A SIMULATION STUDY USING A MIXED MODEL FRAMEWORK TO ANALYZE THE IMPACT OF SAMPLE SIZE AND VARIABILITY ON TYPE I ERROR AND POWER

A Dissertation
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Xiaoxue Gu

In Partial Fulfillment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

Major Department:
Statistics

January 2021

Fargo, North Dakota

NORTH DAKOTA STATE UNIVERSITY

Graduate School

Title

A SIMULATION STUDY USING A MIXED MODEL FRAMEWORK TO ANALYZE THE IMPACT OF SAMPLE SIZE AND VARIABILITY ON TYPE I ERROR AND POWER

By

Xiaoxue Gu

The supervisory committee certifies that this dissertation complies with North Dakota State University's regulations and meets the accepted standards for the degree of

DOCTOR OF PHILOSOPHY

SUPERVISORY COMMITTEE:

Dr. Rhonda Magel
Chair

Curt Doetkott

Dr. Megan Orr

Dr. Changhui Yan

Approved:

09 February 2021
Date

Rhonda Magel
Department Chair

ABSTRACT

Repeated measures design (or longitudinal study) are commonly seen in many research fields, especially in pharmaceutical clinical trials, agricultural research, and psychology. PROC MIXED (SAS Inc.) is a well-known standard tool for analyzing repeated measures data nowadays. The MIXED procedure is based on the standard linear MIXED model, which estimates parameters by maximizing the restricted likelihood. The usual assumption for a standard linear MIXED model is normality. However, the character of data in the real world is hard to tell; it may be non-smoothed, non-symmetric, and having heavy tails, having a small sample size, and so on. Therefore, this simulation study was conducted to check the validity of a MIXED model's statistical inference when violating the underlying assumptions – normality of random errors [Scheffe, 1959], and giving two design features as unbalanced group size and inequality of variance of errors [Scheffe, 1959]. We compare the Type I error rate in different combinations of settings with the Type I error rate under the normal distribution. The power rate is also provided for checking the robustness. The main results in this study show us that the MIXED model is reasonably robust to modest violations of the normal distribution. In the meantime, when small group size combines with large variance, it would cause a severe inflation problem on Type I error rates, which breaks the MIXED model's performance. When the Type I errors were found to be inflated, the Group= option was found to often help with this problem, or sometimes one could use a Sub-Sampling procedure.

ACKNOWLEDGEMENTS

Pursuing a Ph.D. has been a long and arduous journey with many challenges, changes, pitfalls, and opportunities. First and foremost, I want to thank God for helping me through this. It was tough sometimes, but I never lose my faith. Because I know this is God's plans for me, plans to prosper me and not to harm me, plans to give me hope and a future.

Apart from my efforts, the success of this dissertation depends largely on the encouragement and guidelines of many others. I take this opportunity to express my gratitude to the people who have been instrumental in successfully completing this dissertation.

I express my sincere gratitude and most profound appreciation to my advisor, **Dr. Rhonda Magel**, for her guidance, patience, and undivided support through my dissertation and my entire graduate study. I also thank **Mr. Curt Doetkott**. He advised me jointly with Dr. Rhonda Magel and gave me a great deal of insight and expertise that assisted the dissertation. They were always there to listen and give me sincere advice about anything I hoped for. I may not be the easiest student, but they always preserved their confidence in my ability to complete this Ph.D. Without their encouragements and continuous guidance, I could not have finished this dissertation. I would also like to show my gratitude to **Dr. Megan Orr** and **Dr. Changhui Yan** for their valued time and acceptance to serve on my committee.

Finally, I cannot begin to express my unfailing love and gratitude to my husband, **Hanzhe Li**. He has supported and encouraged me through the challenges of graduate study and life. I know there were many sacrifices he made, I am genuinely grateful for having him standing by me. I would also like to thank my amazing son **Ethan Li**, I worked hard because he inspired me to be a better mother. Meanwhile, I owe much gratitude to my parents and parents-in-law — **Guiqin Lu**, **Rui Li**, **Fang Xu**, and **Xianwu Gu** — who supported me emotionally and financially. I salute them for their selfless love, care, and sacrifice they did to shape my life.

It would be difficult to find adequate words to convey how much I owe the people. Lots of love and thank to all of you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Repeated measures design (longitudinal study) is a study in which the outcome variables are repeatedly measured more than once over time for each subject. It is widely used in many research fields, especially in pharmaceutical clinical trials, agricultural research, and psychology. There are three traditional ways to analyze the repeated measures data: ANOVA, MANOVA, and MIXED models, notably using SAS PROC MIXED [Guerin and Stroup, 2000]. Among the three, PROC MIXED allows us to specify the variance/covariance structure and tolerate the missing outcome values, making it a standard tool for repeated measures data nowadays.

This study extends work begun by Taylor King's thesis in 2017 [King, 2017]. In her research, she compared the performance of multivariate analysis with the performance of mixed model and the performance of different covariance structures in a mixed model. According to her suggestions for future research, simulating unbalanced data or following a non-normal distribution would be a good continuation. Therefore, in this study, we violate the underlying assumptions – normality of random errors [Scheffe, 1959], and giving two design features as unbalanced group size and inequality of variance of errors [Scheffe, 1959]. By doing this, the validity of the statistical inference of the MIXED model can be checked. In a repeated measures study, unbalanced sample size features, unequal group variance features, and non-normal distribution are very common. For example, subjects may drop during a longitudinal study, which may cause unbalanced group size; treatments are likely to have heterogeneous variances. Although normal distribution would be an ideal situation, real-world data distribution is unknown, so that non-normal distribution can be expected. Therefore, the results of this study would be useful from a practical point.

For illustration simplification, we first start to view the repeated measures data in a simple linear growth model with only a factor TIME [Kwok et al., 2007]:

$$
\begin{bmatrix} y_{11} \\ : \\ y_{T1} \\ : \\ : \\ : \\ y_{1N} \\ : \\ y_{TN} \end{bmatrix} = \begin{bmatrix} 1 & TIME_1 \\ : & : \\ 1 & TIME_T \\ : & : \\ : & : \\ : & : \\ 1 & TIME_1 \\ : & : \\ 1 & TIME_T \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} 1 & TIME_1 & 0 & : & 0 & 0 & 0 \\ : & : & : & : & : & : & : \\ 1 & TIME_T & 0 & : & 0 & 0 & 0 \\ 0 & 0 & : & : & : & 0 & 0 \\ : & : & : & : & : & : & : \\ 0 & 0 & : & : & : & 0 & 0 \\ 0 & 0 & : & : & : & 1 & TIME_1 \\ : & : & : & : & : & : & : \\ 0 & 0 & 0 & : & 0 & 1 & TIME_T \end{bmatrix} \begin{bmatrix} u_{01} \\ u_{11} \\ u_{02} \\ u_{12} \\ : \\ : \\ : \\ u_{0N} \\ u_{1N} \end{bmatrix} + \begin{bmatrix} e_{11} \\ : \\ e_{T1} \\ : \\ : \\ : \\ e_{1N} \\ : \\ e_{TN} \end{bmatrix} \tag{1.1}
$$

The linear model can be written in matrix form [Littell et al., 2006]

$$
\mathbb{Y} = \mathbb{X}\beta + \mathbb{Z}u + e \tag{1.2}
$$

Where

$$
\begin{pmatrix} u \\ e \end{pmatrix} \sim MVN \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} G & 0 \\ 0 & R \end{pmatrix} \right)
$$

With T time periods and N subjects, the dependent variable y is $TN \times 1$ vector, representing T time measurements for N subjects. $\mathbb{X}$ is TN×2 design matrix containing the average intercept 1 and the slope TIME. β is 2×1 vector having two fixed but unknown parameters β_0 and β_1. $\mathbb{Z}$ is TN×2N design matrix, and u is 2N×1 vector having the random effects of u_{0i} and u_{1i} representing individual subject's difference and follows N(0,G). e is a TN×1 vector containing the random effects for measurement difference and follows N(0,R). The u and e are assumed to be uncorrelated.

There are two components in equation 1.2: fixed effects $\mathbb{X}\beta$ and random effects $\mathbb{Z}u+e$. $\mathbb{X}\beta$ is the mean of y. It is fixed effects because $\mathbb{X}$ is the design matrix, and the parameter β can be fixed. There are two kinds of random effects: between-subject random effects $\mathbb{Z}u$, and within-subject random effects e. The random-effects u_{0i} and u_{1i} in u are between-subject variation, representing the deviation of i_{th} subject's intercept and slope from the averaged intercept and slope. And e is within-subject random error, where the element e_{ti} is the deviation of i_{th} subject at the t_{th}

measurement from the subject's individual regression line. The random-effects u has a covariance matrix G, and error e has a covariance matrix R.

Since the model contains two random effects, the properties of y can be investigated by conditioning on random effects. Therefore, the generalized linear mixed model contains two types of distributions- conditional distribution Equation 1.3 and marginal distribution Equation 1.4, depending on if conditioning on random effects u [Littell et al., 2006]. If there are no random effects (u = 0) in the model, the marginal and conditional variances are identical.

The conditional distribution of y with the following mean and variance

$$y|u \sim MVN(\mathbb{X}\beta + \mathbb{Z}u, R) \tag{1.3}$$

Where R is a block diagonal matrix with N blocks (one per subject), having dimensions T×T.

The marginal distribution of y with the following mean and variance

$$y \sim MVN(\mathbb{X}\beta, V) \tag{1.4}$$

Where the variance V equals to [Kwok et al., 2007]:

$$\begin{aligned} V &= VAR(\mathbb{Z}u + e) \\ &= VAR(\mathbb{Z}u) + VAR(e) \\ &= ZGZ^T + R \\ &= \begin{bmatrix} Z_1 & : & : & 0 \\ 0 & Z_2 & : & 0 \\ : & : & : & : \\ 0 & : & : & Z_N \end{bmatrix}_{N\times N} \begin{bmatrix} T & : & : & 0 \\ 0 & T & : & 0 \\ : & : & : & : \\ 0 & : & : & T \end{bmatrix}_{T\times T} \begin{bmatrix} Z_1 & : & : & 0 \\ 0 & Z_2 & : & 0 \\ : & : & : & : \\ 0 & : & : & Z_N \end{bmatrix}_{N\times N}^T + \begin{bmatrix} \Sigma & : & : & 0 \\ 0 & \Sigma & : & 0 \\ : & : & : & : \\ 0 & : & : & \Sigma \end{bmatrix}_{T\times T} \end{aligned} \tag{1.5}$$

The block structures on the diagonal of matrices in V are defined as follows:

$$Z_i = \begin{bmatrix} 1 & TIME_1 \\ 1 & TIME_2 \\ : & : \\ : & : \\ 1 & TIME_T \end{bmatrix}_{N\times 2} \quad , where \quad i \in (1, ..., N) \tag{1.6}$$

$$T = \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix}_{2\times 2} \tag{1.7}$$

and

$$\Sigma = \begin{bmatrix} \sigma^2_{11} & \sigma^2_{12} & : & \sigma^2_{1T} \\ \sigma^2_{21} & \sigma^2_{22} & : & \sigma^2_{2T} \\ : & : & : & : \\ \sigma^2_{T1} & \sigma^2_{T1} & : & \sigma^2_{TT} \end{bmatrix}_{T\times T} \tag{1.8}$$

T is a 2×2 matrix having variances and covariances for between-subject random effects, and Σ is a T×T block matrix showing the variances and covariances for within-subject random effects.

Since the data are repeatedly measured, the errors in the mixed model are correlated. A common correlation among measurements is assumed for each subject, so there are multiple choices of covariance structures that can be chosen as a common correlation [Guerin and Stroup, 2000, SAS, 2015, Littell et al., 2006, Kincaid, 2005]. The five most common covariance structures using k=4 repeated measures are described as below.

Variance Components(VC) [Littell et al., 2006, Kincaid, 2005] is the default covariance structure for a PROC MIXED procedure in SAS and is also the simplest covariance structure. It has four different subject variances in the diagonal and has zero in all off-diagonals. This structure assumes independence of errors.

$$VC = \begin{pmatrix} \sigma^2_{\gamma 1} & 0 & 0 & 0 \\ 0 & \sigma^2_{\gamma 2} & 0 & 0 \\ 0 & 0 & \sigma^2_3 & 0 \\ 0 & 0 & 0 & \sigma^4_4 \end{pmatrix}$$

First-Order Autoregressive(AR(1)) [Littell et al., 2006, Kincaid, 2005] is used widely in time series data. It can only be used when time intervals between any two measurements are equal in repeated measure. The correlation between two measurements is defined by exponential function ρ^x, so the correlation will decrease when time-space increases.

$$AR(1) = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{pmatrix}$$

Toeplitz(TOEP) [Littell et al., 2006, Kincaid, 2005] has more parameters than VC, AR(1), and CS, but a smaller number of parameters than UN. The measurements taken at closer time intervals have similar correlations.

$$TOEP = \sigma^2 \begin{pmatrix} 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{pmatrix}$$

Compound Symmetry(CS) [Littell et al., 2006, Kincaid, 2005] is used for repeated measures having the same correlation. A constant correlation is assumed between two separate measurements.

$$CS = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{pmatrix}$$

Unstructured(UN) [Littell et al., 2006, Kincaid, 2005] is the most complex covariance structure because each term can be different. It may be the best structure when fitting the real data since the correlation between any two measurements does not have any constraints. However, it may use up many degrees of freedom which would cause the Type I error increasing, especially when the data set is small.

$$UN = \begin{pmatrix} \sigma_{11}^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_{22}^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_{33}^2 & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44}^2 \end{pmatrix}$$

2. LITERATURE REVIEW

2.1. SAS PROC MIXED Procedure

There are lots of guidelines and published papers about how to use SAS PROC MIXED. PROC MIXED is based on REML (restricted maximum-likelihood) approach for parameter estimation [Jennrich and Schluchter, 1986, Littell et al., 2006]. The F test are the default statistical tests in PROC MIXED procedure for the main effects, and interaction effects of repeated measures data, which tends to cause Type I error inflation problems with multiple covariance structures in unbalanced designs, and non-normal data distribution [Keselman et al., 1999b]. However, the Satterthwaite F test, which can define the denominator degrees of freedom of F test through PROC MIXED is fairly robust compared with the default F test on the same condition [Keselman et al., 1999a]. Therefore, the DDFM= option in the MODEL statement is important because it can specify the method for computing the denominator degrees of freedom for the fixed effects tests. There are five methods for DDFM=, which are CONTAIN, BETWITHIN, RESIDUAL, SATTERTH, and KENWARDROGER. Among the five, DDFM=KENWARDROGER adjusts the denominator degrees of freedom based on Satterthwaite-type denominator degrees of freedom [Kenward and Roger, 1997, Prasad and Rao, 1990, Harville and Jeske, 1992], which makes it effectively control the Type I error rate for the repeated measures fixed-effects.

Based on repeated measurement data, there are two statements in PROC MIXED that needed to be specified: REPEATED statement and RANDOM statement[Littell et al., 2006]. The REPEATED statement can specify the variable name of a repeated measure factor. Within a REPEATED statement, the SUBJECT option defines the sets of repeated measures, and the TYPE option names the covariance structure, which must be used when only using REPEATED statement. RANDOM statement can specify the random effects. When repeated measures are modeled with a REPEATED statement, without a RANDOM statement in PROC MIXED, this model is called a conditional model based on the conditional distribution of y. In a conditional model, the TYPE option under the REPEATED statement incorporates the complex covariance structure directly through the variance matrix R. When repeated measures are modeled with both REPEATED statement and RANDOM statement in PROC MIXED, the model is called a marginal model based on the marginal distribution of y. In the marginal model, the TYPE= option under the REPEATED statement specifies the variance matrix R which is typically denoted for variance matrix of random error e, and the TYPE= option under the RANDOM statement specifies the variance matrix G which is typically denoted for variance matrix of random error u [Littell et al., 2006]. Based on our simulation experiment results, these two models would provide the same parameter estimates for fixed effects.

When fitting a model with heterogeneous variance structure, a model with unequal variances can be specified in PROC MIXED under the REPEATED/RANDOM statement with the GROUP= option [Littell et al., 2006]. The GROUP= option allows the parameters of different GROUP effect levels to have different structure parameters despite a covariance structure (TYPE= option) remaining the same. So it will change the covariance parameters from one group to another, which can remarkably increase the number of covariance parameters needing to be estimated [Kincaid, 2005]. Also, GROUP= option is limited to categorical factors, which requires using a CLASS statement. For example, when incorporating between-subject variance heterogeneity, the GROUP= option in the REPEATED statement can be set up. An example code can be viewed as below [SAS, 2015].

```
proc mixed;
class A;
model y = A / ddfm=satterth;
repeated / group=A;
lsmeans A / adjust=smm adjdfe=row;
run;
```

2.2. Previous Simulation Studies

The Behrens–Fisher problem [Fisher, 1938, Kim and Cohen, 1998, Paul et al., 2019]has existed for more than sixty years in the area of statistics. The problem is named after Walter Behrens and Ronald Fisher. It occurs when testing the means of two independent populations without knowing the equality of the variances [Fisher, 1938, Kim and Cohen, 1998, Paul et al., 2019]. The Behrens-Fisher problem considers the basic design features, unequal or unknown variances, under two normally distributed populations. However, data in the real world are more often skewed, non-smoothed, non-symmetric, and having heavy tails [Hill and Dixon, 1982]. The test statistics do not always deal with an ideal situation, like equal sample sizes, equal variances. Therefore, there are lots of studies about the analog of the Behrens-Fisher Problem.

Henry Scheffe talked about the effects of departures from the underlying assumptions in his book "The Analysis of Variance" [Scheffe, 1959], which is one of the analogous problems of the Behrens-Fisher Problem for the non-normal distribution. In this book, he violates the following assumptions[Scheffe, 1959]:

1. normality of errors, and normality of the random effects in the models;

2. equality of variance of the errors;

3. statistical independence of the errors.

Based on his real data examples, he came up with three conclusions[Scheffe, 1959]:

1. nonnormality has minimal impact on inferences about means but substantial impact on inferences about the variances of random effects;

2. Unequal variance has little impact on inferences about means when sample sizes are equal but has notably impact when sample sizes are unequal;

3. Correlated observations can cause severe problems with inferences about means.

According to the underlying violations mentioned above, some methods are recommended for addressing the severe effects when having two population groups[Scheffe, 1959, Paul et al., 2019]. With assumed equal size, the classical Student's T-test is recommended. If two populations have equal group size or if the distributions are symmetrical, the Student's t-test is robust; if two populations have unequal group size and the distributions are skewed, the effects of departure from normality may be a concern; If population distributions are normal but with unequal and unknown variances, either Satterthwaite's t-statistic or Satterthwaite's F test is suggested. However, Satterthwaite's procedure is not robust under most non-normal distributions [Reed III, 2003, Paul et al., 2019].

2.3. Second Order Response Surface Models

Response surface methodology is a statistical technique to investigate, characterize, or optimize a response regarding a set of quantitative variables through linear models and second-order polynomial models [Box and Wilson, 1951]. It has become the standard framework for industrial design and development nowadays. The approximation of response surface model is $y = f(x_1, x_1, ..., x_q) + \epsilon$. There are three types of response surface models: First-Order, Second-Order, and Mixture Models [Johnson and Montgomery, 2009].

The second-order surface model is widely used when the curvature in response surface is detected. The estimation of second-order polynomial model is one of the most important primary tools of response surface methodology, it includes all terms of the first-order of design variables x_1,x_2,...,x_n, quadratic and cross product to present the true response with curvature. In general, a second-order surface model expression[Bradley, 2007, Myers and Montgomery, 1997, Amir et al., 2016] takes the following form:

$$y = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \beta_{ii} x_i^2 + \sum \sum_{i \leq j} \beta_{ij} x_i x_j + \epsilon$$

In SAS, the RSREG procedure is one of the most specialized procedures to conduct a second-order response surface regression model compared to the REG and GLM procedures. It used the method of least squares to fit quadratic response surface models [Inc, 2016]. RSREG procedure with surface plot performance can provide a more explicit and vivid finding and give a better idea of the dependent variables at the various settings of two independent variables [Inc, 2016, Amir et al., 2016].

2.4. Data Transformation

For equal spreads and reducing skewness of distributions, data transformation was usually the first step to deal with data. Transformation is to replace a variable with a function of that variable. After data transformation, the shape of distribution or relationship will be changed. There are many functions for data transformation, such as log(x), square x^2, square root $x^{0.5}$. Among them, rank test [Lehmann and D'Abrera, 1975] is one of the standard tools in applied statistician's tool kit because of its convenience and simplicity. It is to replace the original observations with their respective rank, then compute tests on these ranks.

Aligned rank transformation [Higgins et al., 1990] adds a simple alignment fix-up methodology before ranking, which can be applied to analyzing multi-factor designs when the error distribution is moderately skewed. In the aligned rank test, data are aligned, ranked, then analyzed. The main difference is the alignment. The purpose of alignment is to remove the effect of ”nuisance” parameters when testing the effects of parameters of interests for multi-parameter models. For example, the effect of blocks in testing for effects of treatments can be removed by data alignment in completely randomized block design [Lehmann and D’Abrera, 1975].

2.5. Sub-Sampling and Bootstrap Method

Sub-sampling and Bootstrap are widespread re-sampling methods. Comparing traditional methods, they require fewer assumptions and are more accurate in practice[Hesterberg et al., 2005]. Generally speaking, sub-sampling is the method to draw a subset randomly and without replacement from the original data samples [Efron, 1981, Politis et al., 1999, Schroeder and Martin, 2005]. Bootstrap is to generate a sample with replacement randomly from original data samples, usually of the same size as the original sample [Chernick, 2011].

3. METHODOLOGY

The purpose of this paper is twofold: firstly, to violate the normality assumption of the MIXED model, the methods for this part will be described in Section 3.4; and secondly, to suggest a new approach for the analogous Behrens-Fisher Problem, the methods for this part will be described in Section 3.5.

3.1. Simulation Methodology

3.1.1. Simulation Program

To explore the results of statistics in previous chapters, we used SAS 9.4 (SAS Institute, NC, USA) to perform all simulations and analyses [Wicklin, 2013]. Each simulation was examined using 5000 samples with a 0.05 significance level of α. There are two intervals used as the index for the estimates' precision: Bradley's liberal criterion [Bradley, 1978] and binomial standard error interval [Kowalchuk et al., 2004, Bradley, 1978]. The test robustness can be evaluated by whether the empirical estimate Type I error ($\hat{\alpha}$) stays within the interval of $0.5\alpha \leq \hat{\alpha} \leq 1.5\alpha$, which is $0.025 \leq \hat{\alpha} \leq 0.075$ in this study. The binomial standard error is $[\frac{\hat{\alpha}(1-\hat{\alpha})}{N}]^{0.5}$, where N is the total number of samples. In this study, with a significance level of 0.05 and 5000 samples, the Type I error rate should stay between 0.04396 and 0.05604.

3.1.2. Hypotheses

The Type I error rate was calculated by counting the number of times that the null hypothesis H_0 was rejected when H_0 is true and dividing by the total number of samples. The power rate was calculated by counting the number of times that the null hypothesis H_0 was rejected when H_a is true and dividing by the total number of samples. There are three hypotheses included in this study.

- All Treatment main effects means equal, $H_0 : \tau_1 = \tau_2$
- All Time main effects means equal, $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$
- All Treatment×Time interaction effects means equal,
  $H_0 : \tau\alpha_{11} = \tau\alpha_{12} = \tau\alpha_{13} = \tau\alpha_{14} = \tau\alpha_{21} = \tau\alpha_{22} = \tau\alpha_{23} = \tau\alpha_{24}$

3.1.3. Data Simulation

This simulation study was performed with 5000 samples, and each sample was conducted by a split-plot design assuming equally spaced time intervals. Our split-plot design has 2 treatment groups, 4 repeated time periods, and a First-Order Autoregressive [AR(1)] correlation structure [Littell et al., 2006, Kincaid, 2005] where ρ=0.75. There are two stages in this experiment. In the first stage, subjects are randomly assigned to treatment groups (whole-plot factor); In the subsequent stage, time factor (sub-plot factor) in repeated measures being nested within each of the subjects without randomization. There are 30 subjects in each sample. Each subject was randomly assigned to a treatment group and was repeatedly measured four times. For better understanding, one sample data set can be visualized in Table 3.1. The number in each group changes depending on the specific simulation scenarios.

Table 3.1. Repeated Measures Data with 30 Subjects

Subject ID	Treatment	y1	y2	y3	y4
1	Control	.	.	.	.
2	Control	.	.	.	.
3	Control	.	.	.	.
..	..	..	..	..	..
..	..	..	..	..	..
28	Treatment	.	.	.	.
29	Treatment	.	.	.	.
30	Treatment	.	.	.	.

An effects model for this experiment is

$$Y_{ijk} = \mu_{ij} + \gamma_k + e_{ijk} = \mu + \tau_i + \alpha_j + (\tau\alpha)_{ij} + \gamma_k + e_{ijk} \qquad (3.1)$$

Where

Subjects k=1,2,...,30

Treatments i=1,2

Time periods j= 1,2,3,4

$\mu_{ij} = \mu + \tau_i + \alpha_j + (\tau\alpha)_{ij}$ is the mean μ for treatment i at time j, containing treatment effects τ_i, time effects α_j, and interaction treatment×time effects $(\tau\alpha)_{ij}$, respectively.

