

FROME, Edward Lee, 1942-

NONLINEAR REGRESSION AND SPECTRAL ESTIMATION  
IN BIOMEDICAL DATA ANALYSIS.

Emory University, Ph.D., 1972  
Statistics

University Microfilms, A XEROX Company, Ann Arbor, Michigan

® 1972

Edward Lee Frome

ALL RIGHTS RESERVED

PLEASE NOTE

Some pages may have indistinct print.

Filmed as received

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

NONLINEAR REGRESSION AND SPECTRAL ESTIMATION  
IN BIOMEDICAL DATA ANALYSIS

By

Edward Lee Frome  
B.S., University of Florida, 1964  
M.S., University of Florida, 1966

A Dissertation submitted to the Faculty of the Graduate School  
of Emory University in partial fulfillment  
of the requirements of the degree of  
Doctor of Philosophy

1972

NONLINEAR REGRESSION AND SPECTRAL ESTIMATION  
IN BIOMEDICAL DATA ANALYSIS

Approved for the Department:

Michael H. Kutner  
Adviser

July 25, 1972  
Date

Accepted:

Chas. T. Lester  
Dean of the Graduate School

July 25, 1972  
Date

In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the Dean of the Graduate School when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

Edward J. From

**NOTICE TO BORROWERS**

Unpublished theses deposited in the Emory University Library must be used only in accordance with the stipulations prescribed by the author in the preceding statement.

The author of this dissertation is:

NAME Edward Lee Frome  
ADDRESS 938 Lullwater Rd. NE  
Atlanta, Ga. 30307

The director of this dissertation is:

NAME Michael H. Kutner  
ADDRESS 2572 Circlewood Rd., NE  
Atlanta, Ga. 30345

Users of this dissertation not regularly enrolled as students at Emory University are required to attest acceptance of the preceding stipulations by signing below. Libraries borrowing this dissertation for the use of their patrons are required to see that each user records here the information requested.

Name of user	Address	Date	Type of use (Examination only or copying)
--------------	---------	------	--

NONLINEAR REGRESSION AND SPECTRAL ESTIMATION  
IN BIOMEDICAL DATA ANALYSIS

By

Edward Lee Frome  
B.S., University of Florida, 1964  
M.S., University of Florida, 1966

Michael H. Kutner, Adviser

An Abstract of  
A Dissertation submitted to the Faculty of the Graduate School  
of Emory University in partial fulfillment  
of the requirements of the degree of  
Doctor of Philosophy

1972

## ABSTRACT

In this dissertation we consider the application of statistical theory and methods to two general problems in biomedical data analysis. First, we investigate the application of regression techniques when the observations are counts. In the second problem we propose the use of spectral analysis to deal with certain time series data that occur in the study of respiration. It is assumed throughout that a digital computer will serve as the primary instrument for data analysis, and in the time series problems a computer is also essential for data acquisition.

In Chapter 2 a general regression model is defined. Assuming that the observations are counts that follow the Poisson distribution, the maximum likelihood principle is employed to estimate the unknown parameters in the regression model. A procedure for testing the assumed Poisson variation and 'goodness of fit' of the model is developed. In Chapter 3 we demonstrate that weighted least squares estimates obtained via the Gauss-Newton iterative procedure are identical to the maximum likelihood estimates —which are obtained using the method of scoring. A minimum chi-square estimation procedure is then considered as a method for obtaining best asymptotically normal estimates. We show that when the observations are counts that follow the Poisson distribution, then the maximum likelihood estimates, the least squares estimates, and the modified minimum chi-square estimates are all identical. In Chapter 4 the estimation procedure is applied

to data from biological experiments where the observations are colony counts. The computational procedure is summarized in Section 4.4 and a Fortran IV implementation is provided in the Appendix.

In Chapter 5 we review the statistical considerations of numerical spectrum analysis. Spectral estimation is based on the periodogram —which is computed from the finite Fourier transform of a realization of a time series— and frequency domain smoothing techniques. The computational procedures that we use are summarized in Section 5.3. In Chapter 6 a simple mechanical model of the respiratory system is developed. Cross-spectral analysis is proposed as a data analytic technique that can be used to describe the respiratory system, and an example using pressure-flow time series data is provided. We claim that when the flow-resistive component of the transpulmonary pressure is linearly related to air flow, then the spectrum —which is calculated from the flow measurements— will truly be a power spectrum. In this situation the flow variance provides a single statistic that may be used to describe the mechanical work of breathing.

The impedance pneumograph is a bioimpedance recorder that is used for the indirect measurement of respiration. In Chapter 7 we propose spectral analysis of the pulmonary impedance pneumograph and illustrate with an example how spectra calculated from the impedance tracings reflect changes in the depth, regularity, and frequency of breathing. We then point out certain inadequacies of this approach and propose a new method of analysis which uses the first difference of the pulmonary impedance measurements. We claim that this new method results in a

more meaningful measure of respiratory function. The power concept of Chapter 6 is used to define conditions under which the spectrum calculated from the first difference data will be a power spectrum. An example —which graphically demonstrates the effect of the differencing operation on low frequency 'noise'— is presented.

## ACKNOWLEDGEMENTS

I am grateful to Doctors M. H. Kutner, E. L. Frederickson, and C. D. Stevens for their advise and assistance, and for the improvements that they have suggested during the preparation of this dissertation. I am particularly indebted to Dr. C. C. Lushbaugh for the encouragement and the counsel that he has provided throughout the course of this investigation. The results presented in Chapters 2 through 4 undoubtedly reflect the influence of many stimulating conversations with Doctors J. J. Beauchamp and F. V. Comas. Chapter 7 was made possible through the assistance of Dr. R. C. Ricks and the facilities at the Medical Division, Oak Ridge Associated Universities, Inc., Oak Ridge, Tennessee.

This investigation was supported by a National Institutes of Health Special Fellowship (5 F03 GM 45025 from General Medical Sciences) which was sponsored by Doctors M. E. Turner, Jr. and E. C. Hall. I am grateful to Dr. W. B. Evans and the Emory University Computing Center staff (B. A. Blumenstein in particular) for their contribution to the computational aspects of this study.

I am especially indebted to my parents and to all of my family for many years of tolerant encouragement that made this effort possible, and in particular to Ann for the patient assistance that she has contributed to its final realization.

## CONTENTS

	Page
ACKNOWLEDGEMENTS.....	ii
TABLES.....	v
FIGURES.....	vi
Chapter	
1. INTRODUCTION.....	1
1.1 Regression Analysis	2
1.2 Spectral Analysis	5
2. POISSON COUNT DATA: MAXIMUM LIKELIHOOD ESTIMATION.....	9
2.1 One Parameter Model	9
2.2 Specification of the General Model	10
2.3 ML Estimates for General Regression Model	12
2.4 ML Estimates for Some Special Cases	16
2.4.1 Multiple Linear Regression	17
2.4.2 Exponential Survival Curve	18
2.4.3 Target Survival Curve	19
2.4.4 Weibull Survival Curve	20
2.5 Covariance Matrix and Chi-Square tests	21
3. LEAST SQUARES.....	24
3.1 Least Squares - Constant Variance Case	25
3.2 Weighted Least Squares	28
3.3 Best Asymptotically Normal Estimates	32
3.4 Minimum Chi-Square Estimation	36
4. REGRESSION ANALYSIS WHEN THE DATA ARE COUNTS.....	40
4.1 Applications in Biology	41
4.2 Linear Regression	42
4.3 Nonlinear Regression - Survival Curve Analysis	44
4.3.1 Exponential Survival Curve	47
4.3.2 Target Survival Curve	49
4.3.3 Weibull Survival Curve	54
4.4 Computations	56

	Page
5. SPECTRUM ANALYSIS.....	63
5.1 Spectrum Analysis of a Bivariate Stationary Time Series	65
5.2 Spectral Estimation	69
5.3 Computations	74
6. POWER SPECTRUM OF THE RESPIRATORY SYSTEM.....	81
6.1 Mechanical Model of the Respiratory System	84
6.2 Spectral Analysis of the Respiratory System	87
7. SPECTRAL ANALYSIS OF THE PULMONARY IMPEDANCE PNEUMOGRAPH.....	96
7.1 Pulmonary Impedance Pneumograph Spectrum	97
7.2 PIP First Difference Spectrum	99
APPENDIX.....	104
REFERENCES.....	109

## TABLES

Table		Page
4.1	Experiments That Give Rise to Counts	42
4.2	Plaque Counts	46
4.3	Exponential Survival Curve Data	48
4.4	Chi-Square Values for Exponential Model	48
4.5	Target Survival Curve Data	52
4.6	Chi-Square Values for Target Model	57
4.7	Observed and Expected Values	57

## FIGURES

Figure		Page
4.1	Linear Regression	45
4.2	Exponential Model	50
4.3	Initial Estimates - Target Model	53
4.4	Initial Estimates - Weibull Model	53
4.5	Target and Weibull Models	58
6.1	Model of the Respiratory System	84
6.2	Time Series Plots of Pressure and Flow Data	91
6.3	Pressure-Flow Covariance	92
6.4	Pressure-Flow Spectra	94
6.5	Pressure-Flow Cross-Spectrum	95
7.1	Pulmonary Impedance Tracings and Spectra Before and 15 Minutes After Exposure to 30 R (1.5R/min) on Five Consecutive Days	98
7.2	PIP Spectrum and First Difference Spectrum	103

## 1. INTRODUCTION

The subject of data analysis —especially its relation to statistics (mathematical) and its future— has been investigated by Tukey (1962). He stated that 'theory' will have to guide rather than command data analysis - if it is to be successful - since much of what is done is a matter of judgment. Tukey and Wilk (1966) have observed that data analysis is similar to experimentation and pointed out some common characteristics - emphasizing the interactive, iterative, open-ended nature of the data analysis process. They use the term 'data analysis' to "encompass the techniques, attitudes, interests and concepts which are relevant to the process of learning from organized records of experience". In this dissertation we shall be concerned with the use of statistical theory and methods in data analysis. Two general problems will be considered. The first involves the applications of regression techniques to the analysis of data when the observations are counts. In the second problem we will propose the use of spectral analysis to deal with certain time series data that occur in the study of respiration. It will be assumed throughout that the computations which will be required are to be carried out by a digital computer. Thus we will rely upon the computer as the primary instrument for data analysis —see Muller (1970). This reliance is necessary since we will require extensive computation. In the time series problems we will encounter large volumes of data and the computer will be essential for data acquisition as well as data analysis.

### 1.1 Regression Analysis

Consider the general regression model

$$E(y_{ij}) = f(X_i, \theta) \quad (1.1.1)$$

where  $X_i = (x_{i1}, \dots, x_{im})$  is the  $i$ th set of values of the  $m$  independent variables,  $n_i$  is the number of replications of the  $i$ th experimental condition,  $\theta = (\theta_1, \dots, \theta_p)'$  is a  $p$  dimensional vector of unknown parameters, and  $y_{ij}$ ,  $i = 1, \dots, N$ ,  $j = 1, \dots, n_i$ , is a particular realization of the experiment. The regression function,  $f(X_i, \theta)$ , relates the expected value of the dependent variable to the independent variables and the parameters, and given the experimental conditions and the data we would like to estimate the unknown parameters. The most widely used methods of estimation are maximum likelihood and least squares. A discussion of the assumptions underlying these principles of estimation and the properties of the estimators can be found in Kendall and Stuart's The Advanced Theory of Statistics II.

It is well known that maximum likelihood and least squares estimates are identical when the observations are independent and normally distributed with expectation given by equation (1.1.1) and constant variance. If  $f(X, \theta)$  is linear in the parameters then the estimates are obtained using linear regression analysis. When  $f(X, \theta)$  is not linear in all of the parameters estimation is more difficult, and some type of iterative procedure will usually be required. An introductory discussion of nonlinear regression and the extent to which standard results from linear estimation are applicable has been given by Draper and Smith (1966, chap. 10). The equivalence of maximum

likelihood and least squares methods was demonstrated by Turner, Monroe, and Lucas (1961) for a wide class of nonlinear models when the errors are assumed to be normally distributed and independent with constant variance or variance that depends on the independent variable in a known way. Beauchamp (1966) has presented an extensive discussion of generalized least squares estimation, and properties of **the estimators** obtained using a weighted nonlinear regression procedure. He showed that under normality assumptions with unknown variances weighted least squares and maximum likelihood estimates will be equivalent when the weights are obtained from consistent estimates of the variance-covariance matrix.

Another situation in which maximum likelihood and least squares estimates are found to be equivalent is in the analysis of 'quantal' response data, such as probit analysis. In this situation the observations are assumed to be independent, and to follow the binomial distribution with expected value given by the regression model. Moore and Zeigler (1967) have shown that maximum likelihood estimates are identical to those obtained in a properly weighted least squares analysis in the situation just described. An extension of this result to include a class of models whose expectations are given by inverse polynomials has been dealt with by Nelder (1968).

In Chapters 2 and 3 of this dissertation a general approach to regression analysis —which is applicable when the observed values of the random variable are counts— will be developed. In Chapter 2 the Poisson distribution will be proposed as an alternative to the

normal distribution when the observations are discrete counts rather than continuous measurements. A general regression model will be defined and the maximum likelihood principle will be used to estimate the parameters in the model. For all but the simplest case considered we will find it necessary to use the method of scoring to maximize the logarithm of the likelihood function with respect to the parameters. Methods for testing the assumed Poisson variation and 'goodness of fit' of the model will be developed. Special cases of the general model that have been previously investigated—i.e., linear regression and certain nonlinear models—will be reviewed. In Chapter 4 we will apply the maximum likelihood estimation procedure to data from biological experiments where the observations are colony counts. A particular application that will be developed is the analysis of certain survival curve models.

In Chapter 3 we will review least squares estimation and show that weighted least estimates will be identical to the maximum likelihood estimates when the weights are defined in accord with the assumption of Poisson variation. That is, we will show that when the method of scoring is used to obtain the maximum likelihood estimates and the Gauss-Newton method is used to find the least squares estimates, then the resulting iterative procedures will be identical. We will also consider minimum chi-square estimation as a method for obtaining best asymptotically normal estimates. Using some results that were obtained by Ferguson (1958) it will be shown that when a modified minimum chi-square procedure is followed then

the iteratively obtained estimates are identical to weighted least squares estimates. Consequently, under the conditions which are given in Section 3.3 weighted least squares estimates will be best asymptotically normal whatever the distribution of the observations. When the observations are counts that follow the Poisson distribution then maximum likelihood, least squares, and minimum **chi-square estimates** will all be equivalent —in the sense that they are obtained using identical computational procedures. The computational procedure that we will use is summarized in Section 4.4 and a listing of a Fortran IV implementation will be provided in the Appendix.

## 1.2 Spectral Analysis

The recent developments in the areas of biomedical signal measurement, analog to digital converters, and high speed digital computers has led to a new kind of problem in data analysis. The potential contribution to be made by the statistician to the development of quantitative methods for extracting useful and/or interesting information in this area is exemplified by the recent development of the Cooley-Tukey fast Fourier transform. As Tukey (1967) has pointed out the specter of computational effort has been exorcised from spectrum analysis. Consequently, the data analyst is now confronted with the substantive aspects of the scientific problems which are approachable via the frequency domain. We venture to predict that as the hybrid computer emerges as an important instrument for data acquisition and analysis, so must a new hybrid approach to data analysis evolve. The development of quantitative methods in this area requires

some knowledge of the overall scientific problem, hardware limitations, and response time requirements if meaningful results are to be obtained. We hope to make a contribution in this area by applying existing statistical theory and methods to solve some problems in biomedical data analysis. Following the suggestion of Tukey and Wilk (1966) we will attempt to develop new techniques that are empirically informative and that will provide exposure and summarization of the data. Accordingly, the models and the methods that we will use are intended to be of data analytic value. We do not suggest that any new results of a theoretical nature (either physiological or mathematical) will be presented here.

In Chapter 5 we will discuss the statistical considerations of numerical spectrum analysis, and describe the estimation procedure that will be used in Chapter 6 and 7. The basic theory of bivariate spectral analysis will be reviewed. The method of estimation that we will use is based on the periodogram. The periodogram is computed from the finite Fourier transform of a realization of a time series. A heuristic treatment of the statistical properties of spectrum estimators will be included. The calculations of spectrum analysis will be discussed and two estimation procedures that employ frequency domain smoothing techniques are summarized.

In Chapter 6 a simple mechanical model of the respiratory system will be developed. The model will be used to motivate the definition of the power spectrum that we will present. We will propose spectrum analysis as a data analytic technique that can be used to quantify the

mechanical activity of the respiratory system. When the flow-resistive component of the transpulmonary pressure is linearly related to air flow, we claim that the flow spectrum will be a power spectrum. In this case the flow spectrum will show how the average power dissipated in overcoming the resistance to air flow in the respiratory system is distributed with respect to frequency. A cross-spectrum analysis of some respiratory pressure-flow time series data will be presented. We further propose that the flow variance may be used as a single statistic to describe the mechanical effort of the respiratory system.

In Chapter 7 we will describe a situation in which such a statistic can be employed. The impedance pneumograph is a bioimpedance recorder that is used for the indirect measurement of respiration. It is currently used to monitor the respiratory function of remotely located subjects. Although calibration (in terms of respiratory volume) is theoretically possible, the restrictive requirements are not considered realistic in the monitoring situation. Nevertheless, spectra calculated from pulmonary impedance tracings do reflect changes in the depth, regularity, and frequency of breathing. An example using pulmonary impedance recordings obtained from a patient before and after exposure to total-body radiation will be presented. We will then propose that the first difference of the pulmonary impedance pneumograph should be used to obtain a more meaningful measure of respiratory effort. Argument in favor of this approach will be developed using the power concept which was introduced in Chapter 6. An example will be included which graphically demonstrates the effect of the

differencing operation on the spectrum. That is the influence of the low frequency noise components is greatly reduced when the spectrum is calculated using the first difference data.

## 2. POISSON COUNT DATA: MAXIMUM LIKELIHOOD ESTIMATION

The maximum likelihood (ML) principle has been extensively used as a method of estimation since first being introduced into statistical theory by R. A. Fisher in 1921. In Section 2.1 the ML principle is used to estimate the parameter in a linear regression model. Although this problem has been thoroughly treated by Gart (1964), a brief account is included here since it is the only known case of the general model which does not require iterative methods for solution. In Chapter 3 the same estimate of the parameter will be obtained when a weighted least squares analysis is carried out.

In Section 2.2 a general regression model will be defined, and in Section 2.3 the method of scoring will be used to maximize the logarithm of the likelihood function with respect to the parameters. In Section 2.4 some special cases of the general model will be considered, and we shall obtain some results that will be needed for the applications in Chapter 4.

### 2.1 One Parameter Model

Before obtaining the ML estimates for the parameters in the general regression model—which will be defined in Section 2.2—the special case of one parameter and one independent variable will be considered, i.e.,

$$E(y_{ij}) = f(X_i, \theta) = \theta x_i, \quad \begin{array}{l} i= 1, \dots, N, \\ j= 1, \dots, n_i \end{array} \quad (2.1.1)$$

In the above equation  $\theta$  may represent the concentration of particles

(e.g. bacteria) per unit volume of suspension, and  $x_i$  the concentration of the  $i$ th dilution of the original suspension which is introduced into some nutrient medium. The viable particles each grow and form one visible colony, and  $y_{ij}$  is the number of colonies observed on the  $j$ th replication of the  $i$ th dilution (see Section 4.1). If the counts follow the Poisson distribution and are independent, then the likelihood of  $\theta$  given the observed values of the random variables,  $y_{ij}$ ,  $i = 1, \dots, N$ ,  $j = 1, \dots, n_i$ , is

$$\prod_{i=1}^N \prod_{j=1}^{n_i} \frac{e^{-\theta x_i} (\theta x_i)^{y_{ij}}}{y_{ij}!} . \quad (2.1.2)$$

To find the ML estimate of  $\theta$ , we differentiate the logarithm of the likelihood function with respect to  $\theta$ , set the resulting expression equal to zero and solve for  $\theta$ , denoting the solution by  $\hat{\theta}$ . i.e.,

$$\frac{\partial L(\theta)}{\partial \theta} = \sum_i \sum_j \left( \frac{y_{ij}}{\hat{\theta}} - x_i \right) = 0 . \quad (2.1.3)$$

The ML estimate of  $\theta$  is easily found to be

$$\hat{\theta} = \sum_i \sum_j y_{ij} / \sum_i n_i x_i . \quad (2.1.4)$$

As we shall see in Section 2.3 this special case is the only one considered which does not require some type of iterative method to solve.

## 2.2 Specification of the General Model

The mathematical expression for the Poisson distribution was first discovered by Simeon Denis Poisson in 1837. It was however Bortkiewicz in 1898 who first demonstrated that Poisson's formula

could be used as a probability distribution for certain types of discrete data. In his Handbook of the Poisson Distribution, Haight claims that the Poisson distribution is second in importance to the normal distribution, both in terms of abstract theory and breadth of application. A complete account of the historical development, mathematical properties, and applications of the Poisson distribution as well as the Poisson process can be found in Haight's handbook.

The use of the Poisson distribution to characterize the variability of the observations in a regression analysis has only recently been considered. In this chapter it will be assumed that the observations obtained in an experiment are counts that follow the Poisson distribution with expectation given by

$$E(y_{ij}) = f(X_i, \theta), \quad i = 1, \dots, N, \quad (2.2.1)$$

$$j = 1, \dots, n_i .$$

In the above expression  $X_i = (x_{i1}, \dots, x_{im})$  represents the  $i$ th value taken by a set of  $m$  independent variables, and  $n_i$  is the number of replications of  $X_i$ .  $\theta$  is a  $p$ -dimensional vector of unknown parameters--which belongs to  $\Theta \subset \mathbb{R}^p$  ( $p$  dimensional Euclidean space)--and will be estimated from the data. The quantity  $f(X, \theta)$  may represent, for example, the mean number of bacteria per unit volume of suspension, the expected number of failures of a piece of equipment per unit time, or the average number of purchases of a particular commodity per family. It is assumed that--

1. some general form of the model is known,
2.  $f(X, \theta)$  is a differentiable function of  $\theta$ ,

3.  $N$  values of the independent variables are selected by the experimenter or specified by the situation, and
4.  $N$  is greater than  $p$  to ensure estimability of the parameters.

The problems of estimation and hypothesis testing in simple linear regression were solved by Gart (1964) and Roberts and Coote (1965), and multiple linear regression has been considered by Jorgenson (1963). In Section 2.3 we will use the ML principle to estimate the parameters in the model defined above. In Chapter 3 it will be demonstrated that the ML estimates are equivalent to those obtained in a properly weighted least squares analysis. The least squares estimates are shown to be identical to those obtained by minimizing an appropriately defined quadratic form which will in general result in estimates of the parameters that are best asymptotically normal. Methods of testing the 'goodness of fit' of the model and the assumption of Poisson variation are proposed, and modifications of the analysis are suggested for instances in which the Poisson assumption does not appear to be acceptable.

### 2.3 ML Estimates for General Regression Model

From the assumptions stated in Section 2.2 the likelihood function of the  $p$ -dimensional parameter  $\theta$ , given a particular realization of the experiment is

$$\prod_{i=1}^N \prod_{j=1}^{n_i} \frac{e^{-f(X_i, \theta)} f(X_i, \theta)^{y_{ij}}}{y_{ij}!} \quad (2.3.1)$$

The logarithm of the likelihood is

$$L(\theta) = \sum_i \sum_j \left[ -f(X_i, \theta) + y_{ij} \ln f(X_i, \theta) \right] + C \quad (2.3.2)$$

which may be written, neglecting a constant which does not involve the parameters, as

$$L(\theta) = \sum_i y_{i\cdot} \ln f(x_i, \theta) - \sum_i n_i f(x_i, \theta) \quad , \quad (2.3.3)$$

where  $y_{i\cdot} = \sum_{j=1}^{n_i} y_{ij}$ . The likelihood equations are obtained by differentiating (2.3.3) with respect to  $\theta_k$ ,  $k=1, \dots, p$ , and equating the partial derivatives to zero, i.e.,

$$\begin{aligned} \frac{\partial L(\theta)}{\partial \theta_1} &= \sum_i \left[ \left( \frac{\partial f(x_i, \theta)}{\partial \theta_1} \right) \left( \frac{y_{i\cdot}}{f(x_i, \theta)} - n_i \right) \right] = 0 \\ &\vdots \\ &\vdots \\ \frac{\partial L(\theta)}{\partial \theta_p} &= \sum_i \left[ \left( \frac{\partial f(x_i, \theta)}{\partial \theta_p} \right) \left( \frac{y_{i\cdot}}{f(x_i, \theta)} - n_i \right) \right] = 0 \end{aligned} \quad (2.3.4)$$

The ML estimate of  $\theta$  is the solution of these equations.

