Spring 2014

Comparing k Population Means with No Assumption about the Variances

Tony Yaacoub

COMPARING k POPULATION MEANS WITH NO ASSUMPTION ABOUT THE VARIANCES

by

TONY YAACOUB

(Under the Direction of Charles W. Champ)

ABSTRACT

In the analysis of most statistically designed experiments, it is common to assume equal variances along with the assumptions that the sample measurements are independent and normally distributed. Under these three assumptions, a likelihood ratio test is used to test for the difference in population means. Typically, the assumption of independence can be justified based on the sampling method used by the researcher. The likelihood ratio test is robust to the assumption of normality. However, the equality of variances is often difficult to justify. It has been found that the assumption of equal variances cannot be made even after transforming the data. Our interest is to develop a method for comparing k population means assuming the data are independent and normally distributed but without assuming equal variances. This is the Behrens-Fisher problem for $k = 2$. We propose a method that uses the exact distribution of the likelihood ratio (test) statistic. The data is used to estimate this exact distribution to obtain an estimated critical value or an estimated p-value.

Key Words: ANOVA, fixed effects one-way design, likelihood ratio test, orthogonal designs, statistically designed experiments

2009 *Mathematics Subject Classification*:

COMPARING k POPULATION MEANS WITH NO ASSUMPTION ABOUT THE VARIANCES

by

TONY YAACOUB

B.S. in Mathematics, Clayton State University, 2012

A Thesis Submitted to the Graduate Faculty of Georgia Southern University in Partial Fulfillment
of the Requirement for the Degree

MASTER OF SCIENCE

STATESBORO, GEORGIA

2014

COMPARING k POPULATION MEANS WITH NO ASSUMPTION ABOUT THE VARIANCES

by

TONY YAACOUB

Major Professor: Charles W. Champ

Committee: Broderick O. Oluyede

Hani Samawi

Electronic Version Approved:

May, 2014

DEDICATION

This thesis is dedicated to my family who have always been supportive to me. Without them, I wouldn't have been able to achieve what I have achieved so far.

ACKNOWLEDGMENTS

Firstly, I would like to thank Dr. Charles W. Champ for his continuous encouragement and support, and for the long hours put into this thesis. Secondly, I would like to thank the thesis committee, Dr. Broderick O. Oluyede and Dr. Hani Samawi, for their support and helpful advice. Lastly, I would like to thank Dr. Jimmy Dillies and Dr. Enka Lakuriqi for helping us with one of the proofs in this thesis.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

1.1 Model and Statistical Hypotheses

It is often the interest of researchers to compare two ($k = 2$) or more ($k > 2$) population means. For example, when one is comparing the means of the k levels of a one factor, fixed effects statistically designed experiment. To make these comparisons, the researcher is allowed to sample from the populations and observe the measurement Y on each individual in the samples. The measurement Y_{ij} to be taken on the jth individual from Population i is assumed to have a normal distribution with mean μ_i and variance σ_i^2, for $j = 1, \ldots, n_i$ with $n_i \geq 2$, $i = 1, \ldots, k$ with $k \geq 2$. Also, it is assumed that the Y_{ij}'s are independent. We will refer to this model as the independent normal model.

With the assumption of equal variances ($\sigma_1^2 = \ldots = \sigma_k^2 = \sigma^2$), a method known as "analysis of variance (ANOVA)" is used to analyze these data. The ANOVA test of the null hypothesis $H_0 : \mu_1 = \ldots = \mu_k$ against the alternative hypothesis $H_a :\sim (\mu_1 = \ldots = \mu_k)$ is a likelihood ratio test. In the derivation of the likelihood ratio test statistic, it is observed that

$$\sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(Y_{ij} - \overline{Y}_{..}\right)^2 = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(Y_{ij} - \overline{Y}_{i.}\right)^2 + \sum_{i=1}^{k} n_i \left(\overline{Y}_{i.} - \overline{Y}_{..}\right)^2,$$

where

$$\overline{Y}_{i.} = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij},\ \overline{Y}_{..} = \frac{1}{m} \sum_{i=1}^{k} \sum_{j=1}^{n_i} Y_{ij}, \text{ and } m = n_1 + \ldots + n_K.$$

for $i = 1, \ldots, k$. One form for the decision rule for the likelihood ratio test of size α rejects H_0 in favor of H_a if

$$V = \frac{\sum_{i=1}^{k} n_i \left(\overline{Y}_{i.} - \overline{Y}_{..}\right)^2 / (k-1)}{\sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(Y_{ij} - \overline{Y}_{i.}\right)^2 / (m-k)} \geq F_{k-1, m-k, \alpha},$$

where $F_{k-1,m-k,\alpha}$ is the $100\,(1-\alpha)$th quantile of an F-distribution with $k-1$ numerator and $m-k$ denominator degrees of freedom. We will discuss this statistical procedure in Chapter 4.

A question that arises is "how robust is this test to the model assumptions of independence, normal, and equal variances?" Various authors, as will be discussed in the next section, have shown that the likelihood ratio test is not robust to the assumption of equal variances. In Chapter 5, we address this problem by deriving the exact distribution of V under the independent normal model with not assumption about the equality of the variances. It is shown that the distribution of V under the null hypothesis of equal means depends on the unknown population variances through the parameters

$$\lambda_1^2 = \frac{\sigma_1^2}{\sigma_k^2}, \ldots, \lambda_{k-1}^2 = \frac{\sigma_{k-1}^2}{\sigma_k^2}.$$

In the case in which $k=2$, Welch (1938) developed a test based on an approximation of the distribution of the statistic

$$T = \frac{\overline{Y}_{1.} - \overline{Y}_{2.}}{\sqrt{\frac{1}{n_1(n_1-1)}\sum_{j=1}^{n_1}\left(Y_{1j}-\overline{Y}_{1.}\right)^2 + \frac{1}{n_2(n_2-1)}\sum_{j=1}^{n_2}\left(Y_{2j}-\overline{Y}_{2.}\right)^2}}.$$

He approximated the distribution of T with a noncentral t-distribution in which the degrees of freedom is a function of the ratio $\lambda^2 = \sigma_1^2/\sigma_2^2$ and the sample sizes. The approximate test was then estimated by estimating λ^2 with

$$\widehat{\lambda}^2 = \frac{\frac{1}{n_1-1}\sum_{j=1}^{n_1}\left(Y_{1j}-\overline{Y}_{1.}\right)^2}{\frac{1}{n_2-1}\sum_{j=1}^{n_2}\left(Y_{2j}-\overline{Y}_{2.}\right)^2}.$$

Champ and Hu (2010) derived the exact distribution of T whose null distribution depends only on the parameter λ^2 and developed a test that estimates the test based on the exact null distribution of T under the independent normal model. This estimate of the exact test was demonstrated by them to perform better the approximate/estimate test of Welch (1938). In Chapter 5, we extend Champ and Hu (2010) method to the case in which $k > 2$.

1.2 Literature Review

Scheffee (1959) observed that the level of significance of the test may be significantly different from α if the assumption of equal variances does not hold. Thus, the F-test can result in an increase of the Type I error when the assumption of equal variances is violated. This creates problem in biomedical experiments, where usually large samples are not used. In such experiments each data point can be vital and expensive (Ananda & Weerahandi, 1997).

To avoid such problems, Welch (1951) developed a test that works well under heteroscedasticity, especially when the sample sizes are large. His method involved approximating the distribution of the likelihood ratio test statistic with a chi-square random variable, divided by its corresponding degrees of freedom. Then, the p-value is approximated based on the approximate distribution when the null hypothesis is true. Scott and Smith (1971) developed a test based on redefining the sample standard deviations. Under the null hypothesis, their test statistic follows a Chi-Square distribution. Further, Brown and Forsythe (1974) developed another test that utilizes a new test statistic that follows a central F distribution when the null hypothesis is true. Bishop and Dudewicz (1981), then, developed a test based on a two-stage sampling procedure. To improve on these exact tests, Rice and Gains (1989) extended the argument given by Barnard (1984) for comparing two population means to obtain an exact solution to the one-way ANOVA problem with unequal variances.

Several other methods have been proposed using simulation. Krutchkoff (1988), for example, provided a simulation-based method of obtaining an approximate solution that works fairly well even with small sample sizes. There are several tests done in the literature that provide approximate tests, such as Chen and Chen, 1998; Chen, 2001; Tsui and Weerahandi, 1989; Krishnamoorthy et al., 2006; Xu and Wang, 2007a, 2007b. Krishnamoorthy et al. (2006) developed an approximate test based on parametric bootstrapping. Weerahandi (1995a) developed a generalized p-value based on a generalized F-test to solve such problems with no adverse effect on the size of the test.

Yiğit and Gokpinar (2010) showed using Monte Carlo simulations that Welch's test, Weerahandi's generalized F-test, and Krishnamoorthy et. al 's test appear to be more powerful than other tests in most cases.

1.3 Overview of Proposed Research

In this thesis, we derive a closed-form expressions for the probability density function of a linear combination of non-central chi-squares each with one degree of freedom, of a linear combination of central chi-squares, and of their ratio mutliplied by a certain constant. Furthermore, we look at the case in which variances are equal as a potential guide to our reseach. Laslty, we derive the distributions of SSE, $SSTR$, and the likelihood ratio test statistic Q under the independent normal model when the variances are not assumed equal.

CHAPTER 2

SOME USEFUL MATRIX RESULTS

2.1 Introduction

In the study of the general linear model, vector and matrix notation are indispensable tools. In this chapter, some matrix results are given that are not typically found in the literature that are useful in finding the distribution of various quadratic forms.

2.2 Quadratic Forms

For the vector $\mathbf{y}^{p\times 1}$ and the matrix $\mathbf{A}^{p\times p}$, the scalar value $\mathbf{y}^{\mathbf{T}}\mathbf{A}\mathbf{y}$ is referred to as a quadratic form. For example, the variance S^2 of a sample of measurements $Y_1, \ldots, Y_n$ can be expressed as a quadratic form. The sample variance is defined by

$$S^2 = \frac{1}{n-1} \sum\nolimits_{i=1}^{n} \left(Y_i - \overline{Y}\right)^2,$$

where $\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$. Observe that we can write

$$S^2 = \frac{1}{n-1} \begin{bmatrix} Y_1 - \overline{Y} \\ Y_2 - \overline{Y} \\ \vdots \\ Y_n - \overline{Y} \end{bmatrix}^{\mathbf{T}} \begin{bmatrix} Y_1 - \overline{Y} \\ Y_2 - \overline{Y} \\ \vdots \\ Y_n - \overline{Y} \end{bmatrix}.$$

Next we observe that

$$\begin{bmatrix} Y_1 - \overline{Y} \\ Y_2 - \overline{Y} \\ \vdots \\ Y_n - \overline{Y} \end{bmatrix} = \begin{bmatrix} \frac{n-1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & \frac{n-1}{n} & \cdots & -\frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & \frac{n-1}{n} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}.$$

Note that

$$\begin{bmatrix} \frac{n-1}{n} & -\frac{1}{n} & \dots & -\frac{1}{n} \\ -\frac{1}{n} & \frac{n-1}{n} & \dots & -\frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \dots & \frac{n-1}{n} \end{bmatrix} = \mathbf{I} - \frac{1}{n}\mathbf{J}$$

is symmetrix and idempotent. Hence we see that

$$S^2 = \frac{1}{n-1}\mathbf{Y^T}\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\mathbf{Y},$$

where $\mathbf{I}$ is the $n \times n$ identity matrix, $\mathbf{J}$ the $n \times n$ matrix of ones, and

$$\mathbf{Y} = [Y_1, \dots, Y_n]^{\mathbf{T}}.$$

It is interesting to look at the following factorization of the matrix $\mathbf{I} - \frac{1}{n}\mathbf{J}$. We can write

$$\mathbf{I} - \frac{1}{n}\mathbf{J} = \frac{n-1}{n}\begin{bmatrix} 1 & -\frac{1}{n-1} & \dots & -\frac{1}{n-1} \\ -\frac{1}{n-1} & 1 & \dots & -\frac{1}{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n-1} & -\frac{1}{n-1} & \dots & 1 \end{bmatrix}$$

The following theorem will be useful in examining further this quadratic form.

Therorem 2.1: For any positive real number ρ,

$$\begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \dots & 1 \end{bmatrix}^{n\times n} = \mathbf{VCV^T},$$

where the (i, j)th element of $\mathbf{V}$ is given by

$$v_{ij} = \begin{cases} \frac{\sqrt{j(j+1)}}{j(j+1)}, & i \le j; \\ -\frac{j\sqrt{j(j+1)}}{j(j+1)}, & i = j+1 < n; \\ 0, & \text{otherwise.} \end{cases}$$

and $\mathbf{C} = Diagonal\,(1-\rho,\ldots,1-\rho,1+(n-1)\,\rho)$. Further $\mathbf{V}$ is a normalized orthogonal matrix.

Proof of Theorem 2.1: The proof of this theorem can be found in Champ and Rigdon (2007).■

It follows from Theorem 2.1 that

$$\mathbf{I}-\frac{1}{n}\mathbf{J}=\frac{n-1}{n}\mathbf{V}\mathbf{C}\mathbf{V}^{\mathbf{T}},$$

where

$$\begin{aligned}\mathbf{C} &= Diagonal\left(\frac{n}{n-1},\ldots,\frac{n}{n-1},0\right)\\ &= \frac{n}{n-1}Diagonal\,(1,\ldots,1,0)\\ &= \frac{n}{n-1}\mathbf{H}\end{aligned}$$

with $\mathbf{H} = Diagonal\,(1,\ldots,1,0)$. Thus,

$$\mathbf{I}-\frac{1}{n}\mathbf{J}=\frac{n-1}{n}\mathbf{V}\left(\frac{n}{n-1}\mathbf{H}\right)\mathbf{V}^{\mathbf{T}}=\mathbf{V}\mathbf{H}\mathbf{V}^{\mathbf{T}}.$$

Further, observe that $\mathbf{H}$ is symmetric and idempotent. Thus, we can write

$$S^2=\frac{1}{n-1}\mathbf{Y}^{\mathbf{T}}\left(\mathbf{V}\mathbf{H}\mathbf{V}^{\mathbf{T}}\right)\mathbf{Y}=\frac{1}{n-1}\left[\mathbf{H}\mathbf{V}^{\mathbf{T}}\mathbf{Y}\right]^{\mathbf{T}}\left[\mathbf{H}\mathbf{V}^{\mathbf{T}}\mathbf{Y}\right].$$

Theorem 2.2: If $n_1,\ldots,n_k$ are integers each greater than or equal to 1 with at least one greater than or equal to 2 and $m = n_1+\ldots+n_k$, then the eigenvalues of the matrix

$$\mathbf{I}-\frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}$$

are 0 and 1 with 1 having multiplicity $k-1$, where $\mathbf{I}$ is a $k\times k$ identity matrix, $\mathbf{N} = Diagonal\,(n_1,\ldots,n_k)$, and $\mathbf{J}$ is a $k\times k$ matrix of ones. Further, the rank of this matrix is $k-1$.

