

Summer 2010

Methods for Comparing Two Means with Application in Adaptive Clinical Trials

Fengjiao Hu

Follow this and additional works at: <https://digitalcommons.georgiasouthern.edu/etd>

Recommended Citation

Hu, Fengjiao, "Methods for Comparing Two Means with Application in Adaptive Clinical Trials" (2010). *Electronic Theses and Dissertations*. 660.
<https://digitalcommons.georgiasouthern.edu/etd/660>

This thesis (open access) is brought to you for free and open access by the Jack N. Averitt College of Graduate Studies at Georgia Southern Commons. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Georgia Southern Commons. For more information, please contact digitalcommons@georgiasouthern.edu.

Version: July 21, 2010

**METHODS FOR COMPARING TWO MEANS WITH APPLICATION
IN ADAPTIVE CLINICAL TRIALS**

by

FENGJIAO HU

(Under the Direction of Charles W. Champ)

ABSTRACT

In the design of a clinical trial, the study of the effect of an intervention for a given medical condition is frequently of interest to researcher. Also, in recent years, the use of sequential and adaptive design methods in clinical research and development based on accrued data has become very popular due to its flexibility and efficiency. In this thesis, we derive the Behrens-Fisher distribution, and use the distributional result to examine the effect of an intervention by comparing population means of intervention group and control group. Sample size prediction methods propoing to solve the Behrens-Fisher problem are examined. A new method for solving the Behrens-Fisher problem is proposed. Various sequential and adaptive designs are reviewed.

Key Words: Independent samples, normal model, paired data, random sample, sequential and adaptive methods

2009 Mathematics Subject Classification:

**METHODS FOR COMPARING TWO MEANS WITH APPLICATION
IN ADAPTIVE CLINICAL TRIALS**

by

FENGJIAO HU

B.S. in Huazhong Normal University, China, 2008

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

MASTER OF SCIENCE
IN MATHEMATICS

STATESBORO, GEORGIA

2010

©2010

FENGJIAO HU

All Rights Reserved

**METHODS FOR COMPARING TWO MEANS WITH APPLICATION
IN ADAPTIVE CLINICAL TRIALS**

by

FENGJIAO HU

Major Professor: Charles W. Champ

Committee: Hani Samawi

Broderick Oluyede

Electronic Version Approved:

July 2010

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Charles W. Champ for his insightful guidance and constant encouragement. Without his ideas and help I could not have done this thesis. I would also like to thank Dr. Broderick Oluyede and Dr. Hani Samawi for their suggestions and willingness to be on my committee. I also would like to thank my parents whose encouragement motivated me to achieve my goals in education and in life. I applaud the faculty, staff and fellow graduates of the Department of Mathematics for their support.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	v
LIST OF TABLES	ix
LIST OF FIGURES	x
CHAPTER	
1 INTRODUCTION	1
2 SOME DISTRIBUTIONAL RESULTS	2
2.1 Introduction	2
2.2 Central and Noncentral t -distributions	3
2.3 Distributions of T_1 and T_2	5
2.4 Welsh's Approximation/Estimation Method	13
2.5 Conclusion	15
3 COMPARING TWO POPULATION MEANS: INDEPENDENT SAMPLES CASE	16
3.1 Introduction	16

3.2	Variances Known Case	19
3.3	Unknown Variances	24
3.4	Conclusion	27
4	COMPARING TWO MEANS: PAIRED DATA CASE	28
4.1	Introduction	28
4.2	Inference about μ_D , Variance σ_D^2 Known	29
4.3	Variance σ_D^2 Unknown	31
4.4	Conclusion	32
5	SEQUENTIAL AND ADAPTIVE METHODS	33
5.1	Introduction	33
5.2	Two-Stage Sample Prediction Methods	34
5.3	Paired Data	51
5.4	Multistage Stage Adaptive Methods	52
5.5	Group Sequential Methods	55
5.6	Conclusion	64
6	CONCLUSION	65

6.1	General Conclusions	65
6.2	Areas for Further Research	66
	REFERENCES	68
A	Appendix I	71
B	Appendix II	76
C	Appendix III	82

LIST OF TABLES

Table		Page
3.1	Sample Size Selection Cases	17
3.2	Test Statistics	18
5.1	Behrens-Fisher Procedures	39
5.2	Group Sequential Tests	57

LIST OF FIGURES

Figure		Page
2.1	Central($\nu = 20$) and Noncentral($\nu = 20, \theta = 2.0$) t distribution . . .	4
3.1	$d = \frac{1}{\sqrt{\left[\frac{\lambda^2}{1+\lambda^{10}}\right] + \left[\frac{1}{1+\lambda^{10}}\right]}} - \frac{1}{\sqrt{\left[\frac{\lambda^2}{1+\lambda^{10}}\right] + \left[\frac{1}{1+\lambda^{10}}\right]}}. \dots\dots\dots$	22

CHAPTER 1

INTRODUCTION

The comparison of the means of two populations is frequently of interest to researchers. A variety of statistical methods have been proposed to compare two means. Each of these methods has been developed around the assumptions made about the data to be used in the analysis. In Chapter 2, we discuss some of these modeling assumptions and some relevant distributional results. In particular, we derive the Behrens-Fisher distribution along with another commonly used test statistic, we show that their cumulative distribution (probability density) functions can be expressed as linear combinations of non-central t -distributions. When the data are available in the form of two independent samples, a new method making use of the derived cdf of the Behrens-Fisher distribution is presented in Chapter 3 for comparing two population means when the ratio of the variances is known. When the ratio of the variances is known, a FORTRAN program is given for determining the size and power of the test and the approximation introduced by Welsh(1938).

The “paired data” problem is examined in Chapter 4. A method is given for determining the appropriate sample size. These results depend on the ratio of the means and the variances of the distribution of differences. Various sequential and adaptive methods are given in Chapter 5 for the “paired data” and “independent sample” cases. Two-stage sample size prediction methods are reviewed that purport to provide solutions to the Behrens-Fisher problem. A new method is proposed for solving the Behrens-Fisher problem making use of the Behrens-Fisher distribution derived in Chapter 2. In the last chapter, some conclusions are expressed along with several areas of further research.

CHAPTER 2

SOME DISTRIBUTIONAL RESULTS

2.1 Introduction

The Behrens-Fisher problem is a well known problem in statistics. It is concerned with an interval estimation and testing hypotheses about the difference between two population means when no assumption is made about the equality of the unknown variances. Kim and Cohen (1998) stated that “although a number of methods have been proposed for the Behrens-Fisher problem ... , no definite solutions exists” On the other hand, Dudewicz, et al. (2007) stated that “this problem has three known exact solutions” They claimed that the solutions are due to Chapman (1950), Prokof'yev and Shishkin (1974), and Dudewicz and Ahmed (1998,1999).

In this chapter, we derive the distribution of the statistics

$$T_1 = \frac{\bar{X}_I - \bar{X}_C}{\sqrt{\frac{(n_I-1)S_I^2 + (n_C-1)S_C^2}{n_I+n_C-2}} (1/n_I + 1/n_C)} \text{ and } T_2 = \frac{\bar{X}_I - \bar{X}_C}{\sqrt{S_I^2/n_I + S_C^2/n_C}}$$

under the independent normal model. According to Kim and Cohen (1998), the distribution of the statistic T_2 is referred to as the Behrens-Fisher distribution. Here, \bar{X}_I, \bar{X}_C and S_I^2, S_C^2 are the means and variances of two independent random sample of size n_I, n_C from $N(\mu_I, \sigma_I^2)$ and $N(\mu_C, \sigma_C^2)$ respectively. This model will be referred to as the independent normal model. In this thesis, we do not provide a solution to the Behrens-Fisher problem but give the exact distribution of the statistic T_2 under the independent normal model. The cumulative distribution functions (cdfs) of T_1 and T_2 are shown to be linear combinations of the cdfs of noncentral t -distributions. The exact probability density functions (pdfs) can then be determined from the cdfs. Nel, et al (1990) provided an exact solution to the pdf of T_2 in terms of a hypergeometric

function.

In the next section, we give some useful results about the noncentral t -distribution. The Behrens-Fisher distribution is derived in Section 2.3. Section 2.4 contains Welsh's approximation methods. Some concluding remarks are made in the Section 2.5.

2.2 Central and Noncentral t -distributions

It is useful at this point to examine the central and noncentral t -distributions. The random variable T defined by

$$T = \frac{Z + \theta}{\sqrt{W/\nu}}$$

with $Z \sim N(0, 1)$ and $W \sim \chi_\nu^2$ are independent. T is said to have a noncentral t distribution with $\nu > 0$ degrees of freedom and noncentrality parameter θ . If $\theta = 0$, the distribution of T is known as the central t -distribution. We write $T = t_{\nu, \theta}$ or $T \sim t_{\nu, \theta}$.

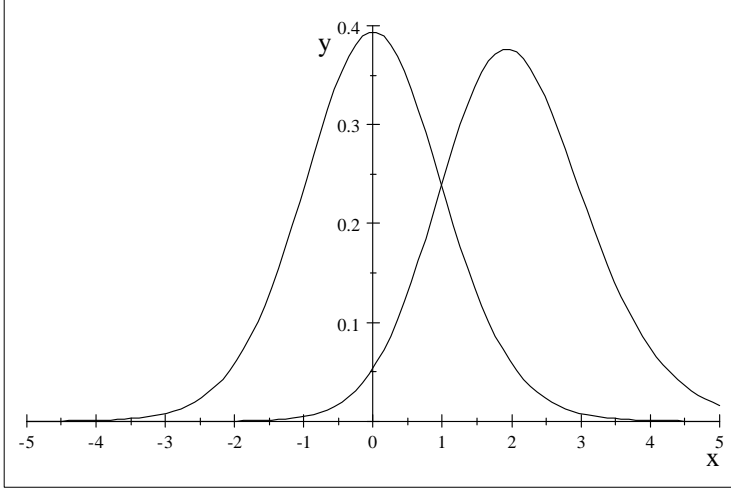
Theorem 2.2.1. *The probability density function $f_T(t|\nu, \theta)$ of a t -distribution is given by*

$$f_T(t|\nu, \theta) = \frac{\Gamma\left(\frac{\nu+1}{2}\right) e^{-\theta^2/2}}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right) (1+t^2/\nu)^{(\nu+1)/2}} \left(1 + \sum_{i=1}^{\infty} \frac{\Gamma\left(\frac{\nu+i+1}{2}\right) (\theta\sqrt{2}t)^i}{\Gamma\left(\frac{\nu+1}{2}\right) (\nu+t^2)^{i/2} i!}\right),$$

where $\nu > 0$ is the degrees of freedom and θ is the noncentrality parameter.

These results for $\theta = 0$ are presented in Bain and Engelhardt (1992) and for $\theta \neq 0$ can be found in Evans, et al. (1993).

In Figure 2.1, the left most curve is the density function of a central t -distribution with $\nu = 20$. The right most curve in this figure is that of a noncentral t -distribution with $\nu = 20$ and $\theta = 2.0$.



Central ($\nu = 20$) and Noncentral ($\nu = 20, \theta = 2.0$) t -Distributions

Figure 2.1: Central($\nu = 20$) and Noncentral($\nu = 20, \theta = 2.0$) t distribution

{f2.1}

The cumulative distribution function (cdf) describing the distribution of T can be expressed by

$$\begin{aligned} F_T(t|\nu, \theta) &= F_{t_{\nu, \theta}}(t) = \int_0^{\infty} \Phi\left(t\sqrt{w/\nu} - \theta\right) f_W(w) dw \\ &= \int_{-\infty}^t \frac{\Gamma\left(\frac{\nu+1}{2}\right) e^{-\theta^2/2}}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right) (1+w^2/\nu)^{(\nu+1)/2}} \\ &\quad \times \left(1 + \sum_{i=1}^{\infty} \frac{\Gamma\left(\frac{\nu+i+1}{2}\right) (\theta\sqrt{2}w)^i}{\Gamma\left(\frac{\nu+1}{2}\right) (\nu+w^2)^{i/2} i!}\right) dw, \end{aligned}$$

where $\Phi(z)$ is the cdf of a standard normal distribution. (See Evans, et al. (1993)).

Benton and Drishnamoorthy (2003) gave the cdf of T as

$$F_T(t|\nu, \theta) = \Phi(-\theta) + \frac{1}{2} \sum_{i=0}^{\infty} \left(P_i I_x\left(i + \frac{1}{2}, \frac{\nu}{2}\right) + \frac{\theta}{\sqrt{2}} Q_i I_x\left(i + 1, \frac{\nu}{2}\right) \right)$$

with

$$P_i = \frac{(\theta^2/2)^i}{i!} e^{-\theta^2/2}, \quad Q_i = \frac{(\theta^2/2)^i}{\Gamma(i + 3/2)} e^{-\theta^2/2}, \quad \text{and } x = \frac{t^2}{\nu + t^2},$$

where

$$I_x(a, b) = \int_0^x \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1} (1-y)^{b-1} dy \text{ and } \Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy.$$

The function $I_x(a, b)$ is the incomplete Beta function.

2.3 Distributions of T_1 and T_2

The following theorem will be useful in determining the distributions of T_1 and T_2 .

Theorem 2.3.1. *If $Z \sim N(0, 1)$, $W_I \sim \chi_{2a}^2$, and $W_C \sim \chi_{2b}^2$ are independent and $\xi, \nu > 0$, then T defined by*

$$T = \frac{Z + \theta}{\sqrt{(\xi W_I + W_C)/\nu}}$$

has cdf

$$\begin{aligned} F_T(t) &= \xi^b F_{t_{2(a+b), \theta}} \left(t \sqrt{(\xi/\nu)(2(a+b))} \right) \\ &\quad + \sum_{k=1}^{\infty} \frac{\xi^b \Gamma(b+k) (1-\xi)^k}{\Gamma(b) k!} F_{t_{2(a+b+k), \theta}} \left(t \sqrt{(\xi/\nu)(2(a+b+k))} \right) \end{aligned}$$

and pdf

$$\begin{aligned} f_T(t) &= \xi^b \sqrt{(\xi/\nu)(2(a+b))} f_{t_{2(a+b), \theta}} \left(t \sqrt{(\xi/\nu)(2(a+b))} \right) \\ &\quad + \sum_{k=1}^{\infty} \frac{\xi^b \Gamma(b+k) (1-\xi)^k \sqrt{(\xi/\nu)(2(a+b+k))}}{\Gamma(b) k!} \\ &\quad \times f_{t_{2(a+b+k), \theta}} \left(t \sqrt{(\xi/\nu)(2(a+b+k))} \right), \end{aligned}$$

where $F_{t_{q, \theta}}(t)$ and $f_{t_{q, \theta}}(t)$ are the cdf and pdf of a noncentral t -distribution with q degrees of freedom and noncentrality parameter θ , respectively.

Proof. We begin by examining the distribution of

$$W = (\xi W_I + W_C) / \nu.$$

For the case in which $\xi = 1$, it is well known that $W \sim \chi_{2(a+b)}^2 / \nu$. Consider now the case in which $\xi \neq 1$. Define the linear transformation

$$Y_1 = \xi W_I + W_C \text{ and } Y_2 = W_C.$$

The inverse transformation is

$$W_I = \xi^{-1} (Y_1 - Y_2) \text{ and } W_C = Y_2$$

with Jacobian ξ^{-1} . The joint probability density function of Y_1 and Y_2 is given by

$$f_{Y_1, Y_2}(y_1, y_2) = f_{W_I}(\xi^{-1}(y_1 - y_2)) f_{W_C}(y_2) \xi^{-1}.$$

Since $W_I \sim \chi_{n_I-1}^2$ and $W_C \sim \chi_{n_C-1}^2$, then

$$\begin{aligned} f_{Y_1, Y_2}(y_1, y_2) &= \xi^{-1} \frac{1}{\Gamma\left(\frac{n_I-1}{2}\right) 2^{(n_I-1)/2}} (\xi^{-1}(y_1 - y_2))^{(n_I-1)/2-1} e^{-\xi^{-1}(y_1-y_2)/2} \\ &\quad \times \frac{1}{\Gamma\left(\frac{n_C-1}{2}\right) 2^{(n_C-1)/2}} (y_2)^{(n_C-1)/2-1} e^{-y_2/2} I_{\{(y_1, y_2) | 0 < y_2 \leq y_1\}}(y_1, y_2). \end{aligned}$$

For convenience, we let $a = (n_I - 1) / 2$ and $b = (n_C - 1) / 2$. Using these notations, we have

$$\begin{aligned} f_{Y_1, Y_2}(y_1, y_2) &= \frac{(\xi^{-1})^a}{\Gamma(a) 2^a} (y_1 - y_2)^{a-1} e^{-\xi^{-1}(y_1-y_2)/2} \\ &\quad \times \frac{1}{\Gamma(b) 2^b} (y_2)^{b-1} e^{-y_2/2} I_{\{(y_1, y_2) | 0 < y_2 \leq y_1\}}(y_1, y_2). \end{aligned}$$

Furthermore, note that we can express this joint density for $0 < y_2 \leq y_1$ as

$$\begin{aligned} f_{Y_1, Y_2}(y_1, y_2) &= \frac{(\xi^{-1})^a}{\Gamma(a+b) 2^{a+b}} y_1^{a+b-2} e^{-y_1/(2\xi)} \\ &\quad \times \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \left(\frac{y_2}{y_1}\right)^{b-1} \left(1 - \frac{y_2}{y_1}\right)^{a-1} e^{-y_1(1-\xi^{-1})(y_2/y_1)/2}. \end{aligned}$$

It is also convenient to let $c = y_1 (\xi^{-1} - 1)$. Thus, we have

$$\begin{aligned} f_{Y_1, Y_2}(y_1, y_2) &= \frac{(\xi^{-1})^a}{\Gamma(a+b) 2^{a+b}} y_1^{a+b-2} e^{-y_1/(2\xi)} \\ &\quad \times \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \left(\frac{y_2}{y_1}\right)^{b-1} \left(1 - \frac{y_2}{y_1}\right)^{a-1} e^{c(y_2/y_1)/2}. \end{aligned}$$

Making the transformation $u = y_2/y_1$, the marginal distribution of Y_1 can be expressed as

$$\begin{aligned} f_{Y_1}(y_1) &= (\xi^{-1})^a \frac{1}{\Gamma(a+b) 2^{a+b}} y_1^{a+b-1} e^{-y_1/(2\xi)} \\ &\quad \times \int_0^1 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} u^{b-1} (1-u)^{a-1} e^{cu/2} du. \end{aligned}$$

Expanding the function $e^{cu/2}$, we have

$$\begin{aligned} f_{Y_1}(y_1) &= (\xi^{-1})^a \frac{1}{\Gamma(a+b) 2^{a+b}} y_1^{a+b-1} e^{-y_1/(2\xi)} \\ &\quad \times \int_0^1 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} u^{b-1} (1-u)^{a-1} \left(\sum_{k=0}^{\infty} \frac{(cu/2)^k}{k!} \right) du \\ &= (\xi^{-1})^a \frac{1}{\Gamma(a+b) 2^{a+b}} y_1^{a+b-1} e^{-y_1/(2\xi)} \\ &\quad \times \sum_{k=0}^{\infty} \frac{1}{k! 2^k} \frac{\Gamma(a+b)\Gamma(b+k)}{\Gamma(b)\Gamma(a+b+k)} c^k \\ &\quad \times \int_0^1 \frac{\Gamma(a+b+k)}{\Gamma(a)\Gamma(b+k)} u^{b+k-1} (1-u)^{a-1} du. \end{aligned}$$

Since

$$\int_0^1 \frac{\Gamma(a+b+k)}{\Gamma(a)\Gamma(b+k)} u^{b+k-1} (1-u)^{a-1} du = 1,$$

then

$$\begin{aligned} f_{Y_1}(y_1) &= (\xi^{-1})^a \frac{1}{\Gamma(a+b) 2^{a+b}} y_1^{a+b-1} e^{-y_1/(2\xi)} \\ &\quad \times \sum_{k=0}^{\infty} \frac{\Gamma(a+b)\Gamma(b+k)}{\Gamma(b)\Gamma(a+b+k) k! 2^k} c^k. \end{aligned}$$

It is useful to express the marginal density of Y_1 in the form

$$\begin{aligned} f_{Y_1}(y_1) &= \frac{(\xi^{-1})^a}{\Gamma(a+b) 2^{a+b}} y_1^{a+b-1} e^{-y_1/(2\xi)} \\ &\quad + \sum_{k=1}^{\infty} \frac{\Gamma(b+k) (\xi^{-1})^a (\xi^{-1} - 1)^k}{\Gamma(b) k!} \\ &\quad \times \frac{1}{\Gamma(a+b+k) 2^{a+b+k}} y_1^{a+b+k-1} e^{-y_1/(2\xi)}. \end{aligned}$$

After some rearrangement of the terms, we have

$$\begin{aligned} f_{Y_1}(y_1) &= \frac{\xi^b}{\Gamma(a+b) (2\xi)^{a+b}} y_1^{a+b-1} e^{-y_1/(2\xi)} \\ &\quad + \xi^b \sum_{k=1}^{\infty} \frac{\Gamma(b+k) (1-\xi)^k}{\Gamma(b) k!} \frac{1}{\Gamma(a+b+k) (2\xi)^{a+b+k}} y_1^{a+b+k-1} e^{-y_1/(2\xi)}. \end{aligned}$$

We see that the density of Y_1 can be expressed as

$$f_{Y_1}(y_1) = \xi^b \left(g(y_1 | 2\xi, a+b) + \sum_{k=1}^{\infty} \frac{\Gamma(b+k) (1-\xi)^k}{\Gamma(b) k!} g(y_1 | 2\xi, a+b+k) \right),$$

where $g(y_1 | 2\xi, a+b+k)$ is the probability density function of a gamma distribution with scale parameter 2ξ and shape parameter $a+b+k$. Since $W = Y_1/\nu$, then for $w > 0$, we have

$$\begin{aligned} f_W(w) &= \nu f_{Y_1}(\nu w) = \frac{\nu \xi^b}{\Gamma(a+b) (2\xi)^{a+b}} (\nu w)^{a+b-1} e^{-\nu w/(2\xi)} \\ &\quad + \nu \xi^b \sum_{k=1}^{\infty} \frac{\Gamma(b+k) (1-\xi)^k}{\Gamma(b) k!} \\ &\quad \times \frac{1}{\Gamma(a+b+k) (2\xi)^{a+b+k}} (\nu w)^{a+b+k-1} e^{-\nu w/(2\xi)} \\ &= \xi^b \frac{1}{\Gamma(a+b) (2\xi/\nu)^{a+b}} w^{a+b-1} e^{-w/(2\xi/\nu)} \\ &\quad + \xi^b \sum_{k=1}^{\infty} \frac{\Gamma(b+k) (1-\xi)^k}{\Gamma(b) k!} \frac{1}{\Gamma(a+b+k) (2\xi/\nu)^{a+b+k}} w^{a+b+k-1} e^{-w/(2\xi/\nu)} \\ &= \xi^b \left(g(w | 2\xi/\nu, a+b) + \sum_{k=1}^{\infty} \frac{\Gamma(b+k) (1-\xi)^k}{\Gamma(b) k!} g(w | 2\xi/\nu, a+b+k) \right). \end{aligned}$$

