IMPROVING COVERAGE OF RECTANGULAR CONFIDENCE REGIONS

by

Hakan Gogtas

Submitted to the Graduate Faculty of
the Art and Sciences in partial fulfillment
of the requirements for the degree of

B.Sc.,Hacettepe University,1994
M.A.,University of Michigan,1998

University of Pittsburgh

2004

UNIVERSITY OF PITTSBURGH

FACULTY OF ART AND SCIENCES

This dissertation was presented

by

Hakan Gogtas

It was defended on

July 6, 2004

and approved by

Professor Leon J. Gleser

Professor Satish Iyengar

Professor Ori Rosen

Professor John W. Wilson

Dissertation Director: Professor Leon J. Gleser

IMPROVING COVERAGE OF RECTANGULAR CONFIDENCE REGIONS

Hakan Gogtas, PhD

University of Pittsburgh, 2004

To find a better confidence region is always of interest in statistics. One way to find better confidence regions is to uniformly improve coverage probability over the usual confidence region while maintaining the same volume. Thus, the classical spherical confidence regions for the mean vector of a multivariate normal distribution have been improved by changing the point estimator for the parameter.

In 1961, James and Stein found a shrinkage estimator having total mean square error, TMSE, smaller than that of the usual estimator. In 1982, Casella and Hwang gave an analytical proof of the dominance of the confidence sphere which uses the James Stein estimator as its center over the usual confidence sphere centered at the sample mean vector. This opened up new possibilities in multiple comparisons.

This dissertation will focus on simultaneous confidence intervals for treatment means and for the differences between treatment means and the mean of a control in one-way and two-way Analysis of Variance, ANOVA, studies. We make use of Stein-type shrinkage estimators as centers to improve the simultaneous coverage of those confidence intervals. The main obstacle to an analytic study is that the rectangular confidence regions are not rotation invariant like the spherical confidence regions.

Therefore, we primarily use simulation to show dominance of the rectangular confidence intervals centered around a shrinkage estimator over the usual rectangular confidence regions

centered about the sample means. For the one-way ANOVA model, our simulation results indicate that our confidence procedure has higher coverage probability than the usual confidence procedure if the number of means is sufficiently large. We develop a lower bound for the coverage probability of our rectangular confidence region which is a decreasing function of the shrinkage constant for the estimator used as center and use this bound to prove that the rectangular confidence intervals centered around a shrinkage estimator have coverage probability uniformly exceeding that of the usual rectangular confidence regions up to an arbitrarily small epsilon when the number of means is sufficiently large. We show that these intervals have strictly greater coverage probability when all the parameters are zero, and that the coverage probability of the two procedures converge to one another when at least one of the parameters becomes arbitrarily large.

To check the reliability of our simulations for the one-way ANOVA model, we use numerical integration to calculate the coverage probability for the rectangular confidence regions. Gaussian quadrature making use of Hermite polynomials is used to approximate the coverage probability of our rectangular confidence regions for n=2, 3, 4. The difference in results between numerical integration and simulations is negligible. However, numerical integration yields values slightly higher than the simulations.

A similar approach is applied to develop improved simultaneous confidence intervals for the comparison of treatment means with the mean of a control. We again develop a lower bound for the coverage probability of our confidence procedure and prove results similar to those that we proved for the one-way ANOVA model.

We also apply our approach to develop improved simultaneous confidence intervals for the cell means for a two-way ANOVA model. We again primarily use simulation to show dominance of the rectangular confidence intervals centered around an appropriate shrinkage estimator over the usual rectangular confidence regions. We again develop a lower bound for the coverage probabilities of our confidence procedure and prove the same results that we proved for the one-way model.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

After rejecting the null hypothesis in one way and two-way analysis of variance,(ANOVA), or in other words, concluding that at least one of the cell means is different from all the other cell means, the next step is to make an inference as to where this difference might be. This is called multiple comparisons or simultaneous statistical inference.

The general principles of multiple comparisons were formed by Duncan, Scheffé and Tukey, creating its current structure. There is no agreement as to which method of multiple comparisons is the best. Simultaneous confidence intervals can be given for the cell means, for comparisons of each mean with with a control, for all pairwise comparisons of means, and for all linear combinations of the cell means. Sample means X_i are used as estimators for population means θ_i in the classical methods. For example, classical simultaneous confidence intervals take the form:

$$\theta_i \in \ [X_i - c\sigma_i \ , \ X_i + c\sigma_i] \ , \ i = 1, .., n$$

where σ_i is the standard deviation of X_i and the half width, c, of the interval depends on the method used (Duncan, Scheffé or Tukey).

Two types of confidence regions are often used:

1.The rectangular confidence regions

2.The spherical confidence regions

We are investigating rectangular confidence regions, but we are motivated by previous results for spherical confidence regions. Rectangular confidence regions have the following form:

$$C_0 = \{\Theta = (\theta_1, ..., \theta_k) : |X_i - \theta_i| \leq c\sigma_i, i = 1..., n\}$$

and c is a tabled constant depending on the method used. Keep in mind that given the confidence procedure, the coverage probabilities of rectangular confidence regions are same as the coverage probabilities of the corresponding simultaneous confidence regions. The spherical confidence regions have the following form:

$$C = \{\|X - \Theta\| \leq s\sigma\} \ , \ X = (X_1, X_2, ..., X_k)^{'}, \quad \Theta = (\theta_1, \theta_2, ..., \theta_k)^{'}$$

where $\|X - \Theta\|$ is the length of the vector, s is the radius of the sphere, and here all $X_i's$ are assumed to have standard deviation σ.

To find better simultaneous confidence intervals for the cell means is always of interest in statistics. One way to find better simultaneous confidence intervals is to uniformly improve the coverage probability of this confidence rectangle over the parameter space while maintaining the same volume. We conjecture that this can be done by changing the point estimator for Θ. Such an approach has been successful for improving the coverage probabilities of the classical spherical confidence regions.

In 1961, James and Stein found a shrinkage point estimator for Θ having total mean square error,(TMSE), smaller than that of the usual estimator, X. In 1982, Casella and Hwang proved that the confidence sphere which uses this point estimator as its center has uniformly higher coverage probability than the usual confidence sphere centered at X while maintaining the same volume. This opened new possibilities in multiple comparisons.

This dissertation focuses on simultaneous confidence intervals for treatment means. We also consider simultaneous confidence intervals for the differences between each of the treatment means and the mean of a control in one-way and two-way ANOVA studies. We make use of Stein-type shrinkage estimators as centers to improve the simultaneous coverage of these confidence intervals. The main obstacle to an analytic demonstration of this improvement in coverage probability is that the rectangular confidence regions are not invariant under rotation, preventing the simplification of integral expressions for coverage probability used in the study of the spherical confidence regions.

In Chapter 2, after introducing one-way and two-way ANOVA models and notations for the models, we do a literature review for the simultaneous confidence regions for θ_i for one-way and two-way ANOVA models, and also for the simultaneous confidence regions for comparisons of cell means with a control mean for one-way ANOVA model. Then, we will outline the main ideas and result of Fabian's (1991) paper. Finally, we will do a literature review for the James Stein estimator and Lindley estimator and generalization of Lindley estimator, and for spherical confidence region centered at the James-Stein estimator, and pretest estimator.

In Chapter 3, we first introduce our confidence procedure for the simultaneous confidence intervals for the cell means. We give a lower bound for the coverage probability of our rectangular confidence region that is a decreasing function of the shrinkage constant used for the point estimator that is the center of the region. We prove that the simultaneous confidence intervals centered around a shrinkage estimator define a rectangular confidence region that dominates the usual rectangular confidence regions up to an arbitrarily small constant for a sufficiently large number of means, n. We show that these intervals have strictly greater coverage probability when all the parameters are zero, and that the coverage probabilities of the two confidence procedures converge to one another when at least one of the parameters becomes arbitrarily large.

In Chapter 4, we introduce our improved simultaneous confidence intervals for the comparison of treatment means with the mean of a control. Since the simultaneous confidence intervals for the comparison of treatment means with the mean of a control have form similar to that of the simultaneous confidence intervals for the cell means, we are able to prove results similar to those that we proved in Chapter 3.

Because we extensively use simulation to demonstrate the domination of our confidence regions over the usual ones, we want to make sure that our simulation results are reliable. To check reliability of our simulations, in Chapter 5 we use numerical integration to calculate the coverage probabilities for the simultaneous confidence regions for the cell means in a one-way ANOVA model.

In Chapter 6, we introduce our rectangular confidence region for two-way ANOVA models. We obtain results for two-way ANOVA models similar to those that we proved for the one-way ANOVA models.

In Chapter 7, we discuss possibilities for future research.

2.0 NOTATION AND LITERATURE REVIEW

For analysis of variance problems, we prefer use of the cell mean model, which has a more straightforward interpretation. We will first introduce our notation for cell mean models for one-way and two-way ANOVA models, then we will briefly review the literature dealing with simultaneous confidence regions for the population means, θ_i, for one-way and two-way ANOVA models and with spherical confidence regions centered at the James Stein,(JS), and Lindley estimators.

2.1 BALANCED ONE-WAY ANALYSIS OF VARIANCE MODEL

In one-way analysis of variance (also known as the one-way classification),we assume that data Y_{ij} are observed according to a model,

$$Y_{ij} = \theta_i + \epsilon_{ij}, \quad j = 1, ..., k,\ i = 1, ..., n, \tag{2.1}$$

where the population means, θ_i, are unknown parameters and ϵ_{ij} are i.i.d N(0,σ^2) error random variables.

We will denote sample cell mean by X_i; thus,

$$X_i = \overline{Y_i} = \sum_{j=1}^{k} \frac{Y_{ij}}{k}, \tag{2.2}$$

and we let $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$, denote the average sample mean (often called Grand mean).

If σ^2 is known, the data can be reduced by sufficiency to $X_1, X_2, ..., X_n$ which are independent normally distributed random variables with means $\theta_1, ..., \theta_n$ and variances $\frac{\sigma^2}{k}, ..., \frac{\sigma^2}{k}$.

In this dissertation, we consider the balanced one-way analysis of variance (ANOVA) model with both known and unknown σ^2. In the *Appendix*, we give a modification of our procedure for the unbalanced one-way ANOVA models with known or unknown σ^2.

2.2 BALANCED TWO-WAY ANALYSIS OF VARIANCE MODEL

Suppose the data Y_{ijk} are obtained from a balanced two-way classification design. In such a two factor experiment, there are I levels of factor A and J levels of factor B and K replications for each treatment combination of the ith level of factor A and jth level of factor B. The model is given by

$$Y_{ijk} = \theta_{ij} + \epsilon_{ijk} = \theta + a_i + b_i + c_i + \epsilon_{ijk} \tag{2.3}$$

where θ_{ij} is the population mean for the $(i, j) - th$ treatment combination, θ is a constant , a_i is the main effect associated with the ith level of factor A, b_j is the main effect associated with factor B, c_{ij} is the interaction of level i of factor A and level j of factor B, and ϵ_{ijk} are the error terms. The ϵ_{ijk} are distributed independently and normally:

$$e_{ijk} \sim N(0, \sigma_e^2), \quad 1 \le i \le I,\ 1 \le j \le J,\ 1 \le k \le K. \tag{2.4}$$

The sample mean of all the observations from the (i,j)th treatment combination will be called the i, j sample cell mean and it is denoted as:

$$\overline{Y_{ij}} = \sum_{k=1}^{K} \frac{Y_{ijk}}{K}. \tag{2.5}$$

Letting X_{ij} be $\overline{Y_{ij}}$, then the reduced data consists of $X_{11}...X_{ij}...X_{IJ}$, and we can represent X_{ij} in the form,

$$X_{ij} = X_{0ij} + W_{ij}, \tag{2.6}$$

In (2.6), X_{0ij} is the additive main effect and W_{ij} is the interaction effect. Here

$$X_{0ij} \sim N(0, \frac{I+J-1}{IJ}),\ W_{ij} \sim N(\theta_{ij}, \frac{(I-1)(J-1)}{IJ})$$

and $\sum_{j=1}^{J} W_{ij} = 0$, i=1,...,I; $\sum_{i=1}^{I} W_{ij} = 0, j = 1, ..., J$, $\sum_{i=1}^{I}\sum_{j=1}^{J} X_{oij} = 0$
The usual confidence rectangle for the population means, θ_{ij}, under the (i,j) treatment combination is

$$D_1 = \{\Theta = (\theta_{11}, ..., \theta_{IJ}) : (|X_{ij} - \theta_{ij}| \leq s, i = 1..., I, j = 1, ..., J\}. \tag{2.7}$$

where $s = \frac{\sigma_e}{\sqrt{K}}\Phi^{-1}(\frac{\beta^{\frac{1}{N}}+1}{2})$, N=IJK and Φ is the standard normal cumulative distribution function.

2.3 LITERATURE REVIEW

We will do a literature review for the simultaneous confidence regions for population cell means (Treatment means) θ_i, and also simultaneous confidence regions for comparisons of the θ_i with a control mean, in one-way ANOVA problems. Then, we will review the literature concerning simultaneous confidence intervals for the population cell means θ_{ij} in two-way ANOVA problems, and also outline the main ideas and result of Fabian's (1991) paper. Then, we review the literature on the James-Stein and Lindley point estimators and their generalizations. Finally, we review results for spherical confidence regions centered at the JS estimator.

2.3.1 One-Way ANOVA Model

As we mentioned in the Introduction, simultaneous inference methods are either for tests or for confidence intervals. We only review the methods for simultaneous confidence regions.

2.3.1.1 The Usual Method

If σ is known:

$$\theta_i \in X_i \pm c\sigma_i, \ i = 1, ..., n \tag{2.8}$$

gives exact $100(1-\alpha)\%$ simultaneous confidence intervals for θ_i, where

$$c = \Phi^{-1}[(\beta^{1/n} + 1)/2],$$

and $\beta = 1 - \alpha$.

If σ is unknown,

The Studentized Maximum Modulus (SMM) method, which gets its name because it is based on the studentized maximum modulus statistic

$$\max_{1 \le i \le n} |\frac{X_i - \theta_i}{\frac{\hat{\sigma}}{\sqrt{k_i}}}|,$$

provides exact $100(1-\alpha)\%$ simultaneous confidence intervals for $\theta_1, ..., \theta_n$, where k_i is the number of replications for the *ith* treatment. These intervals are

$$\theta_i \in X_i \pm |m|_{\alpha,n,\nu} \frac{\hat{\sigma}}{\sqrt{k_i}}, \ i = 1, ..., n \tag{2.9}$$

where $|m|_{\alpha,n,\nu}$ is the $1-\alpha$ quantile of the student maximum modulus distribution, and is computed as the solution of the equation

$$\int_0^\infty [\Phi(|m|_{\alpha,n,\nu}s) - \Phi(-|m|_{\alpha,n,\nu}s)]^n \gamma_\nu(s) ds = 1 - \alpha. \tag{2.10}$$

In (2.10), Φ is the standard normal cdf, γ_ν is the density of $\frac{\hat{\sigma}}{\sigma}$ and $\nu = n - 1$.

2.3.1.2 A Product Inequality Method

The random variables

$$|T_i| = \frac{X_i - \theta_i}{\frac{\hat{\sigma}}{\sqrt{k_i}}},\ i = 1, ..., n$$

are independent except for the common divisor $\hat{\sigma}$. If we assume that they are independent, the appropriate simultaneous confidence intervals for $\theta_1, ..., \theta_n$ would be same as (2.9), except for replacing $|m|_{\alpha,n,\nu}$ by the $1-\alpha$ quantile $t_{[1-(1-\alpha)^{1/n}]/2,\nu}$ of t-distribution with ν degrees of freedom. The resulting simultaneous confidence intervals are

$$\theta_i \in X_i \pm t_{[1-(1-\alpha)^{1/n}]/2,\nu} \frac{\hat{\sigma}}{\sqrt{k_i}},\ i = 1, ..., n. \tag{2.11}$$

The confidence interval (2.11) is conservative; That is,
$P(\theta_i \in X_i \pm t_{[1-(1-\alpha)^{1/n}]/2,\nu} \frac{\hat{\sigma}}{\sqrt{k_i}}, i = 1, ..., n) \geq 1 - \alpha$.

2.3.1.3 The Bonferroni Inequality Method

The familiar Bonferroni inequality states, for any events $E_1, ...E_n$ that,

$$P(\cup_{m=1}^{n} E_m^c) \leq \sum_{m=1}^{n} P(E_m^c) \tag{2.12}$$

Applying the Bonferroni inequality to

$$E_i = \{\theta_i \in \hat{\theta}_i \pm q \frac{\hat{\sigma}}{\sqrt{k_i}}\}$$

then the Bonferroni simultaneous confidence intervals are

$$\theta_i \in X_i \pm t_{\alpha/2n,\nu} \frac{\hat{\sigma}}{\sqrt{k_i}},\quad 1 \leq i \leq n, \tag{2.13}$$

The Bonferroni confidence interval is always more conservative than the product inequality confidence intervals because $(1-\alpha)^{1/n} < 1 - \alpha/n$ for all $\alpha > 0$ and $n > 1$.

2.3.2 Simultaneous Confidence Regions for Comparisons with a Control

2.3.2.1 The Usual Method We are interested in simultaneous confidence regions for comparisons of the θ_i with a control mean in the balanced one-way ANOVA problems with known σ^2. In the *appendix*, we give modifications of our procedure for the balanced one-way ANOVA model with unknown σ^2 and the unbalanced one-way ANOVA model with both known or unknown σ^2 .

If σ^2 is known, we can assume without loss of generality that variance is 1.

The sample cell means, $X_1, ..., X_n$, defined in (2.2), are independent normally distributed random variables with means $\theta_1, ..., \theta_n$ and variance 1. The sample control mean is normally distributed with mean θ_c and variance 1, independent of the sample cell means. The usual confidence interval for multiple comparison with control is

$$E_0 = \{|X_i - X_c - (\theta_i - \theta_c)| \leq c^*, i = 1, .., n\}.$$

where c^* is defined by,

$$P(E_0) = P\{|Y_i| \leq c^*, i = 1, ..., n\} = 0.95. \tag{2.14}$$

Where $Y_i = X_i - X_c - \theta_i + \theta_c$, $Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$ has a n-variate normal distribution $N(0, \Sigma)$, and

$$\Sigma = \begin{pmatrix} 2 & 1 & \cdots & 1 & 1 \\ 1 & 2 & 1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & \cdots & 1 & 2 \end{pmatrix}.$$

2.3.2.2 Dunnett's Method

Dunnett's (1955) two sided method provides the following simultaneous confidence intervals for the difference between each new treatment mean θ_i and the control mean θ_c when σ^2 is unknown:

$\theta_i - \theta_c \in \widehat{\theta_i} - \widehat{\theta_c} \pm |d|\widehat{\sigma}\sqrt{\frac{2}{k}}$ for all $i = 1, ...n$
where $|q|$ is the solution to the equation

$$\int_o^{\infty}\int_{-\infty}^{\infty}[\Phi(z+\sqrt{2}|q|s) - \Phi(z-\sqrt{2}|q|s)]^n|q|\phi(z)\gamma(s)ds = 1-\alpha$$

where Φ is the standard normal distribution function, ϕ is the standard normal density function and γ is the density of $\frac{\hat{\sigma}}{\sigma}$.

2.3.3 Two way analysis of variance

In the context of ANOVA studies, additive effects provide much simpler explanations of the factor effects than do interacting effects. The presence of interacting effects complicates the explanation of the factor effects because they must then be described in terms of the combined effects of the two factors. The following strategy is suggested by most textbooks,

1. Examine whether the factors interact.

2. If they do not interact, examine whether the main effects are important.

3. If the factors do interact, examine if the interactions are important or unimportant.

4. If the interactions are unimportant, proceed as in step 2.

5. If the interactions are important, consider whether they can be made unimportant by a meaningful simple transformation of scale.

6. For important interactions that can not be made unimportant by a simple transfor-

mation, analyze the two factor effects jointly in terms of treatment means θ_{ij}.

2.3.3.1 Analysis of Factor effects when factors do not interact $100(1-\alpha)\%$ simultaneous confidence intervals for $\overline{\theta_i}$ and $\overline{\theta_j}$:

$$\overline{Y_{i..}} \pm t[1-\alpha/2;(k-1)IJ]s\{\overline{Y_{i..}}\}$$

where σ_e^2 is unknown, $s\left\{\overline{Y_{i..}}\right\} = \frac{MSE}{JK}$.

$$\overline{Y_{.j.}} \pm t[1-\alpha/2;(k-1)IJ]s\left\{\overline{Y_{.j.}}\right\}$$

where σ_e^2 is unknown, $s\left\{\overline{Y_{.j.}}\right\} = \frac{MSE}{IK}$. t is the usual t distribution.

Fabian (1991) suggested a confidence rectangle for cell means θ_{ij} that adapts to the extent of interaction estimated by the data in that length of his confidence interval depends on the largest interaction effect. Now, we will give the brief summary of the Fabian's paper (1991).

2.3.3.2 Fabian's Procedure As I mentioned above the usual recommendation is to neglect interactions and do analysis. Fabian (1991) suggested the following recommendation: ignore interactions and do analysis but estimate the error involved in neglecting the interactions from the power of the test. Then he gave the following $(1-\alpha)\%$ simultaneous confidence interval for all cell means;

$$Y_{oij} - s_2 < \theta_{ij} < Y_{oij} + s_2 \; for \; all \; i,j \tag{2.15}$$

where $s_2 = K_\beta(D+\epsilon) + \max|W_{ij}|$, $D = |Z_{oij}|$,and ϵ is a normal(0,(I-1)(J-1)/(IJ)) random variable independent of D.

He also stated that, one way method or usual recommendation for two way method,(1.18), is substantially better than (2.15).

Gleser (1992) criticized the way Fabain constructed his simultaneous confidence interval. He stated that the length of Fabian's confidence interval constructed by triangular inequality and has higher variance. He also argued the question that "Can other methods be found that preserve at least some of the benefits of (additive model) in saving variance,while still accounting for possible interaction. He suggested the confidence rectangle centered at the appropriate shrinkage estimator.

2.3.4 Stein Estimation Procedures in Spherical Confidence Intervals and Bayesian Estimation

We will do literature review for point estimation results of the shrinkage estimator. Then, we will do literature review for spherical confidence intervals centered at the shrinkage estimator.

2.3.4.1 The Point Estimation Results for the Shrinkage Estimator

Stein (1956) showed that the usual maximum likelihood estimator for $\theta = (\theta_1, ..., \theta_n)$, namely
$X = (X_1, ..., X_n)'$, has larger expected loss or risk than the estimator

$$\delta^S(X) = (1 - \frac{a\sigma^2}{b + \|X\|^2})X,$$

where $\|X\|^2 = \sum_{i=1}^n X_i^2$, provided that a is sufficiently small and b is sufficiently large.

James and Stein (1961) showed that each member of the class of estimators

$$\delta^{JS}(X) = (1 - \frac{a\sigma^2}{\|X\|^2})X \tag{2.16}$$

depending on $a > 0$, has a smaller total mean squared error, TMSE, than X for $0 < a < 2(p-2)$ and that a=p-2 gives the uniformly best estimator in this class of estimators, $p \geq 3$.

Baranchik(1964,1970) showed that the James-Stein (JS) positive part estimator has uniformly smaller TMSE than the James-Stein estimator. The positive part estimator is

$$\delta^{JS+}(X) = (1 - \frac{a\sigma^2}{\|X\|^2})^+ X \tag{2.17}$$

where,

$$(1 - \frac{a\sigma^2}{\|X\|^2})^+ = \begin{cases} 1 - \frac{a\sigma^2}{\sum_{i=1}^n X_i^2} & , \quad if \quad \sum_{i=1}^n X_i^2 \geq a\sigma^2 \\ 0 & , \quad if \quad \sum_{i=1}^n X_i^2 \leq a\sigma^2. \end{cases}$$

Brown (1971) observed that the positive part estimator is also inadmissible. An explicit estimator that dominates the JS positive part estimator in total mean squared error has been given by Shao and Strawderman (1994).

Lindley (1962) showed that the estimator $\delta(X) = (\delta_1(X), ..., \delta_n(X)$, where

$$\delta(X) = \overline{X} + (1 - \frac{n-3}{\sum(X_i - \overline{X})^2})(X - \overline{X}) \tag{2.18}$$

and $\overline{X} = (1/n)\sum_{i=1}^n X_i$, has a uniformly smaller mean squared error than the usual estimator X of θ for $n \geq 4$. This estimator shrinks toward an estimate of the average $\overline{\theta}$ of the population means. We briefly explain the main idea of Lindley's estimator. In the one way ANOVA model, the usual estimator of the vector of population cell means is the vector of sample cell means. It is desirable to find the best linear unbiased estimator of θ of the form $\hat{\theta} = a + bX$, where a is a vector. The resulting least square line is $\hat{\theta} = \overline{X} + (1 - \frac{n-1}{\sum(X_i - \overline{X})^2})(X - \overline{X})$. This estimator does not have a smaller total mean squared error, TMSE, than X. The estimator with a smaller TMSE was obtained by Lindley by simply replacing $n-1$ with $n-3$.

Judge, Hill and Bock (1990) note that for shrinkage estimators to achieve significant risk improvements as compared with MLEs, it is necessary to identify the region or subspace where the location vector being estimated is either known or thought to lie as a result of prior information. The best shrinkage estimators are those that shrink toward the correct subspace or region.

Sclove (1968) considers the estimation of the coefficients of a linear model with an column orthogonal design matrix at least three regression variables.He:

1. gives the form of the James-Stein estimator appropriate for this context;

2. observes that the mean square error of prediction(MSEP) when James-Stein estimator is used to estimate model coefficient is uniformly smaller than that resulting from using the Least Squares,LS;

3. observes that when the estimated regression coefficients are significantly different from zero JS is very close to the LS.

The observation that the JS and LS are very close when regression coefficients differ significantly from zero motivates the idea of a *preliminary test estimator*.This kind of estimator is formulated as follows:

1. A test of hypothesis is performed to see which regression coefficients are significantly different from zero.

2. Regression coefficients that are not significantly different from zero are estimated by JS.

3. Regression coefficients that are significantly different from zero are estimated by LS.

Sclove also;

1. mentions a positive part estimator;

2. gives an induction of how to extend his results to the case of non-orthogonal design matrices.

Judge et al.(1980) explain how a JS estimator may be viewed as a pretest estimator that combines the restricted and unrestricted least square estimators. They explain how a Stein like estimator originally formulated by Scolve, Morris, and Radhakrishnan (1972) dominates the usual pretest estimator.

Saleh et al.(1990) give four estimators:

1. the restirected least squares estimator (RLSE);

2. the unrestricted least squares estimator (URLSE);

3. the preliminary test least squares estimator (PTLSE);

4. the shrinkage leasts square estimator (SLSE) (a JS type estimator).

They derive and compare the risk functions. They find that:

1. The RLSE has the smallest risk if the true regression coefficients satisfies the restriction imposed on the estimated coefficients, but is unbounded when parameters move away from the subspace of the restriction.

2. The SLSE generally has the smallest risk, but not when the parameter is in or near the

restriction subspace;in this case PTLSE is better.

The authors recommend the use of PTLSE when the restriction subspace has dimension less than three ; otherwise they advocate the PTLSE and SLSE with the SLSE preferred.

2.3.4.2 Spherical Confidence Intervals Centered at the Shrinkage Estimator

Brown (1966) and Joshi (1967) independently demonstrated the existence of a confidence region for $n \geq 3$ that dominates the usual confidence region $C^o(X) = \{\theta : \|\theta - X\| \leq \chi_{(n;1-\alpha)}\}$. Joshi proved that the set

$$C^J(x) = \left\{\theta : \left\|\theta - \delta^S(X)\right\|^2 \leq c^2\right\},$$

has higher coverage probability than $C^0(x)$ if a is sufficiently small and b is sufficiently large. Olshen(1977) simulated the coverage probability of C^J for selected a, b and $\|\theta\|$. The results indicated that large gains in coverage probability can be achieved.

AS Stein(1962) and others noted, the James-Stein estimator arises in a natural fashion as an "Empirical Bayes Estimator". This result can be seen by following argument:

Assume θ_i are a sample from a prior distribution, where θ_i, for $i = 1, 2, ..., n$ is normally and independently distributed with mean zero and variance τ^2 with joint density in vector form $g(\theta/\lambda^2)$. Then the Bayesian estimator of θ under the squared error loss is $(1 - \frac{1}{1+\tau^2})\widehat{\theta}$.

These Bayesian result assume that $(1 - \frac{1}{1+\tau^2})$ and τ^2 are known. If the investigator does not know τ^2, he cannot use the Bayes rule. However, as Efron and Morris(1973) noted that one can stop short of the Bayesian fold and attempt to estimate $(1 - \frac{1}{1+\tau^2})$ from the data. Morris (1977) also simulated coverage probabilities for certain generalized Bayes estimators resulting in fairly simple confidence region and again, results were good.

Two other important works are those of Faith(1976) and Berger(1980).Faith derives confidence sets from Bayes credible sets and shows, for p=3 or 5, that these sets have small volume and higher coverage probability than $C^0(x)$ except for an interval of middle values of $\|\theta\|^2$. Unfortunately, his confidence regions are difficult to work with, having complicated shape arising from their Bayesian derivation. Berger also proceeds in a Bayesian fashion, but also uses the posterior covariance matrix to construct confidence ellipsoids. Resulting sets are shown to have uniformly smaller volume than $C^o(X)$, and to dominate C^0 in coverage probability for sufficiently large $\|\theta\|^2$.

Casella (1980) extended the method of Faith and derived exact formulas for the coverage probability of spherical confidence regions centered at the James-Stein or positive-part James-Stein estimators. Casella and Hwang (1982) proved that if the usual confidence sphere is recentered at the positive-part James-Stein estimator, then the resulting confidence region has a uniformly higher coverage probability for $n \geq 4$.

Casella and Hwang (1983) studied the spherical confidence set,
$C^{CH}(X) = \{\theta : |\theta - \delta^{JS+}(X)| \leq V(X)\}$ where $\delta^{js+}(X)$ is defined in (2.17) and $V(X)$ is derived through the use of the empirical Bayes argument. They stated conditions on $V(X)$ for the set $C^{CH}(X)$ to have a uniformly higher coverage probability than the usual spherical confidence set at $\theta = 0$. They failed to show numerical evidence for dominance in coverage probabilities for the range of middle values of $\|\theta\|^2$.

Casella and Hwang (1987) derived exact formulas for the coverage probability of the spherical confidence set, $C^{CH1}(X) = \{\Theta : |\Theta - \delta^{(A)}| \leq s\sigma\}$, where A is $n \times n$ idempotent matrix, k is the rank of A, $0 \leq a \leq n-k-2$ and $\delta^{(A)}(X) = AX + (1 - \frac{a}{X'(I-A)X})^+(I-A)X$. They proved that $C^{CH1}(X)$ has uniformly higher coverage than the usual spherical confidence sets for $n \geq 4$. Also, they noted that for $A = (1/n)11'$,$\delta^{(A)}(X)$ becomes the positive-part Lindley's estimator,$\delta^{JSL+}(X)$. They gave a evidence based on design simulations that the spherical confidence set centered at $\delta^{JSL+}(X)$, $C^{CH2}(X) = \{\theta : |\theta - \delta^{JSL+}(X)| \leq V(X)\}$where $V(X)$ is derived through the use of the empirical Bayes argument, has a uniformly higher coverage probability than the usual spherical confidence set.

3.0 ONE-WAY ANOVA MODEL

If the data Y_{ij} are obtained from a balanced one-way classification design,

$$Y_{ij} = \theta_i + \epsilon_{ij},\ \ j = 1, ..., k\ , i = 1, ..., n$$

where θ_i's are the population means and ϵ_{ij} are i.i.d $N(0, \sigma^2)$ error random variables. We first go over the known σ^2 case, and then the unknown σ^2 case. We explain the modification of our procedure for the unbalanced one-way and two-way ANOVA models in the *appendix*.

3.1 KNOWN σ^2 CASE

The sample means, $X_1, ..., X_n$, are independent normally distributed random variables with means $\theta_1, ..., \theta_n$ and variances $\frac{\sigma^2}{k}$. Then, the usual confidence rectangle for the vector Θ of cell means is

$$C_0 = \left\{ \Theta = (\theta_1, ..., \theta_n) : \left|\frac{X_i - \theta_i}{\frac{\sigma}{\sqrt{k}}}\right| \leq c\frac{\sigma}{\sqrt{k}}\ , i = 1, ..., n \right\}. \tag{3.1}$$

, where $c = \Phi^{-1}(\frac{\beta^{1/n}+1}{2})$.

Let

$$
\begin{aligned}
\sigma &= \frac{\sigma}{\sqrt{k}} && (3.2)\\
X_i^* &= \frac{X_i - \overline{X}}{\sigma} && (3.3)\\
\eta_i &= \frac{\theta_i - \overline{\theta}}{\sigma} && (3.4)\\
U_i &= X_i^* - \eta_i && (3.5)\\
V_i &= X_i - \overline{X} && (3.6)\\
Z_i &= \frac{X_i - \theta_i}{\sigma} && (3.7)\\
\overline{Z} &= \frac{\overline{X} - \overline{\theta}}{\sigma} && (3.8)\\
\overline{\theta} &= \frac{\sum_{i=1}^{n} \theta_i}{n} && (3.9)\\
c &= c\sigma
\end{aligned}
$$

The confidence region we are proposing is

$$
\begin{aligned}
C_1(X) &= \left\{\theta = (\theta_1, ..., \theta_k) : |R^+(X_i - \overline{X}) + \overline{X} - \eta_i| \le c, i = 1, ..., n\right\}\\
&= \left\{\theta = (\eta_1, ..., \eta_k) : |R^+ X_i^* + \overline{Z} - \eta_i| \le c, i = 1, ..., n\right\} \qquad (3.10)
\end{aligned}
$$

where R^+ is the positive-part Lindley's shrinkage factor defined by

$$
R^+ = (1 - \frac{a\sigma^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2})^+ = (1 - \frac{a}{\sum_{i=1}^{n} X_i^{*2}})^+. \qquad (3.11)
$$

or equivalently,

$$
R^+ = \begin{cases} 1 - \frac{a\sigma^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2} & : \quad \sum_{i=1}^{n}(X_i - \overline{X})^2 \ge a\sigma^2 \\ 0 & : \quad \sum_{i=1}^{n}(X_i - \overline{X})^2 \le a\sigma^2. \end{cases}
$$

The quantity a is called *the shrinkage constant*.
Also observe that $\sum_i^n U_i = 0$ and $\sum_i^n \eta_i = 0$.

One confidence region A is said to *dominate* another (B) if region A has uniformly higher the coverage probability than B; while maintaining the same or smaller volume.

3.1.1 ANALYTICAL RESULTS

Keep in mind that the coverage probability of $C_1(X)$ is equal to that of $C_0(X)$ when the shrinkage constant a is zero.