γ_k is the whole-plot error effect for subject k, assumed $\overset{\text{iid}}{\sim}$ N(0,σ_γ^2).

e_{ijk} is the sub-plot error effect for j_{th} time measurement of subject k on treatment i, assumed $\overset{\text{iid}}{\sim}$ N(0,σ^2).

γ_k and e_{ijk} are assumed to be independent of one another

The corresponding matrix form of this model is

$$\mathbb{Y} = \mathbb{X}\beta + \mathbb{Z}w + e \tag{3.2}$$

Where

$\mathbb{Y}$ is the vector of observations.

β is the coefficient vector corresponding to the fixed effects μ_{ij}.

$\mathbb{X}$ is the design matrix for the fixed effects.

w is the coefficient vector corresponding to whole-plot errors.

$\mathbb{Z}$ is the design matrix with respect to whole-plot errors.

e is the vector corresponding to split-plot errors.

To obtain the repeated measures y in a simulation study, two parts of this matrix model needed to be provided. The first part is to specify the fixed effects $\mathbb{X}\beta$, and the second part is to generate the two random effects - $\mathbb{Z}w$ and e, respectively.

The first part, the mean effects u_{ij} in the effects model can be obtained by the fixed unknown constant $\mathbb{X}\beta$ in the matrix model, which contains design matrix $\mathbb{X}$ and parameter vector β. Since it is a 2 by 4 split-plot design, the vector β is set as $\beta = (\mu, \tau_1, \tau_2, \alpha_1, \alpha_2, \alpha_3, \alpha_4)'$ with no interaction effects being considered. In this study, the parameter vector $\beta_{7\times1}$ was set as $\beta = (5, 0, 0, 0, 0, 0, 0)'$ when there are no main effects assumed, it was applied when H_0 is true; the parameter vector $\beta_{7\times1}$ was set that $\beta = (5, 0, 1.5, 0, 0.2, 0.4, 0.6)'$ when adding the main effects to treatment effects and time effects, it was applied when H_a is true. The design matrix $\mathbb{X}$ would have 7 columns that correspond to each parameter in β, and have 120 rows that correspond to each measurement of each subject. (30)(4)=120 rows because each sample has 30 subject, and 4 repeated measures per subject. Therefore, the design matrix $\mathbb{X}$ is as follows.

$$\mathbb{X} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ .. & .. & .. & .. & .. & .. & .. \\ .. & .. & .. & .. & .. & .. & .. \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}_{120\times 7}$$

In the second part - two random effects, Ramon Littell [Littell et al., 2006] provided a formula for getting the random effects variance. That is $V = var(y) = \sigma_\gamma^2 J + R$, where J is a matrix of ones. The $\sigma_\gamma^2 J$ is the variance for the between-subject random effect $\mathbb{Z}w$, and the R is the variance for the within-subject random effect e. The two random effects were both assumed with a mean zero. Therefore, $\mathbb{Z}w$ has mean zero and covariance matrix $\sigma_\gamma^2 J$, and e has mean zero and covariance matrix R.

The part J which is matrix of ones was chosen as between-subject covariance structure because the measures are on the same subject, and σ_γ^2 is the variance of treatment groups. The part R represents the covariance due to the proximity of measurements. R is a covariance matrix corresponding to a within-subject variance. In this study, we assumed that the within-subject variances R for all subjects are identical.

Regarding choosing a covariance structure for R, Unstructured(UN) [Kincaid, 2005] is commonly recommended as the initial covariance structure when using the MIXED model for repeated measures data because the right covariance structure is unknown. However, defining a correlation for any pair of terms would be difficult since there would not need to be any pattern for the Unstructured (UN). Meanwhile, the Unstructured (UN) has the most parameters compared with other structures, which may cause loss of power. Therefore, for obtaining a time series structure, First-Order Auto-regressive (AR(1)) [Littell et al., 2006, Kincaid, 2005] was chosen as the right covariance structure with the correlation ρ as 0.75, and the within variance σ^2 was set as 1. The covariance matrix R in this study is presented as follows.

$$R = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0.75 & 0.5625 & 0.0421875 \\ 0.75 & 1 & 0.75 & 0.0525 \\ 0.0525 & 0.75 & 1 & 0.75 \\ 0.0421875 & 0.0525 & 0.75 & 1 \end{pmatrix}$$

3.1.4. Simulation Scenarios

The usual assumption for a standard linear MIXED model is normality. For checking the validity of the MIXED model, we violated the assumption by simulating normal/non-normal distribution of between-subject effects $\mathbb{Z}w$ and within-subject effects e [Scheffe, 1959], specifically. Therefore, four different scenarios were generated: 1.between subject effects $\mathbb{Z}w$ follows a multivariate normal distribution and within-subject effects e follows multivariate non-normal distribution; 2. between-subject effects $\mathbb{Z}w$ follows a multivariate normal distribution and within-subject effects e follows multivariate normal distribution; 3. between-subject effects $\mathbb{Z}w$ follows multivariate non-normal distribution and within-subject effects e follows multivariate normal distribution; 4. between-subject effects $\mathbb{Z}w$ follows multivariate non-normal distribution and within-subject effects e multivariate non-normal distribution. The list of the four basic scenarios are shown as follows.

1. $\mathbb{Z}w$ ~MVN(0, σ_γ^2J),e ~Multi-Skew(μ=0, R, Skew=2, Kurtosis=6)

2. $\mathbb{Z}w$ ~MVN(0, σ_γ^2J),e ~MVN(0, R)

3. $\mathbb{Z}w$ ~Multi-Skew(μ=0, σ_γ^2J, Skew=2, Kurtosis=6),e ~MVN(0, R)

4. $\mathbb{Z}w$ ~Multi-Skew(μ=0, σ_γ^2J, Skew=2, Kurtosis=6),e ~Multi-Skew(μ=0, R, Skew=2, Kurtosis=6)

For simulating the multivariate non-normal distribution data in this study, the univariate distribution was first generated by Fleishman's Cubic Transformation [Fleishman, 1978] with target values of skewness=2 and kurtosis=6, then the method from Vale-Maurelli [Vale and Maurelli, 1983] was used to generate multivariate non-normal data.

Within each scenario, two conditions were applied for checking the stability of the MIXED model. They are the equality of sizes, and the equality of (between-subject) variances for two treatment groups [Scheffe, 1959]. The parameters sets for two conditions were listed as follows. Here, 1 is for the treatment group, and 2 is for the control group.

- Equal group size $n_1 = n_2 = 15$
- Unequal group size $n_1/n_2 = (0.5, 2)$, where
  - For n_1/n_2=0.5, $n_1 = 10$ and $n_2 = 20$
  - For n_1/n_2=2.0, $n_1 = 20$ and $n_2 = 10$
- Equal variances $\sigma_{\gamma 1}^2 = \sigma_{\gamma 2}^2 = (1, 2, 4, 10)$
- Unequal variances $\sigma_{\gamma 1}^2/\sigma_{\gamma 2}^2$=(2,4,6,8,10), where $\sigma_{\gamma 2}^2$=1

The two conditions resulted in four different situation combinations: equal group size with equal variance, equal group size but unequal variance, unequal group size but equal variance, unequal group size and unequal variance.

3.2. Analysis

After these data were generated based on these simulation scenarios and different conditions, PROC MIXED was applied to run the test by sample. In PROC MIXED, the conditional distribution of the mixed model was used. DDFM=KENWARDROGER is to adjust the degrees of freedom. The five most common covariance structures (Variance Components (VC), First-Order Autoregressive (AR(1)), Toeplitz (TOEP), Compound Symmetry (CS), Unstructured (UN) [Littell et al., 2006, Kincaid, 2005]) were applied under the REPEATED statement, respectively. An example code can be viewed as below [Littell et al., 2006].

```
proc mixed data=rm.uv_dsn._p._n;
by sample;
class trt period subj_id;
model stress = trt | period /ddfm=kr;
repeated period / subject=subj_id type=AR(1);
title2 "Repeated Measures ANOVA using Mixed Model Approach -- AR(1)";
run;
```

3.3. Second-Order Response Surface Models

In this study, we would have two independent variables: size ratio and variance ratio. Let us set the size ratio as x_1, variance ratio as x_2. The second-order response surface model would include all quadratic and cross-product terms of x_1 and x_2. Therefore, the second-order response surface model can be expressed as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \epsilon$$

RSREG procedure in SAS is commonly used to create a prediction and generate a surface plot of the prediction for second-order response surface model [Inc, 2016]. The code concerning these two-factor variables is provided below [Inc, 2016]:

```
proc rsreg plots=surface(3D);
model y=x1 x2;
run;
```

3.4. Verifying the Stability of Type I Error Rate

For checking the consistency of the Type I error rates, two methods would be used: increasing the sample sizes and extending the repeated time points.

One scenarios, $\mathbb{Z}w \sim$ Normal $e \sim$ Skew, was set up as an example. Under this scenario, three different group size sets within the same size ratio were added to see if this would affect Type I error rates. According to Taylor's research results [King, 2017], she suggested that a sample size of 30 is needed to produce adequate power with different simulation parameters. Therefore, within the previous condition unchanged, three different group sizes were generated by double, triple, quadruple the original sample size. That is, for size ratio n_1/n_2=0.5, the three different groups sizes are 20vs40, 30vs60, and 40vs80; for size ratio n_1/n_2=1, the three different groups sizes are 30vs30, 45vs45, and 60vs60; for size ratio n_1/n_2=2, the three different groups sizes are 40vs20, 60vs30, 80vs40. The power rates for these three different group sizes were also generated for reference. Besides, the repeated measures would also be extended from 4 times to 6 times to see if this would affect Type I error rates. The corresponding power rates were generated for reference.

3.5. Null Distribution Testing

For characterizing the impact of the Type I error rate inflation on power assessment, we could conduct sample distribution by using Monte Carlo methods.

Firstly, we would simulate a null hypothesis distribution with large enough samples (10,000 samples) and get the null distribution's F values. The 95% F critical value can be found. In the Type I error rate inflation case (we have 5000 samples in our case), by adding an effect size between control and treatment groups (so H_0 is false), we would get the observed power rate and also the F values in all the cases of observed power rates. We would compare the F value in the observed power rate with the F critical value. If the F value in the case of observed power rate is greater than the F critical values, then it is one reject case (Reject H_0 when H_0 is false, which is power). So the theoretical power rate can be calculated by total rejection cases divided by the number of samples in the observed power rate (5,000). Lastly, by subtracting the theoretical power rate from the observed power rate, we would know how the power rate is impacted by the inflated Type I error rate.

Meanwhile, we also check if the Type I error rate was simulated correctly by comparing the F critical value with the F value in the case of the observed Type I error rate. Specifically, if F value in the case of observed Type I error rate is greater or equal to F critical value, then it is a rejection case (Reject H_0 when H_0 is true, which is Type I error). The estimated Type I error rate can be calculated by the total rejection cases divided by the number of times that the samples were replicated in the cases of observed Type I error rate (5,000). If the estimated Type I error rate is approaching 5.00%, the Type I error rate was simulated correctly.

3.6. Methods for Type I Error Rate Inflation Problem

When Type I error rate is inflated, there are four different methods would be used: Rank Test [Lehmann and D'Abrera, 1975], Aligned Rank Test [Higgins et al., 1990], Sub-Sampling Method [Schroeder and Martin, 2005], and MIXED model incorporating GROUP= option under REPEAT statement [Littell et al., 2006]. We would compare the four methods based on their Type I error rate and power rates. The technical details of these four methods would be explained below.

For Rank Test, it is to replace the original observations with their respective rank, then compute tests on these ranks; For Aligned Rank Test [Higgins et al., 1990], the data are ranked after they are aligned. The alignment procedure in this study was introduced as below[Higgins et al., 1990], and the code can be found in Appendix B.

The mathematical model is

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \tag{3.3}$$

Where i=1,2, j=1,2,3,4, k=1,...,30, and the e_{ijk}s are assumed $\overset{\text{iid}}{\sim}$ N(0,σ^2). The α_i and β_j is represented as row and column effects, respectively. The usual assumption on parameters are: $\sum_i \alpha_i$=0, $\sum_j \beta_j$=0, $\sum_i(\alpha\beta)_{ij}$=$\sum_j(\alpha\beta)_{ij}$=0. These estimates are: $\hat{\mu} = \bar{Y}_{...}$, $\hat{\alpha}_i = \bar{Y}_{i..} - \bar{Y}_{...}$, $\hat{\beta}_j = \bar{Y}_{.j.} - \bar{Y}_{...}$, $\hat{(\alpha\beta)}_{ij} = \bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}$

For rows the aligned data are

$$A_{ijk} = Y_{ijk} - (\hat{\mu} + \hat{\beta}_j + \hat{(\alpha\beta)}_{ij}) = Y_{ijk} - \hat{Y}_{ij.} + \hat{Y}_{i..} - \hat{Y}_{...} \tag{3.4}$$

For columns the aligned data are

$$B_{ijk} = Y_{ijk} - (\hat{\mu} + \hat{\alpha}_i + \hat{(\alpha\beta)}_{ij}) = Y_{ijk} - \hat{Y}_{ij.} + \hat{Y}_{.j.} - \hat{Y}_{...} \tag{3.5}$$

The aligned data for testing for interactions have the form

$$AB_{ijk} = Y_{ijk} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j) = Y_{ijk} - \hat{Y}_{i..} - \hat{Y}_{.j.} + \hat{Y}_{...} \tag{3.6}$$

For Sub-Sampling method, since we have unbalanced groups, our aims to make group size equal by drop subjects from the large group. The Sub-Sampling method in this study is to keep the group having a small sample size, then randomly select subjects without replacement from the group having a large sample size until the large group has the same group size as the small one. For GROUP= option, an example code for this method can be viewed as below [SAS, 2015].

```
proc mixed data=rm.uv;
  by sample;
  class trt period subj_id;
  model stress = trt | period /ddfm=kr;
  repeated period / subject=subj_id type=AR(1) group=trt;
  title2 "Repeated Measures ANOVA using Mixed Model Approach -- AR(1)";
  run;
```

4. RESULTS

We estimate Type I error rates by finding the percentage of the cases that reject the null hypothesis H_0 when the H_0 is true. Also, the significance level α is stated as 5.00%. The Type I error rate for all combinations of two conditions: the equality of group size and the equality of variance, under four basic scenarios are presented in the following tables in percentage form. In each table, there are three test parts of different effects: treatment effects, time period effects, and interaction (treatment$\times$time) effects. For better understanding, a diagram which organizes the various elements of the results and shows how the parts are related is presented in Figure 4.1.

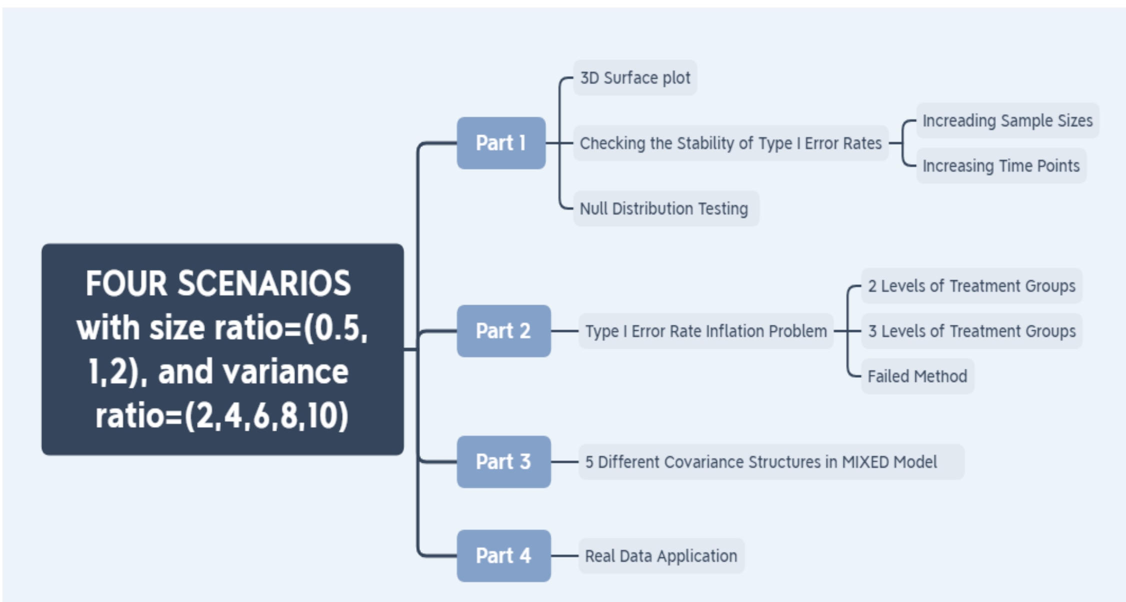

Figure 4.1. Results Diagram

4.1. Four Scenarios

As stated in Chapter 3.4, the four scenarios are listed as follows:

- $\mathbb{Z}w \sim$ Normal & $e \sim$ Skew
- $\mathbb{Z}w \sim$ Normal & $e \sim$ Normal
- $\mathbb{Z}w \sim$ Skew & $e \sim$ Normal

- $\mathbb{Z}w \sim$ Skew & $e \sim$ Skew

Based on the parameters we set, the first two scenarios have normal distribution or distribution with slightly skewed, and the last two scenarios have distributions with heavily skewed.

Table 4.1 displays the Type I error rates of four basic scenarios for treatment and control group with $n_1 = n_1 = 15$ and equal variance $\sigma^2_{\gamma 1} = \sigma^2_{\gamma 2}$ with increased variance as given in (1,2,4,10). The Type I error rates in both period results and interaction results are below the limits of α equals 5.00%, staying between 3.36% and 4.88%, a bit below the binomial threshold (4.396, 5.604). Values of the Type I error rates in treatment results stay within 4.44% to 5.48%, but the highest value, 5.48%, is still below the upper bound of binomial standard error interval (4.396, 5.604). Also, the tests are robust here. The four scenarios have a similar trend, so the violation of normality appears to cause little effect on Type I error rates no matter whether the skewness is in between-subject effects or within-subject effects.

Table 4.1. Type I Error Rate of Four Basic Scenarios for Balanced Group Size $n_1 = n_2 = 15$; and Equal Variances $\sigma^2_{\gamma 1} = \sigma^2_{\gamma 2} = (1, 2, 4, 10)$

σ^2_γ	Distribution Scenarios			
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew	$\mathbb{Z}w \sim$ Normal $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Skew
Treatment Results				
1	4.96	5.48	5.38	5.12
2	4.80	4.66	5.04	4.78
4	5.34	5.46	5.44	4.90
10	4.96	5.14	4.44	4.52
Period Results				
1	3.74	4.20	4.86	3.76
2	3.50	4.12	4.14	3.64
4	3.70	3.84	4.34	4.10
10	3.98	4.58	3.96	3.66
Treatment*Period Results				
1	3.52	3.54	4.22	3.70
2	3.36	4.54	4.02	3.90
4	3.42	4.28	3.86	3.84
10	3.84	4.88	4.02	3.66

Based on the two conditions (the equality of group size and the equality of variance) of Table 4.1, Table 4.2, and Table 4.3, change one condition separately but keep the other one the same. The Type I error rates under the four scenarios keep the same character with only unequal group size, as described in Table 4.3. With only unequal variance, the Type I error rates in period, and interaction tests are all below the limit of 5.00%. However, in the treatment test, when increasing the variance ratio, the distribution which is normal or slightly skewed keeps the Type I error rate staying within 95% binomial standard error interval (4.396, 5.604), but the distribution with high skewness increases the Type I error rates from 4.66% to 7.86% Therefore, unbalanced group size or unequal group variance itself should not be a concern when using a MIXED model. However, when the high skewness presents, the effects of the big difference of group variances should be a concern.

Table 4.2. Type I Error Rate of Four Scenarios for Unequal Group Size $n_1 = 10$ and $n_2 = 20$; Equal Variances $\sigma^2_{\gamma 1} = \sigma^2_{\gamma 2} = (1, 2, 4, 10)$

σ^2_γ	Distribution Scenarios			
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew	$\mathbb{Z}w \sim$ Normal $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Skew
Treatment Results				
1	5.06	4.72	5.14	5.16
2	5.34	5.96	4.98	4.24
4	5.72	4.70	4.48	4.88
10	4.78	5.10	4.38	4.40
Period Results				
1	4.14	4.66	4.16	4.26
2	4.18	3.94	4.52	3.82
4	3.72	4.14	4.30	3.96
10	4.18	4.04	4.62	3.72
Treatment*Period Results				
1	4.04	3.64	4.34	3.94
2	4.04	4.46	4.24	3.84
4	4.34	4.24	4.12	3.84
10	3.96	3.62	4.00	4.26

Table 4.3. Type I Error Rate of Four Scenarios for Equal Group Sizes $n_1 = n_2 = 15$; Unequal Variance as Giving the Variance Ratio $\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$=(2,4,6,8,10), where $\sigma^2_{\gamma 2}$=1

$\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$	Distribution Scenarios			
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew	$\mathbb{Z}w \sim$ Normal $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Skew
Treatment Results				
2	5.32	5.36	5.44	4.66
4	5.16	5.42	5.34	6.28
6	5.34	5.20	6.58	6.32
8	5.08	5.28	7.28	6.90
10	5.24	5.04	7.86	7.76
Period Results				
2	4.02	4.66	3.94	3.96
4	3.58	4.00	4.12	4.00
6	4.04	3.94	3.66	3.20
8	3.46	4.28	3.90	4.02
10	3.96	4.20	3.82	3.88
Treatment*Period Results				
2	3.84	4.00	3.96	3.94
4	3.72	4.14	4.24	3.80
6	3.70	4.12	4.40	3.66
8	3.66	4.34	3.66	3.52
10	3.44	4.04	4.30	3.56

Table 4.4 illustrates how the Type I error rate performs when two conditions change at the same time. The information of size ratio equals to 1 was provided as a reference. In period and interaction tests, Type I error rates stays below the limit of 5.00%. Under the treatment test, when the size ratio equals to 2, Type I error rates go conservative as the variance ratio increases in all four scenarios, below 5.00. When the size ratio equals to 0.5, the Type I error rates are inflated incredibly as the variance ratio increases in all four scenarios, rising from 6.36% to 15.02%, which is above the upper bound of 95% binomial interval (4.396, 5.604). It also means that most treatment tests are not robust when the size ratio equals to 0.5. Meanwhile, all four scenarios have a similar pattern in all size ratios except for size ratio one. According to these results, when small group size combines with large variance, it would cause a severe inflation problem on Type I error rates, which breaks the MIXED model's performance.

By adding the main treatment effects (0,1.5) and time effect (0,0.2,0.4,0.6), Table 4.5 under $H_a = (\mu, \tau_1, \tau_2, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$=(0,1.5,0,0.2,0.4,0.6) is true, was created to present the corresponding power rate for Table 4.4. There are three groups with different size ratios under each scenario but having the same sample size, 30. The power rates for these three groups are assumed to be similar because of the same sample size. However, comparing the power rates in size ratio equals 1, we can see that the power rates in size ratio equal to 0.5 inflated and the power rates in size ratio equal to 2 deflate. So the cases of inflated Type I error rates also inflate the power rates.

Therefore, we came up with two conclusions in this part: 1. the MIXED model is reasonably robust to modest violations of the normal distribution; 2. when a large variance ratio (greater than 8) combines with heavily skew, the MIXED model can not be considered robust anymore. Nevertheless, it should not be a concern since the real data usually would not have such a big variance ratio; 3. When there is a small sample combining with large variance, it will cause serious Type I error inflation problems that need to be paid attention to.