Now consider the distribution of

$$T = \frac{Z + \theta}{\sqrt{(\xi W_I + W_C)/\nu}} = \frac{Z + \theta}{\sqrt{W}}.$$

We have that

$$\begin{aligned} F_T(t) &= P\left(\frac{Z + \theta}{\sqrt{W}} \leq t\right) = P\left(Z \leq t\sqrt{W} - \theta\right) \\ &= \int_0^\infty \Phi(t\sqrt{w} - \theta) f_W(w) dw \\ &= \xi^b \int_0^\infty \Phi(t\sqrt{w} - \theta) \frac{1}{\Gamma(a+b)(2\xi/\nu)^{a+b}} w^{a+b-1} e^{-w/(2\xi/\nu)} dw \\ &\quad + \xi^b \sum_{k=1}^\infty \frac{\Gamma(b+k)(1-\xi)^k}{\Gamma(b)k!} \\ &\quad \times \int_0^\infty \Phi(t\sqrt{w} - \theta) \frac{1}{\Gamma(a+b+k)(2\xi/\nu)^{a+b+k}} w^{a+b+k-1} e^{-w/(2\xi/\nu)} dw. \end{aligned}$$

Consider now the transformation $Y = \nu W/\xi$, we have

$$\begin{aligned} F_T(t) &= \xi^b \int_0^\infty \Phi\left(t\sqrt{\xi y/\nu} - \theta\right) \frac{1}{\Gamma(a+b)(2\xi/\nu)^{a+b}} \left(\frac{\xi y}{\nu}\right)^{a+b-1} e^{-y/2\frac{\xi}{\nu}} dy \\ &\quad + \xi^b \sum_{k=1}^\infty \frac{\Gamma(b+k)(1-\xi)^k}{\Gamma(b)k!} \\ &\quad \times \int_0^\infty \Phi\left(t\sqrt{\xi y/\nu} - \theta\right) \frac{1}{\Gamma(a+b+k)(2\xi/\nu)^{a+b+k}} \left(\frac{\xi y}{\nu}\right)^{a+b+k-1} e^{-y/2\frac{\xi}{\nu}} dy \\ &= \xi^b \int_0^\infty \Phi\left(t\sqrt{(\xi/\nu)(2(a+b))}\sqrt{y/(2(a+b))} - \theta\right) \\ &\quad \times \frac{1}{\Gamma(a+b)2^{a+b}} y^{a+b-1} e^{-y/2} dy \\ &\quad + \sum_{k=1}^\infty \frac{\xi^b \Gamma(b+k)(1-\xi)^k}{\Gamma(b)k!} \\ &\quad \times \int_0^\infty \Phi\left(t\sqrt{(\xi/\nu)(2(a+b+k))}\sqrt{y/(2(a+b+k))} - \theta\right) \\ &\quad \times \frac{1}{\Gamma(a+b+k)2^{a+b+k}} y^{a+b+k-1} e^{-y/2} dy. \end{aligned}$$

It follows from the results in the previous section that $F_T(t)$ can be expressed as

$$\begin{aligned}
F_T(t) &= \xi^b F_{t_{2(a+b),\theta}} \left(t\sqrt{(\xi/\nu)(2(a+b))} \right) \\
&\quad + \sum_{k=1}^{\infty} \frac{\xi^b \Gamma(b+k) (1-\xi)^k}{\Gamma(b) k!} F_{t_{2(a+b+k),\theta}} \left(t\sqrt{(\xi/\nu)(2(a+b+k))} \right) \\
&= \xi^b F_{t_{2(a+b),\theta}} \left(t\sqrt{(\xi/\nu)(2(a+b))} \right) \\
&\quad + \sum_{k=1}^{\infty} \frac{(-1)^k \xi^{b+k} \Gamma(b+k) (1-\xi)^k (-\xi^{-1})^k}{\Gamma(b) k!} \\
&\quad \times F_{t_{2(a+b+k),\theta}} \left(t\sqrt{(\xi/\nu)(2(a+b+k))} \right) \\
&= \xi^b F_{t_{2(a+b),\theta}} \left(t\sqrt{(\xi/\nu)(2(a+b))} \right) \\
&\quad + \sum_{k=1}^{\infty} \frac{(-1)^k \xi^{b+k} \Gamma(b+k) (1-\xi^{-1})^k}{\Gamma(b) k!} \\
&\quad \times F_{t_{2(a+b+k),\theta}} \left(t\sqrt{(\xi/\nu)(2(a+b+k))} \right).
\end{aligned}$$

Hence, the probability density function of T has the form

$$\begin{aligned}
f_T(t) &= \xi^b \sqrt{(\xi/\nu)(2(a+b))} f_{t_{2(a+b),\theta}} \left(t\sqrt{(\xi/\nu)(2(a+b))} \right) \\
&\quad + \sum_{k=1}^{\infty} \frac{\xi^b \Gamma(b+k) (1-\xi)^k \sqrt{(\xi/\nu)(2(a+b+k))}}{\Gamma(b) k!} \\
&\quad \times f_{t_{2(a+b+k),\theta}} \left(t\sqrt{(\xi/\nu)(2(a+b+k))} \right).
\end{aligned}$$

□

It is interesting to note that

$$\xi^b \left(1 + \sum_{k=1}^{\infty} \frac{\Gamma(b+k) (1-\xi)^k}{\Gamma(b) k!} \right) = 1.$$

Thus since

$$\begin{aligned}
1 + \sum_{k=1}^{\infty} \frac{\Gamma(b+k)(1-\xi)^k}{\Gamma(b)k!} &= \sum_{k=0}^{\infty} \frac{\Gamma(b+k)(1-\xi)^k}{\Gamma(b)k!} \\
&= \sum_{k=0}^{\infty} \frac{(b+k-1)(b+k-2)\cdots(b)(1-\xi)^k}{k!} \\
&= \sum_{k=0}^{\infty} \binom{b+k-1}{k} (1-\xi)^k \\
&= \sum_{k=0}^{\infty} (-1)^k \binom{-b}{k} (1-\xi)^k \\
&= \sum_{k=0}^{\infty} \binom{-b}{k} (\xi-1)^k \\
&= \xi^{-b}.
\end{aligned}$$

It can be shown that

$$T_1 = \frac{Z + \theta}{\sqrt{(\xi_1 W_I + W_C)/\nu_1}} \text{ and } T_2 = \frac{Z + \theta}{\sqrt{(\xi_2 W_I + W_C)/\nu_2}},$$

where

$$\begin{aligned}
\theta &= \frac{\mu_I - \mu_C}{\sqrt{\sigma_I^2/n_I + \sigma_C^2/n_C}}, \quad \lambda^2 = \frac{\sigma_I^2}{\sigma_C^2}, \\
\nu_1 &= \frac{(n_I + n_C - 2)(\lambda^2/n_I + 1/n_C)}{1/n_I + 1/n_C}, \quad \xi_1 = \lambda^2, \\
\nu_2 &= n_C(n_C - 1)(\lambda^2/n_I + 1/n_C), \quad \xi_2 = \lambda^2 \frac{n_C(n_C - 1)}{n_I(n_I - 1)}, \\
W_I &= \frac{(n_I - 1)S_I^2}{\sigma_I^2} \sim \chi_{n_I-1}^2, \text{ and } W_C = \frac{(n_C - 1)S_C^2}{\sigma_C^2} \sim \chi_{n_C-1}^2.
\end{aligned}$$

Hence, the distributions of T_1 and T_2 follow from Theorem 1, keeping in mind that

$$a = \frac{n_I - 1}{2} \text{ and } b = \frac{n_C - 1}{2}.$$

It is observed that the distributions of T_1 and T_2 depend in general on the values $\mu_I - \mu_C$, σ_I^2 , σ_C^2 , n_I , and n_C . For the case in which $\mu_I = \mu_C$, then their distributions

depend only on the parameter λ^2 and the sample sizes n_I and n_C . Also, we note that $F_T(t)$ can be thought of as a function of θ , ξ , ν , n_I , and n_C . We observe that

$$\begin{aligned}
F_T(t) &= F_T(t|\theta, \xi, \nu, n_I, n_C) \\
&= P\left(\frac{Z + \theta}{\sqrt{(\xi W_I + W_C)/\nu}} \leq t\right) \\
&= P\left(\frac{Z + \theta}{\sqrt{(\xi^{-1}W_C + W_I)/(\xi^{-1}\nu)}} \leq t\right) \\
&= F_T(t|\theta, \xi^{-1}, \xi^{-1}\nu, n_C, n_I).
\end{aligned}$$

Further we note that $F_T(t)$ can be thought of as a function of δ_C , λ , n_I , and n_C . For the distribution of T_1 , we see that

$$\begin{aligned}
F_{T_1}(t) &= F_{T_1}(t|\delta_C, \lambda, n_I, n_C) \\
&= P\left(\frac{Z + \delta_C/\sqrt{\lambda^2/n_I + 1/n_C}}{\sqrt{\frac{n_I(n_I-1)(1/n_I+1/n_C)(\lambda^2/n_I)W_I}{(n_I+n_C-2)(\lambda^2/n_I+1/n_C)(n_I-1)} + \frac{n_C(n_C-1)(1/n_I+1/n_C)(1/n_C)W_C}{(n_I+n_C-2)(\lambda^2/n_I+1/n_C)(n_C-1)}}} \leq t\right) \\
&= P\left(\frac{Z + \lambda^{-1}\delta_C/\sqrt{\lambda^{-2}/n_C + 1/n_I}}{\sqrt{\frac{n_C(n_C-1)(1/n_I+1/n_C)(\lambda^{-2}/n_C)W_C}{(n_I+n_C-2)(\lambda^{-2}/n_C+1/n_I)(n_C-1)} + \frac{n_I(n_I-1)(1/n_I+1/n_C)(1/n_I)W_I}{(n_I+n_C-2)(\lambda^{-2}/n_C+1/n_I)(n_I-1)}}} \leq t\right) \\
&= F_{T_1}(t|\lambda^{-1}\delta_C, \lambda^{-1}, n_C, n_I).
\end{aligned}$$

Similarly, the distribution of T_2 is

$$\begin{aligned}
F_{T_2}(t) &= F_{T_2}(t|\delta_C, \lambda, n_I, n_C) \\
&= P\left(\frac{Z + \delta_C/\sqrt{\lambda^2/n_I + 1/n_C}}{\sqrt{\frac{\lambda^2/n_I}{\lambda^2/n_I+1/n_C} \frac{W_I}{n_I-1} + \frac{1/n_C}{\lambda^2/n_I+1/n_C} \frac{W_C}{n_C-1}}} \leq t\right) \\
&= P\left(\frac{Z + \lambda^{-1}\delta_C/\sqrt{\lambda^{-2}/n_C + 1/n_I}}{\sqrt{\frac{\lambda^{-2}/n_C}{\lambda^{-2}/n_C+1/n_I} \frac{W_C}{n_C-1} + \frac{1/n_I}{\lambda^{-2}/n_C+1/n_I} \frac{W_I}{n_I-1}}} \leq t\right) \\
&= F_{T_2}(t|\lambda^{-1}\delta_C, \lambda^{-1}, n_C, n_I).
\end{aligned}$$

We suggest that accurate approximations to the distributions of T_1 and T_2 can be obtained by truncating the series for their cdfs and pdfs. The value of k can be

made to depend on the distributional parameters. This is seen by observing that the sequence of coefficients of the series

$$c_k = \frac{\Gamma(b+k)(1-\xi)^k \sqrt{(\xi/\nu)(2(a+b+k))}}{\Gamma(b)k!}$$

for $k > 1$ are decreasing to zero for $\xi \leq 1$. We note that these values can be obtained iteratively by

$$\begin{aligned} c_k &= \frac{(b+k-1)(1-\xi)}{k} \frac{\Gamma(b+k-1)(1-\xi)^{k-1} \sqrt{(\xi/\nu)(2(a+b+k))}}{\Gamma(b)(k-1)!} \\ &= \frac{(b+k-1)(1-\xi)}{k} c_{k-1} = \left(1 + \frac{b-1}{k}\right) (1-\xi) c_{k-1} \end{aligned}$$

with $c_0 = 1$. Taking the limit of c_k as $k \rightarrow \infty$, we have

$$\lim_{k \rightarrow \infty} c_k = 0.$$

This implies that for $\epsilon > 0$, there exist a value of k_0 such that for all $k \geq k_0$, $|c_k| < \epsilon$. A FORTRAN program is given in Appendix I for evaluating the cdf $F_{T_i}(t|\delta_C, \lambda, n_I, n_C)$ of the distribution of T_i for given values of t , δ_C , λ , n_I , and n_C for $i = 1, 2$. This program can be modified to obtain the cdf of T_i given the values of t , $\Delta = \mu_I - \mu_C$, σ_C , λ , n_I , and n_C .

2.4 Welsh's Approximation/Estimation Method

It has been suggested by Welsh (1938) that the distribution of T_2 can be approximated by a t -distribution. This approximation is obtained by assuming that the random quantity

$$W = \frac{S_I^2/n_I + S_C^2/n_C}{\sigma_I^2/n_I + \sigma_C^2/n_C}$$

has approximately the distribution of a chi square random variable that has been divided by its degrees of freedom ν . Setting the variance of W equal to the variance

of a chi square random variable that has been divided by its degrees of freedom ν gives

$$\nu = \frac{(\sigma_I^2/n_I + \sigma_C^2/n_C)^2}{\frac{1}{n_I-1} (\sigma_I^2/n_I)^2 + \frac{1}{n_C-1} (\sigma_C^2/n_C)^2}.$$

Observing that

$$\begin{aligned} \nu &= \frac{\sigma_C^4 \left(\frac{\sigma_I^2}{\sigma_C^2}/n_I + 1/n_C \right)^2}{\sigma_C^4 \left(\frac{1}{n_I-1} \left(\frac{\sigma_I^2}{\sigma_C^2}/n_I \right)^2 + \frac{1}{n_C-1} (1/n_C)^2 \right)} \\ &= \frac{(\lambda^2/n_I + 1/n_C)^2}{\frac{1}{n_I-1} (\lambda^2/n_I)^2 + \frac{1}{n_C-1} (1/n_C)^2}, \end{aligned}$$

we see that ν depends only on the values of n_I , n_C , and λ . If the researcher does not specify a value of λ , Welsh (1938) recommended estimating ν with the statistic

$$\hat{\nu} = \frac{(S_I^2/n_I + S_C^2/n_C)^2}{\frac{1}{n_I-1} (S_I^2/n_I)^2 + \frac{1}{n_C-1} (S_C^2/n_C)^2}.$$

The question that arises at this point is how good is the approximation of Welsh (1938)? This can be examined in two ways. Firstly, for given values of t , λ , n_I , and n_C how well does $F_{t_{\nu,\theta}}(t|\nu,\theta)$ approximate $F_{T_2}(t|\delta_C,\lambda,n_I,n_C)$, where

$$\theta = \delta_C / \sqrt{\lambda^2/n_I + 1/n_C} \text{ and } \nu = \frac{(\lambda^2/n_I + 1/n_C)^2}{\frac{1}{n_I-1} (\lambda^2/n_I)^2 + \frac{1}{n_C-1} (1/n_C)^2}.$$

Secondly, we could ask how well does the 100(1 - α)th percentile $t_{\nu,\theta,\alpha}$ approximate the 100(1 - α)th percentile $t_{\delta_C,\lambda,n_I,n_C,\alpha}$ of the distribution of T_2 for $0 < \alpha < 1$? Another way to ask this question, is how well does $F_{T_2}(t_{\nu,\theta,\alpha}|\delta_C,\lambda,n_I,n_C)$ approximate the value 100(1 - α)%. The FORTRAN program in Appendix I can be used to study these questions. For the case we examined, it was found that Welch's method provides a good approximation to the cdf of the distribution of T_2 .

2.5 Conclusion

An exact distribution was given to Behrens-Fisher problem under the independent normal distribution. It was shown that the cdf (pdf) is an infinite series of the cdfs (pdfs) of noncentral t -distributions. In general, it was observed that the Behrens-Fisher distribution depends on the difference in the means, the variances, and the sample sizes. If the means are equal, then this distribution depends only on the ratio of the variances and the sample sizes. A numerical method is presented for obtaining a good approximation to the cdf and pdf of the Behrens-Fisher distribution. A FORTRAN program was written to implement this method. The approximation and estimation methods presented in Welch (1938) were discussed. These methods appear to provide good approximation of the cdf of the distribution of T_2 . The results presented in this Chapter are useful in interval estimation and hypothesis testing when comparing the two population means.

CHAPTER 3
COMPARING TWO POPULATION MEANS: INDEPENDENT
SAMPLES CASE

3.1 Introduction

It is of interest here to study the effect of an intervention for a given medical condition for individuals in a given population. In this chapter, we consider the effect of the intervention as measured by the mean μ_I of the distribution of a continuous measurement X on each individual in the population to receive the intervention. If μ_C is the mean of the distribution of X when there is no intervention or when another treatment is used, it is typically of interest to know the value of $\mu_I - \mu_C$. If $\mu_I - \mu_C$ is negative, zero, or positive, then the intervention has made the condition worse, has had no effect, or has improved the medical condition of the individuals in the given population, respectively. As is commonly done, we will assume the models for the two distributions of X to be $N(\mu_I, \sigma_I^2)$ over the population with intervention and $N(\mu_C, \sigma_C^2)$ for the control.

In order to make an inference about $\mu_I - \mu_C$, we begin by assuming that samples from the two populations (intervention and control) of sizes n_I and n_C , respectively, are to be taken. The measurement on the individuals in these respective samples will be denoted by $X_{I,1}, \dots, X_{I,n_I}$ and $X_{C,1}, \dots, X_{C,n_C}$. We assume that these measurements are independent random samples with

$$X_{I,i} \sim N(\mu_I, \sigma_I^2) \text{ and } X_{C,j} \sim N(\mu_C, \sigma_C^2),$$

for $i = 1, \dots, n_I$ and $j = 1, \dots, n_C$. The means and variances of these samples are denoted by \bar{X}_I and S_I^2 (provided $n_I > 1$) and \bar{X}_C and S_C^2 (provided $n_C > 1$), respectively. We observe that the statistic $\bar{X}_I - \bar{X}_C$ provides an unbiased estimator for

Table 3.1: Sample Size Selection Cases

{t3.1}

Case	Assumption
1	σ_I^2 and σ_C^2 both known;
2	$\lambda^2 = \sigma_I^2/\sigma_C^2$ is known, and
3	σ_I^2 and σ_C^2 both unknown.

the parameter $\mu_I - \mu_C$ and under our model has a Normal distribution with variance

$$\sigma_{\bar{X}_I - \bar{X}_C}^2 = \sigma_I^2/n_I + \sigma_C^2/n_C.$$

The selection of the sample sizes n_I and n_C is under the control of the researcher. How does a researcher select these values? An answer to this question will be discussed in two parts. Firstly, we examine the selection of these values for a fixed total $m = n_I + n_C$ of the two sample sizes with the restriction that $1 \leq n_I, n_C < m$. Secondly, we investigate methods for selecting the total sample size m . We further divide the problem into three cases listed in Table 3.1.

The cases are individually addressed respectively in the next three sections.

Under the aforementioned scenarios, we are interested in testing the null hypothesis $H_0 : \mu_I = \mu_C$ (no affect due the intervention) versus the alternative (researcher's) hypothesis $H_a : \mu_I \neq \mu_C$ (there is an affect due to the medical intervention). In general, our test rejects the H_0 in favor of H_a if $|T| \geq t^*$, where common selections of the test statistic T in Table 3.2.

where

$$S_p^2 = \frac{(n_I - 1) S_I^2 + (n_C - 1) S_C^2}{n_I + n_C - 2}.$$

The first test statistic listed in Table 3.2 is used if the variances are known. The second

Table 3.2: Test Statistics

{t3.2}

$$\begin{aligned}
 T &= \frac{\bar{X}_I - \bar{X}_C}{\sqrt{\sigma_I^2/n_I + \sigma_C^2/n_C}} \\
 T = T_1 &= \frac{\bar{X}_I - \bar{X}_C}{S_p \sqrt{1/n_I + 1/n_C}} \\
 T = T_2 &= \frac{\bar{X}_I - \bar{X}_C}{\sqrt{S_I^2/n_I + S_C^2/n_C}}
 \end{aligned}$$

statistic in Table 3.2 is often recommended if the researcher is willing to assume the unknown variances are equal. If the researcher is not willing to assume the unknown variances are equal, the third test statistic (whose distribution under the independent normal model is referred to as the Behrens-Fisher distribution) in Table 3.2 is the most commonly recommended test.

The critical value t^* of a test depends on the null distribution of the test statistic T . The general distribution of T is needed to study the power of the test. The distribution of the first test statistic in Table 3.2 under the independent normal model is well known. The distributions of the second and third test statistics are derived in this chapter. These distributions depend on the ratio $\lambda = \sigma_I/\sigma_C$ and hence a test of a given size cannot be selected unless this ratio is given.

If the researcher is willing to assume that $\lambda = 1$, then the test based on T_1 is generally recommended. If the researcher is not willing to make the assumption that $\lambda = 1$, then a test based on T_2 is generally recommended. We examine in this chapter tests when λ is known and when λ is unknown. For each of these tests when λ is known, we can evaluate the size and power of the test for fixed sample sizes as well as a method for determining the sample sizes. For test that use the data to estimate λ , it is shown how these tests perform using simulation.