Lemma 3.1.1. $\lim_{\|\eta\|\to\infty} P(\sum_{i=1}^{n}(U_i+\eta_i)^2 \le a) = 0.$

Proof:

Let $SS = \sum_{i=1}^{n}(U_i+\eta_i)^2$ and $\delta = \sum_{i=1}^{n}\eta_i^2$. SS has a non central chi-square distribution with a non centrality parameter δ. Then,

$$\lim_{\|\delta\|\to\infty} P(SS \le a) = \lim_{\|\delta\|\to\infty} \int_0^a \sum_{j=0}^{\infty} \frac{\delta^j \exp^{-0.5\delta}}{j!} \frac{exp^{-0.5SS} SS^{0.5n+j-1}}{\Gamma(0.5n+j)2^{0.5n}} dSS$$

Since all the conditions are met for the dominated convergence theorem, we can take the limit inside of the integration and summation sign. Therefore;

$$\lim_{\|\delta\|\to\infty} P(SS \le a) = \int_0^a \sum_{j=0}^{\infty} \lim_{\|\delta\|\to\infty} \frac{\delta^j \exp^{-0.5\delta}}{j!} \frac{exp^{-0.5SS} SS^{0.5n+j-1}}{\Gamma(0.5n+j)2^{0.5n}} \to 0.$$

Lemma 3.1.2. $plim_{\|\delta\|\to\infty}(1-R^+)\max_i |U_i| = 0$

Proof:

This follows from the previous lemma 3.1.1.

The following lemma proves that the coverage probabilities of C_0 and C_1 converge to one another when at least one of the η_i becomes arbitrarily large.

Lemma 3.1.3. *if* $\eta'\eta \to \infty$, $P(C_1) \to P(C_0)$.

Proof:

Again let $\delta = \sqrt{\eta'\eta}$. For arbitrary $\epsilon > 0$, Let $A_\epsilon = \{(1-R^+)\max_i |U_i| \le \epsilon\}$. Then from lemma 3.1.1,

$$\lim_{\|\delta\|\to\infty} P(A_\epsilon^c) = 0 \tag{3.12}$$

Note that,

$$\eta \in C_1 \Leftrightarrow -c \le \frac{X_i - \theta_i - (1-R^+)(X_i - \overline{X})}{\sigma} \le c. \tag{3.13}$$

and $\eta = \frac{\theta_i - \bar{\theta}}{\sigma}$. Thus

$$\begin{aligned} P(\eta \in C_1) &= P\left\{\eta \in C_1 \cap A_\epsilon\right\} + P\left\{\eta \in C_1 \cap A_\epsilon^c\right\} \\ &\leq P(\left\{-c-\epsilon \leq \frac{X_i - \theta_i}{\sigma} \leq c+\epsilon\right\} \cap A_\epsilon) + P\left\{A_\epsilon^c\right\} \\ &\leq P\left\{-c-\epsilon \leq Z_i \leq c+\epsilon\right\} + P\left\{A_\epsilon^c\right\} \\ &= (2\Phi(c+\epsilon)-1)^n + P\left\{A_\epsilon^c\right\}. \end{aligned} \tag{3.14}$$

Also

$$\begin{aligned} P(\eta \in C_1) &\geq P\left\{-c+\epsilon \leq \frac{X_i - \theta_i}{\sigma} \leq c-\epsilon \cap A_\epsilon\right\} + P\left\{A_\epsilon^c\right\} \\ &\geq \left\{-c+\epsilon \leq Z_i \leq c-\epsilon\right\} - P\left\{A_\epsilon^c \cap -c+\epsilon \leq Z_i \leq c-\epsilon\right\} \\ &\geq P\left\{-c+\epsilon \leq Z_i \leq c-\epsilon\right\} - P\left\{A_\epsilon^c\right\} \\ &= (2\Phi(c-\epsilon)-1)^n - P\left\{A_\epsilon^c\right\}. \end{aligned} \tag{3.15}$$

It follows from (3.12)-(3.15) that,

$$(2\Phi(c-\epsilon)-1)^n \leq \liminf_{\delta\to\infty} P(\eta \in C_1) \leq \limsup_{\delta\to\infty} P(\eta \in C_1) \leq (2\Phi(c+\epsilon)-1)^n$$

However $\epsilon > 0$ is arbitrary. Taking $\epsilon \to 0$ and noting that $\Phi(.)$ is a continuous function of its argument, the assertion of the lemma follows.

The following lemma says that when all of the population treatment means are zero, our procedure has larger coverage probability than the usual procedure.

Lemma 3.1.4. *For $\eta'\eta = 0$, or equivalently for each $\eta_i = 0$, $i = 1, ..., n$, $P(C_1) \geq P(C_0)$*

Proof:

For each $\eta_i = 0$, R^+ becomes $R^+ = (1 - \frac{a}{\sum_{i=1}^{n} U_i^2})^+ and\ P(C_1)$ becomes

$$\begin{aligned} P(C_1) &= P\left\{\eta : |R^+U_i + \overline{Z}| \leq c, i = 1, ..n\right\} \\ &= \int_{(U_1,...,U_n)} \int_{(\sqrt{n}(-c-R^+U_{[n]})}^{(\sqrt{n}(c-R^+U_{[1]})} f_U(U)\varphi_{\overline{Z}}(\overline{Z})d\overline{Z}dU \end{aligned}$$

where $U_{[1]} = \max_{i=1,...n} U_i$ and $U_{[n]} = \min_{i=1,...n} U_i$. Also, $P(C_0)$ can be written in a similar way:

$$\begin{aligned} P(C_0) &= P\left\{\eta : |Z_i - \overline{Z} + \overline{Z}| \leq c\ , i = 1, ..., n\right\} \\ &= P\left\{\eta : |U_i + \overline{Z}| \leq c\ , i = 1, ..., n\right\} \\ &= \int_{(U_1,...,U_n)} \int_{(\sqrt{n}(-c-U_{[n]})}^{(\sqrt{n}(c-U_{[1]})} f_U(U)\varphi_{\overline{Z}}(\overline{Z})d\overline{Z}dU \end{aligned}$$

Note that $U_{[1]}$ must be ≥ 0 and $U_{[n]}$ must be ≤ 0 since $\sum_i U_i = 0$ because $0 \leq R^+ \leq 1$. It follows that $U_{[1]} \geq R^+U_{[1]}$ and $U_{[n]} \leq R^+U_{[n]}$, and hence

$$\int_{(U_1,...,U_n)} \int_{(\sqrt{n}(-c-R^+U_{[n]})}^{(\sqrt{n}(c-R^+U_{[1]})} f_U(U)\varphi_{\overline{Z}}(\overline{Z})d\overline{Z}dU \geq \int_{(U_1,...,U_n)} \int_{(\sqrt{n}(-c-U_{[n]})}^{(\sqrt{n}(c-U_{[1]})} f_U(U)\varphi_{\overline{Z}}(\overline{Z})d\overline{Z}dU$$

which proves that $P(C_1) \geq P(C_0)$. The following lemma and theorem state that for n=2, there is not an universal domination. In other words, our confidence procedure can not dominate the usual confidence region in every region for a common shrinkage constant.

Lemma 3.1.5. *For n=2, the coverage probability of C_1 is*

$$P(C_1) = \begin{cases} \int_{-\sqrt{a/2}-\eta_1}^{\sqrt{a/2}-\eta_1} (2\Phi(\sqrt{2}(c-|\eta_1|))-1) f_U(U)dU + \\ \int_{\frac{-(\eta_1+c)-\sqrt{(c-\eta_1)^2+2a}}{2}}^{-\sqrt{a/2}-\eta_1} g_{U,\eta_1,c,a}(U) f_U(U)dU \\ + \int_{\sqrt{a/2}-\eta_1}^{\frac{(c-\eta_1)+\sqrt{(c+\eta_1)^2+2a}}{2}} g_{U,\eta_1,c,a}(U) f_U(U)dU & : \quad 0 \le \eta_1 \le \sqrt{a/2} \\ \\ \int_{-\sqrt{a/2}-\eta_1}^{\sqrt{a/2}-\eta_1} (2\Phi(\sqrt{2}(c-|\eta_1|))-1) f_U(U)dU + \\ \int_{\frac{-(\eta_1+c)-\sqrt{(c-\eta_1)^2+2a}}{2}}^{-\sqrt{a/2}-\eta_1} g_{U,\eta_1,c,a}(U) f_U(U)dU \\ + \int_{\sqrt{a/2}-\eta_1}^{\frac{(c-\eta_1)+\sqrt{(c+\eta_1)^2+2a}}{2}} g_{U,\eta_1,c,a}(U) f_U(U)dU & : \quad \sqrt{a/2} \le \eta_1 \le c \\ \\ \int_{\frac{-(\eta_1+c)+\sqrt{(c-\eta_1)^2+2a}}{2}}^{\frac{-(\eta_1-c)+\sqrt{(c+\eta_1)^2+2a}}{2}} g_{U,\eta_1,c,a}(U) f_U(U)dU & : \quad \eta_1 \ge c \end{cases} \tag{3.16}$$

where $g_{U,\eta_1,c,a}(U) = 2\Phi(\sqrt{2}(c - |u - \frac{a}{2(U+\eta_1)}|)) - 1)$.

Proof: See *appendix*.

Theorem 3.1.1. *For n=2, there exist a for each θ_i such that $P(C_1) \ge P(C_0)$.*

Proof: See *appendix*.

We use the following lemma to develop a lower bound for $P(C_1)$.

Lemma 3.1.6. *If $a \le SS$, then $|(1-R^+)(X_i - \overline{X})| = |\frac{a(V_i}{SS}| \le \sqrt{\frac{a \times n-1}{n}} \le \sqrt{a}$, where $SS = \sum_i^n V_i^2$ and V_i is defined in 3.6.*

Proof:

$SS = \sum_i^n V_i^2$, and $V_i = -\sum_{j \ne i} V_j$. Consequently, by the Cauchy-Scharwtz inequality

$$\sum_{j \ne i} V_j^2 \ge \frac{(\sum_{j \ne i} V_j^2}{n-1} = \frac{V_i^2}{n-1}$$

Moreover,

$$SS = \sum_{j \neq i} V_j^2 + V_i^2 \geq \frac{V_i^2}{n-1} + V_i^2 = V_i^2 \frac{n}{n-1}$$
$$Hence,$$
$$|V_i| \leq \sqrt{\frac{SS \times n - 1}{n}}.$$

Then for $a \leq SS$,

$$\begin{aligned}
|(1 - R^{+})(V_i)| &= \frac{a|V_i|}{SS} \\
&\leq \frac{a\sqrt{\frac{SS \times (n-1}{n}}}{SS} \\
&\leq \sqrt{\frac{a}{SS}} \times \sqrt{\frac{a \times (n-1)}{n}} \\
&\leq \sqrt{\frac{a \times (n-1)}{n}} \leq \sqrt{a}.
\end{aligned}$$

completes the proof.

The following theorem states the our lower bound for $P(C_1)$.

Theorem 3.1.2. *The lower bound* $P(C_1)$ *is*

$$P(C_{L1}) = P(-c + \sqrt{a} \leq Z_i \leq c - \sqrt{a}, i = 1, ..., n, a \leq SS).$$

Proof:

Z_i and $\overline{Z}$ are defined in 3.7 and 3.8, V_i is defined in 3.6.

$$\begin{aligned}
P(C_1) &= P\left\{|\overline{Z}-\eta_i| \le c, i=1,...,n, a \le SS\right\} + \\
&\quad P\left\{|R^{+}U_i+\overline{Z}-(1-R^{+})\eta_i| \le c, i=1,...,n, a \ge SS\right\} \\
&\ge P\left\{|R^{+}U_i+\overline{Z}-(1-R^{+})\eta_i| \le c, i=1,...,n, a \le SS\right\} \\
&\ge P\left\{|U_i-\frac{a}{SS}U_i+\overline{Z}-\frac{a}{SS}\eta_i| \le c, i=1,...,n, a \le SS\right\} \\
&\ge P\left\{|Z_i-\overline{Z}-\frac{a}{SS}U_i+\overline{Z}-\frac{a}{SS}\eta_i| \le c, i=1,...,n, a \le SS\right\} \\
&\ge P\left\{|Z_i-\frac{a}{SS}(U_i+\eta_i)| \le c, i=1,...,n, a \le SS\right\} \\
&\ge P\left\{|Z_i-\frac{a}{SS}V_i| \le c, i=1,...,n, a \le SS\right\} \\
&\ge P\left\{-c+\frac{a}{SS}V_i \le Z_i \le c+\frac{a}{SS}V_i, i=1,...,n, a \le SS\right\}.
\end{aligned}$$

Then from Lemma 3.1.6,

$$\begin{aligned}
P(C_1) &\ge P\left\{-c+\frac{a}{SS}V_i \le Z_i \le c+\frac{a}{SS}V_i, i=1,...,n, a \le SS\right\} \\
&\ge P\left\{-c+\sqrt{a} \le Z_i \le c-\sqrt{a}, i=1,...,n, a \le SS\right\}
\end{aligned}$$

Keep in mind that from Lemma 3.1.6 also implies that
$P(C_{L1}) \le P\left\{-c+\sqrt{\frac{a\times n-1}{n}} \le Z_i \le c-\sqrt{\frac{a\times n-1}{n}}, i=1,...,n, a \le SS\right\} \le P(C_1)$.

We need the following lemma for the last theorem.

Lemma 3.1.7. $n \to \infty \Rightarrow c \to \infty$

Proof:

$$\begin{aligned}
c &= \Phi^{-1}(\frac{\beta^{1/n}+1}{2}) \ and \\
\frac{\beta^{1/n}+1}{2} &= \Phi\Phi^{-1}(\frac{\beta^{1/n}+1}{2}) \\
\frac{\partial}{\partial n}\frac{\beta^{1/n}+1}{2} &= \frac{\partial}{\partial n}\Phi\Phi^{-1}(\frac{\beta^{1/n}+1}{2}) \\
-\frac{\beta^{1/n}\ln\beta}{2n^2} &= \phi\left[\Phi^{-1}(\frac{\beta^{1/n}+1}{2})\right] \times (\Phi^{-1}(\frac{\beta^{1/n}+1}{2}))' \\
(\Phi^{-1}(\frac{\beta^{1/n}+1}{2}))' &= \frac{\frac{-\beta^{1/n}\ln\beta}{2n^2}}{\phi\left[\Phi^{-1}(\frac{\beta^{1/n}+1}{2})\right]} \\
c' &= \frac{\frac{-\beta^{1/n}\ln\beta}{2n^2}}{\phi\left[\Phi^{-1}(\frac{\beta^{1/n}+1}{2})\right]} \qquad (3.17)
\end{aligned}$$

$c^{'} \geq 0$ implies that c is an increasing function of n, therefore $n \to \infty \Rightarrow c \to \infty$.

Lemma 3.1.8. $\frac{c}{n} \to 0$ *if* $n \to \infty$ *where* $c = \Phi^{-1}(\frac{0.95^{1/n}+1}{2})$.

Proof:

Observe that,

$$\begin{aligned}
c &= \Phi^{-1}(\frac{\beta^{1/n}+1}{2}) \\
\Phi(c) &= \frac{\beta^{1/n}+1}{2} \\
\frac{1}{n}Ln\beta &= \ln(2\Phi(c)-1) \\
n &= \frac{\ln\beta}{\ln(2\Phi(c)-1)}. \qquad (3.18)
\end{aligned}$$

Then,

$$\begin{aligned}
\lim_{n\to\infty}\frac{c}{n} &= \lim_{c\to\infty}\frac{c}{\frac{\ln\beta}{\ln(2\Phi(c)-1)}} \\
&= \lim_{c\to\infty}\frac{1}{\ln\beta}\frac{\ln(2\Phi(c)-1)}{\frac{1}{c}} = \frac{0}{0}
\end{aligned}$$

we will apply the L'Hospital rule. To do that we need to take derivative of the both numerator and denominator separately.

$$\begin{aligned}\frac{\partial c}{\partial n} &= \frac{-1}{\ln\beta}\frac{2\phi(c)}{\frac{2\Phi(c)-1}{\frac{1}{c^2}}} = \frac{-1}{\ln\beta}\frac{2c^2}{(2\Phi(c)-1)(1/\sqrt{2\pi})\exp(c^2)} \\ &\to 0\end{aligned}$$

completes the proof.

Lemma 3.1.9.

$$\begin{aligned} i-) \quad \lim_{n\to\infty}\frac{\exp(c^2)}{n} \to \infty \\ ii-) \quad \lim_{n\to\infty}\frac{\exp(0.5c^2)}{n} \to 0 \end{aligned}$$

Proof:i-)

$$\begin{aligned}\lim_{n\to\infty}\frac{\exp(c^2)}{n} &= \lim_{c\to\infty}\frac{exp(c^2)}{\frac{\ln\beta}{\ln(2\Phi(c)-1)}} = \lim_{c\to\infty}\frac{1}{\ln\beta}\frac{\ln(2\Phi(c)-1)}{\frac{1}{exp(c^2)}} \\ &= \frac{0}{0}\end{aligned}$$

Applying the L'hospital rule,

$$\begin{aligned}\lim_{n\to\infty}\frac{\exp(c^2)}{n} &= \lim_{c\to\infty}\frac{-1}{\ln\beta}\frac{\frac{2\phi(c)}{(2\Phi(c)-1)}}{\frac{2c}{\exp(c^2)}} \\ &= \lim_{c\to\infty}\frac{-1}{(2\Phi(c)-1)\ln\beta}\frac{\sqrt{2/\pi}\exp(0.5c^2)}{2c} \\ &\to \infty\end{aligned}$$

completes the first part.

ii-)

By similar argument and applying L'hospital rule, it is easy to see that;

$$\begin{aligned}\lim_{n\to\infty}\frac{\exp(0.5c^2)}{n} &= \lim_{c\to\infty}\frac{-1}{\ln\beta}\frac{\frac{2\phi(c)}{(2\Phi(c)-1)}}{\frac{c}{\exp(0.5c^2)}} \\ &= \lim_{c\to\infty}\frac{-1}{(2\Phi(c)-1)\ln\beta}\frac{\sqrt{2/\pi}}{c} \\ &\to 0\end{aligned}$$

completes the second part.

Lemma 3.1.10. $\lim_{n\to\infty} \frac{\sqrt{n}}{c} \to 0$

Proof:

$$\begin{aligned}
\lim_{n\to\infty} \frac{\sqrt{n}}{c} &= \lim_{n\to\infty} \frac{\sqrt{n}}{\Phi^{-1}(\frac{\beta^{1/n}+1}{2})} \\
&\to \frac{\infty}{\infty}
\end{aligned}$$

Applyin L'hospital rule and using the equation 3.17,

$$\begin{aligned}
\lim_{n\to\infty} \frac{\sqrt{n}}{\Phi^{-1}(\frac{\beta^{1/n}+1}{2})} &= \lim_{n\to\infty} \frac{0.5n^{-1/2}}{\frac{\frac{-\beta^{1/n}\ln\beta}{2n^2}}{\phi\left[\Phi^{-1}(\frac{\beta^{1/n}+1}{2})\right]}} \\
&= \lim_{n\to\infty} \frac{n^{3/2}\phi\left[\Phi^{-1}(\frac{\beta^{1/n}+1}{2})\right]}{-\beta^{1/n}Ln\beta} \\
&\to \infty
\end{aligned}$$

completes the proof.

The usual recommendation for the shrinkage constant,a, is $n-3$. Lemma 3.1.10 implies that $\lim_{n\to\infty} \frac{\sqrt{a}}{c} \to \infty$. This means that a has to be bounded above by a constant. In the following proofs, we assume that a is bounded above by a constant,m, such that $a \leq m < n$.

Lemma 3.1.11. *For a large enough* n ,$P(C_{L1}) = 2(\Phi(c-\sqrt{a})-1)^n$.

Proof:

Then from the theorem 3.1.2,

$$P(C_1) \geq P\left\{-c+\sqrt{a} \leq Z_i \leq c-\sqrt{a}, a \leq SS\right\}$$

if $\lim_{sup} \frac{a}{n} \leq 1$, then $P(SS \geq a) = 1$, because SS has a non central chi squared distribution that stochastically increasing in n.

Then let $E : \{-c+\sqrt{a} \leq Z_i \leq c-\sqrt{a}, i = 1, ..., n\}$ and $F = \{a \leq SS\}$. By given $P(F^c) = 0$. Then

$$\begin{aligned}
P(C_1) &\geq P\left\{-c+\sqrt{a} \leq Z_i \leq c-\sqrt{a}, a \leq SS\right\} \\
&\geq P(E) - P(F^c \cap E) \\
&\geq P(-c+\sqrt{a} \leq Z_i \leq c-\sqrt{a}) - P(a \geq SS) \\
&\geq P(-c+\sqrt{a} \leq Z_i \leq c-\sqrt{a}) = 2(\Phi(c-\sqrt{a})-1)^n
\end{aligned}$$

Theorem 3.1.3. *$P(C_1)$ dominate $P(C_0)$ up to an arbitrarily small constant for a sufficiently large number of means, n.*

Proof:

Observe that for a large enough n,$P(SS \leq a) \to 0$, since $\lim_{sup} \frac{a}{n} = 0$. Then,

$$P(C_1) = P\left\{\eta : |R^{+}U_i + \overline{Z} - (1-R^{+})\eta_i| \leq c, SS \geq a\right\}.$$

From Theorem 3.1.2,

$P(C_1) \geq (2\Phi(c-\sqrt{a})-1)^n$, then Let $\Delta = (2\Phi(c)-1)^n - (2\Phi(c-\sqrt{a})-1)^n$. We show that $\lim_{n\to\infty} \Delta = 0$. Keep in mind that $\lim_{n\to\infty}(2\Phi(c)-1)^n = 0.95$. Then,

$$\begin{aligned}
\lim_{n\to\infty} \Delta &= \lim_{n\to\infty} (2\Phi(c)-1)^n [1 - \frac{(2\Phi(c-\sqrt{a})-1)^n}{(2\Phi(c)-1)^n}] \\
&= 0.95 \times [1 - \lim_{n\to\infty} \frac{(2\Phi(c-\sqrt{a})-1)^n}{(2\Phi(c)-1)^n}]
\end{aligned}$$

Also observe that,

$$\lim_{n\to\infty} \tfrac{(2\Phi(c-\sqrt{a})-1)^n}{(2\Phi(c)-1)^n} = \exp^{\lim_{n\to\infty} n \ln \frac{(2\Phi(c-\sqrt{a})-1)}{(2\Phi(c)-1)}}.$$

Then,

$$\begin{aligned}
\lim_{n\to\infty} n \ln \frac{(2\Phi(c-\sqrt{a})-1)}{(2\Phi(c)-1)} &= \lim_{n\to\infty} \frac{\ln\beta}{\ln(2\Phi(c)-1)} \ln \frac{(2\Phi(c-\sqrt{a})-1)}{(2\Phi(c)-1)} \\
&= \lim_{n\to\infty} \frac{\ln\beta}{\ln(2\Phi(c)-1)} \left[\ln(2\Phi(c-\sqrt{a})-1) - \ln(2\Phi(c)-1)\right] \\
&= \lim_{n\to\infty} \ln\beta \left[\frac{\ln(2\Phi(c-\sqrt{a})-1)}{\ln(2\Phi(c)-1)} - 1\right] \\
&\to 0
\end{aligned}$$

Therefore, $\lim_{n\to\infty} \frac{(2\Phi(c-\sqrt{a})-1)^n}{(2\Phi(c)-1)^n} \to 1$. This implies that $\lim_{n\to\infty} \Delta = 0.95 \times [1-1] = 0$.

We showed that the rectangular confidence intervals centered around a shrinkage estimator have coverage probability uniformly exceeding that of the usual rectangular confidence regions up to an arbitrarily small epsilon when the number of means is sufficiently large. We also showed that these intervals have strictly greater coverage probability when all the main effects are zero, and that the coverage probability of the two procedures converge to one another when at least one of the main effects becomes arbitrarily large. We also developed a lower bound for the coverage probability of the rectangular confidence region which is a decreasing function of the shrinkage constant.

3.2 UNKNOWN σ^2 CASE

The usual confidence rectangle for the vector Θ of cell means is

$$C_0 = \{\Theta = (\theta_1, ..., \theta_n) : |X_i - \theta_i| \leq c \ , i = 1, ..., n\}, \tag{3.19}$$

where $c = |m|_{\alpha,n,v} \frac{\hat{\sigma}}{\sqrt{n}}$, and $|m|_{\alpha,n,v}$ is the $1-\alpha$ quantile of the Student maximum modulus statistics, $\hat{\sigma}^2$ is the unbiased estimator of σ^2. It was stated by most of the statistical text book that when n goes to infinity, the quantile of Student maximum modulus statistics can be replaced by the quantile of student t distribution.

We again use the shrinkage estimator with Lindley's shrinkage factor and replace σ with an unbiased estimator, $\hat{\sigma}^2 = \frac{\sum_{i=1}^n \sum j=1^k (Y_{ij}-X_i)^2}{n(k-1)}$. Our confidence interval is the same as (3.10), the only difference is the shrinkage estimator. The confidence interval is,

$$C_1(X) = \left\{\theta = (\theta_1, ..., \theta_k) : |R^+(X_i - \overline{X}) + \overline{X} - \theta_i| \leq c)\right\} \tag{3.20}$$

where R^+ is

$$R^+ = (1 - \frac{a\hat{\sigma}^2}{\sum_{i=1}^k (X_i - \overline{X})^2})^+. \tag{3.21}$$

That is,

$$R^+ = \begin{cases} 1 - \frac{a\hat{\sigma}^2}{\sum_{i=1}^n (X_i - \overline{X})^2} & : \quad \frac{\sum_{i=1}^n (X_i - \overline{X})^2}{\hat{\sigma}^2} \geq a \\ 0 & : \quad \frac{\sum_{i=1}^n (X_i - \overline{X})^2}{\hat{\sigma}^2} \leq a. \end{cases}$$

By using same notations in known σ case, (3.20) becomes

$$C_1(X) = \{\eta = (\eta_1, ..., \eta_n) : |R^+(U_i) + \overline{Z} - (1 - R^+)\eta_i| \leq c)\} \tag{3.22}$$

where

$$R^+ = (1 - \frac{a\hat{\sigma}^2}{\sum_{i=1}^n (U_i + \eta_i)^2})^+.$$

Now, we try to derive same analytical results for the unknown σ case.

3.2.1 Analytical Results

Lemma 3.2.1. $\limsup_{\eta'\eta\to\infty} \frac{a\hat{\sigma}^2}{SS} \leq 1$

Proof:

Let $\gamma_1 = \sqrt{\sum_{i=1}^n \eta_i^2}$, $\epsilon_i = \frac{\eta_i}{\gamma_1}$ and similarly let $\gamma_2 = \sqrt{\sum_{i=1}^n \sum_{j=1}^k \eta_{ij}^2}$, $\epsilon_{ij} = \frac{\eta_{ij}}{\gamma_2}$, where $\eta_i = \theta_i - \overline{\theta}$, and $\eta_{ij} = \theta_{ij} - \theta_i$. Observe that $\sum_{i=1}^n \epsilon_i^2 = 1$, and $\sum_{i=1}^n \sum_{j=1}^k \epsilon_{ij}^2 = 1$.Keep in mind that $\sum_{i=1}^n \eta_i^2 \to \infty$ implies that $\gamma_1, \gamma_2 \to \infty$.

$$\begin{aligned} \limsup_{\eta'\eta\to\infty} \frac{a\hat{\sigma}^2}{SS} &= \limsup_{\gamma_1,\gamma_2\to\infty} \frac{a\frac{\sum_{i=1}^n \sum_{j=1}^k (U_{ij} + \epsilon_{ij}\gamma_2)^2}{n(k-1)}}{\sum_{i=1}^n (U_i + \epsilon_i\gamma_1)^2} \\ &= \limsup_{\gamma_1,\gamma_2\to\infty} a\frac{\gamma_2^2}{\gamma_1^2} \frac{\frac{\sum_{i=1}^n \sum_{j=1}^k (\frac{U_{ij}}{\gamma_2} + \epsilon_{ij})^2}{n(k-1)}}{\sum_{i=1}^n (\frac{U_i}{\gamma_1} + \epsilon_i)^2} \\ &= \frac{a}{n(k-1)} \frac{\sum_{i=1}^n \sum_{j=1}^k \epsilon_{ij}^2}{\sum_{i=1}^n \epsilon_i^2} \\ &= \frac{a}{n(k-1)} \leq 1. \end{aligned}$$

The following lemma proves that, the coverage probability of two procedures converge to one another when at least one of the main effects becomes arbitrarily large.

Lemma 3.2.2. *if* $\eta^{'}\eta \to \infty$, $P(C_1) \to P(C_0)$.

Proof:

Let $\gamma = \sqrt{\eta'\eta}$ and $\epsilon_i = \frac{\eta_i}{\gamma}$. Then $SS = \sum(U_i + \eta_i)^2 = \sum(U_i + \epsilon_i\gamma)^2 \to \infty$. Consider $P(C_1)$,

$$P(C_1) = P(|\overline{Z} - \epsilon_i\gamma| \leq c, i = 1, ..n, \frac{SS}{\hat{\sigma}^2} \leq a) + P(|RU_i + \overline{Z} - (1-R)\epsilon_i\gamma \leq c, i = 1, .., n\frac{SS}{\hat{\sigma}^2} \geq a))$$

Observe that as γ goes to infinity, $P(|\overline{Z} - \epsilon_i\gamma| \leq c, i = 1, .., n)$ goes to zero. Therefore $P(C_1)$ becomes

$$P(C_1) \;=\; P(|U_i - \frac{a\hat{\sigma}^2(U_i + \epsilon_i\gamma)}{\sum_{i=1}^{n}(U_i + \epsilon_i\gamma)^2} + \overline{Z}| \leq c, i = 1, ..., n, \frac{SS}{\sigma^2} \geq a).$$

From the previous lemma $\lim_{\gamma\to\infty} P(\frac{SS}{\sigma^2} \geq a) = 1$, then

$$\begin{aligned}
P(C_1) &= P(|U_i - \frac{a\hat{\sigma}^2(U_i + \epsilon_i\gamma)}{\sum_{i=1}^{n}(U_i + \epsilon_i\gamma)^2} + \overline{Z}| \leq c) \\
P(C_1) &= P(\max_i(-c - U_i + \frac{a\hat{\sigma}^2(U_i + \epsilon_i\gamma)}{\sum_{i=1}^{n}(U_i + \epsilon_i\gamma)^2}) \leq \overline{Z} \leq \min_i(c - U_i + \frac{a\hat{\sigma}^2(U_i + \epsilon_i\gamma)}{\sum_{i=1}^{n}(U_i + \epsilon_i\gamma)^2})) \\
&= \int (T\left[\min_i(c - G(U_i, \epsilon_i\gamma)\right] - T\left[\max_i(-c - G(U_i, \epsilon_i\gamma)\right]) f_U(U)dU
\end{aligned}$$

where

$$\begin{aligned}
G_1(U_i, \epsilon_i\gamma) &= c - \frac{U_i\sqrt{n}}{\hat{\sigma}^2} - \frac{a\sqrt{n}(Ui + \epsilon_i\gamma)}{\sum_{i=1}^{n}(U_i + \epsilon_i\gamma)^2} \\
G_2(U_i, \epsilon_i\gamma) &= -c - \frac{U_i\sqrt{n}}{\hat{\sigma}^2} - \frac{a\sqrt{n}(Ui + \epsilon_i\gamma)}{\sum_{i=1}^{n}(U_i + \epsilon_i\gamma)^2}
\end{aligned}$$

and T is the cdf of t distribution.

Since $T\left[\min_i(G_1(U_i, \epsilon_i\gamma)\right] - T\left[\max_i(G_2(U_i, \epsilon_i\gamma)\right] \leq 1$ and $f_{(U)}(U)$ is absolutely continuous,

$\lim_{\gamma\to\infty} P(C_1) = \int (T\left[\lim_{\gamma\to\infty}(\min_i(G_1(U_i, \epsilon_i\gamma))\right] - T\left[\lim_{\gamma\to\infty}(\max_i(G_2(U_i, \epsilon_i\gamma))\right]) f_U(U)dU.$

Observe that $\lim_{\gamma\to\infty} \frac{aUi}{\sum_{i=1}^{k}(U_i+\epsilon_i\gamma)^2} \to 0$ and $\lim_{\gamma\to\infty} \frac{a\epsilon_i\gamma}{\sum_{i=1}^{k}(U_i+\epsilon_i\gamma)^2} \to 0$. Then $P(C_1)$ becomes

$$
\begin{aligned}
P(|U_i+\overline{Z}| \le c \ \ , i=1,..,n) &= P(|Z_i-\overline{Z}+\overline{Z}| \le c \ , i=1,...,n) \\
&= P(|Z_i|, i=1,...,n) \\
&= P(C_0)
\end{aligned}
$$

Lemma 3.2.3. *For $\eta'\eta = 0$, or for each $\eta_i = 0 \ , i=1,...,n$, $P(C_1) \ge P(C_0)$*

Proof:

$P(C_1) = E_{\hat{\sigma}}P(C_1/\hat{\sigma})$, then consider $P(C_1/\hat{\sigma})$,

$$
\begin{aligned}
P(C_1/\hat{\sigma}) &= \int T(c-R^+U_{[1]}) - T(-c+R^+U_{[n]}) f_{U_{[1]}}(U_{[1]}) f_{U_{[n]}}(U_{[n]}) dU_{[1]} dU_{[n]} \\
&\ge \int T(c-U_{[1]}) - T(-c+U_{[n]}) f_{U_{[1]}}(U_{[1]}) f_{U_{[n]}}(U_{[n]}) dU_{[1]} dU_{[n]}
\end{aligned}
$$

where T is the probability density function for $\overline{Z}$, $U_{[1]} = \max_i U_i$ and $U_{[1]}$ must be positive ,$U_{[n]} = \min_i U_i$ and $U_{[n]}$ must be negative since $\sum_i U_i = 0$, and $0 \le R^+ \le 1$. It follows that,

$$
\begin{aligned}
P(C_1) &= E_{\hat{\sigma}}P(C_1/\hat{\sigma}) \\
&\ge E_{\hat{\sigma}} \int T(c-U_{[n]}) - T(-c+U_{[1]}) f_{U_{[1]}(U_{[1]})} f_{U_{[n]}}(U_{[n]}) dU_{[1]} dU_{[n]} \\
&= P(C_0).
\end{aligned}
$$

Theorem 3.2.1. *A lower bound $P(C_1)$ is*

$$P(C_{L1}) = P(-c+\sqrt{a}\hat{\sigma} \le Z_i \le c-\sqrt{a}\hat{\sigma}, a\hat{\sigma}^2 \le SS)$$

, where $X_i = Z_i + \eta_i$.

Proof:

If $a\hat{\sigma}^2 \leq SS$, then consider $\frac{a\hat{\sigma}^2}{SS}(X_i - \overline{X})$,

$$
\begin{aligned}
|\frac{a\hat{\sigma}^2}{SS}(X_i - \overline{X})| &\leq \frac{a\hat{\sigma}^2}{SS} \max_i |X_i - \overline{X}| \\
&\leq \sqrt{\frac{a\hat{\sigma}^2}{SS}} \sqrt{a\hat{\sigma}^2} \sqrt{\frac{\max_i (X_i - \overline{X})^2}{SS}} \\
&\leq \sqrt{a\hat{\sigma}^2} = \sqrt{a}\hat{\sigma}
\end{aligned}
$$

Observe that if $n \to \infty$, the quantile of maximum modulus can be replaced by the quantile of either t or z distribution. When we prove that the rectangular confidence intervals centered around a shrinkage estimator have coverage probability uniformly exceeding that of the usual rectangular confidence regions up to an arbitrarily small epsilon when the number of means is sufficiently large, we make use of that fact.