Table 4.4. Type I Error Rate of Four Scenarios for Different Size Ratio $n_1/n_2 = (0.5, 1, 2)$, where the Total Group Size is 30; Unequal Variance as Giving Variance Ratio $\sigma^2_{\gamma1}/\sigma^2_{\gamma2}$=(2,4,6,8,10), where $\sigma^2_{\gamma2}$=1

$\sigma^2_{\gamma1}/\sigma^2_{\gamma2}$	Distribution Scenarios											
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew			$\mathbb{Z}w \sim$ Normal $e \sim$ Normal			$\mathbb{Z}w \sim$ Skew $e \sim$ Normal			$\mathbb{Z}w \sim$Skew $e \sim$skew		
	n_1/n_2			n_1/n_2			n_1/n_2			n_1/n_2		
	0.5	1	2	0.5	1	2	0.5	1	2	0.5	1	2
Treatment Results												
2	7.14	5.32	3.70	6.82	5.36	3.88	6.36	5.44	3.62	6.94	4.66	3.86
4	10.56	5.16	2.42	10.06	5.42	2.54	10.82	5.34	2.96	10.18	6.28	3.30
6	11.10	5.34	2.04	11.06	5.20	1.86	12.60	6.58	2.58	12.12	6.32	3.06
8	12.70	5.08	1.84	12.34	5.28	1.70	13.36	7.28	3.38	13.40	6.90	3.20
10	14.38	5.24	1.34	14.08	5.04	1.90	15.02	7.86	3.00	14.66	7.76	2.52
Period Results												
2	3.96	4.02	4.46	4.14	4.66	4.32	4.16	3.94	4.46	4.36	3.96	3.94
4	4.10	3.58	3.94	3.70	4.00	4.68	4.26	4.12	4.14	4.30	4.00	3.76
6	4.20	4.04	3.80	4.14	3.94	4.52	3.92	3.66	4.38	4.10	3.20	4.34
8	4.02	3.46	4.42	3.90	4.28	4.08	4.24	3.90	4.22	4.56	4.02	4.34
10	4.56	3.96	3.86	4.08	4.20	4.18	4.26	3.82	4.02	3.98	3.88	3.74
Treatment*Period Results												
2	3.98	3.84	4.50	4.62	4.00	3.84	3.90	3.96	4.18	3.78	3.94	3.78
4	3.94	3.72	3.76	4.16	4.14	3.86	4.38	4.24	3.88	4.00	3.80	3.74
6	4.16	3.70	4.06	4.12	4.12	4.32	4.56	4.40	4.22	3.86	3.66	4.02
8	4.02	3.66	3.78	4.70	4.34	3.94	4.02	3.66	3.98	3.94	3.52	4.38
10	3.68	3.44	3.72	3.74	4.04	4.06	3.74	4.30	4.20	4.24	3.56	3.76

Table 4.5. Power Rate with $H_a = (\mu, \tau_1, \tau_2, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$=(5,0,1.5,0,0.2,0.4,0.6) for Different Size Ratio $n_1/n_2 = (0.5, 1, 2)$, and Variance Ratio $\sigma^2_{\gamma1}/\sigma^2_{\gamma2}$=(2,4,6,8,10) in Four Scenarios, where $\sigma^2_{\gamma2}$=1

$\sigma^2_{\gamma1}/\sigma^2_{\gamma2}$	Distribution Scenarios											
	$\mathbb{Z}w \sim$ Normal, $e \sim$ Skew			$\mathbb{Z}w \sim$ Normal, $e \sim$ Normal			$\mathbb{Z}w \sim$ Skew, $e \sim$ Normal			$\mathbb{Z}w \sim$Skew, $e \sim$skew		
	n_1/n_2			n_1/n_2			n_1/n_2			n_1/n_2		
	0.5	1	2	0.5	1	2	0.5	1	2	0.5	1	2
Treatment Results												
2	72.90	75.38	69.84	74.18	76.26	68.08	75.56	75.94	69.30	74.16	76.08	69.32
4	62.84	60.72	49.48	60.32	60.34	49.48	63.10	62.78	54.44	62.60	63.04	51.70
6	54.48	49.40	36.60	53.08	48.90	37.76	54.62	54.14	42.68	54.38	53.50	43.20
8	48.80	42.72	28.32	42.34	42.34	27.76	47.32	48.84	37.74	47.08	49.22	37.34
10	44.02	37.12	22.14	44.80	37.44	22.30	41.24	44.80	32.58	42.30	44.30	32.44
Period Results												
2	58.72	62.98	56.78	55.80	60.90	55.86	57.56	61.94	55.10	57.66	64.86	56.56
4	56.50	61.82	56.04	54.24	61.48	53.54	56.24	60.88	53.70	57.02	61.74	55.78
6	57.52	61.38	56.14	54.82	58.86	52.70	54.30	60.24	52.28	56.68	61.12	54.56
8	56.18	60.56	53.02	59.36	59.36	53.10	53.22	58.74	52.90	56.32	60.52	54.88
10	55.18	60.94	53.42	53.00	58.82	51.94	53.06	58.76	52.38	56.14	60.50	54.04
Treatment*Period Results												
2	95.98	96.30	95.98	95.50	95.82	95.96	96.38	95.66	95.66	96.00	96.34	95.66
4	96.24	96.46	96.30	95.62	95.78	95.60	95.76	95.76	95.84	96.22	96.40	96.32
6	95.54	96.10	96.00	95.80	95.88	95.74	95.60	95.38	95.76	95.88	96.14	96.98
8	95.72	96.28	96.00	95.92	95.92	96.14	95.66	95.72	96.00	96.04	96.12	95.30
10	96.18	96.10	96.22	95.68	96.08	95.36	95.68	95.96	95.76	96.28	96.54	96.00

4.1.1. Second-Order Response Surface Models

Based on the results of Treatment Results in Table 4.4, we conducted the second-order response surface model for Type I error rates with different size ratios and variance ratios. The surface model had five independent variables: Variance Ratio, Size Ratio, Variance Ratio∗Size Ratio (interaction), Variance Ratio∗Variance Ratio, and Size Ratio∗Size Ratio.

According to the model outcome in Table 4.6, the P-value for all independent variables is significant (≤ 0.05) except for Variance Ratio*Variance Ratio, which means that Variance Ratio*Variance Ratio is not a significant variable. Meanwhile, the Size Ratio parameter estimate is -14.985996, and the parameter estimate for the Variance Ratio is 1.331589. Also, the Size Ratio's standard error is 1.466766, and the estimated value for the Variance Ratio is 0.254844. With large sample approximation, 2 times standard error was used to calculate the 95% confident interval, so the 95% confident interval for parameter estimate of Size Ratio is around (-17.86,-12.11), and the 95% confident interval for parameter estimate of Variance Ratio is around (0.83,1.83). Therefore, even though the variables of size ratio and variance ratio are both significant, the magnitude of the Size Ratio effect is much larger than the effect of Variance Ratio relative to the parameter estimates from the model. It seems that the size ratio may have more impact on the prediction of Type I error rate inflation comparing to the variance ratio, so balancing the sample size would be worth trying on the inflation problem of Type I error rate. The 3D surface plot is shown in Figure 4.2.

Table 4.6. Second Order Replace Surface Model Outcome

Parameter	DF	Estimate	Standard Error	t Value	Pr > $\|t\|$
Intercept	1	12.639718	1.066612	11.85	<.0001
Variance Ratio	1	1.331589	0.254844	5.23	<.0001
Size Ratio	1	-14.985996	1.466766	-10.22	<.0001
Variance Ratio*Variance Ratio	1	-0.024974	0.019480	-1.28	0.2053
Size Ratio*Variance Ratio	1	-0.633449	0.070449	-8.99	<.0001
Size Ratio*Size Ratio	1	5.184000	0.542210	9.56	<.0001

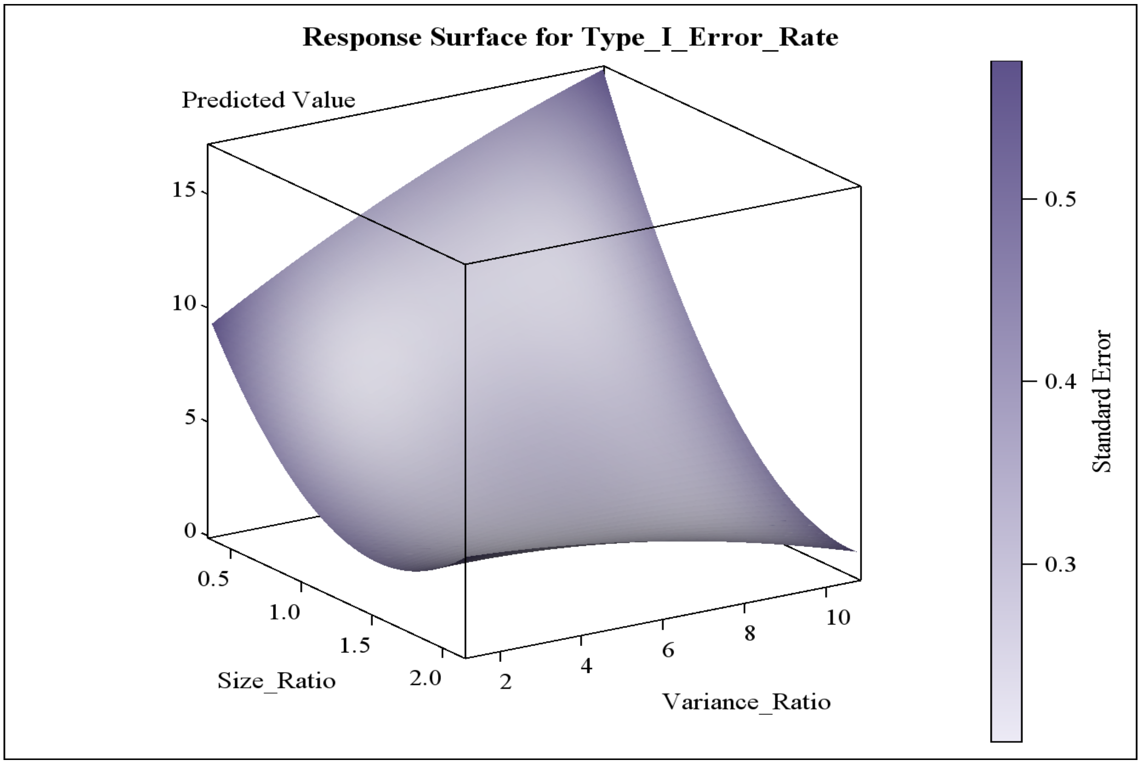

Figure 4.2. RSREG Model 3D Surface Plot for the Predicted Type I Error Rate

4.1.2. Stability of Type I Error Rates

For checking the stability of the Type I error rates, we would increase the sample sizes and the number of repeated time points to see if the Type I error rates keep the consistency, and power rates were provided as a reference.

4.1.2.1. Increasing Sample Sizes

In this section, we chose one scenario, $\mathbb{Z}w \sim$ Normal $e \sim$ Skew, as an example, and generated three different sample sizes under each size ratio. The result of Type I error rates is presented in Table 4.7, and the result of power rates is presented in Table 4.8. According to the two tables, we can see that the Type I error rates keep the same trend under the same size ratio. For example, when the size ratio equals to 0.5 and the variance ratio increases from 2 to 10, the Type I error rate increases from 6.86% to 14.38% regardless of the specific sample size of n_1 and n_2; Also, the Type I error rates have the similar values when they are in the same size ratio and variance ratio. For example, when size ratio equals to 0.5 and variance ratio equals to 10, the Type I error rate keeps around 13 no matter the difference of sample sizes: 13.18% when n_1:n_2=20:40, 12.98% when n_1:n_2=30:60, and

13.54% when n_1:n_2=40:80. It shows us that the Type I error rates across the same sample size ratio was very consistent. The power rates with $H_a = (\mu, \tau_1, \tau_2, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$=(5,0,1.5,0,0.2,0.4,0.6) are provided as a reference in Table 4.7. As we can expect, the larger the sample size, the higher the power. We believe this gives us more confidence that we can look at ratios of sample sizes and ratios of variances without too much concern for the actual sampling effort.

Table 4.7. Type I Error Rates for Four Different Sets of Sample Sizes under ZU∼MVN & e∼exp Scenario; Unequal Variance Ratio $\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$=(2,4,6,8,10), where σ^2_2=1

$\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$	Distribution Scenario: $\mathbb{Z}U$ ∼MVN e ∼skew											
	n_1/n_2=0.5				n_1/n_2=1				n_1/n_2=2			
	10:20	20:40	30:60	40:80	15:15	30:30	45:45	60:60	20:10	40:20	60:30	80:40
Treatment Results												
2	7.14	6.98	6.86	6.96	5.32	5.40	5.10	5.98	3.70	3.96	3.50	4.08
4	10.56	10.38	9.08	9.64	5.16	5.46	5.52	5.48	2.42	2.00	2.08	2.46
6	11.10	11.20	11.04	10.98	5.34	5.44	5.30	5.10	2.04	1.90	1.86	2.08
8	12.70	12.12	13.00	12.06	5.08	5.64	5.02	5.04	1.84	1.48	1.22	1.62
10	14.38	13.18	12.98	13.54	5.24	5.64	5.90	4.68	1.34	1.34	1.38	1.34
Period Results												
2	3.96	4.04	4.82	4.42	4.02	4.50	4.02	4.74	4.46	4.70	4.12	4.28
4	4.10	4.22	4.90	4.18	3.58	4.70	4.26	4.70	3.94	4.38	4.64	4.86
6	4.20	4.72	4.60	4.36	4.04	4.68	4.54	4.82	3.80	4.58	4.14	4.34
8	4.02	4.56	4.56	4.68	3.46	4.00	5.20	4.78	4.42	4.88	4.26	5.12
10 4.56	4.30	4.50	4.72	3.96	4.42	4.54	4.98	3.86	4.10	4.40	4.74	
Treatment*Period Results												
2	3.98	4.36	4.58	4.34	3.84	4.06	4.56	4.72	4.50	4.62	4.34	4.76
4	3.94	4.16	4.64	4.86	3.72	4.42	4.00	4.20	3.76	4.58	4.66	4.60
6	4.16	4.36	4.94	4.24	3.70	4.40	3.98	4.68	4.06	4.80	4.58	4.68
8	4.02	4.72	4.76	4.38	3.66	3.84	4.64	4.64	3.78	4.54	4.82	4.40
10	3.68	4.24	4.60	5.16	3.44	3.96	5.02	4.84	3.72	4.90	4.84	4.40

Table 4.8. Power Rates with $H_a = (\mu, \tau_1, \tau_2, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$=(5,0,1.5,0,0.2,0.4,0.6) for Four Different Sets of Sample Sizes under ZU∼MVN & e∼exp Scenario; Unequal Variance Ratio $\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$=(2,4,6,8,10), where σ^2_2=1

$\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$	Distribution Scenario: $\mathbb{Z}U$ ∼MVN e ∼skew											
	n_1/n_2=0.5				n_1/n_2=1				n_1/n_2=2			
	10:20	20:40	30:60	40:80	15:15	30:30	45:45	60:60	20:10	40:20	60:30	80:40
Treatment Results												
2	72.90	95.62	99.50	99.92	75.38	96.82	99.68	99.96	69.84	94.82	99.20	100.00
4	62.84	86.38	96.26	99.00	60.72	89.78	97.14	99.40	49.48	83.90	95.92	99.10
6	54.48	78.38	91.36	96.70	49.40	78.96	93.22	97.86	36.60	69.90	88.42	96.42
8	48.80	71.76	85.32	92.94	42.72	71.10	87.56	94.46	28.32	57.80	80.62	90.92
10	44.02	66.26	81.24	89.46	37.12	62.60	79.80	90.03	22.14	48.32	70.54	85.42
Period Results												
2	58.72	90.10	98.56	99.82	62.98	93.62	99.14	99.90	56.78	90.24	98.22	99.82
4	56.50	89.84	98.28	99.74	61.82	92.74	99.08	99.92	56.04	88.64	98.22	99.82
6	57.52	89.62	98.32	99.84	61.38	92.82	99.12	99.92	56.14	88.72	98.24	99.84
8	56.18	89.14	98.26	99.66	60.56	92.88	99.22	99.92	53.02	88.58	98.38	99.72
10	55.18	88.66	98.08	99.80	60.94	92.28	98.98	99.92	53.42	88.24	98.10	99.72
Treatment*Period Results												
2	95.98	95.14	95.36	95.72	96.30	95.44	96.04	95.70	95.98	95.10	95.46	95.36
4	96.24	95.40	95.60	95.24	96.46	95.96	95.48	95.02	96.30	94.94	95.20	95.04
6	95.54	95.04	95.62	95.48	96.10	94.98	95.86	95.20	96.00	95.46	95.92	94.90
8	95.72	95.66	95.74	95.02	96.28	95.84	95.64	95.54	96.00	95.34	95.44	95.18
10	96.18	95.90	95.34	95.20	96.10	95.40	95.40	95.38	96.22	95.90	95.00	94.88

4.1.2.2. Increasing Time Points

A real-world longitudinal study is likely more than four times of repeated measures, so the trend consistency of Type I error rates for a different number of times points is also essential. Therefore, in this section, we extended the number of time points from 4 to 6 to check Type I error rates. Based on Table 4.9, we can see that in the treatment test, Type I error rates inflated from 6.94% to 15.68% when the size ratio equals to 0.5, and deflated below 5.00 when the size ratio equals to 2. The trend is the same as Table 4.4 when repeated time points are 4; And in period and interaction tests, Type I error rates all stay below the limit of 5.00. The difference of Type I Error Rate from 6 time points to 4 time points is presented in Table 4.10. The values in Table 4.10 are around 0, which clearly shows us that the difference is quite small.

By adding two main treatment effects (0,1.5) and six-time effects (0,0,0.2,0.4,0.6,0.8), the corresponding power rates were obtained in Table 4.11. The power rates are very similar to Table 4.5 when repeated time points are 4. Specifically, in interaction tests, power rates are all above 95.00%; In period tests, power rates range from 57.06% to 73.48%; In treatment tests, power rates drop from the highest 78.90% to the lowest 22.52% when the variance ratio increased from 2 to 10.

We believe this also gives us more confidence that we can look at ratios of sample sizes and ratios of variances without too much concern for the number of repeated time points.

Table 4.9. Type I Error Rate of Four Scenarios with 6 Time Points for Different Size Ratio $n_1/n_2 = (0.5, 1, 2)$, where the Total Group Size is 30; Unequal Variance Ratio $\sigma^2_{\gamma1}/\sigma^2_{\gamma2}$=(2,4,6,8,10), where $\sigma^2_{\gamma1}/\sigma^2_{\gamma2}$=1

$\sigma^2_{\gamma1}/\sigma^2_{\gamma2}$	Distribution Scenarios											
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew			$\mathbb{Z}w \sim$ Normal $e \sim$ Normal			$\mathbb{Z}w \sim$ Skew $e \sim$ Normal			$\mathbb{Z}w \sim$Skew $e \sim$skew		
	n_1/n_2			n_1/n_2			n_1/n_2			n_1/n_2		
	0.5	1	2	0.5	1	2	0.5	1	2	0.5	1	2
Treatment Results												
2	7.42	5.46	3.92	7.50	4.92	4.16	7.66	5.58	3.52	6.94	5.80	3.96
4	9.92	6.04	2.26	9.32	5.98	2.26	9.66	6.08	3.40	10.40	5.92	3.22
6	11.56	5.34	1.92	12.24	5.20	1.70	13.18	6.78	2.84	13.22	6.84	3.42
8	13.32	5.72	1.56	12.98	6.36	1.58	14.86	7.02	3.14	13.88	7.16	3.54
10	14.12	5.40	1.36	14.64	6.06	1.52	15.68	7.30	3.20	14.38	7.44	3.06
Period Results												
2	3.84	3.70	3.92	3.82	4.24	4.22	3.72	4.46	3.96	3.98	3.96	4.06
4	4.16	3.30	4.10	3.62	4.06	3.98	3.88	4.40	3.80	4.28	3.86	4.02
6	4.60	3.72	4.14	4.04	4.38	4.02	4.32	3.90	4.18	3.70	4.06	4.46
8	3.94	3.74	4.04	4.02	4.36	3.78	3.88	4.38	3.70	4.08	3.48	3.90
10	4.12	3.90	3.96	3.24	3.86	3.86	4.24	4.14	3.98	4.00	3.64	4.20
Treatment*Period Results												
2	3.96	3.68	4.70	4.16	3.96	3.76	4.42	4.08	4.14	4.26	3.78	4.00
4	4.08	3.80	4.10	4.36	4.20	3.68	4.46	4.42	3.84	3.70	3.40	3.76
6	4.34	3.52	4.16	4.10	4.18	3.96	3.58	3.84	4.06	4.32	2.94	3.88
8	3.80	3.70	3.88	3.94	4.28	3.82	4.10	4.36	3.72	4.54	3.34	4.02
10	4.08	3.74	3.82	3.94	4.08	4.04	4.22	4.36	3.68	4.04	3.76	3.86

Table 4.10. The Difference of Type I Error Rate from 6 Time Points to 4 Time Points

$\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$	Distribution Scenarios											
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew			$\mathbb{Z}w \sim$ Normal $e \sim$ Normal			$\mathbb{Z}w \sim$ Skew $e \sim$ Normal			$\mathbb{Z}w \sim$Skew $e \sim$skew		
	n_1/n_2			n_1/n_2			n_1/n_2			n_1/n_2		
	0.5	1	2	0.5	1	2	0.5	1	2	0.5	1	2
Treatment Results												
2	0.28	0.14	0.22	0.68	-0.44	0.28	1.30	0.14	-0.10	0.00	1.14	0.10
4	-0.64	0.88	-0.16	-0.74	0.56	-0.28	-1.16	0.74	0.44	0.22	-0.36	-0.08
6	0.46	0.00	-0.12	1.18	0.00	-0.16	0.58	0.20	0.26	1.10	0.52	0.36
8	0.62	0.64	-0.28	0.64	1.08	-0.12	1.50	-0.26	-0.24	0.48	0.26	0.34
10	-0.26	0.16	0.02	0.56	1.02	-0.38	0.66	-0.56	0.20	-0.28	-0.32	0.54
Period Results												
2	-0.12	-0.32	-0.54	-0.32	-0.42	-0.10	-0.44	0.52	-0.50	-0.38	0.00	0.12
4	0.06	-0.28	0.16	-0.08	0.06	-0.70	-0.38	0.28	-0.34	-0.02	-0.14	0.26
6	0.40	-0.32	0.34	-0.10	0.44	-0.50	0.40	0.24	-0.20	-0.40	0.86	0.12
8	-0.08	0.28	-0.38	0.12	0.08	-0.30	-0.36	0.48	-0.52	-0.48	-0.54	-0.44
10	-0.44	-0.06	0.10	-0.84	-0.34	-0.32	-0.02	0.32	-0.04	0.02	-0.24	0.46
Treatment*Period Results												
2	-0.02	-0.16	0.20	-0.46	-0.04	-0.08	0.52	0.12	-0.04	0.48	-0.16	0.22
4	0.14	0.08	0.34	0.20	0.06	-0.18	0.08	0.18	-0.04	-0.30	-0.40	0.02
6	0.18	-0.18	0.10	-0.02	0.06	-0.36	-0.98	-0.56	-0.16	0.46	-0.72	-0.14
8	-0.22	0.04	0.10	-0.76	-0.06	-0.12	0.08	0.70	-0.26	0.60	-0.18	-0.36
10	0.40	0.30	0.10	0.20	0.04	-0.02	0.48	0.06	-0.52	-0.20	0.20	0.10

Table 4.11. Power Rate with $H_a = (\mu, \tau_1, \tau_2, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6)$=(5,0,1.5,0,0,0.2,0.4,0.6,0.8) for Different Size Ratio $n_1/n_2 = (0.5, 1, 2)$, and Variance Ratio $\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$=(2,4,6,8,10) in Four Scenarios, where $\sigma^2_{\gamma 2}$=1

$\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$	Distribution Scenarios											
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew			$\mathbb{Z}w \sim$ Normal $e \sim$ Normal			$\mathbb{Z}w \sim$ Skew $e \sim$ Normal			$\mathbb{Z}w \sim$Skew $e \sim$skew		
	n_1/n_2			n_1/n_2			n_1/n_2			n_1/n_2		
	0.5	1	2	0.5	1	2	0.5	1	2	0.5	1	2
Treatment Results												
2	75.72	78.78	72.34	76.02	78.90	72.50	77.82	78.14	72.10	78.74	77.50	71.18
4	62.38	63.16	51.16	62.96	61.98	51.34	66.04	65.22	55.26	67.12	64.04	54.32
6	54.56	50.92	37.04	55.34	51.82	38.90	55.70	54.96	44.42	56.06	53.94	44.14
8	49.24	42.86	28.16	50.54	44.24	28.52	48.30	49.98	37.22	48.30	48.94	37.28
10	45.88	37.60	22.52	46.32	37.20	23.28	42.94	45.80	32.92	41.38	44.00	34.00
Period Results												
2	66.98	73.48	66.38	66.06	71.84	64.18	66.92	71.88	64.44	67.88	73.42	65.16
4	65.62	70.96	62.40	63.90	68.90	61.58	63.48	69.90	63.12	66.36	70.30	64.08
6	63.66	68.94	62.12	62.50	67.30	59.90	64.56	69.08	61.90	65.72	69.66	62.90
8	62.62	68.00	60.82	62.04	66.90	59.26	62.50	67.20	59.54	62.46	68.44	60.84
10	62.02	66.68	59.26	61.38	65.62	57.50	61.48	67.50	57.06	62.22	67.06	58.60
Treatment*Period Results												
2	95.88	96.20	95.64	95.62	96.04	95.98	96.38	96.10	95.36	95.68	96.34	96.10
4	95.38	96.12	95.80	96.20	95.94	95.96	96.04	96.16	95.92	96.00	97.04	96.00
6	95.96	96.18	95.96	95.26	95.52	96.22	96.04	95.88	95.66	95.80	96.42	95.52
8	95.98	96.14	95.96	95.46	95.54	95.84	96.02	95.70	96.32	96.28	96.46	95.74
10	96.18	96.04	96.10	95.74	95.48	95.84	95.42	95.50	95.46	96.00	96.36	95.44

4.1.3. Null Distribution Testing

In this section, we would use Monte Carlo methods to get the sampling distributions for Table 4.4. There are two primary purposes:

1. To access the theoretical power rate from the null distribution, and compare it with observed power rate in Table 4.5;

2. To verify if the Type I error rates in Table 4.4 were simulated correctly.

Firstly, we simulated a null hypothesis distribution with 10,000 samples under $H_0 : (\mu, \tau_1, \tau_2, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$=(5,0,0,0,0,0,0) is true. After that, according to this empirical F value distribution from the 10,000 samples, we can find the 95% F critical value.

For obtaining the theoretical power rate, we compare the F critical value obtained from the empirical distribution above with with F values from each of the samples to determine the observed power rate in Table 4.5 where $H_a : (\mu, \tau_1, \tau_2, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$=(5,0,1.5,0,0.2,0.4,0.6) is true: if the F value in the case of observed power rate is greater or equal to the F critical value from null distribution, it is a rejection case (true case, which is power); otherwise, it is not a rejection case (false case). The theoretical power rates were calculated by total rejection cases divided by the number of times that the samples were replicated in the observed power rate (5,000), and the results are presented in Table 4.13.

