples from normal and nonnormal distributions. Unequal sample sizes . .	15
Figure 9. Mean $\alpha$ for LSD test using log-transformed and untrans- formed samples from lognormal distributions . . . . .	16
Figure 10. Mean $\alpha$ for LSD test using rankit-transformed and untransformed samples from gamma distributions . . . . .	16
Figure 11. Power of a <i>t</i> -test for effect size of one standard deviation . . . . .	19
Figure 12. Percent increase in power when one sample from each dredged sediment is reallocated to the reference sediment . . .	21
Figure 13. Change in power of LSD test using untransformed samples from lognormal and gamma distributions compared with untransformed samples from normal distribution . . . . .	22
Figure 14. Change in power of LSD test using transformed lognormal samples compared with untransformed lognormal samples . . .	23
Figure 15. Change in power of LSD test using transformed gamma samples compared with untransformed gamma samples . . . .	24

Figure 16.	Results of Shapiro-Wilk's Test on residuals of untransformed and arcsine-transformed simulated binomial data, balanced design, where $p$ is survival proportion . . . . .	28
Figure 17.	Results of Shapiro-Wilk's Test on residuals of untransformed and arcsine-transformed simulated binomial data, unbalanced design, where $p$ is survival proportion . . . . .	29
Figure 18.	Frequency histograms for example bioassay survival data . . . . .	32
Figure 19.	Power of Shapiro-Wilk's Test to detect nonnormality: lognormal distribution, balanced design . . . . .	33
Figure 20.	Power of Shapiro-Wilk's Test to detect nonnormality: lognormal distribution, unbalanced design . . . . .	34
Figure 21.	Power of Shapiro-Wilk's Test to detect nonnormality: gamma distribution, balanced design . . . . .	34
Figure 22.	Power of Shapiro-Wilk's Test to detect nonnormality: gamma distribution, unbalanced design . . . . .	35
Figure 23.	Change in power and $\alpha$ of LSD test resulting from incorrect identification of parent distribution and application of wrong transformation . . . . .	36
Figure 24.	Change in power and $\alpha$ of LSD test resulting from failure to identify normal distribution and application of wrong transformation . . . . .	37
Figure 25.	Change in power and $\alpha$ of LSD test resulting from use of untransformed lognormal data . . . . .	37
Figure 26.	Change in power and $\alpha$ of LSD test resulting from use of untransformed data from gamma distribution . . . . .	38
Figure 27.	Approximate percent negative values in a normal distribution . . . . .	40
Figure 28.	Change in power and $\alpha$ of LSD test resulting from use of rankits, rather than untransformed data, for samples from a normal distribution . . . . .	40
Figure 29.	Change in power and $\alpha$ of LSD test resulting from use of rankits, rather than log-transformed data, for samples from a lognormal distribution . . . . .	41
Figure 30.	Power of a $t$ -test to detect a difference in means of one unit when $\alpha = 0.05$ and $n = 5$ . . . . .	47
Figure 31.	Mean increase in Type I error rate above the nominal $\alpha$ for Levene's Test using simulated data . . . . .	56
Figure 32.	Power of Levene's Test when sample sizes are equal and one of four variances differs from the other three . . . . .	57

Figure 33.	Power of Levene's Test when sample sizes are unequal and the variance of the largest sample is greater than the variances of the other three samples . . . . .	58
Figure 34.	Type I error rate of Shapiro-Wilk's Test for balanced and unbalanced designs when normal samples include one outlier ranging from 0 to 20 standard deviations from the mean . . . . .	62
Figure 35.	Type I error rate of Levene's Test for balanced design when normal samples include one outlier ranging from 0 to 20 standard deviations from the mean . . . . .	62
Figure 36.	Power of a 1-tailed <i>t</i> -test for unequal variances when normal samples include one outlier ranging from 0 to 20 standard deviations greater than the mean . . . . .	64
Figure 37.	Power of a 1-tailed <i>t</i> -test for unequal variances when normal samples include one outlier in reference sediment ranging from 0 to 20 standard deviations greater than the mean . . . . .	71
Figure 38.	Accuracy of six censored data methods for estimating population mean and standard deviation when $CV \leq 0.25$ . . . . .	92
Figure 39.	Accuracy of six censored data methods for estimating population mean and standard deviation when $CV > 0.25$ and $\leq 0.5$ . . . . .	92
Figure 40.	Accuracy of six censored data methods for estimating population mean and standard deviation when $CV > 0.5$ and $\leq 1$ . . . . .	93
Figure 41.	Accuracy of six censored data methods for estimating population mean and standard deviation when $CV > 1$ . . . . .	93
Figure 42.	Type I error rate for simulated normal, lognormal, and gamma samples in comparisons with an action level . . . . .	107

## List of Tables

---

Table 1.	Possible Consequences of Hypothesis Test Outcomes in Dredged Sediment Evaluations . . . . .	12
Table 2.	Coefficients of Variation for Simulations Comparing Three Dredged Sediments With a Reference Sediment . . . . .	13
Table 3.	Power of a <i>t</i> -test for Various Relative Effect Sizes When $\alpha = 0.05$ . . . . .	18
Table 4.	Balanced Types I and II Error Rates for a <i>t</i> -test . . . . .	20
Table 5.	Survival of Animals Exposed to Great Lakes Sediments . . . . .	30

Table 6.	Results of Shapiro-Wilk's Test for Normality on Example Survival Data . . . . .	31
Table 7.	Bioaccumulation of Selected Metals in Organisms Exposed to New York Bight Sediments . . . . .	42
Table 8.	Results of Shapiro-Wilk's Test for Normality on Example Bioaccumulation Data . . . . .	43
Table 9.	Results of Levene's Test on Example Data . . . . .	49
Table 10.	Results of Bartlett's Test on Example Data . . . . .	50
Table 11.	Results of Hartley's $F_{\max}$ Test on Example Data . . . . .	51
Table 12.	Results of Cochran's Test on Example Data . . . . .	52
Table 13.	Summary of Equality of Variance and Nonnormality Test Significant Results for Example Data . . . . .	53
Table 14.	Steps in Selecting the Most Appropriate Censored Data Method: One Nondetect in One or More Treatments . . . . .	76
Table 15.	Steps in Selecting the Most Appropriate Censored Data Method: More Than One Nondetect in Any Treatment . . . . .	77
Table 16.	Steps in Selecting the Most Appropriate Censored Data Method: Highly Censored Data . . . . .	78
Table 17.	Steps in Selecting the Most Appropriate Censored Data Method: Complete Censoring of a Dredged Sediment Treatment . . . . .	80
Table 18.	Water Quality Data for Zinc . . . . .	84
Table 19.	Preliminary Analysis of Zinc Water Quality Data for Downstream-Upstream Comparison . . . . .	85
Table 20.	Preliminary Analysis of Zinc Water Quality Data for CDF-Harbor Comparisons . . . . .	86
Table 21.	Steps in Selecting the Most Appropriate Censored Data Method: Moderately Censored Data with Multiple Detection Limits . . . . .	87
Table 22.	Steps in Selecting the Most Appropriate Censored Data Method: Moderately Censored Data with Multiple Detection Limits . . . . .	88
Table 23.	Steps in Selecting the Most Appropriate Censored Data Method: Highly Censored Data with Multiple Detection Limits . . . . .	89
Table 24.	Recommended Censored Data Methods for Small Samples To Be Used in Estimation . . . . .	91
Table 25.	Preliminary Analysis for Comparison of Censored Total PCB Bioaccumulation Samples with FDA Action Level . . . . .	95

Table 26.	Comparison of Total PCB Bioaccumulation with FDA Action Level of 2 $\mu\text{g/g}$ Following Application of Censored Data Methods . . . . .	96
Table 27.	Summary of SAS Programs for Preliminary Analysis of Censored Data . . . . .	98

# Preface

---

This report discusses the implications and techniques of statistical data analysis in dredged sediment evaluations and is intended to supplement the Statistics Appendix of the Inland Testing Manual (U.S. Environmental Protection Agency/U.S. Army Corps of Engineers 1994, Appendix D). Emphasis is on types of nonideal data likely to be encountered in real-world testing situations.

Financial support for this work was provided by the Dredging Operations Technical Support (DOTS) Program, managed under the Environmental Effects of Dredging Programs (EEDP) at the U.S. Army Engineer Waterways Experiment Station (WES). Mr. Thomas R. Patin was DOTS Program Manager, and Dr. Robert M. Engler was EEDP Manager. DOTS Technical Monitor was Mr. Dave Mathis. Work was performed by Ms. Joan U. Clarke and Mr. Dennis L. Brandon of the Fate and Effects Branch (FEB), Environmental Processes and Effects Division (EPED), Environmental Laboratory (EL), WES. The work was conducted under the supervision of Dr. Bobby L. Folsom, Jr., Chief, FEB; Mr. Donald L. Robey, Chief, EPED; and Dr. John W. Keeley, Director, EL.

The authors wish to thank Drs. Michael Honeycutt, David Moore, and Thomas Wright, FEB, for technical review. Most of the data included as examples in this report originated from sediment and bioaccumulation testing performed for the U.S. Army Engineer Districts of San Francisco, Chicago, and New York.

At the time of publication of this report, Director of WES was Dr. Robert W. Whalin. Commander was COL Bruce K. Howard, EN.

This report should be cited as follows:

Clarke, J. U., and Brandon, D. L. (1996). "Applications guide for statistical analyses in dredged sediment evaluations," Miscellaneous Paper D-96-2, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.

*The contents of this report are not to be used for advertising, publication, or promotional purposes. Citation of trade names does not constitute an official endorsement or approval of the use of such commercial products.*

# 1 Introduction

---

Evaluation of sediments for dredging and disposal options often requires biological testing when sediment contamination is known or suspected. Biological testing, conducted under a tiered testing framework, is described in U.S. Environmental Protection Agency/U.S. Army Corps of Engineers (USEPA/USACE) (1991) (the Ocean Testing Manual or “Green Book”) and USEPA/USACE (1994) (the Inland Testing Manual). Required tests, usually performed in Tier III, may include water column toxicity, benthic toxicity, benthic bioaccumulation, and steady-state bioaccumulation. Statistical procedures for the analysis of data resulting from these tests are fully described in Appendix D of the Inland Testing Manual. However, inference from these procedures is predicated upon certain assumptions, and data resulting from biological tests may not meet those assumptions for a variety of reasons. Such “nonideal” data occur frequently. The implications of nonideal data for the statistical procedures of the Inland Testing Manual are explored herein. As such, this document complements and should be used along with Appendix D of the Inland Testing Manual.

Topics covered in this document include the following:

- A brief introduction to statistical aspects of sediment sampling (Chapter 2).
- Some basic experimental designs and problems that can arise (Chapter 3).
- Errors in statistical testing and the importance of power (Chapter 4).
- Testing the normality assumption and implications of violations (Chapter 5).
- Testing the equality of variances assumption and implications of violations (Chapter 6).
- The effect of outliers (Chapter 7).

- Methods for analyzing less-than detection limit data (Chapter 8).
- Interpreting statistical test results (Chapter 9).

Topics are illustrated whenever possible using nonideal data from actual dredging projects. Some of the information provided herein arises from statistical simulation work conducted at the U.S. Army Engineer Waterways Experiment Station (WES) and, as such, will not be found in any statistics textbook.

Appendix D of the Inland Testing Manual provides SAS programs for recommended statistical testing procedures in the routine analysis of Tier III toxicity and bioaccumulation data. SYSTAT and SPSS programs for those same procedures are included as Appendixes B and C to this document. SAS programs for a variety of additional analyses discussed herein are provided as Appendix A to this document. These include calculations of power and least significant difference, tests for equality of variances, preliminary analysis of censored data, and analysis of a blocked design.

## 2 Sediment Sampling

---

Collection of sediment samples is the first stage in Tiers II and III testing for dredged sediment evaluations. While a thorough characterization of the dredging, disposal, or reference sites is seldom necessary, it is important to obtain samples that reasonably represent these areas. Carefully constructed plans, based on statistical sampling principles, should be developed before sampling. Section 8.0 of the Inland Testing Manual (USEPA/USACE 1994) describes aspects of sediment sampling, and Section 8.2 lists steps essential to a sampling plan. These steps include subdividing the dredging area into project areas or management units, determining the number of samples to be taken, and selecting sampling locations. Designing a sediment sampling plan is case specific. However, there are general methodologies that should be considered when attempting to achieve the steps listed above. These methodologies are only described briefly herein; an in-depth discussion would require a separate manual.

Subdividing the dredging area into subunits may partition a heterogeneous area into several homogeneous units. The rationale for establishing these subunits vary (e.g., differences in grain size, hydrology, or contaminant concentrations; historical information; and desire to characterize the proposed dredging area both vertically and horizontally). The statistical term used to describe such partitioning is stratification. Even if the dredging area is not partitioned, the entire area may be considered one stratum. Once the area is stratified, several sampling methods may be used, such as random, stratified random, and systematic sampling. Cochran (1977) provides a wealth of information on various sampling methods. However, for adequate coverage of the sediment sampling area, any of the methods may result in a sampling plan that is quite expensive to implement. Pennington et al. (1990) list techniques for reducing the cost of testing dredged material.

Random sampling methods have the most utility in sediment sampling plans where the objective is to make inferences about the entire dredging area, especially when the sediment is fairly homogeneous. Stratified random sampling, which allots a certain number of random samples to be taken in each stratum, should provide more accurate representation of physically stratified sediments. There are numerous ways to determine the number of samples for each stratum. One method is based on proportionality. For instance, if a stratum comprises 25 percent of the dredging area, then 25 percent of the

samples should come from this stratum. The number of samples taken in a stratum could also be inversely proportional to the amount of information known about the stratum. Statistical methods can be used to determine the number of samples. Provost (1984:84-88), for example, describes a method of determining number of samples based on a normal distribution. If the data are not normally distributed, this method provides an approximation that improves as sample size increases. Cochran (1977:96-99) describes methods to optimize the allocation of samples to a stratum while minimizing the cost.

Systematic sampling, for example, taking samples along a transect or in each square of a grid, works well in areas where gradients occur in important sediment characteristics. Gradients in sediment grain size commonly occur from fine-grained material in depositional zones of inner harbors to coarser grained sediments at harbor entrance channels. Contaminant gradients are likely around point sources or where spills have occurred. Gilbert (1987) describes methods for systematic sampling using square, rectangular, or triangular grids. The investigator can select a grid size and calculate the number of samples required to obtain a certain degree of coverage. Alternatively, the investigator can select the number of samples and calculate the grid size and degree of coverage.

Once the number of samples is determined, randomized methods can be used to identify sampling locations. The investigator may choose to composite samples from several locations or from several strata for chemical analysis and biological testing. Compositing is especially appropriate for dredged sediment evaluations, because the dredging and disposal processes are likely to mix sediment strata, in effect compositing large volumes of dredged sediment before final placement. The investigator may also consider a multistage sampling plan. For instance, a physical characterization and chemical analysis of individual sediment cores could be conducted in the first stage. That information could be used to determine a compositing scheme for biological testing in the second stage. Each composited sample is then thoroughly mixed and subsampled for assignment to laboratory test chambers (= experimental units or replicates) using an appropriate experimental design (Chapter 3). Statistical analysis procedures for biological testing in the Inland Testing Manual generally presuppose sediment compositing and the use of laboratory replicates based on subsamples of a sediment composite, rather than field replicates based on individual sediment samples.

# 3 Experimental Design

---

Statistical analysis begins with experimental design. Proper experimental design is critical to ensuring the statistical validity of experiments, for reasons discussed in Section D1.3 of the Inland Testing Manual. The two recommended experimental designs, completely randomized design and randomized blocks design, will be illustrated in this chapter, along with examples of inadequate designs.

## Completely Randomized Design

Treatments may be allocated to all experimental units in a completely randomized manner. One of the simplest ways to do this is to assign a number to each experimental unit, and then use a random numbers table to randomly assign treatments to the experimental units. Figure 1 shows a schematic representation of the Flow-Through Aquatic Toxicology Exposure System (FATES) in use at WES for laboratory experiments such as Tier III bioaccumulation testing. FATES consists of 24 circular aquaria arranged in a double row on two platforms of 12 aquaria each. In Figure 1, three dredged sediment treatments (A, B, C) and one reference sediment treatment (R) have been assigned to six aquaria each. Assignments have been made randomly over the entire experimental setup; i.e., there are no restrictions on randomization. This type of design is the simplest to analyze and works well when exposure conditions can be maintained uniformly over the entire experimental setup.

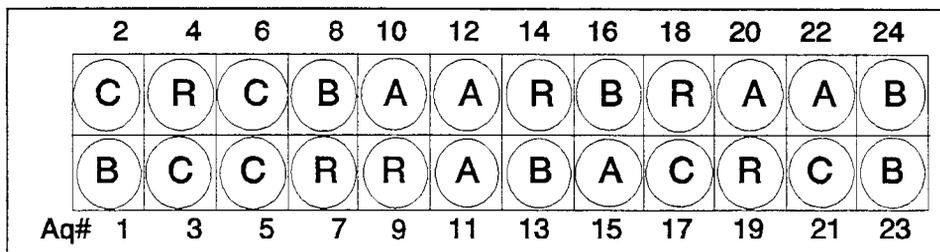


Figure 1. Completely randomized design

## Randomized Blocks Design

Treatments may be assigned randomly within each of a series of blocks of experimental units. A block is a set of adjacent experimental units in which exposure conditions can be considered uniform. In a complete blocks design, each block has an equal number of replicates of each treatment. Figure 2 shows FATES divided into four quadrants of six aquaria each. In each quadrant (block), five dredged sediment treatments (A, B, C, D, E) and the reference sediment (R) have been randomly assigned to one experimental unit each. Because each block is randomized separately and must include all treatments, randomization is said to be restricted. A randomized blocks design is appropriate when groups of experimental units are maintained on separate tables or water baths, for example, and ambient conditions may differ somewhat from one table or water bath to another.

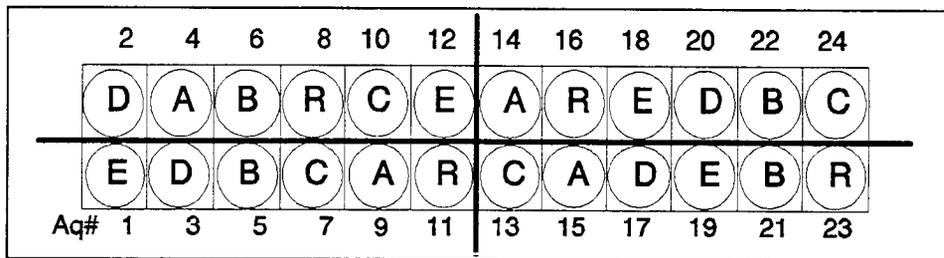


Figure 2. Randomized complete blocks design

Incomplete randomized blocks designs, in which each block has an unequal allocation of treatments to experimental units, may be necessary and even desirable in some circumstances. In Figure 3, four dredged sediment treatments (A, B, C, D) have been randomly assigned to one experimental unit each in each of four blocks, while the reference sediment (R) has been randomly assigned to two experimental units in each block. As discussed in Section D2.2.1 of the Inland Testing Manual and in Chapter 4 of this document, increasing the number of reference replicates and decreasing the dredged sediment replicates can be an effective method of increasing statistical

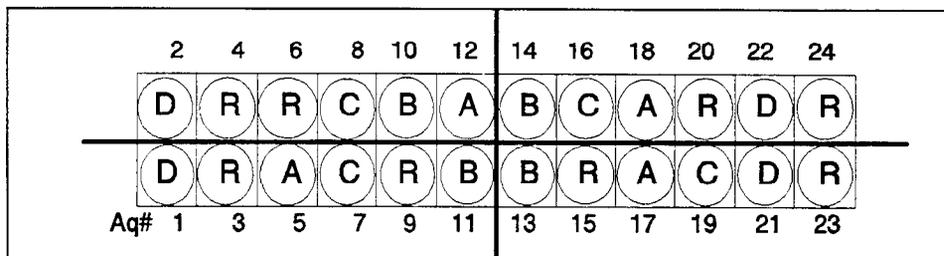


Figure 3. Randomized incomplete blocks design

power for dredged sediment-reference sediment comparisons while maintaining a reasonable overall sample size.

## Unequal Replication, Inadequate Replication, No Replication

Figure 3 showed an example of an acceptable design that incorporated unequal replication. Figure 4 is an undesirable design that includes inadequate replication and no replication as well as unequal replication. In this completely randomized design, four dredged sediment treatments (A, B, C, D) have been assigned to five experimental units each. A fifth dredged sediment (E) has been assigned to only two experimental units, which is inadequate. Worse yet, the reference sediment (R) has been assigned to only a single experimental unit, and thus is unreplicated. A single sand control (S) is also included for chemical quality assurance/quality control (QA/QC), although this treatment will not be part of the statistical data analysis.

	2	4	6	8	10	12	14	16	18	20	22	24
	C	A	B	D	E	A	A	D	B	R	C	S
	A	B	C	D	B	C	D	D	B	A	E	C
Aq#1	3	5	7	9	11	13	15	17	19	21	23	

Figure 4. Randomized design illustrating unequal, inadequate, and no replication

Unequal replication usually will not pose a problem in statistical analysis. The normality test, certain equality of variance tests (e.g., Bartlett's Test, Levene's Test), Least Significant Difference (LSD) test, *t*-tests, and nonparametric tests recommended in Appendix D of the Inland Testing Manual do not require an equal number of treatment replicates. However, some tests, such as Cochran's and Hartley's tests for equality of variance, have not been generalized for use with unequal replication. These two tests have been used in some instances where the number of treatment replicates is nearly equal. Unequal replication alters the degrees of freedom. Winer (1971) provides further details on the use of Cochran's Test and Hartley's Test when treatment replication varies.

Inadequate replication leads to insufficient power in statistical comparisons, i.e., an inability to detect true differences among treatments. The Inland and Ocean Testing Manuals recommend a minimum of five replicates per treatment for toxicity and bioaccumulation testing. However, when treatment variability is large, or when a small difference between dredged sediment and reference sediment end points is biologically significant, five replicates may

not be enough to provide adequate power for statistical comparisons. Statistical power and sample size are discussed more fully in Chapter 4.

When a treatment is unreplicated, there is no estimate of variability in the response to that treatment. If the reference is unreplicated, then the reference response becomes in essence a numerical standard to which the variable dredged sediment responses must be compared. This imposes an unwarranted credibility on the reference response because the single value that measures reference response may be much higher or lower than the true mean reference population response. The same is true for an unreplicated dredged sediment treatment. If all treatments are unreplicated, then statistical comparisons cannot be performed.

If an experimental design has been constructed in which replication is inadequate, or one or more treatments is unreplicated, then the number of experimental units should be increased to accommodate adequate replication for all treatments. If the number of experimental units cannot be increased, for example in a laboratory system such as FATES, it may be necessary to run a sequence of experiments. If 20 aquaria are available, for example, and four dredged sediments need to be compared with a reference sediment, then the first experiment could test three of the dredged sediments with the reference using five replicates per treatment. A second experiment could then be performed testing the remaining dredged sediment and the reference, again using five replicates each. Response data for each experiment should be analyzed separately, as some exposure conditions may have changed from the first experiment to the second.

## Lack of Randomization

Figure 5 illustrates a FATES layout in which treatments have been assigned systematically, rather than randomly. Such a design may fail to control for spatial variability in the experiment. Conditions in a test chamber may in some way influence conditions in neighboring test chambers. Or a chemical spill at one end of the laboratory bench could wipe out an entire

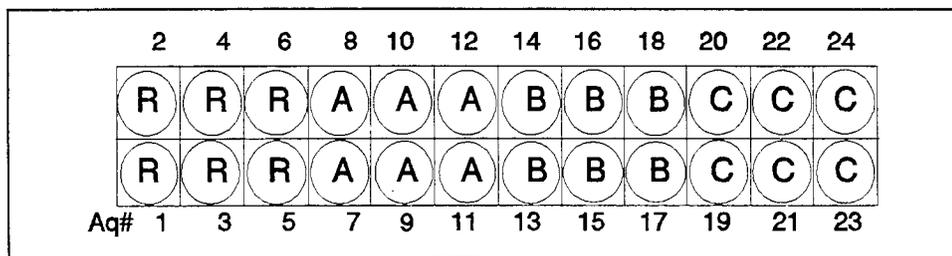


Figure 5. Systematic design

treatment. Systematic designs can increase investigator bias in the assignment of animals to experimental units or in the collection of samples for analysis.

## Inadequate Interspersion

Figure 6 illustrates an experimental design in which a single dredged sediment is compared with a reference using half of the FATES aquaria. Although the six replicates for each treatment were randomly assigned to experimental units, most of the reference replicates are grouped at one end of the bench while most of the dredged sediment replicates are grouped at the other end. Even though the design is random, interspersion of treatments is inadequate and may fail to control effectively for spatial variability. In such a situation, random assignment of treatments to experimental units should be repeated until interspersion is judged to be adequate. Should treatments be systematically interspersed? Generally, they should not because each experimental unit could have a systematic influence on neighboring experimental units.

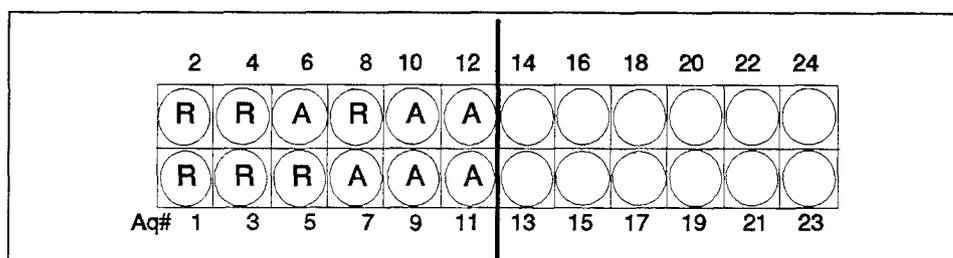


Figure 6. Randomized design with inadequate interspersion

## Control Treatments

In most situations, the reference sediment (or dilution water for water column toxicity testing) will serve as a statistical control to which dredged sediment responses are compared. Therefore, reference (or dilution water) replicates must be included in the experimental design. The number of reference replicates should generally equal the number of dredged sediment replicates. In some cases, it may be desirable to assign more replicates to the reference than to the dredged sediments; refer to Chapter 4 of this document and Section D2.2.1 of the Inland Testing Manual.

Other types of controls such as a clean sand control or clear water control may also be included in the experimental design for QA/QC purposes, for example, to establish the health of the test animals. The test end point is measured to determine whether it falls within acceptable criteria. However, the data for these controls are not included in the statistical analyses that compare treatments.

## Number of Experimental Organisms and Exposure Duration

Assigning an unequal number of experimental animals to treatment containers is undesirable but not unmanageable. This situation requires slight modifications in the toxicity data analysis programs WATTOX.SAS and BENTOX.SAS (Appendix D of USEPA/USACE 1994 or in the corresponding SYSTAT or SPSS programs in Appendixes B and C of this document). Both programs assume the number of organisms per test container is 20. This is denoted in the SAS programs by the statement "M = 20;". One alternative is to add an "IF" statement after "M = 20;" to modify M for specific treatment replicates. For instance, the statement "IF (TRT = 1 AND REP = 2) THEN M = 19;" alters the value of M for treatment 1, replicate 2. A second alternative would be to add M to the input statement. For instance, the statement "INPUT TRT REP SURV @@;" would become "INPUT TRT REP SURV M @@;" and the statement "M = 20;" would be deleted. Of course, the data lines have to be modified to include values for M. Programs WATTOX.SAS and BENTOX.SAS will successfully execute without further modifications.

Samples not collected as described in the experimental design may or may not be a problem during statistical analysis. Toxicity tests are to be terminated at a set time. Altering the termination point of a toxicity test invalidates the test. These toxicity data should be discarded and the test repeated using the appropriate termination point. Tier III bioaccumulation tests are designed for 28 days of exposure. The Tier III exposure length generally should not be altered. Tier IV time-sequenced tests to estimate steady-state bioaccumulation have an exposure period of at least 28 days. Longer exposures may be needed for slowly accumulated contaminants such as dioxin. The time-sequenced tests might include samples at 0, 2, 4, 7, 10, 18, and 28 days. An acceptable alternate time sequence might include samples at 0, 2, 5, 8, 11, 19, 24, and 28 days. BIOACSS.SAS will attempt to estimate steady-state concentrations from any time sequence provided the sampling days are included as data input.

## 4 Statistical Errors, $\alpha$ , and Power

---

Because statistical hypothesis testing is based on probability rather than certainty, it is subject to error. Statistical errors may occur when the samples used in the hypothesis test do not adequately represent the population(s) of interest, when sample size (number of replicates) is insufficient to detect meaningful differences among populations, or when the assumptions of the statistical test are violated. It is impossible to know whether the results of an individual hypothesis test are in error unless the population parameters are known. Such is rarely, if ever, the case for real (as opposed to simulated) populations. However, the rates or probabilities of error can be quantified for ideal data from real populations, and steps can be taken to minimize error rates when data are not ideal. The types of error in statistical hypothesis testing are discussed in Section D1.2 and shown in Table D-1 of the Inland Testing Manual. This chapter examines the probabilities associated with the two types of error and the factors that influence those probabilities. The basic concepts relating to power are explored in detail in Cohen (1988).

### Type I Error and Confidence

When two or more populations do not differ with respect to a given parameter such as the mean, a hypothesis test of samples from those populations is expected to conclude that the populations do not differ. The probability of reaching that conclusion is known as the confidence level of the test and is designated as  $1 - \alpha$ . If the hypothesis test concludes instead that the population parameters differ, then a Type I error has been committed. The probability of Type I error,  $\alpha$ , is also known as the size of the test (Dorfman, Pesti, and Fletcher 1993). For comparisons of treatments,  $\alpha$  is customarily set at 0.05. For ancillary hypothesis tests, such as tests of assumptions,  $\alpha$  may be scaled depending on the number of replicates and on whether the design is balanced or unbalanced (see Table D-2 of the Inland Testing Manual).

## The importance of $\alpha$

Recall that the ultimate purpose of the statistical comparisons in the Inland and Ocean Testing Manuals is to identify the possibility of adverse environmental impact arising from disposal of contaminated dredged material in the aquatic environment. When adverse environmental impact is possible, it may be necessary to consider alternatives to unrestricted aquatic disposal, often at increased cost. A simple scheme of the various consequences of hypothesis testing is presented in Table 1 as an adaptation of Table D-1 from the Inland Testing Manual.

Hypothesis Test Conclusion	True State of Nature	
	Populations Do Not Differ	Populations Differ
Samples do not differ	Correct (probability = $1 - \alpha$ ) No environmental degradation Cost containment	Type II error (probability = $\beta$ ) Possible environmental degradation Cost containment
Samples differ	Type I error (probability = $\alpha$ ) No environmental degradation Possible unnecessary increased cost	Correct (probability = $1 - \beta$ ) No environmental degradation Possible necessary increased cost

Type I error in dredged sediment evaluations means the hypothesis test concludes that the dredged sediment is significantly worse than the reference sediment in terms of organism survival or contaminant bioaccumulation, when in fact the dredged sediment is not worse than the reference. This in turn can lead to erroneous decisions regarding disposal that could result in increased cost compared with unrestricted aquatic disposal. In using the typical significance level of  $\alpha = 0.05$ , the investigator is willing to accept 1 chance in 20 of making this type of error. However, certain characteristics of the data, notably departures from normality and equality of variances, can cause the actual Type I error rate of a statistical test to differ from the nominal  $\alpha$  of 0.05. In most cases, the magnitude and direction of change from the nominal  $\alpha$  are unknown. The consequences for disposal can be especially grave in terms of increased cost and difficulty when departures from the statistical test assumptions result in highly inflated Type I error rates.

Changes in  $\alpha$  can be assessed for given situations using simulations and are described in the following sections for the LSD test. Similar effects would be expected for the  $t$ -test, which is a special case of the LSD test when number of treatments  $k = 2$ . The simulations generally compared three dredged sediment samples with a reference sediment sample. Samples were drawn from populations having known characteristics, including type of frequency distribution (normal, lognormal, or gamma), mean, and coefficient of variation ( $CV = \text{standard deviation} \div \text{mean}$ ). Details of the simulation study are given in Clarke (1995b). CVs for simulations using equal and unequal population variances are provided in Table 2 along with a key to subsequent figures.

**Table 2**  
**Coefficients of Variation for Simulations Comparing Three Dredged Sediments With a Reference Sediment**

Key	Variances	CVs (Reference Sediment, Three Dredged Sediments)			
.1	Equal	0.1	0.1	0.06	0.05
.5	Equal	0.5	0.5	0.31	0.23
1	Equal	1.0	1.0	0.63	0.45
2	Equal	2.0	2.0	1.25	0.91
A	Unequal	0.1	0.7	0.63	0.05
B	Unequal	0.1	1.6	0.56	0.09
C	Unequal	0.1	1.9	0.25	0.27
D	Unequal	0.2	0.2	1.75	0.18
E	Unequal	0.2	1.6	0.44	1.14
F	Unequal	0.3	0.1	0.13	1.91
G	Unequal	0.4	1.8	1.19	0.59
H	Unequal	0.5	1.3	0.88	0.27
I	Unequal	0.6	1.7	1.00	1.14
J	Unequal	0.7	0.1	0.19	1.50
K	Unequal	0.7	0.9	1.19	0.77
L	Unequal	0.7	1.6	1.56	0.18
M	Unequal	0.8	0.1	0.25	0.27
N	Unequal	0.8	0.9	0.88	0.32
O	Unequal	0.9	0.8	0.06	0.59
P	Unequal	1.0	0.1	1.25	0.23
Q	Unequal	1.0	1.0	1.00	1.00
R	Unequal	1.0	1.4	0.06	0.09
S	Unequal	1.2	0.3	1.13	1.68
T	Unequal	1.2	0.7	0.88	0.45
U	Unequal	1.4	1.9	0.31	0.77
V	Unequal	1.5	0.3	0.06	0.68
W	Unequal	1.5	0.5	0.44	1.14
X	Unequal	1.6	0.5	0.13	0.05
Y	Unequal	1.6	1.0	1.38	0.77
Z	Unequal	1.8	1.1	1.06	0.14
AA	Unequal	1.9	0.2	0.56	0.09

## Effect of violations of assumptions on $\alpha$

Certain parametric statistical procedures, such as analysis of variance,  $t$ -test, and LSD test, assume that the data are randomly sampled from normally distributed populations, and that variances are equal among treatments. However, these procedures are often said to be robust to departures from their assumptions, when sample sizes are equal. Robustness here refers to the ability of a test to maintain its prespecified  $\alpha$  regardless of the characteristics of the sample data. Figures 7 and 8 display mean  $\alpha$  of untransformed data from simulations using samples from normal and nonnormal (lognormal and gamma) distributions in the LSD test. When sample sizes are equal (Figure 7) or unequal (Figure 8) and variances are equal among treatments,  $\alpha$  remains approximately 0.05 or less regardless of distribution. Thus, the LSD test could be considered fairly robust to departures from normality when variances are equal among treatments, regardless of whether sample sizes are equal or unequal. When variances are unequal,  $\alpha$  often departs considerably from 0.05, for both normal and nonnormal distributions. Thus, the LSD test is not robust to inequality of variances, whether sample sizes are equal or unequal.

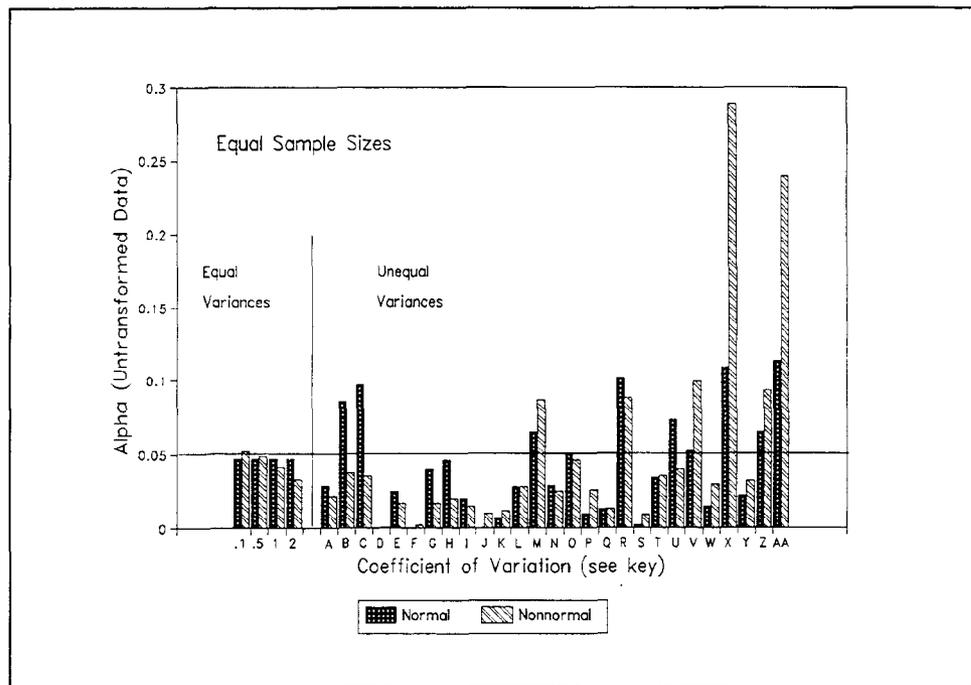


Figure 7. Mean  $\alpha$  for LSD test using untransformed samples from normal and nonnormal distributions. Equal sample sizes (see Table 2 for key to CVs)

## Transformations and $\alpha$

When samples are drawn from nonnormal populations, an appropriate data transformation can help in meeting the assumptions of parametric statistical tests by normalizing the data. The effect of transformations on  $\alpha$  of the LSD

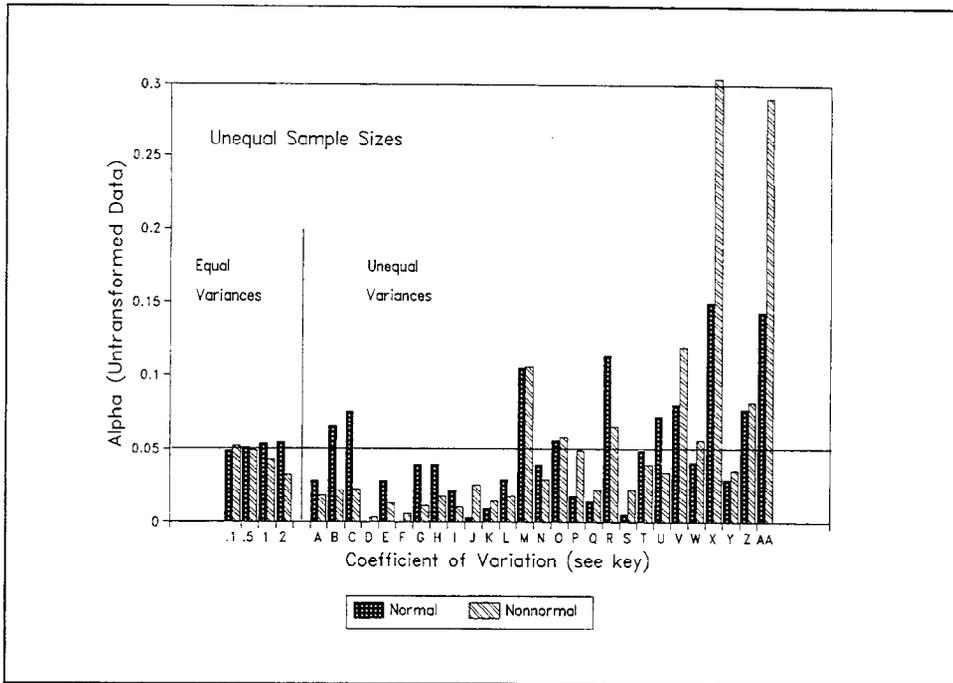


Figure 8. Mean  $\alpha$  for LSD test using untransformed samples from normal and nonnormal distributions. Unequal sample sizes (see Table 2 for key to CVs)

test is shown in Figures 9 and 10. Figure 9 displays mean  $\alpha$  for log-transformed and untransformed data when samples are drawn from lognormal distributions. Figure 10 displays mean  $\alpha$  for rankits and untransformed data when samples are drawn from gamma distributions. Trends in  $\alpha$  are similar for both distributions, although more pronounced for the lognormal. When variances are equal, transformation increases  $\alpha$  slightly compared with untransformed data, especially as the CV increases. When variances are unequal and the reference sediment CV is low ( $< 0.8$ ),  $\alpha$  for both transformed and untransformed data is generally well below 0.05. However, as the reference sediment CV exceeds 0.8,  $\alpha$  for transformed data increases considerably in many situations.  $\alpha$  for untransformed data tends to remain low except when the reference sediment CV is greater than or equal to the CVs of the dredged sediment samples to which the reference is being compared. It is apparent that data transformation of nonnormal samples decreases rather than increases the robustness of the LSD test when variances are unequal, especially when the reference sediment CV is high.

Because inequality of variances can result in a highly inflated Type I error rate in the LSD test, using  $t$ -tests is recommended for individual dredged sediment-reference sediment comparisons when the data fail the test for equality of variances (see Chapter 6). Statistical testing sequences are as shown in Figures D-4 and D-5 of the Inland Testing Manual. If variances for an individual dredged sediment-reference sediment comparison are unequal, then the  $t$ -test for unequal variances should be used. By reducing the degrees of

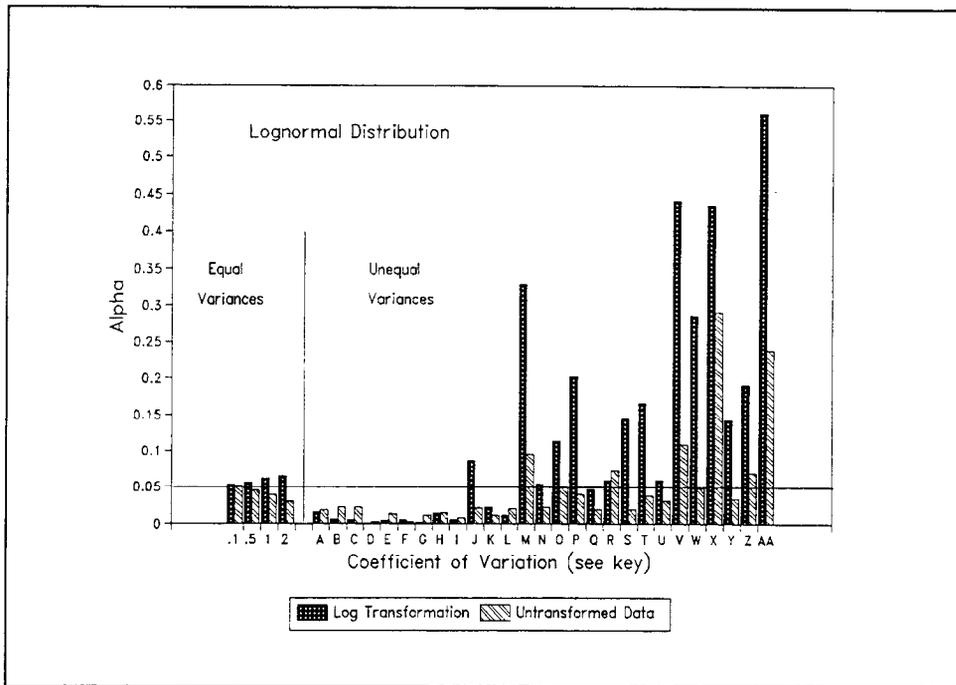


Figure 9. Mean  $\alpha$  for LSD test using log-transformed and untransformed samples from lognormal distributions (see Table 2 for key to CVs)

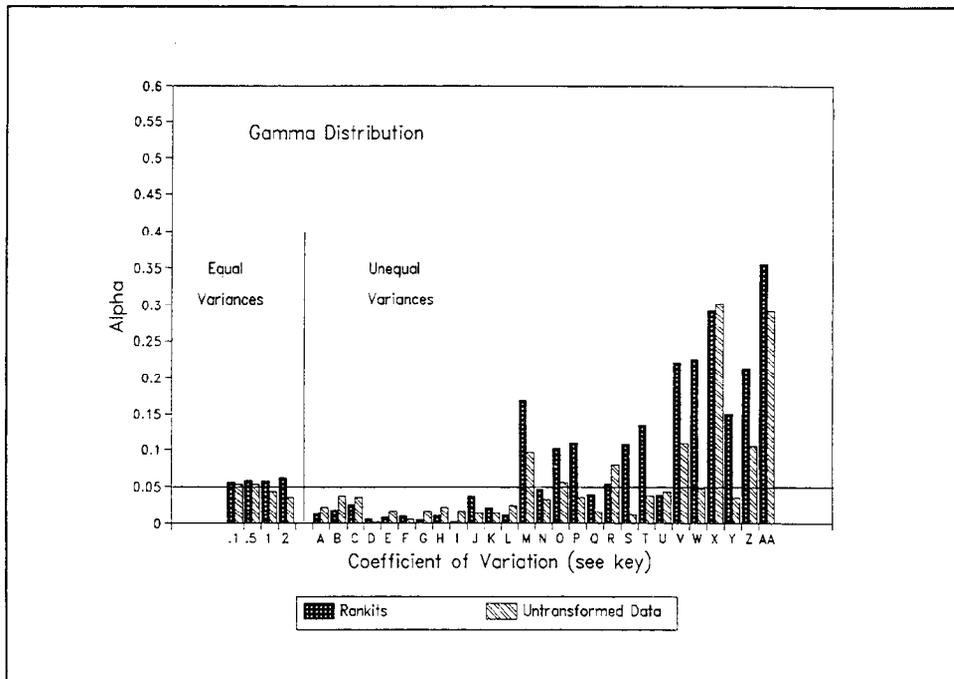


Figure 10. Mean  $\alpha$  for LSD test using rankit-transformed and untransformed samples from gamma distributions (see Table 2 for key to CVs)

freedom in proportion to the inequality of variances, the *t*-test for unequal variances reduces the likelihood of Type I error.

## Type II Error and Power

When two or more populations differ with respect to a given parameter such as the mean, a hypothesis test of samples from those populations is expected to conclude that the populations do in fact differ. The probability of reaching that conclusion is known as the power of the test and is designated as  $1 - \beta$ . If the hypothesis test concludes that the population parameters do not differ, when in reality the population parameters are different, then a Type II error (probability  $\beta$ ) has been committed. While statistical hypothesis tests usually operate on a prespecified  $\alpha$ ,  $\beta$  (or power) must be determined using a formula that incorporates  $\alpha$ , sample size, standard deviation, and effect size (i.e., the amount of difference in population parameters that will be considered significant). For comparison of two samples, the formula is given as Equation 10 in Appendix D of the Inland Testing Manual. Too often, statistical hypothesis tests are conducted with little thought given to a biologically meaningful effect size and no attempt to determine the power of the statistical test to detect that effect size.

### Why power is important

In dredged sediment evaluations, the power of a statistical comparison is the ability of that test to determine that a dredged sediment is worse than the reference sediment in terms of the end point being measured, when that is the true state of nature (see Table 1). When a test lacks sufficient power, the likelihood of Type II error, with possible adverse environmental consequences, is high. Obviously, for protection of the environment, one would like power to be as high as possible. How much power is sufficient? That can only be determined by weighing the relative importance of Types I and II error rates, deciding on a meaningful effect size, and balancing the trade-off between sample size (= cost) and power. The interrelationship of all these factors is discussed in the next section.

### Sample size, effect size, $\alpha$ , and power

If the confidence level of a test is predetermined at 0.95, is it possible to also have power equal to 0.95? Generally, it is not, given the high cost of biological testing and trace contaminant chemical analyses, since power is proportional to sample size. The influence of sample size and effect size (relative to the pooled standard deviation) on the power of a *t*-test is shown in Table 3. Power increases with effect size as well as with sample size.

To detect a difference equal to one standard deviation with a power of 0.95 when  $\alpha = 0.05$ , a *t*-test would require  $n = 23$ . Sample size would drop to 7

**Table 3**  
**Power of a *t*-test for Various Relative Effect Sizes When  $\alpha = 0.05$**

Sample Size	Effect Size Relative to Standard Deviation					
	0.5	0.75	1.0	1.5	2.0	3.0
2	0.068	0.081	0.097	0.146	0.227	0.528
3	0.102	0.146	0.208	0.391	0.617	0.901
4	0.131	0.206	0.308	0.568	0.795	0.969
5	0.158	0.260	0.394	0.689	0.886	0.990
6	0.183	0.309	0.469	0.775	0.935	0.997
7	0.207	0.356	0.535	0.837	0.963	0.999
8	0.230	0.400	0.593	0.882	0.979	0.999
9	0.252	0.439	0.644	0.915	0.988	0.999
10	0.273	0.478	0.689	0.939	0.993	0.999
12	0.314	0.547	0.764	0.968	0.998	0.999
15	0.371	0.637	0.846	0.989	0.999	1.000
20	0.459	0.751	0.926	0.998	0.999	1.000
25	0.536	0.833	0.965	0.999	1.000	1.000
30	0.604	0.889	0.984	0.999	1.000	1.000
40	0.715	0.952	0.997	1.000	1.000	1.000
50	0.798	0.980	0.999	1.000	1.000	1.000
100	0.969	0.999	1.000	1.000	1.000	1.000

if the effect size were twice the standard deviation. To have a power of 0.95 when  $n = 5$  and  $\alpha = 0.05$ , a *t*-test could only detect a difference of nearly three times the standard deviation. To detect a difference of one standard deviation when  $n = 5$  and  $\alpha = 0.05$ , power is only 0.39. If a difference in population means of one standard deviation is considered biologically significant for a particular comparison, then clearly there is a high probability (0.61) that the *t*-test will not be able to detect that difference. In this case, a coin toss would do better.

From Table 3, it is apparent that the effect size must be huge (three standard deviations) for a test to have high power ( $\geq 0.90$ ) when sample size is small ( $n = 3$  to 5) and  $\alpha = 0.05$ . A sample size of 2 is essentially useless because of low power to detect even gross differences. To detect a small effect size of half the standard deviation with high power, more than 50 samples are required for each treatment when  $\alpha = 0.05$ .

## Why $\alpha$ should not be restricted to 0.05

The Ocean Dumping Regulations (40 CFR 227.6(c)(2) and (3)) currently mandate 95-percent confidence ( $\alpha = 0.05$ ) for statistical comparisons in dredged sediment evaluations. For consistency, the Inland Testing Manual also recommends  $\alpha = 0.05$ , although no specific confidence level is required in the Guidelines for Specification of Disposal Sites for Dredged or Fill Material (40 CFR Part 230) under the Clean Water Act. Mandating  $\alpha = 0.05$  can have two undesirable consequences for dredged sediment evaluations: (a) if the cost of testing is kept reasonable, the power of the tests to detect adverse effects could be insufficient, or (b) if tests are sufficiently powerful to detect adverse effects, the cost of testing could be unreasonable.

The influence of  $\alpha$  on the power of a  $t$ -test is shown in Figure 11 for an effect size of one standard deviation and sample sizes of 3, 5, 8, and 10. Raising  $\alpha$  to 0.10 or even 0.20 can have a dramatic effect on power, especially when  $n$  is small. For  $n = 5$ , power increases from 0.39 to 0.57 when  $\alpha$  is raised to 0.10, and power increases to 0.75 if  $\alpha$  is raised to 0.20. For  $n = 5$  and effect size equal to one standard deviation, power and confidence are approximately equal when  $\alpha = 0.25$ .

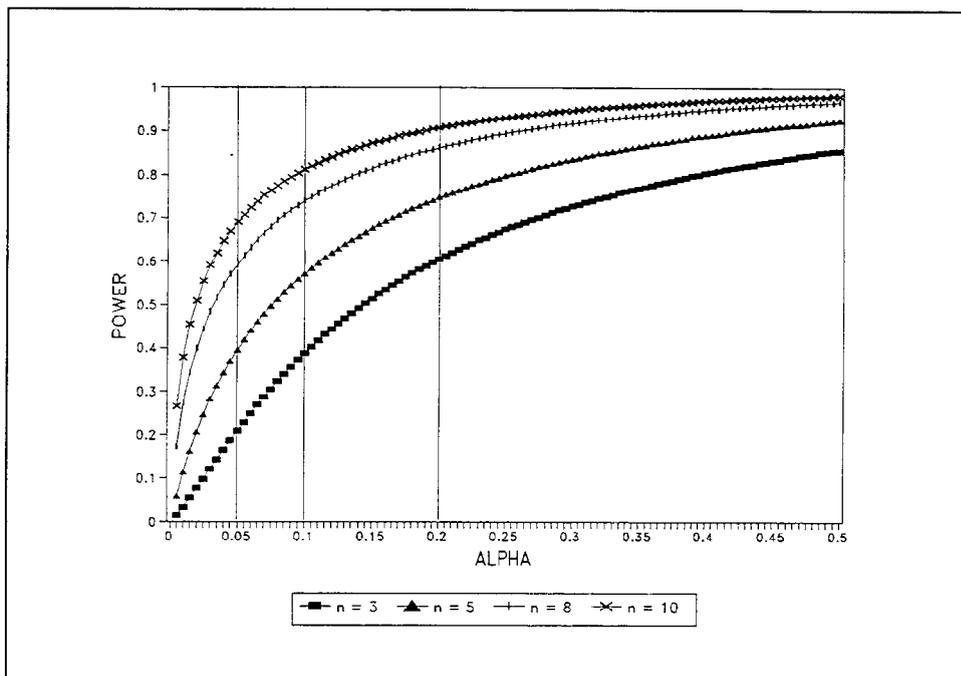


Figure 11. Power of a  $t$ -test for effect size of one standard deviation

Investigators should consider carefully the relative importance of Type I and Type II errors in the statistical comparisons conducted during dredged sediment evaluations. Mandating  $\alpha = 0.05$  practically guarantees that power will be low unless relative effect size is large. The first step is determining a meaningful effect size for a particular comparison, which requires an estimate

of the standard deviation. Such an estimate might be obtained from a pilot study or from previous data. Once the effect size is determined, a logical course might be to select a reasonable sample size and then choose  $\alpha$  such that Types I and II error rates are roughly equal. This strategy would afford equal consideration to protection against adverse environmental impact and to minimizing cost. Examples of balanced Types I and II error rates in a  $t$ -test are given in Table 4 for several sample sizes and relative effect sizes. Shaded cells indicate the effect sizes and sample sizes for which both power and confidence  $\geq 0.90$ .

<b>Table 4</b>								
<b>Balanced Types I and II Error Rates for a <math>t</math>-test<sup>1</sup></b>								
Relative Effect Size	$n = 3$		$n = 5$		$n = 8$		$n = 12$	
	$\alpha$	$\beta$	$\alpha$	$\beta$	$\alpha$	$\beta$	$\alpha$	$\beta$
0.5	0.40	0.37	0.35	0.35	0.30	0.33	0.25	0.30
0.75	0.35	0.32	0.30	0.27	0.25	0.22	0.20	0.17
1.0	0.30	0.27	0.25	0.20	0.15	0.19	0.10	0.14
1.5	0.20	0.21	0.15	0.12	0.10	0.06	0.05	0.03
2.0	0.15	0.14	0.10	0.06	0.05	0.02	0.01	0.01
3.0	0.05	0.10	0.05	0.01	0.01	0.002	0.01	0.00004

<sup>1</sup> For convenience,  $\alpha$  is specified to the nearest 0.05 or to 0.01

In many statistical packages, the default Type I error rate is 0.05 for statistical comparison tests, but other values of  $\alpha$  can be specified. In SAS, any value of  $\alpha$  between 0.0001 and 0.9999 can be specified for the LSD test in PROC GLM or PROC ANOVA (SAS Institute, Inc. 1988a). When a procedure gives the P-value corresponding to the test statistic, as in the SAS TTEST procedure, that probability is simply compared with the desired  $\alpha$  to determine its significance.

### Increasing power by increasing the number of reference replicates

Appendix D of the Inland Testing Manual (Section D2.2.1) mentions that the power of a statistical comparison can be increased prior to conducting the experiment by increasing the number of reference replicates while keeping the number of dredged sediment replicates the same or even decreasing it. This works to increase power when the degrees of freedom for an individual reference sediment-dredged sediment comparison are increased. One simple means of accomplishing this is to decrease each dredged sediment sample size by one and then add those extra samples to the reference sediment, keeping the total number of replicates the same. For example, if  $n = 5$  would normally be used in comparing three dredged sediments with a reference sediment for a

total sample size of  $4 \times 5 = 20$ , each dredged sediment sample size could be decreased to 4 and the reference sediment sample size increased to 8. The degrees of freedom for each dredged sediment-reference sediment comparison,  $n_{ref} + n_{dredged} - 2$ , would then be 10 rather than 8. Power to detect a difference of one standard deviation when  $\alpha = 0.05$  would increase from 0.39 to 0.47, with no increase in expense.

Using this method, the percent increase in power to detect a difference of one standard deviation when  $\alpha = 0.05$  is illustrated in Figure 12 for total number of treatments ranging from 2 to 10 and initial sample sizes of 3, 4, 5, and 6. When there are two treatments, reallocating samples from the dredged sediment to the reference sediment does not change the degrees of freedom, and thus power does not change. When there are three or more treatments, a power increase is realized by this reallocation of samples. The percent increase in power is greatest when initial sample size is very small (e.g.,  $n = 3$ ) and when there are many treatments.

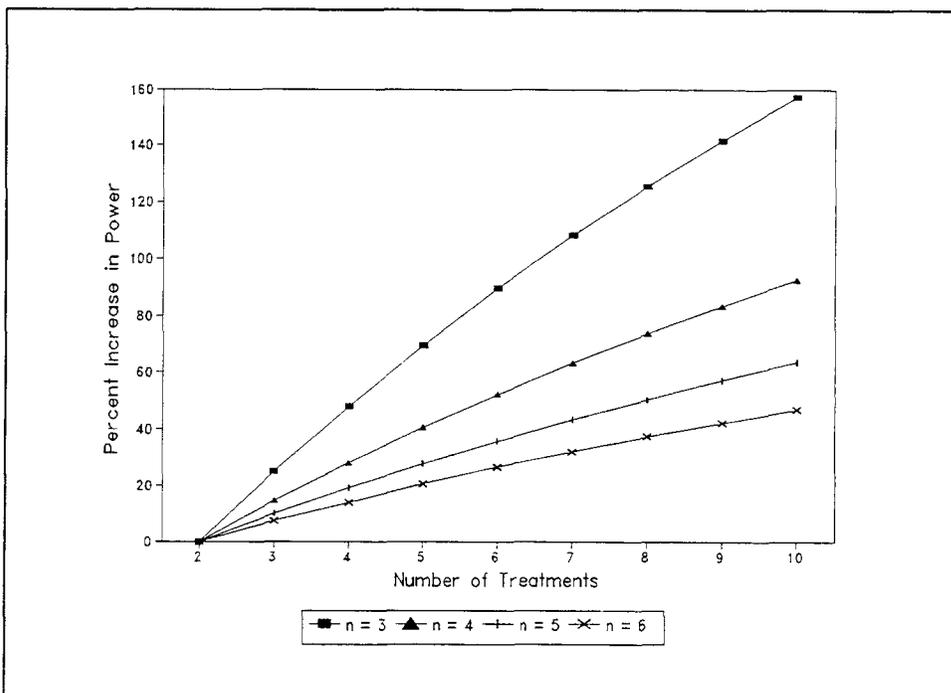


Figure 12. Percent increase in power when one sample from each dredged sediment is reallocated to the reference sediment ( $n$  = initial sample size)

Reallocation of samples from the dredged sediments to the reference sediment, while effective at increasing power, has the drawback of creating an unbalanced design, especially when several dredged sediments will be compared with the reference sediment and initial sample size is small. An unbalanced design can adversely affect Types I and II error rates of a statistical comparison when the normality and equality of variances assumptions are not met. Unbalanced design thus requires higher  $\alpha$  levels for tests of those

assumptions (see Table D-2 of the Inland Testing Manual). The gain in power from reallocating replicates could be offset by a loss in power from having to use *t*-tests for unequal variances or a nonparametric comparison procedure because the data did not pass the tests of assumptions. Nevertheless, the reallocation of replicates is felt to be generally advantageous and should be considered whenever initial power calculations indicate that a design with equal sample sizes may have insufficient power to detect a biologically meaningful effect size.

### Effect of frequency distribution and unequal variances on power

Discussions of power up to this point have been concerned only with ideal data, for which the parametric test assumptions of normality and equality of variances are satisfied. What happens when the data are nonideal and one or both assumptions are violated? Effects on power can be surmised using simulations, as was done for  $\alpha$  earlier in this chapter. Simulations were conducted using the LSD test with  $\alpha = 0.05$ , equal and unequal variances, and sample sizes ranging from 3 to 8. The effect of nonnormal distribution on the power of the LSD test using untransformed data is shown in Figure 13. Power using samples from lognormal or gamma populations was compared with that of samples from normal populations. In almost all cases, power increases, some times substantially, with nonnormality when untransformed data are used. Power increase is greater for the lognormal distribution than for the gamma

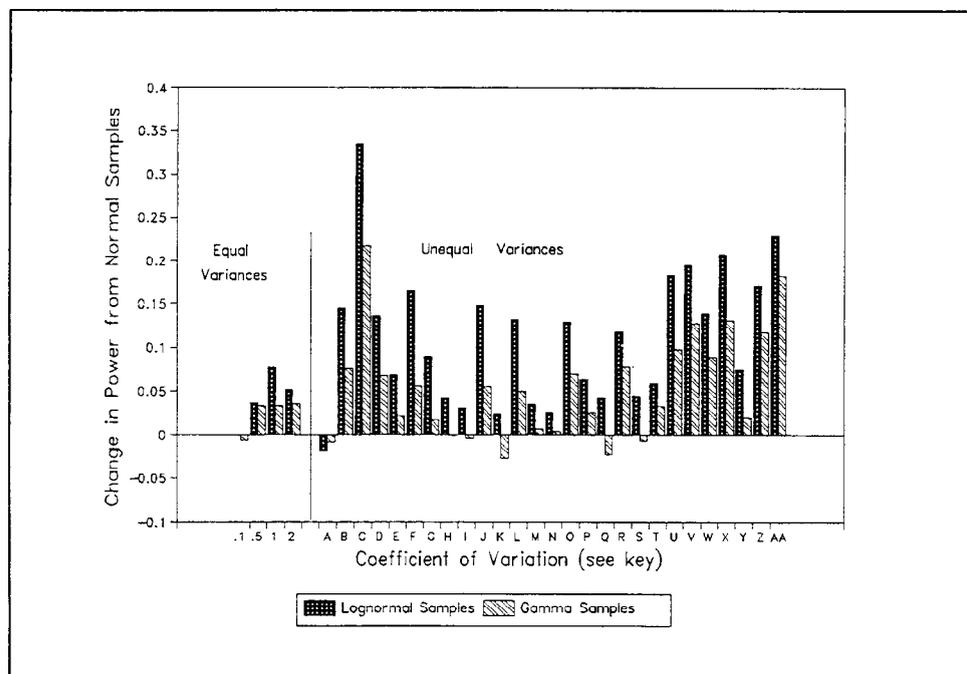


Figure 13. Change in power of LSD test using untransformed samples from lognormal and gamma distributions compared with untransformed samples from normal distribution (see Table 2 for key to CVs)

distribution, and is generally more pronounced for unequal variances than for equal variances. A comparison with Figures 7 and 8 indicates that the increase in power with nonnormality can come at the cost of inflated Type I error rate when variances are unequal and the reference sediment CV is greater than the dredged sediment CVs.

### Effect of transformation on power

In Figures 9 and 10 data transformation of samples from lognormal and gamma distributions had little effect on  $\alpha$  of the LSD test unless variances were unequal and the reference sediment CV was high. Do transformations intended to normalize lognormal or nonnormal data increase the power of LSD comparisons? Data from a lognormal distribution may be normalized using either a log transformation or rankits. When compared with untransformed data from a lognormal distribution, log transformation increases power of the LSD test when variances are equal and CV is moderate to high, and when variances are unequal and the reference sediment CV  $\geq 0.8$  (Figure 14). Rankits increase power in these same circumstances, but to less extent than logs. When variances are unequal and the reference sediment CV is low, log transformation of lognormal data decreases power of the LSD test compared with untransformed lognormal data. Rankits, on the other hand, sometimes increase power when variances are unequal and the reference sediment CV is low.

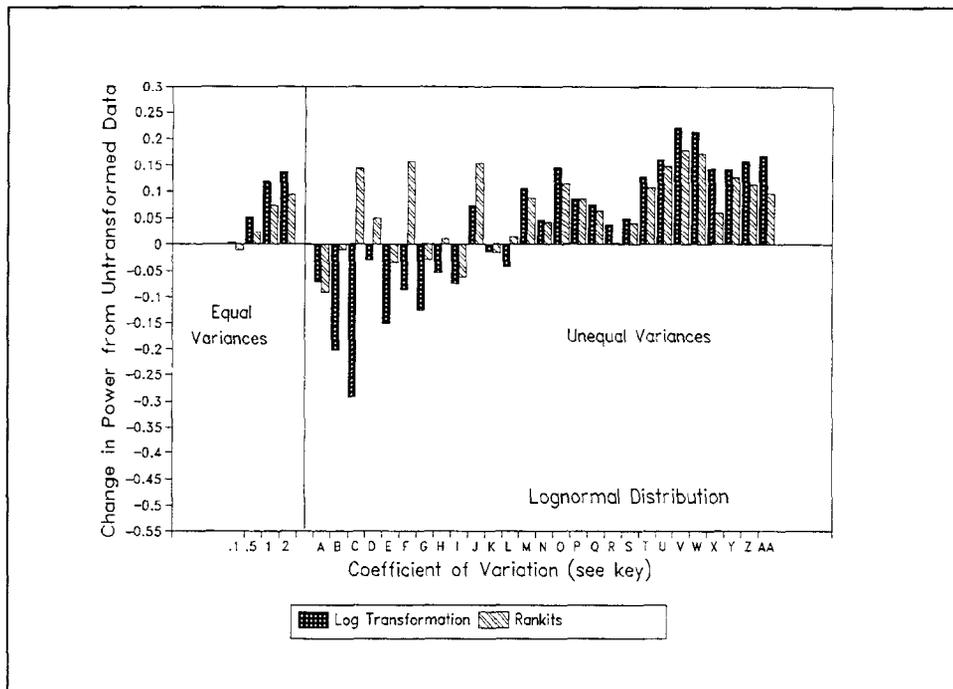


Figure 14. Change in power of LSD test using transformed lognormal samples compared with untransformed lognormal samples (see Table 2 for key to CVs)

The effect of transformation on the power of the LSD test when samples come from gamma distributions (Figure 15) is similar to the effect of transformation of lognormal samples. However, the power loss from log transformation of gamma samples is drastic when variances are unequal and the reference sediment CV is low. In some of these cases, the LSD test had power close to zero. Rankits will normalize data from gamma populations, but log transformation will not.

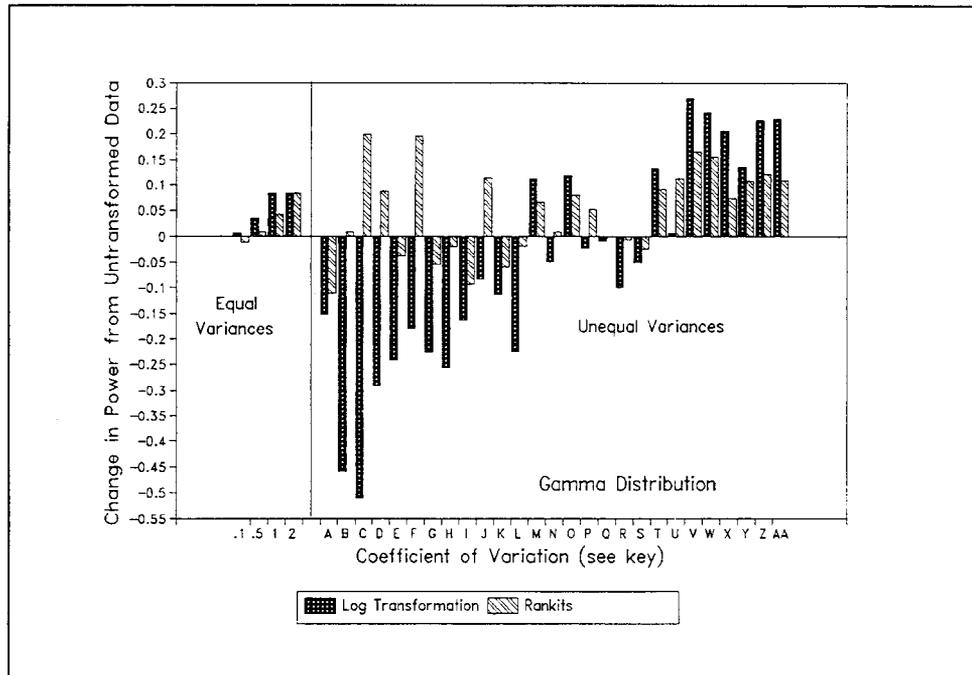


Figure 15. Change in power of LSD test using transformed gamma samples compared with untransformed gamma samples (see Table 2 for key to CVs)

### Calculating power

The desired power of a dredged sediment-reference sediment comparison should be decided at the time the experiment is being designed. Formulae for determining the sample size needed to detect a given effect size at a desired power are provided in Equations 8 and 9 of the Inland Testing Manual (Appendix D, Section D2.1.1.1). Tables 3 and 4 in this chapter can also be useful as quick references to power, relative effect size, and sample size. When sample sizes must be limited, the power of a *t*-test (or LSD test) to detect a difference of the given effect size using a specified sample size can be calculated using Equation 10 of the Inland Testing Manual (Appendix D, Section D2.1.1.1). SAS statements for implementing Equation 10 are given in program POWER.SAS in Appendix A. Remember that power calculated prior to performing a test assumes ideal data and the fulfillment of the statistical test assumptions.

After a statistical test has been performed with actual data, which are often less than ideal, how does one assess the power of the test? One simple way is to look at the least significant difference, which is the difference that can be detected when power is 0.50 and  $t_{1-\beta,df} = 0$ . SAS provides this in the output for the LSD test when sample sizes are equal, or it can be calculated for each dredged sediment-reference sediment comparison using Equation 11 from Appendix D of the Inland Testing Manual. If a log transformation was employed for the test, the least significant difference should be transformed back to the original scale (this is not possible when rankits were used). If the least significant difference is greater than a meaningful effect size for a given comparison, then the statistical test lacks sufficient power. SAS statements for the least significant difference formula are given in program DMIN.SAS in Appendix A.

The least significant difference is the magnitude of difference between two population means that a  $t$ -test could be expected to detect 50 percent of the time given a large number of comparisons with the same sample sizes and variance. To determine the detectable difference when power = 0.95 and  $\alpha = 0.05$ , multiply the least significant difference by 2.

The SAS programs in Appendix D of the Inland Testing Manual, as well as the SYSTAT and SPSS programs in Appendixes B and C of this report, provide statements to calculate the power of an LSD test to detect specified differences from the reference sediment mean, based on the actual sample sizes and standard deviations of the data used in the test. Effect sizes are specified as various percent decreases from reference mean survival or percent increases from reference mean bioaccumulation. Thus, these calculations represent the power of an LSD test to detect a certain percent change from the reference sediment end point, given normally distributed populations with standard deviation equal to the pooled sample standard deviation. For the example single-time point bioaccumulation data given in the Inland Testing Manual, the LSD test can detect a 100-percent increase in contaminant concentration above the reference sediment mean with a power of  $\approx 0.5$ , and a 200-percent increase with a power of  $\approx 0.95$ .

## Summary

Statistical hypothesis testing can result in two types of errors. Type I error (probability  $\alpha$ ) occurs if the test concludes that sample means differ when the populations from which those samples were drawn have identical means. Type II error (probability  $\beta$ ) occurs if the test concludes that sample means are the same when the populations from which those samples were drawn have different means. Confidence is the ability of the test to avoid Type I error, and power is the ability of the test to avoid Type II error. Confidence, power, sample size, and effect size (the amount of difference that the test can detect as significant) are interdependent and should be determined during the design of an experiment.

- Type I error in dredged sediment evaluations would be environmentally protective but could lead to unnecessary expense.
- Type II error in dredged sediment evaluations could result in adverse environmental impact.
- Although  $\alpha = 0.05$  is currently mandated for statistical comparisons in ocean disposal evaluations, it may be more sensible to balance Types I and II error rates.
- Parametric statistical tests assume normality and equality of variances among treatments. The LSD test is robust (i.e.,  $\alpha$  remains constant) to departures from normality when variances are equal, but is not robust to inequality of variances.
- Data transformation can increase Type I error rate of the LSD test when variances are unequal. Log transformation can also increase Type II error rate when variances are unequal and the reference sediment coefficient of variation is low.
- The power of a statistical comparison is increased by increasing the sample size, effect size, or  $\alpha$ .
- When two or more dredged sediments will be simultaneously compared with a reference, the power of the statistical comparison can be increased by using more replicates for the reference sediment than for the dredged sediments.
- When a statistical comparison is not significant, the least significant difference should be calculated to determine whether the test had sufficient power to detect a meaningful effect size.

## 5 Tests of Assumptions: Normality

---

After a properly designed experiment has been completed and the experimental data collected, statistical analyses may be needed to facilitate data interpretation and decision making. The primary statistical tests recommended in the Inland Testing Manual are the *t*-test for comparison of two treatments and the LSD test for comparison of more than two treatments. These parametric tests assume that each treatment is sampled from a normally distributed population and that variances for all treatments are equal or similar. The normality assumption will be examined in this chapter for toxicity test (survival) data and for bioaccumulation (chemical concentration) data.

### Survival Data

Survival data from water column and benthic toxicity tests are not normally distributed. Because the number of organisms used in a toxicity test sample cannot be infinite and the measure of survival is dichotomous (alive or dead), the survival proportion  $p$  from a sample is a discrete variable from a binomial distribution. Unlike a normal distribution, the variance of a binomial distribution is a function of the mean. Nevertheless, a binomial distribution can be adequately approximated by a normal distribution having the same mean  $\mu$  and standard deviation  $\sigma$  providing the interval  $\mu \pm 3\sigma$  lies completely within the range of values for the binomial distribution (McClave and Dietrich 1979). In the case of survival data, if a sample initially contains 20 animals, then  $\mu \pm 3\sigma$  should lie within the interval from 0 to 20; survival ranging from 35 to 65 percent is within that interval. The Inland Testing Manual (Section D2.1.1.1) recommends automatic use of the arcsine-square root transformation for all survival data. This transformation does not necessarily normalize binomial proportion data, but does remove the dependency of the variance on the mean (Sokal and Rohlf 1981).

## Testing for normality

Following arcsine-square root transformation, the residuals of the transformed survival data should be tested for normality using a procedure such as the Shapiro-Wilk's Test. If the transformed survival data fail the normality test, then a nonparametric comparison procedure ( $t$ -test or LSD test on ranks or rankits; see Figures D-1 and D-4B in the Inland Testing Manual) should be used.

Simulated data from binomial distributions (1,000 simulations, total number of replicates  $N$  ranging from 3 to 40) were subjected to the Shapiro-Wilk's Test using SAS (SAS Institute, Inc. 1988b), to gauge the performance of this test with binomial data. Mean survivals of 10, 16, and 19 out of 20 were used ( $p = 0.5, 0.8, \text{ and } 0.95$ ). The proportions of samples passing the Shapiro-Wilk's Test, i.e., closely approximating a normal distribution, are shown in Figures 16 and 17 for balanced and unbalanced designs using the residuals of both raw data and arcsine-transformed data. Significance of the Shapiro-Wilk's Test was determined by comparing the test statistic P-value with the appropriate  $\alpha$  value in Table D-2 of the Inland Testing Manual. Results for mean survivals of 1 and 4 out of 20 may be considered equivalent to results for mean survivals of 19 and 16, respectively. When mean survival is in the range of 4 to 16 ( $p = 0.2$  to  $0.8$ ), a high proportion of samples passes the normality test, especially when number of replicates is equal among treatments (Figure 16, balanced design). However, when mean survival is at

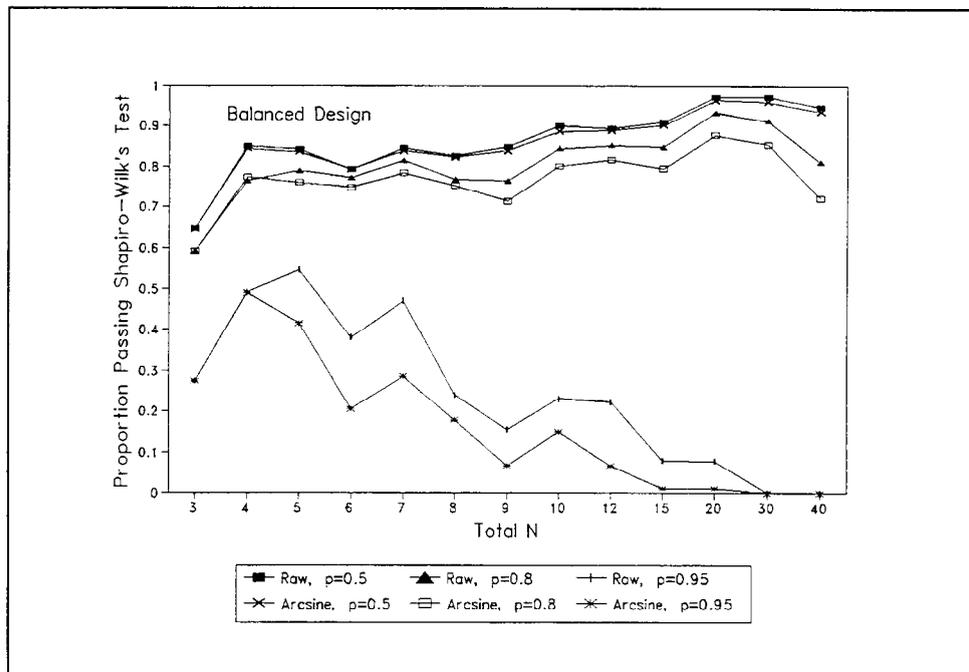


Figure 16. Results of Shapiro-Wilk's Test on residuals of untransformed (raw) and arcsine-transformed simulated binomial data, balanced design, where  $p$  is survival proportion

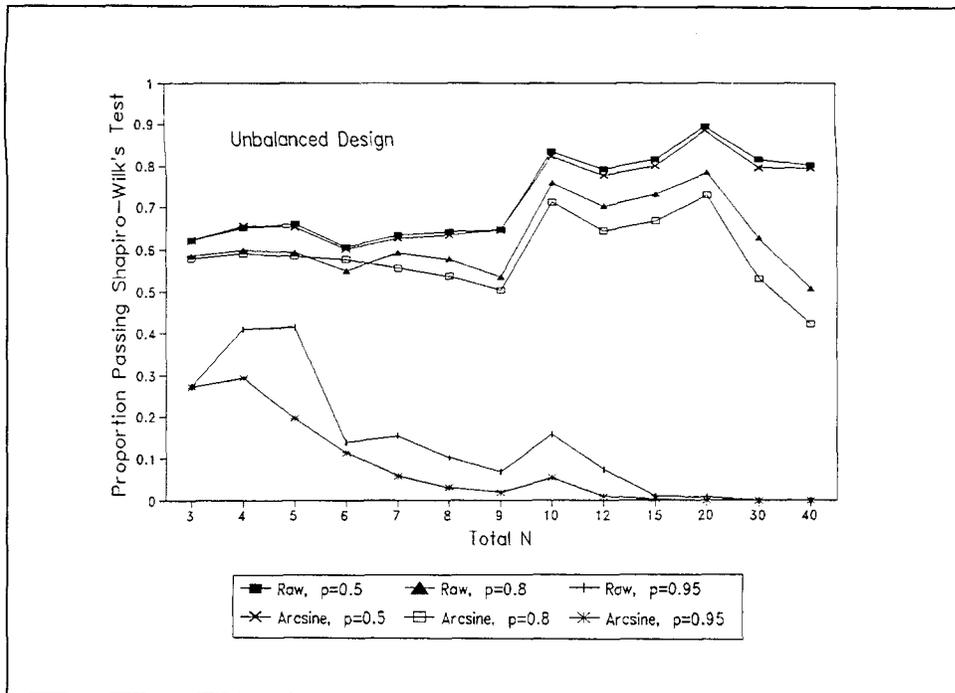


Figure 17. Results of Shapiro-Wilk's Test on residuals of untransformed (raw) and arcsine-transformed simulated binomial data, unbalanced design, where  $p$  is survival proportion

the upper or lower fringes of the distribution ( $p$  around 0.1 or 0.9), fewer samples pass the normality test, especially when data are arcsine-transformed or number of replicates is unequal among treatments (Figure 17, unbalanced design).

Thus, in the middle range of survival proportions, when the binomial distribution approximates the normal distribution, the Shapiro-Wilk's Test does not differentiate the two distributions. Data analysis may proceed using parametric tests with the arcsine-square root transformed data (Figure D-1 in Appendix D of the Inland Testing Manual). However, when survival proportion is very low or very high, the data do not approximate a normal distribution, and the Shapiro-Wilk's Test will most likely reject normality. In this case, conversion to rankits is advised prior to performing statistical comparisons.

### Example data

Three sets of actual survival data are used to illustrate varying results of the Shapiro-Wilk's Test for normality. These data, taken from proposed USEPA dredged material bioassays using Great Lakes sediments (Moore et al. 1994), include balanced and unbalanced designs, sample sizes  $n$  ranging from 1 to 7 replicates, and total replicates  $N$  ranging from 28 to 70 (Table 5).

