Chapter 9

Crossover Designs

9.1 Introduction

Most of the repeated measures designs we used in the previous chapters are sometimes referred to as parallel groups designs in which the subjects in each treatment group receive the same treatment over time. When there is substantial variability between subjects in such a study, it may be very difficult to test whether or not the differences in response means found between the groups are due to the differences in the treatment effects or due to the variability between subjects. In the absence of period effects, the among subject variation is usually much larger than the within subject variation during the period of the experiment. In marketing experiments, where the experimental units are stores selling a certain product, the among unit variation could be extremely large, depending on the variation in store size. In Chapter 7 we briefly addressed this issue by taking a set of historical store sales data to reduce the among unit variation. In a crossover experiment involving two treatments, each subject receives both treatments. This enables tests of the difference in treatment effects using the within subject variability.

In general, crossover designs is a special class of repeated measures designs in which all or some of the subjects receive different treatments in different time periods. For example, in a comparison of a treatment against a placebo with two groups of subjects, the first group could receive the placebo in the first period and the actual treatment in the second period while the second group receiving the treatment in the first period and the placebo in the second period.

With its long history [cf. Federer (1955) and Grizzle (1965)], crossover designs are becoming one of the most popular designs in biomedical experiments as they enable a better control of subject variation. Although the idea of using individual subjects as their own control is an appealing one, analysis of data from crossover designs could be very difficult. Moreover, some alternative models and terminology used in the literature could be confusing and some published results reported in the literature have been found to be erroneous. The terminologies

used in the literature, such as the 'sequence,' 'group,' and 'sequence group,' refer to the same thing. Adding to the confusion, some authors even use alternative terminologies such as *sequence* $\times$ *treatment interaction* in place of the *period* or introduce interaction terms that are not identifiable. In a crossover study, we need to model at least the *treatment effects*, *carryover effects* , and *period effects* . As in other repeated measures experiments, it is important to model the period effect as there might be a trend in the response variable affecting the experiment as a whole. Modeling group effects and the *treatment* $\times$ *period interaction* are also desirable when the design permits us to do so.

Simple designs and models based on unreasonable assumptions could lead to erroneous conclusions. Except for very simple designs that rely on too many assumptions, the solutions available in the literature are asymptotic methods or some other approximate methods. Here we will discuss the underlying problem involving just two treatments and two time periods, and we will demonstrate how the generalized approach could help tackle inference problems in this field. The readers interested in crossover designs involving a number of treatments and periods are referred to Jones and Kenward (1989), Vonesh and Chinchilli (1997), and Senn (2002). Despite much research done in this area, however, the crossover designs remains an area requiring further research to develop methods that do not require very large samples or unreasonable assumptions. The generalized approach provides a promising approach to developing exact methods in such situations.

9.2 Two-Sequence Design

Consider the problem of comparing two treatments A and B in the setting of a crossover design with just two periods. The simplest crossover design is the one with just two sequences AB and BA in which one group of subjects receives Treatment A in the first period and then Treatment B in the second period, while the second group receives Treatment B in the first period and Treatment A in the second period. It is assumed that the subjects are assigned at random to each group to minimize the group effect. The design is further illustrated by the table below.

Group	Period 1	Period 2
1	A	B
2	B	A

An example of a data set from a clinical trial discussed by Senn (2002) is reproduced in Table 9.1. In that experiment, the response variable of interest was the peak expiratory flow (PEF), a measure of lung function, made on 13 children with moderate or severe asthma. The objective of the experiment is to compare two treatments, a single inhaled dose of 200 μg Salbutamol (S) and 12 μg Formoterol (F). In this experiment, a *washout period* of 1 day was included between the two periods of crossover experiment.

Table 9.1: Peak expiratory flow (in liters per minute) measured 8 hours after treatment

Patient #	Sequence	Period 1	Period 2
01	FS	310	270
04	FS	310	260
06	FS	370	300
07	FS	410	390
10	FS	250	210
11	FS	380	350
14	FS	330	365
02	SF	370	385
03	SF	310	400
05	SF	380	410
09	SF	290	320
12	SF	260	340
13	SF	90	220

Suppose that out of a total of $n = n_1 + n_2$ subjects, n_1 subjects are randomly allocated to Group 1 to receive the treatment sequence AB and n_2 subjects are randomly allocated to Group 2 to receive the treatment sequence BA. In the above example, $AB = FS$, $BA = SF$, $n_1 = 7$ and $n_2 = 6$. Let $Y^{(x)}_{ij(i)t}$ denote the observation (or average of observations) taken from $j(i)$th subject in ith group at period t. Assuming only the first order effects, let μ_A, μ_B be the treatment means, and let λ_A, λ_B be their carryover effects from Period 1 to 2. We also need to model sequence and period effects. Let π be the period effect representing the overall trend from Period 1 to Period 2. Note that we are already having a problem with the current design. Although less important, the design allows us to include a parameter γ to represent a possible sequence effect. As clear from the table of expected response means shown below, even without any interaction effects, we already have six parameters to tackle based on the sample means of response data from the four cells.

Group	Sequence	Period 1	Period 2
1	AB	$\mu_A + \gamma$	$\mu_B + \gamma + \pi + \lambda_A$
2	BA	$\mu_B - \gamma$	$\mu_A - \gamma + \pi + \lambda_B$

In comparing the two treatments A and B, the difference in means $\delta = \mu_A - \mu_B$ is usually the parameter of primary importance. Without further assumptions, the current design does not allow us to estimate even δ, the parameter of interest. So assume that $\lambda_A = \lambda_B = \lambda$, an assumption that we will relax later in this chapter. In many applications this assumption might not be very reasonable except when there is no carryover effect. In other words $\lambda_A = \lambda_B = 0$ is perhaps the only sensible case of the assumption. Therefore, one may try to make the carryover effects small by having a washout period between the two periods in which the treatments are administered. A washout

period between the two treatments might minimize the effects of the carryover, but this will not be feasible in experiments involving treatments that can affect the response for a long period or permanently (e.g. education). The idea of having a washout period could also introduce additional complications in conducting the experiment. It also does not resolve the problem of different period effects that exist in many applications.

To be more specific, assume the linear model

$$Y^{(x)}_{ij(i)t} = \mu_x + \mathbf{I}\rho + \mathbf{J}\gamma + \epsilon^{(x)}_{ij(i)t} \tag{9.1}$$

for $i = 1, 2;\ j(i) = 1, 2, \ldots, n_i;\ x = A, B;\ t = 1, 2$, where γ is a nuisance parameter representing the sequence effect,

$$\mathbf{I} = \begin{cases} 0 \text{ for Period } 1 \\ 1 \text{ for Period } 2 \end{cases},$$

and

$$\mathbf{J} = \begin{cases} 1 \text{ for Group } 1 \\ -1 \text{ for Group } 2 \end{cases}.$$

Testing and interval estimation about the parameter requires some distributional assumptions as well. Following with the normal theory, we assume that the ϵ error terms are normally distributed with the terms corresponding to the same subject being correlated and different subjects being independent. We do not make further assumptions on the error structure since the current design does not lead to a useful reduction in the number of unknown parameters of the covariance matrix. The table below further illustrates the structure of various effects, where ρ is the sum of π and λ, which cannot be estimated separately.

Table 9.2: Response means by sequence and period

Group	Sequence	Period 1	Period 2
1	AB	$\mu_A + \gamma$	$\mu_B + \gamma + \rho$
2	BA	$\mu_B - \gamma$	$\mu_A - \gamma + \rho$

In the present design, the sequence numbers and the period numbers uniquely identify the treatment in effect. Therefore, for the sake of simplicity of notation, we shall suppress the treatment index x in the following development. Moreover, for convenience, we shall suppress the dependence of subject index on the sequence index, and simply use j instead of $j(i)$ with the understanding that the values that subscript j depend on the group in question, so we can rewrite the model as

$$Y_{ijt} = \mu_x + \mathbf{I}\rho + \mathbf{J}\gamma + \varepsilon_{ij} + \epsilon_{ijt}.$$

We can also write the model in alternative ways keeping four parameters to represent alternative effects. Widely used alternative models all yield the same point estimate for $\delta = \mu_A - \mu_B$, the quantity of primary interest, as

$$\widehat{\delta} = \frac{1}{2}(\overline{Y}_{11} + \overline{Y}_{22} - \overline{Y}_{12} - \overline{Y}_{21}), \tag{9.2}$$

where

$$\overline{Y}_{it} = \frac{\sum_{j=1}^{n_i} Y_{ijt}}{n_i}$$

is the sample mean computed using the observations from the ith Group and Period t. Table below further illustrates the sample means by sequence and period, which are unbiased estimates of the corresponding parameters appearing in Table 9.2.

Table 9.3: Sample means by sequence and period

Group	Sequence	Period 1	Period 2
1	AB	$\overline{Y}_{11}$	$\overline{Y}_{12}$
2	BA	$\overline{Y}_{21}$	$\overline{Y}_{22}$

For example, the estimate of δ given by (9.2) is established by solving the four equations obtained by equating the cell means. As a further example, $\overline{Y}_{11}$ is an unbiased estimate of $\mu_A + \gamma$ and $\overline{Y}_{21}$ is an unbiased estimate of $\mu_B - \gamma$, and therefore

$$\widehat{\gamma} = \frac{\overline{Y}_{11} - \overline{Y}_{21} - \widehat{\delta}}{2} \tag{9.3}$$

is an unbiased estimate of the sequence effect γ.

9.3 Comparing Treatments

Consider the problem of testing hypotheses concerning the parameter $\delta = \mu_A - \mu_B$. Before we proceed to do testing and interval estimation of the parameters, we need to establish necessary distributional results. If we had data only from the AB sequence or from the BA sequence, we would have performed a paired t-test. So it is intuitive that we should be able to device a t-test in analyzing with both data sets as well. To derive this more formally, assume standard bivariate normal distributions of the form

$$\mathbf{Y}_i = (\ Y_{i1} \quad Y_{i2}\ )' \sim N(\boldsymbol{\mu}_i, \Sigma_i), \quad for\ i = 1, 2. \tag{9.4}$$

The classical approach to making exact inferences on δ fails unless we assume that the two covariance matrices are the same for the data from the two groups in the observed order of data or when the order of data from one sequence is reversed. As illustrated by the Balaam design undertaken in the next section, the case of unequal covariance matrices could be easily tackled by the generalized approach and is left as an exercise (see Exercise 9.2). Here we assume that the two data sets follow normal distributions with the common covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}.$$

Then, the variance parameter we need to tackle is evident from

$$\begin{aligned} \mathrm{Var}(\overline{Y}_{11} - \overline{Y}_{12}) &= \frac{1}{n_1}\mathbf{a}'\Sigma\mathbf{a} \\ &= \frac{1}{n_1}\sigma^2 \end{aligned} \tag{9.5}$$

and from $\mathrm{Var}(\overline{Y}_{22}-\overline{Y}_{21}) = \frac{1}{n_2}\sigma^2$, where $\sigma^2 = \sigma_1^2+\sigma_2^2-2\sigma_{12}$ and $\mathbf{a}' = (\ 1 \quad -1\ )$. As a result, we get

$$\widehat{\delta} = \frac{1}{2}(\overline{Y}_{11} + \overline{Y}_{22} - \overline{Y}_{12} - \overline{Y}_{21}) \sim N(\delta, \frac{\sigma^2}{4}(\frac{1}{n_1} + \frac{1}{n_2})) \tag{9.6}$$

To tackle σ^2, by pairing the data from each subject receiving the sequence AB, we get independent distributions

$$y_{j1} = Y_{1j2} - Y_{1j1} \sim N(\rho - \delta, \sigma^2) \qquad \text{for } \ j = 1, 2, \ldots, n_1,$$

which are also independently distributed from

$$y_2 = Y_{2j2} - Y_{2j1} \sim N(\rho + \delta, \sigma^2) \qquad \text{for } \ j = 1, 2, \ldots, n_2.$$

These results imply that

$$\frac{(n_1 + n_2 - 2)S^2}{\sigma^2} \sim \chi^2_{n_1+n_2-2}, \tag{9.7}$$

where

$$S^2 = \frac{1}{n_1 + n_2 - 2}\sum_{i=1}^{2}\sum_{j=1}^{n_i}(y_{ji} - \overline{y}_i) \tag{9.8}$$

is the pooled unbiased estimator of σ^2. It is now evident from 9.6 and 9.7 that inferences on the parameter can be based on the result

$$\frac{\widehat{\delta} - \delta}{\frac{S}{2}\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})}} \sim t_{n_1+n_2-2}. \tag{9.9}$$

For example, it follows from this result that

$$[\widehat{\delta} - t_\nu \frac{s}{2}\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})},\ \widehat{\delta} + t_\nu \frac{s}{2}\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})}] \tag{9.10}$$

is a $100\gamma\%$ equal-tail confidence interval for $\delta = \mu_A - \mu_B$, where t_ν is the $\nu = (1+\gamma)/2$ quantile of the t distribution with $n_1 + n_2 - 2$ degrees of freedom. Similarly, the p-value for testing hypotheses of the form

$$H_0 : \delta < \delta_0$$

is computed as

$$p = 1 - G(\frac{\widehat{\delta} - \delta_0}{\frac{S}{2}\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})}}), \tag{9.11}$$

where G is the cdf of the t distribution with $n_1 + n_2 - 2$ degrees of freedom. Moreover, point null hypotheses of the form $H_0 : \delta = \delta_0$ are tested using the p-value

$$p = 2G(-\frac{2|\widehat{\delta} - \delta_0|}{S\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})}}), \tag{9.12}$$

Example 9.1. Comparing two diets

Consider the problem of comparing two diets A and B given to pigs using a crossover design. Table 9.4 is a sample of hypothetical data on gain in weights of a sample of 7 pigs during two periods. Four pigs receive the diet sequence A followed by B, and the rest of the pigs receive the diet B followed by A. Observe that in this example the gains in weight are consistently larger in the second period. This is a period effect rather than a carryover effect. Consider the problem of testing whether or not one diet is better than the other in terms of mean gain in weight. In this type of application the best diet plan might actually be any of the four sequences AB, BA, AA, or BB, a question we will address later in this chapter, but here we simply compare the performance of the two diets in a single period.

Table 9.4: Weight gains: Two-sequence case

Pig #	Sequence	Period 1	Period 2	y
01	AB	11.2	17.8	6.6
02	AB	12.7	18.0	5.3
03	AB	9.9	16.8	6.9
04	AB	10.4	17.4	7.0
05	BA	12.0	17.7	5.7
06	BA	11.4	17.1	5.7
07	BA	11.0	15.8	4.8

The table below shows the sample means by sequence and period. In terms of the sample means, the difference in diet means can be estimated using formula (9.2) as

$$\begin{aligned} \widehat{\delta} &= \frac{11.05 + 16.87 - 11.47 - 17.5}{2} \\ &= -0.525. \end{aligned}$$

This indicates the possibility that diet B might be better than diet A, but the question is whether or not the result is statistically significant or the estimate is an artifact of sampling variation.

Mean gains in weight			
Group	Sequence	Period 1	Period 2
1	AB	11.05	17.50
2	BA	11.47	16.87

To test the underlying hypothesis, let us first compute the estimate of σ^2 using (9.2) and the paired differences shown in the last column of Table 9.4 as

$$\begin{aligned} S^2 &= \frac{1}{n_1+n_2-2}\sum_{i=1}^{2}\sum_{j=1}^{n_i}(y_{ji}-\overline{y}_i)^2 \\ &= \frac{3\times \text{Var}(6.6,5.3,6.9,7.0)+2\times Var(5.7,5.7,4.8)}{5} \\ &= \frac{3\times 0.617+2\times 0.27}{5} \\ &= 0.478. \end{aligned}$$

Now consider the hypothesis $H_0:\delta=0$. The p-value for testing the hypothesis is computed from (9.12) as

$$\begin{aligned} p &= 2G(-\frac{2|\widehat{\delta}|}{S\sqrt{(\frac{1}{n_1}+\frac{1}{n_2})}}) \\ &= 2G(-\frac{2\times 0.525}{0.692\sqrt{(\frac{1}{4}+\frac{1}{3})}}) \\ &= 2G(-1.987) \\ &= 0.10, \end{aligned}$$

where G is the cdf of the t distribution with 5 degrees of freedom. So we have some evidence to reject the null hypothesis and conclude that diet B might be better than diet A. Since the evidence is not very strong, it is advisable to conduct the experiment with larger samples.

9.4 Four-Sequence Design

In the above treatment we had to assume that the carryover effects, which are also aliased with the period effects, are equal for the two sequences. This may not be a reasonable assumption in many applications, especially when one of the treatments is a placebo. In fact, in many situations the drawbacks of model may outweigh the benefits of the crossover design. The assumption can be avoided and additional parameters can be introduced to account for interaction effects if we have data from the four sequences AB, BA, AA and BB. The resulting design is known as the Balaam design [see Balaam (1968)]. Let n_1, n_2, n_4, n_4 be the sample sizes from the four groups.

Let $Y^{(x)}_{ij(i)t}$ denote the observation (or average of observations) taken from $j(i)$th experimental unit in ith group at occasion t. Assuming only first-order effects, let μ_A, μ_B be the treatment means in period 1, and let λ_A, λ_B be the carryover effects from period 1 to 2. Let γ_{it} denote other main effects and interaction terms representing effects such as the period effects, the sequence effects, and the interactions between the treatments and the periods. Their structure can be specified in alternative ways subject to a maximum of four unknown parameters. Although the γ_{it} parameters are considered as nuisance parameters in estimating main treatment effects, they might also be parameters of importance in some applications. For example, in an agricultural experiment concerning two diet plans, the best plan might be diet A during period 1 and diet B during period 2. Assume the linear model

$$Y^{(x)}_{ij(i)t} = \mu_x + \mathbf{I}\lambda_x + \gamma_{it} + \varepsilon^{(x)}_{ij(i)} + \epsilon^{(x)}_{ij(i)t} \tag{9.13}$$

for $i = 1, 2, 3, 4$; $j(i) = 1, 2, \ldots, n_i$; $x = A, B$; $t = 1, 2,$ where $\varepsilon^{(x)}_{ij(i)}$ are random effects representing the subject variation, $\epsilon^{(x)}_{ij(i)t}$ are the residual errors, and

$$\mathbf{I} = \begin{cases} 0 \text{ for period } 1 \\ 1 \text{ for period } 2. \end{cases}$$

In making inferences beyond point estimation, we further assume that $\varepsilon^{(x)}_{ij(i)}$ and $\epsilon^{(x)}_{ij(i)t}$ are normally distributed. Table below illustrates the structure of the fixed effects.

Table 9.5: Response means by sequence and period

Group	Sequence	Period 1	Period 2
1	AB	$\mu_A + \gamma_{11}$	$\mu_B + \gamma_{12} + \lambda_A$
2	BA	$\mu_B + \gamma_{21}$	$\mu_A + \gamma_{22} + \lambda_B$
3	AA	$\mu_A + \gamma_{31}$	$\mu_A + \gamma_{32} + \lambda_A$
4	BB	$\mu_B + \gamma_{41}$	$\mu_B + \gamma_{42} + \lambda_B$

Since we allow nuisance parameters in the two periods to be different, without loss of generality as far as the main effects μ_A, μ_B are concerned, we have assumed that the carryover effect in sequence 1 is the same as sequence 3 and that of sequence 2 is the same as sequence 4. In fact the nuisance parameters are measured as deviations from the mean effects and the carry over effects. Therefore, they are normalized to satisfy the usual constraints,

$$\gamma_{11} + \gamma_{22} + \gamma_{31} + \gamma_{32} = 0,$$
$$\gamma_{12} + \gamma_{21} + \gamma_{41} + \gamma_{42} = 0,$$

and

$$\begin{aligned} \gamma_{12} + \gamma_{32} &= 0, \\ \gamma_{22} + \gamma_{42} &= 0, \end{aligned}$$

which is equivalent to

$$\begin{aligned} \gamma_{11} + \gamma_{21} + \gamma_{31} + \gamma_{41} &= 0, \\ \gamma_{12} + \gamma_{22} + \gamma_{32} + \gamma_{42} &= 0. \end{aligned}$$

In the present design, the group numbers and the period numbers uniquely identify the treatment in effect. For example, $i = 3$, $t = 2$ implies that $x = A$. Therefore, as in the previous section, we shall suppress the treatment index x in the following development. Moreover as before, we suppress the dependence of subject index on the group index, and use j instead of $j(i)$ with the understanding that the values that subscript j takes depend on the group in question. Then, model (9.13) can be expressed as

$$Y_{ijt} = \mu_x + \mathbf{I}\lambda_x + \gamma_{it} + \varepsilon_{ij} + \epsilon_{ijt}. \tag{9.14}$$

9.4.1 Point estimates

First consider the problem of estimating fixed effects in Table 9.6 under the constraints assumed above. Let

$$\overline{Y}_{it} = \frac{\sum_{j=1}^{n_i} Y_{ijt}}{n_i}$$

be the sample means computed using observations from ith sequence and period t. Table 9.6 presents cell means by sequence and period.

Table 9.6: Means by sequence and period

Group	Sequence	Period 1	Period 2
1	AB	$\overline{Y}_{11}$	$\overline{Y}_{12}$
2	BA	$\overline{Y}_{21}$	$\overline{Y}_{22}$
3	AA	$\overline{Y}_{31}$	$\overline{Y}_{32}$
4	BB	$\overline{Y}_{41}$	$\overline{Y}_{42}$

From model (9.14) we get

$$E(\overline{Y}_{it}) = \mu_x + \mathbf{I}\lambda_x + \gamma_{it}.$$

Therefore, unbiased estimates of each of the fixed effects can be obtained by equating the cells in the expected means and sample means tables given above and solving the equations. The unbiased estimates of μ_A and μ_B obtained by solving the equations are

$$\widehat{\mu}_A = \frac{3}{8}(\overline{Y}_{11} + \overline{Y}_{31}) + \frac{1}{8}(\overline{Y}_{22} + \overline{Y}_{32} + \overline{Y}_{21} + \overline{Y}_{41}) - \frac{1}{8}(\overline{Y}_{12} + \overline{Y}_{42}) \tag{9.15}$$

and

$$\widehat{\mu}_B = \frac{3}{8}(\overline{Y}_{21} + \overline{Y}_{41}) + \frac{1}{8}(\overline{Y}_{12} + \overline{Y}_{42} + \overline{Y}_{11} + \overline{Y}_{31}) - \frac{1}{8}(\overline{Y}_{22} + \overline{Y}_{32}), \quad (9.16)$$

respectively. Having estimated μ_A and μ_B, the unbiased estimates of the nuisance parameters and the carryover effects can be obtained as

$$\begin{aligned}
\widehat{\gamma}_{11} = \overline{Y}_{11} - \widehat{\mu}_A,\ \widehat{\gamma}_{21} = \overline{Y}_{21} - \widehat{\mu}_B,\\
\widehat{\gamma}_{31} = \overline{Y}_{31} - \widehat{\mu}_A,\ \widehat{\gamma}_{41} = \overline{Y}_{41} - \widehat{\mu}_B,\\
\widehat{\gamma}_{12} = -\widehat{\gamma}_{32} = \frac{1}{2}(\overline{Y}_{12} - \overline{Y}_{32} + \widehat{\mu}_A - \widehat{\mu}_B),\\
\widehat{\gamma}_{22} = -\widehat{\gamma}_{42} = \frac{1}{2}(\overline{Y}_{22} - \overline{Y}_{42} - \widehat{\mu}_A + \widehat{\mu}_B),\\
\widehat{\lambda}_A = \frac{1}{2}\{(\overline{Y}_{12} + \overline{Y}_{32}) - (\widehat{\mu}_A + \widehat{\mu}_B)\}, \quad (9.17)
\end{aligned}$$

and

$$\widehat{\lambda}_B = \frac{1}{2}\{(\overline{Y}_{22} + \overline{Y}_{42}) - (\widehat{\mu}_A + \widehat{\mu}_B)\}. \quad (9.18)$$

It is easily seen that these estimates are also the maximum likelihood estimates .