Proof of Theorem 2.2: Let λ be an eigenvalue of the matrix $\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}$. It follows that the characteristic polynomial $g(\lambda)$ can be expressed as

$$g(\lambda) = \left| \mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2} - \lambda \mathbf{I} \right| = (-1)^k \lambda (1-\lambda)^{k-1}.$$

Thus, the eigenvalues are 0 and 1 with 1 having multiplicity $k-1$. This also shows that the rank of $\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}$ is $k-1$. ■

Theorem 2.3: Suppose $n_1, \ldots, n_k$ are integers each greater than or equal to 1 with at least one greater than or equal to 2 and $m = n_1 + \ldots + n_k$, then we have

$$\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2} = \mathbf{V}\mathbf{H}\mathbf{V}^{\mathbf{T}},$$

where $\mathbf{I}$ is a $k \times k$ diagonal matrix, $\mathbf{N} = \mathit{Diagonal}(n_1, \ldots, n_k)$, and $\mathbf{J}$ is a $k \times k$ matrix of ones, the orthonormal eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are the columns of $\mathbf{V}$ with the first $k-1$ vectors associated with the $k-1$ eigenvalues of 1 of the matrix $\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}$ and the eigenvector $\mathbf{v}_k$ associated with the eigenvalue 0.

Proof of Theorem 2.3: The proof of this theorem can be found in Champ and Jones-Farmer (2007).■

Theorem 2.4: If $\mathbf{\Sigma} = \mathit{Diagonal}(\sigma_1^2, \ldots, \sigma_k^2)$, $\mathbf{I}$ is a $k \times k$ identity matrix, $\mathbf{N} = \mathit{Diagonal}(n_1, \ldots, n_k)$, and $\mathbf{J}$ is a $k \times k$ matrix of ones with $n_1, \ldots, n_k$ positive integers with at least one greater than or equal to 2, $m = n_1 + \ldots + n_k$, and $\sigma_i^2 > 0$ for $i = 1, \ldots, k-1$, then the matrix

$$\mathbf{\Sigma}^{1/2}\left(\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}\right)\mathbf{\Sigma}^{1/2} = \mathbf{V}\mathbf{C}\mathbf{V}^{\mathbf{T}},$$

is of rank $k-1$, where the diagonal elements of $\mathbf{C} = \mathit{Diagonal}(\xi_1, \ldots, \xi_{k-1}, 0)$ are the eigenvalues with $\xi_i > 0$ of the matrix and $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_k]^{\mathbf{T}}$ is the matrix in which the column vectors are the associated orthonormal eigenvectors. Further, we have

$$\mathbf{C}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mathbf{1} = \mathbf{0},$$

where $\mathbf{0}$ and $\mathbf{1}$ are $k \times 1$ vectors of zeros and ones, respectively.

Proof of Theorem 2.4: Since the matrix $\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}$ is a real symmetric matrix of rank $k-1$ and $\mathbf{\Sigma}$ is of rank k, then $\mathbf{\Sigma}^{1/2}\left(\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}\right)\mathbf{\Sigma}^{1/2}$ has rank $k-1$. Thus, the matrix $\mathbf{\Sigma}^{1/2}\left(\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}\right)\mathbf{\Sigma}^{1/2}$ has $k-1$ positive real eigenvalues $\xi_1, \ldots, \xi_{k-1}$ and one zero eigenvalue. It follows that we can express $\mathbf{\Sigma}^{1/2}\left(\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}\right)\mathbf{\Sigma}^{1/2}$ as $\mathbf{V}\mathbf{C}\mathbf{V}^{\mathbf{T}}$. It is convenient in what follows to define $\mathbf{C}^{-1/2} = \mathit{Diagonal}\left(\xi_1^{-1/2}, \ldots, \xi_{k-1}^{-1/2}, 0\right)$. Note that

$$\mathbf{C}^{-1/2}\mathbf{C}^{1/2} = \mathit{Diagonal}\,(1, \ldots, 1, 0) = \mathbf{H} \text{ and } \mathbf{H}\mathbf{C}^{1/2} = \mathbf{C}^{1/2}.$$

Using the results in Dillies and Lakuriqi (2014), we can write

$$\begin{aligned}
\mathbf{C}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mathbf{1} &= \mathbf{H}\mathbf{C}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mathbf{1} \\
&= \left(\mathbf{C}^{-1/2}\mathbf{V}^{\mathbf{T}}\mathbf{V}\mathbf{C}^{1/2}\right)\left(\mathbf{C}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mathbf{1}\right) \\
&= \mathbf{C}^{-1/2}\mathbf{V}^{\mathbf{T}}\left(\mathbf{V}\mathbf{C}\mathbf{V}^{\mathbf{T}}\mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mathbf{1}\right) \\
&= \mathbf{C}^{-1/2}\mathbf{V}^{\mathbf{T}}\left(\mathbf{\Sigma}^{1/2}\left(\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}\right)\mathbf{\Sigma}^{1/2}\right)\mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mathbf{1} \\
&= \mathbf{C}^{-1/2}\mathbf{V}^{\mathbf{T}}\mathbf{\Sigma}^{1/2}\mathbf{N}^{-1/2}\mathbf{N}^{1/2}\left(\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}\right)\mathbf{\Sigma}^{1/2}\mathbf{\Sigma}^{-1/2}\mathbf{N}^{1/2}\mathbf{1} \\
&= \mathbf{C}^{-1/2}\mathbf{V}^{\mathbf{T}}\mathbf{\Sigma}^{1/2}\left(\mathbf{N}\mathbf{1} - \frac{1}{m}\mathbf{N}\mathbf{J}\mathbf{N}\mathbf{1}\right) \\
&= \mathbf{C}^{-1/2}\mathbf{V}^{\mathbf{T}}\mathbf{\Sigma}^{1/2}\left(\mathbf{N}\mathbf{1} - \mathbf{N}\mathbf{1}\right) \\
&= \mathbf{0}.\blacksquare
\end{aligned}$$

As discussed in Champ and Jones-Farmer (2007), it is convenient to define the $p \times p$ matrix $\mathbf{B}$ by

$$\mathbf{B} = \frac{1}{d\left(\delta_1 + d\right)}\left(\delta + d\mathbf{e}_1\right)\left(\delta + d\mathbf{e}_1\right)^{\mathbf{T}} - \mathbf{I},$$

where $\delta = \mathbf{P}_0^{-1}\left(\mu - \mu_0\right)$, δ_1 is the first component of the vector δ, $d^2 = \delta^{\mathbf{T}}\delta = \left(\mu - \mu_0\right)^{\mathbf{T}}\mathbf{\Sigma}_0^{-1}\left(\mu - \mu_0\right)$, and $\mathbf{e}_1$ is a $p \times 1$ vector with first coordinate one and the remaining coordinates zero. It is easy to show that the matrix $\mathbf{B}$ is an orthogonal matrix that transforms δ into

$$\mathbf{B}\delta = [d, 0, \ldots, 0]^{\mathbf{T}} = d\mathbf{e}_1.$$

CHAPTER 3

SOME DISTRIBUTIONAL RESULTS

3.1 Introduction

In the analysis of various designed experiments, the distribution of linear combinations of central and noncentral chi square random variables are of interest as well as the ratio of two such linear combinations. We will examine the distributions of the following linear combination of independent central chi square random variables, the linear combination of independent noncentral chi square random variables each with one degree of freedom, and the ratio of these linear combinations. The linear combinations and ratio are expressed symbolically as

$$(1)\ \sum_{i=1}^{s} a_i \chi^2_{\nu_i};\ (2)\ \sum_{i=1}^{t} c_i \chi^2_{1,\tau_i^2};\ \text{and}\ (3)\ \sum_{i=1}^{t} c_i \chi^2_{1,\tau_i^2} \Big/ \left(\sum_{i=1}^{s} a_i \chi^2_{\nu_i} \right).$$

In the next section, we derive a closed form expression for (1). This is followed by two section in which we give closed form expressions for (2) and (3).

3.2 Joint Distribution of the Sample Mean and Variance

A proof is given in Bain and Engelhart (1991) that the sample mean and variance are independent under the independent normal model. Further, they show that the sample mean has a normal distribution and the distribution of the sample variance is the same as the distribution of a given constant times a chi square random variable. In particular, if $Y_1, \ldots, Y_n$ is a random sample from a $N(\mu, \sigma^2)$ distribution, then

$$\overline{Y} \sim N\left(\mu, \sigma^2/n\right) \text{ and } \frac{(n-1)\,S^2}{\sigma^2} \sim \chi^2_{n-1}.$$

3.3 Linear Combination of Central Chi Squares

First, we recall the form of the probability density function $f_X(x)$ of a central chi square distribution with ν degrees of freedom. We have

$$f_X(x) = \frac{x^{\nu/2-1}e^{-x/2}}{\Gamma\left(\frac{\nu}{2}\right)2^{\nu/2}} I_{(0,\infty)}(x),$$

where $I_{\mathbf{A}}(x) = 1$ if $x \in \mathbf{A}$ and zero otherwise. The mean and variance are, respectively,

$$\mu_X = \nu \text{ and } \sigma_X^2 = 2\nu.$$

The moment generating function $MGF_X(t)$ of the distribution of X is

$$MGF_X(t) = \left(\frac{1}{1-2t}\right)^{\nu/2}.$$

See Bain and Engelhardt (1991) for more details about the chi square distribution.

Theorem 3.1: If $X_1 \sim \chi^2_{\nu_1}, \ldots, X_m \sim \chi^2_{\nu_m}$ are independent random variables with $\nu_1, \ldots, \nu_m$ positive integers, then

$$X_1 + \ldots + X_m \sim \chi^2_\nu, \text{ with } \nu = \nu_1 + \ldots + \nu_m.$$

Proof of Theorem 3.1: Bain and Engelhardt (1991) give a proof of this theorem using the moment generating function method.■

Theorem 3.2: If $X_i \sim \chi^2_{\nu_i}$ are independent random variables for $i = 1, \cdots, k$, and a_i are distinct positive real numbers, then the probability density function of $U = a_1X_1 + a_2X_2 + \cdots + a_kX_k$ is given by

$$\begin{aligned} f_U(u) &= \sum_{r_1=0}^{\infty} \cdots \sum_{r_{k-1}=0}^{\infty} \xi_{r_1,\ldots,r_{k-1}} f_{\chi^2_{\nu_1+\ldots+\nu_k+2r_1+\ldots+2r_{k-1}}}\left(a_k^{-1}u\right)a_k^{-1} \\ &= \sum_{\mathbf{r}} \xi_{\mathbf{r}} f_{\chi^2_{\nu_{\mathbf{r}}}}\left(a_k^{-1}u\right)a_k^{-1} \end{aligned}$$

where

$$\sum\nolimits_{\mathbf{r}} = \sum\nolimits_{r_1=0}^{\infty} \cdots \sum\nolimits_{r_{k-1}=0}^{\infty}; \ \mathbf{r} = [r_1, \ldots, r_{k-1}]^{\mathbf{T}};$$

$$\nu_{\mathbf{r}} = \nu_1 + \ldots + \nu_k + 2r_1 + \ldots + 2r_{k-1}; \ r_0 = 0; \text{ and}$$

$$\begin{aligned} \xi_{\mathbf{r}} &= \xi_{r_1,\ldots,r_{k-1}} \\ &= \frac{(-1)^{r_1+\ldots+r_{k-1}} \left(\prod_{i=1}^{k-1} (a_{i+1} - a_i)^{r_i}\right) a_k^{(\nu_1+\ldots+\nu_{k-1}+2r_1+\ldots+2r_{k-2})/2}}{a_1^{(\nu_1+2r_1)/2} \left(\prod_{i=2}^{k-1} a_i^{(\nu_i+2r_i+2r_{i-1})/2}\right) \left(\prod_{i=1}^{k-1} r_i!\right)} \\ &\times \frac{\prod_{i=1}^{k-1} \Gamma\left(\frac{\nu_1+\ldots+\nu_i+2r_1+\ldots+2r_i}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right) \prod_{i=2}^{k-1} \Gamma\left(\frac{\nu_1+\ldots+\nu_i+2r_1+\ldots+2r_{i-1}}{2}\right)}. \end{aligned}$$

Proof of Theorem 3.2: Let $X_i \sim \chi^2_{\nu_i}$ be independent random variables and a_i be distinct positive real values for $i = 1, \cdots, k$. Consider the one-to-one transformation

$$U_i = \sum\nolimits_{j=1}^{i} a_j X_j$$

for $i = 1, \ldots, k$. Then, we can write the inverse transformation and the Jacobian as

$$X_1 = a_1^{-1} U_1 \text{ and } X_i = a_i^{-1}(U_i - U_{i-1}) \text{ with } J = \prod_{j=1}^{k} a_j^{-1},$$

for $i = 2, \ldots, k$. For $u_i > 0$ $(i = 2, \ldots, k)$, the joint distribution function of $\mathbf{U} =$

$[U_1, U_2, \cdots, U_k]^{\mathbf{T}}$ can be expressed as

$$\begin{aligned}
f_{\mathbf{U}}(\mathbf{u}) &= f_{X_1}(a_1^{-1}u_1)(\prod_{i=2}^{k} f_{X_i}(a_i^{-1}(u_i-u_{i-1})))(\prod_{i=1}^{k} a_i^{-1}) \\
&= \frac{(a_1^{-1}u_1)^{\nu_1/2-1}e^{-a_1^{-1}u_1/2}(\prod_{i=1}^{k} a_i^{-1})}{2^{(\nu_1+\cdots+\nu_k)/2}(\prod_{i=1}^{k}\Gamma(\frac{\nu_i}{2}))} \\
&\times \prod_{i=2}^{k}(a_i^{-1}(u_i-u_{i-1}))^{\nu_i/2-1}e^{-a_i^{-1}(u_i-u_{i-1})/2}) \\
&= \frac{{u_1}^{\nu_1/2}(\prod_{i=2}^{k}(u_i-u_{i-1}))^{\nu_i/2-1})}{(\prod_{i=1}^{k} a_i^{\nu_i/2})(\prod_{i=1}^{k}\Gamma(\frac{\nu_i}{2}))2^{(\nu_1+\cdots+\nu_k)/2}} \\
&\times e^{-a_1^{-1}u_1/2}(\prod_{i=2}^{k} e^{-a_i^{-1}(u_i-u_{i-1})/2}) \\
&= \frac{(\prod_{i=2}^{k}(\frac{u_{i-1}}{u_i})^{(\nu_1+\cdots+\nu_{i-1})/2-1}(1-\frac{u_{i-1}}{u_i})^{\nu_i/2-1}u_i^{-1})u_k^{(\nu_1+\cdots+\nu_k)/2-1}}{(\prod_{i=1}^{k} a_i^{\nu_i/2})(\prod_{i=1}^{k}\Gamma(\frac{\nu_i}{2}))2^{(\nu_1+\cdots+\nu_k)/2}} \\
&\times e^{-a_k^{-1}u_k/2}(\prod_{i=1}^{k-1} e^{-u_i(a_i^{-1}-a_{i+1}^{-1})/2})
\end{aligned}$$

