Analysis of Repeated Measures Data

With Illustrations

by

Frances Garven B.Sc. (Hons.)

A dissertation submitted to the
University of Glasgow for the
degree of

Master of Science

Department of Statistics

November 1986

Acknowledgements

I would like to express my thanks to my supervisor, Dr. Ian Ford, for his patience, help and excellent guidance throughout the period of this research. I am also indebt to Julie Omner and Kathy Rosenberg for permission to use their data, to Myra Smith for the typing of this thesis and finally to the members of the Statistics Department of Glasgow University for their support and understanding.

CONTENTS

Summary

Repeated Measures data arises when a response variable is measured on the same experimental units on two or more occasions. Due to the dependence between observations on the same unit, special statistical methods may be required.

This thesis contains a review of experimental designs, models and methods of analysis which may be appropriate when handling data of this type. These methods are illustrated using three practical problems having different types of repeated measures structure.

Chapter one examines the structure of repeated measures data. Balanced and unbalanced designs of varying complexity are described with several examples. Split-plot designs and designs resulting in growth curve data are identified as special cases.

Chapter two contains some models which may be applied in the analysis of repeated measures data. The univariate and multivariate analysis of variance, the general growth curve and the two stage random effects models are outlined for the analysis of balanced repeated measures data. A generalisation of the general growth curve model is outlined for the analysis of unbalanced data and we note that the two stage random effects model may also be used.

Chapter three reviews some of the approaches and methods which may be used in the analysis of balanced and unbalanced repeated measures data, covering those models contained in chapter two. The methodology discussed includes univariate analysis of variance, approximate and conservative univariate tests, multivariate analysis of variance and some special techniques for the analysis of growth curves. These techniques

involve the modelling of the data and the application of either multivariate analysis of variance or covariance.

Chapter four illustrates some of the models and methods discussed in earlier chapters using three practical problems. The first problem entails the analysis of a balanced two factor repeated measures design with two trial factors. The second involves the analysis of a balanced three factor design with two grouping factors and one trial factor. Finally the third problem is an example of an extremely unstructured set of growth curve data where the underlying problem is discrimination.

Chapter 1

The Structure of Repeated Measures Data

1.1: Repeated Meaures Data

Repeated Meaures (R.M.) data arises when a response variable is measured on the same experimental units on two or more occasions.

There may be one or more response variables measured on each occasion. The experimental units may be grouped according to some grouping factor(s). (For example: sex, age category, treatment type).

The occasions on which the response variable(s) are measured for each experimental unit may be classified according to some trial factor(s). Examples of some trial factors are:-

a) Time - Each experimental unit has measurements taken on the same p occasions. (See Growth Curves later).

b) Location - Each experimental unit may have measurements taken at p different locations on the body.

c) Treatment - Each experimental unit has measurements taken at different levels of some form of treatment or combinations of different forms of treatment.

Sometimes the levels of the trial factor can be randomly assigned to experimental units, say over time. Obviously if time itself is a trial factor there can be no random assignment. The randomisation processes used in any experiment are important since they affect the assumptions underlying, and hence analysis of, the resulting data.

1.2: Balanced Repeated Measures Designs

In studies the design can be broken down into two parts. Namely the design on the trial factor(s) and the design on the grouping factor(s).

A simple form of repeated measures design would have the measurements taken on two occasions, corresponding possibly to a pretest and posttest, with one experimental intervention. In more complex designs there may be multiple measures and multiple treatments or treatment combinations. This trial factor design is sometimes called the design on occasions (Bock, 1975 p.448) or the within-subject design (Huynh and Mandeville, 1979).

There may also be a design on the experimental units or subjects. A simple example would be a single random sample of subjects. In more complex cases there may be m groups of subjects. These groups could arise from any factorial design based on treatment factors or factors related to qualitative categorisations of the subjects (for example, age, sex). The design on the experimental units is sometimes known as the design on the sample (Bock, 1975 p.448) or the across-subject design (Huynh and Mandeville, 1979).

Figures 1 to 4 illustrate various possible balanced repeated measures designs involving progressive degrees of complexity.

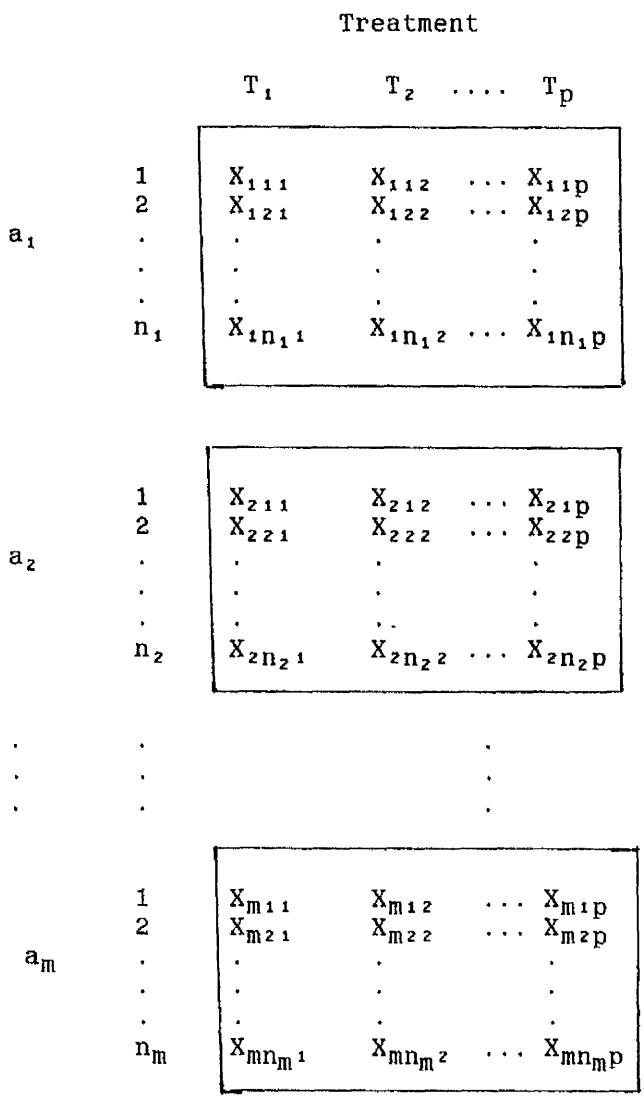
Figure 1 - Single Factor R.M. Design

		Treatment			
		T_1	T_2		T_p
subject	1	X_{11}	X_{12}	...	X_{1p}
	2	X_{21}	X_{22}	...	X_{2p}
	.	.	.		.
	.	.	.		.
	.	.	.	...	.
	N	X_{N1}	X_{N2}	...	X_{Np}

The above figure illustrates a simple repeated measures design with one trial factor and no grouping factor. Data of this type are contained in Winer (1971), Bock (1975) and Goldstein (1979).

For example Winer (1971) examines data from an experiment to study the effects of four drugs upon reaction time to a series of standardised tasks. A random sample of five subjects was taken, where each subject was observed under each of the drugs; the order in which a subject was administered a given drug was randomised.

Figure 2 - Two Factor R.M. Design with R.M. on One Factor

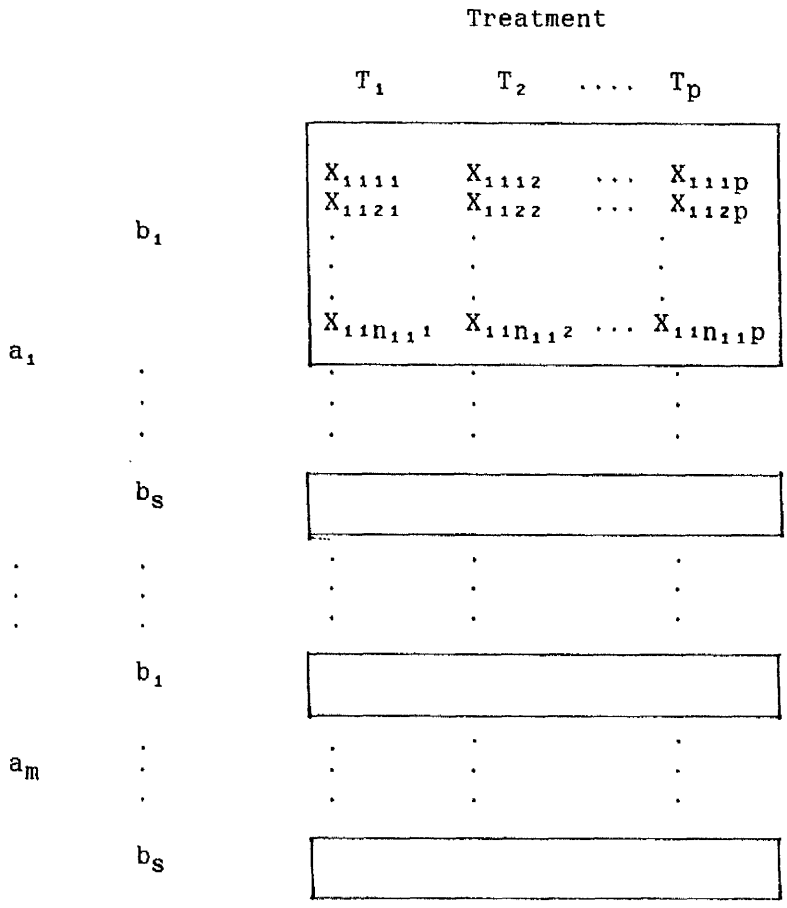


The above figure illustrates a two factor R.M. design with one trial factor and one grouping factor. The grouping factor has m levels with n_k subjects in the k^{th} group (k=1...m). Winer

(1971), Bock (1975) and Goldstein (1979) give examples of data of the above form.

Example: A certain measurement in a dental study was made on each of 11 girls and 16 boys at ages 8, 10, 12 and 14. Obviously sex is our grouping factor with two levels and age is our trial factor with four levels.

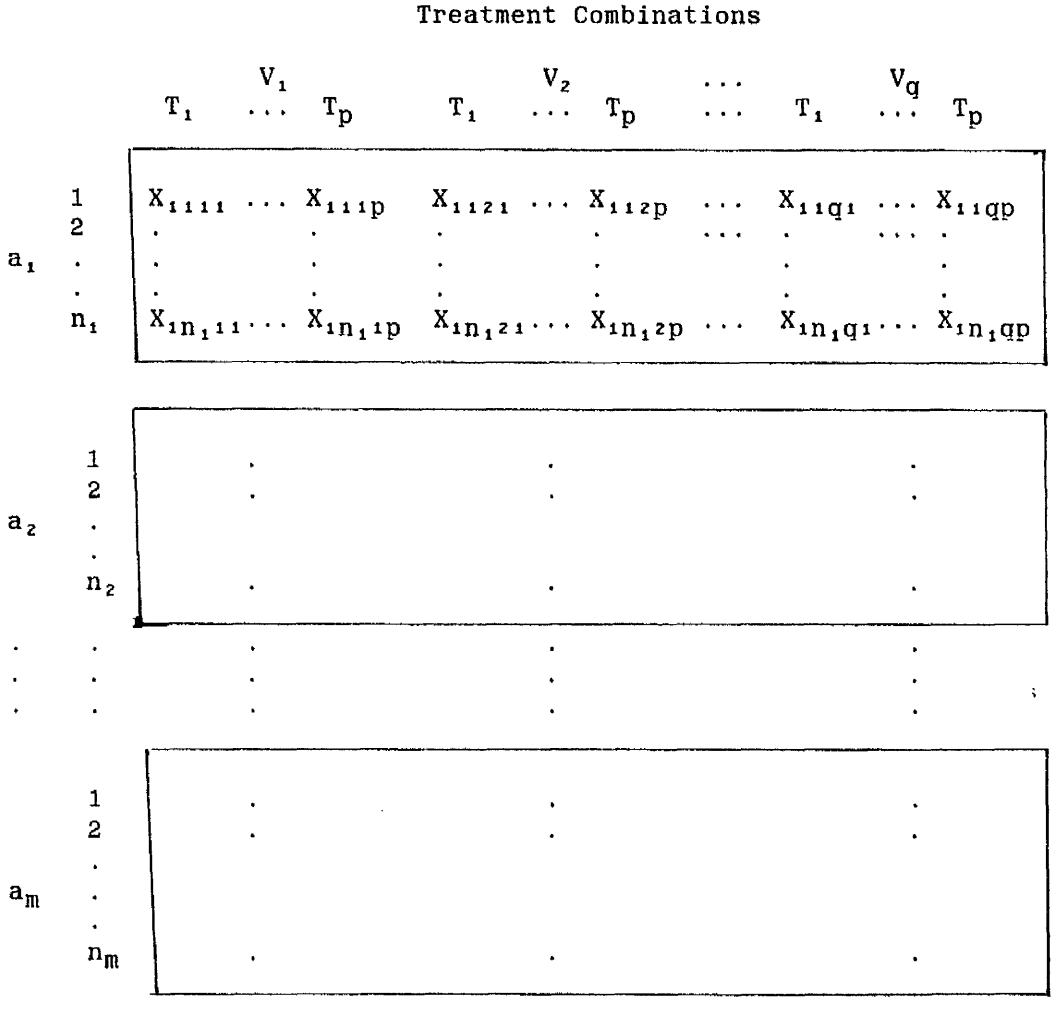
Figure 3 - Three Factor R.M. Design with R.M. on one factor.



The above figure illustrates a three factor R.M. design with two grouping factors (number of levels = m and S) and one trial factor. So the 'across-subject' design is slightly more complicated than that in figure 2. Bock (1975) and Winer (1971 Ch.7) examine data of the above type. Winer considers the following example:

An experimenter is interested in evaluating the effect of anxiety (factor A) and muscular tension (factor B) on a learning task. Subjects are assigned to either level a_1 or level a_2 of the first grouping factor depending on whether they score high or low on a scale measuring manifest anxiety. One half of the subjects in group a_1 are assigned at random to tension condition b_1; the other half are assigned to level b_2. The subjects in group a_2 are divided in a similar manner. Subjects are given four blocks of trials and the number of errors in each is recorded.

Figure 4 - Three factor R.M. Design with R.M. on the last two factors.



The above figure illustrates a three factor R.M. design with two trial factors and one grouping factor. This set up has a

more complicated 'within-subject' design than any of the previous designs. Each subject is measured under one level of the grouping factor and under all pxq combinations of the two trial factors. Designs of the above complexity are discussed by Horton (1978, Ch.5) and Winer (1971, Ch.7) with illustrations.

Example: An experiment was carried out to investigate the effectiveness of four drugs in the treatment of hypertension. Three groups of subjects were taken and classified according to age (young, medium and old). Each subject was administered one of the drugs and their blood pressure was taken on four occasions after taking the drug. After a suitable period of time, to eliminate carry-over effects, one of the other drugs was given and again blood pressure was measured on four occasions. This procedure was carried out until all subjects had been examined under each of the four drugs. Drugs were randomly allocated to subjects.

Useful texts describing repeated measures design structures are Winer (1971) and Bock (1975). Kowalski and Guire (1974) and Goldstein (1979) discuss repeated measures designs as applied to a particular type of data, namely longitudinal data.

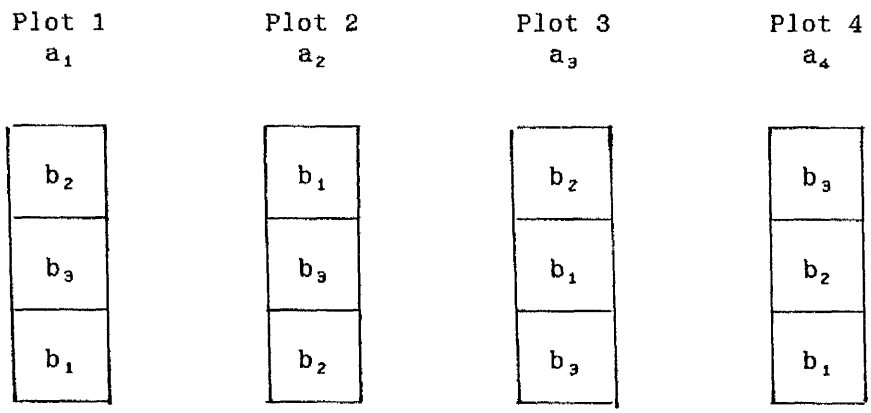
The construction of minimal size designs i.e. those designs which are balanced and require the minimum possible number of experimental units is discussed in Hedayat and Afsarinejad (1975). These authors in a later paper (1978) also examine the optimality of R.M. designs as compared to a large class of competing designs.

Split-Plot Designs

Split-plot designs and certain forms of repeated measures designs have much in common. In particular, they may share the

same analysis under certain circumstances. The term 'split-plot' comes from agricultural experimentation in which a single level of one treatment is applied to a relatively large plot of ground (the whole plot) but all levels of a second treatment are applied to sub-plots within the whole plot. For example, consider the following design (figure 5), in which the levels of factor A are applied to the whole plots and the levels of factor B are applied to the sub-plots (Winer, 1971).

Figure 5 - Split-Plot Design



Note: Levels of factor A are first randomly allocated to the whole plots and then the levels of factor B are randomly allocated to the sub-plots. So the randomisation procedure is a two-stage one, Steel and Torrie (1980).

For the special case in which A and B are fixed factors and the plots are a random sample from a specified population of plots, the analysis would proceed in the same manner as a two factor R.M. design with R.M. on one factor (Winer, 1971). Factor A being a grouping factor and factor B being a trial factor.

Although the analysis proceeds in the same manner for these two designs, there are obvious differences in the experimental set-up of both designs. One of the obvious differences being that experimental units are literally divided into sub-units and

it is these sub-units that are observed under each level of the trial factor for the split-plot design. In R.M. experiments no such sub-division occurs (Monlezun et al, 1984). Another distinction is that in R.M. designs generally (but not always) all except one of the factors are modes of classifying the experimental units rather than treatments which may be randomly allocated to the units by the experimenter (Winer 1971).

Winer (1971), Frane (1980) and Steel and Torrie (1980) discuss split-plot designs and their analysis. A numerical example is given by Steel and Torrie (1980, p.383). Monlezun et al (1984) compare split-plot and R.M. experiments and their analyses. Jennrich (1977) also contrasts R.M. and split-plot designs with respect to randomisation theory.

Growth Curves (Longitudinal Data)

Growth curve or longitudinal data is a special type of R.M. data where time is usually one of the trial factors. Each experimental unit has measurements taken on the same p occasions where p is greater than one. The occasions are defined by a common time scale, for example, historical time or chronological age, Goldstein (1979).

Example 1: The following is an outline of a longitudinal data set discussed by Elston and Grizzle (1962) and Goldstein (1979). The measurements consist of the height of the mandibular ramus bone in a sample of 20 boys taken at four half-yearly intervals from 8.0 years to 9.5 years.

	Age (years)			
Individual	8.0	8.5	9.0	9.5
1	47.8	48.8	49.0	49.7
2	46.4	47.3	47.7	48.4
.				
.				
.				
20	46.3	47.6	51.3	51.8

As in the usual R.M. designs there may be one or more grouping factors.

Example 2: The following is an outline of a longitudinal data set discussed by Potthoff and Roy (1964) and Morrison (1976). The data consists of the distance in millimetres from the centre of the pituitary to the pterygomaxillary fissure on each of 11 girls and 16 boys at ages 8, 10, 12 and 14 years.

		Age (years)			
Sex	Indiv.	8	10	12	14
	1	21	20	21.5	23
	.				
Girls	.				
	.				
	11	24.5	25	28	28
	1	26	25	29	31
	.				
Boys	.				
	.				
	16	22	21.5	23.5	25

Very often longitudinal data may be referred to as R.M. data with no distinction being made. Sometimes the distinction only becomes apparent at the analysis stage. (see later)

Types of longitudinal data and their analyses are examined by Kowalski and Guire (1974), Goldstein (1979) and Woolson and Leeper (1980).

In particular, Kowalski and Guire (1974) discuss six distinct types of longitudinal data.

1.3: Unbalanced R.M. Designs

Quite frequently experiments or studies which involve repeated measurement of the same experimental units result in incomplete data or unstructured data. This is especially true when the repeated measurements are taken over time as in longitudinal studies.

The incompleteness of the data may be due to measurements missing at random or by design, Kleinbaum (1973), Srivastava and McDonald (1974), Machin (1975) and Goldstein (1979).

Some studies might result in one or more of the experimental units having all of their information missing or just some of the observations missing. Reasons for the missing observations could include:

a. Individuals (Exp. Units) leaving before the end of the study e.g. moving house.

b. Individuals not appearing on the set times e.g. inconvenient time.

Woolson, Leeper and Clarke (1978) and Woolson and Leeper (1980) examine the design of longitudinal and mixed longitudinal studies and the analysis of incomplete data arising from these types of study.

Goldstein (1979) discusses the problem of unstructured data where:

a. Individuals come before or after the specified occasions.

b. Measurements are not taken on set occasions but may be taken at quite different times (for example, 'random' time points).

Example of an Incomplete Longitudinal Data Set

Zerbe and Walker (1977) and McCammon (1970) discuss the following example where the Child Research Council (CRC) collected data in a longitudinal study of human growth and development continuously from 1927 to 1969. Fifty-five girl participants were selected for settling a question posed by CRC investigators: whether or not girls heavy at birth differ in average weight from girls light at birth over the age interval 10 to 14 years.

The weights of the sample of 55 girls were recorded during the age interval 9 to 19 years. Examinations were regularly scheduled over this interval, but participants occasionally came early, or late, or missed an appointment altogether. Thus the number of weight measurements per child ranged from 10 to 19.

CHAPTER 2

Models for the Analysis of Repeated Measures Data

Section	Content
2.1	Introduction
2.2	Models for Balanced Repeated Measures Data
2.3	Models for Unbalanced Repeated Measures Data

2.1: Introduction

There are various methods for handling repeated measures data. These could be categorised in terms of;

i) Analysis of Variance, Regression.

ii) Frequentist, Bayesian.

iii) Parametric, Nonparametric.

The approach taken will of course depend on various aspects of the problem in hand. These might include:

i) The structure of the repeated measures data.

ii) The randomisation procedures used in the design.

iii) The assumptions which the investigator is willing to make (especially those concerning variances and independence).

and

iv) Where our interest lies. For example, testing effects or modelling over time.

This chapter outlines some of the models which can be used in the analysis of repeated measures data.

One implicit assumption being made in all of the models outlined here is that of Normality.

Section 2.2 outlines those models which may be used with balanced repeated measures data which includes the analysis of variance model, the general growth curve model and the two stage random effects model.

Section 2.3 outlines those models which may be appropriate when handling unbalanced repeated measures data. This includes Kleinbaum's general growth curve model and again the two stage random effects model.

