

1.1

SAMARAN REPORT

HOLONO CHURCHEN I

MICROCOPY RESOLUTION TEST CHART NATIONAL BUREAU OF STANDARDS-1963-A DESIGN AND ANALYSIS OF CHRONIC AQUATIC TESTS OF TOXICITY

AD

NOV 2

1983

Final Report

By

Paul I. Feder Department of Statistics, The Ohio State University and Applied Statistics Group, Battelle Columbus Laboratories

William J. Collins Department of Entomology, The Ohio State University

(15 September 1979 - 31 October 1980)

Supported by

U.S. Army Medical Research and Development Command Fort Detrick, Frederick, Maryland 21701

Contract No. DAMD 17-79-C-9150

The Ohio State University Research Foundation 1314 Kinnear Road Columbus, Ohio 43212

FILE COPY

40-4134 309

Contract Officer's Technical Representative: Paul H. Gibbs U.S. Army Medical Bioengineering Research and Development Laboratory Fort Detrick, Frederick, Maryland 21701

Approved for public release; distribution unlimited

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

REPORT DOCUMENTAT		READ INSTRUCTIONS BEFORE COMPLETING FORM
. REPORT NUMBER		0. 3. RECIPIENT'S CATALOG NUMBER
	AD. A134	304
4. TITLE (and Subtitio)		5. TYPE OF REPORT & PERIOD COVERE
DESIGN AND ANALYSIS OF CHRONIC	C AQUATIC TESTS	Final Report 9/15/79 - 10/31/80
OF TOXICITY		6. PERFORMING ORG. REPORT NUMBER
		761860/712404
T. AUTHOR()		8. CONTRACT OR SHARK NUMBER(s)
Paul I. Feder		DUM 17 70 0 0150
William J. Collins		DAMD 17-79-C-9150
PERFORMING ORGANIZATION NAME AND AD	DRESS	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
The Ohio State University Rese	earch Foundation	621102A.3E161102BS04.
1314 Kinnear Road		00.048
Columbus, Ohio 43212		12. REPORT DATE
1. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Medical Research and		
Medical Bioengineering Research	-	• •
Fort Detrick Frederick, Maryla	and 21701	415
4. MONITORING AGENCY NAME & ADDRESS(II (	different from Controlling Office	) 15. SECURITY CLASS. (of this report)
		Unclassified
		154. DECLASSIFICATION DOWNGRADING SCHEDULE
		SCHEDULE
Approved for public release; o		
7. DISTRIBUTION STATEMENT (of the abstract o	ntered in Block 20, if different	(rom Report)
7. DISTRIBUTION STATEMENT (of the abstract o 8. SUPPLEMENTARY NOTES 8. KEY WORDS (Continue on reverse side if neces	eary and identify by block numb	from Report)
7. DISTRIBUTION STATEMENT (of the abstract o 8. SUPPLEMENTARY NOTES 9. KEY WORDS (Continue on reverse side if neces chronic aquatic toxicity test	eary and identify by block numb	(rom Report)
7. DISTRIBUTION STATEMENT (of the abstract of 8. SUPPLEMENTARY NOTES 9. KEY WORDS (Continue on reverse side if neces chronic aquatic toxicity tests standardized data reporting tests of hypothesis	eary and identify by block numb b multiple com dose respons computer pro	er) parisons e estimation grams
7. DISTRIBUTION STATEMENT (of the abstract of 8. SUPPLEMENTARY NOTES 9. KEY WORDS (Continue on reverse side if neces chronic aquatic toxicity tests standardized data reporting	eary and identify by block numb dose respons computer pro statistical	(rom Report) er) parisons e estimation grams analysis
7. DISTRIBUTION STATEMENT (of the abstract of 8. SUPPLEMENTARY NOTES 9. KEY WORDS (Continue on reverse side if neces chronic aquatic toxicity tests standardized data reporting tests of hypothesis	eary and identify by block numb by block numb multiple com dose respons computer pro statistical fathead minn	er) parisons e estimation grams analysis ows
<ul> <li>DISTRIBUTION STATEMENT (of the abstract of supplementary notes</li> <li>SUPPLEMENTARY NOTES</li> <li>KEY WORDS (Continue on reverse side if necession of the standardized data reporting tests of hypothesis confidence intervals</li> <li>ABSTRACT (Continue on reverse side if necession)</li> </ul>	eary and identify by block numb some of the series of the	<pre>(rom Report)  er) parisons e estimation grams analysis ows "" of the planning of, reportir arising from chronic aquatic ssion emphasizes analyses and</pre>
<ul> <li>DISTRIBUTION STATEMENT (of the abstract of supplementary notes</li> <li>SUPPLEMENTARY NOTES</li> <li>KEY WORDS (Continue on reverse side if necess chronic aquatic toxicity tests standardized data reporting tests of hypothesis confidence intervals</li> <li>ABSTRACT (Continue on reverse side if necess The present investigation esults from, and the statisticates of toxicity with fathead is ata displays for the interpretata and quantitative weight data and the statist of necess the report consists of number of the statisticates of toxicity ata and quantitative weight data and the statist of the interpretation of the statistic of the statistic of the statistic of the statistic of the interpretation of the statistic of the statistis of the statistic of the statistic of the statistic of the st</li></ul>	sary and identify by block numb sary and identify by block numb s multiple com dose respons computer pro statistical <u>fathead minn</u> sary and identify by block number n considers aspects al analysis of data minnows. The discu ation of qualitativ ta.	<pre>(rom Report) er) parisons e estimation grams analysis ows '' of the planning of, reportin arising from chronic aquatic ssion emphasizes analyses and e mortality and abnormality Section I-IV describe the</pre>
<ul> <li>DISTRIBUTION STATEMENT (of the abstract of supplementary notes</li> <li>SUPPLEMENTARY NOTES</li> <li>KEY WORDS (Continue on reverse side if necess chronic aquatic toxicity tests standardized data reporting tests of hypothesis confidence intervals</li> <li>ABSTRACT (Continue on reverse side if necess The present investigation esults from, and the statisticates of toxicity with fathead rata displays for the interpretata and quantitative weight data</li> </ul>	sary and identify by block numb mitered in Block 20, if different sary and identify by block numb multiple com dose respons computer pro statistical fathead minn sary and identify by block numb n considers aspects al analysis of data minnows. The discu ation of qualitativ ta. > ineteen sections. d present a set of OBSOLETE	<pre>trom Report)  er) parisons e estimation grams analysis ows of the planning of, reportin arising from chronic aquatic ssion emphasizes analyses and e mortality and abnormality Section I-IV describe the standardized data reporting -</pre>
<ul> <li>DISTRIBUTION STATEMENT (of the abstract of supplementary notes</li> <li>SUPPLEMENTARY NOTES</li> <li>KEY WORDS (Continue on reverse side if necess chronic aquatic toxicity tests standardized data reporting tests of hypothesis confidence intervals</li> <li>ABSTRACT (Continue on reverse side if necess The present investigation esults from, and the statisticates of toxicity with fathead is ata displays for the interpretata and quantitative weight data and the statist of necess the report consists of number of the statisticates of toxicity ata and quantitative weight data and the statist of the interpretation of the statistic of the statistic of the statistic of the statistic of the interpretation of the statistic of the statistis of the statistic of the statistic of the statistic of the st</li></ul>	eary and identify by block numbers by and identify by block numbers s multiple com dose respons computer pro statistical fathead minn eary and identify by block numbers a considers aspects al analysis of data minnows. The discu ation of qualitative ta. ) ineteen sections. d present a set of OBSOLETE Unc	<pre>trom Report) er) parisons e estimation grams analysis ows '' of the planning of, reportir arising from chronic aquatic ssion emphasizes analyses and e mortality and abnormality Section I-IV describe the</pre>

# Block 20 (Abstract) - continued

STATES IN

formats, and contain some suggestions for further improving and standardizing the test procedure. Section V briefly describes the early life stage testing procedure and the resulting data. Sections VI-XIX discuss various topics pertaining to the statistical analysis of such data and to statistical aspects in the design of toxicity tests. Techniques discussed include preliminary graphical displays, tests for homogeneity among tanks within treatment groups, adjustments to account for tank to tank heterogeneity within groups, outlier detection procedures, overall tests for treatment effects, multiple comparisons and pairwise confidence intervals of treatment groups with the control group, analysis of variance, regression analysis and parametric and nonparametric dose response curve fits along with associated estimates of "safe concentrations". All suggested procedures are illustrated with real examples based on aquatic toxicity data. Several newly developed, specialized computer programs are described to carry out nonstandard analyses  $\frac{1}{2}$ 

# Unclassified SECURITY CLASSIFICATION OF THIS PAGE(When Date Entered)

# DESIGN AND ANALYSIS OF CHRONIC AQUATIC TESTS OF TOXICITY

AD

Final Report

By

Paul I. Feder Department of Statistics, The Ohio State University and Applied Statistics Group, Battelle Columbus Laboratories

William J. Collins Department of Entomology, The Ohio State University

(15 September 1979 - 31 October 1980)

### Supported by

U.S. Army Medical Research and Development Command Fort Detrick, Frederick, Maryland 21701

Contract No. DAMD 17-79-C-9150

The Ohio State University Research Foundation 1314 Kinnear Road Columbus, Ohio 43212

Contract Officer's Technical Representative: Paul H. Gibbs U.S. Army Medical Bioengineering Research and Development Laboratory Fort Detrick, Frederick, Maryland 21701

Approved for public release; distribution unlimited

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

### EXECUTIVE SUMMARY

### Purpose

During the past several decades problems of environmental contamination have become increasingly important both from the scientific and the legal standpoints. In recent years a great deal of attention has been directed to the potential toxicity to aquatic organisms of chemicals discharged into water bodies.

The U.S. Army, through activities such as munitions manufacturing, operates a number of plants that produce, consume, or discharge a variety of chemical substances. Some of these discharges enter bodies of water inhabited by various aquatic species. Thus the Army must provide the USEPA with safety data concerning the levels of such discharges and the possible extent of resulting surface water and ecological contamination. In order to develop such data the Army conducts both intramural and extramural programs of aquatic toxicity testing.

Considerable amounts of time, money, and manpower are expended by the Army in such aquatic toxicity testing programs. To make these programs more efficient and more effective, the need has been felt for a reexamination of some of the standard methods used. This has been especially true of statistical methods involved in the design of testing programs and the analysis of resulting data. This study is an effort to make some progress in those directions.

The results in this study indicate areas where the conduct of and the summarization and reporting of results from chronic aquatic toxicity tests can be further standardized and made easier to understand. A number of the statistical approaches and procedures discussed and/or developed in this study have not to the authors' knowledge been previously applied to aquatic toxicity data. These improved methods provide increased information, as compared with standard methodology, about the structure, relations, and anomolies in the data. They enhance the sensitivity of statistical analyses, so that greater precision of results can be obtained without increasing the amount of testing. In brief this study provides methods that should improve the reporting and statistical analysis of data from chronic aquatic toxicity tests. This will enhance the sensitivity of conclusions that can be derived from these tests, thereby increasing their efficiency.



### Approach

All the statistical procedures discussed in this report are illustrated with examples based on real data from chronic tests with fathead minnows. At the outset of the study the principal investigators visited the USEPA Environmental Research Laboratory at Duluth to become oriented to the apparatus and procedures used in chronic toxicity tests. Discussions were held with various investigators concerning the details of their studies. Some of these investigators provided us with illustrative data to be used in our subsequent work.

A number of sets of experimental data were received from Duluth. The literature pertaining to chronic toxicity tests in general and to those tests in particular was reviewed and discussed between statistician and toxicologist. Based on an understanding of experimental procedures it was possible to then start considering the statistical aspects of the problems. The statistical procedures discussed in the body of this report represent a combination of methods taken from the statistical literature where appropriate or developed especially for aquatic toxicity testing applications where standard procedures were felt not to be the most appropriate.

### **Results**

Arguments for the use of standardized fish stocks and standardized data reporting formats are presented. Aspects of the statistical analysis of toxicity data are discussed and the suggested procedures are illustrated with examples based on fish toxicity studies. Data analysis topics discussed include: graphical displays, preliminary tests of tank to tank heterogeneity within treatment groups, preliminary outlier detection tests, overall tests of heterogeneity in response rates across treatment groups, treatment group-control group pairwise multiple comparison procedures, the fitting of standard and nonstandard dose response curve models, analysis of variance and multiple regression analyses on quantitative responses, statistical power and estimation precision to be expected for various levels of sample size and suggestions for <u>unequal</u> allocation of experimental effort across treatment groups with greater effort expended on the control group and lower treatment groups.

### Conclusions and Recommendations

- 1. The USEPA should revise and update the standard procedure for life cycle tests on fathead minnows.
- 2. Standardized data reporting sheets are a very useful adjunct to the categorization and analysis of chronic toxicity test data.

- Guidelines on disposal of potentially hazardous effluent from chronic toxicity tests should be incorporated in the procedure.
- 4. Detailed procedures for chemical analysis and quality assurance of chemical data should be incorporated in the procedure.
- 5. Some of the "standard" methods currently used for analyzing data from aquatic toxicity tests can and should be modified. The data should first be graphed, outlying observations or groups of observations should be located and the reason for their aberrant behavior determined, and tests for heterogeneity among tanks within groups should be carried out. Based on the results of these preliminary inferences, the data should be modified or adjusted to account for possible heterogeneity or aberrant values before going on to the inferences of primary interest.
- 6. If hypothesis tests are to be used to compare the treatment group and control group responses they should be one sided tests which are sensitive to monotone alternatives, rather than overall analysis of variance type "shotgun" tests.
- 7. Multiple comparison procedures and confidence intervals procedures should be used to determine specifically which treatment groups have responses which differ from the control group responses and whether the differences are of a specified <u>biological significance</u>. Significance tests, by themselves, are not adequate to define an "MATC" (i.e., maximum acceptable tolerable concentration). Perhaps a confidence bound should be routinely constructed at the MATC to determine just how much worse than the control group the response at that concentration could conceivably be. In general, confidence intervals impart much more information than hypothesis tests and should be routinely used.
- 8. A good way to place monotone response structure on the problem, to amooth the data, and to convert a hypothesis testing problem into an estimation problem is to fit dose response curve models to the data and to define the "safe" concentration as that which results in no more than a specified increment in response from the control group. A number of nonstandard variants on the "standard" dose response models discussed in the literature may be useful. A nonparametric approach to dose response estimation is feasible, has been implemented in a computer program, and may be preferable on occassion to some of the standard parametric dose response models.

iy

9. Statistical power and estimation precision depend both on the number of tanks run per group and the number of fish per tank. In the presence of substantial tank to tank heterogeneity the effective sample size may be more nearly the number of tanks than the number of fish. Thus in the presence of tank to tank heterogeneity, diminishing returns result from increasing the number of fish used without also increasing the number of fish tanks per group.

- 10. Under certain circumstances it is sensible to allocate experimental resources so that the control group and lower concentration groups receive more tanks and fish than the higher concentration groups. This results in greater inference sensitivity in the region of the MATC. Proportional diluters should be modified to permit such asymmetrical allocations of tanks, at the discretion of the investigator.
- 11. Statistical power or statistical precision goals should be stated as part of the protocol for each individual toxicity test and sample sizes should be determined accordingly.

## PUBLICATIONS AND PRESENTATIONS SUPPORTED BY THE CONTRACT

### **Reports and Papers**

- Feder, P. I. and Sherrill, M.C., "A Computer Program to Calculate Nonparametric Lower Confidence Bounds on Safe Concentration in Quantal Response Toxicity Tests, "Technical Report No.215, Department of Statistics, The Ohio State University, Columbus, Ohio, September 1980.
- Feder, P.I. and Willavize, Susan A., "EXAX2 -- A Computer Program to Compare Binomial Proportions, Program Description and Card Input Information, "Technical Report No.216, Department of Statistics, The Ohio State University, Columbus, Ohio, September 1980.
- Feder, P.I. and Collins, W.J., "Considerations in the Design and Analysis of Chronic Aquatic Tests of Toxicity," <u>Aquatic Toxicology</u> and Hazard Assessment: Fifth Conference, ASTM STP 766, J.G. Pearson, R.B. Foster and W.E. Bishop, Eds., American Society for Testing and Materials, 1982, pp.32-68.

### Presentations

- 1. "Considerations in the Design and Analysis of Chronic Aquatic Tests of Toxicity," Feder, P.I. and Collins, W.J. Fifth Annual Symposium on Aquatic Toxicology, American Society for Testing and Materials, Committee E35, October 1980.
- 2. "Design and Analysis of Aquatic Tests of Toxicity," Feder, P.I. 36th Annual Conference on Applied Statistics, American Society for Quality Control and American Statistical Association, December 1980.

# PERSONNEL RECEIVING SUPPORT UNDER THE CONTRACT

Paul I. Feder, Department of Statistics William J. Collins, Department of Entomology Susan Willavize, Department of Statistics Mary Sherrill, Department of Statistics Gerald Skinner, Department of Statistics

### Foreword

The investigators gratefully acknowledge the assistance and expert contributions to the project of the following graduate assistants: Susan Willavize and Mary Sherrill, Department of Statistics and Gerald Skinner, Department of Zoology. Charles Stephan of the USEPA Environmental Research Laboratory, Duluth, Minnesota was kind enough to arr ye for us to talk to a number of investigators at his laboratory about conduct د of toxicity tests. Duane Benoit, Dave Defoe, Gary Holcombe. Ered Jarvinen, Gary Phipps and Robert Spehar of that laboratory an cary Kimball, formerly at the University of Minnesota, sent us a r er of data sets to be used to illustrate the statistical methodolog cussed in the report. EPA, Environmental Research Laboratory - Dult - \_ave permission to use these data for illustrative purposes in the report. Paul Gibbs, Jerry Highfill, J. Gareth Pearson and William van der Schalie of the U.S. Army Medical Bioengineering Research and Development Laboratory provided guidance, suggestions, and coordination throughout the conduct of the study. Wayne Nelson kindly granted permission to reproduce his charts for confidence limits for the ratio of two Poisson means.

# TABLE OF CONTENTS

- - -

		Page
Executiv	e Summary	ii
Publicat	ions and Presentations	vi
Foreword	i	viii
Table of	Contents	ix
Figures		xiv
Tables		xviii
Introduc	tion	xix
Section		
Ι.	Assemble and Evaluate Information on Test Methods	1
11.	Categorizing Data Sets	2
111.	Organization of Data: Standardized Data Records	3
	A. Composite of Experimental Conditions	3
	B. Data Sheets for Separate Categories of Data	5
	C. Transfer of Experimental Data to Standardized Data Sheets	9
IV.	Analysis of and Comments on the Test Procedure	10
	A. Physical System	10
	B. Chemical System	10

ix

Section		Page
v.	Early Life Stage Toxicity Tests	11
	A. Background	11
	B. Data	11
VI.	Preliminary Graphical Data Displays	18
VII.	Testing for Tank to Tank Heterogeneity Within Treatment Groups	27
	A. Background	27
	B. Remarks on Some "Standard" Procedures	27
	C. Separate Heterogeneity Tests Within Treatment Groups	32
VIII.	EXAX2 A Computer Program to Test for Hetero- geneity of Responses Among Tanks Within Treatment Groups	44
IX.	Adjustments to Account for Tank to Tank Hetero- geneity Within Treatment Groups	51
	A. Background, Derivations, and Discussion	51
	B. Illustrations	55
x.	Outlier Detection Procedures	63
	A. Background	63
	B. Applications of Outlier Detection Procedures to Fish Toxicity Data	64
XI.	Testing for Concentration Related Effects	88
	A. Background	88
	B. Chi Square Test of Homogeneity Across Treatment Groups	89
	C. One Sided Tests of Homogeneity Across Treatment Groups	92

Section		Page
XII.	Treatment Group vs Control Group Pairwise Multiple Comparison Procedures	116
XIII.	Confidence Interval Procedures for Comparison of Treatment Group and Control Group Response Rates	120
	A. Introduction	120
	B. Method 1 Asymptotic Approach	122
	C. Method 2 Exact, Small Sample Confidence Intervals	125
	D. Method 3 Poisson Approximation	132
XIV.	Dose Response Curve Estimation Probit Analysis	145
	A. Introduction. Dose Response Estimation vs Hypothesis Testing	145
	B. Probit Analysis Using SAS PROC PROBIT	147
XV.	Dose Response Curve Estimation Maximum Likelihood Estimation by Nonlinear Least Squares Regression	177
XVI.	Nonparametric Lower Confidence Bounds on Safe Concentrations	199
XVII.	Analysis of Quantitative Responses	244
	A. Preliminary Scatterplots	246
	B. Outlier Detection Procedures and Testing for Tank to Tank Heterogeneity Within Groups	246
	C. Multiple Comparison Procedures and Regression Analysis	251
XVIII.	Experimental Design Considerations	267
	A. Assumptions, Additional Variables to Measure, Numbers and Allocations of Tanks	267

xi

÷.

Section		Page
	B. Sample Size and Power Considerations for Qualitative Survival Data	270
	C. Expected Precision for Comparisons of Treatment Group and Control Group Survival Probabilities	277
	D. Power Calculations for Quantitative Weight Response	279
	E. Unequal Allocations of Testing Effort Among Treatment Groups	284
XIX.	Design and Analysis Considerations for Full Life Cycle Tests Distinct from Those for Early Life Stage Tests	288
Reference	es	290
Appendix Examp	AII <sup>*</sup> Early Life Stage Data Sets Used as les in the Body of the Report	294
Probi	AVII* Theoretical Bases of Chi Square and t Based Tests of Tank to Tank Variation Within ment Groups	306
Appendix	AVIII.1* Output from EXAX2 Computer Program	308
	AVIII.2* Description and Instructions for Use AX2 Computer Program	345
	AX <sup>*</sup> Theoretical Bases of Suggested Outlier tion Transformations	362
	AXV* Confidence Interval on Concentration Corresponds to a Given Level of Increase in Resonse Control Group Response	365

X

\* Appendices are numbered so as to indicate their relation to corresponding **Bection** numbers in the body of the report.

# Section

Appendix AXVI.1* Probabilities	Confidence Bounds on Binomial	368
	Description and Instructions for use Lc Dose Response Estimation Computer	
Program	-	369

Page

<sup>\*</sup> Appendices are numbered so as to indicate their relation to corresponding section numbers in the body of the report.

FIGURES

·. \* ·.

Section			Page
V	1.	Standardized data reporting format for results of early life stage test compound C DeFoe	13
	2.	Computerized data format for results of early life stage test compound C DeFoe	16
VI	1.	DeFoe compound C percentage embryo mortality by treatment group	19
	2.	DeFoe compound C percentage fry mortality by treatment group	20
	3.	DeFoe compound C percentage embryo abnormality by treatment group	21
	4.	DeFoe compound C percentage fry abnormality by treatment group	22
	5.	Holcombe and Phipps compound D embryo mortality by treatment group	23
	6.	Holcombe and Phipps compound D fry mortality by treatment group	24
	7.	Holcombe and Phipps compound D embryo abnormality by treatment group	25
	8.	Holcombe and Phipps compound D fry abnormality by treatment group	26
VII	1.	Probit fit to individual responses DeFoe compound C (excluding control group)	34
	2.	Probit fit to responses pooled across tanks within groups DeFoe compound C (excluding control group)	35
	3.	Chi square test of homogeneity based on individual responses DeFoe compound C fry mortality	36
	4.	Chi square test of homogeneity based on responses pooled across tanks within groups DeFoe compound C fry mortality	37

xiv

Ā

Section		Page
5–:	10. Chi square test of homogeneity of percent fry mortality in groups 1-6. Data from Holcombe and Phipps compound D	38
x	<ol> <li>Normal probability plot of standardized values based on DeFoe compound C embryo mortality data to detect outlying responses</li> </ol>	83
	<ol> <li>Normal probability plot of standardized values based on DeFoe compound C fry mortality data to detect outlying responses</li> </ol>	84
	3. Normal probability plot of standardized values based on Holcombe and Phipps compound D embryo mortality data to detect outlying responses.	85
	4. Normal probability plot of standardized values based on Holcombe and Phipps compound D fry mortality data to detect outlying responses	86
	5. Normal probability plot of standardized values based on Jarvinen compound B embryo mortality data to detect outlying responses	87
XI	<ol> <li>SAS PROC FREQ output from chi square test for homogeneity across groups applied to fry mortality data from Holcombe and Phipps test on compound D</li> </ol>	101
2-	-7. EXAX2 output chi square test of homo- geneity across groups	102
	8. Input information required for program BMDP1F	108
<b>9</b> –:	11. BMDP1F output from one sided measure of association test on artificial data	109
12-3	15. BMDP1F output on real data	112
XIII 1-	-2. EXAX2 output from calculation of exact, small sample confidence intervals on the odds ratios between treatments and control	137

<u>Secti</u>	lon		Page
	3-5.	Poisson two sample confidence interval charts 90, 95, 99 percent (from Nelson, 1969)	142
XIV	1.	Schematic representation of dose response curve formulation	159
	2.	Basic data	160
	3-6.	Output from PROC PROBIT fit to DeFoe fry mortality data arithmetic concentration	161
	7-8.	PROC PROBIT fit to DeFoe fry mortality data no background correction	165
	9.	PROC PROBIT fit to DeFoe fry mortality data logarithmic transformation of concentration	167
	10.	Basic data	168
	11-14.	Output from PROC PROBIT fit to Holcombe and Phipps compound D fry mortality data arithmetic	
		concentration, logarithmic concentration	169
	15-18.	Plots of observed and predicted based on probit fit arithmetic concentration, logarithmic concentration	173
XV	1.	Holcombe and Phipps compound D mortality and concentration data pooled	), 190
	2-4.	Output from SAS PROC NLIN applied to Holcombe and Phipps data untransformed concentration	191
	5-7.	Logit fit to Holcombe and Phipps data using SAS PROC NLIN untransformed concentration	194
	8.	Use of SAS PROC NLIN to fit nonstandard dose response model to data	197
	9.	Output from SAS PROC NLIN to fit nonstandard dose response model to data	198
XVI	1.	Schematic description of nonparametric dose response estimation procedure Holcombe and Phipps compound D fry mortality data	241
	2-3.	Benoit compound A fry mortality vs arith- metic concentration, logarithmic concentration	242

• . • . •

۰.

 . . .

.

•

.

÷.,

· ·

xvi

Section		Page	
XVII	1.	Average weight per tank vs group	254
	2.	Standard deviation of weight per tank vs group	255
	3.	Average weight per tank vs concentration	256
	4.	Average weight per tank vs log concentration	257
	5-6.	Analysis of variance fit of individual weights (logarithmic weights) on group and tank within roup. Holcombe and Phipps compound D fry mortality data	258
	7.	Expected mean squares in analysis of variance fits to Holcombe and Phipps fry mortality data	260
	8.	Analysis of variance fit of average weights per tank on group. Holcombe and Phipps compound D fry mortality data	261
	9.	Residuals vs group	262
	10.	Logarithmic residuals vs group	263
	11.	Normal probability plot of residuals from untransformed weights	264
	12.	Normal probability plot of residuals from logarithmic weights	265
	13.	Regression fit of weight gain on log (1 + CONC). Holcombe and Phipps compound D fry mortality data	266
XVIII	1.	Effective sample size per group vs sample size for various values of $\boldsymbol{\theta}$	287

Ī

₹...

٠.

TABLES

• ... ... ... ...

· · ·

Section			Page
VII	1-2	DeFoe compound C fry mortality data probit fit to groups 2, 3, 4, 5, 6, individual tanks, pooled tanks within groups	30
IX	1.	Effective sample sizes and responses in Holcombe and Phipps compound D fry mortality data after adjustment for tank to tank heterogeneity	58
	2.	Effective sample sizes and responses in Jarvinen compound B embryo mortality data after adjustment for tank to tank heterogeneity	62
XVIII	1.	Power of one sided pairwise comparisons of survival rates between control group and treatment groups. $\alpha = 0.05$	272
	2.	Expected half lengths of 95 percent two sided confidence intervals for comparisons between treatment group and control group response rates	277
	3.	Power of one sided pairwise comparisons of average weight gains between control group and treatment group. $\alpha = 0.05$	281

17 . K.

- - - N

# INTRODUCTION

During the 1960's and 1970's, environmental contamination in general and water pollution specifically, became increasingly important as legal and scientific problems. Regulatory agencies needed scientific data to support the notion that a problem existed and also needed factual information for establishing tolerance limits for levels of chemical discharges into surface waters. From that need evolved numerous standard toxicity tests, including a test to determine the long term effects of toxicants on a representative fish, the chronic toxicity test with fathead minnows.

Aquatic toxicologists and biologists evolved and refined an effective fish toxicity test. As data were analyzed and experiments designed, the statistical considerations evolved to a more complex level. It became clear that some of the statistical procedures needed for the design of toxicity tests and for the analysis of chronic toxicity data may be novel or unique and should be developed specifically for fish toxicity tests.

Operational activities of the U.S. Army (e.g. munitions manufacture) involve the production, use, and/or discharge of a variety of commercial chemicals. Safety data must be provided to USEPA concerning surface water contamination due to discharges of chemical intermediates or the final product. In-house research of the U.S.Army with the standard fish chronic toxicity test highlighted the need for a reexamination of the standard procedures, especially regarding statistical techniques. The main goals of this project are to suggest statistical procedures for analyzing data arising from such toxicity tests, to provide recommendations for a more accurate, reliable standard procedure, and to facilitate research in aquatic toxicology in general.

This project was initiated as an interdisciplinary investigation of the EPA chronic toxicity test for fathead minnows. It combined the efforts of a toxicologist, a fish specialist, and statisticians. The biologists functioned as advisors to the statisticians in regard to the characteristics and limitations of the test animal, test procedures, and chronic toxicity data and evaluated the test procedure from a toxicological viewpoint.

The statisticians developed procedures for data storage, transformation and analysis, scrutinized published statistical techniques for their applicability to fish toxicity data, and devised new statistical methods for analyzing data from fish toxicity tests when they were felt to be more applicable than the standard methods.

This final report is the synthesis of a one year effort. It discusses both biological and statistical aspects of the planning, conduct, reporting, and data analyses associated with toxicity tests on fathead minnows. Arguments for the use of standardized fish stocks and standardized data reporting formats are presented. Aspects of the statistical analysis of toxicity data are illustrated with examples based on fish toxicity studies. Data analysis topics discussed include: graphical displays, preliminary tests of tank to tank heterogeneity within treatment groups, preliminary outlier detection tests, adjustments in analysis procedures due to tank to tank heterogeneity, overall tests of heterogeneity in response rates across treatment groups, treatment group-control group pairwise multiple comparison procedures, the fitting of standard and nonstandard dose response curve models, analysis of variance and multiple regression analyses on quantitative responses, statistical power and estimation precision to be expected for levels of sample size and suggestions for <u>unequal</u> allocation of experimental effort across treatment groups with greater effort expended on the control group and lower treatment groups.

It is hoped that the results obtained in this study will contribute to better, more reliable toxicity tests and data analyses. This in turn should provide improved tools for the regulation of toxic chemicals in aquatic environments and should suggest fertile areas for further study and development.

## I. ASSEMBLE AND EVALUATE INFORMATION ON TEST METHODS

Although the scope of work specified that design and analysis of <u>chronic</u> toxicity tests would be researched, it became clear early in the project that many aspects of statistical analysis could be pursued using exemplary data from early life stage tests. Consequently, the review of test methods and our literature search included both chronic life-cycle tests and early life stage tests.

Three literature sets were searched by computer using appropriate key words (fathead minnow, toxicity, chronic tests, etc.): Mechanized Information Center, The Ohio State University; Oak Ridge National Laboratory, Oak Ridge, Tenn.; Ohio Environmental Protection Agency, Columbus, Ohio. The latter search encompassed 13 major data base searches (Amer. Chem. Soc., Biol. Abstracts, etc.). The hundreds of citations received were reviewed for relevancy and the important ones were abstracted and filed. Reprints of copies or articles, 64 in all, that were considered to be directly related to future tasks were assembled, catalogued and reviewed in detail.

These publications provided information on the fathead minnow in regard to biology, life cycle events, duration of developmental stages, nutritional information and reproductive characteristics. The papers on test methods provided details of variation in design among investigators and a large amount of experimental toxicity data for reference and further discussion.

Papers and technical reports from E.P.A.-Duluth describing the apparatus [1] and procedure [2] for chronic toxicity tests were reviewed and studied in detail in order to understand the method of exposure, physical arrangement of the delivery system and important variations among investigators, e.g. the syringe delivery method of DeFoe [3]. The literature search and research paper perusal was essential for the toxicology group to authoritatively interpret biological factors, experimental data or test methods in discussions with the statistical group or to suggest limitations in design due to the animal or technique.

# **II. CATEGORIZING DATA SETS**

Early in the project, a number of sets of experimental data from early life stage toxicity tests and chronic life cycle toxicity tests were received from and discussed with researchers in EPA-Duluth. Various aspects of these data were reviewed with the statistical group to clarify experimental procedures such as types of measurements and how they were acquired, if measurements were destructive or nondestructive, the replication of experiments and variation of chemical concentration in the delivery system. These discussions brought out the need for standardized data reporting sheets so that data could be accurately categorized and recorded in a systematic manner for entry into the computer and statistical analysis. A separate section on standardized data sheets is included in this report.

The examples in this report are based on some of these experimental data sets. In particular the data sets from early life stage toxicity tests by

Benoit - compound A DeFoe - compound C Holcombe and Phipps - compound D Jarvinen - compound B

are used. These data sets are listed in Appendix AII.

# III. ORGANIZATION OF DATA: STANDARDIZED DATA RECORDS

The early life stage data and chronic life cycle toxicity data sets were supplied by six different investigators. Each set of chronic data was received in a unique format. Each set of data was reviewed for the experimental procedure (if available) in order to accurately categorize the data for storage on computer and subsequent statistical analysis. Routine questions, e.g. how many days of exposure, and more complex questions, e.g. are replicate tests genuine replications, were not easily resolved by a review of the data sheets, nor was the comparability of the same categories of data in similar experiments among the array of investigators. Standardized data records have merit if they are sufficiently versatile to meet most needs, clearly summarize the exposure conditions and facilitate transformation, computer storage, and statistical analysis of raw data. The latter task is often done by an individual who is not an expert in biological research and unfamiliar with operational details of toxicity tests.

Good laboratory practice regulations (GLP) have been adopted by FDA for nonclinical laboratory studies [4] to assure the quality of data in support of product safety decisions. One component of the GLP deals with specific record-keeping practices for experimental data. The advantages of these required record-keeping practices have been discussed in regard to vertebrate experiments [5] and would apply equally as well to fish toxicity data to the benefit of investigators, statisticians, and regulatory agencies.

Although there is some variation in the design of fish toxicity tests, certain features are almost universal. For example, in a chronic toxicity test, a flow-through apparatus is always used and standard measurements include hatchability of embryos, fish length and weight, survival (mortality), and spawning data. Consequently, standardized recording sheets could be devised for summarizing exposure methods and experimental data.

Our standardized reporting sheets have two components: (a) a descriptive section summarizing the conditions of the experiment with code words or letters to categorize or define data for the statistician, and (b) the raw data record sheet with no calculations or transformations. These record sheets have been designed for data of early life stage tests or chronic life cycle tests.

A. Composite of Experimental Conditions

- 1. Investigator\_\_\_\_\_
- 2. Toxicant \_\_\_\_\_; Source \_\_\_\_\_ and % purity
- 3. Starting Date of Test (Day Zero): / / \_\_\_\_.

4. Selection of embryos: Embryos selected randomly.

\_\_\_\_Embryos examined, only viable eggs incubated.

Other (specify).

5. Are embryos from paired matings? Yes or no.

- 6. Fish I.D.: for use only with paired matings; use a unique identifier here.
- 7. Generation of embryos or fish? Zero or first?

(Note: In chronic tests, some investigators refer to spawnings of exposed adults as first generation embryos, others, second generation. We define zero generation as any stage or form used to start a test and any stages during that generation, including adults. First generation is any of the stages following zero generation adults. One may argue against this system on a biological basis, but it distinguishes between the same stage of separate generations. With this system there is no second generation in a standard chronic life cycle test.)

- Nominal Concentration: identify tanks by nominal concentration of toxicant (mg/1, ug/1), "Solvent control" by "S", "water only control by "W".
- 9. Identify replicate tanks in Nominal Concentration column by "REP"

e.g. 0.25 REP; identify equivalent tanks by "EQ", e.g. 0.25 EQ.

(Note: It should always be clear in original data sheets the relation of replicate tanks of the same nominal concentration, a crucial factor in deciding what statistical procedures to use. We define <u>replicate</u> as the simultaneous exposure of fish to similar concentrations of a chemical that are delivered independently, i.e. two tanks containing nominal concentrations that originate from separate syringes in the delivery system are replicates. We define <u>equivalent</u> as the simultaneous exposure of fish to the same concentration of a chemical that is delivered from a common origin, i.e. groups of fish in several screened compartments of the same tank are equivalent groups, as are fish in different tanks supplied equally by tubing that is split after the final dilution. Replicate ("repeated experiment") does not distort the conventional meaning of that term. The choice of equivalent ("equal in quantity") was the best approximation of what occurs.

 Tank I.D.: identify multiple tanks of the same nominal concentration and type (REP or EQ) by capital letters, e.g. Rep-A, Rep-B, etc.

- 11. Identify simultaneous incubations (same day, same tank) of embryos from <u>different</u> spawnings by adding "x" to the "No. of days since day zero" entry, e.g. 32x, 32x; for simultaneous incubations of eggs from the <u>same</u> spawning, add a "y", e.g. 32y, 32y.
- 12. Embryo cup I.D.: identify multiple embryo cups in the same tank by Tank I.D. and number, e.g. A-1, A-2.
- Identify multiple spawnings on the same date by adding a lower case letter to the entry, "No. of days since day zero", e.g. 32a, 32b.
- 14. Initial exposure (day zero) as: embryos/fry/juveniles (circle one).
- 15. Data entries on one line are/are not from the same fish.
- B. Data Sheets for Separate Categories of Data
  - 1. Survival Data

- a. Experimental conditions: use entries 1, 2, 3, 7, 8, 9, 10, and 14 from part A to summarize the conditions of the experiment.
- b. Data Sheet

Investigator	Toxicant	% purity
Starting date of Tea	st (day zero):/	
	D M Y	

No. days since day zero	Nominal Concentration Tank I.D.	
	Initial No.	
days	Alive Alive, normal Dead Lost	
	Total for Interval	
days	Alive Alive, normal Dead Lost	
	Total for Interval	
days	Alive Alive, normal Dead Lost	
	Total for Interval	

2. Fry length or weight

- a. Experimental conditions: Use entries, 1, 2, 3, 7, 8, 9, 10 and 14 from Part A to summarize the conditions of the experiment.
- b. Data Sheet

Investigator\_\_\_\_\_

Toxicant	%	purity	

Measurements taken \_\_\_\_\_ days after first exposure of this stage.

Starting Date of Test (day zero): // D M Y

Nominal	Tank										vidual					
Conc.	I.D.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	1

Ļ,

3.	Hatchabil	ity of	embryos
----	-----------	--------	---------

a. Experimental conditions: Use entries 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12 from Part A to summarize the conditions of the experiment.

b. Data Sheet

Investigator \_\_\_\_\_

Toxicant\_\_\_\_\_% purity\_\_\_\_\_

Starting date of Test (day zero): / / D M Y

7

. . . . . . .

	Embryo Cup	Fish	No.days	No.embryos	Cum.no.	Cum. no	Cum. no.
Nominal	I.D.	I.D.	since	at start	hatched	unhatched	unaccounted
Conc.	Tank Cup		day zero		_		for

4. Spawning Data

a. Experimental conditions: use entries 1, 2, 3, 5, 6, 8, 10, and 13 from Part A to summarize the conditions of the experiment.

b. Data Sheet

Investigator \_\_\_\_\_

Toxicant \_\_\_\_\_\_% purity \_\_\_\_\_\_

Starting date of Test (day zero): / / D M Y

Nominal	Tank	Fish	No. days	No. of	No.of Estimated	Embryos used
conc.	I.D.	I.D.	since	spawnings	embryos conditions	
• <del>•••••</del> •••••			day zero		of embryos	Yes, No. Where?

5. Data on surviving adults

Second Content of the second second

- a. Experimental conditions: Use entries 1, 2, 3, 6, 8, 9, 10, 14, and 15 from Part A to summarize the conditions of the experiment.
- b. Data Sheet

Investigator

Toxicant \_\_\_\_\_\_ % purity \_\_\_\_\_

Starting date of test (day zero): / / D M Y

Nominal	Tank	Fish	No.days since	Length	Weight	Sex
conc.	I.D.	I.D.	day zero	(mm)	(mg)	(m, f, im)

- 6. A data report sheet for chemical analysis of water should be provided if detailed instructions for chemical analysis are included in a revised procedure.
- 7. A separate sheet that need not be standardized should be attached to the data records summarizing important conditions, i.e. ph, temperature, photo period, flow rates, type of food and feeding schedule, etc. and any limits of conditions that vary during the test.
- C. Transfer of Experimental Data to Standardized Data Sheets.

After considerable debugging of the data sheets and several trials with actual data, the experimental data were transferred to data sheets for storage on the computer.

## IV. ANALYSIS OF AND COMMENTS ON THE TEST PROCEDURE

The published procedure for chronic toxicity tests [2] has not been revised since 1972. Since that time considerable research on the test <u>per se</u> has been done at EPA - Duluth and elsewhere to improve reliability, reproducibility and accuracy. For example, some conditions specified in the 1972 procedure may be replaced by improved techniques, e.g. handling and selection of embryos, use of paired spawnings, etc. Those changes that could be incorporated as improvements in the test procedure should be made and a revised version published.

Following are comments about specific sections of the procedure, using the number and letter designations of the procedure [2] as a reference.

A. Physical System

4. Flow Rate

Recent USEPA regulations [7] designate certain chemicals as hazardous wastes if and when they are discarded. Guidelines and recommendations on the treatment (clean up) of experimental tank effluent should be included for a test system containing potentially toxic chemicals in ten or more tanks, changing 6 to 10 tank volumes/24 hours in each tank, all operating continuously for months.

14. Where surface water or municipal water is used, a filter system should be considered.

#### B. Chemical System

2. Measurement of toxicant concentration and

5. Methods

μ

A much more detailed procedure should be incorporated in this section in conjunction with a carefully formulated standardized reporting sheet. The essence of this suggestion resides in the absolute need to know the limits of chemical concentration changes and to have assurance that the chemical analysis data are reliable.

# V. EARLY LIFE STACE TOXICITY TESTS

A. <u>Background</u> In recent years there has been some movement in the direction of developing toxicity tests that provide much of the information relating to chronic and sublethal toxicant effects that is obtainable from full life cycle tests, yet which require far fewer resources of time, space, cost and which are simpler to carry out and analyze. To accomplish these aims, the use of early life stage toxicity tests has become more common. For fathead minnows such early life stage tests require about thirty days of effort as compared with 250 to 300 days for full life cycle tests. This permits a great many more compounds to be tested.

A number of guidelines for conducting early life stage tests in a standardized manner have been proposed [8, 9, 10]. In these tests, organisms are exposed during part of the embryonic stage, throughout the larval stage, and during part of the juvenile stage. The rationale is that this represents the period of greatest sensitivity of the fish, and so chronic and sublethal toxicant effects will be revealed.

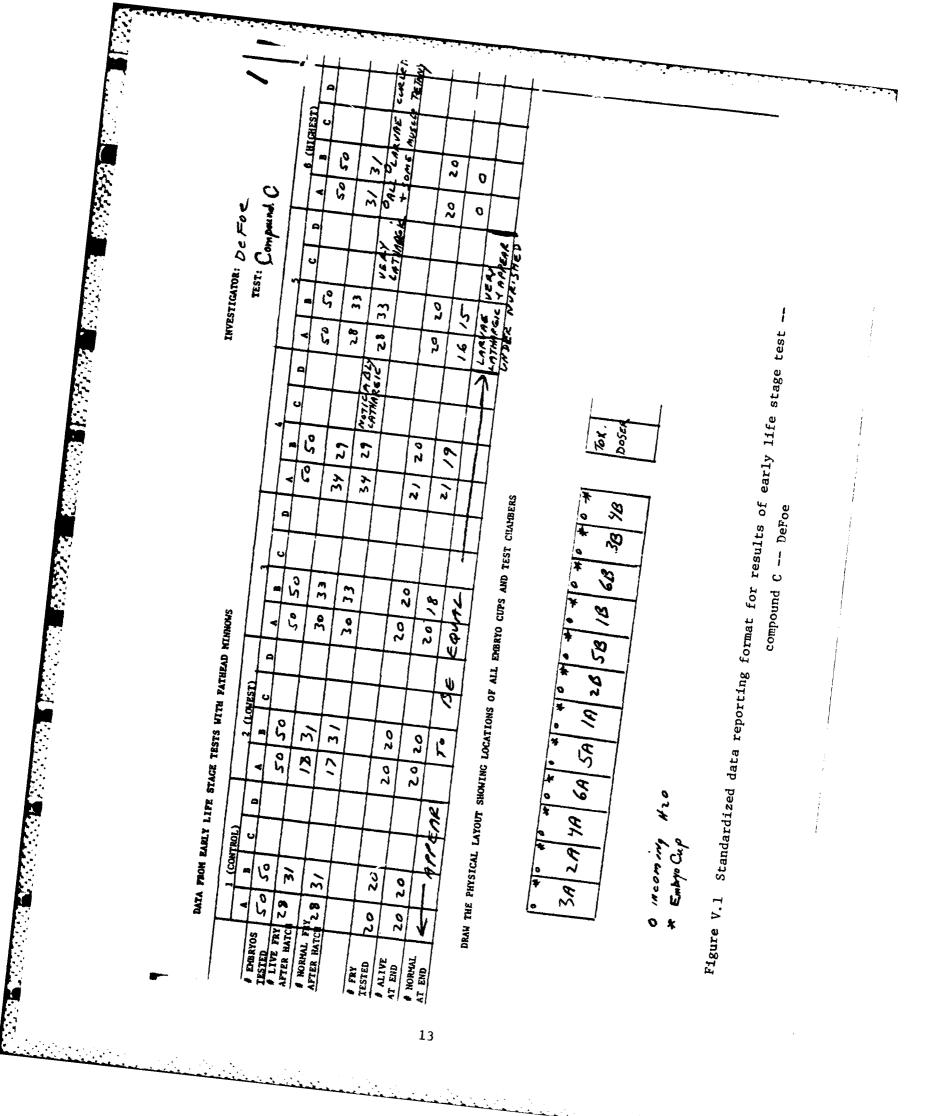
In one version of the test, groups of recently fertilized fish embryos are placed in embryo cups within test chambers. There are generally five or more toxicant groups and one or more control groups. Each (treatment or control) group consists of two or more replicate test chambers. The embryos are kept on test until they hatch (about 5 to 7 days), at which point the live, normal larvae are thinned to the desired number per tank and these are kept on test for about an additional four weeks, at which point the test is terminated. In a variant on this approach, the embryos are thinned after just two days on test. After hatch, all of the live larvae are released into the test chambers for the rest of the test. This avoids handling the newly hatched larvae at a time when they are most sensitive to the toxicant.

B. <u>Data</u> The data recorded in such early life stage tests include number of embryos per embryo cup, number of embryos hatched live and hatched normal, number of fry in each embryo cup after thinning, number of fry live at end of test and number normal, individual weights of all fry alive at end of test, and periodic toxicant concentration measurements within each tank.

Standardized data reporting sheets that facilitate the interpretation of test results and the communication of these results among investigators and laboratories have been developed by investigators at USEPA - Duluth. They have been kind enough to supply us with such sheets from about twenty early life stage tests (personal communication). Figure V. 1 illustrates such a basic data reporting sheet based on the test of compound C carried out by DeFoe. Page 1 contains embryo and fry survival and normality data and a diagram showing the test layout. We see that in this test there was a single control group (1), five treatment groups (2 to 6), two test chambers per group, and a single embryo cup per test chamber. Page 2 contains individual weight measurements on all the fry that survived the test. Page 3 contains the results of the individual toxicant concentration measurements made in each chamber periodically throughout the test. We have found these data reporting sheets to be very easy to understand and very useful.

In order to work with the data it was necessary to put them into computer readable form. The approach that we took to accomplish this is illustrated in Figure V.2 for the data from the test of compound C by DeFoe. The three types of data -- survival, weight, and toxicant concentration-- are represented in three "card types." The data for each "card type" are listed in Figure V.2. Some applications call for use of just one card type while others call for use of two or more card types. The first six entries on each card are the same across card types -- treatment group (col 2), replicate designation (col 4), card type (col 6), card member (cols 7-8), investigator code (cols 9-10), test code (cols 11-12). This provides enough information to sort the cards by investigator, experiment, type, group, and sequence should the data become disarranged. Card type 1 (survival data) contains in addition number of embryos tested (cols 16-20), number hatched live (cols 21-25), number of fry tested (cols 31-35), number live at end of test (cols 36-40), number normal at end of test (cols 41-45). Card type 2 (weight data) contains number of weights recorded from that particular chamber (cols 14-15), individual weights (5 cols per weight, up to 13 weights per card). Card type 3 (toxicant concentration) contains month (cols 16-17), day (cols 18-19), year (cols 20-21), toxicant concentration (cols 32-38) -- one determination per card. At the head of each type of information several lines of descriptive text are given. This text is informative when the data are printed out but is skipped over for purposes of analysis.

We have found this data organization to be easy to prepare, easy to maintain, and easy to use. Such data files represent the "basic data" for all subsequent analyses discussed in this report.



N -		<b>,</b> +	┿		╂──	<del> </del>	<u> </u>			<u> </u>	<b> </b>	<u>}</u>	┝╌╌┥		I	<b> </b>				<b> </b>	<u> </u>				<b> </b>
						ļ	ļ											 							
al	U	EST	V																						
DeFoe	N. N.	HIGHEST		žo																					
à	COMPOUND C		¢	10																					
		Ī	9																						
641	222	-	5																						
1 WVESTI & ATCR	Mgs. 76 ST; (	~   ~	ø	120	2	038	وجح	ہ 35	0 33	050	036	990	ہ لاح	020	068	003	وبرح	038						_	
121	۹ ۲	5	Ł	037	050 050	0 72	0770	056	053 6	0376	090	440	202	0280	047 6	049 0	037 0	1040	650						
	7655						-	~	6	0	_		-	0	0	2		•	0						
	}	ļ,	J																						
	30	- الا 2	0	112	///	020	660	138	084	32	14	64	101	010	160	131	121	063	124	060	108	6/1			
	END			077			127		162 0	089 132	135/14	165	144 102	1220	152	0 23 131	/83	144				1922	150	%	—
	81	-+-	-   	0	1	0	$\Gamma$	'	1	0	/			/	/	0	_	~			1	~	-	-	
	- 1	ł	1											_									-		
	ALIVE	nŀ	+	170	123		165	13/	7	16.	کیا	2	2	120	126	/3/	100	076	-9	12	647				_
		┝	5		6	יז א			0 13%		-	9 162	3 /37				-	<u> </u>	129			~	8		
	FISH	$\downarrow$	4-	111	14	095	179	086	160	128	51	661	ۍ اک	201	146	/37	13	158	144	152	21/2	177	198		
	ALL		)																					_	
			1	_				0	_				_											-	
	5 (mb) 01		9	160	192	56:	214	61	187	095	220	116	164	162	200	190	161	206	200	190	192	/37	181	_	
	NTS .	<u> </u>	:	121	143	2/0	108	141	142	101	121	162	129	153	187	126	181	149	126	103	103	192	080		
	WEIGHTS (MG) OF	٩	7			~																		_	
																								_	
	INDIVIDUAL		17	110	1421	172	173	174	151	002	00/	124	238	123	145	021	180	165	120	083	230	521	120	-	
	INDI			158	601	106	120	166	129	124	132	068	085	1/8	180	260	148	/63	174	138	5	172/	1381	-	

Figure V.1 (Cont)

$\sim$	C'										_							
HEST			1	_		5		-0										
6 (HIG	4 6	1357	141	7.47	2 25	18	14.25	5	31.0	×	×		*	¥	×			$\prec$
2			12.	_		55		5		7		26	<b>6</b>		24			88
		43.)	\$	2.5	2.8	5	02.2	S	12.96	20	1	34	5	8. %	5,		57.63	51.88
	2						ý				5							
*		+	-ţ			2()		e,		7		573	2		-10		/۲	
	ri ri	(%)	~	13.33	(6.8)	2	2.35	×	2.89	5/	412	Ś	ž	66%	کر اک	3.88	77	
	9									┝╼╴┧ ┝╴╺╴╽								
m	5	<u> </u>	72			20		Xo		23		50	2		64		ж	
	R R	2.47	وم.	2.5	4.81	~	1.28	60	6./6	*	بحل	6.	× 	<i>درب</i> ر	é	80.7	د.	
	9																	
KEST.	J	-	99			34		5		Y		00	2				10	
2(10	1	56.	``	28	45.	~	12	<u>к</u>	:87	$ \rightarrow $		2.	~		à	20.	シ	
	0				2		2									2		
TKUD	2		20			4		<u>م</u> لہ		8	-	2	-9		1		3	
1600	\$	250	10,	مرج	74		116	50.	189		523	40,	20.	25	1 4	033	£ 0·	
	L							5				. 0 1	27				m -	
	1((ch Thirl) 2(LOWEST) 3 4 4 5	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	1(C(1), T(1))     2(LONES7)     3     4     6     5       A     B     C     D     A     B     C     D     A       A     B     C     D     A     B     C     D     A       A     B     C     D     A     B     C     D     A       A     B     C     D     A     B     C     D       A     B     C     D     A     C     D       A     C     D     A     C     D       A     C     D     A     C     D       A     C     D     A     C     D       A     C     D     A     C     D       A     C     D     A     D     C       A     C     D     A     D     C       A     C     D     A     D     C       A     C     D     A     D     C       A     C     D     A     D     C       A     C     D     A     D     C       A     C     D     A     D     C       A     C     D <th< td=""><td>/((in Till)     2/LOWEST)     3     4     5       A     B     C     D     A     B     C     D     A       A     B     C     D     A     B     C     D     A       A     B     C     D     A     B     C     D     A       A     B     C     D     A     B     C     D       A     B     C     D     A     C     D       A     B     C     D     A     C     D       A     B     C     D     A     C     D       A     B     C     D     A     D     C     D       A     B     C     D     B     C     D     A       A     B     C     D     A     D     C     D       A     B     C     D     B     C     D     A       A     B     C     D     A     D     C     D       A     D     C     D     A     D     C     D       A     D     C     D     A     D     D       A     D     D     D</td><td>E 1((1), TWL) 2(104E57) 3 F F 5 P 6 C D P 6 C D P 6 C D P 6 C D P 6 C D -77 0(8 1/75 1/75 1/75 1/81) 11.55 1/1.5</td><td><math display="block">\begin{array}{c c c c c c c c c c c c c c c c c c c </math></td><td>E 1(((1) TNLE) 2(104 EST) 3 F F 5 P 6 C D P 6 C D P 6 C D P 6 C D P 6 C D -77 0(8 1) 55 1 1155 143.) -77 0(5 1/35 1/35 143.) 0.074 2.59 7.61 13.32 17.35 17.35 0.074 2.59 7.61 (6.8) 5.76</td><td><math display="block">\begin{array}{c c c c c c c c c c c c c c c c c c c </math></td><td><math display="block">\begin{array}{c ccccccccccccccccccccccccccccccccccc</math></td><td><math display="block">\begin{array}{c ccccccccccccccccccccccccccccccccccc</math></td><td><math display="block">\begin{array}{c ccccccccccccccccccccccccccccccccccc</math></td><td><math display="block">\begin{array}{c c c c} 7 &amp; 7 &amp; 7 &amp; 7 &amp; 7 &amp; 7 &amp; 5 \\ \hline &amp; &amp;</math></td><td><math display="block">\begin{array}{c c c} \hline &amp; 1 \\ \hline &amp; 6 \\ \hline &amp;</math></td><td><math display="block">\begin{array}{c ccccccccccccccccccccccccccccccccccc</math></td><td><math display="block">\begin{array}{c c c c c c c c c c c c c c c c c c c </math></td><td><math display="block">\begin{array}{c c c c c c c c c c c c c c c c c c c </math></td><td><math display="block">\begin{array}{c c c c c c c c c c c c c c c c c c c </math></td><td><math display="block">\begin{array}{c c c c c c c c c c c c c c c c c c c </math></td></th<>	/((in Till)     2/LOWEST)     3     4     5       A     B     C     D     A     B     C     D     A       A     B     C     D     A     B     C     D     A       A     B     C     D     A     B     C     D     A       A     B     C     D     A     B     C     D       A     B     C     D     A     C     D       A     B     C     D     A     C     D       A     B     C     D     A     C     D       A     B     C     D     A     D     C     D       A     B     C     D     B     C     D     A       A     B     C     D     A     D     C     D       A     B     C     D     B     C     D     A       A     B     C     D     A     D     C     D       A     D     C     D     A     D     C     D       A     D     C     D     A     D     D       A     D     D     D	E 1((1), TWL) 2(104E57) 3 F F 5 P 6 C D P 6 C D P 6 C D P 6 C D P 6 C D -77 0(8 1/75 1/75 1/75 1/81) 11.55 1/1.5	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	E 1(((1) TNLE) 2(104 EST) 3 F F 5 P 6 C D P 6 C D P 6 C D P 6 C D P 6 C D -77 0(8 1) 55 1 1155 143.) -77 0(5 1/35 1/35 143.) 0.074 2.59 7.61 13.32 17.35 17.35 0.074 2.59 7.61 (6.8) 5.76	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{c c c c} 7 & 7 & 7 & 7 & 7 & 7 & 5 \\ \hline & & & & & & & & & & & & & & & & & &$	$\begin{array}{c c c} \hline & 1 \\ \hline & 6 \\ \hline &$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	$\begin{array}{c c c c c c c c c c c c c c c c c c c $

M

Figure V. 1 (Cont)

-1

INVE	ST.		- - - -	05.6	:n -	105	۱.	fees		¢01	40013	N Ó A	c.								
JAIA	<u> </u>	<u>លីគ</u>	EAN	LT-	ĬĔ	<u>`šī</u>	<u> </u>	Ξ	TS	11	Th r	AIT	EAD	MINN	045-						
IC' BN:	LËV VIC ER	VELS NUAL OF	WEI WEI	1101 IGH1 GHT:	SNTA ISI 3 8	MG) LIS	•2( 0≓ T Ω	LÚW ALI F W	ST	is in	ALI	(HI VE	GHE AT	ST);2 Eņu C	RÉP F TE	S E	4 1/	CARUS	CRL	.ESS/0	:ELL)
14	<u> </u>	0F 102	<u> </u>	<u>z</u> ų:	158	Ţ	<u>59</u> _	1.6		ŢĔ	i	<u>6</u>	179	124	13	2	Goñ	<u> </u>	_118.	18.0	<u>9</u> 2
		2-2		1.36020	149 011 139		63 47 65	172		138	15:	3. 4. C	179 172 151	138 200 271	10	0	124	235	173	145	17:
<u> </u>			ίĪ.	Ç.	1 4	- 1	43 <u>·</u>	<u>- čí ;</u>		. ė 9 1 <u>6 9</u>	230 14	ĭ	142	<u>ĭċi</u>	12	4	1.52	129	153	197	12-
Z A 22.5	Ž.	<b>4</b> 84	ät :	29	151 161		47 92 66	012 126 199		103	10 151 19	3	135 192 192 187 137	080	22	3	116	164	152	20 G	190
<u>ه د</u>	<u>-</u> Ž`	102	ăŢ.	žă	<u>, Ži</u>	Ī	48 -	095		190 190 192	<u>. 08</u>	6	160	131 128		7	1.79	153	102	146	137
14 15 15 15	222		ŭt (	1921	$\frac{171}{131}$ $\frac{171}{170}$		58 53 96	144 151 125			11 13 04	1	177	198 149	15	2	1 5 2	137	120	129	,131
<u>4</u>	4	1.2	ίI.	1	100		96 3 J	088		156 22 45	130	<u> </u>	152		13	5	5ء 1	14+	122	152	:73
4 8		202	01	19	193		44	145 070	1 2 1	145 199	130	9 R	1 <u>52</u> 226 034	150	) 11			1:02	290	001	
	2	202 102 202		ļć lo	177	C S	11 63 56	1 2 7 4		San.		a –	310				064 ]+4	=	328	091 	131 
	222	2.2		Lo 15 15	037 037 029 045	000	.4 50 36	038 038	e Birt	555	035	5	ü <b>3</b> 3	\$57	C3	6	666	- 845	150	06 E	213
<u>, A</u>	- Ž	1.2	<u> </u>	ίς.								_									
0 5	<u>ک</u>	142	91 I	، <b>لان</b> غ ر		5	و میشود. و در برد ساله م														•1
INVE	ST	GAL	dr ÷	DEF	05	102	<b>N</b> 26	IEST		COM	POUN	D C					•				
- Jin IA SIX	нен Lev	ELS.	CARL	iícc	LIFE MTR	្រុន្តរុ	AGE	TES	ŢŞ	"I I	TH F		EAD	MINN	UWS						
# 3.4 + FK	BR Y	ŌŜ	[ÊS	Ê. A	LĨy	<u>Ē</u> Īý	<u> </u>	FTER NL	1			NŌ	RMAI	[' # ¥ ¥	ÀFT	ER I	i a t c	н			
18	- <b>L</b> -	1020		÷.,	- 50 56		28	Z	). 	žä	.20		: <u>2</u> C							•	
2 4	Ĩ	1020			50			- <u>-</u>		20	20	5	20	-					• •		
38	ł	lug	뷰		50			30			20 15 21		23								
- 4 Ă	ī	1020	1		50		34	34	•	Źĭ	21	Ē	21			<b>.</b> .		-	•		
+ 3 5 A	÷	エレイリ	11		- 54		29		)	31			. 7	•							
5 3	i		1		50		28	28		20020	1		10		•			· ·	•		
A 0 6 8	1	192	bi	•	90 50					2¢	Ģ	1	00	<sup>1</sup> .	•	· ·					
		TAK!	· · · ·		- 20		<u>.</u>			2,11			_96								****
·	(T)		به در.		ne	100		T_CT				-	<u> </u>								
LA TA	°F4		ARL	YE.	IFE.	ST	AGE	'ŦÈ'	ts'		TH FA	ITH	EA O	MINN	ows		•			• •	
SIX	<u>L = V</u>	2 L ` _			IN I R			LOWE	\$1)	9		HI	GHES	ST):2	KEP	S E	1 (A	19).			
A L	≞ ۲∪ دُ	1.02;	21N %2 _ 1	NTR 29	A FI 1157	9 UM 2	UF	;ιυχ	ICA C	Q • 1	TNUI 268	<b>C</b>		10112	GIV	EN)				•••••	
L A	Ę	1,2 2,2 3,2 3,2	ЯĮ.	- Çş	Č77	ý i			- C C		.65										
	5	いいいい	- L' - L	- 69 - 29	127	å.				i9. (	574 116			•						<b>F</b>	
1 A	3	+ 12: 5 12:	51	Ę	147	ģ.	'		្រភ្ញ	č.(	.68									-	
	-	0			157	<del>7</del> —			- 50	فعانيه	23		-					<u></u>			
L A	5		.i	11	5ž7	ģ			čč	C • •	33									•	
13	<u>ج</u> :	1-2	1	22		2			÷9	ا م بر ا د م مک	78					·					
13	5	302	<del>1.</del>	- 39	137	<del>6</del>			- 20	5.7	45		<del></del>	•	·	~					
13	3.	362 462 502		ý	177	2		:	ğğ	QQ	45					• •	• •	•			
		0020	1	_ŭ9	211	<u>í</u>			_ <u>6</u> 6	<u>.</u>	20			··		•		••••		•	•
1 5	-	7.2.				<b>u</b>			ιü	.U	24	_								· · · ·	
2 A	שוטיט	9520		1	57 27	9			 - ມີມີ		25										
-2 Â	3	ZúZ.	ī	ý	<u> 11</u>	<u>9</u>			-	1.	2						-				

Figure V.2 Computerized data format for results of early life stage test -- compound C -- DeFoe

2 A 3 30201	391079	2.54
A = +u2u1	191279	2 • 54 2 • 24 1 • 84 1 • 53
2 A 3 5 2 1	691479	1.84
OVENI		1.53
A 3 (32-1	272919	1.76
	1.0279	2.02
4 2 3 4 4 4	14,679	1.78
<u></u>		2.34
		2.15
2 2 2 4 4 4 4 4	071(15	1 + 94
	222017	1.93
2 3 3 1 2 1		<u>1.024</u>
2 8 7 82281	1 272	2.001
3 A 3 1.2.1	2 470	· · · · · · ·
3 A 3 2 201		1.41
3 A 3 3.201	51075	6 41
3 A 3 4 2.1	391274	
3 A 5 5.231	61479	4.16
3 A 3 60201	ČyľýŻý	- U • 10 6 - 65
JAJILLI	692819	4 . 1 3
3 A 3 30201	1.5279	69
2 2 3 1 2 1 Î	-53679	5.22
3 8 3 2.2.1	141179	7.16
2       2	. 41 379	$ \begin{array}{c} 2.54\\ 2.24\\ 1.84\\ 1.55\\ 1.56\\ 2.02\\ 1.76\\ 2.34\\ 2.15\\ 1.94\\ 1.99\\ 1.54\\ 2.01\\ 7.47\\ 5.66\\ 4.25\\ 4.25\\ 4.25\\ 4.25\\ 4.25\\ 4.25\\ 4.25\\ 4.55\\ 1.55\\ 1.55\\ 1.7.15\\ 1.2.6\\ 8.46\\ 4.73\\ 0.25\\ 4.56\\ 4.73\\ 0.25\\ 4.56\\ 1.7.35\\ 1.2.59\\ 1.4.17\\ 1.5.55\\ 1.7.12\\ 1.6.57\\ 1.7.5\\ 1.7.12\\ 1.6.57\\ 1.7.5\\ 1.7.5\\ 1.7.5\\ 1.2.5\\ 1.7.5\\ 1.2$
3 3 3 46201	191779	4.73
2 5 2 20201	692279	<b>0.</b> 26
	092119	4.54
	1.0112	6.59
	1.44217	6.74
	676374	$14 \cdot 17$
$-\frac{7}{4}$ $\frac{3}{4}$ $\frac{2}{3}$ $\frac{2}{2}$ $\frac{2}{1}$		13.32
	671019	10.8/
4 A 3 56201	- 41470	16.40
+ A 3 6		
+ A 3 /		14.19
+ A 3 5.2.1	1.1.279	12 20
6 3 1.2.1	191679	13.55
A B B 2.201	191176	17.15
4 8 3 34241	091379	16.57
4 8 3 40201	091779	14.76
+ \$ \$ 54211	092479	13.73
4 8 3 60201	<u></u>	14.37
- + 5 3 (vic)T	1.0179	15.1.
4 0 3 3 4 4	1.0379	14.14
	~ <u>~</u> ~ <u>~</u> ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	43.71
	V91119	43.15
	$\begin{array}{c} 91279\\ 91279\\ 91279\\ 91279\\ 91279\\ 91279\\ 91279\\ 91279\\ 91279\\ 91279\\ 91279\\ 91279\\ 91279\\ 91279\\ 91279\\ 91279\\ 912779\\ 912779\\ 99201279\\ 99201279\\ 9912799\\ 9912799\\ 99927799\\ 99927799\\ 10992$	$ \begin{array}{c} 6 \cdot 59 \\ 6 \cdot 59 \\ 6 \cdot 74 \\ 14 \cdot 17 \\ 13 \cdot 52 \\ 16 \cdot 87 \\ 17 \cdot 35 \\ 17 \cdot 35 \\ 12 \cdot 99 \\ 14 \cdot 79 \\ 13 \cdot 58 \\ 13 \cdot 56 \\ 13 \cdot 57 \\ 14 \cdot 76 \\ 13 \cdot 73 \\ 14 \cdot 76 \\ 13 \cdot 73 \\ 14 \cdot 37 \\ 15 \cdot 11 \\ 45 \cdot 71 \\ 45 \cdot 15 \\ 55 \cdot 5 \\ 45 \cdot 25 \\ 45$
5 A 3 5 221 5 A 3 2 221 5 A 3 2 221	041219	22.020
	571479 561676	44040
	シフトフィブ	46.11
	572317 1.270	- <del>4</del> 2.025
5 8 5 1.7.1	していして、	71.000
5 5 3 2 2 1	4174	49.26 51.23 49.21 49.50 51.50 49.24
5 5 5 3.2.1		
5 3 3 40201	241774	→ 4 ● 21 •• 9 → 2 •
5 5 5 5 5 2 . 1	42.74	44, 74
<u>58302.1</u>		43 5
5 3 3 7.201	1.0179	51.99
ာမ်း ဦးစိုပ်ဦးပါ		51.99 51.63 135.9
	-90579 	51.50 43.24 44.27 43.4 51.99 51.59 135.9 117.47
$ \begin{array}{c} 5 \\ 6 \\ 7 \\ 7 \\ 7 \\ 7 \\ 7 \\ 7 \\ 7 \\ 7 \\ 7 \\ 7$		
し み ひ やくどいよう ふ ふ ち ん た ち	5 M L 2 [ M	154.15 150.75 131.
	· · · · · · · · · · · · · · · · · · ·	1010 <u>-</u>
$   \begin{array}{c}                                     $		
0 0 0 0 0000	291179 191379	122.75
the state of the second	671217	上行は、、

#### VI. PRELIMINARY GRAPHICAL DATA DISPLAYS

Graphing the data is generally considered to be a good first step in analyzing data. Graphs provide insights into the structure of the data and reveal the presence of possibly unanticipated relations or anomolies in the data.

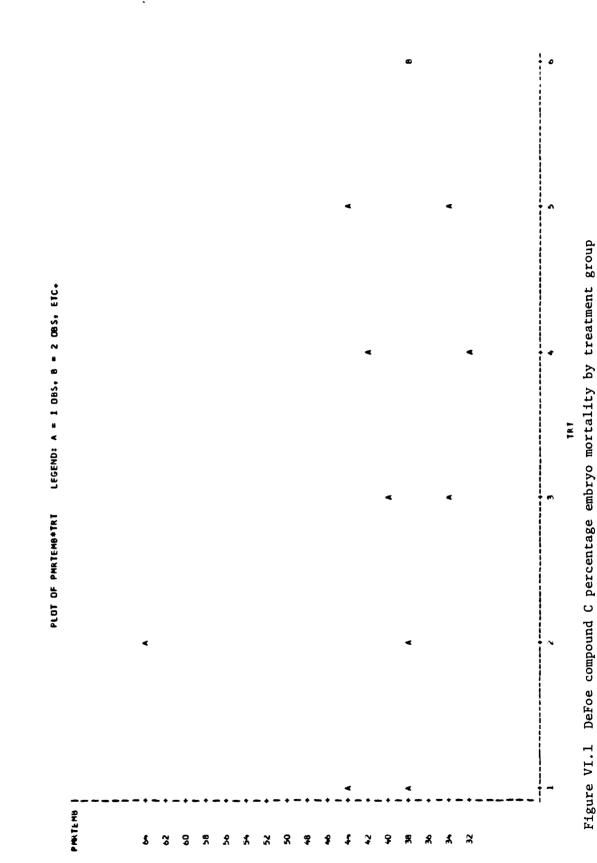
Figures VI.1 and VI.8 illustrate the kinds of information that can be obtained from preliminary plots. They illustrate percentage embryo and fry mortality and abnormality observed in early life stage tests on fathead minnows conducted by DeFoe with compound C and by Holcombe and Phipps with compound D. The tests each consist of a control group (1) and five treatment groups (2 - 6) with toxicant levels in roughly geometric progression. The DeFoe test was run with four chambers per group. The plotting symbol "A" represents a single response, "B" represents two coincident responses, etc.

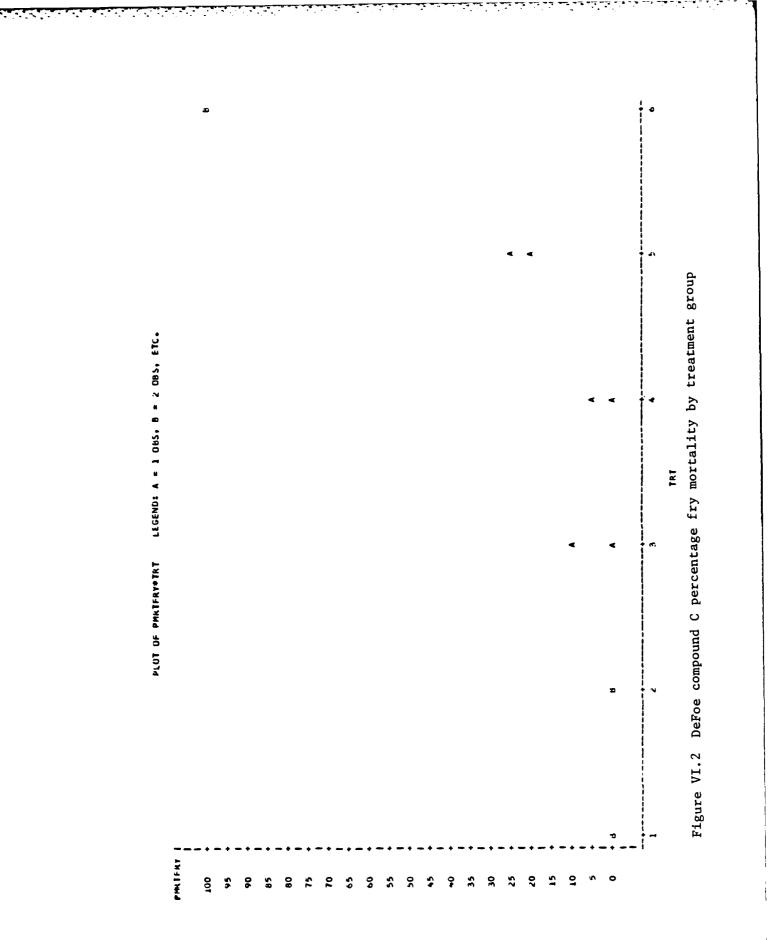
Figures VI.1 and VI.5 reveal no trends in embryo mortality with increasing toxicant level in either test. Tank to tank variation within treatment groups appears to be approximately constant across groups except for a single tank in Group 2 (Figure VI.1) which has about 50 percent greater embryo mortality than all the other tanks in the test. It appears to be an <u>outlier</u>, i.e. its response does not seem to conform to the pattern of the bulk of the data.

Figures VI.2 and VI.6 show increasing trends in fry mortality with toxicant concentration in each test. This pattern is to be expected since the larvae are most toxicant sensitive shortly after hatching. In each test tank to tank variation within groups is greatest in the middle and least at the ends, in conformance with binomial theory. No outlying tanks are evident with respect to fry mortality. Note that in both tests the highest treatment groups experience 100 percent fry mortality.

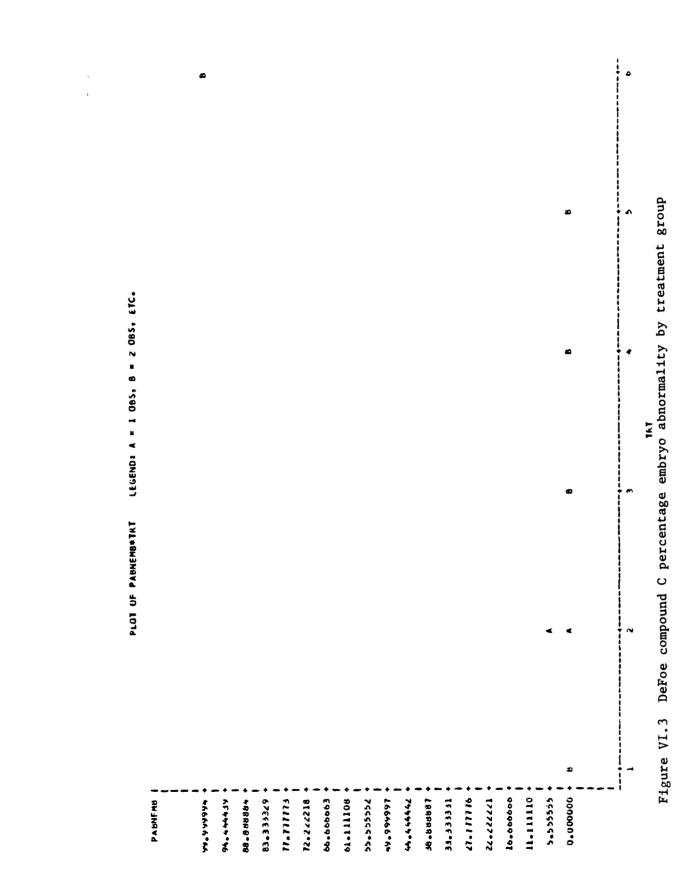
Figures VI.3 and VI.7 exhibit embryo abnormality in the two tests. They are strikingly similar. In the control groups and the four lowest concentration groups there is little or no abnormality among newly hatched live larvae. However in the highest concentration groups there is 100 percent abnormality among newly hatched live larvae. It thus appears that very high concentrations of each of these toxicants will penetrate the embryo.

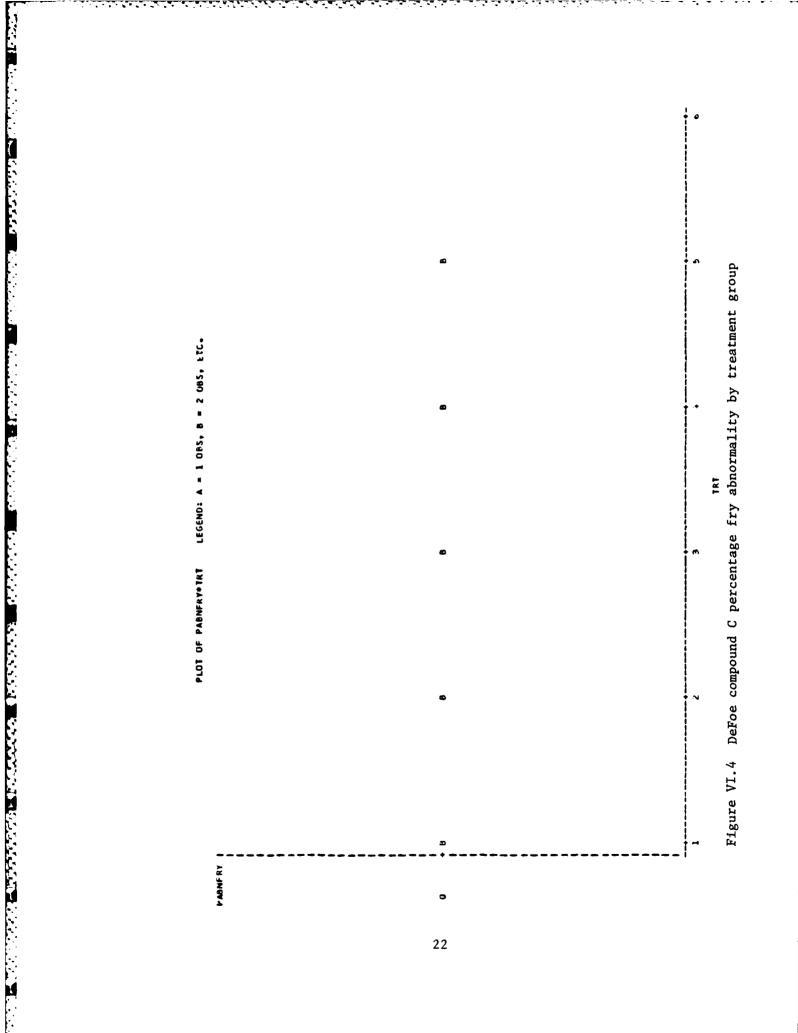
Figures VI.4 and VI.8 exhibit fry abnormality in the two tests. In brief, there is none. After 32 days the fry have either died or are normal. Recall that the highest toxicant groups experience 100 percent mortality and so there is no abnormality data to plot.

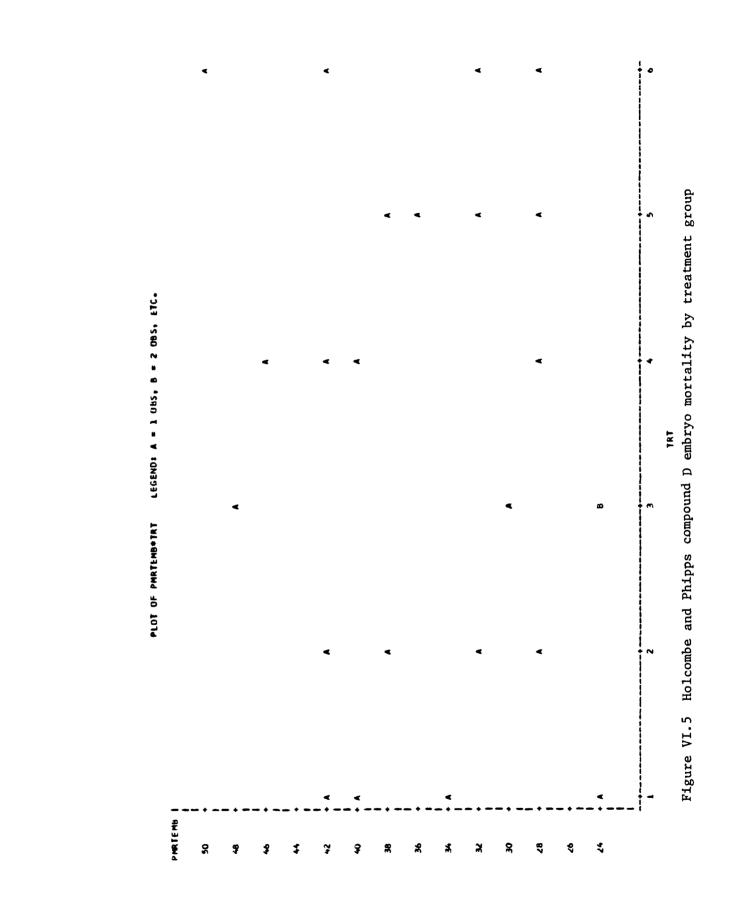


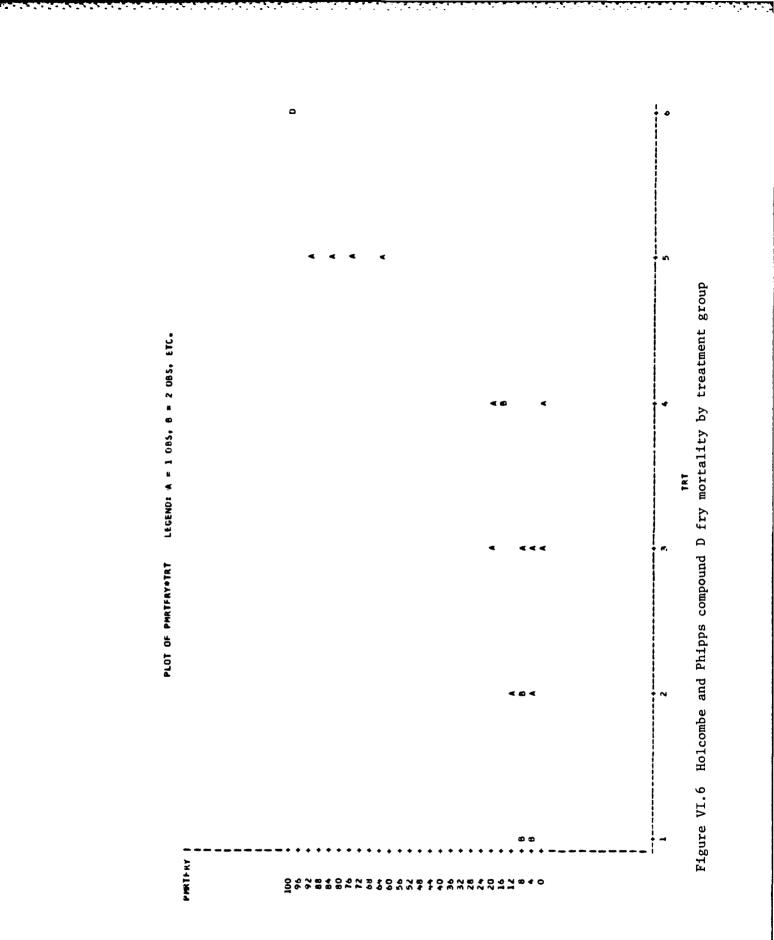


M

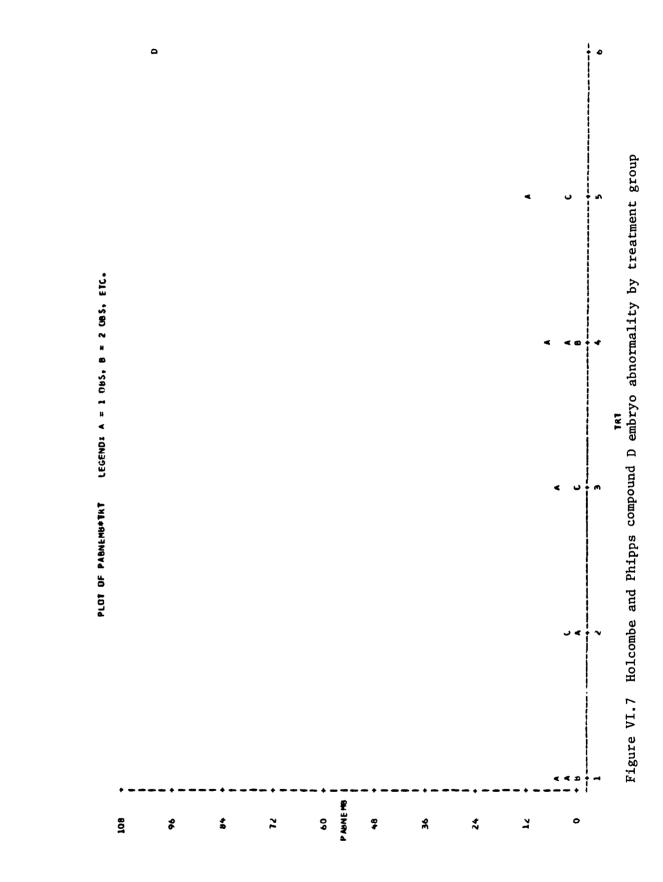




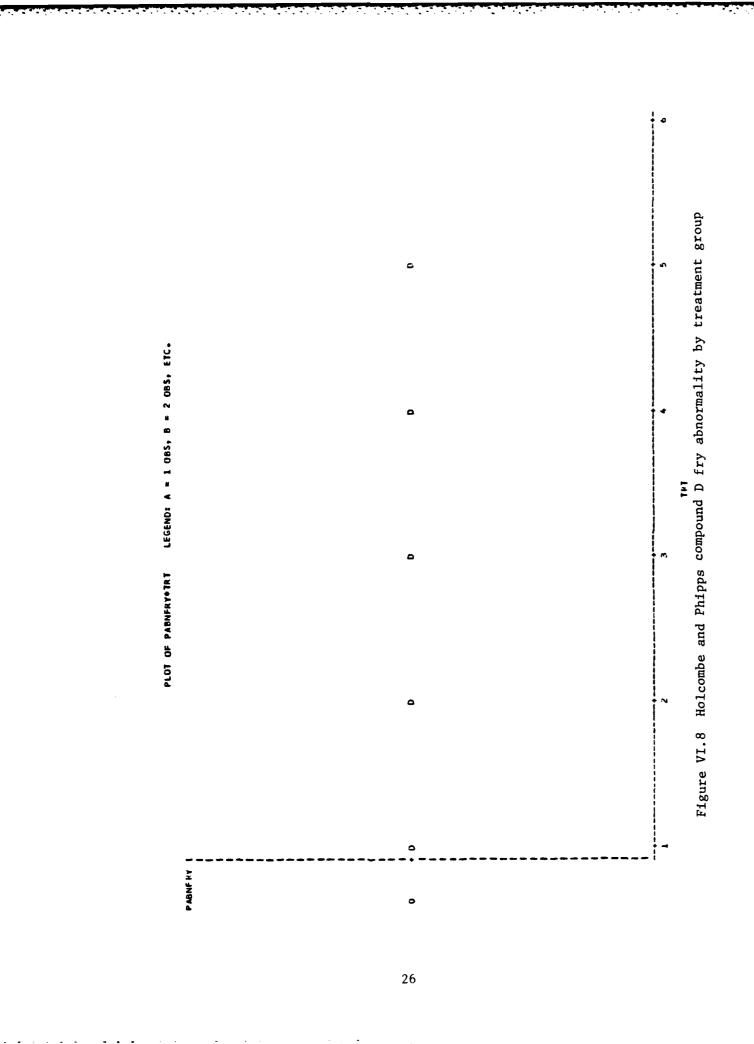








ŗ



#### VII. TESTING FOR TANK TO TANK HETEROGENEITY WITHIN TREATMENT GROUPS

A. <u>Background</u>. In order to assess variability of response, toxicity tests generally include several fish tanks, usually two to four, within each treatment or control group. EPA guidelines for early life stage and full life cycle toxicity tests with fathead minnow [2, 8] call for at least two replicate test chambers for each treatment group. Some tests use more.

An important preliminary inference of interest in toxicology data analyses is to determine if there is any statistical evidence of variation in response among tanks within treatment groups. Such variation might be due to differences in location or handling of individual tanks, to fungus or illnesses that might invade a tank, to unforseen accidents during the test, etc.

If evidence of tank to tank heterogeneity exists then analyses should be carried out on a per tank basis. If no evidence of tank heterogeneity exists then data might be pooled across tanks within groups and analyses carried out on a per fish basis, ignoring the replicate tanks. For example, mortality rates could be compared based on binomial theory. Such per fish analyses would provide many more degrees of freedom to estimate random error than would per tank analyses and so are more sensitive. For example if there are four tanks per group and 25 fish per tank then a per fish analysis might be based on 99 degrees of freedom per group whereas a per tank analysis would be based on just three degrees of freedom per group.

However the validity of per fish analyses rests on the absence of tank to tank heterogeneity. If there is in fact variation in response rate across tanks within treatment groups then variability estimates based on per fish analyses will underestimate the true variability of the estimates and test statistics. This will result in standard error estimates that are too small, confidence intervals that are too short, and hypothesis testing procedures that falsely reject the null hypothesis more often than their nominal rates (i.e. inflated alpha levels). It is thus important to test for the presence of tank to tank heterogeneity within treatment groups before proceeding on to the analysis of primary interest.

## B. Remarks on Some "Standard" Procedures.

Finney [11], section 9.1, pp 175 ff. suggests the following procedure for testing tank to tank heterogeneity. Fit a probit curve to the data based on pooled data across tanks within groups. Fit a probit curve to the data using the individual tanks within groups. The point estimates of the two probit fits will be exactly the same. However the residual chi squares and their respective degrees of freedom will differ. The differences between these two residual chi squares can be interpreted as the chi square for heterogeneity among tanks within treatment groups. Similar considerations hold for the usual chi square test for homogeneity.

We have carried out this procedure using the probit fit and using the usual chi square test for homogeneity. We compare the results of these two tests. The theoretical bases of these heterogeneity tests are discussed in Appendix AVII.

We illustrate these two heterogeneity test procedures on the fry mortality data. First consider the test of compound C by DeFoe.

# Probit Fit

· ·

We fitted the probit model to the treatment groups using the (natural) logarithm of concentration and excluding the control group. (Note that the same fit was obtained when the control group was included). The probit fits were carried out using the PROC PROBIT procedure in the SAS statistical computing system [ 12 ]. The data consist of I=5 treatment groups, J=2 tanks per group. Thus there are 10 responses to which we fit the two parameter probit model.

 $p_i = \Phi(\alpha-5 + \beta \log C_i)$  i = 2,...,6

Figures VII.1, VII.2 contain the results of the probit fits to the individual responses and to the responses pooled across tanks within groups respectively. The analysis of variance table, as suggested by Finney, appears in the bottom portion of Figure VII.1. The upper 0.005 point of the chi square distribution with 5d.f. is 16.75. Thus this test for tank to tank heterogeneity within groups is "highly statistically significant". At face value this suggests strong statistical evidence of variation in response rate across tanks within treatment groups.

## Chi Square Fit

We now carry out a chi square test of heterogeneity in response rates across tanks within groups based on the usual chi square test of homogeneity across groups. Figures VII.3, VII.4 contain the results of the chi square tests based on the individual responses and on responses pooled across tanks within groups respectively. Control group responses are included in these tests. The tests were carried out using the PROC FREQ procedure in the SAS statistical computing system [12]. The analysis of variance table suggested by Finney appears below.

Source	d.f.	<u>SS</u>
Lack of fit of pooled tanks about model (homogeneity)	5	182.785
Variation of individual tanks w/i tmnt groups (by subtraction)	6	0.89
Lack of fit of individual tanks about model (homogeneity)	11	183.675

Thus the heterogeneity chi square is very small. Thus there is <u>no</u> statistical evidence of variation among tanks within treatment groups.

Note that the conclusions arrived at from this heterogeneity chi square test are in direct contradiction to those arrived at from the heterogeneity chi square test based on the probit fit. What is the cause of the discrepancy?

There are two possible sources of difficulties. The first concerns the probability estimates in the denominators of the test statistics. In the probit based statistic the i-th group response rate in the denominator is estimated as  $P_i = \hat{\Phi}_i \equiv \Phi(\hat{\alpha} - 5 + \hat{\beta}\log C_i)$  whereas in the homogeneity chi square based statistic the i-th group response rate in the denominator is estimated as  $\hat{P}_i = \hat{P} \equiv X_{i+}/N_{i+}$  for all i. The assumption of constant P values in the denominator is clearly not justified. The assumption of  $P_i$  values based on the probit model is also not good, as can be seen from the very large residual "chi square" value in Table VII.2. Thus the substantial differences in the response rate estimates that appear in the denominators of the two statistics, along with the probable inadequacies of both sets of estimates, may account for at least a portion of the discrepancy in chi square values.

The second possible source of discrepancy is based on the validity of the chi square assumption itself. The validity of the asymptotic chi square theory is dependent on the cell expected frequencies being large enough. In particular if any responses are observed in cells with very small expected frequencies then very large cell chi squares can result which can greatly inflate the statistic.

Consider the two chi square statistics for lack of fit from the probit model -- one for individual tanks and one for pooled tanks. We break out the individual components of these statistics.

	-						individu Lt to probi		
		Mean	(X <sub>ij</sub> )		(N <sub>ij</sub> )	^	^	~ ^	2
Grp	Tan	k Conc	Dead	Live	Total	Pi	Nj <sup>P</sup> i	N <sub>i</sub> P <sub>Q</sub> i	X <sup>2</sup>
2	Α	1.991	0	20	20	0.000015	0.000301	0.000301	0.000301
2	В	1.991	0	20	20				0.000301
3 3	Α	5.976	0	20	20	0.002542	0.050830	0.050701	0.05096
3	В	5.976	2	18	20			<	74.9340
4	Α	14.812	0	21	21	0.0475	0.9975	0.9501	1.0473
4	В	14.812	1	19	20		0.95	0.9409	0.00276
5	Α	48.307	4	16	20	0.4227	8.454	4.8805	4.0648
5	В	48.307	5	15	20				2.4444
6	Α	146.984	20	0	20	0.88356	17.6713	2.05764	2.6355
6	В	146.984	20	0	20				2.6355
									87.8165

Table VII.1 DeFoe Compound C Fry Mortality Data

(,

i

This compares with  $\chi^2$  = 87.7666 calculated from SAS PROC PROBIT

Table VII.2 Probit fit to groups 2, 3. 4, 5, 6 -- tanks pooled within groups Chi square statistic for lack of fit to the probit model

	Mean	(X <sub>1+</sub> )		(N <sub>i+</sub> )	^	^	^ ^	2
Grp	Conc	Dead	Live	Total	Pi	Ni+ <sup>P</sup> i	<sup>N</sup> i+ <sup>P</sup> i <sup>Q</sup> i	х <sup>2</sup>
2	1.991	0	40	40	0.000015	0.000602	0.000602	0.000602
3	5.976	2	38	40	0.002542	0.10166	0.10142 🤇	35.532
4	14.812	1	40	41	0.0475	1.9475	1.855	0.484
5	48.307	9	31	40	0.4227	16.908	9.761	6.407
6	146.984	40	0	40	0.88356	35.3426	4.11528	5.271
								47.695

This compares with  $\chi^2$  = 47.6775 calculated by SAS PROC PROBIT

Comparison of these two chi square values clearly shows the source of the "significant" chi square for heterogeneity. Namely the tanks from group 3 have very small expected frequencies (NP) yet have observed responses. Thus these component chi square values are large and dominate the overall chi square values.

