TESTS FOR UNEQUAL TREATMENT VARIANCES IN CROSSOVER DESIGNS

by

YOONSUNG JUNG

B.S., Seowon University, Korea, 1995
M.S., Korea University, Korea, 1999
M.S., Texas A&M University, 2003

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas
2009

Abstract

A crossover design is an experimental design in which each experimental unit receives a series of experimental treatments over time. The order that an experimental unit receives its treatments is called a sequence (example, the sequence AB means that treatment A is given first, and then followed by treatment B). A period is the time interval during which a treatment is administered to the experimental unit. A period could range from a few minutes to several months depending on the study. Sequences usually involve subjects receiving a different treatment in each successive period. However, treatments may occur more than once in any sequence (example, ABAB).

Treatments and periods are compared within subjects, i.e. each subject serves as his/her own control. Therefore, any effect that is related to subject differences is removed from treatment and period comparisons.

Carryover effects are residual effects from a previous treatment manifesting themselves in subsequent periods. Crossover designs both with and without carryover are traditionally analyzed assuming that the response due to different treatments have equal variances. The effects of unequal variances on traditional tests for treatment and carryover difference were recently considered in crossover designs assuming that the response due to treatments have unequal variances with a compound symmetry correlation structure.

The likelihood function for the two treatment/two sequence crossover design has closed form maximum likelihood solutions for the parameters at both the null hypothesis, $H_0 : \sigma_A^2 = \sigma_B^2$, and at alternative hypothesis, $H_A : \sigma_A^2 \neq \sigma_B^2$. Under $H_A : \sigma_A^2 \neq \sigma_B^2$, the method of moment estimators and the maximum likelihood estimators of σ_A^2, σ_B^2 and ρ are identical. The dual balanced design, ABA/BAB, which is balanced for carryover effects is also

considered. The dual balanced design has a closed form solution that maximizes the likelihood function under the null hypothesis, $H_0 : \sigma_A^2 = \sigma_B^2$, but not for the alternative hypothesis, $H_A :$ not H_0. Similarly, the three treatment/three sequence crossover design, $ABC/BCA/CAB$, has a closed form solution that maximizes the likelihood function at the null hypothesis, $H_0 : \sigma_A^2 = \sigma_B^2 = \sigma_C^2$, but not for the alternative hypothesis, $H_A :$ not H_0.

An iterative procedure is introduced to estimate the parameters for the two and three treatment crossover designs. To check the performance of the likelihood ratio tests, Type I error rates and power comparisons are explored using simulations.

TESTS FOR UNEQUAL TREATMENT VARIANCES IN CROSSOVER DESIGNS

by

YOONSUNG JUNG

B.S., Seowon University, Korea, 1995
M.S., Korea University, Korea, 1999
M.S., Texas A&M University, 2003

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas
2009

Approved by:	Approved by:
Co-Major Professor Dallas E. Johnson	Co-Major Professor John E. Boyer, Jr.

Copyright

YOONSUNG JUNG

2009

Abstract

A crossover design is an experimental design in which each experimental unit receives a series of experimental treatments over time. The order that an experimental unit receives its treatments is called a sequence (example, the sequence AB means that treatment A is given first, and then followed by treatment B). A period is the time interval during which a treatment is administered to the experimental unit. A period could range from a few minutes to several months depending on the study. Sequences usually involve subjects receiving a different treatment in each successive period. However, treatments may occur more than once in any sequence (example, ABAB).

Treatments and periods are compared within subjects, i.e. each subject serves as his/her own control. Therefore, any effect that is related to subject differences is removed from treatment and period comparisons.

Carryover effects are residual effects from a previous treatment manifesting themselves in subsequent periods. Crossover designs both with and without carryover are traditionally analyzed assuming that the response due to different treatments have equal variances. The effects of unequal variances on traditional tests for treatment and carryover difference were recently considered in crossover designs assuming that the response due to treatments have unequal variances with a compound symmetry correlation structure.

The likelihood function for the two treatment/two sequence crossover design has closed form maximum likelihood solutions for the parameters at both the null hypothesis, $H_0 : \sigma_A^2 = \sigma_B^2$, and at alternative hypothesis, $H_A : \sigma_A^2 \neq \sigma_B^2$. Under $H_A : \sigma_A^2 \neq \sigma_B^2$, the method of moment estimators and the maximum likelihood estimators of σ_A^2, σ_B^2 and ρ are identical. The dual balanced design, ABA/BAB, which is balanced for carryover effects is also

considered. The dual balanced design has a closed form solution that maximizes the likelihood function under the null hypothesis, $H_0 : \sigma_A^2 = \sigma_B^2$, but not for the alternative hypothesis, $H_A :$ not H_0. Similarly, the three treatment/three sequence crossover design, $ABC/BCA/CAB$, has a closed form solution that maximizes the likelihood function at the null hypothesis, $H_0 : \sigma_A^2 = \sigma_B^2 = \sigma_C^2$, but not for the alternative hypothesis, $H_A :$ not H_0.

An iterative procedure is introduced to estimate the parameters for the two and three treatment crossover designs. To check the performance of the likelihood ratio tests, Type I error rates and power comparisons are explored using simulations.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to express my sincere thanks and appreciation to my major advisor, Dr. Dallas E. Johnson for providing guidance, patience, understanding and support during the course of this study. Also, I would like to thank Dr. John E. Boyer Jr. as a co-advisor and committee member, I would like to thank the members of my graduate committee, Dr. James J. Higgins and Dr. Carol W. Shanklin for reviewing my work and providing many helpful comments.

I would like to thank the Department of Statistics for providing support during the course of my studies.

I would like to thank my wife Ryunghee Kim and my son Andrew Jung for support during the course of my study.

I would like to thank my parents, In-Chul Jeong and Ok-Young Yoon, parents-in-law, Han-Soo Kim and Jung-Sook Cho, and my brothers and sisters for their love and support.

I would like to thank my mentor, Dr. Eunhee Kim, for her nice advice and support.

I would like to thank my god-parents, Raymond and Sun-Cha Felix for their love, prayers and support.

Thank you so much to Meishusama.

Preface

A crossover design is an experimental design in which each experimental unit receives a series of experimental treatments over time. The order that an experimental unit receives its treatments is called a sequence (example, the sequence AB means that treatment A is given first, and then followed by treatment B). A period is the time interval during which a treatment is administered to the experimental unit. A period could range from a few minutes to several months depending on the study. Sequences usually involve subjects receiving a different treatment in each successive period. However, treatments may occur more than once in any sequence (example, ABAB).

Treatments and periods are compared within subjects, i.e. each subject serves as his/her own control. Therefore, any effect that is related to subject differences is removed from treatment and period comparisons.

Carryover effects are residual effects from a previous treatment manifesting themselves in subsequent periods. Crossover designs both with and without carryover are traditionally analyzed assuming that the response due to treatments have equal variances. The effects of unequal variances on traditional tests for the difference of treatment and carryover effects were recently considered in crossover designs assuming that the responses due to treatments have unequal variances with compound symmetry correlation structure.

Tests for the equality of variances due to treatments in crossover designs are considered in this study.

Chapter 1

Analysis of Crossover Designs When Treatments have Equal Variances

1.1 Introduction

A crossover design is an experimental design in which each experimental unit receives a series of experimental treatments over time. The order that an experimental unit receives its treatments is called a sequence (example, the sequence AB means that treatment A is given first, and then followed by treatment B). A period is the time interval during which a treatment is administered to the experimental unit. A period could range from a few minutes to several months depending on the study. Sequences usually involve subjects receiving a different treatment in each successive period. However, treatments may occur more than once in any sequence (example, ABAB).

Many crossover designs have been used in agriculture. Brant (1938)[2] used a crossover design for comparing two treatments using two groups of cattle. Fieller (1940)[4] used a two treatment/two sequence (AB/BA) crossover design with rabbits to compare the effects of different doses of insulin in a biological assay. Crossover designs are also used in pharmaceutical studies where subjects with recurrent or chronic conditions, such as high blood pressure, asthma, epilepsy, and angina, are being used to study different treatments. These designs are suitable in situations where a patient(or a subject) is not cured by one of the treatments during the course of the study. Grizzle (1965)[9] describes some of the advantages

of crossover designs having two experimental periods.

In crossover trials, measurements on different treatments are taken from each subject during or at the end of each period of treatment. Thus, treatments and periods are compared within subjects, i.e. each subject serves as his/her own control. Therefore, any effects that are related to subject differences are removed from treatment and period comparisons.

Crossover designs provide an advantage over parallel group designs (a separate group of subjects for each treatment) because the standard error of a within subject contrast depends on within subject variability in a crossover design rather than between subject variability as in parallel group design. Within subject variability is usually much smaller than between subject variability and allows for more precise estimates of treatment differences than can be made using only between subject contrasts. A crossover design generally requires fewer subjects than a design having a separate group of subjects for each treatment. The fewer number of subjects makes crossover designs more economically viable. This is not possible in a parallel groups design.

Crossover designs also have some disadvantages that should not be overlooked. In clinical trials, patients may drop out of a study. Human subjects may tend to withdraw from a study if they feel they are not gaining relief or benefit from one of the treatments. Patients who drop out provide no direct information on the treatments they did not receive which may make it difficult to analyze and interpret the data. Some diseases with a non-negligible chance of death during the course of the trial are not suitable for crossover designs. Crossover designs with long sequences of treatments may be inconvenient for patients. Subjects are required to commit to a certain number of treatments and the total amount of time spent may be too long for some studies. In crossover designs, treatments given in one period may affect a treatment received in a subsequent period. Residual effects from a previous treatment manifesting themselves in subsequent periods are called carryover effects and can be either a prolonged or delayed response from a previous treatment. Carryover effects that affect the nth period treatment are called nth order carryover effects (e.g. 1^{st} and 2^{nd}

order carryover effects affect the response of treatments applied 1 and 2 periods later). The general view for carryover effects is that the effect of 2^{nd} order carryover is much smaller than that of a 1^{st} order carryover effect. Crossover designs often incorporate a washout period. A washout period is also called a rest period. It is a time during which subjects are not given any of the treatments under investigation. It is hoped that the washout period will eliminate or minimize possible carryover effects of a previous treatment in subsequent periods. However, washout periods are not always feasible because, as a example, a patient with a medical condition that requires continuous medical treatment may not allow for a washout period. In an animal science study, animals may be put on different sequences of feed rations to evaluate daily weight gain. In many of these cases, once the treatment is changed, the researcher can leave the animal on the new treatment for a longer period of time before collecting responses in order to minimize carryover effects.

Crossover designs are traditionally analyzed as a split-plot design. A split-plot design has two different sizes of experimental units (whole plot and subplot experimental units). The whole plot in a crossover design is the subject to which a sequence of treatments is assigned and the subplot is the time interval (period). There are both between subject comparisons and within subject comparisons in a crossover design. Treatment and period effects are compared within subjects while sequence effects are compared between subjects.

Multivariate and mixed model analyses of crossover designs with more than two periods were considered by Goad (1994)[6]. Goad considered a crossover experiment as a form of a repeated measures experiment. An appropriate analysis of a repeated measures experiment depends on the form of the variance-covariance matrix, $\mathbf{\Sigma}$, of the repeated measures. Huynh and Feldt (1970)[12] defined type H structure of $\mathbf{\Sigma}$ and proved it to be a necessary and sufficient condition for the within subject analysis of variance F-tests to be valid. In a crossover experiment where $\mathbf{\Sigma}$ does not have type H structure and the analysis of variance tests may not be valid, three alternative approaches were proposed by Goad. The first approach approximates the distribution of the usual analysis of variance F-statistic with

reduced numerator and denominator degrees of freedom proposed by Greenhouse and Geisser (1958,1959)[5 8] and Huynh and Feldt (1976)[13]. The second method gives approximate F-tests for simultaneous orthogonal contrasts and approximate t-tests for simple linear contrasts. The third method uses mixed models techniques to form approximate F-tests and t-tests.

Shanga (2003)[28] studied the effects of unequal treatment variances on the tests for equal treatment effects and/or equal carryover effects in two treatment and three treatment crossover designs. He generated crossover data under four scenarios-(1) equal variances/no carryover, (2) equal variances/carryover, (3) unequal variances/no carryover, and (4) unequal variances/carryover. He then proceeded to analyze the generated data sets using assumptions corresponding to these same four scenarios. Shanga also introduced a method to test the equality of the treatment variances assuming the correlation matrix corresponding to the period measurements satisfied compound symmetry. Shanga proposed using a likelihood ratio type test statistic where the parameters under both the null and the alternative were estimated by the method of moments.

In this paper, a correlation structure is considered on the measurements taken over time for each subject. Since each subject assigned to a particular sequence is given different treatments over time (say, ABC), the correlation between measurements taken in the successive time periods may not be negligible. It is often reasonable to assume a constant correlation for these measurements. Correlations between measurements taken over time that satisfy compound symmetry are considered in this study. Compound symmetry assumes that the correlations between measurements taken at any two time periods are equal.

There are two models usually considered in the analysis of crossover designs. The first model is a model that does not include parameters for carryover effects. Consider the model without carryover used by Milliken and Johnson (1992)[21]. This model is

$$y_{ijkl} = \mu + s_i + \delta_{il} + \tau_j + \pi_k + \epsilon_{ijkl} \tag{1.1}$$

where

μ is effect of an overall mean;

s_i is effect of the ith sequence effect, $i = 1, 2, \cdots, s$;

δ_{il} is the experimental error associated with the lth subject in the ith sequence;

τ_j is effect of the jth treatment effect, $j = 1, 2, \cdots, t$;

π_k is effect of the kth period effect, $k = 1, 2, \cdots, p$;

ϵ_{ijkl} is the error associated with the lth subject in the ith sequence that received the jth treatment in the kth period, $l = 1, 2, \cdots, n_i$.

$\delta_{il} \sim iid\ N(0, \sigma_\delta^2)$ and $\epsilon_{ijkl} \sim iid\ N(0, \sigma_\epsilon^2)$ for all i, j, k, and l.

all δ_{il}'s and ϵ_{ijkl}'s are independent.

A second model includes parameters for carryover effects. The model for the response variable may be written by modifying a notation used by Ratkowsky, Evan, and Alldredge(1993)[27]. The model is

$$y_{ijklm} = \mu + s_i + \delta_{il} + \tau_j + \pi_k + \lambda_m + \epsilon_{ijklm} \tag{1.2}$$

where μ, s_i, τ_j, π_k, and ϵ_{ijklm} are defined as in (1.1)

λ_m is the carryover effect of the mth treatment administered in the previous period. where $m = 1, 2, \cdots, t$.

In both models the values for j and m are determined by the combination of i and k. Also note that there is no carryover parameter associated with the first period. The general form of an ANOVA table for crossover designs analyzed as a split-plot design for (1.1) is given in Table 1.1.

Table 1.1: *ANOVA Table for a Crossover Design Without Carryover Effects*

Source	df
Between Subject Analysis	
Sequence	s-1
Subject(sequence)	N-s
Within Subject Analysis	
Treatment	t-1
Period	p-1
Error	(N-1)(p-1)-(t-1)
Total	Np-1

The general form of an ANOVA table for crossover designs analyzed as a split-plot design for (1.2) is given in Table 1.2.

Table 1.2: *ANOVA Table for a Crossover Design With Carryover Effects*

Source	df
Between Subject Analysis	
Sequence	s-1
Subject(sequence)	N-s
Within Subject Analysis	
Treatment	t-1
Period	p-1
Carryover	t-1
Error	(N-1)(p-1)-2(t-1)
Total	Np-1

Let $\boldsymbol{y}_{il}$ be the $p \times 1$ vector of responses for the lth subject in the ith sequence and let $\boldsymbol{\epsilon}_{il}$ be the corresponding vector of random errors. Models (1.1) and (1.2) can be written as

$$\boldsymbol{y}_{il} = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{\epsilon}_{il},\ i = 1, 2, \cdots, s;\ l = 1, 2, \cdots, n_i \tag{1.3}$$

where $\boldsymbol{\beta} = (\mu, s_1, \cdots, s_s, \tau_1, \cdots, \tau_t, \pi_1, \cdots, \pi_k)'$ for model (1.1) and $\boldsymbol{\beta} = (\mu, s_1, \cdots, s_s, \tau_1, \cdots, \tau_t, \pi_1, \cdots, \pi_k, \lambda_1, \cdots, \lambda_t)'$ for model (1.2). The elements in the design matrix $\boldsymbol{X}_i$ depend on the parameters associated with the ith sequence. In this paper, it is assumed that $\boldsymbol{\epsilon}_{il} \sim iid\ N(\boldsymbol{0}, \boldsymbol{\Sigma}_i)$, $i = 1, 2, \cdots, s$ and $l = 1, 2, \cdots, n_i$.

Carryover effects present problems in the analysis of crossover designs. Sequence, treatment, and period effects may be aliased with carryover effects. The problem of aliasing is encountered in designs such as those in Tables 1.3 and 1.4.

Table 1.3: *Two Treatment/Two Sequence Crossover Design (AB/BA)*

	Period	
Sequence	1	2
1	A	B
2	B	A

Table 1.4: *Three Treatment/Three Sequence Crossover Design (ABC/BCA/CAB)*

	Period		
Sequence	1	2	3
1	A	B	C
2	B	C	A
3	C	A	B

Consider the design in Table 1.4. Note that treatment B always follows treatment A unless treatment B is given during period 1. Treatment A never follows treatment B. As other patterns, treatment A always follows treatment C and treatment C always follows treatment B. Therefore, if treatment C has a 1^{st} order carryover effect, this will always affect the outcome of treatment A, but it will never affect the outcome of treatment B. This makes it impossible to distinguish the effect of treatment A from the carryover effect of treatment C in this design. Therefore, we are interested in crossover designs balanced for carryover effects where direct treatment comparisons can be made when carryover effects exist.

The following sections consider the analysis of several different crossover designs.

1.2 Two Treatment/Two Sequence Crossover Design (AB/BA)

The two treatment/two sequence crossover design is the simplest of all crossover designs. It is also referred to as 2×2 crossover design. In the 2×2 crossover design, each treatment is administered first in one sequence and last in the other sequence. AB is the order of treatments A and B in the first sequence and BA is the order of the treatments in the second sequence. Table 1.3 gives the sequences in a 2×2 crossover design.

Consider the 2×2 crossover design in Table 1.3. Under the assumptions on the random effects given in (1.1), the covariance structure of the measurements on a subject in either sequence 1 or 2 is

$$\boldsymbol{\Sigma}_i = cov(\boldsymbol{y}_{il}) = \begin{pmatrix} \sigma_\epsilon^2 + \sigma_\delta^2 & \sigma_\delta^2 \\ \sigma_\delta^2 & \sigma_\epsilon^2 + \sigma_\delta^2 \end{pmatrix}, \tag{1.4}$$

where $i = 1, 2;\ l = 1, 2, \cdots, n_i$.

The covariance matrices can also be reparameterized as

$$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

where

$$\sigma^2 = \sigma_\epsilon^2 + \sigma_\delta^2 \quad \text{and} \quad \rho = \frac{\sigma_\delta^2}{\sigma_\epsilon^2 + \sigma_\delta^2}.$$

Let $\boldsymbol{\mu}_{ik}$ be the expected response in the kth period of the ith sequence. Table 1.5 shows expected cell means for model (1.1). Each cell mean is estimable if each cell is observed at least once.

Table 1.5: *Expected Cell Means for Model (1.1)*

Sequence	Period	
	1	2
1	$\boldsymbol{\mu}_{11} = \mu + s_1 + \tau_A + \pi_1$	$\boldsymbol{\mu}_{12} = \mu + s_1 + \tau_B + \pi_2$
2	$\boldsymbol{\mu}_{21} = \mu + s_2 + \tau_B + \pi_1$	$\boldsymbol{\mu}_{22} = \mu + s_2 + \tau_A + \pi_2$

The general ANOVA table for the 2×2 crossover design analyzed as a split-plot design for (1.1) is given in Table 1.6 where n_i equals the number of subjects assigned to the i^{th} sequence, $i = 1, 2$.

Table 1.6: *ANOVA Table for a Two Treatment/ Two Sequence Crossover Design Without Carryover Effects (AB/BA)*

Source	df
Between Subject Analysis	
Sequence	1
Subject(sequence)	$(n_1 + n_2)$-2
Within Subject Analysis	
Treatment	1
Period	1
Error	$(n_1 + n_2)$-2
Total	$2(n_1 + n_2)$-1

Table 1.7 shows the expected cell means for model (1.2).

Table 1.7: *Expected Cell Means for Model (1.2)*

Sequence	Period	
	1	2
1	$\boldsymbol{\mu}_{11} = \mu + s_1 + \tau_A + \pi_1$	$\boldsymbol{\mu}_{12} = \mu + s_1 + \tau_B + \pi_2 + \lambda_A$
2	$\boldsymbol{\mu}_{21} = \mu + s_2 + \tau_B + \pi_1$	$\boldsymbol{\mu}_{22} = \mu + s_2 + \tau_A + \pi_2 + \lambda_B$

The general ANOVA table for the 2×2 crossover design analyzed as a split-plot design for (1.2) is given in Table 1.8.

Table 1.8: *ANOVA Table for a Two Treatment/Two Sequence Crossover Design With Carryover Effects (AB/BA)*

Source	df
Between Subject Analysis	
Sequence/Carryover	1
Subject(sequence)	$(n_1 + n_2)$-2
Within Subject Analysis	
Treatment	1
Period	1
Carryover	1
Error	$(n_1 + n_2)$-3
Total	$2(n_1 + n_2)$-1

For the two treatment/ two period crossover design with carryover effect, consider a linear contrast of the form $\sum_{i=1}^{2}\sum_{k=1}^{2} c_{ik}\boldsymbol{\mu}_{ik}$, where the c_{ik}'s are contrasts such that $\sum_{i=1}^{2}\sum_{k=1}^{2} c_{ik} = 0$. The linear contrast to obtain the difference between sequences 1 and 2 is $s_1 - s_2$. For Table 1.7, $\sum_{i=1}^{2}\sum_{k=1}^{2} c_{ik}\boldsymbol{\mu}_{ik}$ is

$$\begin{aligned}
\sum_{i=1}^{2}\sum_{k=1}^{2} c_{ik}\boldsymbol{\mu}_{ik} &= c_{11}\boldsymbol{\mu}_{11} + c_{12}\boldsymbol{\mu}_{12} + c_{21}\boldsymbol{\mu}_{21} + c_{22}\boldsymbol{\mu}_{22} \\
&= c_{12}\left(\mu + s_1 + \tau_A + \pi_1\right) + c_{12}\left(\mu + s_1 + \tau_B + \pi_2 + \lambda_A\right) \\
&\quad + c_{21}\left(\mu + s_2 + \tau_B + \pi_1\right) + c_{22}\left(\mu + s_2 + \tau_A + \pi_2 + \lambda_B\right) \\
&= \left(c_{11} + c_{12} + c_{21} + c_{22}\right)\mu + \left(c_{11} + c_{12}\right)s_1 + \left(c_{21} + c_{22}\right)s_2 + c_{12}\lambda_A + c_{22}\lambda_B \\
&\quad + \left(c_{11} + c_{22}\right)\tau_A + \left(c_{12} + c_{21}\right)\tau_B + \left(c_{11} + c_{21}\right)\pi_1 + \left(c_{12} + c_{22}\right)\pi_2
\end{aligned}$$

Setting $\sum_{i=1}^{2}\sum_{k=1}^{2} c_{ik}\boldsymbol{\mu}_{ik} = s_1 - s_2$ requires that

$$\left(c_{11} + c_{22}\right)\tau_A + \left(c_{12} + c_{21}\right)\tau_B = 0 \quad \text{and} \quad \left(c_{11} + c_{21}\right)\pi_1 + \left(c_{12} + c_{22}\right)\pi_2 = 0.$$

Together these two equations place the following restraints on the c_{ik}'s:

$$c_{11} = c_{12} = -c_{21} = -c_{22}.$$

These restraints also yield

$$\begin{aligned}\sum_{i=1}^{2}\sum_{k=1}^{2} c_{ik}\boldsymbol{\mu}_{ik} &= (c_{11}+c_{12})(s_1-s_2)+c_{12}(\lambda_A-\lambda_B)\\ &= 2c_{11}(s_1-s_2)+c_{11}(\lambda_A-\lambda_B) \quad \text{since} \quad c_{11}=c_{12}\end{aligned}$$

which equals $s_1 - s_2$ only if $\lambda_A = \lambda_B$ and $c_{11} = \frac{1}{2}$. If there are unequal carryover effects from each of the treatments, then the sequence difference cannot be separated from carryover effects. It is said that the sequence and carryover effects are confounded with one another.

By using similar arguments, it can be shown that treatment and carryover effects are also confounded with one another; that is, there exists c_{ik} such that $\sum_{i=1}^{2}\sum_{k=1}^{2} c_{ik}\boldsymbol{\mu}_{ik} = \tau_1 - \tau_2$ if and only if $\lambda_A = \lambda_B$. Also, period and carryover effects are confounded with one another because there exists c_{ik} such that $\sum_{i=1}^{2}\sum_{k=1}^{2} c_{ik}\boldsymbol{\mu}_{ik} = \pi_1 - \pi_2$ if and only if $\lambda_A + \lambda_B = 0$.

The fact that sequence and carryover effects are confounded with one another can also be seen from the expected mean squares in Table 1.9. The carryover effect has been removed from the model because testing for sequence effect is equivalent to testing for carryover effect in the 2×2 crossover design. That is, in the model without carryover, the difference in the two sequence means should be equal to zero and the difference is nonzero only when carryover exists. So without loss of generality, one can remove s_i from the model (1.2).

Note that Table 1.9 also shows that carryover effects are aliased with treatment and period effects. If the AB/BA crossover design has been used and carryover effects are present, Grizzle (1965)[9] used only period one data to estimate the treatment difference since period one data does not have carryover. However, the estimated standard error of the treatment difference may be computed by using both periods. Grizzle (1965)[9] suggested that a test for carryover be done prior to testing for equal treatment effects and that carryover effects be tested at a 10% significant level since the test for carryover is a between subject comparison. The test for carryover is a between subjects test which has more variability and thus is not as powerful as a within subjects test that has smaller variability. If the null hypothesis for equal carryover effects is rejected, then Grizzle recommends that only the

Table 1.9: *ANOVA Table for a Two Treatment/Two Sequence Crossover Design*

Source	df	SS	EMS
Sequence	1	$\frac{2n_1n_2}{n_1+n_2}\left(\overline{y}_{1\ldots}-\overline{y}_{2\ldots}\right)^2$	$\sigma_\varepsilon^2+2\sigma_\delta^2+\frac{2n_1n_2}{n_1+n_2}\left(\lambda_A-\lambda_B\right)^2$
Subject(Sequence)	n_1+n_2-2	$\sum_{i=1}^{2}\sum_{l=1}^{n_i}\frac{y_{i..l}^2}{2}-\sum_{i=1}^{2}\frac{y_{i\ldots}^2}{2n_i}$	$\sigma_\varepsilon^2+2\sigma_\delta^2$
Treatment	1	$\frac{2n_1n_2}{n_1+n_2}\left(\overline{y}_{111.}-\overline{y}_{122.}+\overline{y}_{221.}-\overline{y}_{212.}\right)^2$	$\sigma_\varepsilon^2+\frac{2n_1n_2}{n_1+n_2}\left[2(\tau_1-\tau_2)-(\lambda_A-\lambda_B)\right]^2$
Period	1	$\frac{2n_1n_2}{n_1+n_2}\left(\overline{y}_{111.}-\overline{y}_{122.}+\overline{y}_{221.}-\overline{y}_{212.}\right)^2$	$\sigma_\varepsilon^2+\frac{2n_1n_2}{n_1+n_2}\left[2(\pi_1-\pi_2)-(\lambda_A+\lambda_B)\right]^2$
Error(Time)	n_1+n_2-2	$\sum_{i=1}^{2}\sum_{k=1}^{2}\sum_{l=1}^{n_i}y_{i.kl}^2-\sum_{i=1}^{2}\sum_{l=1}^{n_i}\frac{y_{i..l}^2}{2}$ $-\sum_{i=1}^{2}\sum_{k=1}^{2}\frac{y_{i.k.}^2}{2n_i}+\sum_{i=1}^{2}\frac{y_{i\ldots}^2}{2n_i}$	σ_ε^2
Total	$2(n_1+n_2)-1$	$\sum_{i=1}^{2}\sum_{k=1}^{2}\sum_{l=1}^{n_i}y_{i.kl}^2-\frac{y_{\ldots.}^2}{2(n_1+n_2)}$	

first period data be used for testing for equal treatment effects. If carryover effects are not present in the model, then data from both periods can be used in testing for equal treatment and equal period effects.

If a researcher knows that carryover effects will occur in a 2×2 crossover design, then (s)he can consider some modifications to standard crossover designs to completely separate treatment and carryover effects. The modification usually involves the addition of extra periods or extra sequences, or both. Two treatment crossover designs that avoid confounding between treatment and carryover effects use an extra treatment period. Extra-period crossover designs are known as a dual balanced designs and were used by Patterson and Lucas (1959)[25] and Ratkowsky, Evans and Alldredge (1993)[27]. Dual balanced designs are designs having a dual sequence in which the treatment order is reversed between the two sequences. The construction of these designs simply involves repeating the treatments given to each subject in the last period for one extra period, ABB and BAA. Treatment sequences given by ABA/BAB is another example of a dual balanced design. The ABA/BAB design can be combined with the ABB/BAA design to make the $ABA/ABB/BAB/BAA$ design. Dual balanced designs are able to make a comparison of treatment effects that is free from carryover using within subject contrasts even in the presence of carryover (Jones and Kenward, 2003)[16].

Table 1.10 shows the ABA/BAB crossover design.

Table 1.10: *Two Treatment/Two Sequence Crossover Design (ABA/BAB)*

	Period		
Sequence	1	2	3
1	A	B	A
2	B	A	B

Under the assumptions given in (1.2), the variance-covariance matrices for the two sequences in the two treatment/two period design are given by

$$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

where

$$\sigma^2 = \sigma_\delta^2 + \sigma_\epsilon^2 \text{ and } \rho = \frac{\sigma_\epsilon^2}{\sigma_\delta^2 + \sigma_\epsilon^2}.$$

There is another modification of a two treatment/two period crossover design that is balanced for carryover effects. The design is a special case of two period designs having t^2 sequences proposed by Balaam (1968)[1] where t is the number of treatments. In general, Balaam designs have t treatments, t^2 sequences, and only two periods in which the treatments appear in all combinations (Ratkowsky, Evans and Alldredge, 1993)[27].

Table 1.11 is an example of a Balaam design, $AB/BA/AA/BB$, which has two extra sequences.

Table 1.11: *Two Treatment/Four Sequence Crossover Design (AB/BA/AA/BB)*

	Period	
Sequence	1	2
1	A	B
2	B	A
3	A	A
4	B	B

1.3 Three or More Treatments Crossover Design

This section considers crossover designs with three or more treatments. Table 1.12 shows a basic cyclic three treatment/ three sequence crossover design.

Table 1.12: *Three Treatment/Three Sequence Crossover Design (ABC/BCA/CAB)*

	Period		
Sequence	1	2	3
1	A	B	C
2	B	C	A
3	C	A	B

Under the assumptions on the random effects given in (1.1), the covariance structure of the measurements on a subject in each sequence is

$$\boldsymbol{\Sigma}_i = cov(\boldsymbol{y}_{il}) = \begin{pmatrix} \sigma_\epsilon^2 + \sigma_\delta^2 & \sigma_\delta^2 & \sigma_\delta^2 \\ \sigma_\delta^2 & \sigma_\epsilon^2 + \sigma_\delta^2 & \sigma_\delta^2 \\ \sigma_\delta^2 & \sigma_\delta^2 & \sigma_\epsilon^2 + \sigma_\delta^2 \end{pmatrix}, \tag{1.5}$$

where $i = 1, 2, 3;\ l = 1, 2, \cdots, n_i$.