Meanwhile, we subtracted the theoretical power rates in Table 4.13 from the observed power rates in Table 4.5 to obtain the difference in power rates; the result is presented in Table 4.14. When the size ratio or variance ratio is changed, there are no inflated/deflated trends of Type I error rates in Period Results and Treatment*Period (Interaction) Results in Table 4.4, so we would mainly focus on the part of Treatment Results in Table 4.14. Using size ratio equals to 1 as a reference, the difference in power rate increases from 4.44% to 27.64% when the size ratio equals to 0.5, and the difference in power rate decreases from -5.48% to -17.88% when the size ratio equals to 2. The difference of power rate in Treatment Results can also be visualized in the 3D surface plot of Figure 4.3. The difference between the theoretical and observed powers are greatest when the size ratio is 0.5 and the variance ratio is 10. When the size ratio is 2 and the variance ratio is high, then we see the observed power is worse than expected. This whole figure seems to mimic the Type I error rate inflation where the small size ratio/high variance ratio settings have the highest error inflation and the high size ratio/high variance ratio settings have clearly conservative results. It seems that the magnitude of the power inflation or deflation is greater than the type I error rate inflation or deflation.

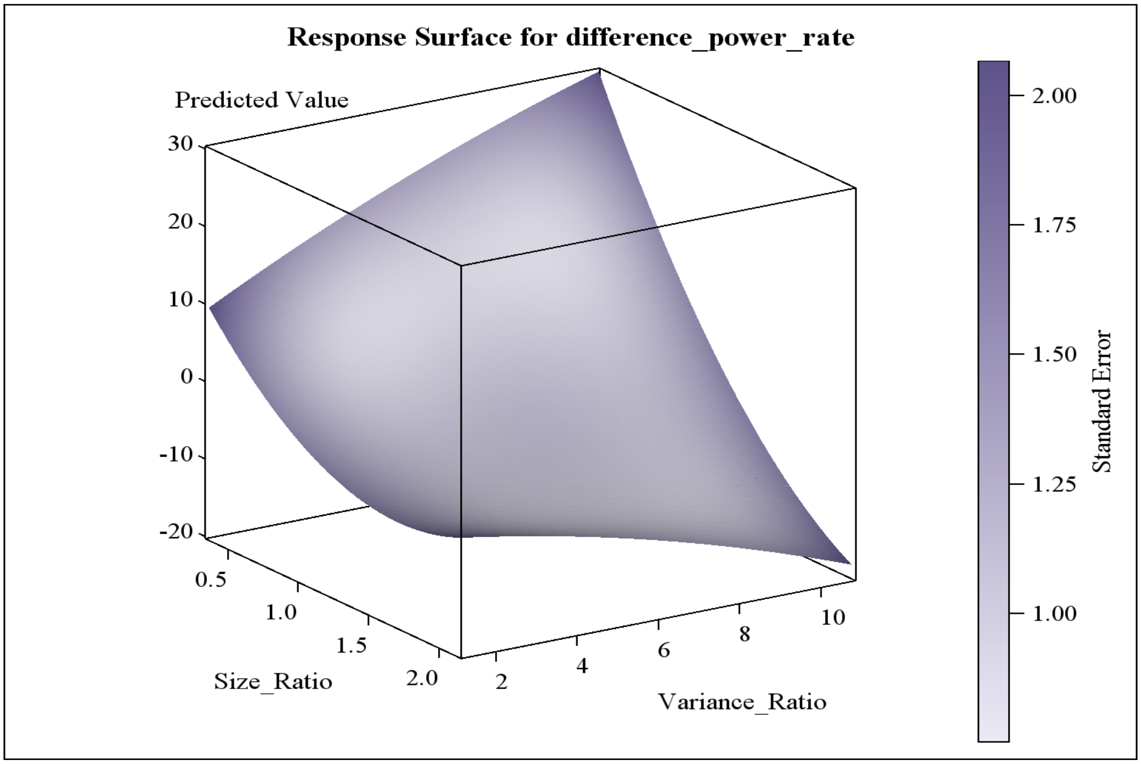

Figure 4.3. RSREG Model 3D Surface Plot for the Difference of Predicted Power Rates

After having the F critical value, we went back to check if the estimated Type I error rates under null distributions stays around 5.00%. We compared the F critical value with F values in the case of observed Type I error rate in Table 4.4 where $H_0 : (\mu, \tau_1, \tau_2, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$=(5,0,0,0,0,0,0) is true: if the F value in the case of observed Type I error rate is greater and equal to the F critical value, then it is a rejection case (false case, which is Type I error); otherwise, it is not a rejection case (true case). The estimated Type I error rate is obtained by the total rejection cases divided by the number of times that the samples were replicated in the cases of observed Type I error rate (5,000), and the result is presented in Table 4.12 in percentage form. In this table, all values are approaching 5.00%, proving that the Type I error rates in Table 4.4 were simulated correctly.

Table 4.12. The Estimated Type I Error Rate from Null Distribution for Different Size Ratio $n_1/n_2 = (0.5, 1, 2)$, where the Total Group Size is 30; Unequal Variance as Giving Variance Ratio $\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$=(2,4,6,8,10), where $\sigma^2_{\gamma 2}$=1

$\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$	Distribution Scenarios											
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew			$\mathbb{Z}w \sim$ Normal $e \sim$ Normal			$\mathbb{Z}w \sim$ Skew $e \sim$ Normal			$\mathbb{Z}w \sim$Skew $e \sim$skew		
	n_1/n_2			n_1/n_2			n_1/n_2			n_1/n_2		
	0.5	1	2	0.5	1	2	0.5	1	2	0.5	1	2
Treatment Results												
2	5.28	5.68	5.42	4.96	5.22	4.38	4.74	5.54	4.92	5.16	4.74	5.06
4	5.90	4.72	5.02	5.34	4.94	5.28	5.80	4.50	4.98	5.46	5.44	5.08
6	5.10	4.80	5.26	4.72	4.40	4.80	5.20	4.96	4.30	3.74	4.60	5.50
8	5.18	4.38	4.64	4.34	4.84	4.68	5.14	5.26	5.26	4.74	4.98	4.92
10	5.50	5.04	4.68	4.76	4.42	5.58	5.38	4.48	5.12	4.60	5.18	4.68
Period Results												
2	4.72	5.00	5.66	4.66	5.16	5.48	4.66	5.18	5.46	4.68	5.54	4.86
4	4.72	4.64	4.88	4.76	5.82	5.46	5.24	5.18	5.68	5.32	5.02	4.78
6	5.36	5.02	4.98	5.44	4.82	5.04	4.80	4.42	4.78	5.40	4.46	5.70
8	5.22	4.76	5.30	4.66	5.12	4.88	5.48	4.42	5.22	5.86	5.14	5.86
10	5.30	5.42	4.96	4.76	5.42	5.48	5.18	4.42	4.82	4.78	5.26	4.54
Treatment*Period Results												
2	5.10	5.42	5.84	5.12	5.22	5.40	4.70	4.84	4.76	4.70	4.58	4.72
4	4.66	5.70	4.72	5.14	4.68	4.42	5.44	5.46	4.80	4.94	4.82	4.44
6	5.12	4.80	4.70	5.16	4.86	4.64	5.30	5.14	4.98	5.02	4.72	4.84
8	5.04	5.02	5.22	6.16	5.34	4.92	4.72	4.16	4.78	4.92	5.04	5.54
10	4.56	4.34	4.58	4.30	4.32	5.14	5.24	5.08	4.92	5.42	4.76	4.74

Table 4.13. The Theoretical Power Rate with H_a : $(\mu, \tau_1, \tau_2, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$=(5,0,1.5,0,0.2,0.4,0.6) for Different Size Ratio n_1/n_2 = $(0.5, 1, 2)$, where the Total Group Size is 30; Unequal Variance as Giving Variance Ratio $\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$=(2,4,6,8,10), where $\sigma^2_{\gamma 2}$=1

$\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$	Distribution Scenarios											
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew			$\mathbb{Z}w \sim$ Normal $e \sim$ Normal			$\mathbb{Z}w \sim$ Skew $e \sim$ Normal			$\mathbb{Z}w \sim$Skew $e \sim$skew		
	n_1/n_2			n_1/n_2			n_1/n_2			n_1/n_2		
	0.5	1	2	0.5	1	2	0.5	1	2	0.5	1	2
Treatment Results												
2	67.84	76.16	74.68	69.34	75.60	73.56	70.80	76.24	73.42	69.72	76.26	72.92
4	50.94	59.38	60.54	47.32	58.52	62.94	48.86	60.36	60.52	47.42	60.72	59.32
6	38.92	47.14	53.52	37.24	46.60	51.42	33.28	50.08	49.62	31.74	48.06	52.52
8	31.64	40.88	45.66	29.26	40.86	44.16	22.90	43.32	46.54	22.08	43.64	45.70
10	26.44	36.60	38.66	25.22	35.06	40.18	16.80	37.76	41.54	14.66	37.10	40.36
Period Results												
2	61.46	66.86	61.58	58.62	63.28	59.88	59.78	66.88	58.40	58.82	69.46	59.98
4	59.86	65.84	59.64	58.54	66.96	56.66	58.64	64.72	59.18	59.74	65.48	59.06
6	62.16	64.84	61.24	59.48	62.30	55.88	58.00	64.56	54.56	60.52	65.90	58.56
8	60.14	65.62	57.64	56.48	62.98	56.28	57.58	61.22	57.32	60.86	65.22	60.16
10	58.62	66.08	57.06	55.40	62.70	57.02	57.12	62.84	54.94	58.66	65.48	58.66
Treatment*Period Results												
2	94.76	95.02	94.36	95.14	94.42	95.24	95.48	94.54	94.94	94.96	95.42	94.74
4	95.42	94.76	95.12	94.50	95.18	95.00	94.70	94.64	95.06	95.24	95.48	95.60
6	94.54	94.90	95.38	94.94	94.84	95.38	94.88	94.70	95.26	95.00	95.12	94.94
8	94.28	95.00	94.62	94.06	95.10	95.24	94.86	95.32	95.22	95.14	94.80	93.96
10	95.40	94.98	95.34	95.20	95.78	94.02	95.00	94.84	94.92	95.02	95.40	94.84

Table 4.14. The Difference of Power Rate (Observed Power Rate minus Theoretical Power Rate) with $H_a : (\mu, \tau_1, \tau_2, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$= (5,0,1.5,0,0.2,0.4,0.6) for Different Size Ratio $n_1/n_2 = (0.5, 1, 2)$, where the Total Group Size is 30; Unequal Variance as Giving Variance Ratio $\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$=(2,4,6,8,10), where $\sigma^2_{\gamma 2}$=1

$\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$	Distribution Scenarios											
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew			$\mathbb{Z}w \sim$ Normal $e \sim$ Normal			$\mathbb{Z}w \sim$ Skew $e \sim$ Normal			$\mathbb{Z}w \sim$Skew $e \sim$skew		
	n_1/n_2			n_1/n_2			n_1/n_2			n_1/n_2		
	0.5	1	2	0.5	1	2	0.5	1	2	0.5	1	2
Treatment Results												
2	5.06	-0.78	-4.84	4.84	0.66	-5.48	4.76	-0.30	-4.12	4.44	-0.18	-3.60
4	11.90	1.34	-11.06	13.00	1.82	-13.46	14.24	2.42	-6.08	15.18	2.32	-7.62
6	15.56	2.26	-16.92	15.84	2.30	-13.66	21.34	4.06	-6.94	22.64	5.44	-9.32
8	17.16	1.84	-17.34	13.08	1.48	-16.40	24.42	5.52	-8.80	25.00	5.58	-8.36
10	17.58	0.52	-16.52	19.58	2.38	-17.88	24.44	7.04	-8.96	27.64	7.20	-7.92
Period Results												
2	-2.74	-3.88	-4.80	-2.82	-2.38	-4.02	-2.22	-4.94	-3.30	-1.16	-4.60	-3.42
4	-3.36	-4.02	-3.60	-4.30	-5.48	-3.12	-2.40	-3.84	-5.48	-2.72	-3.74	-3.28
6	-4.64	-3.46	-5.10	-4.66	-3.44	-3.18	-3.70	-4.32	-2.28	-3.84	-4.78	-4.00
8	-3.96	-5.06	-4.62	2.88	-3.62	-3.18	-4.36	-2.48	-4.42	-4.54	-4.70	-5.28
10	-3.44	-5.14	-3.64	-2.40	-3.88	-5.08	-4.06	-4.08	-2.56	-2.52	-4.98	-4.62
Treatment*Period Results												
2	1.22	1.28	1.62	0.36	1.40	0.72	0.90	1.12	0.72	1.04	0.92	0.92
4	0.82	1.70	1.18	1.12	0.60	0.60	1.06	1.12	0.78	0.98	0.92	0.72
6	1.00	1.20	0.62	0.86	1.04	0.36	0.72	0.68	0.50	0.88	1.02	2.04
8	1.44	1.28	1.38	1.86	0.82	0.90	0.80	0.40	0.78	0.90	1.32	1.34
10	0.78	1.12	0.88	0.48	0.30	1.34	0.68	1.12	0.84	1.26	1.14	1.16

4.2. Type I Error Rate Inflation Problem

In this section, we would focus on one size ratio where there is more clearly a Type I error rate inflation problem caused by small sample size with large variance, so the information with only size rate equals to 0.5 (n_1/n_2=10:20) with four scenarios would be used. Meanwhile, since there is no Type I error rate inflation problem in the part of Periods Test and Treatment×Period (Interaction) Test, we would focus on Treatment Test only. We proposed four different methods as potential solutions. They are the rank test, aligned rank test, Sub-Sampling Method, and MIXED model incorporating GROUP= option under REPEAT statement. Sub-Sampling Method here is to randomly drop ten subjects from the large group (n_2), so the two groups would have an equal sample size (n_1/n_2=10:10). The corresponding power rates were also investigated as a reference. Besides the group test for two-levels of the treatment group, we would also do the group tests for three-level treatment groups.

4.2.1. Methods for Two Levels of Treatment Groups

According to Type I error rates of Table 4.15, we can see that Rank Test and Aligned Rank Test have a similar performance as of Original Test on Type I error rate inflation: for the first two scenarios (normal or slightly skewed distribution), the two methods reduce the inflation a little, but not very helpful; for the last two scenarios (heavily skew distribution), the inflation problem gets worse. Nevertheless, when it comes to the Sub-Sampling Method and MIXED model using GROUP= option method, the inflated Type I error rates in both methods drop substantially compared to the Original Test: In the first two scenarios (normal or slightly skew distribution), two methods both have excellent performance, keeping Type I error rates between 5.10% and 6.10%; In the last two scenarios (heavily skewed distribution), the Type I error rates stay between 5.52% and 8.59%, not as good as the first two scenarios.

Among the four methods, the Sub-Sampling and MIXED model using GROUP= option methods have the best performance on Type I error inflation problem. More specifically, when data has better behaved (normal or slightly skew), the two method of GROUP= option tends to have a quite close performance on Type I error rate. For example, when variance ratio equals to 2 and $\mathbb{Z}w$ & e follow both Normal distribution, the Type I error rate is 5.20% in Sub-Sampling Method and is 5.10% in the MIXED model using GROUP= option Method. But when data has poorly behaved (heavily skew), the two methods have Type I error rate inflation compared to the performance when data is better behaved. Also, the MIXED model using the GROUP= option seems to have a little bit more inflated Type I error rate than the Sub-Sampling method. For example, when variance ratio equals to 10 and $\mathbb{Z}w$ & e follow both skew distribution, the Type I error rate is 7.94% in Sub-Sampling method and is 8.53% in the MIXED model using GROUP= option Method. So the presentation of heavily skew would cause a little bit of inflation on Type I error rate, especially for the MIXED model using the GROUP= option method.

The corresponding power rates with the treatment effect added as $H_a = (\tau_1, \tau_2)$=(0,1.5) is presented in Table 4.16. As we can see, the Rank Test and Aligned Rank Test have a very similar performance on power rate, and the Sub-Sampling Method and the MIXED model method using the GROUP= option have a very similar performance on power rate, too. For example, when variance ratio equals to 2 and $\mathbb{Z}w$ & e both follow the normal distribution, the power rate is 71.68% in Rank Test, and 70.76% in Aligned Rank Test. Comparing to the Original Test, the Sub-Sampling Method and the MIXED model method using the GROUP= option lost lots of power rate, nearly 20 percentage within each case. For example, when variance ratio equals to 2 and $\mathbb{Z}w$ & e both follow the normal distribution, the power rate is 74.18% in Original Test, and 58.52% in Sub-Sampling Method, there is a 15.66% percentage power rate lost. Partly for this reason, power rate in Original Test is inflated by the inflated Type I error rate.

Meanwhile, in Original Test, power falls from 72.90% (the lowest Type I error rate when variance ratio equals to 2) to 44.80% (the highest Type I error rate when variance ratio equals to 10) as the variance ratio increases from 2 to 10. For increasing the power value when the variance ratio is high, we standardize the effect size by adding an effect size η=$\sqrt{\frac{\sigma^2_{\gamma 1}+\sigma^2_{\gamma 2}}{2}}$ to one of the treatment groups, which changes the H_a to (0,η). According to Table 4.17, when increasing the variance ratio from 2 to 10, the power rates increase from 56.14% (the lowest Type I error rate when variance ratio equals 2) to 81.08% (the highest Type I error rate when variance ratio equals 10). It shows us that a standardized effect size (one that tries to account for the increasing variances by increasing the absolute magnitude effect size) will yield higher power.

Table 4.15. Methods for Type I Error Rates Inflation under Treatment Test with Fixed Size Ratio n_1/n_2=0.5 and $\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$=(2,4,6,8,10), where $\sigma^2_{\gamma 2}$=1

$\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$	Distribution Scenarios			
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew	$\mathbb{Z}w \sim$ Normal $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Skew
Original Test				
2	7.14	6.82	6.36	6.94
4	10.56	10.06	10.82	10.18
6	11.10	11.06	12.60	12.12
8	12.70	12.34	13.36	13.40
10	14.38	14.08	15.02	14.66
Rank Test				
2	6.52	6.16	6.46	7.62
4	9.32	8.32	10.66	11.94
6	9.38	9.16	13.50	15.48
8	10.12	9.90	16.08	16.74
10	11.66	10.30	18.24	19.32
Aligned Rank Test				
2	6.62	6.18	6.36	7.66
4	9.32	8.34	10.64	11.70
6	9.46	9.10	13.46	15.40
8	10.02	9.88	16.04	16.76
10	11.72	10.34	17.74	19.10
Sub-Sampling Method				
2	5.64	5.20	4.90	5.36
4	6.36	5.80	6.54	6.18
6	5.28	5.56	7.06	6.92
8	5.60	6.02	7.12	7.42
10	5.68	6.10	7.96	7.94
MIXED model using the GROUP= option				
2	5.62	5.10	5.52	6.44
4	5.52	5.68	7.32	7.36
6	5.22	5.19	7.95	7.98
8	5.14	5.28	8.09	8.12
10	5.23	5.36	8.59	8.53

Table 4.16. Methods for Power Rates with $H_a : (\tau_1, \tau_2)$=(0,1.5) under Treatment Test with Fixed Size Ratio n_1/n_2=0.5 and $\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$=(2,4,6,8,10), where $\sigma^2_{\gamma 2}$=1

$\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$	Distribution Scenarios			
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew	$\mathbb{Z}w \sim$ Normal $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Skew
Original Test				
2	72.90	74.18	75.56	74.16
4	62.84	60.32	63.10	62.60
6	54.48	53.08	54.62	54.38
8	48.80	42.34	47.32	47.08
10	44.02	44.80	41.24	42.30
Rank Test				
2	72.32	71.68	77.32	81.74
4	57.78	54.36	58.78	62.18
6	48.20	45.68	44.04	45.36
8	41.42	40.58	34.08	33.96
10	37.06	38.02	25.64	25.58
Aligned Rank Test				
2	72.14	70.76	76.76	81.94
4	57.82	54.12	58.30	62.22
6	47.58	45.40	43.76	45.08
8	40.98	40.18	33.58	33.28
10	36.88	37.70	25.26	24.90
Sub-Sampling Method				
2	57.72	58.52	59.06	59.92
4	44.74	41.70	44.24	43.54
6	36.48	34.28	32.68	32.68
8	30.54	29.46	25.44	25.38
10	26.16	25.88	19.42	19.84
MIXED model using the GROUP= option				
2	65.38	66.61	69.22	69.46
4	47.72	45.38	47.82	45.23
6	36.87	34.77	32.33	31.34
8	29.33	29.22	23.16	22.89
10	24.70	25.24	17.21	17.19

Table 4.17. Power Rates with H_a : (τ_1, τ_2)=(0,η) under Treatment Test with Fixed Size Ratio n_1/n_2=0.5 and $\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$=(2,4,6,8,10), where $\sigma^2_{\gamma 2}$=1 and $\eta=\sqrt{\frac{\sigma^2_{\gamma 1}+\sigma^2_{\gamma 2}}{2}}$

$\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$	Distribution Scenarios			
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew	$\mathbb{Z}w \sim$ Normal $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Skew
Original Test with H_a : (τ_1, τ_2)=(0,1.5)				
2	72.90	74.18	75.56	74.16
4	62.84	60.32	63.10	62.60
6	54.48	53.08	54.62	54.38
8	48.80	42.34	47.32	47.08
10	44.02	44.80	41.24	42.30
Original Test with H_a : (τ_1, τ_2)=(0,η)				
2	58.62	56.14	59.02	58.42
4	65.36	65.32	67.48	69.08
6	69.16	69.24	74.56	73.38
8	71.36	72.00	77.04	77.96
10	74.02	73.38	80.18	81.08

4.2.2. Methods for Three Levels of Treatment Group

This section increased the treatment group from two-level to three-level to check the Type I error rates and power rates in different methods. The three treatment groups are treatment group A, treatment group B, and control group C. The total sample size is 40. The parameter sets of sample size and variance for the three groups were listed in Table 4.18.

Table 4.18. The Parameters Set for Sample Size and Variances of Three Groups

Treatment Groups	Symbol	Sample size	Variances
Treatment A	γ_A	10	(2,6)
Treatment B	γ_B	10	(2,4,6,8,10)
Control Group C	γ_C	20	1

Specifically, the three-level of treatment group has 10 subjects for group A, 10 subjects for group B, and 20 subjects in group C. Sub-Sampling Method here is to drop 10 subjects from group C randomly, so all three treatment groups would have an equal sample size (10 subjects per group). Since Rank Test and Aligned Rank Test did not have a good performance on Type I error rate inflation problem in the previous section, Section 4.2.1, only the Sub-Sampling and MIXED model methods using the GROUP= option would be presented in this section.

According to the findings in Section 4.1, we already know that when we have a small sample size with large variance, there is a Type I error rate inflation problem. The Type I error rate under the three-level treatment group test is presented in Table 4.19. No matter $\sigma^2_{\gamma_A}$ equals to 2 or 6 when $\sigma^2_{\gamma_B}$ goes from 2 to 10, Type I error rate increased substantially in Original Test, from 6.68% to 12.84%. Nevertheless, the Sub-Sampling method and the MIXED model using the GROUP= option method produce a significant decline in the Type I error rate inflation problem. Specifically, when data is well behaved (normal or slightly skew), Sub-Sampling Method and MIXED model using GROUP= option method both have good performance for keeping Type I error rates between 4.90% and 6.38%, and MIXED model using GROUP= option tends to have more conservative Type I error rate (between 4.90% and 5.65%) than Sub-Sampling method (between 4.96% and 6.38%). However, when data is poorly behaved (heavily skew), the two methods both have a little bit inflated Type I error rate comparing to the performance when data is well behaved, especially for the method of the MIXED model using GROUP= option: the Type I error rates in Sub-Sampling method goes from 5.04% to 8.42%, and the Type I error rates in the MIXED model using GROUP= option method increased from 6.47% to 10.19%.

The corresponding power rates with treatment effect with $H_a = (\tau_1, \tau_2, \tau_3)$=(0,0,1.5) is presented in Table 4.20. As we can see, the Sub-Sampling Method and MIXED model using the GROUP= option method have a very similar performance on power rate. Specifically, when $\mathbb{Z}w$ & e both follow normal distribution, and $\sigma^2_{\gamma_A}$=2 and $\sigma^2_{\gamma_B}$=2, the power rate is 55.08% in Sub-Sampling Method, and 55.30% in the method of MIXED model using GROUP= option.

Meanwhile, in Original Test, power falls from 65.72% (the lowest Type I error rate when $\sigma^2_{\gamma_A}$=2 and $\sigma^2_{\gamma_B}$=2) to 40.08% (the highest Type I error rate when $\sigma^2_{\gamma_A}$=2 and $\sigma^2_{\gamma_B}$=10) as $\sigma^2_{\gamma_B}$ increases from 2 to 10. For increasing the power value, when the variance ratio is high, we standardize the effect size by adding an effect size $\eta=\sqrt{\frac{\sigma^2_{\gamma_A}+\sigma^2_{\gamma_B}+\sigma^2_{\gamma_C}}{3}}$ to one of the treatment groups, which changes the H_a to (0,0,η). According to Table 4.21, under $\sigma^2_{\gamma_A}$ equals to 2, when increasing $\sigma^2_{\gamma_B}$ from 2 to 10, the power rates keeps from 53.74% (the lowest Type I error rate at $\sigma^2_{\gamma_B}$ equals 2) to 64.84% (the highest Type I error rate at $\sigma^2_{\gamma_B}$ equals 10). These results provide a consistent conclusion with the results from the two-level treatment group: standardizing the effect size would yield the power rates for the higher variance ratio.