Since these equations are not linear in the  $\theta_k$ 's (2.3.4) is expanded about a trial value,  $\theta^0$ , in a first-order Taylor series and the following equations result

$$\begin{aligned} \frac{\partial L(\theta)}{\partial \theta_1} &= \frac{\partial L(\theta^0)}{\partial \theta_1} + \delta_1 \frac{\partial^2 L(\theta^0)}{\partial \theta_1^2} + \dots + \delta_p \frac{\partial^2 L(\theta^0)}{\partial \theta_1 \partial \theta_p} \quad , \\ &\vdots \\ &\vdots \\ &\vdots \\ \frac{\partial L(\theta)}{\partial \theta_p} &= \frac{\partial L(\theta^0)}{\partial \theta_p} + \delta_1 \frac{\partial^2 L(\theta^0)}{\partial \theta_1 \partial \theta_p} + \dots + \delta_p \frac{\partial^2 L(\theta^0)}{\partial \theta_p^2} \quad , \end{aligned} \quad (2.3.5)$$

where the first-order and second-order partial derivatives are evaluated at  $\theta = \theta^0$ . This system of equations is linear in the

unknown  $\delta_k^o$ 's and the right hand side may be rewritten as

$$\begin{bmatrix} \frac{\partial L(\theta^o)}{\partial \theta_1} \\ \vdots \\ \frac{\partial L(\theta^o)}{\partial \theta_p} \end{bmatrix} + \begin{bmatrix} \frac{\partial^2 L(\theta^o)}{\partial \theta_1^2} & \dots & \frac{\partial^2 L(\theta^o)}{\partial \theta_1 \partial \theta_p} \\ \vdots & & \vdots \\ \frac{\partial^2 L(\theta^o)}{\partial \theta_1 \partial \theta_p} & \dots & \frac{\partial^2 L(\theta^o)}{\partial \theta_p^2} \end{bmatrix} \begin{bmatrix} \delta_1^o \\ \vdots \\ \delta_p^o \end{bmatrix} = 0 \quad (2.3.6)$$

When equated to zero this system of equations can be solved, but it requires the calculation and evaluation of the matrix of second-order partial derivatives. An alternative approach is to replace the matrix of second-order partial derivatives with its expected value, and solve the new system of equations. This is known as the method of scoring (Rao 1965, p.305). Following this approach we first obtain the 'information' matrix:

$$E \left[ - \frac{\partial^2 L(\theta)}{\partial \theta_r \partial \theta_s} \right] = E \left[ \left( \frac{\partial L(\theta)}{\partial \theta_r} \right) \cdot \left( \frac{\partial L(\theta)}{\partial \theta_s} \right) \right] \quad (2.3.7)$$

Then substituting the expressions for the first-order partial derivatives given by equation (2.3.4) into (2.3.7) we have the element in row r column s equal to

$$E \left[ \left( \sum_i \left( \frac{\partial f(x_i, \theta)}{\partial \theta_r} \right) \left( \frac{y_i}{f(x_i, \theta)} - n_i \right) \right) \left( \sum_i \left( \frac{\partial f(x_i, \theta)}{\partial \theta_s} \right) \left( \frac{y_i}{f(x_i, \theta)} - n_i \right) \right) \right], \quad (2.3.8)$$

which may be rewritten as

$$E \left[ \sum_i \left( \left| \frac{\partial f(X_i, \theta)}{\partial \theta_r} \right| \left| \frac{\partial f(X_i, \theta)}{\partial \theta_s} \right| \left| \frac{y_{i\cdot} - f(X_i, \theta)}{f(X_i, \theta)} \right|^2 \right) + \right. \\ \left. \sum_i \sum_{i \neq k} \left( \left| \frac{\partial f(X_i, \theta)}{\partial \theta_r} \right| \left| \frac{\partial f(X_k, \theta)}{\partial \theta_s} \right| \left| \frac{y_{i\cdot} - n_i f(X_i, \theta)}{f(X_i, \theta)} \right| \left| \frac{y_{k\cdot} - n_k f(X_k, \theta)}{f(X_k, \theta)} \right| \right) \right]. \quad (2.3.9)$$

To evaluate (2.3.9) we use the assumption that the  $y_{ij}$ 's are independent and follow the Poisson distribution with  $E(y_{ij}) = f(X_i, \theta)$  to obtain

$$E(y_{i\cdot}) = E\left(\sum_{j=1}^{n_i} y_{ij}\right) = n_i f(X_i, \theta),$$

$$E\left(y_{i\cdot} - E(y_{i\cdot})\right)^2 = \text{Var}(y_{i\cdot}) = n_i f(X_i, \theta),$$

$$\text{and } E(y_{i\cdot} y_{k\cdot}) = E(y_{i\cdot}) E(y_{k\cdot}). \quad (2.3.10)$$

Evaluating the expectation in (2.3.9) and using the results in (2.3.10) we obtain the following expression for the information matrix

$$\left[ E \left( \frac{\partial L(\theta)}{\partial \theta_r} \right) \left( \frac{\partial L(\theta)}{\partial \theta_s} \right) \right] = \left[ \sum_i \left( \left| \frac{\partial f(X_i, \theta)}{\partial \theta_r} \right| \left| \frac{\partial f(X_i, \theta)}{\partial \theta_s} \right| \left| \frac{n_i}{f(X_i, \theta)} \right| \right) \right], \quad (2.3.11)$$

$$r, s = 1, \dots, p.$$

Equation (2.3.6), when set equal to zero, then becomes

$$\left[ \frac{\partial L(\theta^0)}{\partial \theta_r} \right] - \left[ \sum_i \left( \left| \frac{\partial f(X_i, \theta^0)}{\partial \theta_r} \right| \left| \frac{\partial f(X_i, \theta^0)}{\partial \theta_s} \right| \left| \frac{n_i}{f(X_i, \theta^0)} \right| \right) \right] \left[ \delta_r^0 \right] = \left[ 0 \right], \quad (2.3.12)$$

$$r, s = 1, \dots, p.$$

This system of  $p$ -simultaneous equations is then solved for  $\delta_r^0$ ,  $r=1, \dots, p$  and new values of the parameters are obtained, i.e.,

$$\theta_r^1 = \theta_r^0 + \delta_r^0 \quad r=1, \dots, p. \quad (2.3.13)$$

The procedure just described is then repeated until a stable solution is reached. The criterion that is used to determine when convergence has occurred will be discussed in Chapter 4.

#### 2.4 ML Estimates for Some Special Cases

In this section the results obtained in the preceding section will be used to write down the likelihood equations and the information matrix for several regression models that will be considered in the applications in Chapter 4. A discussion of the independent variables and parameters will be presented at that time. It will, however, become apparent when considering the computational algorithms that only the regression function and the first-order partial derivatives of the regressive function with respect to the parameters are necessary to perform the computations that are required to solve the likelihood equations. In Chapter 3 it will be shown that a general purpose nonlinear regression program can be used in practice to solve the likelihood equations.

The first model is the multiple linear regression model, and then three survival curve models are considered. A survival curve model is obtained by writing equation (2.1.1) in the form

$$f(x_i, \theta) = \theta_1 x_{i1} S(x_{i2}, \bar{\theta}) \quad (2.4.1)$$

In the above expression  $S$  is a function of a second independent

variable  $x_2$  (such as amount of radiation), and one or more unknown parameters,  $\bar{\theta} = (\theta_2, \dots, \theta_p)'$ . To qualify as a survival curve model we require that—

1.  $x \geq 0$ ;
2.  $S(x, \bar{\theta}) = 1$  when  $x=0$ ;
3.  $S(x, \bar{\theta})$  is monotonically decreasing and nonnegative;
4.  $S(x, \bar{\theta})$  is a differentiable function of  $\bar{\theta}$ .

From the above it is apparent that the distribution function of any nonnegative random variable will produce a survival curve if  $S(x, \bar{\theta}) = 1 - F(x, \bar{\theta})$ , where  $F(x, \bar{\theta})$  is a distribution function which is differentiable with respect to  $\bar{\theta}$ .

The exponential and target survival curve models have been widely used by biologists to describe the effects of radiation on the survival of cells and microorganisms. The Weibull model, while not previously considered in this context, appears to be an interesting alternative to the target model. We shall defer further discussion of these models and their applications to Chapter 4.

#### 2.4.1 Multiple Linear Regression

When the regression model defined in equation (2.1.1) is linear in the unknown parameters, i.e.,

$$f(X_i, \theta) = \theta_1 x_{i1} + \dots + \theta_p x_{ip} = \sum_{j=1}^p x_{ij} \theta_j = X_i \theta, \quad (2.4.2)$$

then

$$\frac{\partial f(X_i, \theta)}{\partial \theta_r} = x_{ir}, \quad r = 1, \dots, p, \quad (2.4.3)$$

and the efficient scores obtained using (2.3.4) are

$$\frac{\partial L(\theta)}{\partial \theta_r} = \sum_{i=1}^N \left[ \left( \frac{x_{ir} n_i}{x_i \theta} \right) (y_{i.} - x_i \theta) \right] = 0, \quad r=1, \dots, p. \quad (2.4.4)$$

The information matrix obtained using equation (2.3.11) is

$$\left[ \sum_{i=1}^N \frac{x_{ir} x_{is} n_i}{x_i \theta} \right], \quad \begin{array}{l} r=1, \dots, p, \\ s=1, \dots, p. \end{array} \quad (2.4.5)$$

These results agree with those obtained by Jorgenson (1963) when the appropriate changes in notation are made and provided expected values are taken in his expression for the information matrix. Jorgenson accurately described his procedure for solving the likelihood equations as the Newton-Raphson iterative process, which corresponds to the solution that is obtained when (2.3.6) is set equal to zero. When the method of scoring is used the iterative equations obtained by substituting equations (2.4.4) and (2.4.5) into (2.3.12) will be identical to those obtained when a weighted least squares analysis is carried out (Jorgenson 1963, eq. 3.14). In Chapter 3 it will be shown that this result is true for the general regression model when the Gauss-Newton iterative process is used in the least squares analysis.

#### 2.4.2 Exponential Survival Curve

The exponential survival curve is defined by

$$f(x_i, \theta) = \theta_1 x_{i1} e^{-\theta_2 x_{i2}}, \quad \theta_1, \theta_2 > 0 \quad (2.4.6)$$

and the partial derivatives are

$$\frac{\partial f(x_i, \theta)}{\partial \theta_1} = x_{i1} \exp(-\theta_2 x_{i2}) ,$$

$$\frac{\partial f(x_i, \theta)}{\partial \theta_2} = -x_{i2} \theta_1 x_{i1} \exp(-\theta_2 x_{i2}) . \quad (2.4.7)$$

The elements in the vector of scores are

$$\begin{aligned} \frac{\partial L(\theta)}{\partial \theta_1} &= \frac{1}{\theta_1} \sum_{i=1}^N \left( n_i (y_{i.} - \theta_1 x_{i1} \exp(-\theta_2 x_{i2})) \right) , \\ \frac{\partial L(\theta)}{\partial \theta_2} &= \sum_{i=1}^N \left( x_{i2} n_i (y_{i.} - \theta_1 x_{i1} \exp(-\theta_2 x_{i2})) \right) , \end{aligned} \quad (2.4.8)$$

and the information matrix is

$$\begin{bmatrix} \frac{1}{\theta_1} \sum_i n_i x_{i1} \exp(-\theta_2 x_{i2}) & -\sum_i n_i x_{i1} x_{i2} \exp(-\theta_2 x_{i2}) \\ & \sum_i x_{i2}^2 \theta_1 x_{i1} \exp(-\theta_2 x_{i2}) \end{bmatrix} . \quad (2.4.9)$$

Lellouch and Wambersic (1966) obtained the above results and proposed a method for solving the equations. The results obtained by applying their method will be discussed in Chapter 4.

### 2.4.3 Target Survival Curve

The target survival curve model is

$$\begin{aligned} f(x_i, \theta) &= \theta_1 x_{i1} \left[ 1 - \left( 1 - \exp(-\theta_2 x_{i2}) \right)^{\theta_3} \right] , \\ \theta_1, \theta_2, \theta_3 &> 0 \end{aligned} \quad (2.4.10)$$

and the partial derivatives are

$$\frac{\partial f(X_i, \theta)}{\partial \theta_1} = x_{i1} \left( 1 - A_i^{\theta_3} \right) ,$$

$$\frac{\partial f(X_i, \theta)}{\partial \theta_2} = -\theta_1 \theta_3 x_{i1} A_i^{\theta_3 - 1} x_{i2} \exp(-\theta_2 x_{i2}) ,$$

$$\frac{\partial f(X_i, \theta)}{\partial \theta_3} = -\theta_1 x_{i1} A_i^{\theta_3} \ln A_i \quad (x_{i2} \neq 0) , \quad (2.4.11)$$

where  $A_i = 1 - \exp(-\theta_2 x_{i2})$ . The partial derivatives can be substituted into equations (2.3.4) and (2.3.11) to obtain the efficient scores and the information matrix. As was stated at the beginning of this section only the regression function and the partial derivatives are required to find  $\hat{\theta}$  when the algorithm described in Chapter 4 has been programmed for computation on a digital computer. Methods for finding initial estimates of the parameters will be given in Chapter 4.

#### 2.4.4. Weibull Survival Curve

The Weibull survival curve model is

$$f(X_i, \theta) = \theta_1 x_{i1} \exp \left( -\theta_2 (x_{i2})^{\theta_3} \right) , \quad \theta_1, \theta_2, \theta_3 > 0 \quad (2.4.12)$$

and the partial derivatives are

$$\frac{\partial f(X_i, \theta)}{\partial \theta_1} = x_{i1} \exp \left( -\theta_2 (x_{i2})^{\theta_3} \right) ,$$

$$\frac{\partial f(X_i, \theta)}{\partial \theta_2} = -(x_{i2})^{\theta_3} \theta_1 x_{i1} \exp \left( -\theta_2 (x_{i2})^{\theta_3} \right) ,$$

$$\frac{\partial f(X_i, \theta)}{\partial \theta_3} = -\theta_2 (x_{i2})^{\theta_3} \ln(x_{i2}) \theta_1 x_{i1} \exp \left( -\theta_2 (x_{i2})^{\theta_3} \right) . \quad (2.4.13)$$

The efficient scores and the information matrix can be obtained using equations (2.3.4) and (2.3.11). Methods for finding initial estimates of the parameters will be given in Chapter 4, and the Weibull model will be considered as an alternative to the target model. It is of interest to note that both of these models reduce to the simpler exponential model when  $\theta_3 = 1$  in equations (2.4.10) and (2.4.12).

## 2.5 Covariance Matrix and Chi-Square Tests

The large sample covariance matrix of ML estimators is the inverse of the information matrix, i.e.,

$$V = \left[ \sum_i \left( \left| \frac{\partial f(X_i, \theta)}{\partial \theta_r} \right| \left| \frac{\partial f(X_i, \theta)}{\partial \theta_s} \right| \left( \frac{n_i}{f(X_i, \theta)} \right) \right) \right]^{-1} \quad (2.5.1)$$

If  $\hat{\theta}$  is the solution of the likelihood equations, then estimates of the elements of this matrix may be obtained by replacing  $\theta$  by  $\hat{\theta}$ . The diagonal elements of this matrix may then be used to establish confidence intervals for the parameters in the regression model. Also, the expected number of counts for each value  $X_i$ ,  $i = 1, \dots, N$  of the independent variables may be estimated by

$$E(\hat{y}_{ij}) = f(X_i, \hat{\theta}) \quad , \quad j = 1, \dots, n_i \quad (2.5.2)$$

Then the statistic

$$Q_t = \sum_i \sum_j \left( \frac{(y_{ij} - f(X_i, \hat{\theta}))^2}{f(X_i, \hat{\theta})} \right) \quad (2.5.3)$$

will be distributed approximately as a chi-square with

$D_t = (\sum_i n_i) - p$  d.f. if the regression model and the assumption of Poisson variation are valid.  $Q_t$  may be partitioned into two independent components as follows:

$$\begin{aligned}
 Q_t &= Q_w + Q_r \\
 &= \sum_i \sum_j \left( \frac{(y_{ij} - \bar{y}_{i.})^2}{f(X_i, \hat{\theta})} \right) + \sum_i \left( \frac{n_i}{f(X_i, \hat{\theta})} (\bar{y}_{i.} - f(X_i, \hat{\theta}))^2 \right),
 \end{aligned}
 \tag{2.5.4}$$

with degrees of freedom  $D_w = (\sum_i n_i) - N$  and  $D_r = N - p$ . If the value of the first statistic,  $Q_t$ , is found to be significant, it may be due either to heterogeneity of variance or 'lack of fit' of the regression model, or both. In this situation  $Q_w$  may be compared with the chi-square distribution with  $D_w$  d.f.. If this result is significant, then the ratio  $(Q_r/D_r)/(Q_w/D_w)$  may be compared with the F distribution. If this approximate F test does not indicate significant deviation from the model, then the elements of the estimated covariance matrix are multiplied by a heterogeneity factor,  $H = Q_w/D_w$ .

Although the above results are only based on approximate significance tests, they may provide some useful insight into the sources of errors in an experiment. Further, it should be noted that while the above tests will indicate certain types of inadequacies in the model, it is clearly possible to have an over parameterized model that will fit the data. Consequently, it may be of interest to determine if some subset of the  $\theta_r$ 's satisfy certain restrictions,  $R_i(\theta) = 0$ ,  $i = 1, \dots, k$ . Minus twice the natural logarithm of the Neyman-Pearson likelihood ratio statistic for this type of composite

hypothesis is

$$2 \left( L(\hat{\theta}) - L(\bar{\theta}) \right), \quad (2.5.5)$$

where  $\bar{\theta}$  is the ML estimate of the parameters in the restricted model.

The asymptotic distribution of the above statistic is a chi-square with  $k$  d.f., and may be used to test the hypothesis stated above (see Rao 1965, p.350).

### 3. LEAST SQUARES

In the first two sections of this chapter equation (1.2.1) will be considered in the form

$$y_{ij} = f(X_i, \theta) + e_{ij} \quad ,$$

where  $E(e_{ij}) = 0$ ,  $\text{var}(e_{ij}) = \sigma_i^2$ , and  $\text{cov}(e_{ij}, e_{i'j'}) = 0$ . The problems of estimation and inference when the variance is constant for all  $i$  will be reviewed in Section 3.1. In Section 3.2 we will show how weighted least squares (LS) estimates may be obtained when variances are not assumed constant, and demonstrate that when the observations are Poisson the weighted least squares estimates will be identical to the maximum likelihood estimates of Chapter 2.

In Section 3.4 minimum chi-square (MCS) estimation will be introduced and it will be demonstrated that when a modified minimum chi-square procedure is followed the computations are identical to those of weighted least squares. In Section 3.3 a class of estimates will be defined which are best asymptotically normal (BAN), and a Theorem due to Ferguson (1958) will be presented that shows how BAN estimates may be obtained as roots of certain linear forms. Since the MCS estimates of Section 3.4 will be of this type they are BAN, and since the modified MCS estimates will be computationally equivalent to weighted LS estimates, the latter will also be BAN. Consequently, under the assumptions to be given in Section 3.3, we see that weighted least squares estimates will be BAN whatever the distribution of the observations. When the observations are Poisson the solution of the

likelihood equations results in the same linear form that is defined by Ferguson's Theorem demonstrating that ML estimates are BAN—as should be expected. Further, it will be shown that when the method of scoring, the Gauss-Newton method, and the modified MCS approach are used to obtain ML, LS, and MCS estimates, respectively, the iterative procedures that result will be identical.

### 3.1 Least Squares - Constant Variance Case

In this section it is assumed that the observations have been obtained from a regression model

$$y_i = f(X_i, \theta) + e_i, \quad (3.1.1)$$

where the  $e_i$ 's are error residuals and are assumed to be independent with  $E(e_i) = 0$  and  $\text{var}(e_i) = \sigma^2$ ,  $i = 1, \dots, N$ . In (3.1.1)  $X_i$  is an  $m$ -dimensional vector of independent variables and  $\theta$  is a  $p$ -dimensional vector of unknown parameters. Estimation by the least squares (LS) principle requires minimization of

$$S(\theta) = \sum_{i=1}^N (y_i - f(X_i, \theta))^2. \quad (3.1.2)$$

The least squares equations are obtained by differentiating (3.1.2) with respect to  $\theta_r$ ,  $r = 1, \dots, p$ , and take the form

$$\frac{\partial S(\theta)}{\partial \theta_r} = \sum_i [y_i - f(X_i, \theta)] \left( \frac{\partial f(X_i, \theta)}{\partial \theta_r} \right) = 0, \quad r = 1, \dots, p. \quad (3.1.3)$$

It is well known that when  $f(X, \theta)$  is linear in the parameters, i.e.,

$$f(X_i, \theta) = \theta_1 x_{i1} + \dots + \theta_p x_{ip}, \quad (3.1.4)$$

the LS estimators will be best linear unbiased estimators (BLUE) of  $\theta$ .

If in addition to the above assumptions the  $e_i$ 's are assumed to be normally distributed then minimizing (3.1.2) will be equivalent to maximizing the likelihood function

$$(2\pi\sigma^2)^{-(N/2)} \exp \left[ - \sum_i \left( y_i - f(x_i, \theta) \right)^2 / 2\sigma^2 \right] .$$

Consequently the large sample properties of LS estimators in nonlinear models can be inferred from those of ML estimation. The asymptotic properties of LS estimators in (3.1.1) have been considered by Hartley (1964), Hartley and Booker (1965), Villegas (1969), Jennrich (1969) and Malinvaud (1970).

When the model is linear, (3.1.1) may be written in matrix notation as follows:

$$\begin{array}{c} [y_i] \\ \text{Nx1} \end{array} = \begin{array}{c} [x_{ij}] \\ \text{Nxp} \end{array} \begin{array}{c} [\theta_j] \\ \text{px1} \end{array} + \begin{array}{c} [e_i] \\ \text{Nx1} \end{array} . \quad (3.1.5)$$

If we assume the rank of  $[x_{ij}] = p$ , then the LS estimator is the unique solution of (3.1.3), i.e.

$$\begin{aligned} [\hat{\theta}_j] &= \left( [x_{ij}]' [x_{ij}] \right)^{-1} [x_{ij}]' [y_i] , \\ &= (X'X)^{-1} X'Y . \end{aligned} \quad (3.1.6)$$

$\hat{\theta}$  is BLUE and represents a set of  $p$  statistics jointly sufficient for the estimation of  $\theta$ . Further well known results for the linear case may be summarized as follows:

- 1.)  $\hat{\theta}$  is normally distributed with expected value  $\theta$  and covariance matrix  $(X'X)^{-1} \sigma^2$ ,
- 2.)  $S(\hat{\theta})$  is distributed as  $\sigma^2 \chi^2$  with  $N-p$  degrees of freedom,
- 3.)  $R(\hat{\theta})$ , as defined below, is distributed as  $\sigma^2 \chi^2$  with

p d.f., and

- 4.) the quadratic forms  $S(\hat{\theta})$  and  $R(\hat{\theta})$  are statistically independent.

In 3. above,  $R(\hat{\theta})$  represents the 'regression' sum of squares and  $S(\hat{\theta})$  is the 'residual' sum of squares. Since  $\hat{\theta}$  is a random variable in a p-dimensional linear subspace of the sample space and is orthogonal to the N-p dimensional error space, the relation between the vector  $(Y - X\theta)$ ,  $(Y - \hat{Y})$ , and  $(\hat{Y} - X\theta)$  may be expressed by using the Pythagorean Theorem as follows:

$$\begin{aligned} (Y - X\theta)'(Y - X\theta) &= (\hat{Y} - X\theta)'(\hat{Y} - X\theta) + (Y - \hat{Y})'(Y - \hat{Y}) \\ &= R(\hat{\theta}) + S(\hat{\theta}) \end{aligned} \quad (3.1.7)$$

where  $\hat{Y} = X\hat{\theta}$ , and the regression sum of squares may be written

$$\begin{aligned} R(\hat{\theta}) &= [X'(Y - X\theta)]'(X'X)^{-1}[X'(Y - X\theta)] \\ &= (\theta - \hat{\theta})'(X'X)(\theta - \hat{\theta}) \end{aligned} \quad (3.1.8)$$

An exact  $100\alpha\%$  confidence region for  $\theta$  is given by

$$R(\hat{\theta})/S(\hat{\theta}) \leq pF(\alpha; p, N-p)/(N-p), \quad (3.1.9)$$

where  $F(\alpha; p, N-p)$  is the upper  $100\alpha\%$  point of the "F" distribution with p and N-p degrees of freedom. Hartley (1964) has discussed the extension of this result to nonlinear models.

When the regression model (3.1.1) is not linear in all of the parameters the least squares equations (3.1.3) cannot be solved directly. When this is the case we shall refer to the regression model as being nonlinear, or more precisely 'intrinsically nonlinear' - as defined by Draper and Smith (1966, chap. 10) - if it is not possible to express the model as a linear combination of the parameters by some

type of transformation. To obtain estimates of the parameters in nonlinear regression models various iterative techniques have been developed that involve the approximation of a nonlinear model by a linear one. One of the best known is the Gauss-Newton method which will be briefly described in the next section and more thoroughly considered in Chapter 4. Meeter (1964) has discussed problems encountered in the analysis of nonlinear models, and provides a review of the design problems (i.e., selection of the values of the independent variables). Box and Hunter (1965) have considered the dual problem of generating and analyzing data in situations where the effect of some underlying physical mechanism is expressed through the parameters in a nonlinear regression model. Draper and Hunter (1967) have considered the design problem when the model is nonlinear with normal independent errors, constant variance, and prior information is available on the parameters. Box and Draper (1971) have discussed the  $|X'X|$  design criterion, where  $X = \left[ \frac{\partial f(x_j, \theta)}{\partial \theta_r} \right]$ , and Feder and Mezaki (1971) have proposed variational methods for determining the 'best' experimental design. If the assumption of constant variance is not acceptable, then designs based on a 'variance criterion' require prior information on the parameters.

### 3.2 Weighted Least Squares

In this section we demonstrate that iteratively weighted LS estimates are identical to ML estimates when the LS weights are defined in accord with the assumption of Poisson variation. Consider the problem of obtaining LS estimates by minimizing

$$s(\theta) = \sum_{i=1}^N w_i [z_i - f(X_i, \theta)]^2 \quad (3.2.1)$$

with respect to  $\theta$ . In the above equation  $w_i$  is the weight associated with the  $i$ th observation and we choose it to be  $1/\text{var}(z_i)$ . If the variances of the  $z_i$ 's are not known, we assume that repeated observations have been made for each value of  $X_i$ , that  $z_i = \sum_{j=1}^{n_i} y_{ij}/n_i$ , and that  $w_i = n_i/s_i^2$ , where  $s_i^2$  is the estimated sample variance of the  $y_{ij}$ 's. This is a special case of the generalized least squares estimation procedure that has been developed by Beauchamp (1966), who showed that if the observations are normally distributed and if the variances are known or are replaced by consistent estimates, then the LS estimates will be equivalent to the ML estimates.

Before developing a general procedure for finding weighted LS estimates we shall consider the special case of linear regression through the origin that was discussed in Section 2.1 (i.e., where  $f(X_i, \theta) = \theta x_i$ ). Assuming that the observations follow the Poisson law, so that  $\text{var}(z_i) = \theta x_i/n_i$ , (3.2.1) becomes

$$s(\theta) = \sum_i w_i [z_i - \theta x_i]^2, \quad (3.2.2)$$

and the LS estimate of  $\theta$  is

$$\hat{\theta} = \frac{\sum_i w_i z_i x_i}{\sum_i w_i x_i^2} = \frac{\sum_i n_i z_i}{\sum_i n_i x_i}, \quad (3.2.3)$$

the same result established in Section 2.1 using the ML principle.