Table 5 Survival of Animals Exposed to Great Lakes Sediments			
Species	Sediment	No. of Replicates	No. of Survivors out of 10
<i>Pimephales promelas</i>	Control-2 (0% elutriate)	4	9, 8, 8, 8
<i>Pimephales promelas</i>	RM 1-2 (100% elutriate)	3	4, 3, 3
<i>Pimephales promelas</i>	LM 0-1 (100% elutriate)	1	3
<i>Pimephales promelas</i>	LM 1-2 (100% elutriate)	4	7, 5, 4, 6
<i>Pimephales promelas</i>	LM 2-3 (100% elutriate)	4	6, 4, 5, 2
<i>Pimephales promelas</i>	LM 3-4 (100% elutriate)	4	3, 2, 3, 4
<i>Pimephales promelas</i>	LM 4-5 (100% elutriate)	4	4, 1, 1, 3
<i>Pimephales promelas</i>	LM 7-8 (100% elutriate)	4	4, 3, 4, 3
<i>Pimephales promelas</i>	LM 11-12 (100% elutriate)	4	5, 8, 4, 3
<i>Hyaella azteca</i>	DWR	7	10, 8, 8, 7, 10, 8, 9
<i>Hyaella azteca</i>	SWR	7	9, 10, 9, 8, 10, 10, 10
<i>Hyaella azteca</i>	RM 1-2	7	6, 8, 10, 9, 10, 10, 9
<i>Hyaella azteca</i>	LM 0-1	7	9, 9, 7, 8, 10, 9, 10
<i>Hyaella azteca</i>	LM 1-2	7	0, 0, 0, 2, 0, 7, 10
<i>Hyaella azteca</i>	LM 2-3	7	9, 0, 0, 0, 8, 8, 5
<i>Hyaella azteca</i>	LM 3-4	7	10, 7, 8, 10, 10, 8, 9
<i>Hyaella azteca</i>	LM 4-5	7	9, 9, 8, 8, 10, 8, 10
<i>Hyaella azteca</i>	LM 7-8	7	10, 9, 8, 8, 10, 8, 8
<i>Hyaella azteca</i>	LM 11-12	7	10, 4, 8, 8, 9, 9, 8
<i>Hyaella azteca</i>	MC-4 (REF)	7	9, 10, 10, 7, 10, 10, 10
<i>Hyaella azteca</i>	MC-1	7	10, 9, 10, 10, 10, 9, 10
<i>Hyaella azteca</i>	MC-2	7	3, 9, 10, 10, 10, 9, 10
<i>Hyaella azteca</i>	MC-3	7	6, 9, 9, 9, 8, 9, 10

Results of Shapiro-Wilk's Tests on residuals of the raw and arcsine-transformed data are presented in Table 6.

The data from the *Pimephales promelas* (fathead minnow) bioassay consist of survival from a 0-percent elutriate control, a 100-percent elutriate river mile (RM) sediment, and seven 100-percent elutriate lake mile (LM) sediments. Survival in most replicates is moderate, with no survivals of 0 or 10. RM 1-2 and LM 0-1 have reduced numbers of replicates ( $n = 3$  and  $1$ , respectively); the other sediments each have four replicates. Thus, the design is unbalanced because  $n_{\min}$  is less than half of  $n_{\max}$ ,  $N = 32$ , and  $\alpha = 0.05$  from Table D-2 of the Inland Testing Manual is used for the normality test. The probability associated with Shapiro-Wilk's  $W$  greatly exceeds  $0.05$  regardless of whether the data are untransformed or arcsine-transformed. Therefore, the data pass the normality test and may be considered approximately normally distributed.

Species	Total No. of Replicates	Design	$\alpha$ for Shapiro-Wilk's Test	Untransformed Data			Arcsine-Transformed Data		
				Shapiro-Wilk's W	Probability	Pass/Fail	Shapiro-Wilk's W	Probability	Pass/Fail
<i>Pimephales promelas</i>	32	Unbalanced	0.05	0.976	0.7285	Pass	0.979	0.8207	Pass
<i>Hyaella azteca</i>	70	Balanced	0.01	0.933	0.0013	Fail	0.965	0.1303	Pass
<i>Hyaella azteca</i>	28	Balanced	0.01	0.703	0.0001	Fail	0.879	0.0035	Fail

The data from the first *Hyaella azteca* bioassay include survival from two reference sites (DWR and SWR), one RM sediment, and seven LM sediments. Each sediment has seven replicates; thus, the design is balanced. Survival runs the full range from 0 to 10. Sediments LM 1-2, LM 2-3, and LM 11-12 could be considered to have an unusually broad range of survival. The untransformed data fail the normality test. However, arcsine-transformed data pass the normality test and may be considered approximately normally distributed.

Data from a second *Hyaella azteca* bioassay consist of four Michigan City Harbor (MC) sediments, of which the fourth is a reference. Again, the design is balanced, with each sediment having seven replicates. Survival in most replicates is high, with many replicates having full survival. Whether untransformed or arcsine-transformed, these data fail the normality test. Thus, the data would be converted to rankits or ranks before proceeding with statistical tests.

A look at the frequency histograms (Figure 18) for survival from the three bioassays suggests reasons for the results of the Shapiro-Wilk's Tests. Survival frequencies from the *Pimephales* bioassay, while not symmetrical, look as though they could easily be a random sample from a normal distribution. Survival frequencies from the first *Hyaella* bioassay appear to be bimodal, with clusters around 0 and 7 to 10, and almost nothing in between. Survival frequencies from the second *Hyaella* bioassay are concentrated at 9 and 10. The *Hyaella* histograms offer little to suggest samples from a normal distribution.

## Chemical Concentration Data in Environmental Samples

Bioaccumulation testing in Tiers III and IV produces concentration data for various contaminants of interest in the tissues of the tested organisms. The

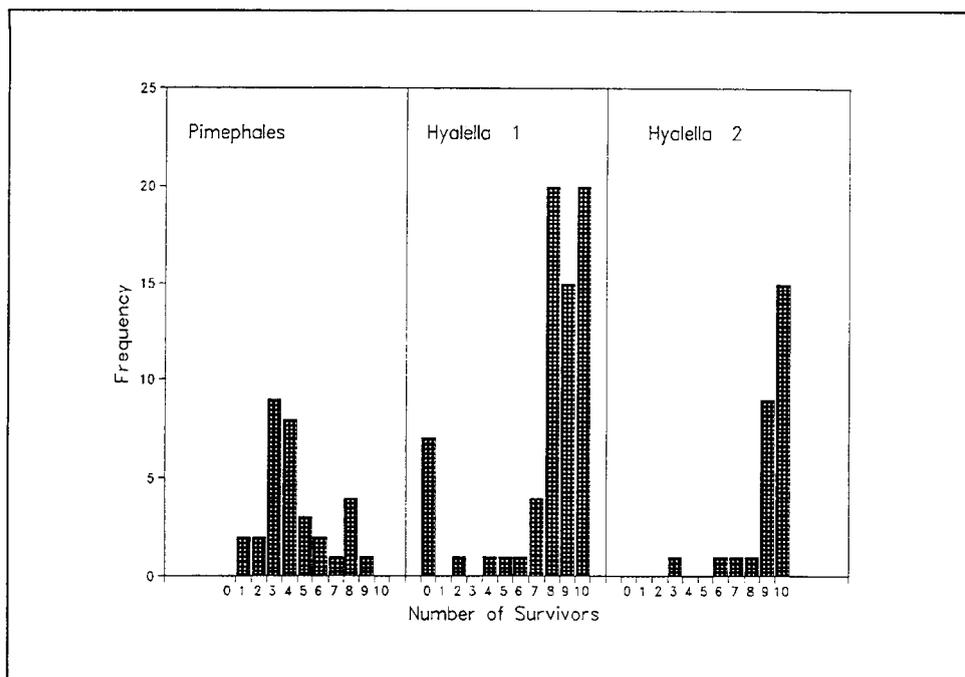


Figure 18. Frequency histograms for example bioassay survival data

true probability distribution for such data is unknown, but normal, lognormal, and gamma distributions are all reasonable candidates (Newman et al. 1989; White 1978). When the CV is low, say less than 0.25, these three distributions are virtually indistinguishable (Parsons 1969; Clarke, unpublished simulations). As the CV increases, differences among the distributions become more pronounced. The normal distribution is represented by the familiar, symmetrical, bell-shaped curve, which can include negative values. Gamma and lognormal distributions become positively skewed as the CV increases, with the bulk of observations near the lower end (left edge) of the frequency curve and no values below zero. Frequency curves for several distributions, including lognormal and gamma, are illustrated by Gilliom and Helsel (1986) at four CVs. The lognormal distribution is often assumed for environmental trace chemical data (El-Shaarawi 1989; Gilliom, Hirsch, and Gilroy 1984; Kushner 1976; Ott and Mage 1976; Porter and Ward 1991; Travis and Land 1990).

If contaminant concentration data can be reasonably assumed to follow a lognormal distribution, and such a distribution can easily be normalized using a log transformation, then why not simply log transform all bioaccumulation data and proceed with parametric statistical comparisons? The reason is that log transformation can produce an excessively high Type I error rate ( $\alpha$ ) with various distributions when variances differ among the treatments. Simulations comparing four treatments, two of which were drawn from populations with identical means, found Type I error rates of the LSD test as high as 0.72 when the standard deviations of the identical-mean populations differed. Furthermore, log transformation can have very low power in some situations,

especially when data are drawn from a gamma distribution and variances are unequal (see Figure 15 in Chapter 4). Therefore, using a test for normality with bioaccumulation data is recommended in an attempt to gain information about the underlying data distribution, and thus the appropriateness of using log-transformed data, untransformed data, or rankits for a particular set of comparisons.

### Performance of Shapiro-Wilk's Test

When samples are drawn from a normal distribution, they are expected to pass the test for normality. In fact, such samples do pass at a rate of  $1 - \alpha$ , i.e., the confidence level of the test. If  $\alpha$  is set at 0.10, then approximately 90 percent of samples from a normal distribution would pass the normality test, assuming the number of samples is large. When samples are drawn from a nonnormal distribution, one would like to see a high proportion of those samples fail the test for normality. If this happens, the normality test has high power.

Simulations were conducted to assess the performance of Shapiro-Wilk's Test with lognormal and gamma distributions at various CVs and total number of replicates ranging from 3 to 40 (Figures 19-22). Power is low for rejecting normality of untransformed lognormal and gamma samples when  $N$  is small and/or the CV is low. Because the distributions are so similar when  $CV = 0.1$ , one would expect most samples to pass the normality test. As the CV

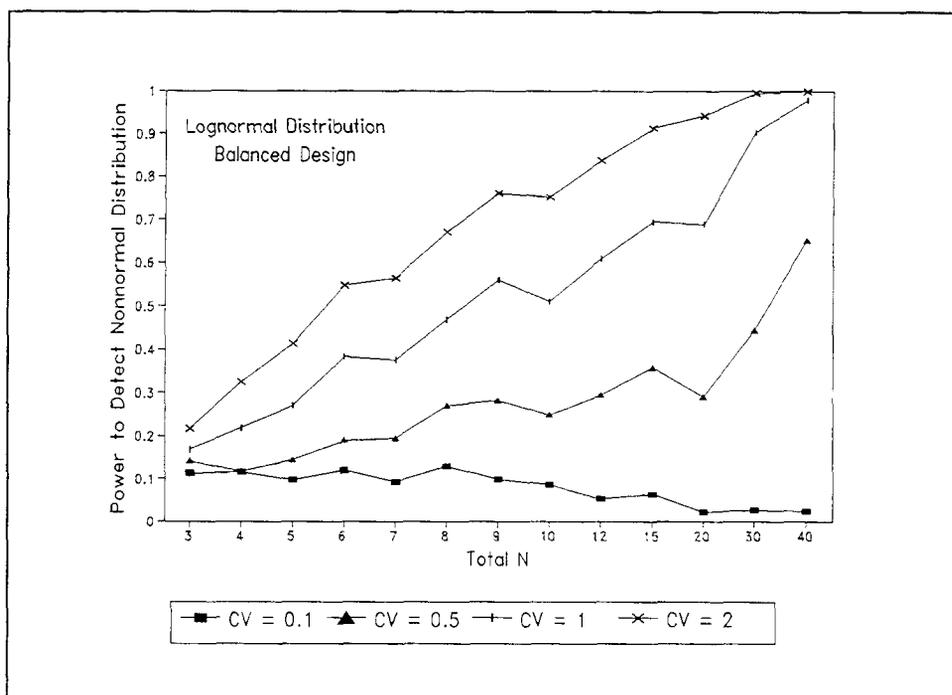


Figure 19. Power of Shapiro-Wilk's Test to detect nonnormality: lognormal distribution, balanced design

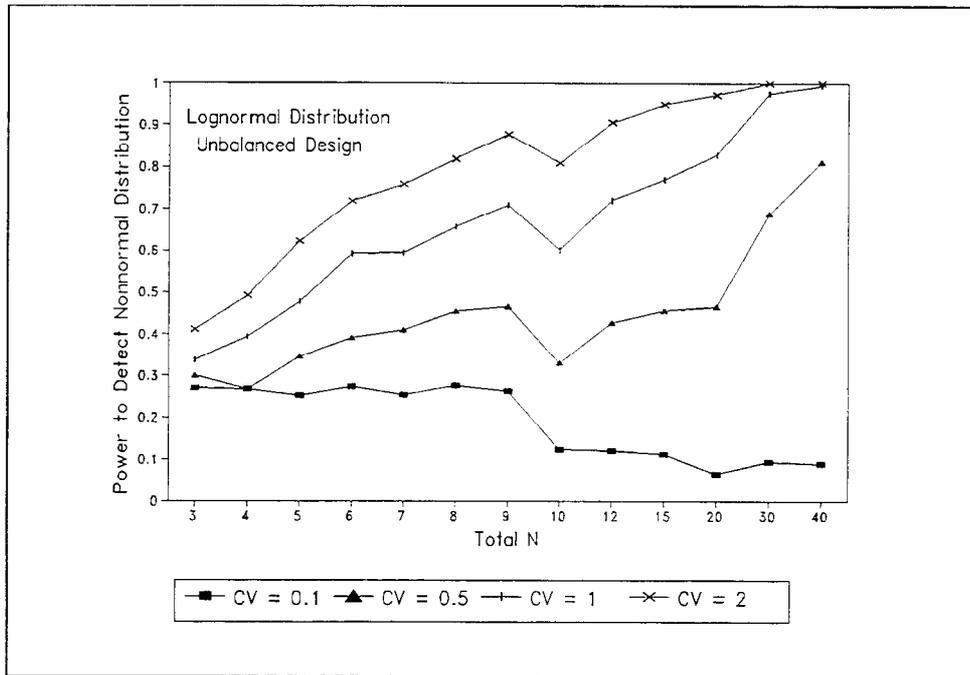


Figure 20. Power of Shapiro-Wilk's Test to detect nonnormality: lognormal distribution, unbalanced design

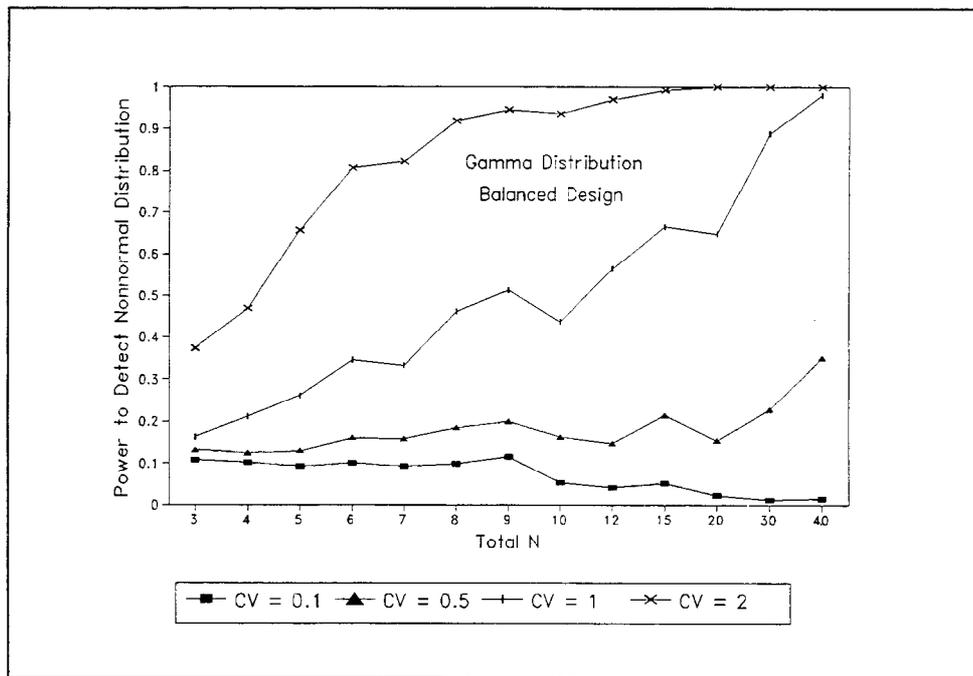


Figure 21. Power of Shapiro-Wilk's Test to detect nonnormality: gamma distribution, balanced design

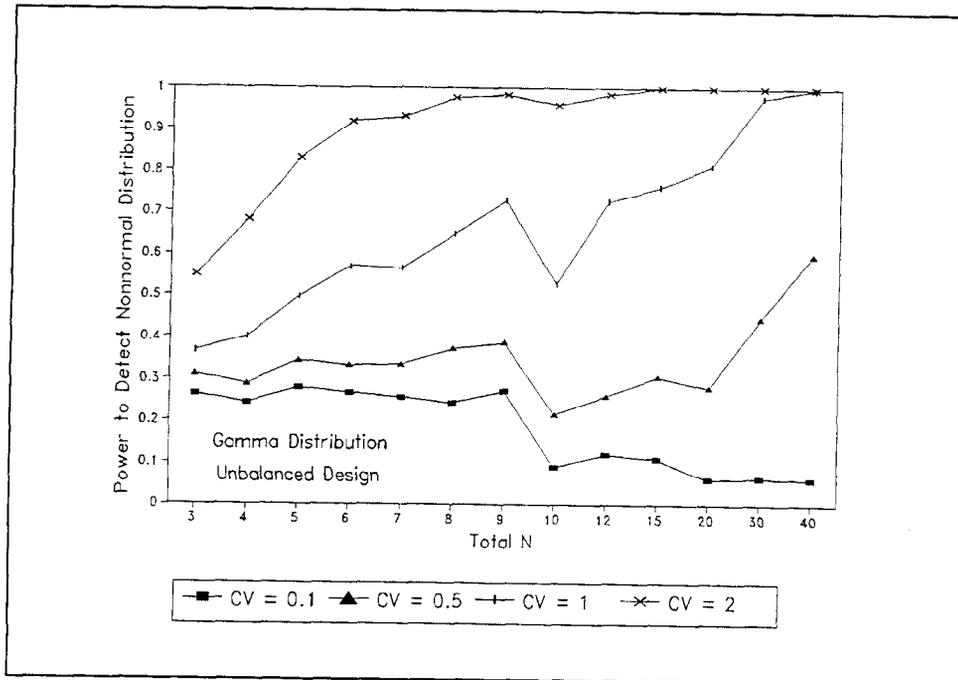


Figure 22. Power of Shapiro-Wilk's Test to detect nonnormality: gamma distribution, unbalanced design

increases, the distributions become more unique, and the power of Shapiro-Wilk's Test to reject normality increases. The test has high power to reject normality of samples from a gamma distribution when  $CV = 2$  and  $N$  exceeds 5, or when  $CV = 1$  and  $N = 30$  or more. Likewise, the test has high power to reject normality of samples from a lognormal distribution when  $CV = 2$  and  $N$  exceeds 10, or when  $CV = 1$  and  $N = 30$  or more. The pronounced "dip" at  $N = 10$ , especially for the unbalanced design (Figures 20 and 22) reflects the sharp change in  $\alpha$  from 0.25 to 0.10. There is a secondary dip for some CVs at  $N = 20$ , where  $\alpha$  changes again (refer to Table D-2 of the Inland Testing Manual). These dips suggest that the changes in  $\alpha$  should occur at  $N = 12$  rather than 10, and  $N = 25$  or 30 rather than 20.

Samples from a lognormal distribution are effectively normalized by a logarithmic transformation. Either natural log or  $\log_{10}$  may be used. The transformed samples will pass the normality test at a rate of  $1 - \alpha$ . Log transformation will not normalize samples from other distributions. Transformation of samples from any distribution to rankits (not ranks) imposes normality. Rankit-transformed samples will also pass the normality test at a rate of approximately  $1 - \alpha$ .

### Consequences of failure to correctly identify sample distribution

What are the consequences of an erroneous outcome of the normality test in terms of subsequent statistical comparisons? One might expect failure to

correctly identify samples as normal (Type I error) or nonnormal (Type II error) to result in increased error rates for statistical comparisons. Often, this does occur. Simulations were performed to assess the changes in power and  $\alpha$  for the LSD test when samples were incorrectly identified regarding distribution and the wrong data transformation was applied.

The effects of Types I and II errors in the normality test on power and  $\alpha$  of a subsequent LSD test are shown in Figures 23 (equal variances) and 24-26 (unequal variances). Type I error in Figures 23-26 relates to the maximum change in power and  $\alpha$  of the LSD test when either a log or a rankit transformation was applied to samples from a normal distribution, compared with the untransformed data. Type II error relates to the change in power and  $\alpha$  when untransformed samples from either a lognormal or a gamma distribution were used in the LSD test, compared with the appropriately transformed data. When variances are equal (Figure 23), failure to identify normal samples (Type I error) results in decreasing power and increasing  $\alpha$  in the LSD test as CV increases. Failure to identify samples as nonnormal (Type II error) results in decreased power and  $\alpha$  in the LSD test as CV increases. Clearly, if the CV of the combined samples is very low, it makes little difference which transformation is applied.

When variances among treatments are unequal, the consequences of error resulting from the normality test are more complicated (Figures 24-26). Simulations performed to assess those consequences for the LSD test

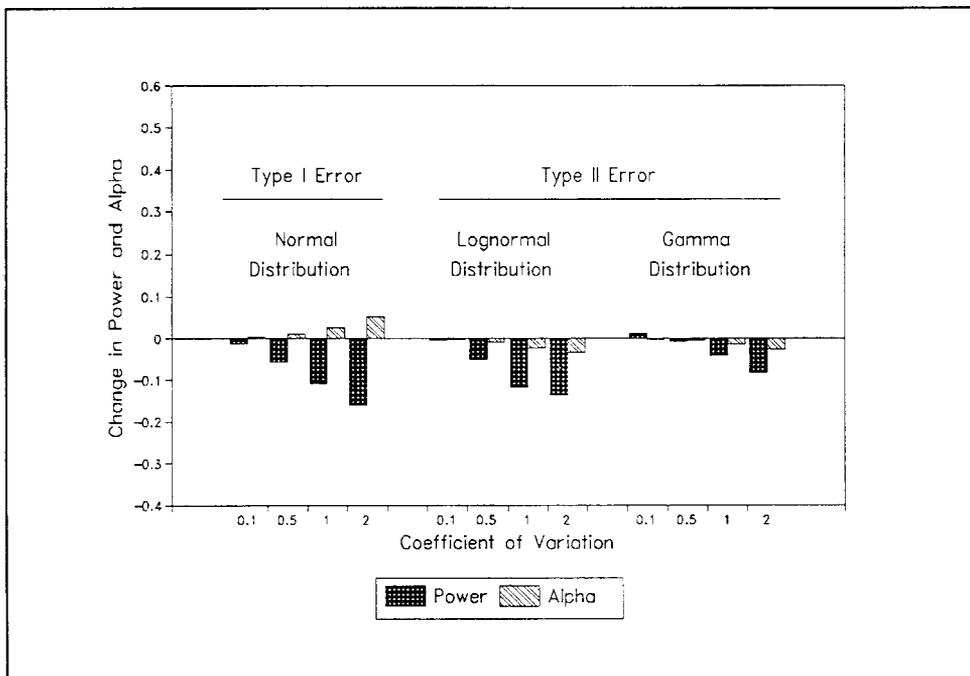


Figure 23. Change in power and  $\alpha$  of LSD test resulting from incorrect identification of parent distribution and application of wrong transformation. Variances equal among treatments

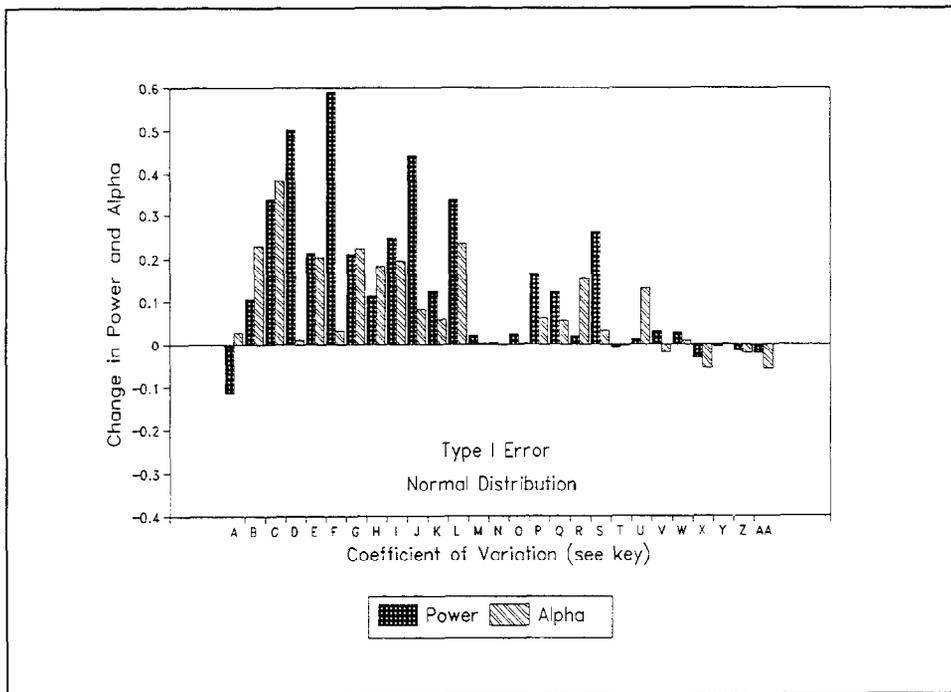


Figure 24. Change in power and  $\alpha$  of LSD test resulting from failure to identify normal distribution and application of wrong transformation. Variances unequal (see Table 2 for key to CVs)

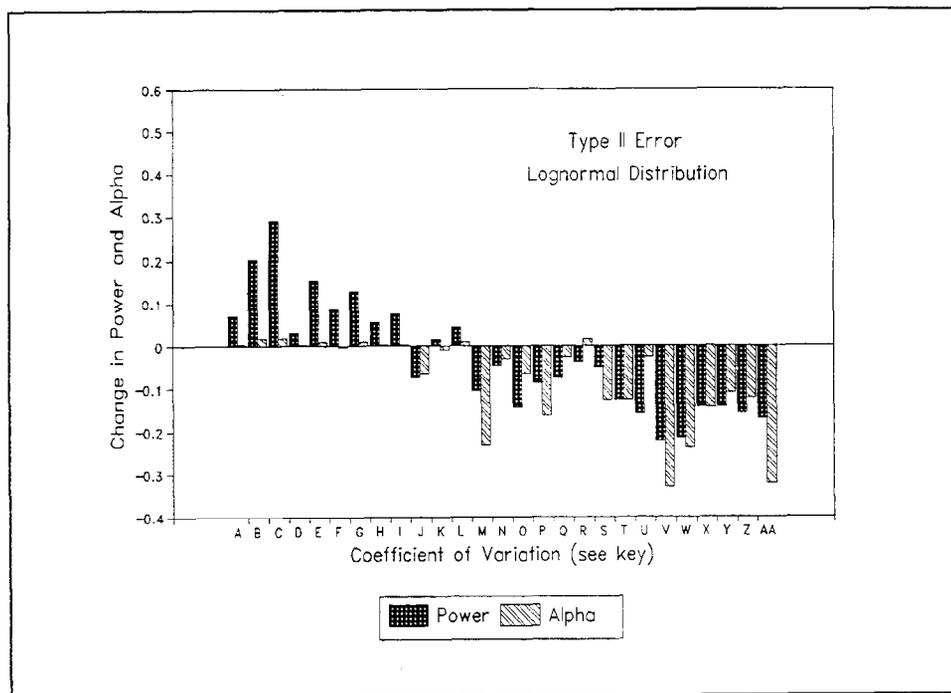


Figure 25. Change in power and  $\alpha$  of LSD test resulting from use of untransformed lognormal data. Variances unequal among treatments (see Table 2 for key to CVs)

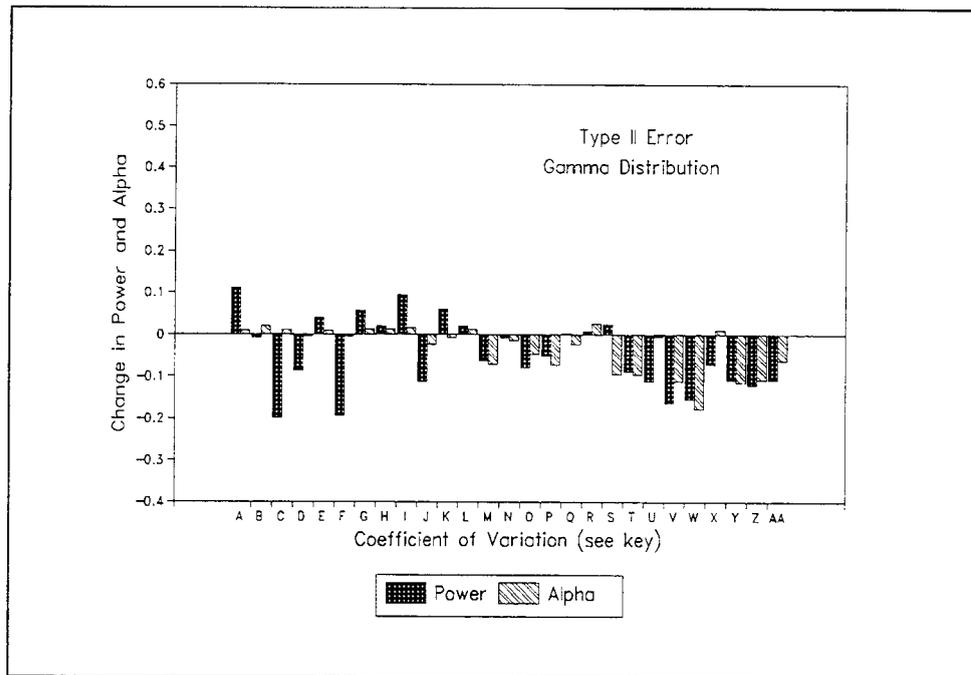


Figure 26. Change in power and  $\alpha$  of LSD test resulting from use of untransformed data from gamma distribution. Variances unequal among treatments (see Table 2 for key to CVs)

compared four sediments having the same or different population means and different population CVs (see Table 2 in Chapter 4). Failure to identify normal samples, and consequent application of a transformation (usually log), often results in increased power when the reference sediment CV is low to moderate, but at the expense of sometimes greatly inflated  $\alpha$  (Figure 24). Recall that  $\alpha$  for the LSD test is expected to be 0.05, but use of log transformation with samples from a normal distribution can raise  $\alpha$  to the range of 0.2 to 0.5 in some cases when variances are unequal. Changes in power and  $\alpha$  become less pronounced, and generally negative, when the reference sediment CV is high.

When variances are unequal and samples from a lognormal distribution pass the normality test, resulting use of untransformed data in the LSD test has an impact on power and  $\alpha$  compared with log-transformed data (Figure 25). Power increases, but there is little change in  $\alpha$  when the reference sediment CV  $\leq 0.5$ . At higher reference sediment CVs, there tends to be a fairly large decrease in both power and  $\alpha$ . Use of untransformed data from a gamma distribution, compared with rankits, also tends to depress both power and  $\alpha$  somewhat in the LSD test when the reference sediment CV is high (Figure 26). At lower reference sediment CVs, there is little effect on  $\alpha$  and mixed effect on power.

It is clear that tests for normality are limited in their ability to correctly identify samples as normal or nonnormal, especially when the total number of

replicates  $N$  is small. When several dredged sediments can be simultaneously compared with a reference, this will increase  $N$  and thus improve the performance of the normality test. Is it better to err in failing too many normal samples (Type I error) or in passing too many nonnormal samples (Type II error) in terms of subsequent effect on statistical comparisons? If high power of the statistical comparison is equated with environmental protection and high  $\alpha$  with unnecessary cost (Table 1 in Chapter 4), then Type I error in the normality test will tend to increase both environmental protection and unnecessary expense, especially in the common situation where variances are unequal and the reference sediment CV is low to moderate. Type II error in the normality test may increase power without increasing unnecessary expense if variances are unequal and the reference sediment CV is low. Type II error will also tend to decrease both environmental protection and unnecessary cost when the reference sediment CV is high. Types I and II errors are determined by the significance level ( $\alpha$ ) of the normality test. Increasing  $\alpha$  will increase Type I error regardless of  $N$ , while lowering  $\alpha$  will increase Type II error if  $N$  is held constant.

### When not to assume normality

Chemical concentration data, such as that resulting from bioaccumulation tests, will at best approximate a known statistical distribution. Reasonable choices are the normal, lognormal, and gamma distributions. The latter two are probably most applicable to concentration data because they cannot include negative values. However, the normal distribution is the basis for most standard statistical procedures, and thus one generally tries to fit samples to a normal distribution whenever possible. Nevertheless, in some situations, the normal distribution is highly improbable for chemical concentration data because such a distribution would include negative values. Assuming  $\mu > 0$ , the proportion of negative values in a normal distribution increases asymptotically to 50 percent as the CV increases to infinity. This relationship is shown in Figure 27 for CVs from 0.1 to 3. A normal distribution will include  $\approx 2$  percent negative values when  $CV = 0.5$ ,  $\approx 17$  percent negative values when  $CV = 1$ , and  $\approx 31$  percent negative values when  $CV = 2$ . Thus, a combined sample CV of 1 or greater could be considered strongly suggestive of nonnormality. Often, a  $CV > 1$  will be the result of one or a few outliers (see Chapter 7).

Statisticians sometimes argue that testing for normality when  $N$  is small is of little value, and nonparametric statistical comparison procedures should be used routinely instead. Simulation results (Figures 28 and 29) suggest that there is little to be lost, and sometimes much to be gained, in following this recommendation for comparisons of chemical concentration data. However, the simulations were based on  $N$  ranging from 18 to 32; in this range, the Shapiro-Wilk's Test has reasonably high power when the CV is high (see Figures 19-22). It is unknown whether the results of the simulations would be similar for small  $N$ , when the Shapiro-Wilk's Test has low power.

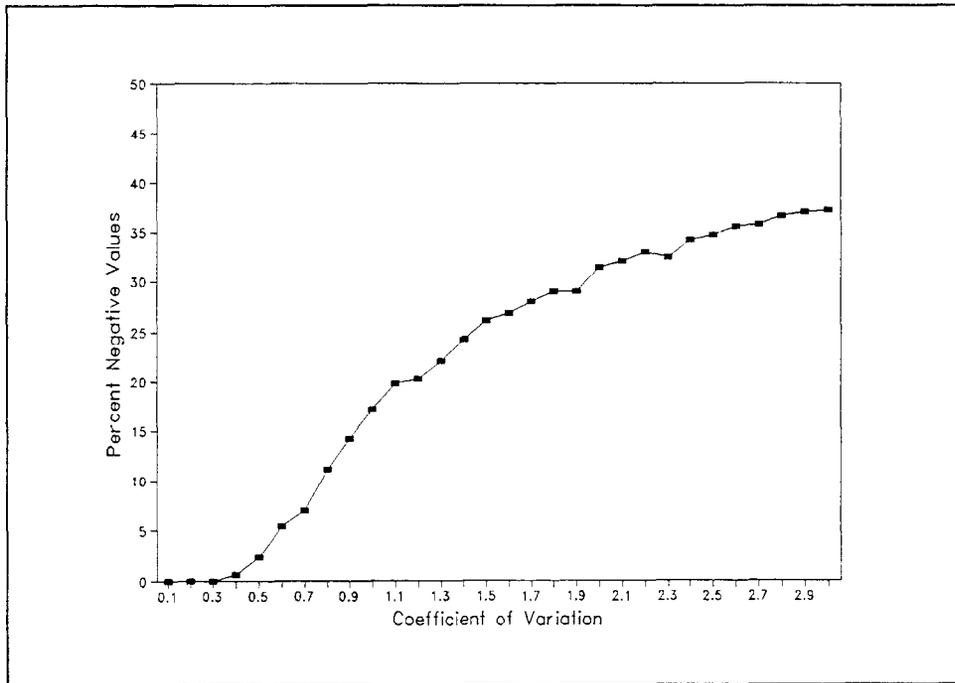


Figure 27. Approximate percent negative values in a normal distribution (based on simulations of  $n = 2,500$ )

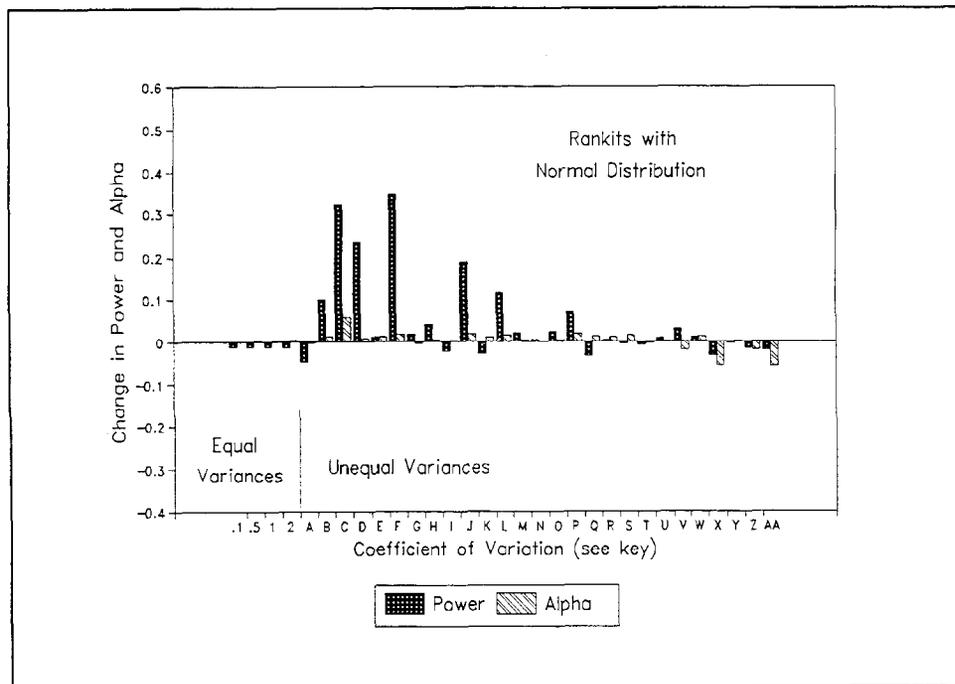


Figure 28. Change in power and  $\alpha$  of LSD test resulting from use of rankits, rather than untransformed data, for samples from a normal distribution (see Table 2 for key to CVs)

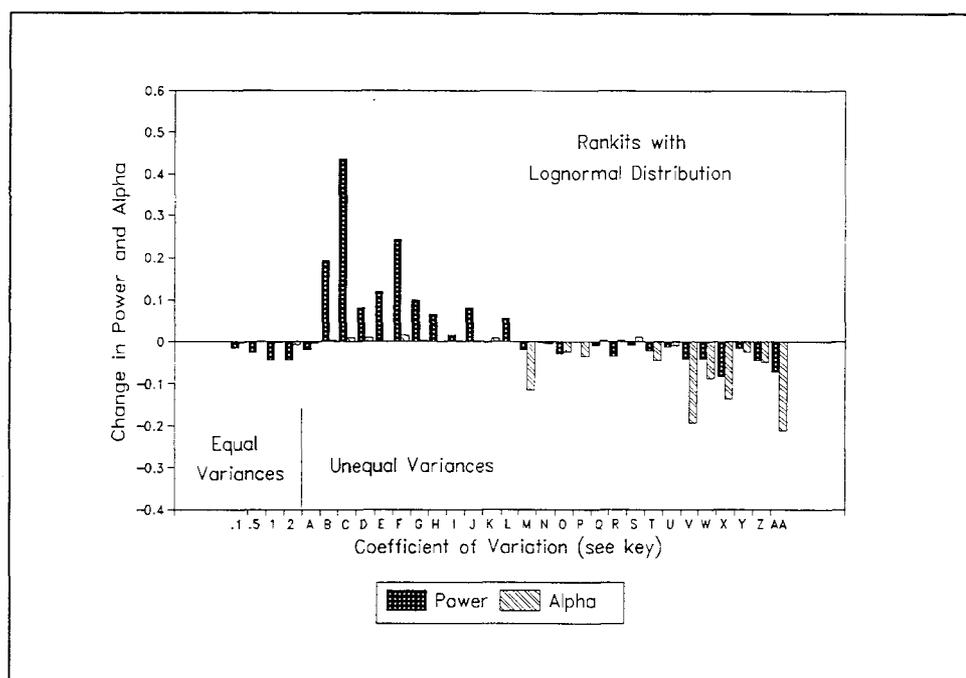


Figure 29. Change in power and  $\alpha$  of LSD test resulting from use of rankits, rather than log-transformed data, for samples from a lognormal distribution (see Table 2 for key to CVs)

Figure 28 illustrates the change in power and  $\alpha$  of an LSD test when rankits rather than untransformed data are used with samples from a normal distribution. When variances are equal, rankits result in a slight loss of power and essentially no change in  $\alpha$  compared with untransformed data. When variances are unequal, rankits sometimes result in a substantial increase in power, with little change in  $\alpha$ , especially when the reference sediment CV is fairly low. Figure 29 illustrates the change in power and  $\alpha$  of an LSD test when rankits rather than log-transformed data are used with samples from a lognormal distribution. Again, when variances are equal, rankits result in a slight loss of power and essentially no change in  $\alpha$  compared with log-transformed data. When variances are unequal, rankits generally result in substantial gain in power, with little change in  $\alpha$ , when the reference sediment CV < 0.8. At higher reference sediment CVs, there tends to be a slight power loss and sometimes large drop in  $\alpha$  (although  $\alpha$  still greatly exceeds 0.05 in most cases) when rankits are used instead of log-transformed data.

### Example data

Metals bioaccumulation data from animals exposed in the laboratory to four sediments from the New York Bight area are used to illustrate varying results of the Shapiro-Wilk's Test for normality (Table 7). Three species (*Macoma nasuta*, *Macoma secta*, and *Nereis virens*) were exposed to three dredged sediments (AK, GOW, and RH) and a reference sediment (SH). Data

<b>Table 7</b>					
<b>Bioaccumulation of Selected Metals (<math>\mu\text{g/g}</math>) in Organisms Exposed to New York Bight Sediments</b>					
Sediment	Cadmium	Cadmium	Mercury	Lead	Zinc
	<i>Macoma nasuta</i>	<i>Nereis virens</i>	<i>Nereis virens</i>	<i>Nereis virens</i>	<i>Macoma secta</i>
AK	0.110	0.031	<0.020	<0.022	28.4
	0.230	0.041	<0.020	0.142	
	0.144	0.039	<0.020	0.171	
	<0.004	0.037	<0.020	0.186	
	<0.004	0.038	<0.020	0.126	
	0.040	0.032	<0.020	0.100	
GOW	0.150	0.037	0.029	0.243	8.01 43.4
	<0.002	0.020	<0.020	0.076	
	0.170	0.035	<0.020	0.039	
	0.170	0.030	0.024	0.112	
	0.090	0.035	<0.020	0.259	
	0.090	0.044	<0.020	0.397	
RH	<0.002	0.034	0.043	0.973	24.3 11.7
	0.110	0.045	0.038	0.066	
	0.140	0.028	<0.020	0.081	
	0.180	0.258	<0.020	0.129	
	<0.004	0.027	<0.020	0.225	
	0.020	0.029	<0.020	0.330	
SH	0.220	0.019	0.029	0.046	36.6 61.4 63.8 19.7 47.6 17.3
	<0.005	0.037	<0.020	0.086	
	0.090	0.023	0.024	0.096	
	0.090	0.036	0.022	0.112	
	0.090	0.020	<0.020	0.115	
	0.020	0.039	<0.020	0.391	
Important features of the data set	Several nondetects CV = 0.82	1 outlier CV = 1.10	Many nondetects CV = 0.64	Mostly low values fewer high values 1 possible outlier CV = 1.05	Unbalanced design CV = 0.58

sets were selected to exemplify some of the factors that can influence normality test results. Shapiro-Wilk's Test results on the residuals of the untransformed data, log-transformed data, and rankits are presented in Table 8.

The first four data sets are balanced designs with six replicates of each of the four sediments. From Table D-2 of the Inland Testing Manual,  $\alpha = 0.01$  for the normality test when the design is balanced and  $N \geq 20$ . The first data set, cadmium bioaccumulation in *Macoma nasuta*, includes six nondetects that were set equal to half the detection limit for this illustration.<sup>1</sup> These data pass the Shapiro-Wilk's Test when untransformed or converted to rankits, but fail when log-transformed. The CV = 0.82, which is high enough that a normal distribution may be unlikely and that rankits may be more appropriate than untransformed data for use in statistical comparisons.

<sup>1</sup> See Chapter 8 for a comprehensive discussion of statistical treatment of less-than detection limit data.

Table 8 Results of Shapiro-Wilk's Test for Normality on Example Bioaccumulation Data							
Metal and Species	Total No. of Replicates	Design	$\alpha$ for Shapiro-Wilk's Test	Transformation	Shapiro-Wilk's $W$	Probability	Pass/Fail
Cadmium <i>Macoma nasuta</i>	24	Balanced	0.01	None	0.948	0.2534	Pass
				Log	0.851	0.0018	Fail
				Rankit	0.985	0.9597	Pass
Cadmium <i>Nereis virens</i>	24	Balanced	0.01	None	0.556	0.0001	Fail
				Log	0.773	0.0001	Fail
				Rankit	0.962	0.4941	Pass
Mercury <i>Nereis virens</i>	24	Balanced	0.01	None	0.891	0.0132	Pass
				Log	0.858	0.0025	Fail
				Rankit	0.872	0.0048	Fail
Lead <i>Nereis virens</i>	24	Balanced	0.01	None	0.807	0.0002	Fail
				Log	0.982	0.9199	Pass
				Rankit	0.966	0.5753	Pass
Zinc <i>Macoma secta</i>	11	Unbalanced	0.10	None	0.941	0.5055	Pass
				Log	0.946	0.5720	Pass
				Rankit	0.954	0.6787	Pass

The second data set, cadmium bioaccumulation in *Nereis virens*, fails the normality test unless the data are converted to rankits. The most likely reason is the presence of an obvious outlier, 0.258, among the RH replicates. Because of the outlier, the CV is a high 1.10. Without the outlier, the CV drops to 0.23, and the data pass the normality test regardless of whether untransformed data, log-transformed data, or rankits are used.

The third data set, mercury bioaccumulation in *Nereis virens*, contains many nondetects, including one treatment (AK) for which the data are entirely below detection limit. Again, nondetects were set equal to half the detection limit. The untransformed data pass the normality test, although marginally ( $P = 0.0132$ ). The log-transformed data and rankits both fail the normality test.  $CV = 0.64$ .

The fourth data set, lead bioaccumulation in *Nereis virens*, contains mostly low values including one nondetect, and fewer high values including one possible outlier (0.973 in RH). This arrangement of values and the high CV ( $= 1.05$ ) suggest a possible lognormal distribution. Indeed, the log-transformed data pass the normality test while the untransformed data fail. Rankits also pass the normality test.

The fifth data set, zinc bioaccumulation in *Macoma secta*, is a highly unbalanced design in which there is only one usable replicate for AK and two

replicates each for GOW and RH. From Table D-2 of the Inland Testing Manual,  $\alpha = 0.10$  for the normality test when the design is unbalanced and  $N = 10$  to 19. Despite the unbalanced design and relatively high  $\alpha$ , the data easily pass the normality test whether untransformed, log-transformed, or converted to rankits. The CV for these data = 0.58.

## Summary

Normality is an important assumption of parametric statistical comparison procedures such as the  $t$ -test and LSD test. Testing for normality, however, has limited power, especially when total number of replicates is small. Nevertheless, testing for normality can provide clues to the underlying data distribution, and therefore, to which data transformation, if any, should be used.

Common knowledge and the simulation results described in this chapter support the following conclusions:

- Survival data are binomially distributed. When survival proportions are neither very high nor very low, the binomial distribution approximates a normal distribution.
- The Shapiro-Wilk's Test will likely reject normality when survival proportions are very high or very low. In such cases, rankits should be used for statistical comparisons instead of the arcsine-square root transformed data.
- Contaminant concentration data in environmental samples may approximate a normal, lognormal, or gamma distribution. Data from a lognormal distribution can be normalized using log transformation, and data from either lognormal or gamma distributions can be normalized using rankits.
- Failure to identify samples from a normal distribution and consequent application of log transformation in the LSD test can result in excessively high Type I error rate ( $\alpha$ ) when variances are unequal.
- Failure to identify samples from nonnormal distributions and consequent use of untransformed data in the LSD test can result in decreased power when variances are unequal.
- Normality of contaminant concentration sample data is unlikely when the coefficient of variation is high ( $\geq 1$ ), as such a distribution would include a fair proportion of negative values.
- Routine use of rankit-based comparisons (conversion of data to rankits followed by  $t$ -test or LSD test) is acceptable and can even increase power of the LSD test in some situations.

Preparation of data for statistical comparisons involves several aspects that are interdependent and should be carried out concurrently. These include the following:

- Calculation of the coefficient of variation for all treatments combined.
- Testing for normality of the treatment residuals of the combined replicates using untransformed data, log-transformed data, and rankits (arcsine-transformed data and rankits for survival proportions).
- Testing for equality of variances among treatments (Chapter 6) using untransformed data, log-transformed data, and rankits (arcsine-transformed data and rankits for survival proportions).
- Determination of the most appropriate method for handling less-than detection limit data (Chapter 8), if any, based on the CV, assumed data distribution, and characteristics of the treatment variances.

Because these aspects are interrelated, iterative determinations may be necessary, especially when less-than detection limit data are involved.

## 6 Tests of Assumptions: Equality of Variances

---

Statistical comparison procedures, whether parametric or nonparametric, generally assume that variances are the same or nearly the same among the treatments being compared. Recall from Chapter 4 that unequal variances can affect both the power and the Type I error rate of statistical comparisons. This chapter will examine the effect of variances on the power of a  $t$ -test and compare several tests for equality of variances using example survival and bioaccumulation data. Levene's Test is the most versatile of the tests for equality of variance mentioned in the Inland Testing Manual, and its performance will be explored using simulated data.

Testing for equality of variances in dredged sediment evaluations follows normality testing in the decision trees of the Inland Testing Manual Appendix D. Note that nonnormality, not unequal variances, leads to nonparametric procedures. When variances are unequal in a two-group comparison, the  $t$ -test for unequal variances is used. When variances are unequal in the comparison of several groups, each dredged sediment is compared separately with the reference sediment using a  $t$ -test for equal or unequal variances as appropriate.

Several tests for equality of variances are recommended in the Inland Testing Manual, although many additional tests are known (Conover, Johnson, and Johnson 1981). However, only the recommended tests, if any, are likely to be available in statistical software packages. SAS provides none of them, except for the  $F'$  test as part of the TTEST procedure for comparison of two groups (SAS Institute, Inc. 1988a). Levene's Test, Bartlett's Test, Cochran's Test, and Hartley's  $F_{\max}$  can easily be programmed in SAS, but the probability of the test statistic can only be calculated for the first two tests. The significance of Cochran's and Hartley's test statistics must be determined using specialized tables available in a few statistics texts.  $\alpha$  levels for equality of variance tests, given in Table D-2 of the Inland Testing Manual, depend on number of replicates per treatment ( $n$ ) and on whether the design is balanced or unbalanced. SAS statements for Levene's Test are included in the programs in Appendix D of the Inland Testing Manual; SAS statements for all four equality of variance tests mentioned above are given in program

EQOFVAR.SAS in Appendix A of this document. Problems inherent in these tests will be discussed in detail later in this chapter.

## Variations and Power

Unequal variances can have a profound and adverse effect upon  $\alpha$  of the LSD test, especially when the reference sediment CV is high, as seen in Chapter 4 (Figures 7-10). The influence of the sample variances, whether equal or unequal, on the power of a  $t$ -test or LSD comparison can be determined using Equation 10 of the Inland Testing Manual Appendix D. Figure 30 shows the power of a  $t$ -test to detect a difference of one unit between two group means when  $\alpha = 0.05$ ,  $n = 5$ , and the group CVs range between 0.05 and 2.0. Power is highest when both group CVs are low, and lowest when one or both group CVs are high. Increasing the variances proportionally increases the effect size that can be detected as significant at a given power. Thus, any measures that can be taken in the laboratory setting to minimize the variability of the test data will work to increase the power of whatever statistical comparisons may be necessary.

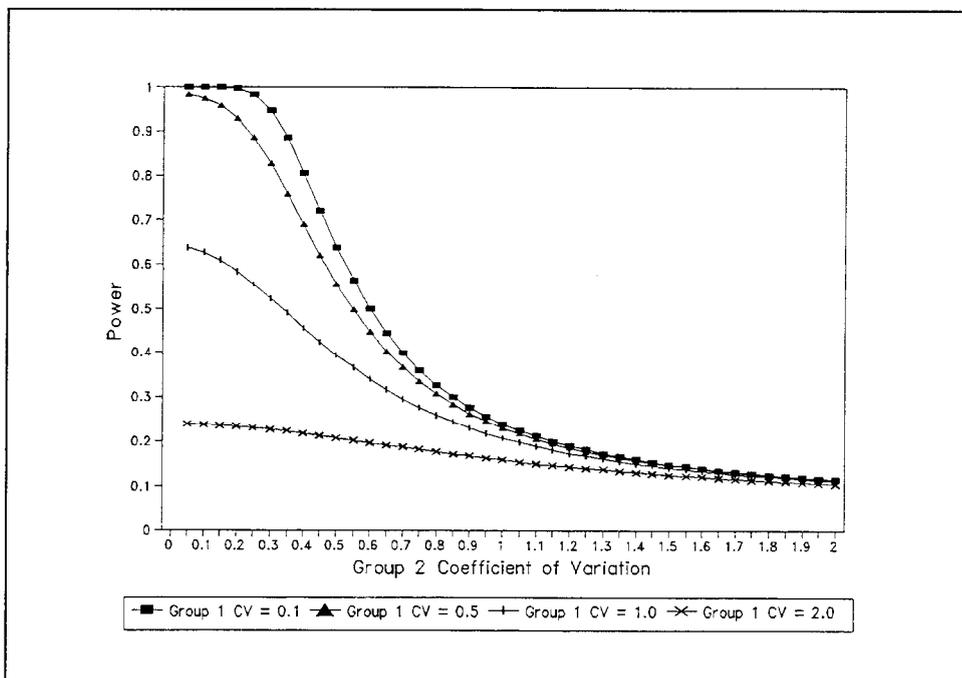


Figure 30. Power of a  $t$ -test to detect a difference in means of one unit when  $\alpha = 0.05$  and  $n = 5$

## Equality of Variance Tests on Example Data

Survival data from water column and benthic toxicity tests are checked for equality of variances according to decision tree steps in Figures D-1 and D-4 of the Inland Testing Manual. Likewise, contaminant concentration data from sediment comparisons or from bioaccumulation tests are checked for equality of variances according to decision tree steps in Figure D-5 of the Inland Testing Manual. If only two treatments are being compared (e.g., dilution water and 100-percent elutriate), then the  $F'$  test is the simplest test for equality of variances and is provided automatically in SAS PROC TTEST (SAS Institute, Inc. 1988a). When more than two treatments are being compared, Levene's Test should be performed if possible. Problems with Bartlett's Test, Hartley's  $F_{\max}$ , and Cochran's Test will become apparent from the analyses of example data.

Equality of variance tests are illustrated using the survival data of Table 5 and the bioaccumulation data of Table 7 from Chapter 5. The tests are performed on the untransformed survival proportions and contaminant concentrations, on the transformed data (arcsine transformation for survival data and log transformation for bioaccumulation data), and on the rankits. In actual practice, the use of untransformed survival data is not recommended. Results of the four equality of variance tests are given in Tables 9-12 and summarized in Table 13.

### Levene's Test

Levene's Test evaluates equality of variances by conducting an analysis of variance on the absolute deviations of treatment observations from the treatment means (Brown and Forsythe 1974; Keppel 1991; Milliken and Johnson 1984; Snedecor and Cochran 1989). A variation of Levene's Test, sometimes called the Brown-Forsythe Test, uses absolute deviations from the treatment medians rather than means. Simulations comparing equality of variance tests (Conover, Johnson, and Johnson 1981) showed that Levene's Test can have a high Type I error rate with some nonnormal (asymmetric) distributions. The Brown-Forsythe variation avoids that problem but at the price of much lower power than Levene's Test. Milliken and Johnson (1984) recommended using Levene's Test rather than Bartlett's or Hartley's tests especially with nonnormal distributions, and using the Brown-Forsythe variation when the data tend to be very skewed. On tests of the example data, Levene's Test found significantly unequal variances in many instances where the other equality of variance tests did not (Table 13). The Brown-Forsythe variation using medians, however, found significant inequality of variances in less than half as many instances as Levene's Test. The Brown-Forsythe variation should not be used in dredged sediment evaluations because of its low power. Note that Levene's Test (and the Brown-Forsythe variation) was able to make a determination on equality of variances in all of the examples, whereas the other three tests were not able to do so in all cases.

**Table 9**  
**Results of Levene's Test on Example Data (Results in parentheses are for Brown-Forsythe variation using medians)**

Data	Transformation	$\alpha$	Test Statistic	Degrees of freedom, $\nu$	Probability	Pass/Fail
Survival <i>Pimephales promelas</i>	Arcsine	0.25	1.777 (1.508)	8, 23	0.1338 (0.2088)	Fail (Fail)
	None	0.25	1.724 (1.549)	8, 23	0.1461 (0.1952)	Fail (Fail)
	Rankit	0.25	1.562 (1.158)	8, 23	0.1910 (0.3649)	Fail (Pass)
Survival <i>Hyaella azteca</i> (1)	Arcsine	0.10	5.304 (1.502)	9, 60	0.0000 (0.1681)	Fail (Pass)
	None	0.10	8.326 (2.646)	9, 60	0.0000 (0.0119)	Fail (Fail)
	Rankit	0.10	0.262 (0.128)	9, 60	0.9823 (0.9988)	Pass (Pass)
Survival <i>Hyaella azteca</i> (2)	Arcsine	0.10	1.257 (0.453)	3, 24	0.3115 (0.7175)	Pass (Pass)
	None	0.10	1.551 (0.531)	3, 24	0.2272 (0.6656)	Pass (Pass)
	Rankit	0.10	1.227 (0.337)	3, 24	0.3215 (0.7988)	Pass (Pass)
Cadmium <i>Macoma nasuta</i>	Log	0.10	0.526 (0.503)	3, 20	0.6698 (0.6846)	Pass (Pass)
	None	0.10	0.535 (0.529)	3, 20	0.6635 (0.6678)	Pass (Pass)
	Rankit	0.10	0.639 (0.714)	3, 20	0.5986 (0.5553)	Pass (Pass)
Cadmium <i>Nereis virens</i>	Log	0.10	2.706 (1.029)	3, 20	0.0727 (0.4008)	Fail (Pass)
	None	0.10	5.082 (0.996)	3, 20	0.0089 (0.4151)	Fail (Pass)
	Rankit	0.10	2.361 (1.003)	3, 20	0.1019 (0.4119)	Pass (Pass)
Mercury <i>Nereis virens</i>	Log	0.10	17.052 (1.435)	3, 20	0.0001 (0.2623)	Fail (Pass)
	None	0.10	16.112 (1.336)	3, 20	0.0001 (0.2909)	Fail (Pass)
	Rankit	0.10	15.166 (1.357)	3, 20	0.0001 (0.2845)	Fail (Pass)
Lead <i>Nereis virens</i>	Log	0.10	0.472 (0.349)	3, 20	0.7055 (0.7899)	Pass (Pass)
	None	0.10	2.470 (1.267)	3, 20	0.0915 (0.3127)	Fail (Pass)
	Rankit	0.10	0.682 (0.657)	3, 20	0.5734 (0.5882)	Pass (Pass)
Zinc <i>Macoma secta</i>	Log	0.25	3.393 (2.536)	3, 7	0.0830 (0.1401)	Fail (Fail)
	None	0.25	2.346 (2.306)	3, 7	0.1591 (0.1634)	Fail (Fail)
	Rankit	0.25	2.289 (2.206)	3, 7	0.1655 (0.1752)	Fail (Fail)

### Bartlett's Test

Bartlett's Test, one of the most widely used tests for equality of variances, can be applied even when sample sizes are unequal and the design is highly unbalanced. However, most statistics texts do not recommend Bartlett's Test except when the distribution is known to be normal, as this test is quite

**Table 10**  
**Results of Bartlett's Test on Example Data**

Data	Transformation	$\alpha$	Test Statistic	Degrees of freedom, $\nu$	Probability	Pass/Fail
Survival <i>Pimephales promelas</i>	Arcsine	0.25	9.033	8	0.3395	Pass
	None	0.25	10.004	8	0.2648	Pass
	Rankit	0.25	6.134	8	0.6323	Pass
Survival <i>Hyaella azteca</i> (1)	Arcsine	0.10	22.269	9	0.0081	Fail
	None	0.10	43.173	9	0.0000	Fail
	Rankit	0.10	2.206	9	0.9878	Pass
Survival <i>Hyaella azteca</i> (2)	Arcsine	0.10	4.446	3	0.2172	Pass
	None	0.10	13.288	3	0.0041	Fail
	Rankit	0.10	1.896	3	0.5943	Pass
Cadmium <i>Macoma nasuta</i>	Log	0.10	0.581	3	0.9009	Pass
	None	0.10	0.477	3	0.9239	Pass
	Rankit	0.10	0.714	3	0.8700	Pass
Cadmium <i>Nereis virens</i>	Log	0.10	17.319	3	0.0006	Fail
	None	0.10	47.528	3	0.0000	Fail
	Rankit	0.10	2.728	3	0.4355	Pass
Mercury <i>Nereis virens</i>	Log	0.10	-17.098	3	--	--
	None	0.10	-44.031	3	--	--
	Rankit	0.10	-5.136	3	--	--
Lead <i>Nereis virens</i>	Log	0.10	0.893	3	0.8272	Pass
	None	0.10	13.168	3	0.0043	Fail
	Rankit	0.10	0.683	3	0.8772	Pass
Zinc <i>Macoma secta</i>	Log	0.25	1.136	3	0.7684	Pass
	None	0.25	0.757	3	0.8597	Pass
	Rankit	0.25	0.697	3	0.8739	Pass

sensitive to nonnormality (Keppel 1991; Milliken and Johnson 1984; Sokal and Rohlf 1981; Winer 1971). This means that Bartlett's Test has a high Type I error rate with skewed distributions and could be detecting nonnormality rather than inequality of variances. Procedures for calculating Bartlett's Test are given in Box 13.1 of Sokal and Rohlf (1981:404), and SAS statements are provided in program EQOFVAR.SAS in Appendix A of this report. Bartlett's Test statistic is distributed approximately as  $\chi^2$ , and its significance

**Table 11**  
**Results of Hartley's  $F_{\max}$  Test on Example Data**

Data	Transformation	$\alpha$	Test Statistic	Number of Treatments	Degrees of freedom, $\nu^1$	Probability	Pass/Fail
Survival <i>Pimep- hales promelas</i>	Arcsine	0.25	14.123	9	2	> 0.25	Pass
	None	0.25	18.667	9	2	> 0.25	Pass
	Rankit	0.25	6.697	9	2	> 0.25	Pass
Survival <i>Hyalella azteca</i> (1)	Arcsine	0.10	9.546	10	6	> 0.10	Pass
	None	0.10	28.385	10	6	< 0.05	Fail
	Rankit	0.10	2.796	10	6	> 0.25	Pass
Survival <i>Hyalella azteca</i> (2)	Arcsine	0.10	5.462	4	6	> 0.25	Pass
	None	0.10	27.600	4	6	< 0.01	Fail
	Rankit	0.10	2.999	4	6	> 0.25	Pass
Cadmium <i>Macoma nasuta</i>	Log	0.10	1.985	4	5	> 0.25	Pass
	None	0.10	1.905	4	5	> 0.25	Pass
	Rankit	0.10	2.130	4	5	> 0.25	Pass
Cadmium <i>Nereis virens</i>	Log	0.10	60.304	4	5	< 0.01	Fail
	None	0.10	536.456	4	5	< 0.01	Fail
	Rankit	0.10	4.605	4	5	> 0.25	Pass
Mercury <i>Nereis virens</i>	Log	0.10	--	4	5	--	--
	None	0.10	--	4	5	--	--
	Rankit	0.10	--	4	5	--	--
Lead <i>Nereis virens</i>	Log	0.10	2.321	4	5	> 0.25	Pass
	None	0.10	29.997	4	5	< 0.01	Fail
	Rankit	0.10	1.931	4	5	> 0.25	Pass
Zinc <i>Macoma secta</i>	Log	0.25	5.345	4	1	--	--
	None	0.25	7.889	4	1	--	--
	Rankit	0.25	6.043	4	1	--	--

<sup>1</sup>  $\nu = (N - k)/k$ , where  $N$  = total number of observations and  $k$  = number of treatments.

can be determined using a  $\chi^2$  table or a computerized probability function such as the PROBCHI function in SAS (SAS Institute, Inc. 1988c).

Bartlett's Test runs into a problem when one of the treatment variances is zero. Consider the example of mercury bioaccumulation in *Nereis virens* (see Table 7). For sediment AK, all observations are less than the same detection limit. For this illustration, one-half the detection limit was substituted,

**Table 12**  
**Results of Cochran's Test on Example Data**

Data	Transformation	$\alpha$	Test Statistic	Number of Treatments	Degrees of freedom, $\nu$	Probability	Pass/Fail
Survival <i>Pimephales promelas</i>	Arcsine	0.25	0.332	9	U <sup>1</sup>	--	--
	None	0.25	0.357	9	U	--	--
	Rankit	0.25	0.269	9	U	--	--
Survival <i>Hyalella azteca</i> (1)	Arcsine	0.10	0.334	10	6	< 0.01	Fail
	None	0.10	0.378	10	6	< 0.01	Fail
	Rankit	0.10	0.173	10	6	> 0.05	Pass
Survival <i>Hyalella azteca</i> (2)	Arcsine	0.10	0.528	4	6	> 0.05	?
	None	0.10	0.676	4	6	< 0.01	Fail
	Rankit	0.10	0.415	4	6	> 0.05	Pass
Cadmium <i>Macoma nasuta</i>	Log	0.10	0.315	4	5	> 0.05	Pass
	None	0.10	0.335	4	5	> 0.05	Pass
	Rankit	0.10	0.315	4	5	> 0.05	Pass
Cadmium <i>Nereis virens</i>	Log	0.10	0.796	4	5	< 0.01	Fail
	None	0.10	0.981	4	5	< 0.01	Fail
	Rankit	0.10	0.397	4	5	> 0.05	Pass
Mercury <i>Nereis virens</i>	Log	0.10	0.505	4	5	> 0.05	?
	None	0.10	0.629	4	5	< 0.05	Fail
	Rankit	0.10	0.539	4	5	> 0.05	?
Lead <i>Nereis virens</i>	Log	0.10	0.335	4	5	> 0.05	Pass
	None	0.10	0.757	4	5	< 0.01	Fail
	Rankit	0.10	0.323	4	5	> 0.05	Pass
Zinc <i>Macoma secta</i>	Log	0.25	0.709	4	U	--	--
	None	0.25	0.565	4	U	--	--
	Rankit	0.25	0.647	4	U	--	--

<sup>1</sup> U = unequal sample sizes.

resulting in uniform values and a variance of zero for treatment AK. Calculation of Bartlett's statistic produced negative values, which do not exist in the  $\chi^2$  distribution. Thus, Bartlett's Test cannot determine whether variances are equal or unequal for that example, unless treatment AK is dropped from the analysis. Dropping a treatment with zero variance from the equality of variances test is undesirable because it artificially narrows the range of

**Table 13**  
**Summary of Equality of Variance and Nonnormality Test Significant Results (\*\*)**  
**for Example Data**

Data	Transformation	$\alpha$	Levene's Test	Brown-Forsythe Test	Bartlett's Test	Hartley's $F_{\max}$ Test	Cochran's Test	Non-normality
Survival <i>Pimephales promelas</i>	Arcsine	0.25	**	**			--	
	None	0.25	**	**			--	
	Rankit	0.25	**				--	
Survival <i>Hyaella azteca</i> (1)	Arcsine	0.10	**		**		**	
	None	0.10	**	**	**	**	**	**
	Rankit	0.10						
Survival <i>Hyaella azteca</i> (2)	Arcsine	0.10					?	**
	None	0.10			**	**	**	**
	Rankit	0.10						
Cadmium <i>Macoma nasuta</i>	Log	0.10						**
	None	0.10						
	Rankit	0.10						
Cadmium <i>Nereis virens</i>	Log	0.10	**		**	**	**	**
	None	0.10	**		**	**	**	**
	Rankit	0.10						
Mercury <i>Nereis virens</i>	Log	0.10	**		--	--	?	**
	None	0.10	**		--	--	**	
	Rankit	0.10	**		--	--	?	**
Lead <i>Nereis virens</i>	Log	0.10						
	None	0.10	**		**	**	**	**
	Rankit	0.10						
Zinc <i>Macoma secta</i>	Log	0.25	**	**		--	--	
	None	0.25	**	**		--	--	
	Rankit	0.25	**	**		--	--	

variances, and because that treatment will still be included in the means comparison (LSD) test. In Levene's Test, the residuals of a zero variance treatment will all be zero, but that treatment still contributes degrees of freedom to the  $F$  test; thus, the effect of the zero variance treatment is not ignored.

From Table 13, it is apparent that Bartlett's Test did not or could not find significant inequality of variances in several cases where Levene's Test did,

namely survival in *Pimephales promelas* (all transformations), mercury bioaccumulation in *Nereis virens* (all transformations), and zinc bioaccumulation in *Macoma secta* (all transformations). Bartlett's Test did find significant inequality of variances in one case where Levene's Test did not—untransformed survival proportions in *Hyaella azteca* (2). In this case, Bartlett's Test may have been responding to nonnormality.

### Hartley's $F_{\max}$ Test

Hartley's  $F_{\max}$  has been described as a "quick and dirty" test because it is simple to perform but perhaps less efficient (powerful) than other tests (Sokal and Rohlf 1981). In fact, Hartley's  $F_{\max}$  has similar power to Bartlett's and Cochran's tests and sometimes higher power than Levene's Test, according to simulations conducted by Conover, Johnson, and Johnson (1981), but can have extremely high Type I error rate when sample sizes are unequal or the distribution is asymmetric. The test statistic is calculated by dividing the largest variance by the smallest variance. The significance of the test statistic must then be determined using an  $F_{\max}$  table. If the test statistic is greater than the table value for the desired  $\alpha$  level, then equality of variances is rejected. The most useful  $F_{\max}$  table, which includes critical values of  $F_{\max}$  for  $\alpha = 0.25, 0.10, 0.05,$  and  $0.01$ , is found in Gill 1978 (Table A.6.1). Gill also provides a table for unequal replication, but only at  $\alpha = 0.05$  (Table A.6.2). An approximate  $F_{\max}$  test can be conducted when sample sizes are unequal by calculating degrees of freedom  $\nu = (N - k)/k$ , where  $N =$  total number of observations and  $k =$  number of treatments, and then using the equal replication table (Table A.6.1) in Gill (1978). Critical values for  $F_{\max}$  (equal replication) are also given in Table A.1 in Milliken and Johnson (1984), Table 17 in Rohlf and Sokal (1981), and Table C.7 in Winer (1971), but only for  $\alpha = 0.05$  and  $0.01$ .

Hartley's  $F_{\max}$  has two limitations besides its reliance on tables in statistics texts for determination of significance and its high Type I error rate with nonnormal distributions and unequal sample sizes. First,  $F_{\max}$  cannot be calculated when a treatment has zero variance as this would entail division by zero. Omitting the zero variance treatment is undesirable for the same reasons cited in the discussion of Bartlett's Test. Second,  $F_{\max}$  tables do not include critical values of  $F_{\max}$  for  $\nu = 1$ . Thus, in our examples,  $F_{\max}$  could not evaluate equality of variances for mercury bioaccumulation in *Nereis virens*, which has a zero variance treatment, or for zinc bioaccumulation in *Macoma secta*, which is a highly unbalanced design requiring an approximate test with calculated degrees of freedom  $\nu = 1$ . For the example data, Hartley's  $F_{\max}$  identified unequal variances in the same cases as Bartlett's Test except for arcsine-transformed survival proportions in *Hyaella azteca* (1) (Table 13).

## Cochran's Test

Cochran's Test is another simple test for equality of variances, in which the largest variance is divided by the sum of the variances. Significance of the test statistic is determined using a table such as Table A-17 in Dixon and Massey (1983) or Table C.8 in Winer (1971). These tables require equal replication and provide critical values for the test statistic,  $C$ , at  $\alpha$  levels of 0.05 and 0.01. If  $C$  is greater than the table value for the desired  $\alpha$  level, then equality of variances is rejected.

Cochran's Test is not affected by a treatment with zero variance. However, it is severely limited by its lack of applicability with unequal replication and by the lack of table values for  $\alpha = 0.10$  and 0.25. Cochran's Test has similar power to Bartlett's Test and Hartley's  $F_{\max}$ , but likewise has high Type I error rate with asymmetric distributions (Conover, Johnson, and Johnson 1981).

Cochran's Test was applied to the example data for the cases having equal replication. Equality of variances was rejected for the same cases as Bartlett's Test, plus untransformed mercury bioaccumulation in *Nereis virens*, which Bartlett's Test was unable to evaluate because of a zero variance treatment. Three additional cases produced uncertain results, indicated by question marks in Tables 12 and 13. In these cases,  $C$  was less than but close to the table value for  $\alpha = 0.05$ . However, the recommended  $\alpha$  level for balanced design,  $n = 2$  to 9 replicates, is 0.10, and the calculated  $C$  could have exceeded the critical value for  $\alpha = 0.10$  had such table values been available.

## Performance of Levene's Test

As seen from the examples above, Levene's Test can be used in more situations than other commonly used tests for equality of variances, including Bartlett's Test, Hartley's  $F_{\max}$ , and Cochran's Test. Levene's Test appears to be less sensitive to nonnormality than the other three tests (Conover, Johnson, and Johnson 1981), although Type I error rates are still high for asymmetric distributions. The performance of Levene's Test is assessed herein using simulations incorporating normal and nonnormal distributions; balanced and unbalanced designs; equal and unequal variances; and untransformed data, log-transformed data, and rankits.

Type I error rate of Levene's Test is shown in Figure 31 as the mean increase in  $\alpha$  above the nominal value ( $\alpha = 0.10$  for balanced design and 0.25 for unbalanced design). Simulations were performed to compare the variances of two samples of 5 and 5 or 8 and 4 replicates and also four samples having 5, 5, 5, and 5 or 8, 4, 4, and 4 replicates. The samples were drawn from normal, lognormal, and gamma populations created using CVs equal to 0.1, 0.5, 1, and 2. A few simulations were also performed using a binomial distribution to represent survival data from toxicity tests.

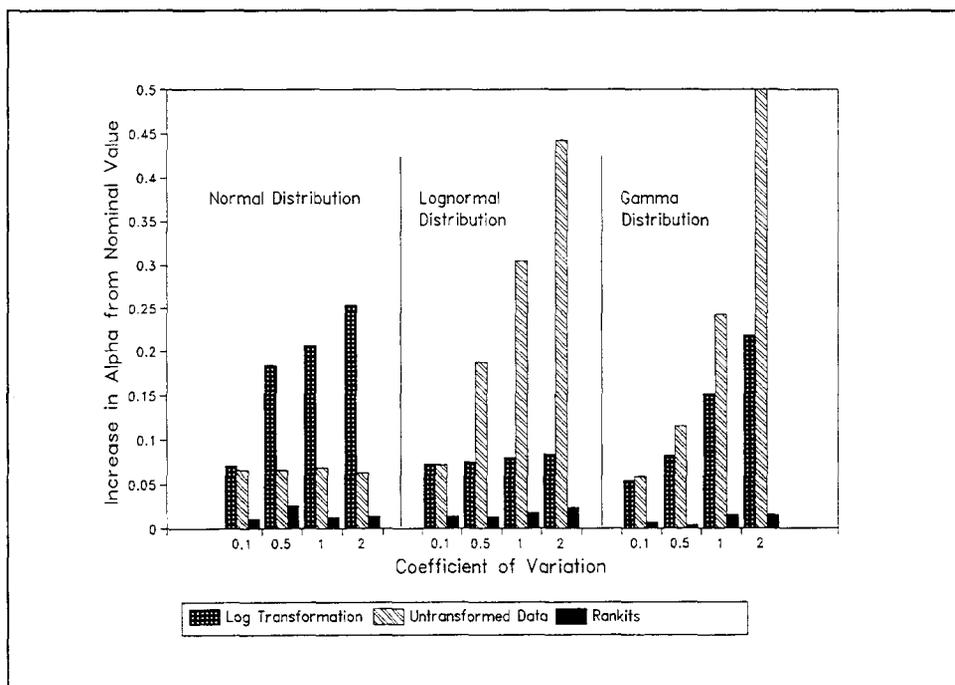


Figure 31. Mean increase in Type I error rate above the nominal  $\alpha$  for Levene's Test using simulated data

Regardless of distribution and CV, use of rankits in Levene's Test results in only a slight increase in Type I error rate over the nominal  $\alpha$ . When samples are drawn from a normal distribution, use of untransformed data in Levene's Test results in an increase in Type I error rate of about 0.05 above the nominal  $\alpha$  at all CVs. However, when normal samples are log-transformed, Type I error rate increases substantially as the CV increases. When samples are drawn from a lognormal distribution, log transformation results in an increase in Type I error rate of about 0.05 above the nominal  $\alpha$ , while use of untransformed data causes highly inflated  $\alpha$  as the CV increases. When samples are drawn from a gamma distribution, Type I error rates of Levene's Test increase greatly with increasing CV for both untransformed and log-transformed data. When samples are drawn from a binomial distribution (not shown in Figure 31), Type I error rate of Levene's Test is about 0.05 to 0.2 above the nominal  $\alpha$  for both untransformed and arcsine-transformed proportions. However, when rankits are used, there is only a slight ( $\leq 0.05$ ) increase in Type I error rate above the nominal  $\alpha$ .

These simulation results clearly support the advantage of testing for normality, especially chemical concentration data, in an attempt to identify the data distribution as normal, lognormal, or nonnormal, prior to testing for equality of variances with Levene's Test. If the data distribution can be determined and the appropriate transformation, if any, is applied to the data, then the Type I error rate of Levene's Test can be kept close to the nominal test value.

The power of Levene's Test to detect inequality of variances is shown in Figure 32 for simulations comparing four samples of  $n = 5$ , where three of the sample variances are equal and the fourth is different. When the ratio of variances is low (largest variance is 2 to 4 times the smallest variance), Levene's Test has low power ( $< 0.5$ ). Power does not approach a high level ( $\geq 0.8$ ) until the ratio of variances is around 16:1. Power is approximately the same for log transformation and for untransformed data, but is much less for rankits. Thus, conversion of the data to rankits, much more so than log transformation, tends to have an equalizing effect on the variances. The underlying population distribution appears to have little effect on the power of Levene's Test.

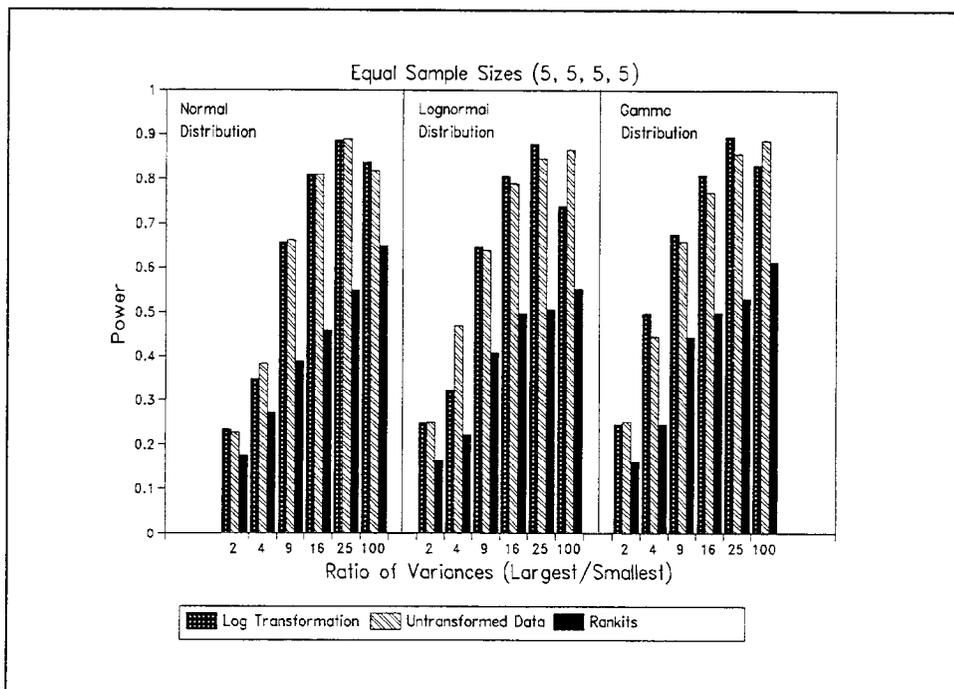


Figure 32. Power of Levene's Test when sample sizes are equal ( $n = 5$ ) and one of four variances differs from the other three

If samples are reallocated from the dredged sediments to the reference sediment, e.g.,  $n = 4, 4, 4,$  and  $8$ , the power of Levene's Test increases, largely because of the use of a higher  $\alpha$  level ( $0.25$  instead of  $0.10$ ). Power also increases when the largest sample has the largest variance. This situation is shown in Figure 33. Levene's Test now has high power when the ratio of variances is about  $9:1$ , and rankits have much less of an equalizing effect on variances than when sample sizes are equal.

For the example data, the ratio of variances (= test statistic in Table 11) ranges from a low of  $1.9:1$ , which was not detected as significantly different by any of the tests, to highs of  $60$  and  $536:1$  for cadmium in *Nereis virens*, which were considered significantly unequal by all of the tests except the Brown-Forsythe variation of Levene's Test. The high variance ratios in this

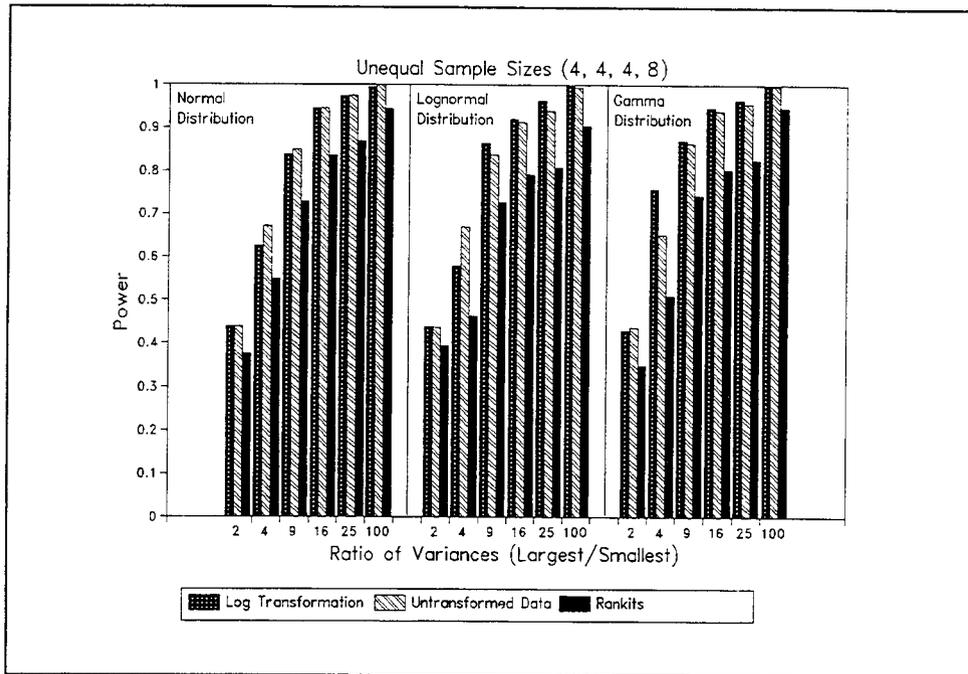


Figure 33. Power of Levene's Test when sample sizes are unequal ( $n = 4, 4, 4,$  and  $8$ ) and the variance of the largest sample is greater than the variances of the other three samples

case are due to a single outlier in one treatment (see Table 7 in Chapter 5). The Brown-Forsythe Test, which is based upon medians, is generally insensitive to outliers. Levene's Test was significant in most cases with the example data when the variance ratio was about 5:1 or higher.

