EQUIVALENCE TESTING FOR MEAN VECTORS OF MULTIVARIATE NORMAL POPULATIONS

A Dissertation by

Elizabeth Clarkson

M.S. Mathematics, Wichita State University, 1991

B.S. Mathematics, Wichita State University, 1983

Submitted to the Department of Mathematics
and the faculty of the Graduate School of
Wichita State University
in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

May 2010

EQUIVALENCE TESTING FOR MEAN VECTORS OF MULTIVARIATE NORMAL POPULATIONS

The following faculty members have examined the final copy of this dissertation for form and content, and recommend that it be accepted in partial fulfillment of the requirement for the degree of Doctor of Philosophy, with a major in Applied Mathematics.

Xiaomi Hu, Committee Chair

Dharam Chopra, Committee Member

Kirk Lancaster, Committee Member

Lop-Hing Ho, Committee Member

John Tomblin, Committee Member

Accepted for the College of Liberal Arts and Sciences

William D. Bischoff, Dean

Accepted for the Graduate School

J. David McDonald, Dean

DEDICATION

I dedicate this work to my husband Mark. He has devoted his life to our family and sacrificed his career for too many years in order that I might pursue my educational goals. He has never wavered in being my partner in all of life's up and downs, of which we have had many. That was true when I was an undergrad and we got married. It was just as true yesterday. I cannot begin to express my gratitude to him for everything he has brought to our lives over the past thirty years. In addition, he created figures 1, 2, and 8 for this work. He also produced nearly a dozen others for me that didn't end up in the final version due to changes in the structure of some of the proofs.

Thank you, Mark, for everything. I love you always.

ACKNOWLEDGEMENTS

Getting a PhD in mathematics while holding down a full-time job and parenting two children is not an easy task. I could not have achieved it without the help of a small army of supporters. I must start with my family; from my parents to my children and everyone in between, they have universally cheered me on the whole way. I am pleased to achieve this if only to honor their belief in my abilities.

My long-time friend, Prof. Kirk Lancaster, was instrumental in my choosing to work for the math department, and then deciding to work on a Ph.D. there. He was one of many such people in the math department, both supporting my efforts and inspiring me to continue. It was one of the best working environments I have ever had the pleasure of participating in. Throughout my nearly ten years of employment at WSU, they have been wonderful to work with: helpful and supportive of my reaching my educational goal. I would like to thank all of my many coworkers and acknowledge the contribution they have made.

My current boss, Yeow Ng, has supported my studies and provided advice and inspiration that was instrumental in the selection of both my advisor and my dissertation topic. I hope that this work will provide NCAMP with what he was hoping for.

Finally, I want to thank Dr. Hu, my advisor for his patience in leading me through the intricate details of the proofs needed for this task was phenomenal. I could never have achieved it without his steady attention to detail and guidance.

Thank you all very very much. I could not have done it alone.

ABSTRACT

This dissertation examines the problem of comparing samples of multivariate normal data from two populations and concluding whether the populations are equivalent; equivalence is defined as the distance between the mean vectors of the two samples being less than a given value.

Test statistics are developed for each of two cases using the ratio of the maximized likelihood functions. Case 1 assumes both populations have a common known covariance matrix. Case 2 assumes both populations have a common covariance matrix, but this covariance matrix is a known matrix multiplied by an unknown scalar value. The power function and bias of each of the test statistics is evaluated. Tables of critical values are provided.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS/NOMENCLATURE

cdf	Cumulative Probability Density Function
CTD	Cold Temperature Dry
ETW	Elevated Temperature Wet
FT	Fill Tension
LRT	Likelihood Ratio Test
pdf	Probability Density Function
RTD	Room Temperature Dry
WC	Warp Compression

SYMBOLS

α	specifies risk of Type I error in hypothesis test
δ	positive constant—specifies largest acceptable difference
σ	positive constant
ε	$\delta / \sqrt{\sigma}$
μ	population mean vector
Δ	difference between population mean vectors
Σ	k Known covariance matrix
$\bar{X}$	qualification sample mean vector
$\bar{Y}$	equivalence sample mean vector

CHAPTER 1

INTRODUCTION

This dissertation explores the use of multivariate analysis to perform acceptance sampling by employing a multivariate equivalence test. This economically feasible approach allows users to specify both the consumer's risk and the producer's risk. Given a new manufacturing facility or a change to a process procedure for a previously qualified material, it will allow engineering basis values to be set for the new procedure with a reduced dataset by making a comparison with the original qualification data. If the new product is sufficiently similar to the original qualification sample, then the two can be considered equivalent in terms of the engineering basis values.

1.1 Composite Materials Testing

Numerous tests are performed on a new composite material in order to compute the engineering basis values for that material. Engineers use these values to determine if a material is appropriate for a specific application. The tests are destructive, so sampling is the only option. The expense in determining engineering basis values is considerable; exacting tests are performed in environmental chambers to simulate the effects of extreme heat or cold on the material, while specialized equipment records precisely what stresses are required to break the specimen.

Data on composite materials from tests in the National Center for Advanced Materials Performance (NCAMP) are used as examples throughout this dissertation. The tests used were "fill compression," which refers to the direction of the material (fill) and the type of stress applied during the test (compressive). Test results analyzed are strength and modulus. Different environmental conditions included cold temperature dry (CTD) at -65°F, room temperature dry (RTD) at 75°F, and elevated temperature wet (ETW) at 200°F.

CHAPTER 2

BACKGROUND

First, it is necessary to understand some of the basic terms and concepts of acceptance sampling, engineering basis values, equivalency testing, and multivariate analysis.

2.1 Terminology

Some key terms relative to acceptance sampling follow:

- Producer's risk: The maximum probability of wrongly rejecting material that actually meets the specified criteria.
- Consumer's risk: The maximum probability of wrongly accepting material that does not actually meet the specified criteria
- B-basis value: An engineering value at the lower end of a 95% confidence interval for the 10^{th} percentile.
- A-basis value: An engineering value at the lower end of a 95% confidence interval for the 1^{st} percentile.
- Null hypothesis: The default assumption used to compute the probabilities above.
- Type I error: Incorrectly rejecting the default assumption when it is actually true.
- Type II error: Incorrectly failing to reject the default assumption when it is actually false.
- Power of a test: Probability of correctly rejecting the default assumption.

2.2 Acceptance Sampling

Acceptance sampling is the practice of accepting or rejecting an entire batch or shipment of material based on testing or inspecting a sample. The two possible default hypotheses are as follows: Either we can assume the new batch is acceptable and check to see if it is not, or we can assume the new batch is not acceptable and check to see if it is. With either one, there are two possible outcomes: Either the batch is accepted and released for use, or the batch is rejected and dispositioned. This leads to only two possible errors that can occur with acceptance sampling: Material is accepted that should have been rejected; the probability of this occurring is called the "consumer's risk." Or material is rejected that should have been accepted; the probability of this occurring is called the "producer's risk."

A puzzling aspect to the current standard practices of acceptance sampling is that, typically, any incoming supply has more than one key characteristic that must be monitored, yet sampling plans are almost universally set up for a single characteristic. A separate sampling plan is needed for each key characteristic being evaluated and makes an assumption that the key characteristics are independent.

Another puzzling aspect of current standard practices is that acceptance plans give probabilities for the producer's risk. This equates to the default hypothesis that the material is acceptable. For example, the sampling plans detailed in Mil-Std-105E, a very widely used set of acceptance sampling plans, are for a single characteristic indexed by the producer's risk. The question remains: Why aren't sampling plans based on the consumer's risk, since acceptance sampling plans are typically constructed by consumers for their own benefit?

2.2.1 Equivalence Testing for Acceptance Sampling

Acceptance sampling that specifies the consumer's risk does so by assuming that the samples are not acceptable. This type of testing is termed 'hypotheses of equivalence' and is rarely mentioned in discussions about acceptance sampling. Most people are unaware of which risk indexes sampling tables such as those found in MIL-STD-105E [1].

One reason that such an approach has not been used is the technical difficulty of computing probabilities for the consumer's risk. The computation requires specifying the largest non-zero difference considered equivalent. When sampling theory was being developed in the first half of the twentieth century, those computations simply were not feasible. But they could certainly have been performed in the past few decades with the necessary computing power that has been widely available.

Another problem is the power of this type of test. Theoretical limitations are imposed on testing equivalence hypotheses. Specifically, the power is limited to a maximum that is dependent not only on the sample size but also on δ. The smaller the value of δ, the lower the maximum achievable power of the test will be for any given set of sample sizes. The lower the power, the higher the producer's risk. This is illustrated in Table 1. For small values of δ, large sample sizes are required to achieve a reasonable producer's risk.

TABLE 1

MAXIMUM POWER OF UMP TEST AND CORRESPONDING PRODUCER'S RISK AT LEVEL A = 0.05 FOR ONE-SAMPLE EQUIVALENCE PROBLEM WITH GAUSSIAN DATA OF UNIT VARIANCE [2]

$\delta\sqrt{n} =$	0.1	0.5	1.0	2.0	3.0
Power	0.05025	0.05665	0.08229	0.32930	0.82465
Producer's Risk	0.94975	0.94335	0.91771	0.67070	0.17535

Because of this issue, plans that focus on the consumer's risk are too expensive to be practical, both for producers and for consumers.

2.3 Engineering Basis Values

Engineering A- and B-basis values are computed for key characteristics of a composite material from tests run under specified conditions, such as the tensile strength of a material in a cold, dry environment. These basis values become the reference for design engineers to use in designs to ensure that a part exposed to stresses, such as a strut in an airplane wing, is composed of materials that will not fail under that level of stress.

2.3.1 Current Computations for Engineering Basis Values of Composite Materials

Basis values are set using the mean and standard deviation of a sample of the material. This sample is referred to as the original qualification sample. Each key property, such as warp compression (WC) modulus or fill tension (FT) strength is tested under various environmental conditions, such as cold temperature dry or elevated temperature wet.

A variety of methods can be used to compute basis values; these include fitting a regression model over the different conditions, normalizing the data and pooling across the environmental conditions, or computing a basis value for each environmental condition individually. Basis values can be computed assuming that the data fit a normal distribution, a Weibull distribution, etc. Each method has certain advantages, and each makes certain assumptions about the distribution of the test data.

How well the data fits the various distributions and assumptions is then tested. The final selection of the mathematical model used to compute the basis values is dependent on the results of those tests. For example, the ANOVA approach is used when between-batch variability is large enough to preclude pooling the batches together within an environmental condition. The assumptions made when using the ANOVA approach are as follows [3]:

1. The data from each batch are normally distributed.
2. The within-batch variance is the same from batch to batch.
3. The batch means are normally distributed.

The model is then set up as follows:

$$x_{ij} = \mu_i + e_{ij}$$

where x_{ij} is test result for the j^{th} specimen in the i^{th} batch, μ_i is the batch mean, and e_{ij} is the error term with $\mu_i \sim n(\mu_., \sigma^2_\mu)$ and $e_{ij} \sim n(0, \sigma_e^2)$. This model assumes that $x_{ij} \sim n(\mu_., \sigma^2_\mu + \sigma_e^{\ 2})$.

The methodology developed in this dissertation relies upon similar assumptions and extends the model to a multivariate normal distribution of test results.

The ANOVA approach uses the population variance to compute the engineering basis values as follows:

$$B-basis = \overline{X} - T_B S$$
$$A-basis = \overline{X} - T_A S$$

In this situation, $\overline{X}$ is the qualification sample mean, *S* represents the estimate of the population variance based on the ANOVA analysis, and *T* is a computed factor [3]. This methodology is relatively robust to deviations from the normality or equal variation assumptions and provides a conservative result when that assumption fails.

2.4 Equivalency Tests for Composite Materials

To determine if a new facility or procedure will produce material capable of meeting the basis values computed from the qualification sample, a smaller 'equivalency sample' is produced, and tests from that sample are compared to the results of the previous tests on the qualification sample.

2.4.1 Current Equivalency Method

For each property tested, a separate comparison is made for each environmental condition. The final decision regarding equivalence is based on using engineering judgment to subjectively assess all test results to arrive at a yes or no decision regarding the equivalence of the new material with the original material.

Tests are conducted as follows: Modulus values are compared using a two-tailed t-test, while strength values are compared with a one-tailed test using the mean and minimum value. [4] Separate independent tests are performed for modulus and strength. Formally, the test hypotheses are set up as follows:

For modulus values:

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_1 : \mu_1 - \mu_2 \neq 0$$

For strength values:

$$H_0 : \mu_1 - \mu_2 \leq 0$$
$$H_1 : \mu_1 - \mu_2 > 0$$

where μ_1 is the mean of the qualification material for the characteristic being tested, and μ_2 is the mean of the material being compared for equivalence. The second material might come from a different manufacturing environment, or it might be that the manufacturer wishes to make a change to the manufacturing process. Either way, before the new process can claim the use of the basis values and other characteristics previously established for the material, the equivalence of the final product to the original material must be established. If the material is not found equivalent, additional testing is required to establish the characteristics of the material coming from the new facility or changed procedure.

Note that the default assumption is that the two samples are from identically distributed populations. If the sample fails the test, this assumption is rejected at the specified level of confidence. If the sample does not fail the test, the probability that the two samples are the same is equivalent to the power of the test. For the sample size typically used in the testing of composite materials, the power is considerably lower than the confidence level used to determine if it is not equivalent. "*A nonsignificant difference must not be confused with significant homogeneity*" [2], yet this is exactly what our current method does.

2.4.2 Disadvantages of the Current Method

One disadvantage of this approach is that it only looks at individual test results for comparison. No use is made of relationships between the characteristics being evaluated. This approach has the unintended side effect of producers benefitting from smaller sample sizes. Smaller sample sizes decrease the power of the test, which means the probability of a Type II error occurring is larger for smaller samples. Since the default assumption is that the new product is equivalent and the test determines whether or not to reject that assumption, a Type II error means to accept material as equivalent when it actually is not. If the null and alternative hypotheses were flipped around and it was assumed that the product was *not* equivalent, this side effect would disappear.

Another problem is that given the number of individual tests compared with a 95 percent level of confidence, the probability that at least one test will fail due to random chance alone is quite high. For example, with 30 tests, the probability of having at least one failure is 0.785. This equates to a producer's risk of more than 20 percent if the material were rejected for a single test failure. In fact, it is extremely uncommon for any equivalency sample to pass all tests that are run. This is why subjective engineering judgment is a major part of the process in deciding whether or not two facilities producing the same material can be considered equivalent.

2.5 Multivariate Tests

Multivariate tests examine multiple characteristics simultaneously and uses the expected relationship between them as part of the criteria used to judge similarity between the two samples. A primary advantage of the multivariate approach is that it allows for the inclusion of information about the relationships between different characteristics, rather than evaluating each characteristic in isolation when making the overall judgment about acceptance. Another advantage of multivariate testing is that it reduces the subjectivity of the overall choice by replacing a decision based on the subjective weighting of many different test results with an objective decision based on the combined results of the different tests.

With the advantages of multivariate testing, it is logical to ask why it is not in use. One reason is the computational difficulty. Another reason that the multivariate approach has not been popular is because, under the traditional null hypothesis of equality, it would result in nearly always rejecting the null since the more information used when comparing two groups, the more likely some minute difference will be found. Since the default assumption is that they are the same, any tiny but statistically significant difference results in a rejection of the null hypothesis. Thus, multivariate acceptance testing has been of limited practical use.

2.5.1 Multivariate Hypotheses of Equivalence

While it seems counter-intuitive, combining multivariate testing with the hypothesis of equivalence can overcome both sets of problems. When acceptance limits are set based on the consumer's risk, there must be some positive value ($\delta > 0$) such that a deviation of less than δ from nominal for the sample being evaluated is considered acceptable. In addition, a measurement of the distance between two multivariate vectors is needed. This measurement will be defined in the next chapter.

One advantage of this approach is that δ can be used to control the producer's risk simultaneously with the consumer's risk. If δ is defined as a multiple of the standard deviation, then a value for δ corresponding to any desired producer's risk can be found.

One consequence of testing hypotheses of equivalence is that the A- and B-basis values used must be adjusted downward. If the mean could possibly deviate from the nominal mean vector by as much as δ and still be considered equivalent, then the engineering basis values must be computed from the lowest possible acceptable mean rather than the qualification sample mean.

At NCAMP, researchers are currently in the process of developing engineering basis values and computing the results of equivalency tests simultaneously. Therefore we are in a unique position to develop and implement a strategy that would set basis values and acceptance limits using a multivariate approach combined with testing hypotheses of equivalence.

2.6 Literature Review

The principle work on equivalence testing is Wellek's *Testing Statistical Hypotheses of Equivalence* [2], of which chapter nine covers the bivariate normal equivalence test and indicates what assumptions are needed for the multivariate approach to equivalence testing and what form an extension of that approach should take. This dissertation extends that work to the multivariate situation and also expands it to include the situation where the common covariance matrix is an unknown multiple of a known covariance matrix.

For analysis of a multivariate normal random variable, the Anderson's venerable *An Introduction to Multivariate Statistical Analysis* [5] is a classic and immensely helpful in understanding the details of multivariate distributions.