Lemma 3.2.4. $\lim_{n\to\infty} P(SS \geq a\hat{\sigma}^2) = 1$.

Proof:

As $n \to \infty$, $\hat{\sigma} \to \sigma$. SS has a non central chi-square distribution ν degrees of freedom, let SS_1 has a central chi-square distribution with ν degrees of freedom. Then,

$$
\lim_{n\to\infty} P(SS \geq a\hat{\sigma}^2) = P(a\sigma \leq SS) = 1 - P(a\sigma \geq SS) \geq 1 - P(a\sigma \geq SS_1) = 1
$$

since a is bounded above by a constant, $\lim_{sup} \frac{a}{n} = 0$.

Theorem 3.2.2. *$P(C_1)$ dominate $P(C_0)$ up to an arbitrarily small constant for a sufficiently large number of means, n.*

Proof: From the previous theorem and lemma,

$$
\begin{aligned}
P(C_1) \geq P(C_{L1}) &= P(|Z_i| \leq c - \sqrt{a}\hat{\sigma}) \\
&= P(|Z_i| \leq c(1 - \frac{\sqrt{a}\hat{\sigma}}{c}))
\end{aligned}
$$

Observe that as $n \to \infty$,

$$
\begin{aligned}
\hat{\sigma} &\to \sigma \\
\frac{\sqrt{a}}{c} &\to 0 \\
(1 - \frac{\sqrt{a}\hat{\sigma}}{c}) &\to 1.
\end{aligned}
$$

Moreover,

$$
\lim_{n\to\infty} P(C_{L1}) = \lim_{n\to\infty} E_{\hat{\sigma}} P(|Z_i| \leq c - \sqrt{a}\hat{\sigma}/\hat{\sigma}) \tag{3.23}
$$

$\hat{\sigma}$ has a chi-square distribution and $|P(|Z_i| \leq c - \sqrt{a}\hat{\sigma})| \leq 1$ is bounded. Therefore, all the conditions for dominated convergence theorem are met, we can take the limit inside of expectation.

$$
\begin{aligned}
\lim_{n\to\infty} E_{\hat{\sigma}} P(|Z_i| \leq c - \sqrt{a}\hat{\sigma}/\hat{\sigma}) &= E_{\hat{\sigma}} \lim_{n\to\infty} P(|Z_i| \leq c - \sqrt{a}\hat{\sigma}/\hat{\sigma}) \\
&= P(C_0).
\end{aligned}
$$

For both a known and an unknown σ cases, we proved that the rectangular confidence intervals centered around a shrinkage estimator have coverage probability uniformly exceeding that of the usual rectangular confidence regions up to an arbitrarily small epsilon when the number of means is sufficiently large. We showed that these intervals have strictly greater coverage probability when all the parameters are zero, and that the coverage probability of two procedures converge to one another when at least one of the parameters becomes arbitrarily large. We also developed a lower bound for the coverage probability of the rectangular confidence region which is a decreasing function of the shrinkage constant.

3.3 SIMULATION RESULTS

Since we want to show that the coverage probability for our rectangular confidence procedure, $P(C_1)$, is uniformly higher than the coverage probability for the usual rectangular confidence procedure, $P(C_0)$, for small number of means, and all the main effects are small, we run the simulation for $n = 2, .., 10$ and we generate the main effects, η, from the following set $H = (-4, -3.75, -3.5, ..., 3.5, 3.75, 4)$.

We can write the coverage probability of our procedure the following way,

$P(C_1) = P\left\{\eta = (\eta_1, ..., \eta_n) : |R^+(Z_i) + \overline{Z} - (1 - R^+)\eta_i| \le c)\right\}$,

where $Z_i's$ are $i.i.d$ $N(0, 1)$, $\overline{Z}$ is $N(0, 1/n)$, and $c = \Phi^{-1}(\frac{0.95^{1/n}+1}{2})$.

Lemma 3.3.1. *$P(C_1)$ is sign invariant.*

Lemma 3.3.2. *$P(C_1)$ is permutation invariant.*

To do simulations, we followed these steps;

Step 1: Generate Z_n^i $i = 1, ..., 10,000$ and store them where Z_n is a $n \times 1$ column matrix.

Step 2: Generate all the possible η_n^j by taking sign and permutation invariance into account from the set above, H, where $j = 1, ...K$, and K is the total number of η_n generated, and η_n is a $n \times 1$ column matrix.

Step 3: Calculate c.

Step 4: For each $j = 1, .., K$, calculate

$$PS_{ij} = (Z_n^i - \overline{Z_n^i})R_i^+ - (1 - R_i^+)\eta_n^i \ for\ each\ i = 1, ..., 10,000.$$

Step 5: For $j = 1, ..., K$ and $i = 1, ..., 10,000$, let $CP_n^{ij} = 1$ if all the values in PS_{ij} are in [-c,c] and 0 o.w., where CP_n^{ij} is a $n \times 1$ column matrix.

Step 6: Calculate the coverage probability for $k = 1, .., K$, $P_k = \sum_{i=1}^{10,000} \frac{CP_n^{ik}}{10,000}$

(See Appendix for the R codes).

Plotting the coverage probabilities against the length of η makes sense for the spherical confidence regions because Casella and Hwang proved that the coverage probability of the spherical region depends only on the length of η. That is not the case for rectangular confidence regions. To have a better understanding of our procedure, we also plot the coverage probabilities against the maximum of $|\eta|$.

For n=2, we try to find the optimum range of the shrinkage constant, a, such that $P(C_1)$ dominates $P(C_0)$. To find the optimum a, we ran simulations for a between 0 and 2 with 0.1 increments. The first thing we notice from simulations for a, when η is zero, is that $P(C_1)$ is higher than $P(C_0)$ and $P(C_1)$ achieves its maximum for any choice of a. Secondly, we notice from the simulation for a that coverage probabilities for our procedure are increasing in η until $\eta = 1$, then the coverage probabilities are decreasing. We could not find a universal a for all the η's such that $P(C_1)$ is higher than $P(C_0)$ in every region. Therefore $P(C_1)$ is not uniformly higher than $P(C_0)$ for n=2. We also plot the coverage probabilities against the length of η for each a. One thing that is common for all values of a is that the coverage probabilities are decreasing until the mid values of the length of η then it starts increasing again. The sharpness of this dip depends on choice of a. As a increases, this dip becomes sharper. (See Figure 8.1 - 8.8)

For n=3, Lindley proved that our shrinkage estimator, fittingly namely Lindley's estimator, did not have a smaller TMSE than the usual estimator. First we want to see if $P(C_1)$ is uniformly higher than $P(C_0)$ for n=3. If so, we want to find the optimum range for a. We run the simulations for different choices of a. Again, when η is around zero, $P(C_1)$ dominates $P(C_0)$ and $P(C_1)$ achieves its maximum when $\eta = 0$ for any choice of a. As in n=2, we plot the coverage probabilities against the length of η for each a. Again the coverage probabilities are decreasing until the mid values of the length of η then it starts increasing again. The sharpness of this dip depends on choice of a. As a gets bigger, this dip becomes sharper. When a is less than 0.05, $P(C_1)$ is uniformly higher than $P(C_0)$. Therefore the optimum range for a is [0,0.05]. The coverage probabilities are not a decreasing function of the length of η for n=3. Since a at 0.05 is very small, we are not gaining very much by using the shrinkage estimator. Therefore we agree with Lindley's result. Since our simulations do not indicate the domination of our procedure over the usual one, we do not plot the coverage probabilities against the $max_i|\eta_i|$(See Figure 8.9 - 8.16)

For n=4, we first plot the coverage probabilities against the length of η for each a, again there is a dip and coverage probabilities are not decreasing function of length of η. Then we plot the coverage probabilities against the maximum $|\eta|$, again there is a dip but the dip is not as sharp as in the first graph. As a gets smaller, this dip is getting smaller like in the first graph. When η is around zero, $P(C_1)$ is uniformly higher than $P(C_0)$ and $P(C_1)$ achieves its maximum at $\eta = 0$ for any choice of a . When a is less than equal to 1, $P(C_1)$ is uniformly higher than $P(C_0)$ everywhere. Therefore the optimum choice for a is 1. The usual recommendation for a is n-3. We agree with the usual recommendation. (See Figure 8.17 - 8.21)

For n=5, we again try to find the optimum shrinkage constant, a. Again, when η is around zero, $P(C_1)$ is uniformly higher than $P(C_0)$ for any choice of a in every region and $P(C_1)$ achieves its maximum at $\eta = 0$. When a is less than equal to 2, $P(C_1)$ is uniformly higher than $P(C_0)$ everywhere. Therefore the optimum choice for a is 2. Again, we agree with the usual recommendation. As in n=4, we plot the coverage probabilities first against the length of η then against the maximum $|\eta|$'s. We observe similar pictures, there is a dip and the sharpness of that dip depends on a. The dip in the second graph is not as sharp as the dip in the first graph. (See Figure 8.22)

For n equal 6,7,8,9, and 10, the coverage probabilities are not a decreasing function of the length of η. We have similar pictures for each n. We again plot the coverage probabilities first against the length η then against the maximum$|\eta|$. Again, there is a dip in the graph and the dip depends on a. The dip in the first graph is sharper than that of the second graph. When $\eta = 0$, $P(C_1)$ achieves its maximum and is uniformly higher than $P(C_0)$, for every choice of a. However the usual recommendation for a did not work for each n. The following table shows the optimum choice of a for each n. We also check the Casella's recommendation for $a = 0.8(n-2)$, even tough Casella's recommendation is for James-Stein shrinkage estimator.(Figure 8.23 - 8.25)

Recommendations for a

n	The Usual Recommendation	Casella's Recommendation	Optimum Choice
2	-	-	-
3	-	0.8	0.05
4	1	1.6	1
5	2	2.4	2
6	3	3.2	2
7	4	4	3
8	5	4.8	4
9	6	5.6	4
10	7	6.4	5

Table 3.1: *Table for the Shrinkage Constant for One-Way ANOVA model*

In conclusion, our simulations indicate that $P(C_1)$ is uniformly higher than $P(C_0)$, for $n = 4, ..., 10$. We plot the coverage probabilities first against the length of η , then against the maximum of the $|\eta|$ for each n. We have similar pictures in both graphs. The only difference is the second graph is smoother than the first graph. There is a small dip, but the dip is a function of a, when a gets smaller, the dip is getting smaller. The coverage probabilities are not a decreasing function of either the length of η nor the maximum $|\eta|$. We also did not agree with the usual recommendation for a for $n \geq 6$. Based on our simulation, our recommendation for a is $[0.6(n-2)]$ for $n \geq 4$, where $[X]$ is the nearest integer function.

4.0 MULTIPLE COMPARISON WITH CONTROL

Dunnett (1955) stated that when a control is present, the comparisons of primary interest may be the comparison of each treatment mean with the mean of a control. For example, the control may be a placebo, or it may be a standard treatment. We call such comparisons multiple comparison with a control. We are interested in simultaneous confidence intervals for the multiple comparison with a control in the balanced one-way ANOVA model with known σ^2. We give the modification of our procedure for the balanced one-way ANOVA model with unknown σ^2 and the unbalanced one-way ANOVA model with both known or unknown σ^2 in *Appendix*. The sample cell means, $X_1, ..., X_n$, defined in (2.2) are independent normally distributed random variables with means $\theta_1, ..., \theta_n$ and variance 1. The sample control mean is normally distributed with mean θ_c and variance 1, independent of the sample cell means. The usual confidence interval for multiple comparison with control is

$$\begin{aligned} E_0 &= \{|X_i - X_c - (\theta_i - \theta_c)| \leq c^*, i = 1, .., n\} \ \textit{and} \\ P(E_0) &= P\{|Y_i| \leq c^*, i = 1, ..., n\} = 0.95. \end{aligned} \tag{4.1}$$

Where $Y_i = X_i - X_c - \theta_i + \theta_c$, $Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$ has normal distribution $N(0, \Sigma)$, $\Sigma = \begin{pmatrix} 2 & 1 & \cdots & 1 & 1 \\ 1 & 2 & 1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & \cdots & 1 & 2 \end{pmatrix}$, and c^* is the solution of (4.1).

The confidence interval we are proposing is

$$\begin{aligned} E_1 &= \left\{|R^+(X_i - X_c - \overline{X} + X_c) + \overline{X} - X_c - (\theta_i - \theta_c)| \le c^*, i = 1, .., n\right\} \\ E_1 &= \left\{|R^+U_i + \overline{Z} - Z_c - (1 - R^+)\eta_i| \le c^*, i = 1, ..., n\right\} \end{aligned} \quad (4.2)$$

Where, $U_i = X_i - \overline{X} - \eta_i$,$\eta_i = \theta_i - \overline{\theta}$, $\overline{Z} = \overline{X} - \overline{\theta}$, $Z_c = X_c - \theta_c$ and

$$R^+ = \begin{cases} 1 - \frac{a}{\sum_{i=1}^{k}(X_i - \overline{X})^2} & : \quad \sum_{i=1}^{k}(X_i - \overline{X})^2 \ge a \\ 0 & : \quad \sum_{i=1}^{k}(X_i - \overline{X})^2 \le a. \end{cases}$$

Since 4.1 is very similar to 3.1, we would like to see if there is any kind of relationship between c and c^*. In the following lemma's, we are going to prove that $c^* \ge c$.

Lemma 4.0.3. $c^* \ge \frac{c}{2}$.

Proof:

See *appendix*.

Lemma 4.0.4. $c^* \ge c$.

Proof:

$$\begin{aligned} P(E_0) &= P\left\{|X_i - X_c - (\theta_i - \theta_j| \le c^*, i = 1, .., n\right\} \\ &= P\left\{-c^* + X_c \le X_i \le c^* + X_c, i = 1, ..., n\right\} \\ &= \int \left[\Phi(c^* + X_c) - \Phi(-c^* + X_c)\right]^n \varphi_{(X_c)}(X_c) dX_c \end{aligned}$$

Where $\varphi_{(X_c)}(X_c)$ is standard normal distribution. We will investigate the properties of $[\Phi(c^* + X_c) - \Phi(-c^* + X_c)]^n$, let $\Delta(X_c) = [\Phi(c^* + X_c) - \Phi(c^* + X_c)]^n$

$$\begin{aligned} \frac{\partial \Delta(X_c)}{\partial X_c} &= n\left[\Phi(c^* + X_c) - \Phi(-c^* + X_c)\right]^{n-1} \frac{1}{\sqrt{2\pi}}(e^{-0.5(c^*+X_c)^2} - e^{-0.5(-c^*+X_c)^2}) \\ &= n(n-1)\left[\Phi(c^* + X_c) - \Phi(-c^* + X_c)\right]^{n-2} \left[\frac{1}{\sqrt{2\pi}}(e^{-0.5(c^*+X_c)^2} - e^{-0.5(-c^*+X_c)^2})\right]^2 \end{aligned}$$

$$+ \tfrac{n}{\sqrt{2\pi}}\left[\Phi(c^* + X_c) - \Phi(-c^* + X_c)\right]^{n-1} \left[-(c^* + X_c)e^{-0.5(c^*+X_c)^2} + (-c^* + X_c)e^{-0.5(-c^*+X_c)^2}\right]$$

Observe that for $X_c = 0$,

$\frac{\partial \Delta(X_c)}{\partial X_c} = 0$ and $\frac{\partial^2 \Delta(X_c)}{\partial^2 X_c} = [\Phi(c^*) - \Phi(-c^*)]^{n-1} \frac{1}{\sqrt{2\pi}}(-2c^* e^{-0.5(c^*)^2}) \leq 0.$
This implies that for $X_c = 0$, $\Delta(X_c)$ achieves its maximum. Moreover,
$[\Phi(c^* + X_c) - \Phi(c^* + X_c)]^n \leq [\Phi(c^*) - \Phi(c^*)]^n$.
Then,

$$
\begin{aligned}
P(E_0) = 0.95 &= \int [\Phi(c^* + X_c) - \Phi(-c^* + X_c)]^n \varphi_{(X_c)}(X_c) dX_c \\
0.95 &\leq [\Phi(c^*) - \Phi(-c^*)]^n \\
\Phi^{-1}(\frac{0.95^{1/n} + 1}{2}) &\leq c^* \\
c &\leq c^*.
\end{aligned}
$$

The confidence interval we are proposing is defined in 4.2,
$P(E_1) = P\left\{|R^+U_i + \overline{Z} - Z_c - (1 - R^+)\eta_i| \leq c^*, i = 1, ..., n\right\}$ and let $Z_1 = \overline{Z} - Z_c$.Then our confidence regions becomes,

$$
P(E_1) = P\left\{|R^+U_i + Z_1 - (1 - R^+)\eta_i| \leq c^*, i = 1, ..., n\right\}. \tag{4.3}
$$

4.3 is very similar to 3.10. The only difference is the constant term Z. In 4.3, Z_1 has a normal distribution with mean 0 and variance $1 + \frac{1}{n}$. In 3.10, $\overline{Z}$ has normal distribution with mean 0 and variance $\frac{1}{n}$.

4.1 ANALYTICAL RESULTS

Lemma 4.1.1. *if $\eta' \eta \to \infty$, $P(E_1) \to P(E_0)$.*

Proof:
Again let $\delta = \sqrt{\eta' \eta}$. For $\epsilon > 0$, let $A_\epsilon = \{(1 - R^+) \max_i |X_i| \leq \epsilon\}$. Then from the lemma 3.1.1,

$$
\lim_{\|\delta\| \to \infty} A_\epsilon^c = 0 \tag{4.4}
$$

Note that,

$$
\eta \in E_1 \Leftrightarrow -c^* \leq X_i - Z_c - \theta_i - (1 - R^+)(X_i - \overline{X}) \leq c^*. \tag{4.5}
$$

and $\eta = \theta_i - \overline{\theta}$. Thus

$$\begin{aligned} P(\eta \in E_1) &= P\{\eta \in E_1 \cap A_\epsilon\} + P\left\{\eta \in E_1 \cap A_\epsilon^{c^*}\right\} \\ &\leq P\{-c^* - \epsilon \leq Z_i - Z_c \leq c^* + \epsilon \cap A_\epsilon\} + P\{A_\epsilon^c\} \\ &\leq P\{-c^* - \epsilon \leq Y_i \leq c^* + \epsilon\} + P\{A_\epsilon^c\} \end{aligned} \tag{4.6}$$

Also

$$\begin{aligned} P(\eta \in E_1) &\geq P\{-c^* + \epsilon \leq X_i - Z_c - \theta_i \leq c^* - \epsilon \cap A_\epsilon\} + P\{A_\epsilon^c\} \\ &\geq \{-c^* + \epsilon \leq Z_i - Z_c \leq c^* - \epsilon\} - P\{A_\epsilon^c \cap -c^* + \epsilon \leq Z_i - Z_c \leq c^* - \epsilon\} \\ &\geq P\{-c^* + \epsilon \leq Y_i \leq c^* - \epsilon\} - P\{A_\epsilon^c\} \end{aligned} \tag{4.7}$$

It follows from (3.12)-(3.15) that,

$$(2\Phi(c-\epsilon)-1)^n \leq \liminf_{\delta\to\infty} P(\eta \in C_1) \leq \limsup_{\delta\to\infty} \leq (2\Phi(c+\epsilon)-1)^n$$

However $\epsilon > 0$ is arbitrary. Taking $\epsilon \to 0$ and noting that $\Phi(.)$ is a continuous function of its argument, the assertion of lemma follows.

Lemma 4.1.2. *For $\eta'\eta = 0$, or for each $\eta_i = 0$, $i = 1, ..., n$, $P(E_1) \geq P(E_0)$*

Proof:

For each $\eta_i = 0$, R^+ becomes $R^+ = (1 - \frac{a}{\sum_{i=1}^n U_i^2})^+ and$ $P(E_1)$ becomes

$$\begin{aligned} P(E_1) &= P\left\{\eta : |R^+U_i + Z_1| \leq c^*, i = 1, ..n\right\} \\ &= \int_{(U_1,\ldots,U_n)} \int_{(\sqrt{n/n+1}(-c^*-R^+U_{[n]})}^{(\sqrt{n/n+1}(c^*-R^+U_{[1]})} f_U(U)\varphi_Z(Z)dZdU \end{aligned}$$

where $U_{[1]} = \max_{i=1,...n} U_i$, $U_{[n]} = \min_{i=1,...n} U_i$ and $\varphi(Z)$ is the p.d.f of the standard normal distribution. Also, $P(E_0)$ can be written in a similar way,

$$\begin{aligned} P(E_0) &= P\left\{\eta : |Z_i - Z_c - \overline{Z} + \overline{Z}| \leq c^* \ , i = 1, ..., n\right\} \\ &= P\left\{\eta : |U_i + Z_1| \leq c^* \ , i = 1, ..., n\right\} \\ &= \int_{(U_1,...,U_n)} \int_{(\sqrt{n/n+1}(-c^* - u_{[n]})}^{(\sqrt{n/n+1}(c^* - u_{[1]})} f_U(U)\varphi_Z(Z)dZdU \end{aligned}$$

Observe that $U_{[1]} \geq R^+U_{[1]}$ and $U_{[n]} \leq R^+U_{[n]}$, and hence,

$$\int\int_{(\sqrt{n/n+1}(-c^* - R^+U_{[n]})}^{(\sqrt{n/n+1}(c^* - R^+U_{[1]})} f_U(U)\varphi_Z(Z)dZdU \geq \int\int_{(\sqrt{n/n+1}(-c^* - U_{[n]})}^{(\sqrt{n/n+1}(c^* - U_{[1]})} f_U(U)\varphi_Z(Z)dZdU$$

which proves that $P(E_1) \geq P(E_0)$.

Theorem 4.1.1. *The lower bound $P(E_1)$ is*
$P(E_{L1}) = P(-c^* + \sqrt{a} \leq Z_i - Z_c \leq c^* - \sqrt{a}, a \leq SS)$.

Proof:

Let $X_i = Z_i + \theta_i$, $\overline{Z} = \overline{X} - \overline{\theta}$, $V_i = X_i - \overline{X}$, and $V_i = U_i + \eta_i$. Then for $a \leq SS$, $P(E_1)$ becomes,

$$\begin{aligned} P(E_1) &= P\left\{|Z_i - Z_1 - \eta_i| \leq c^*, a \leq SS\right\} + P\left\{|R^+U_i + Z_1 - (1 - R^+)\eta_i| \leq c^*, a \geq SS\right\} \\ &\geq P\left\{|R^+U_i + Z_1 - (1 - R^+)\eta_i| \leq c^*, a \leq SS\right\} \\ &\geq P\left\{|U_i - \frac{a}{SS}U_i + Z_1 - \frac{a}{SS}\eta_i| \leq c^*, a \leq SS\right\} \\ &\geq P\left\{|Z_i - Z_c - \frac{a}{SS}(U_i + \eta_i)| \leq c^*, a \leq SS\right\} \\ &\geq P\left\{|Z_i - Z_c - \frac{a}{SS}V_i| \leq c^*, a \leq SS\right\} \\ &\geq P\left\{-c^* + \frac{a}{SS}V_i \leq Z_i - Z_c \leq c^* + \frac{a}{SS}V_i, a \leq SS\right\}. \end{aligned}$$

Then from the lemma 3.1.6,

$$\begin{aligned} P(E_1) &\geq P\left\{-c^* + \frac{a}{SS}V_i \leq Z_i - Z_1 \leq c^* + \frac{a}{SS}V_i, a \leq SS\right\} \\ &\geq P\left\{-c^* + \sqrt{a} \leq Z_i - Z_1 \leq c^* - \sqrt{a}, a \leq SS\right\} = P(E_{L1}) \end{aligned}$$

Keep in mind that $P(E_{L1}) \leq P\left\{-c + \sqrt{\frac{a\times n-1}{n}} \leq Z_i - Z_c \leq c - \sqrt{\frac{a\times n-1}{n}}, a \leq SS\right\} \leq P(E_1)$.$\blacksquare$
From Lemma 3.1.1 If $n \to \infty$, $P(a < SS) \to 0$, then
$P(E_{L1}) = P\{-c^* + \sqrt{a} \leq Z_i - Z_1 \leq c^* - \sqrt{a}, a \leq SS\} = P\{-c^* + \sqrt{a} \leq Z_i - Z_1 \leq c^* - \sqrt{a}\}$.$\blacksquare$

Theorem 4.1.2. *$P(E_1)$ dominate $P(E_0)$ up to an arbitrarily small constant for a sufficiently large number of means, n.*

Proof:

$$\begin{aligned} P(E_1) &= P\left\{\eta : |R^+U_i + Z_1 - (1-R^+)\eta_i| \leq c^*\right\} \\ &= P\{\eta : |Z_1 - \eta_i| \leq c^*, SS \leq a\} + P\left\{\eta : |R^+U_i + Z_1 - (1-R^+)\eta_i| \leq c^*, SS \geq a\right\} \end{aligned}$$

Observe that for a large enough n,$P(SS \leq a) \to 0$ since $a < n$.Then,

$$P(E_1) = P\left\{\eta : |R^+U_i + Z_1 - (1-R^+)\eta_i| \leq c^*\right\}.$$

From Theorem 4.1.1,
$P(E_1) \geq P(-c^* + \sqrt{a} \leq Z_i - Z_c \leq c^* - \sqrt{a}) = E_{Z_c}\left[\Phi(c^* - \sqrt{a} + Z_c) - \Phi(-c^* + \sqrt{a} + Z_c)\right]^n$,
then let $a^* = \sqrt{a}$ and
$\Delta = [\Phi(c^* + Z_c) - \Phi(-c^* + Z_c)]^n - [\Phi(c^* - \sqrt{a} + Z_c) - \Phi(-c^* + \sqrt{a} + Z_c)]^n$.

$$\begin{aligned} \lim_{n\to\infty} P(E_0) - P(E_1) &= \lim_{n\to\infty} E_{Z_c}\Delta \\ &= E_{Z_c} \lim_{n\to\infty} \Delta. \end{aligned}$$

We need to evaluate $\lim_{n\to\infty}\Delta$.

$$\Delta = [\Phi(c^* + Z_c) - \Phi(-c^* + Z_c)]^n \left[1 - \frac{[\Phi(c^* - \sqrt{a} + Z_c) - \Phi(-c^* + \sqrt{a} + Z_c)]^n}{[\Phi(c^* + Z_c) - \Phi(-c^* + Z_c)]^n}\right].$$

Let $A = \frac{\left[\Phi(c^*-\sqrt{a}+Z_c)-\Phi(-c^*+\sqrt{a}+Z_c)\right]^n}{[\Phi(c^*+Z_c)-\Phi(-c^*+Z_c)]^n}$, then let $B = \ln A$,
where $B = n \times \ln\left[\frac{\Phi(c^*-a^*+Z_c)-\Phi(-c^*+a^*+Z_c)}{\Phi(c^*+Z_c)-\Phi(-c^*+Z_c)}\right]$ and $a^* = \sqrt{a}$, our aim is to show that $\lim_{n\to\infty} B = 0$

$$\lim_{n\to\infty} B \quad = \lim_{n\to\infty} \quad n \times \ln\left[\frac{\Phi(c^* - a^* + Z_c) - \Phi(-c^* + a^* + Z_c)}{\Phi(c^* + Z_c) - \Phi(-c^* + Z_c)}\right]$$

and from Lemma 3.18, $n = \frac{\ln\beta}{\ln(2\Phi(c)-1))}$.

$$\begin{aligned}
\lim_{n\to\infty} B &= \lim_{n\to\infty} \frac{\ln\beta}{\ln(2\Phi(c)-1))} \times \ln\left[\frac{\Phi(c^* - a^* + Z_c) - \Phi(-c^* + a^* + Z_c)}{\Phi(c^* + Z_c) - \Phi(-c^* + Z_c)}\right] \\
&= \lim_{n\to\infty} \ln\beta \left[\frac{\ln(\Phi(c^* - a^* + Z_c) - \Phi(-c^* + a^* + Z_c))}{\ln(2\Phi(c)-1))} - \frac{\ln(\Phi(c^* + Z_c) - \Phi(-c^* + Z_c))}{\ln(2\Phi(c)-1))}\right] \\
&\to \ln\beta\left[\frac{0}{0} - \frac{0}{0}\right]
\end{aligned}$$

Applying L'Hospital rule;

$$\lim_{n\to\infty} B \quad = \quad \lim_{n\to\infty} \ln\beta \left[\frac{\frac{\phi(c^*-a^*+Z_c)}{\Phi(c^*-a^*+Z_c)}}{\frac{2\phi(c)}{2\Phi(c)-1}} + \frac{\frac{\phi(-c^*+a^*+Z_c)}{\Phi(-c^*+a^*+Z_c)}}{\frac{2\phi(c)}{2\Phi(c)-1}} - \frac{\frac{\phi(c^*+Z_c)}{\Phi(c^*+Z_c)}}{\frac{2\phi(c)}{2\Phi(c)-1}} + \frac{\frac{\phi(-c^*+Z_c)}{\Phi(-c^*+Z_c)}}{\frac{2\phi(c)}{2\Phi(c)-1}}\right]$$

As $n \to \infty$, $\frac{\phi(c^*)}{\phi(c)} \to 0$ since $c^* \geq c$.Therefore, $B \to 0$. Since $A = exp[B]$, as $n \to \infty$, $A \to 1$. This implies that $\Delta \to 0$. This completes the proof.

We showed that the simultaneous confidence intervals for the differences between treatment means and the mean of a control in one-way ANOVA model centered around a shrinkage estimator, have coverage probability uniformly exceeding that of the usual rectangular confidence regions up to an arbitrarily small epsilon when the number of means is sufficiently large. We also showed that these intervals have strictly greater coverage probability when all the main effects are zero, and that the coverage probability of the two procedures converge to one another when at least one of the main effects becomes arbitrarily large. We also developed a lower bound for the coverage probability of the rectangular confidence region which is a decreasing function of the shrinkage constant.

For $P(E_1)$ and $P(C_1)$, we have the same analytical results. Moreover, $P(E_1)$ and $P(C_1)$ are in the same form (see 4.2 and 3.10). These reasons lead us to conclude that $P(E_1)$ behave like $P(C_1)$. In other words, when we graph $P(E_1)$ against either the length of the vector of the main effects($\|\eta\|$) or the absolute maximum of the main effects($\max_i |\eta_i|$); we expect that the coverage probabilities will be a decreasing or an increasing function of neither $\|\eta\|$ nor $\max_i |\eta_i|$. Similar to the relationship between $P(C_1)$ and $P(C_0)$, $P(E_1)$ will have higher coverage probabilities than $P(E_0)$ for $n \geq 4$.

5.0 NUMERICAL INTEGRATION

Since we used extensive simulation to prove the $P(C_1)$ is uniformly higher than $P(C_0)$, for n = 2,...10, the next step is to make sure our simulation results are reliable. For n=2,3,4, we use the numerical integration method to calculate the coverage probabilities.

The most widely investigated method for approximating a definite integral is

$$\int_a^b w(x)f(x)d(x) \approx \sum_{i=1}^{n} A_i f(x_i). \tag{5.1}$$

here $w(x)$ is a function. The x_i are called the points (or nodes) of the formula and the A_i are called coefficients (or weights). If $w(x)$ is nonnegative in $[a, b]$, then n points and coefficients can be found to make (5.1) exact for all polynomials of degree $\leq 2n - 1$; this is the highest degree of precision which can be obtained using n points. Such formulas are usually called Gaussian quadrature formulas because they were first studied by Gauss. We first need to write our coverage probability to see the form of the integration to identify the weight function, $w(x)$, and function, $f(x)$.

The coverage probability for our procedure can be written in the following way, let n=k, and k=2,3,4.

There are k-1 sets, since $\sum_{i=1}^{n}(X_i - \overline{X}) = 0$ or $\sum_{i=1}^{n} U_i = 0$, those sets are

$$\begin{aligned}
-c - R^{+}U_1 + (1-R^{+})\eta_1 &\leq \overline{Z} \leq c - R^{+}U_1 + (1-R^{+})\eta_1 \\
-c - R^{+}U_2 + (1-R^{+})\eta_1 &\leq \overline{Z} \leq c - R^{+}U_2 + (1-R^{+})\eta_2 \\
&\vdots \\
-c - R^{+}U_{k-1} + (1-R^{+})\eta_{k-1} &\leq \overline{Z} \leq c - R^{+}U_{k-1} + (1-R^{+})\eta_{k-1} \\
-c + R^{+}\sum_{i=1}^{k-1} U_i - (1-R^{+})\sum_{i=1}^{k-1}\eta_i &\leq \overline{Z} \leq c + R^{+}\sum_{i=1}^{k-1} U_i - (1-R^{+})\sum_{i=1}^{k-1}\eta_i
\end{aligned}$$

Let UP be the minimum of the all the upper bounds and LP be the maximum of the all the lower bounds, then

$P(C_1) = P\left\{UP \leq \overline{Z} \leq LP\right\}$ and $U = \begin{pmatrix} U_1 \\ \vdots \\ U_{k-1} \end{pmatrix}$ has normal distribution $N(0, \Sigma)$, where Σ is nonsingular matrix. Then,

$$\begin{aligned}
P(C_1) &= \int \left[\Phi(\sqrt{k}UP) - \Phi(\sqrt{k}LP)\right] f_U(U) 1_{\{LP \leq UP\}} dU \\
&= \int \left[\Phi(\sqrt{k}UP) - \Phi(\sqrt{k}LP)\right] 1_{\{LP \leq UP\}} \frac{1}{\sqrt{2\pi}^{k-1}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}U\Sigma^{-1}U'\right) dU
\end{aligned}$$

Let $W(U) = \left[\Phi(\sqrt{k}UP) - \Phi(\sqrt{k}LP)\right] 1_{\{LP \leq UP\}}$, then

$$P(C_1) = \int W(U) \frac{1}{\sqrt{2\pi}^{k-1}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}U\Sigma^{-1}U'\right) dU.$$

Let $U = \sqrt{2}\Sigma^{1/2}V$ and this implies $dU = \sqrt{2}^{n}|\Sigma^{1/2}|dV = \sqrt{2}^{k-1}|\Sigma|^{1/2}dV$. Then,

$$\begin{aligned}
P(C_1) &= \int W(\sqrt{2}\Sigma^{1/2}V) \frac{1}{\sqrt{2\pi}^{k-1}|\Sigma|^{1/2}} \exp\left(-V'V\right) \sqrt{2}^{k-1}|\Sigma|^{1/2} dV \\
P(C_1) &= \int \frac{W(\sqrt{2}\Sigma^{1/2}V)}{\sqrt{\pi}^{k-1}} \exp\left(-V'V\right) dV.
\end{aligned}$$

Let $W_1(V) = \frac{W(\sqrt{2}\Sigma^{1/2}V)}{\sqrt{\pi}^{k-1}}$, then

$$P(C_1) \;=\; \int W_1(V)\exp\left(-V^{'}V\right)dV \approx \sum_{i_{k-1}}^{M_{k-1}}\cdots\sum_{i_1}^{M_1} A_1\cdots A_{k-1}W_1(a_1,\cdots,a_{k-1}),$$

where $a_i's$ are called nodes, A_i are the Hermite polynomials weights and, $M_i's$ are the number of nodes. We use Hermite polynomials because the function in this form e^{-x^2}. It is stated in most text books that numerical integration method for higher dimensions are not reliable. Therefore, we decided to use numerical integration to calculate $P(C_1)$ for n=2,3,4 (See R codes for numerical integration).