When  $f(X, \theta)$  is nonlinear, (3.2.1) can be minimized by using the Gauss-Newton iterative method. If we let  $\theta^0$  be an initial estimate of

the parameter values, and  $f(X, \theta)$  is replaced with the first order Taylor series  $f(X_i, \theta) = f(X_i, \theta^0) + \sum_{r=1}^p (\theta_r - \theta_r^0) \partial f(X_i, \theta^0) / \partial \theta_r$ , then (3.2.1) becomes

$$S^0(\theta) = \sum_i w_i \left[ z_i - f(X_i, \theta^0) - \sum_r (\theta_r - \theta_r^0) \partial f(X_i, \theta^0) / \partial \theta_r \right]^2 . \quad (3.2.4)$$

Then letting  $d_r^0 = (\theta_r - \theta_r^0)$  and  $p_{ir}^0 = \partial f(X_i, \theta^0) / \partial \theta_r$ , we differentiate (3.2.4) with respect to  $d_s^0$ ,  $s = 1, \dots, p$ , and obtain the LS equations for the  $d_s^0$ 's which may be written as follows:

$$\frac{\partial S^0(\theta)}{\partial d_s^0} = -2 \sum_i w_i \left[ z_i - f(X_i, \theta^0) - \sum_r d_r^0 p_{ir}^0 \right] p_{is}^0 = 0 , \quad s = 1, \dots, p. \quad (3.2.5)$$

This system of  $p$  equations is linear in the unknown  $d_r^0$ 's and may be expressed in matrix form as follows:

$$\left[ \sum_i w_i p_{ir}^0 p_{is}^0 \right] \left[ d_r^0 \right] = \left[ \sum_i w_i p_{ir}^0 \left[ z_i - f(X_i, \theta^0) \right] \right] . \quad (3.2.6)$$

If the observations are Poisson, then  $\text{var}(z_i) = f(X_i, \theta) / n_i$ , which we estimate by replacing  $\theta$  by the initial estimate  $\theta^0$ . Then if we let  $w_i = 1 / \text{var}(z_i)$ , the system of equations (3.2.6) becomes

$$\left[ \sum_i n_i f(X_i, \theta^0)^{-1} p_{ir}^0 p_{is}^0 \right] \left[ d_r^0 \right] = \left[ \sum_i p_{ir}^0 \left( \frac{n_i z_i}{f(X_i, \theta^0)} - n_i \right) \right] , \quad (3.2.7)$$

which is solved for  $[d_r^0]$ . The initial estimates are then 'corrected' by taking  $\theta_r^1 = \theta_r^0 + d_r^0$ ,  $r = 1, \dots, p$ , and the above procedure is repeated with the superscript zero replaced by the superscript one. This iterative procedure is then continued until some convergence criterion is satisfied.

If we compare the system of equations in (3.2.7) with those obtained in Section 2.2 (see equations (2.3.4) and (2.3.7)) where the method of scoring was used to find the ML estimates, we see that the iterative procedures are identical. It therefore follows that when the Gauss-Newton method is used to find the LS estimates (as described in this section) the iterative procedure requires the solution of a system of linear equations which is identical to that obtained when the method of scoring is used to find the ML estimates. Consequently, the weighted LS and ML estimates are identical when the computational procedures described in this section and Section 2.3 are employed.

The choice of weights in (3.2.1) was somewhat arbitrary, although intuitively appealing. In the next two sections justification for this choice will be provided. It will be shown that the LS estimates obtained by using the computational procedure described here will be identical to the estimates obtained using a modified MCS estimation procedure. The MCS estimation procedure will in general result in estimates that are best asymptotically normal.

### 3.3 Best Asymptotically Normal Estimates

The principles of best asymptotically normal (BAN) estimation were first expounded by Neyman (1949). He defined a class of estimates that have the same asymptotic properties as ML estimates, but differ with respect to computational difficulty. Neyman considered the multinomial case and showed that minimization of an appropriately defined chi-square would produce BAN estimates. Barankin and Garland (1950), Chiang (1956), Ferguson (1958), and Wijsman (1959) have presented generalizations of the theory based on a number of independent vectors whose distribution need not be specified. Both Chiang and Ferguson considered methods for generating BAN estimates with a view toward biological applications.

A BAN estimate may be roughly described as being asymptotically normally distributed about the 'true' parameter value with smallest possible variance. In the rest of this section we summarize some of the mathematical results presented by the authors mentioned above. Of special interest is a theorem first given by Ferguson which indicates how a BAN estimate may be obtained as the root of a linear form. In Section 3.4 it will be demonstrated that when a MCS estimation procedure is employed, the linear form that results is of the type specified by Ferguson's Theorem.

Let  $Y_1, Y_2, \dots, Y_n$ , be a sequence of independent random vectors taking values in a  $N$ -dimensional subspace,  $S$  (sample space), of  $N$ -dimensional Euclidean space. The vector  $Y_j = (y_{1j}, \dots, y_{Nj})'$  represents the outcome of the  $j$ th replication of the experimental

conditions  $X_i$ ,  $i = 1, \dots, N$ , which are known (see Section 2.2). The distribution of the  $Y$ 's depends on a parameter  $\theta$  which takes values in an open subset of the 'parameter space',  $\Theta$ , which is a subspace of  $p$ -dimensional Euclidean space. If we let  $\theta^*$  denote the true value of  $\theta$  and  $Z_n = n^{-1} \sum_{j=1}^n Y_j$  be the average of the vectors  $Y_j$ , then it follows that

$$\sqrt{n} [Z_n - F(\theta^*)] \xrightarrow{\mathcal{L}} N(0, V(\theta^*)) \quad , \quad (3.3.1)$$

where  $Z_n \xrightarrow{\mathcal{L}} N(0, V)$  means that the limiting distribution of the random vector  $Z_n$  is multivariate normal with mean 0 and covariance matrix  $V$ . In terms of our previous notation,  $F(\theta) = [f(X_1, \theta), \dots, f(X_N, \theta)]'$ , where  $F$  is a function defined on  $\Theta$  into  $S$ , and  $V$  is a function defined on  $\Theta$  into the space of positive semi-definite matrices of order  $p$ . Let  $U$  be the set  $\{F(\theta), \theta \in \Theta\}$  and further note that for our purposes we may assume that the  $Y_j$ 's are identically distributed, i.e.,

$$E(Y_j) = F(\theta) \text{ and } E[Y_j - F(\theta)] \cdot [Y_j - F(\theta)]' = V(\theta) \quad . \quad (3.3.2)$$

Further assumptions required to obtain the results in this section are: (i)  $F$  is 1-1 and bicontinuous; (ii)  $V(\theta)$  is nonsingular for all values of  $\theta$ ; (iii)  $V$  is continuous and  $F$  has continuous first-order partial derivatives; (iv) the matrix  $P = [\partial f(X_i, \theta) / \partial \theta_r]$  is  $N \times p$  and has rank  $p$  for every value of  $\theta$ . These are the assumptions given by Wijsman (1959a, p.187) and lead to a more general class of estimates which include BAN estimates defined by Ferguson.

We now provide some definitions that will be used in establishing

the results that follow.

Definition 1. Let  $\hat{\theta}_n(Y_1, \dots, Y_n)$ ,  $n = 1, 2, \dots$ , be a sequence of functions of the observations taking their values from the  $p$ -dimensional space containing  $\Theta$ . The sequence  $\hat{\theta}_n$  is said to be a consistent estimate of the parameter point  $\theta$  at the true value  $\theta^*$  if, as  $n \rightarrow \infty$ ,  $\hat{\theta}_n$  tends in probability to  $\theta^*$ .

Definition 2. Let  $B$  be a positive definite matrix such that as  $n \rightarrow \infty$ , the distribution of  $\sqrt{n}B^{-1}(\hat{\theta}_n - \theta^*) \xrightarrow{d} N(0, I)$ ; then the estimate  $\hat{\theta}_n$  is said to be consistent for  $\theta^*$  and asymptotically normal. The asymptotic covariance matrix of  $\hat{\theta}_n$  is  $n^{-1}BB'$ .

Definition 3.  $\hat{\theta}$  is called regular if: (i)  $\hat{\theta}(Z_n)$  converges in probability to  $\theta^*$ , that is,  $\hat{\theta}$  is consistent; (ii)  $\hat{\theta}$  is differentiable in every point  $F(\theta)$  of  $U$ .

The above definition of regular was given by Wisjman (1959a) who pointed out that (ii) implies that for every sequence  $Z_n$  satisfying (3.3.1) we get

$$\sqrt{n}[\hat{\theta}_n(Z_n) - \theta^*] \sim A(\theta^*)\sqrt{n}[Z_n - F(\theta^*)] \quad (3.3.3)$$

In (3.3.3)  $A(\theta)$  is a  $p \times N$  matrix continuous in  $\theta$ , and the notation  $X_n \sim Y_n$  means that  $(X_n - Y_n)$  converges in probability to zero. When  $\hat{\theta}$  is constructed according to the method described in Ferguson's Theorem then  $A = (BP)^{-1}P$  where  $B = P(\theta)'V(\theta)^{-1}$ . Under any circumstance the  $A$  corresponding to BAN estimates is continuous since it is given by  $A = (P'V^{-1}P)^{-1}P'V^{-1}$  (see Wisjman 1959b, p. 1269).

Definition 4. Let  $C$  be a class of symmetric positive definite

matrices of rank  $p$ . A matrix  $H^*$  that belongs to  $C$  is said to be minimal with respect to  $C$  if for every  $H$  belonging to  $C$  the difference  $H - H^*$  is positive semidefinite; i.e., for any  $p \times 1$  vector,  $G$ , and for any  $H$  in  $C$ , the quadratic form  $G'(H - H^*)G$  is nonnegative.

Now suppose that  $C$  is the class of covariance matrices obtained from the limiting distribution of  $\sqrt{n} [\hat{\theta}(Z_n) - \theta^*]$  for some regular estimate  $\hat{\theta}(Z_n)$ .

Definition 5. A regular estimate  $\hat{\theta}(Z_n)$  is said to be best asymptotically normal (BAN) if the covariance matrix of the limiting distribution of  $\sqrt{n} [\hat{\theta}(Z_n) - \theta^*]$  is minimal with respect to the class  $C$ .

The following theorem — first given by Ferguson (1958, Theorem 1 and 2) — was proved by Wijsman (1959a, Theorem 2). The theorem shows us how to generate a BAN estimate as the root of a linear form.

Ferguson's Theorem. Let the  $p \times N$  matrix  $B(Z, \theta)$  be: (i) continuous in  $\theta$  for each  $Z$ ; (ii) continuous in  $Z, \theta$  at each point  $(F(\theta), \theta)$ ; and (iii) such that  $B^*P^*$  is nonsingular whatever  $\theta^*$ . The matrix  $B^* = B(F(\theta^*), \theta^*)$ . Then there exists a neighborhood  $U^* \subset U$  and a function  $\hat{\theta}: U^* \rightarrow \Theta$ . The function  $\hat{\theta}$  is a regular estimate and satisfies the equation

$$B(Z, \theta) [Z - F(\theta)] = 0. \quad (3.3.4)$$

Furthermore, we have

$$\sqrt{n} (\hat{\theta}_n - \theta^*) \sim (B^*P^*)^{-1} B^* \sqrt{n} [Z_n - F(\theta^*)]. \quad (3.3.5)$$

If  $B^* = P^*V^*{}^{-1}$  then  $\hat{\theta}$  is BAN.

We will now find the covariance matrix of  $\hat{\theta}$  using (3.3.5) and

Definition 2. First note that since  $P^*$  has rank  $p$  and  $V^*$  is non-singular,  $B^* = (P^*)'V^{*-1}$  has rank  $p$  and  $B^*P^*$  is nonsingular. Now if we let the  $p \times N$  matrix  $A = (B^*P^*)^{-1}B^*$ , we see that the covariance matrix of (3.3.5) is given by the  $p \times p$  nonsingular matrix  $AV^*A'$ . Consequently there exists a matrix  $M$  such that  $MM' = AV^*A'$ , and it follows that  $(M^{-1}A)\sqrt{n} [Z_n - F(\theta^*)]$  has a limiting normal distribution with mean zero and covariance matrix

$$M^{-1}(AV^*A')(M^{-1})' = M^{-1}(MM')(M^{-1})' = I_p, \quad (3.3.6)$$

where  $I_p$  is a  $p \times p$  identity matrix. Then from Definition 2 we see that  $\hat{\theta}_n$  has an asymptotic covariance matrix which is given by

$$n^{-1}MM' = n^{-1}AV^*A' = n^{-1}(P^*'V^{*-1}P^*)^{-1}. \quad (3.3.7)$$

#### 3.4 Minimum Chi-Square Estimation

In this section we follow a general approach that was developed by Ferguson (1958) for finding BAN estimates. Using the notation of the previous section, the quadratic form

$$n [Z_n - F(\theta)]' V(\theta)^{-1} [Z_n - F(\theta)] \quad (3.4.1)$$

is called a chi-square, and the value of  $\hat{\theta}(Z_n)$  which minimizes it is called a minimum chi-square (MCS) estimate of  $\theta$ . If we let  $W(Z_n)$  be a  $p \times p$  positive definite symmetric matrix depending on  $Z_n$  only, then

$$n [Z_n - F(\theta)]' W(Z_n) [Z_n - F(\theta)] \quad (3.4.2)$$

is called a modified chi-square. Under the condition that

$W(Z_n) \rightarrow V(\theta^*)^{-1}$  as  $n \rightarrow \infty$ , and under the regularity conditions of

the previous section, the modified MCS estimate will be BAN. To minimize (3.4.2) we differentiate with respect to each of the parameters and set the result equal to zero. We obtain the following system of  $p$ -simultaneous linear equations:

$$n\mathbf{P}(\theta)'W(Z_n) [Z_n - F(\theta)] = 0 \quad . \quad (3.4.3)$$

Comparing this with (3.3.4) we see that (3.4.3) is a linear form with  $B(Z_n, \theta) = \mathbf{P}(\theta)'W(Z_n)$ . Since  $W(Z_n)$  converges in probability to  $V(\theta^*)^{-1}$ , and  $W(F(\theta^*)) = V(\theta^*)^{-1}$  we have

$$B(F(\theta^*), \theta^*) = B^* = \mathbf{P}(\theta^*)'V(\theta^*)^{-1} \quad .$$

This implies, by the second part of Ferguson's Theorem, that  $\hat{\theta}$ , the root of (3.4.3), is BAN.

To find  $\hat{\theta}$  we expand  $F(\theta)$  in a first-order Taylor series about an initial estimate  $\theta^0$  and obtain  $F(\theta) = F(\theta^0) + \mathbf{P}(\theta^0)(\theta - \theta^0)$ . Then replacing  $F(\theta)$  in (3.4.3) with this approximation we have

$$\mathbf{P}(\theta)'W(Z_n) [Z_n - F(\theta^0) - \mathbf{P}(\theta^0)(\theta - \theta^0)] = 0 \quad . \quad (3.4.4)$$

Since replacing  $B(Z_n, \theta) = \mathbf{P}(\theta)'W(Z_n)$  with  $B^0(Z_n, \theta) = \mathbf{P}(\theta^0)'W(Z_n)$  will result in a BAN estimate whenever the root to (3.4.3) is a BAN estimate (see Ferguson (1958), p. 1056) we obtain

$$\mathbf{P}(\theta^0)'W(Z_n)\mathbf{P}(\theta^0)(\theta - \theta^0) = \mathbf{P}(\theta^0)'W(Z_n)[Z_n - F(\theta^0)] \quad . \quad (3.4.5)$$

Now if we assume that the elements in the random vectors  $Y_j$  of observations are mutually independent, then  $V(\theta)$  will be a diagonal

matrix. In this situation  $W(Z_n)$  will be diagonal and (3.4.5) is then

$$\begin{bmatrix} p_{ir}^o \end{bmatrix}' \text{diag}(w_i) \begin{bmatrix} p_{is}^o \end{bmatrix} \begin{bmatrix} \theta_r - \theta_r^o \end{bmatrix} = \begin{bmatrix} p_{ir}^o \end{bmatrix}' \text{diag}(w_i) \begin{bmatrix} z_i - f(X_i, \theta^o) \end{bmatrix} , \quad (3.4.6)$$

which may be written as

$$\begin{bmatrix} \sum_i p_{ir}^o w_i p_{is}^o \end{bmatrix} \begin{bmatrix} \theta_r - \theta_r^o \end{bmatrix} = \begin{bmatrix} \sum_i p_{ir}^o w_i \begin{bmatrix} z_i - f(X_i, \theta^o) \end{bmatrix} \end{bmatrix} . \quad (3.4.7)$$

This is the same system of linear equations that was obtained in Section 3.2 when the Gauss-Newton method was used to develop an iterative method for obtaining LS estimates. Consequently, when the  $w_i$ 's are the reciprocals of the variances or consistent estimates of the variances, the weighted LS estimates will be BAN. In particular if the elements of the  $Y_j$ 's follow the Poisson distribution, then  $V(\theta) = \text{diag} [f(X_1, \theta), \dots, f(X_N, \theta)]$ , and we see that the linear form

$$n P(\theta)' V(\theta)^{-1} \begin{bmatrix} z_n - F(\theta) \end{bmatrix} = 0 \quad (3.4.8)$$

is of the type defined in (3.3.4). This is the same system of equations that was obtained in Section 2.3 (see eq. 2.3.4) using the ML principle with  $n_i = n$  for all  $i$ . We conclude that when the observations are Poisson the ML, LS, and MCS estimates obtained by using the method of scoring, the Gauss-Newton method, and the modified MCS approach, respectively, will be identical —i.e., the computational procedures result in the same system of simultaneous linear equations to be solved iteratively.

Also, we may use a result given in the previous section (see equation (3.3.7)) to obtain the covariance matrix of the BAN estimate obtained by minimizing the linear form (3.4.3). The covariance matrix

of  $\hat{\theta}$  is given by

$$n^{-1} \left[ P(\theta^*)' V(\theta^*)^{-1} P(\theta^*) \right]^{-1} . \quad (3.4.9)$$

The elements of this matrix may be estimated by replacing  $\theta^*$  by  $\hat{\theta}$ .

When the observations are assumed to be Poisson (3.4.9) is the information matrix, and for the LS case it is the inverse of the  $p \times p$  matrix on the left hand side of (3.2.6) evaluated at  $\hat{\theta}$ .

#### 4. REGRESSION ANALYSIS WHEN THE DATA ARE COUNTS

In Chapters 2 and 3 methods for estimating the parameters in a general regression model have been developed. It was assumed that the observations - i.e., the  $y_{ij}$ 's - were some type of counts that might be expected to follow the Poisson distribution, and that for a given value of  $X_i$  (the independent variables) the mean value of the  $y_{ij}$ 's was specified by the regression equation  $f(X_i, \theta)$ . In this chapter we will apply the estimation procedure to biological experiments which give rise to counts. After reviewing the biological aspects of the problem several examples involving both linear and nonlinear regression models will be considered. The examples are intended to illustrate how one defines the regression function and how one obtains results when the methods are applied to actual data. It is emphasized that the approach advocated here is intended to achieve conceptual rather than computational simplicity. If the biologist is familiar with regression analysis and the Poisson distribution he may assess the applicability of these methods to his particular problem. If the Poisson assumption is acceptable and a regression model established, then the LS estimation procedure may be used to estimate the parameters in the model. When the weights are defined in accord with the Poisson assumption the LS estimates will be identical to those obtained when the ML principle is used --as was demonstrated in Chapter 3. As a consequence of the iterative nature of nonlinear LS the computations required to solve the equations are difficult to carry out, and for

all but the simplest cases would require a digital computer. The computational procedure that we will use to obtain the LS estimates is outlined in Section 4.4. When the Poisson assumption is not acceptable the analysis may be modified by using empirically defined weights in the LS equations and the resulting estimates will be BAN as has been shown in Chapter 3.

#### 4.1 Applications In Biology

In a variety of experiments the biologist is interested in determining the concentration of microscopic particles in a suspension. The particles may, for example, be bacteria and their presence is established by allowing them to develop into visible colonies in a nutrient medium.

We now propose several definitions to simplify the discussion that follows and to provide a general frame of reference so that the biologist may consider the problem in a context most familiar to his own interests. Therefore, we shall speak of a 'count forming unit' (CFU) and a 'growth medium', the precise meaning of these terms being determined by the situation —for example, see Table 4.1. An experiment is usually carried out by selecting  $N$  values of the independent variable(s),  $X_i$ , and then for each value of  $X_i$  obtaining  $n_i$  'parallel' counts. A count is obtained by introducing a known amount of suspension which contains CFU's into the growth medium, and counting the colonies (plaques) that develop. The following assumptions are made:

- (i) the growth media are homogeneous,
- (ii) the suspensions of CFU's are assigned to growth media at random,

(iii) and each CFU results in only one visible colony (plaque). The first two assumptions depend on the experimental technique and the third depends upon the nature of the CFU's. The Poisson distribution has generally been accepted as a reasonable model for the distribution of colony counts since the early work of Fisher, Thornton, and MacKenzie (1922). In the examples that follow the counts are assumed to follow the Poisson distribution, and the ML principle is used to estimate the parameters in the regression model. All of the regression models that will be considered in this chapter have been discussed in Chapter 2, and the partial derivatives required in the computational procedure of Section 4.4 were defined in Section 2.4.

**TABLE 4.1**  
Experiments That Give Rise to Counts

Type of Particle	Units Counted	Nutrient Medium	References*
bacteria	colonies	agar	Fisher, Thornton and Mackenzie (1922)
virus	plaques	culture of cells	Berg <u>et al.</u> (1963)
bone marrow stem cells	colonies	spleen (of recipient mice)	Till and McCulloch (1961)

\* References for further discussion of experimental technique

#### 4.2 Linear Regression

In this section we consider the problem of estimating the concentration of CFU's per unit volume of suspension. Suppose that  $\theta_1$  is the mean density of CFU's per unit volume of suspension. The experimenter prepares  $N$  independent dilutions of the original suspension, and  $x_{i1}$ , the independent variable, is the  $i$ th dilution factor. The

regression model for this experiment takes the form

$$f(x_i, \theta) = \theta_1 x_{i1} \quad i=1, \dots, N, \quad (4.2.1)$$

and the observations are obtained by making  $n_i$  parallel counts at each dilution. Table 4.2 contains data from an experiment of this type that was originally presented by Berg et al. (1963), and further discussed by Roberts and Coote (1965, Table 3). Then using (2.1.4) we obtain

$$\hat{\theta}_1 = \frac{\sum_{i=1}^N \sum_{j=1}^{n_i} y_{ij}}{\sum_{i=1}^N n_i x_i} = 127.31 .$$

The simple linear regression model is inadequate —as was pointed out by Berg et al.. They attributed the deviations from the model to 'overcrowding and/or clumping'. Roberts and Coote proposed a method for dealing with the problem that uses sequential orthogonal comparisons to establish a region where (4.2.1) is adequate. An alternative approach proposed by Gart (1964) is to formulate a regression model with an additional quadratic term, i.e.,

$$f(x_i, \theta) = \theta_1 x_{i1} + \theta_2 x_{i2}, \quad \text{where } x_{i2} = x_{i1}^2 . \quad (4.2.2)$$

Although (4.2.2) is linear, it is not possible to obtain the ML estimates of  $\theta_1$  and  $\theta_2$  without using the iterative estimation procedure that was described in Chapter 2. Using the method of scoring to solve the likelihood equations (see Section 2.3) and the results established in Section 2.4.1, the ML estimates are found to be  $\hat{\theta}_1 = 162.12$  and  $\hat{\theta}_2 = -92.08$ . The covariance matrix of the estimated parameters is obtained using (2.4.5) and (2.5.1) and is

$$v(\hat{\theta}) = \begin{bmatrix} 74.89 & -169.7 \\ -169.7 & 432.7 \end{bmatrix} .$$

The regression sum of squares is

$$Q_r = \sum \left( \frac{n_i}{f(x_i, \theta)} (\bar{y}_{i\cdot} - f(x_i, \theta))^2 \right) = 14.05 ,$$

with 9 d.f.. The average of the observed counts for each value of  $i$ , and the expected values obtained using the linear and quadratic regression models are shown in Figure 4.1.

The purpose of this example has been to demonstrate the flexibility of the regression approach—i.e., various alternative models may be evaluated—provided a computer is available to do the arithmetic. For example, we could just as well consider a model which allows for a nonzero intercept, i.e.,

$$f(x_i, \theta) = \theta_1 x_{i1} + \theta_3 ,$$

or, more generally, both an intercept and a quadratic term

$$f(x_i, \theta) = \theta_1 x_{i1} + \theta_2 x_{i2} + \theta_3 ,$$

where  $x_{i2} = x_{i1}^2$ .

#### 4.3 Nonlinear Regression—Survival Curve Analysis

In Section 2.4 a general survival curve model was defined as follows:

$$f(x_i, \theta) = \theta_1 x_{i1} S(x_{i2}, \hat{\theta}) , \quad (4.3.1)$$

and several examples were discussed. In this section ML estimates of the parameters in the survival curve model are obtained. In each of

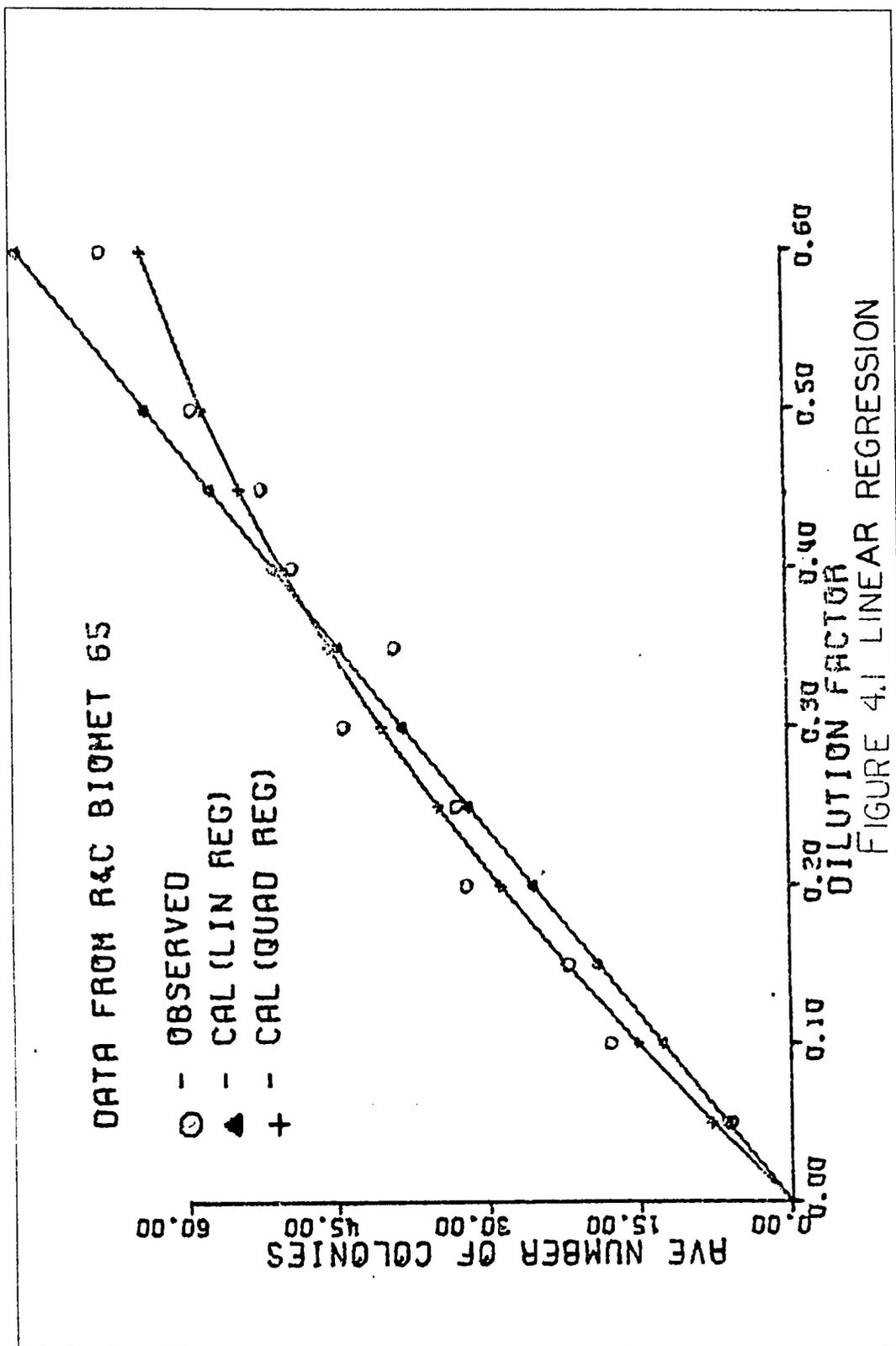


TABLE 4.2  
Plaque Counts\*

i	n <sub>i</sub>	x <sub>i1</sub>	x <sub>i2</sub>	Obs. ȳ <sub>i</sub> .	Expected	
					Linear	Quad
1	4	.05	.0025	6	6.4	7.9
2	5	.10	.0100	18	12.7	15.3
3	5	.15	.0225	22	19.1	22.2
4	5	.20	.0400	32	25.5	28.7
5	5	.25	.0625	33	31.8	34.8
6	5	.30	.0900	44	38.2	40.3
7	5	.35	.1225	39	44.6	45.5
8	4	.40	.1600	49	50.9	50.1
9	5	.45	.2025	52	57.3	54.3
10	3	.50	.2500	59	63.7	58.0
11	3	.60	.3600	68	76.4	64.1

SOURCE: Berg et al. 1963, Table 3.