According to both the results from the two-level treatment group and three-level treatment group, the MIXED model using GROUP= option and Sub-Sampling method have a very similar performance on Type I error rate inflation problem and power rates, even though there is still a little difference depending on if data is well behaved or if group variance difference is small. Specifically, if data is well behaved, two methods get a similar Type I error rate; If data has poorly behaved, the MIXED model method using GROUP= option gets a little more inflated Type I error rate (about 1 to 3 percentage higher) than the Sub-Sampling Method. In general, the MIXED model method using the GROUP= option is recommended for well-behaved data from a practical standpoint since it tries to use all of the data. A method excluding data like Sub-Sampling Methods might be considered an 'interesting academic' result, but not practical. Because the real data is hard to obtain in the first place, and the method would not likely be received as a reliable solution if excluding lots of data.

Table 4.19. Methods for Type I Error Inflation Rates under Treatment Test, where $\sigma^2_{\gamma_A}$=(2,6), $\sigma^2_{\gamma_B}$=(2,4,6,8,10), and σ^2_C=1

$\sigma^2_{\gamma_B}$	$\sigma^2_{\gamma_A}$=2				$\sigma^2_{\gamma_A}$=6			
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew	$\mathbb{Z}w \sim$ Normal $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Skew	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew	$\mathbb{Z}w \sim$ Normal $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Skew
Original Test								
2	6.68	7.18	6.80	6.68	10.00	9.22	9.74	9.76
4	8.62	8.64	8.64	8.22	10.10	9.92	9.74	9.20
6	9.68	10.08	9.60	10.24	10.12	10.90	10.44	10.38
8	10.08	10.68	11.30	11.32	12.10	10.72	11.48	10.84
10	11.64	10.94	12.36	12.84	11.70	11.10	11.52	12.52
Sub-Sampling Method								
2	4.96	5.50	5.18	5.04	6.42	5.50	6.74	7.16
4	5.20	5.84	5.88	5.50	5.68	5.76	5.74	5.08
6	5.40	6.50	6.26	7.04	5.84	5.96	5.70	5.38
8	6.16	6.48	7.06	7.24	6.34	5.62	6.58	5.66
10	6.38	6.20	7.76	8.42	6.16	6.34	6.28	7.00
MIXED model using the GROUP= option								
2	4.90	5.76	6.47	6.72	5.44	5.36	7.36	8.30
4	5.50	5.80	7.18	7.46	5.18	5.76	7.74	8.27
6	4.82	5.86	7.17	7.88	5.11	5.92	9.18	8.77
8	5.20	5.69	8.01	8.50	5.08	5.53	9.47	9.26
10	4.69	5.19	8.33	8.63	5.52	5.65	9.34	10.19

Table 4.20. Methods for Power Rates with $H_a:(\tau_1,\tau_2,\tau_3)$=(0,0,1.5) under Treatment Test, where $\sigma^2_{\gamma_A}$=(2,6), $\sigma^2_{\gamma_B}$=(2,4,6,8,10), and σ^2_C=1

$\sigma^2_{\gamma_B}$	$\sigma^2_{\gamma_A}$=2				$\sigma^2_{\gamma_A}$=6			
	$\mathbb{Z}w\sim$ Normal $e\sim$ Skew	$\mathbb{Z}w\sim$ Normal $e\sim$ Normal	$\mathbb{Z}w\sim$ Skew $e\sim$ Normal	$\mathbb{Z}w\sim$ Skew $e\sim$ Skew	$\mathbb{Z}w\sim$ Normal $e\sim$ Skew	$\mathbb{Z}w\sim$ Normal $e\sim$ Normal	$\mathbb{Z}w\sim$ Skew $e\sim$ Normal	$\mathbb{Z}w\sim$ Skew $e\sim$ Skew
Original Test with Treatment Effect H_a=(0,0,1.5)								
2	65.72	65.84	66.14	68.68	54.38	54.36	57.30	56.52
4	55.60	55.90	56.82	57.04	48.02	48.62	48.40	50.40
6	49.64	48.72	49.60	49.54	43.56	43.16	44.28	43.30
8	43.14	45.34	43.16	41.72	39.94	40.10	39.18	38.56
10	39.34	40.08	38.26	38.36	36.22	36.52	34.40	36.22
Sub-Sampling Method with Treatment Effect H_a=(0,0,1.5)								
2	55.68	55.08	56.68	59.30	38.06	39.18	44.34	44.32
4	44.90	44.92	44.64	45.72	33.26	34.08	36.56	37.04
6	38.58	36.52	36.58	36.44	30.20	28.88	31.14	29.68
8	32.24	33.60	29.54	28.56	26.74	27.82	26.28	25.44
10	29.30	29.12	23.94	24.94	24.28	23.90	21.24	23.34
MIXED model using the GROUP= option with Treatment Effect H_a=(0,0,1.5)								
2	56.09	55.30	60.83	63.90	53.13	54.42	59.18	60.53
4	38.24	38.58	37.23	38.86	35.65	35.95	38.14	38.15
6	29.21	28.51	25.46	25.25	27.95	28.23	26.24	25.47
8	23.21	24.14	18.77	17.39	22.30	23.24	19.07	18.12
10	19.49	20.03	13.28	14.22	18.33	18.90	14.71	15.73

Table 4.21. Power Rates with H_a : (τ_1, τ_2, τ_3)=(0,0,η) under Treatment Test, where $\sigma^2_{\gamma_A}$=(2,6), $\sigma^2_{\gamma_B}$=(2,4,6,8,10), σ^2_C=1, and $\eta=\sqrt{\frac{\sigma^2_{\gamma_A}+\sigma^2_{\gamma_B}+\sigma^2_{\gamma_C}}{3}}$

$\sigma^2_{\gamma_B}$	$\sigma^2_{\gamma_A}$=2				$\sigma^2_{\gamma_A}$=6			
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew	$\mathbb{Z}w \sim$ Normal $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Skew	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew	$\mathbb{Z}w \sim$ Normal $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Normal	$\mathbb{Z}w \sim$ Skew $e \sim$ Skew
Original Test with Treatment Effect H_a=(0,0,1.5)								
2	65.72	65.84	66.14	68.68	54.38	54.36	57.30	56.52
4	55.60	55.90	56.82	57.04	48.02	48.62	48.40	50.40
6	49.64	48.72	49.60	49.54	43.56	43.16	44.28	43.30
8	43.14	45.34	43.16	41.72	39.94	40.10	39.18	38.56
10	39.34	40.08	38.26	38.36	36.22	36.52	34.40	36.22
Original Test with Treatment Effect H_a=(0,0,η)								
2	55.98	54.60	53.74	56.04	66.26	67.06	68.16	66.52
4	56.90	58.38	58.26	58.16	67.44	66.98	67.52	69.16
6	58.34	59.76	60.80	61.28	66.38	66.20	70.02	68.20
8	60.04	60.40	63.52	63.48	65.76	66.34	69.94	69.30
10	59.60	61.60	64.84	64.36	66.82	65.94	69.96	71.30

4.2.3. Other Failed Methods

Besides the four methods above, we also tried the variation of Bootstrap method. The variation of Bootstrap method aims to balance the two group sample size (15 subjects in the treatment group vs. 15 subjects in the control group), though the original sample size (30 subjects) keeping the same. To recall, in this study, there are 5000 simulated samples, and each sample has 30 subjects. Initially, there are 10 subjects in the treatment group and 20 in the control group.

Two ways are used the variation of Bootstrap method to boost the treatment group size from 10 to 15. The first is to randomly select 5 subjects with replacement from the original 10 subjects in the treatment group, then add them to the original treatment group. The second is to randomly select 15 subjects with replacement from the original 10 subjects in the treatment group and treat the new 15 subjects as a new treatment group. There are also two ways to decrease the control group size from 20 to 15: the first is to randomly drop 5 subjects without replacement from the original 20 subjects in the control group; the second is to randomly select 15 subjects with replacement from the original 20 subjects in the control group, and treat them as a new control group. The subsets we created in the treatment and control group can build up to four different sample data-sets. However, none of them can decrease the inflated Type I error rate, and some even make the inflation problem worsen.

4.3. Five Different Covariance Structures

In this paper, we mainly focus on two problems presented in the previous two sections: Section 4.1 and Section 4.2. We also want to provide a general idea of how the incorrect covariance structure affects the Type I error rate and power rate in the treatment test.

To recall, First-Order Autoregressive (AR(1)) [Littell et al., 2006, Kincaid, 2005] was chosen as the correct covariance structure when we generated these datasets. After these datasets were generated, five most common covariance structure [Kincaid, 2005, Littell et al., 2006] (First-Order Autoregressive (AR(1)), Toeplitz (TOEP), Compound Symmetry (CS), Unstructured (UN), Variance Components (VC)) was applied under the REPEATED statement in PROC MIXED to run the test, respectively. Therefore, there are five different Type I error rates, and power rates corresponding to each covariance structure were obtained in every situation. Nevertheless, only the results of Type I error rates, and power rates under First-Order Autoregressive (AR(1)) structure [Kincaid, 2005, Littell et al., 2006] is the correct one, which would be used as the reference for the results under other covariance structures. To sum up, according to all the tables listed below, the Type I error rates and power rates under Toeplitz (TOEP), Compound Symmetry (CS), and Unstructured (UN) are very similar to the results under First-Order Autoregressive (AR(1)) structure: the difference among the four is around 1 percent. However, the results under Variance Components (VC) structure have the worst results among the five: the Type I error rate is 3 to even 17 times than the Type I error rate under First-Order Autoregressive (AR(1)) structure, and the power rate is 20 to 50 percent higher than the power rate under First-Order Autoregressive (AR(1)) structure.

The main tables in the previous sections are presented under 5 different covariance structures in this section, and the guidance for the comparison is listed as below: Table 4.22 is the Type I error rate under 5 different covariance structures of four scenarios for three different size ratios ($n_1/n_2 = (0.5, 1, 2)$) and five different variance ratios ($\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2} = (2, 4, 6, 8, 10)$), which corresponding to the Treatment Test in Table 4.4; Table 4.23 is the power rate under 5 different covariance structures of four scenarios for three different size ratios ($n_1/n_2 = (0.5, 1, 2)$) and five different variance ratios ($\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2} = (2, 4, 6, 8, 10)$), which corresponding to the Treatment Test in Table 4.5; Table 4.24 is the methods of Type I error rate in Treatment Test under 5 different covariance structures of four scenarios for the fixed size ratio ($n_1/n_2 = 0.5$) and five different variance ratios ($\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2} = (2, 4, 6, 8, 10)$), which corresponding to Table 4.15; Table 4.25 is the methods of power rate with H_a=(0,1.5) in Treatment Test under 5 different covariance structures of four scenarios for the fixed size ratio ($n_1/n_2 = 0.5$) and five different variance ratios ($\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2} = (2, 4, 6, 8, 10)$), which corresponding to Table 4.16; Table 4.26 is the methods of Type I error rates for three treatment groups in Treatment Test under 5 different covariance structures of four scenarios for the fixed sample size($n_A = 10$; $n_B = 10$; $n_C = 20$) and the group variance ratios ($\sigma^2_{\gamma_A} = 2; \sigma^2_{\gamma_B} = (2, 4, 6, 8, 10); \sigma^2_{\gamma_C} = 1$), which corresponding to the part $\sigma^2_{\gamma_A} = 2$ in Table 4.19; Table 4.27 is the methods of power rate with H_a:(0,0,1.5) for three treatment groups in Treatment Test under 5 different covariance structures of four scenarios for the fixed sample size($n_A = 10$; $n_B = 10$; $n_C = 20$) and the group variance ratios ($\sigma^2_{\gamma_A} = 2; \sigma^2_{\gamma_B} = (2, 4, 6, 8, 10); \sigma^2_{\gamma_C} = 1$), which corresponding to the part $\sigma^2_{\gamma_A} = 2$ in Table 4.20; Table 4.28 is the methods of Type I error rates for three treatment groups in Treatment Test under 5 different covariance structures of four scenarios for the fixed sample size($n_A = 10$; $n_B = 10$; $n_C = 20$) and the group variance ratios ($\sigma^2_{\gamma_A} = 6; \sigma^2_{\gamma_B} = (2, 4, 6, 8, 10); \sigma^2_{\gamma_C} = 1$), which corresponding to the part $\sigma^2_{\gamma_A} = 6$ in Table 4.19; Table 4.29 is the methods of power rate with H_a:(0,0,1.5) for three treatment groups in Treatment Test under 5 different covariance structures of four scenarios for the fixed sample size($n_A = 10$; $n_B = 10$; $n_C = 20$) and the group variance ratios ($\sigma^2_{\gamma_A} = 6; \sigma^2_{\gamma_B} = (2, 4, 6, 8, 10); \sigma^2_{\gamma_C} = 1$), which corresponding to the part $\sigma^2_{\gamma_A} = 6$ in Table 4.20.

Table 4.22. Type I Error Rate in Treatment Test under 5 Covariance Structures of Four Scenarios for Different Size Ratio $n_1/n_2 = (0.5, 1, 2)$, where the Total Group Size is 30; Unequal Variance as Giving Variance Ratio $\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$=(2,4,6,8,10), where $\sigma^2_{\gamma 2}$=1

$\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$	Distribution Scenarios											
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew			$\mathbb{Z}w \sim$ Normal $e \sim$ Normal			$\mathbb{Z}w \sim$ Skew $e \sim$ Normal			$\mathbb{Z}w \sim$Skew $e \sim$skew		
	n_1/n_2			n_1/n_2			n_1/n_2			n_1/n_2		
	0.5	1	2	0.5	1	2	0.5	1	2	0.5	1	2
First-Order Autoregressive (AR(1))												
2	7.14	5.32	3.70	6.82	5.36	3.88	6.36	5.44	3.62	6.94	4.66	3.86
4	10.56	5.16	2.42	10.06	5.42	2.54	10.82	5.34	2.96	10.18	6.28	3.30
6	11.10	5.34	2.04	11.06	5.20	1.86	12.60	6.58	2.58	12.12	6.32	3.06
8	12.70	5.08	1.84	12.34	5.28	1.70	13.36	7.28	3.38	13.40	6.90	3.20
10	14.38	5.24	1.34	14.08	5.04	1.90	15.02	7.86	3.00	14.66	7.76	2.52
Toeplitz (TOEP)												
2	6.72	5.26	3.56	6.52	5.18	3.70	6.16	5.22	3.42	6.60	4.34	3.58
4	10.04	5.00	2.32	9.86	5.24	2.46	10.54	5.14	2.52	9.74	6.02	3.16
6	10.74	5.10	2.04	10.78	4.98	1.78	12.42	6.40	2.84	11.62	6.04	2.84
8	12.38	4.96	1.72	12.14	5.12	1.62	13.00	7.06	3.12	13.00	6.68	2.96
10	13.80	5.16	1.28	13.76	5.00	1.86	14.60	6.82	2.94	14.34	7.58	2.48
Compound Symmetry (CS)												
2	6.86	5.28	3.62	6.50	5.30	3.68	6.10	5.06	3.46	6.80	4.38	3.84
4	10.16	5.10	2.20	9.92	5.32	2.32	10.56	5.10	2.54	9.84	6.04	3.16
6	10.88	5.18	1.96	11.06	4.92	1.78	12.16	6.50	2.84	11.72	5.96	2.84
8	12.38	5.02	1.68	12.04	5.08	1.60	12.80	7.02	3.18	13.12	6.84	3.04
10	13.92	5.18	1.26	13.64	5.04	1.84	14.56	6.86	2.90	14.36	7.68	2.50
Unstructured (UN)												
2	6.86	5.28	3.62	6.50	5.30	3.68	6.10	5.06	3.46	6.80	4.38	3.84
4	10.16	5.10	2.20	9.92	5.32	2.32	10.56	5.10	2.54	9.84	6.04	3.16
6	10.88	5.18	1.96	11.06	4.92	1.78	12.16	6.50	2.84	11.72	5.96	2.84
8	12.38	5.02	1.68	12.04	5.08	1.60	12.80	7.02	3.18	13.12	6.84	3.04
10	13.92	5.18	1.26	13.64	5.04	1.84	14.56	6.86	2.90	14.36	7.68	2.50
Variance Components (VC)												
2	33.64	30.16	26.26	32.36	30.44	26.92	33.16	30.74	28.00	34.06	30.44	27.60
4	39.74	30.82	23.52	37.92	31.28	24.84	40.00	32.52	22.74	39.60	32.74	23.90
6	10.30	31.36	22.22	39.42	30.64	21.52	41.72	32.86	23.58	42.72	31.40	23.00
8	42.66	31.66	20.90	41.84	31.52	22.04	44.88	33.34	23.30	44.20	32.78	21.80
10	45.62	32.26	19.90	44.44	31.90	22.00	45.36	34.08	21.72	45.22	33.26	22.00

Table 4.23. Power Rate with $H_a : (\mu, \tau_1, \tau_2, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$=(5,0,1.5,0,0.2,0.4,0.6) in Treatment Test under 5 Covariance Structures of Four Scenarios for Different Size Ratio $n_1/n_2 = (0.5, 1, 2)$, and Variance Ratio $\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$=(2,4,6,8,10), where $\sigma^2_{\gamma 2}$=1

$\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$	Distribution Scenarios											
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew			$\mathbb{Z}w \sim$ Normal $e \sim$ Normal			$\mathbb{Z}w \sim$ Skew $e \sim$ Normal			$\mathbb{Z}w \sim$Skew $e \sim$skew		
	n_1/n_2			n_1/n_2			n_1/n_2			n_1/n_2		
	0.5	1	2	0.5	1	2	0.5	1	2	0.5	1	2
First-Order Autoregressive (AR(1))												
2	72.90	75.38	69.84	74.18	76.26	68.08	75.56	75.94	69.30	74.16	76.08	69.32
4	62.84	60.72	49.48	60.32	60.34	49.48	63.10	62.78	54.44	62.60	63.04	51.70
6	54.48	49.40	36.60	53.08	48.90	37.76	54.62	54.14	42.68	54.38	53.50	43.20
8	48.80	42.72	28.32	42.34	42.34	27.76	47.32	48.84	37.74	47.08	49.22	37.34
10	44.02	37.12	22.14	44.80	37.44	22.30	41.24	44.80	32.58	42.30	44.30	32.44
Toeplitz (TOEP)												
2	72.12	74.48	69.08	73.78	75.86	67.46	75.04	75.46	68.62	73.30	75.42	68.50
4	62.16	59.94	48.60	59.86	59.82	49.00	62.60	62.28	53.86	61.84	62.50	51.02
6	53.72	48.82	36.07	52.60	48.42	37.22	54.08	53.66	42.32	53.50	52.96	42.42
8	48.14	42.24	27.87	48.04	41.90	27.26	47.04	48.70	37.30	46.36	48.80	36.90
10	43.64	36.89	21.74	44.80	37.18	21.90	40.58	44.54	32.34	41.62	43.92	32.06
Compound Symmetry (CS)												
2	72.50	74.96	69.12	73.70	75.80	67.56	74.96	75.38	68.74	73.64	75.52	68.56
4	62.28	60.50	48.90	59.84	59.88	49.04	62.50	62.00	53.92	61.80	62.44	51.26
6	53.82	49.06	36.06	52.46	48.78	37.16	53.84	53.72	42.42	53.88	52.86	42.50
8	48.52	42.48	27.86	47.96	41.76	27.02	47.06	48.66	37.36	46.58	49.08	37.12
10	43.78	36.82	21.78	44.42	37.16	21.88	40.72	44.50	32.32	41.82	44.18	32.12
Unstructured (UN)												
2	72.50	74.96	69.12	73.70	75.80	67.56	74.96	75.38	68.74	73.64	75.52	68.56
4	62.28	60.50	48.90	59.84	59.88	49.04	62.50	62.00	53.92	61.80	62.44	51.26
6	53.82	49.06	36.06	52.46	48.78	37.16	53.84	53.72	42.42	53.88	52.86	42.50
8	48.52	42.48	27.86	47.96	41.76	27.02	47.06	48.66	37.36	46.58	49.08	37.12
10	43.78	36.82	21.78	44.42	37.16	21.88	40.72	44.50	32.32	41.82	44.18	32.12
Variance Components (VC)												
2	93.18	95.46	94.00	93.64	95.30	93.52	94.60	94.28	93.48	94.30	94.22	92.82
4	87.92	88.94	87.74	86.32	89.68	87.58	89.28	87.68	85.64	89.32	88.46	84.60
6	82.24	83.16	79.86	81.78	83.32	80.26	84.34	80.80	78.26	84.64	81.68	78.32
8	78.06	79.28	74.08	76.72	78.50	73.84	79.96	78.54	72.70	80.24	76.64	73.62
10	74.80	74.38	68.82	74.82	75.06	68.68	77.78	73.60	69.94	76.80	73.92	69.10

Table 4.24. Methods for Type I Error Rates in Treatment Test under 5 Covariance Structures of Four Scenarios with Fixed Size Ratio n_1/n_2=0.5 and $\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$=(2,4,6,8,10), where $\sigma^2_{\gamma 2}$=1

$\sigma^2_{\gamma 1}$: $\sigma^2_{\gamma 2}$	Distribution Scenarios																			
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew					$\mathbb{Z}w \sim$ Normal $e \sim$ Normal					$\mathbb{Z}w \sim$ Skew $e \sim$ Normal					$\mathbb{Z}w \sim$ Skew $e \sim$ Skew				
	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS
Original Test																				
2	7.14	6.72	6.86	6.86	33.64	6.82	6.52	6.50	6.50	32.36	6.36	6.16	6.10	6.10	33.16	6.94	6.60	6.80	6.80	34.06
4	10.56	10.04	10.16	10.16	39.74	10.06	9.86	9.92	9.92	37.92	10.82	10.54	10.56	10.56	40.00	10.18	9.74	9.84	9.84	39.60
6	11.10	10.74	10.88	10.88	10.30	11.06	10.78	11.06	11.06	39.42	12.60	12.42	12.16	12.16	41.72	12.12	11.62	11.72	11.72	42.72
8	12.70	12.38	12.38	12.38	42.66	12.34	12.14	12.04	12.04	41.84	13.36	13.00	12.80	12.80	44.88	13.40	13.00	13.12	13.12	44.20
10	14.38	13.80	13.92	13.92	45.62	14.08	13.76	13.64	13.64	44.44	15.02	14.60	14.56	14.56	45.36	14.66	14.34	14.36	14.36	45.22
Rank Test																				
2	6.52	6.32	6.36	6.36	32.54	6.16	5.82	5.84	5.84	30.64	6.46	5.96	6.10	6.10	29.94	7.62	7.42	7.38	7.38	33.30
4	9.32	9.10	8.94	8.94	37.10	8.32	8.12	8.02	8.02	34.20	10.66	10.26	10.34	10.34	37.40	11.94	11.44	11.36	11.36	39.26
6	9.38	8.96	9.02	9.02	36.38	9.16	8.76	8.64	8.64	35.72	13.50	13.00	13.14	13.14	42.20	15.48	15.06	15.02	15.02	44.68
8	10.12	9.88	9.80	9.80	37.92	9.90	9.48	9.44	9.44	37.50	16.08	15.42	15.54	15.54	45.50	16.74	16.60	16.50	16.50	46.16
10	11.66	11.26	11.28	11.28	41.06	10.30	10.08	10.00	10.00	38.48	18.24	17.74	17.66	17.66	47.96	19.32	18.74	18.66	18.66	49.24
Aligned Rank Test																				
2	6.62	6.28	6.42	6.42	32.40	6.18	5.84	5.76	5.76	30.74	6.36	6.14	6.00	6.00	29.82	7.66	7.36	7.36	7.36	33.26
4	9.32	9.08	9.00	9.00	36.90	8.34	8.06	8.00	8.00	34.24	10.64	10.14	10.28	10.28	37.60	11.70	11.28	11.24	11.24	39.40
6	9.46	9.08	9.06	9.06	36.40	9.10	8.80	8.60	8.60	35.76	13.46	13.12	13.14	13.14	42.12	15.40	14.96	14.98	14.98	44.64
8	10.02	9.80	9.86	9.86	37.70	9.88	9.46	9.52	9.52	37.56	16.04	15.44	15.36	15.36	45.56	16.76	16.34	16.34	16.34	46.10
10	11.72	11.24	11.26	11.26	41.02	10.34	10.04	10.02	10.02	38.52	17.74	17.48	17.52	17.52	48.10	19.10	18.64	18.58	18.58	49.30
Sub-Sampling Method																				
2	5.64	5.20	5.12	5.12	30.94	5.20	4.82	4.86	4.86	29.10	4.90	4.72	4.62	4.62	29.62	5.36	5.02	5.04	5.04	30.74
4	6.36	5.90	5.90	5.90	33.46	5.80	5.74	5.66	5.66	31.20	6.54	6.32	6.24	6.24	32.54	6.18	5.82	5.82	5.82	33.16
6	5.28	5.10	5.18	5.18	31.28	5.56	5.32	5.24	5.24	31.34	7.06	6.82	6.64	6.64	32.70	6.92	6.62	6.72	6.72	33.14
8	5.60	5.32	5.26	5.26	32.64	6.02	5.90	5.82	5.82	31.40	7.12	6.92	6.82	6.82	34.24	7.42	6.96	7.12	7.12	32.94
10	5.68	5.56	5.62	5.62	34.48	6.10	5.92	5.86	5.86	32.54	7.96	7.78	7.76	7.76	33.46	7.94	7.56	7.54	7.54	33.40
MIXED model using the GROUP= option																				
2	5.62	5.24	5.40	5.40	30.36	5.10	4.92	4.76	4.76	29.50	5.52	5.52	5.34	5.34	30.40	6.44	5.97	6.32	6.32	31.68
4	5.52	5.29	5.38	5.38	32.88	5.68	5.40	5.44	5.44	30.92	7.32	7.24	7.18	7.18	33.60	7.36	6.84	7.00	7.00	32.86
6	5.22	4.92	5.92	4.92	30.88	5.19	5.02	4.92	4.92	30.52	7.95	7.84	7.56	7.56	33.76	7.98	7.53	7.58	7.58	33.26
8	5.14	4.96	5.08	5.08	32.74	5.28	4.98	4.86	4.86	32.02	8.09	7.98	7.86	7.86	34.26	8.12	7.83	7.82	7.82	33.30
10	5.23	4.97	5.02	5.02	34.44	5.36	5.24	5.12	5.12	32.08	8.59	8.43	8.40	8.40	33.72	8.53	8.21	8.28	8.28	33.86