Of special importance are parameters that arise in comparisons such as the difference in treatment means, $\nu = \mu_A - \mu_B$, and the difference in carryover effects, $\eta = \lambda_A - \lambda_B$. Their estimates that follow from the above equations are

$$\begin{aligned}
\widehat{\nu} &= \frac{(\overline{Y}_{11} - \overline{Y}_{12}) + (\overline{Y}_{22} - \overline{Y}_{21}) + (\overline{Y}_{31} + \overline{Y}_{32}) - (\overline{Y}_{41} + \overline{Y}_{42})}{4}, \quad (9.19)\\
&= \frac{(\overline{Y}_{11} - \overline{Y}_{12}) + (\overline{Y}_{22} - \overline{Y}_{21}) + (\overline{Y}_{31} - \overline{Y}_{41}) + (\overline{Y}_{32} - \overline{Y}_{42})}{4}
\end{aligned}$$

and

$$\widehat{\eta} = \frac{1}{2}\{(\overline{Y}_{12} + \overline{Y}_{32}) - (\overline{Y}_{22} + \overline{Y}_{42})\}$$

Some of the above point estimates are valid only under the particular set of constraints placed on the parameters γ_{it}. Nevertheless, other widely used normalizations yield the same estimate for the parameter of special importance, namely $\nu = \mu_A - \mu_B$. Of particular interest is the case where we allow the carryover effects to be different for all four groups. In this case, we can treat $\widetilde{\gamma}_{it} = \gamma_{it} + \lambda_x$ as the nuisance parameters for the second period and impose just the two constraints, namely

$$\begin{aligned}
\gamma_{11} + \widetilde{\gamma}_{22} + \gamma_{31} + \widetilde{\gamma}_{32} = 0\\
\widetilde{\gamma}_{12} + \gamma_{21} + \gamma_{41} + \widetilde{\gamma}_{42} = 0. \quad (9.20)
\end{aligned}$$

It is easily seen that in this case the point estimates of the treatment means become

$$\widehat{\mu}_A = \frac{1}{4}(\overline{Y}_{11} + \overline{Y}_{22} + \overline{Y}_{31} + \overline{Y}_{32}) \quad (9.21)$$

and

$$\widehat{\mu}_B = \frac{1}{4}(\overline{Y}_{12} + \overline{Y}_{21} + \overline{Y}_{41} + \overline{Y}_{42}). \tag{9.22}$$

Clearly these estimates also yield the same estimate as before for $\nu = \mu_A - \mu_B$.

Obviously, all fixed effects of interest, including the individual means and the differences in means, can be expressed as

$$\widehat{\theta} = \sum_{i=1}^{4}\sum_{j=1}^{2} a_{ij}\overline{Y}_{ij}, \tag{9.23}$$

where a_{ij} are known constants as specified above. Therefore, it is convenient to develop distributional results for general values of a_{ij} before we undertake special cases.

9.5 Distributional Results

The classical approach does not provide exact inferences beyond point estimation for θ and Weerahandi and Peterson (2003) showed how the generalized approach could be taken in this context. Before we could proceed to do testing and interval estimation of parameters of interest, we need to establish necessary distributional results. Notice from Model (9.13) that it does not yield any reduction in the number of parameters in the covariance matrix for the data from Group 1 and Group 2. In other words the covariance matrix is unstructured and hence AB and BA data sets follow standard bivariate models of the form

$$\mathbf{Y}_i \sim N(\boldsymbol{\mu}_i, \Sigma_i), \qquad i = 1, 2. \tag{9.24}$$

We can make inferences on θ regardless of whether or not the covariance matrices are equal. Here we develop testing procedures without assuming the equality of covariances and the case of equal covariances is left as an exercise.

It is evident that

$$\text{Var}(a_{11}\overline{Y}_{11} + a_{12}\overline{Y}_{12}) = \frac{1}{n_1}\mathbf{a}_1'\Sigma_1\mathbf{a}_1 \tag{9.25}$$

and that

$$\text{Var}(a_{21}\overline{Y}_{21} + a_{22}\overline{Y}_{22}) = \frac{1}{n_2}\mathbf{a}_2'\Sigma_2\mathbf{a}_2, \tag{9.26}$$

where $\mathbf{a}_1' = \begin{pmatrix} a_{11} & a_{12} \end{pmatrix}$ and $\mathbf{a}_2' = \begin{pmatrix} a_{21} & a_{22} \end{pmatrix}$ are 2×1 vectors of known constants. Notice also that model 9.13 does yield structured covariance matrices for the data from Group 3 and Group 4. Let $\mathbf{Y}_A$ and $\mathbf{Y}_B$ be the random vectors representing the data from these groups following the sequences AA and BB respectively. Then it follows from the one factor repeated measures results that

$$\mathbf{Y}_A \sim N(\boldsymbol{\mu}_{AA}, \begin{pmatrix} \sigma_A^2 + \delta_A^2 & \delta_A^2 \\ \delta_A^2 & \sigma_A^2 + \delta_A^2 \end{pmatrix})$$

and that

$$\mathbf{Y}_B \sim N(\boldsymbol{\mu}_{BB}, \begin{pmatrix} \sigma_B^2+\delta_B^2 & \delta_B^2 \\ \delta_B^2 & \sigma_B^2+\delta_B^2 \end{pmatrix}).$$

We are not making any assumptions such as the equality of variance components, which are unnecessary and unreasonable when we have allowed unequal interaction effects and carryover effects. Now it is evident that

$$\mathrm{Var}(a_{31}\overline{Y}_{31}+a_{32}\overline{Y}_{32}) = \frac{1}{n_3}[(a_{31}^2+a_{32}^2)\sigma_A^2+(a_{31}+a_{32})^2\delta_A^2) \tag{9.27}$$

and that

$$\mathrm{Var}(a_{41}\overline{Y}_{41}+a_{42}\overline{Y}_{42}) = \frac{1}{n_3}[(a_{41}^2+a_{42}^2)\sigma_B^2+(a_{41}+a_{42})^2\delta_B^2). \tag{9.28}$$

Of special importance are the particular cases

$$\mathrm{Var}(\overline{Y}_{31}+\overline{Y}_{32}) = \frac{2}{n_3}(\sigma_A^2+2\delta_A^2) \tag{9.29}$$

and

$$\mathrm{Var}(\overline{Y}_{41}+\overline{Y}_{42}) = \frac{2}{n_4}(\sigma_B^2+2\delta_B^2), \tag{9.30}$$

which imply that

$$\widehat{\nu} \sim N(\nu, \frac{1}{16}(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2})+\frac{1}{8n_3}(\sigma_A^2+2\delta_A^2)+\frac{1}{8n_4}(\sigma_B^2+2\delta_B^2), \tag{9.31}$$

where $\sigma_1^2 = \mathbf{a}'\Sigma_1\mathbf{a}$ and $\sigma_2^2 = \mathbf{a}'\Sigma_2\mathbf{a}$, and $\mathbf{a}' = \begin{pmatrix} 1 & -1 \end{pmatrix}$. Since the variance components appearing in (9.31) are unknown parameters, they also need to be tacked by some statistics.

To handle the variance covariance matrix Σ, consider the samples from the first two groups following a multivariate distribution of the form (9.24).

$$\mathbf{Y}_{1j} = \begin{pmatrix} Y_{1j1} & Y_{1j2} \end{pmatrix} \backsim N(\boldsymbol{\mu}_{AB}, \Sigma_1) \ , \qquad j = 1, 2, \ldots, n_1$$

and

$$\mathbf{Y}_{2j} = \begin{pmatrix} Y_{2j1} & Y_{2j2} \end{pmatrix} \backsim N(\boldsymbol{\mu}_{BA}, \Sigma_2) \, , \qquad j = 1, 2, \ldots, n_2,$$

where $\boldsymbol{\mu}_{AB} = (\mu_A+\gamma_{11} \ \mu_B+\gamma_{12}+\lambda_A)'$ and $\boldsymbol{\mu}_{BA} = (\mu_B+\gamma_{21} \ \mu_A+\gamma_{22}+\lambda_B)'$. Then, it is known from the theory of sampling from multivariate normal distributions that

$$S_1 = \sum_{j=1}^{n_1}(\mathbf{Y}_{1j}-\overline{\mathbf{Y}}_1)(\mathbf{Y}_{1j}-\overline{\mathbf{Y}}_1)' \sim \mathbf{W}(n_1-1, \Sigma_1) \tag{9.32}$$

and that

$$S_2 = \sum_{j=1}^{n_2}(\mathbf{Y}_{2j}-\overline{\mathbf{Y}}_2)(\mathbf{Y}_{2j}-\overline{\mathbf{Y}}_2)' \sim \mathbf{W}(n_2-1, \Sigma_2), \tag{9.33}$$

where $\overline{\mathbf{Y}}_i = \sum_{j=1}^{n_1} \mathbf{Y}_{ij}/n_i$. Moreover, $S_1, \overline{\mathbf{Y}}_1$, S_2, and $\overline{\mathbf{Y}}_2$ are all independently distributed. In particular, (9.32) and (9.33) imply that

$$W_i = \frac{\mathbf{a}' S_i \mathbf{a}}{\mathbf{a}' \Sigma_i \mathbf{a}} = \frac{S_{ii}}{\sigma_i^2} \sim \chi^2_{n_i - 1} \qquad \text{for } i = 1, 2. \tag{9.34}$$

Note that $S_{ii}/(n_i - 1)$ terms are also the same as the variance of the differences of the paired data of the first two groups.

Define

$$\begin{aligned} \overline{Y}_{3j} &= \sum_{t=1}^{2} Y_{3jt}/2\,, \\ \overline{Y}_{3t} &= \sum_{j=1}^{n_3} Y_{3jt}/2\,, \end{aligned}$$

and

$$\overline{Y}_3 = \sum_{j=1}^{n_3} \overline{Y}_{3j}/n_3.$$

From the one-factor repeated measures model we can deduce appropriate statistics and distributions on which inferences of σ_A^2 and δ_A^2 could be based. Recall from Section 7.2 that the appropriate statistics are

$$S_{31} = \sum_{t=1}^{2} \sum_{j=1}^{n_3} (Y_{3jt} - \overline{Y}_{3j} - \overline{Y}_{3t} + \overline{Y}_3)^2$$

and

$$S_{32} = 2 \sum_{j=1}^{n_3} (\overline{Y}_{3j} - \overline{Y}_3)^2, \tag{9.35}$$

and their distributions are given by

$$U_3 = \frac{S_{31}}{\sigma_A^2} \sim \chi^2_{n_3 - 1} \tag{9.36}$$

and

$$V_3 = \frac{S_{32}}{\sigma_A^2 + 2\delta_A^2} \sim \chi^2_{n_3 - 1}, \tag{9.37}$$

respectively, where Similarly, in terms of various sample means as defined above, the variance components σ_B^2 and δ_B^2 can be handled by using the sums of squares

$$S_{41} = \sum_{t=1}^{2} \sum_{j=1}^{n_4} (Y_{4jt} - \overline{Y}_{4j} - \overline{Y}_{4t} + \overline{Y}_4)^2,$$

and

$$S_{42} = 2\sum_{j=1}^{n_4}(\overline{Y}_{4j} - \overline{Y}_4)^2$$

having the distributions

$$U_4 = \frac{S_{41}}{\sigma_B^2} \sim \chi^2_{n_4-1} \tag{9.38}$$

and

$$V_4 = \frac{S_{42}}{\sigma_B^2 + 2\delta_B^2} \sim \chi^2_{n_4-1}. \tag{9.39}$$

Moreover, these random variables are independently distributed.

9.6 Testing and Interval Estimation

We are now in a position to make inferences, beyond point estimation, on the parameters of the Balaam design. Of special interest are the problems of comparing of treatments, comparing the carryover effects, and making inferences about individual treatment means. Inferences on the variance components and the interactions are also of some interest. Although we have fairly simple distributional results as outlined above, the classical approach fails to provide tests and confidence intervals based on exact probability statements. Despite the that the 2-treatment, 2-period Balaam design is the simplest crossover design allowing unequal carryover effects, there are only asymptotic and other approximate methods available in the literature for this problem. The generalized approach allows us to make inferences about any of the parameters. The approach is illustrated below with its application to some important parameters.

9.6.1 Comparing treatments

To compare the two treatments, first consider the problem of testing hypotheses of the form

$$H_0 : \nu \leq \nu_0,$$

where $\nu = \mu_A - \mu_B$. The Z statistic that follows from (9.31), namely

$$Z = \frac{\widehat{\nu} - \nu_0}{\sqrt{\frac{1}{16}(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}) + \frac{1}{8n_3}(\sigma_A^2 + 2\delta_A^2) + \frac{1}{8n_4}(\sigma_B^2 + 2\delta_B^2)}} \sim N(0,1), \tag{9.40}$$

is a standard normal random variable that we can use to test H_0 if the variance components were known. When they are unknown, the substitution method

suggests that the appropriate generalized p-value for testing H_0 is given by

$$\begin{aligned} p &= \Pr(Z \geq \frac{(\overline{y}_{11} - \overline{y}_{12}) + (\overline{y}_{22} - \overline{y}_{21}) + (\overline{y}_{31} + \overline{y}_{32}) - (\overline{y}_{41} + \overline{y}_{42}) - 4\nu_0}{\sqrt{\frac{s_{11}}{n_1 W_1} + \frac{s_{22}}{n_2 W_2} + \frac{2s_{32}}{n_3 V_3} + \frac{2s_{42}}{n_4 V_4}}}) \\ &= 1 - E\Phi(\frac{\Delta y - 4\nu_0}{\sqrt{\frac{s_{11}}{n_1 W_1} + \frac{s_{22}}{n_2 W_2} + \frac{2s_{32}}{n_3 V_3} + \frac{2s_{42}}{n_4 V_4}}}), \end{aligned} \tag{9.41}$$

where $\Delta y = (\overline{y}_{11} - \overline{y}_{12}) + (\overline{y}_{22} - \overline{y}_{21}) + (\overline{y}_{31} + \overline{y}_{32}) - (\overline{y}_{41} + \overline{y}_{42})$, Φ is the cdf of the standard normal distribution, the expected value is taken with respect to the independent random variables W_1, W_2, V_3, V_4, and the lower case letters of sample means and sums of squares denote the observed values of the corresponding random variables. That (9.41) is the probability of an extreme region is evident when the region defined by (9.41) is expressed in the form

$$\left\{ \mathbf{Y} \middle| \begin{array}{c} \frac{\{(\overline{Y}_{11} + \overline{Y}_{22} + \overline{Y}_{31} + \overline{Y}_{32}) - (\overline{Y}_{12} + \overline{Y}_{21} + \overline{Y}_{41} + \overline{Y}_{42})\} - 4\nu_0}{\sqrt{(\frac{\sigma_1^2}{n_1} + \frac{\sigma_1^2}{n_2}) + \frac{2}{n_3}(\sigma_A^2 + 2\delta_A^2) + \frac{2}{n_4}(\sigma_B^2 + 2\delta_B^2)}} \\ \geq \frac{\{(\overline{y}_{11} - \overline{y}_{12}) + (\overline{y}_{22} - \overline{y}_{21}) + (\overline{y}_{31} + \overline{y}_{32}) - (\overline{y}_{41} + \overline{y}_{42})\} - 4\nu_0}{\sqrt{(\frac{\sigma_1^2}{n_1}\frac{s_{11}}{S_{11}} + \frac{\sigma_2^2}{n_2}\frac{s_{22}}{S_{22}}) + \frac{2s_{32}}{n_3 S_{32}}(\sigma_A^2 + 2\delta_A^2) + \frac{2s_{42}}{n_4 S_{42}}(\sigma_B^2 + 2\delta_B^2)}} \end{array} \right\}.$$

Moreover, the probability of the extreme region increases with deviations from the null hypothesis implying that the test given by (9.41) is unbiased. This type of integral are well-behaved and hence is easily evaluated by numerical integration. The p-value can also be well approximated by simulating a large set of chi-squared random numbers and then estimating the expected value in (9.41) by the sample mean of the corresponding quantity.

Procedures for testing point null hypotheses of the form $H_0 : \nu = \nu_0$ could be deduced from the p-value given by (9.41) for one-sided hypotheses. In this case, too small values as well as too large values of the generalized test variable implied by (9.41) constitute the extreme region. From the symmetry of the Z variable, we can thus deduce [cf. Weerahandi (1995)] the generalized p-value for testing H_0 as

$$p = 2E\Phi(\frac{-|\Delta y - 4\nu_0|}{\sqrt{\frac{s_{11}}{n_1 W_1} + \frac{s_{22}}{n_2 W_2} + \frac{2s_{32}}{n_3 V_3} + \frac{2s_{42}}{n_4 V_4}}}), \tag{9.42}$$

The generalized confidence intervals for ν are derived from the generalized pivotal quantity given by the substitution method or deduced from the generalized p-value. For example, the 95% symmetric generalized confidence interval of ν implied by (9.41) is of the form $[(\Delta y - k)/4, (\Delta y + k)/4]$, where k is chosen such that

$$E\Phi(\frac{k}{\sqrt{\frac{s_{11}}{n_1 W_1} + \frac{s_{22}}{n_2 W_2} + \frac{2s_{32}}{n_3 V_3} + \frac{2s_{42}}{n_4 V_4}}}) = 0.975.$$

Example 9.2. Comparing two diets (continued)

Consider again the problem of comparing two diets A and B given to pigs by means of a crossover design. Now suppose some data become available from the

sequences AA and BB as well so that we have a complete data set from the four sequence Balaam design. Table 9.7 is a sample of hypothetical data on gain in weights of a sample of 13 pigs during the two periods.

Table 9.7: Weight gains: Four-sequence case

Pig #	Sequence	Period 1	Period 2	y
01	AB	11.2	17.8	6.6
02	AB	12.7	18.0	5.3
03	AB	9.9	16.8	6.9
04	AB	10.4	17.4	7.0
05	BA	12.0	17.7	5.7
06	BA	11.4	17.1	5.7
07	BA	11.0	15.8	4.8
08	AA	11.6	16.7	5.1
09	AA	10.9	17.0	6.1
10	AA	11.7	16.8	5.1
11	BB	12.0	17.9	5.9
12	BB	11.4	17.7	6.3
13	BB	12.4	18.4	6.0

Mean gain in weights by sequence and period, on which we can base the point estimation of parameters, are also shown below followed by Table 9.7. Of special interest is the difference in mean diet effects estimated as

$$\begin{aligned}\widehat{\nu} &= \frac{(\overline{Y}_{11}-\overline{Y}_{12})+(\overline{Y}_{22}-\overline{Y}_{21})+(\overline{Y}_{31}-\overline{Y}_{41})+(\overline{Y}_{32}-\overline{Y}_{42})}{4}\\ &= \frac{-6.45+5.40-0.53-1.17}{4}\\ &= -0.688\end{aligned}$$

Mean gains in weight

Group	Sequence	Period 1	Period 2
1	AB	11.05	17.50
2	BA	11.47	16.87
3	AA	11.40	16.83
4	BB	11.93	18.00

As in Example 9.1, the point estimate of ν suggests the possibility that diet B is better than diet A. To test the significance of the estimate, consider the hypothesis $H_0 : \nu = 0$ that there is no difference between the mean effects of the two diets. To test this hypothesis using the generalized p-value given by (9.41), let us first compute various sums of squares of deviations appearing in

the formula as

$$\begin{aligned} s_{11} &= 3 \times \text{Var}(6.6, 5.3, 6.9, 7.0) \\ &= 1.85, \\ s_{22} &= 2 \times \text{Var}(5.7, 5.7, 4.8) \\ &= 0.54 \end{aligned}$$

$$\begin{aligned} s_{32} &= 2 \times 2 \times \text{Var}(\frac{11.6+16.7}{2}, \frac{10.9+17}{2}, \frac{11.7+16.8}{2}) \\ &= 0.093, \end{aligned}$$

$$\begin{aligned} s_{42} &= 2 \times 2 \times \text{Var}(\frac{12+17.9}{2}, \frac{11.4+17.7}{2}, \frac{12.4+18.4}{2}) \\ &= 0.723. \end{aligned}$$

Now we can compute the generalized p-value as

$$\begin{aligned} p &= 2E\Phi(\frac{-4 * 0.688}{\sqrt{\frac{1.85}{4W_1} + \frac{0.54}{3W_2} + \frac{2\times 0.093}{3V_3} + \frac{2\times 0.723}{3V_4}}}), \\ &= 2E\Phi(\frac{-2.752}{\sqrt{\frac{0.4625}{W_1} + \frac{0.18}{W_2} + \frac{0.062}{V_3} + \frac{0.482}{V_4}}}), \\ &= 0.014 \end{aligned}$$

where the expectation is computed using 10,000 random digits generated from the independent chi-squared random variables

$$W_1 \sim \chi^2_2, W_2 \sim \chi^2_2, V_3 \sim \chi^2_2, \text{ and } V_4 \sim \chi^2_2.$$

With this p-value we have fairly strong evidence to conclude that diet B is better than diet A.

In this type of application it is possible that change of diet from one period to the next is the best diet plan in maximizing the total gain in weight. If total gain in weight is the quantity of interest in this example, then the underlying problem is nothing but a classical ANOVA problem. In this case, various tests could be based on the data shown in the table below.