2.2: Models for Balanced Data

The Multivariate Analysis of Varance Model

The Multivariate Analysis of Variance (Manova) model is the most general one which can be used to explain repeated measures data. It is the most general in that it requires no assumptions about the mean vector or more importantly about the covariance matrix Σ (Rogan et al, 1979, Arnold, 1981, Morrison, 1976).

Unfortunately the Manova model can only be used when sufficient data is available. The amount of data required, depends on the design on the occasions and hence the number of levels of the trial factor(s) (Greenhouse and Geisser, 1959, Horton, 1978).

Quite frequently, however, we do not have enough individuals in the sample to get a good estimator for the covariance matrix and thus the multivariate procedures are not very powerful (Davidson, 1972, Morrison, 1976, Stevens, 1980, Arnold, 1981).

If certain assumptions about the covariance matrix hold then the Manova model reduces to a univariate analysis of variance model and we get more powerful procedures. The necessary and sufficient condition for the validity of the univariate procedures being 'Circularity' (Rogan et al. 1979, Winer, 1971, Arnold, 1981).

The usual Manova model as given by Morrison (1976), Woolson & Leeper, (1980), Arnold (1981) and Seber (1984) is,

$$\underset{(nxp)}{E(X)} = \underset{(nxm)}{A}\ \underset{(mxp)}{\Phi}$$

where the different rows of X are distributed mutually independently and the p elements in any row follow a multivariate normal distribution with unknown covariance matrix Σ (pxp and positive definite).

A is a (nxm) design matrix of known constants.

Φ is a (mxp) matrix of unknown parameters.

The above is a full model since the number of parameters is equal to the number of responses, p.

Arnold (1981) refers to the above as the generalised repeated measures model.

The univariate anova model is obtained by making stronger assumptions about the covariance matrix Σ. Namely,

Figure 2.1

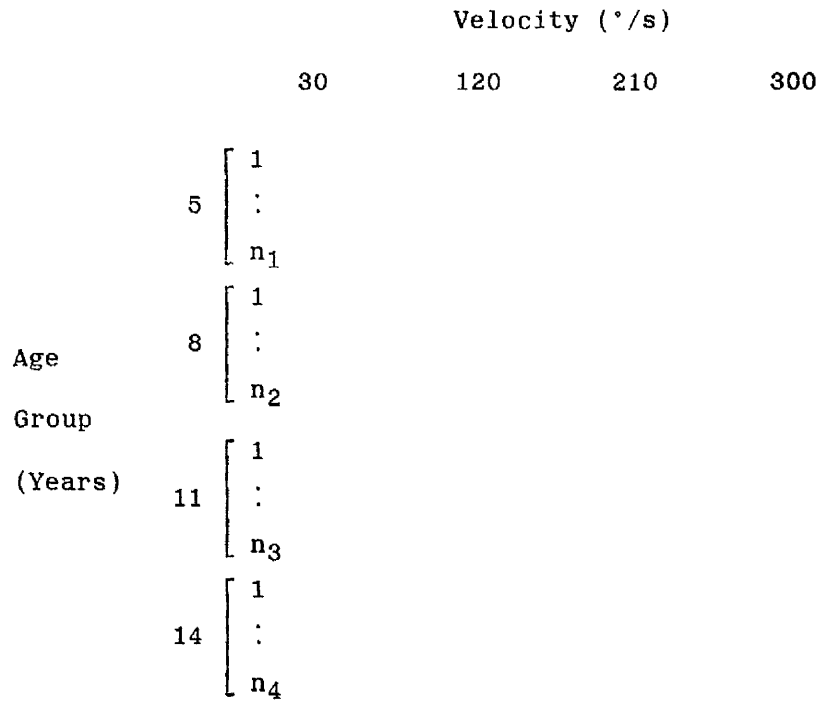
$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \cdot & & & \\ \cdot & & & \\ \rho & \rho & \dots & 1 \end{bmatrix}$$

That is, that all the measurements have the same variance and all the pairs of measurements on the same individual have the same covariance, see Arnold (1981) and Winer (1971). The above form of the covariance matrix satisfies the property of compound symmetry. This is a sufficient but not a necessary condition for the validity of the F-ratios in univariate analysis of variance (Winer 1971).

Example 1: Using the MANOVA Model

A study was carried out to investigate the strength of children in four different age groups. The measurements were obtained using a Cybex Isokinetic Dynamometer at four different velocity settings. The layout of the data can be seen in figure 2.2:

Figure 2.2



Using the terminology introduced in chapter one, the above is a two factor repeated measures design with one grouping factor and one trial factor.

An explicit analysis of variance model for the above, letting X_{ijk} represent the j'th measurement for the i'th individual in the k'th age group would be,

$$X_{ijk} = u + \alpha_k + \Pi_{i(k)} + B_j + \alpha B_{kj} + B\Pi_{ji(k)} + \epsilon_{(ijk)}$$

$$\text{with } i=1 \ldots n_k,$$

$$j=1 \ldots p,$$

$$k=1 \ldots m,$$

with conditions:

$$\sum_k \alpha_k = \sum_j B_j = \sum_k \alpha B_{kj} = \sum_j \alpha B_{kj} = \sum_j B\Pi_{ji(k)} = 0$$

Where

μ is the overall population mean

α_k is the k'th group effect

B_j is the j'th velocity effect

$\pi_{i(k)}$ is a constant associated with the i'th subject in the k'th group

αB_{kj} is the interaction between the k'th group and j'th velocity

$B\pi_{ji(k)}$ is the interaction of velocity j and subject i within treatment group k

and

$\epsilon_{(uk)}$ is the random error component.

In general, when the multivariate model is presented, it is not as fully structured as the model given above. It is sometimes written as follows:

Letting $\underline{X}_{ik}$ be the (px1) vector of observations for the i'th subject in the k'th group.

$$\underline{X}_{ik} = \underline{\mu} + \underline{\tau}_k + \underline{\epsilon}_{ik} \qquad \begin{matrix} i=1 \ldots n_k \\ k=1 \ldots m \end{matrix}$$

where

$\underline{\mu}$ is the overall mean vector, $\underline{\tau}_k$ are the specific (fixed) effects of the groups. The other random effects, unit effects and interactions given in the previous model are submerged in the errors $\underline{\epsilon}_{ik}$ which are assumed to be $N_p(0, \Sigma)$ in each of the groups, Bock (1975) and Davidson (1980).

The General Growth Curve Model

Potthoff and Roy (1964) proposed a generalisation of the usual Manova model for analysing growth curve data. In the General Growth Curve Model of Potthoff and Roy, a second design matrix is included. The addition of this second design matrix

allows the model to be used to describe growth curve situations.

One of the design matrices contains information on the design on the sample and the other, information on the design on the repeated measures.

This model like the Manova model, makes no assumption about the structure of the covariance matrix.

Several authors examine this model and appropriate methods of analysis including Khatri (1966), Rao (1966), Grizzle and Allen (1969), Lee (1974), Tubbs, Lewis and Duran (1975) and Baksalary, Corsten and Kala (1978).

In particular a Bayesian treatment is developed by Geisser (1970, 1971), Lee and Geisser (1972) and Fearn (1975).

The general growth curve model as given by Potthoff and Roy (1964) and Woolson and Leeper (1980) is,

$$\underset{(nxp)}{E(X)} = \underset{(nxm)}{A}\ \underset{(mxq)}{\Phi}\ \underset{(qxp)}{B}$$

where

the rows of X are distributed independently and the p elements in any row follow a multivariate normal distribution with unknown variance matrix Σ_0 (pxp and positive definite).

A is an n x m matrix of known constants

B is a q x p matrix of known constants

Φ is an m x q matrix of unknown parameters.

In this model, q the number of parameters may be less than or equal to p, the number of responses. If p=q then B is a square matrix of full rank.

Example 2: Using the General Growth Curve Model

An illustration is now given to show how the model is constructed for a given set of data where the measurements are taken on some continuum, e.g. time.

We will use the data given in example one where we are interested in the force generated by subjects at four different settings of a Cybex Isokinetic Dynamometer.

The four observations on each subject are not independent, but rather are assumed to be multivariate normal with unknown variance matrix Σ_0 which is the same for all four age groups.

From plots of the average of the logarithm of the force at each velocity, for the four different groups, we can assume that the growth curve is linear, i.e. q=2.

Data:

		Velocity (°/s)			
		30	120	210	300
	5	0.44	0.69	0.87	1.10
Age Group	8	0.72	0.94	1.17	1.43
(years)	11	1.00	1.20	1.42	1.72
	14	1.30	1.50	1.70	1.92

Each point is the average of the logarithm of the force generated at that velocity.

For our example, p=4, m=4, n_1=26, n_2=28, n_3=31 and n_4=29. So for the i'th subject at the j'th velocity in the k'th group we have,

$$E(X_{ijk}) = \alpha_k + \beta_k V_j \quad \text{where} \quad i=1 \ldots n_k, \; j=1 \ldots 4, \; k=1 \ldots 4$$

Using vector notation and letting $\underline{X}_{ik}$ be the p component (p=4) vector of the responses for the i'th subject in the k'th group we have

$$\underline{X}_{ik} = (X_{i1k}, X_{i2k}, X_{i3k}, X_{i4k})$$

and the model above may be rewritten as,

$$E(\underline{X}_{ik}) = [\alpha_k \ \beta_k] \begin{bmatrix} 1 & 1 & 1 & 1 \\ V_1 & V_2 & V_3 & V_4 \end{bmatrix}$$

$$= \Phi_k B$$

If we now let X_k be the (n_k X p) matrix of responses for the n_k subjects in group k we can write the above as.

$$E(X_k) = \begin{bmatrix} 1 \\ 1 \\ . \\ . \\ . \\ . \\ 1 \end{bmatrix} [\alpha_k \ \beta_k] \begin{bmatrix} 1 & 1 & 1 & 1 \\ V_1 & V_2 & V_3 & V_4 \end{bmatrix}$$

$$= 1_{n_k} \Phi_k B$$

where $X_k = [\underline{X}_{1k}, \underline{X}_{2k} \cdots \underline{X}_{n_k k}]^T$

Reformulating this last model into a general framework we have.

$$\underset{(nxp)}{E(X)} = \underset{(nxm)}{A} \ \underset{(mxq)}{\Phi} \ \underset{(qxp)}{B}$$

where

X is the (nxp) matrix of observations and $X = [X_1 \ X_2 \ X_3 \ X_4]^T$

A is the (nxm) matrix describing the design on the sample. It contains n_1 rows consisting of [1, 0, 0, 0], n_2 rows consisting of [0, 1, 0, 0], n_3 rows consisting of [0, 0, 1, 0] and n_4 rows consisting of [0, 0, 0, 1].

Φ is the (mxq) matrix of unknown parameters and

$$\Phi = \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \\ \alpha_3 & \beta_3 \\ \alpha_4 & \beta_4 \end{bmatrix}$$

B is the (qxp) matrix describing the design on the repeated measures. and is given above.

The Conditional Model

Rao (1965, 1966, 1967) gives an alternative reduction of the general growth curve model leading to a conditional model which he examines using an analysis of covariance approach. This alternative format of the general growth curve model and methods of analysis have been studied by Khatri (1966), Grizzle and Allen (1969), Bowden and Steinhorst (1973), Lee (1974), Baksalary, Corsten and Kala (1978) and Seber (1984).

The reduction of the General Growth Curve Model to the Conditional model proceeds as follows:

The first step is to choose a pxq matrix K_1 of rank q such that $BK_1 = I_q$ and a p x (p-q) matrix K_2 of rank p-q such that $BK_2 = 0$.

$K = (K_1, K_2)$ is a pXp nonsingular matrix.

Using the above transformation we can now decompose the general growth curve model into two parts.

Let $Y = XK = (Y_1, Y_2)$ where

$$E(Y_1) = A\Phi BK_1 = A\Phi$$

$$E(Y_2) = A\Phi BK_2 = 0$$

Thus, $E(Y) = [A\Phi: 0]$

with the rows of Y mutually independently normal with covariance matrix,

$$\begin{bmatrix} K_1^T \Sigma K_1 & K_1^T \Sigma K_2 \\ K_2^T \Sigma K_1 & K_2^T \Sigma K_2 \end{bmatrix}$$

Further, Grizzle & Allen (1969) show that the expected value of Y_1 given Y_2 is:

$$E(Y_1|Y_2) = [A : Y_2] \begin{bmatrix} \Phi \\ \Gamma \end{bmatrix}$$

$$= D\Omega$$

where the rows of Y_1 conditionally on Y_2 are mutually independently normally distributed with covariance $(B\Sigma^{-1}B^T)^{-1}$.

Thus the conditional covariance matrix does not depend on the particular choices of K_1 and K_2. As noted by Rao (1966) the matrix K is not unique but the estimates of parametric functions and test criteria will be the same for all choices of K, satisfying the stated conditions.

When the rank of B is q, K may be chosen such that

$$K_1 = G^{-1}B^T(BG^{-1}B^T)^{-1}$$

and K_2 is chosen such that

$$BK_2 = 0$$

The Two Stage Random Effects Model

The explicit feature of this model is its consideration of the individual growth curves as well as the overall one for a given data set, Fearn (1977).

Many authors (Elston and Grizzle (1962), Rao (1965), Fearn (1975, 1977), Darby and Fearn (1979), Laird and Ware (1982), Reinsel (1982, 1984) and Seber (1984)) have examined models of this type where the parameters of the separate growth curve for each individual are assumed to come from a multivariate normal distribution.

In particular, Fearn (1975, 1977) gives a Bayesian analysis of growth curves using the two stage model. Darby and Fearn (1979) utilise this model from a Bayesian viewpoint, in the analysis of a longitudinal study of blood pressure in children.

Rosenberg (1973) considered models of this type in a more general context using maximum likelihood and Bayesian techniques whereas Joreskog (1970, 1973, 1978) investigates a similar model with a more general covariance structure.

Laird and Ware (1982) outline several advantages of two stage models which include the explicit modelling and analysis of

between and within individual variation and discuss the interpretation of the individual parameters.

Unfortunately the special structure assumed for the covariance matrix makes it more limited in its use than for example the general multivariate model where no structure is imposed on the covariance matrix.

The model is presented in two stages. The first stage relates to the individual growth curves. For example, suppose that each of n individuals has a separate growth curve with independent normal observations, that is

$\underline{X}_i \sim N_p(B^T\underline{\gamma}_i, \sigma^2 I_p)$ where $\underline{X}_i$ is the p-dimensional vector of observations on the i'th individual.

Now the second stage assumes that the parameters given above are a random sample from another multivariate normal distribution. For the example given above,

$$\underline{\gamma}_i \sim N_q(\underline{\Phi}, \Lambda)$$

Combining these two stages then gives the marginal distribution of the observations, which is multivariate normal with a special covariance structure, see Laird and Ware (1982) and Seber (1984).

For the above this is,

$$\underline{X}_i \sim N_p(B^T\underline{\Phi}, B^T\Lambda B + \sigma^2 I_p)$$

It is this covariance structure that is ignored by Potthoff and Roy (1984) which the two stage model exploits, Fearn (1975).

Example 3: Using The Two Stage Random Effects Model

Again we will use the data given in example one concerning the force generated by subjects at four different settings of a Cybex Isokinetic Dynamameter. For simplicity we will assume we have one random sample of n individuals.

As shown previously the log of the force is linear with respect to velocity.

Hence our first stage is to suppose that each of the n individuals has a separate growth curve with independent normal observations.

$$X_{ij} = \alpha_i + \beta_i V_j + \epsilon_{ij} \qquad \begin{matrix} i=1 \ldots n \\ j=1 \ldots 4 \end{matrix}$$

where X_{ij} is the force generated at the j'th velocity for the i'th individual and the ϵ_{ij} are independently normally distributed with variance σ^2.

We may write the above in vector form as.

$$\underline{X}_i = B^T \underline{\gamma}_i + \epsilon \qquad \text{for } i=1 \ldots n$$

where

$\underline{X}_i$ is the p-dimensional vector of observations on the i'th individual (p=4).

$\underline{\gamma}_i$ is a q-dimensional vector of unknown parameters. In the above q=2 and $\underline{\gamma}_i = (\alpha_i, \beta_i)^T$.

B^T is a known pxq design matrix of the form.

$$\begin{bmatrix} 1 & V_1 \\ 1 & V_2 \\ 1 & V_3 \\ 1 & V_4 \end{bmatrix}$$

and ϵ is multivariate normal with expectation zero and variance $\sigma^2 I_p$.

We have n separate linear regressions with independent normal errors.

Now we add a second stage by assuming that the regression coefficients are a random sample from another normal distribution with mean $\underline{\Phi}$ and covariance matrix Λ.

That is $\underline{\gamma}_i \sim N_2(\underline{\Phi}, \Lambda)$

Combining the above two stages gives:

$$\underline{X}_i \sim N_p(B^T\underline{\Phi},\ B^T\Lambda B + \sigma^2 I_p)$$

or

$$\underline{X}_i \sim N_p(B^T\underline{\Phi},\ \Sigma_{pp}) \text{ where } \Sigma_{pp} = B^T\Lambda B + \sigma^2 I_p$$

$$\text{for } i=1 \ldots n.$$

2.3: Models for Unbalanced Data

When several responses are obtained from each individual in a study, at different times and possibly under changing experimental conditions it is extremely difficult to control the circumstances under which the measurements are taken and there may be considerable variation among individuals in the number and timing of observations, Goldstein (1979), Woolson et al (1978). The resulting unbalanced data sets are not amenable to analysis using the general multivariate model with unrestricted covariance structure discussed in section 2.2.

Two possible methods of handling this unbalanced data are:

a) To use a model which explicitly takes account of the pattern of the data. Two appropriate models outlined in this chapter are:

   (i) A two stage random effects model

and

   (ii) A generalised multivariate growth curve model.

   Kleinbaum (1973), Woolson et al (1978).

b) To try to estimate the measurements which are missing by using the information from the measurements available, Beale and Little (1975), Dempster et al (1977).

One of the explicit features of the two stage random effects model is its consideration of individual growth curves which means that this model can quite easily be used when handling

unbalanced repeated measures data. This model is of particular use when the repeated measures are obtained at arbitrary or unique times, Laird and Ware (1982). See section 2.2.

When in fact we have observations missing at random, the full multivariate model can be applied by use of multivariate methods for missing observations, Orchard and Woodbury (1972), Laird and Ware (1982).

The Generalised Multivariate Growth Curve Model

Kleinbaum (1973) presented a generalisation of Potthoff and Roy's growth curve model which allows for data which is missing either by chance or by design. Unfortunately this model is limited in its application and may only be used when the data falls into a small number of reasonably large groups.

The model assumes that there are N experimental units and q time points t_1, t_2 ... t_q at which measurements are taken. For simplicity we assume only one variable is measured on each occasion. The N experimental units are divided into u disjoint sets of experimental units S_1, S_2, ..., S_u where S_j has n_j units.

No two different sets S_j and $S_{j'}$ can have measurements at the same q_j time points (although two sets can have the same number of measurements).

X_j and $X_{j'}$ are independent if $j \neq j'$ and the rows of X_j are independent and multinormally distributed for each j.

The model may be written,

$$E(X_j) = A_j \Phi C B_j$$

and

$$V(X_j) = I_{n_j} \otimes B_j \Sigma B_j$$

where

$\otimes$ denotes the right Kronecker product.

X_j is an n_j x q_j matrix of responses.

A_j is a known n_j x m design matrix.

B_j is a known q x q_j incidence matrix of 0's and 1's.

Φ is an m x p matrix of unknown parameters.

C is a p x q known design matrix.

Essentially this is just separating out the sample into sets where the individuals within each set have measurements taken on the same q_j occasions.

Chapter 3

Some Possible Methods of Analysing Repeated Measures Data

Section	Content
3.1	Introduction
3.2	Analysis of Variance Approach
3.3	Modelling Approach
3.4	Other Approaches
3.5	Analysis of Unbalanced Repeated Measures Data

3.1: Introduction

In this chapter we will outline some of the approaches and methods which may be used in the analysis of repeated measures data, covering the models outlined in chapter two.

The following three sections are concerned with the analysis of balanced repeated measures data whereas the final section contains some discussion on how one might handle unbalanced repeated measures data where some of the observations may be missing either by design or at random.

The two main approaches to analysing repeated measures data are the 'Analysis of Variance' approach and the 'Modelling' approach.

The methodology used in the 'Analysis of Variance' approach will be explained first, followed by the 'Modelling' approach.

Finally there will be a brief discussion on Bayesian and Non-parametric approaches.

3.2: The Analysis of Variance Approach

Analysis of Variance is so called because it tests for certain given effects in a set of data by decomposing the total variability in the data into its component parts. This decomposition depends, of course, on the experimental design. In experiments involving repeated measures the total variation is divided into two parts: One part may be called the between-groups component and the other the within-groups component (Winer 1971, Ch.4).

The manner in which the total variation is partitioned in a factorial experiment in which there are no repeated measures is similar to the above. In fact the decomposition into the between-groups component is the same in both cases and the main difference is in the decomposition of the within-groups component (Winer 1971, Ch.7).

A comparison of the two decompositions (one design with repeated measures and one without) will be shown using a specific example of a two factor design. We will consider the case with no repeated measures first.

Example 3.1: Two Factor Design With No Repeated Measures

Random samples of male and female school-children have been drawn from four age groups to investigate alternative methods of measuring strength and power in children. Strength is measured on a Cybex Isokinetic Dynamometer as the Peak Torque (Newton-metres) developed during a maximal concentric contraction. The average peak torque for each sample at a particular velocity setting (300 $deg.s^{-1}$) of the Cybex is shown in table 3.1.

Table 3.1:

	Age (Years)			
Sex:	5	8	11	14
Male	3.1	4.7	10.3	22.4
Female	2.5	5.8	9.6	17.4

The total variability in this design would be partitioned as follows:

$$\text{Total Var.} = \left[\begin{matrix}\text{Between-Group}\\ \text{Var.}\end{matrix}\right] + \left[\begin{matrix}\text{Within-Group}\\ \text{Var.}\end{matrix}\right]$$

$$= \left[\begin{matrix}\text{Var. Due}\\ \text{To Sex}\end{matrix}\right] + \left[\begin{matrix}\text{Var. Due}\\ \text{To Age}\end{matrix}\right] + \left[\begin{matrix}\text{Var. Due To}\\ \text{Sex x Age}\end{matrix}\right] + \left[\begin{matrix}\text{Within-}\\ \text{Group Var.}\end{matrix}\right]$$

Letting X_{ijk} be the peak torque for the i'th individual (i=1...n) of the k'th (k=1...m) sex in the j'th (j=1...ℓ) age group, a summary of the analysis of variance appropriate for this design is given in table 3.2.

Mean squares are obtained from corresponding sums of squares by dividing the latter by their respective degrees of freedom.

