Accounting for Bias and Uncertainty in Power for Multivariate Gaussian Linear Models

Sola Park

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biostatistics, School of Public Health

Chapel Hill

2007

Approved by

Chairman: Keith E. Muller

Reader: Christopher S. Coffey

Reader: C. Ed Davis

Reader: Amy Herring

Reader: Stephen R. Aylward

ABSTRACT

SOLA PARK: Accounting for Bias and Uncertainty in Power

for Multivariate Gaussian Linear Models

(Under the direction of Keith E. Muller)

In choosing a sample size for a study with a Gaussian outcome, scientists can nearly always specify, perhaps with some prodding, mean differences of clinical and scientific importance. Any difficulty in providing a believable power analysis revolves around having a believable value for error variance. Multivariate or repeated measures makes the problem far worse. The uncertainty of the result depends not only on the individual variances of the variables, but also on their covariance. Using an estimate of the covariance introduces uncertainty in power. An estimated covariance may also be biased due to distinct populations in the previous and future studies.

I show how to overcome both problems in multivariate linear models, uncertainty and bias in power due to estimated covariance. Two different methods help, the confidence interval for power and an internal pilot design. Exact confidence intervals for noncentrality, power and sample size are known for the univariate model only. With an internal pilot design, data from the first stage of the study are used to re-calculate the sample size, based on the estimate of error variance. All data may be used in the final analysis, with no interim data analysis. A wide variety of exact and approximate results for internal pilot designs are known for univariate models, but not for multivariate models.

For an important special class of multivariate tests (one "between" degree of freedom), I show how power can be computed from an equivalent univariate linear model. Therefore the theory and application of the univariate results for power confidence intervals can be applied

with proper transformation of the problem. A similar approach allows using univariate results for an internal pilot design. Some additional exact results for confidence intervals are provided for another more general collection of models. Finally, approximations which apply to any general multivariate linear model are described and seen to be accurate in simulations.

To my parents,

Mun Hwan Park

and

Su Jin Kim

Biography

Sola Park was born and grew up in Korea. She received her Bachelors degree in Mathematics from Sogang University, Seoul, Korea, in 1998 and Masters degree in 2000. She had been a Ph. D. student in the Department of Biostatistics at the University of North Carolina at Chapel Hill since August 2001. She received her Ph.D. degree in Biostatistics in 2007.

Acknowledgements

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1. Motivating Examples

1.1 Example 1: When Does Clinically Isolated Syndrome (CIS)

Lead to Multiple Sclerosis (MS); Uncertainty in Power

Multiple sclerosis (MS) occurs chiefly in young adults and is thought to be caused by a defect in the immune system that may be of genetic or viral origin. Valerie Jewells, a neuroradiologist in the Radiology Department in the UNC Hospital and her research group, are interested in using Magnetic Resonance Imaging (MRI) to diagnose the onset of MS. A person with Clinically Isolated Syndrome (CIS) has symptoms suspicious for MS, but does not have laboratory or MRI findings consistent with MS, perhaps because it is too early in the disease development. As always, earlier detection leading to earlier treatment is expected to positively affect outcome.

The new approach uses MRI of a patient's brain as a way of assessing disease status. A new MRI protocol, Diffusion Tensor Imaging (DTI), analyzes the movement (diffusion) of water in the brain. Doing so highlights nerve cell pathways. Average Diffusion (AD) is the average variability across all three directions. Fractional Anisotropy (FA), a commonly used DTI summary variable, quantifies heterogeneity of variability of diffusion (and hence heterogeneity of variability). FA is expected to be higher for MS patients in at least some areas affected by the disease. At the present time, the investigator has images from more than 5 known MS cases, 20 CIS patients and 28 disease-free individuals. The scientists wish to know how many patients are sufficient to allow good power for detecting differences among the three groups.

The power of a test, the probability that it will lead to the rejection of the null hypothesis, may be computed using estimates of some distributional parameters, including an

error covariance. Statisticians may use an estimated error covariance from a previous similar study because knowing an exact error covariance is practically impossible. Statisticians at Frank Porter Graham Child Development Institute (M. Gribbin and M. Poe) have analyzed data for a similar study using FA values in young children. The data were collected by the UNC Center for the Study of Autism. The outcome measure, $\widehat{\delta}$, gives an approximate estimated quantile for a histogram of FA values (as defined by Clement, 2005). The observed error variance is used (with permission from the Autism center) in power calculations for the CIS study. *However, the estimation process introduces uncertainty. Furthermore, the error covariance for a group of young children may not be appropriate for typical CIS patients.*

Table 1.1 Two-Sample t Test Power for $\alpha = 0.05$
as a Function of Error Variance, σ^2,
Mean Difference in $\widehat{\delta}$, a measure of diffusion isotropy,
and Sample Size $N_1 = N_2 = N/2$

Mean Difference	Power		
	$N = 10$	$N = 20$	$N = 40$
0.00	0.050	0.050	0.050
0.01	0.055	0.062	0.076
0.02	0.072	0.100	0.158
0.03	0.099	0.165	0.296
0.04	0.138	0.256	0.475
0.05	0.189	0.370	0.659
0.06	0.251	0.497	0.812
0.07	0.323	0.625	0.913
0.08	0.403	0.740	0.966
0.09	0.486	0.833	0.989
0.10	0.570	0.902	0.997

Table 1.1 and Figures 1.1–1.3 illustrate the uncertainty due to the estimated covariance with various sample sizes. The nonnegative variable $\widehat{\delta}$ is interpreted as the quantile of an F random variable corresponding to the mean plus one standard deviation. The value of $\widehat{\delta}$ roughly indicates the point joining the long positive tail to the large bulge of typical F values. Clement (2005) provided a detailed rationale and evaluation of the measure,

including a justification for assuming it is Gaussian. A total of 21 autistic and 9 non-autistic children had a pooled standard deviation of 0.065. The solid line in Figures 1.1–1.3 is the estimated power for a range of possible mean differences. As derived by Taylor and Muller (1995), the dashed lines in Figures 1.1–1.3 are exact 95% confidence intervals on power, and also provide an exact simultaneous confidence region.

A point estimate does not provide any degree of certainty about the estimate. Therefore a confidence interval for the estimate is used to describe how much the estimated parameter is uncertain. Confidence intervals give us a good solution to dealing with uncertainty due to (valid) estimation of an error variance. However, it cannot deal with the bias which can be solved by an internal pilot.

Medical imaging automatically generates repeated measures of many kinds. However, only confidence intervals for power of univariate linear models have been published. Results for more complicated cases involving multivariate theory, especially repeated measures, are needed for planning a variety of medical imaging studies.
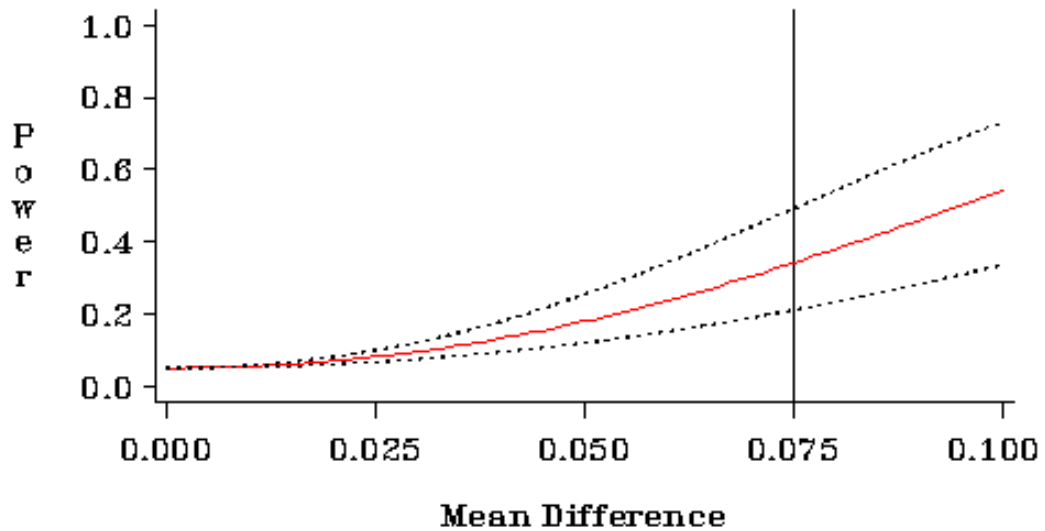
Figure 1.1 Power of mean difference in δ, a measure of diffusion isotropy, with 5 participants in each study design group, $\widehat{\sigma}^2 = 0.065$
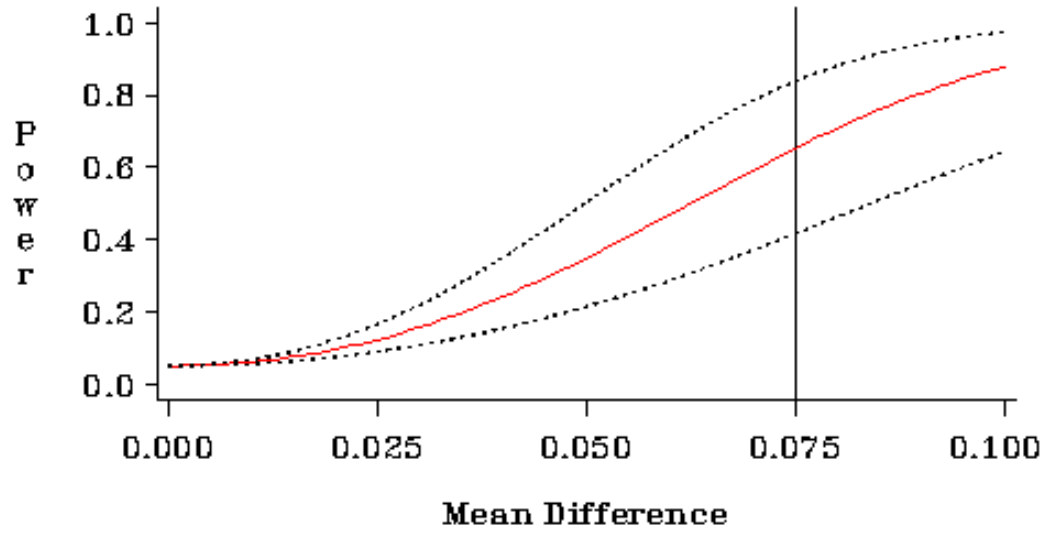
Figure 1.2 Power of mean difference in δ, a measure of diffusion isotropy, with 10 participants in each study de sing group, $\widehat{\sigma}^2 = 0.065$
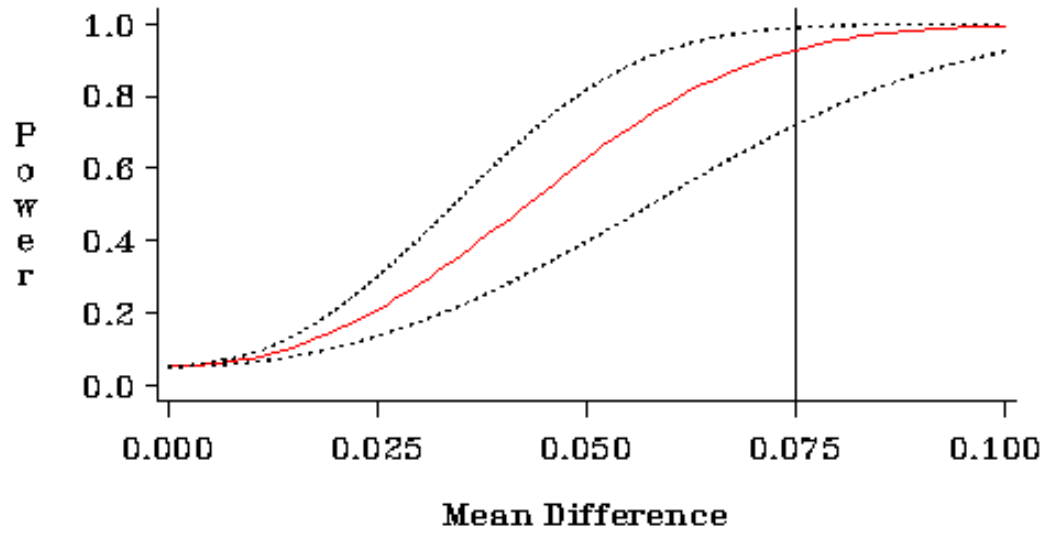
Figure 1.3 Power of mean difference in δ, a measure of diffusion isotropy, with 20 participants in each of two groups, $\widehat{\sigma}^2 = 0.065$, $\alpha = 0.05$,

1.2 Example 2: Does Brain Vessel Tortuosity Vary with Gender and Age?; Bias and Uncertainty in Power

Bullitt et al. (2004a) demonstrated that computer software can measure cerebral vascular tortuosity (bending or twisting rapidly in three dimensions) automatically from MRI data. As in Figure 1, the authors described variation across four different parts of the head: the anterior cerebral, right and left middle cerebral and posterior cerebral circulations. The data supported the assumption of a Gaussian distribution.

Anterior-Posterior   Lateral
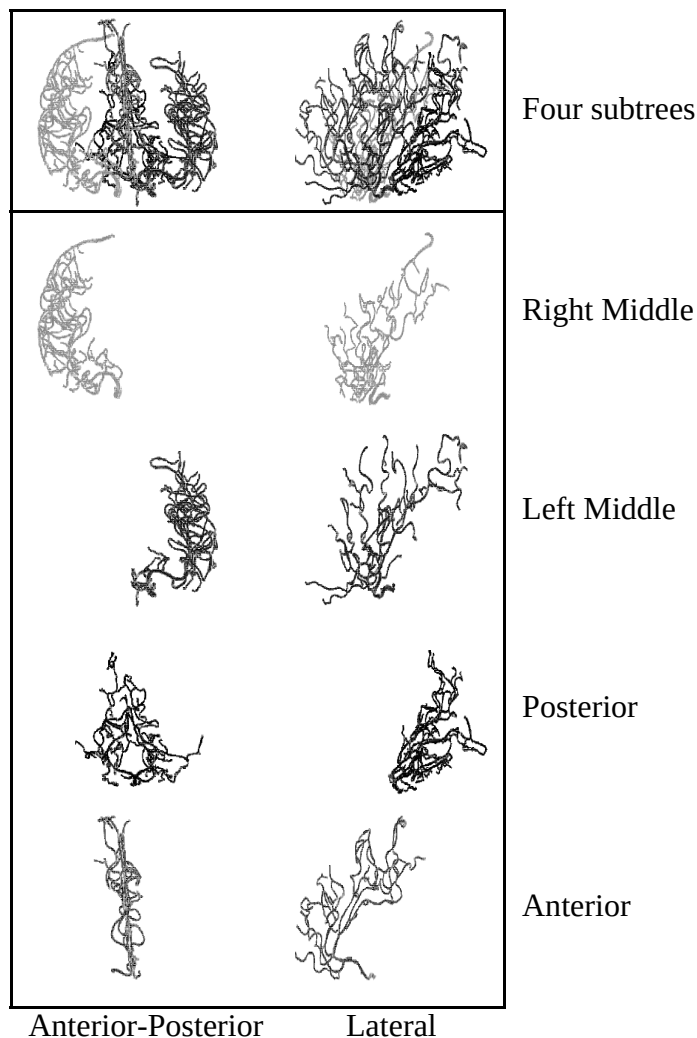
Figure 1.4 Four regions (Right middle, Left middle, Posterior, Anterior)
of cerebral vasculature from two views (Anterior-Posterior, Lateral).

Bullitt et al. (2004b) reviewed a wide range of medical research supporting the principle that blood vessel characteristics often serve as a primary marker of a tumor or disease state. In particular, pathologists use high levels of blood vessel tortuosity (bending or twisting rapidly in three dimensions) to indicate uncontrolled growth and a malignant tumor likely to cause death. The new study will examine the effects of age and gender by recruiting people of both genders across a wide range of ages. Muller, Edwards, Simpson, and Taylor (2007) provided a detailed power analysis for the design. They were motivated to base their power analysis on the univariate approach to repeated measures by the covariance estimate from the previous study appearing to be close to compound symmetric.

Although the covariance matrix was estimated in the previous study, it is not a completely credible value due to the distinct populations in the future study. It seems plausible the estimate may be *biased*. The desire to automatically allow a general covariance structure led to the present research focus on the multivariate approach to repeated measures, in contrast to the work of Muller et al. (2007). In addition to the possibility of bias, estimates are random, and randomness brings uncertainty in small samples (typical for the sample on which the estimate is based, not the target study).

A very large sample size can insure a high likelihood of meeting the study goals. In a clinical trial, it may be impossible in reality because of the limitation of the time and cost of recruiting a large number of participants. Hence, accurate sample size choice seems necessary, with sample size large enough to have good power, but no larger than needed (to control costs and practicality). Typically choosing a sample size depends on knowing nuisance parameters, such as the variance, which may be difficult to specify. A bad choice of nuisance parameters can lead to an underpowered study unlikely to be successful, or an overpowered study, which wastes resources.

Ideally, the scientists would like to adjust the sample size part way though the study, and thereby avoid both problems. Wittes and Brittain (1990) introduced the internal pilot

design, which includes an interim power analysis, without any interim data analysis. Increasing sample size, if needed, avoids an underpowered study. Subsequently, the idea was extended to allow reducing sample size from the original target. The approach has great appeal due to the uncertainty about the appropriateness of the covariance value. Like confidence interval theory, most internal pilot theory has been developed in univariate cases and needs to be extended to repeated measures and multivariate cases.

Chapter 2 Background and Significance

2.1 Notation

2.1.1 Gaussian Multivariate Linear Models

The following notation will be used throughout for convenience. Lower case bold, $\boldsymbol{y}$, indicates a vector and upper case bold, $\boldsymbol{M}$, indicates a matrix. Independent sampling units will be referred to as *participants*. An $n \times 1$ vector $\boldsymbol{x}$ that follows a Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ is denoted by $\boldsymbol{x} \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Also, x having a non-central chi-squared distribution with n degrees of freedom (*df*) and noncentrality ω is indicated $x \sim \chi^2(n, \omega)$. Similarly, x following a non-central F distribution with n_1 numerator *df*, n_2 denominator *df*, and noncentrality ω, is denoted $x \sim F(n_1, n_2, \omega)$. Corresponding cumulative distribution functions are $F_{\chi^2}(n, \omega)$ and $F_F(n_1, n_2, \omega)$, respectively. Also $F_F^{-1}(1 - \alpha; n_1, n_2)$ indicates the $1 - \alpha$ quantile of a central F. In either case, omitting ω indicates a central case with $\omega = 0$.

Notation used in the General Linear Multivariate Model (GLMM) is summarized in Tables 2.1 and 2.0.

Table 2.1
Dimensions

Symbol	Definition
N	Number of independent sampling units
p	Number of response variables (times)
q	Number of predictors and columns in $\boldsymbol{X}$
r	rank($\boldsymbol{X}$)
ν_e	error *df* $= N - r$
a	Number of rows in $\boldsymbol{C}$ = hypothesis *df*
b	Number of columns in $\boldsymbol{U}$, $\boldsymbol{\Sigma}_*$, $\widehat{\boldsymbol{\Sigma}}_*$, $\boldsymbol{S}_h$, $\boldsymbol{S}_e$

Table 2.2
Parameters and Constants

Symbol	Size	Definition and Properties
X	$N \times q$	Fixed, known design matrix
B	$q \times p$	Primary parameters (means)
C	$a \times q$	Between-participant contrasts
U	$p \times b$	Within-participant contrasts
$\Theta = CBU$	$a \times b$	Secondary parameters
Θ_0	$a \times b$	Null values
Σ	$p \times p$	Covariance matrix of $\text{row}_i(E)'$
$\Sigma_* = U'\Sigma U = V\text{Dg}(\lambda)V'$	$b \times b$	Covariance matrix of $\text{row}_i(EU)'$
$M = C(X'X)^-C'$	$a \times a$	Middle matrix
$\Delta = (\Theta - \Theta_0)'M^{-1}(\Theta - \Theta_0)$	$b \times b$	Unscaled noncentrality
$\Omega = \Delta\Sigma_*^{-1}$	$b \times b$	Noncentrality
$\omega = \{\omega_k\}$	$b \times 1$	Eigenvalues of Ω

Here $(N \times p)$ Y is a random matrix of observed responses with independent sampling units, (such as participants), as rows, and multivariate or repeated measures as columns. The model is

$$Y = XB + E, \tag{2.1}$$

with fixed, known design matrix, X, fixed unknown parameter matrix, B, and unobserved errors, E, with independent rows and $\text{row}_i(E)' \sim \mathcal{N}_p(0, \Sigma)$. The usual estimates are $\widetilde{B} = (X'X)^-X'Y$, which is not unique for less than full rank X, and $\widehat{\Sigma} = Y'[I - (X'X)^-X']Y/\nu_e$.

The General Linear Hypothesis (GLH) is

$$H_0 : CBU = \Theta_0, \tag{2.2}$$

for fixed and known Θ_0 ($a \times b$). The C matrix is defined as contrasts between groups or levels of predictors and the U matrix as contrasts within an independent sampling unit, for example, patient, time, etc. The Σ matrix is the covariance matrix among response variables and $\Sigma_* = U'\Sigma U$ is the covariance matrix in the transformed model, $YU = XBU + EU$.

The multivariate hypothesis test statistics can be expressed using noncentrality parameters, Δ (unscaled) and Ω. The rank of Ω, referred to as s_*, plays a key role in theory of multivariate linear models. Only *testable* hypotheses will be considered, which require full rank Σ_*, M, U and $C = C(X'X)^-(X'X)$. The conditions insure $\Theta = CBU$ has a unique and unbiased estimator, and also has a well-defined test (for a fixed sample size).

Generally, we define

$$\widehat{\boldsymbol{\Theta}} = \boldsymbol{C}\widehat{\boldsymbol{B}}\boldsymbol{U}, \tag{2.3}$$

$$\widehat{\boldsymbol{\Sigma}}_* = \boldsymbol{U}'\widehat{\boldsymbol{\Sigma}}\boldsymbol{U} , \tag{2.4}$$

and

$$\widehat{\boldsymbol{\Delta}} = (\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)'\boldsymbol{M}^{-1}(\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0) . \tag{2.5}$$

Here $\boldsymbol{S}_h$ and $\boldsymbol{S}_e$ are independent Wishart matrices with common covariance $\boldsymbol{\Sigma}_*$, and respective degree of freedom a and ν_e:

$$\begin{aligned} \boldsymbol{S}_h &= (\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)'\boldsymbol{M}^{-1}(\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0) = \widehat{\boldsymbol{\Delta}} \\ &\sim \mathcal{W}_b(a, \boldsymbol{\Sigma}_*, \boldsymbol{\Omega}) , \end{aligned} \tag{2.6}$$

and

$$\boldsymbol{S}_e = \boldsymbol{U}'\boldsymbol{\Sigma}\boldsymbol{U} \cdot \nu_e = \widehat{\boldsymbol{\Sigma}}_* \cdot \nu_e \sim \mathcal{W}_b(\nu_e, \boldsymbol{\Sigma}_*) . \tag{2.7}$$

In turn, $\boldsymbol{S}_t = \boldsymbol{S}_h + \boldsymbol{S}_e$ is also Wishart, with

$$\boldsymbol{S}_t \sim \mathcal{W}_b(a + \nu_e, \boldsymbol{\Sigma}_*, \boldsymbol{\Omega}) . \tag{2.8}$$

2.1.2 Internal Pilots

Internal pilot design notation is summarized in Table 2.3.

Table 2.3

Internal Pilot Design Notation for Testing $H_0 : \boldsymbol{\Theta} = \boldsymbol{\Theta}_0$

Symbol	Definition
Dimensions	
a	$\text{rank}(\boldsymbol{C})$
ν_i	error $df = n_i - r$, $i \in \{0, 1, 2, +\}$
Design Parameters	
$\text{Es}(\boldsymbol{X}_+)$	Essence design matrix, of dimension $G \times q$
m_x	Observations taken per replication
α_t	Target test size
P_t	Target probability of interest
$\boldsymbol{\Theta}_*$	'Scientifically important' value of $\boldsymbol{\Theta}$
$\boldsymbol{\Sigma}_{*0}$	Variance value used for planning
n_0	Pre-planned sample size based on $(\alpha_t, P_t, \boldsymbol{\Theta}_*, \boldsymbol{\Sigma}_{*0})$
Sample Size Allocation	
π	Proportion of n_0 used in internal pilot
$n_1 = \pi n_0$	Internal pilot sample size
$n_{+,\min}$	Minimum size of final sample
$n_{+,\max}$	Maximum size of final sample
Fixed, Unknown Parameters	
$\boldsymbol{\Sigma}_*$	Population error variance
$\{\boldsymbol{\Sigma}_*, \boldsymbol{\Sigma}_{*0}\}$	Impact of true vs. planning covariance on $\boldsymbol{\omega} = \{\omega_k\}$
$\boldsymbol{\Theta}$	True value of secondary parameter
Random Variables	
$\widehat{\boldsymbol{\Sigma}}_{*1}$	Internal pilot variance estimate
N_2	Second (random) sample size
$N_+ = n_1 + N_2$	Final (random) sample size
$\widehat{\boldsymbol{\Theta}}_+$	Final estimate of secondary parameter

2.2 Literature Review

2.2.1 Introduction

The Power of a test equals the probability of rejecting the null hypothesis, which measures how a study design is good to test a hypothesis. The power analysis is need to find a sample size. Various method to choose a sample size have been developed until now. In this section, those various method, especially confidence interval for power and internal pilot which had already developed will be introduced.

2.2.2 Multivariate Linear Models

Muller, Lavange, Ramey and Ramey (1992) reviewed the best available approximate and exact power calculations for general linear multivariate models with Gaussian errors. The same methods apply to repeated measures analysis that can be conducted with the multivariate approach to repeated measures.

Throughout, $s = \min(a, b)$ for $\boldsymbol{\Theta} = \boldsymbol{CBU}$ of $a \times b$ dimension. Multivariate test statistics can be expressed as a function of the eigenvalues of $\boldsymbol{S}_h \boldsymbol{S}_e^{-1}$. In the multivariate model, the following four statistic are commonly used: (1) Roy's largest root (RLR), (2) Wilks likelihood ratio statistic (W), (3) Pillai-Bartlett trace (PBT), and (4) Hotelling-Lawley trace (HLT). Table 2.4 summarizes the GLMM test statistics. The univariate approach to repeated measures (UNIREP) statistic is included for comparison.

Table 2.4
GLMM Test Statistic Properties

Name	Statistic	Principle	Association
RLR	$\max \widehat{\rho}_k^2 = \max \text{eval}(\boldsymbol{S}_h \boldsymbol{S}_t^{-1})$	Union-Intersection	RLR
W	$\prod_k (1 - \widehat{\rho}_k^2) = \left\| \boldsymbol{S}_e \boldsymbol{S}_t^{-1} \right\|$	Likelihood Ration	$1 - W^{1/g}$
PBT	$\sum_k \widehat{\rho}_k^2 = \text{tr}(\boldsymbol{S}_h \boldsymbol{S}_t^{-1})$	Total Sqrd Correlation	PBT$/s$
HLT	$\sum_k \widehat{\rho}_k^2 / (1 - \widehat{\rho}_k^2) = \text{tr}(\boldsymbol{S}_h \boldsymbol{S}_e^{-1})$	ANOVA Analog	$(\text{HLT}/s)/(1+\text{HLT}/s)$
UNIREP	$\text{tr}(\boldsymbol{S}_h)/\text{tr}(\boldsymbol{S}_e)$	Sphericity	$(\text{REP})/(1+\text{REP})$

If $s = 1$ (the univariate case), then $\boldsymbol{S}_h \boldsymbol{S}_t^{-1} = \widehat{\rho}^2$ is the scalar ratio of the sum of squares due to the hypothesis and the total sum of squares. Also (if $s = 1$), the four GLMM tests

(but not the UNIREP test, unless $b = 1$) are equivalent in the sense of having the same test size and power. Under the null hypothesis,

$$F_{\text{obs}} = \frac{\widehat{\rho}^2/a}{\left(1 - \widehat{\rho}^2\right)/\widehat{\rho}^2(N - r)} \sim F(a, N - r) \, . \tag{2.9}$$

Under the alternative hypothesis, $F_{\text{obs}} \sim F(a, N - r, \omega)$, with noncentrality $\omega = (\boldsymbol{\theta} - \boldsymbol{\theta}_0)' \boldsymbol{M}^{-1}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)/\sigma^2 = a \cdot F_A$. The parameter F_A is the value of F_{obs} that would occur if $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\sigma^2 = \widehat{\sigma}^2$.

