2017-01-01

Sample Size Estimation for Linear Mixed Models with Dependent End Points

Michael Nsiah-Nimo
University of Texas at El Paso, michael20129648@gmail.com

SAMPLE SIZE ESTIMATION FOR LINEAR MIXED MODELS WITH DEPENDENT END POINTS

MICHAEL NSIAH-NIMO

Master's Program in Mathematical Sciences

APPROVED:

Amy E. Wagler, Ph.D., Chair

Leung Ming-Ying, Ph.D.

Thompson Sarkodie-Gyan, Ph.D.

Charles Ambler, Ph.D.
Dean of the Graduate School

to my

FAMILY

with love

SAMPLE SIZE ESTIMATION FOR LINEAR MIXED MODELS WITH DEPENDENT END POINTS

by

MICHAEL NSIAH-NIMO

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

Master's Program in Mathematical Sciences

THE UNIVERSITY OF TEXAS AT EL PASO

August 2017

Acknowledgements

Praised be the name of the LORD and may His name be lifted up. He that started with me has surely brough it to a successful end. His enduring mercies and loving kindness has brought me this far. A mission and vision worthwhile completing. He did not turn away my prayer, nor my supplications and His grace. I am overwhelmed by what He has done and grateful for His abundant grace, bringing what he begun to a successful end.

I am also very grateful to my supervisor and mentor Dr. Amy Wagler, whose guidance and supervision made this work a success. You are like a mother to me, always encouraging and guiding me through my studies.

I wish to thank the other members of my committee, Dr. Ming-Ying Leung, Director, Computational Science and Dr. Thompson Sarkodie-Gyan, of Department of Electrical Engineering, all at The University of Texas at El Paso. Your coaching and immense assistance was of much significance geared towards the completion of this work.

Again, I wish to thank Dr. Joan Staniswalis, Dr. Ming-Ying Leung, Dr. Panagis Moschopoulos, Dr. Naijun Sha, Dr. Xiaogang Su, Dr. Ori Rosen, and from the Mathematical Sciences Department at The University of Texas at El Paso who taught and mentored me as well through my studies. Additionally, I wish to thank all professors and staff of the Mathematical Sciences Department here in The University of Texas at El Paso for all their support, enabling me to complete my degree.

I appreciate my wonderful family as well, both here in the United States and Ghana. Your prayers, love, support , guidance and advice has made this journey worthwhile travelled. To all my other well-wishers whose names are not mentioned, I say "God richly bless you".

Abstract

The primary objective is sample size estimation in linear mixed model settings. Sample size estimation is an important component of planning a well thought out scientific experiment. Whenever sample size estimation is performed, taking into account a priori model based inferences will provide a sample size estimate that will achieve the desired power without inflating the type I error rate of the study.

One common practice is a traditional approach cited in the literature that uses the largest sample size after you Bonferroni the type I error rate to estimate sample sizes as such. We are going to take into account multiplicity using the tree spanning(graph-based) algorithm and improve on just a Bonferroni correction so that we can estimate somewhat lower sample sizes for linear mixed model settings.

We present a tree spanning algorithm that is a fast simple novel approach for estimating sample size which focuses on controlling the family-wise error in LMM's with arbitrary dependency structures. This method warrants more powerful bounds compared to the Bonferroni which tends to be more conservative for large set of comparisons.

This proposed methodology will yield smaller estimators for sample size may be obtained to make better use of time and resources in experimental settings.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Multiplicity Adjustments And Its Significance

Multiplicity Adjustments play a very significant role in any study with multiple outcomes and measures. In real life settings, when we encounter experiments with several hypotheses evaluated simultaneously either through statistical testing or interval estimation, multiplicity of inferences needs to be corrected. Overall statistical inferences should not be made when these errors are left unattended.

The primary aim of multiple comparisons is to reduce the number of false positives (i.e. falsely rejected true null hypotheses). When you perform a large number of statistical tests, some will have p-values being significant purely by chance, even if all the null hypotheses are indeed true. Hence, anytime we reject a null hypotheses because a p-value is less than an assumed significance level, it's possible that we're wrong in our decision.

Each of the individual tests or confidence intervals has a type I error rate that can be controlled by the experimenter. Considering the tests together as a family, then a combined type I error rate for the family of tests or intervals can be calculated. Multiple comparisons and simultaneous confidence intervals could be used interchangeably, specifically applicable to data with discrete end-points and data with continuous end-points respectively.

Lets consider the folllowng scenario: Say you have a set of hypotheses that you wish to test simultaneously. The first idea that might come to mind is to test each hypothesis separately, using some level of significance α. At first, this doesn't seem like a bad idea. However, consider a case where you have 20 hypotheses to test, and a significance level of 0.05. What's the probability of observing at least one significant result just due to chance?

When you perform a large number of statistical tests, some will have p-values being significant purely by chance, even if all the null hypotheses are really true. Hence, anytime we reject multiple null hypotheses because a p-value is less than a critical value, it's highly possible that we're wrong in our decision and making false positive conclusions. Besides our vital objective with multiple comparisons is to reduce the number of false positives (i.e. falsely rejected true null hypotheses).

$$\begin{aligned} \Pr(\text{at least one significant result}) &= 1 - \Pr(\text{no significant results}) \\ &= 1 - (1 - .05)^{20} \\ &\approx 0.64. \end{aligned}$$

So, with 20 tests being considered, we have a 64% chance of observing at least one significant result, even if all of the tests are actually not significant. The probability of getting a significant result simply due to chance keeps going up, as the number of tests increases.

- Note that each of these test was controlled at an $\varepsilon_i = 0.05$ level. This is the **individual error rate** or **comparison wise error rate** (CWER).
- The calculated overall (combined) rate is called the **experiment-wise error rate** (EER) or **family-wise error rate** (FWER)=0.64.

1.2 Types of Errors

In statistics, a null hypothesis is a statement that one seeks to nullify with evidence to the contrary. Mostly, an experimenter frames a null hypothesis with the intent of rejecting it. The alternative hypothesis is that which is opposed to the null hypothesis. In statistical test theory, the notion of statistical error is an integral part of hypothesis testing. The following tables provide the possible outcomes for both single and multiple hypotheses tests scenario.

Let k=$\{H_{01}$,H_{02},...,$H_{0k}\}$ be a set of null hypotheses to be tested, their combined or overall null hypotheses H_0 can also be defined as;

$$\boldsymbol{H_0 = \{H_{01} \cap H_{02} \cap H_{03} \cap \cdots \cap H_{0k}\}}$$

- H_0 is true if and only if all $H_{0i}s$ are true and false if any of the $H_{0i}s$ is rejected.
- ε_i and ε are the Type 1 error rates for the $\boldsymbol{i=1,...,k}$ test and combined tests respectively.

A Single Hypothesis

Consider the table of outcomes below for a single hypothesis test

Table 1.1: Possible outcomes for a single hypothesis

Decision	**Reality**	
	True Null	**False Null**
Reject Null	False Positive (Type 1 Error)	True Positive
Fail to Reject Null	True Negative	False Negative (Type 2 Error)

In general, we control the probability of making a type I error (ε), and among those procedures that control ε choose one that makes less type II error (β).

Multiple Hypotheses

In any multiple testing problem perform k simultaneous hypothesis tests with a common procedure. For any given procedure, classify the results as follows;

Table 1.2: Possible Decisions (Outcomes) for m hypotheses

	H_0 Rejected	H_0 Retained	Total
H_0 True	$V(\epsilon)$	$S(1-\epsilon)$	k_0
H_0 False	$U(1-\beta)$	$T(\beta)$	k_1
Total	R	R^c	k

For the "k" set family of hypotheses, we have $k = V + S + U + T$. In practice, these counts are unknown but can be worked with theoretically.

k_0 is the number of true nulls out of the set of hypotheses (unknown). k_1 is the number of false nulls out of the set of hypotheses (unknown). R is the number of rejected nulls out of the set of hypotheses (known). R^c is the number of non-rejected nulls out of the set of hypotheses (known). $\frac{V}{k_0}$ is the false positive fraction. $\frac{U}{k_1}$ is the sensitivity fraction. $\frac{U}{R}$ is the true discovery fraction. $\frac{V}{R}$ is the false discovery fraction. $\frac{U+S}{k}$ is the accuracy fraction. Wang and Chen (2004) proposed the use of sensitivity measure u/k_1 for gene selection. These error rates are defined depending on the numbers or fractions of falsely rejected null hypotheses, which will never be known in practice.

1.3 Error Rates

Usually the concept of error rates is appropriately applied to hypothesis testing. It represents the probability of Type I error or equivalently the level of significance.

Per comparison error rate (PCER) is the probability of rejecting a particular H_{0i} in a single test when H_{0i} is true. This is just the error rate usual error rate for a t-test or F-test;

it ignores multiplicity corrections. When the PCER is controlled at ε, it means that the expected fraction of individual tests that reject H_{0i} when H_0 is true is ε i.e.

$$P_r \text{ (Reject each } H_{0i} | H_0 \text{ true)} \leq \varepsilon$$

In scenario above the tests controlled the PCER at 5%.

The Familywise error rate (FWER) is the probability of rejecting a least one of the H_{0is} (hence rejecting H_0) when all H_{0is} are true. When the FWER is controlled at ε, it means that the expected proportion of experiments in which we would reject at least on of the H_{0is} when H_0 is true is ε. i.e.

$$P_r \text{ (Rejectat least one } H_{0i} | H_0 \text{ true)} \leq \varepsilon$$

For example in scenario, the FWER is the fraction of times we would have, observing at least one significant result, even if all of the tests are actually not significant. FWER controls PCER at no more than ε

False discovery rate (FDR); In reality, the assumption that H_0 can reasonably hold, is not always true, since some of the H_{0is} could be actually false. In conducting multiple comparisons the false discovery rate (FDR) is one way of conceptualizing the rate of type I errors in null hypothesis testing. Controlling FDR means making sure that the the expected proportion of "discoveries" (rejected null hypotheses) is at most ε. Assuming all H_{0is} are true (i.e. H_0 true), then all discoveries (rejections) are false and FDR is just the FWER. So FDR also controls the FWER at ε. Now in reference to table

- Let $\dfrac{V}{R}$ be the proportion of rejected $H_{0i}s$ that are actually true. This is also known as the false discovery fraction, where; $\frac{V}{R} = 0$ when $R = 0$.

- Controlling FDR means making sure that; $E\left(\frac{V}{R}\right) \leq \varepsilon$ i.e. the expected fraction of false rejections(discoveries) is at most ε.

- Assuming all $H_{0i}s$ are true (i.e. H_0 true), then all discoveries (rejections) are false and FDR is just the FWER. So FDR also controls the FWER at ε.

Lastly, the more correct rejections you make, the more false rejections FDR lets you make but this ratio is limited.

Strong familywise error rate is the probability of making any false discoveries, that is, the probability that the false discovery fraction is greater than zero. Controlling the strong familywise error rate at ε means that the probability of making any false rejections is ε or less, irrespective of how many correct rejections are made. This allows some of the H_{0is} to be false under the null.i.e.

$$P_r \text{ (Reject any } H_{0i} | some H_{0is} \text{ true)} \leq \varepsilon$$

Thus one true rejection cannot make any false rejections more likely. Controlling the strong familywise error rate at ε controls the FDR at no more than ε.

1.4 Simultaneous confidence intervlas

Let's assume that each of the H_{0is} relates to some parameter (for example, a mean), and we place confidence intervals on all these parameters. In the case where one of our confidence intervals fails to cover the true parameter value, an error occurs. If this true parameter value is also the null hypothesis value, then an error is a false rejection. The simultaneous confidence intervals criterion states that all of our confidence intervals must cover their true parameters simultaneously with confidence $1 - \varepsilon$. Simultaneous $1 - \varepsilon$ confidence intervals also control the strong familywise error rate at no more than ε.

1.5 Some Multiple Testing Methods

Multiple comaprison methods are statistical procedures designed to take into account and control the inflation of the overall probability of Type I error or the deflation overall confidence coefficient. The selection of a type of multiple comparison method depends on the experimenter's knowledge of the type of error rate to control. The ones discussed in this section are Bonferroni -based.

1.5.1 Ordinary Bonferroni

This Bonferroni-based procedure is the simplest and most widely used multiple comparison method.It works for a fixed set of k parameters to estimate or null hypotheses to test. Let p_i be the p-value for testing each null Hoi. All the technique does is to reject H_{oi} and thus Ho if $p_i \leq \frac{\varepsilon}{k}$, i.e each of the test is computed an $\frac{\varepsilon}{k}$. Again this procedure seeks to obtain simultaneous confidence intervals evaluated at $1-\varepsilon$ by constructing individual confidence intervals with Bonferroni coverage $1-\varepsilon/k$. This controls the familywise error rates and produces simultaneous confidence intervals. The tests need not have a relation in any way. The Bonferroni method assumes independence among test endpoints. This a flaw especially in LMMs since there would always be dependecies that exists among test endpoints.