3.2 Variances Known Case

If the population variances σ_I^2 and σ_C^2 are known, then the test we will consider rejects $H_0 : \mu_I = \mu_C$ in favor of $H_a : \mu_I \neq \mu_C$ if $|T| \geq z_{\alpha^*/2}$, where $0 < \alpha^* < 1$ and

$$T = \frac{\bar{X}_I - \bar{X}_C}{\sqrt{\sigma_I^2/n_I + \sigma_C^2/n_C}}.$$

We see that T can be expressed as

$$\begin{aligned} T &= \frac{(\bar{X}_I - \bar{X}_C) - (\mu_I - \mu_C)}{\sqrt{\sigma_I^2/n_I + \sigma_C^2/n_C}} + \frac{\mu_I - \mu_C}{\sqrt{\sigma_I^2/n_I + \sigma_C^2/n_C}} \\ &= Z + \frac{\delta_C}{\sqrt{\lambda^2/n_I + 1/n_C}}, \end{aligned}$$

where

$$Z = \frac{(\bar{X}_I - \bar{X}_C) - (\mu_I - \mu_C)}{\sqrt{\sigma_I^2/n_I + \sigma_C^2/n_C}}, \quad \delta_C = \frac{\mu_I - \mu_C}{\sigma_C}, \quad \text{and } \lambda = \frac{\sigma_I}{\sigma_C}.$$

It will be convenient to define

$$\theta_C = \frac{\delta_C}{\sqrt{\lambda^2/n_I + 1/n_C}}.$$

Note that the null and alternative hypotheses can be expressed as

$$H_0 : \delta_C = 0 \quad \text{and} \quad H_a : \delta_C \neq 0.$$

Under the independent normal model, the random variable Z has a standard normal distribution. If the null hypothesis holds, then T has a standard normal distribution and hence the selection of the critical value $z_{\alpha^*/2}$. The size of the test is α^* and the

power function is

$$\begin{aligned}
\pi(\delta_C, \lambda, n_I, n_C) &= P(|T| \geq z_{\alpha^*/2}) \\
&= 1 - P(-z_{\alpha^*/2} - \theta_C < Z < z_{\alpha^*/2} - \theta_C) \\
&= 1 - \Phi(z_{\alpha^*/2} - \theta_C) + \Phi(-z_{\alpha^*/2} - \theta_C) \\
&= 1 - \Phi(z_{\alpha^*/2} - \theta_C) + 1 - \Phi(z_{\alpha^*/2} + \theta_C),
\end{aligned}$$

when $\delta_C \neq 0$. It is clear that the power function depends on the population means and variances through the parameters δ_C and λ . The power function also depends on the sample sizes n_I and n_C .

Let us assume that $\delta_C \neq 0$ is fixed. For a fixed total sample size $m = n_I + n_C$ and $n = n_I$, we can write

$$\theta_C = \delta_C (\lambda^2 n^{-1} + (m - n)^{-1})^{-1/2}.$$

To determine the value of n that maximizes the power function, we observe that

$$\begin{aligned}
\frac{\partial \pi}{\partial n} &= \phi(z_{\alpha^*/2} - \theta_C) \frac{\partial \theta_C}{\partial n} - \phi(-z_{\alpha^*/2} - \theta_C) \frac{\partial \theta_C}{\partial n} \\
&= (\phi(z_{\alpha^*/2} - \theta_C) - \phi(-z_{\alpha^*/2} - \theta_C)) \frac{\partial \theta_C}{\partial n} \\
&= A \frac{\partial \theta_C}{\partial n},
\end{aligned}$$

where

$$A = \phi(z_{\alpha^*/2} - \theta_C) - \phi(-z_{\alpha^*/2} - \theta_C).$$

Next observe that we can write

$$A = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z_{\alpha^*/2}^2 + 2z_{\alpha^*/2}\theta_C + \theta_C^2)} \left(1 - e^{-2z_{\alpha^*/2}\delta_C/\sqrt{\lambda^2/n+1/(m-n)}}\right).$$

Since $\delta_C \neq 0$, then $A \neq 0$. Hence, a solution to the equation $\partial \pi / \partial n = 0$ is also a solution to the equation $\partial \theta_C / \partial n = 0$.

We see that

$$\frac{\partial \theta_C}{\partial n} = -\frac{1}{2} \delta_C (\lambda^2 n^{-1} + (m-n)^{-1})^{-3/2} (-\lambda^2 n^{-2} + (m-n)^{-2})$$

and since $\delta_C (\lambda^2 n^{-1} + (m-n)^{-1})^{-3/2} \neq 0$, then the equation $\partial \theta_C / \partial n = 0$ has the same solution for n that

$$-\lambda^2 n^{-2} + (m-n)^{-2} = 0.$$

It is easy to see that the real number solution n to this equation is

$$n = \frac{\lambda}{1+\lambda} m.$$

Since we require n to be a positive integer, then we consider selecting $n_I = n$ to be either

$$\begin{aligned} \text{Method (1) } n &= \left\lfloor \frac{\lambda}{1+\lambda} m \right\rfloor \text{ or} \\ \text{Method (2) } n &= \left\lceil \frac{\lambda}{1+\lambda} m \right\rceil. \end{aligned}$$

To maximize the power, the value of n needs to be selected such that

$$\theta_C = \theta_C(\delta_C, \lambda, n, m) = \frac{\delta_C}{\sqrt{\lambda^2/n + 1/(m-n)}}$$

for a fixed value of m and the value $|\theta_C|$ as large as possible. Consider the function

$$d(\delta_C, \lambda, m) = \frac{|\delta_C|}{\sqrt{\left\lceil \frac{\lambda}{1+\lambda} m \right\rceil + \left\lceil \frac{1}{1+\lambda} m \right\rceil}} - \frac{|\delta_C|}{\sqrt{\left\lfloor \frac{\lambda}{1+\lambda} m \right\rfloor + \left\lfloor \frac{1}{1+\lambda} m \right\rfloor}}.$$

The value of n should be selected according to Method (1) if $d(\delta_C, \lambda, m) \leq 0$ and using Method (2) if $d(\delta_C, \lambda, m) \geq 0$. The plot of $d(1, \lambda, 10)$ versus λ (see Figure 3.1) reveals that there does not exist a simple formula for determining when $d(\delta, \lambda, m)$ is negative, zero, or positive. Hence, we recommend simply calculating the value of $d(\delta, \lambda, m)$ for given value of δ , λ , and m to select the value of n .

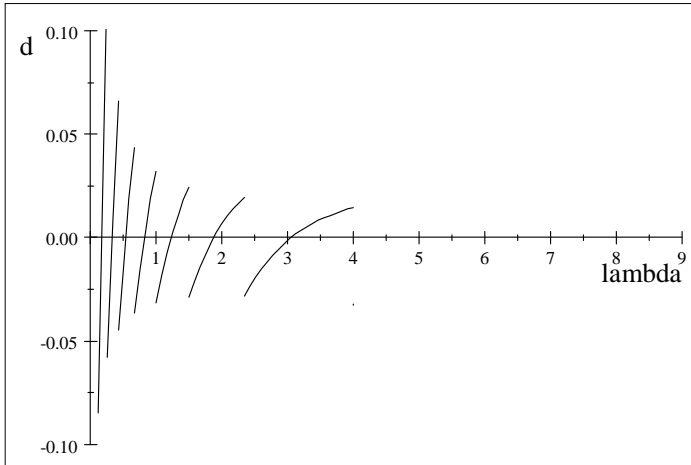


Figure 3.1: $d = \frac{1}{\sqrt{\frac{\lambda^2}{1+\lambda^{10}} + \frac{1}{1+\lambda^{10}}}} - \frac{1}{\sqrt{\frac{\lambda^2}{1+\lambda^{10}} + \frac{1}{1+\lambda^{10}}}}$

Figure 3.1: $d = \frac{1}{\sqrt{\frac{\lambda^2}{1+\lambda^{10}} + \frac{1}{1+\lambda^{10}}}} - \frac{1}{\sqrt{\frac{\lambda^2}{1+\lambda^{10}} + \frac{1}{1+\lambda^{10}}}}$.

{f3.1}

We next consider the selection of m . Assume the researcher has selected a positive value δ_C^* such that for $|\delta_C| \geq \delta_C^*$ it is desired to have the power to be at least $1 - \beta^*$. We stress here that the values α^* , δ_C^* and β^* are to be selected by the researcher. For all δ_C such that $|\delta_C| \geq \delta_C^*$, we have that

$$\begin{aligned} \pi(\delta_C, \lambda, n, m - n) &\geq \pi(\delta_C^*, \lambda, n, m - n) \\ &= 1 - \Phi\left(z_{\alpha^*/2} - \frac{\delta_C^*}{\sqrt{\lambda^2/n + 1/(m - n)}}\right) \\ &\quad + 1 - \Phi\left(z_{\alpha^*/2} + \frac{\delta_C^*}{\sqrt{\lambda^2/n + 1/(m - n)}}\right) \\ &\geq 1 - \beta^*. \end{aligned}$$

We note that the value

$$1 - \Phi\left(z_{\alpha^*/2} + \frac{\delta_C^*}{\sqrt{\lambda^2/n + 1/(m - n)}}\right)$$

decreases as m increases and that

$$\begin{aligned} & 1 - \Phi \left(z_{\alpha^*/2} - \frac{\delta_C^*}{\sqrt{\lambda^2/n + 1/(m-n)}} \right) \\ & + 1 - \Phi \left(z_{\alpha^*/2} + \frac{\delta_C^*}{\sqrt{\lambda^2/n + 1/(m-n)}} \right) \\ & \geq 1 - \Phi \left(z_{\alpha^*/2} - \frac{\delta_C^*}{\sqrt{\lambda^2/n + 1/(m-n)}} \right). \end{aligned}$$

Hence, we elect to choose m such that

$$\begin{aligned} 1 - \Phi \left(z_{\alpha^*/2} - \frac{\delta_C^*}{\sqrt{\lambda^2/n + 1/(m-n)}} \right) & \geq 1 - \beta^* \text{ or} \\ \Phi \left(z_{\alpha^*/2} - \frac{\delta_C^*}{\sqrt{\lambda^2/n + 1/(m-n)}} \right) & \leq \beta^*. \end{aligned}$$

It then follows that we should choose m such that

$$\begin{aligned} z_{\alpha^*/2} - \frac{\delta_C^*}{\sqrt{\lambda^2/n + 1/(m-n)}} & \leq -z_{\beta^*} \text{ or} \\ \frac{\delta_C^*}{\sqrt{\lambda^2/n + 1/(m-n)}} & \geq z_{\alpha^*/2} + z_{\beta^*}. \end{aligned} \tag{3.1} \quad \{\mathbf{m1}\}$$

As was previously shown, we have that

$$\frac{\delta_C^* \sqrt{m}}{1 + \lambda} = \frac{\delta_C^*}{\sqrt{\lambda^2 / \left(\frac{\lambda}{1+\lambda} m\right) + 1 / \left(\frac{1}{1+\lambda} m\right)}} \geq \frac{\delta_C^*}{\sqrt{\lambda^2/n + 1/(m-n)}}.$$

It follows that for any value of m that satisfies the inequality (3.1) must satisfy the compound inequality

$$m \geq \frac{(1 + \lambda)^2}{\lambda^2/n + 1/(m-n)} \geq \frac{(1 + \lambda)^2 (z_{\alpha^*/2} + z_{\beta^*})^2}{(\delta_C^*)^2}.$$

The smallest value of m that is a solution to the compound inequality is

$$\begin{aligned} m & = \left\lceil \frac{(1 + \lambda)^2 (z_{\alpha^*/2} + z_{\beta^*})^2}{(\delta_C^*)^2} \right\rceil \\ & = \left\lceil \frac{(1 + \sigma_I/\sigma_C)^2 (z_{\alpha^*/2} + z_{\beta^*})^2}{(\delta_C^*)^2} \right\rceil. \end{aligned}$$

Once m has been selected, then the previous method can be used for selecting n for a fixed value of m .

3.3 Unknown Variances

Welch (1938) recommended a test based on the statistic T_2 that rejects H_0 in favor of H_a if $|T_2| \geq t_{\nu,0,\alpha/2}$, where value $t_{\nu,0,\alpha/2}$ is the 100 $(1 - \alpha/2)$ th percentile of a central t -distribution with degrees of freedom

$$\nu = \frac{(\sigma_I^2/n_I + \sigma_C^2/n_C)^2}{(\sigma_I^2/n_I)^2 / (n_I - 1) + (\sigma_C^2/n_C)^2 / (n_C - 1)}.$$

The critical value $t_{\nu,0,\alpha/2}$ is an approximation to the 100 $(1 - \alpha/2)$ th percentile of the distribution of T_2 (see Chapter 2 concerning this approximation). In the case in which $\lambda = \sigma_I/\sigma_C$ is not known, he recommended that ν be estimated from the observed data using the statistic

$$V = \frac{(S_I^2/n_I + S_C^2/n_C)^2}{(S_I^2/n_I)^2 / (n_I - 1) + (S_C^2/n_C)^2 / (n_C - 1)}.$$

In this case, the critical value $t_{V,0,\alpha/2}$ is a random variable whose observed value will be used to estimate the 100 $(1 - \alpha/2)$ th percentile of the distribution of T_2 . It then follows that the size A and the power P of the test are random variables.

How well this tests performs depends on how good the t -distribution with ν (V) degrees of freedom approximates (estimates) the distribution of T_2 . Further, one would not expect the test based on λ known to perform as well as one based on an estimated λ . In the case in which λ is known (estimated), the size of the test α and the power of the test π are approximated (estimated) by, respectively,

$$\begin{aligned} a &= 1 - F_{t_{\nu,\delta_C/\sqrt{\lambda^2/n_I+1/n_C}}} (t_{\nu,0,\alpha/2} | 0, \lambda, n_I, n_C) \\ &\quad + F_{t_{\nu,\delta_C/\sqrt{\lambda^2/n_I+1/n_C}}} (-t_{\nu,0,\alpha/2} | 0, \lambda, n_I, n_C), \\ (A &= 1 - F_{t_{V,\delta_C/\sqrt{L^2/n_I+1/n_C}}} (t_{V,0,\alpha/2} | 0, L, n_I, n_C) \\ &\quad + F_{t_{V,\delta_C/\sqrt{L^2/n_I+1/n_C}}} (-t_{V,0,\alpha/2} | 0, L, n_I, n_C)), \end{aligned}$$

and

$$\begin{aligned}
 p &= 1 - F_{t_{\nu, \delta_C / \sqrt{\lambda^2/n_I + 1/n_C}}} (t_{\nu, 0, \alpha/2} | \delta_C, \lambda, n_I, n_C) \\
 &\quad + F_{t_{\nu, \delta_C / \sqrt{\lambda^2/n_I + 1/n_C}}} (-t_{\nu, 0, \alpha/2} | \delta_C, \lambda, n_I, n_C), \\
 (P &= 1 - F_{t_{V, \delta_C / \sqrt{L^2/n_I + 1/n_C}}} (t_{V, 0, \alpha/2} | \delta_C, L, n_I, n_C) \\
 &\quad + F_{t_{V, \delta_C / \sqrt{L^2/n_I + 1/n_C}}} (-t_{V, 0, \alpha/2} | \delta_C, L, n_I, n_C)).
 \end{aligned}$$

We now consider tests that make use of the exact distribution of T_i for $i = 1, 2$. Firstly, we consider tests in which the ratio $\lambda = \sigma_I/\sigma_C$ is known. A test of size α (to be selected by the researcher) rejects null hypothesis $H_0 : \mu_I = \mu_C$ in favor the alternative hypothesis $H_a : \mu_I \neq \mu_C$ if $|T_i| \geq t_i$ (critical value), where t_i is the $100(1 - \alpha/2)$ th percentile of the distribution of T_i when $\mu_I = \mu_C$, for $i = 1$ or 2 . That is,

$$\alpha = 1 - F_{T_i}(t_i | 0, \lambda, n_I, n_C) + F_{T_i}(-t_i | 0, \lambda, n_I, n_C).$$

The power of this test is

$$\pi = 1 - F_{T_i}(t_i | \delta_C, \lambda, n_I, n_C) + F_{T_i}(-t_i | \delta_C, \lambda, n_I, n_C),$$

for $\delta_C \neq 0$. Since the value of the function $F_{T_i}(t | \delta_C, \lambda, n_I, n_C)$ can be determined numerically, then the critical value t_i can be determined under the independent normal model as well as the exact power of the test.

A FORTRAN program is given in Appendix II for determining the critical value t_i of this test for given values of α , λ , n_I , and n_C as well as the power. If the researcher specifies a value δ_C^* and a minimum value $1 - \beta^*$ of the power when $|\delta_C| \geq \delta_C^*$, then the program in Appendix II can be used to determine the minimum value of $n_I + n_C$ to meet these requirements. Also, this program can be used for λ known to compare the method given in Welch (1938) and the exact method presented here. For a few

values of λ we considered, the approximation in Welch (1938) gives fairly accurate critical values.

In general, the researcher will not be able to specify the value λ . In this case, a test can be constructed by first estimating λ and then selecting the critical value as the value C_i that satisfies the equation

$$F_{T_i}(C_i | 0, L, n_I, n_C) = 1 - \alpha/2,$$

where L is an estimate/estimator of the parameter λ . A biased choice for L is S_I/S_C .

An unbiased choice is

$$\frac{\sqrt{n_I - 1} \Gamma\left(\frac{n_I - 1}{2}\right) \Gamma\left(\frac{n_C - 1}{2}\right) S_I}{\sqrt{n_C - 1} \Gamma\left(\frac{n_I}{2}\right) \Gamma\left(\frac{n_C - 2}{2}\right) S_C}.$$

It may at this point seem more reasonable to estimate the value of λ^2 since it is this value that is used directly in the evaluation of $F_{T_i}(t | \delta_C, \lambda, n_I, n_C)$. An unbiased estimator for λ^2 is the statistic

$$\frac{(n_C - 4) S_I^2}{(n_C - 1) S_C^2},$$

provided $n_C > 4$. The size of this test

$$A = 1 - F_{T_i}(C_i | 0, \lambda, n_I, n_C) + F_{T_i}(-C_i | 0, \lambda, n_I, n_C)$$

is a random variable as the power of the test

$$P = 1 - F_{T_i}(C_i | \delta_C, \lambda, n_I, n_C) + F_{T_i}(-C_i | \delta_C, \lambda, n_I, n_C),$$

for fixed values of n_I and n_C . One can study the distributions of A and P using simulation, but only for selected values of λ as the value of λ is required to simulate a value of T_i for $i = 1, 2$. A FORTRAN program is given in Appendix III that can be used to estimate $E(A)$ and $E(P)$ for a given value of λ . This program is only useful in demonstrating how well the test performs for a given value of λ . Since λ is

unknown, one cannot use this program to help the researcher select the values of n_I and n_C . In Chapter 5, a method for sample size selection will be presented.

3.4 Conclusion

In this chapter, we have discussed statistical tests for comparing two means using data in the form of independent random samples. The problem of sample size determination was addressed based on the power of the test. Results are limited to the case in which the ratio of the standard deviations is given. These methods require information from the researcher about the distribution of the test statistic when the alternative hypothesis holds.

CHAPTER 4
COMPARING TWO MEANS: PAIRED DATA CASE

4.1 Introduction

There are a variety of examples in which the intervention can be given to each individual in the treatment group with the individual also serving as the control. This is often referred to as the paired data case. The X measurement is first taken on the individual and once again after the individual is treated with the intervention. We will refer to these measurements as X_C and X_I , respectively. Of interest is to make an inference about the parameter

$$\mu_{X_I - X_C} = \mu_I - \mu_C.$$

A commonly used model for these type of data is the bivariate normal distribution. In particular, we write

$$\begin{bmatrix} X_I \\ X_C \end{bmatrix} \sim N_2 \left(\begin{bmatrix} \mu_I \\ \mu_C \end{bmatrix}, \begin{bmatrix} \sigma_I^2 & \rho\sigma_I\sigma_C \\ \rho\sigma_I\sigma_C & \sigma_C^2 \end{bmatrix} \right).$$

It is not difficult to show under this model that the difference $D = X_I - X_C$ has a normal distribution with mean and variance given by

$$\mu_D = \mu_I - \mu_C \text{ and } \sigma_D^2 = \sigma_I^2 + \sigma_C^2 - 2\text{cov}(X_C, X_I).$$

Let

$$\begin{bmatrix} X_{I,1} \\ X_{C,1} \end{bmatrix}, \dots, \begin{bmatrix} X_{I,n} \\ X_{C,n} \end{bmatrix}$$

be a random sample, then the sample mean vector and covariance matrix are given

by

$$\begin{bmatrix} \bar{X}_I \\ \bar{X}_C \end{bmatrix} \text{ and } \begin{bmatrix} S_I^2 & S_{I,C} \\ S_{I,C} & S_C^2 \end{bmatrix}.$$

Let

$$D_1 = X_{I,1} - X_{C,1}, \dots, D_n = X_{I,n} - X_{C,n}$$

then D_1, \dots, D_n are independent and identically distributed as $N(\mu_D, \sigma_D^2)$ random variables. Thus, we have

$$\bar{D} \sim N(\mu_D, \sigma_D^2/n) \text{ and } \frac{(n-1)S_D^2}{\sigma_D^2} \sim \chi_{n-1}^2.$$

Further, we have that

$$T = \frac{\bar{D}}{S_D/\sqrt{n}} \sim t_{n-1, \sqrt{n}\delta_D},$$

where $\delta_D = \mu_D/\sigma_D$. We note that

$$\bar{D} = \bar{X}_I - \bar{X}_C \text{ and } S_D^2 = S_I^2 + S_C^2 - 2S_{I,C}.$$

4.2 Inference about μ_D , Variance σ_D^2 Known

Assuming that σ_D^2 is known, a test of size α rejects $H_0 : \mu_D = 0$ versus $H_a : \mu_D \neq 0$ (or equivalently, $H_0 : \delta_D = 0$ versus $H_a : \delta_D \neq 0$) if $|T| \geq z_{\alpha/2}$, where

$$T = \frac{\bar{D}}{\sigma_D/\sqrt{n}}.$$

Observe that

$$T = \frac{\bar{D} - \mu_D}{\sigma_D/\sqrt{n}} + \sqrt{n} \frac{\mu_D}{\sigma_D} = Z + \sqrt{n}\delta_D$$

with

$$Z = \frac{\bar{D} - \mu_D}{\sigma_D/\sqrt{n}} \sim N(0, 1).$$

The power $\pi(\delta_D, n)$ of the test for $\delta_D \neq 0$ is

$$\begin{aligned}\pi(\alpha, \delta_D, n) &= P(|T| \geq z_{\alpha/2}) = 1 - P(-z_{\alpha/2} < T < z_{\alpha/2}) \\ &= 1 - P(-z_{\alpha/2} - \sqrt{n}\delta_D < Z < z_{\alpha/2} - \sqrt{n}\delta_D) \\ &= 1 - \Phi(z_{\alpha/2} - \sqrt{n}\delta_D) + \Phi(-z_{\alpha/2} - \sqrt{n}\delta_D),\end{aligned}$$

where $\Phi(z)$ is the cumulative distribution function of a standard normal distribution.