Table 3.2:

Source of Var.	Deg. Free	Sum of Squ.
Between Subjects	$m\ell-1$	$\Sigma(\bar{X}_{.jk}-\bar{X}_{...})^2$
- Sex	$(m-1)$	$\Sigma(\bar{X}_{..k}-\bar{X}_{...})^2$
- Age	$(\ell-1)$	$\Sigma(\bar{X}_{.j.}-\bar{X}_{...})^2$
- Sex x Age	$(m-1)(\ell-1)$	$\Sigma(\bar{X}_{.jk}-\bar{X}_{..k}-\bar{X}_{.j.}+\bar{X}_{...})^2$
Within Subjects	$m\ell(n-1)$	$\Sigma(X_{ijk} - \bar{X}_{.jk})^2$
Total	$m\ell n-1$	$\Sigma(X_{ijk} - \bar{X}_{...})^2$

Each effect is tested for significance by comparing its mean square to the mean square for within groups which is also known as the Residual mean square for fixed effects designs.

Under specific assumptions about the underlying sources of variation the ratio of these mean squares follows an F-distribution with the degrees of freedom corresponding to those mean squares which constitute the ratio. These tests will be exact under the following assumptions (Lindman 1974, Ch.6):

i) The observations are obtained under independent conditions.

ii) The data are normally distributed.

and

iii) Each group has the same underlying variance.

Lindman (1974, Ch.6) discusses the robustness of the above tests when some of the assumptions may be violated. It is the assumption of independence which is violated when analysing repeated measures data and although we may still apply analysis of variance, the necessary assumptions and decomposition of the variability in the data will be different.

Univariate Analysis of Variance for Repeated Measures

Repeated measures analysis of variance is essentially a mixed model analysis of variance, that is, containing both fixed factors and random factors. Usually but not always, for repeated measures designs the levels of the grouping factors and the trial factors are assumed to consist of fixed sets and the subjects are assumed to be a random sample from some larger population. The subjects being crossed with the levels of the trial factors.

Let us first consider the simplest case where we have a random sample of n individuals with p correlated measurements on each (See figure 1 of chapter 1).

The analysis is based on a linear model of the following form (Winer, 1971, Horton, 1978):

$$X_{ij} = \mu + \pi_i + \alpha_j + \pi\alpha_{ij} + \epsilon_{ij}$$

where X_{ij} is the measurement on the i'th (i=1...n) subject at the j'th (j=1...p) occasion and

μ is the grand mean

α_j is the main effect for occasion j.

π_i is a constant associated with subject i.

$\pi\alpha_{ij}$ is the interaction between subject i and occasion j.

ϵ_{ij} is the experimental error associated with X_{ij}.

We assume that the α_j correspond to a fixed factor with

$$\sum_j \alpha_j = 0 \quad \text{and} \quad \sum_j \pi\alpha_{ij} = 0$$

and the terms π_i, $\pi\alpha_{ij}$ and ϵ_{ij} are random effects which are independent and normally distributed with expected values,

$$E(\pi_i) = E(\pi\alpha_{ij}) = E(\epsilon_{ij}) = 0$$

The appropriate analysis of variance table would be:

Table 3.3:

Source of Var.	Deg. Free.	Sum of Squ.
Between Subjects	(n-1)	$\Sigma(\bar{X}_{i.}-\bar{X}_{..})^2$
Within Subjects	n(p-1)	$\Sigma(X_{ij}-\bar{X}_{i.})^2$
- occasions	(p-1)	$\Sigma(\bar{X}_{.j}-\bar{X}_{..})^2$
- occas X subj.	(p-1)(n-1)	$\Sigma(X_{ij}-\bar{X}_{i}-\bar{X}_{.j}+\bar{X}_{..})^2$
Total	(np-1)	$\Sigma(X_{ij}-\bar{X}_{..})^2$

In this analysis of variance for a single factor repeated measures design, there is only one error term denoted by the mean square for the (occasions x subjects) interaction. Hence each effect is tested by constructing an F-ratio using this error term as the denominator. Depending on the repeated measures design the analysis of variance table may contain more than one error term.

Though tables 3.2 and 3.3 are not directly comparable, since they correspond to different experimental designs, it can be seen that in the repeated measures design the within subjects variability is broken down into several component parts. This decomposition of the within-subjects component is specific to repeated measures designs.

For these within-subject F-ratios to follow exact F distributions certain covariance assumptions must be met. Confusion exists regarding these validity conditions. It was originally thought that 'Compound Symmetry' (see figure 2.1) was the necessary and sufficient condition but Huynh & Feldt (1970) and Rouanet & Lepine (1970) have shown that compound symmetry is only a sufficient condition. Therefore for the single factor design the (pxp) covariance matrix Σ may have some other pattern and the within-subject F-ratios will still have the necessary F-distribution.

Huynh & Feldt (1970) proved that for the usual F-ratios to be valid in repeated measures designs the covariance matrix had to have a special form called a type H matrix.

i.e. If Σ has a pattern such that the variance of the difference between all possible pairs of treatment means is constant then the F-ratios will be exactly distributed as F-variates.

A matrix with compound symmetry always posseses this property.

Although the F-ratios will be valid for a broader class of matrices than those having compound symmetry, it is difficult to imagine problems which will generate covariance matrices having this exact structure other than those with compound symmetry.

Let us now look at a slightly more complicated situation where we have a two factor repeated measures design with one grouping factor. The grouping factor has m levels and the trial factor has p levels with n_k (k=1,...,m) subjects in each group (see figure 2).

The linear model on which the analysis is based takes the following form (Winer 1971, Ch.7, Horton 1978, Ch.5):

$$X_{ijk} = \mu + \alpha_j + \pi_{i(k)} + \gamma_k + \alpha\gamma_{jk} + \pi\alpha_{ij(k)} + \epsilon_{ij(k)}$$

where the above is just an extension of the single factor repeated measures design discussed previously with the additions:

X_{ijk} = measurement on the i'th subject (i=1..n_k) in the k'th group (k=1...m) at the j'th occasion (j=1...p).

$\pi_{i(k)}$ = constant associated with the i'th subject in the k'th group.

γ_k = main effect of group k.

$\alpha\gamma_{jk}$ = the interaction between occasion j and the k'th group effect.

$\sum_k n_k = N$

We assume that the γ_k correspond to a fixed factor with

$\sum_k \gamma_k = 0$ and $\sum_k \alpha\gamma_{jk} = 0$.

As before the terms $\pi_{i(k)}$, $\pi\alpha_{ij(k)}$ and $\epsilon_{ij(k)}$ are random effects which are independent and normally distributed with expected values

$$E(\pi_{i(k)}) = E(\pi\alpha_{ij(k)}) = E(\epsilon_{ij(k)}) = 0.$$

The corresponding analysis of variance table would be:

Table 3.4:

Source of Var.	Deg. Free.	E (Mean Square)
Between Subjects	(N-1)	
- Groups	(m-1)	$\sigma_\epsilon^2 + p\sigma_\pi^2 + np\sigma_\gamma^2$
- Subj. W. Groups	m(n-1)	$\sigma_\epsilon^2 + p\sigma_\pi^2$
Within Subjects	N(p-1)	
- Occasions	(p-1)	$\sigma_\epsilon^2 + \sigma_{\pi\alpha}^2 + N\sigma_\alpha^2$
- Occ. X. Groups	(p-1)(m-1)	$\sigma_\epsilon^2 + \sigma_{\pi\alpha}^2 + n\sigma_{\alpha\gamma}^2$
- Occ. X. Subj. W. Gps.	m(p-1)(n-1)	$\sigma_\epsilon^2 + \sigma_{\pi\alpha}^2$
Total		

Note: Assuming $n_k = n = N/m$, for $k=1,\ldots,m$.

The above table may be compared with table 3.2 which is for a two factor design with no repeated measures. The tests for effects which can be classified as part of the between subject variation in the repeated measures case will be the same as in the non repeated measures situation as these tests do not involve any of the repeated measures factors.

However, in the repeated measures case only, there is a breakdown of the within subjects variation into several orthogonal parts. The mean square for the occasions x subjects within groups is sometimes called the mean square error - within i.e. MS (error-within) since it forms the denominator of F ratios used in testing effects which can be classified as part of the within subject variation.

In general, the breakdown of the within-subject variation into several orthogonal parts depends upon the repeated measures design.

If there is more than one trial factor then there will be more than one error term and hence more than one denominator for constructing appropriate F-ratios as shown in Winer (1971).

Validity Conditions

Hence for these F ratios to have exact F distributions further assumptions than those given previously for a one factor repeated measures design must hold.

One further assumption required in order that the F-ratios actually follow an F distribution is that the covariance matrices be homogeneous over the levels of the grouping factor (Winer 1971, Ch.7, Horton 1978 Ch.5).

This assumption is required because the covariance matrix for each group must be pooled over the levels of the grouping factor (Huynh & Feldt 1970).

Huynh & Feldt (1970) and Frane (1980) show that, strictly speaking, it is unnecessary to assume exact equality of covariance matrices for the repeated measures across the levels of the grouping factors and that the necessary conditions involve only the covariance matrices of the orthonormal variables for each test being carried out in the repeated measures analysis of variance. Each cluster of within-subject mean square ratios based on the same error term have an associated set of orthonormal variables. For instance they show that the validity conditions required by the F ratios in a two factor (one grouping factor and one trial factor) design are:-

(i) The covariance matrices associated with each level of the grouping factor, $\Sigma_1,\ldots,\Sigma_2,\ldots,\Sigma_m$ satisfy the relationship:

$C^T\Sigma_1 C = C^T\Sigma_2 C = \ldots = C^T\Sigma_m C$

Where C may be taken as any matrix which defines the (p-1) orthonormal contrasts among the p variates. C is a p x (p-1) matrix.

and

(ii) The common matrix in (i) is of the form λI_{p-1} where λ is a scalar and $\lambda > 0$.

Mendoza, Toothaker and Crain (1976) derive necessary and sufficient conditions for the validity of each F ratio in a three factor design with repeated measures on two factors.

The above conditions are referred to as 'Circularity' by Rogan, Kesselman and Mendoza (1979).

Huynh (1978) and Huynh & Mandeville (1979) give the most general version of the validity conditions required by the F ratios regarding the within subjects effects:

(a) The covariance matrices for the associated set of orthonormal variables are identical across all levels of the grouping factors

and

(b) The common covariance matrix outlined in (a) has a sphericity pattern i.e. Equal variances and zero covariances.

Testing the Validity Conditions

For repeated measures designs that contain no grouping factors, only the Mauchly W criterion (1940) is needed to assess the validity of the conditions associated with testing the within-subject effects. Details of this sphericity test are given by Mauchly (1940), Huynh & Feldt (1970), Huynh & Mandeville

(1979), Anderson (1984) and Morrison (1976).

For repeated measures designs that contain one or more grouping factors, the validity of the necessary conditions may be tested in two stages (Huynh & Feldt (1970), Rogan et al (1979), Huynh & Mandeville (1979) and Kesselman et al (1980)).

First Box's (1949) modified criterion M is used to determine whether the covariance matrices of suitably chosen sets of orthonormal variables are equal across the levels of the grouping factors.

i.e. $C^T\Sigma_1 C = C^T\Sigma_2 C = \ldots = C^T\Sigma_m C$

where

Σ_i is the covariance matrix of the orthonormal variables for the i'th group and C is a p x (p-1) matrix which defines the (p-1) orthonormal variables (Rogan et al (1979)).

Details of the Box test for equality of covariance matrices are given by Winer (1971 Ch.7), Morrison (1976) and Huynh & Mandeville (1979). Box's test is just a multivariate generalisation of Bartletts test for homogeneity of variance (1937). Secondly if Box's test indicates that equality of the covariance matrices across the levels of the grouping factors is tenable, then Mauchlys (1940) W criterion is used to test for sphericity in the pooled covariance matrix.

If equality of the covariance matrices across the levels of the grouping factors holds and sphericity is tenable for the pooled covariance matrix then Rogan et al (1979) say that the conditions required for 'circularity' are satisfied. If either of these two elements are rejected then the required condition of 'circularity' is violated.

Huynh & Mandeville (1979) examine the appropriateness of using the Mauchly W criterion when the variates are not normally distributed. The W criterion was shown to provide a conservative testing procedure for light-tailed distributions and to produce more than the nominal percentage of type I errors for heavy-tailed distributions. Morrison (1976) points out that the Box criterion depends strongly on the assumption of multivariate normality and is thus likely to be sensitive to departures from it (Hopkins & Clay (1963), Korin (1972) and Olson (1974)). Davidson (1972) shows that for small samples, one cannot depend on Box's test to detect serious departures from homogeneity.

Kesselman et al (1980) say that there is no point in trying to assess the validity conditions using the above procedures and suggest alternative steps. Their results indicate that even when data is obtained from normally distributed populations, the tests for circularity are sensitive to all but the most minute departures from the null hypothesis and consequently the circularity hypothesis is not likely to be found tenable.

If the necessary assumptions regarding the covariance matrices in a repeated measures data set are violated then the significance levels associated with the F tests will be too 'liberal' (Box 1954). As alternatives to the univariate analysis of variance the researcher has the option of using a modified (approximate or conservative) univariate test which has a more conservative significance level than that in the usual mixed model analysis or if sufficient data is available the option of using multivariate analysis of variance methods.

Modified Univariate Anova Tests

In this section we discuss some modified tests which may be used when the necessary covariance assumptions for univariate repeated measures analysis of variance tests are violated.

Essentially these modified tests adjust the degrees of freedom in the usual F tests to correct for the lack of the specified form of the covariance matrix. The extent of this correction to the degrees of freedom depends on the extent to which the covariance matrix of the orthonormal variables deviates from the necessary pattern.

Let Θ be a parameter that measures the extent to which the sample covariance matrix deviates from the necessary form. Then when $\Theta = 1$, the covariance matrix either has compound symmetry or some other pattern for which the F-ratio has an F-distribution. Hence the upper bound for Θ is 1 (Winer 1971).

Several authors have proposed estimators for Θ, Huynh & Feldt (1970) and Geisser & Greenhouse (1958) being the authors whose estimators are most well known and applied.

Box (1954) gave an approximate test for a one factor repeated measures design where the degrees of freedom of the F ratio are adjusted by a constant ϵ (figure 3.1) which is estimated, $\hat{\epsilon}$, using the covariance matrix of the orthonormal variables obtained from the data.

If we let n denote the size of the sample and p the number of levels of the trial factor then Box (1954) suggests that an approximate test may be made through the use of the usual F ratio, but the degrees of freedom are taken to be,

$(p-1)\epsilon$ and $(p-1)(n-1)\epsilon$ instead of $(p-1)$ and $(p-1)(n-1)$.

When the necessary covariance assumptions for the orthonormal variables are met, ϵ will equal its upper bound of unity. As the covariance matrix departs from the necessary form, the value of ϵ will decrease from one and the F ratio will be distributed with a reduced number of degrees of freedom.

Figure 3.1:

$$\epsilon = \frac{p^2(\bar{\sigma}_{ii} - \bar{\sigma}_{..})^2}{(p-1)(\Sigma\Sigma\sigma_{ij}^2 - 2p\Sigma\bar{\sigma}_{i.}^2 + p^2\bar{\sigma}_{..}^2)}$$

where

$\bar{\sigma}_{ii}$ = mean of entries on main diagonal of Σ.

$\bar{\sigma}_{..}$ = mean of all entries in Σ.

$\bar{\sigma}_{i.}$ = mean of entries in row i of Σ.

σ_{ij} = entry in row i, column j of Σ.

Σ = (pxp) covariance matrix for orthonormal variables obtained from data.

Geisser and Greenhouse (1958) extended the work of Box (1954) to the two factor repeated measures design with one trial factor and one grouping factor. They showed that the F-ratios for the within-subject effects could be tested by using Box's proposed adjustment to the degrees of freedom. They also showed that for designs with only one trial factor the lower bound of ϵ is $1/(p-1)$. Hence $1/(p-1) \leq \epsilon \leq 1$. But the lower bound of ϵ does depend in general on the repeated measures design.

Note: For tests involving only between group effects the F-ratios are exactly distributed as F-variates, as in a non-repeated measures factorial design, provided the necessary assumptions about the variances are met. Therefore no adjustment to the degrees of freedom is needed for these tests.

The preceeding approximate tests do require the computation of ϵ from the elements of the variance-covariance matrix. Usually though, the variances and covariances are unknown and ϵ must be estimated using the sample variances and covariances. However at the time the effect of using a sample estimated ϵ was unknown and Greenhouse & Geisser (1959) suggested the use of the following conservative test.

They suggested setting the value of ϵ at its lower limit, hence using the maximum reduction in degrees of freedom.

For the two factor design with one trial factor we would have $\epsilon = 1/(p-1)$.

Using this lower bound of ϵ has several advantages including ease of computation and independence from the sample variance-covariance matrix. Unfortunately it does result in the maximum possible reduction in the degrees of freedom which may not always be appropriate. Consequently it can result in a loss of power (Rogan et al 1979).

One important application of these conservative tests identified by Greenhouse & Geisser (1959) is for the situation where we cannot assume the equality of the variance-covariance matrices across the levels of the grouping factor.

As a possible means of handling the problems of the sensitivity of the Box and Mauchly tests for the validity conditions, the estimation of ϵ from unknown population variance-covariance matrices and the over conservativeness of the 'conservative' tests, Greenhouse & Geisser (1959) suggest the following three step approach to testing significance of the F-ratios in the univariate analysis of repeated measures:

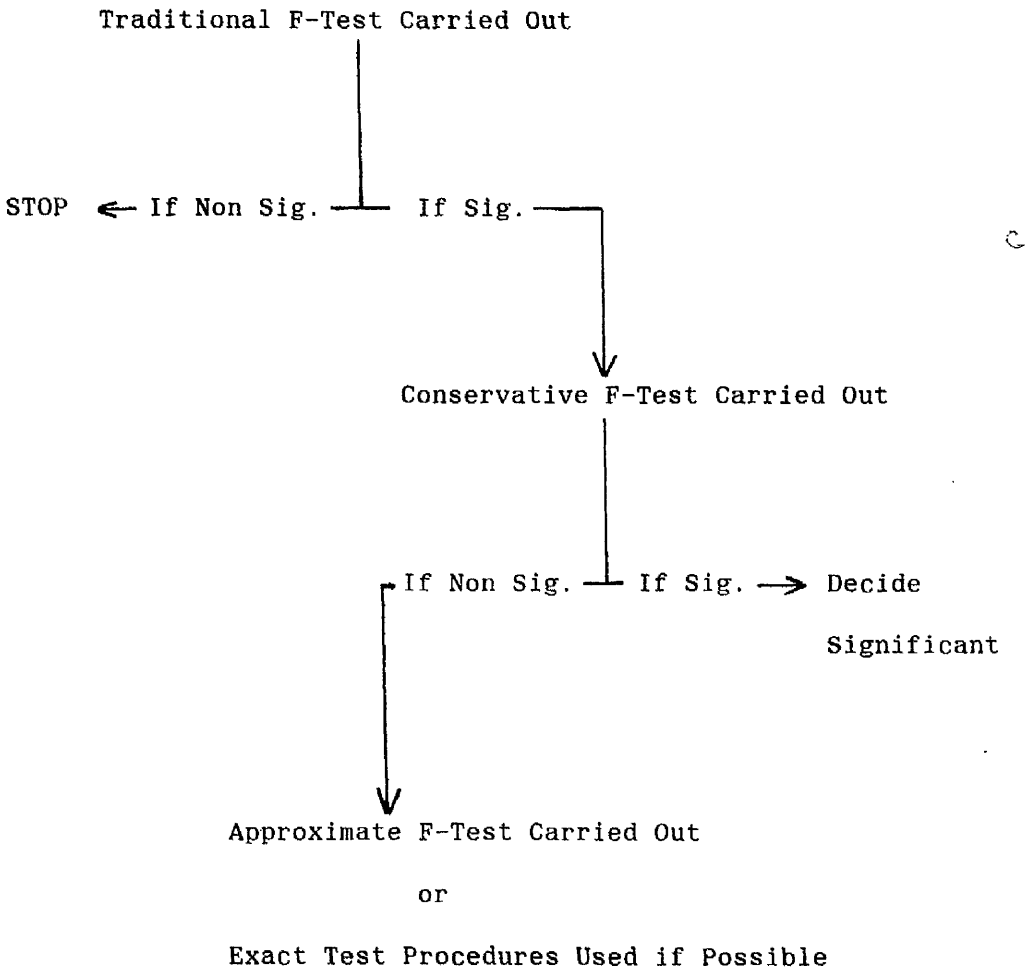
After constructing the traditional analysis of variance F-ratio we test it using the full i.e. unreduced degrees of

freedom. If the F-ratio is smaller than the critical value, one can stop here, for the null hypothesis will not be rejected by reducing the degrees of freedom.

If the observed F-ratio is significant, then one proceeds to the conservative test where the degrees of freedom are reduced by the maximum amount i.e. by the lower bound of ϵ.

If this second test leads to significance, one can at this point reject the null hypothesis without further testing. However, if the conservative test is not significant then it is suggested that ϵ be estimated from the sample variance-covariance matrix and the approximate test be carried out (see Figure 3.2 for summary). If enough data are available exact procedures may be carried out using multivariate methods.

Figure 3.2



The effect of using a sample-estimated value of ϵ, $\hat{\epsilon}$ on the approximate F distribution has been investigated by Collier et al (1967). As noted by Huynh & Feldt (1976) Collier et al's data suggests that when the size of the sample is less than twice the number of levels of the trial factor, $\hat{\epsilon}$ may be seriously biased if ϵ is near or a little above .75. The estimate then tends to over correct the degrees of freedom and produces a more stringent significance level than the nominal level being used.

Thus Huynh & Feldt (1976) suggested adjusting the degrees of freedom of the approximate F-tests by an alternative estimator ω. They show using Monte Carlo methods that their estimator ω is less biased and less dependent on large sample size when the variance-covariance matrix deviates only moderately from the necessary form. They consider the single factor repeated measures design first and then generalise to the two factor design with one trial factor. Formulas for ω in both situations are given in figure 3.3. It is possible for ω to exceed 1.0 but if it does it is equated to 1.0. Some of the other results given by Huynh & Feldt (1976) are that ω is always as large as ϵ ($\omega \geq \epsilon$ for any sample size and number of levels in the trial factor), $\omega = \epsilon$ when $\epsilon = 1/(p-1)$, the difference between ω and ϵ decreases with increasing sample size and the use of ω results in a less conservative test than the use of ϵ.

Figure 3.3:

For a one factor design, $\omega = \dfrac{N(p-1)\epsilon-2}{(p-1)(N-1-(p-1)\epsilon)}$

For a two factor design, $\omega = \dfrac{N(p-1)\epsilon-2}{(p-1)(N-M-(p-1)\epsilon)}$

Huynh & Feldt (1976) also examine the power gained by the test from the use of ω as opposed to ϵ. Frane (1980) examines the sensitivity of ω and ϵ.