- Lets consider two sets of null hypotheses;

$$P_r(H_{01}) = P_r(H_{02}) = \text{probability of Type 1 Error} = \boldsymbol{\varepsilon}$$
$$P_r(H_{01}^c \cap H_{02}^c) = 1 - P_r(H_{01} \cup H_{02}) = 1 - P_r(H_{01}) - P_r(H_{02}) + P_r(H_{01} \cap H_{02})$$
$$P_r(H_{01}^c \cap H_{02}^c) \geq 1 - P_r(H_{01}) - P_r(H_{02}) = 1 - 2\varepsilon$$
$$P_r(\text{at least one Type 1 Error}) \leq 1 - (1 - 2\varepsilon) = 2\varepsilon$$

  Hence by making $P_r(H_{01}) = P_r(H_{02}) = \dfrac{\boldsymbol{\varepsilon}}{2}$;

$$P_r(\text{at least one Type 1 Error}) \leq \varepsilon$$

- It works for a fixed set of k parameters to estimate or null hypotheses to test.Let p_i be the p-value for testing each null Hoi. All the technique does is to reject H_{oi} and thus Ho if $p_i \leq \frac{\varepsilon}{k}$, i.e each of the test is computed at an $\frac{\varepsilon}{k}$.

1.5.2 Holm–Bonferroni Method

This is a modified form of the ordinary bonferroni.Here let $p_1, p_2 ... p_k$ be the p-values for k tests. These tests are sorted in an ascending order. Now the individual null hypothesis,

H_{0is} are sorted together with the p-values.Now begginning from the smallest p-value say p_1, we reject H_{0i} if $p_j \leq \frac{\varepsilon}{k-j+1}$., for all $j = 1, ..i$. We continue to reject Hoi in ascending oreder of p-values till we get to the first non-significant p -value, then we stop. This technique produces more power gain since only the smallest p-value is compared to ε/k.

1.5.3 FDR Modification

This procedure requires tests to be independent. This controls the false dscovery rate. Just as in the Holm-Bonferroni method, we sort the p values and the hypotheses. Now beginiing from the largest p value, say p_1 ,we reject H_{oi} if $p_i < \frac{j\varepsilon}{k}$, for some $j \geq i$. We continue to reject, but now in descending order of p-values. This method does not control the SFWER.

Table 1.3: Summary of Bonferroni-Based Methods

Multiple Comparison Summary		
Reject H_0i if	Name	Control
$P(i) < \frac{\varepsilon}{k}$	Bonferroni	Simultaneous Confidence Intervals and Tests,StrongFWER
$P(j) < \frac{\varepsilon}{k-j+1}$	Holm	Tests, SFWER
$P(i) < \frac{j\varepsilon}{k}$	FDR	Independent Tests, FDR

1.5.4 Pairwise Comparisons

Pariwise comparisons refer to contrasts that study differences between two treatment means $\bar{y}_{i.} - \bar{y}_{j.}$. For k treatment groups there are $\binom{k}{2}$ different pairwise comparisons. There are some paiwise comparison methods that control the type I error rate at ε for all pairwise compaisons. They include Tukey HSD, LSD, SNK and several others. These procedures can be seen as t-tests for the pairwise comparison contrasts. Several pairwise comparisons methods are different based on the definition of the critical value u, and u may depend on

several things, including ε, the degrees of freedom for MSE, the number of treatments and several factors. Pairwise comparisons can be done using tests or confidence intervals.i.e we could either evaluate the confidence intervals for the differences of means $\mu_i - \mu_j$ or $\alpha_i - \alpha_j$ or test the hypothesis,$H_{0ij} : \mu_i = \mu_j$ or $H_{0ij} : \alpha_i = \alpha_j$ equivalently. However confidence bounds for differences of means, generally provide much information hence preferred.

1.6 Critical values of u for t tests

Criteria for Rejection of the Null; The null hypothesis is rejected if

$$\frac{|\bar{y}_{i.} - \bar{y}_{j.}|}{\sqrt{MSE}\sqrt{1/n_i + 1/n_j}} > u_1$$

Equivalently the test would reject the null if

$$|\bar{y}_{i.} - \bar{y}_{j.}| > u\sqrt{MSE}\sqrt{1/n_i + 1/n_j} = D_{ij}$$

where D_{ij} is the significant difference.Thus the confidence interval for the pairwise mean differences is obtained from the pairwise tests through

$$(\bar{y}_{i.} - \bar{y}_{j.}) \pm u\sqrt{MSE}\sqrt{1/n_i + 1/n_j}$$

1.7 Sample Size Estimation

Sample size estimation is a very significant step in planning a statistical study and usually a difficult one. The choice of sample size is significant because we want our experiment to be as small as possible to save time and money, but big enough to get the job done. There are two main ways to effectivley select a sample size.

- Specify a maximium confidence interval width for means or contrasts.
- Hypothesis testing of treatment effects/population parameters
- Specify the effect size for the hypothesis tests.

1. Specify a maximum confidence interval width for means.

   Here we would need the confidence intervals for contrasts be no wider than a specified length. The confidence interval width has dependence on the desired coverage, error variance sample size. Hence one must have an idea of at least one of these information available to plan an appropriate sample size.

2. Hypothesis Tests for Treatment effects. This procedure involves we either reject the null hypothesis or we fail to reject it. A type I error is made when we reject the true null. A type II error is made if we fail to reject a false null. The probability of avoiding a type II error, that is correctly rejecting the null and achieving statistical significance is called power. An important objective in planning a study is to ensure an acceptably high power level.The focus here is to determine a sufficient sample size to achieve a certain desired power.

1.7.1 Taking Multiplicity into Account with Sample Size Estimation

A major criticism frequently posed at multiple testing methods is their lack of power. There is always much difficulty in rejecting hypotheses hence reaching firm conclusions when multiple methods are used. In multiple testing situations, power and sample size is difficult to compute since its test statistics are based on simulations. A most compatible definition of Power with multiple testing issues with the aim of controlling the FWER at α is the probability of correctly rejecting at least one H_i for $i \in K'$. We could assume dependence or independence of tests when approximating power. However, incorporation of dependence structures most likely reduce the p-values ,hence sample sizes determined may be conservative. Again the extent of conservativeness depends upon the nature of correlations (compound symmetry, autoregressive(1),etc) among variables and tests.

1.7.2 Sample Size Formulas

We will consider sample size formulas continuous response random variables.

Continuous Outcomes

For a single hypothesis, the sample size problem is usually formulated in order to ensure the power $(1-\beta)$ of detecting any mean difference (standardized effect size) δ^* at a pre-specified Type 1 error rate ε (CWER). Assuming the hypothesis $H_0 : \mu = \mu_0 \ \ vrs \ \ H_1 : \mu \neq \mu_0$, we define the RR and AR as follows:

$$\text{RR (Rejection Region)} : \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > Z_{1-\varepsilon/2} \quad or \quad \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < Z_{\varepsilon/2}$$

$$\text{AR (Acceptance Region)} : \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < Z_{1-\varepsilon/2} \quad or \quad \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > Z_{\varepsilon/2}$$

To obtain an expression in terms of $(1-\beta)$, we compute the sample size using the lower tail of the standard normal distribution.

$$\beta(\text{Type II error}) = P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > Z_{\varepsilon/2} | H_0 \ \ false\right)$$

$$\beta = 1 - P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < Z_{\varepsilon/2} | H_0 \ \ false\right)$$

$$P\left(Z < \frac{\mu_0 - \mu^*}{\sigma/\sqrt{n}} + Z_{\varepsilon/2}\right) = 1 - \beta \ \ \cdots \ \ w.l.g$$

where μ^* is the observed mean and μ_0 is the hypothesized mean.

Let $\mu_0 - \mu^* = \delta$, the mean difference. Then,

$$\phi\left(\frac{\delta}{\sigma/\sqrt{n}} + Z_{\varepsilon/2}\right) = 1 - \beta \Longrightarrow \frac{\delta}{\sigma/\sqrt{n}} = Z_{1-\beta} - Z_{\varepsilon/2}, \ \ where \ \ \phi^{-1}(1-\beta) = Z_{1-\beta}$$

$$\frac{n\delta^2}{\sigma^2} = (Z_{1-\beta} - Z_{\varepsilon/2})^2$$

$$\therefore n \approx \frac{\sigma^2(Z_{\varepsilon/2} - Z_{1-\beta})^2}{\delta^2} = \frac{(Z_{\varepsilon/2} - Z_{1-\beta})^2}{\delta^{*2}}, \ \ where \ \ \delta^* = \delta/\sigma$$

Again for a two-sample z-test given ε and δ^*, the sample size needed in each group in order to achieve a desired power $(1-\beta)$ can be derived in a similar way to the one-sample test. Consider the hypothesis $H_0 : \mu_1 - \mu_2 = D_0 \ \ vrs \ \ H_1 : \mu_1 - \mu_2 \neq D_0$, we define the RR and AR as follows:

$$\text{RR (Rejection Region)} : \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > Z_{1-\varepsilon/2} \quad or \quad \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < Z_{\varepsilon/2}$$

$$\text{AR (Acceptance Region)} : \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < Z_{1-\varepsilon/2} \quad or \quad \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > Z_{\varepsilon/2}$$

Again in order to obtain an expression in terms of $(1-\beta)$, we calculate the sample size using the lower tail of the standard normal distribution. By assuming equal sample sizes;

$$\beta(\text{Type II error}) = P\left(\frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sigma\sqrt{\frac{1}{n} + \frac{1}{n}}} > Z_{\varepsilon/2} | H_0 \ \ false\right)$$

$$\beta = 1 - P\left(\frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sigma\sqrt{\frac{2}{n}}} < Z_{\varepsilon/2}\right) \ \ \cdots \ \ w.l.o.g$$

$$P\left(Z < \frac{D_0 - \mu^*}{\sigma\sqrt{\frac{2}{n}}} + Z_{\varepsilon/2}\right) = 1 - \beta$$

where μ^* is the observed mean difference and D_0 is the hypothesized mean difference we want to detect.

Again, let $D_0 - \mu^* = \delta$, then;

$$\phi\left(\frac{\delta}{\sigma\sqrt{\frac{2}{n}}} + Z_{\varepsilon/2}\right) = 1 - \beta \Longrightarrow \frac{\delta}{\sigma\sqrt{\frac{2}{n}}} = Z_{1-\beta} - Z_{\varepsilon/2} \quad , where \ \ \phi^{-1}(1-\beta) = Z_{1-\beta}$$

$$\frac{n\delta^2}{2\sigma^2} = (Z_{1-\beta} - Z_{\varepsilon/2})^2$$

$$\therefore n \approx \frac{2\sigma^2(Z_{\varepsilon/2} - Z_{1-\beta})^2}{\delta^2} = \frac{2(Z_{\varepsilon/2} - Z_{1-\beta})^2}{\delta^{*2}} \quad , where \ \ \delta^* = {}^{\delta}/_{\sigma_i} \qquad (1.1)$$

NOTE: Z_ε is the lower ε percentile of a standard normal distribution.

For the simultaneous analysis of a set of hypotheses sample size depends on ε, $(1-\beta)$ and δ^* of each individual hypothesis.

Chapter 2

Linear Mixed Models (LMM's)

2.1 Voilation of the Assumption of Independence of Linear Models

The independence assumption is by far the most important assumption of all statistical tests. If you elicit multiple responses from each subject, then those responses that come from the same subject cannot be regarded as independent from each other. Violating independence may greatly inflate your chance of finding a spurious result and it results in a p-value that is completely meaningless.

A lot of the times, we want to collect more data per subject, such as in repeated measures designs. If you end up with a data set that has non- independencies in it, you need to resolve these non-independencies at the analysis stage. This is where mixed models come in handy.

2.2 Introduction to Linear Mixed Models

Linear Mixed Models are a special case of Generalized Linear Mixed Models. It is composed of both fixed and random effects.This model assumes Gaussian data .i.e the observations conditioned on the random effects have a normal distribution. Is currently a commonplace in any literate conversation about statistical modeling. A complete statistical model of this form requires: A linear predictor $x\beta + Zb$. The distribution of the random model effects, $b \sim N(0, G)$. The distribution of the data conditioned on the random effects, $y|b$. The

link function, $g(E(y|b)) = x\beta + Zb = \eta$ This mixed model assumes a Gaussian data, hence its model being defined as follows: $y|b \sim N(X\beta + Zb, R)$ and $b \sim N(0, G)$ The marginal distribution of y is $N(X\beta, ZGZ' + R)$. A fixed effects-only model defined by $n = X\beta$ and variance $V = var(y) = ZGZ' + R$, which yields the same inference on the fixed effects as inference based on estimable functions from a mixed model. V is a covariance matrix whose structure accounts for both fixed and random effects. Now we notice that the fixed effects only affect the mean of y, while the components of the random effects only affect the variance of y.

2.2.1 Repeated Measures in Mixed Models

A repeated measures design is one in which measurements of the same variable are made on each subject on two or more different occasions. Is the most widely used experimental designs especially in the field of health sciences due to its ability to control for extraneous variation among subjects and also fewer subjecs are involved. An example will be measuring cholesterol or anxiety levels over time for several experimental conditions. Meaurements on each individual would be correlated. In general, pairs of observations adjacent in time are assumed to have a larger correlation than pairs of observations more separated in time. We will write the model for random subject effects and fixed factor A and factor B effects. In the example, factor A is the treatment group, and factor B is time. Let α_j and β_k denote the factor A and factor B main effects, respectively, $(\alpha\beta)_{jk}$ the AB interaction effect and ρ the subject (block) main effect.