If we remove the group 3 values from the chi square statistics we have:

Separate tanks:	$\chi^2 = 87.8165 - 0.05096 - 74.9347$	= 12.831
Pooled tanks:	$\chi^2 = 47.686 - 35.532$	= 12.154

The relation between these two chi square statistics is then just like that of the chi square tests resulting from the contingency table tests.

- Moral: Uncritical use of the chi square test for homogeneity of tanks within concentration groups recommended by Finney can lead to completely incorrect results and results contradictory to those of other homogeneity tests because of:
  - small expected frequencies within cells
  - response rate estimates based on particular (possibly inappropriate) model fitted

We repeated the same calculations on the fry mortality data in the test of compound D by Holcombe and Phipps.

# Probit Fit

Using all six groups, a logarithmic transformation of concentration, and Abbott's correction for background response we obtain:

Source	d.f.	S.S.
Lack of fit of pooled tanks about probit model	3	0.5064
Variation of individual tanks w/i tmnt groups (by subtr)	18	21.7906
Lack of fit of individual tanks about probit model	21	22.2952

The value 21.7906 is at the upper 24 percent point of a chi square distribution with 18 d.f. and so is nonsignificant.

# Chi Square Fit

There are I = 6 groups, J = 4 tanks per group. We carry out chi square tests of heterogeneity in response rates across tanks within groups based on the usual chi square test of homogeneity across groups. We obtain:

Source	d.f.	S.S.
Lack of fit of pooled tanks about model (homogeneity)	5	389.676
Variation of individual tanks /i tmnt groups (by subtr)	18	10.198
Lack of fit of individual tanks about model (homogeneity)	23	399.874

The heterogeneity chi square is again small. There is no statistical evidence of variation among tanks within treatment groups. Note however that there is strong statistical evidence of variation in response rate from group to group, as would be expected. Thus the weights in the denominator of the heterogeneity chi square statistic are suspect.

## C. Separate Heterogeneity Tests Within Treatment Groups

We have seen for DeFoe's fry mortality data that we can obtain diametrically opposite conclusions about heterogeneity of responses within treatment groups depending on whether the test for homogeneity was based on a probit model fit or on a contingency table fit. This was attributed to

- 1. differences in the weights used in the denominators of the chi square statistics (based on the assumed model)
- 2. small expected frequencies within cells that invalidate asymptotic distribution theory.

To account for problem (1), we carry out separate <u>chi square</u> <u>heterogeneity tests</u> within each concentration group, without imposing any structure on the form of the concentration-response relation. We do this by carrying out separate chi square tests within each group and then pooling the results across groups.

There is however, a technical problem associated with this approach. For many (if not most) of the responses of interest the probabilities of occurrence are fairly close to 0 or 1. Therefore the expected frequencies of occurrence can be rather small, thus invalidating the use of asymptotic chi square theory. We illustrate this phenomenon with the Holcombe and Phipps fry mortality data, broken down by group. The output (from SAS PROC FREQ) is shown in Figures VII.5 to VII.10

We see that groups 1, 2, 3, have small expected numbers of dead fry (less equal to 2.0). Group 4 has expected dead = 3.3 and group 6 has expected live per tank = 0. Thus groups 1, 2, 3, 4, 6 have small expected frequencies in at least some of the cells of the table.

We use a (relatively stringent) criterion of applicability of asymptotic chi square theory that requires that there be an expected frequency of at least 5 within each cell of the table. Only group 5 satisfies this criterion within the Holcombe and Phipps data. We must thus base some of the within groups heterogeneity tests on exact, small sample theory.

Thus we wish to pool <u>across</u> groups the results of tests of homogeneity of responses among tanks <u>within</u> groups. Some of these tests are based on asymptotic theory while others are based on exact, small sample theory.

We have developed a computer program, EXAX2, to carry out such a procedure. We discuss this program in detail and illustrate its application in the following section.

<sup>&</sup>lt;sup>1</sup>Dixon and Massey [13] page 233, has a slightly more liberal criterion, namely "...none of the  $F_i$ 's (i.e. expected frequencies) is less than 1 and not more than 20% of the  $F_i$ 's are less than 5..." Again, only group 5 would satisfy this.

0.0111167 (1/mx 2) 0-80174020 SIGM 3 . . ( 1 1 £ 0.02711267 1 211125-PROB > CH1-SQ = 0.0001 11-22-11 COVARIANCE MATRIX 5 1 1 Ì. \*UT22 -23\*4 - 4-02387764 1012775 32 ~ A S 2 NU SIGM i NOT PUBLING TANKS WITHIN RONC 18 2 8 2 8 PROBET ANALYSIS ON LOG (CHEAN) \$7.7666 476775 40.084 5 (ATSC) STOPE(S) 177-1-26.0 1.24728684 į 905 R. LARGE (RI-NU-ARC EV C. 0.10) IN 601 00 UNED IN COMPUTING FLUDUIAL LIMITS. З ار زنور ---m 4 EN MALEIT INTERCEPT (K) 17.7666 WITH SLOP B -1.08 639038 ANALYSIS OF VARIANCE VARIATION OF INDIVIDUAL TANKS W/I TMN'T GROUPS LACK OF FIT OF POLLED TANKS ARUT MOLT MOLE LACK OF EIT OF (BY SURTRACTION) istule 25 1. 130 1 4.1 3 5 77 INDIA TAUKS ABANT - PRUBIT MUDIL Seriece و معافر ال IJ **Zalva i ja j**i CHI-SA 4.40887031 -DUTERCEPT **JTERATION** V. I. HALL ..... COVARIANCE MATAIX ž not pooling teaks which ч. andrel group deleted INTERCEPT SLOPE leg concentration , : l. U. Fry muchility ¥. • NUT: 1.

SIGM

ļ

Figure VII.1 Probit fit to individual responses DeFoe compound C (excluding control group)

l netrone L'avitteix concordiorhène	1200H	WORLT ANALYSES ON LOGICNEAN)	(CHEVN)		
control gump Letter 31 EK AT 10N	INTERCEPT (2)	C. 47 C. 164		SIGM SIGM SIGM SIGM SIGM SIGM SIGM SIGM	
COVARIANCE MATRIX INTERCEPT	91.0PE		COVARIANCE MATAIX		
INT FRCE PT 6.38675097 SLOPE -1.57378025	-1.57376025		MU Signa	0.22229209 0.03927580	08512620-0
CH 2-50	SQ = 47.6775 NTTH	3 DF	PROG > CH1-50 = 0.0001	100	

. • • •

. . .

1

Figure VII. 2 Probit fit to responses pooled across tanks within groups DeFoe compound C (excluding control group)

and the second se

1

( <b>1</b> - 1		,				וער	ועור: חר רואר	а. 1 -		۰.			
EREGUENCY EXPECTED CELL CHIS		4	- - 	- 4	-	- 	L				ŕ	-	
<b>DEA</b> D	0000	440 00,000	34 75 0	14.0 0000	07 70	10.00 10.00 10.00	0.44	44.4	4000	404	5,03	45 54.30	T07AL 52
LIVE	15 20	15.7 15.7 100.001	15.70	15.7	100.00	15.7	16.5 16.5 100.00	15.13	15.15	15.15	00.10	01.31	163
TOTAL	8	20	ຊ	8	8	20	17	20	22	ล	28	20	241
					nesing pyten Netrygann								
ry mortal acts of gran acts acts	Fry murtality data Central group included tanks and pooled within Concentration group	A A A A A A A A A A A A A A A A A A A	1	1 5 1					; • • •	: : . :	÷. :	1	
			Figure VII	÷ ،	Chi square DeF		of homog mpound C	test of homogeneity based o oe compound C fry mortality	oased on ctality.	individual		responses	

•

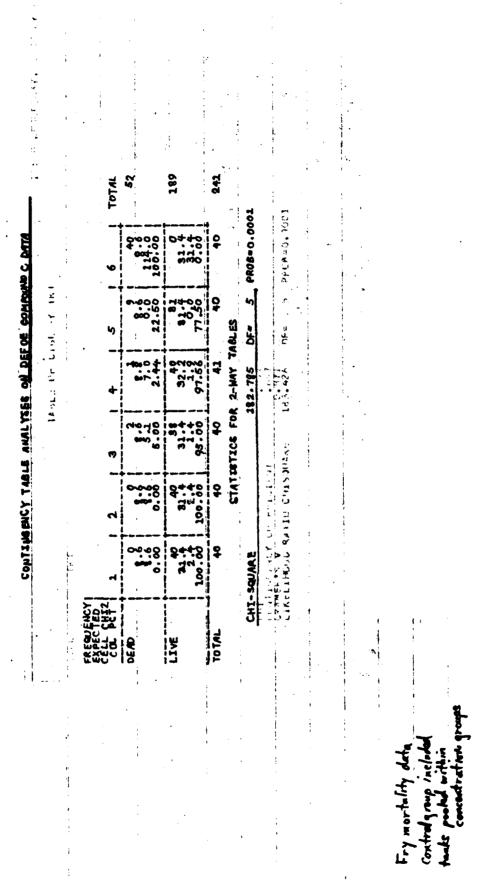
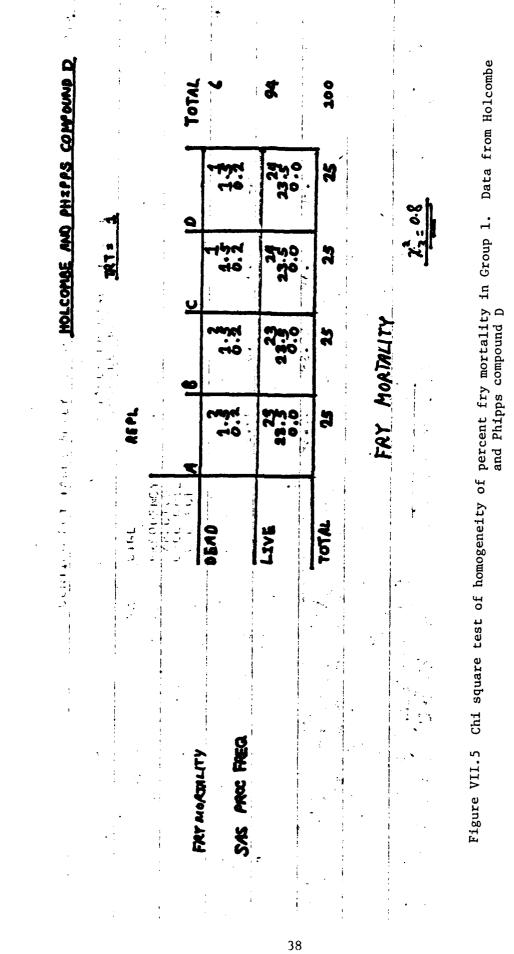
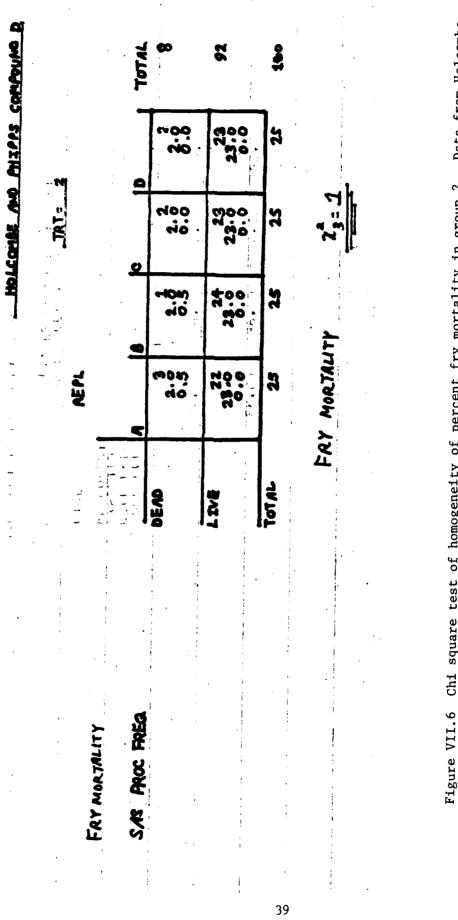


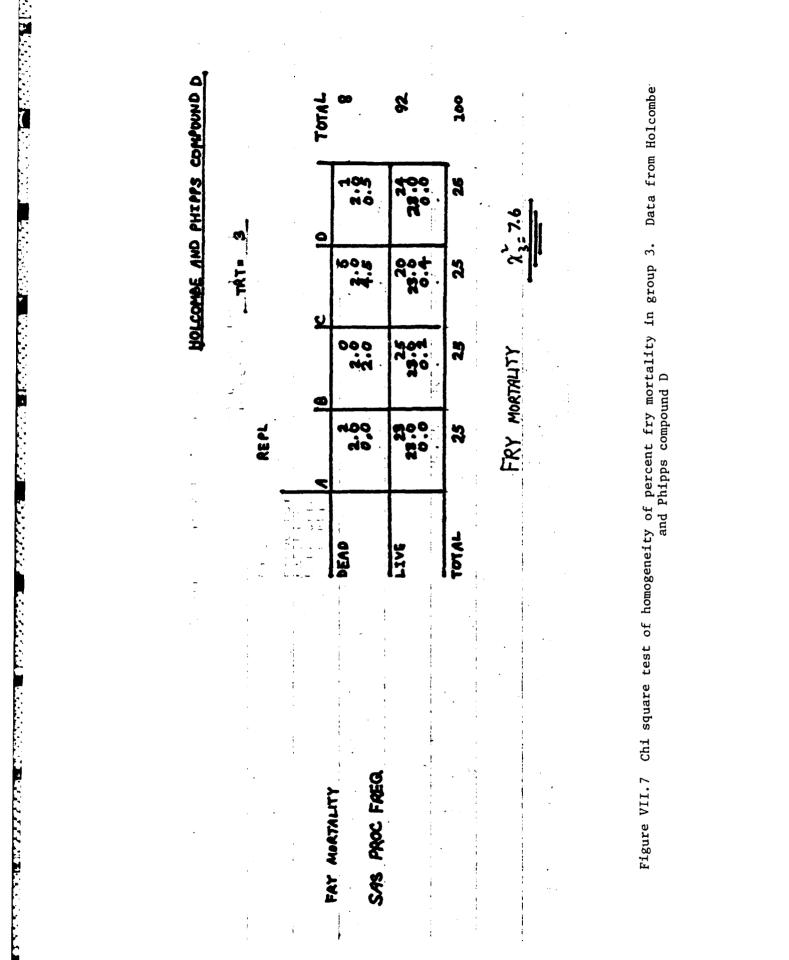
Figure VII.4 Chi square test of homogeneity based on responses pooled across tanks within groups DeFoe compound C fry mortality.

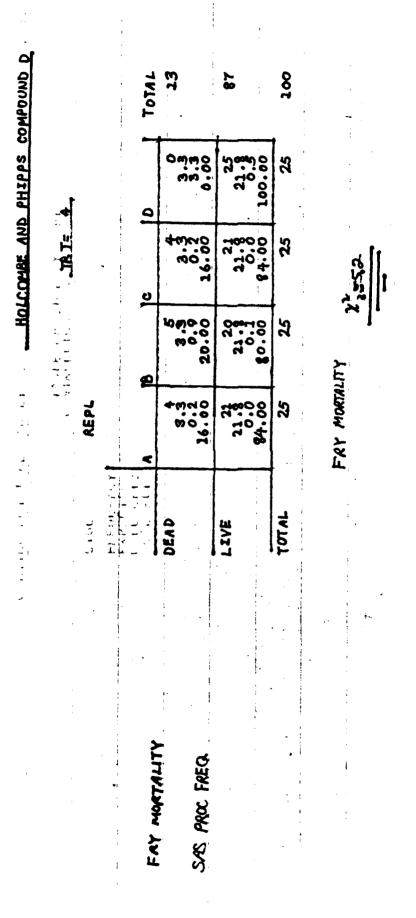
.











1

Chi square test of homogeneity of percent fry mortality in group 4. Data from Holcombe and Phipps compound D Figure VII.8

100 TOTAL 20 2 25 24.00 50 2.2 4.00 36.00 5 FRY MORTALITY 84.001 16.00 50 າມ 1.1 . . REPL 200 000 20.00 3 - بر د د د ر TOTAL LIVE CLLL اللہ لیے۔ ان اللہ ا الہ الم ا DEAD ، ج FRY MORTHLITY SAS PROC FRED

3

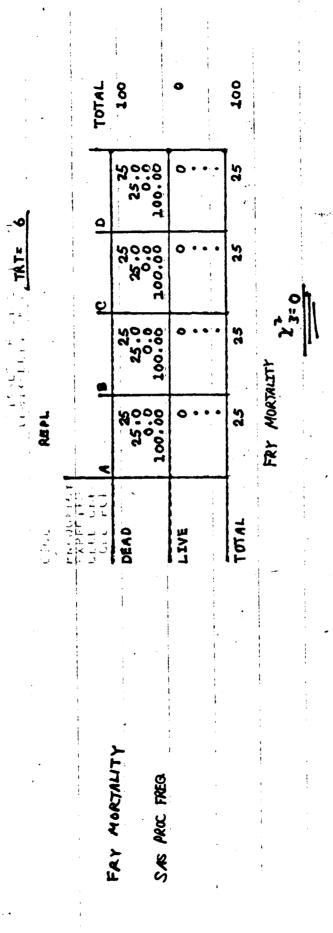
HOLCOMBE AND PHIPPS COMPOUND D

. 2

Data from Holcombe Figure VII.9 Chi square test of homogeneity of percent fry mortality in group 5. and Phipps compound D

TELEVISION AND PHILE COMPOUND D L. INTERPORTED IN T

...



Data from Holcombe Figure VII.10 Chi square test of homogeneity of percent fry mortality in group 6. and Phipps compound D

# VIII. EXAX2 -- A COMPUTER PROGRAM TO TEST FOR HETEROGENEITY OF RESPONSES AMONG TANKS WITHIN GROUPS

We saw in the previous section for DeFoe's fry mortality responses from the test on compound C that we obtained diametrically opposite impressions about the existence of tank to tank heterogeneity within groups depending on whether we based our homogeneity test on a probit model fit or on a congingency table fit. This was attributed to

- 1. Differences in weights used in the denominators of the chi square statistics (based on the assumed model)
- 2. Small expected frequencies that produce substantial departures from the asymptotic distribution theory.

To take account of problem (1), we use the strategy of carrying out separate chi square heterogeneity tests within each concentration group, without imposing any structure on the concentration response relation.

There is, however, a technical problem associated with this scheme. For many (if not most) of the responses of interest the response probabilities of occurrence are fairly close to 0 or 1. Therefore the expected frequencies of occurrence can be rather small, thus invalidating the asymptotic chi square theory, upon which most of the standard tests are based. We saw this in connection with the fry mortality data from the Holcombe and Phipps test on compound D.

We have developed a computer program, EXAX2, that overcomes this problem. It carries out separate chi square tests within each treatment group, based on asymptotic theory when the expected frequencies within cells are large enough and based on <u>exact, small sample</u> theory when the expected frequencies within cells are small. Thus heterogeneity tests using EXAX2 are applicable even with the relatively small samplesizes and relatively extreme response rates encountered in fish toxicity tests. The theory underlying the program and instructions for its use are described in [14] which included as Appendix AVIII.2. In the body of the section we describe the basis of the calculations in EXAX2 and illustrate its application with examples.

EXAX2 pools the results of tests for heterogeneity in each of  $I(I>1) 2 \times K$  independent contingency tables (representing I groups, K tanks per group). The homogeneity test within each group is based on the usual chi square statistic, using either its asymtotic distribution or its exact small sample distribution, as appropriate. The following approach is used.

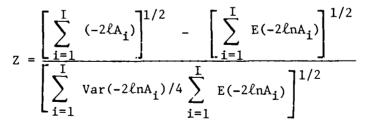
1. Within each concentration group, the chi square for homogeneity among the K tanks is calculated. Let X , N , represent the

number of dead fish and the total number of fish respectively in the j-th tank of the i-th group. Let  $\hat{p}_i$  be a pooled estimate of response probability in the i-th group and  $\hat{q}_i = 1 - \hat{p}_i$ . Then the chi square for group i is

$$\chi_{i}^{2} = \sum_{j=1}^{K} \frac{(x_{ij} - N_{ij}\hat{p}_{i})^{2}}{N_{ij}\hat{p}_{i}\hat{q}_{i}}^{2}$$

If  $\hat{p}_i = 0$  or  $\hat{q}_i = 0$  (corresponding to zero percent or 100 percent observed mortality) the table is degenerate and  $\chi_i^2 = 0$  by definition.

- 2. If all the expected frequencies  $(N_{ij}\hat{p}_i, N_{ij}\hat{q}_i)$  are greater then a specified cutoff value (we currently use five), asymptotic theory is used. Thus the observed significance level of  $\chi^2_i$  is based on the chi square distribution with K - 1 d.f.
- 3. If one or more of the expected frequencies is less than the cutoff level, then the exact distribution of  $\chi_{1}^2$ , conditional on the observed marginal totals, is used. The observed significance level is based on this exact distribution. This approach is described in Agresti and Wackerly [15]. The exact distribution of  $\chi_{1}^2$  is computed by systematically enumerating all possible tables having the given margins using the algorithm in Boulton and Wallace [16] and the associated probabilities due to March [18].
- 4. Let  $A_i$  denote the observed significance level in the i-th group. We pool the  $A_i$ 's over groups to obtain an overall test by an approach analogous to Fisher's method as described in Littell and Folks [19, 20]. For each group we calculate, based on exact or asymptotic theory,  $-2\ell nA_i$  and its mean and variance under the null hypothesis of homogeneity.
- 5. The observed significance levels are pooled into a single statistic by calculating



Z is referred to a standard normal distribution. The null hypothesis of tank to tank homogeneity is rejected for large values of Z. (The square root transformation is used because it represents the variance stabilizing transformation, under asymptotic theory, for  $\Sigma_i(-2\ell_{II}A_i)$  and thus probably improves the normality approximation.)

In addition to calculating preliminary tests of tank to tank heterogeneity within treatment groups, EXAX2 can carry out several other statistical procedures useful in the analysis of data from aquatic toxicity tests. In particular it can:

- Pool data across tanks within groups and test for heterogeneity of response rate <u>across</u> groups by use of the chi square test and either exact small sample theory or asymptotic large samle theory.
- Calculate confidence intervals on the odds ratios of treatment groups to control group using the exact noncentral distribution of Fisher's exact test statistic.

These applications will be discussed in detail in later section.

We now consider several illustrations of the use of EXAX2 for tests of heterogeneity among tanks within groups. The EXAX2 outputs are shown in the referenced figures. The observed and expected cell frequencies are indicated. If any of the expected cell frequencies are lower than the (user-specified) cutoff of 5, exact distribution theory is used. The exact distribution of chi square, conditional on the marginal totals, is enumerated and displayed. The observed value of chi square, the observed significance level,  $-2\ell nA_i$ ,  $E(-2\ell nA_i)$ ,  $Var(-2\ell nA_i)$  are calculated. The six independent tests are combined by summing  $-2\ell nA_i$ ,  $E(-2\ell nA_i)$ ,  $Var(-2\ell nA_i)$  over groups and calculating Z.

DeFoe compound C

- a) Embryo mortality
- b) Fry mortality

Holcombe and Phipps, Compound D

- a) Embryo mortality
- b) Fry mortality

Jarvinen, compound B

- a) Embryo mortality
- b) Fry mortality

#### DeFoe Compound C

# Embryo Mortality

There are two tanks per treatment group, 50 embryo per tank. The results from the EXAX2 calculations are shown in Figures VIII.1 to VIII.6 and are summarized below." The pooled significance level calculatations (using Fisher's method) are presented below the results for group 6. The probability of a standard normal deviate exceeding 0.865 is 0.19.

Embryo	Mortality
--------	-----------

<u>Trt</u>	Method	(Chi sq) XSQØBS	(A <sub>i</sub> ) AI	(-2lnA <sub>i</sub> ) YY	E(-2lnA <sub>i</sub> ) EY	Var(-2lnA <sub>i</sub> ) VARY
1	asympt	0.37205	0.54189	1.22540	2.00	4.00
2	asympt	6.76271	0.00931	9.35371	2.00	4.00
3	asympt	0.38610	0.53436	1.25338	2.00	4.00
4	asympt	1.07250	0.30038	2.40541	2.00	4.00
5	asympt	1.05086	0.30531	2.37286	2,00	4.00
6	asympt	0.0	1.00	0.0	2.00	4.00
	SY = 16.611 Z = 0.86483		YMU = 12.00	SVARY = 24.00		

Except for group 2, where the response from tank 1 appears to be an outlier, there is no statistical evidence of tank to tank heterogeneity within groups.

#### Fry Mortality

There are two tanks per treatment group, 20 fry per tank. The result from the EXAX2 calculations are shown in Figures VIII.7 to VIII.12 and are summarized below. The pooled significance level calculations are presented below the results for group 6. The probability of a standard normal deviate exceeding 0.175 is 0.43.

#### Fry Mortality

Trt	Method	(Chi sq) XSQØBS	(A <sub>i</sub> ) AI	(-2lnA <sub>i</sub> ) YY	E(-2lnA <sub>i</sub> ) EY	Var(-2ℓnA <sub>i</sub> ) VARY
1	(Row total = 0)			0.0	0.0	0.0
2	(Row total = 0)	l .		0.0	0.0	0.0
3	EXACT	2.10526	0.48718	1.43825	0.70068	0.5168
4	EXACT	1.07625	0.48780	1.43568	0.70033	0.51499
5	EXACT	0.14337	1.00	0.0	1.12060	2.7544
6	(Row total = 0)	ł		0.0	0.0	0.0
	SY = 2.8739 Z = 0.17516	YMU = 2.5216		SVARY = 3.7862		

\*Figures VIII.1 to VIII.36 are contained in Appendix AVIII.1.

Trt	Method	(Chi Sq) XSQØBS	(Ai) AI	(-2lnA <sub>i</sub> ) YY	E(-2lnA <sub>i</sub> ) EY	Var(-2lnA <sub>i</sub> ) VARY
1	EXACT	0.70922	1.00	0.0	1.3544	2.7875
2	EXACT	1.08696	0.95647	0.08901	1.543	3.4415
3	EXACT	7.06870	0.07579	5.15948	1.543	3.4415
4	EXACT	5.21662	0.18667	3.35688	1.7349	2.6708
5	asympt	6.44967	0.01967	4.77915	2.0	4.0
6	(Row total = $0$ )		1.0000	0.0	0.0	0.0
	SY = 13.385 Z = 1.09755		YMU = 8	3.1752	SVARY	7 = 17.341

#### Fry Mortality

Groups 1, 2, 4, 6 show no statistical evidence of tank to tank heterogeneity. Groups 3, 6 show some marginal suggestion of tank to tank heterogeneity. It is interesting to note that in direct analogy with the results for embryo mortality, tank 3 of group 3 has about twice the mortality of the other tanks in the group. This "coincidence" should be further investigated to determine if this increased mortality has a systematic cause. Overall, Z = 1.10. The probability of a standard normal random variable exceeding this value by chance is 0.136. Thus there is at most a marginal suggestion of some possible tank to tank variation, but nothing conclusive.

### Jarvinen Compound B

## Embryo Mortality

There are two tanks per treatment group, approximately 50 embryos per tank (actually, between 48 and 57 with an average of 51.2). The results from the EXAX2 calculations are shown in Figures VIII.25 to VIII.30 and are summarized below. The pooled significance level calculations are given below the results for group 6. The probability of a standard normal deviate exceeding 2.54 is 0.005.

#### Embryo Mortality

		(Chi Sq)	(A <sub>i</sub> )	(-2lnA <sub>i</sub> )	E(-2lnA <sub>i</sub> )	Var(-2lnA <sub>i</sub> )
Trt	Method	XSQOBS	AĪ	YY	EY	VARY
1	Asympt	6.51208	0.01071	9.07234	2.0	4.0
2	EXACT	1.78430	0.27477	2.58361	1.52817	3.23307
3	EXACT	3.05250	0.15951	3.67136	1.17120	2.77808
4	EXACT	0.00085	1.0000	0.00000	1.04511	1.38125
5	EXACT	0.74812	0.43704	1.65548	1.45242	2.94750
6	EXACT	4.75938	0.05966	5.63828	1.52125	3.17408
	SY = 22.62107 Z = 2.54488		YMU = 8.	71816	SVARY	= 17.51398

Groups 1 and 6 show significant differences between mortality rates in replicate tanks. Overall (Z = 2.54) the heterogeneity statistic is significant at  $\alpha \cong 0.005$  level. Thus overall there is strong statistical evidence of tank to tank heterogeneity.

Group 1 shows considerable tank to tank heterogeneity in response, group 6 shows moderate tank to tank heterogeneity in response, and group 3 shows marginal tank to tank heterogeneity in response.

If we plot mortality rate by group number we obtain

Embryo mortality rate

.30 -.25 x .20 .15 х х .10 х х .05 х х х х Group 0 1 2 3 5 6

In agreement with the DeFoe and Holcombe and Phipps embryo mortality results, we see no trend in embryo mortality rate with increasing toxicant concentration. We see tank to tank heterogeneity in group 1 and to a lesser extent in groups 3, 6. There is the suggestion that the response from tank 1 of group 1 might be an outlier. This will be considered further in section X.

### Fry Mortality

There are two tanks per treatment group, approximately 15 fry per tank (between 14 and 16 with an average of 14.9). The results from the EXAX2 calculations are shown in Figures VIII.31 to VIII.36 and are summarized below. The pooled significance level calculations are given below the results for group 6. The probability of a standard normal deviate exceeding -0.84 is 0.80.

#### Fry Mortality

		(chi sq)	(A <sub>i</sub> )	(-2lnA <sub>i</sub> )	±	Var(2lnA <sub>i</sub> )
Trt	Method	XSQOBS	AI	YY	EY	VARY
1	(Row total = $0$ )	)		0.0	0.0	0.0
2	(Row total = $0$ )	)		0.0	0.0	0.0
3	(Row total = $0$ )	)		0.0	0.0	0.0
4	(Row total = $0$ )	)		0.0	0.0	0.0
5	asympt	0.13393	0.71439	0.67264	2.0	4.0
6	(Row total = $0$ )	)		0.0	0.0	0.0

SY = 0.67264 YMU = 2.0 SVARY = 4.0Z = -0.84013

This test does not reveal the concentration-response curve very well. In groups 1-4 no fry died while in group 6 all the fry died. Thus the tables are degenerate in 5 of the 6 treatment groups. In group 5 there is no suggestion of tank to tank heterogeneity. Thus overall, there is no suggestion of tank to tank heterogeneity.

In summary we have seen several different degrees of tank to tank heterogeneity within groups in the three toxicity tests studied. With respect to embryo mortality the DeFoe test shows no suggestion of heterogeneity with the exception of an isolated outlier, the Holcombe and Phipps test reveals possible suggestion of heterogeneity, and the Jarvinen test reveals strong suggestion of heterogeneity but this may also be due to an outlier. With respect to fry mortality the DeFoe and Jarvinen tests show no suggestion of tank to tank heterogeneity within groups. The Holcombe and Phipps tests shows possible suggestion of tank to tank heterogeneity but the abberrant looking response originates in precisely the same tank as does the aberrant looking embryo mortality response. This raises questions about both responses. In brief, there does not appear to be very much tank to tank heterogeneity within groups and that which does occur may be due to isolated outlying results.

## IX. ADJUSTMENTS TO ACCOUNT FOR TANK TO TANK HETEROGENEITY WITHIN TREAT-MENT GROUPS

# A. Background, Derivations, and Discussions

We have considered the problem of testing for tank to tank heterogeneity within treatment groups. The results of such tests will influence the way we treat the data in subsequent analyses. Many methods for analyzing qualitative dose response data tacitly assume that there is no tank to tank variation in response rates within groups. Binomial distribution theory is used on data pooled across tanks within groups. Sometimes the assumption of lack of tank to tank heterogeneity is reasonable, as we have seen with the DeFoe and Jarvinen fry mortality data. Sometimes there is borderline statistical evidence of tank to tank heterogeneity, as was the case with the Holcombe and Phipps test (both for the embryo mortality and fry mortality data). In other situations, such as in the embryo mortality data from Jarvinen's test on methyl parathion encapsulated, there is stronger statistical evidence of tank to tank heterogeneity, in at least some of the treatment groups.

In this section we consider methods for accounting for tank to tank heterogeneity when it exists. Three main approaches are possible.

- 1. We can formulate models that explicitly account for tank to tank heterogeneity within groups and fit these models to the data by specialized techniques such as maximum likelihood estimation, using special purpose computer programs. Two such models are the beta binomial [21] and the correlated binomial [22]. This approach requires the formulation of specialized models and development of specialized programs to implement these analyses. Thus such analyses will be difficult for experimenters to carry out and the results of such analyses will be more difficult to interpret.
- 2. We can carry out analyses on a per tank basis rather than on a per fish basis. That is, summary values of such as percent mortality, average weight gain, etc are calculated within each tank and are then used as basic values for subsequent analyses. This is currently the most commonly used approach for analyzing fish toxicity data. While it does correctly account for possible tank to tank heterogeneity, it does so at the cost of considerable reduction in sensitivity. Namely, the data from perhaps 50 to 100 fish or embryos per group are summarized by just two to four summary values. This leaves very few degrees of freedom for estimating error and so diminishes the sensitivity of the subsequent procedures.

3. We can adjust the data to reflect the increased variability due to tank to tank heterogeneity and then use "standard", binomial based techniques on the adjusted "data". This third approach is a workman-like approach and has the dual virtues of being simple to carry out and of permitting the use of "standard" statistical procedures and computer programs for subsequent analyses.

Heterogeneity among tanks within groups can be alternatively regarded as correlation among the responses of the various fish within the same tank. Such correlation is usually positive and this has the effect of increasing the variability of statistics over and above that which would be assumed under a binomial model.

The increased variability can be accounted for by reducing the actual sample size in each tank to an <u>effective sample size</u> and then disregarding the correlation. The number of responses is reduced proportionately so that the observed response rate within each tank remains constant. Suppose for instance there are 40 embryo in a tank and 8 die. We thus have an observed response rate of 0.20. Suppose that the responses within each tank are positively correlated and the variance of  $\hat{p}$  is inflated 20 percent by this correlation. That is

$$Var(\hat{p}) = \frac{1.2p(1 - p)}{40}$$

Then we can regard the effective sample size within that tank as 40/1.2 = 33.33. To maintain the response rate at the observed level of .20 we adjust the number of responses down to a corresponding <u>effective</u> number 8/1.2 = 6.67. We then analyze the data from this tank by ignoring the correlation and treating the data as if we have 6.67 responses in 33.33 trials. All the standard analysis procedures, predicated on the assumption of no tank to tank heterogeneity within groups, can be applied to the modified "data".

The per tank analyses mentioned in paragraph 2 can be regarded as a limiting case of data adjustment where we adjust the effective sample sizes within tanks all the way down to 1.

We now consider the calculation of adjustment factors. Motivation for the adjustment procedure comes from the form of the beta binomial model [21]. Namely suppose  $X_{ij}$  is the number of responses within tank j of treatment group i. The beta binomial model extends the binomial model to allow for tank to tank variation within groups. Thus we assume

$$X_{ij} \sim Binomial (N_{ij}, P_{ij})$$
 conditional on  $P_{ij}$ .  $j = 1, ..., J_{i}$   
 $i = 1, ..., I$ 

where

 $p_{ij} \sim Beta(\alpha_i, \beta_i)$ 

N<sub>ij</sub> are fixed.

and

Let 
$$\mu_{i} \equiv \frac{\alpha_{i}}{\alpha_{i} + \beta_{i}}$$
  $\theta_{i} \equiv \frac{1}{\alpha_{i} + \beta_{i}}$   
Then  $E(p_{ij}) = \mu_{i}$   $Var(p_{ij}) = \mu_{i}(1 - \mu_{i})\frac{\theta_{i}}{1 + \theta_{i}}$ 

When  $\theta_i = 0$ ,  $Var(p_{ij}) = 0$  and we are back to the case of no tank to tank heterogeneity, at least within the i-th group. The larger  $\theta_i$  is, the greater is the extent of tank to tank heterogeneity.  $\theta_i$  varies between 0 and  $\infty$ .

$$L(X_{ij}, p_{ij}) \sim Binomial (N_{ij}, p_{ij})$$
  
 $L(p_{ij}) \sim Beta (\alpha_i, \beta_i).$ 

These two facts imply that  $X_{ij}$  has the marginal beta binomial distribution with probability function

$$P(X_{ij} = x) = {\binom{N_{ij}}{x^j}} \frac{B{\binom{\mu_i}{\theta_i} + x, \frac{1 - \mu_i}{\theta_i} + N_{ij} - x}}{B{\binom{\mu_i}{\theta_i}, \frac{1 - \mu_i}{\theta_i}}} \qquad x = 0, 1, ..., N_{ij}$$

It can be shown directly that

$$E(X_{ij}) = N_{ij}\mu_{i}$$
  
$$Var(X_{ij}) = N_{ij}\mu_{i}(1 - \mu_{i})\left[\frac{1 + N_{ij}\theta_{i}}{1 + \theta_{i}}\right] \qquad 0 < \theta_{i} < \infty$$

We see that the variance of  $X_{ij}$  is inflated over and above a binomial variance by a multiplicative factor.

Suppose  $N_{ij} = N_i$  j = 1, ..., J. This assumption is reasonable in fish toxicity tests where  $N_{ij}$  represents the number of fish or embryos within the j-th tank of the 1-th group. In this case the multiplicative factor becomes  $[(1 + N_i\theta_i)/(1 + \theta_i)] \equiv K_i$ , j = 1, ..., J. Thus  $Var(X_{ij}) = N_i K_i \mu_i (1 - \mu_i)$  where  $1 \le K_i^{\infty}$ . Define  $\hat{p}_{ij} = X_{ij}/N_{ij}$ .  $\hat{p}_{ij}$  is the observed response proportion.

Therefore

$$Var(\hat{p}_{ij}) = \frac{K_i}{N_i} \mu_i (1 - \mu_i) \qquad j = 1..., J$$

Thus the effective sample size is  $N_i/K_i$ . As the extent of tank to tank heterogeneity approaches  $0(i.e. \ \theta_i \rightarrow 0)$ ,  $K_i$  approaches 1 and  $N_i/K_i$  approaches  $N_i$ . As the extent of tank to tank heterogeneity gets greater and greater,  $K_i \rightarrow N_i$  and so  $N_i/K_i$  approaches 1. Thus the two extreme situations are not adjusting the within tank sample sizes at all and adjusting the within tank sample sizes at all and adjusting the within tank basis rather than on a per fish basis. Thus method 2 for accounting for tank to tank heterogeneity can be regarded as an extreme case of method 3. Note that if  $N_i \equiv N$  and  $\theta_i \equiv \theta$  for all i, then  $K_i \equiv K$  for all i.

The procedure suggested here for calculating adjustment factors is motivated by the results based on beta binomial theory, but is simpler to carry out.

Let  $X_{ij}$ ,  $N_i$  denote the number of responses and the total number of fish respectively within tank j of group i,  $j = 1, \ldots, J$ . Let  $\hat{p}_{ij} \equiv X_{ij}/N_i$ . Let  $p_i$  denote the average response rate within the i-th group. The actual variance of  $\hat{p}_{ij}$  is the binomial theory variance multiplied by the inflation factor  $K_i$ . Thus

$$K_{i} = \frac{\operatorname{Var}(\hat{p}_{ij})}{\left[p_{i}(1 - p_{i})/N_{i}\right]}$$

This suggests that we can estimate  $K_i$  by estimating  $Var(\hat{p}_{ij})$  and  $p_i$  by their sample analogues. Let

 $\bar{N}_{i} \equiv J^{-1}\Sigma_{j=1}^{J}N_{ij}, \hat{\bar{p}}_{i} \equiv J^{-1}\Sigma_{j=1}^{J}\hat{p}_{ij}, \tilde{V}ar(\hat{p}_{ij}) \equiv (J-1)^{-1}\Sigma_{j=1}^{J}(\hat{p}_{ij} - \hat{\bar{p}}_{i})^{2}$ denote the average sample size, the average observed response rate, and the sample variance of response rates within the i-th group respectively. Note that the  $N_{ij}$ 's are generally nearly equal in fish toxicity data. We estimate K<sub>i</sub> as

$$\hat{\tilde{K}}_{i} = \hat{\tilde{V}}_{ar}(\hat{\tilde{p}}_{ij}) / [\hat{\tilde{p}}_{i}(1 - \hat{\tilde{p}}_{i}) / \tilde{N}_{i}]$$

The numerator of this ratio is the observed variance among the  $\hat{p}_{ij}$ 's while the denominator is the variance that would be expected just due to binomial variation. We adjust each  $X_{ij}$ ,  $N_{ij}$  in the group downward by a factor  $\hat{k}_i$ .

Notes:

- 1.  $K_i$  is necessarily greater than 1 but  $\hat{K}_i$  may not be. If  $\hat{K}_i < 1$  or if there is no statistical evidence of tank to tank heterogeneity then we should not adjust sample sizes.
- 2. Assuming binomial theory when there is in fact tank to tank heterogeneity results in <u>underestimation</u> of the variabilities of

the various statistics calculated. Thus hypothesis tests comparing treatment group and control group response rates will reject more often than they should, thereby resulting in <u>underestimation of noeffect levels</u>. However the opposite effect occurs with respect to inferences about safe concentrations based on dose response curves. Underestimation of variability results in overly large lower confidence bounds on safe concentration. A nominal 95% lower confidence bound may in fact be just an 80% lower confidence bound.

- 3. The decision as to when to adjust sample sizes downward should be reasonably liberal, perhaps when there is statistical evidence of tank to tank heterogeneity at the  $\alpha$  = .20 or the  $\alpha$  = .25 level. However K, must always be greater than or equal to 1.
- 4. The calculation of  $K_i$  by means of ratios of variances is inefficient. A more precise way of determining  $K_i$  would be to estimate  $\theta_i$  from the data by maximum likelihood estimation and substitute this estimate,  $\hat{\theta}_i$ , into the expression for  $K_i$ . Such an estimate would always be greater than or equal to 1. However such an approach would require special purpose programs. The estimation of  $K_i$  as discussed in this section is simpler and can be carried out by hand calculation. However in the future we will look in calculation of  $\hat{K}_i$ 's by means of maximum likelihood estimation based on the beta binomial model.
- 5. We can calculate separate  $K_1$ 's for each treatment group based on the responses solely from the tanks in that group. Alternatively if  $N_i = N$  for all i and if  $\theta_i = \theta$  for all i then  $K_i = K$  for all i. We can then calculate a common inflation factor  $\hat{K}$  for all treatment groups. The question of whether we should fit separate adjustment factors within each group or a single common factor is a research problem in its own right. We defer the answer to that question to future work and in this report confine attention to fitting common adjustment factors for all groups.

If it is sensible on biological grounds, results on tank to tank heterogeneity observed in previous similar tests might be combined with current results to obtain a more accurate adjustment factor.

6. The adjustment procedure might take into account the statistical precisions of the estimates  $\hat{K_i}$ ,  $\hat{K}$ . A conservative way to do this would be to use upper confidence bounds on  $\hat{K_i}$ ,  $\hat{K}$  as adjustment factors rather than the point estimates. This modification will also await future work.

#### B. Illustrations

We illustrate the application of this adjustment procedure to several sets of data. First consider the fry mortality data of Holcombe and Phipps for compound D. From the preliminary test of tank to tank heterogeneity within treatment groups we conclude that there is at most marginal statistical evidence of tank to tank heterogeneity within treatment groups. (The observed significance level is 0.14). In this example J = 4,  $N_{ij} = 25$  for all i, j.

Group 1: 
$$\hat{p}_{11} = 0.08, \ \hat{p}_{12} = 0.08, \ \hat{p}_{13} = 0.04, \ \hat{p}_{14} = 0.04, \ \hat{p}_{1} = 0.06, \ \bar{N}_{1} = 25, \ \hat{V}ar(\hat{p}_{1j}) = 0.00053, \ \hat{p}_{1}(1 - \hat{p}_{1})/\bar{N}_{1} = 0.00226$$
$$\hat{K}_{1} = \frac{\hat{V}ar(\hat{p}_{1j})}{[\hat{p}_{1}(1 - \hat{p}_{1})/\bar{N}_{1}]} = \frac{0.00053}{0.00226} = 0.236$$

Group 2: 
$$\hat{p}_{21} = 0.12, \ \hat{p}_{22} = 0.04, \ \hat{p}_{23} = 0.08, \ \hat{p}_{24} = 0.08, \ \hat{\bar{p}}_2 = 0.08, \ \bar{\bar{p}}_2 = 0.00107, \ \hat{\bar{p}}_2(1 - \hat{\bar{p}}_2)/\hat{\bar{N}}_2 = 0.00294$$
  
 $\hat{K}_2 = \frac{\hat{Var}(\hat{p}_{2j})}{[\hat{\bar{p}}_2(1 - \hat{\bar{p}}_2)/\hat{\bar{N}}_2]} = \frac{0.00107}{0.00294} = 0.362$ 

Group 3:

$$\hat{p}_{31} = 0.08, \ \hat{p}_{32} = 0.00, \ \hat{p}_{33} = 0.20, \ \hat{p}_{34} = 0.04, \ \hat{\bar{p}}_{3} = 0.08, \ \bar{\bar{N}}_{3} = 25, \ Var(\hat{p}_{3j}) = 0.00747, \ \hat{p}_{3}(1 - \hat{\bar{p}}_{3})/\bar{N}_{3} = 0.00294$$

$$\hat{\mathbf{x}}_{3} = \frac{\hat{\mathbf{var}}(\hat{\mathbf{p}}_{3j})}{[\hat{\bar{\mathbf{p}}}_{3}(1 - \hat{\bar{\mathbf{p}}}_{3})/\bar{\mathbf{N}}_{3}]} = \frac{0.00747}{0.00294} = 2.536$$

Group 4:

$$\hat{p}_{41} = 0.16, \ \hat{p}_{42} = 0.20, \ \hat{p}_{43} = 0.16, \ \hat{p}_{44} = 0.00, \ \hat{\bar{p}}_{4} = 0.13, \ \hat{\bar{N}}_{4} = 25, \ \hat{V}ar(\hat{p}_{4j}) = 0.00787, \ \hat{\bar{p}}_{4}(1 - \hat{\bar{p}}_{4})/\bar{N}_{4} = 0.00452$$

$$\hat{\mathbf{K}}_{4} = \frac{\sqrt{4r(\mathbf{p}_{4j})}}{[\hat{\mathbf{p}}_{4}(1 - \hat{\mathbf{p}}_{4})/\bar{\mathbf{N}}_{4})} = \frac{0.00787}{0.00452} = 1.739$$

Group 5: 
$$\hat{p}_{51} = 0.92, \ \hat{p}_{52} = 0.84, \ \hat{p}_{53} = 0.64, \ \hat{p}_{54} = 0.76, \ \hat{\bar{p}}_{5} = 0.79$$
  
 $\bar{N}_{5} = 25, \ \hat{V}ar(\hat{p}_{5j}) = 0.01427, \ \hat{\bar{p}}_{5}(1 - \hat{\bar{p}}_{5})/\bar{N}_{5} = 0.00664$   
 $\hat{V}ar(\hat{p}_{5j}) = 0.01427$ 

$$\hat{K}_{5} = \frac{\sqrt{ar(p_{5j})}}{[\hat{\bar{p}}_{5}(1 - \hat{\bar{p}}_{5})/\bar{N}_{5}]} = \frac{0.01427}{0.00664} = 2.150$$

Group 6:

 $\hat{p}_{61} = \hat{p}_{62} = \hat{p}_{63} = \hat{p}_{64} = 1.0, \ \hat{p}_{6} = 1.0, \ \hat{\bar{N}}_{6} = 25, \ \hat{V}ar(\hat{p}_{6j}) = 0.0,$  $\hat{\bar{p}}_{6}(1 - \hat{\bar{p}}_{6})/\bar{N}_{6} = 0.0$  $\hat{K}_{c}$  is indeterminate and so we take it to be 1.0.

Thus,

$$\hat{\vec{K}} = \frac{ \stackrel{6}{\Sigma} \hat{\vec{K}}_{i} }{6} = \frac{0.236 + 0.362 + 2.536 + 1.739 + 2.150 + 1.000}{6} = 1.337$$

 $ar{K}$  is the average adjustment factor which is used to adjust the observed sample sizes to effective sample sizes. The sample sizes are adjusted downward so as to maintain the observed response rates within each tank. The results of the adjustment procedure are presented in Table IX.1. These adjusted values are used as basic input "data" for subsequent analyses. We then proceed as if there is no tank to tank variation within groups. The extrabinomial variation has been accounted for by the adjustment procedure.

Group		Tank A	Tank B	Tank C	Tank D
1	Dead	1.50	1.50	0.75	0.75
	Live	17.20	17.20	17.95	17.95
	Total	18.70	18.70	18.70	18.70
2	Dead	2.24	0.75	1.50	1.50
	Live	16.45	17.95	17.20	17.20
	Total	18.70	18.70	18.70	18.70
3	Dead	1.50	0.00	3.74	0.75
	Live	17.20	18.70	14.96	17.96
	Total	18.70	18.70	18.70	18.70
4	Dead	2.99	3.74	2.99	0.00
	Live	15.70	17.96	15.70	18.70
	Total	18.70	18.70	18.70	18.70
5	Dead	17.20	15.70	11.97	14.21
	Live	1.50	2.99	6.73	4.49
	Total	18.70	18.70	18.70	18.70
6	Dead	18.70	18.70	18.70	18.70
	Live	0.00	0.00	0.00	0.00
	Total	18.70	18.70	18.70	18.70

# TABLE IX.1 EFFECTIVE SAMPLE SIZES AND RESPONSES IN HOLCOMBE AND PHIPPS<br/>COMPOUND D FRY MORTALITY DATA AFTER ADJUSTMENT<br/>FOR TANK TO TANK HETEROGENEITY

58

-

We next illustrate the adjustment procedure on the embryo mortality data of Jarvinen for compound B. The preliminary test of tank to tank heterogeneity within treatment groups is highly significant (Z = 2.54, corresponding to an observed significance level of 0.005). This statistical significance is due to group 1 (control) which shows strong tank to tank differences, group 6 which shows a moderate tank to tank difference, and group 3 which shows possible indications -- but at best weak statistical evidence -- of tank to tank heterogeneity.

In this example J = 2,  $N_{ij}$  are close to  $\overline{N}_i$  (within 1) except for group 2. We will assume here that  $N_{ij} = \overline{N}_i$  when calculating  $K_i$ 's. This assumption can be refined somewhat, if necessary, to calculate separate adjustment factors for each tank, but we will not do that here. This will await the development of adjustment procedures based on maximum likelihood estimation.

Group 1: 
$$\hat{p}_{11} = .235$$
,  $\hat{p}_{12} = .04$ ,  $\hat{\bar{p}}_1 = 14/101 = 0.138$ ,  $N_{11} = 51$ ,  $N_{12}$   
50,  $\bar{N}_1 = 50.5$ 

$$\hat{V}ar(\hat{p}_{ij}) = 0.019$$
  $\hat{\bar{p}}_{1}(1 - \hat{\bar{p}}_{1})/\bar{N} = 0.00236$ 

$$\hat{K}_{1} = \frac{\hat{V}ar(\hat{p}_{1j})}{[\hat{\bar{p}}_{1}(1 - \hat{\bar{p}}_{1})/\bar{N}_{1}]} = \frac{.019}{.00236} = 8.056$$

<u>Group 2:</u>  $\hat{p}_{21} = .105$ ,  $\hat{p}_{22} = .038$ ,  $\hat{\bar{p}}_2 = 8/109 = .073$ ,  $N_{21} = 57$ ,  $N_{22} = 52$ ,  $\bar{N}_2 = 54.5$ 

Group 3: 
$$\hat{p}_{31} = .04, \ \hat{p}_{32} = 1.4, \ \hat{\bar{p}}_{3} = 9/100 = .09, \ N_{31} = N_{32} = \bar{N}_{3} = 50$$
  
 $\hat{V}ar(\hat{p}_{3j}) = .005$   
 $\hat{\bar{p}}_{3}(1 - \hat{\bar{p}}_{3})/\bar{N}_{3} = .0016$   
 $\hat{K}_{3} = \frac{\hat{V}ar(\hat{p}_{3j})}{[\hat{\bar{p}}_{3}(1 - \hat{\bar{p}}_{3})/\bar{N}_{3}]} = \frac{.005}{.0016} = 3.053$ 

1

<u>Group 4</u>:  $\hat{p}_{41} = .02, \ \hat{p}_{42} = .021, \ \hat{\bar{p}}_4 = 2/98 = .020, \ N_{41} = 50, \ N_{42} = 48, \ \bar{N}_4 = 49$  $\hat{N}_4 = 49$  $\hat{V}ar(\hat{p}_{4j}) = 3.4 \times 10^{-7}$   $\hat{\bar{p}}_4(1 - \hat{\bar{p}}_4)/\bar{N}_4 = .00041$  $\hat{K}_4 = \frac{\hat{V}ar(\hat{p}_{4j})}{[\hat{\bar{p}}_4(1 - \hat{\bar{p}}_4)/\bar{N}_4]} = \frac{3.4 \times 10^{-7}}{.00041} = 8 \times 10^{-4} = 0$ 

<u>Group 5</u>:  $\hat{p}_{51} = .077, \ \hat{p}_{52} = .038, \ \hat{\bar{p}}_5 = 6/105 = .057, \ N_{51} = 52, \ N_{52} = 53, \ \bar{N}_5 = 52.5$ 

$$\hat{v}_{ar}(\hat{p}_{5j}) = .00077 \qquad \qquad \hat{\bar{p}}_{5}(1 - \hat{\bar{p}}_{5})/\bar{N}_{5} = .00102$$
$$\hat{\kappa}_{5} = \frac{\hat{v}_{ar}(\hat{p}_{5j})}{[\hat{\bar{p}}_{5}(1 - \hat{\bar{p}}_{5})\bar{N}_{5}]} = \frac{.00077}{.00102} = .750$$

<u>Group 6:</u>  $\hat{p}_{61} = .02, \ \hat{p}_{62} = .137, \ \hat{\bar{p}}_{6} = 8/101 = .079, \ N_{61} = 50, \ N_{62} = 51, \ N_{6} = 50.5 \ \hat{v}_{ar}(\hat{p}_{6j}) = .00687 \qquad \hat{\bar{p}}_{6}(1 - \hat{\bar{p}}_{6})/\bar{N}_{6} = .00144$  $\hat{K}_{6} = \frac{\hat{v}_{ar}(\hat{p}_{6j})}{[\hat{\bar{p}}_{6}(1 - \hat{\bar{p}}_{6})/\bar{N}_{6}]} = \frac{.00687}{.00144} = 4.77$ 

Assuming that the relatively high embryo mortality response in group 1 tank 1 is <u>not</u> an outlier and that the inflation in variability is constant across groups, we calculate an <u>average</u> adjustment factor across treatment groups.

Thus,  

$$\hat{\vec{K}} = \frac{\hat{\vec{\Sigma}K}}{6} = \frac{8.056 + 1.799 + 3.053 + 0 + 0.750 + 4.77}{6} = 3.071$$

 $\hat{\vec{K}}$  is the average adjustment factor across groups. Alternatively we might use the separate adjustment factors within groups. The results of the adjustment procedure are presented in Table IX.2. These adjusted values can be used as basic input "data" for subsequent analyses. We proceed with further analyses as if there is no tank to tank heterogeneity within groups.

Group		Tank A	Tank B	······································
1	Dead	3.91	0.65	
	Live	12.70	15.63	
	Total	16.61	16.28	
2	Dead	1.95	0.65	
	Live	16.61	16.28	
	Total	18.56	16.93	
3	Dead	0.65	2.28	
	Live	15.63	14.00	
	Total	16.28	16.28	
4	Dead	0.33	0.33	
	Live	15.95	15.30	
	Total	16.28	15.63	
5	Dead	1.30	0.65	
	Live	15.63	16.61	
	Total	16.93	17.26	
6	Dead	0.33	2.28	
	Live	15.95	14.33	
	Total	16.28	16.61	

## TABLE IX.2 EFFECTIVE SAMPLE SIZES AND RESPONSES IN JARVINEN COMPOUND B EMBRYO MORTALITY DATA AFTER ADJUSTMENT FOR TANK TO TANK HETEROGENEITY

2

M

## X. OUTLIER DETECTION PROCEDURES

#### A. Background

Another preliminary analysis of importance is the detection of responses which do not appear to be in conformance with the substantial majority of responses. Such exceptional responses are often referred to as "outliers". Outlier detection procedures are used to decide how extreme a response must be in order to rule out the possibility that its value is reasonably likely to be due just to random variation. Consider for example the percentage embryo mortality responses from DeFoe's test on compound C that are displayed in Figure VI.1. We remarked that the mortality rate in group 2, tank A appears to be widely separated from the others. Can such a separation be explained by random variation or is there some systematic factor peculiar to this tank? Similarly, the percentage embryo mortality observed in group 1, tank A in Jarvinen's test on compound B is widely separated from the other responses. Can a separation of this magnitude be reasonably explained by random variation or is there some systematic factor peculiar to this tank?

Barnett and Lewis [23] describe a wide class of outlier detection procedures, to screen out those extreme responses that cannot be reasonably attributed to random variation. They include a procedure appropriate for binomial responses (section 3.4, pp 122-124). Their procedure is based on the assumption of n independent responses X1, ..., Xn, each binomially distributed with parameters m and p. They base their outlier test on the exact conditional distribution of maxX<sub>i</sub> given  $\Sigma_i X_i$ . In our data n represents the number of tanks per group, m is the number of embryos or fry per tank (assumed to be equal from tank to tank), and  $X_i$  is the number of responses (e.g. dead embryos) per tank. Their tabulation, Table XIX (pp 320-322) includes only the range of values  $n \ge 3$ ,  $m \ge 10$ ,  $X_{(n)} = m$ , m - 1, m - 2. This is quite inadequate for the ranges of parameters and responses that arise in toxicity tests. Thus their exact conditional test is not too useful for our needs.

Barnett and Lewis state, on page 123, that an alternative, approximate approach to outlier detection in the binomial case is to transform  $\{X_j/m\}$  using the arc sine transformation and then apply normal theory based procedures to these transformed values. This approach, and variants on it, are in the spirit of the methods that we recommend in the remainder of this section. We consider both graphical and numerical procedures.

The theoretical bases of our suggested methods are discussed in Appendix AX.

# B. <u>Application of Outlier Detection Procedures to Fish Toxicity</u> Data

We apply the transformations discussed in Appendix AX to construct graphical outlier detection procedures based on normal probability plotting techniques and associated formal outlier detection tests. We apply these procedures to the following situations:

DeFoe: compound C

The second se

embryo mortality data fry mortality data

Holcombe and Phipps: Compound D embryo mortality data fry mortality data

Jarvinen: compound B

embryo mortality data

DeFoe compound C

7

$$I = 6, J = 2$$

(i.e. 6 groups, 2 tanks per group).

Embryo Mortality

Group 1	$\hat{p}$ = .41; all expected frequencies are greater than 5.
Group 2	$\hat{p}$ = .51; all expected frequencies are greater than 5.
Group 3	$\hat{p}$ = .37; all expected frequencies are greater than 5.
Group 4	$\hat{p}$ = .37; all expected frequencies are greater than 5.
Group 5	$\hat{p}$ = .39; all expected frequencies are greater than 5.
Group 6	$\hat{p}$ = .38; all expected frequencies are greater than 5.

We apply the transformation suggested in case 1. Data summaries are given below.

Calculate	1	$\frac{N}{j}$	$\left(\frac{X_{j} - N_{j}\hat{p}}{\sqrt{N_{j}\hat{p}\hat{q}}}\right)$		for each
-----------	---	---------------	--	--	----------

or each tank within each group.

	x <sub>j</sub>	N j	Ŷ	N p	$\left(N_{j}\hat{p}\hat{q}\right)^{1/2}$	N <sub>j</sub> /N	$\left(1 - \frac{N_{j}}{N}\right)^{1/2} \left(\frac{X_{j} - N_{j}\hat{p}}{\sqrt{N_{j}\hat{p}\hat{q}}}\right)$
Group 1							
Tank A Tank B	22 19	50 50	0.41 0.41	20.5 20.5	3.478 3.478	.50 .50	.610 610
Group 2							
Tank A Tank B	32 19	50 50	0.51 0.51	25.5 25.5	3.535 3.535	.50 .50	2.600 -2.600
<u>Group 3</u>							
Tank A Tank B	20 17	50 50	0.37 0.37	18.5 18.5	3.414 3.414	.50 .50	0.621 -0.621
Group 4							
Tank A Tank B	16 21	50 50	0.37 0.37	18.5 18.5	3.414 3.414	.50 .50	-1.036 1.036
Group 5							
Tank A Tank B	22 17	50 50	0.39 0.39	19.5 19.5	3.449 3.449	.50 .50	1.025 -1.025
Group 6							
Tank A Tank B	19 19	50 50	0.38 0.38	19.0 19.0	3.432 3.432	.50 .50	0 0

To prepare the normal probability plot we order the standardized values and plot the i-th smallest against the plotting position  $100 \times (i - 0.5)/12$  on the probability scale. These values are indicated below.

i	1	2	3	4	5	6
Ordered Value	-2.600	-1.036	-1.025	-0.621	-0.610	0
Plotting Position	4.2	12.5	20.8	29.2	37.5	45.8
i	7	8	9	10	11	12
Ordered Value	0	0.610	0.621	1.025	1.036	2.600
Plotting Position	54.2	62.5	70.8	79.2	87.5	95.8

The normal probability plot of these points is shown in Figure X.1. The plot appears perfectly symmetrical about 0 since J = 2 and so the responses within groups have correlation -1.0. The effective sample size is thus 6(J - 1) = 6(2 - 1) = 6 independent observations. The reference line in the plot is the standard normal distribution function. If the response rates are homogeneous within groups then the standardized values should lie near this line. If there is extrabinomial variation in the data, that is random tank to tank variation within groups, then the points should lie along a line or a curve with steeper slope than the standard distribution function. If there are outlying responses in the data then they should be far removed from the line or curve that typifies the bulk of the data. This latter situation is seen to be the case. The bulk of the data lie very nicely along the standard normal c.d.f. line. The values corresponding to group 2 are far removed from this line.

To determine the extent of statistical evidence that the apparent outliers did not occur just due to chance we calculate the probability that the maximum absolute value of six independent standard normal random variables exceeds 2.600. More precisely let  $Z_1, Z_2, \ldots, Z_6$  be six independent standard normal random variables. Then

 $P[_{j=1,\dots,6} |Z_{j}| \ge 2.600] = P[at \text{ least one } |Z_{j}| \ge 2.600]$ = 1 - P[all |Z\_{j}| \le 2.600] = 1 - {P [|Z\_{1}| \le 2.600]}<sup>6</sup> = 1 - (.9907)<sup>6</sup> = 0.055

This is of borderline statistical significance. We can thus infer that based on this test there is marginal statistical evidence that group 2 contains an outlying tank.

The appearance of Figure VI.1 suggests that the response from Tank 2A is more than a marginal outlier. We can increase the sensitivity of the above outlier test by incorporating additional information. If we assume that there is no trend in response rates across groups then we can estimate the response rate based on all 12 tanks and can ignore the correction factor  $(1 - N_j/N)^{-1/2}$ . In general this assumption will not hold but it seems reasonable in this example based on the appearance of Figure VI.1 and on toxicological considerations (i.e. relatively little penetration of chemical into the embryo). The value of  $\hat{p}$  based on 12 tanks is 243/600 = 0.405. The largest standardized value is that from Tank 2A, namely

$$Z = \frac{32 - (50)(.405)}{[50(.405)(.595)]^{1/2}} = 3.385$$

What is the probability that the most extreme of 11 independent standard normal random variable exceeds 3.385 in absolute value?

 $P[_{j=1,\ldots,12} |Z_j| \ge 3.385] = 1 - \{P[|Z_1| \le 3.385\}^{11} = 1 - (.9997)^{11} = 0.003$ 

We can thus infer that, with the additional assumption of no trend in response with increasing treatment level, there is strong statistical evidence that the response rate in Tank 2A is an outlying value.

Note that just because the Tank 2A response is an outlier does not in and of itself mean that the data should be discarded or disregarded. Rather, the investigator needs to reexamine the records for this tank to determine the reason for the atypical response. If it is due to clerical error, to experimental mishap, to outbreak of a disease unrelated to the toxicant, etc then perhaps the Tank 2A response is inappropriate and should be disregarded. If it represents normal biological variation then the response should be considered with the others. This is a matter for biological judgement. Outlier detection procedures are merely screening devices to direct attention to those places where such biological judgement need be applied.

DeFoe compound C - Fry mortality

Group 1	$\hat{\mathbf{p}} = 0;$			
Group 2	$\hat{\mathbf{p}} = 0;$			
Group 3	$\hat{\mathbf{p}} = 0.05;$			
Group 4	$\hat{p} = 0.024;$			
Group 5	$\hat{\mathbf{p}}$ = 0.225; expected	l frequencies less	than 5 in both	tanks
Group 6	$\hat{p} = 1.00  (\hat{q} = 0);$			

Thus groups 1, 2, 3, 4, 6 correspond to the Poisson case 2. (In group 6 we interchange the roles of p and q). Group 5 corresponds to case 3.

For all groups but 5 we calculate  $\left(1 - \frac{N_j}{N}\right)^{-1/2} 2\left[\sqrt{X_j} - \sqrt{N_j\hat{p}}\right]$ . For group 5 we carry out an arc sine transformation

	x <sub>j</sub>	N j	Ŷ	NjŶ	N j	/N	$(1 - N_j/N)^{-1/2} 2\left[\sqrt{X_j} - \sqrt{N_j\hat{p}}\right]$
Group 1							
Tank A Tank B	0 0	20 20	0 0	0 0	.5 .5		0 0
Group 2							
Tank A	0	20	0	0	.5		0
Tank B	0	20	0	0	.5	0	0
Group 3							
Tank A	0	20	.05	1.0	.5		-2.828
Tank B	2	20	.05	1.0	.5	0	1.172
<u>Group 4</u>							
Tank A	0	21	.024	0.51			-2.049
Tank B	1	20	.024	0.488	3 0.4	88	0.843
	x <sub>j</sub>	N j	ĝ	N ģ	N i	/N	$(1 - N_i/N)^{1/2} 2\left[\sqrt{X_i} - \sqrt{N_i\hat{q}}\right]$
	J	J		J			
<u>Group 6</u>							
Tank A	0	20	0	0	.5		0
Tank B	0	20	0	0	.5	0	0
	Х <sub>ј</sub>	N j	Ŷ	N <sub>j</sub> ∲	N <sub>j</sub> /N	<sup>ĝ</sup> j	$\left(1 - N_j/N\right)^{-1/2} 2\sqrt{N_j} \left[ \arcsin\sqrt{\hat{p}_j} - \arcsin\sqrt{\hat{p}_j} \right]$
Group 5							
Tank A	4	20	0.225	4.5	0.50	0.20	-0.387
Tank B	5	20	0.225	4.5		0.25	0.372

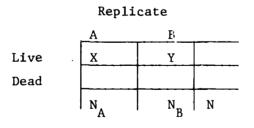
To prepare the normal probability plot we order the standardized values and plot the i-th smallest against the plotting position 100 x (i - 0.5)/12 on the probability scale. These values are indicated below.

i	1	2	3	4	5	6	7	8	9
Ordered Value	-2.828	-2.049	-0.387	0	0	0	0	0	0
Plotting Position		12.5	20.8	29.2	37.5	45.8	54.2	62.5	70.8
i	10	)	11		12				
Ordered Value	0.3	572	0.843	1	.172				
Plotting Position	/4/	· · · · · · · · · · · · · · · · · · ·	87.5	95	5.8				

The normal probability plot of these points appears in Figure X.2. We note two points well below the standard N(0, 1) line. These points correspond to Tanks 3A, 4A. Both correspond to frequencies of 0, where the normal approximation is least reasonable. Furthermore, their companion tanks do <u>not</u> show up as outliers. Thus before we say there is an outlier, we should compare the proportions in the two replicate tanks to see if there is any statistical evidence of differences.

There is an exact test for the equality of two Poisson means. Nelson [24] discusses this test in detail.

In order to use this test for detecting outlier tanks we need to test a slightly more general hypothesis. Namely consider the  $2 \ge 2$  table.



If  $\frac{X}{N_A} < .1$ ,  $\frac{Y}{N_B} < .1$  or if  $\frac{X}{N_A} > .9$ ,  $\frac{Y}{N_B} > .9$  then we're in the Poisson case.

Now  $\lambda_A = N_A P_A$ ,  $\lambda_B = N_B P_B$  implies that if  $P_A = P_B$  then  $\lambda_A / \lambda_B = N_A / N_B$ .

We thus wish to test the hypothesis

$$H_{o}: \frac{\lambda_{A}}{\lambda_{B}} = \rho \qquad \text{where } \rho = \frac{N_{A}}{N_{B}}$$

Nelson's test rejects  $H_{\alpha}$  at level  $\alpha$  if

$$\frac{Y}{X+1} \ge \frac{1}{\rho} F(2X + 2, 2Y; 1 - \alpha/2)$$

or if

$$\frac{Y+1}{X} \leq \frac{1}{\rho} F(2X, 2Y + 2; \alpha/2) \equiv \frac{1}{\rho} \frac{1}{F(2Y + 2, 2X; 1 - \alpha/2)}$$

where  $F(v_1, v_2; \gamma)$  represents the upper  $\gamma$  point of the F c.d.f. with  $v_1$ ,  $v_2$  degrees of freedom. If Y = 0, X>0 we can only carry out the one sided test. (We have just one sided information concerning  $\lambda_{\mathbf{p}}$ ).

$$H_{o}: \lambda_{A}/\lambda_{B} = \rho \qquad vs \qquad H_{1}: \lambda_{A}/\lambda_{B} > \rho$$

Nelson's test rejects H<sub>o</sub> at level  $\alpha$  if  $X \ge \rho F(2, 2X; 1 - \alpha)$ .

We now apply this Poisson test to the outlier detection problem. Consider groups 3, 4. These give rise to the two extreme points on the plot: -2.828, -2.049. Let us see if these should be regarded as outliers.

#### Replicate

In Group 3 we have

oup 3 we have	A	В	
Dead	0	2	2
Live	20	18	38
	20	20	40

Thus X = 0  $X \sim P_o(\lambda_A) \equiv P_o(20P_A)$ 

$$Y = 2$$
  $Y \sim P_{O}(\lambda_{B}) \equiv P_{O}(20p_{B})$ 

Since X = 0, we can carry out only a one sided test. Note that  $\rho = 1$ .

 $H_{O}: \lambda_{B} = \lambda_{A}$  $H_1: \lambda_B > \lambda_A$ 

We reject H<sub>0</sub> at level  $\alpha$  if Y > F(2, 2Y; 1 -  $\alpha$ ). In our example Y = 2. The critical value,  $F(2, 4; .\overline{95})$  is 6.94, which exceeds 2. There is thus no statistical evidence of differences among the two responses in group 3. Thus Tank 3 A is not an outlier.

We now carry out the test for group 4. The situation is less extreme than that in group 3, however we go through with the test for illustrative purposes.

## Replicate

In Group 4 we have

	<u>A</u>	<u> </u>	
Dead	0	1	1
Live	21	19	40
	21	20	41

Thus X = 0  $X \sim P_o(\lambda_A) \equiv P_o(21p_A)$ 

$$Y = 1$$
  $Y \sim P_o(\lambda_B) \equiv P_o(20p_B)$ 

Since X = 0, we can carry out only a one sided test.

$$H_{o}: \frac{\lambda_{B}}{\lambda_{A}} = \frac{N_{B}}{N_{A}} \equiv \frac{20}{21} = 0.952 \equiv \rho.$$
$$H_{1}: \frac{\lambda_{B}}{\lambda_{A}} > 0.952$$

From our previous discussion (interchanging the roles of X and Y)

reject H if Y 
$$\geq \rho$$
 F(2, 2Y; .95) =  $\frac{N_B}{N_A}$  F(2, 2Y; .95)  
= .952F(2, 2Y; .95)

In our case Y = 1, .952F(2, 2Y; .95) = (.952)(19.0) = 18,088

Thus we cannot reject H . There is no statistical evidence of differences in response rates among the tanks.

Thus tank 4A is not an outlier.

Holcombe and Phipps compound D

Embryo Mortality

$$I = 6, J = 4$$

(i.e. 6 groups, 4 tanks per group)

Embryo mortality

·. ·

Group 1	$\hat{p} = 0.35;$	all expected frequencies greater than 5.
Group 2	p̂ = 0.35;	all expected frequencies greater than 5.
Group 3	p̂ = 0.315;	all expected frequencies greater than 5.
Group 4	p̂ = 0.39;	all expected frequencies greater than 5.
Group 5	$\hat{p} = 0.335;$	all expected frequencies greater than 5.
Group 6	$\hat{p} = 0.38;$	all expected frequencies greater than 5.

We apply the transformation suggested in case 1. Data summaries are given below.

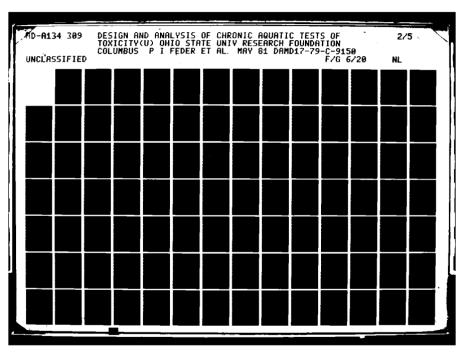
Calculate 
$$\left(1 - \frac{N_j}{N}\right)^{-1/2} \left(\frac{X_j - N_j \hat{p}}{\sqrt{N_j \hat{p} \hat{q}}}\right)$$
 for each tank within each group.

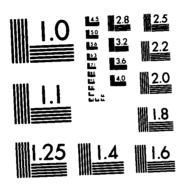
	x <sub>j</sub>	N j	Ŷ	N <sub>j</sub> p̂	√N <sub>j</sub> p̂q̂	Nj/N	$\left(1 - N_{j}/N\right)^{-1/2} \left(\frac{X_{j} - N_{j}\hat{p}}{\sqrt{N_{j}\hat{p}\hat{q}}}\right)$
Group 1							
Tank A Tank B Tank C Tank D	17 21 12 20	50 50 50 50	0.35 0.35 0.35 0.35	17.5 17.5 17.5 17.5	3.373 3.373 3.373 3.373	0.25 0.25 0.25 0.25	-0.171 1.198 -1.883 0.856
Group 2							
Tank A Tank B Tank C Tank D	19 14 16 21	50 50 50 50	0.35 0.35 0.35 0.35	17.5 17.5 17.5 17.5	3.373 3.373 3.373 3.373 3.373	0.25 0.25 0.25 0.25	0.514 -1.198 -0.514 1.198
Group 3							
Tank A Tank B Tank C Tank D	15 12 24 12	50 50 50 50	0.315 0.315 0.315 0.315	15.75 15.75 15.75 15.75	3.293 3.293 3.293 3.293 3.293	0.25 0.25 0.25 0.25	-0.263 -1.315 2.893 -1.315

	x <sub>j</sub>	N j	Ŷ	N <sub>j</sub> p̂	N <sub>j</sub> p̂q̂	√N <sub>j</sub> /N	$\frac{\left(1 - N_{j}/N\right)^{-1/2} \left(\frac{X_{j} - N_{j}\hat{p}}{\sqrt{N_{j}\hat{p}\hat{q}}}\right)}{\sqrt{N_{j}\hat{p}\hat{q}}}$
Group 4							
Tank A Tank B Tank C Tank D	20 21 23 14	50 50 50 50	0.39 0.39 0.39 0.39	19.5 19.5 19.5 19.5	3.449 3.449 3.449 3.449 3.449	0.25 0.25 0.25 0.25	0.167 0.502 1.172 -1.841
Group 5							
Tank A Tank B Tank C Tank D	18 19 14 16	50 50 50 50	0.335 0.335 0.335 0.335	16.75 16.75 16.75 16.75	3.337 3.337 3.337 3.337	0.25 0.25 0.25 0.25	0.433 0.779 -0.952 -0.260
<u>Group 6</u>							
Tank A Tank B Tank C Tank D	14 16 25 21	50 50 50 50	0.38 0.38 0.38 0.38	19.0 19.0 19.0 19.0	3.432 3.432 3.432 3.432	0.25 0.25 0.25 0.25	-1.682 -1.009 2.019 0.673

To prepare the normal probability plot we order the standardized values and plot the i-th smallest against the plotting position 100 x (1 - 0.5)/24 on the probability scale. These values are indicated below

i	1	2	3	4	5	6	7	8
Ordered Value	-1.883	-1.841	-1.682	-1.315	-1.315	-1.198	-1.009	-0.952
Plotting Position	<sup>3</sup> 2.1	6.2	10.4	14.6	18.7	22.9	27.1	31.2
i	9	10	11	12	13	14	15	16
Ordered Value	-0.514	-0.263	-0.260	0.171	0.167	0.433	0.502	0.514
Plotting Position		39.6	43.7	47.9	52.1	56.2	60.4	64.6





MICROCOPY RESOLUTION TEST CHART NATIONAL BUREAU OF STANDARDS-1963-A

<u> </u>	17	18	19	20	21	22	23	24
Ordered Value	0.673	0.779	0.856	1.172	1.198	1.198	2.019	2.893
Plotting Position	68.7	72.9	77.1	81.2	85.4	89.6	93.7	97.9

The normal probability plot of these points is shown in Figure X.3. Due to the within group correlation, the effective sample size is 6(J-1) = 6(4-1) = 18 independent observations. Is the Group 3, Tank C response an outlier? To determine the extent of statistical evidence that the apparent outlier did not occur just due to chance we calculate the probability that the maximum absolute value of 18 independent standard normal random variables does not exceed 2.893. Let  $Z_1, Z_2, \ldots, Z_{18}$  be 18 independent standard normal random variables.

$$P[\max_{j=1,...18} |Z_j| \ge 2.893] = 1 - P[a11 |Z_j| \le 2.893] = 1 - (.9962)^{18}$$
  
= 0.066.

This is of borderline statistical significance. We can thus infer that there is marginal statistical evidence that Group 3 contain an outlying tank.

Note that from the scatter plot of embryo mortality vs treatment in Figure VI.5. It is clear that there is no trend in the data and that the Tank 3C response does not stand out from those of the six groups as a whole. If we knew that there was no trend over groups we could construct a more powerful test by pooling all the tanks and calculating a common  $\hat{p}$ . However there is no point in doing this since we see from the scatter plot that Tank 3C is <u>not</u> out of line with respect to the pooled responses, but rather just with those in Group 3. The reason for this, if any, might be pursued.

Holcombe and Phipps compound D

$$I = 6, J = 4$$

(i.e. 6 groups, 4 tanks per group)

Fry Mortality

Group 1  $\hat{p} = 0.06$ Group 2  $\hat{p} = 0.08$ Group 3  $\hat{p} = 0.08$ Group 4  $\hat{p} = 0.13$ ; expected frequencies less than 5 in each tank. Group 5  $\hat{p} = 0.79$ ; all expected frequencies greater than 5. Group 6  $\hat{p} = 1.00$  ( $\hat{q} = 0$ ). Thus groups 1, 2, 3, 6 correspond to Poisson case 2. Group 4 corresponds

Thus groups 1, 2, 3, 6 correspond to Poisson case 2. Group 4 corresponds to case 3 (possibly to case 2). Group 5 corresponds to case 1.

Thus for groups 1, 2, 3,  
we calculate
$$(1 - N_j/N)^{-1/2} \quad 2 \left[ \sqrt{X_j} - \sqrt{N_j \hat{p}} \right].$$
For group 6 we calculate
$$(1 - N_j/N)^{-1/2} \quad 2 \left[ \sqrt{N_j - X_j} - \sqrt{N_j \hat{q}} \right].$$
For group 4 we calculate
$$(1 - N_j/N)^{-1/2} \quad 2\sqrt{N_j} \left[ \arcsin \sqrt{\hat{p}_j} - \arg \sqrt{N_j \hat{p}} \right].$$
For group 5 we calculate
$$(1 - N_j/N)^{-1/2} \left( \frac{X_j - N_j \hat{p}}{\sqrt{N_j \hat{p} \hat{q}}} \right)$$

	×,	N j	Ŷ	Njp	Nj/N	$\left(1 - N_{j}/N\right)^{-1/2} 2\left[\sqrt{X_{j}} - \sqrt{N_{j}\hat{p}}\right]$
Group 1						
Tank A	2	25	.06	1.5	.25	0.438
Tank B	2	25	.06	1.5	.25	0.438
Tank C	1	25	.06	1.5	.25	-0.519
Tank D	1	25	.06	1.5	.25	-0.519
<u>Group 2</u>						
Tank A	3	25	.08	2.0	.25	0.734
Tank B	1	25	.08	2.0	.25	-0.957
Tank C	2	25	.08	2.0	.25	0
Tank D	2	25	.08	2.0	.25	0
<u>Group 3</u>						
Tank A	2	25	.08	2.0	.25	0
Tank B	0	25	.08	2.0	.25	-3.266
Tank C	5	25	.08	2.0	.25	1.898
Tank D	1	25	.08	2.0	.25	-0.957

							(	(1 - N <sub>j</sub> /N	) <sup>-1/2</sup> 2√N	jarcsinvpj
	x,	и t	Ŷ	1	ĭj₽́	N,/N	<sup>ĝ</sup> j			arcsin√p̂
Group 4										
Tank A Tank B	4 5	25 25	.13 .13		3.25 3.25	.25	.16 .20		.493 1.094	
Tank C		25	.13			.25	.20		.493	
Tank D	0	25	.13		3.25	.25	0		-4.259	
	x <sub>j</sub>	N j	Ŷ	ľ	ı p	и,/N	∕N <sub>j</sub> pq	(1 - N <sub>j</sub>	/N) <sup>-1/2</sup> (	$\left(\frac{x_{j} - N_{j}\hat{p}}{\sqrt{N_{j}\hat{p}\hat{q}}}\right)$
Group 5								· · · · ·	<u> </u>	
Tank A	23	25	.79	19	.75	.25	2.037		1.842	
Tank B	21	25	.79		.75		2.037		0.709	
Tank C Tank D	16 19	25 25	.79 .79		).75 ).75		2.037 2.037		2.126 425	
	Nj-	-x <sub>j</sub>	N б	N á	l Nj	/N (1	- N <sub>j</sub> /N)	<sup>-1/2</sup> 2	Nj <sup>-X</sup> j - √	Ń <sub>j</sub> ĝ
Group 6										<u> </u>
Tank A	(	)	25 0		.2			0		
Tank B		2	25 0		.2			0		
Tank C Tank D		) )	25 C	-	.2			0 0		
values a	To pi nd pi	repar Lot t	e the he 1-t	norma :h sma	1 pro	babil tagai	nst the	ot we ord plottin	g positio	andardized on 100 x ated below.
		1	2		3		4	5	6	7
Ordered Value	-4.	259	-3.2	66	-2.12	.6 –	0.957	-0.957	-0.519	-0.519
Plotting Position	2.	1	6.2		10.4	1	4.6	18.7	22.9	27.1

76

	8	9	10	11	12	13	14	15	16	
Ordered Value	-0.425	0	0	0	0	0	0	0	0.43	8
Plotting Position	31.2	35.4	39.6	43.7	47.9	52.1	56.2	60.4	64.6	
	17	18		19	20	21	2	2	23	24
Ordered Value	0.438	0.493	0.	493	0.709	0.73	4 1.	094	1.842	1.898
Plotting Position	68.7	72.9	77.	1	81.2	85.4	89.	6	93.7	97.9

The normal probability plot appears in Figure X.4. It appears that the lowest 3 points are well below the N(0, 1) line. These three points correspond to Tanks 4D, 3B, 5C. Tanks 4D, 3B each have observed frequencies of 0. This is the region where the normal approximation is the poorest. Thus before we say that there are any outliers, we should compare the tanks within the treatment groups using a more appropriate exact test.

First let's look at tank 3B.

Group 3:

NUMBER OF

Replicate										
	A	<u> </u>	C	D						
Dead	2	0	5	1	8					
Live	23	25	20	24	92					
	25	25	25	25	100					

Tank 3 B is the suspected outlier. Let's compare its results to those in the other 3 tanks.

Rep	1	ic	a	t	e
-----	---	----	---	---	---

	В	A, C, D	
Dead	0	8	8
Live	25	67	92
	25	75	100

We can carry out an exact test by means of the Fisher - Irwin test. (See Lehmann [25], section 4.5, Lieberman and Owen [26]).

To carry out the Fisher - Irwin test we adopt the following notational identifications for the table.

	Grp 1	Grp 2	
Spec.	X		k
Ordin.			K – k
	n	N – n	N

Where

A STATES

k < N - k, n < N - n, k < n.

i.e. k is the smallest marginal entry

n is the smallest marginal entry in the other margin.

X is the cell entry in the cell corresponding to the (n, k) marginal categories.

In our example

_	B	A, C, D		
Dead	$\mathbf{x} = 0$	8	8 = k	("special")
Live	25	67	92	
	25	75	100 = N	
	" n			

Thus k = 8, n = 25, N = 100, X = x = 0. We enter the Lieberman and Owen hypergeometric distribution tables at these parameters. We obtain  $P(X \le 0) = 0.091$ . Thus a two sided probability is 0.182. This is quite marginal, at most.

Now group B was not chosen a priori, but rather as the most extreme of the 4 responses. Thus to get a feeling for how extreme this behavior is we carry out the following approximate calculation. P(most extreme of4 independent responses is more significant than 0.182 level) = 1 - P (all 4 responses less significant than 0.182 level) = 1 - (1 - 0.182)<sup>4</sup> = 0.55. We thus conclude that there is no statistical evidence that the response rate in Tank 3B differs significantly from the responses rates in the other tanks in that group. We conclude that the extreme behavior of the standardized value is due to the inapplicability of the normalizing square root transformation when  $X_i = 0$ . We now consider the responses in Group 4.

Group 4

	nop zzenee					
	A	B	C	D		
Dead	4	5	4	0	13	
Live	21	20	21	25	87	
ľ	25	25	25	25	100	

Replicate

Tank 4D is the suspected outlier. Let's compare its results to those in the other 3 tanks.

D	A, B, C	
0 = x	13	13 = k
25	62	87
25	75	100 = N
	0 = x 25 25	0 = x 13 25 62 25 75

Thus here k = 13, n = 25, N = 100, X = x = 0.

Entering the Lieberman and Owen tables we find that  $P(X \le 0) = 0.018$  this is a one sided probability Thus the observed two sided significance level = 2(0.018) = 0.036

Now tank D was not chosen a priori. Taking selection into account we have P (most extreme of 4 tanks more significant than 0.036 level) = 1 - P (all 4 tanks less significant than 0.036 level)  $= 1 - (1 - .036)^4 = 0.14.$ 

There is thus at most a marginal suggestion that the response rate in Tank 4D differs significantly from the response rates in the other tanks in that group. The very extreme appearance of the standardized value on the normal probability plot is again due to the inapplicability of the normalizing transformation when  $X_i = 0$ .

We now consider the responses in Group 5.

Group 5	Replicate					
-	A	<u> </u>	С	D		
Dead	23	21	16	19	79	
Live	2	4	9	6	21	
t	25	25	25	25	100	

Tank 5C is the suspected outlier. Let's compare its results to those in the other three tanks.

	C	A, B, D		_		
Dead	16	63	79			
Live	9 = x	12	21	=	k	np̂ = 5.25
	25	75	100	=	N	
	n n					

In this case the "special" category is "live". Thus k = 21, n = 25, N = 100, X = x = 9. Entering the Lieberman and Owen tables we find that

 $P(X \ge 9) = 1 - P(X \le 8) = 1 - .9638 = .0362.$ 

South and and and and an and

Thus the two sided significance probability is 2(.0362) = .0724. This is at best marginal. Now tank D was not chosen a priori. Taking selection into account, we have

P (most extreme of 4 tanks most significant than .0724 level) =  $1 - (1 - .0724)^4 = .260$ 

Thus tank D is not significantly different than the others. Since the expected frequencies are fairly large in this example we can also carry out an asymptotic test.

_	C	A, B, D	
Dead	16 (19.65)	63 (59.25)	7 <del>9</del>
Live 	9 (5.25)	12 (15.75)	21
	25	75	100

Expected frequencies are in parentheses.

$$\chi^{2} = \sum \frac{(0 - E)^{2}}{E} = \frac{(16 - 19.75)^{2}}{19.75} + \frac{(63 - 59.25)^{2}}{59.25} + \frac{(9 - 5.25)^{2}}{5.25} + \frac{(12 - 15.75)^{2}}{15.75}^{2}$$
  
Thus  $\chi^{2} = 4.53$ 

Under the hypothesis of homogeneity,  $\chi^2$  is distributed as  $\chi^2_1$ .

Thus  $\chi^2$  is significant at the 0.033 level, in close agreement with the results based on the Fisher - Irwin test, namely lack of statistical evidence of differences when selection is accounted for.

Since the normalizing transformation is appropriate for the range of responses in Group 5, we can also regard the normalized value from Tank C, -2.126, as the minimum of 6 x (4 - 1) - 2 = 16 standard normal deviates (we disregard those responses corresponding to Tanks 3B, 4D). The probability that a standard normal deviate is less than -2.126 is 0.017. The probability that the minimum of 22 independent standard normal deviates is less than -2.126 is thus  $1 - (1 - 0.017)^{16} = 0.25$ . Again there is no statistical evidence that this value is an outlier.

Jarvinen compound B

I = 6, J = 2 (i.e. 6 groups, 2 tank per group)

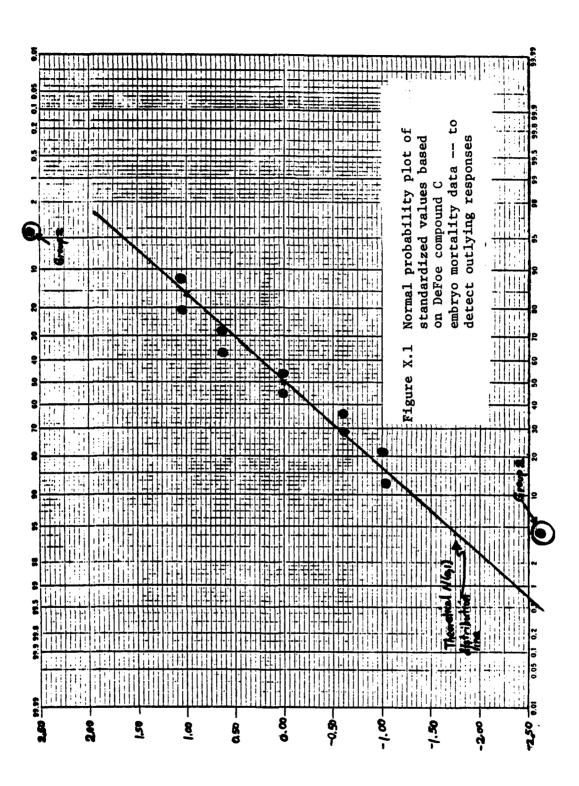
Embryo Mortality

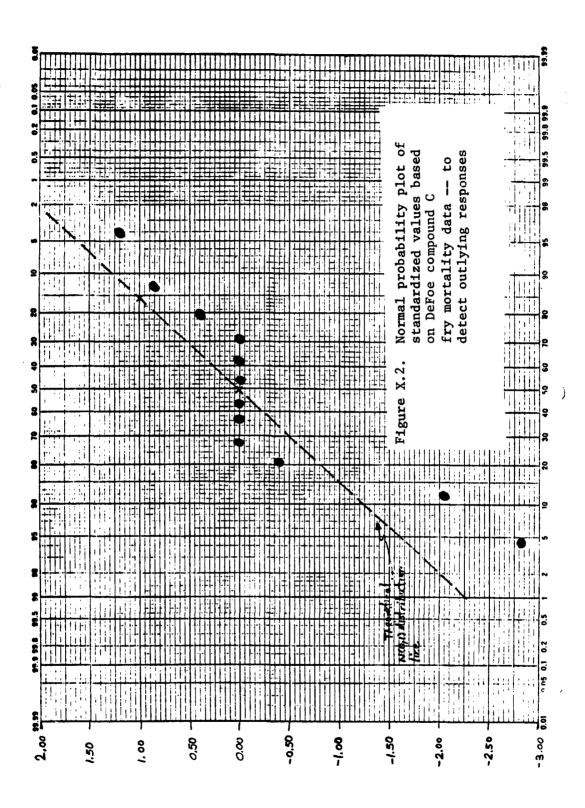
The normal probability plot of standardized values (based on the case 1 transformation) is shown in Figure X.5. The plot appears perfectly symmetric about 0 since J = 2. The effective sample size is 6(J - 1) = 6 independent observations. The reference line in the plot is the standard normal distribution function.

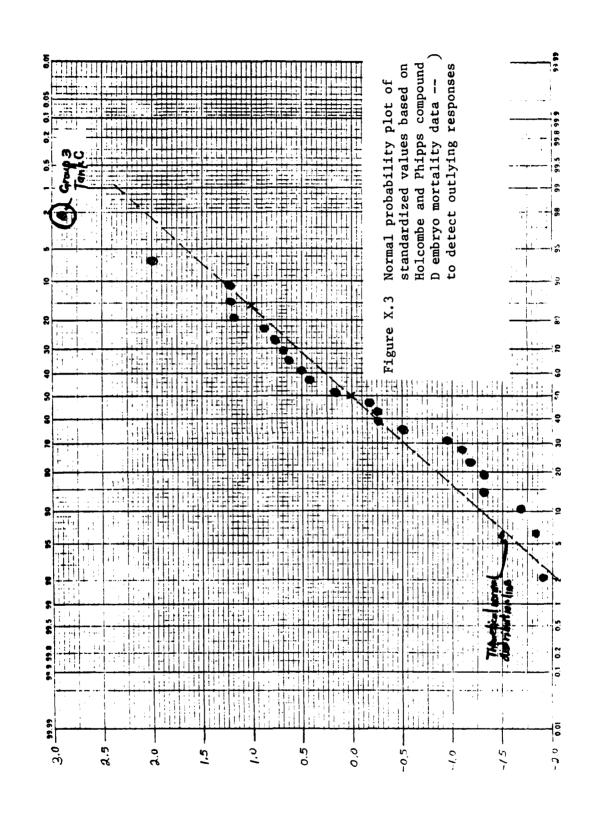
The behavior of the plot suggests the presence of extrabinomial variation (i.e. random tank to tank variation within groups) in the data rather than outliers. This is seen by the fact that the points lie along a curve with steeper slope than that of the standard normal distribution function. The extreme points are <u>not</u> outliers since they lie directly on the curve determined by the other values. Thus the conjecture made in section VIII concerning the presence of outliers is <u>not</u> horne out. Note that this behavior is directly opposite to that observed in Figure X.1.