It is also reparameterized by

$$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_3 = \sigma^2 \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix}$$

where

$$\sigma^2 = \sigma_\epsilon^2 + \sigma_\delta^2 \quad \text{and} \quad \rho = \frac{\sigma_\delta^2}{\sigma_\epsilon^2 + \sigma_\delta^2}.$$

Let $\boldsymbol{\mu}_{ik}$ be the expected response in the kth period of the ith sequence. Table 1.13 shows the expected cell means for model (1.1). Each cell mean is estimable if there is at least one observation in each of the cells.

Table 1.13: *Expected Cell Means for a Three Treatment/Three Sequence Crossover Design (ABC/BCA/CAB) for Model (1.1)*

Sequence	Period		
	1	2	3
1	$\boldsymbol{\mu}_{11} = \mu + s_1 + \tau_A + \pi_1$	$\boldsymbol{\mu}_{12} = \mu + s_1 + \tau_B + \pi_2$	$\boldsymbol{\mu}_{13} = \mu + s_1 + \tau_C + \pi_3$
2	$\boldsymbol{\mu}_{21} = \mu + s_2 + \tau_B + \pi_1$	$\boldsymbol{\mu}_{22} = \mu + s_2 + \tau_C + \pi_2$	$\boldsymbol{\mu}_{23} = \mu + s_2 + \tau_A + \pi_3$
3	$\boldsymbol{\mu}_{31} = \mu + s_3 + \tau_C + \pi_1$	$\boldsymbol{\mu}_{32} = \mu + s_3 + \tau_A + \pi_2$	$\boldsymbol{\mu}_{33} = \mu + s_3 + \tau_B + \pi_3$

The general ANOVA table for the three treatment/ three sequence crossover design analyzed as a split-plot design for model (1.1) is given in Table 1.14 where n_i equals the number of subjects assigned to the i^{th} sequence, $i = 1, 2, 3$.

Table 1.14: *ANOVA Table for Model (1.1) ANOVA Table for a Three Treatment/ Three Sequence Crossover Design (ABC/BCA/CAB) Without Carryover Effects*

Source	df
Between Subject Analysis	
Sequence	2
Subject(sequence)	$(n_1 + n_2 + n_3)$-3
Within Subject Analysis	
Treatment	2
Period	2
Error	$2(n_1 + n_2 + n_3)$-4
Total	$3(n_1 + n_2 + n_3)$-1

The analysis in Table 1.14 is valid only when there are no carryover effects. The expected cell means for model (1.2) are given in Table 1.15.

Table 1.15: *Expected Cell Means for a Three Treatment/Three Sequence Crossover Design (ABC/BCA/CAB) for Model (1.2)*

Sequence	Period		
	1	2	3
1	$\boldsymbol{\mu}_{11} = \mu + s_1 + \tau_A + \pi_1$	$\boldsymbol{\mu}_{12} = \mu + s_1 + \tau_B + \pi_2 + \lambda_A$	$\boldsymbol{\mu}_{13} = \mu + s_1 + \tau_C + \pi_3 + \lambda_B$
2	$\boldsymbol{\mu}_{21} = \mu + s_2 + \tau_B + \pi_1$	$\boldsymbol{\mu}_{22} = \mu + s_2 + \tau_C + \pi_2 + \lambda_B$	$\boldsymbol{\mu}_{23} = \mu + s_2 + \tau_A + \pi_3 + \lambda_C$
3	$\boldsymbol{\mu}_{31} = \mu + s_3 + \tau_C + \pi_1$	$\boldsymbol{\mu}_{32} = \mu + s_3 + \tau_A + \pi_2 + \lambda_C$	$\boldsymbol{\mu}_{33} = \mu + s_3 + \tau_B + \pi_3 + \lambda_A$

The general ANOVA table for model (1.2) with the $ABC/BCA/CAB$ crossover design is given in Table 1.16.

Table 1.16: *ANOVA Table for Model (1.2) ANOVA Table for a Three Treatment/Three Sequence Crossover Design (ABC/BCA/CAB) With Carryover Effects*

Source	df
Between Subject Analysis	
Sequence/Carryover	2
Subject(sequence)	$(n_1 + n_2 + n_3)$-3
Within Subject Analysis	
Treatment	2
Period	2
Carryover	2
Error	$2(n_1 + n_2 + n_3)$-6
Total	$3(n_1 + n_2 + n_3)$-1

In the three treatment/ three sequence crossover design, a Williams' design is recommended. Williams (1949)[30] developed a procedure using Latin square designs that balances crossover designs for carryover effects where the number of periods equals the number of treatments for three or more treatments ($p = t > 2$). The Williams' designs have the characteristic that every treatment follows every other treatment an equal number of times. When implementing the Williams' design procedure, two special Latin squares are generated for all experiments involving three or more treatments. Designs that have an odd number of treatments require both of the Latin squares to achieve balance for carryover effects, and designs that have an even number of treatments require only one Latin square. If the number of treatments is odd, each treatment follows every other treatment twice. If the number of treatments is even, each treatment follow every other treatment once. Matthews (1988)[20] gives a clear description of a general algorithm for developing the Williams' designs. The same description is also given by Jones and Kenward (2003)[16].

Two Latin squares, six different sequences, are used to construct sequences such that

each treatment immediately follows every other treatment exactly twice. Such a design is shown in Table 1.17. Using the design in Table 1.17, treatment contrasts can be constructed that are free from carryover effects.

Table 1.17: *A Three Treatment/Six Sequence Crossover Design (Two Latin Squares' Design: ABC/ACB/BAC/BCA/CAB/CBA)*

	Period		
Sequence	1	2	3
1	A	B	C
2	B	C	A
3	C	A	B
4	A	C	B
5	B	A	C
6	C	B	A

Consider a four treatment/four period Williams' design. A four treatment/ four period design using one Latin square that is balanced for carryover is shown in Table 1.18. Here each treatment immediately follows every other treatment one time.

Table 1.18: *A Four Treatment/Four Sequence Crossover Design (One Latin Square Design: ABDC/BCAD/CDBA/DACB)*

	Period			
Sequence	1	2	3	4
1	A	B	D	C
2	B	C	A	D
3	C	D	B	A
4	D	A	C	B

Shanga (2003)[28] considered the two treatment/two sequence crossover design and the three treatment/six sequence crossover design. Shanga focused on testing for carryover, period, and treatment effects when treatments have unequal variances following Goad's approaches (1994)[6]. Shanga also introduced a test of equality of the two or three treatment variances. Chapters 2 and 3 show and evaluate a likelihood ratio test to test the equality of

variances when there are two or three treatments. The properties of these likelihood ratio tests are explored using simulations.

Chapter 2

Testing for Equal Treatment Variances in Two Treatment Crossover Designs

2.1 Introduction

A crossover design is an experimental design in which each experimental unit receives a series of experimental treatments over time. The order that an experimental unit receives its treatments is called a sequence (example, the sequence ABC means that treatment A is given first, and then followed by treatment B, then followed by treatment C). A period is the time interval during which a treatment is administered to the experimental unit. A period could range from a few minutes to several months depending on the study. Sequences usually involve subjects receiving a different treatment in each successive period. However, treatments may occur more than once in any sequence (example, ABAB).

A two treatment/two sequence crossover design (AB/BA) is the simplest of all crossover designs. It is referred to as a 2×2 crossover design. In the 2×2 crossover design, AB is the order of the treatments A and B in the first sequence and BA is the order of the treatments B and A in the second sequence. Other two treatment designs can be formed by adding extra period(s) and/or sequence(s) to the 2×2 crossover design. ABA/BAB is an example of an extra period crossover design known as a dual balanced design. $AB/BA/AA/BB$ is an example of an extra sequence crossover design known as a Balaam design.

Consider the 2×2 crossover design given in Table 2.1, and suppose there are n_1 subjects assigned to the AB sequence and n_2 subjects assigned to the BA sequence. There are two models usually considered in the traditional analysis of crossover designs. The first model is a model that does not include parameters for carryover effects. Consider the model without carryover used by Milliken and Johnson (1992)[21]. This model is

$$y_{ijkl} = \mu + s_i + \delta_{il} + \tau_j + \pi_k + \epsilon_{ijkl} \tag{2.1}$$

where

μ is effect of an overall mean;

s_i is effect of the ith sequence effect, $i = 1, 2, \cdots, s$;

δ_{il} is the experimental error associated with the lth subject in the ith sequence;

τ_j is effect of the jth treatment effect, $j = 1, 2, \cdots, t$;

π_k is effect of the kth period effect, $k = 1, 2, \cdots, p$;

ϵ_{ijkl} is the error associated with the lth subject in the ith sequence that received the jth treatment in the kth period, $l = 1, 2, \cdots, n_i$.

Also, it is often assumed that $\delta_{il} \sim iid\ N(0, \sigma_\delta^2)$ and $\epsilon_{ijkl} \sim iid\ N(0, \sigma_\epsilon^2)$ for all i, j, k, and l. Finally, it is usually assumed that all δ_{il}'s and ϵ_{ijkl}'s are independent. The value of j is determined by i and k.

A second model includes parameters for carryover effects. The model for the response variable may be written by modifying a notation used by Ratkowsky, Evan, and Alldredge(1993)[27]. The model is

$$y_{ijklm} = \mu + s_i + \delta_{il} + \tau_j + \pi_k + \lambda_m + \epsilon_{ijklm} \tag{2.2}$$

where μ, s_i, τ_j, π_k, and ϵ_{ijklm} are defined as in (2.1) above, and λ_m is the carryover effect of the mth treatment administered in the previous period, where $m = 1, 2, \cdots, t$. There is no carryover parameter associated with the first period.

Let $\boldsymbol{y}_{il}$ be the $p \times 1$ vector of responses for the l^{th} subject in the i^{th} sequence and let $\boldsymbol{\epsilon}_{il}$ be the corresponding vector of random errors. Models (2.1) and (2.2) can be written as

$$\boldsymbol{y}_{il} = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{\epsilon}_{il},\ i = 1, 2, \cdots, s;\ l = 1, 2, \cdots, n_i \tag{2.3}$$

where $\boldsymbol{\beta} = (\mu, s_1, \cdots, s_s, \tau_1, \cdots, \tau_t, \pi_1, \cdots, \pi_k)'$ for model (2.1) and $\boldsymbol{\beta} = (\mu, s_1, \cdots, s_s, \tau_1, \cdots, \tau_t, \pi_1, \cdots, \pi_k, \lambda_1, \cdots, \lambda_t)'$ for model (2.2). The elements in the design matrix $\boldsymbol{X}_i$ depend on the sequence of treatments in the i^{th} sequence. In this paper, it is assumed that $\boldsymbol{\epsilon}_{il} \sim iid\ N(\boldsymbol{0}, \boldsymbol{\Sigma}),\ i = 1, 2, \cdots, s$ and $l = 1, 2, \cdots, n_i$.

Consider a two treatment/two period crossover design (AB/BA) with $s = 2$ in (2.3). Under the assumptions on the random effects given in (2.1), the covariance structure of the measurements on a subject in either sequence 1 or sequence 2 is

$$\boldsymbol{\Sigma} = cov(\boldsymbol{y}_{il}) = \begin{pmatrix} \sigma_\epsilon^2 + \sigma_\delta^2 & \sigma_\delta^2 \\ \sigma_\delta^2 & \sigma_\epsilon^2 + \sigma_\delta^2 \end{pmatrix},\ i = 1, 2;\ l = 1, 2, \cdots, n_i. \tag{2.4}$$

Table 2.1: *Two Treatment/Two Sequence Crossover Design(AB/BA)*

	Period	
Sequence	1	2
1	A	B
2	B	A

Since each subject assigned to a particular sequence is given different treatments over time (say, ABCD), the correlation between measurements taken on a given subject in any two periods can not be assumed to be negligible. It is often reasonable to assume a constant correlation for these measurements when sequences are short. Short sequences will be defined as those having two or three periods. Correlations between measurements taken over time that satisfy compound symmetry are considered. Compound symmetry means that the correlation between measurements from any two periods is constant. The constant correlation between any two periods will be denoted by ρ and the j^{th} treatment variance will be denoted by $\sigma_j^2,\ j = A, B$.

Using the preceding assumptions, the correlation matrix for the 2×2 crossover design will be denoted by $\mathbf{R}$ and $\mathbf{R} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. The correlation between period measurements together with the unequal variances due to the two treatments yield variance-covariance matrices for sequences 1 and 2, respectively, as

$$\mathbf{\Sigma}_1 = cov(\boldsymbol{y}_{1l}) = \begin{pmatrix} \sigma_A^2 & \rho\sigma_A\sigma_B \\ \rho\sigma_A\sigma_B & \sigma_B^2 \end{pmatrix}$$

and

$$\mathbf{\Sigma}_2 = cov(\boldsymbol{y}_{2l}) = \begin{pmatrix} \sigma_B^2 & \rho\sigma_A\sigma_B \\ \rho\sigma_A\sigma_B & \sigma_A^2 \end{pmatrix} \tag{2.5}$$

where σ_A^2 and σ_B^2 are variances due to treatments A and B, respectively, and ρ is the correlation between observations in periods 1 and 2 for the same subject. If $\sigma_A^2 = \sigma_B^2 = \sigma^2$, then $cov(\boldsymbol{y}_{il}) = \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ for $i = 1, 2$ and $l = 1, 2, \cdots, n_i$. This covariance structure is equivalent to the one given in (2.4) with $\sigma^2 = \sigma_\epsilon^2 + \sigma_\delta^2$ and $\rho = \frac{\sigma_\delta^2}{\sigma_\epsilon^2 + \sigma_\delta^2}$.

The general form of the ANOVA table for the 2×2 crossover design when $\sigma_A^2 = \sigma_B^2$ and analyzed as a split-plot design for (2.1) is given in Table 2.2.

Table 2.2: *ANOVA Table for Model 2.1: Two Treatment/ Two Sequence Crossover Design (AB/BA) Without Carryover Effects*

Source	df
Between Subject Analysis	
Sequence	1
Subject(sequence)	$(n_1 + n_2)$-2
Within Subject Analysis	
Treatment	1
Period	1
Error	$(n_1 + n_2)$-2
Total	$2(n_1 + n_2)$-1

Table 2.3 shows the expected cell means for model (2.1). Each cell mean is estimable if there is at least one observation in each cell.

Table 2.3: *Expected Cell Means for a Two Treatment/ Two Sequence Crossover Design (AB/BA) for Model (2.1)*

Sequence	Period	
	1	2
1	$\boldsymbol{\mu}_{11} = \mu + \tau_A + \pi_1$	$\boldsymbol{\mu}_{12} = \mu + \tau_B + \pi_1$
2	$\boldsymbol{\mu}_{21} = \mu + \tau_B + \pi_1$	$\boldsymbol{\mu}_{22} = \mu + \tau_A + \pi_2$

The general form of the ANOVA table for the 2×2 crossover design when $\sigma_A^2 = \sigma_B^2$ and analyzed as a split-plot design for model (2.2) is given in Table 2.4.

Table 2.4: *ANOVA Table for Model 2.2: Two Treatment/Two Sequence Crossover Design (AB/BA) With Carryover Effects*

Source	df
Between Subject Analysis	
Sequence/Carryover	1
Subject(sequence)	$(n_1 + n_2)$-2
Within Subject Analysis	
Treatment	1
Period	1
Carryover	1
Error	$(n_1 + n_2)$-3
Total	$2(n_1 + n_2)$-1

Table 2.5 shows expected cell means for model (2.2).

Table 2.5: *Expected Cell Means for a Two Treatment/ Two Sequence Crossover Design (AB/BA) for Model (2.2)*

Sequence	Period	
	1	2
1	$\boldsymbol{\mu}_{11} = \mu + \tau_A + \pi_1$	$\boldsymbol{\mu}_{12} = \mu + \tau_B + \pi_1 + \lambda_A$
2	$\boldsymbol{\mu}_{21} = \mu + \tau_B + \pi_1$	$\boldsymbol{\mu}_{22} = \mu + \tau_A + \pi_2 + \lambda_B$

2.2 Testing the Equality of the Two Variances due to Treatments

A test for equality of variances due to treatments is given in this section. If the variances are not significantly different, then the traditional analysis assuming equal variances can be used in the analysis of crossover designs. But if the variances are shown to be significantly different, then crossover design experiments can be analyzed assuming unequal variances.

Following Goad (1994)[6], consider writing the crossover design model in a matrix form

$$\boldsymbol{y}_{il} = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{\epsilon}_{il},\ i = 1, 2, \cdots, s;\ l = 1, 2, \cdots, n_i$$

where $\boldsymbol{\beta}$ is the vector of fixed effect parameters; $\boldsymbol{\beta} = (\mu, s_1, \cdots, s_s, \tau_1, \cdots, \tau_t, \pi_1, \cdots, \pi_k)'$ for model (2.1) and $\boldsymbol{\beta} = (\mu, s_1, \cdots, s_s, \tau_1, \cdots, \tau_t, \pi_1, \cdots, \pi_k, \lambda_1, \cdots, \lambda_t)'$ for model (2.2). Assume that $\boldsymbol{\epsilon}_{il} \sim iid\ N(\mathbf{0}, \boldsymbol{\Sigma}),\ i = 1, 2, \cdots, s$ and $l = 1, 2, \cdots, n_i$.

Let

$$N = \sum_{i=1}^{s} n_i,\ \hat{\boldsymbol{\mu}}_i = \frac{\sum_{l=1}^{n_i} \boldsymbol{y}_{il}}{n_i},\ i = 1, 2, \cdots, s,\ \text{and}\ \hat{\boldsymbol{\Sigma}} = \frac{1}{N-s}\sum_{i=1}^{s}\sum_{l=1}^{n_i}(\boldsymbol{y}_{il} - \hat{\boldsymbol{\mu}}_i)(\boldsymbol{y}_{il} - \hat{\boldsymbol{\mu}}_i)'.$$

It can be shown that $\hat{\boldsymbol{\mu}}_1, \hat{\boldsymbol{\mu}}_2, \cdots, \hat{\boldsymbol{\mu}}_s$ and $\hat{\boldsymbol{\Sigma}}$ are sufficient statistics by the factorization theorem in both a model without carryover and a model with carryover. Furthermore, $\hat{\boldsymbol{\Sigma}}$ is an unbiased estimator of the variance-covariance matrix, $\boldsymbol{\Sigma}$. It can be shown that $(N-s)\hat{\boldsymbol{\Sigma}}$ follows a p-variate central Wishart distribution with $N-s$ degrees of freedom and variance-covariance matrix, $\boldsymbol{\Sigma}$; that is, $(N-s)\hat{\boldsymbol{\Sigma}} \sim W_p(N-s, \boldsymbol{\Sigma})$. The Wishart distribution was historically derived to find the joint distribution of sample variances and covariances of a multivariate normal sample. Fisher (1915) derived the probability density function (p.d.f.) for the two dimensional case in order to treat the distribution of a sample correlation coefficient. Wishart (1928)[31] generalized the derivation to the case when $p > 2$. Various multivariate analysis methods for the derivation of Wishart distributions have been discussed by Wishart and Bartlett (1933)[32], Mahalanobis et al. (1937)[18], Madow (1938)[17],

Hsu (1939)[10], Elfving (1947)[3], Sverdrup (1947)[29], Rasch (1948)[26], Ogawa (1953)[24], Olkin and Roy (1954)[23], James (1954)[15], and Jambunathan (1965)[14]. The Wishart distribution is a multivariate generalization of the chi-squared distribution.

Definition 1. *Let* $\boldsymbol{V}_{p\times p} = \boldsymbol{U}\boldsymbol{U}' = \sum_{i=1}^{n} \boldsymbol{u}_i \boldsymbol{u}_i'$, *where* $\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n]'$ *be a normal observation matrix of order* (n, p) *from* $N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$*; that is,* $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ *are independently and identically distributed (i.i.d.) as* $N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$. *Then the distribution of the elements in* $\boldsymbol{V}$ *is called a Wishart distribution with covariance matrix* $\boldsymbol{\Sigma}$ *and* n *degrees of freedom and is written as* $\boldsymbol{V} \sim W_p(n, \boldsymbol{\Sigma})$. *The distribution is nonsingular or singular according to whether* $\boldsymbol{\Sigma}$ *is positive definite (p.d.) or positive semi-definite (p.s.d.). If* $n \geq p$, *then the nonsingular distribution has a probability density function and if* $n < p$, *the distribution is called a pseudo Wishart distribution. When* $\boldsymbol{\Sigma} = \boldsymbol{I}_p$, *the distribution is said to be in standard form. If* $E(\boldsymbol{u}_i) \neq \boldsymbol{0}$ *for at least one* i, *then the distribution of the elements in* $\boldsymbol{V}$ *is called noncentral Wishart distribution. A random matrix having a Wishart distribution is called a Wishart matrix.*

As mentioned, there are several methods for deriving the *p.d.f.* of a Wishart distribution $W_p(n, \boldsymbol{\Sigma})$, $n \geq p$. Here a useful theorem by Hsu (1940)[11] is used.

Theorem 1. *Let* $\boldsymbol{U}$ *be a* $n \times p$ *observation matrix, where* $n \geq p$. *If the density of* $\boldsymbol{U}$ *has the form* $f(\boldsymbol{U}) = g(\boldsymbol{U}\boldsymbol{U}')$, *then the density of* $\boldsymbol{V} = \boldsymbol{U}\boldsymbol{U}'$ *is*

$$h(\boldsymbol{V}) = \frac{\pi^{np/2}}{\Gamma_p\left(\frac{n}{2}\right)} \mid \boldsymbol{V} \mid^{(n-p-1)/2} g(\boldsymbol{V})$$

where $\Gamma_p(t) = \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma\left(t - \frac{i-1}{2}\right)$ *is the generalized gamma function.*

The *p.d.f.* of a Wishart distribution is obtained as a special case, since if $\boldsymbol{U} \sim N_{n,p}(\boldsymbol{0}, \boldsymbol{\Sigma})$, then the *p.d.f.* of $\boldsymbol{U}$ has the form as

$$f(\boldsymbol{U}) = \mid 2\pi\boldsymbol{\Sigma} \mid^{-n/2} exp\left\{-\frac{1}{2} tr\left(\boldsymbol{V}\boldsymbol{\Sigma}^{-1}\right)\right\} = g(\boldsymbol{V}),\ say.$$

Theorem 2. *The p.d.f. of Wishart distribution $W_p(n, \boldsymbol{\Sigma})$, $n \geq p$, is given by*

$$h(\boldsymbol{V}) = \frac{1}{2^{np/2} \mid \boldsymbol{\Sigma} \mid^{n/2} \Gamma_p\left(\frac{n}{2}\right)} \mid \boldsymbol{V} \mid^{(n-p-1)/2} exp\left\{-\frac{1}{2} tr\left(\boldsymbol{V}\boldsymbol{\Sigma}^{-1}\right)\right\}$$

if $\boldsymbol{V} > 0$ is positive definite.

Mardia, Kent and Bibby (1979)[19] give the following Theorem.

Theorem 3. *If $\boldsymbol{V}_1 \sim W_p(m_1, \boldsymbol{\Sigma})$ and $\boldsymbol{V}_2 \sim W_p(m_2, \boldsymbol{\Sigma})$, and if $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ are independent, then $\boldsymbol{V}_1 + \boldsymbol{V}_2 \sim W_p(m_1 + m_2, \boldsymbol{\Sigma})$.*

Consider a two treatment/ two sequence crossover design with two periods (AB/BA) with the covariance structure defined by (2.5). Define the sums of squares for each of the two sequences by

$$\begin{aligned} \boldsymbol{W}_1 &= \sum_{l=1}^{n_1} (\boldsymbol{y}_{1l} - \hat{\boldsymbol{\mu}}_1)(\boldsymbol{y}_{1l} - \hat{\boldsymbol{\mu}}_1)' \\ &= \sum_{l=1}^{n_1} \begin{pmatrix} y_{111l} - \overline{y}_{111.} \\ y_{122l} - \overline{y}_{122.} \end{pmatrix} \begin{pmatrix} y_{111l} - \overline{y}_{111.} \\ y_{122l} - \overline{y}_{122.} \end{pmatrix}' = \begin{pmatrix} w_{111} & w_{112} \\ w_{112} & w_{122} \end{pmatrix} \end{aligned}$$

and

$$\begin{aligned} \boldsymbol{W}_2 &= \sum_{l=1}^{n_2} (\boldsymbol{y}_{2l} - \hat{\boldsymbol{\mu}}_2)(\boldsymbol{y}_{2l} - \hat{\boldsymbol{\mu}}_2)' \\ &= \sum_{l=1}^{n_2} \begin{pmatrix} y_{221l} - \overline{y}_{221.} \\ y_{212l} - \overline{y}_{212.} \end{pmatrix} \begin{pmatrix} y_{221l} - \overline{y}_{221.} \\ y_{212l} - \overline{y}_{212.} \end{pmatrix}' = \begin{pmatrix} w_{211} & w_{212} \\ w_{212} & w_{222} \end{pmatrix}. \end{aligned}$$

By Definition 1, $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ have independent Wishart distribution with the degrees of freedom $n_1 - 1$ and $n_2 - 1$, respectively. That is, $\boldsymbol{W}_1 \sim \boldsymbol{W}_2(n_1 - 1, \boldsymbol{\Sigma}_1)$ and $\boldsymbol{W}_2 \sim \boldsymbol{W}_2(n_2 - 1, \boldsymbol{\Sigma}_2)$.

Consider interchanging data between the two periods in the second sequence.
Let $\boldsymbol{W}_2^* = \begin{pmatrix} w_{222} & w_{212} \\ w_{212} & w_{211} \end{pmatrix}$ be a permutation of $\boldsymbol{W}_2$. Define $\boldsymbol{W}$ as follows

$$\boldsymbol{W} = \boldsymbol{W}_1 + \boldsymbol{W}_2^* = \begin{pmatrix} w_{11} & w_{12} \\ w_{12} & w_{22} \end{pmatrix}. \tag{2.6}$$

Then, by Theorem 3, $\boldsymbol{W}$ has a Wishart distribution with n_1+n_2-2 degrees of freedom and variance-covariance structure $\boldsymbol{\Sigma}$ where $\boldsymbol{\Sigma}=\begin{pmatrix}\sigma_A^2 & \rho\sigma_A\sigma_B\\ \rho\sigma_A\sigma_B & \sigma_B^2\end{pmatrix}=\boldsymbol{D}^{1/2}\boldsymbol{R}\boldsymbol{D}^{1/2}$ where $\boldsymbol{D}^{1/2}=diag(\sigma_A,\sigma_B)$ and $\boldsymbol{R}=\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}$.

The inverse of $\boldsymbol{\Sigma}$ is given by

$$\boldsymbol{\Sigma}^{-1}=\boldsymbol{D}^{-1/2}\boldsymbol{R}^{-1}\boldsymbol{D}^{-1/2}$$

where $\boldsymbol{D}^{-1/2}=diag\left(\frac{1}{\sigma_A},\frac{1}{\sigma_B}\right)$ and $\boldsymbol{R}^{-1}=\frac{1}{1-\rho^2}\begin{pmatrix}1 & -\rho\\ -\rho & 1\end{pmatrix}$.

One can get estimates of $\sigma_A^2,\ \sigma_B^2$ and ρ by the method of moments (MM).

Consider $E\left(\frac{1}{N-2}\boldsymbol{W}\right)=\boldsymbol{\Sigma}=\begin{pmatrix}\sigma_A^2 & \rho\sigma_A\sigma_B\\ \rho\sigma_A\sigma_B & \sigma_B^2\end{pmatrix}$. Then the method of moments estimates of $\sigma_A^2,\ \sigma_B^2$ and ρ are

$$\hat{\sigma}_A^2=\frac{w_{11}}{N-2},\ \hat{\sigma}_B^2=\frac{w_{22}}{N-2},\ \hat{\rho}=\frac{w_{12}}{\sqrt{w_{11}w_{22}}}$$

where $N=n_1+n_2$.

Now consider testing $H_0:\sigma_A^2=\sigma_B^2$ versus $H_A:\sigma_A^2\neq\sigma_B^2$.

A likelihood function based on $\boldsymbol{W}$ is given by

$$L(\boldsymbol{\Sigma})=c\frac{|\boldsymbol{W}|^{\frac{N-5}{2}}exp\left(-\frac{1}{2}tr\left(\boldsymbol{W}\boldsymbol{\Sigma}^{-1}\right)\right)}{|\boldsymbol{\Sigma}|^{\frac{N-2}{2}}}\text{ where }c=\frac{1}{2^{N-2}\pi^{1/2}\prod_{i=1}^{2}\Gamma\left(\frac{N-i-1}{2}\right)}.$$

The log-likelihood function is

$$logL(\boldsymbol{\Sigma})=log(c)+\frac{N-5}{2}log|\boldsymbol{W}|-\frac{N-2}{2}log|\boldsymbol{\Sigma}|-\frac{1}{2}tr\left(\boldsymbol{W}\boldsymbol{\Sigma}^{-1}\right).$$

Under $H_0:\sigma_A^2=\sigma_B^2=\sigma^2$,

$$|\boldsymbol{\Sigma}|=\left(1-\rho^2\right)\sigma^4\text{ and }tr\left(\boldsymbol{W}\boldsymbol{\Sigma}^{-1}\right)=\frac{w_{11}+w_{22}-2\rho w_{12}}{\left(1-\rho^2\right)\sigma^2}.$$

Therefore, the log-likelihood function is

$$\begin{aligned}logL(\boldsymbol{\Sigma})=\ & log(c)+\frac{N-5}{2}log|\boldsymbol{W}|-(N-2)log\left(\sigma^2\right)-\frac{N-2}{2}log\left(1-\rho^2\right)\\ & -\frac{w_{11}+w_{22}-2\rho w_{12}}{2\left(1-\rho^2\right)\sigma^2}.\end{aligned}$$

The derivatives of $logL(\boldsymbol{\Sigma})$ with respect to σ^2 and ρ are

$$\frac{\partial logL(\boldsymbol{\Sigma})}{\partial \sigma^2} = -\frac{N-2}{\sigma^2} + \frac{w_{11}+w_{22}-2\rho w_{12}}{2\left(1-\rho^2\right)\sigma^4}$$
$$\frac{\partial logL(\boldsymbol{\Sigma})}{\partial \rho} = \frac{(N-2)\rho}{1-\rho^2} + \frac{w_{12}\left(1-\rho^2\right)-\rho(w_{11}+w_{22}-2\rho w_{12})}{\left(1-\rho^2\right)^2\sigma^2}.$$

Setting the derivatives of $logL(\boldsymbol{\Sigma})$ with respect to σ^2 and ρ equal to zero, one gets

$$-2(N-2)\sigma^2 + \frac{w_{11}+w_{22}-2\rho w_{12}}{\left(1-\rho^2\right)} = 0$$
$$(N-2)\rho\left(1-\rho^2\right) + \frac{w_{12}\left(1-\rho^2\right)-\rho(w_{11}+w_{22}-2\rho w_{12})}{\sigma^2} = 0.$$

Then, under $H_0 : \sigma_A^2 = \sigma_B^2 = \sigma^2$, the maximum likelihood estimators of σ^2 and ρ are

$$\hat{\sigma}_R^2 = \frac{w_{11}+w_{22}-2\hat{\rho}_R w_{12}}{2(N-2)\left(1-\hat{\rho}_R^2\right)} \text{ and } \hat{\rho}_R = \frac{2w_{12}}{w_{11}+w_{22}}, \text{ respectively.}$$

Under $H_A : \sigma_A^2 \neq \sigma_B^2$,

$$|\boldsymbol{\Sigma}| = \left(1-\rho^2\right)\sigma_A^2\sigma_B^2 \text{ and } tr\left(\boldsymbol{W}\boldsymbol{\Sigma}^{-1}\right) = \frac{1}{1-\rho^2}\left[\frac{w_{11}}{\sigma_A^2} + \frac{w_{22}}{\sigma_B^2} - \frac{2\rho w_{12}}{\sigma_A\sigma_B}\right].$$

Therefore the log-likelihood function is

$$\begin{aligned} logL(\boldsymbol{\Sigma}) = {} & log(c) + \frac{N-5}{2}log|\boldsymbol{W}| - \frac{N-2}{2}log\left(\sigma_A^2\right) - \frac{N-2}{2}log\left(\sigma_B^2\right) \\ & -\frac{N-2}{2}log\left(1-\rho^2\right) - \frac{1}{2\left(1-\rho^2\right)}\left[\frac{w_{11}}{\sigma_A^2} + \frac{w_{22}}{\sigma_B^2} - \frac{2\rho w_{12}}{\sigma_A\sigma_B}\right]. \end{aligned}$$

The derivatives of $logL(\boldsymbol{\Sigma})$ with respect to σ_A^2, σ_B^2, and ρ are

$$\begin{aligned} \frac{\partial logL(\boldsymbol{\Sigma})}{\partial \sigma_A^2} &= -\frac{N-2}{2\sigma_A^2} + \frac{w_{11}}{2\left(1-\rho^2\right)\sigma_A^4} - \frac{\rho w_{12}}{2\left(1-\rho^2\right)\sigma_A^3\sigma_B}, \\ \frac{\partial logL(\boldsymbol{\Sigma})}{\partial \sigma_B^2} &= -\frac{N-2}{2\sigma_B^2} + \frac{w_{22}}{2\left(1-\rho^2\right)\sigma_B^4} - \frac{\rho w_{12}}{2\left(1-\rho^2\right)\sigma_A\sigma_B^3}, \\ \frac{\partial logL(\boldsymbol{\Sigma})}{\partial \rho} &= \frac{(N-2)\rho}{1-\rho^2} - \frac{\rho}{\left(1-\rho^2\right)^2}\left[\frac{w_{11}}{\sigma_A^2} + \frac{w_{22}}{\sigma_B^2} - \frac{\rho w_{12}}{\sigma_A\sigma_B}\right] + \frac{w_{12}}{\left(1-\rho^2\right)^2\sigma_A\sigma_B}. \end{aligned} \tag{2.7}$$

Setting the derivatives of $logL(\mathbf{\Sigma})$ with respect to $\sigma_A^2,\ \sigma_B^2$, and ρ equal to zero, one gets

$$\begin{aligned}
&-(N-2)\sigma_A^2 - \frac{\rho w_{12}}{(1-\rho^2)\,\sigma_B}\sigma_A + \frac{w_{11}}{(1-\rho^2)} = 0,\\
&-(N-2)\sigma_B^2 - \frac{\rho w_{12}}{(1-\rho^2)\,\sigma_A}\sigma_B + \frac{w_{22}}{(1-\rho^2)} = 0,\\
&(N-2)\rho\left(1-\rho^2\right) - \rho\left[\frac{w_{11}}{\sigma_A^2} + \frac{w_{22}}{\sigma_B^2} - \frac{\rho w_{12}}{\sigma_A\sigma_B}\right] + \frac{w_{12}}{\sigma_A\sigma_B} = 0.
\end{aligned} \tag{2.8}$$

Solving these equations for $\sigma_A^2,\ \sigma_B^2$, and ρ, one finds the maximum likelihood estimators of $\sigma_A^2,\ \sigma_B^2$, and ρ as

$$\hat{\sigma}_{A,UR}^2 = \frac{w_{11}}{N-2},\ \hat{\sigma}_{B,UR}^2 = \frac{w_{22}}{N-2},\ \text{and } \hat{\rho}_{UR} = \frac{w_{12}}{\sqrt{w_{11}w_{22}}},\ \text{respectively.}$$

Note that the method of moment estimators and the maximum likelihood estimators of $\sigma_A^2,\ \sigma_B^2$, and ρ are identical.