Using Maclaurin series expansion, we have:

$$\prod_{i=1}^{k-1} e^{-u_i(a_i^{-1}-a_{i+1}^{-1})/2} = \prod_{i=1}^{k-1}\sum_{r_i=0}^{\infty}\frac{(-1)^{r_i}u_i^{r_i}(a_i^{-1}-a_{i+1}^{-1})^{r_i}}{2^{r_i}\, r_i!}$$

Thus, we obtain

$$\begin{aligned}
f_{\mathbf{U}}(\mathbf{u}) &= \sum_{r_1=0}^{\infty}\cdots\sum_{r_{k-1}=0}^{\infty}\frac{(-1)^{r_1+\cdots+r_{k-1}}(\prod_{i=1}^{k-1}(a_i^{-1}-a_{i+1}^{-1})^{r_i})}{(\prod_{i=1}^{k-1} r_i!)(\prod_{i=1}^{k} a_i^{\nu_i/2})(\prod_{i=1}^{k}\Gamma(\frac{\nu_i}{2}))} \\
&\times \frac{e^{-a_k^{-1}u_k/2}\, u_k^{(\nu_1+\cdots+\nu_k+2r_1+\cdots+2r_{k-1})/2-1}}{2^{(\nu_1+\cdots+\nu_k+2r_1+\cdots+2r_{k-1})/2}} \\
&\times (\prod_{i=2}^{k}(\frac{u_{i-1}}{u_i})^{(\nu_1+\cdots+\nu_{i-1}+2r_1+\cdots+2r_{i-1})/2-1}(1-\frac{u_{i-1}}{u_i})^{\nu_i/2-1}u_i^{-1})
\end{aligned}$$

Since

$$\int_0^{u_i}(\frac{u_{i-1}}{u_i})^{c-1}(1-\frac{u_{i-1}}{u_i})^{d-1}u_i^{-1}du_{i-1} = \frac{\Gamma(c)\Gamma(d)}{\Gamma(c+d)}$$

we can write the marginal distribution $f_U(u) = f_{U_k}(u_k)$ as

$$
\begin{aligned}
f_U(u) &= \int_0^{u_k} \cdots \int_0^{u_2} f_{\mathbf{U}}(\mathbf{u})\, du_1 du_2 \cdots du_{k-1} \\
&= \sum_{r_1=0}^{\infty} \cdots \sum_{r_{k-1}=0}^{\infty} \frac{(-1)^{r_1+\cdots+r_{k-1}} (\prod_{i=1}^{k-1} (a_i^{-1} - a_{i+1}^{-1})^{r_i})}{(\prod_{i=1}^{k-1} r_i!)(\prod_{i=1}^{k} a_i^{\nu_i/2})(\prod_{i=1}^{k} \Gamma(\frac{\nu_i}{2}))} \\
&\times \frac{e^{-a_k^{-1} u_k/2}\, (a_k^{-1} u_k)^{(\nu_1+\cdots+\nu_k+2r_1+\cdots+2r_{k-1})/2-1}}{2^{(\nu_1+\cdots+\nu_k+2r_1+\cdots+2r_{k-1})/2}} \\
&\times a_k^{(\nu_1+\cdots+\nu_k+2r_1+\cdots+2r_{k-1})/2-1} \\
&\times \prod_{i=2}^{k} \frac{\Gamma(\frac{\nu_1+\cdots+\nu_{i-1}+2r_1+\cdots+2r_{i-1}}{2})\Gamma(\frac{\nu_i}{2})}{\Gamma(\frac{\nu_1+\cdots+\nu_i+2r_1+\cdots+2r_{i-1}}{2})}
\end{aligned}
$$

Notice that

$$
\begin{aligned}
&\prod_{i=2}^{k} \Gamma(\frac{\nu_1+\cdots+\nu_{i-1}+2r_1+\cdots+2r_{i-1}}{2}) \\
&= \prod_{i=1}^{k-1} \Gamma(\frac{\nu_1+\cdots+\nu_i+2r_1+\cdots+2r_i}{2})
\end{aligned}
$$

and that

$$
\prod_{i=1}^{k-1} (a_i^{-1} - a_{i+1}^{-1})^{r_i} = \frac{\prod_{i=1}^{k-1} (a_{i+1} - a_i)^{r_i}}{\prod_{i=1}^{k-1} (a_i a_{i+1})^{r_i}}.
$$

Also,

$$
(\prod_{i=1}^{k} a_i^{\nu_i/2})(\prod_{i=1}^{k-1} (a_i a_{i+1})^{r_i}) = a_1^{(\nu_1+2r_1)/2} a_k^{(\nu_k+2r_{k-1})/2} (\prod_{i=2}^{k-1} a_i^{(\nu_i+2r_{i-1}+2r_i)/2}).
$$

Thus, we obtain

$$
\begin{aligned}
f_U(u_k) &= \sum_{r_1=0}^{\infty} \cdots \sum_{r_{k-1}=0}^{\infty} \frac{(-1)^{r_1+\cdots+r_{k-1}} (\prod_{i=1}^{k-1} (a_{i+1} - a_i)^{r_i})}{a_1^{(\nu_1+2r_1)/2} (\prod_{i=2}^{k-1} a_i^{(\nu_i+2r_{i-1}+2r_i)/2})(\prod_{i=1}^{k-1} r_i!)\Gamma(\frac{\nu_1}{2})} \\
&\times \frac{\prod_{i=1}^{k-1} \Gamma(\frac{\nu_1+\cdots+\nu_i+2r_1+\cdots+2r_i}{2})}{\prod_{i=2}^{k-1} \Gamma(\frac{\nu_1+\cdots+\nu_i+2r_1+\cdots+2r_{i-1}}{2})} a_k^{(\nu_1+\cdots+\nu_{k-1}+2r_1+\cdots+2r_{k-2})/2-1} \\
&\times \frac{e^{-a_k^{-1} u_k/2}\, (a_k^{-1} u_k)^{(\nu_1+\cdots+\nu_k+2r_1+\cdots+2r_{k-1})/2-1}}{\Gamma(\frac{\nu_1+\cdots+\nu_k+2r_1+\cdots 2r_{k-1}}{2}) 2^{(\nu_1+\cdots+\nu_k+2r_1+\cdots+2r_{k-1})/2}}.
\end{aligned}
$$

The results follow.■

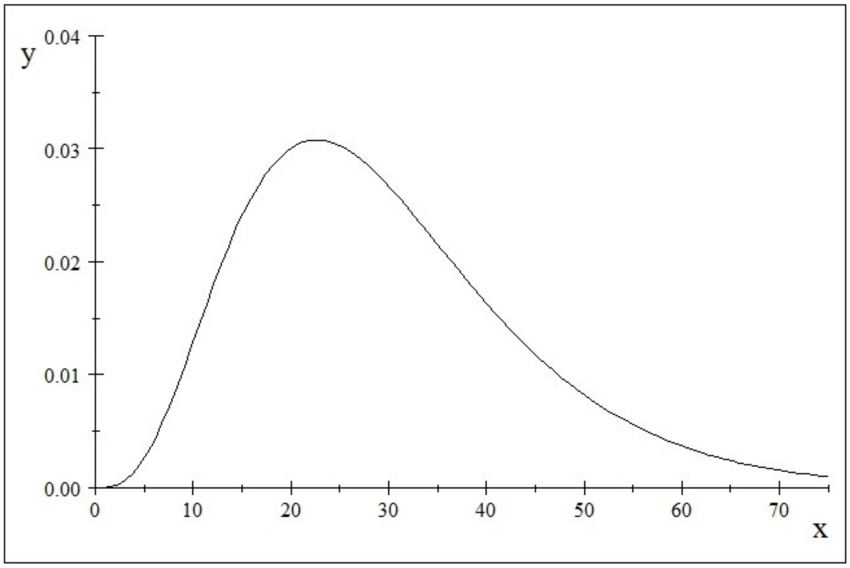

Figure 3.1: Graph of $f_{2\chi_3^2+4\chi_6^2}(u)$

Consider the cases in which $k = 2$ and $k = 3$. For $k = 2$ with $\nu_1 = 3$, $\nu_2 = 6$, $a_1 = 2$, and $a_2 = 4$, the probability density function describing the distribution of

$$U = 2\chi_3^2 + 4\chi_6^2$$

is given by

$$f_U(u) = \left(\frac{4}{2}\right)^{3/2} \sum_{r=0}^{\infty} \frac{(-1)^r \Gamma\left(\frac{3+2r}{2}\right)}{4\Gamma\left(\frac{3}{2}\right) r!} f_{\chi^2_{9+2r}}\left(4^{-1}u\right).$$

The graph of $f_U(u)$ is given in Figure 3.1.

For $k = 3$ with $\nu_1 = 3$, $\nu_2 = 6$, $\nu_3 = 7$, $a_1 = 2$, $a_2 = 4$, and $a_3 = 5$. The probability density function describing the distribution of

$$U = 2\chi_3^2 + 4\chi_6^2 + 5\chi_7^2$$

is given by

$$f_U(u) = \sum_{r=0}^{\infty} \sum_{t=0}^{\infty} \frac{(-1)^{r+t} \left(\frac{1}{4}\right)^r \left(\frac{1}{20}\right)^t \Gamma\left(\frac{3+2r}{2}\right) \Gamma\left(\frac{9+2r+2t}{2}\right) 5^{(7+2r+2t)/2}}{2^{15/2} r! t! \Gamma\left(\frac{3}{2}\right) \Gamma\left(\frac{9+2r}{2}\right)} f_{\chi^2_{16+2r+2t}}\left(5^{-1}u\right)$$

The graph of this density is given in Figure 3.2.

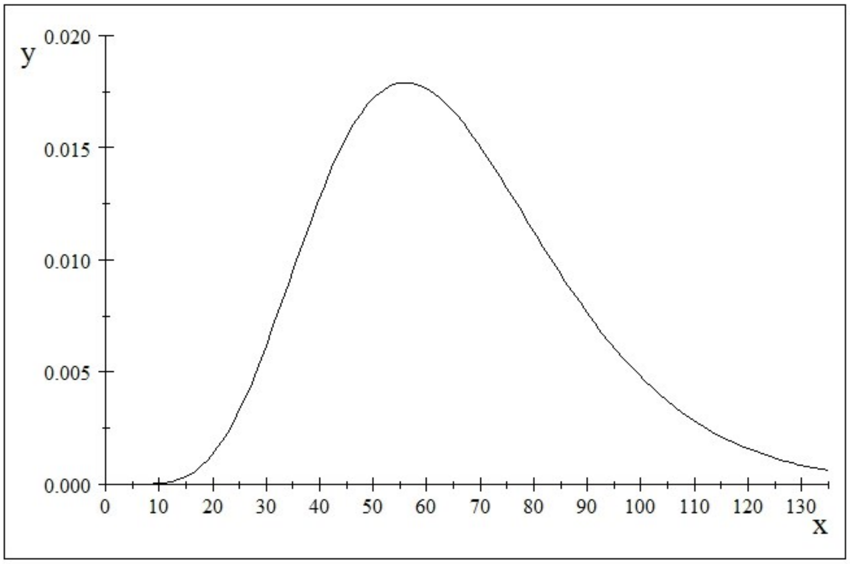

Figure 3.2: Graph of $f_{2\chi_3^2+4\chi_6^2+5\chi_7^2}(u)$

3.4 Linear Combination of Noncentral Chi Squares Each with One Degree of Freedom

In this section, we look at the probability distribution function of a non-central chi-square and some of its properties. We also derive a closed form expression for the probability distribution function describing the distribution of a linear combination of non-central chi-squares, each with one degree of freedom. Several approximations to this distribution have been done in the literature (Press 1966, Davis 1977).