For n=2, $P(C_1)$ becomes

$$P(C_1) \;=\; \int W_1(X)\exp(-X^2)dV \approx \sum_{i_1}^{M_1} A_1W_1(a_i),$$

where $W_1(X) = \frac{1}{\sqrt{2}}[2\Phi(\sqrt{2}(c - |R^+X - (1-R^+)\eta|)) - 1]1_{c\geq|R^+X-(1-R^+)\eta|}$. We use 1500 nodes and weights to approximate $P(C_1)$. We calculate $P(C_1)$ for different choice of the shrinkage constant. The difference between the results from the numerical integration and the simulations is negligible. Our simulation results and the result of the numerical integration method agreed. (See Figure8.1 - 8.8)

For n=3, $P(C_1)$ becomes

$$P(C_1) \;=\; \int W_1(X)\exp(X_1^2 + X_2^2)dX \approx \sum_{i_2}^{M_2}\sum_{i_1}^{M_1} A_1A_2W_1(a_{i_1}, a_{i_2}),$$

where $W_1(X) = \frac{1}{\pi}[\Phi(\sqrt{3}UB) - \Phi(\sqrt{3}LB)]1_{UB\geq LB}$ where UB is the minimum of the all upper bounds and LB is the maximum of all the lower bounds. We use 800 nodes and weights to approximate $P(C_1)$. We calculate $P(C_1)$ for different choices of the shrinkage constant. The difference between the results from the numerical integration and the simulations is negligible. However, the numerical integration results tend to be slightly higher than the simulations results. Also the contour plot shows a clear picture of the domination; the coverage probability achieves its maximum when all the population cell means are zero and the coverage probability achieves its minimum when all the cell means are big that is slight bigger than 0.95. (See Figure 8.9 - 8.16). Since the numerical integration methods are not highly recommended for the higher dimension, we use the numerical integration method for at most n=4 (dimension of integration is 3).

For n=4, $P(C_1)$ becomes

$$P(C_1) \quad = \quad \int W_1(X) \exp(X_1^2 + X_2^2 + X_3^2) dX \approx \sum_{i_3}^{M_2} \sum_{i_3}^{M_2} \sum_{i_1}^{M_1} A_1 A_2 A_3 W_1(a_{i_1}, a_{i_2}, a_{i_3}),$$

where $W_1(X) = \frac{1}{\pi^{3/2}}[\Phi(2UB) - \Phi(2LB)]1_{UB \geq LB}$ where UB is the minimum of the all upper bounds and LB is the maximum of all the lower bounds. We use 40 nodes and weights to approximate $P(C_1)$. We calculate $P(C_1)$ for different choices of the shrinkage constant. The difference between the results from the numerical integration and the simulations is negligible. However, the numerical integration results tend to be slightly higher than the simulations results. (See Figure 8.17-8.19) Since there is no significant difference between our simulation result and the numerical integration result, we are confident saying that our simulation results are reliable. Moreover, since we use 10,000 replication in our simulation, this makes our simulation result to be accurate until the third decimal. Therefore our simulation results are reliable.

6.0 TWO WAY ANALYSIS OF VARIANCE MODEL

Since the two-way ANOVA model a special case of the one-way ANOVA model, we hope to prove the same results for the two- way ANOVA model that we earlier have shown for the one-way ANOVA model. Fabian (1990) gave a simultaneous confidence interval for the cell means in a two-way ANOVA model in which additivity is conjectured but the presence of interaction cannot be ruled out. He suggested the following recommendation: ignore interactions and do analysis but estimate the error involved in neglecting the interactions from the power of the test. He also stated that, one-way method or usual recommendation for two-way method is substantially better than his method. Gleser (1992) pointed out the flaw in Fabaian's recommendation and suggested the confidence rectangle centered at the related shrinkage estimator. Gleser suggested the following point estimator

$$X_{0ij} + R^{+}W_{ij}$$

where $R^{+} = (1 - \frac{a}{\sum i,j W_{ij}^2})^{+}$. Therefore we use Gleser's point estimator to prove that the procedure suggested by Gleser a has uniformly higher coverage probability than the usual procedure.

We consider the balanced two-way ANOVA model but we give modification of our procedure for the unbalanced two-way ANOVA model in an *Appendix*. After briefly restating the model and assumptions, we start with the known σ case and continue with the unknown σ case. We finish with our simulations results for two-way ANOVA model.

6.1 BALANCED TWO WAY ANOVA MODEL

We introduced the model for the balanced two-way ANOVA model in the second chapter. Y_{ijk} are obtained from a balanced two-way classification design. In such a two factor experiment, there are I levels of factor A and J levels of factor B and K replications for each treatment combination of ith level of factor A and jth level of factor B. We defined the cell mean model for two-way ANOVA model in (2.6). The model is

$$X_{ij} = X_{0ij} + W_{ij} \quad \text{i} = 1, ...I \ \text{j} = 1, ..., J$$

In (2.6),X_{ij} is the sample cell mean for ith level of the first factor, A, and the jth level of the second factor, B, X_{0ij} is the main effect for ith level of A and jth level of B and W_{ij} is the interaction effect for the ith level of A with jth level of B. X_{0ij}'s are distributed normally $N(0, \sigma_e \frac{I+J-1}{IJ})$ and W_{ij}'s are normally distributed

$$N(\theta_{ij}, \sigma_e^2 \frac{(I-1)(J-1)}{IJ})$$

where $\sigma_e = \frac{\sigma}{\sqrt{K}}$ and $\sum_{j=1}^{J} W_{ij} = 0$, i=1,...,I; $\sum_{i=1}^{I} W_{ij} = 0$, j=1,...,J.

The usual confidence rectangle for the population means, θ_{ij}, under the (i, j) treatment combination is

$$D_1 = \{\Theta = (\theta_{11}, ..., \theta_{IJ}) : (|X_{ij} - \theta_{ij}| \leq s, i = 1..., I, j = 1, ..., J\} \tag{6.1}$$

where $s = \frac{\sigma_e}{\sqrt{K}} \Phi^{-1}(\frac{\beta^{\frac{1}{N}}+1}{2})$, where N=IJK and Φ is the standard normal distribution function.

The confidence interval we are proposing is

$$D_2 = \left\{\Theta = (\theta_{11}, ..., \theta_{IJ}) : (|X_{oij} + R^{+}W_{ij} - \theta_{ij}| \leq s, i = 1..., I, j = 1, ..., J\right\} \tag{6.2}$$

where

$$R^{+} = (1 - \frac{a\sigma_e}{\sum_{i=1}^{I}\sum_{j=1}^{J} W_{ij}^2})^{+}. \tag{6.3}$$

That is,

$$R^{+} = \begin{cases} 1 - \frac{a\sigma_e}{\sum_{i=1}^{I}\sum_{j=1}^{J}(W_{ij})^2} & : \quad \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} W_{ij}^2}{\sigma_e} \geq a \\ 0 & : \quad \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} W_{ij}^2}{\sigma_e} \leq a. \end{cases}$$

6.2 THE KNOWN σ_E CASE

If the σ_e is known, without loss of generality we assume that σ_e is 1. Then the usual simultaneous confidence interval for the cell mean becomes

$$D_1 = \{\Theta = (\theta_{11}, ..., \theta_{IJ}) : (|X_{ij} - \theta_{ij}| \leq s, i = 1..., I, j = 1, ..., J\} \tag{6.4}$$

where $s = \Phi^{-1}(\frac{\beta^{\frac{1}{N}}+1}{2})$, where N is IJ and Φ is the distribution function of standard normal distribution.

The confidence interval, we are proposing is

$$D_2 = \left\{\Theta = (\theta_{11}, ..., \theta_{IJ}) : (|X_{oij} + R^{+}W_{ij} - \theta_{ij}| \leq s, i = 1..., I, j = 1, ..., J\right\} \tag{6.5}$$

$$R^{+} = \begin{cases} 1 - \frac{a}{\sum_{i=1}^{I}\sum_{j=1}^{J}(W_{ij})^2} & : \quad \sum_{i=1}^{I}\sum_{j=1}^{J} W_{ij}^2 \geq a \\ 0 & : \quad \sum_{i=1}^{I}\sum_{j=1}^{J} W_{ij}^2 \leq a. \end{cases}$$

6.2.1 Analytical Results

Lemma 6.2.1. *if $\theta'\theta \to \infty$, $P(D_2) \to P(D_1)$.*

Proof:

Let $W_{ij} = V_{ij} + \theta_{ij}$, $\gamma = \sqrt{\theta'\theta}$ and $\epsilon_{ij} = \frac{\theta_{ij}}{\gamma}$. Then $\sum_{i=1}^{I}\sum_{j=1}^{J}(V_{ij} + \theta_{ij})^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}(V_{ij} + \epsilon_{ij}\gamma)^2 \to \infty$. This implies that the shrinkage factor, R^{+}, will be positive since $a \leq \sum_{i=1}^{I}\sum_{j=1}^{J}(V_{ij} + \epsilon_{ij}\gamma)^2$. Then,

$$\begin{aligned} P(D_2) &= P(|V_{ij} - \frac{aV_{ij}}{\sum_{i=1}^{I}\sum_{j=1}^{J}(V_{ij} + \epsilon_{ij}\gamma)^2} + X_{0ij} - \frac{a}{\sum_{i=1}^{I}\sum_{j=1}^{J}(V_{ij} + \epsilon_{ij}\gamma)^2}))\epsilon_{ij}\gamma| \leq s) \\ &= P(|V_{ij} - a\frac{V_{ij} + \epsilon_{ij}\gamma}{\sum_{i=1}^{I}\sum_{j=1}^{J}(V_{ij} + \epsilon_{ij}\gamma)^2} + X_{0ij}| \leq s) \end{aligned}$$

Let $X_0' = (X_{011}, ...X_{0IJ})$, $V' = (V_{11}, ..., V_{IJ}$ and $\theta' = (\theta_{11}, ..., \theta_{IJ})$. Then

$$P(D_2) = \int \int_{-s-(1-\frac{a}{\sum_{i=1}^{I}\sum_{j=1}^{J}(V_{ij}+\epsilon_{ij}\gamma)^2})V + \frac{a}{\sum_{i=1}^{I}\sum_{j=1}^{J}(V_{ij}+\epsilon_{ij}\gamma)^2}\theta}^{s-(1-\frac{a}{\sum_{i=1}^{I}\sum_{j=1}^{J}(V_{ij}+\epsilon_{ij}\gamma)^2})V + \frac{a}{\sum_{i=1}^{I}\sum_{j=1}^{J}(V_{ij}+\epsilon_{ij}\gamma)^2}\theta} f_{X_0}(X_0) f_V(V) dX_0 dV$$

. Let $g(V,\theta) = -(1 - \frac{a}{\sum_{i=1}^{I}\sum_{j=1}^{J}(V_{ij}+\epsilon_{ij}\gamma)^2})V + \frac{a}{\sum_{i=1}^{I}\sum_{j=1}^{J}(V_{ij}+\epsilon_{ij}\gamma)^2}\theta$. Observe that

$$\lim_{\gamma\to\infty} g(V,\theta) \to 0.$$

Since all the conditions are met for the dominated convergence theorem, we can take the limit inside of integration.

$$\begin{aligned}
\lim_{\gamma\to\infty} P(D_2) &= \int\int_{-s-g\lim_{\gamma\to\infty}(V,\theta)}^{s-\lim_{\gamma\to\infty} g(V,\theta)} f_{X_0}(X_0) f_V(V) dX_0 dV \\
&= \int\int_{-s}^{s} f_{X_0}(X_0) f_V(V) dX_0 dV \\
&= P(D_1)
\end{aligned}$$

Lemma 6.2.2. *If $a \leq WW$, then $|(1-R^+)W_{ij}| = |\frac{aW_{ij}}{WW}| \leq \sqrt{\frac{a\times N-1}{N}} \leq \sqrt{a}$, where $WW = \sum_{i=1}^{I}\sum_{j=1}^{J} W_{ij}^2$.*

Proof:

We use the fact that $\sum_{i=1}^{I}\sum_{j=1}^{J} W_{ij} = 0$.

$$\begin{aligned}
\sum_{i=1}^{I}\sum_{j=1}^{J} W_{ij} &= \sum_{i=1}^{I}\sum_{j=1}^{J}{}_{i=I,j\neq J} W_{ij} + W_{IJ} = 0 \\
-\sum_{i=1}^{I}\sum_{j=1}^{J}{}_{i=I,j\neq J} W_{ij} &= W_{IJ} \\
\sum_{i=1}^{I}\sum_{j=1}^{J}{}_{i=I,j\neq J} W_{ij}^2 \geq \frac{(\sum_{i=1}^{I}\sum_{j=1} W_{ij})^2}{N-1} &= \frac{W_{IJ}^2}{N-1}
\end{aligned}$$

Moreover,

$$\begin{aligned}
WW &= \sum_{i=1}^{I}\sum_{j=1}^{J}{}_{i=I,j\neq J} W_{ij}^2 + W_{IJ}^2 \\
&\geq \frac{W_{IJ}^2}{N-1} + W_{IJ}^2 \\
&\geq W_{IJ}^2 \frac{N}{N-1} \\
|W_{IJ}| &\leq \sqrt{\frac{WW \times N-1}{N}}.
\end{aligned}$$

Then for $a \leq WW$,

$$\begin{aligned}
(1-R^{+})(W_{ij}) &= \frac{aW_{ij}}{WW} \\
\frac{aW_{ij}}{WW} &\leq \frac{a\sqrt{\frac{WW\times N-1}{N}}}{WW} \\
&\leq \sqrt{\frac{a}{WW}} \times \sqrt{\frac{a\times N-1}{N}} \\
&\leq \sqrt{\frac{a\times N-1}{N}} \leq \sqrt{a}
\end{aligned}$$

and the same way

$$\begin{aligned}
(1-R^{+})(W_{ij}) &= \frac{aW_{ij}}{WW} \\
\frac{aW_{ij}}{WW} &\geq -\frac{a\sqrt{\frac{WW\times N-1}{N}}}{WW} \\
&\geq -\sqrt{\frac{a}{WW}} \times \sqrt{\frac{a\times N-1}{N}} \\
&\geq -\sqrt{\frac{a\times N-1}{N}} \geq -\sqrt{a}.
\end{aligned}$$

This completes the proof.

The following theorem states the our lower bound for $P(D_2)$.

Theorem 6.2.1. *The lower bound $P(D_2)$ is*

$$P(D_{L2}) = P(-s+\sqrt{a} \leq Z_{ij} \leq s-\sqrt{a}, a \leq WW)$$

, where $X_{ij} = Z_{ij} + \theta_{ij}$.

Proof:

Let $W_{ij} = V_{ij} + \theta_{ij}$ and $WW = \sum_{i=1}^{I}\sum_{j=1}^{J} W_{ij}^2$.

$$\begin{aligned}
P(D_2) &= P\{|X_{0ij}-\theta_{ij}| \leq s, a \leq WW\} + P\left\{|R^{+}W_{ij} + X_{0ij} - \theta_{ij}| \leq s, a \geq WW\right\} \\
P(D_2) &\geq P\left\{|R^{+}W_{ij} + X_{0ij} - \theta_{ij}| \leq s, a \leq WW\right\} \\
&\geq P\left\{|W_{ij} - \frac{a}{WW}W_i + X_{0ij} - \theta_{ij}| \leq s, a \leq WW\right\} \\
&\geq P\left\{|V_{ij} + X_{0ij} - \frac{a}{WW}W_i| \leq s, a \leq WW\right\}
\end{aligned}$$

Then from the previous lemma,

$$\begin{aligned} P(D_2) &\geq P\left\{-s + \frac{a}{WW}W_{ij} \leq Z_{ij} \leq s + \frac{a}{WW}W_{ij}, a \leq WW\right\} \\ &\geq P\left\{-s + \sqrt{a} \leq Z_{ij} \leq s - \sqrt{a}, a \leq WW\right\} = P(D_{L2}) \end{aligned}$$

Keep in mind that $P(D_{L2}) \leq P\left\{-s + \sqrt{\frac{a \times N-1}{N}} \leq Z_{ij} \leq s - \sqrt{\frac{a \times N-1}{N}}, a \leq WW\right\} \leq P(D_2)$.

Lemma 6.2.3. *For a large enough N,* $P(D_{L2}) = (2\Phi(s - \sqrt{a}) - 1)^N$.

Proof:

If N is large enough, $\limsup \frac{a}{N} = 0$. Then from Theorem 6.2.1

$$P(D_2) \geq P\left\{-s + \sqrt{a} \leq Z_{ij} \leq s - \sqrt{a}, a \leq WW\right\}$$

Since $\limsup \frac{a}{N} = 0$, then $P(WW \geq a) = 1$, because WW has a non central chi squared distribution which is a schur concave in η.

Then let $E : \{-s + \sqrt{a} \leq Z_{ij} \leq s - \sqrt{a}, i = 1, ..., I\ j = 1, ..., J\}$ and $F = \{a \leq WW\}$. Then

$$\begin{aligned} P(C_1) &\geq P\left\{-s + \sqrt{a} \leq Z_{ij} \leq s - \sqrt{a}, a \leq WW\right\} \\ &\geq P(E) - P(F^c) \geq P(E) - P(F^c) \\ &\geq P(-s + \sqrt{a} \leq Z_{ij} \leq s - \sqrt{a}) + P(a \geq WW) \\ &\geq P(-s + \sqrt{a} \leq Z_{ij} \leq s - \sqrt{a}) = (2\Phi(s - \sqrt{a}) - 1)^N \end{aligned}$$

Theorem 6.2.2. *$P(D_2)$ dominate $P(D_1)$ up to an arbitrarily small constant for a sufficiently large number of means, n.*

Proof:

$$P(D_2) = P\left\{\theta : |R^+W_{ij} + X_{0ij} - \theta_{ij}| \leq s\right\}$$
$$P(D_2) = P\left\{\theta : |X_{0ij} - \theta_{ij}| \leq s, WW \leq a\right\} + P\left\{\theta : |R^+W_{ij} + X_{0ij} - \theta_i| \leq s, WW \geq a\right\}$$

Observe that for a large enough N,$P(WW \leq a) \rightarrow 0$ since a is bounded above by a constant and $a < N$.Then,

$$P(D_2) = P\left\{\theta : |R^+W_i + X_{0ij} - \theta_{ij}| \leq s, WW \geq a\right\}.$$

From Theorem 6.2.1,
$P(D_2) \geq (2\Phi(s - \sqrt{a}) - 1)^N$, the rest of the proof is similar the proof of theorem 3.1.3

6.3 THE UNKNOWN σ_E CASE

The usual confidence rectangle for the vector θ of cell mean is

$$D_1 = \{\theta = (\theta_{11}, ..., \theta_{IJ}) : |X_{ij} - \theta_{ij}| \leq s \text{ , } i = 1, ..., I, j = 1, .., J\}. \tag{6.6}$$

, where $s = |m|_{\alpha,n,v}\frac{\hat{\sigma}}{\sqrt{IJ}}$, and $|m|_{\alpha,n,v}$ is the $1 - \alpha$ quantile of the Student maximum modulus statistics, $\hat{\sigma}^2$ is the unbiased estimator of σ_e^2. Keep in mind that when IJ is large, the quantile of Student maximum modulus statistics can be replaced by the quantile of student t distribution.

We replace σ with an unbiased estimator $\hat{\sigma}^2 = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk} - X_{ij})^2}{IJ(K-1)}$. Our confidence interval is the same as 6.2, the only difference is the shrinkage estimator.

$$D_2(X) = \left\{|(X_{0ij} + R^+W_{ij} - \theta_{ij}| \leq s)\right\} \tag{6.7}$$

where R^+ is

$$R^+ = (1 - \frac{a\hat{\sigma}^2}{\sum_{i=1}^{I}\sum_{j=1}^{J} W_{ij}^2})^+. \tag{6.8}$$

That is,

$$R^+ = \begin{cases} 1 - \frac{a\hat{\sigma}^2}{\sum_{i=1}^{I}\sum_{j=1}^{J} W_{ij}^2} & : \quad \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} W_{ij}^2}{\hat{\sigma}^2} \geq a \\ 0 & : \quad \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} W_{ij}^2}{\hat{\sigma}^2} \leq a. \end{cases}$$

By using same notations in known σ case, 6.7 becomes

$$D_1(X) = \left\{ |X_{oij} + R^+ V_{ij} - (1 - R^+)\theta_{ij}| \leq s) \right\} \tag{6.9}$$

where

$$R^+ = (1 - \frac{a\hat{\sigma}^2}{\sum_{i=1}^{I} \sum_{j=1}^{J} (V_{ij} + \theta_{ij})^2})^+ \tag{6.10}$$

and $W_{ij} = V_{ij} + \theta_{ij}$. Now, we try to derive same analytical results for the unknown σ_e case.

6.4 ANALYTICAL RESULTS

Lemma 6.4.1. *If the shrinkage constant, a, equals to zero, $P(D_2) \to P(D_1)$.*

Proof:

If $a = 0$, then $R^+ = 1$. This implies that $P(D_2) \to P(D_1)$.

This lemma states that if we pick a small enough, $P(D_2)$ will be very close to $P(D_1)$.

Lemma 6.4.2. *if $\theta^{'}\theta \to \infty$, $P(D_2) \to P(D_1)$.*

Proof:

$$\lim_{\theta^{'}\theta \to \infty} P(D_2) = \lim_{\theta^{'}\theta \to \infty} E_{\hat{\sigma}^2} P(D_2/\hat{\sigma}^2)$$

$\hat{\sigma}^2$ has a chi-square distribution and $|P(D_2)| \leq 1$ is bounded. Therefore, all the conditions for dominated convergence theorem are met, we can take the limit inside of expectation. The main and interaction effects have independent multivariate t distribution so we can take the limit inside of the probability too. The rest of the proof of this lemma is very similar for the lemma 6.2.1.

Lemma 6.4.3. *If $a \leq \frac{WW}{\hat{\sigma}^2}$, then $|(1 - R^+)W_{ij|} = |\frac{a\hat{\sigma}^2 W_{ij}}{WW}| \leq \sqrt{\frac{a\hat{\sigma}^2 \times N-1}{N}} \leq \sqrt{a}\hat{\sigma}$, where $WW = \sum_{i=1}^{I} \sum_{j=1}^{J} W_{ij}^2$.*

Proof:

We use the fact that $\sum_{i=1}^{I}\sum_{j=1}^{J} W_{ij} = 0$.

$$\begin{aligned}
\sum_{i=1}^{I}\sum_{j=1}^{J} W_{ij} &= \sum_{i=1}^{I}\sum_{j=1}^{J}{}_{i=I,j\neq J} W_{ij} + W_{IJ} = 0 \\
-\sum_{i=1}^{I}\sum_{j=1}^{J}{}_{i=I,j\neq J} W_{ij} &= W_{IJ} \\
\sum_{i=1}^{I}\sum_{j=1}^{J}{}_{i=I,j\neq J} W_{ij}^2 \geq \frac{(\sum_{i=1}^{I}\sum_{j=1} W_{ij})^2}{N-1} &= \frac{W_{IJ}^2}{N-1}
\end{aligned}$$

Moreover,

$$\begin{aligned}
WW &= \sum_{i=1}^{I}\sum_{j=1}^{J}{}_{i=I,j\neq J} W_{ij}^2 + W_{IJ}^2 \\
&\geq \frac{W_{IJ}^2}{N-1} + W_{IJ}^2 \\
&\geq W_{IJ}^2 \frac{N}{N-1} \\
|W_{IJ}| &\leq \sqrt{\frac{WW \times N-1}{N}}.
\end{aligned}$$

Then for $a\hat{\sigma}^2 \leq WW$,

$$\begin{aligned}
(1-R^+)(W_{ij}) &= \frac{a\hat{\sigma}^2 W_{ij}}{WW} \\
\frac{a\hat{\sigma}^2 W_{ij}}{WW} &\leq \frac{a\hat{\sigma}^2\sqrt{\frac{WW\times N-1}{N}}}{WW} \\
&\leq \sqrt{\frac{a\hat{\sigma}^2}{WW}} \times \sqrt{\frac{a\hat{\sigma}^2 \times N-1}{N}} \\
&\leq \sqrt{\frac{a\hat{\sigma}^2 \times N-1}{N}} \leq \sqrt{a}\hat{\sigma}
\end{aligned}$$

and the same way

$$\begin{aligned}
(1-R^+)(W_{ij}) &= \frac{a\hat{\sigma}^2 W_{ij}}{WW} \\
\frac{a\hat{\sigma}^2 W_{ij}}{WW} &\geq -\frac{a\hat{\sigma}^2\sqrt{\frac{WW\times N-1}{N}}}{WW} \\
&\geq -\sqrt{\frac{a\hat{\sigma}^2}{WW}} \times \sqrt{\frac{a\hat{\sigma}^2 \times N-1}{N}} \\
&\geq -\sqrt{\frac{a\hat{\sigma}^2 \times N-1}{N}} \geq -\sqrt{a}\hat{\sigma}.
\end{aligned}$$

This completes the proof.

The following theorem states the our lower bound for $P(D_2)$.

Theorem 6.4.1. *The lower bound $P(D_2)$ is*
$P(D_{L2}) = P(-s + \sqrt{a}\hat{\sigma} \leq Z_{ij} \leq s - \sqrt{a}\hat{\sigma}, a\hat{\sigma}^2 \leq WW)$, *where* $Z_{ij} = X_{ij} + \theta_{ij}$.

Proof:

Let $W_{ij} = V_{ij} + \theta_{ij}$ and $WW = \sum_{i=1}^{I} \sum_{j=1}^{J} W_{ij}^2$.

$$\begin{aligned} P(D_2) &= P\left\{|X_{0ij} - \theta_{ij}| \leq s, a \leq \frac{WW}{\hat{\sigma}^2}\right\} + P\left\{|R^+W_{ij} + X_{0ij} - \theta_{ij}| \leq s, a \geq \frac{WW}{\hat{\sigma}^2}\right\} \\ &\geq P\left\{|R^+W_{ij} + X_{0ij} - \theta_{ij}| \leq s, a \leq \frac{WW}{\hat{\sigma}^2}\right\} \\ &\geq P\left\{|-(1-R^+)W_{ij} + X_{0ij} + W_{ij} - \theta_{ij}| \leq s, a \leq \frac{WW}{\hat{\sigma}^2}\right\} \\ &\geq P\left\{|-(1-R^+)W_{ij} + Z_{ij}| \leq s, a \leq \frac{WW}{\hat{\sigma}^2}\right\} \end{aligned}$$

Then from the previous lemma,

$$\begin{aligned} P(D_2) &\geq P\left\{-s + (1-R^+)W_{ij} \leq Z_{ij} \leq s + (1-R^+)W_{ij}, a \leq \frac{WW}{\hat{\sigma}^2}\right\} \\ &\geq P\left\{-s + \sqrt{a}\hat{\sigma^2} \leq Z_{ij} \leq s - \sqrt{a}\hat{\sigma^2}, a \leq \frac{WW}{\hat{\sigma}^2}\right\} = P(D_{L2}) \end{aligned}$$

Keep in mind that $P(D_{L2}) \geq P\left\{-s + \sqrt{a\frac{N-1}{N}\hat{\sigma^2}} \leq Z_{ij} \leq s - \sqrt{a\frac{N-1}{N}\hat{\sigma^2}}, a \leq \frac{WW}{\hat{\sigma}^2}\right\}$

Theorem 6.4.2. *$P(D_2)$ dominate $P(D_1)$ up to an arbitrarily small constant for a sufficiently large number of means, N.*

Proof:

Since a is any linear function of N such that $\lim_{sup} \frac{a}{N} = 0$, From lemma 3.2.4 $P(a \leq \frac{WW}{\hat{\sigma}^2}) = 1$. The lower limit $P(D_{L2}) = P\left\{-s + \sqrt{a}\hat{\sigma} \leq Z_{ij} \leq s - \sqrt{a}\hat{\sigma},\right\}$. Then, observe that Z_{ij}'s are independent and have t distribution.

$$P\left\{|Z_{ij}| \leq s - \sqrt{a}\hat{\sigma}\right\} = P\left\{|Z_{ij}| \leq s(1 - \frac{\sqrt{a}\hat{\sigma}}{s})\right\} \tag{6.11}$$

As $N \to \infty$,

$$\begin{aligned} \hat{\sigma} &\to \sigma \\ \frac{\sqrt{a}}{s} &\to 0 \; therefore \\ 1 - \frac{\sqrt{a}\hat{\sigma}}{s} &\to 0. \end{aligned}$$

Then conditioning on $\hat{\sigma}$,

$$\lim_{N\to\infty} E_{\hat{\sigma}} P(|Z_{ij}| \le s(1 - \frac{\sqrt{a}\hat{\sigma}}{s})/\hat{\sigma})$$

Keep in mind that $\hat{\sigma}^2$ has chi-square distribution and $|P(|Z_{ij}| \le s(1 - \frac{\sqrt{a}\hat{\sigma}}{s}))| \le 1$. Then all the conditions for dominated coverage theorem are met. Therefore we can take the limit inside,

$$\begin{aligned} \lim_{N\to\infty} E_{\hat{\sigma}} P(|Z_{ij}| \le s(1 - \frac{\sqrt{a}\hat{\sigma}}{s})/\hat{\sigma}) &= E_{\hat{\sigma}} \lim_{N\to\infty} P(|Z_{ij}| \le s(1 - \frac{\sqrt{a}\hat{\sigma}}{s})/\hat{\sigma}) \\ &= P(|Z_{ij}| \le s) = P(D_1) \end{aligned}$$

6.5 SIMULATION RESULTS

Since we want to show the coverage probability for our rectangular confidence procedure, $P(D_2)$, is uniformly higher than the coverage probability for the usual rectangular confidence procedure, $P(D_1)$, for small N and $\|\theta\|^2$. We run the simulation for the degrees of freedom, *df*, of interaction effect since we shrinkage the interactions toward zero. We use *df*=2,..,6. In other words we run the simulations for the following designs, matrixs, $(2\times 2, 3\times 2, 4\times 2, 5\times 2, 3\times 3, 6\times 2, 7\times 2, 4\times 3)$. Before we explain how we did simulations, we state a couple of $lemma's$ about $P(D_2)$.

Lemma 6.5.1. *$P(D_2)$ is sign invariant.*

Lemma 6.5.2. *$P(D_2)$ is transpose invariant.*

This lemma proves that the coverage probability of a matrix equals the coverage probability of the transpose of the same matrix.

Lemma 6.5.3. *$P(D_2)$ is column invariant.*

This lemma proves that the coverage probability of a matrix is not going to change if you change the position of its columns.

Lemma 6.5.4. *$P(D_2)$ is row invariant.*

This lemma proves that the coverage probability of a matrix is not going to change if you change the position of its rows.

Lemma 6.5.5. *$P(D_2)$ is not row and column invariant.*

This lemma proves that the coverage probability of a matrix is going to change if you change the position of its rows and columns at the same time. D_2 can be written in the following way,

$$D_2 = \left\{ |X_{0ij} + R^+ V_{ij} - (1 - R^+)\theta_{ij}| \leq s \right\}$$

where $X_{0ij} = X_i + X_j - X$, $V_{ij} = X_{ij} - X_{0ij}$, $X_i = \frac{1}{J}\sum_{j=1}^{J} X_{ij}$, $X_j = \frac{1}{I}\sum_{i=1}^{I} X_{ij}$ and $X = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} X_{ij}$. R^+ and s are defined in 6.1 and 6.3. For the simulations for the one-way ANOVA model, we mainly use the following set to create θ vectors. The set is $(-4, -3.75, ..., 3.75, 4)$. Since the simulations for two-way ANOVA model require more computing time than the simulations for one-Way ANOVA model, we use the same set above to create matrixs of θ for small dimensions. For big dimensions, we use a smaller set. This set is $(-3, ..., 3)$.

To do simulations we followed these steps:

Step 1: Generate Z_{IJ}^k ,$k = 1, ..., 10,000$ and store them where Z_{IJ} is a $I \times J$ matrix.

Step 2: Generate all the possible η_{IJ}^l by taking sign, row, transpose, and column invariances into account from the set above where $l = 1, ...L$, and L is the total number of η_{IJ} generated, and η_{IJ} is a $I \times J$ matrix.

Step 3: Calculate s.

Step 4: For each $l = 1, .., L$, calculate

$$PS_{lk} = X_{0IJ} + R^+ V_{IJ} - (1 - R^+)\eta_{IJ}^l \; for \; each \; k = 1, ..., 10,000$$

$,PS_{lk}$ is a $I \times J$ matrix.

Step 5: For $l = 1, ..., L$ and $k = 1, ..., 10,000$, let $CP_l = 1$ if all the values in PS_{lk} are in [-s,s] and 0 o.w.

Step 6: Calculate the coverage probability for $l = 1, .., L$, $P_l = \sum_{k=1}^{10,000} \frac{CP_{lk}}{10,000}$

(See appendix for the R codes).

For $df = 1$, there is only one design matrix which is a 2×2 design matrix. We do not expect to see that $P(D_2)$ dominates $P(D_1)$ in every region, since df is 1. We run simulations for the different choices of the shrinkage constant, a, but $P(D_2)$ could not dominate $P(D_1)$. We are not surprised by this result. We do not except see the domination result until df is 4. We also plot the coverage probabilities against the length of θ, $\sum_{i=1}^{I} \sum_{j=1}^{J} \theta_{ij}^2$. The plot looks like a random plot, there seems to be no relationship between the coverage probabilities and the length of θ. (See Figure 8.26)

For $df = 2$, there is only one design matrix which is a 2×3 design matrix. We run the simulations for different choices of a. When all the population cell means, θ, are zero, $P(D_2)$ dominates $P(D_1)$ and $P(D_2)$ achieves its maximum for any choice of a. However, $P(D_2)$ can not dominate $P(D_1)$ in every region. We plot the coverage probabilities against the length of θ for each a. The coverage probabilities are decreasing until the mid values of the length, then it starts increasing again. The sharpness of this dip depends on choice of a. As a gets bigger, this dip is getting sharper.(See Figure 8.26)

In a one-way ANOVA model, we plot the coverage probabilities against the length of θ and the maximum of $|\eta|$. Those two graphs look very similar; therefore we plot the coverage probabilities against the length of θ in a two-way ANOVA model.