A likelihood ratio test statistic for testing the equality of the two variances is

$$\lambda = \frac{L\left(\hat{\mathbf{\Sigma}}_R\right)}{L\left(\hat{\mathbf{\Sigma}}_{UR}\right)} = \frac{c\frac{|\boldsymbol{W}|^{\frac{N-5}{2}}exp\left(-\frac{1}{2}tr\left(\boldsymbol{W}\hat{\mathbf{\Sigma}}_R^{-1}\right)\right)}{\left|\hat{\mathbf{\Sigma}}_R\right|^{\frac{N-2}{2}}}}{c\frac{|\boldsymbol{W}|^{\frac{N-5}{2}}exp\left(-\frac{1}{2}tr\left(\boldsymbol{W}\hat{\mathbf{\Sigma}}_{UR}^{-1}\right)\right)}{\left|\hat{\mathbf{\Sigma}}_{UR}\right|^{\frac{N-2}{2}}}} = \frac{\left|\hat{\mathbf{\Sigma}}_{UR}\right|^{\frac{N-2}{2}}exp\left(-\frac{1}{2}tr\left(\boldsymbol{W}\hat{\mathbf{\Sigma}}_R^{-1}\right)\right)}{\left|\hat{\mathbf{\Sigma}}_R\right|^{\frac{N-2}{2}}exp\left(-\frac{1}{2}tr\left(\boldsymbol{W}\hat{\mathbf{\Sigma}}_{UR}^{-1}\right)\right)}. \tag{2.9}$$

where

$$c = \frac{1}{2^{N-2}\pi^{1/2}\prod_{i=1}^{2}\Gamma\left(\frac{N-i-1}{2}\right)}$$

$$\hat{\mathbf{\Sigma}}_R = \hat{\sigma}_R^2\begin{pmatrix} 1 & \hat{\rho}_R \\ \hat{\rho}_R & 1 \end{pmatrix},\ \hat{\mathbf{\Sigma}}_{UR} = \begin{pmatrix} \hat{\sigma}_{A,UR}^2 & \hat{\rho}_R\hat{\sigma}_{A,UR}\hat{\sigma}_{B,UR} \\ \hat{\rho}_R\hat{\sigma}_{A,UR}\hat{\sigma}_{B,UR} & \hat{\sigma}_{B,UR}^2 \end{pmatrix}.$$

For large $N = n_1 + n_2$, the test for equal variances is rejected when $Q > \chi_{\alpha,f}^2$ where

$$Q = -2log(\lambda) = -2\left\{log\left(L\left(\hat{\mathbf{\Sigma}}_R\right)\right) - log\left(L\left(\hat{\mathbf{\Sigma}}_{UR}\right)\right)\right\} \tag{2.10}$$

and $f = 3 - 2 = 1$.

2.3 A Simulation Study

In Section 2.2, a likelihood ratio test for testing the equality of the two variances due to treatments was obtained for the two treatment/two period crossover design. The test was based on the distribution of $\boldsymbol{W}$ defined in (2.6). It should be noted that this test is valid whether carryover exists or not.

With a particular probability distribution for a data set, the purpose of determining a set of sufficient statistics is to find functions of the data containing all of the information about the parameters. Therefore, the sufficient statistics contain all of the information in the data necessary for estimation and inference purposes (Mood, Graybill, and Boes, 1974)[22]. When simulating crossover design experiments, one needs only to simulate the sufficient statistics.

For the 2×2 crossover design with unequal treatment variances, a set of sufficient statistics is given by $\hat{\boldsymbol{\mu}}_1 = \frac{1}{n_1}\sum_{n_1}^{l=1} \boldsymbol{y}_{1l}$, $\hat{\boldsymbol{\mu}}_2 = \frac{1}{n_2}\sum_{n_2}^{l=1} \boldsymbol{y}_{2l}$, and $\boldsymbol{W}$ defined in (2.6). The distribution of $\hat{\boldsymbol{\mu}}_1$ and $\hat{\boldsymbol{\mu}}_2$ depend on $\mu,\ \tau_A,\ \tau_B,\ \pi_1$, and π_2 as well as on $\sigma_A^2,\ \sigma_B^2$ and ρ while the distribution of the elements in $\boldsymbol{W}$ depends only on $\sigma_A^2,\ \sigma_B^2$ and ρ. Thus there is no information about $\mu,\ \tau_A,\ \tau_B,\ \pi_1$, and π_2 in $\boldsymbol{W}$ but there may be information about $\sigma_A^2,\ \sigma_B^2$ and ρ in $\hat{\boldsymbol{\mu}}_1$ and $\hat{\boldsymbol{\mu}}_2$. Whether or not there is information in $\hat{\boldsymbol{\mu}}_1$ and $\hat{\boldsymbol{\mu}}_2$ about $\sigma_A^2,\ \sigma_B^2$ and ρ depends on how $\mu,\ \tau_A,\ \tau_B,\ \pi_1$, and π_2 are related to one another.

Consider simulating data for Sequence 1. The variance-covariance matrix for sequence 1 is

$$\boldsymbol{\Sigma}_1 = \begin{pmatrix} \sigma_A^2 & \rho\sigma_A\sigma_B \\ \rho\sigma_A\sigma_B & \sigma_B^2 \end{pmatrix}.$$

The variance-covariance matrix for sequence 1 can be factored as $\boldsymbol{\Sigma}_1 = \boldsymbol{U}'\boldsymbol{U}$ where $\boldsymbol{U}$ is a unique 2×2 upper triangular matrix (Graybill, 1976). The matrix $\boldsymbol{U}$ is given by

$$\boldsymbol{U} = \begin{pmatrix} \sigma_A & \rho\sigma_B \\ 0 & \sigma_B\sqrt{1-\rho^2} \end{pmatrix}.$$

Let

$$\boldsymbol{x} = [x_1, x_2]' \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]$$

and let

$$\boldsymbol{y} = \boldsymbol{U}'\boldsymbol{x} = \begin{pmatrix} \sigma_A & 0 \\ \rho\sigma_B & \sigma_B\sqrt{1-\rho^2} \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Then

$$\boldsymbol{y} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_A^2 & \rho\sigma_A\sigma_B \\ \rho\sigma_A\sigma_B & \sigma_B^2 \end{pmatrix}\right].$$

The elements of $\boldsymbol{y}$ are $y_1 = \sigma_a x_1$ and $y_2 = \rho\sigma_B x_1 + \sigma_B\sqrt{1-\rho^2}x_2$. For the simulations performed, the x's were generated from a standard normal distribution and the above transformations were made to get the y's. Furthermore, appropriate cell parameters were added to the y's to get the expected cell means. For example, appropriate cell parameters for sequence 1 yield $y_1^* = y_1 + \mu + \tau_A + \pi_1$ for period 1 and $y_2^* = y_2 + \mu + \tau_B + \pi_2 + \lambda_A$ for period 2. Without loss of generality, $\mu,\ \pi_1,\ \pi_2,\ \tau_1,\ \tau_2$ and λ_A were all fixed at zero in the simulation study. Data for sequence 2 were similarly generated using the variance-covariance matrix

$$\boldsymbol{\Sigma}_2 = \begin{pmatrix} \sigma_B^2 & \rho\sigma_A\sigma_B \\ \rho\sigma_A\sigma_B & \sigma_A^2 \end{pmatrix}.$$

and appropriate cell mean parameters.

Table 2.6 shows the parameters that were used when simulating data for Type I error rate and power analysis for testing the equality of variances due to treatments.

Table 2.6: *Parameter values used in Type I error rate and Power Analysis for the equality of variances due to Treatment in a Two Treatment/ Two Sequence Crossover Design (AB/BA) for Model (2.1)*

	$\lambda_A = \lambda_B = 0,\ \sigma_A^2 = 1$	
ρ	Equal Variance	Unequal Variance
0	$\sigma_B^2 = 1$	$\sigma_B^2 = 2, 4, 8, 16$
0.1	$\sigma_B^2 = 1$	$\sigma_B^2 = 2, 4, 8, 16$
0.3	$\sigma_B^2 = 1$	$\sigma_B^2 = 2, 4, 8, 16$
0.5	$\sigma_B^2 = 1$	$\sigma_B^2 = 2, 4, 8, 16$
0.7	$\sigma_B^2 = 1$	$\sigma_B^2 = 2, 4, 8, 16$
0.9	$\sigma_B^2 = 1$	$\sigma_B^2 = 2, 4, 8, 16$

The number of subjects assigned to each sequence are $n = 3,\ 6,\ 12,\ 18,\ 24$ for each of the cases described in Table 2.6. The empirical Type I error rates are estimated for the

equal variances case, and the power is estimated for each of the unequal variance cases. To get the empirical Type I error rates and the power, 1000 simulations were done for each n, ρ and σ_B^2.

To estimate parameters, the proposed method described in Section 2.2 is used in R. Using the same data generated in R, restricted maximum likelihood ratio tests using SAS-MIXED were also obtained. The SAS steps used are shown below. The programming statements used in R are shown in Appendix E.

Step 1 Import data

```
INFILE 'C:\Data\KSU PhD THESIS\PRG\CS-2s2p2t-y6-18data.txt' DELIMITER=' ';
INPUT rho sA2 sB2 nsubj sim seq subj y1 y2;
```

Step 2 Define period and arrange treatments

```
DATA one; SET one;
 DROP y1 y2 ;
 per=1; y=y1; output;
 per=2; y=y2; output;
RUN;


DATA one; SET one;
    SUbJ=subj+(seq-1)*nsubj;
    trt='A';
    IF seq=1 and per=2 THEN trt='B';
    IF seq=2 and per=1 THEN trt='B';
RUN;
```

Step 3 Calculate $-2log\left(L\left(\hat{\boldsymbol{\Sigma}}_R\right)\right)$ in (2.10)

```
PROC MIXED ic data=one;
 TITLE 'ANALYSIS USING SAS-MIXED';
```

```
 by rho sim;
 CLASSES seq per trt subj;
 MODEL y=seq trt per/DDFM=SATTERTH;
 REPEATED trt/ SUBJECT=subj TYPE=CS;
 ods listing exclude all;
 ods output infocrit = null COVPARMS=HOPARMS;
RUN;

DATA null; set null;
rename neg2loglike =ho;
drop aic--caic;
```

Step 4 Calculate $-2log\left(L\left(\hat{\mathbf{\Sigma}}_{UR}\right)\right)$ in (2.10)

```
PROC MIXED ic data=one;
 TITLE 'ANALYSIS USING SAS-MIXED';
 by rho sim;
 CLASSES seq per trt subj;
 MODEL y=seq trt per/DDFM=SATTERTH;
 REPEATED trt/ SUBJECT=subj TYPE=CSH;
 ods listing exclude all;
 ods output infocrit = ha COVPARMS=HAPARMS;
RUN;

DATA ha; set ha;
rename neg2loglike =ha;
drop aic--caic;
```

Step 5 Calculate the Type I error rate

```
DATA comb; SET comb;
```

```
 u = ho-ha;
 IF u>3.841459 THEN reject=1; ELSE reject=0;
RUN;

PROC MEANS data=comb;
 by rho;
 ods listing select all;
 var reject;
RUN;
```

More detailed information about using SAS-MIXED can be found in Appendix F. Consider the likelihood ratio test of two treatment crossover design when treatments have unequal variances. Figure A.1 shows the Type I error rates for the likelihood ratio tests of $H_0 : \sigma_A^2 = \sigma_B^2$ as n increases, and Table B.1 displays the computed Type I error rates of the likelihood ratio tests for $H_0 : \sigma_A^2 = \sigma_B^2$. Figure A.1 has six lines. Each line corresponds to a different value of ρ, $\rho = 0.0, 0.1, 0.3, 0.5, 0.7, 0.9$. The blue line shows the results when $\rho = 0$. The green line corresponds to $\rho = 0.1$. The red, light blue, purple, and yellow lines correspond to data satisfying $\rho = 0.3, 0.5, 0.7, 0.9$, respectively. Table B.1 shows the observed Type I error rates of the likelihood ratio tests for $H_0 : \sigma_A^2 = \sigma_B^2$ obtained by both R and SAS. Each row for each method in Table B.1 corresponds to a different value of ρ, $\rho = 0.0, 0.1, 0.3, 0.5, 0.7, 0.9$. Each column corresponds to a different number of subjects assigned to each sequence. Note that the observed Type I error rates converge to 0.05 as the number of subjects in each sequence becomes large. The likelihood ratio tests hold their size when the number of subjects is at least 18 per sequence for each value of ρ. It should be noted that the chi-square approximation to a likelihood ratio test is an asymptotic approximation. Here it appears that this approximation is only valid when $n \geq 18$. This corresponds to the Wishart distribution under the null hypothesis having at least 34 degrees of freedom.

Next consider data that were generated when $\sigma_A^2 \neq \sigma_B^2$. Figure A.2 shows the power of the LRT of $\sigma_A^2 = \sigma_B^2$ when $(\sigma_A^2, \sigma_B^2) = (1, 2)$ for each value of ρ. As the number of subjects is increased from 6 to 24, the power increases sharply towards 1 when $\rho = 0.9$. And, as the value of the correlation, ρ, increases from 0.1 to 0.9, the power also increases. Table B.2 shows the observed power for different variances, $(\sigma_A^2, \sigma_B^2) = (1, 2)$. The performance of the two tests for the equality of variances given by R and SAS are identical. Figures A.3-A.5 show the trend for other values of σ_B^2. When the value of σ_B^2 increases to 4, 8, and 16, the power is very close to 1 for most configurations of the other parameters. And, as the value of correlation, ρ, is increased from 0.1 to 0.9, the power is close to 1. Tables B.3-B.5 display the observed power for as the σ_B^2 increases. The performance to test the equality of two variances by R and SAS are identical. Note that the power value for small numbers of subjects per sequence may be misleading since the desired Type I error rates are not achieved when $n \leq 12$.

2.4 Two Treatment/ Two sequence Crossover Design Balanced for Carryover Effect

When carryover effects exist, they are aliased with period and treatment effects in the 2×2 crossover design. Carryover effects are also aliased with sequence effects. Two treatment crossover designs that are balanced for carryover effects can be used to separate treatment, period and sequence effects from carryover effects. These two treatment crossover designs that can be used to estimate treatment differences free from carryover effects include extra sequence(s) and/or period(s) as described in Section 1.2.

2.4.1 Two Treatment/ Two Sequence Dual Balanced Design

Consider the ABA/BAB dual balanced design. The ABA/BAB crossover design is given in Table 2.7.

Table 2.7: *A Two Treatment/ Two Sequence Crossover Design with an Extra Period (ABA/BAB)*

	Period		
Sequence	1	2	3
1	A	B	A
2	B	A	B

The expected cell means for the ABA/BAB crossover design in Table 2.7 are given in Table 2.8.

Table 2.8: *Expected Cell Means for a Two Treatment/ Two Sequence Crossover Design with an Extra Period (ABA/BAB) for Model (2.2)*

Sequence	Period		
	1	2	3
1	$\boldsymbol{\mu}_{11} = \mu + s_1 + \tau_A + \pi_1$	$\boldsymbol{\mu}_{12} = \mu + s_1 + \tau_B + \pi_2 + \lambda_A$	$\boldsymbol{\mu}_{13} = \mu + s_1 + \tau_A + \pi_3 + \lambda_B$
2	$\boldsymbol{\mu}_{21} = \mu + s_2 + \tau_B + \pi_1$	$\boldsymbol{\mu}_{22} = \mu + s_2 + \tau_A + \pi_2 + \lambda_B$	$\boldsymbol{\mu}_{23} = \mu + s_2 + \tau_B + \pi_3 + \lambda_A$

Define the sums of squares for the two sequences by

$$W_1 = \sum_{l=1}^{n_1} (\boldsymbol{y}_{1l} - \hat{\boldsymbol{\mu}}_1)(\boldsymbol{y}_{1l} - \hat{\boldsymbol{\mu}}_1)^T$$
$$= \sum_{l=1}^{n_1} \begin{pmatrix} y_{111l} - \overline{y}_{111.} \\ y_{122l} - \overline{y}_{122.} \\ y_{113l} - \overline{y}_{113.} \end{pmatrix} \begin{pmatrix} y_{111l} - \overline{y}_{111.} \\ y_{122l} - \overline{y}_{122.} \\ y_{113l} - \overline{y}_{113.} \end{pmatrix}^T = \begin{pmatrix} w_{111} & w_{112} & w_{113} \\ w_{112} & w_{122} & w_{123} \\ w_{113} & w_{123} & w_{133} \end{pmatrix}$$

and

$$W_2 = \sum_{l=1}^{n_2} (\boldsymbol{y}_{2l} - \hat{\boldsymbol{\mu}}_2)(\boldsymbol{y}_{2l} - \hat{\boldsymbol{\mu}}_2)^T$$
$$= \sum_{l=1}^{n_2} \begin{pmatrix} y_{221l} - \overline{y}_{221.} \\ y_{212l} - \overline{y}_{212.} \\ y_{223l} - \overline{y}_{223.} \end{pmatrix} \begin{pmatrix} y_{221l} - \overline{y}_{221.} \\ y_{212l} - \overline{y}_{212.} \\ y_{223l} - \overline{y}_{223.} \end{pmatrix}^T = \begin{pmatrix} w_{211} & w_{212} & w_{213} \\ w_{212} & w_{222} & w_{223} \\ w_{213} & w_{223} & w_{233} \end{pmatrix}.$$

Then $W_i,\ i = 1, 2$ has a Wishart distribution with $n_i - 1$ degrees of freedom and variance-covariance structure $\Sigma_i,\ i = 1, 2$. Assuming compound symmetry for time period correlations together with unequal variances for the treatments, the variance-covariance matrices for the two sequences are given as

$$\Sigma_1 = \begin{pmatrix} \sigma_A^2 & \rho\sigma_A\sigma_B & \rho\sigma_A^2 \\ \rho\sigma_A\sigma_B & \sigma_B^2 & \rho\sigma_A\sigma_B \\ \rho\sigma_A^2 & \rho\sigma_A\sigma_B & \sigma_A^2 \end{pmatrix} = D_1^{1/2} R D_1^{1/2}$$

and

$$\Sigma_2 = \begin{pmatrix} \sigma_B^2 & \rho\sigma_A\sigma_B & \rho\sigma_B^2 \\ \rho\sigma_A\sigma_B & \sigma_A^2 & \rho\sigma_A\sigma_B \\ \rho\sigma_B^2 & \rho\sigma_A\sigma_B & \sigma_B^2 \end{pmatrix} = D_2^{1/2} R D_2^{1/2}.$$

where

$$D_1^{1/2} = diag(\sigma_A, \sigma_B, \sigma_A),\ D_2^{1/2} = diag(\sigma_B, \sigma_A, \sigma_B),\ \text{and } R = \begin{bmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{bmatrix}.$$

Note that

$$D_1^{-1/2} = diag\left(\frac{1}{\sigma_A}, \frac{1}{\sigma_B}, \frac{1}{\sigma_A}\right),\ D_2^{-1/2} = diag\left(\frac{1}{\sigma_B}, \frac{1}{\sigma_A}, \frac{1}{\sigma_B}\right),$$

and

$$R^{-1} = \frac{1}{(1-\rho)(1+2\rho)} \begin{bmatrix} 1+\rho & -\rho & -\rho \\ -\rho & 1+\rho & -\rho \\ -\rho & -\rho & 1+\rho \end{bmatrix}.$$

Then, the inverse and the determinant of $\Sigma_i,\ i=1,2$ are given by

$$\Sigma_1^{-1} = D_1^{-1/2} R^{-1} D_1^{-1/2} = \frac{1}{(1-\rho)(1+2\rho)} \begin{bmatrix} \frac{1+\rho}{\sigma_A^2} & -\frac{\rho}{\sigma_A\sigma_B} & -\frac{\rho}{\sigma_A^2} \\ -\frac{\rho}{\sigma_A\sigma_B} & \frac{1+\rho}{\sigma_B^2} & -\frac{\rho}{\sigma_A\sigma_B} \\ -\frac{\rho}{\sigma_A^2} & -\frac{\rho}{\sigma_A\sigma_B} & \frac{1+\rho}{\sigma_A^2} \end{bmatrix},$$

$$\Sigma_2^{-1} = D_2^{-1/2} R^{-1} D_2^{-1/2} = \frac{1}{(1-\rho)(1+2\rho)} \begin{bmatrix} \frac{1+\rho}{\sigma_B^2} & -\frac{\rho}{\sigma_A\sigma_B} & -\frac{\rho}{\sigma_B^2} \\ -\frac{\rho}{\sigma_A\sigma_B} & \frac{1+\rho}{\sigma_A^2} & -\frac{\rho}{\sigma_A\sigma_B} \\ -\frac{\rho}{\sigma_B^2} & -\frac{\rho}{\sigma_A\sigma_B} & \frac{1+\rho}{\sigma_B^2} \end{bmatrix},$$

$$|\Sigma_1| = \left(\sigma_A^2\right)^2 \sigma_B^2 (1-\rho)^2(1+2\rho),\ |\Sigma_2| = \sigma_A^2 \left(\sigma_B^2\right)^2 (1-\rho)^2(1+2\rho).$$

Now consider testing $H_0 : \sigma_A^2 = \sigma_B^2$ versus $H_A : \sigma_A^2 \neq \sigma_B^2$.

A likelihood function based on $W_i,\ i=1,2$ is given by

$$L = \prod_{i=1}^{2} c_i \frac{|W_i|^{\frac{n_i-5}{2}} exp\left(-\frac{1}{2} tr\left(W_i \Sigma_i^{-1}\right)\right)}{|\Sigma_i|^{\frac{n_i-1}{2}}} \text{ where } c_i = \frac{1}{2^{n_i-1}\pi^3 \prod_{l=1}^{3} \Gamma\left(\frac{n_i-l}{2}\right)}$$

where

$$W_1\Sigma_1^{-1} = \frac{1}{(1-\rho)(1+2\rho)} \begin{bmatrix} w_{111} & w_{112} & w_{113} \\ w_{112} & w_{122} & w_{123} \\ w_{113} & w_{123} & w_{133} \end{bmatrix} \begin{bmatrix} \frac{1+\rho}{\sigma_A^2} & -\frac{\rho}{\sigma_A\sigma_B} & -\frac{\rho}{\sigma_A^2} \\ -\frac{\rho}{\sigma_A\sigma_B} & \frac{1+\rho}{\sigma_B^2} & -\frac{\rho}{\sigma_A\sigma_B} \\ -\frac{\rho}{\sigma_A^2} & -\frac{\rho}{\sigma_A\sigma_B} & \frac{1+\rho}{\sigma_A^2} \end{bmatrix}$$

$$W_2\Sigma_2^{-1} = \frac{1}{(1-\rho)(1+2\rho)} \begin{bmatrix} w_{211} & w_{212} & w_{213} \\ w_{212} & w_{222} & w_{223} \\ w_{213} & w_{223} & w_{233} \end{bmatrix} \begin{bmatrix} \frac{1+\rho}{\sigma_B^2} & -\frac{\rho}{\sigma_A\sigma_B} & -\frac{\rho}{\sigma_B^2} \\ -\frac{\rho}{\sigma_A\sigma_B} & \frac{1+\rho}{\sigma_A^2} & -\frac{\rho}{\sigma_A\sigma_B} \\ -\frac{\rho}{\sigma_B^2} & -\frac{\rho}{\sigma_A\sigma_B} & \frac{1+\rho}{\sigma_B^2} \end{bmatrix}$$

$$\Rightarrow tr\left(W_1\Sigma_1^{-1}\right) = \frac{1}{(1-\rho)(1+2\rho)} \left[\frac{(1+\rho)(w_{111}+w_{133}) - 2\rho w_{113}}{\sigma_A^2} - \frac{2\rho w_{112}}{\sigma_A\sigma_B} + \frac{(1+\rho)w_{122}}{\sigma_B^2} - \frac{2\rho w_{123}}{\sigma_A\sigma_B}\right]$$

$$tr\left(W_2\Sigma_2^{-1}\right) = \frac{1}{(1-\rho)(1+2\rho)} \left[\frac{(1+\rho)(w_{211}+w_{233}) - 2\rho w_{213}}{\sigma_B^2} - \frac{2\rho w_{212}}{\sigma_A\sigma_B} + \frac{(1+\rho)w_{222}}{\sigma_A^2} - \frac{2\rho w_{223}}{\sigma_A\sigma_B}\right]$$

$$\Rightarrow \sum_{i=1}^{2} tr\left(W_i\Sigma_i^{-1}\right) = \frac{1}{(1-\rho)(1+2\rho)} \left[\frac{(1+\rho)(w_{111}+w_{133}+w_{222}) - 2\rho w_{113}}{\sigma_A^2} + \frac{(1+\rho)(w_{122}+w_{211}+w_{233}) - 2\rho w_{213}}{\sigma_B^2} - \frac{2\rho(w_{112}+w_{123}+w_{212}+w_{223})}{\sigma_A\sigma_B}\right].$$

The log-likelihood function is

$$logL = \sum_{i=1}^{2} \left[log(c_i) + \frac{n_i - 5}{2} log|W_i| - \frac{n_i - 1}{2} log|\Sigma_i| - \frac{1}{2} tr\left(W_i \Sigma_i^{-1}\right) \right].$$

Under $H_0 : \sigma_A^2 = \sigma_B^2 = \sigma^2$,

$$\begin{aligned}
|\Sigma_1| = |\Sigma_2| &= (1-\rho)^2(1+2\rho)\sigma^6, \\
tr\left(W_1\Sigma_1^{-1}\right) &= \frac{1}{(1-\rho)(1+2\rho)\sigma^2} \left\{(1+\rho)(w_{111}+w_{122}+w_{133}) - 2\rho(w_{112}+w_{113}+w_{123})\right\}, \\
tr\left(W_2\Sigma_2^{-1}\right) &= \frac{1}{(1-\rho)(1+2\rho)\sigma^2} \left\{(1+\rho)(w_{211}+w_{222}+w_{233}) - 2\rho(w_{212}+w_{213}+w_{223})\right\}.
\end{aligned}$$

Therefore, the log-likelihood function is

$$\begin{aligned}
logL = &\sum_{i=1}^{2} log(c_i) + \sum_{i=1}^{2} \frac{n_i - 5}{2} log|W_i| - \frac{N-2}{2} log\left[(1-\rho)^2(1+2\rho)\right] \\
&- \frac{3(N-2)}{2} log\left(\sigma^2\right) - \frac{1}{2(1-\rho)(1+2\rho)\sigma^2} \left[(1+\rho)A - 2\rho B\right].
\end{aligned}$$

where

$$\begin{aligned}
A &= w_{111} + w_{122} + w_{133} + w_{211} + w_{222} + w_{233} \\
B &= w_{112} + w_{113} + w_{123} + w_{212} + w_{213} + w_{223}, \text{ and } N = \sum_{i=1}^{2} n_i.
\end{aligned}$$

The derivatives of $logL$ with respect to σ^2 and ρ are

$$\begin{aligned}
\frac{\partial logL}{\partial \sigma^2} &= -\frac{3(N-2)}{2\sigma^2} + \frac{1}{2(1-\rho)(1+2\rho)\sigma^4} \left[(1+\rho)A - 2\rho B\right], \\
\frac{\partial logL}{\partial \rho} &= \frac{3(N-2)\rho(1-\rho)}{(1-\rho)^2(1+2\rho)} - \frac{2\left[\rho(2+\rho)A - (1+2\rho^2)B\right]}{2(1-\rho)^2(1+2\rho)^2\sigma^2}.
\end{aligned}$$

Setting the derivatives of $logL$ with respect to σ^2 and ρ equal to zero, one gets

$$\begin{aligned}
-3(N-2)\sigma^2 + \frac{(1+\rho)A - 2\rho B}{(1-\rho)(1+2\rho)} &= 0, \\
3(N-2)\rho(1-\rho)(1+2\rho) - \frac{\rho(2+\rho)A - (1+2\rho^2)B}{\sigma^2} &= 0.
\end{aligned}$$

Then, under $H_0 : \sigma_A^2 = \sigma_B^2 = \sigma^2$, the maximum likelihood estimator of σ^2 and ρ is

$$\hat{\sigma}_R^2 = \frac{A}{3(N-2)} \text{ and } \hat{\rho}_R = \frac{B}{A}, \text{ respectively.}$$

where

$$A = w_{111} + w_{122} + w_{133} + w_{211} + w_{222} + w_{233}$$

$$B = w_{112} + w_{113} + w_{123} + w_{212} + w_{213} + w_{223}.$$

Under $H_A : \sigma_A^2 \neq \sigma_B^2$,

$$\begin{aligned}
&|\Sigma_1| = \left(\sigma_A^2\right)^2 \sigma_B^2 (1-\rho)^2(1+2\rho),\ |\Sigma_2| = \sigma_A^2 \left(\sigma_B^2\right)^2 (1-\rho)^2(1+2\rho),\\
&tr\left(W_1\Sigma_1^{-1}\right) = \frac{1}{(1-\rho)(1+2\rho)}\left[\frac{(1+\rho)(w_{111}+w_{133}) - 2\rho w_{113}}{\sigma_A^2} - \frac{2\rho w_{112}}{\sigma_A\sigma_B}\right.\\
&\qquad\left. + \frac{(1+\rho)w_{122}}{\sigma_B^2} - \frac{2\rho w_{123}}{\sigma_A\sigma_B}\right],\\
&tr\left(W_2\Sigma_2^{-1}\right) = \frac{1}{(1-\rho)(1+2\rho)}\left[\frac{(1+\rho)(w_{211}+w_{233}) - 2\rho w_{213}}{\sigma_B^2} - \frac{2\rho w_{212}}{\sigma_A\sigma_B}\right.\\
&\qquad\left. + \frac{(1+\rho)w_{222}}{\sigma_A^2} - \frac{2\rho w_{223}}{\sigma_A\sigma_B}\right].
\end{aligned}$$