Table 4.25. Methods for Power Rates with $H_a : (\tau_1, \tau_2)$=(0,1.5) in Treatment Test under 5 Covariance Structures with Fixed Size Ratio n_1/n_2=0.5 and $\sigma^2_{\gamma 1}/\sigma^2_{\gamma 2}$=(2,4,6,8,10), where $\sigma^2_{\gamma 2}$=1

$\sigma^2_{\gamma 1}$: $\sigma^2_{\gamma 2}$	Distribution Scenarios																			
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew					$\mathbb{Z}w \sim$ Normal $e \sim$ Normal					$\mathbb{Z}w \sim$ Skew $e \sim$ Normal					$\mathbb{Z}w \sim$ Skew $e \sim$ Skew				
	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS
Original Test																				
2	72.90	72.12	72.50	72.50	93.18	74.18	73.78	73.70	73.70	93.64	75.56	75.04	74.96	74.96	94.60	74.16	73.30	73.64	73.64	94.30
4	62.84	62.16	62.28	62.28	87.92	60.32	59.86	59.84	59.84	86.32	63.10	62.60	62.50	62.50	89.28	62.60	61.84	61.80	61.80	89.32
6	54.48	53.72	53.82	53.82	82.24	53.08	52.60	52.46	52.46	81.78	54.62	54.08	53.84	53.84	84.34	54.38	53.50	53.88	53.88	84.64
8	48.80	48.14	48.52	48.52	78.06	42.34	48.04	47.96	47.96	76.72	47.32	47.04	47.06	47.06	79.96	47.08	46.36	46.58	46.58	80.24
10	44.02	43.64	43.78	43.78	74.80	44.80	44.80	44.42	44.42	74.82	41.24	40.58	40.72	40.72	77.78	42.30	41.62	41.82	41.82	76.80
Rank Test																				
2	72.32	71.40	71.58	71.58	93.18	71.68	70.64	70.58	70.58	92.14	77.32	76.34	76.00	76.00	95.36	81.74	81.12	81.00	81.00	97.22
4	57.78	56.94	56.88	56.88	85.50	54.36	53.54	53.44	53.44	83.02	58.78	57.92	57.56	57.56	86.56	62.18	61.56	61.24	61.24	89.88
6	48.20	47.66	47.60	47.60	77.80	45.68	44.76	44.68	44.68	77.14	44.04	43.34	43.04	43.04	76.42	45.36	44.70	44.50	44.50	77.86
8	41.42	40.80	40.50	40.50	73.28	40.58	40.10	39.84	39.84	70.94	34.08	33.46	33.36	33.36	66.40	33.96	33.32	33.12	33.12	66.18
10	37.06	36.54	36.38	36.38	68.60	38.02	37.26	37.10	37.10	68.60	25.64	24.88	24.76	24.76	59.08	25.58	24.92	24.76	24.76	57.18
Aligned Rank Test																				
2	72.14	71.42	71.32	71.32	93.10	70.76	70.02	70.20	70.20	92.16	76.76	75.98	75.98	75.98	95.36	81.94	81.24	81.28	81.28	97.28
4	57.82	57.12	56.86	56.86	85.46	54.12	53.46	53.28	53.28	82.86	58.30	57.10	57.22	57.22	86.52	62.22	61.26	61.20	61.20	89.82
6	47.58	47.18	47.24	47.24	77.56	45.40	44.74	44.64	44.64	76.98	43.76	42.96	42.76	42.76	76.14	45.08	44.20	44.42	44.42	78.10
8	40.98	40.38	40.32	40.32	72.96	40.18	39.50	39.58	39.58	70.74	33.58	32.66	32.48	32.48	66.10	33.28	32.46	32.48	32.48	65.54
10	36.88	36.22	36.22	36.22	68.24	37.70	37.16	37.04	37.04	68.30	25.26	24.62	24.44	24.44	58.46	24.90	24.34	24.08	24.08	56.80
Sub-Sampling Method																				
2	57.72	56.18	56.88	56.88	88.76	58.52	57.56	57.60	57.60	88.42	59.06	58.26	58.18	58.18	89.54	59.92	58.90	58.86	58.86	89.02
4	44.74	43.90	44.08	44.08	81.08	41.70	40.99	41.08	41.08	79.18	44.24	43.4	43.44	43.44	82.82	43.54	42.44	42.72	42.72	82.80
6	36.48	35.66	35.78	35.78	74.08	34.28	33.80	33.68	33.68	72.82	32.68	31.64	31.80	31.80	76.98	32.68	31.97	32.26	32.26	77.60
8	30.54	29.78	29.74	29.74	68.74	29.46	29.06	29.08	29.08	67.68	25.44	24.96	25.02	25.02	71.44	25.38	24.48	24.68	24.68	71.82
10	26.16	25.78	36.06	26.06	64.36	25.88	25.48	25.66	25.66	64.92	19.42	19.14	19.04	19.04	67.34	19.84	19.28	19.44	19.44	67.16
MIXED model using the GROUP= option																				
2	65.38	63.91	64.76	64.76	92.40	66.61	65.64	65.76	65.76	92.70	69.22	68.07	67.80	67.80	94.90	69.46	67.89	68.02	68.02	95.16
4	47.72	46.48	46.62	46.62	84.46	45.38	44.72	44.58	44.58	82.70	47.82	46.98	46.56	46.56	87.72	45.23	43.42	44.10	44.10	88.76
6	36.87	35.94	36.18	36.18	76.66	34.77	34.32	34.26	34.26	76.10	32.33	31.61	31.26	31.26	81.24	31.34	29.86	30.26	30.26	81.26
8	29.33	28.55	28.88	28.88	71.00	29.22	28.99	28.78	28.78	69.04	23.16	22.78	22.62	22.62	74.64	22.89	21.91	22.12	22.12	74.74
10	24.70	24.11	24.14	24.14	66.14	25.24	34.93	24.92	24.92	66.18	17.21	16.78	16.80	16.80	70.20	17.19	16.38	16.64	16.64	69.60

Table 4.26. Methods for Type I Error Rates in Treatment Test under 5 Covariance Structures with $\sigma^2_{\gamma_A}$=2, $\sigma^2_{\gamma_B}$=(2,4,6,8,10), and σ^2_C=1

$\sigma^2_{\gamma 1}$: $\sigma^2_{\gamma 2}$	Distribution Scenarios																			
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew					$\mathbb{Z}w \sim$ Normal $e \sim$ Normal					$\mathbb{Z}w \sim$ Skew $e \sim$ Normal					$\mathbb{Z}w \sim$ Skew $e \sim$ Skew				
	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS
Original Test																				
2	6.68	6.30	6.26	6.26	46.48	7.18	6.82	6.90	6.90	46.14	6.80	6.58	6.52	6.52	47.08	6.68	6.42	6.62	6.62	47.82
4	8.62	8.20	8.28	8.28	48.70	8.64	8.22	8.20	8.20	49.32	8.64	8.28	8.30	8.30	50.64	8.22	7.90	7.88	7.88	48.86
6	9.68	9.24	9.38	9.38	50.00	10.08	9.78	9.58	9.58	51.08	9.60	9.32	9.48	9.48	51.38	10.24	9.94	9.82	9.82	52.22
8	10.08	9.88	10.00	10.00	51.70	10.68	10.26	10.30	10.30	50.28	11.30	10.86	10.94	10.94	50.46	11.32	10.94	11.02	11.02	53.24
10	11.64	11.38	11.40	11.40	50.96	10.94	10.76	10.74	10.74	51.62	12.36	11.98	11.94	11.94	51.74	12.84	12.42	12.40	12.40	53.08
Sub-Sampling Method																				
2	4.96	4.52	4.48	4.48	43.12	5.50	5.24	5.36	5.36	42.28	5.18	4.86	4.84	4.84	43.06	5.04	4.66	4.68	4.68	44.40
4	5.20	4.88	4.98	4.98	42.58	5.84	5.58	5.48	5.48	43.58	5.88	5.62	5.60	5.60	45.78	5.50	5.26	5.38	5.38	43.48
6	5.40	5.06	5.24	5.24	42.66	6.50	6.14	6.16	6.16	44.30	6.26	6.14	6.02	6.02	44.32	7.04	6.68	6.72	6.72	44.16
8	6.16	5.98	6.04	6.04	43.42	6.48	6.28	6.24	6.24	43.04	7.06	6.80	6.68	6.68	42.76	7.24	7.06	7.14	7.14	44.84
10	6.38	6.28	6.40	6.40	42.44	6.20	6.04	6.04	6.04	41.92	7.76	7.48	7.50	7.50	43.60	8.42	8.12	8.12	8.12	44.78
MIXED model using the GROUP= option																				
2	4.90	4.53	4.76	4.76	43.92	5.76	5.06	5.18	5.18	43.66	6.47	6.02	5.96	5.96	44.76	6.72	6.09	6.28	6.28	45.56
4	5.50	5.15	5.12	5.12	44.44	5.80	5.49	5.22	5.22	44.40	7.18	6.77	6.74	6.74	47.44	7.46	6.98	7.14	7.14	45.68
6	4.82	4.59	4.64	4.64	45.38	5.86	5.49	5.38	5.38	45.92	7.17	6.87	6.54	6.54	47.00	7.88	7.38	7.52	7.52	47.28
8	5.20	4.66	4.76	4.76	45.88	5.69	5.67	5.52	5.52	45.36	8.01	7.70	7.54	7.54	46.74	8.50	7.82	8.00	8.00	48.24
10	4.69	4.23	4.40	4.40	45.94	5.19	4.81	5.00	4.98	45.48	8.33	7.97	7.74	7.74	46.42	8.63	8.35	8.46	8.46	48.30

Table 4.27. Methods for Power Rates with $H_a : (\tau_1, \tau_2, \tau_3)$=(0,0,1.5) for Three Treatment Groups in Treatment Test under 5 Covariance Structures with $\sigma^2_{\gamma_A}$=2, $\sigma^2_{\gamma_B}$=(2,4,6,8,10), and σ^2_C=1

$\sigma^2_{\gamma 1}$: $\sigma^2_{\gamma 2}$	Distribution Scenarios																			
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew					$\mathbb{Z}w \sim$ Normal $e \sim$ Normal					$\mathbb{Z}w \sim$ Skew $e \sim$ Normal					$\mathbb{Z}w \sim$ Skew $e \sim$ Skew				
	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS
Original Test																				
2	65.72	64.72	64.98	64.98	94.74	65.84	65.08	64.94	64.94	94.58	66.14	65.44	65.30	65.30	95.64	68.68	67.86	68.10	68.10	95.98
4	55.60	54.60	54.88	54.88	89.32	55.90	55.24	54.94	54.94	89.94	56.82	56.06	56.12	56.12	92.34	57.04	56.16	56.48	56.48	92.32
6	49.64	48.96	49.10	49.10	85.06	48.72	47.98	47.86	47.86	84.50	49.60	48.82	48.72	48.72	88.06	49.54	48.66	49.16	49.16	87.92
8	43.14	42.42	42.72	42.72	81.58	45.34	44.80	44.72	44.72	81.60	43.16	42.52	42.48	42.48	84.32	41.72	40.80	41.08	41.08	84.28
10	39.34	38.90	39.30	39.30	77.02	40.08	39.72	39.54	39.54	77.70	38.26	37.76	37.92	37.92	81.14	38.36	37.74	37.94	37.94	79.86
Sub-Sampling Method																				
2	55.68	54.58	54.90	54.90	92.70	55.08	54.30	54.42	54.42	92.26	56.68	55.72	55.98	55.98	92.72	59.30	58.06	58.30	58.30	93.60
4	44.90	43.82	44.04	44.04	85.68	44.92	44.30	44.26	44.26	86.34	44.64	43.86	43.72	43.72	89.10	45.72	44.54	45.20	45.20	89.62
6	38.58	37.64	37.42	37.42	80.52	36.52	36.20	36.46	36.46	79.78	36.58	35.68	35.80	35.80	84.72	36.44	35.24	35.54	35.54	84.26
8	32.34	31.80	31.82	31.82	76.32	33.60	32.96	32.82	32.82	76.42	29.54	28.98	28.92	28.92	79.46	28.56	27.68	27.90	27.90	79.54
10	29.30	28.62	28.66	28.66	71.04	29.12	28.92	29.16	29.16	71.50	23.94	23.54	23.70	23.70	75.84	24.94	24.48	24.40	24.40	74.84
MIXED model using the GROUP= option																				
2	56.09	54.09	54.98	54.98	94.24	55.30	53.82	54.00	54.00	93.78	60.83	59.96	58.94	58.94	96.22	63.90	61.19	61.94	61.94	97.10
4	38.24	36.96	37.26	37.26	86.38	38.58	37.18	37.30	37.30	86.28	37.23	35.89	35.68	35.68	91.40	38.86	37.13	37.26	37.26	91.86
6	29.21	28.20	28.14	28.14	80.68	28.51	27.79	27.20	27.20	79.36	25.46	24.35	24.38	24.38	85.50	25.25	23.87	24.58	24.58	84.78
8	23.21	22.10	22.26	22.26	75.88	24.14	23.27	22.98	22.98	75.76	18.77	18.03	17.98	17.98	79.66	17.39	16.29	16.74	16.74	78.92
10	19.49	18.73	19.04	19.04	71.34	20.03	19.50	19.30	19.30	71.58	13.28	12.66	12.40	12.40	74.68	14.22	13.64	13.84	13.84	74.32

Table 4.28. Methods for Type I Error Rates in Treatment Test under 5 Covariance Structures with $\sigma^2_{\gamma_A}$=6, $\sigma^2_{\gamma_B}$=(2,4,6,8,10), and σ^2_C=1

$\sigma^2_{\gamma 1}$: $\sigma^2_{\gamma 2}$	Distribution Scenarios																			
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew					$\mathbb{Z}w \sim$ Normal $e \sim$ Normal					$\mathbb{Z}w \sim$ Skew $e \sim$ Normal					$\mathbb{Z}w \sim$ Skew $e \sim$ Skew				
	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS
Original Test																				
2	10.00	9.56	9.68	9.68	50.00	9.22	8.68	8.72	8.72	50.66	9.74	9.50	9.54	9.54	50.62	9.76	9.20	9.46	9.46	50.50
4	10.10	9.62	9.72	9.72	52.32	9.92	9.74	9.64	9.64	52.86	9.74	9.40	9.54	9.54	54.18	9.20	8.84	8.90	8.90	52.96
6	10.12	9.68	9.88	9.88	53.04	10.90	10.70	10.54	10.54	53.36	10.44	10.22	10.06	10.06	55.30	10.38	10.18	10.24	10.24	56.40
8	12.10	11.82	11.70	11.70	54.16	10.72	10.42	10.46	10.46	53.86	11.48	11.22	11.32	11.32	55.98	10.84	10.42	10.52	10.52	55.30
10	11.70	11.38	11.64	11.64	55.94	11.10	10.80	10.78	10.78	53.48	11.52	11.14	11.14	11.14	56.40	12.52	12.08	12.14	12.14	56.02
Sub-Sampling Method																				
2	6.42	6.10	6.02	6.02	42.58	5.50	5.32	5.40	5.40	43.44	6.74	6.34	6.30	6.30	43.72	7.16	6.74	6.74	6.74	43.12
4	5.68	5.50	5.50	5.50	43.06	5.76	5.54	5.40	5.40	44.34	5.74	5.46	5.42	5.42	45.96	5.08	4.94	4.92	4.92	44.36
6	5.84	5.70	5.72	5.72	43.94	5.96	5.86	5.68	5.68	44.40	5.70	5.56	5.44	5.44	46.00	5.38	5.20	5.18	5.18	46.78
8	6.34	5.96	6.02	6.02	43.86	5.62	5.48	5.64	5.64	43.70	6.58	6.38	6.38	6.38	46.10	5.66	5.32	5.28	5.28	44.92
10	6.16	6.08	6.00	6.00	45.10	6.34	6.24	6.20	6.20	42.82	6.28	6.12	6.08	6.08	46.24	7.00	6.64	6.88	6.88	46.96
MIXED model using the GROUP= option																				
2	5.44	5.11	5.14	5.14	44.94	5.36	4.86	4.64	4.64	45.26	7.36	7.05	6.76	6.76	46.60	8.30	7.45	7.76	7.76	46.34
4	5.18	5.11	5.10	5.10	45.36	5.76	5.47	5.44	5.44	46.36	7.74	7.36	7.38	7.38	48.80	8.27	7.57	7.74	7.74	46.72
6	5.11	4.79	4.88	4.88	46.02	5.92	5.50	5.56	5.56	46.16	9.18	8.61	8.80	8.80	48.38	8.77	8.20	8.26	8.26	48.98
8	5.08	4.73	4.88	4.88	46.90	5.53	5.35	5.28	5.28	45.82	9.47	9.17	9.18	9.18	49.30	9.26	8.61	8.72	8.72	47.82
10	5.52	5.15	5.38	5.38	47.14	5.65	5.37	5.36	5.36	46.18	9.34	8.83	8.82	8.82	49.30	10.19	9.68	9.98	9.98	49.70

Table 4.29. Methods for Power Rates with $H_a : (\tau_1, \tau_2, \tau_3)$=(0,0,1.5) for Three Treatment Groups in Treatment Test under 5 Covariance Structures with $\sigma^2_{\gamma_A}$=6, $\sigma^2_{\gamma_B}$=(2,4,6,8,10), and σ^2_C=1

$\sigma^2_{\gamma 1}$: $\sigma^2_{\gamma 2}$	Distribution Scenarios																			
	$\mathbb{Z}w \sim$ Normal $e \sim$ Skew					$\mathbb{Z}w \sim$ Normal $e \sim$ Normal					$\mathbb{Z}w \sim$ Skew $e \sim$ Normal					$\mathbb{Z}w \sim$ Skew $e \sim$ Skew				
	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS	AR(1)	TOEP	CS	UN	VS
Original Test																				
2	54.38	53.40	53.68	53.68	93.70	54.36	53.82	53.84	53.84	92.72	57.30	56.56	56.58	56.58	94.14	56.52	55.50	55.82	55.82	93.90
4	48.02	47.22	47.42	47.42	88.68	48.62	48.20	48.14	48.14	88.98	48.40	47.74	47.86	47.86	89.50	50.40	49.34	49.42	49.42	90.20
6	43.56	43.12	43.28	43.28	84.74	43.16	42.52	42.72	42.72	84.68	44.28	43.62	43.80	43.80	86.84	43.30	42.64	42.68	42.68	87.20
8	39.94	39.14	39.18	39.18	80.68	40.10	39.56	39.78	39.78	80.66	39.18	38.76	38.66	38.66	83.44	38.56	37.74	37.90	37.90	83.60
10	36.22	35.70	36.18	36.18	79.30	36.52	36.22	36.04	36.04	78.26	34.40	34.00	33.94	33.94	82.16	36.22	35.76	35.98	35.98	81.60
Sub-Sampling Method																				
2	38.34	37.32	37.58	37.58	88.04	39.18	38.42	38.58	38.58	87.84	44.34	43.90	43.88	43.88	88.24	44.32	43.26	43.40	43.40	88.26
4	33.26	32.34	32.70	32.70	82.50	34.08	33.16	33.18	33.18	82.62	36.56	35.92	35.64	35.64	82.62	37.04	36.26	36.34	36.34	84.68
6	30.20	29.58	29.70	29.70	77.74	28.88	28.42	28.66	28.66	78.50	31.14	30.60	30.44	30.44	79.76	29.68	29.24	29.46	29.46	80.38
8	26.74	25.78	26.26	26.26	72.94	27.82	27.32	27.42	27.42	73.52	26.28	25.78	25.68	25.68	75.98	25.44	24.72	24.94	24.94	76.42
10	24.28	23.76	23.92	23.92	71.30	23.90	23.50	23.44	23.44	69.68	21.24	20.88	20.88	20.88	74.42	23.34	22.90	22.92	22.92	74.14
MIXED model using the GROUP= option																				
2	53.13	51.20	52.00	52.00	94.32	54.42	53.24	53.40	53.40	93.22	59.18	57.75	57.56	57.56	96.28	60.53	57.95	58.76	58.76	96.78
4	35.65	34.06	34.40	34.40	86.40	35.95	34.94	34.64	34.64	86.68	38.14	36.92	36.94	36.91	89.76	38.15	36.19	36.24	36.24	91.06
6	27.95	26.72	26.92	26.92	80.76	28.23	27.52	27.42	27.42	80.34	26.24	25.34	25.52	25.52	84.92	25.47	24.31	24.14	24.14	86.04
8	22.30	21.47	21.34	21.34	74.66	23.24	22.75	22.40	22.40	74.62	19.07	18.36	18.24	18.24	79.24	18.12	17.07	17.34	17.34	79.76
10	18.33	17.90	17.70	17.70	72.14	18.90	18.52	18.02	18.02	71.28	14.71	14.13	13.88	13.88	76.42	15.73	14.97	14.76	14.76	75.84

4.4. Real Data Application

To illustrate the PROC MIXED model and the used methods, a real dataset was analyzed. This dataset is from field studies on cocoa trees in central Africa by Owusu Domfeh [Domfeh et al., 2019]. This study aims to determine the number of rows of cocoa that should be inoculated with mild strain N1 for effective cross-protection of core cocoa trees from CSSV severe 1A infection. The data was collected over 8 years (from 2011 to 2018), and we have a dependent variable – the yield data y. Moreover, there are three biological treatment groups: severely infected cocoa trees (CSSV 1A), mildly infected cocoa trees (CSSV N1), uninoculated cocoa trees (Non-inoculated). Therefore, it is a longitudinal study with eight repeated measures as yield data y and heterogeneous variances across three treatments. The visualization of the trend of yield data over time is presented in Figure 4.4. Meanwhile, according to Figure 4.5 and Table 4.30, we can see that there is clear variance differences among treatment groups: the standard deviation of CSSV 1A treatment group is 200.967, the standard deviation of CSSV N1 treatment group is 798.342, and the standard deviation of Non-inoculated treatment group is 339.901.