NOTE: Symbols used in column heading are defined as follows:

n<sub>i</sub> = number of 'parallel' counts for ith dilution

x<sub>i1</sub> = ith dilution factor

x<sub>i2</sub> = x<sub>i1</sub><sup>2</sup>

ȳ<sub>i</sub> = total plaques counted / n<sub>i</sub>

The expected counts are calculated using (see eqs. 4.2.1 and 4.3.3):

Expected Linear =  $\hat{\theta}_1 x_{i1} = 127.31 x_{i1}$

Expected Quad =  $\hat{\theta}_1 x_{i1} + \hat{\theta}_2 x_{i2} = 162.12 x_{i1} - 92.08 x_{i2}$

\*Plaque counts obtained at different dilution levels for Coxsackie A9 virus.

these examples the observations are colony counts, and the CFU's are exposed to some type of ionizing radiation ( $x_{i2}$ ). In these situations the parameter  $\theta_1$  represents the concentration of CFU's/unit volume of suspension when no radiation is present (i.e.,  $x_{i2} = 0$ ).  $S(x_{i2}, \bar{\theta})$  represents the fraction of CFU's —per unit volume of suspension— surviving radiation dose  $x_{i2}$ . Interpretation of the parameters  $\bar{\theta} = (\theta_2, \dots, \theta_p)'$  will depend upon the particular model being considered. The parameters are intended to be descriptive of the effect of radiation of the CFU (i.e., bacteria or stem cell) that is being studied.

#### 4.3.1 Exponential Survival Curve

The exponential survival curve is obtained by letting

$$S(x_{i2}, \bar{\theta}) = \exp(-\theta_2 x_{i2})$$

so that

$$f(x_i, \theta) = \theta_1 x_{i1} \exp(-\theta_2 x_{i2}) \quad , \quad i = 1, \dots, N \quad . \quad (4.3.2)$$

We now consider an example presented by Lellouch and Wambersie (1966) in which E. Coli W 1485 were exposed to radiation doses ( $x_{i2}$ ) which are given in Table 4.3. For each radiation dose  $n_i$  dilutions ( $x_{i1}$  is the concentration for  $x_{i2}$ ) of the irradiated suspension are 'plated' on a petri dish, and  $y_{ij}$  is the number of colonies that result (see Table 4.3). In this situation the parameter  $\theta_2$  describes the radiosensitivity of the cell and its reciprocal is the dose at which 37% survival occurs. Now using the notation of Section 3 we calculate  $z_i = \sum_j y_{ij} / n_i$  and then plot  $\ln(z_i / x_{i1})$  vs.  $x_{i2}$  (see Figure 4.2a). Initial estimates of the parameters are  $\theta_1^0 = 271.0$  and  $\theta_2^0 = .5$ , where  $\theta_1^0 = z_i / x_{i1}$  and  $\theta_2^0$  is obtained graphically as

TABLE 4.3  
Exponential Survival Curve Data

i	$x_{i1}$	$x_{i2}$	$n_i$	$v_{ij}$					$z_i$	counts/unit vol		
										OBS	EXP	
1	1	0	6	299	283	280	246	264	254	271.0	271.0	271.3
2	1	1	2	169	184					176.5	176.5	166.5
3	2	2	5	179	224	188	202	194		197.4	98.7	102.2
4	4	3	5	233	261	229	286	264		254.6	63.6	62.8
5	10	4	4	401	410	356	388			388.75	38.9	38.5
6	4	4	5	157	146	134	161	159		151.4	37.8	38.5

SOURCE: Lellouch and Wambersie 1966, Table 1

NOTE: Column headings are defined as follows:

$x_{i1}$  = concentration of suspension plated

$x_{i2}$  = radiation dose (x4000 rads)

$n_i$  = number of counts made for the  $i$ th experimental condition

$$X_i = (x_{i1}, x_{i2})$$

$v_{ij}$  = number of colonies counted on the  $i$ th replication of  $X_i$

$$z_i = \sum_{j=1}^{n_i} v_{ij} / n_i$$

$$f(x_i, \hat{\theta}) = \hat{\theta}_1 x_{i1} \exp(-\hat{\theta}_2 x_{i2})$$

TABLE 4.4  
Chi-Square Values for Exponential Model

Source	df	$\chi^2$
Deviation from Model	4	$Q_r = 3.03$
Within	21	$Q_w = 30.24$
TOTAL	25	$Q_t = 33.27$

NOTE: The chi-square values are calculated using equations (2.5.3) and (2.5.4) and the data from Table 4.3

indicated in Figure 4.2. Then using the partial derivatives that were given in (2.4.7) and the iterative procedure that was described in Section 2.3 (see Section 4.4) the ML estimates are obtained after three iterations:  $\hat{\theta}_1 = 271.26$ ,  $\hat{\theta}_2 = .4879$ .

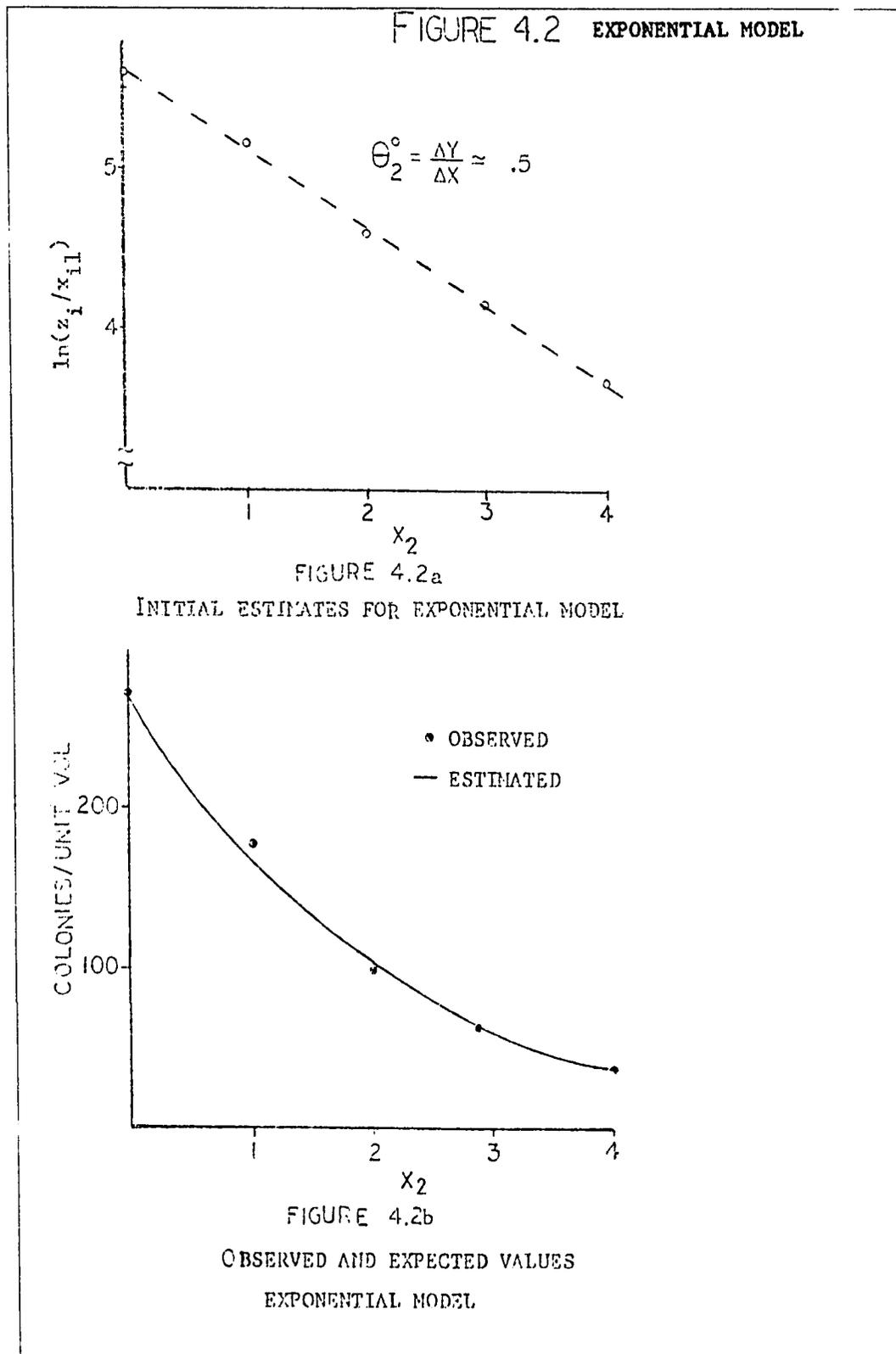
The chi-square statistics are calculated (see Section 2.5) and are displayed in Table 4.4. The elements of the covariance matrix are estimated using (2.5.1), and are available from the last iteration (i.e.,  $V(\hat{\theta}) = C^{-1}$ ) and

$$V(\hat{\theta}) = \begin{bmatrix} 35.79 & .03852 \\ & .6044 \times 10^{-4} \end{bmatrix} .$$

The observed and expected number of counts/unit volume are plotted against radiation dose in Figure 4.2b. From Table 4.4 we conclude that both the exponential survival curve and the Poisson distribution are acceptable —as indicated by the nonsignificant chi-square values— for this example. The estimated parameter values and variances are in close agreement with those obtained by Lellouch and Wambersie (1966, p.677) using a different iterative procedure which requires less computation but is only suitable for this particular model.

#### 4.3.2 Target Survival Curve

In some quantitative studies of cell survival the exponential model is inadequate, and the more general target model is required to describe the relation between radiation dose and the response of the biological system. We now consider a situation in which the target model adequately describes the survival curve data. In this example the CFU's are bone marrow stem cells. The bone marrow stem



cells are irradiated ( $x_{i2}$ ) and then injected into recipient animals. The injected cells locate in the spleen where the viable stem cells divide and produce clonal colonies. The recipient animal is sacrificed after sufficient time has elapsed for the colonies to grow to macroscopic size. The colonies in the excised spleen are then counted.

Since the stem cells constitute a small portion of the bone marrow cells, it is necessary to inject large numbers of bone marrow cells in order to produce colonies. It is convenient to define a unit concentration as  $10^5$  bone marrow cells. Then the  $i$ th experimental condition is given by  $X_i = (x_{i1}, x_{i2})$ , where  $x_{i2}$  = radiation dose and  $x_{i1}$  = concentration of cells injected into the recipient animals. The expected count is then

$$f(X_i, \theta) = \theta_1 x_{i1} \left[ 1 - \left( 1 - \exp(-\theta_2 x_{i2}) \right)^{\theta_3} \right], \quad (4.3.3)$$

and  $y_{ij}$ ,  $i = 1, \dots, N$ ,  $j = 1, \dots, n_i$ , are the observed counts. The number of spleens that are counted for each  $X_i$  is denoted by  $n_i$ .  $\theta_1$  is the concentration of stem cells in the bone marrow under normal conditions (i.e.,  $x_{i2} = 0$ ). The expression

$$s(x_{i2}, \theta) = 1 - \left( 1 - \exp(-\theta_2 x_{i2}) \right)^{\theta_3}$$

represents the fraction of stem cells surviving radiation dose  $x_{i2}$ . The parameter  $\theta_2$  describes the radiosensitivity of the cell, and  $\theta_3$  represents a threshold level or injury required for cell death. The parameter  $\theta_3$  was originally supposed to represent the number of 'targets' per cell and is also referred to as the 'extrapolation number'. Further discussion of the target model and the biological

TABLE 4.5  
Target Survival Curve Data\*

i	$x_{i1}$	$x_{i2}$	$n_i$	$y_{ij}$ 's						$z_i$	$s_i$	
1	1.25	0.0	6	11	10	11	11	9	8		10.00	1.000
2	1.75	.96	7	12	8	9	9	8	9	11	9.43	.673
3	3.00	1.92	4	11	10	11	14				11.50	.479
4	7.20	2.88	9	8	8	9	12	6	10	13	9.11	.158
				10	6							
5	24.0	4.32	11	12	12	14	10	7	10	8	9.55	.048
				11	8	7	6					
6	75.0	5.76	15	7	5	9	4	9	10	7	8.20	.014
				8	9	7	12	7	11	7		
				11								
7	120.0	6.72	4	2	3	3	4				3.00	.003

SOURCE: Till and McCulloch 1961 Table 3

NOTE: Column headings are defined as follows:

$x_{i1}$  = concentration of cells injected

$x_{i2}$  = radiation dose (rad  $\times 10^{-2}$ )

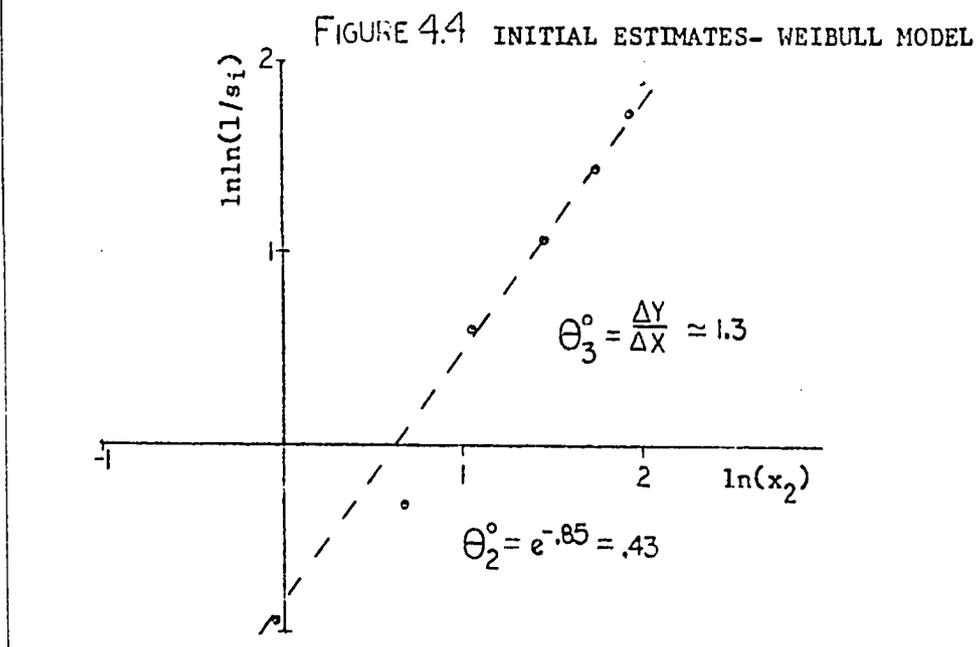
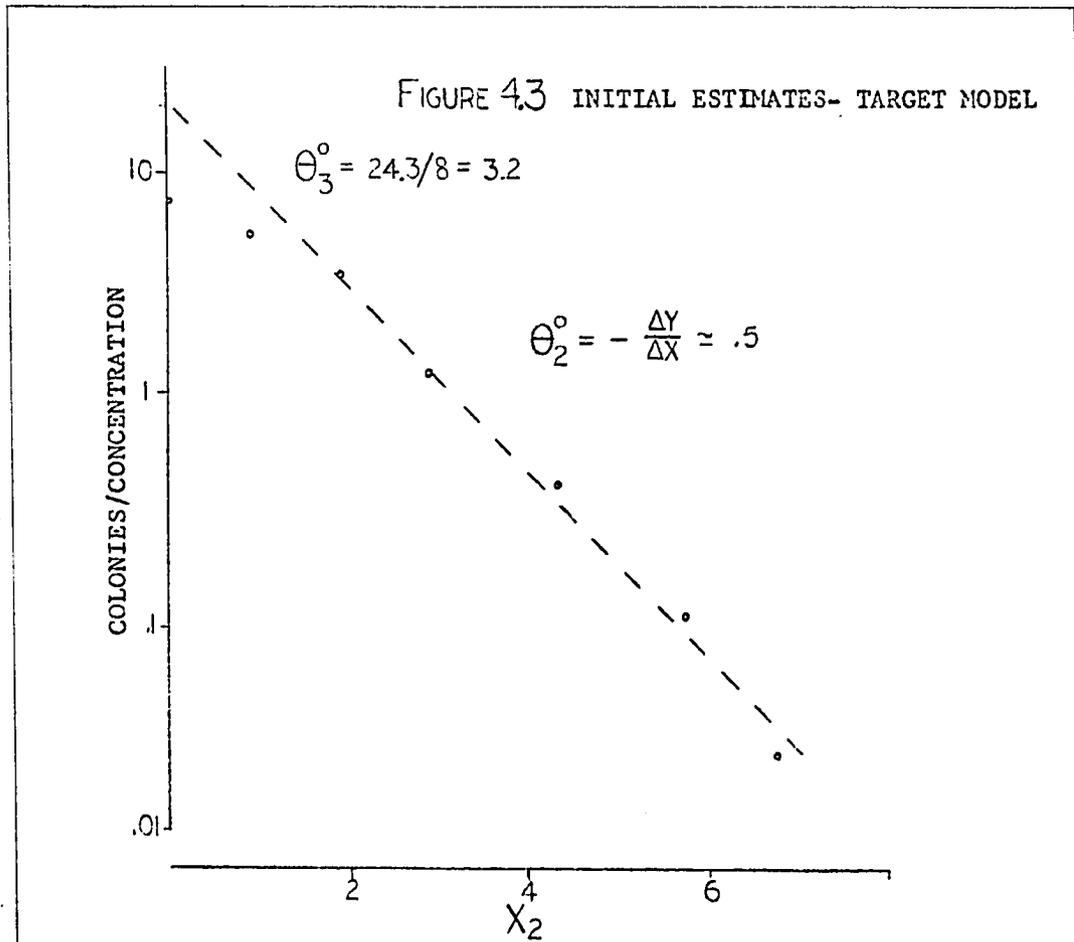
$n_j$  = number of spleens for  $i$ th experimental condition ( $X_i$ )

$y_{ij}$  = number of colonies counted on  $j$ th replication of  $X_i$

$z_i = \sum_j y_{ij} / n_i$

$s_i = z_i / (\theta_1^0 x_{i1})$ , i.e., fraction surviving

\*Bone marrow suspensions (from C57Bl mice) were irradiated in vitro with  $Co^{60}$  gamma-rays and data were obtained using spleen colony counting method.



significance of the parameters have been given by Fowler (1964), Krebs (1967), and Frome and Beauchamp (1968).

The data in Table 4.5 are from an experiment of this type.

An initial estimate of  $\theta_1$  is obtained using the counts at  $x_{i2} = 0$ ;  $\theta_1^0 = z_1/x_{i1} = 8.0$ . Initial estimates of  $\theta_2$  and  $\theta_3$  are obtained graphically by plotting  $\ln(z_1/x_{i1})$  versus  $x_{i2}$  (see Figure 4.3) and are  $\theta_2^0 = 1.0$  and  $\theta_3^0 = 3.1$ . Using the partial derivatives given in (2.4.11) the ML estimates are obtained after six iterations using the method of scoring (see Section 4.4). They are  $\hat{\theta}_1 = 7.636$ ,  $\hat{\theta}_2 = .9341$ , and  $\hat{\theta}_3 = 2.892$ . The estimated covariance matrix is

$$V(\hat{\theta}) = \begin{bmatrix} .8211 & -.0124 & -.5017 \\ & .0016 & .0254 \\ & & .5589 \end{bmatrix} .$$

The approximate chi-square tests (see Table 4.6) indicate that the regression model and the Poisson assumption are acceptable for this set of data. Further applications have been presented by Comas (1970).

#### 4.3.3 Weibull Survival Curve

The target model of Section 4.3.2 was originally obtained as a special case of a more general model derived from the 'target-hit' statistical theory of radiation damage when the number of hits is one. The same general equation has also been obtained from a kinetic model of radiation damage by Dines (1966). Although the target-hit theory remains in doubt as a representation of the lethal process in the cell, most mammalian cell survival curves resemble the target model and can be represented by the same parameters. That is to say, the target

model is accepted because experimentally obtained data 'fit the model', as was demonstrated in Section 4.3.3. We now consider the Weibull model

$$f(X_1, \theta) = \theta_1 x_{11} S(x_{12}, \theta) = \theta_1 x_{11} \exp(-\theta_2 x_{12}^{\theta_3}) \quad (4.3.4)$$

as an alternative to the target model, and use the stem cell survival curve data to illustrate how the ML estimates are obtained. First we suggest a possible interpretation of the parameters.  $\theta_1$  is the expected number of stem cells/unit concentration (i.e., has the same interpretation as  $\theta_1$  in the target model). Since  $S(x_2, \tilde{\theta})$  represents the fraction of cells surviving radiation dose  $x_2$  we consider the following 'destruction rate'

$$r(x_2) = \theta_2 \theta_3 x_2^{\theta_3 - 1}, \quad \theta_2, \theta_3 > 0. \quad (4.3.5)$$

The solution of the first-order differential equation

$$dS(x_2)/dx_2 = -r(x_2)S(x_2)$$

is

$$S(x_2) = \exp(-\theta_2 x_2^{\theta_3}).$$

If  $0 < \theta_3 < 1$ ,  $r(x_2)$  is a monotonically decreasing function of  $x_2$ ; if  $\theta_3 > 1$   $r(x_2)$  is a monotonically increasing function of  $x_2$ ; and if  $\theta_3 = 1$ ,  $r(x_2) = \theta_2$  (i.e., a constant) and the exponential survival curve results. When  $x_2 = 1$  then  $\theta_2$  is the natural logarithm of the fraction of cells surviving at unit dose, i.e.,  $S(1) = \exp(-\theta_2)$ .

Initial estimates are obtained by first calculating  $\theta_1^0 = z_1/x_{11}$ , and then plotting  $\ln \ln(1/s_1)$  vs.  $\ln(x_{12})$  where  $s_1 = z_1/(\theta_1^0 x_{11})$  (i.e.,  $s_1$  is the observed fraction surviving/unit concentration). Initial

estimates of  $\theta_2$  and  $\theta_3$  are then found graphically as shown in Figure 4.4, and are  $\theta_2^0 = .43$  and  $\theta_3^0 = 1.3$ . The ML estimates are obtained after four iterations and are  $\hat{\theta}_1 = 8.134$ ,  $\hat{\theta}_2 = .4206$ , and  $\hat{\theta}_3 = 1.341$ . The estimated covariance matrix is

$$v(\hat{\theta}) = \begin{bmatrix} .7954 & .0512 & -.0573 \\ & .0052 & -.0064 \\ & & .0081 \end{bmatrix} .$$

The regression sum of squares is 7.105 as compared with 7.595 obtained for the target model. The expected number of counts for the Weibull model and for the target model are given in Table 4.7. Figure 4.5 shows the original data and the fitted curves. Since  $\hat{\theta}_3 > 1$  the destruction rate is an increasing function of radiation dose, and  $\ln \hat{\theta}_2 = .657$  is the expected fraction surviving a dose of 100 rads. It is also possible to develop a stochastic version of the Weibull model in which the 'destruction rate' is interpreted as the 'face of mortality' in a pure death process (see Chiang 1968, p.60).

#### 4.4 Computations

In Chapter 2 and 3 it was shown that under certain conditions ML, LS, and MCS estimates were equivalent in the sense that they result in the same iterative procedure. The computational procedure that we have used will now be described in the general context of nonlinear LS. That is, we wish to minimize

$$S(\theta) = \sum_{i=1}^N w_i (z_i - f(x_i, \theta))^2 \quad (4.4.1)$$

with respect to  $\theta_r$ ,  $r = 1, \dots, p$ . In (4.4.1)  $z_i = \sum_{j=1}^{n_i} y_{ij} / n_i$ ,

TABLE 4.6  
Chi-Square Values for Target Model

Source	df	$\chi^2$
Deviation from Model	4	7.595
Within	49	24.442
Total	53	32.037

TABLE 4.7  
Observed and Expected Values\*

$x_{i2}$	Number of colonies per unit volume		
	Observed	Target	Weibull
0.	8.0	7.64	8.13
.96	5.39	5.96	5.46
1.92	3.84	3.13	2.97
2.88	1.27	1.40	1.43
4.32	.398	.384	.408
5.76	.109	.101	.100
6.72	.025	.041	.036

NOTE: Radiation dose =  $x_{i2} \cdot 100$  rads

\* Observed values are obtained from Table 4.5 ( $z_i/x_{i1}$ ). The expected values are calculated using (4.3.3) and (4.3.4) and estimated values of the parameters (see Sections 4.3.2 and 4.3.3).