Therefore the log-likelihood function is

$$\begin{aligned}
logL = &\sum_{i=1}^{2} log(c_i) + \sum_{i=1}^{2} \frac{n_i - 5}{2} log|W_i| - \frac{2n_1 + n_2 - 3}{2} log\left(\sigma_A^2\right)\\
&- \frac{n_1 + 2n_2 - 3}{2} log\left(\sigma_B^2\right) - \frac{N-2}{2} log\left[(1-\rho)^2(1+2\rho)\right]\\
&- \frac{1}{2(1-\rho)(1+2\rho)}\left[\frac{(1+\rho)(w_{111}+w_{133}+w_{222}) - 2\rho w_{113}}{\sigma_A^2}\right.\\
&\left. + \frac{(1+\rho)(w_{122}+w_{211}+w_{233}) - 2\rho w_{213}}{\sigma_B^2} - \frac{2\rho(w_{112}+w_{123}+w_{212}+w_{223})}{\sigma_A\sigma_B}\right].
\end{aligned}$$

The derivatives of $logL$ with respect to $\sigma_A^2,\ \sigma_B^2$, and ρ are

$$\begin{aligned}
\frac{\partial logL}{\partial \sigma_A^2} = &-\frac{2n_1 + n_2 - 3}{2\sigma_A^2} + \frac{(1+\rho)(w_{111}+w_{133}+w_{222}) - 2\rho w_{113}}{2(1-\rho)(1+2\rho)\left(\sigma_A^2\right)^2}\\
&- \frac{2\rho(w_{112}+w_{123}+w_{212}+w_{223})}{4(1-\rho)(1+2\rho)\left(\sigma_A\right)^3\sigma_B},\\
\frac{\partial logL}{\partial \sigma_B^2} = &-\frac{n_1 + 2n_2 - 3}{2\sigma_B^2} + \frac{(1+\rho)(w_{122}+w_{211}+w_{233}) - 2\rho w_{213}}{2(1-\rho)(1+2\rho)\left(\sigma_B^2\right)^2}\\
&- \frac{2\rho(w_{112}+w_{123}+w_{212}+w_{223})}{4(1-\rho)(1+2\rho)\sigma_A\left(\sigma_B\right)^3},
\end{aligned} \tag{2.11}$$

$$\frac{\partial logL}{\partial \rho} = \frac{3(N-2)\rho(1-\rho)}{(1-\rho)^2(1+2\rho)} - \frac{1}{2(1-\rho)^2(1+2\rho)^2} \times$$
$$\left\{ \frac{2\rho(2+\rho)(w_{111}+w_{133}+w_{222})}{\sigma_A^2} + \frac{2\rho(2+\rho)(w_{122}+w_{211}+w_{233})}{\sigma_B^2} \right.$$
$$\left. -2(1+2\rho^2)\left[\frac{w_{113}}{\sigma_A^2} + \frac{w_{213}}{\sigma_B^2} + \frac{(w_{112}+w_{123}+w_{212}+w_{223})}{\sigma_A\sigma_B} \right] \right\}.$$

Setting the derivatives of $logL$ with respect to σ_A^2, σ_B^2, and ρ equal to zero, one gets

$$(2n_1+n_2-3)\sigma_A^2 + \frac{\rho(w_{112}+w_{123}+w_{212}+w_{223})}{(1-\rho)(1+2\rho)\sigma_B}\sigma_A - \frac{(1+\rho)(w_{111}+w_{133}+w_{222})-2\rho w_{113}}{(1-\rho)(1+2\rho)} = 0,$$
$$(n_1+2n_2-3)\sigma_B^2 + \frac{\rho(w_{112}+w_{123}+w_{212}+w_{223})}{(1-\rho)(1+2\rho)\sigma_A}\sigma_B - \frac{(1+\rho)(w_{122}+w_{211}+w_{233})-2\rho w_{213}}{(1-\rho)(1+2\rho)} = 0,$$
$$6(N-2)\rho(1-\rho)(1+2\rho) - \left\{ \frac{2\rho(2+\rho)(w_{111}+w_{133}+w_{222})}{\sigma_A^2} + \frac{2\rho(2+\rho)(w_{122}+w_{211}+w_{233})}{\sigma_B^2} \right.$$
$$\left. -2(1+2\rho^2)\left[\frac{w_{113}}{\sigma_A^2} + \frac{w_{213}}{\sigma_B^2} + \frac{(w_{112}+w_{123}+w_{212}+w_{223})}{\sigma_A\sigma_B} \right] \right\}. \quad (2.12)$$

Unfortunately, the maximum likelihood estimators of σ_A^2, σ_B^2 and ρ do not have a closed form solution. An iterative process to find solutions for $\hat{\sigma}_{A,UR}^2$, $\hat{\sigma}_{B,UR}^2$ and $\hat{\rho}_{UR}$ is discussed in Section 2.4.2. Shanga (2003)[28] incorrectly claimed that the maximum likelihood estimators of σ_A^2, σ_B^2 and ρ were

$$\hat{\sigma}_{A,UR}^2 = \frac{w_{11}}{2n_1+n_2-3},\ \hat{\sigma}_{B,UR}^2 = \frac{w_{22}}{n_1+2n_2-3},\ \hat{\rho}_{UR} = \frac{1}{3}\left[\frac{w_{12}}{\sqrt{w_{11}w_{22}}} + \frac{w_{11}^*}{w_{11}} + \frac{w_{22}^*}{w_{22}} \right]. \quad (2.13)$$

where $w_{11} = w_{111}+w_{133}+w_{222}$, $w_{22} = w_{122}+w_{211}+w_{233}$, $w_{12} = w_{112}+w_{123}+w_{212}+w_{223}$, $w_{11}^* = w_{113}$, and $w_{22}^* = w_{213}$.

The likelihood ratio test statistic for testing the equality of the two variances is

$$\lambda = \frac{L_R}{L_{UR}} = \frac{\prod_{i=1}^{2} c_i \frac{|W_i|^{\frac{n_i-5}{2}} exp\left(-\frac{1}{2}tr\left(W_i\hat{\Sigma}_{i,R}^{-1}\right)\right)}{|\hat{\Sigma}_{i,R}|^{\frac{n_i-1}{2}}}}{\prod_{i=1}^{2} c_i \frac{|W_i|^{\frac{n_i-5}{2}} exp\left(-\frac{1}{2}tr\left(W_i\hat{\Sigma}_{i,UR}^{-1}\right)\right)}{|\hat{\Sigma}_{i,UR}|^{\frac{n_i-1}{2}}}}$$
$$= \frac{\prod_{i=1}^{2}|\hat{\Sigma}_{i,UR}|^{\frac{n_i-1}{2}} \prod_{i=1}^{2} exp\left(-\frac{1}{2}tr\left(W_i\hat{\Sigma}_{i,R}^{-1}\right)\right)}{\prod_{i=1}^{2}|\hat{\Sigma}_{i,R}|^{\frac{n_i-1}{2}} \prod_{i=1}^{2} exp\left(-\frac{1}{2}tr\left(W_i\hat{\Sigma}_{i,UR}^{-1}\right)\right)}.$$

where

$$c_i = \frac{1}{2^{n_i-1}\pi^3 \prod_{l=1}^{3} \Gamma\left(\frac{n_i - l}{2}\right)}$$

$$\hat{\Sigma}_R = \hat{\sigma}_R^2 \begin{pmatrix} 1 & \hat{\rho}_R \\ \hat{\rho}_R & 1 \end{pmatrix},$$

$$\hat{\Sigma}_{UR} = \begin{pmatrix} \hat{\sigma}_{A,UR}^2 & \hat{\rho}_R \hat{\sigma}_{A,UR} \hat{\sigma}_{B,UR} \\ \hat{\rho}_R \hat{\sigma}_{A,UR} \hat{\sigma}_{B,UR} & \hat{\sigma}_{B,UR}^2 \end{pmatrix}.$$

For large $N = n_1 + n_2$, the test for equal variances is rejected when $Q > \chi^2_{\alpha,f}$ where $Q = -2log(\lambda) = -2\{log(L_R) - log(L_{UR})\}$ and $f = 3-2 = 1$. It can be noted that Shanga's estimates are method of moment estimates.

2.4.2 Methods for Estimating Parameters

Since, under the unrestricted condition in the (ABA/BAB) design, solutions of $\sigma_A^2,\ \sigma_B^2$ and ρ for this design do not have a closed form, one needs to consider other methods to find the maximum likelihood estimators under $H_A:\ not\ H_0$.

Consider the derivatives of the log-likelihood functions from (2.11):

$$\begin{aligned}
\frac{\partial logL}{\partial \sigma_A^2} &= -\frac{2n_1+n_2-3}{2\sigma_A^2} + \frac{(1+\rho)(w_{111}+w_{133}+w_{222})-2\rho w_{113}}{2(1-\rho)(1+2\rho)\left(\sigma_A^2\right)^2} \\
&\quad -\frac{2\rho(w_{112}+w_{123}+w_{212}+w_{223})}{4(1-\rho)(1+2\rho)\left(\sigma_A\right)^3\sigma_B}, \\
\frac{\partial logL}{\partial \sigma_B^2} &= -\frac{n_1+2n_2-3}{2\sigma_B^2} + \frac{(1+\rho)(w_{122}+w_{211}+w_{233})-2\rho w_{213}}{2(1-\rho)(1+2\rho)\left(\sigma_B^2\right)^2} \\
&\quad -\frac{2\rho(w_{112}+w_{123}+w_{212}+w_{223})}{4(1-\rho)(1+2\rho)\sigma_A\left(\sigma_B\right)^3}, \\
\frac{\partial logL}{\partial \rho} &= \frac{3(N-2)\rho(1-\rho)}{(1-\rho)^2(1+2\rho)} - \frac{1}{2(1-\rho)^2(1+2\rho)^2}\times \\
&\quad \left\{\frac{2\rho(2+\rho)(w_{111}+w_{133}+w_{222})}{\sigma_A^2} + \frac{2\rho(2+\rho)(w_{122}+w_{211}+w_{233})}{\sigma_B^2}\right. \\
&\quad \left. -2(1+2\rho^2)\left[\frac{w_{113}}{\sigma_A^2}+\frac{w_{213}}{\sigma_B^2}+\frac{(w_{112}+w_{123}+w_{212}+w_{223})}{\sigma_A\sigma_B}\right]\right\}.
\end{aligned} \tag{2.14}$$

Setting the derivatives of $logL$ with respect to $\sigma_A^2,\ \sigma_B^2$ and ρ equal to zero and one can get:

$$\begin{aligned}
&(2n_1+n_2-3)\sigma_A^2 + \frac{\rho(w_{112}+w_{123}+w_{212}+w_{223})}{(1-\rho)(1+2\rho)\sigma_B}\sigma_A - \frac{(1+\rho)(w_{111}+w_{133}+w_{222})-2\rho w_{113}}{(1-\rho)(1+2\rho)} = 0, \\
&(n_1+2n_2-3)\sigma_B^2 + \frac{\rho(w_{112}+w_{123}+w_{212}+w_{223})}{(1-\rho)(1+2\rho)\sigma_A}\sigma_B - \frac{(1+\rho)(w_{122}+w_{211}+w_{233})-2\rho w_{213}}{(1-\rho)(1+2\rho)} = 0, \\
&6(N-2)\rho(1-\rho)(1+2\rho) - \left\{\frac{2\rho(2+\rho)(w_{111}+w_{133}+w_{222})}{\sigma_A^2} + \frac{2\rho(2+\rho)(w_{122}+w_{211}+w_{233})}{\sigma_B^2}\right. \\
&\qquad \left. -2(1+2\rho^2)\left[\frac{w_{113}}{\sigma_A^2}+\frac{w_{213}}{\sigma_B^2}+\frac{(w_{112}+w_{123}+w_{212}+w_{223})}{\sigma_A\sigma_B}\right]\right\}.
\end{aligned} \tag{2.15}$$

Note that the first two equations in (2.15) are quadratic functions of σ_A and σ_B, respectively, and the third equation is a cubic function of ρ. The following definitions can be applied to estimate the parameters, $\sigma_A,\ \sigma_B$ and ρ.

Definition 2. *(Quadratic Equations) Solutions of any quadratic equation,* $ax^2+bx+c=0$, *are given by*

$$x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}.$$

If $a,\ b,$ *and* c *are real, the solutions are as follows:*

- *If* b^2-4ac *is positive, the roots are real and unequal.*
- *If* b^2-4ac *is zero, the roots are real and equal.*
- *If* b^2-4ac *is negative, the roots are imaginary and unequal.*

Definition 3. *(Cubic Equations) When one has a cubic equation,* $y^3+py^2+qy+r=0$, *it can be transformed to* $x^3+ax+b=0$ *by substituting* $x-\frac{p}{3}$ *for* y, *where* $a=\frac{1}{2}\left(3q-p^2\right)$ *and* $b=\frac{1}{27}\left(2p^3-9pq+27r\right)$. *Let* $A=\sqrt[3]{-\frac{b}{2}+\sqrt{\frac{b^2}{4}+\frac{a^3}{27}}}$ *and* $B=\sqrt[3]{-\frac{b}{2}-\sqrt{\frac{b^2}{4}+\frac{a^3}{27}}}$, *then* $x=A+B,\ -\frac{A+B}{2}+\frac{A-B}{2}\sqrt{-3}$, *and* $-\frac{A+B}{2}-\frac{A-B}{2}\sqrt{-3}$ *are solutions.*

If $p,\ q,\ r$ *are real, then*

- *If* $\frac{b^2}{4}+\frac{a^3}{27}>0$, *there will be one real root and two conjugate imaginary roots.*
- *If* $\frac{b^2}{4}+\frac{a^3}{27}=0$, *there will be three real roots of which at least two are equal.*
- *If* $\frac{b^2}{4}+\frac{a^3}{27}<0$, *there will be three real and unequal roots.*

The following iterative procedure can be used to find the maximum likelihood estimates of $\sigma_A,\ \sigma_B$ and ρ under H_A.

Procedure

1. Set starting values using the method of moment estimates from (2.13):

$$\sigma_A^0=\sqrt{\hat{\sigma}^2_{A,UR}},\ \sigma_B^0=\sqrt{\hat{\sigma}^2_{B,UR}},\ \text{and}\ \rho^0=\hat{\rho}_{UR}$$

$\hat{\sigma}^2_{A,UR} = \dfrac{w_{11}}{2n_1+n_2-3}$, $\hat{\sigma}^2_{B,UR} = \dfrac{w_{22}}{n_1+2n_2-3}$, and $\hat{\rho}_{UR} = \dfrac{1}{3}\left[\dfrac{w_{12}}{\sqrt{w_{11}w_{22}}} + \dfrac{w^*_{11}}{w_{11}} + \dfrac{w^*_{22}}{w_{22}}\right]$.

where $w_{11} = w_{111} + w_{133} + w_{222}$, $w_{22} = w_{122} + w_{211} + w_{233}$, $w^*_{11} = w_{113}$, and $w^*_{22} = w_{213}$.

2. Plug σ^0_B, and ρ^0 into the quadratic equation of σ_A:

$$(2n_1+n_2-3)\sigma^2_A + \frac{\rho^0(w_{112}+w_{123}+w_{212}+w_{223})}{(1-\rho^0)(1+2\rho^0)\sigma^0_B}\sigma_A - \frac{(1+\rho^0)(w_{111}+w_{133}+w_{222}) - 2\rho^0 w_{113}}{(1-\rho^0)(1+2\rho^0)} = 0.$$

And obtain a solution for the parameter, σ_A by using Definition 2: call the solution, $\hat{\sigma}_A$.

3. Plug $\hat{\sigma}_A$, and ρ^0 into the quadratic equation of σ_B:

$$(n_1+2n_2-3)\sigma^2_B + \frac{\rho^0(w_{112}+w_{123}+w_{212}+w_{223})}{(1-\rho^0)(1+2\rho^0)\hat{\sigma}_A}\sigma_B - \frac{(1+\rho^0)(w_{122}+w_{211}+w_{233}) - 2\rho^0 w_{213}}{(1-\rho^0)(1+2\rho^0)} = 0.$$

And obtain a solution for the parameter, σ_B by using Definition 2: call the solution, $\hat{\sigma}_B$.

4. Plug $\hat{\sigma}_A$, and $\hat{\sigma}_B$ into the cubic equation of ρ:

$$6(N-2)\rho(1-\rho)(1+2\rho) - \left\{\frac{2\rho(2+\rho)(w_{111}+w_{133}+w_{222}) - (1+2\rho^2)(w_{113}+w_{131})}{\hat{\sigma}^2_A} + \frac{2\rho(2+\rho)(w_{122}+w_{211}+w_{233}) - (1+2\rho^2)(w_{213}+w_{231})}{\hat{\sigma}^2_B} - \frac{2(w_{112}+w_{123}+w_{212}+w_{223})(1+2\rho^2)(1+2\rho^2)}{\hat{\sigma}_A\hat{\sigma}_B}\right\} = 0.$$

And obtain a solution for the parameter, ρ, by using Definition 3: call the solution, $\hat{\rho}$.

5. If $|\hat{\sigma}_A - \sigma^0_A| < \varepsilon$, $|\hat{\sigma}_B - \sigma^0_B| < \varepsilon$, and $|\hat{\rho} - \rho^0| < \varepsilon$ where $\varepsilon = 10^{-5}$, then STOP and use $\hat{\sigma}_A$, $\hat{\sigma}_B$, and $\hat{\rho}$ as maximum likelihood estimates of σ_A, σ_B, and ρ.

   If $|\hat{\sigma}_A - \sigma^0_A| > \varepsilon$ or $|\hat{\sigma}_B - \sigma^0_B| > \varepsilon$ or $|\hat{\rho} - \rho^0| > \varepsilon$, repeat steps 2-4 with $\sigma^0_A \leftarrow \hat{\sigma}_A$, $\sigma^0_B \leftarrow \hat{\sigma}_B$, and $\rho^0 \leftarrow \hat{\rho}$.

No Simulations were carried out for the ABA/BAB design.

2.5 Conclusions

Data were generated for the likelihood ratio test of equal variances for both the null and alternative cases. The Type I error rates for the likelihood ratio test of equal variance is shown in Figure A.1 and Table B.1. The power results for the likelihood ratio tests are shown in Figures A.2-A.5, and Tables B.2-B.5.

Figure A.1 and Table B.1 show that the LRT holds its size when the number of subjects is greater than 18 per sequence. For the case of unequal variance, the power is sharply increased close towards 1 as the number of subjects is increased and the value of variance B is large when compared to the variance of A.

As a special case of two treatment crossover design balanced for carryover effects, ABA/BAB dual balanced design is introduced. Under $H_0 : \sigma_A^2 = \sigma_B^2 = \sigma^2$, the maximum likelihood estimator of σ^2 and ρ has closed form solution, respectively. However, under $H_0 : \sigma_A^2 \neq \sigma_B^2$, the maximum likelihood estimators of σ_A^2, σ_B^2 and ρ do not have closed form solutions. Therefore, a procedure to estimate parameters is introduced in Section 2.4.2. This procedure will be applied in Chapter 3 for a three treatment crossover design, and the performance will be assessed by examining the Type I error rates and power.

2.6 Future Work

In this chapter, two treatment crossover designs with data that have unequal variances are considered. In particular, two kinds of crossover designs were considered; one is a crossover design having the number of treatments equal to the number of periods ($t = p = 2$), AB/BA, and the other is a crossover design balanced for carryover effects known as dual balanced design, ABA/BAB, which is different between the number of treatment and the number of period.

Future work will also consider testing the equality of the variances due to treatments with crossover designs balanced for carryover effects such as the extra period designs, the ABB/BAA and $ABA/ABB/BAB/BAA$, and/or designs with extra sequence(s) such as Balaam design, $AB/BA/AA/BB$.

Chapter 3

Testing for Equal Treatment Variances in a Three Treatment Crossover Design

3.1 Introduction

A crossover design is an experimental design in which each experimental unit receives a series of experimental treatments over time. The order that an experimental unit receives its treatments is called a sequence (for example, the sequence ABC means that treatment A is given first, and then followed by treatment B, then followed by treatment C). A period is the time interval during which a treatment is administered to the experimental unit. A period could range from a few minutes to several months depending on the study. Sequences usually involve subjects receiving a different treatment in each successive period. However, treatments may occur more than once in any sequence (example, $ABAB$).

The basic three treatment crossover design is a cyclic crossover design denoted by $ABC/BCA/CAB$. It means that treatment B always follows treatment A, treatment C always follows treatment B, and treatment A always follows treatment C. In the three treatment/ three sequence crossover design, ABC is the order of the treatments A, B and C in the first sequence, BCA is the order of the treatments B, C and A in the second sequence and CAB is the order of the treatments C, A and B in the third sequence. The three treatment/ three sequence crossover design will be referred to as 3×3 crossover design.

The cyclic crossover design, $ABC/BCA/CAB$, is not balanced for carryover effects.

In the three treatment design, a Williams' design is recommended. Williams (1949)[30] developed a procedure using Latin square designs that balance crossover designs for carryover effects when the number of periods is equal to the number of treatments ($p = t > 2$). The Williams' designs have the characteristic that every treatment follows every other treatment an equal number of times. When implementing the Williams' design procedure, two Latin squares are generated for all experiments involving three or more treatments. Designs that have an odd number of treatments require both the Latin squares to achieve balance for carryover effects, and designs that have an even number of treatments require only one Latin square. If the number of treatments is odd, each treatment follows every other treatment twice. If the number of treatment is even, each treatment follow every other treatment once. Mathews gives a clear description of a general algorithm for developing the Williams' designs. The same description is also given by Jones and Kenward (2003)[16].

Two Latin squares, six different sequences, are used to construct sequences such that each treatment immediately follows every other treatment exactly twice. The three treatment crossover design that is balanced for first order carryover effects has six sequences, $ABC/ACB/BAC/BCA/CAB/CBA$. The six sequence crossover design is recommended since the cyclic three treatment/ three sequence crossover design has carryover effects aliased with treatment effects when carryover is present in the model.

Crossover designs are traditionally analyzed as split-plot designs. A split-plot design is a design that has whole plot and subplot experimental units. The whole plot in a crossover design is the subject to which a sequence of treatments has been assigned and the subplot is the period. There are both between-subject comparisons and within-subject comparisons. One would like treatment comparisons to be within-subject comparisons as within subject variability is usually much smaller than between subject variability. The smaller within subject variability allows for more precise estimators of treatment differences. The within subject comparisons give the crossover design an advantage over a completely randomized

design that could also be used to estimate treatment differences.

In this paper, crossover designs are analyzed when responses due to treatments are assumed to have different variances.

There are two models usually considered in the analysis of crossover designs. The first model is a model that does not include parameters for carryover effects. Consider the model without carryover used by Milliken and Johnson (1992)[21]. This model is

$$y_{ijkl} = \mu + s_i + \delta_{il} + \tau_j + \pi_k + \epsilon_{ijkl} \tag{3.1}$$

where

μ is effect of an overall mean;

s_i is effect of the ith sequence effect, $i = 1, 2, \cdots, s$;

δ_{il} is the experimental error associated with the lth subject in the ith sequence;

τ_j is effect of the jth treatment effect, $j = 1, 2, \cdots, t$;

π_k is effect of the kth period effect, $k = 1, 2, \cdots, p$;

ϵ_{ijkl} is the error associated with the lth subject in the ith sequence that received the jth treatment in the kth period, $l = 1, 2, \cdots, n_i$.

Also, it is often assumed that $\delta_{il} \sim iid\ N(0, \sigma_\delta^2)$ and $\epsilon_{ijkl} \sim iid\ N(0, \sigma_\epsilon^2)$ for all i, j, k, and l. Finally, it is usually assumed that all δ_{il}'s and ϵ_{ijkl}'s are independent.

A second model includes parameters for carryover effects. The model for the response variable may be written by modifying a notation used by Ratkowsky, Evan, and Alldredge(1993)[27]. The model is

$$y_{ijklm} = \mu + s_i + \delta_{il} + \tau_j + \pi_k + \lambda_m + \epsilon_{ijklm} \tag{3.2}$$

where μ, s_i, τ_j, π_k, and ϵ_{ijklm} are defined as in (3.1) above, and λ_k is the carryover effect of the mth treatment administered in period $k-1$, where $m = 1, 2, \cdots, t$. There is no carryover parameter associated with the first period.

Consider a 3×3 crossover design given in Table 3.1.

Table 3.1: *A Three Treatment/Three Sequence Crossover Design(ABC/BCA/CAB)*

	Period		
Sequence	1	2	3
1	A	B	C
2	B	C	A
3	C	A	B

The general ANOVA table for the 3×3 crossover design when $\sigma_A^2 = \sigma_B^2 = \sigma_C^2$ which is analyzed as a split-plot design for (3.1) is given in Table 3.2 where n_i equals the number of subjects assigned to the i^{th} sequence, $i = 1, 2, 3$.

Table 3.2: *ANOVA Table for Model (3.1): a Three Treatment/ Three Sequence Crossover Design (ABC/BCA/CAB) Without Carryover Effects*

Source	df
Between Subject Analysis	
Sequence	2
Subject(sequence)	$(n_1 + n_2 + n_3)$-3
Within Subject Analysis	
Treatment	2
Period	2
Error	$2(n_1 + n_2 + n_3)$-4
Total	$3(n_1 + n_2 + n_3)$-1

The general ANOVA table for the 3×3 crossover design when $\sigma_A^2 = \sigma_B^2 = \sigma_C^2$ and analyzed as a split-plot design for (3.2) is given in Table 3.3 where n_i equals the number of subjects assigned to the i^{th} sequence, $i = 1, 2, 3$.

Two Latin squares, six different sequences, are used to construct sequences such that each treatment immediately follows every other treatment exactly twice. Such a design is shown in Table 3.4. Using the design in Table 3.4, treatment contrasts can be constructed that are free from carryover effects.

Table 3.3: *ANOVA Table for Model (3.2): a Three Treatment/Three Sequence Crossover Design (ABC/BCA/CAB) With Carryover Effects*

Source	df
Between Subject Analysis	
Sequence/Carryover	2
Subject(sequence)	$(n_1 + n_2 + n_3)$-3
Within Subject Analysis	
Treatment	2
Period	2
Carryover	2
Error	$2(n_1 + n_2 + n_3)$-6
Total	$3(n_1 + n_2 + n_3)$-1

Table 3.4: *A Three Treatment/Six Sequence Crossover Design (Two Latin Squares' Design)(ABC/ACB/BAC/BCA/CAB/CBA)*

	Period		
Sequence	1	2	3
1	A	B	C
2	B	C	A
3	C	A	B
4	A	C	B
5	B	A	C
6	C	B	A

The following sections will consider a standard three treatment crossover design to test the equality of variances when three treatments have unequal variances.

3.2 Testing the Equality of the Three Variances due to Treatments

Now consider the three treatment/ three sequence crossover design with three periods (ABC/BCA/CAB) given in Table 3.1. Define the sums of squares for the three sequences

by

$$
\begin{aligned}
W_1 &= \sum_{l=1}^{n_1} (\boldsymbol{y}_{1l} - \hat{\boldsymbol{\mu}}_1)(\boldsymbol{y}_{1l} - \hat{\boldsymbol{\mu}}_1)^T \\
&= \sum_{l=1}^{n_1} \begin{pmatrix} y_{111l} - \bar{y}_{111.} \\ y_{122l} - \bar{y}_{122.} \\ y_{133l} - \bar{y}_{133.} \end{pmatrix} \begin{pmatrix} y_{111l} - \bar{y}_{111.} \\ y_{122l} - \bar{y}_{122.} \\ y_{133l} - \bar{y}_{133.} \end{pmatrix}^T = \begin{pmatrix} w_{111} & w_{112} & w_{113} \\ w_{112} & w_{122} & w_{123} \\ w_{113} & w_{123} & w_{133} \end{pmatrix}, \\
W_2 &= \sum_{l=1}^{n_2} (\boldsymbol{y}_{2l} - \hat{\boldsymbol{\mu}}_2)(\boldsymbol{y}_{2l} - \hat{\boldsymbol{\mu}}_2)^T \\
&= \sum_{l=1}^{n_2} \begin{pmatrix} y_{221l} - \bar{y}_{221.} \\ y_{232l} - \bar{y}_{232.} \\ y_{213l} - \bar{y}_{213.} \end{pmatrix} \begin{pmatrix} y_{221l} - \bar{y}_{221.} \\ y_{232l} - \bar{y}_{232.} \\ y_{213l} - \bar{y}_{213.} \end{pmatrix}^T = \begin{pmatrix} w_{211} & w_{212} & w_{213} \\ w_{212} & w_{222} & w_{223} \\ w_{213} & w_{223} & w_{233} \end{pmatrix},
\end{aligned}
$$

and

$$
\begin{aligned}
W_3 &= \sum_{l=1}^{n_3} (\boldsymbol{y}_{3l} - \hat{\boldsymbol{\mu}}_3)(\boldsymbol{y}_{3l} - \hat{\boldsymbol{\mu}}_3)^T \\
&= \sum_{l=1}^{n_3} \begin{pmatrix} y_{331l} - \bar{y}_{331.} \\ y_{312l} - \bar{y}_{312.} \\ y_{323l} - \bar{y}_{323.} \end{pmatrix} \begin{pmatrix} y_{331l} - \bar{y}_{331.} \\ y_{312l} - \bar{y}_{312.} \\ y_{323l} - \bar{y}_{323.} \end{pmatrix}^T = \begin{pmatrix} w_{311} & w_{312} & w_{313} \\ w_{312} & w_{322} & w_{323} \\ w_{313} & w_{323} & w_{333} \end{pmatrix}.
\end{aligned}
$$

Note that $W_i \sim W_3(n_i - 1, \boldsymbol{\Sigma}_i),\ i = 1, 2, 3$

where

$$
\boldsymbol{\Sigma}_1 = \begin{pmatrix} \sigma_A^2 & \rho\sigma_A\sigma_B & \rho\sigma_A\sigma_C \\ \rho\sigma_A\sigma_B & \sigma_B^2 & \rho\sigma_B\sigma_C \\ \rho\sigma_A\sigma_C & \rho\sigma_B\sigma_C & \sigma_C^2 \end{pmatrix},\ \boldsymbol{\Sigma}_2 = \begin{pmatrix} \sigma_B^2 & \rho\sigma_B\sigma_C & \rho\sigma_A\sigma_B \\ \rho\sigma_B\sigma_C & \sigma_C^2 & \rho\sigma_A\sigma_C \\ \rho\sigma_A\sigma_B & \rho\sigma_A\sigma_C & \sigma_A^2 \end{pmatrix},
$$

and

$$
\boldsymbol{\Sigma}_3 = \begin{pmatrix} \sigma_C^2 & \rho\sigma_A\sigma_C & \rho\sigma_B\sigma_C \\ \rho\sigma_A\sigma_C & \sigma_A^2 & \rho\sigma_A\sigma_B \\ \rho\sigma_B\sigma_C & \rho\sigma_A\sigma_B & \sigma_B^2 \end{pmatrix}.
$$

Define permutation matrices for each sequence by

$$
C_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, C_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, C_3 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},
$$

Define a matrix $\boldsymbol{W}$ by

$$
\boldsymbol{W} = C_1^T W_1 C_1 + C_2^T W_2 C_2 + C_3^T W_3 C_3 = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{12} & w_{22} & w_{23} \\ w_{13} & w_{23} & w_{33} \end{bmatrix}. \tag{3.3}
$$

Note that

$$cov\left(C_i^T \boldsymbol{\varepsilon}_{il}\right) = \boldsymbol{\Sigma},\ i = 1, 2, 3,\ l = 1, 2, \cdots, n_i$$

where

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_A^2 & \rho\sigma_A\sigma_B & \rho\sigma_A\sigma_C \\ \rho\sigma_A\sigma_B & \sigma_B^2 & \rho\sigma_B\sigma_C \\ \rho\sigma_A\sigma_C & \rho\sigma_B\sigma_C & \sigma_C^2 \end{pmatrix}.$$

Thus $\boldsymbol{W} \sim W(N-3, \boldsymbol{\Sigma})$ where $N = n_1+n_2+n_3$ and $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_A^2 & \rho\sigma_A\sigma_B & \rho\sigma_A\sigma_C \\ \rho\sigma_A\sigma_B & \sigma_B^2 & \rho\sigma_B\sigma_C \\ \rho\sigma_A\sigma_C & \rho\sigma_B\sigma_C & \sigma_C^2 \end{pmatrix}$.