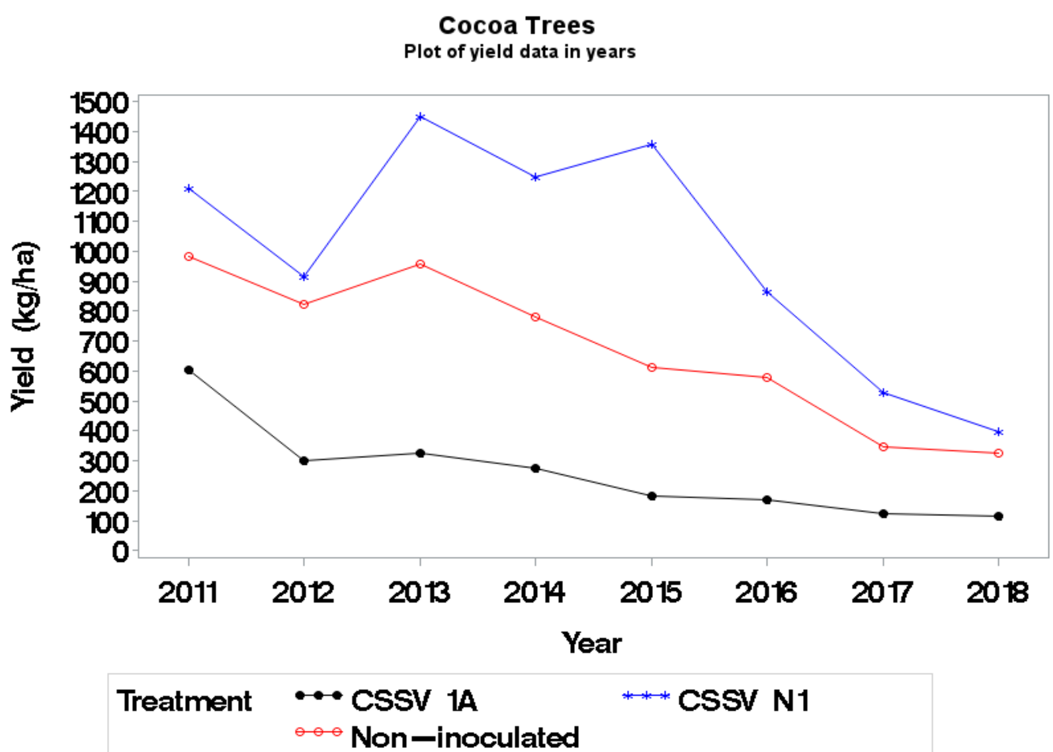

Figure 4.4. Plot of the Yield Data of Cocoa Trees in Years

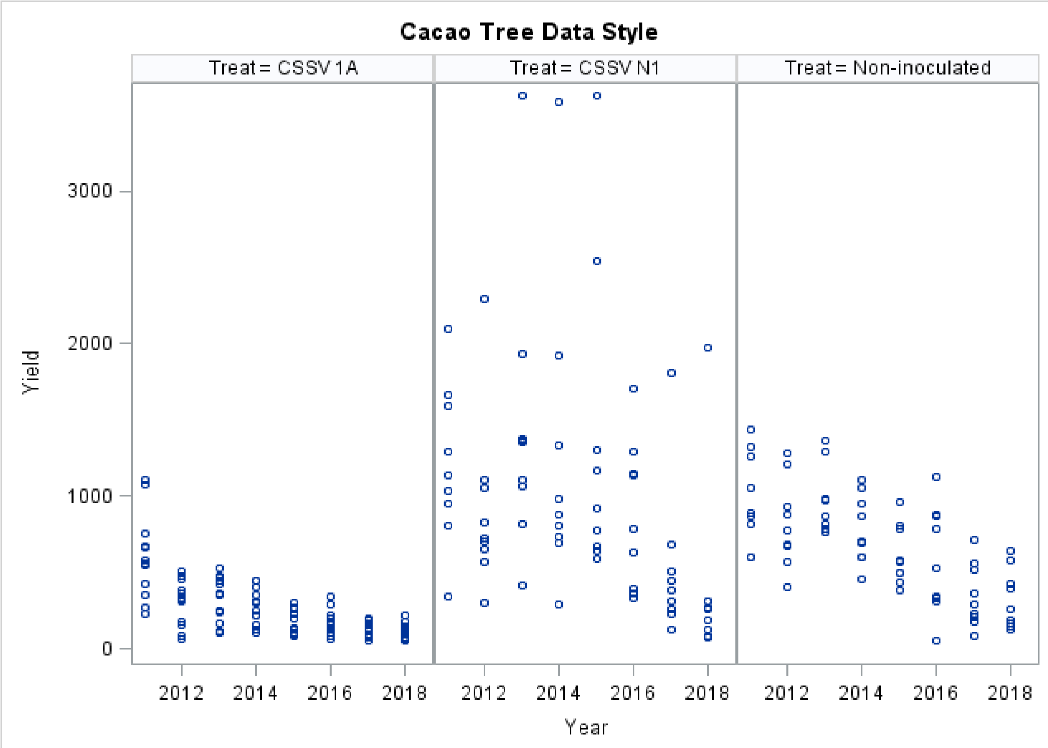

Figure 4.5. Cocao Tree Data Style

Table 4.30. The SAS MEANS Procedure

Analysis Variable: Yield						
Treat Group	N Obs	N	Mean	Std Dev	Minimum	Maximum
CSSV 1A	96	96	262.228	200.967	54.700	1106.000
CSSV N1	72	72	996.062	789.342	70.400	3625.800
Non-inoculated	72	72	675.922	339.901	48.800	1432.000

At first, the dataset was run using the PROC MIXED model with Toeplitz (TOEP) as the covariance structure, and the code is shown in Figure 4.6. Secondly, GROUP= option under REPEATED statement was used, and the code is shown in Figure 4.7. Lastly, the Sub-Sampling method was used, and the code of Sub-Sampling method is shown in Figure 4.8. For this dataset, there are 12 subjects in the treatment group of CSSV 1A per year and 9 subjects in the treatment group of CSSV N1 and Non-inoculated per year. The sub-sampling method here is to randomly select 9 subjects without replacement from the treatment group of CSSV 1A and made the size of three treatment groups equal. Then this updated dataset was run using the PROC MIXED model with TOEP as the covariance structure. The results of Type 3 Tests of Fixed Effects from the MIXED model for these three methods is shown in Table 4.31.

```
proc mixed data=yield;
   class Treat Year RM_Subj;
   model yield = Treat|Year / ddfm=KR;
   repeated Year / subject=RM_Subj Type=TOEP R RCorr;
   title2 'Mixed Model with TOEP covariance structure';
   run;
```

Figure 4.6. PROC MIXED Model with TOEP as the Covariance Structure

```
proc mixed data=yield;
   class Treat Year RM_Subj;
   model yield = Treat|Year / ddfm=KR;
   repeated Year / subject=RM_Subj Type=TOEP group=treat;
   title2 'Mixed Model with GROUP= option';
   run;
```

Figure 4.7. PROC MIXED Model with GROUP= Option

```
data A;
   set yield;
     if treat="CSSV 1A";
     run;

data N;
   set yield;
   if treat="CSSV N1";
   RUN;

data non;
   set yield;
     if treat="Non-inoculated";
     run;

proc surveyselect data=A out=A1
   method=srs  sampsize=9 ;
   strata treat year;
   run;

data yield_equal_group;
   set N non A1;
   run;

proc sort data=yield_equal_group;
   by treat year;
   run;

proc mixed data=yield_equal_group;
   class Treat Year RM_Subj;
   model yield = Treat|Year / ddfm=KR;
   repeated Year / subject=RM_Subj Type=TOEP R RCorr;
   title2 'Sub-Sampling Method';
   run;
```

Figure 4.8. Sub-Sampling Method

Table 4.31. The Results of Type 3 Tests of Fixed Effects from MIXED Model

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr> F
Original MIXED Model				
Treat	2	26.7	9.89	0.006
Year	7	58.6	24.09	<.0001
Treat*Year	14	78.6	6.93	<.0001
MIXED model using the GROUP= option				
Treat	2	19.2	17.36	<.0001
Year	7	35.2	14.50	<.0001
Treat*Year	14	40.5	6.13	<.0001
Sub-Sampling Method				
Treat	2	27.1	11.36	0.0003
Year	7	61.6	17.55	<.0001
Treat*Year	14	77.1	5.36	<.0001

5. CONCLUSION

To sum up, this study proposed four features to simulate longitudinal data, then using the MIXED model to do analysis. The four conditions/features are: 1. Unbalanced sample size; 2. Unequal group variance; 3. Violating normality assumption of the MIXED model; 4. A MIXED model with incorrect covariance structures. This research aims to check how Type I error rates would change within different conditions, and the corresponding power rates would also be provided as references.

The first and main problem in this study is the analogue Behrens-Fisher problem under the MIXED model structure. There are two features in this problem: unbalanced group size and unequal group variance. Only one feature itself should not be a concern when using a MIXED model, but it would be likely to have Type I error rate inflation problem when having the two features simultaneously. The two features were interpreted as the two factors related to the Type I error rate inflation problem in this study: ratios of sample size and ratios of group variance. When the size and variance ratios are fixed, the inflated Type I error rate is consistent no matter the changes of actual sample sizes or the number of repeated measures. According to our Second-Order Response Surface Model in Chapter 4.1.2, the factor of size ratio may have more impact on the prediction of Type I error rate inflation comparing to the factor of variance ratio, so when there is an unbalanced group sample size, there may be a potential Type I error rate inflation/deflation problem when the group variances are different. Furthermore, when a group has a relatively small sample size but relatively large variance, we should be cautious of the Type I error inflation problem. The MIXED model method using the GROUP= option Method and Sub-Sampling Method can be reasonable solutions when having this problem. From a practical point, the method of the MIXED model using the GROUP= option is recommended. However, based on the real data application, we found that the Group= option does not always work on real data. When having these cases, the Sub-Sampling procedure would be a great option. In the meantime, based on the corresponding power rates' performance, when there is Type I error rate inflation, the corresponding power rates would also have detective deficiency. By standardizing the effect size, we could know how much the power rate is impacted by the difference of group variances.

The second problem is regards to violating the normality assumption of the MIXED model. According to Chapter 4, the MIXED model is reasonably robust to modest violations of the normal distribution. However, when data is heavily skewed with a big difference of group variances, the MIXED model's performance would be broken down.

The third problem is how does the incorrect covariance structures affect Type I error rates and power rates. Comparing to the correct covariance structure, First-Order Autoregressive (AR(1)) [Littell et al., 2006, Kincaid, 2005] which choosing the incorrect covariance structure among Toeplitz (TOEP), Compound Symmetry (CS) and Unstructured (UN) [Littell et al., 2006, Kincaid, 2005] does not affect the results of Type I error rate and power rate, the difference among the four is around one percent. Nevertheless, Variance Components (VC) structure [Kincaid, 2005, Littell et al., 2006] would increase the Type I error rate 3 to even 17 times compared to the results of First-Order Autoregressive (AR(1)), as well as creating the highest power rate comparing to other four. As we know, Variance Components (VC) is the simplest covariance structure. It specifies that observations are independent even on the same subjects, which is not realistic for most longitudinal data. So neglecting the correlated measurements in a longitudinal study might be why Variance Components (VC) structure causes the excessive Type I error rate inflation. Therefore, when having longitudinal data, if not sure which covariance structure should be used in the MIXED model, any one of the four following covariance structures would be recommended: First-Order Autoregressive (AR(1)), Toeplitz (TOEP), Compound Symmetry (CS), and Unstructured (UN).

Future research could look into the impact of issues of unbalanced samples and heterogeneous variances in more complex designs. All of the simulations in this thesis assume data from continuous distributions. The mixed model framework generalizes to discrete distributions as well. Perhaps some of these issues such as which variance-covariance structure to use or the impact of unequal variances and unequal sample sizes could be investigated using simulated data from discrete distributions. Thought would need to be given to the assumptions of these generalized mixed models and whether or not the sample size and variance issues that can plague continuous distributions in a mixed model ANOVA would impact the generalized models as well.

REFERENCES

[Amir et al., 2016] Amir, W. M., Shafiq, M., Mokhtar, K., Aleng, N. A., Rahim, H. A., and Ali, Z. (2016). Simple response surface methodology using rsreg (sas). *Journal of Modern Applied Statistical Methods*, 15(1):855–867.

[Box and Wilson, 1951] Box, G. E. and Wilson, K. B. (1951). On the experimental attainment of optimum conditions. *Journal of the royal statistical society: Series b (Methodological)*, 13(1):1–38.

[Bradley, 1978] Bradley, J. V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology*, 31(2):144–152.

[Bradley, 2007] Bradley, N. (2007). *The response surface methodology*. PhD thesis, Indiana University South Bend.

[Chernick, 2011] Chernick, M. R. (2011). *Bootstrap methods: A guide for practitioners and researchers*, volume 619. John Wiley & Sons.

[Domfeh et al., 2019] Domfeh, O., Ameyaw, G., Dzahini-Obiatey, H., and del Río Mendoza, L. (2019). Spatiotemporal spread of cacao swollen shoot virus severe strain 1a in mixed hybrid cacao pre-inoculated with mild strain n1. *Plant Disease*, 103(12):3244–3250.

[Efron, 1981] Efron, B. (1981). Nonparametric estimates of standard error: the jackknife, the bootstrap and other methods. *Biometrika*, 68(3):589–599.

[Fisher, 1938] Fisher, R. A. (1938). The statistical utilization of multiple measurements. *Annals of eugenics*, 8(4):376–386.

[Fleishman, 1978] Fleishman, A. I. (1978). A method for simulating non-normal distributions. *Psychometrika*, 43(4):521–532.

[Guerin and Stroup, 2000] Guerin, L. and Stroup, W. W. (2000). A simulation study to evaluate proc mixed analysis of repeated measures data.

[Harville and Jeske, 1992] Harville, D. A. and Jeske, D. R. (1992). Mean squared error of estimation or prediction under a general linear model. *Journal of the American Statistical Association*, 87(419):724–731.

[Hesterberg et al., 2005] Hesterberg, T., Moore, D., Monaghan, S., Clipson, A., Epstein, R., Moore, D., and McCabe, G. (2005). Bootstrap methods and permutation tests ,(2005).

[Higgins et al., 1990] Higgins, J. J., Blair, R. C., and Tashtoush, S. (1990). The aligned rank transform procedure.

[Hill and Dixon, 1982] Hill, M. and Dixon, W. (1982). Robustness in real life: A study of clinical laboratory data. *Biometrics*, pages 377–396.

[Inc, 2016] Inc, S. (2016). Sas/stat. 14.2 user's guide. *Cary, NC: SAS Institute Inc.*

[Jennrich and Schluchter, 1986] Jennrich, R. I. and Schluchter, M. D. (1986). Unbalanced repeated-measures models with structured covariance matrices. *Biometrics*, pages 805–820.

[Johnson and Montgomery, 2009] Johnson, R. T. and Montgomery, D. C. (2009). Choice of second-order response surface designs for logistic and poisson regression models. *International Journal of Experimental Design and Process Optimisation*, 1(1):2–23.

[Kenward and Roger, 1997] Kenward, M. G. and Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, pages 983–997.

[Keselman et al., 1999a] Keselman, H., Algina, J., Kowalchuk, R. K., and Wolfinger, R. D. (1999a). The analysis of repeated measurements: A comparison of mixed-model satterthwaite f tests and a nonpooled adjusted degrees of freedom multivariate test. *Communications in Statistics-Theory and Methods*, 28(12):2967–2999.

[Keselman et al., 1999b] Keselman, H., Algina, J., Kowalchuk, R. K., and Wolfinger, R. D. (1999b). A comparison of recent approaches to the analysis of repeated measurements. *British Journal of Mathematical and Statistical Psychology*, 52(1):63–78.

[Kim and Cohen, 1998] Kim, S.-H. and Cohen, A. S. (1998). On the behrens-fisher problem: A review. *Journal of Educational and Behavioral Statistics*, 23(4):356–377.

[Kincaid, 2005] Kincaid, C. (2005). Guidelines for selecting the covariance structure in mixed model analysis. In *Proceedings of the thirtieth annual SAS users group international conference*, volume 30, pages 198–130. SAS Institute Inc Cary NC.

[King, 2017] King, T. J. (2017). Power analysis to determine the importance of covariance structure choice in mixed model repeated measures anova. Master's thesis, North Dakota State University.

[Kowalchuk et al., 2004] Kowalchuk, R. K., Keselman, H., Algina, J., and Wolfinger, R. D. (2004). The analysis of repeated measurements with mixed-model adjusted f tests. *Educational and Psychological Measurement*, 64(2):224–242.

[Kwok et al., 2007] Kwok, O.-m., West, S. G., and Green, S. B. (2007). The impact of misspecifying the within-subject covariance structure in multiwave longitudinal multilevel models: A monte carlo study. *Multivariate Behavioral Research*, 42(3):557–592.

[Lehmann and D'Abrera, 1975] Lehmann, E. L. and D'Abrera, H. J. (1975). *Nonparametrics: statistical methods based on ranks.* Holden-Day.

[Littell et al., 2006] Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., and Oliver, S. (2006). *SAS for mixed models.* SAS publishing.

[Myers and Montgomery, 1997] Myers, R. H. and Montgomery, D. C. (1997). A tutorial on generalized linear models. *Journal of Quality Technology*, 29(3):274–291.

[Paul et al., 2019] Paul, S., Wang, Y.-G., and Ullah, I. (2019). A review of the behrens-fisher problem and some of its analogs: Does the same size fit all? *Revstat Statistical Journal*, 17(4):563–597.

[Politis et al., 1999] Politis, D. N., Romano, J. P., and Wolf, M. (1999). *Subsampling.* Springer Science & Business Media.

[Prasad and Rao, 1990] Prasad, N. N. and Rao, J. N. (1990). The estimation of the mean squared error of small-area estimators. *Journal of the American statistical association*, 85(409):163–171.

[Reed III, 2003] Reed III, J. F. (2003). Solutions to the behrens–fisher problem. *Computer methods and programs in biomedicine*, 70(3):259–263.

[SAS, 2015] SAS (2015). Sas/stat® 14.1. user's guide.

[Scheffe, 1959] Scheffe, H. (1959). *The analysis of variance*, volume 72. John Wiley & Sons.

[Schroeder and Martin, 2005] Schroeder, W. J. and Martin, K. M. (2005). Overview of visualization.

[Vale and Maurelli, 1983] Vale, C. D. and Maurelli, V. A. (1983). Simulating multivariate nonnormal distributions. *Psychometrika*, 48(3):465–471.

[Wicklin, 2013] Wicklin, R. (2013). *Simulating data with SAS*. SAS Institute.

