Spring 5-6-1975

Estimation of Growth Curves by Least Square Splines

Dorothy Rybaczyk Pathak
University of New Mexico - Main Campus

THE UNIVERSITY OF NEW MEXICO
ALBUQUERQUE, NEW MEXICO 87106

POLICY ON USE OF THESES AND DISSERTATIONS

Unpublished theses and dissertations accepted for master's and doctor's degrees and deposited in the University of New Mexico Library are open to the public for inspection and reference work. *They are to be used only with due regard to the rights of the authors.* The work of other authors should always be given full credit. Avoid quoting in amounts, over and beyond scholarly needs, such as might impair or destroy the property rights and financial benefits of another author.

To afford reasonable safeguards to authors, and consistent with the above principles, anyone quoting from theses and dissertations must observe the following conditions:

1. Direct quotations during the first two years after completion may be made only with the written permission of the author.
2. After a lapse of two years, theses and dissertations may be quoted without specific prior permission in works of original scholarship provided appropriate credit is given in the case of each quotation.
3. Quotations that are complete units in themselves (e.g., complete chapters or sections) in whatever form they may be reproduced and quotations of whatever length presented as primary material for their own sake (as in anthologies or books of readings) ALWAYS require consent of the authors.
4. The quoting author is responsible for determining "fair use" of material he uses.

This thesis/dissertation by Dorothy Rybaczyk Pathak has been used by the following persons whose signatures attest their acceptance of the above conditions. (A library which borrows this thesis/dissertation for use by its patrons is expected to secure the signature of each user.)

NAME AND ADDRESS	DATE

8/1/72—500

This dissertation, directed and approved by the candidate's committee, has been accepted by the Graduate Committee of The University of New Mexico in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

ESTIMATION OF GROWTH CURVES BY LEAST SQUARE SPLINES

Title

Dorothy Rybaczyk Pathak

Candidate

Department of Mathematics and Statistics

Department

Bernard Spolsky

Dean

May 6, 1975

Date

Committee

William J. Zimmer

Chairman

Herbert T. Davis

Betty Eberle

Steven A. Pruess

ESTIMATION OF GROWTH CURVES BY LEAST SQUARE SPLINES

BY

DOROTHY RYBACZYK PATHAK

B.S., University of Illinois, Chicago, 1970

M.A., University of New Mexico, 1972

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy in Mathematics
in the Graduate School of
The University of New Mexico
Albuquerque, New Mexico
May, 1975

ACKNOWLEDGMENTS

Throughout the course of the study presented in this manuscript, I have received much encouragement and guidance from all the four members of my Ph.D. committee, Professor William Zimmer, Chairman, Professor Herbert T. Davis, Professor Betty J. Eberle and Professor Steven A. Pruess. They read the manuscript at different stages of writing, corrected mistakes and made valuable comments. I am grate-full to them for all their help.

I am particularly thankful to Professor Betty J. Eberle for introducing me to biostatistics; her constant encouragement has done much to convince me that biostatistics is an area well worth working in.

I am much obliged to Professor Steven A. Pruess for his painstaking scrutiny of the final manuscript.

Chapter 5 owes much to Dr. Robert A. Munsick who provided me with the data without which this chapter could not have been written. This dissertation is really an outcome of my efforts to carry out a satisfactory analysis of his data. His help is gratefully acknowledged.

ESTIMATION OF GROWTH CURVES BY LEAST SQUARE SPLINES

Dorothy Rybaczyk Pathak, Ph.D.

Department of Mathematics and Statistics

The University of New Mexico, 1975

The primary object of this dissertation is to present some contributions to the theory of estimation of growth curves by least square splines in the presence of unknown unequal variances. The theoretical developments rest heavily on the standard least square theory and the theory of polynomial spline functions. A modification of the Aitken procedure of weighted least squares is used to estimate regression parameters. It is shown that this modification of the Aitken procedure does not unduly influence the nice least square properties of estimators so obtained; the estimators remain unbiased, consistent and asymptotically efficient.

The techniques developed in this dissertation are then applied to data collected by the University of New Mexico Medical Staff on weight and biparietal diameter of live newborns and fresh abortuses within thirty minutes of birth. A numerical formula is developed for prediction of weight of a fetus or a newborn from its biparietal diameter.

ESTIMATION OF GROWTH CURVES BY LEAST SQUARE SPLINES

Dorothy Rybaczyk Pathak, Ph.D.

Department of Mathematics and Statistics

The University of New Mexico, 1975

The primary object of this dissertation is to present some contributions to the theory of estimation of growth curves by least square splines in the presence of unknown unequal variances. The theoretical developments rest heavily on the standard least square theory and the theory of polynomial spline functions. A modification of the Aitken procedure of weighted least squares is used to estimate regression parameters. It is shown that this modification of the Aitken procedure does not unduly influence the nice least square properties of estimators so obtained; the estimators remain unbiased, consistent and asymptotically efficient.

The techniques developed in this dissertation are then applied to data collected by the University of New Mexico Medical Staff on weight and biparietal diameter of live newborns and fresh abortuses within thirty minutes of birth. A numerical formula is developed for prediction of weight of a fetus or a newborn from its biparietal diameter.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

b	vector of parameters
b_i	ith component of the vector b
$\hat{b}$	estimate of vector of parameters b
b'	transpose of vector b
c_j	scalars
e	unit vector
e'	transpose of vector e
exp	exponential funtion
$f(x)$	regression function
g	constant
$g(x)$	smoothing spline
i	subscript
inf	smallest value
j	subscript
lim	limit
ℓ	vector of scalars
m, n	subscripts
n_i	within group sample size
p	vector of scalars
$P(\ )$	probability of the event within parentheses
p_i	eigenvector with eigenvalue λ_i
q	number of components in the vector b
r	subscript

LIST OF SYMBOLS (Continued)

s_i^2	within group variance
sup	largest value
$s(x)$	polynomial spline function
$s^{(k)}$	kth derivative of s at x
u	vector
v_i	predicted value of a smoothing spline
v	eigenvector
w_i	weights
x	independent variable
x_+^m	truncated power function
x_i	ith knot
$y(x)$	observed value of y-variate at x
y_i	the ith observed value of y-variate
$\bar{y}_{i.}$	the within group average of y-variates
$\hat{\bar{y}}_{i.}$	the estimator of $\bar{y}_{i.}$
z	tail-value from the normal distribution
A	matrix
A^-	generalized inverse of A
A'	transpose of A
B_i	B-splines
C^{m-1}	class of continuous functions with continuous (m-1)st derivative
D	dispersion operator
Df	divided difference operator
E	expectation operator

LIST OF SYMBOLS (Continued)

E_{m+n+1}	Euclidean space of dimension (m+n+1)
F	Variance ratio
H_o	null hypothesis
H, H_r	projection matrices
I	identity matrix
L	matrix
L(x)	Lagrange's polynomial
M_i	divided difference
N	matrix
P	orthogonal matrix
P( )	probability of the event within parentheses
Q	vector of sum of products
Q_i	ith component of vector Q
R_o^2	residual sum of squares under the null hypothesis
R_1^2	residual sum of squares under the alternative hypothesis
R(.)	rank of matrix
S	matrix of within group variances
S_{uv}	matrix of sum of products
T_n	nth random variable
V	true dispersion matrix
$\hat{V}$	estimated dispersion matrix
$\hat{V}^{-1}$	inverse of $\hat{V}$
υ	variance operator

LIST OF SYMBOLS (Continued)

X	design matrix
Xb	regression function of Y
Y	vector of observations
Y_n	nth vector of observations
δ	constant
Δ	finite difference operator
Δ^k	kth order finite difference operator
ϵ_x	the error term at x
ϵ_x	the error term at x
$\varphi(b)$	linear function of vector b
σ^2	variance
$\hat{\sigma}^2$	estimated variance
λ_i	ith eigenvalue of a matrix
θ_i	ith unknown parameter in the density function
$\mathfrak{N}_{2k-1}$	class of natural splines
π	universal constant
π_m	class of polynomials of degree m or less
ω	element of matrix Ω
Σ	summation sign
Γ	matrix
Ω	matrix
Π	product symbol
χ^2	the chi-square distribution

CHAPTER ONE

INTRODUCTION

The subject matter of this dissertation orginated from a study which was recently carried out by the University of New Mexico Medical Staff for the purpose of finding the relation between biparietal diameter (maximum skull breadth) and the weight of a fetus or a newborn. A preliminary examination of the collected data showed that the variance of the observations increased with the biparietal diameter. Since in the literature little is known about the regression analysis of data with unknown unequal variances we have attempted in this dissertation to present a technique of weighted least square regression analysis with polynomial spline functions. We now proceed to describe the results of our investigation.

Chapter 2 contains a review of the known results concerning polynomial spline functions and their representation in terms of B-splines, which are the divided differences of truncated power functions. It is also shown how this basis can be used to construct regression models for growth curves.

Chapter 3 begins with a brief introduction of the customary technique of least square estimation. Then a technique of data reduction through the notion of sufficiency for the purpose of reducing the complexity of a regression model is suggested when more than one observation is available for each value (called cell) of the independent variable and when the within cell variances are heterogeneous. The technique consists in replacing all the observations within each cell by their mean and their sample variance. The theory of sufficiency is used to establish that this data reduction leads to no loss of information in our analysis and much of the analysis in this dissertation heavily depends on this technique of data reduction. We then suggest the use of a modification of the standard Aitken procedure of weighted least squares in which the weights are chosen on the basis of the estimated cell variances and the analysis itself is carried out on the basis of the sample means from all the cells. Owing to the independence of the sample means and the sample variances this procedure preserves many of the nice properties of the standard least square theory. As a preliminary step towards establishing some of these properties some well-known results on generalized inverses of matrices are described. Then it is shown in Theorem 3.1 of this chapter that every estimable parameter $p'b$ under the modified Aitken procedure possesses a unique unbiased estimator. In conclusion we derive in Theorem 3.2 an expression for the variance of these estimators.

Chapter 4 deals with the asymptotic properties of the estimators obtained by the modified Aitken procedure. For this purpose a few well-known results concerning eigenvalues of symmetric matrices are stated and some inequalities concerning the nonzero eigenvalues of a product of matrices are derived in Lemma 2.9. These results are then used to establish Theorem 3.1 which shows that under certain regularity conditions (1.1) and (1.2) the modified Aitken procedure provides consistent estimators of estimable parameters. Theorem 4.1 shows that these estimators are asymptotically equivalent to those obtained under the standard Aitken procedure in which the cell variances are known.

The last chapter gives an actual application of the techniques developed in the earlier chapters to the data collected by the University of New Mexico Medical Staff. Among other things we develop a prediction formula for the weight of a fetus or a newborn in terms of its biparietal diameter; we also provide tables of predicted weight and its 90% and 95% tolerance limits for a given biparietal diameter. The prediction curve and its tolerance limits are also displayed graphically. The chapter concludes with a brief discussion of the advantages and disadvantages of the use of polynomial spline functions in regression analysis of growth data.

CHAPTER TWO

INTRODUCTION TO POLYNOMIAL SPLINE FUNCTIONS

1. Introduction.

Definition 1.1. A polynomial spline function $s(x)$ of degree m with knots $x_1, x_2, \ldots, x_n$ is a function defined over the whole real line having the following two properties:

i) In each interval (x_i, x_{i+1}), $0 \leq i \leq n$, where $x_0 = -\infty$ and $x_{n+1} = \infty$, $s(x)$ is given by a polynomial of degree m or less.

ii) $s(x)$ and its derivatives of orders $1, 2, \ldots, m-1$ are continuous everywhere.

The space of all such splines is denoted by $S_m(x_1, \ldots, x_n)$. For $m > 0$, a spline function of degree m can also be defined as a function in C^{m-1} whose mth derivative is a step function.

One way of representing a spline is by the use of a truncated power function which is defined as:

$$x_+^m = \begin{cases} x^m & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \tag{1.1}$$

Then any spline of degree m admits the representation:

$$s(x) = p(x) + \sum_{j=1}^{n} c_j (x - x_j)_+^m \tag{1.2}$$

where $p \in \pi_m$, π_m being the class of polynomials of degree m or less and the c_j, $1 \leq j \leq n$, are scalars.

From this representation it is clear that one needs $n + m + 1$ independent parameters to specify a spline function of degree m since there are n c_j's and $p(x)$ involves $(m + 1)$ independent parameters.

An alternative way of arriving at this number of parameters is to see that between any two sucessive knots the spline function is a polynomial of degree m; this involves $m + 1$ parameters there and thus over $n + 1$ intervals $(x_o,x_1), (x_1,x_2),\ldots,(x_n,x_{n+1})$ we have in all $(n + 1)(m + 1) = nm + n + m + 1$ parameters. However from condition (ii) of the definition of spline function, $s(x)$ and all of its $m - 1$ derivatives have to agree at internal knots. This yields nm conditions thus reducing the number of independent parameters to

$(n + 1)(m + 1) - nm = \underline{m + n + 1}$.

This consideration also shows that the space of splines of degree m with n internal knots is isomorphic to the $(m + n + 1)$ dimensional Euclidean space E_{m+n+1}.

A special case of some interest is when we restrict our attention only to the interval (x_1,x_n). In this case there are $(n - 1)(m + 1)$ parameters involved and $(n - 2)m$ conditions on $s(x)$ and its derivatives. Thus the number of independent parameters reduces to

$(n - 1)(m + 1) - (n - 2)m = \underline{n + m - 1}$.

<u>2. Natural spline.</u>

<u>Definition 2.1</u>. A polynomial spline $s(x)$ with knots $x_1,\ldots,x_n$ of

odd degree $2k - 1$ is called a natural spline function if it is a polynomial of degree $2k - 1$ in the internal intervals and a polynomial of degree $k - 1$ in the intervals $(-\infty, x_1)$ and (x_n, ∞).

The natural spline assumes the form:

$$s(x) = p(x) + \sum_{j=1}^{n} c_j (x - x_j)_+^{2k-1} \tag{2.1}$$

where $p(x) \in \pi_{k-1}$ and the c_j's satisfy the following conditions:

$$\sum c_j x_j^r = 0 \qquad r = 0,1,\ldots,k-1 \tag{2.2}$$

Unless stated otherwise we denote the class of natural splines by η_{2k-1}.

The conditions on c_j's arise from the fact that a natural spline is a polynomial of degree $k - 1$ on the intervals $(-\infty, x_1)$ and (x_n, ∞). Over the range $(-\infty, x_1)$ the above spline is clearly a polynomial of degree $k - 1$. Now for $x > x_n$ the spline becomes a polynomial of the form:

$$s(x) = p(x) + \sum_{j=1}^{n} c_j (x - x_j)^{2k-1} \tag{2.3}$$

For this to be a polynomial of degree $k-1$, $s^{(k)}(x), \ldots, s^{(2k-1)}(x)$ must all be zero. Thus considering the $(2k - 1)$st derivative we obtain:

$$(2k - 1)! \sum_{j=1}^{n} c_j = 0$$

The preceding equation is equivalent to the condition:

$$\sum_{j=1}^{n} c_j x_j^0 = 0 \qquad (2.4)$$

Next from the $(2k - 2)$nd derivative we have:

$$\frac{(2k-1)!}{2} \sum_{j=1}^{n} c_j (x - x_j) = 0$$

which is equivalent to:

$$\sum_{j=1}^{n} c_j x - \sum_{j=1}^{n} c_j x_j = 0 \qquad (2.5)$$

Since $x \sum_{j=1}^{n} c_j = 0$ from (2.4), this gives us the second condition:

$$\sum_{j=1}^{n} c_j x_j^1 = 0 \qquad (2.6)$$

and so on.

We shall need the following definition.

Definition 2.2. A function $f(x)$ is said to interpolate the data points $(x_1,y_1),\ldots,(x_n,y_n)$ if $f(x_i) = y_i$ for each i, $1 \leq i \leq n$.

The following is an important result concerning natural splines (Cf. Greville [5], p.6, for a proof).

Theorem 2.1. If $1 \leq k \leq n$ and $a=x_1 < x_2 < x_3 \ldots < x_n=b$, then there is a unique $s \in \Pi_{2k-1}(x_1,\ldots,x_n)$ that interpolates any set of data points $(x_1,y_1),\ldots,(x_n,y_n)$.

Suppose now that we wish to find the "smoothest" interpolating function $g(x)$ in C^{2k-2} for n distinct data points, where smoothness is measured by the smallness of the integral

$$\sigma(g) = \int_a^b (g^{(k)}(x))^2\, dx \tag{2.7}$$

Then for $k < n$ the optimum solution is given by the unique natural interpolating spline of degree $2k - 1$ with the abscissas of the given data points as its knots.

If $k = n$, there is a unique polynomial of degree $k - 1$ that interpolates the data points $\{ (x_j,y_j) : 1 \leq j \leq n \}$. This polynomial is given by Lagrange's formula

$$L(x) = \sum_{j=1}^{n} \frac{P_j(x)}{P_j(x_j)}\, y_j \tag{2.8}$$

where $P(x) = (x - x_1)(x - x_2)\ldots(x - x_n)$ and $P_j(x)$ is the product obtained by deleting the factor $(x - x_j)$ from $P(x)$. For $g = L(x)$, $\sigma(g) = 0$ which is the minimum.