Let $D^{1/2} = diag(\sigma_A, \sigma_B, \sigma_C)$ and $\mathbf{R} = \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix}$, then it is noted that $\boldsymbol{\Sigma} = D^{1/2}\mathbf{R}D^{1/2}$.

The inverse of $\boldsymbol{\Sigma}$ is given by

$$\boldsymbol{\Sigma}^{-1} = D^{-1/2}\mathbf{R}^{-1}D^{-1/2}$$

where

$$D^{-1/2} = diag\left(\frac{1}{\sigma_A}, \frac{1}{\sigma_B}, \frac{1}{\sigma_C}\right) \quad \text{and} \quad \mathbf{R}^{-1} = \frac{1}{(1-\rho)(1+2\rho)} \begin{pmatrix} 1+\rho & -\rho & -\rho \\ -\rho & 1+\rho & -\rho \\ -\rho & -\rho & 1+\rho \end{pmatrix}.$$

Consider testing $H_0 : \sigma_A^2 = \sigma_B^2 = \sigma_C^2$ versus $H_A : not\ H_0$. The likelihood function is given by

$$L(\boldsymbol{\Sigma}) = c\frac{|\boldsymbol{W}|^{\frac{N-7}{2}} exp\left(\frac{1}{2}tr\left(\boldsymbol{W}\boldsymbol{\Sigma}^{-1}\right)\right)}{|\boldsymbol{\Sigma}|^{\frac{N-3}{2}}}$$

where

$$c = \frac{1}{2^{3(N-3)/2}\pi^{3/2}\prod_{i=1}^{3}\Gamma\left(\frac{N-(i+2)}{2}\right)} \quad \text{and} \quad N = \sum_{i=1}^{3} n_i.$$

The log-likelihood function is

$$log\left(L(\boldsymbol{\Sigma})\right) = log(c) + \frac{N-7}{2}log|\boldsymbol{W}| - \frac{N-3}{2}log|\boldsymbol{\Sigma}| - \frac{1}{2}tr\left(\boldsymbol{W}\boldsymbol{\Sigma}^{-1}\right).$$

Under $H_0 : \sigma_A^2 = \sigma_B^2 = \sigma_C^2 = \sigma^2$, one has

$$|\boldsymbol{\Sigma}| = \sigma^6(1-\rho)^2(1+2\rho)$$

and

$$tr\left(\boldsymbol{W}\boldsymbol{\Sigma}^{-1}\right) = \frac{(1+\rho)(w_{11}+w_{22}+w_{33})}{(1-\rho)^2(1+2\rho)\sigma^2} - \frac{2\rho(w_{12}+w_{13}+w_{23})}{(1-\rho)^2(1+2\rho)\sigma^2}.$$

The log-likelihood function under restricted condition, H_0, is

$$\begin{aligned} log\left(L(\boldsymbol{\Sigma})\right) = log(c) + \frac{N-7}{2}log|\boldsymbol{W}| - \frac{N-3}{2}\left[3log\sigma^2 + log\left((1-\rho)^2(1+2\rho)\right)\right] \\ -\frac{(1+\rho)(w_{11}+w_{22}+w_{33})}{2(1-\rho)^2(1+2\rho)\sigma^2} + \frac{\rho(w_{12}+w_{13}+w_{23})}{(1-\rho)^2(1+2\rho)\sigma^2}. \end{aligned}$$

The derivatives of $logL(\boldsymbol{\Sigma})$ with respect to σ^2 and ρ are as follows:

$$\begin{aligned} \frac{\partial logL(\boldsymbol{\Sigma})}{\partial\sigma^2} &= -\frac{3(N-3)}{2\sigma^2} + \frac{(1+\rho)(w_{11}+w_{22}+w_{33}) - 2\rho(w_{12}+w_{13}+w_{23})}{2(1-\rho)(1+2\rho)\sigma^4}, \\ \frac{\partial logL(\boldsymbol{\Sigma})}{\partial\rho} &= \frac{3(N-3)\rho}{(1-\rho)(1+2\rho)} - \frac{\rho(2+\rho)(w_{11}+w_{22}+w_{33})}{(1-\rho)^2(1+2\rho)^2\sigma^2} + \frac{(1+2\rho^2)(w_{12}+w_{13}+w_{23})}{(1-\rho)^2(1+2\rho)^2\sigma^2}. \end{aligned}$$

Setting the derivatives of $logL(\boldsymbol{\Sigma})$ with respect to σ^2 and ρ equal to zero, one gets

$$\begin{aligned} &3(N-3)\sigma^2 - \frac{(1+\rho)(w_{11}+w_{22}+w_{33}) - 2\rho(w_{12}+w_{13}+w_{23})}{(1-\rho)(1+2\rho)} = 0, \\ &3(N-3)\rho(1-\rho)(1+2\rho) - \left(\frac{w_{11}+w_{22}+w_{33}}{\sigma^2}\right)\rho(2+\rho) + \left(\frac{w_{12}+w_{13}+w_{23}}{\sigma^2}\right)(1+2\rho^2) = 0. \end{aligned}$$

Under restricted condition, the parameter estimators are

$$\hat{\sigma}_R^2 = \frac{w_{11}+w_{22}+w_{33}}{3(N-3)} \quad \text{and} \quad \hat{\rho} = \frac{w_{12}+w_{13}+w_{23}}{w_{11}+w_{22}+w_{33}}.$$

Thus, the method of moment estimators and the maximum likelihood estimators are identical.

Under $H_A : not\ H_0$, one has

$$|\boldsymbol{\Sigma}| = \sigma_A^2\sigma_B^2\sigma_C^2(1-\rho)^2(1+2\rho)$$

and

$$tr\left(\boldsymbol{W}\boldsymbol{\Sigma}^{-1}\right) = \frac{1+\rho}{(1-\rho)(1+2\rho)}\left[\frac{w_{11}}{\sigma_A^2}+\frac{w_{22}}{\sigma_B^2}+\frac{w_{33}}{\sigma_C^2}\right] - \frac{2\rho}{(1-\rho)(1+2\rho)}\left[\frac{w_{12}}{\sigma_A\sigma_B}+\frac{w_{13}}{\sigma_A\sigma_C}+\frac{w_{23}}{\sigma_B\sigma_C}\right].$$

The log-likelihood function under the unrestricted condition, H_A, is

$$\begin{aligned} logL(\boldsymbol{\Sigma}) &= log(c) + \frac{N-7}{2}log|\boldsymbol{W}| - \frac{N-3}{2}log|\boldsymbol{\Sigma}| - \frac{1}{2}tr\left(\boldsymbol{W}\boldsymbol{\Sigma}^{-1}\right) \\ &= log(c) + \frac{N-7}{2}log|\boldsymbol{W}| - \frac{N-3}{2}\left[log\sigma_A^2 + log\sigma_B^2 + log\sigma_C^2 + log\left\{(1-\rho)^2(1+2\rho)\right\}\right] \\ &\quad -\frac{1+\rho}{2(1-\rho)(1+2\rho)}\left[\frac{w_{11}}{\sigma_A^2}+\frac{w_{22}}{\sigma_B^2}+\frac{w_{33}}{\sigma_C^2}\right] + \frac{\rho}{(1-\rho)(1+2\rho)}\left[\frac{w_{12}}{\sigma_A\sigma_B}+\frac{w_{13}}{\sigma_A\sigma_C}+\frac{w_{23}}{\sigma_B\sigma_C}\right]. \end{aligned}$$

The derivatives of $logL(\boldsymbol{\Sigma})$ with respect to $\sigma_A^2,\ \sigma_B^2,\ \sigma_C^2$ and ρ are

$$\begin{aligned} \frac{\partial logL(\boldsymbol{\Sigma})}{\partial\sigma_A^2} &= -\frac{N-3}{2\sigma_A^2} - \frac{\rho}{2(1-\rho)(1+2\rho)\sigma_A^3}\left(\frac{w_{12}}{\sigma_B}+\frac{w_{13}}{\sigma_C}\right) + \frac{(1+\rho)w_{11}}{2(1-\rho)(1+2\rho)\sigma_A^4}, \\ \frac{\partial logL(\boldsymbol{\Sigma})}{\partial\sigma_B^2} &= -\frac{N-3}{2\sigma_B^2} - \frac{\rho}{2(1-\rho)(1+2\rho)\sigma_B^3}\left(\frac{w_{12}}{\sigma_A}+\frac{w_{23}}{\sigma_C}\right) + \frac{(1+\rho)w_{22}}{2(1-\rho)(1+2\rho)\sigma_B^4}, \\ \frac{\partial logL(\boldsymbol{\Sigma})}{\partial\sigma_C^2} &= -\frac{N-3}{2\sigma_C^2} - \frac{\rho}{2(1-\rho)(1+2\rho)\sigma_C^3}\left(\frac{w_{13}}{\sigma_A}+\frac{w_{23}}{\sigma_B}\right) + \frac{(1+\rho)w_{33}}{2(1-\rho)(1+2\rho)\sigma_C^4}, \\ \frac{\partial logL(\boldsymbol{\Sigma})}{\partial\rho} &= \frac{3(N-3)\rho}{(1-\rho)(1+2\rho)} - \frac{\rho(2+\rho)}{(1-\rho)^2(1+2\rho)^2}\left(\frac{w_{11}}{\sigma_A^2}+\frac{w_{22}}{\sigma_B^2}+\frac{w_{33}}{\sigma_C^2}\right) \\ &\quad + \frac{(1+2\rho^2)}{(1-\rho)^2(1+2\rho)^2}\left(\frac{w_{12}}{\sigma_A\sigma_B}+\frac{w_{13}}{\sigma_A\sigma_C}+\frac{w_{23}}{\sigma_B\sigma_C}\right). \end{aligned} \tag{3.4}$$

Setting the derivatives of $logL(\boldsymbol{\Sigma})$ with respect to $\sigma_A^2,\ \sigma_B^2,\ \sigma_C^2$ and ρ equal to zero and one gets

$$\begin{aligned} &-(N-3)\sigma_A^2 - \frac{\rho}{(1-\rho)(1+2\rho)}\left(\frac{w_{12}}{\sigma_B}+\frac{w_{13}}{\sigma_C}\right)\sigma_A + \frac{(1+\rho)w_{11}}{(1-\rho)(1+2\rho)} = 0, \\ &-(N-3)\sigma_B^2 - \frac{\rho}{(1-\rho)(1+2\rho)}\left(\frac{w_{12}}{\sigma_A}+\frac{w_{23}}{\sigma_C}\right)\sigma_B + \frac{(1+\rho)w_{22}}{(1-\rho)(1+2\rho)} = 0, \\ &-(N-3)\sigma_C^2 - \frac{\rho}{(1-\rho)(1+2\rho)}\left(\frac{w_{13}}{\sigma_A}+\frac{w_{23}}{\sigma_B}\right)\sigma_C + \frac{(1+\rho)w_{33}}{(1-\rho)(1+2\rho)} = 0, \\ &3(N-3)\rho(1-\rho)(1+2\rho) - \rho(2+\rho)\left(\frac{w_{11}}{\sigma_A^2}+\frac{w_{22}}{\sigma_B^2}+\frac{w_{33}}{\sigma_C^2}\right) \\ &\quad +(1+2\rho^2)\left(\frac{w_{12}}{\sigma_A\sigma_B}+\frac{w_{13}}{\sigma_A\sigma_C}+\frac{w_{23}}{\sigma_B\sigma_C}\right) = 0. \end{aligned} \tag{3.5}$$

Unfortunately, the maximum likelihood estimators of $\sigma_A^2,\ \sigma_B^2,\ \sigma_C^2$ and ρ do not have a closed form solution. An iterative process to find solutions for $\hat{\sigma}_{A,UR}^2,\ \hat{\sigma}_{B,UR}^2,\ \hat{\sigma}_{C,UR}^2$ and

$\hat{\rho}_{UR}$ is discussed in Section 2.3.3. Shanga incorrectly claimed that the maximum likelihood estimators of $\sigma_A^2,\ \sigma_B^2,\ \sigma_C^2$ and ρ are

$$\hat{\sigma}_{A,UR}^2 = \frac{w_{11}}{N-3},\ \hat{\sigma}_{B,UR}^2 = \frac{w_{22}}{N-3},\ \hat{\sigma}_{C,UR}^2 = \frac{w_{33}}{N-3}, \tag{3.6}$$
$$\hat{\rho} = \frac{1}{3}\left[\frac{w_{12}}{\sqrt{w_{11}w_{22}}} + \frac{w_{13}}{\sqrt{w_{11}w_{33}}} + \frac{w_{23}}{\sqrt{w_{22}w_{33}}}\right].$$

While these are the method of moment estimates of their respective parameters, they do not maximize the likelihood function.

A likelihood ratio test statistic for testing the equality of the three variances in ABC/BCA/CAB design is given by

$$\lambda = \frac{L\left(\hat{\boldsymbol{\Sigma}}_R\right)}{L\left(\hat{\boldsymbol{\Sigma}}_{UR}\right)} = \frac{c\frac{|W|^{\frac{N-7}{2}} exp\left(-\frac{1}{2}tr\left(W\hat{\boldsymbol{\Sigma}}_R^{-1}\right)\right)}{\left|\hat{\boldsymbol{\Sigma}}_R\right|^{\frac{N-3}{2}}}}{c\frac{|W|^{\frac{N-7}{2}} exp\left(-\frac{1}{2}tr\left(W\hat{\boldsymbol{\Sigma}}_{UR}^{-1}\right)\right)}{\left|\hat{\boldsymbol{\Sigma}}_{UR}\right|^{\frac{N-3}{2}}}} = \frac{\left|\hat{\boldsymbol{\Sigma}}_{UR}\right|^{\frac{N-3}{2}} exp\left(-\frac{1}{2}tr\left(\boldsymbol{W}\hat{\boldsymbol{\Sigma}}_R^{-1}\right)\right)}{\left|\hat{\boldsymbol{\Sigma}}_R\right|^{\frac{N-3}{2}} exp\left(-\frac{1}{2}tr\left(\boldsymbol{W}\hat{\boldsymbol{\Sigma}}_{UR}^{-1}\right)\right)} \tag{3.7}$$

where

$$c = \frac{1}{2^{3(N-3)/2}\pi^{3/2}\prod_{i=1}^{3}\Gamma\left(\frac{N-(i+2)}{2}\right)},\ N = \sum_{i=1}^{3} n_i,$$

$$\hat{\boldsymbol{\Sigma}}_R = \hat{\sigma}_R^2\begin{pmatrix} 1 & \hat{\rho}_R & \hat{\rho}_R \\ \hat{\rho}_R & 1 & \hat{\rho}_R \\ \hat{\rho}_R & \hat{\rho}_R & 1 \end{pmatrix},\ \hat{\boldsymbol{\Sigma}}_{UR} = \begin{pmatrix} \hat{\sigma}_{A,UR}^2 & \hat{\rho}_{UR}\hat{\sigma}_{A,UR}\hat{\sigma}_{B,UR} & \hat{\rho}_{UR}\hat{\sigma}_{A,UR}\hat{\sigma}_{C,UR} \\ \hat{\rho}_{UR}\hat{\sigma}_{A,UR}\hat{\sigma}_{B,UR} & \hat{\sigma}_{B,UR}^2 & \hat{\rho}_{UR}\hat{\sigma}_{B,UR}\hat{\sigma}_{C,UR} \\ \hat{\rho}_{UR}\hat{\sigma}_{A,UR}\hat{\sigma}_{C,UR} & \hat{\rho}_{UR}\hat{\sigma}_{B,UR}\hat{\sigma}_{C,UR} & \hat{\sigma}_{C,UR}^2 \end{pmatrix},$$

and $\boldsymbol{W} = C_1^T W_1 C_1 + C_2^T W_2 C_2 + C_3^T W_3 C_3$.

where

$$W_i = \sum_{k=1}^{3}\left(\boldsymbol{y}_{il} - \hat{\boldsymbol{\mu}}_i\right)\left(\boldsymbol{y}_{il} - \hat{\boldsymbol{\mu}}_i\right)^T,\ i = 1,2,3,\ |W_i| = \left|C_i^T W_i C_i\right|, |C_i| = 1$$

$$C_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, C_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, C_3 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

For large N, the test for equal variances is rejected when $q > \chi_{\alpha,f}^2$ where

$$Q = -2log(\lambda) = -2\left\{log\left(L\left(\hat{\boldsymbol{\Sigma}}_R\right)\right) - log\left(L\left(\hat{\boldsymbol{\Sigma}}_{UR}\right)\right)\right\} \tag{3.8}$$

and $f = 4 - 2 = 2$.

3.3 Methods for Estimating Parameters

Since, under the unrestricted condition in the (ABC/BCA/CAB) design, solutions of $\sigma_A^2,\ \sigma_B^2,\ \sigma_C^2$ and ρ for this design do not have a closed form, one needs to consider other methods to find the maximum likelihood estimators under H_A.

Consider the derivatives of the log-likelihood functions from (3.4):

$$\begin{aligned}
\frac{\partial logL(\mathbf{\Sigma})}{\partial \sigma_A^2} &= -\frac{N-3}{2\sigma_A^2} - \frac{\rho}{2(1-\rho)(1+2\rho)\sigma_A^3}\left(\frac{w_{12}}{\sigma_B}+\frac{w_{13}}{\sigma_C}\right) + \frac{(1+\rho)w_{11}}{2(1-\rho)(1+2\rho)\sigma_A^4},\\
\frac{\partial logL(\mathbf{\Sigma})}{\partial \sigma_B^2} &= -\frac{N-3}{2\sigma_B^2} - \frac{\rho}{2(1-\rho)(1+2\rho)\sigma_B^3}\left(\frac{w_{12}}{\sigma_A}+\frac{w_{23}}{\sigma_C}\right) + \frac{(1+\rho)w_{22}}{2(1-\rho)(1+2\rho)\sigma_B^4},\\
\frac{\partial logL(\mathbf{\Sigma})}{\partial \sigma_C^2} &= -\frac{N-3}{2\sigma_C^2} - \frac{\rho}{2(1-\rho)(1+2\rho)\sigma_C^3}\left(\frac{w_{13}}{\sigma_A}+\frac{w_{23}}{\sigma_B}\right) + \frac{(1+\rho)w_{33}}{2(1-\rho)(1+2\rho)\sigma_C^4},\\
\frac{\partial logL(\mathbf{\Sigma})}{\partial \rho} &= \frac{3(N-3)\rho}{(1-\rho)(1+2\rho)} - \frac{\rho(2+\rho)}{(1-\rho)^2(1+2\rho)^2}\left(\frac{w_{11}}{\sigma_A^2}+\frac{w_{22}}{\sigma_B^2}+\frac{w_{33}}{\sigma_C^2}\right)\\
&\quad + \frac{(1+2\rho^2)}{(1-\rho)^2(1+2\rho)^2}\left(\frac{w_{12}}{\sigma_A\sigma_B}+\frac{w_{13}}{\sigma_A\sigma_C}+\frac{w_{23}}{\sigma_B\sigma_C}\right).
\end{aligned} \tag{3.9}$$

Setting the derivatives of $logL(\mathbf{\Sigma})$ with respect to $\sigma_A^2,\ \sigma_B^2,\ \sigma_C^2$ and ρ equal to zero and one can get:

$$\begin{aligned}
&-(N-3)\sigma_A^2 - \frac{\rho}{(1-\rho)(1+2\rho)}\left(\frac{w_{12}}{\sigma_B}+\frac{w_{13}}{\sigma_C}\right)\sigma_A + \frac{(1+\rho)w_{11}}{(1-\rho)(1+2\rho)} = 0,\\
&-(N-3)\sigma_B^2 - \frac{\rho}{(1-\rho)(1+2\rho)}\left(\frac{w_{12}}{\sigma_A}+\frac{w_{23}}{\sigma_C}\right)\sigma_B + \frac{(1+\rho)w_{22}}{(1-\rho)(1+2\rho)} = 0,\\
&-(N-3)\sigma_C^2 - \frac{\rho}{(1-\rho)(1+2\rho)}\left(\frac{w_{13}}{\sigma_A}+\frac{w_{23}}{\sigma_B}\right)\sigma_C + \frac{(1+\rho)w_{33}}{(1-\rho)(1+2\rho)} = 0,\\
&3(N-3)\rho(1-\rho)(1+2\rho) - \rho(2+\rho)\left(\frac{w_{11}}{\sigma_A^2}+\frac{w_{22}}{\sigma_B^2}+\frac{w_{33}}{\sigma_C^2}\right)\\
&\quad +(1+2\rho^2)\left(\frac{w_{12}}{\sigma_A\sigma_B}+\frac{w_{13}}{\sigma_A\sigma_C}+\frac{w_{23}}{\sigma_B\sigma_C}\right) = 0.
\end{aligned} \tag{3.10}$$

Note that the first three equations in (3.10) are quadratic functions of $\sigma_A,\ \sigma_B$ and σ_C, respectively, and the fourth equation is a cubic function of ρ. The theorem 2 and 3 can be applied to estimate the parameters, $\sigma_A,\ \sigma_B,\ \sigma_C$ and ρ.

The following iterative procedure can be used to find the maximum likelihood estimates of $\sigma_A,\ \sigma_B,\ \sigma_C$ and ρ.

Procedure

1. Set starting values using the method of moment estimates from (3.6):

$$\sigma_A^0 = \sqrt{\hat{\sigma}^2_{A,UR}},\ \sigma_B^0 = \sqrt{\hat{\sigma}^2_{B,UR}},\ \sigma_C^0 = \sqrt{\hat{\sigma}^2_{C,UR}},\ \text{and}\ \rho^0 = \hat{\rho}_{UR}$$

$$\text{where} \quad \hat{\sigma}^2_{A,UR} = \frac{w_{11}}{N-3},\ \hat{\sigma}^2_{B,UR} = \frac{w_{22}}{N-3},\ \hat{\sigma}^2_{C,UR} = \frac{w_{33}}{N-3},$$
$$\hat{\rho} = \frac{1}{3}\left[\frac{w_{12}}{\sqrt{w_{11}w_{22}}} + \frac{w_{13}}{\sqrt{w_{11}w_{33}}} + \frac{w_{23}}{\sqrt{w_{22}w_{33}}}\right].$$

2. Plug $\sigma_B^0,\ \sigma_C^0$, and ρ^0 into the quadratic equation of σ_A:

$$-(N-3)\sigma_A^2 - \frac{\rho^0}{(1-\rho^0)(1+2\rho^0)}\left(\frac{w_{12}}{\sigma_B^0} + \frac{w_{13}}{\sigma_C^0}\right)\sigma_A + \frac{(1+\rho^0)w_{11}}{(1-\rho^0)(1+2\rho^0)} = 0$$

And obtain a solution for the parameter, σ_A by using Definition 2: call the solution, $\hat{\sigma}_A$.

3. Plug $\hat{\sigma}_A,\ \sigma_C^0$, and ρ^0 into the quadratic equation of σ_B:

$$-(N-3)\sigma_B^2 - \frac{\rho^0}{(1-\rho^0)(1+2\rho^0)}\left(\frac{w_{12}}{\hat{\sigma}_A} + \frac{w_{23}}{\sigma_C^0}\right)\sigma_B + \frac{(1+\rho^0)w_{22}}{(1-\rho^0)(1+2\rho^0)} = 0$$

And obtain a solution for the parameter, σ_B by using Definition 2: call the solution, $\hat{\sigma}_B$.

4. Plug $\hat{\sigma}_A,\ \hat{\sigma}_B$, and ρ^0 into the quadratic equations of σ_C:

$$-(N-3)\sigma_C^2 - \frac{\rho^0}{(1-\rho^0)(1+2\rho^0)}\left(\frac{w_{13}}{\hat{\sigma}_A} + \frac{w_{23}}{\hat{\sigma}_B}\right)\sigma_C + \frac{(1+\rho^0)w_{33}}{(1-\rho^0)(1+2\rho^0)} = 0$$

And obtain a solution for the parameter, σ_C by using Definition 2: call the solution, $\hat{\sigma}_C$.

5. Plug $\hat{\sigma}_A,\ \hat{\sigma}_B$, and $\hat{\sigma}_C$ into the cubic equation of ρ:

$$3(N-3)\rho(1-\rho)(1+2\rho) - \rho(2+\rho)\left(\frac{w_{11}}{\sigma_A^2} + \frac{w_{22}}{\sigma_B^2} + \frac{w_{33}}{\sigma_C^2}\right)$$
$$+(1+2\rho^2)\left(\frac{w_{12}}{\sigma_A\sigma_B} + \frac{w_{13}}{\sigma_A\sigma_C} + \frac{w_{23}}{\sigma_B\sigma_C}\right) = 0.$$

And obtain a solution for the parameter, ρ, by using Definition 3: call the solution, $\hat{\rho}$.

6. If $|\hat{\sigma}_A - \sigma_A^0| < \varepsilon,\ |\hat{\sigma}_B - \sigma_B^0| < \varepsilon,\ |\hat{\sigma}_C - \sigma_C^0| < \varepsilon$, and $|\hat{\rho} - \rho^0| < \varepsilon$ where $\varepsilon = 10^{-5}$, then STOP and use $\hat{\sigma}_A,\ \hat{\sigma}_B,\ \hat{\sigma}_C$, and $\hat{\rho}$ as maximum likelihood estimates of $\sigma_A,\ \sigma_B,\ \sigma_C$, and ρ.

   If $|\hat{\sigma}_A - \sigma_A^0| > \varepsilon$ or $|\hat{\sigma}_B - \sigma_B^0| > \varepsilon$ or $|\hat{\sigma}_C - \sigma_C^0| > \varepsilon$ or $|\hat{\rho} - \rho^0| > \varepsilon$, repeat steps 2-5 with $\sigma_A^0 \leftarrow \hat{\sigma}_A,\ \sigma_B^0 \leftarrow \hat{\sigma}_B,\ \sigma_C^0 \leftarrow \hat{\sigma}_C$, and $\rho^0 \leftarrow \hat{\rho}$.

The following section considers the performance of the proposed method to obtain the likelihood ratio test statistics.

3.4 A Simulation Study

Consider simulating data for sequence 1. The variance-covariance matrix for sequence 1 is

$$\boldsymbol{\Sigma}_1 = \begin{pmatrix} \sigma_A^2 & \rho\sigma_A\sigma_B & \rho\sigma_A\sigma_C \\ \rho\sigma_A\sigma_B & \sigma_B^2 & \rho\sigma_B\sigma_C \\ \rho\sigma_A\sigma_C & \rho\sigma_B\sigma_C & \sigma_C^2 \end{pmatrix}.$$

The variance-covariance matrix for sequence 1 can be factored as $\boldsymbol{\Sigma} = \boldsymbol{U}'\boldsymbol{U}$ where $\boldsymbol{U}$ is a unique 2×2 upper triangular matrix (Graybill, 1976)[7]. The matrix $\boldsymbol{U}$ is given by

$$\boldsymbol{U} = \begin{pmatrix} \sigma_A & \rho\sigma_B & \rho\sigma_C \\ 0 & \sigma_B\sqrt{1-\rho^2} & \rho\sigma_C\sqrt{\frac{1-\rho}{1+\rho}} \\ 0 & 0 & \sigma_C\sqrt{\frac{1+\rho-2\rho^2}{1+\rho}} \end{pmatrix}.$$

Let $\boldsymbol{x} = [x_1, x_2, x_3]' \sim N\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right]$ and write

$$\boldsymbol{U}'\boldsymbol{x} = \begin{pmatrix} \sigma_A^2 & 0 & 0 \\ \rho\sigma_B & \sigma_B\sqrt{1-\rho^2} & 0 \\ \rho\sigma_C & \rho\sigma_C\sqrt{\frac{1-\rho}{1+\rho}} & \sigma_C\sqrt{\frac{1+\rho-2\rho^2}{1+\rho}} \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Then

$$\boldsymbol{U}'\boldsymbol{x} \sim N\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_A^2 & \rho\sigma_A\sigma_B & \rho\sigma_A\sigma_C \\ \rho\sigma_A\sigma_B & \sigma_B^2 & \rho\sigma_B\sigma_C \\ \rho\sigma_A\sigma_C & \rho\sigma_B\sigma_C & \sigma_C^2 \end{pmatrix}\right].$$

Let $\boldsymbol{y} = \boldsymbol{U}'\boldsymbol{x}$, then $\boldsymbol{y} = [y_1, y_2, y_3]^T$ where $y_1 = \sigma_a x_1$, $y_2 = \rho\sigma_B x_1 + \sigma_B\sqrt{1-\rho^2}x_2$, and $y_3 = \rho\sigma_C x_1 + \rho\sigma_C\sqrt{\frac{1-\rho}{1+\rho}}x_2 + \sigma_C\sqrt{\frac{1+\rho-2\rho^2}{1+\rho}}x_3$. The x's were generated from a standard normal distribution and the above transformations were made to get y's. Furthermore, appropriate cell parameters were added to the y's to get the expected cell means. For example appropriate cell parameters for sequence 1 yield $y_1^* = y_1 + \mu + \tau_A + \pi_1$ for period 1, $y_2^* = y_2 + \mu + \tau_B + \pi_2 + \lambda_A$ for period 2, and $y_3^* = y_3 + \mu + \tau_C + \pi_3 + \lambda_B$ for period 3. Without loss of generality, μ, π_1, π_2, π_3, τ_A, τ_B, τ_C, λ_A, λ_B and λ_C were all fixed at zero in the simulation study. Data for other two sequences were similarly generated using the

variance-covariance matrix

$$\boldsymbol{\Sigma}_2 = \begin{pmatrix} \sigma_B^2 & \rho\sigma_B\sigma_C & \rho\sigma_A\sigma_B \\ \rho\sigma_B\sigma_C & \sigma_C^2 & \rho\sigma_A\sigma_C \\ \rho\sigma_A\sigma_B & \rho\sigma_A\sigma_C & \sigma_A^2 \end{pmatrix}$$

and

$$\boldsymbol{\Sigma}_3 = \begin{pmatrix} \sigma_C^2 & \rho\sigma_A\sigma_C & \rho\sigma_B\sigma_C \\ \rho\sigma_A\sigma_C & \sigma_A^2 & \rho\sigma_A\sigma_B \\ \rho\sigma_B\sigma_C & \rho\sigma_A\sigma_B & \sigma_B^2 \end{pmatrix}.$$

and appropriate cell mean parameters.