For $df = 3$, there is only one design matrix which is a 4×2 design matrix. We first plot the coverage probabilities against the length of θ for each a, again there is a dip and coverage probabilities are not a decreasing function of length of θ. Then we plot the coverage probabilities against the maximum $|\eta|$, again there is a dip but the dip is not as sharp as in the first graph. As a gets smaller, this dip is getting smaller like in the first graph. When θ is around zero, $P(D_2)$ is uniformly higher than $P(D_1)$ and $P(D_2)$ achieves its maximum at $\theta = 0$ for any choice of a . When a is less than 1, $P(D_2)$ is uniformly higher than $P(D_1)$

everywhere. Therefore the optimum choice for a is 1. We are surprised by this result because in the previous studies, Lindley, Casella and Hwang proved that the shrinkage estimator we use has smaller TMSE than the usual estimator for $df \geq 4$. (See Figure 8.26)

For $df = 4$, there are two design matrixs which are 5×2 and 3×3. We again try to find the optimum shrinkage constant, a. For the two design matrixs; when θ is around zero, $P(D_2)$ is uniformly higher than $P(D_1)$ for any choice of a in every region and $P(D_2)$ achieves its maximum at $\theta = 0$. When a is less than 2, $P(D_2)$ is uniformly higher than $P(D_1)$ everywhere. Therefore the optimum choice for a is 2. Again, we agree with the usual recommendation. We plot the coverage probabilities against the length of θ for the two design matrixs. We have similar pictures, there is a dip and the sharpness of that dip depends on a. The dip in the second graph is not as sharp as the dip in the first graph. Also, we observe that the coverage probabilities for 3×3 is higher than that of 5×2. (See Figure 8.26-8.27)

For $df = 5$, there is only one design matrix which is a 6×2 design matrix. We first plot the coverage probabilities against the length of θ for each a, again there is a dip and coverage probabilities are not decreasing function of length of θ. As a gets smaller, this dip is getting smaller like in the first graph. When θ is around zero, $P(D_2)$ is uniformly higher than $P(D_1)$ and $P(D_2)$ achieves its maximum at $\theta = 0$ for any choice of a . When a is less than 3, $P(D_2)$ is uniformly higher than $P(D_1)$ everywhere. The optimum choice for a is 3. (See Figure 8.26-8.27)

For $df = 6$, there are two design matrixs which are 7×2 and 4×3. We again try to find the optimum shrinkage constant, a. For the two design matrixs; when θ is around zero, $P(D_2)$ is uniformly higher than $P(D_1)$ for any choice of a in every region and $P(D_2)$ achieves its maximum at $\theta = 0$. When a is less than 4, $P(D_2)$ is uniformly higher than $P(D_1)$ everywhere. Therefore the optimum choice for a is 4. Again, we agree with the usual recommendation. We plot the coverage probabilities against the length of θ for the two design matrixs. We have similar pictures, there is a dip and the sharpness of that dip depends on a. The dip in the second graph is not as sharp as the dip in the first graph. Also, we observe that the coverage probabilities for 4×3 is much higher than that of 7×2. (See Figure **??**-8.27)

By using simulations, we showed $P(D_2)$ is uniformly higher than $P(D_1)$, for small design matrixs. We plot the coverage probabilities first against the length of θ , than against the maximum of the $|\theta|$ for each design matrix. We have similar pictures in both graphs. The only difference is the second graph is smoother then the first graph. There is a small dip, but the dip is a function of a, when a gets smaller, the dip is getting smaller. The coverage probabilities are not a decreasing function of either of the the length of θ or the maximum $|\theta|$. We also agree with the usual recommendation for a. Our simulations indicate the domination of our procedure over the usual one when $df = 3$. That is an improvement and we are quite surprised to see the domination result for $df = 3$.

7.0 CONCLUSION AND FUTURE RESEARCH

If the researcher is interested in finding the simultaneous confidence interval for the independent samples normally distributed random variables, our procedure can be applied.

In this dissertation, we concentrate on simultaneous confidence intervals for the cell means, and the comparison of treatment means with the mean of a control. We make use of Stein type Shrinkage estimators as centers to improve the simultaneous coverage of those confidence intervals. Basically, we study the rectangular confidence region centered at a design appropriate shrinkage estimator in one way and two way ANOVA models. The main obstacle to an analytic study of the coverage probabilities of such regions, as compared to studies of coverage probabilities of similarly centered spherical confidence region is that the rectangular confidence regions are not rotation invariant. We briefly state our results and make some suggestions for the future work.

In this dissertation, we primarily use simulation to show dominance of the rectangular confidence intervals centered around a shrinkage estimator over the usual rectangular confidence regions centered about the sample means.

For the one-way ANOVA model, our simulation results indicate that our confidence procedure has higher coverage probability than the usual confidence procedure if the number of means is sufficiently large. We prove that the rectangular confidence intervals centered around a shrinkage estimator have coverage probability uniformly exceeding that of the usual rectangular confidence regions up to an arbitrarily small epsilon when the number of means is sufficiently large. We show that these intervals have strictly greater coverage probability when all the parameters are zero, and that the coverage probability of the two procedures converge to one another when at least one of the parameters becomes arbitrarily large. We also develop a lower bound for the coverage probability of our rectangular confidence region which is a decreasing function of the shrinkage constant for the estimator used as center.

To check the reliability of our simulations for the one-way ANOVA model, we use numerical integration to calculate the coverage probability for the rectangular confidence regions. Gaussion quadrature making use of Hermite polynomials is used to approximate the coverage probability of our rectangular confidence regions for n=2, 3, 4. The difference in results between numerical integration and simulations is negligible. However, numerical integration yields values slightly higher than the simulations.

A similar approach is applied to develop improved simultaneous confidence intervals for the comparison of treatment means with the mean of a control. We again develop a lower bound for the coverage probability of our confidence procedure and prove results similar to those that we proved for one-way model.

We also apply our approach to develop improved simultaneous confidence intervals for the cell means for a two-way ANOVA model. We again primarily use simulation to show dominance of the rectangular confidence intervals centered around an appropriate shrinkage estimator over the usual rectangular confidence regions. We again develop a lower bound for the coverage probabilities of our confidence procedure and prove the same results that we proved for the one-way model. Our simulations indicate that our confidence procedure has higher coverage probability than the usual confidence procedure for, $df \geq 3$. That is an improvement because Lindley, Casella and Hwang proved that the shrinkage estimator with the shrinkage factor that we used in our confidence procedure has a smaller TMSE than usual one for $df \geq 4$.

Since our confidence rectangles are not rotation invariant, it is difficult to come up with a proof for domination result. To calculate coverage probability of our procedure, the integrals must be evaluated. Because of the shrinkage factor we used, the coverage probability is a nonlinear function of the cell means. From the graphs based on our simulation, we see that coverage probability is not a convex or concave function of the cell means. To overcome this and to make the coverage probability function less complex, we tried a shrinkage estimator with the shrinkage factor

$$R^+ = (1 - \frac{a}{n \times max(X_i - \overline{X})^2})^+$$

but the confidence region based on that shrinkage estimator did not have uniformly higher coverage probability than the usual confidence region. We also use the original James-Stein estimator as our shrinkage estimator,

7.1 THE JAMES-STEIN ESTIMATOR

We want to examine the rectangular confidence interval for the cell mean centered at the James-Stein estimator. The James-Stein estimator is,

$$R^{+} = (1 - \frac{a}{\sum_{i=1}^{n} X_i^2})^{+} \tag{7.1}$$

,where R^{+} is defined in 3.11. We hope to prove the same results for the James-Stein estimator for the one-way model that we earlier have shown for the Lindley's estimator for the one-way model.

Let the shrinkage factor to be R^{+}. Keep in mind that $0 \leq R^{+} \leq 1$, that is the only condition we need for the most of the analytical results stated below.

7.1.1 Analytical Results

The rectangular confidence interval centered at the James-Stein estimator is in the following form,

$$C_1 = \left\{ |R^{+} X_i - \theta_i| \leq c \right\}. \tag{7.2}$$

The usual rectangular confidence interval C_0 is The usual confidence interval is

$$C_1 = \{|Z_i| \leq c\}.$$

Theorem 7.1.1. *$P(C_1)$ is uniformly higher than $P(C_0)$ if $\|\theta\|^2 = \sum_{i=1}^{n} \theta_i^2 \leq c^2$.*

Proof:

$\|\theta\|^2 \leq c^2$ implies that $\max_i |\theta_i| \leq c$. Then from triangular inequality,

$$\begin{aligned} |R^{+} X_i - \theta_i| &= |R^{+} Z_i - (1 - R^{+})\theta_i| \\ &\leq |R^{+} Z_i| + |(1 - R^{+})\theta_i| = R^{+}|Z_i| + (1 - R^{+})|\theta_i| \\ &\leq R^{+} c + (1 - R^{+}) c = c. \end{aligned}$$

Lemma 7.1.1. *if* $\theta'\theta \to \infty$, $P(C_1) \to P(C_0)$.

Proof:

Proof of this lemma is almost identical to the proof of lemma 3.1.4.

Lemma 7.1.2. *If* $a \leq SS$, *then* $(1-R^+)X_i = \frac{aX_i}{SS} \leq |\sqrt{\frac{a\times n-1}{n}}| \leq |\sqrt{a}|$, *where* $SS = \sum_i^n X_i^2$.

Proof:

$a \leq SS = \sum_i^n X_i^2$ implies that

$$\begin{aligned} |\frac{aX_i}{\sum_i^n X_i^2}| &\leq \sqrt{\frac{a}{\sum_{i=1}^n X_i^2}}\sqrt{a}\sqrt{\frac{\max_i X_i^2}{\sum_{i=1}^n X_i^2}} \\ &\leq \sqrt{a}. \end{aligned}$$

The following theorem states the our lower bound for $P(C_1)$.

Theorem 7.1.2. *The lower bound* $P(C_1)$ *is*
$P(C_{L1}) = P(-c+\sqrt{a} \leq Z_i \leq c-\sqrt{a}, a \leq SS)$, *where* $X_i = Z_i + \theta_i$.

Proof:

Let $X_i = Z_i + \theta_i$.

$$\begin{aligned} P(C_1) &= P\left\{|\theta_i| \leq c, a \leq SS\right\} + P\left\{|R^+X_i - \theta_i| \leq c, a \geq SS\right\} \\ &\geq P\left\{|R^+X_i - \theta_i| \leq c, a \leq SS\right\} \\ &\geq P\left\{|X_i - \frac{a}{SS}X_i - \theta_i| \leq c, a \leq SS\right\} \\ &\geq P\left\{|Z_i - \frac{a}{SS}X_i| \leq c, a \leq SS\right\} \\ &\geq P\left\{-c + \frac{a}{SS}X_i \leq Z_i \leq c + \frac{a}{SS}X_i, a \leq SS\right\}. \end{aligned}$$

Then from the previous lemma,

$$\begin{aligned} P(C_1) &\geq P\left\{-c + \frac{a}{SS}X_i \leq Z_i \leq c + \frac{a}{SS}X_i, a \leq SS\right\} \\ &\geq P\left\{-c + \sqrt{a} \leq Z_i \leq c - \sqrt{a}, a \leq SS\right\}. \end{aligned}$$

This completes the proof.

Lemma 7.1.3. *As $n \to \infty$, $P(C_{L1}) = (2\Phi(c - \sqrt{a}) - 1)^n$*

Proof:

As $n \to \infty$, $P(a \leq SS) = 1$, and this completes the proof.

Theorem 7.1.3. *$P(C_1)$ dominate $P(C_0)$ up to an arbitrarily small constant for a sufficiently large number of means, n.*

Proof:

As $n \to \infty$, $P(a \leq SS) = 1$ since $a \leq n$ and a is a any linear function of n. Then,

$$\begin{aligned}
P(C_1) &= P\left\{|R^+X_i - \theta_i| \leq c, a \leq SS\right\} \\
&= P\left\{|X_i - \frac{a}{SS}X_i - \theta_i| \leq c\right\} \\
&= P\left\{|Z_i - \frac{a}{SS}X_i| \leq c\right\} \\
&= P\left\{-c + \frac{a}{SS}X_i \leq Z_i \leq c + \frac{a}{SS}X_i\right\}
\end{aligned}$$

Then from previous lemma,

$$\begin{aligned}
P(C_1) &= P\left\{-c + \frac{a}{SS}X_i \leq Z_i \leq c + \frac{a}{SS}X_i\right\} \\
&\geq (2\Phi(c - \sqrt{a}) - 1)^n.
\end{aligned}$$

The rest of the proof is the similar the proof of theorem 3.1.3.

7.1.2 Simulation Results

We run simulations for $n = 3, 4, 5, 6, 7$. We used the simulation method that we described in the previous section. We have exactly matching pictures. The coverage probability is not a monotone decreasing function of the length of θ and the coverage probability achieves its maximum when all the $\theta's$ are zero and the optimum choice for the shrinkage constant is $n - 3$.

One can try different shrinkage factors. If the shrinkage factor makes the coverage probability function a Schure concave function, the proof for the domination result will follow easily. Another way to prove the domination result is to come up with a sharper lower than ours and work on the lower bound to prove the domination result.

Since our confidence procedure has substantially higher coverage probability than the usual confidence procedure for $n \geq 4$ in one way model and $df \geq 3$ in two way model, it may be possible to reduce the volume of the rectangular confidence region while still maintaining superior coverage probability relative to the usual procedure. In other words, it may be possible to permit the length of the interval to be function of the data. In two way model, we let the length of the interval to be a function of data but unfortunately that confidence interval could not dominate the usual one.

It is our hope that our research will contribute to the field of statistical inference and eventually help to applied statisticians.

8.0 APPENDIX

8.1 UNBALANCED ONE AND TWO WAY ANOVA MODELS

If the data Y_{ij} are obtained from an unbalanced one way classification design, then $Y_{ij} = \theta_i + \epsilon_{ij}, \quad j = 1, ..., n_i, \; i = 1, ..., k$. And our notation remains the same as before, except that the sample cell mean is replaced by:

$$X_i = \sqrt{n_i}\overline{Y_i} \tag{8.1}$$

,where $\overline{Y_i} = \sum_{j=1}^{n_i} \frac{Y_{ij}}{n_i}$. Then the X_i's are distributed independently and normally as

$$N(\sqrt{n_i}\theta_i, \sigma^2), \quad i = 1, .., k. \tag{8.2}$$

Let $\gamma = \sqrt{n_i}\theta_i$. The usual confidence rectangle for the vector Θ of cell means now becomes

$$\begin{aligned} C_0 &= \{\Gamma = (\gamma_1, ..., \gamma_k) : |X_i - \gamma_i| \leq c \;, i = 1, ..., k\} \\ &= \left\{\Theta = (\theta_1, ..., \theta_k) : |\overline{Y_i} - \theta_i| \leq \frac{c}{\sqrt{n_i}} \;, i = 1, ..., k\right\} \end{aligned}$$

, where $c = \sigma T^{-1}(\frac{\beta^{1/k} - 1}{2})$ and T is the distribution function of $\frac{\overline{Y_i} - \theta_i}{\frac{\sigma}{\sqrt{n_i}}}$. And the confidence rectangle we are investigating converts into

$$\begin{aligned} C_1 &= \left\{\Gamma = (\gamma_1, ..., \gamma_k) : |R^+(X_i - \overline{X}) + \overline{X} - \gamma_i| \leq c, i = 1, ...k\right\} \\ &= \left\{\Theta = (\theta_1, ..., \theta_k) : |R^+(\overline{Y_i} - \frac{\overline{X}}{\sqrt{n_i}}) + \frac{\overline{X}}{\sqrt{n_i}} - \theta_i| \leq \frac{c}{\sqrt{n_i}}, i = 1, ...k\right\} \\ &= \left\{\Theta : |R^+\overline{Y_i} + (1 - R^+)\frac{\overline{X}}{\sqrt{n_i}} - \theta_i| \leq \frac{c}{\sqrt{n_i}}, i = 1, ...k\right\} \end{aligned}$$

,where R^+ is defined in 3.11. When σ is unknown, we replace σ with an unbiased estimator in C_1 defined above. Similar method will be applied to two way ANOVA model.

8.2 PROOFS

Proof of Lemma 3.1.5:

Let U *be normal*, $N(0,1/2)$, *p.d.f.* For n=2, the coverage probability is

$$P(C_1) = \begin{cases} \int_{-\sqrt{a/2}-\eta_1}^{\sqrt{a/2}-\eta_1} (2\Phi(\sqrt{2}(c-|\eta_1|))-1) f_U(U)dU + \\ \int_{\frac{-(\eta_1+c)-\sqrt{(c-\eta_1)^2+2a}}{2}}^{-\sqrt{a/2}-\eta_1} g_{U,\eta_1,c,a}(U) f_U(U)dU \\ + \int_{\sqrt{a/2}-\eta_1}^{\frac{(c-\eta_1)+\sqrt{(c+\eta_1)^2+2a}}{2}} g_{U,\eta_1,c,a}(U) f_U(U)dU & : \quad 0 \le \eta_1 \le \sqrt{a/2} \\ \\ \int_{-\sqrt{a/2}-\eta_1}^{\sqrt{a/2}-\eta_1} (2\Phi(\sqrt{2}(c-|\eta_1|))-1) f_U(U)dU + \\ \int_{\frac{-(\eta_1+c)-\sqrt{(c-\eta_1)^2+2a}}{2}}^{-\sqrt{a/2}-\eta_1} g_{U,\eta_1,c,a}(U) f_U(U)dU \\ + \int_{\sqrt{a/2}-\eta_1}^{\frac{(c-\eta_1)+\sqrt{(c+\eta_1)^2+2a}}{2}} g_{U,\eta_1,c,a}(U) f_U(U)dU & : \quad \sqrt{a/2} \le \eta_1 \le c \\ \\ \int_{\frac{-(\eta_1+c)+\sqrt{(c-\eta_1)^2+2a}}{2}}^{\frac{-(\eta_1-c)+\sqrt{(c+\eta_1)^2+2a}}{2}} g_{U,\eta_1,c,a}(U) f_U(U)dU & : \quad \eta_1 \ge c \end{cases} \tag{8.3}$$

where $U_{v,\eta_1,c,a}(U) = 2\Phi(\sqrt{2}(c - |U - \frac{a}{2(U+\eta_1)}|)) - 1)$.

Wlog assume $\eta_1 \ge 0$, there three three regions we need to consider;

1.$0 \le \eta_1 \le \sqrt{a/2}$

2.$\sqrt{a/2} \le \eta_1 \le c$

3.$\eta_1 \ge c$

1.The first region ($0 \le \eta_1 \le \sqrt{a/2}$):

Let $u = u_{[1]}$, $\eta = \eta_{[1]}$
$P(C_1) = \int_u \int_{\sqrt{2}(-c+|Ru-(1-R)\eta)}^{\sqrt{2}(c-|Ru-(1-R)\eta)}$

$$R = \begin{cases} 0 & : \quad A = \left\{U : -\sqrt{a/2} - \eta \leq u \leq \sqrt{a/2} - \eta\right\} \\ U - \frac{a}{2(U+\eta)} & : o.w \end{cases}$$

then,
$P(C_1) = \int_{-\sqrt{a/2}-\eta}^{\sqrt{a/2}-\eta} \int_{\sqrt{a/2}(-c+|\eta|)}^{\sqrt{a/2}(c-|\eta|)} \varphi(v)f(u)dvdu + \int_{A^c} \int_{\sqrt{a/2}(-c+|U-\frac{a}{2(U+\eta)}|)}^{\sqrt{a/2}(c-|U-\frac{a}{2(U+\eta)}|)} \varphi(v)f(u)dvdu$

we need to check the upper and lower bounds of integration.

Check1:
$\sqrt{2}(c - |\eta|) > \sqrt{2}(-c + |\eta|)$
$s > \eta$ (yes)
Check2:

$$\sqrt{2}(c - |U - \frac{a}{2(U+\eta)}|) > \sqrt{2}(-c + |U - \frac{a}{2(U+\eta)}|)$$
$$c > |U - \frac{a}{2(U+\eta)}|$$
$$c > U - \frac{a}{2(U+\eta)} > -c$$

Check 2.1:

$$\begin{aligned} c &= U - \frac{a}{2(U+\eta)} \\ 2Uc + 2c\eta &= 2U^2 + 2U\eta - a \\ &= 2U^2 + 2(\eta - c)U - a - 2c\eta \\ 0 &= U^2 + (\eta - c)U - a/2 - c\eta \\ U_{1,2} &= \frac{-(\eta - c) \pm \sqrt{(\eta + c)^2 + 2a}}{2} \end{aligned}$$

Check 2.1.1:
We need to check if $U_{1,2} \in A$.
1.

$U_1 = \frac{-(\eta-c)+\sqrt{(\eta+c)^2+2a}}{2} > \sqrt{a/2} - \eta$

$c - \eta - \sqrt{2a} + 2\eta + \sqrt{(c+\eta)^2 + 2a}$

$c + \eta - \sqrt{2a} + sqrt(c+\eta)^2 + 2a > 0$

so $U_1 \notin A$.

2.

$U_2 = \frac{-(\eta-c)-\sqrt{(\eta+c)^2+2a}}{2} > -\sqrt{a/2} - \eta$

$c - \eta + \sqrt{2a} + 2\eta + \sqrt{(c+\eta)^2 + 2a}$

$c + \eta + \sqrt{2a} - \sqrt{(c+\eta)^2 + 2a} > 0$

so $U_2 \in A$.

Check 2.2:

$$\begin{aligned}
-c &= U - \frac{a}{2(U+\eta)} \\
-2Uc - 2c\eta &= 2U^2 + 2U\eta - a \\
&= 2U^2 + 2(\eta+c)U - a + 2c\eta \\
0 &= U^2 + (\eta+c)U - a/2 + c\eta \\
U_{1,2} &= \frac{-(\eta+c) \pm \sqrt{(\eta-c)^2 + 2a}}{2}
\end{aligned}$$

1:

$U_1 = \frac{-(\eta+c)-\sqrt{(c-\eta)^2+2a}}{2} < -\sqrt{a/2} - \eta$

$-c - \eta + \sqrt{2a} + 2\eta - \sqrt{(c-\eta)^2 + 2a}$

$-(c-\eta) + \sqrt{2a} - \sqrt{(c-\eta)^2 + 2a} < 0$

so $U_1 \notin A$.

2:

$U_2 = \frac{-(\eta+c)+\sqrt{(\eta-c)^2+2a}}{2} > -\sqrt{a/2} - \eta$

$c - \eta + \sqrt{2a} + 2\eta + \sqrt{(c-\eta)^2 + 2a}$

$-(c-\eta) + \sqrt{2a} + \sqrt{(c-\eta)^2 + 2a} > 0$

so $U_2 \in A$.

Then

$$\int_{-\sqrt{a/2}-\eta_1}^{\sqrt{a/2}-\eta_1} (2\Phi(\sqrt{2}(c-|\eta_1|))-1)f_u(u)du \quad + \quad \int_{\frac{-(\eta_1+c)-\sqrt{(c-\eta_1)^2+2a}}{2}}^{-\sqrt{a/2}-\eta_1} g_{u,\eta,c,a}(u)f_u(u)du$$
$$+ \quad \int_{\sqrt{a/2}-\eta_1}^{\frac{(c-\eta_1)+\sqrt{(c+\eta_1)^2+2a}}{2}} g_{u,\eta,c,a}(u)f_u(u)du$$

where,$g_{U,\eta,c,a}(U) = 2\Phi(\sqrt{2}(c-|U-\frac{a}{2(U+\eta)}|))-1)$.
Also observe that $\frac{-(\eta_1+c)-\sqrt{(c-\eta_1)^2+2a}}{2} < -s$ and $\frac{(c-\eta_1)+\sqrt{(c+\eta_1)^2+2a}}{2} > s$.

2.The Second Region($\sqrt{a/2} \leq \eta_1 \leq c$)

Since $\eta < c$, we will get the same integrals as above.

3.The third Region($\eta_1 \geq c$):

Check 3.1:

$\sqrt{2}(c-|\eta|) > \sqrt{2}(-c+|\eta|)$
$c > |\eta|$, No.

Check 3.2:

$c = U - \frac{a}{2(u+\eta)}$ and $U_{12} = \frac{-(\eta-c)\pm\sqrt{(\eta+c)^2+2a}}{2}$

1:
$U_1 = \frac{-(\eta-c)+\sqrt{(\eta+c)^2+2a}}{2} > \sqrt{a/2}-\eta$, YES.
So, $U_1 \notin A$.

2:

$U_2 = \frac{-(\eta-c)-\sqrt{(\eta+c)^2+2a}}{2} < \sqrt{a/2} - \eta$, YES.
$U_2 = \frac{-(\eta-c)-\sqrt{(\eta+c)^2+2a}}{2} > -\sqrt{a/2} - \eta$, YES.

so $U_2 \in A$.

Check 3.3:

$-c = U - \frac{a}{2(u+\eta)}$ and $u_{12} = \frac{-(\eta+c)\pm\sqrt{(\eta-c)^2+2a}}{2}$

1:

$U_1 = \frac{-(\eta+c)-\sqrt{(\eta-c)^2+2a}}{2} < \sqrt{a/2} - \eta$, YES.
$U_1 = \frac{-(\eta+c)-\sqrt{(\eta-c)^2+2a}}{2} > -\sqrt{a/2} - \eta$,YES.

so $U_1 \in A$.

2:

$U_2 = \frac{-(\eta+c)+\sqrt{(\eta-c)^2+2a}}{2} > \sqrt{a/2} - \eta$, YES.

so $U_2 \notin A$.

then, $P(C_1) = \int_{\frac{-(\eta+c)+\sqrt{(\eta-c)^2+2a}}{2}}^{\frac{-(\eta-c)+\sqrt{(\eta+c)^2+2a}}{2}} g_{u,\eta,c,a}(u) f_u(u) du$.

Where, $g_{U,\eta,c,a}(U) = 2\Phi(\sqrt{2}(c - |U - \frac{a}{2(U+\eta)}|)) - 1)$.

Also observe that

$\frac{-(\eta+c)+\sqrt{(\eta-c)^2+2a}}{2} > -s$ and $\frac{-(\eta-c)+\sqrt{(\eta+c)^2+2a}}{2} > s$

Proof of Theorem 3.1.1:

There are three regions we need to consider from the previous lemma;

1.$0 \leq \eta_1 \leq \sqrt{a/2}$

2.$\sqrt{a/2} \leq \eta_1 \leq c$

3.$\eta_1 \geq c$

1.The first region ($0 \leq \eta_1 \leq \sqrt{a/2}$):

$$
\begin{aligned}
P(C_1) - P(C_0) &= 2\int_{-\sqrt{a/2}-|\eta|}^{\sqrt{a/2}-|\eta|} [\phi(\sqrt{2}(c-|\eta|)) - \phi\sqrt{2}(c-|u|))] f_u(u)du \\
&+ 2\int_{-c}^{\sqrt{a/2}-|\eta|} [\phi(\sqrt{2}(c-|u-\frac{a}{2(u+\eta)}|) - \phi\sqrt{2}(c-|u|)] f_u(u)du \\
&+ 2\int_{\sqrt{a/2}-|\eta|}^{c} [\phi(\sqrt{2}(c-|u-\frac{a}{2(u+\eta)}|) - \phi\sqrt{2}(c-|u|)] f_u(u)du \\
&+ \int_{\frac{-(\eta+c)-\sqrt{(c-\eta)^2+2a}}{2}}^{-c} [2\phi(\sqrt{2}(c-|u-\frac{a}{2(u+\eta)}|) - 1] f_u(u)du \\
&+ \int_{c}^{\frac{(c-\eta)+\sqrt{(c+\eta)^2+2a}}{2}} [2\phi(\sqrt{2}(c-|u-\frac{a}{2(u+\eta)}|) - 1] f_u(u)du
\end{aligned}
$$

There are 5 terms in $P(C_1) - P(C_0)$, we will take the first and second derivatives w.r.t $\sqrt{a}$ for term by term.

Let $\kappa = \frac{\eta}{\sqrt{a/2}} \Rightarrow \eta = \kappa\sqrt{a/2}$.

The Derivatives of the first term:

$1 = 2\int_{-\sqrt{a/2}-\kappa\sqrt{a/2}}^{\sqrt{a/2}-\kappa\sqrt{a/2}}[\phi(\sqrt{2}(c-|\kappa\sqrt{a/2}|)) - \phi\sqrt{2}(c-|u|))]f_u(u)du$
$1 = 2\int_{-\sqrt{a/2}(1+\kappa)}^{\sqrt{a/2}(1-\kappa)}[\phi(\sqrt{2}(c-|\kappa\sqrt{a/2}|)) - \phi\sqrt{2}(c-|u|))]f_u(u)du$

$$
\begin{aligned}
\frac{\partial 1}{\partial\sqrt{a}} &= 2[(\frac{1-\kappa}{\sqrt{2}}(\phi(\sqrt{2}(c-|\kappa\sqrt{a/2}|)) - \phi(\sqrt{2}(c-|(1-\kappa)\sqrt{a/2}|)))e^{-(\sqrt{a/2}(1-\kappa))^2} \\
&+ \frac{1+\kappa}{\sqrt{2}}(\phi(\sqrt{2}(c-|\kappa\sqrt{a/2}|)) - \phi(\sqrt{2}(c-|(1+\kappa)\sqrt{a/2}|)))e^{-(\sqrt{a/2}(1+\kappa))^2} \\
&+ \int_{-\sqrt{a/2}(1+\kappa)}^{\sqrt{a/2}(1-\kappa)}(\frac{-\kappa}{\pi})e^{-(c-|\kappa\sqrt{a/2}|)^2-u^2}du \\
&= \sqrt{2/\pi}(1-\kappa)e^{-(\sqrt{a/2}(1-\kappa))^2}[\phi(\sqrt{2}(c-|\sqrt{a/2}\kappa|) - \phi(\sqrt{2}(c-|\sqrt{a/2}(1-\kappa)|))] \\
&+ \sqrt{2/\pi}(1+\kappa)e^{-(\sqrt{a/2}(1-\kappa))^2}[\phi(\sqrt{2}(c-|\sqrt{a/2}\kappa|) - \phi(\sqrt{2}(c-|\sqrt{a/2}(1+\kappa)|))] \\
&- \sqrt{2/\pi}\kappa e^{-(c-|\sqrt{a/2}\kappa|)^2}[\phi(\sqrt{a}(1-\kappa|)) - \phi(-\sqrt{a}(1+\kappa))]
\end{aligned}
$$

$\frac{\partial 1}{\partial\sqrt{a}}|_{\sqrt{a}=0} = 0$

For the second derivative, when we apply product rule into the function above, we will drop the term which is zero for $\sqrt{a} = 0$

$$
\begin{aligned}
\frac{\partial^2 1}{\partial^2\sqrt{a}} &= \sqrt{2/\pi}(1-\kappa)[0 + e^{-(\sqrt{a/2}(1-\kappa))^2}[\frac{-\kappa}{\sqrt{2\pi}}e^{-(c-|\sqrt{a/2}\kappa|)^2} + \frac{1-\kappa}{\sqrt{2\pi}}e^{-(c-(\sqrt{a/2}(1-\kappa)))^2}]] \\
&+ \sqrt{2/\pi}(1+\kappa)[0 + \frac{-k}{\sqrt{2\pi}}e^{(c-|\sqrt{a/2}|)^2} + \frac{1+\kappa}{\sqrt{2\pi}}e^{-(c-|\sqrt{a/2}(1+\kappa)|)^2}] \\
&- \frac{\sqrt{2}\kappa}{\sqrt{\pi}}[e^{-(c-\kappa\sqrt{a/2})^2}(\frac{1-\kappa}{\sqrt{2\pi}}e^{-\sqrt{a}(1-\kappa)^2} + \frac{1+\kappa}{\sqrt{2\pi}}e^{-(\sqrt{a}(1+\kappa))^2} + 0)]
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 1}{\partial^2} 1_{\sqrt{a}=0} &= \sqrt{2/\pi}(1-\kappa)[0+e^0(\frac{-\kappa}{\sqrt{2\pi}}e^{-c^2}+\frac{1+\kappa}{\sqrt{2\pi}}e^{-c^2})] \\
&+ \sqrt{2/\pi}(1+\kappa)[0+e^0(\frac{-\kappa}{\sqrt{2\pi}}e^{-c^2}+\frac{1-\kappa}{\sqrt{2\pi}}e^{-c^2})] \\
&- \sqrt{2/\pi}\kappa e^{-c^2}(\frac{1-\kappa}{\sqrt{2\pi}}+\frac{1+\kappa}{\sqrt{2\pi}}) \\
&= \frac{1-\kappa}{\pi}e^{-c^2}((1-2\kappa)+\frac{(1+\kappa)e^{-c^2}}{\pi}(-\kappa+1+\kappa)-\kappa/\pi e^{-c^2}(1-\kappa+1+\kappa)) \\
&= \frac{(1-\kappa)(1-2\kappa)}{\pi}e^{-c^2}+\frac{1+\kappa}{\pi}e^{-c^2}-\frac{2\kappa}{\pi}e^{-c^2} \\
&= \frac{(1-\kappa)e^{-c^2}}{\pi}[(1-2\kappa)+1] \\
&= \frac{2(1-\kappa)^2e^{-c^2}}{\pi} > 0
\end{aligned}
$$

so $\sqrt{a}=0$ is the local min for the first term.

The Derivatives of the second term:

$2 = 2\int_{-c}^{-\sqrt{a/2}(1+\kappa)}(\phi(\sqrt{2}(c-|u-\frac{a}{2(u+\kappa\sqrt{a/2})}|))-\phi(\sqrt{2}(c-|u|)))f_u(u)du$

$$
\begin{aligned}
\frac{\partial 2}{\partial\sqrt{a}} &= 2[(\frac{-(1+\kappa)}{\sqrt{2\pi}}e^{-(\sqrt{a/2}(1+\kappa))^2}[\phi(\sqrt{2}(c-|-(1+\kappa)\sqrt{a/2} \\
&- \frac{a}{2(-\sqrt{a/2}(1+\kappa)+\kappa\sqrt{a/2}}|))-\phi(\sqrt{2}(c-|-\sqrt{a/2}(1+\kappa)|))] \\
&+ \frac{1}{\pi}\int_{-c}^{-\sqrt{a/2}(1+\kappa)}\frac{2\sqrt{a}(u+\kappa\sqrt{a/2}(-a\kappa/\sqrt{2}}{(u+\kappa\sqrt{a/2})^2}e^{-(c-|u-\frac{a}{2(u+\kappa\sqrt{a/2}}|)^2-u^2}
\end{aligned}
$$

$\frac{\partial 2}{\partial\sqrt{a}}|_{\sqrt{a}=0}=0.$

$$\begin{aligned}
\frac{\partial^2 2}{\partial^2\sqrt{a}}|_{\sqrt{a}=0} &= \frac{-\sqrt{2}}{\sqrt{\pi}}(1+\kappa)(\frac{-\kappa}{\sqrt{2\pi}}e^{-c^2}+\frac{1+\kappa}{\sqrt{2\pi}}e^{-c^2}) \\
&+ \frac{1}{\pi}(1+\kappa)(2+\kappa)e^{-c^2}+\frac{2}{\pi}\int_{-c}^{0}\frac{e^{-(c-|u|)^2-u^2}}{u}du \\
&= \frac{-(1+\kappa)}{\pi}e^{-c^2}+\frac{1}{\pi}(1+\kappa)(2+\kappa)e^{-c^2}+\frac{2}{\pi}\int_{-c}^{0}\frac{e^{-(c-|u|)^2-u^2}}{u}du
\end{aligned}$$

The Derivatives of the third term:

$3 = 2\int_{\sqrt{a/2}(1-\kappa)}^{s}(\phi(\sqrt{2}(c-|u-\frac{a}{2(u+\kappa\sqrt{a/2})}|)-\phi(\sqrt{2}(c-|u|))f_u(u)du$

$$\begin{aligned}
\frac{\partial 3}{\partial\sqrt{a}} &= -\sqrt{2/\pi}(1-\kappa)e^{-(\sqrt{a/2}(1-\kappa))^2}[\phi(\sqrt{2}(c-|-\kappa\sqrt{a/2}|))-\phi(\sqrt{(}c-|\sqrt{a/2}(1-\kappa)|))] \\
&+ \frac{1}{\pi}\int_{\sqrt{a/2}(1-\kappa)}^{s}\frac{2\sqrt{a}(u+\kappa\sqrt{a/2})-a\kappa/\sqrt{2}}{(u+\kappa\sqrt{a/2})^2}e^{-(c-|u-\frac{a}{2(u+\kappa\sqrt{a/2})}|)^2-u^2}
\end{aligned}$$

$\frac{\partial 3}{\partial\sqrt{a}}|_{\sqrt{a}=0} = 0.$

$$\frac{\partial^2 3}{\partial^3\sqrt{a}}|_{\sqrt{a}=0} = -\frac{(1-\kappa)(1-2\kappa)e^{-c^2}}{\pi}-\frac{1}{\pi}(1-\kappa)(2-\kappa)e^{-c^2}+\frac{2}{\pi}\int_{0}^{c}\frac{e^{-(c-|u|)^2-u^2}}{u}du$$

The Derivatives of the fourth term:

$4 = \int_{\Delta}^{-c}(2\phi(\sqrt{2}(c-|u-\frac{a}{2(u+\kappa\sqrt{a/2})}|)-1)f_u(u)du$

$\Delta = -\frac{\kappa\sqrt{a/2}+c-\sqrt{(c-\kappa\sqrt{a/2})^2+2a}}{2}$ and

$\Delta_{\sqrt{a}=0} = -c \ \Delta'_{\sqrt{a}=0} = 0.$

$$\begin{aligned} \frac{\partial 4}{\partial \sqrt{a}}|_{\sqrt{a}=0} &= 0 \\ \frac{\partial^2 4}{\partial^2 \sqrt{a}}|_{\sqrt{a}=0} &= 0 \end{aligned}$$

The Derivatives of the fifth term:

$5 = \int_c^{\Delta}(2\phi(\sqrt{2}(c - |u - \frac{a}{2(u+\kappa\sqrt{a/2})}|) - 1)f_u(u)du$
$\Delta = -frac(c - \kappa\sqrt{a/2}) + \sqrt{(c + \kappa\sqrt{a/2})^2 + 2a2}$ and
$\Delta_{\sqrt{a}=0} = -c \ \Delta'_{\sqrt{a}=0} = 0.$

$$\begin{aligned} \frac{\partial 5}{\partial 5\sqrt{a}}|_{\sqrt{a}=0} &= 0 \\ \frac{\partial^2 5}{\partial^2 5\sqrt{a}}|_{\sqrt{a}=0} &= 0 \end{aligned}$$

Adding the terms;

$sum_{i=1}^{5} \frac{\partial^2 i}{\partial^2 \sqrt{a}}|_{\sqrt{a}=0} > 0$,so that's a local min.