For $k > n$, there are infinitely many interpolating polynomials for which $\sigma(g) = 0$.

3. Data smoothing splines.

In order to smooth a sequence of equally spaced ordinates E. T. Whittaker [12] in 1919 suggested minimizing the quantity:

$$\sum_{i=1}^{n} w_i(v_i - y_i)^2 + g\sum_{i=1}^{n-k}(\Delta^k v_i)^2 \qquad (3.1)$$

where y_i denotes the ith observed ordinate, v_i the corresponding smooth value, w_i a given positive weight, Δ the usual finite-difference operator and g a positive constant to be chosen by the user. If $g = 0$ then letting $v_i = y_i$ provides an interpolatory spline. If $g \to \infty$ and if we wish to make the whole quantity as small as possible, we are then forced into imposing the condition $\Delta^k v_i = 0$, i.e., the set of v_i's must then be values for a polynomial of degree $k - 1$ in which case the solution is given by a least square polynomial of degree $k - 1$. Schoenberg [11] generalized this idea (for data points not necessarily equally spaced) by considering the problem of fitting a function $f \in C^{k-1}$ with $f^{(k)}$ piecewise continuous so that:

$$\sigma_w(f) = \sum_{i=1}^{n} w_i(f(x_i) - y_i)^2 + g\int_a^b (f^k(x))^2\,dx \qquad (3.2)$$

is a minimum. He showed that the optimum solution to the above problem is given by a natural spline of degree $2k - 1$. Also if the natural spline is expressed as:

$$s(x) = p(x) + \sum_{j=1}^{n} c_j(x - x_j)_+^{2k-1}$$

where $p(x) \in \pi_{k-1}$ and c_j's satisfy the following conditions:

$$\sum_{j=1}^{n} c_j x_j^r = 0 \qquad r = 0,\ldots,k-1$$

then $s(x)$ is determined uniquely by the following additional n equations:

$$s(x_i) + (-1)^k(2k - 1)\, c_i w_i = y_i, \qquad i = 1,\ldots,n \qquad (3.3)$$

It should be noted at this stage that a natural spline is completely specified by n independent parameters.

4. B-splines.

The determination of splines through the representation:

$$s(x) = p(x) + \sum_{j=1}^{n} c_j (x - x_j)_+^m$$

where $p \in \pi_m$, generally leads to a system of equations which is ill conditioned. Schoenberg [10] has developed a more satisfactory representation of splines in terms of functions based on divided differences. These functions, which are referred to as B-splines, possess the desirable property that their support (domain over which the function is non-zero) is restricted to only finitely many intervals determined by knots.

Definition 4.1. Define the linear functional D as the divided difference of f of order m based on the arguments $x_i, x_{i+1}, \ldots, x_{i+m}$, i.e.,

$$Df = f(x_i,x_{i+1},\ldots,x_{i+m}) = \frac{f(x_i)}{(x_i-x_{i+1})(x_i-x_{i+2})\ldots(x_i-x_{i+m})} +$$

$$\frac{f(x_{i+1})}{(x_{i+1}-x_i)(x_{i+1}-x_{i+2})\ldots(x_{i+1}-x_{i+m})} +\ldots+ \frac{f(x_{i+m})}{(x_{i+m}-x_i)(x_{i+m}-x_{i+1})\ldots(x_{i+m}-x_{i+m-1})}$$

$$= \sum_{j=i}^{i+m} \frac{f(x_j)}{P_j(x_j)} \tag{3.4}$$

where $P_j(x_j)$ is defined in (2.8). Using this definition of the linear functional D, Schoenberg used the following functions for representation of splines:

$$M_i(x) = DM_x(t) = M_x(x_i,\ldots,x_{i+m}), \quad i = 1,\ldots,n-m \tag{3.5}$$

where $M_x(t) = m(t - x)_+^{m-1}$ and the differencing in (3.5) is carried out with respect to 't'. For example in case of cubic splines (in this case $m = 4$) we have:

$$M_1(x) = M(x;\ x_1,x_2,\ldots,x_5) = \frac{4(x_1-x)_+^3}{(x_1-x_2)(x_1-x_3)(x_1-x_4)(x_1-x_5)} +$$

$$\frac{4(x_2-x)_+^3}{(x_2-x_1)(x_2-x_3)(x_2-x_4)(x_2-x_5)} + \frac{4(x_3-x)_+^3}{(x_3-x_1)(x_3-x_2)(x_3-x_4)(x_3-x_5)} +$$

$$\frac{4(x_4-x)_+^3}{(x_4-x_1)(x_4-x_2)(x_4-x_3)(x_4-x_5)} + \frac{4(x_5-x)_+^3}{(x_5-x_1)(x_5-x_2)(x_5-x_3)(x_5-x_4)}$$

$$M_2(x) = M(x; x_2, x_3, \ldots, x_6) = \frac{4(x_2-x)_+^3}{(x_2-x_3)(x_2-x_4)(x_2-x_5)(x_2-x_6)} +$$

$$\frac{4(x_3-x)_+^3}{(x_3-x_2)(x_3-x_4)(x_3-x_5)(x_3-x_6)} + \frac{4(x_4-x)_+^3}{(x_4-x_2)(x_4-x_3)(x_4-x_5)(x_4-x_6)} +$$

$$\frac{4(x_5-x)_+^3}{(x_5-x_2)(x_5-x_3)(x_5-x_4)(x_5-x_6)} + \frac{4(x_6-x)_+^3}{(x_6-x_2)(x_6-x_3)(x_6-x_4)(x_6-x_5)}$$

and so on. The functions $M_i(x)$ are called B-splines.

Properties of B-splines.

Since for each t, the truncated polynomial $(t - x)_+^{m-1}$ as a function of x is a spline of degree $m - 1$, it follows that the $M_i(x)$'s are all splines of degree $m - 1$. It is easily seen that for $x > x_{i+m}$, $M_i(x)$ vanishes; and for $x < x_i$, $M_i(x)$ becomes the mth divided difference of a polynomial of degree $m - 1$, which is also zero. Thus the support of $M_i(x)$ is contained in the interval (x_i, x_{i+m}). Curry and Schoenberg [4] also established that:

i) $M_i(x)$ is strictly positive in (x_i, x_{i+m}) and

ii) any spline $s \in S_{m-1}(x_1, \ldots, x_n)$ having its support contained in (x_1, x_n) has a unique representation as a linear combination of $M_1, M_2, \ldots, M_{n-m}$. This is because such a spline needs only $m(n - 1) - n(m - 1) = n - m$ parameters in its representation and $M_1, \ldots, M_{n-m}$ are $n - m$ linearly independent splines of degree $m - 1$ which vanish outside (x_1,x_n) and therefore form a basis for such splines, i.e., every polynomial spline of degree $m - 1$ which vanishes outside the interval (x_1,x_n) is uniquely expressible as a linear combination of $M_1,M_2,\ldots,M_{n-m}$.

5. A basis for $S_{m-1}(x_1,x_2,\ldots,x_n)$.

The preceding B-splines play a central role in obtaining a satisfactory basis for $S_{m-1}(x_1,\ldots,x_n)$. Since $S_{m-1}(x_1,\ldots,x_n)$ has dimension $(n + m)$ and since the B-splines are only $(n - m)$ in number, we need an additional $2m$ independent splines to form a basis. The following theorem furnishes these additional splines ([5], p. 26).

Theorem 5.1. Let $m \leq n$ and suppose that $x_1 < x_2 < \ldots < x_n$. Then the following $(n + m)$ splines provide a basis for $S_{m-1}(x_1,\ldots,x_n)$:

$$B_i(x) = M(x;\, x_1,\ldots,x_i), \qquad 1 \leq i \leq m \tag{5.1}$$

$$B_{m+i}(x) = M_i(x), \qquad 1 \leq i \leq n-m$$

$$B_{n+i}(x) = (-1)^{m-i}\, M(x_{n-m+i}, x_{n-m+i+1},\ldots,x_n;\, x), \quad 1 \leq i \leq m$$

where $M(x;\, x_1,\ldots,x_i)$ denotes $(i-1)$st order divided difference of $m(t-x)_+^{m-1}$ with respect to t over the knots $x_1,x_2,\ldots,x_i$; $M(x_{n-m+i},x_{n-m+i+1},\ldots,x_n;\, x)$ the $(m - i)$th order divided difference of $m(x - t)_+^{m-1}$ with respect to t over the knots $x_{n-m+i},\ldots,x_n$ and the $M_i(x)$ are given by (3.5).

For example in the special case of $m = 4$ (cubic spline) and

$n = 5$ the above basis admits the following representation:

$$B_1(x) = M(x;\, x_1) = 4(x_1-x)_+^3 \tag{5.2}$$

$$B_2(x) = M(x;\, x_1,x_2) = \frac{4(x_1-x)_+^3}{(x_1-x_2)} + \frac{4(x_2-x)_+^3}{(x_2-x_1)}$$

$$B_3(x) = M(x;\, x_1,x_2,x_3) = \frac{4(x_1-x)_+^3}{(x_1-x_2)(x_1-x_3)} + \frac{4(x_2-x)_+^3}{(x_2-x_1)(x_2-x_3)} + \frac{4(x_3-x)_+^3}{(x_3-x_1)(x_3-x_2)}$$

$$B_4(x) = M(x;\, x_1,x_2,x_3,x_4) = \frac{4(x_1-x)_+^3}{(x_1-x_2)(x_1-x_3)(x_1-x_4)} +$$

$$\frac{4(x_2-x)_+^3}{(x_2-x_1)(x_2-x_3)(x_2-x_4)} + \frac{4(x_3-x)_+^3}{(x_3-x_1)(x_3-x_2)(x_3-x_4)} + \frac{4(x_4-x)_+^3}{(x_4-x_1)(x_4-x_2)(x_4-x_3)}$$

$$B_5(x) = M_1(x) = M(x;\, x_1,x_2,x_3,x_4,x_5) = \frac{4(x_1-x)_+^3}{(x_1-x_2)(x_1-x_3)(x_1-x_4)(x_1-x_5)} +$$

$$\frac{4(x_2-x)_+^3}{(x_2-x_1)(x_2-x_3)(x_2-x_4)(x_2-x_5)} + \frac{4(x_3-x)_+^3}{(x_3-x_1)(x_3-x_2)(x_3-x_4)(x_3-x_5)} +$$

$$\frac{4(x_4-x)_+^3}{(x_4-x_1)(x_4-x_2)(x_4-x_3)(x_4-x_5)} + \frac{4(x_5-x)_+^3}{(x_5-x_1)(x_5-x_2)(x_5-x_3)(x_5-x_4)}$$

$$B_6(x) = (-1)^3 M(x_2,x_3,x_4,x_5;\ x) =$$

$$- \left[\frac{4(x-x_2)_+^3}{(x_2-x_3)(x_2-x_4)(x_2-x_5)} + \frac{4(x-x_3)_+^3}{(x_3-x_2)(x_3-x_4)(x_3-x_5)} + \right.$$

$$\left. \frac{4(x-x_4)_+^3}{(x_4-x_2)(x_4-x_3)(x_4-x_5)} + \frac{4(x-x_5)_+^3}{(x_5-x_2)(x_5-x_3)(x_5-x_4)} \right]$$

$$B_7(x) = (-1)^2 M(x_3,x_4,x_5;\ x) =$$

$$\frac{4(x-x_3)_+^3}{(x_3-x_4)(x_3-x_5)} + \frac{4(x-x_4)_+^3}{(x_4-x_3)(x_4-x_5)} + \frac{4(x-x_5)_+^3}{(x_5-x_3)(x_5-x_4)}$$

$$B_8(x) = (-1)^1 M(x_4,x_5;\ x) = - \left[\frac{4(x-x_4)_+^3}{(x_4-x_5)} + \frac{4(x-x_5)_+^3}{(x_5-x_4)} \right]$$

$$B_9(x) = (-1)^0 M(x_5;\ x) = 4(x-x_5)_+^3$$

Remarks.

1) It is worthwhile noting that for each i, $1 \leq i \leq m$, B_i is positive for $x < x_i$, zero for $x > x_i$ and is a polynomial of degree $m - i$ for $x < x_1$, and B_{n+i} is zero for $x < x_{n-m+i}$, positive for $x > x_{n-m+i}$ and is a polynomial of degree $m - i$ for $x > x_n$.

2) As will be illustrated in later chapters there are situations in biometry such as in the study of growth curves wherein one needs to fit a spline which vanishes for $x < x_1 = 0$, say. In view of the preceding properties of the B_i's it can be shown that a suitable basis for splines of degree $m - 1$ which vanish for $x < x_1$ is given by:

$$B_{m+i}(x) = M_i(x), \qquad 1 \leq i \leq n-m \qquad \text{and} \tag{5.3}$$

$$B_{n+i}(x) = (-1)^{m-i} M(x_{n-m+i}, x_{n-m+i+1}, \ldots, x_n; x), \quad 1 \leq i \leq m$$

Thus any spline of degree $m - 1$ which vanishes for $x < x_1$ admits the representation:

$$s(x) = b_1 B_{m+1}(x) + b_2 B_{m+2}(x) + \ldots + b_n B_{n+m}(x) \tag{5.4}$$

where $b_1, \ldots, b_n$ are parameters to be determined.

In a similar fashion it can be seen that a basis for splines of degree $m - 1$ which vanish for $x > x_n$ is given by the B-splines $B_1, \ldots, B_m, B_{m+1}, \ldots, B_n$.

3) It should be noted that Theorem 5.1 and above remarks apply only to the case when $m \leq n$. For $m > n$ the representation of splines in terms of the truncated power functions as given in (1.2) should be used.

6. Fitting of splines.

A regression model commonly encountered in biometry has the following form:

$$y(x) = f(x) + \varepsilon_x \qquad (6.1)$$

where x denotes the nonrandom independent variable while $y(x)$ is the dependent variable and ε_x's are independent normally distributed errors. The function $f(x)$ is called the regression function and one of the problems in regression analysis is to estimate the form of f on the basis of the observed values of $y(x)$ for a given set of values of x. In many problems $f(x)$ is assumed to have a polynomial representation such as:

$$f(x) = b_o + b_1 x + b_2 x^2 + \ldots + b_q x^q \qquad (6.2)$$

where $b_o, b_1, \ldots, b_q$ are unknown parameters. Thus in order to ascertain the form of regression function $f(x)$ one needs to estimate the values of parameters $b_o, b_1, \ldots, b_q$. This is usually done by the method of least squares.

We shall see in later chapters that there are situations where a regression function is better explained in terms of a spline rather than a polynomial. Thus if $f(x)$ assumes a representation as a spline of degree $m - 1$ with n knots, say, it then follows from our preceding considerations that $f(x)$ can be written in the following form:

$$f(x) = b_1B_1(x) + b_2B_2(x) + \ldots + b_{n+m}B_{n+m}(x) \qquad (6.2)$$

where $b_1,\ldots,b_{n+m}$ are unknown parameters and the B_i's are the B-splines of Theorem 5.1. Now in order to determine the spline that best fits the data in a given case one can use the standard techniques of least square theory and thus determine the optimum values for $b_1,\ldots,b_{n+m}$. We intend to use the representation in (6.2) for $f(x)$ in the study of growth curves.

CHAPTER THREE

REGRESSION ANALYSIS WITH UNKNOWN UNEQUAL VARIANCES

1. Introduction. The main object of this chapter is to introduce a technique of estimating parameters encountered in regression analysis when the residual errors have unknown unequal variances. For reasons of clarity it is worthwhile to first present a brief outline of the estimation procedure used in case of equal variances. The basis of analysis in this latter case is the assumption of the following model:

$$y_i = b_1 x_{i1} + b_2 x_{i2} + \dots + b_q x_{iq} + \epsilon_i \,, \qquad 1 \leq i \leq n \qquad (1.1)$$

where the observations y_i are assumed to be expressible as linear combinations of known variables $x_{i1}, x_{i2}, \dots, x_{iq}$ and the residual errors ϵ_i's, which are assumed to be independent and normally distributed with zero mean and an unknown common variance σ^2.

The unknown parameters $b_1, b_2, \dots, b_q$ are estimated by the so called technique of least squares which consists of minimizing the following quantity:

$$U = \sum_{i=1}^{n} (y_i - b_1 x_{i1} - b_2 x_{i2} - \dots - b_q x_{iq})^2 \qquad (1.2)$$

The normal equations which determine b's (obtained by differentiating (1.2) with respect to b's and equating to zero) are given by:

$$b_1 \sum x_{i1}^2 + b_2 \sum x_{i2}x_{i1} + \dots + b_q \sum x_{iq}x_{i1} = \sum y_i x_{i1} \qquad (1.3)$$

$$b_1 \sum x_{i1}x_{i2} + b_2 \sum x_{i2}^2 + \dots + b_q \sum x_{iq}x_{i2} = \sum y_i x_{i2}$$

$$\dots$$

$$b_1 \sum x_{i1}x_{iq} + b_2 \sum x_{i2}x_{iq} + \dots + b_q \sum x_{iq}^2 = \sum y_i x_{iq}$$

In the sequel a parameter $\varphi(b)$ (a given function of b's) is said to be estimable if there exists a linear function of the y_i's, say $\sum \ell_i y_i$, such that $E(\sum \ell_i y_i) = \varphi(b)$.