Assume the researcher desires to use this test of size α^* with the power of the test having at least the value $1 - \beta^*$ if $|\delta_D| \geq \delta_D^* > 0$. The values β^* and δ_D^* are to be specified as α^* by the researcher. Since the function $\pi(\delta_D)$ is an increasing function of $|\delta_D|$, then the researcher must select a sample of size n such that

$$\pi(\alpha^*, \delta_D^*, n) \geq 1 - \beta^*.$$

Equivalently, the sample size n must satisfy the inequality

$$\Phi(z_{\alpha^*/2} - \sqrt{n}\delta_D^*) - \Phi(-z_{\alpha^*/2} - \sqrt{n}\delta_D^*) \leq \beta^*. \quad (4.1) \quad \{\text{eq1}\}$$

Since $0 < \Phi(z) < 1$ and $\Phi(z_{\alpha^*/2} - \sqrt{n}\delta_D^*) > \Phi(-z_{\alpha^*/2} - \sqrt{n}\delta_D^*)$ for $z_{\alpha^*/2} > 1/2$ and $\delta_D^* > 0$, then

$$\Phi(z_{\alpha^*/2} - \sqrt{n}\delta_D^*) - \Phi(-z_{\alpha^*/2} - \sqrt{n}\delta_D^*) \leq \Phi(z_{\alpha^*/2} - \sqrt{n}\delta_D^*).$$

Hence any value of n that satisfies the inequality $\Phi(z_{\alpha^*/2} - \sqrt{n}\delta_D^*) \leq \beta^*$ also satisfies the inequality in (4.1). Now as n increases, the probability $\Phi(-z_{\alpha^*/2} - \sqrt{n}\delta_D^*)$ decreases. Thus, the smallest value of n that satisfies the inequality $\Phi(z_{\alpha^*/2} - \sqrt{n}\delta_D^*) \leq \beta^*$ also satisfies the inequality in (4.1). The minimum value of n that satisfies the inequality (4.1) should also satisfies the inequality

$$z_{\alpha^*/2} - \sqrt{n}\delta_D^* \leq -z_{\beta^*} \quad \text{or} \quad n \geq \frac{(z_{\alpha^*/2} + z_{\beta^*})^2}{(\delta_D^*)^2}.$$

Hence a conservative choice for the desired sample size is

$$n \geq \left\lceil \frac{(z_{\alpha^*/2} + z_{\beta^*})^2}{(\delta_D^*)^2} \right\rceil.$$

4.3 Variance σ_D^2 Unknown

Assuming that σ_D^2 is unknown, a commonly recommended test of size α rejects $H_0 : \mu_D = 0$ versus $H_a : \mu_D \neq 0$ if $|T| \geq t_{n-1,0,\alpha/2}$, where

$$T = \frac{\bar{D}}{S_D/\sqrt{n}}$$

and $t_{n-1,0,\alpha/2}$ is the 100(1 - $\alpha/2$) percentile of a central t -distribution with $n - 1$ degrees of freedom. Observe that

$$T = \frac{\frac{\bar{D} - \mu_D}{\sigma_D/\sqrt{n}} + \sqrt{n}\delta_D}{\sqrt{\frac{(n-1)S_D^2}{\sigma_D^2}/(n-1)}} \sim t_{n-1,\sqrt{n}\delta_D},$$

where $t_{n-1,\sqrt{n}\delta_D}$ is a random variable having a noncentral t -distribution with $n - 1$ degrees of freedom and noncentrality parameter $\sqrt{n}\delta_D$. The power $\pi(\alpha, \delta_D, n)$ of the test for $\delta_D \neq 0$ is

$$\begin{aligned} \pi(\alpha, \delta_D, n) &= 1 - P(-t_{n-1,0,\alpha/2} < t_{n-1,\sqrt{n}\delta_D} < t_{n-1,0,\alpha/2}) \\ &= 1 - F_{t_{n-1,\sqrt{n}\delta_D}}(t_{n-1,0,\alpha/2} - \sqrt{n}\delta_D) + F_{t_{n-1,\sqrt{n}\delta_D}}(-t_{n-1,0,\alpha/2} - \sqrt{n}\delta_D), \end{aligned}$$

where $F_{t_{n-1,\sqrt{n}\delta_D}}(t)$ is the cdf of a noncentral t -distribution with $n - 1$ degrees of freedom and noncentrality parameter $\sqrt{n}\delta_D$.

Suppose the researcher desires the test to be of size α^* and the power of the test to be at least $1 - \beta^*$ when $|\delta_D|$ is at least as large as $\delta_D^* > 0$. This requires that n be

selected such that

$$\pi(\alpha^*, \delta_D^*, n) = 1 - F_{t_{n-1}, \sqrt{n}\delta_D^*}(t_{n-1,0,\alpha^*/2} - \sqrt{n}\delta_D^*) \quad (4.2) \quad \{\text{eq3}\}$$

$$+ F_{t_{n-1}, \sqrt{n}\delta_D^*}(-t_{n-1,0,\alpha^*/2} - \sqrt{n}\delta_D^*) \quad (4.3)$$

$$\geq 1 - \beta^*.$$

Since $F_{t_{n-1}, \sqrt{n}\delta_D^*}(-t_{n-1,0,\alpha^*/2} - \sqrt{n}\delta_D^*)$ is approximately zero for relatively small values of n , then the minimum value of n that satisfies the inequality

$$F_{t_{n-1}, \sqrt{n}\delta_D^*}(t_{n-1,0,\alpha^*/2} - \sqrt{n}\delta_D^*) \leq \beta^* \text{ or } n \geq \frac{(t_{n-1,0,\alpha^*/2} + t_{n-1, \sqrt{n}\delta_D^*, \beta^*})^2}{(\delta_D^*)^2}$$

satisfies inequality (4.2) for researcher specified values α^* , $1 - \beta^*$, and δ_D^* .

4.4 Conclusion

An individual in some cases can serve as their own control. For this case, we have presented commonly recommended methods for comparing the means of two populations for paired data. Methods were given both for the case in which the standard deviation of the distribution of differences is given and estimated. Relative simple methods were derived for determining the sample size for the study.

CHAPTER 5

SEQUENTIAL AND ADAPTIVE METHODS

5.1 Introduction

Jennison and Turnbull (2000) noted that “formal application of sequential procedures started in the late 1920s in the area of statistical quality control in manufacturing production.” It was in the early 1920s that Walter A. Shewhart introduced the quality control chart for sequentially analyzing the output of a process for the purpose of improving the quality of the process (see Shewhart 1931). Shewhart (1925) stated “the object of this note is to emphasize what appears to be a comparatively new field of application of statistical methods.” He went on to describe the quality control chart and pointed out it is used sequentially as a statistical tool for improving and maintaining the quality of a production process. Dodge and Romig (1929) introduced acceptance sampling procedures that are sequential methods that today are viewed as methods for improving the quality of the output of a process by removing poor quality items. The ideas of Type I and Type II errors have their origin in the producer’s risk and the customer’s risk. The theory of sequential statistical analysis in designed experiments has its origin in the works of Barnard (1946) and Wald (1947). According to Wald (1947), “an essential feature of the sequential test, as distinguished from the current test procedures, is that the number of observations required by the sequential test depends on the outcome of the observations and is, therefore, not predetermined, but a random variable.” Chow and Chang (2007) discussed adaptive design methods in clinical trials. They stated “the adaptive design methods are usually developed based on observed treatment effects.” This allows for “wider flexibility, adaptations in clinical investigation of treatment regimen may include changes of sample size, inclu-

sion/exclusion criteria, study dose, study endpoints, and methods for analysis (Liu, Proschan, and Pldeger, 2002).” In this chapter, we will examine (some) sequential and adaptive methods applied in clinical trials in which the sample size(s) is (are) a random variable. As of the writing of this thesis, the Food and Drug Administration (FDA) has become receptive to the use of adaptive designs in drug development and testing. They are now active in promoting research in the use of adaptive designs in clinical trials.

Firstly, we discuss some adaptive and sequential methods that use an initial sample to predict the total sample size(s). These methods have been referred to as two-stage sample size prediction methods. It is claimed by Dudewicz, et al. (2007) that three of these methods provide solutions to the Behrens-Fisher problem. The second method causes the researcher to select more data if a conclusion to fail to reject or to reject the null hypothesis is not made with the present data. Since the total sample sizes are not fixed and it is changed based on the prior test it follows that the total sample sizes are random variables for both types of procedures. The attempt in all of these cases is to have a test of a fixed size and to have the required power.

5.2 Two-Stage Sample Prediction Methods

Dantzig (1940) proved that there does not exist any t -test for a fixed sample size(s) in which the power is independent of the population(s) standard deviation(s). Stein (1945) introduced a two-stage test of $H_0 : \mu = \mu_0$ versus $H_a : \mu \neq \mu_0$ in which the power is independent of σ . To use this procedure, the researcher before collecting any data first selects “a priori” a positive number u^* and an initial sample size $n^* \geq 2$.

Next the researcher selects a sample of size n^* from the population. Assume the X measurements X_1, \dots, X_{n^*} on the individuals in this sample constitutes a random samples from a $N(\mu, \sigma^2)$ distribution. Let $S_{n^*}^2$ denote the variance of this sample. The total sample size N for the experiment is the observed value of the random variable defined by

$$N = \max \left\{ \left[\frac{S_{n^*}^2}{u^*} \right] + 1, n^* + 1 \right\},$$

where $[q] =$ largest integer $\leq q$.

Clearly, the researcher using this method will always select at least one more individual from the population than the initial sample size n^* . Using these results the random variables A_1, \dots, A_N are selected subject to the restrictions

$$\sum_{j=1}^N A_j = 1; A_1 = \dots = A_{n^*}; \text{ and } \sum_{j=1}^N A_j^2 = \frac{u^*}{S_{n^*}^2}.$$

The following two theorems are used in the design of the test.

Theorem 5.2.1. *The random variable*

$$T_0 = \frac{\sum_{j=1}^N A_j X_j - \mu}{\sqrt{u^*}} \sim t_{n^*-1}.$$

Proof. First observe that

$$\begin{aligned} E \left(\sum_{j=1}^N A_j X_j \right) &= E \left[E \left(\sum_{j=1}^N A_j X_j \mid S_{n^*}^2 \right) \right] \\ &= E \left(\sum_{j=1}^N A_j E(X_j \mid S_{n^*}^2) \right) \\ &= E \left(\left(\sum_{j=1}^N A_j \right) \mu \mid S_{n^*}^2 \right) \\ &= E(\mu \mid S_{n^*}^2) = \mu. \end{aligned}$$

The variance of $\sum_{j=1}^N A_j X_j$ can be determined as follows.

$$\begin{aligned}
V\left(\sum_{j=1}^N A_j X_j\right) &= E\left[V\left(\sum_{j=1}^N A_j X_j \mid S_{n^*}^2\right)\right] + V\left[E\left(\sum_{j=1}^N A_j X_j \mid S_{n^*}^2\right)\right] \\
&= E\left(\sum_{j=1}^N A_j^2 V(X_j) \mid S_{n^*}^2\right) + V\left(\sum_{j=1}^N A_j E(X_j) \mid S_{n^*}^2\right) \\
&= E\left(\sum_{j=1}^N A_j^2 \sigma^2 \mid S_{n^*}^2\right) + V\left(\sum_{j=1}^N A_j \mu \mid S_{n^*}^2\right) \\
&= E\left(\frac{u^* \sigma^2}{S_{n^*}^2} \mid S_{n^*}^2\right) + V(\mu \mid S_{n^*}^2) \\
&= u^* \sigma^2 / S_{n^*}^2.
\end{aligned}$$

We can now write

$$\begin{aligned}
T_0 &= \frac{\frac{\sum_{j=1}^N A_j X_j - \mu}{\sqrt{u^* \sigma^2 / S_{n^*}^2}}}{\frac{\sqrt{u^*}}{\sqrt{u^* \sigma^2 / S_{n^*}^2}}} = \frac{\frac{\sum_{j=1}^N A_j X_j - \mu}{\sqrt{u^* \nu / W_{n^*}}}}{\frac{\sqrt{u^*}}{\sqrt{u^* \nu / W_{n^*}}}} \\
&= \frac{Z}{\sqrt{W_{n^*} / \nu}},
\end{aligned}$$

where

$$Z = \frac{\sum_{j=1}^N A_j X_j - \mu}{\sqrt{u^* \nu / W_{n^*}}} \text{ and } W_{n^*} = \frac{\nu S_{n^*}^2}{\sigma^2} \sim \chi_\nu^2 \text{ with } \nu = n^* - 1.$$

The cdf $F_{T_0}(t)$ of the distribution of T_0 is determined as follows.

$$\begin{aligned}
F_{T_0}(t) &= \int_{-\infty}^t f_{T_0}(q) dq = \int_{-\infty}^t \int_0^\infty f_{T_0, W_{n^*}}(q, w) dw dq \\
&= \int_{-\infty}^t \int_0^\infty f_{T_0 | W_{n^*}}(q | w) f_{W_{n^*}}(w) dw dq.
\end{aligned}$$

Now observe that the conditional distribution of T_0 given W is a normal distribution with mean 0 and variance ν/w since the conditional distribution of Z given W follows a standard normal distribution. We can now write

$$\begin{aligned}
F_{T_0}(t) &= \int_{-\infty}^t \int_0^\infty \frac{1}{\sqrt{2\pi} \sqrt{\nu/w}} e^{-\frac{1}{2} \left(\frac{q}{\sqrt{\nu/w}}\right)^2} \frac{1}{\Gamma\left(\frac{\nu}{2}\right) 2^{\nu/2}} w^{\nu/2-1} e^{-w/2} dw dq \\
&= \int_{-\infty}^t \int_0^\infty \frac{1}{\sqrt{\nu\pi} \Gamma\left(\frac{\nu}{2}\right) 2^{(\nu+1)/2}} w^{(\nu+1)/2-1} e^{-(1+q^2/\nu)w/2} dw dq.
\end{aligned}$$

Making the change of variable $r = (1 + q^2/\nu)w$, we have

$$\begin{aligned} F_{T_0}(t) &= \int_{-\infty}^t \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)(1+q^2/\nu)^{(\nu+1)/2}} \\ &\quad \times \left(\int_0^\infty \frac{1}{\Gamma\left(\frac{\nu+1}{2}\right)2^{(\nu+1)/2}} r^{(\nu+1)/2-1} e^{-r/2} dr \right) dq \\ &= \int_{-\infty}^t \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)(1+q^2/\nu)^{(\nu+1)/2}} dq. \end{aligned}$$

This is the cdf of a t -distribution with $\nu = n^* - 1$. □

Theorem 5.2.2. *The size of the test that rejects $H_0 : \mu = \mu_0$ in favor of $H_a : \mu \neq \mu_0$ if $|T| \geq t_{n^*-1,0,\alpha/2}$ is α , where $T = \left(\sum_{j=1}^N A_j X_j - \mu_0\right) / \sqrt{u^*}$ and $t_{n^*-1,0,\alpha/2}$ is the $100(1 - \alpha/2)$ th percentile of a central t -distribution with $n^* - 1$ degrees of freedom. The power of the test is*

$$\pi = 1 - F_{t_{n^*-1}}\left(t_{n^*-1,0,\alpha/2} - \Delta/\sqrt{u^*}\right) + F_{t_{n^*-1}}\left(-t_{n^*-1,0,\alpha/2} - \Delta/\sqrt{u^*}\right),$$

where $\Delta = \mu - \mu_0$.

Proof. It is not difficult to see that the size of the test is α . Observe that

$$T = \frac{\sum_{j=1}^N A_j X_j - \mu}{\sqrt{u^*}} + \frac{\mu - \mu_0}{\sqrt{u^*}} = T_0 + \frac{\Delta}{\sqrt{u^*}}.$$

The power of the test can now be expressed as

$$\begin{aligned} \pi &= P\left(\left|T_0 + \frac{\Delta}{\sqrt{u^*}}\right| \geq t_{n^*-1,0,\alpha/2}\right) \\ &= 1 - P\left(-t_{n^*-1,0,\alpha/2} - \frac{\Delta}{\sqrt{u^*}} < T_0 < t_{n^*-1,0,\alpha/2} - \frac{\Delta}{\sqrt{u^*}}\right) \\ &= 1 - P\left(T_0 < t_{n^*-1,0,\alpha/2} - \frac{\Delta}{\sqrt{u^*}}\right) \\ &\quad + P\left(T_0 < -t_{n^*-1,0,\alpha/2} - \frac{\Delta}{\sqrt{u^*}}\right) \end{aligned}$$

for $\Delta \neq 0$. The results now follow. □

The method introduced by Stein (1945) has been used by various authors to develop adaptive procedures for comparing two population means. We provide a general outline for the procedures of those introduced by Chapman (1950), Prokof'yev and Shishkin (1974), and Dudewicz and Ahmed (1998). The procedures begin by selecting “a priori” the positive real number u^* and the integer $n^* \geq 2$. Samples of size n^* are selected from the intervention and control groups with respective measurements $X_{I,1}, \dots, X_{I,n^*}$ and $X_{C,1}, \dots, X_{C,n^*}$. These data and future data are assumed to be stochastically independent with $X_{I,j} \sim N(\mu_I, \sigma_I^2)$ and $X_{C,j} \sim N(\mu_C, \sigma_C^2)$ for $i = 1, 2, \dots$. The total samples sizes from each group are the observed values of the random variables

$$N_I = \max \left\{ \left[\frac{G_I}{u^*} \right] + 1, n^* + 1 \right\} \text{ and } N_C = \max \left\{ \left[\frac{G_C}{u^*} \right] + 1, n^* + 1 \right\},$$

where G_I and G_C are functions of the sample data $X_{I,1}, \dots, X_{I,n^*}$ and $X_{C,1}, \dots, X_{C,n^*}$. Next the observed values of the random variables $A_{I,1}, \dots, A_{I,N_I}$ and $A_{C,1}, \dots, A_{C,N_C}$ are to be selected based on the restrictions

$$\begin{aligned} A_{I,1} &= \dots = A_{I,n^*} \text{ and } A_{C,1} = \dots = A_{C,n^*}; \\ \sum_{j=1}^{N_I} A_{I,j} &= 1 \text{ and } \sum_{j=1}^{N_C} A_{C,j} = 1; \\ \sum_{j=1}^{N_I} A_{I,j}^2 &= \frac{u^*}{H_I} \text{ and } \sum_{j=1}^{N_C} A_{C,j}^2 = \frac{u^*}{H_C}, \end{aligned}$$

where H_I and H_C are functions of the sample data $X_{I,1}, \dots, X_{I,n^*}$ and $X_{C,1}, \dots, X_{C,n^*}$. Chapman (1950) further suggested letting

$$A_{I,n^*+1} = \dots = A_{I,N_I} \text{ and } A_{C,n^*+1} = \dots = A_{C,N_C}$$

Table 5.1: Behrens-Fisher Procedures

{t5.1}

Two-Stage	G_I	H_I
Procedure	G_C	H_C
Chapman	S_{I,n^*}^2	G_I
(1950)	S_{C,n^*}^2	G_C
Prokof'yev and	$S_{Z,n^*}^2 \dagger$	u^*N
Shishkin(1974)	S_{Z,n^*}^2	u^*N
Dudewicz and	$S_{I,n^*} (S_{I,n^*} + S_{C,n^*})$	G_I
Ahmed(1998)	$S_{C,n^*} (S_{I,n^*} + S_{C,n^*})$	G_C
	$\dagger S_{Z,n^*}^2 = \frac{1}{n^*-1} \sum_{j=1}^{n^*} (X_{I,j} - X_{C,j} - \bar{X}_{I,n^*} + \bar{X}_{C,n^*})^2$	

to eliminate the arbitrariness involved in their selection. It is easy to show that

$$\begin{aligned}
 A_{I,n^*+1} &= \frac{1}{N_I} \left(1 + \sqrt{\frac{n^* (N_I u^* - H_I)}{(N_I - n^*) H_I}} \right); \\
 A_{I,1} &= \frac{1 - (N_I - n^*) A_{I,n^*+1}}{n^*}; \\
 A_{C,n^*+1} &= \frac{1}{N_C} \left(1 + \sqrt{\frac{n^* (N_C u^* - H_C)}{(N_C - n^*) H_C}} \right); \\
 A_{C,1} &= \frac{1 - (N_C - n^*) A_{C,n^*+1}}{n^*}.
 \end{aligned}$$

The test rejects $H_0 : \mu_I = \mu_C$ in favor of $H_a : \mu_I \neq \mu_C$ if $|T| \geq t_{\alpha/2}^*$, where

$$T = \frac{\sum_{j=1}^{N_I} A_{I,j} X_{I,j} - \sum_{j=1}^{N_C} A_{C,j} X_{C,j}}{\sqrt{u^*}}$$

with $t_{\alpha/2}^*$ is the 100(1 - $\alpha/2$)th percentile of the null distribution of T . Table 5.1 gives the values of G_I , G_C , H_I , and H_C for the procedures introduced by Chapman (1950), Prokof'yev and Shishkin (1974), and Dudewicz and Ahmed (1998).

Because of the restrictions on the $A_{I,j}$'s, we have that

$$\begin{aligned}
E\left(\sum_{j=1}^{N_I} A_{I,j} X_{I,j}\right) &= E\left[E\left(\sum_{j=1}^{N_I} A_{I,j} X_{I,j} \mid N_I, A_{I,j}, \dots, A_{I,N_I}\right)\right] \\
&= E\left(\sum_{j=1}^{N_I} A_{I,j} E(X_{I,j}) \mid N_I, A_{I,j}, \dots, A_{I,N_I}\right) \\
&= E\left(\sum_{j=1}^{N_I} A_{I,j} \mu_I \mid N_I, A_{I,j}, \dots, A_{I,N_I}\right) \\
&= E(\mu_I \mid N_I, A_{I,j}, \dots, A_{I,N_I}) = \mu_I.
\end{aligned}$$

Similarly, one can show that

$$E\left(\sum_{j=1}^{N_C} A_{C,j} X_{C,j}\right) = \mu_C.$$

We see that T can be expressed as

$$\begin{aligned}
T &= \frac{\left(\sum_{j=1}^{N_I} A_{I,j} X_{I,j} - \sum_{j=1}^{N_C} A_{C,j} X_{C,j}\right) - (\mu_I - \mu_C)}{\sqrt{u^*}} + \frac{\mu_I - \mu_C}{\sqrt{u^*}} \\
&= T_0 + \frac{\Delta}{\sqrt{u^*}},
\end{aligned}$$

where

$$\begin{aligned}
T_0 &= \frac{\left(\sum_{j=1}^{N_I} A_{I,j} X_{I,j} - \sum_{j=1}^{N_C} A_{C,j} X_{C,j}\right) - (\mu_I - \mu_C)}{\sqrt{u^*}} \text{ and} \\
\Delta &= \mu_I - \mu_C.
\end{aligned}$$

The size of the test has been selected to be α with critical value $t_{\alpha/2}^*$ and the power of the test is determined by

$$\pi = 1 - F_{T_0}\left(t_{\alpha/2}^* - \frac{\Delta}{\sqrt{u^*}}\right) + F_{T_0}\left(-t_{\alpha/2}^* - \frac{\Delta}{\sqrt{u^*}}\right).$$

It is desirable select the critical value $t_{\alpha/2}^*$ of the test to be the $100(1 - \alpha/2)$ th percentile of the distribution of T_0 , if possible. Clearly, it can be seen that as u^* is increased the power of the test increases as does the expected values of the sample sizes N_I and N_C (as one would expect).