Total weight gains by diet plan

Pig #	Sequence	Weight Gain
01	AB	29.0
02	AB	30.7
03	AB	26.7
04	AB	27.8
05	BA	29.7
06	BA	28.5
07	BA	26.8
08	AA	28.3
09	AA	27.9
10	AA	28.5
11	BB	29.9
12	BB	29.1
13	BB	30.8

Estimated mean gains in weight due to the four diet plans AB, BA, AA, and BB are 28.55, 28.33, 28.23, and 29.93 respectively, indicating no clear winner. In fact the p-value of the classical F-test for testing the equality of four diet plans is 0.38 and that of the generalized F-test is 0.23. Therefore, we do not have sufficient evidence to reject the null hypothesis of equal effects and proceed to multiple comparisons. However, we know from the above results that the mean effect of Diet B is significantly better than that of A, and so the diet plan BB should be recommended.

9.6.2 Comparing carryover effects

Now consider the problem of making inferences about the difference in the two carryover effects, namely $\eta = \lambda_A - \lambda_B$. From the distributional results of the previous section, we can obtain the distribution of its point estimate as

$$\begin{aligned} \widehat{\eta} &= \frac{(\overline{Y}_{12} + \overline{Y}_{32}) - (\overline{Y}_{22} + \overline{Y}_{42})}{2} \\ &\sim N(\eta, \frac{1}{4}(\frac{\Sigma_{1(22)}}{n_1} + \frac{\Sigma_{2(22)}}{n_2} + \frac{(\sigma_A^2 + \delta_A^2)}{n_3} + \frac{(\sigma_B^2 + \delta_B^2)}{n_4})), \end{aligned} \tag{9.43}$$

where $\Sigma_{i(22)}$ is the lower diagonal element of the matrix Σ_i for $i = 1, 2$. The nuisance parameters $\sigma_A^2, \delta_A^2, \sigma_B^2, \delta_B^2$ could be tackled as before using the statistics and distributions given by(9.36), (9.37), and (9.38). The remaining two parameters should be tackled using the sample variances of the data from the second cells of Group 1 and Group 2, say $S_1^2 = S_{1(2)}/(n_1 - 1)$ and $S_2^2 = S_{2(2)}/(n_2 - 1)$, which are independently distributed as

$$X_1 = \frac{S_{1(2)}}{\Sigma_{1(22)}} \sim \chi^2_{n_1 - 1} \tag{9.44}$$

and

$$X_2 = \frac{S_{2(2)}}{\Sigma_{2(22)}} \sim \chi^2_{n_2-1}, \tag{9.45}$$

where

$$S_{i(2)} = \sum_{j=1}^{n_i} (Y_{ij2} - \overline{Y}_{i2})^2, \qquad i = 1, 2.$$

As in the previous section, we can obtain the generalized p-value for testing hypotheses of the form

$$H_0 : \eta \leq \eta_0,$$

based on the Z statistic

$$Z = \frac{2(\widehat{\eta} - \eta_0)}{\sqrt{\frac{\Sigma_{1(22)}}{n_1} + \frac{\Sigma_{2(22)}}{n_2} + \frac{(\sigma_A^2 + 2\delta_A^2) + \sigma_A^2}{2n_3} + \frac{(\sigma_B^2 + 2\delta_B^2) + \sigma_B^2}{2n_4}}} \sim N(0, 1). \tag{9.46}$$

Now it is clear that the generalized p-value based on above statistics is given by

$$\begin{aligned} p &= \Pr(Z \geq \frac{(\overline{y}_{12} + \overline{y}_{32}) - (\overline{y}_{22} + \overline{y}_{42}) - 2\eta_0}{\sqrt{f(X_1, X_2, V_3, V_4, U_3, U_4)}}), \\ &= 1 - E\Phi(\frac{(\overline{y}_{12} + \overline{y}_{32}) - (\overline{y}_{22} + \overline{y}_{42}) - 2\eta_0}{\sqrt{f(X_1, X_2, V_3, V_4, U_3, U_4)}}), \end{aligned} \tag{9.47}$$

where Φ is the cdf of the standard normal distribution,

$$\begin{aligned} f(X_1, X_2, V_3, V_4, U_3, U_4) &= \frac{s_{1(2)}}{n_1 X_1} + \frac{s_{2(2)}}{n_2 X_2} + \frac{s_{32}}{2n_3 V_3} \\ &\quad + \frac{s_{31}}{2n_3 U_3} + \frac{s_{42}}{2n_4 V_4} + \frac{s_{41}}{2n_4 U_4}, \end{aligned}$$

and the expected value is taken with respect to the independent random variables $X_1, X_2, V_3, V_4, U_3, U_4$. It is straightforward to deduce generalized confidence intervals for η and is left as an Exercise.

Exercises

Exercise 1 *Consider a linear model of the form (9.1) and assume that there is no carryover effect.*

(a) Derive an unbiased estimate of π , the period effect,
(b) Show that it is also the same as the MLE of π,
(c) Construct a 95% confidence interval for π.

Exercise 2 *Consider the linear model (9.1) and assume that the data from each group follows a bivariate normal distribution. When the covariance matrices of the two distributions are not equal, establish generalized procedures for testing hypotheses concerning δ, the difference in the two treatment means. Also establish generalized confidence intervals for δ.*

Exercise 3 *Making the same assumptions as in the previous exercise, construct generalized tests and generalized confidence intervals for ρ, the sum of the common carryover effect and the period effect.*

Exercise 4 *Consider the data set in Table 9.1. Test the hypothesis that there is no difference in the two treatments. Construct left-sided 95% confidence intervals for the parameters δ and ρ when the two covariance matrices are equal.*

Exercise 5 *Consider again the data set in Table 9.1. Construct 95% confidence intervals for the parameters δ and ρ when the two covariance matrices are unequal.*

Exercise 6 *Consider the data set in Table 9.4. Construct left-sided 95% confidence intervals for δ and ρ when the two covariance matrices are equal.*

Exercise 7 *Consider again the data set in Table 9.4. Construct 95% confidence intervals for δ and ρ when the two covariance matrices are unequal.*

Exercise 8 *Consider the generalized p-value given by (9.41). By defining a set of independent random variables*

$$B_1 = \frac{W_1}{W_1+W_2} \sim Beta(\frac{n_1-1}{2}, \frac{n_2-1}{2}),$$
$$B_2 = \frac{W_1+W_2}{W_1+W_2+V_3} \sim Beta(\frac{n_1+n_2-2}{2}, \frac{n_3-1}{2}),$$
$$B_3 = \frac{W_1+W_2+V_3}{W_1+W_2+V_3+V_4} \sim Beta(\frac{n_1+n_2+n_3-3}{2}, \frac{n_4-1}{2}),$$

and

$$X = W_1+W_2+V_3+V_4 \sim \chi^2_{n_1+n_2+n_3+n_4-4},$$

show that the generalized p-value can be expressed as an average of t probabilities as

$$p = 1 - EF_T\left(\frac{(\Delta y - 4\nu_0)\sqrt{(n_1+n_2+n_3+n_4-1)}}{\sqrt{\frac{s_{11}}{n_1B_1B_2B_3}+\frac{s_{22}}{n_2(1-B_1)B_2B_3}+\frac{2s_{32}}{n_3(1-B_2)B_3}+\frac{2s_{42}}{n_4(1-B_3)}}}\right), \quad (9.48)$$

where F_T is the cdf of the t distribution with $n_1+n_2+n_3+n_4-4$ degrees of freedom and the expectation is taken with respect to the Beta random variables B_1, B_2, and B_2.

Exercise 9 *Express the test based on the generalized p-value (9.47) as a generalized t-test.*

Exercise 10 *Find a generalized pivotal quantity for constructing interval estimates for the difference in carryover effects $\eta = \lambda_A - \lambda_B$. Construct 95% generalized confidence intervals for η based on the generalized pivotal.*

Exercise 11 *Consider the two treatment, four sequence Balaam design and assume model (9.13). Establish generalized tests and generalized confidence intervals for the treatment effects μ_A and μ_B.*

Exercise 12 *Consider again the two treatment, four sequence Balaam design and assume model (9.13). Establish generalized tests and generalized confidence intervals for the carryover effects λ_A and λ_B.*

Exercise 13 *Consider again the two treatment, four sequence Balaam design and assume model (9.13). Assuming that $\Sigma_1 = \Sigma_2$, establish procedures for making inferences about the difference in treatment means and carryover effects.*

Exercise 14 *Consider the data set in Table 9.7. If $\Sigma_1 = \Sigma_2$,*

(a) test the hypothesis that there is no difference in treatment means,
(b) construct 95% equal-tail generalized confidence intervals for the difference in treatment means,
(c) construct a 95% equal-tail generalized confidence interval for the difference in carryover effects.

Exercise 15 *Consider again the data set in Table 9.7. Without assuming that the covariance matrices Σ_1 and Σ_2 are equal*

(a) test the hypothesis that there is no difference in carryover effects,
(b) construct 95% left-sided generalized confidence intervals for the difference in treatment means,
(c) construct a 95% right-sided generalized confidence interval for the difference in carryover effects.

Chapter 10

Growth Curves

10.1 Introduction

What is commonly known as growth curves in statistical literature is a special class of multivariate models with a special covariance structure. Some approaches taken in growth curves also lead to a special class of mixed models. In biopharmaceutical applications they deal with groups of subjects observed over time. Hence, this is also a problem of repeated measures. However, in the particular class of growth curves, a certain parametric model is assumed for the growth of the response variable, the quantity on which measurements are taken. The change of the response variable over time is modeled by means of a design matrix, a polynomial growth curve, in particular. To be specific, first consider the case of one group of subjects, such as a cohort of babies observed over time. The observed quantity tracked over time in this case can be the heights or weights of babies. When one studies the growth curves by groups of subjects, the groups in this case might be defined in terms of sex, ethnicity, geographical area, hospital, and so on.

Consider a set of repeated observations taken from subjects of an experiment at a set of common time points. The time points are not necessarily equally spaced. For example, in observing a cohort of babies over time, one could first take the observations every week, then every month, and finally every year. Table 10.1 below (from Elston and Grizzle, 1962) provides a widely referred to and widely analyzed example in the literature on growth curves. In this example a group of 20 boys were observed at four ages and the quantity of interest was their ramus heights.

The response variable here is the ramus height and the problem is to make statistical inference about the growth of ramus heights as a function of age and one may wish to model its growth by a polynomial of some order. Figure 10.1 provides a profile plot of the response variable, the ramus height, as a function of the age of sample subjects. It seems that in this application, a linear growth curve is appropriate for the average child as well as for individual children.

Table 10.1: Ramus heights (in mm) of 20 boys

Boy No.	**Age (years)**			
	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
3	46.3	46.8	47.8	48.5
4	45.1	45.3	46.1	47.2
5	47.6	48.5	48.9	49.3
6	52.5	53.2	53.3	53.7
7	51.2	53.0	54.3	54.5
8	49.8	50.0	50.3	52.7
9	48.1	50.8	52.3	54.4
10	45.0	47.0	47.3	48.3
11	51.2	51.4	51.6	51.9
12	48.5	49.2	53.0	55.5
13	52.1	52.8	53.7	55.0
14	48.2	48.9	49.3	49.8
15	49.6	50.4	51.2	51.8
16	50.7	51.7	52.7	53.3
17	47.2	47.7	48.4	49.5
18	53.3	54.6	55.1	55.3
19	46.2	47.5	48.1	48.4
20	46.3	47.6	51.3	51.8

Moreover, in this application the slope as well as the intercept of growth curves seem to vary substantially from one individual to another, something that should be modeled in analyzing the data.

In the above example there is only one group of subjects under study. More generally we may have to deal with a number of groups of subjects studied over time and the problem of interest might be to compare the mean growth curves of groups. Table 10.2, a data set from Grizzle and Allen (1969) reproduced below, provides an example of a growth curves problem involving four groups of subjects observed at 7 time points. In this example, the data represents the coronary sinus potassium-mil equivalents per liter by time measured in minutes after occlusion. The Group 1 is the control group and groups are three treatment groups as described in Grizzle and Allen (1969).

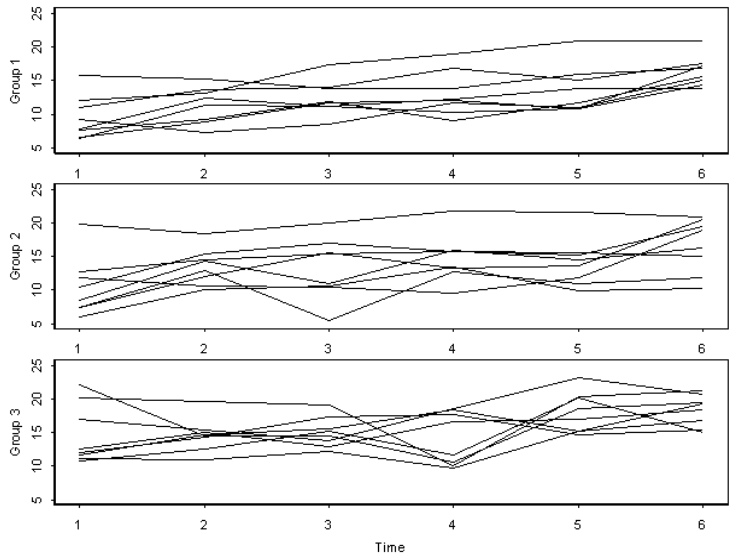

Figure 10.1: Ramus heights by age of childred

Table 10.2 (continued): Treatment dog response after occlusion

	Time(minutes):	*1*	*3*	*5*	*7*	*9*	*11*	*13*
Dog	***Group***	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7
10	*2*	*3.4*	*3.4*	*3.5*	*3.1*	*3.1*	*3.7*	*3.3*
11	*2*	*3.0*	*3.2*	*3.0*	*3.0*	*3.1*	*3.2*	*3.1*
12	*2*	*3.0*	*3.1*	*3.2*	*3.0*	*3.3*	*3.0*	*3.0*
13	*2*	*3.1*	*3.2*	*3.2*	*3.2*	*3.3*	*3.1*	*3.1*
14	*2*	*3.8*	*3.9*	*4.0*	*2.9*	*3.5*	*3.5*	*3.4*
15	*2*	*3.0*	*3.6*	*3.2*	*3.1*	*3.0*	*3.0*	*3.0*
16	*2*	*3.3*	*3.3*	*3.3*	*3.4*	*3.6*	*3.1*	*3.1*
17	*2*	*4.2*	*4.0*	*4.2*	*4.1*	*4.2*	*4.0*	*4.0*
18	*2*	*4.1*	*4.2*	*4.3*	*4.3*	*4.2*	*4.0*	*4.2*
19	*2*	*4.5*	*4.4*	*4.3*	*4.5*	*5.3*	*4.4*	*4.4*
20	*3*	*3.2*	*3.3*	*3.8*	*3.8*	*4.4*	*4.2*	*3.7*
21	*3*	*3.3*	*3.4*	*3.4*	*3.7*	*3.7*	*3.6*	*3.7*
22	*3*	*3.1*	*3.2*	*3.2*	*3.1*	*3.2*	*3.1*	*3.1*
23	*3*	*3.6*	*3.5*	*3.5*	*4.6*	*4.9*	*5.2*	*4.4*
24	*3*	*4.5*	*5.4*	*5.4*	*5.7*	*4.9*	*4.0*	*4.0*
25	*3*	*3.7*	*4.4*	*4.4*	*4.2*	*4.6*	*4.8*	*5.4*
26	*3*	*3.5*	*5.8*	*5.8*	*5.4*	*4.9*	*5.3*	*5.6*
27	*3*	*3.9*	*4.1*	*4.1*	*5.0*	*5.4*	*4.4*	*3.9*
28	*4*	*3.1*	*3.5*	*3.5*	*3.2*	*3.0*	*3.0*	*3.2*
29	*4*	*3.3*	*3.2*	*3.6*	*3.7*	*3.7*	*4.2*	*4.4*
30	*4*	*3.5*	*3.9*	*4.7*	*4.3*	*3.9*	*3.4*	*3.5*
31	*4*	*3.4*	*3.4*	*3.5*	*3.3*	*3.4*	*3.2*	*3.4*
32	*4*	*3.7*	*3.8*	*4.2*	*4.3*	*3.6*	*3.8*	*3.7*
33	*4*	*4.0*	*4.6*	*4.8*	*4.9*	*5.4*	*5.6*	*4.8*
34	*4*	*4.2*	*3.9*	*4.5*	*4.7*	*3.9*	*3.8*	*3.7*
35	*4*	*4.1*	*4.1*	*3.7*	*4.0*	*4.1*	*4.6*	*4.7*
36	*4*	*3.5*	*3.6*	*3.6*	*4.2*	*4.8*	*4.9*	*5.0*

Table 10.2: Control dog response after coronary occlusion

	Time(minutes):	1	3	5	7	9	11	13
Dog	**Group**	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7
1	1	4.0	4.0	4.1	3.6	3.6	3.8	3.1
2	1	4.2	4.3	3.7	3.7	4.8	5.0	5.2
3	1	4.3	4.2	4.3	4.3	4.5	5.8	5.4
4	1	4.2	4.4	4.6	4.9	5.3	5.6	4.9
5	1	4.6	4.4	5.3	5.6	5.9	5.9	5.3
6	1	3.1	3.6	4.9	5.2	5.3	4.2	4.1
7	1	3.7	3.9	3.9	4.8	5.2	5.4	4.2
8	1	4.3	4.2	4.4	5.2	5.6	5.4	4.7
9	1	4.6	4.6	4.4	4.6	5.4	5.9	5.6

Growth curves can be analyzed under alternative assumptions on the covariance structure. For a discussion of various covariance structures and resulting procedures, the reader is referred to Lee (1982, 1991). Analysis of growth curves under unstructured covariance structures were first studied by such authors as Potthoff and Roy (1964), Rao (1965, 1967), and Grizzle and Allen (1969). In that treatment, the model structure and analysis were carried out in the context of Multivariate Analysis (MANOVA) and Generalized Multivariate Analysis of Variance (GMANOVA). Later developments such as that in Lindley and Smith (1972), Fearn (1975), and Laird and Ware (1982) took a different approach involving random coefficient regression models and mixed effects models. This chapter provides an overview of each of these models with greater details on the latter approach leading to structured covariance matrices, a class of problems requiring further research in which generalized inference has the promise to yield solutions with better power and size performance.

10.2 Growth Curve Models

Consider one or more groups of subjects or experimental units on which we have a set of repeated measures on the response variable. The growth of the response variable is to be modeled in terms of a design matrix. In some applications the design matrix may be formed in terms of the time t, at which the measurement is taken. Suppose there are N subjects and they are observed at T time points, say $t_1, t_2, \ldots, t_T$, which are not necessarily equally spaced. Let $Y_{it}, i = 1, \ldots, N$; $t = t_1, t_2, \ldots, t_T$ denote the observation taken on subject i at time t. Let

$\mathbf{Y}_i$ *be the $T \times 1$ vector of responses obtained from subject i,*

$\mathbf{Y}$ *be the $N \times T$ matrix of all responses,*

$\mathbf{B}_t$ *be a $p \times 1$ vector of covariates at time point t,*

$\mathbf{X} = \mathbf{B}'$ *be the $T \times p$ within subject design matrix constructed from covariates. For example, in polynomial growth curves $\mathbf{B}_t^{'} = (1, t, t^2, \ldots, t^{p-1})$ and in turn $\mathbf{X}$ is constructed with the times $t_1, t_2, \ldots, t_T$ that t takes on.*

Assuming a linear regression model for the growth curve of each subject we have

$$\mathbf{Y}_i = \mathbf{X}\boldsymbol{\beta}_i + \epsilon_i \ , \tag{10.1}$$

where $\boldsymbol{\beta}_i$ is a $p \times 1$ vector of unknown parameters, the coefficients of regression models and ϵ_i, $i = 1, \ldots, N$, are N independent $T \times 1$ error vectors with some covariance matrix Σ. The error vectors are assumed to be normally distributed. If the parameter vectors are all different with no structure relating them, then this is just a problem that can be handled by conventional regression procedures. In the context of growth curves, the coefficient vectors have a certain structure. For example, they may all be the same or same within all subjects within a group but possibly different among groups.

Being a special case of repeated measures problems that we discussed in previous chapters, different subjects could belong to the same or different treatment groups. So, we need to setup a between subject design matrix to represent such structures and incorporate it in the model. Potthoff and Roy (1964) formulated a general model with a matrix $\mathbf{A}$ representing the between subject design matrix as

$$\begin{aligned} \mathbf{Y} &= \mathbf{A}\boldsymbol{\gamma}\mathbf{B} + \boldsymbol{\epsilon} \\ &= \mathbf{A}\boldsymbol{\gamma}\mathbf{X}' + \boldsymbol{\epsilon} \ , \end{aligned} \tag{10.2}$$

where $\mathbf{Y} = (\mathbf{Y}_1', \mathbf{Y}_2', \ldots, \mathbf{Y}_N')'$, $\mathbf{A}$ is a full rank matrix of dimension $N \times q$, and $\boldsymbol{\gamma}$ is a $q \times p$ matrix of parameters formed by $\boldsymbol{\beta}_i$ vectors. This model is known as the Generalized Multivariate Analysis of Variance Model and is abbreviated as GMANOVA. With the distributional assumption for error terms made above, we have the matrix normal distribution,

$$\boldsymbol{\epsilon} \sim N_{NT}(0, \Sigma \otimes I_N) \ . \tag{10.3}$$

With polynomial growth curves, the matrix $\mathbf{X}$ has the form

$$\mathbf{X}' = \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ t_1 & t_2 & t_3 & \ldots & t_T \\ t_1^2 & t_2^2 & t_3^2 & \ldots & t_T^2 \\ t_1^3 & t_2^3 & t_3^3 & \ldots & t_T^3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t_1^{p-1} & t_2^{p-1} & t_3^{p-1} & \ldots & t_T^{p-1} \end{pmatrix} \ . \tag{10.4}$$

In general, elements of $\mathbf{X}$ are regressors. Its elements could also be binary indicators representing within subject main effects and interactions.. If $p = T$ and $\mathbf{X} = \mathbf{I}$, any GMANOVA problem reduces to a MANOVA problem, and if $p < T$, then the model has a fewer number of unknown parameters compared to a MANOVA due to the assumed structure.