The estimators of linear functions of b's if estimable have a number of desirable properties. Among other things they possess minimum variance in the class of all estimators which are linear functions of the observations y_i (Cf. Rao, C.R., [7], pp. 181-182).

With this very brief introduction of the estimation technique in the equal variance case let us now turn to an analogous problem of estimation in which the residuals have different unknown variances and for which no satisfactory solution is available in the literature.

Let $y_{11},\dots,y_{1n_1},y_{21},\dots,y_{2n_2},\dots,y_{k1},\dots,y_{kn_k}$ be k sets of a total of $n_1 + \dots + n_k$ observations, with $n_i > 1$ for each i, and suppose that they admit the following model:

$$y_{1j} = b_1 x_{11} + \dots + b_q x_{1q} + \varepsilon_{1j}, \qquad 1 \le j \le n_1 \qquad (1.4)$$

$$y_{2j} = b_1 x_{21} + \dots + b_q x_{2q} + \varepsilon_{2j}, \qquad 1 \le j \le n_2$$

$$\dots$$

$$y_{kj} = b_1 x_{k1} + \dots + b_q x_{kq} + \varepsilon_{kj}, \qquad 1 \le j \le n_k$$

where the x's are known variables, b's unknown parameters and ε's are independent normal variates with zero mean and variance $V(\varepsilon_{ij}) = \sigma_i^2$, $1 \leq i \leq k$. Thus in our model the whole set of observations can be divided into k groups such that the observations in the ith group are independently distributed with a common mean and a common variance of σ_i^2, $1 \leq i \leq k$; the variances σ_i^2 may vary from group to group and are assumed unknown. Our object now is to study the estimation of parameters $b_1, \ldots, b_q$ in the above set up. We first employ the notion of sufficiency to reduce the complexity of this problem. In this connection we need the following well-known theorem (Cf. [7], p. 110).

Theorem 1.1 (Neyman-Fisher). Let $f(y_1, \ldots, y_n;\ \theta_1, \ldots, \theta_k)$ denote the joint density function of n random variables $Y_1, \ldots, Y_n$ with $\theta_1, \ldots, \theta_k$ being the unknown parameters under study. Then a vector-valued statistic t is sufficient if and only if the density function f admits the following factorization:

$$f(y_1, y_2, \ldots, y_n;\ \theta_1, \ldots, \theta_k) = g(y_1, \ldots, y_n)\, h(t, \theta_1, \ldots, \theta_k) \qquad (1.5)$$

where the first function g is free of parameters while the second function h depends on the parameters $\theta_1, \ldots, \theta_k$ only through the statistic t.

Under the model (1.4) and the earlier assumptions the joint density function of the observations $y_{11}, \ldots, y_{1n_1}, \ldots, y_{k1}, \ldots, y_{kn_k}$ is given as follows:

$$f(y_{11},\ldots,y_{1n_1},\ldots,y_{k1},\ldots,y_{kn_k}) =$$

$$\prod_{i=1}^{k} \frac{1}{(2\pi\sigma_i^2)^{n_i/2}} \exp\left\{ -\frac{1}{2\sigma_i^2} \sum_{j=1}^{n_i} (y_{ij} - b_1x_{i1} - \ldots - b_qx_{iq})^2 \right\} =$$

$$\prod_{i=1}^{k} \frac{1}{(2\pi\sigma_i^2)^{n_i/2}} \exp\left\{ -\frac{1}{2\sigma_i^2} \left[\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2 + n_i(\bar{y}_{i.} - b_1x_{i1} - \ldots - b_qx_{iq})^2 \right] \right\}$$

$$= \prod_{i=1}^{k} \frac{1}{(2\pi\sigma_i^2)^{n_i/2}} \exp\left\{ -\frac{1}{2\sigma_i^2} \left[(n_i-1)s_i^2 + n_i(\bar{y}_{i.} - b_1x_{i1} - \ldots - b_qx_{iq})^2 \right] \right\} \qquad \ldots (1.6)$$

where $\bar{y}_{i.} = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}$ and $s_i^2 = \frac{1}{(n_i-1)} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2$ are respectively the sample mean and the sample variance of n_i observations from the ith group. From (1.6) it is clear that the joint density function of the observations depends on the parameters only through the statistics $(\bar{y}_{i.}, s_i^2)$, $1 \leq i \leq k$. Thus the statistics:

$$(\bar{y}_{1.}, s_1^2), \ldots, (\bar{y}_{k.}, s_k^2) \qquad (1.7)$$

form a set of sufficient statistics. As a consequence of the theory of sufficiency it now follows that the statistics given by (1.7) contain all the information that is relevant for the purpose of estimating the parameters $b_1,\ldots,b_q$ and $\sigma_1^2,\ldots,\sigma_k^2$.

It is well-known from the theory of statistical distributions that $\bar{y}_{1.},\ldots,\bar{y}_{k.}$, $s_1^2,\ldots,s_k^2$ are all mutually independent random variables with $\bar{y}_{i.}$ having normal distribution and $(n_i-1)s_i^2/\sigma_i^2$ having χ^2 distribution with $(n_i - 1)$ degrees of freedom ([7], p. 147). It can now be seen that these sufficient statistics admit the following model:

$$\bar{y}_{1.} = b_1 x_{11} + \ldots + b_q x_{1q} + \epsilon_1 ; \quad (n_1-1)s_1^2 \sim \sigma_1^2 \chi^2_{(n_1-1)} \qquad (1.8)$$

$$\bar{y}_{2.} = b_1 x_{21} + \ldots + b_q x_{2q} + \epsilon_2 ; \quad (n_2-1)s_2^2 \sim \sigma_2^2 \chi^2_{(n_2-1)}$$

$$\cdots$$

$$\bar{y}_{k.} = b_1 x_{k1} + \ldots + b_q x_{kq} + \epsilon_k ; \quad (n_k-1)s_k^2 \sim \sigma_k^2 \chi^2_{(n_k-1)}$$

where all the s_i^2 and the ϵ_j's are mutually independent random variables with ϵ_i being normal with zero mean and variance σ_i^2/n_i and $(n_i - 1)s_i^2/\sigma_i^2$ being χ^2 with $(n_i - 1)$ degrees of freedom.

This model can be deduced from (1.4). It follows from the theory of sufficiency that (1.8) possesses as much information about the parameters under study as does (1.4). Actually if desired, one can through randomization obtain a new set of observations y_{ij}'s from $\bar{y}_{i.}$ of (1.8) which would be statistically indistinguishable from those of (1.4). Thus the two models (1.8) and (1.4) are statistically equivalent. From now on we therefore restrict our attention to the simpler model (1.8) rather than (1.4). A very important advantage of the model (1.8) is that it allows one to use s_i^2's as preliminary estimates of σ_i^2's in the estimation of the parameters $b_1,\ldots,b_q$. As we shall see later this usage of s_i^2's does not unduly influence the distributional properties of estimators of $b_1,\ldots,b_q$, because of the independence of s_i^2's from the $\bar{y}_i$'s.

2. Estimation of parameters.

Let us now proceed to the estimation of parameters $b_1,\ldots,b_q$ and $\sigma_1^2,\ldots,\sigma_k^2$ by using the following two step approach.

Step 1. Preliminary estimation of variances.

It is clear from the preceding reduction that the sample variances s_i^2's provide unbiased estimators of the corresponding variances σ_i^2's. It is easily seen that

$$V(s_i^2) = 2\,\sigma_i^4/(n_i-1) \qquad (2.1)$$

Since variance of s_i^2 approaches zero as n_i approaches infinity, it follows that s_i^2's are also consistent estimators of σ_i^2's.

Note: The estimator T_n based on a sample of size n is said to be consistent for a parameter θ if $\lim_{n\to\infty} P(|T_n - \theta| > \epsilon) = 0$ for every $\epsilon > 0$; the estimator T_n is said to converge to θ in quadratic mean if $\lim_{n\to\infty} E(T_n - \theta)^2 = 0$. Convergence in quadratic mean implies convergence in probability and thus consistency ([7], p.281).

If all the σ_i^2's are unequal we suggest the use of s_i^2's as our preliminary estimates of σ_i^2's for our second stage of estimation. If this is not so, then we can divide the k groups into c subgroups, say, such that within each subgroup the variances are equal. The sample variances within each subgroup can then be pooled in the usual way in order to obtain a more efficient pooled estimators of the corresponding common subgroup variances. For example if $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2$, say, then the more efficient estimator of the common variance for the three groups is given by:

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + (n_3 - 1)s_3^2}{(n_1 + n_2 + n_3 - 3)} \tag{2.2}$$

This pooling provides the minimum variance unbiased estimators of the subgroup variances based on $s_1^2, \ldots, s_k^2$. A justification for this is furnished by the following theorem:

<u>Theorem 2.1</u>. Let $m_1 s_1^2, m_2 s_2^2, \ldots, m_r s_r^2$ be independent $\sigma^2 \chi^2$ variables each with $m_1, m_2, \ldots, m_r$ degrees of freedom respectively, where σ^2 is the unknown parameter under estimation. Then $t = m_1 s_1^2 + m_2 s_2^2 + \ldots + m_r s_r^2$ is a complete sufficient statistic for this problem and

$$(m_1 s_1^2 + m_2 s_2^2 + \ldots + m_r s_r^2)/(m_1 + m_2 + \ldots + m_r) \tag{2.3}$$

is the unique minimum variance unbiased estimator of σ^2 based on $s_1^2, s_2^2, \ldots, s_r^2$.

(Note: A sufficient statistic is complete if no function of it has zero expectation for all values of the parameter unless it is itself a zero function. Every function of a complete sufficient statistic is a minimum variance unbiased estimator of its expected value (Cf. [7], p. 261).)

Proof: The joint density function of $m_1 s_1^2, \ldots, m_r s_r^2$ is given by:

$$f(m_1 s_1^2, \ldots, m_r s_r^2) = k \prod_{i=1}^{r} \left(\frac{m_i s_i^2}{\sigma^2}\right)^{\frac{m_i}{2} - 1} \exp(-m_i s_i^2 / 2\sigma^2)$$

$$= h(s_1^2, s_2^2, \ldots, s_r^2) \left(\frac{1}{\sigma^2}\right)^{\sum (m_i - 2)/2} \exp(-1/2 \sum_{i=1}^{r} m_i s_i^2 / \sigma^2) \tag{2.4}$$

By the Neyman-Fisher factorization theorem it follows that $t = \sum_{i=1}^{r} m_i s_i^2$ is a sufficient statistic, and it is also easily seen to have $\sigma^2 \chi^2$ distribution with $\sum_{i=1}^{r} m_i$ degrees of freedom, which is complete since it is a special case of the Gamma distribution. Consequently $\sum_{i=1}^{r} m_i s_i^2$ is a complete sufficient statistic in this problem. Thus $\sum_{i=1}^{r} m_i s_i^2 / \sum_{i=1}^{r} m_i$ is an unbiased estimator of σ^2 based on the complete sufficient statistic and therefore it is the minimum variance unbiased estimator of σ^2. ||

In view of this theorem it is clear that pooling of the sample variances is the best approach for obtaining preliminary estimators of the subgroup variances σ_j^2, $1 \leq j \leq c$, and we shall use this technique of pooling of the variances in Chapter 4.

We now turn to the estimation of the parameters $b_1, \ldots, b_q$.

Step 2. Estimation of $b_1, \ldots, b_q$.

If the residual variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2$ were known, the best way of estimating the parameters $b_1,\ldots,b_q$ would be by minimizing the weighted sum of squares:

$$U = \sum_{i=1}^{k} n_i/\sigma_i^2 \ (\bar{y}_{i.} - b_1x_{i1} - \ldots - b_qx_{iq})^2 \tag{2.5}$$

Since σ_i^2's are not known, we replace these variances by their pooled estimates obtained by the technique in Step 1. We shall denote these by $\hat{\sigma}_i^2$, $1 \leq i \leq k$. Thus in order to estimate $b_1,\ldots,b_q$ we minimize the quantity:

$$V = \sum_{i=1}^{k} n_i/\hat{\sigma}_i^2 \ (\bar{y}_{i.} - b_1x_{i1} - \ldots - b_qx_{iq})^2 \tag{2.6}$$

The normal equations which determine the parameters $b_1,\ldots,b_q$ are as follows:

$$\begin{aligned} b_1 \textstyle\sum (n_i/\hat{\sigma}_i^2)x_{i1}^2 + \ldots + b_q \textstyle\sum (n_i/\hat{\sigma}_i^2)x_{iq}x_{i1} &= \textstyle\sum (n_i/\hat{\sigma}_i^2)\bar{y}_{i.}x_{i1} \\ &\ \cdots \\ b_1 \textstyle\sum (n_i/\hat{\sigma}_i^2)x_{i1}x_{iq} + \ldots + b_q \textstyle\sum (n_i/\hat{\sigma}_i^2)x_{iq}^2 &= \textstyle\sum (n_i/\hat{\sigma}_i^2)\bar{y}_{i.}x_{iq} \end{aligned} \tag{2.7}$$

These normal equations can be used in much the same way as in the equal variance case to obtain estimators of parameters $b_1,\ldots,b_q$. As we shall show later the substitution of $\hat{\sigma}_i^2$ in place of σ_i^2 does not adversely affect the least square properties of the estimators so obtained. The reason for this is that we have succeeded through sufficiency and pooling of sample variances in obtaining initial weights $n_i/\hat{\sigma}_i^2$ which are independent of the sample means $\bar{y}_{i.}$ which contain all the information about the parameters $b_1, \ldots, b_q$.

We now turn to the properties of these estimators.

3. Properties of estimated parameters.

For reasons of clarity in exposition in regard to the properties of estimators briefly introduced in the preceding section, we now reformulate our problem in matrix notation.

Consider the setup:

$$(Y, Xb, V, \hat{V}) \tag{3.1}$$

in which Y is the vector of n independent observations assumed to be normally distributed with $E(Y) = Xb$ and the dispersion matrix $D(Y) = V$, and $\hat{V}$ is an unbiased estimator of V which is independent of Y. The matrix X is assumed known. The parameters under study are the vector b and the dispersion matrix V. For simplicity we shall assume that the matrix V is nonsingular.

When V is nonsingular and known, the standard theory of least squares as formulated by Aitken in 1934 [1] provides the following estimation procedure:

1) Obtain $\hat{b}$ by minimizing the quadratic form

$$U = (Y - Xb)'V^{-1}(Y - Xb) \tag{3.2}$$

2) Let p be a q-vector and suppose that $p'b$ is estimable (i.e., has an unbiased estimator). Then use $p'\hat{b}$ as an unbiased estimator of $p'b$. We refer to this as the Aitken procedure of estimation.

The difficulty in our setup $(Y, Xb, V, \hat{V})$ arises from the fact that V is unknown and only an unbiased estimator of V is available through $\hat{V}$. Consequently we suggest modifying the Aitken procedure by replacing V by $\hat{V}$ in (3.2) and we shall refer to it as the modified Aitken procedure in the sequel.

The following well-known results concerning matrices will be found useful in our study [6].

Definition 3.1. Let $A_{n \times p}$ be an $n \times p$ matrix. Then $A^-_{p \times n}$ is called a g-inverse if the following conditions hold:

1) AA^- is symmetric
2) A^-A is symmetric
3) $AA^-A = A$
4) $A^-AA^- = A^-$

It is easily seen that if A is nonsingular then $A^- = A^{-1}$ and if A is a scalar then $A^- = 1/A$ for $A \neq 0$ and $= 0$ for $A = 0$.

The g-inverse of a matrix exists, is unique and has the following properties [6]:

1) $R(AA^-) = R(A)$ where R denotes rank of a matrix in question.

   Since $A = AA^-A$, $R(A) \leq R(AA^-) \leq R(A)$.

2) $H = AA^-$ is a projection matrix on $\langle H \rangle = \langle A \rangle$, where $\langle . \rangle$ denotes the vector space spanned by the column vectors of the matrix within.

   Since H is symmetric by assumption and $H^2 = AA^-AA^- = AA^- = H$, H is idempotent, hence H is a projection matrix. Further

$Hx = AA^{-}x = A(A^{-}x)$ which is a vector in the column space of A. So $\langle H \rangle \subseteq \langle A \rangle$. Since $R(H) = R(A)$, $\langle H \rangle = \langle A \rangle$.

3) $H_r = A^{-}A$ is a projection matrix onto the row space of $A = \langle A' \rangle$.

By assumption H_r is symmetric and $H_r^2 = A^{-}AA^{-}A = A^{-}A = H_r$ so that H_r is idempotent and is therefore a projection matrix. Also $H_r x = H_r' x = A'A^{-\prime}x = A'(A^{-\prime}x)$ and so $\langle H_r \rangle \subseteq \langle A' \rangle$; also $R(H_r) = R(A')$. Therefore $\langle H_r \rangle = \langle A' \rangle$.

4a) If A is symmetric then so is A^{-}.

The proof follows on showing that the matrix $A^{-\prime}$ satisfies the properties of g-inverse of A. So by uniqueness of the g-inverse $A^{-} = A^{-\prime}$, i.e., A^{-} is symmetric.