For $s > 1$ (the general multivariate case), η_m indicates the measure of multivariate association for $m \in \{\text{W, PBT, HLT}\}$. Under the null hypothesis,

$$F_{\text{obs}}(m) = \frac{\widehat{\eta}_m/df_1(m)}{(1 - \widehat{\eta}_m)/df_2(m)} \, , \tag{2.10}$$

is approximately an F. With $\nu_e = N - r$, the numerator degrees of freedom $\nu_1(m)$ are $\nu_1(\text{HLT}) = ab$, $\nu_1(\text{WLK}) = ab$, and

$$\nu_1(\text{PBT}) = ab\frac{1}{s(\nu_e + a)}\left[\frac{s(\nu_e + s - b)(\nu_e + a + 2)(\nu_e + a - 1)}{\nu_e(\nu_e + a - b)} - 2\right]. \tag{2.11}$$

Also $df_2(\text{W}) = g[(N - r) - (b - a + a)/2] - (ab - 2)/2$, $df_2(\text{PBT}) = s[(N - r) - b + s]$, $df_2(\text{HLT}) = s[(N - r) - b - 1] + 2$, and $g = [(a^2b^2 - 4)/(a^2 + b^2 - 5)]^{1/2}$. Under the alternative hypothesis, noncentral F approximations are available. In general $\boldsymbol{H}$ follows a noncentral Wishart, with a degrees of freedom and $b \times b$ noncentrality matrix

$$\boldsymbol{\Omega} = (\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)' \boldsymbol{M}^{-1}(\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)\boldsymbol{\Sigma}_*^{-1} \, . \tag{2.12}$$

Following Muller et al. (1992), computing approximate power for the tests of the multivariate general linear hypothesis requires just four steps.

1. Specify $\alpha, \boldsymbol{\Sigma}, \boldsymbol{X}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{U}$, and $\boldsymbol{\Theta}_0$.

2. Find the approximate critical value from an inverse (central) F distribution function, say

$$f_{\text{crit}}(m) \approx F_F^{-1}[1 - \alpha; df_1(m), df_2(m)] \, . \tag{2.13}$$

13

3. Compute the noncentrality in terms of $F_A(m)$. Muller and Peterson (1984) suggested that $F_{obs}(m) \overset{d}{\approx} F[df_1(m), df_2(m), \omega_m]$, with

$$\omega_m = (ab) \cdot F_A(m) = \frac{\eta_m}{(1 - \eta_m)/df_2(m)}. \tag{2.14}$$

The specific noncentralities are computed as

$$\omega_W = (ab) \cdot F_A(W) = \frac{1 - W_A^{1/g}}{W_A^{1/g}/df_2(W)}, \tag{2.15}$$

$$\omega_{PBT} = (ab) \cdot F_A(PBT) = \frac{PBT_A/s}{(1 - PBT_A/s)/df_2(PBT)}, \tag{2.16}$$

or

$$\omega_{HLT} = df_1(m) \cdot F_A(HLT) = \frac{HLT_A/s}{(1 - HLT_A/s)/df_2(HLT)}. \tag{2.17}$$

4. Compute approximate power of the multivariate general linear model as

$$\text{Power}(m) \approx 1 - F_F[F_{crit}(m), ab, df_2(m), ab \cdot F_A(m)], \tag{2.18}$$

where $m \in \{W, PBT, HLT\}$ and $F_F(f, df_1, df_2, \omega)$ represents the noncentral F distribution function, $\Pr\{F \leq f\}$ for a noncentral F statistic based on df_1 numerator df_2 denominator degrees of freedom, and noncentrality parameter ω.

The preceding formulas give exact power for $s = 1$. Furthermore, the powers of the four MULTIREP tests coincide for $s = 1$.

2.2.3 Confidence Intervals for Power Based on Parameter Estimates

The confidence interval for unknown parameter θ of interest is defined with U as the upper interval bound and L as the lower interval bound. The event to be interested in, denoted A, can then be defined as the interval containing the parameter, $L \leq \theta \leq U$. Hence the probability that the event A happened is written as

$$\Pr\{A\} = \Pr\{L \leq \theta \leq U\}. \tag{2.19}$$

14

The interval $[L, U]$ can be interpreted as being a $100 \times (1 - \alpha)\%$ confidence interval for θ with α significance level.

The information collected in a sample from a population is not fully informative about the population. Therefore, we estimate the parameter of interest but it may be uncertain. A confidence interval can quantify the uncertainty of the estimation.

The power of a test is the probability of rejecting a false statistical null hypothesis. A well-designed study should (nearly always) ensure reasonably high power. Hence determining the power is an important step in designing a study.

In linear models, power is a function of sample size, covariance, noncentrality and type I-error. For example, power is higher with large sample size. However, we cannot guarantee the "best" power because of restrictions of time and cost. Moreover, parameter estimates may be needed, which can introduce uncertainty and bias. Hence, many methods have been proposed to get better power and reduce the bias within the given time and cost. The fundamental questions are "What is the size of the population effect of interest?" and "How precisely can we estimate it from out sample?". Creating confidence intervals can account for the uncertainty of estimation.

Gillett (1994) discussed about the average power for sample size estimation. He suggested that the average effect size of all previous related experiments be used as the estimate for the current study in place of a single most recent estimate.

Often, fixed means and an estimated variance from a previous study is used for power. Taylor and Muller (1995) described exact confidence intervals for noncentrality, power and sample size in the univariate linear model with fixed means and estimated variance. In the univariate case, $\boldsymbol{S}_h$ reduces to SSH and $\boldsymbol{S}_e$ reduces to SSE, with $\widehat{\sigma}^2 = \boldsymbol{S}_e/\nu_e$. From Gaussian theory, they provided exact confidence intervals for ω, $[\widehat{\omega}_L, \widehat{\omega}_U]$ with

$$\widehat{\omega}_L = \frac{c_{\text{crit}}(\alpha_{cL}; \nu_e)}{\boldsymbol{S}_e/\nu_e} \cdot \boldsymbol{S}_h \qquad (2.20)$$

and

$$\widehat{\omega}_U = \frac{c_{\text{crit}}(1 - \alpha_{cU}; \nu_e)}{\boldsymbol{S}_e/\nu_e} \cdot \boldsymbol{S}_h \qquad (2.21)$$

The fact that the noncentral F distribution function is strictly monotone in terms of noncentrality (Johnson and Kotz, 1970, p. 193) ensures that an exact confidence interval for power can be derived from an exact confidence interval for ω. The confidence interval for power requires the lower bound $\widehat{P}_L$ and the upper bound $\widehat{P}_U$ to satisfy the equations

$$\widehat{P}_L = 1 - F_F[f_{\text{crit}}(1 - \alpha_t)|\nu_1, \nu_2, \widehat{\omega}_L] \qquad (2.22)$$

and

$$\widehat{P}_U = 1 - F_F[f_{\text{crit}}(1 - \alpha_t)|\nu_1, \nu_2, \widehat{\omega}_U] . \qquad (2.23)$$

The (strictly) monotone function of ω ensures that

$$\Pr\{\widehat{\omega}_L \leq \omega \leq \widehat{\omega}_U\} = \Pr\{\widehat{P}_L \leq \omega \leq \widehat{P}_U\} . \qquad (2.24)$$

Approximations for multivariate cases, denoted by $\widetilde{\omega}_L$ and $\widetilde{\omega}_U$, have the parallel property

$$\Pr\{\widetilde{\omega}_L \leq \omega \leq \widetilde{\omega}_U\} \approx \Pr\{\widetilde{P}_L \leq \omega \leq \widetilde{P}_U\} . \qquad (2.25)$$

The confidence interval becomes too wide if ignoring right truncation and too narrow if ignoring left truncation, as proven by Muller and Pasour (1997).

2.2.4 Invariance in Multivariate Linear Models

Muller and Stewart (2006), discussed some invariance properties in multivariate linear models. A multivariate test is said to *linearly invariant* if the hypothesis test does not vary under full rank transformation of the response variables being tested. More formally, a test of $H_0 : \boldsymbol{\Theta} = \boldsymbol{\Theta}_0$ is linearly invariant to applying a full rank $(b \times b)$ $\boldsymbol{T}$, as in $H_0 : \boldsymbol{\Theta T} = \boldsymbol{\Theta}_0\boldsymbol{T}$. However, the UNIREP tests are not invariant to all full rank transformations (Muller and Barton, 1989). They are only invariant to a full rank orthonormal transformation. All linearly invariant tests, including the four "MULTIREP" tests, are also invariant to an orthonormal transformation. The eigenvalues of

$\widehat{\boldsymbol{\Omega}} = \widehat{\boldsymbol{\Delta}}\widehat{\boldsymbol{\Sigma}}_*^{-1} = \nu_e \boldsymbol{S}_h \boldsymbol{S}_e^{-1}$ are invariant to a full rank transformation of the rows of $\boldsymbol{C}$ (with $\boldsymbol{\Theta}_0$ transformed the same), as are the canonical correlations. Furthermore, all multivariate test statistics and associated p-values, are invariant to full rank transformation of the columns of $\boldsymbol{U}$ (with $\boldsymbol{\Theta}_0$ transformed the same).

2.2.5 Internal Pilots

It is often difficult to find a credible estimate of the nuisance parameter, $\boldsymbol{\Sigma}_*$, before a study has been conducted. Usually, the investigators make a guess, or use a value from a previous similar study, which may be far from the true value, and hence give an under- or over-powered study. However, this method to estimate may be far from the truth because the previous study population is not necessarily homogeneous with the population in the current study.

An internal pilot design avoids the problem by basing sample size on an estimate of the error covariance from the first fraction of observations in the current study. Based on this new and improved variance estimate, the sample size may be changed. In this process, no interim data analysis is performed in the study and data analysis will be performed when the study is completed (Jennison and Turnbull (2000))

Most work in internal pilots has concerned the independent groups t test. An internal pilot design has the important disadvantage that it may inflate the type I error rate. Stein's (1945) two sample approach allows the use of observations from the pilot stage to estimate the variance. Spurrier (1982) described two stage tests in the general linear univariate model. He discussed critical values in four different cases: the sizes of two samples were fixed; the result of the first sample determined whether one would take a second sample; interim testing was allowed; and there is no effect of the first sample on the values of the second sample.

Spurrier (1982) described the two stage tests in the GLUM. He discussed critical values in four different cases: the sizes of two samples were fixed; the result of the first sample

17

determined whether one would take a second sample; interim testing was allowed; and there is no effect of first sample to second sample.

Wittes and Brittan (1990) proposed an internal pilot to estimate the variance for a two group study with a Gaussian outcome. They concluded that the bias in the type I error rate is often negligible, at least in restricted designs and moderate to large sample sizes. Thought Stein's final variance estimate was based only on the observations collected during the internal pilot, Wittes and Brittain's final variance estimate is based on all observations. They investigate the one scenario, which the size of the internal pilot (n_1) is half the originally proposed total sample size (n_0). Birkett and Day (1994) suggested that different values of π may be more appropriate. Using the same methodology as Wittes and Brittain, they showed by simulation that the choice of π is not important. The Type I and II error rates and the expected value of the recomputed sample size resulting from the internal pilot calculations were shown to depend on the absolute size of the internal pilot portion of the study. They also argued that by not allowing a reduction in n_0, the final sample size can be wastefully large when $\widehat{\sigma}_{*1}^2$ is less than σ_{*0}^2. In addition, if σ_*^2 is small relative to σ_{*0}^2, the resulting Power is more than that targeted. Sandvik, Erikssen, Mowinckel and Rødland (1996) also proposed a method to choose the size of internal pilots, based on n_1 being proportional to n. This procedure considers both the size of n_0 and the precision of σ_{*0}^2.

Wittes, Schabenberger, Zucker, Brittain and Proschan (1999) proposed a computational method to compute the exact distribution of the test statistic for a two group t test and an internal pilot. Furthermore, they described methods for determining critical values. Gould and Shih (1992) suggested a different method to prevent test size inflation due to interim power analysis. They argued against the internal pilot because it requires unmaking of treatment status at the interim estimation.

Coffey and Muller (1999, 2000a, 2000b, 2001) described exact test size and power for *any* Gaussian error linear model for an internal pilot study. Their results indicate that in

small samples test size can be inflated, especially for designs which allow sample size reduction. Furthermore, the best choice of test (which corresponds primarily to the method for estimating variance) changes with the design features of interest. They recommended a "bounding" test as having the best combination of properties.

Coffey and Muller (2003) extended some of their previous internal pilot results to multivariate linear models using UNIREP analysis and test statistics, and a limited set of designs. Work in progress will extend the results to any possible design (but with the UNIREP tests only). In a different approach, Denne and Zucker (2002) developed approximate methods for two-stage procedures with a special class of the general mixed linear model, allowing for dropouts and missed visits.

2.3 Statement of the Problems To Be Solved

Problem 1. Accurate power analysis is important to determine sample size and an estimate of error covariance is needed to do a power analysis. Uncertainty about the error covariance disturbs accurate power analysis. Hence, it is necessary to recognize how much estimates are uncertain. I propose to develop a mix of exact and approximate results to extend the known results for the univariate linear model to the three MULTIREP tests most often used (Wilks, Pillai-Bartlett, Hotelling-Lawley). There is no UMP-α test in the multivariate linear model, which requires studying all three to allow choosing the best for each situation.

Problem 2. I also propose to provide a mix of exact and approximate results in order to extend internal pilot designs to the GLMM using the (three common) MULTIREP tests. In both problems, exact theory should be available for one and two sample multivariate and repeated measures designs. More general theory will likely need approximations, as for fixed sample power.

Chapter 3 (Paper 1) The Confidence Interval Due to Estimated Covariance for Power of a One or Two Group Test, and Related Models

3.1 Introduction

In a clinical trial, a power analysis is done mainly to help choose a sample size. For a linear model with Gaussian errors, the greatest difficulty usually centers on finding an appropriate value for the error variance in the population. Often an estimate from a previous study is used. In turn, the power becomes an estimate, a random value with uncertainty surrounding it. Taylor and Muller (1995, 1996) described how to create confidence bounds for the estimated power and sample size of the univariate linear model with Gaussian errors.

The many parameters in the error covariance matrix for multivariate or repeated measures models complicate the task of accounting for uncertainty in power due to estimating error variance. We extend the Taylor and Muller (1995, 1996) results to multivariate and repeated measures designs involving one or two groups, as in typical clinical trials. The approach involves transforming the multivariate model to an equivalent univariate model, which allows applying the exact results of Taylor and Muller. Although we provide free software to implement the method, the transformation does specify inputs which can be used with commercial or other univariate software.

3.2 An Equivalent Univariate Model for $s = 1$

All notation used in this chapter is defined in Chapter 2. For $s = 1$, we describe how to convert any multivariate or repeated measures model to an equivalent univariate model. A series of transformations is the basic tool for converting a multivariate linear model and associate hypothesis to an equivalent univariate linear model and associated hypothesis.

Definition. A multivariate general linear model and associated hypothesis, say $Y_1 = X_1 B_1 + E_1$ and $H_{01} : C_1 B_1 U_1 = \Theta_{01}$, are said to be hypothesis equivalent to $Y_2 = X_2 B_2 + E_2$ and $H_{02} : C_2 B_2 U_2 = \Theta_{02}$ if (and only if) the test size and power function of the two tests coincide.

The definition implicitly refers to a variety of invariance and equivalence relations in the parameter space and the sample space. The first type of operations described here arise from full rank linear transformations of contrast and model matrices. Such operations retain the original dimensions and ranks, and have received some attention in the past. Less attention has been paid to hypothesis equivalent transformations that reduce (or enlarge) rank and dimensions. Both types of operations help clarify and simplify analytic properties and can improve computational properties. As used here, the second type allows simplifying the theory so much that a hypothesis with complex or unknown theory can be recognized as equivalent to a hypothesis with simple and known properties.

Lemma 3.1 In general linear multivariate model $Y_A = X_A B_A + E_A$, any testable General Linear Hypothesis (GLH) $H_{0A} : C_A B_A U_A = \Theta_{0A}$ with $U_A \neq I$ and $\Theta_{0,A} \neq 0$ may be expressed in terms of a hypothesis equivalent GLH, $H_{0B} : C_B B_B U_B = \Theta_{0B}$, with $C_B = C_A$, $U_B = I$, and $\Theta_{0,B} = 0$.

Proof. If $C_0 = C_A'(C_A C_A')^{-1}\Theta_{A0}$, then transforming the model gives

$$
\begin{aligned}
Y_A U_A &= X_A B_A U_A &&+ E_A U_A & (3.1)\\
Y_A U_A - X_A C_0 &= X_A(B_A U_A - C_0) &&+ E_A U_A \\
Y_B &= X_B B_B &&+ E_B.
\end{aligned}
$$

Hence $\Theta_B = C_B B_B = \Theta_A - \Theta_{A0}$ and $H_0 : \Theta_B = 0$ is equivalent to H_0: $C_A B_A U_A = \Theta_{0,A}$. $\square$

Lemma 3.2 Any testable GLH in a general linear multivariate model may be transformed to a hypothesis equivalent model and GLH with $C_C = [\,I_a \quad 0\,]$.

Proof. For any $a \times q$ matrix $\boldsymbol{C}_B$ of rank a $(a \le q)$, there exist orthogonal $a \times a$ and $q \times q$ matrices $\boldsymbol{L}_B$ and $\boldsymbol{R}_B$, such that

$$\boldsymbol{C}_B = \boldsymbol{L}_B \boldsymbol{D}_B \boldsymbol{R}'_B = \boldsymbol{L}_B [\text{Dg}(\boldsymbol{\lambda}_B) \quad \boldsymbol{0}] \boldsymbol{R}'_B = \boldsymbol{L}_B \text{Dg}(\boldsymbol{\lambda}_B) \boldsymbol{R}'_{B+}. \tag{3.2}$$

Using a singular value decomposition, here $\boldsymbol{R}_B = [\boldsymbol{R}_{B+} \quad \boldsymbol{R}_{B0}]$ and $\text{Dg}(\boldsymbol{\lambda}_B)$ is an $a \times a$ diagonal matrix with positive diagonal elements of $\text{Dg}^2(\boldsymbol{\lambda}_B)$ the positive eigenvalues of $\boldsymbol{C}'_B \boldsymbol{C}_B$ and $\boldsymbol{C}_B \boldsymbol{C}'_B$. Also, $\boldsymbol{R}'_{B0}$ spans the null space of the rows of $\boldsymbol{C}_B$ and $\boldsymbol{T} = \begin{bmatrix} \boldsymbol{C}_B \\ \boldsymbol{R}'_{B0} \end{bmatrix}$.

Choosing $\boldsymbol{R}_{B0}$ as the orthonormal eigenvectors corresponding to zero eigenvalues of $\boldsymbol{C}'_B \boldsymbol{C}_B$ provides one convenient choice. In turn

$$\boldsymbol{T}^{-1} = [\boldsymbol{R}_{B+} \text{Dg}^{-1}(\boldsymbol{\lambda}_B) \boldsymbol{L}'_B \quad \boldsymbol{R}_{B0}] = [\boldsymbol{C}'_B (\boldsymbol{C}_B \boldsymbol{C}'_B)^{-1} \quad \boldsymbol{R}_{B0}]. \tag{3.3}$$

Using the expressions just defined in the model resulting from Lemma 3.1 gives

$$\boldsymbol{Y}_B = \boldsymbol{X}_B \boldsymbol{B}_B + \boldsymbol{E}_B \tag{3.4}$$
$$\Leftrightarrow \quad \boldsymbol{Y}_B = \boldsymbol{X}_B \boldsymbol{T}^{-1} \boldsymbol{T} \boldsymbol{B}_B + \boldsymbol{E}_B$$
$$\Leftrightarrow \quad \boldsymbol{Y}_B = [\boldsymbol{X}_B \boldsymbol{C}'_B (\boldsymbol{C}_B \boldsymbol{C}'_B)^{-1} \quad \boldsymbol{X}_B \boldsymbol{R}_{B0}] \begin{bmatrix} \boldsymbol{C}_B \boldsymbol{B}_B \\ \boldsymbol{R}'_{B0} \boldsymbol{B}_B \end{bmatrix} + \boldsymbol{E}_B$$
$$\Leftrightarrow \quad \boldsymbol{Y}_B = [\boldsymbol{X}_B \boldsymbol{C}'_B (\boldsymbol{C}_B \boldsymbol{C}'_B)^{-1} \quad \boldsymbol{X}_B \boldsymbol{R}_{B0}] \begin{bmatrix} \boldsymbol{\Theta}_{B1} \\ \boldsymbol{\Theta}_{B\perp} \end{bmatrix} + \boldsymbol{E}_B$$
$$\Leftrightarrow \quad \boldsymbol{Y}_C = [\boldsymbol{X}_{C1} \quad \boldsymbol{X}_{C2}] \boldsymbol{B}_C + \boldsymbol{E}_C = \boldsymbol{X}_C \boldsymbol{B}_C + \boldsymbol{E}_C.$$

Here $\boldsymbol{X}_{C1} = \boldsymbol{X}_B \boldsymbol{C}'_B (\boldsymbol{C}_B \boldsymbol{C}'_B)^{-1}$, $\boldsymbol{X}_{C2} = \boldsymbol{X}_B \boldsymbol{R}_{B0}$, $\boldsymbol{B}_2 = \begin{bmatrix} \boldsymbol{\Theta}_1 \\ \boldsymbol{\Theta}_\perp \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_B \boldsymbol{B}_B \\ \boldsymbol{R}'_{B0} \boldsymbol{B}_B \end{bmatrix}$, and $\boldsymbol{\Theta}_{B1}$ and $\boldsymbol{\Theta}_{B\perp}$ are orthogonal. Hence any testable GLH in $\boldsymbol{Y}_B = \boldsymbol{X}_B \boldsymbol{B}_B + \boldsymbol{E}_B$ with any known constant $\boldsymbol{C}_B$ and $\boldsymbol{U}_B$, is transformed to the equivalent GLH for $\boldsymbol{Y}_C = \boldsymbol{X}_C \boldsymbol{B}_C + \boldsymbol{E}_C$ with $\boldsymbol{C}_C = [\boldsymbol{I}_a \quad \boldsymbol{0}]$ and $\boldsymbol{B}_C = \boldsymbol{T} \boldsymbol{B}_B$. Hence $H_0 : \boldsymbol{C}_B \boldsymbol{B}_B \boldsymbol{U}_B = \boldsymbol{\Theta}_{0,B}$ is equivalent to $H_0 : \boldsymbol{C}_C \boldsymbol{B}_C \boldsymbol{U}_C = \boldsymbol{\Theta}_{0,C}$. $\square$

Lemma 3.3 In a general linear multivariate model any testable GLH with between subject contrast matrix of dimension and rank a may be expressed in terms of a hypothesis equivalent model $\boldsymbol{Y}_D = \boldsymbol{X}_D \boldsymbol{B}_D + \boldsymbol{E}_D$ and GLH with $a \le q$ columns in $\boldsymbol{X}_D$ and $\boldsymbol{C}_D = \boldsymbol{I}_a$. The N rows of the original model become $[N - \text{rank}(\boldsymbol{X}_C) + a]$ rows in $\boldsymbol{X}_D$ and $\boldsymbol{Y}_D$.

Proof. Lemma 3.1 gives $\boldsymbol{Y}_B = \boldsymbol{X}_B\boldsymbol{B}_B + \boldsymbol{E}_B$. For any testable GLH with any known constant $\boldsymbol{C}_B$ and $\boldsymbol{U}_B$, Lemma 3.2 gives the hypothesis equivalent model $\boldsymbol{Y}_C = \boldsymbol{X}_C\boldsymbol{B}_C + \boldsymbol{E}_C$ with $\boldsymbol{C}_C = [\,\boldsymbol{I}_a \quad \boldsymbol{0}\,]$ and $\boldsymbol{B}_C = \boldsymbol{T}\boldsymbol{B}_B$ for $\boldsymbol{T} = [\,\boldsymbol{C}_B' \quad \boldsymbol{R}_{B0}\,]$ with $\boldsymbol{R}_{B0}'$ the orthogonal space of $\boldsymbol{C}_B$. The fixed constant matrix $\boldsymbol{X}$ can be expressed as $[\,\boldsymbol{X}_{C1} \quad \boldsymbol{X}_{C2}\,] = [\,\boldsymbol{X}_B\boldsymbol{C}_B'(\boldsymbol{C}_B\boldsymbol{C}_B')^{-1} \quad \boldsymbol{X}_B\boldsymbol{R}_{B0}\,]$ with rank$(\boldsymbol{X}_B) = r$, while rank$(\boldsymbol{X}_{C1}) = a$ and rank$(\boldsymbol{X}_{C2}) = r - a$. The $N \times (q - a)$ matrix $\boldsymbol{X}_{C2}$ has a singular value decomposition, $\boldsymbol{X}_{C2} = \boldsymbol{L}_C \begin{bmatrix} \mathrm{Dg}(\boldsymbol{\lambda}_{C2}) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \boldsymbol{R}_C'$, with $\boldsymbol{L}_C$ an orthonormal $N \times N$ matrix and $\boldsymbol{R}_C$ an orthonormal $(q - a) \times (q - a)$ matrix. There always exist matrices $\boldsymbol{L}_{C1}$, $\boldsymbol{L}_{C0}$, $\boldsymbol{R}_{C1}'$ and $\boldsymbol{R}_{C0}'$, and $N \times (r - a)$, $N \times (N - r + a)$, $(r - a) \times (q - a)$ and $(q - r) \times (q - a)$ dimensions respectively, such that

$$
\begin{aligned}
\boldsymbol{X}_{C2} &= \boldsymbol{X}_B\boldsymbol{R}_{B0} \\
&= [\,\boldsymbol{L}_{C1} \quad \boldsymbol{L}_{C0}\,] \begin{bmatrix} \mathrm{Dg}(\boldsymbol{\lambda}_{C2}) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{R}_{C1}' \\ \boldsymbol{R}_{C0}' \end{bmatrix} \\
&= \boldsymbol{L}_{C1}\mathrm{Dg}(\boldsymbol{\lambda}_{C2})\boldsymbol{R}_{C1}' \, .
\end{aligned} \tag{3.5}
$$

Therefore

$$
\begin{aligned}
\boldsymbol{Y}_C &= \boldsymbol{X}_C\boldsymbol{B}_C + \boldsymbol{E}_C \\
\Leftrightarrow \quad \boldsymbol{L}_{C0}'\boldsymbol{Y}_C &= \boldsymbol{L}_{C0}'\boldsymbol{X}_C\boldsymbol{B}_C + \boldsymbol{L}_{C0}'\boldsymbol{E}_C \\
\Leftrightarrow \quad \boldsymbol{L}_{C0}'\boldsymbol{Y}_C &= \boldsymbol{L}_{C0}'\boldsymbol{X}_{C1}\boldsymbol{C}_C\boldsymbol{B}_C + \boldsymbol{L}_{C0}'\boldsymbol{E}_C \\
\Leftrightarrow \quad \boldsymbol{Y}_D &= \boldsymbol{X}_D\boldsymbol{B}_D + \boldsymbol{E}_D \, . \\
[(N - r + a) \times b] &= [(N - r + a) \times a](a \times b) + [(N - r + a) \times b]
\end{aligned} \tag{3.6}
$$

with $\boldsymbol{Y}_D = \boldsymbol{L}_{C0}'\boldsymbol{Y}_C$, $\boldsymbol{X}_D = \boldsymbol{L}_{C0}'\boldsymbol{X}_{C1}$, $\boldsymbol{B}_D = \boldsymbol{C}_C\boldsymbol{B}_C$ and $\boldsymbol{E}_D = \boldsymbol{L}_{C0}'\boldsymbol{E}_C$. Hence any testable GLH for $\boldsymbol{Y}_C = \boldsymbol{X}_C\boldsymbol{B}_C + \boldsymbol{E}_C$ with any known constant $a \times q$ matrix $\boldsymbol{C}_C$ and $p \times b$ matrix $\boldsymbol{U}_C$, is transformed to an equivalent GLH for $\boldsymbol{Y}_D = \boldsymbol{X}_D\boldsymbol{B}_D + \boldsymbol{E}_D$ with a columns in $\boldsymbol{Y}_D$, $\boldsymbol{C}_D = \boldsymbol{I}_a$ and $\boldsymbol{B}_D = \boldsymbol{C}_C\boldsymbol{B}_C$. That is, $H_0 : \boldsymbol{C}_C\boldsymbol{B}_C\boldsymbol{U}_C = \boldsymbol{\Theta}_{0,C}$ is equivalent to $H_0 : \boldsymbol{C}_D\boldsymbol{B}_D\boldsymbol{U}_D = \boldsymbol{\Theta}_{0,D}$. $\qquad\square$

The following theorem contains one of the key principles underlying many of the results in the present work. The theorem includes an explicit construction that applies to any multivariate model. It may be summarized as saying that if $s = \min(a, b) = 1$, then a

23

hypothesis equivalent *univariate* model always exists. By definition, it has the same test size and power function as the original model. The theorem has many direct applications in the derivation of multivariate linear model properties. Such applications are ignored here. In the present research the theorem give access to exact results in confidence intervals for multivariate model power based on an estimated covariance and also internal pilot designs for multivariate models.