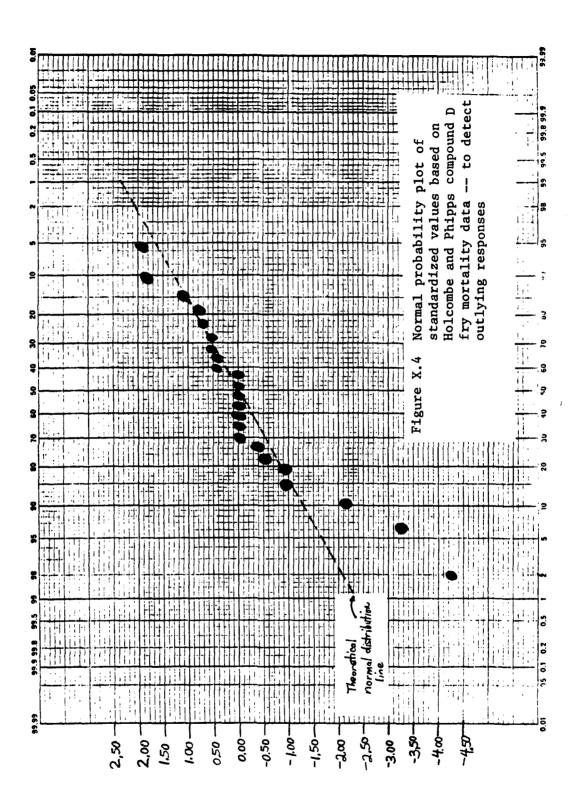
If we draw a line through the plotted points in Figure X.5 the estimated standard deviation (corresponding to the difference between the 84th and 50th percentiles) is about 1.7. Thus the estimated variance is  $(1.7)^2 = 2.9$ . This values is very close to the factor K = 3.07 that we utilized in section IX to adjust these data.

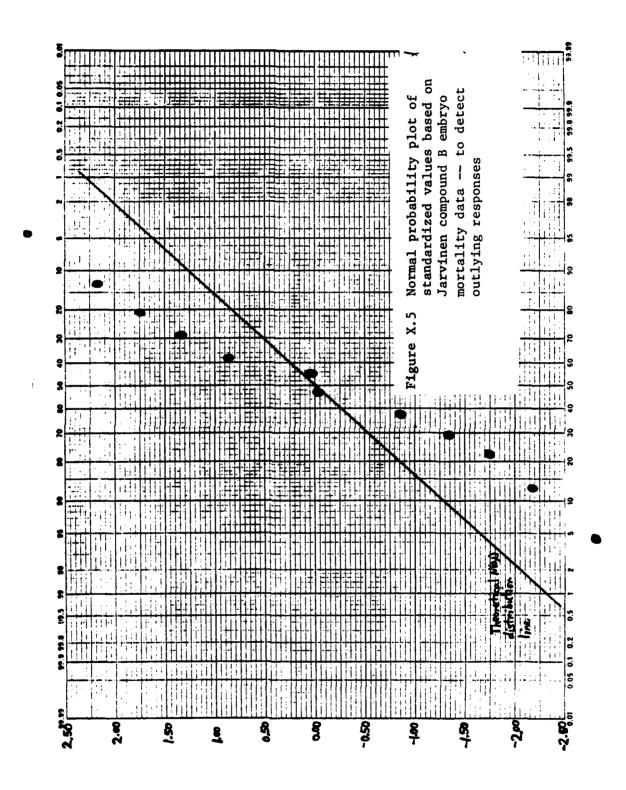
الارد والمسادية











F

## XI. TESTING FOR CONCENTRATION RELATED EFFECTS

## A. Background

After we have completed preliminary graphical displays, tests for tank to tank heterogeneity, and outlier detection procedures we are ready to proceed to the main portion of the data analysis. This involves comparing responses across treatment groups to arrive at an inference about what constitutes an "acceptable" concentration. If no tank to tank heterogeneity is evident in the data then the original data may be pooled across tanks within groups and subsequent analyses carried out on a per fish basis or alternatively the data can first be adjusted to reflect the increased variation and the adjusted "data" can be pooled across tanks and analyzed on a per fish basis. As remarked earlier, we prefer the latter approach.

Before considering statistical procedures to determine acceptable concentrations, we must first define what is considered to be acceptable. According to the guidelines for early life stage tests, [8], "...A lower chronic endpoint is the highest tested concentration...which did not cause the occurrence (which was statistically significantly different from the control at the 95% level) of any specified adverse effect and below which no tested concentration caused such an occurrence...An upper chronic endpoint is the lowest tested concentration...which caused the occurrence (which was statistically significantly different from the control at the 95% level) of any specified adverse effect and above which all tested concentrations caused such an occurence". We are thus interested in determining which concentrations yield (statistically) significantly different results than the control group. In a later section we will present an alternative notion of acceptable concentration.

<u>Opinion</u>: For the purpose of testing hypotheses concerning heterogeneity of response rates across groups or of constructing confidence intervals to compare treatment group and control group response rates, unless there is relatively strong statistical evidence of heterogeneity among tanks within groups (e.g. observed significance level less than 0.05 or 0.10) then act as if there is not heterogeneity of response from tank to tank within groups. Base subsequent tests and confidence intervals on the original (i.e. unadjusted) data, pooled across tanks within groups.

This suggestion reflects a <u>conservative</u> viewpoint with respect to the conclusions drawn from such subsequent analyses. Namely, suppose that such tank to tank variation within treatment groups exists but we do not detect it. Then we proceed as if none exists. Thus the "true" variability of the test statistics that we use will exceed the assumed variability. Thus if we carry out a test at nominal level  $\alpha = 0.05$ , say, the "true" alpha level will in fact be greater than the nominal.

Inflating the actual  $\alpha$  - level over the nominal level makes the test more prone to reject the hypothesis of equality of treatment and control groups than an actual  $\alpha$  - level test. Thus if we err, it will be on the side of declaring a treatment group significantly different from the control group when it in fact is not. This is conservative.

Carrying out analyses on a per tank basis drastically reduces the degrees of freedom available for estimating variability of the test statistics, especially when there are very few tanks per treatment group. This diminishes the power of subsequent tests to compare treatment groups with control group, thus causing acceptance of the null hypothesis more often than necessary when it is false. That is, suppose there is little or no tank to tank heterogeneity within groups but we still carry out our analyses on a per tank basis "to be safe". Then the reduced power of the procedures based on per tank analyses will cause us to miss moderate differences between treatment groups and control groups that we might be able to detect if we were analyzing on a per fish basis. Thus we would be erring on the side of failing to detect departures from the control group response rates. This is unconservative.

It should be noted that exactly the <u>opposite</u> is true when we fit dose response curves to the data. Namely underestimation of variability results in an <u>increased lower bound on safe concent</u>-ration, which is <u>unconservative</u>. Thus for the purpose of fitting dose response curves perhaps we should adjust for tank to tank heterogeneity if the overall significance level is as great as  $\alpha = .20$  or perhaps even  $\alpha = .50$ .

## B. Chi Square Test of Homogeneity Across Treatment Groups

The most commonly used overall test for differences in mortality or abnormality rates across treatment groups is the chi square test for homogeneity. It is analagous to the shotgun analysis of variance F test for quantitative data. Since the preliminary tank to tank heterogeneity tests on the DeFoe and on the Holcombe and Phipps data sets were at most marginal, we use the original data. However the same test could be applied after adjusting the data.

Let  $p_1, p_2, \ldots, p_I$  denote the response probabilities in the I (treatment and control) groups. We wish to test the hypothesis

$$H_0: p_1 = p_2 = \dots = p_1.$$

The chi square test is an <u>overall</u> test of this hypothesis (i.e. a shotgun test). It is based on the following statistic:

Let  $(X_{ij}, N_{ij})$  denote the number of responses and the number of fish, respectively, in the j-th tank within the i-th treatment group

$$i = 1, ..., I; j = 1, ..., J.$$

Let

$$N_{i+} = \sum_{j=1}^{J} N_{ij}, \ X_{i+} = \sum_{j=1}^{J} X_{ij}, \ N_{++} = \sum_{i=1, j=1}^{I} \sum_{j=1}^{J} N_{ij}, \ X_{++} = \sum_{j=1, j=1}^{I} \sum_{j=1, j=1}^{J} X_{ij}.$$
 Then

 $p_{++} \equiv \frac{N_{++}}{N_{++}}$  is the estimate of the common value of p under H<sub>o</sub>.

$$\chi^{2} \equiv \sum_{i=1}^{I} \frac{\left(X_{i+} - N_{i+} \hat{p}_{++}\right)^{2}}{N_{i+} \hat{p}_{++} (1 - \hat{p}_{++})}$$

is the  $\chi^2$  test of  ${\rm H}_0$ . Under  ${\rm H}_0,~\chi^2\sim\chi^2_{\rm I-1}$ . Since the chi square test is sensitive to all kinds of departures from  ${\rm H}_0$ , it is not tailor made to be sensitive to ordered alternatives, of the type most commonly encountered in toxicology. We will discuss this further later.

The chi square test is easy to carry out from a computational standpoint since many standard programs are available. For example the procedure PROC FREQ in the SAS statistical computing system [12] can be used to carry out this test. The program BMDP1F in the BMDP statistical computing system [9] can also be used for this purpose. Figure XI.1 illustrates output from SAS PROC FREQ to test for homogeneity of response rates across groups for the fry mortality data from the Holcombe and Phipps test on compound D. This test is based on data pooled across tanks within groups. Of course, the test rejects  $H_0$  very strongly, as it should based on the appearance of the preliminary scatterplot.

We have also incorporated this test into our EXAX2 computer program [14]. We pool responses across tanks within groups and compute the chi square test. If expected frequencies within each cell exceed the cutoff, we evaluate the significance of chi square based on its asymptotic distribution under  $H_{o}$ .

If any expected frequency is less than the cutoff we evaluate the exact <u>small sample distribution</u> of chi square, conditional on the margins, by enumeration as discussed in the section on the exact chi square program.

We illustrate this feature of EXAX2 with the DeFoe compound C and with the Holcombe and Phipps compound D data. We tested for heterogeneity of response rates across groups for the fry mortality and the embryo mortality data. The results of these tests are shown in Figures XI.2 - XI.5.

Figure XI.2 illustrates the chi square test of homogeneity across treatment groups for the embryo mortality data in the DeFoe compound C experiment. It will be recalled that no tank to tank heterogeneity within treatment groups was found by the preliminary test, and so the data have been pooled across the tanks within treatment groups. The first matrix displays the observed 2 x 6 table. The second matrix displays the expected (under  $H_0$ ) frequencies. Since each expected frequency exceeds 5 (by a great deal in this example) the test is based on asymptotic theory.

NOTE THAT SINCE THIS IS A PRELIMINARY TEST WE SHOULD BE VERY LIBERAL IN DECIDING WHEN TO REJECT H<sub>o</sub> AND TO GO ON TO MORE DE-TAILED COMPARISONS. THUS A LARGE  $\alpha$ -VALUE e.g.  $\alpha$  = .20 or  $\alpha$  = .25 SHOULD BE USED. THIS ENHANCES THE SENSITIVITY OF THE TEST TO DETECT MODERATE DEPARTURES FROM H<sub>o</sub>.

We see from the bottom of Figure XI.2 that the observed significance level is  $\alpha = 0.31$ . Thus even by our liberal yardstick we see no statistical evidence of group to group differences in embryo mortality in the DeFoe data. This agrees with the appearance of the preliminary scatter plots.

The same test was carried out for the DeFoe compound C fry mortality data. The results are given in Figure XI.3. Again, all the expected frequencies exceed 5.0 and so the asymptotic theory is used. This time the chi square statistic is highly significant. Chi square = 182.79 with 5 d.f. Thus there is strong statistical evidence of group to group response rate differences in fry mortality. This again agrees with the appearance of the preliminary scatter plot. Figures XI.4, XI.5 contain the results of the chi square tests of homogeneity across groups for the Holcombe and Phipps compound D embryo mortality and fry mortality data respectively. Again the data have been pooled across tanks within groups.

In both cases the expected frequencies exceed 5 and so asymptotic theory is used. We see for the embryo mortality data the test is nonsignificant even at the liberalized  $\alpha = .20$ . For the fry mortality data the test is very highly significant (chi square = 389.68 with 5 d.f.). Thus again there is no statistical evidence of group to group differences with respect to embryo mortality while there is strong statistical evidence of group to group differences with respect to fry mortality. This is in good agreement with the appearances of the preliminary scatterplots.

Figure XI.6, XI.7 contain the results of the chi square tests of homogeneity across groups for the Jarvinen compound B embryo mortality and fry mortality data respectively. The data have been pooled across tanks within groups. In both cases all the expected frequencies exceed 5 and so asymptotic theory is used.

For the fry mortality data the test is very highly significant, as was the case with the fry mortality data from the other experiments considered. It is quite clear that the last two treatment groups have substantially higher response rates than the first two groups.

In contrast to the cases for the DeFoe and Holcombe and Phipps data sets, there is some statistical evidence of group to group differences in embryo mortality rates.

We also saw strong statistical evidence in Sections VII and X of heterogeneity in response rates among tanks within groups. In Section IX we calculated an adjustment factor of  $\hat{\vec{K}}$  = 3.071 for these responses, to account for the tank to tank heterogeneity. The effect of this adjustment on chi square is to adjust it downward by the factor  $\hat{\vec{K}}$ . Thus with respected to the adjusted."responses", the observed chi square value becomes 10.71426/ $\hat{\vec{K}}$  = 10.71426/3.071 = 3.489. The probability that a chi square random variate with 5 d.f. exceeds 3.489 is 0.625. Thus the tank to tank heterogeneity within groups accounts for the significant chi square across groups. Thus again there is no statistical evidence of variation in embryo mortality rate across groups.

## C. One Sided Tests of Homogeneity Across Treatment Groups

The shotgun chi square test, although the most commonly used test of homogeneity of response rates, is not the most appropriate test for application to toxicity data. The chi square test is an overall test which is not designed to be particularly sensitive to the <u>one sided</u>, <u>monotone alternatives</u> characteristic of dose response tests. More specialized tests have been designed to be more powerful against alternatives of this type.

Several tests of response rate homogeneity against ordered alternatives are discussed in the literature. Snedecor and Cochran [28] section 9.11 and Steel and Torrie [29], section 22.10 extract a single degree of freedom from the overall chi square test to test for linear regression in 2 x K tables where the columns fall in a natural order Scores,  $Z_j$ , are assigned (arbitrarily) to the columns to treat them as values on a continuous scale of measurement. The weighted linear regression coefficient of mortality probability on score  $Z_j$  is calculated and tested for significance. The major drawback of this method is the arbitrariness of the scores. See [28, 29] for details.

An alternative approach to the construction of one sided tests is by means of measures of association for ordered contingency tables. Such measures can be thought of as analogs for qualitative responses to correlation coefficients for quantitative responses. Goodman and Kruskal [ 30, 31 ] have derived and reported on a number of measures. Several commonly used measures of association are Kendall's tau b, Stuart's tau c, Goodman and Kruskal's gamma, just to name a few. For a given table each of these measures yields different numerical values and so it is not clear how to ascribe physical meaning to these values. However for each of the measures a value of zero means no monotone association between categories and positive or negative values mean positive or negative associations respectively. Thus a test of homogeneity of treatment group response rates that is sensitive to monotone, one sided alternatives can be constructed by testing the null hypothesis that these measures are zero against a one sided alternative. Brown and Benedetti [ 32 ] have calculated improved standard error estimates for the various measures that are appropriate studentizing factors to test the null hypothesis that these measures are zero. They show empirically that their new standard errors provide better approximations in small and moderate samples than do the older standard error estimates reported by Goodman and Kruskal [31]. Furthermore they show that a number of measures, each having different numerical values, result in identical "t ratios" when normalized by their respective Brown and Benedetti standard error estimates. This is desirable because we need consider just one "t ratio" rather than five. Proctor [33] shows that tests based on measures of association are in fact much more powerful against one sided, monotone alternatives than is the shotgun chi square test, as would be intuitively expected.

Agresti and Wackerly [15] also discuss one sided tests of homogeneity based on measures of association. They discuss Kendall's  $\gamma_b$  in some detail. They illustrate an instance of the increased sensitivity of such measure of association tests for detecting ordered departures from homogeneity with the following artificial example. Consider the 3 by 3 contingency table with ordered categories:

A	High	Medium	Low	
High Medium	6	4	2	
Medium	4	4	4	
Low	2	4	6	

There is clearly a positive trend in the table however the Fisher-Irwin test (exact) shows significance level  $\alpha = .514$ . This test would thus miss the trend. However Kendall's  $\gamma_b$  test (exact) is significant at  $\alpha = .053$  and so would detect the trend.

Agresti and Wackerly also comment that the asymptotic normal approximation to the distribution of the sample estimate of  $\gamma_{\rm b}$ may be quite poor for small sample sizes. They report that the observed significance level can be substantially greater than the nominal for small sample sizes; i.e. we reject H<sub>o</sub> when it is correct far more than the nominal proportion of times. This is at least in part due to the maximum likelihood estimate of standard error of  $\gamma_{\rm b}$  having a negative bias. Based on Agresti and Wickerly's example, <u>asymptotic normal distribution theory would be suspect</u> at least for N below 50. Agresti and Wackerly suggest that an alternative, exact conditional test against ordered alternatives can be used, based on measures of association and enumeration of tables, when the sample sizes are too small to apply asymptotic normal theory.

The applicability of asymptotic distribution theory for the sample sizes and magnitudes of response proportions encountered in fish toxicity tests is a matter for detailed future study, probably by simulation. This is too involved for us to consider here. However as we use this test only on pooled data (original or adjusted) across tanks within groups, the sample sizes would be expected to be reasonably large (N in excess of 200) and so we utilize asymptotic theory for the remainder of this section.

It should be noted that ordinal measures of association assume that as one variable (e.g. concentration) increases the other variable (e.g. percent mortality or percent abnormality) either increases monotonically or decreases monotonically. Nonmonotone relations (e.g. first increasing and then decreasing) can well result in small or even zero values of the measures. This is analogous to properties of correlation coefficients.

Goodman and Kruskal [ 30 ] define and discuss the properties of a number of measures of association for cross classifications. They include a section on measures for ordered categories (i.e. ordinal data). They propose a measure,  $\gamma$ , which is defined as follows:

Suppose two individuals are drawn at random from a population described jointly by two discrete, ordered categories.

Category 1: i = 1, ..., I Category 2: j = 1, ..., J

In our fish toxicity examples I = 2 (e.g. live, dead), J = C (number of groups)

Let (i, j), (i', j') denote the (random) indices of these two individuals within the two categories. If there is an ordered correspondence between categories, we should see the same (or opposite) orderings of each of the categories, depending on direction of association.

Let 
$$\Pi_s \equiv P[(i>i' \text{ and } j>j') \text{ or } (i  
 $\Pi_d \equiv P[(i>i' \text{ and } j  
 $\Pi_t \equiv P[i = i' \text{ or } j = j']. \equiv P\{\text{tie}\}$$$$

To avoid ambiguity they condition on the absence of ties. The conditional probability of like orders given no ties is  $\Pi_{s}/(1 - \Pi_{t})$ . The conditional probability of unlike orders given no ties is  $\Pi_{d}/(1 - \Pi_{t})$ . The difference of these two probabilities is defined as  $\gamma$ . Namely

Goodman and Kruskal's Gamma

$$\gamma = \frac{\Pi_s - \Pi_d}{1 - \Pi_t} \equiv \frac{\Pi_s - \Pi_d}{\Pi_s + \Pi_t}$$

In the situation when the two categories are independent  $\Pi_s = \Pi_d$  and so  $\gamma = 0$ . However the converse is not necessarily true (except in the 2 x 2 case).

The Kendall's  $\gamma_b$  and Stuart's  $\gamma_c$  measures are related to  $\gamma$ . Let m = min(I, J). Then

$$\gamma_{c} \equiv \frac{\prod_{s} - \prod_{d}}{(m - 1)/m}$$

This modification is made so that  $\gamma_c$  can nearly attain the absolute value 1 for nonsquare tables when the entire population lies on a longest diagonal of the table.

Kendall's  $\gamma_b$  is also related to  $\gamma$ . Namely, let  $p_{ij}$  denote a cell frequency,

$$p_{i} \equiv \sum_{j=1}^{J} p_{ij}, p_{j} \equiv \sum_{i=1}^{I} p_{ij}.$$
 Then

$$\gamma \equiv \frac{\prod_{s} - \prod_{d}}{1 - \prod_{t}} = \frac{\prod_{s} - \prod_{d}}{\prod_{s} - \sum_{t} \prod_{d}}$$
$$\frac{\prod_{s} - \prod_{d}}{1 - \sum_{s} p_{s}^{2} - \sum_{s} p_{s}^{2} + \sum_{s} \sum_{t} p_{s}^{2}}$$
$$\frac{\prod_{s} - \prod_{t} \prod_{s} \prod_{t} \prod_{s} \prod_{t} \prod_{s} \prod_{t} \prod_{s} \prod_{s} \prod_{t} \prod_{s} \prod_$$

$$\gamma_{\mathbf{b}} \equiv \frac{\Pi_{\mathbf{s}} - \Pi_{\mathbf{d}}}{\sqrt{\left[1 - \sum_{i=j}^{\Sigma} \mathbf{p}_{i}^{2}\right] \left[1 - \sum_{j=1}^{\Sigma} \mathbf{p}_{j}^{2}\right]}}$$

 $\gamma_b$  corrects for pairs of observations tied with respect to at least one of the categorizations and ranges between -1 and +1.

Somer's d (asymmetric measure)

$$\mathbf{d}_{\mathbf{R}|\mathbf{C}} = \frac{\Pi_{\mathbf{s}} - \Pi_{\mathbf{d}}}{1 - \Pi_{\mathbf{t}} + Y_{\mathbf{o}}}$$

$$d_{C|R} = \frac{\prod_{s} - \prod_{d}}{1 - \prod_{t} + x_{o}}$$

Where  $Y_{o} \equiv$  probability of tie in row only

 $X_{a} \equiv$  probability of tie in column only

For all of these measures a zero value indicates a complete lack of a monotone relationship between the two variables (no association). A value of +1 indicate a perfect monotone increasing relationship (perfect agreement) and a - 1 indicates a perfect monotone decreasing relationship (perfect disagreement). It should be noted that lack of a monotone relationship is not the same as statistical independence. These measures will equal zero when there is dependence of a complicated form. However, when the variables are independent, the measures will equal zero. Kendall's  ${}_{b}^{\circ}$  differs from the others in that it can reach a value of <u>+</u> 1 only for square tables, otherwise its maximum is lower.

Stuart's  $\gamma_c$  is an adjustment of  $\gamma_b$  that can attain value  $\pm 1$  for non square tables.

There is much discussion in the literature about which measures most realistically portray strengths of monotone relations. In general it is difficult to interpret the magnitudes of these measures in any physically meaningful fashion. However we will be using these measures of association only for tests of significance to detect departures from 0. For this application the situation simplifies considerably since Brown and Benedetti, page 311, show that basing a test of significance on  $\gamma_c$ ,  $\gamma$ ,  $\gamma_b$ ,  $d_{C|R}$ ,  $d_{R|C}$  all lead to exactly the same test statistic, sample for sample. Thus we do not need to be concerned with differences among the values of the measures. That is, Brown and Benedetti, [125] have derived new estimates ASE, of the asymptotic standard errors of the measures of association that are better than those given previously in the literature for testing the null hypothesis that the measure is zero. They report one set of standard errors to use for testing purposes and another set of standard errors to use for constructing confidence intervals. They show that

$$\frac{\hat{\gamma}}{ASE_{o}(\gamma)} \equiv \frac{\hat{\gamma}_{b}}{ASE_{o}(\hat{\gamma}_{b})} \equiv \frac{\hat{\gamma}_{c}}{ASE_{o}(\hat{\gamma}_{c})} \equiv \frac{\hat{d}_{R/C}}{ASE_{o}(\hat{d}_{R/C})} \equiv \frac{\hat{d}_{C/R}}{ASE_{o}(\hat{d}_{C/R})} \equiv T$$

<del>9</del>7

which means that the five measures all give the same test of the null hypothesis of no (monotone) association.

Brown and Benedetti report a simulation study of the use of these T ratios to test the null hypothesis. They compared the T-values to the percentage points of the standard normal distribution. They concluded that

- 1. The ASE 's give empirical type I error rates closer to the nominal significance level and more consistent for different patterns of non-association than do previously reported standard error estimates.
- 2. For N  $\geq$  100 the T value can safely be compared to the percentile points of the standard normal distribution.
- 3. For N  $\leq$  50 the distribution of the T value seems to have heavier than normal tails, and they recommend comparing it to Student's t with approximate degrees of freedom (ADF) = 0.4N.

Their T - ratios for testing the null hypothesis of nonassociation (i.e. monotone) and (asymptotic) standard errors appropriate for constructing confidence intervals on the measures of association are implemented in the BMDP [27] program BMDP1F, measures of association for two way frequency tables. (Note that BMDP1F was extensively rewritten and reissued in August, 1976. Thus only versions of this program dated after August, 1976 are based on the most up to date theory). We will illustrate the use of this program in this section, with both artificial and real data.

Proctor [33] has discussed the relative efficiencies of tests of association for ordered two way contingency tables and has compared these efficiencies with that of the chi square test. He reports that in most cases of ordered alternatives, the efficiencies of tests based on the measures of association are much greater than that of tests based on chi square. For one example of a 6 x 6 ordered contingency table constructed from an underlying **bivariate** normal distribution with correlation  $\rho = 0.80$ , the efficiencies of the tests of association based on measures of association relative to the chi square test were about 3.4. This means that for the chi square test to attain the same power against this alternative as a test based on the measures of association, it would need to be based on more than three times as many observations. In efficiency calculations based on other assumptions about the alternatives, the chi square procedure was consistently generally very much less efficient than test procedures based on measures of association.

To get some further feeling for the sensitivities to ordered alternatives of tests based on measures of association as compared with the chi square test we constructed several artificial sets of data having varying degrees of monotone trend in response probabilities. We tested the null hypothesis of no association between mortality level and treatment group using the one sided tests based on ordinal measures of association and the shot gun chi square test. Both of these tests can be carried out using the BMDP program [27] BMDP1F (versions subsequent to August, 1976). We should note that in these one sided tests we are looking for <u>counterassociation</u>. That is, the probability of being alive decreases as concentration group increases. We are thus testing for departures from 0 in a <u>negative</u> direction.

Figure XI.8 contains instructions for using program BMDP1F. Figures XI.9, 10 and 11 illustrate one sided and chi square tests on tables (based on artificial data) that exhibit linear trends of response probability with concentration group, but with differing slopes. They represent mild, moderate, and strong trends. In each case the one sided test based on measures of association reflects a <u>much</u> stronger association between categories than does the chi square test.

In conclusion, we see that the various ordinal measures of association provide equivalent tests of the null hypothesis of no association between mortality rate and concentration group. Further more all the tests are much more sensitive than the chi square test to alternatives of a monotonic nature.

Appendix A5, pp 778-792, of the 1979 BMDP manual [27] and Brown and Benedetti [32] are helpful in interpreting the output from BMDP1F. Brown and Benedetti calculate two asymptotic standard errors for each measure,  $ASE_0$ ,  $ASE_1$ .  $ASE_1$  is derived assuming the alternative hypothesis is true; i.e. the measure is not zero. It is obtained by the method of Goodman and Kruskal [31] and is appropriate for setting confidence limits on the measures for large samples. Brown and Benedetti discuss the use of  $ASE_1$  in computing confidence limits and power in an unpublished technical report that is available from the Health Sciences Computing Facility at UCLA.

 $\rm ASE_O$  is computed under the null hypothesis that the measure is zero. It was derived by Brown and Benedetti in the 1977 paper cited above.

The T-value for each measure is the ratio of each measure to its  $ASE_0$ . Brown and Benedetti report, based on simulation studies, that the use of  $ASE_0$  in the denominator of the T-statistic rather than  $ASE_1$  or other suggestions made in the literature gives superior results in that the attained type I error rates are closer to nominal and are more consistent for differing patterns of probabilities.

To illustrate the use of these one sided tests of association vs ordered alternatives as compared with the chi square test of significance, we ran the BMDP program, BMDP1F, on mortality data data from the DeFoe compound C test and from the Holcombe and Phipps compound D tests. The results are shown in Figures XI.12-15, both for embryo mortality data and for fry mortality data. Qualitatively there is no difference, in these data, in the conclusions arrived at by each procedure. The relationships between concentration and percent response are so strong in the fry mortality data that the observed significance levels are 0 to many decimal places. For the embryo mortality data, neither procedure reveals a statistically significant relation between concentration and percent mortality. Since the observed mortality in the DeFoe embryo data is smaller at the higher concentrations than at the lower, the measures of association are positive and the observed significance level is higher for the one sided test than for the chi square test. We almost have a significant trend in the wrong direction! Why? It may be due to the outlier tank in group 2.

In conclusion it should be remarked that any overall test for concentration related effects is just a screening device. It merely states whether there is any statistical evidence of concentration related effects but does not provide any indication of which treatment groups have responses that differ from the control group. That is the role of multiple comparisons. The overall test is intended to screen out those data sets for which multiple comparisons would be a futile exercise because no differences exist. In this regard it should be noted that since the overall test is just a preliminary procedure, it makes sense to use a very liberal  $\alpha$ -level, like  $\alpha$  = .20 or perhaps even  $\alpha$  = .50. This improves the sensitivity of the test to detect marginal effects, but at the expense of an increased false rejection rate. However such false rejections of the null hypothesis will be detected later in the multiple comparison phase.

CTGL	TRT						
FREQUENCY! EXPECTED 1 CELL CHI21		FRY	MOR	TALIT	Y		
COL PCT	1	2	3	4	5	16 1	TOTAL
DEAD	6   35.7   24.7   6.00	8   35.7   21.5   8.00	8   35.7   21.5   8.00	13   35.7   14.4   13.00	79 35.7 52.6 79.00	100   35.7   116.0   100.00	214
	94   64.3   13.7   94.00	92   64-3   11-9   92-00	92   64.3   11.9   92.00	87 1 64-3 1 8-0 1 87-00 1	21 64.3 29.2 21.00	0   64.3   64.3   0.00	356
TUTAL	100	100	100	100	100	100	600
		STATI	STICS FOR	2 <del>-1</del> 84 TA	BLES		

ŀ. .

PHI	0.806			
CONTINGENCY CUEFFICIENT	0.627			
CRAMER*S V	0.806			
LIKELIHOOD RATIO CHISQUARE	444.801	DF =	5	PR08=0.0001

389.676

DF=

5 PR08=0.0001

CHI-SQUARE

Figure XI.1 SAS PROC FREQ output from chi square test for homogeneity across groups applied to fry mortality data from Holcombe and Phipps test on Compound D

I     CUNTIAGENCY.IA3LF     Prove the second of the		1 NG FNCY. 1 NG FNCY. 200 3 -7.0C 5 -7.0C			****	59	****
ALIVE $1, 141 + 1, 53, 500$ $1, 24, 500$		1 NG ENCY.					
1 $A_{11}^{11}V_{11}^{11}$ $P_{12}^{12}$ <	4 	377 200 200 200 377.000 5870U_NCI					
AI     AL (V5     3 37,000     3 7,000     100,000       AI     AL (V5     3 37,000     100,000     3 37,000       AI     AL (V5     3 37,000     100,000     3 37,000       AI     AL (V5     3 37,000     101,00,000     3 43,000       AI     AL (V5     3 37,000     100,000     3 43,000       AL (V5     3 50,000     100,000     100,000     3 43,000       AL (V5     3 50,000     100,000     100,000     3 43,000       AL (V5     3 57,000     100,000     3 43,000     3 43,000	а мал. с. с. –	377.000 300 377.000 5870UTNCI					
Al. AL IVE = $3\sqrt{7}$ . $900$ . THEAL DEAL = $743$ . $900$ . C. AND LUEE = $700$ EXPECTED FREQUENCIES X5000 = $7340$ = $743$ . $9000$ . CUTOFF = $6.0$ AL ALIVE = $367$ . $900$ . $900$ . $10000$ AL ALIVE = $367$ . $900$ . $1014$ . $960$ = $743$ . $9000$ AL ALIVE = $367$ . $900$ . $1014$ . $960$ = $743$ . $900$ . $900$	A	IDNEN0285	• TO LAL	t.			
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$		FREDUENCI			• 000 • * * 2	TVIII CAV	100 July
$AL ALIV^{F} = 347,000 + 101,000 + 101,000 + 100,000 + $	<ul> <li>4</li> <li>4</li></ul>		•	= . 3 0:10'SX		i	Úvilu" y
AL ALIV <sup>6</sup> = 347.000 , 191.01 06.000 AL ALIV <sup>6</sup> = 347.000 , 191.01 06.000 AL ALIV <sup>6</sup> = 347.000 , 191.01 06.000	: •		<b>~</b>	TOTAL			
AL ALIVE = 347.000 , 191AL 96AD , = 243.000 , 62349 TRIAL =	ا		3 🐞 🐗 🗶 .				:
AL ALIVE = 347.000 , 19151 98AD . = 243.000 , (83399 TRIAL =	y .	500		100.000		an ann an ann an a ann an ann an ann	-
	AL ALIVE	195		- 17	1 000-542	TOTAL	000 <b>°(</b> 109

Figure XI.2 EXAX2 output -- chi square test of homogeneity across groups

## DEFOE COMPOUND C FIRY MARTALITY

# 

ALIVE       0.00       0.00       40.000         AD.000       20.00       40.000       40.000         AD.000       20.00       40.000       40.000         AD.000       20.000       40.000       40.000         AD.000       20.000       40.000       40.000         AD.000       20.000       40.000       40.000         AD.000       40.000       40.000       40.000         ALIVE       70.000       40.000       40.000         ALIVE       70.000       40.000       40.000         ALIVE       70.000       40.000       40.000         ALIVE       70.000       101.4       50.000         ALIVE       70.000       40.000       40.000         ALIVE       70.000       101.4       50.000         ALIVE       7.000       101.4       50.000         ALIVE       7.000       101.4       50.000         ALIVE       10.40       50.000       60.000         ALIVE       10.40       50.000       60.000         ALIVE       10.40       50.000       60.000         ALIVE       50.000       60.000       60.000         ALIVE	GROUP	•		CONTINGENCY	CY. LARLE				
3       38.000       20.000       40.000         5       31.000       9.000       40.000         5       31.000       9.000       40.000         6       31.000       9.000       40.000         7       1.000       40.000       40.000         7       1.000       40.000       40.000         7       1.000       40.000       1.000         1       31.369       8.631       40.000         31.369       8.631       40.000         5       31.369       8.631       40.000         6       31.369       8.631       40.000         6       31.369       8.631       40.000         6       31.369       8.631       40.000         6       31.369       8.631       40.000         6       31.369       8.631       40.000         6       31.369       8.631       40.000         7       1.111       57.00       57.00			r	ALIVE 47.500 47.600	0 ₹ 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	101×101 40.000			
AL ALIVE = 140.0000 , TUFAL DLAU = 172.70530 (DRAN LATE = 4.00 EXPECTEU FREQUENCISS X50005 = 172.70530 (DHDFF = 4.00 21.369 $B_{0}$ $B_{0}$ $B_{0}$ $B_{0}$ $D_{0}$ $D_{0}$ 21.369 $B_{0}$ $B_{0}$ $B_{0}$ $D_{0}$ $D_{0}$ 21.369 $B_{0}$ $B_{0}$ $B_{0}$ $D_{1}$			12455	340 310 310 310 310 310 30 310 30 30 30 30 30 30 30 30 30 30 30 30 30	0000 0000 0000 0000 000	00000000000000000000000000000000000000			
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$		TOIAL	٦٢،٢٢	. 140	•			, I 	•
I       ALIVE       DTAU       TOTEL         31-369 $B=631$ $49.0000$ 31-369 $B=631$ $49.0000$ 31-369 $B=621$ $40.0000$ 31-369 $B=621$ $40.0000$ $A_1 = 1369$ $B=631$ $40.0000$ $A_1 = 176$ $B=631$ $40.0000$	1	-	E XD :		1.1.6	;	12.705.31	- i	- 17(Jr.
31 - 366 $3 - 531$ $40 - 000$ $32 - 156$ $8 - 531$ $40 - 000$ $4 - 32 - 156$ $8 - 531$ $40 - 000$ $5 - 31 - 369$ $8 - 531$ $40 - 000$ $6 - 31 - 369$ $8 - 531$ $40 - 000$ $6 - 31 - 369$ $8 - 531$ $40 - 000$ $6 - 31 - 369$ $8 - 531$ $40 - 000$ $6 - 31 - 369$ $8 - 531$ $40 - 000$ $6 - 31 - 369$ $8 - 531$ $40 - 000$ $6 - 31 - 369$ $8 - 531$ $40 - 000$ $6 - 31 - 369$ $8 - 531$ $40 - 000$		The	1	21<	0-1) 8-631	1010L 40-000		•	
$A_{1} A_{1} V V = 1^{4} \frac{31 \cdot 349}{31 \cdot 369} - \frac{3 \cdot 531}{8 \cdot 531} - \frac{40 \cdot 000}{40 \cdot 000}$			1795 - 4	1000-1-					ŧ
1 - 1 - 1 - 1 - 0 - 0 - 0 - 1 - 0 - 1 - 0 - 1 - 0 - 0	1		t unico	• m m.	8.631	000-07			
	·	TGFAL	ALIVE	501	10 , TiJTsL	= .67-6	52.000	LUTAL	141 - 21-5

Figure XI.3 --EXAX2 output-- chi square test of homogeneity across groups

GROUP	н	•	TUDIT, AGENCY TARLE	ALE			
		•••**.			701AL 200-000		
•		171-5-10-10 171-5-10-10	137 -000 137 -000 134 -000 124 -000	78-000 78-000 78-000 76-000 76-000	scrca		
	TUTAL	AL 145	= 716.000	. TOTAL	0640 =	424.000 + GUAND	the state of the s
		r XP	CTED FETQUENCI		×SOURS =	3, 39914 CUTOFS	054 = 5.00000
•	COL	; ; ;	ALIV5 129-33	95AU 77-562	200-000		
	รี			70.667	200.000 200.000 200.000 200.000 200.000		
 • •	Col	¢.	10 - 0 - 1 10 - 0 - 1		200.000		
	THIAL	ALIVE	= 776.00U	141111 ·	n540 ≖	424-010 . GRAND	TUTAL = 1: "0. "CO

--EXAX2 output-- chi square test of homogeneity across groups Figure XI.4

--EXAX2 output-- chi square test of homogeneity across groups Figure XI.5

1000000 000°001 1 Ħ 11 SKAND TOTAL 41 GRAND FOTAL CUTUFF 40929\*68E 214.000 214.000 11 11 H Spras DIAU 0430 C C Ì + TOTAL LUTAL 54-57624 0110 35.0 ULV JU COULTNEENCY IABLE EXPERTED FREQUENCIES 396.000 386.729 COMP INFU 210 0 q ALIV 64.44 \* \* \* 7 7 7 4 н +1 5 ALIVE ALIVE HU UNID CHI TURAL TOTAL 200000 333355 v u GROUP \_

NOLCOMBE AND PHIMS COMPOUND D FAY NORTALITY

建立的复数 医生体水水水水 计实代生态 大型的的复数 医白白白白

## \*

Figure XI.6 --EXAX2 output-- chi square test of homogeneity across groups

					COMPOUND B		Mag1aL197	*******		
с <b>џ</b> р	1		CUNTING N	LAL YOU	۲.					
	реррер Кастра, Стрере		1011-000 991-000 994-000 934-0000 934-0000			10000000000000000000000000000000000000		· · · · · ·	:	
	TUÌÀL	ALIVE		67.000	, TOTAL	0€AD =	47,000	. GRAND TOTAL		614.000
		EXP	E XP. ECTED FREQ	EquENCIES			•	CUTOFF	-	5.0000
	TRT		3.26 2.26 2.26	• •	DEAD 7.731					
		19400	00000000000000000000000000000000000000		100021 10002 10002			· · · · · · · · · · · · · · · · · · ·	· · ·	• • • •
	TUIAL	ALIVE	л И	567.000	, TOTAL	DEAD =	47.000	, GRAND TO	TOTAL =	634.067
<u>_</u> 21	THE FUDE	ADENT TESTS C CHT TENTS EU STGWTF1	S COMPIN APEU USS ULEEL E		16-21426	0.057349107	101			

GRUL

\*\*\*\*\*\*\*\*\*\*

-	CUNTINGINCY TAI	TAbL				
	•	00 00 00 00 00 00 00 00 00 00 00 00 00	4000000 4000000 4000000 0000000 0000000 0000000		•	
ALT VE EXPEC	= 133.000 ; CTED_FREQUENCIES	, 101AL	DEAD =	46.090	, GRAND TOTAL = cutoff = 5	179.°čeň 5. nadon
i	2010 2010 2010 2010 2010 2010 2010 2010		10100000000000000000000000000000000000		· ·	
5 6	00 NN 00	7.709 7.709 , TCTAL	00	46.000	, GRAND TOTAL =	179.000

--EXAX2 output-- chi square test of homogeneity across groups Figure XI.7

107

**GR** ()(1)

THE UNAMED CATEGORY NAME & E DEFINED THAN THERE CATEGORY BAJCH TS TREATED AS A MISSING VALUE. SENSITIVITY TESTING OF ORDINAL HEASURES' . SENSITIVITY TESTING OF URDINAL MEASURES FACTLUTT LOS ANGELES EGENTS OF THE UNIVERSITY OF CALLEORNIA. ÷ MAT JONS 1000000 EXTENSIVE CHANGES HAVE BEEN MADE TO THIS PROGRAM. Be sure to use 1977 Manual. CONC. SALIVE , DEAD. KANSFORMATIONS DATA F5.0) . NC . MUR DEFINE OPTION DOES NOT FUNCTION. NUMBER OF VARIABLES TO READ IN NUMBER UF VARIABLES ADDEU BY TRANSFORM NUTEL UMMER OF VARIABLES NUMEL OF CASS TO READ IN UT AL UF CASS TO READ IN UT AL US CARLABLES UN MISSING VALUE CHECKED BEFOR LIMITS ANU MISSING VALUE CHECKED BEFOR LANNS ARE TO THE CHECKED BEFOR TADNS ARE TO THE CHECKED BEFOR TADNS ARE TO THE TO READING. HORT COLUMN I NAMES (2) GAMAA. TITLE I VARIABLE TABLE I FORMAT NAMES A PROGRAM CONTROL INFORMATION VARIABLES TO BE USED /STATEGORY /STATEGORY /END YARIABLE /PRUBLEM PROBLEM TITLE .. INPUT FORMAT BMOPLF -HEALIH S UNIVERSI CUPYRIGH - |

Input information required for program BMDP1F Figure XI.8

Figure XI.9 BMDP1F output from one sided measure of association test on artificial data

TUTAL 100 100 100 100 100 000 DEP. N PRUB. D.F. A SF1 0.02 0.03 SUMER'S D -0.125 0.082 -1.504 1 SOMER'S D ... -0.123 0 VALUE -TOTAL 549 6.00 -100 S TAU-C I VAR **5.00** STATISTIC 00 d=₫(-1.504) = 0.066 STATIST STATIST CONC 4.00 6 21 VS ASE1 1-VALUE DEP 0.098 -1.504 CONC (YAR 1) 3.00 80 6 0.F. PRUB. (VAR 2. One Sided Measure of Association Test 20 70 MOR 1 2.00 93 100 ŗ -0.150 2.250 Observed Significance Levels 1. Chi Spane Test 1.00 64 CELL FREQUENCY COUNTS EQ./EQ. 1.00 PLARSUN CHISQUARE TARLE NU. HUAR 2) ALIVE STATISTIC GAMMA

.....

CELL FREQUENCY COUNTS CONC (YAR EQ./EQ. 1.00 2.00 3. HURT 2 ALIVE 2.00 94 92 9	CELL FREQUENCY COUNTS CELL FREQUENCY COUNTS CONC (YAR 1) EQ./EQ. 1.00 2.00 3.00 4.00 5.00 URT 2) ALIVE 2.00 94 92 90 82 86 10 100 100 100 100 MINIMUM NONZERO EXPECTED VALUE 15 16 16 100 100	6.00 944 100	TO TAL 5 66 6 00	
EQ./EQ. 1.00 2.00 2.00 94 92	11 •00 4.00 90 88 00 100	6.00 946 100	TAL 66 34	
EQ./EQ. 1.00 2.00 1.00 56 8 2.00 94 92	00 4.00 10 12 90 100	6.00 16 100 100	TAL 66 34 00	
2.00 94 9	10 12 90 88 00 100	16 1 100	66 34 00	
	00 100	100	00	
TUTAL 100 100 1		ENCIES TISTIC		
PEAKSON CHISQUARE 7.150 D.F. 0.2	PR 08 5 7A	•	VALUE	0.F. PRCH.
STATISTIC VALUE ASEL T-VA GAMA SOMEA'S D -0.199 0.071 -2.	-VALUE DEP. 514 -2.699 DEP. 510 -2.699 1 SOM	TISTIC ART 5 TAU-C Pails 7 TAU-B ERLS D	VALUE VALUE -0.078 -0.046	6511 T-VALUE DEP. 0.029 -2.699 0.035 -2.699 2
Observed Significance Levels				
1. Chi Source Test	-	-		

 . . .

••••••

. . •••••

Figure XI.10 BMDP1F output from one sided measure of association test on artificial data

TUIAL 100 100 100 100 100 600 DEP. PROB. 0.1. VALUE SGMER'S D -0.250 0.063 -3.738 1 SOMER'S D TOTAL 519 6.00 104 7 STAU-C ( VAR 5.00 **STATISTIC** 18. 92 8=0.01 2) VS CONC 4.00 15 2. One Sideal Measure of Association Test -0.298 0.075 -3.738 DEP I VAR 1) 3.00 0.F. PR08 288 (VAR 2.00 MUR I D Observed Significance Levels 6 CONC 13.487 1. Chi Squar Test 1.00 64 CELL FREQUENCY COUNTS EQ ./EQ. 2.00 PLAKSON CHISQUARE TARLE NO. MURT DEAD (VAR 2) ALIVE STATISTIC CAMMA

H.

H

•••••

••••

Figure XI.11 BMDP1F output from one sided measure of association test on artificial data

-YALVE DEP. PRUB. ϕF• VALUE FOT AL 243 **00**•9 0.503/ \$=0.89 (VAR MINIMUM NONZENO EXPECTED VALUE IS 40.500 ----- ALL STATISTICS ARE COMPUTED USING THE ORIGINAL CELL FREQUENCIES EXCLUDING ROWS AND/OR COLUMNS WHICH ARE ZERO. STATI STIC 5.08 100 2) VS TRTGRP 2. One Sided Measure of Association Test 00°\* Defoe Compound C Embryo Mortality Data 8 VAL & ASEI T-PALUE DEP. THE ANALYSIS IS OF THE DATA USING THE FREQUENCY COUNTS IN FREQ (VAR 3) Observed Significance Levels JRTGRP ... ( VAR. 1) VALUE D.F. PROB. 5.955 D.F. 0.3106 3.00 53 VAR 100 2.00 MURT 404 108 1. Chi Square Test 1. 0d 100 46 CELL FREQUENCY COUNTS EQ./EQ. 2.000 TOTAL STATISTIC TABLE NO. HORT DEAD STATISTIC -GAMAA SOMER 'S U

Figure XI.12 BMDP1F output on real data

DÊP. D.F. PROB. VÄLUE TOTAL 52 00.9 **IVAR** 5°8 MINIMUM NOV JERO EXPECTED VALUE IS ----- ALL STATISTICS ARE COMPUTED-USING THE ORIGINAL CELL FREQUENCIES EXLLUDING ROWS ANU/OR COLUMNS WHICH ARE ZERO. STATI STIC **TRTGRP** \$ =0.000 00.4 2) VS -0.963 0.022 -10.615 DEP 2 2. One Sided Measure of Association Test -0.925 0.029 -10.615 (VAR 3.8 VALUE D.F. PRUB. 182.785 D.F. 0.0 ( VAR ş Defoe Compound C Fry Mortality Data 38 TRIGRP LAR Observed Significance Levels THE AMALYSIS IS OF THE DATA USING THE FREQUENCY COUNTS IN FREQ 2.00 MORT 3 00 1. Chi Square Test 1.00 00 4 CELL FREQUENCY COUNTS 70746 EU ./EQ . 1 •00 2 •00 STATISTIC PEAKSUN CHISQUARE FABLE ND. HORT DEAD STATISTIC GAMMA SUMER'S U

Fig.re XI.13 BMPD1F output on real data

000,000

HOLCOMBE AND PHIPPS COMPOUND D EMBRYD MORTHLITY DATA

P

THE ANALYSIS IS OF THE DATA USING THE FREQUENCY CUUNTS IN FREQ

9 ( VAR

2 2) VS TATGRP (VAR ( VAR MORT -TABLE NO.

CELL FREUVENCY COUNTS

2 TRTGRP (VAR

	Eu «/ Eu	1.00	'E4. 1.00 2.00 3.00 4.00 5.00 6.00 TOTAL	3•00	4 •00	5.00	<b>00</b> •9	TOT AL		
MORT DEAU IVAR 2) ALIVE	2.000	•00 •00 130	130	137	122	193	124	124 1 776	·	
	TOTAL	20.0	200	200	200	200	200	1200	T0TAL 200 200 200 200 200 200 200 200 200 20	
	* * * * * * * * * *			• • • • • • • •			•••••		****	•
MINIMUN NUNZERU EXPECTED VALUE IS USING 70.667 ALL STATISTICS ARE COMPUTED USING THE URIGINAL CELL FREQUENCIES	STICS ARE CO	LUE IS	SING THE	URIGINAL	CELL FREG	DENCIES.				

ALUE DEP. P ROB. 0.F. VALUE TAU-C STATI ŠTIĆ EXCLUDING ROWS AND/OK COLUMNS WHICH ARE ZERO. ASET T-VALUE DEP VALUE D.F. PRUB -0.026 STATISIIC PEARSUN CHISQUARE STATISTIC<sup>-</sup> GAMMA

od = 0.26 ×=0.64 2. One Sided Mensure of Association Test evel - Observed. Significance 1. Chi Square Test

....

..........

+0-0-

0°034

-0-022

SUMLK'S D

Figure XI.14 BMDP1F output on real data

TARE NO. 1 MURT (VAR 2) VS TRTCRP (VAR 1) TARE NO. 1 MURT (VAR 2) VS TRTCRP (VAR 1) TRTCR (VAR 2) VS (RTCRP (VAR 1) TRTCR (VAR 2) VS (RTCRP (VAR 1) RMM 2) RHG (SO - 100 2.00 3.00 4.00 5.00 6.00 TOTAL RMM 2) RHG (SO - 1.00 2.00 3.00 4.00 5.00 6.00 TOTAL RMM 2) RHG (SO - 2.00 3.00 3.00 4.00 5.00 6.00 TOTAL RMM 2) RHG (SO - 2.00 3.00 3.00 4.00 5.00 6.00 TOTAL RMM 2) RHG (SO - 2.00 3.00 5.00 6.00 TOTAL RMM 2) RHG (SO - 2.00 3.00 4.00 5.00 6.00 TOTAL RMM 2) RHG (SO - 2.00 3.00 5.00 100 100 100 000 RMM 2) RHG (SO - 2.00 100 100 100 100 000 RMM 2) RHG (SO - 2.00 100 100 100 000 RMM 2) RHG (SO - 2.00 2.00 5.00 5.00 5.00 5.00 100 100 000 RMM 2) RHG (SO - 2.00 2.00 5.00 5.00 5.00 5.00 5.00 5.00	THE ANALYSIS IS OF THE DATA USING THE FREQUENCY COUNTS IN FR	HE DATA IN	FREG	ter (VAR 3)		5						
UCENCY COUNTS         TRTGRP (MR 1)           PRO-FEO         1.000         2.000         3.000         4.000         5.000         0.00         TOTAL           2.000         9.0         9.0         9.0         3.000         4.000         5.000         701AL           2.000         9.0         9.0         3.000         4.000         5.000         7000         1000         1000         1000         1000         1000         0.00           1.17AL         1.000         1.000         1.000         1.000         1.000         1.000         0.00           1.17AL         1.000         1.000         1.000         1.000         1.000         1.000         1.000           1.17AL         0.4100         0.0100         5.1417         1.000         1.000         0.000           1.17AL         0.41614         0.74145         0.4164         0.414	TABLE NO. 1		MURT		VS	_						
FOO-FEO- 1.000       2.000       3.000       4.000       5.000       6.000       TOTAL         2.000       9.4       9.5       9.9       9.3	CELL FREQUENCY COUNT	5		: : : :					: : ! t		)    -  -	•
F0-FG0       1.00       2.00       3.00       4.00       5.00       6.00       TOTAL         2.000       9.4       9.2       9.2       8.7       2.9       8.7       2.9         2.000       9.4       9.2       9.2       8.7       2.9       8.7       2.9         1011       100       100       100       100       100       100       600         EXPECTED VALUE IS       3.05       0.00       100       100       100       100       0.00         EXPECTED VALUE IS       3.05       0.05       5.60       5.40       0.00       0.00       0.00         CLUGIT MERE COMMENTED VALUE IS       0.00       100       100       100       100       0.00         EXPECTED VALUE IS       0.010       0.00       100       100       100       0.00         CLUGIT MERE COMMENTED VALUE       0.5       0.00       5.140       0.00       0.00         CLUGIT MERE AND/OR SOUGHURE SALED       0.5       0.00       0.00       0.00       0.00         CLUGIT MERE AND/OR SOUGHURE SALED       0.5       0.00       0.00       0.00       0.00       0.00       0.00         CLUGIT MERE AND AND AND AND AND AND AND AND AND		F	4	VAR 1.1				:				
2.000 96 92 92 87 27 100 100 100 366 TUTAL 100 100 100 100 100 000 600 EXPECTED MULEIS LUCOING ROWS AND/OR 2010 MICH ARE CELLOTAR QUENCIES LUCOING ROWS AND/OR 2010 MICH ARE CELLOTAR QUENCIES CLUCOING ROWS AND/OR 2010 MICH ARE CELLOTAR QUENCIES CLUCOING ROWS AND/OR 2010 MICH ARE CELLOTAR QUENCIES CLUCOING ROWS AND/OR 2010 MICH ARE CELLOTAR QUENCIES VALUE D.F. 0.000 STATISTIC VALUE D.F. VALUE D.F. 0.000 STATISTIC VALUE D.F.		1.00	2 <b>0</b> 0	3+00	4.00	5.00	6.00	TOTAL	,			
TUTAL 200 200 200 200 100 100 200 000 600 EXPECTED VALUE IS CLUDIANE ADMS 2014 UNIS WHICH ARE ZERUGA CLUDIANE ADMS 2014 UNIS WHICH ARE ZERUGA CLUDIANE ADMS 2014 UNIS WICH ARE ZERUGA E 360 41 WILLE IS VALUE DoF STATISTIC VALUE 0.F. VALUE 0.F. 0.F. VALU	Z) ALIVE	<b>94</b> 0		88 68	61	21	100	1 214			· · · · · · · · · · · · · · · · · · ·	
EXPECTED MILLE IS CLUODING ROUGHED AND ON SOUTH SUNT CELL FREQUENCIES CLUODING ROWS NUT ON SOUTH SUNT STICT CLUODING ROWS NUT ON SOUTH SUNT STICT SUNT STICT STATISTIC VALUE 0.0525 - 24.2425 DEP. STATISTIC VALUE 0.0525 - 24.2425 DEP. STATISTICT VALUE 0.0525 - 24.242		100	81	100	100	100	100	009	1 - - - - - -	;		
E 369-076 DoF 90000 STATISTIC VALUE DOF. VALUE 0.65 0.000 STATISTIC VALUE 0.6000 STATISTIC VALUE 0.0000 STATISTIC VALUE 0.	ALL STATISTICS AND ADD	NPUTED USI S AND/OR CI	VG THE D DL UMNS N	HIGINAL (	ZERO+	NENCIES				•		•
- VABUE 0.8615 - 2421 - 0.833 0.028 - 2422 Mad Marsure of Association Test	KI STIC KSUN CHI SQUARE	VALUE 389 •676	: 	PR09.	<b>S</b>	TATI STIC			VALUE	0.F.	PROB.	
al Significance Levels a Test Mad Measure of Association Test		VALUE -0.885 -0.833	615 025 028		4 • •	IATISTIC UARISTIC NDALL'S TA	5 <b>1</b> 13		1000 1000 1000	0030 0030 0030 0030 0030 0030 0030 003	T-VALUE DEP -24-242 -24-242	
Test	d Signi	unce Levels				•						•
Test	1. Chi Spune Test		•	Or N	000	-					•	
	2. One Sided Measure o	f Associat	in Test		000	-			•	· •		

Figure XI.15 BMDPlF output on real data

## XII. TREATMENT GROUP VS CONTROL GROUP PAIRWISE MULTIPLE COMPARISON PROCEDURES

If the overall test rejects the hypothesis of no concentration related effects we must determine which treatment group response rates differ from the control group rate. A number of procedures can be used for such inferences, some based on hypothesis testing and some based on confidence interval estimation. In this section we consider several approaches based on tests of hypotheses. The discussion is by no means exhaustive. In the following section we discuss confidence interval procedures.

A common approach to multiple comparisons on qualitative response rate data such as mortality rates is to carry out an arc sine normalizing transformation on the observed response rate within each group and then compare each treatment group with the control using Dunnett's or Williams' procedures [34, 35, 36, 37]. Such procedures are based on asymptotic theory whose validity is questionable if there are a number of small expected frequencies.

An alternative multiple testing approach is to carry out a succession of 2 x 2 contingency table tests of homogeneity between each treatment group and the control group, based on Fisher's exact test [13] or on asymptotic theory depending on expected frequencies. Our EXAX2 program will do this. A treatment group is said to be (statistically) significantly different from the control group at e.g. the  $\alpha = 0.05$  level if the pairwise test rejects the null hypothesis after adjusting for simultaneity by Bonferroni's method. (i.e. If we perform five pairwise comparisons and wish to guarantee an overall  $\alpha = 0.05$  type one error level then each individual comparison must be made at the  $\alpha/5 = 0.01$  level). Note that this approach does not impose any monotonicity structure on the response rates and so may not be most sensitive to detect small to moderate effects.

Dunnett [ 34 , 35 ] presents a procedure for multiple comparison of each of the treatment group responses with the control group response, controlling the overall error rate for all comparisons. His procedure is derived for quantitative responses, assumed to have equal variability. He assumes equal replication among the treatment groups with equal or possibly greater replication of the control group. We might apply this procedure to qualitative response data from toxicity tests after performing an arc sine variance stabilizing related on the observed responses.

A problem with the application of Dunnett's procedure to the analysis of data from toxicity tests is that is does not take full account of the structure of the problem. Namely the various treatment groups correspond to increasing toxicant levels. One might therefore assume a monotone (increasing or decreasing) response level with increasing group number. Since Dunnett did not build such a monotoxicity assumption into his procedure, it loses some sensitivity. Williams [36, 37] assumes a monotone response function. He estimates the treatment and control group response rates under the monotoxicity restraint and uses these estimates for treatment group - control group comparisons. See Williams [36] for details. Chew [38, pp.26-27] briefly describes Williams' method and presents tables for its implementation. Williams [ 36 ] assumes equal replication for all concentrations (including the control). He extends this procedure [ 37 ] to accomodate increased replication in the control group, two sided tests, and modifications to account for unequal replication among the treatment groups. As with Dunnett's procedure, we can apply Williams method to qualitative response data after carrying out an arc sine transformation.

7

We illustrate Williams' method with several examples based on results from fish toxicity tests. Consider first the fry mortality data from the Holcombe and Phipps test on compound D. From the preliminary scatterplot in Figure VI.6 and the overall tests of significance in Section XI, it is quite evident that fry mortality increases with increasing toxicant level. We wish to determine here which treatment groups exhibit significantly greater fry mortality rates than the control group. As the result of the within groups heterogeneity test was marginal ( $\alpha = 0.14$ , see section VIII) we do not adjust the data prior to carrying out Williams' procedure.

The basic and transformed responses, pooled across tanks within groups are:

Group (i)	1	2	3	4	5	6
Sample Size (n <sub>i</sub> )	100	100	100	100	100	100
Response Rate $(\hat{p}_i)$	0.06	0.08	0.08	0.13	0.79	1.00
2Arc Sin $\sqrt{\beta_i} \equiv \hat{\mu}_i$	0.495	0.574	0.574	0.738	2.190	3.142

Since these estimates are already in monotone sequence, they do not need to be modified. We declare the group i response rate to be significantly different from the control rate if

$$\hat{\mu}_{i} - \hat{\mu}_{1} > \bar{t}(2/n)^{1/2}$$

The factor  $\bar{t}$  can be obtained from Williams' tables corresponding to the 5% or the 1% significance level. The yardstick  $\bar{t}(2/n)^{1/2}$  is based on the assumption that the variance of 2 arc sin  $\sqrt{\hat{p}}$  is 1/n. In our example n = 100 and  $\bar{t} = 1.756$  (corresponding to 5 treatment groups and  $\alpha = 0.05$ ). Thus the response in group i is declared to differ significantly from

the control group response if  $\hat{\mu}_i > 0.495 + 1.756 (2/100)^{1/2} = 0.743$ . Groups 5 and 6 differ significantly from the control and group 4 is just on the borderline.

We now examine the effect on the outcome of this procedure of applying an adjustment for tank to tank heterogeneity. From Section IX B we see that this factor is  $\overline{K} = 1.337$  for the above data. Thus the "effective" sample size per group is 100/1.337 = 74.79 and the decision point for Williams' procedure becomes  $0.495 + 1.756(2/74.79)^{1/2} = 0.782$ . Group 4 is no longer borderline.

We now apply this same procedure to the embryo mortality data from the Jarvinen test on compound B. The result of the within groups heterogeneity test was highly significant ( $\alpha = 0.005$ ) and so we first adjust the data prior to carrying out Williams' procedure. From Section IXB the adjustment factor is  $\hat{K} = 3.071$ . The basic and transformed responses along with effective sample sizes, pooled across tanks within groups are:

Group (i)	1	2	3	4	5	6
Effective Sample size (n <sub>i</sub> / <del>k</del> )	32.89	35.49	32.56	31.91	34.19	32.89
Response Rate (p̂ <sub>i</sub> )	0.139	0.073	0.090	0.020	0.057	0.079
2 Arc Sin $\sqrt{\hat{p}_i} \equiv \hat{\mu}_i$	0.764	0.547	0.609	0.284	0.482	0.570

For the sake of simplicity we will utilize an average sample size of 33.32 within each group, but the calculation could alternatively be carried out based on the individual group sample sizes. Since  $\{\mu_i\}$  are not in monotone sequence we must first modify them by an averaging process discussed in Williams [36] or in Chew [38] until the resulting estimates satisfy the monotoxicity constraint. We obtain 0.537, 0.537, 0.537, 0.537, 0.570. We declare group i significantly greater than the control group if  $\mu_i > 0.764 + 1.756 (2/33.32)^{1/2} = 1.194$ . Obviously no treatment groups have significantly greater response than the control group. (Interestingly if we carry out Williams' procedure on these data to look for a monotone decreasing trend in response rate, we arrive at the same conclusion. That is, no group has significantly lower response rate than the control group).

Dunnett's and Williams' procedures are based on asymptotic theory. If the response frequencies do not justify the use of

asymptotic theory we can carry out a succession of <u>exact</u>, <u>small sample</u>  $2 \ge 2$  treatment-control comparisons by means of Fisher's exact test, adjusting for simultaneity by Bonferroni's method. Consider for example the comparison of treatment group 4 with the control group for the fry mortality data from Holcombe and Phipp's test on compound D. We have the following  $2 \ge 2$  table.

	CONTROL	GROUP 4	
DEAD	6	13	19
LIVE	94	87	181
	100	100	200

Here, in the notation of Lieberman and Owen,

k = 19 n = 100 N = 200x = 6

Interpolating in the Lieberman and Owen tables [26]

between N = 100 and N =  $\infty$  we have

N = 100	$P(X \le 6) = .062$	1/N = .01
$N = \infty$	$P(X \le 6) = .0835$	1/N = 0
N = 200	$P(X \le 6) = ?$	1/N = .005

Thus P(X < 6) = 1/2(.0835 + .062) = .073

Thus this table is significant at the .07 level (not accounting for simultaneity).

This exact test procedure is thus seen to be somewhat less sensitive than Williams' procedure applied to the same data. This is understandable since it does not incorporate the monotonicity structure of the response rates.

## XIII. CONFIDENCE INTERVAL PROCEDURES FOR COMPARISON OF TREATMENT GROUP AND CONTROL GROUP RESPONSE RATES

## A. Introduction

We have previously considered overall tests of hypotheses to compare response rates in replicate tanks within treatment groups and to compare response rates across treatment groups. In this section we consider procedures for constructing confidence intervals to compare response rates in the treatment groups with that in the control group on a pairwise basis.

It is well known that hypothesis testing procedures are somewhat limited in their conclusions. They merely state whether the null hypothesis was accepted or rejected but give no indication of the extent of the effect. Thus we have no idea of the <u>biological</u> significance of the effect (as opposed to its statistical significance). The rejection or nonrejection of a null hypothesis is often more a result of sample size than of the biological importance of the effect. The determination of acceptable concentrations should be based on what are biologically significant effects rather than on the power function of a hypothesis testing procedure.

Confidence intervals are more informative than tests of hypothesis. The widths of the confidence intervals indicate the <u>degree</u> of <u>precision</u> in the data concerning the estimates of the quantities of interest in our inferences. Narrow confidence intervals signify precise inferences while wide confidence intervals signify imprecise inferences.

In the discussion in this section we consider the case of no tank to tank heterogeneity within groups. Thus we pool responses across tanks within groups to arrive at average response rates within groups. The presence of tank to tank heterogeneity can be accounted for by

- Fitting a model which explicitely accounts for heterogeneity of response rates across tanks -- for example the beta binomial extension of the binomial model, the negative binomial extension of the Poisson model, or a variance components extension of a fixed effects analysis of variance model for quantitative responses.
- 2. By carrying out analyses on a per tank basis rather than on a per fish basis. This approach is conservative and greatly diminishes the number of degrees of freedom available for error estimation.
- 3. By adjusting the data to account for the extent of tank to tank variation. Namely tank to tank variation can be

regarded as correlated responses within tanks, generally positively correlated. Thus the variability of the average responses within tanks is greater than would be the case if the responses were independent within tanks. Such reduction in variation can be simply accounted for by reducing the "effective" sample size within tanks to a lesser value and then ignoring the within tank correlation and proceeding with binomial based procedures or the like. The reduction in "effective" sample size reduces the precisions of the estimates and test statistics just as does correlation effects.

The procedures discussed in this section, although based on binomial theory, can be used in conjunction with adjustment method 3. Thus they are also relevant in the case when tank to tank heterogeneity exists.

Consider the Holcombe and Phipps compound D fry mortality data. We wish to compare the response rate in treatment group 4 with that in treatment group 1 (the control group). The basic data, pooled across tanks within groups, is

	CONTROL	GROUP 4	
DEAD	6	13	19
LIVE	94	87	181
	100	100	200

Fisher's exact test (without simultaneity adjustment) says that  $p_4$  is "significantly" greater than  $p_1$  at the  $\alpha = 0.07$  level. However a significance statement such as this says nothing about the magnitude of  $p_4/p_1$ . Estimating the value of this ratio is important for assessing whether there is a <u>biologically significant</u> increase in mortality between the control group and group four. Confidence interval procedures enable us to estimate  $p_4/p_1$  and determine the precision of our estimate as well determine whether  $p_4$  is (statistically) significantly greater than  $p_1$ .

There are three approaches to the construction of confidence intervals in the case of quantal response data.

- Large sample normal theory confidence intervals.
- Exact, small sample confidence intervals based on the noncentral distribution of the 2 x 2 contingency table, conditional on the margins. (See Thomas [ 39]) for the theory and the algorithm.

We have implemented this algorithm in EXAX2[14].

 Approximate confidence intervals based on Poisson theory. These intervals are most appropriate when the response probabilities are small (usually under .10).

It should be noted that these procedures do not take the monotonic nature of the response probabilities into account. We consider each of these approaches in turn.

## B. Method 1 Asymptotic Approach

To use the asymptotic normal approach we adopt a conservative yardstick and require that each cell in the  $2 \times 2$  table under consideration contain at least 5 responses.

For most situations of practical interest both  $p_1$  and  $p_4$  will be relatively far from 1. Certainly if  $p_1$ , the mortality rate in the control group is close to 1, the test will be terminated. If  $p_4$  is very close to 1 while  $p_1$  is close to 0, there is no need in calculating confidence intervals on their ratio. Group 4 will be obviously unsatisfactory.

We wish to calculate an asymptotic theory confidence interval on the ratio

$$\theta = p_4/p_1$$

Let

 $\phi \equiv \ln\theta = \ln p_4 - \ln p_1$ 

We estimate  $\phi$  by

$$\hat{\phi} \equiv \ln \hat{p}_4 - \ln \hat{p}_1$$

As  $N_1$ ,  $N_4 \rightarrow \infty$  with  $p_1$ ,  $p_4$  fixed

$$\hat{\phi}$$
 is approximately  $N\left(\phi, \frac{q_1}{N_1 p_1} + \frac{q_4}{N_4 p_4}\right)$ 

Thus an approximate 95% confidence interval on  $\phi$  is

$$\hat{\phi} - 1.96 \left[ \frac{q_1}{N_1 p_1} + \frac{q_4}{N_4 p_4} \right]^{1/2} \le \phi \le \hat{\phi} + 1.96 \left[ \frac{q_1}{N_1 p_1} + \frac{q_4}{N_4 p_4} \right]^{1/2}$$

In the case of the Holcombe and Phipps example

- -. - -

ł

$$N_1 = N_4 = 100$$
  $\hat{q}_1 = 1 - .06 = .94$   
 $\hat{p}_1 = 6/100 = 0.06$   $\hat{q}_4 = 1 - .13 = .87$   
 $\hat{p}_4 = 13/100 = 0.13$ 

Substituting  $\hat{p}_1$ ,  $\hat{q}_1$ ,  $\hat{p}_4$ ,  $\hat{q}_4$  for the corresponding parameters in the standard error formula we have

$$\left[\frac{\hat{q}_1}{N_1\hat{p}_1} + \frac{\hat{q}_4}{N_4\hat{p}_4}\right]^{1/2} = \left[\frac{.94}{100(.06)} + \frac{.87}{100(.13)}\right]^{1/2} = 0.473$$

Thus an approximate 95% confidence interval on  $\phi$  is

$$(0.773 - 1.96(.473), 0.773 + 1.96(.473)) = (-.154, 1.700)$$

Therefore,  $(e^{-.154}, e^{1.700})$  is an asymptotic 95% confidence interval on  $\theta \equiv p_4/p_1$ .

This interval is

(0.857, 5.474)

The conclusions from this confidence interval calculation are

- $p_4$  is <u>not</u> "significantly" different from  $p_1$  at the .05 level since the confidence interval contains 1. (Note that we observed borderline significance with Williams' procedure at  $\alpha = 0.05$ ).
- $p_4$  is not very much smaller than  $p_1$  (at least 86% of  $p_1$ ) but may be much larger than  $p_1$  (as much as 5.5 times  $p_1$ )
- $p_4/p_1$  is not determined very precisely by the data, based on such a comparison.

We have thus quantified the relation between  $p_1$  and  $p_4$ .

We now calculate 95% confidence intervals to compare the response rates in each of the other treatment groups with that in the control group.

	CONTROL	GROUP 2	GROUP 3	GROUP 4	GROUP 5	GROUP 6	
DEAD	6	8	8	13	79	100	
LIVE	94	92	92	87	21	0	
	100	100	100	100	100	100	

## Holcombe and Phipps Compound D Fry Mortality Data Pooled Across Tanks within Treatment Groups

We are not particularly interested in comparing group 6 with the control since it is obviously inferior. We thus compare groups 2, 3, and 5 with the control group by means of asymptotic 95% confidence intervals.

Group 2 vs Control

$$\hat{\theta} = \hat{p}_2/\hat{p}_1$$

$$\hat{\phi} = \hat{l}n\theta = \hat{l}np_2 - \hat{l}np_1$$

$$\hat{\phi} = \hat{l}n\hat{p}_2 - \hat{l}n\hat{p}_1 = \hat{l}n0.08 - \hat{l}n0.06 = 0.288$$

$$\hat{stderr}(\hat{\phi}) = \left[\frac{\hat{q}_2}{N_2\hat{p}_2} + \frac{\hat{q}_1}{N_1\hat{p}_1}\right]^{1/2} = \left[\frac{.92}{100(.08)} + \frac{.94}{100(.08)}\right]^{1/2} = 0.521$$

$$\hat{\phi} + 1.96 \text{ stderr}(\hat{\phi}) = 0.288 + 1.96(0.52) = (-.733, 1.309)$$

Thus an asymptotic 95% confidence interval on  $\theta$  is  $(e^{-.733}, e^{1.309}) = (0.480, 3.703)$ . This implies that there is no statistical evidence at the .05 level of a difference between  $p_2$  and  $p_1$ . Furthermore the present data do not determine this ratio very precisely.

Group 5 vs Control

$$\phi = \ln \hat{p}_5 - \ln \hat{p}_1 = \ln 0.79 - \ln 0.06 = 2.578$$
  
stderr( $\hat{\phi}$ ) =  $\left[\frac{.21}{100(.79)} + \frac{.94}{100(.06)}\right]^{1/2} = (.159)^{1/2} = 0.399$   
 $\hat{\phi} \pm 1.96$  stderr( $\hat{\phi}$ ) = (1.794, 3.362)

Thus an asymptotic 95% confidence interval on  $\theta$  is  $(e^{1.794}, e^{3.362}) = (6.011, 28.859)$ . There is thus overwhelming statistical evidence that the response rate in group 5 is substantially greater than that in the control group, by at least a factor of 6. The interval however is very

wide and so we cannot determine the ratio very precisely.

We may wish to modify these intervals for simultaneity. Since we are calculating 4 confidence intervals we can adjust their levels to attain a <u>familywise confidence level</u> of 0.05. The simplest way to do this is by means of Bonferroni's inequality. Namely we construct each interval at individual confidence level 1 - (.05/4) = .9875. The appropriate normal distribution factor then becomes 2.50.

$$\exp\left\{ \ln(\hat{p}_{j}/\hat{p}_{1}) + 2.50 \left[ \hat{q}_{j}/N\hat{p}_{j} + \hat{q}_{1}/N_{1}\hat{p}_{1} \right]^{1/2} \right\} \qquad J = 2, 3, 4, 5$$

These intervals are:

Group	2 vs	Control	(.363, 4.906)
Group	3 vs	Control	(.363, 4.906)
Group	4 vs	Control	(.664, 7.067)
Group	5 vs	Contro1	(4.857, 35.712)

We thus conclude that there is strong statistical evidence that group 5 has at least 5 times the response rate of group 1 but there is not enough statistical evidence to distinguish the response rates at groups 2, 3, 4, from that at group 1. Furthermore the data are not sufficient to make precise inferences about the ratios of treatment group to control group response rates without putting further structure on the problem such as assuming some sort of dose response relation. We will consider this approach in subsequent sections.

## C. Method 2 Exact, Small Sample Confidence Intervals

If the sample sizes are not sufficiently large to apply the asymptotic confidence interval procedure (method 1) and if response proportions are not sufficiently small to apply Poisson theory (method 3), then confidence interval comparisons between treatment groups and control group can be made by an <u>exact, small sample procedure</u>. This procedure is based on the non null distribution of Fisher's exact test in  $2 \times 2$  contingency tables.

Consider a  $2 \times 2$  contingency table to compare the response rate in a particular treatment group with that in the control group.

	Control	Group 2	Total
Dead	x <sub>1</sub>	x <sub>2</sub>	$x_1 + x_2$
Live	m - X <sub>1</sub>	m - X <sub>2</sub>	$m + n - (X_1 + X_2)$
	m	n	m + n

Let  $p_1$ ,  $p_2$  denote the response probabilities (e.g. probability of death) within the control group and treatment group respectively. We can test the hypothesis

Ho: 
$$p_1 = p_2$$
  
As  $H_1: p_1 \neq p_2$ 

Some and the states of the second second

by means of Fisher's exact test (Lehmann [ 25 ], Lieberman and Owen [ 26 ]), conditional on the margins of the table being fixed. This test is based on the hypergeometric distribution. We reject  $H_0$  if  $X_2$  is too extreme.

The nonnull distribution of  $X_2$  conditional on  $X_1 + X_2 = t$  is

$$P(X_{2} = x | X_{1} + X_{2} = t) = C_{t}(\rho) (t^{m}_{-x})(x) \rho^{x} \qquad x = 0, 1, 2, ..., t$$

$$\rho = \frac{p_1/q_1}{p_2/q_2}$$

The quantity  $\rho$  is known as the <u>odds ratio</u> and is very important for power calculations and for calculating confidence intervals to compare the response rates in the treatment and control groups on a pairwise basis.

The odds ratio is a quantity between 0 and  $\infty$ .  $\rho = 1$  if and only if  $p_1 = p_2$ . If  $\rho > 1$  then  $p_1/p_2 > 1$  and if  $\rho < 1$  then  $p_1/p_2 < 1$ . The size of the confidence interval on  $\rho$  indicates how precisely this quantity can be estimated from the data.

Thomas [39] presents an algorithm for calculating exact, small sample confidence intervals on  $\rho$  based on the distribution of  $X_2$ , conditional on the margins of the table. We have implemented Thomas' algorithm in EXAX2 [14] and illustrate the calculation of the confidence intervals with several examples.

We first consider the Holcombe and Phipps compound D fry mortality data. The output appears in Figure XIII.1. The first page of the output defines the odds ratio explicitly in terms of the order of the groups and the order of the response categories. Subsequent pages present the individual 2 x 2 tables to compare treatment groups with the control group on a pairwise basis, a point estimate and confidence interval on the odds ratio and the one sided significance level of Fisher's exact test for equality of the two response probabilities.

It should be noted that the quantities ALPHAL and ALPHAU, which specify the probability inequalities governing the upper and lower

confidence limits are under the control of the user. They can be adjusted to yield one sided upper or lower confidence bounds in place of two sided intervals or to account for simultaneity by means of Bonferroni's method.

In the present example individual 95% two-sided confidence intervals are calculated on the odds ratios of each treatment group with the control group. The conclusions are similar to those arrived at with the asymptotic intervals. Namely the response rates in groups 2 and 3 cannot be distinguished from that in the control group. The fry mortality rate in group 4 is marginally worse than the control group rate. The lower confidence limit of 0.13 suggests that the fry mortality rate in group 4 could be substantially worse than the control rate. The upper confidence limit of 1.27 is not too far removed from 1.0. This implies that the fry mortality rate in group 4 is not significantly different from that in group 1 at  $\alpha$  = 0.05 but would be significant at a slightly higher  $\alpha$  -level. ( $\alpha$  = 0.07 suffices here). The odds ratios comparing the responses rates in groups 5, 6 to that in the control group are very small and the upper bounds are very small. There is thus strong evidence that these groups have significantly higher fry mortality rates than the control group and substantilly so.

The large widths of the confidence intervals imply that the odds ratios cannot be determined very precisely.

We next consider the Holcombe and Phipps compound D embryo mortality data. The output format is the same as that for the fry mortality data and appears in Figure XIII.2. We see that none of the treatment group response rates are significantly different from the control group rate. The confidence intervals all straddle 1 and so the treatment group response rates cannot be distinguished from the control group response rate. This is in conformance with the results of our preliminary analyses.

The previous discussion pertained to construction of exact, small sample confidence intervals on the odds ratio

$$\rho = \frac{p_1/q_1}{p_2/q_2}$$

However  $\rho$  has no direct physical interpretation. A parameter such as

$$\theta = p_2/p_1$$

is more physically meaningful. How can we construct confidence intervals on  $\theta$  based on the confidence intervals we have constructed on  $\rho$ ? We can express  $\theta$  in terms of  $\rho$  and  $\textbf{p}_1$ . Namely

$$\theta = \frac{1}{\rho + p_1(1 - \rho)}$$

If  $\rho < 1$  then  $\theta$  decreases as  $p_1$  increases from 0 to 1. If  $\rho > 1$  then  $\theta$  increases as  $p_1$  increases from 0 to 1. For fixed  $p_1$ ,  $\theta$  decreases as  $\rho$  increases from 0 to  $\infty$ 

Suppose  $(\rho, \tilde{\rho})$  is a confidence interval on  $\rho$  and suppose  $(\tilde{\rho_1}, \tilde{\rho_1})$  is a confidence interval on  $p_1$ .

Then a conservative confidence interval on  $\theta$  is

$$\left(\frac{1}{\stackrel{\sim}{\rho} + p_{1}^{*}(1-\stackrel{\sim}{\rho})}, \frac{1}{\stackrel{\rho}{\stackrel{\sim}{\rho} + p_{1}^{**}(1-\stackrel{\rho}{\stackrel{\sim}{\rho}})}\right)$$

Where

$$\mathbf{p}_{1}^{\star} = \begin{cases} \mathbf{p}_{1} \\ \mathbf{p}_{1} \\ \mathbf{p}_{1} \end{cases} \quad \mathbf{as} \quad \mathbf{p}_{1}^{\times} \\ <1 \end{cases} \qquad \mathbf{p}_{1}^{\star\star} = \begin{cases} \mathbf{p}_{1} \\ \mathbf{p}_{1} \\ \mathbf{p}_{1} \end{cases} \quad \mathbf{as} \quad \mathbf{p}_{1}^{\times} \\ >1 \end{cases}$$

The confidence interval on the odds ratio  $\rho$  comes from the EXAX2 program output. Confidence intervals on p can be calculated by the Pearson-Clopper method. Namely if

$$\hat{p}_{1} = \frac{x_{1}}{N_{1}} \text{ then}$$

$$\hat{p}_{1} = \left\{ 1 + \frac{N_{1} - X_{1} + 1}{X_{1}} F(2N_{1} - 2X_{1} + 2, 2X_{1}; 1 - \alpha/2) \right\}^{-1}$$

$$= 0 \text{ if } X_{1} = 0$$

$$\overset{\sim}{\mathbf{p}}_{1} = \left\{ 1 + \frac{N_{1} - X_{1}}{X_{1} + 1} \frac{1}{F(2X_{1} + 2, 2N_{1} - 2X_{1}; 1 - \alpha/2)} \right\}^{-1}$$
  
= 1 if  $X_{1} = N_{1}$ 

These confidence intervals are given in chart form. See for example Box, Hunter, and Hunter [40], pages 642, 643 or Dixon and Massey [13], pages 501-504. We apply this conservative procedure to the Holcombe and Phipps compound D fry mortality data and compare the results with those calculated by the asymptotic approach.

In the control group  $X_1 = 6$ ,  $N_1 = 100$ . Thus  $\hat{p}_1 = 0.06$ . A 99% 2 sided confidence interval on  $p_1$  is, from the Pearson-Clopper charts entered at  $\hat{p} = 0.06$ , n = 100,  $(0.02, 0.15) \equiv (p_1, \hat{p}_1)$ .

The 95% confidence intervals on the odds ratio  $\rho$ , namely ( $\rho$ ,  $\overset{\sim}{\rho}$ ) are

Group 2 vs Control(0.2018, 2.5238)Group 3 vs Control(0.2018, 2.5238)Group 4 vs Control(0.1278, 1.2743)Group 5 vs Control(0.0055, 0.0467)Group 6 vs Control(0, 0.0027)

Combining these results as discussed previously, we obtain:

Groups 2, 3 vs Control

Since 
$$\tilde{\rho} > 1$$
,  $\rho < 1$  we have  
 $\vartheta = \frac{1}{\tilde{\rho} + \vartheta_1 (1 - \tilde{\rho})} = \frac{1}{2.5238 + 0.02(1 - 2.5238)} = 0.401$ 

$$\hat{\theta} = \frac{1}{\rho + p_1(1 - \rho)} = \frac{1}{0.2018 + 0.02(1 - 0.2018)} = 4.592$$

This would be a conservative 100(1 - .05 - .01) = 94% confidence interval.

The corresponding 95% confidence interval based on asymptotic normal theory (0.480, 3.703). We see that the two intervals are qualitatively similar but that the conservative interval is longer, as would be expected.

We now compare the conservative small sample with the approximate large sample intervals for comparisons of groups 4, 5, 6 with the control group. The calculations proceed analagously.

			Conservative, Interval		$\frac{\text{Approximate, Large Sample}}{\text{Interval on }\theta}$
Groups 2,3	vs	Control	. (0.401,	4.592)	(0.480, 3.703)
Group 4	vs	Control	. (0.788,	6.885)	(0.857, 5.474)
Group 5	vs	Contro1	. (5.272,	39.856)	(6.011, 28.859)
Group 6	vs	Control	(6.566,	50)	large sample interval not calculated

We see that the two sets of intervals are qualitatively similar however the conservative, small sample intervals are 30%-51% longer than the corresponding asymptotic intervals.

An alternative approximation can be used to calculate conservative confidence intervals on  $\theta \equiv p_2/p_1$ . Consider again the 2 x 2 table.

	CONTROL	GROUP 2
DEAD	x <sub>1</sub>	x <sub>2</sub>
LIVE	Y <sub>1</sub>	<sup>Ү</sup> 2
	N <sub>1</sub>	N <sub>2</sub>

Let  $p_1$ ,  $p_2$  denote the probabilities of death in groups 1, 2 respectively. We wish to construct a confidence interval on  $p_2/p_1 \equiv \theta$ .

Now N<sub>1</sub>, N<sub>2</sub> were fixed by the experimenter. Let  $r \equiv N_2/N_1$ . Suppose we assume the fiction that  $N_1 \sim P_0(\lambda)$ ,  $N_2 \sim P_0(r\lambda)$  and that  $N_1$ ,  $N_2$  in the data are realizations of these two independent random variables. Then  $X_1$ ,  $X_2$ ,  $Y_1$ ,  $Y_2$  can be treated as independent Poisson random variables with means  $p_1\lambda$ ,  $p_2\lambda$ ,  $q_1\lambda$ ,  $q_2r\lambda$  respectively. Confidence intervals on  $p_2/p_1$  can be constructed by methods like those discussed in connection with the Poisson approximation approach, (method 3). Namely

$$P \frac{X_2}{X_1 + 1} \frac{1}{F(2X_1 + 2, 2X_2; 1 - \alpha_1)} \frac{1}{r} \le \frac{P_2}{P_1} \le \frac{X_2 + 1}{X_1} F(2X_2 + 2, 2X_1; 1 - \alpha_2) \frac{1}{r} \ge 1 - \alpha$$

where  $\alpha_1 + \alpha_2 = \alpha$ . Now these confidence intervals are conservative because we are introducing additional variability by assuming that N<sub>1</sub>, N<sub>2</sub> are random variables rather than fixed constants. The variances of X<sub>1</sub>, X<sub>2</sub> are inflated from  $N_1p_1q_1$ ,  $N_2p_2q_2$  to  $N_1p_1$ ,  $N_2p_2$  by this assumption. Thus the greater are  $p_1$ ,  $p_2$ , the more conservative this procedure will be.

We illustrate the application of these intervals with the Holcombe and Phipps compound D fry mortality data and the DeFoe 1, 1, 2, trichloroethane fry mortality data.

First consider the Holcombe and Phipps compound D fry mortality data. The comparisons of Groups 2, 3, 4 vs Control, based on the Poisson approximation, are quite similar to the conservative small sample confidence intervals discussed earlier in this subsection.

Now consider comparisons of Groups 5, 6 with the Control group.

<u>Group 5 vs Control</u>:  $X_5 = 79$ ,  $X_1 = 6$ ,  $N_5 = N_1 = 100$ ,  $\alpha_1 = \alpha_2 = 0.025$ 

Thus

$$\left( \frac{79}{7} \frac{1}{F(14, 158; .975)}, \frac{80}{6} F(160, 12; .975) \right) = \left( \frac{79}{7} (1.94)^{-1}, \frac{80}{6} (2.77) \right) = (5.82, 36.93)$$

is an approximate 95% confidence interval on  $p_5/p_1$ .

<u>Group 6 vs Control</u>:  $X_6 = 100$ ,  $X_1 = 6$ ,  $N_6 = N_1 = 100$ ,  $\alpha_1 = \alpha_2 = 0.025$ 

Thus

$$\left( \frac{100}{7} \frac{1}{F(14, 200; .975)}, \frac{101}{6} F(202, 12; .975) \right) = \left( \frac{100}{7} \frac{1}{1.79}, \frac{101}{6} F(202, 12; .975) \right)$$

is an approximate 95% confidence interval on  $p_6/p_1$ .

These intervals compare with the conservative, small sample intervals calculated earlier as follows:

	Conservative, Small Sample	Approximate Poisson
Group 5 vs Control	(5.27, 39.86)	(5.82, 36.93)
Group 6 vs Control	(6.57, 50)	(7.98, 46.29)

These intervals are seen to be quite similar.

We now consider the DeFoe 1, 1, 2, trichloroethane data and calculate approximate confidence intervals to compare Groups 5, 6 to the control group. Since  $X_1 = 0$  we can only calculate lower confidence bounds.

Group 5 vs Control: 
$$X_5 = 9$$
,  $X_1 = 0$ ,  $N_5 = N_1 = 40$ ,  $\alpha_1 = 0.05$ ,  $\alpha_2 = 0$ .

Thus

$$\frac{9}{1} \frac{1}{F(2, 20; .95)} = \frac{9}{3.49} = 2.58$$

is a 95% lower confidence bound on  $p_5/p_1$ .

<u>Group 6 vs Control</u>  $X_6 = 40$ ,  $X_1 = 0$ ,  $N_6 = N_1 = 40$ ,  $\alpha_1 = 0.05$ ,  $\alpha_2 = 0$ . Thus

$$\frac{40}{1} \frac{1}{F(2, 82; .95)} = \frac{40}{3.12} = 12.82$$

is a 95% lower confidence bound on  $p_6/p_1$ .

Thus there is strong statistical evidence that the response rates in groups 5 and 6 are substantially greater than that in the control group. The response rate in group 5 is at least  $2\frac{1}{2}$  times that in the control group.

### Method 3 Poisson Approximation D.

We now consider method 3 for placing confidence intervals on ratios of parameters. This method is based on the Poisson approximation to the binomial distribution and so requires that each p be less than 0.1 or that each p be greater than 0.9 in order that the Poisson approximation be reasonably accurate. Operationally, we will use this approximation if each  $\hat{p}$  is less than 0.1 or if each  $\hat{p}$  is greater than 0.9. The prototype situation is

	Control	Group 2
Dead	x <sub>1</sub>	x <sub>2</sub>
Live	$N_1 - X_1$	$N_2 - X_2$
	N <sub>1</sub>	N <sub>2</sub>

Let  $p_1$ ,  $p_2$  denote the response probabilities in groups 1, 2 respectively. We wish to construct 1 -  $\alpha$  confidence intervals on  $p_2/p_1$ .

Let 
$$\lambda_1 = N_1 p_1$$
,  $\lambda_2 = N_2 p_2$ .

Then if  $p_1 < .1$ ,  $p_2 < .1 \times \sqrt{p_0}(\lambda_1)$ ,  $\chi_2 \sqrt{p_0}(\lambda_2)$ . We can thus pose the problem as one of placing confidence intervals on the ratio of two Poisson means. If  $p_1 > .9$ ,  $p_2 > .9$  it is probably of more interest to place a confidence interval on the ratio  $q_2/q_1$ , where  $q_1 \equiv 1 - p_1$ ,  $q_2 = 1 - p_2$ . We are then back in the above situation.

Nelson [24] shows that a 1 -  $\alpha$  confidence interval on  $\lambda_2/\lambda_1$  is

$$\left[\frac{x_2}{x_1+1} \frac{1}{F(2x_1+2, 2x_2; 1-\alpha_1)}, \frac{x_2+1}{x_1} F(2x_2+2, 2x_1; 1-\alpha_2)\right]$$

where  $F(v_1, v_2; \gamma)$  represents the upper  $\gamma$  point of the F-distribution with d.f.  $v_1, v_2$  and  $\alpha_1 + \alpha_2 = \alpha$ . Now

$$\lambda_2/\lambda_1 \equiv (N_2P_2)/(N_1P_1) = (N_2/N_1)(P_2/P_1)$$

Thus multiplying the above confidence bounds by the factor  $\rm N_1/N_2$  yields confidence bounds on  $\rm p_2/p_1$ . Namely

$$\begin{bmatrix} \frac{X_2}{X_1 + 1} & \frac{1}{F(2X_1 + 2, 2X_2; 1 - \alpha_1)} & \frac{N_1}{N_2}, \frac{X_2 + 1}{X_1} & F(2X_2 + 2, 2X_1; 1 - \alpha_2)\frac{N_1}{N_2} \end{bmatrix}$$

is a 1 -  $\alpha$  confidence interval on  $p_2/p_1$ . Often we take  $\alpha_1$ ,  $\alpha_2$  to be  $\alpha/2$ . However for one sided confidence intervals we take  $\alpha_1 = \alpha$ ,  $\alpha_2 = 0$  or  $\alpha_1 = 0$ ,  $\alpha_2 = \alpha$ .

If  $X_1 = 0$  or if  $X_2 = 0$  we have only one sided information about  $p_1$ ,  $p_2$  respectively. Thus we can only construct one sided confidence bounds on their ratio. Namely if  $X_1 > 0$ ,  $X_2 = 0$  then set the lower confidence bound equal to 0 and upper confidence bound on  $p_2/p_1$  becomes

$$\frac{N_1}{N_2} \frac{1}{X_1} F(2, 2X_1; 1 - \alpha) \quad \text{if } X_2 = 0, X_1 > 0.$$

If  $X_1 = 0$ ,  $X_2 > 0$ , then we can only get a lower bound on  $p_2/p_1$ . Set the upper bound equal to  $\infty$  and the lower confidence bound becomes

$$X_2 \frac{1}{F(2, 2X_2; 1 - \alpha)} \frac{N_1}{N_2}$$
 if  $X_1 = 0, X_2 > 0$ 

If  $X_1 = 0$ ,  $X_2 = 0$  the problem is indeterminate.

Nelson [24] presents charts which facilitate the construction of two sided 90%, 95% or 99% confidence intervals on  $\lambda_2/\lambda_1$ . However his charts do not apply for the situation when  $X_1 = 0$  or  $X_2 = 0$ . In fact they effectively apply only when  $0.1 \le X_2/X_1 \le 10$ . The charts are shown in in Figures XIII.3, XIII.4, XIII.5.

To use the Nelson charts

- 1. Enter the value of  $X_2/X_1$  on the horizontal axis.
- 2. Go up to the curve labelled with the X<sub>1</sub> value. (There are two sets of curves, corresponding to upper and lower confidence limits).
- 3. Read the upper and lower limits on the vertical scale.
- 4. Multiply the resulting limits by the ratio  $N_1/N_2$ .

We illustrate the use of this Poisson based procedure on several sets of data. First we consider the Holcombe and Phipps compound D fry mortality data. We pool responses across tanks within groups.

GROUP 6

DEAD	6	8	8	13	79	100
LIVE	94	92	92	87	21	0
	100	100	100	100	100	100

CONTROL GROUP 2 GROUP 3 GROUP 4 GROUP 5

We compare various treatment groups with the control group. We will calculate two sided 95 percent, nonsimultaneous confidence intervals. Groups 2 and 3 appear to have response probabilities around 0.10 and group 4 does not seem to be too much beyond this level. We thus stretch our criterion a bit and calculate confidence intervals to compare groups 2, 3, 4 with the control groups.