For understanding how those details fit into an application, such as the one developed here, Johnson and Wichern's *Applied Multivariate Statistical Analysis* [6] was invaluable. *Matrix Analysis for Statistics* by Schott [7] was also a contributing resource to developing this theory.

Hoag and Craig's *Introduction to Mathematical Statistics* [8] and Shorack's *Probability for Statisticians* [9] were used for basic statistical theory.

The main theorem of this dissertation relies on techniques borrowed from "Monotonicity Properties of the Power Functions of Likelihood Ratio Tests for Normal Mean Hypotheses Constrained by a Linear Space and a Cone," by Hu and Wright [10] and "*The Integral of a Symmetric Unimodal Function Over a Symmetric Convex Set and Some Probability Inequalities*" by Anderson [11].

Dr. Hu’s article in the February 2007 issue of *The American Statistician,* “Teacher’s Corner” [12] column, regarding notation for multivariate normal distributions and some theorems that are easily derived using that notation, was extremely useful.

Finally, it is worth mentioning that sites like Mathworld.com and Wikipedia were invaluable for accessing and verifying basic information before going on to the next step in a proof or a program.

CHAPTER 3

THE MATH OF IT ALL

This chapter contains the math of it all, a detailed mathematical description and analysis of the problem of a multivariate equivalence test. It requires the user to specify two things:

- δ, the tolerable difference; a population that differs from the expected mean by a value of less than δ is defined as equivalent (what an engineer would call "close enough").
- α, the maximum probability of incorrectly rejecting the null hypothesis.

3.1 Problem Statement

When comparing two or more samples, some applications need to test (at the α-level of significance) that the samples come from equivalent populations rather than the more typical determination that the two samples are from different populations. This thesis defines a procedure to compare multivariate normal data sampled from two groups and to conclude that the two samples are from equivalent populations; that is, the groups differ by less than the given amount, δ.

Assume that two samples of size n_1 and n_2 of p-vectors come from multivariate normal distributions with means μ_1 and μ_2, and having a common covariance matrix, Σ. In the application of this research, equivalence testing of composite materials, the natural choice for the experimental unit is the panel. Multiple tests of each type are performed on specimens from each panel. The mean result of those tests by panel will have a multivariate normal distribution with a mean vector identical to that of the underlying distribution.

3.1.1 Measurement

The measure of the distance between the two mean vectors will be the norm of the difference between the two mean vectors. This norm is the induced norm from the following inner product:

$$\langle v_1, v_2 \rangle = v_1' \Sigma^{-1} v_2,\ v_1, v_2 \in \mathbf{R}^p \tag{1}$$

3.1.2 Definitions

The following definitions will be used:

$\Delta = \mu_1 - \mu_2$ − The difference vector of the two mean vectors: $\mu_1, \mu_2 \in \mathbf{R}^p$
$m = n_1 + n_2$ − The size of the combined samples: $n_1, n_2 \in \mathbf{N}$ where $\mathbf{N} = \{1, 2, 3, \cdots\}$
$\|\bullet\| = \sqrt{\langle \bullet, \bullet \rangle}$ − The norm is the square root of the inner product defined in (1)

3.1.3 Statement of Hypothesis

Given $\delta > 0$, "equivalent populations" are defined as those populations with mean vectors with a normed difference of less than δ. Define $\Theta_0 = \left\{ v \in \mathbf{R}^p : \|v\| \geq \delta \right\}$. If the null hypothesis is true, then $\Delta \in \Theta_0$. This is formally stated as

$$H_0 : \Delta \in \Theta_0$$
$$H_1 : \Delta \notin \Theta_0$$

This hypothesis flips the typical null, and the alternative hypothesis equality between the two mean vectors is part of the alternative rather than the null hypothesis. Thus, when the null is rejected, we can state with confidence that the difference between the two populations is "close enough" rather than simply failing to reject the null hypothesis that they are the same.

This hypothesis will be tested using a likelihood ratio test (LRT). The LRT requires a test statistic constructed from the ratio of the maximum value of the likelihood function over the entire space $\left(\mathbf{R}^p\right)$ to the maximum value of the likelihood function over the null space (Θ_0).

Figure 1 shows $\mathbf{\Theta_0}$ and its compliment for $\mathbf{R}^2$. $\mathbf{\Theta_0}$ and its compliment are bounded by the solid line with $\mathbf{\Theta_0}$, including the boundary and everything outside of it. Its compliment is the area inside the solid line but not including the boundary.

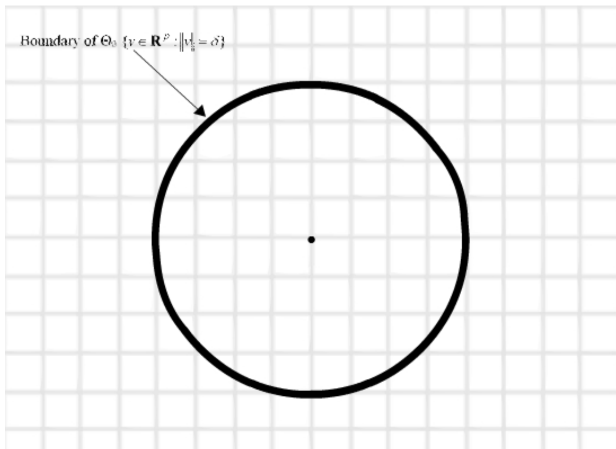

Figure 1. Θ_0 and rejection region in two dimensions.

3.1.4 Case 1 and Case 2

In Case 1, the common covariance matrix is assumed to be a known positive definite matrix, $\mathbf{\Sigma}$. In Case 2, the common covariance matrix is assumed to be an unknown positive scalar multiple of a known positive definite matrix, $\sigma\mathbf{\Sigma}$. Case 1 can be considered a particular instance (σ=1) of the more general problem expressed in Case 2.

3.2 Case 1

3.2.1 Sample Distributions

The formal description of these sample distributions is

$$\left.\begin{array}{l} X_1,\ldots,X_{n_1} \sim N_p\left(\mu_1,\Sigma\right) \\ Y_1,\ldots,Y_{n_2} \sim N_p\left(\mu_2,\Sigma\right) \end{array}\right\} \begin{array}{l} \mu_1,\mu_2 \text{ unknown, } \Sigma \text{ known} \\ \mu_1,\mu_2 \in \mathbf{R}^p, \Sigma \in \mathbf{R}^{pxp} \end{array}$$

3.2.2 Joint Probability Density Function

The joint probability density function (pdf) of the m independent random vectors from the samples is the product of their individual probability density functions (pdfs). For our p-dimensional multinormal sample, this is

$$\prod_{i=1}^{n_1} \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} e^{-\frac{(X_i-\mu_1)'\Sigma^{-1}(X_i-\mu_1)}{2}} \cdot \prod_{j=1}^{n_2} \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} e^{-\frac{(Y_j-\mu_2)'\Sigma^{-1}(Y_j-\mu_2)}{2}}$$

3.2.3 Likelihood Function L(μ₁, μ₂)

The joint probability density function may be regarded as a function of the parameters μ_1 and μ_2. When so regarded, it is denoted by $L(\mu_1, \mu_2)$ and called the likelihood function. [8] This function can be used to determine the likelihood of any particular set of parameter values or to find the parameter values with the largest likelihood given the sample data collected. The likelihood function can be expressed as

$$L(\mu_1, \mu_2) = \frac{1}{(2\pi)^{mp/2}|\Sigma|^{m/2}} e^{-\frac{1}{2}\left(\sum_{i=1}^{n_1}\|X_i-\mu_1\|^2 + \sum_{j=1}^{n_2}\|Y_j-\mu_2\|^2\right)}$$

3.2.4 Maximized Likelihood Function L(μ₁, μ₂) Without Restrictions

The principle of maximum likelihood is the idea that given a particular set of sample values, we can find a function of those sample values such that when the parameter value is set equal to that function of sample values, the likelihood function is maximized [8].

Define $A(\mu_1, \mu_2) = \sum_{i=1}^{n_1}\|X_i-\mu_1\|^2 + \sum_{j=1}^{n_2}\|Y_j-\mu_2\|^2$ (2)

Theorem 1: $\max L(\mu_1, \mu_2) = L(\bar{X}, \bar{Y}) = \frac{1}{(2\pi)^{mp/2}|\Sigma|^{m/2}} e^{-\frac{A(\bar{X},\bar{Y})}{2}}$

Proof: $L(\mu_1,\mu_2)=\frac{1}{(2\pi)^{mp/2}|\Sigma|^{m/2}}e^{-\frac{A(\mu_1,\mu_2)}{2}}$ and it is clear that $L(\mu_1,\mu_2)$ will be maximized when $A(\mu_1,\mu_2)$ is minimized. $A(\mu_1,\mu_2)$ can be further decomposed as

$$A(\mu_1,\mu_2)=\sum_{i=1}^{n_1}\left\|X_i-\bar{X}\right\|^2+\sum_{j=1}^{n_2}\left\|Y_j-\bar{Y}\right\|^2+g(\mu_1,\mu_2)$$

with $$g(\mu_1,\mu_2)=n_1\left\|\bar{X}-\mu_1\right\|^2+n_2\left\|\bar{Y}-\mu_2\right\|^2 \qquad (3)$$

Clearly $g(\mu_1,\mu_2)\geq 0$ with $g(\bar{X},\bar{Y})=0$, and the remaining terms do not contain the parameters μ_1 or μ_2. Thus,

$$A(\mu_1,\mu_2)\geq A(\bar{X},\bar{Y})=\sum_{i=1}^{n_1}\left\|X_i-\bar{X}\right\|^2+\sum_{j=1}^{n_2}\left\|Y_j-\bar{Y}\right\|^2$$

and

$$\max L(\mu_1,\mu_2)=L(\bar{X},\bar{Y})=\frac{1}{(2\pi)^{\frac{pm}{2}}|\Sigma|^{\frac{m}{2}}}e^{-\frac{A(\bar{X},\bar{Y})}{2}} \qquad \square$$

3.2.5 Minimum Distance Projection

Let D be a set on $\mathbf{R}^p$ and v be a vector in $\mathbf{R}^p$. A minimum distance projection of v onto D, denoted by $P(v)$, has the following two characteristics:

- $P(v)\in D$
- $\left\|P(v)-v\right\|\leq\left\|w-v\right\| \quad \forall w\in D$

If D is a closed set in $\mathbf{R}^p$, it is known that the projection, $P(v)$, exists. The projection is unique if D is also convex.

3.2.5.1 Minimum Distance Projection from $\mathbf{R}^p \to \Theta_0$

Figure 2 is a diagram showing a minimum distance projection for the two-dimensional situation.

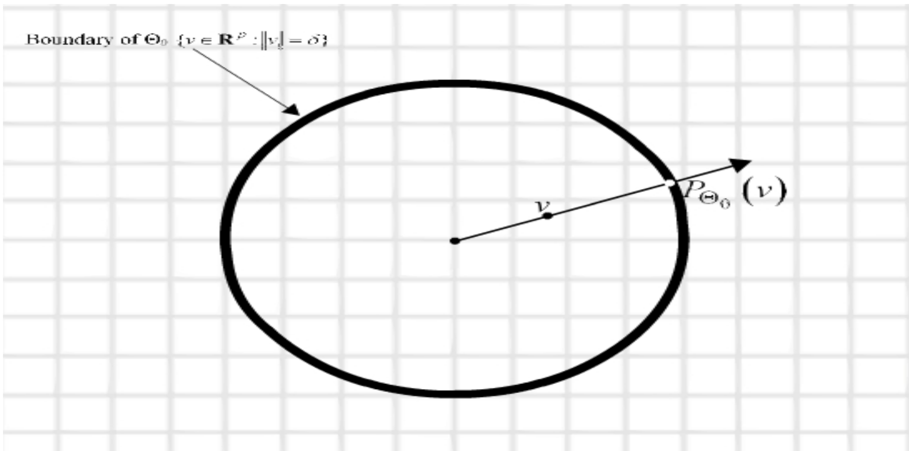

Figure 2. Projection of point v in the compliment of Θ_0 onto Θ_0.

Since Θ_0 is closed, for all $v \in \mathbf{R}^p$, a projection of v onto Θ_0 must exist. Define $P_{\Theta_0}(v)$ as follows:

$$P_{\Theta_0}(v) = \begin{cases} \dfrac{\delta}{\|y\|} y & \text{if } \|v\| = 0 \ \forall\, y \in \mathbf{R}^p s.t. \|y\| \neq 0 \\ v & \text{if } \|v\| \geq \delta \\ \dfrac{\delta}{\|v\|} v & \text{if } 0 < \|v\| < \delta \end{cases}$$

Lemma 1: With $P_{\Theta_0}(v)$ defined as above, $P_{\Theta_0}(v)$ is the minimum distance projection of v onto Θ_0

Proof: The proof is established by considering three cases: $\|v\| = \mathbf{0}$, $\|v\| \geq \delta$, and $0 < \|v\| < \delta$.

Case 1: $\|v\| = 0$

a) $\left\|P_{\Theta_0}(v)\right\| = \left\|\dfrac{\delta}{\|y\|} y\right\| = \delta \Rightarrow P_{\Theta_0}(v) \in \Theta_0$

b) $\left\|P_{\Theta_0}(v) - v\right\| = \left\|\dfrac{\delta}{\|y\|} y - \mathbf{0}\right\| = \delta \leq \|w - \mathbf{0}\| = \|w - v\| \quad \forall w \in \Theta_0$

Case 2: $\|v\| \geq \delta$

a) $\left\|P_{\Theta_0}(v)\right\| = \|v\| \geq \delta \Rightarrow P_{\Theta_0}(v) = v \in \Theta_0$

b) $\left\|P_{\Theta_0}(v) - v\right\| = \|v - v\| = \mathbf{0} \leq \|w - v\| \quad \forall w \in \Theta_0$

Case 3: $0 < \|v\| < \delta$

a) $\left\|P_{\Theta_0}(v)\right\| = \left\|\dfrac{\delta}{\|v\|} v\right\| = \delta \Rightarrow P_{\Theta_0}(v) \in \Theta_0$

b) $\left\|P_{\Theta_0}(v) - v\right\| = \left\|\frac{\delta}{\|v\|} v - v\right\| = \left(\frac{\delta}{\|v\|} - 1\right)\|v\| = \delta - \|v\| \leq \|w\| - \|v\| \, \forall w \in \Theta_0$

By the triangle inequality,

$\|w\| = \|w - v + v\| \leq \|w - v\| + \|v\| \Rightarrow \|w\| - \|v\| \leq \|w - v\|$

$\Rightarrow \left\|P_{\Theta_0}(v) - v\right\| \leq \|w - v\| \, \forall w \in \Theta_0$

Thus, $P_{\Theta_0}(v)$ is the minimum distance projection of v onto Θ_0 . □

This projection is not unique, since when $\|v\| = 0$, any non-zero vector in $\mathbf{R}^p$ can be selected as y. However, $\left\|v - P_{\Theta_0}(v)\right\|$ is unique, regardless of the vector selected.

Lemma 2: Let $P_{\Theta_0}(v)$ be the projection of v onto Θ_0 derived in Lemma 1. Then

$$\left\|v - P_{\Theta_0}(v)\right\| = \begin{cases} 0 & \text{if } v \in \Theta_0 \Rightarrow \|v\| \geq \delta \\ \delta - \|v\| & \text{otherwise} \Rightarrow 0 \leq \|v\| < \delta \end{cases}$$

Proof:

Case 1: When $\|v\| \geq \delta$, then $\left\|v - P_{\Theta_0}(v)\right\| = \|v - v\| = 0$

Case 2: When $0 \leq \|v\| < \delta$

a) If $\|v\| = 0$, then $\left\|v - P_{\Theta_0}(v)\right\| = \left\| v - \frac{\delta}{\|y\|} y \right\| = \frac{\delta}{\|y\|}\|y\| = \delta - \|v\|$

b) If $0 < \|v\| < \delta$,

then $\left\|v - P_{\Theta_0}(v)\right\| = \left\| v - \frac{\delta}{\|v\|} v \right\| = \left|1 - \frac{\delta}{\|v\|}\right| \|v\| = \delta - \|v\|$ □

3.2.6 Maximum Likelihood Function L(μ₁, μ₂) under Restriction that Δ is in Θ_0

Some lemmas will be needed before the maximum of the likelihood function under this restriction can be proven.