APPENDIX A. SIMULATION CODE

The code below was used to perform the simulation study. Various parts of the code were changed to obtain all the simulations based on different conditions. Note that some of the code is commented off to prevent large amounts of output from printing.

```
proc iml;
  load module=_all_;       /* load Fleishman and Vale-Maurelli modules */

  %let seed0=0;
  %let NumSamples=5000;
  t=2;                  *** t defines # of treatment groups  ***;
  k=4;                  *** k defines # of repeated measures ***;
  s=30;                 *** s defines # of subjects          ***;
  rho=0.75;              *** Autocorrelation for AR(1)        ***;
  sigma2_R=1;           *** Residual variance component      ***;
  sigma2_S1=1;           *** Subject variance component-control       ***;
  sigma2_S2=2;           *** Subject variance component-trt    ***;

  NumSamples = &NumSamples;

************ VCV basic structure is AR(1) here ************;
  e=j(k,k,0);           *** Matrix of zeroes.                ***;
  do i=1 to k;
    do j=1 to k;
       if i=j then e[i,j]=1;  *** Convert diagonal to 1s  ***;
         else e[i,j]=rho**abs(i-j); *** AR(1) rhos        ***;
    end;
  end;
  print e;

  do i = 2 to 30;        *** Create block-diagonal matrix   ***;
    R = block(R, e);    ***     with s blocks             ***;
  end;
  R = I(30) @ e;         *** block-diagonal matrix -trt   ***;
  print R;
```

```
************  Between subjects variance structure ************;
   B=j(k,k,1);


  B1 = sigma2_S1*B;       *** Multiply VCV by subject sigma2 ***;
  print B1;
  do i = 2 to 20;        *** Create block-diagonal matrix   ***;
    R1 = block(R1, B1);   ***     with s blocks             ***;
  end;
  R1 = I(20) @ B1;         *** block-diagonal matrix -control   ***;

  B2 = sigma2_S2*B;       *** Multiply VCV by subject sigma2 ***;
  print B2;
  do i = 2 to 10;        *** Create block-diagonal matrix   ***;
    R2 = block(R2, B2);   ***     with s blocks             ***;
  end;
  R2 = I(10) @ B2;         *** block-diagonal matrix -trt  ***;
  print R2;


*************************************************************
 ****************** Fixed Effects *************************
*************************************************************

************  Set up the fixed portion design   ************;
  Trt=j(80,1,1);        *** Set up trt design columns  80*1***;
  do i=2 to t;
  Trt=colvec(Trt//j(40,1,i)); *concatenate vertical //*; *120*1 (contain 1*2)*;
  end;
  Trt = design(Trt);
  print Trt;
  xxx=I(k);
  print xxx;
  *** Set up repeated measures columns ***;
  do i=1 to s;
```

```
  Week=(Week//xxx);
 end;
 print Week;

 X = j(nrow(Week),1,1)||Trt||Week;
 print X;

 ******* Fixed effects parms: mu t1 t2 w1 w2 w3 w4    ********;
 beta = {5, 0, 0, 0, 0, 0, 0};

 ******* Fixed effects eta **********************************;
 eta1 = X*beta;          *** Fixed effects portion of obs. ***;
 eta = repeat(eta1,1,NumSamples);
 print beta;
 *print eta1;
 *print eta;


*********************************************
 ****************** Random Effects ***********
**********************************************

************ Subject Random Error :  zu ~ MVN (0,R) ************;
 call randseed(&seed0);
 zero1 = j(1, 80, 0);             *** the zero vector-control ***;
 zero2 = j(1, 40, 0);             *** the zero vector-trt  ***;
 print zero1 zero2;
 esubj1 = RandNormal(NumSamples, zero1, R1); * esubj ~ MVN(0,R);
 esubject1=t(esubj1);
 esubj2 = RandNormal(NumSamples, zero2, R2); * esubj ~ MVN(0,R);
 esubject2=t(esubj2);
 esubject=(esubject1//esubject2);
 *print esubject;


 ******* Random Error Component -- eps[i,j]~N(0,sigma2_R)Normal *******;
```

```
 call randseed(&seed0);
  zero = j(1, 120, 0);                    *** the zero vector-trt  ***;
  eps1 = RandNormal(NumSamples, zero, R); * esubj ~ MVN(0,R);
  eps=t(eps1);


*****************************************************************
 Conduct the equation with the fixed and random effect we got above ;
*****************************************************************;
  Y = eta + esubject + eps;
  *print Y;



  ******************************************************
  Create MV data= X(design matrix)+ Indiv (subject_ID) + Y
  ******************************************************;
  Indiv = colvec(repeat(T(1:s),1,k));  *** Create Subj_ID ***;
  *print Indiv;

  MVdata=X[,2:(t+k+1)]||Indiv||Y;
  *print mvdata;

  create rm.Subj_MV from MVdata;
  append from MVdata;
  close;

  quit;

proc print data=rm.Subj_MV(obs=10);
  title2 'MV_Data_Structure';
  run;



***************************************************
  *********** Transfer MV data to UN data ***********;
```

```
***************************************************;

  data rm.Subj_UV(keep=Sample Subj_ID Stress Trt Period);
  set rm.Subj_MV;
  Subj_ID=Col7;
  if col1=1 then trt='A';
     else if col1=0 then trt='B';
  if col3=1 then period=1;
     else if col4=1 then period=2;
     else if col5=1 then period=3;
     else if col6=1 then period=4;
  array ys{&NumSamples} col8-col5007; *** ZZZ Adjust with sample size. ***;
  do Sample=1 to &NumSamples;
     Stress=ys{Sample};
     output;
  end;
  run;

proc sort data=rm.Subj_UV;
  by Sample;
  run;
proc print data=rm.Subj_UV (obs=10);
  title2 'UV Data Structure';
  run;

******** ANALYSIS *****************;
%macro mvn(dsn,p,n);

proc transpose data=rm.Subj_UV out=rm.mv_&dsn._&p._&n prefix=Stress;
  by Sample Subj_ID Trt;
  var Stress;
  run;
*proc print;
  run;

*********************************************************************
```

```
*** Repeated Measures MV Analysis Block                              ***
**********************************************************************;
ods listing close;
ods output MultStat=MultStat_&dsn._&p
           ModelANOVA=ModelANOVAmv_&dsn._&p
           Epsilons=Epsilons_&dsn._&p;
proc glm data=rm.mv_&dsn._&p._&n;
  by sample;
  class trt;
  model stress1 - stress&p = trt / nouni;
  repeated period &p;
  title2 'Repeated Measures ANOVA for Effect of Time on Stress Level';
  run;
ods listing;

proc contents noprint;
  run;

data rm.uv_&dsn._&p._&n;
  set rm.Subj_UV;
  run;

**********************************************************************
*** Repeated Measures UV Analysis Block -- Traditional Approach      ***
**********************************************************************;
ods listing close;
ods output ModelANOVA=ModelANOVAuv_&dsn._&p
           FitStatistics=FitStatisticsuv_&dsn._&p;
proc glm data=rm.uv_&dsn._&p._&n;
  by sample;
  class trt period;
  model stress = trt | period;  *** allows subject error combined w/ mse.;
  title2 'Univariate ANOVA of Time on Stress Level (No Subject Effect)';
  run;
ods listing;
```

```
proc contents noprint;
  run;

*********************************************************************
*** Repeated Measures UV Analysis Block -- Mixed Model Approach    ***
*********************************************************************;
%macro mmuv(VCV);

*ods trace on / listing;
ods listing close;
ods output
    CovParms=Cov&VCV._&dsn._&p
    FitStatistics=Fit&VCV._&dsn._&p
    Tests3=Tests3&VCV._&dsn._&p;
proc mixed data=rm.uv_&dsn._&p._&n;
  by sample;
  class trt period subj_id;
  model stress = trt | period /ddfm=kr;
  repeated period / subject=subj_id type=&VCV;
  title2 "Repeated Measures ANOVA using Mixed Model Approach -- &VCV";
  run;
ods listing;

proc contents noprint;
  run;

%mend mmuv;

%mmuv(VC);
%mmuv(CS);
%mmuv(TOEP);
%mmuv(UN);

*ZZZzzzz;
ods listing close;
ods output
```

```
     CovParms=CovAR_&dsn._&p
     FitStatistics=FitAR_&dsn._&p
     Tests3=Tests3AR_&dsn._&p;
proc mixed data=rm.uv_&dsn._&p._&n;
  by sample;
  class trt period subj_id;
  model stress = trt | period /ddfm=kr;
  repeated period / subject=subj_id type=AR(1);
  title2 "Repeated Measures ANOVA using Mixed Model Approach -- AR(1)";
  run;
ods listing;

proc contents noprint;
  run;


title1 "Simulation of Skewed Data, N=&N -- Data Covariance Structure was &dsn";
proc sort data=MultStat_&dsn._&p
(where=(Hypothesis='period' and Statistic="Wilks'_Lambda"))
__________out=Per_mvrm(keep=Sample_value_fvalue_probf);
__by_sample;
__run;
proc_sort_data=MultStat_&dsn._&p
(where=(Hypothesis='period_trt'_and_Statistic="Wilks' Lambda"))
          out=PerTrt_mvrm(keep=Sample value fvalue probf);
  by sample;
  run;
/*
proc sort data=MultStat_&dsn._&p
          out=mvrm;
  by sample;
  run;
*/

proc sort data=ModelANOVAmv_&dsn._&p
(where=(Source='period'))
```

```
        out=Per_mvgghf(keep=Sample FValue ProbF ProbFGG ProbFHF);
  by sample;
  run;
proc sort data=ModelANOVAmv_&dsn._&p
(where=(Source='period*trt'))
        out=PerTrt_mvgghf(keep=Sample FValue ProbF ProbFGG ProbFHF);
  by sample;
  run;
proc sort data=ModelANOVAmv_&dsn._&p
(where=(Source='trt'))
        out=Trt_mvgghf(keep=Sample FValue ProbF ProbFGG ProbFHF);
  by sample;
  run;
/* ANOVA with GG and HF above */

proc sort data=ModelANOVAuv_&dsn._&p
(where=(HypothesisType=3 and source='period'))
        out=Per_uvrm(keep=Sample FValue ProbF SS MS);
  by sample;
  run;
proc sort data=ModelANOVAuv_&dsn._&p
(where=(HypothesisType=3 and source='trt*period'))
        out=PerTrt_uvrm(keep=Sample FValue ProbF SS MS);
  by sample;
  run;
proc sort data=ModelANOVAuv_&dsn._&p
(where=(HypothesisType=3 and source='trt'))
        out=Trt_uvrm(keep=Sample FValue ProbF SS MS);
  by sample;
  run;
/* ANOVA UV above */

proc sort data=Tests3VC_&dsn._&p
(where=(Effect='period'))
        out=Per_uvrmVC(keep=Sample FValue ProbF);
  by sample;
```

```
  run;
proc sort data=Tests3VC_&dsn._&p
(where=(Effect='trt*period'))
          out=PerTrt_uvrmVC(keep=Sample FValue ProbF);
  by sample;
  run;
proc sort data=Tests3VC_&dsn._&p
(where=(Effect='trt'))
          out=Trt_uvrmVC(keep=Sample FValue ProbF);
  by sample;
  run;
/*
proc sort data=Tests3VC_&dsn._&p
          out=uvrmVC;
  by sample;
  run;
*/

proc sort data=Tests3CS_&dsn._&p(where=(Effect='period'))
          out=Per_uvrmCS(keep=Sample FValue ProbF);
  by sample;
  run;
proc sort data=Tests3CS_&dsn._&p(where=(Effect='trt*period'))
          out=PerTrt_uvrmCS(keep=Sample FValue ProbF);
  by sample;
  run;
proc sort data=Tests3CS_&dsn._&p(where=(Effect='trt'))
          out=Trt_uvrmCS(keep=Sample FValue ProbF);
  by sample;
  run;
/* CS Results above */

proc sort data=Tests3TOEP_&dsn._&p(where=(Effect='period'))
          out=Per_uvrmTOEP(keep=Sample FValue ProbF);
  by sample;
  run;
```

```
proc sort data=Tests3TOEP_&dsn._&p(where=(Effect='trt*period'))
          out=PerTrt_uvrmTOEP(keep=Sample FValue ProbF);
  by sample;
  run;
proc sort data=Tests3TOEP_&dsn._&p(where=(Effect='trt'))
          out=Trt_uvrmTOEP(keep=Sample FValue ProbF);
  by sample;
  run;
/* TOEP Results above */

proc sort data=Tests3UN_&dsn._&p(where=(Effect='period'))
          out=Per_uvrmUN(keep=Sample FValue ProbF);
  by sample;
  run;
proc sort data=Tests3UN_&dsn._&p(where=(Effect='trt*period'))
          out=PerTrt_uvrmUN(keep=Sample FValue ProbF);
  by sample;
  run;
proc sort data=Tests3UN_&dsn._&p(where=(Effect='trt'))
          out=Trt_uvrmUN(keep=Sample FValue ProbF);
  by sample;
  run;
/* UN Results above */

proc sort data=Tests3AR_&dsn._&p(where=(Effect='period'))
          out=Per_uvrmAR(keep=Sample FValue ProbF);
  by sample;
  run;
proc sort data=Tests3AR_&dsn._&p(where=(Effect='trt*period'))
          out=PerTrt_uvrmAR(keep=Sample FValue ProbF);
  by sample;
  run;
proc sort data=Tests3AR_&dsn._&p(where=(Effect='trt'))
          out=Trt_uvrmAR(keep=Sample FValue ProbF);
  by sample;
  run;
```

```
/* AR Results above */

/*
proc print data=mvgghf;
  title2 'MVRM_Structure';
  run;
*/

proc format;
  value pow low-.05 = 'Reject'
            .05<-high = 'DNR';
  value mod 1='Wilks'  2='UVCS'  3='GG'  4='HF'  5='UVNoBlk'
            6='MxVC'   7='MxCS'  8='MxTOEP'
            9='MxUN'  10='MxAR';
  value modtrt
            1='UVCS'  2='UVNoBlk'
            3='MxVC'  4='MxCS'  5='MxTOEP'
            6='MxUN'  7='MxAR';
  run;

*** Period tests ***;
data rm.Per_all_ps_&p._&n;
  merge Per_mvrm   (rename=(value=wilks fvalue=wl_f probf=wlf_p))
        Per_mvgghf(rename=(fvalue=per_f probf=per_p))
        Per_uvrm   (rename=(fvalue=uvper_f probf=uvper_p))
        Per_uvrmVC(rename=(fvalue=VCper_f probf=VCper_p))
        Per_uvrmCS(rename=(fvalue=CSper_f probf=CSper_p))
        Per_uvrmTOEP(rename=(fvalue=TOEPper_f probf=TOEPper_p))
        Per_uvrmUN(rename=(fvalue=UNper_f probf=UNper_p))
        Per_uvrmAR(rename=(fvalue=ARper_f probf=ARper_p));
 *by sample;
  run;
data Per_ult(keep=sample i p_value);
  set rm.Per_all_ps_&p._&n;
  array ps{10} wlf_p per_p ProbFGG ProbFHF uvper_p
                VCper_p CSper_p TOEPper_p UNper_p ARper_p;
```

```
  do i=1 to 10;
     p_value=ps{i};
     output;
  end;
  run;
proc freq data=Per_ult;
  tables p_value*i / nopct norow;
  format p_value pow. i mod.;
  title3 'Period_Results';
  run;

*** Period*Trt Tests ***;
data rm.PerTrt_all_ps_&p._&n;
  merge PerTrt_mvrm   (rename=(value=wilks fvalue=wl_f probf=wlf_p))
        PerTrt_mvgghf(rename=(fvalue=per_f probf=per_p))
        PerTrt_uvrm   (rename=(fvalue=uvper_f probf=uvper_p))
        PerTrt_uvrmVC(rename=(fvalue=VCper_f probf=VCper_p))
        PerTrt_uvrmCS(rename=(fvalue=CSper_f probf=CSper_p))
        PerTrt_uvrmTOEP(rename=(fvalue=TOEPper_f probf=TOEPper_p))
        PerTrt_uvrmUN(rename=(fvalue=UNper_f probf=UNper_p))
        PerTrt_uvrmAR(rename=(fvalue=ARper_f probf=ARper_p));
  run;
data PerTrt_ult(keep=sample i p_value);
  set rm.PerTrt_all_ps_&p._&n;
  array ps{10} wlf_p per_p ProbFGG ProbFHF uvper_p
                VCper_p CSper_p TOEPper_p UNper_p ARper_p;
  do i=1 to 10;
     p_value=ps{i};
     output;
  end;
  run;
proc freq data=PerTrt_ult;
  tables p_value*i / nopct norow;
  format p_value pow. i mod.;
  title3 'Period*Trt_Results';
  run;
```

```
*** Trt Tests ***;
data rm.Trt_all_ps_&p._&n;
  merge Trt_mvgghf(rename=(fvalue=per_f probf=per_p))
        Trt_uvrm  (rename=(fvalue=uvper_f probf=uvper_p))
        Trt_uvrmVC(rename=(fvalue=VCper_f probf=VCper_p))
        Trt_uvrmCS(rename=(fvalue=CSper_f probf=CSper_p))
        Trt_uvrmTOEP(rename=(fvalue=TOEPper_f probf=TOEPper_p))
        Trt_uvrmUN(rename=(fvalue=UNper_f probf=UNper_p))
        Trt_uvrmAR(rename=(fvalue=ARper_f probf=ARper_p));
  run;
data Trt_ult(keep=sample i p_value);
  set rm.Trt_all_ps_&p._&n;
  array ps{7} per_p uvper_p VCper_p CSper_p TOEPper_p UNper_p ARper_p;
  do i=1 to 7;
     p_value=ps{i};
     output;
  end;
  run;
proc freq data=Trt_ult;
  tables p_value*i / nopct norow;
  format p_value pow. i modtrt.;
  title3 'Trt_Results';
  run;

%mend mvn;

%macro looper(p);
  %do dim=4 %to &p;
   *** mvn(dsn,p,n) ***;
      %mvn(AR1,4,30);
  %end;
%mend looper;

%looper(4);
*proc print;
```

```
  run;

%let t1 = %sysfunc(datetime());
%let elapsedTime = %sysevalf(&t1-&t0);
%put &elapsedTime;
```

APPENDIX B. ALIGNED RANK TEST CODE

The code below was used to prepare the aligned and ranked data-sets for TREATMENT TEST. Various parts of the code were changed to obtain all the results based on different simulation data-sets. Note that some of the code is commented off to prevent large amounts of output from printing.

```
*******************************************************
  **** Build up calculation for aligned and ranks ****
******************************************************;
data subj_uv;
  set rm.subj_uv;
  run;

********** Get the means for row|column variables ******;
proc means data=subj_uv noprint mean;
  by sample;
  var stress;
  output out=ybar_overall mean=ybar_overall;
  run;

proc means data=subj_uv noprint mean;
  by sample trt;
  var stress;
  output out=ybar_trt mean=ybar_trt;
  run;

proc sort data=subj_uv;
  by sample period;
proc means  noprint mean;
  by sample period;
  var stress;
  output out=ybar_period mean=ybar_period;
  run;

proc sort data=subj_uv;
```

```
  by sample trt period;
proc means  noprint mean;
  by sample trt period;
  var stress;
  output out=ybar_inter mean=ybar_inter;
  run;

***** Align Dataset ****;
Data ybar1 (drop=_freq_ _type_) ;
  merge ybar_period(IN=IN1)
              ybar_overall(IN=IN2) ;
  by sample ;
  if IN1;
  run;

Data ybar2 (drop=_freq_ _type_) ;
  merge ybar_inter(IN=IN1)
              ybar_trt(IN=IN2) ;
  by sample trt ;
  if IN1;
  run;

proc sort data=ybar2;
  by sample period;
Data ybar3 (drop=_freq_ _type_) ;
  merge ybar2(IN=IN1)
              ybar1(IN=IN2) ;
  by sample period ;
  if IN1;
  run;

proc sort data=ybar3;
  by sample trt period;
Data ybar;
  merge subj_uv(IN=IN1)
              ybar3(IN=IN2) ;
```

```
  by sample trt period;
  if IN1;
  run;



***** Aligned datasets ****;
data rm.align (drop=ybar_overall ybar_trt ybar_period ybar_inter);
  set ybar;
  inter_ali=stress-ybar_trt-ybar_period+ybar_overall;
  trt_ali=stress-ybar_inter+ybar_trt-ybar_overall;
  period_ali=stress-ybar_inter+ybar_period-ybar_overall;
  run;

**** Rank Datasets ****;
proc sort data=rm.align;
  by sample trt subj_ID period;
proc rank data=rm.align out=rm.aligned_rank;
  by sample;
  var inter_ali trt_ali period_ali;
  run;

data rm.subj_uv (rename=(trt_ali=stress));
  set rm.aligned_rank (drop=stress period_ali inter_ali);
  format _all_;
  run;

************* ANALYSIS ***************************;

%macro mvn(dsn,p,n);

proc transpose data=rm.Subj_UV out=rm.mv_&dsn._&p._&n prefix=Stress;
  by Sample Subj_ID Trt;
  var Stress;
  run;
*proc print;
```

```
  run;

********************************************************************
*** Repeated Measures MV Analysis Block   ***
********************************************************************;
ods listing close;
ods output MultStat=MultStat_&dsn._&p
           ModelANOVA=ModelANOVAmv_&dsn._&p
           Epsilons=Epsilons_&dsn._&p;
proc glm data=rm.mv_&dsn._&p._&n;
  by sample;
  class trt;
  model stress1 - stress&p = trt / nouni;
  repeated period &p;
  title2 'Repeated Measures ANOVA for Effect of Time on Stress Level';
  run;
ods listing;

proc contents noprint;
  run;

data rm.uv_&dsn._&p._&n;
  set rm.Subj_UV;
  run;

********************************************************************
*** Repeated Measures UV Analysis Block -- Traditional Approach    ***
********************************************************************;
ods listing close;
ods output ModelANOVA=ModelANOVAuv_&dsn._&p
           FitStatistics=FitStatisticsuv_&dsn._&p;
proc glm data=rm.uv_&dsn._&p._&n;
  by sample;
  class trt period;
  model stress = trt | period;  *** allows subject error combined w/ mse.;
  title2 'Univariate ANOVA of Time on Stress Level (No Subject Effect)';
```

```
  run;
ods listing;

proc contents noprint;
  run;

*********************************************************************
*** Repeated Measures UV Analysis Block -- Mixed Model Approach      ***
*********************************************************************;
%macro mmuv(VCV);

*ods trace on / listing;
ods listing close;
ods output
     CovParms=Cov&VCV._&dsn._&p
     FitStatistics=Fit&VCV._&dsn._&p
     Tests3=Tests3&VCV._&dsn._&p;
proc mixed data=rm.uv_&dsn._&p._&n;
  by sample;
  class trt period subj_id;
  model stress = trt | period/ddfm=kr;
  repeated period / subject=subj_id type=&VCV;
  title2 "Repeated Measures ANOVA using Mixed Model Approach -- &VCV";
  run;
ods listing;

proc contents noprint;
  run;

%mend mmuv;

%mmuv(VC);
%mmuv(CS);
%mmuv(TOEP);
%mmuv(UN);
```

```
*ZZZzzzz;
ods listing close;
ods output
    CovParms=CovAR_&dsn._&p
    FitStatistics=FitAR_&dsn._&p
    Tests3=Tests3AR_&dsn._&p;
proc mixed data=rm.uv_&dsn._&p._&n;
  by sample;
  class trt period subj_id;
  model stress = trt | period/ddfm=kr;
  repeated period / subject=subj_id type=AR(1);
  title2 "Repeated Measures ANOVA using Mixed Model Approach -- AR(1)";
  run;
ods listing;

proc contents noprint;
  run;


title1 "Simulation of Skewed Data, N=&N -- Data Covariance Structure was &dsn";
proc sort data=MultStat_&dsn._&p
(where=(Hypothesis='period' and Statistic="Wilks'␣Lambda"))
␣␣␣␣␣␣␣␣␣␣out=Per_mvrm(keep=Sample␣value␣fvalue␣probf);
␣␣by␣sample;
␣␣run;
proc␣sort␣data=MultStat_&dsn._&p
(where=(Hypothesis='period_trt'␣and␣Statistic="Wilks' Lambda"))
          out=PerTrt_mvrm(keep=Sample value fvalue probf);
  by sample;
  run;
/*
proc sort data=MultStat_&dsn._&p
          out=mvrm;
  by sample;
  run;
*/
```

```
proc sort data=ModelANOVAmv_&dsn._&p(where=(Source='period'))
          out=Per_mvgghf(keep=Sample FValue ProbF ProbFGG ProbFHF);
  by sample;
  run;
proc sort data=ModelANOVAmv_&dsn._&p(where=(Source='period*trt'))
          out=PerTrt_mvgghf(keep=Sample FValue ProbF ProbFGG ProbFHF);
  by sample;
  run;
proc sort data=ModelANOVAmv_&dsn._&p(where=(Source='trt'))
          out=Trt_mvgghf(keep=Sample FValue ProbF ProbFGG ProbFHF);
  by sample;
  run;
/* ANOVA with GG and HF above */

proc sort data=ModelANOVAuv_&dsn._&p
(where=(HypothesisType=3 and source='period'))
          out=Per_uvrm(keep=Sample FValue ProbF SS MS);
  by sample;
  run;
proc sort data=ModelANOVAuv_&dsn._&p
(where=(HypothesisType=3 and source='trt*period'))
          out=PerTrt_uvrm(keep=Sample FValue ProbF SS MS);
  by sample;
  run;
proc sort data=ModelANOVAuv_&dsn._&p
(where=(HypothesisType=3 and source='trt'))
          out=Trt_uvrm(keep=Sample FValue ProbF SS MS);
  by sample;
  run;
/* ANOVA UV above */

proc sort data=Tests3VC_&dsn._&p(where=(Effect='period'))
          out=Per_uvrmVC(keep=Sample FValue ProbF);
  by sample;
  run;
```

```
proc sort data=Tests3VC_&dsn._&p(where=(Effect='trt*period'))
          out=PerTrt_uvrmVC(keep=Sample FValue ProbF);
  by sample;
  run;
proc sort data=Tests3VC_&dsn._&p(where=(Effect='trt'))
          out=Trt_uvrmVC(keep=Sample FValue ProbF);
  by sample;
  run;
/*
proc sort data=Tests3VC_&dsn._&p
          out=uvrmVC;
  by sample;
  run;
*/

proc sort data=Tests3CS_&dsn._&p(where=(Effect='period'))
          out=Per_uvrmCS(keep=Sample FValue ProbF);
  by sample;
  run;
proc sort data=Tests3CS_&dsn._&p(where=(Effect='trt*period'))
          out=PerTrt_uvrmCS(keep=Sample FValue ProbF);
  by sample;
  run;
proc sort data=Tests3CS_&dsn._&p(where=(Effect='trt'))
          out=Trt_uvrmCS(keep=Sample FValue ProbF);
  by sample;
  run;
/* CS Results above */

proc sort data=Tests3TOEP_&dsn._&p(where=(Effect='period'))
          out=Per_uvrmTOEP(keep=Sample FValue ProbF);
  by sample;
  run;
proc sort data=Tests3TOEP_&dsn._&p(where=(Effect='trt*period'))
          out=PerTrt_uvrmTOEP(keep=Sample FValue ProbF);
  by sample;
```

```
  run;
proc sort data=Tests3TOEP_&dsn._&p(where=(Effect='trt'))
          out=Trt_uvrmTOEP(keep=Sample FValue ProbF);
  by sample;
  run;
/* TOEP Results above */

proc sort data=Tests3UN_&dsn._&p(where=(Effect='period'))
          out=Per_uvrmUN(keep=Sample FValue ProbF);
  by sample;
  run;
proc sort data=Tests3UN_&dsn._&p(where=(Effect='trt*period'))
          out=PerTrt_uvrmUN(keep=Sample FValue ProbF);
  by sample;
  run;
proc sort data=Tests3UN_&dsn._&p(where=(Effect='trt'))
          out=Trt_uvrmUN(keep=Sample FValue ProbF);
  by sample;
  run;
/* UN Results above */

proc sort data=Tests3AR_&dsn._&p(where=(Effect='period'))
          out=Per_uvrmAR(keep=Sample FValue ProbF);
  by sample;
  run;
proc sort data=Tests3AR_&dsn._&p(where=(Effect='trt*period'))
          out=PerTrt_uvrmAR(keep=Sample FValue ProbF);
  by sample;
  run;
proc sort data=Tests3AR_&dsn._&p(where=(Effect='trt'))
          out=Trt_uvrmAR(keep=Sample FValue ProbF);
  by sample;
  run;
/* AR Results above */

/*
```

```
proc print data=mvgghf;
  title2 'MVRM Structure';
  run;
*/


ods rtf file="ART_10vs20_ZU_(RandFleishman(0,G(trt10,con1))_e_MVN(0,AR(1)).rtf";
title1 " Aligned Rank Test for TREATMENT TEST;
        Unbalanced treatment group (10vs20)";
title2 "Between effect: ZU~RandFleishman(u=0, sigma2=G(trt10,con1),
        skew=2, kurtosis=6)";
title3 "Within effect: e~MVN(0,AR(1))";


proc format;
  value pow low-.05 = 'Reject'
            .05<-high = 'DNR';
  value mod 1='Wilks'  2='UVCS'  3='GG'  4='HF'  5='UVNoBlk'
            6='MxVC'   7='MxCS'  8='MxTOEP'
            9='MxUN'  10='MxAR';
  value modtrt
            1='UVCS'  2='UVNoBlk'
            3='MxVC'  4='MxCS'  5='MxTOEP'
            6='MxUN'  7='MxAR';
  run;

*** Period tests ***;
data rm.Per_all_ps_&p._&n;
  merge Per_mvrm  (rename=(value=wilks fvalue=wl_f probf=wlf_p))
        Per_mvgghf(rename=(fvalue=per_f probf=per_p))
        Per_uvrm  (rename=(fvalue=uvper_f probf=uvper_p))
        Per_uvrmVC(rename=(fvalue=VCper_f probf=VCper_p))
        Per_uvrmCS(rename=(fvalue=CSper_f probf=CSper_p))
        Per_uvrmTOEP(rename=(fvalue=TOEPper_f probf=TOEPper_p))
        Per_uvrmUN(rename=(fvalue=UNper_f probf=UNper_p))
        Per_uvrmAR(rename=(fvalue=ARper_f probf=ARper_p));
```

```
 *by sample;
  run;
data Per_ult(keep=sample i p_value);
  set rm.Per_all_ps_&p._&n;
  array ps{10} wlf_p per_p ProbFGG ProbFHF uvper_p
               VCper_p CSper_p TOEPper_p UNper_p ARper_p;
  do i=1 to 10;
     p_value=ps{i};
     output;
  end;
  run;
proc freq data=Per_ult;
  tables p_value*i / nopct norow;
  format p_value pow. i mod.;
  title4 'Period_Results';
  run;

*** Period*Trt Tests ***;
data rm.PerTrt_all_ps_&p._&n;
  merge PerTrt_mvrm  (rename=(value=wilks fvalue=wl_f probf=wlf_p))
        PerTrt_mvgghf(rename=(fvalue=per_f probf=per_p))
        PerTrt_uvrm  (rename=(fvalue=uvper_f probf=uvper_p))
        PerTrt_uvrmVC(rename=(fvalue=VCper_f probf=VCper_p))
        PerTrt_uvrmCS(rename=(fvalue=CSper_f probf=CSper_p))
        PerTrt_uvrmTOEP(rename=(fvalue=TOEPper_f probf=TOEPper_p))
        PerTrt_uvrmUN(rename=(fvalue=UNper_f probf=UNper_p))
        PerTrt_uvrmAR(rename=(fvalue=ARper_f probf=ARper_p));
  run;
data PerTrt_ult(keep=sample i p_value);
  set rm.PerTrt_all_ps_&p._&n;
  array ps{10} wlf_p per_p ProbFGG ProbFHF uvper_p
               VCper_p CSper_p TOEPper_p UNper_p ARper_p;
  do i=1 to 10;
     p_value=ps{i};
     output;
  end;
```

```
  run;
proc freq data=PerTrt_ult;
  tables p_value*i / nopct norow;
  format p_value pow. i mod.;
  title4 'Period*Trt_Results';
  run;

*** Trt Tests ***;
data rm.Trt_all_ps_&p._&n;
  merge Trt_mvgghf(rename=(fvalue=per_f probf=per_p))
        Trt_uvrm   (rename=(fvalue=uvper_f probf=uvper_p))
        Trt_uvrmVC(rename=(fvalue=VCper_f probf=VCper_p))
        Trt_uvrmCS(rename=(fvalue=CSper_f probf=CSper_p))
        Trt_uvrmTOEP(rename=(fvalue=TOEPper_f probf=TOEPper_p))
        Trt_uvrmUN(rename=(fvalue=UNper_f probf=UNper_p))
        Trt_uvrmAR(rename=(fvalue=ARper_f probf=ARper_p));
  run;
data Trt_ult(keep=sample i p_value);
  set rm.Trt_all_ps_&p._&n;
  array ps{7} per_p uvper_p VCper_p CSper_p TOEPper_p UNper_p ARper_p;
  do i=1 to 7;
     p_value=ps{i};
     output;
  end;
  run;
proc freq data=Trt_ult;
  tables p_value*i / nopct norow;
  format p_value pow. i modtrt.;
  title4 'Trt_Results';
  run;
ods rtf close;
%mend mvn;

%macro looper(p);
  %do dim=4 %to &p;
   *** mvn(dsn,p,n) ***;
```

```
      %mvn(AR1,4,30);
   %end;
%mend looper;

%looper(4);
*proc print;
   run;

%let t1 = %sysfunc(datetime());
%let elapsedTime = %sysevalf(&t1-&t0);
%put &elapsedTime;
```