Note that we could modify the confidence intervals for simultaneity by using Bonferroni's inequality. Groups 2, 3 vs Control:

$$x_2 = 8, x_1 = 6, x_2 = x_1 = 100, \alpha_1 = \alpha_2 = .025$$

Thus

$$\left(\frac{8}{7} \quad \frac{1}{F(14, \ 16; \ .975)}, \frac{9}{6} F(18, \ 12; \ .975) \in \frac{8}{7}(2.83)^{-1}, \frac{9}{6}(3.11)\right) = (0.404, \ 4.665).$$

Group 4 vs Control:

$$x_4 = 13, x_1 = 6, x_2 = x_1 = 100, \alpha_1 = \alpha_2 = .025$$

Thus

$$\left( \frac{13}{7} \quad \frac{1}{F(14, \ 26; \ .975)} , \frac{14}{6} F(28, \ 12; \ .975) = \frac{13}{7} (2.45)^{-1} , \frac{14}{6} (2.98) \right) =$$

$$(0.76, \ 6.95)$$

Comparing the confidence intervals obtained by methods 1, 2, 3 we see that

	Asymptotic	Conservative Small Sample	Poisson Approximation
Group 2, 3 vs Control	(.480, 3.703)	(.401, 4.592)	(.404, 4.665)
Group 4 vs Control	(.857, 5.474)	(.788, 6.885)	(.76, 6.95)

Thus the asymptotic intervals are shorter than either of the small sample intervals. The small sample intervals are thus more conservative.

We next consider the DeFoe compound C fry mortality data. We again pool across tanks within groups.

	CONTROL	GROUP 2	GROUP 3	GROUP 4	GROUP 5	GROUP 6
DEAD	0	0	2	1	9	40
LIVE	40	40	38	40	31	0
	40	40	40	41	40	40

Since there are zero responses in the control group (i.e.  $X_1 = 0$ ), we can only calcualte lower confidence bounds.

<u>Group 2 vs Control</u>: Since Group 2 has 0 responses also, the situation is indeterminate.

<u>Group 3 vs Control</u>: Choose  $\alpha_1 = .05$ ,  $\alpha_2 = 0$   $X_3 = 2$ ,  $X_1 = 0$ ,  $N_3 = N_1 = 40$ Then

$$\frac{2}{1} \frac{1}{F(2, 4; .95)} = \frac{2}{6.94} = 0.29$$

is a 95 percent lower confidence bound on  $p_3/p_1$ . Thus there is no statistical evidence, at the  $\alpha$  = .05 level, that  $p_3 > p_1$ .

<u>Group 4 vs Control</u>:  $X_4 = 1$ ,  $X_1 = 0$   $N_4 = 41$ ,  $N_1 = 40$  Choose  $\alpha_1 = .05$ ,  $\alpha_2 = 0$ .

Then

$$\frac{1}{1} \quad \frac{1}{F(2, 2; .95)} \quad \frac{40}{41} = \frac{40}{41} \quad \frac{1}{19.0} = 0.05$$

is a 95 percent confidence bound on  $p_4/p_1$ . Thus there is no statistical evidence, at the  $\alpha$  = 0.05 level, that  $p_4^{>}p_1$ .

In general the confidence intervals that we have calculated are too wide to determine the ratios of the various probabilities with much precision. We must conclude that the data are not sufficient to estimate these ratios very precisely without placing further structure on the problem. One way of imposing such further structure will be discussed in the following sections.

## CONTRUIND D FRY MACTALITY DATA AND PUNCS HOLCONE EXXX2 OUTPUT

I VE GACH OF THE OTHER PETATMENTS. EXACT 95. 10 % NUMERANDARY CONFIDENCE INTERVALS ON 0015 PATED OF FRE THE ALGUETUM USED IS A MOULTED VIRSIGH EE AS BE GIVEN FY THOMAS (1971) APPLIED STAT. PERIOS-TID. FOR EACH 2 X 2 TABLE Istimates can be interface only gere the table sectom.

ICTAL N2 N2 N1		
С В В С В С С В С С В С	PSI = (P102)/19201) ESTIMATE OF PSI = 96/AD ESTIMATE UF PI = 8/NI	0F P.2 = D/N2 0F 01 = A/N1 0F 02 = C/N2
• • • • • • • • • • • • • • • • • • •	PSI = (P102)/1920 ESTEMATE OF PSI = ESTEMATE OF PI =	ESTIMATE OF P.2 = ESTIMATE OF Q1 = ESTIMATE OF Q2 =

EXAX2 output from calculation of exact, small sample confidence intervals on the odds ratios between treatments and control Figure XIII.1

# EXALS OUTPUT NOLCOMBE AND PARAS CONTINUED FAT MATTHETY DATA

P

i

at Pral = 0.025       at PhAU = 0.00         Tufat 2       92.000         Tufat 2       41.05         Tufat 3       41.05         Tufat 4       41.05         Tufat 5       41.05         Tufat 4       41.05         Tufat 5       41.05         Tufat 4       5         Tufat 5       5         Tufat 5       5		ALPH	STIMATE OF (PIOT) (P201) = 0.734042525 95.00 4 Non-Simultave UNDER INTERIAL = ( 0.201835632, 2.52381802) PRLOABILITY OF 1ASLE UNDER HYPOTHEES OF INDEPLADENCE = 0.130140 FRLOABILITY OF 1ASLE UNDER HYPOTHEES OF INTERVAL DA DDDS RATIO OF TRT' 1 V5 TRT 4	5/17 1/1	ESIMPLE = U E (P I C) / (F C U) = (S A C) C P C A B
---	--	------	---	-------------	---

## Campound o FAY MARTHUT DATA Saran C NOLCONOT AND EXXX3 OUTPUT

I VS TPT EXACT 95+FO % NUN-SIMULTAILEDR'S CUNFIDENCE INTERVAL ON CUDS PATIC OF TRT

ø I VS TRT 200.000 ANNER CODSTRATED ESTIMATE IS 25RO, ONLY AN UPPER CONFIDENCE BOUND CAN BE OBTAINED. TATES FXACI 95.00 & NGN-SIMULTANEGUS CINFIDENCE INTEPVAL ON DODS RATIO OF TRT , GRAND TOTAL = 0,04673290251 -----35,000 ESIIWATE OF (PIJ2)/(2201) = C.0149474384 95.00 × NUX-SIMULTANEOUS CUMPIUENCE INTERVAL = ( C.3054938793, PROBACILITY CE (AML UNDER MYFULTELIS VE INDEPENDENCE = 0.00000 0EAU ≂ 1010-000 100-000 100.000 100.000 . . . . . . . 115.0CO , TUTAL ALTVE DEAD 94.000 0.6 ALPHAL = 0.025 ALPHAU = 0.025 AL PHAU = 0.050 4 ALIVE TRT 1 TRT 5 ALPHAL = 0.0TUTAL

Figure XIII.1 Continued

200-000

. GRAND TOTAL =

106.000

TLTAL ALIVE = 44.000 . TOTAL

6,00269085171

ES 1 4 6 15 0F (P 102) ( 1220 1) = 000 S 45 00 2 NCN-S MULTANE OUS CONFIGENCE INTERVAL = 1 0.0 PRO3AMILITY OF TABLE UNDER HYPOTHESIS OF INDEPENSENCE = 0.00000

EXAX2 output from calculation of exact, small sample confidence intervals on treatmentcontrol odds ratio Figure XIII.2

er, 2 4 I VS TRT I VS TRT I VS TRT 400.000 4 20.000 40C.0CO PTNA 94.00 T Now-Standlations COVEDDINGS INTERNAL CN 1005 RATIS (F. 191 46.00 & NUN-SIMULTANEOUS CONFIDENCE INTERVAL ON CONS RATIO OF TRT ETACH 95.00 % NUN-SIMULIANEUUS CONFIDENCE INTERVAL DN CODS PATIO GF TRT YPACAR u , GRAND TOTAL = 140.000 , GRAND TUTAL = i . GPAND TOTAL -1.20054032 1,54032756 141522560 EMBRYD • 4 . . . . 133.000 1 144.000 ÷ ESTIMETE DF (P102)/(P201) = 1.17043945 95.401 1 4.4-514JLIANECUS CUMFIJ2NUE INTERVAL = 1 0.755745113 . PHERRELLITY OF LABLE UNDER HYRUTHELLS UF INDERENDENCE = 0.2526219 ... ESTEARTE DE (F102)/(P221) = 0.642.0644 V.403 Num-Simultineuus (CRAFLUNU, INFERVEL = ( 0.540400330 PRGAPMILITY DE FABLLUNDER HYPLECIIS DE INDEPENDENFE = 0.22429 ۵ ESTWATE OF (FLU) / (PEU) = 1,000 0000 95.00 + NOM-1/01 (ANLUS CUMEDURU) - 1.154VAL = ( 0.649208069 -PROMENTIT (FLUMEDURU) MYPUTHULLO OF INGEPENDENCE 3.00.54174. Gunedhaa . ŗ # 0040 70-000 - 200-000 63-000 - 200-000 101AL 260.0000 240.0000 701AL 200.000 200.000 đ н ł 0120 UEAD ÷ DEAU AND PHIME . 267.9r0 . TOTAL 252.000 , TOTAL , TUTAL UEAU 76.000 7°.000 9, 20 70,000 70,000 ; HULONOE 250.000 200 8 = 17110 TV ALPHAD = 0.025 ≠LРНАЦ = ^•3с5 В**LРНАU =** 0.025 ALTVC 130-000 122-000 . >LIV5 130-000 137-000 ALTV. 132.000 130.000 EXAX2 OUTPUT łI b 11 ALIVE ALIVE ALTVE 1 2724 ALUMAL = U.UZE 1000 \*LPHAL = 7. 75 -4,44 -- 3 TOIAL TCIAL TUTAL 151 151 22 

٠. .

•

-----

•

## Figure XIII.2 Continued

1

400.000

, GRAND TOTAL =

146.000

a de la comercia a comercia a comercia

254.000 . TOTAL DEAD

ŧŧ

ALIVE

TUTAL

:

ESILATE UF (P122)/(P201) = 0.876542493 95.00 % NON-SIMULTANE OUS CUNFIDENCE INTERVAL = ( 0.572517395 , 1.34771442 PROCADILITY OF TABLE JUNDER HYPUTHESIS OF INDEPENDENCL = 0.30182

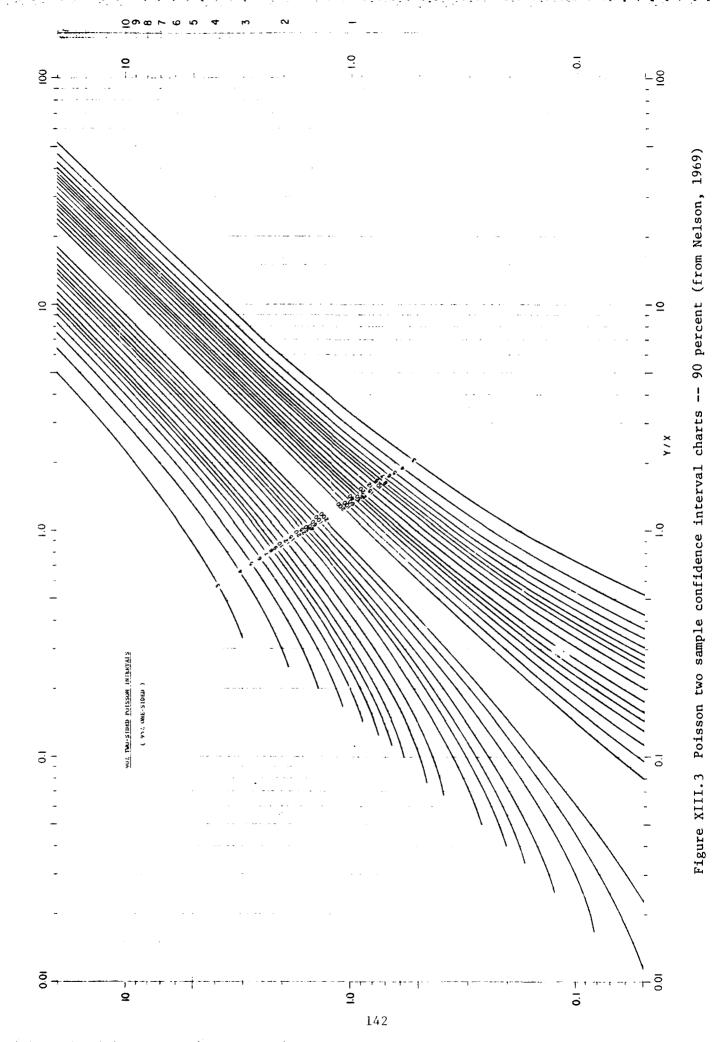
ţ 9. c. I VS TET 1 VS TRT 403.000 EXACT 94.00 % NUN-SIMULTANFULL CONFIDENCE INTRAL ON CODS RATIC OF THE EXACT 95.00 & NUN-SIMULIANE JUS CONFIDENCE INTERVAL ON DDDS RATIO OF THT 137.000 , GRAND TUTAL = 1 .... 1.65059376 ) · · · · · · · · ECTIMATE TE (F132)/(P241) = 1.C6PBHE-U 95-00 %.NUN-SIMULTANEGUS CUNFIUENCE THTERVAL #1 0.652345023. PAUJAHILITY DF TABLE UNDEA HYPUTHESIS. UF INDERENDENCE = .0.441659. 101AL 200-000 707AL 200-000 263.044 , TUTAL DEAD 🚡 : . . 91 AU 79-000 67-000 DE#D 70-000 76-000 ALPHAL = 0.025 ALPHAU = 0.025 ALPHAL = C.(25 ALPHAU = C.C25 ų ALIVE -10 TET ... 1 LUTAL 181 181

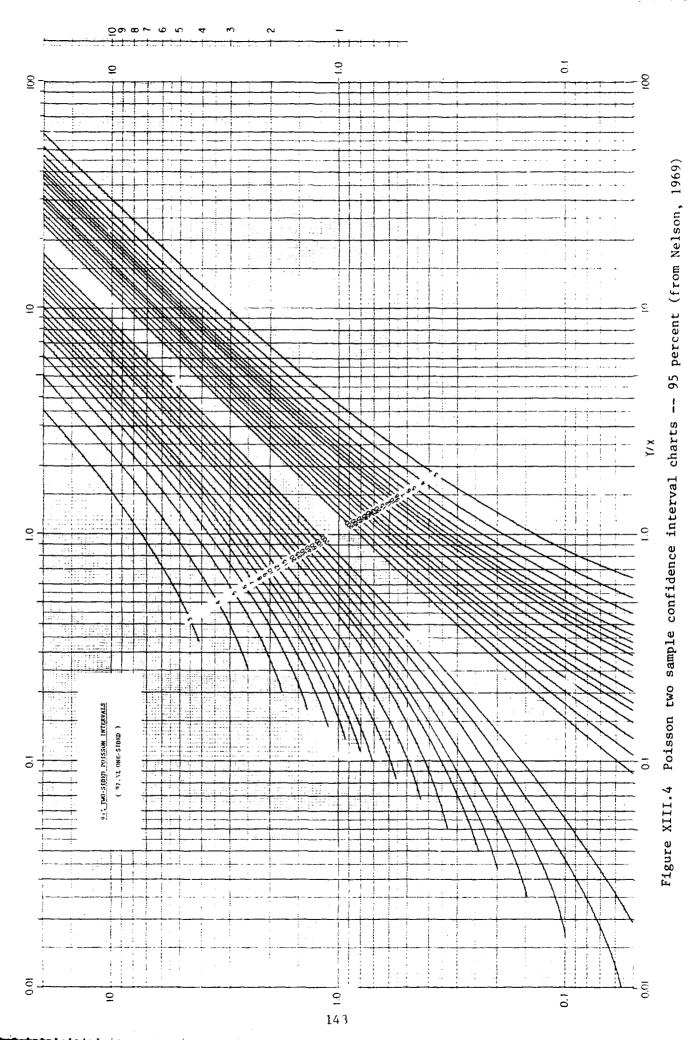
COMPAUND & ENGAYO MONTHLITY DATA

Roccord

DUTPUT

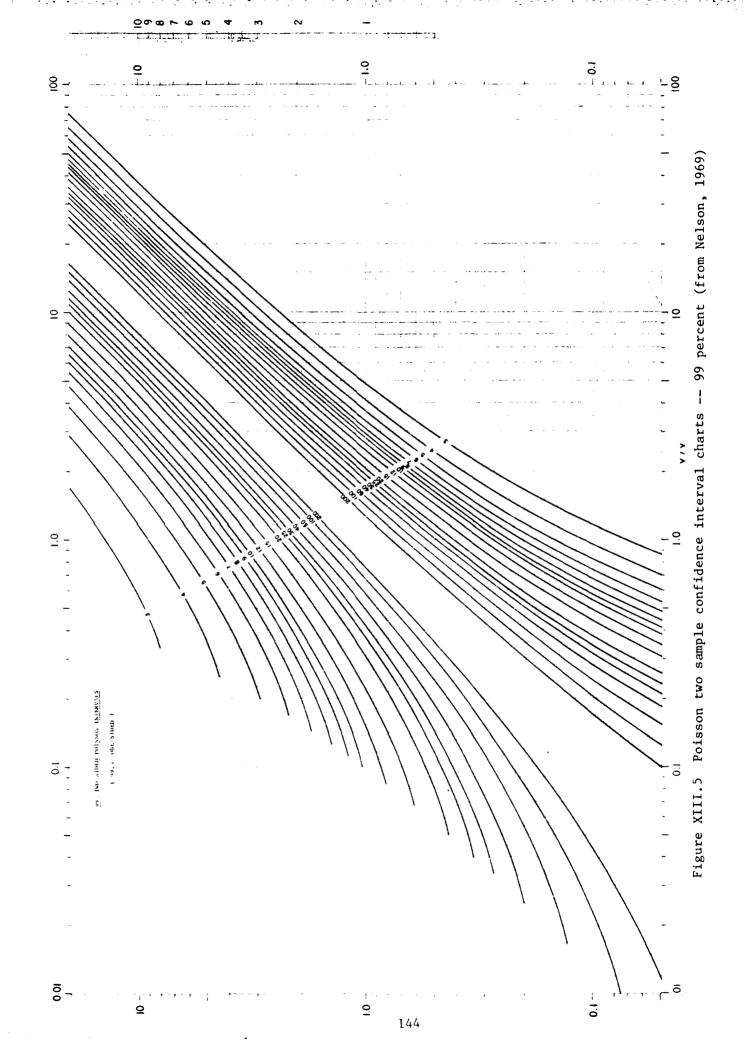
EXAS





۰. ۱

.



## XIV. DOSE RESPONSE CURVE ESTIMATION -- PROBIT ANALYSIS

### A. Introduction. Dose Response Estimation vs Hypothesis Testing

An alternative approach to estimating acceptable concentration levels is based on fitting dose response models to the data and estimating that concentration,  $C_L$ , which results in an increase of at most L in the response rate over and above background level. The dose response curve formulation is pictured schematically in Figure XIV.1. The problem of determining a safe concentration has been transformed from a testing problem (determine which response rates are significantly different than the control rate) to an estimation problem (calculate a lower confidence bound on  $C_T$ ).

The two formulations are conceptually different and lead to different implications. With the classical hypothesis testing formulation the larger and more precise the experiment the more powerful will be the hypothesis test. Thus lower concentration levels will be found significantly different from the control group and so the acceptable concentration will be <u>decreased</u>. By contrast, with the dose response curve estimation formulation the larger and more precise the experiment, the higher will be the lower confidence bound on  $C_L$  and so the acceptable concentration seems more natural to us for two reasons.

- 1. There is no need to specify rigid sample size requirements in the protocol. People could present any level of evidence regarding safe concentrations that they wish. The more extensive the experiment, the higher will be the lower confidence bound on  $C_r$ .
- 2. An investigator conducting toxicity tests in support of petitions to the EPA for discharge permits is induced to carry out more extensive and more precise experimentation by the economics of the situation. He is rewarded for his efforts by demonstrating a greater safe concentration.

OPINION: We feel that increased emphasis should be placed on the fitting and use of dose response curve models in the design of and analysis of data from aquatic toxicity tests.

It should be noted that just because we define  $C_L$  in terms of the concentration associated with an increase in response rate of L units over background does not mean that we consider killing 100L percent of the fish to be "acceptable". No increased mortality is really desirable. However by adopting this formulation we can argue that we are limiting our risk to an <u>upper bound</u> on L. The choice of L in a particular situation would of course need to be a biological and a regulatory decision.

We have fitted (or attempted to fit) a number of dose response models to the embryo and fry mortality data. Some of these models are standard while others are nonstandard. Among the standard models fitted are the probit model (Finney [11]) with either logarithmic or untransformed concentration and the logit model with either logarithmic or untransformed concentration levels. Both of these models classically account for background variation by means of Abbott's correction. For example a probit model with Abbott's correction might state

$$p(\text{conc}) = p_0 + (1 - p_0) \Phi (\beta_0 + \beta_1 \ln (\text{conc}))$$

where  $p_0$ , p(conc) are the response rates at the control and at conc respectively,  $\Phi(\cdot)$  is the normal c.d.f., and  $p_0$ ,  $\beta_0$ ,  $\beta_1$  are unknown parameters to be estimated from the model fit. Such a probit model can easily be fitted to the data using SAS PROC PROBIT [12]. The 1979 version of the BMDP package [27] contains a stepwise logistic regression program.

Among nonstandard dose response models tried are a nonstandard probit type model and a "nonparametric" dose response model. The nonstandard probit type model differs from the standard model in the way it handles background response. One version can be written as

$$p(conc) = \Phi (\alpha_0 + \alpha_1 \ln (conc + c))$$

where p(conc) is the response rate at conc, c accounts for the background response, and  $\alpha_0$ ,  $\alpha_1$ , c are unknown parameters to be estimated from the model fit. A criticism of Abbott's correction is that it tacitly assumes that background related response and toxicant related response are due to different and <u>independent</u> mechanisms. The nonstandard model assumes that background related responses and toxicant related responses are due to similar mechanisms and thus that background acts like an incremental toxicant level c. Which (if either) model is more appropriate in a given situation depends on how well they fit the data and on biological judgement. The nonstandard probit model and a large family of other standard and nonstandard dose response models can be fitted by the use of nonlinear regression programs such as SAS PROC NLIN [12] and BMDP programs BMDP3R, BMDPAR[27] (program versions 1977 or later).

We have developed a "nonparametric" dose response model that differs from the more usual parametric models in a number of ways.

- 1. There is no need to make strong parametric assumptions about the form of the dose response model.
- 2. There is no need to be concerned with transformations of the concentration levels.
- 3. There is no need to worry about the parametric form used to correct for background variation.

4. Exact, small sample theory is used to construct conservative lower bounds on safe concentration.

We have developed a special purpose computer program to carry out such nonparametric dose response analyses. It is described in detail in Feder and Sherrill [41], which is included as an appendix to section XVI.

We now illustrate inferences about safe concentrations based on the various dose response models discussed above.

## B. Probit Analysis Using SAS PROC PROBIT

In this subsection we fit probit models to the fry mortality data from the DeFoe test on compound C and from the Holcombe and Phipps test on compound D. Although we do not adjust the data for tank to tank heterogeneity, the same analyses can be carried out after such adjustments have been made.

We first consider the DeFoe data. The basic data are listed in Figure XIV.2. There are two tanks per treatment group. Concentration values for each group (in units of  $\mu g/liter$ ) have been determined as average values over all determinations and over all tanks within each group. These are denoted as CONCMEAN. Other variables of importance are

DEADESUM  $\equiv$  # dead embryos in the tanks after hatch. (After about 5 days). DEADFSUM  $\equiv$  # dead fry after 32 days.

PRPDEADE, PRPDEADF Ξ proportions of dead embryos and fry respectively. LOGCONC Ξ log<sub>10</sub>(CONC)

Note that the measured concentration in the control group is <u>not</u> zero here and that no fry mortality has occurred in the control group. It is unclear from preliminary plots of proportions of dead fry vs arithmetic and logarithmic concentration (not shown) whether a probit model would better be fitted to arithmetic or to logarithmic concentration. We will try both fits and compare them.

We first fit a standard probit model using arithmatic concentration. The specific model fitted is

 $p(CONC) = c + (1 - c) \Phi(\beta_0 - 5 + \beta_1 CONC)$ 

where  $\Phi(\cdot)$  is the standard normal c.d.f.,  $\beta_0$  and  $\beta_1$  are unknown

model parameters that characterize the shape of the response curve, and c is the unknown model parameter that specifies the background rate. (The quantity 5 in the argument of  $\Phi(\cdot)$  is due to probit convention). We fit this model to the data with SAS PROC PROBIT by maximum likelihood estimation. The output resulting from this fit appears in Figures XIV.3 - XIV.6. The interpretation of this output is as follows:

) Summary of the maximum likelihood iteration process.

Intercept, Slope  $\leftrightarrow \beta_0$ ,  $\beta_1$  respectively in the probit model.

- c ↔ background rate, or threshold rate. (Note that if c goes negative at an iteration step it is set to 0.).
- MU,SIGMA ( $\mu,\sigma$ ) correspond to the mean and standard deviation of the dose-response distribution.
- $\mu$ ,  $\sigma$  are related to  $\beta_0$ ,  $\beta_1$  as  $\beta_0 = 5 - \mu/\sigma$ ,  $\beta_1 = 1/\sigma$

These relations can be verified from the entries given in the output.

2) The estimated asymptotic variance-covariance matrix of  $(\beta_0, \hat{\beta}_1, \hat{c})$ . This is based on the Fisher information's inverse.



) The estimated asymptotic variance-covariance matrix of (μ̂, (ĉ, ĉ). This is based on the inverse of the estimated Fisher information matrix.

Note: These estimated var-cov martices are the basis of the confidence interval calculations made by the program. The validity of these var-cov estimates depends on having the true state of nature and the maximum likelihood estimates interior to the parameter space. In this fit  $\hat{c} = 0.003$ with an estimated standard error of 0.03. We thus might consider dropping c from the model.

4) Chi square test for lack of fit of the probit model.

Degrees of freedom  $\equiv$  number of groups - number of parameters  $\equiv 6 - 3 = 3$ .

Under the null hypothesis of no lack of fit to the model this statistic has a chi square distribution with 3 d.f.

- 5 A plot of the fitted straight line in the probit domain, with the estimated probits of the dose response rates at each concentration level in the data indicated as X's on tho plot.
  - Notes: 1. Probit  $(0) \equiv 5 + \Phi^{-1}(0) = -\infty$ . However it is plotted as 0 because that is the smallest value used. Similarly, Probit  $(1) \equiv 5 + \Phi^{-1}(1) = \infty$ , but it is plotted as 10 because that is the largest value used.
    - 2. The observed thresholds at probit values of 0 and 10 seem far away from the fitted line. The standard errors of these points are also very large, so these points are discounted when determining a probit fit. In particular,

 $Var[probit(\hat{p})] = p(1 - p)/n\phi(\Phi^{-1}(p)) \rightarrow \infty \text{ as}$ p \rightarrow 0 or 1.

Thus these points carry very little weight in the straight line fit in the probit domain.

3. The estimated background response rate has been removed from the plot. Thus estimates represent increments over background.

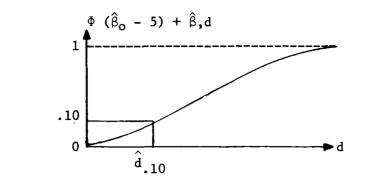
6) Plot of  $\Phi(\hat{\beta}_0 - 5 + \hat{\beta}_1 \text{ CONC})$  vs CONC.

- 7) For various percentiles of the fitted dose response curve (after adjusting for background), the point estimates of CONCMEAN are given as well as 95% lower and upper confidence bounds on these points.
  - Note: These percentiles are percentages of the population responding <u>due to the toxicant</u>, after adjusting for background effects.

The point estimates correspond to the percentiles indicated on the plots.

These lower confidence bounds are just the quantities needed to calculate confidence bounds on safe concentrations. Lower 95% two sided bounds correspond to lower 97.5% one sided bounds. Suppose we are willing to tolerate an increase in response of 10 percent due to toxicant causes. What is a lower confidence bound on safe dose?

Consider the dose response curve (adjusted for background rate).



Now  $\hat{d}_{.10} = \frac{\Phi^{-1}(.10) + 5 - \hat{\beta}_0}{\hat{\beta}_1} = 32.65$  from Figure XIV.6.

「「「ここのため」と、「「」」というというです。

The lower 97.5 percent confidence interval on  $d_{.10}$  is - 45.97. It is thus totally uninformative due to the gentle slope of the dose response curve (0.036) and relatively large standard error of the slope (0.014).

Note that the confidence bounds on the percentiles of the dose response curve are based on Fieller's theorem. See Finney [11], section 4.7 (esp.eqns (4.37), (4.38)) for details.

We now consider the chi square test statistic for goodness of fit in more detail. The chi square statistic can be used for a number of purposes. The statistic given in (4) is

CHI-SQUARE = 
$$\sum_{i=1}^{6} \frac{\left(O_{i} - N_{i}\hat{P}_{i}\right)^{2}}{N_{i}\hat{P}_{i}\hat{q}_{1}}$$

Where 0 = # observed responses in the i-th treatment group

- $N_i = #$  fish in the i-th treatment group
- $\hat{p}_i$  = estimated response probability in the i-th treatment group.

$$\hat{q}_i = 1 - \hat{p}_i$$
.

$$\hat{\mathbf{p}}_{\mathbf{i}} = \hat{\mathbf{c}} + (\mathbf{1} - \hat{\mathbf{c}}) \Phi \left( (\hat{\beta}_{\mathbf{o}} - \mathbf{5}) + \hat{\beta}_{\mathbf{1}} \mathbf{d}_{\mathbf{i}} \right) \equiv$$

$$\hat{c}$$
 +  $(1 - \hat{c}) \Phi\left(\frac{d_1 - \hat{\mu}}{\hat{\sigma}}\right)$   $i = 1, \ldots, 6.$ 

2

The values of these quantities for the 6 treatment groups are as follows:

<u>Trt Grp</u>	<u>0</u>	<u>N</u>	Ŷ	<u>NÊ</u>	NÊQ	$\frac{(0 - N\hat{P})^2}{N\hat{P}\hat{Q}}$
1	0	40	0.0100	0.3996	0.3956	0.4036
2	0	40	0.0115	0.4586	0.4534	0.4639
3	2	40	0.0154	0.6152	0.6058	3.1655
4	1	41	0.0300	1.2300	1.1931	0.0443
5	9	40	0.2397	9.5876	7.2896	0.0474
6	40	40	0.9979	39.9144	0.0854	0.0858
					$\overline{x^2}$	<i>≡</i> 4.2105

We see that this chi square statistic agrees with that calculated in (4) of the PROBIT output. We should break out the cell by cell contributions in order to ensure that a large value of chi square is not due to one or a few cells with very low expected frequency. Just one observed response in such a cell can inflate the chi square statistic tremendously. In our case this does not occur.

Note that the applicability of the asymptotic chi square approximation to the distribution of  $\chi^2$  is doubtful here due to the small expected sample sizes. Namely

			3				
NŶi	0.40	0.46	0.62	1.23	9.59	39.91	_
Nĝi	39.60	39.54	39.38	39.77	30.41	0.09	

Dixon and Massey [13] page 238 state that for the approximate asymptotic  $\chi^2$  distribution to be close "the sample size N must be sufficiently large that none of the  $F_i$ 's (i.e.  $N_i \hat{p}_i$  or  $N_i \hat{q}_i$ ) is less than 1 and not more that 20 per cent of the  $F_i$ 's are less than 5." This criterion is clearly not met in the above example.

Since no control group mortality was observed and since the estimated background rate is compatible with 0 ( $\hat{c} = 0.0031$ , stderr( $\hat{c}$ ) = .0288) it was decided to refit the model specifying that c = 0. This simplification will reduce the standard errors of estimates considerably. The fit is shown in Figure XIV.7 and the associated confidence intervals are given in Figure XIV.8. The point estimates the slope and intercept are seen to be very similar to those based on the threshold model fit in Figure XIV.3. In particular

	β <sub>o</sub>	β <sub>1</sub>	ĉ	ở(β̂ <sub>ο</sub> )	ĉ(β̂ <sub>1</sub> )	σ <b>(ĉ)</b>
Threshold Model	2.537	0.036	0.003	0.834	0.014	0.029
No Threshold Model	2.616	0.035	0 (by defn.)	0.257	0.006	O(by defn.)

We see that the point estimates of  $\beta_0$ ,  $\beta_1$ , c have not changed by much, but the standard errors have decreased markedly. Thus if there is no statistical evidence of background mortality we should eliminate it from the model to increase estimation precision.

Let's see how this affects the percentile point estimates and lower confidence bounds on them.

## Threshold Fit

## No Threshold Fit

Percentile	Point Estimate	Lower 97.5 Per- cent Confiden- ce Bound	Point Estimate	Lower 97.5 Per- cent Confidence Bound
1	3.773	-157.377	1.657	-17.646
3	16.086	-109.518	14.388	0.000
5	22.606	- 84.331	21.130	8.900
10	32.647	- 45.967	31.511	21.622
15	39.421	- 20.657	38.515	29.364
20	44.805	- 1.246	44.082	35.034
30	53.572	27.152	53.146	43.539
50	68.064	54.847	68.130	56.372
70	82.557	67.675	83.115	68.459
80	91.323	73.775	92.179	75.594

We see that the point estimates under the nonthreshold fit are slightly lower than the point estimates under the threshold fit in the lower portion of the curve.

However the increased precision of estimation under the nonthreshhold fit results in substantial increases in the lower confidence bounds.

## Suggestion:

1. If there is no observed response in the control group.

- and 2. If the control group response rate, as estimated from the dose response fit including threshold is nonsignificant.
- and 3. If there is no a priori reason to expect background rate

then, eliminate background threshold parameter from the model.

This raises in a conjecture: Suppose we fit a nonthreshold model even when a non zero background rate exists. We conjecture that the point estimates of nonthreshold response rates will be estimates of quantities lower than the true response rates. However the increased precisions of these estimates may well result in more accurate lower confidence bounds on the "true" response percentiles. This is a bias-variance trade off.

It is interesting to note that Charles Stephan [42], page 78 ff discusses Abbott's correction in connection with the estimation of LC-50 concentration in acute toxicity tests. He comments "...Abbott's formula ... is a statistically sound way of correcting for control mortality if, and only if, the cause of the control mortality does not make the rest of the test organisms more susceptible to the toxicant. This assumption is usually questionable in acute mortality tests with aquatic animals. ... If control mortality is below a specified minimum...it should be reported along with the results of the test, but correction of the LC50 for this mortality would seem to be a meaningless exercise. ..." It is interesting that we arrive at a similar suggestion, based on different reasoning. Our motivation is a bias-variance tradeoff.

The previous PROC PROBIT analyses on the DeFoe data treated concentration without any transformation. We also tried to fit a probit model using log concentration. Folklore states that a probit or logit fit will better fit the response vs logarithmic concentration relation than the response vs arithmetic concentration relation.

Finney [11] page 8-13 recommends using log concentration. Stephan [42] also recommends the use of a logarithmic transformation of concentration on a routine basis.

Finney, pages 9ff states "The frequency distribution of tolerances, as measured on the natural scale (i.e. arithmetic scale - P.F.) is usually markedly skew, but often a simple transformation of the scale of measurement will convert it to a distribution approximately of the familiar Gaussian or normal form ... normalization can often .e effected by expressing the tolerances in terms of the logarithms of the concentrations instead of the absolute vlues. Indeed this transformation is now standard practice ... the justification is the widespread applicability of the normal distribution as an adequate approximation to the truth. ..." Stephan, page 75, states "...Whenever any method is used to analyze concentration mortality data, whether or not a transformation such as probit, logit, or angle is used on the mortality data, the logarithmic transformation should probably be used on the concentration data. All of the methods assume that the concentration-mortality curve is linear, and it seems to be generally accepted that the curve is more likely to be linear if log concentration is used. ..."

We show by example of the DeFoe 1, 1, 2, trichlcroethane data that the logarithmic transformation of concentration provides a much poorer fit of the probit model than does arithmetic concentration. The moral is that each time we fit a probit, logit, or other dose response model we should have an open mind as to using untransformed concentration, logarithmic concentration, or some other function of concentration. We should transform concentration in a manner suitable for each individual data set.

We first tried to fit the probit model with background response to logarithmic concentration. The attempted fit would <u>not</u> converge. To improve convergence performance we refitted the DeFoe data with logarithmic concentration using a <u>specified background rate of 0 (the observed level</u>). The output appears in Figures XIV.9. The maximum likelihood algorithm converges, however the resulting probit model is <u>not</u> an adequate fit to the data as indicated by the highly significant residual chi square statistic. (Chi square = 47.6774 with 4d.f.)

We break out the components of the chi square statistic by group, as discussed previously, to determine whether the very large chi square value is due just to one or two components with very small expected responses but with one or two observed responses. Such components can greatly distort the overall chi square.

0

Log <sub>10</sub> (Conc)	Trt Grp	0	N	Ê —	Nŷ	Np̂q	$\frac{(0 - N\hat{p})^2}{N\hat{p}\hat{q}}$
-1.3098	1	0	40	0.0000	0.0000	0.0000	0.0000
0.29907	2	0	40	0.000015	0.00060	0.00060	0.00060
0.7764	3	2	40	0.0025428	0.10171	0.101455	35.518
1.1706	4	1	41	0.0476	1.9516	1.8587	0.4872
1.6840	5	9	40	0.4227	16.908	9.7610	6.4068
2.1673	6	40	40	0.884	35.36	4.102	5.249
							47.662

We see the third treatment group contributes the most to the statistic. It has a very small expected frequency and two observed responses. If this were the only large deviation between data and model, we might be inclined to consider the possibility of a reasonable probit fit with the exception of an outlier group. However, even if we disregard this group, the components of chi square from the remaining cells sum to 12.144 with 3d.f. This value is still significant at the 0.01 level, even after the largest component has been deleted. We thus conclude that the model does not fit the data well. This, coupled with probit plots suggests the inappropriateness of the probit fit after a logarithmic transformation of concentration. The probit fit to untransformed concentration is superior in this case.

We now consider the Holcombe and Phipps compound D data. A listing of these data is contained in Figure XIV.10. The variable names correspond to those of the DeFoe data. The test consisted of six groups (control + 5 toxicant concentrations) and four tanks per group. Note that the control group concentration is 0 and there is a nonzero threshold response rate.

Probit models were fitted to the fry mortality data after pooling tanks within concentration groups. SAS PROC PROBIT was use to fit probit models both to concentration and to  $\log_{10}$  (concentration). These fits included background effects, to be fitted by maximum likelihood.

The probit fit vs untransformed concentration appears in Figures XIV.11, XIV.12. The residual chi square statistic is quite small (0.3361 with 3d.f.) signifying a good fit to the data. Figure XIV.12 contains the estimated percentiles of the probit response curve, <u>adjusted for</u> <u>background</u>, along with lower and upper 95 percent confidence limits calculated by use of Fieller's Theorem. For example for the 10th percentile the estimate for  $C_{10}$  is 78.77 while a lower 97.5 percent confidence bound is 58.72.

One difference between the fits to the DeFoe and to the Holcombe and Phipps data should be noted. In the DeFoe data no mortality was observed in the control group and the threshold response rate was estimated to be  $\hat{c} = 0.003$  with an asymptotic standard error of 0.029. Thus there was no evidence of background mortality and we markedly improved precision of the fit by deleting the background correction.

In the case of the Holcombe and Phipps compound D data we observe X = 6 deaths within the control group, with each of the 4 tanks exhibiting at least one death. Thus we know that there is background variation. From our probit fit with arithmetic dosage we estimate  $\hat{c} = 0.0718$  with an a-symptotic standard error of 0.016. Thus  $\hat{c}$  is 4.5 asymptotic standard deviations from 0 and so is highly statistically significant.

We now fit a probit model to the same data using  $\log_{10}$  (concentration). The estimated parameter values, their estimated asymptotic variance-covariance matrix, and the residual chi square statistic appear in Figure XIV.13. The residual chi square statistic is 0.5046 with 3d.f., which is very small, thus indicating a good fit to the data<sup>1</sup>. We thus have

<sup>&</sup>lt;sup>1</sup>Note that the chi square value 0.2287, given by SAS in Figure XIV. 13 is incorrect in this case. It seems to be omitting the control group contribution to the chi square statistic. This problem has been brought to the attention of the program developer and has been corrected in later versions of the program.

good probit fits to the data using both arithmetic and logarithmic concentration.

The estimated background level is  $\hat{c} = 0.0738$  with a standard error of 0.0156. Thus  $\hat{c}$  is 4.73 standard deviations from 0 and so is highly statistically significant.

The estimated percentiles (after adjusting for control group mortality) and 95 percent confidence intervals (by Fieller's theorem) are shown in Figure XIV.14. The lower confidence bound is a 97.5 percent one sided bound. This display is analogous to that in Figure XIV.12.

We have thus fitted two distinct models which seem to fit the data well: the probit model with arithmetic concentration (Figure XIV.11) and the probit model with logarithmic concentration (Figure XIV.13). The parameter estimates associated with these two fits are somewhat different. Namely

Arithmetic (StdErr)	Logarithmic (StdErr)
109.916 (3.638)	2.028 (0.016)
24.300 (3.861)	0.1044(0.015)
0.0718(0.016)	0.0738(0.016)

We see that the estimated background levels are somewhat similar, however the estimates  $\hat{\mu}$ ,  $\hat{\sigma}$  are very different.

Even though the parameter estimates differ considerably, the model fits may still be very similar. We compare the estimated response distribution percentiles and associated lower confidence bounds in Figures XIV.12, XIV.14 for the arithmetic and logarithmic concentration fits respectively. These are shown below.

Arithmetic (Figure XIV.12)

μ̂σĉ

i

Logarithmic (Figure XIV.14)

Percentile	Estinate	Lower 97.5% Conf. Bound	Estimate	Lower 97.5 % Conf Bound
1	53.39	22.40	60.97	45.76
2	60.01	31.92	65.10	50.13
3	64.21	37.95	67.87	53.12
5	69.95	46.15	71.83	57.47
10	78.77	58.72	78.38	64.83
15	84.73	67.13	83.14	70.29
20	89.46	73.77	87.12	74.92
30	97.17	84.43	94.03	83.01

The point estimates of the response distribution percentiles corresponding to the arithmetic and logarithmic fits are quite similar beyond the third percentile. However there is <u>considerable discrepancy between</u> <u>corresponding lower confidence bounds on safe dose</u>, based on each of the two fits -- below the 10th percentile! Below the third percentile the discrepancy is fifty percent or more.

How can we choose between the two fits?

#### 1. Prior knowledge or mechanistic information

None here. Since the probit model is an empirical model, no much in the way of mechanistic arguments will distinguish between the two fits.

#### 2. Magnitude of residual chi square

Arithmetic fit	Residual chi square $(\alpha = 0.953)$	=	0.336 with 3d.f.
Logarithmic fit	Residual chi square $(\alpha = 0.918)$	= 0	.505 with 3 d.f.

Both chi square values are quite small and the question of which one is larger is probably just a matter of chance fluctuations. Therefore we should <u>not</u> use these two statistically insignificant values to distinguish between the fits.

## 3. Appearances of plots of predicted and observed responses

Scatterplots of predicted and observed responses vs arithmetic concentration are shown in Figures XIV.15, XIV.16. Similar plots vs logarithmic concentration are shown in Figures XIV.17, XIV.18. The probit plots (Figures XIV.15, XIV.17) indicate greater discrepancies between observed and predicted responses (after adjusting both for background) at the low percentiles of the logarithmic concentration fit than of the arithmetic concentration fit. Similarly at the highest treatment group. Thus the arithmetic concentration fit seems to be a (slightly) better approximation to the data at the low percentiles than does the logarithmic concentration fit.

4. <u>Conservativeness</u>. Below the 25th percentile the lower confidence bounds based on the arithmetic concentration fit are lower than those based on the logarithmic concentration fit. The discrepancy is especially noticeable for the low percentiles, in particular below the 10th percentile. Above the 10th percentile both lower bounds are similar. Thus the arithmetic concentration fit seems to be more conservative than the logarithmic concentration fit at the low percentile.

Opinion I would prefer the arithmetic concentration fit in this case. However further experimentation at the low concentrations would be needed to distinguish between the differing conclusions at the low percentiles.

# Alternative analysis ignoring background

Background Included

We remarked with respect to the analysis of the DeFoe data that an alternative way of handling background mortality is to ignore it. The hope is that the improved precision will offset the downward bias and result in higher values for lower confidence bounds on safe dose.

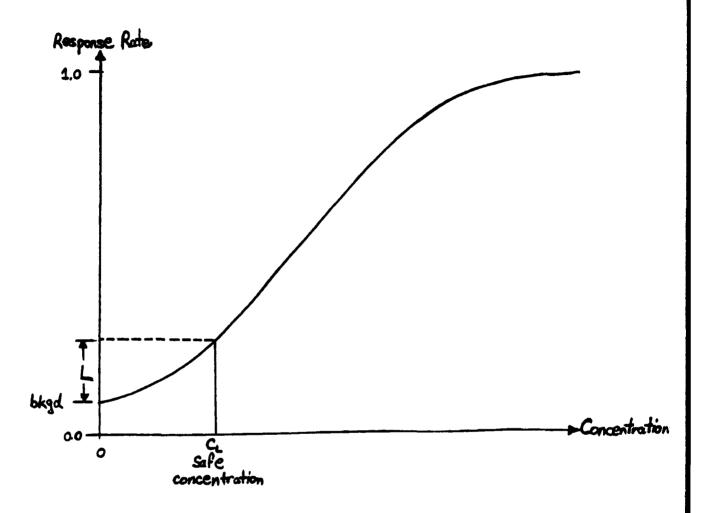
Note: If the control group response rate is significantly different from 0, as is the case with the Holcombe and Phipps data, we would not expect the dose response fit ignoring background to be a good fit to the data. Probit models ignoring background response, fitted to both arithmetic and logarithmic concentrations, show large and highly significant residual chi square statistics 28.96 and 65.03 respectively with 4d.f. The plots of fitted and observed responses vs concentration also show discrepancies.

We now compare the percentile point estimates and their lower confidence bounds based on the fit ignoring background response with those based on the fit including background response. Comparisons pertain to the fit vs untransformed concentration.

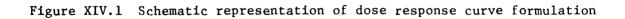
Background Excluded

	<u>v</u>	<u> </u>		
Percentile	Point Estimate	Lower 97.5% Conf. Bound	Point Estimate	Lower 97.5% Conf. Bound
1	53.39	22.403	-7.48	-24.234
3	64.21	37.946	13.37	0.648
5	69.95	46.147	24.41	13.197
7	74.05	52.006	32.33	21.991
10	78.77	58.715	41.42	32.626
15	84.73	67.135	52.89	45.860
20	89.46	73.775	62.01	56.369
25	93.53	79.421	69.83	64.178
50	109.92	101.288	101.40	97.761

We see that in this example, with the background level many standard deviations from 0, the bias-variance trade off is such that it does not pay to reduce the assumed background response level to 0 in order to lessen the standard deviation.



.



PRPDEADE PRPDEADF LOSCONC (My. Come)	0 3 3 4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
PRP	000000000000000000000000000000000000000
DEADFSUM	6000000-4000 NN
DEADESUM	00000000000000000000000000000000000000
FRYSUM	000000000000000000000000000000000000000
EMHSUM	20000000000000000000000000000000000000
sî p	40 40 4 D 40 40 40 40
CENCHEAN	000
TRT	
08 S	

Basic rate. For Arabit Analysis - Delie Compaind C.

Figure XIV.2 Basic data

		Defoe Compound C	nd C			
		Fandri Fll DE F	PADATE FIT DE FRY DATA: OPIC, CEO		11:34 TUFSDAY, JULY 29, 1980	, 1980 2
		FFUCET ANALY	PEUGLI ANALYSIS EN CONCHEAN			
G TERATIVE C	I di ni di di la	SLrpE	ى	٩w	S IGMA	
	00000000000000000000000000000000000000	0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	90. 73507155 69. 30553265 66. 104951779 67. 63555493 67. 63555493 68. 10757133 68. 10757133 68. 95467335 68. 95467335 68. 95457335 68. 95457335 68. 95457355 68. 95457355	34.6560393 34.6560393 34.6560393 32.1152 32409999 32409999 3240999 3240999 2214555 2214555 2214555 2214555 2214555 2214555 2214555 2214555 2214555 2214555 2214555 2214555 2214555 2214555 2214555 2214555 2214555 22145555 2214555 2214555 2214555 22145555 22145555 22145555 22145555 22145555 22145555 22145555 221455555 221455555 22145555555555	
Crvii I AHGE WATKLY INTERCIET	Jaurs		CUVARIANCE MATRIX	fir. XIa	VH51S	ں
С INTERT C. 6956913 StΩrF - 0.01097308 C. 241353		-0,02241393 6,00034624 0,00093146	AUJ SIGMA	59.76577925 40.34127025 -0.03185415	40.34127025 10.59470236 10.2444583	-0.03185415 -0.2544583 0.00093145
	= 05-1HD	4.2108 WITH	3 DF PROB > C	PPOB > CHI-5C = 0.2396		
MOTE: SINCE THE CHI-SOUNDE IS SMALL (P > 0.101.	15 SWALL (P > 0.10)	FIBUCIAL LIMITS WILL	WILL PE COMPUTED US	PL COMPUTED USING A T VALUE OF 1.	1.96	

í.

Figure XIV. 3 Output from PROC PROBIT fit to DeFoe fry mortality data -- arithmetic concentration

٦

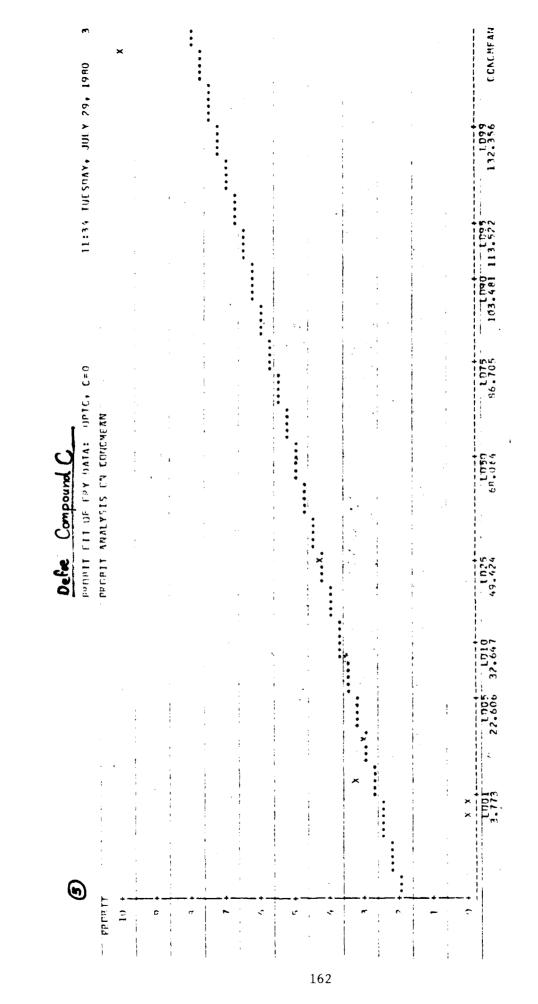
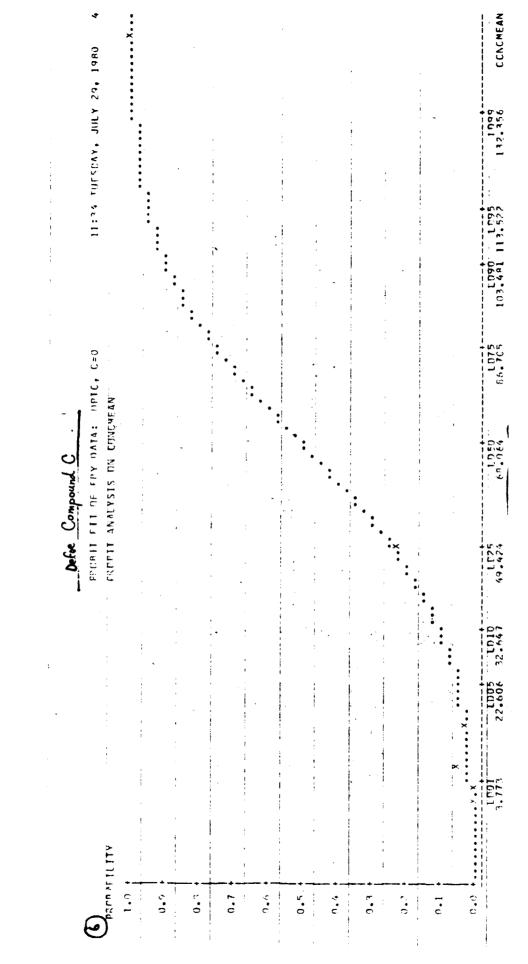
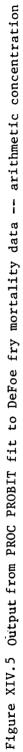


Figure XIV.4 Output from PROC PROBIT fit to DeFoe fry mortality data -- arithmetic concentration





Defee Compound C

TeaPIT FIT OF FRY DATA: JOTC, C=0

ŝ 11:34 TUESDAY, JULY 29, 1980 .

٠. ٦ • ٠.

1				į		
•						
				і	1	•
				:		
			+			
		-	1	•		
			,			!
			•	· ·		
	νx	4-0000	ແມວຈາມພາກ	-posine-ki		: Dinegi
	TML	こうしょうこ しこうしょう	アウルのつれれー えたいのじゅうー	20000000000000000000000000000000000000	とまた のみぎるに どうじゅうであれ だてのじらじ ひょ	1010 1010 1010 1010
	1	-02-44	~N048400	11-040000	SHARCADOK OFEEMAN	
	٩٢	- 0 0 0 x 4	00000000	NHMININ 201		5000
	10G	0.459E 00	24240000	CUL 08-40	ーまつこここここ ~ ひょうよう らん * ひょうひょう	200000 20000 20000
1	F TDUC TAL		:			
	•					
	PTRUENT DMEP		•			1
<u>.</u>	1.1	441.108	00000000000000000000000000000000000000	1511510 860832590 860832590 755580389 75513810 75513810 75513810 757510 7575100 757510000000000		~~~~~
כחיול אי אי	101	04 11 20			10000000000000000000000000000000000000	5 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
ñ "	5		~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	シークきんこうです。	**~00-~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	「なしの」
		~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	000 NIC 0	-4040)-4+0	comments of a	
č		111		in de la la navec 47 des	- <b></b> 0000000	
13				1		i
LYS					1	
PERTT AVILYSTS	Z	2504-0	<u>. – n – n – n – n – n – n – n – n – n – </u>		- EUNNODER	
-	1.1.2		5-4-64 m 4	55400000000000000000000000000000000000	00000000000000000000000000000000000000	1000 - 100 - 1000 - 100 - 1000 - 100
111	CNCMEAN	2005 20 20 20 20 20 20 20 20 20 20 20 20 20	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	41525547 561575194 55112058 55412058 5559258 5555554 7125554 712554 712554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75554 75564 75554 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75564 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 75664 756647 756647 756647 756647 756647 756647 756647 756647 756647 756647 756647 756647 756647 756647 756647 756647 756647 756647 756647 756647 756647 756647 756647 756677 756677 7566777 75667777777777	10040040 10040040	
22	Ξ	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	NCUSESS	10-10000		1 3 C C C
			アローこうかいう	S B C B C C C C C C C C C C C C C C C C	1965-196-	
			•	• •	+	t
			•	· · ·	r	1
i		1		-	· ·	-
	11					i i
	Ξ				1.	
	P¤Cia∧r I	0000000		50000000000000000000000000000000000000	-246969696	0.000-000 0.000
		000000	20000000			00000
	Jac			1	:	1
2	lad		:			-
Ð	lad					

.

.

Figure XIV.6 Output from PROC PROBIT fit to DeFoe fry mortality data -- arithmetic concentration

▼ 11suda			14:12 SUNDAY, AUGUST 3, 1980
	AN		
		SIGMA	:
0.03406555 0.03470469 0.034404695 0.034404695 0.034464552 0.034464552 0.034464552 0.034464552 0.034464552 0.03465924 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.03464552 0.0346552 0.0346552 0.0346552 0.0346552 0.03465552 0.03465552 0.03465552 0.03465552 0.03465552 0.03465552 0.03465552 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.03465555 0.0345555 0.0345555 0.0345555 0.03455555 0.03455555 0.03455555 0.03455555 0.03455555 0.03455555 0.03455555 0.03455555 0.03455555 0.03455555 0.03455555 0.03455555 0.03455555 0.03455555 0.03455555 0.03455555 0.03455555 0.03455555 0.034555555 0.034555555 0.034555555 0.03455555 0.034555555 0.0345555555 0.0345555555 0.03455555555555 0.03455555555555555555555555555555555555	80.78607155 73.53118048 68.54158048 68.54158145 68.131023459 68.130023421	39.64239896 32.64239883398 29.651756619 29.557756619 28.557415660 28.57415660	-
C = 0.000000C			
INTERCEPT SLUPE SLUPE	COVARIANCE MATRIX	MU	SIGMA
ТІ ТЕАСЕРТ <u>0.06660365</u> <u>-0.00118544</u> <u>0.00003515</u> 5.17Р: -0.00118544 0.00018546 <u>0.00003515</u> S	MU	55+25332279 28+22800846	28+22800846
		0	

.

•

•

1

Figure XIV.7 PROC PROBIT fit to DeFoe fry mortality data --- no background correction

Defoe Compound C Assemption of No Theshold

Assamption No Include PROBIT FILLEF FRY DATA GAD.

-

PROPIT ANALYSIS ON CONCMEAN

14:12 SUNDAY, AUGUST 3, 1980

÷

5

7

and the second second

					• • •
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	:				
FIDUCIAL LIMITS	0117993 98520945 38320945 33339780	00000000000000000000000000000000000000	72.99549722 73.9106425 83.413106425 83.413106425 93.548321365 93.548321365 93.79347527 1.0.014054417937	200242000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 20025000 2002500000000	6830735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 66870735 6677075 6677075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 6777075 7777075 7777075 7777075 7777075 7777075 7777075 777775 777775 777775 777775 777775 777777
LOWERCENT	100000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 1000000	100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 100400 1004000 1004000 1004000 10040000 1004000000 100400000000 10040000000000	6 5 7 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5		3-5600276 5-5800276 8-55806736 2-5590803 2-534267
	6567985 4660895 4860895 105951375 105951375 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 1059520 10595200 10595200 10595200 10595	2010 2010 2010 2010 2010 2010 2010 2010	57-12(05569 66-13456569 66-13456569 71-12323417 71-12323417 71-12323417 71-123241272 71-12324205 71-1242425 72-12655		
		00000000000000000000000000000000000000	00000000 00000000000000000000000000000	ng⊇¥,⊴.⊣cin4 Panagogocina D⊇⊇gocina	6 92 3 5 6 6 9 5 7 6 7 6 9 5 7 6 7 7 6 7 6 7 7 6 7 7 6 7 7 7 7

Figure XIV.8 PROC PROBIT fit to DeFoe fry mortality data -- no background correction

•

5 T C M 2		251919174 34819174 34819174			0.00555591	thetics address of bet of Fit	MODEL.
nw.	0 - 5 - 5 - 7 - 5 - 7 - 5	1.75192751 1.75189275 1.75189075 0.3481 0.3481	And a man is a second of the s	COVARIANCE MATRIX	0.031445	PROB > CHI-50 = 0.0001 - Stars . + herber - detains . of herb of F.	2 EACTOR H= 11.919361 H= EMATIC DEPARIDAE FROM THE
CLPT		577939 2.95971407 136657 2.959799079 140211 2.97198394	C = 0.000000	St CPT.	-2-11 772779 MU -1-62647212	47.5775 WITH 4 DF PROB >	THE CEN MULTIPLED BY THE HETRUGENEITY FACTOR H= 11-9193681 . The CEN MULTIPLED BY THE HETRUGENEITY FACTOR H= 11-9193681 . TING FILUCIAL LIMITS.
TTERATION	2 - 27420 0 - 14720 0 - 14720	54 C		COVERIANCE MATRIX INTERCET	4.79506845 -2.	CH1-53 = 47.	NOTE: THE ARUVE VARIANCES AND CUVARIANCES HAVE E CHICK IL SOFETHEL HE LERG CHICSCONFL CLIVELUE OF 22 74 MILLEE USELLIN COMPUTIN

.

6

and the second second

.

- -

•

ø

14:12 SUNDAY, AUGUST 3, 1980

FRUELT EIT GE FRY DATA: C=0, LOGIO

Defoe Compound C -- No threshold

PROC PROBIT it to DeFoe fry mortality data -- logarithmic transformation of concentration Figure XIV.9

:

1

ł

Holemske and Phipps-Compared D

• •

e.	J					1	-1,			0-0	<b>N</b> 0	•	Ċ,	0.00						
+ AUGUST	L DGCDNC		•	•	• •	1.4348	1.4348		1.6522	- 6522	1.6522	1.8630		1.8630	2.1077	2101-2	2.1077	2.0845	2.9480	2.34801
14:31 SUNDAY, AUGUST 3,	PRPDEAUF			40.0	0.04	0.12			0.03		0.04	0.16		00.00	0.92	19	0.76			1.00
	PRPULADE	0.34	0.42	0.24	0.40	0.38				184 ° 0	0.24	04.00		0.28	920	C.28	0.32		- - -	0.44
SURVIVAL AND TEXTCAN	DEADFSUM	2	161	-	<b></b> 1		<b>-</b> 10	łcai	•<	241		<b>4</b> u	· - <b>1</b>	0.0		16	0-10		1040 1101	<i>c</i> 2
•	DEADEGUM		21	12	02	, ,		ی ہے۔ ار کا		4-14 4-14	2		40°.	40	6	14	04	9	500	1.2
GULFRIG UALA	FR YSUM		5-1 1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	() 		10% 414	11-14 (~~{)		100	17-14 (48)	1 1 1 1 1	10. 10.	25	33	1	15	и ( - 1	( <i>1</i> 0	
	F MH 5U M	53	0 4	00	22		20	00 614	         	0	50	- - 	<u>C</u>	5 0 0 0	65	С., И.,		3	⊂، ز ۲۰۱	,
•	d - 1	4	aı	50	2<	4	د:	>د	ta	ى	<u></u> م<	( c:	J	•ت	ш	ە.	•	T	<u>ں</u> ב	•
	$CU^{1}V_{1}V_{2} = A^{1}V_{2}$	- <u>j</u> *;					27.215	5 <b>7 •</b> 5 <b>1</b> 5	000 . 44	4 4 0 C 5		093	223		1 28.150	26.150	2-2-850	250.052	0 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	
	T : T		┛,	-1	40	•*•	. <b>۲</b> . ۱	•.'	   ~~	: <b>ل</b> .	×. 4	4	1	<b>‡</b> 5	1   u'u	հե	ן כי ר	Ŷ	<b>-0</b> -0	
	53 - 2 - 2 -		.16		tur			- J	01	-4[	\	14	للاً - همر ا			5° C 1 (	<b>-</b> ( 4	ru) ru	n) 4 (vr. 4	+ + + + +

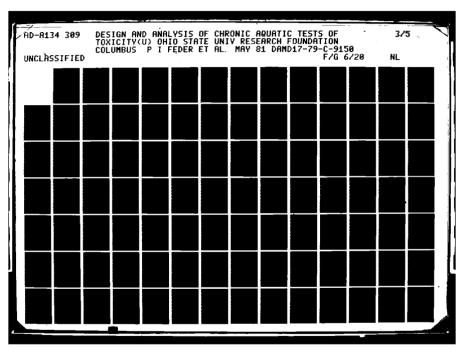
Figure XIV.10 Basic data

Holcombe and Mippe- Congrand O-Arithmetic Countration - Fry Montalty Octa

9, 19A0 Z		e		U	0*02344206 -0*02637706 0*00005637701	
10:44 10FSEAY, JULY 29, 1980		V I C N	34 - 4251 3557 24 - 5272 5338 225 - 5272 5338 225 - 5272 5338 225 - 327255156 225 - 320125535 225 - 3201355535 225 - 3201355535 225 - 3201355535 225 - 320135555 225 - 32013555 225 - 3201555 225 - 3201555 225 - 3201555 225 - 32555 225 - 325555 225 - 3255555 225 - 3255555 225 - 3255555 225 - 3255555555 225 - 3255555555555555555555555555555555	SIGHA	-6.00639777 1062839901 -0.026379	•
10:4		n,.	102.09752235 108.605850566 109.85650566 109.8565196 109.85653265 109.91625245 109.91625245 109.91625245		13.23315612 -6.30638777 0.02344269	PRUB > CHI+SQ = 0.9531
ATA: 0016, C=0.06	S CN CONCHLAN	C.	3.06/000022 0.075/0022 0.071/65233 0.071/65233 0.071/6700 0.071/6700	COVARIANCE MATRIX	SIGNA	3 DF PR08 > CHI-
PECALLE FPY DATA: DATA: C=0.06	RATES AS SISTING AND CONCREMENT	slept	C C 7594 F5 C C 75934 F5 C C 7593 755 C 6511 57 C 6511 57 C 6511 57 C 6511 57 C 6511 51 69 C 6511 151 69 C 6511 151 69	U J	-c.00557447 6.00064447 C.00025335	0.3361 WITH 3
		INTERCED -	199911091 199911901 199929 199929 199727 199727 199727 199929 199929 199929 199929 199929 199929 199929 199929 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 199920 1999200 199920 199920 199920 199920 199920 199920 199920 199920	3ľ£⊅E	-9-000011844 -0-0000475 	CH1-20° = 0'
	ч	1158A1100		THIFPELD	2++18900+0+ \$\$\$11800+0+ €\$\$11800+0+	
		-		C .VIELANCE MATRIX	141197797 51139 6	

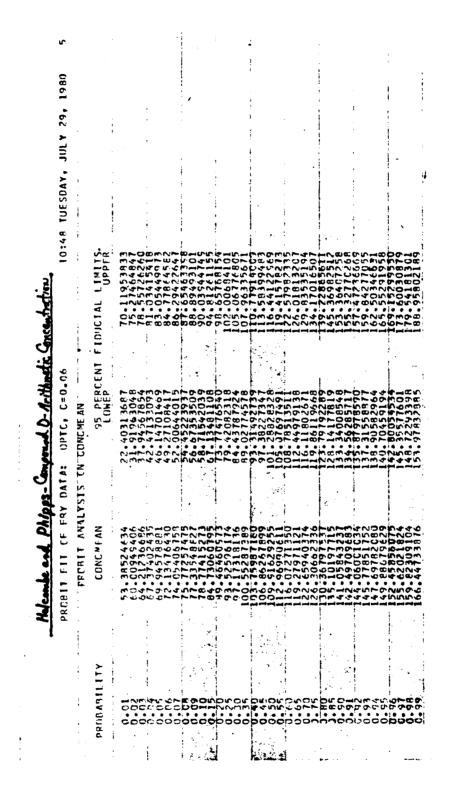
HETER STACE THE CHISSOMARE IN SMALL (P > 0.10), FIDUCIAL LIMITS WILL BE COMPUTED USING A Y VALUE OF 1.96.

Output from PROC PROBIT fit to Holcombe and Phipps fry mortality data -- arithmetic concentration Figure XIV.11



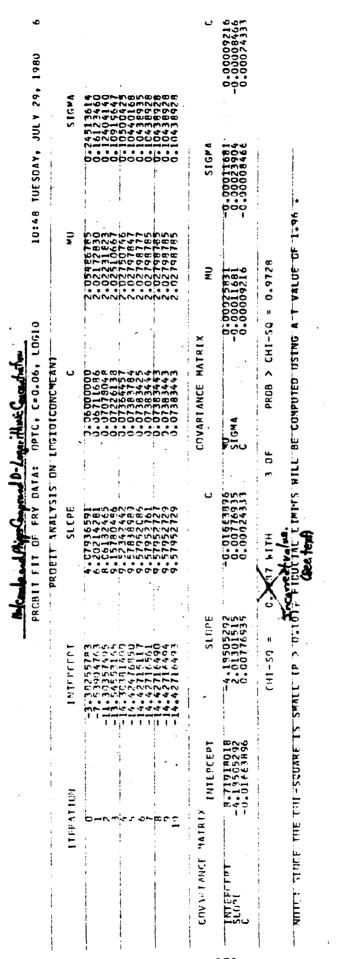


MICROCOPY RESOLUTION TEST CHART NATIONAL BUREAU OF STANDARDS-1963-A



-- arithmetic Output from PROC PROBIT fit to Holcombe and Phipps fry mortality data concentration Figure XIV.12

17	 •	•	•	
Γ.				
L				
F.				
F.				
L-				
Γ.				
Ε.				
Ŀ				
7				
5				
Ы				
6				
1×				
L.				
Ľ				
Ē				
Ľ				
E.				
Þ.,				
Ľ				
E				
Р				
С				
L.				
Ŀ				
٢·				
r٠				
۲·				
Ľ				
*				
5				
Ð				
Ľ				
Γ,				
5				
Þ,				
F.				
F.				
Ľ.				
Ŀ				
r.				
P.				
F.				
Þ				
Ľ٠				
F.				
C				
۲				
F				
24				
F				
5-				
64				
E1				
E-				
5.				
P				
F,				
<b>.</b> '				
<b>.</b> .				
١.				
Ľ				
F.				
<b>1</b>				



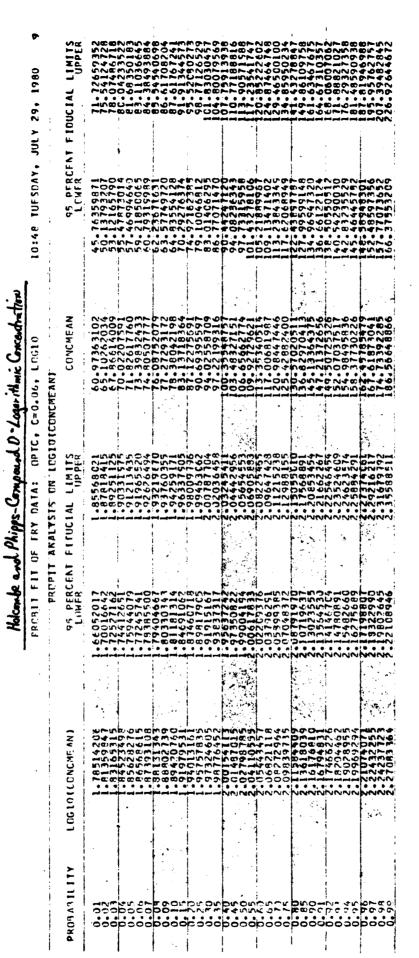
•

Υ.

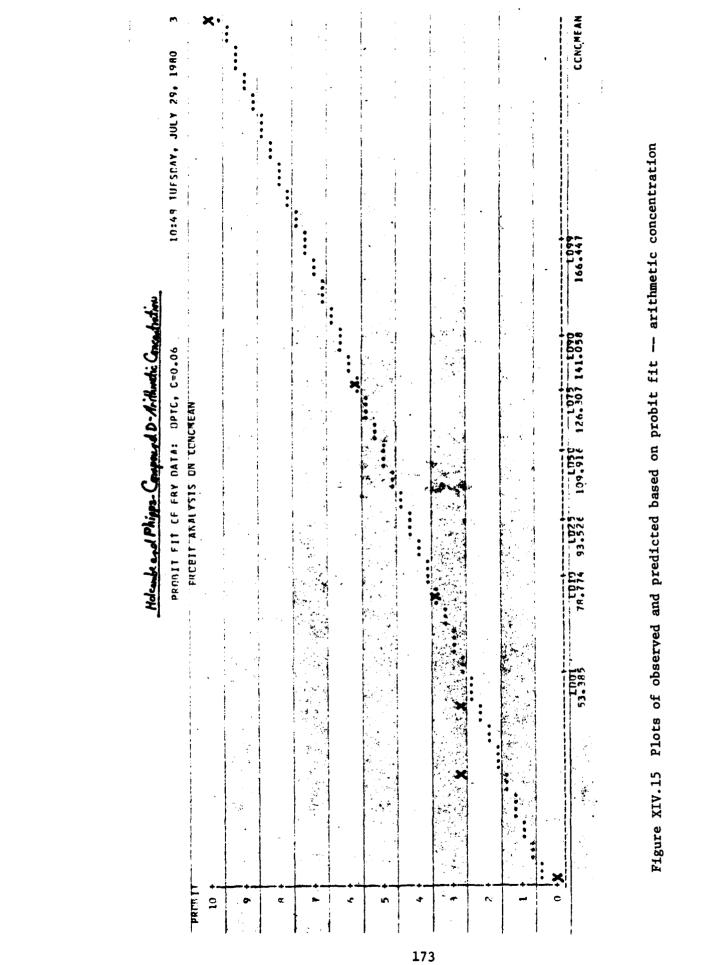
 
> Output from PROC PROBIT fit to Holcombe and Phipps fry mortality data -- logarithmic concentration Figure XIV.13

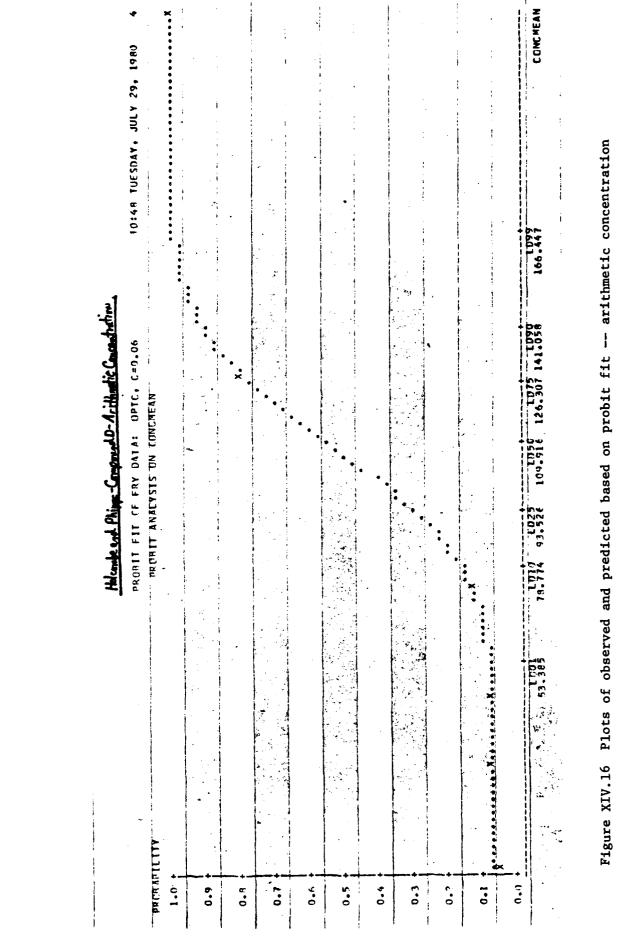
· · · · · ·

. . . .



logarithmic 1 Output from PROC PROBIT fit to Holcombe and Phipps fry mortality data concentration Figure XIV.14





Contract The State

7

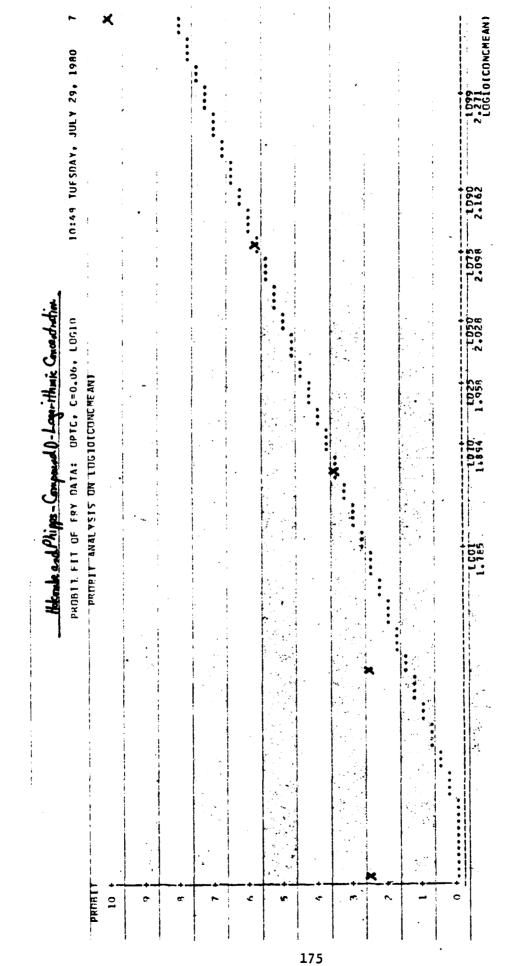
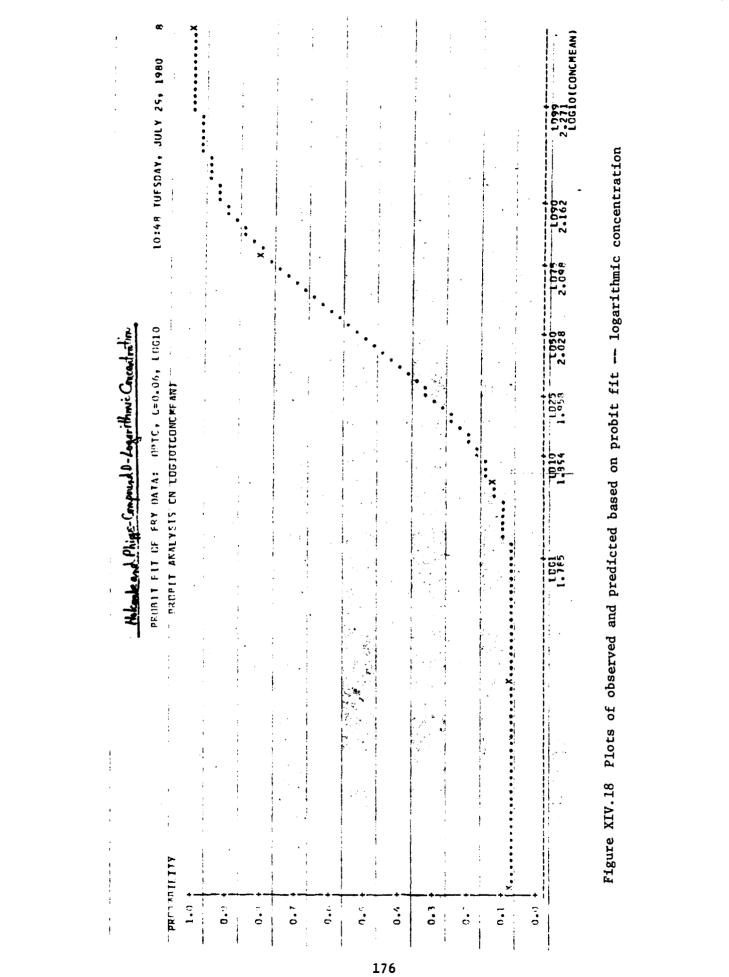


Figure XIV.17 Plots of observed and predicted based on probit fit -- logarithmic concentration



## XV. DOSE RESPONSE CURVE ESTIMATION -- MAXIMUM LIKELIHOOD ESTIMATION BY NONLINEAR LEAST SQUARES REGRESSION

We have seen in the previous section how standard dose response models can be fitted to the data by use of SAS PROC PROBIT. This procedure fits a probit model to the data, with several possible variations. Namely it fits the model

$$p(x) = c + (1 - c)\Phi(\beta_0 + \beta_1 x)$$

where x = concentration or log (concentration), c is the background rate, and  $\beta_0$ ,  $\beta_1$  are unknown parameters to be estimated from the data. p(x) is the response probability corresponding to x. The value of c may be known or unknown. Estimation is done by maximum likelihood, based on binomial theory.

Jennrich and Moore [43] show that for distributions in the exponential family, maximum likelihood calculations can be carried out by means of nonlinear least squares regression calculations. This applies, in particular to models based on the binomial distribution.

Thus dose response curves can be fitted to the data by use of appropriate nonlinear regression programs. Both SAS [12] (PROC NLIN) and BMDP [27] (P3R and PAR) contain nonlinear regression programs that can carry out these calculations. See Jennrich and Moore [43] for a discussion of the theory underlying the relation between maximum likelihood estimation and nonlinear regression in the exponential family. We illustrate the methodology with the use of SAS PROC NLIN. However any nonlinear regression program with capability to carry out iteratively reweighted least squares (i.e. allow weights to be functions of the model paramters) would suffice.

SAS PROC PROBIT also calculates lower and upper confidence bounds on concentration values corresponding to various response curve percentile (after adjusting for background), by use of Fieller's theorem (Finney [ 11 ] pp 78-79). We illustrate how these confidence bounds can be calculated, based on the parameter estimates of the fit and their asymptotic variance-covariance matrix.

Before discussing the details of fitting dose response curves by means of nonlinear regression programs. We discuss some of the reasons that one might wish to do this.

1. The data analyst may have a general purpose nonlinear regression program available but no special purpose dose response estimation program. Thus the general tool can be used without modification instead of having to write a special purpose program. 2. A very wide variety of models can be fitted to the data by use of general nonlinear regression. PROC PROBIT is rather limited in the extent of models it will fit. Namely it will fit only a probit model using concentration or log concentration. It will adjust for background effects only by Abbott's correction.

We may wish to fit models other than the probit, e.g. the logit model, or even more complex models that incorporate both the probit and logit models as special cases. Background effects might be modelled as additive concentrations rather than by Abbott's correction. Namely,

 $p(\mathbf{x}) = \Phi(\beta_0 + \beta_1 \log (\mathbf{c} + \mathbf{c}_0))$ 

where  $c_0$ , the background effect, represents an alternative way of accounting for background. Such a model, although nonstandard, can easily be fitted by nonlinear regression techniques. Also, transformations of concentration levels other than the logarithmic are useful. For example the square root transformation.

- 3. PROC PROBIT automatically inflates variance and covariance estimates estimates and confidence interval bounds by heterogeneity factors whenever the probit model does not fit the data (as determined by the residual chi square statistic). This is not always what we wish to do. PROC NLIN does not inflate variance estimates by heterogeneity factors.
- 4. We can calculate and save predicted and residual values and thus easily construct residual plots.

It should be noted that PROC NLIN will <u>not</u> compute confidence intervals on response curve distribution percentiles by use of Fieller's Theorem, as PROC PROBIT does. However we show in the subsequent discussion how we can carry out these calculations fairly easily, using either a hand calculator or a small computer program, once the parameter estimates and their asymptotic variances and covariances have been determined.

We now consider three examples of fitting dose response models to fry mortality data by use of SAS PROC NLIN. We use the Holcombe and Phipps fry mortality data in all the examples. The models fitted are:

$$p(conc) = c + \bar{c}\Phi(\beta_0 + \beta_1 conc)$$

$$p(\text{conc}) = \frac{\beta_0 + \beta_1 \text{conc}}{1 + e^{\beta_0}} \left( \frac{\beta_0 + \beta_1 \text{conc}}{1 + e^{\beta_0}} \right)$$

$$p(conc) = \Phi(\beta_0 + \beta_1 \log_{10}(conc + c))$$

All the models are fitted based on binomial distribution theory after pooling data across tanks within groups. This is the way that PROC PRO-BIT fits models and is appropriate if there is no tank to tank heterogeneity within groups. In the presence of tank to tank heterogeneity we can first adjust the data by an adjustment factor and then pool across tanks within groups.

The first model fit is a repeat of a model we fitted by PROC PRO-BIT and serves to verify that we can duplicate the PROC PROBIT fits by nonlinear regression. The second model is a logit model and illustrates that we can fit alternative models with PROC NLIN and that the probit and logit model fits result in very similar inferences.

The third model fit treats the background as an additive concentration rather than adjusting for it by Abbott's correction. This sort of model would be appropriate if the mechanism of response due to background sources is the same as the mechanism of response due to the substance under test.

We now discuss the formulation of fitting dose response curves by means of nonlinear regression techniques.

Suppose that there are I concentration groups (both control and treatment) and that the i-th group contains N<sub>i</sub> subjects and has X<sub>i</sub> responses. Let  $p_i(\theta)$  denote the response probability within the i-th group. Then  $X_i \sim \text{Binomial } (N_i, p_i(\theta))$ .

The form of  $p_i(\theta)$  is specified by the form of the dose response model. For example in model 1,  $p_i(\theta) = c + \bar{c} \Phi(\beta_0 + \beta_1 \text{conc}_i)$ , where  $\theta \equiv (\beta_0, \beta_1, c)$  is the unknown parameter vector, to be estimated by least squares.

Under the model assumptions  $X_{i}$  has meen  $\mu_{i}(\theta)$  and variance  $\sigma_{i}^{2}(\theta)$  , where

$$u_{i}(\theta) = N_{i}p_{i}(\theta)$$

$$\sigma_{\mathbf{i}}^{2}(\boldsymbol{\theta}) = N_{\mathbf{i}} P_{\mathbf{i}}(\boldsymbol{\theta}) (1 - P_{\mathbf{i}}(\boldsymbol{\theta})).$$

The nonlinear regression procedure optimizes the function

$$Q(\theta) \equiv \sum_{i=1}^{I} (x_i - \mu_i(\theta))^2 w_i(\theta)$$

where  $w_1(\theta) = 1/\sigma_1^2(\theta)$ . Jennrich and Moore [43] show that optimizing  $Q(\theta)$  by the Gauss-Newton method is equivalent to fitting the dose response

curve by maximum likelihood estimation.

We fit the models to the Holcombe and Phipps compound D data. We first consider,  $p_i(\theta) = c + \overline{c}\Phi(\beta_0 + \beta_1 \text{conc}_1)$ . The results of fitting this model in the standard manner with PROC PROBIT are shown in Figure XIV.11.

We first discuss the NLIN commands needed to produce the output. See the SAS 79 manual [12], pp 317-329, for further details.

The model is

$$\mu_{i}(\theta) = N_{i}[c + \overline{c}\Phi(\beta_{0} + \beta_{1}d_{1})] \equiv N_{i}p_{i}(\theta)$$

where  $(\beta_0, \beta_1, c) \equiv \beta$  are to be estimated by weighted least squares. The weights are  $w_i \equiv w_i(\beta) = 1/[N_i p_i(\beta)(i - p_i(\beta))]$ . The fitting algorithm also uses the derivatives of the mean value function. These, are:

 $\frac{\partial \mu}{\partial \beta_0} = N_i \bar{c} \phi (\beta_0 + \beta_1 d_i)$  $\frac{\partial \mu}{\partial \beta_1} = N_i \bar{c} d_i \phi (\beta_0 + \beta_1 d_i)$  $\frac{\partial \mu}{\partial c} = N_i [1 - \phi (\beta_0 + \beta_i d_i)]$ 

where  $\phi(\mathbf{x})$ ,  $\Phi(\mathbf{x})$  represent the standard normal probability density function and cumulative distribution function respectively.

The SAS commands needed to generate this fit are given below

2	PROC NLIN BEST=20 METHOD=GAUSS; PARAMETERS BC=-5.0 TO -4.0 BY 0.25 51=0.0 TO 0.10 BY 0.025 C=0.03 TO C.11 BY 0.02;	
	BOUND O<=C<=1.0; ARG=BO+B1*CONCMEAN; ARG=MAX(ARG,-5.0); ARG=MIN(ARG,5.0); BIGPHI=PROBNURM(ARG); SMUDITED TO	×
•	SMLPHI=0.3989*EXP(-0.5*ARG**2); PROB=C+(1.0-C)*BIGPHI; MODEL_DEADFSUM=PROB*FRYSUM; DER.BO=(1.0-C)*SMLPHI*FRYSUM; DER.BI=(1.C-C)*CONCMEAN*SMLPHI*FRYSUM; DER.C=(1.0-BIGPHI)*FRYSUM;	
6	, OUTPUT OUT=HOLPHIA PREDICTED=PRDFSM RESIDUAL=RSDFSM; WEIGHT = 1.0/(FRYSUM*PRUB*(1.0-PRUE)); TITLE2 PROBIT MODEL FIT WITH ABBOTT"S CORRECTIONUNTRANSFORMED C 180	ONCENTRATIC
	100	

Line 1 instructs NLIN to fit by the Gauss-Newton method. NLIN can also fit models by use of the method of steepest descent or the Marquardt method. The Marquardt method is a compromise between Gauss-Newton and steepest descent. Sometimes near the optimum, steepest descent or Marquardt methods take smaller steps and in different directions than the Gauss-Newton method and so are less prone to overshoot the optimum. Thus one optimization method will sometimes produce convergence when another one does not. The BEST = 20 command instructs NLIN to print out the locations and error sums of squares values that it calculates in the preliminary grid search to determine starting values for the iterative portion of the search.

Line 2 specifies parameter values and/or ranges of parameter values that NLIN should use for a preliminary grid search to arrive at starting values for the iterative phase.

Line 3 specifies bounds on the parameters. If the parameters exceed these bounds at any time during the iterative process they are forced back into the permissible region.

Line 4 contains the model specification. The variable DEADFSUM represents the mortality within the i-th group. FRYSUM is the total number of fish exposed (pooled over tanks within groups) and PROB is the response probability in the i-th group. The SAS program statements between lines 3 and 4 are used in the specification of the model in line 4.

Line 5 contains expressions for the derivatives  $\partial \mu / \partial \beta_0$ ,  $\partial \mu / \partial \beta_1$ ,  $\partial \mu / \partial c$  respectively.

Line 6 specifies that the predictions and residuals from the fit be calculated and saved for future use.

Line 7 specifies the weights that are to be used in the weighted least squares fit. Note that these weights are functions of the model parameters (through PROB). They are updated following each iteration.

The output from these commands appears in Figures XV.1 to XV.4. Figure XV.1 contains a listing of the Holcombe and Phipps data. Figure XV.2 contains a summary of the parameter values and residual sums of squares associated with the 20 best points in the preliminary grid search. The point with smallest weighted residual sum of squares is used to start the iterative Gauss-Newton search procedure. The results of the Gauss-Newton iteration are summarized in Figure XV.3. It converges after 8 steps. Figures XV.4 contains statistics based on the model converged to in Figure XV.3. The upper portion of the page contains an analysis of variance table based on weighted sums of squares. The middle portion of the page contains parameter estimates and asymptotic standard errors. The bottom of the page contains the asymptotic correlation matrix among the parameter estimates. We compare these results with those in Figure XIV.11.

Several points need to be remembered in making the comparison.

- 1. We are fitting the model  $p(\theta) = c + \bar{c}\Phi(\beta_0 + \beta_1 \text{conc})$  whereas PROC PROBIT parameterizes the model as  $p(\theta) = c + \bar{c}\Phi((\beta_0 - 5) + \beta_1 \text{conc})$ . Thus the estimates  $\hat{\beta}_1$ ,  $\hat{c}$  in the two fits should agree while the PROC NLIN estimate of  $\hat{\beta}_0$  should be 5 smaller than the corresponding PROC PROBIT estimate. Comparison of the estimates shows that this is the case.
- 2. The residual chi square (0.3361) calculated by SAS PROC PROBIT is the same as the (weighted) residual sum of squares in the PROC NLIN fit. Thus this residual sum of squares provides a test of goodness of fit of the model.
- 3. The asymptotic variances and covariances calculated by PROC NLIN need to be adjusted before being compared to those calculated by PROC PROBIT. In particular, combining the asymptotic standard errors and the asymptotic correlation matrix obtained by PROC NLIN, we calculate the asymptotic variance covariance matrix.

 $\begin{pmatrix} 0.066928 - 0.000543 - 0.000618 \\ -0.000543 & 0.00000455 & 0.00000468 \\ -0.000618 & 0.00000468 & 0.00002797 \end{pmatrix}$ 

This matrix looks nothing like the asymptotic variance covariance matrix that is calculated by SAS PROC PROBIT. The reason for this is as follows. We stated that  $X_i$  has mean  $\mu_i(\theta) \equiv N_i p_i(\theta)$  and variance  $\sigma_1^2(\theta) \equiv N_i p_i(\theta) (1 - p_i(\theta))$ . However the weighted least squares fit is carried out assuming that  $Var(X_i) \equiv k\sigma_i^2(\theta)$ . where k is a constant to be estimated from the data. Thus the estimates of the variances and covariances given by PROC NLIN assume  $\hat{Var}(X_i) = \hat{k}\sigma_1^2(\hat{\theta})$ . Thus all variances and covariances are multiplied by  $\hat{k}$ .

How is k estimated? Just as in the case of weighted linear regression, k is estimated by the residual mean square. Namely,

k = weighted residual mean square = 0.11203710

Our maximum likelihood model, though, tells us that k = 1. We thus need to adjust all variances and covariances to this value of k. To do this, we simply divide the above variance covariance matrix by  $\hat{k}$ . When this is done we obtain

 $\begin{pmatrix} 0.5974 & -0.00485 & -0.00551 \\ -0.00485 & 0.0000405 & 0.0000418 \\ -0.00551 & 0.0000418 & 0.000250 \end{pmatrix}$ 

This matrix is nearly the same as that calculated by PROC PROBIT.

An important purpose for fitting the probit model is to calculate lower confidence bounds on safe concentrations by use of Fieller's theorem (Finney [11], pp 78-79). That is, we wish to calculate a lower bound on the concentration such that  $\Phi(\beta_0 + \beta_1 \text{ conc}) = L$ , where L is some specified response rate. Such lower confidence bounds, at (one sided) confidence level 97.5 percent are a standard part of the PROC PROBIT output. They are given in Figure XIV.12 for the Holcombe and Phipps compound D data with untransformed concentration. We indicate below how to calculate these bounds for any confidence level, based on the output from PROC NLIN. The theory underlying these calculations is sketched in Appendix AXV.

The fitted model is 
$$\hat{p}(conc) = \hat{c} + \hat{c}\Phi(\beta_0 + \beta_1 conc)$$
.

We wish to construct a 1 -  $\alpha$  level confidence interval on that CONC such that

$$\Phi(\beta_0 + \beta_1 \text{conc}_L) = L$$

where L is specified (e.g. 0.01, 0.05, 0.10 etc.). L represents the response level attributed to toxicant (i.e. over and above background).

The point estimate, conc, is

$$\operatorname{conc}_{\mathrm{L}} = (\Phi^{-1}(\mathrm{L}) - \hat{\beta}_{0})/\hat{\beta}_{1} \equiv (f_{\mathrm{L}} - \hat{\beta}_{0})/\hat{\beta}_{1}.$$

Let the asymptotic variance-covariance matrix of  $(\beta_0, \beta_1)$  be denoted as

$$\begin{pmatrix} g & h \\ h & j \end{pmatrix} \equiv \begin{pmatrix} \widehat{\operatorname{Var}}(\beta_0) & \widehat{\operatorname{Cov}}(\widehat{\beta}_0, \widehat{\beta}_1) \\ \widehat{\operatorname{Cov}}(\widehat{\beta}_0, \widehat{\beta}_1) & \widehat{\operatorname{Var}}(\widehat{\beta}_1) \end{pmatrix}$$

A 1 -  $\alpha$  confidence interval on conc<sub>1</sub> is shown in Appendix AXV to be

$$\sum_{\text{conc}_{L}} \varepsilon \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}$$

where  $A = \hat{\beta}_1^2 - j z_{\alpha/2}^2$ 

$$B = 2[\hat{\beta}_{1} (\hat{\beta}_{0} - f_{L}) - hz_{\alpha/2}^{2}]$$
$$C = [(\hat{\beta}_{0} - f_{L})^{2} - gz_{\alpha/2}^{2}]$$

 $z_{\alpha/2}$  is the upper  $\alpha/2$  point of the standard normal distribution. The quantities  $\hat{\beta}_0$ ,  $\hat{\beta}_1$ , g, h, j are obtained as output from NLIN. The results of the calculations are given below.

Holcombe and Phipps Compound D Untransformed Concentration						
		om calculation of upper and				
confi	dence bounds of	on various percentiles of 1	PROBIT fit by Fieller's			
		Theorem				
L	$f_L \equiv \Phi^{-1}(L)$	Lower 95% conf.limit	upper 95% conf.limit			
0.01	-2.345	22,958	69,385			
0.03	-1.88	39.00	78.23			
0.05	-1.645	47.07	82.73			
0.07	-1.48	52.73	85.91			
0.10	-1.28	59.96	89.78			
0.15	-1.03	68.06	94.66			
0.20	-0.83	74.80	98.62			
0.25	-0.68	79.82	101.63			
0.50	0	101.69	116.15			

These confidence bounds are seen to agree very closely with the bounds calculated by SAS PROC PROBIT and which appear in Figure XIV.12.