For the Chapman (1950) procedure, we write T_0 as

$$T_0 = \frac{\sum_{j=1}^{N_I} A_{I,j} X_{I,j} - \mu_I}{\sqrt{u^*}} - \frac{\sum_{j=1}^{N_C} A_{C,j} X_{C,j} - \mu_C}{\sqrt{u^*}}.$$

Based on the definitions of $N_I, A_{I,1}, \dots, A_{I,N_I}$ and $N_C, A_{C,1}, \dots, A_{C,N_I}$, we have that

$$T_{I,0} = \frac{\sum_{j=1}^{N_I} A_{I,j} X_{I,j} - \mu_I}{\sqrt{u^*}} \sim t_{n^*-1} \text{ and } T_{C,0} = \frac{\sum_{j=1}^{N_C} A_{C,j} X_{C,j} - \mu_C}{\sqrt{u^*}} \sim t_{n^*-1}$$

and independent. These results follow from Theorem 2. Since $T_0 = T_{I,0} - T_{C,0}$, then one can express the cdf of T_0 by the convolution formula

$$\begin{aligned} F_{T_0}(t) &= P(T_{I,0} - T_{C,0} \leq t) = P(T_{I,0} \leq t + T_{C,0}) \\ &= \int_{-\infty}^{\infty} F_{t_{n^*-1}}(t+q) f_{t_{n^*-1}}(q) dq. \end{aligned}$$

(See Taneja and Dudewicz (1993)). Note that $F_{T_0}(t)$ is only a function of $n^* - 1$. This expression does not reduce to a simple expression and must be evaluated numerically. The value of $F_{T_0}(t) - 0.5$ for various values of t and $n^* - 1 = 2, 4, 6, 8, 10, 12$ are tabled in Chapman (1950). For this procedure, it is possible to select the critical value $t_{\alpha/2}^*$ of the test as the $100(1 - \alpha/2)$ th percentile of the distribution of T_0 .

The random variable T in the Prokof'yev and Shishkin (1974) procedure can be expressed as

$$T = \frac{\bar{X}_{I,N} - \bar{X}_{C,N}}{S_{Z,n^*}/\sqrt{N}}.$$

It is not difficult to show that $T \sim t_{n^*-1} + \Delta/\sqrt{u^*}$ using the results of Theorem 1.

For this test, we have

$$t_{\alpha/2}^* = t_{n^*-1, \alpha/2}.$$

The power of the test for $\Delta \neq 0$ is

$$\begin{aligned} \pi &= 1 - F_{t_{n^*-1}}\left(t_{n^*-1, \alpha/2} - \frac{\Delta}{\sqrt{u^*}}\right) \\ &\quad + F_{t_{n^*-1}}\left(-t_{n^*-1, \alpha/2} - \frac{\Delta}{\sqrt{u^*}}\right). \end{aligned}$$

Taneja and Dudewicz (1993) compared the “solutions of the Behrens-Fisher problem” by Chapman (1950) and Prokof’yev and Shishkin (1974). They recommended the procedure by Chapman (1950).

Dudewicz and Ahmed (1998) stated the following theorem concerning the distribution of T_0 . We provide a proof of this theorem with some added results.

Theorem 5.2.3. *The cdf of T_0 can be expressed as*

$$F_{T_0}(t) = \int_0^\infty 2qF_{t_2(n^*-1)} \left(t \sqrt{\frac{2q(\lambda q + 1)}{(q + \lambda)(1 + q^2)}} \right) f_{F_{n^*-1, n^*-1}}(q^2) dq,$$

where $F_{t_2(n^*-1)}(r)$ and $f_{F_{n^*-1, n^*-1}}(r)$ are, respectively, the cdf and pdf function of a t -distribution with $2(n^* - 1)$ and an F -distribution with numerator and denominator degrees of freedom both $n^* - 1$. The distribution of the random variable T_0 depends only on the values n^* and $\lambda = \sigma_I/\sigma_C$ but not on the value of u^* .

Proof. It is convenient to let

$$D = \sum_{j=1}^{N_I} A_{I,j} X_{I,j} - \sum_{j=1}^{N_C} A_{C,j} X_{C,j}.$$

Observe that

$$\begin{aligned} \mu_D &= E \left(\sum_{j=1}^{N_I} A_{I,j} X_{I,j} - \sum_{j=1}^{N_C} A_{C,j} X_{C,j} \right) \\ &= E \left[E \left(\sum_{j=1}^{N_I} A_{I,j} X_{I,j} - \sum_{j=1}^{N_C} A_{C,j} X_{C,j} \mid S_{I,n^*}, S_{C,n^*} \right) \right] \\ &= E \left(\sum_{j=1}^{N_I} A_{I,j} E(X_{I,j} \mid S_{I,n^*}, S_{C,n^*}) \right) \\ &\quad - E \left(\sum_{j=1}^{N_C} A_{C,j} E(X_{C,j} \mid S_{I,n^*}, S_{C,n^*}) \right) \\ &= E \left(\left(\sum_{j=1}^{N_I} A_{I,j} \right) \mu_I \mid S_{I,n^*}, S_{C,n^*} \right) \\ &\quad - E \left(\left(\sum_{j=1}^{N_C} A_{C,j} \right) \mu_C \mid S_{I,n^*}, S_{C,n^*} \right) \\ &= E(\mu_I \mid S_{I,n^*}, S_{C,n^*}) - E(\mu_C \mid S_{I,n^*}, S_{C,n^*}) \\ &= \mu_I - \mu_C. \end{aligned}$$

The variance of D is given by

$$\begin{aligned}
\sigma_D^2 &= V \left(\sum_{j=1}^{N_I} A_{I,j} X_{I,j} - \sum_{j=1}^{N_C} A_{C,j} X_{C,j} \right) \\
&= E \left[V \left(\sum_{j=1}^{N_I} A_{I,j} X_{I,j} - \sum_{j=1}^{N_C} A_{C,j} X_{C,j} \mid S_{I,n^*}, S_{C,n^*} \right) \right] \\
&\quad + V \left[E \left(\sum_{j=1}^{N_I} A_{I,j} X_{I,j} - \sum_{j=1}^{N_C} A_{C,j} X_{C,j} \mid S_{I,n^*}, S_{C,n^*} \right) \right] \\
&= E \left(\sum_{j=1}^{N_I} A_{I,j}^2 V(X_{I,j}) + \sum_{j=1}^{N_C} A_{C,j}^2 V(X_{C,j}) \mid S_{I,n^*}, S_{C,n^*} \right) \\
&\quad + V(\mu_I - \mu_C \mid S_{I,n^*}, S_{C,n^*}) \\
&= \frac{u^* \sigma_I^2}{S_{I,n^*} (S_{I,n^*} + S_{C,n^*})} + \frac{u^* \sigma_C^2}{S_{C,n^*} (S_{I,n^*} + S_{C,n^*})} \\
&= u \frac{\sigma_C^2 S_{I,n^*} + \sigma_I^2 S_{C,n^*}}{S_{I,n^*} S_{C,n^*} (S_{I,n^*} + S_{C,n^*})}.
\end{aligned}$$

We can now express T_0 as

$$T_0 = \frac{\frac{D - (\mu_I - \mu_C)}{\sqrt{\frac{u(\sigma_C^2 S_{I,n^*} + \sigma_I^2 S_{C,n^*})}{S_{I,n^*} S_{C,n^*} (S_{I,n^*} + S_{C,n^*})}}}{\frac{\sqrt{u}}{\sqrt{u \frac{\sigma_C^2 S_{I,n^*} + \sigma_I^2 S_{C,n^*}}{S_{I,n^*} S_{C,n^*} (S_{I,n^*} + S_{C,n^*})}}}} = \frac{Z}{\sqrt{\frac{S_{I,n^*} S_{C,n^*} (S_{I,n^*} + S_{C,n^*})}{\sigma_C^2 S_{I,n^*} + \sigma_I^2 S_{C,n^*}}}},$$

where

$$Z = \frac{D - (\mu_I - \mu_C)}{\sqrt{\frac{u(\sigma_C^2 S_{I,n^*} + \sigma_I^2 S_{C,n^*})}{S_{I,n^*} S_{C,n^*} (S_{I,n^*} + S_{C,n^*})}}.$$

The conditional distribution of the random variables Z and T_0 given S_{I,n^*} and S_{C,n^*} are respectively, a standard normal distribution and a normal distribution mean zero and variance

$$\sigma_{T_0 \mid S_{I,n^*}, S_{C,n^*}}^2 = \frac{\sigma_C^2 S_{I,n^*} + \sigma_I^2 S_{C,n^*}}{S_{I,n^*} S_{C,n^*} (S_{I,n^*} + S_{C,n^*})}.$$

The cumulative distribution function of T_0 can be expressed as

$$F_{T_0}(t) = \int_0^\infty \int_0^\infty \Phi \left(t \sqrt{\frac{s_I s_C (s_I + s_C)}{\sigma_C^2 s_I + \sigma_I^2 s_C}} \right) f_{s_I}(s_I) f_{s_C}(s_C) ds_C ds_I.$$

Making the change of variables $q_I = s_I/\sigma_I$ and $q_C = s_C/\sigma_C$ with Jacobian $J = \sigma_I\sigma_C$, we have

$$F_{T_0}(t) = \int_0^\infty \int_0^\infty \Phi \left(t \sqrt{\frac{\sigma_I\sigma_C q_I q_C (\sigma_I q_I + \sigma_C q_C)}{\sigma_I\sigma_C^2 q_I + \sigma_I^2 \sigma_C q_C}} \right) \\ \times \sigma_I f_{S_I}(\sigma_I q_I) \sigma_C f_{S_C}(\sigma_C q_C) dq_C dq_I.$$

Next we observe that $F_{S_I}(\sigma_I q_I) = F_{\chi_{n^*-1}^2}((n^*-1)q_I^2)$, hence

$$\begin{aligned} \sigma_I f_{S_I}(\sigma_I q_I) &= 2(n^*-1) q_I f_{\chi_{n^*-1}^2}((n^*-1)q_I^2) \\ &= \frac{2(n^*-1)^{(n^*-1)/2}}{\Gamma\left(\frac{n^*-1}{2}\right) 2^{(n^*-1)/2}} q_I (q_I^2)^{(n^*-1)/2-1} e^{-(n^*-1)q_I^2/2}. \end{aligned}$$

Similarly, we have that

$$\sigma_C f_{S_C}(\sigma_C q_C) = \frac{2(n^*-1)^{(n^*-1)/2}}{\Gamma\left(\frac{n^*-1}{2}\right) 2^{(n^*-1)/2}} q_C (q_C^2)^{(n^*-1)/2-1} e^{-(n^*-1)q_C^2/2}.$$

Hence, we have

$$F_{T_0}(t) = \int_0^\infty \int_0^\infty \Phi \left(t \sqrt{\frac{\sigma_I\sigma_C q_I q_C (\sigma_I q_I + \sigma_C q_C)}{\sigma_I\sigma_C^2 q_I + \sigma_I^2 \sigma_C q_C}} \right) \\ \times \frac{2(n^*-1)^{(n^*-1)/2}}{\Gamma\left(\frac{n^*-1}{2}\right) 2^{(n^*-1)/2}} q_I (q_I^2)^{(n^*-1)/2-1} e^{-(n^*-1)q_I^2/2} \\ \times \frac{2(n^*-1)^{(n^*-1)/2}}{\Gamma\left(\frac{n^*-1}{2}\right) 2^{(n^*-1)/2}} q_C (q_C^2)^{(n^*-1)/2-1} e^{-(n^*-1)q_C^2/2} \\ \times dq_C dq_I.$$

Making the change of variables $q_I = qq_C$ and $q_C = q_C$ with $J = q_C$, we have

$$F_{T_0}(t) = \int_0^\infty \int_0^\infty \Phi \left(t \sqrt{\frac{\sigma_I\sigma_C qq_C q_C (\sigma_I qq_C + \sigma_C q_C)}{\sigma_I\sigma_C^2 qq_C + \sigma_I^2 \sigma_C q_C}} \right) \\ \times \frac{2(n^*-1)^{(n^*-1)/2}}{\Gamma\left(\frac{n^*-1}{2}\right) 2^{(n^*-1)/2}} qq_C (q^2 q_C^2)^{(n^*-1)/2-1} e^{-(n^*-1)q^2 q_C^2/2} \\ \times \frac{2(n^*-1)^{(n^*-1)/2}}{\Gamma\left(\frac{n^*-1}{2}\right) 2^{(n^*-1)/2}} q_C (q_C^2)^{(n^*-1)/2-1} e^{-(n^*-1)q_C^2/2} \\ \times q_C dq_C dq.$$

Simplifying and rearranging terms, we can write

$$\begin{aligned}
F_{T_0}(t) &= \int_0^\infty \int_0^\infty 2q\Phi\left(t\sqrt{\frac{q(\lambda q + 1)}{q + \lambda}q_C^2}\right) \\
&\quad \times \frac{2(n^* - 1)^{(n^* - 1)}}{\Gamma\left(\frac{2(n^* - 1)}{2}\right) 2^{2(n^* - 1)/2}} (q_C^2)^{2(n^* - 1)/2 - 1} e^{-(n^* - 1)(1 + q^2)q_C^2/2} q_C dq_C \\
&\quad \times \frac{\Gamma\left(\frac{2(n^* - 1)}{2}\right)}{\Gamma\left(\frac{n^* - 1}{2}\right) \Gamma\left(\frac{n^* - 1}{2}\right)} (q^2)^{(n^* - 1)/2 - 1} dq.
\end{aligned}$$

Now make the change of variable

$$q_C^2 = \frac{w}{(n^* - 1)(1 + q^2)} \text{ with } q_C dq_C = \frac{dw}{2(n^* - 1)(1 + q^2)}.$$

It follows that

$$\begin{aligned}
F_{T_0}(t) &= \int_0^\infty \int_0^\infty 2q\Phi\left(t\sqrt{\frac{2q(\lambda q + 1)}{(q + \lambda)(1 + q^2)} \frac{w}{2(n^* - 1)}}\right) \\
&\quad \times \frac{1}{\Gamma\left(\frac{2(n^* - 1)}{2}\right) 2^{2(n^* - 1)/2}} w^{2(n^* - 1)/2 - 1} e^{-w/2} dw \\
&\quad \times \frac{\Gamma\left(\frac{2(n^* - 1)}{2}\right)}{\Gamma\left(\frac{n^* - 1}{2}\right) \Gamma\left(\frac{n^* - 1}{2}\right)} (q^2)^{(n^* - 1)/2 - 1} (1 + q^2)^{-2(n^* - 1)/2} dq.
\end{aligned}$$

Therefore, we have

$$F_{T_0}(t) = \int_0^\infty 2qF_{t_{2(n^* - 1)}}\left(t\sqrt{\frac{2q(\lambda q + 1)}{(q + \lambda)(1 + q^2)}}\right) f_{F_{n^* - 1, n^* - 1}}(q^2) dq.$$

□

Theorem 5.2.4. *The cdf of T is given by*

$$F_T(t) = F_{T_0}\left(t - \Delta/\sqrt{u^*}\right),$$

where $\Delta = \mu_I - \mu_C$.

Theorem 5.2.5. *The following hold for the cdf of T_0 :*

$$\begin{aligned} \lim_{\lambda \rightarrow 0} F_{T_0}(t) &= G_{n^*,0}(t) = \int_0^\infty 2qF_{t_2(n^*-1)}\left(t\sqrt{\frac{2}{1+q^2}}\right) f_{F_{n^*-1,n^*-1}}(q^2) dq \text{ and} \\ \lim_{\lambda \rightarrow \infty} F_{T_0}(t) &= G_{n^*,\infty}(t) = \int_0^\infty 2qF_{t_2(n^*-1)}\left(t\sqrt{\frac{2q^2}{1+q^2}}\right) f_{F_{n^*-1,n^*-1}}(q^2) dq. \end{aligned}$$

The test suggested by Dudewicz and Ahmed (1998) is given in the following theorem.

Theorem 5.2.6. *A test of size α rejects $H_0 : \mu_I = \mu_C$ in favor of $H_a : \mu_I \neq \mu_C$ if $|T| \geq h_{n^*,\alpha/2}$, where $h_{n^*,\alpha/2}$ is the solution to the equation*

$$\inf_{\lambda > 0} \{h | F_{T_0}(h | n^*, \lambda) = 1 - \alpha/2\}.$$

Proof. At the writing of this thesis, we have not been able to prove this theorem. \square

An equivalent form of the test of Dudewicz and Ahmed (1998) is presented in the following theorem.

Theorem 5.2.7. *A test of size α rejects $H_0 : \mu_I = \mu_C$ in favor of $H_a : \mu_I \neq \mu_C$ if $|T| \geq h_{n^*,\alpha/2}$, where $h_{n^*,\alpha/2}$ is the solution to the equation*

$$G_{n^*,\infty}(t) = \int_0^\infty 2qF_{t_2(n^*-1)}\left(t\sqrt{\frac{2q^2}{1+q^2}}\right) f_{F_{n^*-1,n^*-1}}(q^2) dq = 1 - \alpha/2.$$

Proof. At the writing of this thesis, we have not been able to prove this theorem although the solutions to the equation $G_{n^*,\infty}(h) = \int_0^\infty 2qF_{t_2(n^*-1)}\left(h\sqrt{\frac{2q^2}{1+q^2}}\right) f_{F_{n^*-1,n^*-1}}(q^2) dq = 1 - \alpha/2$ are those values tabled in Table I of Dudewicz and Ahmed (1998). \square

The size of the test is α and the power of the test is determined by

$$\pi = 1 - F_{T_0}\left(h_{n^*,\alpha/2} - \Delta/\sqrt{u^*}\right) + F_{T_0}\left(-h_{n^*,\alpha/2} - \Delta/\sqrt{u^*}\right).$$

At the writing of this thesis, we have not been able to show that there is a solution $h_{n^*,\alpha/2}$ to the equation

$$\inf_{\lambda>0} F_{T_0}(t|n^*, \lambda) = 1 - \alpha/2.$$

Consequently, it is not clear that their claim is true. They go on to state that this procedure is asymptotic optimal in the sense that it “achieves asymptotically what the fixed sample does with λ known.”

Proschan (2005) described two-stage sample prediction procedure to test for the difference between means assuming that the population variances are equal. A modified version of this procedure has the researcher selecting

$$u^* = \frac{(\Delta^*)^2}{2 \left(t_{2(n_1-1),0,\alpha/2} + t_{2(n_1-1),0,\beta} \right)^2}$$

and $n^* \geq 2$. The values α^* , β^* , and $\Delta^* > 0$ are to be specified by the researcher, where Δ^* is a value chosen such that for all $\Delta = \mu_I - \mu_C$ the power of the test is at least $1 - \beta^*$ for $|\Delta| \geq \Delta^*$. The values of $A_{I,j} = A_{C,j} = 1/N$ with

$$\sum_{j=1}^N A_{I,j}^2 = \sum_{j=1}^N A_{C,j}^2 = \frac{u^*}{S_{p,n^*}^2},$$

where $S_{p,n^*} = (S_{I,n^*} + S_{C,n^*})/2$. The test rejects the null hypothesis of equal means in favor of the alternative hypothesis of unequal means if

$$\left| \frac{\bar{X}_{I,N} - \bar{X}_{C,N}}{\sqrt{2S_{p,n^*}^2/N}} \right| \geq t_{2(n^*-1),\alpha/2}.$$

The test is of size α and has power function

$$\begin{aligned} \pi &= 1 - F_{t_{n^*-1}} \left(t_{2(n^*-1),\alpha/2} - \frac{\Delta}{\sqrt{u^*}} \right) \\ &\quad + F_{t_{n^*-1}} \left(-t_{2(n^*-1),\alpha/2} - \frac{\Delta}{\sqrt{u^*}} \right). \end{aligned}$$

As discussed in Chapter 3, the sample size problem could only be solved if λ is known. For the case in which λ is not known, we propose a method for determining the intervention and control sample sizes. This method requires the researcher provide the following values: (1) the desired size α^* of the test; (2) an initial sample size n^* for both the intervention and the control groups; (3) a value $\Delta^* > 0$ of Δ ; (4) a value $1 - \beta^*$ such that the power of the test is at least this value for all $|\Delta| \geq \Delta^*$, and (5) a value $u^* > 0$. The proposed method is as follows. Obtain the observed values of the random sample of X measurements

$$X_{I,1}, X_{I,2}, \dots, X_{I,n^*} \text{ and } X_{C,1}, X_{C,2}, \dots, X_{C,n^*}$$

to be taken on individuals from the intervention and control groups, respectively. Determine the variances S_{I,n^*}^2 and S_{C,n^*}^2 from the respective sample values. Calculate the observed value of the estimator

$$L = u^* S_{I,n^*} / S_{C,n^*}$$

of λ . Find the sample sizes N_I and N_C such that total sample size $N = N_I + N_C$ is a minimum, $F_{T_2}(t_{0,L,N_I,N_C,\alpha^*/2} |\Delta^* / S_{C,n^*}, L, N_I, N_C)$ is a maximum over all pairs of sample sizes that have a total of N , and

$$F_{T_2}(t_{0,L,N_I,N_C,\alpha^*/2} |\Delta^* / S_{C,n^*}, L, N_I, N_C) \geq 1 - \beta^*.$$

The FORTRAN program given in Appendix II can be used to select the values of N_I and N_C . Note that the sample sizes N_I and N_C are random variables. Next select the test based on the observed values of the random variables $A_{I,1}, \dots, A_{I,N_I}$, $A_{C,1}, \dots, A_{C,N_C}$, and test statistic

$$T = \frac{\sum_{j=1}^{N_I} A_{I,j} X_{I,j} - \sum_{j=1}^{N_C} A_{C,j} X_{C,j}}{\sqrt{u^*}}$$

using one of the procedures by Chapman (1950), Prokof'yev and Shishkin (1974), or Dudewicz and Ahmed (1998).