## Summary

In dredged sediment evaluations, statistical comparisons are performed according to the Inland Testing Manual using either the LSD test to compare all treatments simultaneously, or  $t$ -tests to compare each dredged sediment individually with the reference. The LSD test is more powerful than  $t$ -tests when more than one dredged sediment is involved because the LSD error term incorporates more degrees of freedom. However, the LSD test, like most multiple comparison procedures, assumes that all treatment variances are approximately equal. The  $t$ -test also assumes equal variances for the two treatments being compared, but a  $t$ -test for unequal variances is available. In order to know which test to use, it is essential to first determine whether variances are statistically equal or unequal. Other important points from this chapter:

- As variances increase, the effect size (amount of difference that can be detected as significant) of a statistical comparison also increases.

- Unequal variances can inflate the Type I error rate of a statistical comparison.
- Among the tests for equality of variances recommended in the Inland Testing Manual, Levene's Test is the most versatile and should be preferred.
- SAS statements for Levene's Test, Bartlett's Test, Hartley's  $F_{\max}$ , and Cochran's Test are given in Appendix A.
- Bartlett's Test can be used with unequal sample sizes, but may be unusable when one or more treatments has zero variance.
- Hartley's  $F_{\max}$  is simple to calculate but requires looking up significance in a specialized table of critical values. An approximate test can be conducted for unequal sample sizes. Hartley's  $F_{\max}$  Test cannot be conducted when a treatment has zero variance and tables do not include critical values when there is only one degree of freedom for the test.
- Cochran's Test is also simple to calculate and can be used with zero variance treatments. However, significance must be determined using specialized tables that are limited to  $\alpha = 0.05$  or  $0.01$ . Sample sizes must be equal.
- Normality should be tested and, if necessary, an appropriate transformation applied to the data, prior to checking equality of variances.
- Bartlett's, Hartley's, and Cochran's tests all may have very high Type I error rates with asymmetric nonnormal distributions.
- Levene's Test generally has acceptable Type I error rate when a data transformation appropriate to the underlying distribution is used.
- None of the tests will have much power to detect inequality of variances when sample size is small and the ratio of largest to smallest variance is low (about 2 to 4:1).
- The power of Levene's Test is increased by reallocation of dredged sediment samples to the reference sediment, as discussed in Chapter 4, especially when the largest sample also has the largest variance.

## 7 Outliers

---

Outliers are sample observations that are exceptionally high or low, lying outside the general range of the remaining sample data. Outliers have been the subject of considerable statistical interest and much has been written about them. Barnett and Lewis (1984) defined the term as “an observation (or subset of observations) which appears to be inconsistent with the remainder of that set of data.” Although there are a variety of statistical techniques for detecting outliers, simple visual inspection of the data can be sufficient to identify obvious outliers. The outlier in treatment RH for cadmium bioaccumulation in *Nereis virens* (Table 7 in Chapter 5) is a good example. Obvious outliers may differ from the rest of the sample observations by an order of magnitude or more.

Outliers can arise from human or mechanical error, e.g., in chemical analysis or data entry. The resultant values are often quite different from the remaining data and may even be impossible. Consider as an example the following percent lipid data for sanddabs used in an experiment:

Treatment A:	6.4	5.5	5.8	5.2	5.8	7.5	41.0
Treatment B:	8.5	8.3	7.9				
Treatment C:	5.7	8.8	6.0	5.9			

Clearly, 41-percent lipid is not only well beyond the range of the other lipid data, but is highly unlikely for this fish species. The 41.0 observation could have arisen from contamination of the sample or from a decimal place or other data entry error. Outliers that are unquestionably erroneous should be corrected if possible. Often, such values can be corrected by retracing the data development process; if not, they should be discarded.

On the other hand, outliers can be perfectly valid observations that would fall in the outer tails of the underlying probability distribution. Outliers may also indicate that the samples arise from a mixture of two or more underlying distributions rather than a single parent distribution. If no explanation of error can be found for an outlier and it lies within the realm of possibility, it should be treated as a valid observation.

## Effect of Outliers on Statistical Tests

A single outlier can have a considerable effect upon the outcome of a statistical test. A classic example is simple linear (least squares) regression, where one outlier can dramatically change the slope of the regression line. Outliers can certainly affect analysis of variance type procedures, including  $t$ -tests and multiple comparison tests, as well as their concomitant tests of assumptions. This section presents simulation results showing the effect of an outlier on the Shapiro-Wilk's Test for normality, Levene's Test for equality of variances, and the  $t$ -test. Outlier effects are also illustrated using actual data examples.

### Effect on tests of assumptions

Simulations were conducted using four samples of five replicates each, randomly drawn from one normal population. In each simulation, one observation from one sample was replaced by an outlier having a distance from the population mean  $\mu$  ranging from 0 to 20 standard deviations ( $\sigma$ ). Shapiro-Wilk's Test and Levene's Test were then performed using untransformed data, log-transformed data, and rankits.

The effect of an outlier on the test for normality is shown in Figure 34 for balanced ( $P < 0.01$ ) and unbalanced ( $P < 0.05$ ) designs. Plotted points indicate the Type I error rate of the test, as the samples were drawn from a normal population. Conversion of the data to rankits results in essentially no failure of the normality test regardless of the magnitude of the outlier. The untransformed data have Type I error rates approximately equal to  $\alpha$  when the outlier is less than three standard deviations from the mean. At a distance of  $3\sigma$ , the Type I error rate begins to increase until all samples fail the normality test when the outlier is  $10\sigma$  from the mean. Log transformation results in inflated Type I error rate, which decreases somewhat as the outlier distance from the mean increases and causes the samples to resemble more closely a lognormal distribution.

The effect of an outlier on the test for equality of variances is shown in Figure 35 for the balanced design ( $P < 0.10$ ). Again, plotted points indicate the Type I error rate of the test, as the four samples being compared in each simulation were drawn from the same population. Patterns are similar to those observed for the normality test. Untransformed data have a Type I error rate somewhat above the nominal 0.10, but  $< 0.20$ , when the outlier is less than  $5\sigma$  from the mean. At an outlier distance of  $5\sigma$  from the mean, the Type I error rate begins to rise until equality of variances is rejected for all samples when the outlier is  $20\sigma$  from the mean. Log-transformed data have an initially high Type I error rate (0.35 to 0.40), which falls slightly as outlier distance from the mean increases. Rankits have a Type I error rate approximately equal to the nominal  $\alpha$  regardless of the magnitude of the outlier.

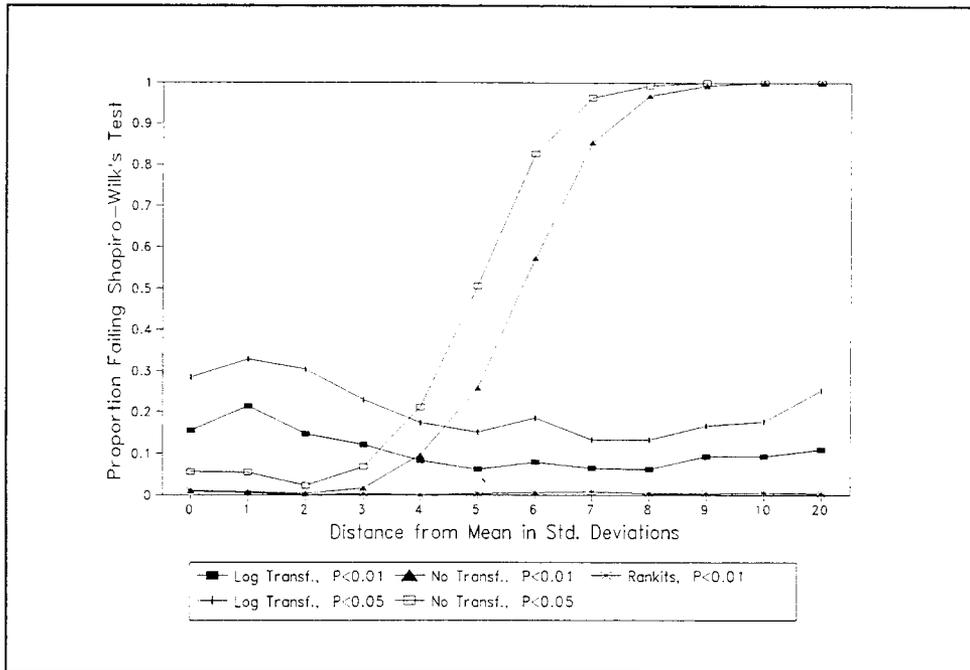


Figure 34. Type I error rate of Shapiro-Wilk's Test for balanced ( $P < 0.01$ ) and unbalanced ( $P < 0.05$ ) designs when normal samples include one outlier ranging from 0 to 20 standard deviations from the mean

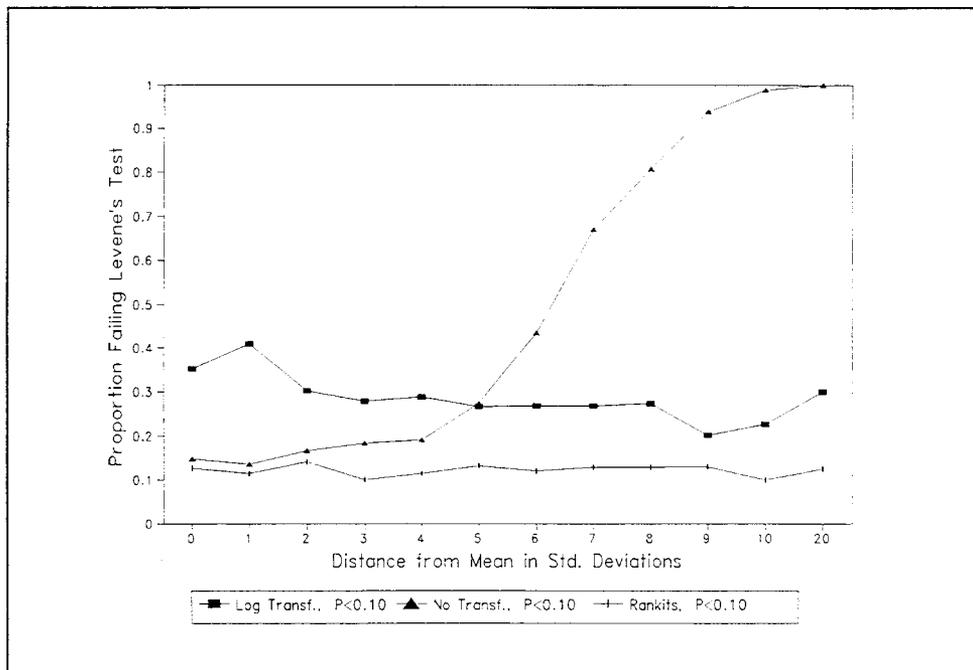


Figure 35. Type I error rate of Levene's Test for balanced ( $P < 0.10$ ) design when normal samples include one outlier ranging from 0 to 20 standard deviations from the mean

The simulations show that a single erroneous outlier can greatly increase the Type I error rate of tests of assumptions, especially if the difference between the outlier and the mean is large relative to the standard deviation. Conversion of the data to rankits effectively negates the influence of the outlier. A gross outlier will likely cause failure of the normality test for untransformed data; however, the log-transformed data may still pass. Therefore, visual inspection of the data is important, and a single large (or small) outlier should be treated with suspicion. If the outlier can be traced to error, it should be corrected or eliminated. If the outlier cannot be traced to error, then the data should be converted to rankits before proceeding with statistical analysis. If the sample data include many small values and several large values, this could indicate a lognormal distribution; testing should proceed as usual according to the decision trees in Appendix D of the Inland Testing Manual.

### Effect on *t*-tests

Staudte and Sheather (1990) have stated that the *t*-test is not robust to outliers. However, simulations using two random samples from the same normal distribution, with a high outlier substituted for one observation, do not bear this out. Simulations conducted with both the *F*-test (= a two-tailed *t*-test for equal variances) and a one-tailed *t*-test for unequal variances show no elevation of Type I error rate above the nominal  $\alpha$ , even when the outlier is  $20\sigma$  from the mean. This is the case regardless of whether the data are untransformed, log-transformed, or converted to rankits.

However, an outlier can obscure a difference that would be significant without the outlier, as will be seen later in the lipid example. Simulations were performed using random samples from two normal populations having different means and the same variance. The population with the lower mean is considered to represent a reference sediment, and the population with the higher mean is considered to represent a dredged sediment. One-tailed *t*-tests for unequal variances are used to compare the two samples in each simulation. When an observation from the reference sample is replaced by an outlier, the power of the *t*-test declines rapidly as the outlier distance above the mean increases (Figure 36). When the outlier is  $4\sigma$  greater than the mean, the power of the *t*-test is zero. Power in the presence of a reference sediment outlier is approximately the same for untransformed data, log-transformed data, and rankits.

Substituting an outlier for an observation from the dredged sediment sample increases the power of the *t*-test when the outlier is 1 or  $2\sigma$  greater than the mean, but then power begins to decline, especially for untransformed and log-transformed data, as outlier distance from the mean increases (Figure 36). These trends occur because, as the outlier increases, the dredged sediment mean increases, but the variance also increases, resulting in progressively more unequal variances and thus a progressively less powerful test as the adjusted degrees of freedom decrease. When data are untransformed, power

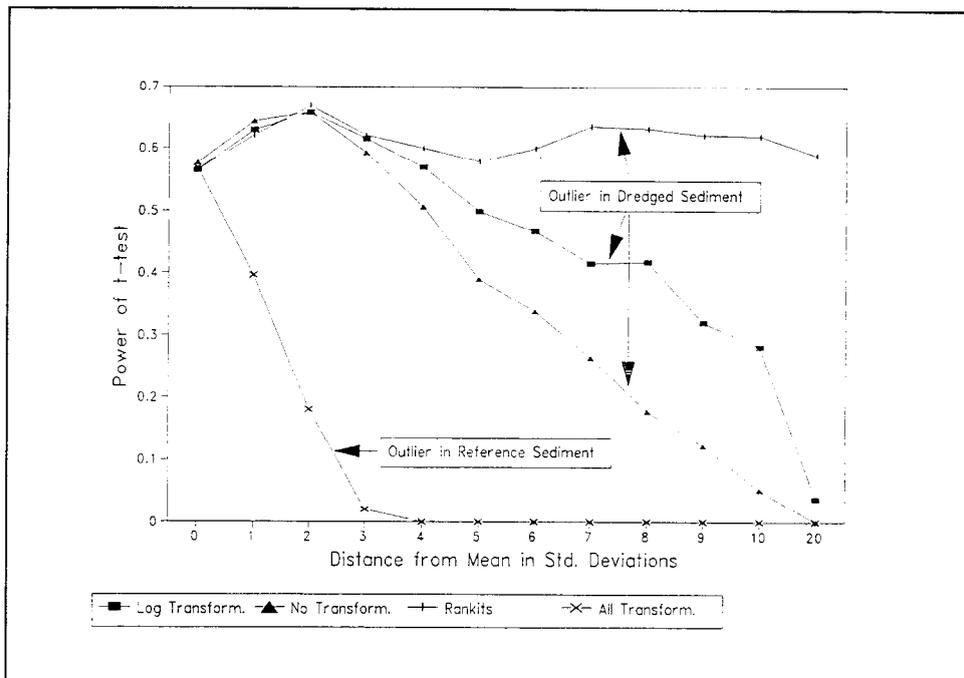


Figure 36. Power of a 1-tailed  $t$ -test for unequal variances when normal samples include one outlier ranging from 0 to 20 standard deviations greater than the mean

drops steadily to zero at an outlier distance of  $20\sigma$  above the mean. When data are log transformed, the decline in power is less steep. However, when data are converted to rankits, the effect of the outlier is largely negated and power remains approximately constant regardless of the magnitude of the outlier.

### Example data

Several data sets that include an outlier or possible outlier are analyzed with and without the outlier, following the decision tree procedures in Figures D-4 and D-5 of the Inland Testing Manual, to illustrate the effects of the outlier on tests of assumptions and the LSD test. Data sets include survival in *Hyalella azteca* (2) (Table 5 in Chapter 5), cadmium and lead bioaccumulation in *Nereis virens* (Table 7 in Chapter 5), and the sanddab lipid data given at the beginning of this chapter.

**Example 1: survival in *Hyalella azteca* (2).** Survival of 3 in the first replicate of sediment MC-2 might be considered an outlier because survival in all other replicates of this treatment is 9 or 10, and survival in the rest of the data set ranges from 6 to 10. Water quality data collected for this experiment indicate no discernible problem that might have caused high mortality in one replicate, and the survival of 3 does not appear to be an error. Thus, there would be no reason to exclude the outlier from the data analysis. Results are

given below with the outlier excluded merely to illustrate the effect of the outlier. The whole data set (survival expressed as proportions) may be summarized as follows:

Outlier included:  $\bar{x} = 0.911$   $s^2 = 0.0240$   $s = 0.1548$   $CV = 0.17$   $N = 28$   
 Outlier excluded:  $\bar{x} = 0.933$   $s^2 = 0.01$   $s = 0.1$   $CV = 0.11$   $N = 27$

Exclusion of the outlier increases the mean slightly, halves the variance, and nearly halves the CV. Analysis of arcsine-transformed survival proportions shows failure of normality, regardless of the outlier ( $\alpha$  values for tests of assumptions are taken from Table D-2 of the Inland Testing Manual):

Outlier included: Shapiro-Wilk's  $W = 0.879$   $P = 0.0035$  (fail,  $\alpha = 0.01$ )  
 Outlier excluded: Shapiro-Wilk's  $W = 0.859$   $P = 0.0014$  (fail,  $\alpha = 0.01$ )

Thus, the data would be converted to rankits and the analysis completed:

Outlier included: Shapiro-Wilk's  $W = 0.911$   $P = 0.0220$  (pass,  $\alpha = 0.01$ )  
 Outlier excluded: Shapiro-Wilk's  $W = 0.855$   $P = 0.0012$  (fail,  $\alpha = 0.01$ )

Outlier included: Levene's  $F = 1.23$   $P = 0.3215$  (pass,  $\alpha = 0.10$ )  
 Outlier excluded: Levene's  $F = 0.34$   $P = 0.7987$  (pass,  $\alpha = 0.10$ )

Outlier included: LSD results: MC-3 survival significantly lower than reference

Outlier excluded:  $t$ -test results: MC-3 survival significantly lower than reference

With the outlier included, the rankits pass the normality and equality of variances assumptions, so the LSD test is used with rankits to compare treatments. When the outlier is excluded, the rankits fail the normality assumption, so each treatment is compared with the reference using a  $t$ -test. In either case, the comparison outcome is ultimately the same.

**Example 2: cadmium bioaccumulation in *Nereis virens*.** In this example from Table 7 of Chapter 5, sediments AK, GOW, and RH are the dredged sediments, while SH is the reference sediment. Observation 0.258 in sediment RH is clearly an outlier, as it is about an order of magnitude greater than any observation in sediment RH or in the other sediments as well. No information is available to indicate why this observation differs from the others. Therefore, there appears to be no justification for excluding 0.258 from the data analysis. Again, results are given below with the outlier excluded as well as included to illustrate the influence of the outlier. Bioaccumulation for the whole data set may be summarized as follows:

Outlier included:  $\bar{x} = 0.042$   $s^2 = 0.0022$   $s = 0.0465$   $CV = 1.10$   $N = 24$   
 Outlier excluded:  $\bar{x} = 0.033$   $s^2 = 0.00006$   $s = 0.0074$   $CV = 0.23$   $N = 23$

As in the previous example, exclusion of the outlier has a pronounced effect upon the variance and CV. Analysis of the untransformed data shows failure of normality when the outlier is included but not when it is excluded:

Outlier included: Shapiro-Wilk's  $W = 0.556$   $P = 0.0001$  (fail,  $\alpha = 0.01$ )

Outlier excluded: Shapiro-Wilk's  $W = 0.974$   $P = 0.7769$  (pass,  $\alpha = 0.01$ )

If the outlier were excluded, the analysis would continue as follows using untransformed data:

Levene's  $F = 2.18$   $P = 0.1234$  (pass,  $\alpha = 0.10$ )

LSD results: no dredged sediment significantly greater than reference

However, since there is no apparent reason to exclude the outlier, in practice the data would be transformed to logs and the normality test rerun:

Shapiro-Wilk's  $W = 0.773$   $P = 0.0001$  (fail,  $\alpha = 0.01$ )

The log-transformed data still fail normality, so the data would then be converted to rankits:

Shapiro-Wilk's  $W = 0.962$   $P = 0.4941$  (pass,  $\alpha = 0.01$ )

Levene's  $F = 2.36$   $P = 0.1019$  (pass,  $\alpha = 0.10$ )

LSD results: no dredged sediment significantly greater than reference

When the outlier is not excluded, the rankits pass normality and equality of variances, and again the LSD test shows no statistically significant elevation in cadmium bioaccumulation from the dredged sediments compared with bioaccumulation from the reference sediment. Thus, the outlier does not affect the comparison outcome, just the data transformation necessary to pass the tests of assumptions.

**Example 3: lead bioaccumulation in *Nereis virens*.** As seen in Chapter 5, this data set is suggestive of a lognormal distribution. If this is not the case, then the highest value, 0.973 in sediment RH, looks like a possible outlier. Data analysis results are given as in the previous examples. For the whole data set:

Outlier included:  $\bar{x} = 0.188$   $s^2 = 0.0390$   $s = 0.1974$   $CV = 1.05$   $N = 24$

Outlier excluded:  $\bar{x} = 0.154$   $s^2 = 0.0115$   $s = 0.1073$   $CV = 0.70$   $N = 23$

Decrease in variance and CV with exclusion of the putative outlier is not as pronounced as in the cadmium example. Analysis of the untransformed data shows failure of normality when the outlier is included but not when it is excluded:

Outlier included: Shapiro-Wilk's  $W = 0.807$   $P = 0.0002$  (fail,  $\alpha = 0.01$ )  
Outlier excluded: Shapiro-Wilk's  $W = 0.925$   $P = 0.0849$  (pass,  $\alpha = 0.01$ )

If the outlier were excluded, the analysis would continue as follows using untransformed data:

Levene's  $F = 1.26$   $P = 0.3176$  (pass,  $\alpha = 0.10$ )  
LSD results: no dredged sediment significantly greater than reference

Not excluding the outlier, the data would be transformed to logs and the analysis continued:

Shapiro-Wilk's  $W = 0.982$   $P = 0.9199$  (pass,  $\alpha = 0.01$ )  
Levene's  $F = 0.47$   $P = 0.7055$  (pass,  $\alpha = 0.10$ )  
LSD results: no dredged sediment significantly greater than reference

When the outlier is included, the log-transformed data pass the tests of assumptions, indicating that the sample is drawn from a lognormal distribution. Without the outlier, there is insufficient evidence to reject normality of the untransformed data. As in the previous examples, the outlier affects the tests of assumptions and the data transformation required, but not the statistical comparison outcome.

**Example 4: lipid concentration of sanddabs.** The lipid data given at the beginning of this chapter are actual data that have been arbitrarily assigned to treatments to illustrate the influence of an outlier on statistical comparisons. In this example, all treatments will be compared with each other. The 41-percent lipid observation in Treatment A is indisputably an outlier and almost certainly erroneous. If possible, the error should be corrected; if not, the outlier should be dropped prior to data analysis.

Summary statistics for the whole data set are as follows:

Outlier included:  $\bar{x} = 9.164$   $s^2 = 85.484$   $s = 9.246$   $CV = 1.01$   $N = 14$   
Outlier excluded:  $\bar{x} = 6.715$   $s^2 = 1.651$   $s = 1.285$   $CV = 0.19$   $N = 13$

Here, inclusion of the outlier inflates the variance and CV by an order of magnitude. Analysis of the untransformed data shows failure of normality regardless of whether the outlier is included:

Outlier included: Shapiro-Wilk's  $W = 0.560$   $P = 0.0001$  (fail,  $\alpha = 0.10$ )  
Outlier excluded: Shapiro-Wilk's  $W = 0.837$   $P = 0.0181$  (fail,  $\alpha = 0.10$ )

Transforming the data to logs also results in failure of the normality test:

Outlier included: Shapiro-Wilk's  $W = 0.637$   $P = 0.0001$  (fail,  $\alpha = 0.10$ )  
Outlier excluded: Shapiro-Wilk's  $W = 0.856$   $P = 0.0280$  (fail,  $\alpha = 0.10$ )

The data would thus be converted to rankits and the analysis would proceed:

Outlier included: Shapiro-Wilk's  $W = 0.945$   $P = 0.4630$  (pass,  $\alpha = 0.10$ )  
 Outlier excluded: Shapiro-Wilk's  $W = 0.968$   $P = 0.8195$  (pass,  $\alpha = 0.10$ )

Outlier included: Levene's  $F = 1.52$   $P = 0.2612$  (pass,  $\alpha = 0.25$ )  
 Outlier excluded: Levene's  $F = 1.25$   $P = 0.3264$  (pass,  $\alpha = 0.25$ )

Outlier included: LSD results: treatments do not differ significantly from each other  
 Outlier excluded: LSD results: Treatment B is significantly greater than Treatment A

In this example, the outlier does not affect the tests of assumptions, but does mask a significant difference between Treatments A and B.

## Detecting Outliers

Often, outliers will be glaring, and the need to recheck the data and correct errors if possible will be obvious. Sometimes, however, identifying suspicious observations as outliers may be more difficult. Figure 36 shows that a high outlier differing from the population mean by only one or two standard deviations can have a considerable effect upon the power of a statistical comparison, especially when the outlier is in the reference sediment data. At times it may be cumbersome to retrace the data development process, and one would like to know whether a suspicious value really is an outlier before attempting to retrace the data. A procedure to identify outliers statistically would be desirable for such ambiguous situations. Many such procedures, known as discordancy tests, have been developed for a variety of distributions (Barnett and Lewis 1984). Only Dixon's Test (Dixon 1950) will be considered here. Dixon's Test is applicable to normal distributions and also to lognormal distributions when the data have been transformed to logarithms. Dixon's Test serves as a quick screen for outliers when sample size  $n \leq 25$ . For  $n > 25$ , the suspected outlier is normalized; the procedure is described in Sokal and Rohlf (1981).

To illustrate Dixon's Test, reconsider the data for lead bioaccumulation in Table 7 of Chapter 5. To conduct Dixon's Test, order the data for each treatment from lowest to highest so that the first and last values will be the suspected low and high outliers;  $X_1$  will be the lowest value and  $X_n$  will be the highest value. A critical ratio for each treatment is calculated using:

$$\text{Suspected low outlier} \quad r = (X_2 - X_1)/(X_n - X_1) \quad (1)$$

$$\text{Suspected high outlier} \quad r = (X_n - X_{n-1})/(X_n - X_1) \quad (2)$$

Using data from Sediment AK (and substituting one-half the detection limit for the nondetect), Equation 1 yields:

$$r = (0.100 - 0.011)/(0.186 - 0.011) = 0.089/0.175 = 0.509$$

and Equation 2 yields:

$$r = (0.186 - 0.171)/(0.186 - 0.011) = 0.015/0.175 = 0.086$$

The critical values are listed in Table 36 of Rohlf and Sokal (1981), Table 2 of Grubbs (1969), and Table A 16 of Snedecor and Cochran (1989). When  $n = 6$  and  $\alpha = 0.05$ , the critical value is 0.560. Data points whose ratio  $r$  exceeds the critical value are considered outliers. In this example, neither 0.509 nor 0.086 exceeds the critical value, so one concludes that the extreme observations of treatment AK are not outliers.

For sediment GOW, the ratios are 0.103 and 0.385. These do not exceed the critical value. For sediment RH, the ratios are 0.017 and 0.709. Since 0.709 exceeds the critical value, one can consider observation 0.973 to be an outlier. For sediment SH, the ratios are 0.116 and 0.800. Since 0.800 exceeds the critical value, observation 0.391 could also be considered an outlier.

As previously seen, the untransformed lead bioaccumulation data did not pass the test for normality, whereas the log-transformed data did. Therefore, Dixon's Test should be performed on the log-transformed data. Recalculating the ratios for AK using log-transformed data, Equation 1 yields:

$$r = (-1 - -1.959)/(-0.730 - -1.959) = 0.959/1.228 = 0.781$$

and Equation 2 yields:

$$r = (-0.730 - -0.767)/(-0.730 - -1.959) = 0.037/1.228 = 0.030$$

Now, 0.781 exceeds the critical value of 0.560, so the less-than detection limit observation in treatment AK could be considered a low outlier, assuming the sample is drawn from a lognormal distribution.

For log-transformed data of sediment GOW, the ratios are 0.288 and 0.184, indicating no outliers. For log-transformed data of sediment RH, the ratios are 0.076 and 0.402. The observation that was previously seen to be an outlier when untransformed, is no longer an outlier when log-transformed. For log-transformed data of sediment SH, the ratios are 0.292 and 0.572. The observation that was an outlier when untransformed remains an outlier when log-transformed.

Two consecutive values may be tested by repeated use of Dixon's Test. The second largest or second smallest values are tested after deleting the

largest or smallest value from the data set. Sometimes a second outlier can mask the first outlier.

Routine use of Dixon's Test should not be necessary in dredged sediment evaluations. If the need arises for statistical identification of outliers, Dixon's Test can easily be performed using a hand calculator.

## What To Do When Outliers Occur

It is most important to identify outliers due to error. Such values are often obviously extreme and may even be impossible. In situations where erroneous outliers can be corrected by retracing the data development process, the corrected data are then used in all statistical analyses. If an outlier is certainly due to error but there is no way of finding the correct value, the outlier should be deleted before conducting statistical analyses.

When outliers cannot be unequivocally attributed to errors, the investigator must decide what, if anything, to do about them. Special procedures are available to analyze data that include outliers, such as trimming and Winsorization (Dixon and Massey 1983; Dixon and Tukey 1968; Winer 1971). Symmetrical trimming discards the highest and lowest observations in a sample containing an outlier, whereas symmetrical Winsorization substitutes the second highest and second lowest values for the highest and lowest observations, respectively. The calculated  $t$  statistic and degrees of freedom are then adjusted (formulas are given in Dixon and Massey 1983:381-382).

Simulations were conducted as before to compare the effect of trimming, Winsorization, and simple deletion of a high outlier in the reference sediment on the power of a one-tailed  $t$ -test for unequal variances when  $n = 5$ . Figure 36 shows power declines precipitously as the outlier distance above the mean increases. When the reference sediment data are symmetrically trimmed or Winsorized, the magnitude of the outlier becomes irrelevant; but the power of the  $t$ -test is low for small  $n$  (Figure 37). Trimming and Winsorization thus appear to be inappropriate for small sample sizes because these methods unnecessarily sacrifice information and degrees of freedom, resulting in less powerful tests. Simply deleting the outlier results in power nearly as high as in the absence of outliers (i.e., outlier distance of  $0\sigma$  from the mean in Figure 37). Therefore, if a high outlier in the reference sediment (or low outlier in the case of survival data) is merely suspected of being due to error, the investigator may choose to delete that value in order to preserve power in the statistical comparisons. However, if the outlier is thought to be a valid observation, it should be retained.

The best procedure for dealing with nonerror outliers among dredged sediment replicates is to simply follow the decision trees in Appendix D of the Inland Testing Manual. Usually the presence of extreme outliers will result in failure of the normality test, in which case the data would be converted to

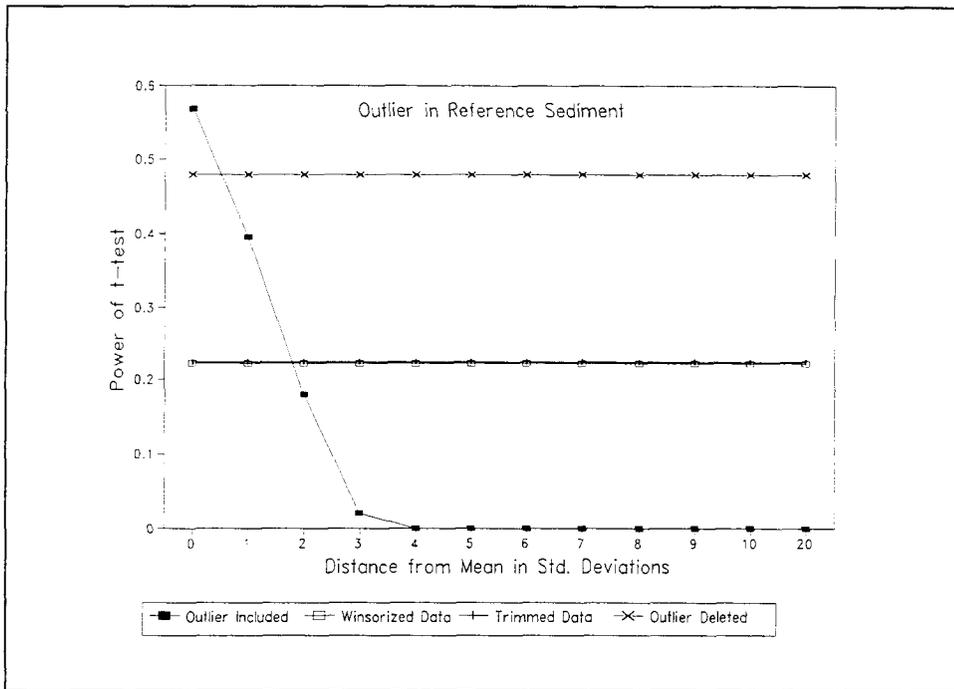


Figure 37. Power of a 1-tailed  $t$ -test for unequal variances when normal samples include one outlier in reference sediment ranging from 0 to 20 standard deviations greater than the mean

rankits. As seen earlier in this chapter, rankits effectively negate the influence of outliers, in tests of assumptions as well as statistical comparisons. Thus, when outliers are identified but cannot be traced to error, it would be acceptable to simply convert the data to rankits and proceed with Step 11 in Figure D-1 or with the steps in Figures D-4B or D-5B of the Inland Testing Manual. As noted above, extra caution is advised when the suspect observation is a high outlier in the reference sample (or low outlier for survival data). Conversion to rankits in such cases will not preserve the power of the  $t$ -test (see Figure 36).

## Summary

The sequences of statistical procedures established in the Inland Testing Manual are generally well equipped to handle outliers, and thus concern over outliers is usually unnecessary unless the outliers can be traced to error. In such cases, the errors should be corrected or the outliers eliminated prior to data analysis.

Other important points of this chapter may be summarized as follows:

- Visual inspection of the data is often sufficient to identify outliers. If statistical identification of questionable outliers is desired, Dixon's Test may be used on normal data or on lognormal data that are log-transformed.

- Extreme values should be traced through the data development process to determine if they can be attributed to error, and corrected if possible.
- Outliers result in increasing failure rates for tests of assumptions using untransformed data, as the relative distance of the outlier from the mean increases. Outliers have less influence on tests of assumptions when data are log-transformed and no influence when data are converted to rankits.
- Outliers do not appear to affect the Type I error rate of the  $t$ -test, but can substantially decrease the power of the test, especially when an outlier occurs among the reference sediment replicates.
- An outlier among the reference sediment replicates should be treated with caution. If there is any suspicion that the outlier is due to error, it should be deleted (or corrected if possible).
- Trimming and Winsorization are not recommended for dealing with outliers when sample size is small (e.g.,  $n = 5$ ) because subsequent statistical comparisons may lose too much power.

# 8 Less-Than Detection Limit (Censored) Data

---

Less-than detection limit observations in samples of contaminant concentration data present a problem for statistical analysis because such “censored” observations represent unknown values. Although various authors have argued strongly that concentration data not be censored by analytical laboratories (Cressie 1994; Porter, Ward, and Bell 1988), the reality is that censored measurements will continue to be reported. Therefore, data sets that include censored observations must be manipulated in some way before the data can be summarized or statistical comparisons performed. A series of publications describes the problem of less-than detection limit data in dredged sediment evaluations (Clarke 1992; Clarke and Brandon 1994) and a simulation study designed to assess methods for statistical treatment of less-than detection limit data to permit statistical comparisons (Clarke 1994; Clarke 1995b). Guidelines developed from the simulation study are presented in Clarke (1995a) and included in Appendix D of the Inland Testing Manual. This chapter will focus on application of the recommended censored data methods and give examples for a number of common situations.

## Methods for Censored Data

A variety of methods have been applied to the analysis of left-censored data sets. Left censoring refers to an area of unknown values below some cut-off point, such as a detection limit, in the left tail of a probability distribution. Methods discussed below that were evaluated in the simulation study (Clarke 1995a) will be named in bold. The simplest censored data method is to delete the unknown observations. However, deletion sacrifices valuable information, leading to biased estimation and less powerful statistical comparisons because of loss of degrees of freedom. Less-than detection limit concentration data are known to lie within the interval from zero to the detection limit, and that information should be incorporated into statistical analyses.

- Many censored data methods substitute a value from the interval for each less-than detection limit observation (hereafter referred to for convenience as

“nondetect”).<sup>1</sup> Simple substitution methods replace each nondetect with a constant, usually the detection limit (DL), one-half the detection limit (DL/2), or zero (ZERO).<sup>2</sup>

Other common methods substitute an observation from an assumed type of probability distribution, such as normal, lognormal, uniform, Weibull, etc. The substituted value may be drawn randomly from the assumed distribution, or it may correspond to a given quantile of that distribution. For example, if one assumes that a nondetect has an equal probability of occurring anywhere between zero and the detection limit, the interval from zero to the detection limit follows a uniform distribution. For each nondetect, a substitution value may be picked at random using a random numbers table or a computerized random number generator (UNIFR). Alternatively, values at evenly spaced intervals between zero and the detection limit can be substituted for the nondetects in a sample (UNIF). In a sample with four nondetects, for example, the nondetects would be replaced by zero, one-third DL, two-thirds DL, and DL.

Quantiles of distributions other than the uniform are more difficult to estimate and require more sophisticated techniques such as maximum likelihood estimation. In the SAS LIFEREG procedure, maximum likelihood methods are applied to the known (uncensored) observations in a sample, given a specified distribution, to determine below detection quantile values that can then be substituted for the nondetects. This procedure was used in the simulation study assuming a normal (MLE NORM), lognormal (MLE LOGN), or Weibull (MLE WEIB) distribution. Survival analysis procedures that can accommodate left-censored data, such as SAS LIFEREG, can also be applied directly for statistical comparisons of samples with nondetects, without first having to substitute values for the nondetects (Slymen, de Peyster, and Donohoe 1994).<sup>3</sup>

Values for nondetects can be estimated using least-squares regression techniques, which also take advantage of the information provided by the uncensored observations in a sample. The regression methods assume either a normal (NR) or lognormal (LR) distribution depending on whether the untransformed or log-transformed data are used. The above-detection limit data are regressed against their normal scores (rankits) to estimate regression

---

<sup>1</sup> Some analytical laboratories distinguish between less-than detection limit observations, in which a signal was observed but fell below predetermined limits of precision, and nondetects, in which no signal was observed. For more information concerning detection limits, see Glaser et al. (1981); Lambert, Peterson, and Terpenning (1991); Parsons (1969); and Porter, Ward, and Bell (1988).

<sup>2</sup> When ranks or rankits are used for nonparametric statistical comparisons, the censored data methods DL, DL/2, and ZERO are equivalent, and are simply referred to as CONST (for substitution of any constant between zero and the detection limit).

<sup>3</sup> Slymen, de Peyster, and Donohoe (1994) propose that their technique, called tobit analysis, avoids the bias inherent in substitution methods. This technique, which could not be evaluated in the simulation study, shows promise for hypothesis testing, although questions remain to be addressed.

parameters, which are then used to calculate substitution values for the nondetects. If desired, the detection limit may be included in the regression.

Much has been written concerning estimation of population parameters using censored data, but few studies have considered the application of censored data methods in hypothesis testing problems. Maximum likelihood and regression methods have generally been recommended for estimation problems, but these techniques have limited utility with small sample sizes. Therefore, a simulation study was undertaken to determine the most appropriate censored data methods for hypothesis testing using small samples, as in dredged sediment bioaccumulation comparisons. The 10 methods listed above, along with uncensored data (method = NONE), were evaluated for performance with the LSD test. Sample sizes ranged from three to eight replicates. Complete details of the simulation study are given in Clarke (1995b). Recommendations based on the simulation results are presented in Table D-12 of the Inland Testing Manual and in Clarke (1995a). Verifications using a large number of actual chemical concentration data sets support the simulation study conclusions (Clarke 1995a).

## Applying Censored Data Methods

Examples are given in the following sections illustrating a variety of common situations involving nondetects, along with the steps in selecting and applying the most appropriate censored data method for a given situation. The steps of this process are outlined in Clarke (1995a) and in Section D3.1.1.1 of the Inland Testing Manual. These steps include checking for equality of variances and normality and calculating the CV of the combined samples. However, such steps cannot be accomplished unless values are first substituted for the nondetects. Two or more censored data methods may be applied as a preliminary step to assess the type of distribution and obtain a range of possible values for the variances and CV. If the recommended method from Table D-12 of the Inland Testing Manual then turns out to be one of the methods used in the preliminary analysis, the initial tests of assumptions are already completed; one simply proceeds at the appropriate step in the decision tree of Figure D-5. SAS program statements for all recommended censored data methods are given in Section D4.5 of the Inland Testing Manual, and most of them are also provided in preliminary analysis programs in Appendix A of this report. Information in the following example sections applies to single detection limit censored data unless otherwise stated.

### One nondetect in one or more treatments

When each treatment includes no more than one nondetect, two or three simple substitution methods may be used for the preliminary analysis. A SAS program to apply DL, DL/2, and ZERO, followed by the other preliminary steps in selecting the best censored data method, is given in Appendix A (PRELIM1.SAS). The data for lead bioaccumulation in *Nereis virens* (Table 7

in Chapter 5) include a single nondetect in one treatment. Results of the preliminary steps for these example data are shown in Table 14.

**Table 14**

**Steps in Selecting the Most Appropriate Censored Data Method: One Nondetect in One or More Treatments**

	Preliminary Method			Consensus
	DL	DL/2	ZERO	
Combined CV	1.05	1.05	1.06	> 1
Sampling Distribution	Lognormal	Lognormal	Lognormal	Lognormal
Variances of Untransformed Data	Unequal; increase as means increase			
Data Transformation Required				Log
Most Appropriate Censored Data Method (from Table D-12 of the Inland Testing Manual)				DL
Steps Remaining After Applying Censored Data Method: Test equality of variances of logs; perform comparisons using logs				
Note: Example Data: Lead bioaccumulation in <i>Nereis virens</i> . Amount of Censoring: 4 percent (1 out of 24 total replicates).				

For this example, all three preliminary methods are in agreement that the combined CV > 1, the data are lognormally distributed, and variances increase as means increase. Checking the appropriate section of Table D-12 for log transformation and ≤20-percent censoring, one finds the recommended censored data method to be DL. The test of normality has already been done for logs using DL, so the only steps remaining are to test equality of variances of the logs, and then perform statistical comparisons using logs in the LSD test if variances are equal, or in *t*-tests if variances are unequal (Figure D-5A of the Inland Testing Manual).

**Several nondetects in one or more treatments (<50-percent censoring)**

When sample data contain more than one nondetect in any treatment, at least one of the uniform distribution or maximum likelihood methods should be included in the preliminary analysis. A SAS program to apply DL/2, DL, UNIF, and MLE WEIB is given in Appendix A (PRELIM2.SAS). As an example, consider the following data for mercury bioaccumulation in *Macoma nasuta*, which include four nondetects in one treatment and two nondetects in another:

Sediment AK:	0.033	0.066	0.028	0.034	0.034	0.030
Sediment GOW:	<0.02	<0.02	<0.02	0.160	<0.02	0.036
Sediment RH:	0.028	<0.02	<0.02	0.207	0.032	0.032
Sediment SH:	0.036	0.028	0.085	0.023	0.023	0.040

Preliminary analysis results for these data are given in Table 15.

<b>Table 15</b>					
<b>Steps in Selecting the Most Appropriate Censored Data Method: More Than One Nondetect in Any Treatment</b>					
	Preliminary Method				Consensus
	DL	DL/2	MLE WEIB	UNIF	
Combined CV	1.02	1.12	1.18	1.13	> 1
Sampling Distribution	Nonnormal	Nonnormal	Nonnormal	Nonnormal	Nonnormal
Variances of Untransformed Data	Equal	Equal	Equal	Equal	Equal
Data Transformation Required					Rankit
Most Appropriate Censored Data Method (from Table D-12 of the Inland Testing Manual)					CONST or UNIF
Steps Remaining After Applying Censored Data Method: Test normality and equality of variances of rankits; perform comparisons using rankits					
Note: Example Data: Mercury bioaccumulation in <i>Macoma nasuta</i> . Amount of Censoring: 25 percent (6 out of 24 total replicates).					

In this example, once again all preliminary methods were in agreement concerning the CV, sampling distribution, and equality of variances. The most appropriate censored data method is either CONST or UNIF; both were shown to give satisfactory results in the simulations for this type of situation. After applying CONST or UNIF to the nondetects, the data are converted to rankits and tested for normality and equality of variances prior to comparing treatment bioaccumulation.

### More than 50-percent censoring

When data sets are highly censored, any analysis will be tenuous. PRELIM2.SAS can still be used for preliminary analysis and identification of the most appropriate censored data method, providing no treatment is completely censored. Consider the following data for polychlorinated biphenyl (PCB) congeners 95+66 in *Mytilus edulis*, in which two-thirds of the observations are nondetects. Exposures were to a surficial San Francisco Bay sediment via two routes—bedded sediment (BS) and 50-mg/l sediment suspension (S50). Bioaccumulation from each exposure was to be compared with background (Day0) tissue levels.

BS: <5.0 <5.0 <5.0 7.7 <5.0 5.1  
 S50: <5.0 5.3 <5.0 <5.0 <5.0 <5.0  
 Day0: <5.0 5.8 5.4

Note that all uncensored tissue concentrations are fairly close to the detection limit and that the design is unbalanced. Results from PRELIM2.SAS are reported in Table 16.

<b>Table 16</b>					
<b>Steps in Selecting the Most Appropriate Censored Data Method: Highly Censored Data</b>					
	Preliminary Method				Consensus
	DL	DL/2	UNIF	MLE WEIB	
Combined CV	0.13	0.48	0.34	0.62	Probably between 0.26 and 1
Sampling Distribution	Nonnormal	Nonnormal	Nonnormal	Normal	Probably nonnormal
Variances of Untransformed Data	Unequal; mixed	Unequal; mixed	Unequal; mixed	Equal	Probably unequal; mixed
Data Transformation Required					Rankits
Most Appropriate Censored Data Method (from Table D-12 of the Inland Testing Manual)					CONST
Steps Remaining After Applying Censored Data Method: Test normality and equality of variances of rankits; perform comparisons using rankits					
Note: Example Data: PCB 95 + 66 bioaccumulation in <i>Mytilus edulis</i> . Amount of Censoring: 67 percent (10 out of 15 total replicates).					

Three of the preliminary methods (DL, DL/2, UNIF) are in agreement with each other concerning sampling distribution and variances, but disagree with preliminary results using MLE WEIB. Therefore, a best guess is made concerning the properties of the data set, based on the three methods in agreement. The data are most likely nonnormal and variances most likely mixed, leading to selection of CONST with rankits from Table D-12 of the Inland Testing Manual for data that are 61- to 80-percent censored.<sup>1</sup> Accepting the preliminary results of MLE WEIB would lead to use of DL/2 or ZERO with untransformed data. Table 16 has already shown that application of DL/2 to these data leads to rejection of normality and equality of variances. Thus the end result, analysis of rankits after substituting a constant, would come out the same.

<sup>1</sup> Note that the Type I error rate for simulations of this situation is between 0.06 and 0.10; i.e., there is a slightly greater than expected chance of finding a significant difference among treatments where none exists in reality.

### Severe (> 80-percent) censoring

When more than 80 percent of the observations in a data set are nondetects, the data may be summarized or described, and perhaps tentative conclusions can be drawn; but statistical analysis generally should not be attempted. Consider the following data for PCB congener 100 bioaccumulation in the mussel *Mytilus edulis* and sanddabs *Citharichthys stigmaeus* exposed to BS and S50 compared with Day0:

Mussels	BS:	<0.5	<0.5	<0.5	<0.5	<0.5	<0.5
	S50:	<0.5	<0.5	<0.5	<0.5	1.4	<0.5
	Day0:	<0.5	<0.5	3.0			
Sanddabs	BS:	<0.5	<0.5	<0.5	<0.5	<0.5	<0.5
	S50:	<0.5	<0.5	<0.5	<0.5	15.6	<0.5
	Day0:	<0.5	<0.5	<0.5			

The data for mussels, which are 87-percent censored, suggest that neither exposure treatment resulted in elevated tissue concentrations of PCB 100 compared with background. On the other hand, the data for sanddabs, in which the only nondetect is a relatively high value of 15.6 in treatment S50, might suggest possible bioaccumulation from S50 compared with background. However, 15.6 is more than an order of magnitude higher than all the other observations, and thus is likely an outlier that should be checked for error or possible quality control problems. Statistical analysis of such highly censored data sets is fraught with problems, including low power, high Type I error rate, and high susceptibility to the influence of outliers.

### Complete censoring of a dredged sediment treatment

When all of the observations for a treatment are nondetects, the maximum likelihood and regression methods for censored data cannot be used. Instead, preliminary analysis of the data should be done using at least one simple substitution method (DL, DL/2, or ZERO), and at least one uniform distribution substitution method (UNIF or UNIFR). Because UNIFR is not the recommended method for any situation in Table D-12 of the Inland Testing Manual, the most practical candidates are the other four methods; these are included in PRELIM3.SAS (Appendix A). Note that ZERO and UNIF result in the elimination of all or some nondetects during log transformation, so the test for lognormality should be ignored if a large proportion of the data are eliminated. The data for mercury bioaccumulation in *Nereis virens* (Table 7 in Chapter 5) include a large proportion (71 percent) of nondetects, and all six observations for sediment AK are below detection limit. Results of PRELIM3.SAS on these data are shown in Table 17.

The four preliminary methods applied to these data differ widely in estimated CV and do not agree regarding whether the data are normally or non-normally distributed. The three simple substitution methods result in

**Table 17**  
**Steps in Selecting the Most Appropriate Censored Data Method: Complete Censoring of a Dredged Sediment Treatment**

	Preliminary Method				Consensus
	DL	DL/2	UNIF	ZERO	
Combined CV	0.27	0.64	0.75	1.66	Probably between 0.26 and 1
Sampling Distribution	Nonnormal	Normal	Normal	Nonnormal	?
Variances of Untransformed Data	Unequal; increase as means increase	Unequal; increase as means increase	Equal	Unequal; increase as means increase	Probably unequal; increase as means increase
Data Transformation Required					Untransformed data or rankits
Most Appropriate Censored Data Method (from Table D-12 of the Inland Testing Manual)					DL/2 with untransformed data or CONST with rankits
Steps Remaining After Applying Censored Data Method: If DL/2 is used with untransformed data, perform comparisons using <i>t</i> -tests. If CONST is used with rankits, test normality and equality of variances of rankits; perform comparisons using rankits.					
Note: Example Data: Mercury bioaccumulation in <i>Nereis virens</i> . Amount of Censoring: 71 percent (17 out of 24 total replicates).					

agreement that the variances are unequal and increase as means increase; however, using UNIF, the variances are not significantly unequal. Assuming that the data are nonnormal, the most appropriate censored data method is CONST with rankits in any case when >40 percent of the data are censored. If one assumes that the data are normally distributed, the most appropriate censored data method for untransformed data, whether variances are equal or increase with means, is DL/2.

Because there is no clear consensus on which method should be used with these data, the wisest course might be to perform comparisons using both DL/2 with untransformed data and CONST with rankits. If the comparison results using the two methods disagree, then the statistical analysis should be considered inconclusive. Using DL/2 with untransformed data and performing *t*-tests, no treatment had significantly greater mercury bioaccumulation than the reference (SH). Using CONST with rankits, the data were found to be nonnormal with unequal variances; *t*-tests using rankits again showed no treatment to have significantly greater mercury bioaccumulation than the reference. Because the two method results agree, one can reasonably conclude that mercury bioaccumulation from the dredged sediments does not appear to be of any concern compared with the reference.

## Reference sediment data completely censored

If the reference sediment data are completely censored, and the dredged sediment data include some nondetects, the methods described in the previous section (PRELIM3.SAS) should be used. Occasionally, a situation will arise in which all of the reference sediment data, and none of the dredged sediment data, are below detection limit. When this occurs, it is not necessary to apply any censored data methods. Instead, the following steps should be performed for each dredged sediment that will be compared with the reference:

- Calculate one-sided, 95-percent lower confidence limit (LCL) for the dredged sediment sample.
- Is  $LCL < 0$ ? If yes, conclude that the dredged sediment concentrations do not exceed the lowest possible reference sediment concentrations. If no, continue with the next steps.

For single detection limits:

- Is  $LCL >$  detection limit? If yes, conclude that the dredged sediment concentrations exceed the highest possible reference sediment concentrations.
- Is LCL between 0 and detection limit? If yes, conclude that there is not enough information to determine whether the dredged sediment concentrations significantly exceed the reference sediment concentrations.

For multiple detection limits:

- If  $LCL \geq 0$ , follow the decision tree procedures in Figure D-5 of the Inland Testing Manual to compare each dredged sediment sample with the reference sediment detection limits. If the test is significant, conclude that the dredged sediment concentrations exceed the highest possible reference sediment concentrations. If the test is not significant, conclude that there is not enough information to determine whether the dredged sediment concentrations exceed the reference sediment concentrations.

As examples involving single detection limits, consider the following data comparing PCB congener concentrations in sediment RH with those in the reference sediment SH, which are all nondetects:

PCB 60	RH:	12.019	10.577	5.769	9.135	
	SH:	< 8.2	< 8.2	< 8.2	< 8.2	< 8.2
PCB 101	RH:	17.788	9.615	13.462	10.096	
	SH:	< 8.2	< 8.2	< 8.2	< 8.2	< 8.2

PCB 170	RH:	1.539	0.673	7.692	1.539	
	SH:	<8.2	<8.2	<8.2	<8.2	<8.2

The one-sided, 95-percent LCLs for RH are as follows:

PCB 60 LCL = 6.23

PCB 101 LCL = 8.30

PCB 170 LCL = -0.959

The LCL for PCB 60 is between 0 and the detection limit, so one concludes that there is not enough information to determine whether RH significantly exceeds SH. The LCL for PCB 101 is greater than the detection limit, so one concludes that RH significantly exceeds SH. The LCL for PCB 170 is less than 0, so one concludes that RH does not exceed SH.

The following data illustrate the procedure for multiple detection limits. The polynuclear aromatic hydrocarbon indeno-1,2,3-cd-pyrene (I123PY) and dibutyltin (DBT) in *Macoma nasuta* from BS and S50 exposures are to be compared with Day0 concentrations, which are all less than various detection limits:

I123PY	BS:	3.70	3.79	5.35	4.81	3.52	3.72
	S50:	3.73	6.20	4.42	4.20	3.19	3.60
	Day0:	<1.41	<1.67	<0.70			
DBT	BS:	5.8	2.9	2.3	2.5	6.6	6.1
	S50:	2.5	2.7	2.9	2.5	2.0	2.4
	Day0:	<2.4	<2.7	<2.5			

The first step is to calculate the one-sided, 95-percent LCLs for each exposure treatment:

I123PY BS LCL = 3.54  
S50 LCL = 3.35

DBT BS LCL = 2.72  
S50 LCL = 2.25

None of the LCLs is  $<0$ , so one proceeds with the bioaccumulation decision tree procedures from the Inland Testing Manual using the treatment data and Day0 detection limits. For I123PY, the residuals are lognormally distributed and the variances are unequal, so each exposure treatment is compared with Day0 detection limits by  $t$ -test using log-transformed data. Mean contaminant concentrations in both BS and S50 exposures are significantly greater than the Day0 detection limits. For DBT, the residuals are normally distributed and the variances are unequal, so the untransformed data for each exposure treatment are compared with Day0 detection limits by  $t$ -test. Mean contaminant

concentrations in BS, but not S50, are significantly greater than the Day0 detection limits. Thus, one concludes that DBT bioaccumulation from BS probably exceeds background tissue levels, but there is insufficient evidence to make that determination for DBT bioaccumulation from S50.

### Complex problems

At times, problems will arise in dredged sediment evaluations that are more complicated than the straightforward comparison of dredged sediment test results with those of a reference sediment. When such problems involve below detection limit contaminant concentrations, the complexity of analysis is further increased. It is imperative to determine exactly what questions need to be addressed, and then if possible break the complex problem down into simpler component parts.

Consider the following example. A Corps District must determine whether periodic dredging and disposal operations within a confined disposal facility (CDF) negatively impact water quality of the surrounding harbor and adjacent river. Contaminant concentrations within the CDF, from wells in the CDF dike wall, and from outside the CDF next to the dike will be compared with background concentrations from two sites in the harbor. Also, contaminant concentrations in an adjacent river downstream from the CDF filter discharge outlet will be compared with background concentrations in the river upstream from the outlet. Water samples were collected on several dates before, during, and after a dredging event, and a variety of contaminants were analyzed. The data for zinc, most of which are below detection limit, are presented in Table 18.

First, the data should be scanned for outliers. Observation 3.200 in Sample 6 is one to two orders of magnitude greater than the other observations and should be double-checked for error. Then, the problem should be broken into component parts. The downstream-upstream (CDF outlet) comparison should be considered separately from the rest, as all of the other samples will be compared with the harbor background (Samples 8a and 8b).

**Downstream-upstream comparison.** The data are divided into three discrete sampling times, which could be expected to affect the variability of the observed contaminant concentrations. Therefore, the time blocks should be included in the data analysis as follows:

- Censored data methods should be applied to the data within each time block in each treatment rather than over the treatment as a whole. This has no effect on simple substitution methods or UNIFR but does affect UNIF, the maximum likelihood methods, and the regression methods.

**Table 18**  
**Water Quality Data for Zinc (mg/ℓ); Detection Limit (dl) = 0.010 mg/ℓ**

Sampling Date	Inside CDF	Dike Wall			Adjacent to CDF			Harbor		CDF Outlet		
		Well			Sample No.						Downstream	Upstream
		4	7	9	5	6	7	8a	8b			
Before	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	
Before	< dl	< dl	0.020	< dl	< dl	< dl	< dl	< dl	< dl	0.010	< dl	
Before	0.022	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	0.044	
Before	0.028	< dl	0.027	< dl	< dl	0.012	0.016	0.012	< dl	< dl	< dl	
During	< dl	< dl	0.021	< dl	< dl	< dl	< dl	0.012	< dl	< dl	< dl	
During	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	0.011	< dl	
During	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	
During	< dl	< dl	0.030	< dl	< dl	< dl	0.013	0.016	0.011	0.034	0.012	
During	0.024	< dl	0.140	0.014	0.050	3.200	0.360	0.064	0.022	0.023	0.420	
After	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	
After	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	0.011	< dl	
After	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	
After	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	< dl	

- When the treatments are compared, a randomized block design with the sampling times as blocks should be used in the analysis of variance prior to the LSD test. This accounts for the variability due to time blocks before the mean square error term, which is used in the LSD test, is calculated. This method is illustrated in Chapter 9.

Because two of the blocks (before downstream and after upstream) are entirely  $\leq$  dl, maximum likelihood methods cannot be used and preliminary analysis should be done with the censored data methods in PRELIM3.SAS. However, UNIF and the tests of assumptions must be modified to include the time blocks; statements are provided in PRELIM3A.SAS (Appendix A). The results are given in Table 19.

After applying three simple substitution methods and UNIF in the preliminary analysis, these data are found to be nonnormally distributed with a high CV ( $>2$ ). Variances are not significantly unequal. For this situation, all censored data methods have unacceptably low power and/or high Type I error rate (from Table D-12 of the Inland Testing Manual), and statistical analysis of the data should not be performed. If variances were considered unequal and increasing as means increase, then CONST with rankits would be

**Table 19**  
**Preliminary Analysis of Zinc Water Quality Data for Downstream-Upstream Comparison**

	Preliminary Method				Consensus
	DL	DL/2	UNIF	ZERO	
Combined CV	2.83	3.29	3.30	3.92	> 1
Sampling Distribution	Nonnormal	Nonnormal	Nonnormal	Nonnormal	Nonnormal
Variances of Untransformed Data	Equal	Equal	Equal	Equal	Equal
Data Transformation Required					Rankit
Most Appropriate Censored Data Method (from Table D-12 of the Inland Testing Manual)					None
Steps Remaining After Applying Censored Data Method: Statistical analysis should not be attempted					
Note: Amount of Censoring: 71 percent (20 out of 28 total replicates).					

appropriate; the rankits would be tested for normality and equality of variances prior to comparing the two treatments.

**CDF-harbor comparisons.** Again, three time blocks are involved (Table 18). However, all data for the after time block are < dl, so this block can be eliminated from the data analysis. One concludes simply that zinc was not detected in any sample from any CDF or harbor location after the dredging event. The remainder of the data could be analyzed by comparing each sample with each of the two harbor reference samples, after blocking by date. Or the general locations (dike wall, adjacent to CDF, etc.) could be compared after blocking by date (see Chapter 9). The preliminary data analysis is presented in Table 20. The time blocks are included in the preliminary analysis as in the downstream-upstream comparison. Because some of the samples or time blocks within samples are completely censored, the methods in PRELIM3A.SAS should be used.

These data have an extremely high CV, mainly because of the outlier in Sample 6. The preliminary methods are in agreement that the residuals are nonnormal and variances unequal, increasing as means increase. The most appropriate method for these highly censored data is CONST with rankits.<sup>1</sup>

Can the CDF-harbor comparisons be further simplified prior to analysis? Perhaps they can, depending upon the objectives of the analysis. It may be reasonable to consider Wells 4, 7, and 9, indicative of the dike wall as a whole; Samples 5, 6, and 7, indicative of all the water immediately adjacent

<sup>1</sup> When sample sizes are reasonably large (10 or more replicates per treatment, for example), and each treatment includes at least three uncensored observations, regression methods for censored data, such as LR, should be considered (see Gilliom and Helsel 1986; Helsel 1990).

<b>Table 20 Preliminary Analysis of Zinc Water Quality Data for CDF-Harbor Comparisons</b>					
	Preliminary Method				Consensus
	DL	DL/2	UNIF	ZERO	
Combined CV	6.33	6.83	6.83	7.41	> 1
Sampling Distribution	Nonnormal	Nonnormal	Nonnormal	Nonnormal	Nonnormal
Variances of Untransformed Data	Unequal; increase as means increase				
Data Transformation Required					Rankit
Most Appropriate Censored Data Method (from Table D-12 of the Inland Testing Manual)					CONST
Steps Remaining After Applying Censored Data Method: Test normality and equality of variances of rankits; perform blocked comparisons using rankits					
Note: Amount of Censoring: 77 percent (69 out of 90 total replicates).					

to the CDF; and samples 8a and 8b, sufficiently representative of background conditions in the harbor as a whole. In this case, only three treatments (Inside CDF, Dike Wall, and Adjacent to CDF) need be compared with a single combined reference.

## Multiple Detection Limits

Some analytical laboratories will determine detection limits individually for each replicate depending upon the amount of sample available for analysis. A treatment for which several nondetects are reported may then include several different detection limits. Multiple detection limits were not included in the simulation study used to assess censored data methods. Until censored data methods can be statistically evaluated for multiple detection limits, the Inland Testing Manual recommends using the same procedures as for single detection limits. SAS programs for the censored data methods discussed in this chapter can be performed for multiple as well as single detection limit data; modification of program statements is unnecessary for multiple detection limits. Earlier in this chapter a procedure was described for multiple detection limits in the case where all of the reference sediment data and none of the dredged sediment data are censored. The following sections provide procedures for more general situations of moderately to highly censored data involving multiple detection limits.

## Moderately censored data

The moderately censored data for cadmium bioaccumulation in *Macoma nasuta* (Table 7 in Chapter 5) are an example of multiple detection limits. Preliminary analysis for these data was performed using PRELIM2.SAS; results are given in Table 21.

<b>Table 21</b>					
<b>Steps in Selecting the Most Appropriate Censored Data Method: Moderately Censored Data with Multiple Detection Limits</b>					
	Preliminary Method				Consensus
	DL	DL/2	MLE WEIB	UNIF	
Combined CV	0.81	0.82	0.82	0.82	Between 0.26 and 1
Sampling Distribution	Normal	Normal	Normal	Normal	Normal
Variances of Untransformed Data	Equal	Equal	Equal	Equal	Equal
Data Transformation Required					None
Most Appropriate Censored Data Method (from Table D-12 of the Inland Testing Manual)					DL/2
Steps Remaining After Applying Censored Data Method: Compare untransformed data using LSD test					
Note: Example Data: Cadmium bioaccumulation in <i>Macoma nasuta</i> . Amount of Censoring: 25 percent (6 out of 24 total replicates).					

The preliminary methods are in agreement that the residuals are normally distributed and variances are equal, with a combined CV of about 0.8. The censored data method recommended for this situation with single detection limit is DL/2, so this method should be used.

As another example of moderately censored data with multiple detection limits, consider the following data for PCB 52 bioaccumulation in *Nereis virens*:

Sediment AK:	12.0	<0.64	8.7	16.0	8.0	
Sediment GOW:	12.0	<0.90	13.0	11.0	8.9	
Sediment RH:	5.9	<0.96	4.5	<0.98	<0.92	<0.75
Sediment SH:	4.3	3.5	<1.2	<0.82	<1.0	<0.64

The methods in PRELIM2.SAS may be used; however, MLE WEIB must be modified because sample sizes are unequal. For these data, MLE WEIB is run first for AK and GOW using quantiles for five replicates ( $Q = .01 .25 .5 .75 .99$ ), and then rerun for RH and SH using quantiles for six replicates ( $Q = .01 .21 .4 .6 .79 .99$ ). The output data sets are then combined. SAS statements are provided in PRELIM2A.SAS in Appendix A. Table 22 shows results of the preliminary analysis.

**Table 22**  
**Steps in Selecting the Most Appropriate Censored Data Method: Moderately Censored Data with Multiple Detection Limits**

	Preliminary Method				Consensus
	DL	DL/2	UNIF	MLE WEIB	
Combined CV	0.95	1.02	1.05	1.01	Approximately 1
Sampling Distribution	Normal	Normal	Normal	Normal	Normal
Variances of Untransformed Data	Equal	Equal	Equal	Equal	Equal
Data Transformation Required					None
Most Appropriate Censored Data Method (from Table D-12 of the Inland Testing Manual)					DL/2
Steps Remaining After Applying Censored Data Method: Perform comparisons using LSD test					
Note: Example Data: PCB 52 bioaccumulation in <i>Nereis virens</i> . Amount of Censoring: 45 percent (10 out of 22 total replicates).					

The preliminary methods are in agreement that the data are normally distributed with equal variances and a combined CV of approximately 1. No data transformation is required, and DL/2 is the most appropriate censored data method. Nevertheless, the P-values for Shapiro-Wilk's Test using untransformed data (and log-transformed data as well) were all just slightly greater than  $\alpha$  (0.01, balanced design,  $N = 20$  or more). Thus, considering the high CV, it might be more appropriate to assume that these data are non-normal, in which case the recommended censored data method is CONST with rankits. The rankits would then be tested for normality and equality of variances prior to comparing treatments.

### Highly censored data

Consider the following highly censored data with multiple detection limits for PCB 44 bioaccumulation in *Macoma nasuta*:

Sediment AK:	<0.71	<0.98	4.2	4.9	<1.0	3.0
Sediment GOW:	1.4	<0.72	<1.1	<0.64	2.4	<0.81
Sediment RH:	<0.79	<0.67	<0.69	<0.75	<0.53	<0.83
Sediment SH:	<0.96	<0.59	<0.64	<0.70	<0.56	<0.58

The data for two treatments (RH and SH) are entirely below detection limits, so PRELIM3.SAS should be used for preliminary analysis. Results are given in Table 23.

**Table 23**  
**Steps in Selecting the Most Appropriate Censored Data Method: Highly Censored Data with Multiple Detection Limits**

	Preliminary Method				Consensus
	DL	DL/2	UNIF	ZERO	
Combined CV	0.93	1.36	1.33	2.18	Probably > 1
Sampling Distribution	Lognormal	Lognormal	Normal	Nonnormal	?
Variances of Untransformed Data	Unequal; increase as means increase				
Data Transformation Required					?
Most Appropriate Censored Data Method (from Table D-12 of the Inland Testing Manual)					DL/2 with untransformed or log-transformed data; CONST with rankits
Steps Remaining After Applying Censored Data Method: Apply DL/2 and log-transform; test equality of variances of logs; perform comparisons using logs					
Note: Example Data: PCB 44 bioaccumulation in <i>Macoma nasuta</i> . Amount of Censoring: 79 percent (19 out of 24 total replicates).					

The preliminary methods applied to these data are in agreement that variances are unequal, increasing as means increase, with an overall CV of approximately 1 or greater. However, there is no agreement concerning the underlying data distribution, which is not surprising considering the severity of censoring. Nevertheless, knowledge of the distribution may not really be necessary in this case. From Table D-12 of the Inland Testing Manual, when data are >60-percent censored and variances increase as means increase, the recommended method is DL/2 for either untransformed normal data or log-transformed lognormal data. One has seen in the preliminary analysis that applying DL/2 to these data results in a finding of lognormality. The data should now be log-transformed and tested for equality of variances prior to comparing treatments.

## Nondetects and Estimation

When contaminant concentrations are being compared among treatments, as in dredged sediment-reference sediment bioaccumulation comparisons, it is generally not necessary to estimate the population parameters' such as mean and standard deviation. However, other situations may arise in which parameter estimates are desired. When tissue contaminant concentrations must be compared with an action level, for example, accurate estimates of the mean and standard deviation are needed. If the data are uncensored, the sample

mean  $\bar{x}$  and standard deviation  $s$  are unbiased estimates of the population mean  $\mu$  and standard deviation  $\sigma$ . “Unbiased” means that over all possible random samples from a population, the average difference between the sample estimate and the population parameter will be zero. If the data include non-detects, then estimating  $\mu$  and  $\sigma$  becomes more difficult, necessitating use of a censored data method. The resulting estimates will not be unbiased, and as censoring increases, the amount of bias introduced will also tend to increase.

A number of earlier studies have examined statistical estimation with censored data. Those that compared censored data methods for estimation using actual or simulated chemical concentration data generally evaluated relatively large samples of 10 or more replicates (e.g., El-Shaarawi 1989; Gilliom and Helsel 1986; Haas and Scheff 1990; Helsel and Cohn 1988; Helsel and Gilliom 1986; Newman et al. 1989). The methods that performed best for estimation in these studies were generally maximum likelihood or log regression techniques. Gleit (1985) examined some censored data methods for estimating the mean and variance of normal populations from samples as small as  $n = 5$  and recommended an iterative method using expected values of order statistics.

### Evaluating censored data methods for estimation

The 10 censored data methods described earlier in this chapter were evaluated for estimation accuracy as part of the simulation study for small samples from normal and nonnormal populations. After a censored data method is applied to random samples from a censored simulated population, how well do the sample  $\bar{x}$  and  $s$  approximate  $\mu$  and  $\sigma$ ? Estimation was evaluated using two measures: the average bias, defined as the average difference between the sample estimate and the population parameter, and the root mean square error (rmse), defined as the average

$$\sqrt{\sum \left( \frac{\bar{x} - \mu}{\mu} \right)^2} \quad (1)$$

The rmse is a measure of the amount of deviation from the population parameter, regardless of the direction of the deviation, summed over all samples from the population. Bias indicates whether on average the sample statistic overestimates or underestimates the population parameter. The best censored data method for estimation in a given situation will result in lowest rmse and bias closest to zero.

Average bias and rmse were determined for three distributions (normal, lognormal, gamma), four CV ranges ( $\leq 0.25$ , 0.25 to 0.5, 0.5 to 1, and  $> 1$ ), and five ranges of censoring ( $\leq 20$  percent, 21 to 40 percent, 41 to 60 percent, 61 to 80 percent, and  $> 80$  percent). Sample sizes ranged from three to eight replicates. In general, results differ less among distributions than among

the CV ranges. Estimation accuracy tends to decline as the CV increases and as censoring increases. The best methods over all distributions are presented in Table 24; recommendations are based primarily on the methods having lowest rmse. Note that in most cases, different methods are recommended for estimating mean and standard deviation. Furthermore, the censored data methods recommended for estimation are generally not the ones recommended for hypothesis testing.

**Table 24**  
**Recommended Censored Data Methods for Small Samples To Be Used in Estimation**

Amount of Censoring	Coefficient of Variation	Estimation of Mean	Estimation of Standard Deviation
≤20 percent	≤0.25	DL	DL
	0.25 - 0.5	DL	DL/2
	0.5 - 1	DL	ZERO (DL/2)
	>1	DL/2	UNIFR
21 - 40 percent	≤0.25	DL	DL
	0.25 - 0.5	MLE LOGN (DL)	DL/2
	0.5 - 1	MLE LOGN (DL/2)	ZERO (UNIFR)
	>1	DL/2	ZERO (UNIFR)
41 - 60 percent	≤0.25	MLE LOGN (DL)	MLE LOGN (DL)
	0.25 - 0.5	MLE LOGN (DL/2)	DL/2
	0.5 - 1	MLE WEIB (DL/2)	UNIFR
	>1	DL/2	ZERO (UNIFR)
61 - 80 percent	≤0.25	MLE LOGN (DL)	MLE LOGN (DL)
	0.25 - 0.5	MLE WEIB (DL/2)	UNIFR
	0.5 - 1	DL/2	UNIFR
	>1	DL/2	UNIFR

Situations may arise in which parameter estimates are desired over several treatments combined, but one or more of the treatments are entirely below detection limit. MLE methods cannot be used for those treatments, and the method listed in parentheses in Table 24 should be used instead. Likewise, ZERO cannot be used when the data will be transformed to logarithms, and the method listed in parentheses in Table 24 should be substituted.

Rmse and bias of the six methods in Table 24 are shown in Figures 38-41 for estimation of population mean and standard deviation when results are averaged over all distributions. When the CV is low (Figure 38), the rmse for estimating  $\sigma$  differs greatly among the methods. As the CV increases, the

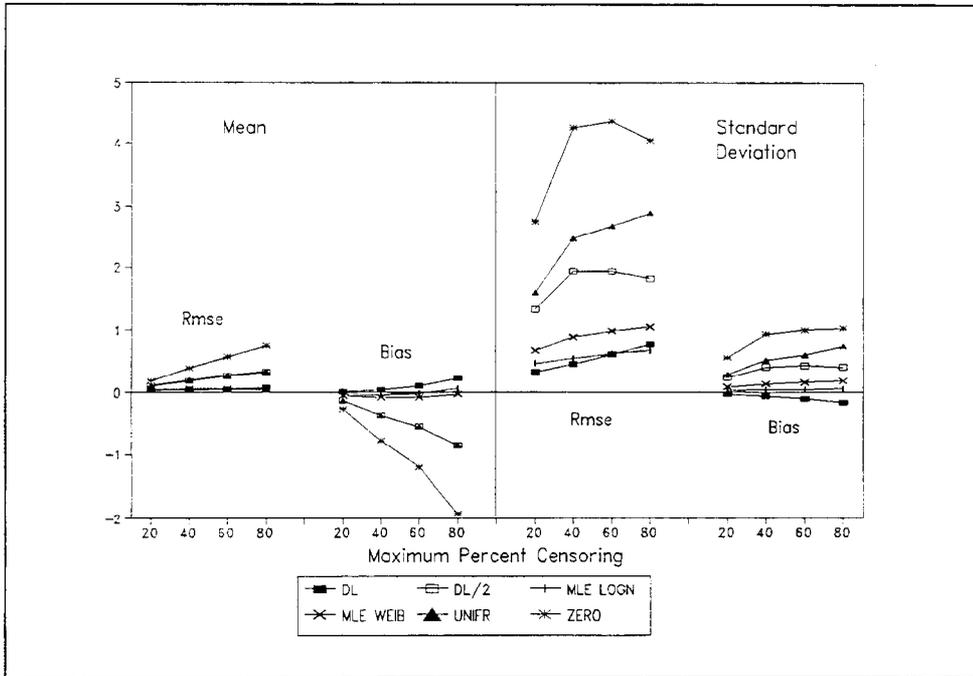


Figure 38. Accuracy of six censored data methods for estimating population mean and standard deviation when  $CV \leq 0.25$

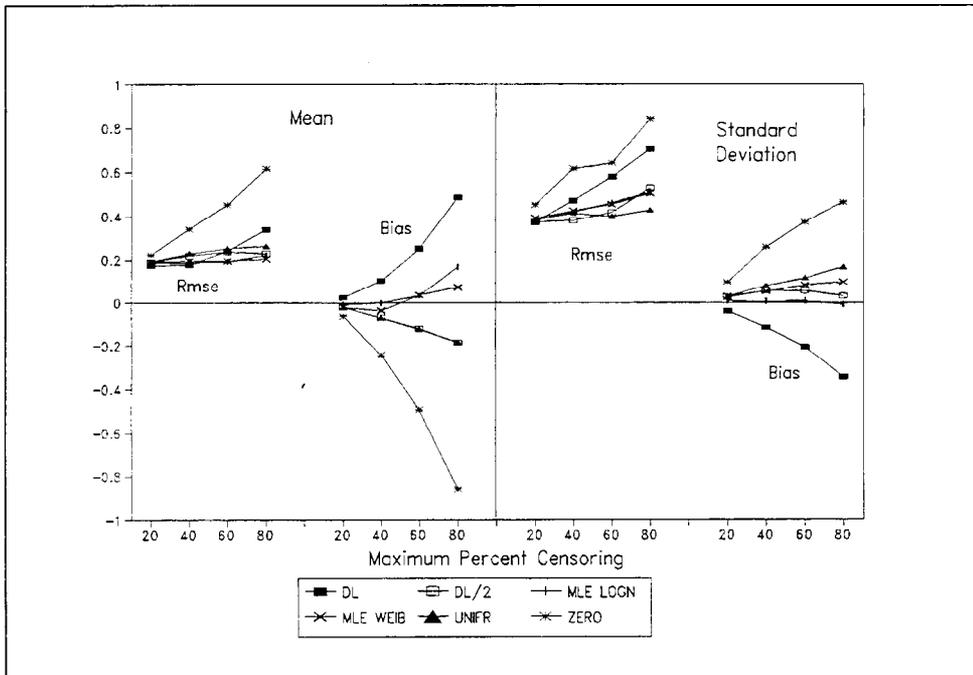


Figure 39. Accuracy of six censored data methods for estimating population mean and standard deviation when  $CV > 0.25$  and  $\leq 0.5$

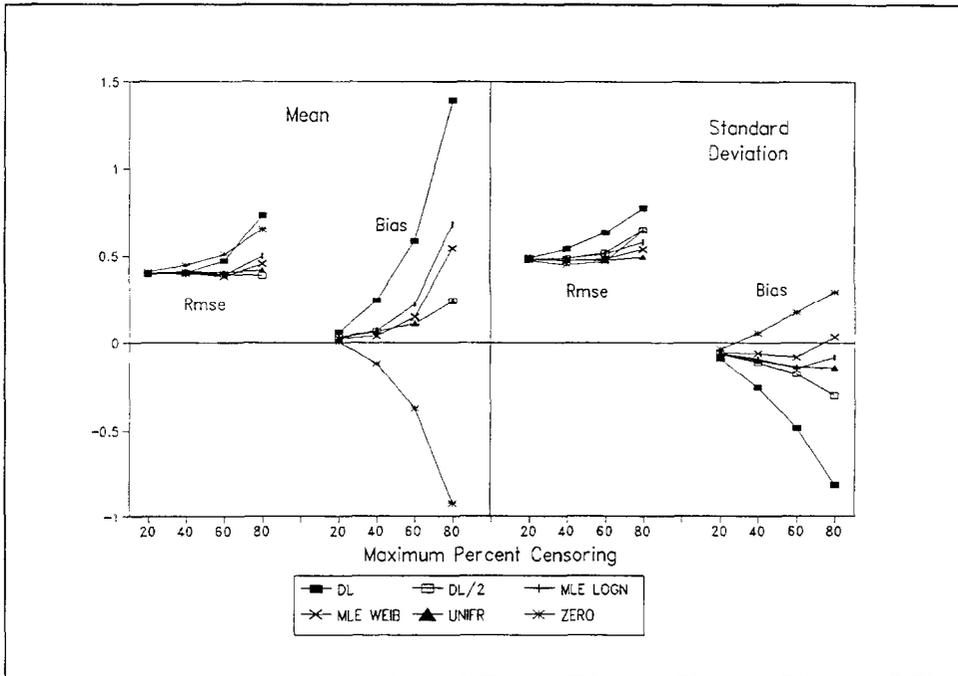


Figure 40. Accuracy of six censored data methods for estimating population mean and standard deviation when  $CV > 0.5$  and  $\leq 1$

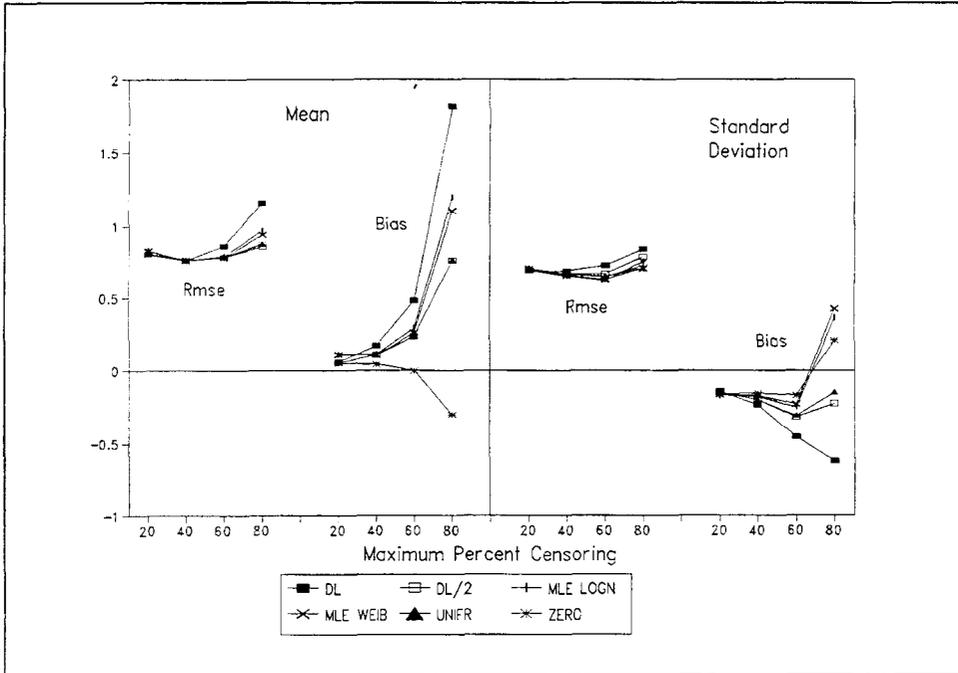


Figure 41. Accuracy of six censored data methods for estimating population mean and standard deviation when  $CV > 1$

differences in rmse among methods becomes much less for estimating both  $\mu$  and  $\sigma$  (Figure 41). Bias of some methods, notably DL and ZERO, becomes greatly exaggerated as censoring increases, especially when the CV exceeds 0.5.