The previous dose response model fitted by use of PROC NLIN was a repeat of a model that has also been fitted by PROC PROBIT. Comparisons of the PROC PROBIT and PROC NLIN outputs verified that dose response models can in fact be fitted by nonlinear regression programs and helped to interpret the various features of the PROC NLIN output.

We now consider the fitting two dose response models that cannot be fitted by PROC PROBIT. This of course is the reason for considering the application of PROC NLIN for dose response estimation in the first place. We first consider the logistic model and then look at an alternative to Abbott's correction for accounting for background response.

The logistic model is a commonly used dose response model and gives results very similar to probit fits, at least between the 2nd and 98th percentiles. The logistic c.d.f. is

$$F(x) = \frac{e^{x}}{1 + e^{x}} \qquad -\infty < x < \infty$$

and is a symmetric unimodel distribution like the normal, but has heavier tails. We fit the dose response model

$$p(conc) = c + \bar{c}F(\beta_0 + \beta_1 conc)$$

in direct analogy to the probit fit that appears in Figures XIV.11.

The results of the Gauss-Newton interative process are given in Figure XV.6. The Marquardt algorithm converges whereas the Gauss-Newton algorithm does not because the Marquardt algorithm can take smaller steps and is more flexible in direction. However both algorithms arrive at nearly the same parameter estimates. The summary of the fitted dose response model appears in Figure XV.7. We can compare this fit to the probit fit in Figure XV.4.

We see that both the logit model and the probit model fit the data quite well (residual sums of squares are quite small). The background mortality rate is estimated to be about 0.07 by each model. The asymptotic variance-covariance matrix of the logit fit parameters is estimated to be

$$\widehat{\operatorname{Var}} = \frac{1}{0.093958} \begin{pmatrix} 0.48086 & 0 & 0 \\ 0 & 0.00388 & 0 \\ 0 & 0 & 0.00518 \end{pmatrix} \begin{pmatrix} 1.0000 & -0.988351 & -0.550366 \\ -0.988351 & 1.0000 & 0.520791 \\ -0.55036 & 0.520791 & 1.0000 \end{pmatrix} \cdot \\ \begin{pmatrix} 0.48086 & 0 & 0 \\ 0 & 0.00388 & 0 \\ 0 & 0 & 0.00518 \end{pmatrix} = \begin{pmatrix} 2.46095 & -0.019626 & -0.01459 \\ -0.019626 & 0.000160 & 0.000111 \\ -0.01459 & 0.000111 & 0.000286 \end{pmatrix} \\ \equiv \begin{pmatrix} g & h & * \\ h & j & * \\ * & * & * \end{pmatrix}$$

We now apply Fieller's procedure for calculating lower end upper confidence bounds on distribution percentiles of the dose response fit. We need only modify the calculations done for the probit fit by defining

$$f_{L} = F^{-1}(L) = \ln \frac{L}{1-L} \equiv \log t (L)$$

and using the appropriate point estimates and variance-covariance matrix.

Holcombe and Phipps Compound D Untransformed Concentration				
Results from calculations of upper and lower 95 percent				
confidence bounds on various percentiles of LOGIT fit by Fieller's				
Theorem				

L	$f_L \equiv \ln \frac{L}{1-L}$	lower 95% conf.limit	upper 95% conf.limit	LOGIT point estimate	PROBIT Point estimate
0.01	-4.5951	2.515	66.209	45.31	52.931
0.03	-3.4761	26.859	78.151	61.205	64.231
0.05	-2.9444	38.390	83.863	68.757	69.942
0.07	-2.5867	46.127	87.726	73.838	73.952
0.10	-2.1972	54.526	91.957	79.370	78.812
0.15	-1.7346	64.454	97.031	85.940	84.887
0.20	-1.3863	71.879	100.900	90.888	89.747
0.25	-1.0986	77.967	104.142	94.974	93.392
0.50	0	100.441	117.293	110.578	109.916

The point estimates of the probit and logit fit percentiles are presented side by side for comparison. Except at L = 0.01 they are very close and even at L = 0.01 they are similar. The situation is a bit different with respect to confidence bounds on the safe concentration. The logit confidence bounds are to be compared with the probit confidence bounds. We see that the <u>upper</u> logit and probit confidence bounds are very similar at each percentile. However the <u>lower</u> confidence bounds for the logit fit are somewhat lower than the lower confidence bounds for the probit fit at the low distribution percentiles.

For L $\geq$ 0.07, the lower confidence bounds for the logit and probit fits are rather similar, the lower logit bounds being constantly smaller than the lower probit bounds. For L below 0.05 this phenomenon is accentuated, especially at L = 0.01. This the region in which mortality due to background is the first order effect while toxicant related mortality is secondary. Thus the data and the fitted model reflect primarily the background effects and provide little direct evidence about toxicant related mortality. Since the tails of the logistic distribution are heavier and steeper than the tails of the normal distribution, changes in parameter values perturb percentile estimates in the normal distribution. Thus the lower logistic confidence limits become much wider than the corresponding lower normal limits as  $L \rightarrow 0$ . This phenomenon holds very strongly in this example at L = 0.01 and to some extent at L = 0.03, 0.05. In this region the data provide little basis to choose between the logit and probit fits. Both models fit the data well and yield very similar point estimates. Thus we learn the following lesson:

THE LOWER CONFIDENCE BOUNDS ON "SAFE" CONCENTRATIONS CORRESPONDING TO LOW PERCENTILES OF THE DOSE RESPONSE DISTRIBUTION MAY BE SENSITIVE TO THE PARTICULAR FORM ASSUMED FOR THE DOSE RESPONSE RELATION, EVEN THOUGH SEVERAL MODELS MAY FIT THE DATA EQUALLY WELL AND PROVIDE SIMILAR POINT ESTIMATES OF PERCENTILES. THE DATA MAY NOT BE SUFFICIENT TO DISTINGUISH BETWEEN THE MODELS.

This phenomenon is observed quite frequently in very low dose extrapolation based on results of carcinogenesis experiments. However in those applications the extrapolation is much more extreme than in fish toxicology applications. However this example illustrates that even in fish toxicology situations the inference about safe dose can be very sensitive to model assumptions, even at the first to the third percentile. The extent of background effects may prevent us from distinguishing among alternative models which fit the data about equally but which yield qualitatively different inferences about safe concentrations corresponding to low distribution percentiles.

To partially circumvent this problem we consider an alternative approach to dose response estimation based on fewer assumptions about the shape of the response distribution. See the following section for a discussion on this nonparametric approach to dose response estimation.

We consider now a third example of fitting dose response models by means of nonlinear regression. This example involves a nonstandard model which provides an alternative to Abbott's correction to account for background response. Abbott's correction is appropriate when the mechanism associated with background effects is <u>independent</u> of the mechanism associated with toxicant effects. For example toxicant mortality may be due to chemical effects whereas background mortality may be due to increased handling of the fish.

However Stephan [42] criticizes the assumption that the control mortality mechanism is totally independent of the toxicant mortality mechanism. He states that stressing the fish during the acclimation or testing periods may make them more susceptible to the toxicant. Thus background effects may act like additions to the toxicant concentrations. Stephan suggests <u>not</u> correcting for control mortality when assessing the effects of various toxicant concentrations.

An alternative way to reflect the dependence between background and toxicant mortality mechanisms is to fit a model which reflects the fact that background may function as an addition to the effective toxicant concentration. Assume that background effects are equivalent to an addition of c  $\mu$ g/liter in toxicant concentration. The quantity c is a model parameter to be estimated from the data. (Note that the usage of the notation c is completely different in this example than in the previous examples in this section. Here it is being used as a concentration whereas in previous examples the symbol c represented a probability.

Consider the following models based on an assumed normal dose response curve.

- (1)  $p(conc) = \Phi[\beta_0 + \beta_1(conc + c)]$
- (2)  $p(conc) = \Phi[\beta_0 + \beta_1(\log_{10}(conc + c) 3.0)].$

These models are to be fitted to the data by maximum likelihood estimation, based on binomial distribution theory. The parameters  $\beta_0$ ,  $\beta_1$ , c are to be estimated from the data. The first model is over parameterized, in that  $\beta_0$  and c cannot be separated from one another. Thus to fit model (1) we fit

$$p(conc) = \Phi[(\beta_0 + \beta_1 c) + \beta_1 conc] \equiv \Phi(\alpha_0 + \beta_1 conc)$$

using PROC PROBIT with untransformed concentration and no "background" effect included.

The centering constant 3.0 in model (2) is intended to reduce the correlation between  $\hat{\beta}_0$ ,  $\hat{\beta}_1$  thereby improving the covergence properties of the fitting algorithms. To fit model (2) we carry out a maximum like-lihood analysis using PROC NLIN. The output from this analysis is shown in Figures XV.8, XV.9. The Marquardt algorithm again achieves convergence whereas the Gauss-Newton algorithm does not. Note however that the Gauss-Newton algorithm attains a smaller residual sum of squares due to the difference in weighting. (The distinction between attaining the smallest residual sum of squares and attaining a stationary point corresponds to the difference between minimum chi square estimation and maximum like-lihood estimation. This distinction is discussed in Jennrich and Moore [43], page 10, and in the BMDP manual [27]. Both of these methods are asymptotically equivalent. The summary of the Marquardt algorithm fit is presented in Figure XV.9. The residual sum of squares represents a chi square test for goodness of fit of the model. We see that

residual chi square = 59.32 with 3d.f.

Thus the model does not seem to fit the data. We break down this residual chi square into individual cell components to determine whether the large residual chi square represents consistent lack of fit or the contribution of a single aberrant cell.

We see that the expected frequencies,  $N\hat{p}$ ,  $N\hat{q}$ , are quite large under the model fit and that there is a systematic discrepancy between model and data. Namely the model underestimates at the lower and upper ends and overestimates p in the middle.

We thus conclude that model (2) is not appropriate for this set of data. However this or similar models may be appropriate for other sets of data. The point is that the use of nonlinear regression techniques to fit dose response curves greatly expands the variety of models that we can fit to the data.

Since model (2) does <u>not</u> fit the data well we do not use it to calculate lower confidence bounds on "safe" dose. However these calculations can easily be made by use of asymptotic maximum likelihood theory.

**1** 980 FITTING DUSE RESPONSE MOUELS TO MOLLOMBE AND PHIPPS COMPOUND D DATA BY MEANS OF MONLINGAR RIGRESSION Embryd and Fry Murtality Uata Merged, with Average concentration data--this is now basic uata

• • •

;

Ł

100 100 100 100 100 100 100	OBS	TRT	EMB SUM	FRYSUM	<b>DE ADE SUM</b>	DEADFSUM	CONC MEAN	
3 200 100 63 5 200 100 63 13 78 13 78 13 78	 1		202	100	•	-00	000.0	•
5 200 100 100 100 100 100 100 100 100 100	1177) -	410-	200		- 10	U CD	000 · 44	
	 	* AC 1	200	1000	- 28-	<b>EC</b>	128-150	

Figure XV.1 Holcombe and Phipps compound D mortality and concentration data -- pooled

10-5944 S-P1	1		• • •	•	
	•			r	( # # 1
2N = 44	t				
CONCENTRATIO	•	1	· · ·		
CONCE ARLF D	່. ເກ	0-0-4-00-	14-00-0000	00000	
ÌT♥S CÖRRÉCTIONUNTRANSFORMED CONCENTRATIO Grid search dependent variarle derufsum	RESIDUAL SS	5000000 100000 100000 100000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 1000000		00000000000000000000000000000000000000	
PENDEN			10000-400-00	1000-400	
10N101	WEIGHTED	, , ,			
JRRECT SARCH	C N	200100			
LITES CORREC Grid Search		000000		000000	
<b>a</b> i -	ßl		000000000000000000000000000000000000000		
FLT MITH ABB	BÖ	00000r	100000	1000000	
- <b>J</b>			111111	111111	
RUBIT MUDE Non-Linfar	.   	•			
		•			
PRG NC					
		i • •	1		ŀ

Main 2, 1480

r.

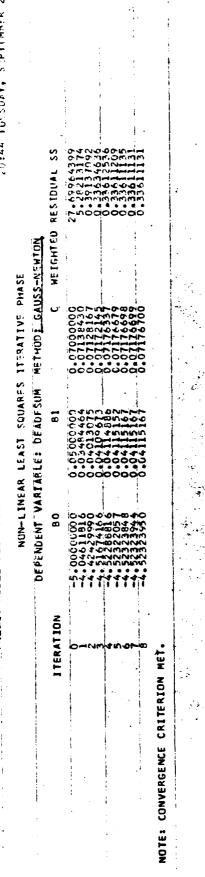
,

;

t

L'AND CALL

Output from SAS proc NLIN applied to Holcombe and Phipps data -- untransformed concentration Figure XV.2





MODELS TO HULCOMBE AND PHIPPS CUMPOUND D DATA BY MEARS OF MONLIN AR WICKLEETEN 3 Mudel fit with abbuiltes currection--untransformed concentratien 20:44 TUSSDAY, SUPTEMBER 2, 1980 FITTING DUSE RESPONSE

FITING DOSE RESPONSE MODELS TO HOLCOMBE AND PHIPPS COMPOUND D DATA BY MEANS LE VUMLINHAR FREFECTION . . . . . PRUBIT MODEL FIT MITH.ABBDITMS CORRECTION--UNTRANSFORMED CONCENTRATIC 20144 TUBER NODEL FIT MITH.ABBDITMS CORRECTION--UNTRANSFORMED CONCENTRATIC ASYMPTUTIC 95 & CONFIDENCE INTERV DEPENDENT VARIAGLE DEADESUM WEIGHTED MS 21367553.5597845 0.11203710 1 -5-34656562 0-03436492 0-05436492 ASYMPTOTIC CORRELATION MATRIX OF THE PARAMETERS 451540 000000 WEIGHTED SS 64102660.67813536 4498.36044442 81 414906 NON-LINEAR LEAST SQUARES SUMMARY STATISTICS 0.00528838 0.00528838 ASYMPTOTIC STD. ERRUR -----80 40 mm o ۰. ۲ REGRESSION RESIDUAL UNCORRECTED TUTAL (CORRECTED TOTAL) 220 ESTIMATE 8 ... 8 ... 1 ... SOURCE PARAMETER 010

Output from SAS PROC NLIN applied to Holcombe and Phipps data -- untransformed concentration Figure XV 4.

193

-----

-- untransformed concentration Figure XV.5 Logit fit to Holcombe and Phipps data using SAS PROC NLIN

NOTE: USING THE ABOVE PARAMETERS, THE PARAMETER MODIFICATION VECTOR (DELTA) WAS HALVED TO TIMES. NO REDUCTION IN THE RESERVEL SE WAS FOUND (CONVERGENCE ASSUMED). WEIGHTED RESIDUAL 55 METHOD: GAUSS-NEWTON NON-LINEAR, LEAST SQUARES IT RATIVE PHASE 0.030-000 1.7347235-18 1.7347235-18 1.7347295594 0.0565784594 0.065784594 0.065982377 0.06904976 0.06904976 υ DEPENDENT VARIARLE: DEADESUM 8 0 ŝ I TERATION 0 1 ļ í ÷ 194

¢, FITTING DOSE RESPONSE MUDELS TO HOLCOMPE AND PHIPPS COMPANNO D'DATA BY M'AME OF NOVELN AF PEGGERSEUN 12 Lugistic Mudel Fit with Abbottms, conkection—-Untransformed curcentration 20144 Thesury, Surtember 2, 1980

			•													
														:		
			;											;		
	·	HEIGHTED RESIDUAL SS	164.29824192	I37.56748230	40.14723427	20-32299412	3.08186413	0.45655573	0.25230946	0-28187490	0-28187365	0.28187383	0.26187382	0.28187381	0.2 6 187380	0.28187380
HÀSE	QUARDI	WEIGHI	£ 1													
ES ITERATIVE P	METHUD: MAR	<b>ں</b>	0.000000	3.469447E-18	3.469447E-18	0-08107047	0.06662428	0.06981510	0.06896285	0.06904992	0.06905248	0.06905251	0.06905254	0.06905257	0 • 0 6 9 0 5 2 6 0	0.06905263
NUN-LINEAR LEAST SQUARES ITERATIVE PHÁSE	UEPENDENT VARIABLE: UEADFSUM METHUD: MARQUARDI	. 81	0.0110000	U-U1191808	U.U2438035	0.03863356	0.05423863	0-06609747	0.06991643	0.07040193	0.07040485	0.07040488	0.07040491	0.07040494	0+010+0497	0.07040500
NUN-LINI	UEP ENDENT VAI	09	-2.0000000	-1.41496226	-2-85943212	1625524-	-5-96693004	-7 -2 109 2454	-7.72383965		-7.76521738	-7.18522188	-7.78522632	-1.7 6523069	-7.18523502	-7.78523930
	•	ITERATION	0	-		c)	*	ĥ	٩		80	3	10		12	ET -
	•								-	•						

FITTING WUSE RESPUNSE MOVELS TO MOLCOMBE AND PHIPPS COMPOUND D DATA BY MFANS OF NUNLINEAR REGPESSIUN Lugistic mudel fit with Abbuilt's correction---Uniransformed concentration 20:47 Tuesday, September 2, 1980

NGTE : LOWERGENCE CRITERION NET.

> untransformed concentration ł Logit fit to Holcombe and Phipps data using SAS PROC NLIN Figure XV.6

FITTING UOSE RESPUNSE MUDELS TO MULLUMBE AND PHIPPS COMPOUND D DATA BY MEANS OF NONLINEAR REGRESSION 15 Logistic model fit with abbuiltes currection---Untransfurmed concentration 20:47 Tuesday, September 2, 1940

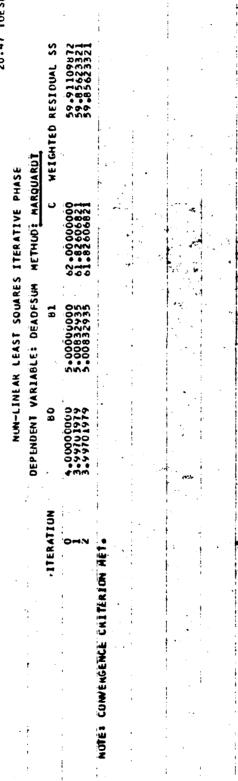
and a set DEPENDENT VAPT NUN-LINEAR LEAST SUURSES SUMMARY STATISTICS

	SOURCE	SOURCE DE WEIGHTED SS WEIGHTED MS	WEIGHTED MS
· · ·	REGRESSION RESIDUAL Uncorrected tutal (corrected tutal)	3 291669-41199674 3 291669-69387054 6 291669-69387054 5 4424-13968364	97223 <b>.1</b> 3733225 0.09395793
PAKAMETER	ESTIMATE	ASYMPTOTIC STU. ERKUR	SYMPTOT FIDENCE
580 580		0.448085523 0.00388025 0.00517702	-9-31556108 -6.25491753 0.05805609 0.08275391753 0.05257678 0.08552849
•	AS IMPTOTIC	ASTMPTOTIC CORRELATION MATRIX OF THE PARAMETERS	DARAMETERS
	<b>8</b> 97 97	1.000000 -0.988351 -0.550366 -0.4509251 1.000000 -0.550366 -0.4550346 1.000000	550366 550366 5503791

Figure XV.7 Logit fit to Holcombe and Phipps data using SAS PROC NLIN -- untransformed concentration

ŀ.





FITING DOSE RESPONSE MODELS TO HULCOMBE AND PHIPPS CUMPOUND D DATA BY MEANS OF NUMLINEAM REGRESSION Y PROBIT MUDEL WITH BACKGROUND TREATED AS AUDITIVE DOSE--LOGARITHMIC CONCENTRATION 20:47 TUESDAY, SEPTEMBER 2, 1960

			SUUAKES SU	COLICIIAIS TAATTOS			VAR ANCE URAUPSUM
	SOURCE REGRES UNCORRI	SION SION ECTED TCTCD	TUTAL DTAL )		WEIGHTFD 55 1344.22184920 59.32465332 1403.54659253 1005.24690533		WT 1 GHT 50 MS
PARAMETER Bo CI		ESTIMATE ESTIMATE 4.13381039 5.29783127 65.06065330	MATE 1039 3127 5330	ASY MPT STD., F 3.0254 7.23554 82.0107	TUTIC EFROR 149811 149811 113320	-195.49483734	ASYMPTOTIC 95 % CONFIDENCE INTERVAL 64 15 7624501 326.355981170
		WASA	SYMPTOTIC CC	CORRELATION MATRIX	ATRIX OF THE	IE PARAMETERS	
			CB30	1.00000 0.965543 0.883880	0.965543	0.883880	· · · · · · · · · · · · · · · · · · ·
	•					i - 1	
9.0 	ny (ite) Conc	Breek	Brakdown of resid	recident ou	<b>Suer</b> e 2 2	\$\$ <b>X</b>	(v.w)
0.0	.000	100	9	2510.0	/368	121	/8.725
27.2	26	100	90	0.0111	8.176	8-088	0.095
14,905	505	00/	00	0.1723	122.51	1424	5.978
र्ज्यत्/ स्	20	2007	54	0.3364	38.646	22.525	19.0%1 10.068
	8	100.	/00/	0.1978	81.78	9.173	11.379

Figure XV.9 Output from SAS PROC NLIN to fit nonstandard dose response model to data

198

10 PT-MAE2, 1980

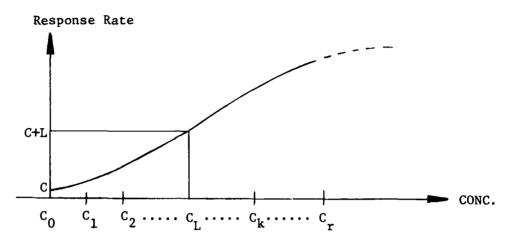
HOLLOWOE AND PNITATE COMMUND FRY MUNTALITY DATA

•

XVI. NONPARAMETRIC LOWER CONFIDENCE BOUNDS ON SAFE CONCENTRATIONS

In this section we again consider the estimation of "safe" concentrations based on fitted dose response curves. We wish to estimate a lower confidence bound on that concentration,  $c_r$ , for which the respo-

nse rate is no more than L greater than the control group rate. The value of L is specified by the user. We present this situation pictorially below.



The dose response curve is assumed to be concave upward at the lower percentiles. For a logit or probit fit this would be below the median of the distribution. The solid portion of the illustration represents the region where the dose response curve is concave upward.

The upper bound on the upward concavity region is denoted UCR. UCR is specified by the user.

Let  $c_0 \le c_1 \le c_2 \le \ldots \le c_r \le UCR$  denote the test concentrations (treatment and control) in the upward concavity region.  $c_0$ , the control group concentration, would often be 0.

The standard method of estimating  $c_L$  by means of dose response curves is to assume a specific form for the dose response curve such as probit or logit in concentration or in log concentration and then fit the model by means of maximum likelihood estimation, based on all the data. SAS PROC PROBIT or a nonlinear regression package can be used to fit such models.

The procedure discussed in this section has a number of important differences from these standard parametric dose response models. Among these are

> 1. Inferences about safe concentrations can sometimes be rather sensitive to the particular form of the dose response curve

assumed. Yet it may not be possible to distinguish among such competing models based on the data at hand. The need for such strong parametric assumptions is alleviated with the procedure in this section.

- 2. Once a functional form is chosen for the dose response curve, there is still the question of the dose metameter. Different lower bounds may result depending on whether the probit (say) model is chosen with respect to concentration, log (concentration), or some other function of concentration. There is no need to worry about the specific form of the dose metameter with the nonparametric procedure of this section.
- 3. The parametric dose response models assume a specific functional form for the correction for background responses; Abbott's formula (Finney, [11], ppl25-126) is commonly used. However estimates of the low percentiles of the dose response curve can be sensitive to the specific form of background correction used. The procedure in this section does not require the specification of any particular functional form for background response.
- 4. The standard method of fitting a parametric dose response model is by means of maximum likelihood estimation. The theoretical justification is based on the assumptions of large samples and asymptotic normality. These assumptions may not be entirely satisfied in the case of relatively small sample sizes or of many response group probabilities at or near 0 percent or 100 percent. By contrast, the method discussed in this section is based on <u>exact</u> small sample theory and so is appropriate irrespective of small sample sizes or extreme response rates. We also present an alternative confidence bound calculation which may yield higher lower bounds, however this alternative approach depends on large sample theory and asymptotic normality. Both estimates are routinely calculated by our computer program.
- 5. The standard parametric probit or logit dose response curve fits utilize the information from all the test concentrations. including those high concentrations at the upper end of the dose response curve, far away from the safe concentration. In fact, these upper concentrations, with high response rates are very instrumental in determination of the slope estimate and associated precision estimate. These high concentrations, thus carry considerable weight, through the specification of the model, in estimating response behavior at the low concentrations. This is not desirable, since the same functional form may not be appropriate throughout the entire range of concentrations. By contrast, the method in this section uses information only from those concentration groups where the dose response curve is concave upward. This is generally in the region below the median of the dose response curve.

One assumption made throughout this section is that the response results can be modelled with the binomial distribution within each tank and that there is no evidence of tank to tank heterogeneity within treatment groups. The responses can then be pooled across tanks within treatment groups and we can assume a single binomial distribution for the pooled responses within each treatment group. This distributional assumption is made in our program.

What do we do if there in fact is evidence of heterogeneity within tanks? There are three approaches to account for this situation. See Section IX for detailed discussion. Briefly these approaches are:

- Carry out analyses on a per tank basis rather than on a per fish basis. This is the approach that is currently being used by some researchers. However this approach greatly reduces the number of degrees of freedom available for analysis. I feel it is too conservative.
- 2. Fit distributional models to the data that explicitely account for such tank to tank heterogeneity. Several such models proposed are the beta binomial model (Williams [21]) or the correlated binomial model (Kupper and Haseman [22]). These models generalize the binomial distribution model and can be incorporated into a dose response curve estimation model. The fitting would be by maximum likelihood estimation and the statistical inferences would be based on asymptotic normal distribution theory.
- 3. We can adjust the data to reflect the within tank correlation. Namely tank to tank heterogeneity reflects itself as variation in response rate from tank to tank within treatment groups. This can also be regarded as correlation of responses within individual tanks. The effect of such correlation is to reduce precisions of estimates as compared to what they would be in a binomial model, since the correlations will usually be positive. This reduced precision can be accounted for in a workmanlike manner by reducing the effective sample size within each tank. Namely suppose we have 4 tanks per group, 25 fry per tank, and responses 1, 3, 8, 7 respectively. The effect of assuming a binomial model would be to pool data across tanks within groups, so that we have 100 fry and 19 responses. Thus  $\hat{p} = .19$  and  $\sqrt{Var(\hat{p})} =$  $\sqrt{(.19)(1 - .19)/100} = .039$ . However correlation within tanks inflates the variability by a factor h. (h>1). Reduce the assumed sample size within each tank from 25 to 25/h. Correspondingly reduce the effective number of responses within each tank to 1/h, 3/h, 8/h, 7/h, for a tot<u>al of 19/h.</u> Thus  $\hat{p} = (19/h)/(100/h) = .19 \text{ still.}$  However  $\sqrt{\hat{V}ar(\hat{p})}$  $\sqrt{(.19)(1 - .19)/(100/h)} = .039\sqrt{h}$ . We then disregard the

tank to tank heterogeneity and utilize the binomial based procedures, such as the computer program discussed in this section.

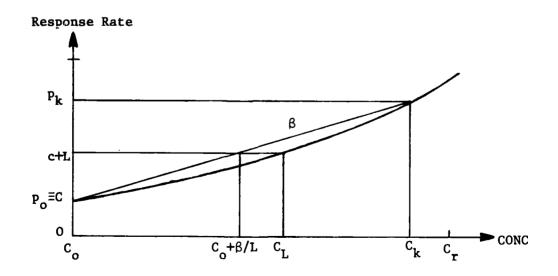
This method of adjustment to effective sample sizes is approximate and somewhat crude, however it has the advantage of simplicity and no special computer programs need be used. Namely, the same methods that are utilized in the absence of tank to tank heterogeneity are used in the presence of such heterogeneity, only with reduced sample sizes. This allows for the use of standard analysis tools in nonstandard situations.

Looked at from the perspective of reducing sample sizes to an effective sample size, carrying out analyses on a per tank basis is like reducing the effective sample size in a tank all the way down to 1. I feel that this is going a bit too far.

In particular the methods discussed in this section can be utilized following such adjustments to account for tank to tank heterogeneity. Thus from now on in this section we ignore the question of tank to tank heterogeneity within groups and discuss our procedure, based on binomial distribution theory, as if there were no tank to tank heterogeneity.

We now consider the details of the nonparametric dose response procedure.

Assume that k is such that  $c_0 \le c_L \le c_k \le UCR$ . The value of k is specified by the user of the program.



Let  $p_0$ ,  $p_k$  denote the true response rates at  $c_0$ ,  $c_k$  respectively. Draw a chord joining the points  $(c_0, p_0)$ ,  $(c_k, p_k)$ . Let  $\beta$  denote the slope of this chord. Thus

$$\beta = \frac{p_k - p_o}{c_k - c_o}$$

Upward concavity implies that the chord lies above the dose response curve throughout the region  $(c_0, c_k)$ . The concentration at which this chord crosses the value c + L on the response scale is  $c_0 + L/\beta$ . If  $c_0 = 0$ , this concentration is  $L/\beta$ . Thus

$$c_{o} + L/\beta \leq c_{L}$$

and so it we can place a lower confidence bound on  $c_0 + L/\beta$ , then this also serves as a nonparametric lower confidence bound on  $c_r$ .

Let  $\hat{p}_0$ ,  $\hat{p}_1$ , ...,  $\hat{p}_r$  be the estimated response rates based on binomial theory at concentrations c, c<sub>1</sub>, ..., c<sub>r</sub> respectively. Then  $N_i \hat{p}_i \sim B(N_i, p_i)$  i = 0, 1, ..., r when  $N_i$  is the number of animals on test in the i-th treatment group pooled across tanks and  $p_i$  is the true response probability within the i-th treatment group.

Let  $p_0$  denote a lower confidence bound on  $p_0$ ,  $p_k$  denote an upper confidence bound on  $p_k$ . Such exact lower and upper confidence bounds were derived by Clopper and Pearson [46] and are valid for small sample sizes. Expressions for them are contained in a number of sources, including Hollander and Wolfe [47], pages 23, 24. Charts for these confidence intervals are given in a number of places, including Dixon and Massey [13] pp 501-504. Expressions for these confidence bounds are given in Appendix AXVI.1.

An upper confidence bound,  $\beta_{\eta}$ , on  $\beta$  is

$$\beta_{U} = \frac{\hat{p}_{k} - \hat{p}_{o}}{c_{k} - c_{o}}$$

Thus a lower confidence bound on c is c +  $L/\beta_U$ . This confidence bound is valid in small samples.

The results in the concentration groups  $c_{k+1}$ ,  $c_{k+2}$ ,  $\cdots$ ,  $c_r$  can be used to improve on the confidence bound discussed above. The details of this procedure, along with a description of an alternative confidence bound, valid in large samples, are discussed in the writeup "A Computer Program to Calculate Nonparametric Lower Confidence Bounds on

Safe Concentrations in Quantal Response Toxicity Tests" by Feder and Sherrill [41]. This writeup also describes in detail the use of a computer program to implement this procedure. This document is included as an appendix to this section. We illustrate the results of this program by example in the remainder of this section and compare the results of the nonparametric estimates of safe concentration with those based on the logit or probit fits.

We first consider the DeFoe compound C fry mortality data.

We have seen from previous sections that there is no evidence of tank to tank heterogeneity within treatment groups.

The various portions of the computer program output are numbered and we discuss them in detail.

As a number of the parameter values used in the program were chosen rather arbitrarily (e.g. UCR) we should regard the output as illustrative of the algorithm's working rather than as a definitive answer in this particular case. We know that the algorithm will give conservative answers. The question is just how conservative the algorithm is.

We know from the preliminary plots and tests of homogeneity that there is no concentration related trend in embryo mortality. Such preliminary analyses are very important to carry out, in order to gain an understanding of the structure of the data. This helps us to interpret the results of the procedures such as the one in this section.

The numbered descriptions below refer to the similarly numbered descriptions in the computer printout for the DeFoe fry mortality data.

- 1. The title of the output. This title appears at the head of every page.
- 2. The basic data are presented for each tank within each concentration group (treatment and control). Numbers of fry per tank, numbers survived, and toxicant concentration are given.
- 3. The number and the proportion of dead fry within each group are given. These values are calculated by pooling across tanks within groups.
- 4. Basic parameter values for the procedure.
  - L = response rate, over and above the control rate, at the "safe" concentration.

k = the index of assumed upper bound on the "safe" concentration, c<sub>1</sub>.

i.e.  $c_0 \leq c_L \leq c_k$ .

UCR  $\equiv$  upper limit of the concave upward region in the dose response curve.

A number of confidence bound calculations are carried out for differing combinations of (L, <sup>k</sup>, UCR). In this example UCR is specified as 50. This places it just above the fifth treatment group. ( $\hat{p}_5 = .225$   $\hat{p}_6 = 1.000$ ).

Thus r = 5 and  $c_r = 48.3074$  in this problem.

L and k are varied.

L = .01, .05, .10k = 3, 4, 5.

L = .01, k = 3, UCR = 50.0 in the first calculation.

- 5. Simultaneous confidence interval adjustments are made in this run by means of Bonferroni's inequality with familywise confidence level 0.95. Thus all small sample confidence intervals are calculated at individual confidence level 1 - (.05/4) = 0.9875.
- 6. Upper and lower confidence intervals are calculated at each concentration group. These are exact, small sample confidence intervals, calculated as discussed by Clopper and Pearson using the expressions in Appendix AXVI.1.
- 7. Straight line approximations to the dose response curve are calculated using the combinations of treatment groups shown. The specific method of calculation of the slopes is discussed in the program documentation in Appendix AXVI.2. (Feder and Sherrill [41]). For each combination of concentrations CONC MEAN is the arithmetic average of the concentrations, slope (normal approx) and slope (small sample) are the calculated values of  $\beta_U$  based on either asymptotic theory or exact small sample theory. See the program documentation for details.
- 8. Lower confidence bounds on  $c_L$  are calculated using the minimum of the slopes in paragraph 7 (in this case .0056 for the normal approximation and .0084 for the exact approach). The values given under "calculated safe dose" are  $c_1 + L/\beta_U$ . These are .0494 + .01/.0056 = 1.845 and .0494 + .01/.0084 = 1.24 respectively for the normal theory and small sample calculation.

Since we have taken k = 3,  $c_k = c_3 = 5.9762$  is an upper bound on  $c_r$ , by assumption. Thus

 $\hat{c}_{L}(normal) = min(1.85, 5.98) = 1.85$  $\hat{c}_{L}(small sample) = min(1.24, 5.98) = 1.24$ 

Since the response rates are so extreme in 5 of the 6 groups (i.e. close to 0 or 1) we have small expected frequencies in many of the cells and asymptotic normal theory is suspect here. We will thus confine attention to calculations based on exact small sample theory for the remainder of the section.

- 9. We modify the parameters defining the procedure. The values k, UCR remain at 3, 50.0 respectively, however L is changed to 0.05. We thus define the "safe" concentration as that which yields a response rate of .05 above control.
- 10. Proceeding through the same calculations as before we find that the minimum slope (small sample) is .0084. Thus

 $c_1 + L/\beta_{TT} = .0494 + .05/.0084 = 5.98$ 

Since  $c_k \equiv c_3 = 5.9762$  we estimate

$$\hat{c}_{L} = \min (c + L/\beta_{U}, c_{k}) = \min(5.98, 5.98) = 5.98$$

- 11. We now alter L to 0.10, leaving k and UCR as before.
- 12.  $\hat{c}_{L} = \min(c_{1} + L/\beta_{U}, c_{k}) = \min(.0494 + .10/.0084, 5.9762) = \min(11.954, 5.9762) = 5.9762.$

Thus  $\hat{c}_L$  is constrained by overly conservative assumption about  $c_L$ .

- 13. We now set k = 4 ( $c_{L} = 14.8125$ ) and set L back to 0.01
- 14. We now calculate slopes, but we have fewer to work with. Namely we use  $c_1$ ,  $c_4$ ,  $c_5$  in various combinations.
- 15.  $\hat{c}_{L} = \min(c_{1} + L/\beta_{U}, c_{k}) = \min(.0494 + .01/.0082, 14.8125) = 1.2625$

- 16. We now change L to 0.05, leaving the other parameters as before.
- 17.  $\hat{c}_{L} = \min(c_{1} + L/\beta_{U}, c_{k}) = \min(.0494 + .05/.0082, 14.8125) = \min(6.1145, 14.8125) = 6.1145$
- 18. We change L to 0.10 leaving k, UCR unchanged.
- 19.  $\hat{c}_{L} = \min(c_{1} + L/\beta_{U}, 14.8125) = \min(.0494 + .10/.0082, 14.8125) = \min(12.1796, 14.8125) = 12.1796$
- 20. We now change k to 5 and set L back to 0.01. Thus  $c_k \equiv c_5 = 48.3074$ .
- 21. L = 0.01, k = 5

 $\hat{c}_{L} = \min(c_{1} + L/\beta_{U}, c_{k}) = \min(.0494 + .01/.0080, 48.3074) = 1.3045$ 

- 22. Change L to 0.05
- 23. L = 0.05.

 $\hat{c}_{L} = \min(c_{1} + L/\beta_{U}, c_{k}) = \min(.0494 + .05/.0080, 48.3074) = 6.3246$ 

25. L = 0.10, k = 5

 $\hat{c}_{L} = \min(c_{1} + L/\beta_{U}, c_{k}) = \min(.0494 + .10/.0080, 48.3074) = 12.5998$ 

We thus conclude that k = 3 is too conservative. Setting k = 4 or 5 yields nearly the same lower bound on c. In particular for k = 5

L	=	.01	ĉ,	=	1.3045
L	=	.05	ĉ,	=	6.3246
L	=	.10	ĉL	=	1.3045 6.3246 12.5998

the second secon ан и ја Риши Марија Ист ALL PLEAUS • • • • • • • • • . • • • . . 1 - - 1.5 E. 4 - 1 - 1 - 1 - 1 s×ExInski lütml.: . • • • . : • • -. · **.** .

ن بر ا بر ا

<u>ن</u>یں • • • بر بر

••

4

· · / · ·

1 - 1 - 1 1 - 1 - 1 1 - 1 - 1

; ;

••

to attack the target

• ,

• •

ہ . 1944ء

.

TALEN NI NI ALLE LUE

۲. د بر از د ا

• Θ

Lease and a second

.

-

\_

. 1ر

•

-

Ì

. Juist form .. 2 E CARANA CARANA ۰ . • :

Contract of an entraction to appreciation State and the second second a first second

, . . . . TAULANDER CONTINUED LEVELAND MADE AT SUPER 

:

# 1 · · 1 Communication of the second sec

								,	
				÷				: 4	
				÷				and the second matrix of the second	
				:		:		t	
				!			•		
							•		
				i					
				1.	-			ł	
				:				1	
							-		
				•				•	
				i					
				-					
								•	
						:		•	
	· · ; •:)	6 4 4	<b>ار ۱۰</b>		5.4.4	2.450.5			
, : ,	5	. •	•	: 5		. •	:	•	
				į ••		•••			
•	1.1.N. S					•	. N. 11.		
	1.1					•			
		٠	ί,	<b>†•</b> :	1	• •	Ļ		
				•					
•									

i

### 

SECRETE WALL WITH I			0.50 ite				
TENT (NOT THE MERICA) SECT						10 n0 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 +	
	: • : • :	7.24			i - i i i i	2.1.255.2	
· · · · · · · · · · ·		г • •	,	· <del>*</del>	, ,	11 1. 1.	•

A1: U. ଚ

:

(1) A set of the se 1..... :

.

111 . 1

### ۲. <u>۱</u>.۱ P. E. F. 1 . L.

•• • ••; . •

CLEAR & WIND REPORTED A CONTRACT ve te tra Di 

COLLENUAL OF STATES AND E VILLEN TAL & STATE FAMLY LIVELULE L'VER ALTRO INversioner einerungen eine Einer Die Generationen

# CUMPUTION CONTINUES OF A PLACE OF PORTUNES OF CONTRACT 25

t 1				•				
	•							
				•				
POLYO,	U111 - X	•			U. 445C	、よいにする	1.000	
TENTE STREET BUILD	۲ ( م : ۲	114 • •	•	1400.	1.458 N. 458			•
1:1214151		1 -		1 1 1 1 1	r	•••		

## PRILIM C KI ... CLUIT

;				•			
Ņ					•		
SUPPORTED AND SUBPLIEVAL BASEL 1	2 - 6 3 - 2		0+004.			۰. • · •	
SUCPTINGPAL HERICKI						•	
トチョー・シング	2 · · 2 · ·	7.4.25		t. • T • · ·	- 1.0.12	A Real of the	
E. APP AFT C. V	· 1	. J	20 14	1 	 : ;; ; ; ;	in an	
						. (	

L'ENLL SAME

LILLING THE MES

1 · · · · ·

SAFE LUCE THEFT

.

1. J. J. 1.

-. .

•• لأمامان PL+2 - 1.

.. 7 Vulta Junio

As a bottom of the second s . . . we have a feet

. Plank and a second of a A BIC of a Best of the .

CUMPULID LUDITULE CORPUSED FOR LUNCE CONTRACTOR -.

THIR MUNIT	20101 AND 3		
	1 GM	T GM (M) M G G G G G G G G G G G G G G G G G G	-
\$***			•
	•		
T	. 5400° · ·	1.0048.11.10.44.04.1.1.	
t	ていいい		
1 	-G 		
ð	206210	1.×205 1.×10.	

ļ

, (X			
•	シー・シー・ション しょうしょう シー・シー・シー		
	121111	5. J.	
	<b>.</b>		
0.1 (Compared to 1.1) (Compare	C+1.3 × G	またのい。	•-
	. č. č. 5.		-
	5 • · · · · ·		
	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1		1 1
		-	
<pre>/ ALAULAL CONTACT STATE (CONTACT CONTACT) CONTACT (CONTACT)</pre>			:

. . . .

. . . . .

. . . . .

1-+1:25

-

. 100 : KAN, I K V. LU. . . : : ଳ

and the terms of the second of the second of the second of VULE S MULLAN, SU LUNIJUNC

2 ; Thursdown Contraction Lovel 6300 MEE - PREParate FAMLEY LEGRIC NULLE VER UP 555

TECHE MIZH TELEPITHERAL APPLOXI SLEPE(SHAIL SAMPLT) j 4600000 -2801 10 . . . 0.1.029 () 3 () • • • 1633. 31 0.1.5 1 CUMPUTED FURTHER FOR THANK REPORTED AND 0.1904 TENELSCALT REVES 712611 0.1201 1.3000 1 · 3 · 1 215707 1241175 1.4316 21.6505 د. درزه 1.14.1 1.0044 ..... LLNIN 15 • • • TREFTMENTS USED PR\_LAPANARY AROCHUSE LINIALS SL. ्त - २ - २ ы Н 7 -E . .

(G) "TREAVEL PPACKT TO UNCERT L'EPACK) UNCULTED AND TO THE ACT OF AND SAME TOTAL 1.26.7 1.25.11 4 2112 ; • SAFE LUS HASHLINE 1.00 D - - -

. . . . . AND AND AL THE TO THE LEVEL (NOTMAL APPELATE + - ucderPr0 u Brc Picht + chaig + cuhant\_\_\_\_ O MELLER PERENTER OFFICER OF CARE TERRENT CARELES SAFELS OF CHELES SAFELS 2411-3 CONT NUM TO SLOPUTUREL REPRESE O SLOPUTSARLUSARPLUT 0. (DAT 440J\*9 9.0000 : ł 1 į 1.1145 9103.16 1.63.5 t .t. 5 . FAMILY COVERDING LUVILE - 040000 Individual augrafonat asvealantal fener 1: -447.00 CUMPUTED CONFIDENCE POUNDED ON F. LUNCLE DMPL ): NUIC: SIMULIANELUS LEAFLO AL IFILLE : 10 30.4 LE GLELE KE & UCKE SUBLE 0.397: 1 9-1-51 J. 173 515000 1,000.1 1. F. L - 1. LYS UFFCK 221222 -----21.1702 كعتمعة i A RUNDA 5000 +000+0 210111 5 20 6 . 5 TREATMONTS UTED .... ر «ر PRELIM\_MAAY P. SULISE 4 1 PARANILLY VILVEDT 4.51 3. <del>9</del> ∽ + -1 • SAFE DUS. E

1

. ..... •••••• , 0...LI 1 1. PANALEL & VALLES

of the provident of the she was favorably to 

FAMLEY CENFILLED LEV LE VOUSSION INUIVIUSE LONESCING - EXALCONSEL CORFEED) - C.V.C.

• APULATION AL AL ALL'ANDA LAVAL (NOAMAL D'A ALTE CUMPUTED Catherine Lucions and Electric Senter 14

TREATHERT - TONFIDENCE PLANDS 5141 K. K.

0.1 . .... i.1367 -----~15 200 40004 . . 

: ..... 372.5.5 1.0000 1 54962 110100

i . . . PRELIMINERY RESULTER

SLUPE (SRALE SARPER) 28:0:13 410010 COAC MURN' " (LUP. (NURMAL REFEAR) 0.6376 610010 -24.57.54 72422.0 ThEATMENTS LOUD t • ----

-0.0024 21.4560 Safe unde Kebullur Č t A

12.1750 240625 19.5620

	•	المالية والإرادة			
PAPAMILL VELUES					
• 101 119 11 • 1	* ± * 2 · · · * *	•			
NUTES SEALEFAILS		للمالية المراجع	the contraction of	יין אוריאסערעיין ארעיין ארעיין ארעיין ארעין א	
FAMALY LUNFULING LAULE		ווייניין בייינב טיאייני 1:		and the state of the state and the state and service	
CLARUTED CONFLUENCE FILMER ON FULLE SAFE	úf tunn. T	A LEMIT S		•	
			• • • •		
	LUN . A	CE F. 3	·	-	
	•	- インシン・キット	•	•	
i. d	÷	1.4623			
	- I 92 2º J		· · · · · · · · · · · · · · · · · · ·	na manananana a perintahan kerena tahun	
.7	າຈານເຊ	L.1285 .			
uı	1,1361,1		• •		•
2	1	1. 1. 1	;		-
PRELIMINARY REGULTUE	H- -7	-	•		
TREATMENTS (ISTU	i	COPC MILW 1 34	JLOPETNERMEL APPRENT	I SLEFE (SMALL SAMPLY) " "	
		+ 2 / 2 + 7	C. tu 54	0.40.0	
SAFE LUSE RESULTS			•	1	
C CALCULITE SAFL UPS INCAMET TEPRUAT	CX1 065:	CRUUN - 285 70050 TRURANE SPRUX 7 1 2000		CHC3E	•
100 + 1		2044*7	n+20+1	のまつれる。	

١.

ttern - witry is not not be

1

•• PAPAM 1." VALU

, م. م. د ا 1 11 1 :: 8

CELEVANIE CONTRACT PRESE LANCE LANCE G P ut les uni funteur hui an 1 pV- an e R ニト

APLENS. UPL COPPENSION OF NOR WAL APPSOND FAMLY CONFLUENCE LEVEL - 01204 INDAVIOUSE CONFLUENCE - UNDER SALE - UNFUEDE

SLUPTI, MALL STAPLET CORD. MURNING CLUPS (PUEMNE SPPREME AMEL 1: CLMPULLU CURFILELLE FURNES A F LUMLEL CEMPTURNET SULMET 0.159T ...... • ;;;; • ; ; ; • ; 4000 1012\*\* ete : : \*30.X. 1502.0 0000.0 ( • • • • • ز LOW.N 1) • • ., TELEVENTER USER PRELEMINARY RESULTOT THINE SITE THINK

CALULATED SAFE LOSE INDER ABORN AFAUDESE 6.3246 0+11-0 . • 1.5.40 540**3**•3 ţ CALLULATO JAFC JUSE CHOLL N. 245 JUSE (NURMEL STRREAT STORE) TREAMEL SPECK) ł 7.22 i i 24.2184 SAFE LUSE MUSULICE 1.272 : - ----`-ନ୍ତ

<pre>cd ul = million introduct introduct using a divident is lis bolate as ours prover a million using a divident introduct introduct divident bolation of a divide particular using a divident interval divident bolation introduct and in a visual interval divident bolation introduct and interval realize contractions realize contractions is and interval interval interval of the second is and interval bolation interval of the second is and interval bolation interval of the second interval of the second i</pre>							
<pre>Addition Vell : (: 6.14.2 K = 0.000 ACT ACT ACT ACT ACT ACT ACT ACT ACT ACT</pre>			- ሥትላቲ የራደ		·		
MCT FECKUS UND FE	PARAMIEL VE	. *					
WUTER LEMER FUNNUL UN FLOMALL ANH.]: TELATMENT CONFTOENCE FULKUE LEMER UNPERA LEMER UNPER UNPERA LEMER UNPERA LEMER UNPERA LEMER UNPER U	40 1.1 5 4.400. FAMLEY CONFER	LANANDS AUNTLE	atta she ave setter alamanta an	ה מנגר יין יי	USAL NULLA	1	
THEATMGRET CONFIDENCY FECADE LEWIR UPPLA 1 T. C.C.C. C.C.C. 2 2 3 T.C.C. 5.159: 4 U.C.U.S. 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	נפארטדנט בטאו	Fitterie tundu.					•
1     T.E.     UPPLA       2      T.E.       2      T.E.       3     T.E.     T.E.       5     T.E.     T.E.       5     T.E.     T.E.       6     U.G.     U.G.       7     T.E.     U.G.       6     U.G.     U.G.       7     T.E.     U.G.       8     U.G.     U.G.       9     T.E.     U.G.       10     T.E.     U.G.       11     U.G.     U.G.       12     T.E.     U.G.       13     U.G.     U.G.       14.10     U.G.     U.G.	TESATME	FT C.MFT021	VCT FOUNDER				a the state of the
<pre>7 T.E. C.T.C. 2 1.0001 5.1892 4</pre>		LCW.R	UPF.K			-	
<ul> <li>2</li> <li>3 T.GUET 5.159: "</li></ul>		1		•			
$\frac{3}{6} = \frac{1}{10000000000000000000000000000000000$		ن. د	1				·
<ul> <li>CALCUC ULES</li> <li>C.10F. 0.5645</li> <li>C.10F. 0.5645</li> <li>C.10F. 0.5645</li> <li>C.10F. C.10F. M. R. R. M. R. M. M.</li></ul>	9- <b>1</b>	3 T.CEL					The second s
5 Т.1054 5.5645 0	<b>.</b>	0 • 2	1 . I I I I I I I I I I I I I I I I I I			•	
$\frac{1}{12} \frac{1}{12} \frac$		•	0+3645	•			
$\begin{aligned} \mathbf{I} = \left\{ $	- -	0 • ~					
<pre>III.I.A.A.D. 02.0</pre>	RILLWINSEY F						
t ube albures Catures Catures Inconstration Inconstration Inconstration	711 C 1867	0. E		Churs (Surges - mrs4.	(X) (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1	יינייט, אינייע, אינייע אינייע, אינייע,	
e euse algures Calcorde Frank were frank 21 frank (Laurter and ) Tarent Frank were frank frank frank) (Laurter and ) Tarent Frank			6 4 6 1 C 4	1.1042			
CALCULATE AND	AFE EUSE RE	ļ i	•	:		•	
	CALLUCIT	IS				Contraction and a second	
		1.100	1				

ţ

We now compare the lower confidence bounds on  $c_{\rm L}$  obtained from the nonparametric dose response curve fits to these obtained by more classical probit fits in Section XIV.

Nonparametric	(k = 5, UCR = 50.0) Small Sample	Asymptotic
L = 0.01	$c_{L} = 1.3045$	1.496
L = 0.05	$c_{\rm L}^{-} = 6.3246$	7.282
L = 0.10	$c_{\rm L}^2 = 12.5998$	14.515

We fitted a probit model to the data (untransformed concentration) using SAS PROC PROBIT. We corrected for background with Abbott's correction and obtained

Point estimate	Lower 95% Conf.Bnd.
3.7725	-157.377
22.6065	- 84.3307
32.6468	- 45.9667
	22.6065

The lower bounds based on the probit fit with background are thus useless. We refitted the model with the assumption of no background response. This is therefore a more restrictive model. Results:

	Point estimate	Lower 95% Conf.Bnd.
L = 0.01	1.6567	-17.6457
L = 0.05	21.1299	8.8997
L = 0.10	31.5110	21.2620

Thus at L = .05 and especially at L = 0.10 the nonparametric bounds are more conservative. However they are based on many fewer assumptions.

An attempt was made to fit the probit model to log concentration, as suggested by Finney. The probit program would not converge at all.

We now apply the nonparametric dose response program to the <u>Holcombe and Phipps compound D</u> fry mortality data and compare the estimates of safe concentration with those based on probit and on logit fits. The logic underlying the procedure is indicated in Figure XVI.1

Refer to the computer printout (nonparametric). In this example the control group is at concentration 0 so we do not have to adjust for its affects. However we do have significant background effect.

We see from the listing of the data ( (1) on computer printout following) and mortality rates by group that group 5 has an observed fry mortality rate of 0.79 while group 4 has an observed fry mortality rate

of 0.13. Confidence intervals on these values of p clearly confirm that group 4 is below the median of the dose response curve while group 5 is above the median. Therefore UCR lies somewhere between group 4 and group 5. We have taken it at concentration  $100\mu g/liter$ , about midway between the two groups. Thus  $c_r = c_4 = 72.9499$ .

If we define the "safe" values of L to be below 0.10 (over and above control) we should try  $c_k = c_3 = 44.9049$  or  $c_k = c_4 = 72.9499$  as upper bounds for  $c_L$ . We consider the results of the small sample calculations.

First trying k = 3:

L = 0.01

$$\hat{c}_{L} = \min(L/\hat{\beta}_{U}, c_{k}) = \min(0.01/0.0027, 44.9049) = 3.663$$

L = 0.05

$$\hat{c}_{L} = \min(1/\beta_{U}, c_{k}) = \min(0.05/0.0027, 44.9049) = 18.3315$$

L = 0.10

$$\hat{c}_{L} = \min(L/\hat{\beta}_{U}, c_{k}) = \min(0.10/.0027, 44.9049) = 36.6629$$

If we next try k = 4, we have less of an adjustment for simultaneity and so we get slightly shorter intervals in this case. Namely

k = 4:	
L = 0.01	$\hat{c}_{1} = 3.8453$
L = 0.05	$\hat{c}_{\tau}^{L} = 19.2264$
L = 0.10	$\hat{c}_{L} = 3.8453$ $\hat{c}_{L}^{L} = 19.2264$ $\hat{c}_{L}^{L} = 38.4529$

Thus k = 3 and k = 4 yield essentially the same results for all practical purposes.

Let's now compare these results with those obtained by fitting probit and logit models to the data.

Probit models were fitted to the data vs CONC (untransformed) and  $\log_{10}$  (CONC). Both probit fits have nonsignificant residual chi square of about the same magnitude and so are judged to fit the data about equally well. The following 95 percent lower confidence bounds on safe concentration were obtained from these fits.

### PROBIT FITS

L	UNTNSFMD CONC	LOG10 CONC
0.01	22.403	45.764
0.05	46.147	57.470
0.10	58.715	64.836

We see that the nonparametric fit is much more conservative than the probit fits since it is based on low dose linearity. Note that especially at the low percentiles 'ne probit inference is quite sensitive to whether untransformed concentration or log concentration are used.

We also fitted a logistic dose response to the data using the SAS nonlinear regression program, PROC NLIN. Only untransformed concentrations were used with the logistic fit. Background response was adjusted for the Abbott's correction, just as we did with the probit fit. The following summarizes the results of the logistic fit and compare the inferences with those based on the probit fits.

Chi square for goodness of fit 0.282 with 3d.f.

Thus there is no evidence of lack of fit of the model to the data and the logit and probit models fit about equally well.

### Point estimates of percentiles (after adjusting for background)

L	Probit(Untransformed CONC)	Probit(log CONC)	Logit(untransformed CONC)
0.01	69.946	60.974	45.31
0.05		71.826	68.757
0.10		78.380	79.370

Thus, except for L = 0.01 where there is a bit of disparity among the models (although not of practical concern) the three models yield essentially the same percentile estimates.

We now consider lower confidence bounds on these same percentiles based on the three parametric fits and on the nonparametric fit.

				HIG - ANALY				
_	TRT KE	P NE	MR NHATC	H NNCRMH	NFKY	NALIVE	NNUR ME	TOXIERNT COM
D	1	5.6	0. 33.	23.	25.	23.	23.	
	1 3	3	0. 38. U. 30.	36.	25	74.	24 • 24 •	0.C 0.0
	TREATMEN	T TOIÀ	LS: NDEAD NDEAU		PDEA	DIMB# UFRY=	0.350	
		5	0 31 0 0 36 0 0 34 0 0 29	31. 25. 33. 28.	25.	22. 24. 23. 23.	22 • 24 • 23 • 23 •	27 • 21 49 27 • 21 49 27 • 21 49 27 • 21 49 27 • 21 49
1	TREATMEN				PDEA	DEMU= DFRY=	0.350 0.080	
	3 1	. <b>5</b>	0. 35.	35.	25.	23.	23.	44.90.49
	3 3		0. <u>39</u> 0. 26 0. 35	38. 26. 36.	25 •	20.	20 • 24 •	44,5049 44,9049 44,9049
	IRLATMEN	<u>1_1014</u>	LS: NDEAD NDEAD	MB= 63. FRY= 8.	PDEA	DEMAE DFRY=	0.315	
					an Link ku ku		* · ·	
	4 1 4 2 4 3	5555	0. 30. 20. 20. 0. 27. 0. 36.	30. 20, 25.	25. 25. 25.	21. 20. 21. 25.	21 • 20 • 21 • 25 •	72.9409 72.0400 72.0409 72.9499
1	TREATMEN	T TOTA	LS: ND EAD	RV= 13:	POEA	DEMA= DFKY=	8.340	
-	5 1	5	0. <u>3</u> 2. 0. <u>31</u> .	31.	25.	2.	2	128.1500 128.1500
	3 3	55	0. 36. 0. 34.	35.	25.	9 <b>6</b> •	9. 5.	128.1500 128.1500 128.1500 128.1500 128.1500
	TREATMEN	L.IUTA	LS: ND EA D NU EA D	EMH= 67. FRY= 79.	PDEA PDEA	DEMB#	0.325	
	6 1 6 2 6 3	5	0. 36. 0. 34. U. 25.	C. U. C.	25.	C. C.		
	a. <u>4</u> .		2 24.			. 0.	0.	and a state of a UNI

· .•

~.

HOLPHIA - ANALYCIS OF FRY DETA

PARAMETER VALUES: L: 0.0100 K: 3 UCK: 100.1400 Nute: Similian dus cumfinence tevel is callulated using benferreni inteud

NUTET SIMULIANUUS CUNFIDERCELEVEL IS CALCULATED USING ERRERINT INFUDELIY CONFIDÊNCE LEVEL: 0.9833

0.023 0.0031 0.0027 0\*(022 0.0021 0.0022 CUMPUTED CONFIDENCE BOUNDS ON P (SMALL SAMPLE): 0.2193 0-8704 0.9599 1.0000 CUNFIDENCE, BUUNDS 0.0203 0.1321 0.1591 0.1521 UPPER 22.4525 34.2849 36.4750 0.6391 0.0326 LOWER 0.0326 0.0672 PRELIMINARY RESULTS: TREATMENTS USED ... , IREATMENT Ś 134 13 4

CHUSEN SAFL DOGE CALCULATED SAFE DOSE CHUSEN SAFE DOSF INURMAL APPROXI ISMALL SAMPLEI DOSE ISMALL SAMPLIL

3.6663

4.8295

CALCULATED SAFE DCSE INURMAL APPRUX

4.8295

SAFE DOSE RESULTS:

3.6663

HOLPHIS - AVALY IS OF FRY ULTA

1

-CHOSEN SAFE DUSE 3155.315 NDIE: SIMULTANEEUS LUNFIGENCE-LEVEL IS LALCULATED USING BUNFLARUND IN-QUALETY 0.0031 0.0027 3200.0 CALCULATEU SAFE LOSE 18.3315 1 0.(022 0.0021 0.0522 CUMPUTED CUNFIDENCE BOUNDS UN P (EMALL SAMPLE): CHUSEN SAFE DUST 24.1473 3 UCR: 100.0000 CONFIDENCE BOUNDS 0-8704 1.0090 C.2193 1251-0 0.1581 0.1551 UPPER 22.4525 36-4750 39.2849 0.0203 0.0326 1.29 4.0 0.4549 0.0672 0-0326 LOWER CUNFIDENCE LEVEL: 0.9833 CALCULAT D JAFT NO SC INURMAL APPRUXI PRELIMINARY RESULTS: THEATMENTS. USED. 0.04CC K: SAFE DUSE RESULTSE PARAMETER VALUES: 24.1473 ...IKEAIMENT ÷ ۲**:** 4 134 .«, 4 :

ł

. . . .

CHOSEN SAFS 1975 --Lemall Sampley. SLOPEINUMMAL 'APPROXI ... SLOPEISMALL SAMPLEI. 36.6623 NOTE: SIMULIANEOUS CUMETURACE LEVEL IS CALULATED USING BUILS BUILS .0.0027 0.0031 0.0028 LALL SAMPLEL UUSE 36.6629 0.0021 0.0022 U.L022. HUGPHIG - ANALYLIL OF HEY SIIA COMPUTED CONFIDENCE BOUNDS ON P (SMALL SAMPLE): . ...... CHOSEN SAF, DOSE INCRMAL APPROXI 44.9049 2 UC9: 100.010 0.2193 CUNEIDENCE BUUNUS 0.1321. 1.0000 0.1581 0-1581-0.8704. UPPER CONC MEAN 2. 22.4525 36.4750 39.2849 LOWER D. D.203. 0.0326 0-0326 0.0672 0.6891 0°9599 0.9833 CALCULETED SAFE DOSE - (NURMAL APPROX) . PRELIMINARY RESULTSE TREATMENTS USED 0.100J K: SAFE UUSE RESULTS: CONFIDENCE LEVEL: PARAMETER VALUESE 1462.84 **TREATMENI** ÷ 134 ••• 4 .. L

CHOSEN SAFE DUSE 3.9453 ...SLOPEINCRMAL. APPROX) ....SLOPEISMALL ..SAMPLE) NOTE: SIMULIANEUUS CUMFILIENCE LEVLE IS CALCULATED USING EPHIFERRONI INCOURLEY 0.0026 CALCULATED SAFE DOSE 3.8453 0.0021 -----HOLPHIK - EVELYCIS OF ENY UATA COMPUTED CONFIDENCE FOUNDS ON P (SMALL SAMPLE): . . . . . CHUSEN SAFEDUSE 4.8410 4 UCR: 100.0000 0.1260 1.0000 0.1516 0.2120 0.8651. CUNFIDENCE BUUNDS CONC. MEAN 0.1516 UPPER 36.4750 0.0711 0.0223 0.49638 0.0352 0.0352 1794.0 LOWER CONFIDENCE LEVEL: 0.9750 CALCULATE O SAFE DO SE INGRAAL APPROXI PRELIMINARY RESULTSE ... TREATMENTS. USED. L: 0.0100 K: SAFE DUSE RESULTS: 4.8410 PARAMETER VALUES: IREAIMENT. ٥ 4

HCLPHIG - AMELYSI' LF FRY DAIA

PARAMETER\_VALUES:

SLOPE(NCKMAL APPROX) SLUPE(SMALL SAMPLE) NOTE 1... SIMULTANERUS LUNFLUENCELLEVEL IS. CALCULATED - USING - EDAFTRRUNT-INE GUALITY 0.0024 CHOSEM SAFE DOSE CALCULATEN SAMPLES DOSE 0.0021 1 COMPUTED CONFIDENCE ROUNDS ON P (SMALL SAMPLE): -----A UCC: 100 4 -CENETDENCH-BUUNDS 0.1260 0.1516 0.2120 1.0000 CONC. MEAN 0.151.0 0-8651 11PP - R 36.4750 -----0.0223 0.0352 0.6571 9638.0 0.1352 1110.3 LOWER 0.9750 .TREATMENTS.USED PRELIMINARY RESULTS: 0.0500 K: SAFE DOSE RESULTS: CONFIDINCE LEVEL: IREAIMENT 1 4 -

CHOSEN SAFE USSE

19.2264

19.2264

24.2049

CALCULATED SAFE DOSE

24.2049

	IPPS COMPOUND D	FRY MORTALITY DATA	9
PARAMITER VALUES:	• • ··· • ·· ··· ··· ··		
L: 0.1000 K: 4 U	CR: 106.0000		
-NUTE: SIMULIAN.OUS-LUNFLD	ENCE LEVIL 15	alwied - No Ing, Builflarian	IL INCOUNCEF#
CONFIDENCE LEVEL: 0.9750		,	
COMPUTED CONFIDENCE HOUNDS	ON P ISMALL SAMPLE	<b>():</b>	
TREATMENT CONFIDE	NCE BOUNDS		
LOWER	UPPER		· ·
	0.1260	·	
2 010352	0.1516		
	Č.1516		
	0.2120	· · · · · · · · · · · · · · · · · · ·	
	0.8651	· · · · ·	•
1	1.0000		
		, •	·
PRELIMINARY RESULTS:			
TREATMENTS USED		NORMAL APPROXI. SLOP	ECSMALL SAMPLES
1 4	36.4750	0.0021	0.0026
• •	0004120		0.0020
SAFE DUSE RESULTS:			
CALCULATED SAFE DOSE	CHUSEN SAFE DOSE	CALCULATED SAFE DOS	E CHOSEN SAFE DOSE
			· · · · · · · · · · · · · · · · · · ·
48.4098	48.4098	38+4529	38.4529
······	·····		

.

Output from special purpose program to calculate lower bounds on safe concentration based on nonparametric dose response curve fit

### Lower confidence bounds\*on percentiles (after adjusting for background)

L	Probit(Untransformed CONC)	Probit (log CONC)	Logit(Untransformed CONC)	Non- parametric
0.01	22.403	45.764	2,515	3.845
0.05	46.147	57.470	38.390	19.226
0.10	58.715	64.836	54.526	38.453

\*Confidence intervals for parametric fits are based on 95% two sided nonsimultaneous procedures using Fieller's theorem. Nonparametric confidence intervals are based on simultaneous procedure with familywise confidence level of 95%.

Comparing the lower confidence bounds on safe concentration we see the following relations.

- 1. The nonparametric procedure is more conservative than the parametric procedures, especially at the very low percentiles. This is because the nonparametric procedure is built on the assumption of linear approximations in the low concentration region, while the logit and probit fits approach their limiting values through exponential decay. Which estimate is more appropriate would need to be a matter of biological judgement. Since responses in the region of concentration for which  $L \leq 0.05$  are dominated by background response, the data themselves do not provide much empirical evidence.
- 2. The results at L = 0.01 are surprisingly inconsistent across model fits. There is at least an order of magnitude difference in confidence bounds based on the logit and probit fits, despite the fact that the point estimates are in good agreement. There is even a factor of two difference between the bounds based on the two probit fits, despite good agreement of the point estimates. Thus the inferences at low percentiles are very sensitive to the model assumed. The nonparametric procedure, while conservative, is based on many fewer assumptions.

In this example the probit model fitted the data quite nicely and yielded more liberal confidence bounds on c<sub>1</sub> than did the nonparametric procedure. It is our experience that this is not always the case. In some situations the probit or logit model does not fit the data well. In other situations Fieller's method may yield lower confidence bounds on c<sub>1</sub> which are negative! We saw this in the case of the DeFoe data. In such cases the nonparametric procedure can provide more liberal bounds than the parametric procedures. We will see this in the following example.

We now place lower confidence bounds on the safe concentration for the Benoit compound A fry mortality data. We first consider the nonparametric procedure and then compare results with bounds based on the probit model. Figures XVI.2, XVI.3 contain plots of proportions of dead fry vs untransformed concentration and vs  $\log_{10}(\text{concentration})$  respectively.

Based on the appearance of these plots, a probit dose response model does not seem to hold very well, especially with respect to untransformed concentration. Furthermore, there is some question about homogeneity of response rates within tanks at the highest concentration.

We first consider the nonparametric procedure. We see that the average concentration at the control group is 0.0809 and there was no observed fry mortality there. Based on the plots of proportion dead fry vs concentration and based on fry mortality proportions printed out by the program, we set UCR, the upper bound on the upward concavity region, to be somewhere above the 5th treatment group. In particular we set UCR = 15.0. Then  $c_r \equiv c_5 = 13.3182$ .

Note that for the purpose of illustration we are assuming that there is no tank to tank heterogeneity within groups. This assumption needs to be checked and appropriate modifications made, if necessary.

Because of the many sample proportions close to 0 the asymptotic normality assumption is questionable and so we use the small sample confidence bounds.

Based on the observed proportions of dead fry in the various groups, if L is less than or equal to 0.10 it makes sense to choose  $c_k$ , the upper bound on safe concentrations to be  $c_4$  or  $c_5$ . For definiteness we choose  $c_4$  here. Thus we have

UCR = 15.0 k = 4

L = 0.01  $\hat{c}_L = \min(c_1 + L/\hat{\beta}_U, c_k) = \min(0.0809 + 0.01/0.0423, 6.6020)$ = min(0.0809 + 0.2364, 6.6020) = 0.3137

L = 0.05  $\hat{c}_L = \min(c_1 + L/\hat{\beta}_U, c_k) = \min(0.0809 + 1.1819, 6.6020)$ = 1.2628

L = 0.10 
$$\hat{c}_{L} = \min(c_{1} + L/\hat{\beta}_{U}, c_{k}) = \min(0.0809 + 2.3638, 6.6020)$$
  
= 2.4447

We now compare these results with those based on the probit fit, using SAS PROC PROBIT.

First models including background variation were assumed. Namely

$$p(\text{CONC}) = c + \overline{c}\Phi\left((\beta_0 - 5) + \beta_1 \text{ CONC}\right)$$
$$p(\text{CONC}) = c + \overline{c}\Phi\left((\beta_0 - 5) + \beta_1 \log_{10}(\text{CONC})\right).$$

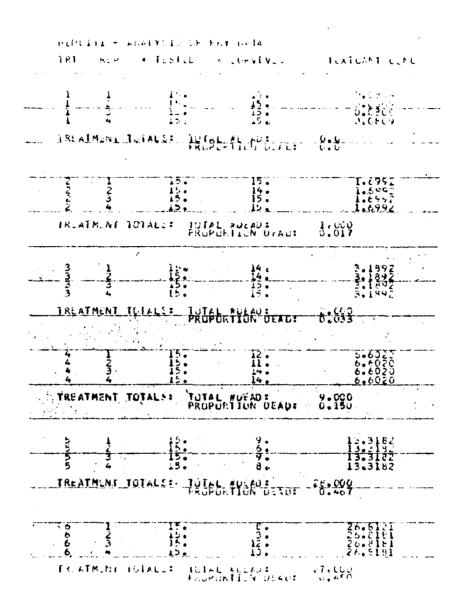
The maximum likelihood fitting algorithm was unable to converge with either of these three parameter models!

Next the background rate was specified to be 0 and two parameter probit models were fitted to the data. The plot of proportion dead fry vs CONC suggests that the probit model does not fit the untransformed concentration and in fact the model converged to shows substantial lack of fit to the data. We therefore consider the probit fit in log concentration. This fit is better, but still exhibits marginal statistical evidence of lack of fit (Residual chi square = 8.00 with 4d.f., which is significant at  $\alpha$  = 0.09). Based on this fit, the 95% lower confidence bounds on response distribution percentiles (using Fieller's theorem and adjusted by Finney's heterogeneity factor) are:

L	0.01	$\hat{c}_{L} = 0.082$
	0.05	$\hat{c}_{L} = 0.523$
	0.10	$\hat{c}_{L} = 1.362$

Thus in this example these bounds are lower than those based on the nonparametric procedure. They are also based on a much more restrictive model.

The nonparametric procedure seems quite superior in this example.



WENTER ATALYSE, UP PEY UATA

i 3 UCH: 15.1610 PAKAMETER VILU. :

L: 0.0100 K:

1.1.1.2. \_\_\_NUTES\_\_\_SIMULTERLEUS\_SUMPLEPTING\_IND\_ATARANAL\_BAT\_UNUTATAU\_USING NUTERAURI, INSOUALITY

APOLVAUUAL CONFIDENCE DEVELINGRMAL APPREX : 6.9929 FAMILY CUNFIDENCE LIVILE UNSOON INDIVIDUAL CUNFIDENCE LIVELESMALL SAMPLEDE FROM 75 i

CUMPUTED CUNFIURNCE SHUNDS UN F (EMALL SAMPLE):

TREATMENT	C. NF ID ENCE BUCKDS	E BUUNS	n		A I I I I I I I I I I I I I I I I I I I
	LUWER	UPPER	<b>-</b>		
	<b>c.</b> r	25152	and the second second and the second s	a ser anna a marta a marta anna an an anna anna anna anna anna	
۰.	J. 5 6 6 2	C.1017			
	820011	U-1286			
<b>*</b>	<sup>4</sup> 0.0631	0.282B			
1 1 2 2 2	52250	0.6175	and a submersion of the submer		
÷	U . 3049	410°°U+			
and a support of the second			and a first of the second state		
PRELIMINARY HESULIS:	15:	ſ	•	<b>1</b>	
THERTHINTS USED		CURCHINN LUDY	LUP TRUKKAL APPRUXF SI	SLOPETSMALL SAMPLET	and the second se
5 1		1.451	76 23 • 1	Ú. 0414	
<b>b</b> I		3+3415	1.0403 ·	<u>0*v</u> 434	rannen a er hann
<b>u</b> ) 		6.5445	C+C+72	0.0466	,
		12:20	<u> </u>	C+40	5 
135		5.5.4	0.00.0	109016	
		h.ot7C	9.413	C. C.453	and a second of the second
4 T I	• ,	5.7970	C + 0+ 3 T	0+0503	
SALE VUSE RESULTS					· · · · ·
CALULATE DURING APPRCAT		CHUSEN SAF BUSS	I CAL UIL 21 . U SAF UUSE	UNE CHUSEN STEE UUSC	

0.3225 

0.3225

F. . 42.59

8274-0

INLIVIDUAL CURFIDINGE LUVIL (MSPMAL AFFROX): CUMPEN ŀ ł ł 1 ULEUUSEL ZPIECX) ULE - LEULE ALE APPECY - LUL GUL CLEU ENEL VUSE - UNSE - LEUSEN SAFE UUSE - INUERAL SAMPLE) ULUPE (NUMMAL APPREX) SLOPE (SMALL SAMPLE) 1.2891 0440.0 C. 0454 0.0501 C.1503 C. (414 0.0466 0.1468 ..... 1.20%1 : 0.0290 0.0463 6.0472 10+0-0 0.0500 E (413. FAMLEY COMPLIANCE LIVEL ONNOC ANDLY OMELLOWER ON LOWER SUMPLIE ONNOT SQUELL - ALLER DE FREESE CUMPUTED CONFLUTECE FUUNDE DE L'UMALE (FMEL.): TREATMENT TOTHETOTHOUNDE TOTH 1. R. 5 0.10**1**44 0821.0 J.6175 0.011.4 7.11.17 0.2825 0.0645 ULA MULH UPER 1400.1 6.667L 10.211 0121.4 3-3415 6.6455 \*\*\*\* c ULA: 120751 243044 0.0325 v.ů631 1.000 LCWER T-EATY NT . W. LU PRILITINEY RESULTSE ; SAFE TESE RESULTS 1.8050 PAFARELE. VALUES 1 т в о, П 14 ÷, \* .. ł -----ī

4246.0 INDIVICIAL COMPLOENCE LEVEL INCRMAL APPREXIS CHOUS MALE TOPPOST - CALGULE (SUMPLE) - DUSA - CHOSEN SAME UP 35-<u>ፈና ታሁህ ቀላቸት ወጣ የሆኑ ያን በታይ የድድረት በ ግድ ፕሮሮሀፅዬ ቤሮን ...</u> 215502 SCOPTINDEMAL TEPPROXI" - SUCPEISMALL SAMPLEY 0.6440 C. C414 C. C434 0.0466 0.0503 0,0501 0.04.68 2104.2 5.5408 U-- L-2 9 U 0.0405 U.C.472 0.65.10 0.6491 C- 04 73 PLAUTT - 287 LYLL LF FIX LIRA · · · · 75 CUPPENED CONFINETIES OUTINS ON F LUMMER URMEL ): 3 FAMIER CERFIGERCE E-VOLE - 0.4500 INEARDORE SUBJECTER (SMALE UNAVEC): 10 H T T CUNFIGENCE RUUNS 0.2828 3.0175 1010.1 0.1735 110100 0.6610 MX\_M\_DNDD... - 0191° C UFPER 3.3415 1650.1 1052.5 4456.6 3669.9 1133.3 1 ILR: 7:3252 620010 U+1631+U 0.3049 UNLANZE PPRUXI DULL : • • • • • 1.000 LUW: K TREATH RTU COUDT PRILITINARY RELUCESS ..1∷(... K. SAFL UISE K SULLS +05.3•1 PALANCI - VALU. : TREATMENT in o 4 6 7 ы .•, --J r ŝ ... ł

PURCHER - UNITED OF THE UNITED

þ

N

PANEN I NALU ...

I

and a submit of we down a spire of share a state of the 1. 1. r v . STAUL FARELUS CONFICE 201.3

-----1901410041 CUNPLUENCE LÉVELINERMAL APPRIXIE i FAMSLY CLARTCHARCE LURCE - 0.5500 Intevender Libradi Neur Lovuel Librel - 0.00000 - 0.53

CUMPUTED LUNFLERGE EUDINDS ON F (LARLE GAMPLE):

i SLOPETHICKAL APPREXT SLOPETSMALL SAMPLET 0.0423 1543.0 0. U463 0.93.60-0 2.74.55 1 - M. M. P. J. CAFTJERCE FJUNDS 5 • 1 ÷ 0 ¢ 2512.0 0.6105 C.1.32 L.5444 \*:340 CUNC REAR 6.0010 ۇيېڭىنى. 616465 ł .0662 r.5168. 0.5115 (1) 262.2\* LÜWcR TREPIMENTS USED -PRILIMINARY RESULLS THEMEMERT + + 4 ....

-CHUSENSEFAPLESC LAL NUL SAMPLE UNSE 0.31/3 - INUT NAL TEPROXT 0.3438 --- CALLULATS DEPENDENCE SAFL LULE RUSULISE 0.3435

INULV.UUAL (UNPLUENCE LEVELENDRMAL AFPRUX): 0.9833 A FULL AND LUNCTURE THE FASTER AND LACTOR AND A TARGET AND THE AND THE ADDRESS FAMELY COMPLONATON INVILLE OFYSION INDEVENDAL COMPLONATON VENERCYCONSTRUCTUR SAMELUTE OFEDD CLAREBIU CUMPLACA L POUNDU UN P (IMPLE AMPL.): 6.Jeef CONFIGENCE SUCRES .104.5.1 0.1232 Y : 440 - '' 1.0003 1 202.7 LUWUN .n. PARAM. I. Y VALL SE TELATAENT 2.0 . . . . . 2010

alco Art ar o

....

47 - [[], 44 ·

ULING MEAN SUUPE (NURMAL APPRUAT SUUPE LUMPLE SAMPLET 2424 1940.0 0.1423 0.0463 ¥ • ذ > د ≚ 0960.0 0.0456 10.0457 1225.4 C. 6 135 22154 \*\*5 ?\* 1 0.667U 3+2+2+5 5069. 7 ....... 133 C. J 2115.0 1.0062 PRILIPINALY REUCEDS IF-ALT-NTS WED SAFE LUDE REDULT 10360

..... 6.5523-0 FRUTVIDUAL CENTIDENCE LEVEL(NOSMAL APPROA): Leaded AL D. VOINE BLACK J'READAL J'REQUALLIN.
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A

A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A 2+++2 SLUPETEMALL SAMPLE 0-0463 0.6423 1340.0 SLUPEINUM SAL APPROXI 2 + 4 4 4 7 0+0380 6,04 57 0.0456 FAMLEY CONFICENCE ELVIL: 24500 INUIVICUAL CONFICE ELVILLMAEL SAMEEL): 0.49833 LUNDER - ANALT ... . FRY WERA CUMPULU CUMPLORME POURD UN P (SMALL PRAFE): - Chick Hall Apples 2 .7045 U..... U.1237 CUNC MEAN 0.1966 CUNFIDENCE BUDNUS 0.2754 4464°N UFFER .6142.5 019910 6.6995 3 7.32.68 0.0033 C+0003 0.3115 5300.0 CALLUASTE FFEEDY CLASS LUWSK TREATMENTS USED PRELIMENARY RESULTS: SAFE ULSE RESULTS: PARAMITCH VALUES 6.1043 TREATHERT . . . I. э m 3 843 ۲. ۲ :: Nuls:

## A NOTEL - ANALYSIS OF FAY WER

. INCLAILUAL CONFIDENCE LEVILINUN MAL AFPRUXI: 0.0510 CALCULATE VERE DUST CHUSAN SAFE DUSE CALGULATED SAFE DUSE CHOSEN SAFE LOSE NUTER STRUE FARTUS CONFLUTZON ALL LYARS ARE CARDERED BURSTAN AND FARMAN LYANAARTIK. CORC WERN SECRETICKARL APPLIED SLUPEISMALL SAMPLED 0.3015 0-0453 J ł :` 0.2015 0.04.00 FAMILY CONTINUES BOYLE, 0.5500 Instruction Fight of Version El Samples 1: (49350 CUNPUTE CONFIGUE GURDE OF PORTE SAMPLEST J.3121 E ULK: ISTONO 0.550 CURFICENCE BUCKE 34.1.1 L.2657 G.1 153 2-35-0 LEPLK C. C E 94 0.69+5 .. i 1:00:0 Ũ.3367 0110.1 - 211 1.4064 0.3214 ې، سال نا THEFTHINT OF USED PRELIMINIKY AS JULISE SAFE LUSE RESULTS: · · · · · 6.2421 PARAMER VILLE TFELTMENT" ي. م Ξ.

## COLTLE - PEALYLE, EP FUX DAMA

1

Parkana bar yi bu aat

· · · · · · • • • • • • • :: :

Nulst - 3 Phelmanult for elane and a Meli an Chantur de Calive Ruthard Meli (182304011)

INCLUSIONAL CUMPADENCE LIVEL PURMAL MARCENERS: LASSIC 

GUMPUTED COTTOM CONTRACT CONTRACT STARTED

ţ CLUE: INERAL APPLET SLUPE (SMALL SAMPLE) U•0453 じょしゅい 1 CONFIDERLE EQUAD. --N.C.M. 2102 C.1155 すかりょう 1 30 3.0 3936.0 0,52234 0++ . E 22223 ن مَح ق. تَ 1. • ( · · h 512540 1+0210 L. MCH . TRE 21 4: 51 5 0.20 PRELIMANTY ALCOLAT Tre21M.NF J ية 1

1

SAFE BUSE - HELDS

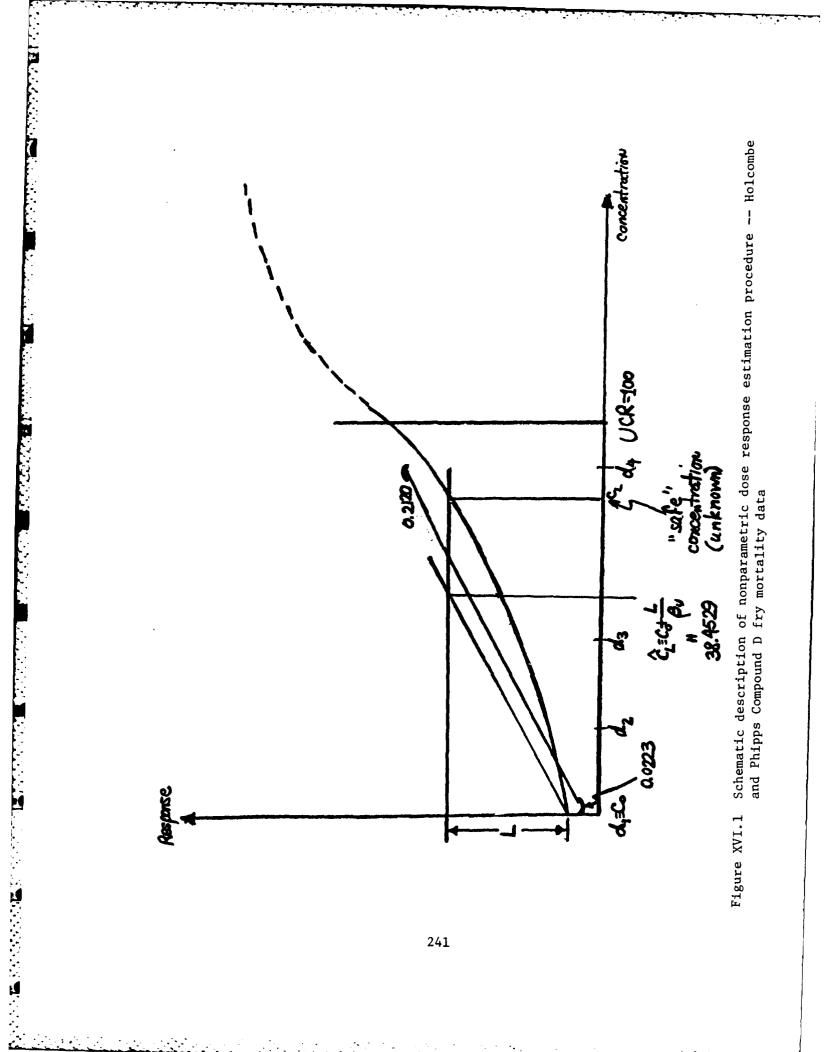
1.1940 1.1546 1.276 4 ...... Le stit

PRAVALLA VILLE I: "DU X: - UX: "5.0000 NUL: SIMULA UD: UNITE VIENAL AN LALUAL MALAN VALIT NUL: SIMULA UD: UNITE VIENAL AN LALUAL UNIT UNTURE LEVEL (NOMEL AFFRONT: UNIT FAMILY CONFLUENCE UDING UT 1.1 UNALL SAMEL D. ULTUS LINFULNE LEVEL (NOMEL AFFRONT: UNIT COMPOLU CLAFTURE BUDNUS UT 1.1 UMALL SAMEL D. UNITS BUDY UNE LEVEL (NOMEL AFFRONT: UNITS COMPOLU CLAFTURE BUDNUS UT 1.1 UMALL SAMEL D. UNITS BUDY UNE LEVEL (NOMEL AFFRONT: UNITS COMPOLU CLAFTURE BUDNUS UT 1.1 UMALL SAMEL D. UNITS COMPOLU CLAFTURE BUDNUS UT 1.1 UMALL SAMEL D. UNITS COMPULU CLAFTURE BUDNUS UT 1.1 UMALL SAMEL D. UNITS COMPULU CLAFTURE BUDNUS UT 1.1 UMALL SAMEL D. UNITS COMPULUE CONFERENCE COMPULATION D. UNITS COMPULATION D. UNITS PRALEMENT RESULTS PRALEMENT RESULTS PRALEMENT D. D. UNITS SUBFLICATION D. UNITS	<ul> <li>HALL SAMELA UNING SURLET AL ALTUALIT</li> <li>SIMILIS UNES INVITUUAL CLAFIULINE LEVELTAMAL APPROXIS</li> <li>MALL SAMELIS INLIVUUAL CLAFIULINE LEVELTAMAL APPROXIS</li> <li>MALL SAMELIS INLIVUUAL CLAFIULINE LEVELTAMAL APPROXIS</li> <li>MALL SAMELIS UNLYLUUAL CLAFILIS SAMELES</li> <li>ULA SUMELIS UNLYLUUAL LEVELTAMELI SAMELLIS</li> <li>ULA SUMELI SERVILIS</li> <li>ULA SUMELIS UNLYLUUAS UNSEL UNSEL UNLYLUIA</li> </ul>		beidil - ANALYSis Lr	मा मार्ज्य के मिल्		
SAMIL IN LAILAI U USING BUITHALAN AN JUALIY SAMIL I COPES INLAVIULAL CLAFIULNCE LEVEL (NUMAL AFFACA): MALL EA PLI I: 60 155 155 155 155 155 155 155 155 155 15	ARMEL AN LATURAL & USING BUTTANA AN JUNJULIT SAMELT COVES INVIRTUAL CONTINUE LEVELTADIA ARL SAPART: AR ARL SAPART: 55 55 55 55 55 55 55 55 55 55 55 55 55	ריי <b>:</b> יפר א:	: (') ;	:		
MALL SAPALIJ: 161 153 153 153 153 153 153 153 153 153 15	MALL EATHLIJ: 100 153 153 153 153 153 153 153 153 154 154 154 154 154 154 154 154 154 154	JANA USAL	UNERDANGE ARENALD als Orgeon Ebeltomate Same c	PANE (ANIMAL 2) US	INC. Suttates NA. AN AUALAIY	AFFKUX):
LUNFILIALE CUMPLA LUNER UPPER C.CUMER UPPER C.CUMER C.0555 C.0710 C.0557 C.0557 C.0557 C.0557 C.0557 C.0557 C.0557 C.0529 FS: C.0505 C.0529 FS: C.0505 C.0529 FS: C.0529 C.0529 FS: C.0529 FS: C.0529 FS: C.0525 FS: C.0555 FS: FS: FS: FS: FS: FS: FS: FS: FS: FS:	UNFILIALE CUMOL LONER UPPER C.O. D.C. D.C. D.C. C.C. C.C. L.C.C.C. C.C. L.C.C.C. C.C.C. L.C.C.C. C.C.C.C.C.C. C.C.C.C.C.C. C.C.C.C.C.C. C.C.C.C.C.C. C.C.C.C.C.C. C.C.C.C.C.C. C.C.C.C.C.C. C.C.C.C.C.C. C.C.C.C.C.C. C.C.C.C.C.C. C.C.C.C.C.C. C.C.C.C.C.C. C.C.C.C.C.C. C.C.C.C.C.C.C. C.C.C.C.C.C.C. C.C.C.C.C.C.C. C.C.C.C.C.C.C. C.C.C.C.C.C.C. C.C.C.C.C.C.C. C.C.C.C.C.C.C. C.C.C.C.C.C.C. C.C.C.C.C.C.C. C.C.C.C.C.C.C.C. C.C.C.C.C.C.C.C. C.C.C.C.C.C.C.C. C.C.C.C.C.C.C.C.C. C.C.C.C.C.C.C.C. C.C.C.C.C.C.C.C.C. C.C.C.C.C.C.C.C. C.C.C.C.C.C.C.C.C.C.C. C.C.C.C.C.C.C.C.C.C.C.C.C. C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.	t diversitie	NUMUS LI I SMALL S			
C.0.0.0.0.000 C.0.0.0.0.0.000 C.0.010 C.0010 C.0010 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050 C.0050	СССС	<b>ب</b>	NFIGENCE COMPLE CRER UPPER			
C.C.C. V.C.C.44 C.C.T.L C.1153 C.371C C.2657 C.3367 C.2657 C.3212 U.5239 C.3212 U.5239	UNDER UNERA UNDE UNER UNDE UNE UNDE UNER UN					•
C.3312 C.2657 C.3357 C.2000 C.3212 U.5829 C.3212 U.5829	U-0710 U-2657 C-3312 U-5229 C-3212 U-5229 C-05C W-1N SLUPTINGAAL SIPSUX) SLUPEICA U-05 CP65A SLUPTINGAAL SIPSUX) SLUPEICA					
C.3367 C.0000 C.3212 U.5829 C.3212 U.5829	(13367 21000 (13212 015829 (13212 015829 (1432 015829 (1432 01688 Na Price (Nicenal Stratus) SLOPE (SA 011688 Na Prices) (1432 016 15 016 15 016 15	ر ۲			•	•
0.3212 0.5239 0.527 0.529	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1		•			
5-6-5-5 <u>19-19-00</u>	LUNC WIN SLUPPINGAL SIPSUX) SLUPEICA U.1432 U.1432 U.1432 U U UUSE URGANATO PERUS	-	212			
5550	t x buss chick hat represent the chick have a chick hat represent the chick hat have been a chick ha	K. SULTS:	•			
	(x) wase chilis hat repeat of head the children wase			0.1432		

ذرك

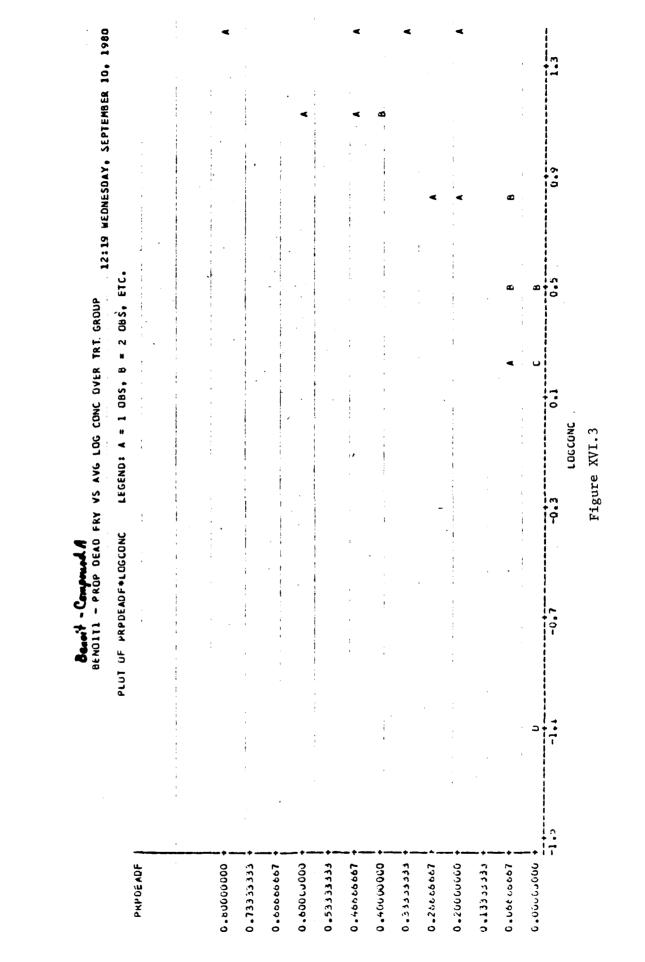
....

۰.



### Figure XVI.2

12:19 WEDNESUAY, SEPTEMBER 10, 1980 29 28 .27 26 52 23 24 12 . 21 LEGEND: A = 1 095, 8 = 2 085, ETC. 18 . 19 . 20 Benoit - Company A BENOILI - PROP DEAD FRY VS AVG CONC DVER TRT GROUP 17 16 CONCMEAN 13 14 15 ia PLUT UF PRPUEADF + CONCMEAN 10 . 11 œ 2 PRFLEADF G.80000CUO 25522525.0 V • 66666666 0.0000004.0 ÉÉÉÉÉÉÉÉ? 0 0.46666667 0.4400000000 5515555.0 0.20066667 0.26000600 25262661.0 0.36666667



### XVII. ANALYSIS OF QUANTITATIVE RESPONSES

Most of the preceeding discussion has been concerned with aspects of the analysis of quantal survival data. This includes preliminary graphical displays, tests for heterogeneity among tanks within groups, outlier detection procedures, adjustments for tank to tank heterogeneity, analysis of variance and multiple comparison procedures, and various types of dose response estimation procedures.