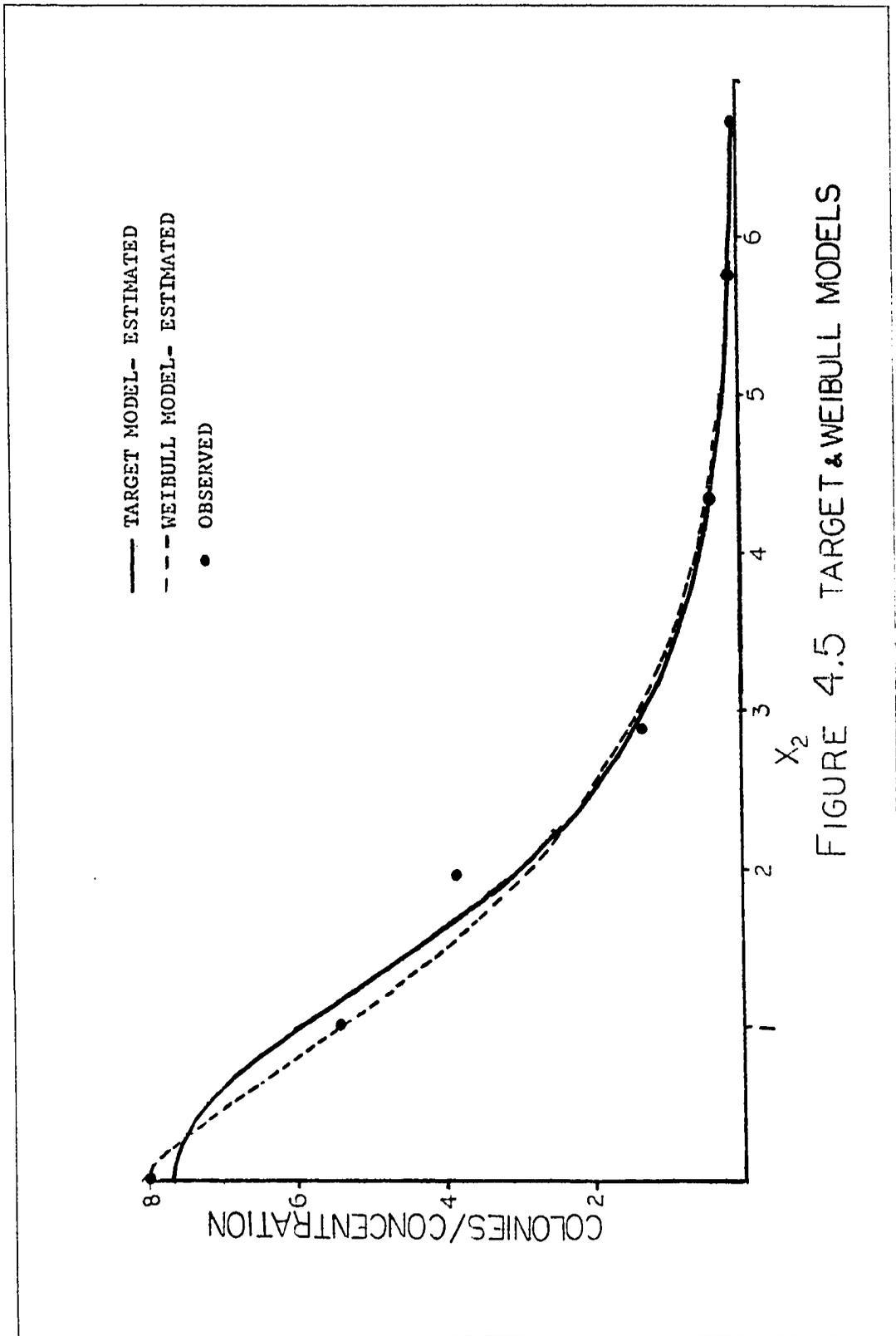


FIGURE 4.5 TARGET &amp; WEIBULL MODELS

$f(X_i, \theta)$  is the regression function, and  $w_i$  is the weight associated with the  $i$ th observation (see Section 3.2). The procedure used to minimize (4.4.1) is iterative, that is, if  $\theta^k$  is the value of  $\theta$  on the  $k$ th iteration, we compute a search vector  $D^k$  and set  $\theta^{k+1} = \theta^k + D^k$ . The procedure is repeated until a stable solution,  $\hat{\theta}$ , is reached according to some convergence criterion. Various approaches to the numerical problem based on gradient methods have been discussed by Smith and Shanno (1971).

The gradient of  $S$  is written in matrix notation as

$$\nabla S = 2(WP)' [Z - F(\theta)] \quad , \quad (4.4.2)$$

where  $W = \text{diag}(w_1, \dots, w_N)$ ,  $Z = (z_1, \dots, z_N)'$ ,  $F(\theta) = [f(X_1, \theta), \dots, f(X_N, \theta)]'$  and  $P = [p_{ij}] = [\partial f(X_i, \theta) / \partial \theta_j]$ . A gradient method for minimizing  $S$  is one which calculates a search vector,  $D = (d_1, \dots, d_p)'$ , defined by

$$D = A(WP)' [Z - F(\theta)] \quad (4.4.3)$$

where  $A$  is a  $p \times p$  matrix. When  $A = (P'WP)^{-1}$ ,  $D$  is the Gauss-Newton vector, and if the starting value,  $\theta^0$ , is good this procedure will have desirable convergence properties. If good initial estimates are not available, then some other search procedure such as Marquardt's (1963) may be useful. Marquardt's vector combines the best features of steepest descent and the Gauss-Newton method, avoiding singularities that may occur when  $P'WP$  is ill-conditioned.

Another approach to function optimization that does not use a gradient method has been developed by Bremermann (1970) using a

random search technique so that only the function being minimized need be evaluated and no derivatives are required. Bremermann's method may prove to be of considerable value in large scale problems (i.e., many parameters) when no starting values are available. A method for generating (pseudo random) normal deviates --which are required in Bremermann's program-- on 32-bit-word computers has been discussed by Chen (1971).

In the applications in Sections 4.2 and 4.3 methods for finding starting values were given for the particular problem being considered and the Gauss-Newton method was used to minimize  $S(\theta)$ . In the examples considered the starting values, final estimates of the parameters, number of iterations to convergence, and estimated covariance matrix were given. Convergence was defined to have occurred when the relative change in all of the parameters was less than  $10^{-5}$ . The computations were carried out in single precision arithmetic, and the Gauss-Jordan method was used for matrix inversion.

In general if a stable solution,  $\hat{\theta}$ , is found, it will be the ML estimate of  $\theta$  when the observations are counts that follow the Poisson distribution. When the number of parameters is large double precision arithmetic and special care in matrix inversion are advisable. In such situations Marquardt's algorithm would be desirable unless good initial estimates of the parameters are available. The following is a summary of the steps in the computational procedure:

1. [Input]  $N, m, p, X_i, z_i, n_i, (i=1, \dots, N)$  as defined in Section 3.2;  $\theta_r, r=1, \dots, p$ , starting values for

- parameters; convergence criteria ( $\epsilon$ ), and maximum number of iterations.
2. [Define regression function and derivatives] The regression function (see Section 2.2) and partial derivatives with respect to the parameters are defined by  $f_i = f(X_i, \theta)$ ,  $p_{ir} = \partial f(X_i, \theta) / \partial \theta_r$ ,  $r = 1, \dots, p$ .
  3. [Define weights] Under the Poisson assumption weights are defined iteratively by  $w_i = n_i / f_i$ , or as described in Section 3.2.
  4. [Calculate C and G] C is  $p \times p$  (symmetric) matrix and G is the  $p \times 1$  vector both of which appear in the system of equations defined in (3.2.6). They are defined by  $C = [\sum_i w_i p_{ir} p_{is}]$ ,  $G = [\sum_i w_i p_{ir} (z_i - f_i)]$ , for  $r, s = 1, \dots, p$ , and the subscript  $i$  assumes the values  $1, 2, \dots, N$ .
  5. [Solve linear equations] Obtain the 'correction' vector  $D = (d_1, \dots, d_p)'$ , where  $D = C^{-1}G$ . This is the most important step in the program and requires considerable care in large problems. If the matrix C is nearly singular Marquardt's algorithm or some other procedure may be used.
  6. [Check for convergence] If the  $(|d_r| / |\theta_r|) < \epsilon$  for  $r = 1, \dots, p$ , or the maximum number of iterations has been reached, go to step 8.
  7. [Update estimates] Put  $\theta_r = (\theta_r + d_r)$ ,  $r = 1, \dots, p$ , then return to step 2.
  8. [Output] Output should include all input data, final

parameter values, number of iterations, estimated variance-covariance matrix ( $C^{-1}$ ), minimum value of sum of squares in (4.4.1) and maximum value of logarithm of the likelihood function --which is given by (2.3.3) when  $y_{ij}$ 's are independent Poisson counts.

A Fortran IV implementation of this computational procedure will be given in the Appendix.

## 5. SPECTRUM ANALYSIS

The primary purpose of this chapter is to summarize the computational procedures and statistical considerations of numerical spectrum analysis. Most of the results presented here have been discussed in detail in Spectral Analysis and Its' Application by Jenkins and Watts (1968). Jenkins and Watts were primarily concerned with engineering applications of spectral analysis, where the methods have a rather natural appeal as an extension of Fourier transform techniques. In a biomedical context, spectral analysis has been extensively used in electroencephalography (Brazier 1965), and in electrocardiography (Murthy, et al. 1971). Randall (1958) developed an approach to the analysis of pulsatile pressure and flow (in the femoral artery of the dog) using cross-spectral analysis. Tick and Woodbury (1965) have suggested that an important function of spectral techniques in biomedical signal processing is in the development of data reduction systems. In the applications to be presented we hope to demonstrate the potential value of spectrum analysis, not only as a method of data reduction, but also as an intermediate step in the development of patient monitoring techniques and in furthering the understanding of the biological phenomena being considered. The inherent nonlinearity and nonstationarity of many biological processes makes a formal mathematical treatment virtually impossible. In Chapter 6 a simple mechanical model of the respiratory system will be developed which serves to illustrate this point. The model will be considered primarily as a

conceptual aid in the interpretation of the spectrum analysis. When the frequency domain has a natural interpretation the nonparametric spectral approach has great intuitive appeal.

Other methods of time series analysis that follow a parametric time domain approach have been developed by Box and Jenkins (1970). Their approach is based on a mixed autoregressive-moving average model where estimation is accomplished via nonlinear regression methods similar to those described in Chapter 3. Hannan (1969) has also considered this type of model and has developed an estimation procedure based upon a Fourier transformation of the data. This approach will be most useful if the purpose of the analysis is to predict future values of the series. However, it appears to be rather difficult to attach any biological meaning to the parameters in this kind of model. Jones et al. (1970) has presented an application of this approach to multivariate biological time series analysis.

In Section 5.1 of this chapter a review of the basic theory of bivariate time series analysis will be presented. In discussing the theory only the discrete time case will be considered. It will be assumed that the data is obtained by sampling continuous signals at equi-spaced intervals of time. Practical considerations and problems encountered in analog to digital conversion are discussed in detail by Bendat and Piersol (1971, chap 7), and in biological applications by Macy (1965). The choice of an appropriate sampling rate is of great practical importance since it determines the Nyquist

frequency. If the sampling interval is too large aliasing will occur in the sample spectrum.

The method of estimation that will be used in the applications is based on the periodogram and is outlined in Section 5.2. The periodogram is computed from the finite Fourier transform (FT) of a realization of a time series. If viewed as a linear transformation the finite FT can be used to develop a heuristic treatment of the statistical properties of spectrum estimators. The development of the Cooley-Tukey fast FT algorithm has had a dramatic effect on spectrum analysis and digital signal processing. The calculations of spectrum analysis will be reviewed in Section 5.3.

### 5.1 Spectrum Analysis of a Bivariate Stationary Time Series

Consider the sequences  $X(t)$ ,  $Y(t)$ , of real-valued random variables, where the indexing variable  $t$  is integer valued and denotes time. It is assumed that  $X(t)$  and  $Y(t)$  are weakly stationary time series with  $EX(t) = EY(t) = \mu$  for all values of  $t$ , and for notational convenience we assume  $\mu=0$ . The autocovariance and cross-covariance functions are defined by

$$\begin{aligned} R_x(k) &= E[X(t)X(t+k)] \quad , \\ R_y(k) &= E[Y(t)Y(t+k)] \quad , \\ R_{yx}(k) &= E[X(t)Y(t+k)] \quad , \quad k = 0, \pm 1, \pm 2, \dots \quad (5.1.1) \end{aligned}$$

The spectra of  $X$  and  $Y$  are defined to be the Fourier transforms of their autocovariance functions and are given by

$$S_x(f) = \sum_{k=-\infty}^{\infty} R_x(k) \exp(-2\pi i f k) \quad , \quad -\frac{1}{2} \leq f < \frac{1}{2} \quad ,$$

$$S_y(f) = \sum_{k=-\infty}^{\infty} R_y(k) \exp(-2\pi i f k) \quad , \quad -\frac{1}{2} \leq f < \frac{1}{2} \quad . \quad (5.1.2)$$

The cross-spectrum is defined to be the Fourier transform of the cross-covariance function and is given by

$$S_{yx}(f) = \sum_{k=-\infty}^{\infty} R_{yx}(k) \exp(-2\pi i f k) \quad , \quad -\frac{1}{2} \leq f < \frac{1}{2} \quad . \quad (5.1.3)$$

Both  $S_x(f)$  and  $S_y(f)$  will be real-valued and symmetric about zero since they are Fourier transforms of real even functions. The cross-covariance function is not necessarily symmetric about the origin so the cross-spectrum will in general be complex-valued.

Since the spectrum and the autocovariance function form a transform pair it follows that

$$R_x(k) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \exp(2\pi i f k) S_x(f) df \quad . \quad (5.1.4)$$

If we set  $k=0$  in the above equation we see that

$$R_x(0) = E[X^2(t)] = \sigma_x^2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} S_x(f) df \quad . \quad (5.1.5)$$

Consequently,  $S_x(f)$  is interpreted as representing the distribution of the process variance with respect to frequency. In a similar manner we may invert (5.1.3) to obtain

$$R_{yx}(k) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \exp(2\pi i k f) S_{yx}(f) df, \quad (5.1.6)$$

and letting  $k=0$  in (5.1.6) we find that

$$R_{yx}(0) = \int_{-\frac{1}{2}}^{\frac{1}{2}} S_{yx}(f) df = E[X(t)Y(t)]. \quad (5.1.7)$$

If  $X(t)$  represents the potential difference across a two terminal device, and  $Y(t)$  the resulting current then (5.1.7) is the expected value of the power delivered to the device (see Papoulis 1965, chap. 10).

To gain some further insight into equations (5.1.2) and (5.1.3) it is useful to consider the spectral representations of the processes  $X(t)$  and  $Y(t)$  which are given by (see e.g. Cox and Miller 1965, chap. 8)

$$X(t) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \exp(2\pi i f t) dZ_x(f), \quad \text{and}$$

$$Y(t) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \exp(2\pi i f t) dZ_y(f), \quad (5.1.8)$$

where  $dZ_x(f)$  and  $dZ_y(f)$  are complex-valued processes of orthogonal increments and are also cross-orthogonal. Using (5.1.8) in (5.1.1) and comparing the result with (5.1.4) and (5.1.6) we find that

$$S_x(f) df = E[|dZ_x(f)|^2],$$

$$S_y(f) df = E[|dZ_y(f)|^2], \quad \text{and}$$

$$S_{yx}(f) df = E[dZ_y(f) dZ_x^*(f)], \quad (5.1.9)$$

where \* indicates complex conjugate. The cross-spectrum may be written in terms of its real (cospectrum) and imaginary (quadrature spectrum or 'quadspectrum') components as  $S_{yx}(f) = C_{yx}(f) + iQ_{yx}(f)$  where  $i$  is the complex number  $(0,1)$ . From the cross-spectrum we obtain the amplitude spectrum

$$|S_{yx}(f)|^2 = C_{yx}^2(f) + Q_{yx}^2(f) \quad , \quad (5.1.10)$$

and the phase spectrum

$$\phi_{yx}(f) = \arctan [Q_{yx}(f)/C_{yx}(f)] \quad . \quad (5.1.11)$$

Then following Priestly (1971), we write

$$dZ_x(f) = |dZ_x(f)| \exp(i\phi_x(f)), \quad dZ_y(f) = |dZ_y(f)| \exp(i\phi_y(f)),$$

and suppose that  $|dZ_x(f)|$  is independent of  $\phi_x(f)$  and that  $|dZ_y(f)|$  is independent of  $\phi_y(f)$ . Then the cross-spectrum is given by

$$\begin{aligned} S_{yx}(f) \exp i\phi_{yx} &= E [dZ_y(f) dZ_x^*(f)] \\ &= E [ |dZ_y(f)| \cdot |dZ_x(f)| ] E [ \exp i(\phi_y(f) - \phi_x(f)) ] \quad . \end{aligned} \quad (5.1.12)$$

Consequently, the phase spectrum  $\phi_{yx}(f)$  may be interpreted as the average value of the phase shift  $\phi_y(f) - \phi_x(f)$  between the components  $X(t)$  and  $Y(t)$  at frequency  $f$ . The squared coherency spectrum between  $X(t)$  and  $Y(t)$  is defined by

$$\rho^2(f) = |S_{yx}(f)|^2 / S_x(f) S_y(f) \quad , \quad (5.1.13)$$

and the coherency may be interpreted as a correlation coefficient defined at each frequency.

## 5.2 Spectral Estimation

The use of the smoothed periodogram in spectral analysis was reviewed by Jones (1965) who pointed out the advantages of this approach, and suggested that its apparent neglect was attributable to computational considerations. In the same year the now famous fast FT algorithm was published by Cooley and Tukey (1965), and as a consequence any computational limitations imposed by the periodogram were eliminated. In this section the univariate case is reviewed, and in Section 5.3 we will deal with smoothing techniques and cross-spectrum analysis. The purpose of this discussion is to describe the computational procedure that will be used in the following chapters, and to provide some insight into the statistical aspects of this type of data analysis. A general treatment of the mathematical aspects of multiple time series analysis, and in particular the finite FT and spectral analysis has been presented by Hannan (1970).

The periodogram of a finite realization of a time series  $X(t)$   $t=0,1,\dots,N-1$ , is obtained by computing

$$J(f) = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} X(t) \exp(2\pi i f t / N) \quad , \quad f=0,\dots,n, \quad (5.2.1)$$

and then calculating

$$I(f) = |J(f)|^2 \quad , \quad f=1,\dots,n \quad , \quad (5.2.2)$$

where  $n$  equals  $N/2$  if  $N$  is even and  $(N/2)-1/2$  if  $N$  is odd. It is well known (Bartlett 1966, chap. 9) that when the  $X(t)$ 's are independent normal random variables with mean zero and variance  $\sigma^2$ ,

then the random variable  $2I(f)$ ,  $0 < f < n$ , has an exponential distribution. Consequently, the fluctuations in the  $I(f)$ 's are on the same order of magnitude as their expectations. It can be further shown that a similar result holds for linear processes in general, i.e., the  $I(f)$ 's have asymptotically independent exponential distributions with means equal to the spectral density function at the frequencies  $f/N$ . A heuristic demonstration of this result is presented at the end of this section. When the  $I(f)$ 's are plotted against  $f$  (as is usually done in this type of analysis), the resulting erratic behavior of the graph suggests that a smoothing operation would be desirable. A method first suggested by Daniell (1946) is to average the periodogram over adjacent frequencies. An alternative method proposed by Bartlett (1966) is to divide the original series into  $m$  subseries each of length  $N/m$ , and average the estimates obtained from each subseries. More generally, smoothed spectral estimators may be calculated from

$$\hat{S}(f) = \sum_{k=1-N}^{N-1} w(k) \hat{R}(|k|) \exp(-2\pi i f k / N) ,$$

where

$$w(k) = \begin{cases} 1 - |k| / (N/m) , & k \leq N/m \\ 0 & \text{otherwise} , \end{cases}$$

and

$$\hat{R}(k) = (N-k)^{-1} \sum_{t=0}^{N-k-1} X(t+k)X(t) . \quad (5.2.3)$$

The  $w(k)$ 's are referred to as the 'lag window', and their Fourier transform is called the 'spectral window'. When  $m$  is large this approach has an obvious computational advantage. It also produces

consistent estimators of the spectrum. Further investigation of the above mentioned methods of spectrum smoothing has been provided by Jenkins and Watts (1968, chap. 6). If in (5.2.3) we take  $m=1$  it is possible to show that

$$\hat{S}(f) = \sum_{k=1-N}^{N-1} \left[ \frac{1}{N} \sum_{t=0}^{N-k-1} X(t+k)X(t) \right] \exp(-2\pi i f k / N) = I(f), \quad (5.2.4)$$

where  $I(f)$  is the periodogram as defined in (5.2.2). The development of the fast FT has renewed interest in direct computation of the periodogram followed by smoothing in the frequency domain. Bartlett (1967) has reviewed the advantages of this approach as well as the effects of nonnormality and nonstationarity on spectrum analysis. It is also possible to use the periodogram to develop hypothesis testing procedures which are useful when a parametric representation of a linear process is being considered --see Davis (1968) and Durbin (1969).

To gain some insight into how the finite FT occurs in a natural way in time series analysis we consider the case of a circularly defined stationary time series with  $E X(t) = 0$  for all  $t$  (Hannan 1960, chap 1). Here, the stationarity assumption implies that

$$R(k) = E [X(t+k)X(t)] = E [X(s+t+k)X(s+t)] \quad , \\ \text{for } k = 0, \dots, N-1 \quad , \quad (5.2.5)$$

where  $s$  takes on the values  $0, \pm 1, \pm 2, \dots$ . If the  $X(t)$ 's are Gaussian then a realization of the process will have a multivariate normal distribution with mean vector zero and covariance matrix

$$R = \begin{bmatrix} R(0) & R(1) & R(2) & \dots & R(N-2) & R(N-1) \\ R(-1) & R(0) & R(1) & \dots & R(N-3) & R(N-2) \\ R(-2) & R(-1) & R(0) & \dots & R(N-4) & R(N-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ R(2-N) & \dots & \dots & \dots & R(0) & R(1) \\ R(1-N) & R(2-N) & \dots & \dots & R(-1) & R(0) \end{bmatrix}. \quad (5.2.6)$$

From (5.2.5) it follows that  $R$  is a circulant since it satisfies  
(see Bellman 1960, p.242)

$$R(j,k) = R(j+1,k+1), \quad (5.2.7)$$

where  $j$  and  $k$  are integer valued indexing variables and are reduced mod  $N$ , and  $R$  is an  $N \times N$  matrix of complex numbers. The eigenvalues of  $R$  are

$$\beta_j = \sum_{k=0}^{N-1} R(k) \exp(2\pi i j k / N), \quad j = 0, \dots, N-1, \quad (5.2.8)$$

and the corresponding eigenvectors are

$$B_j = \frac{1}{\sqrt{N}} [\exp(2\pi i j k / N)]^T, \quad \begin{matrix} k = 0, \dots, N-1, \\ j = 0, \dots, N-1. \end{matrix} \quad (5.2.9)$$

Then observing that for  $j = 0, \dots, N-1$ ,

$$B_j^{*'} = [1/\sqrt{N}] [\exp(-2\pi i j k / N)] \quad , \quad k = 0, \dots, N-1, \quad (5.2.10)$$

it follows that  $B^{-1} = B^{*'}$  where  $B$  is the  $N \times N$  matrix with columns  $B_j$ , and  $B^{*'}$  is the  $N \times N$  matrix with rows  $B_j^{*'}$ . Further it follows that

$$B^{-1} R B = \text{diag}(\beta_0, \dots, \beta_j, \dots, \beta_{N-1}) \quad , \quad (5.2.11)$$

and that  $B$  is a unitary matrix since  $B^{*' } B = I$ , where  $I$  is the  $N \times N$  identity matrix. Consequently we see that the eigenvectors of a

circulix are independent of the circulix and are given by the columns of the matrix B. The above results and other interesting properties of matrices of this type have been discussed by Good (1950). He further notes that when R is real and symmetric the eigenvalues are real and that  $\beta_j = \beta_{N-j}$  where

$$\beta_j = \begin{cases} R(0) + 2 \sum_{k=1}^{n-1} R(k) \cos(2\pi jk/N) + (-1)^j & (N \text{ even}) \\ R(0) + 2 \sum_{k=1}^n R(k) \cos(2\pi jk/N) & (N \text{ odd}) \end{cases} \quad (5.2.12)$$

Now consider the linear transformation  $J = HX$ , where

$$H = \begin{bmatrix} B'_1 \\ \vdots \\ B'_j \\ \vdots \\ B'_n \end{bmatrix}, \quad \text{and} \quad X = \begin{bmatrix} X(0) \\ \vdots \\ X(t) \\ \vdots \\ X(N-1) \end{bmatrix}. \quad (5.2.13)$$

The rows of H are the transposed eigenvectors defined in (5.2.9) and n was defined in (5.2.2). Since X is multivariate normal with mean zero and covariance matrix R, it follows that  $E(J) = 0$  and that

$$\text{Var}(J) = E(JJ^{*'}) = E(HXX^{*'}H^{*'}) = HRH^{*'} = \text{diag}(\beta_1, \dots, \beta_n). \quad (5.2.14)$$

Then recalling the definition of the periodogram (5.2.2), i.e.,

$$I(f) = |J(f)|^2$$

and using (5.2.14) it follows that  $I(f)/\beta_f$ ,  $f = 1, \dots, n$ , will be mutually independent chi-square random variables with 2 d.f..

In general if  $X(t)$ ,  $t = 0, \pm 1, \pm 2, \dots$ , is a noncircularized discrete Gaussian time series with mean zero and covariance sequence

$$R(k), k = 0, \pm 1, \pm 2, \dots, \text{ where } \sum_{k=-\infty}^{\infty} |R(k)| < \infty, \quad (5.2.15)$$

then the joint distribution of the  $J(f)$ 's  $f = 0, \dots, n$ , where

$$J(f) = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} X(t) \exp(2\pi i f t / N), \quad (5.2.16)$$

can be approximated (as  $N \rightarrow \infty$ ) by the distribution of  $n$  complex normally distributed random variables. The  $J(f)$ 's are mutually independent with zero mean and  $\text{Var } J(f) \simeq S_x(f)$ . Since  $I(f) = |J(f)|^2$  it follows that the joint distribution of the  $I(f)$ 's can be approximated by that of independent scaled chi-square variables. Further discussion of the above results and investigation of the asymptotic distributional properties of the periodogram have been given by Davis (1968).