The above models could be used in formulating inference procedures concerning parameter vectors of one group or a number of groups of subjects. In case

of repeated measurements from a random sample of subjects following a single growth curve, model (10.1) reduces to

$$\mathbf{Y}_i = \mathbf{X}\boldsymbol{\beta} + \epsilon_i \quad \text{for all } i. \tag{10.5}$$

In this case, the model can be rewritten in form (10.2) by defining

$$\mathbf{A} = \mathbf{1} \quad \text{and} \quad \boldsymbol{\gamma} = \boldsymbol{\beta}', \tag{10.6}$$

where $\mathbf{1}$ *is a* $N \times 1$ *vector of 1s. Even in the one group case,* $\mathbf{A}$ *could be a matrix of data on a set of explanatory variables. For example, if a is an explanatory variable representing the age of subjects, then* $\mathbf{A}$ *is a matrix of the form*

$$\mathbf{A} = \begin{pmatrix} 1 & a_1 \\ 1 & a_2 \\ \vdots & \vdots \\ 1 & a_n \end{pmatrix}$$

To provide the form of $\mathbf{A}$ *and* $\boldsymbol{\gamma}$ *for problems of comparing parameter vectors of a number of groups, suppose there are G groups and gth group has* n_g *subjects so that* $\sum n_g = N$*. Assume without loss of generality that the index i for subjects belonging to a certain group occurs next to each other. Then, the equivalent two forms of the model can be written explicitly. Model (10.1) can then be written as*

$$\mathbf{Y}_i = \mathbf{X}\boldsymbol{\beta}_g + \epsilon_i \quad \text{for } i \epsilon g \tag{10.7}$$

and model (10.2) remains valid with $q = G$*, when the between-subject design matrix and the parameter matrix defined are as*

$$\mathbf{A} = \begin{pmatrix} \mathbf{1}_{n_1} & \mathbf{0}_{n_1} & \dots & \mathbf{0}_{n_1} \\ \mathbf{0}_{n_2} & \mathbf{1}_{n_2} & \dots & \mathbf{0}_{n_2} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0}_{n_G} & \mathbf{0}_{n_G} & \dots & \mathbf{1}_{n_G} \end{pmatrix}, \tag{10.8}$$

and

$$\boldsymbol{\gamma} = \begin{pmatrix} \boldsymbol{\beta}_1' \\ \boldsymbol{\beta}_2' \\ \vdots \\ \boldsymbol{\beta}_G' \end{pmatrix},$$

where $\mathbf{1}_{n_i}$ *is a* $n_i \times 1$ *vector of 1s and* $\mathbf{0}_{n_i}$ *is a* $n_i \times 1$ *vector of* 0*'s. Here also* $\mathbf{A}$ *is allowed to be quite general. For example, in place of* $\mathbf{1}_{n_i}$ *in (10.8) we could use matrices of data on some explanatory variables.*

10.3 Inference with Unstructured Covariances

Potthoff and Roy (1964), Khatri(1966), Rao (1965, 1967), and Grizzle and Allen (1969) were among pioneering researchers who developed inference methods for the above models. Here we provide an overview of some of the main results and the reader is referred to the above articles for details and for additional results.

First consider the problem of point estimation of parameters of model (10.2) with the distribution of the error term $\boldsymbol{\epsilon}$ *given by (10.3). Here we do not address the question of prediction based on growth curves. The readers interested in prediction methods in growth curves are referred to Lee and Geisser (1972a, 1972b).*

Define

$$\overline{Y}_{gt} = \frac{\sum_{i \in g} Y_{it}}{n_g}$$

and

$$\overline{\mathbf{Y}}_g = \begin{pmatrix} \overline{Y}_{g1} \\ \overline{Y}_{g2} \\ \vdots \\ \overline{Y}_{gT} \end{pmatrix}.$$

In the particular case of model (10.7), it is intuitive that the covariance matrix Σ *can be estimated by the pooled sample covariance matrix* $\mathbf{S}/(N-G)$*, which takes advantage of the data available from all* G *groups. In turn, the parameter* $\boldsymbol{\gamma}$ *matrix can be estimated by GLSE as*

$$\widehat{\boldsymbol{\gamma}} = \overline{\mathbf{Y}}\mathbf{S}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{S}^{-1}\mathbf{X})^{-1}, \tag{10.9}$$

where $\mathbf{S} = \sum_{g=1}^{G} \sum_{i \epsilon g} (\mathbf{Y}_i - \overline{\mathbf{Y}}_g)(\mathbf{Y}_i - \overline{\mathbf{Y}}_g)'$ *is a* $T \times T$ *matrix and*

$$\overline{\mathbf{Y}} = \begin{pmatrix} \overline{\mathbf{Y}}'_1 \\ \overline{\mathbf{Y}}'_2 \\ \vdots \\ \overline{\mathbf{Y}}'_G \end{pmatrix}, \tag{10.10}$$

is a $G \times T$ *matrix of* $T \times 1$ *sample mean vectors* $\overline{\mathbf{Y}}_g$ *computed from independent observation from* n_j *subjects in each of the* G *groups.*

Example 10.1*. The point estimates of growth curve parameters*

Consider the data set in Table 10.3 involving 3 groups of subjects observed at 6 time points. Suppose linear growth curves are adequate for each of the 3 groups.

Table 10.3: Responses of 3 groups of subjects observed at 6 time points

	Time:	t_1	t_2	t_3	t_4	t_5	t_6
Group	**Subject**	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
1	1	15.74	15.22	13.85	13.79	15.95	16.75
1	2	6.53	8.85	11.74	11.99	10.84	14.41
1	3	11.99	13.1	17.3	18.93	20.95	20.89
1	4	9.21	7.28	8.56	11.76	10.97	15.16
1	5	6.43	11.26	11.15	10.24	10.76	17.13
1	6	7.81	12.49	11.13	12.22	13.89	13.88
1	7	10.98	13.69	13.95	16.75	15.08	17.52
1	8	7.59	9.24	11.96	9.06	11.67	15.55
2	9	11.88	10.61	10.38	9.56	11.88	19.02
2	10	12.64	14.55	15.31	15.76	15.2	19.5
2	11	5.94	10.09	10.66	13.44	9.88	10.24
2	12	19.91	18.4	20.05	21.71	21.61	20.9
2	13	7.45	12.03	15.59	13.33	13.56	20.6
2	14	8.44	14.42	10.87	15.93	14.56	16.21
2	15	10.44	15.4	16.99	15.66	15.55	15.04
2	16	7.37	12.95	5.47	12.65	10.9	11.82
3	17	16.96	15.34	13.75	18.57	23.11	20.75
3	18	11.6	14.65	15.51	18.44	15.13	19.23
3	19	11.07	10.94	12.19	9.65	15.19	16.73
3	20	20.21	19.54	19.03	10.01	20.4	21.15
3	21	12.03	14.39	17.39	17.61	14.64	15.3
3	22	10.8	12.52	15.24	11.64	20.17	15.06
3	23	22.02	14.53	14.43	10.55	18.63	19.39
3	24	12.47	15.01	12.92	16.58	17.01	18.41

The means for the each group at the 6 time points are the pooled covariance matrix are shown below:

$$\overline{\mathbf{Y}} = \begin{pmatrix} 9.535 & 11.3912 & 12.455 & 13.0925 & 13.7638 & 16.4112 \\ 10.5088 & 13.5563 & 13.165 & 14.755 & 14.1425 & 16.6662 \\ 14.645 & 14.615 & 15.0575 & 14.1313 & 18.035 & 18.2525 \end{pmatrix},$$

$$\frac{\mathbf{S}}{21} = \begin{pmatrix} 16.688 & 7.44417 & 7.27353 & 3.79836 & 9.85837 & 7.61947 \\ 7.44417 & 6.99782 & 5.74013 & 4.92179 & 6.36276 & 3.90188 \\ 7.27353 & 5.74013 & 11.1629 & 6.00876 & 7.62439 & 5.9018 \\ 3.79836 & 4.92179 & 6.00876 & 13.1102 & 6.81696 & 3.48661 \\ 9.85837 & 6.36276 & 7.62439 & 6.81696 & 11.7371 & 6.44863 \\ 7.61947 & 3.90188 & 5.9018 & 3.48661 & 6.44863 & 8.92281 \end{pmatrix}.$$

Then the 3×2 *matrix of growth curves parameters,*$\boldsymbol{\gamma}$ *estimated by applying (10.9) is*

$$\widehat{\boldsymbol{\gamma}} = \begin{pmatrix} 9.40283 & 1.15992 \\ 12.0991 & 0.88738 \\ 12.0985 & 0.95336 \end{pmatrix}.$$

More generally, Potthoff and Roy (1964) provided a class of estimates, and Rao (1965, 1967), and Khatri(1966) independently provided specific estimates such as that in (10.9) for the growth parameter vector $\boldsymbol{\gamma}$ and also for the covariance matrix Σ of the model (10.2). The results of Potthoff and Roy (1964) lead to exact solutions only if LSE instead of GLSE is used. It could also be employed to deduce (10.9), but the two approaches lead to different test statistics. Here we confine our attention to only the exact solutions provided by Rao (1965, 1967), Khatri (1966), and Grizzle and Allen (1967), which can be presented in a unified and general manner.

10.3.1 Case of general A design matrix

Formula (10.9) is valid only for the special A matrix defined by (10.8). Khatri (1966) and Rao (1965, 1967) obtained the maximum likelihood estimate (MLE) of $\boldsymbol{\gamma}$ with a general A matrix. As further illustrated by Grizzle and Allen (1967), the estimate can be expressed as

$$\widehat{\boldsymbol{\gamma}} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}\mathbf{S}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{S}^{-1}\mathbf{X})^{-1}, \tag{10.11}$$

where

$$\mathbf{S} = \mathbf{Y}'[\mathbf{I}_N - \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}']\mathbf{Y}. \tag{10.12}$$

A proof of this result is given in Appendix B.2. In terms of the $\mathbf{S}$ matrix, the MLE of Σ is obtained as

$$\widehat{\Sigma} = \frac{1}{N}[\mathbf{S} + \mathbf{W}'\mathbf{Y}'\mathbf{A}(\mathbf{A}'\mathbf{A})\mathbf{A}'\mathbf{Y}\mathbf{W}], \tag{10.13}$$

where

$$\mathbf{W} = \mathbf{I}_N - \mathbf{S}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{S}^{-1}\mathbf{X}')^{-1}\mathbf{X}'. \tag{10.14}$$

Moreover, Gleser and Olkin (1970) showed that $\widehat{\boldsymbol{\gamma}}$ and $\widehat{\Sigma}$ are sufficient statistics for making inferences about the parameters $\boldsymbol{\gamma}$ and Σ. If unbiased estimates are desired, then N appearing in the above equations should be replaced by $N-q$. In particular, with model (10.7) and $q = G$, an unbiased estimate of the covariance can be obtained as

$$\widehat{\Sigma} = \frac{\mathbf{S}}{N-G}. \tag{10.15}$$

It is easily verified that this is the usual pooled sampled covariance matrix computed from data from all G groups. Moreover, in this case notice that

$$\mathbf{A}'\mathbf{A} = \begin{pmatrix} n_1 & 0 & \cdots & 0 \\ 0 & n_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & n_G \end{pmatrix}$$

and

$$\mathbf{A}'\mathbf{Y} = \begin{pmatrix} n_1\overline{Y}_{11} & n_1\overline{Y}_{12} & \ldots & n_1\overline{Y}_{1T} \\ n_2\overline{Y}_{21} & n_2\overline{Y}_{22} & \ldots & n_2\overline{Y}_{2T} \\ \vdots & \vdots & \vdots & \vdots \\ n_G\overline{Y}_{G1} & n_G\overline{Y}_{G2} & \ldots & n_G\overline{Y}_{GT} \end{pmatrix},$$

implying, in particular, that

$$(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y} = \overline{\mathbf{Y}} \tag{10.16}$$

and hence $\widehat{\boldsymbol{\gamma}}$ given by (10.11) reduces to GLSE given by (10.9).

10.4 Inferences on general linear contrasts

Now consider the problems of constructing confidence regions and testing the parameters of the general model (10.2) with unstructured covariances. As in Chapter 6, most problems of practical importance in this context involving parameters among growth curves and within growth curves could be handled by considering double linear combinations of the form

$$\boldsymbol{\theta} = \mathbf{C}\boldsymbol{\gamma}\mathbf{D}\ , \tag{10.17}$$

where $\mathbf{C}$ is a $c \times G$ matrix of known constants and $\mathbf{D}$ is a $p \times d$ matrix of known constants, constructed using desired individual contrasts, where $c \leq G$ and $d \leq p$. For example, if the equality of growth curves, say

$$H_0 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \cdots = \boldsymbol{\beta}_G \tag{10.18}$$

is the hypothesis of interest we can define the required two matrices as

$$\mathbf{C} = \begin{pmatrix} 1 & -1 & 0 & \ldots & 0 \\ 0 & 1 & -1 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & 0 & \ldots & -1 \end{pmatrix}, \qquad \mathbf{D} = \mathbf{I}_p,$$

or as

$$\mathbf{C} = \begin{pmatrix} \mathbf{1}_{G-1} & -\mathbf{I}_{G-1} \end{pmatrix}, \qquad \mathbf{D} = \mathbf{I}_p,$$

where $\mathbf{1}_{G-1}$ *is a* $G-1\times 1$ *vector of* 1*'s. In this case,* $c=G-1$ *and* $d=p$. *Similarly to test the hypothesis of parallel polynomial growth curves with* $\mathbf{X}$ *defined by (10.4) we can use (10.17) with*

$$\mathbf{C}=\begin{pmatrix}\mathbf{1}_{G-1} & -\mathbf{I}_{G-1}\end{pmatrix},\qquad \mathbf{D}=\begin{pmatrix}0\\ \mathbf{I}_{p-1}\end{pmatrix} \tag{10.19}$$

and to make inferences concerning an individual parameter vector, say $\boldsymbol{\beta}_g$, *we can define*

$$\mathbf{C}=\mathbf{i}_g',\qquad \mathbf{D}=\mathbf{I}_p, \tag{10.20}$$

where $\mathbf{i}_g$ *is a* $G\times 1$ *vector having a* 1 *as the* g*th element and* 0*'s elsewhere. In the latter case, we have* $c=1$ *and* $d=p$.

10.4.1 Exact likelihood ratio test

Khatri (1966) extended the MANOVA likelihood ratio test to handle hypotheses of the form

$$H_0:\mathbf{C}\boldsymbol{\gamma}\mathbf{D}=\mathbf{0}. \tag{10.21}$$

At level α, *the hypothesis is to be rejected if the likelihood ratio is*

$$\Lambda=\frac{|\mathbf{E}|}{|\mathbf{E}+\mathbf{H}|}<\lambda_\alpha, \tag{10.22}$$

where

$$\mathbf{E}=\mathbf{D}'(\mathbf{X}'\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{D},$$
$$\mathbf{H}=(\mathbf{D}'\widehat{\boldsymbol{\gamma}}'\mathbf{C}')\mathbf{F}^{-1}(\mathbf{C}\widehat{\boldsymbol{\gamma}}\mathbf{D}),$$

λ_α *is the critical point that must be computed using the distribution of* Λ, *and the* $\mathbf{F}$ *matrix is defined as*

$$\mathbf{F}=\mathbf{C}[(\mathbf{A}'\mathbf{A})^{-1}+\mathbf{Y}_\mathbf{A}(\mathbf{S}^{-1}-\mathbf{S}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{S}^{-1})\mathbf{Y}_\mathbf{A}']\mathbf{C}', \tag{10.23}$$

where $\mathbf{Y}_A=(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}$. *Recall that in the particular case of MANOVA with unstructured covariances,* $\mathbf{X}=\mathbf{D}=\mathbf{I}$ *and* $\mathbf{F}$ *reduces to* $\mathbf{F}=\mathbf{C}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{C}'$. *Although one may use the asymptotic distribution of* Λ *in this case as well, there is no need to resort to such approximations. With a bit of coding in SAS or SPlus, one can carry out exact tests using the* U *distribution of the random variable* Λ. *Without any coding, one can perform the test using the XPro software package.*

That Λ *has a* U *distribution under the null hypothesis, as was the case in the MANOVA problem considered in Chapter 6, follows from the fact that* $\mathbf{E}$ *and* $\mathbf{H}$ *matrices having* d*-dimensional central Wishart distributions,*

$$\begin{aligned}\mathbf{E} &= \mathbf{D}'(\mathbf{X}'\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{D}\ \tilde{}\ W_d(e,\mathbf{D}'(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{D})\\ \mathbf{H} &= (\mathbf{D}'\widehat{\boldsymbol{\gamma}}'\mathbf{C}')\mathbf{F}^{-1}(\mathbf{C}\widehat{\boldsymbol{\gamma}}\mathbf{D})\ \tilde{}\ W_d(h,\mathbf{D}'(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{D}),\end{aligned} \tag{10.24}$$

where $h = c$ and $e = N - T - q + p$, provided that $d \le \min(h, e)$. As in the MANOVA, the necessary quantiles λ_α can now be found using the U distribution,

$$\Lambda \sim U_{d,h,e}, \tag{10.25}$$

If $c < d$, the required condition $d \le \min(h, e)$ will be violated. In this case, the problem can be still handled by re-parametrization based on the fact that the distributions $U_{d,h,e}$ and $U_{h,d,e+h-d}$ are identical. Some authors set $h = \min(c, d)$ and provide alternative ways of evaluating the distribution of Λ.

As before, the probabilities and quantiles of the U distribution are calculated by expressing it as a product of independent beta random variables as

$$U = B_1 B_2 \cdots B_d \ , \text{ where } B_j \sim Beta(\frac{e-j+1}{2}, \frac{h}{2}) \ , \tag{10.26}$$

provided that $e \ge d$. In testing the equality of G growth curves, we use the U distribution,

$$\Lambda \sim U_{p,G-1,N-G-T+p} \ .$$

In testing whether or not the G growth curves are parallel, we use the U distribution,

$$\Lambda \sim U_{p-1,G-1,N-G-T+p} \ .$$

Also recall that when $d \le 2$ or $h \le 2$, the U distribution can be transformed into an F distribution:

$$\text{When } d = 1, \quad \frac{1-\Lambda}{\Lambda}\frac{e}{h} \sim F_{h,e} \ . \tag{10.27}$$

$$\text{When } d = 2 \text{ and } h \ge 2, \quad \frac{1-\Lambda^{1/2}}{\Lambda^{1/2}}\frac{e-1}{h} \sim F_{2h,2(e-1)} \ . \tag{10.28}$$

$$\text{When } h = 1, \quad \frac{1-\Lambda}{\Lambda}\frac{e+1-d}{d} \sim F_{d,e+1-d} \ . \tag{10.29}$$

$$\text{When } h = 2 \text{ and } d \ge 2, \quad \frac{1-\Lambda^{1/2}}{\Lambda^{1/2}}\frac{e+1-d}{d} \sim F_{2d,2(e+1-d)} \ . \tag{10.30}$$

For example, if we had planned only the comparisons of the coefficients of the first two groups, we can then set $d = 1$. Similarly, if all pairwise comparisons of only one coefficient, say the slope parameter, from all growth curves had been planned, we can set $c = h = 1$. When d is large, perhaps the best way to compute the p-value is to generate a large number of random numbers from each of the independent beta distributions and compute the fraction of times that the inequality in (10.22) is satisfied. Asymptotically the statistic $-(e-(d-h+1)/2)$ $\log(\Lambda)$ has a chi-squared distribution with dh degrees of freedom. Although this approximation is good for certain critical values, in this computer age there is

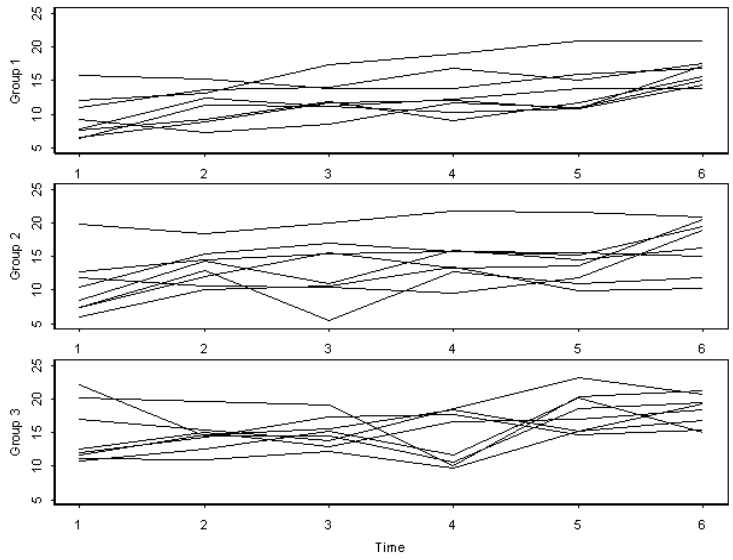

Figure 10.2: Profile plots by time and group

no need to resort to such asymptotic results, especially in computing p-values regardless of the observed value of the statistic.

Finally, the p-value for testing the null hypotheses of the form (10.21) in general and the particular hypothesis (10.18) of special interest can be computed as

$$p = 1 - F_U(\Lambda) \ , \tag{10.31}$$

where F_U is the cdf of the U distribution with d, h and e degrees of freedom. The hypothesis could also be tested using other test statistics used in MANOVA such as the Roy's largest root test statistic, whose distribution could also be derived from that of $\mathbf{E}$ and $\mathbf{H}$. However, except in special cases, close-form solutions are available for the test statistic defined in terms of the likelihood ratio (10.22) only.

Example 10.2. *Comparing growth curve parameters*

Continuing with the example involving 3 groups of subjects, let us to test the hypotheses of equal growth curves and parallel growth curves. In the latter case we set

$$\mathbf{C} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}, \qquad \mathbf{D} = \begin{pmatrix} 0 \\ \mathbf{1} \end{pmatrix}.$$

Figure 10.2 provides profile plots for each of the three groups, as a function of time. It seems that the assumption of linear growth curves is a reasonable one. Despite fair differences in point estimates of growth curve parameters computed

in Example 10.1, given the large variances in within subject and among subject variations, it is not clear from the figure whether they are statistically significant. Each hypothesis can be tested using (10.21) and (10.31). For the problem of testing for identical growth curves, the observed value of the likelihood ratio statistic is 0.8343, with the distribution $\Lambda \sim U_{2,2,17}$ *and hence*

$$F = 8\frac{1-\Lambda^{1/2}}{\Lambda^{1/2}} \sim F_{4,32} \ .$$

The observed value of the F Statistic is 0.7587. Hence the p-value for testing equal growth curves is 0.5599 providing no evidence to reject the null hypothesis. The observed value of the likelihood ratio for testing parallel growth curves, i.e. the equality of the three slopes, is 0.9416. In this case, $\Lambda \sim U_{1,2,17}$ *and*

$$F = 8.5\frac{1-\Lambda}{\Lambda} \sim F_{2,17}$$

with the observed value of the F-statistic 0.5271. Hence, the p-value for testing the hypothesis is .5996, which leads to the same conclusion.