The model is

$$Y_{ijk} = \mu + \rho_{ij} + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \epsilon_{ijk},$$

where:

μ is a constant

$\rho_{i(j)}$ are independent $N(0, \sigma^2_\rho)$

α_j are constants subject to $\sum \alpha_j = 0$

β_k are constants subject to $\sum \beta_k = 0$

$(\alpha\beta)_{jk}$ are constants subject to $\sum_j (\alpha\beta)_{jk} = 0$ for all k and $\sum_k (\alpha\beta)_{jk} = 0$ for all j

ϵ_{ijk} are independent $N(0, \sigma^2)$

ρ_{ij} and ϵ_{ijk} are independent

$i = 1, \ldots, n$; $j = 1, \ldots, a$; $k = 1, \ldots, b$.

From these assumptions, we can see that $\sigma_Y^2 = \sigma_\rho^2 + \sigma^2$ and $Cov(Y_{ijk}, Y_{ijk'}) = \sigma_\rho^2$, $k \neq k'$.

2.3 Covariance Structures for Repeated Measures in Mixed Models

The variance-covariance matrix of the repeated observations on the four observations for each subject within each (treatment) group has the form

Compound Symmetric Structure

$$\Sigma_{subject} = \frac{1}{\sigma_\rho^2 + \sigma^2} \begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix}.$$

Autoregressive Structure

$$\Sigma_{subject} = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$$

Some other structures include **Toeplitz** and **Unstructured** covariance structures.

2.4 Estimation And Inference of LMM'S

2.4.1 Estimation

The objective here, is to estimate parameters to construct inference statistics.This will involve how to;

1. Develop the estimating equations for the linear predictor effects, β, also referred to as mixed model equations

2. Develop estimating equations for the components of the variance-covariance matrices G and R seen earlier

The mixed model equations are maximum likelihood, however equivalent to generalized least squares for fixed effects, β. The estimating equations for σ i.e for the components of the variance-covariance matrices G and R could based on the full likelihood or residual likelihood. Full likelihood yields biased estimates which are generally avoided ,whiles the latter yield unbiased estimates for the covariance components, i.e REML(Restricted Maximum Likelihood) estimates of σ. Estimating equations for LMM's could be written exactly but their solutions require iteration.

2.4.2 Mixed Model Equations for β and b

The log-likelihood obtained from the distributional assumptions for $y|b$ and b respectively are:

$$\ell(y|b) = -\left(\frac{-n}{2}\right)log(2\pi - \left(\frac{1}{2}\right)log(|R|) - \left(\frac{1}{2}\right)(y - X\beta - Zb)'R^{-1}(y - X\beta - Zb)$$

and

$$\ell(\beta) = \left(- \frac{-b}{2}\right)log(2\pi) - \left(- \frac{1}{2}\right)log(|G|) - \left(- \frac{1}{2}\right)b'G^{-1}b$$

where b represents the total number of levels of the random effects. We now obtain a joint log-likelihood as

$$\ell(y, b) = \left(- \frac{1}{2}\right)(y - X\beta - Zb)'R^{-1}(y - X\beta - Zb)\left(- \frac{1}{2}\right)b'G^{-1}$$

The maximum likelihood estimator is obtained by setting $\partial\ell(y, b)/\partial\beta'$ and $\partial\ell(y, b)/\partial b'$ to zero and solving the set of equations for β and b. The derivatives are

$$\frac{\partial[\ell(y, b)]}{\partial\beta'} = X'R^{-1}y - X'R{-}1X\beta - Z'R{-}1X\beta$$

and

$$\frac{\partial[\ell(y,b)]}{\partial b'} = Z'R^{-1}y - X'R^{-1}Z\beta - Z'R^{-1}Zb - G^{-1}b$$

The mixed model equations for LMM is obatined as

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G{-}1 \end{bmatrix} \begin{bmatrix} \beta \\ b \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

Exact estimates of β and b are obtained when G and R are known. Otherwise we would have to estimate G and R as well for unknown components of G and R resorting to an iterative process.

2.4.3 Restricted Maximum Likelihood

This idea was presented comprehensively by Patterson nd Thompson(1971) for LMMs. The basic idea is to maximize the likelihood after taking into consideration the fixed effects of the model.We obtain estimators of the from the likelihood of $K'y$ where K is a matrix such that $E(K'y) = 0$, then $K'y \sim N(0, K'VK)$, instead of maximizing the likelihood of $y \sim N(X\beta, V)$. This process tends to effectively remove fixed effects from the estimation of σ.The likelihood $K'y$ is called the REML likelihood.

The REML log-likelihood is :

$$\ell(\sigma; y) = \left(\frac{n-p}{2}\right) log(2\pi) - \left(\frac{1}{2}\right) log\Big(|V(\sigma)|\Big) - \left(\frac{1}{2}\right) log\Big(|X'[V(\sigma)]^{-1}X|\Big) - \left(\frac{1}{2}\right) r'[V(\sigma)]^{-1}r$$

where $p = rank(X)$ and $r = y - X(X'[V(\sigma)]^{-1}X) - X'[V(\sigma)]^{-1}y$ r could be seen as $y - X\hat{\beta}_{ML}$, where $\hat{\beta}_{ML}$,is the maximum likelihood of β

2.4.4 Hypothesis Testing and Interval Estimation

Wald and Approximate F-Statistics

There are generally two approaches to hypothesis testing. They include the likelihood ratio and the Wald-Based. We focus on the Wald -based for linear mixed model settings. This

includes the Wald statistics and approximate F-statistics which is obtained by dividing the Wald statistics by the rank of the estimable or predictable functions. The LR statistics are most often used with fixed effects only.i.e General Linear Models.

Considering a hypothesis test about estimable and predictable functions we have; H_0 : $\Psi = \Psi_0$ where $\Psi = K'\beta + M'b$. Assuming $K'\beta$ satisfies estimability criteria,$\hat{\Psi} = K'\hat{\beta} + M'\hat{b}$ is the e-BLUP(estimated best linear unbiased predictor of Ψ).When $M = 0$,$K'\beta$ is the BLUE(Best Linear Unbiased estimator) for Gaussian models as in Linear mixed model settings.

The general form of the Wald statistic is $\hat{\Psi}[Var(\hat{\Psi}]^{-1}\hat{\Psi}$ where $Var(\hat{\Psi})$ is known. However in practice $Var(\hat{\Psi})$ is unknown hence estimated. We eventually have the result $\left[\hat{\Psi}'\left[Var(\hat{\Psi})^{-1}\hat{\Psi}\right]\right]/rank\Psi$ which has an approximate and in some cases exact F -distribution. Wald statistics are assumed to have a chi-squared distribution with degrees of freedom determined by rank of Ψ which is th rank of matrix K that defines Ψ

The LMM estimating equations are special cases of GLMM estimating equations where the LMM likelihood in its conditional and marginal forms are Gaussian and also uses the identity link hence easy to work wih using standard likelihood methodolgy. LMM estimating eqautions are seen earlier. Hence From GLMM estimating equations;

$$\begin{bmatrix} X'WX & X'WZ \\ Z'WX & Z'WZ + G^{-1} \end{bmatrix} \begin{bmatrix} \beta \\ b \end{bmatrix} = \begin{bmatrix} X'Wy* \\ Z'Wy* \end{bmatrix}$$

where $W = \left(DV_{\mu}^{\frac{1}{2}}AV_{\mu}^{\frac{1}{2}}D\right)^{-1}$, $D = \partial\mu/\partial\mu$ and $y^* = g(\mu^{\sim}) + D(y - \mu^{\sim})$ The following results are obtained for inference on $K'\beta + M'b$ assuming $K'\beta$ satisfies the criteria of estimability. In a more realistic case where covariance components are unknown, the covariance components must be estimated. Let C denote the generalized inverse of the left hand side of the GLMM estimating equations

$$\begin{bmatrix} X'WX & X'WZ \\ Z'WX & Z'WZ + G^{-1} \end{bmatrix}$$

and $L' = [K'M']$ Replacing C by $\hat{C}$ using $\hat{G}$, $\hat{R}$ and $\hat{W}$ yields

- $Var\left[K'\hat{\beta} + M'(\hat{b} - b)\right] = Var\left(L'\left[\begin{bmatrix}\hat{\beta}\\(\hat{b}-b)\end{bmatrix}\right)\right) \cong L'\hat{C}L$
- For scalar L, $t = \frac{\hat{\Psi}-\Psi_0}{\sqrt{L'\hat{C}L}} \sim t_{v2}$

A two sided confidence interval obtained for Ψ is $\hat{\Psi} = K'\hat{\beta} + M'\hat{b} \pm t_{v2,\alpha/2}\sqrt{L'CL}$ The fundamental hypothesis testing results is

$$\frac{L'\left[\begin{bmatrix}\hat{\beta}\\(\hat{b}-b)\end{bmatrix}\right](L'\hat{C}L)^{-1}\begin{bmatrix}\hat{\beta}\\(\hat{b}-b)'\end{bmatrix}L}{rank(L)} \sim F_{v1,v2,\varphi}$$

v_1=rank(L)

v_2=degrees of freedom used in estimating C

φ denotes non-centrality parameter.

2.5 Inferences on Fixed Effects of LMM's

As seen previously, we considered a general inference on both the fixed and random effects of the LMM. however in our study we will focus on just the fixed effect in the model i.e. $\hat{\Psi} = K_i'\hat{\beta}$ assuming $M = 0$. As such we obtain the BLUE as $K_i'\hat{\beta}$ for linear mixed model settings.

The formal tests and confidence intervals are as follows;

- Inferences made are based on the appropriate t- test statistic and/or confidence intervals in the form;
- **Test Statistic** $t = \frac{K_i'\hat{\beta}}{\hat{\sigma}_i}$ where $i = 1, ..., k^*$ where k^* is the number of test end points and k is the total number of variables (timepoints, treatments, etc.)

- k^* could be $\binom{k}{2}$ or $k-1$ for all-pairwise comparisons or comparison to a control among test endpoints respectively.
- **Confidence interval** $\mathrm{R(C)} = K_i'\hat{\beta} \pm t_{n-k(\frac{\alpha}{2k^*})}, \frac{\hat{\sigma}_i}{\sqrt{n-k}}$
- where $\hat{\sigma}_i$=diag($\sqrt{K_i'\hat{\Sigma}K_i'}$) hence a scalar quantity
- Also for $K_i'\hat{\beta}$ we pick the the i_{th} component of that matrix hence also a scalar as represented in the formular above.

Illustration

Let's illustrate the idea above for a clearer picture. $y = X\beta + Zu$, $\hat{\Psi} = K_i'\hat{\beta}$, $\hat{\Sigma}_{K_i'\hat{\beta}} = K_i'\hat{\Sigma}K_i$

- The test end endpoints are ;
  $H_0 : \mu_1 - \mu_4 = 0$, $H_0 : \mu_2 - \mu_4 = 0$, $H_0 : \mu_3 - \mu_4 = 0$ for a comparison to a control where our reference level is μ_4
- We could interchangably use μ for β_i beacause $\mu_i = \mu + \beta_i$,
  $\mu_i - \mu_j = (\mu + \beta i) - (\mu + \beta j) = \beta_i - \beta_j$

$$\text{Hence if } \hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \hat{\beta}_4 \end{bmatrix} \text{ then } K_i'\hat{\beta} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \hat{\beta}_4 \end{bmatrix} = \begin{bmatrix} \hat{\beta}_1 - \hat{\beta}_4 \\ \hat{\beta}_2 - \hat{\beta}_4 \\ \hat{\beta}_3 - \hat{\beta}_4 \end{bmatrix}$$

$$\hat{\Sigma}_{K_i'\hat{\beta}} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \hat{\sigma_1^2} & \hat{\sigma_{12}} & \hat{\sigma_{13}} & \hat{\sigma_{14}} \\ & \hat{\sigma_2^2} & \hat{\sigma_{23}} & \hat{\sigma_{24}} \\ & & \hat{\sigma_3^2} & \hat{\sigma_{34}} \\ & & & \hat{\sigma_4^2} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \end{bmatrix}$$

$$\hat{\Sigma}_{K_i'\hat{\beta}} = \begin{bmatrix} \hat{\sigma_1^2} - 2\hat{\sigma_{14}} + \hat{\sigma_4^2} & \hat{\sigma_{12}} - \hat{\sigma_{24}} - \hat{\sigma_{14}} + \hat{\sigma_4^2} & \hat{\sigma_{13}} - \hat{\sigma_{14}} - \hat{\sigma_{34}} + \hat{\sigma_4^2} \\ & \hat{\sigma_2^2} - 2\hat{\sigma_{24}} + \hat{\sigma_4^2} & \hat{\sigma_{23}} - \hat{\sigma_{24}} - \hat{\sigma_{34}} + \hat{\sigma_4^2} \\ & & \hat{\sigma_3^2} - 2\hat{\sigma_{34}} + \hat{\sigma_4^2} \end{bmatrix}$$

The t test statistic and confidence interval for $H_0 : \mu_1 - \mu_4 = 0$ will be

- **Test Statistic** $t = \frac{\hat{\beta}_1 - \hat{\beta}_4}{\hat{\sigma}_i}$
- **Confidence interval** $\mathrm{R(C)} = \hat{\beta}_1 - \hat{\beta}_4 \pm t_{n-k(\frac{\alpha}{2k^*})}, \frac{\hat{\sigma}_i}{\sqrt{n-k}}$
- where $\hat{\sigma_i} = \sqrt{\hat{\sigma_1^2} - 2\hat{\sigma_{14}} + \hat{\sigma_4^2}}$

2.6 Sample Size Methods For Linear Mixed Models

Power and sample size estimation for variety of linear models have a higher level of complexity relative to simple hypothesis testing. Here we describe a number of steps involved in obtaining the required information to perform these computations.