Theorem 3.1 Multivariate general linear model $Y_A = X_A B_A + E_A$ has fixed $N \times q$ X_A of rank r, $N \times p$ Y_A and $\boldsymbol{E}_A$ with independent rows, and $\text{row}_i(\boldsymbol{E}_A) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_A)$. A corresponding testable general linear hypothesis about $a \times b$ $\boldsymbol{\Theta}_A = \boldsymbol{C}_A \boldsymbol{B}_A \boldsymbol{U}_A$ is stated $H_0 : \boldsymbol{\Theta}_A = \boldsymbol{\Theta}_{0,A}$. **a)** A hypothesis equivalent model, $\boldsymbol{Y}_E = \boldsymbol{X}_E \boldsymbol{B}_E + \boldsymbol{E}_E$ with $(N - r + a)$ rows, always exists with $\boldsymbol{C}_E = \boldsymbol{I}_a$, $\boldsymbol{U}_A = \boldsymbol{I}_b$, $\boldsymbol{\Theta}_{0,E} = \boldsymbol{0}$ so that testing $H_0 : \boldsymbol{B}_E = \boldsymbol{0}$ is equivalent to testing $H_0 : \boldsymbol{\Theta}_A = \boldsymbol{\Theta}_{0,A}$. With $\boldsymbol{R}_0$ the $q \times (r - a)$ orthonormal eigenvectors corresponding to zero eigenvalues of $\boldsymbol{C}_A' \boldsymbol{C}_A$, the $N \times (r - a)$ matrix $\boldsymbol{X}_{A0} = \boldsymbol{X}_A \boldsymbol{R}_0$ has rank $r - a$. Singular value decomposition gives $\boldsymbol{X}_{A0} = [\boldsymbol{T}_+ \quad \boldsymbol{T}_r] \begin{bmatrix} \text{Dg}(\boldsymbol{\lambda}_{A0}) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_+' \\ \boldsymbol{V}_0' \end{bmatrix}$ with $N \times (r - a)$ matrix $\boldsymbol{T}_+$, and $N \times (N - r + a)$ matrix $\boldsymbol{T}_r$. The $N \times N$ matrix $\boldsymbol{T} = [\boldsymbol{T}_+ \quad \boldsymbol{T}_r]$ is such that $\boldsymbol{T}'\boldsymbol{T} = \boldsymbol{I}_N$. Also $\boldsymbol{T}_c = \boldsymbol{C}_A'(\boldsymbol{C}_A \boldsymbol{C}_A')^{-1}$ and

$$\boldsymbol{Y}_E = \boldsymbol{T}_r' \boldsymbol{Y}_A \boldsymbol{U}_A - \boldsymbol{T}_r \boldsymbol{X}_A \boldsymbol{T}_c^{-1} \boldsymbol{\Theta}_{0,A} \tag{3.7}$$
$$\boldsymbol{X}_E = \boldsymbol{T}_r' \boldsymbol{X}_A \boldsymbol{T}_c \tag{3.8}$$
$$\boldsymbol{B}_E = \boldsymbol{\Theta}_A - \boldsymbol{\Theta}_{0,A} \tag{3.9}$$
$$\boldsymbol{E}_E = \boldsymbol{T}_r' \boldsymbol{E}_A \boldsymbol{U}_A \tag{3.10}$$
$$\boldsymbol{\Sigma}_E = \boldsymbol{U}_A' \boldsymbol{\Sigma}_A \boldsymbol{U}_A . \tag{3.11}$$

Proof. The theorem is proven by combining Lemmas 3.1, 3.2, and 3.3. Testability insures $a \times q$ $\boldsymbol{C}_A$ has $\text{rank}(\boldsymbol{C}_A) = a$ and orthogonal complement $\boldsymbol{R}_0'$ has rank $q - a$, which insure $\boldsymbol{X}_r = \boldsymbol{X}_A \boldsymbol{R}_0$ exists of rank $(r - a)$. $\square$

Corollary 3.1 If $b = 1$, then $\boldsymbol{Y}_E = \boldsymbol{X}_E \boldsymbol{B}_E + \boldsymbol{E}_E$ is a *univariate* model, which may be written $\boldsymbol{y}_E = \boldsymbol{X}_E \boldsymbol{\beta}_E + \boldsymbol{e}_E$ for clarity.

Corollary 3.2 If $a = 1$, then a hypothesis equivalent *univariate* model always exists.

If $\omega_E = \text{tr}(\boldsymbol{\Omega}_E) = (\boldsymbol{x}'_{E1}\boldsymbol{x}_{E1})\text{tr}(\boldsymbol{B}_E\boldsymbol{\Sigma}_E^{-1}\boldsymbol{B}'_E)$, then an equivalent univariate model is

$\boldsymbol{y}_F = \boldsymbol{X}_F\boldsymbol{\beta}_F + \boldsymbol{e}_F$, with $n_F = (N - r + a)$ rows, $\boldsymbol{e}_F \sim \mathcal{N}_{n_F}(\boldsymbol{0}, \boldsymbol{I})$, $\boldsymbol{X}'_F = \begin{bmatrix} \boldsymbol{I}_b & \boldsymbol{0} \end{bmatrix}'$, $b \times 1$

$\boldsymbol{\beta}_F = \begin{bmatrix} \sqrt{\omega_E} & \boldsymbol{0} \end{bmatrix}'$, $\boldsymbol{C}_F = \boldsymbol{I}_b$, $U_F = \{1\}$, and $\boldsymbol{\theta}_{0F} = \boldsymbol{0}$.

Proof. Model $\boldsymbol{Y}_E = \boldsymbol{X}_E\boldsymbol{B}_E + \boldsymbol{E}_E$ has $(N - r + a)$ rows, $\boldsymbol{Y}_E$ has b columns, and $\boldsymbol{X}_E$ has

$a = 1$ columns, $\boldsymbol{x}_{E1}$, $\boldsymbol{C}_E = 1$, and $\boldsymbol{U}_E = \boldsymbol{I}_b$. Testing $\boldsymbol{B}_E = \boldsymbol{0}$ gives $\boldsymbol{\Delta}_E =$

$\boldsymbol{B}'_E[1(\boldsymbol{x}'_{E1}\boldsymbol{x}_{E1})^{-1}1']^{-1}\boldsymbol{B}_E = (\boldsymbol{x}'_{E1}\boldsymbol{x}_{E1})\boldsymbol{B}'_E\boldsymbol{B}_E$. With $1 \times b$ matrix $\boldsymbol{B}_E$ necessarily of rank 1

under the alternative and rank 0 under the null, $\boldsymbol{B}'_E\boldsymbol{B}_E$ has the same rank, as does $b \times b$

$\boldsymbol{\Omega}_E = (\boldsymbol{x}'_{E1}\boldsymbol{x}_{E1})\boldsymbol{B}'_E\boldsymbol{B}_E\boldsymbol{\Sigma}_E^{-1}$. Linear model results (Muller and Stewart, 2006, section 16.8,

Corollary 16.11) and properties of the trace ensure $\boldsymbol{\Omega}_E$ and $(\boldsymbol{x}'_{E1}\boldsymbol{x}_{E1})\boldsymbol{B}_E\boldsymbol{\Sigma}_E^{-1}\boldsymbol{B}'_E$ have the

same eigenvalues, with the only one that is possibly nonzero given by the scalar $\omega_E =$

$\text{tr}(\boldsymbol{\Omega}_E) = (\boldsymbol{x}'_{E1}\boldsymbol{x}_{E1})\text{tr}(\boldsymbol{B}_E\boldsymbol{\Sigma}_E^{-1}\boldsymbol{B}'_E)$.

Model $\boldsymbol{y}_F = \boldsymbol{X}_F\boldsymbol{\beta}_F + \boldsymbol{e}_F$ has $(N - r + a)$ rows, with $\boldsymbol{e}_F \sim \mathcal{N}_{N-r+a}(\boldsymbol{0}, \boldsymbol{I})$, $\boldsymbol{X}'_F = \begin{bmatrix} \boldsymbol{I}_b & \boldsymbol{0} \end{bmatrix}'$ and $b \times 1$ $\boldsymbol{\beta}_F = \begin{bmatrix} \omega_E^{1/2} & \boldsymbol{0} \end{bmatrix}'$, $\boldsymbol{C}_F = \boldsymbol{I}_b$ and $\boldsymbol{\theta}_{0F} = \boldsymbol{0}$. Hence $\boldsymbol{X}'_F\boldsymbol{X}_F = \boldsymbol{I}_b = (\boldsymbol{X}'_F\boldsymbol{X}_F)^{-1}$ and $\omega_F = \boldsymbol{\Delta}_F\boldsymbol{\Sigma}_{*F}^{-1} = \boldsymbol{\beta}'_F(\boldsymbol{I}_b\boldsymbol{I}_b\boldsymbol{I}'_b)^{-1}\boldsymbol{\beta}_F 1^{-1} = \omega_E$. The degrees

of freedom are $\{b, N - r + a\}$. $\square$

Although robustness issues are not discussed in detail, many results for Gaussian data also

will apply to data which is not Gaussian but otherwise meet the assumptions, as least in large

samples. The transformation approach uses only linear transformations, which allows a form

of the central limit theorem to operate with sufficient sample size.

3.3 An Exact Confidence Interval of Power with $\widehat{\boldsymbol{\Sigma}}$ for $s = 1$

Muller et al. (1992) mentioned that for $s = 1$ the exact noncentral F random variables

are the same for all multivariate statistics and give us a UMP-α test. However, for $s > 1$

there is no UMP-α test. Taylor and Muller (1995, 1996) and Muller and Pasour (1997)

derived exact confidence intervals for power due to $\widehat{\sigma}^2$ *in univariate* linear models. They

derived the confidence interval for power from the confidence interval for noncentrality, ω.

They observed that

$$\Pr\{c_{\text{crit}}(\alpha_{cL}|\nu_{2e}) < \frac{SSE}{\sigma^2} < c_{\text{crit}}(1 - \alpha_{cU}|\nu_{2e})\} = 1 - \alpha_{cL} - \alpha_{cU} \tag{3.12}$$

and that the exact confidence interval for ω is provided by

$$\widehat{\omega}_L = \frac{c_{\text{crit}}(\alpha_{cL}|\nu_{2e})}{SSE} \cdot SSH(\boldsymbol{\theta}, N_t) \tag{3.13}$$

and

$$\widehat{\omega}_U = \frac{c_{\text{crit}}(1 - \alpha_{cU}|\nu_{2e})}{SSE} \cdot SSH(\boldsymbol{\theta}, N_t) . \tag{3.14}$$

Therefore, the lower ($\widehat{P}_L$) and upper ($\widehat{P}_U$) bounds for power may be computed directly from the interval for noncentrality because the power function increases strictly as a function of noncentrality:

$$\Pr\{\widehat{\omega}_L \le \omega \le \widehat{\omega}_U\} = \Pr\{\widehat{P}_L \le \omega \le \widehat{P}_U\} . \tag{3.15}$$

Also

$$\widehat{P}_L = 1 - F_F[f_{\text{crit}}(1 - \alpha_t)|\nu_1, \nu_2, \widehat{\omega}_L] \tag{3.16}$$

and

$$\widehat{P}_U = 1 - F_F[f_{\text{crit}}(1 - \alpha_t)|\nu_1, \nu_2, \widehat{\omega}_U] . \tag{3.17}$$

A confidence interval for power which accounts for the uncertainty induced by using an estimated variance or covariance matrix helps the investigator find the appropriate sample size. Figure 3.1 illustrates the idea in the univariate model. In the next section the same type of figure is provided after extending the result to the multivariate model for $s = 1$.
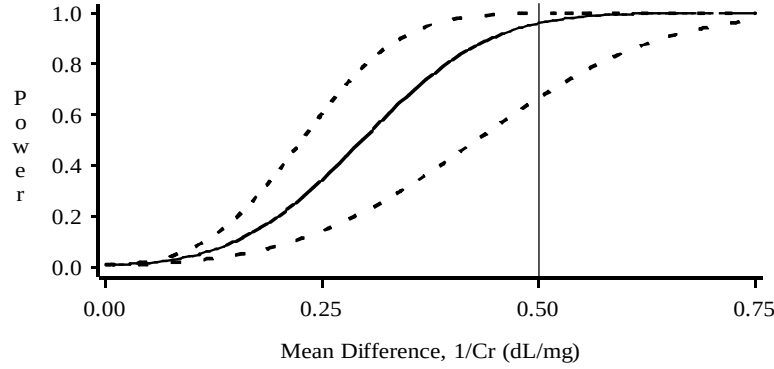
Figure 3.1 Two-sided 95% confidence region due to $\widehat{\sigma}^2$

The noncentrality is invariant under the transformations described in the lemmas and theorem. Hence the numerical size of a confidence region for noncentrality will not be changed even though the transformations are applied and the model looks different. For $s = 1$ the multivariate general linear model can be expressed as a univariate linear model. Therefore $s = 1$ allows applying the exact confidence interval for power of the univariate general linear model from Taylor and Muller (1996). The result is formalized in the following corollary to Theorem 3.1.

Corollary 3.3 If $s = 1$, then an exact confidence interval for MULTIREP noncentrality or power due to an estimated covariance can be expressed as a function of a scalar parameter.

In all cases, the original multivariate model has a $b \times b$ noncentrality matrix. At first glance it therefore seems incorrect to describe a confidence *interval*, rather than confidence region, if $b > 1$. The problem is resolved by recognizing that Theorem 3.1 ensures if $b > 1$ and $s = 1$ because $a = 1$ then the $b \times b$ noncentrality matrix $\mathbf{\Omega}$ has rank 1 and the scalar tr$(\mathbf{\Omega})$ suffices to describe the noncentral distribution. The confidence interval in the corollary refers to tr$(\mathbf{\Omega})$, not $\mathbf{\Omega}$.

3.4 Tortuosity Study Example: Two Between Factors, One Within

3.4.1 Using a Multivariate Model

Tortuosity, bending, twisting, or winding of a vessel in the brain can be measured automatically from magnetic resonance imaging (MRI). It is hoped that many diseases may

be diagnosed by examining the feature of tortuosity as an indicator of vessel abnormality. Bullitt, et al. (2004b) studied vessel abnormality in the brain. The assumption of a Gaussian distribution was supported by the data. A new study is desired to examine the effects of age and gender across a wide range of ages. A power analysis for the new study should be done using multivariate power analysis. The purpose of the next study is to create a pool of normal brains for subsequent studies. Even though tortuosity seems likely to vary across age and gender, the nature of the differences is unknown. Therefore the study is being designed to have good power for the most complex hypothesis of concern, namely Gender $\times$ Region.

Power analysis was done for a variable, SOAM1, to be computed separately in the four regions of the brain identified by neurosurgeons. The name SOAM1 indicates the Sum of all positive Angles between successive trios of equally spaced vessel points, divided by total path length (radians/cm), for all vessels in a region. In the multivariate model for the study, $\boldsymbol{Y}$ has four columns of region of brain, Anterior, Left Middle, Posterior and Right Middle (Ant, LMid, Post, RMid) with N rows (participants). A design matrix is composed of ten columns for Gender $\times$ Age group (20-30, 30-40, 40-50, 50-60, 60+ years of age). The balanced design has $N/10$ participants in each cell. With a cell mean coding, $\boldsymbol{B}$ (10×4) contained mean tortuosity for each combination of Age, Gender and brain Region. Since age seems like a natural source of variation, age factor will be not considered in this dissertation to apply our theorem. Hence the study was designed with simpler hypothesis of the most concern, namely Gender $\times$ Region.

The covariance matrix of SOAM1 (radians/cm) in four regions is provided from data with $\nu_e = 12$ ($N = 13$) participants from Bullitt, Muller, Jung, Lin and Aylward (2004a):

$$\widehat{\boldsymbol{\Sigma}} = \begin{bmatrix} 0.0838 & 0.0502 & 0.0356 & 0.0533 \\ 0.0502 & 0.0537 & 0.0325 & 0.0333 \\ 0.0356 & 0.0325 & 0.0441 & 0.0386 \\ 0.0533 & 0.0333 & 0.0386 & 0.0722 \end{bmatrix}. \tag{3.18}$$

An appropriate sample size is needed to get the desirable power for the Gender $\times$ Region

interaction hypothesized. Using reference cell coding to conduct the power analysis gives

$$\boldsymbol{B} = \mu \cdot \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} + \qquad (3.19)$$

$$\delta_{\mathrm{G}} \cdot \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} +$$

$$\delta_{\mathrm{R}} \cdot \begin{bmatrix} -1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \end{bmatrix} +$$

$$\delta_{\mathrm{GR}} \cdot \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Here μ represents the grand mean while δ_{G}, δ_{R}, and δ_{GR} correspond to the effects of Gender, Region, and the Gender $\times$ Region interaction respectively.

For the sake of brevity, the reduced model is used with the assumption that no Age, Age $\times$ Gender, Age $\times$ Region, or Age $\times$ Gender $\times$ Region effect occur. The reduced model has only Gender as a between effect (and Region within), so $\mathrm{Es}(\boldsymbol{X}) = \boldsymbol{I}_2$. For a balanced design $\boldsymbol{X} = \boldsymbol{I}_2 \otimes \boldsymbol{1}_m$ with m the cell size (REPN in POWERLIB software). For $= \boldsymbol{C}\boldsymbol{B}\boldsymbol{U}$ with $H_0 : \boldsymbol{\Theta} = \boldsymbol{0}, \boldsymbol{C} = [\,1 \; {-}1\,]$ compares genders. One choice for testing Region differences (and Gender $\times$ Region) uses $\boldsymbol{U} = [\,\boldsymbol{u}_1 \; \boldsymbol{u}_2 \; \boldsymbol{u}_3\,]$ with $\boldsymbol{u}_1' = [\,{-}1 \; 1 \; 0 \; 0\,]$, $\boldsymbol{u}_2' = [\,{-}1 \; 0 \; 1 \; 0\,]$, and $\boldsymbol{u}_3' = [\,{-}1 \; 0 \; 0 \; 1\,]$.

3.4.2 The Equivalent Univariate Model Using The Transformations

The example was computed in SAS/IML (SAS Institute, 1999). The free software POWERLIB 2.3, which may be downloaded at no cost at *http://www.bios.unc.edu/~muller*, was used for power calculation within the simulations. The appropriate $\boldsymbol{\Sigma}, \boldsymbol{X}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{U}, \boldsymbol{\Theta}_0$ and α completely determine a power analysis. A data analysis and transformation starting from an original model is explained in detail in Appendix A. A complete and balanced mixed model (no missing or mistimed data) and with no repeated covariate may use the strategy.

The reduced design and the test for Region $\times$ Gender will be used to illustrate the transformation to a univariate for a case with $b > 1$ and $a = 1$ so $s = 1$. Here

$$\boldsymbol{C}_A = [\,1 \quad -1\,],\, \boldsymbol{U}_A = \begin{bmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\, \text{and } \boldsymbol{\Theta}_A = [\,0 \quad 0 \quad 0\,].\ \text{For ten observations in}$$

each group $\boldsymbol{X}_A = \begin{bmatrix} \mathbf{1}_{10} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{10} \end{bmatrix}$. For computational convenience, the example used

$\boldsymbol{B}_A = \begin{bmatrix} 0 & 0 & \delta & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, and $\boldsymbol{\Sigma}_A$ is given as $\boldsymbol{\Sigma}$ in equation 3.18. The unscaled noncentrality

parameter $\boldsymbol{\Delta}_A$ is the 3×3 matrix

$$\begin{aligned} \boldsymbol{\Delta}_A &= (\boldsymbol{\Theta}_A - \boldsymbol{\Theta}_{A0})' \boldsymbol{M}_A^{-1} (\boldsymbol{\Theta}_A - \boldsymbol{\Theta}_{A0}) \\ &= (\boldsymbol{\Theta}_A - \boldsymbol{\Theta}_{A0})' [\boldsymbol{C}_A (\boldsymbol{X}_A' \boldsymbol{X}_A)^{-} \boldsymbol{C}_A']^{-1} (\boldsymbol{\Theta}_A - \boldsymbol{\Theta}_{A0}) , \end{aligned} \tag{3.20}$$

and the noncentrality parameter $\boldsymbol{\Omega}_A$ is the 3×3 matrix

$$\begin{aligned} \boldsymbol{\Omega}_A &= \boldsymbol{\Delta}_A \boldsymbol{\Sigma}_{A*}^{-1} \\ &= (\boldsymbol{\Theta}_A - \boldsymbol{\Theta}_{A0})' [\boldsymbol{C}_A (\boldsymbol{X}_A' \boldsymbol{X}_A)^{-} \boldsymbol{C}_A']^{-1} (\boldsymbol{\Theta}_A - \boldsymbol{\Theta}_{A0}) \boldsymbol{\Sigma}_{A*}^{-1} . \end{aligned} \tag{3.21}$$

Here $\boldsymbol{\Sigma}_{A*} = \boldsymbol{U}_A' \boldsymbol{\Sigma}_A \boldsymbol{U}_A$. The hypothesis test for Gender $\times$ Region has $s = \min(a, b) = 1$, which allows applying Theorem 3.1. The hypothesis, $H_0: \boldsymbol{\Theta}_A = \boldsymbol{\Theta}_{A0}$, considers $\boldsymbol{\Theta}_A = \boldsymbol{C}_A \boldsymbol{B}_A \boldsymbol{U}_A$ (1×3), for fixed and known $\boldsymbol{C}_A, \boldsymbol{U}_A$, and $\boldsymbol{\Theta}_{A0}$. Here $\boldsymbol{C}_A$ (1×2) gives contrasts between groups, and $\boldsymbol{U}_A$ (4×3) gives contrasts within person. A brief summary of the transformation detailed in Appendix A is shown next.

Step A. The original model is identified as $\boldsymbol{Y}_A = \boldsymbol{X}_A \boldsymbol{B}_A + \boldsymbol{E}_A$ which has 20 rows (10 observations in each group). Here $\boldsymbol{Y}_A$ has 4 columns, and $\boldsymbol{X}_A$ has 2 columns, with $\boldsymbol{E}_A \sim \mathcal{N}_{20,4}(\mathbf{0}, \boldsymbol{I}_{20}, \boldsymbol{\Sigma}_A)$. The parameters and constants in terms of the original multivariate model and hypothesis test must be given. Here they are $\boldsymbol{X}_A, \boldsymbol{B}_A, \boldsymbol{C}_A, \boldsymbol{U}_A, \boldsymbol{\Sigma}_A$, and $\boldsymbol{\Theta}_A$. In practice, for some applications considered

Step B. Doing the left linear transformation by multiplying $\boldsymbol{U}_A$ to the original model. i.e. $\boldsymbol{Y}_A \boldsymbol{U}_A = \boldsymbol{X}_A \boldsymbol{B}_A \boldsymbol{U}_A + \boldsymbol{E}_A \boldsymbol{U}_A$. The new model is $\boldsymbol{Y}_B = \boldsymbol{X}_B \boldsymbol{B}_B + \boldsymbol{E}_B$ which has 20 rows, $\boldsymbol{Y}_B$ has 3 columns, and $\boldsymbol{X}_B$ has 2 columns, with $\boldsymbol{E}_B \sim \mathcal{N}_{20,3}(\mathbf{0}, \boldsymbol{I}_{20}, \boldsymbol{\Sigma}_B)$ where

$$\boldsymbol{\Sigma}_B = \boldsymbol{U}_A' \boldsymbol{\Sigma}_A \boldsymbol{U}_A = \begin{bmatrix} 0.0371 & 0.0305 & 0.0136 \\ 0.0305 & 0.0567 & 0.0335 \\ 0.0136 & 0.0335 & 0.0494 \end{bmatrix},\, \boldsymbol{C}_B = \boldsymbol{C}_A = [\,1 \quad -1\,],\, \text{and } \boldsymbol{U}_B = \boldsymbol{I}_3.$$

Step C. This step allows using a simpler contrast matrix, C_C. The only changes are to X_B, B_B and C_B. The matrix of $X_B B_B$ equals to $(X_B T^{-1})(T B_B)$ with specific matrix T. Producing T using singular value decomposition of C_B is detailed in Appendix A and Theorem 3.1. The result is model C, $Y_C = X_C B_C + E_C$ which has 20 rows, while Y_C has 3 columns, and $X_C = \begin{bmatrix} 0.5 & 0.7071 \\ -0.5 & 0.7071 \end{bmatrix} \otimes \mathbf{1}_{10}$, with $E_C \sim \mathcal{N}_{20,3}(\mathbf{0}, I_{20}, \Sigma_C)$, $\Sigma_C = \Sigma_B$, $C_C = [\,1 \quad 0\,]$, and $U_C = I_3$.

Step D. Next model C is transformed by a multiplication on the right. The result is model D, $Y_D = X_D B_D + E_D$, which has $N_A - 1 = 19$ rows. Also Y_C has 3 columns, and X_C has 1 column, with $E_D \sim \mathcal{N}_{19,3}(\mathbf{0}, I_{19}, \Sigma_C)$, $\Sigma_D = \Sigma_C$, $C_D = 1$, and $U_C = I_3$.

Step E. The final step is to transform to the univariate model, $y_F = X_F \beta_F + e_F$. It has $N_A - 1 = 19$ rows, Y_C has 3 columns, and X_C has 1 column, with $E_D \sim \mathcal{N}_{19,3}(\mathbf{0}, I_{19}, \Sigma_C)$, $\Sigma_D = \Sigma_C$, $C_D = 1$, and $U_C = I_3$. Also $e_F \sim \mathcal{N}_{19}(\mathbf{0}, I)$, $X_F' = [\,I_3 \quad \mathbf{0}\,]'$, 3×1, $\beta_F = [\,\sqrt{\omega_E} \quad \mathbf{0}\,]' = [\,15.65 \quad \mathbf{0}\,]'$, $C_F = I_3$, $U_F = \{1\}$, and $\theta_{0F} = \mathbf{0}$ with $\omega_E = \text{tr}(\Omega_E) = (x_{E1}' x_{E1})\text{tr}(B_E \Sigma_E^{-1} B_E') = 244.98$.

The exact power and the confidence interval for power are exactly same for the original multivariate model (before being transformed) and the corresponding univariate case. The traditional approach to power analysis computes a single number. A single number fails to capture the uncertainty due to estimating the covariance matrix. The uncertainty varies greatly with the estimation sample size, and the power value. The larger the estimation sample size, the narrower the confidence interval because the larger sample size reduce the uncertainty of error covariance.

Even though the prior study had the information for μ ($= 3.2$) and δ_R ($= 0.30$), there was no information available about the effect of Gender or the interaction of Gender with Region. As seen in the model, the interaction parameter only corresponds to a localized Posterior region difference. Full rank coding schemes give the advantage of making explicit specification of such parameter matrices straightforward for power analysis. Muller et al.

(2007) displays the pattern of interaction means that results if $\delta_{GR} = 0.16$ while $\delta_G = 0$.

Using $\delta_{GR} = 0.16$, gives exact power of 0.917 with $N = 51$.

The power confidence interval accounting for having used an estimate of the error covariance with $\nu_e = 12$ is summarized with various sample sizes and δ_{GR} values in Table 3.1. Values include $N \in \{40, 80\}$, which corresponds to $m \in \{20, 40\}$ participants per gender, with $\delta_{GR} \in [0, 0.60]$. In addition, the same cases are considered but under the (false) assumption that $\nu_e = 36$ in order the illustrate the impact of estimation sample size. In POWERLIB, the table is produced by choosing RANKEST=1; and NEST=13; or NEST=37. As shown in Table 3.1, power is only changed based upon target sample size, not based upon N_{est}. However, the confidence interval for power is affected by N_{est}. In other words, the results illustrate the changes in uncertainty due to changes in population properties affecting power.