### 5.3 Computations

A historical introduction to the computations of numerical spectrum analysis has been given by Tukey (1967). The development of the fast Fourier transform (FT) algorithm (Cooley and Tukey, 1965) has dramatically reduced the computational effort of spectrum analysis. It has also stimulated interest in the theory of the finite FT. Cooley, Lewis, and Welch (1969) have used the finite FT to develop a very simple and illuminating derivation of the fast FT algorithm. They also derived a number of other widely used results (e.g., Parseval's Theorem) and we use their definition of the finite FT in this section. The finite FT of  $N$  finite valued complex numbers

$X(t)$ ,  $t = 0, \dots, N-1$ , is given by

$$Z_x(f) = \frac{1}{N} \sum_{t=0}^{N-1} X(t) \exp(-2\pi i f t / N) = \frac{1}{N} \sum_{t=0}^{N-1} X(t) W_N^{-ft},$$

$$f = 0, \dots, N-1, \quad (5.3.1)$$

where  $W_N = \exp(2\pi i / N)$ . The sequence  $X(t)$  can be expressed as the inverse finite FT of  $Z_x(f)$ , i.e.,

$$X(t) = \sum_{f=0}^{N-1} Z_x(f) W_N^{tf}, \quad t = 0, \dots, N-1. \quad (5.3.2)$$

The sequences  $X(t)$  and  $Z_x(f)$  are referred to as a transform pair.

For mathematical purposes the finite sequences are extended periodically to all the integers as follows:

$$X(t) = X(kN + t),$$

$$Z_x(f) = Z_x(kN + f), \quad (5.3.3)$$

and  $k = 0, \pm 1, \pm 2, \dots$  in (5.3.3). The finite sequences are always recovered by considering the values of the infinite sequence at the points  $0, 1, \dots, N-1$ .

It is possible to compute the transform by using equation (5.3.1). This requires  $N^2$  operations (an operation being a complex multiplication and addition). When  $N$  is an integer power of two the Cooley-Tukey algorithm requires  $N \log N$  operations. In the applications that will be presented  $N$  will be an integer power of two. The fast FT algorithm for arbitrary  $N$  (mixed radix) has been discussed by Singleton (1969), who has developed a Fortran subroutine to carry out the computations. Singleton (1969) has also provided a bibliography of the fast FT. It

should be noted that when the sequence being transformed is real-valued an additional step is required for efficient use of the Cooley-Tukey algorithm. When transforms are being computed for two real-valued sequences (as will be the case in Chapter 6) the Cooley-Tukey algorithm is efficiently used by defining a new series as follows:

$$W(t) = X(t) + iY(t) \quad , \quad t = 0, \dots, N-1 \quad ,$$

where  $X(t)$  and  $Y(t)$  are real-valued. The transform of  $W(t)$ , namely  $Z_w(f)$ , is then computed and the transforms of  $X(t)$  and  $Y(t)$  are recovered by using the complex conjugate symmetry of real-valued series.

That is, since

$$Z_x(f) = Z_x^*(N-f) \quad ,$$

$$Z_y(f) = Z_y^*(N-f) \quad ,$$

we may use the linearity of the finite FT to obtain

$$\begin{aligned} Z_x(f) &= \left( Z_w(f) + Z_w^*(N-f) \right) / 2 \\ Z_y(f) &= \left( Z_w(f) - Z_w^*(N-f) \right) / 2i \quad . \end{aligned} \quad (5.3.4)$$

One method of estimating the spectra and cross-spectrum for two real-valued data sets is now summarized. Let  $X(t)$  and  $Y(t)$ ,  $t = 0, \dots, N-1$ , denote the data. First we compute  $Z_x(f)$  and  $Z_y(f)$ .

Then we compute

$$\begin{aligned} \hat{S}_x(f) &= \frac{1}{M} \sum_k |Z_x(k)|^2 \quad , \\ \hat{S}_y(f) &= \frac{1}{M} \sum_k |Z_y(k)|^2 \quad , \\ \hat{S}_{yx}(f) &= \frac{1}{M} \sum_k Z_y(k) Z_x^*(k) \quad , \end{aligned} \quad (5.3.5)$$

where  $f = 1, \dots, N'$ , and  $N' = \left[ \frac{N}{2} \right] - m/M$ .  $[A]$  denotes the largest positive integer that is not greater than  $A$ , and  $m = (M-1)/2$  where  $M$  is taken to be a positive odd integer. For fixed  $f$  the index  $k$  in (5.3.5) takes on the integer values  $(M \cdot f - m), \dots, (M \cdot f + m)$ . When  $f=0$  in (5.3.5) then  $k=1, \dots, m$ , and  $M$  is replaced by  $m$ . When  $M' = \left( \left[ \frac{N}{2} \right] - m \right) \bmod M$  is not equal to zero an additional estimate is obtained for  $f = N' + 1$ , and is given as

$$\begin{aligned}\hat{S}_x(N'+1) &= \frac{1}{M'} \sum_k |Z_x(k)|^2, \\ \hat{S}_y(N'+1) &= \frac{1}{M'} \sum_k |Z_y(k)|^2, \\ \hat{S}_{yx}(N'+1) &= \frac{1}{M'} \sum_k Z_y(k) Z_x^*(k),\end{aligned}\quad (5.3.6)$$

where  $k = M \cdot N' + 1, \dots, \left[ \frac{N}{2} \right]$ . If  $N$  is even then  $Z_x(N/2)$  and  $Z_y(N/2)$  are each multiplied by  $\frac{1}{2}$ .

Estimates of the cospectrum and the quadrspectrum are obtained from the real and imaginary components of  $S_{yx}(f)$ , i.e.,

$$\begin{aligned}\hat{C}_{yx}(f) &= \text{Real } \hat{S}_{yx}(f), \\ \hat{Q}_{yx}(f) &= \text{Imag } \hat{S}_{yx}(f), \quad f = 0, \dots, N'.\end{aligned}\quad (5.3.7)$$

Estimates of the phase spectrum and coherency are then given by

$$\hat{\phi}_{yx} = \arctan(\hat{Q}_{yx}(f)/\hat{C}_{yx}(f)), \quad (5.3.8)$$

and

$$\hat{\phi}_{yx}^2(f) = |\hat{S}_{yx}(f)|^2 / \hat{S}_y(f) \hat{S}_x(f). \quad (5.3.9)$$

To clarify the above procedure —which will be referred to as method D (since it is similar to the method first proposed by P.J. Daniell)— we consider the univariate case. From Parseval's Theorem

we have

$$\frac{1}{N} \sum_{t=0}^{N-1} X(t)^2 = \sum_{f=0}^{N-1} |Z_x(f)|^2 . \quad (5.3.10)$$

Since  $Z(0) = (\bar{X}, 0)$  we write (5.3.10) as

$$\text{var}(X) = \frac{1}{N} \sum_{t=0}^{N-1} (X(t) - \bar{X})^2 = \sum_{f=1}^{N-1} |Z_x(f)|^2 . \quad (5.3.11)$$

Since  $X(t)$  is real-valued,  $Z_x(f) = Z_x^*(N-f)$  so we may write (5.3.11) as

$$\text{var}(X) = \begin{cases} \frac{N-1}{2} \sum_{f=1}^{N-1} |Z_x(f)|^2 , & N \text{ odd} \\ 2 \sum_{f=1}^{\frac{N}{2}-1} |Z_x(f)|^2 + |Z_x(N/2)|^2 , & N \text{ even} . \end{cases} \quad (5.3.12)$$

Recalling the definition of the periodogram (5.2.2) we see that

$I(f) = N |Z_x(f)|^2$ . To illustrate how method D works we take  $N = 32$  and

$M = 5$  and obtain  $N' = 2$ ,  $m = 2$ ,  $M' = 4$ , and

$$\hat{S}_x(0) = \frac{1}{2} \sum_{k=1}^2 |Z_x(k)|^2 ,$$

$$\hat{S}_x(1) = \frac{1}{5} \sum_{k=3}^7 |Z_x(k)|^2 ,$$

$$\hat{S}_x(2) = \frac{1}{5} \sum_{k=8}^{12} |Z_x(k)|^2 ,$$

$$\hat{S}_x(3) = \frac{1}{4} \left[ \sum_{k=13}^{15} |Z_x(k)|^2 + \frac{1}{2} |Z_x(16)|^2 \right] .$$

From (5.3.12) we see that

$$\frac{1}{2}\text{var}(X) = m \cdot \hat{S}(0) + M \sum_{f=1}^{N'} \hat{S}_x(f) + M' \hat{S}_x(N'+1) \quad (5.3.13)$$

so that  $S_x(f)$  is the proportion of the  $\text{var}(X)$  in the  $f$ th frequency band.

In (5.3.5) through (5.3.9) the scale factor was chosen to maintain the identity given in (5.3.13), and the sampling interval was assumed to be unity. In a particular problem the appropriate scaling may be introduced by indicating a scale factor for the ordinate when the estimated quantities are plotted against frequency in cycles/unit time.

Another straight forward smoothing procedure that utilizes the computational advantage of the fast FT we call method B (since it is essentially the method that was first proposed by Bartlett). Let  $X_k(t)$ ,  $Y_k(t)$ , denote the  $N$  data values obtained on the  $k$ th realization of a bivariate time series;  $t=0, \dots, N-1$ , and  $k=1, \dots, M$ . Then for each  $k$  we compute the finite FT's  $Z_x(k,f)$ ,  $Z_y(k,f)$  which are then used to obtain

$$\begin{aligned} \hat{S}_x(f) &= \frac{1}{M} \sum_{k=1}^M |Z_x(k,f)|^2, \\ \hat{S}_y(f) &= \frac{1}{M} \sum_{k=1}^M |Z_y(k,f)|^2, \\ \hat{S}_{yx}(f) &= \frac{1}{M} \sum_{k=1}^M Z_y(k,f) Z_x^*(k,f). \end{aligned} \quad (5.3.14)$$

The cospectrum, quadrspectrum, phase, and coherency may then be calculated as in (5.3.6) - (5.3.8). From (5.3.14) we see that method B consists of averaging at a fixed frequency over  $M$  realizations of the process. In method D we average over  $M$  adjacent frequencies of a single realization.

It has been pointed out by Tick (1967) that the estimation of functions of the cross-spectrum (i.e., coherency) presents problems (e.g., badly biased estimates). Further discussion concerning bias and variability of phase and coherency estimates and methods for obtaining confidence intervals has been given by Jenkins and Watts (1968, chap 9). More recently results concerning the estimation of coherency have been obtained by Jones (1969) and Hannan and Thompson (1971).

## 6. POWER SPECTRUM OF THE RESPIRATORY SYSTEM

In the discussion of respiration that follows, many important factors will not be explicitly represented in the model that is developed. An accurate biophysical model of the respiratory process would necessarily be complicated. One attempt at such a model of the human respiratory system (RS) has been presented by Milhorn (1966, chap. 16). The model is based on chemical factors in the blood ( $\text{CO}_2$ ,  $\text{H}^+$  Ion, and  $\text{O}_2$  concentrations) that affect aveolar ventilation. A three compartment model of the respiratory system and a number of other assumptions are made to obtain a 'steady-state' model. It would appear that the primary motivation for the assumptions is the resulting system of first-order linear differential equations, which are convenient for analog simulation.

Milhorn's analog computer model does not directly include the dynamic aspect of respiration, namely breathing. Although ventilation is eventually affected by chemical factors in the blood, the direct effect of these and other factors will be on the transpulmonary pressure via the respiratory muscles. If an individual's ability to maintain homeostasis under stress is of interest, then the physical characteristics of the pulmonary system are of obvious importance. These characteristics include lung elasticity, airway resistance, and the diffusing capacity of the lung. Consequently, any attempt to describe respiration should consider the mechanical properties of the respira-

tory system (RS), as well as the chemical factors in the blood which are indirectly controlling the process. Further discussion of these ideas and their clinical application (i.e. diagnosis of pulmonary disease) can be found in The Lung (Comroe et al., 1962). An introductory discussion of respiration —which emphasizes the application of respiratory physiology to normal man exposed to stress— has been presented by Dejours (1966).

The use of the concept of feedback control systems in representing the human body was originally proposed over a century ago by the French physiologist, Claude Bernard. Bernard's idea of a constant 'internal environment' was later developed by the American physiologist Walter Cannon, and a new concept —homeostatic mechanism— emerged. The concept of homeostasis is currently used in experimental medicine. One can think of disease in terms of the failure of homeostatic mechanisms which lead to the instability of the controlling system. The principle of feedback control in simple mechanical systems and instruments has been traced back over two thousand years by Mayr (1970). In recent years feedback control theory has been successfully used to solve many practical engineering problems. Norbert Wiener (1949) developed the mathematics that is the basis of modern communications and control theory. His theory is based on a statistical characterization of the response of a system to nondeterministic inputs. The system is represented by its transfer function and the input by its spectral density function. Wiener's contributions —e.g. the design of optimum linear systems— to communication and

control engineering have been reviewed by Lee (1964). Wiener (1951) also suggested that the concept of feedback control may be useful in the design of biomechanical control systems. One example of such a system is a mechanical anesthetist that would regulate the depth of anesthesia. Continuously monitored physiological variables —that reflect the current level of anesthesia— would be used to control the administration of anesthetic.

Although the concept of homeostasis is a very appealing one, recent research indicates that biological rhythms may be important in maintaining the health of an individual, and understanding disease —see e.g., Volume 2 (1971), Number 2 and 3 of the Journal of Interdisciplinary Cycle Research. In a review of chronobiology Halberg (1969) has dealt with various quantitative methods that have been used in the analysis of biological time series. Further discussions of the methodological considerations involved in a parametric approach to the analysis of bioperiodicity have been presented by Bliss (1970, chap. 17) and Sollberger (1970).

In the preceding discussion we have attempted to summarize some of the important concepts that have influenced the method of data analysis that will now be developed. In the rest of this chapter we will primarily be concerned with a mechanical model of the RS. We have pointed out that there are many important factors that will not be explicitly represented in the RS model. Their influence is, however, present at all times through their control of the transpulmonary pressure which is the 'driving force' of the RS. In Section 6.1

a differential equation model of the RS will be discussed. In Section 6.2 we will propose cross-spectrum analysis of pressure and flow time series recordings as a data analytic technique that is useful in summarizing and describing this kind of data. The model developed in Section 6.1 will be used to motivate the definition of the power spectrum that will be proposed in Section 6.2. An example will be presented with the intention of illustrating the implementation of the computational procedure that was described in Chapter 5, and to demonstrate how graphical methods may be used to summarize the results of the analysis.

### 6.1 Mechanical Model of the Respiratory System

As a first step in obtaining a model of the RS consider Figure 6.1. Here we represent  $V(t)$  —the volume of the lungs at time  $t$ — as the output of the RS which is being driven by  $P(t)$  - the transpulmonary pressure. How we proceed from here depends upon the purpose of the analysis and the variables that may be conveniently measured.

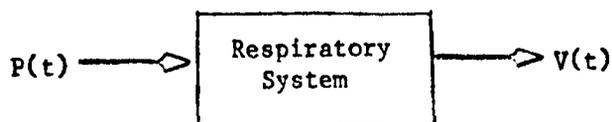


Figure 6.1

A general mechanical model of the pressure-volume relation in the RS has been presented by Mead (1961). According to Mead, the RS may be described using Newton's third law of motion, i.e., in terms

of its position (volume), velocity (air flow), and acceleration. The transpulmonary pressure is the force applied to the system, and the opposing forces which develop may be attributed to physical properties of the RS which are elasticity ( $F_1$ ), resistance ( $F_2$ ) and inertia ( $F_3$ ). Although it may not be possible to specify how these properties are distributed through the system, they are represented in the 'equation of motion' of the RS. In its most general form this equation is given by

$$F_1 [V(t)] + F_2 [dV(t)/dt] + F_3 [d^2V(t)/dt^2] = P(t) \quad (6.1.1)$$

In equation (6.1.1)  $P$  is the pressure difference between the pleural space and the mouth or nose. It is equal to the sum of the opposing pressures developed by the RS which are represented by  $F_1$ ,  $F_2$ , and  $F_3$ . A detailed discussion of the assumptions made to obtain this model and further restrictions which will specify the form of the functions  $F_1$ ,  $F_2$ , and  $F_3$  has been given by Mead (1961). One simpler form of the model is obtained by letting

$$\begin{aligned} F_1 [V(t)] &= B_1 (V(t) - V_0) \quad , \\ F_2 [V'(t)] &= B_2 V'(t) + B_3 [V'(t)]^2 \quad , \\ \text{and } F_3 [V''(t)] &= B_4 V''(t) \quad . \end{aligned} \quad (6.1.2)$$

In (6.1.2)  $V_0$  represents end expiratory volume, and the  $B$ 's are parameters that characterize the system. In Mead's notation  $B_1 = 1/C$  ( $C$  is the 'compliance');  $B_2 = K_1$ ,  $B_3 = K_2$ , where  $K_1$  and  $K_2$  describe the resistive property of the system; and  $B_4 = I$  is an inertial parameter.

During quiet breathing it is assumed that flow resistance is nearly linear and the inertial pressure is negligible which leads to

$$B_1 V(t) + B_2 V'(t) = P(t) + B_1 V_0 \quad (6.1.3)$$

In (6.1.3)  $B_1$  is the reciprocal of the compliance and  $B_2$  is the resistance (i.e.,  $B_1 = 1/C$  and  $B_2 = R$ ). A detailed discussion of the clinical measurement and interpretation of these parameters, their range of values in healthy individuals, and their use in diagnosing respiratory disease and evaluating a patient's response to treatment have been given by Comroe et al. (1962).

Although the model defined by (6.1.3) is only a first approximation, it can be used to describe the RS. Stacy and Peters (1965) have developed a method that can be used to calculate lung compliance and airway resistance using transpulmonary pressure and air flow recordings. They also calculate the mechanical 'work' of respiration which is defined by

$$\text{work} = \int P(t) V'(t) dt \quad (6.1.4)$$

It should be noted that this definition does not depend on the assumptions that were made to obtain (6.1.3). Since it is easily calculated by either analog or digital methods, the work concept may be very useful in detecting changes in a particular individual, which is an important practical problem when a patient is being continuously monitored in an intensive care unit. This of course requires the measurement of both air flow and transpulmonary pressure, which may not be practically possible.

The second-order differential equation given in (6.1.2) may be written as a two dimensional system of differential equations

$$\begin{aligned} V' &= G_1(V, V', t) \\ V'' &= G_2(V, V', t) \quad , \end{aligned} \quad (6.1.5)$$

which describes the behavior of the system when no external disturbances (i.e., forcing function =  $P(t)$ ) are present. McShane (1969) has presented some general mathematical results for systems governed by differential equations of this type for a wide class of forcing functions (which includes both Lipschitzian and Brownian-motion processes). The forcing function  $P(t)$  is stochastic in nature and any mechanical model of the RS based on differential equations can be considered as a special case of (6.1.5).

## 6.2 Spectral Analysis of the Respiratory System

In this section we assume that the pressure difference between the mouth and alveoli and the rate of air flow into the lungs are the variables of interest. We let  $X(t)$  denote the pressure difference in cm H<sub>2</sub>O and  $Y(t)$  the flow rate in liters/minute. The product of pressure and flow has the physical dimension of power. Since a centimeter of water (4°C) is (approximately) equal to 98.1 newton/meter<sup>2</sup> we obtain the power at time  $t$  in units of .0981 joules/minute. The quantity

$$\frac{1}{T} \int_0^T X(t)Y(t)dt \quad , \quad (6.2.1)$$

then represents the mechanical energy required to overcome the resistive

component of the RS per unit time (i.e., we may think of  $X(t)Y(t)$  as the power dissipated in overcoming the resistance to air flow).

We now consider an example in which time recordings were obtained from an adult male. The pressure difference was calculated by subtracting esophageal pressure from airway pressure —which are measured with pressure transducers. A Fisher-Porter flowmeter was used to calibrate the air flow measurements which are obtained with a differential pressure flowmeter. The data were originally recorded on two channels of an FM data tape at the Anesthesia Research Laboratory, Emory University Medical School. The continuous pressure and flow time series were then sampled using the A/D conversion facilities at the Emory University Computing Center. Following the procedure described in Chapter 5 we let  $X(t)$ ,  $Y(t)$ ,  $t = 0, \dots, N-1$  denote the values obtained in the A/D conversion. In this example  $N = 1028$  and the sampling rate is 240 points/minute. The first two minutes of the pressure and flow time series are shown in Figure 6.2. The auto and cross-covariance functions were defined in (5.1.1) and Figures 6.3 a-c show the sample estimates obtained for the pressure-flow data.

The pressure spectrum, flow spectrum, and pressure-flow cross-spectrum are estimated using Method D of Section 5.3 (see eq. 5.3.5) with a smoothing span of  $M = 15$ . The spectral window has a band width of approximately 3.5 (cycles/min). The spectrum estimates have 30 degrees of freedom (except for  $f=0$  and  $f=N'+1$ ) and are uncorrelated. The pressure and flow spectra are plotted versus frequency in Figure 6.4. The estimated cospectrum and quadrspectrum are plotted

versus frequency in Figures 6.5 a&b. The phase spectrum is then obtained using (5.3.7) and is shown in Figure 6.5c. The coherency function is then estimated using (5.3.8) and is displayed in Figure 6.5d.

The results obtained in this analysis indicate that the two signals are highly correlated in the frequency range where normal breathing occurs —i.e., 10-30 cyc/min (see Figure 6.5d). Figure 6.5c shows a negative phase shift over this range, indicating that pressure is 'leading' flow. The cross-covariance function at lag zero — $R_{yx}(0) = 7.1(\text{cmH}_2\text{O})(\text{L}/\text{min}) = .7 \text{ joules}/\text{min}$ — represents the mechanical work done in overcoming the resistive component of the RS. The cospectrum in Figure 6.4 shows how the power is distributed with respect to frequency (see Section 5.1). Comparing Figure 6.5a with the estimated spectra in Figure 6.4 we see that aside from a scale factor the graphs are quite similar. All of this leads us to hypothesize that the pressure-flow relation is approximately linear —the constant of proportionality being the parameter R in (6.1.3). If this assumption is made the flow spectrum will be a 'power' spectrum. In practice the linearity assumption will only be a crude approximation, —in fact we would expect R to be frequency dependent. Whether the pressure-flow relation is linear or not the flow spectrum does provide a useful quantitative method for describing the RS. This suggests that the flow variance —i.e.,  $\frac{1}{N} \sum_t Y(t)^2$ — should be an important statistic to consider when relative changes in the RS are of interest. It may, for example, be desirable to detect changes in human respiratory

function that portend the onset of fatigue or other symptoms that are of clinical importance. Such a situation will be encountered in Chapter 7 when we consider spectrum analysis of the pulmonary impedance pneumograph.