Quantitative responses such as length and weight are also recorded during and at the conclusion of toxicity tests. Weight measurements (in mgs) on all surviving fish at the conclusion of early life stage tests are standard. The statistical analysis of such responses is discussed in this section. This discussion is also directly relevant to the statistical analysis of the length and weight measurements that are taken at periodic intervals (e.g. at 30 day or 60 day intervals) in full life cycle tests.

The approach to analyzing weight and length data is directly analagous to the approach to analyzing survival data. Each aspect of analyzing quantal data, mentioned above, has a direct counterpart for analyzing quantitative data. In fact the procedures appropriate for analyzing quantitative responses are more "standard" and more familiar to most users of statistical methodology than those appropriate for analyzing quantal responses.

In addition to weight and length data a quantitative response often recorded and analyzed in mammalian toxicology tests is time to death (or time to tumor or time to any apriori specified event). Such time to death data provide more information than the 30 day, 60 day, 120 day, etc., survival rates that are commonly reported in aquatic toxicity tests. In particular knowledge of time to death of each embryo or fry yields the percent survival responses as a byproduct. However 30 day survival data will not reveal whether the fish that died did so on day 1, day 15, or day 29. Such information is important for understanding the mechanisms by which the toxicants act.

The analysis of time to death data involves working with censored responses. Parametric and nonparametric approaches to the analysis of such data are discussed in a number of books such as Gross and Clark [49] and Kalbfleisch and Prentice [50].

Time to death (to the nearest day) data is usually collected as part of the day to day test procedure since the tanks are examined daily (or at least on weekdays) for dead fish and these are removed and recorded. Unfortunately, time to death data is not routinely reported as part of the experimental results. In particular time to death of individual fish was not included in any of the data sets made available to us. As such, the analysis of censored life data is not discussed here. We would make a strong recommendation that time to death of individual fish be routinely reported in the future instead of or in addition to 30 day, 60 day, etc., percent survival data. This would require little additional cost or effort but could possibly provide valuable additional information.

In the remainder of the section we consider the various aspects of analyzing the weights recorded on survivors of 30 day early life stage tests. Before getting down to the technical and methodological issues. several conceptual points should be discussed. The primary difficulty in the interpretation of weight data is the confounding of weight gain with survival. A number of scenarios can be postulated leading to different conclusions about the relationships to be anticipated. Death can be thought of as the first order effect of the toxicant and weight loss (or lesser weight gain) as a secondary effect. If death and weight loss represent different degrees of severity of the same mechanism then one would expect that average weights of survivors would decrease as the mortality rate increases and lack of observation of such a decrease might be interpreted as lack of effect. However, since weights are measured just on survivors a selection phenomenon may be occurring. Presumably, in the various treatment groups the stronger fish survive while the weaker fish die. Presumably the weaker fish would have gained less weight on average than the stronger fish, had they survived (e.g., if they were in the control group). Since greater numbers of weak fish survive in the control and low treatment level groups than in the higher treatment level groups, these weak survivors might decrease the average weight gain relative to the strong survivors in the treatment groups. Thus an increase in observed average weight with treatment level might be possible, or if the toxicant reduces weight gain in the treatment groups, the selection and reduction effects may offset one another. therby resulting in no observable trend. Therefore the biological meaning of observed trends in weight gain with increasing concentration or the lack of observed trends depends very much on the biological assumptions about toxicant mechanisms and about association between survival and weight loss.

One way to reduce or eliminate the confounding of the survival and weight gain responses is to confine weight gain comparisons to those concentration groups whose mortality rates are not significantly (either biologically or statistically) greater than that in the control groups. The rationale for this viewpoint is that mortality is a first order effect while weight gain is a second order effect. Thus in groups with significant mortality, the question of reduction in weight gain is not of concern. Only when the mortality rate approximates that in the control group is the comparison of weight gains important.

We illustrate the analysis of the weight responses with fry data from the Holcombe and Phipps test on compound D. Recall that the observed mortality rates in the control group and the five treatment groups were 0.06, 0.08, 0.08, 0.13, 0.79, 1.00. It was shown in Section XII that the mortality rates in groups 5 and 6 differ significantly from that in the control group and group 4 is borderline (statistically) significant according to Williams' procedure. Thus it might be reasonable to confine weight gain comparisons with the control group to treatment groups 2, 3, and possibly 4.

### A. Preliminary Scatterplots

A number of plots were prepared showing the means and standard deviations of weight by tank plotted against group number, concentration, log concentration, proportion of survivors, and other variables. Several of these plots are shown in Figures XVII.1 to XVII.4. Figure XVII.1 shows average weight per tank vs group number. Group 6 does not appear in this plot since it had 100 percent mortality. A downward trend in average weight with increasing group number is evident. Note that the total range of variation in average weights is not very great -- between 100 mg and 145 mg. Figure XVII.2 shows standard deviation of weight per tank vs group number. With the exception of one tank in group 5 there appears to be no trend in standard deviation with group. Note that the scandard deviation estimates in group 5 are less stable than those in the other groups since they are based on many fewer observations. Figures XVII.3, XVII.4 show the average weight per tank vs concentration and vs log concentration respectively (more specifically log(1 + CONC)). The decreasing trend in average weight with increasing concentration is again evident. In Figure XVII.4 a linear or quadratic trend in log concentration can be seen among the treatment groups.

### B. <u>Outlier Detection Procedures and Testing for Tank to Tank Heteroge</u>neity Within Groups

Analysis of variance models were fitted to the individual weights and logarithmic weights to determine if there is any statistical evidence of tank to tank heterogeneity within groups or of differences in average weights across groups. The two way mixed model

$$W_{ijk} = \mu + \alpha_i + \tau_{j(i)} + \varepsilon_{ijk} \qquad i = 1, \dots, I \qquad j = 1, \dots, J$$
  
$$k = 1, \dots, n_{ij}$$

was specified where  $W_{ijk}$  corresponds to the weight or to the log weight of the k-th fish within the j-th tank of the i-th group,  $\alpha_i$  is the fixed group effect,  $\tau_{j(i)}$  is the random effect of the j-th tank within the i-th group, and  $\varepsilon_{ijk}$  is the experimental variation. It is assumed that  $\tau_{j(i)}$ are independent N(0,  $\sigma_{\tau}^2$ ) and  $\varepsilon_{ijk}$  are independent N(0,  $\sigma_{\varepsilon}^2$ ) and the  $\{\tau_{j(i)}\}, \{\varepsilon_{ijk}\}$  are independent. In the case of the Holcombe and Phipps Compound D fry mortality data I = 5, J = 4,  $n_{ij}$  varies with tank and with tank and with group but is nearly constant in groups 1 to 4. The model was fitted to the data using PROC GLM in the SAS statistical computing system [12]. The results are shown in Figures XVII.5 to XVII.7. Figures XVII.5 and XVII.6 show the analysis of variance tables for the responses weight and log weight respectively. Figure XVII.7 shows the expected mean squares of the entries in the analysis of variance tables in Figures XVII.5 and XVII.6. The conclusions from the fits to the untransformed weights and to the logarithmic weights are very similar. From the expected mean square of TANK (GRP) indicated in Figure XVII.7 it is seen that the hypothesis  $H_0: \sigma_T^2 = 0$  is tested by comparing the TANK (GRP) mean square with the error mean square. The resulting F tests are nonsignificant for both the untransformed and logarithmic weight responses (observed significance levels 0.64, 0.70 respectively). The estimated variance components are:

untransformed weights

$$\hat{\sigma}_{\tau}^2 = \left(\frac{11324.61}{15} - 900.35\right)/19.19 = -7.58$$

logarithmic weights

$$\hat{\sigma}_{\tau}^2 = \left(\frac{1.136}{15} - 0.097\right)/19.19 = -0.001$$

Thus in both cases  $\hat{\sigma}_{\tau}^2$  is set equal to zero. Therefore in this example the error mean square may be used as an error yardstick against which to compare the fixed effect mean squares for group effects.

In general  $\sigma_{\tau}^2$  will not be equal to zero and so an appropriate error yardstick will be a linear combination of the error mean square and the tank (group) mean square. To see how this works consider the test of the hypothesis H<sub>0</sub>: $\alpha_i = 0$ , that is no group effects. This null hypothesis is obviously false and the test given by PROC GLM, based on the error mean square with 366 d.f. rejects H<sub>0</sub> very strongly. However the error mean square underestimates the variability of the group mean square if  $\sigma_{\tau}^2$  is is greater than zero. The type IV expected mean square for group is shown in Figure XVII.7 to be  $\sigma_{e}^2 + 18.2048\sigma_{\tau}^2$ . Thus the tank (group) mean square with 15 d.f. is a more appropriate error yardstick than the error mean square estimates  $\sigma_{e}^2$ . In general a linear combination of these two mean squares would be an even better yardstick.

The classical approach to combining expected mean squares is based on choosing that linear combination which yields an unbiased estimator of  $\sigma_e^2$  + 18.2048 $\sigma_\tau^2$ . Namely

$$w[\sigma_e^2 + 19.1915\sigma_\tau^2] + (1 - w)\sigma_e^2 = \sigma_e^2 + 18.2048\sigma_\tau^2$$

Thus

$$w = \frac{18.2048}{19.1915} = 0.95$$

Note that w does <u>not</u> depend on  $\sigma_{\tau}^2$ .

If the design had been completely balanced, this would have led to using the tank (group) mean square with 15d.f. irrespective of the value of  $\sigma_T^2$  Such an approach is analagous to carrying out analyses on a per tank basis rather than on a per fish basis.

For the untransformed weights

.95 MS TANK (GRP) + .05 MS ERROR = .95(754.97) + .05(900.35) = 762.26

For the logarithmic weights

.95 MS TANK (GRP) + .05 MS ERROR = .95(0.0757) + .05(.0972) = 0.0768

To calculate the effective number of degrees of freedom of this linear combination, assume that

.95 MS TANK (GRP) + .05 MS ERROR 
$$\sim (\sigma_e^2 + 18.2048\sigma_{\tau}^2)\chi_v^2/v$$

where v is unknown. Equating the variances of the two sides we obtain:

$$(.95)^{2}(\sigma_{e}^{2} + 19.1915\sigma_{\tau}^{2})^{2}\frac{2}{15} + (.05)^{2}\sigma_{e}^{2}\frac{2}{366} = (\sigma_{e}^{2} + 18.2048\sigma_{\tau}^{2})^{2}\frac{2}{\nu}$$

Approximating the expectations by the mean squares yields

(a)untransformed weights

$$(.95)^{2}(754.97)^{2}\frac{1}{15} + (.05)^{2}(900.35)^{2}\frac{1}{366} = (762.26)^{2}\frac{1}{\nu}$$

or v = 16.9

(b)logarithmic weights

$$(.95)^{2}(.0757)^{2} \frac{1}{15} + (.05)^{2}(.0972)^{2} \frac{1}{366} = (.0768)^{2} \frac{1}{v}$$

or v = 17.1

Comparison of the type IV group mean square with this error yardstick yields

untransformed weights: 7649.96/762.26 = 10.04 logarithmic weights: 0.6743/0.0768 = 8.78

Although in the case of unbalanced data the numerator and denominator mean squares are not strictly independent and the denominator "mean square" is not strictly distributed as chi square, an approximate test is usually constructed by treating these ratios as F ratios with degrees of freedom 4 and 17. These ratios are of course very highly statistically significant according to this yardstick. Thus there is very strong statistical evidence of differences in average weights among groups.

It should be noted that this approach of using essentially the tank (group) mean square with 15d.f. as error yardstick is very close to carrying out group to group comparisons on a per tank basis rather than on a per fish basis. Figures XVII.8 shows the output from a one way analysis of variance to compare group means using the tank means as basic input data. This corresponds to a quantification of the relations seen in Figure XVII.1. Again the group effects are highly significant. The expected mean square for tank (group) as shown in Figure XVII.7 is as if each tank contained 19.19 fish on average. (The average is actually 19.3). If we divide MS TANK (GRP) by 19.19 we obtain 39.34 which agrees quite well with the error mean square of 39.40 in Figure XVII.8.

An alternative approach to pooling mean squares in the analysis of variance is based on finding that linear combination of tank (group) and error mean squares which minimizes the mean square difference from  $\sigma^2 + 18.2048\sigma_{\tau}^2$ . Namely choose w so that

$$E[w \text{ MS TANK (GRP)} + (1 - w) \text{ MS ERROR} - (\sigma_e^2 + 18.2048\sigma_\tau^2)]^2$$

is minimized. The resulting choice of w will depend on the relative magnitudes of  $\sigma_e^2$  and  $\sigma_\tau^2$ . As  $\sigma_\tau^2/\sigma_e^2$  approaches zero, more and more emphasis will be placed on the error mean square because its reduced variance will more than compensate for its bias. The above expectation can be calculated to be

$$w^{2}(\sigma_{e}^{2} + 19.19\sigma_{\tau}^{2})^{2}\frac{2}{15} + (1 - w)^{2}\sigma_{e}^{4}\frac{2}{366} + [(19.19w - 18.2048)\sigma_{\tau}^{2}]^{2}$$

We wish to choose  $w(0 \le w \le 1)$  to minimize this expression. If we substitute in the estimates of the mean squares and of the variance components we obtain

$$w^{2}(754.97)\frac{2}{15} + (1 - w)^{2}(900.35)^{2}\frac{2}{366} + 0 = 75997.3w^{2} + 4429.7(1 - w)^{2}$$

which has its minimum in the interval  $0 \le w \le 1$  at w = 0.055. Thus the second approach leads to the error yardstick

.055 MS TANK (GRP) + .945 MS ERROR = 892.35

with approximate degrees of freedom obtained by solving the equation

$$(.055)^{2}(754.97)^{2}\frac{1}{15} + (.945)^{2}(900.35)^{2}\frac{1}{366} = (892.35)^{2}\frac{1}{v}$$

or v = 380.

Thus this allocation effectively leads to the use of the error mean square in this case and is very different from that obtained by equating expected mean squares. This alternative pooling scheme is most useful when there are an inadequate number of degrees of freedom for estimating MS TANK (GRP) because it then puts more weight on the error mean square. The criterion used does not of course, insure that the resulting mean square is an unbiased estimate of  $\sigma_{e}^{2}$  + 18.2048 $\sigma_{\gamma}^{2}$ .

It should be noted that the GLM procedure permits the decomposition of the model sum of squares into individual degree of freedom components. This feature is illustrated in Figure XVII.8. The linear and quadratic components of trend are defined by the contrasts (-2, -1, 0, 1, 2) and (2, -1, -2, -1, 2) respectively. These contrasts single out physically important comparisons among the groups to test and estimate and thereby increase the sensitivity of the analysis of variance tests. This approach is analagous to carrying out a one sided measure of association test with qualitative survival data.

The residuals from the analysis of variance fits can be used to check distributional assumptions and to detect outliers. Figures XVII.9 - XVII.12 display the arithmetic and logarithmic residuals from the fits in Figures XVII.5 and XVII.6 respectively. Figures XVII.9 and XVII.10 show the residuals plotted vs group. No outliers are obvious. The variability seems constant with group. The residuals from the untransformed weights appear to be much more symmetric about zero than those from the logarithmic weights. Figures XVII.11 and XVII.12 show normal probability plots of the arithmetic and logarithmic residuals respectively. The plot in Figure XVII.11 looks much more nearly normal than that in Figure XVII.12.

The lowest two residuals in Figure XVII.11 lie below the line through the remainder of the data. To determine whether there is any statistical evidence that these observations are outliers we can test whether the most extreme of 386 independent normally distributed random variables with mean 0 and standard deviation 30 is likely to exceed 118 in absolute value. (The two extreme residuals correspond to observations 209 and 220, are from group 3, tanks A and B, have values -112.9 and -117.9, and are associated with fish having reported weights of 15 mgs and 10 mgs. I am assuming that these are the correct weights, but this should be checked).

P[most extreme of 386 observations greater than 118 in absolute value] =

$$1 - \left[P(-118 < X < 118)\right]^{386} = 1 - \left[2\Phi\left(\frac{118}{30}\right) - 1\right]^{386} = 1 - (.999950)^{386}$$
$$= 1 - .98 = .02$$

Thus there is statistical evidence that this extreme residual does not conform to the others. Whether this represents a clerical error or natural biological variation would need to be determined.

Assuming that the extreme observation is an outlier, the second most extreme observation can be compared to the extreme of 385 observations.

P[most extreme of 385 obsvns greater than 113 in absolute value] =

$$1 - [P(-113 < x < 113)]^{385} = 1 - [2\Phi(\frac{113}{30}) - 1]^{385} = 1 - (.99990)^{385}$$
  
= 0.037

There is thus statistical evidence that this second most extreme observation is also an outlier.

Basic records should be examined to determine if these observations are valid. If not, they should be corrected or deleted and the modified data reanalyzed. If they in fact represent natural biological variation then biological judgement should be used to determine whether or not to retain these observations with the remainder.

### C. Multiple Comparison Procedures and Regression Analyses

Based on the results of the analysis of variance calculations previously discussed, we can carry out comparisons of average weight gains across groups. The average weight gains and numbers of animals per group are:

Group	1	2	3	4	5
N	94	92	92	87	21
Average	131.2	135.59	127.9	113.1	108.1

The standard errors of these averages based on averaging the responses from four tanks and varying numbers of fish per group, are  $1/4(\sigma_{\tau}^2 + \sigma_e^2/23.5)$ ,  $1/4(\sigma_{\tau}^2 + \sigma_e^2/23)$ ,  $1/4(\sigma_{\tau}^2 + \sigma_e^2/23)$ ,  $1/4(\sigma_{\tau}^2 + \sigma_e^2/21.75)$ ,  $1/4(\sigma_{\tau}^2 + \sigma_e^2/5.25)$ . The variance components  $\sigma_{\tau}^2$ ,  $\sigma_e^2$  can be estimated by appropriate linear combinations of the tank (group) and error mean squares displayed in Figure XVII.5. For the Holcombe and Phipps fry mortality data  $\hat{\sigma}_{\tau}^2 = 0$  and so we can estimate the standard errors using the error mean square.

Thus

$$\hat{\sigma}_{e}^{2} = 900.355$$
 with 366 d.f.

Alternative standard error estimates with appropriate degrees of freedom can be constructed using approaches analagous to those discussed in the previous subsection.

We apply Williams' procedure [ 36 ] to determine which groups have (statistically) significantly lower weight gain than the control group. We first need to adjust these mean values so that they are in monotone decreasing sequence. We simply calculate weighted average (94 x 131.2 + 92 x 135.59)/186 = 133.37 of the averages in groups 1 and 2. The modified averages are now in monotone decreasing order. We declare the group i average weight gain to be significantly smaller than the control average if

$$\bar{X}_{i,adj} - \bar{X}_{1} < -\bar{t}\hat{\sigma}_{e}(1/N_{i} + 1/N_{1})^{1/2}$$

Note that  $\bar{X}_{i,adj}$  is the adjusted average whereas  $\bar{X}_{1}$  is the unadjusted average. We use the factor  $\bar{t}$  obtained from Williams' table, which is derived under the assumption of equal group sample sizes. This assumption is quite reasonable for groups 1-4. We assume that  $\hat{\sigma}_{e}^{2}$  is estimated with 366 d.f. A more conservative assumption might be use 15 d.f. since this is the amount of information concerning  $\sigma_{\tau}^{2}$ . This would raise  $\bar{t}$  from 1.75 to 1.88.

The group i mean is declared to be significantly lower than the control mean if

$$\bar{x}_{i,adj} < \bar{x}_{1} - \bar{t}\hat{\sigma}_{e}(1/N_{i} + 1/N_{1})^{1/2} = 131.2 - (1./30)(30.01)(1/N_{i} + 1/N_{1})^{1/2}$$

i = 2:	$\bar{X}_{2,adj} = 133.37$	critical value = 123.50
i = 3:	$\bar{x}_{3,adj} = 127.9$	critical value = 123.50
i = 4:	$\bar{x}_{4,adj} = 113.1$	critical value = 123.39
i = 5:	$\bar{x}_{5,adj} = 108.1$	critical value = 118.52

Thus the average weight gains in groups 4 and 5 are declared by Williams' procedure to be (statistically) significantly lower than the control group average at the 5 percent level. With respect to mortality, group 5 is obviously much different than the control group and group 4 is border-line (statistically) significantly different. Thus the quantal survival and quantitative weight responses yield essentially the same conclusions.

We can fit regression models to the weight gain data to quantify the trends in averages across groups. Figure XVII.4 shows average weight gain plotted against log(1 + CONC). The responses in the four treatment groups show a definite trend, mostly linear but possibly with some second order curvature. Figure XVII.13 shows the results of fitting a cubic polynomial in log (1 + CONC) to the treatment groups and an indicator function to the control group. Namely the model

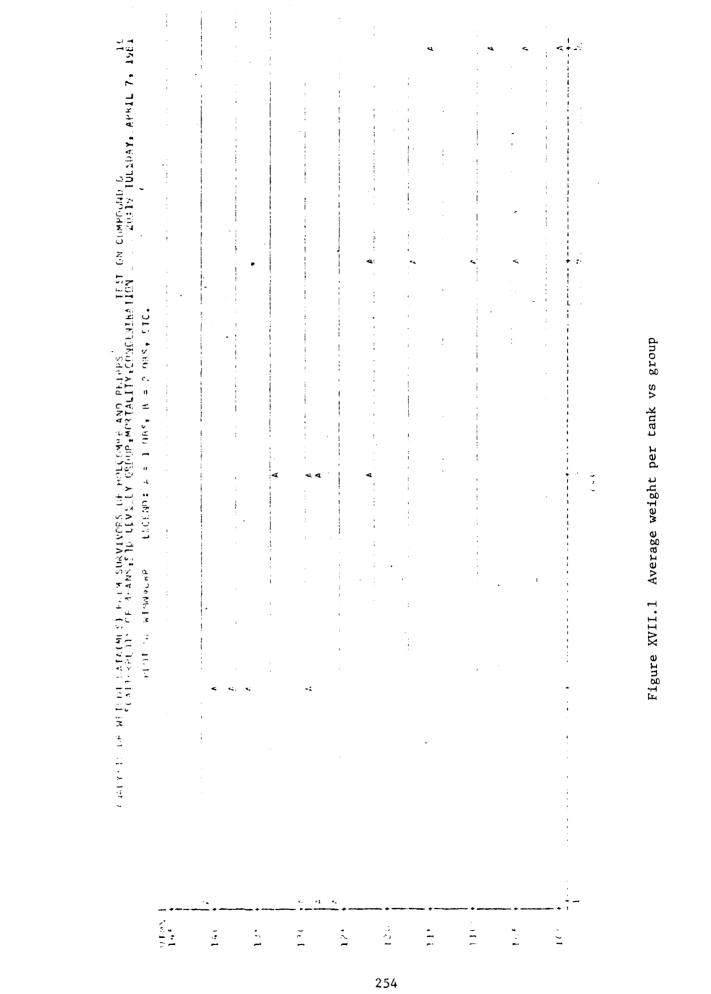
$$W_{ijk} = \beta_0 + \beta_1 \text{ ICTL} + \beta_2 \text{ LCONC} + \beta_3 \text{ LCONC}^2 + \beta_4 \text{ LCONC}^3 + \varepsilon_{ijk}$$

was fitted to the weight data where ICTL = 1 if GRP = 1 and 0 otherwise and LCONC = log(1 + CONC). This model fits a cubic polynomial to the treatment groups. The parameter  $\beta_1$ , represents the difference between the control group mean and the extrapolation back to LCONC = 0 along the cubic polynomial. The contrasts estimated at the bottom of the figure correspond to the differences between the mean responses at the treatment groups, based on the polynomial fit, and the control group response.

A complication in inference procedures arises if there is tank to tank heterogeneity within groups. Observations within the same tank are then dependent due to a common tank effect. The variation of the type IV mean square for GRPS is seen in Figure XVII.7 to be inflated from  $\sigma_e^2$  to  $\sigma_e^2$  + 18.2048 $\sigma_\tau^2$  due to such heterogeneity. The standard errors in Figure XVII.13 might then be inflated by the factor  $[1 + 18.2048\sigma_\tau^2/\sigma_e^2]^{1/2}$  to account for such heterogeneity. The quantity  $\sigma_e^2$  + 18.2048 $\sigma_\tau^2/\sigma_e^2]^{1/2}$  to account for such heterogeneity. The quantity  $\sigma_e^2$  + 18.2048 $\sigma_\tau^2$  can be estimated by pooling mean squares in Figure XVII.5 in at least one of two different ways, as discussed in the previous subsection. These yields estimates of 726.26 with 17 d.f. or 892.35 with 380 d.f. Alternatively, regression fits can be carried out on a per tank basis, as is commonly done. This turns out to be very similar to using the 17d.f. variance estimate.

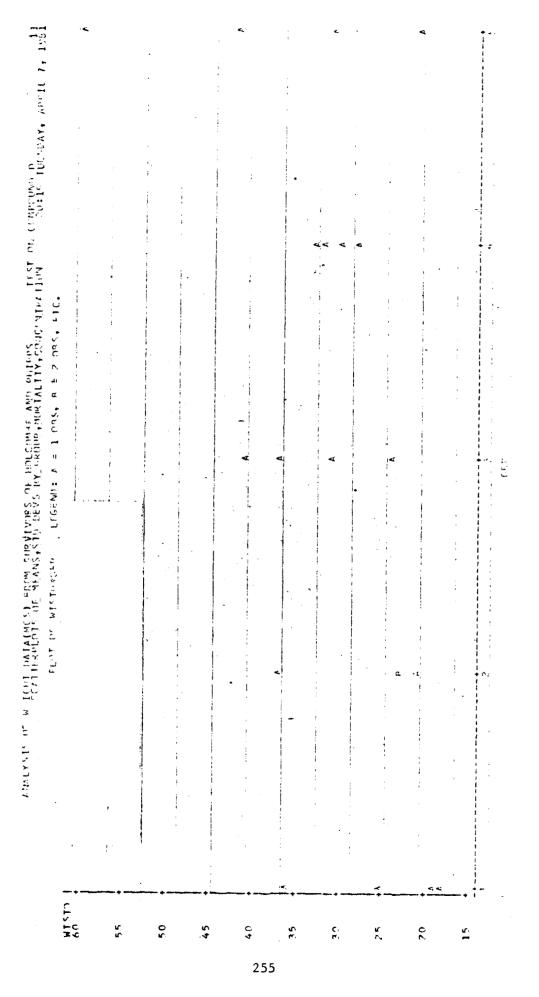
Since  $\sigma_{\tau}^2$  was nonsignificant in the previous ANOVA fit, since  $\hat{\sigma}_{\tau}^2 = 0$ , and since the variance estimate with 380 d.f. is very similar to the error mean square in Figure XVII.13, we will use the error mean square as the basis of standard error calculations in this example. It should be noted however that this is appropriate only if  $\sigma_{\tau}^2 = 0$ .

We see from the type I sums of squares in Figure XVII.13 that the linear component of trend is highly significant while the quadratic and cubic trends are nonsignificant over and above the linear trend. This agrees with the appearance of Figure XVII.4. The quadratic and cubic terms should be deleted and the model refitted. The contrasts at the bottom of Figure XVII.13 show, in agreement with the results from Williams' procedure, that groups 4 and 5 differ significantly from the control group while groups 2 and 3 do not.

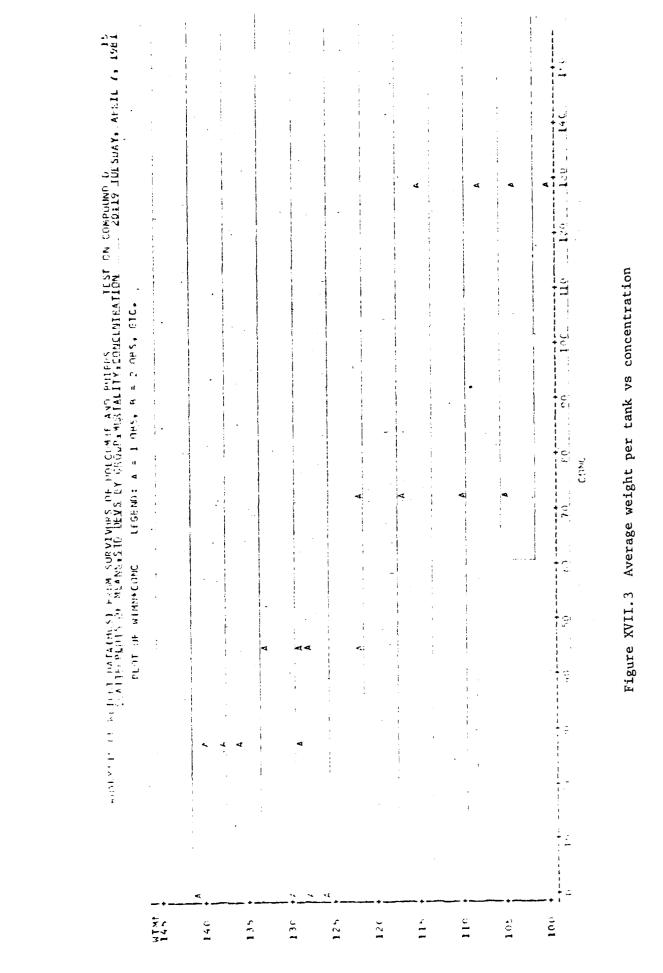


.

.

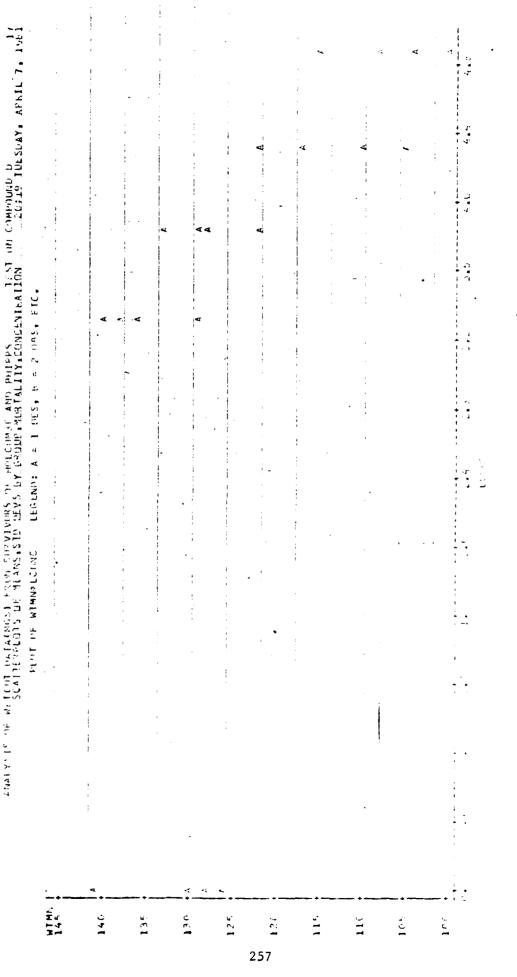


# Figure XVII.2 Standard deviation of weight per tank vs group



k





Ċ

j

TEST ON CUMPOUNU D 13:38 MONDAY, APRAL 13, 1981 ANALYSIS OF WEIGHT DATA(MGS) FROM SURVIVORS OF HOLCOMBE AND PHIPPS ANOVA ON URICINAL OCSERVATIONS TO DETECT QUILIERS \_ GENERAL LINEAR MODELS PROCEDURE

SOURCE DF SUM 0 MODEL 19 43953							
51	SUM OF SQUARES	MEAN SQUARE		F VALUE	PR & F	R-SUARE	C.V.
	43953.20526379	2313, 32659283	9283	2.57	0.004	V.117665	23+7569
ERKUR 366 329529	329529.78955487	900.35461627	1627		STG DEV	٠	WGHT MEAN
CORRECTED TOTAL 385 373482	373482.94401865	•			30.00550269	120	126-13471503
SOURCE DF	TYPE I SS	F VALUE	PR > F	СF Г	TYPE IV SS	F VALUE	<b>FX &gt; F</b>
GR / 4 7 32628 TANK (GRP) 15 11324	32628 • 6044 3368 11324 • 6008 3011	40° 40° 40°	0.0001	15	30599.85439173 11324.66083011	H.50 0.84	0.0001

Analysis of variance fit of individual weights on group and tank within group. Holcombe and Phipps Compound D fry mortality data Figure XVII.5

TEST UN CUMPOUNU D 22 131.38 MONLAY, APRIL 14, 1941 ANALYSIS OF WEIGHT DATAIMGS) FROM SURVIVORS OF HOLCOMBE AND PHIPPS ANGVA ON ORIGINAL OBSERVATIONS TO DETECT OUTLIERS GENERAL LINEAR MODELS PROCEDURE

чт (л.

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	QUARE	F VALUE	PR > F	R-SQUARE	C.V.
MODEL	12	12400120-4	0.211	0.21194786	2.18	0.0032	212101-0	0.4966
ERKUR .	366	35 . 56 34 9376	160-0	0.09716802		STD DEV		LWGHT MEAN
CORRECTED TOTAL	385	£0£05065*6£				0.31171785		-4_24465 <u>244</u> .
SOURCE	DF	TYPE 1 SS	F VALUE	PR > F	DF	TYPE IV SS	F VALUE	FK > F
GRP Tankickp)	451	2.09106672	7.44	0.0001	15	2.6970731	6-94 0-78	6.0001

Figure XVII.6 Analysis of variance fit of individual logarithmic weights on group and tank within group. Holcombe and Phipps Compound D fry mortality data

	AMERTALS OF MELTAN VANNOS ON CALGINAL OUSERVATIONS TO DETECT OUTLIERS	LESI UN LUMPUONU UN APRIL 13. 1981	APRIL 13, 1981
DEPENDENT VARIABLES: WGHT LWGHT	ES: WOHT LNGHT		
SOURCE GRP	TYPE, I EXPECTED MEAN SQUARE VAR(ERROR) + 19.0254986 VAR(TANK(CRP)) + Q(GRP)	•	
TANKIGRP)			•
SOURCE	TYPE IV EXPECTED MEAN SQUARE		•
GRP	VARIERROK) + 18,2048497 VARITANKIGRP)) + 01GRP)		
TANK (GRP )	VAR(ERKCH) + 19.1914733 VAR(TANK(GRP))		

•.

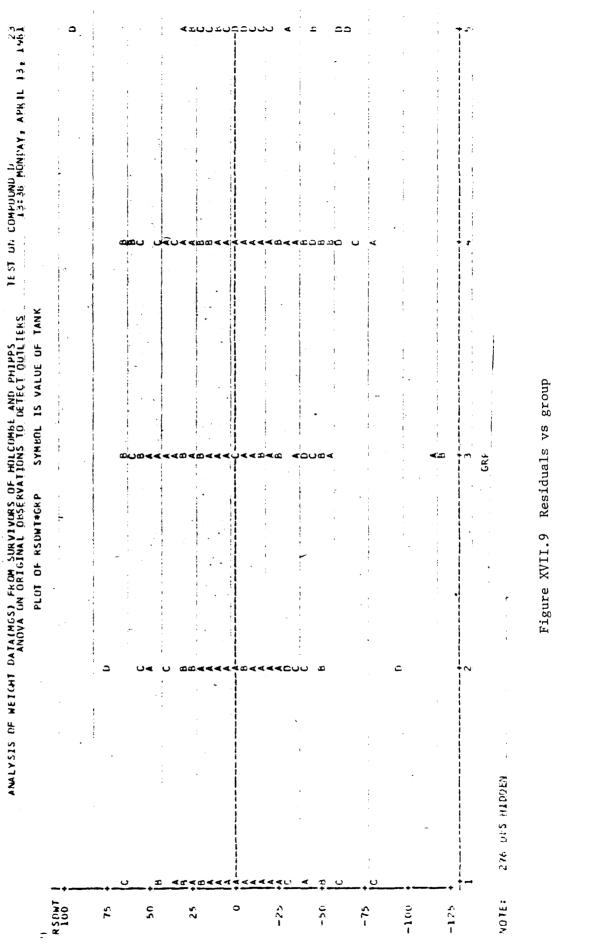
· .

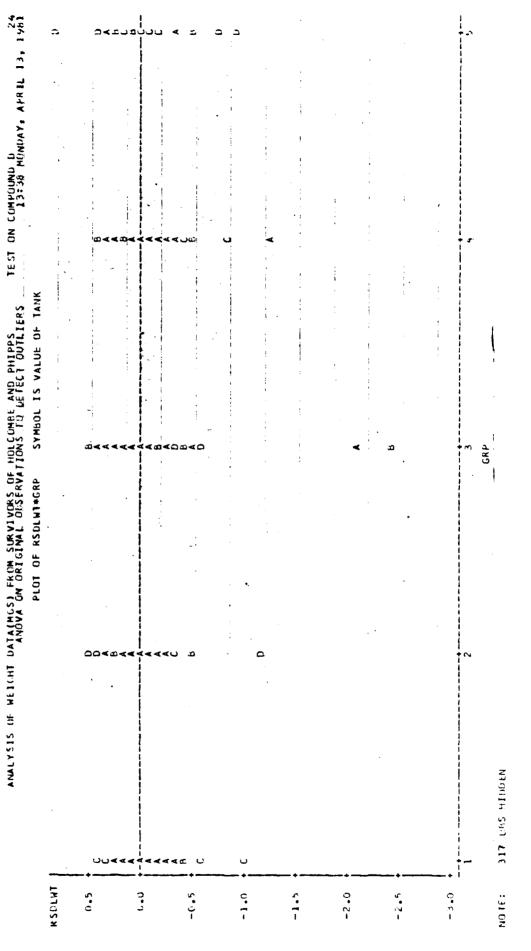
-

Expected mean squares in analysis of varinace fits to Holcombe and Phipps fry mortality data Figure XVII.7

د.۷.) 1.1030 WIME MLAN 123.00105100 4 **~** ~ ~ 1 TI ST ON CUMPUUNU U 13538 MUNUAY, AFK IL 13, 1961 1009.0 VALUE 15.04 I UNS . ALAMEU SY ALAMETEKS IMABEI F-SUUA+ L 4543944.V. TYPE IV SS 2465.21503461 PR > F STU DEV 51211312.00 1007-0 HAS BEEN EMPLOYED TO SOLVE HUT HE NORMEL EVENTIONS ES BLUE FON SOME LINEAR CUMBINS BLUE FON SOME LINEAR CUMBINS BIALNEU FROM THE NEWER IN BARE BLUE ANTON AND THE PARAMETER 4.43830786 4.43830786 4.43830786 4.43830786 4.43830786 3.13635759 9.92435808 1.74265864 ö STD FARDE OF STD FRROR ( ESTIMATE ANALYSIS OF WEIGHT DATA(MGS) FROM SURVIVERS OF HOLCOMBE AND PHIPPS 5 F VALUE 15-64 GENLRAL LINEAR MODELS PROCEDURE 0.0201 PR > F PR > F 0.0001 Ξ DI ESTIVALUTHE VARAMETER BUI ARE BOR DF THE RELACE ESTIMATORS MAY BE TORS THE RELACE STAND OF THE B TORS NUT FOLLOWIN FY THE LETTER IMATES NUT FOLLOWIN FY THE LETTER 29957502,010. MEAN SQUARE 39.39715335 0.000.0 PR > 111 0.0223 FK > 11 0.127 SQ G. NFKAL 12ED 52.29 F VALUE 15.64 VALUE u, E S -7.23 T FOR HQ: Parameter=D 33.91 5.61 1.60 1.60 1.60 T FOR HO : PARAMETER=0 MANY , t 2465,21503461 AND AND TYPE I SS ŝŝ 2465.21503461 590.95730026 3056.17233487 255.60134117 SUM OF SQUARES INGUI 106-4166667 B 24-91983696 B 2909528710 B 21-7452899 B 7-16726199 B 0-0000000 B EN DEEMED S CTED VALU -11.76570911 ESTIMATE **ESTIMATE** 5 200 00" MATOR 61 ЧÖ 51 ٢ ď DEPENDENT VARJABLE: WIMN EIBLASED ESTI ž CORRECTED TOTAL N m + m GRP LINEAR GRP QUADRATIC GRP LINEAR GRP QUADRATIC THE X•X THE ABO THE ABO THE LET TOR ARE FUNCTIO INTERCIPT GRP PAKA4E TER PAKAME TER **CONTRAST** SOURCE SOURCE NUTE: MOPEL ERKIN Gar

Holcombe and Phipps Analysis of variance fit of average weights per tank on group. Compound D fry mortality data Figure XVII.8







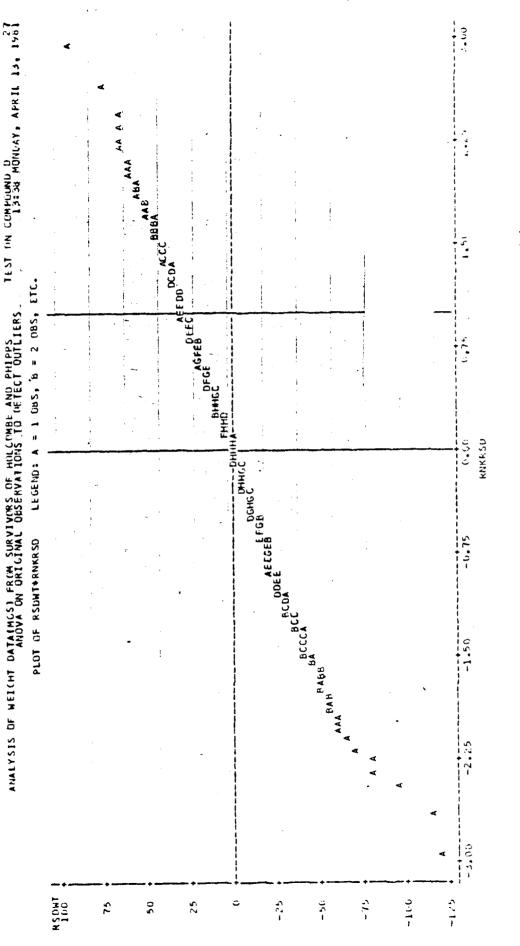
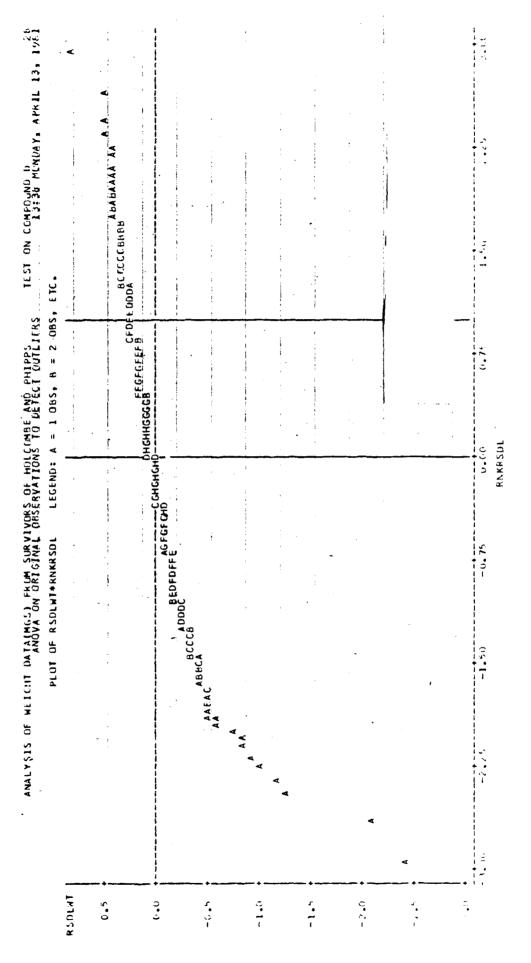


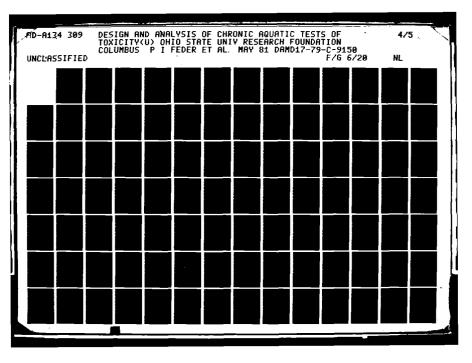
Figure XVII.11 Normal probability plot of residuals from untransformed weights

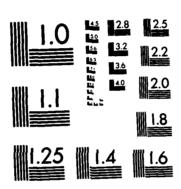


d

Figure XVII.12 Normal probability plot of residuals from logarithmic weights

٦





MICROCOPY RESOLUTION TEST CHART NATIONAL BUREAU OF STANDARDS-1963-A

ANALYSIS UF WEICHT DATA(MGS) FRUM SURVIYORS OF HILCONDE AND PHIPPS TEST ON COMPOUND D 	SOURCE UF SUM OF SQUARES MEAN SQUARE F VALUE PR > F ~-SQUARE C.V. MDEL 4 22628.60440373 8157.15110093 9.12 0.0001 0.087363 23.7130 ERMM 381 340854.39041492 894.63094597 5.10 510 05V 0.087363 23.7130 CONRECTIONINTAL 395 373482.99481865 126.13471503	SQUACE DF TYPE'L SS F VALUE PR > F 7 0F TYPE IV SS F VALUE PR > F 7 0F TYPE IV SS F VALUE PR > F 171.67790459 1.31 0.2532 1.201 1.21 0.2532 1.21 0.2532 0.05459 1.31 0.2532 0.2345 0.2345 0.2345 0.2345 0.2345 0.2226 1.31 0.2194 1.42 0.2226 1.42 0.2226 1.42 0.2226 1.42 0.2294 1.51 0.2294 1.51 0.2194 1.45 0.2226 1.45 0.2226 1.45 0.2294 1.51 0.2194 1.45 0.2226 1.45 0.2226 1.45 0.2294 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.2194 1.51 0.5194 1.51 0.5194 1.55 0.50 0.50 0.51 0.51 0.51 0.51 0.51	PAKAMI LEK         T FOR HO:         P. VIT         STD E970 CF           INIFACUPT         -1368-84174         -1606         P. VIT         STD E970 CF           INIFACUPT         -1368-84174         -164         0.2970         1310.6785176           INIFACUPT         -1368-174         -1.064         0.2970         1310.6785176           ICOVC         -246.50040278         -1.12         0.2970         1310.77316503           ICOVC         -246.50040278         -1.12         0.2970         1310.7731650           ICOVC         -246.50040278         -1.12         0.29365         244.500           ININECUDICAL         -3365545         0.23365         0.23365         244.501           ININECUDICAL         -3365545         0.23365         0.24401         0.2346           ININECUDICAL         -3365545         0.23365         0.24401         0.2441           ININECUDICAL         -3365545         0.2421         0.2441         0.2441           ININECUNICAL         -3365545         0.2421         0.2441         0.2724           ININECUNICAL         -3365545         0.2421         0.2421         0.2724           ININECUNICAL         -3407         0.2421         0.2421         0.2421
-------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

9

Regression fit of weight gain on log(1 + CONC). Holcombe and Phipps Compound D fry mortality data Figure XVII.13

#### XVIII. EXPERIMENTAL DESIGN CONSIDERATIONS

In this section we consider a number of issues pertaining to the design and conduct of aquatic toxicity tests such as precision to be expected as a function of sample size, allocation of tanks among treatment groups, additional variables to measure, numbers of tanks to be run.

### A. <u>Assumptions, Additional Variables to Measure, Numbers and Allocations</u> of Tanks

We briefly discuss several assumptions and recommendations associated with planning toxicity tests.

It is assumed that the constraint on size of test is the number of fish tanks that can be run within cost, manpower, and apparatus limitations. The cost of running a test is assumed to be directly proportional to the number of tanks run. Once a tank is put on test, the fish are essentially free. Thus a sample size strategy for early life stage tests is to run as many tanks as can be afforded and fill each tank with the maximum number of embryos or fry that is biologically sensible. This may differ for full life cycle tests.

A sufficient number of fish tanks should be included in the tests to be able to detect the presence of tank to tank heterogeneity within groups. If we have enough degrees of freedom among tanks within groups, the need for pooling mean squares in the analysis of variance in order to improve the sensitivity of tests or estimates of treatment effects is diminished. We can analyze the data on a per tank basis if we wish to, which is appropriate whether or not tank to tank heterogeneity is present. The main difficulty with per tank analyses is when the lack of adequate degrees of freedom diminishes precision of inference. We recommend that there be at least 12 d.f. to estimate tank to tank variation within groups. This would correspond to an average of 3 tanks per group if 6 treatment groups were being run. However to account for the possibility zero percent mortality in the control group and 100 percent mortality in the highest treatment group, it is suggested that 4 replicate tanks per group be run. This would provide 12d.f. for estimating variability just based on the results of the four intermediate groups. A glance at the charts of the noncentral t or noncentral F distribution shows that the power of analysis of varaince type tests based on 12d.f. is nearly as great as that with the infinite number of degrees of freedom. For example for  $\alpha = 0.05$  and one d.f. in the numerator

noncentrality ( $\emptyset$ )	1.5	2.0	2.25	2.5	2.75	3.0
power, 12 d.f.	0.50	0.74	0.83	0.90	0.945	0.972
power, ∞ d.f.	0.56	0.81	0.89	0.94	0.973	0.989

The differences in power are of little qualitative importance.

An important distinction should be recognized between the numbers of tanks needed to estimate the variability of treatment group responses as opposed to the numbers of tanks needed to reduce the tank to tank variability. Suppose that we run J tanks per treatment group, n fish per tank, and that  $\sigma_T^2$ ,  $\sigma_e^2$  represent the components of variation between tanks and between fish within tanks respectively. Then the variance of a treatment group average is

$$\frac{\sigma_{e}^{2}}{\sigma_{J}^{2}} + \frac{\sigma_{\tau}^{2}}{J}$$

If  $\sigma_T^2$  is large relative to  $\sigma_e^2/n$ , the only way to <u>reduce</u> the variance is by increasing J. However for fixed J, the ability to <u>estimate</u> this variance with 12 d.f. yields nearly as much sensitivity of tests and confidence intervals as an estimate with infinite d.f. This is also reflected in the fact that the upper 97.5 percentile of the t distribution is 2.78 with 4 d.f., 2.18 with 12 d.f., and 1.96 with infinite d.f. Thus 12 d.f., is most of the way between 4 d.f., and infinite d.f.

It should be noted that while we recommend at least 12 d.f., for estimating variablity, the tanks do <u>not</u> necessarily need to be equally replicated across groups. In fact we will suggest an unequal allocation later in this section.

Another important aspect of planning experiments is specifying classes of variables. At least six classes of variables can be distinguished: responses to be measured, controlled experimental variables, blocking variables, variables to be held constant, covariates to be measured, and variables to be randomized over. Rather than present a detailed discussion of each class of variable, we will emphasize those aspects which either vary from practice or are less obvious. In addition to the responses currently measured, it was argued in the previous section that individual times to death should be reported. This response must be measured but usually is neither reported nor analyzed. Obvious blocking variables are fish tank or test series. Other, less obvious blocking variables that might be incorporated into investigations are homogeneous subsets of fish (e.g., offspring of common parents, fish raised in the same breeding chamber, fish purchased from a single supplier at a single time, etc), investigators, laboratories, time period when test was conducted, technician, and many others. Some of the latter blocking factors would most naturally occur in round robin tests. In any test program a number of variables are held constant, at least nominally. Examples are water temperature, pH, hardness, levels of additives or impurities; type, amount, and frequency of food; type of fish tank; photoperiod. All these variables must be reported with the experimental results so that experimental conditions can be repeated and results compared across laboratories. Differences in variables held constant will sometimes account for discrepancies

in results. Covariates are factors which cannot be controlled but which can be measured and taken into account when analyzing the data. Covariates commonly reported are deviations from nominal in either controlled experimental variables or in variables held constant. The most obvious covariate is actual test concentration. This should be determined periodically in each tank and reported. The analysis of the data should be based on actual toxicant concentrations rather than nominal. The question of how to summarize toxicant levels has biological as well as statistical aspects. For example is effective level the average, the median, the maximum or some other? The question of frequency of measurement pertains to the short term effects of fluctuations in levels. The greater the effect of short term fluctuations in toxicant levels, the more frequently they need be measured. This aspect will not be considered further here. Other covariates to be measured and used for analysis might be water temperature. hardness, or pH, and measures of the size or health of the brood stock from which the test fish were taken. The remaining class of variables -- those to be randomized over -- is perhaps the most numerous. However the variables thought to be most important were included in the other five categories. These variables, many of which are not explicitely known, are randomized over. Their effects thus enter into the experimental variability. It is hoped that their effects are not too great.

#### B. Sample Size and Power Considerations for Quantal Survival Data

We first assume that there is no tank to tank variation within groups. We later modify the results to account for tank effects by adjusting sample sizes downward to "effective sample sizes."

If there is no tank to tank variation then it suffices to consider just the number of fish run per control or treatment group. Suppose that there is a control group (group 0) and I treatment groups (groups 1, 2, ..., I). In standard practice I = 5. Suppose that we run N<sub>0</sub> fish in the control group and N fish in each treatment group. Then test then involves a total of N<sub>0</sub> + IN  $\equiv$  C fish. If we carry out pairwise comparisons of treatment and control groups based on the arcsine transformation of observed response rates, the variances of 2 arc sin  $\sqrt{\hat{p}_1} - 2$  arc sin  $\sqrt{\hat{p}_0}$ are  $1/N_0 + 1/N$ . We wish to allocate fish to treatment and control groups so as to minimize  $1/N_0 + 1/N$  subject to N<sub>0</sub> + IN = C, fixed. This is a Lagrange multiplier problem whose solution is N<sub>0</sub> = NI<sup>1/2</sup>. Thus the more treatment groups, the greater is the sample size in the control group relative to that in the treatment groups. This is because the control group enters into all pairwise comparisons whereas each treatment group enters into just one. The suggested sample sizes are then

 $N = C/(I + I^{1/2})$  $N_{o} = CI^{1/2}(I + I^{1/2})$ 

This implies that for every 100 fish tested, the allocation between treatment and control groups would be

<u>    I     </u>	<u>Control (per 100 fish</u> )	Treatment(per 100 fish)
1	50	50
2	41.4	29.3
3	36.60	21.13
4	33.33	16.67
5	30.90	13.82
6	28.99	11.84
7	27.43	10.36
8	26.12	9.23
9	25	8.33
10	24.03	7.60
11	23.16	6.98
12	22.40	6.47

We see that the allocation is far from equal if I is moderate. For example if I = 5, the control group gets 2.5 times as many fish as any of the treatment groups.

How effective in increasing sensitivity of inferences is this departure from equal allocation? To determine the sensitivity of various sizes of tests to detect increases in response rates between control group and treatment groups, we carried out a series of power calculations. The null and alternative hypotheses considered were:

$$H_o: p = p_o$$

 $H_1: p > p_0$ 

We estimate  $p_0, \ p$  by the sample response rates  $\hat{p}_0, \ \hat{p}$  and reject  $H_0$  at  $\alpha$  = 0.05 if

2 arc sin 
$$\sqrt{p}$$
 - 2 arc sin  $\sqrt{p}$  > 1.645(1/N + 1/N)<sup>1/2</sup>

Simultaneity considerations are ignored in this calculation. The power of this test is calculated for various levels of N, p,  $p_0$ . The expression for the power is

$$1 - \Phi \left[ 1.645 - \frac{(\phi_{1} - \phi_{0})}{\sqrt{1/N_{0} + 1/N}} \right]$$

where  $\Phi(\cdot)$  is the standard normal c.d.f.  $\phi_i = 2 \operatorname{arc} \sin \sqrt{p_i}$  $\phi_o = 2 \operatorname{arc} \sin \sqrt{p_o}$ 

Calculations were made for the cases of equal allocation (i.e. N fish per group) and "optimal" allocation (i.e.  $N(I + 1)/(I + I^{1/2})$  fish in each treatment group and  $N(I + 1)I^{1/2}/(I + I^{1/2})$  fish in the control group. The usual situation, I = 5, is considered. The results are shown in Table XVIII.1.

TABLE XVIII.1	POWE	R OF ONE S	IDED PAIRWI	SE COMPARIS	ONS OF SURVIN	AL RATES
	BETW	EEN CONTRO	L GROUP AND	TREATMENT (	GROUPS. $\alpha =$	0.05.
$p_0 = 0.001$	<u>N</u>	50	75	100	150	200
	<u>n</u>					
p = .05		0.62*	0.77	0.87	0.96	0.988
		0.67*	0.82	0.90	0.98	0.995
.10		0.87	0.96	0.99	0.999	1.000
		0.91	0.98	0.995	1.000	1.000
.15		0.96	0.995	0.999		
		0.98	0.998	1.000		
.20		0.99	1.000			
		0.996	1.000			
. 30		1.000				
. 50		1.000				
		1.000				
p = .05						
p = .05	N	50	75	100	150	200

<u>p05</u>	<u>N</u>	50	75	100	150	200
p = .10		0.25 0.27	0.32 0.35	0.39 0.42	0.51 0.55	0.61 0.66
.15		0.53 0.58	0.68 0.73	0.79 0.83	0.91 0.94	0.96 0.98
.20		0.77 0.82	0.90 0.93	0.96 0.98	0.993 0.997	0.999 1.000
. 30		0.97 0.98	0.996 0.999	1.000 1.000	1.000 1.000	
.40		0.998 1.000	1.000 1.000			

\*Top number represents power under equal allocation. Bottom number represents power under optimal allocation, I = 5.

$p_0 = .10$	<u>N</u>	50	75	100	150	200
p = .15		0.19	0.24	0.28	0.37	0.45
		0.20	0.26	0.31	0.41	0.49
.20		0.41	0.54	0.64	0.79	0.88
		0.45	0.58	0.69	0.84	0.92
. 30		0.83	0.93	0.98	0.998	1.000
		0.87	0.96	0.99	0.999	1.000
.40		0.98	0.997	1.000	1.000	
		0.99	0.999	1.000	1.000	
.50		0.999	1.000			
		1.000	1.000			

$p_0 = 0.15$	<u>N</u>	50	75	100	150	200
p = .20		0.16 0.18	0.20 0.22	0.24 0.26	0.31 0.34	0.37 0.41
. 30		0.57 0.62	0.72 0.77	0.82 0.86	0.93 0.96	0.98 0.99
.40		0.89 0.92	0.97 0.98	0.992 0.997	1.000 1.000	1.000 1.000
.50		0.99 0.994	0.999 1.000	1.000 1.000		

The following conclusions can be drawn from Table XVIII.1.

- The ability to discriminate between treatment group and control group mortality rates varies considerably with N, p<sub>0</sub>, p. Thus these calculations provide some idea of the discrimination capability of the test as a function of size. Remember of course that these calculations don't account for the effects of heterogeneity among tanks.
- 2. The effect of allocating more fish to the control group than to the treatment groups is minor. Equal allocation yields nearly as good power as "optimal" allocation and is logistically much simpler.
- 3. If <u>no</u> assumptions can be made about the magnitudes of survival rates to be expected at the various concentration groups then equal allocations should be used.
- 4. If we can say something a priori about the survival rates at the various treatment groups, then we should have larger sample sizes in the lower concentration groups and smaller sample sizes in the higher concentration groups. The aim should be to even out the power of treatment group-control group comparisons as much as possible across groups.

We now consider adjustments in power calculations to take account of tank to tank heterogeneity. We do this by adjusting downward the effective sample size in each group and then entering Table XVIII.1 with the effective sample size. We calculate the adjustment by use of the beta binomial model [21].

Suppose that the test consists of I groups, J tanks per group, and n organisms (fish or embryos) per tank. Thus the actual sample size per group is N  $\equiv$  Jn. Let X<sub>ij</sub> denote the number of responses (dead, abnormal, etc) in the j-th tank of group i. We assume that X<sub>i</sub> is binomially distributed with parameters (n, p<sub>ij</sub>) and p<sub>ij</sub> is in turn<sup>j</sup> beta distributed with parameters ( $\alpha_i$ ,  $\beta_i$ ). This model allows for random variation of the p<sub>ij</sub>'s within the i-th group. Let

 $\mu_{i} \equiv \alpha_{i} / (\alpha_{i} + \beta_{i})$  $\theta_{i} \equiv 1 / (\alpha_{i} + \beta_{i})$ 

Then  $E(p_{ij}) = \mu_i$ 

$$Var(p_{ij}) = \mu_i (1 - \mu_i) \theta_i / (1 + \theta_i)$$

The unconditional distribution of  $X_{ij}$  is beta binomial with mean and variance

$$E(X_{ij}) = n\mu_i$$
  
Var (X<sub>ij</sub>) =  $n\mu_i(1 - \mu_i) \frac{1 + n\theta_i}{1 + \theta_i}$   $0 \le \theta_i < \infty$ 

Assume that  $\theta_{i} \equiv \theta$  is constant across groups. Let

$$K \equiv \frac{1 + n\theta}{1 + \theta}$$

This is the variance inflation factor due to tank to tank heterogeneity. The effective sample size per tank is then n/K and so the effective sample size per group is

$$N_{eff} = Jn/K = N \frac{1+\theta}{1+n\theta}$$

As  $\theta \rightarrow 0$ ,  $N_{eff} \rightarrow N$ , the number of organisms. As  $\theta \rightarrow \infty$ ,  $N_{eff} \rightarrow J$ , the number of tanks. As  $\sigma \rightarrow \infty$ ,  $N_{eff} \rightarrow J(1 + \theta)/\theta$ . Thus the effective number of organisms per tank asymptotes out as the actual number increases. Figure XVIII.1 shows a plot of  $n_{eff} \equiv n(1 + \theta)/(1 + n\theta)$  vs n for various values of  $\theta$ . We see the diminishing returns of placing more and more fish per tank in the presence of tank to tank heterogeneity. However under the cost structure assumed in this section we still place the maximum number of organisms within each tank, which we assume is 50 for embryos and 25 for fry. These numbers of course are only working assumptions.

To get some feeling for the meaning of  $\theta$  in terms of variance inflation factors, we calculate the factors corresponding to n = 25, n = 50 for various values of  $\theta$ .

θ	0	0.01	0.025	0.05	0.10	0.50	1.00
Var.infl.fact.,n=25	1	1.24	1.61	2.19	3.18	9	13
Var.infl.fact.,n=50	1	1.5	2.2	3.3	5.5	17.3	25.5

We calculated variance inflation factors for several sets of Fathead Minnow data in Section IX. These were

1. Holcombe and Phipps compound D fry mortality

$$n = 25$$
,  $\hat{K} = 1.337 \equiv (1 + 25\theta)/(1 + \theta)$ 

Thus  $\hat{\theta} = 0.014$  and  $n_{eff} = 18.78$ . Thus  $N_{eff} = 75$ .

2. Jarvinen compound B embryo mortality

LARKE LANDER

$$n = 50, K = 3.071 \equiv (1 + 50\theta)/(1 + \theta)$$

Thus  $\hat{\theta} = 0.044$  and  $n_{eff} = 16.31$ . Thus  $N_{eff} = 65$ .

# C. <u>Expected Precision for Comparison of Treatment Group and Control</u> <u>Group Survival Probabilities</u>

In the previous subsection we calculated the power to be expected for pairwise treatment group-control group comparisons of survival probabilities as a function of  $p_0$ , p, N. In this subsection we calculate the expected half lengths of two sided 95 percent confidence intervals on  $p - p_0$  for these same combinations of  $p_0$ , p, N. We again assume no tank to tank heterogeneity within groups and account for such heterogeneity by reducing the effective sample sizes, as discussed in the previous subsection. We base the precision calculations on asymptotic normal theory. Namely the confidence interval half length is calculated as  $1.96[(p q + pq)/N]^{1/2}$ . Asymptotic normality may not be a very good assumption when N = 50 or when  $p_0 = 0.001$ , but it is only being used for planning purposes.

# TABLE XVIII.2 EXPECTED HALF LENGTHS OF 95 PERCENT TWO SIDED CONFIDENCE INTERVALS FOR COMPARISONS BETWEEN TREATMENT GROUP AND CONTROL RESPONSE RATES

$p_0 = 0.001$	<u>N</u>	50	75	100	150	200
p = .05		.06	.05	.04	.04	.03
.10		.08	.07	.06	.05	.04
.15		.10	.08	.07	.06	.05
.20		.11	.09	.08	.06	.06
.30		.13	.10	.09	.07	.06
.40		.14	.11	.10	.08	.07
.50		.14	.11	.10	.08	.07
$p_0 = 0.05$						
p = .10		.10	.08	.07	.06	.05
.15		.12	. 09	.08	.07	.06
.20		.13	.10	.09	.07	.06
.30		.14	.11	.10	.08	.07
.40		.15	.12	.10	.09	.07
.50		.15	.12	.11	.09	.08
$\frac{p_0 = 0.10}{2}$						
p = .15		.13	.11	.09	.07	.06
.20		.14	.11	.10	.08	.07
. 30		.15	.12	.11	.09	.08
.40		.16	.13	.11	.09	.08
.50		.16	.13	.11	.09	.08

$p_0 = 0.15$	<u>N</u>	50	75	100	150	200
p = .20		.15	.12	.11	.09	.07
.30		.16	.13	.11	.09	.08
.40		.17	.14	.12	.10	.08
.50		.17	.14	.12	.10	.09

#### D. Power Calculations for Quantitative Weight Response

In previous subsections we calculated expected power and expected estimation precision for the quantal survival response. In this subsection we carry out similar calculations for the quantitative weight response. Distributional agsumptions are based on the results of analyzing the Holcombe and Phipps compound D fry weights in Section XVII. There was considerable group to group variation in survival proportions but not as much variation in the weights of the surviving fry. In particular

	Control Group	Group 4	Group 5	<u>Group 6</u>
Survivors	94/100	87/100	21/100	0/100
Avg.Wt.of Survivors(mg)	131.2	113.1	108	
Max. Wt.	190	186	195	
Min. Wt.	45	29	34	
Std. Dev.	25.8	34.0	36.7	

Variability is not too dependent on concentration group or on survival rate. The variance components are assumed to be constant across treatment groups.

The power and precision calculations below are based on a number of assumptions.

- 1. There is no tank to tank variation within treatment groups. We discuss corrections for such factors later in the subsection.
- 2. Equal sample sizes among treatment and control groups.

This assumption is reasonable if we confine comparisons of weight gains to treatment groups with mortality rates not greatly in excess of the control rate. Otherwise an average or minimum N might be used.

3. Constant variability across treatment groups.

This assumption might hold for the weights themselves or for some function of the weights such as log weights.

4. There are enough observations to have effectively an infinite number of degrees of freedom. The power obtained with 12 d.f. is nearly that obtained with infinite d.f.

#### 5. No simultaneity correction is applied.

Table XVIII.3 shows the power of a one sided normal theory test of

$$H_{o}: \mu = \mu_{o}$$
vs
$$H_{1}: \mu < \mu_{o}$$

where  $\mu$ ,  $\mu_0$  are the average weights in the treatment and control groups respectively. In the absence of tank effects the individual weights are assumed to have standard deviation  $\sigma$ . The bottom portion of the table contains factors C, for constructing 95 percent lower confidence bounds on  $\mu_0 - \mu$ . Namely  $\mu_0 - \mu \geq \bar{X}_0 - \bar{X} - C\sigma$ .

1	GAINS	BETWEI	EN CON	TROL GI	ROUP AI	ND TREA	ATMENT	GROUPS		= 0.05
~	<u> </u>	10	<u>15</u>	<u>25</u>	<u>35</u>	<u>50</u>	75	100	150	200
$d = (\mu_o - \mu)/\sigma$	.02	0.05	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.07
	.04	0.06	0.06	0.07	0.07	0.07	0.08	0.09	0.10	0.11
	.06	0.07	0.07	0.08	0.08	0.09	0.10	0.11	0.13	0.15
	.08	0.07	0.08	0.09	0.10	0.11	0.12	0.14	0.17	0.20
	•1	0.08	0.09	0.10	0.11	0.13	0.15	0.17	0.22	0.26
	.2	0.12	0.14	0.18	0.21	0.26	0.33	0.41	0.54	0.64
	.3	0.17	0.21	0.28	0.35	0.44	0.58	0.68	0.83	0.91
	.4	0.23	0.29	0.41	0.51	0.64	0.79	0.88	0.97	0.99
	.5	0.30	0.39	0.55	0.67	0.80	0.92	0.97	0.996	1.000
	.6	0.38	0.50	0.68	0.83	0.91	0.98	0.995	1.000	
	.7	0.47	0.61	0.80	0.90	0.97	0.996	1.000		
	.8	0.56	0.71	0.88	0.96	0.991	0.999			
	.9	0.64	0.79	0.94	0.98	0.998	1.000			
	1	0.72	0.86	0.97	0.995	1.000				
	1.2	0.85	0.95	0.995	1.000					
	1.4	0.93	0.99	1.000						
	1.6	0.97	0.997							
	1.8	0.991	1.000							
	2	0.998								
	2.2	1.000								
	С	.74	.60	.47	. 39	. 33	.27	.23	.19	.16

TABLE XVIII.3 POWER OF ONE SIDED PAIRWISE COMPARISONS OF AVERAGE WEIGHT

وتعتقدت

The power calculations and precision factors in Table XVIII.3 need to be adjusted for tank to tank variation within groups. Suppose that there are J tanks per group, n organisms per tank, and that  $\sigma_{\tau}^2$ ,  $\sigma_{e}^2$  represent the between and within tank components of variation. The variance of the average weight in the group is then

$$\sigma_{\tau}^{2} + \frac{\sigma_{e}^{2}}{N} = \frac{\sigma_{e}^{2} + n\sigma_{\tau}^{2}}{N}$$

Table XVIII.3 is entered at N and at d =  $(\mu_o - \mu)/(\sigma_e^2 + n\sigma_\tau^2)^{1/2}$ 

The precision factors at the bottom of the table remain the same but the confidence bound is  $\mu_0 - \mu \ge \bar{X}_0 - X - C(\sigma_e^2 + n\sigma_\tau^2)^{1/2}$ 

We now apply these relations to the weight gain data from the Holcombe and Phipps test on compound D. In that example there is no statistical evidence of tank to tank variation within groups. Namely

$$\hat{\sigma}_{e}^{2} = 900.354$$
 with 366 d.f.  
 $\hat{\sigma}_{e}^{2} + 19.191\hat{\sigma}_{\tau}^{2} = 754.97$  with 15 d.f.

Thus

 $\hat{\sigma}_{\tau}^2 = -7.58$ 

and we assume it is 0. Since  $\sigma_e^2$  is estimated with 366 degrees of freedom, we assume it is known exactly. Thus  $\hat{\sigma}_e = 30.0$ . The average N in groups 1-4 is 91.24 while the sample size in group 5 is 21. Assume for the purpose of power calculations that  $\mu_0 = \bar{X}_0$ ,  $\mu = \bar{X}$ . Thus  $\mu_0 = 131.245$ ,  $\mu_3 = 127.924$ ,  $\mu_4 = 113.080$ ,  $\mu_5 = 108.095$ .

Therefore,

$$d_{3} = \frac{\mu_{0} - \mu_{3}}{\sigma} = \frac{131.245 - 127.924}{30.0} = 0.111$$
$$d_{4} = \frac{\mu_{0} - \mu_{4}}{\sigma} = \frac{131.245 - 113.080}{30} = 0.606$$
$$d_{5} = \frac{\mu_{0} - \mu_{5}}{\sigma} = \frac{131.245 - 108.095}{30} = 0.772$$

Interpolating (approximately) in Table XVIII.3 with N = 90 and  $d_3$ ,  $d_4$ ,  $d_5$  yields

Control vs	group	3	Power =	0.16
Control vs	group	4	Power =	0.99

\_

For group 5 the assumption of equal N is not reasonable and so we calculate the noncentrality parameter for the test as  $d_5/(1/N_0 + 1/N_5)^{1/2} =$ 3.186. Thus power =  $\Phi$ (noncentrality - 1.645) =  $\Phi$ (3.186 - 1.645) =  $\Phi$ (1.540) = 0.94.

#### E. Unequal Allocations of Testing Effort Among Treatment Groups

Standard test guidelines call for equal numbers of tanks to be run at each treatment group. Such a design would be sensible only if prior to running the test there was total ignorance about response levels to be expected. That is suppose it was thought a priori that the mortality rate for each treatment group could be anywhere between 0 and 100 percent. Then it would make good sense to allocate experimental effort equally among treatment groups to assure specified power whenever and wherever the mortality rate exceeds that in the control group by a specified amount. However if on the basis of either a priori scientific information or previous testing some information was available concerning mortality rates to be expected at the various treatment groups, then unequal allocation of experimental effort would be preferable. In particular at the higher treatment groups, where mortality would be expected to be substantially higher than the control rate, it is easy to detect differences from the control. Thus the experimental effort should be decreased at these groups. At the lower experimental groups, where it is more difficult to detect differences from the control group, the experimental effort should be increased to improve sensitivity. Thus the degree of experimental effort should in general decrease as the toxicant level increases.

Details of a procedure for arriving at an unequal allocation will be discussed in the report on phase 2, for Daphnia magna. For the purpose of this subsection, consider the following illustration of the effects on sensitivity of unequal replication. Suppose that the experiment is to consist of a control group and I = 5 treatment groups. Suppose that cost and logistical restraints limit the number of tanks to 24, that n = 25fry will be exposed in each tank and that tank to tank heterogeneity is such that the variance inflation factor is 1.5. Suppose it is felt that the control group response rate will be about 0.05 and the mortality rates in the treatment groups will be about 0.10, 0.15, 0.20, 0.40, and 0.80 respectively. The classical allocation would be to run J = 4 tanks per group. Thus N = 100 fish per group. The effective sample sizes would be  $N_{eff} = 100/1.5 = 66.67$  fish per group. Suppose further that it is considered important (biologically and/or legally) to detect increases in mortality of 10 percent above the control rate. That is, we wish to detect differences between 5 percent and 15 percent mortality.

Under the classical allocation scheme the power to be expected for each treatment group-control group comparison would be:

Group 2 vs control

$$1 - \beta = \Phi[(2 \text{ arc } \sin\sqrt{.10} - 2 \text{ arc } \sin\sqrt{.05})/(2/66.67)^{1/2} - 1.645] = \Phi[(.64 - .45)/(2/66.67)^{1/2} - 1.645] = \Phi(-0.53) = 0.30$$

Group 3 vs control

$$1 - \beta = \Phi[(.80 - .45)/(2/66.67)^{1/2} - 1.645] = \Phi(0.38) = 0.65$$

Group 4 vs control

$$1 - \beta = \Phi[(.93 - .45)/(2/66.67)^{1/2} - 1.645] = \Phi(1.13) = 0.87$$

Group 5 vs control

$$1 - \beta = \Phi[(1.37 - .45)/(2/66.67)^{1/2} - 1.645] = \Phi(3.67) = 1.000$$

Group 6 vs control

$$1 - \beta = \Phi[(2.21 - .45)/(2/66.67)^{1/2} - 1.645] = \Phi(8.52) = 1.000$$

Consider the modified e ocation of 7, 6, 6, 3, 1, 1 tanks in the control group and in each of the treatment groups respectively. Then  $N_0 = 175$ ,  $N_2 = 150$ ,  $N_3 = 150$ ,  $N_4 = 75$ ,  $N_5 = N_6 = 25$ . The effective sample sizes,  $N_{eff} = N/1.5$ , are then 116.67, 100, 100, 50, 16.67, 16.67. The power to be expected for each treatment group-control group comparison would then be:

Group 2 vs control

$$1 - \beta = \Phi[(2 \operatorname{arc} \sin \sqrt{.10} - 2 \operatorname{arc} \sin \sqrt{.05})/(1/116.67 + 1/100)^{1/2} - 1.645] =$$

$$\Phi[(.64 - .45)/0.136 - 1.645] = \Phi(-0.25) = 0.40$$

Group 3 vs control

 $1 - \beta = \Phi[(.80 - .45)/(1/116.67 + 1/100)^{1/2} - 1.645] = \Phi(0.923) = 0.82$ 

Group 4 vs control

$$1 - \beta = \Phi[(.93 - .45)/(1/116.67 + 1/50)^{1/2} - 1.645] = \Phi(1.195) = 0.88$$

Group 5 vs control

$$1 - \beta = \Phi[(1.37 - 0.45)/(1/116.67 + 1/16.67)^{1/2} - 1.645] = \Phi(1.87) = 0.97$$

Group 6 vs control

$$1 - \beta = \Phi[(2.21 - 0.45)/(1/116.67 + 1/16.67)^{1/2} - 1.645] = \Phi(5.08) = 1.00$$

Comparison of the two sets of calculations shows that we have improved the power of the comparisons at the low concentration end of the test without sacrificing any appreciable power at the high concentration end of the test. This has been done without increasing the size of the test. If we were willing to add several additional tanks we could do even better. The power for the comparison of the mortality rate 0.15 vs the control rate has been increased from 0.65 to 0.82. This is a substantial improvement because the chance of not detecting such an increase in mortality has diminished from 1 in 2.86 to 1 in 5.56. This just about halves the type 2 error probability.

The previously suggested unequal allocation of experimental effort to concentration groups is intuitively sensible, improves sensitivity at the low end of the experiment where it is most needed, does not diminish sensitivity at the high end of the experiment, and does not increase the overall size of the test. It requires specifying prior beliefs about the response levels to be expected at the various concentration levels. A scheme for doing this will be discussed in the phase 2 report. However, it is pretty clear that the more definitive the prior information, the more unequal should the allocation be. As prior information diminishes to complete ignorance the design should approach equal allocation. Of course the efficacy of the design depends on the accuracy of the prior information. If it is believed a priori that group 6 will have between a 75 percent and 100 percent mortality rate but it in fact has a 15 percent mortality rate, then an unequal tank allocation scheme will probably do worse than an equal tank allocation scheme.

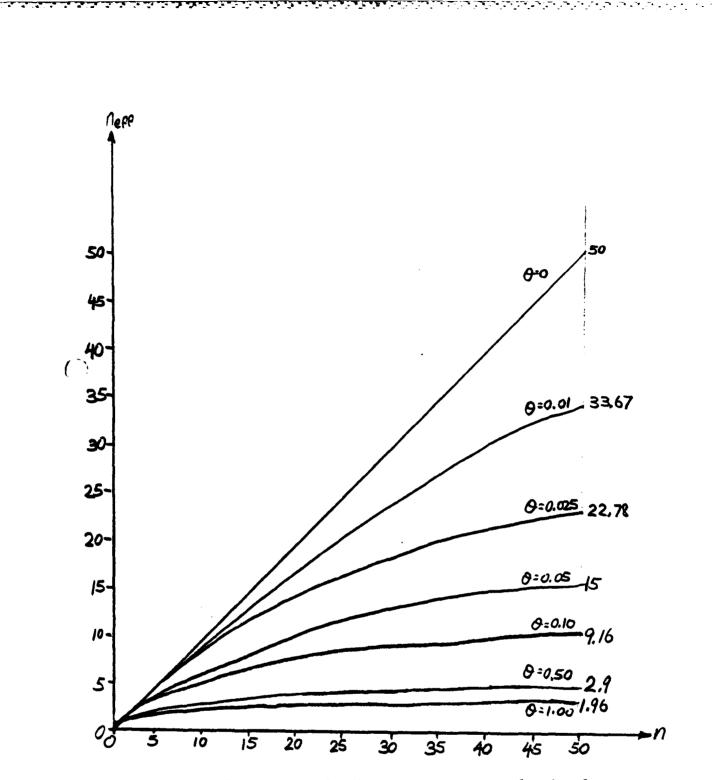


Figure XVIII.1 Effective sample size per group vs sample size for various values of  $\boldsymbol{\theta}$ 

•

## XIX <u>DESIGN AND ANALYSIS CONSIDERATIONS FOR FULL LIFE CYCLE TESTS</u> DISTINCT FROM THOSE FOR EARLY LIFE STAGE TESTS

In recent years, efforts in aquatic toxicity testing have been shifting more and more to early life stage tests and away from full life cycle tests. The results obtained from a full life cycle test are directly analagous to those obtained from early life stage tests, only much more of them are accumulated. Mortality rates are recorded periodically, length and weight measurements are obtained periodically, and fecundity responses such as embryos per spawn, total numbers of embryos, spawns per female are recorded. With respect to the survival, weight, and length responses the design and analysis considerations discussed in the previous sections are directly applicable and require no amplifications or modifications.

Some differences in design and analysis considerations may be called for with respect to the fecundity responses. Questions of homogeneity of variances, tank to tank heterogeneity within groups, form of dose response relation are handled in essentially the same manner as the analagous questions for the mortality and weight responses. Similarly, sample size determination and tank allocation design calculation need to be made in the same manner as those carried out for mortality and for weight responses. The statistical issues are the same, but the numbers may turn out to be different.

One design consideration associated with fecundity responses may well introduce important differences as compared with those for mortality and weight considerations. Namely it has been assumed that the cost structure is such that there is a certain incremental cost associated with adding an additional tank to the test but once the tank is added, the fish are free. This leads to the recommendation to run as many tanks as can be afforded and fill each tank with the maximum number of embryos or fry that is biologically sensible. This cost structure may not hold for fecundity respon-There is a considerable amount of operational and clerical effort ses. associated with accumulating the hatched embryos, counting them. associating them with the appropriate fish or groups of fish, and properly recording the data. The numbers of embryos produced are related to the numbers of fish rather than to the numbers of tanks. Thus fish can no longer be considered free. Effects of competition on production must also be considered. In the presence of tank effects it may thus be sensible to increase the number of tanks in the test and decrease the number of fish per tank. This may improve the precision of statistical inferences without incurring additional expense. The particular trade off between number of tanks and number of fish per tank would of course depend on the extent of tank to tank heterogeneity and on the tank and fish costs.

All of these issues arise in the design and analysis of toxicity tests on <u>Daphnia magna</u>. Survival, length, and fecundity responses are reported at periodic intervals. The fecundity responses reported by various investigators include average embryos per surviving female per chamber, embryos per surviving female, total number of embryos depending on the design of the test. Questions of multiple daphnids or individual daphnids per beaker are commonly posed. Thus most of the design and analysis considerations involved in full life cycle tests that are distinct from those in the early life stage tests will be addressed in the course of the discussion of the design and analysis of the Daphnia toxicity tests.

#### REFERENCES

- Manual for Construction and Operation of Toxicity Testing Proportional Diluters. USEPA Environmental Research Laboratory- Duluth, Minnesota, July, 1978 (EPA 600/3-78-072).
- [2] "Recommended Bioassay Procedure for Fathead Minnow <u>Pimephales Promelas</u> <u>Rafinesque</u> Chronic Tests," "SEPA Environmental Research Laboratory, Duluth, Minnesota, January, 1972.
- [3] DeFoe, D. L., "Multichannel Toxicant Injection System For Flowthrough Bioassays," <u>Journal Fisheries Research Board of Canada</u>, Vol. 32, 1975, pp. 544-546.
- [4] Federal Register, Vol. 44, No. 145, July 24, 1979, pp. 44054-44093.
- [5] Fishbeck, R. D., American Laboratory, September, 1980, pp. 125-130.
- [6] F. McKim, J. M., "Evaluation of Tests With Early Life Stages of Fish for Predicting Long Term Toxicity, <u>Journal Fisheries Research</u> <u>Board of Canada</u>, Vol. 34, 1977, pp. 1148-1154.
- [7] Federal Register, Vol. 45, No. 98, May 19, 1980, pp. 33124-127.
- [8] Stephan, C. et al, "Guidelines for Conducting Flow-Through Early Life Stage Toxicity Tests with Fathead Minnows for use in the USEPA, OTS-ORD Round Robin Test," USEPA Environmental Research Laboratory, Duluth, Minnesota, November, 1979.
- [9] Stephan, C. et al, "Early Life Stage Toxicity Tests with Fathead Minnows," USEPA Environmental Research Laboratory, Duluth, Minnesota memorandum, October, 1979.
- [10] "Standard Practice for Conducting Toxicity Tests on the Early Life Stages of Fishes," Working Draft of ASTM Committee E-35 on Pesticides, September, 1979.
- [11] Finney, D. J., Probit Analysis, Third Edition, Cambridge University Press, London, 1971, pp. 175 ff., page 125.
- [12] Helwig, J. T. and Council, K. A., Editors, <u>SAS User's Guide</u>, 1979 Edition, SAS Institute, Inc. Raleigh, 1979.
- [13] Dixon, W. J. and Massey, F. J., Jr., <u>Introduction to Statistical</u> Analysis, McGraw-Hill, New York, 1969, pp. 238, 240-243.
- [14] Feder, P. I. an Villavize, Susan A., "EXAX2 -- A Computer Program to Compare Binomial Proportions, Program Description and Card Input Information," Technical Report No. 216, Department of Statistics, The Ohio State University, Columbus, Ohio, September 1980.

- [15] Agresti, A. and Wackerly, D., "Some Exact Conditional Tests of Independence for RxC Crossclassification Tables," <u>Psychometrika</u>, Vol. 42, No. 1, March, 1977, pp. 111-125.
- [16] Boulton, D. M. and Wallace, C. S., "Occupancy of a Rectangular Array," <u>Computer Journal</u>, Vol. 16, 1973, pp. 57-63.
- [17] Boulton, D. M., "Remark on Algorithm 434," <u>Communications of the ACM</u>, Vol. 17, 1974, pp. 326.
- [18] March, D. L., "Algorithm 434, Exact Probabilities for rxc Contingency Tables," Communications of ACM, Vol. 15, 1972, pp. 991-992
- [19] Littell, R. C. and Folks, J. L., "Asymptotic Optimality of Fishers Method of Combining Independent Tests," <u>Journal of the American</u> Statistical Association, Vol. 66, No. 336, December, 1971, pp. 802-806.
- [20] Littell, R. C. and Folks, J. L., "Asymptotic Optimality of Fisher's Method of Combining Independent Tests II," <u>Journal of the American</u> <u>Statistical Association</u>, Vol. 68, No. 341, March, 1973, pp. 193-194.
- [21] Williams, D. A., "The Analysis of Binary Responses from Toxicological Experiments Involving Reproduction and Teratogenicity," <u>Biometrics</u>, Vol. 31, 1975, pp. 949-952.
- [22] Kupper, L. L. and Haseman, J. K., "On the Use of a Correlated Binomial Model for the Analysis of Certain Toxicological Experiments, Biometrics, Vol. 34, No. 1, 1978, pp. 69-76.
- [23] Barnett, V. and Lewis, F., <u>Outliers in Statistical Data</u>, Wiley, New York, 1978.
- [24] Nelson, W., "Confidence Intervals for the Ratio of Two Poisson Means and Poisson Predictor Intervals," <u>General Electric Company</u> <u>Research and Development Center TIS Report 69C118</u>, February, 1969. IEEE Transactions on Reliability, 1970, Vol. R-19, pp. 42-49.
- [25] Lehmann, E. L., <u>Testing Statistical Hypotheses</u>, 1959, John Wiley and Sons, Inc., <u>Section</u> 4.5.
- [26] Lieberman, G. J. and Owen D. B., <u>Tables of the Hypergeometric</u> Probability Distribution, 1961, Stanford University Press.
- [27] Dixon, W. J. and Brown, M. B., Editors, <u>BMDP Biomedical Computer</u> <u>Programs P-Series 1979</u>, University of California Press, Berkeley, 1979.
- [28] Snedecor, G. W. and Cochran, W. G., <u>Statistical Methods</u>, <u>Sixth</u> Edition, Iowa State University Press, Ames, 1967, p. 246.

- [29] Steel, R. G. D. and Torrie, J. H., <u>Principles and Procedures of</u> Statistics - <u>A Biometrical Approach</u>, <u>Second Edition</u>, <u>McGraw-Hill</u>, New York, 1980 p. 514.
- [30] Goodman, L. A. and Kruskal, W. H., "Measures of Association for Crossclassifications," Journal of the American Statistical Association. Vol. 49, December, 1954, pp. 732-764.
- [31] Goodman, L. A. and Kruskal, W. H., "Measures of Association for Crossclassifications, IV: Simplification of Asymptotic Variances," Journal of the American Statistical Association, Vol. 67, No. 338, June, 1972, pp. 415-421.
- [32] Brown, M. B., and Benedetti, J. K., "Sampling Behavior of Tests for Correlation in Two-Way Contingency Tables, <u>Journal of the American Statistical Association</u>, Vol. 72, No. 358, June, 1977, pp. 309-315.
- [33] Proctor, C. H., "Relative Efficiencies of Some Measures of Association for Ordered Two-Way Contingency Tables Under Varying Intervalness of Measurement Errors," <u>Proceedings of the Social Statistics Section</u>, American Statistical Association, 1973, pp. 372-379.
- [34] Dunnett, C. W., "A Multiple Comparisons Procedure for Comparing Several Treatments With a Control, "Journal of the American Statistical Association, Vol. 50 pp. 1096-1121.
- [35] Dunnett, C. W., "New Tables for Multiple Comparisons With a Control," Biometrics, Vol. 20, 1964, pp. 482-491.
- [36] Williams, D. A., "A Test for Differences Between Treatment Means When Several Dose Levels are Compared With a Zero Dose Control," <u>Biometrics</u>, Vol. 27, 1971, pp. 103-117.
- [37] Williams, D. A., "The Comparison of Several Dose Levels With a Zero Dose Control," <u>Biometrics</u>, Vol. 28, 1972, pp. 519-531.
- [38] Chew, V., <u>Comparisons Among Means in an Analysis of Variance</u>, Agricultural Research Service Publication ARS/H/6, October, 1977.
- [39] Thomas, D. G., "Exact Confidence Limits for the Odds Ratio in a 2x2 Table," <u>Applied Statistics</u>, Vol. 20, 1971, pp. 105-110.
- [40] Box, G. E. P., Hunter, J. S., and Hunter, W. G., <u>Statistics for</u> <u>Experimenters. An Introduction to Design, Data Analysis, and Model</u> <u>Building</u>, Wiley, New York, 1978.
- [41] Feder, P. I. and Sherrill, M. E., "A Computer Program to Calculate Nonparametric Lower Confidence Bounds on Safe Concentration in Quantal Response Toxicity Tests," Technical Report No. 215, Department of Statistics, The Ohio State University, Columbus, Ohio, September, 1980.

[42] Stephan, C. E., "Methods for Calculating an LC<sub>50</sub>," in <u>Aquatic Toxicology and Hazard Evaluation</u>, F. L. Mayer and J. L. Hamelink, Editors, ASTM Special Technical Publication 634, American Society for Testing and Materials, 1977, pp. 65-84.

P.

Ľ

M

- [43] Jennrich, R. I. and Moore, R. H., "Maximum Likelihood Estimation by Means of Nonlinear Least Squares," in <u>Proceedings of the Statistical</u> <u>Computing Section</u>, American Statistical Association, 1975, pp. 57-65
- [44] Mandel, John, <u>The Statistical Analysis of Experimental Data</u>, John Wiley and Sons, Inc., 1964.
- [45] Graybill, F. A., <u>An Introduction to Linear Statistical Models</u>, <u>Volume I</u>, McGraw-Hill Book Company, 1961, pp. 186-194.
- [46] Clopper, C. J. and Pearson, E. S., "The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial, <u>Biometrics</u> Vol. 26, 1934, pp. 404-413.
- [47] Hollander, M. and Wolfe, D. A., <u>Nonparametric Statistical Methods</u>, John Wiley and Sons, Inc., 1973.
- [48] Natrella, M. G., <u>Experimental Statistics</u>, National Bureau of Standards Statistics Handbook No. 91, U. S. Government Printing Office, 1963.
- [49] Gross, A. J. and Clark, V. A., <u>Survival Distributions: Reliability</u> <u>Applications in the Biomedical Sciences</u>, John Wiley and Sons, Inc., 1975.
- [50] Kalbfleisch, J. D. and Prentice, R. L., <u>The Statistical Analysis of</u> Failure Time Data, John Wiley and Sons, Inc., 1980.

### APPENDIX AII **EARLY LIFE STAGE DATA SETS USED AS EXAMPLES IN THE BODY OF** THE REPORT

This appendix contains listings of data sets from four early life stage toxicity tests. These data are used for illustrations of the procedures discussed in the body of the report. These data sets are

> Benoit - compound A DeFoe - compound C Holcombe and phipps - compound D Jarvinen - compound B.

The three types of data -- survival, weight and toxicant concentration -- are represented in three "card types." The first six entries on each card are the same across card types -- treatment group (col 2), replicate designation (col 4), card type (col 6), card member (cols 7=8), investigator code (cols 9-10), test code (cols 11-12). This provides enough information to sort the cards by investigator, experiment, type, group, and sequence should the data become disarranged. Card type 1 (survival data) contains in addition number of embryos tested (cols 16-20), number hatched live (cols 21-25), number of fry tested (cols 31-35), number live at end of test (cols 36-40), number normal at end of test (cols 41-45). Card type 2 (weight data) contains number of weights recorded from that particular chamber (cols 14-15), individual weights (5 cols per weight, up to 13 weights per card). Card type 3 (toxicant concentration) contains month (cols 16-17), day (cols 18-19), year (cols 20-21), toxicant concentration (cols 32-38) -- one determination per card. At the head of each type of information several lines of descriptive text are given.

Appendix AII is the appendix for Section II.