Huynh (1978) extended the work by Huynh & Feldt (1976) on the ω approximate test to cover the case where we cannot assume equality of the covariance matrices across all levels of the grouping factor. For the two factor design they propose two additional approximate tests for assessing the significance of within-subject effects. Both tests make use of a theorem by Box (1954). He first proposes the General Approximate (GA) test for arbitrary covariance matrices and secondly the Improved General Approximate (IGA) test which is more sensitive than the GA test for situations which nearly display the required sphericity.

Multivariate Analysis of Repeated Measures

The previous procedures present approximate and conservative tests of significance for repeated measures data, when the necessary assumptions required for univariate analysis of variance are suspect. There are exact procedures available through the use of multivariate analysis of variance. Unfortunately these exact procedures do require a certain amount of data and as Greenhouse & Geisser (1959) noted a lot more computation, but with the facilities available today the computation can be coped with fairly easily.

Cole & Grizzle (1966) assert that repeated measures data is essentially multivariate and should be analysed as such. Bock (1963) exploits this fact when discussing the analysis of repeated measures data and points out that this is a special class of multivariate data in which the observations on each occasion are assumed to be measurements on the same scale with the same origin and unit, that is commensurable.

Both the univariate and the multivariate procedures rest on the assumption that the population random error components are normally distributed. However unlike the univariate procedures which stipulate a particular form for the population covariance matrix Σ, the multivariate approach makes no such specification as to the form of Σ (Rogan et al 1979).

One disadvantage of the multivariate approach is its potential lack of sensitivity when compared to the univariate approach. If the necessary assumptions required by the univariate analysis are tenable then the univariate tests will be more powerful than the tests in the multivariate analysis of variance. Comparison of the power of the two approaches is discussed by Mendoza et al (1974), Davidson (1972), Rogan et al

(1979) and Stevens (1980).

As said previously one other disadvantage which makes the multivariate analysis less applicable, is the need for a reasonable amount of data. For example, in a one-way multivariate analysis of variance context, the estimated variance-covariance matrix will be singular unless the total sample size minus the number of groups is greater than the number of levels in the trial factor. Hence if this requirement is not met, the multivariate analysis cannot be used.

The exact procedures in multivariate analysais of variance for testing hypotheses depend on the repeated measures design. For example, in a two factor design with one trial factor and one grouping factor we might be interested in testing for,

(i) the existence of a group-response interaction.

(ii) the existence of a group effect.

(iii) the existence of a response effect.

The multivariate approach to the analysis of repeated measures when the measurements are taken on some continuum is often referred to as 'Profile Analysis' Mager (1973), Kowalski & Guire (1974), Morrison (1976). This is because the data may be presented graphically by plotting the average response of each group at each level of the trial factor on a graph. The adjacent means for each group can then be joined to form profiles for each group. The above hypotheses are concerned with these profiles. The methodology for the analysis of repeated measures using multivariate analysis of variance will now be outlined for a two factor design with one grouping factor which has m levels. For further details see Greenhouse & Geisser (1959), Morrison (1976) and Seber (1984).

Multivariate Analysis for a Two Factor Design

Suppose that we have m independent random samples of individuals or some other experimental unit with p correlated measurements on each subject. There are n_k (k=1,...,m) subjects in the k'th group.

For example, consider the following data layout;

	Velocity Setting (Units)			
	300	210	120	30
Age Group (years) 5				
8				
11				
14				

Letting X_{ijk} represent the j'th response on the i'th individual in the k'th group for $i=1...n_k$

$j=1...p$

$k=1...m$

where $\sum_{k=1}^{m} n_k = N$

and letting $\underline{X}_{ik}$ represent the (px1) vector of observations for the i'th subject ($i=1...n_k$) from the k'th group (k=1...m) we can write,

$$\underline{X}_{ik} = \underline{\mu}_k + \underline{e}_{ik}$$

where the $\underline{e}_{ik}$ are independently $N_p(\underline{0}, \Sigma)$.

Then we have,

$$\begin{bmatrix} \underline{X}_{11}^T \\ \underline{X}_{21}^T \\ \vdots \\ \underline{X}_{n_1 1}^T \\ \vdots \\ \underline{X}_{m1}^T \\ \underline{X}_{m2}^T \\ \vdots \\ \underline{X}_{m n_m}^T \end{bmatrix} = \begin{bmatrix} \underline{1}_{n_1} & 0 & \dots & 0 \\ 0 & \underline{1}_{n_2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \underline{1}_{n_m} \end{bmatrix} \begin{bmatrix} \underline{\mu}_1^T \\ \underline{\mu}_2^T \\ \vdots \\ \underline{\mu}^T{}_m \end{bmatrix} + \begin{bmatrix} \underline{e}_{11}^T \\ \underline{e}_{21}^T \\ \vdots \\ \underline{e}_{n_1 1}^T \\ \vdots \\ \underline{e}_{m1}^T \\ \underline{e}_{m2}^T \\ \vdots \\ \underline{e}_{m n_m}^T \end{bmatrix}$$

$$X = A\,\Phi + U$$

$$E(X) = A\,\Phi$$

where A is the appropriate design matrix

and Φ is the matrix of unknown parameters. (See section 2.2).

Note: $\underline{1}_{n_K}^T = [1 \dots 1]$ contains n_K ones for K=1 ... m.

Now for the two factor design under this model, the hypothesis of parallelism may be written.

$$H_{01}: M^T\underline{\mu}_1 = M^T\underline{\mu}_2 = \dots = M^T\underline{\mu}_m$$

or in matrix form

$$H_{01}: \quad J\,\Phi\,M = 0$$

where

J is a (m-1) x m matrix and M is a px(p-1) matrix given below.

$$J = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 \\ 0 & 1 & -1 & \dots & 0 \\ . & . & . & \dots & . \\ 0 & 0 & 0 & \dots & -1 \end{bmatrix} \qquad M = \begin{bmatrix} 1 & 0 & \dots & 0 \\ -1 & 1 & \dots & 0 \\ . & . & \dots & . \\ 0 & . & \dots & -1 \end{bmatrix}$$

Now assuming that the above hypothesis is not rejected and parallelism is tenable then the hypothesis of no group effect can be expressed as:

$$H_{02}: \quad J\Phi \underline{V} = 0$$

where J is as given above and

$\underline{V}^T$ = [1....1] contains p ones.

Finally the test for the hypothesis of equal response effects assuming that parallelism is tenable may be expressed as:

$$H_{03}: \quad \underline{b}^T\Phi M = 0$$

where $\underline{b}^T$ = [1,...,1] contains m ones and M is the px(p-1) matrix given above.

Note: In H_{01} the matrix J allows the generation of hypotheses on the between-group effects whereas the matrix M allows the generation of hypotheses on the within-subject effects.

We will first discuss tests of the hypthesis H_{01} and then H_{02} and H_{03}.

The test of H_{01} amounts to a one-way multivariate analysis of variance on the p-1 differences of the observations of the adjacent responses from each subject. Several procedures have been developed for testing H_{01} under the multivariate analysis of variance model. Unfortunately these different procedures result in different forms for the test statistic. We will sketch the derivation of the union-intersection test of Roy.

The multivariate hypothesis H_{01} is true if and only if the univariate hypotheses

$$H_0: \quad J \Phi M \underline{a} = 0$$

for all non-null (p-1) component vectors $\underline{a}$.

The test statistic for any one of these univariate hypotheses is given by

$$F(\underline{a}) = \frac{(N-m)\underline{a}^T M^T X^T A(A^TA)^{-1}J^T[J(A^TA)^{-1}J^T]^{-1}J(A^TA)^{-1}A^TXM\underline{a}}{(m-1)\underline{a}^T M^T X^T[I-A(A^TA)^{-1}A^T]XM\underline{a}}$$

For a univariate test at significance level β we accept

H_0: $J\Phi M\underline{a} = 0$

if $F(a) \leq F(\beta;\ m-1, N-m)$

and accept the original multivariate hypothesis

H_{01}: $J\Phi M = 0$ at some other level α if

$$\bigcap_{\underline{a}} [F(a) \leq F(\beta;\ m-1, N-m)]$$

for all non-null $\underline{a}$.

This acceptance region is equivalent to that defined by

$$\max_{a} F(a) \leq F(\beta;\ m-1, N-m)$$

for if the greatest F-ratio falls in the acceptance region, so must those of all other compounding vectors.

This maximum value of F(a) can be shown to be proportional to the greatest root of the determinantal equation

$$|H - \lambda E| = 0$$

where $H = M^TX^TA(A^TA)^{-1}J^T[J(A^TA)^{-1}J^T]^{-1}J(A^TA)^{-1}A^TXM$

$E = M^TX^T[I-A(A^TA)^{-1}A^T]XM$

Let C_s = greatest root of $|H-\lambda E|=0$ where $s = \min(m-1, p-1)$ or the smaller of the parameters $(m-1)$ and $(p-1)$.

To use available tables we must use the test statistic.

$$\Theta_s = \frac{C_s}{1 + C_s}$$

where the parameters for the distribution of Θ_s under the null hypothesis are

$$s = \min\ (m-1,\ p-1)$$

$$f = \frac{|m-p|-1}{2}$$

$$g = \frac{N-m-p}{2}$$

and for a significance level, α, the acceptance region is

$$\theta_s \leq R(\alpha;\ s,f,g)$$

where

$R(\alpha;\ s,f,g)$ is obtained from appropriate tables. For details see Morrison (1976, Ch.5).

Note: The nonzero roots of $|H-\lambda E|=0$ are equal to the nonzero characteristic roots of HE^{-1}, and in practice it is usually more efficient to extract the roots from this matrix.

Two common alternative procedures to the Union-intersion approach (Roy (1953) are the Wilks Λ Criterion and the Lawley-Hotelling Trace statistic which will be outlined. All three of the test procedures mentioned so far use as their test criteria some function of the roots of the determinantal equation $|H-\lambda E|=0$.

The Wilks Λ Criterion

Wilks (1932) developed test criteria through the generalised likelihood ratio principle. This approach led to the test statistic

$$\Lambda = \frac{|E|}{|H + E|} = \frac{1}{|HE^{-1} + I|}$$

Λ is the reciprocal of the product of all the characteristic roots of $HE^{-1}+I$. When the null hypothesis is true, the large-sample distribution theory of likelihood statistics implies that

$$\chi^2 = -[N - m - \tfrac{1}{2}(p - m + 1)] \ln \Lambda$$

is distributed as a chi-squared variate with $(p-1)(m-1)$ degrees of freedom as N tends to infinity.

The Lawley-Hotelling Trace Statistic

Lawley (1938) and Hotelling (1947, 1951) proposed the sum of the roots of HE^{-1} as a test criterion.

The exact distribution of

$$T_0^2 = \text{trace } (HE^{-1})$$

is complicated, but when the null hypothesis is true, NT_0^2 tends to a chi-squared variate with $(m-1)(p-1)$ degrees of freedom as the number of independent sampling units N becomes large.

Other test criteria have been proposed by Pillai (1955) and others. For further details on the tests mentioned see Mardia, Kent and Bibby (1979 Ch.12), Morrison (1976, Ch.5), Seber (1984, Ch.8) and Anderson (1984, Ch.8).

We note here that when we have only two groups (m=2) then all four of the test statistics mentioned for testing H_{01} are equivalent to Hotellings T_0^2 test for comparing two means. For further details on the two sample case see Seber (1984) and Morrison (1976).

Power comparisons of the different test criteria and discussion of the use of the tests may be found in Mardia, Kent & Bibby (1979, Ch.5), Morrison (1976, Ch.5), Seber (1984, Ch.9) and Anderson (1984, Ch.8).

The hypothesis test, H_{02} of no group effect or identical profile heights, may be carried out by a one-way univariate analysis of variance on the sums of the responses for each subject across the m groups. Although the matrices H and E could be computed for $\underline{V}$, Morrison (1976) notes that it may be more efficient to transform the data.

$$R = X\underline{V}$$

and carry out the test as an analysis of variance on the response totals.

Finally the test for the hypothesis of equal response effects may be carried out using,

$$T_o^2 = N\,\bar{\underline{X}}^T\, M(M^T S M)^{-1}\, M^T\, \bar{\underline{X}}$$

where

$\bar{\underline{X}}$ is the grand mean vector

and

S is the usual pooled estimator of the covariance matrix Σ.

$$S = \frac{E}{N-m}$$

When H_{03} is true,

$$F = \frac{N-m-p+2}{(N-m)(p-1)}\, T_o^2 \sim F(p-1,\ N-m-p+2)$$

If the hypothesis of parallelism is rejected, it will be necessary to test the equality of the group effects separately for each response by p univariate analyses of variance.

3.3: The Modelling Approach

Apart from using the better known methods for analysing repeated measures data which fall under the heading of 'Analysis of Variance', there are alternatives which make use of patterns or relationships that exist within the data.

For these alternative methods to be appropriate the measurements have to be taken over some continuous scale such as time, dose or age. Since the development of these methods has been motivated by modelling time series of the sizes or weights of an organism they are said to fall under the heading of 'Analysis of Growth Curves' (Morrison 1976). In this thesis we refer to this as the 'Modelling' approach.

Morrison (1976) noted that a natural sequel to hypothesis testing, in the Analysis of Variance approach, might be the fitting of some simple polynomial function to sample means. The modelling of repeated measures data using orthogonal polynomials originated in the work of Wishart (1938) who as Rao (1972) noted was the forerunner in this approach. Wishart (1938) used orthogonal polynomials to transform raw data into unweighted least square estimates of parameters of a linear model. Having reduced the dimension of the data from seventeen to a few parameters, he then analysed these parameters separately using univariate analysis of variance, looking for mean differences between groups. Thus, one of the advantages of using this modelling approach is that a large number of repeated measures may be reduced to a few fitted coefficients of a polynomial model. However, Goldstein (1979) noted that Wishart's (1938) methodology was lacking in several aspects including a procedure for testing whether a low-order polynomial fitted the data well, or whether higher order terms were needed to provide a completely

adequate description. Kowalski & Guire (1974) and Woolson & Leeper (1980) have reviewed extensions and developments to the Wishart approach. Some of these developments being as follows:

(i) Box (1950) examined the general growth curve problem for data vectors arising from a multivariate normal distribution with a uniform covariance structure and provided appropriate significance tests,

(ii) Rao (1958) gave a method of estimating a transformation for the time scale, so that growth with respect to the new time scale was linear. He then showed that a Wishart-type analysis remains valid when polynomials in the estimated time metameter are fitted and that comparison between treatment groups could be reduced to examining differences in linear growth rates with respect to this transformed time. Goldstein (1979) identifies several difficulties with this approach including the existence of a common transformation and the interpretation of differences using this form of transformation.

(iii) Rao (1965, 1966) developed multivariate methods for analysing growth curve data using analysis of covariance. These will be considered in detail later.

(iv) Elston & Grizzle (1962) compared three methods of analysis under different assumptions (including a univariate analysis of the problems considered by Rao (1959)).

(v) Hoel (1964) studied the effect of ignoring the dependence of observations taken at different time points on the validity of the statistical inference.

(vi) Further developments have also been made by Potthoff & Roy (1964) and Grizzle & Allen (1969).

We now examine some of the methodology for the General Growth Curve model developed by Potthoff & Roy (1964) including hypothesis testing, estimation and construction of confidence bounds. These methods and some alternatives are also discussed in Grizzle & Allen (1969) and Roy, Gnanadesikan & Srivastava (1971).

These methods are derived from the theory of multivariate normal analysis of variance and are based on polynomial models for the growth curves.

We will first examine the problem of testing hypotheses, of the following form, under the growth curve model (see section 2.2):

H_0: $J\Phi V = 0$ against H_1: $J\Phi V \neq 0$

where

J is a known (sxm) matrix of rank s and

V is a known (qxu) matrix of rank u.

Potthoff & Roy's (1964) solution is to reduce the general model to the usual multivariate analysis of variance model and hence make use of the standard results already outlined for that model (section 3.2).

Assuming $q \leq p$ and that B is of full rank q, Potthoff & Roy use the following transformation which introduces an arbitrary non-singular matrix G.

Transform X to X_0 where $X_0 = XK_1$ and

$$K_1 = G^{-1}B^T (BG^{-1}B^T)^{-1}$$

This gives,

$$X_0 = XG^{-1}B^T(BG^{-1}B^T)^{-1}$$

It then follows that X_0 conforms to the usual MANOVA model,

$$E(X_0) = A\Phi$$

The matrix X_0 (nxq) will be such that the different rows of X_0 will be distributed mutually independently and the q elements in any row will follow a multivariate normal distribution with (unknown) positive definite covariance matrix,

$$\Sigma_0 = [B(G^T)^{-1}B^T]^{-1}B(G^T)^{-1}\ \Sigma G^{-1}B^T(BG^{-1}B^T)^{-1}$$

As long as G is non-singular and as long as $BG^{-1}B^T$ is non-singular, the above procedure is valid for any choice of the matrix G (Potthoff & Roy 1964).

From the theory of multivariate analysis of variance outlined in section 3.2 the appropriate sums of squares and cross products matrices are,

$$H_1 = (J\hat{\Phi}_1V)^T[J(A^TA)^{-1}J^T]^{-1}(J\hat{\Phi}_1V)$$

and

$$E_1 = V^TX_0^T[I_n - A(A^TA)^{-1}A^T]X_0V$$

where

$$\hat{\Phi}_1 = (A^TA)^{-1}A^TX_0$$
$$= (A^TA)^{-1}A^T\ X\ G^{-1}B^T(BG^{-1}B^T)^{-1}$$

From the general theory, H_1 and E_1 are independently distributed as $W_u(S,\ V^T\Sigma_oV)$ and $W_u(n-m,\ V^T\Sigma_oV)$ respectively when H_0 is true. To test the hypothesis H_0 we may use one of the test statistics and procedures described in section 3.2.

Some methodological problems with the above analysis were discussed by Rao (1966). The main problem being the choice of the arbitrary matrix G. If q=p, then B is non-singular and we simply make the transformation

$$X_0 = XB^{-1}$$

but for q<p there are several choices. The simplest choice is $G = I_p$ as noted by Seber (1984) so that,

$$X_0 = XB^T(BB^T)^{-1}$$

This is equivalent to using the estimated regression coefficients of the associated polynomials instead of the original data. The calculations will be simplified if normalised orthogonal polynomials are used in B. We then have

$$BB^T = I_q \text{ and } X_0 = XB.$$

This is essentially the method adopted in the earlier papers of Wishart (1938) and Leech & Healy (1959).

G may also be chosen on the basis of prior information or estimated from previous data.

Using a minimum variance criterion, Potthoff & Roy (1964) showed that the optimal choice of G is $G=\Sigma$, so that the variances increase as G moves away from Σ. However Σ is unknown and most estimates of Σ are statistically dependent on X. A natural choice is G=S where

$$S = \frac{X^T R X}{n - m}$$

$$= \frac{X^T[I_n - A(A^TA)^{-1} A^T]X}{n - m}$$

However S depends on X so that the theory of multivariate analysis of variance is no longer applicable. Rao (1966) criticises Potthoff & Roy's procedures on two points. One is the arbitrariness of the matrix G and the other is the loss of information as a result of reducing the matrix of observations X of order nxp to X_0 of order nxq where q<p. However, if the covariance matrix Σ is known and G is chosen to be this known matrix then there will be no loss of information.

Rao (1966) showed further that the additional information that is not used in Potthoff & Roy's approach could be used by incorporating it into the model in the form of concomitant information.

Hence Rao (1965-67) and Khatri (1966) independently proposed an alternative reduction of the general growth curve model which leads to a conditional model. Tests of hypotheses, estimators and confidence bounds may then be obtained for this conditional model using analysis of covariance. Grizzle & Allen (1969) develop further some of the procedures suggested by Rao (1966). For the alternative reduction of the general growth curve model see chapter 2. Grizzle & Allen (1969) define a general growth curve model as

$$E(Y_1|Y_2) = [A: Y_2] \begin{bmatrix} \Phi \\ \Gamma \end{bmatrix}$$

$$= D\Omega$$

where

the rows of Y_1 conditionally on Y_2 are mutually independently normally distributed with covariance matrix $(B\Sigma^{-1}B^T)^{-1}$.

Potthoff & Roy's methodology just involves the use of the marginal distribution of Y_1 and completely ignores any information in Y_2 whereas Rao and Khatri's involves the use of some or all of Y_2 as covariates as can be seen above.

Some of the general results using the theory of analysis of covariance are now given (Grizzle & Allen, 1969; Morrison, 1976; Seber, 1984).

The least squares (and maximum likelihood) estimates of Φ and Γ for the conditional model are:

$$\hat{\Gamma} = (Y_2^T R Y_2)^{-1} Y_2^T R Y_1$$

$$= (K_2^T X^T R XK_2)^{-1} K_2^T X^T R XK_1$$

$$= (K_2^T SK_2)^{-1} K_2^T SK_1$$

where $R = I_n - A(A^TA)^{-1}A^T$

and $S = X^T R X$

$$\hat{\Phi} = (A^TA)^{-1} A^T (Y_1 - Y_2 \hat{\Gamma})$$

$$= (A^TA)^{-1} A^T (XK_1 - XK_2 \hat{\Gamma})$$

$$= (A^TA)^{-1} A^TX (K_1 - K_2(K_2^TSK_2)^{-1} K_2^T SK_1$$

Since $BK_2 = 0$ and $BK_1 = I_q$ then

$$K_2(K_2^T SK_2)^{-1} K_2^T = S^{-1} - S^{-1} B^T(BS^{-1} B^T)^{-1} BS^{-1}$$

$$\hat{\Phi} = (A^TA)^{-1} A^T X S^{-1} B^T(BS^{-1} B^T)^{-1}$$

To test the general hypothesis given by:

$$H_0: \quad J\Phi V = 0$$

where

J is a known (sxm) matrix of rank s

and

V is a known (qxu) matrix of rank u,

we need to obtain the error and hypothesis matrices given by,

$$E = (n-m) V^T (BS^{-1} B^T)^{-1} V$$

and

$$H = (J\hat{\Phi}V)^T [J(A^TA)^{-1} J^T + J R J^T]^{-1} (J\hat{\Phi}V)$$

where

$$R = \frac{1}{n - m} (A^TA)^{-1}A^TX[S^{-1}-S^{-1}B^T(BS^{-1}B^T)^{-1}BS^{-1}]X^TA(A^TA)^{-1}$$

To test H_0 we may use one of the test criteria discussed previously. Both the union-intersection approach and the Wilks Λ test criteria are functions of the characteristic roots of HE^{-1}.