Let $g(\mu_1, \mu_2)$ be defined as in equation (3). Then

Lemma 3: $g(\mu_1,\mu_2) = \frac{n_1 n_2}{m}\left\|\bar{X} - \bar{Y} - \Delta\right\|^2 + \frac{1}{m}\left\|n_1\left(\bar{X} - \mu_1\right) + n_2\left(\bar{Y} - \mu_2\right)\right\|^2$

Proof: Define $u = \frac{n_1\bar{X} + n_2\bar{Y} + n_2\Delta}{m}$ and $v = \frac{n_1\bar{X} + n_2\bar{Y} - n_1\Delta}{m}$

Then from equation (3): $g(\mu_1,\mu_2) = n_1\left\|\bar{X} - \mu_1\right\|^2 + n_2\left\|\bar{Y} - \mu_2\right\|^2$

$$= n_1\left\|\bar{X} - u + u - \mu_1\right\|^2 + n_2\left\|\bar{Y} - v + v - \mu_2\right\|^2$$

$$= t_1(\mu_1,\mu_2) + t_2(\mu_1,\mu_2) + t_3(\mu_1,\mu_2)$$

where $t_1(\mu_1,\mu_2) = n_1\left\|\bar{X} - u\right\|^2 + n_2\left\|\bar{Y} - v\right\|^2$

$t_2(\mu_1,\mu_2) = n_1\left\|u - \mu_1\right\|^2 + n_2\left\|v - \mu_2\right\|^2$

$t_3(\mu_1,\mu_2) = 2n_1\left\langle \bar{X} - u, u - \mu_1\right\rangle + 2n_2\left\langle \bar{Y} - v, v - \mu_2\right\rangle$

But $t_1(\mu_1,\mu_2) = n_1\left\|\bar{X} - u\right\|^2 + n_2\left\|\bar{Y} - v\right\|^2$

$$= n_1 \left\| \frac{m\bar{X}}{m} - \frac{n_1\bar{X} + n_2\bar{Y} + n_2\Delta}{m} \right\|^2 + n_2 \left\| \frac{m\bar{Y}}{m} - \frac{n_1\bar{X} + n_2\bar{Y} - n_1\Delta}{m} \right\|^2$$

$$= n_1 \left\| \frac{n_2\bar{X} - n_2\bar{Y} - n_2\Delta}{m} \right\|^2 + n_2 \left\| \frac{-n_1\bar{X} + n_1\bar{Y} + n_1\Delta}{m} \right\|^2$$

$$= \frac{n_1 n_2^2}{m^2} \left\| \bar{X} - \bar{Y} - \Delta \right\|^2 + \frac{n_1^2 n_2}{m^2} \left\| \bar{X} - \bar{Y} - \Delta \right\|^2$$

$$= \frac{n_1 n_2}{m} \left\| \bar{X} - \bar{Y} - \Delta \right\|^2,$$

$$t_2(\mu_1, \mu_2) = n_1 \|u - \mu_1\|^2 + n_2 \|v - \mu_2\|^2$$

$$= n_1 \left\| \frac{n_1\bar{X} + n_2\bar{Y} + n_2\mu_1 - n_2\mu_2}{m} - \frac{m\mu_1}{m} \right\|^2 + n_2 \left\| \frac{n_1\bar{X} + n_2\bar{Y} - n_1\mu_1 + n_1\mu_2}{m} - \frac{m\mu_2}{m} \right\|^2$$

$$= \frac{n_1}{m^2} \left\| n_1\bar{X} + n_2\bar{Y} - n_1\mu_1 - n_2\mu_2 \right\|^2 + \frac{n_2}{m^2} \left\| n_1\bar{X} + n_2\bar{Y} - n_1\mu_1 - n_2\mu_2 \right\|^2$$

$$= \frac{n_1 + n_2}{m^2} \left\| n_1\bar{X} + n_2\bar{Y} - n_1\mu_1 - n_2\mu_2 \right\|^2$$

$$= \frac{1}{m} \left\| n_1\left(\bar{X} - \mu_1\right) + n_2\left(\bar{Y} - \mu_2\right) \right\|^2,$$

and $t_3(\mu_1, \mu_2) = 2n_1 \left\langle \bar{X} - u, u - \mu_1 \right\rangle + 2n_2 \left\langle \bar{Y} - v, v - \mu_2 \right\rangle$

$$= 2n_1 \left\langle \bar{X} - \frac{n_1\bar{X} + n_2\bar{Y} + n_2(\mu_1 - \mu_2)}{m}, \frac{n_1\bar{X} + n_2\bar{Y} + n_2(\mu_1 - \mu_2)}{m} - \mu_1 \right\rangle$$

$$+ 2n_2 \left\langle \bar{Y} - \frac{n_1\bar{X} + n_2\bar{Y} - n_1(\mu_1 - \mu_2)}{m}, \frac{n_1\bar{X} + n_2\bar{Y} - n_1(\mu_1 - \mu_2)}{m} - \mu_2 \right\rangle$$

$$= 2n_1 \left\langle \frac{m\bar{X} - \left(n_1\bar{X} + n_2\bar{Y} + n_2(\mu_1 - \mu_2)\right)}{m}, \frac{n_1\bar{X} + n_2\bar{Y} + n_2(\mu_1 - \mu_2) - m\mu_1}{m} \right\rangle$$

$$+ 2n_2 \left\langle \frac{m\bar{Y} - \left(n_1\bar{X} + n_2\bar{Y} - n_1(\mu_1 - \mu_2)\right)}{m}, \frac{n_1\bar{X} + n_2\bar{Y} - n_1(\mu_1 - \mu_2) - m\mu_2}{m} \right\rangle$$

$$= \frac{2n_1}{m^2} \left\langle n_2\bar{X} - n_2\bar{Y} + n_2(\mu_1 - \mu_2), n_1\bar{X} + n_2\bar{Y} - n_1\mu_1 - n_2\mu_2 \right\rangle$$

$$+ \frac{2n_2}{m^2} \left\langle -n_1\bar{X} + n_1\bar{Y} + n_1(\mu_1 - \mu_2), n_1\bar{X} + n_2\bar{Y} - n_1\mu_1 - n_2\mu_2 \right\rangle$$

$$= \frac{2n_1n_2}{m^2} \left\langle \bar{X} - \bar{Y} - (\mu_1 - \mu_2), n_1(\bar{X} - \mu_1) + n_2(\bar{Y} - \mu_2) \right\rangle$$

$$- \frac{2n_1n_2}{m^2} \left\langle \bar{X} - \bar{Y} - (\mu_1 - \mu_2), n_1(\bar{X} - \mu_1) + n_2(\bar{Y} - \mu_2) \right\rangle$$

$$= 0.$$

The conclusion follows from the fact that

$$g(\mu_1, \mu_2) = t_1(\mu_1, \mu_2) + t_2(\mu_1, \mu_2) + t_3(\mu_1, \mu_2). \qquad \square$$

Lemma 4: $\min_{\Delta \in \Theta_0} g(\mu_1,\ \mu_2)$ occurs at the following parameter values:

$$\hat{\mu}_1 = \frac{n_1\bar{X} + n_2\bar{Y} + n_2 P_{\Theta_0}(\bar{X} - \bar{Y})}{m},\ \hat{\mu}_2 = \frac{n_1\bar{X} + n_2\bar{Y} - n_1 P_{\Theta_0}(\bar{X} - \bar{Y})}{m}$$

Proof: First we claim $\hat{\mu}_1 - \hat{\mu}_2 \in \Theta_0$ since

$$\hat{\mu}_1 - \hat{\mu}_2 = \frac{n_1\bar{X} + n_2\bar{Y} + n_2 P_{\Theta_0}(\bar{X} - \bar{Y})}{m} - \frac{n_1\bar{X} + n_2\bar{Y} - n_1 P_{\Theta_0}(\bar{X} - \bar{Y})}{m}$$

$$= \frac{(n_1 + n_2) P_{\Theta_0}(\bar{X} - \bar{Y})}{m} = P_{\Theta_0}(\bar{X} - \bar{Y}) \in \Theta_0$$

Second, the first term in $g(\mu_1,\ \mu_2)$ is minimized at $\hat{\mu}_1$ and $\hat{\mu}_2$ since

$$t_1(\mu_1, \mu_2) = \frac{n_1n_2}{m} \left\| \bar{X} - \bar{Y} - \Delta \right\|^2 \geq \frac{n_1n_2}{m} \left\| \bar{X} - \bar{Y} - P_{\Theta_0}(\bar{X} - \bar{Y}) \right\|^2 = t_1(\hat{\mu}_1, \hat{\mu}_2)$$

Finally, the second term of $g(\mu_1,\ \mu_2)$ is also minimized at $\hat{\mu}_1$ and $\hat{\mu}_2$:

$$t_2(\hat{\mu}_1, \hat{\mu}_2) = \frac{1}{m}\left\| n_1\left(\bar{X} - \frac{n_1\bar{X} + n_2\bar{Y} + n_2 P_{\Theta_0}(\bar{X} - \bar{Y})}{m}\right) + n_2\left(\bar{Y} - \frac{n_1\bar{X} + n_2\bar{Y} - n_1 P_{\Theta_0}(\bar{X} - \bar{Y})}{m}\right)\right\|^2$$

$$= \frac{1}{m}\left\|\left(\frac{n_1 n_2\bar{X} - n_1 n_2\bar{Y} - n_1 n_2 P_{\Theta_0}(\bar{X} - \bar{Y})}{m}\right) + \left(\frac{-n_1 n_2\bar{X} + n_1 n_2\bar{Y} + n_1 n_2 P_{\Theta_0}(\bar{X} - \bar{Y})}{m}\right)\right\|$$

$$= 0 \le t_2(\mu_1, \mu_2). \qquad \square$$

Theorem 2: $\max\left[L(\mu_1, \mu_2) : \Delta \in \Theta_0\right] = L(\hat{\mu}_1, \hat{\mu}_2)$

$$= \frac{1}{(2\pi)^{\frac{pm}{2}} |\Sigma|^{\frac{m}{2}}} \exp\left[-\frac{1}{2}\left(\sum_{i=1}^{n_1}\left\|X_i - \bar{X}\right\|^2 + \sum_{j=1}^{n_2}\left\|Y_j - \bar{Y}\right\|^2 + \frac{n_1 n_2}{m}\left\|\bar{X} - \bar{Y} - P_{\Theta_0}(\bar{X} - \bar{Y})\right\|^2\right)\right]$$

Proof: In section 3.2.4, we established that $L(\mu_1, \mu_2)$ depends on μ_1 and μ_2 only through $g(\mu_1, \mu_2)$ and that $L(\mu_1, \mu_2)$ is a decreasing function of $g(\mu_1, \mu_2)$. Lemma 4 shows that $g(\mu_1, \mu_2)$ is minimized at $\hat{\mu}_1, \hat{\mu}_2$ under the restriction $\Delta \in \Theta_0$. Thus, $L(\mu_1, \mu_2)$ is maximized at $\hat{\mu}_1, \hat{\mu}_2$ under $\Delta \in \Theta_0$, i.e., $\max_{\Delta \in \Theta_0} L(\mu_1, \mu_2) = L(\hat{\mu}_1, \hat{\mu}_2)$. Direct computation shows that

$$L(\hat{\mu}_1, \hat{\mu}_2) = \frac{1}{(2\pi)^{\frac{pm}{2}} |\Sigma|^{\frac{m}{2}}} \exp\left[-\frac{1}{2}\left(\sum_{i=1}^{n_1}\left\|X_i - \bar{X}\right\|^2 + \sum_{j=1}^{n_2}\left\|Y_j - \bar{Y}\right\|^2 + \frac{n_1 n_2}{m}\left\|\bar{X} - \bar{Y} - P_{\Theta_0}(\bar{X} - \bar{Y})\right\|^2\right)\right].$$

Thus, the theorem is established. $\square$

3.2.7 Ratio of Maximized Likelihood Functions

Let Λ represent the ratio of the two maximized likelihood functions:

$$\Lambda = \frac{\max L(\mu_1, \mu_2)}{\max\left(L(\mu_1, \mu_2) : \Delta \in \Theta_0\right)}$$

By Theorems 1 and 2,

$$\Lambda=\frac{L(\bar{X},\bar{Y})}{L(\hat{\mu}_1,\hat{\mu}_2)}=\frac{\frac{1}{(2\pi)^{\frac{pm}{2}}|\Sigma|^{\frac{m}{2}}}\exp\left[-\frac{1}{2}\left(\sum_{i=1}^{n_1}\left\|X_i-\bar{X}\right\|^2+\sum_{j=1}^{n_2}\left\|Y_j-\bar{Y}\right\|^2\right)\right]}{\frac{1}{(2\pi)^{\frac{pm}{2}}|\Sigma|^{\frac{m}{2}}}\exp\left[-\frac{1}{2}\left(\sum_{i=1}^{n_1}\left\|X_i-\bar{X}\right\|^2+\sum_{j=1}^{n_2}\left\|Y_j-\bar{Y}\right\|^2+\frac{n_1n_2}{m}\left\|\bar{X}-\bar{Y}-P_{\Theta_0}\left(\bar{X}-\bar{Y}\right)\right\|^2\right)\right]}$$

$$=\exp\left[\frac{n_1n_2}{2m}\left\|\bar{X}-\bar{Y}-P_{\Theta_0}\left(\bar{X}-\bar{Y}\right)\right\|^2\right].$$

Hence, Λ is a non-decreasing function of $\left\|\bar{X}-\bar{Y}-P_{\Theta_0}\left(\bar{X}-\bar{Y}\right)\right\|^2$.

3.2.8 Likelihood Ratio Test (LRT) Statistic

By the relationship expressed in Lemma 2:

$$\left\|\bar{X}-\bar{Y}-P_{\Theta_0}\left(\bar{X}-\bar{Y}\right)\right\|^2=\begin{cases}0 & \bar{X}-\bar{Y}\in\Theta_0\\ \left(\delta-\left\|\bar{X}-\bar{Y}\right\|\right)^2 & \bar{X}-\bar{Y}\notin\Theta_0\end{cases}$$

$\left\|\bar{X}-\bar{Y}-P_{\Theta_0}\left(\bar{X}-\bar{Y}\right)\right\|^2$ is a non-increasing function of $\left\|\bar{X}-\bar{Y}\right\|$, as shown in Figure 3.

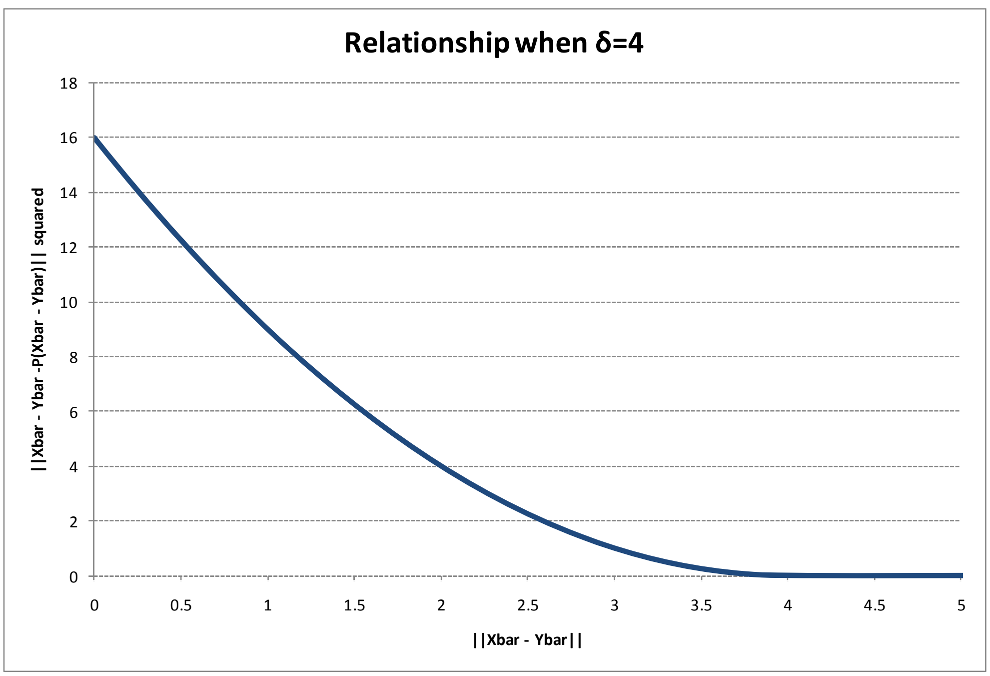

Figure 3. Relationship of $\left\|\bar{X}-\bar{Y}\right\|$ with $\left\|\bar{X}-\bar{Y}-P_{\Theta_0}\left(\bar{X}-\bar{Y}\right)\right\|^2$.

Λ is a non-decreasing function of $\left\| \bar{X} - \bar{Y} - P_{\Theta_0}\left(\bar{X} - \bar{Y}\right)\right\|^2$, which is itself a non-increasing function of $\left\| \bar{X} - \bar{Y} \right\|$, and since $\left\| \bar{X} - \bar{Y} \right\|$ is non-negative, it is also a non-decreasing function of $\left\| \bar{X} - \bar{Y} \right\|^2$. This means that $T = \frac{n_1 n_2}{m}\left\| \bar{X} - \bar{Y} \right\|^2$ can be used as our test statistic. We can reject the null hypothesis when T is sufficiently small.