1. Define the study design. This is the structure of the planned study.It must be clearly and completely specified.It involves treatment groups, primary outcome variables to be assessed.

2. Propose a mathematical model(scenario model). This explains and gives a much more detailed picture of the nature of data being collected. This is to help capture significant features of the study design and existing relationships amongst response measures and study factors

3. Make specific conjectures about the parameters in the model

4. Choose statistical methods to address your research question best. It includes defining statistical models and procedures for the data analysis,and inferences.

5. Clearly express your aim of assessment so that the power and sample size for the study is met.

6. In the case of a study involving multiple inferences,identify the set of comparisons(all pairwise or comparison to control) to be made.

7. Identify the variance and correlation patterns between test endpoints. Failing to specify variance and correlation patterns among observations can lead to incorrect sample size estimates. The various correlation structures include; 1. zero correlation (independent observations) 2. equal correlations 3. rule-based patterns(AR-1 structure or compound symmetry) 4. unstructured correlatons(no specific pattern)

8. Generate sample based on desired settings: Estimate the power for a given n or vice versa.

Mixed Models, i.e. models with both fixed and random effects arise in a variety of research situations. Split plots, strip plots, repeated measures, multi-site clinical trials, hierarchical linear models, random coefficients, analysis of covariance are all special cases of the mixed model In linear mixed model settings such as repeated measures and longitudinal studies, the fixed effects could be considered as time and the random effects could be the variation existing among the subjects. Traditionally we would test the hypothesis of the effect of a time-treatment interaction. These models study changes over time or space and the effect of treatments on these changes. Hence there is a key characteristic common to these models which is the possibilty of correlation between observations on the same subject. The responses in longitudinal studies are therefore highly correlated since they are performed on the same subjects multiple times. Sample sizes for these studies should therefore account for this high degree of correlation when they are created. Also in the case of multiple inferences made, multiplcity of inferences is warranted. This study considers linear mixed models, which incoporates the evaluation of multiple end points(several hypopothesis) hence etimates of the sample size for a given power should account for the multipilicty and

also the kind of correlation structure existing amongst observations Generally, for cases involving multiple comparisons,power considerations require you to specify exactly, the kind of inferences to be made for power.

- **Simulation** This remains a viable approach for evaluating sample size analysis irrespective of the availability of approximate or exact formulas for power computations. With a simulation approach, an analyst would first specify a data-generating distribution that corresponds to her research design and model assumptions, design ,effects and variability. Then compute test statistics, and determine when the null hypothesis is rejected. Then repeat the process and estimate power as proportion of rejections.

Chapter 3

Proposed Methodology

3.1 Multiple Correction for Dependent Test Endpoints

When multiple comparisons are simultaneously evaluated, a procedure which controls the familywise error rate is warranted.There are several methods that address multiplicity corrections for independent or normally distributed outcomes. Most of these methods take into consideration the correlation structure existing among these test endpoints.(e.g., Hunter, 1976; McCann & Edwards, 1996; Naiman, 1987; Sun & Loader, 1994; Worsley, 1982. However, when the endpoints are not normally distributed, these methods are not directly applicable.

Resampling-based methods can be applied for non normally corrlated otcomes. An example is (e.g. Westfall and Troendle, 2008) its application to correlated chi-square outcomes. Resampling-based procedures provide strong solutions in varied settings but are at times criticized for being "computationally demanding" (Pan, W., 2009, p.3) and sometimes do not allow for computation of interval-based inferences. In linear mixed model settings controlling multiplicity is achieved with reasonable power and sample size estimates using the Bonferroni procedure for a small set of comparisons. However, when the set of inferences is large, the Bonferroni procedure is generally too conservative and more powerful bounds are warranted. Also the Bonferroni procedure assumes independence among testendpoints which is a flaw for LMMs. This is because there would always be some sort of dependencies existing for LMM's (longitudinal studies). Hence we will apply the method described by Hunter and Worsley using the minimal spanning tree algorithm which seeks to control at least the FWER, for these end points, which are t dependent random variables (normally

distributed outocomes in the case of linear mixed model setiings)and also accounts for the complex dependency structure among these endpoints.

Hence, the objective is to;

- Bound the probability of the union of all m sets, i.e. for all events $A_i = \{Y_i > c\}, \quad i = 1, 2, \ldots m.$ where Y_i is a random variable and c a constant.
- Develop a computationally efficient algorithm for estimating sample sizes based on this bound while controlling the FWER.

The Bonferroni inequality, given my the FORMULA, is utilized.

$$P\Big(\bigcup_{i=1}^{m} A_i\Big) \leq \sum_{i=1}^{m} P\Big(A_i\Big) = \varepsilon,$$

However the Bonferroni procedure tends to be too conservative, hence more powerful bounds are needed. Hunter (1970) and Worsley (1982) improved this inequality by tightening the upper bound on $P\Big(\cup_{i=1}^{m} A_i\Big)$ using a tree spanning algorithm. The use of the tree algorithm is justified by the inequality

$$P\Big(\bigcup_{i=1}^{m} A_i\Big) \leq \sum_{i=1}^{m} P\Big(A_i\Big) - \sum_{(i,j)\in\tau} P\Big(A_i \cap A_j\Big). \tag{3.1}$$

where is a subgraph (tree) of G, defined by the sets, A_i , $i = 1, ..., k$. This methodology involves choosing a point c such that the right-hand side of (3.1) equals α. The set of nodes for the tree are the $A_i s$ and the set of branches are the intersections of $A_i A_j s$.The algorithm finds the shortest path within G that optimizes the spanning tree τ until the R.H.S is as close to ε as possible and this yields c_{hw} (the hunter-worsely critical value). This provides an improved solution over Bonferroni in all settings.

3.1.1 Hunter- Worsely Bound

Given a complete graph whose vertices are identified by the events. The weight of the edge connecting A_i and A_j is $P\Big(A_i \cap A_j\Big)$. Let $T = (G, \tau)$ be the maximum weight spanning

tree on the complete graph. τ=edges of the maximum weight spanning tree.

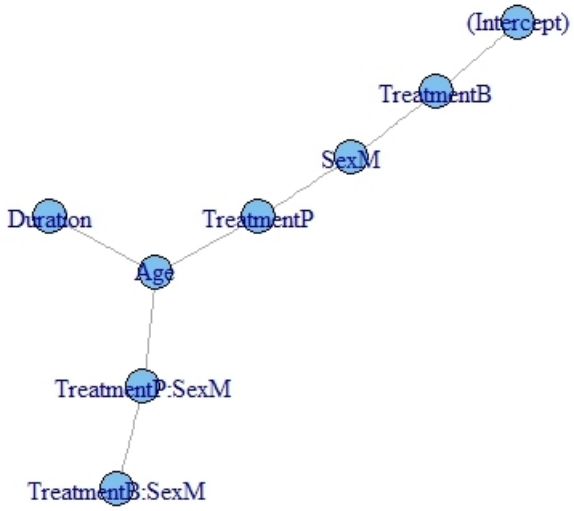

Figure 3.1: Graph of Tree Based Algorithm

Eventually $c_{hw} < c_{bon}$. Again $P(Y_i > c_{hw}) = \varepsilon_{hw} \neq \frac{\varepsilon}{k^*}$ yields $\varepsilon_{adj} = k^* \times \varepsilon_{hw} \geq \varepsilon$. This allows for an upward adjusted error rate under observed dependency. It only demands the computations of the univariate and bivariate cumulative distributions of the Y_i interest. However, in order to evaluate the joint probability on the right hand side of equation (3.1), calculation of bivariate t probabilities are necessary. The aim is to find c such that $P\Big(\cup_{i=1}^{m} A_i\Big)$ is bounded by α (here $\mathbf{Y} \sim \ \boldsymbol{t_m(\nu, R)}$, where m is the dimension, ν is the degrees of freedom and $\mathbf{R}$ is the correlation structures of endpoints). $\mathbf{R}$ would be assumed to have a compound symmetric (CSS) or an autoregressive correlation structure (AR).

Compound Symmetric Structure **Autoregressive Structure**

$$\begin{bmatrix} 1 & \rho_a & \rho_a & \rho_a \\ \rho_a & 1 & \rho_a & \rho_a \\ \rho_a & \rho_a & 1 & \rho_a \\ \rho_a & \rho_a & \rho_a & 1 \end{bmatrix}. \qquad \begin{bmatrix} 1 & \rho_a & \rho_a^2 & \rho_a^3 \\ \rho_a & 1 & \rho_a & \rho_a^2 \\ \rho_a^2 & \rho_a & 1 & \rho_a \\ \rho_a^3 & \rho_a^2 & \rho_a & 1 \end{bmatrix}$$

Dunnett and Sobel (1954) gave an expression for the joint t density function of m variates $(Y_1, Y_2, \ldots, Y_m)$. The bivariate t density function is given by

$$P\Big(Y_1 = c_1, Y_2 = c_2; \rho_{12}\Big) = \frac{1}{2\pi\sqrt{1-\rho_{12}^2}}\left[1 + \frac{c_1^2 - 2\rho_{12}^2 c_1 c_2 + c_2^2}{\nu(1-\rho_{12}^2)}\right]^{-\frac{1}{2}(\nu+2)},$$

where $Corr(Y_1, Y_2) = \rho_{12}$.

For either the normal or t distributions, the bound given in (3.1) allows for a single evaluation of the optimal that maximizes the regardless of c. Thus, different values of c may be evaluated for a single optimal spanning tree until the right hand side of (3.1) is as close as possible to α .

3.2 Advantages of Graph Based MCs over Bonferroni Method

Using Equations (3.1) and (3.2), improvements to a Bonferroni critical value in linear mixed model settings(normally distributed outcomes) may be obtained. This is an application of the minimal spanning tree algorithm to bivariate t. Each spanning tree, we obtain depends on the relative differences of the ρ_{ij}. Consequently, the tree only needs to be found once for a particular covariance structure and can subsequently be used for evaluating (3.1) for any value c. Simulation of critical values is also an approach that would be useful in a variety of settings. However, it is not always clear how to go about the simulation, particularly to researchers in fields other than statistics, and implementation is far from

direct and can be computationally intensive. Consequently, we present a flexible and simple method for multiplicity correction that often competes well even with a simulation-based approach. Computational efficiency is considerably obtained when many types of sets of inferences may be of interest, particularly since simulation or resampling-based methods would demand an inefficient recalculation of the critical point for each set of comparisons while this method would allow for very efficient recalculation of each critical point.

3.3 Sample Size Estimation For Dependent Testendpoints in LMM settings

Power and sample size analysis for linear mixed models is currently an active area of reearch. Is also particularly complicated for mixed models because of the availability of a wide range of statistical tests. Generally there is no accepted standard for these calculations in linear mixed models, with both fixed and random effects. A recommended approach is simulation. Several research has been done in these settings, however did not take into account multiplcity. O'Brien and Muller(1993) demonstrate an exact poer calculation for a one-way random effects ANOVA using a multiple cantra F distribution. Lenth(2000) also computes approximate power for balanced ANOVA designs with fixed and random effectswhere the random effects have a compound symmetric correlation structure.(i.e random effects are mutually independent)

Considering a study involving multiple inferences(more than one primary hypothesis), there arises a need to account for multiplicity of errors to make valid conclusions. In most cases each power analysis must be dependent on one specific hypothesis using a preplanned data analysis method. With a small number of multiple comparisons, a simple Bonferroni correction is applied to control the type I error rate. For instance with 10 primary hypothesis, a type I error rate of $\alpha = 0.05/10 = 0.005$ would be used.Conducting 10 power analysis leads to 10 different power values or ideal sample sizes. Eventually the

largest sample size to guarantee power for all the 10 tests is chosen.

Chapter 4

Results and Analysis

This chapter presents the results from our case studies. We use the single factor as a simple case and increase complexity of model to the and 2 factor repeated measures design as our linear mixed model settings. Analysis and results outcomes will provide grounds for comparing the proposed method with the naive (bonferroni) approach of sample size calculation. We would compare the critical values, type 1 error rate used in obtaining sample sizes in the Bonferroni approach and the proposed Hunter worsely method.

For the naive (bonferroni) aproach , consider k^* multiple dependent endpoints, bonferroni the type one error rate and choose the largest sample size warranted by our adjusted type one error. For our proposed method, we apply the Hunter Worsley method which eventually warrants an upward adjusted type 1 error rate. Now bonferroni the hunter-worsley adjusted type 1 error rate and use in calculating sample sizes which warrants an improvement over the Bonferroni- based (naive) approach.