Theorem 3.3: If Z has a standard normal distribution and τ is a real number, then the probability density function describing the distribution of

$$W = (Z + \tau)^2$$

is

$$f_W(w) = e^{-\tau^2/2} \sum\nolimits_{r=0}^{\infty} \frac{\theta_{r,\tau^2} \Gamma\left(\frac{1+2r}{2}\right) 2^r}{\sqrt{\pi}\,(2r)!} f_{\chi^2_{1+2r}}(w)$$
$$= e^{-\tau^2/2} \left(f_{\chi^2_1}(w) + \sum\nolimits_{r=1}^{\infty} \frac{(\tau^2)^r \Gamma\left(\frac{1+2r}{2}\right) 2^r}{\sqrt{\pi}\,(2r)!} f_{\chi^2_{1+2r}}(w) \right),$$

where

$$\theta_{r,\tau^2} = \begin{cases} 1, & \text{if } r = 0; \\ (\tau^2)^r, & \text{if } r > 0. \end{cases}$$

Proof of Theorem 3.3: The cumulative distribution function describing the distribution of W for is given by

$$F_W(w) = P\left((Z+\tau)^2 \leq w\right) = P\left(-\sqrt{w} - \tau \leq Z \leq \sqrt{w} - \tau\right) I_{(0,\infty)}(w)$$
$$= \left[\Phi\left(\sqrt{w} - \tau\right) - \Phi\left(-\sqrt{w} - \tau\right)\right] I_{(0,\infty)}(w),$$

where $\Phi(z)$ is the cumulative distribution function of a standard normal distribution. It follows that the probability density function describing the distribution of W is

$$f_W(w) = \left[\frac{1}{2} w^{-1/2} \phi\left(\sqrt{w} - \tau\right) + \frac{1}{2} w^{-1/2} \phi\left(-\sqrt{w} - \tau\right)\right] I_{(0,\infty)}(w)$$
$$= \left[\frac{1}{2\sqrt{2\pi}} w^{-1/2} e^{-\left(\sqrt{w}-\tau\right)^2/2} + \frac{1}{2\sqrt{2\pi}} w^{-1/2} e^{-\left(-\sqrt{w}-\tau\right)^2/2}\right] I_{(0,\infty)}(w)$$
$$= \frac{e^{-\tau^2/2}}{2\sqrt{2\pi}} w^{-1/2} e^{-w/2} \left(e^{\tau\sqrt{w}} + e^{-\tau\sqrt{w}}\right) I_{(0,\infty)}(w),$$

where $\phi(z)$ is the probability density function of a standard normal distribution. Next we observe that

$$e^{\tau\sqrt{w}} + e^{-\tau\sqrt{w}} = \sum\nolimits_{r=0}^{\infty} \frac{\tau^r w^{r/2}}{r!} + \sum\nolimits_{r=0}^{\infty} \frac{(-1)^r \tau^r w^{r/2}}{r!}$$
$$= \sum\nolimits_{r=0}^{\infty} \left(1 + (-1)^r\right) \frac{\tau^r w^{r/2}}{r!}$$
$$= 2 \sum\nolimits_{r=0}^{\infty} \frac{(\tau^2)^r w^r}{(2r)!}.$$

Hence, we have

$$\begin{aligned} f_W(w) &= \frac{e^{-\tau^2/2}}{2\sqrt{2\pi}} w^{-1/2} e^{-w/2} \left(2 \sum_{r=0}^{\infty} \frac{(\tau^2)^r w^r}{(2r)!} \right) I_{(0,\infty)}(w) \\ &= e^{-\tau^2/2} \sum_{r=0}^{\infty} \frac{(\tau^2)^r}{\sqrt{2\pi}\,(2r)!} w^{(2r-1)/2} e^{-w/2} I_{(0,\infty)}(w) \\ &= e^{-\tau^2/2} \sum_{r=0}^{\infty} \frac{(\tau^2)^r}{\sqrt{2\pi}\,(2r)!} w^{(1+2r)/2-1} e^{-w/2} I_{(0,\infty)}(w) \\ &= e^{-\tau^2/2} \sum_{r=0}^{\infty} \frac{(\tau^2)^r \Gamma\left(\frac{1+2r}{2}\right) 2^{(1+2r)/2}}{\sqrt{2\pi}\,(2r)!} \\ &\times \frac{1}{\Gamma\left(\frac{1+2r}{2}\right) 2^{(1+2r)/2}} w^{(1+2r)/2-1} e^{-w/2} I_{(0,\infty)}(w) \\ &= e^{-\tau^2/2} \sum_{r=0}^{\infty} \frac{(\tau^2)^r \Gamma\left(\frac{1+2r}{2}\right) 2^{(1+2r)/2}}{\sqrt{2\pi}\,(2r)!} f_{\chi^2_{1+2r}}(w) \\ &= e^{-\tau^2/2} \sum_{r=0}^{\infty} \frac{(\tau^2)^r \Gamma\left(\frac{1+2r}{2}\right) 2^r}{\sqrt{\pi}\,(2r)!} f_{\chi^2_{1+2r}}(w) . \blacksquare \end{aligned}$$

The random variable W is said to have a noncentral chi square distribution with 1 degree of freedom and noncentrality parameter τ^2 or more compactly stated as $W \sim \chi^2_{1,\tau^2}$.

Theorem 3.4: If $W \sim \chi^2_{\alpha,\tau^2}$, then the moment generating function of the distribution of W is

$$MGF_W(t) = (1-2t)^{-\alpha/2} e^{-\tau^2 t/(1-2t)}$$

for $t \in (-1/2, 1/2)$.

Proof of Theorem 3.4: The proof of this theorem can be found in Tanizaki (2004).■

Theorem 3.5: If $Z_1, \ldots, Z_\nu$ are ν independent random variables each having a standard normal distribution and $\tau_1, \ldots, \tau_\nu$ are real numbers, then the distribution of

$$W = (Z_1 + \tau_1)^2 + \ldots + (Z_\nu + \tau_\nu)^2$$

is a noncentral chi square distribution with ν degrees of freedom and noncentrality parameter

$$\tau^2 = \tau_1^2 + \ldots + \tau_\nu^2.$$

Proof of Theorem 3.5: The moment generating function of W is

$$MGF_W(t) = \prod_{i=1}^{\nu} MGF_{(Z_i+\tau_i)^2}(t) = \prod_{i=1}^{\nu} (1-2t)^{-1/2} e^{-\tau^2 t/(1-2t)}$$
$$= (1-2t)^{-\nu/2} e^{-\left(\tau_1^2+\cdots+\tau_\nu^2\right)t/(1-2t)}.$$

This the moment generating function of a noncentral chi square distribution with ν degrees of freedom and noncentral parameter $\tau_1^2+\cdots+\tau_\nu^2$.■

Theorem 3.6: If $W_1 \sim \chi^2_{1,\tau_1^2},\ldots,W_\eta \sim \chi^2_{1,\tau_\eta^2}$ are stochasticly independent noncentral chi square random variables each with one degree of freedom and $c_1,\ldots,c_\eta$ are distinct positive real numbers for $\eta \geq 2$, then the probability density function describing the distribution of

$$W = c_1 W_1 +,\ldots + c_\eta W_\eta \sim c_1\chi^2_{1,\tau_1^2} + \ldots + c_\eta \chi^2_{1,\tau_\eta^2}$$

is for $w > 0$ given by

$$f_W(w) = \sum_{t_1=0}^{\infty}\cdots\sum_{t_\eta=0}^{\infty}\sum_{s_1=0}^{\infty}\cdots\sum_{s_{\eta-1}=0}^{\infty}\zeta_{t_1,\ldots,t_\eta,s_1,\ldots,s_{\eta-1}}$$
$$\times f_{\chi^2_{\nu+2t_1+\ldots+2t_\eta+2s_1+\ldots+2s_{\eta-1}}}\left(c_\eta^{-1}w\right)c_\eta^{-1}$$
$$= \sum_{\mathbf{t},\mathbf{s}} \zeta_{\mathbf{t},\mathbf{s}} f_{\chi^2_{\nu_{\mathbf{t},\mathbf{s}}}}\left(c_\eta^{-1}w\right)c_\eta^{-1},$$

for $w > 0$ and $f_W(w) = 0$ otherwise, where

$$\zeta_{\mathbf{t},\mathbf{s}} = \zeta_{t_1,\cdots,t_\eta,s_1,\cdots,s_{\eta-1}} = \frac{(\prod_{i=1}^{\eta}\theta_{t_i,\tau_i^2})(-1)^{s_1+\cdots+s_{\eta-1}}2^{t_1+\cdots+t_\eta}e^{-(\tau_1^2+\cdots+\tau_\eta^2)/2}}{c_1^{(1+2t_1+2s_1)/2}(\sqrt{\pi})^\eta(\prod_{i=1}^{\eta}2t_i!)(\prod_{i=1}^{\eta-1}s_i!)}$$
$$\times\frac{(\prod_{i=1}^{\eta-1}(c_{i+1}-c_i)^{s_i})\,c_\eta^{(\eta-1+2t_1+\cdots+2t_{\eta-1}+2s_1+\cdots+2s_{\eta-2})/2}}{(\prod_{i=2}^{\eta-1}c_i^{1+2t_i+2s_{i-1}+2s_i})}$$
$$\times\frac{(\prod_{i=1}^{\eta-1}\Gamma(\frac{i+2t_1+\cdots+2t_i+2s_1+\cdots+2s_i}{2}))(\prod_{i=2}^{\eta}\Gamma(\frac{1+2t_i}{2}))}{(\prod_{i=2}^{\eta-1}\Gamma(\frac{i+2t_1+\cdots+2t_i+2s_1+\cdots+2s_{i-1}}{2}))};$$

$$\theta_{t_i,\tau_i^2} = \begin{cases} 1, & \text{if } t_i = 0; \\ \left(\tau_i^2\right)^{t_i}, & \text{if } t_i > 0. \end{cases}\ ;\text{ and } \mathbf{t} = (t_1,\ldots,t_\eta)\ \text{ and } \mathbf{s} = (s_1,\ldots,s_{\eta-1}).$$

Proof of Theorem 3.6: Let $W_i \sim \chi^2_{1,\tau_i^2}$be independent random variables and c_i be positive real values for $i = 1, \cdots, \eta$. Consider the one-to-one transformation

$$W_i = \sum_{j=1}^{i} c_j X_j$$

for $i = 1,\ldots,\eta$. Note that $W = W_\eta$. Then, we can write the inverse transformation and the Jacobian as

$$X_1 = c_1^{-1} W_1 \text{ and } X_i = c_i^{-1}(W_i - W_{i-1}) \text{ with } J = \prod_{j=1}^{\eta} c_j^{-1},$$

for $i = 2,\ldots,\eta$. For $w_i > 0$ $(i = 1,\ldots,\eta)$, the joint distribution function of $\mathbf{W} = [W_1, W_2, \cdots, W_\eta]^{\mathbf{T}}$ can be expressed as

$$\begin{aligned}
f_{\mathbf{W}}(\mathbf{w}) &= f_{X_1}(c_1^{-1}w_1)(\prod_{i=2}^{\eta} f_{X_i}(c_i^{-1}(w_i - w_{i-1})))(\prod_{i=1}^{\eta} c_i^{-1}) \\
&= e^{-\tau_1^2/2} \sum_{t_1=0}^{\infty} \frac{\theta_{t_1,\tau_1^2} \Gamma\left(\frac{1+2t_1}{2}\right) 2^{t_1} (c_1^{-1}w_1)^{(1+2t_1)/2-1} e^{-c_1^{-1}w_1/2}}{\sqrt{\pi}\,(2t_1)!\Gamma\left(\frac{1+2t_1}{2}\right) 2^{(1+2t_1)/2}} \\
&\times \prod_{i=2}^{\eta} \sum_{t_i=0}^{\infty} \frac{\theta_{t_i,\tau_i^2} \Gamma\left(\frac{1+2t_i}{2}\right) 2^{t_i} (c_i^{-1}(w_i - w_{i-1}))^{(1+2t_i)/2-1}}{(2t_i)!\Gamma\left(\frac{1+2t_i}{2}\right) 2^{(1+2t_i)/2}} \\
&\times \prod_{i=2}^{\eta} \frac{e^{-\tau_i^2/2} e^{-c_i^{-1}(w_i - w_{i-1})/2}}{(\sqrt{\pi})^i} \times \prod_{i=1}^{\eta} c_i^{-1} \\
&= \sum_{t_1=0}^{\infty} \cdots \sum_{t_\eta=0}^{\infty} \frac{(\prod_{i=1}^{\eta} \theta_{t_i,\tau_i^2}) w_1^{(1+2t_1)/2-1} (\prod_{i=2}^{\eta} (w_i - w_{i-1})^{(1+2t_i)/2-1})}{2^{\eta/2} (\sqrt{\pi})^{\eta} (\prod_{i=1}^{\eta} (2t_i)!) (\prod_{i=1}^{\eta} c_i^{(1+2t_i)/2})} \\
&\times e^{-(\tau_1^2+\cdots+\tau_\eta^2)/2} e^{-c_\eta^{-1} w_\eta} (\prod_{i=1}^{\eta-1} e^{-w_i(c_i^{-1} - c_{i+1}^{-1})/2}) \\
&= \sum_{t_1=0}^{\infty} \cdots \sum_{t_\eta=0}^{\infty} \frac{(\prod_{i=1}^{\eta} \theta_{t_i,\tau_i^2}) e^{-(\tau_1^2+\cdots+\tau_\eta^2)/2} e^{-c_\eta^{-1} w_\eta}}{2^{\eta/2} (\sqrt{\pi})^{\eta} (\prod_{i=1}^{\eta} (2t_i)!) (\prod_{i=1}^{\eta} c_i^{(1+2t_i)/2})} \\
&\times (\prod_{i=2}^{\eta} (\frac{w_{i-1}}{w_i})^{(i-1+2t_1+\cdots+2t_{i-1})/2-1} (1 - \frac{w_{i-1}}{w_i})^{(1+2t_i)/2-1} w_i^{-1}) \\
&\times (\prod_{i=1}^{\eta-1} e^{-w_i(c_i^{-1} - c_{i+1}^{-1})/2})\, w_\eta^{(\eta+2t_1+\cdots+2t_\eta)/2-1}.
\end{aligned}$$

Using Maclaurin series expansion, we have:

$$\prod_{i=1}^{\eta-1} e^{-w_i(c_i^{-1} - c_{i+1}^{-1})/2} = \prod_{i=1}^{\eta-1} \sum_{s_i=0}^{\infty} \frac{(-1)^{s_i} w_i^{s_i} (c_i^{-1} - c_{i+1}^{-1})^{s_i}}{2^{s_i}\, s_i!}.$$

Hence,

$$f_{\mathbf{W}}(\mathbf{w}) = \sum_{t_1=0}^{\infty}\cdots\sum_{t_\eta=0}^{\infty}\sum_{s_1=0}^{\infty}\cdots\sum_{s_{\eta-1}=0}^{\infty}\frac{(\prod_{i=1}^{\eta}\theta_{t_i,\tau_i^2})e^{-(\tau_1^2+\cdots+\tau_\eta^2)/2}(-1)^{s_1+\cdots+s_{\eta-1}}}{(\sqrt{\pi})^{\eta}(\prod_{i=1}^{\eta}(2t_i)!)(\prod_{i=1}^{\eta}c_i^{(1+2t_i)/2})}$$
$$\times\frac{e^{-c_\eta^{-1}w_\eta/2}(\prod_{i=1}^{\eta-1}(c_i^{-1}-c_{i+1}^{-1})^{s_i})\,w_\eta^{(\eta+2t_1+\cdots+2t_\eta+2s_1+\cdots+2s_{\eta-1})/2-1}}{(\prod_{i=1}^{\eta-1}s_i!)\,2^{(\eta+2s_1+\cdots+2s_{\eta-1})/2}}$$
$$\times(\prod_{i=2}^{\eta}(\frac{w_{i-1}}{w_i})^{(i-1+2t_1+\cdots+2t_{i-1}+2s_1+\cdots+2s_{i-1})/2-1}(1-\frac{w_{i-1}}{w_i})^{(1+2t_i)/2-1}w_i^{-1}).$$