Note: $\hat{\Phi}$, E, R and H do not depend on K, so that the above test of H_0 does not depend on K. Therefore if the transformation suggested by Potthoff & Roy is used,

$$K_1 = G^{-1} B^T(BG^{-1} B^T)^{-1}$$

then the above test is independent of G and we can set $G=I_p$. This would give us

$$K_1 = B^T(BB^T)^{-1}$$

and choosing K_2 such that $B^TK_2 = 0$ implies that

$K_1^T K_2 = 0$. Therefore K_2^T may be chosen to be the linearly independent rows of the projection matrix,

$$I_p - B^T(BB^T)^{-1} B$$

For example, if normalised orthogonal polynomials are used then $BB^T = I_q$, $K_1 = B^T$. So we could choose K_1 to be the matrix of normalised orthogonal polynomials of degrees 0 to q-1, and K_2 to be a similar matrix for degrees q to p-1.

The above method uses all the p-q concomitant variables in Y_2. In the context of the single growth curve example, Rao (1965) suggests that a better procedure might be to select only a subset of the concomitant variables, particularly if the correlations between any of the columns of Y_1 and Y_2 are small. If the covariance of each column of Y_2 with each of Y_1 is zero, then Y_2 provides no information about Y_1 and Y_2 should be discarded. In this case the method suggested by Potthoff & Roy with $G = I_p$ is appropriate. Rao (1967) and Grizzle & Allen (1969) discuss the issue of selection of covariates in the conditional model and in particular discuss the possibility of using fewer then p-q covariables. As these authors note, for certain patterned covariance matrices, a certain subset of the set of (p-q) covariates contains all of the concomitant information. An example by Grizzle & Allen (1969) illustrates the effect of using the entire set of covariables versus a subset.

Grizzle & Allen (1969) also discuss estimators employed by Potthoff & Roy (1964), Khatri (1966) and Rao (1965-67) and they note that,

Rao's (1966) estimator with p-q covariables used	=	Potthoff & Roy's estimator with G = S	=	Khatri's estimator

This was also noted independently by Lee (1974) and extended by Baksalary et al (1978) to the case when not all the covariates are used.

The set of simultaneous confidence intervals corresponding to the union-intersection test of H_0: $J\Phi V = 0$ with probability $1-\alpha$ is now given.

$\underline{a}^T J\Phi V\underline{b}$ is contained in the interval,

$$\underline{a}^T J\hat{\Phi} V\underline{b} \pm [(\gamma/(1-\gamma)).\ \underline{a}^T F\underline{a}\ .\ \underline{b}^T E\underline{b}]^{\frac{1}{2}}$$

for all $\underline{a}$ and $\underline{b}$ where,

$$F = J(A^TA)^{-1}\ J^T + JRJ^T$$

and γ is obtained from tables for Roy's Maximum Root Statistic. (Union-intersection approach of Roy).

When only a few specific linear contrasts are of interest, shorter intervals can be obtained by using the Bonferroni intervals.

Gafarian (1978) develops two methods for constructing confidence bands for growth curve data assuming that there is a polynomial trend of known degree. Tolerance bands for the population growth curve are derived by Bowden & Steinhorst (1973).

Tubbs, Lewis & Duran (1975) generalise the general growth curve model by relaxing the assumption of independence of the rows of X and consider a general covariance matrix, $\Sigma \oplus \Psi$, for the vector x (pn x 1), a vector composed of the transposed rows of X stacked on top of each other. The (nxn) matrix Ψ is assumed to be positive definite. They derive directly the maximum likelihood estimator for Φ under the general growth curve model, for this general covariance structure. Estimators are derived

under the restriction $J\Phi V = Q$ (a fixed matrix) and under no linear constraints on Φ. Aspects of testing H_0: $J\Phi V = 0$ are discussed and the results are applied to a problem initially analysed by Beauchamp & Hoel (1974).

Note: $\oplus$ denotes the Kronecker product.

There is an alternative procedure that considers the individual growth curves as well as the overall one for a given data set. Essentially this means making use of the two-stage model outlined in section 2.

Letting $\underline{X}_i$ be the p-dimensional vector of observations on the i'th individual then,

$$\underline{X}_i \sim N_p(B^T \underline{\Phi},\ B^T \wedge B + \sigma^2 I_p) \text{ for } i=1...n$$

This model is similar to the general growth curve model but it has a more structured covariance matrix as can be seen above.

Inference can be based either on Least Squares and Maximum Likelihood or on empirical Bayes methodology.

For this model, the minimum variance unbiased estimate of $\underline{\Phi}$ is $(BB^T)^{-1} B \overline{\underline{X}}$ where

$$\overline{\underline{X}} = 1/n\ 1_n{}^T \begin{bmatrix} \underline{X}_1{}^T \\ \vdots \\ \underline{X}n^T \end{bmatrix}$$

If we choose $K_1 = B^T(BB^T)^{-1}$ which we can do since the analysis of covariance method does not depend on the choice of K, then, we will have a special case of the situation where the covariances between corresponding rows of Y_1 and Y_2 are zero as shown by Seber (1984). As mentioned previously, Y_2 may then be discarded since it provides no concomitant information. Hence we can therefore proceed, using the methods suggested by Potthoff & Roy (1964).

The two-stage model can also be written in the form of a regression model with stochastic coefficients which arise

naturally in random effects analysis of variance models. Bowden & Steinhorst (1973) use the two-stage model to construct a tolerance band so that a given proportion of individuals have their (conditional) expected growth curves $(1, t, t^2, \ldots)\ \underline{\gamma}$ lying in the band for all t, with an overall probability of approximately $1-\alpha$.

Joreskog (1970) investigates a model similar to the two-stage model but with a more general covariance structure.

The main problem Joreskog considers is the estimation of parameters of his general model. He also considers hypothesis testing and gives a number of examples including growth curves with serially correlated errors. Laird & Ware (1982) discuss a general family of random-effects models which includes both growth models and repeated measures models as special cases. The family of two-stage models for repeated measurements which they introduce is based on the work of Harville (1977). They describe a unified approach to inference using these models discussing both maximum likelihood and empirical Bayes estimation.

Fearn (1975, 1977) gives a Bayesian analysis for the two-stage model. In his 1975 paper he also considers the prediction problem and compares several Bayesian predictors with those given by Lee and Geisser (1972). The model as given by Fearn (1975) takes the following form.

$$\underline{X}_i/\underline{\Phi},\ \sigma^2{}_i,\ \Lambda \ \sim\ N(B_i{}^T\underline{\Phi},\ B_i{}^T \Lambda B_i + \sigma_i{}^2 I_p)$$

for i=1...n.

Fearn gives the posterior distributions of the first stage parameters $\underline{\gamma}_i$ given $X^T = (\underline{X}_1{}^T \ldots \underline{X}_n{}^T)$ for known variances. They were found to be normal with mean vectors

$$\underline{\gamma}_i{}^* = W_i\ \hat{\underline{\gamma}}_i + (I - W_i)\ \left(\sum_{j=1}^{p} W_j\right)^{-1} \left(\sum_{j=1}^{n} W_j\ \hat{\underline{\gamma}}_j\right)$$

and covariance matrices

$$V^*(\underline{\gamma}_i) = \{W_i + (I-W_i)(\sum_{j=1}^{n} W_j)^{-1} W_i\}\sigma_i^2(B_iB_i^T)^{-1}$$

where

$$W_i = (\sigma_i^{-2} B_iB_i^T + \Lambda^{-1})^{-1} \sigma_i^{-2} (B_iB_i^T)$$

$$\hat{\underline{\gamma}}_i = (B_iB_i^T)^{-1} B_i\underline{X}_i$$

As noted by Fearn, the estimation of the second stage parameter $\underline{\Phi}$ was considered by Smith (1973) who derived the posterior distribution of $\underline{\Phi}$ given X, again for known variances, to be normal with mean vector

$$\underline{\Phi}^* = \left[\sum_{i=1}^{n} W_i\right]^{-1} \left[\sum_{i=1}^{n} W_i \hat{\gamma}_i\right]$$

and covariance matrix

$$V^* (\underline{\Phi}) = \left[\sum_{i=1}^{n} \{\sigma_i^2 (B_iB_i^T)^{-1} + \Lambda\}^{-1}\right]^{-1}$$

For the case when we have equal B_i's, σ_i^2's and W_i's, the above expressions reduce to,

$$\underline{\gamma}_i^* = W \hat{\underline{\gamma}}_i + (I-W)n^{-1} \sum_{j=1}^{n} \hat{\underline{\gamma}}_j$$

$$V^*(\underline{\gamma}_i) = [W + n^{-1} (I-W)] \sigma^2 (BB^T)^{-1}$$

$$\underline{\Phi}^* = n^{-1} \sum_{i=1}^{n} \hat{\underline{\gamma}}_i$$

$$V^*(\underline{\Phi}) = n^{-1}[\sigma^2(BB^T)^{-1} + \Lambda]$$

For the situation where the variances σ_i^2 and the covariance matrix Λ are unknown Fearn approximates the posterior distributions of Φ and the γ's by substituting estimates of the unknown variances in the appropriate expressions given previously. This procedure will give reasonable estimates for the posterior means, so long as the estimates of variance are reasonable, but will underestimate the posterior variances. except when the B_i's and σ_i's are equal. Further results and applications of these results may be found in Fearn (1975, 1977).

Darby & Fearn (1979) utilise the Bayesian analysis of the two-stage model in a longitudinal study of blood pressure in

children. Dunsmore (1981) compares two different approaches to the analysis of repeated measurements two-period change-over designs. One of these approaches being Fearns (1975) Bayesian approach using the two-stage model.

3.4: Other Approaches

There are several alternative approaches for analysing repeated measures data but these are less frequently applied and hence less occurent in the literature. Two approaches that we mention here are the Bayesian and non-parametric.

A Bayesian treatment of the generalised growth curve model is developed by Geisser (1970, 1971) and Lee & Geisser (1972). The problem of prediction of individual growth curves is addressed by these authors in addition to their discussion of the Bayes' estimators. Geisser (1970) provides a Bayesian justification for Rao's (1967) covariance adjusted estimator.

Gosh, Grizzle & Sen (1973) present generally applicable rank based statistical methods for repeated measurements, appropriate either when multivariate normality does not hold or when the measurements take on discrete values from some nominal scale.

Goldstein (1979) illustrates Gosh et al's (1973) results using some artificial data. Univariate and multivariate nonparametric techniques are used to test for differences among the observed growth patterns. The rank sum test given by Gosh et al (1973) is in fact an adaption of that developed by Chatterjee & Sen (1966) to the problem of comparing mean growth curves.

For the special case of a completely randomised design, Zerbe & Walker (1977) introduce a randomisation test for comparing mean growth curves over an interval of time specified by the investigator. The order of the polynomial does not have to be the same in each group. Further work on this randomisation test is done by Zerbe (1979).

The analysis of longitudinal categorical data is examined by Koch et al (1977) and Plewis (1981).

Koziol et al (1981) present a distribution-free test for tumour-growth curve analyses.

Finally, Koch et al (1980) review some general approaches to the analysis of repeated measures data and present some views on parametric and non-parametric methods.

3.5: Analysis of Unbalanced Repeated Measures Data

As said in Chapter 2, the two-stage random effects model (see section 2.2 and 3.3) may be used to handle unbalanced repeated measures data because of its consideration of individual growth curves.

Most of the literature however, is concerned with the model given by Kleinbaum (1973) and Schwertman et al (1981). Kleinbaum (1973) gives a generalisation of the growth curve model that allows for data which is missing, either by accident or by design (see section 2.3).

Several authors have examined this model including Woolson, Leeper and Clarke (1978) and Woolson and Leeper (1980). For this model Kleinbaum (1973) presents a BAN (Best Asymptotically Normal) estimator of an estimable linear function $M^1\underline{b}$, a straightforward and consistent estimator of the covariance matrix Σ, an asymptotic Wald test of the hypothesis H_0: $M^1\underline{b} = 0$ and a competing test statistic. He derives these by first rewriting the data in the form of a univariate linear model. Following Kleinbaum, the columns of X_j are stacked on top of each other into the n_jq_j vector $\underline{z}_j$. Then setting $Z = (\underline{z}_1^T,\ldots,\underline{z}_u^T)$ it follows that

$$Z = D\,\underline{b} + \underline{\delta}$$

where

$\underline{b}$ is a (mpx1) vector consisting of the columns of Φ stacked on top of each other and

$$\underset{(n \times mp)}{D} = \begin{bmatrix} B_1^T P^T \oplus A_1 \\ \vdots \\ B_u^T P^T \oplus A_u \end{bmatrix}$$

The vector $\underline{\delta}$ follows a $\sum_{j=1}^{u} n_j q_j$ ~ variate normal distribution

with mean $\underline{0}$ and covariance matrix Ω where:

$$\Omega = \text{diag}[B_1^T \Sigma B_1 \oplus I_{n_1}, \ldots\ldots, B_u^T \Sigma B_u \oplus I_{n_u}]$$

Letting $M^T = V^T \oplus J$ then $M^T\underline{b}$ contains the same linear compounds as $J\Phi V$ in a re-arranged fashion. Kleinbaum then applies a theory by Wald (1943) and derives the following results as outlined by Woolson & Leeper (1980):

(i) A best asymptotic normal estimator of $M^T\underline{b}$ is given by

$$M^T\hat{b} = M^T (D^T\hat{\Omega}^{-1}D)^{-1} D^T \hat{\Omega}^{-1} Z$$

where

$\hat{\Omega}$ is an estimator of Ω obtained by substituting any consistent estimator for Σ in the equation for Ω given previously.

(ii) The asymptotic covariance matrix of $\widehat{M^T\underline{b}}$ is estimated by

$$T = M^T (D^T \hat{\Omega}^{-1}D)^{-1}M$$

(iii) Under the null hypothesis H_0: $J\Phi V = 0$ the statistic,

$$(\widehat{M^T\underline{b}})^T T^{-1} (\widehat{M^T\underline{b}})$$

follows a chisquare distribution with sc degrees of freedom as $n \to \infty$ where $s = \text{rank}(J)$ and $c = \text{rank}(V)$.

(iv) If the data are all complete then the statistic in (iii) reduces algebraically to n times the trace of H^*E^{*-1}

where

$$H^* = (J\hat{\Phi}V)^T [J (A^TA)^{-1}J^T]^{-1} (J\hat{\Phi}V)$$

$$E^* = V^T (BS^{-1}B^T)V$$

$$\hat{\Phi} = (A^TA)^{-1} A^T X S^{-1} B^T (BS^{-1}B^T)^{-1}$$

$$nS = X^T[I - A(A^TA)^{-1} A^T] X$$

Further details and discussion may be obtained from Kleinbaum (1973).

Other workers have also investigated the incomplete growth curve problem.

Srivastava & McDonald (1974) study the hierarchical growth curve model and develop methods which utilise a procedure similar to Roy's (1958) step down procedure for all sample sizes. The application of their results is limited primarily by the restricted pattern of missing data required.

Machin (1975) discusses a situation rather less general than those covered by Kleinbaum (1973) in which certain observations are omitted by design but compensated for by the introduction of new subjects so that the total number of observations remain constant.

In growth curves, it is usually assumed that subjects are measured at identical times, modelled with polynomials of identical degree and multivariate normality of the measurements can be assumed. A method which may be used when these assumptions are relaxed is given by Zerbe & Walker (1977).

Koziol et al (1981) describe a distribution-free procedure for the comparison of growth curves which may be used with incomplete data.

CHAPTER 4

Illustrations

Section	Content
4.1	Introduction
4.2	A study of the role of Prostaglandin I_2.
4.3	An investigation into the measurement of strength in children from four age bands.
4.4	The efficacy of fundal height in the identification of growth retardation.

4.1: Introduction

Having examined the structure of repeated measures data, outlined some repeated measures designs, presented some models and possible methods of analysis, we will now illustrate some of this methodology using three different sets of data.

The repeated measures designs from which these data sets arise vary in their complexity. We first examine a two factor repeated measures design with two trial factors and then a three factor design with two grouping factors and one trial factor. The last data set that we examine arises from a discrimination problem where no design was used in the collection of the data. Hence this resulted in an extremely unbalanced set of growth curve data.

It should be noted that these examples are being used to illustrate the methodology given in earlier chapters and hence the analysis reported here may not necessarily be a full analysis of the problem. Indeed some of the problems may be analysed using several methods, only some of which may be appropriate.

4.2: The Role of Prostaglandin I_2

4.2.1: Introduction to the Problem

Prostaglandin I_2 (PGI_2) is known to be synthesised in the blood vessels and in the kidney. Its action is to cause vasodilation and hence decrease systemic blood pressure. In the kidney it increases renal blood flow and causes sodium excretion. Several disease states (e.g. hypertension) are thought to be partly due to insufficient synthesis of PGI_2 and it has been suggested that PGI_2 acts as a circulating hormone to control blood pressure and renal function.

More recent evidence, however, indicates that this may be an exaggeration and that PGI_2 may be no more than a local modulator akin to histamine.

The elucidation of the role of PGI_2 is complicated by its short half-life, making measurement of this compound impossible. The stable metabolite of PGI_2, 6-Keto-Pgf is therefore routinely measured as an indicator of PGI_2 production. However, previous to this study no information was available as to the precise relationship between PGI_2 and its metabolite.

Therefore a study was carried out to investigate the relationship between PGI_2 and its metabolite and to examine the effects of intravenous infusion of PGI_2 on systemic and renal haemodynamics and electrolyte excretion to evaluate the possibility of PGI_2 acting as a circulating hormone.

In the study eight male dogs were used and each animal received four intravenous infusions. One of these infusions was a control and the other three were different concentrations of PGI_2. The concentrations of PGI_2 used were

7.5 ng/kg/min,

15.0 ng/kg/min,

30.0 ng/kg/min.

Each animal received only one infusion per day and each infusion took place over four hours. Several variables were measured at thirty minute intervals over the infusion period. The first time point at which measurements were taken was thirty minutes into the infusion period. For ease of presentation we will use $T_1, T_2, \ldots, T_8$ to represent the time points during the infusion period when measurements were taken.

T_1 represents 30 minutes into infusion

T_2 " 60 " " "

. . .

. . .

. . .

T_8 " 240 " " "

Measurements were taken at each time point on the following variables,

6-Keto-Pgf,

Sodium Excretion,

Renal Blood Flow,

and Systemic Blood Pressure.

Figure 4.2.1 shows the repeated measures design and hence the format of the measurements obtained. This format is the same for each of the above variables.

Figure 4.2.1: Repeated Measures Design

	Solution (PGI_2)			
	Control	(7.5)	(15)	(30)
	$T_1, T_2, \ldots, T_8$	$T_1, T_2, \ldots, T_8$	$T_1, T_2, \ldots T_8$	$T_1, T_2, \ldots, T_8$
1				
Animal 2				
.				
.				
.				
N				

Using the terminology introduced in chapter one we can see that the above is a two factor experiment with two trial factors. These being dose and time with four and eight levels respectively.

From the background information given previously we summarise the following points about the action of PGI_2:

(i) Systemic blood pressure should decrease

(ii) Renal blood flow should increase

and (iii) Sodium should be excreted.

We will now examine the data obtained for these variables and the relationship (if any) between PGI_2 and its metabolite.

4.2.2: Analysis of the Data

For 6-Keto-Pgf measurements were available on only six out of the eight animals and for sodium excretion, renal blood flow and systemic blood pressure measurements were available on seven of the animals.

In general there was a lot of between animal variability and to keep diagrams clear and uncluttered only the means are plotted against time for each dose of PGI_2.

For each of the variables mentioned above, a univariate repeated measures analysis of variance was carried out using the BMDP (1983) program P2V. Included in the output from this program are results from applying the Greenhouse & Geisser and Huynh & Feldt approximate tests (section 3.2) as well as the standard univariate F-tests.

Since there were no grouping factors in this repeated measures design only the Mauchly W criterion was needed to assess the validity conditions required by the within-subject F-tests. It is interesting to note that although the doses of PGI_2 were not randomly allocated, the covariance matrix for this effect did not differ significantly from the necessary form for each of the four variables. This could be partly due to a lack of power in the test due to extremely small sample sizes. However the covariance matrix for the time effect did differ significantly from the necessary form. It seems more likely that this covariance matrix would exhibit some serial correlation.

Hence for those within-subject effects concerning time (main effect and dose by time interaction) we must use the results from either the Greenhouse & Geisser or the Huynh & Feldt approximate tests.

For details of the univariate analysis of variance model and methodology see sections 2.2 and 3.2 respectively.

In the following section the results obtained from applying univariate analysis of variance are presented for each of the four variables. In the tables containing these results abbreviated headings are given where,

F-Value represents the value of the univariate F test statistic obtained by the ratio of the appropriate mean squares.

P-Value represents the probability of obtaining a value greater than the observed value of the F test statistic under the null hypothesis of no difference.

and

G-G and H-F represent the probability of obtaining a value greater than the observed value of the test statistic when the Greenhouse & Geisser and Huynh & Feldt approximate tests are being used respectively.

4.2.3: Results

The Effect of PGI_2 on Sodium Excretion

A graph of the average sodium excretion against time for each of the four doses (control = dose zero) is contained in figure 4.2.2. Given the large variability between individual animals there would seem to be no clear difference between the sodium excreted at different levels of PGI_2. However there may be a slight time effect.

The results from carrying out a univariate repeated measures analysis of variance can be seen in table 4.2.1.

Table 4.2.1: Univariate Anova-Sodium Excretion

Source	F-value	P-Value	G-G	H-F
Between Subjects	14.46	0.0089*		
Within Subjects				
- Dose	1.06	0.3918	0.3820	0.3918
- Time	5.19	0.0003*	0.0389*	0.0269*
- Dose by Time	1.56	0.0691	0.2193	0.1274

Note: *indicates significance for α=0.05

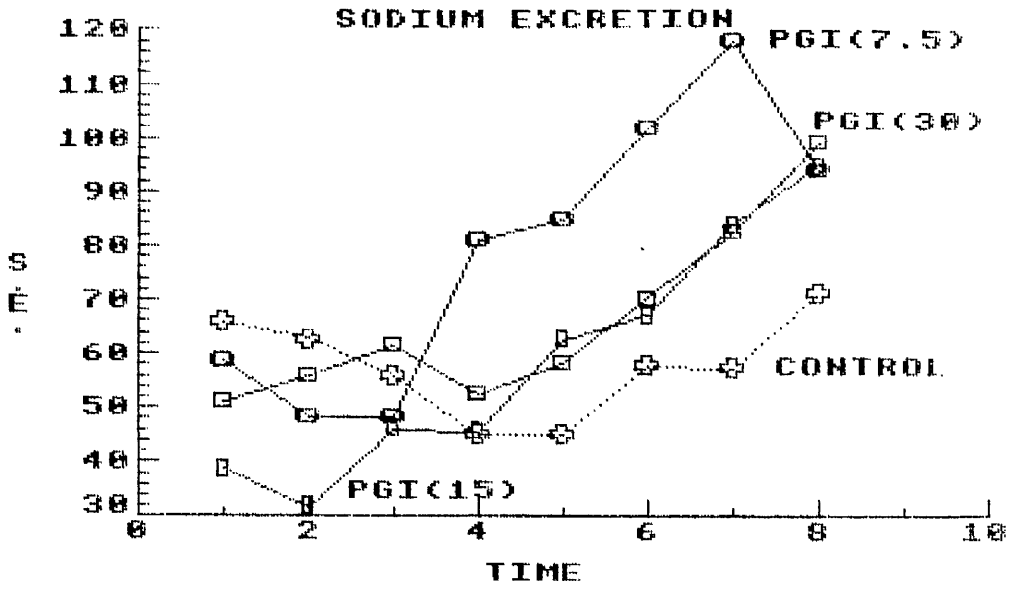

SODIUM EXCRETION
PGI(7.5)
PGI(30)
CONTROL
PGI(15)
120
110
100
90
80
70
60
50
40
30
0
2
4
6
8
10
TIME

Figure 4.2.2.