As an alternative to the methods in Table 24, one may wish to use the “fill-in with expected values” method recommended by Gleit (1985) for estimation from small samples. This method is one of several censored data methods available in the public domain software program UNCENSOR (Newman and Dixon 1990).

### Comparisons with an action level

On occasion it may be necessary to compare tissue contaminant concentrations with an FDA action level (or with some other numeric criterion), as described in Section D3.1.2 of the Inland Testing Manual. When some contaminant concentrations are below detection limit, the data should first be assessed for percent of censoring and probable CV. If necessary, two or more censored data methods may be applied in a preliminary analysis to determine the likely CV. Then the most appropriate censored data methods to use in estimating mean and standard deviation should be selected from Table 24. SAS program statements for the six recommended methods are provided in ESTIMATE.SAS (Appendix A). The estimated mean and standard deviation may then be used in a one-sample *t*-test, or equivalently, in determining the one-sided, 95-percent upper confidence limit (UCL) for comparison with the action level.

As an example, the following data for total PCB bioaccumulation in *Macoma nasuta* and *Nereis virens* might be compared with the FDA action level of 2  $\mu\text{g/g}$  total PCBs in edible tissue:<sup>1</sup>

#### *Macoma nasuta*:

Sediment AK	1.10	3.00	2.50	2.00	1.70	2.60
Sediment GOW	<0.06	2.00	2.00	1.00	1.80	2.00
Sediment RH	2.00	0.90	1.30	0.90	1.90	1.60
Sediment SH	<0.06	0.80	0.80	0.40	<0.06	0.50

#### *Nereis virens*:

Sediment AK	7.50	1.90	<0.06	3.60	1.80	
Sediment GOW	4.00	2.30	2.30	3.00	<0.06	
Sediment RH	2.00	12.00	1.20	<0.06	1.90	1.40
Sediment SH	0.90	0.80	<0.06	<0.06	<0.06	<0.06

<sup>1</sup> The actual data for this example were multiplied by 10 to put the data in the same range as the action level.

Total PCB bioaccumulation in *Macoma nasuta* exposed to sediments AK and RH is above detection limit for all replicates, so these treatments are compared directly with the action level using either the one-sample *t*-test or the UCL according to Section D3.1.2 of the Inland Testing Manual. PCB bioaccumulation in *Macoma nasuta* from sediments GOW and SH, and PCB bioaccumulation in *Nereis virens* from all four sediments include nondetects. Preliminary analysis is recommended for these treatments to determine the best censored data methods for estimation. This analysis is conducted for each censored treatment using ESTIMATE.SAS (Appendix A). Results are provided in Table 25.

Species	Sediment	Number of Replicates	Percent Censored	Coefficient of Variation Using Censored Data Method					
				DL	DL/2	MLE LOGN	MLE WEIB	UNIFR	ZERO
<i>Macoma nasuta</i>	AK	6	0	0.32	0.32	0.32	0.32	0.32	0.32
	GOW	6	16.7	0.54	0.55	0.55	0.54	0.54	0.56
	RH	6	0	0.34	0.34	0.34	0.34	0.34	0.34
	SH	6	33.3	0.76	0.81	0.81	0.81	0.80	0.86
<i>Nereis virens</i>	AK	5	20	0.95	0.95	0.96	0.96	0.96	0.96
	GOW	5	20	0.62	0.63	0.63	0.63	0.63	0.63
	RH	6	16.7	1.43	1.43	1.43	1.44	1.43	1.44
	SH	6	66.7	1.27	1.40	1.48	1.47	1.32	1.55

For the uncensored treatments, application of censored data methods is irrelevant, and the CV of course is the same regardless of method. For the treatments that include nondetects, all six censored data methods result in similar CVs. The censored data method selected and the results of the UCL-action level comparisons are shown in Table 26. ESTIMATE.SAS calculates mean and variance for each treatment using each of the six censored data methods and provides a table *t* value for  $1 - \alpha = 0.95$  and  $n - 1$  degrees of freedom. The UCL (Equation 21 in Appendix D of the Inland Testing Manual) can then be determined easily with a hand calculator using the *t* value provided by ESTIMATE.SAS, along with the mean and variance calculated using the most appropriate censored data method(s).

For several treatments (*Macoma nasuta*, AK; *Nereis virens*, AK, GOW, RH), the mean exceeds the action level regardless of which censored data method is used, so no further analysis is really necessary. For three treatments (*Macoma nasuta*, RH, SH; *Nereis virens*, SH), the mean and the UCL are less than the action level, resulting in the conclusion that mean PCB

**Table 26**  
**Comparison of Total PCB Bioaccumulation With FDA Action Level of 2 µg/g**  
**Following Application of Censored Data Methods**

Species	Sediment	Censoring Percent	CV Range	Recommended Censored Data Method		Mean Concentration µg/g	Std. Dev. µg/g	95 Percent One-Sided UCL	Comparison Conclusion
				Mean	Std. Dev.				
<i>Macoma nasuta</i>	AK	0	0.25 - 0.5	none	none	2.150	0.689	2.72	Mean > action level
	GOW	17	0.5 - 1	DL	ZERO	1.477	0.817	2.15	Mean not significantly < action level
	RH	0	0.25 - 0.5	none	none	1.433	0.480	1.83	Mean significantly < action level
	SH	33	0.5 - 1	MLE LOGN	ZERO	0.427	0.360	0.72	Mean significantly < action level
<i>Nereis virens</i>	AK	20	0.5 - 1	DL	ZERO	2.972	2.840	5.68	Mean > action level
	GOW	20	0.5 - 1	DL	ZERO	2.332	1.472	3.38	Mean > action level
	RH	17	> 1	DL/2	UNIFR	3.088	4.423	6.73	Mean > action level
	SH	67	> 1	DL/2	UNIFR	0.303	0.415	0.64	Mean significantly < action level

bioaccumulation from these treatments is significantly less than the action level. For *Macoma nasuta* exposed to GOW, mean bioaccumulation is below the action level, but the UCL exceeds the action level. The conclusion would be that PCB bioaccumulation from this sediment may exceed the action level.

## Summary

Statistical analyses required as part of dredged sediment evaluations cannot be applied directly to contaminant concentration data when some observations

are reported as less than detection limit (“censored” data). Generally, the unknown observations must first be replaced by some numeric value using a censored data method. Many such methods are available. A simulation study was conducted to determine which methods work best for statistical comparisons involving small sample size.

- Censored data methods recommended for statistical comparisons of small samples include simple substitution methods (DL, DL/2, ZERO), substitution of evenly spaced values between zero and the detection limit (UNIF), maximum likelihood estimation methods (MLE NORM and MLE WEIB), and a log regression method (LR).
- The most appropriate censored data method for a given situation depends upon the amount of censoring, the underlying probability distribution, the coefficient of variation, and the pattern of variances among the treatments (equal, increasing as means increase, or mixed).
- Recommended censored data methods are listed in Section D3.1.1.1 (Table D-12) of the Inland Testing Manual Appendix D, and in Clarke (1995a), along with the steps for selecting the most appropriate method.
- To best accomplish the steps for selecting the most appropriate censored data method, two or more of those methods should be applied to the data in a preliminary analysis to estimate the most likely data distribution (normal, lognormal, or nonnormal), and to obtain a range of possible CVs and variances.
- When the data are nonnormal and must be converted to rankits for statistical comparisons, the censored data methods DL, DL/2, and ZERO are equivalent and may be referred to as CONST for substitution of any constant between zero and the detection limit.
- SAS programs for doing the preliminary analysis are provided in Appendix A. The situations in which each program should be used are summarized in Table 27.
- Censored data methods recommended for estimating mean and standard deviation of small censored samples include DL, DL/2, ZERO, MLE LOGN, MLE WEIB, and UNIFR. The best method depends upon the amount of censoring and the sample CV. Recommendations are given in Table 25. Recommended methods are generally not the same as the ones that would be used for comparing treatments.
- Recommended methods for estimation should be used in comparisons of censored contaminant concentration data with an action level.

**Table 27**  
**Summary of SAS Programs (Appendix A) for Preliminary Analysis of Censored Data**

Data Situation	SAS Program <sup>1</sup>	Censored Data Methods Included
One nondetect in one or more treatments	PRELIM1.SAS	DL, DL/2, ZERO
Several nondetects in one or more treatments (< 50 percent censoring) •Sample sizes unequal	PRELIM2.SAS •PRELIM2A.SAS	DL, DL/2, MLE WEIB, UNIF
Many nondetects (> 50 percent censoring); no treatment completely censored •Sample sizes unequal	PRELIM2.SAS •PRELIM2A.SAS	DL, DL/2, MLE WEIB, UNIF
Mostly nondetects (> 80 percent censoring)	No statistical analysis	--
One or more treatments completely censored	PRELIM3.SAS	DL, DL/2, UNIF, ZERO
Data blocked within treatments	PRELIM3A.SAS	DL, DL/2, UNIF, ZERO
Reference sediment data completely censored; dredged sediment data uncensored	Use LCL approach	--
Comparisons with action level or other problems involving estimation of mean and standard deviation	ESTIMATE.SAS	DL, DL/2, MLE LOGN, MLE WEIB, UNIFR, ZERO

<sup>1</sup> Programs are applicable for either single or multiple detection limits.

# 9 Interpreting Statistical Test Results

---

In most cases, the interpretation of test results for the types of statistical analyses recommended in the Inland Testing Manual is reasonably straightforward. By looking at the test statistic and its associated probability (P-value) one can usually determine without difficulty whether the null hypothesis should be rejected. This chapter is devoted to the interpretation of statistical test results in the occasional situations where the test statistics or SAS program output may be confusing or ambiguous. Examples showing analysis and interpretation of the blocked design are also presented.

## Interpreting P-Values

When statistical tests were performed without the aid of computers, the significance of a test statistic was determined by comparing the statistic with the table value at a given significance level, e.g.,  $\alpha = 0.05$  or  $\alpha = 0.01$ . If the calculated  $t$  statistic, for example, was greater than the table value for the appropriate degrees of freedom and  $\alpha = 0.05$ , the test was said to be “significant at the 5-percent level.” Now most statistical tests are performed using computer software packages, in which computation of exact P-values is common. The computed P-values, which range from 0 to 1, can be used as an indication of the degree of confidence in rejecting or accepting the null hypothesis. A P-value of 0.0001 is a strong indication of a “highly significant” test result, i.e., a high amount of confidence in rejecting the null hypothesis. On the other hand, a P-value of 0.9999 indicates not only that the test result is not significant, but that the sample data provide strong evidence for accepting the null hypothesis.

P-values should not be confused with the power of the test. P actually refers to the probability of getting a more extreme value than the calculated statistic if the null hypothesis is true. For example, a very high (or very low negative)  $t$  statistic will have a P-value close to zero. When P is very low, there is little chance of obtaining a more extreme value of the test statistic than the one calculated from the sample data, if the null hypothesis is true. In other words, the null hypothesis is most likely false. Conversely, when P is

very high, there is a large probability of obtaining a more extreme value of the test statistic than the one calculated, assuming the null hypothesis is true. Thus, the null hypothesis most likely is true.

Interpretation problems arise when the P-value is close to the predetermined significance level of the test ( $\alpha$ ). If  $\alpha = 0.05$  and  $P = 0.049$  or  $0.051$ , is the test significant or not? A test where  $P = 0.049$  is obviously not as significant as a test where  $P = 0.0001$ . On the other hand, a test with  $P = 0.051$ , while strictly speaking not significant, should leave the investigator feeling uneasy. For illustration, consider the Chapter 8 preliminary analysis of zinc water quality data for the downstream-upstream comparison (Table 19). When equality of variances was tested after applying four censored data methods in the preliminary analysis, the test statistics and P-values for Levene's Test were as follows:

DL	$F = 2.82$	$P = 0.1061$
DL/2	$F = 2.89$	$P = 0.1022$
UNIF	$F = 2.88$	$P = 0.1028$
ZERO	$F = 2.95$	$P = 0.0989$

Suppose the significance level for this test was 0.10 instead of 0.05 (from Table D-2 of the Inland Testing Manual,  $n \geq 10$ , balanced design, or  $n < 10$ , unbalanced design). Then, regardless of which censored data method is used, the P-value for the equality of variances test would be approximately equal to  $\alpha$ . Would  $P$  be  $> \alpha$  (variances equal) or would  $P$  be  $< \alpha$  (variances unequal)? Concluding that variances were equal would result in a recommendation that no statistical analysis be performed, whereas concluding that variances were unequal would result in a recommendation that statistical comparisons be performed on rankits following substitution of CONST for the nondetects. However,  $P$  is so close to  $\alpha$  in this case that variances could reasonably be considered unequal. One would then proceed with the analysis using rankits.

As another example, consider the following data for bioaccumulation of PCB congeners 56+60 in *Mytilus edulis* exposed to a contaminated Oakland Harbor sediment. Uptake (ng/g) from bedded sediment (BS) was compared with uptake from 50-mg/l suspended sediment (S50):

BS:	2.1	1.8	2.4	0.6	2.2	2.1	$\bar{x} = 1.87$	$s = 0.65$
S50:	3.0	3.4	2.2	3.2	1.8	2.6	$\bar{x} = 2.70$	$s = 0.62$

These data were found to be normally distributed, with equal variances. The  $t$ -test results were  $t = -2.278$  and  $P = 0.0459$ , with the conclusion that mean bioaccumulation from S50 was significantly greater than mean bioaccumulation from BS. However, a P-value that is so close to  $\alpha$  (0.05) should signal caution in interpreting the test results. The biological or ecological significance of the amount of difference between treatment means should certainly be considered. In this case, one might reasonably conclude that suspended sediment (S50) exposure does enhance PCB 56+60 uptake in mussels

compared with bedded sediment exposure, and this is consistent with the fact that the mussels filter suspended particles but do not ingest bedded sediment. When statistical comparisons are performed as part of dredged sediment evaluations, the magnitude of the statistical test P-values should be included as part of the information input into decision making concerning dredging and disposal options.

## When Different Tests on the Same Data Have Different Results

When two or more equivalent tests are performed on the same data, the test conclusions can and often will differ from each other, especially when the P-values are close to the significance level of the tests. For example, one sees in Chapter 6 that four tests are generally acceptable for checking the equality of variances assumption. Because each test uses different calculations, test statistics and P-values will certainly differ among the tests, and the conclusions may differ as well. This was quite clear from the analysis of eight example data sets (Chapter 6, Table 13). There is usually no reason to perform more than one test, and the investigator will logically choose whatever test is provided in the available statistical software. However, sometimes the software package will provide results for more than one test. SAS, for example, offers a wide variety of multiple comparison tests, any number of which can be requested in the analysis of variance (ANOVA and GLM) procedures (SAS Institute, Inc. 1988a).

If the results of two or more equivalent tests corroborate each other, confidence in the test results is strengthened. However, when results disagree, several factors should be considered:

- Is one test more appropriate for the particular data set than the other tests? For example, when the data are nonnormal, Levene's Test is more appropriate for checking equality of variances than Bartlett's, Hartley's, or Cochran's tests, providing the data have been appropriately transformed for the type of distribution (see Chapter 6).
- Is one test known to be more powerful than the alternative tests? For example, the LSD test is considered more appropriate for dredged sediment-reference sediment comparisons than other tests such as Dunnett's Test, because the LSD test does not control experimentwise error rate and therefore has more power.
- Is one test more appropriate than another because a test assumption is violated? For example, some statistical packages provide *t*-test results for both equal and unequal variances, and those results will often differ from each other. If the data have failed the test for equality of variances, then the *t*-test results for unequal variances should be accepted.

## When LSD Is Significant and ANOVA Is Not or Vice Versa

Analysis of variance (ANOVA) is a global test for differences among treatment means, although it does not identify which means differ if the test is significant. Therefore, means comparison tests are commonly employed following a significant ANOVA to determine which treatments differ. ANOVA, by definition, controls experimentwise error rate, as do many of the means comparison procedures. However, in dredged sediment evaluations, it is unnecessary to control experimentwise error rate when an independent decision will be made for each dredged sediment or “management unit” included in a comparison test. The LSD test controls pairwise or comparisonwise error rate, rather than experimentwise error rate, and thus has more power for individual comparisons. Therefore, the LSD test will sometimes identify significant differences when the ANOVA that preceded it was not significant.

On the other hand, the ANOVA results will sometimes be significant when the LSD test has not identified any dredged sediment treatment as significantly worse than the reference sediment or control treatment. This can easily happen because the ANOVA  $F$  test is a two-tailed test that will be significant whenever there is a sufficient difference in either direction between any two or more treatments. The LSD test should also identify all such significant differences, but in dredged sediment evaluations the only differences of interest are for one-directional comparisons between the reference sediment and each dredged sediment. Differences that are not of interest are simply ignored.

For dredged sediment evaluations, the ANOVA is generally needed only as an easy means to calculate the mean square error (MSE), an estimate of pooled variance across all treatments. The MSE and its associated degrees of freedom are then used in the calculation of the LSD  $t$  statistic for each comparison of interest. The ANOVA  $F$  statistic and its  $P$ -value can safely be ignored.

### Interpreting LSD Output

Some statistical software packages, including SAS, will print letters next to the treatment means to indicate groups of means (called “T Groupings” in SAS) in the LSD test output. Means with the same letter are not significantly different from each other. However, when sample sizes are unequal, SAS provides confidence intervals for the difference between means, instead of letters (SAS Institute, Inc. 1988a). Consider the following data for lead bioaccumulation in *Macoma nasuta* and *Macoma secta* exposed to three dredged sediments (AK, GOW, RH) and a reference sediment (SH):

<i>Macoma nasuta</i>	AK:	2.70	2.23	2.12	2.51	1.31	1.82
	GOW:	1.50	1.52	1.37	1.43	0.95	1.23
	RH:	1.01	2.32	1.63	2.63	1.57	1.27
	SH:	1.97	0.52	2.15	1.52	0.35	2.24
<i>Macoma secta</i>	AK:	0.24					
	GOW:	0.80	0.80				
	RH:	0.76	1.64				
	SH:	0.68	0.40	0.22	0.54	0.30	0.28

For both species, the data residuals are normally distributed and variances are unequal among treatments. Thus, bioaccumulation from the dredged sediments would be compared with bioaccumulation from the reference sediment using *t*-tests. However, LSD results will be presented here for the sake of illustration.

For *Macoma nasuta*, the SAS LSD output is as follows:

```

General Linear Models Procedure

T tests (LSD) for variable: CONC

NOTE: This test controls the type I comparisonwise
      error rate not the experimentwise error rate.

Alpha= 0.1  df= 20  MSE= 0.343233
      Critical Value of T= 1.72
      Least Significant Difference= 0.5834

Means with the same letter are not significantly
different.
```

T Grouping	Mean	N	SEDIMENT
A	2.115	6	AK
B	1.738	6	RH
B	1.458	6	SH
B	1.333	6	GOW

The output provides  $\alpha$ , the degrees of freedom (df), and the MSE used in calculating *t* for each comparison. In the SAS GLM and ANOVA procedures, the default  $\alpha = 0.05$  is for a two-tailed test; this corresponds to  $\alpha = 0.025$  in most textbook *t* tables. For a one-tailed test at the 5-percent significance level in SAS, ALPHA = 0.1 must be specified as an option. The critical value of T is the value that the calculated  $|t|$  must exceed for a comparison to be considered significant. The least significant difference is the amount of difference between means that the test can detect with a power of 0.5. From the output, it is clear that dredged sediment AK (T Grouping A) is significantly greater than the reference sediment SH (T Grouping B), while SH and dredged sediment GOW do not differ significantly. SH and dredged sediment RH, because they share the letter B, also do not differ significantly. AK is significantly greater than GOW but not RH; however, these comparisons between dredged sediments are not of interest. The ANOVA *F* for these data

was 2.10 with  $P = 0.1322$ , illustrating a case in which the ANOVA would not be considered significant although the LSD test does identify significant differences among treatment means.

For *Macoma secta*, the SAS LSD output is as follows:

```

General Linear Models Procedure

T tests (LSD) for variable: CONC

NOTE: This test controls the type I comparisonwise
      error rate not the experimentwise error rate.

Alpha= 0.1  Confidence= 0.9  df= 7  MSE= 0.077419
      Critical Value of T= 1.89458

Comparisons significant at the 0.1 level are indicated by
'***'.

```

SEDIMENT Comparison	Lower Confidence Limit	Difference Between Means	Upper Confidence Limit	
RH - GOW	-0.127	0.400	0.927	
RH - SH	0.366	0.797	1.227	***
RH - AK	0.314	0.960	1.606	***
GOW - RH	-0.927	-0.400	0.127	
GOW - SH	-0.034	0.397	0.827	
GOW - AK	-0.086	0.560	1.206	
SH - RH	-1.227	-0.797	-0.366	***
SH - GOW	-0.827	-0.397	0.034	
SH - AK	-0.406	0.163	0.733	
AK - RH	-1.606	-0.960	-0.314	***
AK - GOW	-1.206	-0.560	0.086	
AK - SH	-0.733	-0.163	0.406	

Again, the output displays  $\alpha$ ,  $df$ ,  $MSE$ , and the critical value of  $t$ . The least significant difference is not provided because it is different for each comparison due to unequal sample sizes. Means and letters are replaced by the difference between means and 90-percent confidence intervals about that difference for each pairwise comparison. Treatments are presented in order of decreasing means. Differences that are considered significant are indicated by \*\*\*. Any confidence interval (i.e., lower confidence limit to upper confidence limit) that does not include zero will be significant. Here, RH is significantly greater than SH (and AK, although that comparison is not of interest). The last two treatment groups simply repeat the same information, namely, that SH and AK are significantly less than RH. For this data set, the ANOVA  $F$  was 5.00 with  $P = 0.0367$  so the ANOVA would have been considered significant as well.

## When Transformation Changes the Order of Treatment Means

Data transformation changes the scale on which data points are ordered and will sometimes change the order of treatment means as well, especially when

the untransformed means are not widely separated. Occasionally, this will change which treatments differ significantly in the LSD test. An example is the data for lead bioaccumulation in *Macoma nasuta* presented in the previous section. LSD outputs for untransformed data (CONC), log-transformed data (LOGCONC), and rankits (RANKIT) follow:

General Linear Models Procedure

T tests (LSD) for variable: **CONC**

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.1 df= 20 MSE= 0.343233  
 Critical Value of T= 1.72  
 Least Significant Difference= 0.5834

Means with the same letter are not significantly different.

T Grouping	Mean	N	SEDIMENT
A	2.115	6	AK
B	1.738	6	RH
B	1.458	6	SH
B	1.333	6	GOW

General Linear Models Procedure

T tests (LSD) for variable: **LOGCONC**

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.1 df= 20 MSE= 0.041384  
 Critical Value of T= 1.72  
 Least Significant Difference= 0.2026

Means with the same letter are not significantly different.

T Grouping	Mean	N	SEDIMENT
A	0.314	6	AK
B	0.217	6	RH
B	0.120	6	GOW
B	0.070	6	SH

General Linear Models Procedure

T tests (LSD) for variable: **RANKIT**

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.1 df= 20 MSE= 0.815612  
 Critical Value of T= 1.72

Least Significant Difference= 0.8993

Means with the same letter are not significantly different.

T Grouping	Mean	N	SEDIMENT
A	0.676	6	AK
A			
B	0.162	6	RH
B			
B	-0.283	6	SH
B			
B	-0.555	6	GOW

When the untransformed data or the rankits are used in the LSD test, the treatment means in decreasing order are AK, RH, SH, and GOW, with AK significantly greater than SH and GOW. However, when the data are transformed to logs, the treatment means in decreasing order are now AK, RH, GOW, and SH, with AK significantly greater than SH only. Interpretation of test results is generally not complicated by situations such as these, because the investigator should select the most appropriate transformation and statistical comparison tests based upon the tests of assumptions in the decision trees of the Inland Testing Manual. The results of those comparisons are then accepted, while results using any other transformations or tests are ignored.

## When Data for an Action Level Comparison Fail the Normality Test

If mean tissue contaminant concentration does not exceed an applicable action level or other numeric standard, then statistical testing is recommended to determine whether the mean concentration is significantly less than the action level. The null hypothesis for the action level comparison is that the mean tissue contaminant concentration is equal to the action level. The alternative hypothesis is that the mean contaminant concentration is less than the action level. The Inland Testing Manual provides two equivalent procedures for comparing replicate bioaccumulation data with an action level or other criterion. The two methods are the one-sample *t*-test and the upper confidence limit (UCL) approach, both of which assume normality.

What if the bioaccumulation data have failed the normality test? If the data are nonnormal, there is an increased likelihood of Type I error, i.e., falsely concluding that the action level is not exceeded. Simulations of normal, lognormal, and gamma samples show that the Type I error rate for normal samples is approximately 5 percent regardless of coefficient of variation (Figure 42). Type I error rate for lognormal and gamma samples is approximately 5 percent when the CV is very low (0.1), but increases to about 30 percent as the CV increases to 2. Power for comparisons with an action level follows a pattern similar to Type I error rate for normal and nonnormal distributions. That is, power remains approximately constant regardless of CV when samples are normal, but increases with CV when samples are nonnormal.

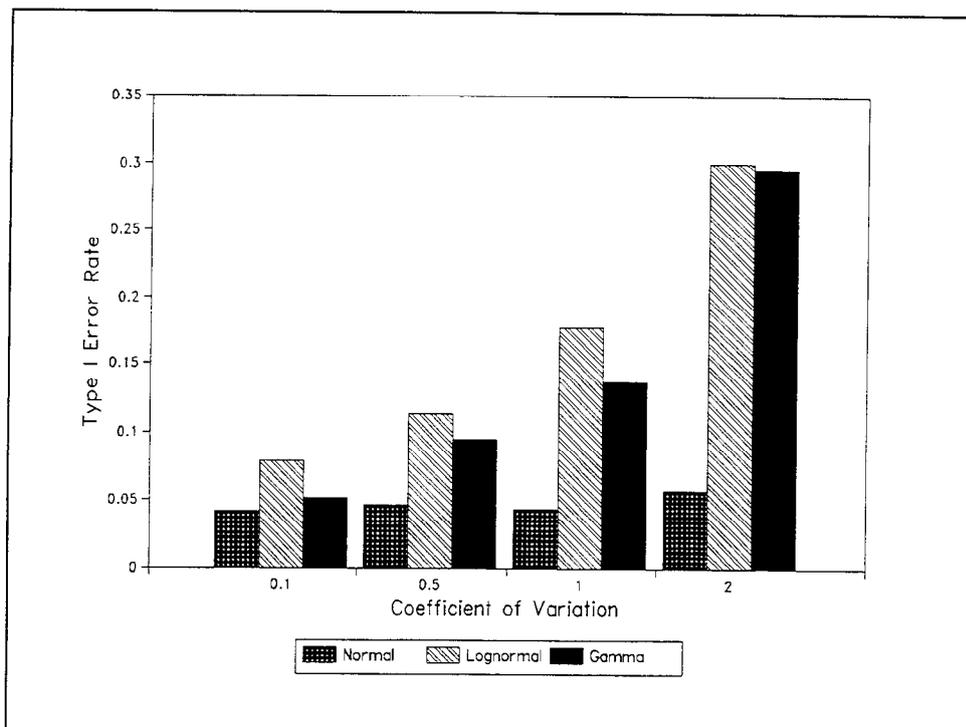


Figure 42. Type I error rate for simulated normal, lognormal, and gamma samples ( $n = 5$ ) in comparisons with an action level

In summary, when bioaccumulation samples derive from an underlying nonnormal population, such as lognormal or gamma, both power and Type I error rate of action level comparisons tend to increase as the CV increases. In other words, nonnormality increases the ability of the action level comparison to identify bioaccumulation samples as being significantly less than the action level regardless of whether the population mean bioaccumulation is in reality below the action level. Therefore, extra caution is advised in action level comparisons when the data have failed the normality test. The investigator may wish to use a more conservative significance level, such as 0.025 or 0.01, especially if the CV is high. Alternatively, the investigator could employ a nonparametric test comparing median bioaccumulation with the action level. Two such procedures, the Sign Test and the Wilcoxon Signed-Ranks Test, are simple to compute by hand and are described in nearly all general and nonparametric statistics texts.

## Analyzing a Blocked Design

Most dredged sediment evaluations can be performed using a completely randomized experimental design and routine statistical procedures as described in the Inland Testing Manual. Occasionally, however, a dredged sediment evaluation may require more complicated statistical treatment, as seen in the CDF water quality example of Chapter 8. There is a wide array of complex experimental designs and statistical analyses, described in texts such as Gad

and Weil (1988); Hicks (1982); Keppel (1991); and Winer (1971). The only topic to be covered in this manual is the analysis of data that can be grouped or blocked by some criterion, such as location of experimental units or sampling time, prior to comparison of sediment testing end points.

Blocking is appropriate when the blocking criterion can be expected to contribute to variability among the data, but comparisons among the blocks are not of interest. A simple example is the arrangement of laboratory test chambers on a series of benches or in a series of water baths. Because spatial variability in environmental conditions could contribute to variability in the test end point, the benches or water baths should be considered blocks (see Chapter 3 for experimental design using randomized blocks). By including the blocks in the data analysis, any variability attributable solely to differences among the blocks can be statistically removed prior to testing for differences among the treatments.

In another example, an investigator might wish to compare two treatments in which tissue contaminant concentrations have been analyzed for a number of related contaminants, such as PCB congeners. However, the investigator desires only an overall PCB comparison, not wanting to interpret comparison results for a multitude of individual congeners. In this situation, a variable identifying the individual congeners may be used as a blocking variable in the analysis.

Consider the following data for metals bioaccumulation in *Mytilus edulis* exposed for 28 days to a contaminated Oakland Harbor sediment in bedded sediment (BS) and 50-mg/l suspended sediment (S50) treatments. Three metals were analyzed, but the investigator only wishes to know whether the route of exposure influences bioaccumulation of the metals as a group. Therefore, a comparison is performed between BS and S50 with the individual metals as a blocking variable. The bioaccumulation data (milligrams/kilogram) are:

Cadmium	BS:	6.84	6.94	6.24	7.76	6.56	4.82
	S50:	5.80	6.00	6.31	7.76	6.95	7.69
Chromium	BS:	5.10	2.30	6.10	2.60	4.50	3.50
	S50:	2.90	3.20	3.20	4.40	3.30	4.60
Mercury	BS:	0.289	0.296	0.263	0.283	0.296	0.282
	S50:	0.264	0.315	0.366	0.318	0.266	0.325

SAS program statements are given in BLOCKS.SAS in Appendix A. Note that each observation in the data input step includes three variables: the contaminant concentration (CONC), a treatment identifier (TRT), and a block identifier (CONTAM). The data set was found to be normally distributed ( $W = 0.964$ ,  $P = 0.3691$ ), and variances of the two treatments were equal after adjusting for the blocks (Levene's  $F = 1.27$ ,  $P = 0.2682$  for TRT Type III sums of squares). Metals bioaccumulation did not differ between the

two treatments after adjusting for blocks (ANOVA  $F = 0.04$ ,  $P = 0.8431$  for TRT Type III sums of squares).

Another example of blocking is the zinc water quality data described earlier (Chapter 8, Table 18). Recall that samples were collected from several locations during three time intervals: before, during, and after a dredging event. The Corps District was interested in comparing water quality among locations but not necessarily among the time intervals. Nevertheless, the time interval could certainly influence the amount of variability in the data set as a whole. Therefore, the three time intervals should be considered blocks in the statistical analysis.

Preliminary analysis of the zinc water quality data for the CDF-harbor comparisons indicated that this highly censored data set should be analyzed using rankits after substituting a constant for the nondetects (Table 20 in Chapter 8). Because the time intervals will be used as blocks in the analysis, the variable that specifies the time intervals should also be included as a blocking variable to calculate residuals for the test of normality and in the Levene's Test for equality of variances. The analysis can be conducted substituting the appropriate data and variable names in BLOCKS.SAS (Appendix A). Tests of assumptions for the zinc data indicated that the rankits are not normally distributed ( $W = 0.833$ ,  $P = 0.0$ ), and variances are unequal among the locations even after adjusting for blocks (Levene's  $F = 4.12$ ,  $P = 0.0002$ ). When the assumptions are violated for rankits, the blocked ANOVA and LSD test may still be used, but the investigator should be aware of the possibility of increased Type I error rate (see Chapter 4, Figure 10). Alternatively, a ranks test such as the Friedman test (Conover 1980) may be used.

The LSD results for the zinc water quality data (rankits) are shown below. The after-dredging time block is included in this analysis.

```

WATER QUALITY DATA FOR ZINC
LSD TEST

General Linear Models Procedure

T tests (LSD) for variable: RANKIT

NOTE: This test controls the type I comparisonwise
      error rate not the experimentwise error rate.

Alpha= 0.1  df= 115  MSE= 0.429205
Critical Value of T= 1.66
Least Significant Difference= 0.4106

Means with the same letter are not significantly
different.

```

T Grouping	Mean	N	TRT
A	0.373	14	WELL7
B	0.160	14	SAMPLESA

B	A				
B	A	C	0.093	14	SAMPLE7
B	A	C			
B	A	C	0.089	14	CDF
B	A	C			
B	A	C	0.005	14	SAMPLE6
B		C			
B		C	-0.084	14	SAMPLE8B
B		C			
B		C	-0.148	14	SAMPLE5
B		C			
B		C	-0.193	14	WELL9
B		C			
		C	-0.297	14	WELL4

Remember that the only comparisons of interest are between the two harbor reference samples (Sample 8a and Sample 8b) and each of the other locations in and near the CDF. No location was significantly greater than reference Sample 8a, and only Well 7 was significantly greater than reference Sample 8b.

To simplify this problem somewhat, the investigator could look at general locations rather than individual samples and compare the CDF locations (Inside, Dike, and Adjacent) with the reference location (Harbor), again using the collection dates as a blocking variable. In this analysis, the rankits still are not normally distributed ( $W = 0.732$ ,  $P = 0.0$ ), but the variances among locations are equal after adjusting for date (Levene's  $F = 0.55$ ,  $P = 0.6504$ ). LSD output is as follows:

WATER QUALITY DATA FOR ZINC  
LSD TEST

General Linear Models Procedure

T tests (LSD) for variable: RANKIT

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.1 Confidence= 0.9 df= 120 MSE= 0.448616  
Critical Value of T= 1.65765

Comparisons significant at the 0.1 level are indicated by '\*\*\*\*'.

LOCATION Comparison	Lower Confidence Limit	Difference Between Means	Upper Confidence Limit
INSIDE - HARBOR	-0.312	0.051	0.415
INSIDE - ADJACENT	-0.237	0.106	0.449
INSIDE - DIKE	-0.215	0.128	0.471
HARBOR - INSIDE	-0.415	-0.051	0.312
HARBOR - ADJACENT	-0.216	0.055	0.326
HARBOR - DIKE	-0.194	0.077	0.348
ADJACENT - INSIDE	-0.449	-0.106	0.237
ADJACENT - HARBOR	-0.326	-0.055	0.216
ADJACENT - DIKE	-0.220	0.022	0.264
DIKE - INSIDE	-0.471	-0.128	0.215
DIKE - HARBOR	-0.348	-0.077	0.194
DIKE - ADJACENT	-0.264	-0.022	0.220

No location differs significantly from any other.

## Summary

Statistical tests have little meaning unless the test results are properly interpreted. This chapter has described a number of situations in which statistical test results are ambiguous or the interpretation of results is less straightforward than the examples given in the Inland Testing Manual. Guidance on these interpretation problems generally is not presented in statistical texts.

- The P-value of a test statistic may be considered an indication of confidence in accepting the null hypothesis based on the sample data. The closer the P-value is to zero, the more likely the null hypothesis should be rejected, especially when the P-value is much lower than the significance level ( $\alpha$ ) of the test.
- When P is close to  $\alpha$ , conclusions should be drawn with caution. The P-value of the test should be considered in subsequent decisions.
- If two or more equivalent statistical tests result in different conclusions, the investigator should consider carefully the relative power of the tests and their appropriateness for the particular data situation. If there is no statistical reason to favor one test over another, then inferences from the tests should be considered inconclusive.
- LSD significance may differ from that of the ANOVA that precedes the LSD test. When an independent decision will be made for each dredged sediment or management unit included in the statistical comparison, the ANOVA results should be ignored.
- When LSD output provides letters next to treatment means, means having the same letter are not significantly different from each other.
- When sample sizes are unequal, SAS provides confidence intervals for the difference between means in the LSD output. A confidence interval that does not include zero indicates that the difference between means is significant (denoted in SAS by \*\*\*).
- Data transformation can change the order of treatment means as well as the groupings of means that do not differ significantly in the LSD output. Results should be accepted for the most appropriate transformation, if any, based on tests of assumptions in the decision trees of the Inland Testing Manual.
- When nonnormal data are used in statistical comparison with an action level, the Type I error rate tends to increase as the coefficient of variation increases. Use of a lower significance level than 0.05 may be warranted to increase environmental protectiveness of the action level comparison when the data have failed the normality test, especially if the CV is high. Alternatively, a nonparametric procedure such as the Sign Test or the Wilcoxon Signed-Ranks Test may be used.

- A blocked design should be used in data analysis when the blocking criterion could contribute to variability among the data, but comparisons among the blocks are not of interest. The blocks are used in the calculation of residuals for the tests of assumptions and are included in the ANOVA model for Levene's Test and the LSD test. The variability attributable to the blocks is statistically removed, leaving a more accurate analysis of variability attributable to the treatments of interest.

# References

---

- Barnett, V., and Lewis, T. (1984). *Outliers in statistical data*. 2nd ed., John Wiley and Sons, New York.
- Brown, M. B., and Forsythe, A. B. (1974). "Robust tests for the equality of variances," *Journal of the American Statistical Association* 69, 364-67.
- Clarke, J. U. (1992). "Analyzing "less-than" data: An example from sediment bioaccumulation testing." *Water Quality '92 Proceedings of the 9th Seminar*. Miscellaneous Paper W-92-3, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS, 178-91.
- \_\_\_\_\_. (1994). "Evaluating methods for statistical analysis of less than detection limit data using simulated small samples. 2. General results." *Dredging '94, Proceedings of the Second International Conference on Dredging and Dredged Material Placement, Vol. 1*. E. C. McNair, Jr., ed., American Society of Civil Engineers, New York, 747-55.
- \_\_\_\_\_. (1995a). "Guidelines for statistical treatment of less than detection limit data in dredged sediment evaluations," Environmental Effects of Dredging Technical Note EEDP-04-23, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.
- \_\_\_\_\_. (1995b). "Evaluating methods for statistical analysis of less than detection limit data using simulated small samples. 1. Study description." *Water Quality '94 Proceedings of the 10th Seminar, 14-18 February 1994, Savannah, Georgia*. Miscellaneous Paper W-95-1, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS, 330-35.
- Clarke, J. U., and Brandon, D. L. (1994). "Less than detection limit data—a problem for statistical analysis and decisionmaking in dredged material disposal evaluations," Environmental Effects of Dredging Information Exchange Bulletin, Vol. D-94-1, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.
- Cochran, W. G. (1977). *Sampling techniques*. 3rd ed., John Wiley and Sons, New York.

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. 2nd ed., Lawrence Erlbaum Associates, Hillsdale, NJ.
- Conover, W. J. (1980). *Practical nonparametric statistics*. 2nd ed., John Wiley and Sons, New York.
- Conover, W. J., Johnson, M. E., and Johnson, M. M. (1981). "A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data," *Technometrics* 23, 351-61.
- Cressie, N. (1994). "Spatial chemometrics," *Environmental statistics, assessment, and forecasting*. C. R. Cothorn and N. P. Ross, ed., Lewis Publishers, Boca Raton, FL.
- Dixon, W. J. (1950). "Analysis of extreme values," *Annals of Mathematical Statistics* 21, 488-506.
- Dixon, W. J., and Massey, F. J., Jr. (1983). *Introduction to statistical analysis*. 4th ed., McGraw-Hill Book Company, New York.
- Dixon, W. J., and Tukey, J. W. (1968). "Approximate behavior of the distribution of Winsorized  $t$  (trimming/Winsorization 2)," *Technometrics* 10, 83-98.
- Dorfman, J. H., Pesti, G. M., and Fletcher, D. L. (1993). "Searching for significance: The perils of excluding pseudo-outliers," *Poultry Science* 72, 37-41.
- El-Shaarawi, A. H. (1989). "Inferences about the mean from censored water quality data," *Water Resources Research* 25(4), 685-90.
- Gad, S., and Weil, C. S. (1988). *Statistics and experimental design for toxicologists*. 2nd ed., The Telford Press, Caldwell, NJ.
- Gilbert, R. O. (1987). *Statistical methods for environmental pollution monitoring*. Van Nostrand Reinhold, New York.
- Gill, J. L. (1978). *Design and analysis of experiments in the animal and medical sciences*. Vol. 3, Appendices, The Iowa State University Press, Ames, IA.
- Gilliom, R. J., and Helsel, D. R. (1986). "Estimation of distributional parameters for censored trace level water quality data. 1. Estimation techniques," *Water Resources Research* 22(2), 135-46.
- Gilliom, R. J., Hirsch, R. M., and Gilroy, E. J. (1984). "Effect of censoring trace-level water-quality data on trend-detection capability," *Environmental Science and Technology* 18(7), 530-35.

- Glaser, J. A., Foerst, D. L., McKee, G. D., Quave, S. A., and Budde, W. L. (1981). "Trace analyses for wastewaters," *Environmental Science and Technology* 15(12), 1426-35.
- Gleit, A. (1985). "Estimation for small normal data sets with detection limits," *Environmental Science and Technology* 19(12), 1201-06.
- Grubbs, F. E. (1969). "Procedures for detecting outlying observations in samples," *Technometrics* 11(1), 1-21.
- Haas, C. N., and Scheff, P. A. (1990). "Estimation of averages in truncated samples," *Environmental Science and Technology* 24(6), 912-19.
- Helsel, D. R. (1990). "Less than obvious," *Environmental Science and Technology* 24(12), 1766-74.
- Helsel, D. R., and Cohn, T. A. (1988). "Estimation of descriptive statistics for multiply censored water quality data," *Water Resources Research* 24(12), 1997-2004.
- Helsel, D. R., and Gilliom, R. J. (1986). "Estimation of distributional parameters for censored trace level water quality data. 2. Verification and applications," *Water Resources Research* 22(2), 147-55.
- Hicks, C. R. (1982). *Fundamental concepts in the design of experiments*. 3rd ed., Holt, Rinehart and Winston, New York.
- Keppel, G. (1991). *Design and analysis, a researcher's handbook*. Prentice Hall, Englewood Cliffs, NJ.
- Kushner, E. J. (1976). "On determining the statistical parameters for pollution concentration from a truncated data set," *Atmospheric Environment* 10, 975-79.
- Lambert, D., Peterson, B., and Terpenning, I. (1991). "Nondetects, detection limits, and the probability of detection," *Journal of the American Statistical Association* 86(414), 266-77.
- McClave, J. T., and Dietrich, F. H., II. (1979). *Statistics*. Dellen Publishing Company, San Francisco, CA.
- Milliken, G. A., and Johnson, D. E. (1984). *Analysis of messy data. Volume 1: Designed experiments*. Van Nostrand Reinhold, New York.

- Moore, D. W., Gibson, A. B., Dillon, T. M., Bridges, T. S., Gamble, E. W., Gray, B. R., Wright, R. B., and Baggett, L. H. (1994). "Evaluation of proposed U.S. Environmental Protection Agency dredged material bioassays using Great Lakes sediments," Miscellaneous Paper EL-94-11, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.
- Newman, M. C., and Dixon, P. M. (1990). "UNCENSOR: A program to estimate means and standard deviations for data sets with below detection limit observations," *American Environmental Laboratory* April 1990, 26-30.
- Newman, M. C., Dixon, P. M., Looney, B. B., and Pinder, J. E., III. (1989). "Estimating mean and variance for environmental samples with below detection limit observations," *Water Resources Bulletin* 25(4), 905-16.
- Norusis, M. (1992a). *SPSS/PC+ base system user's guide version 5.0*. SPSS, Inc., Chicago, IL.
- \_\_\_\_\_. (1992b). *SPSS/PC+ advanced statistics version 5.0*. SPSS, Inc., Chicago, IL.
- Ott, W. R., and Mage, D. T. (1976). "A general purpose univariate probability model for environmental data analysis," *Computers and Operations Research* 3, 209-16.
- Parsons, M. L. (1969). "The definition of detection limits," *Journal of Chemical Education* 46(5), 290-92.
- Pennington, J. C., Higgins, T. R., Folsom, B. L., Jr., and Brandon, D. L. (1990). "Considerations for reducing the cost of testing dredged material," Technical Report D-90-7, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.
- Porter, P. S., and Ward, R. C. (1991). "Estimating central tendency from uncensored trace level measurements," *Water Resources Bulletin* 27(4), 687-700.
- Porter, P. S., Ward, R. C., and Bell, H. F. (1988). "The detection limit," *Environmental Science and Technology* 22(8), 856-61.
- Provost, L. P. (1984). "Statistical methods in environmental sampling." *Environmental sampling for hazardous wastes*. G. E. Schweitzer and J. A. Santolucito, ed., ACS Symposium Series 267, 79-96, American Chemical Society, Washington, DC.
- Rohlf, F. J., and Sokal, R. R. (1981). *Statistical tables*. 2nd ed., W. H. Freeman and Company, San Francisco, CA.

- SAS Institute, Inc. (1988a). *SAS/STAT user's guide, release 6.03 edition*. SAS Institute, Inc., Cary, NC.
- \_\_\_\_\_. (1988b). *SAS procedures guide, release 6.03 edition*. SAS Institute, Inc., Cary, NC.
- SAS Institute, Inc. (1988c). *SAS language guide for personal computers, release 6.03 edition*. SAS Institute, Inc., Cary, NC.
- Slymen, D. J., de Peyster, A., and Donohoe, R. R. (1994). "Hypothesis testing with values below detection limit in environmental studies," *Environmental Science and Technology* 28(5), 898-902.
- Snedecor, G. W., and Cochran, W. G. (1989). *Statistical methods*. 8th ed., The Iowa State University Press, Ames, IA.
- Sokal, R. R., and Rohlf, F. J. (1981). *Biometry, the principles and practice of statistics in biological research*. 2nd ed., W. H. Freeman and Company, San Francisco, CA.
- Staudte, R. G., and Sheather, S. J. (1990). *Robust estimation and testing*. John Wiley and Sons, Inc., New York.
- Travis, C. C., and Land, M. L. (1990). "Estimating the mean of data sets with nondetectable values," *Environmental Science and Technology* 24(7), 961-62.
- U.S. Environmental Protection Agency/U.S. Army Corps of Engineers. (1991). "Evaluation of dredged material proposed for ocean disposal: Testing manual," EPA-503/8-91/001, Environmental Protection Agency, Office of Marine and Estuarine Protection, and Department of the Army, U.S. Army Corps of Engineers, Washington, DC.
- \_\_\_\_\_. (1994). "Evaluation of dredged material proposed for discharge in waters of the U.S.—Testing manual (draft)," EPA-823-B-94-002, U.S. Environmental Protection Agency Office of Water, Washington, DC.
- White, G. C. (1978). "Estimation of expected value and coefficient of variation for lognormal and gamma distributions," Informal Report LA-7393-MS, Los Alamos Scientific Laboratory, Los Alamos, NM.
- Wilkinson, L. (1990a). *SYSTAT: The system for statistics*. SYSTAT, INC., Evanston, IL.
- \_\_\_\_\_. (1990b). *SYGRAPH: The system for graphics*. SYSTAT, Inc., Evanston, IL.

Winer, B. J. (1971). *Statistical principles in experimental design*. 2nd ed., McGraw-Hill Book Company, New York.

# Appendix A

## SAS Programs

---

### Program POWER.SAS to Calculate Power for a *t*-Test

The following statements implement Equation 10 from Appendix D of the Inland Testing Manual, for calculating power given a fixed sample size and effect size. The data step defines the following variables: ALPHA ( $\alpha$ ), CONF (confidence or  $1 - \alpha$ ), EFFSIZE (effect size relative to standard deviation), NREF (sample size for the reference sediment), and NDREDGE (sample size for the dredged sediment). DF (degrees of freedom), TALPHA (*t* value for  $\alpha$  and DF), TBETA (*t* value for  $\beta$  and DF), and POWER are calculated. In this example program, power is calculated for two  $\alpha$  levels (0.05 and 0.10), equal sample sizes of  $n = 5$ , and several relative effect sizes.

#### POWER.SAS program statements

```
DATA A;
  NREF=5; NDREDGE=5;
  DF=NREF+NDREDGE-2;
  DO ALPHA=.05,.1;
    CONF=1-ALPHA;
    TALPHA=TINV(CONF,DF);
    DO EFFSIZE=.5,.75,1,2;
      TBETA=(SQRT((NREF+NDREDGE)/2)/SQRT(2))*EFFSIZE-TALPHA;
      POWER=PROBT(TBETA,DF);
      OUTPUT;
    END;
  END;
PROC PRINT;
```

#### POWER.SAS program output

OBS	NREF	NDREDGE	DF	ALPHA	CONF	TALPHA	EFFSIZE	TBETA	POWER
1	5	5	8	0.05	0.95	1.85955	0.50	-1.06898	0.15814
2	5	5	8	0.05	0.95	1.85955	0.75	-0.67369	0.25975
3	5	5	8	0.05	0.95	1.85955	1.00	-0.27841	0.39388
4	5	5	8	0.05	0.95	1.85955	2.00	1.30273	0.88554
5	5	5	8	0.10	0.90	1.39682	0.50	-0.60625	0.28058
6	5	5	8	0.10	0.90	1.39682	0.75	-0.21096	0.41910
7	5	5	8	0.10	0.90	1.39682	1.00	0.18432	0.57083
8	5	5	8	0.10	0.90	1.39682	2.00	1.76546	0.94226

## Program DMIN.SAS to Calculate Least Significant Difference for a *t*-Test

The following statements implement Equation 11 from Appendix D of the Inland Testing Manual to calculate least significant difference for a statistical comparison. The data step defines the following variables: ALPHA ( $\alpha$ ), CONF (confidence or  $1 - \alpha$ ), S2 (variance), NREF (sample size for the reference sediment), and NDREDGE (sample size for the dredged sediment). S2 is the pooled variance for the two treatments being compared. If more than two treatments are included in the test and variances are equal among treatments, then the mean square error (MSE) from the analysis of variance should be used as S2. DF (degrees of freedom), TALPHA (*t* value for  $\alpha$  and DF), and DMIN (least significant difference) are calculated. TBETA (*t* value for  $\beta$  and DF) = 0. In this example program, least significant difference is calculated for  $\alpha = 0.05$ , equal sample sizes of  $n = 5$ , and variance = 1.

### DMIN.SAS program statements

```
DATA A;
  NREF=5;
  NDREDGE=5;
  DF=NREF+NDREDGE-2;
  ALPHA=.05;
  CONF=1-ALPHA;
  S2=1;
  TALPHA=TINV(CONF,DF);
  DMIN=TALPHA*2*SQRT(S2/(NREF+NDREDGE));
PROC PRINT;
```

### DMIN.SAS program output

OBS	NREF	NDREDGE	DF	S2	ALPHA	CONF	TALPHA	DMIN
1	5	5	8	1	0.05	0.95	1.85955	1.17608

## Program EQOFVAR.SAS to Perform Equality of Variance Tests

This program provides SAS statements for Bartlett's Test, Hartley's  $F_{\max}$ , Cochran's Test, and Levene's Test for equality of variances. The survival data in Table 5 of Chapter 5 are used to illustrate the procedures. Equality of variance tests are conducted on the untransformed survival proportions, arcsine-transformed survival proportions, and rankits. The probabilities associated with Bartlett's  $\chi^2$  and Levene's  $F$  statistics are calculated. The significance of Hartley's  $F_{\max}$  and Cochran's  $C$  statistics must be determined using tables such as those provided in Dixon and Massey (1983); Gill (1978); Rohlf and Sokal (1981); and Winer (1971).<sup>1</sup>

---

<sup>1</sup> References cited in this appendix are located at the end of the main text.

## EQOFVAR.SAS program statements

```
OPTIONS PAGESIZE=500 LINESIZE=79 NODATE NONUMBER;
LIBNAME Q 'C:\SAS';
TITLE 'GREAT LAKES SURVIVAL DATA';

/* Recall and sort the data, calculate rankits, split and rearrange the data
set by transformation. The variable P is the survival proportion in the
original data set. */

DATA A;
  SET Q.GLSURV;
PROC SORT; BY SPECIES SEDIMENT;
PROC RANK NORMAL=BLOM OUT=B;
  BY SPECIES;
  VAR P; RANKS RANKIT;
DATA B1; SET B;
  TRANSFRM='ARCSIN';
  SURVIVAL=ARCP;
  KEEP SPECIES SEDIMENT SURVIVAL TRANSFRM;
DATA B2; SET B;
  TRANSFRM='NONE';
  SURVIVAL=P;
  KEEP SPECIES SEDIMENT SURVIVAL TRANSFRM;
DATA B3; SET B;
  TRANSFRM='RANKIT';
  SURVIVAL=RANKIT;
  KEEP SPECIES SEDIMENT SURVIVAL TRANSFRM;
DATA ALL; SET B1 B2 B3;
PROC SORT; BY SPECIES TRANSFRM SEDIMENT;

/* Calculate statistics for Bartlett's Test. Procedures are given in Sokal
and Rohlf (1981, p. 404). */

PROC MEANS NOPRINT;
  BY SPECIES TRANSFRM SEDIMENT;
  VAR SURVIVAL;
  OUTPUT OUT=O VAR=VAR N=N;
DATA O1; SET O;
  LNVAR=LOG(VAR);
  DF=N-1;
  DFLNVAR=DF*LNVAR;
  DFVAR=DF*VAR;
  INVDF=1/DF;
PROC MEANS NOPRINT;
  BY SPECIES TRANSFRM;
  VAR DF DFVAR DFLNVAR INVDF VAR;
  OUTPUT OUT=P SUM=SUMDF SUMDFVAR SUMLNVAR SUMINVDF SUMVAR;
DATA P1; SET P;
  S2=SUMDFVAR/SUMDF;
  LNS2=LOG(S2);
  X2=(SUMDF*LNS2)-SUMLNVAR;
  DF1=FREQ_1;
  C=1+(1/(3*DF1))*(SUMINVDF-(1/SUMDF));
  CHI2=X2/C;
  P=1-PROBCHI(CHI2,DF1);
PROC PRINT LABEL NOOBS;
  VAR SPECIES TRANSFRM X2 C DF1 CHI2 P;
  LABEL TRANSFRM='DATA TRANSFOR- MATION'
        X2='UNCORRECTED CHI-SQUARE'
        C='CORRECTION FACTOR'
        DF1='DEGREES OF FREEDOM FOR CHI- SQUARE'
        CHI2='CORRECTED CHI- SQUARE'
        P='PROBA- BILITY';
  TITLE2 'BARTLETT'S TEST';

/* Calculate Hartley's Fmax and Cochran's C statistics. */

PROC MEANS NOPRINT DATA=O;
  BY SPECIES TRANSFRM;
  VAR VAR;
  OUTPUT OUT=OO MAX=MAXVAR MIN=MINVAR SUM=SUMVAR N=K;
DATA OO1; SET OO;
  HARTLEY=MAXVAR/MINVAR;
  COCHRAN=MAXVAR/SUMVAR;
PROC PRINT LABEL NOOBS;
  VAR SPECIES TRANSFRM K MAXVAR MINVAR SUMVAR HARTLEY COCHRAN;
  LABEL TRANSFRM='DATA TRANSFOR- MATION'
        K='NUMBER OF TREAT- MENTS'
        MAXVAR='MAXIMUM VARIANCE'
        MINVAR='MINIMUM VARIANCE'
        SUMVAR='SUM OF VARI- ANCES';
  TITLE2 'HARTLEY'S Fmax AND COCHRAN'S TEST';

/* Perform Levene's Test. SAS Statements for this test are also given in
programs BENTOX.SAS, BIOACC.SAS, and BIOACCSS.SAS of the Inland Testing
Manual Appendix D. */

PROC GLM DATA=ALL NOPRINT;
  BY SPECIES TRANSFRM;
```

```

CLASS SEDIMENT;
MODEL SURVIVAL=SEDIMENT;
OUTPUT OUT=Z R=RESID;
DATA Z1; SET Z;
ABSRESID=ABS(RESID);
PROC GLM OUTSTAT=Y NOFRINT;
BY SPECIES TRANSFRM;
CLASS SEDIMENT;
MODEL ABSRESID=SEDIMENT;
DATA Y1; SET Y;
IF TYPE='SS1';
KEEP SPECIES TRANSFRM F PROB;
PROC PRINT LABEL NOOBS;
VAR SPECIES TRANSFRM F PROB;
LABEL TRANSFRM='DATA TRANSFORMATION'
      F='LEVENE'S F'
      PROB='PROBABILITY';
TITLE2 'LEVENE'S TEST';

```

## EQOFVAR.SAS program output

GREAT LAKES SURVIVAL DATA BARTLETT'S TEST						
SPECIES	DATA TRANSFORMATION	UNCORRECTED CHI-SQUARE	CORRECTION FACTOR	DEGREES OF FREEDOM FOR CHI-SQUARE	CORRECTED CHI-SQUARE	PROBABILITY
HYALELLA 1	ARCSIN	23.6299	1.06111	9	22.2690	0.00806
HYALELLA 1	NONE	45.8109	1.06111	9	43.1726	0.00000
HYALELLA 1	RANKIT	2.3412	1.06111	9	2.2064	0.98777
HYALELLA 2	ARCSIN	4.7545	1.06944	3	4.4457	0.21718
HYALELLA 2	NONE	14.2105	1.06944	3	13.2877	0.00405
HYALELLA 2	RANKIT	2.0273	1.06944	3	1.8957	0.59434
PIMEPHALES	ARCSIN	10.0829	1.11624	8	9.0329	0.33953
PIMEPHALES	NONE	11.1668	1.11624	8	10.0039	0.26475
PIMEPHALES	RANKIT	6.8468	1.11624	8	6.1338	0.63225

GREAT LAKES SURVIVAL DATA HARTLEY'S Fmax AND COCHRAN'S TEST							
SPECIES	DATA TRANSFORMATION	NUMBER OF TREATMENTS	MAXIMUM VARIANCE	MINIMUM VARIANCE	SUM OF VARIANCES	HARTLEY	COCHRAN
HYALELLA 1	ARCSIN	10	0.39284	0.04115	1.17708	9.5462	0.33374
HYALELLA 1	NONE	10	0.17571	0.00619	0.46524	28.3846	0.37769
HYALELLA 1	RANKIT	10	1.12248	0.40151	6.47736	2.7956	0.17329
HYALELLA 2	ARCSIN	4	0.13462	0.02465	0.25494	5.4615	0.52804
HYALELLA 2	NONE	4	0.06571	0.00238	0.09714	27.6000	0.67647
HYALELLA 2	RANKIT	4	1.11219	0.37091	2.68035	2.9986	0.41494
PIMEPHALES	ARCSIN	8	0.05198	0.00368	0.15647	14.1226	0.33220
PIMEPHALES	NONE	8	0.04667	0.00250	0.13083	18.6667	0.35669
PIMEPHALES	RANKIT	8	0.91126	0.13608	3.39261	6.6966	0.26860

GREAT LAKES SURVIVAL DATA LEVENE'S TEST			
SPECIES	DATA TRANSFORMATION	LEVENE'S F	PROBABILITY
HYALELLA 1	ARCSIN	5.30353	0.00003
HYALELLA 1	NONE	8.32553	0.00000
HYALELLA 1	RANKIT	0.26204	0.98231
HYALELLA 2	ARCSIN	1.25660	0.31151
HYALELLA 2	NONE	1.55095	0.22715
HYALELLA 2	RANKIT	1.22707	0.32153
PIMEPHALES	ARCSIN	1.77744	0.13380
PIMEPHALES	NONE	1.72424	0.14613
PIMEPHALES	RANKIT	1.56208	0.19096

# SAS Programs for Preliminary Analysis of Censored Data

The following programs provide SAS statements for a preliminary analysis of censored data to determine the likely data distribution, CV, and pattern of variances among treatments. This information will enable selection of the most appropriate censored data method for subsequent statistical comparisons.

## PRELIM1.SAS program statements

PRELIM1.SAS should be used when there is one nondetect in one or more treatments. This program includes the simple substitution methods DL, DL/2, and ZERO. Example data used in the program are for lead bioaccumulation in *Nereis virens* (Table 7 in Chapter 5).

```
OPTIONS LINESIZE=79 PAGESIZE=500 NODATE NONUMBER;

/* Input the data here or read in an existing permanent SAS data set. We
recommend that nondetects be coded as -dl where dl is the numeric detection
limit. To avoid changing statements after the first data step each time
the program is run, name or rename the contaminant concentration variable
CONC and the treatment variable TRT. */

DATA A;
  INPUT TRT $ CONC @@;
  CARDS;
AK -.022 AK .142 AK .171 AK .186 AK .126 AK .1
GOW .243 GOW .076 GOW .039 GOW .112 GOW .259 GOW .397
RH .973 RH .066 RH .081 RH .129 RH .225 RH .33
SH .046 SH .086 SH .096 SH .112 SH .115 SH .391
;
TITLE 'LEAD BIOACCUMULATION IN NEREIS VIRENS';

/* Apply ZERO and count number of nondetects */

DATA ZERO;
  SET A;
  IF CONC<0 THEN COUNT=1; ELSE COUNT=0;
  IF CONC<0 THEN CONC=0;
  METHOD='ZERO';
  PROC PRINT;

/* Apply DL/2 */

DATA DL2;
  SET A;
  IF CONC<0 THEN CONC=ABS(CONC)/2;
  METHOD='DL/2';
  PROC PRINT;

/* Apply DL */

DATA DL;
  SET A;
  IF CONC<0 THEN CONC=ABS(CONC);
  METHOD='DL';
  PROC PRINT;

/* Determine percent of data that are censored */

PROC MEANS DATA=ZERO NOPRINT;
  VAR COUNT;
  OUTPUT OUT=O SUM=SUM N=N;
DATA O1; SET O;
  PROPCENS=SUM*100/N;
PROC PRINT LABEL NOOBS;
  VAR PROPCENS N;
  LABEL PROPCENS='PERCENT OF DATA THAT ARE CENSORED'
        N='TOTAL NUMBER OF REPLICATES';

/* Combine data sets and sort by method. Calculate logs. */

DATA ALL;
  SET ZERO DL2 DL;
  LOGCONC=LOG10(CONC);
```

```

LABEL TRT='SEDIMENT';
PROC SORT; BY METHOD TRT;

/* Determine CV of combined samples */

PROC MEANS NOPRINT;
VAR CONC; BY METHOD;
OUTPUT OUT=O CV=CV;
DATA O1; SET O;
CV=CV/100;
PROC PRINT LABEL NOOBS;
VAR METHOD CV;
LABEL CV='CV OF COMBINED SAMPLES';

/* Test normality of combined sample residuals using Shapiro-Wilk's Test */

PROC GLM NOPRINT DATA=ALL;
BY METHOD;
CLASS TRT;
MODEL CONC LOGCONC=TRT;
OUTPUT OUT=Z R=RESID LOGRESID;
PROC UNIVARIATE NORMAL;
BY METHOD;
VAR RESID LOGRESID;

/* Test equality of variances using Levene's Test */

DATA Z1; SET Z;
DEV=ABS(RESID);
PROC GLM;
BY METHOD;
CLASS TRT;
MODEL DEV=TRT;

/* Calculate sample variances */

PROC MEANS NOPRINT DATA=ALL;
BY METHOD TRT;
VAR CONC;
OUTPUT OUT=O VAR=VARI MEAN=MEAN N=N;
PROC SORT; BY METHOD VARI;
PROC PRINT LABEL NOOBS;
VAR METHOD TRT N MEAN VARI;
LABEL VARI='SAMPLE VARIANCES'
MEAN='SAMPLE MEANS';

```

## PRELIM1.SAS program output

```

LEAD BIOACCUMULATION IN NEREIS VIRENS

```

OBS	TRT	CONC	COUNT	METHOD
1	AK	0.000	1	ZERO
2	AK	0.142	0	ZERO
3	AK	0.171	0	ZERO
4	AK	0.186	0	ZERO
5	AK	0.126	0	ZERO
6	AK	0.100	0	ZERO
7	GOW	0.243	0	ZERO
8	GOW	0.076	0	ZERO
9	GOW	0.039	0	ZERO
10	GOW	0.112	0	ZERO
11	GOW	0.259	0	ZERO
12	GOW	0.397	0	ZERO
13	RH	0.973	0	ZERO
14	RH	0.066	0	ZERO
15	RH	0.081	0	ZERO
16	RH	0.129	0	ZERO
17	RH	0.225	0	ZERO
18	RH	0.330	0	ZERO
19	SH	0.046	0	ZERO
20	SH	0.086	0	ZERO
21	SH	0.096	0	ZERO
22	SH	0.112	0	ZERO
23	SH	0.115	0	ZERO
24	SH	0.391	0	ZERO

OBS	TRT	CONC	METHOD
1	AK	0.011	DL/2
2	AK	0.142	DL/2
3	AK	0.171	DL/2
4	AK	0.186	DL/2
5	AK	0.126	DL/2
6	AK	0.100	DL/2
7	GOW	0.243	DL/2
8	GOW	0.076	DL/2
9	GOW	0.039	DL/2

10	GOW	0.112	DL/2
11	GOW	0.259	DL/2
12	GOW	0.397	DL/2
13	RH	0.973	DL/2
14	RH	0.066	DL/2
15	RH	0.081	DL/2
16	RH	0.129	DL/2
17	RH	0.225	DL/2
18	RH	0.330	DL/2
19	SH	0.046	DL/2
20	SH	0.086	DL/2
21	SH	0.096	DL/2
22	SH	0.112	DL/2
23	SH	0.115	DL/2
24	SH	0.391	DL/2

OBS	TRT	CONC	METHOD
1	AK	0.022	DL
2	AK	0.142	DL
3	AK	0.171	DL
4	AK	0.186	DL
5	AK	0.126	DL
6	AK	0.100	DL
7	GOW	0.243	DL
8	GOW	0.076	DL
9	GOW	0.039	DL
10	GOW	0.112	DL
11	GOW	0.259	DL
12	GOW	0.397	DL
13	RH	0.973	DL
14	RH	0.066	DL
15	RH	0.081	DL
16	RH	0.129	DL
17	RH	0.225	DL
18	RH	0.330	DL
19	SH	0.046	DL
20	SH	0.086	DL
21	SH	0.096	DL
22	SH	0.112	DL
23	SH	0.115	DL
24	SH	0.391	DL

LEAD BIOACCUMULATION IN NEREIS VIRENS

PERCENT OF DATA THAT ARE CENSORED	TOTAL NUMBER OF REPLICATES
4.16667	24

LEAD BIOACCUMULATION IN NEREIS VIRENS

METHOD	CV OF COMBINED SAMPLES
DL	1.04540
DL/2	1.05016
ZERO	1.05508

LEAD BIOACCUMULATION IN NEREIS VIRENS  
SHAPIRO WILK'S TEST

----- METHOD=DL -----

Variable=RESID

UNIVARIATE PROCEDURE

N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.184052	Variance	0.033875
Skewness	2.230779	Kurtosis	7.260387
USS	0.779128	CSS	0.779128
CV	.	Std Mean	0.037569
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-33	Prob> S	0.3567
Num ^ = 0	24		
W:Normal	0.80449	Prob<W	0.0002

Variable=LOGRESID

N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.34366	Variance	0.118102
Skewness	-0.05364	Kurtosis	-0.22618
USS	2.716343	CSS	2.716343
CV	.	Std Mean	0.070149
T:Mean=0	0	Prob> T	1.0000

Sgn Rank	2	Prob> S	0.9559
Num ^ = 0	24		
W:Normal	0.985065	Prob<W	0.9603

----- METHOD=DL/2 -----

UNIVARIATE PROCEDURE

Variable=RESID

N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.18433	Variance	0.033978
Skewness	2.218603	Kurtosis	7.198329
USS	0.781484	CSS	0.781484
CV	.	Std Mean	0.037626
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-33	Prob> S	0.3567
Num ^ = 0	24		
W:Normal	0.806615	Prob<W	0.0002

Variable=LOGRESID

N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.372859	Variance	0.139024
Skewness	-0.46143	Kurtosis	0.438804
USS	3.197553	CSS	3.197553
CV	.	Std Mean	0.07611
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	8	Prob> S	0.8247
Num ^ = 0	24		
W:Normal	0.981971	Prob<W	0.9199

----- METHOD=ZERO -----

UNIVARIATE PROCEDURE

Variable=RESID

N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.184631	Variance	0.034089
Skewness	2.205263	Kurtosis	7.132011
USS	0.784041	CSS	0.784041
CV	.	Std Mean	0.037688
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-32	Prob> S	0.3717
Num ^ = 0	24		
W:Normal	0.809082	Prob<W	0.0003

Variable=LOGRESID

N	23	Sum Wgts	23
Mean	0	Sum	0
Std Dev	0.31417	Variance	0.098703
Skewness	0.367923	Kurtosis	0.102775
USS	2.171467	CSS	2.171467
CV	.	Std Mean	0.065509
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-4	Prob> S	0.9063
Num ^ = 0	23		
W:Normal	0.978207	Prob<W	0.8614

LEAD BIOACCUMULATION IN NEREIS VIRENS  
LEVENE'S TEST

----- METHOD=DL -----

General Linear Models Procedure

Dependent Variable: DEV

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.12236044	0.04078681	2.52	0.0868
Error	20	0.32322076	0.01616104		
Corrected Total	23	0.44558120			

```

----- METHOD=DL/2 -----
                          General Linear Models Procedure

Dependent Variable: DEV
Source                    DF          Sum of
                          Squares          Mean
                          Square          F Value          Pr > F
Model                     3          0.12017104          0.04005701          2.47          0.0915
Error                     20          0.32429896          0.01621495
Corrected Total           23          0.44447000

```

```

----- METHOD=ZERO -----
                          General Linear Models Procedure

Dependent Variable: DEV
Source                    DF          Sum of
                          Squares          Mean
                          Square          F Value          Pr > F
Model                     3          0.11803541          0.03934514          2.42          0.0964
Error                     20          0.32550713          0.01627536
Corrected Total           23          0.44354254

```

LEAD BIOACCUMULATION IN NEREIS VIRENS  
SAMPLE VARIANCES

METHOD	SEDIMENT	N	SAMPLE MEANS	SAMPLE VARIANCES
DL	AK	6	0.12450	0.00347
DL	SH	6	0.14100	0.01562
DL	GOW	6	0.18767	0.01845
DL	RH	6	0.30067	0.11828
DL/2	AK	6	0.12267	0.00394
DL/2	SH	6	0.14100	0.01562
DL/2	GOW	6	0.18767	0.01845
DL/2	RH	6	0.30067	0.11828
ZERO	AK	6	0.12083	0.00445
ZERO	SH	6	0.14100	0.01562
ZERO	GOW	6	0.18767	0.01845
ZERO	RH	6	0.30067	0.11828

## PRELIM2.SAS program statements

PRELIM2.SAS should be used when there are several nondetects in one or more treatments, no treatment is completely censored, and sample sizes are equal. This program includes DL, DL/2, MLE WEIB, and UNIF. Example data used in the program are for mercury bioaccumulation in *Macoma nasuta* (Chapter 8).

```

OPTIONS LINESIZE=79 PAGESIZE=500 NODATE NONUMBER;

/* Input the data here or read in an existing permanent SAS data set. We
   recommend that nondetects be coded as -dl where dl is the numeric
   detection limit. To avoid changing statements after the first data step
   each time the program is run, name or rename the contaminant concentration
   variable CONC and the treatment variable TRT. */

DATA A0;
  INPUT TRT $ CONC @@;
  CARDS;
AK .033 AK .066 AK .028 AK .034 AK .034 AK .030
GOW -.020 GOW -.020 GOW .160 GOW -.020 GOW .036
RH .028 RH -.020 RH -.020 RH .207 RH .032 RH .032
SH .036 SH .028 SH .085 SH .023 SH .023 SH .040
;
DATA A; SET A0;
  IF CONC<0 THEN DL=ABS(CONC);
  IF CONC<0 THEN COUNT=1; ELSE COUNT=0;
  ABSCONC=ABS(CONC);
TITLE 'MERCURY BIOACCUMULATION IN MACOMA NASUTA';
PROC SORT; BY TRT ABSCONC;

/* Apply DL/2 */

```

```

DATA DL2;
  SET A;
  IF CONC<0 THEN CONC=ABS(CONC)/2;
  METHOD='DL/2';
PROC PRINT;

/* Apply DL */

DATA DL;
  SET A;
  IF CONC<0 THEN CONC=ABS(CONC);
  METHOD='DL';
PROC PRINT;

/* Apply UNIF */

PROC MEANS NOPRINT DATA=A;
  BY TRT;
  VAR COUNT;
  OUTPUT OUT=B0 SUM=NC N=NREP;
DATA B; SET B0;
  NUC=NREP-NC;
  DROP TYPE _FREQ_;
DATA UNIF;
  MERGE A B; BY TRT;
  IF FIRST.TRT THEN I=1;
  IF CONC<0 THEN DO;
    CONC=DL*(I-1)/(NC-1);
    IF NC=1 THEN CONC=DL/2;
    I+1;
  END;
  METHOD='UNIF';
PROC PRINT;

/* Apply MLE WEIB. The output statement for PROC LIFEREG assumes n=6
  replicates for each treatment. If n=5 replicates for each treatment,
  substitute Q=.01 .25 .5 .75 .99 in the OUTPUT statement. See Section
  D4.5.3 of the Inland Testing Manual for quantiles to use with other
  sample sizes. */

DATA C; SET A;
  IF CONC<0 THEN LOWER=.; ELSE LOWER=CONC;
  IF CONC<0 THEN CONC=DL;
PROC LIFEREG NOPRINT;
  BY TRT;
  MODEL (LOWER,CONC)= /D=WEIBULL SHAPE1=1;
  OUTPUT OUT=C1 P=PRED Q=.01 .21 .4 .6 .79 .99;
PROC SORT; BY TRT _PROB_;
DATA D; SET C1;
  BY TRT _PROB_;
  IF FIRST._PROB_;
  KEEP TRT _PROB_ PRED;
DATA MLEWEIB;
  MERGE D A; BY TRT;
  IF CONC<DL AND PRED>=DL THEN CONC=DL;
  IF CONC<DL AND PRED<DL THEN CONC=PRED;
  METHOD='MLE WEIB';
PROC PRINT;

/* Determine percent of data that are censored */

PROC MEANS NOPRINT DATA=A;
  VAR COUNT;
  OUTPUT OUT=O SUM=SUM N=N;
DATA O1; SET O;
  PROPCENS=SUM*100/N;
PROC PRINT LABEL NOOBS;
  VAR PROPCENS N;
  LABEL PROPCENS='PERCENT OF DATA THAT ARE CENSORED'
        N='TOTAL NUMBER OF REPLICATES';

/* Combine data sets and sort by method. Calculate logs. */

DATA ALL;
  SET DL2 DL UNIF MLEWEIB;
  LOGCONC=LOG10(CONC);
  LABEL TRT='SEDIMENT';
PROC SORT; BY METHOD TRT;

/* Determine CV of combined samples */

PROC MEANS NOPRINT;
  VAR CONC; BY METHOD;
  OUTPUT OUT=O CV=CV;
DATA O1; SET O;
  CV=CV/100;
PROC PRINT LABEL NOOBS;
  VAR METHOD CV;
  LABEL CV='CV OF COMBINED SAMPLES';

/* Test normality of combined sample residuals using Shapiro-Wilk's Test */

```

```

PROC GLM NOPRINT DATA=ALL;
  BY METHOD;
  CLASS TRT;
  MODEL CONC LOGCONC=TRT;
  OUTPUT OUT=Z R=RESID LOGRESID;
PROC UNIVARIATE NORMAL;
  BY METHOD;
  VAR RESID LOGRESID;
  TITLE2 'SHAPIRO-WILK'S TEST';

/* Test equality of variances using Levene's Test */

DATA Z1; SET Z;
  DEV=ABS(RESID);
PROC GLM;
  BY METHOD;
  CLASS TRT;
  MODEL DEV=TRT;
  TITLE2 'LEVENE'S TEST';

/* Calculate sample variances */

PROC MEANS NOPRINT DATA=ALL;
  BY METHOD TRT;
  VAR CONC;
  OUTPUT OUT=O VAR=VARI MEAN=MEAN N=N;
PROC SORT; BY METHOD VARI;
PROC PRINT LABEL NOOBS;
  VAR METHOD TRT N MEAN VARI;
  LABEL VARI='SAMPLE VARIANCES'
        MEAN='SAMPLE MEANS';
  TITLE2 'SAMPLE VARIANCES';

```

## PRELIM2.SAS program output

MERCURY BIOACCUMULATION IN MACOMA NASUTA						
OBS	TRT	CONC	DL	COUNT	ABSCONC	METHOD
1	AK	0.028	.	0	0.028	DL/2
2	AK	0.030	.	0	0.030	DL/2
3	AK	0.033	.	0	0.033	DL/2
4	AK	0.034	.	0	0.034	DL/2
5	AK	0.034	.	0	0.034	DL/2
6	AK	0.066	.	0	0.066	DL/2
7	GOW	0.010	0.02	1	0.020	DL/2
8	GOW	0.010	0.02	1	0.020	DL/2
9	GOW	0.010	0.02	1	0.020	DL/2
10	GOW	0.010	0.02	1	0.020	DL/2
11	GOW	0.036	.	0	0.036	DL/2
12	GOW	0.160	.	0	0.160	DL/2
13	RH	0.010	0.02	1	0.020	DL/2
14	RH	0.010	0.02	1	0.020	DL/2
15	RH	0.028	.	0	0.028	DL/2
16	RH	0.032	.	0	0.032	DL/2
17	RH	0.032	.	0	0.032	DL/2
18	RH	0.207	.	0	0.207	DL/2
19	SH	0.023	.	0	0.023	DL/2
20	SH	0.023	.	0	0.023	DL/2
21	SH	0.028	.	0	0.028	DL/2
22	SH	0.036	.	0	0.036	DL/2
23	SH	0.040	.	0	0.040	DL/2
24	SH	0.085	.	0	0.085	DL/2

OBS	TRT	CONC	DL	COUNT	ABSCONC	METHOD
1	AK	0.028	.	0	0.028	DL
2	AK	0.030	.	0	0.030	DL
3	AK	0.033	.	0	0.033	DL
4	AK	0.034	.	0	0.034	DL
5	AK	0.034	.	0	0.034	DL
6	AK	0.066	.	0	0.066	DL
7	GOW	0.020	0.02	1	0.020	DL
8	GOW	0.020	0.02	1	0.020	DL
9	GOW	0.020	0.02	1	0.020	DL
10	GOW	0.020	0.02	1	0.020	DL
11	GOW	0.036	.	0	0.036	DL
12	GOW	0.160	.	0	0.160	DL
13	RH	0.020	0.02	1	0.020	DL
14	RH	0.020	0.02	1	0.020	DL
15	RH	0.028	.	0	0.028	DL
16	RH	0.032	.	0	0.032	DL
17	RH	0.032	.	0	0.032	DL
18	RH	0.207	.	0	0.207	DL
19	SH	0.023	.	0	0.023	DL
20	SH	0.023	.	0	0.023	DL
21	SH	0.028	.	0	0.028	DL
22	SH	0.036	.	0	0.036	DL
23	SH	0.040	.	0	0.040	DL

OBS	TRT	CONC	DL	COUNT	ABSCONC	NC	NREP	NUC	I	METHOD
1	AK	0.02800	.	0	0.028	0	6	6	1	UNIF
2	AK	0.03000	.	0	0.030	0	6	6	1	UNIF
3	AK	0.03300	.	0	0.033	0	6	6	1	UNIF
4	AK	0.03400	.	0	0.034	0	6	6	1	UNIF
5	AK	0.03400	.	0	0.034	0	6	6	1	UNIF
6	AK	0.06600	.	0	0.066	0	6	6	1	UNIF
7	GOW	0.00000	0.02	1	0.020	4	6	2	2	UNIF
8	GOW	0.00667	0.02	1	0.020	4	6	2	3	UNIF
9	GOW	0.01333	0.02	1	0.020	4	6	2	4	UNIF
10	GOW	0.02000	0.02	1	0.020	4	6	2	5	UNIF
11	GOW	0.03600	.	0	0.036	4	6	2	5	UNIF
12	GOW	0.16000	.	0	0.160	4	6	2	5	UNIF
13	RH	0.00000	0.02	1	0.020	2	6	4	2	UNIF
14	RH	0.02000	0.02	1	0.020	2	6	4	3	UNIF
15	RH	0.02800	.	0	0.028	2	6	4	3	UNIF
16	RH	0.03200	.	0	0.032	2	6	4	3	UNIF
17	RH	0.03200	.	0	0.032	2	6	4	3	UNIF
18	RH	0.20700	.	0	0.207	2	6	4	3	UNIF
19	SH	0.02300	.	0	0.023	0	6	6	1	UNIF
20	SH	0.02300	.	0	0.023	0	6	6	1	UNIF
21	SH	0.02800	.	0	0.028	0	6	6	1	UNIF
22	SH	0.03600	.	0	0.036	0	6	6	1	UNIF
23	SH	0.04000	.	0	0.040	0	6	6	1	UNIF
24	SH	0.08500	.	0	0.085	0	6	6	1	UNIF