Since

$$\int_0^{w_i}(\frac{w_{i-1}}{w_i})^{a-1}(1-\frac{w_{i-1}}{w_i})^{b-1}w_i^{-1}dw_{i-1}=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)},$$

we can write the marginal distribution of W_η as

$$f_{W_\eta}(w_\eta)=\int_0^{w_\eta}\cdots\int_0^{w_2}f_{\mathbf{W}}(\mathbf{w})\,dw_1dw_2\cdots dw_{\eta-1}$$
$$=\sum_{t_1=0}^{\infty}\cdots\sum_{t_\eta=0}^{\infty}\sum_{s_1=0}^{\infty}\cdots\sum_{s_{\eta-1}=0}^{\infty}\frac{(\prod_{i=1}^{\eta}\theta_{t_i,\tau_i^2})e^{-(\tau_1^2+\cdots+\tau_\eta^2)/2}(-1)^{s_1+\cdots+s_{\eta-1}}}{(\sqrt{\pi})^{\eta}(\prod_{i=1}^{\eta}(2t_i)!)}$$
$$\times\frac{e^{-c_\eta^{-1}w_\eta/2}(\prod_{i=1}^{\eta-1}(c_i^{-1}-c_{i+1}^{-1})^{s_i})\,(c_\eta^{-1}w_\eta)^{(\eta+2t_1+\cdots+2t_\eta+2s_1+\cdots+2s_{\eta-1})/2-1}}{(\prod_{i=1}^{\eta}c_i^{(1+2t_i)/2})(\prod_{i=1}^{\eta-1}s_i!)}$$
$$\times\frac{2^{t_1+\cdots+t_\eta}c_\eta^{(\eta+2t_1+\cdots+2t_\eta+2s_1+\cdots+2s_{\eta-1})/2-1}}{2^{(\eta+2t_1+\cdots+2t_\eta+2s_1+\cdots+2s_{\eta-1})/2}}$$
$$\times\prod_{i=2}^{\eta}\frac{\Gamma(\frac{i-1+2t_1+\cdots+2t_{i-1}+2s_1+\cdots+2s_{i-1}}{2})\Gamma(\frac{1+2t_i}{2})}{\Gamma(\frac{i+2t_1+\cdots+2t_i+2s_1+\cdots+2s_{i-1}}{2})}.$$

Notice that

$$\prod_{i=2}^{\eta}\Gamma(\frac{i-1+2t_1+\cdots+2t_{i-1}+2s_1+\cdots+2s_{i-1}}{2})$$
$$=\prod_{i=1}^{\eta-1}\Gamma(\frac{i+t_1+\cdots+t_i+2s_1+\cdots+2s_i}{2})$$

and that

$$\prod_{i=1}^{\eta-1}(c_i^{-1}-c_{i+1}^{-1})^{s_i}=\frac{\prod_{i=1}^{\eta-1}(c_{i+1}-c_i)^{s_i}}{\prod_{i=1}^{\eta-1}(c_ic_{i+1})^{s_i}}.$$

Also,

$$\begin{aligned}
&(\prod_{i=1}^{\eta} c_i^{(1+2t_i)/2})(\prod_{i=1}^{\eta-1}(c_i c_{i+1})^{s_i}) \\
&= c_1^{(1+2t_1+2s_1)/2} c_\eta^{(1+2t_\eta+2s_{\eta-1})/2}(\prod_{i=2}^{\eta-1} c_i^{(1+2t_i+2s_{i-1}+2s_i)/2}).
\end{aligned}$$

Hence, the marginal distribution $f_W(w)$ becomes

$$\begin{aligned}
f_W(w) = &\sum_{t_1=0}^{\infty}\cdots\sum_{t_\eta=0}^{\infty}\sum_{s_1=0}^{\infty}\cdots\sum_{s_{\eta-1}=0}^{\infty} \frac{(\prod_{i=1}^{\eta}\theta_{t_i,\tau_i^2})e^{-(\tau_1^2+\cdots+\tau_\eta^2)/2}(-1)^{s_1+\cdots+s_{\eta-1}}}{(\sqrt{\pi})^\eta(\prod_{i=1}^{\eta}(2t_i)!)(\prod_{i=1}^{\eta-1} s_i!)} \\
&\times \frac{2^{t_1+\cdots+t_\eta}(\prod_{i=1}^{\eta-1}(c_{i+1}-c_i)^{s_i})\, c_\eta^{(\eta-1+2t_1+\cdots+2t_{\eta-1}+2s_1+\cdots+2s_{\eta-2})/2-1}}{c_1^{(1+2t_1+2s_1)/2}(\prod_{i=2}^{\eta-1} c_i^{(1+2t_i+2s_{i-1}+2s_i)/2})} \\
&\times \frac{(\prod_{i=1}^{\eta-1}\Gamma(\frac{i+2t_1+\cdots+2t_i+2s_1+\cdots+2s_i}{2}))(\prod_{i=2}^{\eta}\Gamma(\frac{1+2t_i}{2}))}{(\prod_{i=2}^{\eta-1}\Gamma(\frac{i+2t_1+\cdots+2t_i+2s_1+\cdots+2s_{i-1}}{2}))} \\
&\times \frac{e^{-c_\eta^{-1}w_\eta/2}\,(c_\eta^{-1}w_\eta)^{(\eta+2t_1+\cdots+2t_\eta+2s_1+\cdots+2s_{\eta-1})/2-1}}{2^{(\eta+2t_1+\cdots+2t_\eta+2s_1+\cdots+2s_{\eta-1})/2}\Gamma(\frac{\eta+2t_1+\cdots+2t_\eta+2s_1+\cdots+2s_{\eta-1}}{2})}.
\end{aligned}$$

The results follow.■

3.5 Ratio of a Linear Combination of Noncentral Chi Square and Central Chi Square Random Variables

It is useful to examine the distribution of a constant times the ratio of W and U. The probability density function describing the distribution of this ratio is given in the following theorem.

Theorem 3.7: Let $X_i \sim \chi^2_{\nu_i}$are independent random variables for $i = 1, \cdots, k$, and a_i are distinct positive real numbers, and let $U = a_1X_1 + a_2X_2 + \cdots + a_kX_k$. Let $W_1 \sim \chi^2_{1,\tau_1^2}, \ldots, W_k \sim \chi^2_{1,\tau_\eta^2}$ be stochasticly independent noncentral chi square random variables each with one degree of freedom and $c_1, \ldots, c_\eta$ are distinct positive real numbers for $\eta \geq 2$, and let $W = c_1W_1+, \ldots + c_\eta W_\eta$. Then for positive real numbers d_w and d_u the distribution of

$$Q = \frac{d_u}{d_w}\frac{W}{U} = \frac{W/d_w}{U/d_u}$$

is given by

$$f_Q(q) = \sum\nolimits_{\mathbf{t},\mathbf{s}} \sum\nolimits_{\mathbf{r}} \frac{\zeta_{\mathbf{t},\mathbf{s}} \xi_{\mathbf{r}} \left(\frac{a_k d_w}{c_\eta d_u}\right)^{\nu_{\mathbf{t},\mathbf{s}}/2} \Gamma\left(\frac{\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}}}{2}\right) q^{\nu_{\mathbf{t},\mathbf{s}}/2-1}}{\Gamma\left(\frac{\nu_{\mathbf{t},\mathbf{s}}}{2}\right) \Gamma\left(\frac{\nu_{\mathbf{r}}}{2}\right) \left(1+\frac{a_k d_w}{c_\eta d_u} q\right)^{(\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}})/2}}$$

where $\sum_{\mathbf{t},\mathbf{s}}$, $\sum_{\mathbf{r}}$, $\nu_{\mathbf{t},\mathbf{s}}$, $\nu_{\mathbf{r}}$, $\zeta_{\mathbf{t},\mathbf{s}}$, and $\xi_{\mathbf{r}}$ are defined as in theorems 3.2 and 3.6.

Proof of Theorem 3.7:

The distributions of U and W are given in theorems 3.2 and 3.6 as

$$f_U(u) = \sum\nolimits_{\mathbf{r}} \xi_{\mathbf{r}} f_{\chi^2_{\nu_{\mathbf{r}}}}\left(a_k^{-1} u\right) a_k^{-1} \text{ and } f_W(w) = \sum\nolimits_{\mathbf{t},\mathbf{s}} \zeta_{\mathbf{t},\mathbf{s}} f_{\chi^2_{\nu_{\mathbf{t},\mathbf{s}}}}\left(c_\eta^{-1} w\right) c_\eta^{-1}.$$

Define

$$Q = \frac{W/d_w}{U/d_u} \text{ and } Q_2 = U/d_u$$
$$W = d_w Q Q_2 \text{ and } U = d_u Q_2 \text{ and } J = d_w d_u q_2$$

$$\begin{aligned}
f_{Q,Q_2}(q,q_2) &= f_W(d_w q q_2) f_U(d_u q_2) d_w d_u q_2 \\
&= \sum\nolimits_{\mathbf{t},\mathbf{s}} \sum\nolimits_{\mathbf{r}} \zeta_{\mathbf{t},\mathbf{s}} \xi_{\mathbf{r}} f_{\chi^2_{\nu_{\mathbf{t},\mathbf{s}}}}\left(c_\eta^{-1} d_w q q_2\right) f_{\chi^2_{\nu_{\mathbf{r}}}}\left(a_k^{-1} d_u q_2\right) \\
&\times c_\eta^{-1} a_k^{-1} d_w d_u q_2 \\
&= \sum\nolimits_{\mathbf{t},\mathbf{s}} \sum\nolimits_{\mathbf{r}} \zeta_{\mathbf{t},\mathbf{s}} \xi_{\mathbf{r}} \frac{\left(c_\eta^{-1} d_w q q_2\right)^{\nu_{\mathbf{t},\mathbf{s}}/2-1} e^{-\left(c_\eta^{-1} d_w q q_2\right)/2}}{\Gamma\left(\frac{\nu_{\mathbf{t},\mathbf{s}}}{2}\right) 2^{\nu_{\mathbf{t},\mathbf{s}}/2}} \\
&\times \frac{\left(a_k^{-1} d_u q_2\right)^{\nu_{\mathbf{r}}/2-1} e^{-\left(a_k^{-1} d_u q_2\right)/2}}{\Gamma\left(\frac{\nu_{\mathbf{r}}}{2}\right) 2^{\nu_{\mathbf{r}}/2}} c_\eta^{-1} a_k^{-1} d_w d_u q_2 \\
&= \sum\nolimits_{\mathbf{t},\mathbf{s}} \sum\nolimits_{\mathbf{r}} \frac{\zeta_{\mathbf{t},\mathbf{s}} \xi_{\mathbf{r}} \left(c_\eta^{-1} d_w\right)^{\nu_{\mathbf{t},\mathbf{s}}/2} \left(a_k^{-1} d_u\right)^{\nu_{\mathbf{r}}/2} q^{\nu_{\mathbf{t},\mathbf{s}}/2-1}}{\Gamma\left(\frac{\nu_{\mathbf{t},\mathbf{s}}}{2}\right) \Gamma\left(\frac{\nu_{\mathbf{r}}}{2}\right) 2^{(\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}})/2} \left(a_k^{-1} d_u + c_\eta^{-1} d_w q\right)^{(\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}})/2}} \\
&\times \left(\left(a_k^{-1} d_u + c_\eta^{-1} d_w q\right) q_2\right)^{(\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}})/2-1} \\
&\times e^{-\left(a_k^{-1} d_u + c_\eta^{-1} d_w q\right) q_2/2} \left(a_k^{-1} d_u + c_\eta^{-1} d_w q\right)
\end{aligned}$$

It follows that

$$
\begin{aligned}
f_Q(q) &= \int_0^\infty f_{Q,Q_2}(q,q_2)\,dq_2 \\
&= \sum\nolimits_{\mathbf{t},\mathbf{s}} \sum\nolimits_{\mathbf{r}} \frac{\zeta_{\mathbf{t},\mathbf{s}}\xi_{\mathbf{r}} \left(c_\eta^{-1} d_w\right)^{\nu_{\mathbf{t},\mathbf{s}}/2} \left(a_k^{-1} d_u\right)^{\nu_{\mathbf{r}}/2} q^{\nu_{\mathbf{t},\mathbf{s}}/2-1}}{\Gamma\left(\frac{\nu_{\mathbf{t},\mathbf{s}}}{2}\right)\Gamma\left(\frac{\nu_{\mathbf{r}}}{2}\right) 2^{(\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}})/2} \left(a_k^{-1} d_u + c_\eta^{-1} d_w q\right)^{(\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}})/2}} \\
&\times \int_0^\infty \left(\left(a_k^{-1} d_u + c_\eta^{-1} d_w q\right) q_2\right)^{(\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}})/2-1} \\
&\times e^{-\left(a_k^{-1} d_u + c_\eta^{-1} d_w q\right) q_2/2} \left(a_k^{-1} d_u + c_\eta^{-1} d_w q\right) dq_2 \\
&= \sum\nolimits_{\mathbf{t},\mathbf{s}} \sum\nolimits_{\mathbf{r}} \frac{\zeta_{\mathbf{t},\mathbf{s}}\xi_{\mathbf{r}} \left(c_\eta^{-1} d_w\right)^{\nu_{\mathbf{t},\mathbf{s}}/2} \left(a_k^{-1} d_u\right)^{\nu_{\mathbf{r}}/2} q^{\nu_{\mathbf{t},\mathbf{s}}/2-1}}{\Gamma\left(\frac{\nu_{\mathbf{t},\mathbf{s}}}{2}\right)\Gamma\left(\frac{\nu_{\mathbf{r}}}{2}\right) 2^{(\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}})/2} \left(a_k^{-1} d_u + c_\eta^{-1} d_w q\right)^{(\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}})/2}} \\
&\times \Gamma\left(\frac{\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}}}{2}\right) 2^{(\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}})/2} \\
&= \sum\nolimits_{\mathbf{t},\mathbf{s}} \sum\nolimits_{\mathbf{r}} \frac{\zeta_{\mathbf{t},\mathbf{s}}\xi_{\mathbf{r}} \left(c_\eta^{-1} d_w\right)^{\nu_{\mathbf{t},\mathbf{s}}/2} \left(a_k^{-1} d_u\right)^{\nu_{\mathbf{r}}/2} \Gamma\left(\frac{\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}}}{2}\right) q^{\nu_{\mathbf{t},\mathbf{s}}/2-1}}{\Gamma\left(\frac{\nu_{\mathbf{t},\mathbf{s}}}{2}\right)\Gamma\left(\frac{\nu_{\mathbf{r}}}{2}\right) \left(a_k^{-1} d_u + c_\eta^{-1} d_w q\right)^{(\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}})/2}} \\
&= \sum\nolimits_{\mathbf{t},\mathbf{s}} \sum\nolimits_{\mathbf{r}} \frac{\zeta_{\mathbf{t},\mathbf{s}}\xi_{\mathbf{r}} \left(\frac{a_k d_w}{c_\eta d_u}\right)^{\nu_{\mathbf{t},\mathbf{s}}/2} \Gamma\left(\frac{\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}}}{2}\right) q^{\nu_{\mathbf{t},\mathbf{s}}/2-1}}{\Gamma\left(\frac{\nu_{\mathbf{t},\mathbf{s}}}{2}\right)\Gamma\left(\frac{\nu_{\mathbf{r}}}{2}\right) \left(1 + \frac{a_k d_w}{c_\eta d_u} q\right)^{(\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}})/2}}. \blacksquare
\end{aligned}
$$

3.6 Conclusion

Some distributional results were given. A closed form expression was derived for a linear combination of (central) chi square variates. Also, a closed form expression was derived for a linear combination of noncentral chi square variates each with one degree of freedom. In both cases, the resulting forms of the probability density functions are convex convolutions of probability density functions of central chi squares. This was followed by a closed form expression for the probability density function describing the distribution of a constant times the ratio of a linear combination of noncentral chi square variates each with one degree of freedom and a linear combination of central chi squares.