Multiple comparisons of the coefficients of the growth curves can also be carried out with appropriate choice of $\mathbf{C}$ *and* $\mathbf{D}$ *in (10.17). In this example there seem no point in proceeding to multiple comparisons as the hypothesis of equal growth curves was not rejected. As we will see later in this chapter, however, the main reason for insignificance of differences in growth curves is partly due to the assumption of unstructured covariance matrix with too many unknown parameters. Continuing with the analysis for the purpose of illustration, the F statistic based on the U statistic can be employed in testing, since we have only 3 groups to compare. The appropriate values of* $\mathbf{C}$ *and* $\mathbf{D}$ *are set according to the coefficients being compared. For example, to compare the slopes of growth curves 1 and 2 we set*

$$\mathbf{C} = \begin{pmatrix} 1 & -1 & 0 \end{pmatrix}, \qquad \mathbf{D} = \begin{pmatrix} 0 \\ \mathbf{1} \end{pmatrix}.$$

If we are comparing the coefficients of only one pair of growth curves, we can set $c = h = 1$ *and use the F statistic. If we are comparing only one coefficient of any pair of growth curves we can set* $d = 1$ *and again use the F statistic. Some results are summarized in Table 10.4. Multiple comparisons are valid for a pair of growth curves. Despite the fact that these are not even simultaneous tests, as expected, no pair of growth curves have significantly different parameters.*

Table 10.4: Comparing individual coefficients of growth curves

	Testing equal intercepts		Testing equal slopes	
Growth Curves	Observed F	p-value	Observed F	p-value
1 and 2	2.4220	.1381	0.9407	.3457
1 and 3	2.0401	.1713	0.4555	.5088
2 and 3	0.0000	.9998	0.0402	.8436

10.5 Simultaneous Confidence Intervals

The above results can be employed to deduce simultaneous confidence intervals of double linear combinations by considering, $\theta_{\mathbf{ab}} = \mathbf{a}'\mathbf{C}\boldsymbol{\gamma}\mathbf{D}\mathbf{b}$*, where* $\mathbf{a}$ *is a* $c \times 1$ *vector and* $\mathbf{b}$ *is a* $d \times 1$ *vector. As in MANOVA,* $100\gamma\%$ *confidence intervals for a set of linear combinations of parameters can be constructed by alternative approaches.*

Khatri (1966) derived confidence intervals based on the maximum root criterion, which has a direct relationship with simultaneous confidence intervals. Let $\alpha = 1 - \gamma$ *be the corresponding critical level. Then the* $100\gamma\%$ *confidence bounds for* $\theta_{\mathbf{ab}}$ *given by the maximum root method is*

$$\widehat{\theta}_{\mathbf{ab}} \pm \left[\frac{k_\alpha}{1-k_\alpha}(\mathbf{a}'\mathbf{F}\mathbf{a})(\mathbf{b}'\mathbf{E}\mathbf{b})\right]^{1/2},$$

where $\widehat{\theta}_{\mathbf{ab}} = \mathbf{a}'\mathbf{C}\widehat{\boldsymbol{\gamma}}\mathbf{D}\mathbf{b}$*, and* k_α *is the* $(1-\alpha)$*th percentile of the largest root test criterion with degrees of freedom* $\min(c,d)$*,* $(|c-d|-1)/2$*, and* $(e-d-1)/2$*. The confidence level remains valid for any number of intervals that can be deduced from* $\mathbf{C}\boldsymbol{\gamma}\mathbf{D}$*. In two important special cases, namely when* $c = 1$ *or* $d = 1$*,* κ_α *can be obtained from the* F *distribution. In this case, the maximum root-based procedure is the same as the ones discussed below. This is, for instance, the case if we had planned comparing only the parameters of two particular groups only so that* $c = h = 1$*. Similarly, if all pairwise comparisons of only one coefficient, say the slope parameter, from all growth curves had been planned, we can set* $d = 1$*. In these cases, a counterpart based on the* U *statistic also leads to an* F *statistic*

$$F = \frac{e_1}{h_1}\frac{1-U}{U} = \frac{e_1}{h_1}\frac{(\widehat{\theta}_{\mathbf{ab}})^2(\mathbf{a}\mathbf{F}\mathbf{a})^{-1}}{\mathbf{b}'\mathbf{E}\mathbf{b}} \sim F_{h_1,e_1}$$

and the $100\gamma\%$ *confidence intervals*

$$\widehat{\theta}_{\mathbf{ab}} \pm \kappa_\gamma \left[\frac{h_1}{e_1}(\mathbf{a}'\mathbf{F}\mathbf{a})(\mathbf{b}'\mathbf{E}\mathbf{b})\right]^{1/2}, \tag{10.32}$$

where κ_γ *is the* $(1-\gamma)$*th quantile of the* F *distribution with* h_1 *and* e_1 *degrees of freedom and*

$$\begin{aligned} h_1 &= h \quad \text{and} \quad e_1 = e \quad \text{if } d = 1 \\ h_1 &= d \quad \text{and} \quad e_1 = e + 1 - d \quad \text{if } h = 1. \end{aligned}$$

If the F *statistic has already been computed in hypotheses testing, the interval given by (10.32) can be computed using the formula*

$$\widehat{\theta}_{\mathbf{ab}} \pm \kappa_\gamma |\widehat{\theta}_{\mathbf{ab}}| F^{-1/2}.$$

If only a few prespecified contrasts are to be tested, shorter intervals than those given by the above methods can be obtained in terms of the t *distribution*

by applying the Bonferroni method. If r tests or intervals of linear contrasts are pre-planned, then the Bonferroni intervals of $\theta_{\mathbf{ab}}$ are computed as

$$\widehat{\theta}_{\mathbf{ab}} \pm \frac{t_e(1-\frac{\alpha}{2r})}{\sqrt{e}} \left[(\mathbf{a}'\mathbf{F}\mathbf{a})(\mathbf{b}'\mathbf{E}\mathbf{b})\right]^{1/2}, \tag{10.33}$$

where $e = N - q - (T - p)$ is the degrees of freedom of the t distribution.

In computing confidence intervals by applying any of the above methods, we need to be concerned with the matrices $\mathbf{C}$ and $\mathbf{D}$ only to figure out the appropriate degrees of freedom and the contrasts that they can generate. Having done that, the linear combination on which the intervals are to be constructed can be written in terms of $\boldsymbol{\gamma}$ as $\theta = \theta_{\mathbf{ab}} = \mathbf{a}'\mathbf{C}\boldsymbol{\gamma}\mathbf{D}\mathbf{b} = \mathbf{c}'\boldsymbol{\gamma}\mathbf{d}$, where $\mathbf{c} = \mathbf{C}'\mathbf{a}$ is a $G \times 1$ vector and $\mathbf{d} = \mathbf{D}\mathbf{b}$ is a $p \times 1$ vector. If there are no special set of contrasts directly based on the parameters of the growth curves, then $\mathbf{c} = \mathbf{a}$ and $\mathbf{d} = \mathbf{b}$. In either case, if the linear combinations of interest is expressed as $\theta = \mathbf{c}'\boldsymbol{\gamma}\mathbf{d}$, the $100\gamma\%$ confidence intervals given by each of the above methods is of the form

$$\widehat{\theta} \pm c_\gamma \left[(\mathbf{F}_d)(\mathbf{E}_c)\right]^{1/2}, \tag{10.34}$$

where

$$\mathbf{E}_d = \mathbf{d}'(\mathbf{X}'\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{d}$$

and

$$\mathbf{F}_c = \mathbf{c}[(\mathbf{A}'\mathbf{A})^{-1} + \mathbf{Y}_A(\mathbf{S}^{-1} - \mathbf{S}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{S}^{-1})\mathbf{Y}_\mathbf{A}']\mathbf{c}'.$$

With the common problem of comparing the growth curves of a number of groups, namely with model (10.1), $\mathbf{F}_c$ further reduces to

$$\mathbf{F}_c = \sum c_i^2/n_i + \mathbf{c}\overline{\mathbf{Y}}(\mathbf{S}^{-1} - \mathbf{S}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{S}^{-1})\overline{\mathbf{Y}}'\mathbf{c}' \ .$$

For example, if we are interested in the difference in intercepts of the first two growth curves, we apply the above formulae with,

$$\mathbf{c} = (\ 1 \quad -1 \quad 0 \quad \cdots \quad 0)$$

and

$$\mathbf{d} = (\ 1 \quad 0 \quad 0 \quad \cdots \quad 0)'.$$

Example 10.3. *Comparing growth curve parameters (continued)*

Continuing with the example involving 3 groups of subjects, let us now construct confidence intervals for the differences in intercepts and slopes of each pair of growth curves. Table below shows the 95% simultaneous confidence intervals constructed under the assumption that intervals had been planned only for comparing either the intercepts or the slopes. In applying formula (10.33) to ensure the confidence level for all three pairs, we set $r = 3$. As a result, the confidence interval for one interval is 98.33%. Despite the fact that the intervals are constructed to ensure the level of only one parameter (either the slope or the

intercept), all the confidence intervals are still too wide and include 0 providing no support for rejecting hypotheses of equal coefficients. The difference in the two intercepts of growth curves 1 and 2 is one which is close to being significant.

95% Confidence intervals for the difference in parameters

Parameter:	*Intercept difference*		*Slope difference*	
Growth Curves	*Lower limit*	*Upper limit*	*Lower limit*	*Upper limit*
1 and 2	*-7.2944*	*1.9019*	*-0.4733*	*1.0184*
1 and 3	*-7.7041*	*2.3127*	*-0.6058*	*1.0189*
2 and 3	*-5.3871*	*5.3882*	*-0.9398*	*0.8079*

10.5.1 Case of one group

Recall that in the case of one group of subjects, model (10.2) and the foregoing results remain valid. The model is equivalent to (10.5) when we set the parameters as $\mathbf{A} = \mathbf{1}$ *and* $\boldsymbol{\gamma} = \boldsymbol{\beta}'$. *In particular, the point estimates given by (10.9) and (10.11) and reduces to*

$$\widehat{\boldsymbol{\beta}}=\widehat{\boldsymbol{\gamma}}'= (\mathbf{X}'\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{S}^{-1}\overline{\mathbf{y}} \ , \tag{10.35}$$

where $\overline{\mathbf{y}}$ *is the* $T\times 1$ *vector of* T *sample means responses from* n *subjects observed at each of the time points. In this case the* $T \times T$ *matrix* $\mathbf{S}$ *reduces to*

$$\mathbf{S} =\sum_{i=1}^{N}(\mathbf{y}_i - \overline{\mathbf{y}})(\mathbf{y}_i - \overline{\mathbf{y}})' \ . \tag{10.36}$$

Tests of hypotheses of the form

$$H_0 : \mathbf{D}'\boldsymbol{\beta} = \mathbf{0} \tag{10.37}$$

is deduced from (10.21) by setting $\mathbf{A} = \mathbf{1}$ *and* $\mathbf{C} =1$. *In this case the* $\mathbf{H}$ *and* $\mathbf{E}$ *matrices reduces to*

$$\mathbf{E} = \mathbf{D}'(\mathbf{X}'\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{D} \ \tilde{} \ W_d(e, \mathbf{D}'(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{D}) \tag{10.38}$$

$$\mathbf{H} = \frac{(\mathbf{D}'\widehat{\boldsymbol{\gamma}}'\widehat{\boldsymbol{\gamma}}\mathbf{D})}{F} \ \tilde{} \ W_d(1, \mathbf{D}'(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{D}), \tag{10.39}$$

where $e = N - 1 - T + p$ *and* F *is now a scalar parameter given by*

$$F = 1/N+\overline{\mathbf{y}}'(\mathbf{S}^{-1}-\mathbf{S}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{S}^{-1})\overline{\mathbf{y}}.$$

Since $h = 1$ *in this case, the test provided by the* U *statistic reduces to an* F*-test, because under the null hypothesis*

$$\frac{e+1-d}{d}\left(\frac{1-\Lambda}{\Lambda}\right) \sim \mathbf{F}_{d,e+1-d} \tag{10.40}$$

where $\Lambda = |\mathbf{E}|/|\mathbf{E}+\mathbf{H}|$.

10.6 Mixed Models in Growth Curves

The GMANOVA treatment we studied above does not allow the covariance matrix to have any special structure, as we often encounter in many growth curves applications. Yet we assumed that each group under comparison has the same covariance matrix and that the design matrix was the same for all subjects. A more flexible general model which overcome such drawbacks of models of the form (10.2) was introduced by Laird and Ware (1982) in the context of mixed models. The general form of the linear mixed effects growth curves model is

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{Z}_i\mathbf{b}_i + \epsilon_i \qquad \text{for } i = 1, \ldots, N, \tag{10.41}$$

where $\mathbf{y}_i$ is the $T \times 1$ vector of responses from ith subject, $\mathbf{X}_i$ is a known design matrix of dimension $T \times p$, $\mathbf{Z}_i$ is another design matrix of dimension $T \times q$, $\boldsymbol{\beta}_i$ is a vector of fixed effects, and other variables are jointly independent random variables distributed as

$$\mathbf{b}_i \sim \mathbf{N}_q(\mathbf{0}, \boldsymbol{\Psi}) \tag{10.42}$$

and

$$\epsilon_i \sim \mathbf{N}_T(\mathbf{0}, \boldsymbol{\Lambda}_i) \ ,$$

where $\boldsymbol{\Lambda}_i$ is a within-subject covariance matrix of dimension $T \times T$ and $\boldsymbol{\Psi}$ is usually a between-subject covariance matrix of dimension $q \times q$. Of course the model can also be rewritten in the form of a structured covariance matrix as

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{e}_i, \tag{10.43}$$

where

$$\mathbf{e}_i \sim \mathbf{N}_T(\mathbf{0}, \boldsymbol{\Lambda}_i + \mathbf{Z}_i\boldsymbol{\Psi}\mathbf{Z}_i').$$

If the parameters are all different and if the covariance matrices are known except for a scalar parameter, they can be analyzed separately by regression methods. In more important applications of the model, some parameters are common and some are different, especially depending on certain groups that they belong to.

As in the previous section, now consider the particular problem that arise in comparing G groups of subjects. In this case, denoting ith subject in group g by $i(g)$, assuming some common fixed effects for subjects in a single group, we can rewrite model (10.41) as

$$\begin{aligned} \mathbf{y}_{i(g)} &= \mathbf{X}_{i(g)}\boldsymbol{\beta}_g + \mathbf{Z}_{i(g)}\mathbf{b}_{i(g)} + \epsilon_{i(g)} \\ i(g) &= 1, \ldots, n_g; \ g = 1, \ldots, G \ , \end{aligned} \tag{10.44}$$

where the random variables have common within-group distributions,

$$\mathbf{b}_{i(g)} \sim \mathbf{N}_q(\mathbf{0}, \boldsymbol{\Psi}) \tag{10.45}$$

and

$$\epsilon_{i(g)} \sim \mathbf{N}_T(\mathbf{0}, \boldsymbol{\Lambda}_g).$$

The model is equivalent to

$$\mathbf{y}_{i(g)} = \mathbf{X}_{i(g)}\boldsymbol{\beta}_g + \boldsymbol{\epsilon}_{i(g)},$$

where

$$\boldsymbol{\epsilon}_{i(g)} \sim \mathbf{N}_T(\mathbf{0}, \boldsymbol{\Sigma}_g), \quad \boldsymbol{\Sigma}_\mathbf{g} = \boldsymbol{\Lambda}_\mathbf{g} + \mathbf{Z}_{\mathbf{i(g)}}\boldsymbol{\Psi}\mathbf{Z}'_{\mathbf{i(g)}}.$$

If $\boldsymbol{\Sigma}_g$ were known or has been estimated, the parameters of the growth curves are estimated by MLE as

$$\widehat{\boldsymbol{\beta}}_g = \left(\sum_{i\epsilon g}\mathbf{X}'_{i(g)}\boldsymbol{\Sigma}_g^{-1}\mathbf{X}_{i(g)}\right)^{-1}\sum_{i\epsilon g}\mathbf{X}'_{i(g)}\boldsymbol{\Sigma}_g^{-1}\mathbf{y}_{i(g)}. \tag{10.46}$$

Usually it is assumed that the structure of matrix $\boldsymbol{\Sigma}_g$ is known except for a few parameters. For example, with independent residuals it is assumed that $\boldsymbol{\Lambda}_g = \sigma_g^2\mathbf{I}_T$. The between-subject covariance matrix $\boldsymbol{\Psi}$ is treated as unstructured or structured. In the next section we will consider an important case of the problems in which $\boldsymbol{\Psi}$ is a structured and has only one unknown parameter. When it is unstructured, it is estimated by sample covariance using $\widehat{\mathbf{b}}_{i(g)}$ along with other unknown parameters including $\boldsymbol{\beta}_g$ iteratively. The reader is referred to Laird and Ware (1982) for such estimation methods and for asymptotic results concerning other types of inferences. In any case, there are no general results available for exact inference on the parameters of model (10.44) and this is an area requiring much research. Next we consider the problem when each covariance matrix has a particular structure, and provide a solution to a simple, yet very important, class of problems. The solution is obtained by taking the generalized approach.

10.7 Exact Inference under Structured Covariances

When the covariance matrix of a growth curves model has a special structure, classical approaches do not provide exact solutions to inference problems even for a situation of a single growth curve. Hence this is an area requiring much research. In this section, we consider one of the simplest, yet important, class of problems involving one group of subjects and show how the generalized approach could help find exact solutions. Later in this chapter we will consider the problem of comparing a number of treatment groups. Specifically, let us consider the class of compound symmetric covariance matrices, which can be derived from a linear model with one random effect as described below. This is also known as inference problem in growth curves under intraclass correlation structure.

To derive the model under a linear structure with one random effect, consider one group of subjects following a simple linear growth curve model of the form (10.1) [see Weerahandi and Berger (1999) for details], in which the difference in parameter vectors is a single random effect associated with the subjects. For ith subject we have

$$Y_{it} = \alpha_i + \mathbf{X}_t'\boldsymbol{\beta} + \epsilon_{it}, \tag{10.47}$$

where $\mathbf{X}_t' = \mathbf{V}_t$ *is the* $p \times 1$ *design vector,* $\boldsymbol{\beta}$ *is a* $p \times 1$ *vector of parameters common for all subjects,* α_i *is a random effect due to subjects, and* ϵ_{it} *is the error term. Under the usual normality assumption for each random variable, we get*

$$\alpha_i \sim N(0, \sigma_\alpha^2) \tag{10.48}$$

and

$$\epsilon_{it} \sim N(0, \sigma_e^2), \tag{10.49}$$

where σ_α^2 *and* σ_e^2 *are variance components of the model. Moreover,* α_i *and all* ϵ_{it} *terms are assumed to be independently distributed. Collecting data from* i*th subject, the model for the* $T \times 1$ *vector of responses,* $\mathbf{Y}_i$*, can be written in vector form in terms of the* $T \times p$ *design matrix* $\mathbf{X} = \mathbf{V}'$ *as*

$$\mathbf{Y}_i = \alpha_i \mathbf{1}_T + \mathbf{X}\boldsymbol{\beta} + \epsilon_i \ , \tag{10.50}$$

where $\mathbf{1}_T$ *is a* $T \times 1$ *vector of 1s. It is easily seen from (10.48) that* $Var(Y_{it}) = \sigma_\alpha^2 + \sigma_e^2$ *and that* $Cov(Y_{it}, Y_{it'}) = \sigma_\alpha^2$*, and hence*

$$\mathbf{Y}_i \sim N_T(\mathbf{X}\boldsymbol{\beta}, \Sigma) \text{ with the covariance matrix } \Sigma = \sigma_\alpha^2 \mathbf{1}_T \mathbf{1}_T^{'} + \sigma_e^2 \mathbf{I}_T \tag{10.51}$$

This means that the covariance matrix of the observations vector has the intraclass structure, which is also known as the compound symmetric structure. Notice that model (10.50) is a special case of model (10.41) with

$$\mathbf{Z}_i \boldsymbol{\Psi}_i \mathbf{Z}_i' = \sigma_\alpha^2 \mathbf{1}_T \mathbf{1}_T^{'} \qquad \text{and} \qquad \boldsymbol{\Lambda}_i = \sigma_e^2 \mathbf{I}_T$$

Being a matrix with intraclass structure, the inverse Σ *of is also an intraclass matrix. More specifically,*

$$\Sigma^{-1} = \sigma_e^{-2} \left[\mathbf{1}_T - \frac{\sigma_\alpha^2}{\sigma_e^2 + T\sigma_\alpha^2} \mathbf{1}_T \mathbf{1}_T^{'} \right]. \tag{10.52}$$

Lin and Lee (2003) considered model (10.47) under the more general intraclass covariance matrix

$$\Sigma = (\rho_1 \sigma_\alpha^2 + \rho_2 \sigma_e^2) \mathbf{1}_T \mathbf{1}_T^{'} + [(1 - \rho_1)\sigma_\alpha^2 + (1 - \rho_2)\sigma_e^2] \mathbf{I}_T \ , \tag{10.53}$$

where $1 \geq \rho_1 \geq -1/(T-1)$ *and* $1 \geq \rho_2 \geq -1/(T-1)$*. The inverse of this matrix follows from formula (10.52).*

10.7.1 Inference with intraclass correlation structure

The problem is to make inferences about the unknown parameters $\boldsymbol{\beta}$ *and the variance components* σ_α^2 *and* σ_e^2*. It follows from (10.51) that the maximum likelihood estimate (MLE) of* $\boldsymbol{\beta}$ *is the weighted least-squares estimate (WLSE)*

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\overline{\mathbf{Y}}, \tag{10.54}$$

which is also known as the generalized least squares estimate (GLSE) of $\boldsymbol{\beta}$*, where* $\overline{\mathbf{Y}}=\sum \mathbf{Y}_i/N$ *is a* $T \times 1$ *vector. It is easily seen that the distribution of* $\widehat{\boldsymbol{\beta}}$ *is given by*

$$\widehat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta},(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}/\mathbf{N}) \tag{10.55}$$

Rao (1967) showed that, if the columns of $\Sigma\mathbf{X}$ *is a subspace of the vector space spanned by the columns of* $\mathbf{X}$*, then the GLSE reduces to the ordinary least-squares estimate (OLSE), regardless of what* Σ *is. When* Σ *is as in (10.51) and the first column of* $\mathbf{X}$ *is a vector of* 1*'s (i.e., an intercept term is present in the growth curve model), this condition is satisfied and consequently (10.54) reduces to the OLSE,*

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\overline{\mathbf{Y}} \tag{10.56}$$

The result is summarized by Theorem 1 below. A direct proof of this result was given by McElroy (1967).