4.1 Case Study 1

4.1.1 Single Factor Repeated Measures Design

Let's consider the Sleep Study Experiment. This laboratory experiment measured the effect of sleep deprivation on cognitive performance. There were 18 subjects, chosen from the population of interest (long-distance truck drivers), in the 10 day trial. These subjects were restricted to 3 hours sleep per night during the trial. Again we see that time is nested within subject, hence we incoporate that in fitting a model for this study. How many

subjects are nedded to achieve 80% , 90% power with 5% error?

4.1.2 Model Fitting

Here we would construct a random intercept and slope model. Our fixed effect is Days and the random effects would be the random intercepts for the subjects and also the by-subject random slopes for the effect of Days. Hence we would expect differing baseline-levels of reaction(the intercept) as well as differing responses for different subjcets for the effect of Days. For the single-factor repeated measures design

$$Y_{ij} = \mu + \beta_i + \rho_j + \epsilon_{ij}, \quad i = 1, \ldots, n \quad j = 1, \ldots, k,$$

where n is the number of subjects and k is the number of treatments. It makes sense to assume here that $\beta_i \sim N(0, \sigma^2_\beta)$.

4.1.3 Hypothesis Testing and Sample Size Estimation

Our interest is to test the difference in the mean reaction across the various time points(Days). In this study we have $1 - 10$ days ($k = 10$). We would employ "Comparison to a Control". Day 10 would be our reference level.

Hence we would have $9(k-1)$ set of comparisons, $H_0 : \mu_1 - \mu_{10} = 0$, $H_0 : \mu_2 - \mu_{10} = 0$, $H_0 : \mu_3 - \mu_{10} = 0$, $H_0 : \mu_4 - \mu_{10} = 0$, ... , $H_0 : \mu_9 - \mu_{10} = 0$, which yields multiple test dependent endpoints, which warrants the problem of multipilicity. We would apply the naive (traditional bonferroni) approach and the proposed approach to estimate sample sizes, based on different error rates from the two approaches.

Table 4.1: Sample size Estimates - Case Study 1, Comparison to a Control , 80% Power

Method	**Sample Size**	$\delta_1 = 45$	$\delta_2 = 68$	$\delta_3 = 136$
Naive	n_a	59	26	6
Proposed	n_p	54	24	6

Table 4.2: Sample size Estimates - Case Study 1, Comparison to a Control , 90% Power

Method	**Sample Size**	$\delta_1 = 45$	$\delta_2 = 68$	$\delta_3 = 136$
Naive	n_a	74	33	8
Proposed	n_p	69	30	8

Comments

In comparison of both approaches we see an increase in the type 1 error rate warranted by the Hunter worsley critical value as against the bonferroni error rate. Also sample size estimates are smaller for our proposed methodology. For insatnce for $\delta = 68$ and 80% power we obtained a sample of 24 as against 26 (a decrease of 2) which would be studied across time points.

4.2 Case Study 2

4.2.1 Two Factor Repeated Measures Design

Yucha conducted a study to determine if nursing students who were assigned to a home hospital (HH) experience differed from those traditionally placed (TP) in hospitals throughout their nursing training. Anxiety, as measured by Spielberger's State Anxiety Scale (where

higher scores suggest higher levels of anxiety), is the within-subjects variable and is provided at four points in time during nursing training. How many subjects are needed in each treatment-time combination to achieve 80% , 90% power and 5% error?

There is a significant interaction between placement type and time as seen in nmerical results and Figure 4.3 as expected. Hence we would construct a random intercept and slope model. Our fixed effect is the interaction between time and treatment and the random effects would be the random intercepts for the subjects and also the by-subject random slopes for the effect of time . We will write the model for random subject effects and fixed factor A and factor B effects. In the example, factor A is the treatment group(hospital placement), and factor B is time. Let α_j and β_k denote the factor A and factor B main effects, respectively, $(\alpha\beta)_{jk}$ the AB time-treatment interaction effect and ρ the subject (block) main effect.

The model is

$$Y_{ijk} = \mu + \rho_{ij} + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \epsilon_{ijk},$$

μ is a constant, ρ_{ij} are independent $N(0, \sigma_\rho^2)$

α_j are constants subject to $\sum \alpha_j = 0$

β_k are constants subject to $\sum \beta_k = 0$

$(\alpha\beta)_{jk}$ are constants subject to $\sum_j (\alpha\beta)_{jk} = 0$ for all k and $\sum_k (\alpha\beta)_{jk} = 0$ for all j

ϵ_{ijk} are independent $N(0, \sigma^2)$

$\rho_{i(j)}$ and ϵ_{ijk} are independent

$i = 1, \ldots, n$; $j = 1, \ldots, a$; $k = 1, \ldots, b$.

From these assumptions, we can see that $\sigma_Y^2 = \sigma_\rho^2 + \sigma^2$ and $Cov(Y_{ijk}, Y_{ijk'}) = \sigma_\rho^2$, $k \neq k'$.

4.2.2 Hypothesis Testing and Sample Size Estimation

Actually we could test several hypothesis based on our model, however our hypothesis of interest is to test the difference in the mean anxiety levels between those in a home hospital placement versus traditional placementthrough time for these nursing students.

There were four time points during the nursing training. Hence we would have these set of comparisons, $H_0 : \mu_{HH_{ii}} - \mu_{TP_{ii}} = 0$, where$i = 1, ..., 4$, Again we have multiple dependent test endpoints and will follow as previous to obtain sample size estimates.

Table 4.3: Sample size Estimates- Case Study 2, 80% Power

Method	**Sample Size**	$\delta_1 = 11$	$\delta_2 = 16$	$\delta_3 = 32$
Naive	n_a	30	14	4
Proposed	n_p	30	14	4

Table 4.4: Sample size Estimates- Case Study 2, 90% Power

Method	**Sample Size**	$\delta_1 = 11$	$\delta_2 = 16$	$\delta_3 = 32$
Naive	n_a	39	19	5
Proposed	n_p	38	18	5

Comments

In comparison of both approaches we see an increase in the type 1 error rate warranted by the Hunter worsley critical value as against the bonferroni error rate. Also sample size estimates are smaller for our proposed methodology. For insatnce for $\delta = 16$ we obtained a sample of 14 vrs 15(for each treatment group;hospital placement) studied across time points. Again we see this as a counter example. There are 3 set of comparisons which is a few hence bonferroni approach will yield quite reasonable sample size estimates. Even though our proposed method works better, the difference in the sample size estimates for both approaches is not quite significant. However for lager set of comparisons in the previous we noticed an increase in difference in sample size estimates. On this basis we will say that as the set of comparisons become large the proposed methodolgy yields much smaller

sample sizes showing a much significant difference in sample size estimates compared to the bonferroni since the bonferroni becomes much more conservative.

4.3 Sample size estimation and Tests for Various Covariance structures

In all cases we fit a LMM just as seen for the Sleep study case and analyse results for different covariance structures existing among response measures across time points. On a general case, lets consider an **AR(1) covariance structure** with $\rho = 0.5$. Our interest is to test the difference in the mean response measure across the various time points. In this study we have $1 - 20(k = 20)$ time points. We would employ "Comparison to a Control. Time point 20 would be our reference level. Hence we would have $19(k - 1)$ set of comparisons, as seen earlier.

Table 4.5: AR(1) Covariance Strucrure, Comparison to a control, 80% Power

Method	**Sample Size**	$\delta_1 = 6$	$\delta_2 = 8$	$\delta_3 = 10$
Naive	n_a	31	18	11
Proposed	n_p	29	17	11

Table 4.6: AR(1) Covariance Structure), Comparison to a control, 90% Power

Method	**Sample Size**	$\delta_1 = 6$	$\delta_2 = 8$	$\delta_3 = 10$
Naive	n_a	39	22	14
Proposed	n_p	37	21	13

Now let's consider a **Compound Symmetric Covariance structure** with $\rho = 0.5$

Table 4.7: Compound Symmetric Structure, Comparison to a control, 80% Power

Method	**Sample Size**	$\delta_1 = 6$	$\delta_2 = 8$	$\delta_3 = 10$
Naive	n_a	16	9	6
Proposed	n_p	15	9	5

Table 4.8: Compound Symmetric Structure, Comparison to a control, 90% Power

Method	**Sample Size**	$\delta_1 = 6$	$\delta_2 = 8$	$\delta_3 = 10$
Naive	n_a	19	11	7
Proposed	n_p	19	11	7

Now let's consider a **Independent Covariance structure** with $\rho = 0$

Table 4.9: Zero correlation (Independent Covariance Structure), Comparison to a control, 80% Power

Method	**Sample Size**	$\delta_1 = 6$	$\delta_2 = 8$	$\delta_3 = 10$
Naive	n_a	31	18	11
Proposed	n_p	30	17	11

Table 4.10: Zero correlation (Independent Covariance Structure), Comparison to a control, 90% Power

Method	**Sample Size**	$\delta_1 = 6$	$\delta_2 = 8$	$\delta_3 = 10$
Naive	n_a	38	22	14
Proposed	n_p	38	21	14

4.4 Case Study 3

Our main concern as stated earlier on is to know what happens to sample size estimates as power of the test and dependecies among test end points increases. We illustrate with this extensive case study. We would consider the various covariance structures (AR(1), Compound Symmetry, Indepenent) as well as the various form of contrasts for the means (all pairwise, comparison to control)

Case Study

Improved aethestics are often developed by first studying their effects on animals. In one study 20 dogs were initially given the drug pentobarbitol. Each dog was then administered carbon dioxide CO_2 at each of two pressure levels. Next halothane (H) was added, and the administration of CO_2 was repeated. The response, milliseconds between heartbeats, was measured for the four treatment combinations.

Our question then is how many subjects are nedded to achieve 80% and 90% power with 5% error? The experimenter desires to analyse the anesthetizing effects of carbon dioxide from this repeated measures design. In all cases we fit a LMM just as seen for the Sleep study case and analyse results for different covariance structures existing among response measures across the treatment combinations. On a general case, lets consider an **AR(1) covariance structure** with $\rho = 0.5$. In this study we have $1 - 4(k = 4)$ treatments. We would employ "Comparison to a Control". Treatment 4 would be our reference level.

Hence we would have $3(k-1)$ set of comparisons, as seen earlier.

Table 4.11: AR(1) Covariance Structure- Case Study 3, Comparison to a control, 80% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	12	4	2
Proposed	n_p	12	4	2

Table 4.12: AR(1) Covariance Structure) - Case Study 3, Comparison to a control, 90% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	16	6	2
Proposed	n_p	15	5	2

Now let's consider a **Compound Symmetric Covariance structure** with $\rho = 0.5$

Table 4.13: Compound Symmetric Structure - Case Study 3, Comparison to a control,80% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	7	3	1
Proposed	n_p	7	2	1

Table 4.14: Compound Symmetric Structure- Case Study 3, Comparison to a control, 90% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	9	3	1
Proposed	n_p	9	3	1

Now let's consider an **Independent Covariance structure** with $\rho = 0$

Table 4.15: Zero correlation (Independent Covariance Structure) - Case Study 3, Comparison to a control, 80% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	14	5	2
Proposed	n_p	14	5	2

Table 4.16: Zero correlation (Independent Covariance Structure) - Case Study 3, Comparison to a control, 90% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	18	6	3
Proposed	n_p	18	6	2

Now lets consider **all pairwise** contrasts of means , and for an **AR(1) covariance structure** with $\rho = 0.5$. Hence we would have $\binom{4}{2}$=6 set of comparisons, $H_0 : \mu_1 - \mu_2 = 0$, $H_0 : \mu_1 - \mu_3 = 0$, $H_0 : \mu_1 - \mu_4 = 0$, $H_0 : \mu_2 - \mu_3 = 0$, $H_0 : \mu_2 - \mu_4 = 0$, $H_0 : \mu_3 - \mu_4 = 0$.

Now let's consider a **Compound Symmetric Covariance structure** with $\rho = 0.5$

Table 4.17: AR(1) Covariance Structure - Case Study 3, All-Pairwise Comparison, 80% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	28	10	4
Proposed	n_p	27	10	4

Table 4.18: AR(1) Covariance Structure- Case Study 3, All-Pairwise Comparison, 90% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	36	13	5
Proposed	n_p	34	12	5

Table 4.19: Compound Symmetric Structure - Case Study 3, All-Pairwise Comparison, 80% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	16	6	2
Proposed	n_p	16	6	2

Table 4.20: Compound Symmetric Structure - Case Study 3, All-Pairwise Comparison, 90% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	20	7	3
Proposed	n_p	20	7	3

Now let's consider a **Independent Covariance structure** with $\rho = 0$

Table 4.21: Zero correlation (Independent Covariance Structure) - Case Study 3, All-Pairwise Comparison, 80% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	32	12	5
Proposed	n_p	31	11	4

Table 4.22: Zero correlation (Independent Covariance Structure), - Case Study 3, All-Pairwise Comparison, 90% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	41	15	6
Proposed	n_p	40	14	6

4.4.1 Application to higher order models

Now assuming the experimenter desires to analyze the anesthetizing effects of CO_2 pressure and halothane from this repeated measures design. The effects include, Treatment 1= high CO_2 pressure without H, Treatment 2= low CO_2 pressure without H, Treatment 3= high CO_2 pressure with H, Treatment 4= low CO_2 pressure with H, There are three treatment contrasts that might be of interest in the experiment. Let $\mu_1, \mu_2, \mu_3, \mu_4$ correspond to the mean responses for treatments 1,2,3, and 4 respectively.