CHAPTER 4

ANALYSIS ASSUMING EQUAL VARIANCES

4.1 Model

It is often of interest to compare the means of more than two populations. This occurs in the design of experiments when the research wishes to investigate the effect of one factor on a response variable in which the factor has more than two levels. The levels of this factor that are of interest to the researcher may be considered to be fixed or a sample from the possible values of the factor. We are interested in this thesis in examining the case in which (1) the researcher is interested in comparing the means of k fixed populations or (2) a designed experiment with one factor in which the k levels of the factor are fixed.

A commonly used statistical method for comparing these means is known as the ANalysis Of VAriance (ANOVA). However, this method was constructed by making three assumptions about the data. The firstly of these assumes that the response variables $Y_{i,j}$ to be taken on the jth individual from the ith population are uncorrelated. Secondly, it is assumed that $Y_{i,j} \sim N\left(\mu_i, \sigma_i^2\right)$. These first two assumptions imply that the $Y_{i,j}$'s are independent. Thirdly, it is assumed that the k population variances are equal. That is,

$$\sigma_1^2 = \ldots = \sigma_k^2 = \sigma^2.$$

As has been demonstrated by Brown and Forsythe (1974), the size of the test is "markedly" different than the desired size. As stated by Rice and Gains (), an ANOVA analysis of the data is most affected when the variances are unequal.

In this chapter, we derive the mathematical properties of the ANOVA procedure under the previously mentioned assumptions. While these results are derived in the literature, we will use this chapter as an outline for our method for comparing k

means without the assumption of equal variances.

4.2 Sources of Variability

The method used to analyze the data begins by examining the statistic

$$SST = \sum\nolimits_{i=1}^{k} \sum\nolimits_{j=1}^{n_i} \left(Y_{i,j} - \overline{Y}_{..}\right)^2$$

that is a measure of the variability in the responses. The statistic SST is referred to as sum of squares total. This variability can be partitioned as $SST = SSE + SSTR$, where

$$SSE = \sum\nolimits_{i=1}^{k} \sum\nolimits_{j=1}^{n_i} \left(Y_{ij} - \overline{Y}_{i.}\right)^2 \text{ and } SSTR = \sum\nolimits_{i=1}^{k} \sum\nolimits_{j=1}^{n_i} \left(\overline{Y}_{i.} - \overline{Y}_{..}\right)^2.$$

The statistics SSE and $SSTR$ are commonly referred to as the sum of squares error and the sum of squares treatment, respectively.

Observe that we can write the $SSTR$ as

$$\begin{aligned} SSTR &= \sum\nolimits_{i=1}^{k} \sum\nolimits_{j=1}^{n_i} \left(\overline{Y}_{i.} - \overline{Y}_{..}\right)^2 = \sum\nolimits_{i=1}^{k} n_i \left(\overline{Y}_{i.} - \overline{Y}_{..}\right)^2 \\ &= \begin{bmatrix} \sqrt{n_1}\left(\overline{Y}_{1.} - \overline{Y}_{..}\right) \\ \sqrt{n_2}\left(\overline{Y}_{2.} - \overline{Y}_{..}\right) \\ \vdots \\ \sqrt{n_k}\left(\overline{Y}_{k.} - \overline{Y}_{..}\right) \end{bmatrix}^{\mathbf{T}} \begin{bmatrix} \sqrt{n_1}\left(\overline{Y}_{1.} - \overline{Y}_{..}\right) \\ \sqrt{n_2}\left(\overline{Y}_{2.} - \overline{Y}_{..}\right) \\ \vdots \\ \sqrt{n_k}\left(\overline{Y}_{k.} - \overline{Y}_{..}\right) \end{bmatrix}. \end{aligned}$$

It is not difficult to show that

$$SSTR = \left(\mathbf{N}^{1/2}\overline{\mathbf{Y}}\right)^{\mathbf{T}} \left(\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}\right) \left(\mathbf{N}^{1/2}\overline{\mathbf{Y}}\right),$$

where $\mathbf{I}$ is a $k \times k$ identity matrix, $\mathbf{J}$ is a $k \times k$ matrix of ones, $m = n_1 + \ldots + n_k$, $\mathbf{N} = Diagonal\left(n_1, \ldots, n_k\right)$; and

$$\overline{\mathbf{Y}} = \left[\overline{Y}_{1.}, \ldots, \overline{Y}_{k.}\right]^{\mathbf{T}}.$$

As one can see, this is the quadratic form of $SSTR$. It was shown in Chapter 2 that the $k \times k$ matrix $\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}$ has $k-1$ eigenvalues that are one and one eigenvalue that is zero. Further, it was shown that the matrix can be expressed as

$$\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2} = \mathbf{V}\mathbf{H}\mathbf{V}^{\mathbf{T}},$$

where the $k \times k$ matrix $\mathbf{H}$ has ones in the first $k-1$ diagonal components and zeroes elsewhere, and the first $k-1$ columns of the $k \times k$ matrix $\mathbf{V}$ are the normalized eigenvectors associated with the $k-1$ eigenvalues of 1 and the k column is the normalized eigenvector associated with the eigenvalue of zero. It is not difficult to show that $\mathbf{H}$ is idempotent.

We can now write

$$SSTR = \left(\mathbf{N}^{1/2}\overline{\mathbf{Y}}\right)^{\mathbf{T}} \mathbf{V}\mathbf{H}\mathbf{V}^{\mathbf{T}} \left(\mathbf{N}^{1/2}\overline{\mathbf{Y}}\right) = \left[\mathbf{H}\mathbf{V}^{\mathbf{T}} \left(\mathbf{N}^{1/2}\overline{\mathbf{Y}}\right)\right]^{\mathbf{T}} \left[\mathbf{H}\mathbf{V}^{\mathbf{T}} \left(\mathbf{N}^{1/2}\overline{\mathbf{Y}}\right)\right].$$

Further, we can express $\mathbf{N}^{1/2}\overline{\mathbf{Y}}$ as

$$\mathbf{N}^{1/2}\overline{\mathbf{Y}} = \sigma \begin{bmatrix} \frac{\overline{Y}_{1.}-\mu_1}{\sigma/\sqrt{n_1}} + \sqrt{n_1}\frac{\mu_1}{\sigma} \\ \frac{\overline{Y}_{2.}-\mu_2}{\sigma/\sqrt{n_2}} + \sqrt{n_2}\frac{\mu_2}{\sigma} \\ \vdots \\ \frac{\overline{Y}_{k.}-\mu_k}{\sigma/\sqrt{n_k}} + \sqrt{n_k}\frac{\mu_k}{\sigma} \end{bmatrix} = \sigma \begin{bmatrix} Z_1 + \sqrt{n_1}\delta_1 \\ Z_2 + \sqrt{n_2}\delta_2 \\ \vdots \\ Z_k + \sqrt{n_k}\delta_k \end{bmatrix} = \sigma\left(\mathbf{Z} + \mathbf{N}^{1/2}\delta\right),$$

where $Z_i = \left(\overline{Y}_{i.} - \mu_i\right) / \left(\sigma/\sqrt{n_1}\right) \sim N(0,1)$, $\delta_i = \mu_i/\sigma$,

$$\mathbf{Z} = [Z_{1.}, \ldots, Z_{k.}]^{\mathbf{T}} \text{ and } \delta = [\delta_{1.}, \ldots, \delta_{k.}]^{\mathbf{T}}.$$

It is not difficult to see that Z_i's are independent. It follows that

$$\begin{aligned} SSTR &= \sigma^2 \left[\mathbf{H}\mathbf{V}^{\mathbf{T}}\left(\mathbf{Z} + \mathbf{N}^{1/2}\delta\right)\right]^{\mathbf{T}} \left[\mathbf{H}\mathbf{V}^{\mathbf{T}}\left(\sigma\left(\mathbf{Z} + \mathbf{N}^{1/2}\delta\right)\right)\right] \\ &= \sigma^2 \left[\left(\mathbf{H}\mathbf{V}^{\mathbf{T}}\mathbf{Z} + \mathbf{H}\mathbf{V}^{\mathbf{T}}\mathbf{N}^{1/2}\delta\right)\right]^{\mathbf{T}} \left[\left(\mathbf{H}\mathbf{V}^{\mathbf{T}}\mathbf{Z} + \mathbf{H}\mathbf{V}^{\mathbf{T}}\mathbf{N}^{1/2}\delta\right)\right]. \end{aligned}$$

The random vector

$$\mathbf{Z}^* = \mathbf{H}\mathbf{V}^{\mathbf{T}}\mathbf{Z} = \left[Z_{1.}^*, \ldots, Z_{k-1.}^*, 0\right]^{\mathbf{T}} \sim N_k(0, \mathbf{H}).$$

Hence, the first $k-1$ components of $\mathbf{Z}^*$ are independent standard normal random variables. Letting

$$\tau = \mathbf{H}\mathbf{V}^{\mathbf{T}}\mathbf{N}^{1/2}\delta,$$

we have

$$SSTR = \sigma^2 \left(\mathbf{Z}^* + \tau\right)^{\mathbf{T}} \left(\mathbf{Z}^* + \tau\right) = \sigma^2 \sum\nolimits_{i=1}^{k-1} \left(Z^*_{i.} + \tau_i\right)^2 .$$

Applying Theorem 3.5 in Chapter 3, we have

$$SSTR \sim \sigma^2 \chi^2_{k-1,\tau^2},$$

where

$$\begin{aligned}
\tau^2 &= \sum\nolimits_{i=1}^{k-1} \tau_i^2 = \tau^{\mathbf{T}}\tau = \left(\mathbf{H}\mathbf{V}^{\mathbf{T}}\mathbf{N}^{1/2}\delta\right)^{\mathbf{T}} \left(\mathbf{H}\mathbf{V}^{\mathbf{T}}\mathbf{N}^{1/2}\delta\right) \\
&= \left(\mathbf{N}^{1/2}\delta\right)^{\mathbf{T}} \left(\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}\right) \left(\mathbf{N}^{1/2}\delta\right) \\
&= \sum\nolimits_{i=1}^{k} n_i \left(\delta_i - \overline{\delta}\right)^2 = \sum\nolimits_{i=1}^{k} \frac{n_i \left(\mu_i - \overline{\mu}\right)^2}{\sigma^2},
\end{aligned}$$

where

$$\overline{\delta} = \frac{1}{m} \sum\nolimits_{i=1}^{k} n_i \delta_i = \frac{1}{m} \sum\nolimits_{i=1}^{k} n_i \mu_i / \sigma = \overline{\mu}/\sigma.$$

Further, we observe that

$$\begin{aligned}
SSE &= \sum\nolimits_{i=1}^{k} \sum\nolimits_{j=1}^{n_i} \left(Y_{ij} - \overline{Y}_{i.}\right)^2 \\
&= \sum_{i=1}^{k} (n_i - 1)\ s_i^2 \\
&= \sum_{i=1}^{k} \sigma^2 \frac{(n_i - 1)\ s_i^2}{\sigma^2}.
\end{aligned}$$

Thus,

$$\frac{SSE}{\sigma^2} \sim \sum_{i=1}^{k} \chi^2_{n_i - 1} \sim \chi^2_{m-k}$$

4.3 The Classical F test

In the previous section, we have derived the distribution of $SSTR/\sigma^2$and SSE/σ^2. It is important to note that SSE and $SSTR$ are independent. Thus,

$$Q = \frac{SSTR/(k-1)}{SSE/(m-k)} \sim \frac{\chi^2_{k-1,\tau^2}/(k-1)}{\chi^2_{m-k}/(m-k)}.$$

The F distribution is a ratio of two independent Chi-squares, each divided by its degrees of freedom (Bain and Engelhardt 1991). Thus, $Q \sim F_{k-1,m-k,\tau^2}$. Under the null hypothesis when all the population means are equal, the non-centrality parameter τ^2 becomes zero. Thus, the distribution of Q becomes a central F distribution with $k-1$ and $m-k$ degrees of freedom. Therefore, p-value $= P(F_{k-1,m-k} \geq q_{obs})$, where q_{obs} is the observed value of the ratio Q. The null hypothesis is rejected when p-value is less than some significance level α.

CHAPTER 5

ANALYSIS WITH NO ASSUMPTION OF THE EQUALITY OF VARIANCES

5.1 Introduction

In this chapter, we use the methods and theorems provided in chapters 2 through 4 to derive the distribution of the likelihood ratio test statistic under the independent normal model without the assumption of equal population variances.

5.2 Distribution of SSE with No Assumption about the Variances

Theorem 5.1: Under the independent normal model with no assumption of equality of the population variances,

$$SSE/\sigma_k^2 \sim \sum\nolimits_{i=1}^{k} \lambda_i^2 \chi_{n_i-1}^2,$$

where $\lambda_i^2 = \sigma_i^2/\sigma_k^2$.

Proof of Theorem 5.1:

$$\begin{aligned} SSE &= \sum\nolimits_{i=1}^{k} \sum\nolimits_{j=1}^{n_i} \left(Y_{ij} - \overline{Y}_{i.}\right)^2 = \sum\nolimits_{i=1}^{k} (n_i - 1)\, s_i^2 \\ &= \sum\nolimits_{i=1}^{k} \sigma_i^2 \, \frac{(n_i - 1)\, S_i^2}{\sigma_i^2} \sim \sum\nolimits_{i=1}^{k} \sigma_i^2 \, \chi_{n_i-1}^2, \end{aligned}$$

since $(n_i - 1)\, S_i^2/\sigma_i^2 \sim \chi_{n_i-1}^2$ for $i = 1, \ldots, k$. By dividing SSE by σ_k^2, the results follows.■

The following theorem gives a closed form expression for the probability density function describing the distribution of SSE/σ_k^2.