Theorem 1 *If* Σ *is a covariance matrix with the intraclass structure and if* $\mathbf{X}$ *is design matrix with first column being a vector of 1s, then*

$$(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\overline{\mathbf{Y}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\overline{\mathbf{Y}}.$$

Of course, there is no difficulty constructing confidence regions or performing statistical tests concerning the parameter vector $\boldsymbol{\beta}$ *if* Σ *is a known matrix. But when* σ_α^2 *and* σ_e^2 *are unknown parameters, as usually the case, this is not straightforward. In fact the classical approaches to the problem do not provide exact solutions to making inferences beyond the point estimation.*

Now consider the problem of making exact inferences on coefficients of $\boldsymbol{\beta}$*, including the problems of testing whether one or more of these regression coefficients equals some pre-specified value such as zero or one sided hypotheses such as*

$$H_0 : \beta_j \leq \beta_{j0} \tag{10.57}$$

Application of asymptotic methods for testing this type of hypothesis could lead to results with poor size performance even when the sample is large. To illustrate the generalized approach in testing hypotheses concerning a single slope coefficient of the growth curves, consider the particular hypothesis (10.57). The

search for most powerful tests can be confined to the class of procedures based on the sufficient statistics,

$$\widehat{\beta}_j, \qquad j = 1, \cdots p,$$

$$S_{e1}^2 = \sum_i \sum_t (Y_{it} - \mathbf{X}_t' \widehat{\boldsymbol{\beta}} - (\overline{Y_i} - \overline{Y}))^2,$$

and

$$S_{e2}^2 = T \sum_i (\overline{Y_i} - \overline{Y})^2,$$

where $\widehat{\beta}_j$ *is the* j*th component of* $\widehat{\boldsymbol{\beta}}$ *given by (10.56),* $\overline{Y_i}$ *denotes the sample mean for subject* i *and* $\overline{Y}$ *is the sample mean computed from all* Y_{it} *data. It follows from (10.55) and from the literature on variance components that the distribution of these quantities are given by*

$$\widehat{\beta}_j \sim N(\beta_j, \frac{(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}{}_{jj}}{N}, \tag{10.58}$$

$$W_1 = \frac{S_{e1}^2}{\phi_1^2} \sim \chi_{\nu_1}^2, \qquad \textit{where } \nu_1 = N(T-1) - p + 1, \tag{10.59}$$

and

$$W_2 = \frac{S_{e2}^2}{\phi_2^2} \sim \chi_{\nu_2}^2, \qquad \textit{where } \nu_2 = N - 1 \tag{10.60}$$

where $\phi_1^2 = \sigma_e^2$, $\phi_2^2 = \sigma_e^2 + T\sigma_\alpha^2$ *, and* $(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}{}_{jj}$ *is the* jj*th element of the covariance matrix* $(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}$*. Lin and Lee (2003) pointed out that the distributions in (10.59) and (10.60) remains valid under the more general covariance matrix (10.53) if* ϕ_1^2 *and* ϕ_2^2 *are redefined as*

$$\phi_1^2 = (1 - \rho_1)\sigma_\alpha^2 + (1 - \rho_2)\sigma_e^2,$$

and

$$\phi_2^2 = \phi_1^2 + T(\rho_1 \sigma_\alpha^2 + \rho_2 \sigma_e^2).$$

In terms of the general parameters, define

$$S_j(\phi_1^2, \phi_2^2) = (\frac{(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}{}_{jj}}{n})^{1/2} \tag{10.61}$$

a well defined random variable involving nuisance parameters. Given the two arguments of S_j*, we evaluate it first by computing* Σ *and then performing the matrix operations. To derive a test appropriate for testing (10.57) consider the potential extreme region given by*

$$\left\{ \frac{\widehat{\beta}_j - \beta_j}{S_j(\phi_1^2, \phi_2^2)} \geq \frac{b_j - \beta_j}{S_j(\phi_1^2 \frac{s_{e1}^2}{S_{e1}^2}, \phi_2^2 \frac{s_{e2}^2}{S_{e2}^2})} = \frac{b_j - \beta_j}{S_j(\frac{s_{e1}^2}{W_1}, \frac{s_{e2}^2}{W_2})} \right\}, \tag{10.62}$$

where s_{e1}^2, s_{e2}^2, *and* b_j *are the observed values of* S_{e1}^2, S_{e2}^2, *and* $\widehat{\beta}_j$, *respectively. The inequality appearing in (10.62) could also be expressed as*

$$Z \geq \frac{b_j - \beta_j}{S_j(\frac{s_{e1}^2}{W_1}, \frac{s_{e2}^2}{W_2})}, \qquad \text{with } Z \sim N(0,1), \tag{10.63}$$

Moreover, Z, W_1, W_2 *are mutually independent, and as in other similar applications, we can also use the independent random variables* Z,

$$W = W_1 + W_2 \sim \chi_\nu^2 \tag{10.64}$$

and

$$B \sim Beta\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right)$$

to specify the extreme region, where $\nu = NT - p$. *It follows from the structure of* S_j *that*

$$S_j(k\phi_1^2, k\phi_2^2) = k^{1/2} S_j(\phi_1^2, \phi_2^2),$$

a property we can exploit to factor out W *term from* $S_j(s_{e1}^2/W_1, s_{e2}^2/W_2)$. *Noting that the probability of the above inequality is an increasing function of the parameter of interest, thus making it define a true extreme region, the generalized p-value, the maximum probability of the extreme region, is computed as*

$$p = \Pr\left(t = \frac{Z}{\sqrt{W/\nu}} \geq \sqrt{\nu}\frac{(b_j - \beta_j)}{S_j(\frac{s_{e1}^2}{B}, \frac{s_{e2}^2}{1-B})}\right) \tag{10.65}$$

$$= 1 - E\left\{G_\nu\left[\sqrt{\nu}\frac{(b_j - \beta_j)}{S_j(\frac{s_{e1}^2}{B}, \frac{s_{e2}^2}{1-B})}\right]\right\}, \tag{10.66}$$

where G_ν *is the cumulative distribution function of the* t *distribution with* $\nu = NT - p$ *degrees of freedom and the expectation is taken with respect to the beta random variable in (10.64).*

This p-value could be evaluated by exact numerical integrating using formula (10.66) or by Monte Carlo integration using (10.65). As Weerahandi and Berger(1999) argued, the test obtained in this manner in fact the unique unbiased test based on sufficient statistics. Two-sided $100\gamma\%$ *confidence limits for* β_j *are deduced from (10.66), as in the previous chapters, by equating* p *to* $(1+\gamma)/2$.

Weerahandi and Berger(1999) also discussed the problem of testing hypotheses involving several or all of the parameters. A general test of hypothesis of the form $H_0 : \boldsymbol{\beta}_1 = \mathbf{0}$ *involving a subset of the parameter vector is carried out based on the error sums of squares* S_{e2}^2 *and* S_{e1}^2, *with latter obtained by applying generalized regressions with and without the null hypothesis. Inferences for linear combinations of the parameters of the form* $\theta = \mathbf{k}'\boldsymbol{\beta}$ *could be based on the result*

$$\widehat{\theta} = \mathbf{k}'\widehat{\boldsymbol{\beta}} \sim \mathbf{N}(\theta, \frac{1}{N}\mathbf{k}'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{k}) \tag{10.67}$$

and the above results. Since we can employee S_{e1}^2 *and* S_{e2}^2 *to tackle unknown parameters in (10.67), the corresponding tests could be deduced from the above results.*

Analysis of data from a growth curve model with structured covariance matrix is an area requiring further research. In particular, there is a need to extend the above results to the case when σ_e^2 varies over time. Other extensions of interest include the cases of multiple variance components and AR type covariance structures that arise in time series data.

10.8 Comparing Growth Curves

Now suppose there are a number of treatment groups following possibly different growth curves. For convenience, let Y_{gi_gt}, or rather, suppressing the subscript g in i_g, Y_{git} denote the observation taken from ith subject in gth group at tth time point. Assuming the same model as in the above section for subjects in each group we have

$$Y_{git} = \alpha_{gi} + \mathbf{X}'_{gt}\boldsymbol{\beta}_g + \epsilon_{git}, \tag{10.68}$$
$$g = 1, \ldots, G;\ i = 1, \ldots, n_g,$$

where $\boldsymbol{\beta}_g$'s are the parameter vectors of particular interest. In matrix notation, the counterpart of (10.50) in this case is

$$\mathbf{Y}_{gi} = \alpha_{gi}\mathbf{1}_T + \mathbf{X}_g\boldsymbol{\beta}_g + \epsilon_{gi}\ , \tag{10.69}$$

where

$$\alpha_{gi} \sim N(0, \sigma_\alpha^2),$$

and

$$\epsilon_{git} \sim N(0, \sigma_g^2),$$

for all $g = 1, \ldots, G$; $i = 1, \ldots, n_g$. They are also mutually independent in the treatment of Chi and Weerahandi (1998), implying that

$$Var(\mathbf{Y}_{gi}) = \Sigma_g = \sigma_\alpha^2\mathbf{1}_T\mathbf{1}_T^{'} + \sigma_g^2\mathbf{I}_T.$$

Lin and Lee (2003) argued that the results remain valid under the more general covariance structure,

$$\Sigma_g = (\rho_1\sigma_\alpha^2 + \rho_2\sigma_g^2)\mathbf{1}_T\mathbf{1}_T^{'} + [(1-\rho_1)\sigma_\alpha^2 + (1-\rho_2)\sigma_g^2]\mathbf{I}_T\ . \tag{10.70}$$

In either case, the inverse of the covariance matrix is of the form

$$\Sigma_g^{-1} = \frac{1}{\phi_{1g}^2}\left[\mathbf{I}_T - \frac{(\phi_{2g}^2 - \phi_{1g}^2)}{T\phi_{2g}^2}\mathbf{1}_T\mathbf{1}_T^{'}\right],$$

where

$$\phi_{1g}^2 = (1-\rho_1)\sigma_\alpha^2 + (1-\rho_2)\sigma_g^2 \tag{10.71}$$

and

$$\phi_{2g}^2 = \phi_{1g}^2 + T(\rho_1\sigma_\alpha^2 + \rho_2\sigma_g^2) \tag{10.72}$$

with the particular case $\rho_1 = 1$ and $\rho_2 = 0$ in the Chi and Weerahandi (1998) treatment. The problem of primary importance is that of comparing growth curves of different groups, and especially the problem of testing the hypothesis

$$\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \cdots = \boldsymbol{\beta}_G \ . \tag{10.73}$$

Define

$$\begin{aligned} \overline{\mathbf{Y}}_{g.} &= \sum_{i=1}^{n_g} \mathbf{Y}_{gi}/n_g, \\ \overline{Y}_{gi} &= \sum_{t=1}^{T} Y_{git}/T, \end{aligned}$$

and

$$\overline{Y}_{g} = \sum_{i=1}^{n_g} \overline{Y}_{gi}/n_g.$$

It follows from the results in the previous section that the MLE and the GLSE of each parameter vector is the same as the LSE and its distribution is normal:

$$\widehat{\boldsymbol{\beta}}_g = (\mathbf{X}_g'\mathbf{X}_g)^{-1}\mathbf{X}_g'\overline{\mathbf{Y}}_{g.} \sim N(\boldsymbol{\beta}_g,((\mathbf{X}_g'\Sigma_g^{-1}\mathbf{X}_g)^{-1})/n_g) \tag{10.74}$$

First consider the simpler problem of testing the hypothesis (10.73) under the assumption that group residual variances are all equal; that is,

$$\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_G^2 = \sigma_e^2 \tag{10.75}$$

so that $\Sigma_g = \Sigma$ for all g. Classical approach to the problem does not provide exact solutions or good approximations to even this case where there are only two nuisance parameters. The solution by the generalized approach is quite straightforward. To derive a test, decompose the residual sum of squares as

$$SSE = \sum_{g=1}^{G}\sum_{i=1}^{n_g}\sum_{t=1}^{T}(Y_{git} - \mathbf{X}_{gt}'\widehat{\boldsymbol{\beta}}_g)^2 = S_{e1}^2 + S_{e2}^2 \ ,$$

where

$$S_{e1}^2 = \sum_g\sum_i\sum_t (Y_{git} - \mathbf{X}_{gt}'\widehat{\boldsymbol{\beta}}_g - (\overline{Y}_{gi} - \overline{Y}_g))^2, \tag{10.76}$$

and

$$S_{e2}^2 = \ T\sum_g\sum_i (\overline{Y_i} - \overline{Y})^2.$$

These random variables are independently distributed as

$$W_1 = \frac{S_{e1}^2}{\phi_1^2} \sim \chi_{\nu_1}^2 \tag{10.77}$$

and

$$W_2 = \frac{S_{e2}^2}{\phi_2^2} \sim \chi_{\nu_2}^2,$$

where $\nu_1 = N(T-1) - G(p-1)$, $\nu_2 = N - G$, $N = \sum n_g$,

$$\phi_1^2 = (1-\rho_1)\sigma_\alpha^2 + (1-\rho_2)\sigma_e^2,$$

and

$$\phi_2^2 = \phi_1^2 + T(\rho_1\sigma_\alpha^2 + \rho_2\sigma_e^2).$$

Letting $\Sigma^{-1/2}$ *denote a positive definite square root matrix of* Σ^{-1}*, multiplying both sides of (10.69) we can rewrite the model as*

$$\widetilde{\mathbf{Y}}_{gi} = \widetilde{\mathbf{X}}_g\boldsymbol{\beta}_g + \mathbf{e}_{gi}, \tag{10.78}$$

where $\widetilde{\mathbf{Y}}_{gi} = \Sigma^{-1/2}\mathbf{Y}_{gi}$ *and* $\widetilde{\mathbf{X}}_g = \Sigma^{-1/2}\mathbf{X}_g$*. Let*

$$\widetilde{S}_{1,2}^2(\phi_1^2, \phi_2^2) = \widetilde{S}_1^2(\phi_1^2, \phi_2^2) + \cdots + \widetilde{S}_G^2(\phi_1^2, \phi_2^2) \tag{10.79}$$

be the residual sum of squares obtained using model (10.78) when ϕ_1^2 *and* ϕ_2^2 *have been specified to enable the computation of* $\widetilde{\mathbf{Y}}_{gi}$ *and* $\widetilde{\mathbf{X}}_g$*. Let* $\widetilde{S}_{12}^2(\phi_1^2, \phi_2^2)$ *denote the residual sum of squares obtained under the null hypothesis of identical growth curves. The distributions of these sums of squares are given by*

$$\widetilde{S}_{1,2}^2(\phi_1^2, \phi_2^2) \sim \chi_{\nu_3}^2 \tag{10.80}$$

and

$$\widetilde{S}_{12}^2(\phi_1^2, \phi_2^2) \sim \chi_{\nu_4}^2,$$

where $\nu_3 = NT - pG = \nu_1 + \nu_2$*, and* $\nu_4 = NT - p$*. It is also easily seen that*

$$\begin{aligned} \widetilde{S}_{1,2}^2(\phi_1^2, \phi_2^2) &= \frac{S_{e1}^2}{\phi_1^2} + \frac{S_{e2}^2}{\phi_2^2} && (10.81) \\ &= W_1 + W_2 \sim \chi_{\nu_3}^2 && (10.82) \end{aligned}$$

and that it is distributed independently of $\widetilde{S}_{12}^2(\phi_1^2, \phi_2^2) - \widetilde{S}_{1,2}^2(\phi_1^2, \phi_2^2)$*.*

To derive the generalized test for testing the hypothesis (10.73), consider the potential extreme region defined by the inequality

$$\left\{ \widetilde{S}_{12}^2(\phi_1^2, \phi_2^2) \geq \widetilde{s}_{12}^2\left(\frac{s_{e1}^2}{S_{e1}^2}\phi_1^2, \frac{s_{e2}^2}{S_{e2}^2}\phi_2^2\right) \right\}, \tag{10.83}$$

a well-defined subset of the sample space, where s_{e1}^2 *and* s_{e2}^2 *are the observed values of the sums of squares* S_{e1}^2 *and* S_{e2}^2*, respectively. Obviously, the observed sample point falls on the boundary of the extreme region and the probability of the inequality is greater under the alternative hypothesis compared to that under the null hypothesis, as further clarified below. Using the property* $\widetilde{S}_{12}^2(k\phi_1^2, k\phi_2^2) =$

$k^{-1}\widetilde{S}^2_{12}(\phi^2_1,\phi^2_2)$ and the identity (10.81) and defining a beta random variable as in the previous section, the generalized p-value for testing the hypothesis of identical growth curves can be computed as

$$\begin{aligned} p &= \Pr\left(\frac{\widetilde{S}^2_{12}(\phi^2_1,\phi^2_2)}{W_1+W_2} \geq \widetilde{s}^2_{12}\left(\frac{s^2_{e1}}{W_1/(W_1+W_2)}, \frac{s^2_{e2}}{W_2/(W_1+W_2)}\right)\right. \quad (10.84)\\ &= \Pr\left\{\frac{\widetilde{S}^2_{12}(\phi^2_1,\phi^2_2)}{\widetilde{S}^2_{1,2}(\phi^2_1,\phi^2_2)} \geq \widetilde{s}^2_{12}\left(\frac{s^2_{e1}}{B}, \frac{s^2_{e2}}{1-B}\right)\right\}\\ &= \Pr\left\{\frac{\widetilde{S}^2_{12}(\phi^2_1,\phi^2_2) - \widetilde{S}^2_{1,2}(\phi^2_1,\phi^2_2)}{\widetilde{S}^2_{1,2}(\phi^2_1,\phi^2_2)} \geq \widetilde{s}^2_{12}\left(\frac{s^2_{e1}}{B}, \frac{s^2_{e2}}{1-B}\right) - 1\right\}\\ &= 1 - E\left\{H_{\nu,\vartheta}\left[\frac{\vartheta}{\nu}\left(\widetilde{s}^2_{12}\left(\frac{s^2_{e1}}{B}, \frac{s^2_{e2}}{1-B}\right) - 1\right)\right]\right\}, \quad (10.85) \end{aligned}$$

where $B \sim Beta(\nu_1/2, \nu_2/2)$, and $H_{\nu,\vartheta}$ is the cdf of F distribution with $\nu = p(G-1)$ and $\vartheta = NT - pG$ degrees of freedom. Using widely used software packages such as SAS and SPlus, the p-value could be computed by Monte Carlo integration. Without having to code the underlying formulas, the p-value can be conveniently computed using the XPro software package. It also provides necessary input for making inferences concerning variance components σ^2_α and σ^2_e.

Example 10.4. *Comparing growth curves with structured covariances*

Consider again the data set given in Table 10.3 and let us compare the growth curves assuming model (10.69). The estimates of growth curve parameters, $\boldsymbol{\gamma}$, estimated now by applying (10.74) for ordinary LSE, is

$$\widehat{\boldsymbol{\gamma}} = \begin{pmatrix} \boldsymbol{\beta}'_1 \\ \boldsymbol{\beta}'_2 \\ \boldsymbol{\beta}'_3 \end{pmatrix} = \begin{pmatrix} 8.5612 & 1.2039 \\ 10.385 & 0.9753 \\ 13.052 & 0.7820 \end{pmatrix}$$

The p-value for testing the equality of growth curves computed using (10.85) is 0.1288. Note that, as a result of using a structured covariance matrix this p-value is much less than the one computed using (10.31).

10.9 Case of Unequal Group Variances

Chi and Weerahandi (1998) also provided a procedure for testing the assumption of equal group variances is reasonable or not. When the assumption is not reasonable, it is better to drop the assumption rather than risking the accuracy of tests. Now consider the problem of testing the hypothesis (10.73) without the assumption of equal error variances made in the above derivation. In this case, unknown variances can be tackled by means of the residual sum of squares

$$\begin{aligned} S_g^2 &= \sum_{i=1}^{n_g}\sum_{t=1}^{T}(Y_{git} - \mathbf{X}'_{gt}\widehat{\boldsymbol{\beta}}_g - (\overline{Y}_{gi} - \overline{Y}_g))^2, \\ g &= 1, \cdots, G. \end{aligned}$$

The distribution of S_g^2 *is given by*

$$U_g = \frac{S_g^2}{\phi_{1g}^2} \sim \chi^2_{\nu_{1g}}, \tag{10.86}$$

where $\nu_{1g} = n_g(T-1) - p + 1$. *Moreover, these random variables are distributed independently from that of*

$$V_g = \frac{\acute{S}_g^2}{\phi_{2g}^2} \sim \chi^2_{\nu_{2g}}, \tag{10.87}$$

where $\nu_{2g} = n_g - 1$, *and*

$$\acute{S}_g^2 = T\sum_{i=1}^{n_g}(\overline{Y}_{gi} - \overline{Y}_g)^2.$$

Now we can proceed as before with the same data transformation as in (10.78) by defining

$$\widetilde{S}^2_{1,2}(\phi_{11}^2, \ldots, \phi_{1G}^2, \phi_{21}^2, \ldots, \phi_{2G}^2) = \widetilde{S}^2_1(\phi_{11}^2, \phi_{21}^2) + \ldots + \widetilde{S}^2_G(\phi_{1G}^2, \phi_{2G}^2), \tag{10.88}$$

the residual sum of squares obtained using model (10.78) when the parameters ϕ_{1g}^2 *and* ϕ_{2g}^2 *for all* g *have been specified to enable computation of* $\widetilde{\mathbf{Y}}_{gi}$ *and* $\widetilde{\mathbf{X}}_g$. *It is easily seen using (10.86), (10.87), and (10.88) that*

$$\widetilde{S}^2_{1,2}(\phi_{11}^2, \ldots, \phi_{1G}^2, \phi_{21}^2, \ldots, \phi_{2G}^2) = \sum_g (U_g + V_g) \sim \chi^2_{NT-pG} \tag{10.89}$$

Similarly, let $\widetilde{S}^2_{12}(\phi_{11}^2, \cdots, \phi_{1G}^2, \phi_{21}^2, \cdots, \phi_{2G}^2)$ *denote the residual sum of squares obtained under the null hypothesis of identical growth curves. The distributions of these sums of squares are given by*

$$\widetilde{S}^2_{12} \sim \chi^2_{\nu_4} \text{ and } \widetilde{S}^2_{12} - \widetilde{S}^2_{1,2} \sim \chi^2_{\nu} \tag{10.90}$$

are independently distributed, where $\nu_3 = NT - pG$, $\nu_4 = NT - p$, *and* $\nu = p(G-1) = \nu_4 - \nu_3$ *are the same degrees of freedom as in the equal variances case, because the nuisance parameters have been specified. Hence, proceeding as before we can compute the generalized p-value for testing identical growth curves can be computed as*

$$p = \Pr\{\widetilde{S}_{12}^2(\phi_{11}^2,\ldots,\phi_{1G}^2,\phi_{21}^2,\ldots,\phi_{2G}^2) \geq \tag{10.91}$$

$$\widetilde{s}_{12}^2(\frac{s_1^2}{S_1^2}\phi_{11}^2,\ldots,\frac{s_G^2}{S_G^2}\phi_{1G}^2,\frac{\acute{s}_1^2}{\acute{S}_1^2}\phi_{21}^2,\ldots,\frac{\acute{s}_G^2}{\acute{S}_G^2}\phi_{2G}^2)\} \tag{10.92}$$

$$= \Pr\left(\frac{\widetilde{S}_{12}^2}{\widetilde{S}_{1,2}^2} \geq \frac{1}{\widetilde{S}_{1,2}^2}\widetilde{s}_{12}^2(\frac{s_1^2}{U_1},\ldots,\frac{s_G^2}{U_G},\frac{\acute{s}_1^2}{V_1},\ldots,\frac{\acute{s}_G^2}{V_G})\right). \tag{10.93}$$

To express this p-value also as integration with respect to some beta random variables, define $2G$ *chi-squared random variables observed values as*

$$R_k = \begin{cases} U_k & \text{if } k \leq G \\ V_{k-G} & \text{if } k > G \end{cases} \tag{10.94}$$

and

$$z_k = \begin{cases} s_k^2 & \text{if } k \leq G \\ \acute{s}_{k-G}^2 & \text{if } k > G \end{cases} . \tag{10.95}$$

Since $R_k \sim \chi_{r_k}^2$, $k = 1,\ldots 2G$ *are all independent random variables, we can define* $2G-1$ *independent beta random variables as*

$$B_l = \frac{\sum_k^l R_k}{\sum_k^{l+1} R_k} \sim Beta(\frac{\sum_k^l r_k}{2}, \frac{r_{l+1}}{2}) \ , \ l = 1,\ldots,,2G-1 \ , \tag{10.96}$$

where

$$r_k = \begin{cases} \nu_{1k} & \text{if } k \leq G \\ \nu_{2(k-G)} & \text{if } k > G \end{cases}$$

With this notation, the p-value corresponding to the extreme region defined by (10.91) can be computed as

$$p = \Pr\left(\frac{\widetilde{S}_{12}^2 - \widetilde{S}_{1,2}^2}{\widetilde{S}_{1,2}^2} \geq \frac{\widetilde{s}_{12}^2(\frac{z_1}{R_1},\frac{z_2}{R_2},\ldots,\frac{z_{2G}}{R_{2G}})}{R_1+\cdots+R_{2G}} - 1\right)$$

$$= \Pr\{F \geq \widetilde{s}_{12}^2(\frac{z_1}{B_1B_2\cdots B_{2G-1}},\ldots,$$

$$\frac{z_k}{(1-B_{k-1})B_k\cdots B_{2G-1}},\ldots,\frac{z_{2G}}{(1-B_{2G-1})}) - 1\}$$

$$= \Pr(1 - E\{H_{\nu,\vartheta}\widetilde{s}_{12}^2(\frac{z_1}{B_1B_2\cdots B_{2G-1}},\ldots,$$

$$\frac{z_k}{(1-B_{k-1})B_k\cdots B_{2G-1}},\ldots,\frac{z_{2G}}{(1-B_{2G-1})}) - 1\}) \ , \tag{10.97}$$

where $H_{\nu,\vartheta}$ *is the cdf of* F *distribution with* $\nu = p(G-1)$ *and* $\vartheta = NT - pG$ *degrees of freedom of* F *random variable.*

Perhaps the simplest way to evaluate this probability is by Monte Carlo Integration. This involve generating large number of beta random variates from

each of the beta random variables, the value of $\widetilde{s}^2_{12}$ at for each set of values, and then the cdf of the F distribution at each value of $\widetilde{s}^2_{12}$. Finally, the expected value is estimated by the sample mean of F-values and in turn the p-value. In this case also the p-value can be conveniently computed using the XPro software package. Based on formulae given by Chi and Weerahandi(1998), it also provides various input for making inferences concerning variance components σ^2_α and σ^2_g; $g = 1, \cdots, G$.