Case 1 Test Statistic: $$T = \frac{n_1 n_2}{m}\left\| \bar{X} - \bar{Y} \right\|^2 \tag{4}$$

3.2.9 Distribution of *T*

Theorem 3: $T \sim \chi_p^2\left(\frac{n_1 n_2}{m}\left\| \Delta \right\|^2\right)$

Proof: $T = \frac{n_1 n_2}{m}\left\| \bar{X} - \bar{Y} \right\|^2 = \frac{n_1 n_2}{m}\left(\bar{X} - \bar{Y}\right)' \Sigma^{-1}\left(\bar{X} - \bar{Y}\right) = \left(\bar{X} - \bar{Y}\right)'\left(\frac{m}{n_1 n_2}\Sigma\right)^{-1}\left(\bar{X} - \bar{Y}\right)$

Since we know that $\bar{X} \sim N\left(\mu_1, \frac{1}{n_1}\Sigma\right)$ and $\bar{Y} \sim N\left(\mu_2, \frac{1}{n_2}\Sigma\right)$ are independent, then

$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\Sigma\right) = N\left(\Delta, \frac{m}{n_1 n_2}\Sigma\right)$ because it is a linear transformation of multinormal random variables. [6]

Using Theorem 10.12 from Schott [7]:

> *Let $x \sim N_m(\mu, \Omega)$, where Ω is a positive definite matrix and let A be an m x m symmetric matrix. If $A\Omega$ is idempotent and rank $(A\Omega)$ = r, Then $x'Ax \sim \chi^2_r(\lambda)$, where $\lambda = \mu' A \mu$.*

Checking those conditions, we have $x = \bar{X} - \bar{Y} \sim N\left(\Delta, \frac{m}{n_1 n_2}\Sigma\right)$ and $r = p$. *Let* $A = \left(\frac{m}{n_1 n_2}\Sigma\right)^{-1}$; then $A\Omega = I_p$ is idempotent with rank $(A\Omega) = p$ and $\Delta' A \Delta = \frac{n_1 n_2}{m}\left\| \Delta \right\|^2$; thus, $T \sim \chi_p^2\left(\frac{n_1 n_2}{m}\left\| \Delta \right\|^2\right)$. □

3.2.10 Stochastic Monotonicity of Distribution of *T*

A monotonic function is either entirely non-increasing or non-decreasing. If $P\left(T \le t \middle| \Delta_1\right) \ge P\left(T \le t \middle| \Delta_2\right)\ \forall t > 0$ when $\|\Delta_1\| < \|\Delta_2\|$, then the distribution of T is stochastically non-decreasing in $\|\Delta\|$.

Theorem 4: $\chi_p^2(\theta)$ **is stochastically non-decreasing in** θ

Proof: Let Z be a random vector with Z~$N(0, I_p)$ and let $f(z)$ be the pdf of Z.

Define X = $Z+\mu$. Then X~$N(\mu, I_p)$ and has pdf $f(x-\mu)$. Note that $X'X \sim \chi_p^2(\mu'\mu)$.

Define $E = \left\{v \in \mathbf{R}^p : v'v \le t\right\}$. This leads to the following equalities:

$$P\left(\chi_p^2(\mu'\mu) \le t\right) = P\left(X'X \le t\right) = P\left(X \in E\right) = \int_E f\left(x-\mu\right)dx$$

Theorem 1 in Anderson [11] states the following:

Let E be a convex set in n-space, symmetric about the origin. Let $f(x) \ge 0$ be a function such that:

(i) $f(x)=f(-x)$,

(ii) $\left\{x \middle| f(x) \ge u\right\} = K_u$ *is convex for every u* $(0 < u < \infty)$, *and*

(iii) $\int_E f(x)dx < \infty$ *(in the Lebesgue sense).*

Then $\int_E f\left(x+ky\right)dx \ge \int_E f\left(x+y\right)dx$ *for* $0 \le k \le 1$.

Since y is an arbitrary vector, the conclusion actually claims $\int_E f(x+ky)dx$ is a non-increasing function of $k \in [0, \infty)$.

Our E is a convex set, which is symmetric about the origin. The pdf of Z~$N(0, I_p)$, $f(z) = f(-z)$ with $\{z \mid f(z) \geq u\} = K_u$ being convex for every positive real number u. Since it is a pdf, the integral is finite in the Legesgue sense, so this theorem can be applied to $f(x)$ and E as defined above. Thus, $P\left(\chi_p^2\left(k^2\mu'\mu\right) \leq t\right) = \int_E f\left(x - k\mu\right)dx$ is a non-increasing function of k. Hence, $\chi_p^2\left(\theta\right)$ is stochastically non-decreasing in θ. □

3.2.11 Properties of the Test

Let $\beta(\Delta) = P\left(T < t \middle| \Delta\right)$, with t being the critical value associated with α and H_0. This function plays an important role in the study of the properties of the test because $\beta(\Delta)$ is the probability that we will reject H_0 given a value for Δ.

A Type I error is the probability of rejecting H_0 when it is actually true. The probability of a Type I error is $\beta(\Delta)$ when Δ is in Θ_0 .

A Type II error is the probability of failing to reject H_0 when it is actually false. The probability of a Type II error is $1-\beta(\Delta)$ under that restriction that Δ is not in Θ_0.

3.2.11.1 Power of the Test

The power of the test is the probability of rejecting H_0 given that H_0 is actually false, i.e., $P\left(T \geq t \mid \Delta \notin \Theta_0\right)$. So $\beta(\Delta)$ under the restriction that Δ is not in Θ_0 is the power function for this test.

3.2.11.2 Least-Favorable Points in H_0

The least favorable points in H_0 are those that maximize the probability of rejecting H_0, i.e., those points that maximize the probability of a Type I error. To find the least favorable points in Θ_0 , we find the maximum value of $\beta(\Delta)$ over all possible values of Δ in Θ_0. As shown in Theorem 4, when the true value of the difference, $\|\Delta\|$, increases, the probability of rejecting H_0 decreases. So $\beta(\Delta)$ is maximized at that lowest possible value of $\|\Delta\| \in \Theta_0$. This is δ by definition, that is

$$\max_{\Delta \in \Theta_0} \beta(\Delta) = \max_{\|\Delta\| \geq \delta} \beta(\Delta)$$

Thus, the set of least favorable points is $\{\Delta \in \Theta_0 : \|\Delta\| = \delta\}$.

3.2.11.3 Setting the Critical Value

If t = α^{th} percentile of a chi-squared distribution with p degrees of freedom and non-central parameter of $\theta = \frac{n_1 n_2}{m}\delta^2$, then $P(T < t) \leq \alpha$ for all Δ in Θ_0. Thus, the maximum probability of a Type I error is α.

3.2.11.4 Unbiasedness of the Test

An unbiased test has a higher probability of rejecting the null hypothesis when it is false than when it is true. This test is unbiased because if $\Delta_1 \notin \Theta_0$ and $\Delta_2 \in \Theta_0$, then $\|\Delta_1\| < \delta \leq \|\Delta_2\|$. By Theorem 4, this implies that $\beta(\Delta_1) \geq \beta(\Delta_2)$. Since the probability of rejecting H_0 when Δ lies in Θ_0 is smaller than the probability of rejecting H_0 when Δ does not lie in Θ_0, this test is unbiased for H_0.

3.3 Case 2

For Case 2, we tackle the situation where the two populations are assumed to have a common covariance of σΣ, with $\sigma > 0$. Case 1 can be considered a particular instance ($\sigma=1$) of the more general problem expressed in Case 2. The same measurement and definitions given in sections 3.1.1 and 3.1.2 will be used.

3.3.1 Sample Distributions

The formal description of these sample distributions is

$$\left.\begin{array}{l} X_1,\ldots,X_{n_1} \sim N_p\left(\mu_1,\sigma\Sigma\right) \\ Y_1,\ldots,Y_{n_2} \sim N_p\left(\mu_2,\sigma\Sigma\right) \end{array}\right\} \begin{array}{l} \mu_1,\mu_2,\sigma \text{ unknown, } \Sigma \text{ known} \\ \sigma>0, \mu_1,\mu_2 \in \mathbf{R}^p, \Sigma \in \mathbf{R}^{pxp} \end{array}$$

3.3.2 Joint Probability Density Function

The joint probability density function of the m independent random vectors from the samples is the product of their individual pdfs. For Case 2, this is

$$\prod_{i=1}^{n_1} \frac{1}{(2\pi)^{p/2}\left|\sigma\Sigma\right|^{1/2}} e^{-\frac{(x_i-\mu_1)'\Sigma^{-1}(x_i-\mu_1)}{2\sigma}} \cdot \prod_{j=1}^{n_2} \frac{1}{(2\pi)^{p/2}\left|\sigma\Sigma\right|^{1/2}} e^{-\frac{(y_j-\mu_2)'\Sigma^{-1}(y_j-\mu_2)}{2\sigma}}$$

3.3.3 Likelihood Function L(μ1, μ2,σ)

The likelihood function can be expressed with $A(\mu_1, \mu_2)$ and defined as in equation (2).

$$L(\mu_1,\mu_2,\sigma) = \frac{1}{(2\pi)^{mp/2}\left|\sigma\Sigma\right|^{m/2}} e^{-\frac{A(\mu_1,\mu_2)}{2\sigma}}$$

3.3.4 Maximized Likelihood Function *L(μ1, μ2,σ)* Without Restrictions

Lemma 5: The likelihood function is maximized over all $\sigma > 0$ at $\sigma = \frac{1}{m}\mathbf{A}(\mu_1,\mu_2)$

Proof: We can use the derivative test to establish a value for σ that maximizes $\boldsymbol{L}(\mu_1, \mu_2,\sigma)$.

Since this function and its natural log are maximized at the same values of σ, the technique of maximizing the natural log of the function is used.

$$\ln\left(L(\mu_1,\mu_2,\sigma)\right) = \ln\frac{1}{(2\pi)^{mp/2}\left|\Sigma\right|^{m/2}} - \frac{m}{2}\ln(\sigma) - \frac{\mathbf{A}(\mu_1,\mu_2)}{2\sigma}$$

Taking the derivative of the log with respect to σ yields

$$\frac{d}{d\sigma}\left\{\ln\left(L(\mu_1,\mu_2,\sigma)\right)\right\} = -\frac{m}{2\sigma} + \frac{\mathbf{A}(\mu_1,\mu_2)}{2\sigma^2}$$

Setting the derivative equal to zero and solving for σ yields

$$-\frac{m}{2\sigma}+\frac{\mathbf{A}(\mu_1,\mu_2)}{2\sigma^2}=0 \Rightarrow \sigma=\tfrac{1}{m}\mathbf{A}(\mu_1,\mu_2)$$

Since $-\frac{m}{2\sigma}+\frac{A(\mu_1,\mu_2)}{2\sigma^2}=-\frac{m}{2\sigma^2}\left(\sigma-\frac{A(\mu_1,\mu_2)}{m}\right) \Rightarrow \begin{cases} \frac{d}{d\sigma}\left\{\ln\left(L(\mu_1,\mu_2,\sigma)\right)\right\}<0 \text{ when } \sigma>\frac{A(\mu_1,\mu_2)}{m} \\ \frac{d}{d\sigma}\left\{\ln\left(L(\mu_1,\mu_2,\sigma)\right)\right\}=0 \text{ when } \sigma=\frac{A(\mu_1,\mu_2)}{m} \\ \frac{d}{d\sigma}\left\{\ln\left(L(\mu_1,\mu_2,\sigma)\right)\right\}>0 \text{ when } \sigma<\frac{A(\mu_1,\mu_2)}{m} \end{cases}$

This establishes that the maximum of the likelihood function will occur at $\sigma=\frac{A(\mu_1,\mu_2)}{m}$. □

Lemma 6: The unrestricted likelihood function is maximized over all possible values of $\boldsymbol{\mu_1}$, $\boldsymbol{\mu_2}$ and σ at $\mu_1=\bar{X}$, $\mu_2=\bar{Y}, \sigma=\tfrac{1}{m}\mathbf{A}\left(\bar{X},\bar{Y}\right)$.

Proof: As established in Lemma 5, $L(\mu_1,\mu_2,\sigma)\le L\left(\mu_1,\mu_2,\tfrac{1}{m}\mathbf{A}(\mu_1,\mu_2)\right)$

But $\mathbf{L}\left(\mu_1,\ \mu_2,\ \tfrac{1}{m}\mathbf{A}(\mu_1,\mu_2)\right) = \dfrac{1}{(2\pi)^{mp/2}\left|\dfrac{\mathbf{A}(\mu_1,\mu_2)}{m}\Sigma\right|^{m/2}} e^{-\frac{1}{2}\frac{\mathbf{A}(\mu_1,\mu_2)}{\left(\frac{\mathbf{A}(\mu_1,\mu_2)}{m}\right)}}$

$$=\frac{e^{-\frac{m}{2}}}{(2\pi)^{mp/2}|\Sigma|^{m/2}}\cdot\left(\frac{\mathbf{A}(\mu_1,\mu_2)}{m}\right)^{-\frac{m}{2}},$$

which is maximized when $\mathbf{A}(\mu_1,\mu_2)$ is minimized. As shown in section 3.2.4, if there are no restrictions on μ_1 and μ_2, $\mathbf{A}(\mu_1,\mu_2)$ is minimized at $\mu_1=\bar{X}$, $\mu_2=\bar{Y}$. Thus,

$$\max L(\mu_1,\mu_2,\sigma)=L\left(\bar{X},\bar{Y},\frac{A\left(\bar{X},\bar{Y}\right)}{m}\right)=\frac{e^{-\frac{m}{2}}}{(2\pi)^{mp/2}|\Sigma|^{m/2}}\cdot\left(\frac{A\left(\bar{X},\bar{Y}\right)}{m}\right)^{-\frac{m}{2}}. \qquad □$$

3.3.5 Maximized Likelihood Function L(μ_1, μ_2,σ) under Restriction that Δ is in Θ_0

Next, we need to maximize the likelihood function under the restriction of the null hypothesis. Define

$$\hat{\mu}_1 = \frac{n_1\bar{X} + n_2\bar{Y} + n_2 P_{\Theta_0}\left(\bar{X} - \bar{Y} \middle| \Delta \in \Theta_0\right)}{m},\ \hat{\mu}_2 = \frac{n_1\bar{X} + n_2\bar{Y} - n_1 P_{\Theta_0}\left(\bar{X} - \bar{Y} \middle| \Delta \in \Theta_0\right)}{m}$$

Lemma 7: The restricted likelihood function is maximized over all possible values of $\boldsymbol{\mu_1}$, $\boldsymbol{\mu_2}$ and σ at $\mu_1 = \hat{\mu}_1,\ \mu_2 = \hat{\mu}_2, \sigma = \frac{1}{m}\mathbf{A}\left(\hat{\mu}_1, \hat{\mu}_2\right)$.

Proof: Since $\boldsymbol{L}(\mu_1, \mu_2,\sigma) \le \boldsymbol{L}(\mu_1, \mu_2, A(\mu_1, \mu_2)/m)$, as shown in Lemma 5, and

$\mathbf{L}\left(\mu_1,\ \mu_2,\ \frac{1}{m}\mathbf{A}\left(\mu_1,\mu_2\right)\right) = \dfrac{e^{-\frac{m}{2}}}{(2\pi)^{mp/2}\left|\Sigma\right|^{m/2}} \cdot \left(\dfrac{\mathbf{A}\left(\mu_1,\mu_2\right)}{m}\right)^{-\frac{m}{2}}$, the problem of finding the maximum

under the restriction can be reduced to finding $\min\limits_{\Delta\in\Theta_0} A\left(\mu_1,\ \mu_2\right)$. Recall from section 3.2.4 that

this is accomplished by finding $\min\limits_{\Delta\in\Theta_0} g\left(\mu_1,\ \mu_2\right)$ with $g\left(\mu_1,\ \mu_2\right)$ defined as in equation (3). By

Lemma 4 this occurs at $\hat{\mu}_1$ and $\hat{\mu}_2$.

Thus, $\max\limits_{\Delta\in\Theta_0} L\left(\mu_1,\mu_2,\sigma\right) = \dfrac{e^{-\frac{m}{2}}}{(2\pi)^{mp/2}\left|\Sigma\right|^{m/2}}\left(\dfrac{A\left(\hat{\mu}_1,\hat{\mu}_2\right)}{m}\right)^{-\frac{m}{2}}$. □

3.3.6 Ratio of Maximized Likelihood Functions

As in Case 1, we define Λ as the ratio of the two maximum likelihood functions, with the function for the restricted domain in the denominator: $\Lambda = \dfrac{\max L\left(\mu_1,\mu_2,\sigma\right)}{\max\limits_{\Delta\in\Theta_0} L\left(\mu_1,\mu_2,\sigma\right)}$. Putting the expression found for the maximized likelihood function from Lemmas 6 and 7, we get

$$\Lambda = \frac{L\left(\bar{X},\bar{Y},\frac{A\left(\bar{X},\bar{Y}\right)}{m}\right)}{L\left(\hat{\mu}_1,\hat{\mu}_2,\frac{A\left(\hat{\mu}_1,\hat{\mu}_2\right)}{m}\right)} = \frac{\frac{1}{(2\pi)^{mp/2}|\Sigma|^{m/2}} e^{-\frac{m}{2}} \cdot \left(\frac{A\left(\bar{X},\bar{Y}\right)}{m}\right)^{-\frac{m}{2}}}{\frac{1}{(2\pi)^{mp/2}|\Sigma|^{m/2}} e^{-\frac{m}{2}} \cdot \left(\frac{A\left(\hat{\mu}_1,\hat{\mu}_2\right)}{m}\right)^{-\frac{m}{2}}} = \left(\frac{A\left(\hat{\mu}_1,\hat{\mu}_2\right)}{A\left(\bar{X},\bar{Y}\right)}\right)^{\frac{m}{2}}$$

Define $T = \dfrac{\left\|\bar{X}-\bar{Y}-P_{\Theta_0}\left(\bar{X}-\bar{Y}\,|\Delta\in\Theta_0\right)\right\|^2}{\sum_{i=1}^{n_1}\left\|X_i-\bar{X}\right\|^2+\sum_{j=1}^{n_2}\left\|Y_j-\bar{Y}\right\|^2}$.