Theorem 5.2: Under the independent normal model with no assumption about the population variances with distinct values of $\lambda_i^2 = \sigma_i^2/\sigma_k^2$ for $i = 1, \ldots, k$,

$$f_{SSE/\sigma_k^2}(u) = \sum\nolimits_{r_1=0}^{\infty} \cdots \sum\nolimits_{r_{k-1}=0}^{\infty} \xi_{r_1,\ldots,r_{k-1}} f_{\chi^2_{n_1+\ldots+n_k-k+2r_1+\ldots+2r_{k-1}}}(u),$$

where

$$\xi_{r_1,\ldots,r_{k-1}} = \frac{(-1)^{r_1+\ldots+r_{k-1}}\left(\prod_{i=1}^{k-1}\left(\lambda_{i+1}^2-\lambda_i^2\right)^{r_i}\right)}{\lambda_1^{\nu_1+2r_1}\left(\prod_{i=2}^{k-1}\lambda_i^{n_i-1+2r_i+2r_{i-1}}\right)\left(\prod_{i=1}^{k-1}r_i!\right)} \times \frac{\prod_{i=1}^{k-1}\Gamma\left(\frac{n_1+\ldots+n_i-i+2r_1+\ldots+2r_i}{2}\right)}{\Gamma\left(\frac{n_1-1}{2}\right)\prod_{i=2}^{k-1}\Gamma\left(\frac{n_1+\ldots+n_i-i+2r_1+\ldots+2r_{i-1}}{2}\right)}$$

with $\lambda_i^2 = \sigma_i^2/\sigma_k^2$.

Proof of Theorem 5.2: Subsituting a_i by λ_i^2 and $\nu_i = n_i - 1$ $(i = 1, \ldots, k)$ in Theorem 3.2 gives the result.■

5.3 Distribution of $SSTR$

Theorem 5.3: Under the independent normal model and no assumption about the population variances,

$$SSTR/\sigma_k^2 \sim \begin{cases} \sum_{i=1}^{k-1}\gamma_i\left(\chi_1^2\right)_i, & \text{if } H_0 \text{ holds;} \\ \sum_{i=1}^{k-1}c_i\chi_{1,\Delta_i^2}^2, & \text{if } H_a \text{ holds.}\end{cases},$$

where $\Delta_i = \sum_{j=1}^{k} v_{ji}\mu_j\sqrt{n_j}/\sigma_j$, $\gamma_i = c_i\sigma_k^2$, c_i are the eigenvalues and $\mathbf{v}_i$ are the orthonormal eigenvectors of the matrix $\boldsymbol{\Lambda}\left(\mathbf{I}-\frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}\right)\boldsymbol{\Lambda}$ with $\boldsymbol{\Lambda} = Diagonal(\lambda_1,\ldots,\lambda_{k-1},1)$ and $\lambda_i = \sigma_i/\sigma_k$ for $i = 1, ..., k$.

Proof of Theorem 5.3:

$$\begin{aligned} SSTR &= \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\overline{Y}_{i.}-\overline{Y}_{..}\right)^2 = \sum_{i=1}^{k}n_i\left(\overline{Y}_{i.}-\overline{Y}_{..}\right)^2 \\ &= \begin{bmatrix}\sqrt{n_1}\left(\overline{Y}_{1.}-\overline{Y}_{..}\right)\\ \sqrt{n_2}\left(\overline{Y}_{2.}-\overline{Y}_{..}\right)\\ \vdots \\ \sqrt{n_k}\left(\overline{Y}_{k.}-\overline{Y}_{..}\right)\end{bmatrix}^{\mathbf{T}} \begin{bmatrix}\sqrt{n_1}\left(\overline{Y}_{1.}-\overline{Y}_{..}\right)\\ \sqrt{n_2}\left(\overline{Y}_{2.}-\overline{Y}_{..}\right)\\ \vdots \\ \sqrt{n_k}\left(\overline{Y}_{k.}-\overline{Y}_{..}\right)\end{bmatrix} \\ &= \left(\mathbf{N}^{1/2}\overline{\mathbf{Y}}\right)^{\mathbf{T}}\left(\mathbf{I}-\frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}\right)\left(\mathbf{N}^{1/2}\overline{\mathbf{Y}}\right)\end{aligned}$$

$$\mathbf{N}^{1/2}\overline{\mathbf{Y}} = \begin{bmatrix} \sqrt{n_1}\overline{Y}_{1.} \\ \sqrt{n_2}\overline{Y}_{2.} \\ \vdots \\ \sqrt{n_k}\overline{Y}_{k.} \end{bmatrix} = \begin{bmatrix} \sigma_1\left(\frac{\overline{Y}_{1.}-\mu_1}{\sigma_1/\sqrt{n_1}} + \sqrt{n_1}\frac{\mu_1}{\sigma_1}\right) \\ \sigma_2\left(\frac{\overline{Y}_{2.}-\mu_2}{\sigma_2/\sqrt{n_2}} + \sqrt{n_2}\frac{\mu_2}{\sigma_2}\right) \\ \vdots \\ \sigma_k\left(\frac{\overline{Y}_{k.}-\mu_k}{\sigma_k/\sqrt{n_k}} + \sqrt{n_k}\frac{\mu_k}{\sigma_k}\right) \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1\left(Z_1 + \sqrt{n_1}\frac{\mu_1}{\sigma_1}\right) \\ \sigma_2\left(Z_2 + \sqrt{n_2}\frac{\mu_2}{\sigma_2}\right) \\ \vdots \\ \sigma_k\left(Z_k + \sqrt{n_k}\frac{\mu_k}{\sigma_k}\right) \end{bmatrix}$$

$$= \mathbf{\Sigma}^{1/2}\left(\mathbf{Z} + \mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mu\right)$$

It follows that

$$\begin{aligned} SSTR &= \left(\mathbf{\Sigma}^{1/2}\left(\mathbf{Z} + \mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mu\right)\right)^{\mathbf{T}}\left(\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}\right)\left(\mathbf{\Sigma}^{1/2}\left(\mathbf{Z} + \mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mu\right)\right) \\ &= \left(\mathbf{Z} + \mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mu\right)^{\mathbf{T}}\left[\mathbf{\Sigma}^{1/2}\left(\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}\right)\mathbf{\Sigma}^{1/2}\right]\left(\mathbf{Z} + \mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mu\right). \end{aligned}$$

Then,

$$\frac{SSTR}{\sigma_k^2} = \left(\mathbf{Z} + \mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mu\right)^{\mathbf{T}}\left[\mathbf{\Lambda}\left(\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}\right)\mathbf{\Lambda}\right]\left(\mathbf{Z} + \mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mu\right)$$

where $\mathbf{\Lambda} = diagonal(\lambda_1, \ldots, \lambda_{k-1}, 1)$ with $\lambda_i^2 = \sigma_i^2/\sigma_k^2$.
Using Theorem 2.4, $\mathbf{\Lambda}\left(\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}\right)\mathbf{\Lambda} = \mathbf{V}\mathbf{C}\mathbf{V}^{\mathbf{T}}$, with $\mathbf{C} = Diagonal\left(c_1, \ldots, c_{k-1}, 0\right)$. The diagonal elements of $\mathbf{C}$ are the eigenvalues with corresponding column vectors of $\mathbf{V}$ the normalized orthogonal eigenvectors of the matrix $\mathbf{\Lambda}\left(\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}\right)\mathbf{\Lambda}$.

We can now write $SSTR/\sigma_k^2$ as

$$\begin{aligned} \frac{SSTR}{\sigma_k^2} &= \left(\mathbf{Z} + \mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mu\right)^{\mathbf{T}}\left[\mathbf{V}\mathbf{C}\mathbf{V}^{\mathbf{T}}\right]\left(\mathbf{Z} + \mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mu\right) \\ &= \left[\mathbf{C}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{Z} + \mathbf{C}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mu\right]^{\mathbf{T}}\left[\mathbf{C}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{Z} + \mathbf{C}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mu\right] \end{aligned}$$

With $\mathbf{Z}^* = \mathbf{V^T Z}$, observe that

$$\mathbf{C}^{1/2}\mathbf{V^T Z} = \mathbf{C}^{1/2}\mathbf{Z}^* = \left[c_1^{1/2} Z_1^*, \ldots, c_{k-1}^{1/2} Z_{k-1}^*, 0\right]^{\mathbf{T}}$$

where $\mathbf{Z}^* \sim N_k\left(\mathbf{0}, \mathbf{I}\right)$. For convenience, we define

$$\delta = \mathbf{N}^{1/2}\boldsymbol{\Sigma}^{-1/2}\mu \text{ and } \boldsymbol{\Delta} = \mathbf{V^T N}^{1/2}\boldsymbol{\Sigma}^{-1/2}\mu = \mathbf{V^T}\delta.$$

For clarity, we observed that

$$\frac{SSTR}{\sigma_k^2} = \begin{bmatrix} \sqrt{c_1}\left(Z_1^* + \Delta_1\right) \\ \vdots \\ \sqrt{c_{k-1}}\left(Z_{k-1}^* + \Delta_{k-1}\right) \\ 0 \end{bmatrix}^{\mathbf{T}} \begin{bmatrix} \sqrt{c_1}\left(Z_1^* + \Delta_1\right) \\ \vdots \\ \sqrt{c_{k-1}}\left(Z_{k-1}^* + \Delta_{k-1}\right) \\ 0 \end{bmatrix}$$

with $\Delta_i = \sum_{j=1}^{k} v_{ji}\mu_j\sqrt{n_j}/\sigma_j$ $(i = 1, \ldots, k)$, where v_{ji} is the jth element of the eigenvector $\mathbf{v}_i$.. It then follows that

$$\frac{SSTR}{\sigma_k^2} = \sum\nolimits_{i=1}^{k-1} c_i\left(Z_i^* + \Delta_i\right)^2.$$

As has been shown,

$$\left(Z_i^* + \Delta_i\right)^2 \sim \chi^2_{1,\Delta_i^2}.$$

Hence,

$$\frac{SSTR}{\sigma_k^2} \sim \sum\nolimits_{i=1}^{k-1} c_i\chi^2_{1,\Delta_i^2}.$$

The null hypothesis of equal means can be expressed as

$$H_0 : \mu = \mu\mathbf{1},$$

where μ is an unknown constant and $\mathbf{1}$ is a $k \times 1$ vector of ones. Note that

$$\boldsymbol{\Sigma}^{1/2}\left(\mathbf{I} - \frac{1}{m}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{1/2}\right)\boldsymbol{\Sigma}^{1/2} = \mathbf{VDV^T},$$

where $\mathbf{D} = Diagonal\left(\gamma_1, \ldots, \gamma_{k-1}, 0\right)$ with $\gamma_i = c_i\sigma_k^2$. Using the results of Theorem 2.4, we see that

$$\mathbf{D}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mu = \mu\mathbf{D}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mathbf{1} = \mathbf{0}.$$

Thus, we can write

$$\frac{SSTR}{\sigma_k^2} = \left(\mathbf{D}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{Z}\right)^{\mathbf{T}}\left(\mathbf{D}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{Z}\right) = \sum\nolimits_{i=1}^{k-1}\gamma_i\left(\chi_1^2\right)_i . \blacksquare$$

We can also note from the previous theorem that

$$\begin{aligned}
\sum\nolimits_{i=1}^{k-1} c_i\Delta_i^2 &= \left(\mathbf{C}^{1/2}\Delta\right)^{\mathbf{T}}\left(\mathbf{C}^{1/2}\Delta\right) \\
&= \left(\mathbf{C}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mu\right)^{\mathbf{T}}\left(\mathbf{C}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{N}^{1/2}\mathbf{\Sigma}^{-1/2}\mu\right) \\
&= \left(\mathbf{C}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{\Sigma}^{-1/2}\mathbf{N}^{1/2}\mu\right)^{\mathbf{T}}\left(\mathbf{C}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{\Sigma}^{-1/2}\mathbf{N}^{1/2}\mu\right) \\
&= \left(\mathbf{N}^{1/2}\mu\right)^{\mathbf{T}}\left(\mathbf{C}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{\Sigma}^{-1/2}\right)^{\mathbf{T}}\left(\mathbf{C}^{1/2}\mathbf{V}^{\mathbf{T}}\mathbf{\Sigma}^{-1/2}\right)\left(\mathbf{N}^{1/2}\mu\right) \\
&= \left(\mathbf{N}^{1/2}\mu\right)^{\mathbf{T}}\left(\mathbf{\Sigma}^{-1/2}\left(\mathbf{V}\mathbf{C}\mathbf{V}^{\mathbf{T}}\right)\mathbf{\Sigma}^{-1/2}\right)\left(\mathbf{N}^{1/2}\mu\right) \\
&= \left(\mathbf{N}^{1/2}\mu\right)^{\mathbf{T}}\left(\mathbf{\Sigma}^{-1/2}\left(\mathbf{\Sigma}^{1/2}\left(\mathbf{I} - \frac{1}{n}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{-1/2}\right)\mathbf{\Sigma}^{1/2}\right)\mathbf{\Sigma}^{-1/2}\right)\left(\mathbf{N}^{1/2}\mu\right) \\
&= \left(\mathbf{N}^{1/2}\mu\right)^{\mathbf{T}}\left(\mathbf{I} - \frac{1}{n}\mathbf{N}^{1/2}\mathbf{J}\mathbf{N}^{-1/2}\right)\left(\mathbf{N}^{1/2}\mu\right) \\
&= \sum\nolimits_{i=1}^{k} n_i\left(\mu_i - \overline{\mu}\right)^2,
\end{aligned}$$

where

$$\overline{\mu} = \frac{1}{m}\sum\nolimits_{i=1}^{k} n_i\mu_i.$$

5.4 Distribution of the Likelihood Ratio Test Statistic

In this section, we derive the distribution of the likelihood ratio test statistic under the null and alternative hypotheses.