As can be seen from table 4.2.1 time is the only significant effect. The orthogonal polynomial breakdown of the total variation due to time indicated significant linear and quadratic components.

The Effect of PGI_2 on Renal Blood Flow

Figure 4.2.3 shows the average renal blood flow against time for each of the four doses of PGI_2 (control = dose zero). Given the large variability between the animals there would seem to be no clear differences between the renal blood flow over time or dose.

The results from carrying out a univariate repeated measures analysis of variance can be seen in table 4.2.2.

Table 4.2.2: Univariate Anova-Renal Blood Flow

Source	F-value	P-value	G-G	H-F
Between Subjects	66.62	0.0002*		
Within Subjects				
- Dose	1.67	0.2099	0.2178	0.2099
- Time	1.36	0.2465	0.2916	0.2788
- Dose by Time	0.82	0.6872	0.5337	0.6798

Note: *indicates significance for $\alpha = 0.05$

From table 4.2.2 we can see that there are no significant effects.

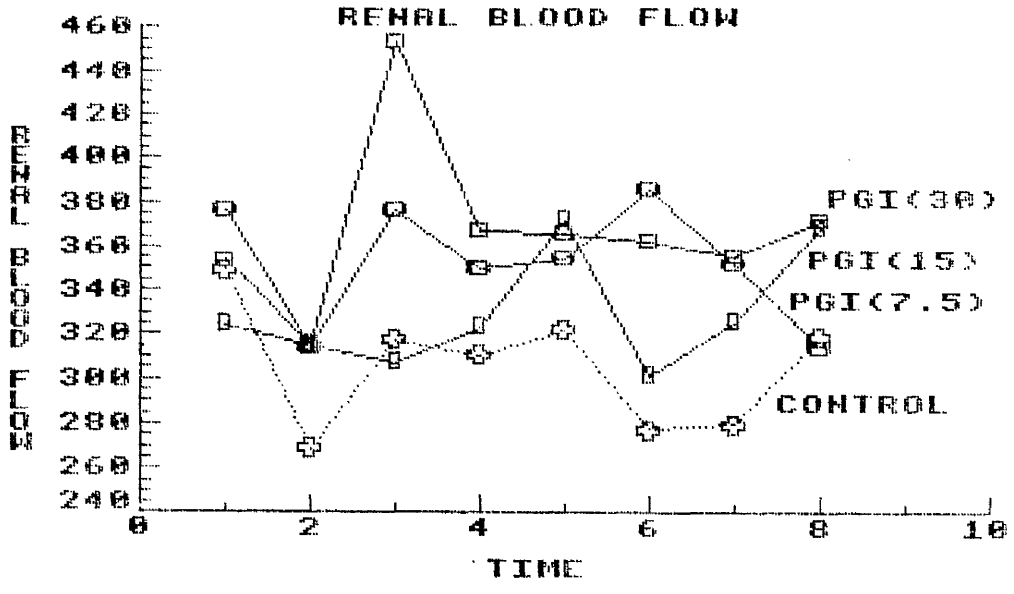

RENAL BLOOD FLOW
RENAL BLOOD FLOW
460
440
420
400
380
360
340
320
300
280
260
240
0
2
4
6
8
10
TIME
PGI(30)
PGI(15)
PGI(7.5)
CONTROL

Figure 4.2.3

The Effect of PGI_2 on Systemic Blood Pressure

Figure 4.2.4 shows the average systemic blood pressure over time for each of the four doses of PGI_2 (control = dose zero). As can be seen from this graph there does appear to be some evidence of a general linear trend over time as well as some dose effect. The lines for doses 15 and 30 never overlap on the graph with the lines for the control and dose 7.5.

The results from the univariate repeated measures analysis of variance can be seen in table 4.2.3.

Table 4.2.3: Univariate Anova-Systemic Blood Pressure

Source	F-value	P-value	G-G	H-F
Between Subjects	1484.62	0.0000*		
Within Subjects				
- Dose	10.95	0.0003*	0.0026*	0.0005*
- Time	15.77	0.0000*	0.0001*	0.0000*
- Dose by Time	1.99	0.0105*	0.1237	0.0238*

Note: * indicates significance for $\alpha = 0.05$

Here we are faced with an interesting dilemma. Looking at table 4.2.3 the interaction term, dose by time, is significant for the Huynh & Feldt approximate test but not for the Greenhouse & Geisser. The conclusion to draw from the results in this instance is uncertain. Here we have to remember that our sample size is quite small (n=7) and that the tests being used are approximate.

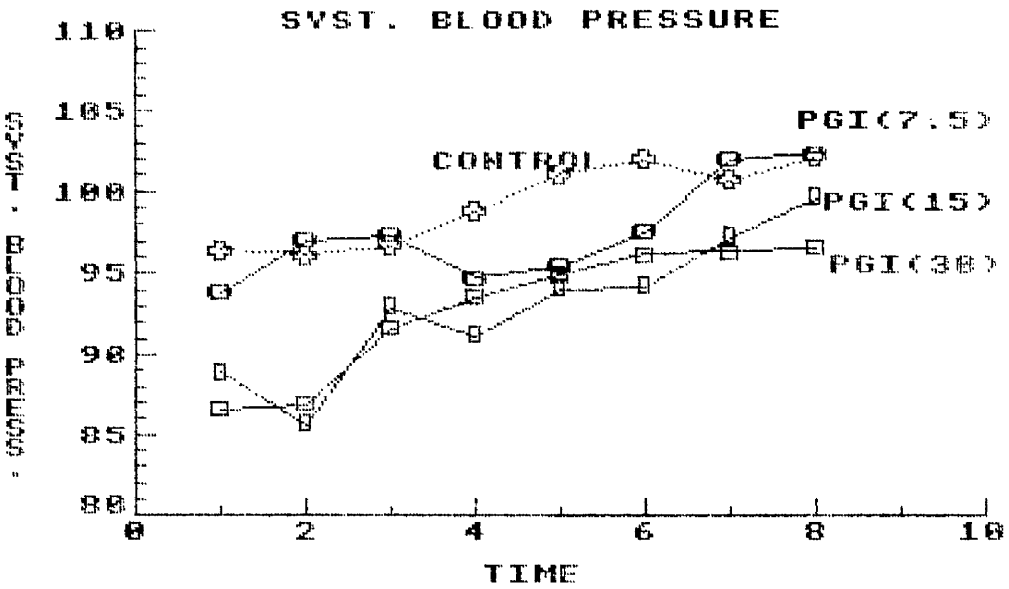

SYST. BLOOD PRESSURE
110
105
100
95
90
85
80
SYST. BLOOD PRESS.
0
2
4
6
8
10
TIME
CONTROL
PGI(7.5)
PGI(15)
PGI(30)

Figure 4.2.4

If we assume that the Greenhouse & Geisser approximate test is appropriate then since the interaction term is not significant we can go on to test the main effects for dose and time. From table 4.2.3 we can see that both of these main effects are significant.

From the orthogonal polynomial breakdown of the total variation due to time only the linear component was found to be significant.

The orthogonal decomposition for the unequally spaced dose factor was also examined and only the linear component was found to be significant.

If we were to assume that the Huynh & Feldt approximate test was appropriate then we would have a significant dose by time interaction. This significant interaction term makes it difficult to interpret any of the other dose and time main effects. Further analyses would have to be done to examine these effects jointly. For example examining the time effect for each of the four doses individually using a one factor repeated measures analysis of variance.

PGI$_2$ and its metabolite 6-Keto-Pgf

Figure 4.2.5 shows the average of 6-Keto-Pgf over time for the four doses (control = dose zero). Even taking into consideration the large variability between animals from this graph there would appear to be some evidence of a dose effect but the time effect is less clear.

Table 4.2.4 summarises the results obtained from applying univariate repeated measures analysis of variance.

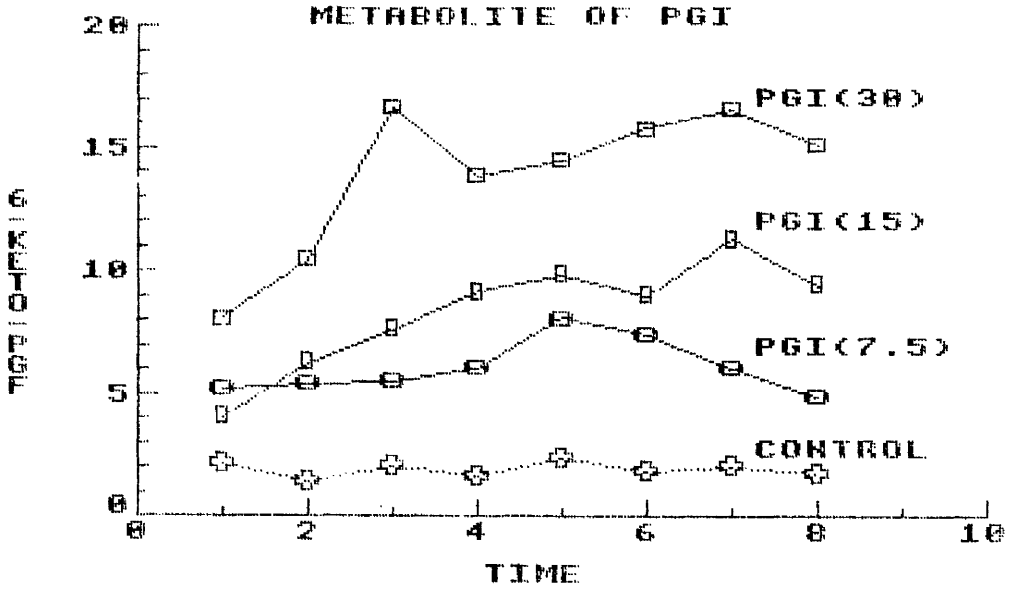

METABOLITE OF PGI
PGI(30)
PGI(15)
PGI(7.5)
CONTROL
6 KETO PGF
TIME
0
5
10
15
20
0
2
4
6
8
10

Figure 4.2.5

Table 4.2.4: Univariate Anova - 6-Keto-Pgf

Source	F-value	P-value	G-G	H-F
Between Subjects	50.89	0.0008*		
Within Subjects				
- Dose	23.52	0.0000*	0.0016*	0.0006*
- Time	6.21	0.0001*	0.0095*	0.0004*
- Dose by time	2.22	0.0043*	0.1138	0.0177*

Note: * indicates significance for $\alpha = 0.05$.

Again we are faced with a dilemma. For 6-Keto-Pgf the Huynh & Feldt approximate test shows a significant result for the dose by time interaction and the Greenhouse & Geisser does not. As before the situation is unclear. Some of this confusion may possibly be caused by having a very small sample (n=6) and using approximate tests.

If we assume that the Greenhouse & Geisser approximate test is apropriate then we have a non-significant interaction and two significant main effects. From the orthogonal polynomial breakdown of the total variation due to time both the linear and quadratic components were found to be significant.

The orthogonal decomposition for the unequally spaced dose effect showed only a significant linear component.

If we assume that the Huynh & Feldt approximate test is appropriate then we would have a significant dose by time interaction effect and hence our main effects for dose and time would have to be examined jointly in a further analysis since their interpretation from table 4.2.4 is difficult.

Given that our interest here lies in the relationship between PGI_2 and its metabolite 6-Keto-Pgf a univariate analysis of variance was also carried out on the data using only the three doses of PGI_2 and eliminating the results obtained for the control. These results are given in table 4.2.5.

Table 4.2.5:

Source	F-value	P-value	G-G	H-F
Between Subjects	45.95	0.0011*		
Within Subjects				
- Dose	18.11	0.0005*	0.0025*	0.0008*
- Time	6.18	0.0001*	0.0107*	0.0007*
- Dose by Time	1.75	0.0651	0.1902	0.0775

Note: * indicates significance for $\alpha = 0.05$

Having removed the control we can see that the dose by time interaction term is no longer significant although it is borderline according to the Huynh & Feldt approximate test. Assuming that there is no significant interaction we can go on to examine the main effects which are both significant on looking at our two approximate tests.

Both the linear and quadratic components for the orthogonal breakdown of the total variation due to time are significant, but only the linear component is significant for dose (orthogonal polynomial breakdown for unequally spaced factor).

4.2.4: Summary

Univariate repeated measures analysis of variance was used to analyse the measurements obtained for the four different variables. Each variable being examined separately. When assessing the validity conditions for this analysis, the covariance matrix associated with the dose effect did not differ significantly from the necessary form, whereas the covariance matrix for the time effect did. It seems more likely that for this effect, observations closer together in time would be more highly correlated than observations further apart. Hence the covariance matrix may exhibit some serial correlation.

Unfortunately the sample sizes obtained were quite small. For 6-Keto-pgf only six animals had a full set of measurements and only seven animals for the other three variables. This has to be noted since small sample sizes affect the power of the symmetry test used in assessing the validity conditions. Also from examining plots of the measurements over time for the individual animals there appeared to be a large amount of between animal variation. This again can affect the power of the symmetry test and the tests of the effects in the analysis of variance.

Due to the lack of the required form of the covariance matrix for the time effect, the Greenhouse & Geisser and Huynh & Feldt approximate tests were used to determine the significance of the dose by time interaction and the main effect for time. Unfortunately these two approximate tests did not always agree in the final conclusions. For two of the variables, namely systemic blood pressure and 6-Keto-pgf there was some confusion over the significance of the dose by time interaction. In both cases the Huynh & Feldt approximate test found the interaction term to be

significant, while the Greehnouse & Geisser did not.

Using the results obtained from applying the Greenhouse & Geisser approximate test we would conclude that PGI_2 has only had a significant effect on systemic blood pressure and the production of 6-Keto-pgf.

Alternative methods of analysing this data could have been used. Since both of the trial factors have levels taken over a continuous scale, the modelling approach could have been applied. This is reinforced by the orthogonal polynomial breakdown for the variation due to time which showed significant linear and quadratic terms for sodium excretion and 6-Keto-pgf and a significant linear term for systemic blood pressure. The breakdown for the variation due to dose (unequally spaced) showed a significant linear trend for both systemic blood pressure and 6-Keto-pgf.

Due to the small sample sizes the multivariate approach could not be used.

4.3. : Strength in Children

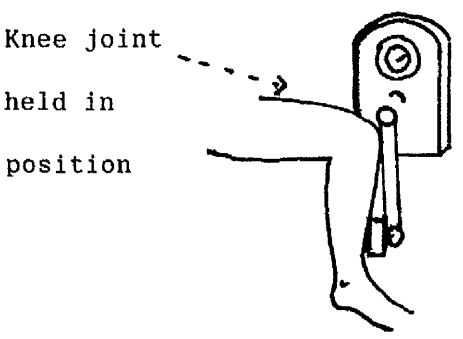
4.3.1.: Introduction

At present, development of strength and power in children is monitored using various isometric tests such as grip strength. A study was carried out to investigate the strength of children in four different age bands using measurements obtained from a Cybex Isokinetic Dynamometer. The four age bands being 5-5½, 8-8½, 11-11½ and 14-14½ years. Samples of male and female children in each age band were obtained from local schools.

Strength is usually measured on a Cybex as the Peak Torque (Newton-metres) developed during a maximal concentric contraction. In this study the peak torque obtained through knee extension contractions of the left leg was measured at four different velocity settings of the Cybex. Since the peak torque varies through the range of movement, a recorder was attached to the Cybex from which a graph of the strength curve could be obtained. The actual peak torque was then read from this graph.

To reduce variation in positioning which could affect the readings obtained, each subject was strapped in to the Cybex with the lever of the Cybex strapped to their left leg (see figure 4.3.1).

Figure 4.3.1: Cybex Isokinetic Dynamometer



Measurements were obtained at four different velocity settings of the Cybex. Namely 300, 210, 120 and 30 deg.s^{-1}. Each subject had three warm up tries at each velocity followed by four maximal test contractions where they were encouraged to kick their leg as quickly and as strongly as possible. Appropriate rest periods were given between contractions.

This procedure was repeated one week later with the same subject and the maximum peak torque out of the eight measured was taken at each velocity.

Subjects all started at the fastest velocity namely 300 deg.s^{-1} which gives the greatest peak torque and progressed down to the slowest velocity, 30 deg.s^{-1} which is the more difficult for knee extensions. This procedure was used to try and minimise any carryover effects. For further information on the use of the Cybex see MacDougall, Wenge and Green (1982).

It was not possible to obtain an equal number of subjects in each group and hence table 4.3.1 outlines the sample sizes obtained.

Table 4.3.1: Sample Sizes

		Age Group (Years)			
		5-5½	8-8½	11-11½	14-14½
Sex	Male	14	15	16	14
	Female	12	13	15	15

Table 4.3.2 shows the structure of the repeated measures design and contains the average peak torque at each velocity setting for each of the groups.

Table 4.3.2: Average peak torque

Sex	Age	Velocity (deg.S^{-1})			
		300	210	120	30
Male	5-5½	3.1	5.0	7.3	12.6
	8-8½	4.7	8.0	13.2	24.5
	11-11½	10.3	16.8	28.0	57.0
	14-14½	22.4	35.3	53.5	85.7
Female	5-5½	2.5	4.9	7.4	12.4
	8-8½	5.8	9.5	16.1	29.5
	11-11½	9.6	15.3	25.3	47.8
	14-14½	17.4	27.8	46.4	79.6

To examine the relationship between the peak torque and velocity for different groups figure 4.3.1 was constructed. As can be seen from this graph the relationship between these two variables is not linear but curved slightly. To simplify this relationship and hence the handling of the data, a log transformation of the peak torque was used. Figure 4.3.2 shows the average log of the peak torque against velocity for each group. From examining this graph, there appears to be a fairly strong linear relationship between the log of the average peak torque and velocity. This relationship was also found to exist for the individuals data.

The profiles in figure 4.3.2 suggest that this linear relationship is not the same for the four age groups and possibly not the same for males and females but the difference between the sexes is less pronounced.

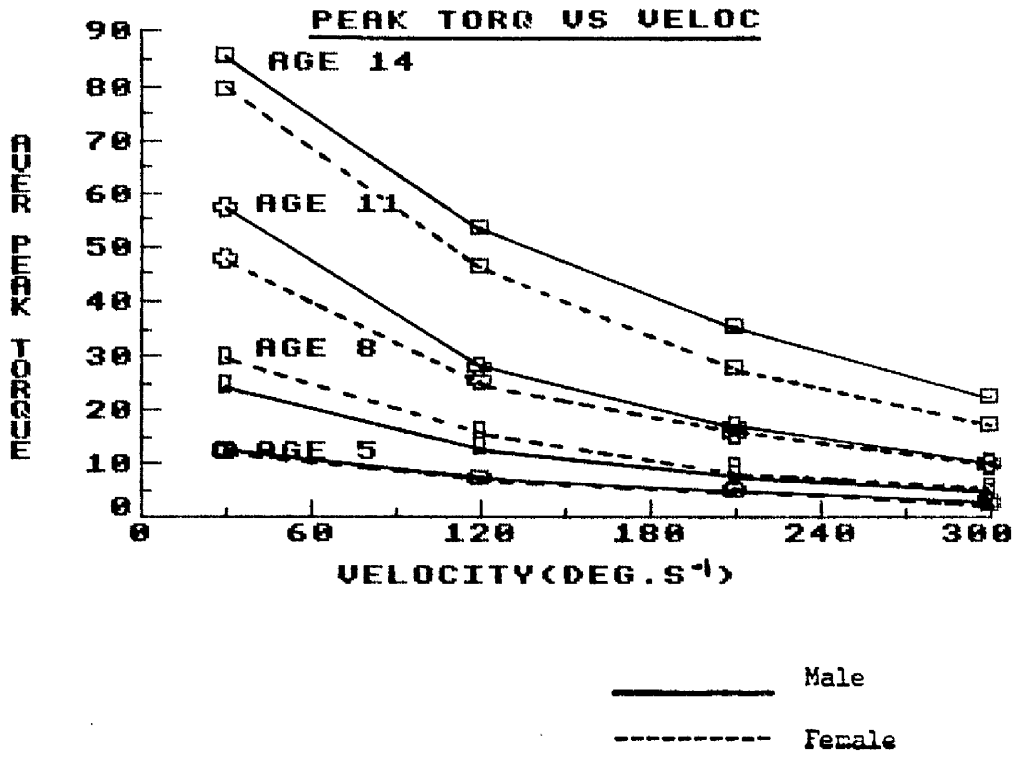


Figure 4.3.1: Peak torque against velocity

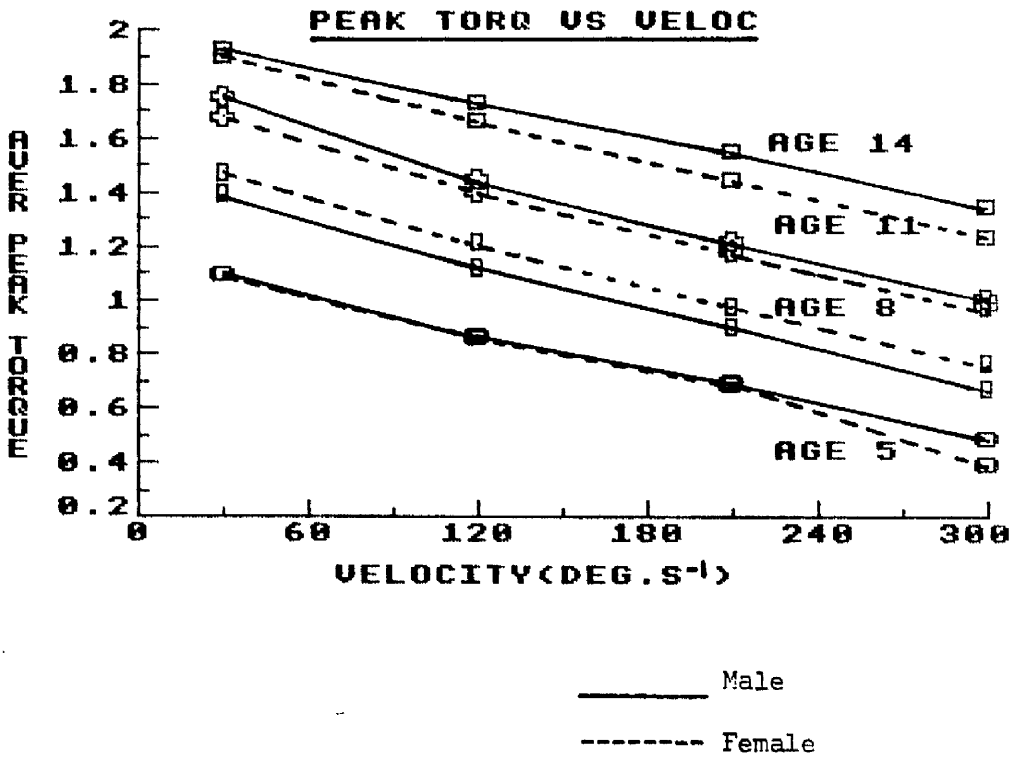


Figure 4.3.2: Log of peak torque against velocity

4.3.2: Analysis

Using the terminology introduced in chapter one, we have a three factor repeated measures design with two grouping factors (sex and age) and one trial factor (velocity). All three factors being fixed factors. The trial factor having levels taken over a continuous scale.