Then $(\mu_3 + \mu_4) - (\mu_1 + \mu_2) =$ (Halothane contrast representing the difference between the presence and absence of halothane)

Then $(\mu_1 + \mu_3) - (\mu_2 + \mu_4) =$ (Halothane contrast representing the difference between the presence and absence of halothane)

Then $(\mu_1 + \mu_4) - (\mu_2 + \mu_3)$ = (Halothane contrast representing the difference between the presence and absence of halothane)

With μ'=$[\mu_1, \mu_2, \mu_3, \mu_4]$ the contrast matrix s $K = \begin{bmatrix} -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$

We now have an **a priori** (unique) contrast matrix based on the researcher's interests. What sample size estimates will be required to guarantee 80% and 90% power incorporating the various forms of covariance structures?

Table 4.23: AR(1) Covariance Structure, - Case Study 3, A priori Contrasts, 80% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	26	9	4
Proposed	n_p	25	9	4

Table 4.24: AR(1) Covariance Structure)- Case Study 3, A priori Contrasts, 90% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	34	12	5
Proposed	n_p	32	12	5

Now let's consider a **Compound Symmetric Covariance structure** with $\rho = 0.5$

Table 4.25: Compound Symmetric Structure, - Case Study 3, A priori Contrasts, 80% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	14	5	2
Proposed	n_p	14	5	2

Table 4.26: Compound Symmetric Structure, - Case Study 3, A priori Contrasts, 90% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	18	7	3
Proposed	n_p	18	6	2

Now let's consider a **Independent Covariance structure** with $\rho = 0$

Table 4.27: Zero correlation (Independent Covariance Structure), - Case Study 3, A priori Contrasts, 80% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	28	10	4
Proposed	n_p	27	9	4

Table 4.28: Zero correlation (Independent Covariance Structure), - Case Study 3, A priori Contrasts, 90% Power

Method	**Sample Size**	$\delta_1 = 3$	$\delta_2 = 5$	$\delta_3 = 8$
Naive	n_a	36	13	5
Proposed	n_p	35	13	5

Chapter 5

Conclusion and Recommendation

5.1 Summary

There are power and sample size analyses methods that have been proposed, however for specific cases. Power and sample size analysis is particularly complicated for linear mixed models, due to the wide variety of statistical tests that exists.

The Bonferroni approach is widely used in obtaining sample sizes for multiple endpoints in such settings where the maximum sample size is chosen. However, this thesis has led to the development of a novel approach for sample size estimation in linear mixed model settings, taking into account the dependency structure of test endpoints whiles controlling for multiplicity, which improves over the maximum sample size in the Bonferroni based approach.

From the case study results, our proposed method performs relatively better compared to the naive approach. However, we would compare these two mehodologies based on different situations and settings.

From our numerical analysis, we would conclude that as power increases, there is little to no effect on the effectiveness of the proposed method compared to the bonferroni. Power of the test is therefore not a factor to be considered in comparing the efficacy of the proposed method to bonferroni. Secondly we see an advantage in using the proposed method as effect size decreases, as smaller effect sizes yields much smaller sample size estimates. Again we see no difference in both methods when ρ increases. This is beacause the correlations are induced by contrast matrix K. Lastly all-pairwise comparisons among test endpoints our proposed methodology yields more gains compared to bonferroni, whereas there is little to

no effect on the efficacy of the proposed method compared to bonferroni for comaprison to a control among test end points.

5.2 Recommendations

To an experimenter who seeks the best course of action under situations of sample size estimation in linear mixed model settings (repeated measures or longitudinal studies) we propose the following;

1. When considering smaller effect sizes, the proposed methodology yields more gains hence has an advantage over the bonferroni
2. When taking into account all-pairwise comparisons among test endpoints, the proposed methodology again yields much gains over bonferroni
3. For power and correlation considerations, both methods are efficient and effective as they warrant similar sample size estimates.

References

[1] Gary W. Oehlert, University of Minnesota. *A First Course in Design and Analysis of Experiments*, 2010.

[2] Keith E. Muller, Paul W. Stewart. *Linear Model Theory .Univariate, Multivariate and Mixed Models*, 2006

[3] Walter W. Stroup. *Generalized Linear Mixed Models.Modern Concepts Methods and Applications*, 2013

[4] Peter H. Westfall and S.stanley Young. *Resampling Based Multipe Testing ; Examples and Methods For P- Value Adjustment*, 1993

[5] Robert O. Kheul *Statistical Principles of Research Design and Analysis, Second Edition*, 2000

[6] Guo et.al (2013). *Selecting a sample size for studies with repeated measures*, BMC Biomedical Research Methodolgy.

[7] Ralph G. O'Brien et.al (1998) *Power and Sample Size Determination for Linear Models* Cleveland Clinic Foundation, Cleveland, OH

[8] Casella, G and Berger, R (2001). *Statistical Inference* Brooks/Cole, Boston, MA, USA, second edition, pg 188-194.

[9] Wayne W. Daniel and Chad L.Cross (2013) *Biostatistics A Foundation for Analysis in the Health Sciences*

[10] Dunnett, C. W., & Sobel, M. (1954). *A bivariate generalization of Student's t-distribution, with tables for certain special cases.* Biometrika, 41(1-2), 153-169.

[11] Hunter D.(1976) *An Upper Bound for the Probability of Union*Journal of Applied Probability Vol. 13, No. 3 (Sep., 1976), pp. 597-603

[12] Wagler, A. & McCann, M. (in review). *A multiplicity adjustment for chi-square distributed endpoints.* In review at Journal of Statistical Computation and Simulation.

[13] Melinda McCann and Don Edwards (1996) *A Path Length Inequality for the Multivariate-t Distribution, With Applications to Multiple Comparisons* Journal of the American Statistical AssociationVol. 91, No. 433 pp. 211-216

[14] Worsley K.J. (1982) *An improved Bonferroni Inequality and Applications* Biometrika 69, 297-302

Appendix

R CODES

```
# Load packages & set working directory

# -------------------------------------

library(igraph)
library(cubature)
library("ape")
library("RBGL")
#library("graph")
library("MSBVAR")
library("dlm")
library("Matrix")
library("nlme")
library(MASS)
library("mvtnorm")


setwd("C:/Users/Documents/rstudiofiles")


library(TSP)
library(scatterplot3d )
```

```
library(MVA )
library(igraph)



#################### Hunter-Worsley ######################
hw<-function(C,alpha,nu,V){
#V=V_hat
R=cov2cor(C%*%V%*%t(C))
acosR<-acos(R)


####### Step One: Find minimum spanning tree ####################
dis<-graph.adjacency(acosR,mode="max",weighted=TRUE)
mst <- minimum.spanning.tree(dis)
#lay <- layout.reingold.tilford(dis, mode="all")
#plot(mst, layout=lay)
#plot(mst)


####### Step Two: Numerical Integration Method######
p<-nrow(C)
r=qr(C)$rank

intg<- function(d)
{

fx <- function(x){
```

```
B<-p*pf(((d*x)^(-2)-1)/(r-1), r-1, 1)

if (r==2) {for (i in 1:(p-1)) {
phi<-E(mst)$weight[i]
B<-B-max(-phi/pi+2*acos(x*d)/pi,0)}
}

else {
for (i in 1:(p-1)) {
phi<-E(mst)$weight[i]

gx<-function(w)
{acos(d*x*sqrt(1/(1-w)))*w^(r/2-2)}
#print(d)
gx<-Vectorize(gx)
if (((cos(phi/2)/(x*d))^2-1) >0 ) {
B<-B+phi/pi*pf(2*max(0,(cos(phi/2)/(x*d))^2-1)/(r-2),r-2,2)-(r-2)/pi*(integrate(gx,
0,1-1/(1+max(0,(cos(phi/2)/(x*d))^2-1)))$value) }
else B<-B+phi/pi*pf(2*max(0,(cos(phi/2)/(x*d))^2-1)/(r-2),r-2,2)

}
}

return (B*df(r*x^2, nu, r)*2*r*x)
```

```
}



fx<-Vectorize(fx)
gf <- function(k) {integrate(fx, lower=0, upper=1/k)$value}
#d=qt(1-.05,nu);d=sqrt(r*qf(1-alpha,r, nu))
return (gf(d)-alpha)}

######### Step Three:  Root Finding Algorithm ############
secant <- function(fun, x0, x1, tol=1e-4, niter=500){
for ( i in 1:niter ) {#fun=intg
x2 <- x1-fun(x1)*(x1-x0)/(fun(x1)-fun(x0))
if (abs(fun(x2)) < tol)
return(x2)
x0 <- x1
x1 <- x2