Lemma 8: Λ is an increasing function of T

Proof: Λ is clearly an increasing function of

$$\frac{A\left(\hat{\mu}_1,\hat{\mu}_2\right)}{A\left(\bar{X},\bar{Y}\right)} = \frac{\sum_{i=1}^{n_1}\left\|X_i-\bar{X}\right\|^2+\sum_{j=1}^{n_2}\left\|Y_j-\bar{Y}\right\|^2+\frac{n_1 n_2}{m}\left\|\bar{X}-\bar{Y}-P_{\Theta_0}\left(\bar{X}-\bar{Y}\,|\Delta\in\Theta_0\right)\right\|^2}{\sum_{i=1}^{n_1}\left\|X_i-\bar{X}\right\|^2+\sum_{j=1}^{n_2}\left\|Y_j-\bar{Y}\right\|^2}$$

Then $\Lambda = \left(\frac{n_1 n_2}{m}T+1\right)^{\frac{m}{2}}$. Since $T \geq 0$, Λ is an increasing function of T. □

3.3.7 Likelihood Ratio Test (LRT) Statistic

Since Λ is an increasing function of T, T can serve as our test statistic. We will reject H_0 when T is sufficiently large. Substituting in the projection from section 3.2.5, we can express the test statistic as follows:

Case 2 Test Statistic: $$T = \frac{\left(\delta-\left\|\bar{X}-\bar{Y}\right\|\right)^2 \cdot I\left(\left\|\bar{X}-\bar{Y}\right\|<\delta\right)}{\sum_{i=1}^{n_1}\left\|X_i-\bar{X}\right\|^2+\sum_{j=1}^{n_2}\left\|Y_j-\bar{Y}\right\|^2}$$

where $I\left(\left\|\bar{X}-\bar{Y}\right\|<\delta\right)$ is an indicator function such that $I\left(\left\|\bar{X}-\bar{Y}\right\|<\delta\right) = \begin{cases} 1 & \text{if } \left\|\bar{X}-\bar{Y}\right\|<\delta \\ 0 & \text{if } \left\|\bar{X}-\bar{Y}\right\|\geq\delta \end{cases}$.

Some lemmas will be needed to establish the properties of this test statistic.

Define $T_N = \left[\left(\delta-\left\|\bar{X}-\bar{Y}\right\|\right)^2 \cdot I\left(\left\|\bar{X}-\bar{Y}\right\|<\delta\right)\right]$ and $T_D = \sum_{i=1}^{n_1}\left\|X_i-\bar{X}\right\|^2+\sum_{j=1}^{n_2}\left\|Y_j-\bar{Y}\right\|^2$.

Lemma 9: Distributions of T_N and T_D are independent

Proof: T_N is a function of $\bar{X}-\bar{Y}$ only.

T_D is a function of $\left[X_1-\bar{X},\cdots,X_{n_1}-\bar{X},Y_1-\bar{Y},\cdots,Y_{n_2}-\bar{Y}\right]$.

If $\bar{X}-\bar{Y}$ is independent of $\left[X_1-\bar{X},\cdots,X_{n_1}-\bar{X},Y_1-\bar{Y},\cdots,Y_{n_2}-\bar{Y}\right]$, then the distributions of T_N and T_D are independent.

Let D be the matrix of data values for the two samples.

$$\mathbf{D}=\left[X_1,\cdots,X_{n_1},Y_1,\cdots,Y_{n_2}\right] \text{ then } \mathbf{D}\sim N_{p,m}\left(\mathbf{M},\sigma\boldsymbol{\Sigma}\right)$$

$$\text{with } \mathbf{M}_{pxm}=\left(\mu_1,\mu_2\right)\begin{bmatrix}\mathbf{1}'_{n_1} & 0\\ 0 & \mathbf{1}'_{n_2}\end{bmatrix}$$

where $\mathbf{D}\sim N_{p,m}\left(\mathbf{M},\sigma\boldsymbol{\Sigma}\right)$ is a notation that indicates **D** is a p x m random matrix, E(**D**) = **M**, and the columns of **D** are independent normal vectors with a common covariance matrix σ**Σ**. This is the notation used by Hu [12].

Define **A** and **B** as follows: $\mathbf{A}=\begin{bmatrix}\dfrac{\mathbf{1}_{n_1}}{n_1}\\ -\dfrac{\mathbf{1}_{n_2}}{n_2}\end{bmatrix}_{mx1}$ $\quad \mathbf{B}=I_m-\begin{bmatrix}\dfrac{\mathbf{1}_{n_1}\mathbf{1}'_{n_1}}{n_1} & \mathbf{0}\\ \mathbf{0} & \dfrac{\mathbf{1}_{n_2}\mathbf{1}'_{n_2}}{n_2}\end{bmatrix}_{mxm}$

$$\text{Then } \mathbf{DA}=\left[X_1,\cdots,X_{n_1},Y_1,\cdots,Y_{n_2}\right]\begin{bmatrix}\dfrac{\mathbf{1}_{n_1}}{n_1}\\ -\dfrac{\mathbf{1}_{n_2}}{n_2}\end{bmatrix}= \quad \bar{X}-\bar{Y},$$

$$\mathbf{DB}_{pxm}=\left[X_1,\cdots,X_{n_1},Y_1,\cdots,Y_{n_2}\right]\left(I_m-\begin{bmatrix}\dfrac{\mathbf{1}_{n_1}\mathbf{1}'_{n_1}}{n_1} & \mathbf{0}\\ \mathbf{0} & \dfrac{\mathbf{1}_{n_2}\mathbf{1}'_{n_2}}{n_2}\end{bmatrix}\right)$$

$$=\left[X_1,\cdots,X_{n_1},Y_1,\cdots,Y_{n_2}\right]-\left[X_1,\cdots,X_{n_1},Y_1,\cdots,Y_{n_2}\right]\begin{bmatrix}\frac{\mathbf{1}_{n_1}\mathbf{1}'_{n_1}}{n_1} & \mathbf{0}\\ \mathbf{0} & \frac{\mathbf{1}_{n_2}\mathbf{1}'_{n_2}}{n_2}\end{bmatrix}$$

$$=\left[X_1-\bar{X},\cdots,X_{n_1}-\bar{X},Y_1-\bar{Y},\cdots,Y_{n_2}-\bar{Y}\right].$$

Clearly, $\mathbf{A}'\mathbf{B}=\left[\frac{\mathbf{1}'_{n_1}}{n_1},-\frac{\mathbf{1}'_{n_2}}{n_2}\right]\left(I_m-\begin{bmatrix}\frac{\mathbf{1}_{n_1}\mathbf{1}'_{n_1}}{n_1} & \mathbf{0}\\ \mathbf{0} & \frac{\mathbf{1}_{n_2}\mathbf{1}'_{n_2}}{n_2}\end{bmatrix}\right)=\left[\frac{\mathbf{1}'_{n_1}}{n_1},-\frac{\mathbf{1}'_{n_2}}{n_2}\right]-\left[\frac{\mathbf{1}'_{n_1}}{n_1},-\frac{\mathbf{1}'_{n_2}}{n_2}\right]=\mathbf{0}.$

By (b) of Lemma 2 according to Hu [12], which states the following:

Suppose $\mathbf{Y} \sim N_{pxn}(\mathbf{M},\boldsymbol{\Sigma})$ *If* $\mathbf{A}'\mathbf{B} = \mathbf{0}$*, where* **A** *has n rows, then* **YA** *and* **YB** *are independent.*

With the data matrix **D** taking the part of **Y** in the lemma, all criterion are met for **A** and **B**; thus, $\bar{X}-\bar{Y}$ and $\left[X_1-\bar{X},\cdots,X_{n_1}-\bar{X},Y_1-\bar{Y},\cdots,Y_{n_2}-\bar{Y}\right]$ are independent. Therefore, T_N and T_D are independent. □

Lemma 10: Distribution of $T_D=\sum_{i=1}^{n_1}\|X_i-\bar{X}\|^2+\sum_{j=1}^{n_2}\|Y_j-\bar{Y}\|^2$ **is free of** $\boldsymbol{\mu}_1$ **and** $\boldsymbol{\mu}_2$

Proof: It is helpful to note the following:

$$\begin{pmatrix}1'_{n_1} & 0\\ 0 & 1'_{n_2}\end{pmatrix}\mathbf{B}=\begin{pmatrix}1'_{n_1} & 0\\ 0 & 1'_{n_2}\end{pmatrix}\left(I_m-\begin{bmatrix}\frac{\mathbf{1}_{n_1}\mathbf{1}'_{n_1}}{n_1} & \mathbf{0}\\ \mathbf{0} & \frac{\mathbf{1}_{n_2}\mathbf{1}'_{n_2}}{n_2}\end{bmatrix}\right)=\begin{pmatrix}1'_{n_1} & 0\\ 0 & 1'_{n_2}\end{pmatrix}-\begin{pmatrix}1'_{n_1} & 0\\ 0 & 1'_{n_2}\end{pmatrix}=\mathbf{0}$$

By part (d) of Hu's Lemma 2, that rank (**B**) = Tr(**B**) = $n_1\frac{n_1-1}{n_1}+n_2\frac{n_2-1}{n_2}=m-2$.

Since **B** is idempotent, by Theorem 4.5 of Schott [7], there exists a matrix **P** (not unique) where $\mathbf{P}\in\mathbf{R}^{mx(m-2)}$ and rank(**P**) **=** $m-2$ s with $\mathbf{P}'\mathbf{P} = \mathbf{I}_{m-2}$ such that $\mathbf{B} = \mathbf{P}\mathbf{P}'$.

Pre-multiplying B by $\begin{pmatrix} 1'_{n_1} & 0 \\ 0 & 1'_{n_1} \end{pmatrix}$ and post-multiplying by **P** yields

$$\begin{aligned}
\begin{pmatrix} 1'_{n_1} & 0 \\ 0 & 1'_{n_2} \end{pmatrix}\mathbf{BP} &= \begin{pmatrix} 1'_{n_1} & 0 \\ 0 & 1'_{n_2} \end{pmatrix}\mathbf{PP'P} \\
\mathbf{0}\cdot\mathbf{P} &= \begin{pmatrix} 1'_{n_1} & 0 \\ 0 & 1'_{n_2} \end{pmatrix}\mathbf{PI}_{m-2} \\
\Rightarrow \begin{pmatrix} 1'_{n_1} & 0 \\ 0 & 1'_{n_1} \end{pmatrix}\mathbf{P} &= \mathbf{0}
\end{aligned}$$

Invoking part (a) of Lemma 2 from Hu [12], which states the following:

For an n x m matrix P with P′P = I, YP~N_{pxm}(MP, Σ).

Thus, $\mathbf{DP} \sim N\left((\mu_1, \mu_2)\begin{bmatrix} \mathbf{1}'_{n_1} & 0 \\ 0 & \mathbf{1}'_{n_2} \end{bmatrix}\mathbf{P}, \sigma\Sigma \right) = N_{px(m-2)}\left(0, \sigma\Sigma\right)$ is a distribution free of μ_1 and μ_2. Since T_D is a function of **DB**, which is a function of **DP**, the distribution of T_D is free of the parameters μ_1 and μ_2. □

Lemma 11: Distribution of numerator of T depends on μ_1 and μ_2 only through $\|\Delta\|$

Proof: Recall that: $T_N = \left[\left(\delta - \left\| \bar{X} - \bar{Y} \right\| \right)^2 \cdot I\left(\left\| \bar{X} - \bar{Y} \right\| < \delta \right) \right]$

and that by Theorem 3, above, the following product has a non-central chi-squared distribution:

$$\begin{aligned}
&\left(\frac{m\sigma}{n_1 n_2} \right)^{-1} \left\| \bar{X} - \bar{Y} \right\|^2 = \left(\bar{X} - \bar{Y} \right)' \left(\frac{m\sigma}{n_1 n_2}\mathbf{\Sigma} \right)^{-1} \left(\bar{X} - \bar{Y} \right) \sim \chi_p^2\left(\left(\frac{n_1 n_2}{m\sigma} \|\Delta\|^2 \right) \right) \\
&\Rightarrow \left\| \bar{X} - \bar{Y} \right\|^2 \sim \frac{m\sigma}{n_1 n_2} \chi_p^2\left(\left(\frac{n_1 n_2}{m\sigma} \|\Delta\|^2 \right) \right)
\end{aligned}$$

Since the pdf of $\left\| \bar{X} - \bar{Y} \right\|^2$ depends on μ_1 and μ_2 only through $\|\Delta\|$, the pdf of $\left\| \bar{X} - \bar{Y} \right\|$ also depends on μ_1 and μ_2 only through $\|\Delta\|$. Since T_N is a function of $\left\| \bar{X} - \bar{Y} \right\|$, this establishes that the distribution of T_N is dependent on μ_1 and μ_2 only through $\|\Delta\|$. □

Theorem 5: PDF of T depends on μ₁ and μ₂ only through $\|\Delta\|$

Proof: By Lemma 11, we can denote the pdf of T_N by $f\left(t_n, \|\Delta\|\right)$. By Lemma 10, we can denote the pdf of T_D by $g(t_n)$. By their independence, established in Lemma 9, the joint pdf of T_N and T_D is $f\left(t_n, \|\Delta\|\right) \cdot g(t_n)$, which depends on μ_1 and μ_2 only through $\|\Delta\|$. But $T = \dfrac{T_N}{T_D}$ is a function of T_D and T_N. Therefore, the distribution of T depends on on μ_1 and μ_2 only through $\|\Delta\|$.

□

3.3.8 Stochastic Monotonicity of Distribution of *T*

Some lemmas are needed before we can establish that the distribution of T is monotonically non-decreasing in $\|\Delta\|$.

Define $\gamma\left(\Delta, R\right) = P\left(\left(\delta - \left\|\bar{X} - \bar{Y}\right\|\right)^2 \cdot I\left(\left\|\bar{X} - \bar{Y}\right\| < \delta\right) > tR\right)$.