Table 3.1
Confidence Interval for power, $s = 1$
Gender × Region Tortuosity

$\nu_{e,est}$	N	δ	Lower Limit	Power Estimate	Upper Limit
12	40	0.12	0.219	0.543	0.850
		0.16	0.370	0.813	0.984
		0.30	0.907	1.000	1.000
		0.54	0.997	1.000	1.000
	80	0.12	0.439	1.000	0.995
		0.16	0.704	0.991	1.000
		0.30	0.999	1.000	1.000
		0.54	1.000	1.000	1.000
36	40	0.12	0.338	0.543	0.741
		0.16	0.567	0.813	0.948
		0.30	0.990	1.000	1.000
		0.54	1.000	1.000	1.000
	80	0.12	0.657	0.886	0.978
		0.16	0.903	0.991	1.000
		0.30	1.000	1.000	1.000
		0.54	1.000	1.000	1.000

The same phenomenon can be see graphically. Power values with exact confidence regions for $\delta_{GR} \in [0, 0.60]$ are presented for $N = 51$ (power of 0.917 for $\delta_{GR} = 0.16$) in Figure 3.2 for $\nu_{e,est} = 12$ (the correct estimation degrees of freedom) and in Figure 3.3 for $\nu_{e,est} = 36$.
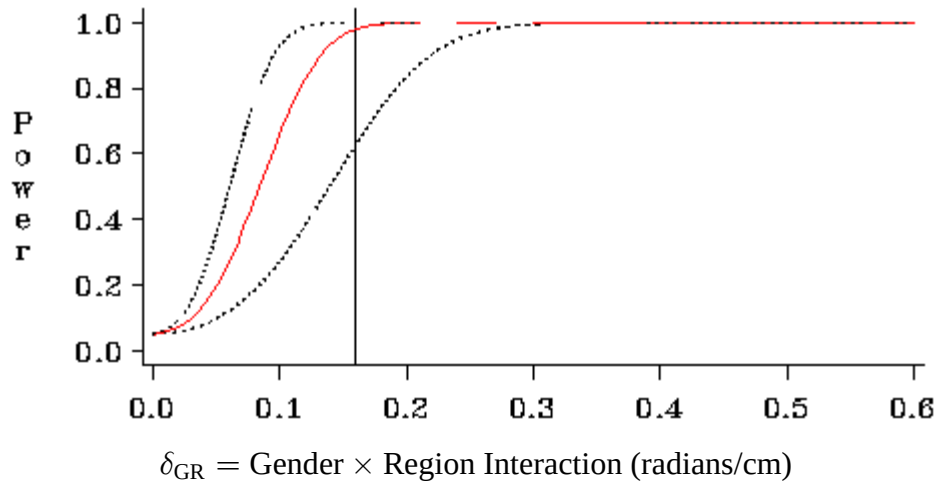
$\delta_{GR} =$ Gender $\times$ Region Interaction (radians/cm)
Figure 3.2 Confidence Interval for Power to the test Gender $\times$ Region at $\alpha = 0.05/6$ with 51 participants , $\nu_{e,est} = 12$
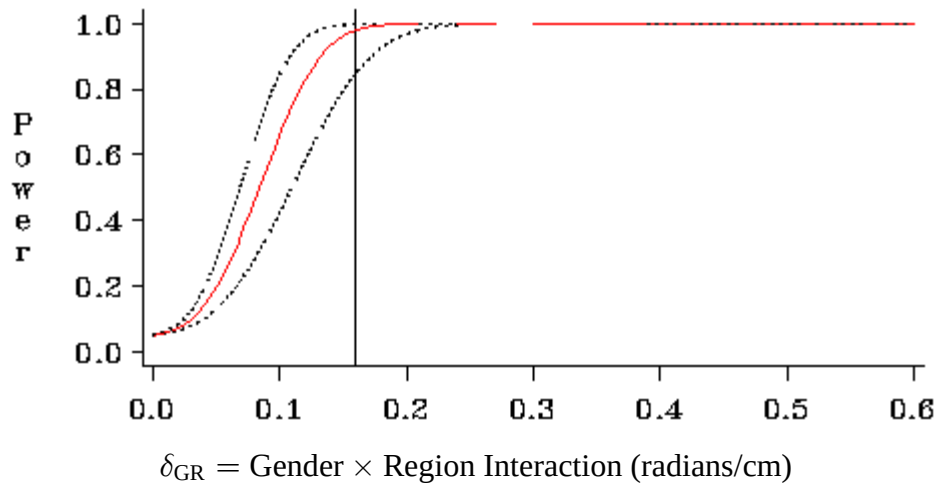
$\delta_{GR} =$ Gender $\times$ Region Interaction (radians/cm)
Figure 3.3 Confidence Interval for Power to the test Gender $\times$ Region at $\alpha = 0.05/6$ with 51 participants, $\nu_{e,est} = 36$

Chapter 4 (Paper 2) Internal Pilots for a One or Two Group Test, and Related Models

4.1 Introduction

4.1.1 Motivation

Blood vessels are affected by many diseases and Figure 4.1 displays a vessel map for a normal person, with roughly 25-50 segments in four regions of the brain (anterior, posterior, left middle, and right middle). Bullitt et al. (2004a) recently demonstrated that computer software can measure cerebral vascular tortuosity (bending, twisting, or winding) automatically from MRI data. Furthermore, the approach appears to allow automatic discrimination between benign and malignant tumors.

Anter-Post   Lateral   Basal
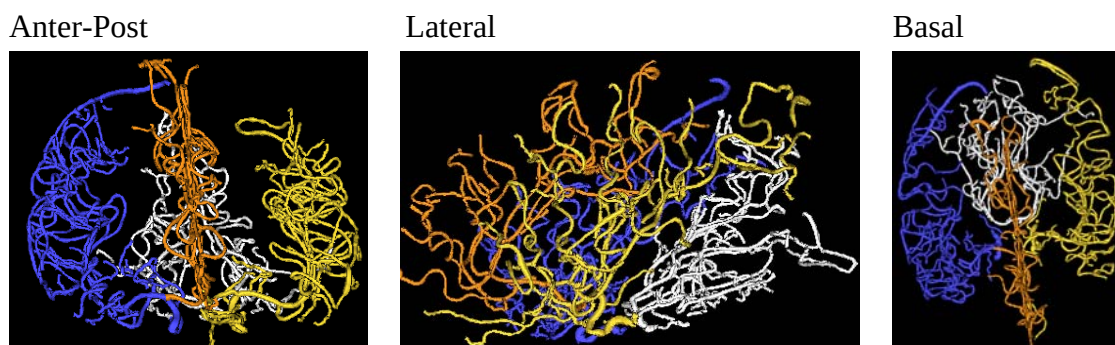
Figure 4.1 Human cerebral vessel 3D *tortuosity*

Information from the study supports the assumption of Gaussian distributions. A desire to examine the effects of age and gender led the investigators to plan to recruit a new group of participants. The estimated covariance matrix from the previous study was used to determine the sample size. However, the new study includes a much wider range of ages than previously. Hence an internal pilot has great appeal due to the uncertainty about the covariance value and due to concern for bias. Although extensive exact results have been developed for the univariate case, only limited results are available for repeated measures

and multivariate models. We restrict attention to hypotheses involving one or two groups, as in typical clinical trials. For the "multivariate" approach, we describe how to exactly transform any such repeated measures model to an equivalent univariate model. In turn, known exact results from univariate internal pilot theory apply, and provide the advantages of internal pilots to a useful class of repeated measures and multivariate linear models.

4.1.2 Adaptive Designs and Group Sequential Designs

Designs with interim analyses have been developed in clinical trials since they were introduced at first in 1970's (Pocock, 1977). Interim analysis is a good method because it can detect early benefits and potential harmful effects. Recently, sample size re-estimation or internal pilot studies have been popular. They consider re-estimation of sample size based on interim information about the values of nuisance parameters like variances.

Appropriate sample size is very important in a design of clinical trial. Economic pressure may lead to an underpowered study. An overpowered study may waste resources. Therefore, a considerable amount of research in sample size adjustment has been one of top issue in clinical trials recently.

Group sequential designs, and internal designs, are special cases of adaptive designs. Group sequential designs involve one or more interim analyses. The most common approach allows stopping early only under the alternative or after the last planned group. However, some designs also allow stopping early under the null. Discussion of adaptive design has involved more discussion of increasing or decreasing the sample size, depending on the interim data. An internal pilot allows either, but without any interim data analysis (only interim power analysis). A group sequential design appeals with ethical concerns pushing towards early termination and saving exposure to an ineffective treatment (and saving resources). Groups sequential designs have fixed maximum sample sizes. More general adaptive designs, including internal pilots, can allow modifying the sample size to avoid an underpowered or overpowered study.

35

4.1.3 Why Use an Internal Pilot?

Using confidence interval based sample size calculation addresses uncertainty but does not address bias. An internal pilot is one of the adaptive designs which allows modifying the variance value used to choose sample size and thereby change sample size, based on fixed means. It can be a good method when there are economic pressures because it starts with a smaller sample size and increases the sample size until the minimum meaningful treatment effect corresponds to good power when combined with the observed variance estimate. An internal pilot does not allow the investigator to see in estimated means from the interim data.

It is hard to specify the nuisance parameter which is needed to choose the sample size and power. Therefore, internal pilots can be used when sample size and nuisance parameters need to be re-estimated.

4.2 Theory of Internal Pilots for Special Cases

Coffey and Muller (1999, 2000, 2001) described much exact theory for an internal pilot for the general linear univariate model with Gaussian errors. Coffey and Muller (2003) studied properties of internal pilots with the univariate approach to repeated measures. Given that no UMP-α test exists for general cases of the multivariate model, in some case a MULTIREP approach will be preferred to a UNIREP approach.

Table 4.1
Steps for an Internal Pilot Study with a GLUM

1. Specify test size (α), target power (P_t), design, hypothesis (H_0), mean difference (δ), original variance estimate $(\widehat{\sigma}_0^2)$, and proportion for internal pilot (π)
2. Use $\widehat{\sigma}_0^2$ to find a total sample size, n_0, to get a target power (P_t)
3. Use the first observations $(n_1 = \pi n_0)$ to find internal pilots variance estimate $(\widehat{\sigma}_1^2)$
4. Use $\widehat{\sigma}_1^2$ to find the number of second sample size (n_2) to find P_t.
5. Test the hypothesis on the $n_1 + n_2 = n_+$ observations

Table 4.2

Steps for an Internal Pilot Study with Repeated Measures

1. Choose a sample size, n_0, typically determined from an initial value of the $p \times p$ covariance, $\mathbf{\Sigma}_0$. Next compute the $b \times b$ matrix $\mathbf{\Sigma}_0 = \mathbf{U}'\mathbf{\Sigma}_0\mathbf{U}$.

2. Collect $n_1 = \pi n_0$ observations as the internal pilot sample, and obtain $\widehat{\mathbf{\Sigma}}_{*1} = \mathbf{U}'\widehat{\mathbf{\Sigma}}_1\mathbf{U}$.

3. Ignore the randomness of the final sample size, N_+, and the use $\widehat{\mathbf{\Sigma}}_{*1}$ with the M-B approx. to choose the N_+ needed to achieve the target power.

Steps for an internal pilot study with the univariate model and with a multivariate model (such as for repeated measure) are summarized in Table 4.1 and Table 4.2. Thee steps were described in Coffey and Muller (1999, 2003).

It is very convenient that internal pilot results for the univariate model can be used in some important special case of the multivariate model. The transformation shown in Chapter 3 tells us that power for the multivariate linear model can be expressed in terms of an equivalent univariate linear model for $s = 1$. Therefore, the theory and application of an internal pilot for a univariate linear model can be applied in this case.

Corollary 4.4 If $s = 1$, then all exact and approximate internal pilot methods for a univariate linear model can be applied to any multivariate linear model.

Applying an internal pilot to a multivariate linear model in the special case $s = 1$ involves three steps:

1) apply the transformation in Theorem 3.1 to the univariate model;

2) follow the steps for an internal pilot study with a univariate model;

3) interpret the results in terms of the multivariate model.

The first step can be done easily by following the lemmas introduced in Chapter 3. If the transformation mentioned in the first step has been completed, then the procedures for step 2 (for a univariate model) are mentioned in Table 4.1. It will be shown in next section how to interpret the results.

After completing the transformation, the easiest way to apply the method is to use the internal pilot program (available free at http://www.soph.uab.edu/ccoffey). The internal pilot

design has been developed thoroughly in the univariate linear model. Given a univariate model version of the multivariate model of interest, the existing free software to compute the power and sample size at the internal pilot stage can be applied. Doing so allows taking advantage of an internal pilot design with full confidence in accuracy even with a small sample size.

4.3 A Practical Implementation Process

Simulation will not be done in this chapter. We do not need any simulations due to the exact nature of the transformation to an equivalent univariate model for $s = 1$ case. Hence the exact and approximate properties known in the univariate case apply to the multivariate case. This section contains a description of internal pilot model formulation in the multivariate case with $s = 1$.

The notation used is the General Linear Multivariate Model (GLMM) notation in Muller, Edwards, Taylor and Simpson (2005) and the internal pilot study notation in Coffey and Muller (2003). In the original model random observed $\boldsymbol{Y}$ ($N \times p$, independent sampling units as rows, repeated measures as columns), fixed observed $\boldsymbol{X}$, and unobserved $\boldsymbol{E}$ such that $\text{row}_i(\boldsymbol{E})' \sim \mathcal{N}_p(\boldsymbol{0}, \boldsymbol{\Sigma})$, independent of $\text{row}_j(\boldsymbol{E})$ for $i \neq j$. The internal pilot design has two models, n_1 first sample used in the internal pilot and N_2 second samples. Therefore, the model for the final analysis is

$$
\begin{bmatrix} \boldsymbol{Y}_1 \\ n_1 \times p \\ \boldsymbol{Y}_2 \\ N_2 \times p \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_1 \\ n_1 \times q \\ \boldsymbol{X}_2 \\ N_2 \times q \end{bmatrix} q \underset{\times}{\boldsymbol{B}} p + \begin{bmatrix} \boldsymbol{E}_1 \\ n_1 \times p \\ \boldsymbol{E}_2 \\ N_2 \times p \end{bmatrix} \tag{4.22}
$$
$$
\begin{array}{ccccc} \boldsymbol{Y}_+ & = & \boldsymbol{X}_+\boldsymbol{B} & + & \boldsymbol{E}_+ \\ N_+ \times p & & N_+ \times q \times p & & N_+ \times p \end{array},
$$

with partitioning corresponding to the n_1 and, random, N_2 observations in the internal pilot and second samples, respectively. The total sample size, $N_+ = n_1 + N_2$, may or may not be increased. We require $\text{Es}(\boldsymbol{X}_1) = \text{Es}(\boldsymbol{X}_2) = \text{Es}(\boldsymbol{X}_+)$, which means $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ and $\boldsymbol{X}_+$ are all span the same space, and have the same rank, r.

The usual test statistic for the GLMM are $\tilde{\boldsymbol{B}}(n_k) = (\boldsymbol{X}_k'\boldsymbol{X}_k)^-\boldsymbol{X}_k'\boldsymbol{Y}_k$, $\boldsymbol{H}_k = \boldsymbol{X}_k(\boldsymbol{X}_k'\boldsymbol{X}_k)^-\boldsymbol{X}_k'$, $\widehat{\boldsymbol{\Sigma}}(n_k) = \boldsymbol{Y}_k'(\boldsymbol{I}_{n_k} - \boldsymbol{H}_k)\boldsymbol{Y}_k/\nu_k$. The General Linear Hypothesis is H$_0$: $\boldsymbol{\Theta} = \boldsymbol{CBU} = \boldsymbol{\Theta}_0$, with $\boldsymbol{C}$ a fixed between-subject contrast matrix and $\boldsymbol{U}$ a fixed within-subject contrast and $\nu_k = n_k - r$ where $k \in \{1, +\}$. No matter how the original model is the multivariate case, we have already shown that the multivariate model can be transformed to the univariate case in special case, $s = 1$. Therefore known exact univariate results (Coffey and Muller, 1999; 2000a; 2000b; 2001; 2004) apply. The model transformation creates a univariate model with $N_{F+} = (N_+ - r + a)$ observations. The final model can be expressed as the univariate form

$$\underset{N_{F+} \times 1}{\boldsymbol{y}_+} = \underset{N_{F+} \times b \times 1}{\boldsymbol{X}_+\boldsymbol{\beta}} + \underset{N_{F+} \times 1}{\boldsymbol{e}_+}, \tag{4.23}$$

for b the number of column of $\boldsymbol{U}$ in the original multivariate model.

The following lemma states the cumulative probability of N_{F+} by using a lemma given by Coffey and Muller (1999) for power in the GLUM with an internal pilot design. It can be applied to compute the desired probability, $\Pr\{N_{F+} = n_{F+}\}$.

Lemma 4.1 Let $\widehat{\sigma}_1^2$ be the internal pilot variance estimate for the equivalent univariate model, and N_{F+} be defined as previously, with n_{F+} denoting a particular value of the random final sample size. Then, with $n_{F1} = n_1 - r + a$ the number of row in the transformed model at the time of interim power analysis,

$$\begin{aligned}
\Pr\{N_{F+} \leq n_{F+}\} &= \Pr\{\widehat{\sigma}_1^2 \leq \sigma^2(n_{F+})\} \\
&= \Pr\{\chi^2(n_{F1} - b) \leq q_2(n_{F+})\} .
\end{aligned} \tag{4.24}$$

The value of $q_2(n_{F+})$ is a chi square quantile satisfying the equation in terms of expressions given in Coffey and Muller (1999) in terms of the *univariate model equivalent* to the multivariate model.

By taking advantage of the discreteness of sample size, the probability mass associated with a given value of sample size can be computed as

$$\Pr\{N_{F+} = n_{F+} + 1\} = \Pr\{N_{F+} \leq n_{F+} + 1\} - \Pr\{N_{F+} \leq n_{F+}\} . \qquad (4.25)$$

4.4 Numerical Examples

4.4.1 Overview

Exact power and internal pilot properties were computed with the free software for internal pilots mentioned earlier. All examples compare an internal pilot to a fixed sample size design with $\alpha = 0.05$, and a desired power of $P_t, = 0.90$. Both use the same mean difference of δ, and a variance value of σ_0^2. The internal pilot is applied with the same values for sample size allocation rules indicated in Coffey and Muller(1999): (i) the choice of π; (ii) a bound on $\max(N_{F+})$, and (iii) a bound on $\min(N_{F+})$. Moreover, two different scenarios for bounding $\min(N_{F+})$: (i) $\min(N_{F+}) = n_0$, are considered in which there is no decrease in the original sample size, and (ii) $\min(N_{F+}) = n_0$, in which there may be a decrease in the original sample size.

4.4.2 Example: Tortuosity Study

An internal pilot can be applied to find an appropriate sample size in a new study of tortuosity. The scientists' interest is on the effect of age and gender on vessel abnormality of tortuosity in the brain. Therefore, new subjects need to be recruited. An internal pilot is a good method to deal with the uncertainty of the error variance due to the possibility of population shift. As mentioned in the previous chapter, the study is to be designed to have good power for the most complex hypothesis of concern, namely Gender $\times$ Region. The initial study design is the example in the previous chapter. To test the null hypothesis, we can apply Theorem 3.1 because $s = 1$ ($a = 1$), and transform to an appropriate univariate model. Hence an internal pilot for a univariate model can be used in this example.

In planning the fixed sample study, the required sample size was 51 with $\alpha_t = 0.05$, $P_t = 0.90$, $\boldsymbol{\Theta}_{cr} = \begin{bmatrix} 0 & \delta & 0 \end{bmatrix}$, and $\widehat{\boldsymbol{\Sigma}}_*$ from an earlier study. However, owing to uncertainty surrounding the estimated covariance matrix used to determine the sample size, an internal pilot has great appeal.

Theorem 3.1 defines a transformation from the multivariate linear model to a univariate linear model. The transformation defines the new form of the design which is used in the univariate internal pilot design.

The result of Chapter 3 suggested the 51 sample were required to get the target power, 0.90. A current internal pilot program (available free at http://www.soph.uab.edu/ccoffey) can be used the restricted sample size, which must be the integer multiplication of the number of row of $\boldsymbol{X}$. The number of $\boldsymbol{X}$ equals to the number of columns in $\boldsymbol{U}$ in the transformed univariate model. The hypothesis test of tortuosity study has 3 columns in and the required sample size is not divided by the number of row of $\boldsymbol{X}$. The uncorrected (naive fixed sample) test was used in all computations.

Table 4.3 summarizes the results on test size, power, and expected sample size with $\delta = 0.16$. One half of the originally planned sample size ($\pi = 0.47$, $n_1 = 24$) was used for the internal pilot study. Disallowing any reduction in sample size (requiring $\min(N_{F+}) = n_0$) takes away the opportunity for an internal pilot to do better if the original design was pessimistic. The last few rows of the table illustrates that it does allow retaining power despite overly optimistic planning, while the fixed sample loses power.

Table 4.3
Example, Tortuosity Study:
$\min(N_{F+}) = n_0, \sigma_0^2 = 1, n_0 = 51, \ n_1 = 24, \delta = 0.16, \pi = 0.47$

σ^2/σ_0^2	Internal Pilot			Fixed sample size design		
	$\mathcal{E}(N_{F+})$	α	Power	N_{F+}	α	Power
0.50	51	0.05	0.998	51	0.05	0.998
0.75	51.2	0.05	0.976	51	0.05	0.975
1.00	54.1	0.05	0.941	51	0.05	0.917
1.50	72.2	0.05	0.915	51	0.05	0.762
2.00	94.1	0.05	0.907	51	0.05	0.624

Table 4.4 summarizes same conditions as in Table 4.3 but with $\min(N_{F+}) = n_1$. When $\min(N_{F+}) = n_1$, the sample size is more variable compared to $\min(N_{F+}) = n_0$. Therefore, it allows reducing the sample size and saving costs. Furthermore, the ability to avoid power loss due to optimistic planning is retained, giving the scientist essentially the best features of both optimistic and pessimistic planning.

Table 4.4

Example, Tortuosity Study:

$\min(N_{F+}) = n_1, \sigma_0^2 = 1, n_0 = 51, \; n_1 = 24, \delta = 0.16, \pi = 0.5$

σ^2/σ_0^2	Internal Pilot			Fixed sample size design		
	$\mathcal{E}(N_{F+})$	α	Power	N_{F+}	α	Power
0.50	28.8	0.05	0.947	51	0.05	0.998
0.75	38.9	0.05	0.924	51	0.05	0.975
1.00	49.5	0.05	0.910	51	0.05	0.917
1.50	65.2	0.05	0.875	51	0.05	0.762
2.00	71.8	0.05	0.801	51	0.05	0.624

Table 4.5 illustrates the same point by using for $n_1 = 24$ and $\pi = 0.3$ with two different bounding conditions for the internal pilot study. Although the ratio of sample size of internal pilot design is changed, the result tells us that the internal pilot design has chosen a better sample size than the fixed design. The internal pilot has many advantages over a fixed sample design when the variance is misspecified.

Table 4.5

Example, Tortuosity Study:

Internal Pilot, $\sigma_0^2 = 1, n_0 = 51, n_1 = 15, \delta = 0.16, \pi = 0.3$

σ^2/σ_0^2	$\min(N_{F+}) = n_0$			$\min(N_+) = n_1$		
	$\mathcal{E}(N_{F+})$	α	Power	N_{F+}	α	Power
0.50	51.1	0.05	0.998	27.9	0.05	0.927
0.75	52.7	0.05	0.979	38.5	0.05	0.903
1.00	57.6	0.05	0.944	47.5	0.05	0.885
1.50	74.4	0.05	0.892	58.4	0.05	0.826
2.00	95	0.05	0.870	62.8	0.05	0.735

Table 4.6 illustrates that the choice of π affect test size. As the minimum total sample size n_1 decreases, test size tends to have greater inflation. Such test size inflation has already been mentioned in the Coffey and Muller (1999). The results in Table 4.6 tells us that the

inflation of α must be considered in small samples with an internal pilot study in the special case multivariate models. Hence one of the tests that control test size, as discussed in Coffey and Muller (2001) must be used. They particularly recommended a bounding test, which is implemented in their software, among some other popular choices.

Table 4.6
Example, Tortuosity Study:
Test size, $\sigma_0^2 = 1$, $n_0 = 51$, $\delta = 0.16$

σ^2/σ_0^2	$\min(N_{F+}) = n_0$			$\min(N_+) = n_1$		
	$n_1 = 12$	$n_1 = 24$	$n_1 = 36$	$n_1 = 12$	$n_1 = 24$	$n_1 = 36$
0.50	0.050	0.050	0.050	0.069	0.055	0.050
0.75	0.051	0.050	0.050	0.066	0.060	0.053
1.00	0.052	0.052	0.051	0.061	0.059	0.056
1.50	0.053	0.054	0.055	0.056	0.055	0.056
2.00	0.053	0.054	0.054	0.053	0.052	0.053

4.5 Elaborations and Conclusions

An artful choice of linear model parameter formulations greatly simplifies the interpretation. The tables were produced after running the program for the internal pilot study based on the transformation to a univariate model. Therefore, a backwards step must be made to interpret the results. Fortunately, the interpretation can always be simplified whenever $s = 1$. The special cases of interest have $a = 1$, which guarantees that the unscaled noncentrality matrix $\mathbf{\Delta}$ always has rank 1 (under the alternative). Therefore the $\mathbf{B}$ matrix always has rank 1 (with $s = 1$, under the alternative), which implies one can always choose a $\mathbf{B}$ matrix with only one nonzero element and the rest zero.

Having chosen a $\mathbf{B}$ matrix with one nonzero element, choosing simple design matrices and contrast matrices (such as orthonormal ones) allows simple expressions for $\mathbf{\Theta}$ and noncentrality. In the formulation chose here, there is no difference in δ, power and α with the transformation. Scalar multiples may occur, depending the particular choices made. Only the expected the sample size needs to be explained in terms of the multivariate linear model. The final sample size after transformation is $N_{F+} - r + a$ with N_{F+} total sample

size, r the rank of $\boldsymbol{X}$, and a the number of row in $\boldsymbol{C}$. As mentioned in Chapter 3, the hypothesis test for the δ_{GR} interaction has $r = 2$ and $a = 1$. Consequently the final sample size after transformation is $N_{F+} - 2 + 1 = N_{F+} - 1$ in our example. If $\sigma^2/\sigma_0^2 = 1$ and $\min(N_{F+}) = n_0$ in Table 4.3, the expected sample size, $\mathcal{E}(N_+)$, is 54.1 to get 0.941 power in the univariate model. In turn the expected sample size is 55.1 in the multivariate linear model. Therefore at least 27.5 participants are needed per gender.

Except for the additional complication of interpretation, other previously mentioned properties of the internal pilot design with univariate models naturally also hold for the special case multivariate model with $s = 1$. An internal pilot design helps prevent wasting resources and helps avoid low power due to uncertainty about error covariance. Moreover, the bias can be avoided (roughly on the average) because a portion of the target population is used to estimate the error covariance matrix. However, test size inflation can occur in small sample sizes, which requires using special internal pilot software (available free) to control it by using special tests, such as the bounding approach.

Chapter 5 (Paper 3) The Confidence Interval Due to Estimated Covariance for Power of a Gaussian Multivariate General Linear Model

5.1 Introduction

Power analysis helps choose an appropriate sample size. Multivariate linear models are widely used for a variety of studies. As discussed in previous chapters, using an estimated covariance matrix causes uncertainty about the power value. Inappropriate sample size can waste resources or make the study useless due to low power.

In Chapter 3 exact results for confidence limits of in the univariate case were proven to apply to multivariate linear models whenever $s = 1$ (1 or 2 sample designs as well as any single degree of freedom between, $a = 1$). In the univariate case, we have a UMP-α test. An equivalent univariate model can be found in a multivariate case with $s = 1$, which implies there is also a UMP-α test. However, the multivariate general linear model with $s > 1$ does not have a UMP-α test. Which multivariate statistic is most powerful varies with the population covariance pattern.