Table 3.5: *Setup of Parameters for Type I error rate and Power Analysis for the equality of variances due to Treatment at a Three Treatment/ Three Sequence Crossover Design (ABC/BCA/ CAB) for Model*

	$\lambda_A = \lambda_B = \lambda_C = 0,\ \tau_A = \tau_B = \tau_C = 0,\ \sigma_A^2 = \sigma_B^2 = 1$	
ρ	Equal Variance	Unequal Variance
0	$\sigma_C^2 = 1$	$\sigma_C^2 = 2, 4, 8, 16$
0.1	$\sigma_C^2 = 1$	$\sigma_C^2 = 2, 4, 8, 16$
0.3	$\sigma_C^2 = 1$	$\sigma_C^2 = 2, 4, 8, 16$
0.5	$\sigma_C^2 = 1$	$\sigma_C^2 = 2, 4, 8, 16$
0.7	$\sigma_C^2 = 1$	$\sigma_C^2 = 2, 4, 8, 16$
0.9	$\sigma_C^2 = 1$	$\sigma_C^2 = 2, 4, 8, 16$
	$\lambda_A = \lambda_B = \lambda_C = 0,\ \tau_A = \tau_B = \tau_C = 0,\ \sigma_C^2 = 1$	
ρ	Equal Variance	Unequal Variance
0	$(\sigma_A^2, \sigma_B^2) = (1, 1)$	$(\sigma_A^2, \sigma_B^2) = (2, 2), (4, 4), (8, 8), (16, 16)$
0.1	$(\sigma_A^2, \sigma_B^2) = (1, 1)$	$(\sigma_A^2, \sigma_B^2) = (2, 2), (4, 4), (8, 8), (16, 16)$
0.3	$(\sigma_A^2, \sigma_B^2) = (1, 1)$	$(\sigma_A^2, \sigma_B^2) = (2, 2), (4, 4), (8, 8), (16, 16)$
0.5	$(\sigma_A^2, \sigma_B^2) = (1, 1)$	$(\sigma_A^2, \sigma_B^2) = (2, 2), (4, 4), (8, 8), (16, 16)$
0.7	$(\sigma_A^2, \sigma_B^2) = (1, 1)$	$(\sigma_A^2, \sigma_B^2) = (2, 2), (4, 4), (8, 8), (16, 16)$
0.9	$(\sigma_A^2, \sigma_B^2) = (1, 1)$	$(\sigma_A^2, \sigma_B^2) = (2, 2), (4, 4), (8, 8), (16, 16)$
	$\lambda_A = \lambda_B = \lambda_C = 0,\ \tau_A = \tau_B = \tau_C = 0,\ \sigma_A^2 = 1$	
ρ	Equal Variance	Unequal Variance
0	$(\sigma_B^2, \sigma_C^2) = (1, 1)$	$(\sigma_B^2, \sigma_C^2) = (2, 4),\ (4, 8)$
0.1	$(\sigma_B^2, \sigma_C^2) = (1, 1)$	$(\sigma_B^2, \sigma_C^2) = (2, 4),\ (4, 8)$
0.3	$(\sigma_B^2, \sigma_C^2) = (1, 1)$	$(\sigma_B^2, \sigma_C^2) = (2, 4),\ (4, 8)$
0.5	$(\sigma_B^2, \sigma_C^2) = (1, 1)$	$(\sigma_B^2, \sigma_C^2) = (2, 4),\ (4, 8)$
0.7	$(\sigma_B^2, \sigma_C^2) = (1, 1)$	$(\sigma_B^2, \sigma_C^2) = (2, 4),\ (4, 8)$
0.9	$(\sigma_B^2, \sigma_C^2) = (1, 1)$	$(\sigma_B^2, \sigma_C^2) = (2, 4),\ (4, 8)$

Consider the likelihood ratio test of the three treatment crossover design when treatments have unequal variances. Table 3.5 shows the parameters that were used when simulating data for Type I error rate and power analyses for testing the equality of variances due to treatments. To get the empirical Type I error rates and the power, 1000 simulations were done for each n, ρ, σ_A^2, σ_B^2 and σ_C^2.

To estimate parameters, the proposed method mentioned in Section 2.2 is used in R. And, with the same data generated in R, the likelihood ratio tests using SAS-MIXED to calculate Type I error rates and powers are obtained using the following SAS commands.

Step 1 Import data

```
INFILE 'C:\Data\KSU PhD THESIS\PRG\y12-148data.txt' DELIMITER=' ';
INPUT rho sA2 sB2 sC2 nsubj sim seq subj y1 y2 y3;
```

Step 2 Define periods and arrange treatments

```
DATA one; SET one;
 DROP y1 y2 y3;
 per=1; y=y1; output;
 per=2; y=y2; output;
 per=3; y=y3; output;
RUN;


DATA one; SET one;
    SUbJ=subj+(seq-1)*nsubj;
    trt='A';
    IF seq=1 and per=2 THEN trt='B';
    IF seq=1 and per=3 THEN trt='C';
    IF seq=2 and per=1 THEN trt='B';
    IF seq=2 and per=2 THEN trt='C';
    IF seq=3 and per=1 THEN trt='C';
```

```
    IF seq=3 and per=3 THEN trt='B';
RUN;
```

Step 3 Calculate $-2log\left(L\left(\hat{\boldsymbol{\Sigma}}_R\right)\right)$ in (3.8)

```
PROC MIXED ic data=one;
 TITLE 'ANALYSIS USING SAS-MIXED';
 by rho sim;
 CLASSES seq per trt subj;
 MODEL y=seq trt per/DDFM=SATTERTH;
 REPEATED trt/ SUBJECT=subj TYPE=CS;
 ods listing exclude all;
 ods output infocrit = null COVPARMS=HOPARMS;
RUN;

DATA null; set null;
rename neg2loglike =ho;
drop aic--caic;
```

Step 4 Calculate $-2log\left(L\left(\hat{\boldsymbol{\Sigma}}_{UR}\right)\right)$ in (3.8)

```
PROC MIXED ic data=one;
 TITLE 'ANALYSIS USING SAS-MIXED';
 by rho sim;
 CLASSES seq per trt subj;
 MODEL y=seq trt per/DDFM=SATTERTH;
 REPEATED trt/ SUBJECT=subj TYPE=CSH;
 ods listing exclude all;
 ods output infocrit = ha COVPARMS=HAPARMS;
RUN;
```

```
DATA ha; set ha;
rename neg2loglike =ha;
drop aic--caic;
```

Step 5 Calculate the Type I error rate

```
DATA comb; SET comb;
 u = ho-ha;
 IF u>5.991465 THEN reject=1; ELSE reject=0;
RUN;


PROC MEANS data=comb;
 by rho;
 ods listing select all;
 var reject;
RUN;
```

All of the SAS steps for the three treatment design are shown in Appendix F. Figure C.1 shows the Type I error rates of the likelihood ratio tests for $H_0 : \sigma_A^2 = \sigma_B^2 = \sigma_C^2$ and Table D.1 shows the observed Type I error of the likelihood ratio tests for $H_0 : \sigma_A^2 = \sigma_B^2 = \sigma_C^2$. Figure C.1 has six lines. Each line corresponds to a different value of ρ, $\rho = 0.0, 0.1, 0.3, 0.5, 0.7, 0.9$. Figure C.1 shows that the Type I error with different correlation converge to 0.05 as the number of subjects in each sequence becomes large. Table D.1 shows the observed Type I error rates of the likelihood ratio tests for $H_0 : \sigma_A^2 = \sigma_B^2 = \sigma_C^2$ by both R and SAS-MIXED. Each row for each method in Table D.1 corresponds to a different value of ρ, $\rho = 0.0, 0.1, 0.3, 0.5, 0.7, 0.9$. Each column corresponds to a different number of subjects assigned to each sequence. The likelihood ratio tests generally hold their size when the number of subjects is at least 12 per sequence for each value of ρ. The two analyses differ in their Type I error rates only when the number of subjects is 3 and correlation is 0, 0.1, 0.3, 0.5, and 0.7. When the value of correlation is 0.1 and 0.3, Type I error rate by R is

larger than SAS. In the case of that correlation is 0, 0.5, and 0.7, Type I error rate by SAS is larger than R. In no case are the differences large enough to be of any concern.

Next consider data that were generated for the cases when $\sigma_A^2 = \sigma_B^2 = 1$ and $\sigma_C^2 = 2, 4, 8, 16$. Figure C.2 and Table D.2 show the observed power for each ρ and n when $\sigma_C^2 = 2$. As the number of subjects increase from 6 to 24, the power sharply increases towards 1. Figures C.3-C.5 show the trends and Tables D.3-D.5 show the observed power as σ_C^2 increases. When σ_C^2 increases to 4, 8, and 16, the power sharply increases towards 1. The power is close to 1 as the value of correlation become larger. In every case the power increases as ρ increases.

Figures C.6-C.9 and Tables D.6-D.9 use data that were generated in the case of changing the values of variances A and B and holding $\sigma_C^2 = 1$. Figures C.6-C.9 show the trends and Tables D.6-D.9 show the observed powers as σ_A^2 and σ_B^2 increase. As the number of subjects are increased from 6 to 24, the power is sharply increased towards 1. When the values of variances A and B are increased to 4, 8, and 16, the power is sharply increased towards 1.

The final case is for data when all three variances are different. The trends are similar to the two previous cases. See Figures C.10 and C.11, and Tables D.10 and D.11.

3.5 Conclusions

Data were generated for likelihood ratio test of equal variances for both the equal variance case and the unequal variance case. The type I error rate results of the likelihood ratio test for data with equal variance are shown in Figure C.1 and Table D.1. The power results of the likelihood ratio tests for data with unequal variances are shown in Figures C.2-C.11, and Tables D.2-D.11.

Generally, one needs at least 12 subjects per sequence to achieve acceptable Type I error rates for each value of ρ. When one has at least 12 subjects per sequence, the power of the likelihood ratio tests tend to increase as ρ increases.

It is important to note that while these simulation results were performed for the three treatment/three period crossover design without carryover in a three sequence design, the simulations performed using R are much more general than that. The simulations performed using R were based on the sufficient statistics. In particular, they were based on the distribution of $\boldsymbol{W}$ defined in (3.4). Such a $\boldsymbol{W}$ has a Wishart distribution whether there is carryover or not. Also, if one should have the three treatment/three period crossover design with or without carryover in a six sequence design, the distribution of $\boldsymbol{W}$ will be Wishart with degrees of freedom equal to $N-6$ where $N = n_1+n_2+n_3+n_4+n_5+n_6$. So the simulation results reported in Appendices C and D are valid for the six sequence design if one views them as to the number of degrees of freedom associated with the underlying Wishart distribution. For example, six subjects per sequence in the three sequence design will have results similar to three subjects per sequence in the six sequence design as the former will have 15 degrees of freedom for the Wishart distribution and the latter will have 12 degrees of freedom for the Wishart distribution, and 9 subjects per sequence in a six sequence design will perform similarly to 18 subjects per sequence in a three sequence design as the degrees of freedom in the corresponding Wishart distributions are 48 and 51, respectively.

3.6 Future Work

In this chapter, crossover designs with data that have unequal variances are considered. In particular, a crossover design having the number of treatments equal to the number of periods ($t = p = 3$), $ABC/BCA/CAB$, was considered in this chapter. A constant compound symmetry correlation structure under null hypothesis was considered.

As described in the previous section, the simulation results provided in this dissertation can also be used to evaluate tests for equal variances when carryover is present and for the Williams' design for three treatments in three periods for the likelihood ratio tests based on the Wishart distribution. However, analyses using SAS-MIXED in the three sequence design cannot be generalized to the carryover case nor to Williams' designs. Simulations when carryover is present in the six sequence three treatment Williams design using SAS-MIXED should be obtained. Tests for equal variances should also be obtained for Williams' designs that involve more than three treatments using both SAS-MIXED and the appropriate likelihood ratio test statistics.

All of the results in this dissertation assume that the underlying correlation structure for the period measurements is compound symmetry. One should also consider other correlation structures such as an autoregressive of lag 1 structure, a Toeplitz structure, or a banded Toeplitz structure. In each of these cases, maximum likelihood estimators will need to be obtained for both the null and alternative cases when one is basing tests on the appropriate Wishart matrices.

Another topic for future study is to determine the robustness of these tests for equal variances when the data are not normal.

Bibliography

[1] Balaam, L. N. (1968). A two-period design with t^2 experimental units. *Biometrics*, 24:61–67.

[2] Brant, A. E. (1938). Tests of significance in reversal or switchback trials. *Research Bulletin*, 234:Iowa agriculture experimental Station.

[3] Elfving, G. (1947). The asymptotical distribution of range in samples from a normal population. *Biometrika*, 34:111–119.

[4] Fieller, E. C. (1940). The biological standardization of insulin (with discussion). *Supplement to the Journal of Royal Statistical Society*, 7:1–64.

[5] Geisser, S. and Greenhouse, S. W. (1958). Extension of box's results on the use of the f distribution in multivariate analysis. *Annals of Mathematical Statistics*, 29:885–891.

[6] Goad, C. (1994). Multivariate and mixed models analysis of crossover designs with more than two periods.

[7] Graybill, F. A. (1976). *Theory and Application of the Linear Model*. Duxbury Press: North Scituate, Massachusetts.

[8] Greenhouse, S. W. and Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24:95–112.

[9] Grizzle, J. E. (1965). The two-period changeover design and its use in clinical trials. *Biometrics*, 21:467–480.

[10] Hsu, P. L. (1939). A new proof of thejoint product moment distribution. *Proceedings of the Cambridge Philosophical Society*, 35:336.

[11] Hsu, P. L. (1940). On generalized analysis of variance. *Biometrika*, 31:221–237.

[12] Huynh, H. and Feldt, L. S. (1970). Condition under which mean square ratios in repeated measurements design have exact f distributions. *Journal of the American Statistical Association*, 65(332):1582–1589.

[13] Huynh, H. and Feldt, L. S. (1976). Estimation of the box correction for degrees of fredom from sample data in randomized block and split-plot designs. *Journal of Education Statistics*, 1:69–82.

[14] Jambunathan, M. V. (1965). A quick method of deriving wishart's distribution. *Current Series*, 34:78.

[15] James, A. T. (1954). Normal multivariate analysis and the orthogonal group. *Annals of Mathematical Statistics*, 25:40–75.

[16] Jones, B. and Kenward, M. (2003). *Design and Analysis of Cross-Over Trials*. London: Chapman and Hall.

[17] Madow, W. G. (1938). Contributions to the theory of multivariate statistical analysis. *Trans. of the Amer. Math. Soc.*, 44:476.

[18] Mahalanobis, P. C., Bose, R. C., and Roy, S. N. (1937). Normalisation of statistical variates and the use of rectangular co-ordinates in the theory of sampling distributions. *Sankhyā*, 3:1–40.

[19] Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979). *Multivariate Ananysis*. California: Academic Press.

[20] Matthews, J. N. (1988). Recent development in crossover design. *International Statistical Review*, 56(2):117–127.

[21] Milliken, G. A. and Johnson, D. E. (1992). *Ananysis of Messy Data*. New York: Chapman and Hill.

[22] Mood, A. M., Graybill, F. A., and Boes, D. C. (1974). *Introduction to the Theory of Statistics*. McGraw-Hill, Inc: New York.

[23] Ogawa, I. and Roy, S. N. (1954). On multivariate distribution theory. *Annals of Mathematical Statistics*, 25:329–339.

[24] Ogawa, J. (1953). On the sampling distributions of classical statistics in multivariate analysis. *Osaka Mathematical Journal*, 5:13–52.

[25] Patterson, H. D. and Lucas, H. L. (1959). Extra-period change-over designs. *Biometrics*, 15:116–132.

[26] Rasch, G. (1948). A functional equation for wishart's distribution. *Annals of Mathematical Statistics*, 19:262–266.

[27] Ratkowsky, D. A., Evans, M. A., and Alldredge, J. R. (1993). *Crossover Experiments, Design, Analysis, and Application*. New York: MArcel Dekker, Inc.

[28] Shanga, G. M. (2003). Analysis of crossover designs when treatments have unequal variances.

[29] Sverdrup, E. (1947). Derivation of the wishart distribution of second order mements by straight forward integration of a multiple integral. *Skand. Aktuarietdskr.*, 30:151–166.

[30] Williams, E. J. (1949). Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Scientific Research*, 2:149–164.

[31] Wishart, J. (1928). Sampling errors in the theory of two factors. *British Journal of Psychology*, 19:180–187.

[32] Wishart, J. and Bartlett, M. S. (1933). The distribution of second order moment statistics in a normal system. *Proceedings of the Cambridge Philosophical Society*, 28:455–459.

Appendix A

Chapter 2: Figures for Two Treatment Crossover Design

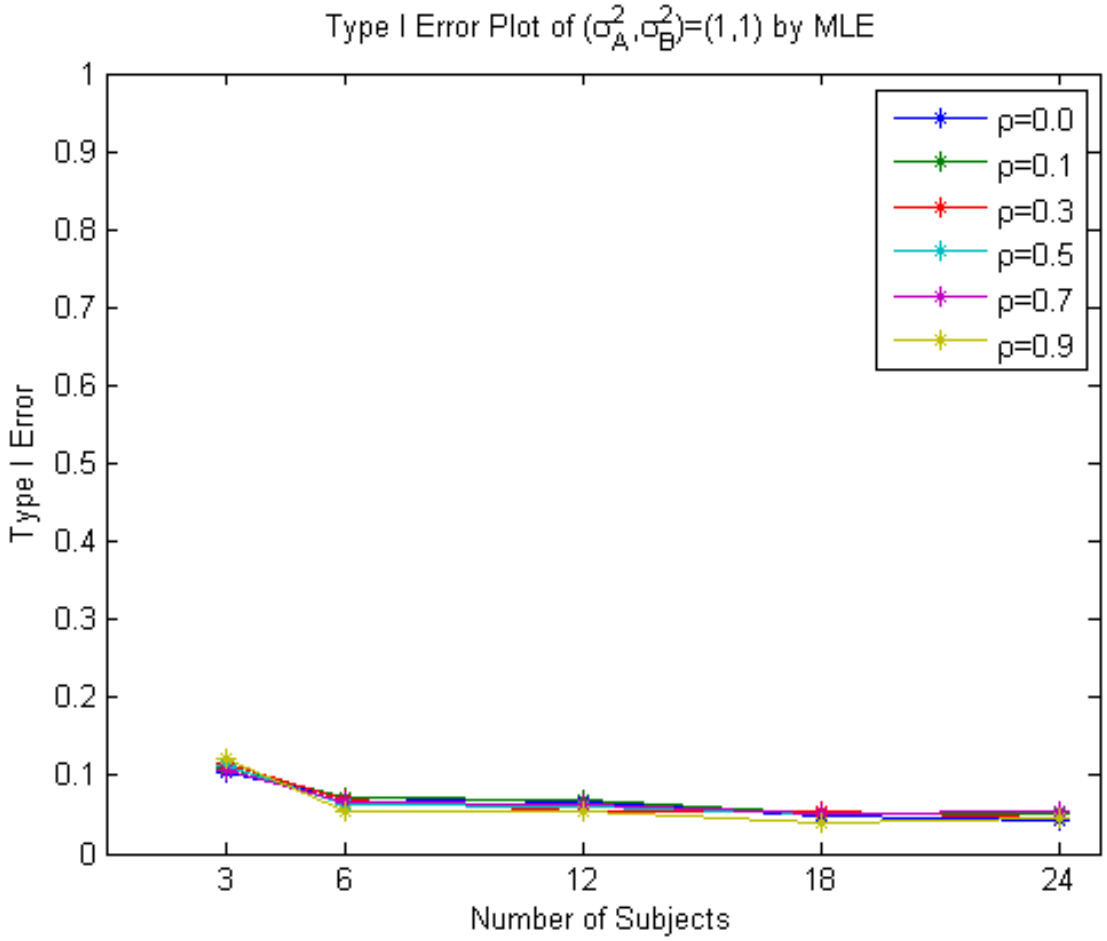

Figure A.1: *Type I Error Plot at* (σ_A^2, σ_B^2)*=(1,1)*

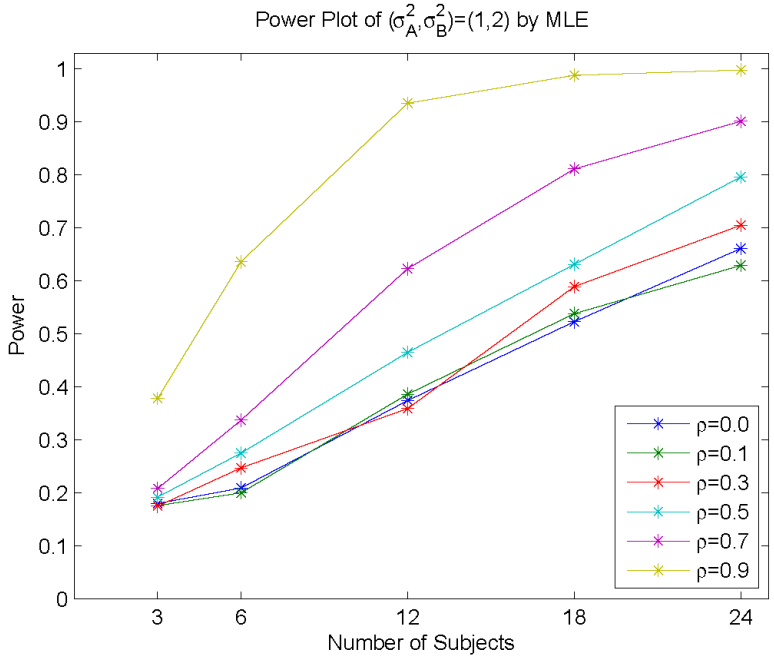

Figure A.2: *Power Plot at* (σ_A^2, σ_B^2)*=(1,2)*

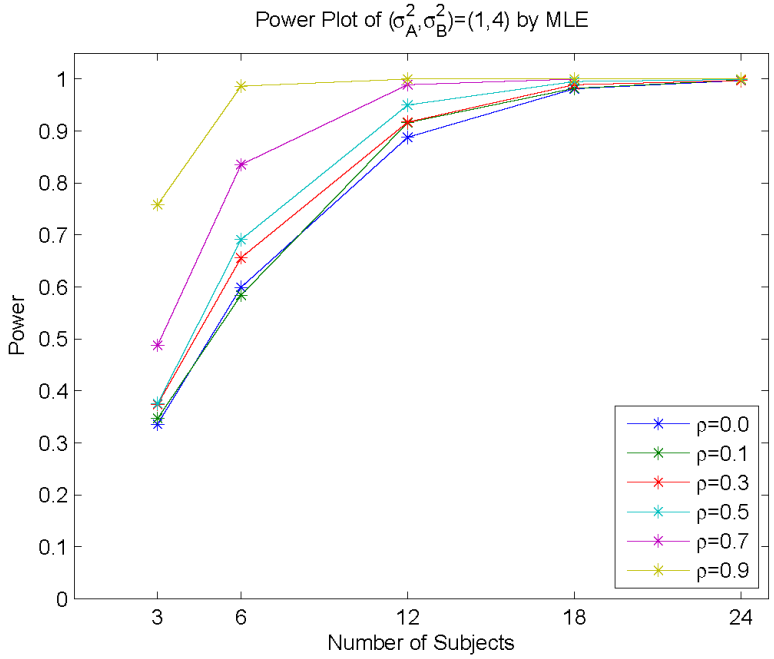

Figure A.3: *Power Plot at* (σ_A^2, σ_B^2)*=(1,4)*

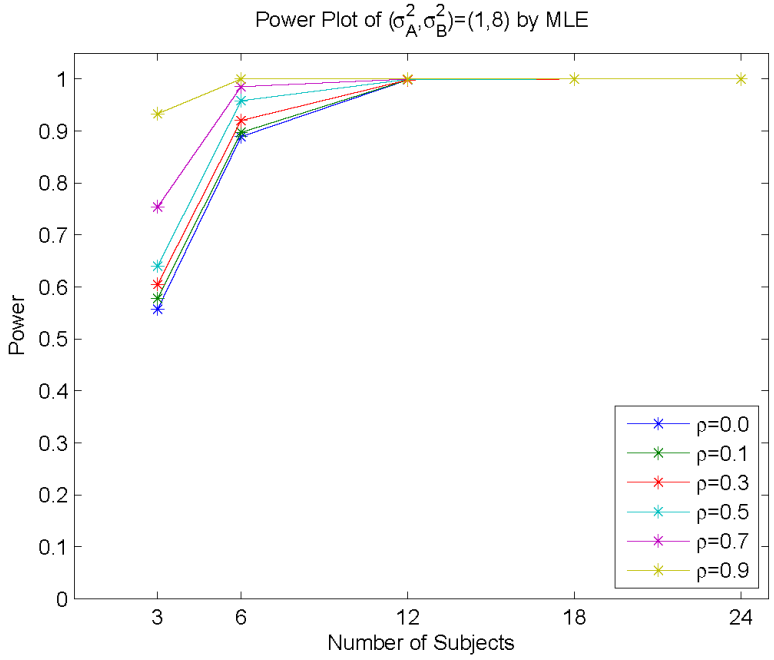

Figure A.4: *Power Plot at* (σ_A^2, σ_B^2)*=(1,8)*

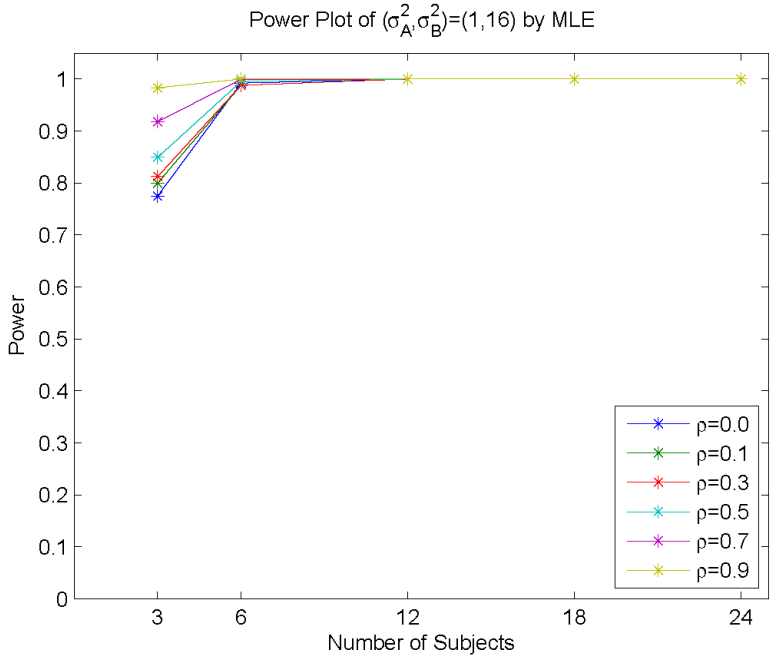

Figure A.5: *Power Plot at* (σ_A^2, σ_B^2)*=(1,16)*

Appendix B

Chapter 2: Tables for Two Treatment Crossover Design

Table B.1: *Type I Error* $\alpha = 0.05$ *at* (σ_A^2, σ_B^2)*=(1,1)*

		Number of Subject				
	ρ	3	6	12	18	24
MLE R	0	0.103	0.072	0.065	0.048	0.042
	0.1	0.110	0.072	0.068	0.050	0.052
	0.3	0.116	0.068	0.055	0.053	0.046
	0.5	0.112	0.062	0.060	0.052	0.053
	0.7	0.108	0.065	0.062	0.052	0.054
	0.9	0.121	0.053	0.055	0.041	0.044
REML SAS	0	0.103	0.072	0.065	0.048	0.042
	0.1	0.110	0.072	0.068	0.050	0.052
	0.3	0.116	0.068	0.055	0.053	0.046
	0.5	0.112	0.062	0.060	0.052	0.053
	0.7	0.108	0.065	0.062	0.052	0.054
	0.9	0.121	0.053	0.055	0.041	0.044

Table B.2: *Power at* (σ_A^2, σ_B^2)*=(1,2)*

		Number of Subject				
	ρ	3	6	12	18	24
MLE R	0	0.180	0.209	0.374	0.523	0.661
	0.1	0.176	0.200	0.386	0.538	0.629
	0.3	0.175	0.247	0.359	0.589	0.705
	0.5	0.192	0.275	0.465	0.631	0.796
	0.7	0.208	0.337	0.623	0.811	0.901
	0.9	0.378	0.636	0.935	0.988	0.998
REML SAS	0	0.180	0.209	0.374	0.523	0.661
	0.1	0.176	0.200	0.386	0.538	0.629
	0.3	0.175	0.247	0.359	0.589	0.705
	0.5	0.192	0.275	0.465	0.631	0.796
	0.7	0.208	0.337	0.623	0.811	0.901
	0.9	0.378	0.636	0.935	0.988	0.998

Table B.3: *Power at* (σ_A^2, σ_B^2)*=(1,4)*

		Number of Subject				
	ρ	3	6	12	18	24
MLE R	0	0.336	0.599	0.888	0.981	0.997
	0.1	0.347	0.584	0.916	0.982	0.998
	0.3	0.374	0.656	0.917	0.989	0.997
	0.5	0.376	0.691	0.950	0.995	0.999
	0.7	0.488	0.835	0.989	1.000	1.000
	0.9	0.758	0.986	1.000	1.000	1.000
REML SAS	0	0.336	0.599	0.888	0.981	0.997
	0.1	0.347	0.584	0.916	0.982	0.998
	0.3	0.374	0.656	0.917	0.989	0.997
	0.5	0.376	0.691	0.950	0.995	0.999
	0.7	0.488	0.835	0.989	1.000	1.000
	0.9	0.758	0.986	1.000	1.000	1.000

Table B.4: *Power at* (σ_A^2, σ_B^2)*=(1,8)*

		Number of Subject				
	ρ	3	6	12	18	24
MLE R	0	0.557	0.889	0.998	1.000	1.000
	0.1	0.578	0.897	0.998	1.000	1.000
	0.3	0.605	0.920	0.998	1.000	1.000
	0.5	0.640	0.958	0.999	1.000	1.000
	0.7	0.754	0.985	1.000	1.000	1.000
	0.9	0.933	1.000	1.000	1.000	1.000
REML SAS	0	0.557	0.889	0.998	1.000	1.000
	0.1	0.578	0.897	0.998	1.000	1.000
	0.3	0.605	0.920	0.998	1.000	1.000
	0.5	0.640	0.958	0.999	1.000	1.000
	0.7	0.754	0.985	1.000	1.000	1.000
	0.9	0.933	1.000	1.000	1.000	1.000

Table B.5: *Power at* (σ_A^2, σ_B^2)*=(1,16)*

		Number of Subject				
	ρ	3	6	12	18	24
MLE R	0	0.775	0.992	1.000	1.000	1.000
	0.1	0.800	0.988	1.000	1.000	1.000
	0.3	0.812	0.988	1.000	1.000	1.000
	0.5	0.849	0.995	1.000	1.000	1.000
	0.7	0.918	0.999	1.000	1.000	1.000
	0.9	0.983	1.000	1.000	1.000	1.000
REML SAS	0	0.775	0.992	1.000	1.000	1.000
	0.1	0.800	0.988	1.000	1.000	1.000
	0.3	0.812	0.988	1.000	1.000	1.000
	0.5	0.849	0.995	1.000	1.000	1.000
	0.7	0.918	0.999	1.000	1.000	1.000
	0.9	0.983	1.000	1.000	1.000	1.000