4b) $(A')^{-} = (A^{-})'$.

The proof follows on showing that $A^{-\prime}$ satisfies the properties of the unique g-inverse of A'.

4c) If A is a projection matrix then $A^{-} = A$.

Here A satisfies the properties of its own generalized inverse.

4d) $A^{--} = A$.

This result follows from the symmetry in the definition of g-inverse with respect to A and A^{-}.

4e) $(AB)^{-} \neq B^{-}A^{-}$ in general.

For example let $A = [1, 0]$, $B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Then $A^{-} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $B = [1/2, 1/2]$. Thus $(AB)^{-} = 1$ and $B^{-}A^{-} = 1/2$.

5) $A^- = (A'A)^- A'$

Let $A^* = (A'A)^- A'$. Then AA^* and A^*A are easily seen to be symmetric matrices because of the symmetry of $(A'A)^-$. Further $A^*AA^* = (A'A)^- A'A(A'A)^- A' = (A'A)^- A' = A^*$. Also $AA^*A = A(A'A)^- A'A = AH_r$, where H_r denotes projection onto the row space of $A'A$ = column space of $A'A$, because of symmetry of $A'A$, = column space A'. Now $AH_r = (H_r'A')' = (H_rA')' = (A')' = A$. Thus $AA^*A = A$. Thus A^* satisfies all the properties of the unique g-inverse of A.

6) $(A'A)^- = A^- A'^-$

We have $A^- A'^- = (A'A)^- A'A'^- = (A'A)^-(A^- A)' = (A'A)^-(A^- A) = (A'A)^- H_r = [\ H_r'(A'A)^-]' = [\ H_r(A'A)^-]' = [\ A^- A(A'A)^-]' = [\ (A'A)^- A'A(A'A)^-]' = [\ (A'A)^-]' = (A'A)^-$.

7) The system of linear equations: $Ax = b$ is consistent (has at least one solution) if and only if $Hb = b$ where $H = AA^-$.

The system is consistent if and only if b is in $< A >$. Since H is a projection onto $< A >$, $b \in < A >$ if and only if $Hb = b$.

8) Let $Ax = b$ be a consistent system of linear equations. Then the system admits the following general solution:

$$x = A^- b + (I - H_r)y \tag{3.3}$$

where $H_r = A^- A$ and y is an arbitrary vector. Since the system is consistent $b = A(A^- b)$. So $A^- b$ is a particular solution of the system $Ax = b$. Further $A(I - H_r)y = (A - AH_r)y = (A - AA^- A)y = 0$. So

$(I - H_r)y$ provides solutions to the homogeneous case. Since $R(I - H_r) = n - \text{Rank}(A)$ which is the nullity of A, it follows that $< I - H_r >$ provides all solutions to the homogeneous case. Thus $x = A^- b + (I - H_r)y$ provides all possible solutions of the consistent system $Ax = b$.

We now turn to the study of estimable parameters in our previous setup (modified Aitken procedure).

Lemma 3.1. A necessary and sufficient condition that $p'b$ admits a linear unbiased estimator is that $p \in < X' >$.

Proof: Let $\ell'Y$ be an unbiased estimator of $p'b$. Then $E(\ell'Y) = \ell'Xb = p'b$ for all b if and only if $\ell'X = p'$ or $p = X'\ell$, i.e., if and only if $p \in < X' >$. ||

Returning to the modified Aitken procedure as outlined in the earlier part of this section, we see that the quadratic form

$$(Y - Xb)'\hat{V}^{-1}(Y - Xb)$$

is minimized when

$$X'\hat{V}^{-1}Xb = X'\hat{V}^{-1}Y \tag{3.4}$$

We refer to (3.4) as the normal equations for estimation of unknown parameters b. Clearly the general solution of the system of equations in (3.4) is given by:

$$b = (X'\hat{V}^{-1}X)^- X'\hat{V}^{-1}Y + (I - H_r)z \tag{3.5}$$

where $H_r = (X'\hat{V}^{-1}X)^- X'\hat{V}^{-1}X$ and z is an arbitrary vector.

It should be noted that H_r represents the projection matrix onto the row space of $(X'\hat{V}^{-1}X)$ = column space of X' = row space of X. Also $(I - H_r)$ represents projection onto the subspace orthogonal to the column space of X'.

The following theorem shows that all linear estimable parameters under the modified Aitken procedure possess a unique unbiased estimator.

Theorem 3.1. Let $p'b$ be a linear estimable parameter. Then for any solution $\hat{b}$ of (3.4), $p'\hat{b}$ is unique and unbiased for $p'b$.

Proof:

To prove the uniqueness of $\hat{b}$ we note that $p \in \langle X' \rangle$ so that p is orthogonal to $\langle I - H_r \rangle$. Therefore $p'(I - H_r)z = 0$ for all z. Thus from (3.5) we have

$$p'\hat{b} = p'(X'\hat{V}^{-1}X)^{-}X'\hat{V}^{-1}Y \tag{3.6}$$

This establishes the uniqueness of the estimator $p'\hat{b}$.

Now to prove the unbiasedness we make use of the independence of Y and $\hat{V}$. Clearly

$$E(p'\hat{b}) = E_{\hat{V}}\,[E(p'\hat{b}|\hat{V})] = E_{\hat{V}}\,[E(p'(X'\hat{V}^{-1}X)^{-}X'\hat{V}^{-1}Y|\hat{V})] \tag{3.7}$$

Now, because of the independence of Y and $\hat{V}$, the conditional distribution of Y is same as its unconditional distribution and therefore the above equation equals

$$E_{\hat{V}}\,[p'(X'\hat{V}^{-1}X)^{-}X'\hat{V}^{-1}E(Y)] = E_{\hat{V}}\,[p'(X'\hat{V}^{-1}X)^{-}X'\hat{V}^{-1}Xb] \tag{3.8}$$

Since for a given nonsingular V, $(X'\hat{V}^{-1}X)^{-}X'\hat{V}^{-1}X$ represents the projection matrix onto the row space of $X'\hat{V}^{-1}X$ = column space of X' = row space of X, the equation (3.8) gives

$$E(p'\hat{b}) = E_{\hat{V}}(p'X^{-}Xb) \tag{3.9}$$

Since $p \in < X' >$, this implies $X^{-}Xp = p$ so that $p'X^{-}X = (X^{-}Xp)' = p'$ and consequently (3.9) becomes

$$E(p'\hat{b}) = E_{\hat{V}}(p'b) = p'b \tag{3.10}$$

This completes the proof.

The theorem below furnishes the sampling variance of the estimator considered in the preceding theorem.

<u>Theorem 3.2.</u> Let $p'b$ be a linear estimable parameter and $p'\hat{b}$ its unbiased estimator obtained by the modified Aitken procedure. Then

$$\mathcal{V}(p'b) = p'E_{\hat{V}}\ [(X'\hat{V}^{-1}X)^{-}X'\hat{V}^{-1}V\hat{V}^{-1}X(X'\hat{V}^{-1}X)^{-}]\ p \tag{3.11}$$

<u>Proof</u>:

We need the following well-known result for our proof (Cf. [7], p.79).

$$\mathcal{V}(T) = \mathcal{V}[E(T|W)] + E_{W}[\mathcal{V}(T|W)] \tag{3.12}$$

Since $p'b$ is linearly estimable it follows from the preceding theorem that:

$$p'\hat{b} = p'(X'\hat{V}^{-1}X)X'\hat{V}^{-1}Y \tag{3.13}$$

and

$$E(p'\hat{b}|\hat{V}) = p'b \tag{3.14}$$

Therefore we now have from (3.12)

$$\mathcal{V}(p'\hat{b}) = E_{\hat{V}}\,[\,\mathcal{V}(p'\hat{b}|\hat{V})] + \mathcal{V}[E(p'\hat{b}|\hat{V})] = \tag{3.15}$$

$$E_{\hat{V}}\,[\,\mathcal{V}(p'(X'\hat{V}^{-1}X)^{-}X'\hat{V}^{-1}Y|\hat{V})] + \mathcal{V}(p'b) =$$

$$E_{\hat{V}}\,[p'(X'\hat{V}^{-1}X)^{-}X'\hat{V}^{-1}D(Y|\hat{V})\hat{V}^{-1}X(X'\hat{V}^{-1}X)^{-}p] + 0$$

where $D(Y|\hat{V})$ denotes the conditional dispersion matrix of Y given $\hat{V}$. Since Y and $\hat{V}$ are independent, $D(Y|\hat{V}) = D(Y) = V$. So that

$$\mathcal{V}(p'\hat{b}) = E_{\hat{V}}\,[p'(X'\hat{V}^{-1}X)^{-}X'\hat{V}^{-1}V\hat{V}^{-1}X(X'\hat{V}^{-1}X)^{-}p] \tag{3.16}$$

$$= p'E_{\hat{V}}\,[(X'\hat{V}^{-1}X)^{-}X'\hat{V}^{-1}V\hat{V}^{-1}X(X'\hat{V}^{-1}X)^{-}]p$$

This completes the proof.

The above expression for variance of $p'\hat{b}$ cannot be further simplified without additional information about the probability distribution of $\hat{V}$. Nonetheless, we will later study under certain regularity conditions the asymptotic behavior of the variance of $p'\hat{b}$.

Corollary 1. Let $p'b$ and $q'b$ be two linear estimable parameters. Then

$$\text{Cov}(p'\hat{b}, q'\hat{b}) = p'E_{\hat{V}}\,[(X'\hat{V}^{-1}X)^{-}X'\hat{V}^{-1}V\hat{V}^{-1}X(X'\hat{V}^{-1}X)^{-}]q \tag{3.17}$$

<u>Corollary 2</u>. Let the design matrix X be of full rank. Then $X'V^{-1}X$ is nonsingular and in this case the vector of parameters b is linearly estimable with

$$\hat{b} = (X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}Y \qquad (3.18)$$

and

$$D(\hat{b}) = E_{\hat{V}}\,[(X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}V\hat{V}^{-1}X(X'\hat{V}^{-1}X)^{-1}] \qquad (3.19)$$

We now turn to the asymptotic properties of the estimators derived under the modified Aitken procedure.

CHAPTER FOUR

ASYMPTOTIC PROPERTIES OF ESTIMATORS

1. Introduction. In the preceding chapter we introduced the modified Aitken procedure under the setup $(Y, Xb, V, \hat{V})$ for estimation of unknown parameters b. This procedure furnishes unbiased estimators of all the linear estimable parameters. In this chapter we study the asymptotic properties of the estimators so obtained. We begin this study in a general manner.

For each integer $n \geq 1$, we consider the setup $(Y_n, Xb, V_n, \hat{V}_n)$ where Y_n denotes the vector of k normally distributed observations with $E(Y_n) = Xb$ and $D(Y_n) = V_n$, $\hat{V}_n$ denotes an unbiased estimator of the matrix V_n and is assumed independent of Y_n. Under the modified Aitken procedure estimates of parameters are obtained by minimizing the quantity $(Y_n - Xb)'\hat{V}_n^{-1}(Y_n - Xb)$. It should be noted that here we are introducing the modified Aitken procedure through a sequence of regression models $[\ \{Y_n, Xb, V_n, \hat{V}_n\} : n \geq 1\ ]$ in which the unknown parameters b and the design matrix X are held fixed, while the observations Y_n, and dispersion matrices V_n and $\hat{V}_n$ vary with n. In practice however we have only a single model. The sequence of models considered here is introduced for the sole purpose of studying the asymptotic properties of our estimators. Before we impose the necessary regularity conditions on the sequence $[\ \{Y_n, Xb, V_n, \hat{V}_n\} :\ n \geq 1\ ]$, we need the following definition.

<u>Definition 1.1</u>. Let V be a $q{\times}q$ symmetric matrix and x be any q-vector. Then $x'Vx$ considered as a function of x is called the quadratic form of the matrix V. The matrix V is then said to be

a) positive definite if $x'Vx > 0$ for all nonnull x and negative definite if $-x'Vx$ is positive definite and

b) semipositive definite if $x'Vx \geq 0$ for all x and seminegative definite if $x'Vx \leq 0$ for all x ([7], p. 31).

The following regularity conditions will now be imposed on the sequence $[\ \{Y_n, Xb, V_n, \hat{V}_n\} : n \geq 1\ ]$:

<u>Regularity conditions for the modified Aitken model</u>.

1) For each n, the matrix V_n is nonsingular and there exists a positive definite matrix V and numbers a_n with $\lim_{n\to\infty} a_n = \infty$ such that:

$$V_n \leq V/a_n \tag{1.1}$$

where $A \leq B$ means that the matrix $(B - A)$ is semipositive definite.

2) For each n the matrix $\hat{V}_n$ is nonsingular and there exists numbers b_n with $\lim_{n\to\infty} b_n = \infty$ such that:

$$\hat{V}_n \geq \delta I/b_n \qquad \text{and} \qquad \lim_{n\to\infty} b_n a_n^{-2} = 0 \tag{1.2}$$

where I is the identity matrix and δ a positive number free of n.

The condition (1.2) may at first appear to be slightly arbitrary, nonetheless this condition plays a vital role in our analysis. It is motivated by the fact that in practice V is a diagonal matrix $\mathrm{diag}(\sigma_1^2,\ldots,\sigma_k^2)$ and $V_n = \mathrm{diag}(\sigma_1^2/n_1, \ldots, \sigma_k^2/n_k)$ and we assume that estimates of $\sigma_1^2,\ldots,\sigma_k^2$ are at least as large as δ; the a_n then corresponds to the smallest n_i and the b_n corresponds to the largest n_i.

In order to present the consistency of the modified Aitken model we need the following preliminary results regarding matrices and their eigenvalues.

2. Mathematical preliminaries.

Definition 2.1. Let A be a given square $n \times n$ matrix. Then a scalar λ is said to be an eigenvalue of A if and only if there exists an n-vector p of unit length $(p'p=1)$ such that $Ap = \lambda p$. The unit vector p is then said to be an eigenvector corresponding to the eigenvalue λ.

Lemma 2.1. Every square matrix $A_{n \times n}$ has n eigenvalues; these are given by the roots of the equation $|A - \lambda I| = 0$ and they may be zero, real or complex. If A is symmetric then all of its eigenvalues are real.

Definition 2.2. Let $A = (a_{ij})$ be a given $n \times n$ square matrix. Then $\mathrm{trace}(A) = \mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$ = sum of the diagonal elements of A.

Theorem 2.1 (Spectral Theorem [6], p. 5). Let A be an $n \times n$ symmetric matrix. Then there exist n unique real numbers $\lambda_1, \ldots, \lambda_n$ (some of which may be equal) and n mutually orthogonal unit vectors $p_1, \ldots, p_n$ such that:

a) $P'AP = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ where $P = (p_1, p_2, \ldots, p_n)$ is the orthogonal matrix of the eigenvectors of A.

b) $A = \lambda_1 p_1 p_1' + \lambda_2 p_2 p_2' + \ldots + \lambda_n p_n p_n'$

c) $|A - \lambda I| = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \ldots (\lambda_n - \lambda)$

d) $\mathrm{tr}(A) = \lambda_1 + \lambda_2 + \ldots + \lambda_n$

Lemma 2.2. Let $A_{n \times n}$ be a symmetric matrix. Then A is positive definite (pd), semipositive definite (spd), negative definite (nd), seminegative definite (snd) according as its eigenvalues are all positive, nonnegative, nonpositive or negative respectively ([7], p.35 & 50).

The following results refer to symmetric matrices [6].

Lemma 2.3.

$$A^- = \sum_{i=1}^{n} \lambda_i^- \, p_i p_i' \qquad (2.1)$$

where $\lambda_i^- = 1/\lambda_i$ if $\lambda_i \neq 0$ and 0 otherwise.

Lemma 2.4. Every spd matrix A has a unique spd square root $A^{1/2}$ which is given by $A^{1/2} = \sum_{i=1}^{n} \lambda_1^{1/2} \, p_i p_i'$.

Lemma 2.5. Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be the eigenvalues of an spd matrix A. We denote by $\|A\|$ the largest eigenvalue of A. The following propositions hold:

$$\sup_{u\neq 0} \frac{u'Au}{u'u} = \sup_{u'u=1} u'Au = \|A\| = \lambda_1 \tag{2.2}$$

and

$$\inf_{u\neq 0} \frac{u'Au}{u'u} = \inf_{u'u=1} u'Au = \lambda_n \tag{2.3}$$

See [7], p. 50, for details.