In this chapter, I extend the previous results and describe approximate confidence intervals for power for $s > 1$. With $\mathbf{\Omega} = \mathbf{\Delta}\mathbf{\Sigma}_*^{-1}$ and $s_* = \text{rank}(\mathbf{\Omega})$, some partially exact results are available for $s_* = 1$.

Muller and Peterson (1984) provided power approximations based upon noncentral F distributions for multivariate statistics HLT, PBT and W. Muller et al. (1992) surveyed power methods for the general linear multivariate model with Gaussian errors and recommended how to choose power analysis designs. They discussed the benefits of power analysis and also mentioned obstacles to power analysis of the multivariate model. The biggest obstacle is choosing $\mathbf{\Sigma}$, which often includes uncertainty due to using an estimate.

5.2 Approximate Results

5.2.1 Some Useful Wishart and Matrix Properties

The definition and following three Wishart theorems are in Chapter 10 of Muller and Stewart (2006). The two matrix theorems are in Chapter 1 of Muller and Stewart (2006).

Definition 5.1 Suppose a random matrix S of a dimension $p \times p$ is distributed as Wishart, with ν degrees of freedom and parameter matrix Σ of a dimension $p \times p$, i.e., $S \sim W_p(\nu, \Sigma)$. Then $V = S^{-1}$ is said to be distributed as inverted Wishart, denoted by $V \sim IW_p(\nu, \Sigma^{-1})$ and the density of V is given by

$$f(V) = \frac{2^{-(\nu-p-1)p/2}|\Sigma^{-1}|^{(\nu-p-1)2}}{\Gamma_p[(\nu-p-1)/2]|V|^{\nu/2}} \exp\left[-\frac{1}{2}\text{tr}\left(V^{-1}\Sigma^{-1}\right)\right]. \tag{5.1}$$

Theorem 5.1 If $S \sim W_p(\nu, \Sigma)$, then, for any full rank constant T of dimension $p \times q$, with $q \leq p$,

$$T'ST \sim W_q(\nu, T'\Sigma T). \tag{5.2}$$

Corollary. Diagonal elements of an inverted Wishart matrix are distributed proportional to an inverted gamma.

Proof. If $T = [0, \cdots, 0, 1, 0, \cdots, 0]'$ is a $p \times 1$ vector with 1 in element i. Applying Theorem 5.1 gives

$$\begin{aligned} T'ST &= s_{ii} \\ &\sim W_1\left(\nu, \sigma_i^2\right) \\ &\sim \chi^2(\nu) \end{aligned} \tag{5.3}$$

for s_{ii} and σ_i^2 are the i-th diagonal elements of S and Σ, respectively. Hence diagonal elements of S^{-1} is $s_{ii}^{-1} \sim c/G$, with gamma distribution G and constant c. □

Theorem 5.2 The marginal sum of squares matrix from a Wishart is Wishart. In particular, if

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \sim \mathcal{W}_p(\nu, \Sigma),$$

with Σ partitioned to match as $\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$, S_{11} $r \times r$ positive definite, then

$$S_{11} \sim \mathcal{W}_r(\nu, \Sigma_{11}) . \tag{5.4}$$

Theorem 5.3 The conditional sum of squares matrix from a Wishart is Wishart. If $S_{1.2} = S_{11} - S_{12}S_{22}^{-1}S_{21}$, then

$$S_{1.2} \sim \mathcal{W}_r(\nu - p + r, \Sigma_{1.2}) . \tag{5.5}$$

Theorem 5.4 Let the $m \times m$ matrix A that is partitioned into the 2×2 block form given by $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ where A_{11} is $m_1 \times m_1$, A_{12} is $m_1 \times m_2$, A_{21} is $m_2 \times m_1$, and A_{22} is $m_2 \times m_2$, and suppose that A, A_{11}, and A_{12} are nonsingular matrices. Then the inverse matrix of A, $B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$ where the submatices of B are of the same sizes as the corresponding submatrices of A. Then we have

$$\text{(a) } B_{11} = \left(A_{11} - A_{12}A_{22}^{-1}A_{21} \right)^{-1} \tag{5.6}$$
$$\text{(b) } B_{12} = -A_{11}^{-1}A_{12}B_{22} \tag{5.7}$$
$$\text{(c) } B_{21} = -A_{22}^{-1}A_{21}B_{11} \tag{5.8}$$
$$\text{(d) } B_{22} = \left(A_{22} - A_{21}A_{11}^{-1}A_{12} \right)^{-1} . \tag{5.9}$$

Theorem 5.5 Suppose A and B are nonsingular matrices, with A being $m \times m$ and B being $n \times n$. For any $m \times n$ matrix C and any $n \times m$ matrix D, it follows that if $A + CBD$ is nonsingular then

$$(A + CBD)^{-1} = A^{-1} - A^{-1}C\left(B^{-1} + DA^{-1}C \right)^{-1}DA^{-1} . \tag{5.10}$$

5.2.2 Results for Concentrated Noncentrality (One Nonzero Canonical Correlation)

If $s_* = 1$ and $s > 1$, an exact result for the distribution of estimated noncentrality can be derived for the power approximations of at least two MULTIREP tests. The case of interest has estimated covariance, $\widehat{\Sigma}$, and fixed B giving fixed Δ. The distributions of the resulting

estimated noncentrality scalar parameter for the power approximations of the Hotelling-Lawley trace (HLT) and Pillai-Bartlett Trace (PBT) are provided in the following theorems.

The functions lie at the heart of power approximations based on $\boldsymbol{\Delta}$ (fixed) and $\widehat{\boldsymbol{\Sigma}}_* = \boldsymbol{U}'\widehat{\boldsymbol{\Sigma}}\boldsymbol{U}$ (random). The results cover the case with $s > 1$ but only one nonzero canonical correlation. i.e. $s_* = \text{rank}(\boldsymbol{\Omega}) = 1$ with $\boldsymbol{\Omega} = \boldsymbol{\Delta}\boldsymbol{\Sigma}_*^{-1}$.

Theorem 5.6 For $\boldsymbol{\Delta}$ (fixed) and $\widehat{\boldsymbol{\Sigma}}$ (random), if $s_* = 1$, then approximate noncentrality function for HLT, $\text{tr}[\boldsymbol{\Delta}(\widehat{\boldsymbol{\Sigma}}_*\nu_e)^{-1}]$, is distributed proportional to an inverted gamma variable.

Proof. We can use a spectral decomposition to write

$$
\begin{aligned}
\text{tr}\big[\boldsymbol{\Delta}\big(\widehat{\boldsymbol{\Sigma}}_*\nu_e\big)^{-1}\big] &= \text{tr}\big[\boldsymbol{\Delta}\big(\widehat{\boldsymbol{\Sigma}}_*\nu_e\big)^{-1}\big] \qquad\qquad (5.11)\\
&= \text{tr}\big[\boldsymbol{F}_{\boldsymbol{\Delta}}\boldsymbol{F}_{\boldsymbol{\Delta}}'\big(\widehat{\boldsymbol{\Sigma}}_*\nu_e\big)^{-1}\big]\\
&= \text{tr}\big[\boldsymbol{F}_{\boldsymbol{\Delta}}'\big(\widehat{\boldsymbol{\Sigma}}_*\nu_e\big)^{-1}\boldsymbol{F}_{\boldsymbol{\Delta}}\big]
\end{aligned}
$$

for symmetric matrix $\boldsymbol{\Delta} = \boldsymbol{V}\boldsymbol{D}\boldsymbol{V}' = \begin{bmatrix} \boldsymbol{V}_+ & \boldsymbol{V_0} \end{bmatrix}\begin{bmatrix} \boldsymbol{D}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{V}_+' \\ \boldsymbol{V_0}' \end{bmatrix} = \boldsymbol{V}_+\boldsymbol{D}_1\boldsymbol{V}_+',$

$\boldsymbol{F}_{\boldsymbol{\Delta}} = \boldsymbol{V}_+\boldsymbol{D}_1^{1/2}$ with a full rank constant of dimension $b \times s_*$, with $s_* \leq b$, and a diagonal matrix $\boldsymbol{D}_1$ under the null hypothesis. We know that $\widehat{\boldsymbol{\Sigma}}_*\nu_e$ is a Wishart distribution, which implies $(\widehat{\boldsymbol{\Sigma}}_*\nu_e)^{-1}$ is distributed as an inverted Wishart. i.e. $(\widehat{\boldsymbol{\Sigma}}_*\nu_e)^{-1} \sim I\mathcal{W}_b(\nu_e, \boldsymbol{\Sigma}_*)$. In turn $\boldsymbol{S}_F^{-1} = \boldsymbol{F}_{\boldsymbol{\Delta}}'(\widehat{\boldsymbol{\Sigma}}_*\nu_e)^{-1}\boldsymbol{F}_{\boldsymbol{\Delta}} \sim I\mathcal{W}_{s_*}(\nu_e, \boldsymbol{F}_{\boldsymbol{\Delta}}'\boldsymbol{\Sigma}_*\boldsymbol{F}_{\boldsymbol{\Delta}})$. Diagonal elements of an inverse Wishart are distributed proportional to $1/G$, with $G \sim \gamma(,)$ (Gupta and Nagar, 2000). In particular, for $\langle \boldsymbol{S}^{-1}\rangle_{i,i}$, for $1 \leq i \leq s_*$, indicating the diagonal element i, $\langle \boldsymbol{S}^{-1}\rangle_{i,i} = 1/G_i$ is scaled chi square, $1/G_i \sim c_i/X$ with constant c_i and $X \sim \chi^2(\nu_e)$. Hence, if $s_* = 1$,

$$
\text{tr}[\boldsymbol{\Delta}(\widehat{\boldsymbol{\Sigma}}_*\nu_e)^{-1}] = 1/G \qquad\qquad (5.12)
$$

with $G \sim \gamma(,)$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Muller and Stewart (Chapter 3 and Chapter 21, 2007) give explicit forms for transforming HLT to an approximate central or noncentral F (for the null and alternative cases respectively). We can compute approximate power as

$$\text{Power} \approx 1 - F_F[f_{\text{crit}}(\text{HLT}), \nu_1(\text{HLT}), \nu_2(\text{HLT}), \omega(\text{HLT})] . \tag{5.13}$$

Here $\omega(\text{HLT})$ is a 1–1 function (specifically a linear transformation) of $\text{tr}(\boldsymbol{\Delta}\boldsymbol{\Sigma}_*^{-1})$.

The theorem allows concluding that the Hotelling-Lawley trace approximate noncentrality function, denoted by $\omega(\text{HLT})$, is a 1–1 function of an inverted gamma if the hypothesis test has one nonzero canonical correlation. Therefore an approximate confidence interval for power can be computed in terms of lower (α_{cL}) and upper (α_{cU}) tail probabilities and the α_{cL} and α_{cU} quantiles of a gamma. The quantiles imply corresponding quantiles for $\text{tr}[\boldsymbol{\Delta}(\widehat{\boldsymbol{\Sigma}}_*\nu_e)^{-1}]$, which in turn imply corresponding quantiles for $\widehat{\omega}(\text{HLT})$, which in turn imply corresponding quantiles for power. The chain does not break because all transformations are smooth and 1–1. Thus an approximate confidence interval for power is

$$\Pr\{\widehat{P}_L \le \text{Power} \le \widehat{P}_U\}. \tag{5.14}$$

Theorem 5.7 For $\boldsymbol{\Delta}$ (fixed) and $\widehat{\boldsymbol{\Sigma}}$ (random), if $s_* = 1$, then the Pillai-Bartlett Trace approximate noncentrality function $\text{tr}(\boldsymbol{\Delta}\boldsymbol{S}_t^{-1})$, has density

$$f_P(p) = f_W(w)\frac{(1+\lambda w)^2}{\lambda} = \frac{(1-p)^{(\tau/2)-1}e^{-\tau/2}}{p(\lambda p)^{(\tau/2)}2^{\tau/2}\boldsymbol{\Gamma}(\tau/2)} \quad . \tag{5.15}$$

Proof. Here

$$\begin{aligned}
\text{tr}(\boldsymbol{\Delta}\boldsymbol{S}_t^{-1}) &= \text{tr}\left[\boldsymbol{V}\boldsymbol{D}\boldsymbol{V}'(\boldsymbol{V}\boldsymbol{D}\boldsymbol{V}'+\boldsymbol{S}_e)^{-1}\right] \tag{5.16} \\
&= \text{tr}\{\boldsymbol{V}\boldsymbol{D}\boldsymbol{V}'[\boldsymbol{V}(\boldsymbol{D}+\boldsymbol{V}'\boldsymbol{S}_e\boldsymbol{V})\boldsymbol{V}']^{-1}\} \\
&= \text{tr}\left[\boldsymbol{V}\boldsymbol{D}\boldsymbol{V}'(\boldsymbol{V}')^{-1}(\boldsymbol{D}+\boldsymbol{V}'\boldsymbol{S}_e\boldsymbol{V})^{-1}\boldsymbol{V}^{-1}\right] \\
&= \text{tr}\left[\boldsymbol{D}\boldsymbol{V}'(\boldsymbol{V}')^{-1}(\boldsymbol{D}+\boldsymbol{V}'\boldsymbol{S}_e\boldsymbol{V})^{-1}\boldsymbol{V}^{-1}\boldsymbol{V}\right] \\
&= \text{tr}\left[\boldsymbol{D}(\boldsymbol{D}+\boldsymbol{V}'\boldsymbol{S}_e\boldsymbol{V})^{-1}\right],
\end{aligned}$$

with

$$V'S_eV = \begin{bmatrix} V'_+ \\ V'_0 \end{bmatrix} S_e [V_+ \quad V_0] \tag{5.17}$$

$$= \begin{bmatrix} V'_+S_eV_+ & V'_+S_eV_0 \\ V'_0S_eV_+ & V'_0S_eV_0 \end{bmatrix}$$

$$= \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

$$\sim \mathcal{W}_{s_*}(\nu_e, V'\Sigma_*V)$$

and

$$D_T = \begin{bmatrix} D_1^{1/2} & 0 \\ 0 & I \end{bmatrix} \tag{5.18}$$

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} = D_T^{-1} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} D_T^{-1} \tag{5.19}$$

$$= \begin{bmatrix} D_1^{-1/2}S_{11}D_1^{-1/2} & D_1^{-1/2}S_{12} \\ S_{21}D_1^{-1/2} & S_{22} \end{bmatrix}$$

In turn

$$\operatorname{tr}(\Delta S_t^{-1}) = \operatorname{tr}\left[D\left(\begin{bmatrix} D_1^{1/2} & 0 \\ 0 & 0 \end{bmatrix}^2 + D_T T D_T \right)^{-1} \right] \tag{5.20}$$

$$= \operatorname{tr}\left[D\left(\begin{bmatrix} D_1^{1/2} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} D_T + D_T T D_T \right)^{-1} \right]$$

$$= \operatorname{tr}\left[\begin{bmatrix} D_1^{1/2} & 0 \\ 0 & 0 \end{bmatrix} \left(\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} + T \right)^{-1} D_T^{-1} \right]$$

$$= \operatorname{tr}\left[\left(\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} + T \right)^{-1} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \right]$$

$$= \operatorname{tr}\left(\begin{bmatrix} I + T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}^{-1} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \right)$$

$$= \operatorname{tr}\left(\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \right)$$

$$= \operatorname{tr}(B_{11}).$$

Using the formula for the inverse of a partitioned matrix,

$$\begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} + \boldsymbol{T}_{11} & \boldsymbol{T}_{12} \\ \boldsymbol{T}_{21} & \boldsymbol{T}_{22} \end{bmatrix}^{-1} \tag{5.21}$$

$$= \begin{bmatrix} [(\boldsymbol{I} + \boldsymbol{T}_{11}) - \boldsymbol{T}_{12}\boldsymbol{T}_{22}^{-1}\boldsymbol{T}_{21}]^{-1} & -(\boldsymbol{I} + \boldsymbol{T}_{11})^{-1}\boldsymbol{T}_{12}\boldsymbol{B}_{22} \\ -\boldsymbol{T}_{22}^{-1}\boldsymbol{T}_{21}\boldsymbol{B}_{11} & \left(\boldsymbol{T}_{22} - \boldsymbol{T}_{21}(\boldsymbol{I} + \boldsymbol{T}_{11})^{-1}\boldsymbol{T}_{12}\right)^{-1} \end{bmatrix}.$$

Recalling $\begin{bmatrix} \boldsymbol{T}_{11} & \boldsymbol{T}_{12} \\ \boldsymbol{T}_{21} & \boldsymbol{T}_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{D}_1^{-1/2}\boldsymbol{S}_{11}\boldsymbol{D}_1^{-1/2} & \boldsymbol{D}_1^{-1/2}\boldsymbol{S}_{12} \\ \boldsymbol{S}_{21}\boldsymbol{D}_1^{-1/2} & \boldsymbol{S}_{22} \end{bmatrix},$

$$\text{PBT} = \text{tr}(\boldsymbol{B}_{11}) \tag{5.22}$$

$$= \text{tr}\left[\left((\boldsymbol{I} + \boldsymbol{T}_{11}) - \boldsymbol{T}_{12}\boldsymbol{T}_{22}^{-1}\boldsymbol{T}_{21} \right)^{-1} \right]$$

$$= \text{tr}\left[\left(\boldsymbol{I} + \boldsymbol{D}_1^{-1/2}\boldsymbol{S}_{11}\boldsymbol{D}_1^{-1/2} - \boldsymbol{D}_1^{-1/2}\boldsymbol{S}_{12}\boldsymbol{S}_{22}^{-1}\boldsymbol{S}_{21}\boldsymbol{D}_1^{-1/2} \right)^{-1} \right]$$

$$= \text{tr}\left[\left(\boldsymbol{I} + \boldsymbol{D}_1^{-1/2}(\boldsymbol{S}_{11} - \boldsymbol{S}_{12}\boldsymbol{S}_{22}^{-1}\boldsymbol{S}_{21})\boldsymbol{D}_1^{-1/2} \right)^{-1} \right]$$

$$= \text{tr}\left[\left(\boldsymbol{I} + \boldsymbol{D}_1^{-1/2}\boldsymbol{S}_{1.2}\boldsymbol{D}_1^{-1/2} \right)^{-1} \right].$$

We know that the conditional distribution of Wishart is also Wishart and then

$$\boldsymbol{S}_{1.2} \sim \mathcal{W}_{s_*}(\nu_e - b + s_*, \boldsymbol{\Sigma}_{1.2}) \tag{5.23}$$

$$\Leftrightarrow \quad \boldsymbol{D}_1^{-1/2}\boldsymbol{S}_{1.2}\boldsymbol{D}_1^{-1/2} \sim \mathcal{W}_{s_*}\left(\nu_e - b + s_*, \boldsymbol{D}_1^{-1/2}\boldsymbol{\Sigma}_{1.2}\boldsymbol{D}_1^{-1/2} \right)$$

If $\boldsymbol{W} = \boldsymbol{D}_1^{-1/2}\boldsymbol{S}_{1.2}\boldsymbol{D}_1^{-1/2} \sim \mathcal{W}_{s_*}(\tau, \boldsymbol{\Sigma_W})$ with $\tau = \nu_e - b + s_*$, $\boldsymbol{\Sigma_W} = \boldsymbol{D}_1^{-1/2}\boldsymbol{\Sigma}_{1.2}\boldsymbol{D}_1^{-1/2} = \boldsymbol{D}_1^{-1/2}(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})\boldsymbol{D}_1^{-1/2}$, and $s_* = 1$, then $W \sim \chi^2(\tau)$ with $\tau = \nu_e - b + 1$ and $\text{tr}(\boldsymbol{\Delta}\boldsymbol{S}_t^{-1}) = 1/(1 + \lambda W)$. If $P = 1/(1 + \lambda W)$ then $W = (1 - P)/(\lambda P)$, $dW = -(\lambda P^2)^{-1}dP$ and the distribution of P, $f_P(p)$, is

$$f_P(p) = f_W(w)\frac{(1 + \lambda w)^2}{\lambda} \tag{5.24}$$

$$= \frac{(1 - p)^{(\tau/2)-1}e^{-\tau/2}}{p(\lambda p)^{(\tau/2)}2^{\tau/2}\Gamma(\tau/2)} = \frac{e^{-\tau/2}}{(2\lambda)^{\tau/2}\Gamma(\tau/2)}\left[\frac{(1 - p)^{\tau/2-1}}{p^{1+\tau/2}} \right].$$

$\square$

Similar to Hotelling-Lawley trace, the Pillai-Bartlett Trace (PBT) has an easily computed exact distribution for the approximate noncentrality function if the hypothesis test has one nonzero canonical correlation. Numerical integration to compute the CDF, and numerical inversion of the CDF will produce quantiles of the noncentrality function $\text{tr}(\boldsymbol{\Delta} \boldsymbol{S}_t^{-1})$, and in turn $\widehat{\omega}(\text{PBT})$ needed for the F approximation. In parallel to the computations for HLT, the confidence interval endpoints for approximate noncentrality give confidence interval endpoints for approximate power. The simple form of the density may allow closed form integration.

5.2.3 Results for The General Case

A simple approximation is always available. For test $M \in \{\text{HLT}, \text{PBT}, \text{W}\}$, and $\widehat{\boldsymbol{\Sigma}}$ with fixed $\boldsymbol{\Delta}$, the approximation takes the form

$$\text{Power}(M) \approx 1 - F_F[f_{\text{crit}}(M), \nu_1(M), \nu_2(M), \widehat{\omega}(M)] . \tag{5.25}$$

Here $\widehat{\omega}(\text{HLT})$ is a 1–1 function of $\text{tr}(\boldsymbol{\Delta} \widehat{\boldsymbol{\Sigma}}_*^{-1})$, while $\widehat{\omega}(\text{PBT})$ is a 1–1 function of $\text{tr}[\boldsymbol{\Delta}(\boldsymbol{\Delta} + \nu_{et} \widehat{\boldsymbol{\Sigma}}_*)^{-1}]$ for ν_{et} the error *df* for the target study, and $\widehat{\omega}(\text{W})$ is a 1–1 function of $|\nu_{et} \widehat{\boldsymbol{\Sigma}}_* (\boldsymbol{\Delta} + \nu_{et} \widehat{\boldsymbol{\Sigma}}_*)^{-1}|$. Muller and Stewart (2006) and Muller give the explicit forms. If $\widehat{\boldsymbol{\Sigma}}$ is based on ν_{ee} error *df* for the estimation study, then

$$\widehat{\omega}_L(M) = \widehat{\omega}(M) \cdot c_{\text{crit}}(\alpha_{cL}|\nu_{ee})/\nu_{ee} \tag{5.26}$$

and

$$\widehat{\omega}_U(M) = \widehat{\omega}(M) \cdot c_{\text{crit}}(1 - \alpha_{cU}|\nu_{ee})/\nu_{ee} . \tag{5.27}$$

A GLMM power analysis can be done by specifying seven values: $\alpha, \boldsymbol{X}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{U}, \boldsymbol{\Sigma}$, and $\boldsymbol{\Theta}_0$, as mentioned in Muller et al (1992). Once the factors are chosen the degrees of freedom are fixed because they depend on the dimensions of the model. To do the power analysis, $\boldsymbol{\Theta} = \boldsymbol{CBU}$, $\boldsymbol{\Theta}_0$ and $\boldsymbol{\Sigma}_* = \boldsymbol{U}'\boldsymbol{\Sigma}\boldsymbol{U}$ are sufficient instead of specifying all factors. As a further simplication, $\boldsymbol{\Omega} = (\boldsymbol{\Theta} - \boldsymbol{\Theta}_0)'\boldsymbol{M}^{-1}(\boldsymbol{\Theta} - \boldsymbol{\Theta}_0)\boldsymbol{\Sigma}_*^{-1}$ in addition to the degrees of freedom suffice. In turn, the degrees of freedom and the eigenvalues of $\boldsymbol{\Omega}$ are the minimal

information required (hence the adjective "canonical"). With eigenvalue k indicated by $\omega_k = N\rho_k^2/(1 - \rho_k^2)$, the value ρ_k^2 is a squared canonical correlation. In consequence, the minimal sufficient factors to do power analysis are a set of canonical correlation and the dimensions. The next theorem contains the result that for a testable hypothesis implies a hypothesis equivalent model with nonzero $\boldsymbol{B}$ values expressed as simple functions of $\{\omega_k\}$, or equivalently, $\{\rho_k\}$.

Theorem 5.8 Any combination of multivariate linear model and testable general linear hypothesis has an associated hypothesis equivalent model and hypothesis in which $\boldsymbol{B}$ is a diagonal matrix. The associated model may be chosen so that the diagonal element k is the square root of eigenvalue k of the original noncentrality matrix, $\omega_k^{1/2}$, giving $q \times p$ matrix $\boldsymbol{B} = \begin{bmatrix} \mathrm{Dg}^{1/2}(\boldsymbol{\omega}) & \boldsymbol{0} \\ & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}$. As always $\omega_k = N\rho_k^2/(1-\rho_k^2)$ for ρ_k the canonical correlation k and N the total sample size. The associated model has $\boldsymbol{\Sigma} = \boldsymbol{I}_p$ and uses $\boldsymbol{C} = [\,\boldsymbol{I}_a \ \boldsymbol{0}\,]$, $\boldsymbol{U} = [\,\boldsymbol{I}_b \ \boldsymbol{0}\,]'$ and $\boldsymbol{\Theta}_0 = \boldsymbol{0}$ to test the hypothesis.

Proof. The associated pair is not unique. We need only construct a convenient choice and show that the degrees of freedom and noncentrality matrix coincide with the original. Here $\boldsymbol{Y}_M = \boldsymbol{X}_M\boldsymbol{B}_M + \boldsymbol{E}_M$ for $M \in \{A, B\}$ indicates the original or equivalent model, with corresponding secondary parameter $\boldsymbol{\Theta}_M = \boldsymbol{C}_M\boldsymbol{B}_M\boldsymbol{U}_M$ and hypotheses $H_0 : \boldsymbol{\Theta}_M = \boldsymbol{\Theta}_{0,M}$. Some of the zero submatrices may be of dimension zero, depending on the relatives sizes of $\{q, p, a, b\}$. The SVD gives $\boldsymbol{X}_A = [\,\boldsymbol{T}_+ \quad \boldsymbol{T}_0\,]\begin{bmatrix} \mathrm{Dg}(\boldsymbol{s}) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{V}_+' \\ \boldsymbol{V}_0' \end{bmatrix}$ with $N \times r$ matrix $\boldsymbol{T}_+$, and $N \times (N - r)$ matrix $\boldsymbol{T}_0$. The $N \times N$ matrix $\boldsymbol{T} = [\,\boldsymbol{T}_+ \quad \boldsymbol{T}_r\,]$ is such that $\boldsymbol{T}'\boldsymbol{T} = \boldsymbol{I}_N$. If $\boldsymbol{X}_B = [\,\boldsymbol{T}_+ \quad \boldsymbol{0}\,]$ is $N \times q$ with $q - r$ columns of zeros, then $\boldsymbol{X}_B'\boldsymbol{X}_B = \begin{bmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}$. In the following ρ_k is canonical correlation k and $\omega_k = N\rho_k^2/(1 - \rho_k^2)$. Also $\mathrm{Dg}(\boldsymbol{\omega})$ is a $b \times b$ diagonal matrix with the $s_* \le s = \min(a, b)$ nonzero $\{\omega_k\}$ in positions $(1, 1)$ to (s_*, s_*). In turn

$$\underset{(a\times q)(q\times p)(p\times b)}{\boldsymbol{\Theta}_B = \boldsymbol{C}_B\boldsymbol{B}_B\boldsymbol{U}_B} = [\,\boldsymbol{I}_a \quad \boldsymbol{0}\,]\begin{bmatrix}\begin{bmatrix}\mathrm{Dg}^{1/2}(\boldsymbol{\omega}) & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{I}_b\\ \boldsymbol{0}\end{bmatrix} = \underset{(a\times b)}{\begin{bmatrix}\mathrm{Dg}^{1/2}(\boldsymbol{\omega}) & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}}$$

$$\boldsymbol{M}_B = \boldsymbol{C}_B(\boldsymbol{X}_B'\boldsymbol{X}_B)^-\boldsymbol{C}_B' = [\,\boldsymbol{I}_a \quad \boldsymbol{0}\,]\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}[\,\boldsymbol{I}_a \quad \boldsymbol{0}\,]' = \boldsymbol{I}_a$$

$$\boldsymbol{\Delta}_B = (\boldsymbol{\Theta}_B - \boldsymbol{\Theta}_{0,B})'\boldsymbol{M}_B^{-1}(\boldsymbol{\Theta}_B - \boldsymbol{\Theta}_{0,B}) = \underset{(a\times b)}{\begin{bmatrix}\mathrm{Dg}^{1/2}(\boldsymbol{\omega}) & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}}'\boldsymbol{I}_a\underset{(a\times b)}{\begin{bmatrix}\mathrm{Dg}^{1/2}(\boldsymbol{\omega}) & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}} = \underset{(b\times b)}{\mathrm{Dg}(\boldsymbol{\omega})}$$

$$\boldsymbol{\Sigma}_{*B} = \boldsymbol{U}_B'\boldsymbol{\Sigma}_B\boldsymbol{U}_B = [\,\boldsymbol{I}_b \quad \boldsymbol{0}\,]'\boldsymbol{I}_p[\,\boldsymbol{I}_b \quad \boldsymbol{0}\,] = \boldsymbol{I}_b$$

$$\boldsymbol{\Omega}_B = \boldsymbol{\Delta}_B\boldsymbol{\Sigma}_{*B}^{-1} = \begin{bmatrix}\mathrm{Dg}(\boldsymbol{\omega}) & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\boldsymbol{I}_b^{-1} = \underset{(b\times b)}{\begin{bmatrix}\mathrm{Dg}(\boldsymbol{\omega}) & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}}.$$

Observing that all dimensions of corresponding model and hypothesis matrices are the same, and that eigenvalues of $\boldsymbol{\Omega}_B$ coincide with those of $\boldsymbol{\Omega}_B$ completes the proof. $\qquad\square$

5.3 Simulations

5.3.1 Simulation Methods

Which multivariate test is best? All multivariate test statistics are functions of the eigenvalues of the $b \times b$ matrix $\widehat{\boldsymbol{\Omega}} = \widehat{\boldsymbol{\Delta}}\widehat{\boldsymbol{\Sigma}}_*^{-1}$. If $s = 1$, then the four multivariate test statistics have the same p-values and power. However, if $s > 1$, there does not exist a most powerful unbiased test (among similarly invariant tests). The decision about the most powerful test depends upon the set of eigenvalues of $\boldsymbol{\Omega}$, $\{\omega_k\}$.