2.The second region $(\sqrt{a/2} \leq \eta \leq c)$:

$$
\begin{aligned}
P(C_1) - P(C_0) &= 2\int_{-\sqrt{a/2}-|\eta|}^{\sqrt{a/2}-|\eta|} [\phi(\sqrt{2}(c-|\eta|)) - \phi\sqrt{2}(c-|u|))] f_u(u)du \\
&+ 2\int_{-c}^{\sqrt{a/2}-|\eta|} [\phi(\sqrt{2}(c-|u-\frac{a}{2(u+\eta)}|) - \phi\sqrt{2}(c-|u|)] f_u(u)du \\
&+ 2\int_{\sqrt{a/2}-|\eta|}^{c} [\phi(\sqrt{2}(c-|u-\frac{a}{2(u+\eta)}|) - \phi\sqrt{2}(c-|u|)] f_u(u)du \\
&+ \int_{\frac{-(\eta+c)-\sqrt{(c-\eta)^2+2a}}{2}}^{-c} [2\phi(\sqrt{2}(c-|u-\frac{a}{2(u+\eta)}|) - 1] f_u(u)du \\
&+ \int_{c}^{\frac{(c-\eta)+\sqrt{(c+\eta)^2+2a}}{2}} [2\phi(\sqrt{2}(c-|u-\frac{a}{2(u+\eta)}|) - 1] f_u(u)du
\end{aligned}
$$

There are 5 terms in $P(C_1) - P(C_0)$, we will take the first and second derivatives w.r.t $\sqrt{a}$ for term by term.

The Derivatives of the first term:

$1 = 2\int_{-\sqrt{a/2}-\eta}^{\sqrt{a/2}-\eta} [\phi(\sqrt{2}(c-|eta|)) - \phi\sqrt{2}(c-|u|))] f_u(u)du$

$$
\begin{aligned}
\frac{\partial 1}{\partial\sqrt{a}} &= \sqrt{2/\pi}e^{-(\sqrt{a/2}-\eta)^2}[\phi(\sqrt{2}(c-|\eta|)) - \phi\sqrt{2}(c-|\sqrt{a/2}-|\eta||))] \\
&+ \sqrt{2/\pi}e^{-(\sqrt{a/2}+\eta)^2}[\phi(\sqrt{2}(c-|\eta|)) - \phi\sqrt{2}(c-|\sqrt{a/2}+|\eta||))]
\end{aligned}
$$

$\frac{\partial 1}{\partial\sqrt{a}}|_{\sqrt{a}=0} = 0$

$\frac{\partial^2 1}{\partial^2\sqrt{a}}|_{\sqrt{a}=0} = \frac{2}{\pi}e^{-\eta^2-(c-\eta)^2} > 0$ so it's a local min.

The Derivatives of the second term:

$2 = 2\int_{-c}^{-\sqrt{a/2}-\eta}(\phi(\sqrt{2}(c-|u-\frac{a}{2(u+\eta)}|)) - \phi(\sqrt{2}(c-|u|)))f_u(u)du$

$$\begin{aligned}\frac{\partial 2}{\partial\sqrt{a}} &= -\sqrt{2/\pi}e^{-(\sqrt{a/2}+\eta)^2}(\phi(\sqrt{2}(c-|\eta|)) - \phi(\sqrt{2}(c-|\sqrt{a/2}+\eta|))\\ &+ \frac{2}{\pi}\int_{-c}^{-\sqrt{a/2}-\eta}\frac{\sqrt{a}}{u+\eta}e^{-(c-|u-\frac{a}{2(u+\eta)}|)^2-v^2}\end{aligned}$$

$\frac{\partial 2}{\partial\sqrt{a}}|_{\sqrt{a}=0} = 0$
$\frac{\partial^2 2}{\partial\sqrt{a}}|_{\sqrt{a}=0} = -\frac{1}{\pi}e^{-\eta^2-(c-\eta)^2} + \frac{2}{\pi}\int_{-c}^{-\eta}\frac{e^{-(s-|u|)^2-u^2}}{u+\eta}du$

$3 = 2\int_{\sqrt{a/2}-\eta}^{s}(\phi(\sqrt{2}(c-|u-\frac{a}{2(u+\eta)}|) - \phi(\sqrt{2}(c-|u|))f_u(u)du$

$$\begin{aligned}\frac{\partial 3}{\partial\sqrt{a}} &= -\sqrt{2/\pi}e^{-(\sqrt{a/2}-\eta))^2}[\phi(\sqrt{2}(c-|\eta|)) - \phi(\sqrt{(}c-|\sqrt{a/2}-\eta)|))]\\ &+ \frac{2}{\pi}\int_{\sqrt{a/2}-\eta}^{c}\frac{\sqrt{a}}{u+\eta}e^{-(s-|u-\frac{a}{2(u+\eta)}|)^2-u^2}\end{aligned}$$

$\frac{\partial 3}{\partial\sqrt{a}}|_{\sqrt{a}=0} = 0$
$\frac{\partial^2 3}{\partial\sqrt{a}}|_{\sqrt{a}=0} = -\frac{1}{\pi}e^{-\eta^2-(c-\eta)^2} + \frac{2}{\pi}\int_{-\eta}^{c}\frac{e^{-(s-|u|)^2-u^2}}{u+\eta}du$

The Derivatives of the fourth term:

$4 = \int_{c}^{\Delta}(2\phi(\sqrt{2}(c-|u-\frac{a}{2(u+\eta)}|) - 1)f_u(u)du$

$\Delta = \frac{c-\eta+\sqrt{(c+\eta)^2+2a}}{2}$ and

$\Delta_{\sqrt{a}=0} = c \; \Delta'_{\sqrt{a}=0} = 0.$

$$\begin{aligned} \frac{\partial 4}{\partial \sqrt{a}}|_{\sqrt{a}=0} &= 0 \\ \frac{\partial^2 4}{\partial^2 \sqrt{a}}|_{\sqrt{a}=0} &= 0 \end{aligned}$$

The Derivatives of the fifth term:

$$\begin{aligned} \frac{\partial 5}{\partial \sqrt{a}}|_{\sqrt{a}=0} &= 0 \\ \frac{\partial^2 5}{\partial^2 \sqrt{a}}|_{\sqrt{a}=0} &= 0 \end{aligned}$$

Adding the terms;

$sum_{i=1}^{5} \frac{\partial^2 i}{\partial^2 \sqrt{a}}|_{\sqrt{a}=0} = \frac{2}{\pi} \int_{-c}^{c} \frac{e^{-(s-|u|)^2-u^2}}{u+\eta} du > 0$,so that's a local min.

3.The Third region $\eta_1 \geq c$:

$$
\begin{aligned}
P(C_1) - P(C_0) &= 2\int_{\frac{-(\eta+c)+\sqrt{(\eta-c)^2+2a}}{2}}^{s} [[2\phi(\sqrt{2}(c - |u - \frac{a}{2(u+\eta)}|) - \phi\sqrt{2}(c - |u|))] f_u(u) du \\
&+ 2\int_{c}^{\frac{-(\eta-c)+\sqrt{(\eta+c)^2+2a}}{2}} [2\phi(\sqrt{2}(c - |u - \frac{a}{2(u+\eta)}|) - 1] f_u(u) du \\
&+ 2\int_{-c}^{\frac{-(\eta+c)+\sqrt{(\eta-c)^2+2a}}{2}} [(2\phi(\sqrt{2}(c - |u|)]) - 1) f_u(u) du]
\end{aligned}
$$

There are 3 terms in $P(C_1) - P(C_0)$, we will take the first and second derivatives w.r.t $\sqrt{a}$ for term by term.

The Derivatives of the first term:

$1 = \int_{\Delta}^{c} [2\phi(\sqrt{2}(c - |u - \frac{a}{2(u+\eta)}|) - \phi\sqrt{2}(c - |u|))] f_u(u) du$
$\Delta = \frac{-(c+\eta)+\sqrt{(c-\eta)^2+2a}}{2}$ and
$\Delta_{\sqrt{a}=0} = -c \ \Delta'_{\sqrt{a}=0} = 0.$

$$
\frac{\partial^2 4}{\partial^2 \sqrt{a}}|_{\sqrt{a}=0} = \frac{2}{\pi}\int_{-c}^{c} \frac{e^{-(c-|u-\frac{a}{2(u+\eta)}|)^2}}{u+\eta} du
$$

The Derivatives of the second term:

$2 = \int_{c}^{\Delta} [2\phi(\sqrt{2}(c - |u - \frac{a}{2(u+\eta)}|) - 1] f_u(u) du$
$\Delta = \frac{-(\eta-c)+\sqrt{(\eta+c)^2+2a}}{2}$ and
$\Delta_{\sqrt{a}=0} = c \ \Delta'_{\sqrt{a}=0} = 0.$

$$
\begin{aligned}
\frac{\partial 2}{\partial\sqrt{a}}|_{\sqrt{a}=0} &= 0 \\
\frac{\partial^2 2}{\partial^2\sqrt{a}}|_{\sqrt{a}=0} &= 0
\end{aligned}
$$

The Derivatives of the third term:

$3 = \int_{-c}^{\Delta}[(2\phi(\sqrt{2}(c-|u|)]) - 1)f_u(u)du]$
$\Delta = \frac{-(\eta+c)+\sqrt{(\eta-c)^2+2a}}{2}$ and
$\Delta_{\sqrt{a}=0} = c\ \Delta'_{\sqrt{a}=0} = 0.$

$$
\begin{aligned}
\frac{\partial 3}{\partial\sqrt{a}}|_{\sqrt{a}=0} &= 0 \\
\frac{\partial^2 3}{\partial^2\sqrt{a}}|_{\sqrt{a}=0} &= 0
\end{aligned}
$$

$sum_{i=1}^{3}\frac{\partial^2 i}{\partial^2\sqrt{a}}|_{\sqrt{a}=0} = \frac{2}{\pi}\int_{-c}^{c}\frac{e^{-(s-|u|)^2-u^2}}{u+\eta}du > 0$,so that's a local min.

This completes the proof.

Proof of Lemma 4.0.3:

We first need to find Σ^{-1}, **Finding** Σ^{-1}

$$
\Sigma \times \begin{bmatrix} a & b & \cdots & b & b \\ b & a & b & \cdots & b \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b & b & \cdots & b & a \end{bmatrix} = I \text{ , then}
$$

$$
\begin{aligned}
2a + (n-1)b &= 1 \\
a + nb &= 0
\end{aligned}
$$

From the equations above, $b = \frac{-1}{n+1}$ and, $a = \frac{n}{n+1}$.

Finding $\Sigma^{-1/2}$

$$\begin{bmatrix} a & b & \cdots & b & b \\ b & a & b & \cdots & b \\ \multicolumn{5}{c}{\cdots\cdots\cdots\cdots\cdots\cdots} \\ b & b & \cdots & b & a \end{bmatrix} \times \begin{bmatrix} a & b & \cdots & b & b \\ b & a & b & \cdots & b \\ \multicolumn{5}{c}{\cdots\cdots\cdots\cdots\cdots\cdots} \\ b & b & \cdots & b & a \end{bmatrix} = \begin{bmatrix} \frac{n}{n+1} & \frac{-1}{n+1} & \cdots & \frac{-1}{n+1} & \frac{-1}{n+1} \\ \frac{-1}{n+1} & \frac{n}{n+1} & \frac{-1}{n+1} & \cdots & \frac{-1}{n+1} \\ \multicolumn{5}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\ \frac{-1}{n+1} & \frac{-1}{n+1} & \cdots & \frac{-1}{n+1} & \frac{n}{n+1} \end{bmatrix}$$ From above,

$$\begin{aligned} a^2 + (n-1)b^2 &= \frac{n}{n+1} \\ 2ab + (n-2)b^2 &= \frac{-1}{n+1} \end{aligned}$$

and working on the first equation,

$$\begin{aligned} a^2 + (n-1)b^2 \mp (n-1)^2 b^2 \mp 2ab(n-1) &= \frac{n}{n+1} \\ (a + (n-1)b)^2 + (n-1)b^2 - (n-1)^2 b^2 - 2ab(n-1) &= \frac{n}{n+1} \\ (a + (n-1)b)^2 + (n-1)\left[b^2 - (n-1)b^2 - 2ab\right] &= \frac{n}{n+1} \\ (a + (n-1)b)^2 + (n-1)\left[b^2 - (n-1)b^2 \mp (n-2)b^2 - 2ab\right] &= \frac{n}{n+1} \\ (a + (n-1)b)^2 + (n-1)\left[b^2 - (n-1)b^2 + (n-2)b^2 - (n-2)b^2 - 2ab\right] &= \frac{n}{n+1} \\ (a + (n-1)b)^2 + (n-1)\left[b^2 - nb^2 + b^2 + nb^2 - 2b^2 + \frac{1}{n+1}\right] &= \frac{n}{n+1} \\ (a + (n-1)b)^2 + \frac{n-1}{n+1} &= \frac{n}{n+1} \\ (a + (n-1)b)^2 &= \frac{1}{n+1} \\ a + (n-1)b &= \pm \frac{1}{\sqrt{n+1}}. \end{aligned}$$

Assume $a+(n-1)b=\frac{1}{\sqrt{n+1}}$, then $a=\frac{1}{\sqrt{n+1}}-(n-1)b$,

$$
\begin{aligned}
a^2+(n-1)b^2 &= \frac{n}{n+1}\\
(\frac{1}{\sqrt{n+1}}-(n-1)b)^2+(n-1)b^2 &= \frac{n}{n+1}\\
\frac{1}{n+1}+(n-1)^2b^2-\frac{2(n-1)b}{\sqrt{n+1}}+(n-1)b^2 &= \frac{n}{n+1}\\
(n-1)\left[(n-1)b^2-\frac{2b}{\sqrt{n+1}}+b^2\right] &= \frac{n-1}{n+1}\\
nb^2-\frac{2b}{\sqrt{n+1}} &= \frac{1}{n+1}\\
nb^2-\frac{2}{\sqrt{n+1}}b-\frac{1}{n+1} &= 0.
\end{aligned}
$$

Solving this equation w.r.t b yields, $b_{12}=\frac{1}{n}\left(\frac{1}{\sqrt{n+1}}\pm 1\right)$.Then
$b_1=\frac{1}{n}\left(\frac{1}{\sqrt{n+1}}+1\right)$ and $a_1=\frac{1}{\sqrt{n+1}}-\frac{n-1}{n}\left(\frac{1}{\sqrt{n+1}}+1\right)$.Moreover,
$b_2=-\frac{1}{n}\left(1-\frac{1}{\sqrt{n+1}}\right)$ and $a_2=\frac{1}{\sqrt{n+1}}+\frac{n-1}{n}\left(1-\frac{1}{\sqrt{n+1}}\right)$.
Now, assume $a+(n-1)b=\frac{-1}{\sqrt{n+1}}$ and carry out same calculation above, we derived,
$b_3=\frac{-1}{n}\left(\frac{1}{\sqrt{n+1}}+1\right)$ and $a_3=\frac{-1}{\sqrt{n+1}}+\frac{n-1}{n}\left(\frac{1}{\sqrt{n+1}}+1\right)$.Moreover,
$b_4=\frac{1}{n}\left(1-\frac{1}{\sqrt{n+1}}\right)$ and $a_4=\frac{-1}{\sqrt{n+1}}-\frac{n-1}{n}\left(1-\frac{1}{\sqrt{n+1}}\right)$.

$$
\begin{aligned}
P(D_0) &= P(|Y|\leq c^*)\\
P(D_0) &\leq P(-c^*(|a|+|b|(n-1))\leq \Sigma^{-1/2}Y^*\leq c^*(|a|+|b|(n-1))\\
P(D_0)=0.95 &\leq P\left\{|Z_i|\leq c^*(|a|+|b|(n-1)), i=1,...,n\right\}\\
0.95 &\leq [2\Phi(c^*|a|+c^*|b|(n-1))]^n\\
\frac{0.95^{1/n}+1}{n} &\leq \Phi(c^*(|a|+|b|(n-1)))\\
\Phi^{-1}(\frac{0.95^{1/n}+1}{2}) &\leq c^*(|a|+|b|(n-1))\\
c &\leq c^*(|a|+|b|(n-1))\\
\frac{c}{|a|+(n-1)|b|} &\leq c^*
\end{aligned}
$$

Then $\frac{c}{|a|+(n-1)|b|} \geq \frac{c}{2}$, since $|a| = \frac{-1}{\sqrt{n+1}} + (n-1)b$ and $|b| = \frac{1}{n}(1 + \frac{1}{\sqrt{n+1}})$.This implies that,