Lemma 12: $\gamma\left(\Delta_1, R\right) \geq \gamma\left(\Delta_2, R\right)\ \forall R > 0$ **when** $\|\Delta_1\| < \|\Delta_2\|$

Proof: Let $D = \left\{x \in \mathbf{R}^p : \|x\| \leq \delta - \sqrt{tR}\ \right\}$ and let $g\left(\|x - \Delta\|^2\right)$ be the multinormal pdf of $\bar{X} - \bar{Y} \sim N_p\left(\Delta, \dfrac{m\sigma}{n_1 n_2}\Sigma\right)$. Then

$$\gamma\left(\Delta, R\right) = P\left(\left(\delta - \left\|\bar{X} - \bar{Y}\right\|\right)^2 \cdot I\left(\left\|\bar{X} - \bar{Y}\right\| < \delta\right) > tR\right)$$

$$= P\left(\left(\delta - \left\|\bar{X} - \bar{Y}\right\|\right)^2 > tR \text{ and } \mathrm{P}\left(\left\|\bar{X} - \bar{Y}\right\| < \delta\right)\right)$$

$$= P\left(\left\|\bar{X} - \bar{Y}\right\| < \delta - \sqrt{tR} \text{ and } \mathrm{P}\left(\left\|\bar{X} - \bar{Y}\right\| < \delta\right)\right)$$

$$= P\left(\left\|\bar{X} - \bar{Y}\right\| < \delta - \sqrt{tR}\right)$$

$$= P\left(\bar{X} - \bar{Y} \in D\right)$$

$$= P\left(\bar{X} - \bar{Y} \in D\right)$$

$$= \int_D g\left(\left\|x - \Delta\right\|^2\right) dx$$

By Theorem 1 from Anderson [11], since D is a convex set symmetric about the origin and $g\left(\left\|v\right\|^2\right)$ is symmetric about the origin with $\left\{v \in R^p : g\left(\left\|v\right\|^2\right) \geq u\right\} = K_u$ being convex for every positive real number *u* and its integral always non-negative and finite in the Lebesgue sense, $\gamma\left(k\Delta, R\right)$ is a non-increasing function of $k \in [0, \infty)$ for all $\Delta \in \mathbf{R}^p$. By Lemma 11, $\gamma\left(k\Delta, R\right)$ depends on $k\Delta$ only through $\left\|k\Delta\right\| = k\left\|\Delta\right\|$. So we conclude that $\gamma\left(\Delta, R\right)$ is a non-increasing function of $\left\|\Delta\right\|$, i.e. $\left\|\Delta_1\right\| < \left\|\Delta_2\right\| \Rightarrow \gamma\left(\Delta_1, R\right) \geq \gamma\left(\Delta_2, R\right)\ \forall R > 0$ □

Theorem 6: T is stochastically non-increasing with respect to $\left\|\Delta\right\|$, i.e.

$$P\left(T > t \middle| \Delta_1\right) \geq P\left(T > t \middle| \Delta_2\right)\ \forall t > 0 \text{ when } \left\|\Delta_1\right\| < \left\|\Delta_2\right\|.$$

Proof: Note that

$$P\left(T > t\right) = P\left[\left(\delta - \left\|\bar{X} - \bar{Y}\right\|\right)^2 \cdot I\left(\left\|\bar{X} - \bar{Y}\right\| < \delta\right) > t\left(\sum_{i=1}^{n_1}\left\|X_i - \bar{X}\right\|^2 + \sum_{j=1}^{n_2}\left\|Y_j - \bar{Y}\right\|^2\right)\right]$$

$$= E\left[P\left(\left(\delta - \left\|\bar{X} - \bar{Y}\right\|\right)^2 \cdot I\left(\left\|\bar{X} - \bar{Y}\right\| < \delta\right) > t\left(\sum_{i=1}^{n_1}\left\|X_i - \bar{X}\right\|^2 + \sum_{j=1}^{n_2}\left\|Y_j - \bar{Y}\right\|^2\right) \middle| \sum_{i=1}^{n_1}\left\|X_i - \bar{X}\right\|^2 + \sum_{j=1}^{n_2}\left\|Y_j - \bar{Y}\right\|^2\right)\right]$$

By Lemma 9, which establishes the independence of $\bar{X} - \bar{Y}$ and $\sum_{i=1}^{n_1}\left\|X_i - \bar{X}\right\|^2 + \sum_{j=1}^{n_2}\left\|Y_j - \bar{Y}\right\|^2$, and the definition of $\gamma\left(\Delta, R\right)$,

$$P\left[\left(\delta - \left\|\bar{X} - \bar{Y}\right\|\right)^2 \cdot I\left(\left\|\bar{X} - \bar{Y}\right\| < \delta\right) > t\left(\sum_{i=1}^{n_1}\left\|X_i - \bar{X}\right\|^2 + \sum_{j=1}^{n_2}\left\|Y_j - \bar{Y}\right\|^2\right) \middle| \sum_{i=1}^{n_1}\left\|X_i - \bar{X}\right\|^2 + \sum_{j=1}^{n_2}\left\|Y_j - \bar{Y}\right\|^2\right]$$

$$= \gamma\left(\Delta, \sum_{i=1}^{n_1}\left\|X_i - \bar{X}\right\|^2 + \sum_{j=1}^{n_2}\left\|Y_j - \bar{Y}\right\|^2\right).$$

So $P\left(T > t\right) = E\left[\gamma\left(\Delta, \sum_{i=1}^{n_1}\left\|X_i - \bar{X}\right\|^2 + \sum_{j=1}^{n_2}\left\|Y_j - \bar{Y}\right\|^2\right)\right].$

By Lemma 12, $\gamma\left(\Delta, \sum_{i=1}^{n_1}\left\|X_i - \bar{X}\right\|^2 + \sum_{j=1}^{n_2}\left\|Y_j - \bar{Y}\right\|^2\right)$ is a non-increasing function of $\|\Delta\|$ with probability 1. So $P\left(T > t\right)$ is a non-increasing function of $\|\Delta\|$. □

3.3.9 Properties of the Test

We will reject H_0 when T is greater than the critical value. Let $\beta(\Delta) = P\left(T > t \middle| \Delta\right)$, with t being the critical value of the test for a given α. Thus, $\beta(\Delta)$ is the probability of rejecting H_0. When the null hypothesis is true, $\beta(\Delta)$ gives the probability of Type I error. When the null hypothesis is false, $1-\beta(\Delta)$ gives the probability of a Type II error .

3.3.9.1 Least-Favorable Points in H_0

The significant level of the test is the maximum probability that H_0 is rejected when H_0 is actually true. This level is denoted by α. When $\{T > t\}$ being the region of the rejection of H_0,

$$\alpha = \max\left[P\left(T > t\right) : \Delta \in \Theta_0\right]$$

By the definition of $\beta\left(\Delta\right)$,

$$\alpha = \max\left[\beta\left(\Delta\right) : \|\Delta\| \geq \delta\right].$$

By Theorem 6, $\beta\left(\Delta\right)$ is a non-increasing function of $\|\Delta\|$. So

$$\max\left[\beta\left(\Delta\right) : \|\Delta\| \geq \delta\right] = \beta\left(\Delta^*\right) \text{ where } \|\Delta^*\| = \delta$$

This point, Δ^* is called a least-favorable point in H_0. Clearly, $\left\{\Delta \in \Theta_0 : \|\Delta\| = \delta\right\}$ gives the collection of least-favorable points in H_0.

3.3.9.2 Unbiasedness of Test

If $\beta(\Delta)$,then the probability of rejecting the null hypothesis is always larger for $\Delta \notin \Theta_0$ than for $\Delta \in \Theta_0$, and the test is unbiased. Let $\Delta_1 \notin \Theta_0 \Rightarrow \|\Delta_1\| < \delta$ and $\Delta_2 \in \Theta_0 \Rightarrow \|\Delta_2\| \geq \delta$

Then by Theorem 6, $\beta(\Delta_1) = P(T > t|\Delta_1) \geq P(T > t|\Delta_2) = \beta(\Delta_2)$.

Thus. this test is unbiased.

3.3.9.3 Evaluation of Power Function

As we saw earlier, many properties of the test were obtained through the established monotonicity of the rejection probability function $\beta(\Delta)$ with respect to $\|\Delta\|$. The expression

$$\beta(\Delta) = P(T > t) = P\left(\frac{T_N}{T_D} > t\right) = P(T_N > T_D \cdot t),$$

however, does not have a closed form.

Let $D^* = \left\{\begin{pmatrix} t_n \\ t_d \end{pmatrix} : t_n > t_d \cdot t\right\}$ and $f(t_n, t_d)$ be the joint probability density function of $(T_N, T_D)'$. Then

$$\beta(\Delta) = \iint_{D^*} f(t_n, t_d)\, dt_n dt_d.$$

From the proof of Theorem 5, $f(t_n, t_d) = f(t_n, \|\Delta\|) \cdot g(t_d)\, dt_n dt_d$ where $f(t_n, \|\Delta\|)$ is the probability density function of T_N , and $g(t_d)$ is the probability density function of T_D. With given Δ, both $f(t_n, \|\Delta\|)$ and $g(t_d)$ can be numerically determined through χ^2- distributions. Therefore,

$$\beta(\Delta) = \iint_{D^*} f(t_n, \|\Delta\|) \cdot g(t_d)\, dt_n dt_d$$

can be computed by numerical integration method. In such a computation, Δ with different norm values can be selected on a ray staring at the origina in any convenient direction.

3.3.10 Setting the Critical Value

While the Case 2 test statistic does not fit a known distribution, its distribution can be simulated using the distributions of $\left\|\bar{X}-\bar{Y}\right\|^2$ and $\sum_{i=1}^{n_1}\left\|X_i-\bar{X}\right\|^2+\sum_{j=1}^{n_2}\left\|Y_j-\bar{Y}\right\|^2$. The critical values can be found via such a simulation using the least favorable value for Δ – i.e. $\left\{\Delta\in\Theta_0:\left\|\Delta\right\|=\delta\right\}$. The distribution of $\left\|\bar{X}-\bar{Y}\right\|^2$ was established in Theorem 3. Before we can establish the distribution of T_D, another lemma is needed.

Lemma 13: If $\mathbf{X}=\left(X_1,\cdots,X_n\right)\sim N\left(\mu\cdot 1_n',\sigma\Sigma\right)$, then $\sum_{i=1}^{n}\left\|X_i-\bar{X}\right\|^2\sim\sigma\chi^2_{np-p}$

Proof: $\mathbf{X}=\left(X_1,\cdots,X_n\right)\sim N\left(\mu\cdot 1_n',\sigma\Sigma\right)$ implies $Vec\left(\mathbf{X}\right)\sim N\left(1_n\otimes\mu,\Omega\right)$ where $\Omega=I_n\otimes\sigma\Sigma$

Since $\left(X_1-\bar{X},\ldots,X_n-\bar{X}\right)=\mathbf{X}\left(I_n-\frac{1_n1_n'}{n}\right)=I_p\mathbf{X}\left(I_n-\frac{1_n1_n'}{n}\right)$,

$$Vec\left(X_1-\bar{X},\ldots,X_n-\bar{X}\right)=Vec\left(I_p\mathbf{X}\left(I_n-\frac{1_n1_n'}{n}\right)\right)=\left[\left(I_n-\frac{1_n1_n'}{n}\right)\otimes I_p\right]Vec(\mathbf{X}).$$

This implies that $\frac{1}{\sigma}\sum_{i=1}^{n}\left\|X_i-\bar{X}\right\|^2=\frac{1}{\sigma}\sum_{i=1}^{n}(X_i-\bar{X})'\Sigma^{-1}(X_i-\bar{X})$

$$=\sum_{i=1}^{n}(X_i-\bar{X})'\left(\sigma\Sigma\right)^{-1}(X_i-\bar{X})$$

$$=Vec\left(X_1-\bar{X},\ldots,X_n-\bar{X}\right)'\left[I_n\otimes\left(\sigma\Sigma\right)^{-1}\right]Vec\left(X_1-\bar{X},\ldots,X_n-\bar{X}\right)$$

$$=\left[Vec\left(\mathbf{X}\right)\right]'\left[\left(I_n-\frac{1_n1_n'}{n}\right)\otimes I_p\right]'\left[I_n\otimes\left(\sigma\Sigma\right)^{-1}\right]\left[\left(I_n-\frac{1_n1_n'}{n}\right)\otimes I_p\right]\left[Vec\left(\mathbf{X}\right)\right]$$

$$= \left[Vec(\mathbf{X})\right]' \left[\left(I_n - \frac{1_n 1_n'}{n}\right) \otimes (\sigma\Sigma)^{-1}\right] \left[Vec(\mathbf{X})\right]$$

$$= \left[Vec(\mathbf{X})\right]' A \left[Vec(\mathbf{X})\right] \text{ with } A = \left[\left(I_n - \frac{1_n 1_n'}{n}\right) \otimes (\sigma\Sigma)^{-1}\right].$$

Once again applying Theorem 10.12 from Schott [7], given in section 3.2.9 , we find that since $Vec(\mathbf{X}) \sim N(1_n \otimes \mu, \Omega)$ and

$$A'\Omega A = \left[\left(I_n - \frac{1_n 1_n'}{n}\right) \otimes (\sigma\Sigma)^{-1}\right] \left[I_n \otimes \sigma\Sigma\right] \left[\left(I_n - \frac{1_n 1_n'}{n}\right) \otimes (\sigma\Sigma)^{-1}\right]$$

$$= \left[\left(I_n - \frac{1_n 1_n'}{n}\right) I_n \left(I_n - \frac{1_n 1_n'}{n}\right)\right] \otimes \left[I_n (\sigma\Sigma)^{-1} I_n\right] = \left(I_n - \frac{1_n 1_n'}{n}\right) \otimes (\sigma\Sigma)^{-1} = A$$

we can conclude that $\frac{1}{\sigma}\sum_{i=1}^{n} \left\| X_i - \bar{X} \right\|^2 \sim \chi^2(\lambda)$ with

$$\lambda = [1_n \otimes \mu]' \left[\left(I_n - \frac{1_n 1_n'}{n}\right) \otimes (\sigma\Sigma)^{-1}\right] [1_n \otimes \mu]$$

$$= \left[1_n \left(I_n - \frac{1_n 1_n'}{n}\right) 1_n\right]' \otimes \left[\mu' (\sigma\Sigma)^{-1} \mu\right] = 0 \otimes \left[\mu' (\sigma\Sigma)^{-1} \mu\right] = 0,$$

and degrees of freedom equal to

$$TR(A\Omega) = TR\left(\left[\left(I_n - \frac{1_n 1_n'}{n}\right) \otimes (\sigma\Sigma)^{-1}\right] \left[I_n \otimes \sigma\Sigma\right]\right)$$

$$= TR\left(\left(I_n - \frac{1_n 1_n'}{n}\right) \otimes I_p\right) = np\left(\frac{n-1}{n}\right) = p(n-1).$$

Thus, $\frac{1}{\sigma}\sum_{i=1}^{n} \left\| X_i - \bar{X} \right\|^2 \sim \chi^2_{np-p} \Rightarrow \sum_{i=1}^{n} \left\| X_i - \bar{X} \right\|^2 \sim \sigma\chi^2_{np-p}$. □

Theorem 7: $T_D \sim \sigma\chi^2_{n_1 p - p} + \sigma\chi^2_{n_2 p - p} = \sigma\chi^2_{mp-2p}$

Proof: By Lemma 13, $\sum_{i=1}^{n_1} \left\| X_i - \overline{X} \right\|^2 \sim \sigma\chi^2_{n_1 p-p}$ and $\sum_{j=1}^{n_2} \left\| Y_j - \overline{Y} \right\|^2 \sim \sigma\chi^2_{n_2 p-p}$. These are independent central chi-squared distributions.

The sum of independent central chi-squared distributions is well known to be a central chi-squared distribution with degrees of freedom equal to the sum of the degrees of freedom of the distributions being added. The conclusion follows immediately.

3.3.11 Simulation of Case 2 Test Statistic Distribution

A random variable with the distribution of T can be simulated by substituting in randomly generated values from the appropriate distributions for $\left\| \overline{X} - \overline{Y} \right\|^2$ and $\sum_{i=1}^{n_1} \left\| X_i - \overline{X} \right\|^2 + \sum_{j=1}^{n_2} \left\| Y_j - \overline{Y} \right\|^2$. The simulation can be computed specifying values for n_1, n_2, p, α, and δ.

Let U be a randomly generated value from a $\chi^2_p\left(\left(\frac{n_1 n_2 \varepsilon^2}{m}\right)\right)$ distribution, and let W be a randomly generated value from a χ^2_{mp-2p} distribution. If δ is defined as a multiple of the square root of σ, i.e., $\delta = \varepsilon\sqrt{\sigma}$, we can simulate values for T as follows:

$$T = \begin{cases} 0 & \sqrt{\dfrac{m\sigma}{n_1 n_2}U} \geq \varepsilon\sqrt{\sigma} \\ \dfrac{1}{W}\left(\varepsilon\sqrt{\sigma} - \sqrt{\dfrac{m\sigma}{n_1 n_2}U}\right)^2 & \sqrt{\dfrac{m\sigma}{n_1 n_2}U} < \varepsilon\sqrt{\sigma} \end{cases}$$

This simplifies to

$$T = \begin{cases} 0 & \sqrt{U} \geq \varepsilon\sqrt{\dfrac{n_1 n_2}{m}} \\ \dfrac{1}{W}\left(\varepsilon - \sqrt{\dfrac{m}{n_1 n_2}U}\right)^2 & \sqrt{U} < \varepsilon\sqrt{\dfrac{n_1 n_2}{m}} \end{cases}$$

Because this formulation eliminates σ from the non-centrality parameter, critical values for T will be stable for a given ε, whereas they will vary for a given δ. It makes more sense to provide tables for values of ε rather than δ.

Table 2 in Appendix A shows the results of the simulation for critical values with $n_1 = 6$; $n_2 = 2$; $\alpha=0.05$, p = 3, 4, 5, and 6; and $\varepsilon = 0.1$ to 5.0. One million test statistics were randomly generated for each set of parameters. The 95th percentile of those 1,000,000 test statistics was computed to determine the critical value. This was done twice in order to verify the accuracy of the resulting statistics. The results were consistent to the first three decimal places.

The simulation results are shown graphically in Figure 4. The SAS code used to generate those values is provided in Appendix B. The null hypothesis is rejected and the new material considered equivalent when the test statistic is greater than the critical value.