The testable hypotheses can be restricted with full rank $\boldsymbol{C}, \boldsymbol{U}, \boldsymbol{\Sigma}_*, \boldsymbol{M}$ and $\boldsymbol{C} = \boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^-(\boldsymbol{X}'\boldsymbol{X})$. Muller et al. (1992) discussed that the GLM power calculation with fixed predictors, $\alpha, \boldsymbol{\Sigma}, \boldsymbol{X}, \boldsymbol{C}, \boldsymbol{U}$ and $\boldsymbol{\Theta}_0$. In the end, the eigenvalues of $\boldsymbol{\Omega}$ give the minimally sufficient information to do the power analysis. Therefore, we can produce a new set of predictors with simpler forms depending upon the canonical correlation. The only additional values needed to do the test are a, b, and $\nu_e = N - r$. Therefore, the tests are invariant if you have the same statistics, $\{\rho_k\}$, a, b, and ν_e.

The simulations that follow use the same basic design. It has $\mathrm{Es}(\boldsymbol{X}) = \boldsymbol{I}_3$, and a balanced design so $\boldsymbol{X} = \boldsymbol{I}_3 \otimes \boldsymbol{1}_n$ with $q = r = \mathrm{rank}(\boldsymbol{X}) = 3$. Furthermore $\boldsymbol{\Sigma} = \boldsymbol{I}_p$, $\boldsymbol{C} = [\boldsymbol{I}_2 \quad \boldsymbol{0}]$ so $a = 2$ and $\boldsymbol{U}[\boldsymbol{I}_{p-1} \quad \boldsymbol{0}]'$ so $b = 3$. Since the Hotelling-Lawley (HLT) statistic is $N\sum_k \widehat{\rho}_k^2 / (1 - \widehat{\rho}_k^2) = \mathrm{tr}(\boldsymbol{S}_h \boldsymbol{S}_e^{-1})$, we can set the new hypothesis after reduction without changing the a, b, and ν_e. Then, with n the cell size (replication factor),

$$
\boldsymbol{B} = \sqrt{\frac{N}{n}} \cdot
\begin{bmatrix}
\sqrt{\rho_1^2/(1-\rho_1^2)} & 0 & 0 & 0 \\
0 & \sqrt{\rho_2^2/(1-\rho_2^2)} & 0 & 0 \\
0 & 0 & \sqrt{\rho_3^2/(1-\rho_3^2)} & 0
\end{bmatrix}
\tag{5.28}
$$

expresses the parameter matrix in terms of the canonical correlation.

All simulations were conducted in SAS/IML. Without loss of generality, $\boldsymbol{\Sigma} = 1$ for all simulations. The matrix $\boldsymbol{B}$ was chosen with four different canonical correlation matrix, i) decreasing with zero elements, ii) decreasing with nonzero elements, iii) equal values and iv) equal values with zero elements. i.e. $\rho^2 = \mathrm{Dg}(.7, .4, 0, 0)$, $\mathrm{Dg}(.5, .03, .01, .001)$, $\mathrm{Dg}(.1, .1, .1, .1)$ and $\mathrm{Dg}(.5, .5, 0, 0)$. The SAS Normal function was used to generate pseudo-random independent, identically distributed Gaussian data. In computing observed power, the free software LINMOD 3.3, which may be downloaded at no cost at *http://www.bios.unc.edu/~muller*, was used for all linear models computations within the simulations.. Power was tabulated for 50,000 replications per condition.

For tabulating confidence interval coverage, $\boldsymbol{\Delta}$ was fixed and a pseudo-random sample of N_{est} observations was collected and $\widehat{\boldsymbol{\Sigma}}$ was calculated. For each replication a binary value was computed to indicate whether the observed confidence interval covered the population predict or observed power.

5.3.2 Simulation Results

Table 5.1 through Table 5.3 contain observed and predicted power (approximated with the free software POWERLIB) for the three different tests, HLT, PBT and W. Values of β_P for simulation are reported to eight digits to allow others to use the same conditions in future

work. Mean/max absolute difference between observed and expected power were 0.004/0.02 for HLT, 0.014/0.107 for PBT and 0.011/0.028 for W. Overall, HLT test power approximated more accurately than W power, which was more accurate approximated that PBT power. Overall, the simulations of power are consistent with Olson's (1974, 1976, 1979) results.

Table 5.4 contains the results of simulations based upon the predicted power where $N_{\text{est}} = 30$, $\nu_{\text{est}} = 27$ and target power coverage level $= 0.95$. Coverage is somewhat high except for correlation pattern case 2, when it is somewhat low for HLT. Not surprisingly, accuracy increases with increasing target sample size (N) and fixed estimation sample size (N_{est}). Overall, the coverage is reasonably accurate, and certainly adequate for most study planning applications.

Table 5.5 through Table 5.7 compare the observed power coverage with N_{est} of 30 or 60. The results for N_{est} of 30 are very similar to the predicted power coverages due to the accuracy of the approximations. Surprisingly, with N fixed, as N_{est} increases the observed power coverage level increases. The nature of the approximations seems to ensure that the increase reflects an error of programming either in the simulation or in the implementation of the approximation in POWERLIB.

Table 5.1 *HLT Observed and Predicted Power*
Std. Err. of Observed < 0.0023
Canonical Correlation Matrix Indexed by ρ.
Target Power Implied by β_P

N	ρ	Target power	β_P	Obs. power
15	1	0.2	1.2403995	0.199
		0.5	2.0121646	0.510
		0.8	2.7367126	0.815
	2	0.2	2.1159630	0.199
		0.5	3.4324956	0.496
		0.8	4.6684820	0.798
	3	0.2	4.5575186	0.198
		0.5	7.3931647	0.513
		0.8	10.055324	0.820
	4	0.2	1.5191729	0.196
		0.5	2.4643882	0.510
		0.8	3.3517748	0.818
30	1	0.2	1.0809277	0.200
		0.5	1.7391939	0.504
		0.8	2.3471213	0.805
	2	0.2	1.8439245	0.203
		0.5	2.9668425	0.501
		0.8	4.0038890	0.794
	3	0.2	3.9715819	0.199
		0.5	6.3902063	0.503
		0.8	8.6238744	0.804
	4	0.2	1.3238606	0.203
		0.5	2.1300688	0.502
		0.8	2.8746248	0.805
60	1	0.2	1.0274175	0.199
		0.5	1.6510441	0.500
		0.8	2.2254904	0.800
	2	0.2	1.7526429	0.203
		0.5	2.8164703	0.499
		0.8	3.7964021	0.799
	3	0.2	3.7749729	0.202
		0.5	6.0663235	0.499
		0.8	8.1769737	0.804
	4	0.2	1.2583243	0.199
		0.5	2.0221078	0.501
		0.8	2.7257579	0.803

Table 5.2 *PBT Observed and Predicted Power*
Std. Err. of Observed < 0.0023
Canonical Correlation Matrix Indexed by ρ.
Target Power Implied by β_P

N	ρ	Target power	β_P	Obs. power
15	1	0.2	1.1438613	0.183
		0.5	1.8821430	0.483
		0.8	2.5997122	0.803
	2	0.2	2.0524526	0.177
		0.5	3.7351729	0.506
		0.8	6.0677481	0.907
	3	0.2	4.1106372	0.187
		0.5	6.5797578	0.482
		0.8	8.8512159	0.783
	4	0.2	1.3702124	0.184
		0.5	2.1932526	0.479
		0.8	2.9504053	0.782
30	1	0.2	1.0516693	0.195
		0.5	1.7074546	0.498
		0.8	2.3326156	0.810
	2	0.2	1.8269215	0.194
		0.5	3.0621623	0.511
		0.8	4.3914229	0.856
	3	0.2	3.8294050	0.193
		0.5	6.1371552	0.489
		0.8	8.2578685	0.790
	4	0.2	1.2764683	0.196
		0.5	2.0457184	0.488
		0.8	2.7526228	0.789
60	1	0.2	1.0156269	0.197
		0.5	1.6395923	0.497
		0.8	2.2248055	0.804
	2	0.2	1.7462787	0.199
		0.5	2.8563368	0.505
		0.8	3.9489891	0.828
	3	0.2	3.7164016	0.198
		0.5	5.9625074	0.493
		0.8	8.0277601	0.797
	4	0.2	1.2388005	0.196
		0.5	1.9875025	0.496
		0.8	2.6759200	0.796

Table 5.3 *Wilks' Observed and Predicted Power*
Std. Err. of Observed < 0.0023
Canonical Correlation Matrix Indexed by ρ.
Target Power Implied by β_P

N	ρ	Target power	β_P	Observed power
15	1	0.2	1.1462073	0.185
		0.5	1.8694368	0.475
		0.8	2.5569624	0.785
	2	0.2	1.9913092	0.184
		0.5	3.3391332	0.474
		0.8	4.7253111	0.801
	3	0.2	4.1734892	0.187
		0.5	6.7293625	0.475
		0.8	9.1001311	0.778
	4	0.2	1.3911631	0.185
		0.5	2.2431209	0.472
		0.8	3.0333770	0.777
30	1	0.2	1.0500220	0.194
		0.5	1.6977031	0.490
		0.8	2.3044671	0.795
	2	0.2	1.8052506	0.195
		0.5	2.9544561	0.492
		0.8	4.0739127	0.805
	3	0.2	3.8424393	0.192
		0.5	6.1767619	0.486
		0.8	8.3284689	0.787
	4	0.2	1.2808131	0.196
		0.5	2.0589207	0.484
		0.8	2.7761563	0.786
60	1	0.2	1.0144494	0.196
		0.5	1.6346650	0.493
		0.8	2.2110201	0.796
	2	0.2	1.7368897	0.200
		0.5	2.8149719	0.497
		0.8	3.8364750	0.806
	3	0.2	3.7201038	0.198
		0.5	5.9770431	0.490
		0.8	8.0551883	0.796
	4	0.2	1.2400346	0.195
		0.5	1.9923477	0.493
		0.8	2.6850628	0.797

Table 5.4 *Predicted Power Coverage,*
Canonical Correlation Matrix Indexed by ρ.
$N_{\text{est}} = 30$ and $\nu_{\text{est}} = 27$ *Target Coverage level $= 0.95$*

N	ρ	Predicted Power	HLT	PBT	WILKS
15	1	0.2	0.958	0.979	0.967
		0.5	0.957	0.989	0.974
		0.8	0.957	0.994	0.980
	2	0.2	0.933	0.976	0.950
		0.5	0.929	0.999	0.965
		0.8	0.929	0.999	0.977
	3	0.2	0.971	0.979	0.977
		0.5	0.971	0.984	0.978
		0.8	0.972	0.986	0.980
	4	0.2	0.970	0.977	0.973
		0.5	0.971	0.984	0.978
		0.8	0.972	0.987	0.980
30	1	0.2	0.957	0.968	0.962
		0.5	0.956	0.978	0.966
		0.8	0.958	0.987	0.972
	2	0.2	0.930	0.949	0.938
		0.5	0.930	0.973	0.949
		0.8	0.929	0.9794	0.960
	3	0.2	0.970	0.974	0.972
		0.5	0.971	0.979	0.975
		0.8	0.972	0.982	0.977
	4	0.2	0.971	0.974	0.972
		0.5	0.971	0.979	0.975
		0.8	0.971	0.981	0.976
60	1	0.2	0.958	0.962	0.960
		0.5	0.957	0.968	0.963
		0.8	0.955	0.974	0.964
	2	0.2	0.932	0.941	0.936
		0.5	0.931	0.952	0.941
		0.8	0.930	0.967	0.947
	3	0.2	0.971	0.973	0.972
		0.5	0.972	0.976	0.974
		0.8	0.972	0.976	0.975
	4	0.2	0.971	0.973	0.972
		0.5	0.972	0.976	0.974
		0.8	0.972	0.977	0.975

Table 5.5 *HLT Observed Power Coverage,*
Std. Err. of Observed < 0.0023
Canonical Correlation Matrix Indexed by ρ.
Target Coverage level $= 0.95$

N	ρ	Target power	Obs. power	$N_{\text{est}} = 30$	$N_{\text{est}} = 60$
15	1	0.2	0.199	0.957	0.971
		0.5	0.510	0.962	0.977
		0.8	0.815	0.966	0.978
	2	0.2	0.199	0.931	0.942
		0.5	0.496	0.926	0.941
		0.8	0.798	0.928	0.943
	3	0.2	0.198	0.968	0.983
		0.5	0.513	0.977	0.990
		0.8	0.820	0.981	0.991
	4	0.2	0.196	0.962	0.978
		0.5	0.510	0.976	0.989
		0.8	0.818	0.980	0.991
30	1	0.2	0.200	0.957	0.973
		0.5	0.504	0.959	0.974
		0.8	0.805	0.960	0.975
	2	0.2	0.203	0.934	0.950
		0.5	0.501	0.930	0.946
		0.8	0.794	0.925	0.941
	3	0.2	0.199	0.968	0.984
		0.5	0.503	0.976	0.987
		0.8	0.804	0.974	0.987
	4	0.2	0.203	0.974	0.989
		0.5	0.502	0.972	0.985
		0.8	0.805	0.973	0.987
60	1	0.2	0.199	0.956	0.970
		0.5	0.500	0.957	0.974
		0.8	0.800	0.955	0.973
	2	0.2	0.203	0.936	0.949
		0.5	0.499	0.930	0.944
		0.8	0.799	0.929	0.945
	3	0.2	0.202	0.974	0.988
		0.5	0.499	0.971	0.985
		0.8	0.804	0.974	0.987
	4	0.2	0.199	0.970	0.984
		0.5	0.501	0.973	0.986
		0.8	0.803	0.973	0.987

Table 5.6 *PBT Observed Power Coverage level*
Std. Err. of Observed < 0.0023
Canonical Correlation Matrix Indexed by ρ.
Target Coverage level $= 0.95$

N	ρ	Target power	Observed power	$N_{\text{est}} = 30$	$N_{\text{est}} = 60$
15	1	0.2	0.183	0.947	0.943
		0.5	0.483	0.984	0.988
		0.8	0.803	0.994	0.996
	2	0.2	0.177	0.927	0.901
		0.5	0.506	0.999	0.999
		0.8	0.907	0.991	0.903
	3	0.2	0.187	0.957	0.963
		0.5	0.482	0.976	0.981
		0.8	0.783	0.981	0.985
	4	0.2	0.184	0.946	0.950
		0.5	0.479	0.975	0.980
		0.8	0.782	0.981	0.984
30	1	0.2	0.195	0.959	0.971
		0.5	0.498	0.977	0.984
		0.8	0.810	0.989	0.993
	2	0.2	0.194	0.939	0.950
		0.5	0.511	0.977	0.981
		0.8	0.856	0.989	0.974
	3	0.2	0.193	0.963	0.976
		0.5	0.489	0.973	0.983
		0.8	0.790	0.978	0.984
	4	0.2	0.196	0.968	0.981
		0.5	0.488	0.972	0.981
		0.8	0.789	0.977	0.984
60	1	0.2	0.197	0.957	0.970
		0.5	0.497	0.967	0.979
		0.8	0.804	0.975	0.986
	2	0.2	0.199	0.940	0.951
		0.5	0.505	0.958	0.964
		0.8	0.828	0.976	0.975
	3	0.2	0.198	0.970	0.984
		0.5	0.493	0.972	0.984
		0.8	0.797	0.976	0.986
	4	0.2	0.196	0.967	0.981
		0.5	0.496	0.974	0.985
		0.8	0.796	0.976	0.986

Table 5.7 *Wilks' Observed Power Coverage level*
Std. Err. of Observed < 0.0023
Canonical Correlation Matrix Indexed by ρ.
Target Coverage level = 0.95

N	ρ	Target power	Observed power	$N_{\text{est}} = 30$	$N_{\text{est}} = 60$
15	1	0.2	0.185	0.935	0.936
		0.5	0.475	0.957	0.966
		0.8	0.785	0.973	0.980
	2	0.2	0.184	0.914	0.908
		0.5	0.474	0.948	0.952
		0.8	0.801	0.977	0.984
	3	0.2	0.187	0.951	0.959
		0.5	0.475	0.962	0.972
		0.8	0.778	0.969	0.978
	4	0.2	0.185	0.942	0.950
		0.5	0.472	0.961	0.967
		0.8	0.777	0.969	0.975
30	1	0.2	0.194	0.951	0.964
		0.5	0.490	0.960	0.972
		0.8	0.795	0.970	0.980
	2	0.2	0.195	0.930	0.943
		0.5	0.492	0.944	0.955
		0.8	0.805	0.962	0.971
	3	0.2	0.192	0.958	0.973
		0.5	0.486	0.966	0.980
		0.8	0.787	0.971	0.980
	4	0.2	0.196	0.967	0.980
		0.5	0.484	0.965	0.977
		0.8	0.786	0.968	0.980
60	1	0.2	0.196	0.952	0.966
		0.5	0.493	0.959	0.973
		0.8	0.796	0.961	0.976
	2	0.2	0.200	0.936	0.949
		0.5	0.497	0.939	0.951
		0.8	0.806	0.950	0.961
	3	0.2	0.198	0.970	0.984
		0.5	0.490	0.968	0.981
		0.8	0.796	0.973	0.984
	4	0.2	0.195	0.964	0.978
		0.5	0.493	0.970	0.983
		0.8	0.797	0.973	0.985

Chapter 6 LOOKING FORWARD, LOOKING BACK

6.1 Looking Back

The aims of this dissertation were introduced at the beginning, which were to find better ways to deal with uncertainty and bias in power calculations for MULTIREP tests. A summary of what has been achieved is described below, along with the strengths and limitations of the research. Moreover, the results show the way to interesting questions and solution strategies for future work.

6.1.1 Chapter 3 (Paper 1) The Confidence Interval Due to Estimated Covariance for Power of a One or Two Group Test, and Related Models

To do a power analysis, we often need to estimate the nuisance parameter, error covariance. An estimate from a prior similar study may be used. The uncertainty has been a big barrier to computing an accurate power and sample size. A confidence interval has been used to show the uncertainty of the estimation. Taylor and Muller (1995) described a confidence interval for power in the univariate linear model to deal with the uncertainty of estimating the variance. Their theory extends exactly to the multivariate linear model in the special case, $s = \min(a, b) = 1$, by creating an equivalent univariate linear model. The application of the confidence interval for the univariate model was used in the transformed equivalent model. Interpretation is simplified by recognizing canonical representations of the parameter matrices based on the fact that $\text{rank}(\boldsymbol{B}) = \text{rank}(\boldsymbol{\Delta}) = 1$ for $s = 1$.

6.1.2 (Paper 2) Internal Pilots for a One or Two Group Test, and Related Models

If a prior study is not comparable to the new study, the variance estimate from the prior study may be biased. For example, the target population of a new study may be for adults

though the prior study was for children. An internal pilot study may be a good solution to reduce the risk of inaccurate variance estimation and sample size choice.

A fraction of the observations in the target study is used to estimate the variance in an internal pilot design. Then the sample size is adjusted, based upon the new estimate. There is no interim data analysis in internal pilot designs. Most internal pilot work has been to chose a sample size to achieve a target power in hypothesis testing in the univariate model. The method derived in Chapter 3 can be used in internal pilot designs in the multivariate model special case of $s = 1$. Internal pilot designs are thereby extended to the multivariate linear model. Many exact results are available. Very conveniently, little or no new software is needed.

Tables were provided to show how the misspecification of γ, the ratio of the population error variance to the initial value used for planning, affects the value of power and sample size. Misspecification of the population error variance was shown to lead to potentially large differences in power. The need for control of test size by using adjusted tests was also illustrated.

6.1.3 (Paper 3) The Confidence Interval Due to Estimated Covariance for Power of a Gaussian Multivariate General Linear Model

If $s > 1$ then there does not exist a most powerful unbiased test in the multivariate model. Also distribution theory is more complicated which leads to using approximations for power and confidence intervals.

In the special case with $s > 1$ and $s_* = 1$, HLT and PBT approximate noncentrality is a 1–1 function of an inverted gamma and a specific known random variable, respectively. Hence, an approximate confidence interval for power can be computed.

Moreover, any combination of multivariate linear model and testable general linear hypothesis has an associated hypothesis equivalent model and hypothesis in which B is a

diagonal matrix which the diagonal element k is the square root of eigenvalue k of the original noncentrality matrix, $\omega_k^{1/2}$. The result helps simplify simulations.

Simulations were done with four patterns of canonical correlations. Simulation results illustrated that there is no most powerful unbiased test. Accuracy increased with increasing target sample size. Overall the approximations worked quite well. Anomalous results were found with accuracy decreasing with increasing estimate sample size. The source of the anomaly is unknown at this time.

6.2 Looking Forward

Research often brings forward more questions than were answered. All results derived in this dissertation are limited to the general linear multivariate with Gaussian errors and fixed means.

In Chapter 3 and 5 of this dissertation the univariate results of Taylor and Muller (1995) were extended to the multivariate model. They derived exact confidence intervals for noncentrality, power, and sample size based on a variance estimate with fixed means. However, Taylor and Muller (1996) considered the estimation of both means and variance in a power calculation. They also described the bias arising from conducting a study depending on the results of the previous study. Muller and Pasour (1997) described the same bias but with fixed means. Extending both papers to the multivariate model has great appeal. Moreover, considering extensions to data that are not Gaussian has great appeal for clinical trials.

Chapter 4 described the internal pilot design in special case of the multivariate model. It was an application of the univariate model with transformation. Therefore, the derivation of the results were restricted, not applicable to all multivariate linear models. Practically, the study design may be more complex. Hence, the internal pilot design for the general multivariate linear model looks very interesting as a future result. The simulation results for

confidence intervals for the general case are encouraging because the basic theory of the internal pilot depends directly on such results. The anomalous results need to be resolved.

Although hypothesis testing has been developed for MULTIREP models with missing data (Catellier and Muller, 2000), power has not been. In turn, confidence intervals for estimated power and internal pilots would be very appealing in the same setting.

Appendix A: Analytic Results

A.1 Transformation

As described in Theorem 3.1,a Gaussian multivariate linear model and associated testable hypothesis may be transformed to a hypothesis equivalent univariate linear model in $s = \min(a, b) = 1$ case. The following example illustrates the details of the process.

Step A. Specify (original) model A, $\boldsymbol{Y}_A = \boldsymbol{X}_A \boldsymbol{B}_A + \boldsymbol{E}_A$, α, $\boldsymbol{C}_A$, $\boldsymbol{U}_A$, and $\boldsymbol{\Theta}_{0,A}$. The design of the original tortuosity study with $\boldsymbol{X}_A = \begin{bmatrix} \mathbf{1}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_m \end{bmatrix}$ (with n and m varied, which indicates the sample size in benign and malign groups), $\boldsymbol{C}_A = \begin{bmatrix} 1 & -1 \end{bmatrix}$,

$$\boldsymbol{U}_A = \begin{bmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \boldsymbol{B}_A = \begin{bmatrix} \mu - \delta_{\text{R}} & \mu & \mu + \delta_{\text{R}} + \delta_{\text{GR}} & \mu \\ \mu + \delta_{\text{G}} - \delta_{\text{R}} & \mu + \delta_{\text{G}} & \mu + \delta_{\text{G}} + \delta_{\text{R}} & \mu + \delta_{\text{G}} \end{bmatrix},$$

$\boldsymbol{\Theta}_{A0} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$, and $\boldsymbol{\Sigma}_A = \begin{bmatrix} 0.0838 & 0.0502 & 0.0356 & 0.0533 \\ 0.0502 & 0.0537 & 0.0325 & 0.0333 \\ 0.0356 & 0.0325 & 0.0441 & 0.0386 \\ 0.0533 & 0.0333 & 0.0386 & 0.0722 \end{bmatrix}$. Here

$\boldsymbol{E}_A \sim \mathcal{N}_{N_A, p_A}(\mathbf{0}, \boldsymbol{I}_N, \boldsymbol{\Sigma}_A)$, $\boldsymbol{\Theta}_A = \boldsymbol{C}_A \boldsymbol{B}_A \boldsymbol{U}_A$, $\boldsymbol{\Sigma}_{*A} = \boldsymbol{U}_A' \boldsymbol{\Sigma}_A \boldsymbol{U}_A$, and

$\boldsymbol{M}_A = \boldsymbol{C}_A (\boldsymbol{X}_A' \boldsymbol{X}_A)^- \boldsymbol{C}_A'$. The unscaled noncentrality parameter is

$$\begin{aligned} \boldsymbol{\Delta}_A &= (\boldsymbol{\Theta}_A - \boldsymbol{\Theta}_{A0})' M_A^{-1} (\boldsymbol{\Theta}_A - \boldsymbol{\Theta}_{A0}) \\ &= (\boldsymbol{\Theta}_A - \boldsymbol{\Theta}_{A0})' [\boldsymbol{C}_A (\boldsymbol{X}_A' \boldsymbol{X}_A)^- \boldsymbol{C}_A']^{-1} (\boldsymbol{\Theta}_A - \boldsymbol{\Theta}_{A0}), \end{aligned}$$

and the noncentrality parameter is

$$\boldsymbol{\Omega}_A = (\boldsymbol{\Theta}_A - \boldsymbol{\Theta}_{0,A})' \boldsymbol{M}_A^{-1} (\boldsymbol{\Theta}_A - \boldsymbol{\Theta}_{0,A}) \boldsymbol{\Sigma}_{*A}^{-1}.$$

The rank of $\boldsymbol{X}_A$ and the trace of $\boldsymbol{\Omega}_A$ give the scalar noncentrality when $s = 1$. Here $a = 1$, which implies $s = 1$ and $s_* = \text{rank}(\boldsymbol{\Omega}_A) = 1$ under the alternative.