$$c^* \geq \frac{c}{2}.$$

8.3 GRAPHS

8.3.1 The One-Way ANOVA Model

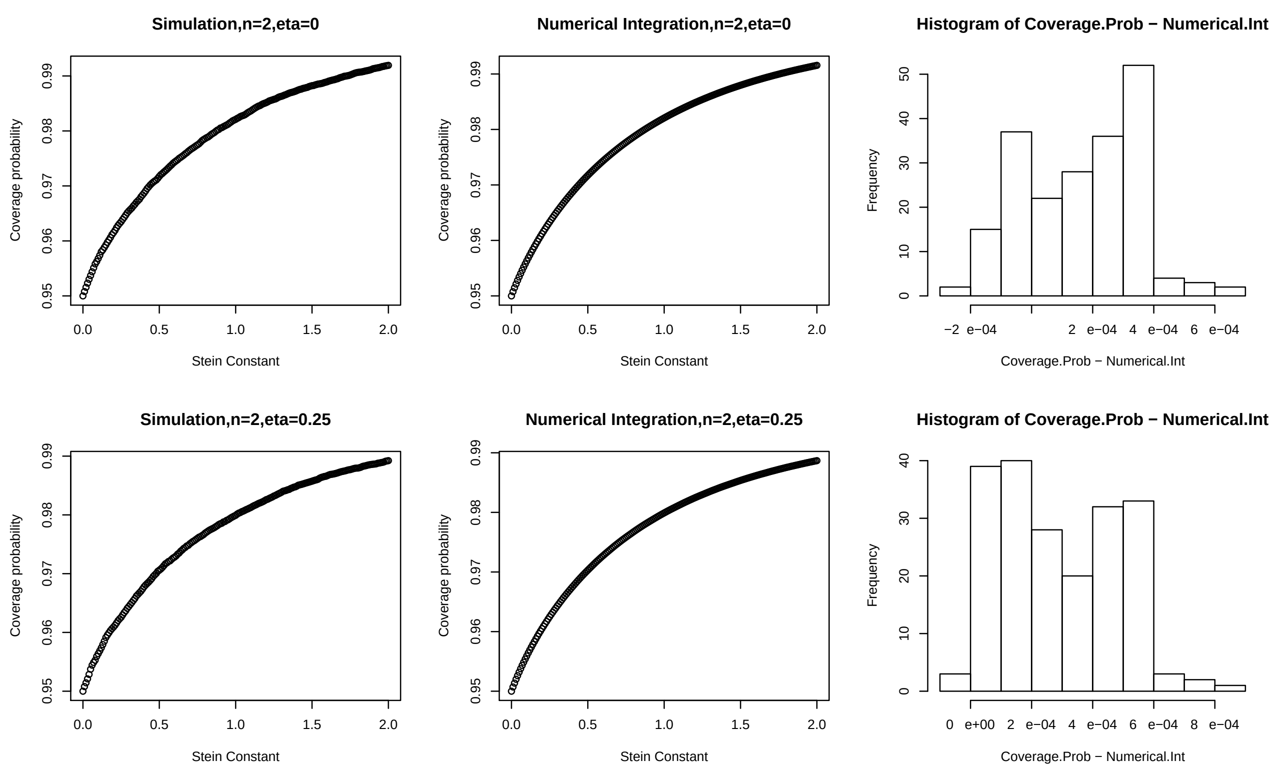

Figure 8.1: *Coverage probabilities for n=2 One-Way ANOVA model*

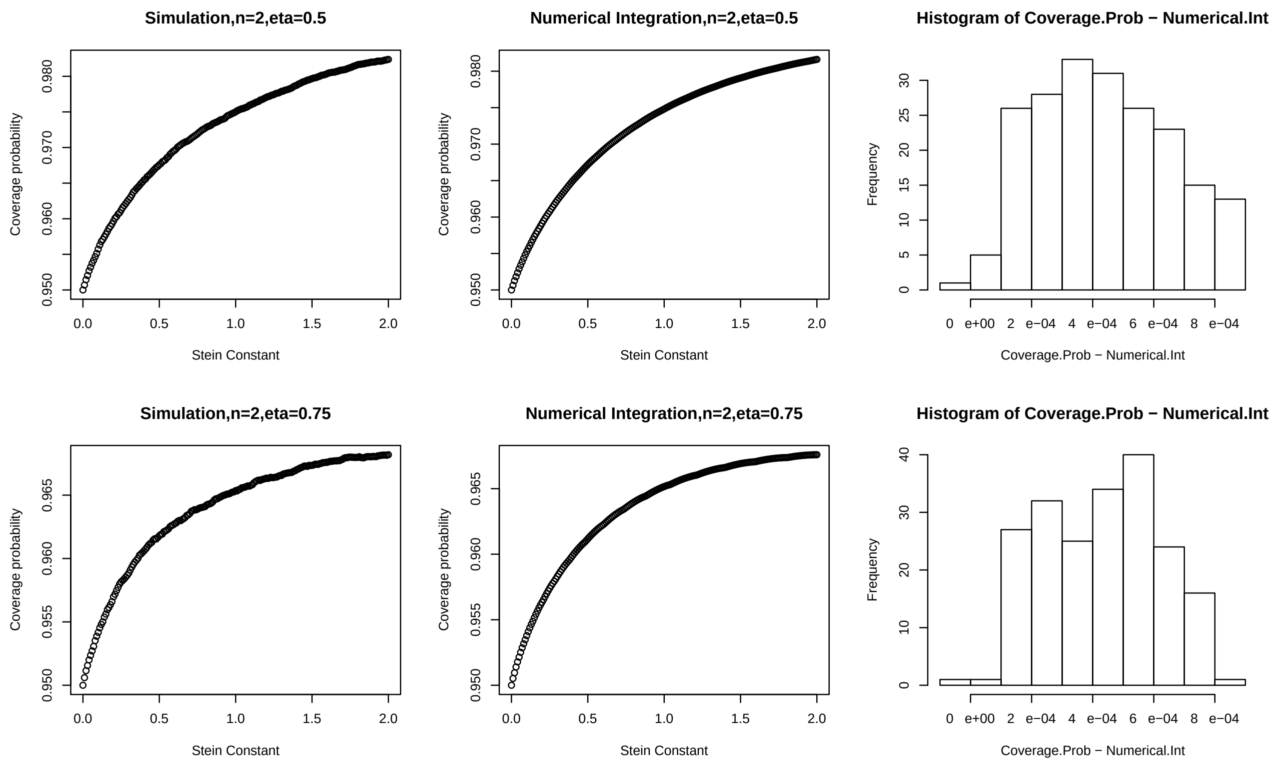

Figure 8.2: *Coverage probabilities for n=2 One-Way ANOVA model*

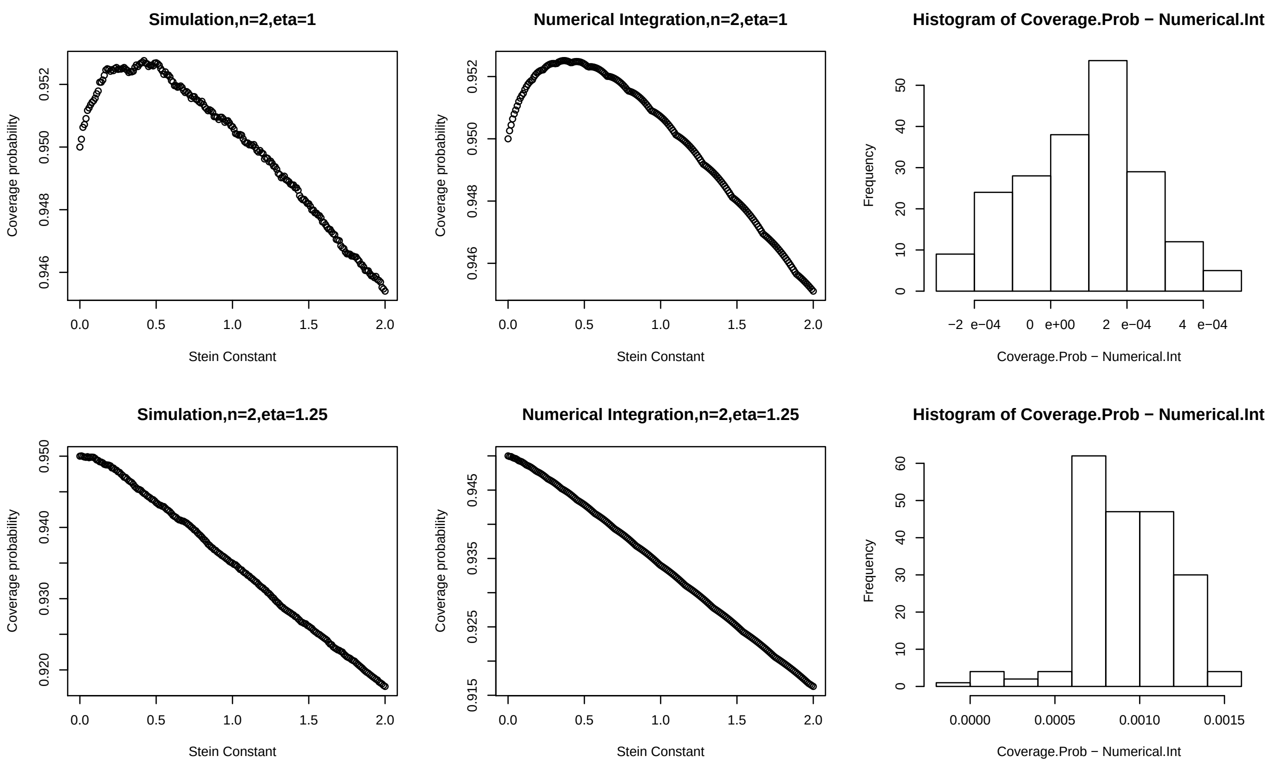

Figure 8.3: *Coverage probabilities for n=2 One-Way ANOVA model*

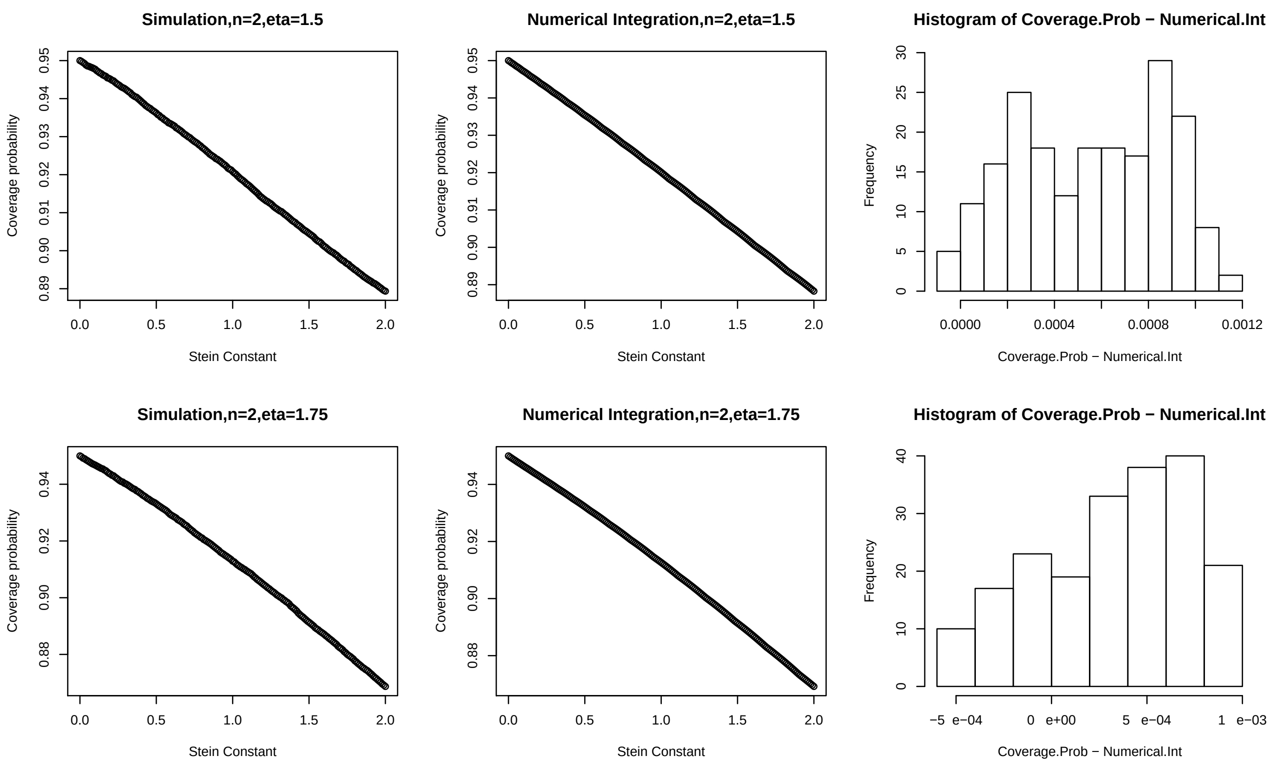

Figure 8.4: *Coverage probabilities for n=2 One-Way ANOVA model*

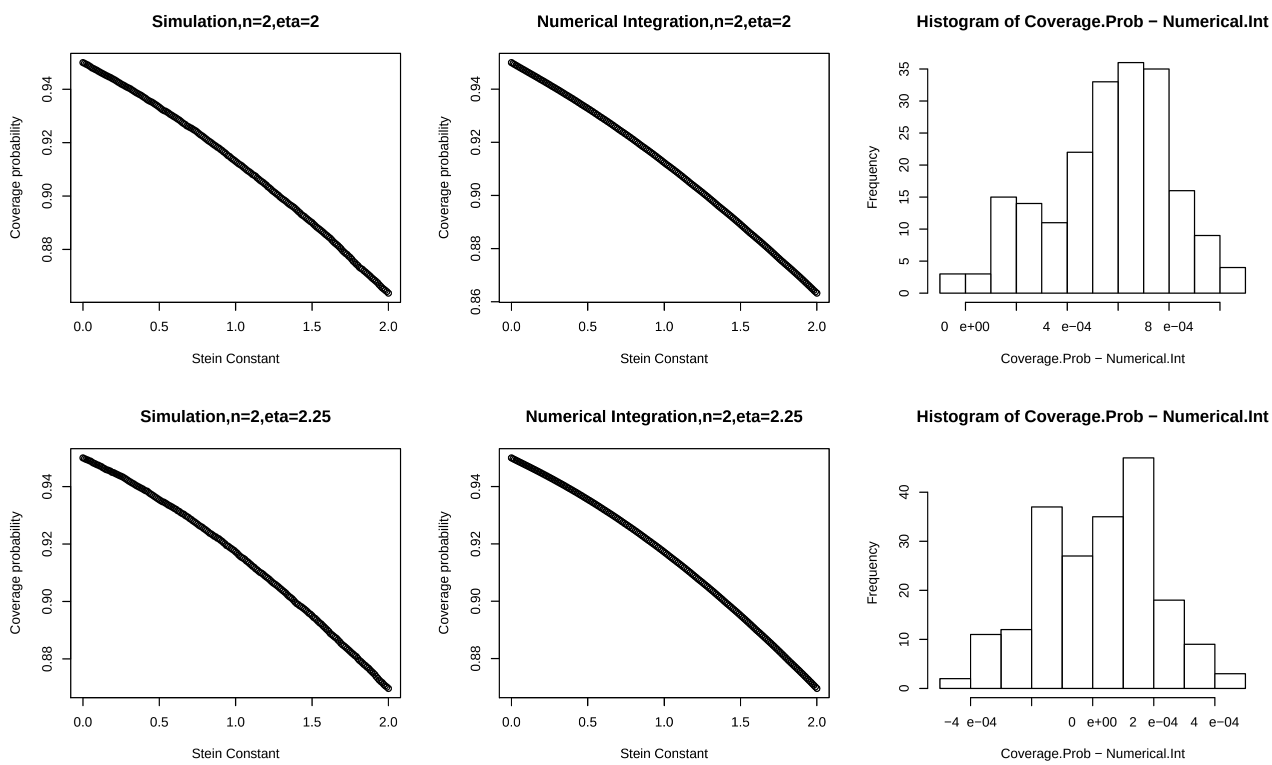

Figure 8.5: *Coverage probabilities for n=2 One-Way ANOVA model*

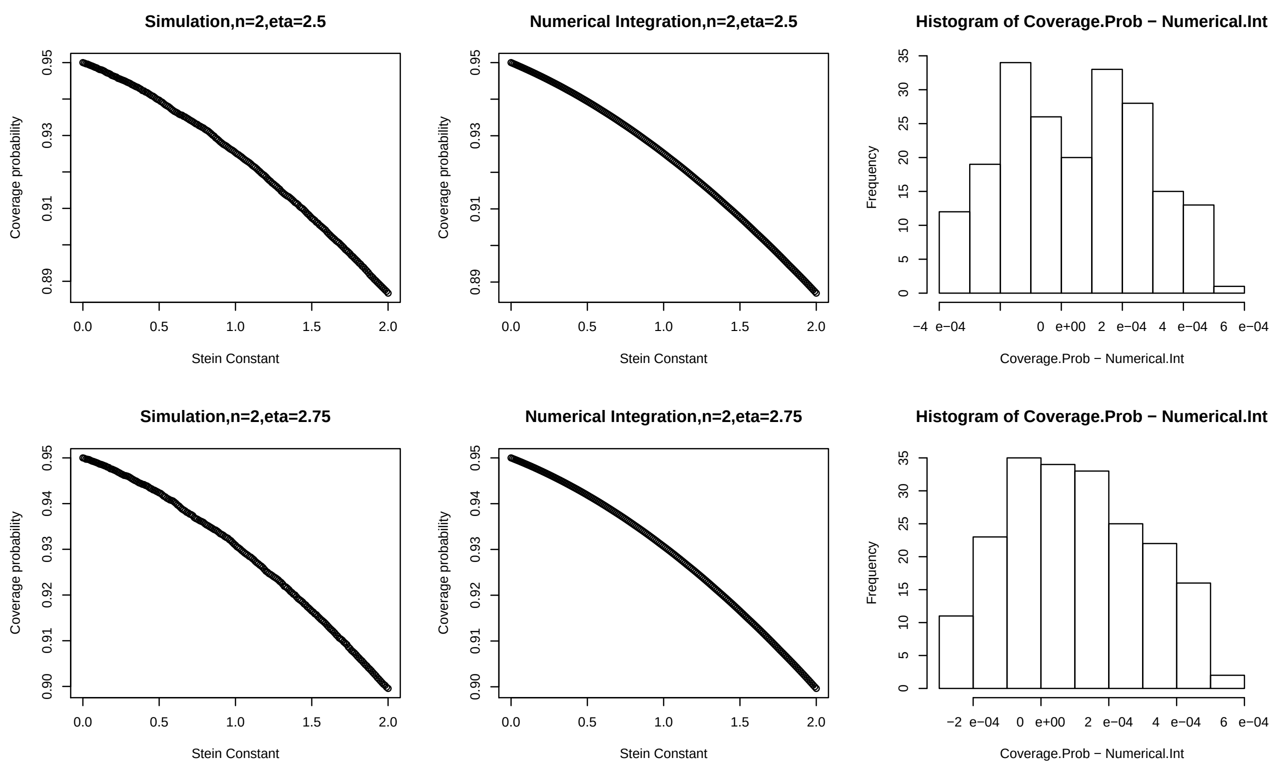

Figure 8.6: *Coverage probabilities for n=2 One-Way ANOVA model*

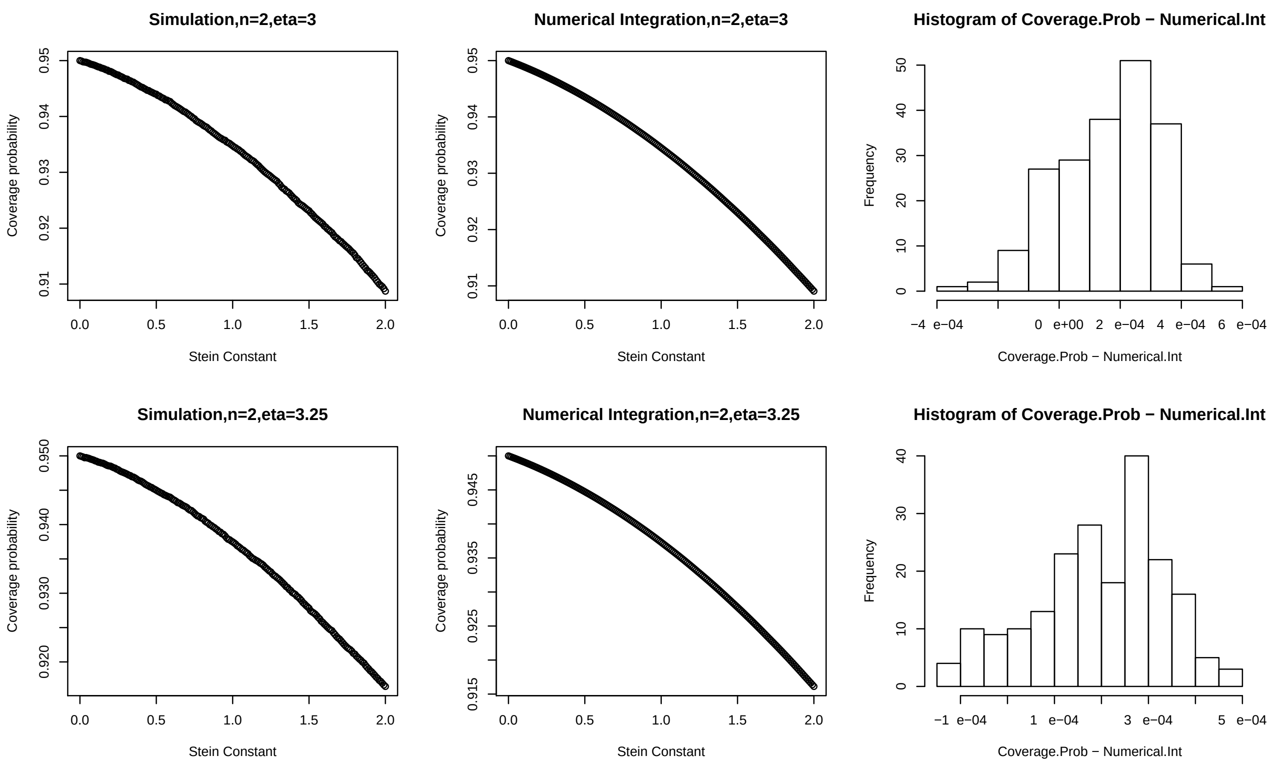

Figure 8.7: *Coverage probabilities for n=2 One-Way ANOVA model*

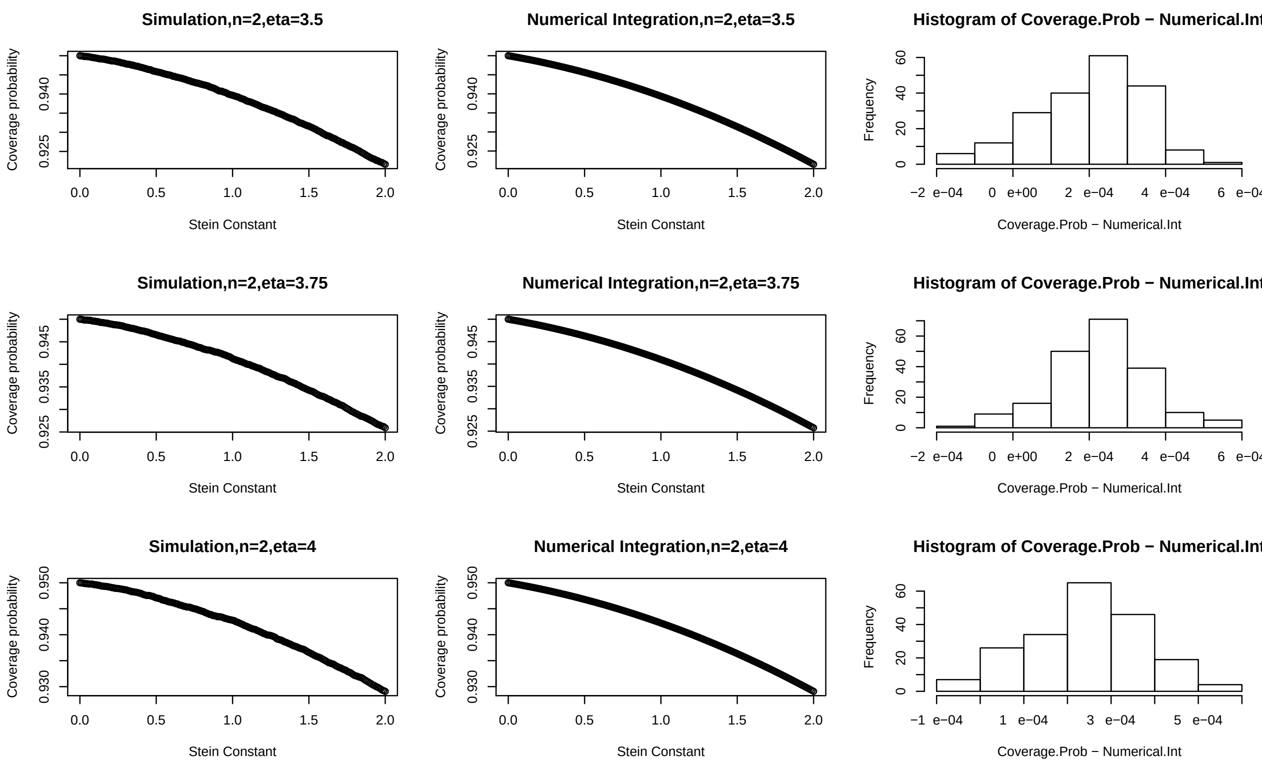

Figure 8.8: *Coverage probabilities for n=2 One-Way ANOVA model*

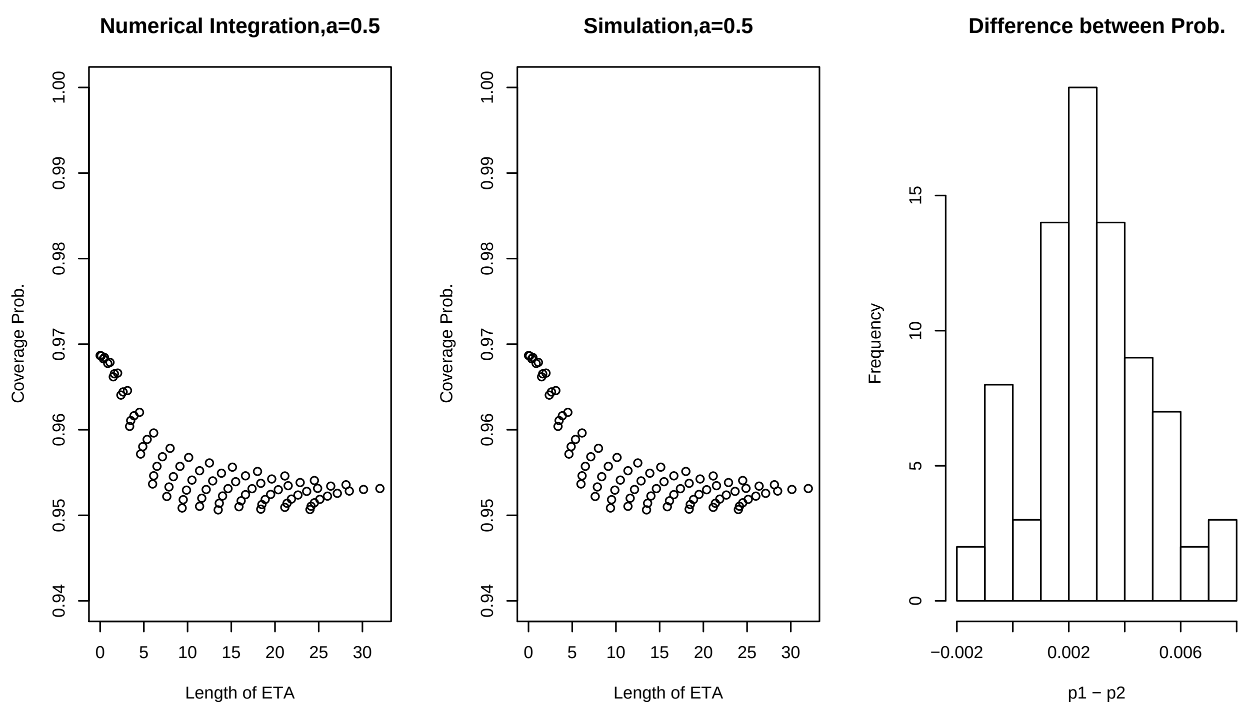

Figure 8.9: *Coverage probabilities for n=3 One-Way ANOVA model*

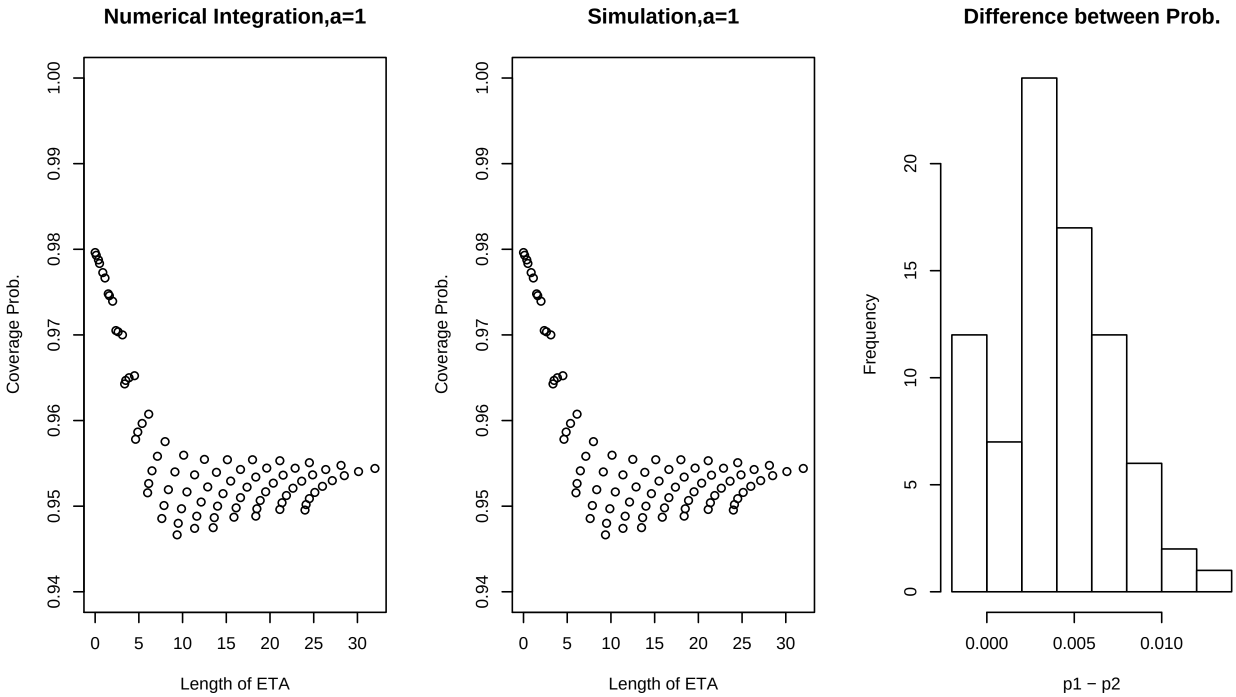

Figure 8.10: *Coverage probabilities for n=3 One-Way ANOVA model*

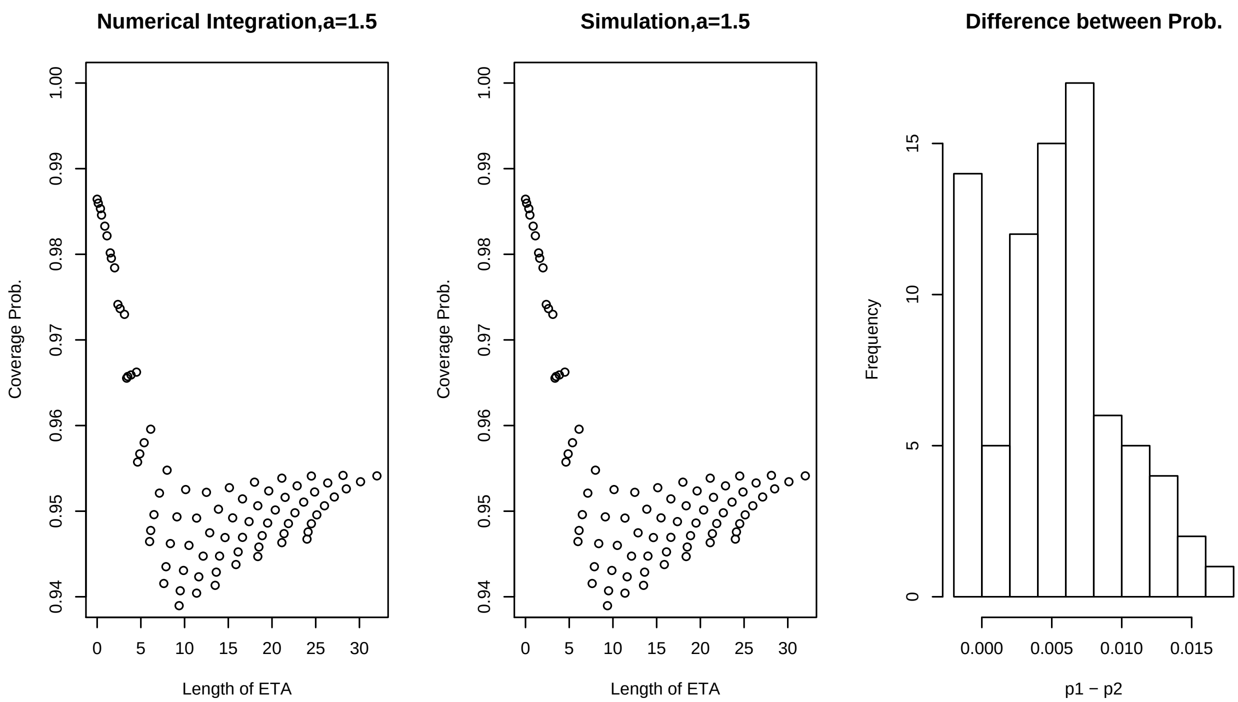

Figure 8.11: *Coverage probabilities for n=3 One-Way ANOVA model*

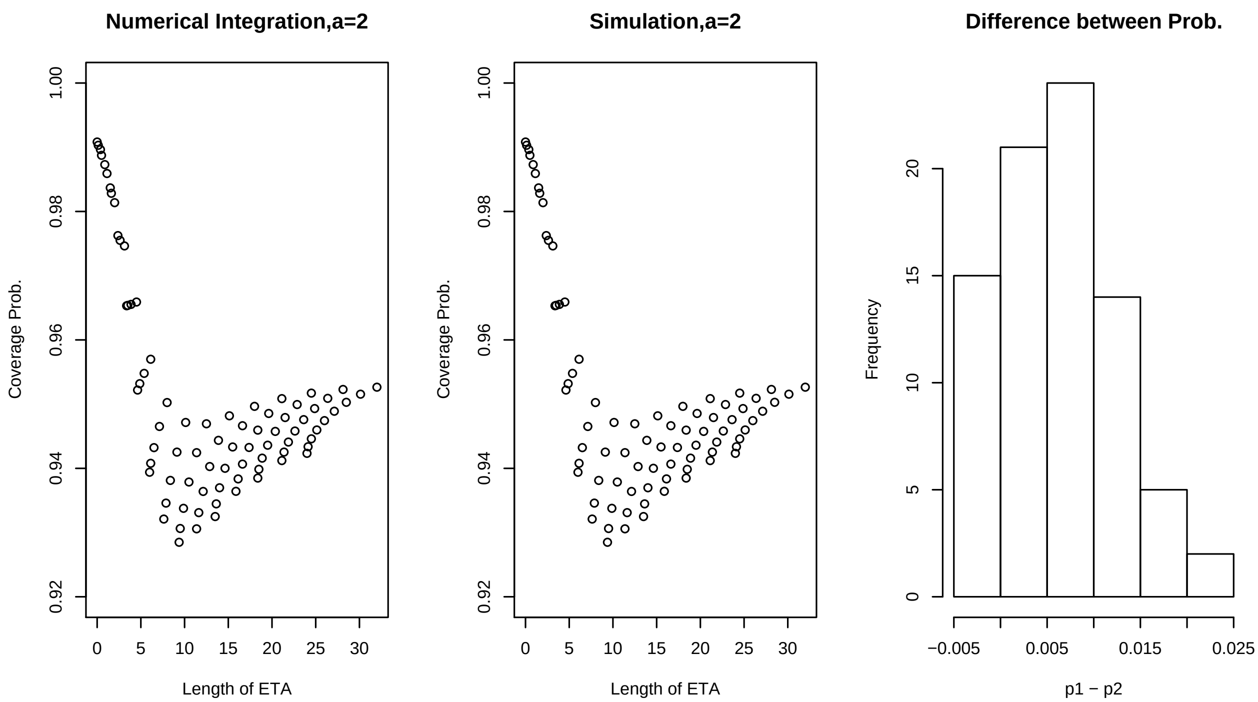

Figure 8.12: *Coverage probabilities for n=3 One-Way ANOVA model*

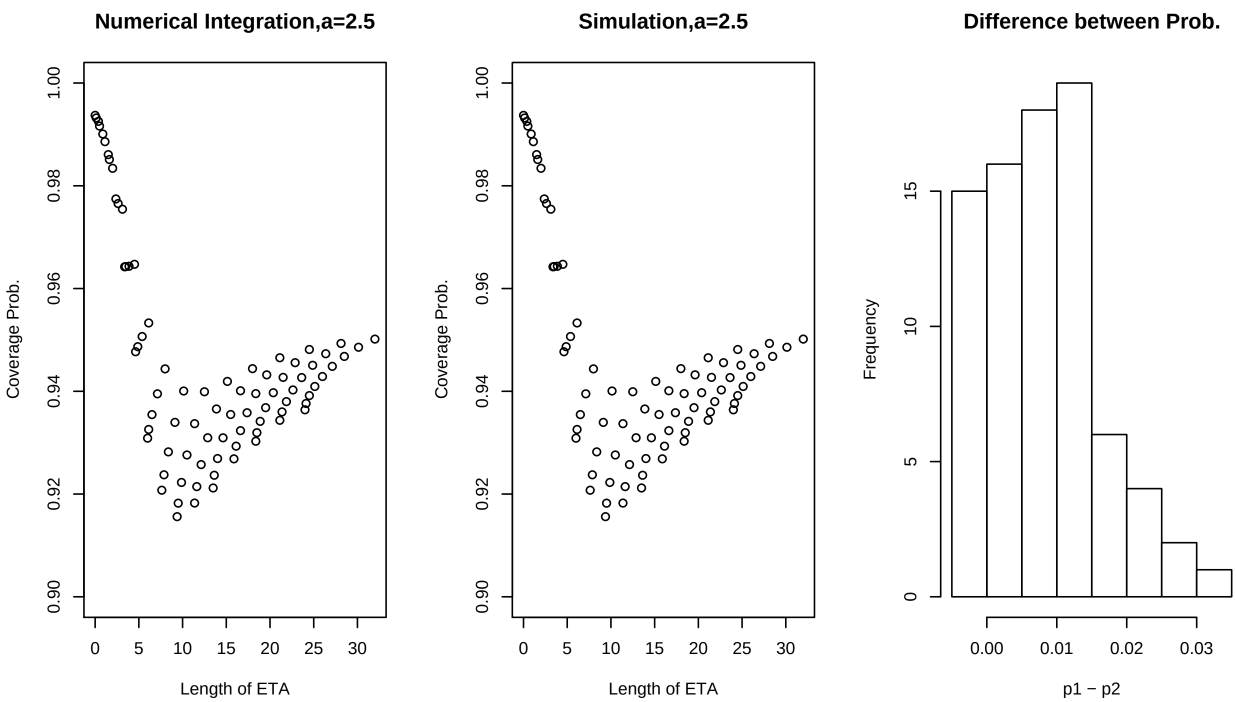

Figure 8.13: *Coverage probabilities for n=3 One-Way ANOVA model*

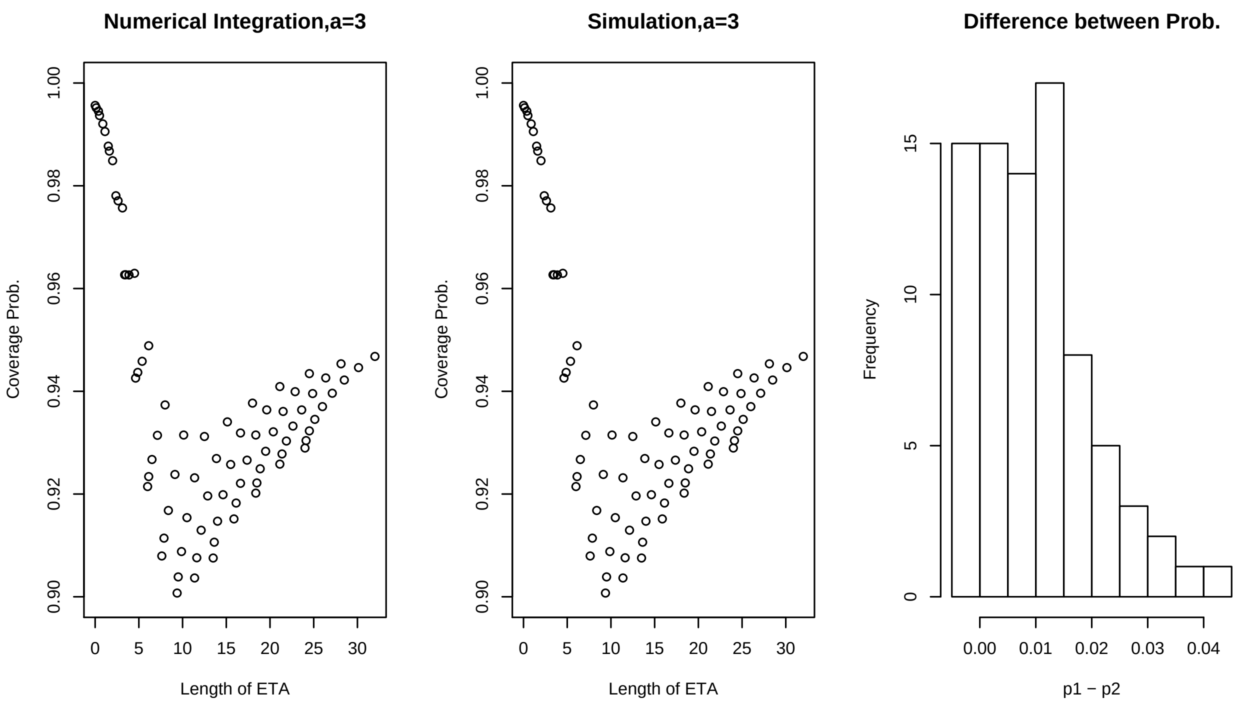

Figure 8.14: *Coverage probabilities for n=3 One-Way ANOVA model*

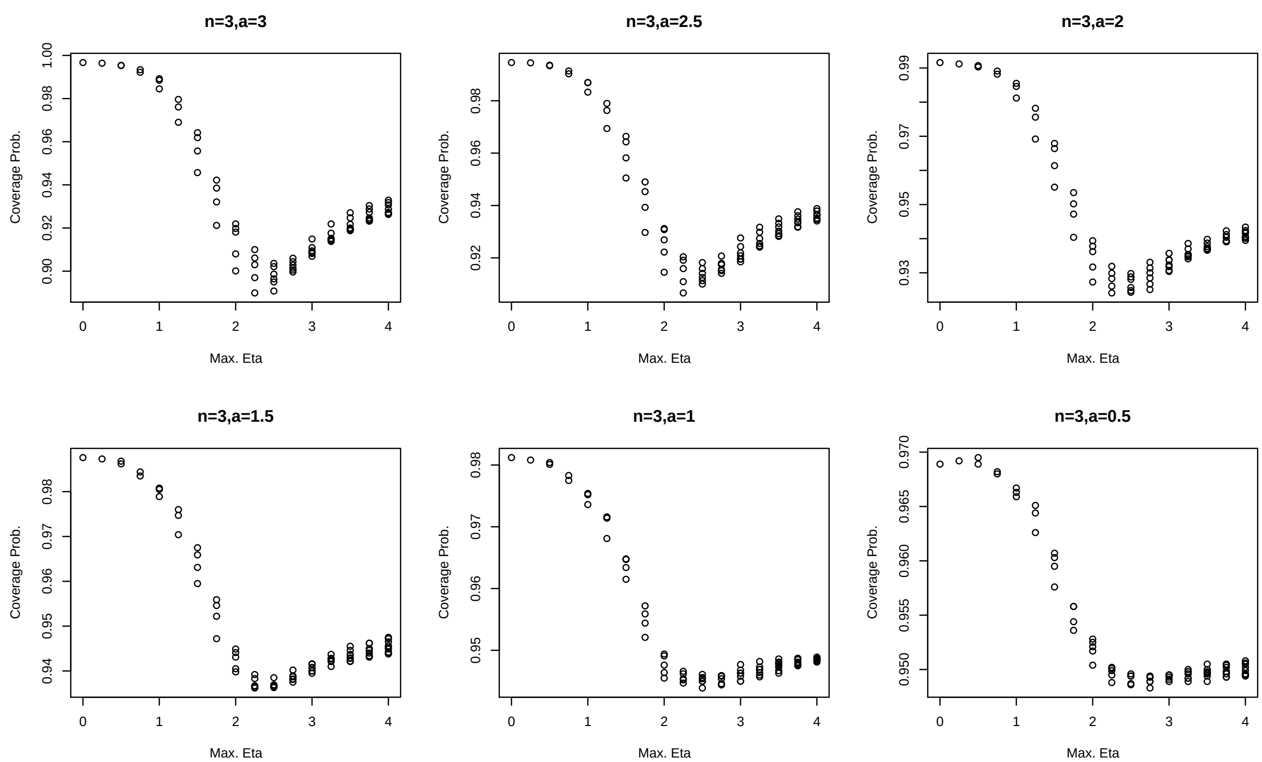

Figure 8.15: *Coverage probabilities for n=3 One-Way ANOVA model*

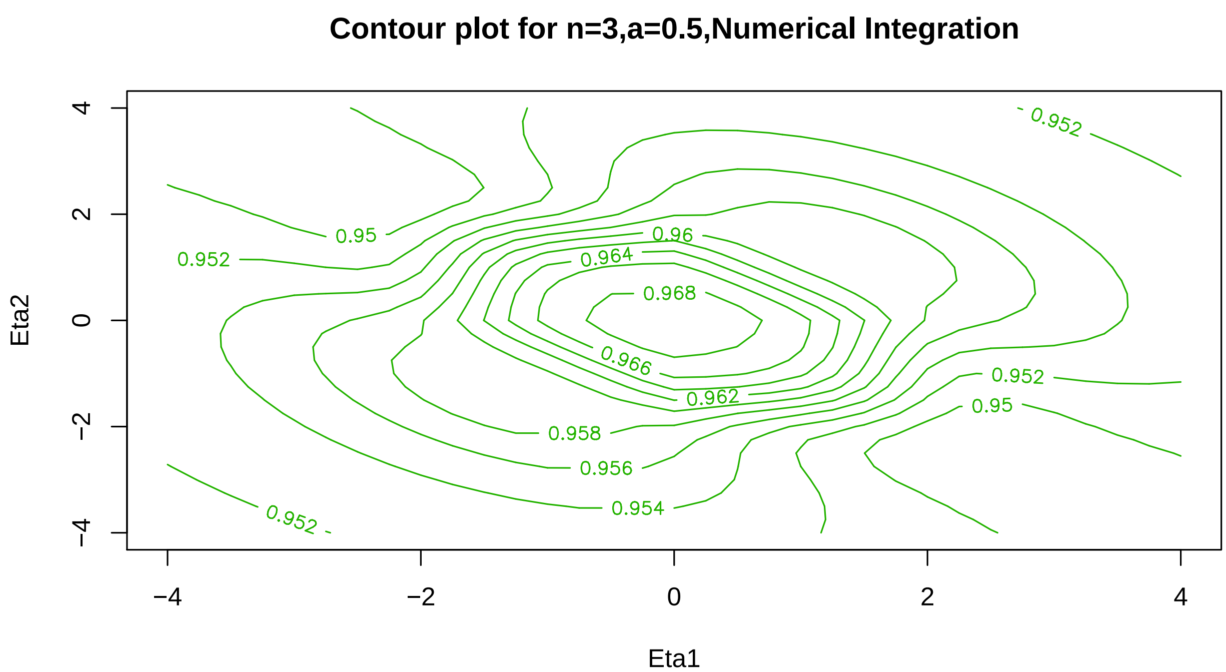

Figure 8.16: *Coverage probabilities for n=3 One-Way ANOVA model*

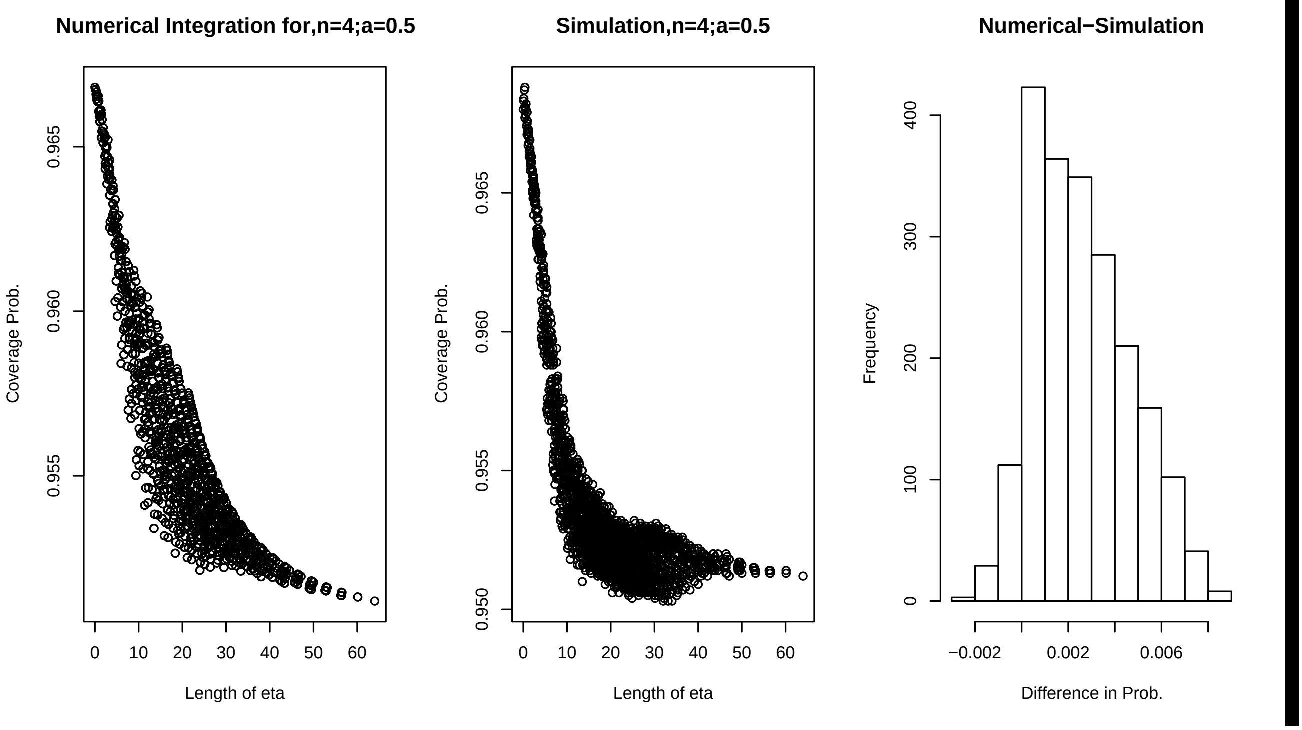

Figure 8.17: *Coverage probabilities for n=4 One-Way ANOVA model*

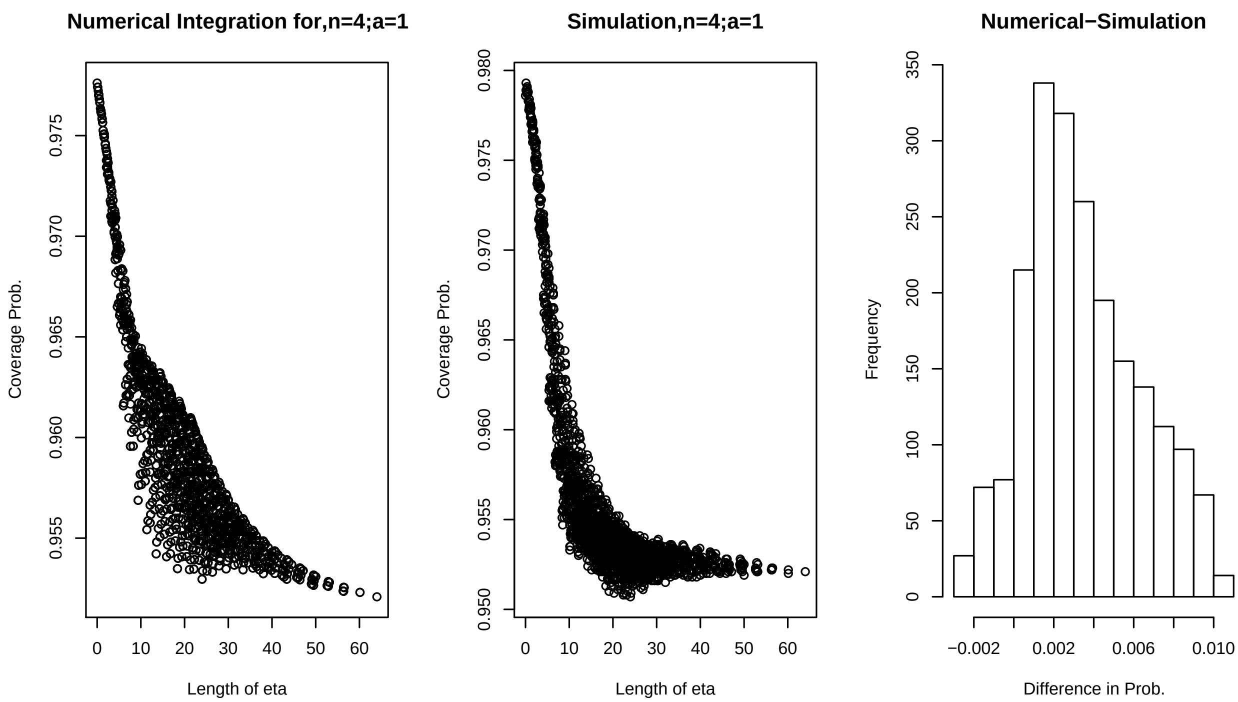

Figure 8.18: *Coverage probabilities for n=4 One-Way ANOVA model*

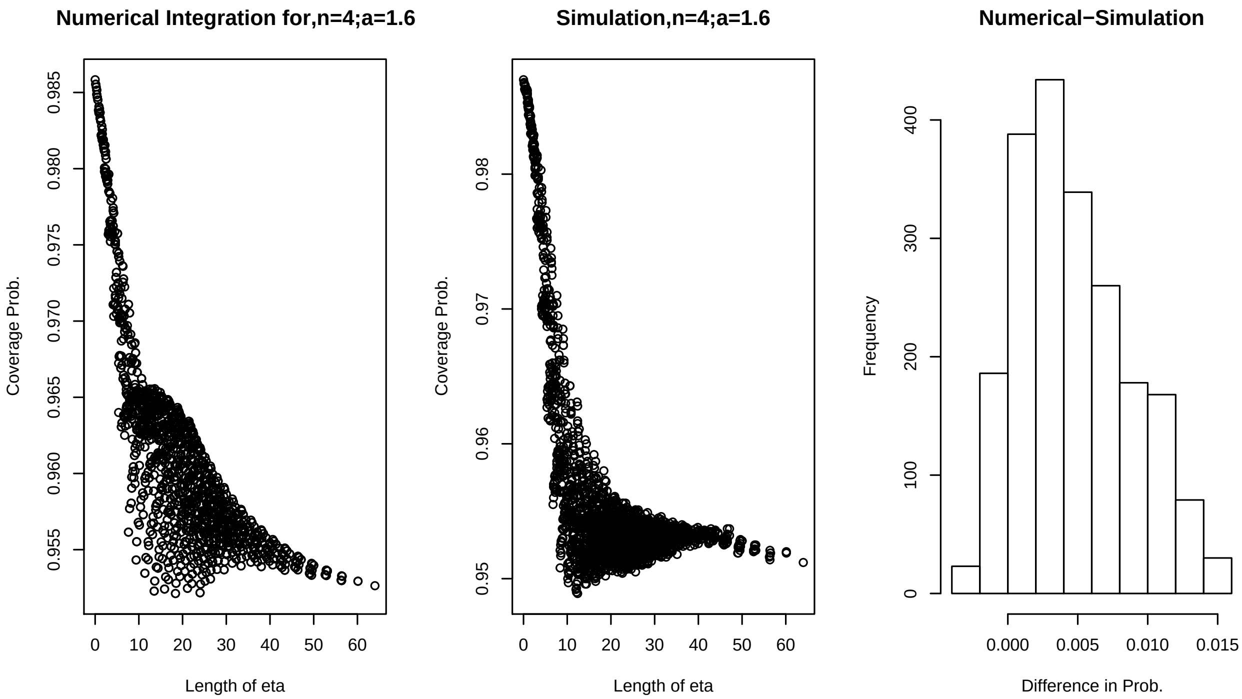

Figure 8.19: *Coverage probabilities for n=4 One-Way ANOVA model*

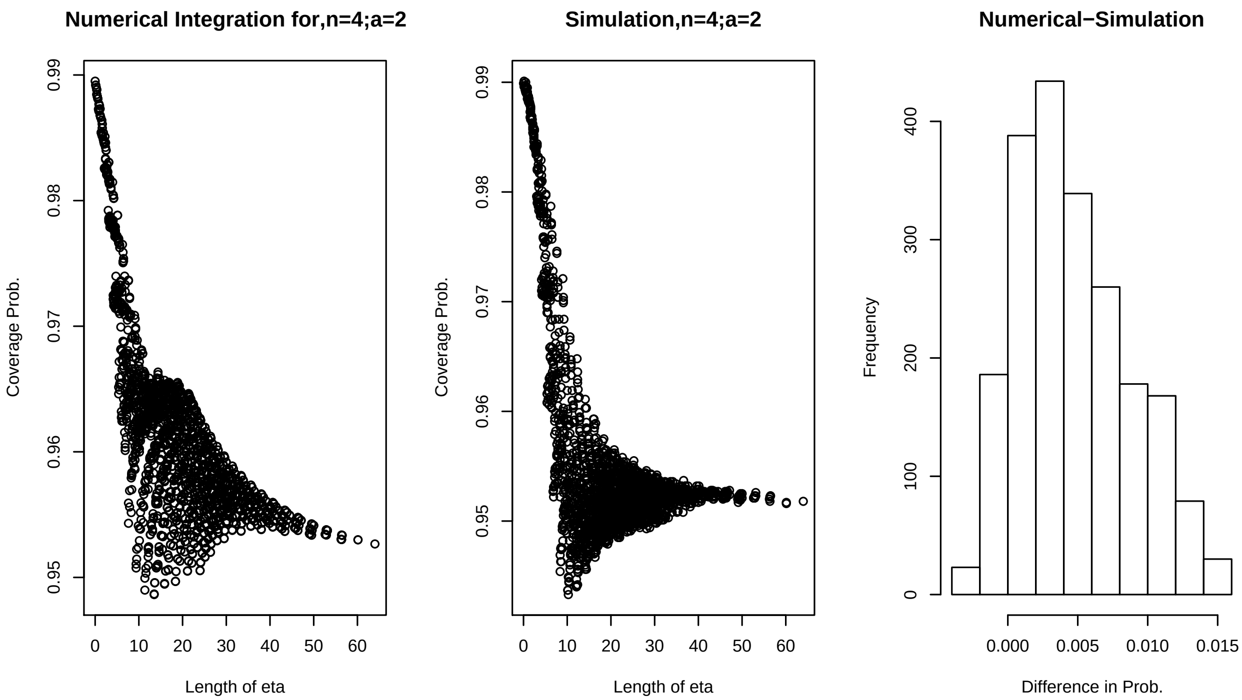

Figure 8.20: *Coverage probabilities for n=4 One-Way ANOVA model*

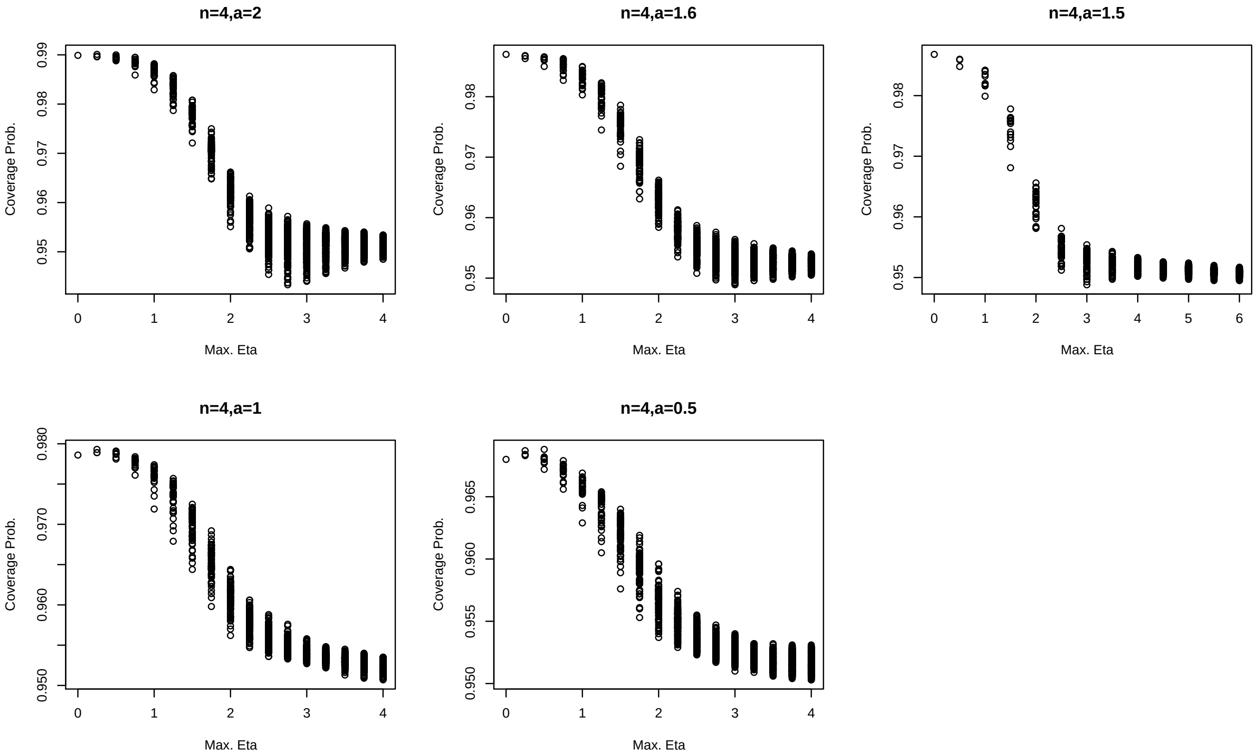

Figure 8.21: *Coverage probabilities for n=4 One-Way ANOVA model*

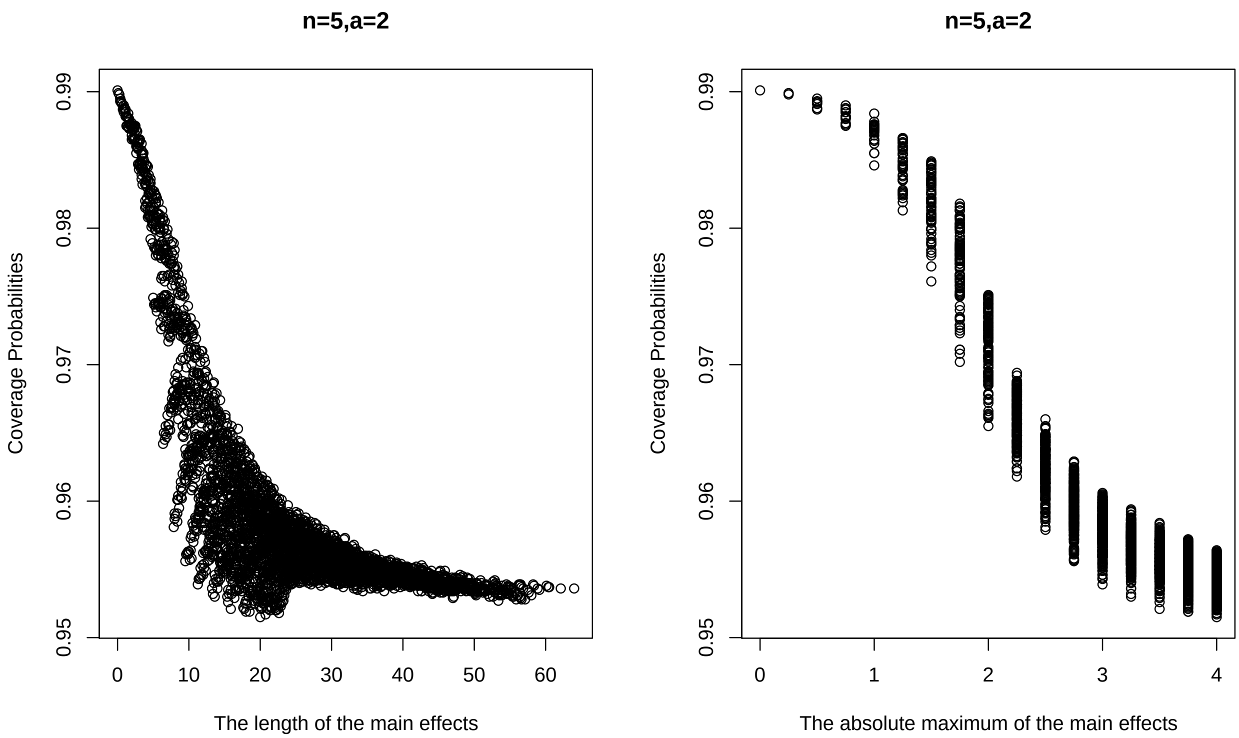

Figure 8.22: *Coverage probabilities for n=5 One-Way ANOVA model*

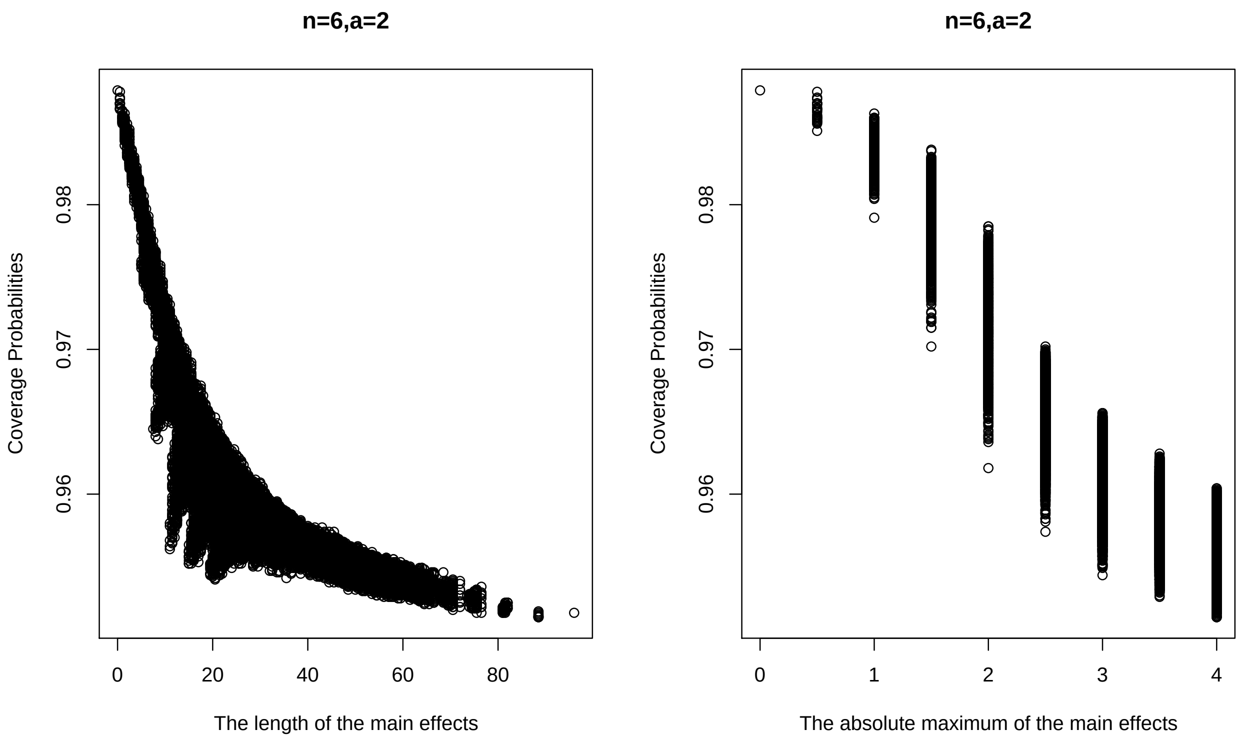

Figure 8.23: *Coverage probabilities for n=6 One-Way ANOVA model*

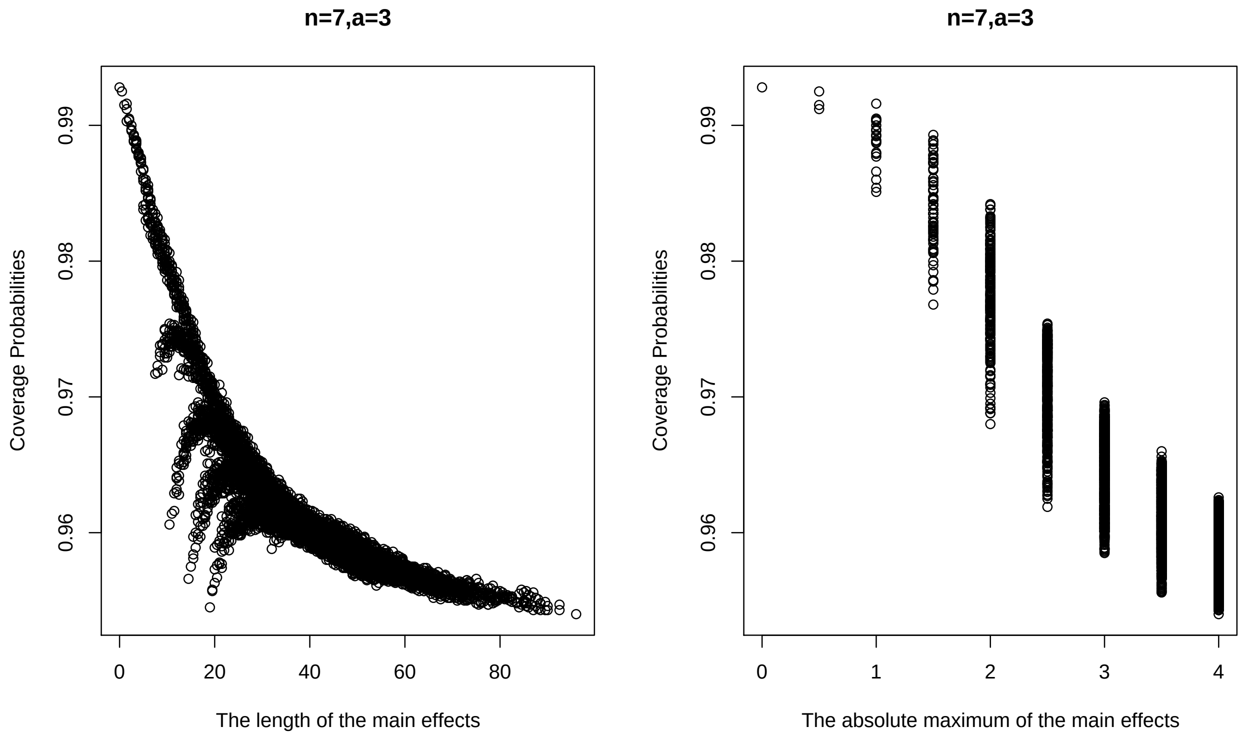

Figure 8.24: *Coverage probabilities for n=7 One-Way ANOVA model*

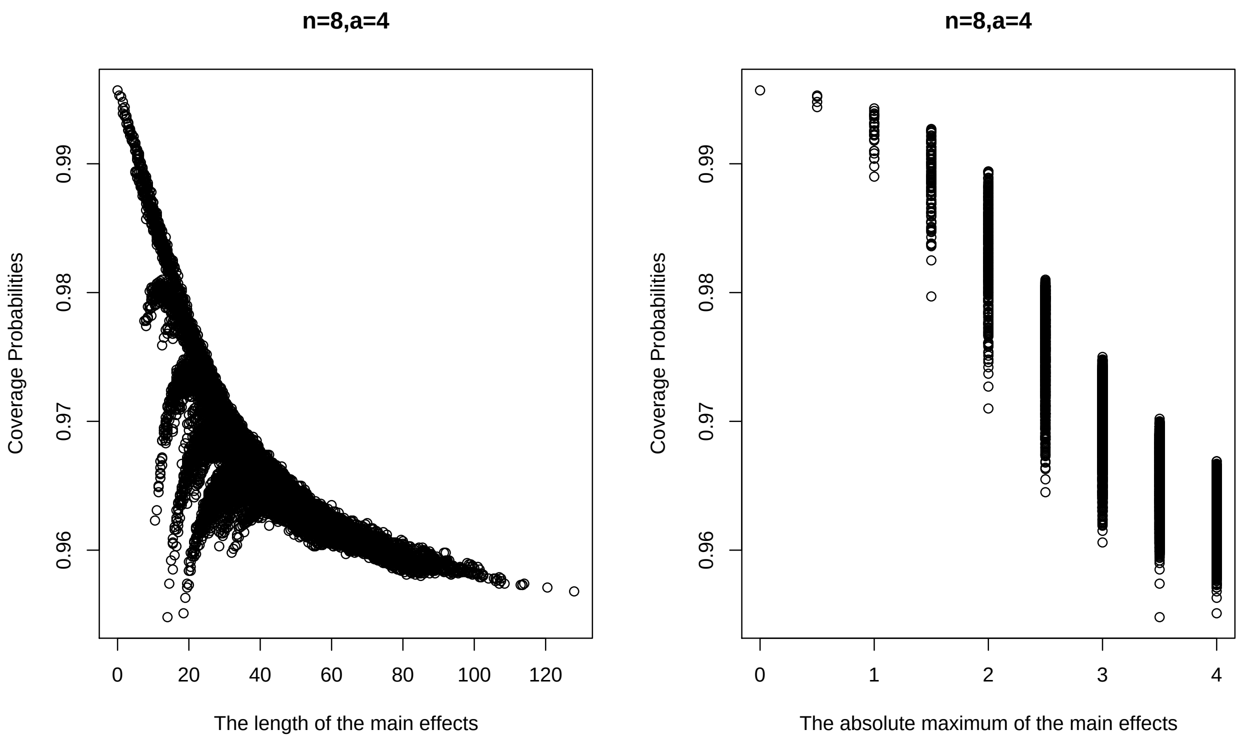

Figure 8.25: *Coverage probabilities for n=8 One-Way ANOVA model*

8.3.2 The Two-Way ANOVA Model

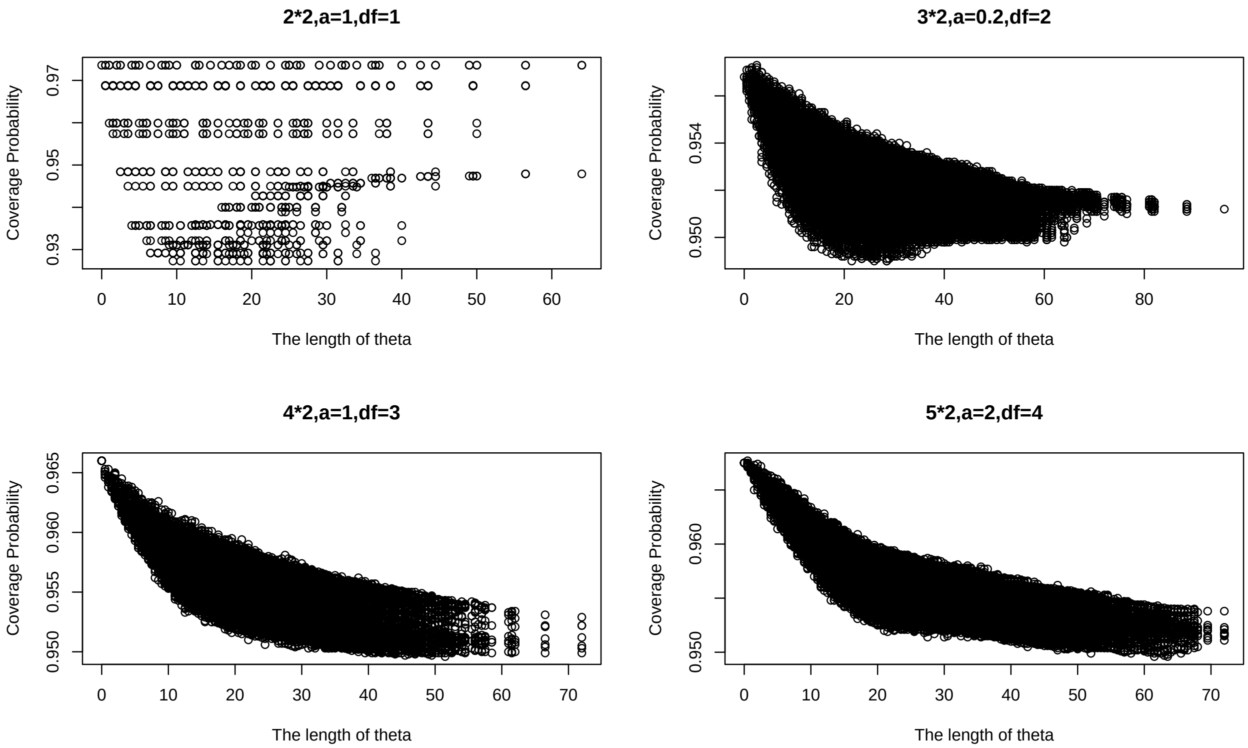

Figure 8.26: *Coverage probabilities for 2×2,3×2,4×2,5×2, Two-way ANOVA model*

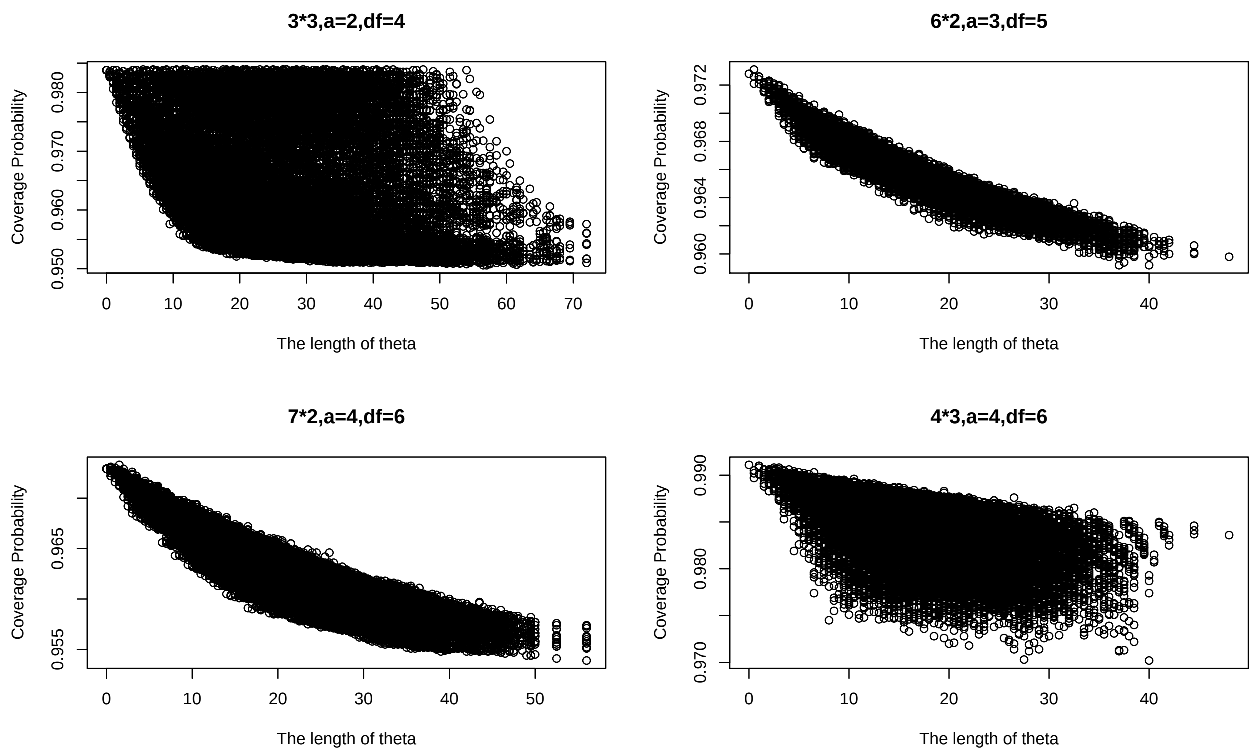

Figure 8.27: *Coverage probabilities for 5×2,6×2,7×2,4×3, Two-way ANOVA model*

BIBLIOGRAPHY

[1] Anderson, T.W (1971), *An Introduction to Multivariate Statistical Analysis, John Wiley and Sons.*

[2] Baranchik,A.J.(1964), *Multiple Regression and Estimation of the Mean of a Multivariate Normal Distribution,Technical Report 51. Department of Statistics, Stanford University.*

[3] Baranchik,A.J.(1970), A family of Minimax Estimators, of the Mean of a Multivariate Normal Distributions, Annals of Mathematical Statistics 41,642-645.

[4] Brandwein, A.C and Strawderman, W.E.(1990), *Stein Estimation:The Spherically Symmetric Case,Statistical Science,5(3),356-369.*