Example 10.5. *Comparing growth curves under unequal error variances*

Consider again the data set given in Table 10.3 and let us compare the growth curves assuming the model (10.69). The estimates of growth curve parameters, $\boldsymbol{\gamma}$, still computed using (10.74) and remains as.

$$\widehat{\boldsymbol{\gamma}} = \begin{pmatrix} \boldsymbol{\beta}_1' \\ \boldsymbol{\beta}_2' \\ \boldsymbol{\beta}_3' \end{pmatrix} = \begin{pmatrix} 8.5612 & 1.2039 \\ 10.385 & 0.9753 \\ 13.052 & 0.7820 \end{pmatrix}$$

However, their standard errors are now different. As a result, the p-value for testing the equality of growth curves computed using (10.97) now becomes 0.0834 and the null hypothesis can be rejected at the 0.1 level of significance. Note that as a result of dropping the unreasonable assumption of equal error variances, despite the fact that now we have 2 additional nuisance parameters, this p-value is less than the one computed under the assumption that they are equal. This further demonstrates the repercussions of unreasonable assumptions on the power of a test.

Exercises

Exercise 16 *Consider model (10.2) in generalized MANOVA. Write down the form of*

(a) $\mathbf{A}$ *matrix if growth curves all subjects are different,*
(b) parameter matrix $\boldsymbol{\gamma}$*, and the design matrix* $\mathbf{X}$ *if exponential growth curves with one nuisance parameter, representing the coefficient of time variable are assumed*
(c) $\mathbf{A}$*,* $\boldsymbol{\gamma}$*, and* $\mathbf{X}$ *matrices in ordinary MANOVA.*

Exercise 17 *Consider model (10.2) and hypotheses of the form (10.21). Write down the form of* $\mathbf{C}$ *and* $\mathbf{D}$ *matrices if*

(a) growth curves of group 1 and group 2 are to be compared,
(b) the hypothesis that all polynomial growth curves have the same intercept term is to be tested,
(c) the hypothesis of identical linear coefficients of polynomial growth curves is to be tested,
(d) the hypothesis that first two growth curves are parallel is to be tested.

Exercise 18 *Consider again model (10.2) in generalized MANOVA and assume straight lines for the growth curves. Write down the particular formulae for computing p-values for testing the hypothesis that*

(a) the intercepts of the first and the second growth curves are identical,
(b) the slopes of the first and the second growth curves are identical,
(c) intercept of ith growth curve is positive,
(d) slope of ith growth curve is positive.

Exercise 19 *Repeat the inference procedures carried out in previous exercise when the covariance matrices has a common intraclass covariance structure. Repeat the inferences when the covariance structure is of the intraclass form and yet the error variances of the three growth curves are unequal.*

Exercise 20 *Consider again the generalized MANOVA model with straight lines for the growth curves. Write down the particular formulae for Scheffe and Bonferroni confidence intervals for the*

(a) difference in intercepts of the first and the second growth curves,
(b) difference in slopes of the first and the second growth curves,
(c) intercept of ith growth curve,
(d) slope of ith growth curve.

Exercise 21 *Repeat the inference procedures carried out in the previous exercise when the covariance matrices have a common intraclass covariance structure. Repeat the inferences when the covariance structure is of the intraclass form and yet the error variances of the three growth curves are unequal.*

Exercise 22 *Consider the data set in Table 10.1 involving one group of subjects. Fit a second-order polynomial to the data. Assuming that the covariance matrix in unstructured, report*

(a) estimates of parameters of the growth curve,
(b) estimates of their standard errors,
(c) p-value for testing the hypothesis that the intercept of the growth curve is 48,
(d) p-value for testing the hypothesis that the growth curve is actually a straight line,
(e) 95% confidence intervals for each coefficient of the growth curve.

Exercise 23 *Repeat the analyses carried out in the previous exercise when the covariance matrix has the intraclass covariance structure.*

Exercise 24 *Consider the data set in Table 10.2 involving three groups of subjects. Fit second-order polynomials to each group and report the estimated coefficients, assuming an unstructured common covariance matrix.*

(a) Test the hypothesis that the growth curves are identical.
(b) Test the hypothesis that the growth curves are parallel.
(c) Test the hypothesis that the first growth curve is actually a straight line.
(d) Find a 95% confidence interval for the difference in the slopes of the first two growth curves.
(e) Find a 95% confidence intervals for coefficients of the first growth curve.

Exercise 25 *Repeat the analyses carried out in the previous exercise when the covariance matrices have a common intraclass covariance structure. Repeat the analyses when the covariance structure is of the intraclass form and yet the error variances of the three growth curves are unequal.*

Exercise 26 *Consider the data set in Table 10.3 involving four groups of subjects. Assuming a linear model with an unstructured common covariance matrix,*

(a) test the hypothesis that the growth curves are identical,
(b) test the hypothesis that the growth curves are parallel,
(c) 95% confidence intervals for difference in slopes of the first two growth curves.

Exercise 27 *Repeat the analyses carried out in the previous exercise when the covariance matrices have a common intraclass covariance structure.*

Univariate Technical Arguments

.1 Derivation of the Generalized F-test in One-Way ANOVA

To express the generalized test in Chapter 2 as a generalized F-test, consider the p-value derived as

$$p \ = \Pr(\widetilde{S}_B \geq \widetilde{s}_B(\frac{n_1 s_1^2}{Y_1}, \frac{n_2 s_2^2}{Y_2}, \ldots, \frac{n_k s_k^2}{Y_k}),$$

where

$$\widetilde{S}_B \ = \ \widetilde{S}_B(\sigma_1^2, \ldots, \sigma_k^2) \ \sim \ \chi^2_{k-1},$$

$\widetilde{s}_B$ is the observed value of $\widetilde{S}_B$, s_i^2 is the observed values of sample variance S_i^2, and

$$Y_i = n_i S_i^2/\sigma_i^2 \sim \ \chi^2_{n_i-1}. \tag{98}$$

To express the test in terms of the F distribution and to obtain a computationally superior representation of the underlying integral, define the independent beta random variables,

$$B_1 = \frac{Y_1}{Y_1 + Y_2} \sim Beta(\frac{(n_1 - 1)}{2} \ , \ \frac{n_2 \ - \ 1}{2})$$

$$\begin{aligned} B_j \quad &= \quad \frac{Y_1 + Y_2 + \cdots + Y_j}{Y_1 + Y_2 + \ldots + Y_j + Y_{j+1}} \sim \ Beta(\sum_{i=1}^{j} \frac{(n_i - 1)}{2} \ , \ \frac{n_{j+1} \ - \ 1}{2}), \\ j \quad &= \quad 2, \ldots, k-1. \end{aligned}$$

Moreover, the sum of all chi-squared random variables is distributed as

$$W = Y_1 + Y_2 + \ldots + Y_{jk} \sim \ \chi^2_{N-k}$$

independently of the beta random variables, where $N = \sum n_i$. It is also known that $W, \ B_j, \ j \ = \ 1, 2, \ldots, k-1$ are mutually independent random variables. Note also that the chi-squared random variables Y_i's can be expressed as

$$Y_1 \ = \ W B_1 B_2 \cdots B_{k-1},$$

$$Y_i \ = W(1 - B_{i-1}) B_i \cdots B_{k-1} \qquad \textit{for } i = 2, \ldots, k,$$

and

$$Y_k \ = \ W(1 \ - \ B_{k-1}),$$

Therefore, the p-value can be expressed as

$$p = \ \Pr(\widetilde{S}_B \geq \ \widetilde{s}_B(\frac{n_1 s_1^2}{W B_1 B_2 \cdots B_{k-1}}, \ldots, \frac{n_k s_k^2}{W(1 \ - \ B_{k-1})}) \ )$$

But it follows from the definition of $\widetilde{S}_B$ *that it has the property that, for any given positive constant* c *and a vector* $\mathbf{x}$ *with positive elements,*

$$\widetilde{S}_B(cx_1, \ldots, cx_k) \;=\; \widetilde{S}_B(x_1, \ldots, x_k)/c.$$

Moreover, under the null hypothesis,

$$W = \frac{\widetilde{S}_B/(k-1)}{W/(N-k)} \;\sim\; F_{k-1,N-k}.$$

Hence, the p-value can be expressed as

$$\begin{aligned} p &= \Pr(\frac{\widetilde{S}_B}{W} \geq \widetilde{s}_B(\frac{n_1 s_1^2}{B_1 B_2 \cdots B_{k-1}}, \ldots, \frac{n_k s_k^2}{(1 - B_{k-1})})) \\ &= 1 - E(H_{k-1,N-k}(\frac{N-k}{k-1}\widetilde{s}_B[\frac{n_1 s_1^2}{B_1 B_2 \cdots B_{k-1}}, \frac{n_2 s_2^2}{(1-B_1)B_2 \cdots B_{k-1}}, \\ &\quad \frac{n_3 s_3^2}{(1-B_2)B_3 \cdots B_{k-1}}, \ldots, \frac{n_k s_k^2}{(1-B_{k-1})}])), \end{aligned} \tag{99}$$

where $H_{k-1,N-k}$ *is the cdf of* F *distribution with* $k-1$ *and* $N-k$ *degrees of freedom, and the expectation is taken with respect to the independent beta variables defined above.*

.2 ANOVA for the Two-Way Layout without Replicates

Suppose we are interested in the fixed effects of two factors, A *and* B*. Let* $A_1,\ A_2, \ldots,\ A_k$ *be the levels of factor* A*, and let* $B_1,\ B_2, \ldots,\ B_n$ *be the levels of factor* B*. In the two-way layout with no replicates we have just one data for each combination of factor levels. The available data can be set out as in Table 2.9. To enable analysis of data from an experiment involving the two-way cross classified design involving no interactions, consider the linear model*

$$\begin{gathered} Y_{ij} \;=\; \mu \;+\; \alpha_i \;+\; \beta_j \;+\; \epsilon_{ij}, \\ i = 1, \ldots, k,\ j = 1, \ldots, n, \end{gathered} \tag{100}$$

where α_i *is the* i*th effect of factor* A *and* β_j *is the* j*th effect of factor* B *standardized such that* $\sum_{i=1}^{k} \alpha_i \;=\; 0$ *and* $\sum_{j=1}^{n} \beta_j \;=\; 0$*, respectively. We also assume that the residuals are normally distributed with a common variance; i.e.* $\epsilon_{ij} \;\sim\; N(0, \sigma^2)$*. Note that, to avoid over parameterization, we need to assume in the current problem that error variances are all equal. If the assumption is not reasonable it is necessary that we obtain more than one observation from each cell.*

We are interested in testing the equality of factor levels, namely testing the hypotheses

$$H_{0A}: \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0 \tag{101}$$

and

$$H_{0B}: \beta_1 = \beta_2 = \cdots = \beta_n = 0 \tag{102}$$

against the obvious alternative hypotheses.

Consider the sample means $\overline{Y}_{i.}$ *corresponding to factor* A_i *and the sample means* $\overline{Y}_{.j}$ *corresponding to factor* B_j. *Let* $\overline{Y}$ *be the grand mean of all data and let* $N = nk$. *It is straightforward to see that* $\widehat{\alpha}_i = \overline{Y}_{i.} - \overline{Y}$ *and* $\widehat{\beta}_j = \overline{Y}_{.j} - \overline{Y}$ *are unbiased estimates of the parameters* α_i *and* β_j, *respectively. They are also the maximum likelihood estimates of the parameters. Consider the sums of squares*

$$\begin{aligned} S_A &= n\sum_{i=1}^{k}(\overline{Y}_{i.} - \overline{Y})^2 , \\ S_B &= k\sum_{j=1}^{n}(\overline{Y}_{.j} - \overline{Y})^2, \end{aligned}$$

and

$$S_E = \sum_{i=1}^{k}\sum_{j=1}^{n}(Y_{ij} - \overline{Y}_{i.} - \overline{Y}_{.j} + \overline{Y})^2.$$

As in the one-way ANOVA, various testing procedures can be based on a decomposition of the total sum of squares $S_T = \sum\sum(Y_{ij} - \overline{Y})^2$ *given by*

$$S_T = S_A + S_B + S_E.$$

The result is easily seen by taking the sum of squares of each side of the identity

$$(Y_{ij} - \overline{Y}) = (\overline{Y}_{i.} - \overline{Y}) + (\overline{Y}_{.j} - \overline{Y}) + (Y_{ij} - \overline{Y}_{i.} - \overline{Y}_{.j} + \overline{Y}). \tag{103}$$

and showing that the cross products of terms on the right hand side sum to zero. In this case, the mean sums of squares are defined as

$$\begin{aligned} MSA &= \frac{S_A}{(k-1)}, \\ MSB &= \frac{S_B}{(n-1)}, \end{aligned}$$

and

$$MSE = \frac{S_E}{(n-1)(k-1)}.$$

It also follows from zero sums of cross products that the decomposition is orthogonal. In turn it follows from the normal theory results that the sums of squares S_A, S_B, and S_E are mutually independent and that they are distributed as

$$\frac{S_E}{\sigma^2} \sim \chi^2_{(n-1)(k-1)}. \tag{104}$$

and

$$\frac{S_A}{\sigma^2} \sim \chi^2_{k-1} \quad \text{and} \quad \frac{S_B}{\sigma^2} \sim \chi^2_{n-1} \tag{105}$$

The distributional result (104) is true in general, and the results in (105) are valid under H_{0A} and H_{0B}, respectively. Hence, under H_{0A},

$$F_A = \frac{MSA}{MSE} = \frac{S_A/a}{S_E/e} \sim F_{a,e}\ . \tag{106}$$

where $a = k-1$ and $e = (n-1)(k-1)$. Furthermore, if H_{0A} is not true the distributions in (105) and (106), respectively, become non-central chi-squared and non-central F_A distributions. Consequently, unbiased tests of H_{0A} can be based on the p-value

$$p_a = 1 - H_{a,e}(\ \frac{s_A/a}{s_E/e}), \tag{107}$$

where $H_{a,b}$is the cdf of the F distribution with a and b degrees of freedom.

Similarly, testing of the hypothesis H_{0B} can be based on the result

$$F_B = \frac{MSB}{MSE} = \frac{S_B/(n-1)}{S_E/(n-1)(k-1)} \sim F_{n-1,(n-1)(k-1)}. \tag{108}$$

leading to the p-value

$$p_b = 1 - H_{b,e}(\ \frac{s_B/b}{s_E/e}\ ), \tag{109}$$

where $b = n-1$ and $e = (n-1)(k-1)$. Various quantities playing a role in the computation these p-values are summarized by the ANOVA table shown below, where various sums of squares are denoted by SS and their mean sums of squares are denoted by MS.

Table 5: Two-way ANOVA table

Source	D.F.	SS	MS	F-statistic
A	$k-1$	s_A	MSA	$\frac{MSA}{MSE}$
B	$n-1$	s_B	MSB	$\frac{MSB}{MSE}$
Error	$(n-1)(k-1)$	s_E	MSE	
Total	$nk-1$	s_T		

.3 Derivation of F-tests for Two-Way ANOVA with Replications

Consider the linear model

$$\begin{aligned} Y_{ijk} &= \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}, \\ i &= 1,\ldots,I,\ j=1,\ldots,J,\ k=1,\ldots,n_{ij}, \end{aligned} \tag{110}$$

where γ_{ij} terms represent the interactions and

$$\epsilon_{ijk} \sim N(0,\sigma^2).$$

Consider the problem of testing zero interactions

$$H_{0AB}:\ \gamma_{ij} = 0 \ \text{ for all } \ i=1,\ldots,I, \quad j=1,\ldots,J. \tag{111}$$

It is also of interest to test the hypotheses that there is no difference in main effects as specified by H_{0A} and H_{0B} in the previous section. Testing of hypotheses could be based on a decomposition of the total sum of squares into sums of squares that are attributed to factor A, factor B, the interaction between factors A and B, and the random error. Consider the partition

$$S_T = S_A + S_B + S_I + S_E, \tag{112}$$

where

$$S_T = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk} - \overline{Y})^2,$$

$$S_A = JK\sum_{i=1}^{I}(\overline{Y}_{i.} - \overline{Y})^2,\ S_B = IK\sum_{j=1}^{J}(\overline{Y}_{.j} - \overline{Y})^2,$$

$$S_I = K\sum_{i=1}^{I}\sum_{j=1}^{J}(\overline{Y}_{ij} - \overline{Y}_{i.} - \overline{Y}_{.j} + \overline{Y})^2,$$

and

$$S_E = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk} - \overline{Y}_{ij})^2 = K\sum_{i=1}^{I}\sum_{j=1}^{J}S_{ij}^2;$$

where S_{ij}^2 is the sample variance (MLE of σ_{ij}^2) of the data from ijth cell.

By the orthogonality of the above decomposition, the sums of squares S_A, S_B, and S_E are all independently distributed and

$$\frac{S_E}{\sigma^2} \sim \chi^2_{N-IJ}.$$

Moreover, respectively under H_{0A} and H_{0B}, we have

$$\frac{S_A}{\sigma^2} \sim \chi^2_{I-1}, \quad \frac{{S_B}^2}{\sigma} \sim \chi^2_{J-1}$$

and under H_{0AB}, we have

$$\frac{{S_I}^2}{\sigma} \sim \chi^2_{(I-1)(J-1)}.$$

Hence, if H_{0AB} is true we get the F-statistic

$$F_I = \frac{MSI}{MSE} = \frac{S_I/(I-1)(J-1)}{S_E/(N-IJ)} \sim F_{(I-1)(J-1),N-IJ}; \tag{113}$$

that is, F_I has an F distribution with $(I-1)(J-1)$ and $N-IJ$ degrees of freedom. Moreover, F_I tends to take large values for deviations from H_{0AB}.

Therefore, H_{0AB} is tested on the basis of the p-value

$$p_I = 1 - H_{i,e}\left(\frac{s_I/i}{s_E/e}\right), \tag{114}$$

where $i = (I-1)(J-1)$, $e = N-IJ$, and in general, the notation $H_{a,b}$ stands for the cdf of the F distribution with a and b degrees of freedom. At fixed-level α, H_{0AB} is rejected if the observed value of the F_I statistic is greater than $F_{i,e}(\alpha)$, the $(1-\alpha)$th quantile of the F distribution with i and e degrees of freedom.

Similarly, testing of the hypotheses H_{0B} and H_{0AB} can be based on the p-values

$$p_A = 1 - H_{a,e}\left(\frac{s_A/a}{s_E/e}\right) \text{ and } p_B = 1 - H_{b,e}\left(\frac{s_B/b}{s_E/e}\right), \tag{115}$$

respectively, where $a = I-1$ and $b = J-1$. The main effects and interactions can also be tested jointly. For instance, to test whether H_{0A} and H_{0AB} are both true, the sum of squares $s_A + s_I$ and its degrees of freedom $J(I-1)$ can be used in the numerator of the appropriate F-statistic. The computation of the F-statistics and the p-values is facilitated by the analysis of variance table as shown below.