No guidance is given by Stein (1945), Chapman (1950), Prokof'yev and Shishkin (1974), and Proschan (2005) in the selection of the sample size n^* . Seelbinder (1953) suggested for the procedure by Stein (1945) to select n^* such that $E(N)$ is minimized. Moshman (1958) proposed the use of an upper percentage point of the distribution of N in conjunction with $E(N)$ to guide in the selection of n^* when selecting a confidence interval for $\mu_I - \mu_C$.

For the procedure given in Stein (1945), the variability in the distribution of N is due to the variability the statistic $S_{n^*}^2$. It is not difficult to show, as previously stated, that

$$S_{n^*}^2 \sim \frac{\sigma^2}{n^* - 1} \chi_{n^* - 1}^2.$$

We can then express N as

$$N = \max \left\{ \left[\frac{\sigma^2 W / (n^* - 1)}{u^*} \right] + 1, n^* + 1 \right\},$$

where

$$W = \frac{(n^* - 1) S_{n^*}^2}{\sigma^2} \sim \chi_{n^* - 1}^2.$$

We see that the probability mass function describing the distribution of N is given by

$$P(N = n^* + k) = \begin{cases} F_W(a_1), & \text{for } k = 1; \\ F_W(a_k) - F_W(a_{k-1}) & \text{for } k > 1, \end{cases}$$

where

$$a_k = \frac{(n^* - 1)(n^* + k)u^*}{\sigma^2}.$$

Note that the distribution of the sample size is a function of u^* , n^* , and σ^2 .

There are several parameters of the distribution of N that would be of interest to the researcher. The most noted of these are the mean μ_N , the standard deviation σ_N , and the 100γ th percentage point $N_{1-\gamma}$ for various values of $0 < \gamma < 1$. As is well known, the mean of the distribution can be expressed as

$$\mu_N = E(N) = \sum_{k=0}^{\infty} (n^* + k) P(N = n^* + k).$$

We observe that we can write

$$\begin{aligned} \mu_N &= n^* \sum_{k=0}^{\infty} P(N = n^* + k) + \sum_{k=1}^{\infty} kP(N = n^* + k) \\ &= n^* + \sum_{k=1}^{\infty} kP(N = n^* + k). \end{aligned}$$

To obtain the standard deviation of the distribution of N , we first find $E(N^2)$ which can be expressed as

$$\mu_{N^2} = E(N^2) = \sum_{k=0}^{\infty} (n^* + k)^2 P(N = n^* + k).$$

Expanding the term $(n^* + k)^2$ and simplifying, we have

$$\mu_{N^2} = n^* (2\mu_N - n^*) + \sum_{k=0}^{\infty} k^2 P(N = n^* + k).$$

It then follows that

$$\begin{aligned} \sigma_N^2 &= E(N^2) - (E(N))^2 \\ &= \sum_{k=0}^{\infty} (n^* + k)^2 P(N = n^* + k) \\ &\quad - \left(\sum_{k=0}^{\infty} (n^* + k) P(N = n^* + k) \right)^2 \\ &= n^* (2\mu_N - n^*) + \sum_{k=0}^{\infty} k^2 P(N = n^* + k) \\ &\quad - \left(n^* + \sum_{k=1}^{\infty} kP(N = n^* + k) \right)^2. \end{aligned}$$

The 100γ th percentage point $N_{1-\gamma}$ can be determined by

$$P(N < N_{1-\gamma}) < \gamma \leq P(N \leq N_{1-\gamma}).$$

It follows that

$$\sum_{k=0}^{N_{1-\gamma}-1} P(N = n^* + k) < \gamma \leq \sum_{k=0}^{N_{1-\gamma}} P(N = n^* + k).$$

Under the assumption that the null hypothesis holds, it may be of interest to the researcher to obtain a prediction interval for N . One method for obtaining a $100(1 - \alpha)\%$ prediction interval for N is the interval

$$(N_{1-\tau}, N_{\alpha-\tau}],$$

where $0 < \tau < \alpha$.

5.3 Paired Data

In the paired data case, a test of size α^* (specified by the researcher) to be use to test the hypotheses $H_0 : \mu_D = 0$ versus $H_a : \mu_D \neq 0$ rejects H_0 if $|T| \geq t_{n-1, \alpha^*/2}$, where

$$T = \frac{\bar{D}}{S/\sqrt{n}}.$$

First we consider selecting the sample size n if the researcher can provide values $1 - \beta^*$ (the desired minimum power of the test) and $\delta_D^* > 0$ of $\delta_D = \mu_D/\sigma_D$ such that for all values of δ_D such that $|\delta_D| \geq \delta_D^*$ the power of the test is at least $1 - \beta^*$. In this case, we can represent the power function by

$$\pi(\alpha^*, \delta_D, n) = 1 - F_{t_{n-1, \sqrt{n}\delta_D}}(t_{n-1, 0, \alpha^*/2}) + F_{t_{n-1, \sqrt{n}\delta_D}}(-t_{n-1, 0, \alpha^*/2}).$$

It is not difficult to show that the desired sample size n is the smallest positive integer that satisfies the inequality

$$\pi(\alpha^*, \delta_D^*, n) \geq 1 - \beta^*.$$

In general, the power function is a function of α , μ_D , σ_D , and n . If the researcher can specify a value $\mu_D^* > 0$ such that

$$\pi(\alpha^*, \mu_D^*, \sigma_D, n) \geq 1 - \beta^*,$$

then the desired sample size is the minimum value of n that satisfies this inequality. The problem with this method is that it depends on the unknown value σ_D .

Using the procedure by Stein (1945), the researcher would select the positive real number u^* and an initial sample size $n^* \geq 2$. Using the estimator S_{D,n^*}^2 from this initial sample values D_1, \dots, D_{n^*} to estimate σ_D^2 , we predict the total sample size N by

$$N = \max \left\{ \left[\frac{S_{D,n^*}^2}{u^*} \right] + 1, n^* + 1 \right\}.$$

Assume the measurements D_1, \dots, D_{n^*} is a random sample from a $N(\mu_D, \sigma_D^2)$ distribution. Let $S_{n^*}^2$ denote the variance of this sample. Using these results the random variables A_1, \dots, A_N are selected subject to the restrictions

$$\sum_{j=1}^N A_j = 1; A_1 = \dots = A_{n^*}; \text{ and } \sum_{j=1}^N A_j^2 = \frac{u^*}{S_{D,n^*}^2}.$$

The test rejects $H_0 : \mu_D = 0$ in favor of $H_a : \mu_D \neq 0$ if $|T| \geq t_{n^*-1, \alpha/2}$, where

$$T = \frac{\sum_{j=1}^N A_j D_j}{\sqrt{u^*}}.$$

This test has size α and power function given by

$$\begin{aligned} \pi &= 1 - F_{t_{n^*-1}} \left(t_{n^*-1, \alpha/2} - \frac{\mu_D^*}{\sqrt{u^*}} \right) \\ &\quad + F_{t_{n^*-1}} \left(-t_{n^*-1, \alpha/2} - \frac{\mu_D^*}{\sqrt{u^*}} \right). \end{aligned}$$

5.4 Multistage Stage Adaptive Methods

The hypotheses to be tested are

$$H_0 : \mu_D = 0 \text{ versus } H_0 : \mu_D \neq 0.$$

The data will be available in the form of two random samples $D_{1,1}, \dots, D_{1,n_1}$ and $D_{2,1}, \dots, D_{2,n_2}$ with respective means \bar{D}_1 and \bar{D}_2 . The first sample will be used to decide (1) to fail to reject H_0 , (2) reject H_0 , or (3) observe the second sample. If the measurements on the second sample are taken, then the samples are combined and this information is used to decide to either fail to reject the null hypothesis or reject it. The first test we will consider assumes that σ_D is known. The null hypothesis is rejected if

$$(1) |T_1| \geq z_{\alpha_1/2} \text{ or } (2) z_{\alpha_0/2} \leq |T_1| < z_{\alpha_1/2} \text{ and } |T| \geq z_{\alpha_2/2},$$

where

$$T_1 = \frac{\bar{D}_1}{\sigma_D/\sqrt{n_1}} \text{ and } T = \frac{\bar{D}}{\sigma_D/\sqrt{n}} \text{ with } \bar{D} = \frac{n_1\bar{D}_1 + n_2\bar{D}_2}{n}.$$

Here $n = n_1 + n_2$. This testing method is a two-stage sampling method. The size α of the test is determined by

$$\begin{aligned} \alpha &= P\left(\left|\frac{\bar{D}_1}{\sigma_D/\sqrt{n_1}}\right| \geq z_{\alpha_1/2} \mid \mu_D = 0\right) \\ &\quad + P\left(z_{\alpha_0/2} \leq \left|\frac{\bar{D}_1}{\sigma_D/\sqrt{n_1}}\right| < z_{\alpha_1/2}, \left|\frac{\bar{D}}{\sigma_D/\sqrt{n}}\right| \geq z_{\alpha_2/2} \mid \mu_D = 0\right). \end{aligned}$$

The power of the test is determined by

$$\begin{aligned} \pi &= P\left(\left|\frac{\bar{D}_1}{\sigma_D/\sqrt{n_1}}\right| \geq z_{\alpha_1/2} \mid \mu_D \neq 0\right) \\ &\quad + P\left(z_{\alpha_0/2} \leq \left|\frac{\bar{D}_1}{\sigma_D/\sqrt{n_1}}\right| < z_{\alpha_1/2}, \left|\frac{\bar{D}}{\sigma_D/\sqrt{n}}\right| \geq z_{\alpha_2/2} \mid \mu_D \neq 0\right). \end{aligned}$$

It is of interest to be able to determine the size and power of the test. To do so we examine the statistics T_1 and T . We see that we can express T_1 as

$$T_1 = \frac{\bar{D}_1 - \mu_D}{\sigma_D/\sqrt{n_1}} + \sqrt{n_1} \frac{\mu_D}{\sigma_D} = Z_1 + \sqrt{n_1} \delta_D,$$

where

$$Z_1 = \frac{\bar{D}_1 - \mu_D}{\sigma_D/\sqrt{n_1}} \text{ and } \delta_D = \frac{\mu_D}{\sigma_D}.$$

The random variable T can be expressed as

$$\begin{aligned} T &= \sqrt{\frac{n_1}{n}} \frac{\bar{D}_1 - \mu_D}{\sigma_D/\sqrt{n_1}} + \sqrt{\frac{n_2}{n}} \frac{\bar{D}_2 - \mu_D}{\sigma_D/\sqrt{n_2}} + \sqrt{n} \frac{\mu_D}{\sigma_D} \\ &= \sqrt{\frac{n_1}{n}} Z_1 + \sqrt{\frac{n_2}{n}} Z_2 + \sqrt{n} \delta_D, \end{aligned}$$

where

$$Z_2 = \frac{\bar{D}_2 - \mu_D}{\sigma_D/\sqrt{n_2}}.$$

Further, we observe that we can express T as

$$T = \sqrt{\frac{n_2}{n}} \left(Z_2 + \sqrt{\frac{n_1}{n_2}} T_1 + \frac{n}{\sqrt{n_2}} \delta \right).$$

We can now express the power function as

$$\begin{aligned} \pi &= 1 - \Phi(z_{\alpha_1/2} - \sqrt{n_1} \delta_D) + \Phi(-z_{\alpha_1/2} - \sqrt{n_1} \delta_D) \\ &\quad - \int_{z_{\alpha_0/2} - \sqrt{n_1} \delta_D}^{z_{\alpha_1/2} - \sqrt{n_1} \delta_D} \left(\Phi\left(\frac{\sqrt{n} z_{\alpha_2/2} - h}{\sqrt{n_2}}\right) - \Phi\left(\frac{-\sqrt{n} z_{\alpha_2/2} - h}{\sqrt{n_2}}\right) \right) \\ &\quad \times \phi(t_1) dt_1 \\ &\quad - \int_{-z_{\alpha_1/2} - \sqrt{n_1} \delta_D}^{-z_{\alpha_0/2} - \sqrt{n_1} \delta_D} \left(\Phi\left(\frac{\sqrt{n} z_{\alpha_2/2} - h}{\sqrt{n_2}}\right) - \Phi\left(\frac{-\sqrt{n} z_{\alpha_2/2} - h}{\sqrt{n_2}}\right) \right) \\ &\quad \times \phi(t_1) dt_1, \end{aligned}$$

where $h = \sqrt{n_1}t_1 + n\delta_D$. The size α is equal to π when $\delta_D = 0$. Hence, we can write

$$\begin{aligned} \alpha &= \alpha_1 + \\ &\quad - \int_{z_{\alpha_0/2}}^{z_{\alpha_1/2}} \left(\Phi \left(\frac{\sqrt{n}z_{\alpha_2/2} - \sqrt{n_1}t_1}{\sqrt{n_2}} \right) - \Phi \left(\frac{-\sqrt{n}z_{\alpha_2/2} - \sqrt{n_1}t_1}{\sqrt{n_2}} \right) \right) \\ &\quad \times \phi(t_1) dt_1 \\ &\quad - \int_{-z_{\alpha_1/2}}^{-z_{\alpha_0/2}} \left(\Phi \left(\frac{\sqrt{n}z_{\alpha_2/2} - \sqrt{n_1}t_1}{\sqrt{n_2}} \right) - \Phi \left(\frac{-\sqrt{n}z_{\alpha_2/2} - \sqrt{n_1}t_1}{\sqrt{n_2}} \right) \right) \\ &\quad \times \phi(t_1) dt_1. \end{aligned}$$

A more general setting of this problem allows for upto K samples to be examined with a decision to fail to reject or reject the null hypothesis to be made at sampling stage K . A even more general test does not fix K but allows the data to determine the value of K . This multistage method can be designed as follows. At sampling stage k , a decision to reject H_0 is made if

$$\bigcap_{i=1}^{k-1} \{z_{\alpha_i,0/2} \leq |T_{n_1,\dots,n_i}| < z_{\alpha_i,1/2}\} \text{ and } |T_{n_1,\dots,n_i}| \geq z_{\alpha_k,2/2},$$

where

$$T_{n_1,\dots,n_i} = \frac{\bar{D}_{n_1,\dots,n_i}}{\sigma_D/\sqrt{n_1 + \dots + n_i}} \text{ with } \bar{D}_{n_1,\dots,n_i} = \frac{n_1\bar{D}_1 + \dots + n_i\bar{D}_i}{n_1 + \dots + n_i}$$

for $i = 1, 2, 3, \dots$. For the case in which σ_D is unknown, we would replace $z_{\alpha_i,0/2}$ with $t_{\alpha_i,j/2}$ for $j = 0, 1, 2$ and

$$T_{n_1,\dots,n_i} = \frac{\bar{D}_{n_1,\dots,n_i}}{S_{n_1,\dots,n_i,p}/\sqrt{n_1 + \dots + n_i}},$$

where

$$S_{n_1,\dots,n_i,p}^2 = \frac{\sum_{j=1}^i (n_j - 1) S_{n_j}^2}{\sum_{j=1}^i (n_j - 1)}.$$

5.5 Group Sequential Methods

Although Elfring and Schultz (1973), McPherson (1974) and Canner (1977) provided some of the earliest sequential medical studies, the methods proposed by Pocock (1977), O'Brien and Fleming (1979), and Wang and Tsiatis (1987) are typically the ones most often cited. A family of group sequential tests that include (see Table 5.1) those of Pocock (1977), O'Brien and Fleming (1979), and Wang and Tsiatis (1987) can be described as follows. Measurements are to be taken sequentially on samples from the intervention and control in groups each of sizes $n_1, n_2, \dots, n_k, \dots$. We denote the measurements taken at time k by

$$X_{I,k,1}, \dots, X_{I,k,n_k}, X_{C,k,1}, \dots, X_{C,k,n_k}$$

for $k = 1, 2, 3, \dots$. Our test assuming $\sigma_I^2 = \sigma_C^2 = \sigma^2$ is a sequence of decision rules based on the sequence of statistics $\{T_{(k)}\}$ with $T_{(k)}$ defined in general by

$$T_{(k)} = \frac{\bar{X}_{I^{(k)}} - \bar{X}_{C^{(k)}}}{S_{p^{(k)}} \sqrt{2/n^{(k)}}},$$

where

$$\begin{aligned} \bar{X}_{I^{(k)}} &= \frac{1}{n^{(k)}} \sum_{j=1}^k n_j \bar{X}_{I,j}, \quad \bar{X}_{C^{(k)}} = \frac{1}{n^{(k)}} \sum_{j=1}^k n_j \bar{X}_{C,j}, \text{ and} \\ S_{p^{(k)}}^2 &= \frac{1}{2(n^{(k)} - k)} \sum_{j=1}^k 2(n_j - 1) S_{p,j}^2 \text{ with } n^{(k)} = \sum_{j=1}^k n_j. \end{aligned}$$

At time k , a decision is made to either fail to reject $H_0 : \mu_I = \mu_C$ or reject H_0 in favor of $H_a : \mu_I \neq \mu_C$ if

$$0 \leq |T_{(k)}| < \xi_{0(k)} \text{ or } |T_{(k)}| \geq \xi_{1(k)},$$

respectively, where $0 \leq \xi_{0(k)} \leq \xi_{1(k)}$. Otherwise, data is collected at sampling stage $k + 1$ and the decision rule is applied at this stage. The event the test will fail to

reject the null hypothesis is given by

$$\begin{aligned} \{\text{fail to reject}\} &= \{0 \leq |T_{(1)}| < \xi_{0(1)}\} \\ &\cup \left(\bigcup_{k=2}^{\infty} \{\xi_{0(k-1)} \leq |T_{(k-1)}| < \xi_{1(k-1)}, 0 \leq |T_{(k)}| < \xi_{0(k)}\} \right). \end{aligned}$$

The power of the test can now be expressed as

$$\begin{aligned} \pi &= 1 - P(0 \leq |T_{(1)}| < \xi_{0(1)}) \\ &\quad - \sum_{k=2}^{\infty} P(\xi_{0(k-1)} \leq |T_{(k-1)}| < \xi_{1(k-1)}, 0 \leq |T_{(k)}| < \xi_{0(k)}) \end{aligned}$$

if the alternative hypothesis holds. The size α of the test is functionally equivalent to π when the null hypothesis is true.

In the unequal variances case, the sequence of test statistics for the procedures for Pocock (1977), O'Brien and Fleming (1979), and Wang and Tsatis (1987) would have the k th test statistic defined by

$$T_{(k)} = \frac{\bar{X}_{I,k} - \bar{X}_{C,k}}{\sqrt{(\sigma_I^2 + \sigma_C^2)/n_k}} \text{ or } T_{(k)} = \frac{\bar{X}_{I,k} - \bar{X}_{C,k}}{\sqrt{(S_{I,k}^2 + S_{C,k}^2)/n_k}}.$$

If the decision is to collect more data, then the decision rule is applied at sampling stage $k + 1$. Otherwise, the test is applied at time $k + 1$, for $j = 1, \dots, k - 1$. At time k , if the test does not reject the null hypothesis then a decision is made to fail to reject H_0 . Here we are assuming a common and known variance ($\sigma_I^2 = \sigma_C^2 = \sigma^2$). In the case in which σ is not known, we replace the value of σ in the expressions for our test statistic with

$$S_{p,j} = \sqrt{\frac{S_{I,j}^2 + S_{C,j}^2}{2}},$$

where $S_{I,j}^2$ and $S_{C,j}^2$ are the sample variances of the intervention and control data, respectively. For the case in which the variances are unequal and known (unknown),

Table 5.2: Group Sequential Tests

{t5.2}

Authors	Critical Values	$T_{(k)}$
	$\xi_{0(1)} = \dots = \xi_{0(K-1)}$ $= 0;$	
Pocock (1977)	$\xi_{0(K)} = \xi_{1(1)} = \dots$ $= \xi_{1(K)} = \xi;$ $n_1 = n_2 = \dots$ $= n_K = n.$	$T_{(k)} = \frac{\bar{X}_{I,k} - \bar{X}_{C,k}}{\sigma\sqrt{2/n}}$ or $T_{(k)} = \frac{\bar{X}_{I,k} - \bar{X}_{C,k}}{S_{p,k}\sqrt{2/n}}$
	$\xi_{0(1)} = \dots = \xi_{0(K-1)}$ $= 0;$	
O'Brien and Fleming (1979)	$\xi_{1(1)} = \dots = \xi_{1(K-1)}$ $= \xi\sqrt{K/k};$ $\xi_{0(K)} = \xi_{1(K)} = \xi;$ $n_1 = n_2 = \dots$ $= n_K = n.$	$T_{(k)} = \frac{\bar{X}_{I,k} - \bar{X}_{C,k}}{\sigma\sqrt{2/n}}$ or $T_{(k)} = \frac{\bar{X}_{I,k} - \bar{X}_{C,k}}{S_{p,k}\sqrt{2/n}}$
Wang and Tsiatis (1987)	$\xi_{0(1)} = \dots = \xi_{0(K-1)} = 0;$ $n_1 = n_2 = \dots = n_K = n.$	$T_{(k)} = \frac{\bar{X}_{I,k} - \bar{X}_{C,k}}{\sigma\sqrt{2/n}}$ or $T_{(k)} = \frac{\bar{X}_{I,k} - \bar{X}_{C,k}}{S_{p,k}\sqrt{2/n}}$
Champ and Hu (2009)	$0 \leq \xi_{0(k)} \leq \xi_{1(k)}.$	$T_{(k)} = \frac{\bar{X}_{I(k)} - \bar{X}_{C(k)}}{\sigma\sqrt{2/n_{(k)}}}$ or $T_{(k)} = \frac{\bar{X}_{I(k)} - \bar{X}_{C(k)}}{S_{p(k)}\sqrt{2/n_{(k)}}}$

$$\bar{X}_{I(k)} = \frac{1}{n_{(k)}} \sum_{j=1}^k n_j \bar{X}_{I,j}, \quad \bar{X}_{C(k)} = \frac{1}{n_{(k)}} \sum_{j=1}^k n_j \bar{X}_{C,j}, \text{ and}$$

$$S_{p(k)}^2 = \frac{1}{2(n_{(k)} - k)} \sum_{j=1}^k 2(n_j - 1) S_{p,j}^2 \text{ with } n_{(k)} = \sum_{j=1}^k n_j.$$

we would replace the decision inequality with

$$\left| \frac{\bar{X}_{I,j} - \bar{X}_{C,j}}{\sqrt{(\sigma_I^2 + \sigma_C^2)/n_j}} \right| \geq \xi_j$$

and if the variances are unknown with

$$\left| \frac{\bar{X}_{I,j} - \bar{X}_{C,j}}{\sqrt{(S_{I,j}^2 + S_{C,j}^2)/n_j}} \right| \geq \xi_j.$$

The values ξ_1, \dots, ξ_k are known as critical values of the test.