OBS	TRT	PROB_	PRED	CONC	DL	COUNT	ABSCONC	METHOD
1	AK	0.01	0.00871	0.02800	.	0	0.028	MLE WEIB
2	AK	0.21	0.02563	0.03000	.	0	0.030	MLE WEIB
3	AK	0.40	0.03338	0.03300	.	0	0.033	MLE WEIB
4	AK	0.60	0.04077	0.03400	.	0	0.034	MLE WEIB
5	AK	0.79	0.04891	0.03400	.	0	0.034	MLE WEIB
6	AK	0.99	0.07081	0.06600	.	0	0.066	MLE WEIB
7	GOW	0.01	0.00000	0.00000	0.02	1	0.020	MLE WEIB
8	GOW	0.21	0.00092	0.00092	0.02	1	0.020	MLE WEIB
9	GOW	0.40	0.00455	0.00455	0.02	1	0.020	MLE WEIB
10	GOW	0.60	0.01518	0.01518	0.02	1	0.020	MLE WEIB
11	GOW	0.79	0.04552	0.03600	.	0	0.036	MLE WEIB
12	GOW	0.99	0.42359	0.16000	.	0	0.160	MLE WEIB
13	RH	0.01	0.00013	0.00013	0.02	1	0.020	MLE WEIB
14	RH	0.21	0.00723	0.00723	0.02	1	0.020	MLE WEIB
15	RH	0.40	0.01930	0.02800	.	0	0.028	MLE WEIB
16	RH	0.60	0.04051	0.03200	.	0	0.032	MLE WEIB
17	RH	0.79	0.07962	0.03200	.	0	0.032	MLE WEIB
18	RH	0.99	0.31432	0.20700	.	0	0.207	MLE WEIB
19	SH	0.01	0.00439	0.02300	.	0	0.023	MLE WEIB
20	SH	0.21	0.02152	0.02300	.	0	0.023	MLE WEIB
21	SH	0.40	0.03177	0.02800	.	0	0.028	MLE WEIB
22	SH	0.60	0.04265	0.03600	.	0	0.036	MLE WEIB
23	SH	0.79	0.05577	0.04000	.	0	0.040	MLE WEIB
24	SH	0.99	0.09618	0.08500	.	0	0.085	MLE WEIB

MERCURY BIOACCUMULATION IN MACOMA NASUTA

PERCENT  
OF DATA  
THAT ARE  
CENSORED

TOTAL  
NUMBER OF  
REPLICATES

25

24

METHOD

CV OF  
COMBINED  
SAMPLES

DL 1.02234  
DL/2 1.12057  
MLE WEIB 1.18322  
UNIF 1.12514

MERCURY BIOACCUMULATION IN MACOMA NASUTA  
SHAPIRO-WILK'S TEST

----- METHOD=DL -----

UNIVARIATE PROCEDURE

Variable=RESID

N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.045152	Variance	0.002039
Skewness	2.461642	Kurtosis	5.992476
USS	0.04689	CSS	0.04689
CV	.	Std Mean	0.009217
T:Mean=0	0	Prob> T	1.0000

Sgn Rank	-61	Prob> S	0.0809
Num ^= 0	24		
W:Normal	0.666305	Prob<W	0.0001

Variable=LOGRESID

N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.271824	Variance	0.073889
Skewness	1.845132	Kurtosis	2.98889
USS	1.699436	CSS	1.699436
CV	.	Std Mean	0.055486
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-44	Prob> S	0.2156
Num ^= 0	24		
W:Normal	0.764796	Prob<W	0.0001

----- METHOD=DL/2 -----

Variable=RESID			
UNIVARIATE PROCEDURE			
N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.046949	Variance	0.002204
Skewness	2.387347	Kurtosis	5.722725
USS	0.050697	CSS	0.050696
CV	.	Std Mean	0.009583
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-65	Prob> S	0.0616
Num ^= 0	24		
W:Normal	0.690165	Prob<W	0.0001

Variable=LOGRESID			
N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.343952	Variance	0.118303
Skewness	1.351413	Kurtosis	2.110561
USS	2.720962	CSS	2.720962
CV	.	Std Mean	0.070209
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-32	Prob> S	0.3717
Num ^= 0	24		
W:Normal	0.86895	Prob<W	0.0042

----- METHOD=MLE WEIB -----

Variable=RESID			
UNIVARIATE PROCEDURE			
N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.048084	Variance	0.002312
Skewness	2.315833	Kurtosis	5.517361
USS	0.053177	CSS	0.053177
CV	.	Std Mean	0.009815
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-64	Prob> S	0.0661
Num ^= 0	24		
W:Normal	0.712983	Prob<W	0.0001

Variable=LOGRESID			
N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.98422	Variance	0.968689
Skewness	-1.76278	Kurtosis	5.565335
USS	22.27985	CSS	22.27985
CV	.	Std Mean	0.200903
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	23	Prob> S	0.5227
Num ^= 0	24		
W:Normal	0.798226	Prob<W	0.0002

----- METHOD=UNIF -----

Variable=RESID			
UNIVARIATE PROCEDURE			
N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.047144	Variance	0.002223
Skewness	2.337253	Kurtosis	5.611203
USS	0.051119	CSS	0.051119
CV	.	Std Mean	0.009623
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-64	Prob> S	0.0661
Num ^= 0	24		
W:Normal	0.707571	Prob<W	0.0001

Variable=LOGRESID			
N	22	Sum Wgts	22
Mean	0	Sum	0

```

Std Dev      0.312334  Variance    0.097552
Skewness    1.155918  Kurtosis   1.943256
USS         2.048597  CSS        2.048597
CV          .          Std Mean   0.06659
T:Mean=0    0          Prob>|T|   1.0000
Sgn Rank   -30.5     Prob>|S|   0.3336
Num ^= 0    22
W:Normal    0.878027  Prob<W     0.0095

```

MERCURY BIOACCUMULATION IN MACOMA NASUTA  
LEVENE'S TEST

```

----- METHOD=DL -----
General Linear Models Procedure

Dependent Variable: DEV

Source          DF          Sum of Squares          Mean Square          F Value          Pr > F
Model              3          0.00652861          0.00217620          2.06          0.1379
Error             20          0.02113231          0.00105662
Corrected Total   23          0.02766093

```

```

----- METHOD=DL/2 -----
General Linear Models Procedure

Dependent Variable: DEV

Source          DF          Sum of Squares          Mean Square          F Value          Pr > F
Model              3          0.00709898          0.00236633          2.04          0.1409
Error             20          0.02321972          0.00116099
Corrected Total   23          0.03031870

```

```

----- METHOD=MLE WEIB -----
General Linear Models Procedure

Dependent Variable: DEV

Source          DF          Sum of Squares          Mean Square          F Value          Pr > F
Model              3          0.00743350          0.00247783          2.00          0.1460
Error             20          0.02473968          0.00123698
Corrected Total   23          0.03217318

```

```

----- METHOD=UNIF -----
General Linear Models Procedure

Dependent Variable: DEV

Source          DF          Sum of Squares          Mean Square          F Value          Pr > F
Model              3          0.00709898          0.00236633          2.00          0.1462
Error             20          0.02364194          0.00118210
Corrected Total   23          0.03074093

```

MERCURY BIOACCUMULATION IN MACOMA NASUTA  
SAMPLE VARIANCES

METHOD	SEDIMENT	N	SAMPLE MEANS	SAMPLE VARIANCES
DL	AK	6	0.037500	.0002007
DL	SH	6	0.039167	.0005518
DL	GOW	6	0.046000	.0031600
DL	RH	6	0.056500	.0054655
DL/2	AK	6	0.037500	.0002007
DL/2	SH	6	0.039167	.0005518
DL/2	GOW	6	0.039333	.0036027
DL/2	RH	6	0.053167	.0057842
MLE WEIB	AK	6	0.037500	.0002007
MLE WEIB	SH	6	0.039167	.0005518

MLE WEIB	GOW	6	0.036110	.0038649
MLE WEIB	RH	6	0.051061	.0060180
UNIF	AK	6	0.037500	.0002007
UNIF	SH	6	0.039167	.0005518
UNIF	GOW	6	0.039333	.0036471
UNIF	RH	6	0.053167	.0058242

## PRELIM2A.SAS program statements

PRELIM2A.SAS should be used when there are several nondetects in one or more treatments, no treatment is completely censored, and sample sizes are unequal. This program includes DL, DL/2, MLE WEIB, and UNIF. MLE WEIB must be modified for the appropriate sample sizes. The MLE methods are probably not worth the extra programming effort if there are more than two different sample sizes. Example data used in the program are for PCB 52 bioaccumulation in *Nereis virens* (Chapter 8).

```

OPTIONS LINESIZE=79 PAGESIZE=500 NODATE NONUMBER;

/* Input the data here or read in an existing permanent SAS data set. We
   recommend that nondetects be coded as -dl where dl is the numeric
   detection limit. To avoid changing statements after the first data step
   each time the program is run, name or rename the contaminant concentration
   variable CONC and the treatment variable TRT. */

DATA A0;
  INPUT TRT $ CONC @@;
  CARDS;
AK 12.0 AK -.64 AK 8.7 AK 16.0 AK 8.0
GOW 12.0 GOW -.90 GOW 13.0 GOW 11.0 GOW 8.9
RH 5.9 RH -.96 RH 4.5 RH -.98 RH -.92 RH -.75
SH 4.3 SH 3.5 SH -1.2 SH -.82 SH -1.0 SH -.64
;
DATA A; SET A0;
  IF CONC<0 THEN DL=ABS(CONC);
  IF CONC<0 THEN COUNT=1; ELSE COUNT=0;
  ABSCONC=ABS(CONC);
TITLE 'PCB 52 BIOACCUMULATION IN NEREIS VIRENS';
PROC SORT; BY TRT ABSCONC;

/* Apply DL/2 */

DATA DL2;
  SET A;
  IF CONC<0 THEN CONC=ABS(CONC)/2;
  METHOD='DL/2';
PROC PRINT;

/* Apply DL */

DATA DL;
  SET A;
  IF CONC<0 THEN CONC=ABS(CONC);
  METHOD='DL';
PROC PRINT;

/* Apply UNIF */

PROC MEANS NOPRINT DATA=A;
  BY TRT;
  VAR COUNT;
  OUTPUT OUT=B0 SUM=NC N=NREP;
DATA B; SET B0;
  NUC=NREP-NC;
  DROP TYPE _FREQ_;
DATA UNIF;
  MERGE A B; BY TRT;
  IF FIRST.TRT THEN I=1;
  IF CONC<0 THEN DO;
    CONC=DL*(I-1)/(NC-1);
    IF NC=1 THEN CONC=DL/2;
    I+1;
  END;
  METHOD='UNIF';
PROC PRINT;

/* Apply MLE WEIB. This method is run first for the treatments with 5
   replicates and then for the treatments with 6 replicates. The
   resulting data are then combined. See Section D4.5.3 of the Inland
   Testing Manual for quantiles to use in the OUTPUT statements for
   other sample sizes. */

```

```

DATA C; SET A;
  IF CONC<0 THEN LOWER=.; ELSE LOWER=CONC;
  IF CONC<0 THEN CONC=DL;
DATA CA; SET C;
  IF TRT='AK' OR TRT='GOW';
PROC LIFEREG NOPRINT;
  BY TRT;
  MODEL (LOWER,CONC)= /D=WEIBULL SHAPE1=1;
  OUTPUT OUT=C1 P=PRED Q=.01 .25 .5 .75 .99;
DATA CB; SET C;
  IF TRT='RH' OR TRT='SH';
PROC LIFEREG NOPRINT;
  BY TRT;
  MODEL (LOWER,CONC)= /D=WEIBULL SHAPE1=1;
  OUTPUT OUT=C2 P=PRED Q=.01 .21 .4 .6 .79 .99;
DATA C3; SET C1 C2;
PROC SORT; BY TRT _PROB_;
DATA D; SET C3;
  BY TRT _PROB_;
  IF FIRST._PROB_;
  KEEP TRT _PROB_ PRED;
DATA MLEWEIB;
  MERGE D A; BY TRT;
  IF CONC<DL AND PRED>=DL THEN CONC=DL;
  IF CONC<DL AND PRED<DL THEN CONC=PRED;
  METHOD='MLE WEIB';
PROC PRINT;

/* Determine percent of data that are censored */
PROC MEANS NOPRINT DATA=A;
  VAR COUNT;
  OUTPUT OUT=O SUM=SUM N=N;
DATA O1; SET O;
  PROPCENS=SUM*100/N;
PROC PRINT LABEL NOOBS;
  VAR PROPCENS N;
  LABEL PROPCENS='PERCENT OF DATA THAT ARE CENSORED'
        N='TOTAL NUMBER OF REPLICATES';

/* Combine data sets and sort by method. Calculate logs. */
DATA ALL;
  SET DL2 DL UNIF MLEWEIB;
  LOGCONC=LOG10(CONC);
  LABEL TRT='SEDIMENT';
PROC SORT; BY METHOD TRT;

/* Determine CV of combined samples */
PROC MEANS NOPRINT;
  VAR CONC; BY METHOD;
  OUTPUT OUT=O CV=CV;
DATA O1; SET O;
  CV=CV/100;
PROC PRINT LABEL NOOBS;
  VAR METHOD CV;
  LABEL CV='CV OF COMBINED SAMPLES';

/* Test normality of combined sample residuals using Shapiro-Wilk's Test */
PROC GLM NOPRINT DATA=ALL;
  BY METHOD;
  CLASS TRT;
  MODEL CONC LOGCONC=TRT;
  OUTPUT OUT=Z R=RESID LOGRESID;
PROC UNIVARIATE NORMAL;
  BY METHOD;
  VAR RESID LOGRESID;
  TITLE2 'SHAPIRO-WILK'S TEST';

/* Test equality of variances using Levene's Test */
DATA Z1; SET Z;
  DEV=ABS(RESID);
PROC GLM;
  BY METHOD;
  CLASS TRT;
  MODEL DEV=TRT;
  TITLE2 'LEVENE'S TEST';

/* Calculate sample variances */
PROC MEANS NOPRINT DATA=ALL;
  BY METHOD TRT;
  VAR CONC;
  OUTPUT OUT=O VAR=VARI MEAN=MEAN N=N;
PROC SORT; BY METHOD VARI;
PROC PRINT LABEL NOOBS;
  VAR METHOD TRT N MEAN VARI;
  LABEL VARI='SAMPLE VARIANCES'

```

MEAN='SAMPLE MEANS';  
 TITLE2 'SAMPLE VARIANCES';

### PRELIM2A.SAS program output

#### PCB 52 BIOACCUMULATION IN NEREIS VIRENS

OBS	TRT	CONC	DL	COUNT	ABSCONC	METHOD
1	AK	0.320	0.64	1	0.64	DL/2
2	AK	8.000	.	0	8.00	DL/2
3	AK	8.700	.	0	8.70	DL/2
4	AK	12.000	.	0	12.00	DL/2
5	AK	16.000	.	0	16.00	DL/2
6	GOW	0.450	0.90	1	0.90	DL/2
7	GOW	8.900	.	0	8.90	DL/2
8	GOW	11.000	.	0	11.00	DL/2
9	GOW	12.000	.	0	12.00	DL/2
10	GOW	13.000	.	0	13.00	DL/2
11	RH	0.375	0.75	1	0.75	DL/2
12	RH	0.460	0.92	1	0.92	DL/2
13	RH	0.480	0.96	1	0.96	DL/2
14	RH	0.490	0.98	1	0.98	DL/2
15	RH	4.500	.	0	4.50	DL/2
16	RH	5.900	.	0	5.90	DL/2
17	SH	0.320	0.64	1	0.64	DL/2
18	SH	0.410	0.82	1	0.82	DL/2
19	SH	0.500	1.00	1	1.00	DL/2
20	SH	0.600	1.20	1	1.20	DL/2
21	SH	3.500	.	0	3.50	DL/2
22	SH	4.300	.	0	4.30	DL/2

OBS	TRT	CONC	DL	COUNT	ABSCONC	METHOD
1	AK	0.64	0.64	1	0.64	DL
2	AK	8.00	.	0	8.00	DL
3	AK	8.70	.	0	8.70	DL
4	AK	12.00	.	0	12.00	DL
5	AK	16.00	.	0	16.00	DL
6	GOW	0.90	0.90	1	0.90	DL
7	GOW	8.90	.	0	8.90	DL
8	GOW	11.00	.	0	11.00	DL
9	GOW	12.00	.	0	12.00	DL
10	GOW	13.00	.	0	13.00	DL
11	RH	0.75	0.75	1	0.75	DL
12	RH	0.92	0.92	1	0.92	DL
13	RH	0.96	0.96	1	0.96	DL
14	RH	0.98	0.98	1	0.98	DL
15	RH	4.50	.	0	4.50	DL
16	RH	5.90	.	0	5.90	DL
17	SH	0.64	0.64	1	0.64	DL
18	SH	0.82	0.82	1	0.82	DL
19	SH	1.00	1.00	1	1.00	DL
20	SH	1.20	1.20	1	1.20	DL
21	SH	3.50	.	0	3.50	DL
22	SH	4.30	.	0	4.30	DL

OBS	TRT	CONC	DL	COUNT	ABSCONC	NC	NREP	NUC	I	METHOD
1	AK	0.3200	0.64	1	0.64	1	5	4	2	UNIF
2	AK	8.0000	.	0	8.00	1	5	4	2	UNIF
3	AK	8.7000	.	0	8.70	1	5	4	2	UNIF
4	AK	12.0000	.	0	12.00	1	5	4	2	UNIF
5	AK	16.0000	.	0	16.00	1	5	4	2	UNIF
6	GOW	0.4500	0.90	1	0.90	1	5	4	2	UNIF
7	GOW	8.9000	.	0	8.90	1	5	4	2	UNIF
8	GOW	11.0000	.	0	11.00	1	5	4	2	UNIF
9	GOW	12.0000	.	0	12.00	1	5	4	2	UNIF
10	GOW	13.0000	.	0	13.00	1	5	4	2	UNIF
11	RH	0.0000	0.75	1	0.75	4	6	2	2	UNIF
12	RH	0.3067	0.92	1	0.92	4	6	2	3	UNIF
13	RH	0.6400	0.96	1	0.96	4	6	2	4	UNIF
14	RH	0.9800	0.98	1	0.98	4	6	2	5	UNIF
15	RH	4.5000	.	0	4.50	4	6	2	5	UNIF
16	RH	5.9000	.	0	5.90	4	6	2	5	UNIF
17	SH	0.0000	0.64	1	0.64	4	6	2	2	UNIF
18	SH	0.2733	0.82	1	0.82	4	6	2	3	UNIF
19	SH	0.6667	1.00	1	1.00	4	6	2	4	UNIF
20	SH	1.2000	1.20	1	1.20	4	6	2	5	UNIF
21	SH	3.5000	.	0	3.50	4	6	2	5	UNIF
22	SH	4.3000	.	0	4.30	4	6	2	5	UNIF

OBS	TRT	_PROB_	PRED	CONC	DL	COUNT	ABSCONC	METHOD
1	AK	0.01	0.2254	0.2254	0.64	1	0.64	MLE WEIB
2	AK	0.25	3.4319	8.0000	.	0	8.00	MLE WEIB
3	AK	0.50	7.0076	8.7000	.	0	8.70	MLE WEIB
4	AK	0.75	12.3010	12.0000	.	0	12.00	MLE WEIB
5	AK	0.99	32.5983	16.0000	.	0	16.00	MLE WEIB

6	GOW	0.01	0.4785	0.4785	0.90	1	0.90	MLE WEIB
7	GOW	0.25	4.3134	8.9000	.	0	8.90	MLE WEIB
8	GOW	0.50	7.6766	11.0000	.	0	11.00	MLE WEIB
9	GOW	0.75	12.0920	12.0000	.	0	12.00	MLE WEIB
10	GOW	0.99	26.5631	13.0000	.	0	13.00	MLE WEIB
11	RH	0.01	0.0001	0.0001	0.75	1	0.75	MLE WEIB
12	RH	0.21	0.0469	0.0469	0.92	1	0.92	MLE WEIB
13	RH	0.40	0.2395	0.2395	0.96	1	0.96	MLE WEIB
14	RH	0.60	0.8212	0.8212	0.98	1	0.98	MLE WEIB
15	RH	0.79	2.5241	4.5000	.	0	4.50	MLE WEIB
16	RH	0.99	24.7188	5.9000	.	0	5.90	MLE WEIB
17	SH	0.01	0.0003	0.0003	0.64	1	0.64	MLE WEIB
18	SH	0.21	0.0731	0.0731	0.82	1	0.82	MLE WEIB
19	SH	0.40	0.2904	0.2904	1.00	1	1.00	MLE WEIB
20	SH	0.60	0.8231	0.8231	1.20	1	1.20	MLE WEIB
21	SH	0.79	2.1274	3.5000	.	0	3.50	MLE WEIB
22	SH	0.99	14.6504	4.3000	.	0	4.30	MLE WEIB

PCB 52 BIOACCUMULATION IN NEREIS VIRENS

PERCENT  
OF DATA  
THAT ARE  
CENSORED

TOTAL  
NUMBER OF  
REPLICATES

45.4545                      22

METHOD

CV OF  
COMBINED  
SAMPLES

DL                      0.94659  
DL/2                    1.02010  
MLE WEIB               1.04572  
UNIF                    1.01378

PCB 52 BIOACCUMULATION IN NEREIS VIRENS  
SHAPIRO-WILK'S TEST

----- METHOD=DL -----

UNIVARIATE PROCEDURE

Variable=RESID

N	22	Sum Wgts	22
Mean	0	Sum	0
Std Dev	3.528519	Variance	12.45045
Skewness	-0.79011	Kurtosis	1.682036
USS	261.4594	CSS	261.4594
CV	.	Std Mean	0.752283
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	17.5	Prob> S	0.5819
Num ^ = 0	22		
W:Normal	0.889263	Prob<W	0.0163

Variable=LOGRESID

N	22	Sum Wgts	22
Mean	0	Sum	0
Std Dev	0.41402	Variance	0.171413
Skewness	-0.85768	Kurtosis	0.423371
USS	3.599672	CSS	3.599672
CV	.	Std Mean	0.088269
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	14.5	Prob> S	0.6486
Num ^ = 0	22		
W:Normal	0.91975	Prob<W	0.0728

----- METHOD=DL/2 -----

UNIVARIATE PROCEDURE

Variable=RESID

N	22	Sum Wgts	22
Mean	0	Sum	0
Std Dev	3.675792	Variance	13.51145
Skewness	-0.79843	Kurtosis	1.509502
USS	283.7404	CSS	283.7404
CV	.	Std Mean	0.783681
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	17.5	Prob> S	0.5819
Num ^ = 0	22		
W:Normal	0.889369	Prob<W	0.0164

Variable=LOGRESID

N	22	Sum Wgts	22
Mean	0	Sum	0
Std Dev	0.54443	Variance	0.296404
Skewness	-0.67036	Kurtosis	0.009256
USS	6.224482	CSS	6.224482

```

CV . Std Mean 0.116073
T:Mean=0 0 Prob>|T| 1.0000
Sgn Rank 7.5 Prob>|S| 0.8140
Num ^= 0 22
W:Normal 0.926628 Prob<W 0.1023

```

----- METHOD=MLE WEIB -----

```

UNIVARIATE PROCEDURE
Variable=RESID
N 22 Sum Wgts 22
Mean 0 Sum 0
Std Dev 3.712581 Variance 13.78326
Skewness -0.77896 Kurtosis 1.391436
USS 289.4484 CSS 289.4484
CV . Std Mean 0.791525
T:Mean=0 0 Prob>|T| 1.0000
Sgn Rank 17.5 Prob>|S| 0.5819
Num ^= 0 22
W:Normal 0.90111 Prob<W 0.0290

```

```

Variable=LOGRESID
N 22 Sum Wgts 22
Mean 0 Sum 0
Std Dev 1.25576 Variance 1.576933
Skewness -1.47097 Kurtosis 2.295472
USS 33.11559 CSS 33.11559
CV . Std Mean 0.267729
T:Mean=0 0 Prob>|T| 1.0000
Sgn Rank 28.5 Prob>|S| 0.3669
Num ^= 0 22
W:Normal 0.84414 Prob<W 0.0020

```

----- METHOD=UNIF -----

```

UNIVARIATE PROCEDURE
Variable=RESID
N 22 Sum Wgts 22
Mean 0 Sum 0
Std Dev 3.677161 Variance 13.52151
Skewness -0.80508 Kurtosis 1.502431
USS 283.9518 CSS 283.9518
CV . Std Mean 0.783973
T:Mean=0 0 Prob>|T| 1.0000
Sgn Rank 15.5 Prob>|S| 0.6261
Num ^= 0 22
W:Normal 0.900724 Prob<W 0.0285

```

```

Variable=LOGRESID
N 20 Sum Wgts 20
Mean 0 Sum 0
Std Dev 0.548097 Variance 0.30041
Skewness -1.05128 Kurtosis 0.149698
USS 5.707791 CSS 5.707791
CV . Std Mean 0.122558
T:Mean=0 0 Prob>|T| 1.0000
Sgn Rank 12 Prob>|S| 0.6742
Num ^= 0 20
W:Normal 0.87515 Prob<W 0.0138

```

PCB 52 BIOACCUMULATION IN NEREIS VIRENS  
LEVENE'S TEST

----- METHOD=DL -----

General Linear Models Procedure

```

Dependent Variable: DEV
Source DF Sum of Squares Mean Square F Value Pr > F
Model 3 24.85583147 8.28527716 1.60 0.2239
Error 18 93.09944653 5.17219147
Corrected Total 21 117.95527800

```

----- METHOD=DL/2 -----

General Linear Models Procedure

```

Dependent Variable: DEV
Source DF Sum of Squares Mean Square F Value Pr > F
Model 3 22.03579727 7.34526576 1.31 0.3033

```

Error	18	101.28499111	5.62694395
Corrected Total	21	123.32078839	

----- METHOD=MLE WEIB -----

General Linear Models Procedure

Dependent Variable: DEV					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	20.65247063	6.88415688	1.20	0.3382
Error	18	103.28651631	5.73813980		
Corrected Total	21	123.93898694			

----- METHOD=UNIF -----

General Linear Models Procedure

Dependent Variable: DEV					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	22.61966077	7.53988692	1.32	0.2974
Error	18	102.46296575	5.69238699		
Corrected Total	21	125.08262652			

PCB 52 BIOACCUMULATION IN NEREIS VIRENS  
SAMPLE VARIANCES

METHOD	SEDIMENT	N	SAMPLE MEANS	SAMPLE VARIANCES
DL	SH	6	1.91000	2.4747
DL	RH	6	2.33500	5.1275
DL	GOW	5	9.16000	23.6230
DL	AK	5	9.06800	32.2391
DL/2	SH	6	1.60500	3.2329
DL/2	RH	6	2.03417	6.2111
DL/2	GOW	5	9.07000	25.5220
DL/2	AK	5	9.00400	33.6081
MLE WEIB	SH	6	1.49781	3.6093
MLE WEIB	RH	6	1.91795	6.7446
MLE WEIB	GOW	5	9.07571	25.3992
MLE WEIB	AK	5	8.98508	34.0205
UNIF	SH	6	1.65667	3.2464
UNIF	RH	6	2.05444	6.2399
UNIF	GOW	5	9.07000	25.5220
UNIF	AK	5	9.00400	33.6081

### PRELIM3.SAS program statements

PRELIM3.SAS should be used when a treatment is completely censored. This program includes DL, DL/2, UNIF, and ZERO. Example data used in the program are for mercury bioaccumulation in *Nereis virens* (Table 7 in Chapter 5).

```

OPTIONS LINESIZE=79 PAGESIZE=500 NODATE NONUMBER;

/* Input the data here or read in an existing permanent SAS data set. We
  recommend that nondetects be coded as -dl where dl is the numeric
  detection limit. To avoid changing statements after the first data step
  each time the program is run, name or rename the contaminant concentration
  variable CONC and the treatment variable TRT. */

DATA A0;
  INPUT TRT $ CONC @@;
  CARDS;
AK -.02 AK -.02 AK -.02 AK -.02 AK -.02 AK -.02
GOW .029 GOW -.020 GOW -.020 GOW .024 GOW -.020 GOW -.02
RH .043 RH .038 RH -.020 RH -.02 RH -.02 RH -.02
SH .029 SH -.02 SH .024 SH .022 SH -.02 SH -.02
;

```

```

DATA A; SET A0;
  IF CONC<0 THEN DL=ABS(CONC);
  IF CONC<0 THEN COUNT=1; ELSE COUNT=0;
  ABSCONC=ABS(CONC);
TITLE 'MERCURY BIOACCUMULATION IN NEREIS VIRENS';
PROC SORT; BY TRT ABSCONC;

/* Apply DL/2 */

DATA DL2;
  SET A;
  IF CONC<0 THEN CONC=ABS(CONC)/2;
  METHOD='DL/2';
PROC PRINT;

/* Apply DL */

DATA DL;
  SET A;
  IF CONC<0 THEN CONC=ABS(CONC);
  METHOD='DL';
PROC PRINT;

/* Apply ZERO */

DATA ZERO;
  SET A;
  IF CONC<0 THEN CONC=0;
  METHOD='ZERO';
PROC PRINT;

/* Apply UNIF */

PROC MEANS NOPRINT DATA=A;
  BY TRT;
  VAR COUNT;
  OUTPUT OUT=B0 SUM=NC N=NREP;
DATA B; SET B0;
  NUC=NREP-NC;
  DROP _TYPE_ _FREQ_;
DATA UNIF;
  MERGE A B; BY TRT;
  IF FIRST.TRT THEN I=1;
  IF CONC<0 THEN DO;
    CONC=DL*(I-1)/(NC-1);
    IF NC=1 THEN CONC=DL/2;
    I+1;
  END;
  METHOD='UNIF';
PROC PRINT;

/* Determine percent of data that are censored */

PROC MEANS NOPRINT DATA=A;
  VAR COUNT;
  OUTPUT OUT=O SUM=SUM N=N;
DATA O1; SET O;
  PROPCENS=SUM*100/N;
PROC PRINT LABEL NOOBS;
  VAR PROPCENS N;
  LABEL PROPCENS='PERCENT OF DATA THAT ARE CENSORED'
        N='TOTAL NUMBER OF REPLICATES';

/* Combine data sets and sort by method. Calculate logs. */

DATA ALL;
  SET DL2 DL UNIF ZERO;
  LOGCONC=LOG10(CONC);
  LABEL TRT='SEDIMENT';
PROC SORT; BY METHOD TRT;

/* Determine CV of combined samples */

PROC MEANS NOPRINT;
  VAR CONC; BY METHOD;
  OUTPUT OUT=O CV=CV;
DATA O1; SET O;
  CV=CV/100;
PROC PRINT LABEL NOOBS;
  VAR METHOD CV;
  LABEL CV='CV OF COMBINED SAMPLES';

/* Test normality of combined sample residuals using Shapiro-Wilk's Test */

PROC GLM NOPRINT DATA=ALL;
  BY METHOD;
  CLASS TRT;
  MODEL CONC LOGCONC=TRT;
  OUTPUT OUT=Z R=RESID LOGRESID;
PROC UNIVARIATE NORMAL;
  BY METHOD;

```

```

VAR RESID LOGRESID;
TITLE2 'SHAPIRO-WILK'S TEST';

/* Test equality of variances using Levene's Test */

DATA Z1; SET Z;
DEV=ABS (RESID);
PROC GLM;
BY METHOD;
CLASS TRT;
MODEL DEV=TRT;
TITLE2 'LEVENE'S TEST';

/* Calculate sample variances */

PROC MEANS NOPRINT DATA=ALL;
BY METHOD TRT;
VAR CONC;
OUTPUT OUT=O VAR=VARI MEAN=MEAN N=N;
PROC SORT; BY METHOD VARI;
PROC PRINT LABEL NOOBS;
VAR METHOD TRT N MEAN VARI;
LABEL VARI='SAMPLE VARIANCES'
      MEAN='SAMPLE MEANS';
TITLE2 'SAMPLE VARIANCES';

```

### PRELIM3.SAS program output

MERCURY BIOACCUMULATION IN NEREIS VIRENS						
OBS	TRT	CONC	DL	COUNT	ABSCONC	METHOD
1	AK	0.010	0.02	1	0.020	DL/2
2	AK	0.010	0.02	1	0.020	DL/2
3	AK	0.010	0.02	1	0.020	DL/2
4	AK	0.010	0.02	1	0.020	DL/2
5	AK	0.010	0.02	1	0.020	DL/2
6	AK	0.010	0.02	1	0.020	DL/2
7	GOW	0.010	0.02	1	0.020	DL/2
8	GOW	0.010	0.02	1	0.020	DL/2
9	GOW	0.010	0.02	1	0.020	DL/2
10	GOW	0.010	0.02	1	0.020	DL/2
11	GOW	0.024	.	0	0.024	DL/2
12	GOW	0.029	.	0	0.029	DL/2
13	RH	0.010	0.02	1	0.020	DL/2
14	RH	0.010	0.02	1	0.020	DL/2
15	RH	0.010	0.02	1	0.020	DL/2
16	RH	0.010	0.02	1	0.020	DL/2
17	RH	0.038	.	0	0.038	DL/2
18	RH	0.043	.	0	0.043	DL/2
19	SH	0.010	0.02	1	0.020	DL/2
20	SH	0.010	0.02	1	0.020	DL/2
21	SH	0.010	0.02	1	0.020	DL/2
22	SH	0.022	.	0	0.022	DL/2
23	SH	0.024	.	0	0.024	DL/2
24	SH	0.029	.	0	0.029	DL/2

OBS	TRT	CONC	DL	COUNT	ABSCONC	METHOD
1	AK	0.020	0.02	1	0.020	DL
2	AK	0.020	0.02	1	0.020	DL
3	AK	0.020	0.02	1	0.020	DL
4	AK	0.020	0.02	1	0.020	DL
5	AK	0.020	0.02	1	0.020	DL
6	AK	0.020	0.02	1	0.020	DL
7	GOW	0.020	0.02	1	0.020	DL
8	GOW	0.020	0.02	1	0.020	DL
9	GOW	0.020	0.02	1	0.020	DL
10	GOW	0.020	0.02	1	0.020	DL
11	GOW	0.024	.	0	0.024	DL
12	GOW	0.029	.	0	0.029	DL
13	RH	0.020	0.02	1	0.020	DL
14	RH	0.020	0.02	1	0.020	DL
15	RH	0.020	0.02	1	0.020	DL
16	RH	0.020	0.02	1	0.020	DL
17	RH	0.038	.	0	0.038	DL
18	RH	0.043	.	0	0.043	DL
19	SH	0.020	0.02	1	0.020	DL
20	SH	0.020	0.02	1	0.020	DL
21	SH	0.020	0.02	1	0.020	DL
22	SH	0.022	.	0	0.022	DL
23	SH	0.024	.	0	0.024	DL
24	SH	0.029	.	0	0.029	DL

OBS	TRT	CONC	DL	COUNT	ABSCONC	METHOD
1	AK	0.000	0.02	1	0.020	ZERO
2	AK	0.000	0.02	1	0.020	ZERO

3	AK	0.000	0.02	1	0.020	ZERO
4	AK	0.000	0.02	1	0.020	ZERO
5	AK	0.000	0.02	1	0.020	ZERO
6	AK	0.000	0.02	1	0.020	ZERO
7	GOW	0.000	0.02	1	0.020	ZERO
8	GOW	0.000	0.02	1	0.020	ZERO
9	GOW	0.000	0.02	1	0.020	ZERO
10	GOW	0.000	0.02	1	0.020	ZERO
11	GOW	0.024	.	0	0.024	ZERO
12	GOW	0.029	.	0	0.029	ZERO
13	RH	0.000	0.02	1	0.020	ZERO
14	RH	0.000	0.02	1	0.020	ZERO
15	RH	0.000	0.02	1	0.020	ZERO
16	RH	0.000	0.02	1	0.020	ZERO
17	RH	0.038	.	0	0.038	ZERO
18	RH	0.043	.	0	0.043	ZERO
19	SH	0.000	0.02	1	0.020	ZERO
20	SH	0.000	0.02	1	0.020	ZERO
21	SH	0.000	0.02	1	0.020	ZERO
22	SH	0.022	.	0	0.022	ZERO
23	SH	0.024	.	0	0.024	ZERO
24	SH	0.029	.	0	0.029	ZERO

OBS	TRT	CONC	DL	COUNT	ABSCONC	NC	NREP	NUC	I	METHOD
1	AK	0.000000	0.02	1	0.020	6	6	0	2	UNIF
2	AK	0.004000	0.02	1	0.020	6	6	0	3	UNIF
3	AK	0.008000	0.02	1	0.020	6	6	0	4	UNIF
4	AK	0.012000	0.02	1	0.020	6	6	0	5	UNIF
5	AK	0.016000	0.02	1	0.020	6	6	0	6	UNIF
6	AK	0.020000	0.02	1	0.020	6	6	0	7	UNIF
7	GOW	0.000000	0.02	1	0.020	4	6	2	2	UNIF
8	GOW	0.006667	0.02	1	0.020	4	6	2	3	UNIF
9	GOW	0.013333	0.02	1	0.020	4	6	2	4	UNIF
10	GOW	0.020000	0.02	1	0.020	4	6	2	5	UNIF
11	GOW	0.024000	.	0	0.024	4	6	2	5	UNIF
12	GOW	0.029000	.	0	0.029	4	6	2	5	UNIF
13	RH	0.000000	0.02	1	0.020	4	6	2	2	UNIF
14	RH	0.006667	0.02	1	0.020	4	6	2	3	UNIF
15	RH	0.013333	0.02	1	0.020	4	6	2	4	UNIF
16	RH	0.020000	0.02	1	0.020	4	6	2	5	UNIF
17	RH	0.038000	.	0	0.038	4	6	2	5	UNIF
18	RH	0.043000	.	0	0.043	4	6	2	5	UNIF
19	SH	0.000000	0.02	1	0.020	3	6	3	2	UNIF
20	SH	0.010000	0.02	1	0.020	3	6	3	3	UNIF
21	SH	0.020000	0.02	1	0.020	3	6	3	4	UNIF
22	SH	0.022000	.	0	0.022	3	6	3	4	UNIF
23	SH	0.024000	.	0	0.024	3	6	3	4	UNIF
24	SH	0.029000	.	0	0.029	3	6	3	4	UNIF

MERCURY BIOACCUMULATION IN NEREIS VIRENS

PERCENT  
OF DATA  
THAT ARE  
CENSORED

TOTAL  
NUMBER OF  
REPLICATES

70.8333                      24

METHOD                      CV OF  
COMBINED  
SAMPLES

DL                      0.26617  
DL/2                      0.63681  
UNIF                      0.75279  
ZERO                      1.65734

MERCURY BIOACCUMULATION IN NEREIS VIRENS  
SHAPIRO-WILK'S TEST

----- METHOD=DL -----

Variable=RESID		UNIVARIATE PROCEDURE	
N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.005537	Variance	0.000031
Skewness	1.35701	Kurtosis	2.377675
USS	0.000705	CSS	0.000705
CV	.	Std Mean	0.00113
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-18.5	Prob> S	0.4367
Num ^= 0	18		
W:Normal	0.853674	Prob<W	0.0020

Variable=LOGRESID

N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.085821	Variance	0.007365
Skewness	1.228296	Kurtosis	1.407239
USS	0.169399	CSS	0.169399
CV	.	Std Mean	0.017518
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-14.5	Prob> S	0.5428
Num ^ = 0	18		
W:Normal	0.86626	Prob<W	0.0037

----- METHOD=DL/2 -----

Variable=RESID

UNIVARIATE PROCEDURE

N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.009306	Variance	0.000087
Skewness	0.965349	Kurtosis	0.231249
USS	0.001992	CSS	0.001992
CV	.	Std Mean	0.0019
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-4	Prob> S	0.9119
Num ^ = 0	24		
W:Normal	0.891486	Prob<W	0.0132

Variable=LOGRESID

N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.205756	Variance	0.042336
Skewness	0.771359	Kurtosis	-0.65472
USS	0.973718	CSS	0.973718
CV	.	Std Mean	0.042
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	5.5	Prob> S	0.8229
Num ^ = 0	18		
W:Normal	0.858454	Prob<W	0.0025

----- METHOD=UNIF -----

Variable=RESID

UNIVARIATE PROCEDURE

N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.01126	Variance	0.000127
Skewness	0.04844	Kurtosis	-0.57005
USS	0.002916	CSS	0.002916
CV	.	Std Mean	0.002299
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	1	Prob> S	0.9779
Num ^ = 0	24		
W:Normal	0.986046	Prob<W	0.9699

Variable=LOGRESID

N	20	Sum Wgts	20
Mean	0	Sum	0
Std Dev	0.244163	Variance	0.059615
Skewness	-0.61263	Kurtosis	-0.60698
USS	1.132691	CSS	1.132691
CV	.	Std Mean	0.054596
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	6	Prob> S	0.8408
Num ^ = 0	20		
W:Normal	0.930838	Prob<W	0.1693

----- METHOD=ZERO -----

Variable=RESID

UNIVARIATE PROCEDURE

N	24	Sum Wgts	24
Mean	0	Sum	0
Std Dev	0.013371	Variance	0.000179
Skewness	0.803258	Kurtosis	-0.49144
USS	0.004112	CSS	0.004112
CV	.	Std Mean	0.002729
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	5.5	Prob> S	0.8229
Num ^ = 0	18		
W:Normal	0.869991	Prob<W	0.0045

Variable=LOGRESID

N	7	Sum Wgts	7
Mean	0	Sum	0
Std Dev	0.045358	Variance	0.002057
Skewness	0.393417	Kurtosis	-1.53458
USS	0.012344	CSS	0.012344
CV	.	Std Mean	0.017144

T:Mean=0            0   Prob>|T|            1.0000  
 Sgn Rank            1   Prob>|S|            0.9375  
 Num ^ = 0            7  
 W:Normal    0.928654   Prob<W            0.5580

MERCURY BIOACCUMULATION IN NEREIS VIRENS  
 LEVENE'S TEST

----- METHOD=DL -----  
 General Linear Models Procedure  
 Dependent Variable: DEV

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.00026815	0.00008938	15.63	0.0001
Error	20	0.00011435	0.00000572		
Corrected Total	23	0.00038250			

----- METHOD=DL/2 -----  
 General Linear Models Procedure  
 Dependent Variable: DEV

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.00055379	0.00018460	16.11	0.0001
Error	20	0.00022915	0.00001146		
Corrected Total	23	0.00078294			

----- METHOD=UNIF -----  
 General Linear Models Procedure  
 Dependent Variable: DEV

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.00018057	0.00006019	1.69	0.2015
Error	20	0.00071293	0.00003565		
Corrected Total	23	0.00089350			

----- METHOD=ZERO -----  
 General Linear Models Procedure  
 Dependent Variable: DEV

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.00103268	0.00034423	17.30	0.0001
Error	20	0.00039804	0.00001990		
Corrected Total	23	0.00143072			

MERCURY BIOACCUMULATION IN NEREIS VIRENS  
 SAMPLE VARIANCES

METHOD	SEDIMENT	N	SAMPLE MEANS	SAMPLE VARIANCES
DL	AK	6	0.020000	.00000000
DL	SH	6	0.022500	.00001270
DL	GOW	6	0.022167	.00001377
DL	RH	6	0.026833	.00011457
DL/2	AK	6	0.010000	.00000000
DL/2	SH	6	0.017500	.00007270
DL/2	GOW	6	0.015500	.00007510
DL/2	RH	6	0.020167	.00025057
UNIF	AK	6	0.010000	.00005600
UNIF	SH	6	0.017500	.00011270
UNIF	GOW	6	0.015500	.00011954
UNIF	RH	6	0.020167	.00029501

ZERO	AK	6	0.000000	.00000000
ZERO	GOW	6	0.008833	.00018977
ZERO	SH	6	0.012500	.00019270
ZERO	RH	6	0.013500	.00043990

## PRELIM3A.SAS program statements

PRELIM3A.SAS should be used when there are two or more blocks of data within each treatment. This program includes DL, DL/2, UNIF, and ZERO. Example data used in the program are zinc water quality data (Table 18 in Chapter 8).

```

OPTIONS LINESIZE=79 PAGESIZE=500 NODATE NONUMBER;
LIBNAME Q 'C:\SAS';

/* Input the data here or read in an existing permanent SAS data set. We
recommend that nondetects be coded as -dl where dl is the numeric
detection limit. To avoid changing statements after the first data step
each time the program is run, name or rename the contaminant concentration
variable CONC and the treatment variable TRT. */

DATA A; SET Q.CDF;
IF CONC<0 THEN DL=ABS(CONC);
IF CONC<0 THEN COUNT=1; ELSE COUNT=0;
ABSCONC=ABS(CONC);
TITLE 'WATER QUALITY DATA FOR ZINC';
PROC SORT; BY LOCATION TRT DATE ABSCONC;

/* Apply DL/2 */

DATA DL2;
SET A;
IF CONC<0 THEN CONC=ABS(CONC)/2;
METHOD='DL/2';
PROC PRINT;

/* Apply DL */

DATA DL;
SET A;
IF CONC<0 THEN CONC=ABS(CONC);
METHOD='DL';
PROC PRINT;

/* Apply ZERO */

DATA ZERO;
SET A;
IF CONC<0 THEN CONC=0;
METHOD='ZERO';
PROC PRINT;

/* Apply UNIF. Here, DATE is the blocking variable within each
treatment. */

PROC MEANS NOPRINT DATA=A;
BY LOCATION TRT DATE;
VAR COUNT;
OUTPUT OUT=B0 SUM=NC N=NREP;
DATA B; SET B0;
NUC=NREP-NC;
DROP _TYPE_ _FREQ_;
DATA UNIF;
MERGE A B; BY LOCATION TRT DATE;
IF FIRST.DATE THEN I=1;
IF CONC<0 THEN DO;
CONC=DL*(I-1)/(NC-1);
IF NC=1 THEN CONC=DL/2;
I+1;
END;
METHOD='UNIF';
PROC PRINT;

/* Determine percent of data that are censored */

PROC MEANS NOPRINT DATA=A;
VAR COUNT;
OUTPUT OUT=O SUM=SUM N=N;
DATA O1; SET O;
PROPCENS=SUM*100/N;
PROC PRINT LABEL NOOBS;
VAR PROPCENS N;
LABEL PROPCENS='PERCENT OF DATA THAT ARE CENSORED'
N='TOTAL NUMBER OF REPLICATES';

```

```

/* Combine data sets and sort by method. Calculate logs. */
DATA ALL;
  SET DL2 DL UNIF ZERO;
  LOGCONC=LOG10(CONC);
  LABEL TRT='SEDIMENT';
PROC SORT; BY METHOD TRT;

/* Determine CV of combined samples */
PROC MEANS NOPRINT;
  VAR CONC; BY METHOD;
  OUTPUT OUT=O CV=CV;
DATA O1; SET O;
  CV=CV/100;
PROC PRINT LABEL NOOBS;
  VAR METHOD CV;
  LABEL CV='CV OF COMBINED SAMPLES';

/* Test normality of combined sample residuals using Shapiro-Wilk's Test.
   DATE is included in the model as a blocking variable. */
PROC GLM NOPRINT DATA=ALL;
  BY METHOD;
  CLASS TRT DATE;
  MODEL CONC LOGCONC=TRT DATE;
  OUTPUT OUT=Z R=RESID LOGRESID;
PROC UNIVARIATE NORMAL;
  BY METHOD;
  VAR RESID LOGRESID;
  TITLE2 'SHAPIRO-WILK'S TEST';

/* Test equality of variances using Levene's Test. DATE is included in the
   model as a blocking variable. */
DATA Z1; SET Z;
  DEV=ABS(RESID);
PROC GLM;
  BY METHOD;
  CLASS TRT DATE;
  MODEL DEV=TRT DATE;
  TITLE2 'LEVENE'S TEST';

/* Calculate sample variances */
PROC MEANS NOPRINT DATA=ALL;
  BY METHOD TRT;
  VAR CONC;
  OUTPUT OUT=O VAR=VARI MEAN=MEAN N=N;
PROC SORT; BY METHOD VARI;
PROC PRINT LABEL NOOBS;
  VAR METHOD TRT N MEAN VARI;
  LABEL VARI='SAMPLE VARIANCES'
         MEAN='SAMPLE MEANS';
  TITLE2 'SAMPLE VARIANCES';

```

## PRELIM3A.SAS program output

WATER QUALITY DATA FOR ZINC								
OBS	DATE	LOCATION	CONC	TRT	DL	COUNT	ABSCONC	METHOD
1	BEFORE	ADJACENT	0.005	SAMPLE5	0.01	1	0.010	DL/2
2	BEFORE	ADJACENT	0.005	SAMPLE5	0.01	1	0.010	DL/2
3	BEFORE	ADJACENT	0.005	SAMPLE5	0.01	1	0.010	DL/2
4	BEFORE	ADJACENT	0.005	SAMPLE5	0.01	1	0.010	DL/2
5	DURING	ADJACENT	0.005	SAMPLE5	0.01	1	0.010	DL/2
6	DURING	ADJACENT	0.005	SAMPLE5	0.01	1	0.010	DL/2
7	DURING	ADJACENT	0.005	SAMPLE5	0.01	1	0.010	DL/2
8	DURING	ADJACENT	0.005	SAMPLE5	0.01	1	0.010	DL/2
9	DURING	ADJACENT	0.005	SAMPLE5	0.01	1	0.010	DL/2
10	DURING	ADJACENT	0.050	SAMPLE5	.	0	0.050	DL/2
11	BEFORE	ADJACENT	0.005	SAMPLE6	0.01	1	0.010	DL/2
12	BEFORE	ADJACENT	0.005	SAMPLE6	0.01	1	0.010	DL/2
13	BEFORE	ADJACENT	0.005	SAMPLE6	0.01	1	0.010	DL/2
14	BEFORE	ADJACENT	0.012	SAMPLE6	.	0	0.012	DL/2
15	DURING	ADJACENT	0.005	SAMPLE6	0.01	1	0.010	DL/2
16	DURING	ADJACENT	0.005	SAMPLE6	0.01	1	0.010	DL/2
17	DURING	ADJACENT	0.005	SAMPLE6	0.01	1	0.010	DL/2
18	DURING	ADJACENT	0.005	SAMPLE6	0.01	1	0.010	DL/2
19	DURING	ADJACENT	0.005	SAMPLE6	0.01	1	0.010	DL/2
20	DURING	ADJACENT	3.200	SAMPLE6	.	0	3.200	DL/2
21	BEFORE	ADJACENT	0.005	SAMPLE7	0.01	1	0.010	DL/2
22	BEFORE	ADJACENT	0.005	SAMPLE7	0.01	1	0.010	DL/2
23	BEFORE	ADJACENT	0.005	SAMPLE7	0.01	1	0.010	DL/2
24	BEFORE	ADJACENT	0.016	SAMPLE7	.	0	0.016	DL/2
25	DURING	ADJACENT	0.005	SAMPLE7	0.01	1	0.010	DL/2
26	DURING	ADJACENT	0.005	SAMPLE7	0.01	1	0.010	DL/2
27	DURING	ADJACENT	0.005	SAMPLE7	0.01	1	0.010	DL/2

28	DURING	ADJACENT	0.005	SAMPLE7	0.01	1	0.010	DL/2
29	DURING	ADJACENT	0.013	SAMPLE7	.	0	0.013	DL/2
30	DURING	ADJACENT	0.360	SAMPLE7	.	0	0.360	DL/2
31	BEFORE	DIKE	0.005	WELL4	0.01	1	0.010	DL/2
32	BEFORE	DIKE	0.005	WELL4	0.01	1	0.010	DL/2
33	BEFORE	DIKE	0.005	WELL4	0.01	1	0.010	DL/2
34	BEFORE	DIKE	0.005	WELL4	0.01	1	0.010	DL/2
35	DURING	DIKE	0.005	WELL4	0.01	1	0.010	DL/2
36	DURING	DIKE	0.005	WELL4	0.01	1	0.010	DL/2
37	DURING	DIKE	0.005	WELL4	0.01	1	0.010	DL/2
38	DURING	DIKE	0.005	WELL4	0.01	1	0.010	DL/2
39	DURING	DIKE	0.005	WELL4	0.01	1	0.010	DL/2
40	DURING	DIKE	0.005	WELL4	0.01	1	0.010	DL/2
41	BEFORE	DIKE	0.005	WELL7	0.01	1	0.010	DL/2
42	BEFORE	DIKE	0.005	WELL7	0.01	1	0.010	DL/2
43	BEFORE	DIKE	0.020	WELL7	.	0	0.020	DL/2
44	BEFORE	DIKE	0.027	WELL7	.	0	0.027	DL/2
45	DURING	DIKE	0.005	WELL7	0.01	1	0.010	DL/2
46	DURING	DIKE	0.005	WELL7	0.01	1	0.010	DL/2
47	DURING	DIKE	0.005	WELL7	0.01	1	0.010	DL/2
48	DURING	DIKE	0.021	WELL7	.	0	0.021	DL/2
49	DURING	DIKE	0.030	WELL7	.	0	0.030	DL/2
50	DURING	DIKE	0.140	WELL7	.	0	0.140	DL/2
51	BEFORE	DIKE	0.005	WELL9	0.01	1	0.010	DL/2
52	BEFORE	DIKE	0.005	WELL9	0.01	1	0.010	DL/2
53	BEFORE	DIKE	0.005	WELL9	0.01	1	0.010	DL/2
54	BEFORE	DIKE	0.005	WELL9	0.01	1	0.010	DL/2
55	DURING	DIKE	0.005	WELL9	0.01	1	0.010	DL/2
56	DURING	DIKE	0.005	WELL9	0.01	1	0.010	DL/2
57	DURING	DIKE	0.005	WELL9	0.01	1	0.010	DL/2
58	DURING	DIKE	0.005	WELL9	0.01	1	0.010	DL/2
59	DURING	DIKE	0.005	WELL9	0.01	1	0.010	DL/2
60	DURING	DIKE	0.014	WELL9	.	0	0.014	DL/2
61	BEFORE	HARBOR	0.005	SAMPLE8A	0.01	1	0.010	DL/2
62	BEFORE	HARBOR	0.005	SAMPLE8A	0.01	1	0.010	DL/2
63	BEFORE	HARBOR	0.005	SAMPLE8A	0.01	1	0.010	DL/2
64	BEFORE	HARBOR	0.012	SAMPLE8A	.	0	0.012	DL/2
65	DURING	HARBOR	0.005	SAMPLE8A	0.01	1	0.010	DL/2
66	DURING	HARBOR	0.005	SAMPLE8A	0.01	1	0.010	DL/2
67	DURING	HARBOR	0.005	SAMPLE8A	0.01	1	0.010	DL/2
68	DURING	HARBOR	0.012	SAMPLE8A	.	0	0.012	DL/2
69	DURING	HARBOR	0.016	SAMPLE8A	.	0	0.016	DL/2
70	DURING	HARBOR	0.064	SAMPLE8A	.	0	0.064	DL/2
71	BEFORE	HARBOR	0.005	SAMPLE8B	0.01	1	0.010	DL/2
72	BEFORE	HARBOR	0.005	SAMPLE8B	0.01	1	0.010	DL/2
73	BEFORE	HARBOR	0.005	SAMPLE8B	0.01	1	0.010	DL/2
74	BEFORE	HARBOR	0.005	SAMPLE8B	0.01	1	0.010	DL/2
75	DURING	HARBOR	0.005	SAMPLE8B	0.01	1	0.010	DL/2
76	DURING	HARBOR	0.005	SAMPLE8B	0.01	1	0.010	DL/2
77	DURING	HARBOR	0.005	SAMPLE8B	0.01	1	0.010	DL/2
78	DURING	HARBOR	0.005	SAMPLE8B	0.01	1	0.010	DL/2
79	DURING	HARBOR	0.011	SAMPLE8B	.	0	0.011	DL/2
80	DURING	HARBOR	0.022	SAMPLE8B	.	0	0.022	DL/2
81	BEFORE	INSIDE	0.005	CDF	0.01	1	0.010	DL/2
82	BEFORE	INSIDE	0.005	CDF	0.01	1	0.010	DL/2
83	BEFORE	INSIDE	0.022	CDF	.	0	0.022	DL/2
84	BEFORE	INSIDE	0.028	CDF	.	0	0.028	DL/2
85	DURING	INSIDE	0.005	CDF	0.01	1	0.010	DL/2
86	DURING	INSIDE	0.005	CDF	0.01	1	0.010	DL/2
87	DURING	INSIDE	0.005	CDF	0.01	1	0.010	DL/2
88	DURING	INSIDE	0.005	CDF	0.01	1	0.010	DL/2
89	DURING	INSIDE	0.005	CDF	0.01	1	0.010	DL/2
90	DURING	INSIDE	0.024	CDF	.	0	0.024	DL/2
OBS	DATE	LOCATION	CONC	TRT	DL	COUNT	ABSCONC	METHOD
1	BEFORE	ADJACENT	0.010	SAMPLE5	0.01	1	0.010	DL
2	BEFORE	ADJACENT	0.010	SAMPLE5	0.01	1	0.010	DL
3	BEFORE	ADJACENT	0.010	SAMPLE5	0.01	1	0.010	DL
4	BEFORE	ADJACENT	0.010	SAMPLE5	0.01	1	0.010	DL
5	DURING	ADJACENT	0.010	SAMPLE5	0.01	1	0.010	DL
6	DURING	ADJACENT	0.010	SAMPLE5	0.01	1	0.010	DL
7	DURING	ADJACENT	0.010	SAMPLE5	0.01	1	0.010	DL
8	DURING	ADJACENT	0.010	SAMPLE5	0.01	1	0.010	DL
9	DURING	ADJACENT	0.010	SAMPLE5	0.01	1	0.010	DL
10	DURING	ADJACENT	0.050	SAMPLE5	.	0	0.050	DL
11	BEFORE	ADJACENT	0.010	SAMPLE6	0.01	1	0.010	DL
12	BEFORE	ADJACENT	0.010	SAMPLE6	0.01	1	0.010	DL
13	BEFORE	ADJACENT	0.010	SAMPLE6	0.01	1	0.010	DL
14	BEFORE	ADJACENT	0.012	SAMPLE6	.	0	0.012	DL
15	DURING	ADJACENT	0.010	SAMPLE6	0.01	1	0.010	DL
16	DURING	ADJACENT	0.010	SAMPLE6	0.01	1	0.010	DL
17	DURING	ADJACENT	0.010	SAMPLE6	0.01	1	0.010	DL
18	DURING	ADJACENT	0.010	SAMPLE6	0.01	1	0.010	DL
19	DURING	ADJACENT	0.010	SAMPLE6	0.01	1	0.010	DL
20	DURING	ADJACENT	3.200	SAMPLE6	.	0	3.200	DL
21	BEFORE	ADJACENT	0.010	SAMPLE7	0.01	1	0.010	DL
22	BEFORE	ADJACENT	0.010	SAMPLE7	0.01	1	0.010	DL
23	BEFORE	ADJACENT	0.010	SAMPLE7	0.01	1	0.010	DL
24	BEFORE	ADJACENT	0.016	SAMPLE7	.	0	0.016	DL
25	DURING	ADJACENT	0.010	SAMPLE7	0.01	1	0.010	DL

26	DURING	ADJACENT	0.010	SAMPLE7	0.01	1	0.010	DL
27	DURING	ADJACENT	0.010	SAMPLE7	0.01	1	0.010	DL
28	DURING	ADJACENT	0.010	SAMPLE7	0.01	1	0.010	DL
29	DURING	ADJACENT	0.013	SAMPLE7	.	0	0.013	DL
30	DURING	ADJACENT	0.360	SAMPLE7	.	0	0.360	DL
31	BEFORE	DIKE	0.010	WELL4	0.01	1	0.010	DL
32	BEFORE	DIKE	0.010	WELL4	0.01	1	0.010	DL
33	BEFORE	DIKE	0.010	WELL4	0.01	1	0.010	DL
34	BEFORE	DIKE	0.010	WELL4	0.01	1	0.010	DL
35	DURING	DIKE	0.010	WELL4	0.01	1	0.010	DL
36	DURING	DIKE	0.010	WELL4	0.01	1	0.010	DL
37	DURING	DIKE	0.010	WELL4	0.01	1	0.010	DL
38	DURING	DIKE	0.010	WELL4	0.01	1	0.010	DL
39	DURING	DIKE	0.010	WELL4	0.01	1	0.010	DL
40	DURING	DIKE	0.010	WELL4	0.01	1	0.010	DL
41	BEFORE	DIKE	0.010	WELL7	0.01	1	0.010	DL
42	BEFORE	DIKE	0.010	WELL7	0.01	1	0.010	DL
43	BEFORE	DIKE	0.020	WELL7	.	0	0.020	DL
44	BEFORE	DIKE	0.027	WELL7	.	0	0.027	DL
45	DURING	DIKE	0.010	WELL7	0.01	1	0.010	DL
46	DURING	DIKE	0.010	WELL7	0.01	1	0.010	DL
47	DURING	DIKE	0.010	WELL7	0.01	1	0.010	DL
48	DURING	DIKE	0.021	WELL7	.	0	0.021	DL
49	DURING	DIKE	0.030	WELL7	.	0	0.030	DL
50	DURING	DIKE	0.140	WELL7	.	0	0.140	DL
51	BEFORE	DIKE	0.010	WELL9	0.01	1	0.010	DL
52	BEFORE	DIKE	0.010	WELL9	0.01	1	0.010	DL
53	BEFORE	DIKE	0.010	WELL9	0.01	1	0.010	DL
54	BEFORE	DIKE	0.010	WELL9	0.01	1	0.010	DL
55	DURING	DIKE	0.010	WELL9	0.01	1	0.010	DL
56	DURING	DIKE	0.010	WELL9	0.01	1	0.010	DL
57	DURING	DIKE	0.010	WELL9	0.01	1	0.010	DL
58	DURING	DIKE	0.010	WELL9	0.01	1	0.010	DL
59	DURING	DIKE	0.010	WELL9	0.01	1	0.010	DL
60	DURING	DIKE	0.014	WELL9	.	0	0.014	DL
61	BEFORE	HARBOR	0.010	SAMPLE8A	0.01	1	0.010	DL
62	BEFORE	HARBOR	0.010	SAMPLE8A	0.01	1	0.010	DL
63	BEFORE	HARBOR	0.010	SAMPLE8A	0.01	1	0.010	DL
64	BEFORE	HARBOR	0.012	SAMPLE8A	.	0	0.012	DL
65	DURING	HARBOR	0.010	SAMPLE8A	0.01	1	0.010	DL
66	DURING	HARBOR	0.010	SAMPLE8A	0.01	1	0.010	DL
67	DURING	HARBOR	0.010	SAMPLE8A	0.01	1	0.010	DL
68	DURING	HARBOR	0.012	SAMPLE8A	.	0	0.012	DL
69	DURING	HARBOR	0.016	SAMPLE8A	.	0	0.016	DL
70	DURING	HARBOR	0.064	SAMPLE8A	.	0	0.064	DL
71	BEFORE	HARBOR	0.010	SAMPLE8B	0.01	1	0.010	DL
72	BEFORE	HARBOR	0.010	SAMPLE8B	0.01	1	0.010	DL
73	BEFORE	HARBOR	0.010	SAMPLE8B	0.01	1	0.010	DL
74	BEFORE	HARBOR	0.010	SAMPLE8B	0.01	1	0.010	DL
75	DURING	HARBOR	0.010	SAMPLE8B	0.01	1	0.010	DL
76	DURING	HARBOR	0.010	SAMPLE8B	0.01	1	0.010	DL
77	DURING	HARBOR	0.010	SAMPLE8B	0.01	1	0.010	DL
78	DURING	HARBOR	0.010	SAMPLE8B	0.01	1	0.010	DL
79	DURING	HARBOR	0.011	SAMPLE8B	.	0	0.011	DL
80	DURING	HARBOR	0.022	SAMPLE8B	.	0	0.022	DL
81	BEFORE	INSIDE	0.010	CDF	0.01	1	0.010	DL
82	BEFORE	INSIDE	0.010	CDF	0.01	1	0.010	DL
83	BEFORE	INSIDE	0.022	CDF	.	0	0.022	DL
84	BEFORE	INSIDE	0.028	CDF	.	0	0.028	DL
85	DURING	INSIDE	0.010	CDF	0.01	1	0.010	DL
86	DURING	INSIDE	0.010	CDF	0.01	1	0.010	DL
87	DURING	INSIDE	0.010	CDF	0.01	1	0.010	DL
88	DURING	INSIDE	0.010	CDF	0.01	1	0.010	DL
89	DURING	INSIDE	0.010	CDF	0.01	1	0.010	DL
90	DURING	INSIDE	0.024	CDF	.	0	0.024	DL

OBS	DATE	LOCATION	CONC	TRT	DL	COUNT	ABSCONC	METHOD
1	BEFORE	ADJACENT	0.000	SAMPLE5	0.01	1	0.010	ZERO
2	BEFORE	ADJACENT	0.000	SAMPLE5	0.01	1	0.010	ZERO
3	BEFORE	ADJACENT	0.000	SAMPLE5	0.01	1	0.010	ZERO
4	BEFORE	ADJACENT	0.000	SAMPLE5	0.01	1	0.010	ZERO
5	DURING	ADJACENT	0.000	SAMPLE5	0.01	1	0.010	ZERO
6	DURING	ADJACENT	0.000	SAMPLE5	0.01	1	0.010	ZERO
7	DURING	ADJACENT	0.000	SAMPLE5	0.01	1	0.010	ZERO
8	DURING	ADJACENT	0.000	SAMPLE5	0.01	1	0.010	ZERO
9	DURING	ADJACENT	0.000	SAMPLE5	0.01	1	0.010	ZERO
10	DURING	ADJACENT	0.050	SAMPLE5	.	0	0.050	ZERO
11	BEFORE	ADJACENT	0.000	SAMPLE6	0.01	1	0.010	ZERO
12	BEFORE	ADJACENT	0.000	SAMPLE6	0.01	1	0.010	ZERO
13	BEFORE	ADJACENT	0.000	SAMPLE6	0.01	1	0.010	ZERO
14	BEFORE	ADJACENT	0.012	SAMPLE6	.	0	0.012	ZERO
15	DURING	ADJACENT	0.000	SAMPLE6	0.01	1	0.010	ZERO
16	DURING	ADJACENT	0.000	SAMPLE6	0.01	1	0.010	ZERO
17	DURING	ADJACENT	0.000	SAMPLE6	0.01	1	0.010	ZERO
18	DURING	ADJACENT	0.000	SAMPLE6	0.01	1	0.010	ZERO
19	DURING	ADJACENT	0.000	SAMPLE6	0.01	1	0.010	ZERO
20	DURING	ADJACENT	3.200	SAMPLE6	.	0	3.200	ZERO
21	BEFORE	ADJACENT	0.000	SAMPLE7	0.01	1	0.010	ZERO
22	BEFORE	ADJACENT	0.000	SAMPLE7	0.01	1	0.010	ZERO
23	BEFORE	ADJACENT	0.000	SAMPLE7	0.01	1	0.010	ZERO

24	BEFORE	ADJACENT	0.016	SAMPLE7	.	0	0.016	ZERO
25	DURING	ADJACENT	0.000	SAMPLE7	0.01	1	0.010	ZERO
26	DURING	ADJACENT	0.000	SAMPLE7	0.01	1	0.010	ZERO
27	DURING	ADJACENT	0.000	SAMPLE7	0.01	1	0.010	ZERO
28	DURING	ADJACENT	0.000	SAMPLE7	0.01	1	0.010	ZERO
29	DURING	ADJACENT	0.013	SAMPLE7	.	0	0.013	ZERO
30	DURING	ADJACENT	0.360	SAMPLE7	.	0	0.360	ZERO
31	BEFORE	DIKE	0.000	WELL4	0.01	1	0.010	ZERO
32	BEFORE	DIKE	0.000	WELL4	0.01	1	0.010	ZERO
33	BEFORE	DIKE	0.000	WELL4	0.01	1	0.010	ZERO
34	BEFORE	DIKE	0.000	WELL4	0.01	1	0.010	ZERO
35	DURING	DIKE	0.000	WELL4	0.01	1	0.010	ZERO
36	DURING	DIKE	0.000	WELL4	0.01	1	0.010	ZERO
37	DURING	DIKE	0.000	WELL4	0.01	1	0.010	ZERO
38	DURING	DIKE	0.000	WELL4	0.01	1	0.010	ZERO
39	DURING	DIKE	0.000	WELL4	0.01	1	0.010	ZERO
40	DURING	DIKE	0.000	WELL4	0.01	1	0.010	ZERO
41	BEFORE	DIKE	0.000	WELL7	0.01	1	0.010	ZERO
42	BEFORE	DIKE	0.000	WELL7	0.01	1	0.010	ZERO
43	BEFORE	DIKE	0.020	WELL7	.	0	0.020	ZERO
44	BEFORE	DIKE	0.027	WELL7	.	0	0.027	ZERO
45	DURING	DIKE	0.000	WELL7	0.01	1	0.010	ZERO
46	DURING	DIKE	0.000	WELL7	0.01	1	0.010	ZERO
47	DURING	DIKE	0.000	WELL7	0.01	1	0.010	ZERO
48	DURING	DIKE	0.021	WELL7	.	0	0.021	ZERO
49	DURING	DIKE	0.030	WELL7	.	0	0.030	ZERO
50	DURING	DIKE	0.140	WELL7	.	0	0.140	ZERO
51	BEFORE	DIKE	0.000	WELL9	0.01	1	0.010	ZERO
52	BEFORE	DIKE	0.000	WELL9	0.01	1	0.010	ZERO
53	BEFORE	DIKE	0.000	WELL9	0.01	1	0.010	ZERO
54	BEFORE	DIKE	0.000	WELL9	0.01	1	0.010	ZERO
55	DURING	DIKE	0.000	WELL9	0.01	1	0.010	ZERO
56	DURING	DIKE	0.000	WELL9	0.01	1	0.010	ZERO
57	DURING	DIKE	0.000	WELL9	0.01	1	0.010	ZERO
58	DURING	DIKE	0.000	WELL9	0.01	1	0.010	ZERO
59	DURING	DIKE	0.000	WELL9	0.01	1	0.010	ZERO
60	DURING	DIKE	0.014	WELL9	.	0	0.014	ZERO
61	BEFORE	HARBOR	0.000	SAMPLE8A	0.01	1	0.010	ZERO
62	BEFORE	HARBOR	0.000	SAMPLE8A	0.01	1	0.010	ZERO
63	BEFORE	HARBOR	0.000	SAMPLE8A	0.01	1	0.010	ZERO
64	BEFORE	HARBOR	0.012	SAMPLE8A	.	0	0.012	ZERO
65	DURING	HARBOR	0.000	SAMPLE8A	0.01	1	0.010	ZERO
66	DURING	HARBOR	0.000	SAMPLE8A	0.01	1	0.010	ZERO
67	DURING	HARBOR	0.000	SAMPLE8A	0.01	1	0.010	ZERO
68	DURING	HARBOR	0.012	SAMPLE8A	.	0	0.012	ZERO
69	DURING	HARBOR	0.016	SAMPLE8A	.	0	0.016	ZERO
70	DURING	HARBOR	0.064	SAMPLE8A	.	0	0.064	ZERO
71	BEFORE	HARBOR	0.000	SAMPLE8B	0.01	1	0.010	ZERO
72	BEFORE	HARBOR	0.000	SAMPLE8B	0.01	1	0.010	ZERO
73	BEFORE	HARBOR	0.000	SAMPLE8B	0.01	1	0.010	ZERO
74	BEFORE	HARBOR	0.000	SAMPLE8B	0.01	1	0.010	ZERO
75	DURING	HARBOR	0.000	SAMPLE8B	0.01	1	0.010	ZERO
76	DURING	HARBOR	0.000	SAMPLE8B	0.01	1	0.010	ZERO
77	DURING	HARBOR	0.000	SAMPLE8B	0.01	1	0.010	ZERO
78	DURING	HARBOR	0.000	SAMPLE8B	0.01	1	0.010	ZERO
79	DURING	HARBOR	0.011	SAMPLE8B	.	0	0.011	ZERO
80	DURING	HARBOR	0.022	SAMPLE8B	.	0	0.022	ZERO
81	BEFORE	INSIDE	0.000	CDF	0.01	1	0.010	ZERO
82	BEFORE	INSIDE	0.000	CDF	0.01	1	0.010	ZERO
83	BEFORE	INSIDE	0.022	CDF	.	0	0.022	ZERO
84	BEFORE	INSIDE	0.028	CDF	.	0	0.028	ZERO
85	DURING	INSIDE	0.000	CDF	0.01	1	0.010	ZERO
86	DURING	INSIDE	0.000	CDF	0.01	1	0.010	ZERO
87	DURING	INSIDE	0.000	CDF	0.01	1	0.010	ZERO
88	DURING	INSIDE	0.000	CDF	0.01	1	0.010	ZERO
89	DURING	INSIDE	0.000	CDF	0.01	1	0.010	ZERO
90	DURING	INSIDE	0.024	CDF	.	0	0.024	ZERO

OBS	DATE	LOCATION	CONC	TRT	DL	COUNT	ABSCONC	NC	NREP	NUC	I	METHOD
1	BEFORE	ADJACENT	0.00000	SAMPLE5	0.01	1	0.010	4	4	0	2	UNIF
2	BEFORE	ADJACENT	0.00333	SAMPLE5	0.01	1	0.010	4	4	0	3	UNIF
3	BEFORE	ADJACENT	0.00667	SAMPLE5	0.01	1	0.010	4	4	0	4	UNIF
4	BEFORE	ADJACENT	0.01000	SAMPLE5	0.01	1	0.010	4	4	0	5	UNIF
5	DURING	ADJACENT	0.00000	SAMPLE5	0.01	1	0.010	5	6	1	2	UNIF
6	DURING	ADJACENT	0.00250	SAMPLE5	0.01	1	0.010	5	6	1	3	UNIF
7	DURING	ADJACENT	0.00500	SAMPLE5	0.01	1	0.010	5	6	1	4	UNIF
8	DURING	ADJACENT	0.00750	SAMPLE5	0.01	1	0.010	5	6	1	5	UNIF
9	DURING	ADJACENT	0.01000	SAMPLE5	0.01	1	0.010	5	6	1	6	UNIF
10	DURING	ADJACENT	0.05000	SAMPLE5	.	0	0.050	5	6	1	6	UNIF
11	BEFORE	ADJACENT	0.00000	SAMPLE6	0.01	1	0.010	3	4	1	2	UNIF
12	BEFORE	ADJACENT	0.00500	SAMPLE6	0.01	1	0.010	3	4	1	3	UNIF
13	BEFORE	ADJACENT	0.01000	SAMPLE6	0.01	1	0.010	3	4	1	4	UNIF
14	BEFORE	ADJACENT	0.01200	SAMPLE6	.	0	0.012	3	4	1	4	UNIF
15	DURING	ADJACENT	0.00000	SAMPLE6	0.01	1	0.010	5	6	1	2	UNIF
16	DURING	ADJACENT	0.00250	SAMPLE6	0.01	1	0.010	5	6	1	3	UNIF
17	DURING	ADJACENT	0.00500	SAMPLE6	0.01	1	0.010	5	6	1	4	UNIF
18	DURING	ADJACENT	0.00750	SAMPLE6	0.01	1	0.010	5	6	1	5	UNIF
19	DURING	ADJACENT	0.01000	SAMPLE6	0.01	1	0.010	5	6	1	6	UNIF
20	DURING	ADJACENT	3.20000	SAMPLE6	.	0	3.200	5	6	1	6	UNIF
21	BEFORE	ADJACENT	0.00000	SAMPLE7	0.01	1	0.010	3	4	1	2	UNIF

22	BEFORE	ADJACENT	0.00500	SAMPLE7	0.01	1	0.010	3	4	1	3	UNIF
23	BEFORE	ADJACENT	0.01000	SAMPLE7	0.01	1	0.010	3	4	1	4	UNIF
24	BEFORE	ADJACENT	0.01600	SAMPLE7	.	0	0.016	3	4	1	4	UNIF
25	DURING	ADJACENT	0.00000	SAMPLE7	0.01	1	0.010	4	6	2	2	UNIF
26	DURING	ADJACENT	0.00333	SAMPLE7	0.01	1	0.010	4	6	2	3	UNIF
27	DURING	ADJACENT	0.00667	SAMPLE7	0.01	1	0.010	4	6	2	4	UNIF
28	DURING	ADJACENT	0.01000	SAMPLE7	0.01	1	0.010	4	6	2	5	UNIF
29	DURING	ADJACENT	0.01300	SAMPLE7	.	0	0.013	4	6	2	5	UNIF
30	DURING	ADJACENT	0.36000	SAMPLE7	.	0	0.360	4	6	2	5	UNIF
31	BEFORE	DIKE	0.00000	WELL4	0.01	1	0.010	4	4	0	2	UNIF
32	BEFORE	DIKE	0.00333	WELL4	0.01	1	0.010	4	4	0	3	UNIF
33	BEFORE	DIKE	0.00667	WELL4	0.01	1	0.010	4	4	0	4	UNIF
34	BEFORE	DIKE	0.01000	WELL4	0.01	1	0.010	4	4	0	5	UNIF
35	DURING	DIKE	0.00000	WELL4	0.01	1	0.010	6	6	0	2	UNIF
36	DURING	DIKE	0.00200	WELL4	0.01	1	0.010	6	6	0	3	UNIF
37	DURING	DIKE	0.00400	WELL4	0.01	1	0.010	6	6	0	4	UNIF
38	DURING	DIKE	0.00600	WELL4	0.01	1	0.010	6	6	0	5	UNIF
39	DURING	DIKE	0.00800	WELL4	0.01	1	0.010	6	6	0	6	UNIF
40	DURING	DIKE	0.01000	WELL4	0.01	1	0.010	6	6	0	7	UNIF
41	BEFORE	DIKE	0.00000	WELL7	0.01	1	0.010	2	4	2	2	UNIF
42	BEFORE	DIKE	0.01000	WELL7	0.01	1	0.010	2	4	2	3	UNIF
43	BEFORE	DIKE	0.02000	WELL7	.	0	0.020	2	4	2	3	UNIF
44	BEFORE	DIKE	0.02700	WELL7	.	0	0.027	2	4	2	3	UNIF
45	DURING	DIKE	0.00000	WELL7	0.01	1	0.010	3	6	3	2	UNIF
46	DURING	DIKE	0.00500	WELL7	0.01	1	0.010	3	6	3	3	UNIF
47	DURING	DIKE	0.01000	WELL7	0.01	1	0.010	3	6	3	4	UNIF
48	DURING	DIKE	0.02100	WELL7	.	0	0.021	3	6	3	4	UNIF
49	DURING	DIKE	0.03000	WELL7	.	0	0.030	3	6	3	4	UNIF
50	DURING	DIKE	0.14000	WELL7	.	0	0.140	3	6	3	4	UNIF
51	BEFORE	DIKE	0.00000	WELL9	0.01	1	0.010	4	4	0	2	UNIF
52	BEFORE	DIKE	0.00333	WELL9	0.01	1	0.010	4	4	0	3	UNIF
53	BEFORE	DIKE	0.00667	WELL9	0.01	1	0.010	4	4	0	4	UNIF
54	BEFORE	DIKE	0.01000	WELL9	0.01	1	0.010	4	4	0	5	UNIF
55	DURING	DIKE	0.00000	WELL9	0.01	1	0.010	5	6	1	2	UNIF
56	DURING	DIKE	0.00250	WELL9	0.01	1	0.010	5	6	1	3	UNIF
57	DURING	DIKE	0.00500	WELL9	0.01	1	0.010	5	6	1	4	UNIF
58	DURING	DIKE	0.00750	WELL9	0.01	1	0.010	5	6	1	5	UNIF
59	DURING	DIKE	0.01000	WELL9	0.01	1	0.010	5	6	1	6	UNIF
60	DURING	DIKE	0.01400	WELL9	.	0	0.014	5	6	1	6	UNIF
61	BEFORE	HARBOR	0.00000	SAMPLE8A	0.01	1	0.010	3	4	1	2	UNIF
62	BEFORE	HARBOR	0.00500	SAMPLE8A	0.01	1	0.010	3	4	1	3	UNIF
63	BEFORE	HARBOR	0.01000	SAMPLE8A	0.01	1	0.010	3	4	1	4	UNIF
64	BEFORE	HARBOR	0.01200	SAMPLE8A	.	0	0.012	3	4	1	4	UNIF
65	DURING	HARBOR	0.00000	SAMPLE8A	0.01	1	0.010	3	6	3	2	UNIF
66	DURING	HARBOR	0.00500	SAMPLE8A	0.01	1	0.010	3	6	3	3	UNIF
67	DURING	HARBOR	0.01000	SAMPLE8A	0.01	1	0.010	3	6	3	4	UNIF
68	DURING	HARBOR	0.01200	SAMPLE8A	.	0	0.012	3	6	3	4	UNIF
69	DURING	HARBOR	0.01600	SAMPLE8A	.	0	0.016	3	6	3	4	UNIF
70	DURING	HARBOR	0.06400	SAMPLE8A	.	0	0.064	3	6	3	4	UNIF
71	BEFORE	HARBOR	0.00000	SAMPLE8B	0.01	1	0.010	4	4	0	2	UNIF
72	BEFORE	HARBOR	0.00333	SAMPLE8B	0.01	1	0.010	4	4	0	3	UNIF
73	BEFORE	HARBOR	0.00667	SAMPLE8B	0.01	1	0.010	4	4	0	4	UNIF
74	BEFORE	HARBOR	0.01000	SAMPLE8B	0.01	1	0.010	4	4	0	5	UNIF
75	DURING	HARBOR	0.00000	SAMPLE8B	0.01	1	0.010	4	6	2	2	UNIF
76	DURING	HARBOR	0.00333	SAMPLE8B	0.01	1	0.010	4	6	2	3	UNIF
77	DURING	HARBOR	0.00667	SAMPLE8B	0.01	1	0.010	4	6	2	4	UNIF
78	DURING	HARBOR	0.01000	SAMPLE8B	0.01	1	0.010	4	6	2	5	UNIF
79	DURING	HARBOR	0.01100	SAMPLE8B	.	0	0.011	4	6	2	5	UNIF
80	DURING	HARBOR	0.02200	SAMPLE8B	.	0	0.022	4	6	2	5	UNIF
81	BEFORE	INSIDE	0.00000	CDF	0.01	1	0.010	2	4	2	2	UNIF
82	BEFORE	INSIDE	0.01000	CDF	0.01	1	0.010	2	4	2	3	UNIF
83	BEFORE	INSIDE	0.02200	CDF	.	0	0.022	2	4	2	3	UNIF
84	BEFORE	INSIDE	0.02800	CDF	.	0	0.028	2	4	2	3	UNIF
85	DURING	INSIDE	0.00000	CDF	0.01	1	0.010	5	6	1	2	UNIF
86	DURING	INSIDE	0.00250	CDF	0.01	1	0.010	5	6	1	3	UNIF
87	DURING	INSIDE	0.00500	CDF	0.01	1	0.010	5	6	1	4	UNIF
88	DURING	INSIDE	0.00750	CDF	0.01	1	0.010	5	6	1	5	UNIF
89	DURING	INSIDE	0.01000	CDF	0.01	1	0.010	5	6	1	6	UNIF
90	DURING	INSIDE	0.02400	CDF	.	0	0.024	5	6	1	6	UNIF