Step B. Transformed model B has identity $\boldsymbol{U}_B$ and zero $\boldsymbol{\Theta}_{0,B}$ matrices. That is $\boldsymbol{Y}_B = \boldsymbol{X}_B \boldsymbol{B}_B + \boldsymbol{E}_B$, α, such that $\boldsymbol{C}_B = \boldsymbol{C}_A$, $\boldsymbol{U}_B = \boldsymbol{I}_3$, and $\boldsymbol{\Theta}_{0,B} = \mathbf{0}$ (1×3). For model B,

$$Y_B = Y_A U_A - X_A C_A'(C_A C_A')^{-1}\Theta_{0,A} \quad (N_A \times 3)$$
$$X_B = X_A \quad (N_A \times 2)$$
$$B_B = B_A U_A - C_A'(C_A C_A')^{-1}\Theta_{0,A} \quad (2 \times 3)$$
$$\Sigma_B = U_A'\Sigma_A U_A = \Sigma_{*A} \quad (3 \times 3)$$
$$E_B = E_A U_A \sim \mathcal{N}_{N_A,3}(\mathbf{0}, I_{N_A}, \Sigma_B) \quad (N_A \times 3).$$

Also

$$C_B = C_A \quad (1 \times 2)$$
$$U_B = I_b \quad (3 \times 3)$$
$$\Theta_B = \Theta_A - \Theta_{0,A} \quad (1 \times 3)$$
$$\Sigma_{*B} = U_A'\Sigma_A U_A = \Sigma_B \quad (3 \times 3)$$
$$M_B = C_A(X_A'X_A)^- C_A' = M_A \quad (1 \times 1)$$
$$\Omega_B = (\Theta_A - \Theta_{0,A})' M_A^{-1}(\Theta_A - \Theta_{0,A})\Sigma_{*A}^{-1} = \Omega_A \quad (3 \times 3)$$
$$\mathrm{tr}(\Omega_B) = \mathrm{tr}(\Omega_A). \quad (1 \times 1).$$

Step C. Specify the transformed model C. Here $Y_C = X_C B_C + E_C$, such that

$C_C = [1 \quad 0]$, $U_C = I_3$, and $\Theta_{0,C.} = \mathbf{0}$ $(1 \times b)$. With $C_B = L_A D_A R_A' =$
$[L_{A1} \quad L_{A0}]\begin{bmatrix} D_A & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} R_{A1}' \\ R_{A0}' \end{bmatrix}$, $T = \begin{bmatrix} C_B \\ R_{A0}' \end{bmatrix}$ and $T^{-1} = \begin{bmatrix} C_B'(C_B C_B')^{-1} & R_{A0} \end{bmatrix}$. Also

$$Y_C = Y_B \quad (N_A \times 3)$$
$$X_C = X_B T^{-1} = X_B\begin{bmatrix} C_B'(C_B C_B')^{-1} & R_{A0} \end{bmatrix} \quad (N_A \times 2)$$
$$B_C = T B_B = [C_B' \quad R_{A0}]' B_B \quad (2 \times 3)$$
$$\Sigma_C = \Sigma_B \quad (3 \times 3)$$
$$E_C = E_B \sim \mathcal{N}_{N_A,b}(\mathbf{0}, I_{N_A}, \Sigma_C) \quad (N_A \times 3).$$

Furthermore

$$C_C = [1 \quad 0] \quad (1 \times 2)$$
$$U_C = I_3 \quad (3 \times 3)$$
$$\Theta_C = \Theta_B \quad (1 \times 3)$$
$$\Sigma_{*C} = \Sigma_B \quad (3 \times 3)$$
$$M_C = C_C(X_C'X_C)^- C_C' \quad (1 \times 1)$$
$$\Omega_C = \Theta_C' M_C^{-1}\Theta_C \Sigma_{*C}^{-1} \quad (3 \times 3)$$
$$\mathrm{tr}(\Omega_C) = \mathrm{tr}(\Omega_B) = \mathrm{tr}(\Omega_A)$$
$$\text{numerator df} = ab = 3$$
$$\text{denominator df} = N_A - r = N_A - 2.$$

Step D. Specify model D, $\boldsymbol{Y}_D = \boldsymbol{X}_D \boldsymbol{B}_D + \boldsymbol{E}_D$, α, such that $\boldsymbol{C}_D = 1$, $\boldsymbol{U}_D = \boldsymbol{I}_3$, and

$\boldsymbol{\Theta}_{0,D} = \boldsymbol{0}$ (1×3). For model D,

with $\boldsymbol{X}_C = [\begin{array}{cc} \boldsymbol{X}_{C1} & \boldsymbol{X}_{C2} \end{array}]$, $\boldsymbol{X}_C \begin{bmatrix} \boldsymbol{1}_{1\times 2} \\ \boldsymbol{0}_{1\times 2} \end{bmatrix} = \boldsymbol{X}_B \left[\boldsymbol{C}_B'(\boldsymbol{C}_B \boldsymbol{C}_B')^{-1} \right]$,

$\boldsymbol{X}_{C2} = \boldsymbol{X}_C \begin{bmatrix} \boldsymbol{0}_{1\times 2} \\ \boldsymbol{1}_{1\times 2} \end{bmatrix} = \boldsymbol{X}_B \boldsymbol{R}_{A0} = [\begin{array}{cc} \boldsymbol{L}_{C1} & \boldsymbol{L}_{C0} \end{array}] \begin{bmatrix} \boldsymbol{D}_{C2} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{R}_{C1}' \\ \boldsymbol{R}_{C0}' \end{bmatrix}$ and

$N_D = N_A - \text{rank}(\boldsymbol{X}_{C2}) = N_A - $ # of columns in $\boldsymbol{L}_{C1} = N_A - 1$, and

$$
\begin{aligned}
\boldsymbol{Y}_D &= \boldsymbol{L}_{C1}' \boldsymbol{Y}_C && (N_D \times 3) \\
\boldsymbol{X}_D &= \boldsymbol{L}_{C1}' \boldsymbol{X}_{C1} && (N_D \times 2) \\
\boldsymbol{B}_D &= \boldsymbol{C}_C \boldsymbol{B}_C && (N_D \times 1) \\
\boldsymbol{\Sigma}_D &= \boldsymbol{\Sigma}_C && (3 \times 3) \\
\boldsymbol{E}_D &= \boldsymbol{L}_{C1}' \boldsymbol{E}_C \sim \mathcal{N}_{N_D,3}(\boldsymbol{0}, \boldsymbol{L}_{C1}' \boldsymbol{L}_{C1}, \boldsymbol{\Sigma}_D) && (N_D \times 3).
\end{aligned}
$$

Also

$$
\begin{aligned}
\boldsymbol{C}_E &= 1 && (1 \times 1) \\
\boldsymbol{U}_E &= \boldsymbol{I}_3 && (3 \times 3) \\
\boldsymbol{\Theta}_E &= \boldsymbol{\Theta}_D && (1 \times 3) \\
\boldsymbol{\Sigma}_{*E} &= \boldsymbol{\Sigma}_E && (3 \times 3) \\
\boldsymbol{m}_E &= (\boldsymbol{X}_E' \boldsymbol{X}_E)^{-} && (1 \times 1) \\
\boldsymbol{\Omega}_E &= \boldsymbol{\Theta}_E' \boldsymbol{m}_E^{-1} \boldsymbol{\Theta}_E \boldsymbol{\Sigma}_{*E}^{-1} && (3 \times 3) \\
\text{tr}(\boldsymbol{\Omega}_E) &= \text{tr}(\boldsymbol{\Omega}_D) && (1 \times 1).
\end{aligned}
$$

Step F. The final model is $\boldsymbol{y}_F = \boldsymbol{X}_F \boldsymbol{\beta}_F + \boldsymbol{e}_F$ which has $n_F = N_A - 1$ rows, $\boldsymbol{e}_F \sim \mathcal{N}_{N_A-1}(\boldsymbol{0}, \boldsymbol{I})$, $\boldsymbol{X}_F' = [\begin{array}{cc} \boldsymbol{I}_3 & \boldsymbol{0} \end{array}]'$, 3×1, $\boldsymbol{\beta}_F = [\begin{array}{cc} \sqrt{\omega_E} & \boldsymbol{0} \end{array}]'$, $\boldsymbol{C}_F = \boldsymbol{I}_3$, $\boldsymbol{U}_F = \{1\}$, and $\boldsymbol{\theta}_{0F} = \boldsymbol{0}$ with $\omega_E = \text{tr}(\boldsymbol{\Omega}_E) = (\boldsymbol{x}_{E1}' \boldsymbol{x}_{E1}) \text{tr}(\boldsymbol{B}_E \boldsymbol{\Sigma}_E^{-1} \boldsymbol{B}_E')$.

A.2 SAS Code for Transformation

The following SAS IML code may be used to transform from the multivariate linear model to a hypothesis equivalent univariate linear model in the special case $s = 1$.

To transform a univariate model to the multivariate model, please note the following general principles.

1) Specify the TRANS and POWERLIB IML path on the top of your program.

2) The basic 5 inputs need to be specified, ESSENCE, BETA, SIGMA, and C. REPN (fractional REPN is also available) is needed here because it has a default value of 1.

Noncentrality will be printed in each step. It must remain exactly the same. Moreover, C, U, X, Σ and B will be provided in each steps. If your model has s which is greater than 1, then the process of transformation will be stopped because this program cannot deal with the case. A simple example will show how to use the program to calculate the power for the multivariate linear model.

Example. An essence matrix, $\mathrm{Es}(X)$, and a replication factor (REPN in the program) is needed to express the design matrix, $X = \mathrm{Es}(X) \otimes \mathbf{1}_{\mathrm{REPN}}$. For example, if there are two groups with 20 observations in each group, then $X = I_{32} \otimes \mathbf{1}_{20}$ with $\mathrm{Es}(X) = I_2$ and REPN $= 20$. Next, the following seven variables have to be provided: $\Sigma, X, B, C, U, \alpha$ and Θ_0. The following code transforms from the multivariate linear model to the a hypothesis equivalent univariate model and checks the invariance of noncentrality in every step of transformation.

```
PROC IML WORKSIZE=1000 SYMSIZE=2000;
%INCLUDE "&ROOT\Iml\POWERLIB203.IML"/NOSOURCE2;
%INCLUDE "&ROOT\Iml\NONCEN.IML"/NOSOURCE2;
%INCLUDE "&ROOT\Iml\TRANS.IML"/NOSOURCE2;


OPT_OFF = {ALPHA};       *Turn options off;
OPT_ON  = {NOPRINT FRACREPN}; *Turn options on;
_ZERO_=1E-12;


* INPUT SIGMA, C, U, BETA, THETA0, X *;
ALPHA = .05/6;
SIGMA = {
     0.0838     0.0502     0.0356     0.0533,
     0.0502     0.0537     0.0325     0.0333,
     0.0356     0.0325     0.0441     0.0386,
     0.0533     0.0333     0.0386     0.0722};
ESSENCEX = I(2);


N1=20;      /*TARGET SAMPLE SIZE*/
X=ESSENCEX@J(20,1,1);
BETA = {0 0 1 0,
```

```
        0 0 0 0};

    Q=NROW(BETA);
    P=NCOL(BETA);
    C = {1 -1};
    U= ( J(P-1,1,-1) || I(3) )` ;
    THETA0 = J(NROW(C),NCOL(U),0);
    RUN TRANS;
    CLOSE;
    QUIT;
```

The output will produce Σ, $\text{Es}(X)$, B, C, U and Ω in each step. The following output is for the final model.

```
                        C                        BETA        U
Model 5 (Final) ,       1        0       0 -5.768277          1

                        0        1       0 0.8443463

                        0        0       1 1.9243963

                     SIGMA                  ESSENCEX

                       1                  1        0        0

                                          0        1        0

                                          0        0        1




                    OMEGA_F              OMEGA_ALL
         Original   489.96011       A    489.96011

                                    B    489.96011

                                    C    489.96011

                                    D    489.96011

                                    E    489.96011

                                    F    489.96011
```

As shown in the above output, the final model is the univariate model and there is no loss of information to calculate the power because the noncentrality is invariant to the transformation. Power and an exact confidence interval of power for estimated covariance can be calculated using the POWERLIB program (or any univariate power program) for the univariate linear model.

```
PROC IML WORKSIZE=1000 SYMSIZE=2000;
%INCLUDE "&ROOT\IML\POWERLIB203.IML"/NOSOURCE2;

OPT_OFF = {ALPHA};      *Turn options off;
OPT_ON  = {FRACREPN}; *Turn options on;
```

```
_ZERO_=1E-12;


* INPUT SIGMA, C, U, BETA, THETA0, X *;

ALPHA = .05/6;


    BETA={-5.768277,0.8443463,1.9243963};

    U = {1};

    C = I(B);

    SIGMA = 1;

    A=NROW(C);

    B=NCOL(U);

    N=40;

    R=B;

    ESSENCEX = I(B);

    REPN=(N-R+A)/B;

    X=ESSENCEX@J(REPN,1,1);


    BETASCAL = {0.3}#DO(0,2.0, 0.25);


*Statements to create confidence limits;


CLTYPE=1;

N_EST=20;          *# Obs for variance estimate;

RANK_EST=1;        *# model df for study giving variance estimate;

ALPHA_CL=.025;    *Lower confidence limit tail size;

ALPHA_CU=.025;    *Upper confidence limit tail size;

RUN POWER;

QUIT;
```

SAS produces the following output, with results exactly same as that of the multivariate power analysis for the original model.

HOLDPOWER

SIGSCAL	BETASCAL	TOTAL_N	CLTYPE	ALPHA_CL	ALPHA_CU	POWER_L	POWER	POWER_U
1	0	39	1	0.025	0.025	0.008	0.008	0.008
1	0.075	39	1	0.025	0.025	0.031	0.073	0.152
1	0.15	39	1	0.025	0.025	0.171	0.479	0.807
1	0.225	39	1	0.025	0.025	0.51	0.921	0.997
1	0.3	39	1	0.025	0.025	0.847	0.998	1
1	0.375	39	1	0.025	0.025	0.979	1	1
1	0.45	39	1	0.025	0.025	0.999	1	1
1	0.525	39	1	0.025	0.025	1	1	1
1	0.6	39	1	0.025	0.025	1	1	1

The SAS IML code for transformation is below.

```
START TRANS;

CALL SVD(LX0,DX0,RX0, X);

r=SUM(DX0>_ZERO_);  /*RANK, C*/

a = NROW(C); *# of rows, C *;

b = NCOL(U); *# of cols, C *;

s = MIN(a, b);

N = NROW(X);        /*NUMBER OF ROW, X*/


/* STEP A */

PRINT "Original, Model 0", C BETA U THETA0 ,

a b s N "   " SIGMA ,

X;


****Compute original noncentrality, for model A;

OMEGA_A=NONCEN(X,C,BETA,U,THETA0,SIGMA);

PRINT "Original  " OMEGA_A;

OMEGA_ALL=OMEGA_ALL//OMEGA_A;


IF a>{1} THEN DO;

      PRINT "stopping due to a>1 not handled here";

      STOP;

      END;

****************************************************************;

/* STEP B */

* IF U ^= I AND THETA0 ^=0 *;

* Y = XB + E ==> Y1 = X1B1 + E1, row(E1)~ N(0,SIGMA1) *;

* WITH Y1 = Y*U - C`*INV(CC`)*THETA0 and SIGMA1 = U`*SIGMA*U;*;
```

```
BETA_B = BETA*U-C`*INV(C*C`)*THETA0;

SIGMA_B = U`*SIGMA*U;

C_B = C;

U_B = I(b);

X_B = X;

THETA0_B = J(NROW(C_B),NCOL(U_B),0);


PRINT / "Model 1 , " C_B BETA_B U_B,

         SIGMA_B , X_B;

OMEGA_B= NONCEN(X_B,C_B,BETA_B,U_B,THETA0_B,SIGMA_B);

PRINT "Original  " OMEGA_B;

OMEGA_ALL=OMEGA_ALL//OMEGA_B;

**************************************************************;

/* STEP C */

* CHANGE FROM (axb) C_A TO C_B =[I(a) 0] *;

* C1 = [L1 L0]D[R1 R0]` = L1*D1*R1` *;

* Y1 = XB1 + E1 ==> Y2 = X2B2 + E2 row(E2)~ N(0,SIGMA2) *;

* WITH X2 = [X21 X22] = [XC1`*INV(C1C1`) XR0], B2 = [C1*B1 R0*B1]`*;

* Y2 = Y1, SIGMA2 = SIGMA1 *;

IF NROW(C_B)^=NCOL(C_B) THEN DO;

CALL SVD(L_B,D_B,R_B,C_B);/* L1: axb, D1: bx1, R = [R1 R0]: bxb */


RANK_C_B=SUM(D_B>_ZERO_);

IF NCOL(R_B)>=1 THEN R0_B=R_B[,NCOL(R_B)-RANK_C_B+1:NCOL(R_B)];

ELSE IF NCOL_R0_B=0 THEN PRINT "RO=0";


T=C1//R0_B`;

INV_T=C_B`*INV(C_B*C_B`)||R0_B;

X_C = X_B*INV_T;

X_C1=X_B*C_B`*INV(C_B*C_B`);

X_C2=X_B*R0_B;


BETA_C=(C_B*BETA_B)//(R0_B`*BETA_B);

C_C = I(a)||J(a,NROW(BETA_C)-a,0);

U_C = U_B;

SIGMA_C=SIGMA_B;

THETA0_C = J(NROW(C_C),NCOL(U_C),0);


PRINT / "Model 2 , " C_C BETA_C U_C,

         SIGMA_C , X_C;

OMEGA_C= NONCEN(X_C,C_C,BETA_C,U_C,THETA0_C,SIGMA_C);

PRINT "Original  " OMEGA_C;

OMEGA_ALL=OMEGA_ALL//OMEGA_C;

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

```
/* STEP D */

* REDUCE # OF ROWS X and Y, from N to N-rank(X)+a*;

* Y2 = X2B2 + E2 ==> Y3 = X3B3 + E3, row(E2)~ N(0,SIGMA3) *;

* WITH X22 = LDR`=L2*D2*R2` where L=[L2 L0], B3 = C2B2, X3=L0`*X21*;

* Y3 = L0`*Y2, SIGMA3 = SIGMA2 *;


CALL SVD(L_C2,D_C2,R_C2,X_C2);


NROW_X_C2=NROW(X_C2);

LTL=I(NROW_X_C2)-L_C2*L_C2`;


CALL EIGEN(EVAL, EVEC, LTL);

NONZERO= (EVAL>_ZERO_);

RANK_LTL=SUM(NONZERO);


CALL SVD(L_LTL,D_LTL,R_LTL,LTL);

D4=DIAG(D_LTL);

D=D4[,1:RANK_LTL];

L0=L_LTL*D;


CALL SVD(LX0,DX0,RX0, X);

NEW_NROW=N-r+a;

LPL=I(N)-L_C2*L_C2`;

CALL SVD(L20,D20,R20,LPL);

D=DIAG(D20);

DL0=D[,1:NEW_NROW];

L_D0=L20*DL0;


X_D=L_D0`*X_C1;

BETA_D=C_C*BETA_C;

SIGMA_D = SIGMA_C;


C_D = I(a);

U_D = U_C;

THETA0_D = J(NROW(C_D),NCOL(U_D),0);


PRINT / "Model 3 , " C_D BETA_D U_D,

        SIGMA_D , X_D;

OMEGA_D= NONCEN(X_D,C_D,BETA_D,U_D,THETA0_D,SIGMA_D);

PRINT "Original  " OMEGA_D;

OMEGA_ALL=OMEGA_ALL//OMEGA_D;

END;


IF C_B={1} THEN DO;

X_D=X_B;
```

76

```
BETA_D=BETA_B;

SIGMA_D = SIGMA_B;


C_D = C_B;

U_D = U_B;

THETA0_D = THETA0_B;

END;

*****************************************************************;

/* STEP E for a=1*/

*****Transform to unvariate model;

* Define inputs to power program;

SIGMA_DINV=INV(SIGMA_D);

UPPER=HALF(SIGMA_DINV);

NX_D=NROW(X_D);

ESSENCEX_E = UPPER//J( (NROW(X_D)-b) ,b,0);

REPN = {1};

X_E=ESSENCEX_E;

BETA_E = BETA_D`;

U_E = C_D;

C_E = U_D;

THETA0_E = J(NROW(BETA_E),1,0);

SIGMA_E = {1}/(X_D`*X_D);


PRINT / "Model 4  , " C_E BETA_E U_E,

          SIGMA_E , ESSENCEX_E;


OMEGA_E= NONCEN(X_E,C_E,BETA_E,U_E,THETA0_E,SIGMA_E);

PRINT "Original  " OMEGA_E;

OMEGA_ALL=OMEGA_ALL//OMEGA_E;

*****************************************************************;

M=X_E`*X_E;

CALL SVD(LM,DM,RM,M);

DM2=DIAG(DM);

PSIM=LM*(SQRT(DM2));

M2M=PSIM*PSIM`;

ESSENCEX = I(b);

REPN = (N-R+A)/B;

X=ESSENCEX@J(REPN,1,1);

NXX=NROW(X);

MMX=X`*X;

K=MMX[1,1];

BETA = PSIM`*BETA_E/SQRT(SIGMA_E*K);

BETAREPN = PSIM`*BETA_E/SQRT(SIGMA_E*REPN);

U = U_E;

C = C_E;
```

```
THETA0 = J(NROW(BETA),1,0);

SIGMA = 1;


PRINT / "Model 5 (Final) , " C BETA U,

        SIGMA , ESSENCEX;


OMEGA_F= NONCEN(X,C,BETA,U,THETA0,SIGMA);

PRINT "Original  " OMEGA_F;


OMEGA_ALL=OMEGA_ALL//OMEGA_F;

ROWNM={A B C D E F};

PRINT OMEGA_ALL[ROWNAME=ROWNM FORMAT=10.5];

SIGMA_INV=INV(SIGMA);

OMEGADALSO=TRACE(BETA`*(X`*X)*BETA*SIGMA_INV);

PRINT OMEGADALSO;

FINISH;
```

Appendix B: Code for Chapter 4

This code is to run the internal pilot using the free SAS code for internal pilots.

```
PROC IML WORKSIZE=1000 SYMSIZE=300;
 RESET FUZZ FW=5;
 %INCLUDE "&ROOT\IML\GLUMIP20.IML" / NOSOURCE2;
USE INOUT.P0301;                                      *Data after transformation*;
READ ALL VAR{N R A B BETA1 BETA2 BETA3} INTO T;
CLOSE INOUT.P0301;
DO I=2 TO 6 BY 2 ;
 ESSENCEX = I(3);
 ALPHAT   = .05;
 POWERT   = .90;
 C        = I(3);
 BETASCAL = {0.3}#DO(0,2.0, 0.25);
   DO J=1 TO NCOL(BETASCAL);
   DELTA=BETASCAL[1,J];
   BETA = T[I,5:7];
   BETA_PLN = DELTA#T[I,5:7]`;

 SIGMA0   = 1;
 N1       = T[I,1];
 GAMLIST  ={.5 .75 1 1.5 2};
 BETA_ALT = DELTA#T[I,5:7]`;
 TEST     = 0;
 RULE     = 0;
 NPLUSMIN = 48;
 RUN GLUMIP;
 PRINT DELTA BETA;
 PRINT _IPCALCS[COLNAME=_IPNAMES];
 END;
 END;
QUIT;
```

The results indicate the expected sample size and power with $\delta_{\text{GR}} = 0.16$.

	ALPHAT	CRITVAL	POWERT	GAMMA	N1	NPLUSMIN	NPLUSMAX	RULE	TEST	E(N)	POWER
ROW1	0.05	0.05	0.9	0.5	36	36	108	0	0	36.2	0.05
ROW2	0.05	0.05	0.9	0.75	36	36	108	0	0	40.2	0.053
ROW3	0.05	0.05	0.9	1	36	36	108	0	0	49.2	0.056
ROW4	0.05	0.05	0.9	1.5	36	36	108	0	0	70.3	0.056
ROW5	0.05	0.05	0.9	2	36	36	108	0	0	88.9	0.054

	ALPHAT	CRITVAL	POWERT	GAMMA	N1	NPLUSMIN	NPLUSMAX	RULE	TEST	E(N)	POWER
ROW1	0.05	0.05	0.9	0.5	72	72	I	0	0	72	0.05
ROW2	0.05	0.05	0.9	0.75	72	72	I	0	0	72	0.05
ROW3	0.05	0.05	0.9	1	72	72	I	0	0	72	0.05
ROW4	0.05	0.05	0.9	1.5	72	72	I	0	0	75.7	0.051
ROW5	0.05	0.05	0.9	2	72	72	I	0	0	92.5	0.053

The results indicate the choice of π affect test size.

1) $\min(N_{F+}) = n_0 = 72$ and $n_1 = 12$

	ALPHAT	CRITVAL	POWERT	GAMMA	N1	NPLUSMIN	NPLUSMAX	RULE	TEST	E(N)	POWER
ROW1	0.05	0.05	0.9	0.5	12	72	I	0	0	72	0.05
ROW2	0.05	0.05	0.9	0.75	12	72	I	0	0	72.3	0.05
ROW3	0.05	0.05	0.9	1	12	72	I	0	0	74	0.05
ROW4	0.05	0.05	0.9	1.5	12	72	I	0	0	83.3	0.051
ROW5	0.05	0.05	0.9	2	12	72	I	0	0	98.8	0.052

2) $\min(N_{F+}) = n_0 = 72$ and $n_1 = 24$

	ALPHAT	CRITVAL	POWERT	GAMMA	N1	NPLUSMIN	NPLUSMAX	RULE	TEST	E(N)	POWER
ROW1	0.05	0.05	0.9	0.5	24	72	I	0	0	72	0.05
ROW2	0.05	0.05	0.9	0.75	24	72	I	0	0	72	0.05
ROW3	0.05	0.05	0.9	1	24	72	I	0	0	72.4	0.05
ROW4	0.05	0.05	0.9	1.5	24	72	I	0	0	79.2	0.051
ROW5	0.05	0.05	0.9	2	24	72	I	0	0	94.9	0.053

3) $\min(N_{F+}) = n_0 = 72$ and $n_1 = 36$

	ALPHAT	CRITVAL	POWERT	GAMMA	N1	NPLUSMIN	NPLUSMAX	RULE	TEST	E(N)	POWER
ROW1	0.05	0.05	0.9	0.5	36	72	I	0	0	72	0.05
ROW2	0.05	0.05	0.9	0.75	36	72	I	0	0	72	0.05
ROW3	0.05	0.05	0.9	1	36	72	I	0	0	72.1	0.05
ROW4	0.05	0.05	0.9	1.5	36	72	I	0	0	77.6	0.051
ROW5	0.05	0.05	0.9	2	36	72	I	0	0	93.6	0.053

4) $\min(N_{F+}) = n_1 = 12$

	ALPHAT	CRITVAL	POWERT	GAMMA	N1	NPLUSMIN	NPLUSMAX	RULE	TEST	E(N)	POWER
ROW1	0.05	0.05	0.9	0.5	12	12	84	0	0	27.3	0.069
ROW2	0.05	0.05	0.9	0.75	12	12	84	0	0	38	0.066
ROW3	0.05	0.05	0.9	1	12	12	84	0	0	48	0.062
ROW4	0.05	0.05	0.9	1.5	12	12	84	0	0	63.3	0.057
ROW5	0.05	0.05	0.9	2	12	12	84	0	0	72.3	0.055

5) $\min(N_{F+}) = n_1 = 24$

	ALPHAT	CRITVAL	POWERT	GAMMA	N1	NPLUSMIN	NPLUSMAX	RULE	TEST	E(N)	POWER
ROW1	0.05	0.05	0.9	0.5	24	24	96	0	0	28.4	0.054
ROW2	0.05	0.05	0.9	0.75	24	24	96	0	0	38.2	0.06
ROW3	0.05	0.05	0.9	1	24	24	96	0	0	48.9	0.06
ROW4	0.05	0.05	0.9	1.5	24	24	96	0	0	69.1	0.057
ROW5	0.05	0.05	0.9	2	24	24	96	0	0	83.2	0.054

6) $\min(N_{F+}) = n_1 = 36$

	ALPHAT	CRITVAL	POWERT	GAMMA	N1	NPLUSMIN	NPLUSMAX	RULE	TEST	E(N)	POWER
ROW1	0.05	0.05	0.9	0.5	36	36	108	0	0	36.2	0.05
ROW2	0.05	0.05	0.9	0.75	36	36	108	0	0	40.2	0.053
ROW3	0.05	0.05	0.9	1	36	36	108	0	0	49.2	0.056
ROW4	0.05	0.05	0.9	1.5	36	36	108	0	0	70.3	0.056
ROW5	0.05	0.05	0.9	2	36	36	108	0	0	88.9	0.054

Appendix C: Code for Chapter 5

1) Code1: Creates betascale, β_P.