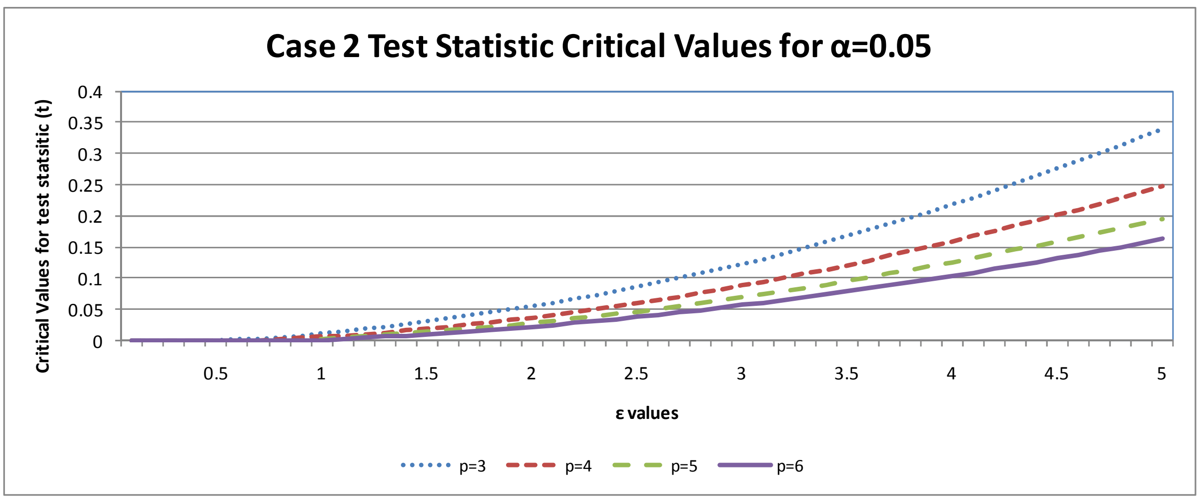

Figure 4. Critical values of Case 2 test statistic with $\alpha = 5\%$, $n_1 = 6$ and $n_2 = 2$.

CHAPTER 4

EXAMPLE APPLICATION

4.1 Example Data

To demonstrate the use of this approach, we will use NCAMP test results. Table 2 [13] show strength and modulus results for a fiberglass epoxy composite material. The qualification sample had two panels, cured separately, from each of three different batches of material, for a total of six panels. In addition, nine companies produced smaller equivalency samples of this material. Each equivalency sample consisted of two panels, cured separately and usually from a single batch. Each row in Table 2 gives the results for a single panel of material. Multiple test specimens were cut from each panel. Each value is the mean of multiple destructive tests from a minimum of three. Since all the values are means, the panel mean vectors have a multinormal distribution centered on the true mean vector.

TABLE 2

GLASS 6781 FILL TENSION PANEL DATA

Company Code	Panel ID	Mean Vectors					
		Strength CTD (ksi)	**Modulus CTD (msi)**	**Strength RTD (ksi)**	**Modulus RTD (msi)**	**Strength ETW (ksi)**	**Modulus ETW (msi)**
A0	A1	86.8256	4.1342	81.9492	4.0615	58.3693	3.7252
A0	A2	90.5834	4.1229	83.6153	4.0303	59.3905	3.7333
A0	B1	93.5993	4.1052	82.5799	4.0336	57.9774	3.7246
A0	B2	92.6419	4.1026	83.2296	4.0310	56.9364	3.7374
A0	C1	91.6543	4.1567	77.3917	4.0975	52.6690	3.8361
A0	C2	86.1103	4.2412	73.7584	4.1754	57.6251	3.8522
A1	H8	104.7720	4.3109	93.1665	4.2907	66.2259	3.8234
A1	H9	102.3422	4.3259	93.6730	4.2642	66.2396	3.8532
A2	G8	100.7421	4.2101	92.4146	4.1596	64.1142	3.7933
A2	G9	102.5788	4.1789	91.0630	4.1324	63.7832	3.7749
A3	E1	91.4111	4.1400	84.6080	4.0335	58.1473	3.7413
A3	E2	95.4070	4.1375	84.7847	4.0411	58.8351	3.7159

Company Code	Panel ID	Mean Vectors					
		Strength CTD (ksi)	**Modulus CTD (msi)**	**Strength RTD (ksi)**	**Modulus RTD (msi)**	**Strength ETW (ksi)**	**Modulus ETW (msi)**
A4	F1	101.5407	4.1214	91.3569	4.0616	60.1149	3.6535
A4	F2	103.3725	4.1850	89.9774	4.0893	60.8018	3.8006
A5	F1	93.7389	4.1293	85.9663	4.0944	57.3603	3.7727
A5	F2	89.7994	4.1654	81.7455	4.0942	58.7655	3.7617
A6	E1	102.0095	4.2496	94.9075	4.2011	63.5386	3.9018
A6	E2	101.0647	4.2748	97.5123	4.1656	66.6956	3.8939
A7	F5	100.7955	4.1914	87.5858	4.1444	61.7685	3.8369
A7	F6	99.3658	4.2207	88.5357	4.1221	63.5465	3.8330
A8	F7	101.4659	4.1396	90.5325	4.0954	62.8581	3.7360
A8	F8	99.5110	4.1478	90.0073	4.1408	62.0124	3.7273
A9	E5	99.6023	4.3338	88.3391	4.2965	62.6129	3.9849
A9	E6	99.5025	4.2903	89.5656	4.2779	61.7814	3.9429

Table 3 shows the vectors of company means. The qualification sample (A0) is $\overline{X}$, while each equivalency company below is $\overline{Y}$. The panel is the experimental unit. The sample sizes are six for the qualification sample and two for each of the equivalency samples. Six different test results are listed for each panel.

TABLE 3

FILL TENSION MEAN VECTORS

Company Code	Mean Vectors					
	Strength CTD (ksi)	**Modulus CTD (msi)**	**Strength RTD (ksi)**	**Modulus RTD (msi)**	**Strength ETW (ksi)**	**Modulus ETW (msi)**
A0	90.236	4.144	80.421	4.072	57.161	3.768
A1	103.557	4.318	93.420	4.277	66.233	3.838
A2	101.660	4.195	91.739	4.146	63.949	3.784
A3	93.409	4.139	84.696	4.037	58.491	3.729
A4	102.457	4.153	90.667	4.075	60.458	3.727
A5	91.769	4.147	83.856	4.094	58.063	3.767
A6	101.537	4.262	96.210	4.183	65.117	3.898
A7	100.081	4.206	88.061	4.133	62.658	3.835
A8	100.488	4.144	90.270	4.118	62.435	3.732
A9	99.552	4.312	88.952	4.287	62.197	3.964

In Figure 5, the mean vectors are displayed graphically. The ETW values are plotted in the lower left, the CTD values in the upper right, and the RTD values in the middle. Lines connect the mean values of the different environments by company.

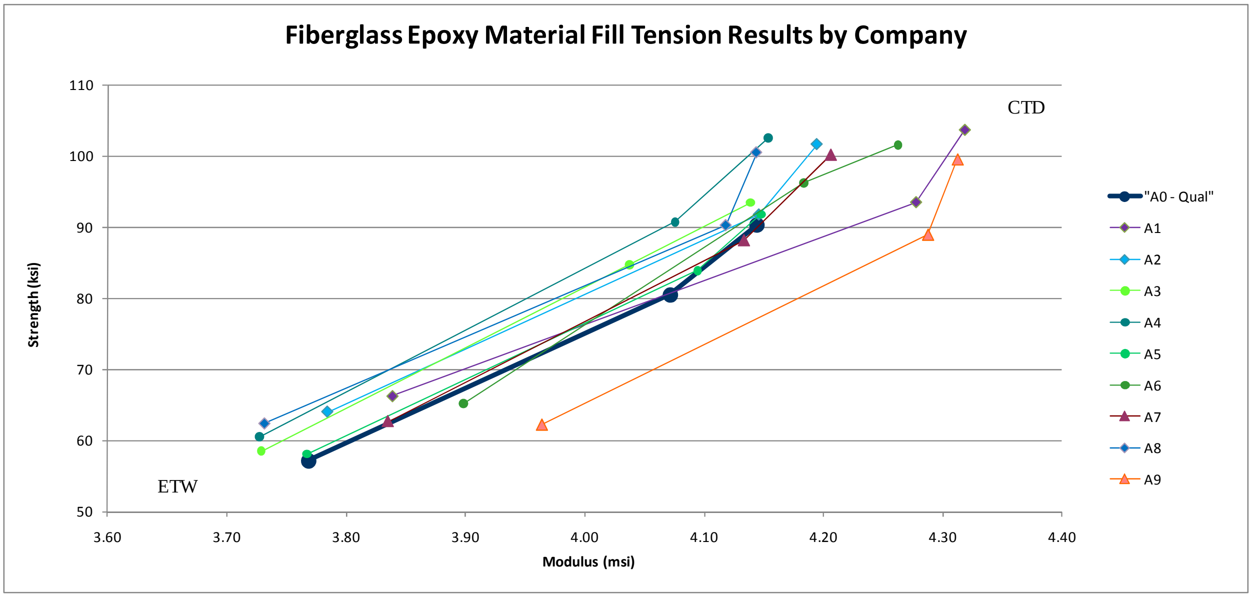

Figure 5. Fill tension mean vectors.

The assumed covariance matrix, Σ, is defined as follows:

$$\Sigma = \begin{bmatrix} 31.07 & 0.18 & 26.83 & .23 & 14.43 & .12 \\ .18 & .005 & .176 & .006 & .16 & .005 \\ 26.83 & .176 & 32.70 & .205 & 16.21 & .103 \\ .23 & .006 & .205 & .007 & .18 & .005 \\ 14.43 & .16 & 16.21 & .18 & 12.32 & .10 \\ .12 & .005 & 0.103 & .005 & .10 & .007 \end{bmatrix}$$

This covariance matrix was constructed using the data from all 24 panels available for this material; therefore, it includes the variance attributable to different producers as well as different cure cycle recipes for this material. Separating the sources of variability will prove useful when constructing basis values to accompany the equivalency criteria, but that is beyond the scope of this paper.

Theoretical issues arise when substituting an estimated matrix constructed from the sample data. These issues will be discussed in Chapter 5. For the example analysis, this matrix will be treated as the known covariance matrix for the population of fill compression test results for the fiberglass epoxy material.

4.2 Setting δ or Defining 'Close Enough'

Before a comparison can be made, δ must be determined. Recall that δ represents the largest allowable difference that is considered 'close enough.' One choice is to key δ to the producer's risk. At this point, it will not be an exact computation, since the final acceptance limit will fall inside the ellipsoid, with the boundary of points exactly δ from the center of the ellipse. Thus, the producer's risk based on δ is only approximate, although an exact value can be computed later.

Since this is a multivariate normal distribution, the norm has a chi-squared distribution. So a value can be found for δ that will correspond to a specified producer's risk. For example, to achieve a producers risk of approximately 5 percent, set $\delta = \frac{1.635}{\sqrt{\hat{\sigma}}}$, the value of a χ_6^2 for α = 5%. The area defined as the compliment of Θ_0, which represents acceptable product, will correspond to approximately 95 percent of the expected output.

4.3 Test Statistics and Results for Case 1

Differences between the mean vector of the qualification sample and the mean vector of each equivalency sample $\left(\bar{X} - \bar{Y}\right)$ are shown in Table 4. The non-centrality parameter of the distribution of T will be the same for all companies but will vary with δ. For this example, the non-centrality parameter is $\frac{n_1 n_2}{n_1 + n_2}\delta^2 = 4.0344$.

TABLE 4

DIFFERENCES OF MEAN VECTORS

Company Code	Differences of Equivalency Mean Vectors with Qualification Mean Vectors					
	Strength CTD (ksi)	**Modulus CTD (msi)**	**Strength RTD (ksi)**	**Modulus RTD (msi)**	**Strength ETW (ksi)**	**Modulus ETW (msi)**
A1	(13.3213)	(0.1746)	(12.9991)	(0.2059)	(9.0715)	(0.0702)
A2	(11.4246)	(0.0507)	(11.3181)	(0.0745)	(6.7874)	(0.0160)
A3	(3.1733)	0.0051	(4.2757)	0.0343	(1.3299)	0.0395
A4	(12.2208)	(0.0094)	(10.2465)	(0.0039)	(3.2970)	0.0411
A5	(1.5334)	(0.0036)	(3.4352)	(0.0228)	(0.9016)	0.0010
A6	(11.3013)	(0.1184)	(15.7893)	(0.1118)	(7.9558)	(0.1297)
A7	(9.8448)	(0.0622)	(7.6401)	(0.0617)	(5.4962)	(0.0669)
A8	(10.2526)	0.0001	(9.8492)	(0.0466)	(5.2739)	0.0365
A9	(9.3166)	(0.1682)	(8.5317)	(0.2157)	(5.0359)	(0.1958)

Critical values for the Case 1 test statistic for this example ($n_1 = 6$, $n_2 = 2$, $p = 6$) are shown in Table A-1 of Appendix A for values of δ from 0.1 to 6.0 and values of α ranging from 0.01 to 0.20. The null hypothesis is rejected when the test statistic is less than the critical value. When the null hypothesis is rejected, the two samples can be said to be equivalent—that is, they have a difference of less than δ—at the $1-\alpha$ confidence level.

The Case 1 test statistic for each company was computed, and the results are shown in Table 5 along with the value of δ required for the company to be considered equivalent to the qualification sample with $\alpha = 0.05$. These results indicate that only companies A5 and A3 were able to produce material sufficiently close to the qualification sample to fall within an acceptance ellipse for a value of δ less than two. However, company A5 passes even with the consumer's risk set at 1 percent.

TABLE 5

CASE 1 TEST STATISTIC EXAMPLE RESULTS

Company Code	**Case 1 Test Statistic T**	**Passes Equivalence at α = 0.10 for δ ≥**	**Passes Equivalence at α = 0.05 for δ ≥**	**Passes Equivalence at α = 0.01 for δ ≥**
A5	1.5908	0.1	0.1	1.6
A3	2.9431	1.2	1.6	2.4
A7	6.8096	2.5	2.8	3.5
A2	7.5873	2.6	3.0	3.6
A8	8.9699	2.9	3.2	3.8
A4	10.9461	3.2	3.5	4.1
A9	11.9352	3.4	3.7	4.3
A6	14.6322	3.7	4.0	4.6
A1	15.9013	3.9	4.2	4.8

Some options are available with this approach. We could center the ellipse at the mean of the combined values of all panels, rather than that of the qualification sample. Another option is to allow our ellipse to be stretched out toward the high end of strength and modulus as acceptable, as long as they increase at the proper proportion to each other, rather than insisting on the modulus associated with the mean strength of the qualification sample.

4.4 Test Statistics and Results for Case 2

The test statistic, which varies with ε when $\varepsilon = \frac{\delta}{\sqrt{\sigma}}$, was computed for each company for various values of ε. These results are shown in Table 6. A value above the critical value indicates that H_0 can be rejected. A value of zero for the test statistic indicates that it was larger than the value of δ. Table 7 indicates the smallest ε for each company that will reject H_0.

TABLE 6

CASE 2 TEST STATISTICS WITH α = 0.05

Company Code	Critical Value for α = .05								
	0	**0**	**0.00059**	**0.00073**	**0.00089**	**0.00104**	**0.00138**	**0.00155**	**0.00176**
	ε = 0.7	**ε = 0.9**	**ε = 1.4**	**ε = 1.5**	**ε = 1.6**	**ε = 1.7**	**ε = 1.9**	**ε = 2.0**	**ε = 2.1**
A1	0	0	0	0	0	0	0	0.00025	0.00262
A2	0	0	0.00014	0.00224	0.00683	0.01393	0.03562	0.05022	0.06731
A3	0	0.00036	0.03828	0.05337	0.07095	0.09104	0.13871	0.16630	0.19638
A4	0	0	0	0	0.00059	0.00355	0.01698	0.02745	0.04041
A5	0.001186	0.01106	0.07948	0.10067	0.12436	0.15054	0.21041	0.24410	0.28028
A6	0	0	0	0	0	0	0.00090	0.00427	0.01015
A7	0	0	0.00220	0.00676	0.01382	0.02339	0.05002	0.06708	0.08664
A8	0	0	0	0.00018	0.00236	0.00705	0.02392	0.03611	0.05080
A9	0	0	0	0	0	0	0.00438	0.01031	0.01874

TABLE 7

CASE 2 TEST STATISTICS EXAMPLE RESULTS

Company Code	Class II Test Statistic Passes Equivalence at α = 0.05 for ε >
A5	0.7
A3	0.9
A7	1.6
A2	1.8
A8	1.9
A4	2.0
A9	2.2
A6	2.4
A1	2.6

The Case 2 critical values and test statistics for a consumer's risk of 5 percent are shown graphically in Figure 6. When the test statistic is above the critical value, the sample for that company can be considered equivalent to the qualification sample for that value of ε.