[5] Brown, L.D.(1966), *On the Admissibility of Invariant Estimators of One or More Location Parameters, Annals of Mathematical Statistics 37,1087-1135.*

[6] Brown, L.D.(1971), *Admissible Estimators,Recurrent Diffusion and Insoluble Boundary Value Problems, Annals of Mathematical Statistics,42,855-903.*

[7] Casella, G. and Berger, R.L (1990), *Statistical Inference, Duxbury Press.*

[8] Casella, G and Hwang, J.T.(1983),*Empirical Bayes Confidence Sets for the Mean of a Multivariate Normal Distribution,The Journal American Statistical Association, 78,688-698.*

[9] Efron, B and Morris, C.(1973), *Stein's Estimation Rule and Its Competitors, Journal of the American Statistical Association, 65,117-130.*

[10] Fabian,V.(1991), *On the Problem of Interactions in the Analysis of Variance, Journal of the American Statistical Association,Vol 86,No 414,362-367.*

[11] Gleser,Leon Jay.(1992), *Accounting for Interactions,Journal of the American Statistical Association,Letters to the Editor,September,Vol.87,No. 419,912-913.*

[12] Graybill, F.A (1977), *Theory and Application of the Linear Model, Wadsworth and Brooks.*

[13] Gruber, Marvin, H.J.(1998),*Improving Efficiency by Shrinkage, New York:Marcel Dekker Incorporation.*

[14] Hsu,Jason C.(1996),*Multiple Comparisons Theory and Methods,Chapman and Hall.*

[15] Hwang, J.T,Casella, G.(1982), *Minimax Confidence Sets for the Mean of a Multivariate Normal Distribution,Annals of Statistics,10,868-881.*

[16] Hwang, J.T, Ullah,A.(1994), *Confidence Sets Centered at the James Stein Estimators: A Surprise Concerning the Unknown Variance Case,Journal of Econometrics,60,145-156.*

[17] James,W.and Stein ,C.(1961),*Estimation With Quadratic Loss,Proceeding of the Fourth Berkeley Symposium on Mathematics and Statistics.Berkeley:University of California Press 1,361-379.*

[18] Joshi,V.M.(1967), *Inadmissibility of the Usual Confidence Sets for the Mean of a Multivariate population, Annals of Mathematical Statistics,38,1868-1875.*

[19] Judge,G.G and Bock,M.E. (1978), *The Statistical Implications of the Pretest and Stein-Rule Estimators,New York: North Holland Publishing Company.*

[20] Judge, G.G,Griffiths, W.E,Hill,R.C and Lee, T.C.(1980), *The Theory and Practice of Econometrics,New York:John Wiley and Sons.*

[21] Lindley,D.V.(1962), *Discussion of Professor Stein's paper, Confidence Sets for the Mean of a Multivariate Normal Distribution, Journal of Royal Statistical Society Series B 242,265-296.*

[22] Marshall A.W. and Olkin I.(1979), *Inequalities: Theory of Majorization and its Applications, Academic Press.*

[23] Miller, R.G.(1966), *Simultaneous Statistical Inference, McGraw Hill Book Company.*

[24] Morris, C.N.(1983), *Parametric empirical Bayes Inference: Theory and applications, Journal of the American Statistical Association,78,47-65.*

[25] Neter,Kutner,Nachtsheim,Wasserman (1996), *Applied Linear Statistical Models, The McGraw-Hill Companies Incorporation*

[26] Saleh,MD,A.K.E and Han.(1990), *Shrinkage Estimation in Regression Analysis, Estadistica,42(139),40-63.*

[27] Seber, G.A.F(1976), *Linear Regression Analysis, John Willey and Sons.*

[28] Sclove, S.L.(1976), *Improved Estimators for the Coefficients in Linear Regression, Journal of the American Statistical Association, 63,596-606.*

[29] Stein, J.(1956), *Inadmissibility of the Usual Estimator for the Mean of a Multivariate Distribution, Berkeley: University of California Press,197-206.*

[30] Stroud, A.H and Secrest, D. (1966), *Gaussion Quadrature Formulas, Prentice-Hall.*

[31] Sun, L.A.(1981) *A Bayesian Approach to the Symmetric Multiple Comparisons Problem in the Two-Way Balanced Design, Unpublished Doctoral Dissertation, Purdue University, Department of Statistics.*