Our interest here is in investigating the effects of velocity, age, sex and their interactions on the peak torque. In our investigation of these effects we may make use of the linear relationship between the log of the peak torque and velocity. Hence several of the approaches discussed in chapter three may be applied. In total four different methods of analysis have been used to analyse this data and the results are presented below.

The following methods have been applied:

(i) Univariate repeated measures analysis of variance.

(ii) Multivariate analysis of variance

(iii) The modelling approach using the two-stage model outlined in section 2.2.

(iv) The modelling approach using the conditional model given by Rao (1965-67) as outlined in section 2.2.

Comparison of the results from using these different methods will be made in a later section.

Having discovered a strong linear relationship between the log of the peak torque and velocity, all analyses were carried out using the log of the peak torques and the appropriate programs in the BMDP (1983) statistics package. The program P2V was used to carry out the univariate repeated measures analysis of variance and the program P4V for the other three methods. For further information on the use of these two BMDP programs in

univariate and multivariate repeated measures analysis of variance see Davidson (1980) and Frane (1980).

4.3.3: Results

Univariate Repeated Measures Analysis of Variance

The univariate analysis of variance model for this three factor design is written out below where we let $X_{ijk\ell}$ represent the observation on the i'th ($i=1,\ldots,n_{k\ell}$) individual for the j'th ($j=1,\ldots,4$) velocity in the k'th ($k=1,2$) sex and ℓ'th ($\ell=1,\ldots,4$) age group.

$$X_{ijk\ell} = \mu + \alpha_k + \beta_\ell + \alpha\beta_{k\ell} + \pi_{i(k\ell)} + \gamma_j + \alpha\gamma_{kj} + \beta\gamma_{\ell j} + \alpha\beta\gamma_{k\ell j} + \gamma\pi_{ji(k\ell)} + \xi_{ji(k\ell)}$$

Before carrying out the univariate repeated measures analysis of variance the required validity conditions for this analysis were assessed. Since this design contains two grouping factors both the Box (1940) M criterion and the Mauchly (1940) W criterion are needed to assess the validity conditions (see section 3.2).

Unfortunately the necessary and sufficient conditions for the univariate analysis of variance do not hold and we must use either the Greenhouse & Geisser or the Huynh & Feldt approximate test. The results from this analysis are outlined in table 4.3.3.

Table 4.3.3: Univariate Repeated Measures Anova

Source	F-value	P-value	G-G P-value	H-F P-value
Between Subjects	15442.07	0.0000*		
- Sex	0.51	0.4766		
- Age	345.02	0.0000*		
- Sex by Age	4.07	0.0088*		
Within Subjects				
- Velocity	3754.54	0.0000*	0.0000*	0.0000*
- Velocity by Sex	2.77	0.0419*	0.0661	0.0613
- Velocity by Age	6.84	0.0000*	0.0000*	0.0000*
- Velocity by Sex and by Age	2.31	0.0159*	0.0361*	0.0315*

Note: * indicates significance for sig. level = 0.05

In table 4.3.3 F-value represents the value of the test statistic calculated from the ratio of the appropriate mean squares. There are three columns in this table headed P-value, G-G P-value and H-F P-value which represent the probability of obtaining a more extreme value of the test statistic than that observed, when the null hypothesis is true, for the standard univariate F test, the Greenhouse & Geisser and the Huynh & Feldt approximate tests respectively.

From table 4.3.3 we can see that the velocity by sex by age interaction term is just significant which makes it difficult to interpret any of the other within subject effects.

Looking at the between-subjects results there is a significant sex by age interaction term which means again we cannot interpet the main effects without examining them jointly in a further analysis. For example examining the age effect for each level of the sex effect.

Multivariate Analysis of Variance

Since we have a reasonable amount of data we can make use of the multivariate approach to the analysis of repeated measures. For details of multivariate analysis of variance for repeated measures see section 3.2.

When carrying out this analysis using the BMDP program P4V several multivariate test statistics are calculated but only Roy's largest root statistic and Wilk's likelihood ratio statistic will be reported to ease the interpretation of the results. The results are outlined in table 4.3.4.

Table 4.3.4: Multivariate Analysis of Variance

Source	Test Statistic	P-value
Velocity	- Hotellings T^2	0.0000*
Velocity by Sex	- Hotellings T^2	0.0663
Velocity by Age	- Wilks L. Ratio	0.0000*
	Roy Max. Root	0.0000*
Velocity by Sex and by Age	- Wilks L. Ratio	0.0625
	Roy Max. Root	0.1618

Note: * indicates significance for sig. level = 0.05

Note that the Between-subjects tests on the grouping factors and their interactions are essentially univariate tests and hence the results will be the same as in table 4.3.3 so we have not reported them in table 4.3.4. Only the within subjects tests are reported here.

In table 4.3.4 for testing the velocity and the velocity by sex effects, the Hotellings T^2 test was carried out. This is because when we have either one or two groups, all the multivariate test statistics mentioned in section 3.2 are equivalent to the Hotelling's T^2. Under the heading 'Test Statistic' in table 4.3.4 we have identified which test statistics were calculated where

Wilks L. Ratio represents Wilks Likelihood ratio statistic, Lambda

and

Roy Max. Root represents Roy's maximum root statistic derived using the union intersection approach discussed in section 3.2.

The probability of obtaining a more extreme value than that observed for the test statistic is given in the last column i.e. the P-value.

From the results we can see that the velocity by sex by age interaction term is non-significant as is the velocity by sex interaction but the velocity by age term is very significant.

The Modelling Approach Using The Two-Stage Model

Since the measurements were taken on successively decreasing velocities and from figure 4.3.2 and individual graphs there appears to be a very strong linear relationship between the log of the peak torque and velocity, the modelling approach discussed

in section 3.3 may be applied. Here we use the two-stage model as outlined by Rao (1965-67) which is equivalent in this application, to the Potthoff & Roy model with $G=I_p$ (see section 3.3). We will use the same notation, $X_{ijk\ell}$ as was introduced at the beginning of this section for the univariate repeated measures analysis of variance.

The first stage of this approach involved obtaining the least squares regression coefficients for each subject where we assumed a linear relationship existed for each individual.

$$E(X_{ijk\ell}) = \alpha_{ik\ell} + \beta_{ik\ell} V_j$$

where V_j represents the j'th (j=1,...,4) velocity setting.

Letting $\underline{X}_{ik\ell} = (x_{i1k\ell}, x_{i2k\ell}, x_{i3k\ell}, x_{i4k\ell})$ we may write

$$E(\underline{x}_{ik\ell}) = (\alpha_{ik\ell}, \beta_{ik\ell}) \begin{bmatrix} 1 & 1 & 1 & 1 \\ V_1 & V_2 & V_3 & V_4 \end{bmatrix}$$

$$= \underline{\theta}_{ik\ell} D$$

for $i = 1,\ldots,n_{k\ell}$

$k = 1,2$

$\ell = 1,2,3,4.$

The mean and standard deviation (S.D.) of the slopes and intercepts for each group are given in table 4.3.5. The regression coefficients for each subject were calculated using the statistics package Minitab.

Table 4.3.5: Descriptive Statistics for the Regression Coefficients

Sex	Age	Intercept		Slope	
		Mean	S.D.	Mean	S.D.
Males	5-5½	1.1464	0.0890	-0.00226	0.00048
	8-8½	1.4279	0.1676	-0.00261	0.00038
	11-11½	1.7999	0.0938	-0.00271	0.00023
	14-14½	1.9878	0.0811	-0.00216	0.00021
Females	5-5½	1.1664	0.1247	-0.00252	0.00048
	8-8½	1.5280	0.0877	-0.00263	0.00038
	11-11½	1.7301	0.0810	-0.00260	0.00030
	14-14½	1.9637	0.0593	-0.00248	0.00031

The assumptions required by this approach are that each individual has a separate regression line and that for each individual, the errors about the regression line are independently distributed.

The second stage of the two stage model then assumes that for each individual the regression coefficients come from underlying normal distributions.

Essentially the analysis for this two-stage model (see section 2.2 and 3.3) consists of replacing the four peak torque measurements for each individual by their least squares regression coefficients and then applying multivariate analysis of variance to test for group effects and interactions. Table 4.3.6 contains the results from applying multivariate analysis of variance on the obtained regression coefficients.

Table 4.3.6: Two-Stage Model Results

Effect	Test Statistic	P-value
Slope	Univariate F	0.0000*
Intercept	Univariate F	0.0000*
Sex	Hotellings T^2	0.1875
- Slope	Univariate F	0.0689
- Intercept	Univariate F	0.4766
Age	Wilks L. Ratio	0.0000*
	Roy Max. Root	0.0000*
- Slope	Univariate F	0.0005*
- Intercept	Univariate F	0.0000*
Sex by Age	Wilks L. Ratio	0.0056*
	Roy Max. Root	0.0373*
- Slope	Univariate F	0.0717
- Intercept	Univariate F	0.0088*

Note: * indicates significance for sig. level = 0.05

There are three columns in this table where the first column headed 'effects' just identifies which effect is being tested, the third column represents the p-value as defined earlier and the second column identifies which test statistic is being used to test each effect where,

'Univariate F' represents the F-ratio obtained in applying univariate analysis of variance.

'Wilks L. Ratio' represents Wilks likelihood ratio test statistic and

'Roy Max. Root' represents Roy's maximum root test statistic.

We first examine the results obtained from the multivariate tests given in table 4.3.6 to get an overall view of any significant effects and then if appropriate, examine the results from the univariate tests to try and simplify the interpretation.

From table 4.3.6 we can see that there is a significant sex by age interaction term. On looking at the univariate results which are essentially treating the slope and intercept as two univariate measurements, we see that for the intercept the interaction effect is significant but for the slope it is borderline. The significance of the interaction term makes it difficult to interpret the main effects for age and sex from this table. Further analysis would have to be done, although the age effect from table 4.3.6 would appear to be very significant.

The Modelling Approach Using The Conditional Model

In the previous analysis we replaced each subjects four measurements with two regression coefficients obtained from fitting individual linear models. As noted by Rao (1965-67) this could result in the throwing away of useful information. Possibly, some of the information being thrown away may be of use, not necessarily all of it. Hence we have applied the conditional model and its analysis as proposed by Rao (1965-67) (see sections 2.2 and 3.3).

This approach involves taking each individual's slope and intercept as before and two other othogonal linear combinations of the four observations on each individual. Multivariate analysis of covariance is then applied on the subjects regression coefficients using the two other orthogonal linear combinations as covariates.

These covariates may be easily obtained from two orthogonal linear combinations of the residuals for each individual. An alternative method for obtaining them if orthogonal polynomials were being used, would be to use the cubic and quadratic orthogonal polynomials as covariates.

Since we were not using orthogonal polynomials the two other orthogonal linear combinations were derived from the residual space. Multivariate analysis of covariance was then applied using the two regression coefficients and covariates. The results are presented in table 4.3.7. The format of this table i.e. the headings and abbreviations are the same as for table 4.3.6.

Table 4.3.7: The Conditional Model Results

Source	Test	P-value
Covariates	Wilks L. Ratio	0.0582
	Roy Max. Root	0.0642
- Slope	Univariate F	0.5005
- Intercept	Univariate F	0.0179*
Grand Mean	Hotellings T^2	0.0000*
- Slope	Univariate F	0.0000*
- Intercept	Univariate F	0.0000*
Sex	Hotellings T^2	0.1353
- Slope	Univariate F	0.0480*
- Intercept	Univariate F	0.4017
Age	Wilks L. Ratio	0.0000*
	Roy Max. Root	0.0000*
- Slope	Univariate F	0.0030*
- Intercept	Univariate F	0.0000*
Sex x Age	Wilks L. Ratio	0.0041*
	Roy Max. Root	0.0279*
- Slope	Univariate F	0.0750
- Intercept	Univariate F	0.0066*
* indicates significance for $\alpha = 0.05$		

As with table 4.3.6 we examine the results from the multivariate tests first.

From table 4.3.7 the sex by age interaction term is significant. Examining the slope and intercept separately using univariate analysis of covariance, we see a significant interaction effect for the intercepts but borderline for the slopes.

Due to the significant interaction term it is again difficult to interpret the main effects for age and sex, although there does seem to be a very significant result for the age effect.

It is worthwhile noting that the test on the need for the covariates is not significant when the two variables are looked at jointly but that the univariate test when considering intercept alone is significant.

4.3.3: Summary

For this problem four different methods of analysis were used. The results obtained from these different methods being very similar with only a few contrasting results. We will compare the univariate and multivariate procedures first, then the two-stage model analysis against the conditional model and finally all four procedures.

Comparing the univariate analysis to the multivariate, the main discrepancy lay in whether the sex by age interaction term for the within subjects effects was significant. This interaction term was found to be significant for both of the approximate tests used in the univariate analysis but not for the multivariate analysis. All other results were comparable. It should be noted here that the tests used in the univariate analysis are approximate, and the multivariate analysis will give more exact results and is more powerful when the univariate assumptions do not hold. In contrast, of course, it will be less

powerful when the univariate assumptions do hold.

Comparing the two-stage model analysis with that for the conditional model the results were comparable. Since in both cases, the sex by age interaction term was found to be significant and in looking at the univariate analysis for the slopes and intercepts separately, only the interaction term for the intercepts was found to be significant. Further conclusions are more difficult to arrive at due to the difficulties in interpretation arising from the significance of the interaction term. This may indicate that the two covariates included in the analysis using the conditional model provide little useful additional information since the results are similar (though see the note in the previous section). Hence the two-stage model which is equivalent to the Potthoff & Roy method with $G=I_p$ appears to be a reasonable approach although we might have some misgivings about the need for a covariate for the slope.

Comparing all four procedures the only discrepancy is in the significance of the sex by age interaction term for the within subjects effects. Given that this effect requires different assumptions depending on the procedure used this might be expected unless there is a very clear cut effect. For this problem, there does not appear to be any obvious advantage to using either the analysis of variance approach or the modelling approach except in the practicalities of carrying out the analysis and interpretation of the results for non-statisticians, though it would seem inefficient not to use any relationships that existed in the data.

4.4: Fundal Height and Growth Retardation

4.4.1: Introduction

The accurate identification of the growth retarded foetus remains a problem in spite of a wide range of clinical, biochemical and ultrasonographic techniques now available.

Ultrasonographic measurements including serial cephalometry, the crown-rump length and trunk area are the most sensitive techniques for diagnosing IUGR (Intra Uterine Growth Retardation) available, but the expertise involved in their application precludes their use as a widespread screening test outside of teaching centres.

The need remains for a sensitive screening procedure for growth retardation which can be applied easily and at low cost in the course of routine ante-natal care.

A study was carried out to examine the efficacy of the formal measurement of fundal height in the identification of growth retardation.

During 1978 and 1979 measurements of fundal height were made routinely from twenty weeks gestation until delivery on all women attending one of Bellshill Maternity Hospital's peripheral antenatal clinics. The method of measuring the fundal height was taught to junior medical staff and midwives.

Since the data was being collected routinely it was not possible to have the same person recording a patient's measurements throughout the pregnancy and hence patients were likely to have their fundal height measured by a different person at each visit. The case notes of all women attending the clinic during 1978 and 1979 were examined retrospectively. Women whose gestation was known with certainty, either on the basis of a careful menstrual history or an early scan (less than 26 weeks)

were selected for the study sample. Women with uncertain gestation were excluded from the study, as were all cases of multiple pregnancy. For further details see Rosenberg et al (1982).

In total there were 761 women included in the study. The values of fundal height, as well as basic information such as age, height, parity, smoking habits and previous obstetric history were recorded for these women. In the sample of 761 women, 51 babies were born who were growth retarded. The definition of growth retardation used being a weight less than the tenth centile for gestation according to the standards of Thomson et al (1968).

Unfortunately the number of visits to the clinic and hence the number of fundal height measurements were not the same for all of the women in the sample. Table 4.4.1 shows the number of visits to the clinic for the two groups (Growth-retarded and Normal). Only the measurements obtained up to 36 weeks into the pregnancy are used since this was identified as a gestation at which suspicion of growth-retardation would permit effective clinical surveillance and allow for necessary intervention.

Table 4.4.1: Number of visits up to 36'th week

	Number of Visits to Clinic														
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	All
G.R.	0	0	3	5	5	13	14	6	3	1	0	0	0	1	51
Norm	1	2	29	56	120	150	150	117	55	20	6	4	0	0	710
All	1	2	32	61	125	163	164	123	58	21	6	4	0	1	761

4.4.2: Analysis of the Data

Essentially this is a discrimination problem where using the available information - fundal height, age, smoking habits etc. we are trying to correctly identify a woman who is likely to have a growth-retarded baby. Obviously in identifying these women correctly we may misclassify some women who are likely to have a normal weight baby. A high number of misclassified normals may be expensive in costs to the hospital and hence must be minimised.

One of the main objectives of this study was to examine the usefulness of fundal height measurements in the identification of growth-retarded babies. Hence for each mother we have a set of growth curve data ie. the fundal height measurements at various time points during the pregnancy. Unfortunately the data is extremely unstructured since the mothers do not all have the same number of visits and the timing of the visits could not be controlled.

There are several other difficulties with this problem including the crudeness of the measurement of fundal height and the variability in the measurements caused by different examiners in the clinic.

Since we had a very complex set of growth curve data plots of fundal height against gestation (weeks into pregnancy) were obtained for a large sample of the subjects. The reasoning behind this was to see if there was any relationship between fundal height and gestation which would allow us to model the data and hence simplify the structure. From the plots it appeared that a roughly linear relationship existed. Hence the least squares regression line for each individual was obtained and the complex set of measurements was replaced by the intercept

and slope for each regression line.

The means and standard deviations of the intercept and slope for the two groups are given in table 4.4.2.

Table 4.4.2: Descriptive Statistics for the Intercept and Slope

Mean ± St. Dev.

	Intercept	Slope
GR. (N = 47)	4.973 ± 6.820	0.741 ± 0.235
Norm. (N = 671)	1.964 ± 9.570	0.908 ± 0.306

Having simplified the complexity of the fundal height measurements we then used the values for the intercept and slope and the other covariates for each individual in a linear discriminant analysis using the stepwise linear discriminant program P7M of BMDP (1983). The covariates included measurements on the mothers such as age, height, number of cigarettes smoked per day and number of previous growth retarded babies. The results are given in table 4.4.3.

In carrying out this linear discriminant analysis equal prior probabilities were used. The reason for this was that when one used the natural prior probabilities 0.1 and 0.9 there was an unacceptably high number of growth retardeds classified as normals. Indeed very few of the growth retardeds were correctly identified. Prior probabilities of 0.5 and 0.5 gave a more acceptable balance of correctly identified growth retardeds and normals. When running this program, P7M in stepwise mode, all other parameters to control the stepping and selecting of variables were left at the default values. The obtained jack-knifed classifications are presented in table 4.4.3 for several fitted models. The features used and the order in which

they were entered in the program is given in table 4.4.3 under the heading 'features used'.

Table 4.4.3: Results Using Linear Discriminant Analysis

Prior Probabilities: 0.5, 0.5

: Jack-knifed Classifications

Features Used	% Growth Retarded Correctly Identified	% Normal Correctly Identified
Least squares slope and intercept.	70.5 (31)	73.3 (462)
Height	68.2 (30)	62.9 (396)
Smoking Habits	68.2 (30)	64.6 (407)
Height, smoking	72.7 (32)	68.3 (430)
Height, least squares slope and intercept.	70.5 (31)	76.2 (480)
Least squares slope, intercept and smoking habits.	72.7 (32)	75.7 (477)
Height, least squares slope, intercept and smoking	72.7 (32)	76.7 (483)
GOR, height, least squares slope, intercept and age.	68.2 (30)	80.2 (505)

Notes: The figures in brackets represent the number of cases correctly identified.

GOR represents the number of previous growth retarded babies.

4.4.3: Discussion

As can be seen from table 4.3.3 the group of variables which best discriminated between growth retarded and normal weight babies were the height of the mother, the fundal height slope and intercept and the number of cigarettes smoked per day by the mother. This group of variables correctly identifying 32/44 growth retardeds and 483/630 normals.

It is worthy of note that all of the models provide similar results in the number of growth retardeds correctly identified irrespective of whether the fundal height slope and intercept are used. Although height and smoking do quite well on their own, there does seem some advantage in using the slope and intercept as well.

The above approach to this problem consists of a simple procedure for using a very complex set of growth curve data in a discrimination problem. This procedure involved a potentially inefficient use of the two-stage model (see section 2.3). This was because all of the parameter estimates were treated equally, irrespective of the number of visits and the timing of these visits which affects the accuracy of estimation of parameters for different individuals. Thus the assumption for the second stage of the model where we assume the parameters arise from the same distribution is strictly speaking not tenable.

A more complex approach and hence analysis may have given an improved discrimination between the two groups but this would seem unlikely. In practice the use of the fundal height slope and intercept in a simple discriminant function would provide a practically useful technique which can be applied by junior medics and other staff.

CHAPTER 5

RECENT DEVELOPMENTS

Chapter 5: Recent Developments

In this chapter we summarise, briefly, some additional developments and more recent contributions to the design and analysis of studies involving repeated measures data. Some of these references may not necessarily have been mentioned previously.