VESTIGATOR: JARVINON (04), TEST: COMPOUND B

Jers, Mark, Mark, M

A

b

2 20401 15

2 10401 15

TA FROM EARLY LIFE STAGE TESTS WITH FATHEAD MINNOWS X LEVELS: 1(CUNTRUL),2(LOWEST),...,6(HIGHEST); 2 REPS EA (A,B) EMBRYOS TESTED+ # LIVE AFTER HATCH+ # NORMAL FRY AFTER HATCH FRY TESTED, # ALIVE AT END, # NORMAL AT END Α 1 10401 1 10401 5û Ы 1 10401 A đ 1 10401 - 15 A 1 10401 . 48 --48 1 10401 ð 1 10401 Α 1 10401 .48 4.7 . ы 1 10401 Α 1 10401 d 6 A 1 10401 a 1 10401 VESTIGATOR: JARVINON (04), TEST: COMPOUND 8 (01) f TA FROM EARLY LIFE STAGE TESTS WITH FATHEAD MINNOWS X LEVELS: 1(CONTROL)+2(LOWEST)+...+6(HIGHEST); 2 REPS EA (A+B) DIVIDUAL WEIGHTS (MG) OF ALL FISH ALIVE AT END OF TEST (1-2 CARDS/TANK) MEER OF WEIGHTS & LIST OF WEIGHTS 2 10401 16 Α 2 20401 16 А 2 10401 14 ø 2 20401 14 ы 096 - 096 2 10401 15 Δ 2 20401 15 A 2 10401 14 Ġ 2 20401 14 b C87 2 10401 15 C68 A 2 20401 15 A 2 10401 15 ....088 ь ы 2 20401 15 A 2 10401 15 

ы 2 20401 15 2 10401 08 - 060 A 2 10401 06 d 2 10401 00 Α н 2 10401 00

C77

06C

C 5 C

(01)

### INVESTIGATOR: JARVINON (04), TEST: COMPOUND 8

<b>.</b>	<b>+</b> .	<b>r</b>		· · · · · · · · · · · · · · · · · · ·	
					WITH FATHEAD MINNOWS
				NTRATIONS OF TOXIC	++++++ (A+B) +2 REPS EA (A+B)
					•
1			1040ì 20401	Ú30879 -	0.00
1				031579	0.00
1			30401	032279	0.00
1			40401	032979	0.00
1			10401	030679	0.00
1			20401	031379	0.00
1			30401	032079	0.00
1		_	40401	032779	0.00
2			10401	030879	0.22
2			20401	031579	0-24
2	A		30401	032979	0.19
2	B		10401	030679	C•29
	6		20401	031379	0.23
	В		30401	032079	0.28
2			40401	032779	0-18
3	A		10401	030879	0.34
3	A		20401	031579	0.39
د	A		30401	032279	-0-53
3	A		40401	032979	0.25
3	Б		10401	030679	0•48
3	đ		20401	031379	0.36
3	В		30401	032079	0.42
3	8		40401	032779	0.29
	A		10401	030879	0.46
4	A		20401	031579	0.55
4	A		30401	032279	0.89
4	<b>A</b>		40401	032979	.0.33
4	8		10401	030679	0.72
4	D		20401	031379	0.52
4	8		30401	032079	0.88
4	B		40401	032779	0.35
5	A		10401	030879	0.60
5	A		20401	031579	0.73
5	A		30401	032279	0.91
5	<b>A</b>		40401	032979	0.61
			10401	030679	1.10
	ь		20401	031379	C•59
5	ø		30401	032079	1.10
- 5	ø	3			Q+51
6	A		10401	030879	1.10
6	<b>A</b>		20401	031579	0.96
6 -		3	30401	032279	1.90
6	A	٤	40401	032979	1.10
Ð	b		10401	030579	1.40
6	B		20401	031379	0.69
	8		30401	032079	1.60
6	ø	٤	40401	032779	1+10

296

(01)

}

DATA COOM CADLY		OIT (01				M T 845 (							
DATA FROM EARLY SIX LEVELS: 1(CO								EA (A	• B• C•	D)			
# EMBRYOS TESTED	I∎ # ALI	VE AFTE	R HATC	Н. 💋	NORMA								
# FRY TESTED. #													
1 A 1 10101 1 m 1 10101	30 30	24 23 26 26	-	15 15	15								
1 C 1 10101		20 20		15	15								
1 D 1 10101	30	24 23		15	15								
2 A 1 10101		22 21	15	15	15								
2 В 1 10101 2 С 1 10101	30 30	21 21 21 21	15 15	14 15	14 15								•
2 0 1 10101	30	16 16		15	15								
3 A 1 10101	30	17 16	15	14	14								,
3 8 1 10101	30	20 20		14	14								`•
3 - 1 10101	-	22- 21	-	15	15								
3 D 1 10101 4 A 1 10101	30 30	22 21 12 11	15 15	15 12	15 12								1
4 8 1 10101		-18 18		11	11								
4 C 1 10101	30	22 22		14	14								•
4 D 1 10101	30	25 24	-	14	14								F
5 A 1 10101 5 5 1 10101	- 30 30	-2020 17 17		09- 06									
5 C 1 10101	30	17 16	-	05	09								۲.
5 0 1 10101	30	24 24	-	06	08								
6 A 1 10101	30	19 18		80	08								\$
6 В 1 10101 6 С 1 10101	30 30	20 19 26 25	_	03 12	03								•
6 C 1 10101	30	20 25		10	12 10								-
INVESTIGATOR: DU													1
DATA FROM EARLY										<b>_</b> .			
SIX LEVELS: 1(CC													
											<b>1 1</b>		\$
INDIVIDUAL WEIGH	ITS (MG)	OF ALL	FISH	ALIVE	AT E						ik )		١
	ITS (MG) S-6-4-15	OF ALL	FISH <del>IGMTS</del>	ALIVE	AT E						к) 079	110	-
INDIVIDUAL #EIGH NUMBER OF # <del>EIGH</del> 1 a 2 10101 14 1 a 2 20101 14	ITS (MG) S-6-LIS 164 1 139	0F ALL 1-0F-48 52 164	FISH <del>IGHTS</del> 123	ALIVE 13C	AT E	ND DF	TEST 150	(1-2 205	128	085	070		•
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 10101 -15	ITS (MG) S-6-LIS 164 1 139 -0901	) OF ALL 57-05-48 52 164	FISH <del>IGHTS</del> 123	ALIVE 13C	AT E	ND DF	TEST	(1-2	CARD	S/TAN		118 162	-
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 10101 -15 1 B 2 20101 15	ITS (MG) IS <u>6</u> LIS 164 1 139 -0901 139 1	0F ALL 57-0F-WE 52 164 25180 06	FISH IGMTS 123 129	ALIVE 13C 102	AT E 094 -165	ND DF 141 103	TEST 150 135	(1-2 205 170	128	085 085 140	070 092	162	-
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 10101 15 1 B 2 20101 15 1 C 2 10101 15 1 C 2 20101 15	ITS (MG) S-6-L-IS 164 1 139 -0901 139 1 139 1 175 1 -1521	) OF ALL 57 - OF HE 52 164 -25 - 180 06 .30 102 18	FISH IGHTS 123 129 143	ALIVE 13C 102 121	AT E 094 -165 131	ND DF 141 103 172	TEST 150 135 120	(1-2 2C5 170 150	CARD 128 180 090	085 140 133	070		•
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 10101 15 1 B 2 20101 15 1 C 2 10101 15 1 C 2 20101 15 1 D 2 10101 15	ITS (MG) S <u>6 LIS</u> 164 1 139 090 1 139 1 139 1 175 1 <u>152 1</u> 136 0	0 OF ALL 57 OF WE 52 164 25 180 06 .30 102 18	FISH IGHTS 123 129 143	ALIVE 13C 102 121	AT E 094 -165 131	ND DF 141 103 172	TEST 150 135	(1-2 205 170	128 180	085 085 140	070 092	162	•
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 10101 15 1 B 2 20101 15 1 C 2 10101 15 1 C 2 20101 15 1 D 2 10101 15 1 U 2 20101 15	ITS (MG) S 6-L-IS 164 1 139 090 -1 139 1 139 1 175 1 -152 1 136 0 152 1	) OF ALL T OF WE 52 164 25 180 06 30 102 18	FISH 104TS 123 129 143 116	ALIVE 13C 102 121 133	AT E 094 -165 131 100	ND DF 141 103 172 100	TEST 150 135 120 090	(1-2 2C5 170 150 183	CARD 128 180 090 190	085 140 133 073	079 092 125 084	162 121 090	•
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 10101 15 1 B 2 20101 15 1 C 2 10101 15 1 C 2 20101 15 1 D 2 10101 15	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	0 OF ALL 57 OF WE 52 164 25 180 06 .30 102 18	FISH 104TS 123 129 143 116	ALIVE 13C 102 121 133	AT E 094 -165 131 100	ND DF 141 103 172 100	TEST 150 135 120	(1-2 2C5 170 150 183	CARD 128 180 090	085 140 133	070 092 125	162 121	•
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 10101 15 1 C 2 10101 15 1 C 2 10101 15 1 C 2 20101 15 1 U 2 10101 15 1 U 2 20101 15 2 A 2 20101 15 2 B 2 10101 14	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	) OF ALL T OF WE 52 164 25 180 06 30 102 18 06 18 00 123 60 240-125	FISH 123 129 143 116 	ALIVE 13C 102 121 133	AT E 094 -165 131 100	ND DF 141 103 172 100	TEST 150 135 120 090	(1-2 2C5 170 150 183	CARD 128 180 090 190	085 140 133 073	079 092 125 084	162 121 090	•
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 10101 15 1 C 2 10101 15 1 C 2 10101 15 1 C 2 20101 15 1 U 2 10101 15 1 U 2 20101 15 2 A 2 20101 15 2 A 2 20101 14 2 B 2 20101 14	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	0 OF ALL 0 OF WE 52 164 25 180 06 30 102 18 00 123 60 40 125 60 45 112	FISH IUMTS 123 129 143 116 	ALIVE 13C 102 121 133 -084 083	AT E 094 -165 131 100 120- 124	ND DF 141 103 172 100 190 100	TEST 150 135 120 090 114 155	(1-2 2C5 170 150 183 120 111	2 CARD 128 180 090 190 108 107	S/TAN 086 140 133 073 107 090	070 092 125 084 158 161	162 121 090 128 142	•
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 10101 15 1 C 2 10101 15 1 C 2 10101 15 1 C 2 20101 15 1 U 2 10101 15 2 A 2 10101 15 2 A 2 20101 15 2 B 2 10101 14 2 B 2 20101 14 2 C 2 10101 15	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	) OF ALL T-OF-WE 52 164 25 180 06 30 102 18 00 123 50 240-125 160 45 112 34 114	FISH IUMTS 123 129 143 116 	ALIVE 13C 102 121 133 -084	AT E 094 -165 131 100 120- 124	ND DF 141 103 172 100 190	TEST 150 135 120 090 114	(1-2 2C5 170 150 183 120	2 CARD 128 180 090 190 108	S/TAN 086 140 133 073 107	070 092 125 084 158	162 121 090 128	•
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 10101 15 1 B 2 20101 15 1 C 2 10101 15 1 C 2 20101 15 1 U 2 20101 15 2 A 2 20101 15 2 B 2 10101 15 2 B 2 10101 14 2 B 2 20101 14 2 C 2 10101 15 2 C 2 20101 15	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	0 OF ALL 0 OF WE 52 164 25 180 06 30 102 18 00 123 60 40 125 60 45 112	FISH IWHTS 123 129 143 116 	ALIVE 13C 102 121 133 -084 083 18C	AT E 094 -165 131 100 120 -124 160	ND DF 141 103 172 100 190 100 130	TEST 150 135 120 090 114 155 180	(1-2 2C5 170 150 183 120 111 132	2 CARD 128 180 090 190 108 107 190	S/TAN 085 140 133 073 107 090 084	070 092 125 084 158 161 120	162 121 090 128 142 132	•
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 10101 15 1 C 2 10101 15 1 C 2 10101 15 1 C 2 20101 15 1 U 2 10101 15 2 A 2 10101 15 2 A 2 20101 15 2 B 2 10101 14 2 B 2 20101 14 2 C 2 10101 15	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	OF       ALL         IT       OF       HE         52       164         -25       180         06       30       102         18	FISH IWHTS 123 129 143 116 	ALIVE 13C 102 121 133 -084 083 18C	AT E 094 -165 131 100 120 -124 160	ND DF 141 103 172 100 190 100 130	TEST 150 135 120 090 114 155 180 092	(1-2 2C5 170 150 183 120 111	2 CARD 128 180 090 190 108 107	S/TAN 086 140 133 073 107 090	070 092 125 084 158 161	162 121 090 128 142 132 115	•
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 10101 15 1 C 2 10101 15 1 C 2 10101 15 1 C 2 10101 15 1 U 2 10101 15 2 A 2 20101 15 2 A 2 20101 15 2 B 2 10101 14 2 B 2 20101 15 2 C 2 10101 15 2 C 2 10101 15 2 D 2 20101 15 3 A 2 10101 14	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	OF ALL         IT OF WE         52       164         25       180         06         30       102         18       123         60       125         60       125         40       125         60       112         34       114         09       152         37       59       094	FISH IWHTS 123 129 143 116 082 180 118 180 148	ALIVE 13C 102 121 133 084 083 18C 173 08C	AT E 094 -165 131 100 120 - 124 160 -138 - 151	ND DF 141 103 172 100 190 100 130 130 130 121	TEST 150 135 120 090 114 155 180 092 096	(1-2 2C5 170 150 183 120 111 132	2 CARD 128 180 090 190 108 107 190	S/TAN 085 140 133 073 107 090 084	070 092 125 084 158 161 120	162 121 090 128 142 132	•
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 20101 15 1 C 2 10101 15 1 C 2 10101 15 1 C 2 20101 15 1 U 2 10101 15 2 A 2 20101 15 2 A 2 20101 15 2 B 2 10101 14 2 B 2 20101 15 2 C 2 10101 15 2 C 2 20101 15 2 D 2 20101 15 3 A 2 20101 14 - 3 A 2 20101 14	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	OF       ALL         IT       OF       HE         S2       164         25       180         06       102         18       102         18       123         60       125         60       125         60       125         34       114         09       152         37       59       094	FISH IWHTS 123 129 143 116 082 180 118 180 148	ALIVE 13C 102 121 133 084 083 18C 173 08C	AT E 094 -165 131 100 120 - 124 160 - 138 - 151	ND DF 141 103 172 100 190 100 130 130 130 121	TEST 150 135 120 090 114 155 180 092 096	<pre>(1-2 2C5 170 150 183 120 111 132 133 144</pre>	2 CARD 128 180 090 190 108 107 190 130 132	S/TAN 085 140 133 073 107 090 084 085 100	079 092 125 084 158 161 120 140 112	162 121 090 128 142 132 115 095	• • •
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 20101 15 1 C 2 10101 15 1 C 2 10101 15 1 C 2 20101 15 1 U 2 20101 15 2 A 2 20101 15 2 A 2 20101 15 2 B 2 10101 15 2 B 2 20101 15 2 C 2 20101 15 2 C 2 20101 15 2 D 2 20101 15 3 A 2 20101 14 3 B 2 10101 14 3 B 2 10101 14	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	OF       ALL         IT       OF       HE         S2       164         25       180         06       102         18       102         18       123         60       125         60       125         60       125         34       114         09       152         37       59       094	FISH IWHTS 123 129 143 116 082 180 118 180 148	ALIVE 13C 102 121 133 084 083 18C 173 08C	AT E 094 -165 131 100 120 - 124 160 - 138 - 151	ND DF 141 103 172 100 190 100 130 130 130 121	TEST 150 135 120 090 114 155 180 092 096	(1-2 2C5 170 150 183 120 111 132 133	2 CARD 128 180 090 190 108 107 190 130	085 140 133 073 107 090 084 085	079 092 125 084 158 161 120 140	162 121 090 128 142 132 115	•
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 20101 15 1 C 2 10101 15 1 C 2 10101 15 1 C 2 20101 15 1 U 2 10101 15 2 A 2 20101 15 2 A 2 20101 15 2 B 2 10101 14 2 B 2 20101 15 2 C 2 10101 15 2 C 2 20101 15 2 D 2 20101 15 3 A 2 20101 14 - 3 A 2 20101 14	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	OF         ALL           IT         OF         WE           S2         164           25         180           06         30         102           18         00         123           00         123         60           440         125           160         45         112           34         114         09           W2         152         37           59         094         62           62         070         62	FISH IWHTS 123 129 143 116 082 180 118 180 148 141	ALIVE 13C 102 121 133 084 083 18C 173 08C 134	AT E 094 -165 131 100 120 -124 160 -138 -151 090	ND DF 141 103 172 100 190 100 130 130 130 121 124	TEST 150 135 120 090 114 155 180 092 096 248	(1-2 2C5 170 150 183 120 111 132 133 144	2 CARD 128 180 090 190 108 107 190 130 132	S/TAN 085 140 133 073 107 090 084 085 100	079 092 125 084 158 161 120 140 112	162 121 090 128 142 132 115 095	• • •
INDIVIDUAL WEIGH NUMBER OF HEIGHT 1 A 2 10101 14 1 B 2 20101 14 1 B 2 20101 15 1 C 2 10101 15 1 C 2 10101 15 1 C 2 10101 15 1 U 2 10101 15 1 U 2 20101 15 2 A 2 20101 15 2 B 2 10101 15 2 B 2 20101 15 2 C 2 10101 15 2 C 2 10101 15 2 C 2 20101 15 2 D 2 20101 15 3 A 2 10101 14 3 B 2 20101 14 3 B 2 20101 14 3 C 2 10101 15 3 C 2 20101 15	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	OF       ALL         IT       OF       WE         52       164         25       180         06       30       102         18       00       123         080       123       60         400       125       160         440       125       112         34       114       09         182       152       37         59       094       62       070         73       240       18       18	FISH IWHTS 123 129 143 116 	ALIVE 13C 102 121 133 -084 083 18C -173 08C 134 -13C	AT E 094 -165 131 100 120 - 124 160 - 138 - 151 090 110	ND DF 141 103 172 100 190 100 130 130 130 121 124 090	TEST 150 135 120 090 114 155 180 092 096 248 084	(1-2 2C5 170 150 183 120 111 132 133 144 182 080	2 CARD 128 180 090 190 108 107 190 130 132 138 193	S/TAN 085 140 133 073 107 090 084 085 100 094 123	079 092 125 084 158 161 120 140 112 098 086	162 121 090 128 142 132 115 095 131 150	• • •
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 20101 15 1 C 2 10101 15 1 C 2 10101 15 1 C 2 10101 15 1 C 2 20101 15 2 A 2 20101 15 2 A 2 20101 15 2 A 2 20101 15 2 B 2 10101 14 2 B 2 20101 15 2 C 2 10101 15 2 D 2 20101 15 2 D 2 20101 15 3 A 2 10101 14 3 B 2 10101 14 3 B 2 20101 14 3 C 2 10101 15 3 C 2 20101 15 3 C 2 20101 15 3 C 2 20101 15 3 C 2 20101 14 3 C 2 10101 14 3 C 2 10101 15 3 C 2 1010	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	OF       ALL         IT       OF       HE         S2       164         25       180         06       30       102         18	FISH IWHTS 123 129 143 116 082 180 118 180 148 141	ALIVE 13C 102 121 133 084 083 18C 173 08C 134	AT E 094 -165 131 100 120 -124 160 -138 -151 090	ND DF 141 103 172 100 190 100 130 130 130 121 124	TEST 150 135 120 090 114 155 180 092 096 248	(1-2 2C5 170 150 183 120 111 132 133 144 182	2 CARD 128 180 090 190 108 107 190 130 132 138	S/TAN 085 140 133 073 107 090 084 085 100 094	079 092 125 084 158 161 120 140 112 098	162 121 090 128 142 132 115 095 131	• • •
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 20101 15 1 C 2 10101 15 1 C 2 10101 15 1 C 2 10101 15 1 C 2 20101 15 2 A 2 20101 15 2 A 2 20101 15 2 A 2 20101 15 2 B 2 10101 14 2 B 2 20101 15 2 C 2 10101 15 2 C 2 10101 15 3 A 2 10101 14 3 B 2 10101 14 3 B 2 10101 14 3 C 2 10101 15 3 C 2 20101 15 3 D 2 10101 15 3 C 2 20101 15 3 D 2 10101 15 3 D 2 1010	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	OF       ALL         IT       OF       HE         S2       164         25       180         06       102         18       123         60       125         400       125         60       125         34       114         09       152         37       59       094         .62       070         173       240         18       200         73       240	FISH 10415 123 129 143 116 082 180 118 180 148 141 - 119 078	ALIVE 13C 102 121 133 -084 -083 18C -173 -08C 134 -13C -093	AT E 094 -165 131 100 120 - 124 160 - 138 - 151 090 110 147	ND DF 141 103 172 100 190 100 130 130 130 121 124 090 115	TEST 150 135 120 090 114 155 180 092 096 248 084 122	<pre>(1-2 2C5 170 150 183 120 111 132 133 144 182 080 101</pre>	2 CARD 128 180 090 190 108 107 190 130 132 138 193 126	S/TAN 085 140 133 073 107 090 084 085 100 094 123 092	070 092 125 084 158 161 120 140 112 098 086 153	162 121 090 128 142 132 115 095 131 150	
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 20101 15 1 C 2 10101 15 1 C 2 10101 15 1 C 2 10101 15 1 C 2 20101 15 2 A 2 20101 15 2 A 2 20101 15 2 A 2 20101 15 2 B 2 10101 14 2 B 2 20101 15 2 C 2 10101 15 2 D 2 20101 15 2 D 2 20101 15 3 A 2 10101 14 3 B 2 10101 14 3 B 2 20101 14 3 C 2 10101 15 3 C 2 20101 15 3 C 2 20101 15 3 C 2 20101 15 3 C 2 20101 14 3 C 2 10101 14 3 C 2 10101 15 3 C 2 1010	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	OF       ALL         IT       OF       HE         S2       164         25       180         06       30       102         18	FISH IWHTS 123 129 143 116 082 180 118 180 148 141 - 119 078 132	ALIVE 13C 102 121 133 -084 083 18C -173 08C 134 -13C	AT E 094 -165 131 100 120 - 124 160 - 138 - 151 090 110	ND DF 141 103 172 100 190 100 130 130 130 121 124 090	TEST 150 135 120 090 114 155 180 092 096 248 084	(1-2 2C5 170 150 183 120 111 132 133 144 182 080	2 CARD 128 180 090 190 108 107 190 130 132 138 193	S/TAN 085 140 133 073 107 090 084 085 100 094 123	079 092 125 084 158 161 120 140 112 098 086	162 121 090 128 142 132 115 095 131 150	• • •
INDIVIDUAL WEIGH NUMBER OF WEIGHT 1 A 2 10101 14 1 A 2 20101 14 1 B 2 20101 15 1 C 2 10101 15 1 C 2 10101 15 1 C 2 10101 15 1 C 2 20101 15 2 A 2 20101 15 2 A 2 20101 15 2 A 2 20101 15 2 B 2 20101 15 2 C 2 10101 15 2 C 2 20101 15 2 D 2 20101 15 3 A 2 20101 15 3 A 2 20101 14 3 B 2 20101 14 3 B 2 20101 14 3 B 2 20101 14 3 C 2 10101 14 3 D 2 20101 15 3 C 2 20101 15 3 D 2 20101 15 3 D 2 10101 15 3 D 2 1010	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	OF       ALL         IT       OF       WE         S2       164         25       180         06       30       102         18	FISH 10415 123 129 143 116 082 180 118 180 148 141 - 119 078 132 060	ALIVE 13C 102 121 133 084 083 18C 173 08C 134 13C 093 092	AT E 094 -165 131 100 120 -124 160 -138 -151 090 110 147 132	ND DF 141 103 172 100 190 100 130 130 130 121 124 090 115 112	TEST 150 135 120 090 114 155 180 092 096 248 084 122 099	(1-2 2C5 170 150 183 120 111 132 133 144 182 080 101 106	2 CARD 128 180 090 190 108 107 190 130 132 138 193 126 140	S/TAN 085 140 133 073 107 090 084 088 100 094 123 092 125	070 092 125 084 158 161 120 140 112 098 086 153	162 121 090 128 142 132 115 095 131 150	

4 V 2 10101	14 121	140	250	132	132	114	123	105	110	154	120	131	072
4 D 2 20101						-				-			-
5 A 2 10101		111	100	064	07°C	111	090	050	150				
5 8 2 10101		117	127	098	121	157							
5 C 2 10101		055	184	144	125	084	112	071	106				
5 0 2 10101			071	166	110	078	137	078	-				
6 A 2 10101	• • • • •	-	025	010	021	014	051	018					
6 8 2 10101			007										
6 C 2 10101			051	031	035	049	033	033	052	030	034	052	
ο U 2 10101			048	013		082	037	024	042	011	•••		
INVESTIGATOR:									•	••••			
DATA FROM EAR							MINNO	) w S					
SIX LEVELS: 1										• 8• C •	DI		
SEE DATA													
MEASURED CONC	ENTRATIO	NS OF	TOXIC	ANT	(MICR	0 6/L	)						
1 4 3 10101	- 061179			0.0									
1 A 3 10101 1 A 3 20101	062579	1		0.0	8								
1 A 3 30101	070979	t i		C+1	5								
1 8 3 10101				0.0	0								
1 В 3 20101	061879			0.0	8								
1 8 3 30101	070770			0.1									
1 6 3 10101 -	010219	-··		0.0	7								
I C 3 20101	062179			0.0	6								
1 C 3 30101	070579			C+1	3								
1 0 3 10101	061479			0.0	6								
1 0 3 20101	062879			0.1	0								
2 A 3 10101	061179			1.6	8								
2 A 3 20101	062579			1.5	9								
2 A 3 30101	070979			2.0	0								
2 В З 10101	060679			1.8	3								
2 8 3 20101	-061479	• • • • • • • • • • • • • • • • • • •		-1-6	9		· _ ·		-				
2 B 3 30101	061879 070279			1.2	7								
2 в 3 40101	070279			1.5									
2 6 3 10101-													
2 C 3 20101	061879 062179			1.4	0								
2 C 3 30101	062179			1.5	2								
2 6 3 40101	070579			- 2+1-	5								
2 D 3 10101	061479			1.8	4								
2 U 3 20101	062879			1.8									
3 A 3 10101	-061179					-							
3 A 3 20101	062579			2.6									
3 A 3 30101	070979			4.0									
3 d 3 10101	060679			3+2	6								
з в 3 20101	061479			3.5	8								
3 в 3 30101													
3 8 3 40101			• • •										
3 C 3 10101	060879			2.8									
3 C 3 20101	061879			2.3									
	062179			- 2+7									
3 C 3 40101	070579 061479			3.6									
3 0 3 10101	061479			5.7	4								
3 0 3 20101	062879			5 +4	4								
4 A 3 10101	060879			5.9									
4 A 3 20101	061179			6.3									
4 A 3 30101	-062579			7.3	7 <u>.</u>								
4 A 5 40101	070979			7.0									
4 ± 3 10101	060679												
4 8 3 20101	060679			6+8 7+1									
4 B 3 30101	061179			5.9									
4 5 3 40101 4 5 3 50101	061879												
4 6 3 50101 4 6 3 60101	070279			6.0									
4 0 3 00101	010219			6•4	7								

.

V

 .....

4 B 3 70101	076579	7.92
4 8 3 80101	070979	7.80
4 C 3 10101	060879	5.66
4 C 3 20101	061179	6.21 '
4 C 3 30101	062179	5.23
4 L 3 40101	062179	5.14
4 L 3 50101	062579	5.44
4 C 3 60101	070279	5.95
4 C 3 70101	070579	7.02
4 C 3 30101	û7û979	7.50
4 U 3 10101	061179	6.72
4 0 3 20101	061479	7.57
	- 262579	5.93
4 Ŭ 3 40101	062879	7.76
4 D 3 50101	670979	8.40
		-13+40
5 A 3 20101	062579	12.00
5 A 3 30101	070979	18.00
5 B 3 10101		17.20
5 b 3 20101	061879	10.80
5 B 3 30101	070279	11.80
5 6 3 10101		-10+60
5 C 3 20101	062179	10.70
5 C 3 30101	070579 '	13.10
5 Ú 3 10101	661479	13.60
5 D 3 20101	062879	15.30
6 A 3 10101	361179	24.10
6 A 3 20101	062579	23-30
6 A 3 30101	070979	30.00
6 H 3 10101	360679	29.80
	061879	-28-80
6 6 3 30101	070279	32.70
6 C 3 10101	060879	24.20
	- 062179	-24+20
6 C 3 30101	070579	33.40
6 D 3 10101	061479	21.20
6 D 3 20101	062879	23+30

۰.

.

DATA FROM EARLY SIX LEVELS & 11C # Embryds tested	FOE (D2), TESTI C LIFE STAGE TESTS VIT DNTRDLJ,2(LIVEST), , # LIVE AFTER MATCH ALIVE AF END, # NOPM 50 29 29 20		EPS EA (4+3)	
1 0 1 10201 2 A 1 10201 2 B 1 10201 3 A 1 10201 3 B 1 10201 4 A 1 10201 4 B 1 10201 5 A 1 10201	50       31       31       20         50       1 F       17       20         50       31       31       20         50       30       30       20         50       33       33       20         50       34       34       21         50       29       23       20         50       36       34       21         50       28       24       20	20 20 20 20 20 20 20 20 19 18 21 21 19 13 16 15		
	50 33 33 20 50 31 33 20 50 31 30 20 50 31 00 20 FOE (02), TEST: C LIFE STAGE F3STS W(T	15 15 30 00 03 00		
SIX LEVELS # 1(C) Individual Weigh	DNTROLD,2(LO4EST), TS (NG) OF ALL FISH S ~ LIST OF 4ELGHTS	.,6(HIGHEST);2 4	EPS EA (AJB)	DR LESS/CELL)
1 A 2 10201 23 1 A 2 20201 20	159 139 105 170 149 153 174 138	166 177 124 153 172 134	132 068 095	113 180 092
1 B 2 10201 20 1 B 2 20201 20 2 A 2 10201 20	011 147 172 173 190 155 153 089 124 143 312 109	174 151 200 230 135 071 141 142 101	100 124 238 124 162 124	173 145 170 153 187 128
2 A 2 20201 20 2 B 2 10201 20	191 199 125 103 091 192 195 214	103 192 080 193 187 095	220 116 164	162 200 170
2 8 2 20201 20 3 A 2 10201 20 3 A 2 20201 20	161 206 200 190 171 148 075 177 131 159 144 152	197 137 101 096 160 129 112 177 198	157 179 153	102 145 137
3 8 2 10201 18 3 8 2 20201 18	170 153 151 165 100 076 129 155	131 134 149 047	152 162 137	120 129 131
4 A 2 10201 21 4 A 2 20201 21 4 B 2 10201 19	097 133 099 122 193 144 145 145 115 111 070 097	197 162 089 137 225 150 133 084 132	135 165 144 045 114 <sup>3</sup> 064 102	122 152 073 090 091 131
4 B 2 20201 19 5 A 2 10201 16	177         053         124         090           037         056         072         047	109 119 056 053 037	050 044 007	023 047 049
5 A 2 20201 16 5 B 2 10201 15 5 B 2 20201 15 6 A 2 10201 00 6 B 2 10201 00	037 004 057 029 050 039 055 045 038	035 033 057	035 056 045	050 06 <del>3</del> 0J3
DATA FROM EARLY SIX LEVELS = 11C Measured concent	FOE (O2), TESTI C LIFE STAGE TESTS WIT INTROL),2(LD4EST), RATIONS OF TOXICANT 20579 000.0	., Staighest); 2 # (Note no units o	EP3 34 (4+3)	
1 A 3 30201 0 1 A 3 40201 0 1 A 3 50201 0 1 A 3 50201 0 1 A 3 60201 0	93779     300.3       91373     000.0       91273     000.1       91473     000.3       91973     000.3	74 16 68 23		
1 A 3 80201 19 1 B 3 10201 09	00279 000.0 90679 000.0 91179 000.0	3 3 7 8		

€.

••••

1 8 3 30201	091379	000.045
1 8 3 40201	091779	000.018
1 8 3 50201	072077	000.042
1 🕴 3 60201	072777	000.026
1 R 3 70201	100177	000.024
1 8 3 80201	100379	000.033
2 A 3 10201	070579	001.75
2 A 3 20201	090774	1.02
2 A 3 30201	071077	2.54
2 A 3 40201	091272	2.24
2 A 3 50201	091479	1.34
2 A 3 60201	091977	1.03
2 A 3 70201	092979	1.96
2 A 3 80201	100277	2.02
2 B 3 10201	090679	1.78
2 B 3 20201	091179	2.34
2 8 3 30201	091377	2,15
2 8 3 40201	091779	1.74
2 B 3 50201	072077	1.99
2 B 3 60201	072777	1.54
2 8 3 70201	100179	2.01
2 R 3 90501	100377	2.01
3 A 3 10201	070579	7.47
3 A 3 20201	070779	5.66
3 A 3 30201	071077	4.91
3 A 3 40201	091277	4.28
3 A 3 50201	091479	6.15
3 A 3 60201	091979	6.55
3 A 3 70201	092819	4.73
3 A 3 80201	100279	6.03
3 8 3 10201	070677	5.22
3 B 3 20201	071179	7.06
3 8 3 30201	071377	8.40
3 B 3 40201	091779	4.73
3 8 3 50201	092079	6.20
3 B 3 60201	092779	4.54
3 8 3 70201	100179	6.99
3 B 3 80201	100377	6.74
4 A 3 10201	070579	14+17
4 A 3 20201	070777	13.32
4 A 3 30201	071077	16.87
4 A 3 40201	091279	17.35
4 A 3 50201	091477	12.89
4 A 3 60201	071977	14.19
4 A 3 70201	072977	14.77
4 A 3 80201 4 8 3 10201	100279	13,33
	090679	13.55
	071179	17.12
4 P 3 30201 4 9 3 40201	091377	16.57
4 8 3 50201	091779 092079	14.76
4 B 3 60201		13.73
4 9 3 70201	092777	14.37
4 8 3 80201	100177 100377	15.10
5 A 3 10201	090579	14.14
	090779	43.71 43.15
5 A 3 20201 5 A 3 30201	091377	43.15 55.80
5 A 3 40201	071277	52.20
5 A 3 40201 5 A 3 50201	091479	42.46
5 A 3 60201	091979	46.11
5 A 3 70201	092379	48.36
5 A 3 80201	100377	51.53
5 9 3 10201	070577	48.21
-		

Ĉ.

5	A	3	20201	071177	49.00
5	P	3	30201	071377	51.5)
5		3	40201	091777	43.24
5	9	3	50201	092079	44,23
5	8	3	60201	072777	43.10
5	4	3	70201	100177	51.44
5	Ą	3	80201		51.33
6		3	10201	070577	135.4
6		3	20201	070777	117.57
6	٨	3	30201	011077	154.25
6		3	40231	091279	154.75
6	A	3	50201	071477	131.)
6	8	3	10201	090577	1 4.15
6	8	3	20201	071177	1075
6	8	3	30201	071377	153.)

٠.

1

L' CL L'LL

Ŀ

	IN	VE	51	110	A T	ŪŔ I	H:	LĊO	MBE		PHIP	, s	1251	↓ ŤE	šti o	(0))						
	DA	T A	. F		)M	EAR	LY	LIF	E SI	T A G	E TEC		VIT		THEAD		vs					
	51	X.	LE	VE	LŚ	1 1	100	INTRI	36.31	, Ż	11349	: \$ \$	1		6141G	HESTIT	1 B	POS -	ĒĀ	i i a i a	r.bi	
-	1	ĒM	0	140	15	TES	TEO	), ¥	LI	/ E	AFTS	<b>ч</b>	ATCH		NORNA	LFRY	AFTE	Ř 44	TCH.		.,	
		FŔ	Ŷ	ŤĒ	ST	FD,		ALIN	VE I	IT	END.		NJRM	AL A								
	1		1	1	06	03		- 59	3	33	31	۱.	25	23	23							
	1	Ą	1	1	05	03		- 5(		27	) <u>7</u> 9	۹.	25	23	23							
	1	C	1	1	04	03		5:	)	38		5	25	24	24							
	1	3	1	1	06	03		50	ò	30	30	)	25	24	Ž4	-	-					
	Z	A	1	1	06	03		50	)	31	31	L	25		22							
	2	- Ņ	1	1	65	03		50	5	36	19	5	25	24	24							
	2	C	1		05			50	)	34	33		25	23	23							
	2	D	1	1	06	01		50	1.00 1.00	žģ			25	23	23	· •						·- ·
	3		1	1	05	03		50		35			25	23	23							
	3	R	1		05			50		38	19		25	25	25							-
	3	C	Ĩ		06			50		26	24		25	20	zó							
	- j	Ō	ī		Ō6			50		38			-25-		24			<u> </u>				
	- Ă	Ň	ī		05			50		30			25	- 21	21							
	4	9	ī		06			50		29			25	23	20							
	6	c	ī		06			50		27	25		25	21	21							
	- <b>i</b>	Ď	ī	-	06			- śó		36			24	25	žŚ							
	. 5	Ă	ĩ	-	06			50		32			25	02	δŻ							
• • • • • • • •	5	8	ī		06	_		50	_	31	27		25	04	04							
	5	č	î		05			50		36	35		25	09	07							
	ś	ŏ	î		06			50		34	11		25	ČŠ								
	í	ž	÷		050			50		36	00		29		25							
	6	1	-		05			50		34	00			00	00							
	6	~	1		051								25	00	00							
			÷		050			50		25	22		- 25		00							
	0	U			ומט	15		50		29	00		25	00	00							

303

.

			-	-	-		-	STAGE			-			-					
								L) . 21											
								461 OF			ALTVE	AT 5	NO OF	1621	(2	CARDS	ં ગુરુ (	E2210	ειι
	-			-	-			LISTO	-										
-			100				21	104	114	147	119	144	136	167	110	090	118	145	10
			200				40	129	167	132	155	132	124	147	120	125			
			104				47	173	142	177	117	115	145	127	136	152	138	185	12
			200				41	116	$\frac{175}{132}$	-151	127	147	035	165	127	103		096	05
			200				55 73	127	129	155	195 198	159	134	146	152	150	045	070	07
			100				29	175	177	117	151	120	142	100	134	099	114	144	13
			206				23	108	177	135	147	105	137	121	103	124	116	[44	13
			10				44	101	135	115	117	127	150	131	170	103	120	146	13
			20				15	112	175	178	137	137	110	116	116	105	120	140	
			100				29	160	155	111	-145-	094	-110	-123-	145	147	158	145	12
-		-	20		-		45	133	115	141	163	112	149	113	384	122	154		
			10				51	145	in	132	157	101	146	136	140	131	119	175	14
			201				55	101	169	iii	150	123	123	117	103	147	••	•	•••
			10				77	511	123	193	134	124	153	165	241	110	140	151	12
2	ŋ.	S	204	03	23	i i	02	111	123	112	142	115	128	153	154	101		•	
			100				33	145	-125	175	120	177	110	135	125	074	094	111	17
3	A	Ż	201	03	23	1	53	166	107	177	145	119	145	117	137	C15			
3	Я	Ż	100	50ġ	25	1	71	189	107	197	155	137	123	154	しいえ	131	010	124	15
3	ŋ.	2	204	01	25	1	37	149	141	123	113	141	117	124	079	695	074	079	
7	0	2	10/	603	20	i 1	68	137	171	112	เวร	113	140	120	140	1°3	145	120	- 15
 3	Ç_	2	206	03	20	1	43	1 73	140	າຈາ	Jus	113	(19.9						
3	ŋ	2	106	03	-24	ા	17	1 37	799	117	117	167	151	102	130	143	124	122	11
			2.04				43	144	134	113	115	153	101	105	125	668	ંગ્ય 3		
			106				10	149	147	151	114	101	117	15?	1 3 4	142	じつマ	118	11
			204				34	103	175	125	244		032	029					
			106				90	130	110	150	141	349	216	122	195	149	100	114	14
			506				15	123	213	112	127	372	694						
			104				17	171	119	124	137	115	153	114	135	153	109	084	13
			204				36	110	101	044	174	390	075	049			1.10		
			126				30	136		- 14? - 187-	120	122	105	104	137	115	179	173	11
			205				34	043		17.7	113	382	051	זינ	367	677	643	056	
			-				17	117	136	352									
			104				18	123	100	111	135	131	147	094	394				
			105				35	129	102	097	041	034	1.47	744	104				
- -		ſ	102				• •	1 6 4	107	(147	10.4.1	034							
		2	105																
4			10																

. .

.

٠.

.

X

5

.....

STIGATORI HOLCOMIE & P	11995 (35), TESTI D (33)		
FPON FARLY LIFE STAGE	TESTS UTTH FATHEAD MINNING		
LEVELS# 1(CONTROLID 2() SUPPO CONCENTRATIONS OF	TWESTIN 6(HIGHESTIN 4 TOVICANT (MICPO G / L)	4243 24 [\$,F.C.D]	
3 10603	9 <b>.</b>		
8 3 10403 3 10403			
10403	ŷ.		
1 3 10403 022777 1 3 20403 090479	25.1 26.7		
3 30403 091079	24.5		
3 40503 391779	28.1 24.7		
1 3 50603 092479 1 3 10403 092377	<u> </u>		
3 206 3 040579	**. 3		
3 30503 991177 3 40603 891977	26 · 1 32 · 2		
3 50403 092579	30.0		
<u>3 10603 082479</u> 3 20503 090579	<u></u>		
3 20503 090579 3 36503 091277	11+4 25+9	5 A 3 36503 - 071077 5 A 3 40603 - 091779	141.
3 40503 1091979	37.7	5 A 3 56503 J72477	171.
3 50403 092579 3 16403 083077	52°d 50°Ü	5 3 3 14503 - JP2472 5 9 3 20603 - 495579	124, 120,
3 20603 090777	25.4	5 8 3 30603 091172	111.
3 30603 301373 3 40503 392077	27.5	5 R 3 4(503 091377	127,
3 50603 892779	<u>?</u> *•7 ?*•7	5 8 9 50503 092574 5 C 3 10403 092379	120.
10603 082779	45.l	5 C 3 20603 090579	117.
3 20603 096479 3 30603 091979	47. <u>)</u> 41.7	5 C 3 30403 091779 5 C 3 40603 091979	140. 122.
1 40401 101777	67.1	5 6 3 56 503 092573	
3 5CAD3 072477 3 10403 082377	44.7 43.2	5 D 3 1(603 083)77 5 D 3 20603 080772	123.
3 20403 091579	57.3	5 D 3 20603 040779 5 D 3 30403 040779	125. 190.
3 30503 JALL7A 3 40403 091974	44.7 44.7	5 0 3 46503 092077	115.
3 50503 192572	4.7.5	<u>     50 3 50603 042772     64 3 10603 082772 </u>	
3 10403 JR2977	30.0	6 A 3 20603 490477	276.
3 20603 390579 3 30603 991279	5°.1 69.3	5 4 3 30603 - 741379 5 4 3 40603 - 941777	221. 213.
3 46603 991979	39.7	6 A 3 40603 092477	217+
<u>3 50603 092573</u> 1 10403 283373	47.5 47.7	<u>6 9 3 10603 )82379</u>	
3 20563 903779	47.1	5 B 3 20503 - 393577 6 B 3 30603 - 041177	212.
3 30603 041379	45.7	5 B 3 40503 391479	224.
3 40563 C92379 3 50603 092774	17.4 46.0	6 8 3 50603 092579 6 C 3 10603 092377	175. 239.
3 10503 082777	40.7	6 C 3 20603 09067:	216.
1 3 20603 090479 1 3 30603 091079	75.5 59.2	6 C 3 3C603 091277 6 C 3 40503 C91977	······································
3 40403 991774	75.4		205.
1 × 50603 - 992479 1 3 10603 - 082879	73.7 43.1	6 D 3 10603 083079	724.
3 70403 390579	77.6	6 D 7 20603 092777 6 D 3 30603 091377	***. 731.
3 30403 991172	11.0	6 9 3 40603 092072	2.3F.
1 3 40603 - 341477 1 3 56503 - 392577	77.2 h7.2	6 D 3 50603 J92773	،۱۰
1 10403 342477	65.0		
1 3 20603 - 040677 1 3 30603 - 091277	7 . ) 7 . 7		
1 40403 091477	73.3		
3 56603 072577	73.1 75.0		
1 7 10407 - 393077 3 7 70403 - 997777	75.5		
1 3 30503 091377	74.0		
3 4(603 )92077 3 50403 092777	77 • 3 55 • 5		
1 10403 082777	11,		
1 7 20503 0 00677	143.		

F

APPENDIX AVII<sup>\*</sup> THEORETICAL BASES OF CHI SQUARE AND PROBIT BASED TESTS OF TANK TO TANK VARIATION WITHIN TREATMENT GROUPS

### A. Chi Square Test for Heterogeneity of Tanks Within Treatment Groups

Suppose that there are I treatment (or control) groups and J tanks per treatment group. Let  $p_i$  denote the probability of death within the i-th group, i = 1, 2, ..., I assuming homogeneity of tanks within groups. The heterogeneity chi square procedure tests the hypothesis

 $H_{o}: p_{1} = p_{2} = \dots = p_{I} = p.$ 

Let N<sub>ij</sub>, X<sub>ij</sub> denote the number of fish and number of dead fish respectively in the j-th tank of the i-th group. Let  $X_{i+} = \Sigma_j X_{ij}, X_{++} = \Sigma_i \Sigma_j X_{ij}, N_{i+} = \Sigma_j N_{ij}, N_{++} = \Sigma_i \Sigma_j N_{ij}, \hat{p}_{ij} = X_{ij}/N_{ij}, \hat{p}_i = X_{i+}/N_{i+}, \hat{p} = X_{++}/N_{++}.$ 

The chi test of homogeneity based on each tank separately is based on the statistic

$$\chi_{I}^{2} = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(X_{ij} - N_{ij}\hat{p})}{N - p(1 - \hat{p})} \quad \text{with IJ} - 1 \, d.f.$$

The chi square test of homogeneity based on tanks pooled within treatment groups is based on the statistic

$$\chi_{II}^{2} \equiv \sum_{i=1}^{I} \frac{(X_{i+} - N_{i+}\hat{p})^{2}}{N_{i+}\hat{p}(1 - \hat{p})} \quad \text{with } I - 1 \text{ d.f}$$

The difference of these two statistics,  $\chi$  -  $\chi$ , is

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(X_{ij} - N_{ij}\hat{p}_{i.})^2}{N_{ij}\hat{p}(1 - \hat{p})} \quad \text{with } I(J - 1) \text{ d.f.}$$

This difference is thus a test of homogeneity of response rates across tanks within treatment groups and has a nominal chi square distribution with I(J - 1) degrees of freedom if such homogeneity exists. Note however that the weights in the denominator of this "chi square" statistic have been calculated under the null hypothesis that  $p_1 = p_2 = \ldots = p_I = p$ . Thus if the response probabilities differ among treatment groups, as we have seen in the case with respect to fry mortality, then the weights are incorrect and the nominal null distribution is inappropriate. Thus this procedure leaves something to be desired.

\*Appendix AVII is the appendix for Section VII.

### B. <u>Probit Model Based Test for Heterogeneity of Tanks Within</u> Treatment Groups

Let  $X_{ij}$ ,  $N_{ij}$ ,  $X_{i+}$ ,  $N_{i+}$ ,  $\hat{p}_{ij}$ ,  $\hat{p}_i$  have the same interpretation as in subsection A. Let  $\Phi_i \equiv \Phi(\alpha + \beta C_i)$  denote the response probability in the i - th group where  $C_i$  is some function of the toxicant concentration in the i - th group,  $\Phi$  is the standard normal c.d.f. and  $\alpha$ ,  $\beta$  are model parameters to be estimated from the data. Let  $\Phi_i \equiv \Phi(\hat{\alpha} + \hat{\beta} C_i)$  denote the maximum likelihood estimate of  $\Phi_i$ .

Finney's suggestion is to compare test statistics for lack of fit to the probit model based on each tank separately and based on tanks pooled within treatment groups. The lack of fit test based on each tank separately is

$$\chi_{I}^{2} \equiv \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(X_{ij} - N_{ij}\hat{\phi}_{i})^{2}}{N_{ij}\hat{\phi}_{i}(1 - \hat{\phi}_{i})} \quad \text{with IJ - 2 d.f.}$$

The lack of fit test based on tanks pooled within treatment groups is

$$\chi_{II}^{2} = \sum_{i=1}^{I} \frac{(X_{i+} - N_{i+}\hat{\Phi})^{2}}{N_{i+}\hat{\Phi}_{i}(1 - \hat{\Phi}_{i})} \quad \text{with } I - 2 \text{ d.f.}$$

The difference is these two statistics,

$$\chi_{I}^{1} - \chi_{II}^{2} = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(X_{ij} - N_{ij}\hat{p}_{i})^{2}}{N_{ij}\hat{\Phi}_{i}(1 - \hat{\Phi}_{i})} \quad \text{with } I(J - 1)d.f.$$

This difference is again a test of homogeneity of response rates across tanks within treatment groups and has a nominal chi square distribution with I(J - 1) degrees of freedom if such homogeneity exists. Note that the weights in the denominator of the statistic are based on the fitted probit model and depend for their validity on the validity of this model. APPENDIX AVIII.1\* OUTPUT FROM EXAX2 COMPUTER PROGRAM

<sup>\*</sup> Appendix AVIII.1 is an appendix for Section V\* ...

Figure VIII.1 EXAX2 output

$ \begin{array}{cccccccccccccccccccccccccccccccccccc$		•		DEFOE		COMPOUND C	Prento mental ITY	۲۲		
Y TAPLE Y TAPLE 272,070, 50,090, 50,090, 10,010, 50,090, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00, 50,00,0,0,0,	, <b>.</b>									
$\begin{array}{rclccccccccccccccccccccccccccccccccccc$	CASE	7	i	-€0NFI+4Ω	⊁	ب ب لول الج		• •	:	
$\frac{1}{20.37205} \cdot 101 k = 0.40 = 41.000 \cdot 52040 151 k = 5.0 \\ \frac{1}{20.500} \cdot 101 k = 41.000 \cdot 52040 \cdot 510 k = 5.0 \\ \frac{1}{20.500} \cdot 101 k = 50.000 \\ \frac{1}{20.500} \cdot 50.000 = 41.000 \cdot 520 k = 41.000 \cdot 520 k = 5.0 \\ \frac{1}{20.500} \cdot 101 k = 540 = 41.000 \cdot 520 k = 41.000 \cdot 520 k = 5.0 \\ \frac{1}{20.500} \cdot 101 k = 0.54189 \cdot 50.000 \\ \frac{1}{20.54189} \cdot 50.000 \cdot 520 k = 41.000 \cdot 520 k = 5.0 \\ \frac{1}{20.54189} \cdot 50.000 \cdot 520 k = 41.000 \cdot 520 k = 5.0 \\ \frac{1}{20.54189} \cdot 50.000 \cdot 520 k = 41.000 \cdot 520 k = 5.0 \\ \frac{1}{20.54189} \cdot 50.000 \cdot 520 k = 41.000 \cdot 520 k = 5.0 \\ \frac{1}{20.54189} \cdot 50.000 \cdot 500 k = 41.000 \cdot 520 k = 5.0 \\ \frac{1}{20.54189} \cdot 50.000 \cdot 500 k = 5.000 \cdot 500 k = 5.0 \\ \frac{1}{20.54189} \cdot 50.000 \cdot 500 k = 5.000 \cdot 500 k = 5.0$		н н С О С (С			:	06.40 22.000 4.006	101AL 50.000			
UNCTES XSQUAS = $0.3720^{\circ}$ CUTOFF = 5.7 26.600 - 101AL 20.500 - 50.000 20.500 - 50.000 0.37205 0.37205 0.54189 0.37205 0.54189 107AL 0.54189 107AL 0.54189 107AL 20.54189 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 107AL 10		LILL		)     	<i>.</i> 900	TUTAL	9. AU = UA	41.309	• 53AV7 TSTAL =	
20.530 50.000 20.530 50.000 .003 . TUTAL DEAD = 41.000 .0.37205 54199 		:	י א א י	ALIV.	U NC	D S A D	XSQCRS = TYTAL	0.37206	:	<b>5.</b> 75327
0.37205 ct = At = 0.54189 0000a			ר אר	- 29.500 29.500 =	Co0.	000 +	\$0,500 50,000 D∉AD =	000 i 4	- GUAND TUTA: -	
0000	< >		CHI 520 D CHI 520 S I H 50	148.50 UCC UARED = UARED =	0 6.0.37 6.Vat	i 👘				
	>		14 R 12-)	000	000 4•00000			-		

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

				•	
. COL 2	15.60 32.000 19.000	101#L 50.000 			
THEAL ALIVE = 49.000	. TUTAL	DEAU =	51.000	SRAND TOTAL =	0v0*0v1
F XPECTEU FRENUENCT	₽ S DE AD	XSRUGS = TUTAL	17537.42	CUTCHE = A	vere.
VE = 4	25.500	53.300 53.300	51.000 ·	GRAND TOTAL =	600°001
ASYMPTUTIC CHI SUUARED USED GASERVID CHI SQUARED = 6.7627 UBSERVID SISHIFTCANGE LEVEL = 41 YY = 21LV(AI) = 9.3571 SY = 21-2LN(AI) = 9.3571	÷271 ∺Å1	1660			:
VARY = VARI-TLM(AL)) = 4,0000			, ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ;		

### Figure VIII.3 EXAX2 output

$\begin{array}{llllllllllllllllllllllllllllllllllll$	•	000-001	000005				
$r_1 = r_1 r_1$ $r_1 = r_1 r_1 r_1$ $2^{n_1} + 3^{n_2} r_1$ $r_1 + r_1 r_1$ $2^{n_1} + 3^{n_2} r_1$ $r_1 + r_1 r_1$ $r_1 - r_2 r_1$ $r_2 + r_2 r_1$ $r_1 + r_2 r_1$ $r_2 + r_2 r_1$ $r_1 + r_2 r_2$ $r_2 + r_2 r_2$ $r_2 + r_2 r_2$ $r_2 + r_2 r_2$ $r_1 + r_2 r_2$ $r_2 + r_2 r_2$ <td></td> <td>GRAND TUTAL =</td> <td></td> <td></td> <td>GRAND TOTAL =</td> <td></td> <td>-</td>		GRAND TUTAL =			GRAND TOTAL =		-
<pre>x tatt: 20500 201000 17.000 17.000 100 1000 1000 1000 1</pre>		-	0.38610		•		
$\begin{array}{rcl} & C_{0} & F_{1} & W_{0} & W_{0} & V_{1} & V_{1} \\ & & C_{0} & C_{0} & C_{0} & C_{0} & C_{0} \\ & & C_{0} & C_{0} & C_{0} & C_{0} & C_{0} \\ & & C_{0} & C_{0} & C_{0} & C_{0} & C_{0} & C_{0} \\ & & & C_{0} \\ & & & & & & & & & & & & & & & & & & $	TUTEL 50.000	DEAD =		Tulat 50.000 50.000	9513 =	34 <del>3</del> 6	
	х	TUTAL ALIVE = 23.030 + TOTAL	EXPECTED FREQUENCIES	; ; ;	63.090	SOUARED USED SQUARED = 0.38610 NIFICANEE LEVEL = AI = 0.55	1 = 1 + 2 + 2 + 3 + 3 + 3 + 3 + 3 + 3 + 3 + 3