.4 Probability Coverage of GCI to the Behrens−Fisher Problem

The $100\gamma\%$ left-sided generalized confidence interval is of the form

$$(\overline{x}_1 - \overline{x}_2) - c_\gamma(s_1^2, s_2^2) \leq \delta, \tag{116}$$

Table 6: Two-way ANOVA with interactions

Source	D.F.	SS	MS	F-statistic
A	$I - 1$	S_A	MSA	MSA/MSE
B	$J - 1$	S_B	MSB	MSB/MSE
Interac.	$(I-1)(J-1)$	s_I	MSI	MSI/MSE
Error	$N - IJ$	s_E	MSE	
Total	$N - 1$	s_T		

where $c_\gamma(s_1^2, s_2^2)$ *is the solution of the equation*

$$E_B G_{n_1+n_2-2}(c_\gamma(s_1^2, s_2^2)\sqrt{\frac{n_1+n_2-2}{\frac{s_1^2}{B}+\frac{s_2^2}{(1-B)}}}) = \gamma, \tag{117}$$

where the expectation is taken with respect to the beta random variable

$$B \sim Beta(\frac{n_1-1}{2}, \frac{n_2-1}{2}).$$

Notice that $c_\gamma(s_1^2, s_2^2)$ *function has the property,* $c(ks_1^2, ks_2^2) = \sqrt{k}c(s_1^2, s_2^2)$ *for positive real number* k*. Akahira (1999) obtained a general formula for the probability coverage and hence the actual size of any two-sided confidence interval having this property. Using the argument of Akahira (1999), the actual size of confidence intervals of the form (116), can be obtained in terms of* $\rho = \frac{\sigma_1^2/n_1}{\sigma_1^2/n_1+\sigma_2^2/n_2}$ *as*

$$\begin{aligned} p(\rho) &= \Pr((\overline{X}_1 - \overline{X}_2) - c(S_1^2, S_2^2) \le \delta) \\ &= \Pr(Z \le \frac{c(\sigma_1^2 Y_1/n_1, \sigma_2^2 Y_2/n_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}) \\ &= \Pr(Z \le c(\rho Y_1, (1-\rho)Y_2) \\ &= \Pr(Z \le \sqrt{Y}c(\rho B, (1-\rho)(1-B))) \\ &= EG_{n_1+n_2-2}(\sqrt{n_1+n_2-2}c(\rho B, (1-\rho)(1-B))), \end{aligned} \tag{118}$$

where

$$Y_1 = \frac{n_1 S_1^2}{\sigma_1^2} \sim \chi^2_{n_1-1},$$

$$Y_2 = \frac{n_2 S_2^2}{\sigma_2^2} \sim \chi^2_{n_2-1},$$

and

$$Y = Y_1 + Y_2 \sim \chi^2_{n_1+n_2-2}.$$

.5 Distribution of Sums of Squares of Random Effects Model

Consider the simplest random effects model

$$Y_{ij} \ = \mu \ + \ \alpha_i \ + \ \epsilon_{ij}, \qquad \textit{for } i = 1, \ldots, k, \ j = 1, \ldots, n,$$

where ϵ_{ij} is the error term representing the deviation of the response of the jth observation from the mean of observations from A_i, and α_i represents the random effect. Assume that

$$\epsilon_{ij} \ \sim \ N(0, \sigma_\epsilon^2), \qquad \alpha_i \ \sim \ N(0, \sigma_\alpha^2) \tag{119}$$

and that α_i, ϵ_{ij}; $i = 1, \ldots, k, \ j = 1, \ldots, n$ are mutually independent. Let $\overline{Y}_i$, $i = 1, \ldots, k$ be the sample means corresponding to the k random effects and let $\overline{Y}$ be the mean of all data. Consider the problem of finding the distributions of the sums of squares

$$S_E \ = \ \sum_{i=1}^{k} \sum_{j=1}^{n} (Y_{ij} \ - \ \overline{Y}_{i\cdot})^2 \tag{120}$$

and

$$S_B \ = \ n \sum_{i=1}^{k} (\overline{Y}_{i\cdot} \ - \ \overline{Y})^2 \tag{121}$$

that play an important role in one-way ANOVA models.

It immediately follows from

$$\overline{Y}_i = \mu \ + \alpha_i + \overline{\epsilon}_i$$

that

$$\overline{Y}_i \sim \ N(\mu, \ \sigma_\alpha^2 \ + \ \sigma_\epsilon^2/n), \qquad i = 1, \ldots, k \ . \tag{122}$$

From standard results valid for sampling a normal population, we also get k independent and identically distributed chi-squared random variables

$$\frac{\sum_{j=1}^{n} (Y_{ij} \ - \ \overline{Y}_{i\cdot})^2}{\sigma_\epsilon^2} \sim \chi^2_{n-1}, \qquad i = 1, \ldots, k \ . \tag{123}$$

It follows from (122) that

$$\frac{S_E}{\sigma_\epsilon^2} \sim \ \chi^2_{k(n-1)} \tag{124}$$

and it follows from (122) that

$$\frac{S_B}{\sigma_\epsilon^2 \ + \ n\sigma_\alpha^2} \sim \ \chi^2_{k-1}. \tag{125}$$

Moreover, these random variables are independently distributed.

.6 Distributions in Three-Way Random Effects Model

Consider the three-way random effects model

$$\begin{aligned} Y_{ijkl} &= \mu + (\alpha_i + \beta_j + \theta_k) + \\ &\quad (\alpha\beta_{ij} + \alpha\theta_{ik} + \beta\theta_{jk}) + \alpha\beta\theta_{ijk} + \epsilon_{ijkl}, \\ i &= 1, \ldots, I;\ j = 1, \ldots, J;\ k = 1, \ldots, K;\ l = 1, \ldots, L, \end{aligned} \tag{126}$$

where α_i is the effect due to the ith random level of A, β_j is the random effect due to the jth level of B, and θ_k is the random effect due to kth level of C. Terms such as $\alpha\beta_{ij}$ represents the two way interactions and $\alpha\beta\theta$ denotes the interaction between all three factors. As in the two-way mixed effects model, assume that the fixed effects are measured as deviations from the overall mean so that they satisfy the equation $\sum_{j=1}^{J} \beta_j \ = \ 0$. In three-way random effects it is assumed that

$$\begin{aligned} \alpha_i &\sim N(0, \sigma_\alpha^2), \\ \beta_j &\sim N(0, \sigma_\beta^2), \\ \theta_k &\sim N(0, \sigma_\theta^2) \end{aligned}$$

$$\begin{aligned} \alpha\beta_{ij} &\sim N(0, \sigma_{\alpha\beta}^2), \\ \alpha\theta_{ik} &\sim N(0, \sigma_{\alpha\theta}^2), \\ \beta\theta_{jk} &\sim N(0, \sigma_{\beta\theta}^2) \end{aligned}$$

$$\alpha\beta\theta_{ijk} \sim N(0, \sigma_{\alpha\beta\theta}^2),$$

and

$$\epsilon_{ijkl} \sim N(0, \sigma_\epsilon^2).$$

ANCOVA in three-way random effects models and mixed models are based on the sum of squares decomposition

$$S_T \ = (S_\alpha + S_\beta + S_\theta) + (S_{\alpha\beta} + S_{\alpha\theta} + S_{\beta\theta}) + S_{\alpha\beta\theta} + S_e$$

with definitions given in Chapter 4. To describe the approach to deriving the distribution of the sums of squares, consider for example

$$S_\alpha \ = \ JKL \sum_{i=1}^{I} (\overline{Y}_{i..} \ - \ \overline{Y})^2.$$

To derive its distribution, consider the identity

$$\begin{aligned}(\overline{Y}_{i..} - \overline{Y}) &= (\alpha_i - \overline{\alpha}) + (\overline{\alpha\beta}_i - \overline{\alpha\beta}) + (\overline{\alpha\beta\theta}_i - \overline{\alpha\beta\theta}) + (\overline{\epsilon}_i - \overline{\epsilon}),\\ &= (\nu_i - \overline{\nu}),\end{aligned}$$

where

$$\nu_i = \alpha_i + \overline{\alpha\beta}_i + \overline{\alpha\beta\theta}_i + \overline{\alpha\beta\theta}_i. \tag{127}$$

Since

$$\begin{aligned}\overline{\alpha\beta}_i &\sim N(0, \frac{\sigma^2_{\alpha\beta}}{J}),\\ \overline{\alpha\beta\theta}_i &\sim N(0, \frac{\sigma^2_{\alpha\beta\theta}}{JK}),\end{aligned}$$

and

$$\overline{\epsilon}_i \sim N(0, \frac{\sigma^2_\epsilon}{JKL})$$

we have

$$\nu_i \sim N(0, \sigma^2_\alpha + \frac{\sigma^2_{\alpha\beta}}{J} + \frac{\sigma^2_{\alpha\beta\theta}}{JK} + \frac{\sigma^2_\epsilon}{JKL}). \tag{128}$$

Consequently, from basic results of sampling from a normal distribution, we get

$$\frac{\sum_{i=1}^{I}(\nu_i - \overline{\nu})^2}{\sigma^2_\alpha + \frac{\sigma^2_{\alpha\beta}}{J} + \frac{\sigma^2_{\alpha\beta\theta}}{JK} + \frac{\sigma^2_\epsilon}{JKL}} \sim \chi^2_{I-1}. \tag{129}$$

Hence,

$$\frac{S_\alpha}{JKL\sigma^2_\alpha + KL\sigma^2_{\alpha\beta} + L\sigma^2_{\alpha\beta\theta} + \sigma^2_\epsilon} = \frac{\sum_{i=1}^{I}(\overline{Y}_{i..} - \overline{Y})^2}{\sigma^2_\alpha + \frac{\sigma^2_{\alpha\beta}}{J} + \frac{\sigma^2_{\alpha\beta\theta}}{JK} + \frac{\sigma^2_\epsilon}{JKL}} \sim \chi^2_{I-1}. \tag{130}$$

The distribution of S_α given by (130) also implies that

$$E(MS_\alpha) = E(\frac{S_\alpha}{I-1}) = JKL\sigma^2_\alpha + KL\sigma^2_{\alpha\beta} + L\sigma^2_{\alpha\beta\theta} + \sigma^2_\epsilon.$$

Multivariate Technical Arguments

.1 Form of $\widetilde{\mathbf{H}}_1(\mathbf{X}; \mathbf{\Lambda_1}, \mathbf{\Lambda_2})$ in the Behrens–Fisher Problem

Consider the generalized test variable

$$\begin{aligned}\widetilde{\mathbf{H}}_1(\mathbf{X}; \mathbf{\Lambda_1}, \ldots, \mathbf{\Lambda_I}) &= \sum_{i=1}^{I} \mathbf{T}_i'\mathbf{T}_i \qquad (131)\\ &= \sum_{i=1}^{I}(\overline{\mathbf{X}}_i - \overline{\widetilde{\mathbf{X}}})'\mathbf{\Lambda}_i^{-1}(\overline{\mathbf{X}}_i - \overline{\widetilde{\mathbf{X}}})\end{aligned}$$

used in the derivation of the generalized MANOVA test in Section 6.5 with the weights

$$\mathbf{W}_i(\mathbf{\Lambda}) = \left(\sum_{i=1}^{I} \mathbf{\Lambda}_i\right) \mathbf{\Lambda}_i^{-1},$$

where

$$\mathbf{T}_i = \mathbf{W}_i(\mathbf{\Lambda})(\overline{\mathbf{X}}_i - \overline{\widetilde{\mathbf{X}}}) = \mathbf{W_i}\mathbf{\Lambda_i^{1/2}}\mathbf{Z}_i - \mathbf{W_i}\left(\sum_{i=1}^{I} \mathbf{\Lambda}_i^{-1}\right)^{-1} \sum_{i=1}^{I} \mathbf{\Lambda}_i^{-1/2}\mathbf{Z}_i.$$

*When $I = 2$, the generalized test given by (131) reduces to the test given by (**??**) in the two-sample case. To see this, starting with the identity*

$$(\mathbf{\Lambda}_1^{-1} + \mathbf{\Lambda}_2^{-1})^{-1} = \mathbf{\Lambda}_1 (\mathbf{\Lambda}_1 + \mathbf{\Lambda}_2)^{-1} \mathbf{\Lambda}_2 = \mathbf{\Lambda}_2 (\mathbf{\Lambda}_1 + \mathbf{\Lambda}_2)^{-1} \mathbf{\Lambda}_1,$$

let us express $\mathbf{T}_i$ as

$$\begin{aligned}
\mathbf{T}_i &= \mathbf{W}_i(\overline{\mathbf{X}}_i - (\mathbf{\Lambda}_1^{-1} + \mathbf{\Lambda}_2^{-1})^{-1}(\mathbf{\Lambda}_1^{-1}\overline{\mathbf{X}}_1 + \mathbf{\Lambda}_2^{-1}\overline{\mathbf{X}}_2)) \\
&= \mathbf{W}_i(\overline{\mathbf{X}}_i - \mathbf{\Lambda}_2 (\mathbf{\Lambda}_1 + \mathbf{\Lambda}_2)^{-1} \overline{\mathbf{X}}_1 - \mathbf{\Lambda}_1 (\mathbf{\Lambda}_1 + \mathbf{\Lambda}_2)^{-1} \overline{\mathbf{X}}_2) \\
&= \pm\mathbf{W}_i\mathbf{\Lambda}_1 (\mathbf{\Lambda}_1 + \mathbf{\Lambda}_2)^{-1} (\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2) \\
&= \pm(\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2) \\
&= \pm (\mathbf{\Lambda}_1 + \mathbf{\Lambda}_2)^{1/2} \mathbf{Z},
\end{aligned}$$

where $\mathbf{Z} = (\mathbf{\Lambda}_1 + \mathbf{\Lambda}_2)^{-1/2} (\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2) \sim N(\mathbf{0}, \mathbf{I})$, and $\pm$ stands for the signs we get when $i = 1$ and 2. Therefore,

$$\begin{aligned}
\widetilde{\mathbf{H}}_1(\mathbf{X}; \mathbf{\Lambda_1}, \ldots, \mathbf{\Lambda_I}) &= \sum_{i=1}^{I} \mathbf{T}_i'\mathbf{T}_i \\
&= 2\mathbf{Z}' (\mathbf{\Lambda}_1 + \mathbf{\Lambda}_2) \mathbf{Z},
\end{aligned}$$

*which leads to the same test as in (**??**).*

.2 Estimating the Parameters of GMANOVA Model

Consider the Generalized Multivariate Analysis of Variance model

$$\mathbf{Y} = \mathbf{A}\boldsymbol{\gamma}\mathbf{B} + \boldsymbol{\epsilon} \ , \tag{132}$$

where $\mathbf{Y}$ is a $N \times T$ matrix of responses from N subjects taken over T time points, $\mathbf{A}$ is an among subject design matrix of dimension $N \times q$, $\mathbf{B}$ is a $p \times T$ within subject design matrix, and $\boldsymbol{\gamma}$ is a $q \times p$ matrix of parameters. Assume that

$$\boldsymbol{\epsilon} \sim N_{NT}(0, \Sigma \otimes I_N) \ .$$

Then the MLE of $\boldsymbol{\gamma}$ *is*

$$\widehat{\boldsymbol{\gamma}} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}\mathbf{S}^{-1}\mathbf{B}(\mathbf{B}'\mathbf{S}^{-1}\mathbf{B})^{-1},$$

where

$$\mathbf{S} = \mathbf{Y}'(\mathbf{I} - \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}')\mathbf{Y}.$$

To establish this result, let us first choose components of a $T \times T$ *nonsingular matrix* $\mathbf{C} = (\ \mathbf{C}_1 \quad \mathbf{C}_2\ )$ *such that*

$$\mathbf{B}\mathbf{C}_1 = \mathbf{I}_p \text{ and } \mathbf{B}\mathbf{C}_2 = \mathbf{0},$$

where $\mathbf{C}_1$ *is a* $T \times p$ *matrix of rank* p *and* $\mathbf{C}_2$ *is a* $T \times (T-p)$ *matrix of rank* $(T-p)$ *That is such that* $\mathbf{B}\mathbf{C} = (\ \mathbf{I}_p \quad \mathbf{0}\ )$. *For example, we could set* $\mathbf{C}_1 = \mathbf{B}'(\mathbf{B}\mathbf{B}')^{-1}$ *and* $\mathbf{C}_2$ *to be basis of* $\mathbf{I} - \mathbf{B}'(\mathbf{B}\mathbf{B}')^{-1}\mathbf{B}$. *Now consider the transformed observations given by*

$$\begin{aligned}\widetilde{\mathbf{Y}} &= \mathbf{Y}\mathbf{C} \\ &= (\ \mathbf{Y}\mathbf{C}_1 \quad \mathbf{Y}\mathbf{C}_2\ ) \\ &= (\ \mathbf{Y}_1 \quad \mathbf{Y}_2\ ).\end{aligned}$$

The transformed data allows us to obtain a classical linear model, because

$$\begin{aligned}E(\mathbf{Y}_1) &= \mathbf{A}\boldsymbol{\gamma}\mathbf{B}\mathbf{C}_1 \\ &= \mathbf{A}\boldsymbol{\gamma} \qquad (133)\end{aligned}$$

and

$$\begin{aligned}E(\mathbf{Y}_2) &= \mathbf{A}\boldsymbol{\gamma}\mathbf{B}\mathbf{C}_2 \\ &= \mathbf{0}. \qquad (134)\end{aligned}$$

Hence, we can express model (117) as

$$(\ \mathbf{Y}_1 \quad \mathbf{Y}_2\ ) = (\ \mathbf{A} \quad \mathbf{0}\ )\boldsymbol{\gamma} + \mathbf{E},$$

where $\mathbf{E} \sim N_{NT}(0, \widetilde{\Sigma} \otimes I_N)$ *and*

$$\widetilde{\Sigma} = \begin{pmatrix} \mathbf{C}_1'\Sigma\mathbf{C}_1 & \mathbf{C}_2'\Sigma\mathbf{C}_1 \\ \mathbf{C}_1'\Sigma\mathbf{C}_2 & \mathbf{C}_2'\Sigma\mathbf{C}_2 \end{pmatrix}.$$

Therefore, the joint density function of $(\mathbf{Y}_1, \mathbf{Y}_2)$ *can be expressed as the product of the conditional density of* $\mathbf{Y}_1$ *given* $\mathbf{Y}_2$, *and the marginal density of* $\mathbf{Y}_1$ *given by*

$$\mathbf{Y}_1|\mathbf{Y}_2 \sim N_{N(T-p)}(\mathbf{A}\boldsymbol{\gamma} + \mathbf{Y}_2\boldsymbol{\lambda}, \Sigma_{12} \otimes I_N) \qquad (135)$$

and

$$\mathbf{Y}_2 \sim N_{Np}(\mathbf{0}, \mathbf{C}_2'\Sigma\mathbf{C}_2 \otimes I_N), \qquad (136)$$

where

$$\boldsymbol{\lambda} = (\mathbf{C}_2'\Sigma\mathbf{C}_2)^{-1}\mathbf{C}_2'\Sigma\mathbf{C}_1$$

and

$$\begin{aligned}\Sigma_{12} &= \mathbf{C}_1'\Sigma\mathbf{C}_1 - \mathbf{C}_1'\Sigma\mathbf{C}_2(\mathbf{C}_2'\Sigma\mathbf{C}_2)^{-1}\mathbf{C}_2'\Sigma\mathbf{C}_1 \quad (137)\\ &= (\mathbf{B}\boldsymbol{\Sigma}^{-1}\mathbf{B}')^{-1}. \quad (138)\end{aligned}$$

Since the marginal distribution does not involve $\boldsymbol{\gamma}$, the MLE of $\boldsymbol{\gamma}$ is the same as the one obtained by maximizing the log likelihood given by (135). By writing the linear model underlying (135) as

$$\begin{aligned}E(\mathbf{Y}_1|\mathbf{Y}_2) &= (\ \mathbf{A}\quad \mathbf{Y}_2\ )\begin{pmatrix}\boldsymbol{\gamma}\\ \boldsymbol{\lambda}\end{pmatrix}\\ &= \mathbf{Z}\boldsymbol{\theta},\end{aligned}$$

we can obtain the MLE of $\boldsymbol{\gamma}$ by partitioning that of $\boldsymbol{\theta}$, which is the same as the usual least squares estimate

$$\begin{aligned}\widehat{\boldsymbol{\theta}} &= \begin{pmatrix}\widehat{\boldsymbol{\gamma}}\\ \widehat{\boldsymbol{\lambda}}\end{pmatrix} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}_1\\ &= \begin{pmatrix}\mathbf{A}'\mathbf{A} & \mathbf{A}'\mathbf{Y}_2\\ \mathbf{Y}_2'\mathbf{A} & \mathbf{Y}_2'\mathbf{Y}_2\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{A}'\mathbf{Y}_1\\ \mathbf{Y}_2'\mathbf{Y}_1\end{pmatrix}.\end{aligned}$$

Define matrices

$$\begin{aligned}\mathbf{R} &= \mathbf{I} - \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}',\\ \mathbf{S} &= \mathbf{Y}'\mathbf{R}\mathbf{Y},\end{aligned}$$

and

$$\begin{aligned}\mathbf{W} &= (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}_2(\mathbf{Y}_2'\mathbf{Y}_2 - \mathbf{Y}_2'\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}_2)^{-1}.\\ &= (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}_2(\mathbf{Y}_2'\mathbf{R}\mathbf{Y}_2)^{-1}\end{aligned}$$

By applying the formula for inverting a partitioned matrix and using matrix algebra, thus we get the MLE of $\boldsymbol{\gamma}$

$$\begin{aligned}\widehat{\boldsymbol{\gamma}} &= (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}_1 + \mathbf{W}\mathbf{Y}_2'\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}_1 - \mathbf{W}\mathbf{Y}_2'\mathbf{Y}_1\\ &= (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}_1 - \mathbf{W}\mathbf{Y}_2'\mathbf{R}\mathbf{Y}_1\\ &= (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'(\mathbf{Y}_1 - \mathbf{Y}_2(\mathbf{Y}_2'\mathbf{R}\mathbf{Y}_2)^{-1}\mathbf{Y}_2'\mathbf{R}\mathbf{Y}_1)\\ &= (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}(\mathbf{C}_1 - \mathbf{C}_2(\mathbf{C}_2'\mathbf{S}\mathbf{C}_2)^{-1}\mathbf{C}_2'\mathbf{S}\mathbf{C}_1).\end{aligned}$$

But from (135), which is valid for any $\boldsymbol{\Sigma}$,

$$\mathbf{C}_2(\mathbf{C}_2'\mathbf{S}\mathbf{C}_2)^{-1}\mathbf{C}_2' = \mathbf{S}^{-1} - \mathbf{S}^{-1}\mathbf{B}(\mathbf{B}'\mathbf{S}^{-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{S}^{-1}.$$

Therefore, $\widehat{\boldsymbol{\gamma}}$ can be expressed as

$$\begin{aligned}\widehat{\boldsymbol{\gamma}} &= (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}(\mathbf{C}_1 - [\mathbf{S}^{-1} - \mathbf{S}^{-1}\mathbf{B}(\mathbf{B}'\mathbf{S}^{-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{S}^{-1}]\mathbf{S}\mathbf{C}_1)\\ &= (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}\mathbf{S}^{-1}\mathbf{B}(\mathbf{B}'\mathbf{S}^{-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{C}_1\\ &= (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}\mathbf{S}^{-1}\mathbf{B}(\mathbf{B}'\mathbf{S}^{-1}\mathbf{B})^{-1}.\end{aligned}$$