WATER QUALITY DATA FOR ZINC

PERCENT OF DATA THAT ARE CENSORED	TOTAL NUMBER OF REPLICATES
76.6667	90

METHOD	CV OF COMBINED SAMPLES
DL	6.32748
DL/2	6.82723
UNIF	6.82756
ZERO	7.41106

WATER QUALITY DATA FOR ZINC  
SHAPIRO-WILK'S TEST

----- METHOD=DL -----

Variable=RESID		UNIVARIATE PROCEDURE	
N	90	Sum Wgts	90
Mean	0	Sum	0
Std Dev	0.321204	Variance	0.103172
Skewness	7.875328	Kurtosis	70.79399
USS	9.182335	CSS	9.182335
CV	.	Std Mean	0.033858
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-258.5	Prob> S	0.3008
Num ^ = 0	90		
W:Normal	0.327697	Prob<W	0.0

Variable=LOGRESID		UNIVARIATE PROCEDURE	
N	90	Sum Wgts	90
Mean	0	Sum	0
Std Dev	0.337437	Variance	0.113864
Skewness	4.132627	Kurtosis	22.40555
USS	10.13386	CSS	10.13386
CV	.	Std Mean	0.035569
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-743.5	Prob> S	0.0023
Num ^ = 0	90		
W:Normal	0.625065	Prob<W	0.0

----- METHOD=DL/2 -----

Variable=RESID		UNIVARIATE PROCEDURE	
N	90	Sum Wgts	90
Mean	0	Sum	0
Std Dev	0.321721	Variance	0.103504
Skewness	7.869586	Kurtosis	70.72175
USS	9.211898	CSS	9.211898
CV	.	Std Mean	0.033912
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-278.5	Prob> S	0.2647
Num ^ = 0	90		
W:Normal	0.32941	Prob<W	0.0

Variable=LOGRESID		UNIVARIATE PROCEDURE	
N	90	Sum Wgts	90
Mean	0	Sum	0
Std Dev	0.42012	Variance	0.176501
Skewness	3.070173	Kurtosis	13.58329
USS	15.70856	CSS	15.70856
CV	.	Std Mean	0.044285
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-545.5	Prob> S	0.0273
Num ^ = 0	90		
W:Normal	0.744427	Prob<W	0.0

----- METHOD=UNIF -----

Variable=RESID		UNIVARIATE PROCEDURE	
N	90	Sum Wgts	90
Mean	0	Sum	0
Std Dev	0.321739	Variance	0.103516
Skewness	7.868259	Kurtosis	70.70575
USS	9.212901	CSS	9.212901
CV	.	Std Mean	0.033914
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-333.5	Prob> S	0.1811
Num ^ = 0	90		
W:Normal	0.330868	Prob<W	0.0

Variable=LOGRESID		UNIVARIATE PROCEDURE	
N	72	Sum Wgts	72
Mean	0	Sum	0
Std Dev	0.456752	Variance	0.208623
Skewness	2.23372	Kurtosis	9.231596
USS	14.81222	CSS	14.81222
CV	.	Std Mean	0.053829
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-188	Prob> S	0.2946
Num ^ = 0	72		
W:Normal	0.850556	Prob<W	0.0001

----- METHOD=ZERO -----

UNIVARIATE PROCEDURE

Variable=RESID

N	90	Sum Wgts	90
Mean	0	Sum	0
Std Dev	0.322249	Variance	0.103845
Skewness	7.863021	Kurtosis	70.6394
USS	9.242161	CSS	9.242161
CV	.	Std Mean	0.033968
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-307.5	Prob> S	0.2178
Num ^ = 0	90		
W:Normal	0.331324	Prob<W	0.0

Variable=LOGRESID

N	21	Sum Wgts	21
Mean	0	Sum	0
Std Dev	0.441495	Variance	0.194918
Skewness	-0.01389	Kurtosis	0.312789
USS	3.898358	CSS	3.898358
CV	.	Std Mean	0.096342
T:Mean=0	0	Prob> T	1.0000
Sgn Rank	-3.5	Prob> S	0.9066
Num ^ = 0	21		
W:Normal	0.983466	Prob<W	0.9542

WATER QUALITY DATA FOR ZINC  
LEVENE'S TEST

----- METHOD=DL -----

General Linear Models Procedure

Dependent Variable: DEV

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	8	2.51373062	0.31421633	4.37	0.0002
DATE	1	0.07380001	0.07380001	1.03	0.3141

----- METHOD=DL/2 -----

General Linear Models Procedure

Dependent Variable: DEV

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	8	2.51872049	0.31484006	4.36	0.0002
DATE	1	0.07563674	0.07563674	1.05	0.3089

----- METHOD=UNIF -----

General Linear Models Procedure

Dependent Variable: DEV

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	8	2.51862863	0.31482858	4.36	0.0002
DATE	1	0.07553894	0.07553894	1.05	0.3093

----- METHOD=ZERO -----

General Linear Models Procedure

Dependent Variable: DEV

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	8	2.52375451	0.31546931	4.36	0.0002
DATE	1	0.07749605	0.07749605	1.07	0.3038

WATER QUALITY DATA FOR ZINC  
SAMPLE VARIANCES

METHOD	SEDIMENT	N	SAMPLE MEANS	SAMPLE VARIANCES
DL	WELL4	10	0.0100	0.00000
DL	WELL9	10	0.0104	0.00000
DL	SAMPLE8B	10	0.0113	0.00001
DL	CDF	10	0.0144	0.00005
DL	SAMPLE5	10	0.0140	0.00016
DL	SAMPLE8A	10	0.0164	0.00028

DL	WELL7	10	0.0288	0.00159
DL	SAMPLE7	10	0.0459	0.01218
DL	SAMPLE6	10	0.3292	1.01747
DL/2	WELL4	10	0.0050	0.00000
DL/2	WELL9	10	0.0059	0.00001
DL/2	SAMPLE8B	10	0.0073	0.00003
DL/2	CDF	10	0.0109	0.00009
DL/2	SAMPLE5	10	0.0095	0.00020
DL/2	SAMPLE8A	10	0.0134	0.00033
DL/2	WELL7	10	0.0263	0.00170
DL/2	SAMPLE7	10	0.0424	0.01247
DL/2	SAMPLE6	10	0.3252	1.02031
UNIF	WELL4	10	0.0050	0.00001
UNIF	WELL9	10	0.0059	0.00002
UNIF	SAMPLE8B	10	0.0073	0.00004
UNIF	CDF	10	0.0109	0.00010
UNIF	SAMPLE5	10	0.0095	0.00022
UNIF	SAMPLE8A	10	0.0134	0.00034
UNIF	WELL7	10	0.0263	0.00171
UNIF	SAMPLE7	10	0.0424	0.01248
UNIF	SAMPLE6	10	0.3252	1.02032
ZERO	WELL4	10	0.0000	0.00000
ZERO	WELL9	10	0.0014	0.00002
ZERO	SAMPLE8B	10	0.0033	0.00006
ZERO	CDF	10	0.0074	0.00014
ZERO	SAMPLE5	10	0.0050	0.00025
ZERO	SAMPLE8A	10	0.0104	0.00040
ZERO	WELL7	10	0.0238	0.00182
ZERO	SAMPLE7	10	0.0389	0.01277
ZERO	SAMPLE6	10	0.3212	1.02316

## ESTIMATE.SAS program statements

This program provides SAS statements for a preliminary analysis of censored data to determine the most appropriate censored data methods for estimation of the mean and standard deviation. ESTIMATE.SAS may be used when a treatment is no more than 80 percent censored. This program includes all six censored data methods recommended for estimation when sample size is small: DL, DL/2, MLE LOGN, MLE WEIB, UNIFR, and ZERO. Example data used in the program are for total PCB bioaccumulation in *Macoma nasuta* and *Nereis virens* exposed to four sediments (Chapter 8). For each species, treatment, and method, the program calculates CV, mean, variance, standard deviation,  $n$ , and the table  $t$  value. The appropriate mean, variance,  $n$ , and  $t$  value may then be used in determining the one-sided, 95-percent upper confidence limit for comparison with an action level.

```

OPTIONS LINESIZE=79 PAGESIZE=500 NODATE NONUMBER;

/* Input the data here or read in an existing permanent SAS data set. We
recommend that nondetects be coded as -dl where dl is the numeric
detection limit. To avoid changing statements after the first data step
each time the program is run, name or rename the contaminant concentration
variable CONC and the treatment variable TRT. */
/* The following data are entered as two separate data sets, one for each
organism, simply for ease of entry so the species name does not have
to be repeated for each observation. */

DATA A00;
  INPUT TRT $ CONC @@;
  CARDS;
AK 1.1 AK 3.0 AK 2.5 AK 2.0 AK 1.7 AK 2.6
GOW -.06 GOW 2.0 GOW 2.0 GOW 1.0 GOW 1.8 GOW 2.0
RH 2.0 RH 0.9 RH 1.3 RH 0.9 RH 1.9 RH 1.6
SH -.06 SH .80 SH .80 SH .40 SH -.06 SH .50
;
DATA A0; SET A00;
  ORG='MACOMA';
DATA A11;
  INPUT TRT $ CONC @@;
  CARDS;
AK 7.5 AK 1.9 AK -.06 AK 3.6 AK 1.8
GOW 4.0 GOW 2.3 GOW 2.3 GOW 3.0 GOW -.06
RH 2.0 RH 12.0 RH 1.2 RH -.06 RH 1.9 RH 1.4
SH 0.9 SH 0.8 SH -.06 SH -.06 SH -.06 SH -.06

```

```

;
DATA A1; SET A11;
  ORG='NEREIS';
DATA A; SET A0 A1;
  IF CONC<0 THEN DL=ABS(CONC);
  IF CONC<0 THEN COUNT=1; ELSE COUNT=0;
  ABSCONC=ABS(CONC);
TITLE 'TOTAL PCB BIOACCUMULATION IN MACOMA NASUTA AND NEREIS VIRENS';
PROC SORT; BY ORG TRT ABSCONC;

/* Apply DL/2 */

DATA DL2;
  SET A;
  IF CONC<0 THEN CONC=ABS(CONC)/2;
  METHOD='DL/2';
PROC PRINT;

/* Apply DL */

DATA DL;
  SET A;
  IF CONC<0 THEN CONC=ABS(CONC);
  METHOD='DL';
PROC PRINT;

/* Apply ZERO */

DATA ZERO;
  SET A;
  IF CONC<0 THEN CONC=0;
  METHOD='ZERO';
PROC PRINT;

/* Apply UNIFR */

DATA UNIFR;
  SET A;
  SEED=0;
  IF CONC<0 THEN CONC=DL*RANUNI(SEED);
  METHOD='UNIFR';
PROC PRINT;

/* Apply MLE WEIB. MLE WEIB and MLE LOGN must each be run twice with these
data, once for the 6-replicate treatments and once for the 5-replicate
treatments, as the quantiles (Q=) in the OUTPUT statement depend upon
the number of replicates. */

DATA CA; SET A;
  IF ORG='NEREIS' AND TRT='AK' THEN DELETE;
  IF ORG='NEREIS' AND TRT='GOW' THEN DELETE;
DATA C; SET CA;
  IF CONC<0 THEN LOWER=.; ELSE LOWER=CONC;
  IF CONC<0 THEN CONC=DL;
PROC LIFEREG NOPRINT;
  BY ORG TRT;
MODEL (LOWER,CONC)= /D=WEIBULL SHAPE1=1;
OUTPUT OUT=C1 P=PRED Q=.01 .21 .4 .6 .79 .99;
PROC SORT; BY ORG TRT _PROB_;
DATA D1; SET C1;
  BY ORG TRT _PROB_;
  IF FIRST._PROB_;
  KEEP ORG TRT _PROB_ PRED;
DATA E1;
  MERGE D1 CA; BY ORG TRT;
  IF CONC<DL AND PRED>=DL THEN CONC=DL;
  IF CONC<DL AND PRED<DL THEN CONC=PRED;

DATA CB; SET A;
  IF ORG='MACOMA' THEN DELETE;
  IF ORG='NEREIS' AND TRT='RH' THEN DELETE;
  IF ORG='NEREIS' AND TRT='SH' THEN DELETE;
DATA C; SET CB;
  IF CONC<0 THEN LOWER=.; ELSE LOWER=CONC;
  IF CONC<0 THEN CONC=DL;
PROC LIFEREG NOPRINT;
  BY ORG TRT;
MODEL (LOWER,CONC)= /D=WEIBULL SHAPE1=1;
OUTPUT OUT=C2 P=PRED Q=.01 .25 .5 .75 .99;
PROC SORT; BY ORG TRT _PROB_;
DATA D2; SET C2;
  BY ORG TRT _PROB_;
  IF FIRST._PROB_;
  KEEP ORG TRT _PROB_ PRED;
DATA E2;
  MERGE D2 CB; BY ORG TRT;
  IF CONC<DL AND PRED>=DL THEN CONC=DL;
  IF CONC<DL AND PRED<DL THEN CONC=PRED;
DATA MLEWEIB;
  SET E1 E2;
  METHOD='MLE WEIB';

```

```

PROC SORT; BY ORG TRT;
PROC PRINT;

/* Apply MLE LOGN */
DATA CA; SET A;
  IF ORG='NEREIS' AND TRT='AK' THEN DELETE;
  IF ORG='NEREIS' AND TRT='GOW' THEN DELETE;
DATA C; SET CA;
  IF CONC<0 THEN LOWER=.; ELSE LOWER=CONC;
  IF CONC<0 THEN CONC=DL;
PROC LIFEREG NOPRINT;
  BY ORG TRT;
  MODEL (LOWER,CONC)= /D=LNORMAL;
  OUTPUT OUT=C1 P=PRED Q=.01 .21 .4 .6 .79 .99;
PROC SORT; BY ORG TRT _PROB_;
DATA D1; SET C1;
  BY ORG TRT _PROB_;
  IF FIRST._PROB_;
  KEEP ORG TRT _PROB_ PRED;
DATA E1;
  MERGE D1 CA; BY ORG TRT;
  IF CONC<DL AND PRED>=DL THEN CONC=DL;
  IF CONC<DL AND PRED<DL THEN CONC=PRED;

DATA CB; SET A;
  IF ORG='MACOMA' THEN DELETE;
  IF ORG='NEREIS' AND TRT='RH' THEN DELETE;
  IF ORG='NEREIS' AND TRT='SH' THEN DELETE;
DATA C; SET CB;
  IF CONC<0 THEN LOWER=.; ELSE LOWER=CONC;
  IF CONC<0 THEN CONC=DL;
PROC LIFEREG NOPRINT;
  BY ORG TRT;
  MODEL (LOWER,CONC)= /D=LNORMAL;
  OUTPUT OUT=C2 P=PRED Q=.01 .25 .5 .75 .99;
PROC SORT; BY ORG TRT _PROB_;
DATA D2; SET C2;
  BY ORG TRT _PROB_;
  IF FIRST._PROB_;
  KEEP ORG TRT _PROB_ PRED;
DATA E2;
  MERGE D2 CB; BY ORG TRT;
  IF CONC<DL AND PRED>=DL THEN CONC=DL;
  IF CONC<DL AND PRED<DL THEN CONC=PRED;
DATA MLELOGN;
  SET E1 E2;
  METHOD='MLE LOGN';
PROC SORT; BY ORG TRT;
PROC PRINT;

/* Determine percent of data that are censored */
PROC MEANS NOPRINT DATA=A;
  BY ORG TRT;
  VAR COUNT;
  OUTPUT OUT=O SUM=SUM N=N;
DATA O1; SET O;
  PROPCENS=SUM*100/N;
PROC PRINT LABEL NOOBS;
  VAR ORG TRT PROPCENS N;
  LABEL ORG='SPECIES'
         TRT='SEDIMENT'
         PROPCENS='PERCENT OF DATA THAT ARE CENSORED'
         N='NUMBER OF REPLICATES';

/* Combine data sets and sort */
DATA ALL;
  SET DL2 DL ZERO UNIFR MLEWEIB MLELOGN;
PROC SORT; BY ORG TRT METHOD;

/* Calculate CV, mean, variance, standard deviation, and n for each
species, treatment and method. Calculate table t value */
PROC MEANS NOPRINT;
  VAR CONC; BY ORG TRT METHOD;
  OUTPUT OUT=O CV=CV MEAN=MEAN VAR=VAR STD=STD N=N;
DATA O1; SET O;
  CV=CV/100;
  TALPHA=TINV(.95,N-1);
PROC PRINT LABEL NOOBS;
  VAR ORG TRT METHOD CV N MEAN VAR STD TALPHA;
  LABEL ORG='SPECIES'
         TRT='SEDI- MENT'
         VAR='VARIANCE'
         STD='STD. DEV.'
         TALPHA='T VALUE FOR (1-ALPHA= .95,N-1)';

```

# ESTIMATE.SAS program output

## TOTAL PCB BIOACCUMULATION IN MACOMA NASUTA AND NEREIS VIRENS

OBS	TRT	CONC	ORG	DL	COUNT	ABSCONC	METHOD
1	AK	1.10	MACOMA	.	0	1.10	DL/2
2	AK	1.70	MACOMA	.	0	1.70	DL/2
3	AK	2.00	MACOMA	.	0	2.00	DL/2
4	AK	2.50	MACOMA	.	0	2.50	DL/2
5	AK	2.60	MACOMA	.	0	2.60	DL/2
6	AK	3.00	MACOMA	.	0	3.00	DL/2
7	GOW	0.03	MACOMA	0.06	1	0.06	DL/2
8	GOW	1.00	MACOMA	.	0	1.00	DL/2
9	GOW	1.80	MACOMA	.	0	1.80	DL/2
10	GOW	2.00	MACOMA	.	0	2.00	DL/2
11	GOW	2.00	MACOMA	.	0	2.00	DL/2
12	GOW	2.00	MACOMA	.	0	2.00	DL/2
13	RH	0.90	MACOMA	.	0	0.90	DL/2
14	RH	0.90	MACOMA	.	0	0.90	DL/2
15	RH	1.30	MACOMA	.	0	1.30	DL/2
16	RH	1.60	MACOMA	.	0	1.60	DL/2
17	RH	1.90	MACOMA	.	0	1.90	DL/2
18	RH	2.00	MACOMA	.	0	2.00	DL/2
19	SH	0.03	MACOMA	0.06	1	0.06	DL/2
20	SH	0.03	MACOMA	0.06	1	0.06	DL/2
21	SH	0.40	MACOMA	.	0	0.40	DL/2
22	SH	0.50	MACOMA	.	0	0.50	DL/2
23	SH	0.80	MACOMA	.	0	0.80	DL/2
24	SH	0.80	MACOMA	.	0	0.80	DL/2
25	AK	0.03	NEREIS	0.06	1	0.06	DL/2
26	AK	1.80	NEREIS	.	0	1.80	DL/2
27	AK	1.90	NEREIS	.	0	1.90	DL/2
28	AK	3.60	NEREIS	.	0	3.60	DL/2
29	AK	7.50	NEREIS	.	0	7.50	DL/2
30	GOW	0.03	NEREIS	0.06	1	0.06	DL/2
31	GOW	2.30	NEREIS	.	0	2.30	DL/2
32	GOW	2.30	NEREIS	.	0	2.30	DL/2
33	GOW	3.00	NEREIS	.	0	3.00	DL/2
34	GOW	4.00	NEREIS	.	0	4.00	DL/2
35	RH	0.03	NEREIS	0.06	1	0.06	DL/2
36	RH	1.20	NEREIS	.	0	1.20	DL/2
37	RH	1.40	NEREIS	.	0	1.40	DL/2
38	RH	1.90	NEREIS	.	0	1.90	DL/2
39	RH	2.00	NEREIS	.	0	2.00	DL/2
40	RH	12.00	NEREIS	.	0	12.00	DL/2
41	SH	0.03	NEREIS	0.06	1	0.06	DL/2
42	SH	0.03	NEREIS	0.06	1	0.06	DL/2
43	SH	0.03	NEREIS	0.06	1	0.06	DL/2
44	SH	0.03	NEREIS	0.06	1	0.06	DL/2
45	SH	0.80	NEREIS	.	0	0.80	DL/2
46	SH	0.90	NEREIS	.	0	0.90	DL/2

OBS	TRT	CONC	ORG	DL	COUNT	ABSCONC	METHOD
1	AK	1.10	MACOMA	.	0	1.10	DL
2	AK	1.70	MACOMA	.	0	1.70	DL
3	AK	2.00	MACOMA	.	0	2.00	DL
4	AK	2.50	MACOMA	.	0	2.50	DL
5	AK	2.60	MACOMA	.	0	2.60	DL
6	AK	3.00	MACOMA	.	0	3.00	DL
7	GOW	0.06	MACOMA	0.06	1	0.06	DL
8	GOW	1.00	MACOMA	.	0	1.00	DL
9	GOW	1.80	MACOMA	.	0	1.80	DL
10	GOW	2.00	MACOMA	.	0	2.00	DL
11	GOW	2.00	MACOMA	.	0	2.00	DL
12	GOW	2.00	MACOMA	.	0	2.00	DL
13	RH	0.90	MACOMA	.	0	0.90	DL
14	RH	0.90	MACOMA	.	0	0.90	DL
15	RH	1.30	MACOMA	.	0	1.30	DL
16	RH	1.60	MACOMA	.	0	1.60	DL
17	RH	1.90	MACOMA	.	0	1.90	DL
18	RH	2.00	MACOMA	.	0	2.00	DL
19	SH	0.06	MACOMA	0.06	1	0.06	DL
20	SH	0.06	MACOMA	0.06	1	0.06	DL
21	SH	0.40	MACOMA	.	0	0.40	DL
22	SH	0.50	MACOMA	.	0	0.50	DL
23	SH	0.80	MACOMA	.	0	0.80	DL
24	SH	0.80	MACOMA	.	0	0.80	DL
25	AK	0.06	NEREIS	0.06	1	0.06	DL
26	AK	1.80	NEREIS	.	0	1.80	DL
27	AK	1.90	NEREIS	.	0	1.90	DL
28	AK	3.60	NEREIS	.	0	3.60	DL
29	AK	7.50	NEREIS	.	0	7.50	DL
30	GOW	0.06	NEREIS	0.06	1	0.06	DL
31	GOW	2.30	NEREIS	.	0	2.30	DL
32	GOW	2.30	NEREIS	.	0	2.30	DL
33	GOW	3.00	NEREIS	.	0	3.00	DL
34	GOW	4.00	NEREIS	.	0	4.00	DL
35	RH	0.06	NEREIS	0.06	1	0.06	DL

36	RH	1.20	NEREIS	.	0	1.20	DL
37	RH	1.40	NEREIS	.	0	1.40	DL
38	RH	1.90	NEREIS	.	0	1.90	DL
39	RH	2.00	NEREIS	.	0	2.00	DL
40	RH	12.00	NEREIS	.	0	12.00	DL
41	SH	0.06	NEREIS	0.06	1	0.06	DL
42	SH	0.06	NEREIS	0.06	1	0.06	DL
43	SH	0.06	NEREIS	0.06	1	0.06	DL
44	SH	0.06	NEREIS	0.06	1	0.06	DL
45	SH	0.80	NEREIS	.	0	0.80	DL
46	SH	0.90	NEREIS	.	0	0.90	DL

OBS	TRT	CONC	ORG	DL	COUNT	ABSCONC	METHOD
1	AK	1.1	MACOMA	.	0	1.10	ZERO
2	AK	1.7	MACOMA	.	0	1.70	ZERO
3	AK	2.0	MACOMA	.	0	2.00	ZERO
4	AK	2.5	MACOMA	.	0	2.50	ZERO
5	AK	2.6	MACOMA	.	0	2.60	ZERO
6	AK	3.0	MACOMA	.	0	3.00	ZERO
7	GOW	0.0	MACOMA	0.06	1	0.06	ZERO
8	GOW	1.0	MACOMA	.	0	1.00	ZERO
9	GOW	1.8	MACOMA	.	0	1.80	ZERO
10	GOW	2.0	MACOMA	.	0	2.00	ZERO
11	GOW	2.0	MACOMA	.	0	2.00	ZERO
12	GOW	2.0	MACOMA	.	0	2.00	ZERO
13	RH	0.9	MACOMA	.	0	0.90	ZERO
14	RH	0.9	MACOMA	.	0	0.90	ZERO
15	RH	1.3	MACOMA	.	0	1.30	ZERO
16	RH	1.6	MACOMA	.	0	1.60	ZERO
17	RH	1.9	MACOMA	.	0	1.90	ZERO
18	RH	2.0	MACOMA	.	0	2.00	ZERO
19	SH	0.0	MACOMA	0.06	1	0.06	ZERO
20	SH	0.0	MACOMA	0.06	1	0.06	ZERO
21	SH	0.4	MACOMA	.	0	0.40	ZERO
22	SH	0.5	MACOMA	.	0	0.50	ZERO
23	SH	0.8	MACOMA	.	0	0.80	ZERO
24	SH	0.8	MACOMA	.	0	0.80	ZERO
25	AK	0.0	NEREIS	0.06	1	0.06	ZERO
26	AK	1.8	NEREIS	.	0	1.80	ZERO
27	AK	1.9	NEREIS	.	0	1.90	ZERO
28	AK	3.6	NEREIS	.	0	3.60	ZERO
29	AK	7.5	NEREIS	.	0	7.50	ZERO
30	GOW	0.0	NEREIS	0.06	1	0.06	ZERO
31	GOW	2.3	NEREIS	.	0	2.30	ZERO
32	GOW	2.3	NEREIS	.	0	2.30	ZERO
33	GOW	3.0	NEREIS	.	0	3.00	ZERO
34	GOW	4.0	NEREIS	.	0	4.00	ZERO
35	RH	0.0	NEREIS	0.06	1	0.06	ZERO
36	RH	1.2	NEREIS	.	0	1.20	ZERO
37	RH	1.4	NEREIS	.	0	1.40	ZERO
38	RH	1.9	NEREIS	.	0	1.90	ZERO
39	RH	2.0	NEREIS	.	0	2.00	ZERO
40	RH	12.0	NEREIS	.	0	12.00	ZERO
41	SH	0.0	NEREIS	0.06	1	0.06	ZERO
42	SH	0.0	NEREIS	0.06	1	0.06	ZERO
43	SH	0.0	NEREIS	0.06	1	0.06	ZERO
44	SH	0.0	NEREIS	0.06	1	0.06	ZERO
45	SH	0.8	NEREIS	.	0	0.80	ZERO
46	SH	0.9	NEREIS	.	0	0.90	ZERO

OBS	TRT	CONC	ORG	DL	COUNT	ABSCONC	SEED	METHOD
1	AK	1.1000	MACOMA	.	0	1.10	0	UNIFR
2	AK	1.7000	MACOMA	.	0	1.70	0	UNIFR
3	AK	2.0000	MACOMA	.	0	2.00	0	UNIFR
4	AK	2.5000	MACOMA	.	0	2.50	0	UNIFR
5	AK	2.6000	MACOMA	.	0	2.60	0	UNIFR
6	AK	3.0000	MACOMA	.	0	3.00	0	UNIFR
7	GOW	0.0131	MACOMA	0.06	1	0.06	0	UNIFR
8	GOW	1.0000	MACOMA	.	0	1.00	0	UNIFR
9	GOW	1.8000	MACOMA	.	0	1.80	0	UNIFR
10	GOW	2.0000	MACOMA	.	0	2.00	0	UNIFR
11	GOW	2.0000	MACOMA	.	0	2.00	0	UNIFR
12	GOW	2.0000	MACOMA	.	0	2.00	0	UNIFR
13	RH	0.9000	MACOMA	.	0	0.90	0	UNIFR
14	RH	0.9000	MACOMA	.	0	0.90	0	UNIFR
15	RH	1.3000	MACOMA	.	0	1.30	0	UNIFR
16	RH	1.6000	MACOMA	.	0	1.60	0	UNIFR
17	RH	1.9000	MACOMA	.	0	1.90	0	UNIFR
18	RH	2.0000	MACOMA	.	0	2.00	0	UNIFR
19	SH	0.0508	MACOMA	0.06	1	0.06	0	UNIFR
20	SH	0.0131	MACOMA	0.06	1	0.06	0	UNIFR
21	SH	0.4000	MACOMA	.	0	0.40	0	UNIFR
22	SH	0.5000	MACOMA	.	0	0.50	0	UNIFR
23	SH	0.8000	MACOMA	.	0	0.80	0	UNIFR
24	SH	0.8000	MACOMA	.	0	0.80	0	UNIFR
25	AK	0.0073	NEREIS	0.06	1	0.06	0	UNIFR
26	AK	1.8000	NEREIS	.	0	1.80	0	UNIFR
27	AK	1.9000	NEREIS	.	0	1.90	0	UNIFR
28	AK	3.6000	NEREIS	.	0	3.60	0	UNIFR

29	AK	7.5000	NEREIS	.	0	7.50	0	UNIFR
30	GOW	0.0459	NEREIS	0.06	1	0.06	0	UNIFR
31	GOW	2.3000	NEREIS	.	0	2.30	0	UNIFR
32	GOW	2.3000	NEREIS	.	0	2.30	0	UNIFR
33	GOW	3.0000	NEREIS	.	0	3.00	0	UNIFR
34	GOW	4.0000	NEREIS	.	0	4.00	0	UNIFR
35	RH	0.0598	NEREIS	0.06	1	0.06	0	UNIFR
36	RH	1.2000	NEREIS	.	0	1.20	0	UNIFR
37	RH	1.4000	NEREIS	.	0	1.40	0	UNIFR
38	RH	1.9000	NEREIS	.	0	1.90	0	UNIFR
39	RH	2.0000	NEREIS	.	0	2.00	0	UNIFR
40	RH	12.0000	NEREIS	.	0	12.00	0	UNIFR
41	SH	0.0545	NEREIS	0.06	1	0.06	0	UNIFR
42	SH	0.0111	NEREIS	0.06	1	0.06	0	UNIFR
43	SH	0.0085	NEREIS	0.06	1	0.06	0	UNIFR
44	SH	0.0520	NEREIS	0.06	1	0.06	0	UNIFR
45	SH	0.8000	NEREIS	.	0	0.80	0	UNIFR
46	SH	0.9000	NEREIS	.	0	0.90	0	UNIFR

OBS	TRT	ORG	_PROB_	PRED	CONC	DL	COUNT	ABSCONC	METHOD
1	AK	MACOMA	0.01	0.7628	1.1000	.	0	1.10	MLE WEIB
2	AK	MACOMA	0.21	1.6645	1.7000	.	0	1.70	MLE WEIB
3	AK	MACOMA	0.40	2.0153	2.0000	.	0	2.00	MLE WEIB
4	AK	MACOMA	0.60	2.3287	2.5000	.	0	2.50	MLE WEIB
5	AK	MACOMA	0.79	2.6564	2.6000	.	0	2.60	MLE WEIB
6	AK	MACOMA	0.99	3.4715	3.0000	.	0	3.00	MLE WEIB
7	GOW	MACOMA	0.01	0.0485	0.0485	0.06	1	0.06	MLE WEIB
8	GOW	MACOMA	0.21	0.5216	1.0000	.	0	1.00	MLE WEIB
9	GOW	MACOMA	0.40	0.9339	1.8000	.	0	1.80	MLE WEIB
10	GOW	MACOMA	0.60	1.4501	2.0000	.	0	2.00	MLE WEIB
11	GOW	MACOMA	0.79	2.1655	2.0000	.	0	2.00	MLE WEIB
12	GOW	MACOMA	0.99	4.8915	2.0000	.	0	2.00	MLE WEIB
13	RH	MACOMA	0.01	0.4705	0.9000	.	0	0.90	MLE WEIB
14	RH	MACOMA	0.21	1.0865	0.9000	.	0	0.90	MLE WEIB
15	RH	MACOMA	0.40	1.3339	1.3000	.	0	1.30	MLE WEIB
16	RH	MACOMA	0.60	1.5575	1.6000	.	0	1.60	MLE WEIB
17	RH	MACOMA	0.79	1.7938	1.9000	.	0	1.90	MLE WEIB
18	RH	MACOMA	0.99	2.3902	2.0000	.	0	2.00	MLE WEIB
19	SH	MACOMA	0.01	0.0020	0.0020	0.06	1	0.06	MLE WEIB
20	SH	MACOMA	0.21	0.0763	0.0600	0.06	1	0.06	MLE WEIB
21	SH	MACOMA	0.40	0.1862	0.4000	.	0	0.40	MLE WEIB
22	SH	MACOMA	0.60	0.3651	0.5000	.	0	0.50	MLE WEIB
23	SH	MACOMA	0.79	0.6745	0.8000	.	0	0.80	MLE WEIB
24	SH	MACOMA	0.99	2.3479	0.8000	.	0	0.80	MLE WEIB
25	AK	NEREIS	0.01	0.0067	0.0067	0.06	1	0.06	MLE WEIB
26	AK	NEREIS	0.25	0.5290	1.8000	.	0	1.80	MLE WEIB
27	AK	NEREIS	0.50	1.6602	1.9000	.	0	1.90	MLE WEIB
28	AK	NEREIS	0.75	4.0895	3.6000	.	0	3.60	MLE WEIB
29	AK	NEREIS	0.99	19.4893	7.5000	.	0	7.50	MLE WEIB
30	GOW	NEREIS	0.01	0.0216	0.0216	0.06	1	0.06	MLE WEIB
31	GOW	NEREIS	0.25	0.6528	2.3000	.	0	2.30	MLE WEIB
32	GOW	NEREIS	0.50	1.5961	2.3000	.	0	2.30	MLE WEIB
33	GOW	NEREIS	0.75	3.2293	3.0000	.	0	3.00	MLE WEIB
34	GOW	NEREIS	0.99	10.9448	4.0000	.	0	4.00	MLE WEIB
35	RH	NEREIS	0.01	0.0031	0.0031	0.06	1	0.06	MLE WEIB
36	RH	NEREIS	0.21	0.3043	1.2000	.	0	1.20	MLE WEIB
37	RH	NEREIS	0.40	0.9329	1.4000	.	0	1.40	MLE WEIB
38	RH	NEREIS	0.60	2.1749	1.9000	.	0	1.90	MLE WEIB
39	RH	NEREIS	0.79	4.7042	2.0000	.	0	2.00	MLE WEIB
40	RH	NEREIS	0.99	22.5570	12.0000	.	0	12.00	MLE WEIB
41	SH	NEREIS	0.01	0.0000	0.0000	0.06	1	0.06	MLE WEIB
42	SH	NEREIS	0.21	0.0007	0.0007	0.06	1	0.06	MLE WEIB
43	SH	NEREIS	0.40	0.0083	0.0083	0.06	1	0.06	MLE WEIB
44	SH	NEREIS	0.60	0.0530	0.0530	0.06	1	0.06	MLE WEIB
45	SH	NEREIS	0.79	0.2872	0.8000	.	0	0.80	MLE WEIB
46	SH	NEREIS	0.99	8.8872	0.9000	.	0	0.90	MLE WEIB

OBS	TRT	ORG	_PROB_	PRED	CONC	DL	COUNT	ABSCONC	METHOD
1	AK	MACOMA	0.01	0.943	1.1000	.	0	1.10	MLE LOGN
2	AK	MACOMA	0.21	1.563	1.7000	.	0	1.70	MLE LOGN
3	AK	MACOMA	0.40	1.879	2.0000	.	0	2.00	MLE LOGN
4	AK	MACOMA	0.60	2.224	2.5000	.	0	2.50	MLE LOGN
5	AK	MACOMA	0.79	2.673	2.6000	.	0	2.60	MLE LOGN
6	AK	MACOMA	0.99	4.432	3.0000	.	0	3.00	MLE LOGN
7	GOW	MACOMA	0.01	0.026	0.0264	0.06	1	0.06	MLE LOGN
8	GOW	MACOMA	0.21	0.261	1.0000	.	0	1.00	MLE LOGN
9	GOW	MACOMA	0.40	0.602	1.8000	.	0	1.80	MLE LOGN
10	GOW	MACOMA	0.60	1.294	2.0000	.	0	2.00	MLE LOGN
11	GOW	MACOMA	0.79	2.980	2.0000	.	0	2.00	MLE LOGN
12	GOW	MACOMA	0.99	29.511	2.0000	.	0	2.00	MLE LOGN
13	RH	MACOMA	0.01	0.641	0.9000	.	0	0.90	MLE LOGN
14	RH	MACOMA	0.21	1.049	0.9000	.	0	0.90	MLE LOGN
15	RH	MACOMA	0.40	1.255	1.3000	.	0	1.30	MLE LOGN
16	RH	MACOMA	0.60	1.479	1.6000	.	0	1.60	MLE LOGN
17	RH	MACOMA	0.79	1.770	1.9000	.	0	1.90	MLE LOGN
18	RH	MACOMA	0.99	2.896	2.0000	.	0	2.00	MLE LOGN
19	SH	MACOMA	0.01	0.005	0.0050	0.06	1	0.06	MLE LOGN
20	SH	MACOMA	0.21	0.056	0.0563	0.06	1	0.06	MLE LOGN
21	SH	MACOMA	0.40	0.136	0.4000	.	0	0.40	MLE LOGN

22	SH	MACOMA	0.60	0.307	0.5000	.	0	0.50	MLE LOGN
23	SH	MACOMA	0.79	0.743	0.8000	.	0	0.80	MLE LOGN
24	SH	MACOMA	0.99	8.447	0.8000	.	0	0.80	MLE LOGN
25	AK	NEREIS	0.01	0.010	0.0102	0.06	1	0.06	MLE LOGN
26	AK	NEREIS	0.25	0.297	1.8000	.	0	1.80	MLE LOGN
27	AK	NEREIS	0.50	1.174	1.9000	.	0	1.90	MLE LOGN
28	AK	NEREIS	0.75	4.651	3.6000	.	0	3.60	MLE LOGN
29	AK	NEREIS	0.99	135.300	7.5000	.	0	7.50	MLE LOGN
30	GOW	NEREIS	0.01	0.013	0.0129	0.06	1	0.06	MLE LOGN
31	GOW	NEREIS	0.25	0.304	2.3000	.	0	2.30	MLE LOGN
32	GOW	NEREIS	0.50	1.106	2.3000	.	0	2.30	MLE LOGN
33	GOW	NEREIS	0.75	4.020	3.0000	.	0	3.00	MLE LOGN
34	GOW	NEREIS	0.99	94.759	4.0000	.	0	4.00	MLE LOGN
35	RH	NEREIS	0.01	0.015	0.0152	0.06	1	0.06	MLE LOGN
36	RH	NEREIS	0.21	0.254	1.2000	.	0	1.20	MLE LOGN
37	RH	NEREIS	0.40	0.707	1.4000	.	0	1.40	MLE LOGN
38	RH	NEREIS	0.60	1.808	1.9000	.	0	1.90	MLE LOGN
39	RH	NEREIS	0.79	5.036	2.0000	.	0	2.00	MLE LOGN
40	RH	NEREIS	0.99	84.168	12.0000	.	0	12.00	MLE LOGN
41	SH	NEREIS	0.01	0.000	0.0000	0.06	1	0.06	MLE LOGN
42	SH	NEREIS	0.21	0.002	0.0016	0.06	1	0.06	MLE LOGN
43	SH	NEREIS	0.40	0.009	0.0092	0.06	1	0.06	MLE LOGN
44	SH	NEREIS	0.60	0.045	0.0452	0.06	1	0.06	MLE LOGN
45	SH	NEREIS	0.79	0.257	0.8000	.	0	0.80	MLE LOGN
46	SH	NEREIS	0.99	30.565	0.9000	.	0	0.90	MLE LOGN

TOTAL PCB BIOACCUMULATION IN MACOMA NASUTA AND NEREIS VIRENS

SPECIES	SEDIMENT	PERCENT OF DATA THAT ARE CENSORED	NUMBER OF REPLICATES
MACOMA	AK	0.0000	6
MACOMA	GOW	16.6667	6
MACOMA	RH	0.0000	6
MACOMA	SH	33.3333	6
NEREIS	AK	20.0000	5
NEREIS	GOW	20.0000	5
NEREIS	RH	16.6667	6
NEREIS	SH	66.6667	6

TOTAL PCB BIOACCUMULATION IN MACOMA NASUTA AND NEREIS VIRENS

SPECIES	SEDI-MENT	METHOD	CV	N	MEAN	VARIANCE	STD. DEV.	T VALUE FOR (1-ALPHA=.95, N-1)
MACOMA	AK	DL	0.32056	6	2.15000	0.4750	0.68920	2.01505
MACOMA	AK	DL/2	0.32056	6	2.15000	0.4750	0.68920	2.01505
MACOMA	AK	MLE LOGN	0.32056	6	2.15000	0.4750	0.68920	2.01505
MACOMA	AK	MLE WEIB	0.32056	6	2.15000	0.4750	0.68920	2.01505
MACOMA	AK	UNIFR	0.32056	6	2.15000	0.4750	0.68920	2.01505
MACOMA	AK	ZERO	0.32056	6	2.15000	0.4750	0.68920	2.01505
MACOMA	GOW	DL	0.53839	6	1.47667	0.6321	0.79503	2.01505
MACOMA	GOW	DL/2	0.54750	6	1.47167	0.6492	0.80574	2.01505
MACOMA	GOW	MLE LOGN	0.54860	6	1.47107	0.6513	0.80703	2.01505
MACOMA	GOW	MLE WEIB	0.54188	6	1.47475	0.6386	0.79914	2.01505
MACOMA	GOW	UNIFR	0.55267	6	1.46885	0.6590	0.81179	2.01505
MACOMA	GOW	ZERO	0.55670	6	1.46667	0.6667	0.81650	2.01505
MACOMA	RH	DL	0.33508	6	1.43333	0.2307	0.48028	2.01505
MACOMA	RH	DL/2	0.33508	6	1.43333	0.2307	0.48028	2.01505
MACOMA	RH	MLE LOGN	0.33508	6	1.43333	0.2307	0.48028	2.01505
MACOMA	RH	MLE WEIB	0.33508	6	1.43333	0.2307	0.48028	2.01505
MACOMA	RH	UNIFR	0.33508	6	1.43333	0.2307	0.48028	2.01505
MACOMA	RH	ZERO	0.33508	6	1.43333	0.2307	0.48028	2.01505
MACOMA	SH	DL	0.76169	6	0.43667	0.1106	0.33261	2.01505
MACOMA	SH	DL/2	0.81158	6	0.42667	0.1199	0.34628	2.01505
MACOMA	SH	MLE LOGN	0.81139	6	0.42688	0.1200	0.34637	2.01505
MACOMA	SH	MLE WEIB	0.81101	6	0.42700	0.1199	0.34630	2.01505
MACOMA	SH	UNIFR	0.80869	6	0.42733	0.1194	0.34558	2.01505
MACOMA	SH	ZERO	0.86422	6	0.41667	0.1297	0.36009	2.01505
NEREIS	AK	DL	0.95020	5	2.97200	7.9749	2.82399	2.13185
NEREIS	AK	DL/2	0.95474	5	2.96600	8.0188	2.83175	2.13185
NEREIS	AK	MLE LOGN	0.95775	5	2.96204	8.0479	2.83689	2.13185
NEREIS	AK	MLE WEIB	0.95828	5	2.96135	8.0530	2.83779	2.13185
NEREIS	AK	UNIFR	0.95819	5	2.96145	8.0523	2.83765	2.13185
NEREIS	AK	ZERO	0.95930	5	2.96000	8.0630	2.83954	2.13185
NEREIS	GOW	DL	0.62114	5	2.33200	2.0981	1.44849	2.13185
NEREIS	GOW	DL/2	0.62780	5	2.32600	2.1324	1.46027	2.13185
NEREIS	GOW	MLE LOGN	0.63162	5	2.32258	2.1521	1.46699	2.13185
NEREIS	GOW	MLE WEIB	0.62969	5	2.32431	2.1421	1.46359	2.13185

NEREIS	GOW	UNIFR	0.62426	5	2.32918	2.1142	1.45401	2.13185
NEREIS	GOW	ZERO	0.63451	5	2.32000	2.1670	1.47207	2.13185
NEREIS	RH	DL	1.42829	6	3.09333	19.5203	4.41817	2.01505
NEREIS	RH	DL/2	1.43194	6	3.08833	19.5568	4.42231	2.01505
NEREIS	RH	MLE LOGN	1.43375	6	3.08586	19.5750	4.42436	2.01505
NEREIS	RH	MLE WEIB	1.43523	6	3.08386	19.5898	4.42603	2.01505
NEREIS	RH	UNIFR	1.42832	6	3.09329	19.5206	4.41821	2.01505
NEREIS	RH	ZERO	1.43561	6	3.08333	19.5937	4.42647	2.01505
NEREIS	SH	DL	1.26550	6	0.32333	0.1674	0.40918	2.01505
NEREIS	SH	DL/2	1.39986	6	0.30333	0.1803	0.42463	2.01505
NEREIS	SH	MLE LOGN	1.48001	6	0.29268	0.1876	0.43317	2.01505
NEREIS	SH	MLE WEIB	1.47280	6	0.29368	0.1871	0.43253	2.01505
NEREIS	SH	UNIFR	1.39398	6	0.30436	0.1800	0.42428	2.01505
NEREIS	SH	ZERO	1.55321	6	0.28333	0.1937	0.44008	2.01505

## Program BLOCKS.SAS to Analyze Blocked Design

The following statements perform tests of assumptions (Shapiro-Wilk's Test for normality and Levene's Test for equality of variances) and the LSD test when the data include a blocking variable. In the program output, the *F* and P-value for the treatment variable (TRT) Type III sums of squares are used to determine the significance of Levene's Test.

### BLOCKS.SAS program statements

```

OPTIONS LINESIZE=79 PAGESIZE=500 NODATE NONUMBER;
LIBNAME Q 'C:\SAS';

/* Data analysis for a blocked design using untransformed data (CONC),
   log-transformed data (LOGCONC) and rankits (RANKIT). The treatment
   variable is TRT and the block variable is CONTAM. */

TITLE 'BLOCKED DESIGN ANALYSIS';
DATA A1;
INPUT TRT $ CONTAM $ CONC;
CARDS;
BS CD 6.84
BS CD 6.94
BS CD 6.24
BS CD 7.76
BS CD 6.56
BS CD 4.82
S50 CD 5.80
S50 CD 6.00
S50 CD 6.31
S50 CD 7.76
S50 CD 6.95
S50 CD 7.69
BS CR 5.10
BS CR 2.30
BS CR 6.10
BS CR 2.60
BS CR 4.50
BS CR 3.50
S50 CR 2.90
S50 CR 3.20
S50 CR 3.20
S50 CR 4.40
S50 CR 3.30
S50 CR 4.60
BS HG .289
BS HG .296
BS HG .263
BS HG .283
BS HG .296
BS HG .282
S50 HG .264
S50 HG .315
S50 HG .366
S50 HG .318
S50 HG .266
S50 HG .325
;
DATA A;
SET A1;
LOGCONC=LOG10(CONC);

```

```

PROC RANK NORMAL=BLOM OUT=B;
VAR CONC;
RANKS RANKIT;

/* Perform Shapiro-Wilk's Test for normality */

PROC GLM NOPRINT;
CLASS TRT CONTAM;
MODEL CONC LOGCONC RANKIT=TRT CONTAM;
OUTPUT OUT=Z R=RESID RESIDLOG RRANKIT;
PROC UNIVARIATE NORMAL;
VAR RESID RESIDLOG RRANKIT;
TITLE2 'SHAPIRO-WILKS TEST';

/* Perform Levene's Test for equality of variances */

DATA Z1; SET Z;
ABSRESID=ABS(RESID);
ABSRANK=ABS(RRANKIT);
ABSLOG=ABS(RESIDLOG);
PROC GLM;
CLASS TRT CONTAM;
MODEL ABSRESID ABSLOG ABSRANK=TRT CONTAM;
TITLE2 'LEVENES TEST';

/* Perform LSD test */

PROC GLM DATA=B;
CLASS TRT CONTAM;
MODEL CONC LOGCONC RANKIT=TRT CONTAM;
MEANS TRT/LSD;
TITLE2 'LSD TEST';

```

## BLOCKS.SAS program output

*Note: output is given for untransformed data only.*

```

BLOCKED DESIGN ANALYSIS
SHAPIRO-WILKS TEST

UNIVARIATE PROCEDURE

Variable=RESID

```

	N	36	Sum Wgts	36
Mean	0	0	Sum	0
Std Dev	0.802699	Variance	0.644326	
Skewness	0.28859	Kurtosis	1.234563	
USS	22.5514	CSS	22.55141	
CV	.	Std Mean	0.133783	
T:Mean=0	0	Prob> T	1.0000	
Sgn Rank	-12	Prob> S	0.8535	
Num ^= 0	36			
W:Normal	0.964422	Prob<W	0.3691	

```

BLOCKED DESIGN ANALYSIS
LEVENES TEST

General Linear Models Procedure

Dependent Variable: ABSRESID

```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	5.37267250	1.79089083	9.39	0.0001
Error	32	6.10444268	0.19076383		
Corrected Total	35	11.47711518			

```

R-Square          C.V.          Root MSE          ABSRESID Mean
0.468120          78.74833          0.4367652          0.55463426

```

Source	DF	Type I SS	Mean Square	F Value	Pr > F
TRT	1	0.24225537	0.24225537	1.27	0.2682
CONTAM	2	5.13041713	2.56520857	13.45	0.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	1	0.24225537	0.24225537	1.27	0.2682
CONTAM	2	5.13041713	2.56520857	13.45	0.0001

BLOCKED DESIGN ANALYSIS  
LSD TEST

General Linear Models Procedure

Dependent Variable: CONC

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	242.29925397	80.76641799	114.61	0.0001
Error	32	22.55140500	0.70473141		
Corrected Total	35	264.85065897			

R-Square	C.V.	Root MSE	CONC Mean
0.914852	23.43960	0.8394828	3.58147222

Source	DF	Type I SS	Mean Square	F Value	Pr > F
TRT	1	0.02805625	0.02805625	0.04	0.8431
CONTAM	2	242.27119772	121.13559886	171.89	0.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	1	0.02805625	0.02805625	0.04	0.8431
CONTAM	2	242.27119772	121.13559886	171.89	0.0001

BLOCKED DESIGN ANALYSIS  
LSD TEST

General Linear Models Procedure

T tests (LSD) for variable: CONC

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 32 MSE= 0.704731  
Critical Value of T= 2.04  
Least Significant Difference= 0.57

Means with the same letter are not significantly different.

T Grouping	Mean	N	TRT
A	3.609	18	BS
A	3.554	18	S50

# Appendix B

## SYSTAT Programs

---

SAS programs provided in Appendix D of the Inland Testing Manual (U.S. Environmental Protection Agency/U.S. Army Corps of Engineers 1994)<sup>1</sup> are duplicated herein using SYSTAT version 5.0 for DOS (Wilkinson 1990a). SYSTAT is a registered trademark of SYSTAT, Inc. Recently, SYSTAT, Inc., was acquired by SPSS, Inc., and SYSTAT version 6.0 for DOS is now available. The use of this product name does not constitute official endorsement or approval of this or any other product. Other equally acceptable software products are commercially available and may be used to perform these analyses.

The interpretation of test results is described in Appendix D of the Inland Testing Manual. There are several differences between the SAS and SYSTAT programs. The SYSTAT programs calculate a normality test called Lilliefors' Test, while SAS programs calculate the Shapiro-Wilk's Test. The decision rule for Lilliefors' Test is to reject the null hypothesis of normality at the appropriate level of significance if the maximum distance calculated exceeds the  $1 - \alpha$  quantile. Acceptance of the null hypothesis does not mean that the parent population is normal, only that the normal distribution does not seem to be an unreasonable approximation to the true unknown distribution (Conover 1980). The algorithms used in SYSTAT provide values that are slightly different from those shown in Table A15 of Conover (1980). The approximations used in SYSTAT differ from the table values by less than 0.01 (Wilkinson 1990a:397). The SYSTAT Lilliefors' option automatically standardizes the variable tested. Lilliefors' Test is illustrated in the analysis of water column toxicity data below (WATCOL.CMD).

SYSTAT produces the LSD results in a matrix of probabilities, whereas SAS denotes differences with letters of the alphabet. LSD results are illustrated in the analysis of benthic toxicity data below (BENTOX.CMD). Another difference between the SYSTAT and SAS programs is the use of ranks instead of rankits (see node 10 of Figure D-1 in the Inland Testing Manual). For example, Conover's *T* Test, which is an LSD-type test using

---

<sup>1</sup> References cited in this appendix are located at the end of the main text.

ranks, is included in the programs instead of the LSD test on rankits. Test results using ranks and rankits are interpreted similarly. SYSTAT statements can be constructed to calculate rankits. Rankits are illustrated using benthic toxicity data (BENTOX.CMD).

Output values from one step that are used as input in subsequent steps must sometimes be manually inserted in the subsequent steps. Variable values that were inserted are underlined when they appear in output and when they are used as input. Comment statements in the following format:

```
/* Comment line */
```

were added for clarity and must be removed before executing the SYSTAT code. Several lines of output have been deleted from each program to reduce the volume of output.

## Program WATCOL.CMD for Water Column Toxicity Test Data Analysis

The following program is quite similar to WATTOX.SAS in Appendix D of the Inland Testing Manual. WATCOL.CMD is a program to compare water column toxicity data, control survival versus 100-percent elutriate survival, using arcsine-square root transformation of the survival proportions. Analyses include mean survival for control and all elutriate dilutions, Kolmogorov-Smirnov test for normality using probabilities developed by Lilliefors, and *t*-tests for equal or unequal variances.

### WATCOL.CMD program statements

```
DATA
FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
SAVE WATCOL / 'Input the toxicity test data after the RUN statement,',
'list the treatment code, replicate and number of survivors.'
INPUT TRT REP SURV \
LABEL TRT/0='DIL. WATER ', 1='100% ELUT. ',
2='50% ELUT.', 3='25% ELUT.', 4='12.5% ELUT.'
FORMAT=5
RUN
0 1 20 0 2 19 0 3 20 0 4 20 0 5 19
1 1 6 1 2 7 1 3 9 1 4 5 1 5 8
2 1 8 2 2 8 2 3 9 2 4 10 2 5 11
3 1 12 3 2 18 3 3 15 3 4 14 3 5 13
4 1 17 4 2 17 4 3 18 4 4 16 4 5 18
~

/* Input no. of organisms (M) per test container at start of test. */
/* Format, print, sort the data. Print no. of observations, mean, and */
/* standard error for survival in each treatment. */

USE WATCOL
LET M=20
LET ARCSURV=ASN(SQR(SURV/M))
PRINT TRT, REP, M, SURV, ARCSURV
SAVE A0
RUN
USE A0
SORT TRT
RUN
STATS
USE A0
BY TRT
STATISTICS SURV / SUM N MEAN SEM
```

```

/* Delete data not needed for the dilution water-100% elutriate comparison. */
/* Print descriptive stats. */

DATA
USE A0
IF TRT>1 THEN DELETE
SAVE A
RUN
STATS
USE A
BY TRT
STATISTICS ARCSURV / N MEAN VARIANCE SD SEM

/* Test normality using Lilliefors' Test */

NPAR
USE A
KS ARCSURV / LILLIEFORS

/* Conduct t-test, and F' test for equality of variances */

STATS
USE A
TTEST ARCSURV*TRT
DATA
LET N0=5
LET N1=5
LET VAR0=0.01526
LET VAR1=0.00699
IF VAR0>=VAR1 THEN LET DFNUM=N0-1
IF VAR0>=VAR1 THEN LET DFDEN=N1-1
IF VAR0<VAR1 THEN LET DFNUM=N1-1
IF VAR0<VAR1 THEN LET DFDEN=N0-1
IF VAR0>=VAR1 THEN LET FMAX=((N0-1)*VAR0)/((N1-1)*VAR1)
IF VAR0<VAR1 THEN LET FMAX=((N1-1)*VAR1)/((N0-1)*VAR0)
LET PROBF=(1-PCF(FMAX,DFNUM,DFDEN))*2
PRINT FMAX, DFNUM, DFDEN, PROBF
RUN

/* Convert data to ranks, conduct t-test, and F' test for
/* equality of variances. */

DATA
USE A
LET RANKSURV=SURV
RANK RANKSURV
SAVE A1
RUN
REPEAT 10
LIST TRT, SURV
RUN
STATS
USE A1
TTEST RANKSURV*TRT
DATA
LET N0=5
LET N1=5
LET VAR0=1.369^2
LET VAR1=1.581^2
IF VAR0>=VAR1 THEN LET DFNUM=N0-1
IF VAR0>=VAR1 THEN LET DFDEN=N1-1
IF VAR0<VAR1 THEN LET DFNUM=N1-1
IF VAR0<VAR1 THEN LET DFDEN=N0-1
IF VAR0>=VAR1 THEN LET FMAX=((N0-1)*VAR0)/((N1-1)*VAR1)
IF VAR0<VAR1 THEN LET FMAX=((N1-1)*VAR1)/((N0-1)*VAR0)
LET PROBF=(1-PCF(FMAX,DFNUM,DFDEN))*2
PRINT FMAX, DFNUM, DFDEN, PROBF
RUN

/* Calculate minimum significant difference and power of a t-test to detect */
/* true population differences of 10, 20, 30, 40, and 50% below mean */
/* dilution water survival. */

DATA
LET MEANPCT=19.6/20
LET N0=5
LET S20=0.01526
LET N1=5
LET S21=0.00699
LET DF=N0+N1-2
LET N=(N0+N1)/2
LET S2POOL=(S20*(N0-1)+S21*(N1-1))/DF
LET TALPHA=TIF(.95,DF)
LET DMIN=TALPHA*SQR(2*S2POOL/N)

SAVE B2
PRINT M, N, MEANPCT, S2POOL, DF, TALPHA, DMIN
RUN

DATA
USE B2

```

```

FOR I=10 TO 50 STEP 10
LET SEDSURV=MEANPCT-I/100
LET ARCSURV=ASN(SQR(SED SURV))
LET ARCDIFF=1.48059-ARCSURV
LET TBETA=(SQR(N)*ARCDIFF)/SQR(2*S2POOL)-TALPHA
LET POWER=TCF(TBETA,DF)
PRINT I, SEDSURV, ARCSURV, ARCDIFF, TBETA, POWER
NEXT
STOP
RUN

```

## WATCOL.CMD program output

### WATER COLUMN TOXICITY DATA

TRT	REP	M	SURV	ARCSURV
0.00000	1.00000	20.00000	20.00000	1.57080
0.00000	2.00000	20.00000	19.00000	1.34528
0.00000	3.00000	20.00000	20.00000	1.57080
0.00000	4.00000	20.00000	20.00000	1.57080
0.00000	5.00000	20.00000	19.00000	1.34528
1.00000	1.00000	20.00000	6.00000	0.57964
1.00000	2.00000	20.00000	7.00000	0.63305
1.00000	3.00000	20.00000	9.00000	0.73531
1.00000	4.00000	20.00000	5.00000	0.52360
1.00000	5.00000	20.00000	8.00000	0.68472
2.00000	1.00000	20.00000	8.00000	0.68472
2.00000	2.00000	20.00000	8.00000	0.68472
2.00000	3.00000	20.00000	9.00000	0.73531
2.00000	4.00000	20.00000	10.00000	0.78540
2.00000	5.00000	20.00000	11.00000	0.83548
3.00000	1.00000	20.00000	12.00000	0.88608
3.00000	2.00000	20.00000	18.00000	1.24905
3.00000	3.00000	20.00000	15.00000	1.04720
3.00000	4.00000	20.00000	14.00000	0.99116
3.00000	5.00000	20.00000	13.00000	0.93774
4.00000	1.00000	20.00000	17.00000	1.17310
4.00000	2.00000	20.00000	17.00000	1.17310
4.00000	3.00000	20.00000	18.00000	1.24905
4.00000	4.00000	20.00000	16.00000	1.10715
4.00000	5.00000	20.00000	18.00000	1.24905

### WATER COLUMN TOXICITY SUMMARY DATA

THE FOLLOWING RESULTS ARE FOR:

TRT = 0.00000

SURV

N OF CASES	5
MEAN	19.60000
STD. ERROR	0.24495
SUM	98.00000

THE FOLLOWING RESULTS ARE FOR:

TRT = 1.00000

SURV

N OF CASES	5
MEAN	7.00000
STD. ERROR	0.70711
SUM	35.00000

THE FOLLOWING RESULTS ARE FOR:

TRT = 2.00000

SURV

N OF CASES	5
MEAN	9.20000
STD. ERROR	0.58310
SUM	46.00000

THE FOLLOWING RESULTS ARE FOR:

TRT = 3.00000

SURV

N OF CASES	5
MEAN	14.40000
STD. ERROR	1.02956
SUM	72.00000

THE FOLLOWING RESULTS ARE FOR:

TRT = 4.00000

```

SURV
N OF CASES      5
MEAN            17.20000
STD. ERROR     0.37417
SUM            86.00000

```

ARCSINE-SQUARE ROOT TRANSFORMATION

THE FOLLOWING RESULTS ARE FOR:  
TRT = 0.00000

ARCSURV

```

N OF CASES      5
MEAN            1.48059
VARIANCE        0.01526
STANDARD DEV    0.12352
STD. ERROR      0.05524

```

THE FOLLOWING RESULTS ARE FOR:  
TRT = 1.00000

ARCSURV

```

N OF CASES      5
MEAN            0.63126
VARIANCE        0.00699
STANDARD DEV    0.08358
STD. ERROR      0.03738

```

/\* Normality test \*/

KOLMOGOROV-SMIRNOV ONE SAMPLE TEST USING STANDARD NORMAL DISTRIBUTION

VARIABLE	N-OF-CASES	MAXDIF	LILLIEFORS	PROBABILITY (2-TAIL)
ARCSURV	10.00000	0.25779	0.05838	

/\* The decision rule is to reject the null hypothesis at the appropriate level of significance if the maximum distance calculated exceeds the 1 -  $\alpha$  quantile. The recommended level of significance is  $\alpha = 0.05$  (Table D-2 of the Inland Testing Manual). Quantiles are provided in Table A15 (Conover 1980). The maximum distance is 0.25779. The 1 -  $\alpha$  quantile is 0.258. We conclude that the normal distribution does not seem to be an unreasonable approximation to the binomial distribution. However, since the Lilliefors probability 0.05838 is greater than 0.05, we would reach the same conclusion (i.e., fail to reject the null hypothesis) without using Table A15. \*/

ARCSINE-SQUARE ROOT TRANSFORMATION

INDEPENDENT SAMPLES T-TEST ON ARCSURV GROUPED BY TRT

GROUP	N	MEAN	SD
0.00000	5	1.48059	0.12352
1.00000	5	0.63126	0.08358

SEPARATE VARIANCES T = 12.73397 DF = 7.0 PROB = 0.00000  
POOLED VARIANCES T = 12.73397 DF = 8 PROB = 0.00000

F'	DF1	DF2	PROB>F'
2.18312	4.00000	4.00000	0.46815

CONOVER'S T-TEST USING RANKS

INDEPENDENT SAMPLES T-TEST ON RANKSURV GROUPED BY TRT

GROUP	N	MEAN	SD
0.00000	5	8.00000	1.36931
1.00000	5	3.00000	1.58114

SEPARATE VARIANCES T = 5.34522 DF = 7.8 PROB = 0.00074  
POOLED VARIANCES T = 5.34522 DF = 8 PROB = 0.00069

F'	DF1	DF2	PROB>F'
1.33370	4.00000	4.00000	0.78698

POWER OF T-TEST TO DETECT A TRUE POPULATION DIFFERENCE (D)  
FROM MEAN DILUTION WATER SURVIVAL USING ARCSINE TRANSFORMATION

# OF REPLICATE	N	MEAN DIL. WATER SURVIVAL	POOLED VARIANCE	DEGREES OF FREEDOM	T-VALUE FOR (.95, DF)	MINIMUM SIGNIFICANT DIFFERENCE
20.00	5.0	0.980	0.01113	8.00	1.85955	0.12405

% REDUCTION IN SURVIVAL FROM DIL. WATER	ARSINE		D	T-VALUE FOR (1-BETA, DF)	POWER
	100% ELUTRIATE SURVIVAL	100% ELUTRIATE SURVIVAL			
10.00000	0.88000	1.21705	0.26354	2.09101	0.96505
20.00000	0.78000	1.08259	0.39800	4.10671	0.99830
30.00000	0.68000	0.96953	0.51106	5.80153	0.99980
40.00000	0.58000	0.86574	0.61485	7.35739	0.99996
50.00000	0.48000	0.76539	0.71520	8.86171	0.99999

## Program BENTOX.CMD for Benthic Toxicity Test Data Analysis

The following program is quite similar to BENTOX.SAS in Appendix D of the Inland Testing Manual. BENTOX.CMD is a program to compare benthic toxicity data, reference survival versus survival from one or more test sediments using arcsine-square root transformation of the data. Analyses include mean survival for reference and test sediment(s), Kolmogorov-Smirnov (KS) test for normality using probabilities developed by Lilliefors, and *t*-tests for equal or unequal variances. The test results are interpreted as described in Appendix D of the Inland Testing Manual.

### BENTOX.CMD program statements

```

/* Input the toxicity test data after the RUN statement. */
/* List the treatment code, replicate and number of survivors. */
/* FORMAT=5 requests output displaying 5 decimals. */

DATA
FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
SAVE BENTHIC
INPUT TRT REP SURV \
FORMAT=5
RUN
1 1 20 1 2 20 1 3 19 1 4 19 1 5 20
2 1 17 2 2 16 2 3 18 2 4 17 2 5 15
3 1 15 3 2 16 3 3 13 3 4 17 3 5 11
4 1 17 4 2 12 4 3 10 4 4 16 4 5 13
~

/* Input no. of organisms (M) per test container at start of test. */
/* Format, print, sort the data. Print no. of observations, mean, */
/* variance, and standard error for each treatment. */
/* Treatment code 1=Reference, 2=Sediment 1, 3=Sediment 2, 4=Sediment 3. */

USE BENTHIC
LET M=20
LET ARCSURV=ASN(SQR(SURV/M))
SAVE A0
PRINT TRT, REP, M, SURV, ARCSURV
RUN
USE A0
SORT TRT
RUN
STATS
USE A0
BY TRT
STATISTICS SURV ARCSURV / SUM N MEAN VARIANCE SEM

/* Perform KS normality test */

NPAR
USE A0
KS ARCSURV / LILLIEFORS

/* Perform Bartlett's homogeneity of variance Test */

STATS
USE A0
BY TRT
PRINT=LONG
STATISTICS ARCSURV

```

```

/* LSD test using ARCSURV */

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
MGLH
USE A0
CATEGORY TRT
ANOVA ARCSURV
ESTIMATE
HYPOTHESIS
POST TRT / LSD
TEST

/* t-test comparing the Reference and Sediment 1 */

DATA
USE A0
IF TRT>2 THEN DELETE
SAVE T1
RUN
STATS
USE T1
TTEST ARCSURV*TRT
DATA
LET N0=5
LET N1=5
LET VAR0=0.01526
LET VAR1=0.00582
IF VAR0>=VAR1 THEN LET DFNUM=N0-1
IF VAR0>=VAR1 THEN LET DFDEN=N1-1
IF VAR0<VAR1 THEN LET DFNUM=N1-1
IF VAR0<VAR1 THEN LET DFDEN=N0-1
IF VAR0>=VAR1 THEN LET FMAX=((N0-1)*VAR0)/((N1-1)*VAR1)
IF VAR0<VAR1 THEN LET FMAX=((N1-1)*VAR1)/((N0-1)*VAR0)
LET PROBF=(1-FCF(FMAX,DFNUM,DFDEN))*2
PRINT FMAX, DFNUM, DFDEN, PROBF
RUN

/* t-test comparing the Reference and Sediment 2 */

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE A0
IF (TRT=2 OR TRT=4) THEN DELETE
SAVE T2
RUN
STATS
USE T2
TTEST ARCSURV*TRT
DATA
LET N0=5
LET N1=5
LET VAR0=0.01526
LET VAR1=0.01815
IF VAR0>=VAR1 THEN LET DFNUM=N0-1
IF VAR0>=VAR1 THEN LET DFDEN=N1-1
IF VAR0<VAR1 THEN LET DFNUM=N1-1
IF VAR0<VAR1 THEN LET DFDEN=N0-1
IF VAR0>=VAR1 THEN LET FMAX=((N0-1)*VAR0)/((N1-1)*VAR1)
IF VAR0<VAR1 THEN LET FMAX=((N1-1)*VAR1)/((N0-1)*VAR0)
LET PROBF=(1-FCF(FMAX,DFNUM,DFDEN))*2
PRINT FMAX, DFNUM, DFDEN, PROBF
RUN

/* t-test comparing the Reference and Sediment 3 */

DATA
USE A0
IF (TRT=2 OR TRT=3) THEN DELETE
SAVE T3
RUN
STATS
USE T3
TTEST ARCSURV*TRT
DATA
LET N0=5
LET N1=5
LET VAR0=0.01526
LET VAR1=0.02548
IF VAR0>=VAR1 THEN LET DFNUM=N0-1
IF VAR0>=VAR1 THEN LET DFDEN=N1-1
IF VAR0<VAR1 THEN LET DFNUM=N1-1
IF VAR0<VAR1 THEN LET DFDEN=N0-1
IF VAR0>=VAR1 THEN LET FMAX=((N0-1)*VAR0)/((N1-1)*VAR1)
IF VAR0<VAR1 THEN LET FMAX=((N1-1)*VAR1)/((N0-1)*VAR0)
LET PROBF=(1-FCF(FMAX,DFNUM,DFDEN))*2
PRINT FMAX, DFNUM, DFDEN, PROBF
RUN

USE A0
SORT TRT
RUN

```

```

/* Calculate ranks */
DATA
USE A0
LET RANKSURV=SURV
RANK RANKSURV
SAVE RAO
RUN

/* Calculate mean of ranks */

STATS
USE RAO
BY TRT
STATISTICS RANKSURV / N MEAN VARIANCE

/* Calculate residuals for Levene's Test */

DATA
USE RAO
IF TRT=1 THEN LET ABSDEV=ABS(RANKSURV-18.0)
IF TRT=2 THEN LET ABSDEV=ABS(RANKSURV-11.1)
IF TRT=3 THEN LET ABSDEV=ABS(RANKSURV-6.9)
IF TRT=4 THEN LET ABSDEV=ABS(RANKSURV-6.0)
SAVE LEVENE
RUN

/* Calculate Levene's Test */

MGLH
USE LEVENE
CATEGORY TRT
ANOVA ABSDEV
ESTIMATE

/* Conover T-test (i.e., LSD on ranks) */

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE A0
LET RANKSURV=SURV
RANK RANKSURV
SAVE RBENTOX
RUN

MGLH
USE RBENTOX
CATEGORY TRT
ANOVA RANKSURV
ESTIMATE
HYPOTHESIS
POST TRT / LSD
TEST

/* t-tests using ranks */

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE A0
IF TRT>2 THEN DELETE
SAVE BENT1
RUN

DATA
USE A0
IF (TRT=2 OR TRT=4) THEN DELETE
SAVE BENT2
RUN

DATA
USE A0
IF (TRT=2 OR TRT=3) THEN DELETE
SAVE BENT3
RUN

/* Calculate ranks and t-test (Reference and Sediment 1) */

DATA
USE BENT1
LET RANKSURV=SURV
RANK RANKSURV
SAVE RBENT1
RUN

STATS
USE RBENT1
TTEST RANKSURV*TRT

/* Calculate ranks and t-test (Reference and Sediment 2) */

```

```

DATA
USE BENT2
LET RANKSURV=SRV
RANK RANKSURV
SAVE RBENT2
RUN

STATS
USE RBENT2
TTEST RANKSURV*TRT

/* Calculate ranks and t-test (Reference and Sediment 3) */

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE BENT3
LET RANKSURV=SRV
RANK RANKSURV
SAVE RBENT3
RUN

STATS
USE RBENT3
TTEST RANKSURV*TRT

/* Calculate power of LSD test to detect true population differences */
/* of 10, 20, 30, 40, and 50% below mean (arcsine-transformed) reference */
/* sediment survival. */

DATA
LET M=20
LET N=5
LET MEANPCT=19.6/M
LET MSE=0.016
LET DF=16
LET TALPHA=TIF(.95,DF)
SAVE B2
PRINT M, N, MEANPCT, MSE, DF, TALPHA
RUN

DATA
USE B2
FOR I=10 TO 50 STEP 10
LET SEDSURV=MEANPCT-I/100
LET ARCSURV=ASN(SQR(SEDSURV))
LET ARCDIFF=1.48059-ARCSURV
LET TBETA=ARCDIFF*SQR(N/(2*MSE))-TALPHA
LET POWER=TCF(TBETA,DF)
PRINT I, SEDSURV, ARCSURV, ARCDIFF, TBETA, POWER
NEXT
STOP
RUN

```

## BENTOX.CMD program output

```

BENTHIC TOXICITY DATA

```

TRT	REP	M	SURV	ARCSURV
1.00000	1.00000	20.00000	20.00000	1.57080
1.00000	2.00000	20.00000	20.00000	1.57080
1.00000	3.00000	20.00000	19.00000	1.34528
1.00000	4.00000	20.00000	19.00000	1.34528
1.00000	5.00000	20.00000	20.00000	1.57080
2.00000	1.00000	20.00000	17.00000	1.17310
2.00000	2.00000	20.00000	16.00000	1.10715
2.00000	3.00000	20.00000	18.00000	1.24905
2.00000	4.00000	20.00000	17.00000	1.17310
2.00000	5.00000	20.00000	15.00000	1.04720
3.00000	1.00000	20.00000	15.00000	1.04720
3.00000	2.00000	20.00000	16.00000	1.10715
3.00000	3.00000	20.00000	13.00000	0.93774
3.00000	4.00000	20.00000	17.00000	1.17310
3.00000	5.00000	20.00000	11.00000	0.83548
4.00000	1.00000	20.00000	17.00000	1.17310
4.00000	2.00000	20.00000	12.00000	0.88608
4.00000	3.00000	20.00000	10.00000	0.78540
4.00000	4.00000	20.00000	16.00000	1.10715
4.00000	5.00000	20.00000	13.00000	0.93774

THE FOLLOWING RESULTS ARE FOR:  
TRT = 1.00000

	SURV	ARCSURV
N OF CASES	5	5
MEAN	19.60000	<u>1.48059</u>
VARIANCE	0.30000	<u>0.01526</u>

```

STD. ERROR      0.24495      0.05524
SUM             98.00000      7.40295

```

```

THE FOLLOWING RESULTS ARE FOR:
TRT = 2.00000

```

```

          SURV      ARCSURV
N OF CASES          5          5
MEAN             16.60000      1.14992
VARIANCE         1.30000      0.00582
STD. ERROR       0.50990      0.03412
SUM              83.00000      5.74959

```

```

THE FOLLOWING RESULTS ARE FOR:
TRT = 3.00000

```

```

          SURV      ARCSURV
N OF CASES          5          5
MEAN             14.40000      1.02013
VARIANCE         5.80000      0.01815
STD. ERROR       1.07703      0.06024
SUM              72.00000      5.10067

```

```

THE FOLLOWING RESULTS ARE FOR:
TRT = 4.00000

```

```

          SURV      ARCSURV
N OF CASES          5          5
MEAN             13.60000      0.97789
VARIANCE         8.30000      0.02548
STD. ERROR       1.28841      0.07138
SUM              68.00000      4.88947

```

---

```

/* Normality Test Results */

```

```

KOLMOGOROV-SMIRNOV ONE SAMPLE TEST USING STANDARD NORMAL DISTRIBUTION

```

VARIABLE	N-OF-CASES	MAXDIF	LILLIEFORS	PROBABILITY (2-TAIL)
ARCSURV	20.00000	0.17276	0.12031	

---

```

/* Perform Bartlett's homogeneity of variance Test and LSD Test */

```

```

SUMMARY STATISTICS FOR ARCSURV

```

```

BARTLETT TEST FOR HOMOGENEITY OF GROUP VARIANCES = 2.045

```

```

APPROXIMATE F = 0.616 DF = 3, 460 PROBABILITY = 0.605

```

```

ANALYSIS OF VARIANCE

```

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F	PROBABILITY
BETWEEN GROUPS	0.778	3	0.259	16.030	0.000
WITHIN GROUPS	0.259	16	0.016		

---

```

/* LSD using ARCSURV */

```

```

POST HOC TEST OF ARCSURV

```

```

USING MODEL MSE OF .016 WITH 16. DF.

```

```

FISHER'S LEAST-SIGNIFICANT-DIFFERENCE TEST.
MATRIX OF PAIRWISE COMPARISON PROBABILITIES:

```

	1	2	3	4
1	1.000			
2	0.001	1.000		
3	0.000	0.126	1.000	
4	0.000	0.048	0.607	1.000

```

/* If the probability associated with corresponding treatments is less than 0.05, we
conclude that the means are significantly different. For instance, the probability
associated with the Reference and Sediment 1 is 0.001. Hence, we conclude the Refer-
ence and Sediment 1 means are significantly different. Also, means from Sediment 2
and Sediment 3 are significantly different from the Reference mean; Sediment 1 mean
is not different from the Sediment 2 mean but it is different from the Sediment 3
mean; Sediment 2 mean is different from the Sediment 3 mean. */

```

---

```

/* t-test comparing the Reference and Sediment 1 */

```

INDEPENDENT SAMPLES T-TEST ON ARCSURV GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	1.481	0.124
2.000	5	1.150	0.076

SEPARATE VARIANCES T = 5.093 DF = 6.7 PROB = 0.002  
 POOLED VARIANCES T = 5.093 DF = 8 PROB = 0.001

DATA VERSION 5.02

F'	DF num	DF den	Prob >F'
2.622	4.000	4.000	0.373

/\* t-test comparing the Reference to Sediment 2 \*/

INDEPENDENT SAMPLES T-TEST ON ARCSURV GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	1.481	0.124
3.000	5	1.020	0.135

SEPARATE VARIANCES T = 5.634 DF = 7.9 PROB = 0.001  
 POOLED VARIANCES T = 5.634 DF = 8 PROB = 0.000

F'	DF num	DF den	Prob>F'
1.189	4.000	4.000	0.871

/\* t-test comparing the Reference to Sediment 3 \*/

INDEPENDENT SAMPLES T-TEST ON ARCSURV GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	1.481	0.124
4.000	5	0.978	0.160

SEPARATE VARIANCES T = 5.569 DF = 7.5 PROB = 0.001  
 POOLED VARIANCES T = 5.569 DF = 8 PROB = 0.001

F'	DF num	DF den	Prob>F'
1.67	4.000	4.000	0.632

THE FOLLOWING RESULTS ARE FOR:  
 TRT = 1.00000

RANKSURV

N OF CASES	5
MEAN	18.00000
VARIANCE	1.87500

THE FOLLOWING RESULTS ARE FOR:  
 TRT = 2.00000

RANKSURV

N OF CASES	5
MEAN	11.10000
VARIANCE	11.17500

THE FOLLOWING RESULTS ARE FOR:  
 TRT = 3.00000

RANKSURV

N OF CASES	5
MEAN	6.90000
VARIANCE	16.42500

THE FOLLOWING RESULTS ARE FOR:  
 TRT = 4.00000

RANKSURV

N OF CASES	5
MEAN	6.00000
VARIANCE	21.87500

/\* Levene's Test \*/

DEP VAR: ABSDEV N: 20 MULTIPLE R: 0.545 SQUARED MULTIPLE R: 0.297

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
TRT	18.02200	3	6.00733	2.25331	0.12148
ERROR	42.65600	16	2.66600		

/\* Conover T-Test \*/

DEP VAR:RANKSURV N: 20 MULTIPLE R: 0.828 SQUARED MULTIPLE R: 0.686

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
TRT	449.10000	3	149.70000	11.66115	0.00027
ERROR	205.40000	16	12.83750		

POST HOC TEST OF RANKSURV

USING MODEL MSE OF 12.837 WITH 16. DF.  
MATRIX OF PAIRWISE MEAN DIFFERENCES:

	1	2	3	4
1	0.00000			
2	-6.90000	0.00000		
3	-11.10000	-4.20000	0.00000	
4	-12.00000	-5.10000	-0.90000	0.00000

FISHER'S LEAST-SIGNIFICANT-DIFFERENCE TEST.  
MATRIX OF PAIRWISE COMPARISON PROBABILITIES:

	1	2	3	4
1	1.00000			
2	0.00772	1.00000		
3	0.00016	0.08235	1.00000	
4	0.00007	0.03883	0.69649	1.00000

/\* t-test using ranks \*/

THE FOLLOWING RESULTS ARE FOR:

TRT = 1.00000

RANKSURV

N OF CASES 5  
MEAN 8.00000  
VARIANCE 1.87500

THE FOLLOWING RESULTS ARE FOR:

TRT = 2.00000

RANKSURV

N OF CASES 5  
MEAN 3.00000  
VARIANCE 2.37500

INDEPENDENT SAMPLES T-TEST ON RANKSURV GROUPED BY TRT

GROUP	N	MEAN	SD
1.00000	5	8.00000	1.36931
2.00000	5	3.00000	1.54110

SEPARATE VARIANCES T = 5.42326 DF = 7.9 PROB = 0.00066  
POOLED VARIANCES T = 5.42326 DF = 8 PROB = 0.00063

THE FOLLOWING RESULTS ARE FOR:

TRT = 3.00000

RANKSURV

N OF CASES 5  
MEAN 3.00000  
VARIANCE 2.50000

INDEPENDENT SAMPLES T-TEST ON RANKSURV GROUPED BY TRT

GROUP	N	MEAN	SD
1.00000	5	8.00000	1.36931
3.00000	5	3.00000	1.58114

SEPARATE VARIANCES T = 5.34522 DF = 7.8 PROB = 0.00074  
 POOLED VARIANCES T = 5.34522 DF = 8 PROB = 0.00069

INDEPENDENT SAMPLES T-TEST ON RANKSURV GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	8.000	1.369
4.000	5	3.000	1.581

SEPARATE VARIANCES T = 5.345 DF = 7.8 PROB = 0.001  
 POOLED VARIANCES T = 5.345 DF = 8 PROB = 0.001

BENTHIC TOXICITY DATA  
 POWER TO DETECT A TRUE POPULATION DIFFERENCE (D)  
 FROM MEAN REFERENCE SURVIVAL USING ARSINE TRANSFORMATION

# OF ORGANISMS AT START OF TEST	# OF REPLICATE	MEAN REFERENCE SURVIVAL	MEAN SQUARE ERROR	DEGREES OF FREEDOM	T-VALUE FOR (.95, DF)
20.00000	5.00000	0.98000	0.01600	16.00000	1.74588

% REDUCTION IN SURVIVAL FROM REFERENCE	DREDGED SEDIMENT SURVIVAL	ARSINE DREDGED SEDIMENT SURVIVAL	D	T-VALUE FOR (1-BETA, DF)	POWER
10.00000	0.88000	1.21705	0.26354	1.54831	0.92945
20.00000	0.78000	1.08259	0.39800	3.22910	0.99738
30.00000	0.68000	0.96953	0.51106	4.64234	0.99986
40.00000	0.58000	0.86574	0.61485	5.93970	0.99999
50.00000	0.48000	0.76539	0.71520	7.19408	1.00000

/\* Note: Rankits can be computed using the formula  
 $rankit = Z((rank - 0.375) / (N + 0.25))$ .