Lemma 2.6. Let $p_1,\ldots,p_m,p_{m+1},\ldots,p_q$ be mutually orthogonal unit eigenvectors corresponding to the eigenvalues $\lambda_1 \geq \ldots \geq \lambda_m > \lambda_{m+1} = \ldots \ldots = \lambda_q = 0$. Then

$$\inf_{\substack{u\neq 0 \\ u \in < p_1,\ldots,p_m >}} \frac{u'Au}{u'u} = \inf_{\substack{u'u=1 \\ u \in < p_1,\ldots,p_m >}} u'Au = \lambda_m \tag{2.4}$$

Proof:

By the spectral decomposition theorem 2.1 $A = \sum_{i=1}^{n} \lambda_i \, p_i p_i'$. Let $u = c_1p_1 + c_2p_2 + \ldots + c_mp_m$ be a given nonnull vector in $< p_1,\ldots,p_m >$. Then

$$u'Au = \sum_j^n \sum_i^m \sum_k^n c_j c_k \lambda_i \; p_j' p_i p_i' p_k \tag{2.5}$$

Since the p's are orthonormal,

$$\sum_j^n \sum_i^m \sum_k^n c_j c_k \lambda_i \; p_j' p_i p_i' p_k = \sum_i^m c_i^2 \lambda_i \geq \lambda_m \sum_i^m c_i^2 = \lambda_m u'u \tag{2.6}$$

where the equality holds if $c_m = 1$ and $c_1,\ldots,c_{m-1} = 0$. Therefore $\inf_{u'u=1} u'Au = \lambda_m$. ||

Lemma 2.7. The largest eigenvalue of A^- is the reciprocal of the smallest nonzero eigenvalue of A, where A is spd.

It follows from Lemma 2.3.

Lemma 2.8. Let $A \leq B$ be spd matrices. Then $\|A\| \leq \|B\|$ and the smallest eigenvalue of A is $\leq$ the smallest eigenvalue of B.

The following lemma plays a vital role in the next section.

Lemma 2.9. Let $D_{k \times k}$ be a symmetric positive definite matrix and X be $k \times q$, $k > q$ matrix of rank m. Then the smallest nonzero eigenvalue of $X'DX$ is at least as large as $d_k w_m$ where d_k denotes the smallest eigenvalue of D and w_m the smallest nonzero eigenvalue of $X'X$.

Proof:

Without loss of generality we can assume D is diagonal for otherwise we can express D in a diagonal form by applying a suitable orthogonal transformation. Now let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m > \lambda_{m+1} = \ldots = \lambda_q = 0$ be the eigenvalues of $X'DX$. Let $(p_1, \ldots, p_m, p_{m+1}, \ldots, p_q)$ be mutually orthogonal eigenvectors corresponding to $\lambda_1, \ldots, \lambda_q$. Then by the spectral theorem 2.1

$$X'DX = \sum_{i=1}^{q} \lambda_i \, p_i p_i' \qquad (2.7)$$

Let $e \in \langle p_1, \ldots, p_m \rangle$ and $e'e = 1$ and suppose that $e = c_1 p_1 + \ldots + c_m p_m$. Then from Lemma 2.6 it follows that

$$\inf_{\substack{e'e=1 \\ e \in \langle p_1,\ldots,p_m \rangle}} e'X'DXe = \lambda_m \tag{2.8}$$

where λ_m is the smallest nonzero eigenvalue of $X'DX$ and $p_1,\ldots,p_m$ are mutually orthogonal eigenvectors of all the nonzero eigenvalues of $X'DX$. Since D is pd, the quadratic form $e'X'DXe = (Xe)'D(Xe)$ is never zero unless $Xe = 0$. Since the vectors p_j, $m + 1 \leq j \leq q$ are orthogonal eigenvectors with eigenvalues zero, it follows from (2.7) that $e'X'DXe = 0$ when $e = p_j$, $m + 1 \leq j \leq q$. This implies that $Xp_j = 0$ for $m + 1 \leq j \leq q$. Consequently the row vectors of X are orthogonal to the space $\langle p_{m+1},\ldots,p_q \rangle$. So $\langle X' \rangle \subset \langle p_{m+1},\ldots,p_q \rangle^{\perp} = \langle p_1,\ldots,p_m \rangle$. Since $R(X') = R(\langle p_1,\ldots,p_m \rangle) = m$, $\langle X' \rangle = \langle p_1,\ldots,p_m \rangle$. In view of this last equality we have from (2.8) the following result:

$$\inf_{\substack{e'e=1 \\ e \in \langle X' \rangle}} e'X'DXe = \lambda_m \tag{2.9}$$

Next, let $D = \mathrm{diag}(d_1,\ldots,d_k)$ and $Xe = \ell = (\ell_1,\ldots,\ell_k)'$. Then

$$e'X'DXe = \sum_{i=1}^{k} d_i \ell_i^2 \geq d_k \sum_{i=1}^{k} \ell_i^2 = d_k(e'X'Xe) \tag{2.10}$$

From (2.9) and (2.10) it follows that

$$\lambda_m = \inf_{\substack{e'e=1 \\ e \in \langle X' \rangle}} e'X'DXe \geq d_k \inf_{\substack{e'e=1 \\ e \in \langle X' \rangle}} e'X'Xe \tag{2.11}$$

We now proceed to analyze the eigenvalues of $X'X$. Let $w_1 \geq w_2 \geq \ldots \geq w_m > = \ldots = w_q = 0$ be the eigenvalues of $X'X$. Let $v_1, \ldots, v_q$ be the corresponding eigenvectors. It is now easily seen that $Xv_j = 0$ for $m + 1 \leq j \leq q$. Therefore the rows of X are orthogonal to the vectors $v_{m+1}, \ldots, v_q$ so that $< X' > \subset < v_{m+1}, \ldots, v_q >^{\perp}$ $= < v_1, \ldots, v_m >$. Since $R(X') = m = R(< v_1, \ldots, v_m >)$, $< X' > = < v_1, \ldots, v_m >$. In view of (2.11) this gives

$$\lambda_m \geq d_k \inf_{\substack{e'e=1 \\ e \in < v_1, \ldots, v_m >}} e'X'Xe \tag{2.12}$$

From the preceding considerations it follows that

$$X'X = w_1 v_1 v_1' + w_2 v_2 v_2' + \ldots + w_m v_m v_m' \tag{2.13}$$

Now let $e = k_1 v_1 + \ldots + k_m v_m$ be a given but quite arbitrary unit vector in $< v_1, \ldots, v_m >$. Then from Lemma 2.6 and (2.13) it follows that

$$\inf_{\substack{e'e=1 \\ e \in < v_1, \ldots, v_m >}} e'X'Xe \geq w_m \tag{2.14}$$

for all unit vectors belonging to $< v_1, \ldots, v_m >$. The equality is actually achived when $k_m = 1$ and $k_1 = \ldots = k_{m-1} = 0$. From (2.12) and (2.14) we have

$$\lambda_m \geq d_k w_m$$

This completes the proof.

With these preliminary results we are now in a position to present the consistency of estimators under the modified Aitken procedure.

3. Consistency of estimators.

We consider the sequence of regression models introduced in Section 1 and unless stated to the contrary we assume that they satisfy regularity conditions (1.1) and (1.2). We recall that if $p'b$ is a linear estimable parameter then under the modified Aitken procedure its unique unbiased estimator is

$$p'\hat{b} = p'(X'\hat{V}_n^{-1}X)^{-}X'\hat{V}_n^{-1}Y_n \tag{3.1}$$

with

$$(p'\hat{b}) = p'E_{\hat{V}_n}[(X'\hat{V}_n^{-1}X)^{-}X'\hat{V}_n^{-1}V_n\hat{V}_n^{-1}X(X'\hat{V}_n^{-1}X)^{-}]p \tag{3.2}$$

In order to prove the consistency of every linear estimable parameter we need the following well-known result ([7], p. 281).

<u>Lemma 3.1</u>. Let $\{t_n : n \geq 1\}$ be a sequence of estimators of a parameter θ. Then a sufficient condition for the sequence $\{t_n : n \geq 1\}$ to be consistent is that $\lim_{n\to\infty} E(t_n - \theta)^2 = 0$.

<u>Theorem 3.1</u>. Under the regularity conditions (1.1) and (1.2) every unbiased estimator $p'\hat{b}$ as given by (3.1) is a consistent estimator of $p'b$.

<u>Proof</u>:

Since $p'\hat{b}$ is an unbiased estimator, by virtue of Lemma 3.1 it suffices to show that $(p'\hat{b})$ tends to zero as n tends to infinity.

This will be so if we show that $E[(X'\hat{V}_n^{-1}X)^- X'\hat{V}_n^{-1}V_n\hat{V}_n^{-1}X(X'\hat{V}_n^{-1}X)^-]$ tends to zero as n tends to infinity. Now

$$\| E[(X'\hat{V}_n^{-1}X)^- X'\hat{V}_n^{-1}V_n\hat{V}_n^{-1}X(X'\hat{V}_n^{-1}X)^-] \| \tag{3.3}$$

$$= E[\ \|(X'\hat{V}_n^{-1}X)^- X'\hat{V}_n^{-1}V_n\hat{V}_n^{-1}X(X'\hat{V}_n^{-1}X)^-\|\ ]$$

$$= E[\ \|(X'\hat{V}_n^{-1}X)^- X'\hat{V}_n^{-1/2}\hat{V}_n^{-1/2}V_n\hat{V}_n^{-1/2}\hat{V}_n^{-1/2}X(X'\hat{V}_n^{-1}X)^-\|\ ]$$

$$= E[\ \|L'\hat{V}_n^{-1/2}V_n\hat{V}_n^{-1/2}L\|\ ]$$

where $L = \hat{V}_n^{-1/2}X(X'\hat{V}_n^{-1}X)^-$. Since $L'L = [(X'\hat{V}_n^{-1}X)^- X'\hat{V}_n^{-1/2}\hat{V}_n^{-1/2}X(X'\hat{V}_n^{-1}X)^-]$ $= [(X'\hat{V}_n^{-1}X)^- X'\hat{V}_n^{-1}X(X'\hat{V}_n^{-1}X)^-] = [(X'\hat{V}_n^{-1}X)^-]$, the last expression in (3.3) satisfies the following inequality:

$$\le E[\ \|\hat{V}_n^{-1/2}V_n\hat{V}_n^{-1/2}\|\ \|L'L\|\ ] = E[\ \|\hat{V}_n^{-1/2}V_n\hat{V}_n^{-1/2}\|\ \|(X'\hat{V}_n^{-1}X)^-\|\ ]$$
$$\le E[\ \|V_n\|\ \|V_n^{-1}\|\ \|(X'\hat{V}_n^{-1}X)^-\|\ ] \tag{3.4}$$

Now from regularity conditions (1.1) and (1.2) and Lemma 2.8 it follows that

$$\|V_n\|\ \|\hat{V}_n^{-1}\| \le \frac{\|V\|}{a_n}\frac{b_n}{\delta} = \frac{b_n\ \|V\|}{\delta a_n} \tag{3.5}$$

Since $\|(X'\hat{V}_n^{-1}X)^-\|$ = [reciprocal of the smallest nonzero eigenvalue of $(X'\hat{V}_n^{-1}X)$] , it follows from Lemma 2.9 that

$$\|(X'\hat{V}_n^{-1}X)^-\| \leq [w_m/\|\hat{V}_n\|]^{-1} = \|\hat{V}_n\|/w_m \tag{3.6}$$

where w_m denotes the smallest nonzero eigenvalue of $X'X$. From (3.5) and (3.6) it follows that

$$E[\ \|(X'\hat{V}_n^{-1}X)^-X'\hat{V}_n^{-1}V_n\hat{V}_n^{-1}X(X'\hat{V}_n^{-1}X)^-\|\ ] \tag{3.7}$$

$$\leq E[\ \|V\|\ \|\hat{V}_n^{-1}\|\ \|(X'\hat{V}_n^{-1}X)^-\|\ ] \leq E\left[\ \frac{b_n\ \|V\|}{\delta a_n}\ \frac{\|\hat{V}_n\|}{w_m}\ \right]$$

$$= \frac{b_n}{\delta\ a_n\ w_m}\ \|V\|\ E[\|\hat{V}_n\|] \leq \frac{b_n\ \|V\|}{\delta\ a_n\ w_m}\ E[\ tr\ \hat{V}_n\ ]$$

$$= \frac{b_n\ \|V\|}{\delta\ a_n\ w_m}\ tr(V_n) \leq \frac{b_n\ \|V\|}{\delta\ a_n\ w_m}\ k\|V_n\| \leq \frac{b_n\ k\ \|V\|^2}{\delta\ a_n^2\ w_m}$$

$$= (b_n\ a_n^{-2})\left[\frac{k\ \|V\|^2}{\delta\ w_m}\right] \tag{3.8}$$

from (1.1), where k denotes the number of diagonal elements in the matrix V_n.

Since by assumption (1.2) $\lim_{n\to\infty} b_n a_n^{-2} = 0$, it follows that

$$\lim_{n\to\infty} E[\ \|(X'\hat{V}_n^{-1}X)^-X'\hat{V}_n^{-1}V_n\hat{V}_n^{-1}X(X'\hat{V}_n^{-1}X)^-\|\ ] = 0 \tag{3.9}$$

This completes the proof of consistency.

4. Asymptotic optimality of the modified Aitken procedure.

The object of this section is to demonstrate that under certain conditions the modified Aitken procedure approaches the standard Aitken procedure as the sample size increases. For reasons of simplicity we shall confine our attention to the model (1.8) of Chapter 3, which is basically the regression model for our data in the succeeding chapter.

As a first step towards showing the optimality of the modified Aitken procedure we note that the estimator s_i^2 of the within group variance σ_i^2 admits the following representation:

$$s_i^2 = \frac{n_i}{n_i-1}\left[\frac{1}{n_i}\sum_{j=1}^{n_i} y_{ij}^2 - \bar{y}_{i.}\right] \tag{4.1}$$

and as a consequence of the strong law of large numbers, the s_i^2 converges with probability one to its expected value, namely, σ_1^2, $1 \leq i \leq k$. We also note that the normal equations (2.7) of Chapter 3 which determine the parameters $b_1,\ldots,b_q$ can be equivalently written as follows:

$$(X'NS_n^{-1}X)b = (X'NS_n^{-1})Y \tag{4.2}$$

where $N = \mathrm{diag}(n_1,\ldots,n_k)$ and $S_n^{-1} = \mathrm{diag}(s_1^{-2},\ldots,s_k^{-2})$. We also assume that $X'X$ is nonsingular. In this case every linear function of b is estimable and the unique estimator of parameters $b_1,\ldots,b_q$ is given by

$$\hat{b} = (X'NS_n^{-1}X)^{-1}X'NS_n^{-1}Y \tag{4.3}$$

Had the true dispersion matrix $V = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_k^2)$ been known the estimator of b under the standard Aitken procedure would be

$$\hat{b}_t = (X'N\, V^{-1}X)^{-1}X'N\, V^{-1}Y \tag{4.4}$$

We now proceed to show that under certain regularity conditions with probability one $\hat{b}$ and $\hat{b}_t$ are asymptotically equivalent. The proof of this involves a number of steps.

Definition 4.1. Let $\{A_n : n \geq 1\}$ be a given sequence of matrices of the same order. This sequence is said to converge to matrix A and we write $\lim_{n\to\infty} A_n = A$ if every element of A_n converges to the corresponding element of A as n tends to infinity.

The following propositions concerning convergence of matrices hold.

Lemma 4.1. Let $\{A_n : n \geq 1\}$ be a sequence of square matrices with $\lim_{n\to\infty} A_n = A$ and suppose that A is nonsingular. Then

a) $\lim_{n\to\infty} |A_n| = |A|$, where $|.|$ denotes the determinant of the matrix.

   This is so since determinant is a continuous operation involving the elements of the matrix.

b) every minor of A_n, A_n^{ij}, say, converges to the corresponding minor A^{ij} of A.

c) $\{A_n^{-1} : n \geq 1\}$ converges to A^{-1}.

   This is clear from (a) and (b).

Lemma 4.2. Suppose that $\lim_{n\to\infty} A_n = A$ and $\lim_{n\to\infty} B_n = B$. Then $\lim_{n\to\infty} A_n B_n = AB$.

The proof follows on observing that matrix multiplication is a continuous operation.

We now present the asymptotic equivalence of the two procedures.

Theorem 4.1. Suppose that the regression model (1.8) of Chapter 3 satisfies the following conditions:

1) $X'X$ is nonsingular
2) the within group sample sizes n_i increase to infinity in such a way that $\lim n_i/(n_1 + \ldots + n_k) = \omega_i$, $0 < \omega_i < 1$.

Then with probability one

$$\lim\ (X'NS_n^{-1}X)^{-1}X'NS_n^{-1} = (X'\Omega V^{-1}X)^{-1}X'\Omega V^{-1} \tag{4.5}$$

and

$$\lim\ (X'N V^{-1}X)^{-1}X'N V^{-1} = (X'\Omega V^{-1}X)^{-1}X'\Omega V^{-1} \tag{4.6}$$

where these matrices have been defined in (4.2), (4.3) and (4.4) and $\Omega = \mathrm{diag}(\omega_1,\ldots,\omega_k)$.

Proof:

It is easily seen that

$$(X'NS_n^{-1}X)^{-1}X'NS_n^{-1} = (X'\Gamma S_n^{-1}X)^{-1}X'\Gamma S_n^{-1} \tag{4.7}$$

where $\Gamma = \mathrm{diag}(n_1/(n_1+\ldots+n_k),\ldots,n_k/(n_1+\ldots+n_k))$. By assumption

$\lim \Gamma = \Omega$ and as we have pointed out in the remark following (4.1) that with probability one $\lim S_n = V$. Therefore $\lim S_n^{-1} = V^{-1}$ since V is nonsingular. Thus we have with probability one

$$\lim (X'\Gamma S_n^{-1} X) = (X'\Omega V^{-1} X)$$

and (4.8)

$$\lim (X'\Gamma S_n^{-1}) = X'\Omega\, V^{-1}$$

Combining these two equations and applying Lemma 4.1 and Lemma 4.2 it follows that

$$\lim (X'NS_n^{-1}X)^{-1}X'NS_n^{-1} = \lim (X'\Gamma S_n^{-1}X)^{-1}X'\Gamma S_n^{-1} \tag{4.9}$$

$$= (X'\Omega\, V^{-1}X)^{-1}X'\Omega\, V^{-1}$$

with probability one.