CASE

EMBAYD NORTALITY DEFOE CONPORTE C

~~~~~

ø

| CONFIGUENCY TARLE  |                 |                                 |
|--|-----------------|---------------------------------|
| 1     34.000     05.000     50.000       2     34.000     56.000     50.000       2     29.000     50.000     50.000       AL     AL     -     6.3.000     707AL       P     P     -     6.3.000     707AL     27.000       AL     AL     -     -     6.3.000     707AL     27.000       AL     AL     -     -     6.3.000     707AL     27.000       AL     AL     -     -     5.3.000     7.07AL     27.000       AL     AL     -     -     7.07AL     7.07AL     37.000 |                 |                                 |
| AL ALIVE = 63.000 . TOTAL DEAD = 37.<br>EXPECTED FREQUENCIES X52065 = 1<br>ALIVE DEAD DEAD   |                 |                                 |
| ARECTED FREQUINCIES 'XSQUES = 1<br>ALIVE DEAD FREQUINCIES 'XSQUES = 1  | - CRAND TOTAL - | 100-001                         |
| ALIV DEAU  | CUTOFF = 6      | 0.00.0                          |
|  |                 | :<br>:<br>:<br>:<br>:<br>:<br>: |
| 005-16   |                 |                                 |
| FUTAL ALEVE = 63.000 + TOTAL DEAD = 37.000 +   | - CAANT TOTAL - | 000.001                         |
| ASYMPTOTIC CHI SUUARED USED<br>095FRV-D CHI SQUARED = 1.07250<br>085ERVED SIGNTFICANCE LEVEL = AT = 0.30039  |                 | •                               |
| 2LN(AI) = 2.40541<br>[-2LN(AI)) = 2.60000  |                 | •                               |
| $V \tilde{A} \tilde{A} (-2 L \tilde{A} (A L)) = -4.00000$  | •               |                                 |

I REPORTED IN THE REPORT

312

.

CASE

### Figure VIII.5 EXAX2 output

| <b>3</b><br>504<br>504<br>701AL<br>701AL<br>701AL |
|---|
|---|

.

Second Bar

# DEFOE COMPOUND C BINARYO NORTALITY

# 

### CENTING NCY-TAHL S CASE

| COL     1     ALIVL     DEAD     TOTAL       COL     2     31.000     19.000     50.000       COL     2     31.000     19.000     50.000       TOTAL     ALIVE     62.000     19.000     50.000       TOTAL     ALIVE     62.000     19.000     50.000 |
|--|
| ALIVE DEAD<br>31.000 19.000<br>= 62.000 19.000<br>= 62.000 TDTAL D<br>ALIVE ALIVE<br>31.000 - 19.000<br>-31.000 - 19.000   |
| ALIVE<br>31.000<br>= 62.<br>= 62.<br>ALIVE<br>31.000   |
| 1<br>2<br>2<br>2<br>2<br>2<br>2<br>2<br>2<br>2<br>2<br>2<br>2<br>2<br>2<br>2<br>2<br>2<br>2<br>2   |
|  |

| STANTOTIC CUT STADED HOLD   | - |
|---|---|
|   |   |
| $\frac{165 \text{ FV} = 0}{21 \text{ M} \text{ FT} \text{ ST} = 0}$ |   |
| FY = (-2LN(AT)) = 2.00000   |   |
| $v_{A} = v_{A} R (-2L N(AI)) = 4.00060$                             |   |
| IN DE DENDENT TESTS COMBINED  |   |
| SY = SUM(YY) = 16.61076   |   |
|   |   |
|   |   |

0.96493 Z = [SURT (SY) - SURT (YMU))/SORT (SVARY/ (4\*YMU)) = UNDER HA Z HAS A STANDARD NORMAL DISTRIBUTION

Ø

Figure VIII.6 EXAX2 output

Figure VIII.7 EXAX2 output

.....

\*

PAY PRATALITY DEFOE COMPOIND C ÷

CONFINEENCY -FABLE

-1

CA BE

...

-00-05 1ı GRAND TOTAL 0.0 701AL 20.000 Ħ OEAD 40.000 . TOTAL 0:40 0.0 -----Ħ TOTAL ALIVE niu 1 -2 C 0 5 50 00 TAULE

= 1.0 1 V = SIGNTE ICANCE, LEV'L 4. 0.0 Ч С

0.0 H ((IV) YY II -2LN A FY II -2LN A FY II -2LN

Figure VIII.8 EXAX2 output

DEVE COMPOUND O FRY MORTALITY

\*\*\*\*\*\*\*\*

•

GENTINGENCY TANLE

÷

CASE 2

|             |          | -          | 20+999  |                |  |                                       |   |                 |        |
|-------------|----------|------------|---------|----------------|--|---------------------------------------|---|-----------------|--------|
| TETAL ALIVE |          | 41         | 40*000  | 40.000 . TOTAL | D7AD =   | 0.0                                   |   | GRANG TOTAL = . | CC0-04 |
| AHLE IS DE  | 170 V 39 | 1FICANC    | E LEVUL | - AI = 1.0     | TAMLE IS DEGENTRATE<br>DASSERVED SIGNIFICANCE LEVOL = AI = 1.0 | · · · · · · · · · · · · · · · · · · · | -<br>-<br>-<br>-                        |                 |        |
|             | (VI) = ( | 0.0<br>0.0 |         |                |  |                                       |   | ×               | -      |
| VARY = VAR  |          |            | 0-0     |                |  |                                       | a - a - a - a - a - a - a - a - a - a - | •               |        |

### Figure VIII.9 EXAX2 output

11 J. CV · ) 25.05 . GRAND TOTAL = 11 , GRAWN TOTAL 81 CUTCEE 2.10E74 2.600 2.003 . = 38069X 101AL 20.000 \*{ -70+044 05.40 . D E A D U 1194 0 2.000-----30.700 , TDTAL 38.000 . . TDIAL - 0000 - 0000 0640 0.10 -CONTINCENCY -TAPLE ---07:0 \*+0526 TXPECTED FREQUENCIES 1 - 0. 51 480 ISTRIBUTION US • ` | | N C C EXSCT X 50 DISTRIGUTION X 50 0.0 2.17526 0.46718 0.46718 • 2005 • ALIV. -19.000 ALIVE 20.000 29.000 . 433.75 u **H** . 11 -24 H ( 41-) TUTAL ALEVE TUTAL ALIVE COL - 1 -2LNIAJ YY = -2[N[A Y = -2[N]A VARY = -2[N GEI **с** Сос 145540 145540 FXAU

CASE

·\*

FRY MORTALITY

DEFOE COMPOUND C

Figure VIII.10 EXAX2 output

0.70033

: CDD. [A 0/5 17 500000 u h ATUT AND TUTEL , GRAND TOTAL CUTOFF = 1 1 1 1 1 1.07625 1.000 1.000 ; . H 707AL 21.700 20.000-101%L 20.000 u H 2 ទល្**ខ**េ ភ \*\*\* \* \* \* \* DEAD 0540 0.49790 40.000 . TBIAL 40.000 , TUTAL . . . . . 0.488 0.49780 01A U •07425 FUNTING NCY TARK F XPTCTED FRIQUENCIAS DISTRIBUTION USED ۰ ۷ ت. **L** 20.488 TXACT YSU DISTRIBUTION XSD PRUE . 0.37619 0.51220 ALTY ÷ Ħ 11 UBŠEAV:D ŠIGNIFJC Y = -2LN(1) = 1. - Y = [(-2LN(A]) = 1. VARY = VAP(-2LN(A]) = -AULVE ALEV ビジタつけん TOTAL TUTAL 20 50 50 50 -F XAC I

CASE

PRY MORTALLITY

DEFOE COMPOUND C

- ----

ŧ

### Figure VIII.11 EXAX2 output

じりじょしゃ ŀ . GRAND TOTAL 000.9 ŧ1 DEAD 0.000011 , TOTAL 1 h 00010 0000 Ø JISTRIBUTION USED 4 3.3 4-00 S16 ¢ è. t 31.006 TXAUT XSU DISTRIBUTION 0-12-46-15 COD C 1 U HA C H ALIVE 0414237 1.629032 1.629032 2.62500 1.662500 I ۲. ر ۲ LGIAL 1 2 Y 2

40.000

!!

, GRAND TUTAL

060.0 

11

06490

31.000 , TUIAL

Ħ

26175

TOTAL

!

55

1 1 1

TCTAL ≷0.000 20.000

ປ⊧≜ປ 4• 000 --5•005 -

ALIV: 16.000 -- 15.000

ł ţ

ţ

- CONFINGENCY TAPLE

CASE

---\*

\*

FAY NORTALITY

DEFOE COMPOUND C

2000000

cutcet =

1.5541.0

Seriesx

101AL 20.000

4.500

۵LIVE 15.509-----15.500

UK:0.

EXPECTED FREQUENCIES

1

,

555 XAt Ξ λ

0670

= ((14)

VA 50

### EXAX2 output Figure VIII.12

000.03 • JATOT CUAS 40.000 ---0.17516 \*\*\*\*\*\*\*\*\*\*\*\*\*\* PRY PORTNLITY 101AE # 50 = 0 CG 01:40 DEFOE COMPOUND G ≂ 1.0 • נטוער 76≜3 21€33 20€30€-I A . = CONFINCTINGY FAREN GAT FLOANCE LEVEL c•0 INDEPENDENT TESTS COMBINED SY = SUM((Y) = --2,87342 - ---YMU = SUM((Y) = -2,62157 SVAFY = SUM((VARY) = -3,78617 ALIV 19.0 11 21 i V? رد: **س**ر TUTAL CASE

٠.

•

\*\*\*\*\*\*\*\*\*\*

. ---

## Figure VIII.13 EXAX2 output

| GROUP     L     CONTINUE OFY TABLE       File     L     ALLV       File     ALLV     L       File     ALLV     L  |                                  |
|---|----------------------------------|
| 1     1     1     1     1     1       1     1     1     1     1     1     1       1     1     1     1     1     1     1       1     1     1     1     1     1     1       1     1     1     1     1     1     1       1     1     1     1     1     1     1       1     1     1     1     1     1     1   | ارت<br>ا                         |
| TH     2< |                                  |
| L ALIVE = 129.374 , TUVL  | 000<br>000                       |
|   | = 73,000 + 22440 TOTAL = 500,000 |
| יין איז   |                                  |
| 1 32.550<br>32.550<br>17.550<br>17.550  | 0000<br>0000                     |
| 17.500  |                                  |
| ICIAL ALIVE - 130-DLU - IUIAL USAD  |                                  |

٠,

i

.

.

## Figure VIII.14 EXAX2 output

| 2     1 <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>                 |                  |                  |         |   |   |                           |                  |                                   |          |
|---|------------------|------------------|---------|---|---|---------------------------|------------------|-----------------------------------|----------|
| 1     1 <th></th> <th></th> <th></th> <th></th> <th></th> <th>*******</th> <th>****</th> <th>***</th> <th></th>   |                  |                  |         |   |   | *******                   | ****             | ***                               |          |
| 1     1 <th>R DUP 2</th> <th>-</th> <th></th> <th>1</th> <th>V. L.</th> <th></th> <th>:</th> <th>•</th> <th></th> | R DUP 2          | -                |         | 1   | V. L.   |                           | :                | •                                 |          |
| 1     3     34.000     50.000     50.000     50.000       1     4     5     50.000     50.000     50.000     50.000       1     4     5     50.000     50.000     50.000     50.000       1     4     5     50.000     10.14     50.000       1     4     50.000     10.14     50.000       1     4     50.000     50.000       2     50.000     50.000   |                  | 1                |         |   |   |                           |                  |                                   |          |
| <pre>(AL ALIVE = 130.000 • TUTAL 0:40 = 70.000 • SARD TUTAL<br/>EXPECTED PERDUALITY<br/>ALIVE 0540 101AL<br/>2 225500 17500 50.000<br/>17.500 50.000<br/>17.500 50.000</pre>  |                  |                  | . in st | 1000<br>1000<br>1000<br>1000<br>100<br>100<br>100<br>100<br>100 | 14<br>14<br>14<br>14<br>14<br>14<br>14<br>14<br>14<br>14<br>14<br>14<br>14<br>1 | * • • •                   | :<br>:<br>:<br>: | · .<br>•<br>•                     |          |
| EXPECTED PEROURNELLE<br>EXPECTED PEROURNELLE<br>1 32: 500<br>17:500<br>17:500<br>17:500<br>17:500<br>50:000<br>17:500<br>50:000<br>17:500<br>50:000   | 3<br>3<br>4<br>3 | ¥.               | ALTV:   | 130   | • 1011  | :                         | 70.000           | . SARKI TUTAL =                   |          |
| 4 17 500<br>17 500<br>17 500<br>17 500<br>17 500<br>17 500<br>17 500<br>17 500<br>17 500  |                  |                  | T dX T  | CIEL FEEDUENC   |   |                           |                  |                                   | 5 +00000 |
| 4 24 500 17 500<br>17 500   |                  | ،<br>بيني<br>۲ ۲ | •       | ALIV:<br>32.05<br>505<br>505                                    | 01 A0<br>17 - 500   | 101AC<br>50.000<br>50.000 |                  |                                   |          |
|   |                  | xa               |         | 005-26  | 17.500  | 50.000                    |                  |                                   |          |
| TELAL ALINE - 130.001 . TULAL DEAD - 70.001 . GRAND TELAL   |                  | TLIAL            | LLVE    | 132.  |   | 0 € A 0 =                 |                  | <ul> <li>GRAND TCIAL =</li> </ul> |          |

.

. 

í.

•

322

۰.

| -<br>1<br>-<br>5  |                 | 5.00000             |                           | <b>,</b>                                   | .200.000.         |             |  |         |
|---|-----------------|---------------------|---------------------------|--|-------------------|-------------|--|---------|
|   | + 62AN0 TUTEL = | CUITEF_=5.          |                           |  | - GELAND TUTAL =  |             |  |         |
|   | 63.010          |                     |                           |  | CU0-E3            | -           | an an Indean and a state of some state of the source of th |         |
| 1011<br>1011<br>1010<br>1010<br>1010<br>1010<br>1010<br>101 | 0440 =          |                     | 5017L<br>50-000<br>50-000 |  | <u> 0 = 40 = </u> |             | 2 Lė   |         |
|   |                 |                     | ne AU<br>15-750<br>15-750 | 15.75C<br>15.750                           | U . TOTAL         | 177 XU - P  | = 1: = 0.02972   | 4.2(123 |
|   | n i             | EXPLOITU ER UULIGTE | ALTVE<br>24-250<br>24-250 | 44 • 1 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 | = :37.0C          | 960 ((\$EU• | 7.01102  | -       |
| -1-5015   | -               | T dX T              |                           | ন্ধ  | ALAV              | CHL SQUARED | 2212   |         |
|   |                 |                     | 161                       | 2 J  | TUTUT             | A CYMPTC LI |  | V-5-1   |
| 1   |                 |                     | 1                         |  |                   |             |  |         |

: • •

CONTRACT FOR

GADUP 3

HOLCOMBE AND PHIPES COMPOUND ENDING MORTALITY

•

大子女教学女子女女子女子女子女子女子女子女子女子女女子女女子女子

Figure VIII.15 EXAX2 output

Figure VIII.16 EXAX2 output

|                    | ***                   | ***************           | *******                 |                                       |
|--------------------|-----------------------|---------------------------|-------------------------|---------------------------------------|
|                    |                       |                           |                         |                                       |
| GROUP 4. CUNTINGEN | GUNCY TANL            |                           |                         | · · · · · · · · · · · · · · · · · · · |
| -                  | 0. AU<br>10. AU       |                           |                         |                                       |
| TK1 32 29 000      |                       | 2000                      |                         |                                       |
| TUTAL ALTVE =      | 122.000 . 10141       | - U:AU =                  | 70.000 . GRAND TUTAL    | AL = 200.000                          |
| EXPLOTED F         | EKENUENCIES           |                           | CUIDEE                  |                                       |
| •                  | 0<br>16-500<br>14-500 | 107AL<br>50.000<br>50.000 |                         |                                       |
|                    | ·                     | 50.000<br>50.000          |                         | •                                     |
|                    | 111 - JO-271          | DêAD ≠                    | 72.000 . GRAND . T CVAI | AL = 200.000                          |

324

25222

Figure VIII.17 EXAX2 output

| G         C(1)11 [01, b)CY 1,   | S         CONTINUE Conforme & DMAND MONTALTY           S         CONTINUE MONTANT         MAND MONTALTY           S         CONTINUE MONTANT         MAND MONTALTY           S         CONTINUE MONTANT         MAND MONTANT           ALIV         ALIVE         MAND MONTANT           TOIR         ALIVE         TOIR         TOIR           TOIR         ALIVE         TOIR         TOIR         TOIR           TOIR         ALIVE         TOIR         TOIR         TOIR         TOIR           TOIR         ALIVE         TOIR         TOIR         ALIVE         TOIR         TOIR           TOIR         ALIVE <thalive< th="">         TOIR         TOIR<th>C     C     C     C     C     C     C       T     T     T     C     C     C     C     C       T     T     T     C     C     C     C     C     C       T     T     T     C     C     C     C     C     C       T     T     T     C     C     C     C     C     C       T     T     T     C     C     C     C     C     C       T     T     T     C     C     C     C     C       T     T     T     C     C     C     C     C       T     T     T     C     C     C     C     C       T     T     T     T     C     C     C     C       T     T     T     T     T     C     C     C       T     T     T     T     T     T     C     C       T     T     T     T     T     T     T     C       T     T     T     T     T     T     T     T       T     T     T     T     T     T     T<th></th><th></th><th></th><th>-</th><th></th><th></th><th></th><th></th></th></thalive<>  | C     C     C     C     C     C     C       T     T     T     C     C     C     C     C       T     T     T     C     C     C     C     C     C       T     T     T     C     C     C     C     C     C       T     T     T     C     C     C     C     C     C       T     T     T     C     C     C     C     C     C       T     T     T     C     C     C     C     C       T     T     T     C     C     C     C     C       T     T     T     C     C     C     C     C       T     T     T     T     C     C     C     C       T     T     T     T     T     C     C     C       T     T     T     T     T     T     C     C       T     T     T     T     T     T     T     C       T     T     T     T     T     T     T     T       T     T     T     T     T     T     T <th></th> <th></th> <th></th> <th>-</th> <th></th> <th></th> <th></th> <th></th>   |  |               |  | -                           |   |        |          |             |
|---|--|--|--|---------------|--|-----------------------------|---|--------|----------|-------------|
| CITIENCENCY LANCE         TOTAL MELANCY LANCE           ALIVE         N * D         TOTAL           ALIVE         TOTAL         TOTAL         TOTAL           ALIVE         TOTAL         TOTAL         TOTAL         TOTAL         TOTAL           ALIVE         TOTAL         TOTAL         TOTAL         TOTAL         TOTAL   | <b>5</b> CUTING THE TARK TARE<br><b>6</b> CONTING TO THE TARE<br><b>1</b> ALIV = $\frac{1}{123}$ CONTING TO $\frac{1}{12}$ CONTINUE   |  |  | Veriou        | 19E AND FHIME CO   | M Powe D E                  | HOATO HOATAL  | 27     |          |             |
| $f_{11}$ $f_{$   | S     C(1) T [M(1) M(2) M(2)     TO TALE       141     1     21000     1000     1014L       141     21000     1000     1000     1014L       141     34.000     1000     1000     50.000       141     21000     1000     1000     50.000       141     2100     1000     1000     50.000       141     2100     1000     1000     50.000       141     2100     1000     1000     50.000       141     2100     1000     1000     50.000       141     2100     1000     1000     50.000       141     2100     1000     1010     50.000       141     2100     1010     1010     50.000       141     2100     1010     1010     50.000       141     2100     1010     1010     50.000       141     2100     1010     1010     50.000       141     2100     1010     1010     1010       141     2100     1010     1010     1010       141     2100     1010     1010     1000       141     2100     10100     1000     10000       141     2000     10000  | S     C(-1) [M(-MCV ]M(-V) [M(-MCV ]M(-V)]       181     1     ALIV     1000       141     21000     1000     1000       141     21000     1000     1000       141     2     2000     1000       141     2     2000     1000       141     34.000     1000     1000       141     34.000     1000     1000       141     34.000     1000     1000       141     34.000     1000     1000       141     4LIV     1000     1000       141     2     1000     1000       141     33.250     1000     1000       161     1000     1000     1000       161     1000     1000     1000       101     1010     1000   |  | -             |  | ** ***                      | ****  | *****  | ***      |             |
| 1     ALTVE     1000     1000     1000     1000       2     5000     1000     1000     50000     5000       3     5000     1000     50000     5000     50000       4     000     16.000     50.000     50.000     50.000       4     1000     16.000     50.000     50.000     50.000       4     4     1000     7014     5000     50.000       4     4     5000     50.000     50.000     50.000       5     50000     1014     50000     50.000       5     50000     50.0000     50.000     50.000       5     550     16.750     50.000     50.000       5     50000     50.000     50.000     50.000       5     550     16.750     50.000     50.000       5     550     16.750     50.000     50.000       5     550     16.750     50.000     50.000       5     550     16.750     50.000     50.000       5     550     16.750     50.000     50.000       5     550     16.750     50.000     50.000       5     5     50.000     50.000     50.000 <td><math display="block">\begin{array}{cccccccccccccccccccccccccccccccccccc</math></td> <td>13.000     13.000     10.000     10.000       13.000     14.000     10.000     10.000       13.000     16.000     10.000     10.000       13.000     16.000     10.00     10.000       13.000     16.000     10.00       13.000     16.000     10.00       13.000     16.000     10.00       13.000     16.000     10.00       13.000     10.00     10.00</td> <td></td> <td></td> <td>X</td> <td></td> <td></td> <td>-</td> <td></td> <td></td>  | $\begin{array}{cccccccccccccccccccccccccccccccccccc$   | 13.000     13.000     10.000     10.000       13.000     14.000     10.000     10.000       13.000     16.000     10.000     10.000       13.000     16.000     10.00     10.000       13.000     16.000     10.00       13.000     16.000     10.00       13.000     16.000     10.00       13.000     16.000     10.00       13.000     10.00     10.00  |  |               | X  |                             |   | -      |          |             |
| 3       34.000       14.000       50.000  | $\frac{133.000}{14.000} = \frac{5000}{50.000} = 67.000 + 6400 \text{ Intal} = 2000 \\ \frac{133.000}{16.000} = 7010 = 67.000 + 6400 \text{ Intal} = 2000 \\ \frac{1000}{16.750} = \frac{101}{50.000} = 67.000 + 6400 \text{ Intal} = 2000 \\ \frac{1000}{16.750} = \frac{101}{50.000} = 67.000 + 6400 \text{ Intal} = 2000 \\ \frac{1000}{16.750} = \frac{1000}{50.000} = 67.000 + 6400 \text{ Intal} = 2000 \\ \frac{1000}{16.750} = \frac{1000}{50.000} = 67.000 + 6400 \text{ Intal} = 2000 \\ \frac{1000}{16.750} = \frac{1000}{50.000} = 67.000 + 6400 \text{ Intal} = 2000 \\ \frac{1000}{16.750} = \frac{1000}{50.000} = 67.000 + 6400 \text{ Intal} = 2000 \\ \frac{1000}{16.750} = \frac{1000}{50.000} = 67.000 + 6400 \text{ Intal} = 2000 \\ \frac{1000}{16.750} = \frac{1000}{50.000} = \frac{1000}{50.000} = 67.000 + 6400 \text{ Intal} = 2000 \\ \frac{1000}{16.750} = \frac{1000}{50.000} = \frac{1000}{50.000} = 67.000 + 6400 \text{ Intal} = 2000 \\ \frac{1000}{16.750} = \frac{1000}{50.000} = \frac{10000}{50.000} = \frac{10000}{50.000} = \frac{10000}{50.000} = \frac{1000}{50.$  | 00     15.000     50.000       15.000     16.000     50.000       15.100     16.000     50.000       15.100     16.120     50.000       15.150     16.120     50.000       125.000     16.121     0.20   |  | -             | ALIV:<br>2000  |                             | 101AL<br>50.000   |        |          |             |
| AL       AL <td< td=""><td><math display="block">\frac{153.000}{1000} \cdot \frac{1013L}{1000} = 67.000 \cdot 630M0 TrTAL = 200</math></td><td>153.060 • TETAL DEAD =</td><td><br/>YYX<br/></td><td>্ৰজন্ম</td><td>46 000<br/>246 000<br/>246 000<br/>246 000</td><td>-15+000<br/>14+000<br/>16+000</td><td>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>50000<br/>500000<br/>5000000</td><td></td><td></td><td>-<br/>-<br/>-</td></td<> | $\frac{153.000}{1000} \cdot \frac{1013L}{1000} = 67.000 \cdot 630M0 TrTAL = 200$   | 153.060 • TETAL DEAD =   | <br>YYX<br>  | ্ৰজন্ম        | 46 000<br>246 000<br>246 000<br>246 000                      | -15+000<br>14+000<br>16+000 | 50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>50000<br>500000<br>5000000 |        |          | -<br>-<br>- |
| L XPECT J FRANUFACTUS<br>1 811 VF<br>2 83.2550 16.750 50.000<br>3 2.2550 16.750 50.000<br>5 32.2550 16.750 50.000<br>6 750 16.750 50.000<br>16.750 50.000<br>17.750 50.0000<br>17.750 50.000<br>17.750 50.0000<br>17.750 50.0000<br>17.750 50.000<br>1  | Kuluikties     Curuet =     5.0000       0     10.150     101AL       0     16.750     50.000       0     16.750     50.000       0     16.750     50.000       125.00     10.20     67.000       125.00     10.20     67.000       125.00     1.32.421     0.20       125.00     1.32.421     0.20  | 0     10.00     10.00       0     10.00     10.00       0     10.00     10.00       0     10.00     10.00       0     10.11     0.00       10     10.11     0.00   | -  | :<br>E        | 4 . T.   | 14101 4                     |   | 67.000 |          | 200,000     |
| 1 211VF (140)<br>2 33.250 16.750 50.000<br>3 3.250 16.750 50.000<br>3 3.250 16.750 50.000<br>5 3.250 16.750 50.000<br>16.750 50.000<br>17.750 50.000<br>17.7500 50.0000<br>17.7500 50.0000<br>17.7500 50.000                            | $\frac{1}{125} \frac{1}{125} \frac{1}$ | 60 166-756 50,000<br>166-750 50,000<br>50 166-750<br>50 50,000<br>166-750 50,000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>120-000<br>1 |  | -dX-1         |  | 11.5                        |   |        | CUTOEL = | • 0,000 0   |
| 5 32.250 16.750 50.000<br>6 3.250 16.750 50.000<br>AL ALIVE = 123.000 - TOTAL DIAD = 67.000 - GAANO-JUTAL = 300   | $\frac{125 - 360}{125 - 360} = \frac{50 - 000}{101 - 30} = -67 - 600 - 62 - 62 - 62 - 62 - 62 - 62 - 6$  | 50 15-750 50.000<br>123-304 - 10121 0120 =   | ана рад<br>1921 год<br>1925 год<br>1920 год<br>192 | <b>***</b> ** | 9.00<br>9.00<br>9.00<br>0.00<br>0.00<br>0.00<br>0.00<br>0.00 | $< \cdot \cdot$             | 10101<br>50.000   |        |          |             |
| ALIVÉ = 133.300 . TOTAL DEAD = 67.000 . GRAND TUTAL = 335   | $\frac{125 - 3C_{1}}{120} + \frac{1013L}{1 - 32423} = \frac{67 - 600}{1 - 32423} + \frac{6200}{1 - 3242} + \frac{6200}{1 - 3242} + \frac{6200}{1 - 3242} + 6$   | 123-300 - TOIAL 0140 =<br>120.   |  | in.t          | 0.92<br>2.02<br>2.02<br>2.02<br>2.02<br>2.02<br>2.02<br>2.02 |                             | 50 • 000<br>50 • 000  |        |          |             |
|   | 10.<br>1.2.2.<br>1.2.2.<br>1.2.2.  | •  | TUTA .   | LILLY         | 1  | -ICLIL                      | 0140 =  |        |          | 300-700     |

C

. . .

## Figure VIII.18 EXAX2 output

|  | Norcomat  | HOLCOMBE AND PHIPPE COM                | CAMPANIO O BAGATO         | ATS MORTALETY                         |         | -               |   |
|--|---|--|---------------------------|---------------------------------------|---------|-----------------|---|
|  |   |  | ****                      | · · · · · · · · · · · · · · · · · · · | ****    | ****            |   |
| GROUP 6  | •   | CLATINGESS 1                           | LANLE                     |                                       |         |                 |   |
| 1.1  | `-  |  | 5-1 <b>8</b> 5            | INTAL<br>N. DAN                       |         |                 |   |
| 1997<br>1997<br>1997   | (41) <b>4</b>   | 0000<br>000<br>• • •<br>• • •<br>• • • |                           | 50 000<br>50 000<br>50 000            |         | - N - A - F     | ·<br>•<br>•<br>•                                  |
| T. 1A  | AL ALV  | = 1.4.01                               | 101.6                     | . 95.80 =                             | 76.091  | - TULUL UNVER   | 55 <b>5</b> - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - |
| and the second | <u></u>   | עריכ ויע הייצענוונ                     | ) - T -                   |                                       |         | CUTPEE.=        | . JOCOL .   |
|  | ,<br>,<br>,<br>,<br>,<br>,<br>,<br>,<br>,<br>,<br>,<br>,<br>,<br>,<br>,<br>,<br>,<br>,<br>, |  | 5-000<br>19-000<br>19-000 | 16 [AL<br>50.000                      | - I     | -               |   |
| - 1  | er <b>4</b>   |  |                           | 50•000<br>50•000                      |         |                 |   |
| TE   | 1.1.4   | <u>= 124+705</u>                       | INI OI INI                | DEAL                                  | 76.000  | - TRINC TRIAL - | 332.4.625   |
| A LEAD A   | VEJ.CHI STO   | Studale USED.                          | 64152-1                   |                                       |         |                 |   |
| 21 - 3<br>   | Y =   | • • • • • • • • • • • • • • • • • • •  | 20.010 - 0.02.4           | 1986 e                                |         |                 | -   |
|  | TNOEPENDENT TESTS COMPLUI<br>= SUV(YY) = TESTS COMPLUI<br>= SUV(YY) = TESTS COMPLUI         | STS COMBINED                           |                           | •                                     |         |                 |   |
| Z = (S 241 (2Y)  | 5.Un(VIRY) = 1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1  | + T ( YMU ) ) Z = 11                   | 1 VASA (444               | Y M(1) ) = <b>1.3</b>                 | 1.31060 |                 | •   |

## Figure VIII.19 FXAX2 output

= (((X))) =

• ٠.

100-001 100-001 - 5. 00000 . GRAND JUTAL = 0-000 + 6KANL TUTAL = - CUTUFF = 6 • UUO 11 300 25.000 25.000 TOTAL JUTAL DEAD 1.0 000 94.000 . TUIAL 15181001100 0F X50 574115110 USEV-VLU ASV = 0.70922 VLU ASV = 0.20922 VLU ASV = 0.20001 54211 = ---0.00000 DC 00000 000-1 2010105 ULAU CUNIINGENCY TABLE 9606 68 --- EAPLETED FREQUENTES 42446 +1 EAALI ASU UISTRIBUTION X54 0.10722 U.54213 S 2.10722 U.24213 U 2.245106 U.24117 U 2.2451740 U.247117 0.247117 U 2.240142 U.247117 U 19.24244 U.200054 U 19.14374 U.200554 U ALIV 23.5000 23.5000 23.5000 23.5000 23.5000 23.5000 AL IVE 24-0000-24-0000-24-0000-24-0000-24-0000-u - TUÍAL ALLY. ALAVE TUFAL XXXX EEEE EXALI 1 **BRUP** 

\*

Convine o Pry Partial 17

NOLCOMOE AND PHINE

(

## Figure VIII.20 EXAX2 output

•

## \*

. . .

. . .

# HOLOWIE AND PHINE COMPOUND O PAY NORTALITY

# 

.

• • `

|  |  |  | CUNTINGENCY TABLE                             | TABLE  |   |             |               |           |            |
|--|--|--|---|--|---|-------------|---------------|-----------|------------|
|  | Tkl  | -  | AL1VE<br>22-000                               | UEAU<br>3-000  | 101AL   |             |               |           |            |
|  |  | 174  | 24+000<br>23+0000<br>23+0000                  | 0000   | 0000  |             |               |           | • .<br>• . |
|  | TUTAL  | ALIVE  | 000 • 76                                      | U . TUTAL  | DtAU =  | 8 •000      | . UKAND IUTAL | 17L =     | 100.000    |
|  |  | 3484   | LAKECIED-FREQUENCIES                          | ctes   |   |             |               | LUTOFF .= | 0000       |
|  | 111  | <b>1</b> 971 \$                              | ALIVE<br>25.000<br>23.000<br>23.000<br>23.000 | C C C C C C C C C C C C C C C C C C C                                      | 101<br>25.000<br>25.000<br>25.000<br>25.000<br>25.000 | •           | :<br>:        | · .       |            |
| i<br>i<br>i  | IUTAL  | ALIVE  |   | 1.1  | DEAD =  | d •U00      | -+            | IAL = .   | 100-001    |
|  | E AAC I  |  | NULIU   | c t c  | ·   |             |               |           |            |
|  |  |  | v.04453 1                                     | 1.0000   | 1<br>   |             |               |           | •          |
|  |  |  |   | U-6227c  |   |             |               |           |            |
| :<br>:<br>!  | 4 1 4<br>4 4 4<br>5 5 5<br>1   |  |   | 2.2008.  | ann an an ann an ann an ann ann ann                   | :<br>:<br>: |               |           | •          |
|  |  |  |   |  |   |             |               |           |            |
|  | 20   |  |   | 04470-0  |   |             |               |           |            |
|  | - , - , - ,<br>, - , - , - ,<br>, - , - , - ,  | ¥•/8601                                      |   | C=U1424<br>C=C1136   |   |             |               |           |            |
| -<br>-<br>-  |  | ļ  |   | 0.00444<br>0.000444<br>0.000202  | 4-<br>4   |             |               |           |            |
| т<br>К<br>К<br>С<br>С<br>С<br>С<br>С<br>С<br>С<br>С<br>С<br>С<br>С<br>С<br>С<br>С<br>С<br>С<br>С | EXACT . ULSTREBUT<br>UDSERVEU 354<br>UDSERVEU 314<br>TT = -LLM(ALL 314<br>ET = -LLM(ALL 314) | CI.ULSTRIBUTAUN UF. XJ<br>UDSENVEU ASU =<br> | ים" ר   | STATISTIC WSED.<br>STATISTIC WSED.<br>Laboyo<br>Laboyo<br>Laboyo<br>Laboyo | 1+90  |             |               |           |            |

-

### Figure VIII.21 EXAX2 output

100.000 000-007 00000-5 . GRAND IUTAL = ... CUTOFF = 8.000 1 8 + UUC 00000 UEAU = IVIN -- ULAU JU = 1.000870 JUNIFICANUE LEVEL = A1 = 0.01575 = 2.15948 ł 1 TUIAL -+- TUTAL ULATABUTICN UP ASH STATISTIC USEU. ţ 2000 2000 00000 ×ו000 ו•0000 ו•0000 2-000 3=44]40 0\*077700000 0\*077700000 0\*077700000 0\*0777000000 0\*0777000000 LUNTING NCY TABLE ----EXMELTED FREQUENCIES \*\*\*\* 1424601 74-000 EXALT XSQ UISTHIBUTION ALIVE 23-1000 243-0000 243-0000 AL 14 25+000 25+000 24+000 1100000 0000 ~~~~ 5 H >
- (IŢVIN]
- ALIVE 10-01-01-01 10-4 1020 20-00070 06000. このつの・ 070 Tulat XXXX UBSER CBSE EAACT **Group** 

COMPOUND FRY HONTALSTY

And Phipp

HOLOMOR

•

|       | Holcome                  | E ANO   | PH IPPE                                       | COMPOU   | NOD FRY   | HORTAL                                   | ETY            |           |             |       |         |
|-------|--------------------------|---|---|----------|---|--|----------------|-----------|-------------|-------|---------|
|       |                          |   | · · · ·                                       | ****     | *******   | ******                                   | ***            | *******   | ******      |       |         |
| GROUP | ₩                        |   | CONTIN  | GENLY 1/ | ABLE  |  |                |           |             |       |         |
|       | TRI<br>TRT<br>TRT        | 4<br>4  | ALIVE<br>24.000<br>20.000<br>21.000<br>25.000 | 0<br>U   | UEAD<br>- 4+000<br>5+000<br>4+000<br>0+0          | 101A<br>25.00<br>25.00<br>25.00<br>25.00 | 00<br>UU<br>00 | <b></b> . |             |       |         |
|       | FUFAL                    | ALIVE   |   | 87.000   | . 101 AL  | UEAD                                     | #              | 13.000    | . GRAND TO  | IAL - | 100-000 |
|       |                          | . <b>2</b> .XP  | LCTED FI                                      | KEQUENC. | 1£5   |  |                |           | LUTOFI      | - = : | 00000   |
|       | IRI<br>TRI<br>IRI<br>Tri | 1<br>2<br>3<br>4  | ALIVE<br>21.75<br>21.75<br>21.75<br>21.75     | 0        | 0EAU<br>3+250<br>3+250<br>3+250<br>3+250<br>3+250 | TU IA<br>25+0<br>25+0<br>25+0<br>25+0    | 00<br>00<br>00 |           |             |       |         |
|       | TUTAL                    | ALÍVZ   |   | 87.000   | + TUTAL   | DEAD                                     | <b>E</b>       | 13.000-   | + GRANU TUI | AL =  | 100.000 |
| Ē     |                          | 612743<br>1744<br>1744<br>1744<br>23574<br>2449<br>2449<br>2449<br>2449<br>2449<br>2449<br>2449<br>24 |   |          | >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>                  |  |                |           | •<br>•      |       |         |

Figure VIII.22 EXAX2 output

Figure VIII.23 EXAX2 output

|                   |          |                             | 200              |                      | -  | 30          |  |
|-------------------|----------|-----------------------------|------------------|----------------------|--|-------------|--|
|                   |          |                             | 100 - 001        | 00000-4              |  | 100-001     |  |
|                   |          | ·                           | d                | ) • ¢ -              |  | 14          |  |
|                   |          |                             |                  | भ<br>भ               |  |             |  |
|                   |          |                             | GKAND ILIAL      | cuTurf =             |  | יו וי       | ۲  |
|                   |          |                             | uk Al            |                      |  | GRANU FUTAL |  |
|                   |          | ÷                           |                  | !                    |  |             |  |
|                   |          |                             | 19.00.41         | 1<br>-<br>-<br>-     |  | 19-400      | · · · · · · · · · · · · · · · · · · ·  |
|                   |          |                             |                  |                      |  | t           | · ·  |
|                   | AL<br>AL |                             | : <b>1</b> 4     | ł                    | 1010<br>25.0000<br>25.0000<br>25.0000<br>25.0000 | P <b>F</b>  |  |
|                   | Į OI     | 1000                        | DFAU             |                      | 123222   | UEAU        | 1.9  |
| •                 | _        |                             | AL               | İ                    |  | Å           | AL = 0.07167   |
|                   | EAU      | 140000<br>140000<br>1400000 | , TUTAL          |                      | UEAU<br>144750<br>144750<br>144750               | + 1UT AL    |  |
| Y TAB             | 21       |                             | 000              | ENCLE                | 3-4-4-A  | • 000 •     |  |
| ic .NC            | 0        | 222                         | 2 <b>1</b> .     | REQUI                | 0003   | 77          | 1.0.0  |
| CUNITNG NCY TABLE | ALIVE    | 2000<br>2000                | :<br>:<br>:<br>: | EARECTEU FREQUENCIES | AU000  | 1           | AS FMPL UL LU CHA SUUAKEU USEU-<br>UESERVEU CHA SUUAKEU USEU-<br>UESERVEU SIGNIFICANUE LEVEL =<br>Y = -2LN(A) = -2.00000<br>EY = VAK(-2LN(A)) = -2.00000 |
|                   |          |                             | ALIVE            | ANEC.                | •<br>•<br>•                                      | ALIVE       |  |
| •                 | -        | 11111                       | N.               |                      |  | í           |  |
|                   | y<br>Y   |                             | <b>I</b> JLAL    | 1                    | 777<br>777<br>777<br>777                         | TUTAL -     | 21 UI 10 CH1 SUUAI<br>35 ER VEU CH1 SUUAI<br>35 ER VEU SIGNIFLI<br>25 ER VEU SIGNIFLI<br>25 EN (12 LN (12 L) =<br>= - V AR (-2 LN (AL)                   |
| 3                 | -        |                             | ١ <u>ـ</u>       |                      |  | I           |  |
| GROU P            |          |                             |                  | 4<br>4               |  |             | AV<br>VAV  |
| 3 <b>8</b> 1      | I        |                             |                  | !                    | 1  | ļ           | t t  |

COMPOUND PRY NONTALETY HOLONIE AND PHIPPE

\*\*\*\*\*\*\*\*\*\*\*

11 . GRAND JUTAL \* IAULE IS WEGUNERATES RUM SUM 15 LERU UK AT LEAST KAL LUL SUMS AKE ZERD. Ubserved Significance Level = AI = 1.0 IV = -2clnica) = 0.0 EV = EI-2clotai) = 0.0 Vaky = Vaki-2clutai) = 0.0 i 100-001 1.09755 10100 2550000 2550000 2550000 2550000 2550000 2550000 Ħ DEAU H Z = (SUATISY) + SUATITAU)/SUATISVAARY/(4\*YAU)) UNUER HU 2 HAS A STANDARD NURMAL UISTREBUTEUN 1 4 4 . TUTAL 0000 67 67 7 UEAU 25-000-LUNIINGENCY TABLE 13-38451 8-11241 = 11-34129 0.0 INDEPENDENT TESTS COMBINED ALIVE 3009 3330 11 n ALAVE THU H SUR(ET) H VARY H SUR(VARY) n 3 IULAL XXXX GROUP 1

\*

COMPONIO D FRY MORTALETY

AND PHIAG

NOLONDE

100.000

Figure VIII.24 EXAX2 output

Figure VIII.25 EXAX2 output

|       |                |           | JARVINEN                          | O                 |  | Compaulo & BNDAYO MORTALITY  | *           | ***     | · 1             |          |
|-------|----------------|-----------|-----------------------------------|-------------------|--|--|-------------|---------|-----------------|----------|
| GROUP | ч              |           | LUNTINI                           | LUNTINLENCY TABLE | 18 <b>6</b> 5  |  |             |         |                 |          |
|       | IXI<br>IXI     |           | AL IVE<br>59.000                  | •                 | UEAU<br>12.000   | 101AL<br>51.000<br>50.000  |             | 1       | •               |          |
|       | IUÍAL          | ALAVE     | -<br>11                           | 81.060            | 81.000 . TUTAL   | UEAU =   | 14.000      |         | • GRAND TUTAL = | 101-000  |
| •     |                | EXPE      | EAPECIEU -FREQUENCIES             | LEGUENLI          | lÊ3  |  |             | (U)     | CUTOFF = .      | 5. 00000 |
|       | 1 X 1<br>1 X 1 |           | AL I VE<br>43.931<br>- 43.069     | i                 | UEAU<br>7.009<br>6.433   | 101AL<br>51.000<br>50.000  | •           | • *     | ·               |          |
|       | IUTAL          | ALIVE     | H                                 | 87.000            | · TUTAL  | UEAU =   | 14+000      | • GRANU | • GRANU TUTAL = | 101-000  |
| **    | r* chí su      | JUANEN HA | N) BEEN                           | LUKKEL            | IEU FUK CU   | **** CHÍ SUUANEU MAS BEEN LUKKELTEU FUK LUNTINUITY, SINCE U.F. = 1. **** | iŭE 0.eF. = | 1. **** | •               |          |
|       |                | CH1 5005  | KKEU US.<br>JAKED US.<br>ICANCE L | cU.<br>EVEL =     | ASYMPTUTIC CHI JUUAKEU USZU.<br>UBSCKVED CHI SUUAKEU = 0.51206<br>UBSCKVEU SLONIFICANCE LEVEL = A1 = 0.01071 | 11 01  |             | -       | :               | •        |
|       |                |           |                                   | 0000              | - 4 • 0000   |  |             |         | i<br>           |          |

\*

**)**.

19

Figure VIII.26 EXAX2 output

109-000 109.600 8.000 GKAND TUTAL = • GRAND TOTAL = \* 9.000 ; 101AL 57-000 ų ŧ UEAU ULAU LSFKABUTIUM OF ASU STATISTIC USED. 24ED X29 = 1-78430 24ED X16NIFICANLE LEVEL = A1 = 0-27477 24ED 31 = 2-58361 = 101.000 + TUTAL , **TUTAL** DEAU 6.000 2.000 1065201 00000 EXPECTED FREQUENCIES LUNTINGENCY TABLE 1.024.1 ALIVE 52.817 48.103 101-000 **S16** EXACT ASQ UISTALBUTION AL LYE 51.004 0-02614 0-02614 0-02031 16 ALIVE ALIVE 11 JUIAL TUT AL TAT TAT TX TXT d EXACI GLOUP

\*\*\*\*\*\*\*\*\*\*\*\*\*\*

TAXYARA COMPANIE BHARTO MONTALITY

.

:

. 

.\*

## \*

\$

. . .

# ZARVINEN CONCOURS ENDAYD MONTALITY

| IKI         Z         *3*000           IUIAL         ALIVE         91.           IVIAL         EXPELIEU FKEQU         91.           IXI         2         45.500           IXI         2         45.500           IXI         2         45.500           IXI         2         95.500           IVIAL         2         45.500           IXI         2         95.500           IVIAL         2         91.01           IXI         2         91.01           IVIAL         2         91.01 |   | UIAL UEAD - 000<br>000 500-000<br>000 500-000<br>000 500-000<br>000 500-000<br>000 500-000<br>000 10 10 10 10 10 10 10 10 10 10 10 10 | 000 | <ul> <li>GRANU TUTAL = 100.</li> <li>GUTUFF = 5.00000</li> <li>GRAND TUTAL = 100.</li> </ul> | 100.000<br>5.00000<br>100.000 |
|---|---|---|-----|--|-------------------------------|
| TO ULST WILL US UN  | ALLE LEVEL = AL = AL = AL = AL = AL = AL = LEVEL = 0.02089<br>•02822 0.02089<br>•02822 0.02089<br>•02263 0.00268<br>•02263 0.02086<br>•02263 0.02086<br>•02629 0.02086<br>•00266<br>•02629 0.02086<br>•02629 0.02086<br>•02620 0.02086<br>•02600 00000000000000000000000000000000 | 1.001.001<br>1.001<br>1.001   |     |  |                               |

Figure VIII.27 EXAX2 output

į

### Figure VIII.28 EXAX2 output

| - 🕿            |   |
|----------------|---|
| - #            |   |
| - #            | ٠ |
| - 2            |   |
|                |   |
|                | • |
| - <b>R</b>     |   |
| ÷              |   |
| - #            |   |
| Ä              |   |
| - 2            |   |
| <u>ज</u>       |   |
| - <del></del>  |   |
| - 🛪            |   |
| - ¥            |   |
| - #            |   |
| - <del>4</del> |   |
| 프              |   |
|                |   |
| - <u>T</u>     |   |
| . 🖀            |   |
| - <b>*</b>     |   |
| *              |   |
| -#             |   |
| - ä            |   |
|                |   |
|                |   |
| - <del></del>  |   |
| -#             | • |
| -#             |   |
| *              |   |
| ÷.             |   |
| 1              |   |
| Σ              |   |
| ۳ <b>۳</b>     |   |
| -#             |   |
| -#             |   |
|                |   |
| Ŧ              |   |
|                |   |
|                |   |
| *              |   |
| *              |   |
| *              |   |
| #              |   |
|                |   |
| -              |   |
|                |   |
|                |   |

۳.

ţ.

•

## JARVALA COMPAND & ENDAYO MORTALITY

### \*\*\*\*\*\*\*\*\*\*\*\*\*\*\* •• •, .

1

| TKT         2         **1.VE<br>**7.000         Teato<br>1.000         TUTAL           1UTAL         ALIVE         **7.000         1.000         \$8.000           1UTAL         ALIVE         **7.000         1.000         \$8.000           1UTAL         ALIVE         **7.000         1.000         \$0.000           1UTAL         ALIVE         **0.000         10.1AL         \$8.000           1NTAL         ALIVE         C40         \$0.000         \$0.000           1NTAL         ALIVE         C40         \$0.000         \$0.000           1NTAL         ALIVE         C40         \$0.000         \$0.000           1NT         A         ALIVE         C400         \$0.000           1NT         A         ALIVE         C400         \$0.000           1NT         A         \$0.000         \$0.000         \$0.000           1UTAL         ALIVE         \$0.000         \$0.000         \$0.000           1UTAL         ALIVE         \$0.000         \$0.000         \$0.000           1UTAL         ALIVE         \$0.000         \$0.000         \$0.000           \$0.000         \$0.000         \$0.000         \$0.000         \$0.000 | Group | 4       |  |         | CUNTINGLNCY TABLE          | INCY IA            | bLE                     |                                |          |       | -                                       |                |   |
|---|-------|---------|--|---------|----------------------------|--------------------|-------------------------|--------------------------------|----------|-------|---|----------------|---|
| UEAD = 2.000 . GRANU TUTAL =<br>JUTAL<br>50.000<br>0.000 . GRAND TUTAL =<br>2.000 . GRAND TUTAL =   |       | TX<br>X | ;<br>;<br>;  | 4.7     | 111VE                      | •                  | LcAD<br>1.000<br>1.000  | 101<br>102<br>000              | 00<br>00 |       |   | -              | •   |
| JUTAL       JUTAL       >•00         50.000       • 6.000       • 6.000         00000       • 6.000       • 6.000   | -     | 10      | TAL  | ALIVE   | *<br>*<br>*                | 96.000             | , 101AL                 | UEAD                           | N        | 2.000 | . GRAND TI                              | UTAL =         | 48 • U00  |
| 1014L<br>50.600<br>68.000 = 2.000 , GRAND TUTAL =<br>00000  |       |         |  | E KPE   | Cleu-FKI                   | CULENC L           | <b>F 3</b>              |                                |          |       | יטוט                                    | 11<br>11<br>11 |   |
| UEAD = 2.000 , GRAND TUTAL =<br>00000   |       | 77      | :<br>فير مور   |         | ALIVE<br>48.980            | •                  | DEAU<br>1.020<br>-0.580 | 107A                           | 00C      |       |   | •              | •<br>•<br>•   |
| EXACT ASH WISTRIBUTION<br>0.500 WES WISTRIBUTION<br>1.960 WES WISTRIBUTION<br>1.960 WES WISTRIBUTION<br>1.960 WESTRIBUTION<br>1.960 WESTRIBUTION WE X54 STATISTIC USE D.<br>EAALI WISTRIBUTION WE X54 STATISTIC USE D.<br>WESERVED ASH FILANCE LEVEL = AI = 1.000000<br>TT = -LANEND = 0.00000  |       | n I     | TAL  | ALAVE   |                            | 46 <b>-</b> 000    | , TUTAL                 | UEAD                           | 44       | 2.000 | , GRAND TI                              | UTAL =         | 98 <b>•</b> 000   |
| <pre>c.ico74 0.23732 0.23732<br/>c.ico74 0.23732 0.23732<br/>cadi uis[ribuiiun uf xS4 STafistic UseD.<br/>ubserveu asu = ubserveu significance level = af = 1.00000<br/>rt = -2LN(af) = 0.00000</pre>   |       |         | EXACT<br>SUCT  | LASH UI | STR16U1<br>PR08<br>U-50494 | 10N<br>515<br>1.00 | 0000                    |                                |          |       |   |                |   |
| EAACI UISTATBUTTUN UF XS4 STATISTIC USED.<br>UBSERVEU A34 = G.OCOUS G.OCOUS UBSERVEU SIGNIFICANCE LEVEL = AI = 1.00000<br>YT = -2LNIAL) = 0.00000   | •     | :<br>:  | 2 - 1 - 7<br>2 - 1 - 7   | 014     | 0.23732                    | 0.0                | 3732                    | ****************************** | 1        |       | 5 · · · · · · · · · · · · · · · · · · · | •              | :   |
| $\begin{array}{l} ubserved significance \ Level = \ AI = 1.00000 \\ vr = -2Ln(AI) = 0.00000 \end{array}$  | L.    | AACI U  | 1.51.41  |         | UF XS4                     | STATIST<br>COCCERS | IC USED.                |                                |          |       |   |                |   |
|   | λ     |         | K<br>N<br>N<br>N<br>N<br>N<br>N<br>N<br>N<br>N<br>N<br>N<br>N<br>N<br>N<br>N<br>N<br>N<br>N<br>N | SLUNIFL | LANCE L                    | EVEL =             | AI = 1.0(               | 0000                           | -        |       |   |                | -<br>-<br>-<br>-<br>-<br>-<br>-<br>-<br>-<br>-<br>-<br>-<br>-<br>-<br>-<br>-<br>-<br>-<br>- |

UBSERVEU SIGNAT ...... = -2LN(AI) = = E(-2LN(AI)) = K = VAR(-2LN(AI)) =

÷

.64511 ..... 1.38125

### Figure VIII.29 EXAX2 output

105-000 100-000 ł 00000.4 . GKAND IOTAL = ١ culore = . GKAND TUTAL -----000.0 6.000 ..... - 1 • 101AL 52-000 Ħ Ħ DÉAU DEAU JARVINEN CHADIND & BHANYE NUNTALETY . IUTAL 52.000 , 101AL ULSIMICUIUM UN ANY SIMILSIAC USEU. DEAD 2.000 2.000 UEAU 2+571 3+029 021+5+3 0000 212 CUNTINGENCY TABLE EAPELIEU FREUUENLIES . 000. W 24264. 0.63 EXACT X34 UISINIGUTIUN ALIVE 44.971 AL 4VE 0-24167 0-24127 0-24127 0-24127 C-11+21 C-1202 v c c a O . u I UNITICANCE VARI-ALNIALII B 3 'n . ALIVL ALI VL • **8** 1 TULAL **FULAL** TAT TAT 2 Txľ Ivľ TT = -21 FT = -21 VANY = -1 EXALI 3 -

avon

\*

Sector.

### Figure VIII.30 EXAX2 output

101-000 101-000 H , 6KANU TUTAL = -CUTUFE.E GRANU TUTAL \* -----. 9000 R 8.000 2.54488 101AL 50.000 101 × 101 × 101 × 101 × 1000 4 DEAU UEAU Z # LILARIATI - LUNITAUITAUITAURIAVANTIATAUI # •75436 /el = Al = U•05406 N DAN , lula 43.000 . Tul AL ATSSIL USED. ĎEAU 3+960 --- 4+646--DEAD 1.000 125 3•1 1400 0000 14441. CUNTINGENCY TABLE EXCELTED FREQUENCIES 000-64 200 ALIVE 40.040 40.9960 8461C-11 EXALT XSQ UISTRIBUTION INDEPENDENT TESTS COMBINED AL1VE 44.0000 1142 002000-0 7 222000 Der in mitugatio Bellalles 0.100 1500-0 UBSERVEN XLU = UBSERVEN SJUITE ALANLE #1 = +ZLHEALA = ET = E1-2LEALA = VART = VART-2LAIAL) = 3 , ų 1 -INU = JUALEY) = J JAAT = JUALATI = J ALIVE ALAVE 47 JARVAN 1 x54 0.60085 0.50085 101101 101114 101114 101114 101114 2.08053 1403-2 TUTAL トーラク TUIAL INT ... TX TX EALT 4 19 GROUP

1



•••

|                    | '               |
|--------------------|-----------------|
|                    | *****           |
| WY NORTHLY         | *************** |
| KVINEN COMPOND & 1 |                 |
| JARVI W            |                 |

•

:

1

| 1: 20 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: | 1. 2. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. | $\begin{bmatrix} 1 & 15 & 57 \\ 1 & 15 & 57 \\ 1 & 15 & 57 \\ 1 & 15 & 57 \\ 1 & 15 & 57 \\ 1 & 15 & 57 \\ 1 & 15 & 15 \\ 1 & 15 & 15 \\ 1 & 15 & 15$   |          |       |           |    | L.         |                           |          |             |                    |
|--|--|---|----------|-------|-----------|----|------------|---------------------------|----------|-------------|--------------------|
| .ο   | ουί - ΤυΤΑΙ - ΓΈΑΝ                           |   |          |       | ,         |    |            | 101AL<br>16.000<br>14.000 | 5<br>    |             |                    |
|  |  | T. A. J. D.GEN AMER at A SUM J. ZAMERS AT LEASE KET LEE SUMM ARE ZIRU.<br>AMER V.O. STOMFORMANCHER E. A. H. A.  | 4        | F. C. | -W[]A-    | )f | 6 + 10 FAL | = U≠±u                    |          | - 17101 UNE | ن <b>ا در</b> در ا |
|  |  | $\begin{array}{c} \mathbf{v}_{1} \in \mathbf{v}_{1} \cup \mathbf{v}_{2} \in \mathbf{v}_{1} \in \mathbf{v}$ | <b>P</b> |       | Ser. sof. |    |            | 1.1.1.4.5.1.4.1.(1.1      | STIME AD |             |                    |

Ĩ.''

Figure VIII.31 EXAX2 output

339

: . ..

Figure VIII.32 EXAX2 output

|         | 1<br>1<br>1<br>1 |                                      |                              |                           |             |                 |  |
|---------|------------------|--------------------------------------|------------------------------|---------------------------|-------------|-----------------|--|
| GROUP 2 | 2                | CONTANGARY TRALE                     | 12.21.                       |                           |             |                 |  |
| •       |                  | <br>ALIVE<br>1917<br>19100<br>149000 | 1000<br>1000<br>1000<br>1000 | 101AL<br>15.000<br>14.000 |             |                 |  |
|         |                  |                                      |                              |                           | <u>ن</u> •ر | , GRAND TOTAL = |  |

. . . .

١.

500.03

-ي '

我在我外生我来来这些我这些女子的人生我这么么这些我去的人,只是这个这次这个的 大手手 法说

# \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

• JARYINEN COMPOND & FAY MARTALETY

\*\*\*\*\*\*\*\*\*

|                 | •                         | GRAND TUTAL = 34.000 |  |
|-----------------|---------------------------|----------------------|--|
|                 |                           | CeAb                 |  |
|                 |                           | e<br>C               | A SMUS -   |
|                 | 15000<br>15000<br>15000   | = U4_C               | 1.15 ZEHU UR AT LEAST K-L CUL SUMS ARE ZERU.   |
|                 | 0.00<br>0.00<br>0.0       | 30.300 . 1014L 02AU  |  |
| CUNTING ACY LAU | 417VE<br>15.000<br>15.000 |                      | U-CANCE LEVEL  |
| Ĵ               | e-154                     | 4                    | $I(A) = \frac{1}{2} \frac{1}{$ |
| 64              | I XI                      | TUTAL ALAV-          |  |
| CADUP           |                           |                      |  |

1

Figure VIII.33 EXAX2 output

Figure VIII.34 EXAX2 output

| *          |
|------------|
| £ .        |
| ₩.         |
| <b>#</b> ` |
| *          |
| *          |
| 8          |
| ***        |
| <u>.</u>   |
| ж.         |
| Ξ.         |
| *          |
| 꾼.         |
| *          |
| ****       |
|            |
| #          |
| *          |
| ***        |
|            |
| ¥.         |
| ÷          |
| *          |
| Ξ.         |
| **         |
| *          |
| *          |
|            |
| *          |
|            |
| *          |
| ÷          |
| ***        |
| ¥          |
| *          |
| 4          |
| 5          |
| *          |
|            |
| 춙          |
|            |
| ***        |
| #          |
| - <b>3</b> |
|            |
|            |

| ×. | •                                 |             |
|----|-----------------------------------|-------------|
| ŀ. |                                   |             |
|    |                                   | ł           |
| i. | ÷ .                               |             |
| Ϊ. | •                                 | <u>,</u> `• |
| Ľ  |                                   |             |
| ١. |                                   | . 1         |
| ł  |                                   | 1           |
| Γ. |                                   |             |
| 1  |                                   |             |
| Ľ  | Ξ.                                | -           |
| İ. | •                                 |             |
| ł  |                                   | ۰,          |
| ŀ  |                                   | ÷,          |
| Ŀ  | ` مو                              |             |
| ł  |                                   |             |
| 1  | H                                 |             |
| Ŀ  | . فلب                             |             |
| ŀ  | <b>K</b>                          | . •         |
| μ  |                                   |             |
| ļ  | <b>K</b> .                        |             |
| Ľ  | <u>9</u> .                        | ·           |
| H  | Ζ.                                |             |
| Ļ  | -                                 |             |
| L  | 5                                 |             |
| ĥ  |                                   |             |
| ľ  |                                   |             |
| ł  |                                   |             |
| ł  | 0                                 | -           |
| ł  | TARY WEN COMPOUND & FAY MORTALITY | ĩ           |
| Į. | 2                                 | ٢.          |
| ł  | 5                                 | ~           |
| ł  | a                                 |             |
| i  | 7                                 |             |
| ł  | 8                                 | -           |
| ł  | Q                                 | ۰.          |
|    |                                   |             |
| ł  | 3                                 | 1           |
| 1  | ¥                                 |             |
| ł  | 3                                 | 1           |
| 1  | 2                                 | ;           |
| ļ  | Ş                                 | Ľ           |
| t  | 9                                 | 5           |
|    |                                   |             |
|    |                                   |             |

| . 1      |        |        |
|----------|--------|--------|
| 1        |        |        |
| 1        | ĥ      | •      |
|          | H      | ÷      |
| ł        | ×      | •      |
| -        | K      |        |
|          | ž      | •      |
|          | 7      |        |
| ۰.,      | 1      | Ē      |
|          | e      | •      |
|          | Ņ      | ŀ      |
| 1        | 1      | •      |
|          | 2      | •      |
| 1        | 7.<br> | •      |
| - 1      |        | ,<br>  |
|          |        | •      |
|          | 1      | Þ      |
|          | +      | ŀ      |
|          | ł      | 2      |
|          | 1      |        |
|          | 2      | 5      |
|          | 1      | ŀ      |
|          | •      | Þ      |
|          | 4      | ŀ      |
|          | 1      | Ę      |
|          | *      | r<br>2 |
|          | ł      | ÷      |
|          | ÷      | -      |
|          | Ì      | ŀ      |
|          | ,      | ł      |
|          | ł      | Ź      |
|          | •      | ĩ      |
| μ.       |        |        |
|          | 4      | ÷      |
| <b>.</b> |        |        |
|          | 2      | •      |
|          |        | .,     |
|          |        |        |
| Ċ        |        | •      |
|          |        |        |
|          |        |        |

| BRUUF 4 | •           |              |                              |                           |     |                 |
|---------|-------------|--------------|------------------------------|---------------------------|-----|-----------------|
|         |             | - A-1 VE<br> | <b>5</b> 00                  | 10101<br>15-000<br>15-000 |     | •               |
| i       | ורודר שרוא. | ł            | <u>56.000 , 10161 - DEAU</u> | Ú≧∆U =                    | 0.0 | . GRAND TOTAL = |

CANCE LEV

-

(C) T

----

(AI) = 3.0

110.02

t ,

### Figure VIII.35 EXAX2 output

1

.

٠,

٠.

|       | JARVINGO COMPOUND B PAY PUNIALIT   |
|-------|--|
| erout | S. CLMT INLINCY TAR.   |
|       | The I         I         I = 1         I =  |
|       | TUAL ALIVE - 14.00. TUAL DEAD - 16.000 . 64400 TUTAL - 20.000  |
| -     | CUTCFF =   |
| •     | TRY 1 441V5 UPAU 701AL 101AL 15.000  |
|       | TUTAL ALIVE = 14.000 + TUTAL 0'AU = 16.000 + GRAND TUTAL = 30,000  |
| **    | **** Chil SyurgED HAS 51 H LUPR CLID FUN CONTINUITY, SINCE D.F. = 1. ****  |
| 7     | ASYMPTCFIC (HI NUUAPUD USEN)<br>DESERVEN CHI SEUAREL = DAIBJSJ<br>VPUERVEN CHI SEUAREL = DAIBJSJ<br>VY = TELM(AI) =<br>VY = TEL |
| Y     | .Y = E(-2L4(AI)) = / Z. G(UU<br>V23Y = VAS(-2LALZI) = / 2.00000  |

医化学学者 网络带山北北北北北北北北北北北北北北北北北北北部的东西的东西

Figure VIII.36 EXAX2 output

000.05 a JATAT DIAL CIL SUMS ARE ZERD. \* 30.000 -0.84013 1 1018L 15.000 15.000 UN 1 = ( (I)WA()) - TUTAL 0640 15.050 15.050 Criat House CY Avela 
 IABLE IS UNCONSENTE AND SUM SUM IS Z

 PLASERVED SERVICE SERVICE

 YY = -2LN(AI)

 YY = -2LN(AI)

 YY = -2LN(AI)

 YY = -2LN(AI)
 0• 0 INDEPENDENT TESTS COMBINED ·# SVAFY = SUN(VAPY) = JLLV 1 LIAL אט ב בטא • GADUP ł 

JARVINEN COMPONIO B FAY MONTALITY

唐云在北京市市市市市大学大学大学大学中学学生大学大学大学大学大学学生 法法国政会

APPENDIX AVIII.2\* DESCRIPTION AND INSTRUCTIONS FOR USE OF EXAX2 COMPUTER PROGRAM

J.

<sup>\*</sup> Appendix AVIII.2 is an appendix for Section .III.

EXAX2--A COMPUTER PROGRAM TO COMPARE BINOMIAL PROPORTIONS\*

Program Description and Card Input Information

Paul I. Feder and Susan A. Willavize Department of Statistics The Ohio State University

September 12, 1980

\* This work was performed with the support of the U.S. Army Medical Bioengineering Research and Development Laboratory, Frederick, Maryland under Contract DAMD17-79-C-9150 at The Ohio State University.

### EXAX2--A Computer Program to Compare Binomial Proportions

### I. Introduction

Dichotom us data arise in many fields of application. In fish toxicology experiments it is of interest whether an embryo hatches into a live fry, whether the fry has survived 27 days post hatching, whether the surviving fry are normal or abnormal. In toxicology experiments on higher animals, such as mammals, the presence or absence of specific types of tumors or deformities is of importance. Such responses naturally give rise to 0-1 or success-failure type data. Such 0-1 data are encountered also in many other fields. For example in industrial applications it is noted whether or not a unit meets design specifications or whether or not a unit lasts beyond the warranty period. In sociological applications it is of interest to note whether an individual exhibits certain behavior patterns, has specific opinions, etc. A myriad of additional applications could be cited.

Part of the analysis of success-failure data involves estimating the probabilities of "success" within various groups and comparing these probabilities across groups. EXAX2 is a computational tool to assist in carrying out such comparisons. Since the need for the program was motivatied by problems arising in aquatic toxicology, the remainder of the section centers around such applications.

In toxicity tests on fish or daphnids, a number of test concentrations are run along with one or more control groups. Within each concentration group (treatment or control) several tanks or beakers are run, each chamber containing a number of the organisms under study. Within each chamber the numbers of embryos, numbers of fry hatched live or normal, numbers of fry surviving or normal at the conclusion of the experiment are recorded. It is desired to compare the proportions of live or normal embryos or fry in the various treatment groups with corresponding proportions in the control groups.

A preliminary inference of importance is to test the response proportions among the tanks or beakers within each group for homogeneity. If there is no evidence of tank to tank heterogeneity within concentration groups, the data can be pooled across tanks within groups and further analyses carried out based on binomial theory. However if evidence exists of tank to tank heterogeneity within groups, subsequent analyses must be adjusted to reflect this, either by adjusting the model or the data or by carrying out analyses on a per tank basis.

EXAX2 carries out tests of homogeneity of tanks within groups based on the chi square statistic. If the expected response frequencies are "large enough" the distribution of the test statistic is approximated by large sample chi square theory. If the expected response frequencies are not large enough for asymptotic theory to be applicable, the test statistic is evaluated based on its exact small sample distribution, derived from the exact small sample distribution of the contingency table, conditional on the margins (March, 1972). Individual tests of homogeneity within groups are combined by means of Fisher's method (Littell and Folks, 1971, 1973) to obtain an overall test of homogeneity.

EXAX2 also has the capability to test for heterogeneity of response rates across treatment groups based on responses pooled within groups. Either the exact small sample or approximate large sample distribution of the chi square statistic is utilized. Heterogeneity among tanks within groups can be accounted for either by modifying the chi square statistic by a heterogeneity factor (Finney, 1971) or by modifying the data to "effective frequencies" to reflect within tank correlation.

EXAX2 can calculate exact confidence intervals on the odds ratio of a treatment group response rate as compared with a control group response rate, based on the traull distribution of Fisher's exact test for 2x2 tables (Thomas, 1971). The traulence interval calculations are based on responses pooled across tanks within The response frequencies would need to be modified to "effective frequencies"

to reflect within tank correlation.

Section II discusses program organization and capabilities and provides a more detailed description of the program's procedures. Section III contains detailed instructions for card input.

### II Program Organization and Capabilities

Suppose that the aquatic toxicity test consists of N treatment groups (including both test and control groups) and K tanks or beakers per group. (EXAX2 can handle different values of K for each group, however we assume a single K value for notational convenience.) Thus the responses within each treatment group can be summarized as a 2xK contingency table. The rows represent the response category (e.g. dead, live or abnormal, normal, etc.) Each column represents the responses from an individual tank. We wish to compare response probabilities across columns.

It is first necessary to specify the entries in the tables. A 2xK contingency table can be specified as a 3x(K+1) matrix partitioned as follows:

|                             |                      | <b>Å</b> 3 6                       |          |            | Row              |          |
|-----------------------------|----------------------|------------------------------------|----------|------------|------------------|----------|
| Bow                         | $\frac{COL}{V(1-1)}$ | $\frac{\text{CoJ}}{\text{X}(1,2)}$ |          |            | - Totals<br>R(T) |          |
|                             | A(1, 1)              | A(1, 27                            | •••      | A(1, A)    | 11(1)            |          |
| Row 2                       | X(2, 1)              | X(2, 2)                            | • • •    | X(2, K)    | R(2)             |          |
| Row 1<br>Row 2<br>Col Total | C(1)                 | <b>じ(2)</b>                        | • • •    | С(к)       | XN               |          |
| where the $X($              |                      | se the "body"                      | " of the | table R(1) | and $R(2)$       | ane mu 1 |

where the X(I,J) comprise the "body" of the table; R(I) and R(2) are row totals; C(1), C(2), ..., C(K) are the column totals; and XN is the grand total.

The information in the first K columns of each matrix is inputted one column at a time either by specifying the body of the table (X(1, 1), X(2, 1), ..., X(1, K),X(2, K)) or the column totals and the first row (C(1), X(1, 1), ..., C(K), X(1, K)),or the column totals and the second row (C(1), X(2, 1), ..., C(K), X(2, K)) (See section III, intructions for card input). The remaining elements are computed and the complete matrix is printed. A single matrix or several matrices must be inputted, depending on the purpose of the analysis. Each inputted matrix is first examined to detect and adjust for the following conditions -- if they exist:

1.) If any column totals are zero those columns are deleted, K is reduced accordingly, and the new matrix is printed.

2.) If any row total is zero or if only one column total is nonzero, then that table is the only possible table with the given row and column totals. As such it is defined to be degenerate and the observed level of significance (for future heterogeneity tests) is set to 1. A message to this effect is printed.

3.) Steps 1 and 2 above are repeated for each succeeding inputted contingency table.

We first consider tests of homogeneity within treatment groups and later we will discuss tests across treatment groups.

If a table is not degenerate, the expected cell frequencies (EX(I, J)) are calculated as EX(I, J) = R(I) C(J) /XN (I = 1, 2; J = 1, ..., K) and XSQ is calculated as  $XSQ = \sum_{j=1}^{K} \sum_{l=1}^{2} (X(I, J) - EX(I, J))^2 / EX(I, J)$ . If K = 2 a correction for continuity is applied to improve the convergence of the distribution of XSQ to its asymptotic chi square form. Namely if K = 2  $XSQ = \sum_{j=1}^{2} \sum_{l=1}^{2} (|X(I, J) - EX(I, J)| - 1/2)^2 / EX(I, J)$ . The table of expected frequencies and the user specified cutoff value, CUTOFF (for what constitutes a "large" expected frequency within each cell), are printed.

The table of expected frequencies is then examined to see if any of the expected frequencies are less than CUTOFF, e.g. 5. If not XSQ is considered to be asymptotically distributed as  $\chi^2$  with K-1 degrees of freedom, and its significance level (A<sub>i</sub>) is calculated based on chi square theory. The observed chi square value and its significance level (A<sub>i</sub>) are printed. If one or more of the expected frequencies is less than CUTOFF, the <u>exact</u> distribution of the XSQ statistic is calculated. This is done by enumerating all possible tables with the given row

and column totals (algorithm due to Boulton and Wallace, 1973) and their associated chi square values. Under the assumption of homogeneity of response probabilities across columns the probability of each possible table, conditional on the row and column margins fixed, is (March, 1972)

> ПС(J)! R(1)! R(2)!J=1 XN! ПХ(I, J)! I,J

From this exact distribution over possible tables the exact distribution of XSQ is derived. Based on this derived distribution, the significance level  $(A_i)$  of the XSQ value associated with the observed table (XSQOBS) is calculated as the probability of a XSQ value greater than or equal to XSQOBS. The observed XSQ value and significance level  $(A_i)$  are printed and optionally (see instructions for card input) the exact XSQ distribution is printed. The entire process is then repeated on the next contingency table.

After tests of homogeneity (asymptotic or exact) have been carried out on each treatment group, the significance levels  $A_1, A_2, \ldots, A_n$  summarize the results of the independent tests of the homogeneity on each table. To obtain an overall significance level these independent  $A_i$ 's are combined as follows:

For groups where the distribution of XSQ has been approximated by its asymptotic chi square form, the null distribution of  $A_i$  is approximately uniform (0, 1). Thus  $Y_i \equiv -2 \ln (A_i)$  has an approximate chi square distribution with 2 degrees of freedom, mean  $E(Y_i) \equiv EY_i \equiv E(-2 \ln (A_i)) = 2$ , and variance Var  $(Y_i) \equiv VARY_i \equiv Var (-2 \ln (A_i)) = 4$ .

For groups where the exact small sample distribution of XSQ has been used,  $A_{\underline{i}}$  and  $Y_{\underline{i}} \equiv -2 \ln (A_{\underline{i}})$  have discrete null distributions derived from the null distribution of the contingency table. The mean  $E(Y_{\underline{i}}) \equiv EY_{\underline{i}}$  and variance Var  $(Y_{\underline{i}}) \equiv VARY_{\underline{i}}$  are calculated from the exact distribution of  $Y_{\underline{i}} = -2 \ln (A_{\underline{i}})$ .

In either case (i.e. exact or asymptotic) the values of  $Y_i$ ,  $EY_i$ , and  $VARY_i$  are calculated and printed along with the other resluts for each table. The test  $\cdot$  statistic for the overall test of tank to tank homogeneity within groups is

$$Z = \frac{\sqrt{\Sigma Y_{i}} - \sqrt{\Sigma E Y_{i}}}{\sqrt{\Sigma VARY_{i}} / (4 \Sigma E Y_{i})}$$

It is calculated and printed at the end of the output. Under the null hypothesis that the tables are all homogeneous, Z has an approximate standard normal distribution. The null hypothesis is rejected for large values of Z.

We now consider additional applications of EXAX2. The program has the capability to carry out chi square tests of homogeneity across treatment groups and to construct exact confidence intervals on the odds ratios of treatment groups compared with the control. These applications are discussed in turn, beginning with tests of homogeneity across groups.

If preliminary tests do not reveal heterogeneity among tanks within treatment groups, then it is appropriate to sum the observed frequencies in the individual tanks within each treatment group. This results in a new 2xN contingency table which can be tested for homogeneity across treatment groups. EXAX2 can perform the appropriate summing within groups and then proceed with the chi square test across groups, based either on exact or asymptotic theory, as discussed above. Since just one contingency table is involved in this application, the Z statistic is not computed.

EXAX2 has the capability to compare, on a pairwise basis, the odds (p/(1-p)) within each treatment group to the odds in a user specified group; e.g. the control group. Using an algorithm given by Thomas (1971) an exact confidence interval is computed for each odds ratio (one per treatment group). The user specifies both the upper and lower alpha levels, thus permitting either one sided or two sided confidence intervals.

To illustrate how this works consider the frequencies for treatment 1 (the control group) and treatment T as forming a 2x2 contingency table:

|            | Col 1   | Col T    | Row Totals |
|------------|---------|----------|------------|
| Row 1      | X(1, 1) | X(1, T)  | Y(1)       |
| Row 2      | X(2, 1) | (2, 'I') | Y(2)       |
| Col Totals | C(1)    | C(I)     | YN         |

Y(1), Y(2), YN designate respectively the two row totals and the grand total of this new table. The odds ratio PSI is defined as PSI =  $(P_1Q_1) / (P_TQ_1)$ where  $P_1$ ,  $P_T$  are the category 2 probabilities (i.e. "success") within treatment groups 1, T respectively and  $Q_1 \equiv 1-P_1$ ,  $Q_T \equiv 1-P_T$  are the category 1 probabilities within treatment groups 1, T respectively.

We estimate these quantities by

 $\hat{P}_{1} = X(2, 1) / C(1)$   $\hat{P}_{T} = X(2, T) / C(T)$   $\hat{Q}_{1} = X(1, 1) / C(1) \equiv 1 - P_{1}$   $\hat{Q}_{T} = X(1, T) / C(T) \equiv 1 - P_{T}$   $\hat{PSI} = (X(2, 1) X (1, T)) / (X(1, 1) X (2, T)).$ 

Thomas' algorithm is an iterative technique for finding upper and lower conficence bounds on PSI. It is based on the noncentral distribution of Fisher's exact test statistic.

### III Instructions for Card Input

In this section we present detailed instructions for card input to EXAX2.

The card input consists of 5-12 Program Information Cards followed by the Data Cards and the Alpha Card. The Program Information Cards are, in order: one Input Option Card, one Parameter Card, one Format Card, one Header Card, one Labels Card, and from one to six Title Cards. These cards must be punched as described below:

1. INPUT OPTION CARD This card should have a 1 in card column 1 if the

subsequent data cards represent K tanks per treatment group and the tank data is to be pooled (summed) within each treatment group to produce one 2XN contingency table which will then be analyzed. There should be a 2 in card column 1 if no pooling is to occur. In this case, the data will be tested for homogeneity among tanks within treatment groups.

### 2. PARAMETER CARD

| Card Cols. | Description  |
|------------|--|
| 1-5        | K = the number of columns in each<br>contingency table (a right justified<br>integer) (K is the number of tanks<br>per treatment group.) K should be less<br>than or equal to 12 if the number on the<br>Input Option Card is 2.   |
| 6-15       | CUIDFF = the smallest expected cell<br>frequency with which the use of the<br>asymptotic chi square approximation is<br>permitted. If one or more expected cell<br>frequencies are smaller than CUTOFF then<br>the exact small sample distribution of<br>XSQ is used. (CUTOFF is a real number<br>with decimal point.) |
| 16-19      | blanks   |
| 20         | NTITLE = the number of Title Cards used<br>(an integer from 1 to 6)  |
| 21-25      | N = the number of 2XK contingency tables,<br>i.e. the number of groups (both treatment<br>and control.)<br>If the number on the Input Option Card is<br>is 1, then N should be less than or equal<br>to 12.  |
| 26-29      | blanks   |
| 30         | IOPT = 1, if the exact distribution of<br>XSQ is to be printed<br>= 0, otherwise.  |
| 31-34      | blanks   |
| 35         | IDATA = 1, if data card input is of the form<br>C(I), X(1, I) (see the description of the<br>Data Cards given below for an explanation of<br>this notation)<br>= 2 if data card input is C(I), X(2, I)<br>= 3 if data card input is X(1, I), X(2, I)   |

### Description

Card Cols.

40

IC = group number corresponding to group (control or treatment) to which each other treatment group is to be compared when calculating confidence intervals on the odds ratios (e.g. the number of the control group).

It should be noted with respect to the specification of K in columns 1-5 of the parameter card that some groups may have L<K tanks. The associated contingency tables for these groups thus have L<K columns. These tables must be augmented with K-L columns of zeros by inputting K-L Data Cards containing zeros. The program will later delete these dummy columns and perform its computations based only on the original L columns.

### 3. FORMAT CARD

This card contains the format in which the subsequent data cards have been punched, e.g. (F5.0, F5.0), (see the description of the data cards given below). Note: <u>No I or A formats may be used!</u> All 80 card columns may be used. The format statement may be positioned anywhere on the card and must include the parentheses (see any FORTRAN text for an explanation of **formats**).

### 4. HEADER CARD

This contains a special heading (it may be blanks) which will be printed above the output for each contingency table. Any or all of the 80 columns may be used. (e.g. Embryo Mortality)

### 5. LABELS CARD

| Card Cols. | Description  |
|------------|--|
| ⊥-2        | blanks   |
| 3-10       | A labeı for the first row of the contin-<br>gency table                        |
| 11-12      | blanks   |
| 13-20      | A label for the second row of the contin-<br>gency table (e.g. "live", "dead") |

### 6. TITLE CARDS

These cards contain titles which will be printed above the output for each contingency table. Any or all of the 80 columns on each card may be used. The number of Title Cards must agree with the entry NTITLE on the Parameter Card. If no titles are desired there must be a blank card here and NTITLE must equal 1.

### 7. DATA CARDS

An observed 2xK contingency table can be specified as a 3x(K+1) matrix partitioned as follows:

| _           | Col 1   | Col 2   | ••• | COT K   | Row<br>Totals |
|-------------|---------|---------|-----|---------|---------------|
| Row 1       | X(1, 1) | X(1, 2) | ••• | х(1, к) | R(1)          |
| Row 2       | X(2, 1) | X(2, 2) | ••• | х(2, К) | R(2)          |
| Col. Totals | C(1)    | C(2)    | ••• | С(К)    | XN            |

The X(I, J) comprise the "body" of the table; R(1) and R(2) are row totals; C(1), C(2), ..., C(K) are the column totals; and XN is the grand total. Each Data Card contains the content of one of the first K columns of this matrix. This information Lar be specified in one of three ways. If IDATA = 1 (in card column 35 of the "arameter Card), the I-th Data Card must contain C(I), X(1, I). If IDATA = 2, the I-th Tata Card must contain C(I), X(2, I). If IDATA = 3, the I-th Data Card must contain X(1, I), X(2, I). EXAX2 will compute the other matrix elements. The format given on the second Program Information Card specifies the format under which this information will be read.

The K Data Cards for the first contingency table are followed by the K Data Cards for the second contingency table and so on until all N contingency tables have been given. There will thus be NxK Data Cards in all. The form of the data for each contingency table must be consistent with the entries on the Program Information Cards. 8. ALPHA CARD

| Card Cols | Description   |
|-----------|---|
| 1-5       | ALPHAL = the lower alpha level for the odds ratio confidence intervals  |
| 6-10      | ALPHAU = the upper alpha level for the odds ratio confidence intervals. |

These must be decimal numbers less than or equal to one and their sum must be less than or equal to one.

If the data consist of just one tank per group or if EXAX2 is instructed to pool data across tanks within groups, then confidence intervals on the odds ratios are automatically computed. If the data consist of more than one tank per group (i.e. K>1) and EXAX2 is instructed <u>not</u> to pool across tanks within groups then confidence intervals are not computed and this card may be omitted.

The input cards are placed at the end of the program deck between the //GØ.SYSIN DD \* card and the/\* card. The example below illustrates the input stream.

(15x+F5.0+F5.0)

### EMBRYO MORTALITY

1

1 1

| ALIVE D              | EAD        |  |   |     |        |                           |
|----------------------|------------|--|---|-----|--------|---------------------------|
| INVESTIGATUN: H      | JLCOMBE &  | PHIPPS                                       | (06),   | TE  | ST: D  | (03)                      |
| UATA FRUM EARLY      |            |  |   |     |        |                           |
| SIX LEVELS: 1(C      | UNTRULI, 2 | 2 (LUWES                                     | T),   | • • | 6(HIGH | EST); 4 REPS EA (A+B+C+D) |
| # EMBRYUS TESTE      |            |  |   |     |        | FRY AFTER HATCH           |
| # FRY TESTED, #      |            | END, #                                       |   |     | T END  |                           |
| 1 A 1 10003          | 50 33      | 3 33   | 25  | 23  | 23     |                           |
| 1 0 1 10003          | 50 29      | -  | 25  | 23  | 23     |                           |
| 1 C 1 10603          | 50 38      |  | 25  | 24  | 24     |                           |
| 1 U 1 10003          | <u> </u>   | 0 30   | 25  | 24  | 24     |                           |
| 2 A 1 10003          | 50 31      |  | 25  | 22  | 22     |                           |
| 2 0 1 10003          | 50 36      |  | 25  | 24  | 24     |                           |
| 2 C 1 10603          | 50 34      | والمكالة والتستكيب بتهيير البواسات المحاجبين | 25  | 23  | 23     |                           |
| 2 Ú 1 10603          | 50 29      |  | 25  | 23  | 23     |                           |
| 3 A 1 10603          | 50 35      |  | 25  | 23  | 23     |                           |
| 30110003             | 50 39      |  | 25  | 25  | 25     |                           |
| د 1 1000 د C د C د C | 50 26      |  | 25  | 20  | 20     |                           |
| 1 10603 ن د          | 50 38      |  | 25  | 24  | 24     |                           |
| 4 A 1 10003          | 50 30      |  | and the second se | 21  | 21     |                           |
| 4 0 1 10003          | 50 29      |  | 25  | 20  | 20     |                           |
| 4 L 1 10003          | 50 27      |  | 25  | 21  | 21     |                           |
| 4 0 1 1 2003         | 50 36      |  | 25  | 25  | 25     |                           |
| 5 A 1 10003          | 50 32      |  | 25  | 02  | 02     | -                         |
| 5 0 1 10003          | 50 31      |  | 25  | 04  | 04     |                           |
| 5 Č 1 10003          | <u> </u>   |  | 25  | 09  | 09     |                           |
| 5 U 1 10003          | 50 34      | · <u>-</u> -                                 | 25  | 06  | 06     |                           |
| 6 A 1 10603          | 50 36      | -  | 25  | 00  | 00     |                           |
| o o l 10003          | 50 34      | · • -  | 25  | 00  | 00     |                           |
| o L 1 10603          | 50 25      |  | 25  | 00  | 00     |                           |
| 6 J 1 10603          | 50 29      | 9 OC   | 25  | 00  | 00     |                           |
| 0.0250.025           |            |  | <b>.</b>  |     | ·      |                           |

In this example the tank data will be pooled within treatments and 95% confidence intervals (ALPHAL = 0.025, ALPHAU = 0.025) will be computed on the odds ratios between treatment one and every other treatment. There are six treatments with 4 tanks each. The exact distribution of XSQ will be printed if it is used in the homogeneity test; i.e., if any of the expected frequencies is less than 5.0 (CUTOFF). There are five title cards preceeding the data cards from which the total number of embryos and the number alive will be read.

### IV Program Limitations

.

ĺ

ŗ,

EXAX2 has limitations of time and space. The space limitations are due to the dimensioned size of various arrays. When pooling tanks within treatments, N must be less than or equal to 12 and when not pooling tanks, K must be less than or equal to 12. This limit can be raised by increasing the sizes of the arrays in the following table:

### Program Location

### Arrays

na nime i su stati a nime (ma su s

| MAIN program      | X, C, EX, ICHECK           |
|-------------------|----------------------------|
| SUBROUTINE TABLE  | Х, С                       |
| SUBROUTINE INPUT  | Х, С                       |
| SUBROUTINE LNPUTI | Х, С                       |
| SUBROUTINE INPUT2 | Х, С                       |
| SUBROUTINE EXAX   | Y, RX, XSQT, D, V, U, DL1M |

Depending upon the number of tables (with the given margins) that are enumerated in generating the exact distribution of XSQ, it may also be necessary to increase the sizes of the arrays SIG, XSQ, PROB and IPOINT in SUBROUTINE EXAX. Also in FUNCTION FUN (a part of Thomas' confidence interval algorithm) there is a machinedependent constant (DPC) which should be set in the DATA statement to the largest real number the machine can hold.

In some cases the CPU time required by the program may be quite large. No systematic study has been performed to determine when this is so, but the following two examples may be of interest:

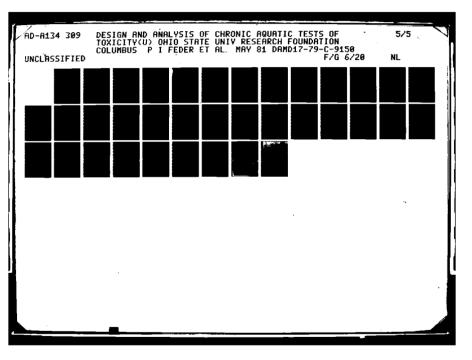
1.) For a 2x10 table with column margins all equal to 50 and row margins of 46 and 454, the smallest expected cell frequency is 4.6. If CUTOFF = 5.0, the exact distribution of XSQ will be generated requiring the enumberation of over 435,000 tables and a CPU time of over 10 minutes on the AMDAHL 470 V6 computer.

2.) For a 2x6 table with column margins all equal to 60 and row margins of 333 and 27, the smallest expected frequency is 4.5. If CUTOFF = 5.0, the exact distribution of XSQ will be generated requiring a CPU time of over 50 seconds.

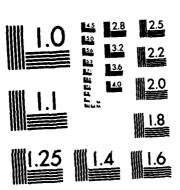
It should also be noted that, as Agresti and Wackerly (1977) point cut, generally for a fixed grand total the number of tables enumerated (and hence the CPU time) in generating the exact distribution is much higher when the column margins are equal than when they are unequal.

### References

- Agresti, A. and Wackerly, D. (1977). Some Exact Conditional Tests of Independence for RxC Classification Tables. <u>Psychometrika</u>, 42, pp. 111-125.
- Boulton, D. M. and Wallace, C. S. (1973). Occupancy of a Rectangular Array. <u>Computer Journal</u>, 16, pp. 57-63.
- 3. Finney, D. J. (1971). <u>Probit Analysis</u>, 3rd ed., Cambridge University Press.
- Littell, R. C. and Folks, J. L. (1971, 1973) Asymptotic Optimality of Fisher's Method of Compining Independent Tests, I and II, Journal of the Amer. Statist Assn, 66, pp. 802-806 and 68. pp. 193-194.
- 5. March, D. L. (1972). Algorithm 434: Exact probabilities for RXC contingency tables. <u>Comm. of ACM.</u> 15, pp. 991-992.
- 6. Thomas, D. G. (1971). Exact Confidence Limits for the Odds Ratio in a 2x2 Table. Applied Statistics, 20, pp. 105-110.







MICROCOPY RESOLUTION TEST CHART NATIONAL BUREAU OF STANDARDS-1963-A APPENDIX AX.\* THEORETICAL BASES OF SUGGESTED OUTLIER DETECTION TRANSFORMATIONS

In this appendix we discuss the motivations and theoretical bases underlying the outlier detection procedures that are illustrated in the body of the section.

Suppose that the data originate from I treatment groups, J tanks per group. Consider the i-th group. Let  $X_{ij}$ ,  $N_{ij}$  denote the number of responses and the total number of fish in the j-th tank, j = 1, ..., J, and let  $\hat{p}_i$  denote the pooled estimated response rate  $(\hat{p}_i = \sum_j X_{ij} / \sum_j N_{ij})$ ,  $\hat{q}_i = 1 - \hat{p}_i$ .

In subsequent discussion we omit the subscript i for notational convenience and so these quantities are denoted as  $\hat{p}$ ,  $\hat{q}$ ,  $X_j$ ,  $N_j$ , respectively. Let  $N \equiv \sum_{j} N_j$ .

Consider the chi square statistic for testing for tank to tank heterogeneity within groups

$$\chi^{2} \equiv \sum_{j=1}^{J} \frac{(x_{j} - N_{j}\hat{p})^{2}}{N_{j}\hat{p}\hat{q}}$$

This statistic is distributed as chi square with J-1 d.f. under the null hypothesis of no tank to tank heterogeneity. For the purpose of detecting outlying responses we consider three cases:

Case 1: All expected frequencies within the group are greater than a specific cutoff, e.g. 5.

- Case 2:  $\hat{p} < 0.1$  or  $\hat{p} > 0.9$ .
- Case 3:  $0.1 \le \hat{p} \le 0.9$  and one or more expected frequencies of responses is less than the cutoff, e.g.5.

We suggest somewhat different transformation in each case.

Case 1. All the expected frequencies within the group exceed the cutoff

Consider the individual terms in the chi square test statistic,  $(X_j - N_j\hat{p})/[N_j\hat{p}\hat{q}]^{1/2}$ . Assume that the weights in the denominator are "correct" and "fixed". The quantities  $X_j$ ,  $\hat{p}$  in the numerator are correlated since  $\hat{p}$  includes  $X_i$ . It can be shown that the variance of

Appendix AX is an appendix for Section X.

 $(X_j - N_j \hat{p})/[N_j \hat{p} \hat{q}]^{1/2}$  is  $1 - (N_j/N)$ . If all the N<sub>j</sub>'s are equal, this variance is 1 - 1/J. If all the expected frequencies in the table exceed 5, as is the case with the Defoe compound C embryo mortality data, then the quantities

$$(1 - N_j/N)^{-1/2} \left( \frac{X_j - N_j \hat{p}}{[N_j \hat{p} \hat{q}]^{1/2}} \right) \qquad j = 1, ..., J$$

can be treated as having an approximate standard normal distribution. Graphical and numerical outlier detection procedures are based on these standardized ratios. We pool them across groups and plot the resulting IJ values on normal probability paper. However, for the purpose of formal inference we account approximately for the correlation among terms within groups (approximately -1/(J - 1)) by treating the J values within each group as if they were J - 1 independent values. This adjustment of course has the most impact when J = 2. The normality assumption might be enhanced by first carrying out an arc sine variance stabilizing transformation.

### Case 2. Expected response probability within the group is small e.g. $\hat{p} < 0.1$

Case 2 is also applicable to the situation when  $\hat{p}$ >.9, by concidering the complementary response.

The distribution of X<sub>j</sub> can be approximated by a Poisson distribution with mean  $\lambda_j \equiv N_j p$ . The variance stabilizing transformation in the Poisson case is well known to be the square root transformation. In particular,  $2(X_j^{1/2} - \lambda_j^{1/2})$  has an approximate standard normal distribution. We estimate  $\lambda_j^{j}$  by  $N_j \hat{p} \equiv \hat{\lambda}_j$ . Now  $X_j$ ,  $\hat{\lambda}_j$  are positively correlated since  $\hat{\lambda}_j$  includes X<sub>j</sub>. It can be shown by a Taylor expansion argument that the variance of  $2(X_j^{1/2} - \hat{\lambda}_j^{1/2})$  is approximately  $(1 - N_j/N)$ . If all the N<sub>j</sub>'s are equal, this variance is 1 - 1/J.

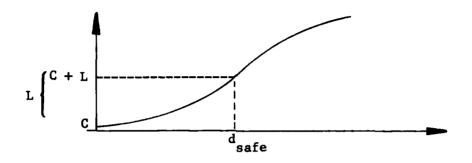
Thus if  $\hat{p}<0.1$  (or if  $\hat{q}<0.1$ ), the quantities  $(1 - N_j/N)^{1/2} \times 2(X_j^{1/2} - (N_j\hat{p})^{1/2})$ , j = 1, ..., J, can be treated as having an approximate standard normal distribution. Graphical and numerical outlier detection procedures are based on these standardized values. We carry out the same types of analyses with these values as with the standardized ratios calculated under case 1. For formal inferences we account approzimately for the correlation among terms within groups (approximately - 1/(J - 1)), as we did in case 1, by treating the J values within each group as if they were J - 1 independent values

# Case 3. Expected response probability is moderate ( i.e. $0.1 \le \hat{p} \le 0.9$ ) but some expected frequencies do not exceed the cutoff (e.g. 5)

We follow the suggestion of Barnett and Lewis and carry out the arcsine variance stabilizing transformation. In particular  $2N_{j}^{1/2}$  [arc sin  $(\hat{p}_{j}^{1/2})$  - arc sin  $(\hat{p}^{1/2})$ ] has an approximate standard normal distribution as  $N_{j} \rightarrow \infty$ . We estimate p by  $\hat{p}$ . It can be shown that the variance of  $2N_{j}^{1/2}$  [arc sin  $(\hat{p}_{j}^{1/2})$  - arc sin  $(\hat{p}^{1/2})$ ] is approximately  $(1 - N_{j}/N)$ . If all the N's are equal this variance is 1 - 1/J. Thus the quantities  $2N_{j}^{1/2}(1 - N_{j}/N)^{-1/2}$  [arc sin  $(\hat{p}_{j}^{1/2})$  - arc sin $(\hat{p}^{1/2})$ ],  $j = 1, \ldots, J$  can be treated as having an approximate standard normal distribution. Graphical and numerical outlier detection procedures are based on these standardized values. We carry out the same types of analyses with these values as with the standardized ratios calculated under case 1. For formal inferences we account approximately for the correlation among terms within groups (approximately - 1/(J - 1)), as we did in the previous cases, by treating the J values within each group as if they were <u>J - 1 independent</u> values.

APPENDIX AXV. CONFIDENCE INTERVAL ON CONCENTRATION THAT CORRESPONDS TO A GIVEN LEVEL OF INCREASE IN RESPONSE OVER CONTROL GROUP RESPONSE.

After we have fitted the nonlinear regression model we wish to calculate confidence bounds on the safe dose. We use Fieller's theorem.



Suppose we're willing to tolerate a response rate L above the control group rate, C.

We want a confidence interval on d safe such that

$$\Phi(\beta_0 + \beta_1 d_{safe}) = L \text{ and } \theta \equiv (\beta_0, \beta_1, c).$$

The standard probit fit assumes that

$$p(\theta; d) = c + (1 - c)\Phi(\beta_0 + \beta_1 d).$$
 (1)

where c is the background rate.

We obtain  $\beta$  by a maximum likelihood fit of the model using a nonlinear regression program or using SAS PROC PROBIT.

We then wish to solve the equation

$$\Phi(\beta_0 + \beta_1 d_{safe}) = L.$$
 (2)

Thus,  $\beta_0 + \beta_1 d_{safe} = \Phi^{-1}(L) \equiv f_L$  (3)

Appendix AXV is the appendix for Section XV.

The point estimate for 
$$d_{safe}$$
 is  $\hat{d}_{safe} = \frac{\phi^{-1}(L) - \hat{\beta}_0}{\hat{\beta}_1}$  (4)

Placing a confidence interval on d<sub>safe</sub> is now a direct application of Fieller's Theorem. See Mandel [44], page 279 or Graybill [45], pages 126-127

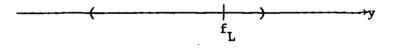
Thus for fixed d, a 1 -  $\alpha$  confidence interval on  $y_d \equiv \beta_0 + \beta_1 d$  is

$$y_{d} \varepsilon \hat{\beta}_{0} + \hat{\beta}_{1} d \pm z_{\alpha/2} \sqrt{g + 2hd + jd^{2}}$$
(5)

where  $\mathbf{g} \equiv \hat{\mathbf{V}}_{ar}(\hat{\boldsymbol{\beta}}_{0}), \mathbf{j} \equiv \mathbf{V}_{ar}(\hat{\boldsymbol{\beta}}_{1}), \mathbf{h} \equiv \mathbf{Cov}(\hat{\boldsymbol{\beta}}_{0}, \hat{\boldsymbol{\beta}}_{1}).$ 

The confidence interval on d includes all d's such that safe

 $f_{L} \hat{\epsilon \beta_{0}} + \hat{\beta}_{1} d + z_{\alpha/2} \sqrt{g + 2h + jd^{2}}$ i.e.  $(\hat{\beta}_{0} + \hat{\beta}_{1} d - f_{L})^{2} \leq z_{\alpha/2}^{2} (g + 2h + jd^{2})$ 



Thus the limits on d are obtained by solving the equations

$$f_{L} = \hat{\beta}_{0} + \hat{\beta}_{1}d + z_{\alpha/2}\sqrt{g + 2hd + jd^{2}}$$
(6)

Thus

$$(\hat{\beta}_0 + \hat{\beta}_1 d - f_L)^2 = z_{\alpha/2}^2 (g + 2hd + jd^2)$$
 (7)

A necessary and sufficient condition for Fieller's Theorem to yield a valid confidence interval for  $d_{safe}$  is that

$$\hat{\beta}_1^2 - j z_{\alpha/2}^2 > 0$$
 (8)

That is,  $\beta_1^2$  must be "many" std error units away from 0. Under the condition in equation (8) a 1 -  $\alpha$  two sided confidence interval on d<sub>safe</sub> is

$$d_{safe} \varepsilon \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}$$
(9)

where A, B, C are

$$A = \hat{\beta}_{1}^{2} - j z_{\alpha/2}^{2}$$
  

$$B = 2 [\hat{\beta}_{1} (\hat{\beta}_{0} - f_{L}) - h z_{\alpha/2}^{2}]$$
(10)  

$$C = [(\hat{\beta}_{0} - f_{L})^{2} - g z_{\alpha/2}^{2}]$$

Suppose now we wish to calculate a  $1 - \alpha$  level <u>lower bound</u> on d safe. This is the value that could be used for regulatory purposes. By an argument similar to the one above, it can be shown that the form of the confidence interval is the same as above, except that  $z_{\alpha/2}$  is replaced by  $z_{\alpha}$  and the smaller root in equation (9) is used. Thus to calculate a lower 95 percent confidence bound on d safe we use the lower end point of a 90 percent two sided confidence interval and d safe, etc. APPENDIX AXVI.1. CONFIDENCE BOUNDS ON BINOMIAL PROBABILITIES

We observe  $\hat{p} = \frac{Y}{n}$   $Y \vee Bi(n, p)$ .

We wish to construct a  $1 - \alpha/2$  sided confidence interval on p.

These limits can be obtained easily using the <u>Clopper-Pearson charts</u>.

See Dixon and Massey [13] pp 501-504.

Tables of such confidence limits are given in Natrella [48].

$$P \{ p \leq p \leq \widetilde{p} \} = 1 - \alpha \quad \text{where}$$

$$p = \left[ 1 + \frac{n - Y + 1}{Y} F(2n - 2Y + 2, 2Y; 1 - \alpha/2) \right]^{-1} \quad p = 0 \text{ if } Y = 0$$

$$p = \left[ 1 + \frac{n - Y}{Y + 1} \frac{1}{F(2Y + 2, 2n - 2Y; 1 - \alpha/2)} \right]^{-1} \quad \widetilde{p} = 1 \text{ if } Y = n$$

Appendix AXVI.1 is an appendix for Section XVI.

APPENDIX AXVI.2. DESCRIPTION AND INSTRUCTIONS FOR USE OF NONPARAMETRIC DOSE RESPONSE ESTIMATION COMPUTER PROGRAM

Sector States

www.wijith withights strateging second of

. . . . . . · · · ·

Appendix AXVI.2 is an appendix for Section XVI.

# A COMPUTER PROGRAM TO CALCULATE NONPARAMETRIC LOWER CONFIDENCE BOUNDS ON SAFE CONCENTRATIONS IN QUANTAL RESPONSE TOXICITY TESTS\*

by

Paul I. Feder and Mary E. Sherrill The Ohio State University

Technical Report No. 215

September 22, 1980

This work was carried out with the support of the U. S. Army Medical Bioengineering Research and Development Laboratory, Frederick, Maryland under Contract DAMD17-79-C-9150 at The Ohio State University.

# I. INTRODUCTION

Dose response experimentation has many applications such as toxicity tests, bioassay, engineering stress tests, tests of response to advertising campaigns, just to name a few. This writeup considers application of dose response experimentation to tests of toxicity on fish and other water species. However the procedures and the computer program discussed are relevant to many other applications.

In aquatic toxicity tests samples of fish or other water species are exposed in tanks or beakers to the substance or substances under study. The exposures are carried out at a succession of concentrations starting at or near zero (i.e. the control groups) and progressing to relatively high and lethal concentrations. Responses of these creatures to the toxicant are recorded and degrees of response are compared across concentration groups.

Many different responses are recorded. Some of these are percent of embryos hatched live or hatched normal, percent of fry surviving for a fixed duration or to the conclusion of the test, body weight or body length of surviving fish, numbers of eggs laid, numbers of eggs hatched. These various responses give rise to several different types of data-quantitative (e.g. body weights), count (e.g. number of eggs laid), quantal response (e.g. hatch/no hatch, live/die, normal/abnormal). The discussion here pertains to quantal responses.

Quantal response toxicity tests often give rise to binomial distribution data. Namely within each test chamber a certain number of organisms are placed on test. Under reasonable assumptions the numbers of "successes" (e.g. numbers of fish per tank that die before the conclusion of the test) follow the binomial distribution. Many standard methods have been proposed over the years to fit models to such binomial dose response data. Finney (1971) and Cox (1970) discuss two of the most commonly used empirical models, namely the probit and logit dc response functions. Background response is commonly accounted for in t models by means of Abbot's correction (see Finney, Chapter 7). Imber of other parametric dose response models have been proposed, based on empirical or on mechanistic considerations. See Kalbfleisch and Prentice (1980), pp. 195-198 for a description of a general family of dose response models.

Determining safe concentrations by inferences based on dose response curves has an important advantage over determinations based on hypothesis tests to compare treatment and control groups. Namely if a particular test is either too small or too variable then a hypothesis test comparing treatment and control group response rates may not be sufficiently powerful. It may thus not be able to detect moderate sized changes in response rate from the control group. This has the effect of <u>raising</u> the estimated "no effect" level, which is unconservative. Such a problem might well arise in the presence of a reasonable sized background response rate. By contrast, decreased sample size or increased variability <u>reduce</u> lower confidence bounds on safe

372

concentrations derived from dose response curve fits. In this sense, inferences about safe concentrations based on dose response curves are more conservative than inferences based on hypothesis tests.

However parametric dose response models have the common drawbacks that

- Inferences about percentiles, especially low percentiles, of the response distribution can be very sensitive to the specific functional form assumed.
- Inference procedures are generally based on asymptotic normal maximum likelihood theory and may thus be inappropriate for data sets with small sample sizes or with many treatment group response rates near 0% or near 100%.
- 3. Results of the high concentration treatment groups, far away from the safe concentrations, have important influence on the estimation of the low percentiles of the dose response curve.
- 4. Background responses are accounted for in a structured parametric manner, such as by Abbott's correction. The form of this background correction may influence the determination of safe concentrations.

The procedure discussed in this writup avoids many of these problems. It does not require assumption of a specific functional form of the dose response curve. It is based on exact small sample distribution theory. It uses information only from the lower half of the dose response curve. It does not require a structured, parametric form of correction for background.

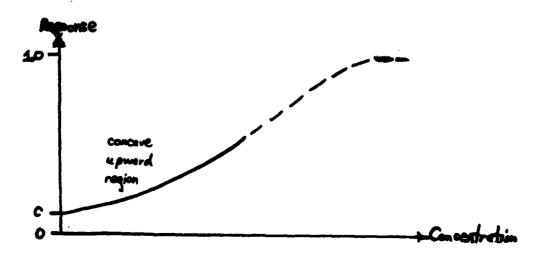
One assumption made throughout this writeup is that no tank to tank heterogeneity is present within treatment groups. This implies that the responses can be pooled across tanks within treatment groups and summarized by a single binomial distribution per group. If tank to tank heterogeneity exists it can be accounted for by adjusting the data to reflect within tank correlation, by fitting more complex models which explicitly account for such heterogeneity, or by carrying out analyses on a per tank basis. The procedure discussed in this writeup can be used in conjunction with the first adjustment approach.

#### II. PROCEDURE

The aim is to calculate a lower confidence bound on the "safe" concentration. The "safe" concentration is defined to be the greatest concentration for which the response rate is at most 100L% above the control group response rate. Note that this does <u>not</u> imply that 100L% response to toxicant is considered to be "acceptable". We wish to eliminate risk altogether. However, we can be confident with such a criterion that at worst we have limited the risk to this level.

The procedure described in this section is a nonparametric approach to determining such a safe concentration. Thus it is not necessary to specify a particular parametric form for the dose-response curve. The procedure was motivated by one discussed in Gross, Fitzhugh, and Mantel (1970) for quantitative response, but differs from it in a number of respects.

Consider a dose response curve relating percent response (e.g. mortality) to toxicant concentration.



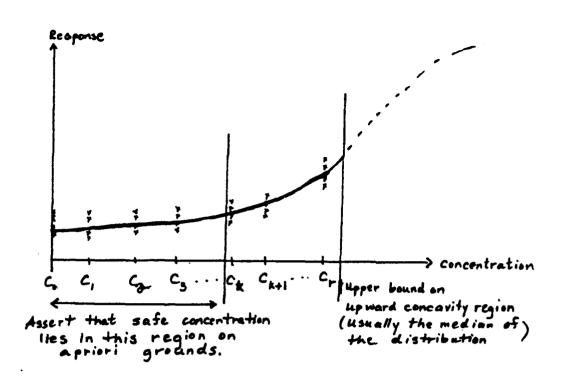
We concentrate on the portion of the curve that is <u>concave upward</u>. In the case of a probit or a logit model this would be all concentrations below the EC50, if there is no background and even beyond the EC50 if there is background response.

Suppose that the toxicity test involves several tanks per concentration group and that we have determined which concentration are in the concave upward portion of the curve (e.g. below the EC50). We might do this by looking at a graph or by a preliminary analysis. Let  $c_0$ ,  $c_1$ ,  $c_2$ ,  $c_3$ , ...,  $c_r$  denote these concentrations. Since we assume the absence of tank to tank heterogeneity within groups, we pool the responses across tanks within groups to obtain the estimated response rates  $\hat{p}_0$  at  $c_0$ (control),  $\hat{p}_1$  at  $c_1$ ,  $\hat{p}_2$  at  $c_2$ ,  $\hat{p}_3$  at  $c_3$ , ...,  $\hat{p}_r$  at  $c_r$ . Let  $p_0$ ,  $p_1$ ,  $p_2$ ,  $p_3$ , ...,  $p_r$  denote the "true" response rates at these concentrations.

We wish to construct a lower confidence bound on the "safe" concentration.

# Definition. A safe concentration is one that increases the response at most L (limit) units above background.

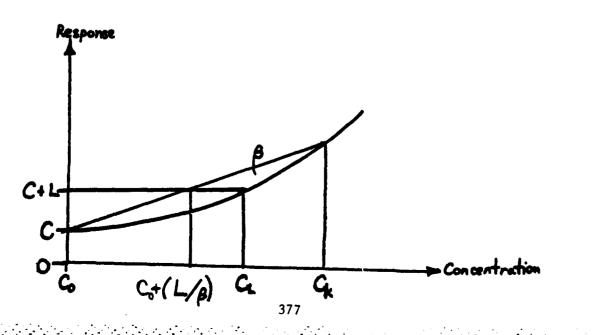
Let  $C_L$  denote this "safe" concentration. Suppose it can be asserted on a priori grounds that  $C_L$  lies in the interval (0,  $C_k$ ) where  $C_1 \leq C_k \leq C_r$ .



We construct upper bounds on the dose response curve. These upper bounds will be used to construct lower bounds on  $C_L$ . First construct a chord joining  $(C_0, p_0)$  and  $(C_k, p_k)$ . Let  $\beta$  denote its slope.

$$\beta = \frac{P_{k} - P_{0}}{C_{k} - C_{0}}$$
(1)

Concavity implies that the chord lies above the dose response curve



throughout the interval ( $C_0$ ,  $C_k$ ). Thus L/ $\beta$  is a lower bound on  $C_L$ .

Let  $\tilde{p}_k$  be an exact upper confidence bound on  $p_k$  and let  $p_0$  be an exact lower confidence bound on p. Expressions for such exact confidence bounds on  $p_0$ ,  $p_k$  were derived by Clopper and Pearson (1934) and are contained in Hollander and Wolfe (1973), pp. 23-24. Set

$$\hat{\beta}_{u} = \frac{P_{k} - Q_{0}}{C_{k} - C_{0}}$$
(2)

Then  $\hat{\beta}_{u}$  is an upper confidence bound on  $\beta$ . Thus  $C_{0} + L/\hat{\beta}_{u}$  is a lower confidence bound on  $L/\beta$  and therefore also on  $C_{L}$ . Since we assume that  $C_{L} \leq C_{k}$ , our final lower confidence bound on  $C_{L}$  is

$$\hat{C}_{L} = \min(C_{0} + L/\hat{\beta}_{u}, C_{k})$$
(3)

It may be possible to increase  $\hat{C}_{L}$  by including information from  $C_{k+1}, c_{k+2}, \ldots, C_{r}$  in addition to that from  $C_{0}, C_{k}$ . Namely, we fit straight lines using  $(C_{0}, C_{k}), (C_{0}, C_{k+1}), (C_{0}, C_{k+2}), \ldots, (C_{0}, C_{r}), (C_{0}, C_{k}, C_{k+1}), (C_{0}, C_{k}, C_{k+2}), \ldots, (C_{0}, C_{k}, C_{r}), (C_{0}, C_{k+1}, C_{k+2}), (C_{0}, C_{k+1}, C_{k+2}), \ldots, (C_{0}, C_{k}, C_{r}), (C_{0}, C_{k+1}, C_{k+2}), (C_{0}, C_{k+1}, C_{k+2}), \ldots, (C_{0}, C_{k}, C_{k+1}), (C_{k+2}), \ldots, (C_{0}, C_{k}, C_{k+1}), (C_{k+2}), \ldots, (C_{0}, C_{k}, C_{k+1}), (C_{k+2}), \ldots, (C_{0}, C_{k}, C_{k+1}), (C_{0}, C_{k+1}, C_{k+2}), \ldots, (C_{0}, C_{k}, C_{k+1}), (C_{k+2}, C_{k+2}), \ldots, (C_{0}, C_{k}, C_{k+1}), C_{k+2}, \ldots, C_{r})$ . That is, we include  $C_{0}$  and all possible combinations of  $(C_{k}, C_{k+1}, \ldots, C_{r})$ . Thus there will be  $2^{(r-k+1)} - 1$  lines calculated altogether. The lines are fitted by ordinary least squares.

Consider the caluclation of an upper bound on the slope of the line based on the subset of concentrations:  $C_0$ ,  $C_{k_1}$ ,  $C_{k_2}$ , ...,  $C_{k_j}$  where  $C_k \leq C_{k_1} \leq C_{k_2} \leq C_{k_j} \leq C_r$ . Denote this as the "s-th subset" (in some ordering of subsets). Define

$$\overline{C}_{s} = (C_{0} + \sum_{j=1}^{J} C_{k_{j}})/(J+1)$$
 (4)

$$\hat{\beta}_{u,s} = \frac{(C_0 - \overline{C}_s)p_0^* + \sum_{j=1}^{J} (C_{k_j} - \overline{C}_s)p_{k_j}^*}{(C_0 - \overline{C}_s)^2 + \sum_{j=1}^{J} (C_{k_j} - \overline{C}_s)^2}$$
(5)

 $P_0^* = P_0$ 

$$p_{j}^{*} = \begin{pmatrix} p_{k_{j}} & \text{if } C_{k_{j}} < \overline{C}_{s} \\ & & & \\ & & & \\ & & & \\ & & & p_{k_{j}} & \text{if } C_{k_{j}} > C_{s} \end{pmatrix}$$
(6)

 $s = 1, 2, ..., 2^{(r-k+1)} - 1$ . We let

$$\hat{\beta}_{u} = \min_{S} \hat{\beta}_{u,S}$$
(7)

$$\hat{C}_{L} = \min(L/\hat{\beta}_{u}, C_{k})$$
 (8)

There is a problem associated with simultaneous inference. We wish to assert with a given level of confidence that  $\hat{\beta}_{u} \geq \beta$ . It can be proved that if  $p_{j} \leq p_{j} \leq \hat{p}_{j}$ , j = 0, k, k+1, k+2, ..., r then  $\hat{\beta}_{u} \geq \beta$ . To guarantee this, with specified probability 1 -  $\alpha$ , we use Bonferroni's inequality. Namely we calculate r-k+2 simultaneous two-sided 1 -  $\alpha$  confidence intervals by calculating each two-sided confidence interval at level 1 -  $\alpha/(r-k+2)$ . Thus, with probability 1 -  $\alpha$ , at least, all the intervals are assured of containing their "true" response rates.

In the case of "large" samples we can calculate alternative upper bounds on  $\beta$  by use of the normal approximation to the binomial distribution. Assume for definiteness that the minimum cell frequency is at least 5. That is,  $N_j \hat{p}_j \ge 5$ ,  $N_j \hat{q}_j \ge 5$ , j = 0, k, k+l, ..., r. (Actually Dixon and Massey (1969), page 238, state a more liberal standard.) Then

$$\hat{\beta}_{s} = \frac{(c_{0} - \overline{c}_{s})\hat{p}_{0} + \sum_{j=1}^{J} (c_{k_{j}} - \overline{c}_{s})\hat{p}_{k_{j}}}{(c_{0} - \overline{c}_{s})^{2} + \sum_{j=1}^{J} (c_{k_{j}} - \overline{c}_{s})^{2}}$$
(9)

is a point estimate of

$$\beta_{0,s} = \frac{(C_0 - \overline{C}_s)p_0 + \sum_{j=1}^{J} (C_{k_j} - \overline{C}_s)p_{k_j}}{(C_0 - \overline{C}_s)^2 + \sum_{j=1}^{J} (C_{k_j} - \overline{C}_s)^2} \ge \beta$$
(10)

and

$$\hat{s.e.}(\hat{\beta}_{s}) \equiv std.err. (\hat{\beta}_{s}) = \left[ \frac{(C_{0} - \overline{C}_{s})^{2} \frac{\hat{p}_{0}\hat{q}_{0}}{N_{0}} + \int_{j=1}^{J} (C_{k_{j}} - \overline{C}_{s})^{2} \frac{\hat{p}_{k_{j}}\hat{q}_{k_{j}}}{N_{j}}}{[(C_{0} - \overline{C}_{s})^{2} + \int_{j=1}^{J} (C_{k_{j}} - \overline{C}_{s})^{2}]^{2}} \right]^{1/2}$$
(11)

Thus

$$\hat{\beta}_{u,s} \equiv \hat{\beta}_{s} + z_{1-\alpha} \hat{s.e.} (\hat{\beta}_{s})$$
 (12)

is an upper 1 -  $\alpha$  level normal theory confidence bound on  $\beta$ .

To obtain simultaneous upper confidence intervals we use Bonferroni's inequality for family size  $2^{(r-k+1)} - 1$ ; i.e., individual intervals at level  $1 - \alpha/(2^{r-k+1} - 1)$ . As before

$$\hat{\beta}_{u} = \min_{S} \hat{\beta}_{u,s}$$
(13)

$$\hat{C}_{L} = \min(L/\hat{\beta}, C_{k})$$
 (14)

# III. INSTRUCTIONS FOR PROGRAM INPUT

Mortality or Abnormality Data:

In this program, Subroutine RDSURV handles mortality data input. The mortality or abnormality data is read from <u>Fortran file 4</u>. Each input card of the mortality data must supply the program with the following two numbers (in this order):

- The number of organisms tested in each replicate of each treatment, and
- the number of organisms in each replicate surviving the test (or surviving normally).

The data must be inputted according to increasing order of treatments.

<u>Important</u>. The user supplies the input format for the card image of the survival data (see input card 4 in Fortran file 5) with variable format statements.

# Concentration (Dose) Data:

In this program, Subroutine RDCONC handles concentration data input. The concentration data is read from <u>Fortran file 9</u>. Each input card from this file must supply the program with the following two numbers (in this order):

- 1. The treatment number, and
- the concentration measurement corresponding to this treatment number.

# Notes to the user:

- 1. The data must be inputted in increasing order of treatments.
- 2. The user may input many concentration measurements for a
  - given treatment since the program calculates an <u>average</u> concentration for each treatment.
- <u>Important</u>. The input format for the card image of the concentration data is user-supplied with variable format statements. (See input card 6 in Fortran file 5).

• IDENTIFICATION SIGNENCE • . LADERCISO MUMBER ••• ----i ž \_ • • • • : -. 13 PS 11 15 15 76 0P 14P 1 ---• • • -. 3 3 CLAPHIC PLINCH -- -\_ \_ . -. . . .... 1 L. Con HILLS Husbord, Husburg FORTRAN STATEMENT • 0861 53 ----•------UAS DATA 12. # -... FILE FRY -2 -INPUT 0 1 ļ 0 14 1 - 10 0.1 · · · · · · · 4 CARDS ON ANALYSIS a proved to fur how how state -2 2 0.05 5X 2 F5.0 ره. + 1 1 1 1 \_ :: = 1 T30 F12 2 Physics 10015 Ś --EXAMPLE ļ -0 ---. ᆋ e 3 . 114 • BENOI (116) 0 STAIgard MT 12 Public and IBM POORM -Bur.,  $\mathcal{O} \odot \mathcal{O} \odot \mathcal{O} \odot \mathcal{O} \odot \mathcal{O}$ 

Example of Program Input Stream. Figure l.

1

·\_•

where they be used ž 10-01 GX28-7327-6 U/M 050\* Primed in U.S.A.

FORTRAN Coding Form

. . . . .

/

~

USER'S JOB CARD REGION=206K.MSULEVEL=(1.1) FJUBPAKA LINES=2000.DISKIU=2500 EXEL FUREXLU, PARM. - JRT= "UPT=2" // FukTestaliv ت÷ رن PROGRAM DECK IS PLACED HERE 1/GU.SYSLID Ji. 11 UU UISP=SHR+DSN=SYS1+IMSL+OUUBLE ---------- DU -UISHESHROSMESYSLOIMSLOSINULE // Ju + Tu5+001 Du + BENUITI - ANALYSIS OF FRY DATA 5 Information From FORTRAN (110+15X+2+5+3) Coding form in Figure 1. 112, 130, +12.4) 15.0 0.05 1 3. J. 4 . 5. . . - . . . . 0.01 0.30 0.10 3 14 -After Tuter Gil ... UL (INVESTIGATOR: DUANE BENJIT (01). TEST: A (01) five description DATA FRUM EARLY LIFE STAGE TESTS WITH FATHEAD MINNOWS lines SIA LEVELS - ILCUNTRULA + 4 LONESTLA .... + CHIGHESTL: 4 REPS EA LA.B. C.D. " EMBNYUS TESTED, & ALIVE AFTER HATCH, & NORMAL FRY AFTER HATCH 1 10101 26 25 15 15 15 36 الك 1 10101 كان 20 كان 15 15 15 <u>1 10101 بال 10</u>20 <u>15 15 15 15</u> Mortality deta set The two columns with \*'s at the top are 21 15 15 22 15 2 A 1 10131 30 30 21 15 14 L D L LÜICL - 21 14 <u>-30 -21 - 21 - 15 - 15 - 15 - 15 - </u> 2 0 1 10101 30 lo lo 15 15 15 17 ไอ 15 3 A 1 10101 30 14 14 these being used in this example, as specified in -1-01-04 -----30 22 21 15 15 15 1 10101 با د 21 15 15 JU 22 15 T TOTOT Cond #4 in Figure 1. -16+61 ----13 15 22 15 1 10101 30 13 11 11 1 10101 22 30 14 14 4 L 4 U 1 10101 -14 14 20 20 15 19 ۔ڏ 0a J W T TOTOT 15 30 17 17 00 D D I IGICI 06 5 6 1 10101 -- 30 - - 17 - -- 10 - 15 - 39 09 24 10 15 7 C 101 JL 24 19 04 30 09 ĴЗ 30 6 A 1 10101 08 19 30 20 0 0 1 10101 15 I <u>^3</u> 03 , 30 25 25 15 0 L I 10101 12 12 15 0 0 X YJYCY 35 21 21 10 10

| SIX LEVELS: 1   | CUNTRULI +2(L)                        |                        | 1; 4 KEPS FA (A. 5. (. 0) Cords a |  |  |
|---|---------------------------------------|------------------------|-----------------------------------|--|--|
| SEE JAIA SMEET FUR MUTLS<br>MEASUREU LÜNLENIRATIONS OF TOXICANTE TOTORO U/E) deta set |                                       |                        |                                   |  |  |
| T W D TOTAT   | 20111                                 |                        |                                   |  |  |
| 1 A 3 20101   | 062079                                | 0•03                   |                                   |  |  |
|   | · · · · · · · · · · · · · · · · · · · | <b></b> . <b>.</b>     |                                   |  |  |
|   |                                       |                        |                                   |  |  |
| 10105 6 A 1   | 070979                                | 0.15                   |                                   |  |  |
| 1 0 3 10101   | 00019<br>( 810                        | 0.00                   |                                   |  |  |
|   | J7J217                                | <u></u><br>0•10        |                                   |  |  |
| 1 0 5 50101<br>1 0 5 10101  | Jous 79                               | 0.07                   |                                   |  |  |
| المكاملات ما الم  |                                       | <u> </u>               |                                   |  |  |
| 1 6 3 30101   | J76577                                | 0+13                   |                                   |  |  |
| 1 0 3 10101   | 361419                                | 0.05                   |                                   |  |  |
| -1 U A 2010 -   |                                       | 0.10                   |                                   |  |  |
| 2 4 3 10101<br>2 A 3 20101  | J61179<br>J62579                      | 1.69                   |                                   |  |  |
| - 2 A 5 20101   | <u> </u>                              |                        |                                   |  |  |
| 2 2 3 10101   | 0005/9                                | 1.33                   |                                   |  |  |
| 2 0 2 20101   | J51419                                | 1.69                   |                                   |  |  |
| - 2020-20202  |                                       |                        |                                   |  |  |
| 2 3 3 401CL   | 575279                                | 1.59                   |                                   |  |  |
| د ت 10101<br>10101  | 060019<br><u>- 061879</u>             | 1.70                   | Concentration data set.           |  |  |
| 2 L 3 30101   | 062179                                | 1.52                   | •••                               |  |  |
| 2 6 3 40101   | 376579                                | 2.15                   | The two columns with              |  |  |
| . 2. U. S. 19194  |                                       | 1                      |                                   |  |  |
| 200520101   | J02319                                | 1.34                   | *'s at the top are those          |  |  |
| 10101 C A C<br>-10102 C-A C-  | Joil19                                | 3.33                   | being used in this example,       |  |  |
|   |                                       | 4.07                   | very user in two examples         |  |  |
| 2 2 2 1 J L L L   | 160019                                | 3 • 26                 | as specified in cord #6 in        |  |  |
| المفتقع داهد.   | _ 16147.9                             |                        |                                   |  |  |
| 10106 6 8 6   | 401219                                | 2.51                   | Figure 1.                         |  |  |
| 40101 د م د<br><u>12101 د ب د</u>   | J70279<br>36J319                      | 3.32<br>2.05           | J -                               |  |  |
| <u>ــــــــــــــــــــــــــــــــــــ</u>   | J61379                                | 2.35                   |                                   |  |  |
|   | 362179                                | 2.72                   |                                   |  |  |
| 3 4 3 40101   | UTU519                                | 3.51                   |                                   |  |  |
| 10151 C U C   | JO1414                                | 3.79                   |                                   |  |  |
| 2 0 2 2022 C  | -62819                                | 44 <b>د د</b><br>د ۲ م |                                   |  |  |
| 4 A J L J L J L J L J L J L J L J L J L J   | <u>603/9</u><br>361179                | 5.647                  |                                   |  |  |
| 4 A 3 30131   | JO2517                                | 5.57                   |                                   |  |  |
| 4 A 2 40101   | 176979                                | 7.32                   |                                   |  |  |
| + 0 J 10101   | 100017                                | 1.05                   |                                   |  |  |
| 4 0 3 20101   | 060519                                | 0.d3                   |                                   |  |  |
|   | <u>1611 19</u>                        | <u> </u>               |                                   |  |  |
| 4 0 3 40101<br>1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0                                    | 081379<br>082379                      | 0.07                   |                                   |  |  |
| 4 0 3 00101   | u70279                                | 6.49                   |                                   |  |  |
| 4 L 3 /JLJL   | J70579                                | 7.92                   |                                   |  |  |
| 4 0 5 00101   | 276719                                | 7•8C                   |                                   |  |  |
| 4 6 9 10101   | 060479                                | 5.06                   |                                   |  |  |
|   |                                       | 386                    |                                   |  |  |

| •           | -<br>L    | -     | 20101          | 561179           |                    |
|-------------|-----------|-------|----------------|------------------|--------------------|
|             |           |       | 70101          | 52119            | 5.21<br>5.23       |
|             |           |       |                | 552119           | 5.14               |
|             |           |       | 40101          | JO2179<br>JO2579 |                    |
|             |           |       | 20101<br>20101 | JBL577<br>J7J279 | 5 • 4 4<br>5 • 9 5 |
|             |           |       | 10101          |                  | 7.02               |
|             |           |       | 10101          |                  | 7.50               |
| 4           |           |       |                | 361179           |                    |
| 4           |           |       | 10101<br>20101 |                  | c • 12             |
|             |           |       | 30101          |                  |                    |
|             |           |       | 43101          | 062079           | 7.76               |
|             |           |       | - 40101<br>    |                  | (• / 7<br>         |
|             |           |       | 10101          |                  | 13.40              |
|             |           |       | 20101          |                  | 12.00              |
| 9           | -         | J     | 201 71         | 002019           | 12.00              |
| · · · · · · |           |       |                |                  |                    |
|             |           |       |                |                  |                    |
| ۲<br>۲      | <br>A     | <br>3 | 36,01          | j76919           | 10.0?              |
|             |           |       |                | 060019           | 17.20              |
|             |           |       |                |                  | 1-0-30             |
|             |           |       |                | J70279           | 11.00              |
|             |           |       |                | 060379           | 10.60              |
|             |           |       |                |                  |                    |
|             |           |       | 30101          |                  | 13.10              |
|             |           |       |                | 361479           | 13.00              |
|             |           |       |                |                  |                    |
|             |           |       |                | 061179           | 24.10              |
|             |           |       |                | 202579           | 23.30              |
|             |           |       |                |                  | 30.00              |
|             |           |       | 10101          | 565579           | 08.65              |
|             |           |       | 20101          |                  | 28.30              |
|             |           |       |                | <u></u>          | 32.70              |
|             |           |       | 10101          |                  | 24.20              |
|             |           |       | 20101          |                  | 24.20 -            |
|             |           |       | 20121          |                  |                    |
| 6           |           |       | 10101          | 061479           | 21.20              |
| ь           |           |       | c0101          | J62319           | 23.30              |
|             | <b></b> _ |       |                |                  | · ·                |
| 11          |           |       |                |                  |                    |
|             |           |       |                |                  | •                  |
|             |           |       |                |                  |                    |
|             |           |       |                |                  |                    |
|             |           |       |                |                  |                    |
|             |           |       |                |                  |                    |
|             |           |       |                |                  |                    |
|             |           |       |                |                  |                    |
|             |           |       |                |                  |                    |



# DATA INPUT

# Values Used by the Program:

The input cards are read from <u>Fortran file 5</u>. There <u>must</u> be 9 input cards in this file. An example input stream is illustrated in Figure 1 and a sample deck suitable for card input on the AMDAHL 470 V6 computer is illustrated in Figure 2. The following descriptions refer to the numbered cards in Figure 1.

<u>Card #1</u> is the title card. The user is to choose a title or message and place it anywhere in the first 72 columns of this card. If the user does not wish to supply a title, this first card must be blank.

> A typical title is illustrated on the Fortran Coding Form, line l.

<u>Card #2</u> contains the number of treatments and the number of replicate tanks within each treatment. Suppose N1 = number of treatments and N2 = number of replicates. N1 and N2 must be <u>integer</u> values.

> If  $1 \le N1 \le 9$ , place N1 in column 2 of this card. If N1 = 10, place N1 in columns 1 and 2 of the card. If  $1 \le N2 \le 9$ , place N2 in column 4. If N2 = a two-digit number, place N2 in columns 3 and 4.

(See program limitation 1.)

An example of this card is found on Fortran Coding Form, line 2. The numbers '6' and '4' indicate that there are 6 treatments with 4 replicates in each treatment.

- <u>Card # 3</u> supplies the program with the number of descriptive cards (lines) at the beginning of the mortality dataset. (A descriptive card may be any card at the beginning of the dataset which is not data; e.g., a typical card might contain information about the contents of the dataset.) The number on this card must be an <u>integer</u> from 0 to 99. Suppose NS = # of descriptive cards. If  $0 \le NS \le 9$ , place NS in column 2 of this card. If  $10 \le NS \le 99$ , place NS in columns 1 and 2. Note: If the user has no descriptive cards in the mortality dataset, place a '0' in column 2 of the card. An example of this card is found on the Fortran Coding Form, line 3. The number '5' indicates that there are 5 descriptive cards at the head of the mortality dataset.
- <u>Card #4</u> contains the format for the input line image of the mortality data. This format must supply the program with the following two values (in <u>this</u> order): The number of objects tested within a given replicate of a given treatment, and the number of these objects surviving the test (or surviving normally). Line 4 on the Fortran Coding Form contains a typical variable format statement. This format indicates to the program to tabulate to column 16 of the card, skip 15 spaces, and read the next 2 values from the card.

- <u>Card #5</u> contains the number of descriptive cards at the head of the concentration dataset. This card is similar to Card # 3. Suppose NC = number of descriptive cards. If  $0 \le NC \le 9$ , place NC in column 2. If  $10 \le NC \le 99$ , place NC in columns 1 and 2. An example of this card is line 5 on the Fortran Coding Form. The number '5' indicates there are 5 descriptive cards at the head of the concentration dataset.
- <u>Card #6</u> contains the format for the input image of the concentration (dose) data. This card is similar to Card #4. The format of this card must supply the program with the following two numbers (in <u>this</u> order): an <u>integer</u> value for the treatment number, and a <u>real</u> value for the concentration measurement within that treatment.

Note line 6 on the Fortran Coding Form. This variable format instructs the program to read an integer value from the first 2 columns, tabulate to column 30, and read the real value beginning in that column.

Card # 7 contains the user's choices for 3 parameter values:

- 1. the upper bound on the upward concavity region (UCR),
- 2. the level of significance, or alpha level (ALEVEL),
- 3. the value of a flag indicating the user's desire for simultaneous confidence intervals. (IFLG1)

0tc

The user should place UCR in columns 1-8 with a decimal point in column 4. This allows the user to specify a number as large as 999.9999. ALEVEL is placed in columns 11-16 with a decimal point in column 12. Note that this value must be a number between 0 and 1. If the user desires simultaneous confidence intervals, the value of IFLG1 must be 1. Place a 'l' in column 19, if this is the case; otherwise, leave the column blank. Note line 7 on the Fortran Coding Form. The number '15.0' indicates that the upward concavity region lies between 0 and 15.0 (0 and 15.0 are possible concentration values.) The number '0.05' indicates the level of significance, and the 'l' in column 19 tells the program that a simultaneity adjustm is desired. (Thus 0.05 is interpreted as the familywise error rat

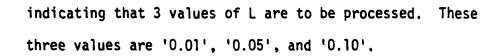
<u>Card #8</u> contains the number of values of k to be considered in the analysis, followed by the actual k values. (Recall that the safe concentration is assumed to lie between  $C_0$  and  $C_k$ .) The values given on this card must be <u>integers</u>. Suppose NK = number of k's to be considered. If  $1 \le NK \le 9$ , place NK in column 2 of this card. If NK = 10, place NK in columns 1 and 2. Leave columns 3 and 4 blank. Columns 5-24 contain the actual values of k. Each value is allotted 2 columns; e.g., columns 5 and 6 contain the first value of k, and columns 7 and 8 contain the second value, etc.

If the value for k is a treatment number from 1 to 9, place the value in the rightmost column of the field. Otherwise, the value for k would equal 10 and both columns are used for this two-digit number.

Note: NK <u>must</u> equal the number of actual values placed on this card.

Line 8 on the Fortran Coding Form is a typical example of Card Number 8. Column 2 contains the number '3', indicating that 3 values for k are to be considered. The numbers '3', '4', '5', are these k values.

Card #9 contains the number of L's to be considered followed by the actual values for L. (Recall that L represents the incremental response rate over background, associated with the "safe" concentration.) The values on this last card are as follows: Columns 1 and 2 contain an integer value for the number of L's. Suppose NL = number of L's. If  $1 \le NL \le 9$ , place NL in column 2. If NL = 10, place NL in columns 1 and 2. Columns 3 and 4 are to be left blank. Beginning with column 5, the remaining columns are used to specify the desired values of L. Each value of L uses 6 columns; e.g., the first value appears in columns 5-10 with a decimal point in column 6. A decimal point must be placed in the second column of a given field. Note that L must be a number from 0 to 1. The user is allowed at most 4 digits to the right of the decimal point. The example of Card Number 9 is found on the Fortran Coding Form, line 9. A '3' appears in column 2



(

# PROGRAM LIMITATIONS

# Array Size Limitations:

- The number of treatments may not exceed 10 and (the number of treatments) X (the number of replicates per treatment) may not exceed 40. For example, the program could handle as many as 10 treatments with 4 replicates each, or 8 treatments with 5 replicates each.
- 2. The user can supply this program with no more than 10 values for k (input card 8 in Fortran file 5). Similarly, the user is allowed no more than 10 values for L (input card 9 in Fortran file 5).

# Output Limitations:

- Only one title card is allowed and the user must restrict his title to the first 72 columns of the title card. (See input card 1. in Fortran file 5.)
- 2. Subroutine WRITE1 prints the data summary found on the first page of output. The number of objects tested and the number of survivals are each printed in format F6.0. Therefore, any quantity greater than 99999. will <u>not</u> print correctly. Similarly, the concentration for each treatment is printed in F8.4 format. Numbers greater than 999.9999 will not print correctly.

# DISTRIBUTION LIST

| 25 copies | Commander<br>US Army Medical Bioengineering<br>Research and Development Laboratory<br>ATTN: SGRD-UBG<br>Fort Detrick, Frederick, MD 21701        |
|-----------|--|
| 4 copies  | Commander<br>US Army Medical Research and Development<br>Command<br>ATTN: SGRD-RMS<br>Fort Detrick, Frederick, MD 21701                          |
| 12 copies | Defense Technical Information Center (DTIC)<br>ATTN: DTIC-DDA<br>Cameron Station<br>Alexandria, VA 22314   |
| 1 сору    | Dean<br>School of Medicine<br>Uniformed Services University of the<br>Health Sciences<br>4301 Jones Bridge Road<br>Bethesda, MD 20014            |
| 1 сору    | Commandant<br>Academy of Health Sciences, US Army<br>ATTN: AHS-CDM<br>Fort Sam Houston, TX 78234   |
| 1 copy    | Commander<br>US Army Medical Bioengineering Resear<br>and Development Laboratory<br>ATTN: SGRD-UBD-A/Librarian<br>Fort Detrick, Frederick, MD 21 |



Sector And Party 