The calculation of rankits is illustrated using Benthic Toxicity Data. \*/

```
DATA
FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
SAVE BENTHIC
INPUT TRT REP SURV \
FORMAT=5
RUN
1 1 20 1 2 20 1 3 19 1 4 19 1 5 20
2 1 17 2 2 16 2 3 18 2 4 17 2 5 15
3 1 15 3 2 16 3 3 13 3 4 17 3 5 11
4 1 17 4 2 12 4 3 10 4 4 16 4 5 13
~
USE BENTHIC
LET M=20
LET ARCSURV=ASN(SQR(SURV/M))
LET RANKSURV=SURV
RANK RANKSURV
SAVE A7
RUN

DATA
USE A7
LET RANKIT=ZIP((RANKSURV-.375)/(20+.25))
PRINT TRT, REP, SURV, ARCSURV, RANKSURV, RANKIT
RUN
```

BENTHIC TOXICITY DATA

TRT	REPLICATE	NUMBER OF SURVIVORS	ARCSINE TRANSFORMATION	RANK OF SURVIVORS	NORMALIZED RANK FOR SURVIVAL
1.00000	1.00000	20.00000	1.57080	19.00000	1.40341
1.00000	2.00000	20.00000	1.57080	19.00000	1.40341
1.00000	3.00000	19.00000	1.34528	16.50000	0.82846
1.00000	4.00000	19.00000	1.34528	16.50000	0.82846
1.00000	5.00000	20.00000	1.57080	19.00000	1.40341
2.00000	1.00000	17.00000	1.17310	12.50000	0.25015
2.00000	2.00000	16.00000	1.10715	9.00000	-0.18676
2.00000	3.00000	18.00000	1.24905	15.00000	0.58946
2.00000	4.00000	17.00000	1.17310	12.50000	0.25015
2.00000	5.00000	15.00000	1.04720	6.50000	-0.51731
3.00000	1.00000	15.00000	1.04720	6.50000	-0.51731
3.00000	2.00000	16.00000	1.10715	9.00000	-0.18676
3.00000	3.00000	13.00000	0.93774	4.50000	-0.82846
3.00000	4.00000	17.00000	1.17310	12.50000	0.25015
3.00000	5.00000	11.00000	0.83548	2.00000	-1.40341
4.00000	1.00000	17.00000	1.17310	12.50000	0.25015
4.00000	2.00000	12.00000	0.88608	3.00000	-1.12814
4.00000	3.00000	10.00000	0.78540	1.00000	-1.86824

4.00000	4.00000	16.00000	1.10715	9.00000	-0.18676
4.00000	5.00000	13.00000	0.93774	4.50000	-0.82846

## Program BIOACC.COMD for Single-Time Point Bioaccumulation Test Data Analysis

BIOACC.COMD is a program to compare Tier III bioaccumulation data from dredged sediments versus reference sediment, using raw data and a  $\log_{10}$  transformation. Analyses include mean bioaccumulation for reference and test sediment(s), Kolmogorov-Smirnov test for normality using probabilities developed by Lilliefors,  $t$ -tests for equal or unequal variances, and Conover's  $T$ -Test. The test results are interpreted as described in Appendix D of the Inland Testing Manual.

### BIOACC.COMD program statements

```

DATA
FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
SAVE BIOACC
INPUT TRT REP CONC \
FORMAT=5
RUN
1 1 .06 1 2 .05 1 3 .05 1 4 .08 1 5 .09
2 1 .16 2 2 .19 2 3 .18 2 4 .22 2 5 .31
3 1 .24 3 2 .10 3 3 .13 3 4 .18 3 5 .30
4 1 .13 4 2 .05 4 3 .17 4 4 .08 4 5 .22
~
/* Treatment code 1=Reference, 2=Sediment 1, 3=Sediment 2, 4=Sediment 3 */

USE BIOACC
LET LOGCONC=LOG(CONC)/LOG(10)
SAVE A0
PRINT TRT, REP, CONC, LOGCONC
RUN
USE A0
SORT TRT
RUN
STATS
USE A0
BY TRT
STATISTICS CONC LOGCONC / SUM N MEAN VARIANCE SEM

/* Normality test */

NPAR
USE A0
KS CONC LOGCONC / LILLIEFORS

/* Bartlett's homogeneity of variance Test */

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
STATS
USE A0
BY TRT
PRINT=LONG
STATISTICS CONC

STATS
USE A0
BY TRT
PRINT=LONG
STATISTICS LOGCONC

/* LSD test on untransformed and log10-transformed data */

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
MGLH
USE A0
CATEGORY TRT
ANOVA CONC
ESTIMATE
HYPOTHESIS
POST TRT / LSD
TEST

```

```

MGLH
USE A0
CATEGORY TRT
ANOVA LOGCONC
ESTIMATE
HYPOTHESIS
POST TRT / LSD
TEST

/* t-test comparing reference and sediment 1 */

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET

DATA
USE A0
IF TRT>2 THEN DELETE
SAVE T1
RUN
STATS
USE T1
TTEST CONC*TRT
DATA
LET N0=5
LET N1=5
LET VAR0=0.00033
LET VAR1=0.00347
IF VAR0>=VAR1 THEN LET DFNUM=N0-1
IF VAR0>=VAR1 THEN LET DFDEN=N1-1
IF VAR0<VAR1 THEN LET DFNUM=N1-1
IF VAR0<VAR1 THEN LET DFDEN=N0-1
IF VAR0>=VAR1 THEN LET FMAX=((N0-1)*VAR0)/((N1-1)*VAR1)
IF VAR0<VAR1 THEN LET FMAX=((N1-1)*VAR1)/((N0-1)*VAR0)
LET PROBF=(1-FCF(FMAX,DFNUM,DFDEN))*2
PRINT FMAX, DFNUM, DFDEN, PROBF
RUN

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE A0
IF TRT>2 THEN DELETE
SAVE T1L
RUN
STATS
USE T1L
TTEST LOGCONC*TRT
DATA
LET N0=5
LET N1=5
LET VAR0=0.01377
LET VAR1=0.01226
IF VAR0>=VAR1 THEN LET DFNUM=N0-1
IF VAR0>=VAR1 THEN LET DFDEN=N1-1
IF VAR0<VAR1 THEN LET DFNUM=N1-1
IF VAR0<VAR1 THEN LET DFDEN=N0-1
IF VAR0>=VAR1 THEN LET FMAX=((N0-1)*VAR0)/((N1-1)*VAR1)
IF VAR0<VAR1 THEN LET FMAX=((N1-1)*VAR1)/((N0-1)*VAR0)
LET PROBF=(1-FCF(FMAX,DFNUM,DFDEN))*2
PRINT FMAX, DFNUM, DFDEN, PROBF
RUN

/* t-test comparing reference and sediment 2 */

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE A0
IF (TRT=2 OR TRT=4) THEN DELETE
SAVE T2
RUN
STATS
USE T2
TTEST CONC*TRT
DATA
LET N0=5
LET N1=5
LET VAR0=0.00033
LET VAR1=0.00660
IF VAR0>=VAR1 THEN LET DFNUM=N0-1
IF VAR0>=VAR1 THEN LET DFDEN=N1-1
IF VAR0<VAR1 THEN LET DFNUM=N1-1
IF VAR0<VAR1 THEN LET DFDEN=N0-1
IF VAR0>=VAR1 THEN LET FMAX=((N0-1)*VAR0)/((N1-1)*VAR1)
IF VAR0<VAR1 THEN LET FMAX=((N1-1)*VAR1)/((N0-1)*VAR0)
LET PROBF=(1-FCF(FMAX,DFNUM,DFDEN))*2
PRINT FMAX, DFNUM, DFDEN, PROBF
RUN

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE A0
IF (TRT=2 OR TRT=4) THEN DELETE
SAVE T2L

```

```

RUN
STATS
USE T2L
TTEST LOGCONC*TRT
DATA
LET N0=5
LET N1=5
LET VAR0=0.01377
LET VAR1=0.03738
IF VAR0>=VAR1 THEN LET DFNUM=N0-1
IF VAR0>=VAR1 THEN LET DFDEN=N1-1
IF VAR0<VAR1 THEN LET DFNUM=N1-1
IF VAR0<VAR1 THEN LET DFDEN=N0-1
IF VAR0>=VAR1 THEN LET FMAX=((N0-1)*VAR0)/((N1-1)*VAR1)
IF VAR0<VAR1 THEN LET FMAX=((N1-1)*VAR1)/((N0-1)*VAR0)
LET PROBF=(1-FCF(FMAX,DFNUM,DFDEN))*2
PRINT FMAX, DFNUM, DFDEN, PROBF
RUN

```

```
/* t-test comparing reference and sediment 3 */
```

```

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE A0
IF (TRT=2 OR TRT=3) THEN DELETE
SAVE T3
RUN
STATS
USE T3
TTEST CONC*TRT
DATA
LET N0=5
LET N1=5
LET VAR0=0.00033
LET VAR1=0.00465
IF VAR0>=VAR1 THEN LET DFNUM=N0-1
IF VAR0>=VAR1 THEN LET DFDEN=N1-1
IF VAR0<VAR1 THEN LET DFNUM=N1-1
IF VAR0<VAR1 THEN LET DFDEN=N0-1
IF VAR0>=VAR1 THEN LET FMAX=((N0-1)*VAR0)/((N1-1)*VAR1)
IF VAR0<VAR1 THEN LET FMAX=((N1-1)*VAR1)/((N0-1)*VAR0)
LET PROBF=(1-FCF(FMAX,DFNUM,DFDEN))*2
PRINT FMAX, DFNUM, DFDEN, PROBF
RUN

```

```

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE A0
IF (TRT=2 OR TRT=3) THEN DELETE
SAVE T3L
RUN
STATS
USE T3L
TTEST LOGCONC*TRT
DATA
LET N0=5
LET N1=5
LET VAR0=0.01377
LET VAR1=0.06666
IF VAR0>=VAR1 THEN LET DFNUM=N0-1
IF VAR0>=VAR1 THEN LET DFDEN=N1-1
IF VAR0<VAR1 THEN LET DFNUM=N1-1
IF VAR0<VAR1 THEN LET DFDEN=N0-1
IF VAR0>=VAR1 THEN LET FMAX=((N0-1)*VAR0)/((N1-1)*VAR1)
IF VAR0<VAR1 THEN LET FMAX=((N1-1)*VAR1)/((N0-1)*VAR0)
LET PROBF=(1-FCF(FMAX,DFNUM,DFDEN))*2
PRINT FMAX, DFNUM, DFDEN, PROBF
RUN

```

```
/* Calculate ranks */
```

```

DATA
USE RBIOACC
LET RANKCONC=CONC
RANK RANKCONC
SAVE RBIOACC
RUN

```

```

STATS
USE RBIOACC
BY TRT
STATISTICS RANKCONC / N MEAN VARIANCE

```

```

DATA
USE RBIOACC
IF TRT=1 THEN LET ABSDEV=ABS(RANKCONC-4.1)
IF TRT=2 THEN LET ABSDEV=ABS(RANKCONC-15.2)
IF TRT=3 THEN LET ABSDEV=ABS(RANKCONC-13.6)
IF TRT=4 THEN LET ABSDEV=ABS(RANKCONC-9.1)
SAVE BLEVENE
RUN

```

```

/* Levene's Test using ranks */
MGLH
USE BLEVENE
CATEGORY TRT
ANOVA ABSDEV
ESTIMATE

/* Conover T-Test */

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE BIOACC
LET RANKCONC=CONC
RANK RANKCONC
SAVE RRBIACC
RUN

MGLH
USE RRBIACC
CATEGORY TRT
ANOVA RANKCONC
ESTIMATE
HYPOTHESIS
POST TRT / LSD
TEST

/* t-tests using ranks */

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE BIOACC
IF TRT>2 THEN DELETE
SAVE BIOT1
RUN

DATA
USE BIOACC
IF (TRT=2 OR TRT=4) THEN DELETE
SAVE BIOT2
RUN

DATA
USE BIOACC
IF (TRT=2 OR TRT=3) THEN DELETE
SAVE BIOT3
RUN

/* Calculate ranks and t-test (reference and sediment 1) */

DATA
USE BIOT1
LET RANKCONC=CONC
RANK RANKCONC
SAVE RRBIOT1
RUN

STATS
USE RRBIOT1
TTEST RANKCONC*TRT

/* Calculate ranks and t-test (reference and sediment 2) */

DATA
USE BIOT2
LET RANKCONC=CONC
RANK RANKCONC
SAVE RRBIOT2
RUN

STATS
USE RRBIOT2
TTEST RANKCONC*TRT

/* Calculate ranks and t-test (reference and sediment 3) */

DATA
USE BIOT3
LET RANKCONC=CONC
RANK RANKCONC
SAVE RRBIOT3
RUN

STATS
USE RRBIOT3
TTEST RANKCONC*TRT

/* Calculate power of LSD test to detect true population differences
10, 25, 50, and 100% above the reference mean contaminant concentration. */

```

```

REPEAT 1
DATA
LET N=5
LET MEANCONC=0.066
LET SS=0.06020
LET DF=16
LET MSE=SS/DF
LET TALPHA=TIF(.95,DF)
SAVE BIO2
PRINT N, MEANCONC, MSE, DF, TALPHA
RUN

DATA
REPEAT 1
USE BIO2
5 LET I=10
10 LET SIZE=10
15 LET SEDCONC=MEANCONC+((I/100)*MEANCONC)
25 LET D=SEDCONC-MEANCONC
30 LET TBETA=D*SQR(N/(2*MSE))-TALPHA
35 LET POWER=TCF(TBETA,DF)
38 PRINT I, SEDCONC, D, TBETA, POWER
40 IF I=10 THEN LET I=SIZE+15
42 IF I=25 AND SIZE=10 THEN GOTO 59
45 IF I=25 THEN LET I=SIZE+25
47 IF I=50 AND SIZE=25 THEN GOTO 59
49 IF I=50 THEN LET I=SIZE+50
50 IF I=100 AND SIZE=50 THEN GOTO 59
51 IF I=100 THEN LET I=SIZE+100
52 IF I=200 AND SIZE=100 THEN GOTO 59
53 IF I=200 THEN LET I=SIZE+100
55 IF I=300 AND SIZE=200 THEN GOTO 59
57 IF I=300 AND SIZE=300 THEN GOTO 70
59 LET SIZE=I
60 GOTO 15
70 STOP
RUN

REPEAT 1
DATA
USE BIO2
110 LET POWER=.5
112 LET SIZE=.5
115 LET TBETA=TIF(POWER,DF)
120 LET D=((TBETA+TALPHA)*SQR(2*MSE))/SQR(N)
125 LET SEDCONC=MEANCONC+D
130 LET PCTDIFF=(D*100)/MEANCONC
135 PRINT POWER, D, SEDCONC, PCTDIFF, TBETA
140 IF POWER<.9 THEN LET POWER=SIZE+.1
142 IF POWER=.9 AND SIZE=.8 THEN GOTO 155
145 IF POWER=.9 THEN LET POWER=SIZE+.05
147 IF POWER=.95 AND SIZE=.9 THEN GOTO 155
150 IF POWER=.95 THEN LET POWER=SIZE+.04
152 IF POWER=.99 AND SIZE=.99 THEN GOTO 170
155 LET SIZE=POWER
160 GOTO 115
170 STOP
RUN

/* Calculation of upper confidence limits (UCL) for comparison of mean */
/* dredged sediment bioaccumulation with an action level. */

DATA
FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
SAVE POWER
INPUT TRT MEANCONC S2
FORMAT=5
RUN
2 .212 .00347
3 .190 .0066
4 .130 .00465
~
DATA
LET N=5
LET SS=0.06020
LET DF=16
LET MSE=SS/DF
LET TALPHA1=TIF(.95,DF)
LET TALPHA2=TIF(.95,N-1)
LET UCL1=MEANCONC+TALPHA1*(SQR(MSE/N))
LET UCL2=MEANCONC+TALPHA2*(SQR(S2/N))
LET DMIN1=TALPHA1*SQR(MSE/N)
LET DMIN2=TALPHA2*SQR(S2/N)
SAVE BIO3
PRINT MEANCONC, UCL1, MSE, TALPHA1, DF, DMIN1
PRINT MEANCONC, UCL2, S2, TALPHA2, N, DMIN2
RUN

/* Calculate power of dredged sediment-action level comparisons using */
/* MSE given 10, 20, 30, 40, and 50% decreases in mean concentration. */

```

```

DATA
USE BIO2
LET ACTION=.2
FOR PCTDIFF=10 TO 50 STEP 10
LET D=PCTDIFF*ACTION/100
LET SEDCONC=ACTION-D
LET TBETA=D*SQR(N/MSE)-TALPHA
LET POWER=TCF(TBETA,DF)
PRINT PCTDIFF, SEDCONC, D, TBETA, POWER
NEXT
STOP
RUN

```

## BIOACC.CMD program output

### SINGLE TIME-POINT CONTAMINANT BIOACCUMULATION DATA

TRT	REPLICATE	CONC	LOGCONC
1.00000	1.00000	0.06000	-1.22185
1.00000	2.00000	0.05000	-1.30103
1.00000	3.00000	0.05000	-1.30103
1.00000	4.00000	0.08000	-1.09691
1.00000	5.00000	0.09000	-1.04576
2.00000	1.00000	0.16000	-0.79588
2.00000	2.00000	0.19000	-0.72125
2.00000	3.00000	0.18000	-0.74473
2.00000	4.00000	0.22000	-0.65758
2.00000	5.00000	0.31000	-0.50864
3.00000	1.00000	0.24000	-0.61979
3.00000	2.00000	0.10000	-1.00000
3.00000	3.00000	0.13000	-0.88606
3.00000	4.00000	0.18000	-0.74473
3.00000	5.00000	0.30000	-0.52288
4.00000	1.00000	0.13000	-0.88606
4.00000	2.00000	0.05000	-1.30103
4.00000	3.00000	0.17000	-0.76955
4.00000	4.00000	0.08000	-1.09691
4.00000	5.00000	0.22000	-0.65758

THE FOLLOWING RESULTS ARE FOR:

TRT = 1.00000

	CONC	LOGCONC
N OF CASES	5	5
MEAN	0.06600	-1.19332
VARIANCE	0.00033	0.01377
STD. ERROR	0.00812	0.05248
SUM	0.33000	-5.96658

THE FOLLOWING RESULTS ARE FOR:

TRT = 2.00000

	CONC	LOGCONC
N OF CASES	5	5
MEAN	0.21200	-0.68561
VARIANCE	0.00347	0.01226
STD. ERROR	0.02634	0.04951
SUM	1.06000	-3.42807

THE FOLLOWING RESULTS ARE FOR:

TRT = 3.00000

	CONC	LOGCONC
N OF CASES	5	5
MEAN	0.19000	-0.75469
VARIANCE	0.00660	0.03737
STD. ERROR	0.03633	0.08645
SUM	0.95000	-3.77345

THE FOLLOWING RESULTS ARE FOR:

TRT = 4.00000

	CONC	LOGCONC
N OF CASES	5	5
MEAN	0.13000	-0.94223
VARIANCE	0.00465	0.06667
STD. ERROR	0.03050	0.11547
SUM	0.65000	-4.71113

---

/\* Normality Test Results \*/

KOLMOGOROV-SMIRNOV ONE SAMPLE TEST USING STANDARD NORMAL DISTRIBUTION

VARIABLE	N-OF-CASES	MAXDIF	LILLIEFORS	PROBABILITY (2-TAIL)
CONC	20.00000	0.12941	0.52466	
LOGCONC	20.00000	0.14648	0.31559	

SUMMARY STATISTICS FOR CONC

BARTLETT TEST FOR HOMOGENEITY OF GROUP VARIANCES = 6.964  
 APPROXIMATE F = 2.118 DF = 3, 460 PROBABILITY = 0.097

ANALYSIS OF VARIANCE

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F	PROBABILITY
BETWEEN GROUPS	0.064	3	0.021	5.714	0.007
WITHIN GROUPS	0.060	16	0.004		

SUMMARY STATISTICS FOR LOGCONC

BARTLETT TEST FOR HOMOGENEITY OF GROUP VARIANCES = 3.911  
 APPROXIMATE F = 1.182 DF = 3, 460 PROBABILITY = 0.316

ANALYSIS OF VARIANCE

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F	PROBABILITY
BETWEEN GROUPS	0.774	3	0.258	7.932	0.002
WITHIN GROUPS	0.520	16	0.033		

DEP VAR: CONC N: 20 MULTIPLE R: 0.719 SQUARED MULTIPLE R: 0.517

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
TRT	0.064	3	0.021	5.714	0.007
ERROR	0.060	16	0.004		

POST HOC TEST OF CONC

USING MODEL MSE OF .004 WITH 16. DF.

FISHER'S LEAST-SIGNIFICANT-DIFFERENCE TEST.  
 MATRIX OF PAIRWISE COMPARISON PROBABILITIES:

	1	2	3	4
1	1.000			
2	0.002	1.000		
3	0.006	0.579	1.000	
4	0.118	0.051	0.142	1.000

DEP VAR: LOGCONC N: 20 MULTIPLE R: 0.773 SQUARED MULTIPLE R: 0.598

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
TRT	0.774	3	0.258	7.932	0.002
ERROR	0.520	16	0.033		

POST HOC TEST OF LOGCONC

USING MODEL MSE OF .033 WITH 16. DF.

MATRIX OF PAIRWISE MEAN DIFFERENCES:

FISHER'S LEAST-SIGNIFICANT-DIFFERENCE TEST.  
 MATRIX OF PAIRWISE COMPARISON PROBABILITIES:

TRT	1	2	TRT	3	4
1	1.000				
2	0.000	1.000			
3	0.001	0.553	1.000		
4	0.043	0.039	0.120	1.000	

INDEPENDENT SAMPLES T-TEST ON CONC GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	0.066	0.018
2.000	5	0.212	0.059

SEPARATE VARIANCES T = -5.296 DF = 4.8 PROB = 0.004  
 POOLED VARIANCES T = -5.296 DF = 8 PROB = 0.001

F'	DFnum	DFden	PROB>F'
10.515	4.000	4.000	0.043

INDEPENDENT SAMPLES T-TEST ON LOGCONC GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	-1.193	0.117
2.000	5	-0.686	0.111

SEPARATE VARIANCES T = -7.037 DF = 8.0 PROB = 0.000  
 POOLED VARIANCES T = -7.037 DF = 8 PROB = 0.000

F'	DFnum	DFden	PROB>F'
1.123	4.000	4.000	0.913

INDEPENDENT SAMPLES T-TEST ON CONC GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	0.066	0.018
3.000	5	0.190	0.081

SEPARATE VARIANCES T = -3.331 DF = 4.4 PROB = 0.025  
 POOLED VARIANCES T = -3.331 DF = 8 PROB = 0.010

F'	DFnum	DFden	PROB>F'
20.000	4.000	4.000	0.013

INDEPENDENT SAMPLES T-TEST ON LOGCONC GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	-1.193	0.117
3.000	5	-0.755	0.193

SEPARATE VARIANCES T = -4.337 DF = 6.6 PROB = 0.004  
 POOLED VARIANCES T = -4.337 DF = 8 PROB = 0.002

F'	DFnum	DFden	PROB>F'
2.715	4.000	4.000	0.357

INDEPENDENT SAMPLES T-TEST ON CONC GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	0.066	0.018
4.000	5	0.130	0.068

SEPARATE VARIANCES T = -2.028 DF = 4.6 PROB = 0.104  
 POOLED VARIANCES T = -2.028 DF = 8 PROB = 0.077

F'	DFnum	DFden	PROB>F'
14.091	4.000	4.000	0.025

INDEPENDENT SAMPLES T-TEST ON LOGCONC GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	-1.193	0.117
4.000	5	-0.942	0.258

SEPARATE VARIANCES T = -1.980 DF = 5.6 PROB = 0.099  
 POOLED VARIANCES T = -1.980 DF = 8 PROB = 0.083

F'	DFnum	DFden	PROB>F'
4.841	4.000	4.000	0.156

THE FOLLOWING RESULTS ARE FOR:  
TRT = 1.00000

RANKCONC

N OF CASES 5  
MEAN 4.10000  
VARIANCE 4.80000

THE FOLLOWING RESULTS ARE FOR:  
TRT = 2.00000

RANKCONC

N OF CASES 5  
MEAN 15.20000  
VARIANCE 11.32500

THE FOLLOWING RESULTS ARE FOR:  
TRT = 3.00000

RANKCONC

N OF CASES 5  
MEAN 13.60000  
VARIANCE 24.17500

THE FOLLOWING RESULTS ARE FOR:  
TRT = 4.00000

RANKCONC

N OF CASES 5  
MEAN 9.10000  
VARIANCE 31.67500

/\* Levene's Test (Ranks) \*/

DEP VAR: ABSDEV N: 20 MULTIPLE R: 0.477 SQUARED MULTIPLE R: 0.227

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
TRT	22.02200	3	7.34067	1.56776	0.23609
ERROR	74.91600	16	4.68225		

/\* Conover T-Test \*/

DEP VAR:RANKCONC N: 20 MULTIPLE R: 0.751 SQUARED MULTIPLE R: 0.564

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
TRT	373.100	3	124.367	6.912	0.003
ERROR	287.900	16	17.994		

POST HOC TEST OF RANKCONC

USING MODEL MSE OF 17.994 WITH 16. DF.  
MATRIX OF PAIRWISE MEAN DIFFERENCES:

	1	2	3	4
1	0.000			
2	11.100	0.000		
3	9.500	-1.600	0.000	
4	5.000	-6.100	-4.500	0.000

FISHER'S LEAST-SIGNIFICANT-DIFFERENCE TEST.  
MATRIX OF PAIRWISE COMPARISON PROBABILITIES:

	1	2	3	4
1	1.000			
2	0.001	1.000		
3	0.003	0.559	1.000	
4	0.081	0.037	0.113	1.000

/\* t-test using ranks \*/

INDEPENDENT SAMPLES T-TEST ON RANKCONC      GROUPED BY      TRT

GROUP	N	MEAN	SD
1.000	5	3.000	1.541
2.000	5	8.000	1.581

SEPARATE VARIANCES T =      -5.064 DF =      8.0 PROB =      0.001  
 POOLED VARIANCES T =      -5.064 DF =      8 PROB =      0.001

INDEPENDENT SAMPLES T-TEST ON RANKCONC      GROUPED BY      TRT

GROUP	N	MEAN	SD
1.000	5	3.000	1.541
3.000	5	8.000	1.581

SEPARATE VARIANCES T =      -5.064 DF =      8.0 PROB =      0.001  
 POOLED VARIANCES T =      -5.064 DF =      8 PROB =      0.001

INDEPENDENT SAMPLES T-TEST ON RANKCONC      GROUPED BY      TRT

GROUP	N	MEAN	SD
1.000	5	4.100	2.191
4.000	5	6.900	3.209

SEPARATE VARIANCES T =      -1.611 DF =      7.1 PROB =      0.151  
 POOLED VARIANCES T =      -1.611 DF =      8 PROB =      0.146

SINGLE-TIME POINT CONTAMINANT BIOACCUMULATION DATA  
 POWER OF LSD TO DETECT A TRUE DIFFERENCE (D)  
 ABOVE REFERENCE MEAN CONTAMINANT CONCENTRATION

# OF REPLICATES	REFERENCE MEAN CONTAMINANT CONCENTRATION	MEAN SQUARE ERROR	DEGREES OF FREEDOM	T VALUE FOR (1-ALPHA=0.95, DF)
5.00000	0.06600	0.00376	16.00000	1.74588

POWER OF LSD TO DETECT % INCREASE IN CONCENTRATION ABOVE REFERENCE MEAN CONTAMINANT CONCENTRATION GIVEN N, MSE AND DF SHOWN ABOVE

% INCREASE IN CONC. ABOVE REFERENCE	DREDGED SEDIMENT BIOACCUMULATION	D	T VALUE FOR (1-BETA,DF)	POWER (1-BETA)
10.00000	0.07260	0.00660	-1.57577	0.06732
25.00000	0.08250	0.01650	-1.32060	0.10261
50.00000	0.09900	0.03300	-0.89531	0.19195
100.00000	0.13200	0.06600	-0.04475	0.48243
200.00000	0.19800	0.13200	1.65639	0.94144
300.00000	0.26400	0.19800	3.35753	0.99800

MINIMUM DREDGED SEDIMENT BIOACCUMULATION THAT CAN BE DETECTED BY LSD AS SIGNIFICANT GIVEN SPECIFIED POWER, N, MSE, AND DF SHOWN ABOVE

POWER (1-BETA)	D	DREDGED SEDIMENT BIOACCUMULATION	% INCREASE IN CONC. ABOVE REF.	T VALUE FOR (1-BETA, DF)
0.50000	0.06777	0.13377	102.68773	0.00098
0.60000	0.07773	0.14373	117.77307	0.25760
0.70000	0.08849	0.15449	134.08046	0.53501
0.80000	0.10128	0.16728	153.45905	0.86467
0.90000	0.11960	0.18560	181.21048	1.33676
0.95000	0.13547	0.20147	205.26064	1.74588
0.99000	0.16797	0.23397	254.49847	2.58349

COMPARISON OF MEAN DREDGED SEDIMENT BIOACCUMULATION WITH ACTION LEVEL: UPPER CONFIDENCE LIMITS (UCL) WHEN VARIANCES ARE EQUAL

MEAN BIOACCUMULATION	UCL (EQUAL VARIANCES)	MEAN SQUARE ERROR	T VALUE FOR (1-ALPHA=.95,DF)	DF	MINIMUM SIGNIFICANT DIFFERENCE
0.21200	0.25990	0.00376	1.74588	16.0	0.04790
0.19000	0.23790	0.00376	1.74588	16.0	0.04790
0.13000	0.17790	0.00376	1.74588	16.0	0.04790

COMPARISON OF MEAN DREDGED SEDIMENT BIOACCUMULATION WITH ACTION LEVEL: UPPER CONFIDENCE LIMITS (UCL) WHEN VARIANCES ARE UNEQUAL

MEAN BIOACCUMULATION	UCL (UNEQUAL VARIANCES)	VARIANCE	T VALUE FOR (1-ALPHA=.95,N-1)	DF	MINIMUM SIGNIFICANT DIFFERENCE
0.21200	0.26816	0.00347	2.13185	5.0	0.05616
0.19000	0.26745	0.00660	2.13185	5.0	0.07745
0.13000	0.19501	0.00465	2.13185	5.0	0.06501

POWER TO DETECT % DECREASE IN CONCENTRATION BELOW ACTION LEVEL OF .2 ug/g GIVEN N, MSE, AND DF SHOWN ABOVE

% DECREASE BELOW ACTION LEVEL	MEAN DREDGED SEDIMENT BIOACCUMULATION	D	T VALUE FOR (1-BETA, DF)	POWER (1-BETA)
10.00000	0.18000	0.020	-1.01686	0.16218
20.00000	0.16000	0.040	-0.28784	0.38858
30.00000	0.14000	0.060	0.44118	0.66751
40.00000	0.12000	0.080	1.17020	0.87047
50.00000	0.10000	0.100	1.89923	0.96214

## Program BIOACSS.CMD for Time-Sequenced Bioaccumulation Test Data Analysis

This program is designed to compare Tier IV estimated steady-state bioaccumulation data from dredged sediments versus reference sediment, using raw data and a  $\log_{10}$  transformation. Analyses include mean bioaccumulation from each sediment exposure, Kolmogorov-Smirnov test for normality, Levene's and Bartlett's tests for homogeneity of variance, *t*-tests for equal or unequal variances, LSD test, and Conover *T*-Test. The test results are interpreted as described in Appendix D of the Inland Testing Manual. The SYSTAT NONLIN algorithm has two options, Quasi-Newton and Simplex. Quasi-Newton requires numerical estimates of the first and second derivatives. Simplex cannot make use of the information in the second derivative (Wilkinson 1990a). SYGRAPH (Wilkinson 1990b) may be used to produce graphs of time-sequenced contaminant bioaccumulation (statements necessary to produce graphs are provided in the following program, but the graphic output is not included in this appendix). The program includes power calculations for an LSD test on untransformed  $C_{ss}$  estimates.

The user may find it convenient to divide this program into several segments, which can be executed independently. The original program was constructed in that manner. The first two statements in each independent segment are shown below:

```
DATA
FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
```

### BIOACSS.CMD program statements

```
DATA
FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
SAVE BIOACSS
INPUT DAY REP TRT CONC \
FORMAT=5
RUN
2 1 1 .054 2 2 1 .163 2 3 1 .391 2 4 1 .234 2 5 1 .034
2 1 2 .159 2 2 2 .292 2 3 2 .428 2 4 2 .558 2 5 2 .256
2 1 3 .869 2 2 3 .726 2 3 3 .394 2 4 3 1.232 2 5 3 .977
2 1 4 .745 2 2 4 1.703 2 3 4 2.045 2 4 4 1.855 2 5 4 1.135
4 1 1 .441 4 2 1 .797 4 3 1 .203 4 4 1 .564 4 5 1 .018
4 1 2 .516 4 2 2 .158 4 3 2 .743 4 4 2 .324 4 5 2 .126
```

```

4 1 3 .838 4 2 3 .633 4 3 3 .452 4 4 3 .728 4 5 3 1.314
4 1 4 1.316 4 2 4 .930 4 3 4 2.141 4 4 4 1.150 4 5 4 1.621
7 1 1 .687 7 2 1 .177 7 3 1 .862 7 4 1 .413 7 5 1 .029
7 1 2 .881 7 2 2 .317 7 3 2 .270 7 4 2 .562 7 5 2 .603
7 1 3 1.246 7 2 3 .816 7 3 3 .897 7 4 3 1.639 7 5 3 .688
7 1 4 1.583 7 2 4 2.715 7 3 4 1.016 7 4 4 2.221 7 5 4 2.134
10 1 1 .037 10 2 1 .549 10 3 1 .884 10 4 1 .787 10 5 1 .294
10 1 2 .278 10 2 2 .485 10 3 2 .051 10 4 2 .909 10 5 2 .718
10 1 3 1.767 10 2 3 1.272 10 3 3 1.003 10 4 3 1.158 10 5 3 1.415
10 1 4 1.578 10 2 4 2.268 10 3 4 1.756 10 4 4 2.899 10 5 4 .890
18 1 1 .856 18 2 1 .598 18 3 1 .016 18 4 1 .806 18 5 1 .119
18 1 2 .904 18 2 2 1.300 18 3 2 .671 18 4 2 .934 18 5 2 1.173
18 1 3 1.631 18 2 3 1.877 18 3 3 1.487 18 4 3 1.216 18 5 3 1.280
18 1 4 2.822 18 2 4 2.607 18 3 4 3.414 18 4 4 1.319 18 5 4 1.866
28 1 1 .514 28 2 1 .839 28 3 1 .793 28 4 1 .899 28 5 1 .226
28 1 2 .172 28 2 2 1.049 28 3 2 .476 28 4 2 .712 28 5 2 1.245
28 1 3 1.178 28 2 3 1.721 28 3 3 1.366 28 4 3 1.513 28 5 3 1.843
28 1 4 1.295 28 2 4 2.964 28 3 4 2.109 28 4 4 2.820 28 5 4 3.325
~
/* Treatment code 1=Reference, 2=Sediment 1, 3=Sediment 2, 4=Sediment 3 */

USE BIOACCS
  IF TRT=1 THEN LET CS=.45
  IF TRT=2 THEN LET CS=4
  IF TRT=3 THEN LET CS=33
  IF TRT=4 THEN LET CS=44
SAVE AA
SORT TRT, REP, DAY
RUN

USE AA
PRINT TRT, REP, DAY, CONC, CS
RUN

SYGRAPH
OUTPUT @
USE AA
BY TRT
PLOT CONC*DAY / SYMBOL=REP, YMAX=4,
YLABEL ='CONC IN TISSUE',
TITLE='TIME-SEQUENCED BIOACCUMULATION'
OUTPUT *

/* Note: OUTPUT @ directs the graphic output to the printer and OUTPUT * directs
subsequent output to the screen. */

/* Fit nonlinear model */

NONLIN
USE AA
BY TRT, REP
MODEL CONC=CS*((K1=>0 AND K1<=3)/(K2=>.01 AND K2<=2)),
* (1-EXP(-(K2=>.01 AND K2<=2)*DAY))
SAVE REGPARMS
ESTIMATE /INTER=50,SIMPLEX,START=.07,.11,TOL=1E-15,PRINT

/* Input values of k1 and k2 */

DATA
FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
INPUT TRT REP K1 K2 \
SAVE PARMS
FORMAT=5
RUN
1 1 .07 .165 1 2 .07 .2475 1 3 .07 .37125 1 4 .07 .55688 1 5 .07 .83531
2 1 .05 .21 2 2 .05 .315 2 3 .05 .4725 2 4 .05 .70875 2 5 .05 1.06313
3 1 .04 .18 3 2 .04 .27 3 3 .04 .405 3 4 .04 .6075 3 5 .04 .91125
4 1 .09 .345 4 2 .09 .5175 4 3 .09 .77625 4 4 .09 1.16438 4 5 .09 1.45547
~
USE PARMS
SORT TRT, REP
SAVE PARMS2
RUN

USE AA
IF DAY<28 THEN DELETE
SAVE A
RUN

USE A PARMS2/TRT, REP
SAVE APARMS
RUN

/* Calculate and print Css and log-transformed Css. */

USE APARMS
LET CSS=CS*K1/K2
LET LOGCSS=LOG(CSS)/LOG(10)
DROP DAY CONC
SAVE APCS

```

```

PRINT TRT, REP, K1, K2, CSS, LOGCSS
RUN

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE ACSS
DROP K1 K2
SAVE ACSS2
RUN

STATS
USE ACSS2
BY TRT
STATISTICS / MEAN VARIANCE SEM

/* Normality test on untransformed and log10-transformed data */

NPAR
USE ACSS
KS CSS LOGCSS / LILLIEFORS

/* Bartlett's Test on untransformed data */

STATS
USE ACSS
BY TRT
PRINT=LONG
STATISTICS CSS

/* Bartlett's Test on log10-transformed data */

STATS
USE ACSS
BY TRT
PRINT=LONG
STATISTICS LOGCSS

/* LSD test on untransformed and log-transformed Css */

MGLH
USE ACSS
CATEGORY TRT
ANOVA CSS
ESTIMATE
HYPOTHESIS
POST TRT / LSD
TEST

MGLH
USE ACSS
CATEGORY TRT
ANOVA LOGCSS
ESTIMATE
HYPOTHESIS
POST TRT / LSD
TEST

/* Perform t-tests for each dredged sediment-reference sediment comparison */
/* using untransformed and log-transformed Css. */

/* t-test comparing the Reference and Sediment 1 */

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE ACSS
IF TRT>2 THEN DELETE
SAVE TCSS1
RUN
STATS
USE TCSS1
TTEST CSS*TRT
DATA
LET N0=5
LET N1=5
LET VAR0=0.003721
LET VAR1=0.093636
IF VAR0>=VAR1 THEN LET DFNUM=N0-1
IF VAR0>=VAR1 THEN LET DFDEN=N1-1
IF VAR0<VAR1 THEN LET DFNUM=N1-1
IF VAR0<VAR1 THEN LET DFDEN=N0-1
IF VAR0>=VAR1 THEN LET FMAX=((N0-1)*VAR0)/((N1-1)*VAR1)
IF VAR0<VAR1 THEN LET FMAX=((N1-1)*VAR1)/((N0-1)*VAR0)
LET PROBF=(1-FCF(FMAX,DFNUM,DFDEN))*2
PRINT FMAX, DFNUM, DFDEN, PROBF
RUN

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET

DATA
USE ACSS
IF TRT>2 THEN DELETE

```

```

SAVE TCSS1L
RUN
STATS
USE TCSS1L
TTEST LOGCSS*TRT
DATA
LET N0=5
LET N1=5
LET VAR0=0.077284
LET VAR1=0.077284
IF VAR0>=VAR1 THEN LET DFNUM=N0-1
IF VAR0>=VAR1 THEN LET DFDEN=N1-1
IF VAR0<VAR1 THEN LET DFNUM=N1-1
IF VAR0<VAR1 THEN LET DFDEN=N0-1
IF VAR0>=VAR1 THEN LET FMAX=((N0-1)*VAR0)/((N1-1)*VAR1)
IF VAR0<VAR1 THEN LET FMAX=((N1-1)*VAR1)/((N0-1)*VAR0)
LET PROBF=(1-FCF(FMAX,DFNUM,DFDEN))*2
PRINT FMAX, DFNUM, DFDEN, PROBF
RUN

/* t-test comparing the Reference and Sediment 2 */

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE ACSS
IF (TRT=2 OR TRT=4) THEN DELETE
SAVE TCSS2
RUN
STATS
USE TCSS2
TTEST CSS*TRT
DATA
LET N0=5
LET N1=5
LET VAR0=0.003721
LET VAR1=5.53661
IF VAR0>=VAR1 THEN LET DFNUM=N0-1
IF VAR0>=VAR1 THEN LET DFDEN=N1-1
IF VAR0<VAR1 THEN LET DFNUM=N1-1
IF VAR0<VAR1 THEN LET DFDEN=N0-1
IF VAR0>=VAR1 THEN LET FMAX=((N0-1)*VAR0)/((N1-1)*VAR1)
IF VAR0<VAR1 THEN LET FMAX=((N1-1)*VAR1)/((N0-1)*VAR0)
LET PROBF=(1-FCF(FMAX,DFNUM,DFDEN))*2
PRINT FMAX, DFNUM, DFDEN, PROBF
RUN

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE ACSS
IF (TRT=2 OR TRT=4) THEN DELETE
SAVE TCSS2L
RUN
STATS
USE TCSS2L
TTEST LOGCSS*TRT
DATA
LET N0=5
LET N1=5
LET VAR0=0.077284
LET VAR1=0.077284
IF VAR0>=VAR1 THEN LET DFNUM=N0-1
IF VAR0>=VAR1 THEN LET DFDEN=N1-1
IF VAR0<VAR1 THEN LET DFNUM=N1-1
IF VAR0<VAR1 THEN LET DFDEN=N0-1
IF VAR0>=VAR1 THEN LET FMAX=((N0-1)*VAR0)/((N1-1)*VAR1)
IF VAR0<VAR1 THEN LET FMAX=((N1-1)*VAR1)/((N0-1)*VAR0)
LET PROBF=(1-FCF(FMAX,DFNUM,DFDEN))*2
PRINT FMAX, DFNUM, DFDEN, PROBF
RUN

/* t-test comparing the Reference and Sediment 3 */

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE ACSS
IF (TRT=2 OR TRT=3) THEN DELETE
SAVE TCSS3
RUN
STATS
USE TCSS3
TTEST CSS*TRT
DATA
LET N0=5
LET N1=5
LET VAR0=0.003721
LET VAR1=12.75918
IF VAR0>=VAR1 THEN LET DFNUM=N0-1
IF VAR0>=VAR1 THEN LET DFDEN=N1-1
IF VAR0<VAR1 THEN LET DFNUM=N1-1
IF VAR0<VAR1 THEN LET DFDEN=N0-1
IF VAR0>=VAR1 THEN LET FMAX=((N0-1)*VAR0)/((N1-1)*VAR1)

```

```

IF VAR0<VAR1 THEN LET FMAX=((N1-1)*VAR1)/((N0-1)*VAR0)
LET PROBF=(1-FCF(FMAX,DFNUM,DFDEN))*2
PRINT FMAX, DFNUM, DFDEN, PROBF
RUN

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE ACSS
IF (TRT=2 OR TRT=3) THEN DELETE
SAVE TCSS3L
RUN
STATS
USE TCSS3L
TTEST LOGCSS*TRT
DATA
LET N0=5
LET N1=5
LET VAR0=0.077284
LET VAR1=0.06503
IF VAR0>=VAR1 THEN LET DFNUM=N0-1
IF VAR0>=VAR1 THEN LET DFDEN=N1-1
IF VAR0<VAR1 THEN LET DFNUM=N1-1
IF VAR0<VAR1 THEN LET DFDEN=N0-1
IF VAR0>=VAR1 THEN LET FMAX=((N0-1)*VAR0)/((N1-1)*VAR1)
IF VAR0<VAR1 THEN LET FMAX=((N1-1)*VAR1)/((N0-1)*VAR0)
LET PROBF=(1-FCF(FMAX,DFNUM,DFDEN))*2
PRINT FMAX, DFNUM, DFDEN, PROBF
RUN

/* Css converted to ranks */

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET

DATA
USE ACSS
LET RANKCSS=CSS
RANK RANKCSS
SAVE RACSS
RUN

STATS
USE RACSS
BY TRT
STATISTICS RANKCSS / N MEAN VARIANCE

DATA
USE RACSS
IF TRT=1 THEN LET ABSDEV=ABS(RANKCSS-3.2)
IF TRT=2 THEN LET ABSDEV=ABS(RANKCSS-7.8)
IF TRT=3 THEN LET ABSDEV=ABS(RANKCSS-14.2)
IF TRT=4 THEN LET ABSDEV=ABS(RANKCSS-16.8)
SAVE BSLEVENE
RUN

/* Levene's Test */

MGLH
USE BSLEVENE
CATEGORY TRT
ANOVA ABSDEV
ESTIMATE

/* Conover T-Test using ranks */

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE ACSS
LET RANKCSS=CSS
RANK RANKCSS
SAVE RRACSS
RUN

MGLH
USE RRACSS
CATEGORY TRT
ANOVA RANKCSS
ESTIMATE
HYPOTHESIS
POST TRT / LSD
TEST

/* Perform t-tests for each dredged sediment-reference */
/* sediment comparison using ranks */

FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
DATA
USE ACSS
IF TRT>2 THEN DELETE
SAVE ACSST1
RUN

```

```

DATA
USE ACSS
IF (TRT=2 OR TRT=4) THEN DELETE
SAVE ACSST2
RUN

DATA
USE ACSS
IF (TRT=2 OR TRT=3) THEN DELETE
SAVE ACSST3
RUN

/* t-test comparing the Reference and Sediment 1 */

DATA
USE ACSST1
LET RANKCSS=CSS
RANK RANKCSS
SAVE RRACSST1
RUN

STATS
USE RRACSST1
TTEST RANKCSS*TRT

/* t-test comparing the Reference and Sediment 2 */

DATA
USE ACSST2
LET RANKCSS=CSS
RANK RANKCSS
SAVE RRACSST2
RUN

STATS
USE RRACSST2
TTEST RANKCSS*TRT

/* t-test comparing the Reference and Sediment 3 */

DATA
USE ACSST3
LET RANKCSS=CSS
RANK RANKCSS
SAVE RRACSST3
RUN

STATS
USE RRACSST3
TTEST RANKCSS*TRT

/* Calculate power of LSD test to detect true population differences
of 10, 25, 50, and 100% above the reference mean Css. */

DATA
REPEAT 1
LET N=5
LET MEANCSS=0.099
LET SS=73.558
LET DF=16
LET MSE=SS/DF
LET TALPHA=TIF(.95,DF)
SAVE BIOCS2
PRINT N, MEANCSS, MSE, DF, TALPHA
RUN

DATA
REPEAT 1
USE BIOCS2
5 LET I=10
10 LET SIZE=10
15 LET SEDCSS=MEANCSS+((I/100)*MEANCSS)
25 LET D=SEDCSS-MEANCSS
30 LET TBETA=D*SQR(N/(2*MSE))-TALPHA
35 LET POWER=TCF(TBETA,DF)
38 PRINT I, SEDCSS, D, TBETA, POWER
40 IF I=10 THEN LET I=SIZE+15
42 IF I=25 AND SIZE=10 THEN GOTO 59
45 IF I=25 THEN LET I=SIZE+25
47 IF I=50 AND SIZE=25 THEN GOTO 59
49 IF I=50 THEN LET I=SIZE+50
50 IF I=100 AND SIZE=50 THEN GOTO 59
51 IF I=100 THEN LET I=SIZE+100
52 IF I=200 AND SIZE=100 THEN GOTO 59
53 IF I=200 THEN LET I=SIZE+100
55 IF I=300 AND SIZE=200 THEN GOTO 59
57 IF I=300 AND SIZE=300 THEN GOTO 70
59 LET SIZE=I
60 GOTO 15
70 STOP
RUN

```

```

REPEAT 1
DATA
USE BIOCS2
110 LET POWER=.5
112 LET SIZE=.5
115 LET TBETA=TIF(POWER,DF)
120 LET D=((TBETA+TALPHA)*SQR(2*MSE))/SQR(N)
125 LET SEDCSS=MEANCSS+D
130 LET PCTDIFF=(D*100)/MEANCSS
135 PRINT POWER, D, SEDCSS, PCTDIFF, TBETA
140 IF POWER<.9 THEN LET POWER=SIZE+.1
142 IF POWER=.9 AND SIZE=.8 THEN GOTO 155
145 IF POWER=.9 THEN LET POWER=SIZE+.05
147 IF POWER=.95 AND SIZE=.9 THEN GOTO 155
150 IF POWER=.95 THEN LET POWER=SIZE+.04
152 IF POWER=.99 AND SIZE=.99 THEN GOTO 170
155 LET SIZE=POWER
160 GOTO 115
170 STOP
RUN

/* Calculation of upper confidence limits (UCL) for comparison of mean */
/* dredged sediment Css with action level. */

DATA
FPATH 'C:\SYSTAT' / SAVE USE OUTPUT GET
SAVE POWERCSS
INPUT TRT MEANCSS S2
FORMAT=5
RUN
2 .496 .09364
3 3.821 5.53661
4 6.071 12.75918
~
DATA
LET N=5
LET SS=73.558
LET DF=16
LET MSE=SS/DF
LET TALPHA1=TIF(.95,DF)
LET TALPHA2=TIF(.95,N-1)
LET UCL1=MEANCSS+TALPHA1*(SQR(MSE/N))
LET UCL2=MEANCSS+TALPHA2*(SQR(S2/N))
LET DMIN1=TALPHA1*SQR(MSE/N)
LET DMIN2=TALPHA2*SQR(S2/N)
SAVE BIOCS3
PRINT TRT, MEANCSS, UCL1, MSE, TALPHA1, DF, DMIN1
PRINT TRT, MEANCSS, UCL2, S2, TALPHA2, N, DMIN2
RUN

/* Calculate power of dredged sediment-action level comparisons using MSE */
/* given 10, 20, 30, 40, and 50% decreases in mean Css below action level */

DATA
USE BIOCS2
LET ACTION=.2
FOR PCTDIFF=10 TO 50 STEP 10
LET D=PCTDIFF*ACTION/100
LET SEDCSS=ACTION-D
LET TBETA=D*SQR(N/MSE)-TALPHA
LET POWER=TCF(TBETA,DF)
PRINT PCTDIFF, SEDCSS, D, TBETA, POWER
NEXT
STOP
RUN

```

## BIOACCS.S.CMD program output

```

-----TIME-SEQUENCED BIOACCUMULATION-----
-----TREATMENT GROUP=REFERENCE-----

```

1.00000	1.00000	2.00000	0.05400	0.45000
1.00000	1.00000	4.00000	0.44100	0.45000
1.00000	1.00000	7.00000	0.68700	0.45000
1.00000	1.00000	10.00000	0.03700	0.45000
1.00000	1.00000	18.00000	0.85600	0.45000
1.00000	1.00000	28.00000	0.51400	0.45000
1.00000	2.00000	2.00000	0.16300	0.45000
1.00000	2.00000	4.00000	0.79700	0.45000
1.00000	2.00000	7.00000	0.17700	0.45000
1.00000	2.00000	10.00000	0.54900	0.45000
1.00000	2.00000	18.00000	0.59800	0.45000
1.00000	2.00000	28.00000	0.83900	0.45000
1.00000	3.00000	2.00000	0.39100	0.45000
1.00000	3.00000	4.00000	0.20300	0.45000
1.00000	3.00000	7.00000	0.86200	0.45000
1.00000	3.00000	10.00000	0.88400	0.45000

1.00000	3.00000	18.00000	0.01600	0.45000
1.00000	3.00000	28.00000	0.79300	0.45000
1.00000	4.00000	2.00000	0.23400	0.45000
1.00000	4.00000	4.00000	0.56400	0.45000
1.00000	4.00000	7.00000	0.41300	0.45000
1.00000	4.00000	10.00000	0.78700	0.45000
1.00000	4.00000	18.00000	0.80600	0.45000
1.00000	4.00000	28.00000	0.89900	0.45000
1.00000	5.00000	2.00000	0.03400	0.45000
1.00000	5.00000	4.00000	0.01800	0.45000
1.00000	5.00000	7.00000	0.02900	0.45000
1.00000	5.00000	10.00000	0.29400	0.45000
1.00000	5.00000	18.00000	0.11900	0.45000
1.00000	5.00000	28.00000	0.22600	0.45000

-----TREATMENT GROUP=SEDIMENT 1-----

2.00000	1.00000	2.00000	0.15900	4.00000
2.00000	1.00000	4.00000	0.51600	4.00000
2.00000	1.00000	7.00000	0.88100	4.00000
2.00000	1.00000	10.00000	0.27800	4.00000
2.00000	1.00000	18.00000	0.90400	4.00000
2.00000	1.00000	28.00000	0.17200	4.00000
2.00000	2.00000	2.00000	0.29200	4.00000
2.00000	2.00000	4.00000	0.15800	4.00000
2.00000	2.00000	7.00000	0.31700	4.00000
2.00000	2.00000	10.00000	0.48500	4.00000
2.00000	2.00000	18.00000	1.30000	4.00000
2.00000	2.00000	28.00000	1.04900	4.00000
2.00000	3.00000	2.00000	0.42800	4.00000
2.00000	3.00000	4.00000	0.74300	4.00000
2.00000	3.00000	7.00000	0.27000	4.00000
2.00000	3.00000	10.00000	0.05100	4.00000
2.00000	3.00000	18.00000	0.67100	4.00000
2.00000	3.00000	28.00000	0.47600	4.00000
2.00000	4.00000	2.00000	0.55800	4.00000
2.00000	4.00000	4.00000	0.32400	4.00000
2.00000	4.00000	7.00000	0.56200	4.00000
2.00000	4.00000	10.00000	0.90900	4.00000
2.00000	4.00000	18.00000	0.93400	4.00000
2.00000	4.00000	28.00000	0.71200	4.00000
2.00000	5.00000	2.00000	0.25600	4.00000
2.00000	5.00000	4.00000	0.12600	4.00000
2.00000	5.00000	7.00000	0.60300	4.00000
2.00000	5.00000	10.00000	0.71800	4.00000
2.00000	5.00000	18.00000	1.17300	4.00000
2.00000	5.00000	28.00000	1.24500	4.00000

-----TREATMENT GROUP=SEDIMENT 2-----

3.00000	1.00000	2.00000	0.86900	33.00000
3.00000	1.00000	4.00000	0.83800	33.00000
3.00000	1.00000	7.00000	1.24600	33.00000
3.00000	1.00000	10.00000	1.76700	33.00000
3.00000	1.00000	18.00000	1.63100	33.00000
3.00000	1.00000	28.00000	1.17800	33.00000
3.00000	2.00000	2.00000	0.72600	33.00000
3.00000	2.00000	4.00000	0.63300	33.00000
3.00000	2.00000	7.00000	0.81600	33.00000
3.00000	2.00000	10.00000	1.27200	33.00000
3.00000	2.00000	18.00000	1.87700	33.00000
3.00000	2.00000	28.00000	1.72100	33.00000
3.00000	3.00000	2.00000	0.39400	33.00000
3.00000	3.00000	4.00000	0.45200	33.00000
3.00000	3.00000	7.00000	0.89700	33.00000
3.00000	3.00000	10.00000	1.00300	33.00000
3.00000	3.00000	18.00000	1.48700	33.00000
3.00000	3.00000	28.00000	1.36600	33.00000
3.00000	4.00000	2.00000	1.23200	33.00000
3.00000	4.00000	4.00000	0.72800	33.00000
3.00000	4.00000	7.00000	1.63900	33.00000
3.00000	4.00000	10.00000	1.15800	33.00000
3.00000	4.00000	18.00000	1.21600	33.00000
3.00000	4.00000	28.00000	1.51300	33.00000
3.00000	5.00000	2.00000	0.97700	33.00000
3.00000	5.00000	4.00000	1.31400	33.00000
3.00000	5.00000	7.00000	0.68800	33.00000
3.00000	5.00000	10.00000	1.41500	33.00000
3.00000	5.00000	18.00000	1.28000	33.00000
3.00000	5.00000	28.00000	1.84300	33.00000

-----TREATMENT GROUP=SEDIMENT 3-----

4.00000	1.00000	2.00000	0.74500	44.00000
4.00000	1.00000	4.00000	1.31600	44.00000
4.00000	1.00000	7.00000	1.58300	44.00000
4.00000	1.00000	10.00000	1.57800	44.00000
4.00000	1.00000	18.00000	2.82200	44.00000
4.00000	1.00000	28.00000	1.29500	44.00000
4.00000	2.00000	2.00000	1.70300	44.00000
4.00000	2.00000	4.00000	0.93000	44.00000

4.00000	2.00000	7.00000	2.71500	44.00000
4.00000	2.00000	10.00000	2.26800	44.00000
4.00000	2.00000	18.00000	2.60700	44.00000
4.00000	2.00000	28.00000	2.96400	44.00000
4.00000	3.00000	2.00000	2.04500	44.00000
4.00000	3.00000	4.00000	2.14100	44.00000
4.00000	3.00000	7.00000	1.01600	44.00000
4.00000	3.00000	10.00000	1.75600	44.00000
4.00000	3.00000	18.00000	3.41400	44.00000
4.00000	3.00000	28.00000	2.10900	44.00000
4.00000	4.00000	2.00000	1.85500	44.00000
4.00000	4.00000	4.00000	1.15000	44.00000
4.00000	4.00000	7.00000	2.22100	44.00000
4.00000	4.00000	10.00000	2.89900	44.00000
4.00000	4.00000	18.00000	1.31900	44.00000
4.00000	4.00000	28.00000	2.82000	44.00000
4.00000	5.00000	2.00000	1.13500	44.00000
4.00000	5.00000	4.00000	1.62100	44.00000
4.00000	5.00000	7.00000	2.13400	44.00000
4.00000	5.00000	10.00000	0.89000	44.00000
4.00000	5.00000	18.00000	1.86600	44.00000
4.00000	5.00000	28.00000	3.32500	44.00000

/\* NOTE: The following NONLIN output is given as an example only for the reference sediment replicate 1. NONLIN output for the other replicates and sediments has been deleted to reduce volume. \*/

THE FOLLOWING RESULTS ARE FOR:

TRT = 1.00000  
 REP = 1.00000

ITERATION	LOSS	PARAMETER	VALUES
0	.5081398D+00	.7000D-01	.1100D+00
1	.5081398D+00	.7007D-01	.1649D+00
2	.5081398D+00	.7000D-01	.1650D+00
3	.5081398D+00	.7000D-01	.1650D+00
4	.5081398D+00	.7000D-01	.1650D+00
5	.5081398D+00	.7000D-01	.1650D+00
6	.5081398D+00	.7000D-01	.1650D+00

DEPENDENT VARIABLE IS CONC

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE
REGRESSION	1.15616	2	0.57808
RESIDUAL	0.50814	4	0.12703
TOTAL	1.66767	6	
CORRECTED	0.55051	5	

RAW R-SQUARED (1-RESIDUAL/TOTAL) = 0.69530  
 CORRECTED R-SQUARED (1-RESIDUAL/CORRECTED) = 0.07697

PARAMETER	ESTIMATE
K1	<u>0.07000</u>
K2	<u>0.16500</u>

/\* Time-Sequenced Bioaccumulation \*/

TRT	REPLICATE	UPTAKE RATE CONSTANT K1	DEPURATION RATE CONSTANT K2	STEADY STATE CONC. C <sub>ss</sub>	LOG10 C <sub>ss</sub>
1.00000	1.00000	0.07000	0.16500	0.19091	-0.71917
1.00000	2.00000	0.07000	0.24750	0.12727	-0.89526
1.00000	3.00000	0.07000	0.37125	0.08485	-1.07136
1.00000	4.00000	0.07000	0.55688	0.05657	-1.24745
1.00000	5.00000	0.07000	0.83531	0.03771	-1.42354
2.00000	1.00000	0.05000	0.21000	0.95238	-0.02119
2.00000	2.00000	0.05000	0.31500	0.63492	-0.19728
2.00000	3.00000	0.05000	0.47250	0.42328	-0.37337
2.00000	4.00000	0.05000	0.70875	0.28219	-0.54946
2.00000	5.00000	0.05000	1.06313	0.18812	-0.72556
3.00000	1.00000	0.04000	0.18000	7.33333	0.86530
3.00000	2.00000	0.04000	0.27000	4.88889	0.68921
3.00000	3.00000	0.04000	0.40500	3.25926	0.51312
3.00000	4.00000	0.04000	0.60750	2.17284	0.33703
3.00000	5.00000	0.04000	0.91125	1.44856	0.16094
4.00000	1.00000	0.09000	0.34500	11.47826	1.05988
4.00000	2.00000	0.09000	0.51750	7.65217	0.88378
4.00000	3.00000	0.09000	0.77625	5.10145	0.70769
4.00000	4.00000	0.09000	1.16438	3.40095	0.53160
4.00000	5.00000	0.09000	1.45547	2.72077	0.43469

/\* Time-Sequenced Bioaccumulation \*/

THE FOLLOWING RESULTS ARE FOR:  
 TRT = 1.000

	CSS	LOGCSS
N OF CASES	<u>5</u>	<u>5</u>
MEAN	<u>0.099</u>	-1.071
VARIANCE	<u>0.004</u>	<u>0.078</u>
STD. ERROR	<u>0.027</u>	<u>0.125</u>

THE FOLLOWING RESULTS ARE FOR:  
TRT = 2.000

	CSS	LOGCSS
N OF CASES	<u>5</u>	<u>5</u>
MEAN	<u>0.496</u>	-0.373
VARIANCE	<u>0.093</u>	<u>0.078</u>
STD. ERROR	<u>0.137</u>	<u>0.125</u>

THE FOLLOWING RESULTS ARE FOR:  
TRT = 3.000

	CSS	LOGCSS
N OF CASES	<u>5</u>	<u>5</u>
MEAN	<u>3.821</u>	0.513
VARIANCE	<u>5.534</u>	<u>0.078</u>
STD. ERROR	<u>1.052</u>	<u>0.125</u>

THE FOLLOWING RESULTS ARE FOR:  
TRT = 4.000

	CSS	LOGCSS
N OF CASES	<u>5</u>	<u>5</u>
MEAN	<u>6.071</u>	0.724
VARIANCE	<u>12.758</u>	<u>0.065</u>
STD. ERROR	<u>1.597</u>	<u>0.114</u>

/\* Normality test on untransformed and log-transformed data \*/

---

KOLMOGOROV-SMIRNOV ONE SAMPLE TEST USING STANDARD NORMAL DISTRIBUTION

VARIABLE	N-OF-CASES	MAXDIF	LILLIEFORS PROBABILITY (2-TAIL)
CSS	20.00000	0.20972	0.02133
LOGCSS	20.00000	0.14145	0.37020

/\* Bartlett's Test on untransformed data \*/

SUMMARY STATISTICS FOR CSS

BARTLETT TEST FOR HOMOGENEITY OF GROUP VARIANCES = 39.208

APPROXIMATE F = 12.741 DF = 3, 460 PROBABILITY = 0.000

ANALYSIS OF VARIANCE

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F	PROBABILITY
BETWEEN GROUPS	121.063	3	40.354	8.778	0.001
WITHIN GROUPS	73.558	16	4.597		

/\* Bartlett's Test on log-transformed data \*/

SUMMARY STATISTICS FOR LOGCSS

BARTLETT TEST FOR HOMOGENEITY OF GROUP VARIANCES = 0.046

APPROXIMATE F = 0.014 DF = 3, 460 PROBABILITY = 0.998

ANALYSIS OF VARIANCE

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F	PROBABILITY
BETWEEN GROUPS	10.316	3	3.439	46.250	0.000
WITHIN GROUPS	1.190	16	0.074		

/\* LSD test on untransformed data \*/

---

DEP VAR: CSS N: 20 MULTIPLE R: 0.789 SQUARED MULTIPLE R: 0.622

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
TRT	121.063	3	40.354	8.778	0.001
ERROR	<u>73.558</u>	16	4.597		

POST HOC TEST OF CSS

USING MODEL MSE OF 4.597 WITH 16. DF.  
MATRIX OF PAIRWISE MEAN DIFFERENCES:

	1	2	3	4
1	0.000			
2	0.397	0.000		
3	3.721	3.324	0.000	
4	5.971	5.575	2.250	0.000

FISHER'S LEAST-SIGNIFICANT-DIFFERENCE TEST.  
MATRIX OF PAIRWISE COMPARISON PROBABILITIES:

	1	2	3	4
1	1.000			
2	0.774	1.000		
3	0.014	0.026	1.000	
4	0.000	0.001	0.117	1.000

/\* LSD test on log-transformed data \*/

DEP VAR: LOGCSS N: 20 MULTIPLE R: 0.947 SQUARED MULTIPLE R: 0.897

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
TRT	10.316	3	3.439	46.250	0.000
ERROR	1.190	16	0.074		

POST HOC TEST OF LOGCSS

USING MODEL MSE OF .074 WITH 16. DF.  
MATRIX OF PAIRWISE MEAN DIFFERENCES:

	1	2	3	4
1	0.000			
2	0.698	0.000		
3	1.584	0.886	0.000	
4	1.795	1.097	0.210	0.000

FISHER'S LEAST-SIGNIFICANT-DIFFERENCE TEST.  
MATRIX OF PAIRWISE COMPARISON PROBABILITIES:

	1	2	3	4
1	1.000			
2	0.001	1.000		
3	0.000	0.000	1.000	
4	0.000	0.000	0.240	1.000

INDEPENDENT SAMPLES T-TEST ON CSS GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	0.099	0.061
2.000	5	0.496	0.306

SEPARATE VARIANCES T = -2.847 DF = 4.3 PROB = 0.043  
POOLED VARIANCES T = -2.847 DF = 8 PROB = 0.022

F'	DF1	DF2	PROB>F
25.164	4.000	4.000	0.009

INDEPENDENT SAMPLES T-TEST ON LOGCSS GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	-1.071	0.278
2.000	5	-0.373	0.278

SEPARATE VARIANCES T = -3.964 DF = 8.0 PROB = 0.004  
POOLED VARIANCES T = -3.964 DF = 8 PROB = 0.004

F'	DF1	DF2	PROB>F
1.000	4.000	4.000	1.000

INDEPENDENT SAMPLES T-TEST ON CSS GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	0.099	0.061
3.000	5	3.821	2.353

SEPARATE VARIANCES T = -3.536 DF = 4.0 PROB = 0.024  
 POOLED VARIANCES T = -3.536 DF = 8 PROB = 0.008

F'	DF1	DF2	PROB>F
1487.936	4.000	4.000	0.000

INDEPENDENT SAMPLES T-TEST ON LOGCSS GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	-1.071	0.278
3.000	5	0.513	0.278

SEPARATE VARIANCES T = -8.998 DF = 8.0 PROB = 0.000  
 POOLED VARIANCES T = -8.998 DF = 8 PROB = 0.000

F'	DF1	DF2	PROB>F
1.000	4.000	4.000	1.000

INDEPENDENT SAMPLES T-TEST ON CSS GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	0.099	0.061
4.000	5	6.071	3.572

SEPARATE VARIANCES T = -3.738 DF = 4.0 PROB = 0.020  
 POOLED VARIANCES T = -3.738 DF = 8 PROB = 0.006

F'	DF1	DF2	PROB>F
3428.965	4.000	4.000	0.000

INDEPENDENT SAMPLES T-TEST ON LOGCSS GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	-1.071	0.278
4.000	5	0.724	0.255

SEPARATE VARIANCES T = -10.638 DF = 7.9 PROB = 0.000  
 POOLED VARIANCES T = -10.638 DF = 8 PROB = 0.000

F'	DF1	DF2	PROB>F
1.188	4.000	4.000	0.871

/\* Time-sequenced bioaccumulation C<sub>ss</sub> converted to ranks \*/

THE FOLLOWING RESULTS ARE FOR:  
 TRT = 1.000

RANKCSS

N OF CASES	5
MEAN	<u>3.200</u>
VARIANCE	<u>3.700</u>

THE FOLLOWING RESULTS ARE FOR:  
 TRT = 2.000

RANKCSS

N OF CASES	5
MEAN	<u>7.800</u>
VARIANCE	<u>3.700</u>

THE FOLLOWING RESULTS ARE FOR:  
 TRT = 3.000

RANKCSS

N OF CASES	5
MEAN	<u>14.200</u>
VARIANCE	<u>8.200</u>

THE FOLLOWING RESULTS ARE FOR:  
 TRT = 4.000

RANKCSS

N OF CASES 5  
 MEAN 16.800  
 VARIANCE 8.200

/\* Levene's Test on ranks \*/

DEP VAR: ABSDEV N: 20 MULTIPLE R: 0.341 SQUARED MULTIPLE R: 0.116

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
TRT	3.200	3	1.067	0.703	0.564
ERROR	24.288	16	1.518		

/\* Conover T-Test \*/

DEP VAR: RANKCSS N: 20 MULTIPLE R: 0.926 SQUARED MULTIPLE R: 0.857

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
TRT	569.800	3	189.933	31.922	0.000
ERROR	95.200	16	5.950		

POST HOC TEST OF RANKCSS

USING MODEL MSE OF 5.950 WITH 16. DF.  
 MATRIX OF PAIRWISE MEAN DIFFERENCES:

	1	2	3	4
1	0.000			
2	4.600	0.000		
3	11.000	6.400	0.000	
4	13.600	9.000	2.600	0.000

FISHER'S LEAST-SIGNIFICANT-DIFFERENCE TEST.  
 MATRIX OF PAIRWISE COMPARISON PROBABILITIES:

	1	2	3	4
1	1.000			
2	0.009	1.000		
3	0.000	0.001	1.000	
4	0.000	0.000	0.111	1.000

/\* t-test on ranks \*/

INDEPENDENT SAMPLES T-TEST ON RANKCSS GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	3.200	1.924
2.000	5	7.800	1.924

SEPARATE VARIANCES T = -3.781 DF = 8.0 PROB = 0.005  
 POOLED VARIANCES T = -3.781 DF = 8 PROB = 0.005

INDEPENDENT SAMPLES T-TEST ON RANKCSS GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	3.000	1.581
3.000	5	8.000	1.581

SEPARATE VARIANCES T = -5.000 DF = 8.0 PROB = 0.001  
 POOLED VARIANCES T = -5.000 DF = 8 PROB = 0.001

INDEPENDENT SAMPLES T-TEST ON RANKCSS GROUPED BY TRT

GROUP	N	MEAN	SD
1.000	5	3.000	1.581
4.000	5	8.000	1.581

SEPARATE VARIANCES T = -5.000 DF = 8.0 PROB = 0.001  
 POOLED VARIANCES T = -5.000 DF = 8 PROB = 0.001

TIME-SEQUENCED BIOACCUMULATION POWER OF LSD TO DETECT A TRUE DIFFERENCE  
(D) ABOVE REFERENCE MEAN C<sub>SS</sub>

# OF REPLICATES	REFERENCE MEAN C <sub>SS</sub>	MEAN SQUARE ERROR	DEGREES OF FREEDOM	T VALUE FOR (1-ALPHA=0.95, DF)
5.00000	0.09900	4.59738	16.00000	1.74588

POWER OF LSD TO DETECT % INCREASE IN C<sub>SS</sub> ABOVE REFERENCE MEAN C<sub>SS</sub> GIVEN N, MSE AND DF SHOWN ABOVE

% INCREASE IN C <sub>SS</sub> ABOVE REFERENCE	DREDGED SEDIMENT C <sub>SS</sub>	D	T VALUE FOR (1-BETA, DF)	POWER (1-BETA)
10.00000	0.10890	0.00990	-1.73858	0.05065
25.00000	0.12375	0.02475	-1.72763	0.05165
50.00000	0.14850	0.04950	-1.70938	0.05335
100.00000	0.19800	0.09900	-1.67288	0.05689
200.00000	0.29700	0.19800	-1.59987	0.06459
300.00000	0.39600	0.29700	-1.52687	0.07316

MINIMUM DREDGED SEDIMENT C<sub>SS</sub> THAT CAN BE DETECTED BY LSD AS SIGNIFICANT GIVEN SPECIFIED POWER, N, MSE, AND DF SHOWN ABOVE

POWER (1-BETA)	D	DREDGED SEDIMENT C <sub>SS</sub>	% INCREASE IN C <sub>SS</sub> ABOVE REF.	T VALUE FOR (1-BETA, DF)
0.50000	2.36888	2.46788	2392.80838	0.00098
0.60000	2.71688	2.81588	2744.32407	0.25760
0.70000	3.09307	3.19207	3124.31554	0.53501
0.80000	3.54011	3.63911	3575.87140	0.86467
0.90000	4.18030	4.27930	4222.52944	1.33676
0.95000	4.73511	4.83411	4782.94141	1.74588
0.99000	5.87097	5.96997	5930.27118	2.58349

COMPARISON OF MEAN DREDGED SEDIMENT C<sub>SS</sub> WITH ACTION LEVEL: UPPER CONFIDENCE LIMITS (UCL) WHEN VARIANCES ARE EQUAL

TRT	MEAN DREDGED SEDIMENT C <sub>SS</sub>	UCL (EQUAL VARIANCES)	MEAN SQUARE ERROR	T VALUE FOR (1-ALPHA=.95, DF)	DF	MINIMUM SIGNIFICANT DIFFERENCE
2	0.49600	2.17011	4.59738	1.74588	16.0	1.67411
3	3.82100	5.49511	4.59738	1.74588	16.0	1.67411
4	6.07100	7.74511	4.59738	1.74588	16.0	1.67411

COMPARISON OF MEAN DREDGED SEDIMENT C<sub>SS</sub> WITH ACTION LEVEL: UPPER CONFIDENCE LIMITS (UCL) WHEN VARIANCES ARE UNEQUAL

TRT	MEAN DREDGED SEDIMENT C <sub>SS</sub>	UCL (UNEQUAL VARIANCES)	VARIANCE	T VALUE FOR (1-ALPHA=.95, N-1)	DF	MINIMUM SIGNIFICANT DIFFERENCE
2	0.49600	0.78774	0.09364	2.13185	5.0	0.29174
3	3.82100	6.06433	5.53661	2.13185	5.0	2.24333
4	6.07100	9.47651	12.75918	2.13185	5.0	3.40551

POWER TO DETECT % DECREASE IN C<sub>SS</sub> BELOW ACTION LEVEL OF .2 ug/g GIVEN N, MSE, AND DF SHOWN ABOVE

% DECREASE BELOW ACTION LEVEL	MEAN DREDGED SEDIMENT C <sub>SS</sub>	D	T VALUE FOR (1-BETA, DF)	POWER (1-BETA)
10.00000	0.18000	0.02000	-1.72503	0.05189
20.00000	0.16000	0.04000	-1.70417	0.05384
30.00000	0.14000	0.06000	-1.68331	0.05586
40.00000	0.12000	0.08000	-1.66245	0.05794
50.00000	0.10000	0.10000	-1.64160	0.06009

# Appendix C

## SPSS Programs

---

SAS programs provided in Appendix D of the Inland Testing Manual (U.S. Environmental Protection Agency/U.S. Army Corps of Engineers 1994)<sup>1</sup> are duplicated herein using SPSS/PC+ version 5.0 for DOS (Norusis 1992a,b). SPSS is a registered trademark of SPSS, Inc. The use of this product name does not constitute official endorsement or approval of this or any other product. Other equally acceptable software products are commercially available and may be used to perform these analyses.

The interpretation of test results is described in Appendix D of the Inland Testing Manual. There are minor differences between the SAS and SPSS programs. The SPSS programs calculate standard deviations instead of standard errors. SPSS LSD output uses (\*) to denote pairs of groups significantly different, whereas SAS denotes differences with letters of the alphabet. The SPSS LSD output is discussed in the analysis of benthic toxicity data below (BENTOX.INC). Another difference between the SPSS and SAS programs is that the algorithms to calculate rankits produce slightly different values.

Output values from one step that are used as input in subsequent steps must sometimes be manually inserted in the subsequent steps. Variable values that were inserted are underlined when they appear in output and when they are used as input. Comment statements in the programs begin with an asterisk (\*). Statements in SPSS end with a period. Several lines of output have been deleted from each program to reduce the volume of output.

### Program WATCOL.INC for Water Column Toxicity Test Data Analysis

The following program is quite similar to WATTOX.SAS in Appendix D of the Inland Testing Manual. WATCOL.INC is a program to compare water column toxicity data, control survival versus 100-percent elutriate survival,

---

<sup>1</sup> References cited in this appendix are located at the end of the main text.

using arcsine-square root transformation of the survival proportions. Analyses include mean survival for control and all elutriate dilutions, Shapiro-Wilk's Test for normality, and *t*-tests for equal or unequal variances.

## WATCOL.INC program statements

```

DATA LIST FREE / TRT REP SURV.
* Input the toxicity test data after the BEGIN DATA statement
  List the treatment code, replicate and number of survivors.
VALUE LABELS TRT 0 'DILUTION WATER' 1 '100% ELUTRIATE'
2 '50% ELUTRIATE' 3 '25% ELUTRIATE' 4 '12.5% ELUTRIATE'.
VARIABLES LABEL TRT 'TREATMENT GROUP' /
REP 'REPLICATE' /
SURV 'NUMBER OF SURVIVORS'.

COMPUTE M=20.
COMPUTE X=SQRT(SURV/M).
COMPUTE ARCSURV=ARTAN(X/SQRT(1-X*X)).
* -1 < ARTAN ARGUMENT < 1.
* When SURV=20 this is undefined.
IF (SURV EQ 20) ARCSURV=1.57080.
IF (TRT EQ 0) RESID=ARCSURV-1.48059.
IF (TRT EQ 1) RESID=ARCSURV-0.63126.
IF (TRT GT 1) RESID=0.0.
BEGIN DATA.
0 1 20 0 2 19 0 3 20 0 4 20 0 5 19
1 1 6 1 2 7 1 3 9 1 4 5 1 5 8
2 1 8 2 2 8 2 3 9 2 4 10 2 5 11
3 1 12 3 2 18 3 3 15 3 4 14 3 5 13
4 1 17 4 2 17 4 3 18 4 4 16 4 5 18
END DATA.
* Input number of organisms (M) per test container at start of test.
* Format, print, sort the data.
* Print number of observations, mean, and standard deviation for survival in each
treatment.

TITLE 'WATER COLUMN TOXICITY DATA'.
FORMATS ARCSURV (F7.5).
LIST VARIABLES=TRT REP M SURV ARCSURV.
SAVE OUTFILE 'C:\SPSS\A0.SYS' /DROP=X.
GET FILE 'C:\SPSS\A0.SYS'.
SORT CASES BY TRT.
MEANS TABLES=SURV BY TRT.
GET FILE 'C:\SPSS\A0.SYS'.
SELECT IF (TRT LE 1).
SORT CASES BY TRT.
MEANS TABLES=ARCSURV BY TRT.
EXAMINE VARIABLES=RESID /PLOT=NPLOT.
T-TEST GROUPS=TRT(0,1) /VARIABLES=ARCSURV.
RANK VARIABLES=SURV /NORMAL INTO RANKIT /FRACTION=BLOM /TIES=MEAN.
T-TEST GROUPS=TRT(0,1) /VARIABLES=RANKIT.

* POWER ANALYSIS AND MINIMUM SIGNIFICANT DIFFERENCE.
* DWMEAN= Mean survival in the dilution water.
* S0= standard deviation and N0= # replicates in dilution water.
* S1= standard deviation and N1= # replicates in 100% elutriate.
* TALPHA= value from TINV function.

DATA LIST FREE / DWMEAN N0 S0 N1 S1 TALPHA.
COMPUTE MEANPCT=DWMEAN/20.
COMPUTE S20=S0*S0.
COMPUTE S21=S1*S1.
COMPUTE DF=N0+N1-2.
COMPUTE N=(N0+N1)/2.
COMPUTE S2POOL=(S20*(N0-1)+S21*(N1-1))/DF.
COMPUTE DMIN=TALPHA*SQRT(2*S2POOL/N).
BEGIN DATA.
19.6 5 0.12352 5 0.08361 1.85955
END DATA.
FORMATS S2POOL (F8.5) TALPHA (F8.5) DMIN (F8.5).
REPORT FORMAT=AUTOMATIC LIST
/VARIABLES= N 'NUMBER' 'OF' 'REPLICATES'
MEANPCT 'MEAN' 'DILUTION' 'WATER' 'SURVIVAL'
S2POOL 'POOLED' 'VARIANCE' DF 'DEGREES' 'OF' 'FREEDOM' 'DF'
TALPHA 'T VALUE' 'FOR' '(1-ALPHA=0.95,DF)'
DMIN 'MINIMUM' 'SIGNIFICANT' 'DIFFERENCE'.

* POWER OF T-TEST TO DETECT A TRUE POPULATION DIFFERENCE.
* Calculate power from external source using tbeta and df as input to
function.
* The function computes the probability that a random variable with a
Student's t distribution with df degrees of freedom falls below the
tbeta value given.

```

```

DATA LIST FREE / PCTDIFF POWER.
COMPUTE MEANPCT=19.6/20.
COMPUTE S2POOL=0.011121.
COMPUTE TALPHA=1.85955.
COMPUTE N=5.
COMPUTE SEDSURV=MEANPCT-PCTDIFF/100.
COMPUTE X=SQRT(SEDSURV).
COMPUTE ARCSURV=ARTAN(X/SQRT(1-X*X)).
COMPUTE ARCDIFF=1.48059-ARCSURV.
COMPUTE TBETA=(SQRT(N)*ARCDIFF)/SQRT(2*S2POOL)-TALPHA.
BEGIN DATA.
10 .96508 20 .99830 30 .99980 40 .99996 50 .99999
END DATA.
FORMATS ARCSURV (F8.5) ARCDIFF (F8.5) TBETA (F8.5) POWER (F8.5).
REPORT FORMAT=AUTOMATIC LIST
/VARIABLES= PCTDIFF '% REDUCTION' 'IN' 'SURVIVAL' 'FROM' 'DIL. WATER'
SEDSURV '100%' 'ELUTRIATE' 'SURVIVAL'
ARCSURV 'ARSINE' '100%' 'ELUTRIATE' 'SURVIVAL'
ARCDIFF 'D' TBETA 'T VALUE' 'FOR' '(1-BETA,DF)' POWER.
FINISH.