The proof of equation (4.6) follows analogously. ||

It is clear from equations (4.3), (4.4), (4.5) and (4.6) that under the regularity conditions of the above theorem, the modified Aitken procedure leads to the same estimators in the limit as does the standard Aitken procedure. Thus in this sense the two procedures become asymptotically equivalent as the sample sizes increase.

CHAPTER FIVE

ESTIMATION OF GROWTH CURVES BY LEAST SQUARE SPLINES

1. Statement of the problem.

A few years ago the University of New Mexico Medical Staff collected data on weight and biparietal diameter of live newborns and fresh abortuses within thirty minutes of birth. The primary object of this study was to develop a formula which could be used to predict the weight of a fetus or a newborn from its biparietal diameter (maximum skull breadth). Such weight prediction plays an important role in determination of fetal maturity in maternal diabetes and in cases of repeat caesarian section. It should be added that the biparietal diameter of a developing fetus is determined through the use of ultrasound without any known toxic effect to the mother of the fetus (Cf. [9]).

The basic data given in Table 1.1 of the Appendix consists of complete records on weight measured in grams and biparietal diameter measured in centimeters on 295 newborns and abortuses. In the succeeding sections we use the data from Table 1.1 to estimate the prediction formula.

Growth curves generally display different growth patterns in different regions and consequently have the handicap that knowledge of their form in one region does not adequately determine them over the entire region of interest. Since polynomial spline functions represent different polynomials in different regions determined by their knots, some of their characteristics are analogous to those of growth curves. Consequently a growth curve is better explained in terms of a polynomial spline rather than a single polynomial over the entire region of interest. We therefore carry out regression analysis of the data in Table 1.1 in the Appendix by the technique of weighted least square splines. This is based on the assumption of the following model:

$$y(x) = f(x) + \epsilon_x \tag{1.1}$$

where $y(x)$ denotes the weight, x the biparietal diameter, $f(x)$ a suitable spline function and the error ϵ_x is assumed to be independently normally distributed with zero mean and variance σ^2 which increases with x.

2. Examination and estimation of error variances.

Initial analysis of the data was carried out by first grouping the observations into 2 mm intervals of biparietal diameter; means and standard deviations for the body weight were then calculated within each interval which are summarized in the table on the next page.

Table 2.1

Within Group Means and Standard Deviations

Serial Number	Biparietal Diameter (cm)	Sample Size	Mean Weight (gm)	Standard Deviation
1	0.00		0.0	
2	2.85	1	113.0	
3	3.65	2	114.0	1.4
4	3.85	2	288.5	212.8
5	4.05	1	115.0	
6	4.25	2	206.5	12.0
7	4.45	4	319.5	83.9
8	4.65	2	383.0	100.4
9	4.85	6	376.3	188.9
10	5.05	1	765.0	
11	5.45	5	550.2	51.6
12	5.65	2	553.0	19.8
13	5.85	1	709.0	
14	6.05	2	701.5	290.6
15	6.25	4	800.8	43.0
16	6.45	1	1000.0	
17	6.85	3	982.7	184.2
18	7.05	2	1139.5	232.6
19	7.25	5	1116.4	215.9
20	7.45	6	1224.0	184.1
21	7.65	6	1456.3	329.0
22	7.85	10	1881.7	415.6
23	8.05	14	2541.4	575.1
24	8.25	24	2559.8	282.8
25	8.45	17	2574.9	371.8
26	8.65	22	2576.0	378.3
27	8.85	44	2717.7	275.4
28	9.05	8	2661.4	353.9
29	9.25	13	2975.8	373.6
30	9.45	15	3041.3	342.9
31	9.65	10	3512.7	282.3
32	9.85	7	3317.0	439.1
33	10.05	17	3750.6	385.2
34	10.25	9	3518.8	575.8
35	10.45	21	3788.1	592.5
36	10.65	5	4145.0	604.9
37	10.90	1	4338.0	

As is evident from the above table within group variance for body weight increases with biparietal diameter. Consequently we turned to the technique of weighted least square splines for analysis of the data.

First the whole region of interest for the biparietal diameter was divided into the following intervals [0, 7.65), [7.65, 10.25), [10.25 -) since an application of Bartlett's test ([3], pp. 159-167) showed that within each region the variances were homogeneous. Application of formula (2.2) of Chapter 3 gave us the following estimates of the three group variances.

Table 2.2

Estimated Pooled Variances

Biparietal Diameter (cm)	Sample Size	Center of Gravity	Pooled Variance Estimate
Group 1 [0 - 7.65)	47	5.85	25274.6
Group 2 [7.65 - 10.25)	208	8.85	129143.6
Group 3 [10.25 -)	35	10.45	348083.4

The center of gravity for each group was calculated by the formula $\bar{x} = \sum(n_i - 1)x_i / \sum(n_i - 1)$; the estimated pooled variances s_i^2 are assumed to be estimates of the error variance at the corresponding $\bar{x}_i$. This method provides a correction for any linear trend that might exist in the error variance within each group. Linear interpolation of these three within group variances yielded an estimate of σ_x^2 for each x in the region of interest.

3. Representation of regression equation.

We assume that the regression function $f(x)$ given by equation (1.1) of this chapter is representable by a cubic spline over the knots 0.0, 1.0, 3.65, 5.65, 7.65, 8.95, 10.25. The justification for the choice of these knots is as follows. First the knots 0.0, 7.65, 10.25 were chosen because this divided the entire range of biparietal diameter into groups of homogeneous within group variances. Since a cubic spline with the end conditions $f(0) = f'(0) = f''(0) = 0$ admits the form dx^3 in the first region which is not desirable for the whole region from 0.0 to 7.65, we introduced a new knot at 1.0. Further since our real region of interest begins with 4.0 cm we added an additional knot at 3.65 to obtain a better approximation in the region to the right. We then chose two additional knots at 5.65 and 8.95 since these are the midpoints of the intervals 3.65 - 7.65 and 7.65 - 10.25. These assumptions gave rise to the following regression model for the data under study:

$$y(x) = b_1B_1(x) + b_2B_2(x) + b_3B_3(x) + b_4B_4(x) +$$

$$b_5B_5(x) + b_6B_6(x) + b_7B_7(x) + \epsilon_x \qquad (3.1)$$

where $y(x)$ denotes the weight at biparietal diameter x, ϵ_x the error term with zero mean and variance σ_x^2, $b_1,\ldots,b_7$ the unknown parameters to be estimated, and $B_1,\ldots,B_7$ the spline functions given below

$$B_1(x) = (4/157.76)(x)_+^3 - (4/81.94)(1 - x)_+^3 + (4/77.38)(3.65 - x)_+^3 - (4/105.09)(5.65 - x)_+^3 + (4/406.98)(7.65 - x)_+^3$$

$$B_2(x) = (4/651.46)(1 - x)_+^3 - (4/112.36)(3.65 - x)_+^3 + (4/61.38)(5.65 - x)_+^3 - (4/69.96)(7.65 - x)_+^3 + (4/180.76)(8.95 - x)_+^3$$

$$B_3(x) = (4/279.84)(3.65 - x)_+^3 - (4/60.72)(5.65 - x)_+^3 + (4/27.04)(7.65 - x)_+^3 - (4/29.56)(8.95 - x)_+^3 - (4/102.62)(10.25 - x)_+^3$$

$$B_4(x) = (4/30.36)(x - 5.65)_+^3 - (4/6.76)(x - 7.65)_+^3 + (4/5.58)(x - 8.95)_+^3 - (4/15.55)(x - 10.25)_+^3$$

$$B_5(x) = (4/3.38)(x - 7.65)_+^3 - (4/1.69)(x - 8.95)_+^3 + (4/3.38)(x - 10.25)_+^3$$

$$B_6(x) = (4/1.3)(x - 8.95)_+^3 - (4/1.3)(x - 10.25)_+^3$$

$$B_7(x) = 4(x - 10.25)_+^3 \qquad (3.2)$$

We now turn to the estimation of the parameters $b_1,\ldots,b_7$.

4. Estimation of parameters.

As explained in Chapter 3, we now use the modified Aitken procedure for the purpose of estimating the parameters $b_1,\ldots,b_7$. The normal equations determining the parameters are obtained by minimizing

$$\sum_{i=1}^{37}(n_i/\hat{\sigma}_i^2)(\bar{y}_{i.} - b_1B_1(x) - \ldots - b_7B_7(x))^2 \qquad (4.1)$$

where n_i is the number of observations with biparietal diameter x_i, $\bar{y}_{i.}$ the average weight at x_i, and $\hat{\sigma}_i^2$ the estimate of the error variance at x_i obtained by the method described following Table 2.2 of this chapter.

The normal equations are given as follows:

$$\begin{aligned} b_1S_{11} + b_2S_{12} + \ldots + b_7S_{17} &= Q_1 \\ b_1S_{21} + b_2S_{22} + \ldots + b_7S_{27} &= Q_2 \\ &\;\;\cdots \\ b_1S_{71} + b_2S_{72} + \ldots + b_7S_{77} &= Q_7 \end{aligned} \qquad (4.2)$$

where S_{uv} and Q_u denote the following sum of products:

$$S_{uv} = \sum_{i=1}^{k}(n_i/\hat{\sigma}_i^2)\, B_u(x_i)B_v(x_i) \qquad (4.3)$$

$$Q_u = \sum_{i=1}^{k}(n_i/\hat{\sigma}_i^2)\, B_u(x_i)\,\bar{y}(x_i)$$

$$u, v = 1,\ldots,7 \text{ and } k = 37.$$

Using the program given in the Appendix at the end of this chapter, the following values are computed:

$$(S_{uv}) = \begin{matrix} 7\cdot98 & & & & & & \\ 6\cdot91 & 11\cdot21 & & & & & \\ 0\cdot68 & 3\cdot96 & 11\cdot46 & & & & \\ 0\cdot02 & 2\cdot97 & 64\cdot22 & 1443\cdot97 & & & \\ 0\cdot0 & 0\cdot10 & 18\cdot23 & 2000\cdot88 & 5661\cdot96 & & \\ 0\cdot0 & 0\cdot0 & 0\cdot14 & 575\cdot48 & 2388\cdot68 & 1306\cdot23 & \\ 0\cdot0 & 0\cdot0 & 0\cdot0 & 3\cdot21 & 16\cdot70 & 12\cdot98 & 0\cdot39 \end{matrix} \tag{4.4}$$

where (S_{uv}) is a symmetric matrix, and

$$\begin{pmatrix} Q_1 \\ Q_2 \\ Q_3 \\ Q_4 \\ Q_5 \\ Q_6 \\ Q_7 \end{pmatrix} = \begin{pmatrix} 10126.65 \\ 23438.79 \\ 79476.31 \\ 1322118.00 \\ 1783669.00 \\ 543366.00 \\ 3307.62 \end{pmatrix} \tag{4.5}$$

Using the values in (4.6) and (4.7) we obtain the following estimates of the parameters:

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \\ b_7 \end{pmatrix} = \begin{pmatrix} 235.97 \\ 978.74 \\ 2141.47 \\ 793.34 \\ -3.35 \\ 76.99 \\ -466.61 \end{pmatrix} \tag{4.6}$$

Thus in term of the above estimates the prediction formula is:

$$y(x) = 235.97\ B_1(x) + 978.74\ B_2(x) + 2141.47\ B_3(x) + 793.34\ B_4(x) - 3.35\ B_5(x) + 76.99\ B_6(x) - 466.61\ B_7(x) \quad (4.7)$$

The residual sum of squares in this case is:

$$R_o^2 = \sum_{i=1}^{k} (n_i / \hat{\sigma}_i^2)\bar{y}(x_i)^2 - b_1Q_1 - \dots - b_7Q_7 \quad (4.8)$$
$$= 8920834.0$$

An examination of the values of parameters indicated that the spline function $B_5(x)$ does not make a singificant contribution to the prediction formula. Thus we proceeded to test the hypothesis $H_o : b_5 = 0$. Under H_o the regression model reduces to

$$y(x) = b_1B_1(x) + \dots + b_4B_4(x) + b_6B_6(x) + b_7B_7(x) + \epsilon_x \quad (4.9)$$

A procedure similar to the earlier one yielded the following estimates of the parameters:

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_6 \\ b_7 \end{pmatrix} = \begin{pmatrix} 242.19 \\ 969.95 \\ 2158.01 \\ 789.92 \\ 71.80 \\ -408.81 \end{pmatrix} \quad (4.10)$$

and the residual sum of squares $R_1^2 = 8924111.0$. To test the significance of the hypothesis H_o, we calculate the variance ratio

$$F_{(1,30)} = (R_1^2 - R_o^2)/(R_o^2/k-7) = 0.011 \qquad (4.11)$$

This computed value of $F_{(1,30)}$ is found to be insignificant. Consequently the spline function $B_5(x)$ plays no significant role in the prediction equation.

Thus as far as the problem of prediction is concerned the following formula (without $B_5(x)$) should be used:

$$y(x) = 242.19\, B_1(x) + 969.95\, B_2(x) + 2158.01\, B_3(x) +$$
$$789.92\, B_4(x) + 71.80\, B_6(x) - 408.81\, B_7(x) \qquad (4.12)$$

5. Large sample tolerance limits.

We now proceed to set up large sample tolerance limits. The usual cases of interest are the 95th and 90th percentile tolerance limits. The formula that we use is as follows:

$$\text{Lower Tolerance Limit} = y(x) - z\,\hat{\sigma}_x$$
$$\text{Upper Tolerance Limit} = y(x) + z\,\hat{\sigma}_x \qquad (5.1)$$

where $\hat{\sigma}_x^2$ denotes an estimate of $\mathcal{V}(\varepsilon_x)$ and $z = 1.96$ for the 95th percentile and 1.645 for the 90th percentile.

We now turn to the estimation of σ_x for each x.

Estimation of σ_x.

Table 2.2 of this chapter gives us preliminary estimates of the within group variances σ_1^2, σ_2^2 and σ_3^2. Now a second estimate of these variances is also available through the residual variances. The quantities

$$\sum_{i=1}^{20} n_i(\bar{y}_{i.} - \hat{\bar{y}}_{i.})^2/16 , \quad \sum_{i=21}^{33} n_i(\bar{y}_{i.} - \hat{\bar{y}}_{i.})^2/11 \quad \text{and} \quad \sum_{i=34}^{37} n_i(\bar{y}_{i.} - \hat{\bar{y}}_{i.})^2/3 \qquad \ldots(5.2)$$

are asymptotically unbiased estimators of σ_1^2, σ_2^2 and σ_3^2 respectively. We were motivated to use the divisors 16, 11 and 3 since in the standard regression analysis one uses the divisor $k - q$ where k denotes the number of observations and q is the number of parameters estimated within that group. In our case the first group contributed to estimation of four parameters, the second to only one new parameter and the third group to an additional new parameter. Thus in the second group we subtracted an additional parameter to account for the effect of the other two splines that were already accounted for in the first interval. This is somewhat arbitrary but in no way influences the large sample properties of these estimators.

Pooling of these estimates from Table 2.1 and (5.2) gives the following combined estimates of variances for the three groups:

$$\hat{\sigma}^2_{1c} = [\sum_{i=1}^{20}(n_i-1)s_i^2 + \sum_{i=1}^{20} n_i(\bar{y}_{i.}-\hat{\bar{y}}_{i.})^2]/[\sum_{i=1}^{20}(n_i-1) + 16]$$

$$\hat{\sigma}^2_{2c} = [\sum_{i=21}^{33}(n_i-1)s_i^2 + \sum_{i=21}^{33} n_i(\bar{y}_{i.}-\hat{\bar{y}}_{i.})^2]/[\sum_{i=21}^{33}(n_i-1) + 11] \qquad (5.3)$$

$$\hat{\sigma}^2_{3c} = [\sum_{i=34}^{37}(n_i-1)s_i^2 + \sum_{i=34}^{37} n_i(\bar{y}_{i.}-\hat{\bar{y}}_{i.})^2]/[\sum_{i=34}^{37}(n_i-1) + 3]$$

Using these formulas we obtain the following estimates of the within group standard deviations:

$$\hat{\sigma}_1 = 205.3 \quad \text{at} \quad \bar{x} = 5.85$$

$$\hat{\sigma}_2 = 405.37 \quad \text{at} \quad \bar{x} = 8.85 \qquad (5.4)$$

$$\hat{\sigma}_3 = 573.4 \quad \text{at} \quad \bar{x} = 10.45$$

Again by the technique of linear interpolation we obtained estimates of σ_x at each x and then fitted a least square cubic spline to these estimates to obtain a smooth curve for $\hat{\sigma}_x$. These estimates were then used to obtain the lower and upper tolerance limits which are presented in the tables and figures that follow.