For the case in which $\sigma_I^2 = \sigma_C^2 = \sigma^2$ and σ is known, the power of the K -stage test of Pocock (1977) can be expressed as

$$\pi(\xi, \delta) = 1 - \left[\Phi\left(\xi - \sqrt{n/2\delta}\right) - \Phi\left(-\xi - \sqrt{n/2\delta}\right) \right]^K,$$

for $\delta \neq 0$. The size of the test α is

$$\alpha = \pi(\xi, 0) = 1 - [2\Phi(\xi) - 1]^K.$$

Solving this equation for ξ in terms of α , we have

$$\xi = z_{(1+(1-\alpha)^{1/K})/2}.$$

If $\lambda = \sigma_I/\sigma_C$, then either σ_I or σ_C or both are not equal to σ . It then follows that

$$\begin{aligned} T_{(j)} &= \frac{\bar{X}_{I,j} - \bar{X}_{C,j}}{\sigma\sqrt{2/n}} \\ &= \frac{\sqrt{(\sigma_I^2 + \sigma_C^2)/n}}{\sigma\sqrt{2/n}} \left(\frac{(\bar{X}_{I,j} - \bar{X}_{C,j}) - (\mu_I - \mu_C)}{\sqrt{(\sigma_I^2 + \sigma_C^2)/n}} + \frac{\mu_I - \mu_C}{\sqrt{(\sigma_I^2 + \sigma_C^2)/n}} \right) \\ &= \frac{\sqrt{\lambda^2 + 1}}{(\sigma/\sigma_C)\sqrt{2}} \left(Z_j + \frac{\delta_C}{\sqrt{(\lambda^2 + 1)/n}} \right) \\ &= \frac{\sigma_C}{\sigma} \sqrt{(\lambda^2 + 1)/2} \left(Z_j + \sqrt{n/(\lambda^2 + 1)}\delta_C \right). \end{aligned}$$

The power of the test is then given by

$$\begin{aligned}\pi &= 1 - \prod_{j=1}^k P \left(\frac{-\xi(\sigma/\sigma_C) - \sqrt{n/2}\delta_C}{\sqrt{(\lambda^2 + 1)/2}} < Z_j < \frac{\xi(\sigma/\sigma_C) - \sqrt{n/2}\delta_C}{\sqrt{(\lambda^2 + 1)/2}} \right) \\ &= 1 - \left[\Phi \left(\frac{\xi(\sigma/\sigma_C) - \sqrt{n/2}\delta_C}{\sqrt{(\lambda^2 + 1)/2}} \right) - \Phi \left(\frac{-\xi(\sigma/\sigma_C) - \sqrt{n/2}\delta_C}{\sqrt{(\lambda^2 + 1)/2}} \right) \right]^K.\end{aligned}$$

It then follows that the size of the test is given by

$$\alpha = 1 - \left[2\Phi \left(\frac{\xi}{\sqrt{(\lambda^2 + 1)/2}} \right) - 1 \right]^K.$$

Solving this equation for ξ in terms of α and λ , we have

$$\xi = \frac{\sigma_C}{\sigma} \sqrt{(\lambda^2 + 1)/2} z_{(1+(1-\alpha)^{1/K})/2}.$$

When $\lambda = 1$, then

$$\begin{aligned}\pi &= 1 - \prod_{j=1}^k P \left(-\xi(\sigma/\sigma_C) - \sqrt{n/2}\delta_C < Z_j < \xi(\sigma/\sigma_C) - \sqrt{n/2}\delta_C \right) \\ &= 1 - \left[\Phi \left(\xi(\sigma/\sigma_C) - \sqrt{n/2}\delta_C \right) - \Phi \left(-\xi(\sigma/\sigma_C) - \sqrt{n/2}\delta_C \right) \right]^K.\end{aligned}$$

The group sequential methods of Pocock (1977) and O'Brien and Fleming (1979) do not allow for a decision to fail to reject the null hypothesis until all k samples are measured. In what follows, we propose a family of tests that allow the researcher to make the decision to fail to reject the null hypothesis on or before all k samples are measured. Also, this testing procedure allows at time j for all the data collected to this point in time to be used in the decision making process. A member of this family of test at sampling stage $j = 1, \dots, k - 1$

- (1) fails to reject H_0 if $|T_j| < \xi_{0,j}$; or (2) reject H_0 if $|T_j| \geq \xi_{1,j}$;

or (3) decides to include more data in the decision making process if $\xi_{0,j} \leq |T_j| < \xi_{1,j}$.
At sampling stage k , the test

(1) fails to reject H_0 if $|T_k| < \xi_k$; or (2) reject H_0 if $|T_k| \geq \xi_k$.

The test statistics are define by

$$T_j = \frac{(n_1 \bar{X}_{I,1} A_1 + \dots + n_j \bar{X}_{I,j} A_j) / m_j - (n_1 \bar{X}_{C,1} A_1 + \dots + n_j \bar{X}_{C,j} A_j) / m_j}{\sigma \sqrt{2/m_j}}$$

where

$$m_j = n_1 A_1 + \dots + n_j A_j$$

with $A_i = 1$ if the i th sample ($i = 1, \dots, j$) is to be included in the decision making process at time j .

We first examine the case in which we wish to test $H_0 : \mu_I = \mu_C$ in favor of the alternative (researcher's) hypothesis $H_a : \mu_I \neq \mu_C$ assuming equal variances ($\sigma_I^2 = \sigma_C^2 = \sigma^2$). Two samples are taken, one of size n_1 and the second of size n_2 . We represent the measurements on these samples by $X_{I,1,1}, \dots, X_{I,1,n_1}, X_{C,1,1}, \dots, X_{C,1,n_1}$ and $X_{I,2,1}, \dots, X_{I,2,n_2}, X_{C,2,1}, \dots, X_{C,2,n_2}$. The means and variances of these samples are represented by $\bar{X}_{I,1}, S_{I,1}^2$ and $\bar{X}_{C,1}, S_{C,1}^2$, respectively. First let us assume that σ^2 is known. A two stage test rejects H_0 if

$$\left| \frac{\bar{X}_{I,1} - \bar{X}_{C,1}}{\sigma \sqrt{2/n_1}} \right| \geq z_{\alpha_1/2}$$

or if

$$z_{\alpha_0/2} \leq \left| \frac{\bar{X}_{I,1} - \bar{X}_{C,1}}{\sigma \sqrt{2/n_1}} \right| < z_{\alpha_1/2} \text{ and } \left| \frac{\bar{X}_I - \bar{X}_C}{\sigma \sqrt{2/(n_1 + n_2)}} \right| \geq z_{\alpha_2/2}.$$

Here, \bar{X}_I and \bar{X}_C are the means of the combined intervention and control data,

respectively. The power of the test is

$$\begin{aligned}
\pi &= 1 - \Phi\left(z_{\alpha_1/2} - \sqrt{\frac{n_1}{2}}\delta\right) + \Phi\left(-z_{\alpha_1/2} - \sqrt{\frac{n_1}{2}}\delta\right) \\
&\quad - \int_{z_{\alpha_0/2} - \sqrt{n_1/2}\delta}^{z_{\alpha_1/2} - \sqrt{n_1/2}\delta} \Phi\left(z_{\alpha_2/2}\sqrt{\frac{n_1+n_2}{n_2}} - \sqrt{\frac{n_1}{n_2}}z_1 - \frac{n_1+n_2}{\sqrt{2n_2}}\delta\right) \phi(z_1) dz_1 \\
&\quad + \int_{z_{\alpha_0/2} - \sqrt{n_1/2}\delta}^{z_{\alpha_1/2} - \sqrt{n_1/2}\delta} \Phi\left(-z_{\alpha_2/2}\sqrt{\frac{n_1+n_2}{n_2}} - \sqrt{\frac{n_1}{n_2}}z_1 - \frac{n_1+n_2}{\sqrt{2n_2}}\delta\right) \phi(z_1) dz_1 \\
&\quad - \int_{-z_{\alpha_1/2} - \sqrt{n_1/2}\delta}^{-z_{\alpha_0/2} - \sqrt{n_1/2}\delta} \Phi\left(z_{\alpha_2/2}\sqrt{\frac{n_1+n_2}{n_2}} - \sqrt{\frac{n_1}{n_2}}z_1 - \frac{n_1+n_2}{\sqrt{2n_2}}\delta\right) \phi(z_1) dz_1 \\
&\quad + \int_{-z_{\alpha_1/2} - \sqrt{n_1/2}\delta}^{-z_{\alpha_0/2} - \sqrt{n_1/2}\delta} \Phi\left(-z_{\alpha_2/2}\sqrt{\frac{n_1+n_2}{n_2}} - \sqrt{\frac{n_1}{n_2}}z_1 - \frac{n_1+n_2}{\sqrt{2n_2}}\delta\right) \phi(z_1) dz_1,
\end{aligned}$$

where $\phi(z)$, $\Phi(z)$, and $z_{1-\gamma}$ are the probability density function, the cumulative distribution function, and the 100 γ th percentile of a standard normal distribution and $\delta = (\mu_I - \mu_C) / \sigma$. We see that the power function is a function of α_0 , α_1 , α_2 , n_1 , n_2 , and δ with the restriction that $\alpha_0 > \alpha_1$. It follows that the size α of the test can be expressed as

$$\begin{aligned}
\alpha &= 2(1 - \Phi(z_{\alpha_1/2})) \\
&\quad - 2 \int_{z_{\alpha_0/2}}^{z_{\alpha_1/2}} \Phi\left(z_{\alpha_2/2}\sqrt{\frac{n_1+n_2}{n_2}} - \sqrt{\frac{n_1}{n_2}}z_1\right) \phi(z_1) dz_1 \\
&\quad + 2 \int_{z_{\alpha_0/2}}^{z_{\alpha_1/2}} \Phi\left(-z_{\alpha_2/2}\sqrt{\frac{n_1+n_2}{n_2}} - \sqrt{\frac{n_1}{n_2}}z_1\right) \phi(z_1) dz_1.
\end{aligned}$$

The size α is functionally equivalent to π when $\delta = 0$.

Suppose we select $n_1 = n_2 = n$. Further, suppose we select $z_{\alpha_0/2} = 0$. The power

function can then be expressed as

$$\begin{aligned}
\pi &= 1 - \Phi\left(z_{\alpha_1/2} - \sqrt{\frac{n}{2}}\delta\right) + \Phi\left(-z_{\alpha_1/2} - \sqrt{\frac{n}{2}}\delta\right) \\
&\quad - \int_{-\sqrt{n/2}\delta}^{z_{\alpha_1/2} - \sqrt{n/2}\delta} \Phi\left(z_{\alpha_2/2}\sqrt{2} - z_1 - \sqrt{2n}\delta\right) \phi(z_1) dz_1 \\
&\quad + \int_{-\sqrt{n/2}\delta}^{z_{\alpha_1/2} - \sqrt{n/2}\delta} \Phi\left(-z_{\alpha_2/2}\sqrt{2} - z_1 - \sqrt{2n}\delta\right) \phi(z_1) dz_1 \\
&\quad - \int_{-\sqrt{n/2}\delta}^{-z_{\alpha_0/2} - \sqrt{n/2}\delta} \Phi\left(z_{\alpha_2/2}\sqrt{2} - z_1 - \sqrt{2n}\delta\right) \phi(z_1) dz_1 \\
&\quad + \int_{-\sqrt{n/2}\delta}^{-z_{\alpha_0/2} - \sqrt{n/2}\delta} \Phi\left(-z_{\alpha_2/2}\sqrt{2} - z_1 - \sqrt{2n}\delta\right) \phi(z_1) dz_1,
\end{aligned}$$

and the size as

$$\alpha_0 = 2\left(1 - \Phi\left(z_{\alpha/2}\right)\right).$$

It then follows that

$$z_{\alpha/2} = \Phi^{-1}\left(1 - \alpha_0/2\right) = z_{\alpha_0/2}.$$

Next we consider the case in which σ is unknown. In this case, we consider the two stage test that rejects H_0 if

$$(1) |T_1| \geq t_{n_1-1,0,\alpha_1/2} \text{ or } (2) \text{ if } t_{n_1-1,0,\alpha_0/2} \leq |T_1| < t_{n_1-1,0,\alpha_1/2} \text{ and } |T| \geq t_{n_1+n_2-2,0,\alpha_2/2}.$$

where

$$\begin{aligned}
T_1 &= \frac{\bar{X}_{I,1} - \bar{X}_{C,1}}{S_{p,1}\sqrt{2/n_1}}, \quad T = \frac{\bar{X}_I - \bar{X}_C}{S_p\sqrt{2/(n_1+n_2)}}, \quad S_{p,1}^2 = \frac{S_{I,1}^2 + S_{C,1}^2}{2}, \\
S_{p,2}^2 &= \frac{S_{I,2}^2 + S_{C,2}^2}{2}, \text{ and } S_p^2 = \frac{(n_1-1)S_{p,1}^2 + (n_2-1)S_{p,2}^2}{n_1+n_2-2}.
\end{aligned}$$

The power of the test can be expressed as

$$\begin{aligned}
\pi &= 1 - \int_0^\infty \Phi \left(t_{2(n_1-1),0,\alpha_1/2} \sqrt{\frac{y_{p,1}}{2(n_1-1)}} - \theta_1 \right) f_{Y_{p,1}}(y_{p,1}) dy_{p,1} \\
&+ \int_0^\infty \Phi \left(-t_{2(n_1-1),0,\alpha_1/2} \sqrt{\frac{y_{p,1}}{2(n_1-1)}} - \theta_1 \right) f_{Y_{p,1}}(y_{p,1}) dy_{p,1} \\
&- \int_0^\infty \int_{-t_{2(n_1-1),0,\alpha_0/2} \sqrt{y_{p,1}/(2(n_1-1))} - \theta_1}^{-t_{2(n_1-1),0,\alpha_0/2} \sqrt{y_{p,1}/(2(n_1-1))} - \theta_1} \\
&\times \int_0^\infty G(t_{n_1+n_2-2,0,\alpha_2/2}, n_1, n_2, z_1, y_{p,1}, y_{p,2}, \theta) \\
&\times f_{Y_{p,1}}(y_{p,2}) \phi(z_1) f_{Y_{p,1}}(y_{p,1}) dy_{p,2} dz_1 dy_{p,1} \\
&- \int_0^\infty \int_{t_{2(n_1-1),0,\alpha_0/2} \sqrt{y_{p,1}/(2(n_1-1))} - \theta_1}^{t_{2(n_1-1),0,\alpha_1/2} \sqrt{y_{p,1}/(2(n_1-1))} - \theta_1} \\
&\times \int_0^\infty G(t_{n_1+n_2-2,0,\alpha_2/2}, n_1, n_2, z_1, y_{p,1}, y_{p,2}, \theta) \\
&\times f_{Y_{p,1}}(y_{p,2}) \phi(z_1) f_{Y_{p,1}}(y_{p,1}) dy_{p,2} dz_1 dy_{p,1},
\end{aligned}$$

where $\theta_1 = \sqrt{n_1/2}\delta$, $\theta = \sqrt{(n_1+n_2)/2}\delta$, $y_{p,1} = 2(n_1-1)S_{p,1}^2/\sigma^2$, and

$$\begin{aligned}
G(t, n_1, n_2, z_1, y_1, y_2, \theta) &= \Phi \left(t \sqrt{\frac{(n_1+n_2)(y_1+y_2)}{n_2(n_1+n_2-2)}} - \sqrt{\frac{n_1}{n_2}} z_1 - \sqrt{\frac{n_1+n_2}{n_2}} \theta \right) \\
&- \Phi \left(-t \sqrt{\frac{(n_1+n_2)(y_1+y_2)}{n_2(n_1+n_2-2)}} - \sqrt{\frac{n_1}{n_2}} z_1 - \sqrt{\frac{n_1+n_2}{n_2}} \theta \right).
\end{aligned}$$

Note that π is a function of α_0 , α_1 , α_2 , n_1 , n_2 , and δ . Since the size of the test α is functionally equivalent to the power when $\delta = 0$, we can express the size of the test

as

$$\begin{aligned}
\alpha &= 1 - \int_0^\infty \left(2\Phi \left(t_{2(n_1-1),0,\alpha_1/2} \sqrt{\frac{y_{p,1}}{2(n_1-1)}} \right) - 1 \right) f_{Y_{p,1}}(y_{p,1}) dy_{p,1} \\
&\quad - \int_0^\infty \int_{-t_{2(n_1-1),0,\alpha_1/2} \sqrt{y_{p,1}/(2(n_1-1))}}^{-t_{2(n_1-1),0,\alpha_0/2} \sqrt{y_{p,1}/(2(n_1-1))}} \\
&\quad \times \int_0^\infty G(t_{n_1+n_2-2,0,\alpha_2/2}, n_1, n_2, z_1, y_{p,1}, y_{p,2}, 0) \\
&\quad \times f_{Y_{p,1}}(y_{p,2}) \phi(z_1) f_{Y_{p,1}}(y_{p,1}) dy_{p,2} dz_1 dy_{p,1} \\
&\quad - \int_0^\infty \int_{t_{2(n_1-1),0,\alpha_0/2} \sqrt{y_{p,1}/(2(n_1-1))}}^{t_{2(n_1-1),0,\alpha_1/2} \sqrt{y_{p,1}/(2(n_1-1))}} \\
&\quad \times \int_0^\infty G(t_{n_1+n_2-2,0,\alpha_2/2}, n_1, n_2, z_1, y_{p,1}, y_{p,2}, 0) \\
&\quad \times f_{Y_{p,1}}(y_{p,2}) \phi(z_1) f_{Y_{p,1}}(y_{p,1}) dy_{p,2} dz_1 dy_{p,1}.
\end{aligned}$$

5.6 Conclusion

In this chapter, various sequential and adaptive methods for comparing two population means were presented. One is a proposed new method for solving the Behrens-Fisher problem.

CHAPTER 6

CONCLUSION

6.1 General Conclusions

Various of statistical methods based on the assumptions about the data has been discussed. Using these methods to comparing two population means and also to compute the appropriate sample sizes. An exact solution to the Behrens-Fisher distribution was given in Chapter 2, showing that the pdf and cdf functions can be expressed as linear combinations of non-central t -distributions. A FORTRAN program was written to present the numerical method for obtaining a good approximation to the cdf and pdf of the Behrens-Fisher distribution. Also it shown that the method proposed by Welsh(1938) provides a good approximation.

In Chapter 3, methods for comparing two population means were examined either under the assumption the variances are known or that their ratio is given. Methods are discussed for selecting the sample sizes based on certain requirements imposed by the researcher. These methods are based on the power of the test.

Besides the independent sample case, sometimes the intervention can be given to each individual in the population with the individual also serving as the control. This is the paired data case. For this case, methods for comparing two population means were presented and methods for deriving sample sizes were given.

Various sequential and adaptive methods were presented in Chapter 5 for the “paired data” and “independent sample” cases. These methods included the two-stage sample size prediction methods, sequential methods, and group sequential methods. A new method was presented for solving the Behrens-Fisher problem.

6.2 Areas for Further Research

While in clinical trials, it is usually the case the hypotheses to be tested are that the means are equal versus them not being equal. In other applications, the appropriate hypotheses that are one-sided. We are interested in developing our method for this case.

Missing data is a common problem in designed experiments. We wish to examine how our method can be adapted to account for missing data. Very little research has been done in the area of missing data in sequential methods.

We wish to compare our method to the methods of Chapman (1950), Prokof'yev and Shishkin (1974), and Dudewicz and Ahmed (1998,1999). It has been stated by Dudewicz, E.J., Ma, Y., Mai, S.E., and Su, H. (2007) that the latter three methods are solutions to the Behrens-Fisher problem.

Often the response variable is a multivariate measurement. It would be of interest to study how the univariate methods we have examined could be extended to the multivariate case.

There is much work to be done in the area of group sequential methods. It would be interesting to see if our method could be extended to this area.

In actual practice, clinical trials sometimes are expensive and dangerous, so making a decision about the size of the sample just based on power may not be feasible. It may be then necessary do the experiment step by step with a smaller overall expected sample size.

One can provide a bioequivalence analysis based on control and intervention

groups, then obtain a confidence interval for the difference in the two means. Based on the observed confidence interval a decision can be made about the effect of the intervention. This bioequivalence analysis idea will give us a much better decision than the method we use now.

REFERENCES

- [1] A. Wald, *Sequential Analysis*, Dover, 1947.
- [2] B.L. Welch, *The Significance of the Difference Between Two Means When the Population Variances Are Unequal*, *Biometrika*.29(1938), 350-362.
- [3] B.M. Seelbinder, *On Stein's Two-stage Sampling Scheme*, *Annals of Mathematical Statistics*. 24(1953), 640-649.
- [4] B.K. Taneja, E.J. Dudewicz, *Exact Solutions to the Behrens-Fisher Problem*, Chapter 27 in *Multiple Comparisons, Selection, and Applications in Biometry*, A Festschrift in Honor of Charles W. Dunnett, edited by F.M. Hoppe(1993), Marcel-Dekker, Inc.: New York.
- [5] C. Stein, *A Two-Sample Test for a Linear Hypothesis Whose Power is Independent of the Variance*, *Annals of Mathematical Statistics*. 16(1945), 243-258.
- [6] C. Jennison, B.W. Turnbull, *Group Sequential Methods with Applications to Clinical Trials*, Chapman & Hall/CRC, 2000.
- [7] D.G. Chapman, *Some Two Sample Tests*, *Annals of Mathematical Statistics*. 21(1950), 601-606.
- [8] D.G. Nel, C.A. van der Merwe, B.K. Moser, *The Exact Distributions of the Univariate and Multivariate Behrens-Fisher Statistics with a Comparison of Several Solutions in the Univariate Case*, *Communications in Statistics - Theory and Methods*. 19(1990), 279-298.
- [9] D. Benton, K. Drishnamoorthy, *Computing Discrete Mixtures of Continuous Distributions: Noncentral Chi Square, Noncentral t, and the Distribution of the Square of the Sample Multiple Correlation Coefficient*, *Computational Statistics and Data Analysis*, 43(2003), 249-267.
- [10] E.J. Dudewicz, S.U. Ahmed, *New Exact and Asymptotically Optimal Solution to the Behrens-Fisher Problem, with Tables*, *American Journal of Mathematical and Management Sciences*, 18(1998), 359-426.