```

## WATCOL.INC program output

### WATER COLUMN TOXICITY DATA

TRT	REP	M	SURV	ARCSURV
.00	1.00	20.00	20.00	1.57080
.00	2.00	20.00	19.00	1.34528
.00	3.00	20.00	20.00	1.57080
.00	4.00	20.00	20.00	1.57080
.00	5.00	20.00	19.00	1.34528
1.00	1.00	20.00	6.00	.57964
1.00	2.00	20.00	7.00	.63305
1.00	3.00	20.00	9.00	.73531
1.00	4.00	20.00	5.00	.52360
1.00	5.00	20.00	8.00	.68472
2.00	1.00	20.00	8.00	.68472
2.00	2.00	20.00	8.00	.68472
2.00	3.00	20.00	9.00	.73531
2.00	4.00	20.00	10.00	.78540
2.00	5.00	20.00	11.00	.83548
3.00	1.00	20.00	12.00	.88608
3.00	2.00	20.00	18.00	1.24905
3.00	3.00	20.00	15.00	1.04720
3.00	4.00	20.00	14.00	.99116
3.00	5.00	20.00	13.00	.93774
4.00	1.00	20.00	17.00	1.17310
4.00	2.00	20.00	17.00	1.17310
4.00	3.00	20.00	18.00	1.24905
4.00	4.00	20.00	16.00	1.10715
4.00	5.00	20.00	18.00	1.24905

Number of cases read = 25      Number of cases listed = 25

Summaries of SURV      NUMBER OF SURVIVORS  
By levels of TRT      TREATMENT GROUP

Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population					
TRT	.00	DILUTION WATER	<u>19.6000</u>	.5477	5
TRT	1.00	100% ELUTRIATE	7.0000	1.5811	5
TRT	2.00	50% ELUTRIATE	9.2000	1.3038	5
TRT	3.00	25% ELUTRIATE	14.4000	2.3022	5
TRT	4.00	12.5% ELUTRIATE	17.2000	.8367	5

Total Cases = 25

Summaries of ARCSURV  
By levels of TRT      TREATMENT GROUP

Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population					
TRT	.00	DILUTION WATER	<u>1.4805932</u>	<u>.1235208</u>	5
TRT	1.00	100% ELUTRIATE	<u>.6312648</u>	<u>.0835823</u>	5

Total Cases = 10

```

-----
RESID                      Statistic      df              Significance
Shapiro-Wilks              .8463              10              .0579

```

t-tests for independent samples of TRT      TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
ARCSURV				
DILUTION WATER	5	1.4806	.124	.055
100% ELUTRIATE	5	.6313	.084	.037

Mean Difference = .8493

Levene's Test for Equality of Variances: F= 3.932 P= .083

Variances	t-test for Equality of Means			SE of Diff	95% CI for Diff
	t-value	df	2-Tail Sig		
Equal	12.73	8	.000	.067	(.695, 1.003)
Unequal	12.73	7.03	.000	.067	(.692, 1.007)

t-tests for independent samples of TRT TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
RANKIT NORMAL of SURV using BLOM				
DILUTION WATER	5	.6991	.413	.185
100% ELUTRIATE	5	-.7401	.557	.249

Mean Difference = 1.4392

Levene's Test for Equality of Variances: F= .238 P= .639

Variances	t-test for Equality of Means			SE of Diff	95% CI for Diff
	t-value	df	2-Tail Sig		
Equal	4.64	8	.002	.310	(.724, 2.154)
Unequal	4.64	7.38	.002	.310	(.706, 2.172)

#### WATER COLUMN TOXICITY DATA

POWER OF T-TEST TO DETECT A TRUE DIFFERENCE (D) FROM MEAN DILUTION WATER SURVIVAL USING THE ARCSINE TRANSFORMATION

NUMBER OF REPLICATES	MEAN DILUTION WATER SURVIVAL	POOLED VARIANCE	DEGREES OF FREEDOM DF	T VALUE FOR (1-ALPHA=0.95, DF)	MINIMUM SIGNIFICANT DIFFERENCE
5.00	.98	.01112	8.00	1.85955	.12404

% REDUCTION IN SURVIVAL FROM DIL. WATER	100% ELUTRIATE SURVIVAL	ARSINE 100% ELUTRIATE SURVIVAL	D	T VALUE FOR (1-BETA, DF)	POWER
10.00	.88	1.21705	.26354	2.09172	.96508
20.00	.78	1.08259	.39800	4.10778	.99830
30.00	.68	.96953	.51106	5.80291	.99980
40.00	.58	.86574	.61485	7.35905	.99996
50.00	.48	.76539	.71520	8.86364	.99999

## Program BENTOX.INC for Benthic Toxicity Test Data Analysis

The following program is quite similar to BENTOX.SAS in Appendix D of the Inland Testing Manual. BENTOX.INC is a program to compare benthic toxicity data, reference survival versus survival from one or more test sediments, using arcsine-square root transformation of the survival proportions. Analyses include mean survival for reference and test sediment(s), Shapiro-Wilk's Test for normality, and *t*-tests for equal or unequal variances. The test

results are interpreted as described in Appendix D of the Inland Testing Manual.

## BENTOX.INC program statements

```
DATA LIST FREE / TRT REP SURV.
* Input the toxicity test data after the BEGIN DATA statement
  List the treatment code, replicate and number of survivors.
VALUE LABELS TRT
  1 'Reference' 2 'Sediment 1' 3 'Sediment 2' 4 'Sediment 3'.

VARIABLES LABEL TRT 'TREATMENT GROUP' /
  REP 'REPLICATE' /
  SURV 'NUMBER OF SURVIVORS'.

COMPUTE M=20.
COMPUTE X=SQRT(SURV/M).
COMPUTE ARCSURV=ATAN(X/SQRT(1-X*X)).
IF (SURV EQ 20) ARCSURV=1.57080.
IF (TRT EQ 1) RESID=ARCSURV-1.4806.
IF (TRT EQ 2) RESID=ARCSURV-1.1499.
IF (TRT EQ 3) RESID=ARCSURV-1.0201.
IF (TRT EQ 4) RESID=ARCSURV-0.9779.
BEGIN DATA.
1 1 20 1 2 20 1 3 19 1 4 19 1 5 20
2 1 17 2 2 16 2 3 18 2 4 17 2 5 15
3 1 15 3 2 16 3 3 13 3 4 17 3 5 11
4 1 17 4 2 12 4 3 10 4 4 16 4 5 13
END DATA.
* Input number of organisms (M) per test container at start of test.
* Format, print, sort the data.
* Print number of observations, mean, and standard deviation for survival in each
treatment.

TITLE 'BENTHIC TOXICITY DATA'.
FORMATS ARCSURV (F7.5).
RANK VARIABLES=SURV /NORMAL INTO RANKIT /FRACTION=BLOM /TIES=MEAN.
SAVE OUTFILE 'C:\SPSS\A0.SYS' /DROP=X.
LIST VARIABLES=TRT REP M SURV ARCSURV RANKIT.
GET FILE 'C:\SPSS\A0.SYS'.
SORT CASES BY TRT.
MEANS TABLES=SURV ARCSURV RANKIT BY TRT.
EXAMINE VARIABLES=RESID /PLOT=NPLOT.
ONEWAY ARCSURV BY TRT(1,4) /STATISTICS 3 /RANGES=LSD(.1).
T-TEST GROUPS=TRT(1,2) /VARIABLES=ARCSURV.
T-TEST GROUPS=TRT(1,3) /VARIABLES=ARCSURV.
T-TEST GROUPS=TRT(1,4) /VARIABLES=ARCSURV.

GET FILE 'C:\SPSS\A0.SYS'.
SORT CASES BY TRT.
IF (TRT EQ 1) RRESID=RANKIT-1.173434.
IF (TRT EQ 2) RRESID=RANKIT-0.077139.
IF (TRT EQ 3) RRESID=RANKIT-(-0.537159).
IF (TRT EQ 4) RRESID=RANKIT-(-0.752290).
EXAMINE VARIABLES=RRESID /PLOT=NPLOT.
ONEWAY RANKIT BY TRT(1,4) /STATISTICS 3 /RANGES=LSD(.1).
T-TEST GROUPS=TRT(1,2) /VARIABLES=RANKIT.
T-TEST GROUPS=TRT(1,3) /VARIABLES=RANKIT.
T-TEST GROUPS=TRT(1,4) /VARIABLES=RANKIT.

* POWER ANALYSIS AND MINIMUM SIGNIFICANT DIFFERENCE.
* Power of t-test to detect a true population difference.
* Calculate power from external source using tbeta and df as input to
function.
* The function computes the probability that a random variable with a
Student's t distribution with df degrees of freedom falls below the tbeta
value given.
DATA LIST FREE / PCTDIFF POWER.
COMPUTE MEANPCT=19.6/20.
COMPUTE MSE=0.0162.
COMPUTE DF=16.
COMPUTE TALPHA=1.74588.
COMPUTE N=5.
COMPUTE SEDSURV=MEANPCT-PCTDIFF/100.
COMPUTE X=SQRT(SEDSURV).
COMPUTE ARCSURV=ATAN(X/SQRT(1-X*X)).
COMPUTE ARCDIFF=1.48059-ARCSURV.
COMPUTE TBETA=(SQRT(N/(2*MSE))*ARCDIFF)-TALPHA.
BEGIN DATA.
10 .92728 20 .99722 30 .99985 40 .99999 50 1.0000
END DATA.
FORMATS MSE (F8.5) TALPHA (F8.5).
REPORT FORMAT=AUTOMATIC LIST
/VARIABLES= N 'NUMBER' 'OF' 'REPLICATES'
MEANPCT 'MEAN' 'REFERENCE' 'SURVIVAL'
MSE 'MEAN' 'SQUARE' 'ERROR'
```

```

DF 'DEGREES' 'OF' 'FREEDOM' 'DF'
TALPHA 'T VALUE' 'FOR' '(1-ALPHA=0.95,DF)'.
FORMATS PCTDIFF (F8.5) ARCSURV (F8.5) ARCDIFF (F8.5) TBETA (F8.5) POWER (F8.5).
REPORT FORMAT=AUTOMATIC LIST
/VARIABLES= PCTDIFF '% REDUCTION' 'IN' 'SURVIVAL' 'FROM' 'REFERENCE'
SEDSURV 'DREDGED' 'SEDIMENT' 'SURVIVAL'
ARCSURV 'ARSINE' 'DREDGED' 'SEDIMENT' 'SURVIVAL'
ARCDIFF 'D' TBETA 'T VALUE' 'FOR' '(1-BETA,DF)' POWER.
FINISH.

```

## BENTOX.INC program output

### BENTHIC TOXICITY DATA

TRT	REP	M	SURV	ARCSURV	RANKIT
1.00	1.00	20.00	20.00	1.57080	1.4034
1.00	2.00	20.00	20.00	1.57080	1.4034
1.00	3.00	20.00	19.00	1.34528	.8285
1.00	4.00	20.00	19.00	1.34528	.8285
1.00	5.00	20.00	20.00	1.57080	1.4034
2.00	1.00	20.00	17.00	1.17310	.2502
2.00	2.00	20.00	16.00	1.10715	-.1868
2.00	3.00	20.00	18.00	1.24905	.5895
2.00	4.00	20.00	17.00	1.17310	.2502
2.00	5.00	20.00	15.00	1.04720	-.5173
3.00	1.00	20.00	15.00	1.04720	-.5173
3.00	2.00	20.00	16.00	1.10715	-.1868
3.00	3.00	20.00	13.00	.93774	-.8285
3.00	4.00	20.00	17.00	1.17310	.2502
3.00	5.00	20.00	11.00	.83548	-1.403
4.00	1.00	20.00	17.00	1.17310	.2502
4.00	2.00	20.00	12.00	.88608	-1.128
4.00	3.00	20.00	10.00	.78540	-1.868
4.00	4.00	20.00	16.00	1.10715	-.1868
4.00	5.00	20.00	13.00	.93774	-.8285

Number of cases read = 20      Number of cases listed = 20

Summaries of SURV      NUMBER OF SURVIVORS  
By levels of TRT      TREATMENT GROUP

Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population			16.0500	2.9996	20
TRT	1.00	Reference	<u>19.6000</u>	.5477	5
TRT	2.00	Sediment 1	<u>16.6000</u>	1.1402	5
TRT	3.00	Sediment 2	<u>14.4000</u>	2.4083	5
TRT	4.00	Sediment 3	<u>13.6000</u>	2.8810	5

Total Cases = 20

### BENTHIC TOXICITY DATA

Summaries of ARCSURV  
By levels of TRT      TREATMENT GROUP

Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population			1.1571343	.2335849	20
TRT	1.00	Reference	<u>1.4805932</u>	.1235208	5
TRT	2.00	Sediment 1	<u>1.1499172</u>	.0762914	5
TRT	3.00	Sediment 2	<u>1.0201339</u>	.1347090	5
TRT	4.00	Sediment 3	<u>.9778931</u>	.1596151	5

Total Cases = 20

### BENTHIC TOXICITY DATA

Summaries of RANKIT      NORMAL of SURV using BLOM  
By levels of TRT      TREATMENT GROUP

Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population			-.009719	.935343	20
TRT	1.00	Reference	<u>1.173434</u>	.314912	5
TRT	2.00	Sediment 1	<u>-.077139</u>	.431668	5
TRT	3.00	Sediment 2	<u>-.537159</u>	.628150	5
TRT	4.00	Sediment 3	<u>-.752290</u>	.824187	5

Total Cases = 20

RESID	Statistic	df	Significance
Shapiro-Wilks	.9460	20	.3667

BENTHIC TOXICITY DATA

----- O N E W A Y -----

Variable ARCSURV  
By Variable TRT TREATMENT GROUP

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	.7779	.2593	16.0300	.0000
Within Groups	16	.2588	<u>.0162</u>		
Total	19	1.0367			

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
1.7434	3	16	.198

----- O N E W A Y -----

Variable ARCSURV  
By Variable TRT TREATMENT GROUP

Multiple Range Test

LSD Procedure  
Ranges for the .100 level -

2.47 2.47 2.47

- \* In the output below, (\*) denotes pairs of groups significantly different at the  $\alpha = 0.100$  level.
- \* For instance, (\*) denotes that Grp 4 and Grp 2 are significantly different.
- \* Hence, we conclude that Sediment 1 and Sediment 3 means are significantly different.
- \* Also, means from Sediment 1, Sediment 2 and Sediment 3 are significantly different from the Reference mean; Sediment 1 mean is not different from the Sediment 2 mean.

BENTHIC TOXICITY DATA

		G G G G
		r r r r
		p p p p
Mean	Group	4 3 2 1
.9779	Grp 4	
1.0201	Grp 3	
1.1499	Grp 2	*
1.4806	Grp 1	* * *

t-tests for independent samples of TRT TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
ARCSURV				
Reference	5	1.4806	.124	.055
Sediment 1	5	1.1499	.076	.034

Mean Difference = .3307

Levene's Test for Equality of Variances: F= 5.701 P= .044

t-test for Equality of Means				95%	
Variances	t-value	df	2-Tail Sig	SE of Diff	CI for Diff
Equal	5.09	8	.001	.065	(.181, .480)
Unequal	5.09	6.66	.002	.065	(.177, .484)

t-tests for independent samples of TRT TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
ARCSURV				
Reference	5	1.4806	.124	.055
Sediment 2	5	1.0201	.135	.060

Mean Difference = .4605

Levene's Test for Equality of Variances: F = .002 P = .963

t-test for Equality of Means					95%
Variances	t-value	df	2-Tail Sig	SE of Diff	CI for Diff
Equal	5.63	8	.000	.082	(.272, .649)
Unequal	5.63	7.94	.001	.082	(.272, .649)

t-tests for independent samples of TRT TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
ARCSURV				
Reference	5	1.4806	.124	.055
Sediment 3	5	.9779	.160	.071

Mean Difference = .5027

Levene's Test for Equality of Variances: F = .461 P = .516

t-test for Equality of Means					95%
Variances	t-value	df	2-Tail Sig	SE of Diff	CI for Diff
Equal	5.57	8	.001	.090	(.295, .711)
Unequal	5.57	7.53	.001	.090	(.295, .711)

RRESID	Statistic	df	Significance
Shapiro-Wilks	.9817	20	.9340

----- O N E W A Y -----

Variable RANKIT NORMAL of SURV using BLOM  
By Variable TRT TREATMENT GROUP

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	11.1850	3.7283	10.9708	.0004
Within Groups	16	5.4375	.3398		
Total	19	16.6225			

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
1.3288	3	16	.300

BENTHIC TOXICITY DATA

----- O N E W A Y -----

Variable RANKIT NORMAL of SURV using BLOM  
By Variable TRT TREATMENT GROUP

Multiple Range Test

LSD Procedure  
Ranges for the .100 level -

2.47 2.47 2.47

(\*) Denotes pairs of groups significantly different at the .100 level

BENTHIC TOXICITY DATA

G G G G  
r r r r  
P P P P

Mean      Group      4 3 2 1  
-.7523      Grp 4  
-.5372      Grp 3  
.0771      Grp 2      \*  
1.1734      Grp 1      \* \* \*

t-tests for independent samples of TRT      TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
RANKIT    NORMAL of SURV using BLOM				
Reference	5	1.1734	.315	.141
Sediment 1	5	.0771	.432	.193

Mean Difference = 1.0963

Levene's Test for Equality of Variances: F= .528    P= .488

Variances	t-value	df	2-Tail Sig	SE of Diff	95% CI for Diff
Equal	4.59	8	.002	.239	(.545, 1.647)
Unequal	4.59	7.32	.002	.239	(.531, 1.662)

t-tests for independent samples of TRT      TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
RANKIT    NORMAL of SURV using BLOM				
Reference	5	1.1734	.315	.141
Sediment 2	5	-.5372	.628	.281

Mean Difference = 1.7106

Levene's Test for Equality of Variances: F= 1.340    P= .280

Variances	t-value	df	2-Tail Sig	SE of Diff	95% CI for Diff
Equal	5.44	8	.001	.314	(.986, 2.435)
Unequal	5.44	5.89	.002	.314	(.941, 2.480)

t-tests for independent samples of TRT      TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
RANKIT    NORMAL of SURV using BLOM				
Reference	5	1.1734	.315	.141
Sediment 3	5	-.7523	.824	.369

Mean Difference = 1.9257

Levene's Test for Equality of Variances: F= 3.220    P= .110

Variances	t-value	df	2-Tail Sig	SE of Diff	95% CI for Diff
Equal	4.88	8	.001	.395	(1.016, 2.836)
Unequal	4.88	5.14	.004	.395	(.911, 2.940)

BENTHIC TOXICITY DATA

POWER OF LSD TO DETECT A TRUE POPULATION DIFFERENCE (D) FROM MEAN REFERENCE SURVIVAL USING THE ARCSINE TRANSFORMATION

NUMBER OF REPLICATES	MEAN REFERENCE SURVIVAL	MEAN SQUARE ERROR	DEGREES OF FREEDOM DF	T VALUE FOR (1-ALPHA=0.95, DF)
5.00	.98	.01600	16.00	1.74588

% REDUCTION IN SURVIVAL FROM REFERENCE	DREDGED SEDIMENT SURVIVAL	ARSINE DREDGED SEDIMENT SURVIVAL	D	T VALUE FOR (1-BETA, DF)	POWER
10.00000	.88	1.21705	.26354	1.54831	.92728
20.00000	.78	1.08259	.39800	3.22911	.99722
30.00000	.68	.96953	.51106	4.64234	.99985
40.00000	.58	.86574	.61485	5.93970	.99999
50.00000	.48	.76539	.71520	7.19408	1.00000

## Program BIOACC.INC for Single-Time Point Bioaccumulation Test Data Analysis

BIOACC.INC is a program to compare Tier III bioaccumulation data from dredged sediments versus reference sediment, using raw data and a  $\log_{10}$  transformation. Analyses include mean bioaccumulation for reference and test sediment(s), Shapiro-Wilk's Test for normality, LSD test, and  $t$ -tests for equal or unequal variances. The test results are interpreted as described in Appendix D of the Inland Testing Manual.

### BIOACC.INC program statements

```

DATA LIST FREE / TRT REP CONC.
* Input the bioaccumulation data after the BEGIN DATA statement
  List the treatment code, replicate and concentration.
VALUE LABELS TRT
  1 'Reference' 2 'Sediment 1' 3 'Sediment 2' 4 'Sediment 3'.

VARIABLES LABEL TRT 'TREATMENT GROUP' /
  REP 'REPLICATE' /
  CONC 'CONTAMINANT CONCENTRATION, ug/g'.

COMPUTE LOGCONC=LG10(CONC).
BEGIN DATA.
1 1 .06 1 2 .05 1 3 .05 1 4 .08 1 5 .09
2 1 .16 2 2 .19 2 3 .18 2 4 .22 2 5 .31
3 1 .24 3 2 .10 3 3 .13 3 4 .18 3 5 .30
4 1 .13 4 2 .05 4 3 .17 4 4 .08 4 5 .22
END DATA.
* Format, print, sort the data.
* Print number of observations, mean, and standard deviation.
* Calculate rankits.
TITLE 'SINGLE-TIME POINT CONTAMINANT BIOACCUMULATION DATA'.
RANK VARIABLES=CONC /NORMAL INTO RANKIT /FRACTION=BLOM /TIES=MEAN.
SAVE OUTFILE 'C:\SPSS\BIOACC.SYS'.
LIST VARIABLES=TRT REP CONC LOGCONC RANKIT.
GET FILE 'C:\SPSS\BIOACC.SYS'.
SORT CASES BY TRT.
MEANS TABLES=CONC LOGCONC RANKIT BY TRT.

GET FILE 'C:\SPSS\BIOACC.SYS'.
SORT CASES BY TRT.
IF (TRT EQ 1) RESID=CONC-0.066.
IF (TRT EQ 2) RESID=CONC-0.212.
IF (TRT EQ 3) RESID=CONC-0.19.
IF (TRT EQ 4) RESID=CONC-0.13.
IF (TRT EQ 1) LRESID=LOGCONC-(-1.1933).
IF (TRT EQ 2) LRESID=LOGCONC-(-0.6856).
IF (TRT EQ 3) LRESID=LOGCONC-(-0.7547).
IF (TRT EQ 4) LRESID=LOGCONC-(-0.9422).
IF (TRT EQ 1) RRESID=RANKIT-(-0.967708).
IF (TRT EQ 2) RRESID=RANKIT-0.745684.
IF (TRT EQ 3) RRESID=RANKIT-0.494642.
IF (TRT EQ 4) RRESID=RANKIT-(-0.235421).
EXAMINE VARIABLES=RESID LRESID RRESID /PLOT=NPLOT.
ONEWAY /VARIABLES= CONC LOGCONC RANKIT BY TRT(1,4)
  /STATISTICS 3 /RANGES=LSD(.1).
T-TEST GROUPS=TRT(1,2) /VARIABLES=CONC LOGCONC RANKIT.
T-TEST GROUPS=TRT(1,3) /VARIABLES=CONC LOGCONC RANKIT.
T-TEST GROUPS=TRT(1,4) /VARIABLES=CONC LOGCONC RANKIT.

* POWER OF T-TEST TO DETECT A TRUE POPULATION DIFFERENCE.

```

- \* Calculate power from external source using tbeta and df as input to function.
- \* The function computes the probability that a random variable with a Student's t distribution with df degrees of freedom falls below the tbeta value given.
- \* Calculate power of an LSD test to detect true population differences of 10, 25, 50, and 100% above the reference mean contaminant concentration.

```
DATA LIST FREE / PCTDIFF POWER.
COMPUTE MEANCONC=0.066.
COMPUTE SS=0.0602.
COMPUTE DF=16.
COMPUTE MSE=SS/DF.
COMPUTE TALPHA=1.74588.
COMPUTE N=5.
COMPUTE SEDCONC=MEANCONC+((PCTDIFF/100)*MEANCONC).
COMPUTE D=SEDCONC-MEANCONC.
COMPUTE TBETA=D*SQRT(N/(2*MSE))-TALPHA.
BEGIN DATA.
10 .06732 25 .10261 50 .19196 100 .48249 200 .94147 300 .99800
END DATA.
FORMATS MEANCONC (F9.5) MSE (F8.5) TALPHA (F7.5).
REPORT FORMAT=AUTOMATIC LIST
/VARIABLES= N 'NUMBER' 'OF' 'REPLICATES'
MEANCONC 'REFERENCE' 'MEAN' 'CONTAMINANT' 'CONCENTRATION'
MSE 'MEAN' 'SQUARE' 'ERROR'
DF 'DEGREES' 'OF' 'FREEDOM' 'DF'
TALPHA 'T VALUE' 'FOR' '(1-ALPHA=0.95,DF)'
FORMATS SEDCONC (F9.5) POWER (F8.5) TBETA (F8.5) D (F7.4).
REPORT FORMAT=AUTOMATIC LIST
/VARIABLES= PCTDIFF '% INCREASE' 'IN CONC.' 'ABOVE' 'REFERENCE'
SEDCONC 'DREDGED' 'SEDIMENT' 'BIOACCUMULATION' D
TBETA 'T VALUE' 'FOR' '(1-BETA,DF)' POWER 'POWER' '(1-BETA)'.
```

```
DATA LIST FREE / POWER TBETA.
COMPUTE MEANCONC=0.066.
COMPUTE SS=0.0602.
COMPUTE DF=16.
COMPUTE MSE=SS/DF.
COMPUTE TALPHA=1.74588.
COMPUTE N=5.
COMPUTE D=((TBETA+TALPHA)*SQRT(2*MSE))/SQRT(N).
COMPUTE SEDCONC=MEANCONC+D.
COMPUTE PCTDIFF=(D*100)/MEANCONC.
BEGIN DATA.
.5 0.0 .6 .2576 .7 .53501 .80 .86467 .90 1.33676 .95 1.74588 .99 2.58349
END DATA.
FORMATS SEDCONC (F9.5) POWER (F8.5) TBETA (F8.5) D (F8.5).
REPORT FORMAT=AUTOMATIC LIST
/VARIABLES= POWER 'POWER' '(1-BETA)' D
SEDCONC 'DREDGED' 'SEDIMENT' 'BIOACCUMULATION'
PCTDIFF '% INCREASE' 'IN CONC.' 'ABOVE' 'REFERENCE'
TBETA 'T VALUE' 'FOR' '(1-BETA,DF)'.
```

- \* Calculation of upper confidence limits (UCL) for comparison of mean dredged sediment bioaccumulation with an action level.

```
DATA LIST FREE / TRT MEANCONC SD.
VALUE LABELS TRT
1 'Reference' 2 'Sediment 1' 3 'Sediment 2' 4 'Sediment 3'.
COMPUTE N=5.
COMPUTE SS=0.0602.
COMPUTE DF=16.
COMPUTE MSE=SS/DF.
COMPUTE TALPHA1=1.74588.
COMPUTE TALPHA2=2.13185.
COMPUTE S2=SD*SD.
COMPUTE UCL1=MEANCONC+TALPHA1*(SQRT(MSE/N)).
COMPUTE UCL2=MEANCONC+TALPHA2*(SQRT(S2/N)).
COMPUTE DMIN1=TALPHA1*SQRT(MSE/N).
COMPUTE DMIN2=TALPHA2*SQRT(S2/N).
BEGIN DATA.
2 .212 .0589
3 .190 .0812
4 .130 .0682
END DATA.
FORMATS MEANCONC (F9.5) MSE (F8.5) TALPHA1 (F8.5) TALPHA2 (F8.5)
UCL1 (F8.5) UCL2 (F8.5) DMIN1 (F8.5) DMIN2 (F8.5) S2 (F8.5).
LIST VARIABLES= TRT MEANCONC UCL1 MSE TALPHA1 DF DMIN1.
LIST VARIABLES= TRT MEANCONC UCL2 S2 TALPHA2 N DMIN2.
```

- \* Calculate power of dredged sediment-action level comparisons using MSE given 10, 20, 30, 40, and 50% decreases in mean concentration below action level.

```
DATA LIST FREE / PCTDIFF POWER.
COMPUTE MEANCONC=0.066.
COMPUTE SS=0.0602.
COMPUTE DF=16.
COMPUTE MSE=SS/DF.
COMPUTE TALPHA=1.74588.
```

```

COMPUTE N=5.
COMPUTE ACTION=0.2.
COMPUTE D=PCTDIFF*ACTION/100.
COMPUTE SEDCONC=ACTION-D.
COMPUTE TBETA=D*SQRT(N/MSE)-TALPHA.
BEGIN DATA.
10 .16219 20 .38863 30 .66757 40 .87052 50 .96216
END DATA.
FORMATS SEDCONC (F9.5) POWER (F8.5) TBETA (F8.5).
REPORT FORMAT=AUTOMATIC LIST
/VARIABLES= PCTDIFF '% DECREASE' 'BELOW' 'ACTION' 'LEVEL'
SEDCONC 'MEAN' 'DREDGED' 'SEDIMENT' 'BIOACCUMULATION' D
TBETA 'T VALUE' 'FOR' '(1-BETA,DF)' POWER 'POWER' '(1-BETA)'.
FINISH.

```

## BIOACC.INC program output

### SINGLE-TIME POINT CONTAMINANT BIOACCUMULATION DATA

TRT	REP	CONC	LOGCONC	RANKIT
1.00	1.00	.06	-1.22	-.9191
1.00	2.00	.05	-1.30	-1.403
1.00	3.00	.05	-1.30	-1.403
1.00	4.00	.08	-1.10	-.6648
1.00	5.00	.09	-1.05	-.4478
2.00	1.00	.16	-.80	.0619
2.00	2.00	.19	-.72	.5895
2.00	3.00	.18	-.74	.3803
2.00	4.00	.22	-.66	.8285
2.00	5.00	.31	-.51	1.8682
3.00	1.00	.24	-.62	1.1281
3.00	2.00	.10	-1.00	-.3146
3.00	3.00	.13	-.89	-.1241
3.00	4.00	.18	-.74	.3803
3.00	5.00	.30	-.52	1.4034
4.00	1.00	.13	-.89	-.1241
4.00	2.00	.05	-1.30	-1.403
4.00	3.00	.17	-.77	.1868
4.00	4.00	.08	-1.10	-.6648
4.00	5.00	.22	-.66	.8285

Number of cases read = 20      Number of cases listed = 20

Summaries of CONC      CONTAMINANT CONCENTRATION, ug/g  
By levels of TRT      TREATMENT GROUP

Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population			.1495	.0810	20
TRT	1.00	Reference	<u>.0660</u>	.0182	5
TRT	2.00	Sediment 1	<u>.2120</u>	.0589	5
TRT	3.00	Sediment 2	<u>.1900</u>	.0812	5
TRT	4.00	Sediment 3	<u>.1300</u>	.0682	5

Total Cases = 20

Summaries of LOGCONC  
By levels of TRT      TREATMENT GROUP

Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population			-.8940	.2610	20
TRT	1.00	Reference	<u>-1.1933</u>	.1174	5
TRT	2.00	Sediment 1	<u>-.6856</u>	.1107	5
TRT	3.00	Sediment 2	<u>-.7547</u>	.1933	5
TRT	4.00	Sediment 3	<u>-.9422</u>	.2582	5

Total Cases = 20

### ----- SINGLE-TIME POINT CONTAMINANT BIOACCUMULATION DATA

Summaries of RANKIT      NORMAL of CONC using BLOM  
By levels of TRT      TREATMENT GROUP

Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population			.009299	.938833	20
TRT	1.00	Reference	<u>-.967708</u>	.431313	5
TRT	2.00	Sediment 1	<u>.745684</u>	.687822	5

TRT	3.00	Sediment 2	.494642	.754641	5
TRT	4.00	Sediment 3	<u>-.235421</u>	.847210	5

-----  
SINGLE-TIME POINT CONTAMINANT BIOACCUMULATION DATA

RESID	Statistic	df	Significance
Shapiro-Wilks	.9579	20	.4889
LRESID	Statistic	df	Significance
Shapiro-Wilks	.9802	20	.9154
RRESID	Statistic	df	Significance
Shapiro-Wilks	.9715	20	.7495

SINGLE-TIME POINT CONTAMINANT BIOACCUMULATION DATA

----- O N E W A Y -----

Variable By Variable    CONC TRT    CONTAMINANT CONCENTRATION, ug/g TREATMENT GROUP

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	.0645	.0215	5.7138	.0074
Within Groups	<u>16</u>	<u>.0602</u>	.0038		
Total	19	.1247			

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
2.1501	3	16	.134

Multiple Range Test

LSD Procedure

Ranges for the .100 level -

2.47    2.47    2.47

(\*) Denotes pairs of groups significantly different at the .100 level

-----  
SINGLE-TIME POINT CONTAMINANT BIOACCUMULATION DATA

G G G G  
r r r r  
P P P P

Mean	Group	
.0660	Grp 1	
.1300	Grp 4	
.1900	Grp 3	*
.2120	Grp 2	* *

SINGLE-TIME POINT CONTAMINANT BIOACCUMULATION DATA

----- O N E W A Y -----

Variable By Variable    LOGCONC TRT    TREATMENT GROUP

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	.7737	.2579	7.9320	.0018
Within Groups	16	.5202	.0325		
Total	19	1.2940			

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
2.1891	3	16	.129

Multiple Range Test

LSD Procedure

Ranges for the .100 level -

2.47 2.47 2.47

(\*) Denotes pairs of groups significantly different at the .100 level

-----  
SINGLE-TIME POINT CONTAMINANT BIOACCUMULATION DATA

G G G G  
r r r r  
p p p p

Mean	Group	
-1.1933	Grp 1	
-.9422	Grp 4	*
-.7547	Grp 3	*
-.6856	Grp 2	* *

----- O N E W A Y -----

Variable RANKIT NORMAL of CONC using BLOM  
By Variable TRT TREATMENT GROUP

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	8.9612	2.9871	6.1388	.0056
Within Groups	16	7.7855	.4866		
Total	19	16.7468			

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
.6997	3	16	.566

Multiple Range Test

LSD Procedure

Ranges for the .100 level -

2.47 2.47 2.47

(\*) Denotes pairs of groups significantly different at the .100 level

G G G G  
r r r r  
p p p p

Mean	Group	
-.9677	Grp 1	
-.2354	Grp 4	
.4946	Grp 3	*
.7457	Grp 2	* *

t-tests for independent samples of TRT TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
CONC CONTAMINANT CONCENTRATION, ug/g				
Reference	5	.0660	.018	.008
Sediment 1	5	.2120	.059	.026

Mean Difference = -.1460

Levene's Test for Equality of Variances: F= 2.927 P= .125

t-test for Equality of Means

95%

Variances	t-value	df	2-Tail Sig	SE of Diff	CI for Diff
Equal	-5.30	8	.001	.028	(-.210, -.082)
Unequal	-5.30	4.75	.004	.028	(-.217, -.075)

Variable	Number of Cases	Mean	SD	SE of Mean
LOGCONC				
Reference	5	-1.1933	.117	.052
Sediment 1	5	-.6856	.111	.050

Mean Difference = -.5077

Levene's Test for Equality of Variances: F= .212 P= .657

t-test for Equality of Means				
Variances	t-value	df	2-Tail Sig	95% CI for Diff
Equal	-7.04	8	.000	(-.674, -.341)
Unequal	-7.04	7.97	.000	(-.674, -.341)

t-tests for independent samples of TRT TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
RANKIT NORMAL of CONC using BLOM				
Reference	5	-.9677	.431	.193
Sediment 1	5	.7457	.688	.308

Mean Difference = -1.7134

Levene's Test for Equality of Variances: F= .412 P= .539

t-test for Equality of Means				
Variances	t-value	df	2-Tail Sig	95% CI for Diff
Equal	-4.72	8	.002	(-2.551, -.876)
Unequal	-4.72	6.72	.002	(-2.572, -.855)

t-tests for independent samples of TRT TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
CONC CONTAMINANT CONCENTRATION, ug/g				
Reference	5	.0660	.018	.008
Sediment 2	5	.1900	.081	.036

Mean Difference = -.1240

Levene's Test for Equality of Variances: F= 7.828 P= .023

t-test for Equality of Means				
Variances	t-value	df	2-Tail Sig	95% CI for Diff
Equal	-3.33	8	.010	(-.210, -.038)
Unequal	-3.33	4.40	.025	(-.227, -.021)

SINGLE-TIME POINT CONTAMINANT BIOACCUMULATION DATA

Variable	Number of Cases	Mean	SD	SE of Mean
LOGCONC				
Reference	5	-1.1933	.117	.052
Sediment 2	5	-.7547	.193	.086

Mean Difference = -.4386

Levene's Test for Equality of Variances: F= 1.298 P= .288

t-test for Equality of Means				
Variances	t-value	df	2-Tail Sig	95% CI for Diff
Equal	-4.34	8	.002	(-.672, -.205)
Unequal	-4.34	6.60	.004	(-.678, -.199)

t-tests for independent samples of TRT TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
RANKIT NORMAL of CONC using BLOM				
Reference	5	-.9677	.431	.193
Sediment 2	5	.4946	.755	.337

Mean Difference = -1.4623

Levene's Test for Equality of Variances: F = 2.815 P = .132

t-test for Equality of Means					
Variances	t-value	df	2-Tail Sig	SE of Diff	95% CI for Diff
Equal	-3.76	8	.006	.389	(-2.359, -.566)
Unequal	-3.76	6.36	.008	.389	(-2.414, -.511)

t-tests for independent samples of TRT TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
CONC CONTAMINANT CONCENTRATION, ug/g				
Reference	5	.0660	.018	.008
Sediment 3	5	.1300	.068	.030

Mean Difference = -.0640

Levene's Test for Equality of Variances: F = 5.164 P = .053

t-test for Equality of Means					
Variances	t-value	df	2-Tail Sig	SE of Diff	95% CI for Diff
Equal	-2.03	8	.077	.032	(-.137, .009)
Unequal	-2.03	4.56	.104	.032	(-.145, .017)

Variable	Number of Cases	Mean	SD	SE of Mean
LOGCONC				
Reference	5	-1.1933	.117	.052
Sediment 3	5	-.9422	.258	.115

Mean Difference = -.2511

Levene's Test for Equality of Variances: F = 3.678 P = .091

t-test for Equality of Means					
Variances	t-value	df	2-Tail Sig	SE of Diff	95% CI for Diff
Equal	-1.98	8	.083	.127	(-.544, .041)
Unequal	-1.98	5.58	.099	.127	(-.562, .059)

t-tests for independent samples of TRT TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
RANKIT NORMAL of CONC using BLOM				
Reference	5	-.9677	.431	.193
Sediment 3	5	-.2354	.847	.379

Mean Difference = -.7323

Levene's Test for Equality of Variances: F = 1.745 P = .223

t-test for Equality of Means					
Variances	t-value	df	2-Tail Sig	SE of Diff	95% CI for Diff
Equal	-1.72	8	.123	.425	(-1.713, .248)
Unequal	-1.72	5.94	.136	.425	(-1.773, .308)

SINGLE-TIME POINT CONTAMINANT BIOACCUMULATION DATA

POWER OF LSD TO DETECT A TRUE POPULATION DIFFERENCE (D) ABOVE REFERENCE MEAN CONTAMINANT CONCENTRATION

NUMBER OF REPLICATES	REFERENCE MEAN CONTAMINANT CONCENTRATION	MEAN SQUARE ERROR	DEGREES OF FREEDOM DF	T VALUE FOR (1-ALPHA=0.95, DF)
5.00	.06600	.00376	16.00	1.74588

POWER OF LSD TO DETECT % INCREASE IN CONCENTRATION ABOVE REFERENCE MEAN CONTAMINANT CONCENTRATION GIVEN N, MSE AND DF SHOWN ABOVE

% INCREASE IN CONC. ABOVE REFERENCE	DREDGED SEDIMENT BIOACCUMULATION	D	T VALUE FOR (1-BETA, DF)	POWER (1-BETA)
10.00	.07260	.0066	-1.57575	.06732
25.00	.08250	.0165	-1.32056	.10261
50.00	.09900	.0330	-.89524	.19196
100.00	.13200	.0660	-.04460	.48249
200.00	.19800	.1320	1.65668	.94147
300.00	.26400	.1980	3.35796	.99800

MINIMUM DREDGED SEDIMENT BIOACCUMULATION THAT CAN BE DETECTED BY LSD AS SIGNIFICANT GIVEN SPECIFIED POWER AND N, MSE AND DF SHOWN ABOVE

POWER (1-BETA)	D	DREDGED SEDIMENT BIOACCUMULATION	% INCREASE IN CONC. ABOVE REFERENCE	T VALUE FOR (1-BETA, DF)
.50000	.06773	.13373	102.62	0.00000
.60000	.07772	.14372	117.76	.25760
.70000	.08849	.15449	134.07	.53501
.80000	.10127	.16727	153.45	.86467
.90000	.11959	.18559	181.20	1.33676
.95000	.13546	.20146	205.24	1.74588
.99000	.16796	.23396	254.48	2.58349

COMPARISON OF MEAN DREDGED SEDIMENT BIOACCUMULATION WITH ACTION LEVEL: UPPER CONFIDENCE LIMITS (UCL) WHEN VARIANCES ARE EQUAL

TRT	MEANCONC	UCL1	MSE	TALPHA1	DF	DMIN1
2.00	.21200	.25989	.00376	1.74588	16.00	.04789
3.00	.19000	.23789	.00376	1.74588	16.00	.04789
4.00	.13000	.17789	.00376	1.74588	16.00	.04789

COMPARISON OF MEAN DREDGED SEDIMENT BIOACCUMULATION WITH ACTION LEVEL: UPPER CONFIDENCE LIMITS (UCL) WHEN VARIANCES ARE UNEQUAL

TRT	MEANCONC	UCL2	S2	TALPHA2	N	DMIN2
2.00	.21200	.26816	.00347	2.13185	5.00	.05616
3.00	.19000	.26745	.00660	2.13185	5.00	.07745
4.00	.13000	.19501	.00465	2.13185	5.00	.06501

POWER TO DETECT % DECREASE IN CONCENTRATION BELOW ACTION LEVEL OF 0.2 ug/g GIVEN N, MSE, AND DF SHOWN ABOVE

% DECREASE BELOW ACTION LEVEL	MEAN DREDGED SEDIMENT BIOACCUMULATION	D	T VALUE FOR (1-BETA, DF)	POWER (1-BETA)
10.00	.18000	.02	-1.01680	.16219
20.00	.16000	.04	-.28771	.38863
30.00	.14000	.06	.44137	.66757
40.00	.12000	.08	1.17045	.87052
50.00	.10000	.10	1.89953	.96216

# Program BIOACSS.INC for Time-Sequenced Bioaccumulation Test Data Analysis

This program is designed to compare Tier IV estimated steady-state bioaccumulation data from dredged sediments versus reference sediment, using raw data and a  $\log_{10}$  transformation. Analyses include mean bioaccumulation from each sediment exposure, Shapiro-Wilk's Test for normality, Levene's Test for equality of variances, *t*-tests for equal or unequal variances, LSD test, and nonparametric LSD test using rankits. The test results are interpreted as described in Appendix D of the Inland Testing Manual. The nonlinear regression procedure NLR is part of Advanced Statistics 5.0 (Norusis 1992b). Statements necessary to produce graphs are provided in the following program, but the graphic output is not included in this appendix.

## BIOACSS.INC program statements

```
DATA LIST FREE / DAY REP TRT CONC.
*   Input the bioaccumulation data after the BEGIN DATA statement.
*   List day, replicate, treatment code, and contaminant concentration.

VALUE LABELS TRT
  1 'Reference' 2 'Sediment 1' 3 'Sediment 2' 4 'Sediment 3'.

VARIABLES LABEL TRT 'TREATMENT GROUP' /
              REP 'REPLICATE' /
              CONC 'CONTAMINANT CONCENTRATION, ug/g'.

  IF (TRT=1) CS=.45.
  IF (TRT=2) CS=4.
  IF (TRT=3) CS=33.
  IF (TRT=4) CS=44.

BEGIN DATA.
2 1 1 .054 2 2 1 .163 2 3 1 .391 2 4 1 .234 2 5 1 .034
2 1 2 .159 2 2 2 .292 2 3 2 .428 2 4 2 .558 2 5 2 .256
2 1 3 .869 2 2 3 .726 2 3 3 .394 2 4 3 1.232 2 5 3 .977
2 1 4 .745 2 2 4 1.703 2 3 4 2.045 2 4 4 1.855 2 5 4 1.135
4 1 1 .441 4 2 1 .797 4 3 1 .203 4 4 1 .564 4 5 1 .018
4 1 2 .516 4 2 2 .158 4 3 2 .743 4 4 2 .324 4 5 2 .126
4 1 3 .838 4 2 3 .633 4 3 3 .452 4 4 3 .728 4 5 3 1.314
4 1 4 1.316 4 2 4 .930 4 3 4 2.141 4 4 4 1.150 4 5 4 1.621
7 1 1 .687 7 2 1 .177 7 3 1 .862 7 4 1 .413 7 5 1 .029
7 1 2 .881 7 2 2 .317 7 3 2 .270 7 4 2 .562 7 5 2 .603
7 1 3 1.246 7 2 3 .816 7 3 3 .897 7 4 3 1.639 7 5 3 .688
7 1 4 1.583 7 2 4 2.715 7 3 4 1.016 7 4 4 2.221 7 5 4 2.134
10 1 1 .037 10 2 1 .549 10 3 1 .884 10 4 1 .787 10 5 1 .294
10 1 2 .278 10 2 2 .485 10 3 2 .051 10 4 2 .909 10 5 2 .718
10 1 3 1.767 10 2 3 1.272 10 3 3 1.003 10 4 3 1.158 10 5 3 1.415
10 1 4 1.578 10 2 4 2.268 10 3 4 1.756 10 4 4 2.899 10 5 4 .890
18 1 1 .856 18 2 1 .598 18 3 1 .016 18 4 1 .806 18 5 1 .119
18 1 2 .904 18 2 2 1.300 18 3 2 .671 18 4 2 .934 18 5 2 1.173
18 1 3 1.631 18 2 3 1.877 18 3 3 1.487 18 4 3 1.216 18 5 3 1.280
18 1 4 2.822 18 2 4 2.607 18 3 4 3.414 18 4 4 1.319 18 5 4 1.866
28 1 1 .514 28 2 1 .839 28 3 1 .793 28 4 1 .899 28 5 1 .226
28 1 2 .172 28 2 2 1.049 28 3 2 .476 28 4 2 .712 28 5 2 1.245
28 1 3 1.178 28 2 3 1.721 28 3 3 1.366 28 4 3 1.513 28 5 3 1.843
28 1 4 1.295 28 2 4 2.964 28 3 4 2.109 28 4 4 2.820 28 5 4 3.325
END DATA.

TITLE 'TIME-SEQUENCED BIOACCUMULATION'.
SORT CASES BY TRT REP DAY.
LIST VARIABLES=TRT REP DAY CONC CS.
SAVE OUTFILE 'C:\SPSS\BIOACSS.SYS'.

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 1).
PLOT FORMAT=DEFAULT
/TITLE='TIME-SEQUENCED BIOACCUMULATION'
/PLOT=CONC WITH DAY BY REP.
GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 2).
PLOT FORMAT=DEFAULT
/TITLE='TIME-SEQUENCED BIOACCUMULATION'
```

```

/PLOT=CONC WITH DAY BY REP.
GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 3).
PLOT FORMAT=DEFAULT
/TITLE='TIME-SEQUENCED BIOACCUMULATION'
/PLOT=CONC WITH DAY BY REP.
GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 4).
PLOT FORMAT=DEFAULT
/TITLE='TIME-SEQUENCED BIOACCUMULATION'
/PLOT=CONC WITH DAY BY REP.

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 1 AND REP EQ 1).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 1 AND REP EQ 2).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 1 AND REP EQ 3).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 1 AND REP EQ 4).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 1 AND REP EQ 5).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 2 AND REP EQ 1).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 2 AND REP EQ 2).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 2 AND REP EQ 3).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

```

```

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 2 AND REP EQ 4).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

```

```

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 2 AND REP EQ 5).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

```

```

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 3 AND REP EQ 1).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

```

```

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 3 AND REP EQ 2).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

```

```

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 3 AND REP EQ 3).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

```

```

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 3 AND REP EQ 4).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

```

```

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 3 AND REP EQ 5).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

```

```

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 4 AND REP EQ 1).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

```

```

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 4 AND REP EQ 2).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

```

```

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 4 AND REP EQ 3).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 4 AND REP EQ 4).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).

GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (TRT EQ 4 AND REP EQ 5).
MODEL PROGRAM K1=0.23 K2=0.17.
COMPUTE EX=EXP(-K2*DAY).
COMPUTE PRED=CS*(K1/K2)*(1-EX).
DERIVATIVES.
COMPUTE D.K1=CS/K2*(1-EX).
COMPUTE D.K2=CS*(K1/K2)*(DAY*EX-(1-EX)/K2).
NLR CONC WITH CS DAY /CRITERIA=CKDER(0).
GET FILE 'C:\SPSS\BIOACSS.SYS'.
SELECT IF (DAY EQ 28).
SAVE OUTFILE 'C:\SPSS\A.SYS'.
DATA LIST FREE / TRT REP K1 K2.

*   Input values of k1 and k2.

BEGIN DATA.
1 1 .2373 .1757 1 2 .3056 .2002 1 3 .5397 .4067 1 4 .3180 .1621 1 5 .0452 .0867
2 1 .0592 .4270 2 2 .0192 .0468 2 3 .2426 2.201 2 4 .0506 .2429 2 5 .0242 .0605
3 1 .0144 .3191 3 2 .0065 .1131 3 3 .0055 .1196 3 4 .0341 .8710 3 5 .0232 .5675
4 1 .0112 .2502 4 2 .0149 .2362 4 3 .0937 1.976 4 4 .0235 .4578 4 5 .0084 .1390
END DATA.
SORT CASES BY TRT REP.
JOIN MATCH FILE='A.SYS' /DROP=DAY
      /FILE=* /BY TRT REP.
COMPUTE CSS=CS*K1/K2.
COMPUTE LOGCSS=LG10(CSS).
COMPUTE CSS=CS*K1/K2.
COMPUTE LOGCSS=LG10(CSS).
IF (TRT EQ 1) RESID=CSS-0.6018.
IF (TRT EQ 2) RESID=CSS-1.0139.
IF (TRT EQ 3) RESID=CSS-1.5089.
IF (TRT EQ 4) RESID=CSS-2.3499.
IF (TRT EQ 1) LRESID=LOGCSS-(-0.2574).
IF (TRT EQ 2) LRESID=LOGCSS-(-0.0543).
IF (TRT EQ 3) LRESID=LOGCSS-0.1747.
IF (TRT EQ 4) LRESID=LOGCSS-0.3671.

*   Test normality of Css rankits.

RANK VARIABLES=CSS /NORMAL INTO RANKIT /FRACTION=BLOM /TIES=MEAN.
SAVE OUTFILE 'C:\SPSS\PARMS.SYS'.
GET FILE 'C:\SPSS\PARMS.SYS'.
IF (TRT EQ 1) RRESID=RANKIT-(-0.887109).
IF (TRT EQ 2) RRESID=RANKIT-(-0.443397).
IF (TRT EQ 3) RRESID=RANKIT-0.117891.
IF (TRT EQ 4) RRESID=RANKIT-1.212615.
FORMATS K1 (F8.5) K2 (F8.5) CSS (F8.5) LOGCSS (F8.5) RANKIT (F8.5).
LIST VARIABLES=TRT REP K1 K2 CSS LOGCSS RANKIT.
SAVE OUTFILE 'C:\SPSS\PARMS2.SYS'.
MEANS TABLES=CSS LOGCSS RANKIT BY TRT.
EXAMINE VARIABLES=RESID LRESID RRESID/PLOT=NPLOT.
ONEWAY /VARIABLES= CSS LOGCSS RANKIT BY TRT(1,4)
      /STATISTICS 3 /RANGES=LSD(.1).
T-TEST GROUPS=TRT(1,2) /VARIABLES=CSS LOGCSS RANKIT.
T-TEST GROUPS=TRT(1,3) /VARIABLES=CSS LOGCSS RANKIT.
T-TEST GROUPS=TRT(1,4) /VARIABLES=CSS LOGCSS RANKIT.

*   POWER OF LSD TO DETECT A TRUE POPULATION DIFFERENCE (D).
*   Calculate power from external source using tbeta and df
    as input to function.
*   The function computes the probability that a random
    variable with a Student's t distribution with df degrees
    of freedom falls below the tbeta value given.
*   Calculate power of an LSD test to detect true population
    differences of 10, 25, 50, and 100% above the reference
    mean Css.

DATA LIST FREE / PCTDIFF POWER.
COMPUTE MEANCSS=0.6018.
COMPUTE SS=2.2523.

```

```

COMPUTE DF=16.
COMPUTE MSE=SS/DF.
COMPUTE TALPHA=1.74588.
COMPUTE N=5.
COMPUTE SEDCSS=MEANCSS+((PCTDIFF/100)*MEANCSS).
COMPUTE D=SEDCSS-MEANCSS.
COMPUTE TBETA=D*SQRT(N/(2*MSE))-TALPHA.
BEGIN DATA.
10 .07746 25 .14097 50 .31848 100 .77767 200 .99780 300 .99999
END DATA.
FORMATS MEANCSS (F9.5) MSE (F8.5) TALPHA (F7.5).
REPORT FORMAT=AUTOMATIC LIST
/VARIABLES= N 'NUMBER' 'OF' 'REPLICATES'
MEANCSS 'REFERENCE' 'MEAN' 'Css'
MSE 'MEAN' 'SQUARE' 'ERROR'
DF 'DEGREES' 'OF' 'FREEDOM' 'DF'
TALPHA 'T VALUE' 'FOR' '(1-ALPHA=0.95,DF)'.
FORMATS SEDCSS (F9.5) POWER (F8.5) TBETA (F8.5) D (F7.4).
REPORT FORMAT=AUTOMATIC LIST
/VARIABLES= PCTDIFF '% INCREASE' 'IN Css' 'ABOVE' 'REFERENCE'
SEDCSS 'DREDGED' 'SEDIMENT' 'Css' D
TBETA 'T VALUE' 'FOR' '(1-BETA,DF)' POWER 'POWER' '(1-BETA)'.

DATA LIST FREE / POWER TBETA.
COMPUTE MEANCSS=0.6018.
COMPUTE SS=2.2523.
COMPUTE DF=16.
COMPUTE MSE=SS/DF.
COMPUTE TALPHA=1.74588.
COMPUTE N=5.
COMPUTE D=((TBETA+TALPHA)*SQRT(2*MSE))/SQRT(N).
COMPUTE SEDCSS=MEANCSS+D.
COMPUTE PCTDIFF=(D*100)/MEANCSS.
BEGIN DATA.
.5 0.0 .6 .2576 .7 .53501 .80 .86467 .90 1.33676 .95 1.74588 .99 2.58349
END DATA.
FORMATS SEDCSS (F9.5) POWER (F8.5) TBETA (F8.5) D (F8.5).
REPORT FORMAT=AUTOMATIC LIST
/VARIABLES= POWER 'POWER' '(1-BETA)' D
SEDCSS 'DREDGED' 'SEDIMENT' 'Css'
PCTDIFF '% INCREASE' 'IN Css' 'ABOVE' 'REFERENCE'
TBETA 'T VALUE' 'FOR' '(1-BETA,DF)'.

* Calculation of upper confidence limits (UCL) for comparison of mean
dredged sediment Css with an action level.

DATA LIST FREE / TRT MEANCSS SD.
VALUE LABELS TRT
1 'Reference' 2 'Sediment 1' 3 'Sediment 2' 4 'Sediment 3'.
COMPUTE N=5.
COMPUTE SS=2.2523.
COMPUTE DF=16.
COMPUTE MSE=SS/DF.
COMPUTE TALPHA1=1.74588.
COMPUTE TALPHA2=2.13185.
COMPUTE S2=SD*SD.
COMPUTE UCL1=MEANCSS+TALPHA1*(SQRT(MSE/N)).
COMPUTE UCL2=MEANCSS+TALPHA2*(SQRT(S2/N)).
COMPUTE DMIN1=TALPHA1*SQRT(MSE/N).
COMPUTE DMIN2=TALPHA2*SQRT(S2/N).
BEGIN DATA.
2 1.0139 .5720043
3 1.5089 .2362959
4 2.3499 .3532433
END DATA.
FORMATS MEANCSS (F9.5) MSE (F8.5) TALPHA1 (F8.5) TALPHA2 (F8.5)
UCL1 (F8.5) UCL2 (F8.5) DMIN1 (F8.5) DMIN2 (F8.5) S2 (F8.5).
LIST VARIABLES= TRT MEANCSS UCL1 MSE TALPHA1 DF DMIN1.
LIST VARIABLES= TRT MEANCSS UCL2 S2 TALPHA2 N DMIN2.

* Calculate power of dredged sediment-action level comparisons using MSE
given 10, 20, 30, 40, and 50% decreases in mean Css below action level.

DATA LIST FREE / PCTDIFF POWER.
COMPUTE MEANCSS=0.6018.
COMPUTE SS=2.2523.
COMPUTE DF=16.
COMPUTE MSE=SS/DF.
COMPUTE TALPHA=1.74588.
COMPUTE N=5.
COMPUTE ACTION=2.0.
COMPUTE D=PCTDIFF*ACTION/100.
COMPUTE SEDCSS=ACTION-D.
COMPUTE TBETA=D*SQRT(N/MSE)-TALPHA.
BEGIN DATA.
10 .29268 20 .73192 30 .95634 40 .99585 50 .99966
END DATA.
FORMATS SEDCSS (F9.5) POWER (F8.5) TBETA (F8.5).
REPORT FORMAT=AUTOMATIC LIST
/VARIABLES= PCTDIFF '% DECREASE' 'BELOW' 'ACTION' 'LEVEL'
SEDCSS 'DREDGED' 'SEDIMENT' 'Css' D

```

TBETA 'T VALUE' 'FOR' '(1-BETA,DF)' POWER 'POWER' '(1-BETA)'.  
FINISH.

## BIOACSS.INC program output

### TIME-SEQUENCED BIOACCUMULATION

TRT	REP	DAY	CONC	CS
1.00	1.00	2.00	.05	.45
1.00	1.00	4.00	.44	.45
1.00	1.00	7.00	.69	.45
1.00	1.00	10.00	.04	.45
1.00	1.00	18.00	.86	.45
1.00	1.00	28.00	.51	.45
1.00	2.00	2.00	.16	.45
1.00	2.00	4.00	.80	.45
1.00	2.00	7.00	.18	.45
1.00	2.00	10.00	.55	.45
1.00	2.00	18.00	.60	.45
1.00	2.00	28.00	.84	.45
1.00	3.00	2.00	.39	.45
1.00	3.00	4.00	.20	.45
1.00	3.00	7.00	.86	.45
1.00	3.00	10.00	.88	.45
1.00	3.00	18.00	.02	.45
1.00	3.00	28.00	.79	.45
1.00	4.00	2.00	.23	.45
1.00	4.00	4.00	.56	.45
1.00	4.00	7.00	.41	.45
1.00	4.00	10.00	.79	.45
1.00	4.00	18.00	.81	.45
1.00	4.00	28.00	.90	.45
1.00	5.00	2.00	.03	.45
1.00	5.00	4.00	.02	.45
1.00	5.00	7.00	.03	.45
1.00	5.00	10.00	.29	.45
1.00	5.00	18.00	.12	.45
1.00	5.00	28.00	.23	.45
2.00	1.00	2.00	.16	4.00
2.00	1.00	4.00	.52	4.00
2.00	1.00	7.00	.88	4.00
2.00	1.00	10.00	.28	4.00
2.00	1.00	18.00	.90	4.00
2.00	1.00	28.00	.17	4.00
2.00	2.00	2.00	.29	4.00
2.00	2.00	4.00	.16	4.00
2.00	2.00	7.00	.32	4.00
2.00	2.00	10.00	.49	4.00
2.00	2.00	18.00	1.30	4.00
2.00	2.00	28.00	1.05	4.00
2.00	3.00	2.00	.43	4.00
2.00	3.00	4.00	.74	4.00
2.00	3.00	7.00	.27	4.00
2.00	3.00	10.00	.05	4.00
2.00	3.00	18.00	.67	4.00
2.00	3.00	28.00	.48	4.00
2.00	4.00	2.00	.56	4.00
2.00	4.00	4.00	.32	4.00
2.00	4.00	7.00	.56	4.00
2.00	4.00	10.00	.91	4.00
2.00	4.00	18.00	.93	4.00
2.00	4.00	28.00	.71	4.00
2.00	5.00	2.00	.26	4.00
2.00	5.00	4.00	.13	4.00
2.00	5.00	7.00	.60	4.00
2.00	5.00	10.00	.72	4.00
2.00	5.00	18.00	1.17	4.00
2.00	5.00	28.00	1.25	4.00
3.00	1.00	2.00	.87	33.00
3.00	1.00	4.00	.84	33.00
3.00	1.00	7.00	1.25	33.00
3.00	1.00	10.00	1.77	33.00
3.00	1.00	18.00	1.63	33.00
3.00	1.00	28.00	1.18	33.00
3.00	2.00	2.00	.73	33.00
3.00	2.00	4.00	.63	33.00
3.00	2.00	7.00	.82	33.00
3.00	2.00	10.00	1.27	33.00
3.00	2.00	18.00	1.88	33.00
3.00	2.00	28.00	1.72	33.00
3.00	3.00	2.00	.39	33.00
3.00	3.00	4.00	.45	33.00
3.00	3.00	7.00	.90	33.00
3.00	3.00	10.00	1.00	33.00
3.00	3.00	18.00	1.49	33.00
3.00	3.00	28.00	1.37	33.00
3.00	4.00	2.00	1.23	33.00

3.00	4.00	4.00	.73	33.00
3.00	4.00	7.00	1.64	33.00
3.00	4.00	10.00	1.16	33.00
3.00	4.00	18.00	1.22	33.00
3.00	4.00	28.00	1.51	33.00
3.00	5.00	2.00	.98	33.00
3.00	5.00	4.00	1.31	33.00
3.00	5.00	7.00	.69	33.00
3.00	5.00	10.00	1.42	33.00
3.00	5.00	18.00	1.28	33.00
3.00	5.00	28.00	1.84	33.00
4.00	1.00	2.00	.75	44.00
4.00	1.00	4.00	1.32	44.00
4.00	1.00	7.00	1.58	44.00
4.00	1.00	10.00	1.58	44.00
4.00	1.00	18.00	2.82	44.00
4.00	1.00	28.00	1.30	44.00
4.00	2.00	2.00	1.70	44.00
4.00	2.00	4.00	.93	44.00
4.00	2.00	7.00	2.72	44.00
4.00	2.00	10.00	2.27	44.00
4.00	2.00	18.00	2.61	44.00
4.00	2.00	28.00	2.96	44.00
4.00	3.00	2.00	2.05	44.00
4.00	3.00	4.00	2.14	44.00
4.00	3.00	7.00	1.02	44.00
4.00	3.00	10.00	1.76	44.00
4.00	3.00	18.00	3.41	44.00
4.00	3.00	28.00	2.11	44.00
4.00	4.00	2.00	1.86	44.00
4.00	4.00	4.00	1.15	44.00
4.00	4.00	7.00	2.22	44.00
4.00	4.00	10.00	2.90	44.00
4.00	4.00	18.00	1.32	44.00
4.00	4.00	28.00	2.82	44.00
4.00	5.00	2.00	1.14	44.00
4.00	5.00	4.00	1.62	44.00
4.00	5.00	7.00	2.13	44.00
4.00	5.00	10.00	.89	44.00
4.00	5.00	18.00	1.87	44.00
4.00	5.00	28.00	3.33	44.00

- \* Note: The following NLR output is given as an example only for the Reference Sediment replicate 1.
- \* NLR output for the other replicates and sediments has been deleted.

#### TIME-SEQUENCED BIOACCUMULATION

There are 6 cases. There is enough memory for them all.

Iteration	Residual SS	K1	K2
1	.4001133841	.2300000000	.1700000000
1.1	.3999837931	.236601107	.174992891
2	.3999837931	.236601107	.174992891
2.1	.3999828889	.237238249	.175595606
3	.3999828889	.237238249	.175595606
3.1	.3999828767	.237309469	.175667422
4	.3999828767	.237309469	.175667422
4.1	.3999828765	.237317874	.175675916

Run stopped after 8 model evaluations and 4 derivative evaluations. Iterations have been stopped because the relative reduction between successive residual sums of squares is at most SSCON = 1.000E-08

#### Nonlinear Regression Summary Statistics      Dependent Variable CONC

Source	DF	Sum of Squares	Mean Square
Regression	2	1.26768	.63384
Residual	4	.39998	.10000
Uncorrected Total	6	1.66767	
(Corrected Total)	5	.55051	

R squared = 1 - Residual SS / Corrected SS = .27344

Parameter	Estimate	Asymptotic Std. Error	Asymptotic 95 % Confidence Interval	
			Lower	Upper
K1	.237317874	.225010285	-.387410829	.862046577
K2	.175675916	.217227321	-.427443816	.778795649

#### TIME-SEQUENCED BIOACCUMULATION

TRT	REP	K1	K2	CSS	LOGCSS	RANKIT
1.00	1.00	.23730	.17570	.60777	-.21626	-.74414
1.00	2.00	.30560	.20020	.68691	-.16310	-.58946

1.00	3.00	.53970	.40670	.59716	-.22391	-.91914
1.00	4.00	.31800	.16210	.88279	-.05414	-.31457
1.00	5.00	.04520	.08670	.23460	-.62967	-1.86824
2.00	1.00	.05920	.42700	.55457	-.25605	-1.12814
2.00	2.00	.01920	.04680	1.64103	.21512	.44777
2.00	3.00	.24260	2.20100	.44089	-.35567	-1.40341
2.00	4.00	.05060	.24290	.83326	-.07922	-.44777
2.00	5.00	.02420	.06050	1.60000	.20412	.31457
3.00	1.00	.01440	.31910	1.48919	.17295	.06193
3.00	2.00	.00650	.11310	1.89655	.27796	.58946
3.00	3.00	.00550	.11960	1.51756	.18115	.18676
3.00	4.00	.03410	.87100	1.29196	.11125	-.18676
3.00	5.00	.02320	.56750	1.34907	.13004	-.06193
4.00	1.00	.01120	.25020	1.96962	.29438	.74414
4.00	2.00	.01490	.23620	2.77561	.44336	1.86824
4.00	3.00	.09370	1.97600	2.08644	.31941	.91914
4.00	4.00	.02350	.45780	2.25863	.35384	1.12814
4.00	5.00	.00840	.13900	2.65899	.42472	1.40341

TIME-SEQUENCED BIOACCUMULATION

Summaries of By levels of	CSS TRT	TREATMENT GROUP	Mean	Std Dev	Cases
Variable	Value	Label			
For Entire Population			1.3686307	.7516114	20
TRT	1.00	Reference	.6018465	.2351011	5
TRT	2.00	Sediment 1	<u>1.0139495</u>	<u>.5720043</u>	5
TRT	3.00	Sediment 2	<u>1.5088673</u>	<u>.2362959</u>	5
TRT	4.00	Sediment 3	<u>2.3498593</u>	<u>.3532433</u>	5

Summaries of By levels of	LOGCSS TRT	TREATMENT GROUP	Mean	Std Dev	Cases
Variable	Value	Label			
For Entire Population			.0575139	.2904762	20
TRT	1.00	Reference	-.2574161	.2188628	5
TRT	2.00	Sediment 1	<u>-.0543394</u>	<u>.2605337</u>	5
TRT	3.00	Sediment 2	<u>-.1746692</u>	<u>.0646702</u>	5
TRT	4.00	Sediment 3	<u>.3671419</u>	<u>.0649487</u>	5

Summaries of By levels of	RANKIT TRT	NORMAL of TREATMENT GROUP	Mean	Std Dev	Cases
Variable	Value	Label			
For Entire Population			-8.882E-17	.9633703	20
TRT	1.00	Reference	-.8871094	.5917094	5
TRT	2.00	Sediment 1	<u>-.4433969</u>	<u>.8305451</u>	5
TRT	3.00	Sediment 2	<u>-.1178911</u>	<u>.2980743</u>	5
TRT	4.00	Sediment 3	<u>1.2126151</u>	<u>.4412995</u>	5

TIME-SEQUENCED BIOACCUMULATION  
TEST FOR NORMALITY

RESID	Statistic	df	Significance
Shapiro-Wilks	.9637	20	.5938
LRESID	Statistic	df	Significance
Shapiro-Wilks	.9417	20	.3230
RRESID	Statistic	df	Significance
Shapiro-Wilks	.9701	20	.7226

TIME-SEQUENCED BIOACCUMULATION

- - - - - O N E W A Y - - - - -

Variable	CSS	TREATMENT GROUP
By Variable	TRT	

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	8.4812	2.8271	20.0829	.0000
Within Groups	16	2.2523	.1408		
Total	19	10.7335			

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
4.8545	3	16	.014

----- O N E W A Y -----

Variable CSS  
By Variable TRT TREATMENT GROUP

Multiple Range Test

LSD Procedure  
Ranges for the .100 level -

2.47 2.47 2.47

The ranges above are table ranges.  
The value actually compared with Mean(J)-Mean(I) is..  
.2653 \* Range \* Sqrt(1/N(I) + 1/N(J))

(\*) Denotes pairs of groups significantly different at the .100 level

Mean	Group	1	2	3	4
.6018	Grp 1				
1.0139	Grp 2				
1.5089	Grp 3	*	*		
2.3499	Grp 4	*	*	*	

TIME-SEQUENCED BIOACCUMULATION

----- O N E W A Y -----

Variable LOGCSS  
By Variable TRT TREATMENT GROUP

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	1.1064	.3688	11.8800	.0002
Within Groups	16	.4967	.0310		
Total	19	1.6032			

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
3.6906	3	16	.034

----- O N E W A Y -----

Variable LOGCSS  
By Variable TRT TREATMENT GROUP

Multiple Range Test

LSD Procedure  
Ranges for the .100 level -

2.47 2.47 2.47

The ranges above are table ranges.  
The value actually compared with Mean(J)-Mean(I) is..  
.1246 \* Range \* Sqrt(1/N(I) + 1/N(J))

(\*) Denotes pairs of groups significantly different at the .100 level

```

          G G G G
          r r r r
          p p p p

Mean      Group    1 2 3 4
-.2574   Grp 1
-.0543   Grp 2      *
.1747    Grp 3      * *
.3671    Grp 4      * *

```

TIME-SEQUENCED BIOACCUMULATION

----- ONEWAY -----

Variable RANKIT NORMAL of CSS using BLOM  
By Variable TRT TREATMENT GROUP

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	12.3395	4.1132	12.4310	.0002
Within Groups	16	5.2941	.3309		
Total	19	17.6336			

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
1.8772	3	16	.174

----- ONEWAY -----

Variable RANKIT NORMAL of CSS using BLOM  
By Variable TRT TREATMENT GROUP

Multiple Range Test

LSD Procedure

Ranges for the .100 level -

2.47 2.47 2.47

The ranges above are table ranges.

The value actually compared with Mean(J)-Mean(I) is..  
.4067 \* Range \* Sqrt(1/N(I) + 1/N(J))

(\*) Denotes pairs of groups significantly different at the .100 level

```

          G G G G
          r r r r
          p p p p

Mean      Group    1 2 3 4
-.8871   Grp 1
-.4434   Grp 2
.1179    Grp 3      *
1.2126   Grp 4      * * *

```

TIME-SEQUENCED BIOACCUMULATION

t-tests for independent samples of TRT TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
CSS				
Reference	5	.6018	.235	.105
Sediment 1	5	1.0139	.572	.256

Mean Difference = -.4121

Levene's Test for Equality of Variances: F= 9.363 P= .016

Variances	t-value	df	2-Tail Sig	SE of Diff	95% CI for Diff
Equal	-1.49	8	.175	.277	(-1.050, .226)
Unequal	-1.49	5.31	.193	.277	(-1.123, .299)

Variable	Number of Cases	Mean	SD	SE of Mean
LOGCSS				
Reference	5	-.2574	.219	.098
Sediment 1	5	-.0543	.261	.117

Mean Difference = -.2031

Levene's Test for Equality of Variances: F= .600 P= .461

t-test for Equality of Means					95%
Variances	t-value	df	2-Tail Sig	SE of Diff	CI for Diff
Equal	-1.33	8	.219	.152	(-.554, .148)
Unequal	-1.33	7.77	.220	.152	(-.554, .148)

t-tests for independent samples of TRT TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
RANKIT NORMAL of CSS using BLOM				
Reference	5	-.8871	.592	.265
Sediment 1	5	-.4434	.831	.371

Mean Difference = -.4437

Levene's Test for Equality of Variances: F= 1.113 P= .322

t-test for Equality of Means					95%
Variances	t-value	df	2-Tail Sig	SE of Diff	CI for Diff
Equal	-.97	8	.359	.456	(-1.496, .608)
Unequal	-.97	7.23	.362	.456	(-1.522, .635)

#### TIME-SEQUENCED BIOACCUMULATION

t-tests for independent samples of TRT TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
CSS				
Reference	5	.6018	.235	.105
Sediment 2	5	1.5089	.236	.106

Mean Difference = -.9070

Levene's Test for Equality of Variances: F= .009 P= .926

t-test for Equality of Means					95%
Variances	t-value	df	2-Tail Sig	SE of Diff	CI for Diff
Equal	-6.08	8	.000	.149	(-1.251, -.563)
Unequal	-6.08	8.00	.000	.149	(-1.251, -.563)

#### TIME-SEQUENCED BIOACCUMULATION

Variable	Number of Cases	Mean	SD	SE of Mean
LOGCSS				
Reference	5	-.2574	.219	.098
Sediment 2	5	.1747	.065	.029

Mean Difference = -.4321

Levene's Test for Equality of Variances: F= 2.510 P= .152

t-test for Equality of Means					95%
Variances	t-value	df	2-Tail Sig	SE of Diff	CI for Diff
Equal	-4.23	8	.003	.102	(-.668, -.197)
Unequal	-4.23	4.69	.009	.102	(-.695, -.170)

t-tests for independent samples of TRT TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
RANKIT NORMAL of CSS using BLOM				
Reference	5	-.8871	.592	.265
Sediment 2	5	.1179	.298	.133

Mean Difference = -1.0050

Levene's Test for Equality of Variances: F= 1.020 P= .342

Variances	t-test for Equality of Means			SE of Diff	95% CI for Diff
	t-value	df	2-Tail Sig		
Equal	-3.39	8	.009	.296	(-1.688, -.322)
Unequal	-3.39	5.91	.015	.296	(-1.730, -.280)

TIME-SEQUENCED BIOACCUMULATION

t-tests for independent samples of TRT TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
CSS				
Reference	5	.6018	.235	.105
Sediment 3	5	2.3499	.353	.158

Mean Difference = -1.7480

Levene's Test for Equality of Variances: F= 2.375 P= .162

Variances	t-test for Equality of Means			SE of Diff	95% CI for Diff
	t-value	df	2-Tail Sig		
Equal	-9.21	8	.000	.190	(-2.186, -1.310)
Unequal	-9.21	6.96	.000	.190	(-2.197, -1.299)

Variable	Number of Cases	Mean	SD	SE of Mean
LOGCSS				
Reference	5	-.2574	.219	.098
Sediment 3	5	.3671	.065	.029

Mean Difference = -.6246

Levene's Test for Equality of Variances: F= 2.184 P= .178

Variances	t-test for Equality of Means			SE of Diff	95% CI for Diff
	t-value	df	2-Tail Sig		
Equal	-6.12	8	.000	.102	(-.860, -.389)
Unequal	-6.12	4.70	.002	.102	(-.887, -.362)

t-tests for independent samples of TRT TREATMENT GROUP

Variable	Number of Cases	Mean	SD	SE of Mean
RANKIT NORMAL of CSS using BLOM				
Reference	5	-.8871	.592	.265
Sediment 3	5	1.2126	.441	.197

Mean Difference = -2.0997

Levene's Test for Equality of Variances: F= .113 P= .745

t-test for Equality of Means					95%
Variances	t-value	df	2-Tail Sig	SE of Diff	CI for Diff
Equal	-6.36	8	.000	.330	(-2.861, -1.338)
Unequal	-6.36	7.40	.000	.330	(-2.881, -1.319)

TIME-SEQUENCED BIOACCUMULATION

POWER OF LSD TO DETECT A TRUE POPULATION DIFFERENCE (D)  
ABOVE REFERENCE MEAN C<sub>SS</sub>

NUMBER OF REPLICATES	REFERENCE MEAN C <sub>SS</sub>	MEAN SQUARE ERROR	DEGREES OF FREEDOM DF	T VALUE FOR (1-ALPHA=0.95, DF)
5.00	.60180	.14080	16.00	1.74588

POWER OF LSD TO DETECT % INCREASE IN C<sub>SS</sub> ABOVE REFERENCE MEAN C<sub>SS</sub> GIVEN N, MSE AND DF SHOWN ABOVE

% INCREASE IN C <sub>SS</sub> ABOVE REFERENCE	DREDGED SEDIMENT C <sub>SS</sub>	D	T VALUE FOR (1-BETA, DF)	POWER (1-BETA)
10.00	.66198	.0602	-1.49230	.07746
25.00	.75225	.1504	-1.11192	.14097
50.00	.90270	.3009	-.47796	.31848
100.00	1.20360	.6018	.78995	.77767
200.00	1.80540	1.2036	3.32579	.99780
300.00	2.40720	1.8054	5.86162	.99999

MINIMUM DREDGED SEDIMENT C<sub>SS</sub> THAT CAN BE DETECTED BY LSD AS SIGNIFICANT GIVEN SPECIFIED POWER AND N, MSE AND DF SHOWN ABOVE

POWER (1-BETA)	D	DREDGED SEDIMENT C <sub>SS</sub>	% INCREASE IN C <sub>SS</sub> ABOVE REFERENCE	T VALUE FOR (1-BETA, DF)
.50000	.41433	1.01613	68.85	0.00000
.60000	.47546	1.07726	79.01	.25760
.70000	.54130	1.14310	89.95	.53501
.80000	.61953	1.22133	102.95	.86467
.90000	.73157	1.33337	121.56	1.33676
.95000	.82866	1.43046	137.70	1.74588
.99000	1.02744	1.62924	170.73	2.58349

COMPARISON OF MEAN DREDGED SEDIMENT C<sub>SS</sub> WITH ACTION LEVEL: UPPER CONFIDENCE LIMITS (UCL) WHEN VARIANCES ARE EQUAL

TRT	MEANCSS	UCL1	MSE	TALPHA1	DF	DMIN1
2.00	1.01390	1.30688	.14080	1.74588	16.00	.29298
3.00	1.50890	1.80188	.14080	1.74588	16.00	.29298
4.00	2.34990	2.64288	.14080	1.74588	16.00	.29298

COMPARISON OF MEAN DREDGED SEDIMENT C<sub>SS</sub> WITH ACTION LEVEL: UPPER CONFIDENCE LIMITS (UCL) WHEN VARIANCES ARE UNEQUAL

TRT	MEANCSS	UCL2	S2	TALPHA2	N	DMIN2
2.00	1.01390	1.55924	.32718	2.13185	5.00	.54534
3.00	1.50890	1.73419	.05584	2.13185	5.00	.22529
4.00	2.34990	2.68664	.12475	2.13185	5.00	.33674

POWER TO DETECT % DECREASE IN C<sub>SS</sub> BELOW ACTION LEVEL OF 2 µg/g GIVEN N, MSE, AND DF SHOWN ABOVE

% DECREASE BELOW ACTION LEVEL	DREDGED SEDIMENT C <sub>SS</sub>	D	T VALUE FOR (1-BETA, DF)	POWER (1-BETA)
10.00	1.80000	.20	-.55405	.29268
20.00	1.60000	.40	.63778	.73192
30.00	1.40000	.60	1.82960	.95634
40.00	1.20000	.80	3.02143	.99585
50.00	1.00000	1.00	4.21326	.99966

# REPORT DOCUMENTATION PAGE

*Form Approved*  
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

<b>1. AGENCY USE ONLY (Leave blank)</b>	<b>2. REPORT DATE</b> July 1996	<b>3. REPORT TYPE AND DATES COVERED</b> Final report	
<b>4. TITLE AND SUBTITLE</b> Applications Guide for Statistical Analyses in Dredged Sediment Evaluations		<b>5. FUNDING NUMBERS</b>	
<b>6. AUTHOR(S)</b> Joan U. Clarke, Dennis L. Brandon			
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> U.S. Army Engineer Waterways Experiment Station 3909 Halls Ferry Road Vicksburg, MS 39180-6199		<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b> Miscellaneous Paper D-96-2	
<b>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> U.S. Army Corps of Engineers Washington, DC 20314-1000		<b>10. SPONSORING/MONITORING AGENCY REPORT NUMBER</b>	
<b>11. SUPPLEMENTARY NOTES</b>  Available from National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161.			
<b>12a. DISTRIBUTION / AVAILABILITY STATEMENT</b>  Approved for public release; distribution is unlimited.		<b>12b. DISTRIBUTION CODE</b>	
<b>13. ABSTRACT (Maximum 200 words)</b>  Dredged sediment evaluations often require statistical analysis of chemical or biological test results. However, the resulting data are frequently problematic for standard statistical procedures because of improper experimental design, insufficient replication, failure to meet statistical test assumptions, outliers, and missing or below detection limit observations. Such nonideal data can seriously affect the error rates of statistical tests. This in turn increases the likelihood of drawing false inferences concerning the potential of a dredged sediment for adverse biological effects. This report explores the impact of nonideal data on the performance of statistical tests recommended for dredged sediment evaluations. Procedures are discussed for basic experimental design, testing assumptions and outliers, assessing statistical power, analyzing less-than detection limit data, and interpreting statistical test results.			
<b>14. SUBJECT TERMS</b> See reverse.		<b>15. NUMBER OF PAGES</b> 242	
		<b>16. PRICE CODE</b>	
<b>17. SECURITY CLASSIFICATION OF REPORT</b> UNCLASSIFIED	<b>18. SECURITY CLASSIFICATION OF THIS PAGE</b> UNCLASSIFIED	<b>19. SECURITY CLASSIFICATION OF ABSTRACT</b>	<b>20. LIMITATION OF ABSTRACT</b>

#### 14. Subject Terms (Concluded).

Dredged sediment evaluation  
Equality of variances assumption  
Experimental design  
Less-than detection limit data  
Nonideal data  
Normality assumption  
Outliers  
Power  
Statistical analysis  
Statistical error rates