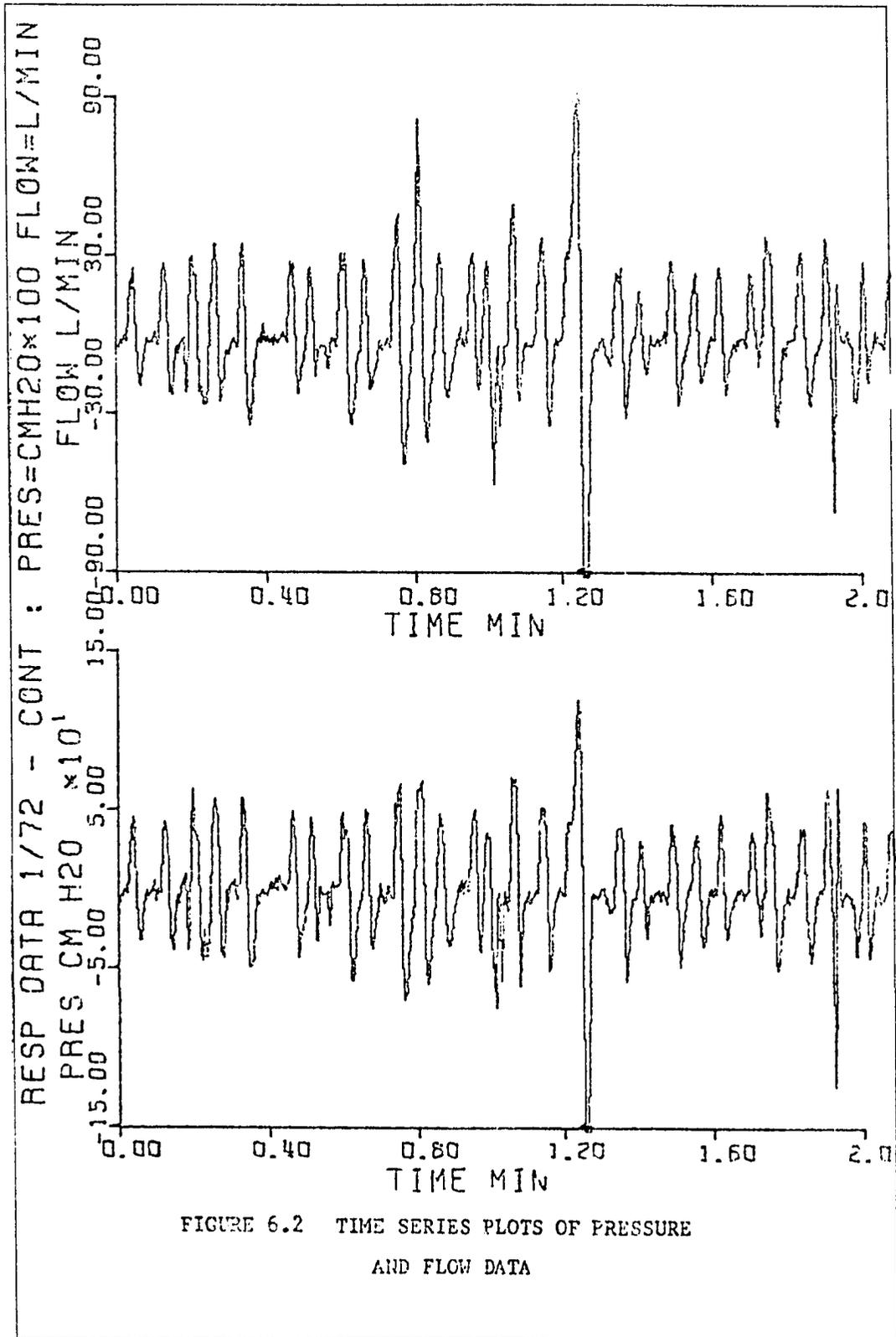


FIGURE 6.2 TIME SERIES PLOTS OF PRESSURE  
AND FLOW DATA

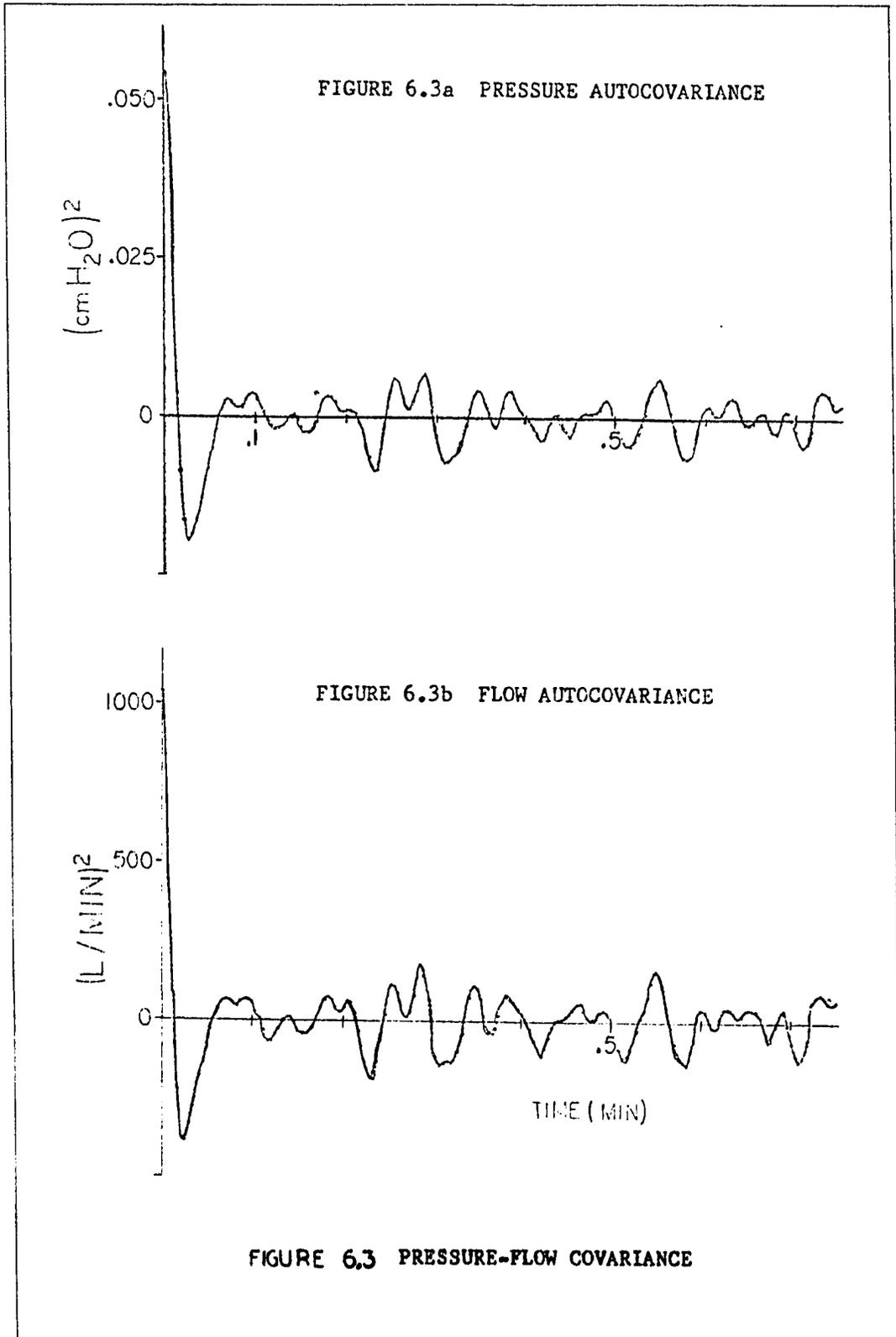


FIGURE 6.3 PRESSURE-FLOW COVARIANCE

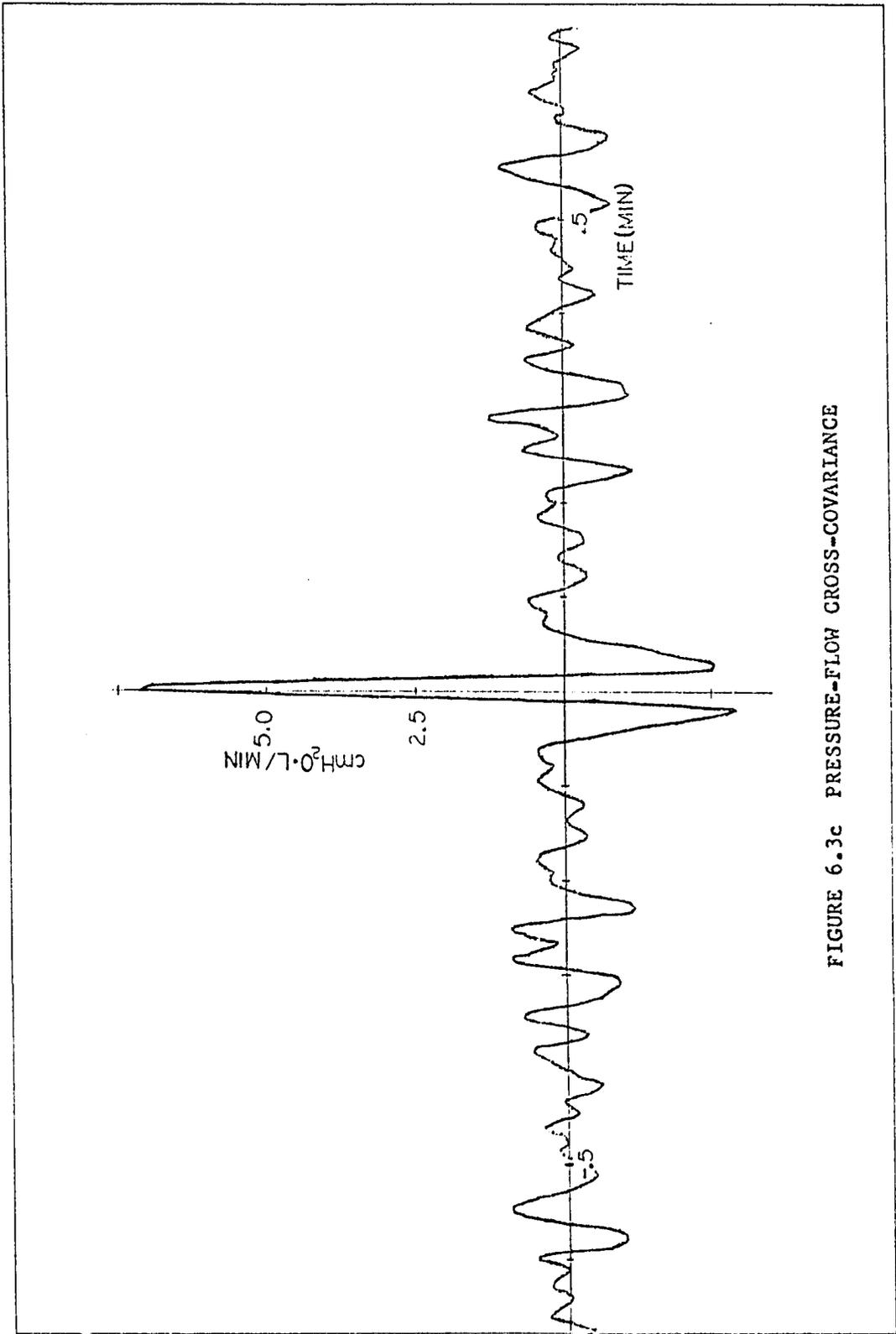


FIGURE 6.3c PRESSURE-FLOW CROSS-COVARIANCE

FIGURE 6.4a PRESSURE SPECTRUM

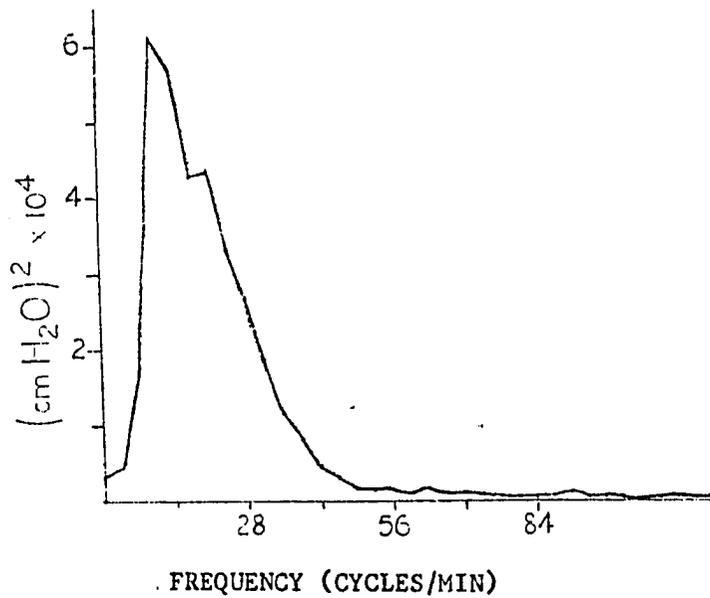


FIGURE 6.4b FLOW SPECTRUM

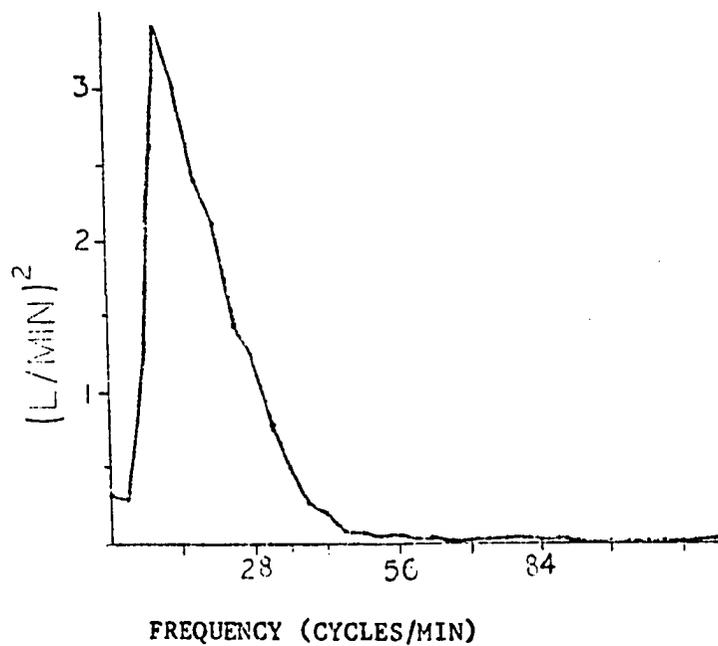


FIGURE 6.4 PRESSURE-FLOW SPECTRA

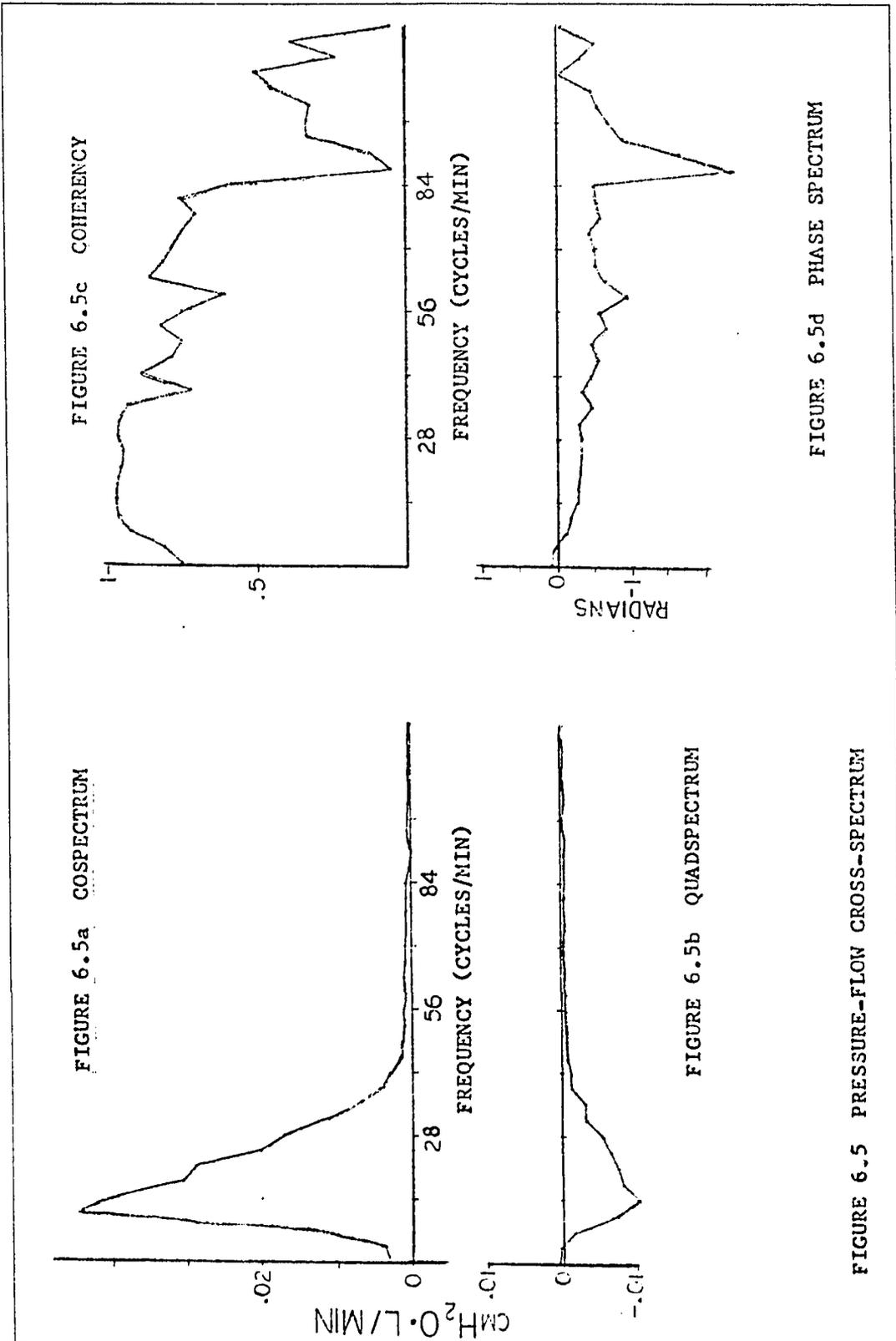


FIGURE 6.5 PRESSURE-FLOW CROSS-SPECTRUM

FIGURE 6.5a COSPECTRUM

FIGURE 6.5b QUADSPECTRUM

FIGURE 6.5c COHERENCY

FIGURE 6.5d PHASE SPECTRUM

## 7. SPECTRAL ANALYSIS OF THE PULMONARY IMPEDANCE PNEUMOGRAPH

The pulmonary impedance pneumograph (PIP) is a bioimpedance recorder that is used for the indirect measurement of respiration. Pacela (1966) has reviewed the development of impedance pneumography and the various instrumentation techniques that have been developed. If a base line calibration is made the thoracic impedance can be related to respiratory volume. While this calibration is important in the study of respiratory physiology it is not essential when the pulmonary impedance pneumograph is used to monitor the respiratory function of remotely located persons over extended periods of time. The latter situation occurs, for example, during therapeutic exposure to total-body radiation at low dose rates or during manned-space flights. In these situations clinically useful information can be obtained from the PIP tracings by counting the number of respiratory cycles per minute and observing their amplitude and regularity.

In Section 7.1 a statistical method (spectrum analysis) for analyzing the analog tracings that are obtained in impedance pneumography will be reviewed. This PIP spectrum has been used as part of a computerized data reduction system (see Ricks et al. 1972). In this situation the PIP spectrum reflects the onset and course of gastrointestinal distress induced pharmacologically or by total-body irradiation. This method for quantitating respiratory effort also appears to provide a useful method for evaluating the decrease in exercise capability which reflects increased fatigue occurring during bicycle ergometry.

In Section 7.2 we will use the results of Chapter 6 to propose a modification of the PIP spectrum analysis. The new analysis will utilize the first difference of the PIP tracing which provides a measure of flow rather than volume. We claim that the spectrum obtained will be a 'power' spectrum. Examples will be provided.

### 7.1 Pulmonary Impedance Pneumograph Spectrum

The electronic physiologic monitoring system used to obtain the pulmonary impedance data that are analyzed here (and in Section 7.2) has been described by Morris, Barclay and Lushbaugh (1967). Briefly, the pulmonary impedance coupler (Beckman) produces an analog signal which is obtained from surface electrodes attached to the subject. The signal can be analyzed in real time or recorded on tape for retrospective analysis. The analog signal is sampled at a rate of 256 points per minute by an IBM-1800 computer. Four-minute data segments are then analyzed using method B of Chapter 6 (i.e.,  $N=256$ ,  $M=4$ ) to compute the PIP spectrum.

We now consider an example using data obtained from a patient who received rapidly-delivered fractionated exposures of 30R/day (1.5R/min) on each of 5 consecutive days. Figure 7.1a shows portions of the pre- and postexposure pulmonary impedance strip chart recordings that were obtained each day. The spectra obtained from the corresponding time periods are shown in Figure 7.1b. The increasing irregularity of the postexposure tracings is apparent in the spectra (see Figure 7.1b) which reflect the changes in the depth and frequency of breathing.

FIGURE 7.1a STRIP CHART TRACINGS

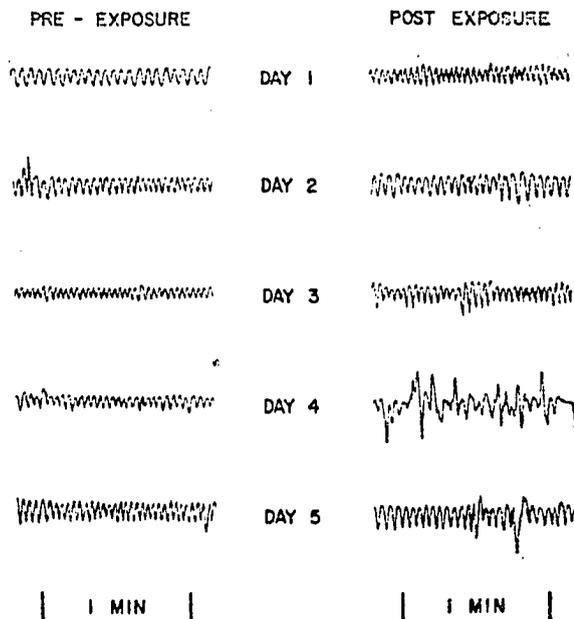


FIGURE 7.1b PIP SPECTRA

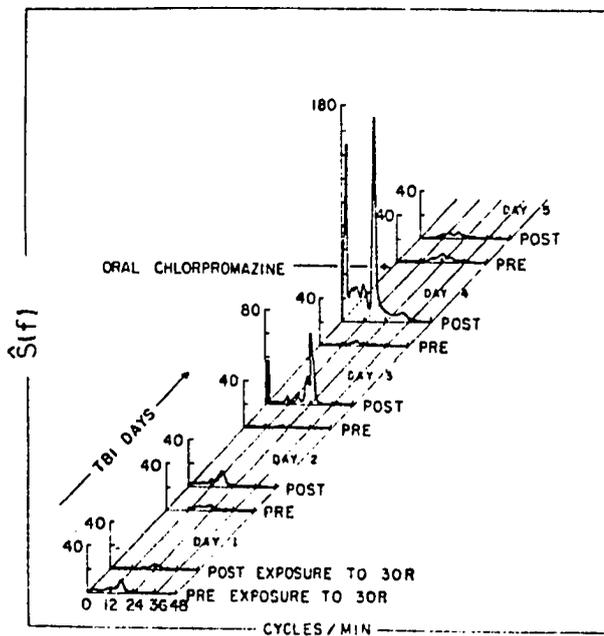


FIGURE 7.1 PULMONARY IMPEDANCE TRACINGS AND SPECTRA BEFORE AND 15MINUTES AFTER EXPOSURE TO 30R ( 1.5R/MIN) ON FIVE CONSECUTIVE DAYS.

The patient reported increasingly severe levels of gastrointestinal distress following treatment on days 1-4. On the last day of treatment 20 mg of chlorpromazine was administered prior to exposure to prevent radiation sickness. The apparent affect of the chlorpromazine can be seen on day 5 since the patient's postexposure symptoms were greatly reduced. Further discussion of this example and other applications of spectrum analysis of the PIP as a method of detecting and quantifying the effect of stress on human respiratory function have been presented by Ricks et al. (1972). They also discuss how the analysis may be used to detect changes in an individual's ability to perform under controlled exercise conditions. It is concluded that the PIP variance —i.e., the area under the PIP spectrum— is of primary interest in this regard. In the rest of this chapter we propose that the first difference of the PIP (FDPIP) should be used to obtain a quantitative measurement of respiratory effort.

## 7.2. PIP First Difference Spectrum

In dealing with spectrum analysis of the PIP we originally described (see Lushbaugh, et al. 1969) the PIP spectrum as a power spectrum. The adjective 'power' was used in the general sense —i.e., to describe the method of analysis (see Blackman and Tukey 1958, p.8). It was stated that no implication of a relation to the energy involved in the respiratory effort that produces the impedance traces was intended. We now propose a situation in which the energy concept is specified physically. In this situation the spectrum obtained may be interpreted as reflecting the mechanical work done in overcoming the

flow-resistive component of the respiratory system.

We let  $V(t)$ ,  $t = 0, \dots, N$ , denote the values obtained (in volts) by sampling the PIP. If a base-line calibration has been made by simultaneously recording pulmonary impedance and respiratory volume (as measured by a spirometer) then pulmonary impedance can be converted into respiratory volume. Then if we let

$$Y(t) = [V(t) - V(t-1)] / \Delta t, \quad (7.2.1)$$

$Y(t)$  measures the change in the volume of the RS per unit time.

If (7.2.1) is multiplied by the scale factor  $K$  (that converts volts to liters), then  $K \cdot Y(t)$  will have units of L/min. Consequently, the spectrum obtained using  $Y(t)$ ,  $t = 0, \dots, N-1$  will be in  $(L/min)^2$ . If we further assume that the pressure developed by the flow-resistive component of the RS is proportional to flow (see Section 6.2) the spectrum is truly a power spectrum as defined in Chapter 6. Further we see that

$$\frac{R}{N} \sum_{t=0}^{N-1} (KY(t))^2 = \text{work} . \quad (7.2.2)$$

Then if  $R$  is known, —see eq. (6.1.3)— the quantity defined in (7.2.2) is an estimate of the energy expended per unit time in overcoming the flow-resistive component of the respiratory system.

In the practical monitoring situation the assumptions given above are only first approximations. The impedance-volume relation will depend on the subject's bodily position and may be affected by 'motion artifacts'. Other factors that cause changes in the thorax impedance will also produce 'noise' in the PIP recording. These factors must

be taken into account if meaningful comparisons of spectra obtained from the PIP are to be made. The noise artifacts mentioned above are often of a low frequency nature —i.e., they affect the spectrum in the frequency range below about 10 cycles/min. This points to another advantage of the first difference spectrum. The difference operator  $(1-B)$  —where  $BX(t) = X(t-1)$ — has a transfer function of the form  $\sin(2\pi f/\Delta t)$ , so that the low frequencies are selectively filtered by the differencing operation. To illustrate this point consider Figure 7.2a which shows a sampled PIP tracing ( $N = 512$ , sampling rate = 240/minute). The spectra obtained from the PIP data and the first difference data are displayed in Figure 7.2b and 7.2c. It was pointed out by Lushbaugh et al. (1969) that yawning and sighing produce dramatic changes in pulmonary impedance that appear as high amplitude low frequency components in the PIP tracings. Use of the FDPIP clearly eliminates this low frequency noise. In calculating the spectra we used method D with  $M = 7$ . The PIP data were adjusted for the mean (i.e., the DC component is subtracted from the signal) before the spectrum was calculated.

In summary we conclude that the statistic

$$W(k) = \frac{1}{N} \sum_{t=0}^{N-1} Y_k(t)^2 \quad (7.2.3)$$

may be used to describe respiratory effort. When  $Y(t)$  is proportional to air flow, then  $W(k)$  will be proportional to the mechanical effort expended in overcoming the flow-resistive component of the RS. In this

situation the spectrum of  $Y(t)$  is a power spectrum. In any case the statistic  $W(k)$  may be useful in assessing the direction and relative magnitude of changes in the 'mechanical' activity of the RS. In (7.2.3) the index  $k$  is used to indicate the time at which  $W(k)$  is calculated relative to some arbitrary origin (i.e., the time at which monitoring began). If, for example, we choose  $N \cdot \Delta t =$  one minute, then  $W(k)$  describes the  $k$ th minute from the beginning of monitoring. Thus we have reduced the continuous record to a single statistic which is easily calculated and appears to provide clinically useful information.  $W(k)$  may then be used —along with other variables (e.g., heart rate)— to describe the 'state' of the patient. This approach is being used in retrospective studies and is of potential value in the real time monitoring situation.