Appendix C

Chapter 3: Figures for Three Treatment Crossover Design

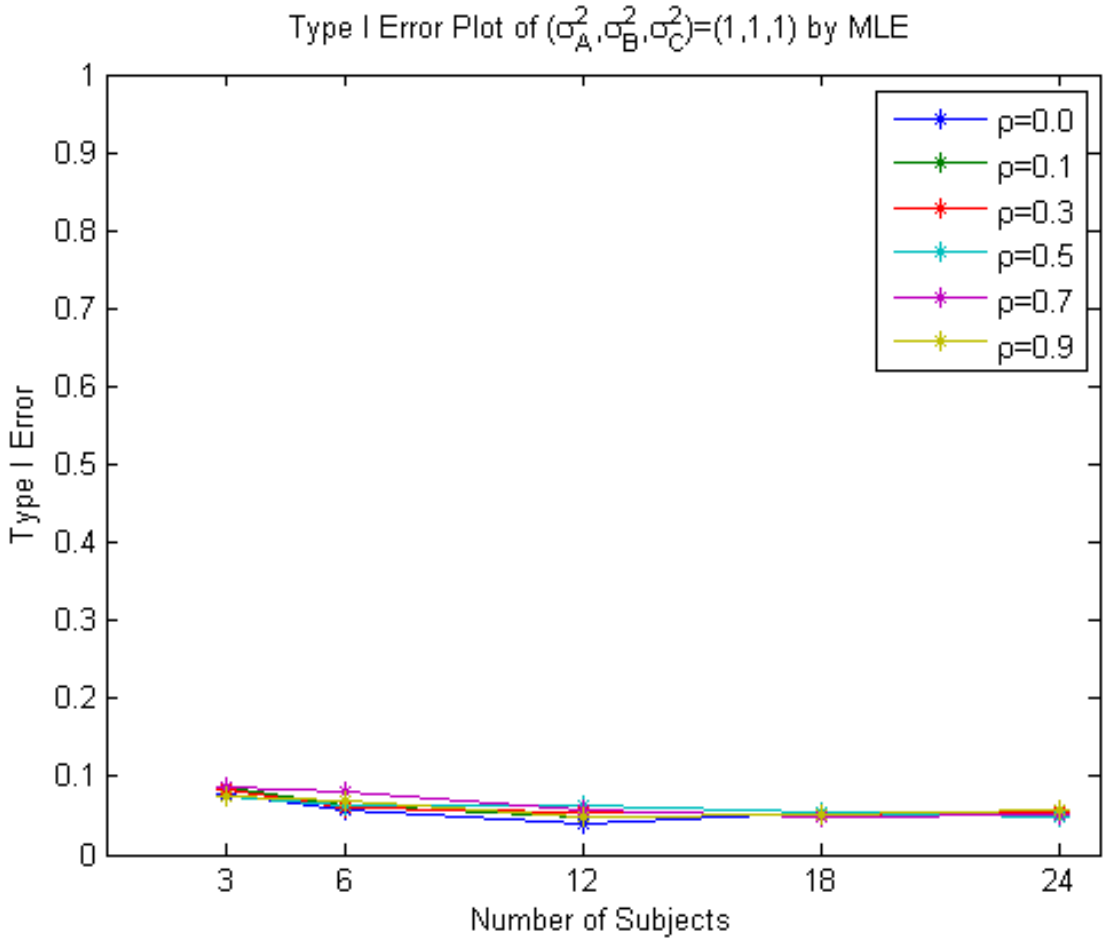

Figure C.1: *Type I Error Plot at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(1,1,1)*

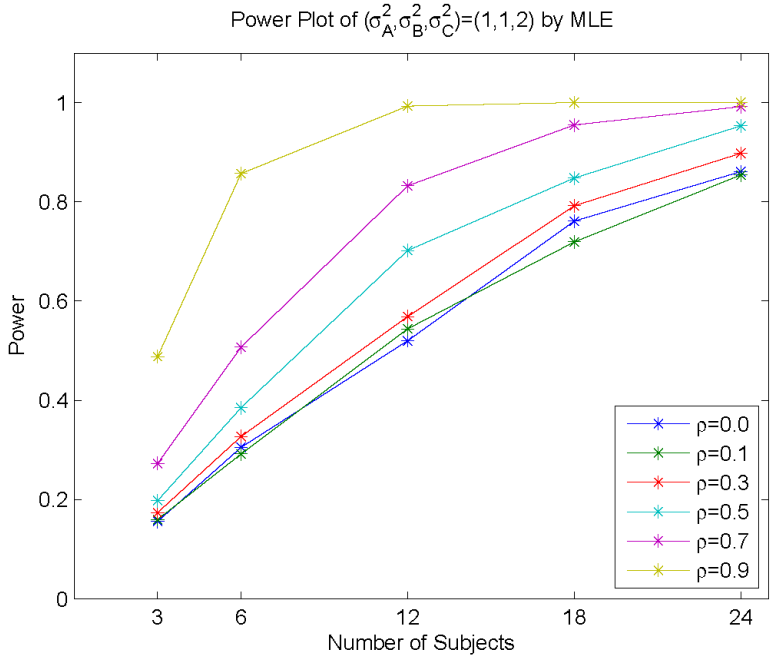

Figure C.2: *Power Plot at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(1,1,2)*

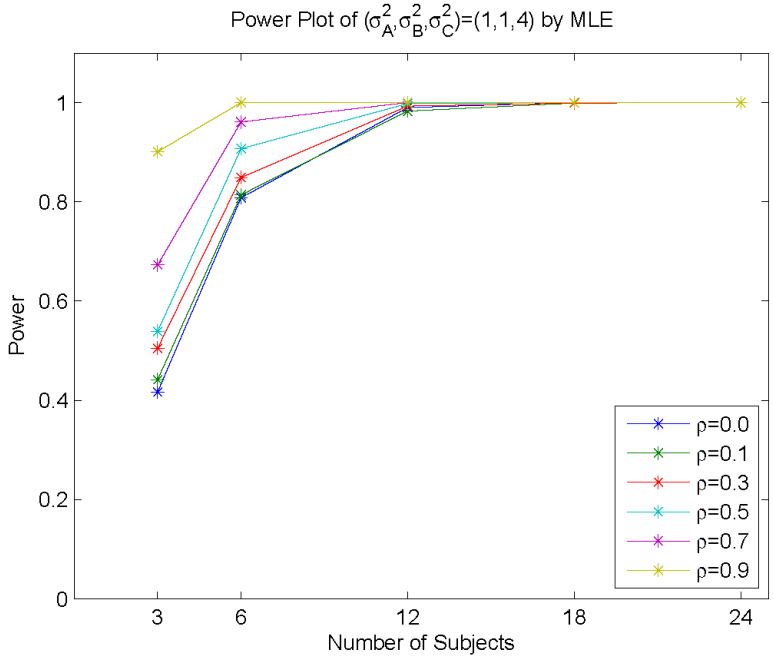

Figure C.3: *Power Plot at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(1,1,4)*

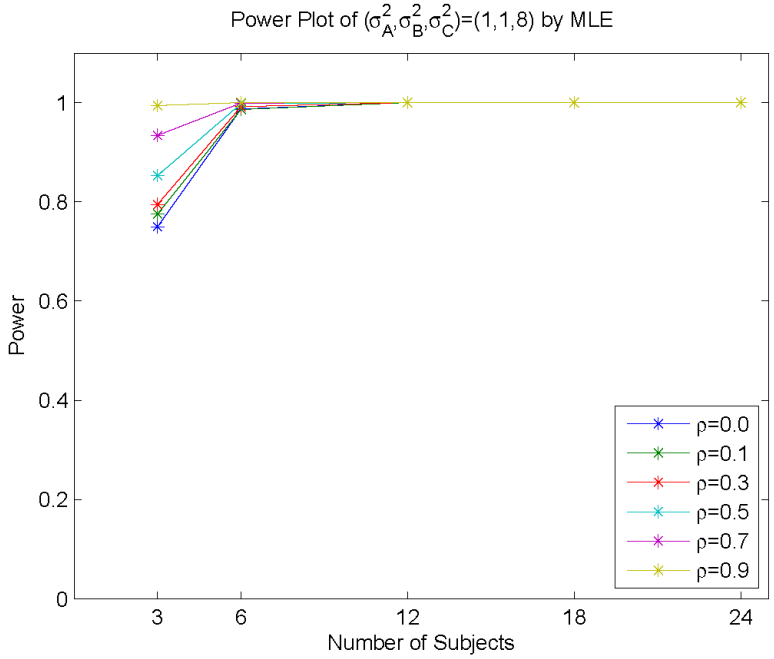

Figure C.4: *Power Plot at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(1,1,8)*

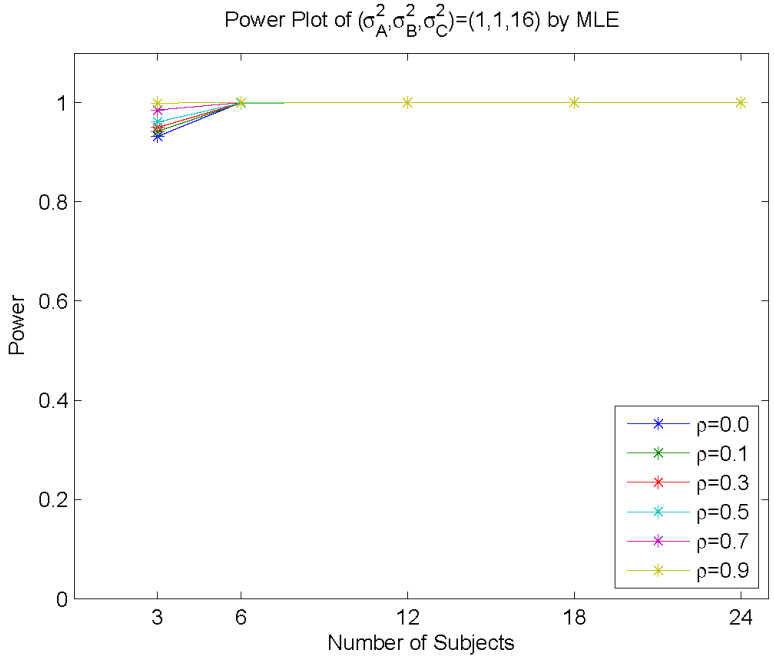

Figure C.5: *Power Plot at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(1,1,16)*

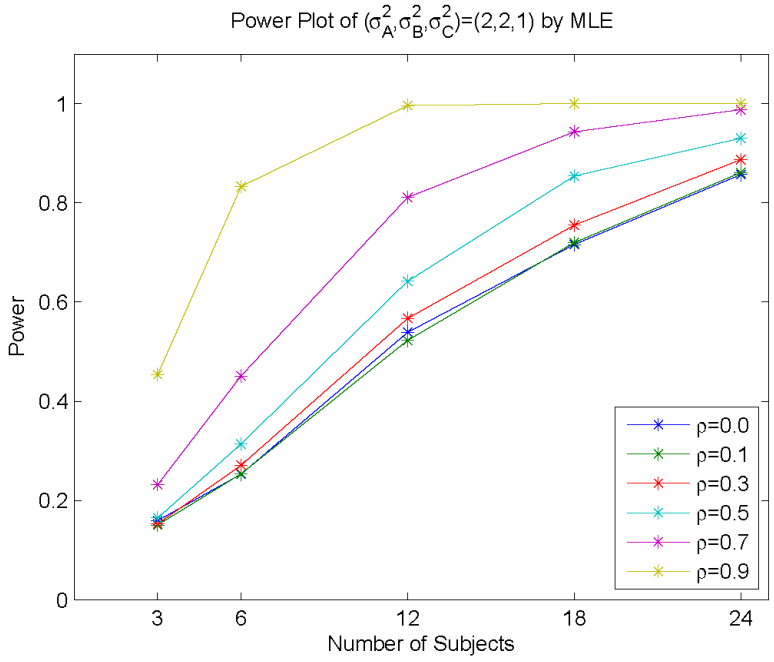

Figure C.6: *Power Plot at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(2,2,1)*

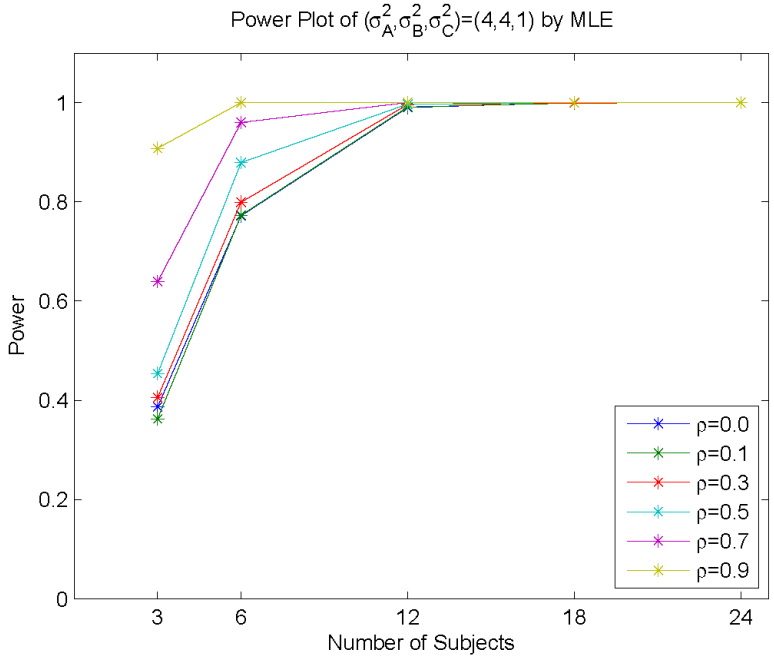

Figure C.7: *Power Plot at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(4,4,1)*

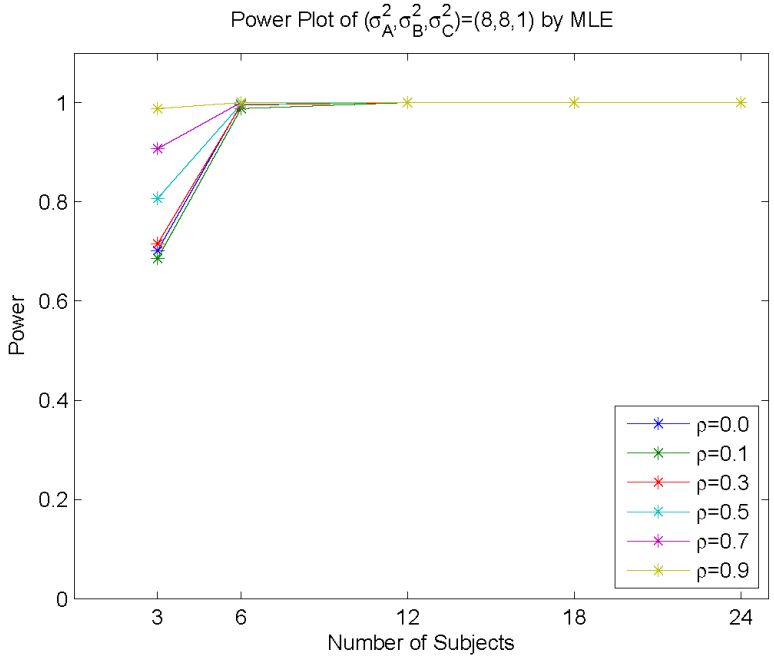

Figure C.8: *Power Plot at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(8,8,1)*

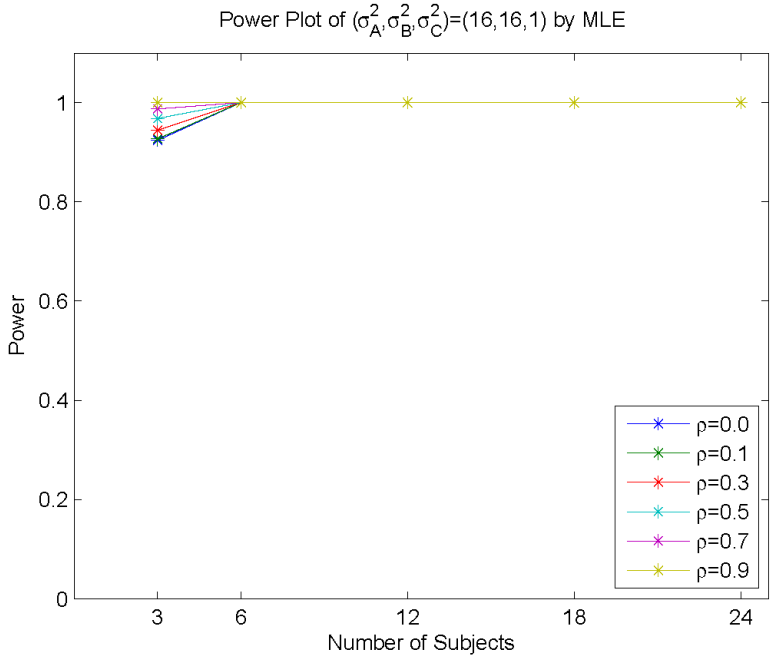

Figure C.9: *Power Plot at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(16,16,1)*

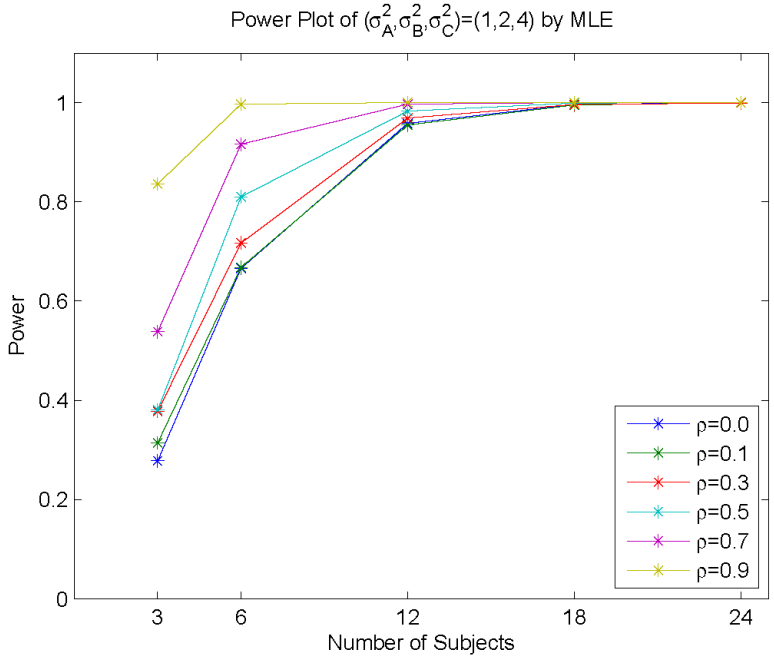

Figure C.10: *Power Plot at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(1,2,4)*

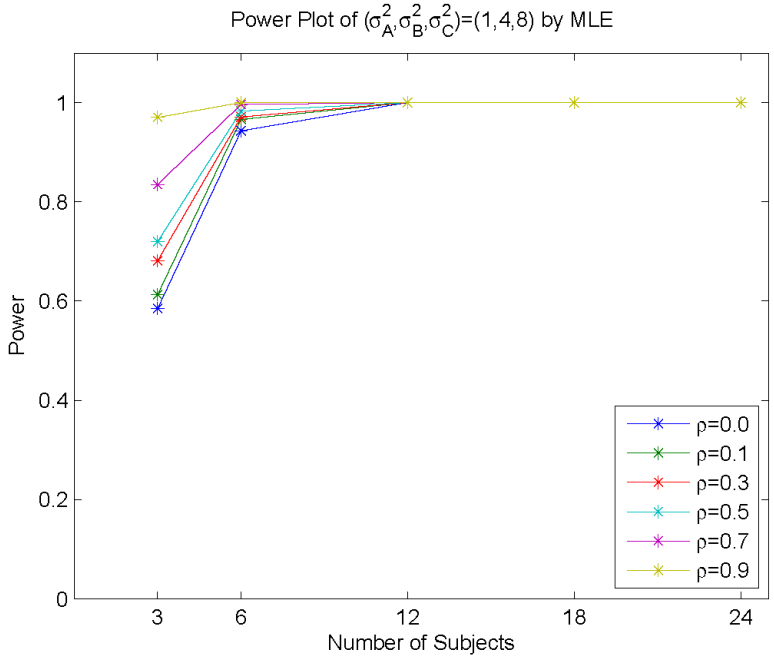

Figure C.11: *Power Plot at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(1,4,8)*

Appendix D

Chapter 3: Tables for Three Treatment Crossover Design

Table D.1: *Type I Error* $\alpha = 0.05$ *at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(1,1,1)*

		Number of Subject				
	ρ	3	6	12	18	24
MLE R	0	0.077	0.058	0.041	0.054	0.047
	0.1	0.087	0.062	0.048	0.052	0.057
	0.3	0.084	0.061	0.055	0.051	0.054
	0.5	0.075	0.064	0.064	0.053	0.049
	0.7	0.086	0.080	0.056	0.048	0.051
	0.9	0.074	0.068	0.047	0.052	0.057
REML SAS	0	0.078	0.058	0.041	0.054	0.047
	0.1	0.086	0.062	0.048	0.052	0.057
	0.3	0.083	0.061	0.055	0.051	0.054
	0.5	0.076	0.064	0.064	0.053	0.049
	0.7	0.088	0.080	0.056	0.048	0.051
	0.9	0.074	0.068	0.047	0.052	0.057

Table D.2: *Power at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(1,1,2)*

		Number of Subject				
	ρ	3	6	12	18	24
MLE R	0	0.156	0.305	0.520	0.761	0.861
	0.1	0.160	0.292	0.544	0.719	0.854
	0.3	0.174	0.327	0.569	0.792	0.898
	0.5	0.198	0.385	0.702	0.848	0.953
	0.7	0.273	0.507	0.832	0.955	0.992
	0.9	0.488	0.857	0.993	1.000	1.000
REML SAS	0	0.151	0.306	0.522	0.761	0.861
	0.1	0.162	0.291	0.545	0.719	0.854
	0.3	0.180	0.327	0.568	0.792	0.899
	0.5	0.199	0.384	0.702	0.847	0.953
	0.7	0.272	0.506	0.832	0.955	0.992
	0.9	0.490	0.857	0.993	1.000	1.000

Table D.3: *Power at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(1,1,4)*

		Number of Subject				
	ρ	3	6	12	18	24
MLE R	0	0.417	0.808	0.990	1.000	1.000
	0.1	0.442	0.814	0.983	0.999	1.000
	0.3	0.505	0.849	0.993	0.999	1.000
	0.5	0.539	0.907	0.998	1.000	1.000
	0.7	0.673	0.961	1.000	1.000	1.000
	0.9	0.901	1.000	1.000	1.000	1.000
REML SAS	0	0.396	0.807	0.990	1.000	1.000
	0.1	0.440	0.815	0.983	0.999	1.000
	0.3	0.509	0.852	0.993	0.999	1.000
	0.5	0.541	0.906	0.998	1.000	1.000
	0.7	0.674	0.961	1.000	1.000	1.000
	0.9	0.905	1.000	1.000	1.000	1.000

Table D.4: *Power at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(1,1,8)*

		Number of Subject				
	ρ	3	6	12	18	24
MLE R	0	0.750	0.987	1.000	1.000	1.000
	0.1	0.776	0.986	1.000	1.000	1.000
	0.3	0.795	0.992	1.000	1.000	1.000
	0.5	0.853	0.998	1.000	1.000	1.000
	0.7	0.934	0.999	1.000	1.000	1.000
	0.9	0.994	1.000	1.000	1.000	1.000
REML SAS	0	0.724	0.987	1.000	1.000	1.000
	0.1	0.763	0.986	1.000	1.000	1.000
	0.3	0.791	0.992	1.000	1.000	1.000
	0.5	0.851	0.998	1.000	1.000	1.000
	0.7	0.936	0.998	1.000	1.000	1.000
	0.9	0.994	1.000	1.000	1.000	1.000

Table D.5: *Power at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(1,1,16)*

		Number of Subject				
	ρ	3	6	12	18	24
MLE R	0	0.932	1.000	1.000	1.000	1.000
	0.1	0.942	0.999	1.000	1.000	1.000
	0.3	0.950	0.999	1.000	1.000	1.000
	0.5	0.961	0.999	1.000	1.000	1.000
	0.7	0.985	1.000	1.000	1.000	1.000
	0.9	0.999	1.000	1.000	1.000	1.000
REML SAS	0	0.883	1.000	1.000	1.000	1.000
	0.1	0.903	0.999	1.000	1.000	1.000
	0.3	0.945	0.999	1.000	1.000	1.000
	0.5	0.959	0.999	1.000	1.000	1.000
	0.7	0.983	1.000	1.000	1.000	1.000
	0.9	0.999	1.000	1.000	1.000	1.000

Table D.6: *Power at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(2,2,1)*

		Number of Subject				
	ρ	3	6	12	18	24
MLE R	0	0.160	0.253	0.539	0.716	0.857
	0.1	0.151	0.254	0.523	0.720	0.861
	0.3	0.153	0.271	0.567	0.755	0.887
	0.5	0.165	0.314	0.642	0.854	0.930
	0.7	0.232	0.451	0.811	0.943	0.988
	0.9	0.454	0.833	0.996	1.000	1.000
REML SAS	0	0.158	0.253	0.539	0.717	0.857
	0.1	0.148	0.254	0.525	0.720	0.862
	0.3	0.155	0.273	0.568	0.755	0.887
	0.5	0.168	0.314	0.644	0.854	0.930
	0.7	0.232	0.451	0.811	0.943	0.988
	0.9	0.456	0.834	0.996	1.000	1.000

Table D.7: *Power at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(4,4,1)*

		Number of Subject				
	ρ	3	6	12	18	24
MLE R	0	0.387	0.772	0.990	1.000	1.000
	0.1	0.363	0.773	0.991	0.999	1.000
	0.3	0.406	0.799	0.996	0.999	1.000
	0.5	0.454	0.879	0.997	1.000	1.000
	0.7	0.639	0.960	1.000	1.000	1.000
	0.9	0.908	1.000	1.000	1.000	1.000
REML SAS	0	0.367	0.775	0.990	1.000	1.000
	0.1	0.353	0.777	0.991	0.999	1.000
	0.3	0.407	0.801	0.996	0.999	1.000
	0.5	0.460	0.880	0.997	1.000	1.000
	0.7	0.640	0.960	1.000	1.000	1.000
	0.9	0.908	1.000	1.000	1.000	1.000

Table D.8: *Power at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(8,8,1)*

		Number of Subject				
	ρ	3	6	12	18	24
MLE R	0	0.702	0.997	1.000	1.000	1.000
	0.1	0.686	0.988	1.000	1.000	1.000
	0.3	0.716	0.995	1.000	1.000	1.000
	0.5	0.807	0.997	1.000	1.000	1.000
	0.7	0.907	1.000	1.000	1.000	1.000
	0.9	0.988	1.000	1.000	1.000	1.000
REML SAS	0	0.668	0.997	0.999	1.000	1.000
	0.1	0.663	0.989	1.000	1.000	1.000
	0.3	0.713	0.995	1.000	1.000	1.000
	0.5	0.810	0.997	1.000	1.000	1.000
	0.7	0.908	1.000	1.000	1.000	1.000
	0.9	0.988	1.000	1.000	1.000	1.000

Table D.9: *Power at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(16,16,1)*

		Number of Subject				
	ρ	3	6	12	18	24
MLE R	0	0.924	1.000	1.000	1.000	1.000
	0.1	0.927	1.000	1.000	1.000	1.000
	0.3	0.944	1.000	1.000	1.000	1.000
	0.5	0.968	1.000	1.000	1.000	1.000
	0.7	0.987	1.000	1.000	1.000	1.000
	0.9	1.000	1.000	1.000	1.000	1.000
REML SAS	0	0.859	1.000	1.000	1.000	1.000
	0.1	0.890	1.000	1.000	1.000	1.000
	0.3	0.936	1.000	1.000	1.000	1.000
	0.5	0.969	1.000	1.000	1.000	1.000
	0.7	0.988	1.000	1.000	1.000	1.000
	0.9	1.000	1.000	1.000	1.000	1.000

Table D.10: *Power at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(1,2,4)*

		Number of Subject				
	ρ	3	6	12	18	24
MLE R	0	0.278	0.666	0.958	0.996	1.000
	0.1	0.314	0.668	0.955	0.995	1.000
	0.3	0.378	0.717	0.969	0.996	0.999
	0.5	0.381	0.810	0.983	0.999	1.000
	0.7	0.538	0.916	0.997	1.000	1.000
	0.9	0.836	0.997	1.000	1.000	1.000
REML SAS	0	0.262	0.665	0.958	0.996	1.000
	0.1	0.314	0.668	0.955	0.995	1.000
	0.3	0.377	0.722	0.969	0.996	0.999
	0.5	0.383	0.810	0.983	0.999	1.000
	0.7	0.539	0.917	0.997	1.000	1.000
	0.9	0.838	0.997	1.000	1.000	1.000

Table D.11: *Power at* $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$*=(1,4,8)*

		Number of Subject				
	ρ	3	6	12	18	24
MLE R	0	0.585	0.943	1.000	1.000	1.000
	0.1	0.613	0.966	1.000	1.000	1.000
	0.3	0.681	0.971	1.000	1.000	1.000
	0.5	0.720	0.983	1.000	1.000	1.000
	0.7	0.835	0.996	1.000	1.000	1.000
	0.9	0.970	1.000	1.000	1.000	1.000
REML SAS	0	0.553	0.943	1.000	1.000	1.000
	0.1	0.591	0.966	1.000	1.000	1.000
	0.3	0.674	0.972	1.000	1.000	1.000
	0.5	0.721	0.983	1.000	1.000	1.000
	0.7	0.834	0.996	1.000	1.000	1.000
	0.9	0.971	1.000	1.000	1.000	1.000