```
PROC IML  WORKSPACE=2000 SYMSIZE=4000;

&LINMOD;

%INCLUDE "&ROOT\IML\POWERLIB21.IML"/  SOURCE2;

*INCLUDE "&ROOT\IML\QLIB01.IML"  /NOSOURCE2;

%INCLUDE "&ROOT\IML\ARLIB1.IML"  /NOSOURCE2;

%INCLUDE "&ROOT\IML\BETASCALE0203.IML" /NOSOURCE2;


*POWERLIB inputs to chose target predicted power methods;

ALPHA=&ALPHA;


OPT_ON = {NOPRINT HLT};

OPT_OFF= {GG UN HF PBT WLK   SIGSCAL RHOSCAL ALPHA TOTAL_N MAXRHOSQ
            COLLAPSE    WARN };

ROUND=10; *DEFAULT IS 3;

SIGSCAL={1};

RHOSCAL={1};

*BETASCL1 inputs below;

DIFFOK=10E-9;  *Difference tolerated, |target - achieved pwr|;

*TARGET=   is varied below in DO loop;

DEBUG="NO";  *DEBUG={"YES"};


ESSENCEX=I(3);

REPNLIST={5 10 20};*# subjects per group in a balanced design;


DO INREPN = 1 TO NCOL(REPNLIST);

   REPN = REPNLIST[,INREPN];

Q=NCOL(ESSENCEX);


RHOSQLIST1={.7 .4 0 0};

DLIST1=SQRT(RHOSQLIST1/(1-RHOSQLIST1)/REPN);

RHOSQLIST2={.5 .03 .01 .001};

DLIST2=SQRT(RHOSQLIST2/(1-RHOSQLIST2)/REPN);

RHOSQLIST3={.1 .1 .1 .1};

DLIST3=SQRT(RHOSQLIST3/(1-RHOSQLIST3)/REPN);

RHOSQLIST4={.5 .5 0 0};

DLIST4=SQRT(RHOSQLIST4/(1-RHOSQLIST4)/REPN);


RHOSQLIST=RHOSQLIST1//RHOSQLIST2//RHOSQLIST3//RHOSQLIST4;

DLIST=DLIST1//DLIST2//DLIST3//DLIST4;

P=4;

Q=3;
```

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

N=REPN#NROW(ESSENCEX);*Total # of subjects;

C=I(Q-1)||J(Q-1,1,0);

U=I(P-1)//J(1,P-1,0);

SIGMA=I(P);

BETA=J(Q,P,0);

DMAX=MIN(NCOL(DLIST),P,Q);

DROW=NROW(DLIST);
 DO IN2=1 TO DROW;
  DO D=1 TO DMAX;
  BETA[D,D]=DLIST[IN2,D];
  END;
     DO TARGET=.20 TO .80 BY .30; *target power;
      *RUN BETASCL1;  *Creates OUTSCAL;
    CALL _BETASCL(OUTSCAL);
     C1=INREPN+2;
     C2=IN2+2;
     C3=0;
    HOLD=HOLD // (C1||C2||C3||N||P||Q||TARGET||OUTSCAL);
     END;
   END;
END;

*PRINT HOLD;

HOLDNM={C1  C2  C3  N  P  Q  TARGET  BETASCAL};

*CREATE &JOB VAR HOLDNM;

CREATE OUT01.&JOB VAR HOLDNM;

APPEND FROM HOLD;
```

2) Code 2: HLT Oberserved Power Coverage Level, Target Coverage = .95

```
%LET JOB = P1001;

%LET ALPHA = .05;


TITLE1 "OBSERVED POWER CONFIDENCE LIMITS CALCULATIONS (TWO-SIDED), HLT S>1 N_EST=60";

TITLE2 "TARGET COVERAGE = %SYSEVALF(1-&ALPHA)";


%LET ROOT = E:\THESIS\SolaCI;


LIBNAME  INOUT01 "&ROOT\PROG\SIM";


%LET LMDIRECT = &ROOT\IML\ ;          *LINMOD VERSION 3.3;

%INCLUDE "&LMDIRECT.MACROLIB.MAC" /NOSOURCE2;


PROC IML  WORKSPACE=2000 SYMSIZE=4000;

&LINMOD;

%INCLUDE "&ROOT\IML\POWERLIB21.IML"/  SOURCE2;

*INCLUDE "&ROOT\IML\QLIB01.IML"  /NOSOURCE2;

%INCLUDE "&ROOT\IML\ARLIB1.IML"  /NOSOURCE2;

%INCLUDE "&ROOT\IML\BETASCALE0203.IML" /NOSOURCE2;

%INCLUDE "&ROOT\IML\GAUSS01.IML" / NOSOURCE2;


***COMPUTE CONFIDENCE LIMITS FOR POWER FOR CLAHE EXAMPLES IN MEST***;


****SPECIFY NUMBER OF POWERS TO COMPUTE;

NREP=1;

SEED=43421;

ALPHA=&ALPHA;


****READ IN CONDITIONS FOR BETASCAL, N INTO HOLDIN MATRIX;

USE INOUT01.P0701;

READ ALL   VAR _ALL_   INTO HOLDIN[COLNAME =INNM];

CLOSE INOUT01.P0701;

HLTNM = "HLT";

USE INOUT01.HLT;

READ ALL   VAR _ALL_   INTO HLT[COLNAME =HLTNM];

CLOSE INOUT01.HLT;


PRINT HOLDIN[COLNAME =INNM]

HLT[COLNAME =HLTNM];


****SPECIFY ALPHA, THETA, U, C, BETA;

ALPHA=&ALPHA;
```

```
P=4;
Q=3;


C=I(Q-1)||J(Q-1,1,0);
U=I(P-1)//J(1,P-1,0);



****SPECIFY ADDITIONAL INPUTS TO POWERLIB;
ROUND=10;


REPN=1;
SIGSCAL=1;
RHOSCAL=1;
TESTSON = {HLT};
TESTSOFF = {GG HF UN  BOX  WLK PBT};


OUTDATALBL = "POWER_L" || "POWER_U";



  DO TESTIND=1 TO 1 BY 1;


  OPT_ON =  TESTSON||{NOPRINT };
  OPT_OFF = TESTSOFF || {  ALPHA SIGSCAL WARN TOTAL_N BETASCAL};


    DO ICASE=1 TO 36 BY 1;    ***36 BY 1;
     ****CREATE BETASCAL, N, ESSENCEX;
    BETASCAL=HOLDIN[ICASE,8];
    *BETASCAL=1;
    N=HOLDIN[ICASE,4];
    N1=N/Q;
    OFFSET=MIN(HOLDIN[*,2]);
    ESSENCEX=I(Q)@J(N1,1,1);
    SIGMA=I(P);
    BETA=J(Q,P,0);
    DMAX=MIN(NCOL(DLIST),P,Q);
    DROW=NROW(DLIST);

    ****CREATE BETA;
    RHOSQLIST1={.7 .4 0 0};
    DLIST1=SQRT(RHOSQLIST1/(1-RHOSQLIST1)/N1);
    RHOSQLIST2={.5 .03 .01 .001};
    DLIST2=SQRT(RHOSQLIST2/(1-RHOSQLIST2)/N1);
    RHOSQLIST3={.1 .1 .1 .1};
    DLIST3=SQRT(RHOSQLIST3/(1-RHOSQLIST3)/N1);
    RHOSQLIST4={.5 .5 0 0};
```

85

```
DLIST4=SQRT(RHOSQLIST4/(1-RHOSQLIST4)/N1);


RHOSQLIST=RHOSQLIST1//RHOSQLIST2//RHOSQLIST3//RHOSQLIST4;

DLIST=DLIST1//DLIST2//DLIST3//DLIST4;

DLIST4=SQRT(Q#RHOSQLIST4/(1-RHOSQLIST4));


IVAR=HOLDIN[ICASE,2]-OFFSET+1;

BETA1=DIAG(DLIST[IVAR,*]);

BETA=BETA1[1:Q,];

****COMPUTE POPULATION POWER;

THETA=C*BETA*U;

SIGMASTAR=U`*SIGMA*U;


*RUN POWER;

MESTPOWER=HLT[ICASE,1];;

*PRINT "A:HLT - MESTPOWER" MESTPOWER;

****INITIALIZE COVERAGE COUNTER;

COUNTCOVER=J(3,1,0); *cover MEST power, low, in interval, high;


****LOOP TO SIMULATE SIGMAHAT AND COMPUTE POWER;

DEBUG=0;

CLTYPE=1;

ALPHA_CL= (ALPHA/2);

ALPHA_CU= (ALPHA/2);


N_EST=60;*training sample;

N_ESTG=N_EST/Q;

RANK_EST=2; *training sample;

NU_EST = N_EST - RANK_EST;


MUMATEST=J(N_EST,NCOL(SIGMA),0);

X1=I(Q) @ J(N_ESTG,1,1);

XPX1=X1`*X1;

IH1=I(N_EST)-X1*INV(XPX1)*X1`;

*RESET FUZZ;

F_SIGMA=I(P);


        DO REPLICAT=1 TO NREP BY 1;

        Y=GAUSS1(N_EST,MUMATEST,F_SIGMA,SEED);

        SIGMAHAT=Y`*IH1*Y/NU_EST;

        SIGMA=SIGMAHAT;

        FREE _HOLDPOWER;

        RUN POWER;


        POWER_L=_HOLDPOWER[1,4];
```

```
            POWER_U=_HOLDPOWER[1,6];


        PRINT / _HOLDPOWER[COLNAME=_HOLDPOWERLBL]
              ICASE POWER_L POWER_U;


        IF (MESTPOWER > POWER_U) THEN
              COUNTCOVER[1,1]=COUNTCOVER[1,1]+{1};
     IF (POWER_L<=MESTPOWER)&(MESTPOWER<=POWER_U) THEN
        COUNTCOVER[2,1]=COUNTCOVER[2,1]+{1};
        IF (MESTPOWER < POWER_L) THEN
              COUNTCOVER[3,1]=COUNTCOVER[3,1]+{1};


        END; *REPLICAT;


  COVER=COUNTCOVER/NREP;


  OUTMATROW = TESTSON || COMPRESS( CHAR(N  ||  MESTPOWER ||  COVER` ) );
  IF NROW(OUTMAT)=0 THEN OUTMAT=OUTMATROW;
  ELSE OUTMAT = OUTMAT // OUTMATROW;


  FREE N_EST RANK_EST CLTYPE _HOLDPOWER;


        END; *ICASE;


  END; *TESTIND;


OUTMATLBL = {"MTEST"  "N"  "OBSPOWER" "COVER_L" "COVER"  "COVER_U"} ;
OUTMATNM = {MTEST  N  POPPOWER COVER_L COVER  COVER_U} ;
PRINT OUTMAT[COLNAME=OUTMATLBL];
CREATE INOUT01.&JOB FROM  OUTMAT [COLNAME=OUTMATLBL];
APPEND FROM OUTMAT;


QUIT;
```

3) Code 3: HLT Predicted Power Coverage level, Target Coverage = .95

```
%LET PROG = P0706;

%LET ALPHA = .05;

%LET POWER = .90;


TITLE1 "POWER CONFIDENCE LIMITS CALCULATIONS (TWO-SIDED), HLT S>1";

TITLE2 "TARGET COVERAGE = &POWER";


*%LET ROOT = C:\THESIS\SolaCI;

%LET ROOT = E:\THESIS\SolaCI;


LIBNAME  INOUT01 "&ROOT\DATA";


%LET LMDIRECT = &ROOT\IML\ ;          *LINMOD VERSION 3.3;

%INCLUDE "&LMDIRECT.MACROLIB.MAC" /NOSOURCE2;


PROC IML  WORKSPACE=2000 SYMSIZE=4000;

&LINMOD;

OPT_ON={MULTTEST NOPRINT };

OPT_OFF={CHKMISS MPARMS MSS

        BETA UNIBETA EXBETA SIGMA SCORR LINDEP PARMOUT

        C U THETA0 THETA EXTHETA UNITHETA ECORR CANVEC CANRSQ EVEC2

        UNIREP RSQUARED UNIRPRNT};

RUN SETOPT;

FREE OPT_ON OPT_OFF;

DISPLAY=1;

%INCLUDE "&ROOT\IML\POWERLIB21.IML"/  NOSOURCE2;

*INCLUDE "&ROOT\IML\QLIB01.IML"  /NOSOURCE2;

%INCLUDE "&ROOT\IML\ARLIB1.IML"  /NOSOURCE2;

%INCLUDE "&ROOT\IML\BETASCALE0203.IML" /NOSOURCE2;

%INCLUDE "&ROOT\IML\GAUSS01.IML" / NOSOURCE2;


***COMPUTE CONFIDENCE LIMITS FOR POWER FOR CLAHE EXAMPLES IN MEST***;


****SPECIFY NUMBER OF POWERS TO COMPUTE;

NREP=50000;

SEED=43421;

ALPHA=&ALPHA;


****READ IN CONDITIONS FOR BETASCAL, N INTO HOLDIN MATRIX;

USE INOUT01.P0705;

READ ALL  VAR _ALL_  INTO HOLDIN[COLNAME =INNM];

CLOSE INOUT01.P0705;
```

```
*SIZE=NCOL(HOLDIN);
*PRINT SIZE;
*PRINT HOLDIN[COLNAME=INNM];


****SPECIFY ALPHA, THETA, U, C, BETA;
ALPHA=&ALPHA;
P=4;
Q=3;


C=I(Q-1)||J(Q-1,1,0);
U=I(P-1)//J(1,P-1,0);



****SPECIFY ADDITIONAL INPUTS TO POWERLIB;
ROUND=10;


REPN=1;
SIGSCAL=1;
RHOSCAL=1;
TESTSON = {HLT};
TESTSOFF = {GG HF UN  BOX  PBT WLK};


  DO TESTIND=1 TO 1 BY 1;


  OPT_ON =  TESTSON||{NOPRINT };
  OPT_OFF = TESTSOFF || {  ALPHA SIGSCAL WARN TOTAL_N BETASCAL};


    DO ICASE=1 TO 36 BY 1;   ***36 BY 1;
****CREATE BETASCAL, N, ESSENCEX;
   BETASCAL=HOLDIN[ICASE,8];
    *BETASCAL=1;
   N=HOLDIN[ICASE,4];
   HOLDCASE=HOLDIN[ICASE,*];
   N1=N/Q;
   OFFSET=MIN(HOLDIN[*,2]);
   ESSENCEX=I(Q)@J(N1,1,1);
   SIGMA=I(P);
   BETA=J(Q,P,0);
   DMAX=MIN(NCOL(DLIST),P,Q);
   DROW=NROW(DLIST);


****CREATE BETA;
RHOSQLIST1={.7 .4 0 0};
DLIST1=SQRT(RHOSQLIST1/(1-RHOSQLIST1)/N1);
RHOSQLIST2={.5 .03 .01 .011};
```

```
DLIST2=SQRT(RHOSQLIST2/(1-RHOSQLIST2)/N1);

RHOSQLIST3={.1 .1 .1 .1};

DLIST3=SQRT(RHOSQLIST3/(1-RHOSQLIST3)/N1);

RHOSQLIST4={.5 .5 0 0};

DLIST4=SQRT(RHOSQLIST4/(1-RHOSQLIST4)/N1);


   RHOSQLIST=RHOSQLIST1//RHOSQLIST2//RHOSQLIST3//RHOSQLIST4;

   DLIST=DLIST1//DLIST2//DLIST3//DLIST4;

   *DLIST4=SQRT(Q#RHOSQLIST4/(1-RHOSQLIST4));


   IVAR=HOLDIN[ICASE,2]-OFFSET+1;

   BETA1=DIAG(DLIST[IVAR,*]);

   BETA=BETA1[1:Q,];
****COMPUTE POPULATION POWER;
THETA=C*BETA*U;
SIGMASTAR=U`*SIGMA*U;


   RUN POWER;

   *PRINT _HOLDPOWERLBL;

   HLTPOWER=_HOLDPOWER[1,1];

   *PRINT "A:HLT - HLTPOWER" HLTPOWER;
****INITIALIZE COVERAGE COUNTER;
   COUNTCOVER=J(3,1,0); *cover HLT power, low, in interval, high;


   ****LOOP TO SIMULATE SIGMAHAT AND COMPUTE POWER;

   DEBUG=0;

   CLTYPE=1;


   ALPHA_CL= 0.05;

   ALPHA_CU= 0.05;


   *ALPHA_CL= (ALPHA/2);

   *ALPHA_CU= (ALPHA/2);


   N_EST=N;*training sample;

   N_ESTG=N_EST/Q;

   RANK_EST=3; *training sample;

   NU_EST = N_EST - RANK_EST;


   *MUMATEST=J(N_EST,NCOL(SIGMA),0);

   X1=I(Q) @ J(N_ESTG,1,1);

   MUMATEST=BETASCAL#(X1*BETA);


   XPX1=X1`*X1;

   IH1=I(N_EST)-X1*INV(XPX1)*X1`;
```

```
*RESET FUZZ;
   REJECTS=J(1,1,0);


       DO REPLICAT=1 TO NREP BY 1;
      F_SIGMA=I(P);
        Y=GAUSS1(N_EST,MUMATEST,F_SIGMA,SEED);
INDVARS={"X1" "X2" "X3"};
DEPVARS={"Y1" "Y2" "Y3" "Y4"};
ZNAMES=INDVARS||DEPVARS;
     Z=X1||Y;
     RUN MAKESS;
     RUN FITMODEL;
     RUN TESTGLH;
     PVALUE=_STMAT1_[2,5];
    IF (PVALUE <= ALPHA) THEN REJECTS=REJECTS+{1};
    SIGMAHAT=Y`*IH1*Y/NU_EST;
    SIGMA=SIGMAHAT;


    RUN POWER;


    END; *REPLICAT;


  IF DISPLAY THEN DO;
     PRINT / HOLDCASE[COLNAME=INNM];
     ROWPNM = {"HLT"};
     POWERHAT=REJECTS/NREP;
     STDPOW=SQRT(POWERHAT#(1-POWERHAT)/NREP);
     Z=PROBIT(.95);
     CLPOW=(POWERHAT-Z#STDPOW)||(POWERHAT+Z#STDPOW);
     PRINT POWERHAT  [ROWNAME=ROWPNM FORMAT=7.3]
          "  95% CI"  CLPOW [FORMAT=7.3];
    PRINT STDPOW;
     END;



   FREE N_EST RANK_EST CLTYPE _HOLDPOWER;

   END; *ICASE;


  END; *TESTIND;



QUIT;
```

References

Bain, L. J. and Engelhardt, M. (1992) *Introduction to Probability and Mathematical Statistics*. PWS-KENT; Boston, MA.

Bauer, P. and Kieser, M. (1996) A unifying approach for confidence intervals and testing of equivalence and difference, *Biometrika*, **83**, 934-937.

Beal, S. L. (1989) Sample size determination for confidence intervals on the population mean and on the difference between two population means, *Biometrics*, **45**, 969-977.

Berry, D. A. (1996) *Statistics: A Bayesian Perspective*. Duxbury Press; Belmont, CA.

Bickel, P.J. and Doksum, K.A. (1977) *Mathematical Statistics: Basic Ideas and Selected Topics*. Prentice Hall; Englewood Cliffs, NJ.

Birkett, M. A. and Day, S. J. (1994) Internal pilot studies for estimating sample size, *Statistics in Medicine*, **13**, 2455-2463.

Bristol D. R. (1989) Sample sizes for constructing confidence intervals and testing hypotheses, *Statistics in Medicine*, **8**, 803-811.

Bullitt E., Muller K. E., Jung I., Lin W. and Aylward S. R. (2004a) Analyzing attributes of vessel populations, *Medical Image Analysis*, **9**, 39-49.

Bullitt E, Jung I, Muller K, Gerig G, Aylward S, Joshi S, Smith K and Ewend M. (2004b) Determining malignancy of brain tumors by analysis of vessel shape, *MICCAI (LNCS)*.

Carlin, B. P. and Louis, T. A. (2000) *Bayes and Empirical Bayes Methods for Data Analysis*, 2nd edition. Chapman & Hall/CRC; Boca Raton, FL.

Catellier D.J. and Muller K.E. (2000) Tests for Gaussian repeated measures with missing data in small samples, *Statistics in Medicine*, **19**, 1101-1114.

Clement, M. E. (2005) *Analyzing Diffusion Tensor Imaging Data*, unpublished masters paper, Department of Biostatistics, University of North Carolina.

Coffey, C. S. and Muller, K. E. (1999) Exact test size and power of a Gaussian error linear model for an internal pilot study, *Statistics in Medicine*, **18**, 1199-1214.

Coffey, C. S. and Muller, K. E. (2000a) Properties of doubly-truncated Gamma variables, *Communications in Statistics - Theory & Methods*, **29**, 2677-2691.

Coffey, C. S. and Muller, K. E. (2000b) Some distributions and their implications for an internal pilot study with a univariate linear model, *Communications in Statistics - Theory & Methods*, **29**, 2677-2691.

Coffey, C. S. and Muller, K. E. (2001) Controlling test size while gaining the benefits of an internal pilot design, *Biometrics*, **57**, 625-631.

Denne, J. S. and Jennison, C. (1999) Estimating the sample size for a t-test using an internal pilot, *Statistics in Medicine*, **18**, 1575-1585.

Gatsonis, C. and Sampson, A. R. (1989) Multiple correlation: exact power and sample size calculations, *Psychological Bulletin*, **106**, 516-524.

Gillett, R. (1994) An average power criterion for sample size estimation, *Statistician*, **43**, 389-394

Glueck, D. H. (1995) *Power for a generalization of the GLMM with fixed and random predictors,* Ph.D. dissertation, Department of Biostatistics, University of North Carolina, Chapel Hill.

Gould, A. L. (1992) Interim analyses for monitoring clinical trials that do not materially effect the type I error rate, *Statistics in Medicine*, **11**, 55-66.

Gould, A. L. and Shih, W. J. (1992) Sample size reestimation without unblinding for normally distributed outcomes with unknown variance, *Communications in Statistics - Theory & Methods*, **21**, 2833-2853.

Gould, A. L. (1995) Planning and revising the sample size for a trial, *Statistics in Medicine*, **14**, 1039-1051.

Grieve, A. P. (1991) Confidence intervals and sample sizes, *Biometrics*, **47**, 1597-1603.

Guenther, W. C. (1973) Determination of sample size for tests concerning means and variances of normal distributions, *Statistica Neerlandica*, **27**, 103-113.

Gupta, A. K. and Nagar, D. K. (2000) *Matrix Variate Distributions*. Chapman Hall/CRC; Boca Raton, FL.

Helms, R. W. (1988) Comparisons of Parameter and Hypothesis Definitions in a General Linear Model, *Communications in Statistics, Theory and Methods*, **17**, 2725-2753.

Helms, R. W. (1992) Intentionally incomplete longitudinal designs: I. Methodology and comparison of some full span designs, *Statistics in Medicine*, **11**, 1889-1993.

Hogg, R. V. and Tanis, E. A. (1988) *Probability and Statistical Inference, 3rd edition.* Macmillian Publishing Company; New York, NY.

Hsu, J. C., Hwang, J. T. G., Liu H. K. and Ruberg, S. J. (1994) Confidence intervals associated with tests for bioequivalence, *Biometrika*, **81**, 103-114.

Huque, M. F., |Sankoh, A. J. and Rashid, M. M. (1997) Large Sample Inference for Non-Inferiority and Clinical Equivalence Trials and the Impact of Multiple Endpoints. *Proceedings of 1997 ASA Conference - Biopharmceutical Section*, 224-229.

Jennison, C. and Turnbull, B. W. (2000) *Group Sequential Methods with Applications to Clinical Trials*. Chapman & Hall; Boca Raton, FL.

Johnson, N. L., Kotz S. and Balakrishnan, N. (1994) *Continuous Univariate Distributions, Vol. 1, 2nd edition.* John Wiley & Sons, Inc; New York, NY.

Kieser, M. and Friede, T. (2000) Re-calculating the sample size in internal pilot study designs with control of the type I error rate, *Statistics in Medicine*, **19**, 901-911.

Kupper, L. L. and Hafner, K. B. (1989) How appropriate are popular sample size formulas?, *American Statistician*, **43**, 101-105.

Lehmann, E. L. (1959) *Testing Statistical Hypotheses*. Wiley; New York.

Lindley, D. V. (1997) The choice of sample size, *Statistician*, **46**, 129-138.

Muller, K. E. and Barton, C. N. (1989) Approximate power for repeated measures ANOVA lacking sphericity, *Journal of the American Statistical Association*, **84**, 549-555.

Muller, K. E., LaVange, L. M., Ramey, S. L and Ramey, C. T. (1992) Power calculations for general linear multivariate models including repeated measures applications, *Journal of the American Statistical Association*, **87**, 1209-1226.

Muller, Keith E., Benignus, Vernon (1992) Increasing scientific power with statistical power, *Neurotoxicology and Teratology,* **Vol. 14**, 211-219

Muller, K. E. and Pasour, V. B. (1997) Bias in linear model power and sample size due to estimating variance, *Communications in Statistics - Theory & Methods*, **26**, 839-851.

Muller, K. E. and Peterson L. Bercedis (1984) Practicl methods for computing power in testing the multivariate general linear hypothesis, *Computational Statistics & Data Analysis,* **2**, 143-158.

Muller K.E. and Stewart P.W. (2006), *Linear Model Theory: Univariate, Multivariate and Mixed Models*. Wiley; New York.

Muller K.E., Lloyd J. Edwards, Sean L. Simpson and Douglas J. Taylor (2007), Statistical tests with accurate size and power for balanced linear mixed models, *Statistics in Medicine*, **26**, in press.

Olson, C. L. (1974) Comparative robustness of six tests in multivariate analysis of variance, *Journal of the American Statistical Association*, **69**, 894–908.

Olson, C. L. (1976) On choosing a test statistic in multivariate analysis, *Psychological Bulletin*, **83**, 579–586.

Olson, C. L. (1979) Practical considerations in choosing a MANOVA test statistic: a rejoinder to Stevens, *Psychological Bulletin*, **86**, 1350–1352.

Pan, Z. and Kupper, L. L. (1999) Sample size determination for multiple comparison studies treating confidence interval width as random, *Statistics in Medicine*, **18**, 1475-1488.

Pratt, J. W. (1961) Length of confidence intervals, *Journal of the American Statistical Association*, **56**, 549-567.

Proschan, M. A. and Wittes, J. (2000) An improved double sampling procedure based on the variance, *Biometrics*, **56**, 1183-1187.

Rashid, M. M. (2000) Rank-Based Procedures for Non-inferiority and Equivalence Hypotheses in Clinical Trials When the Centers are Chosen at Random. *Proceedings of 2000 ASA Conference - Biopharmceutical Section*, 127-132.

Sampson, A. R. (1974) A tale of two regressions, *Journal of the American Statistical Association*, **69**, 682-689.

Sandvik, L., Erikssen, J., Mowinckel, P. and Rødland, A. (1996) A method for determining the size of internal pilot studies, *Statistics in Medicine*, **15**, 1587-1590.

Stein, C. (1945) A two-sample test for a linear hypothesis whose power is independent of the variance, *Annals of Mathematical Statistics*, **16**, 243-258.

Taylor, D. J. and Muller, K. E. (1995) Computing confidence bounds for power and sample size of the general linear univariate model, *American Statistician*, **49**, 43-47.

Taylor, D. J. and Muller, K. E. (1996) Bias in linear model power and sample size calculation due to estimating noncentrality, *Communications in Statistics: Theory and Methods*, **25**, 1595-1610.

Wang Y. and Kupper L. L. (1997) Optimal samples sizes for estimating the difference in means between two normal populations treating confidence interval length as a random variable, *Communications in Statistics - Theory & Methods*, **26**, 727-741.

Wittes, J. and Brittain, E. (1990) The role of internal pilot studies in increasing the efficiency of clinical trials, *Statistics in Medicine*, **9**, 65-72.

Wittes, J., Schabenberger, O., Zucker, D. M., Brittain, E. and Proschan, M. (1999) Internal pilot studies I: type I error rate of the naive t-test, *Statistics in Medicine*, **18**, 3481-3491.

Zucker, D. M., Wittes, J. T., Schabenberger, O. and Brittain, E. (1999) Internal pilot studies II: comparison of various procedures, *Statistics in Medicine*, **18**, 3493-3509.