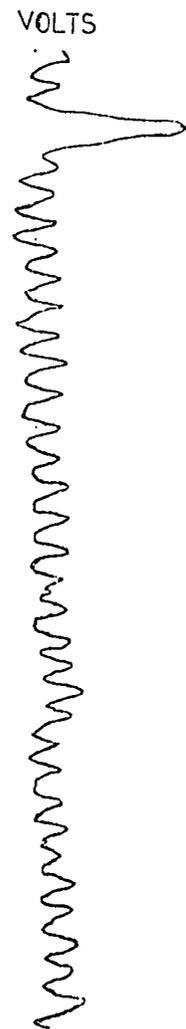


FIGURE 7.2a PULMONARY IMPEDANCE PNEUMOGRAPH

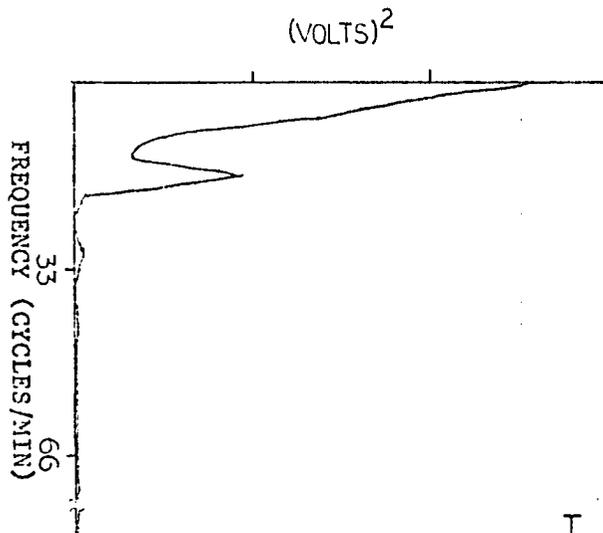


FIGURE 7.2b PIP SPECTRUM

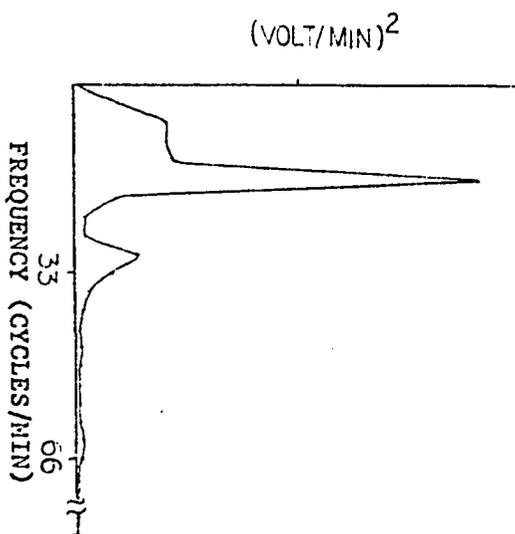


FIGURE 7.2c FDP-IP SPECTRUM

1 MIN

FIGURE 7.2 PIP SPECTRUM AND FIRST DIFFERENCE SPECTRUM

14-38

APPENDIX

```

C   VERSION A 7/10/72
C   THIS PROGRAM PERFORMS THE COMPUTATIONS THAT WERE
C   SUMMARIZED IN SECTION 4.4.  THE FINAL SUMMARY OUTPUT
C   IS FOR THE SURVIVAL CURVE MODELS.  THE ONLY SUBROUTINE
C   THAT IS NOT INCLUDED IN THIS LISTING IS GAUSJ1-- ANY
C   SUBROUTINE THAT INVERTS AN NP X NP REAL SYMMETRIC
C   MATRIX MAY BE USED IN ITS PLACE.  THE VARIABLES IN THIS
C   PROGRAM ARE DEFINED AS FOLLOWS:
C   X(I,J), J=1,M IS THE ITH SET OF VALUES OF THE M INDEP-
C   ENDENT VARIABLES ( I= 1,...,N )
C   FN(I)= NUMBER OF REPLICATIONS OF THE ITH EXPERIMENTAL
C   CONDITION.
C   Y(I), I=1,...,N IS THE OBSERVED AVERAGE RESPONSE FOR
C   THE ITH CONDITION.
C   T(I), I=1,...,NP ARE THE PARAMETERS TO BE ESTIMATED IN
C   THE REGRESSION MODEL WHERE THE EXPECTATION OF THE JTH
C   REPLICATION OF THE ITH EXPERIMENTAL CONDITION IS GIVEN
C   BY F( X(I, ),T).
C   P(I), I=1,...,NP DENOTES THE PARTIAL DERIVATIVE OF F(X,T)
C   WITH RESPECT TO T(I).
C   C IS THE NP X NP MATRIX THAT WAS DEFINED IN SEC 4.4
C   G " " " VECTOR " " " " "
C   WEIGHT OPTION IWF=0 FOR ITERATIVE (ML) WEIGHTS
C   =1 FOR FIXED WEIGHTS
C   EPS= CONVERGENCE CRITERION
C   THIS PROGRAM WAS WRITTEN AT THE EMORY UNIVERSITY
C   COMPUTING CENTER.  THE PROGRAM WAS TESTED USING
C   THE WATFOR COMPILER ON THE UNIVAC SERIES 70
C   MODEL 46 COMPUTER.
C   DIMENSION X(50,2),Y(50),FN(50),WF(50),NAME(15)
C   DIMENSION T(5),G(5),D(5),P(5),C(5,5)
C   WRITE(6,9)
C >  READ IN DATA AND INITIAL ESTIMATES OF PARMETERS
C   READ(5,19)NAME
C   READ(5,20)N,M,NP,IWF,IMAX
C   READ(5,21)(T(I),I=1,NP)
C   DO 66 I=1,N
66  READ(5,21)Y(I),FN(I),(X(I,J),J=1,M)
C   IF(IMAX,EQ.0)IMAX=10
C >  WRITE OUT INPUT DATA
C   WRITE(6,8)NAME
C   WRITE(6,26)N,M,NP,IWF
C   DO 67 I=1,N
67  WRITE(6,24)I,Y(I),FN(I),(X(I,J),J=1,M)
C >  DEFINE WEIGHTS IF OTHER THAN ML
C   IF(IWF)80,80,85
85  DO 82 I=1,N
82  WF(I)=FN(I)/Y(I)
80  CONTINUE

```

```

ITER=0
C >> BEGIN ITERATIVE PROCEDURE
100 ITER=ITER+1
C >> INITIALIZE C & G
DO 25 J=1,NP
G(J)=0,0
DO 25 I=1,NP
25 C(I,J)=0,0
SDS=0,0
FL=0,0
C >> DEFINE C & G FOR THIS ITERATION
DO 50 I=1,N
CALL FND(I,F,P,X,T)
IF(IWF)95,95,90
90 W=WF(I)
GO TO 96
95 W=FN(I)/F
96 FBAR=Y(I)
FL=FL+FN(I)*(Y(I)*ALOG(F)-F)
SDS=SDS+W*(FBAR-F)**2
DO 45 J=1,NP
DO 40 K=1,NP
40 C(J,K)=C(J,K)+P(J)*P(K)*W
45 G(J)=G(J)+W*P(J)*(FBAR-F)
50 CONTINUE
C >> SOLVE LINEAR SYSTEM OF EQUATIONS
CALL GAUSJ1(NP,NP,C,0,0,DET)
CALL MVEC(D,C,G,NP)
C ***** OUTPUT SDS T & D FOR EACH ITER *****
WRITE(6,11)ITER,SDS,FL
WRITE(6,12)
DO 150 L=1,NP
AP=ABS(D(L))/T(L)
150 WRITE(6,10)L,T(L),D(L),AP
C >>> UPDATE T ... CHECK FOR CONVERGENCE
CALL UPDAT(T,D,NP,EPS,IND)
IF(IND)200,200,310
200 IF(ITER=IMAX)100,300,300
300 WRITE(6,220)ITER
GO TO 315
310 WRITE(6,230)ITER
C ***** WRITE OUT VAR-COV MATRIX *****
WRITE(6,8)NAME
WRITE(6,210)
315 DO 260 I=1,NP
260 WRITE(6,215)I,(C(I,J),J=1,NP)
C ***** OUTPUT YBAR AND YHAT *****
WRITE(6,405)
C THE FOLLOWING OUTPUT IS FOR SURVIVAL CURVE DATA
DO 400 I=1,N

```

```

CALL FND(I,F,P,X,T)
Y1=Y(I)/X(I,1)
Y2=F/X(I,1)
F1=Y1/T(1)
F2=Y2/T(1)
CHI=(FN(I)*(Y(I)-F)**2)/F
400 WRITE(6,410)X(I,2),Y1,Y2,F1,F2,CHI
C ----- FORMAT ----->>>>>>>>>
8  FORMAT(1H1,5X,15A4/)
9  FORMAT(1H1, ' NLLSP.1 ')
10 FORMAT(1X,I4,3F20,10)
11 FORMAT(1X, //10X, ' ITER=',I3, ' SDS=',E15.6,
1'  LF=',E15.6/)
12 FORMAT(1X, ' I ',11X, 'T(I)',11X, 'D(I)'5X, 'ABS(D)/T'/)
19  FORMAT(15A4)
20  FORMAT(5I4)
21  FORMAT(7F10,5)
24  FORMAT(1X,I5,5F12,4)
26  FORMAT(1X, //20X, ' INPUT DATA FOR THIS PROBLEM .../'
110X, ' N=',I4, //10X, ' M=',I4, //10X, ' NP=',I4/
2,10X, ' WEIGHT OPTION=',I4//,
31X, ' I',7X, ' Y(I)',7X, ' N(I)',7X,
4' X(I,J), J=1,..,M'//)
210 FORMAT(1X, //10X, ' VARIANCE COVARIANCE MATRIX'//)
215 FORMAT(1X, // ' ROW ',I2,4E15,7//)
220 FORMAT(5X, ' NO CONVERGENCE AFTER ',I3, 'ITERATIONS'//)
230 FORMAT(5X, ' CONVERGENCE AFTER ',I3, ' ITERATIONS'//)
405 FORMAT(//,5X, ' RADIATION COLONIES/UNIT CONC',
13X, ' FRACTION SURVIVAL CHI-SQUARE',
2/8X, ' DOSE',4X, 'OBSERVED EXPECTED',
33X, 'OBSERVED EXPECTED',/1X)
410 FORMAT(5X,6F10.5)
STOP
END
C <<<<< SUBROUTINES FOR NLLSP.1A >>>>>
SUBROUTINE MVEC(D,A,B,N)
DIMENSION D(1),A(1,1),B(1)
C D=A*B
DO 10 K=1,N
D(K)=0.0
DO 10 I=1,N
10 D(K)=D(K)+A(K,I)*B(I)
RETURN
END
SUBROUTINE FNDLIN(I,F,P,X,T,NP)
DIMENSION P(1),X(1,1),T(1)
C THIS SUBROUTINE IS FOR MULTIPLE LINEAR REGRESSION
F=0.0
DO 10 J=1,NP
F=F+T(J)*X(I,J)

```

```

10 P(J)=X(I,J)
    RETURN
    END
    SUBROUTINE UPDAT(T,D,NP,EPS,IND)
    DIMENSION D(1),T(1)
C DETERMINE IF ABS(D(I)).LT, EPS*T(I)
C SET T(I)=T(I)+ SL*D(I)
    EPS=.00001
    SL=1.0
    IND=1
    DD 20 I=1,NP
    CK=EPS- ABS(D(I))/ABS(T(I))
    IF(CK)19,20,20
19 IND=0
20 T(I)=T(I)+SL*D(I)
    RETURN
    END
    SUBROUTINE FNDEX(I,F,P,X,T)
    DIMENSION P(1),X(1,1),T(1)
C SUB FOR SINGLE EXPONENTIAL F(X,T)=T1*EXP(-T2*X)
    P(1)=X(I,1)*EXP(-T(2)*X(I,2))
    F=T(1)*P(1)
    P(2)=-X(I,2)*F
    RETURN
    END
    SUBROUTINE FNDW(I,F,P,X,T)
C THIS SUBROUTINE IS FOR 3 PARAMETER WEIBUL MODEL
    DIMENSION P(1),X(1,1),T(1)
    IF(X(I,2))10,10,5
5 A=-T(2)*X(I,2)**T(3)
    B=EXP(A)
    P(1)=X(I,1)*B
    F=T(1)*P(1)
    P(2)=-T(1)*X(I,1)*R*X(I,2)**T(3)
    P(3)=F*A*ALOG(X(I,2))
    RETURN
10 P(1)=X(I,1)
    F=T(1)*X(I,1)
    P(2)=0.0
    P(3)=0.0
    RETURN
    END
    SUBROUTINE FND(I,F,P,X,T)
    DIMENSION P(1),X(1,1),T(1)
C EVALUATE F(X,T) AND FIRST PARTIALS AT ITH VALUE OF X
C FOR TARGET SURVIVAL CURVE MODEL
    EX=EXP(-T(2)*X(I,2))
    A=1.0-EX
    F=T(1)*X(I,1)*(1.0-A**T(3) )
    P(1)=F/T(1)

```

```

      IF(X(I,2))10,10,20
10  P(2)=0,0
      P(3)=0,0
      GO TO 500
20  P(2)=-T(1)*X(I,1)*T(3)*A**(T(3)-1,0)*X(I,2)*EX
      P(3)=-T(1)*X(I,1)*A**(T(3))*ALOG(A)
500  RETURN
      END
SAMPLE RUN      INPUT DATA
      NLLSP, TM  TILL & MCH (RAD. RESEARCH 161)-TARGET
      7  2  3  0  10
8,0    1,0    3,1
10.0000 6,0    1,25    0,0
9.4286  7,0    1,75    .96
11.5000 4,0    3,0     1,920
9.1111  9,0    7,2     2,880
9.5455  11,0  24,0    4,320
8.2000  15,0  75,0    5,760
3.0000  4,0   120,0   6,720
SAMPLE RUN ... PARTIAL OUTPUT
      ITER= 5 SDS= 0.759514E 01 LF= 0,590639E 03

```

I	T(I)	D(I)	ABS(D)/T
1	7,6364900000	-0,0000558684	0,0000073
2	0,9341030000	-0,0000010532	0,0000011
3	2,8922830000	0,0000276566	0,0000095

CONVERGENCE AFTER 5 ITERATIONS

VARIANCE COVARIANCE MATRIX.

1	0.8206457E 00	-0,1238492E-01	-0,5017006E 00
2	-0,1238476E-01	0,1590302E-02	0,2543391E-01
3	-0,5016983E 00	0,2543397E-01	0,5588594E 00

RADIATION DOSE	COLONIES/UNIT CONC		FRACTION SURVIVAL		CHI-S
	OBSERVED	EXPECTED	OBSERVED	EXPECTED	
0.00000	8,00000	7,63643	1,04761	1,00000	0,12982
0.96000	5,38777	5,95921	0,70553	0,78037	0,67126
1.92000	3,83333	3,12511	0,50198	0,40924	1,92602
2.88000	1,26543	1,40466	0,16571	0,18394	0,89423
4.32000	0,39773	0,38400	0,05208	0,05029	0,12960
5.76000	0,10933	0,10129	0,01432	0,01326	0,71929
6.72000	0,02500	0,04142	0,00327	0,00542	3,12467

#### REFERENCES

- Barankin, E. W. and Garland, J. (1950). On asymptotically normal, efficient estimators: I. In U. of California Publications in Statistics 1, 89-130.
- Bartlett, M. S. (1966). An Introduction to Stochastic Processes. The University Press, Cambridge, England.
- Bartlett, M. S. (1967). Some remarks on the analysis of time-series. Biometrika 54, 25-38.
- Beauchamp, J. J. (1966). Simultaneous Nonlinear Estimation. Ph.D. Dissertation, Florida State University. University Microfilm Inc., 67-6460, Ann Arbor, Michigan.
- Bellman, R. (1960). Introduction to Matrix Analysis. McGraw-Hill Book Company, New York.
- Berg, G., Harris, E. K., Chang, S. L. and Busch, K. A. (1963). Quantitation of viruses by the plaque technique. Journal Bacteriol. 85, 691-700.
- Blackman, R. B. and Tukey, J. W. (1958). The Measurement of Power Spectra. Dover Publications, Inc., New York.
- Bliss, C. I. (1970). Statistics in Biology, vol. 2. McGraw-Hill Book Company, New York.
- Box, M. J. and Draper, N. R. (1971). Factorial designs, the  $|X'X|$  criterion, and some related matters. Technometrics 13, 731-741.
- Box, G. E. P. and Hunter, W. G. (1965). The experimental study of physical mechanisms. Technometrics 7, 23-42.
- Box, G. E. P. and Jenkins, G. (1970). Time Series Analysis. Holden-Day, San Francisco.
- Brazier, M. A. B. (1965). The application of computers to electroencephalography. Computers in Biomedical Research, vol. 1. Edited by Ralph W. Stacy and Bruce D. Waxman. 295-318. Academic Press, New York.
- Bremermann, H. (1970). A method of unconstrained global optimization. Mathematical Biosciences 9, 1-15.

- Chen, E. H. (1971). Random normal number generator for 32-bit-word computers. J. Amer. Statist. Ass. 66, 400-403.
- Chiang, C. L. (1956). On regular best asymptotically normal estimates. Ann. Math. Stat. 27, 336-351.
- Chiang, C. L. (1968). Introduction to Stochastic Processes in Biostatistics. John Wiley & Sons, Inc., New York.
- Comas, F. (1970). The radiosensitivity of rat marrow cells. International Journal of Radiation Biology 17, 549-557.
- Comroe, J. H., Forster, R. E., Debois, A. B., Briscoe, W. A. and Carlson, E. (1962). The Lung. Year Book Medical Publishers, Inc., Chicago.
- Cooley, J. W. Lewis, P. A. W. and Welch, P. D. (1969). The finite Fourier transform. IEEE Trans., AU-17, 77-86.
- Cooley, J. W. and Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. Math. Comp. 19, 297-301.
- Cox, D. R. and Miller, H. D. (1968). The Theory of Stochastic Processes. John Wiley & Sons, Inc., New York.
- Daneill, P. J. (1946). Discussion of symposium on autocorrelation in time-series. J. Roy. Statist. Soc. 8, 88-90.
- Davis, H. T., III (1968). Inference Based on the Periodogram. Ph.D. Dissertation, Johns Hopkins University. University Microfilm, Inc., 68-16,408, Ann Arbor, Michigan.
- Dejours, P. (1966). Translated by Fari, L. E. Respiration. Oxford University Press, New York.
- Dienes, G. J. (1966). A kinetic model of biological radiation response. Radiation Res. 28, 183-202.
- Draper, N. R. and Hunter, W. G. (1967). The use of prior distribution in the design of experiments for parameter estimation in non-linear situations. Biometrika 54, 147-153.
- Draper, N. R. and Smith, H. (1966). Applied Regression Analysis. John Wiley & Sons, Inc., New York.
- Durbin, J. (1969). Tests for serial correlation in regression analysis based on the periodogram of least-squares residuals. Biometrika 56, 1-15.

- Feder, P. I. and Mezaki, R. (1971). An application of variational methods to experimental design. Technometrics 13, 771-793.
- Ferguson, T. S. (1958). A method of generating best asymptotically normal estimates with application to the estimation and bacterial densities. Ann. Math. Stat. 29, 1046-1062.
- Fisher, R. A., Thornton, H. G. and MacKenzie, W. A. (1922). The accuracy of the plating method of estimating the density of bacterial populations. Annals of Applied Biology 9, 325-359.
- Fowler, J. F. (1964). Differences in survival curve shape for formal multi-target and multi-hit models. Phys. Med. Biol. 9, 177-188.
- Frome, E. L. and Beauchamp, J. J. (1968). Maximum likelihood estimation of survival curve parameters. Biometrics 24, 595-605.
- Gart, J. J. (1964). The analysis of Poisson regression with an application in virology. Biometrika 51, 517-521.
- Good, I. J. (1950). On the inversion of circulant matrices. Biometrika 37, 185-186.
- Haight, F. A. (1967). Handbook of the Poisson Distribution. John Wiley & Sons, Inc., New York.
- Halberg, F. (1969). Chronobiology. Annual Review of Physiology 31, 675-725.
- Hannan, E. J. (1960). Time Series Analysis. Science Paperbacks and Methuen & Co. Ltd., Canberra, Australia.
- Hannan, E. J. (1969). The estimation of mixed moving average autoregressive systems. Biometrika 56, 579-593.
- Hannan, E. J. and Thomson, P. J. (1971). The estimation of coherency and group delay. Biometrika 58, 469-481.
- Hartley, H. O. (1964). Exact confidence regions for the parameters in non-linear regression laws. Biometrika 51, 347-353.
- Hartley, H. O. and Booker, A. (1965). Non-linear least squares estimation. Ann. Math. Stat. 36, 638-650.
- Jenkins G. M. and Watts, D. G. (1968). Spectral Analysis and Its Applications. Holden-Day, San Francisco.
- Jennerich, R. I. (1969). Asymptotic properties of non-linear least squares estimators. Ann. Math. Stat. 40, 633-643.

- Jones, R. H. (1965). A reappraisal of the periodogram in spectral analysis. Technometrics 7, 531-542.
- Jones, R. H. (1969). Phase free estimation of coherence. Ann. Math. Stat. 40, 540-548.
- Jones, R. H., Crowell, D. H. and Kapuniak, L. E. (1970). Change detection model for serially correlated multivariate data. Biometrics 26, 269-280.
- Jorgenson, D. W. (1963). Multiple regression analysis of a Poisson process. J. Amer. Statist. Ass. 56, 235-245.
- Kendall, M. G. and Stuart, A. (1967). The Advanced Theory of Statistics, vol. 2. 2nd edition. Hafner Publishing Co., New York.
- Krebs, J. S. (1967). The response of erythropoietic stem cells of mice to irradiation with fission neutrons. Radiation Res. 31, 796-807.
- Lee, Y. W. (1964). Contributions of Norbert Wiener to linear theory and nonlinear theory in engineering. In Selected Papers of Norbert Wiener. 17-33. M.I.T. Press, Cambridge, Mass.
- Lellouch, J. and Wamberale, A. (1966). Estimation par la methode du maximum de vraisemblance des courbes de survie de micro-organismes irradies. Biometrics 22, 673-683.
- Lushbaugh, C. C., Frome, E. L., Davis, H. T., and Bibler, D. S. (1969). Power spectrum of the impedance pneumograph: a data reduction system producing an analytical parameter of potential clinical usefulness. Aerospace Medicine 40, 425-429.
- Macy, J. (1965). Analog-digital conversion systems. In Computers In Biomedical Research, vol. 2. Edited by Ralph W. Stacy and Bruce D. Waxman. 3-34. Academic Press, New York.
- Malinvaud, E. (1970). The consistency of nonlinear regression. Ann. Math. Stat. 41, 956-969.
- Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Indust. Appl. Math. 11, 431-441.
- Mayr, O. (1970). The origins of feedback control. Scientific American 223, 110-118.
- McShane, E. J. (1969). Toward a stochastic calculus, I. In Proceedings of the National Academy of Sciences 63, 275-280.

- Mead, J. (1961). Mechanical properties of lungs. Physiological Review 41, 281-330.
- Meeter, D. A. (1964). Problems in the Analysis of Nonlinear Models by Least Squares. Ph.D. Dissertation, University of Wisconsin. University Microfilms, Inc., 64-12736, Ann Arbor, Michigan.
- Milhorn, H. T. (1966). The Application of Control Theory to Physiological Systems. W. B. Saunders Company, Philadelphia.
- Moore, R. H. and Zeigler, R. K. (1967). The use of nonlinear regression methods for analyzing sensitivity and quantal response data. Biometrics 23, 563-567.
- Morris, A. C., Barclay, T. R. and Lushbaugh, C. C. (1967). Physiologic monitoring in the low-exposure-rate total body irradiation facility. In Annual Research Report - Medical Division, Oak Ridge Associated Universities, 23-26.
- Muller, M. E. (1970). Computers as an instrument for data analysis. Technometrics 12, 259-293.
- Murthy, V. K., Haywood, L. J., Richardson, J., Kalaba, R., Saltzberg, S., Harvey, G., and Vereede, D. (1971). Analysis of power spectral densities of electrocardiograms. Mathematical Biosciences 12, 41-51.
- Nelder, J. A. (1968). Weighted regression, quantal response data, and inverse polynomials. Biometrics 13, 979-985.
- Neyman, J. (1949). Contribution to the theory of the  $X^2$  test. In Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, 239-273, U. of Calif. Press.
- Pacela, A. F. (1966). Impedance pneumography - a survey of instrumentation techniques. Med. and Biol. Engng. 4, 1-15.
- Papoulis, A. (1965). Probability, Random Variables, and Stochastic Processes. Mc-Graw Hill Book Company, New York.
- Priestly, M. B. (1971). Fitting Relationships Between Time Series. Invited paper (No. 9C) at the 38th session of the International Statistical Institute, Washington, D.C.
- Randall, J. E. (1958). Statistical properties of pulsatile pressure and flow in the femoral artery of the dog. Circulation Res. 6, 689-98.
- Rao, C. R. (1958). Linear Statistical Inference and Its Applications. John Wiley & Sons, Inc., New York.

- Ricks, R. C., Lushbaugh, C. C., McDow, E., and Frome, E. L. (1972). Pulmonary-impedance power spectral analysis: a facile means of detecting radiation-induced gastrointestinal distress and performance decrement in man. In Proceedings of the National Symposium on Natural and Manmade Radiation in Space. Edited by E. A. Warman. NASA TM X- 2440.
- Roberts, E. A. and Coote, G. C. (1965). The estimation of concentration of viruses and bacteria from dilution counts. Biometrics 21, 600-615.
- Singleton, R. C. (1969). An algorithm for computing the mixed radix fast Fourier transform. IEEE Trans. AU-17, 93-103.
- Singleton, R. C. (1969). A short bibliography on the fast Fourier transform. IEEE Trans. AU-17, 166-169.
- Smith, F. B. and Shanno, D. F. (1971). An improved Marquardt procedure for nonlinear regression. Technometrics 21, 63-74.
- Sollberger, A. (1970). Problems in the statistical analysis of short periodic time series. Journal of Interdisciplinary Cycle Research 1, 49-88.
- Stacy, R. W. and Peters, R. M. (1965). Computations of respiratory mechanical parameters. In Computers in Biomedical Research, vol. 2. Academic Press, New York.
- Tick, L. J. (1967). Estimation of coherency. In Spectral Analysis of Time Series, Edited by Bernard Harris, 133-152. John Wiley & Sons, Inc., New York.
- Tick, L. J. and Woodbury, M. A. (1965). Some time series techniques useful in life sciences. In Proceedings IBM Scientific Computing Symposium Statistics. Data Processing Div., White Plains, New York.
- Till, J. E. and McCulloch, E. A. (1961). A direct measurement of radiation sensitivity of normal mouse bone marrow cells. Radiation Research 14, 213-222.
- Tukey, J. W. (1962). The future of data analysis. Ann. Math. Stat. 33, 1-67.
- Tukey, J. W. (1967). An introduction to the calculations of numerical spectrum analysis. In Spectral Analysis of Time Series. Edited by Bernard Harris, 35-46. John Wiley & Sons, Inc., New York.

- Tukey, J. W. and Wilks, M. B. (1966). Data analysis and statistics: an expository overview. In Proceedings-Fall Joint Computer Conference. 695-709.
- Turner, M. E., Monroe, R. J. and Lucas, Jr., H. L. (1961). Generalized regression and non-linear path analysis. Biometrics 17, 120-143.
- Villegas, C. (1969). On the least squares estimation of non-linear relations. Ann. Math. Stat. 40, 462-466.
- Wiener, N. (1949). The Extrapolation, Interpolation and Smoothing of Stationary Time Series. John Wiley & Sons, Inc., New York.
- Wiener, N. (1951). Homeostasis in the individual and society. J. of the Franklin Institute 251, 65-68.
- Wijsman, R. A. (1959a). On the theory of BAN estimates. Ann. Math. Stat. 30, 185-191.
- Wijsman, R. A. (1959b). Correction to "On the theory of BAN estimates". Ann. Math. Stat. 30, 1268-1270.

#### VITA

Edward L. Frome was born on April 19, 1942 in Baltimore, Maryland. He attended primary and secondary school in Baltimore, and was graduated from Winter Park High School, Winter Park, Florida in 1960. He received a Bachelor of Science degree in 1964 and a Master of Statistics degree in 1966, both from the University of Florida. He was employed as a statistician at the Medical Division, Oak Ridge Associated Universities from 1966 to 1968, and in the Marketing Research Department of the Coca Cola Co. from 1968 to 1969. In 1969 he was awarded a three year Special Fellowship by the National Institutes of Health for graduate study and research in the Department of Statistics and Biometry at Emory University.