Appendix E

R Code

```
#=========================================================================#
#       Three Treatment/ Three Sequence Crossover Data (ABC/BCA/CAB)      #
#                  With Compound Symmetry Covariance Structure            #
#=========================================================================#

#-------------------------------------------------------------------------#
#                     Definition of Initialized values                    #
#-------------------------------------------------------------------------#
#   nsubj  : number of subject at each sequence(3,6,12,18,24)             #
#   nseq : number of sequence                                             #
#   np     : number of period                                             #
#   N : total number of subject used in 3x3 crossover design              #
#   rho : values of rho # W.size : iteration size                         #
#-------------------------------------------------------------------------#

nsubj <- 24; nseq <- 3; np <- 3; N <- nseq*nsubj;
R <- 10000; muA <- 0;
rho <- c(0,0.1,0.3,0.5,0.7,0.9)
W.size <- 1000

#*************************************************************************#
#                                START                                    #
#       Step to obtain W after generating matrix W per each sequence      #
#       with fixed rho value among 0, 0,1, 0.5, 0.7                       #
#       fixed carryover value lambda A=0, lambda B=0 , lambda C=0         #
#*************************************************************************#

#-------------------------------------------------------------------------#
#  Generate W1, W2, W3 With Sigma A = 1, Sigma B = 1, Sigma C = sqrt(2),  #
#                           lambda A = 0, lambda B = 0 , lambda C =0      #
#-------------------------------------------------------------------------#
```

```
# Part to generate Matrices of W, W1, W2, and W3

y.data <- matrix(0,length(rho)*length(muA)*W.size*nseq*nsubj,11)

W.mu <- array(0,c(3,3,W.size,length(muA),length(rho)))

W1.mu <- array(0,c(3,3,W.size,length(muA),length(rho)))

W2.mu <- array(0,c(3,3,W.size,length(muA),length(rho)))

W3.mu <- array(0,c(3,3,W.size,length(muA),length(rho)))

mu.mat <- matrix(0,length(rho)*length(muA)*W.size,6)

diff.AB <- matrix(0,length(rho)*length(muA)*W.size,4)

# Define Permutation matrices Ci, i=1,2,3,4,5,6

c <- array(0,c(3,3,3))

c[,,1] <- matrix(c(1,0,0,0,1,0,0,0,1),3,3)

c[,,2] <- matrix(c(0,0,1,1,0,0,0,1,0),3,3)

c[,,3] <- matrix(c(0,1,0,0,0,1,1,0,0),3,3)

set.seed(9702467)

for(jj in 1:length(rho)){
  for(k in 1:length(muA)){

    mu.a <- muA[k]; mu.b <- 0; mu.c <- 0
    lambda.a <- 0; lambda.b <- 0; lambda.c <- 0
    sigma.a <- 1; sigma.b <- 1; sigma.c <- sqrt(2)
    rho1 <- rho[jj]; rho2 <- rho1^2

    for(kk in 1:W.size){
    seq1 <- matrix(0,nsubj,np)
    seq2 <- matrix(0,nsubj,np)
    seq3 <- matrix(0,nsubj,np)

    # Generate values of y_ij at each sequences
    # seq1[i,j] means that ith subject value for jth treatment(A,B,or C)
    # in sequence 1
```

```
for(i in 1:nseq){
  for(j in 1:nsubj){

     x1 <- rnorm(1,mean=0,sd=1)
     x2 <- rnorm(1,mean=0,sd=1)
     x3 <- rnorm(1,mean=0,sd=1)

    # Sequence ABC
      if (i==1){
        seq1[j,1]<-sigma.a*x1+mu.a
        seq1[j,2]<-rho1*sigma.b*x1+sigma.b*sqrt(1-rho2)*x2+mu.b+lambda.a
        seq1[j,3]<-rho1*sigma.c*x1+rho1*sigma.c*sqrt((1-rho1)/(1+rho1))*x2+
                sigma.c*sqrt((1+rho1-2*rho2)/(1+rho1))*x3+mu.c+lambda.b

    y.data[(jj-1)*length(muA)*W.size*nseq*nsubj+(k-1)*W.size*nseq*nsubj
          +(kk-1)*nseq*nsubj+(i-1)*nsubj+j,] <- c(rho[jj],sigma.a^2,
          sigma.b^2,sigma.c^2,nsubj,kk,i,j,seq1[j,1],seq1[j,2],seq1[j,3])

    }
    # Sequence BCA
    if (i==2){
        seq2[j,1]<-sigma.b*x1+mu.b
        seq2[j,2]<-rho1*sigma.c*x1+sigma.c*sqrt(1-rho2)*x2+mu.c+lambda.b
        seq2[j,3]<-rho1*sigma.a*x1+rho1*sigma.a*sqrt((1-rho1)/(1+rho1))*x2+
                sigma.a*sqrt((1+rho1-2*rho2)/(1+rho1))*x3+mu.a+lambda.c

    y.data[(jj-1)*length(muA)*W.size*nseq*nsubj+(k-1)*W.size*nseq*nsubj
          +(kk-1)*nseq*nsubj+(i-1)*nsubj+j,] <- c(rho[jj],sigma.a^2,
          sigma.b^2,sigma.c^2,nsubj,kk,i,j,seq2[j,1],seq2[j,2],seq2[j,3])

    }
    # Sequence CAB
    if (i==3){
        seq3[j,1]<-sigma.c*x1+mu.c
        seq3[j,2]<-rho1*sigma.a*x1+sigma.a*sqrt(1-rho2)*x2+mu.a+lambda.c
        seq3[j,3]<-rho1*sigma.b*x1+rho1*sigma.b*sqrt((1-rho1)/(1+rho1))*x2+
                sigma.b*sqrt((1+rho1-2*rho2)/(1+rho1))*x3+mu.b+lambda.a

    y.data[(jj-1)*length(muA)*W.size*nseq*nsubj+(k-1)*W.size*nseq*nsubj
          +(kk-1)*nseq*nsubj+(i-1)*nsubj+j,] <- c(rho[jj],sigma.a^2,
          sigma.b^2,sigma.c^2,nsubj,kk,i,j,seq3[j,1],seq3[j,2],seq3[j,3])

    }
  }
}
```

```
# Calculation for (each value - cell mean)
for (i in 1:3){
    seq1[,i]<-seq1[,i]-mean(seq1[,i])
    seq2[,i]<-seq2[,i]-mean(seq2[,i])
    seq3[,i]<-seq3[,i]-mean(seq3[,i])
}

seq1
seq2
seq3


# Calculation for elements of W1, W2, W3, and W matrix

w1.11 <- sum(seq1[,1]^2)
w1.22 <- sum(seq1[,2]^2)
w1.33 <- sum(seq1[,3]^2)
w1.12 <- sum(seq1[,1]*seq1[,2])
w1.13 <- sum(seq1[,1]*seq1[,3])
w1.23 <- sum(seq1[,2]*seq1[,3])

w2.11 <- sum(seq2[,1]^2)
w2.22 <- sum(seq2[,2]^2)
w2.33 <- sum(seq2[,3]^2)
w2.12 <- sum(seq2[,1]*seq2[,2])
w2.13 <- sum(seq2[,1]*seq2[,3])
w2.23 <- sum(seq2[,2]*seq2[,3])

w3.11 <- sum(seq3[,1]^2)
w3.22 <- sum(seq3[,2]^2)
w3.33 <- sum(seq3[,3]^2)
w3.12 <- sum(seq3[,1]*seq3[,2])
w3.13 <- sum(seq3[,1]*seq3[,3])
w3.23 <- sum(seq3[,2]*seq3[,3])


w1 <- c(w1.11,w1.12,w1.13,w1.12,w1.22,w1.23,w1.13,w1.23,w1.33)
w2 <- c(w2.11,w2.12,w2.13,w2.12,w2.22,w2.23,w2.13,w2.23,w2.33)
w3 <- c(w3.11,w3.12,w3.13,w3.12,w3.22,w3.23,w3.13,w3.23,w3.33)

W1.mu[,,kk,k,jj] <- matrix(w1,3,3)
W2.mu[,,kk,k,jj] <- matrix(w2,3,3)
W3.mu[,,kk,k,jj] <- matrix(w3,3,3)
```

```
    # Case of 3 Sequences

    temp.1 <- t(c[,,1])%*%W1.mu[,,kk,k,jj]%*%c[,,1]
    temp.2 <- t(c[,,2])%*%W2.mu[,,kk,k,jj]%*%c[,,2]
    temp.3 <- t(c[,,3])%*%W3.mu[,,kk,k,jj]%*%c[,,3]

    W.mu[,,kk,k,jj] <- temp.1+temp.2+temp.3

    }
  }
}

y.data <- data.frame(y.data)

colnames(y.data) <-
c("rho","sigmaA^2","sigmaB^2","sigmaC^2","nsubj","Sim","Seq","subj","y1","y2","y3")

write.table(y.data, file = "y24-112data.txt", sep = " ", row.names =
F,qmethod = "double",quote=FALSE)

W.mean <- array(0,c(3,3,length(muA),length(rho)))

for(i in 1:length(rho)){
  for(j in 1:length(muA)){

    W.temp <- matrix(0,3,3)

    for(k in 1:W.size){
        W.temp <- W.temp + W.mu[,,k,j,i]
    }

    W.mean[,,j,i] <- W.temp/W.size

  }
}

#*****************************************************************************#
#*****************************************************************************#
#                                                                             #
#              Part of parameter estimation to make Sigma matrix              #
#                                                                             #
#*****************************************************************************#
#*****************************************************************************#
```

```
#=============================================================================#
#              By polyroot Function in package base in R                      #
#                                                                             #
#              Note:  Find zeros of a real or complex polynomial.             #
#=============================================================================#

#-----------------------------------------------------------------------------#
#          Unrestricted Condition Case (Ha): Part of Denominator in LRT       #
#-----------------------------------------------------------------------------#
#-----------------------------------------------------------------------------#
#                      Part to estimate Parameters in Sigma                   #
#-----------------------------------------------------------------------------#

library(base)

errbnd <- 0.00001

#====================== Define functions of each parameter ====================#
ur.funA <- function(N,sB,sC,sR){
    a3ur <- N-3
    a2ur <- (sR/((1-sR)*(1+2*sR)))*(w12/sB + w13/sC)
    a1ur <- -((1+sR)*w11)/((1-sR)*(1+2*sR))

    sigA.ur <- 0

    sigA.ur <- polyroot(c(a1ur,a2ur,a3ur))
    sigA.ur
}

ur.funB <- function(N,sA,sC,sR){
    b3ur <- N-3
    b2ur <- (sR/((1-sR)*(1+2*sR)))*(w12/sA + w23/sC)
    b1ur <- -((1+sR)*w22)/((1-sR)*(1+2*sR))

    sigB.ur <- 0

    sigB.ur <- polyroot(c(b1ur,b2ur,b3ur))
    sigB.ur
}

ur.funC <- function(N,sA,sB,sR){
    c3ur <- N-3
    c2ur <- (sR/((1-sR)*(1+2*sR)))*(w13/sA + w23/sB)
    c1ur <- -((1+sR)*w33)/((1-sR)*(1+2*sR))
```

```
    sigC.ur <- 0

    sigC.ur <- polyroot(c(c1ur,c2ur,c3ur))
    sigC.ur
}

ur.funR <- function(N,sA0,sB0,sC0){

    sA <- sA0; sB <- sB0; sC <- sC0

    r4ur <- -6*(N-3)
    r3ur <- 3*(N-3) - (w11/sA^2 + w22/sB^2 + w33/sC^2)
            + 2*(w12/(sA*sB) + w13/(sA*sC) + w23/(sB*sC))
    r2ur <- 3*(N-3) - 2*(w11/sA^2 + w22/sB^2 + w33/sC^2)
    r1ur <- (w12/(sA*sB) + w13/(sA*sC) + w23/(sB*sC))

    rho.ur <-0

    temp1 <- Re(polyroot(c(r1ur,r2ur,r3ur,r4ur)))

    if(abs(temp1[1]-temp1[2])< 0.0000000001) rho.ur <- temp1[3]
    if(abs(temp1[1]-temp1[3])< 0.0000000001) rho.ur <- temp1[2]
    if(abs(temp1[2]-temp1[3])< 0.0000000001) rho.ur <- temp1[1]

    rho.ur
}


logL <- function(N,c,W,sA0,sB0,sC0,sR0){

  f <-  log(c)+(N-7)*log(det(W))/2 -(N-3)/2*(log(sA0^2)+log(sB0^2)
    +log(sC0^2)+log((1+2*sR0)*(1-sR0)^2))
    -(1+sR0)/(2*(1-sR0)*(1+2*sR0))*(W[1,1]/sA0^2 +W[2,2]/sB0^2 +W[3,3]/sC0^2)
    +sR0/((1-sR0)*(1+2*sR0))*(W[1,2]/(sA0*sB0)+W[1,3]/(sA0*sC0)+W[2,3]/(sB0*sC0))
  f
}

#============================== Procedure 2 ==============================#


ini.ur2  <- matrix(0,length(rho)*length(muA)*W.size,7)

est.ur2  <- matrix(0,length(rho)*length(muA)*W.size,7)

ini.val2  <- rep(0,7); err.val2  <- rep(0,7); logL.val2 <- rep(0,4)
```

```
Con2<-2^(3*(N-3)/2)*pi^(3/2)*gamma((N-3)/2)*gamma((N-4)/2)*gamma((N-5)/2)

# Start - FOR statement for rho for(i in 1:length(rho)){

 # Start - FOR statement for muA
 for(j in 1:length(muA)){

  # Start - FOR statement for W.size
  for(k in 1:W.size){

    w11 <- W.mu[1,1,k,j,i]; w22 <- W.mu[2,2,k,j,i]; w33 <- W.mu[3,3,k,j,i]
    w12 <- W.mu[1,2,k,j,i]; w13 <- W.mu[1,3,k,j,i]; w23 <- W.mu[2,3,k,j,i]

    sA0 <- sqrt(w11/(N-3)); sB0 <- sqrt(w22/(N-3)); sC0 <- sqrt(w33/(N-3))
    sR0 <- (w12/sqrt(w11*w22) + w13/sqrt(w11*w33) + w23/sqrt(w22*w33))/3
    ini.ur2[(i-1)*length(muA)*W.size+(j-1)*W.size+k,]
            <- c(rho[i],muA[j],k,sA0,sB0,sC0,sR0)

    logL1 <- c(rho[i],muA[j],k,logL(N,1/Con2,W.mu[,,k,j,i],sA0,sB0,sC0,sR0))

    sigA <- c(rho[i],muA[j],k,sA0)
    sigB <- c(rho[i],muA[j],k,sB0)
    sigC <- c(rho[i],muA[j],k,sC0)
    sigR <- c(rho[i],muA[j],k,sR0)

    #err2.ur1 <- c(rho[i],muA[j],k,0)
    err2.ur1 <- c(rho[i],muA[j],k,0,0,0,0)
    ini2.ur1 <- c(rho[i],muA[j],k,sA0,sB0,sC0,sR0)

    # Start - REPEAT statement
    repeat{

    sA1 <- Re(ur.funA(N,sB0,sC0,sR0))[1]
    sigA <- rbind(sigA,c(rho[i],muA[j],k,sA1))

    sB1 <- Re(ur.funB(N,sA1,sC0,sR0))[1]
    sigB <- rbind(sigB,c(rho[i],muA[j],k,sB1))

    sC1 <- Re(ur.funC(N,sA1,sB1,sR0))[1]
    sigC <- rbind(sigC,c(rho[i],muA[j],k,sC1))

    sR1 <- ur.funR(N,sA1,sB1,sC1)
    sigR <- rbind(sigR,c(rho[i],muA[j],k,sR1))
```

```
    logL0 <- c(rho[i],muA[j],k,logL(N,1/Con2,W.mu[,,k,j,i],sA1,sB1,sC1,sR1))
    logL1 <- rbind(logL1,logL0)

    ini2.ur1 <- rbind(ini2.ur1,c(rho[i],muA[j],k,sA1,sB1,sC1,sR1))

    e1 <- abs(sA1-sA0)
    e2 <- abs(sB1-sB0)
    e3 <- abs(sC1-sC0)
    e4 <- abs(sR1-sR0)

    err2.ur0 <- c(rho[i],muA[j],k,e1,e2,e3,e4)
    err2.ur1 <- rbind(err2.ur1,err2.ur0)

    # Start - IF statement
    if((e1 <= errbnd) && (e2 <= errbnd) && (e3 <= errbnd) && (e4 <= errbnd)){
       est.ur2[(i-1)*length(muA)*W.size+(j-1)*W.size+k,]
              <- c(rho[i],muA[j],k,sA1,sB1,sC1,sR1)
       break
    }
    # End - IF statement

    sA0 <- sA1
    sB0 <- sB1
    sC0 <- sC1
    sR0 <- sR1
    }
    # End - REPEAT statement

    ini.val2  <- rbind(ini.val2,ini2.ur1)
    err.val2  <- rbind(err.val2,err2.ur1)
    logL.val2 <- rbind(logL.val2,logL1)

  }
  # End - FOR statement for W.size

 }
 # End - FOR statement for muA

} # End - FOR statement for rho


est.ur2 <- data.frame(est.ur2)

colnames(est.ur2)<-c("rho","muA","Rep","Est-sigma A","Est-sigma
B","Est-sigma C","Est-rho")
```

```
est.ur2.mean <- matrix(0,length(rho)*length(muA),6)

for(i in 1:length(rho)){
  for(j in 1:length(muA)){
    t.4 <- est.ur2[(((i-1)*length(muA)*W.size+(j-1)*W.size+1):
                    ((i-1)*length(muA)*W.size+(j-1)*W.size+W.size)),4]
    t.5 <- est.ur2[(((i-1)*length(muA)*W.size+(j-1)*W.size+1):
                    ((i-1)*length(muA)*W.size+(j-1)*W.size+W.size)),5]
    t.6 <- est.ur2[(((i-1)*length(muA)*W.size+(j-1)*W.size+1):
                    ((i-1)*length(muA)*W.size+(j-1)*W.size+W.size)),6]
    t.7 <- est.ur2[(((i-1)*length(muA)*W.size+(j-1)*W.size+1):
                    ((i-1)*length(muA)*W.size+(j-1)*W.size+W.size)),7]

    t1<-sqrt(sum(t.4^2)/W.size)
    t2<-sqrt(sum(t.5^2)/W.size)
    t3<-sqrt(sum(t.6^2)/W.size)
    t4<-mean(t.7)

    est.ur2.mean[(i-1)*length(muA)+j,] <-
                 c(Re(rho[i]),Re(muA[j]),Re(t1),Re(t2),Re(t3),Re(t4))

  }
}

est.ur2.mean <- data.frame(est.ur2.mean)

colnames(est.ur2.mean) <- c(" rho "," muA ","Est-Sigma A","Est-Sigma
B","Est-Sigma C","Est-Rho")


#-------------------------------------------------------------------------------#
#           Restricted Condition Case (H0): Part of Numerator in LRT             #
#-------------------------------------------------------------------------------#
#-------------------------------------------------------------------------------#
#                     Part to estimate Parameters in Sigma                      #
#-------------------------------------------------------------------------------#


ini.r  <- matrix(0,length(rho)*length(muA)*W.size,5)

sig.r <- matrix(0,length(muA)*length(rho)*W.size,4)

rho.r <- matrix(0,length(muA)*length(rho)*W.size,4)
```

```
temp.r <- matrix(0,length(muA)*length(rho)*W.size,3)


for(i in 1:length(rho)){
  for(j in 1:length(muA)){
    for(k in 1:W.size){

    w11 <- W.mu[1,1,k,j,i]; w22 <- W.mu[2,2,k,j,i]; w33 <- W.mu[3,3,k,j,i]
    w12 <- W.mu[1,2,k,j,i]; w13 <- W.mu[1,3,k,j,i]; w23 <- W.mu[2,3,k,j,i]

    # sA and sR are Shanga's MLE of parameter sigma A and rho
    # under the restricted condition

    sA <- (w11+w22+w33)/(3*(N-3))
    sR <- (w12+w13+w23)/(w11+w22+w33)
    ini.r[(i-1)*length(muA)*W.size+(j-1)*W.size+k,] <- c(rho[i],muA[j],k,sA,sR)

    # We can get sigma^2 directly because there is no 1st order term

    a3r <- 3*(N-3)
    a2r <- 0
    a1r <- ((1+sR)*(w11+w22+w33)-2*sR*(w12+w13+w23))/((1-sR)*(1+2*sR))

    # We have to choose one real root from polyroot function

    r4r <- -6*(N-3)
    r3r <- 3*(N-3) - (w11+w22+w33)/sA + 2*(w12+w13+w23)/sA
    r2r <- 3*(N-3) - 2*(w11+w22+w33)/sA
    r1r <- (w12+w13+w23)/sA

    sig.r[(i-1)*length(muA)*W.size+(j-1)*W.size+k,] <- c(rho[i],muA[j],k,a1r/a3r)

    temp.r[(i-1)*length(muA)*W.size+(j-1)*W.size+k,] <- Re(polyroot(c(r1r,r2r,r3r,r4r)))
    temp1 <- temp.r[(i-1)*length(muA)*W.size+(j-1)*W.size+k,]

    if(abs(temp1[1]-temp1[2])< 0.0000000001)
      {rho.r[(i-1)*length(muA)*W.size+(j-1)*W.size+k,] <- c(rho[i],muA[j],k,temp1[3])}
    if(abs(temp1[2]-temp1[3])< 0.0000000001)
      {rho.r[(i-1)*length(muA)*W.size+(j-1)*W.size+k,] <- c(rho[i],muA[j],k,temp1[1])}
    if(abs(temp1[1]-temp1[3])< 0.0000000001)
      {rho.r[(i-1)*length(muA)*W.size+(j-1)*W.size+k,] <- c(rho[i],muA[j],k,temp1[2])}

   }
  }
}
```

```
sig.r <- data.frame(sig.r)

colnames(sig.r) <- c("rho","muA","Rep","Est-SigSq")

rho.r <- data.frame(rho.r)

colnames(rho.r) <- c("rho","muA","Rep","Est-rho")

est.r <- cbind(abs(sig.r[,1:3]),Re(sig.r[,4]),Re(rho.r[,4]))

est.r <- data.frame(est.r)

colnames(est.r) <- c("rho","muA","Rep","Est-SigSq","Est-rho")


est.r2.mean <- matrix(0,length(rho)*length(muA),4)

for(i in 1:length(rho)){
  for(j in 1:length(muA)){

    t1<-sqrt(sum(est.r[(((i-1)*length(muA)*W.size+(j-1)*W.size+1):
                        ((i-1)*length(muA)*W.size+(j-1)*W.size+W.size)),4])/W.size)
    t2<-mean(est.r[(((i-1)*length(muA)*W.size+(j-1)*W.size+1):
                    ((i-1)*length(muA)*W.size+(j-1)*W.size+W.size)),5])
    est.r2.mean[(i-1)*length(muA)+j,] <- c(Re(rho[i]),Re(muA[j]),Re(t1),Re(t2))

  }
}

est.r2.mean <- data.frame(est.r2.mean)

colnames(est.r2.mean) <- c(" rho "," muA ","Est-Sigma^2","Est-Rho")


#=============================================================================#
#                          LIKELIHOOD RATIO TEST                              #
#=============================================================================#
#-----------------------------------------------------------------------------#
#  New: Likelihood Ratio Test with MLEs of parameter Sigma-A,-B,-C, and rho   #
#-----------------------------------------------------------------------------#


# Part to compute Determinant of W1, W2, W3, and W

W0.det <- rep(0,length(rho)*length(muA)*W.size)
```

```
for(i in 1:length(rho)){
  for(j in 1:length(muA)){
    for(k in 1:W.size){

    W0.det[(i-1)*length(muA)*W.size+(j-1)*W.size+k] <- det(W.mu[,,k,j,i])

    }
  }
}

W0.det


# Under Ha(Unrestricted condition): sigh1.mat, sigh2.mat, sigh3.mat
# Under H0(Restricted condition): sigh.mat


sigha.mat <- array(0,c(3,3,W.size,length(muA),length(rho)))
sigh0.mat <- array(0,c(3,3,W.size,length(muA),length(rho)))

for(i in 1:length(rho)){
  for(j in 1:length(muA)){
    for(k in 1:W.size){

    e11 <- est.ur2[(i-1)*length(muA)*W.size+(j-1)*W.size+k,4]
    e22 <- est.ur2[(i-1)*length(muA)*W.size+(j-1)*W.size+k,5]
    e33 <- est.ur2[(i-1)*length(muA)*W.size+(j-1)*W.size+k,6]
    r <- est.ur2[(i-1)*length(muA)*W.size+(j-1)*W.size+k,7]

    e0 <- est.r[(i-1)*length(muA)*W.size+(j-1)*W.size+k,4]
    r0 <- est.r[(i-1)*length(muA)*W.size+(j-1)*W.size+k,5]

    d.a <- c(e11,0,0,0,e22,0,0,0,e33)
    d.0 <- c(sqrt(e0),0,0,0,sqrt(e0),0,0,0,sqrt(e0))

    w.a <- c(1,r,r,r,1,r,r,r,1)
    w.0 <- c(1,r0,r0,r0,1,r0,r0,r0,1)

    sigha.mat[,,k,j,i] <- matrix(d.a,3,3)%*% matrix(w.a,3,3)%*%matrix(d.a,3,3)
    sigh0.mat[,,k,j,i] <- matrix(d.0,3,3)%*%matrix(w.0,3,3)%*%matrix(d.0,3,3)

    }
  }
}
```

```
sigha.mat
sigh0.mat

# Part to compute Determinant of sigma_1_hat, sigma_2_hat,
#        sigma_3_hat, and sigma_hat


Sh0.det <- rep(0,length(rho)*length(muA)*W.size)

Sha.det <- rep(0,length(rho)*length(muA)*W.size)

for(i in 1:length(rho)){
  for(j in 1:length(muA)){
    for(k in 1:W.size){

    Sh0.det[(i-1)*length(muA)*W.size+(j-1)*W.size+k] <- det(sigh0.mat[,,k,j,i])
    Sha.det[(i-1)*length(muA)*W.size+(j-1)*W.size+k] <- det(sigha.mat[,,k,j,i])

    }
  }
}

Sh0.det
Sha.det


# Part to compute Inverse Matrix of sigma_1_hat, sigma_2_hat,
#                       sigma_3_hat, and sigma_hat
# Here used Generalized Inverse from MASS Library


library(MASS)

Sh0.inv <- array(0,c(3,3,W.size,length(muA),length(rho)))

Sha.inv <- array(0,c(3,3,W.size,length(muA),length(rho)))

for(i in 1:length(rho)){
  for(j in 1:length(muA)){
    for(k in 1:W.size){

    Sh0.inv[,,k,j,i] <- ginv(sigh0.mat[,,k,j,i])
    Sha.inv[,,k,j,i] <- ginv(sigha.mat[,,k,j,i])

    }
  }
}
```

```
Sh0.inv Sha.inv


# Part to compute tr(W*sigma_hat_inverse) under H0
# Part to compute tr(W*sigma1_hat_inverse),tr(W*sigma2_hat_inverse)
# and tr(W*sigma3_hat_inverse) under Ha


tr.0 <- rep(0,length(rho)*length(muA)*W.size)

tr.a <- rep(0,length(rho)*length(muA)*W.size)

for(i in 1:length(rho)){
  for(j in 1:length(muA)){
    for(k in 1:W.size){

    tr.0[(i-1)*length(muA)*W.size+(j-1)*W.size+k]
         <- sum(diag(W.mu[,,k,j,i]%*%Sh0.inv[,,k,j,i]))
    tr.a[(i-1)*length(muA)*W.size+(j-1)*W.size+k] <-
         <- sum(diag(W.mu[,,k,j,i]%*%Sha.inv[,,k,j,i]))

    }
  }
}

tr.0
tr.a


# Part to compute the value of Gamma function multiplication
# under H0 and Ha


gam.1 <- gamma((3*nsubj-3)/2)

gam.2 <- gamma((3*nsubj-4)/2)

gam.3 <- gamma((3*nsubj-5)/2)

gam.1;gam.2;gam.3

# Part to construct lambda function to test H0 vs Ha

lam.num <- rep(0,length(rho)*length(muA)*W.size)

lam.den <- rep(0,length(rho)*length(muA)*W.size)
```

```
LRT.lambda <- rep(0,length(rho)*length(muA)*W.size)

Q.nLRT <- matrix(0,length(muA)*W.size,length(rho))

lam.num <- rep(0,length(rho)*length(muA)*W.size)

rej.nLRT <- matrix(0,length(rho)*length(muA),4)

for(i in 1:length(rho)){
  for(j in 1:length(muA)){
    cnt <- 0
    for(k in 1:W.size){

    # Part to compute the numerator part of lambda

    part1.num <- W0.det[(i-1)*length(muA)*W.size+(j-1)*W.size+k]^((3*nsubj-7)/2)
                   *exp(-tr.0[(i-1)*length(muA)*W.size+(j-1)*W.size+k]/2)
    part2.num <- Sh0.det[(i-1)*length(muA)*W.size+(j-1)*W.size+k]^((3*(nsubj-1))/2)

    lam.num[(i-1)*length(muA)*W.size+(j-1)*W.size+k] <- part1.num/
       (part2.num*(gam.1*gam.2*gam.3)*(2^((3*(3*nsubj-3))/2))*(pi^((np*(np-1))/4)))

    l.num <- lam.num[(i-1)*length(muA)*W.size+(j-1)*W.size+k]

    # Part to compute the denominator part of lambda

    part1.den <- W0.det[(i-1)*length(muA)*W.size+(j-1)*W.size+k]^((3*nsubj-7)/2)
                 *exp(-tr.a[(i-1)*length(muA)*W.size+(j-1)*W.size+k]/2)
    part2.den <- Sha.det[(i-1)*length(muA)*W.size+(j-1)*W.size+k]^((3*(nsubj-1))/2)

    lam.den[(i-1)*length(muA)*W.size+(j-1)*W.size+k] <- part1.den/
       (part2.den*(gam.1*gam.2*gam.3)*(2^((3*(3*nsubj-3))/2))*(pi^((np*(np-1))/4)))

    l.den <- lam.den[(i-1)*length(muA)*W.size+(j-1)*W.size+k]

    LRT.lambda[(i-1)*length(muA)*W.size+(j-1)*W.size+k] <- l.num/l.den

    Q.nLRT[(j-1)*W.size+k,i] <- -2*log(LRT.lambda[(i-1)*length(muA)*W.size+(j-1)*W.size+k])

    if(Q.nLRT[(j-1)*W.size+k,i] > qchisq(1-0.05,df=2, ncp=0, log = FALSE))
         cnt <- cnt+1
    }

    rej.nLRT[(i-1)*length(muA)+j,] <- c(rho[i],muA[j],cnt,cnt/W.size)
```

```
  }
}

lam.num
lam.den
LRT.lambda
Q.nLRT

rej.nLRT <- data.frame(rej.nLRT)

colnames(rej.nLRT) <- c(" rho "," muA ","# Rej/1000"," Ratio ")

est.r2.mean
est.ur2.mean
rej.nLRT
```

Appendix F

SAS Code

- **Two Treatment Case**

```
OPTIONS NODATE PAGENO=1 noteS SOURCE;

DATA ONE;
 INFILE 'C:\Data\KSU PhD THESIS\PRG\CS-2s2p2t-y6-18data.txt' DELIMITER=' ' ;
 INPUT rho sA2 sB2 nsubj sim seq subj y1 y2 ;
RUN;

DATA one; SET one;
 DROP y1 y2 ;
 per=1; y=y1; output;
 per=2; y=y2; output;
RUN;

DATA one; SET one;
    SUbJ=subj+(seq-1)*nsubj;
    trt='A';
    IF seq=1 and per=2 THEN trt='B';
    IF seq=2 and per=1 THEN trt='B';
RUN;

proc sort data=one;
by rho sim ;
Run;

PROC MIXED ic data=one;
 TITLE 'ANALYSIS USING SAS-MIXED';
 by rho sim;
 CLASSES seq per trt subj;
 MODEL y=seq trt per/DDFM=SATTERTH;
 REPEATED trt/ SUBJECT=subj TYPE=CS;
```

```
 ods listing exclude all;
 ods output infocrit = null COVPARMS=HOPARMS;
RUN;

PROC SORT DATA=HOPARMS;
BY RHO COVPARM;

PROC MEANS;
BY RHO COVPARM;
RUN;

DATA null; set null;
rename neg2loglike =ho;
drop aic--caic;

PROC MIXED ic data = one;
 TITLE 'ANALYSIS USING SAS-MIXED';
 by rho sim;
 CLASSES seq per trt subj;
 MODEL y=seq trt per/DDFM=SATTERTH;
 REPEATED trt/ SUBJECT=subj TYPE=CSH;
 ods listing exclude all;
 ods output infocrit=ha COVPARMS=HAPARMS;
RUN;

PROC SORT DATA=HAPARMS;
BY RHO COVPARM;

PROC MEANS;
BY RHO COVPARM;
RUN;

DATA ha; SET ha;
rename neg2loglike =ha;
drop aic--caic;

DATA comb;
 merge null ha;
 by rho sim;
RUN;

DATA comb;
 SET comb;
 u = ho-ha;
 IF u>3.841459 THEN reject=1; ELSE reject=0;
```

```
RUN;

PROC MEANS data=comb; by rho;
 ods listing select all;
 var reject;
RUN;

QUIT;
```

- **Three Treatment Case**

```
OPTIONS NODATE PAGENO=1 noteS SOURCE;

DATA ONE;
 INFILE 'C:\Data\KSU PhD THESIS\PRG\y12-148data.txt' DELIMITER=' ';
 INPUT rho sA2 sB2 sC2 nsubj sim seq subj y1 y2 y3;
RUN;

DATA one; SET one;
 DROP y1 y2 y3;
 per=1; y=y1; output;
 per=2; y=y2; output;
 per=3; y=y3; output;
RUN;

DATA one; SET one;
    SUbJ=subj+(seq-1)*nsubj;
    trt='A';
    IF seq=1 and per=2 THEN trt='B';
    IF seq=1 and per=3 THEN trt='C';

    IF seq=2 and per=1 THEN trt='B';
    IF seq=2 and per=2 THEN trt='C';

    IF seq=3 and per=1 THEN trt='C';
    IF seq=3 and per=3 THEN trt='B';
RUN;

proc sort data=one;
by rho sim ;
Run;

PROC MIXED ic data=one;
 TITLE 'ANALYSIS USING SAS-MIXED';
 by rho sim;
 CLASSES seq per trt subj;
```

```
 MODEL y=seq trt per/DDFM=SATTERTH;
 REPEATED trt/ SUBJECT=subj TYPE=CS;
 ods listing exclude all;
 ods output infocrit = null COVPARMS=HOPARMS;
RUN;

PROC SORT DATA=HOPARMS;
BY RHO COVPARM;

PROC MEANS;
BY RHO COVPARM;
RUN;

DATA null; set null;
rename neg2loglike =ho;
drop aic--caic;

PROC MIXED ic data = one;
 TITLE 'ANALYSIS USING SAS-MIXED';
 by rho sim;
 CLASSES seq per trt subj;
 MODEL y=seq trt per/DDFM=SATTERTH;
 REPEATED trt/ SUBJECT=subj TYPE=CSh;
 ods listing exclude all;
 ods output infocrit=ha COVPARMS=HAPARMS;
RUN;

PROC SORT DATA=HAPARMS;
BY RHO COVPARM;

PROC MEANS;
BY RHO COVPARM;
RUN;

DATA ha; set ha;
rename neg2loglike =ha;
drop aic--caic;

DATA comb;
 merge null ha;
 by rho sim;
RUN;

DATA comb;
 SET comb;
```

```
 u = ho-ha;
 IF u>5.991465 THEN reject=1; ELSE reject=0;
RUN;

PROC MEANS data=comb;
by rho;
ods listing select all;
var reject;
RUN;

QUIT;
```