Table 5.1

The Predicted Body Weight from Biparietal Diameter,

Its Standard Deviation and 95% Tolerance Limits

BIPARIETAL DIAMETER (cm)	ESTIMATED WEIGHT (gm)	STANDARD DEVIATION	LOWER TOLERANCE LIMIT	UPPER TOLERANCE LIMIT
4·0	234	143	0	515
4·1	249	146	0	536
4·2	264	149	0	557
4·3	279	152	0	579
4·4	296	155	0	601
4·5	313	158	2	624
4·6	331	161	15	647
4·7	350	164	28	672
4·8	370	167	42	698
4·9	391	170	58	725
5·0	413	173	74	753
5·1	437	176	91	783
5·2	461	179	109	814
5·3	487	183	127	847
5·4	514	187	147	881
5·5	542	191	167	918
5·6	572	195	189	956
5·7	604	200	211	997
5·8	637	205	234	1039
5·9	671	210	258	1084
6·0	708	216	284	1132
6·1	746	222	311	1182
6·2	787	228	340	1234
6·3	830	234	371	1290
6·4	876	240	404	1348
6·5	924	247	439	1409
6·6	975	254	477	1473
6·7	1029	260	518	1540
6·8	1086	267	561	1611
6·9	1146	274	608	1685

Table 5.1 (Continued)

BIPARIETAL DIAMETER(cm)	ESTIMATED WEIGHT(gm)	STANDARD DEVIATION	LOWER TOLERANCE LIMIT	UPPER TOLERANCE LIMIT
7·0	1210	281	658	1762
7·1	1277	288	711	1843
7·2	1348	295	768	1927
7·3	1422	302	829	2016
7·4	1501	309	895	2108
7·5	1584	316	964	2204
7·6	1671	322	1038	2304
7·7	1762	329	1117	2408
7·8	1858	335	1200	2516
7·9	1956	342	1286	2627
8·0	2056	348	1373	2739
8·1	2157	354	1462	2853
8·2	2258	361	1550	2966
8·3	2357	367	1637	3078
8·4	2454	374	1720	3188
8·5	2548	381	1800	3295
8·6	2637	388	1875	3399
8·7	2720	396	1944	3497
8·8	2797	404	2005	3589
8·9	2866	412	2057	3675
9·0	2927	421	2100	3753
9·1	2979	431	2134	3824
9·2	3025	440	2161	3889
9·3	3067	451	2182	3951
9·4	3106	461	2201	4011
9·5	3146	472	2220	4071
9·6	3187	483	2240	4134
9·7	3233	494	2264	4201
9·8	3284	505	2294	4274
9·9	3344	516	2332	4355
10·0	3413	527	2380	4446
10·1	3495	537	2441	4549
10·2	3591	548	2516	4665
10·3	3703	558	2608	4797
10·4	3826	568	2712	4941
10·5	3953	578	2819	5086
10·6	4072	588	2919	5225
10·7	4174	598	3001	5347
10·8	4249	609	3055	5444
10·9	4288	620	3071	5505

Table 5.2

The Predicted Body Weight from Biparietal Diameter,

Its Standard Deviation and 90% Tolerance Limits

BIPARIETAL DIAMETER (cm)	ESTIMATED WEIGHT (gm)	STANDARD DEVIATION	LOWER TOLERANCE LIMIT	UPPER TOLERANCE LIMIT
4·0	234	143	0	470
4·1	249	146	8	490
4·2	264	149	17	510
4·3	279	152	28	530
4·4	296	155	40	552
4·5	313	158	52	574
4·6	331	161	66	596
4·7	350	164	80	620
4·8	370	167	95	645
4·9	391	170	111	671
5·0	413	173	128	698
5·1	437	176	146	727
5·2	461	179	165	757
5·3	487	183	185	789
5·4	514	187	206	822
5·5	542	191	228	857
5·6	572	195	250	894
5·7	604	200	274	933
5·8	637	205	299	975
5·9	671	210	325	1018
6·0	708	216	352	1064
6·1	746	222	381	1112
6·2	787	228	412	1163
6·3	830	234	445	1216
6·4	876	240	480	1272
6·5	924	247	517	1331
6·6	975	254	557	1393
6·7	1029	260	600	1458
6·8	1086	267	646	1527
6·9	1146	274	694	1598

Table 5.2 (Continued)

BIPARIETAL DIAMETER(cm)	ESTIMATED WEIGHT(gm)	STANDARD DEVIATION	LOWER TOLERANCE LIMIT	UPPER TOLERANCE LIMIT
7·0	1210	281	746	1673
7·1	1277	288	802	1752
7·2	1348	295	862	1834
7·3	1422	302	925	1920
7·4	1501	309	992	2010
7·5	1584	316	1064	2104
7·6	1671	322	1140	2202
7·7	1762	329	1221	2304
7·8	1858	335	1306	2410
7·9	1956	342	1393	2519
8·0	2056	348	1483	2629
8·1	2157	354	1574	2741
8·2	2258	361	1664	2852
8·3	2357	367	1753	2962
8·4	2454	374	1838	3070
8·5	2548	381	1920	3175
8·6	2637	388	1998	3276
8·7	2720	396	2068	3372
8·8	2797	404	2132	3462
8·9	2866	412	2187	3545
9·0	2927	421	2233	3620
9·1	2979	431	2270	3688
9·2	3025	440	2299	3750
9·3	3067	451	2324	3809
9·4	3106	461	2347	3866
9·5	3146	472	2369	3923
9·6	3187	483	2392	3982
9·7	3233	494	2420	4045
9·8	3284	505	2453	4115
9·9	3344	516	2494	4193
10·0	3413	527	2546	4280
10·1	3495	537	2610	4379
10·2	3591	548	2689	4493
10·3	3703	558	2784	4621
10·4	3826	568	2891	4761
10·5	3953	578	3002	4904
10·6	4072	588	3104	5040
10·7	4174	598	3189	5159
10·8	4249	609	3247	5252
10·9	4288	620	3266	5309

Figure 5.1

The Prediction Curve for Body Weights from Biparietal Diameter and Upper and Lower 95% Tolerance Limits

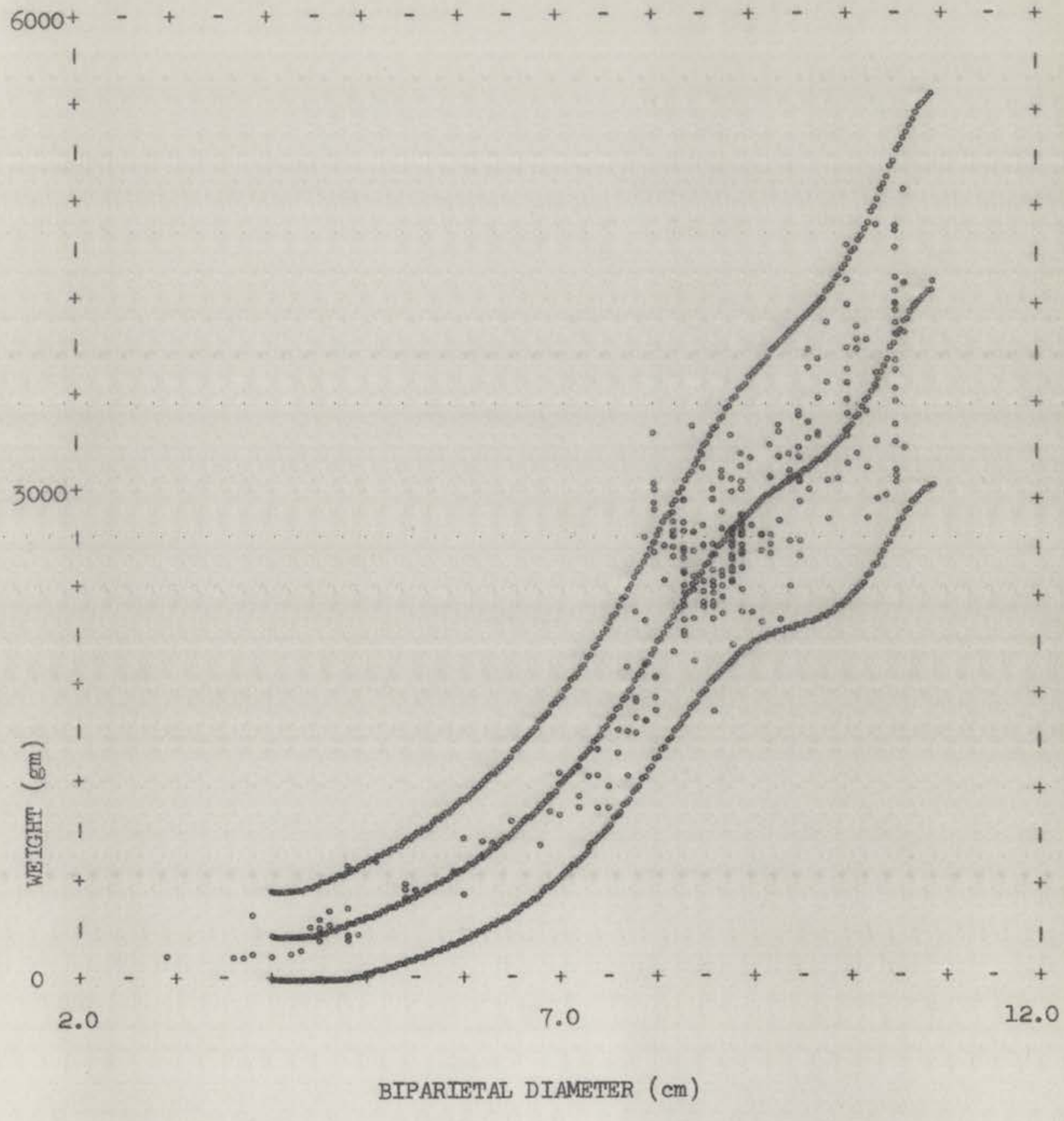
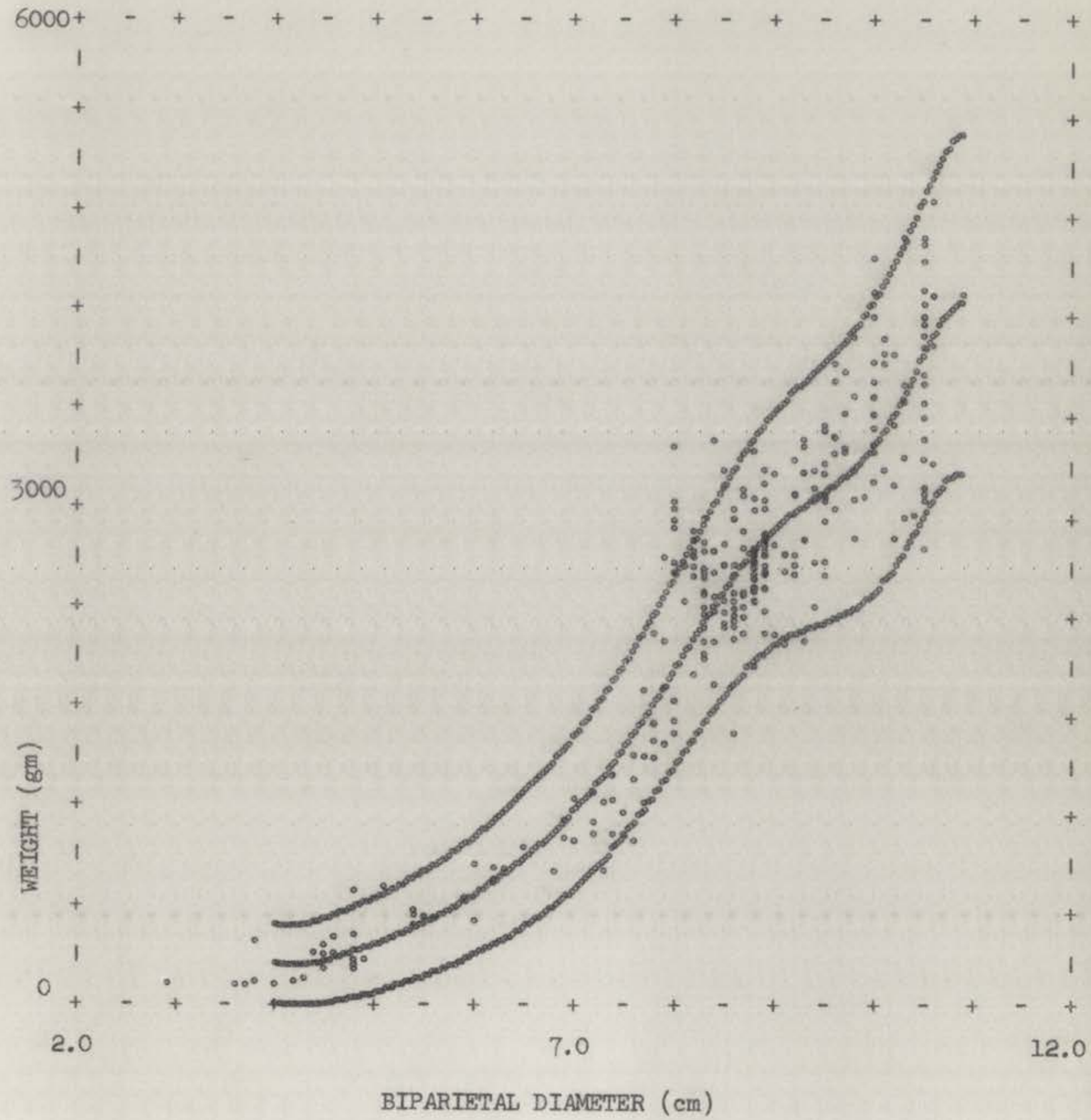


Figure 5.2

The Prediction Curve for Body Weights from Biparietal Diameter and Upper and Lower 90% Tolerance Limits



6. Conclusions.

It is worthwhile to conclude this chapter by pointing out some of the advantages and disadvantages of the use of spline functions in the study of growth curves. As we have noticed earlier growth curves generally display different growth patterns in different regions and consequently have the handicap that knowledge of their form in one region does not adequately determine them over the entire region of interest. Since polynomial spline functions represent different polynomials in different regions determined by their knots, their characteristics are analogous to those of growth curves. Consequently a growth curve is better explained in terms of a polynomial spline rather than a single polynomial over the entire region of interest. At the same time polynomial spline functions have local properties which are analogous to those of polynomial functions, for example they can be differentiated and integrated at any given point. The least square theory with spline functions is no more complicated than it is for instance with orthogonal polynomials. In practice electronic computers can be used to compute values of spline functions of a given basis accurately and rapidly.

One of the disadvantages with the use of splines is the somewhat arbitrary nature in which the knots are chosen. The simplest guideline that can be offered is to suggest the use of as few knots as possible thus ensuring greater smoothness and less overfitting. In addition the knots should be chosen close to inflection points. The reason for this is that cubic spline functions tend to have inflection points in the neighborhood of knots. Also with cubic splines one should try not to have more than one inflection point per interval. The choice of knots for our data was greatly influenced by these considerations. If the knots are poorly chosen, there is a tendency for the residuals to be autocorrelated. This can be used as a measure of how well the knots have been chosen. In practice the best rule of thumb is perhaps to obtain a scatter diagram of a sample of data, examine it very carefully and then carry out a trial analysis to arrive at the optimum choice of knots.

For those interested in further pursuing this line of research the next step would be to write a general purpose program using the codes of Carl de Boor based on a simple recursive formula for computation of B-splines [2]. As with Schoenberg's basis in order to incorporate the behavior of growth curves near zero the first few basis functions obtained by Carl de Boor's codes should be deleted.

Another line of investigation worth further pursuing would be a comparative study of the relative merits and demerits of regression analysis by least square polynomial splines and by least square polynomials; and to actually determine what added reduction in residual variance can be achieved by the use of polynomial splines.

Finally how the knots should be chosen in a given case is also an important area that needs further investigation.