#print(c(x0,x1,x2))
}
stop("exceeded allowed number of iteractions")
}
secant(intg, x0=qt(1-alpha/2,nu), x1=qt(1-alpha/(2*choose(k,2)),nu))
}

#### Example #####

## Parameter setup
k=10 #number of variables
```

```
times <- 1:k
rho <- 0.9

#AR(1) structure
H <- abs(outer(times, times, "-"))
V_hat <-  (k-1)*(rho)^H

#CS structure
V_hat=matrix(rho,k,k)
diag(V_hat)=1
C<-cbind(diag(k-1),rep(-1,k-1)) #comparison to a control

alpha=0.05
n=20 #sample size

nu=n-k #effective df

## Compute Hunter-Worsley critical value
(hw.t=hw(C,alpha,nu,V_hat))
(a.adj=nrow(C)*2*pt(hw.t,nu,lower.tail=F))
qt(1-alpha/2,nu);qt(1-alpha/(2*(k-1)),nu)

#new value for epsilon
old method episilon (PCER)= alpha/(nrow(C))
a.adj/(nrow(C))

#complete the code to incorporate strategies 1 and 2
```

```
> hw<-function(C,alpha,nu,V){
+   #V=V_hat
+   R=cov2cor(C%*%V%*%t(C))
+   acosR<-acos(R)
+
+
+   ####### Step One: Find minimum spanning tree ####################
+   dis<-graph.adjacency(acosR,mode="max",weighted=TRUE)
+   mst <- minimum.spanning.tree(dis)
+   #lay <- layout.reingold.tilford(dis, mode="all")
+   #plot(mst, layout=lay)
+   #plot(mst)
+
+
+   ####### Step Two: Numerical Integration Method######
+   p<-nrow(C)
+   r=qr(C)$rank
+
+   intg<- function(d)
+   {
+
+     fx <- function(x){
```

```
+
+       B<-p*pf(((d*x)^(-2)-1)/(r-1), r-1, 1)
+
+       if (r==2) {for (i in 1:(p-1)) {
+         phi<-E(mst)$weight[i]
+         B<-B-max(-phi/pi+2*acos(x*d)/pi,0)}
+       }
+
+       else {
+         for (i in 1:(p-1)) {
+           phi<-E(mst)$weight[i]
+
+           gx<-function(w)
+           {acos(d*x*sqrt(1/(1-w)))*w^(r/2-2)}
+           #print(d)
+           gx<-Vectorize(gx)
+           if (((cos(phi/2)/(x*d))^2-1) >0 ) {
+             B<-B+phi/pi*pf(2*max(0,(cos(phi/2)/(x*d))^2-1)/(r-2),r-2,2)-(r-2)/pi*(int
+
}
+           else B<-B+phi/pi*pf(2*max(0,(cos(phi/2)/(x*d))^2-1)/(r-2),r-2,2)
+
+         }
+       }
+
+       return (B*df(r*x^2, nu, r)*2*r*x)
+
+
```

```
+
+     }
+
+
+
+     fx<-Vectorize(fx)
+     gf <- function(k) {integrate(fx, lower=0, upper=1/k)$value}
+     #d=qt(1-.05,nu);d=sqrt(r*qf(1-alpha,r, nu))
+     return (gf(d)-alpha)}
+
+   ######### Step Three:  Root Finding Algorithm #############
+   secant <- function(fun, x0, x1, tol=1e-4, niter=500){
+     for ( i in 1:niter ) {#fun=intg
+       x2 <- x1-fun(x1)*(x1-x0)/(fun(x1)-fun(x0))
+       if (abs(fun(x2)) < tol)
+         return(x2)
+       x0 <- x1
+       x1 <- x2
+       #print(c(x0,x1,x2))
+     }
+     stop("exceeded allowed number of iteractions")
+   }
+   secant(intg, x0=qt(1-alpha/2,nu), x1=qt(1-alpha/(2*choose(k,2)),nu))
+ }
> (hw.t=hw(C,alpha=0.05,nu,V=CovCS))
[1] 2.926557
```

```
#Sleep Study Data

> summary(fm1_a)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: Reaction ~ Days + (1 + Days | Subject)
Data: sleepstudy

AIC      BIC   logLik deviance df.resid
1763.9   1783.1   -876.0   1751.9      174

Scaled residuals:
Min      1Q  Median      3Q     Max
-3.9416 -0.4656  0.0289  0.4636  5.1793

Random effects:
Groups   Name        Variance Std.Dev. Corr
Subject  (Intercept) 565.52   23.781
Days         32.68    5.717   0.08
Residual             654.94   25.592
Number of obs: 180, groups:  Subject, 18

Fixed effects:
Estimate Std. Error t value
(Intercept)  251.405      6.632   37.91
Days          10.467      1.502    6.97

Correlation of Fixed Effects:
(Intr)
```

```
Days -0.138

#80% power; Sleep study

#just alpha with bonferroni(traditional)
naive(alpha=0.05/9,beta=0.2,delta=68,V=CovCS,C=C)
[1] 18.645062 17.840637 20.821315 17.247020 12.537202  7.669204 25.902889 14.557690
[9]  5.665958
naive(alpha=0.05/9,beta=0.2,delta=45,V=CovCS,C=C)
[1] 42.57519 40.73832 47.54457 39.38282 28.62816 17.51230 59.14813 33.24186 12.93797
> naive(alpha=0.05/9,beta=0.2,delta=136,V=CovCS,C=C)
[1] 4.661265 4.460159 5.205329 4.311755 3.134301 1.917301 6.475722 3.639423 1.416489

n=18
k=10
n
[1] 18
k
[1] 10
nu=n-k
nu
[1] 8
(hw.t=hw(C,alpha=0.05,nu,V=cov2cor(CovCS)))
[1] 3.438117

qt(1-0.05/18,8)
[1] 3.758586
```

```
p_hw=1-pt(hw.t,nu)
P_hw= 2*p_hw
alpha_hw=nrow(C)*P_hw
alpha_hw
[1] 0.07963489
alpha_hw/9
[1] 0.008848322

#with adjusted alpha
naive(alpha_hw/9,beta=0.2,delta=68,V=CovCS,C)
[1] 17.079651 16.342765 19.073189 15.798987 11.484598  7.025309 23.728122
[8] 13.335449  5.190253
naive(alpha_hw/9,beta=0.2,delta=45,V=CovCS,C)
[1] 39.00064 37.31800 43.55280 36.07630 26.22458 16.04199 54.18214 30.45092
[9] 11.85172
naive(alpha_hw/9,beta=0.2,delta=136,V=CovCS,C)
[1] 4.269913 4.085691 4.768297 3.949747 2.871149 1.756327 5.932031 3.333862
[9] 1.297563


#90% Power; Sleep Study
#just alpha with bonferroni(traditional)
naive(alpha=0.05/9,beta=0.1,delta=68,V=CovCS,C=C)
[1] 23.459888 22.447732 26.198128 21.700822 15.774759  9.649669 32.591948
[8] 18.317011  7.129112
naive(alpha=0.05/9,beta=0.1,delta=45,V=CovCS,C=C)
[1] 53.56964 51.25843 59.82229 49.55289 36.02098 22.03460 74.42231 41.82610
[9] 16.27902
```

```
naive(alpha=0.05/9,beta=0.1,delta=136,V=CovCS,C=C)
[1] 5.864972 5.611933 6.549532 5.425205 3.943690 2.412417 8.147987 4.579253
[9] 1.782278

nu=8
#with adjusted alpha
naive(alpha_hw/9,beta=0.1,delta=68,V=CovCS,C)
[1] 21.699773 20.763556 24.232572 20.072684 14.591233  8.925687 30.146686
[8] 16.942748  6.594239
naive(alpha_hw/9,beta=0.1,delta=45,V=CovCS,C)
[1] 49.55049 47.41268 55.33403 45.83511 33.31845 20.38142 68.83865 38.68803
[9] 15.05766
naive(alpha_hw/9,beta=0.1,delta=136,V=CovCS,C)
[1] 5.424943 5.190889 6.058143 5.018171 3.647808 2.231422 7.536672 4.235687
[9] 1.648560



##Anxiety data##

fit=ezANOVA(data=anx_data,dv=anxiety,wid=SUBJ,within=TIMED,between = TREAT,type = 3,deta
> print(fit)
$ANOVA
Effect DFn DFd          SSn      SSd             F            p p<.05
1 (Intercept)   1  18 172329.6125 7797.125 397.83035734 1.009516e-13     *
2       TREAT   1  18     32.5125 7797.125   0.07505651 7.872311e-01
3       TIMED   3  54    591.4375 3475.175   3.06340688 3.568792e-02     *
4 TREAT:TIMED   3  54    685.1375 3475.175   3.54873496 2.029868e-02     *
```

```
ges
1 0.938604670
2 0.002875988
3 0.049852544
4 0.057298021

$‘Mauchly’s Test for Sphericity‘
Effect         W         p p<.05
3       TIMED 0.5900661 0.1169751
4 TREAT:TIMED 0.5900661 0.1169751

$‘Sphericity Corrections‘
Effect       GGe      p[GG] p[GG]<.05      HFe      p[HF] p[HF]<.05
3       TIMED 0.8169178 0.04702471         * 0.955337 0.03816227         *
4 TREAT:TIMED 0.8169178 0.02909222         * 0.955337 0.02215315         *


 mfm1=lmer(anxiety~TREAT*TIME+(1+TIME|SUBJ),anx_data)
> summary(mfm1)
Linear mixed model fit by REML [’lmerMod’]
Formula: anxiety ~ TREAT * TIME + (1 + TIME | SUBJ)
Data: anx_data

REML criterion at convergence: 581.2

Scaled residuals:
Min       1Q   Median       3Q      Max
-2.52765 -0.47718  0.03952  0.56368  1.71640
```

```
Random effects:
Groups   Name        Variance Std.Dev. Corr
SUBJ     (Intercept) 131.824  11.481
TIME          4.944   2.224   -0.54
Residual              57.093   7.556
Number of obs: 80, groups:  SUBJ, 20

Fixed effects:
Estimate Std. Error t value
(Intercept)   58.200      4.663  12.480
TREAT2       -12.750      6.595  -1.933
TIME          -4.460      1.279  -3.487
TREAT2:TIME    4.590      1.809   2.537

Correlation of Fixed Effects:
(Intr) TREAT2 TIME
TREAT2      -0.707
TIME        -0.709  0.501
TREAT2:TIME  0.501 -0.709 -0.707
> mfm2=lmer(anxiety~TREAT+TIME+(1+TIME|SUBJ),anx_data)
> summary(mfm2)
Linear mixed model fit by REML ['lmerMod']
Formula: anxiety ~ TREAT + TIME + (1 + TIME | SUBJ)
Data: anx_data

REML criterion at convergence: 590
```

```
Scaled residuals:
Min       1Q   Median       3Q      Max
-2.57265 -0.58254  0.01055  0.57108  1.80681

Random effects:
Groups   Name        Variance Std.Dev. Corr
SUBJ     (Intercept) 163.102  12.771
TIME          9.627   3.103   -0.65
Residual              57.093   7.556
Number of obs: 80, groups:  SUBJ, 20

Fixed effects:
Estimate Std. Error t value
(Intercept)  52.2688     4.2244  12.373
TREAT2       -0.8877     4.6514  -0.191
TIME         -2.1650     1.0258  -2.111

Correlation of Fixed Effects:
(Intr) TREAT2
TREAT2 -0.551
TIME   -0.628  0.000
```