- [11] E.J. Dudewicz, S.U. Ahmed, *New Exact and Asymptotically Optimal Heteroscedastic Statistical Procedures and Tables, II*, American Journal of Mathematical and Management Sciences, 19(1999), 157–180.
- [12] E.J. Dudewicz, Y. Ma, S.E. Mai, H. Su, *Exact Solutions to the Behrens–Fisher Problem: Asymptotically Optimal and Finite Sample Efficient Choice Among*, Journal of Statistical Planning and Inference, 137(2007), 1584–1605.
- [13] G.B. Dantzig, *On the Non-existence of Tests of Student’s Hypothesis Having Power Functions Independent of σ* , Annals of Mathematical Statistics 11(1940), 186-192.
- [14] G.A. Barnard, *Sequential Tests in Industrial Statistics*, Journal of the Royal Statistical Society, Series B 8(1946), 1-21.
- [15] G.L. Elfring, J.R. Schultz, *Group Sequential Designs for Clinical Trials*, Biometrics, 29(1973), 471-477.
- [16] H.F. Dodge, H.G. Roming, *A Method for Sampling Inspection*, Bell Syst. Tech. J., 8(1929), 613-631.
- [17] J. Moshman, *A Method for Selecting the Size of the Initial Sample in Stein’s Two Sample Procedure*, The Annals of Mathematical Statistics, 29(1958), 1271-1275.
- [18] K. McPherson, *Statistics: The Problem of Examining Accumulating Data More Than Once*, New Engl. J. Med., 290(1974), 501-502.
- [19] L. Bain, M. Engelhardt, *Introduction to Probability and Mathematical Statistics*, Second Edition, Duxbury, 1992.
- [20] M. Evans, N. Hastings, B. Peacock, *Statistical Distributions*, Second Edition, John Wiley & Sons, 1993.
- [21] M.A. Proschan, *Two-stage Sample Size Re-estimation Based on a Nuisance Parameter: A Review*, Journal of Biopharmaceutical Statistics, 15 (2005), 539-574.
- [22] P.L. Canner, *Monitoring Treatment Differences in Long-Term Clinical Trials*, Biometrics, 33(1977), 603-615.

- [23] P.C. O'Brien, T.R. Fleming, *A Multiple Testing Procedure for Clinical Trials*, *Biometrics* 35(1979), 549-556.
- [24] Q. Liu, M.A. Proschan, G.W. Pldeger,, *A Unified Theory of Two-stage Adaptive Designs*, *Journal of American Statistical Association*, 97(2002), 1034-1041.
- [25] S.J. Pocock, *Group Sequential Methods in the Design and Analysis of Clinical Trials*, *Biometrika* 64(1977), 191-199.
- [26] S.K. Wang, A.A.Tsiatis, *Approximately Optimal One-parameter Boundaries for Group Sequential Trials*, *Biometrics* 43(1987), 193-200.
- [27] S.-H. Kim, A.S. Cohen, *On the Behrens-Fisher Problem: A Review*, *Journal of Educational and Behavioral Statistics* 23(1998), 356-377.
- [28] S.C. Chow, M. Chang, *Adaptive Design Methods in Clinical Trials*, Chapman & Hall/CRC, 2007.
- [29] V.N. Prokof'yev, A.D. Shishkin, *Successive Classification of Normal Sets with Unknown Variances*, *Radio Engineering Electronics Physics* 19(1974), 141-143.
- [30] W.A. Shewhart, *The Application of Statistics as an Aid in Maintaining Quality of a Manufactured Product*, *Journal of the American Statistical Association*, 20(1925), 546-548.
- [31] W.A. Shewhart, *Economic control of quality of manufactured product*, Van Nostrand, (1931).

Appendix A
APPENDIX I

```
C*
C*-----*
C* THE CUMULATIVE DISTRIBUTION FUNCTION FOR THE *
C* STATISTICS T_1 AND T_2 EVALUATED AT THE VALUE T *
C* WITH PARAMETERS *
C* *
C* DELTAC = (MU_I-MU_C)/SIGMAC *
C* LAMBDA = SIGMAI/SIGMAC *
C* NI = INTERVENTION GROUP SAMPLE SIZE *
C* NC = CONTROL GROUP SAMPLE SIZE *
C* *
C* WHERE *
C* *
C*          XBARI - XBARC *
C* T_1 = ----- *
C*          S_P*SQRT(1/NI+1/NC) *
C* *
C* AND *
C* *
C*          XBARI - XBARC *
C* T_2 = ----- *
C*          SQRT((S_I)^2/NI+(S_C)^2/NC) *
C* *
C* WITH *
C* *
```

```

C*          (NI-1)*(S_I)^2+(NC-1)*(S_C)^2          *
C* S_P = ----- *
C*          (NI-1)+(NC-1) *
C* *
C* XBARI AND S_I THE MEAN AND STANDARD DEVIATION OF *
C* THE INTERVENTION GROUP DATA AND XBARC AND S_C THE *
C* MEAN AND STANDARD DEVIATION OF CONTROL GROUP *
C* DATA. *
C* N(MUI,SIGMAI^2) AND N(MUC,SIGMAC^2) *
C* DISTRIBUTIONS, RESPECTIVELY *
C* *
C* AUTHORS: CHARELS W. CHAMP AND FENGJIAO HU *
C* *
C*-----*
C*
      DOUBLE PRECISION FUNCTION DFTCDF(T,DELTAC,
      & LAMBDA,NI,NC,TI)
C*
      INTEGER DF,K,NI,NII,NC,NCC,TI
      DOUBLE PRECISION B,CK,DELTAC,DTNDF,EP,LAMBDA,
      & NU,ONE,T,THETA,TWO,U,XI,ZERO
C*
C* CONSTANT VALUES
C*
      ZERO=0.0DO
      ONE=1.0DO
      TWO=2.0DO

```

```

      EP=0.0000000000000001D0
C*
C* PARAMETER CONSTANTS
C*
      THETA=DELTAC
      & /DSQRT(LAMBDA*LAMBDA/NI+ONE/NC)
      NII=NI
      NCC=NC
C*
C* T_1: THE VALUES OF XI AND NU ARE CALCULATED
C* (SEE THE RESULTS IN CHAPTER 2)
C*
      IF (TI.EQ.1) THEN
          XI=LAMBDA*LAMBDA
          NU=NII+NCC-TWO
          NU=NU*(LAMBDA*LAMBDA/NII+ONE/NCC)
          NU=NU/(ONE/NII+ONE/NCC)
      ENDIF
C*
C* T_2: THE VALUES OF XI AND NU ARE CALCULATED
C* (SEE THE RESULTS IN CHAPTER 2)
C*
      IF (TI.EQ.2) THEN
          XI=LAMBDA*LAMBDA*NCC*(NCC-ONE)
          XI=XI/(NII*(NII-ONE))
          NU=NCC*(NCC-ONE)
          NU=NU*(LAMBDA*LAMBDA/NII+ONE/NCC)

```

```
        ENDIF

C*
C* IF XI > 1, THEN THE CDF IS EVALUATED AS FUNCTIONS
C* OF THE PARAMETERS 1/XI FOR XI, NU/XI FOR NU, AND
C* THE SAMPLE SIZES ARE INTERCHANGED.
C*
        IF (XI.GT.ONE) THEN
            XI=ONE/XI
            NU=XI*NU
            NII=NC
            NCC=NI
        ENDIF

C*
C* SOME SIMPLIFYING CONSTANTS
C*
        B=(NCC-ONE)/TWO
        DF=NII+NCC-2
        U=T*DSQRT((XI/NU)*DF)

C*
C* THE CDF IS INITIALIZED IN TERMS OF THE CDF OF A
C* NON-CENTRAL T-DISTRIBUTION WITH PARAMETERS DF AND
C* THETA EVALUATED AT THE VALUE U
C*
C* THE IMSL ROUTINE DTNDF IS USED TO DETERMINE THE
C* CDF OF A NON-CENTRAL T-DISTRIBUTION
C*
        DFTCDF=(XI**B)*DTNDF(U,DF,THETA)
```

```
C*
C* IF XI NOT EQUAL TO ONE, THEN A FINITE NUMBER OF
C* OF TERMS OF THE SERIES REPRESENTATION OF THE
C* DISTRIBUTION OF T_I ARE USED TO APPROXIMATE THE
C* CDF. THE NUMBER OF TERMS USED DEPENDS ON THE
C* VALUES OF XI AND NU.
```

```
C*
      K=0
      CK=ONE
      IF (XI.NE.ONE) THEN
          CDF=ZERO
1      K=K+1
          DF=NII+NCC-2+2*K
          U=T*DSQRT((XI/NU)*DF)
          CK=(B+K-ONE)*(ONE-XI)*CK/K
          CDF=CDF+CK*DTNDF(U,DF,THETA)
          IF (DABS(CK).GT.EP) GOTO 1
      ENDIF
      DFTCDF=DFTCDF+(XI**B)*CDF
C*
      RETURN
      END
```

Appendix B

APPENDIX II

```

C*-----*
C* THE SIZE OR POWER OF TEST BASED ON THE STATISTICS *
C* T_1 AND T_2 ARE CALCULATED UNDER GIVE ASSUMPTIONS.*
C*
C* INPUT
C* ALPHA = DESIRED SIZE OF THE TEST
C* DELTAC = (MU_I-MU_C)/SIGMAC
C* LAMBDA = SIGMAI/SIGMAC
C* NI = INTERVENTION GROUP SAMPLE SIZE
C* NC = CONTROL GROUP SAMPLE SIZE
C*
C*          XBARI - XBARC
C* T_1 = -----
C*          S_P*SQRT(1/NI+1/NC)
C*
C* AND
C*
C*          XBARI - XBARC
C* T_2 = -----
C*          SQRT((S_I)^2/NI+(S_C)^2/NC)
C*
C* WITH
C*
C*          (NI-1)*(S_I)^2+(NC-1)*(S_C)^2

```



```

C* S_P = ----- *
C*          (NI-1)+(NC-1) *
C* AND *
C* *
C*          ((S_I)^2/NI+(S_C)^2/NC)^2 *
C* V = ----- *
C*          ((S_I)^2/NI)^2/(NI-1)+((S_C)^2/NC)^2/(NC-1) *
C* *
C* FOR WELCH'S APPROXIMATION/ESTIMATION METHOD. *
C* *
C* XBARI AND S_I THE MEAN AND STANDARD DEVIATION OF *
C* THE INTERVENTION GROUP DATA AND XBARC AND S_C THE *
C* MEAN AND STANDARD DEVIATION OF CONTROL GROUP *
C* DATA. *
C* *
C* THE INTERVENTION AND CONTROL GROUP SAMPLES ARE *
C* TO BE INDEPENDENT RANDOM SAMPLES FROM *
C* N(MUI,SIGMAI^2) AND N(MUC,SIGMAC^2) *
C* DISTRIBUTIONS, RESPECTIVELY *
C* *
C* AUTHORS: CHARELS W. CHAMP AND FENGJIAO HU *
C* *
C*-----*
      USE MSIMSL
C*
      INTEGER CK,NI,NC,TI
      DOUBLE PRECISION ALPHA,ALPHA0,DELTAC,DF,DFTCDF,

```

```

& LAMBDA,ONE,POWER,T,TA,TB,TP,TWO,V,ZERO
C*
C* CONSTANT VALUES
C*
ONE=1.0DO
TWO=2.0DO
ZERO=0.0DO
C*
WRITE(*,*) 'INPUT ALPHAO'
READ(*,*) ALPHAO
WRITE(*,*) 'INPUT DELTAC'
READ(*,*) DELTAC
WRITE(*,*) 'INPUT LAMBDA'
READ(*,*) LAMBDA
WRITE(*,*) 'INPUT NI'
READ(*,*) NI
WRITE(*,*) 'INPUT NC'
READ(*,*) NC
WRITE(*,*) 'INPUT (1) T_1 EQUAL VARIANCES'
WRITE(*,*) 'INPUT (2) T_1'
WRITE(*,*) 'INPUT (3) T_2'
WRITE(*,*) 'INPUT (4) T_2 WELCH"S APPROXIMATION'
READ(*,*) TI
C*
IF (TI.EQ.1) T=DTIN(ONE-ALPHAO/TWO,NI+NC-TWO)
IF ((TI.EQ.2).OR.(TI.EQ.3)) THEN
    IF (TI.EQ.2) TI=1

```

```

IF (TI.EQ.3) TI=2
  DF=NI-1.0D0
  IF (NI.GT.NC) DF=NC-1.0D0
    TA=ZERO
    TB=DTIN(1.0D0-ALPHA0/10.0D0,DF)
1    CK=0
    T=(TA+TB)/2.0D0
    ALPHA=1.0D0-DFTCDF(T,ZERO,LAMBDA,NI,NC,TI)
&    +DFTCDF(-T,ZERO,LAMBDA,NI,NC,TI)
    IF (DABS(ALPHA0-ALPHA).GT.0.000001D0) THEN
      IF (ALPHA.LT.ALPHA0) TB=T
      IF (ALPHA.GT.ALPHA0) TA=T
      CK=1
    ENDIF
    IF (CK.EQ.1) GOTO 1
  ENDIF

C*
C*-----*
C* WELSH(1938) SUGGESTED THAT USING T_2 FOR THE      *
C* VARIANCE UNKNOWN CASE, AND THE DEGREES OF FREEDOM *
C* IS ESTIMATED BY                                  *
C*          (SIGMAI^2/NI+SIGMAC^2/NC)^2              *
C* V = ----- *
C*          (SIGMAI^2/NI)^2/(NI-1)+(SIGMAC^2/NC)^2/(NC-1) *
C*
C*-----*
C*

```

```

      IF (TI.EQ.4) THEN
          TI=2
          TP=LAMBDA*LAMBDA/NI
          V=TP+ONE/NC
          V=V*V
          TP=TP*TP/(NI-ONE)+ONE/(NC*NC)/(NC-ONE)
          V=V/TP
          T=DTIN(ONE-ALPHA0/TWO,V)
      ENDIF
C*
      ALPHA=1.0DO-DFTCDF(T,ZERO,LAMBDA,NI,NC,TI)
      & +DFTCDF(-T,ZERO,LAMBDA,NI,NC,TI)
C*
      POWER=1.0DO-DFTCDF(T,DELTAC,LAMBDA,NI,NC,TI)
      & +DFTCDF(-T,DELTAC,LAMBDA,NI,NC,TI)
C*
      WRITE(*,61) ' DELTAC =',DELTAC
      WRITE(*,61) ' LAMBDA =',LAMBDA
      WRITE(*,62) ' NI =',NI
      WRITE(*,62) ' NC =',NC
      WRITE(*,61) ' T =',T
      WRITE(*,61) ' ALPHA =',ALPHA
      IF (DELTAC.NE.ZERO)
      & WRITE(*,61) ' POWER =',POWER
61  FORMAT(A10,F9.5)
62  FORMAT(A10,I3)
C*

```

STOP

END

Appendix C
APPENDIX III

```
C*-----*
C*
C*-----*
C*
      USE MSIMSL
C*
      INTEGER CK, ISIM, NI, NC, NR, NSIM, TI, TII
      DOUBLE PRECISION ALPHA, ALPHAO, DELTAC, DF, DFTCDF,
& LHAT, LAMBDA, ONE, P, R(1), T, TA, TB, THETAC, TP,
& TSIM, TWO, VHAT, W, WC, WI, Z, ZERO
C*
C* CONSTANT VALUES
C*
      NR=1
      ONE=1.0DO
      TWO=2.0DO
      ZERO=0.0DO
C*
      WRITE(*,*) 'INPUT ALPHAO'
      READ(*,*) ALPHAO
      WRITE(*,*) 'INPUT DELTAC'
      READ(*,*) DELTAC
      WRITE(*,*) 'INPUT LAMBDA'
      READ(*,*) LAMBDA
      WRITE(*,*) 'INPUT NI'
```

```

READ(*,*) NI
WRITE(*,*) 'INPUT NC'
READ(*,*) NC
WRITE(*,*) 'INPUT (1) T_1 EQUAL VARIANCES'
WRITE(*,*) 'INPUT (2) T_1'
WRITE(*,*) 'INPUT (3) T_2'
WRITE(*,*) 'INPUT (4) T_2 WELCH"S APPROXIMATION'
READ(*,*) TI
WRITE(*,*) 'INPUT NUMBER OF SIMULATIONS'
READ(*,*) NSIM

C*
C* PARAMETER CONSTANTS
C*

      THETAC=DELTAC/DSQRT(LAMBDA*LAMBDA/NI+ONE/NC)
      TP=LAMBDA*LAMBDA/NI

C*
      P=ZERO

C*
C* COMPUTE THE POWER BY SIMULATION
C*

      DO 2 ISIM=1,NSIM

C*

      CALL DRNNOR(NR,R)

      Z=R(1)

      DF=NI-ONE

      CALL DRNCHI(NR,DF,R)

```

```

      WI=R(1)
      DF=NC-ONE
      CALL DRNCHI(NR,DF,R)
      WC=R(1)
      TSIM=Z+THETAC

C*
C*-----*
C* THE CUMULATIVE DISTRIBUTION FUNCTION FOR THE *
C* STATISTICS T_1 AND T_2 CAN BE EXPRESSED AS *
C* *
C* WHERE *
C* *
C*          Z+THETAC *
C* T_1 = ----- *
C*          (1/NI+1/NC) (LAMBDA^2*WI+WC) *
C*          SQRT(-----) *
C*          (NI+NC-2) (LAMBDA^2/NI+1/NC) *
C* *
C* AND *
C*          Z+THETAC *
C* T_2 = ----- *
C*          LAMBDA^2*WI      WC *
C*          -----+----- *
C*          NI(NI-1)      NC(NC-1) *
C*          SQRT(-----) *
C*          LAMBDA^2/NI+1/NC *
C* WITH *

```



```

C*
C* THETAC=DELTAC/SQRT(LAMBDA*LAMBDA/NI+ONE/NC)
C*
C* AND
C*
C* WI IS CHI-SQUARE DISTRBUTION WITH NI-1 DEGREES OF
C* FREEDOM
C*
C* WC IS CHI-SQUARE DISTRBUTION WITH NC-1 DEGREES OF
C* FREEDOM
C*-----*
C*
      IF ((TI.EQ.1).OR.(TI.EQ.2)) THEN
      W=ONE/NI+ONE/NC
      W=W/(NI+NC-TWO)
      W=W/(TP+ONE/NC)
      W=W*(LAMBDA*LAMBDA*WI+WC)
      TSIM=TSIM/DSQRT(W)
      ENDIF
C*
      IF ((TI.EQ.3).OR.(TI.EQ.4)) THEN
      W=(TP/(TP+ONE/NC))*WI/(NI-ONE)
      W=W+(ONE/NC/(TP+ONE/NC))*WC/(NC-ONE)
      TSIM=TSIM/DSQRT(W)
      ENDIF
C*
C*-----*

```

```

C* ESTIMATE LAMBDA BY L, WHERE *
C* *
C*      (NC-1)(NC-4)WI *
C* L^2 = ----- *
C*      (NI-1)^2*WC *
C*-----*
C*
      IF (TI.EQ.1) T=DTIN(ONE-ALPHA0/TWO,NI+NC-TWO)
      IF ((TI.EQ.2).OR.(TI.EQ.3)) THEN
          IF (TI.EQ.2) TII=1
          IF (TI.EQ.3) TII=2
          LHAT=(NC-ONE)*(NC-4.0DO)/((NI-ONE)*(NI-ONE))
          LHAT=LHAT*WI/WC
          LHAT=LHAT*LAMBDA*LAMBDA
          LHAT=DSQRT(LHAT)
          DF=NI-1.0DO
          IF (NI.GT.NC) DF=NC-1.0DO
          TA=ZERO
          TB=DTIN(1.0DO-ALPHA0/10.0DO,DF)
1      CK=0
          T=(TA+TB)/2.0DO
          ALPHA=1.0DO-DFTCDF(T,ZERO,LHAT,NI,NC,TII)
&      +DFTCDF(-T,ZERO,LHAT,NI,NC,TII)
      IF (DABS(ALPHA0-ALPHA).GT.0.00001DO) THEN
          IF (ALPHA.LT.ALPHA0) TB=T
          IF (ALPHA.GT.ALPHA0) TA=T
          CK=1

```

```

ENDIF
IF (CK.EQ.1) GOTO 1
ENDIF

C*
C*-----*
C* WELSH(1938) SUGGESTED THAT USING T_2 FOR THE *
C* VARIANCE UNKNOWN CASE, AND THE DEGREES OF FREEDOM *
C* IS ESTIMATED BY *
C*          (NC-1)*WI      1 *
C*          (LAMBDA^2-----+-----)^2 *
C*          NI(NI-1)*WC    NC *
C* V = ----- *
C*          (NC-1)*WI      1 *
C*          (LAMBDA^2-----)^2/(NI-1)+----- *
C*          NI(NI-1)*WC          NC^2(NC-1) *
C* *
C*-----*
C*
IF (TI.EQ.4) THEN
TII=2
TP=LAMBDA*LAMBDA*WI/(NI-ONE)
TP=TP/(WC/(NC-ONE))
VHAT=TP/NI+ONE/NC
VHAT=VHAT*VHAT
TP=(TP/NI)*(TP/NI)/(NI-ONE)
TP=TP+(ONE/NC)*(ONE/NC)/(NC-ONE)
VHAT=VHAT/TP

```

```
T=DTIN(ONE-ALPHA0/TWO,VHAT)
ENDIF
C*
IF (DABS(TSIM).GE.T) P=P+ONE
C*
2 CONTINUE
C*
P=P/NSIM
C*
WRITE(*,61) ' DELTAC =',DELTAC
WRITE(*,61) ' LAMBDA =',LAMBDA
WRITE(*,62) ' NI =',NI
WRITE(*,62) ' NC =',NC
IF (DELTAC.EQ.ZERO)
& WRITE(*,61) ' ESTIMATED ALPHA =',P
IF (DELTAC.NE.ZERO)
& WRITE(*,61) ' POWER =',P
61 FORMAT(A10,F9.5)
62 FORMAT(A10,I3)
C*
STOP
END
```