APPENDIX

1) Table 1.1

2) Computer Program

Table 1.1

Data on Weight and Biparietal Diameter

Serial Number	Biparietal Diameter(cm)	Weight (gm)
1	2.9	113
2	3.6	115
3	3.7	113
4	3.8	439
5	3.8	138
6	4.0	115
7	4.2	198
8	4.3	215
9	4.4	369
10	4.5	300
11	4.5	399
12	4.5	210
13	4.6	454
14	4.6	312
15	4.8	227
16	4.8	737
17	4.8	420
18	4.8	250
19	4.8	312
20	4.9	312
21	5.1	765
22	5.4	482
23	5.4	624
24	5.4	539
25	5.5	567
26	5.5	539
27	5.6	567
28	5.6	539
29	5.9	709
30	6.0	496
31	6.0	907
32	6.2	765
33	6.2	765
34	6.2	851
35	6.3	822
36	6.5	1000
37	6.8	794
38	6.9	992
39	6.9	1162
40	7.0	975
41	7.0	1304
42	7.2	1021
43	7.2	1162
44	7.2	1335
45	7.3	794
46	7.3	1270
47	7.4	1021
48	7.4	1219
49	7.4	1418
50	7.4	1446
51	7.5	1021
52	7.5	1219
53	7.6	970
54	7.6	1418
55	7.6	1559
56	7.7	1276
57	7.7	1559
58	7.7	1956

Table 1.1 (Continued)

Serial Number	Biparietal Diameter(cm)	Weight (gm)
59	7.8	1474
60	7.8	1588
61	7.8	1673
62	7.8	1758
63	7.8	2296
64	7.9	1550
65	7.9	1559
66	7.9	1985
67	7.9	2183
68	7.9	2751
69	8.0	1616
70	8.0	1701
71	8.0	1814
72	8.0	1899
73	8.0	2608
74	8.0	2665
75	8.0	2948
76	8.0	2977
77	8.0	3062
78	8.0	3090
79	8.0	3430
80	8.1	2410
81	8.1	2637
82	8.1	2722
83	8.2	1985
84	8.2	2778
85	8.2	2920
86	8.2	2693
87	8.2	2637
88	8.2	2580
89	8.2	2863
90	8.2	2750
91	8.3	2353
92	8.3	2778
93	8.3	3005
94	8.3	2948
95	8.3	2778
96	8.3	2608
97	8.3	2552

Serial Number	Biparietal Diameter(cm)	Weight (gm)
98	8.3	2466
99	8.3	2070
100	8.3	2183
101	8.3	2693
102	8.3	2268
103	8.3	2183
104	8.3	2410
105	8.3	2552
106	8.3	2381
107	8.4	2523
108	8.4	3459
109	8.4	2608
110	8.4	2353
111	8.4	2807
112	8.4	2807
113	8.4	1928
114	8.4	2268
115	8.5	2835
116	8.5	2637
117	8.5	2637
118	8.5	3281
119	8.5	2296
120	8.5	3204
121	8.5	2410
122	8.5	2353
123	8.5	2268
124	8.6	2722
125	8.6	3005
126	8.6	2296
127	8.6	2183
128	8.6	3005
129	8.6	1616
130	8.6	2410
131	8.6	2466
132	8.6	2268
133	8.6	2552
134	8.6	2665
135	8.6	2410
136	8.6	3062
137	8.7	2466

Table 1.1 (Continued)

Serial Number	Biparietal Diameter(cm)	Weight (gm)
138	8.7	2920
139	8.7	2211
140	8.7	3147
141	8.7	2325
142	8.7	2466
143	8.7	2552
144	8.7	3175
145	8.7	2750
146	8.8	2807
147	8.8	2013
148	8.8	3402
149	8.8	2637
150	8.8	2475
151	8.8	2495
152	8.8	2655
153	8.8	2495
154	8.8	2495
155	8.8	2495
156	8.8	2778
157	8.8	2778
158	8.8	2637
159	8.8	2750
160	8.8	2778
161	8.8	2570
162	8.8	2608
163	8.8	2750
164	8.8	2410
165	8.8	2807
166	8.8	2410
167	8.8	3119
168	8.8	2977
169	8.9	2722
170	8.9	2637
171	8.9	2835
172	8.9	3062
173	8.9	3204
174	8.9	2580
175	8.9	2580
176	8.9	2863
177	8.9	2211
178	8.9	2892
179	8.9	2609
180	8.9	2778

Serial Number	Biparietal Diameter(cm)	Weight (gm)
181	8.9	2778
182	8.9	2835
183	8.9	3289
184	8.9	2608
185	8.9	2892
186	8.9	2495
187	8.9	2523
188	8.9	2353
189	8.9	3090
190	9.0	2807
191	9.0	2211
192	9.0	2835
192	9.0	2722
193	9.1	3260
194	9.1	2693
195	9.1	2608
196	9.1	2155
197	9.2	2722
198	9.2	3147
199	9.2	2693
200	9.2	3345
201	9.2	3062
202	9.2	2580
203	9.3	3416
204	9.3	3204
205	9.3	3204
206	9.3	3459
207	9.3	2183
208	9.3	2863
209	9.3	2807
210	9.4	3289
211	9.4	2381
212	9.4	3119
213	9.4	2665
214	9.4	3175
215	9.5	3062
216	9.5	2665
217	9.5	3317
218	9.5	3515

Table 1.1 (Continued)

Serial Number	Biparietal Diameter(cm)	Weight (gm)
219	9.5	2892
220	9.5	3260
221	9.5	2580
222	9.5	3090
223	9.5	3547
224	9.5	3062
225	9.6	3374
226	9.6	3912
227	9.6	3742
228	9.6	2977
229	9.6	3487
230	9.6	3487
231	9.7	3317
232	9.7	3856
233	9.7	3345
235	9.7	3629
236	9.8	2807
237	9.8	4082
238	9.8	3033
239	9.8	3289
240	9.8	3232
241	9.8	3714
242	9.9	3062
243	10.0	3430
244	10.0	3714
245	10.0	3686
246	10.0	3459
247	10.0	4564
248	10.0	4139
249	10.0	3544
250	10.0	3317
251	10.0	3799
252	10.0	4338
253	10.0	3544
254	10.0	3912
255	10.0	3147
256	10.1	3289
257	10.1	3912
258	10.1	4054
259	10.1	3912
260	10.2	3912
261	10.2	2807
262	10.2	3544
263	10.2	3119
264	10.3	3771
265	10.3	3629
266	10.3	2920
267	10.3	4678
268	10.3	3289
269	10.4	2948
270	10.5	3090
271	10.5	3459
272	10.5	4196
273	10.5	3090
274	10.5	3686
275	10.5	3062
276	10.5	4167
277	10.5	4281
278	10.5	4649
279	10.5	4678
280	10.5	3771
281	10.5	3515
282	10.5	4026
283	10.5	2750
284	10.5	4082
285	10.5	4366
286	10.5	3714
287	10.5	4536
288	10.5	3289
289	10.5	4196
290	10.6	4905
291	10.6	4253
292	10.6	4309
293	10.6	3232
294	10.6	4026
295	10.9	4338

Computer Program of Regression Analysis with B-splines.

```
DIMENSION A(40,8),BS(7),T(7),X(40),W(40),S(7,7),Y(40)
DIMENSION YY(40),XX(400),YG(400),Q(7)
DATA T/0.0,1.0,3.65,5.65,7.65,8.95,10.25/
CALL OPEN(1,'DATT5','INPUT')
CALL OPEN(2,'PATHAK','OUTPUT')
READ(1,*) N
DO 10 I=1,N
READ(1,*) A(I,1),Y(I),W(I)
CALL SPLINE(T,A(I,1),BS,I)
DO 11 J=2,8
11 A(I,J)=BS(J-1)
10 CONTINUE
DO 15 K=2,8
DO 15 J=K,8
SUM=0.0
DO 16 I=1,N
16 SUM=SUM+A(I,K)*A(I,J)*W(I)
15 S(J-1,J-K+1)=SUM
DO 17 J=2,8
SUM=0.0
DO 50 I=1,N
50 SUM=SUM+Y(I)*A(I,J)*W(I)
17 Q(J-1)=SUM
DO 18 I=1,7
18 WRITE(6,19) Q(I),(S(I,I-J+1),J=1,I)
19 FORMAT(8F9.2)
CALL SYMMBE(7,7,S,7,7)
CALL SOLVE(7,7,S,7,7,Q)
WRITE(6,19)(Q(I),I=1,7)
DO 20 I=1,N
CALL SPLINE(T,A(I,1),BS,I)
20 YY(I)=Q(1)*BS(1)+Q(2)*BS(2)+Q(3)*BS(3)+Q(4)*BS(4) -
+Q(5)*BS(5)+Q(6)*BS(6)+Q(7)*BS(7)
WRITE(6,21)
21 FORMAT(//' TABLE OF RESIDUALS'/)
WRITE(6,22)
22 FORMAT(8X,'X',13X,'Y',10X,'YHAT',10X,'RESID'/)
```

Computer Program (Continued)

```
SS=0.0
DO 23 I=1,N
R=Y(I)-YY(I)
SS=SS+R*R*W(I)
23 WRITE(6,24) A(I,1),Y(I),YY(I),R
24 FORMAT(4G15.5)
WRITE(6,25) SS
25 FORMAT(//' SUM OF SQUARED ERRORS',G15.8//)
DEL=10.9/327
DO 26 I=1,208
XX(I)=4.0+(I-1)*DEL
CALL SPLINE(T,XX(I),BS,I)
26 YG(I)=Q(1)*BS(1)+Q(2)*BS(2)+Q(3)*BS(3)+Q(4)*BS(4)+ -
Q(5)*BS(5)+Q(6)*BS(6)+Q(7)*BS(7)
DO 99 I=1,208
99 WRITE(2) XX(I),YG(I)
CALL CLOSE(1)
STOP
END
SUBROUTINE SPLINE(T,Z,BS,NCALL)
DIMENSION T(7),BS(7),TP(7,7)
IF(NCALL.GT.1) GO TO 3
DO 1 I=1,7
DO 2 J=1,I
2 TP(I,J)=T(I)-T(J)
1 CONTINUE
3 A=AMAX1(0., (Z-T(6))**3)
B=AMAX1(0., (Z-T(7))**3)
BS(7)=4*B
A=-4*A/TP(7,6)
B=4*B/TP(7,6)
BS(6)=-(A+B)
A=A/TP(6,5)
B=B/TP(7,5)
C=AMAX1(0., (Z-T(5))**3)
C=4*C/(TP(6,5)*TP(7,5))
BS(5)=0.0
A=A/TP(6,4)
B=B/TP(7,4)
C=C/TP(5,4)
D=AMAX1(0., (Z-T(4))**3)
D=-4*D/(TP(5,4)*TP(6,4)*TP(7,4))
BS(4)=-(A+B+C+D)
E=AMAX1(0.,(T(7)-Z)**3)
E1=4*E/(TP(7,3)*TP(7,4)*TP(7,5)*TP(7,6))
F=AMAX1(0., (T(6)-Z)**3)
F1=-4*F/(TP(6,3)*TP(6,4)*TP(6,5)*TP(7,6))
```

Computer Program (Continued)

```
G=AMAX1(0., (T(5)-Z)**3)
G1=4*G/(TP(5,3)*TP(5,4)*TP(6,5)*TP(7,5))
H=AMAX1(0., (T(4)-Z)**3)
H1=-4*H/(TP(4,3)*TP(5,4)*TP(6,4)*TP(7,4))
P=AMAX1(0., (T(3)-Z)**3)
P1=4*P/(TP(4,3)*TP(5,3)*TP(6,3)*TP(7,3))
BS(3)=E1+F1+G1+H1+P1
F2=4*F/(TP(6,2)*TP(6,3)*TP(6,4)*TP(6,5))
G2=-4*G/(TP(5,2)*TP(5,3)*TP(5,4)*TP(6,5))
H2=4*H/(TP(4,2)*TP(4,3)*TP(5,4)*TP(6,4))
P2=-4*P/(TP(3,2)*TP(4,3)*TP(5,3)*TP(6,3))
R=AMAX1(0., (T(2)-Z)**3)
R2=4*R/(TP(3,2)*TP(4,2)*TP(5,2)*TP(6,2))
BS(2)=F2+G2+H2+P2+R2
G3=4*G/(TP(5,1)*TP(5,2)*TP(5,3)*TP(5,4))
H3=-4*H/(TP(4,1)*TP(4,2)*TP(4,3)*TP(5,4))
P3=4*P/(TP(3,1)*TP(3,2)*TP(4,3)*TP(5,3))
R3=-4*R/(TP(2,1)*TP(3,2)*TP(4,2)*TP(5,2))
U=AMAX1(0., (T(1)-Z)**3)
U3=4*U/(TP(2,1)*TP(3,1)*TP(4,1)*TP(5,1))
BS(1)=G3+H3+P3+R3+U3
RETURN
END
   SUBROUTINE SYMMBE(NDIM,MDIM,A,N,M)
   REAL*4 A(NDIM,MDIM),EM
"
" THIS ROUTINE USES SYMMETRIC ELIMINATION (IN PLACE) TO FACTOR A
" POSITIVE DEFINITE N BY N BAND MATRIX A (BAND WIDTH = 2*M+1) INTO
" L*D*L**T.  ONLY THE LOWER TRIANGLE OF A IS USED AND THIS BECOMES
" L AND D ON OUTPUT.
" A(1,1),....,A(N,1) CONTAINS THE DIAGONAL BAND
" A(2,2),....,A(N,2) CONTAINS THE FIRST SUBDIAGONAL BAND
" A(M+1,M+1),....,A(N,M+1) CONTAINS THE LOWEST SUBDIAGONAL BAND
"
   NM1=N-1
   DO 30 K=1,NM1
   KP1=K+1
   L1=MIN0(K+M,N)
   DO 20 I1=KP1,L1
   I=L1+KP1-I1
   EM=A(I,I-K+1)/A(K,1)
   DO 10 J=KP1,I
10 A(I,I-J+1)=A(I,I-J+1)-EM*A(J,J-K+1)
20 A(I,I-K+1)=EM
30 CONTINUE
   RETURN
   END
```

Computer Program (Continued)

```
      SUBROUTINE SOLVE(NDIM,MDIM,A,N,M,B)
      REAL*4 A(NDIM,MDIM),B(NDIM)
"
" THIS ROUTINE USES BACK SUBSTITUTION TO SOLVE THE TRIANGULAR
" BANDED SYSTEM A*X=B WHERE A IS THE OUTPUT OF SYMMBE, B ON INPUT
" IS THE RIGHT HAND SIDE AND ON OUTPUT IS THE ANSWER X.  AS IN
" SYMMBE THE ORIGINAL COEFFICIENT MATRIX IS ASSUMED TO HAVE BAND
" WIDTH 2*M+1.
"
      DO 10 I=2,N
      L=MAX0(1,I-M)
      IM1=I-1
      DO 5 J=L,IM1
    5 B(I)=B(I)-A(I,I-J+1)*B(J)
   10 CONTINUE
      B(N)=B(N)/A(N,1)
      NM1=N-1
      DO 20 I1=1,NM1
      I=N-I1
      IP1=I+1
      B(I)=B(I)/A(I,1)
      L=MIN0(N,I+M)
      DO 20 J=IP1,L
   20 B(I)=B(I)-A(J,J-I+1)*B(J)
      RETURN
      END
```

SELECTED REFERENCES

1. Aitken, A.C., On least squares and linear combination of observations. Proc. Roy. Soc. Edinb. A, 55 (1934), 42-47.

2. de Boor, C., Calculating with B-splines. J. of Approx. Theory, 6 (1972), 50-62.

3. Campbell, R.C., Statistics for Biologists. Cambridge University Press, 1967.

4. Curry, H.B., and Schoenberg, I.J., On Polya frequency functions. IV. The fundamental spline functions and their limits. J. Analyse Math., 17 (1966), 71-107.

5. Greville, T.N.E., (ed), Theory and Applications of Spline Functions. Academic Press, 1969.

6. Mitra, S.K., and Rao, C.R., Generalized Inverse of Matrices and its Applications. Wiley, 1971.

7. Rao, C.R., Linear Statistical Inference and its Applications. Wiley, 1965.

8. Rao, C.R., Unified theory of linear estimation. Sankhya A, 33 (1971), 371-394.

9. Taylor, E.S., et al, Clinical use of ultrasound in obstetrics and gynecology. Am. J. Obst & Gync., 99 (1967), 671-682.

10. Schoenberg, I.J., On spline functions, Inequalities. (Proceedings of a Symposium held at Wright-Patterson Air Force Base, Ohio, Aug. 19-27, 1965), O. Shisha, ed., Academic Press, 1967, 255-291.

11. Schoenberg, I.J., Spline functions and the problem of graduation. Proc. Nat. Acad. Sci. U.S.A., 52 (1964), 947-950.

12. Whittaker, E.T., On a new method of graduation. Proc. Edinburgh
M Math. Soc., 41 (1923), 63-75.

13. Wold, S., Spline functions in data analysis. Techonometrics, 16 (1974), 1-11.

CURRICULUM VITAE

Name	Dorothy Rybaczyk Pathak
Date of birth	November 26th, 1947
Place of birth	Dlugosiodlo, Poland
Citizenship	U.S.

Educational Background

University	Duration	Degree	Subjects
Univ. of Illinois, Chicago	1966-70	B.S.	Math. & Physics
Univ. of New Mexico	1970-72	M.A.	Math. & Statist.
Univ. of New Mexico	1973-75	Ph.D.	Statistics

Positions Held

University	Duration	Position
Univ. of New Mexico	1970-71	Graduate Teaching Assistant
Univ. of New Mexico, School of Medicine, Dept. of Family and Community Medicine	1971-	Statistical Assistant

Publications

1. (With Kellner, R., M.D., et al) A Pilot Study of the Short-Term Antianxiety Effects of Molindone HCl. Journal of Clinical Pharmacology (1972) 12, No. 11&12, 474-476.

2. (With Kellner, R., M.D., et al) The Short-Term Antianxiety Effects of Propranolol HCl. Journal of Clinical Pharmacology (1974) 12, No. 5&6, 301-304.

3. (With Kellner, R., M.D., et al) Intensive Design Study of Chloridediazepoxide and Placebo in Chronic Schizophrenia Patients Maintained on Phenothianes. (Submitted to Archives of Psychiatry.)