##Apply hunter worsley on anxiety

```
> k=4
> n=14
> nu=n-k
> nu
[1] 10
> View(CX3)
> CX33 = matrix( c(k-3,0,0,0,0,0,k-3,0,0,0,0,k-3),nrow=3,ncol=4,byrow = TRUE)
> CX33
[,1] [,2] [,3] [,4]
[1,]    1    0    0    0
[2,]    0    0    1    0
[3,]    0    0    0    1


#80% power; Anxiety study

#just alpha with bonferroni(traditional)
naive(alpha=0.05/3,beta=0.2,delta=32,V=Ssa,C=CX33)
[1] 3.553773 3.593377 2.759120
naive(alpha=0.05/3,beta=0.2,delta=16,V=Ssa,C=CX33)
[1] 14.21509 14.37351 11.03648
naive(alpha=0.05/3,beta=0.2,delta=11,V=Ssa,C=CX33)
[1] 30.07491 30.41007 23.34990


k=4
n=20
nu=n-k
```

```
k
[1] 4
n
[1] 20
nu
[1] 16

(hw.t=hw(C=CX33,alpha=0.05,nu,V=cov2cor(Ssa)))
[1] 2.619978

qt(1-0.05/6,16)
2.673032

p_hw=1-pt(hw.t,nu)
> P_hw= 2*p_hw
> alpha_hw=nrow(CX33)*P_hw
> alpha_hw
[1] 0.05571308
> alpha_hw/3
[1] 0.01857103
> 0.05/3
[1] 0.01666667

#with adjusted alpha
> naive(alpha=alpha_hw/3,beta=0.2,delta=16,V=Ssa,C=CX33)
[1] 13.86625 14.02078 10.76564
> naive(alpha=alpha_hw/3,beta=0.2,delta=11,V=Ssa,C=CX33)
[1] 29.33686 29.66379 22.77689
```

```
> naive(alpha=alpha_hw/3,beta=0.2,delta=32,V=Ssa,C=CX33)
[1] 3.466562 3.505194 2.691410


#90% power Anxiety data

#just alpha with bonferroni(traditional)
naive(alpha=0.05/3,beta=0.1,delta=11,V=Ssa,C=CX33)
[1] 38.80920 39.24169 30.13113
naive(alpha=0.05/3,beta=0.1,delta=16,V=Ssa,C=CX33)
[1] 18.34341 18.54783 14.24167
naive(alpha=0.05/3,beta=0.1,delta=32,V=Ssa,C=CX33)
[1] 4.585852 4.636958 3.560417


#with adjusted alpha
> naive(alpha=alpha_hw/3,beta=0.1,delta=32,V=Ssa,C=CX33)
[1] 4.486710 4.536710 3.483443
> naive(alpha=alpha_hw/3,beta=0.1,delta=11,V=Ssa,C=CX33)
[1] 37.97017 38.39332 29.47972
> naive(alpha=alpha_hw/3,beta=0.1,delta=16,V=Ssa,C=CX33)
[1] 17.94684 18.14684 13.93377


##AR(1) Structure power(80%)
k=20 #number of variables
times <- 1:k
rho <- 0.5
```

```
H <- abs(outer(times, times, "-"))
V <-  (k-1)*(rho)^H
V

#just alpha with bonferroni(traditional)
naive(alpha=0.05/19,beta=0.2,delta=10,V=V,C=C20)
[1] 11.261613 11.261592 11.261549 11.261463 11.261291 11.260947 11.260260
[8] 11.258885 11.256136 11.250637 11.239639 11.217644 11.173653 11.085672
[15] 10.909709 10.557782  9.853930  8.446226  5.630817
naive(alpha=0.05/19,beta=0.2,delta=8,V=V,C=C20)
[1] 17.596271 17.596237 17.596170 17.596036 17.595767 17.595230 17.594156
[8] 17.592008 17.587712 17.579120 17.561936 17.527569 17.458833 17.321362
[15] 17.046420 16.496535 15.396766 13.197228  8.798152
naive(alpha=0.05/19,beta=0.2,delta=6,V=V,C=C20)
[1] 31.28226 31.28220 31.28208 31.28184 31.28136 31.28041 31.27850 31.27468
[9] 31.26704 31.25177 31.22122 31.16012 31.03793 30.79353 30.30475 29.32717
[17] 27.37203 23.46174 15.64116

#with adjusted alpha
(hw.t=hw(C20,alpha=0.05,nu,V=cov2cor(V)))
[1] 3.749694
> p_hw=1-pt(hw.t,nu)
> P_hw= 2*p_hw
> alpha_hw=nrow(C20)*P_hw
> alpha_hw
[1] 0.07192866
> 0.05/19
[1] 0.002631579
```

```
> alpha_hw/19
[1] 0.003785719
> qt(1-0.05/38,10)
[1] 3.972759

> naive(alpha=alpha_hw/19,beta=0.2,delta=10,V=V,C=C20)
[1] 10.614118 10.614097 10.614057 10.613976 10.613814 10.613490 10.612842
[8] 10.611546 10.608955 10.603772 10.593407 10.572676 10.531215 10.448292
[15] 10.282446  9.950754  9.287371  7.960603  5.307069
> naive(alpha=alpha_hw/19,beta=0.2,delta=8,V=V,C=C20)
[1] 16.584559 16.584527 16.584464 16.584337 16.584084 16.583578 16.582566
[8] 16.580541 16.576492 16.568394 16.552198 16.519807 16.455023 16.325456
[15] 16.066322 15.548053 14.511516 12.438443  8.292295
> naive(alpha=alpha_hw/19,beta=0.2,delta=6,V=V,C=C20)
[1] 29.48366 29.48360 29.48349 29.48327 29.48282 29.48192 29.48012 29.47652
[9] 29.46932 29.45492 29.42613 29.36855 29.25337 29.02303 28.56235 27.64098
[17] 25.79825 22.11279 14.74186



# AR(1) 90% Power
#just alpha with bonferroni(traditional)

naive(alpha=0.05/19,beta=0.1,delta=10,V=V,C=C20)
[1] 13.98277 13.98275 13.98269 13.98259 13.98237 13.98195 13.98109 13.97939
[9] 13.97597 13.96915 13.95549 13.92818 13.87356 13.76432 13.54584 13.10888
[17] 12.23495 10.48710  6.99140
> naive(alpha=0.05/19,beta=0.1,delta=8,V=V,C=C20)
[1] 21.84808 21.84804 21.84796 21.84779 21.84746 21.84679 21.84546 21.84279
```

```
[9] 21.83746 21.82679 21.80545 21.76278 21.67744 21.50675 21.16537 20.48262
[17] 19.11711 16.38609 10.92406
> naive(alpha=0.05/19,beta=0.1,delta=6,V=V,C=C20)
[1] 38.84104 38.84096 38.84082 38.84052 38.83993 38.83874 38.83637 38.83163
[9] 38.82215 38.80318 38.76525 38.68939 38.53767 38.23422 37.62733 36.41354
[17] 33.98597 29.13083 19.42056

#with adjusted alpha
naive(alpha=alpha_hw/19,beta=0.1,delta=10,V=V,C=C20)
[1] 13.260183 13.260158 13.260108 13.260006 13.259804 13.259399 13.258590
[8] 13.256971 13.253734 13.247259 13.234310 13.208411 13.156613 13.053018
[15] 12.845827 12.431446 11.602683  9.945157  6.630104
naive(alpha=alpha_hw/19,beta=0.1,delta=8,V=V,C=C20)
[1] 20.71904 20.71900 20.71892 20.71876 20.71844 20.71781 20.71655 20.71402
[9] 20.70896 20.69884 20.67861 20.63814 20.55721 20.39534 20.07161 19.42413
[17] 18.12919 15.53931 10.35954
naive(alpha=alpha_hw/19,beta=0.1,delta=6,V=V,C=C20)
[1] 36.83384 36.83377 36.83363 36.83335 36.83279 36.83167 36.82942 36.82492
[9] 36.81593 36.79794 36.76197 36.69003 36.54615 36.25838 35.68285 34.53179
[17] 32.22967 27.62543 18.41696



#Compound Symmetry 80% power
rho
[1] 0.5
k
[1] 20
V_cs=matrix(rho,k,k)
```

```
diag(V_cs)=1
CV_cs=(k-1)*V_cs

#just alpha with bonferroni(traditional)
naive(alpha=0.05/19,beta=0.2,delta=6,V=CV_cs,C=C20)
[1] 15.64116 15.64116 15.64116 15.64116 15.64116 15.64116 15.64116 15.64116
[9] 15.64116 15.64116 15.64116 15.64116 15.64116 15.64116 15.64116 15.64116
[17] 15.64116 15.64116 15.64116
> naive(alpha=0.05/19,beta=0.2,delta=8,V=CV_cs,C=C20)
[1] 8.798152 8.798152 8.798152 8.798152 8.798152 8.798152 8.798152 8.798152
[9] 8.798152 8.798152 8.798152 8.798152 8.798152 8.798152 8.798152 8.798152
[17] 8.798152 8.798152 8.798152
> naive(alpha=0.05/19,beta=0.2,delta=10,V=CV_cs,C=C20)
[1] 5.630817 5.630817 5.630817 5.630817 5.630817 5.630817 5.630817 5.630817
[9] 5.630817 5.630817 5.630817 5.630817 5.630817 5.630817 5.630817 5.630817
[17] 5.630817 5.630817 5.630817

nu
[1] 10
> (hw.t=hw(C20,alpha=0.05,nu,V=cov2cor(CV_cs)))
[1] 3.879992
> p_hw=1-pt(hw.t,nu)
> P_hw= 2*p_hw
> alpha_hw=nrow(CV_cs)*P_hw
> alpha_hw
[1] 0.06117944
alpha_hw/19
[1] 0.003219971
```

```
0.05/19
[1] 0.002631579
qt(1-0.05/38,10)
[1] 3.972759

#with adjusted alpha
naive(alpha=alpha_hw/19,beta=0.2,delta=6,V=CV_cs,C=C20)
[1] 15.14243 15.14243 15.14243 15.14243 15.14243 15.14243 15.14243 15.14243
[9] 15.14243 15.14243 15.14243 15.14243 15.14243 15.14243 15.14243 15.14243
[17] 15.14243 15.14243 15.14243
> naive(alpha=alpha_hw/19,beta=0.2,delta=8,V=CV_cs,C=C20)
[1] 8.517616 8.517616 8.517616 8.517616 8.517616 8.517616 8.517616 8.517616
[9] 8.517616 8.517616 8.517616 8.517616 8.517616 8.517616 8.517616 8.517616
[17] 8.517616 8.517616 8.517616
> naive(alpha=alpha_hw/19,beta=0.2,delta=10,V=CV_cs,C=C20)
[1] 5.451274 5.451274 5.451274 5.451274 5.451274 5.451274 5.451274 5.451274
[9] 5.451274 5.451274 5.451274 5.451274 5.451274 5.451274 5.451274 5.451274
[17] 5.451274 5.451274 5.451274



#Compound Symmetry 90% power
##just alpha with bonferroni(traditional)
naive(alpha=0.05/19,beta=0.1,delta=6,V=CV_cs,C=C20)
[1] 19.42056 19.42056 19.42056 19.42056 19.42056 19.42056 19.42056 19.42056
[9] 19.42056 19.42056 19.42056 19.42056 19.42056 19.42056 19.42056 19.42056
[17] 19.42056 19.42056 19.42056
> naive(alpha=0.05/19,beta=0.1,delta=8,V=CV_cs,C=C20)
[1] 10.92406 10.92406 10.92406 10.92406 10.92406 10.92406 10.92406 10.92406
```

```
[9] 10.92406 10.92406 10.92406 10.92406 10.92406 10.92406 10.92406 10.92406
[17] 10.92406 10.92406 10.92406
> naive(alpha=0.05/19,beta=0.1,delta=10,V=CV_cs,C=C20)
[1] 6.9914 6.9914 6.9914 6.9914 6.9914 6.9914 6.9914 6.9914 6.9914 6.9914 6.9914
[12] 6.9914 6.9914 6.9914 6.9914 6.9914 6.9914 6.9914 6.9914


#with adjusted alpha


naive(alpha=alpha_hw/19,beta=0.1,delta=6,V=CV_cs,C=C20)
[1] 18.86437 18.86437 18.86437 18.86437 18.86437 18.86437 18.86437 18.86437
[9] 18.86437 18.86437 18.86437 18.86437 18.86437 18.86437 18.86437 18.86437
[17] 18.86437 18.86437 18.86437
> naive(alpha=alpha_hw/19,beta=0.1,delta=8,V=CV_cs,C=C20)
[1] 10.61121 10.61121 10.61121 10.61121 10.61121 10.61121 10.61121 10.61121
[9] 10.61121 10.61121 10.61121 10.61121 10.61121 10.61121 10.61121 10.61121
[17] 10.61121 10.61121 10.61121
> naive(alpha=alpha_hw/19,beta=0.1,delta=10,V=CV_cs,C=C20)
[1] 6.791172 6.791172 6.791172 6.791172 6.791172 6.791172 6.791172 6.791172
[9] 6.791172 6.791172 6.791172 6.791172 6.791172 6.791172 6.791172 6.791172
[17] 6.791172 6.791172 6.791172


#Independence Structure power(80%)
rho=0
[1] 0
k
[1] 20
```

```
V_ind=matrix(rho,k,k)
diag(V_ind)=1
CV_ind=(k-1)*V_ind

##just alpha with bonferroni(traditional)

naive(alpha=0.05/19,beta=0.2,delta=6,V=CV_ind,C=C20)
[1] 31.28232 31.28232 31.28232 31.28232 31.28232 31.28232 31.28232 31.28232
[9] 31.28232 31.28232 31.28232 31.28232 31.28232 31.28232 31.28232 31.28232
[17] 31.28232 31.28232 31.28232
> naive(alpha=0.05/19,beta=0.2,delta=8,V=CV_ind,C=C20)
[1] 17.5963 17.5963 17.5963 17.5963 17.5963 17.5963 17.5963 17.5963 17.5963
[10] 17.5963 17.5963 17.5963 17.5963 17.5963 17.5963 17.5963 17.5963 17.5963
[19] 17.5963
> naive(alpha=0.05/19,beta=0.2,delta=10,V=CV_ind,C=C20)
[1] 11.26163 11.26163 11.26163 11.26163 11.26163 11.26163 11.26163 11.26163
[9] 11.26163 11.26163 11.26163 11.26163 11.26163 11.26163 11.26163 11.26163
[17] 11.26163 11.26163 11.26163

(hw.t=hw(C20,alpha=0.05,nu,V=cov2cor(CV_ind)))
[1] 3.879992
alpha_hw
[1] 0.06117944
alpha_hw/19
[1] 0.003219971
0.05/19
[1] 0.002631579
qt(1-0.05/38,10)
```

```
[1] 3.972759

#with adjusted alpha
> naive(alpha=alpha_hw/19,beta=0.2,delta=6,V=CV_ind,C=C20)
[1] 30.28486 30.28486 30.28486 30.28486 30.28486 30.28486 30.28486 30.28486
[9] 30.28486 30.28486 30.28486 30.28486 30.28486 30.28486 30.28486 30.28486
[17] 30.28486 30.28486 30.28486
> naive(alpha=alpha_hw/19,beta=0.2,delta=8,V=CV_ind,C=C20)
[1] 17.03523 17.03523 17.03523 17.03523 17.03523 17.03523 17.03523 17.03523
[9] 17.03523 17.03523 17.03523 17.03523 17.03523 17.03523 17.03523 17.03523
[17] 17.03523 17.03523 17.03523
> naive(alpha=alpha_hw/19,beta=0.2,delta=10,V=CV_ind,C=C20)
[1] 10.90255 10.90255 10.90255 10.90255 10.90255 10.90255 10.90255 10.90255
[9] 10.90255 10.90255 10.90255 10.90255 10.90255 10.90255 10.90255 10.90255
[17] 10.90255 10.90255 10.90255

#Independence Structure power(90%)

##just alpha with bonferroni(traditional)
naive(alpha=0.05/19,beta=0.1,delta=10,V=CV_ind,C=C20)
[1] 13.9828 13.9828 13.9828 13.9828 13.9828 13.9828 13.9828 13.9828 13.9828
[10] 13.9828 13.9828 13.9828 13.9828 13.9828 13.9828 13.9828 13.9828 13.9828
[19] 13.9828
> naive(alpha=0.05/19,beta=0.1,delta=8,V=CV_ind,C=C20)
[1] 21.84813 21.84813 21.84813 21.84813 21.84813 21.84813 21.84813 21.84813
[9] 21.84813 21.84813 21.84813 21.84813 21.84813 21.84813 21.84813 21.84813
[17] 21.84813 21.84813 21.84813
> naive(alpha=0.05/19,beta=0.1,delta=6,V=CV_ind,C=C20)
```

```
[1] 38.84111 38.84111 38.84111 38.84111 38.84111 38.84111 38.84111 38.84111
[9] 38.84111 38.84111 38.84111 38.84111 38.84111 38.84111 38.84111 38.84111
[17] 38.84111 38.84111 38.84111


#with adjusted alpha
[1] 13.58234 13.58234 13.58234 13.58234 13.58234 13.58234 13.58234 13.58234
[9] 13.58234 13.58234 13.58234 13.58234 13.58234 13.58234 13.58234 13.58234
[17] 13.58234 13.58234 13.58234
> naive(alpha=alpha_hw/19,beta=0.1,delta=8,V=CV_ind,C=C20)
[1] 21.22241 21.22241 21.22241 21.22241 21.22241 21.22241 21.22241 21.22241
[9] 21.22241 21.22241 21.22241 21.22241 21.22241 21.22241 21.22241 21.22241
[17] 21.22241 21.22241 21.22241
> naive(alpha=alpha_hw/19,beta=0.1,delta=6,V=CV_ind,C=C20)
[1] 37.72873 37.72873 37.72873 37.72873 37.72873 37.72873 37.72873 37.72873
[9] 37.72873 37.72873 37.72873 37.72873 37.72873 37.72873 37.72873 37.72873
[17] 37.72873 37.72873 37.72873
```

Curriculum Vitae

Michael Nsiah-Nimo, born on July 2, 1992 is the first son of Edward Nsiah-Nimo and Evelyn Odame. After completing St.Peter's Senior High School, Nkwatia-Kwahu in June 2009, he continued his college education a year later in Kwame Nkrumah University of Science and Technology (KNUST), Kumasi where he pursued a bachelor's degree in Actuarial Science, graduating with First Class Honors. At KNUST, Michael Nsiah-Nimo showed active involvement in student fellowships and organisations which broadened his horizon in leadership roles and responsibilities. He worked as a Teaching Assistant, in KNUST due to his excellent academic standards and achievements

In Fall 2015, he entered the Graduate School of The University of Texas at El Paso (UTEP) pursuing a master's degree in Statistics. While in UTEP, he worked as a Teaching Assistant and coordinator at the Math Tututoring Center. He also engaged in Pre-Calculus and Summer school workshops where he was mainly involved in tutuoring. He is currently working under the supervision and mentorship of Dr. Amy Wagler, conducting research in Sample Size Estimation for Linear Mixed Model Settings with Dependent End Points. He plans to continue his studies in Ph.D Biostatistics in the near future.

Present address: 803 Upson Dr, Apt # 2,
El Paso, Texas 79902.