Kunert (1983, 1984, 1985) presents further findings on the optimality of balanced uniform repeated measurement designs which was previously examined by Hedayat and Afsarinejad (1978), Cheng and Wu (1980) and Afsarinejad (1983).

Hearne et al (1983) examine the robustness of univariate analysis of variance when the covariance matrices for the data have serial correlation. They also present a likelihood ratio test for testing for patterns in the covariance matrices as have Chinchilli and Carter (1984). Kenny and Judd (1986) have very recently presented a comprehensive discussion of the consequences of violating the assumptions required in univariate analysis of variance.

Mitzel and Games (1981) examined the assumptions required in carrying out multiple comparison tests in repeated measures designs. Foutz (1985) using the results of Zerbe & Walker (1977) and Zerbe (1979) also examines multiple comparison procedures, in the randomisation analysis of growth curve responses.

Verbyla (1986) and Kenward (1985) continue examination of the use of covariates in the growth curve model as previously discussed by Rao (1965-67). In particular Kenward (1985) examines the use of fitted higher-order polynomial coefficients as covariates.

Liang and Zeger (1986) and Ware (1985) contribute further work on the analysis of repeated measures data using generalised linear models.

Both Hui et al (1983) and Strenio et al (1983) use empirical Bayes estimation techniques, Hui et al (1983) in the estimation of rates of change in longitudinal studies and Strenio et al (1983) in the estimation of growth-curve parameters. Hui follows the same approach as Fearn (1975) and Laird & Ware (1982).

Katz and McSweeney (1983) develop some non-parametric tests for analysing ranked data in repeated measures designs and Gasser et al (1984) compare parametric and non-parametric regression analysis of growth curves.

Woolson and Clark (1984), Wei and Johnson (1985) and Crépeau et al (1985) discuss methods of analysing incomplete data from repeated measurements experiments. Woolson and Clark (1984) examine categorical incomplete longitudinal data and use a simple modification of Grizzle et al (1969)'s methodology. Crépeau et al (1985) discuss incomplete data where the missing observations always arise at the end of the data. Hui (1984) discusses curve fitting for repeated measurements made at irregular time-points.

Zeger et al (1985) examine the analysis of binary longitudinal data with time-independent covariates.

Various models and methods of analysis for repeated measures data are also discussed in Plewis (1985). In particular Plewis discusses models for categorical repeated measures data.

REFERENCES

AFSARINEJAD, K. (1983). Balanced repeated measurements designs. Biometrika, 70, 199-204.

ANDERSON, T.W. (1984). An Introduction to Multivariate Statistical Analysis. John Wiley & Sons.

ARNOLD, S.F. (1981). The Theory of Linear Models and Multivariate Analysis. Wiley: New York.

BAKSALARY, J.K., CORSTEN, L.C.A. and KALA, R. (1978). Reconciliation of two different views on estimation of growth curve parameters. Biometrika, 65, 662-665.

BARTLETT, M.S. (1937). The statistical conception of mental factors. British Journal of Psychology, 28, 97-104.

BEALE, E.M.L. and LITTLE, R.J.A. (1975). Missing values in multivariate analysis. Journal of the Royal Statistical Society, B, 37, 129-145.

BEAUCHAMP, J.J. and HOEL, D.G. (1974). Some investigation and simulation studies of economical analysis. J. Statist. Comp. and Simulation, 2, 197-209.

BOCK, R.D. (1963). Multivariate Analysis of Variance of repeated measurements. In Harris, C.W. (Ed.) Problems in Measuring Change. Univ. of Wisconsin Press, Madison.

BOCK, R.D. (1975). Multivariate Statistical Methods in Behavioural Research. McGraw-Hill.

BOWDEN, D.C. and STEINHORST, R.K. (1973). Tolerance bands for growth curves. Biometrics, 29, 361-371.

BOX, G.E.P. (1949). A general distribution theory for a class of likelihood criteria. Biometrika, 36, 317-346.

BOX, G.E.P. (1950). Problems in the analysis of growth and wear curves. Biometrics, 6, 362-389.

BOX, G.E.P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems. The Annals of Mathematical Statistics, 25, 290-302, 484-498.

CHATTERJEE, S.K. and SEN, P.K. (1966). Nonparametric tests for the Multisample Multivariate Location problem. In Bose, R.C., Chakravorti, I.M., Mahalanobis, B.C., Rao, C.R. and Smith, K.J.C. Essays in Probability and Statistics, 197-228 University of North Carolina Press.

CHENG, C.S. and WU, C.F. (1980). Balanced repeated measurements designs. Annals of Statistics, 8, 1272-1283.

CHINCHILLI, V.M. and CARTER, W.H., Jrs. (1984). A likelihood ratio test for a patterned covariance matrix in a multivariate growth-curve model. Biometrics, 40, 151-156.

COLE, J.W.L. and GRIZZLE, J.E. (1966). Applications of multivariate analysis of variance to repeated measurements experiments. Biometrics, 22, 810-828.

COLLIER, R.O., Jr., BAKER, F.B., MANDEVILLE, G.K. and HAYES, T.F. (1967). Estimates of test size for several test procedures based on variance ratios in the repeated measures design. Psychometrika, 32, 339-353.

CREPEAU, H., KOZIOL, J., REID, N. and YUH, Y.S. (1985). Analysis of incomplete multivariate data from repeated measurement experiments. Biometrics, 41, 505-514.

DARBY, S.C. and FEARN, T. (1979). The Chatham blood pressure study. An application of Bayesian growth curve models to a longitudinal study of blood pressure in children. International Journal of Epidemiology, 8, 15-21.

DAVIDSON, M.L. (1980). Univariate versus multivariate tests in repeated-measures experiments. Psychological Bulletin, 77, 446-452.

DAVIDSON, M.L. (1975). The multivariate approach to repeated measures. BMDP Statistical Software. Dept. of Biomathematics, Univ. of California, Los Angeles. Technical Report No.75.

DEMPSTER, A.P., LAIRD, N.M. and RUBIN, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). Journal of the Royal Statistical Society, B, 39, 1-38.

DIXON, W.J. (chief editor) (1983). BMDP Statistical Software. Berkeley: University of California Press.

DUNSMORE, I.R. (1981). Growth curves in two-period change over models. Applied Statistics, 30, 223-229.

ELSTON, R.C. and GRIZZLE, J.E. (1962). Estimation of time-response curves and their confidence bands. Biometrics, 18, 148-159.

FEARN, T. (1975). A Bayesian approach to growth curves. Biometrika, 62, 89-100.

FEARN, T. (1977). A two-stage model for growth curves which leads to Rao's covariance adjusted estimators. Biometrika, 64, 141-143.

FOUTZ, R.V., JENSEN, D.R. and ANDERSON, G.W. (1985). Multiple comparisons in the randomisation analysis of designed experiments with growth curve responses. Biometrics, 41, 29-37.

FRANE, J.W. (1980). The univariate approach to repeated measures. BMDP Statistical Software, Dept. of Biomathematics, University of California, Los Angeles, Technical Report No.69.

GAFARIAN, A.V. (1978). Confidence bands in multivariate polynomial regression. Technometrics, 20, 141-149.

GASSER, T., MULLER, H-G. and KOHLER, W. (1984). Non-parametric regression analysis of growth curves. Annals of Statistics, 12, 210-229.

GEISSER, S. and GREENHOUSE, S.W. (1958). An extension of Box's results on the use of the F-distribution in multivariate analysis. The Annals of Mathematical Statistics, 29, 885-891.

GEISSER, S. (1970). Bayesian analysis of growth curves. Sankhyā: The Indian Journal of Statistics: Series A, 53-64.

GEISSER, S. (1971). The inferential use of predictive distributions. In Foundations of Statistical Inference. Holt, Rinehart, and Winston, 456-469.

GHOSH, M., GRIZZLE, J.E. and SEN, P.K. (1973). Nonparametric methods in longitudinal studies. Journal of the American Statistical Association, 68, 29-36.

GOLDSTEIN, H. (1979). The Design and Analysis of Longitudinal Studies. Academic Press.

GREENHOUSE, S.W. and GEISSER, S. (1959). On the methods in the analysis of profile data. Psychometrika, 24, 95-112.

GRIZZLE, J.E. and ALLEN, D.M. (1969). Analysis of growth and dose response curves. Biometrics, 25, 357-381.

HARVILLE, D.A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. Journal of the American Statistical Association, 72, 320-340.

HEARNE, E.M., CLARK, G.M. and HATCH, J.P. (1983). A test for serial correlation in univariate repeated-measures analysis. Biometrics, 39, 237-243.

HEDAYAT, A. and AFSARINEJAD, K. (1975). Repeated measurements designs, I. In J.N. Srivastava (Ed.), A Survey of Statistical Design and Linear Models, 229-242. North-Holland, Amsterdam.

HEDAYAT, A. and AFSARINEJAD, K. (1978). Repeated measurements designs, II. The Annals of Statistics, 6, 619-628.

HOEL, P.G. (1964). Methods for comparing growth type curves. Biometrics, 20, 859-872.

HOPKINS, J.W. and CLAY, P.P.F. (1963). Some empirical distributions of bivariate T^2 and homoscedasticity criterion M under unequal variance and leptokurtosis. Journal of the American Statistical Association, 58, 1048-1053.

HORTON, R.L. (1978). The General Linear Model. McGraw Hill.

HOTELLING, H. (1947). Multivariate quality control, illustrated by the air testing of sample bombsights. In Techniques of Statistical Analysis. Eisenhart, C., Hastay, M. and Wallis, W.A. (Eds.) McGraw-Hill, 111-184.

HOTELLING, H. (1951). A generalised T test and measure of multivariate dispersion. Proceedings of the second Berkeley Symposium on Mathematical Statistics and Probability. Neyman, J. ed., 23-41.

HUI, S.L. and BERGER, J.O. (1983). Empirical Bayes estimation of rates in longitudinal studies. Journal of the American Statistical Association, 78, 753-760.

HUI, S.L. (1984). Curve fitting for repeated measurements made at irregular time-points. Biometrics, 40, 691-697.

HUYNH, H. and FELDT, L.S. (1970). Conditions under which mean square ratios in repeated measurements designs have exact F-distributions. Journal of the American Statistical Association, 65, 1582-1589.

HUYNH, H. and FELDT, L.S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomised block and split-plot designs. Journal of Educational Statistics, 1, 69-82.

HUYNH, H. (1978). Some approximate tests for repeated measurement designs. Psychometrika, 43, 161-175.

HUYNH, H. and MANDEVILLE, G.K. (1979). Validity conditions in repeated measures designs. Psychological Bulletin, 86, 964-973.

JENNRICH, R.I. (1977). Letting BMDP8V tell us something about randomised blocks repeated measures, and split-plots. BMDP Statistical Software, Dept. of Biomathematics, Univ. of California, Los Angeles, Technical Report No.34.

JORESKOG, K.G. (1970). A general method for analysis of covariance structures. Biometrika, 57, 239-251.

JORESKOG, K.G. (1973). Analysis of covariance structures. In P.R. Krishnaiah (Ed.), Multivariate Analysis, Vol III, 263-285. Academic Press.

JORESKOG, K.G. (1978). Structural analysis of covariance and correlation matrices. Psychometrika, 43, 443-477.

KATZ, B.M. and McSWEENEY, M. (1983). Some non-parametric tests for analysing ranked data in multi group repeated measures designs. British Journal of Mathematical and Statistical Psychology, 36, 145-156.

KENNY, D.A. and JUDD, C.M. (1986). Consequences of violating the independence assumption in analysis of variance. Psychological Bulletin, 99, 422-431.

KENWARD, M.G. (1985). The use of fitted higher-order polynomial coefficients as covariates in the analysis of growth curves. Biometrics, 41, 19-28.

KESELMAN, H.J., ROGAN, J.C., MENDOZA, J.L. and BREEN, L.J. (1980). Testing the validity conditions of repeated measures F tests. Psychological Bulletin, 87, 479-481.

KHATRI, C.G. (1966). A note on a manova model applied to problems in growth curve. Annals of the Institute of Statistical Mathematics, 18, 75-86.

KLEINBAUM, D.G. (1973). A generalisation of the growth curve model which allows missing data. Journal of Multivariate Analysis, 3, 117-124.

KOCH, G.G., LANDIS, J.R., FREEMAN, J.L., FREEMAN, D.H. Jr. and LEHNEN, R.G. (1977). A general methodology for the analysis of experiments with repeated measurement of categorical data. Biometrics, 33, 133-158.

KOCH, G.G., AMARA, I.A., STOKES, M.E. and GILLINGS, D.B. (1980). Some views on parametric and non-parametric analysis for repeated measurements and selected bibliography. International Statistical Review, 48, 249-265.

KORIN, B.P. (1972). Some comments on the homoscedasticity criterion M and the multivariate analysis of variance tests T^2, W and R. Biometrika, 59, 215-216.

KOWALSKI, C.J. and GUIRE, K.E. (1974). Longitudinal data analysis. Growth, 38, 131-169.

KOZIOL, J.A., MAXWELL, D.A., FUKUSHIMA, M., COLMERAUER, M.E.M. and PILCH, Y.H. (1981). A distribution-free test for tumour-growth curve analyses with application to an animal tumour immunotherapy experiment. Biometrics, 37, 383-390.

KUNERT, J. (1983). Optimal design and refinement of the linear model with applications to repeated measurements designs. Annals of Statistics, 11, 247-257.

KUNERT, J. (1984). Optimality of balanced uniform repeated measurements designs. Annals of Statistics, 12, 1006-1017.

KUNERT, J. (1985). Optimal repeated measurements designs for correlated observations and analysis by weighted least squares. Biometrika, 72, 375-389.

LAIRD, N.M. and WARE, J.H. (1982). Random-effects models for longitudinal data. Biometrics, 38. 963-974.

LAWLEY, D.N. (1938). A generalisation of Fisher's Z-test. Biometrika, 30, 180-187.

LEE. Y.-H.K. (1974). A note on Rao's reduction of Potthoff & Roy's generalised linear model. Biometrika, 61, 349-351.

LEE, J.C. and GEISSER, S. (1972). Growth curve prediction. Sankhyā: The Indian Journal of Statistics: Series A, 34, 393-412.

LEECH, F.B. and HEALY, M.J.R. (1959). The analysis of experiments on growth rate. Biometrics, 15, 98-106.

LIANG, K.-Y. and ZEGER, S.L. (1986). Longitudinal data analysis using generalised linear models. Biometrika, 73, 13-22.

LINDMAN, H.R. (1974). Analysis of Variance in Complex Experimental Designs. W.H. Freeman and Co.

McCAMMON, R.W. (1970). Human Growth and Development. Thomas, Springfield.

MACDOUGALL, J.D., WENGER, H.A. and GREEN, H.J. (1982). Physiological Testing of the Elite Athlete. Mutual Press Ltd.

MACHIN, D. (1975). On a design problem in growth studies. Biometrics, 31, 749-753.

MAGER, P.P. (1973). Repeated measurement designs. Activitas nervosa superior, 15, 269-277.

MARDIA, K.V., KENT, J.T. and BIBBY, J.M. (1979). Multivariate Analysis. Academic Press.

MAUCHLY, J.W. (1940). Significance test for sphericity of a normal n-variate distribution. The Annals of Mathematical Statistics, 29, 204-209.

MENDOZA, J.L., TOOTHAKER, L.E. and NICEWANDER, W.A. (1974). A Monte Carlo comparison of the univariate and multivariate methods for the groups by trials repeated measures design. Multivariate Behavioural Research, 9, 165-178.

MENDOZA, J.L., TOOTHAKER, L.E. and CRAIN, B.R. (1976). Necessary and sufficient conditions for F-ratios in the L X J X K factorial design with two repeated factors. Journal of the American Statistical Association, 71, 992-993.

MITZEL, H.C. and GAMES, P.A. (1981). Circularity and multiple comparisons in repeated measure designs. British Journal of Mathematical and Statistical Psychology, 34, 253-259.

MONLEZUN, C.J., BLOUIN, D.C. and MALONE, L.C. (1984). Constrasting split plot and repeated measures experiments and analyses. The American Statistician, 38, 21-27.

MORRISON, D.F. (1976). Multivariate Statistical Methods. McGraw-Hill.

OLSON, C.L. (1974). Comparative robustness of six tests in multivariate analysis of variance. Journal of the American Statistical Association, 69, 894-908.

ORCHARD, T. and WOODBURY, M.A. (1972). A missing information principle: theory and applications. In Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, 1. Le Cam, L.M., Neyman, J. and Scott, E.L. (eds.), 697-715. University of California Press.

PILLAI, K.C.S. (1955). Some new test criteria in multivariate analysis. The Annals of Mathematical Statistics, 26. 117-121.

PLEWIS, I. (1981). A comparison of approaches to the analysis of longitudinal categoric data. British Journal of Mathematical and Statistical Psychology, 34, 118-123.

PLEWIS, I. (1985). Analysing Change: measurement and explanation using longitudinal data. John Wiley & Sons.

POTTHOFF, R.F. and ROY, S.N. (1964). A generalised multivariate analysis of variance model useful especially for growth curve problems. Biometrika, 51, 313-326.

RAO, C.R. (1958). Some statistical methods for comparison of growth curves. Biometrics, 14, 1-17.

RAO, C.R. (1959). Some problems involving linear hypotheses in multivariate analysis. Biometrika, 46, 49-58.

RAO, C.R. (1965). The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. Biometrika, 52, 447-458.

RAO, C.R. (1966). Covariance adjustment and related problems in multivariate analysis. In Krishnaiah, P.R. (Ed.), Multivariate Analysis, Vol. II, pp 87-103, Academic Press.

RAO, C.R. (1967). Least squares theory using an estimated dispersion matrix and its application to measurement of signals. In the Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1, 355-372.

RAO, C.R. (1972). Recent trends of research work in multivariate analysis. Biometrics, 28, 3-22.

REINSEL, G. (1982). Multivariate repeated-measurement or growth curve models with multivariate random-effects covariance structure. Journal of the American Statistical Association, 77, 190-195.

REINSEL, G. (1984). Estimation and prediction in a multivariate random effects generalised linear model. Journal of the American Statistical Association, 79, 406-414.

ROGAN, J.C., KESELMAN, H.J. and MENDOZA, J.L. (1979). Analysis of repeated measurements. British Journal of Mathematical and Statistical Psychology, 32, 269-286.

ROSENBERG, B. (1973). Linear regression with randomly dispersed parameters. Biometrika, 60, 65-72.

ROSENBERG, K., GRANT, J.M., AITCHISON, T. and GALLAGHER, F. (1982). Measurement of fundal height as a screening test for fetal growth retardation. British Journal of Obstetrics and Gynaecology, 89, 447-450.

ROUANET, H. and LEPINE, D. (1970). Comparison between treatments in a repeated-meaures design: ANOVA and multivariate methods. The British Journal of Mathematical and Statistical Psychology, 23, 147-163.

ROY, J. (1958). Step-down procedure in multivariate analysis. The Annals of Mathematical Statistics, 29, 1177-1187.

ROY, S.N. (1953). On a heuristic method of test construction and its use in multivariate analysis. The Annals of Mathematical Statistics, 24, 220-238.

ROY, S.N., GNANADESIKAN, R. and SRIVASTAVA, J.N. (1971). Analysis and Design of Certain Quantitative Multiresponse Experiments. Pergamon Press.

RYAN, B.F., JOINER, B.L., RYAN, T.A. (1985). Minitab Handbook (2nd edition). Duxbury Press.

SCHWERTMAN, N.C., FRIDSHAL, D. and MAGREY, J.M. (1981). On the analysis of incomplete growth curve data, a Monte Carlo study of the two nonparametric procedures. Communications in Statistics, B, 10, 51-66.

SEBER, G.A.F. (1984). Multivariate Observations. John Wiley & Sons.

SMITH, A.F.M. (1973). A general Bayesian linear model. Journal of the Royal Statistical Society, B. 35, 67-75.

SRIVASTAVA, J.N. and MCDONALD, L.L. (1974). Analysis of growth curves under the hierarchical models. Sankhyā : The Indian Journal of Statistics: Series A, 36, 251-260.

STEEL, R.G.D. and TORRIE, J.H. (1980). Principles and Procedures of Statistics A Biometrical Approach (2nd edition). McGraw-Hill.

STEVENS, J.P. (1980). Power of the multivariate analysis of variance tests. Psychological Bulletin, 88, 728-737.

STRENIO, J.F., WEISBERG, H.I. and BRYK, A. (1983). Empirical Bayes estimation of individual growth-curve parameters and their relationship to covariates. Biometrics, 39, 71-86.

THOMSON, A.M., BILLEWICZ, W.Z. and HYTTEN, F.E. (1968). The assessment of fetal growth. British Journal of Obstetrics and Gynaecology, 75, 903-916.

TUBBS, J.D., LEWIS, T.O. and DURAN, B.S. (1975). A note on the analysis of the manova model and its application to growth curves. Communications in Statistics, A, 4, 643-653.

VERBYLA, A.P. (1986). Conditioning in the growth curve model. Biometrika, 73, 475-483.

WALD, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. Transactions of the American Mathematical Society, 54, 426-482.

WARE, J.H. (1985). Linear models for the analysis of longitudinal studies. The American Statistician, 39, 95-101.

WEI, L.J. and JOHNSON, W.E. (1985). Combining dependent tests with incomplete repeated measurements. Biometrika, 72, 359-364.

WILKS, S.S. (1932). Certain generalisations in the analysis of variance. Biometrika, 24, 471-494.

WINER, B.J. (1971). Statistical Principles in Experimental Design (Second Edition). McGraw-Hill.

WISHART, J. (1938). Growth-rate determinations in nutrition studies with the bacon pig, and their analysis. Biometrika, 30, 16-28.

WOOLSON, R.F., LEEPER, J.D. and CLARKE, W.R. (1978). Analysis of incomplete data from longitudinal and mixed longitudinal studies. Journal of the Royal Statistical Society, A, 141, 242-252.

WOOLSON, R.F. and LEEPER, J.D. (1980). Growth curve analysis of complete and incomplete longitudinal data. Communications in Statistics, A, 9, 1491-1513.

WOOLSON, R.F. and CLARKE, W.R. (1984). Analysis of categorical incomplete longitudinal data. Journal of the Royal Statistical Society, A, 147, 87-99.

ZEGER, S.L., LAING, K.-Y. and STEVEN, G.S. (1985). The analysis of binary longitudinal data with time-independent covariates. Biometrika, 72, 31-38.

ZERBE, G.O. and WALKER, S.H. (1977). A randomization test for comparison of groups of growth curves with different polynomial design matrices. Biometrics, 33, 653-657.

ZERBE, G.O. (1979). Randomization analysis of the completely randomised design extended to growth and response curves. Journal of the American Statistical Association, 74, 215-221.