Theorem 5.4: Under the independent normal model and no assumption about the variances, the distribution of

$$Q = \frac{MSTR}{MSE} = \frac{SSTR/(k-1)}{SSE/(m-k)}$$

is given by

$$f_Q(q) = \sum\nolimits_{\mathbf{t},\mathbf{s}} \sum\nolimits_{\mathbf{r}} \frac{\zeta_{\mathbf{t},\mathbf{s}}\xi_{\mathbf{r}} \left(\frac{k-1}{c_{k-1}(m-k)}\right)^{\nu_{\mathbf{t},\mathbf{s}}/2} \Gamma\left(\frac{\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}}}{2}\right) q^{\nu_{\mathbf{t},\mathbf{s}}/2-1}}{\Gamma\left(\frac{\nu_{\mathbf{t},\mathbf{s}}}{2}\right)\Gamma\left(\frac{\nu_{\mathbf{r}}}{2}\right)\left(1+\frac{k-1}{c_{k-1}(m-k)}q\right)^{(\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}})/2}},$$

where $MSTR = SSTR/(k-1)$, $MSE = SSE/(m-k)$, and

$$\begin{aligned}
\sum\nolimits_{\mathbf{t},\mathbf{s}} &= \sum\nolimits_{t_1=0}^{\infty}\cdots\sum\nolimits_{t_{k-1}=0}^{\infty}\sum\nolimits_{s_1=0}^{\infty}\cdots\sum\nolimits_{s_{k-2}=0}^{\infty};\\
\sum\nolimits_{\mathbf{r}} &= \sum\nolimits_{r_1=0}^{\infty}\cdots\sum\nolimits_{r_{k-1}=0}^{\infty};\\
\nu_{\mathbf{t},\mathbf{s}} &= k-1+2t_1+\ldots+2t_{k-1}+2s_1+\ldots+2s_{k-2};\\
\nu_{\mathbf{r}} &= n_1+\ldots+n_k+2r_1+\ldots+2r_{k-1}-k;\\
\xi_{\mathbf{r}} &= \frac{(-1)^{r_1+\ldots+r_{k-1}}\left(\prod_{i=1}^{k-1}\left(\lambda_{i+1}^2-\lambda_i^2\right)^{r_i}\right)}{\lambda_1^{(n_1-1+2r_1)}\left(\prod_{i=2}^{k-1}\lambda_i^{(n_i-1+2r_i+2r_{i-1})}\right)\left(\prod_{i=1}^{k-1}r_i!\right)}\\
&\times\frac{\prod_{i=1}^{k-1}\Gamma\left(\frac{n_1+\ldots+n_i+2r_1+\ldots+2r_i-i}{2}\right)}{\Gamma\left(\frac{n_1-1}{2}\right)\prod_{i=2}^{k-1}\Gamma\left(\frac{n_1+\ldots+n_i+2r_1+\ldots+2r_{i-1}-i}{2}\right)};\\
\zeta_{\mathbf{t},\mathbf{s}} &= \frac{(\prod_{i=1}^{k-1}\theta_{t_i,\Delta_i^2})(-1)^{s_1+\cdots+s_{k-2}}2^{t_1+\cdots+t_{k-1}}e^{-(\Delta_1^2+\cdots+\Delta_{k-1}^2)/2}}{c_1^{(1+2t_1+2s_1)/2}(\sqrt{\pi})^{k-1}(\prod_{i=1}^{k-1}2t_i!)(\prod_{i=1}^{k-2}s_i!)}\\
&\times\frac{(\prod_{i=1}^{k-2}(c_{i+1}-c_i)^{s_i})\,c_{k-1}^{(k-2+2t_1+\cdots+2t_{k-2}+2s_1+\cdots+2s_{k-3})/2}}{(\prod_{i=2}^{k-2}c_i^{1+2t_i+2s_{i-1}+2s_i})}\\
&\times\frac{(\prod_{i=1}^{k-2}\Gamma(\frac{i+2t_1+\cdots+2t_i+2s_1+\cdots+2s_i}{2}))(\prod_{i=2}^{k-1}\Gamma(\frac{1+2t_i}{2}))}{(\prod_{i=2}^{k-1}\Gamma(\frac{i+2t_1+\cdots+2t_i+2s_1+\cdots+2s_{i-1}}{2}))};\text{ and}\\
\theta_{t_i,\Delta_i^2} &= \begin{cases}1, & \text{if } t_i=0;\\ (\Delta_i^2)^{t_i}, & \text{if } t_i>0.\end{cases}.
\end{aligned}$$

Proof Theorem 5.4: Using Theorems 3.2, 3.6, 3.7, 5.2 and 5.3, and subtituting ν_i by $n_i - 1$, a_i by λ_i^2, τ by Δ, and η by $k-1$, give the result. ■

Corollary 5.4.1: The cumulative distribution function describing the distribution of Q is given as

$$F_Q(q) = \int_0^q f_Q(x)\,dx$$
$$= \sum\nolimits_{\mathbf{t},\mathbf{s}} \sum\nolimits_{\mathbf{r}} \frac{\zeta_{\mathbf{t},\mathbf{s}}\xi_{\mathbf{r}} \left(\frac{k-1}{c_{k-1}(m-k)}\right)^{\nu_{\mathbf{t},\mathbf{s}}/2} \Gamma\left(\frac{\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}}}{2}\right)}{\Gamma\left(\frac{\nu_{\mathbf{t},\mathbf{s}}}{2}\right)\Gamma\left(\frac{\nu_{\mathbf{r}}}{2}\right)}$$
$$\times \int_0^q \frac{x^{\nu_{\mathbf{t},\mathbf{s}}/2-1}}{\left(1+\frac{k-1}{c_{k-1}(m-k)}x\right)^{(\nu_{\mathbf{t},\mathbf{s}}+\nu_{\mathbf{r}})/2}} dx.$$

Proof of Corollary 5.4.1: The proof is straight forward.

Theorem 5.5: Under the independent normal model and no assumption about the variances, and under the null hypothesis that the population means are equal, the probability distribution function describing the distribution of

$$Q = \frac{MSTR}{MSE} = \frac{SSTR/(k-1)}{SSE/(m-k)}$$

is given by

$$f_Q(q) = \sum\nolimits_{\mathbf{s}} \sum\nolimits_{\mathbf{r}} \frac{\zeta_{\mathbf{s}}\xi_{\mathbf{r}} \left(\frac{k-1}{\gamma_{k-1}(m-k)}\right)^{\nu_{\mathbf{s}}/2} \Gamma\left(\frac{\nu_{\mathbf{s}}+\nu_{\mathbf{r}}}{2}\right) q^{\nu_{\mathbf{s}}/2-1}}{\Gamma\left(\frac{\nu_{\mathbf{s}}}{2}\right)\Gamma\left(\frac{\nu_{\mathbf{r}}}{2}\right)\left(1+\frac{k-1}{\gamma_{k-1}(m-k)}q\right)^{(\nu_{\mathbf{s}}+\nu_{\mathbf{r}})/2}},$$

where $MSTR = SSTR/(k-1)$, $MSE = SSE/(m-k)$, and

$$\sum_{\mathbf{s}} = \sum_{s_1=0}^{\infty} \cdots \sum_{s_{k-2}=0}^{\infty};$$

$$\sum_{\mathbf{r}} = \sum_{r_1=0}^{\infty} \cdots \sum_{r_{k-1}=0}^{\infty};$$

$$\nu_{\mathbf{s}} = k - 1 + 2s_1 + \ldots + 2s_{k-2};$$

$$\nu_{\mathbf{r}} = n_1 + \ldots + n_k + 2r_1 + \ldots + 2r_{k-1} - k;$$

$$\xi_{\mathbf{r}} = \frac{(-1)^{r_1+\ldots+r_{k-1}} \left(\prod_{i=1}^{k-1} \left(\lambda_{i+1}^2 - \lambda_i^2\right)^{r_i}\right)}{\lambda_1^{n_1-1+2r_1} \left(\prod_{i=2}^{k-1} \lambda_i^{n_i-1+2r_i+2r_{i-1}}\right) \left(\prod_{i=1}^{k-1} r_i!\right)} \times \frac{\prod_{i=1}^{k-1} \Gamma\left(\frac{n_1+\ldots+n_i+2r_1+\ldots+2r_i-i}{2}\right)}{\Gamma\left(\frac{n_1-1}{2}\right) \prod_{i=2}^{k-1} \Gamma\left(\frac{n_1+\ldots+n_i+2r_1+\ldots+2r_{i-1}-i}{2}\right)};$$

$$\zeta_{\mathbf{s}} = \frac{(-1)^{s_1+\cdots+s_{k-2}} (\prod_{i=1}^{k-2} (\gamma_{i+1} - \gamma_i)^{s_i}) \gamma_{k-1}^{(k-2+2s_1+\cdots+2s_{k-3})/2}}{\gamma_1^{(1+2s_1)/2} (\prod_{i=1}^{k-2} s_i!) (\prod_{i=1}^{k-2} \gamma_i^{1+2s_{i-1}+2s_i})} \times \frac{\prod_{i=1}^{k-2} \Gamma(\frac{i+2s_1+\cdots+2s_i}{2})}{\Gamma(\frac{1}{2}) \left(\prod_{i=2}^{k-2} \Gamma(\frac{i+2s_1+\cdots+2s_{i-1}}{2})\right)}.$$

Proof Theorem 5.5: Using the same procedure as in Theorems 3.7, and using Theorems 5.3, we get the result. ■

Corollary 5.5.1: Under the null hypothesis that the population means are equal, the cumulative distribution function describing the distribution of Q is given as

$$F_Q(q) = \int_0^q f_Q(x)\, dx = \sum_{\mathbf{s}} \sum_{\mathbf{r}} \frac{\zeta_{\mathbf{s}} \xi_{\mathbf{r}} \left(\frac{k-1}{\gamma_{k-1}(m-k)}\right)^{\nu_{\mathbf{s}}/2} \Gamma\left(\frac{\nu_{\mathbf{s}}+\nu_{\mathbf{r}}}{2}\right)}{\Gamma\left(\frac{\nu_{\mathbf{s}}}{2}\right) \Gamma\left(\frac{\nu_{\mathbf{r}}}{2}\right)} \times \int_0^q \frac{x^{\nu_{\mathbf{s}}/2-1}}{\left(1 + \frac{k-1}{\gamma_{k-1}(m-k)} x\right)^{(\nu_{\mathbf{s}}+\nu_{\mathbf{r}})/2}} dx.$$

Proof of Corollary 5.5.1: The proof is straight forward.

The value of $F_Q(q)$ for a given value of q must be obtained numerically as would determining the $100(1-\alpha)$th percentile of the distribution of Q. Under the hypothesis that the population means are equal (null hypothesis), this could only be done for given values of $\lambda_1^2 = \sigma_1^2/\sigma_k^2, \ldots, \lambda_{k-1}^2 = \sigma_{k-1}^2/\sigma_k^2$. On the other hand, the null distribution of Q can be estimated using the data (see Welch (1938)). For example, one could estimate the paramter $\lambda_i^2 = \sigma_i^2/\sigma_k^2$ by estimating the population variance σ_i^2 with the sample variance S_i^2 for $i = 1, \ldots, k$. One could then obtain an estimate of the $100\,(1-\alpha)$th percentile of the distribution of Q as well as an estimate of the p-value of the test for the equality of means.

5.5 Conclusion

Closed form expressions were given for the probability density function describing the distribution of $SSTR/\sigma_k^2$ and SSE/σ_k^2 under the independent normal model with no assumption about the equality of the variances. Further, closed form expressions for the probability density and cumulative distribution functions describing the distribution of the likelihood ratio test statistic $MSTR/MSE$ were derived.

CHAPTER 6

CONCLUSION

6.1 General Conclusions

To conclude, closed form expressions for the probability density function and the cumulative distribution function describing the distribution of the likelihood ratio test statistic, under the independent normal model with no assumption of equality of population variances, are derived. However, values of $F_Q(q)$ for given values of q must be obtained numerically as would determining critical values. An estimate of the p-value of the test for the equality of the population means must also be obtained numerically based on the exact distribution of the likelihood ratio test statistic.

6.2 Areas for Further Research

We intend to develop a MATLAB program that computes an estimate of the $100(1-\alpha)$th percentile of the distribution of Q under the null hypothesis, as well as an estimate of the p-value of the test for the equality of means. Furthermore, we intend to compare our test to other tests found in the literature. This method can also be extended to comparing levels of more than one factor. Furthermore, further research can be made when the population means are ordered or when two population variances are the same.

REFERENCES

[1] Ananda, M. and Weerahandi, S. (1997), "Two-Way ANOVA with Unequal Cell Frequencies and Unequal Variances," *Statistica Sinica* **7**, 631-646.

[2] Barnard, G. A. (1984), "Comparing the Means of Two Independent Samples," *Applied Statistics* **33**, 266-271.

[3] Bishop, T.A. and Dudewicz, E.J. (1978), "Exact Analysis of Variance with Unequal Variances: Test Procedures and Tables," *Technometrics* **20**, 419-430.

[4] Brown, M. B., and A. B. Forsythe. 1974. Robust tests for the equality of variances. *Journal of the American Statistical Association* **69**: 364–367.

[5] Champ, C.W. and Jones-Farmer, L.A. (2007), "Properties of Multivariate Control Chart with Estimated Parameters," *Sequential Analysis* **26(2)**, 153-169.

[6] Champ, C.W. and Hu, F. (2010), " An Exact Expression for the Univariate Behrens-Fisher Distribution," unpublished manuscript.

[7] Champ, C.W. and Rigdon, S.E. (2007), personal communication.

[8] Davis, A.W. (1977) "A Differential Equation Approach to Linear Combinations of Independent Chi-Squares," *Journal of the American Statistical Association* **72**: 212–214.

[9] Dillies, J. and Lakuriqi, E. (2014), personal communication.

[10] Krutchkoff, R. G. (1988), "One-way Fixed Effects Analysis of Variance When the Error Variances May Be Unequal," *Journal of Statistical Computation and Simulation* **30**, 259-271.

[11] Press, S.J. (1966), "Linear Combinations of Non-central Chi-squared Variates," *Annals of Mathematical Statistics* **37(2)**, 480–487.

[12] Rice, W. R. and Gains, S. D. (1989), "One-way Analysis of Variance with Unequal Variances," *Proceedings of the National Academia Sciences* **86**, 8183-8184.

[13] Scheffe, H. (1959), *The Analysis of Variance*, Wiley, New York (pp. 351-358).

[14] Tanizaki, H. (2004), *Computational Methods in Statistics and Econometrics,* Marcel Dekker Inc., New York (pp. 118-119)

[15] Tsui, K. and Weerahandi, S. (1989), "Generalized p-values in Significance Testing of Hypotheses in the Presence of Nuisance Parameters," *Journal of the American Statistical Association* **84**, 602-607.

[16] Wang, Y.Y. (1971), "Probabilities of the Type I Errors of the Welch Tests for the Behrens-Fisher Problem," *Journal of the American Statistical Association* **66**, 605-608.

[17] Weerahandi, S. (1995a). " ANOVA under unequal error variances," *Biometrics* **51**, 589-599.

[18] Welch, B.L. (1938), "The Significance of the Difference Between Two Means When the Population Variances Are Unequal," *Biometrika* **29**, 350–362.

[19] Welch, B.L. (1951), "On the Comparison of Several Mean Values: An Alternative Approach," *Biometrika* **38**, 330-336.

[20] Yiğit E, Gokpinar F. (2010), "A Simulation Study on Tests for One-Way ANOVA Under the Unequal Variance Assumption," *Communications Series A1 Mathematics & Statistics* **59**(2), 15-34.