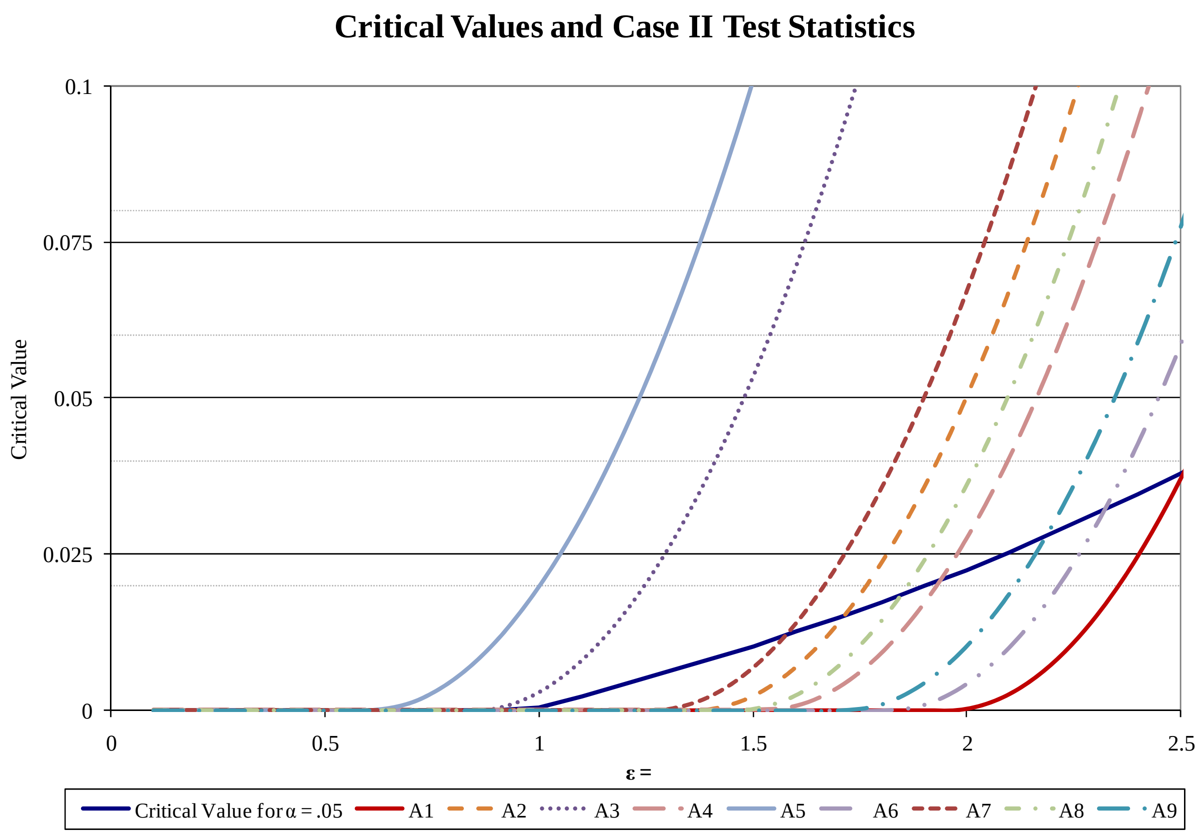

Figure 6. Critical values and Case 2 test statistics for FT data with α = .05.

4.5 Comparison with Current Method Results

Results of the current equivalency tests are displayed graphically in Figure 7. The open-top black rectangles represent the current acceptance limits for a producer's risk of 5 percent for the three environments. Points that lie outside the boxes have failed the equivalency. Table 8 shows the individual results for each company and each test using the current methodology.

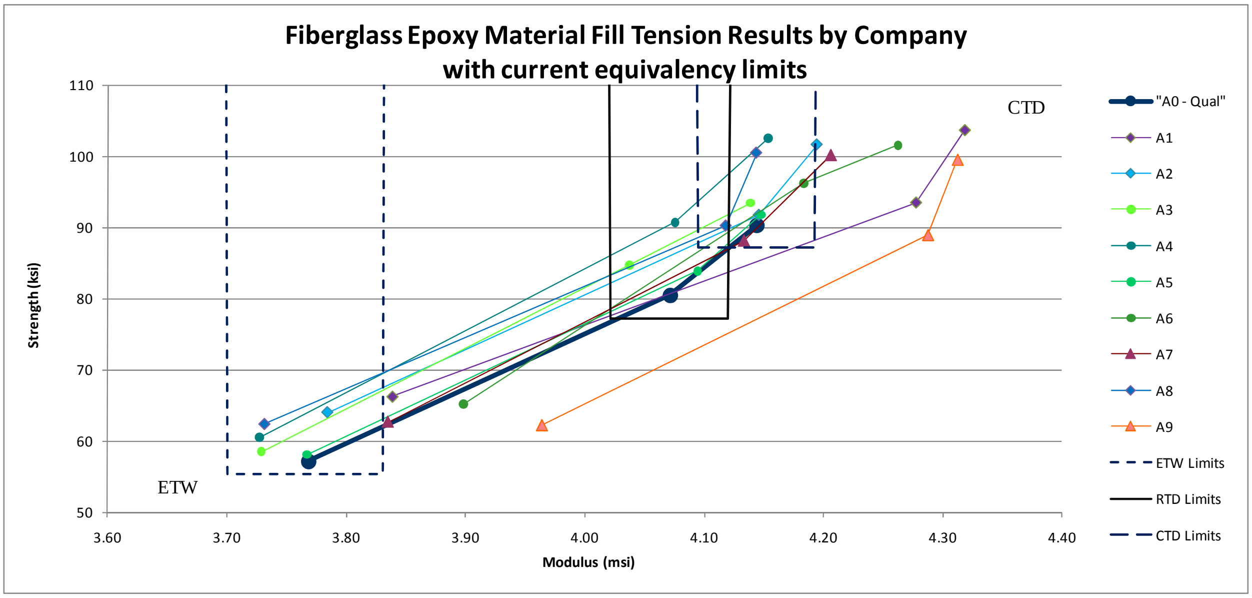

Figure 7. Fill tension mean vectors with current acceptance limits.

TABLE 8

FIBERGLASS EPOXY FILL TENSION TEST RESULTS AT α = 0.05 CURRENT METHOD

Company Code	Current Equivalency Test Results					
	Strength CTD (ksi)	Modulus CTD (msi)	Strength RTD (ksi)	Modulus RTD (msi)	Strength ETW (ksi)	Modulus ETW (msi)
A1	PASS	FAIL	PASS	FAIL	PASS	FAIL
A2	PASS	FAIL	PASS	FAIL	PASS	PASS
A3	PASS	PASS	PASS	PASS	PASS	PASS
A4	PASS	PASS	PASS	PASS	PASS	PASS
A5	PASS	PASS	PASS	PASS	PASS	PASS
A6	PASS	FAIL	PASS	FAIL	PASS	FAIL
A7	PASS	FAIL	PASS	FAIL	PASS	FAIL
A8	PASS	PASS	PASS	FAIL	PASS	PASS
A9	PASS	FAIL	PASS	FAIL	PASS	FAIL

None of the companies had any difficulties passing the strength tests, but modulus tests were problematic. Only three companies—A3, A4, and A5—lie within the equivalency limits for all environments. The remaining companies fall outside of it for at least one test result. While A3 and A5 were the two companies that were ranked closest to the qualification sample according to both Case 1 and Case 2 test statistics, companies A7, A2, and A8 all scored closer to the qualification sample than A4. There is an explanation.

Recall that the strength tests are evaluated with respect to a one-sided test. A material with higher strength values is not going to be rejected even though it may differ significantly from the strength values of the qualification sample. Company A4 has higher strength values, and that is the reason for the large distance measurement from the qualification sample. For this equivalence approach to be a viable alternative to the current method of assessing composite materials, an adjustment must be made in order to accommodate the one-sided hypothesis of the strength tests.

For the example data, those higher strength values are the reason that a company may require a large value of δ for equivalence using the Case 1 and Case 2 test statistics. The sample data must be checked to determine whether they fall inside the union of the acceptance ellipsoid with the original acceptance box.

Figure 8 shows an artist's rendition of what the various acceptance regions would look like in three dimensions. The mean vectors of various samples are displayed as white dots. The larger black dot is actually a very small ellipsoid centered on the qualification mean vector. This black dot is the acceptance region for a consumer's risk of 5 percent and $\delta = 0$. The blue and green ellipsoids represent acceptance regions for a producer's risk of 5 and 1 percent, respectively.

The blue box is the open-ended acceptance region using current methods. The sides represent the limits for the mean of the modulus, both upper and lower. The bottom and back represent the minimum value for the mean of two different strength tests, but there is no top or front because there is no maximum placed on the strength test results.

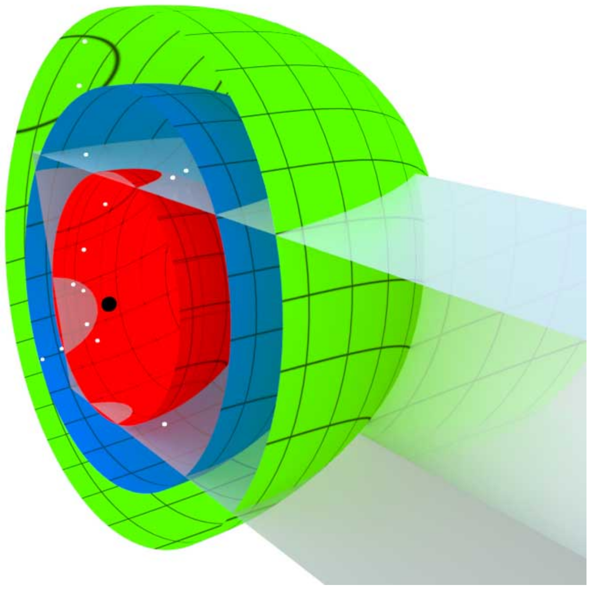

Figure 8. Artist's rendition of multivariate acceptance regions.

The red ellipse is contained inside the box representing current acceptance limits, so it is the ellipse that corresponds with current limits.   This diagram shows the problem regarding the acceptance ellipses. Essentially, any sample mean vectors that within the blue box are accepted as equivalent by current standards. This mismatch is the reason an adjustment will need to be made to accommodate the one-sided hypothesis of the strength tests. This is not a difficult adjustment to make.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

5.1 Engineering Basis Values

Table 9 [14] shows the basis values that are computed using currently accepted methods for the example data. Due to large batch-to-batch variability, the ANOVA method was required for three of the four environmental conditions. This method requires five independent batches, so only estimates are available for those conditions. A-basis values require five independent batches for all methods, so only estimates are provided. The modified CV method approach inflates the variation of the qualification batch when the coefficient of variation is small (under 8 percent). This attempts to make the basis values more realistic and to compensate for the variation over time and between producers, which the qualification sample does not include.

TABLE 9

BASIS VALUES FOR GLASS 6781 FILL TENSION

Fill Tension Strength Basis Values and Statistics				
Env	CTD	RTD	ETW	ETW2
Mean	90.06	80.50	57.22	55.32
Stdev	3.31	3.81	2.29	1.74
CV	3.67	4.73	4.00	3.15
Mod CV	6.00	6.37	6.00	6.00
Min	84.24	72.21	52.09	52.03
Max	95.81	84.60	59.77	59.62
No. batches	3	3	3	3
No. spec.	19	19	19	22
Basis Values and Estimates				
B-basis value	83.61			
B-estimate		55.47	46.28	47.58
A-estimate	79.03	37.60	38.47	42.05
Method	Normal	ANOVA	ANOVA	ANOVA
Modified CV Basis Values and Estimates				
B-basis value	79.52	NA	NA	49.06
A-estimate	72.06	NA	NA	44.59
Method	normal	NA	NA	normal

When the equivalence approach discussed in this thesis is used, these basis values will not be appropriate when the value of δ exceeds the difference between the qualification mean and the minimum acceptable value of the mean strength. This is due to any value within the acceptable ellipsoid being considered acceptable, which will include values that fall below the original acceptance limits computed from the qualification sample in that case. This method allows for additional variation with large values of δ, but this must be reflected in the engineering basis values. Fortunately, this is not a difficult computation.

5.2 Engineering Basis Values to Accompany δ

Since any value within the acceptable ellipsoid is possible, to compute basis values it is necessary to find the point on the ellipse with the minimum value for that property (x). Then the basis value for that property is computed by assuming that x is the mean of the qualification sample. Figure 9 shows warp compression RTD qualification and equivalency data for Glass 6781, corresponding basis values, and acceptance ellipses for the bivariate distribution of strength and modulus. The B-basis value computed from the qualification sample results in the same B-basis value as with $\delta = 1.1$.

5.3 Advantages of Multivariate Hypothesis Test of Equivalence

This approach begins with an acceptance region that lies inside the δ-ellipsoid (which is the boundary of the maximum possible acceptance region) and expands toward that boundary as the sample size increases. As the database of material test results increases, the expected variance decreases due to the larger sample size. Thus, the acceptance region of each grade can be expected to increase as the boundary of the acceptance region moves closer to the maximum acceptance region, which is defined by δ-ellipsoid around the qualification mean vector.

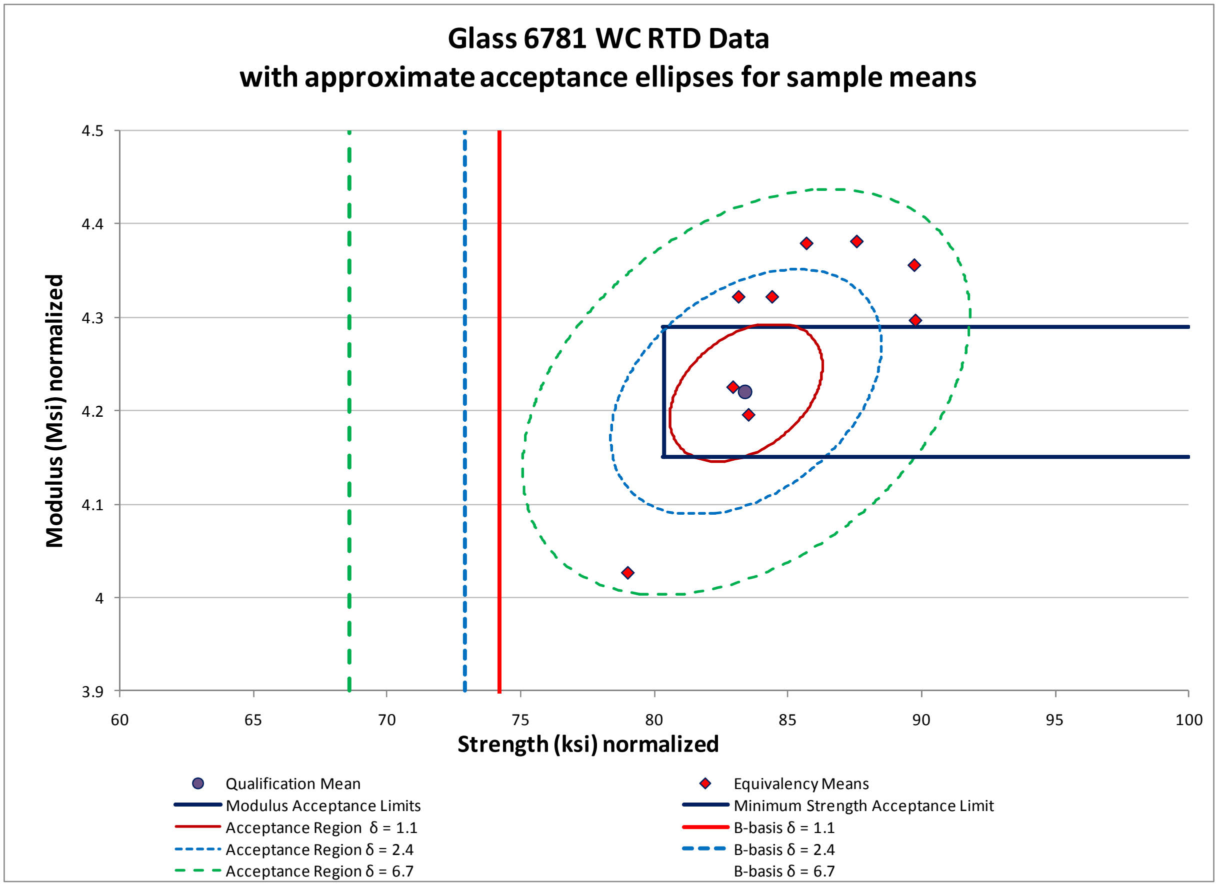

Figure 9. Glass 6781 warp compression RTD strength and modulus results.

As mentioned in Chapter 2, this approach also eliminates the side effect of producers being benefitted by smaller sample sizes and larger uncertainty about their product's test results. Instead, larger sample sizes will result in a larger ellipsoidal acceptance area.

In addition, the basis values can be expected to climb upward as the variance decreases. This means that over time, as the database accumulates more information, basis values may increase, and those higher basis values will retroactively include all previously accepted material for that grade.

Producers would be able to both select an acceptable producer's risk and provide their customers with a specified probability that their material will meet those basis values. These are guarantees that do not exist with the current methodology.

5.4 Checking Assumption of Equal Covariance Matrices

Since a primary assumption of this analysis is that the covariance matrices are the same, those covariance matrices will need to be verified as similar before materials can be compared in this manner. Anderson [5] established a method to accomplish this. It remains to be seen if this is a useful method or if it will nearly always classify two panels as having "different" co-variance matrices. If it is the latter, a similar approach for 'close enough,' will need to be developed for testing the equality of co-variance matrices before the results of applying it to the mean vectors of composite test results can be considered sound.

5.5 Recommendations

I recommend that an analysis of NCAMP materials be done using this technique to create the following categories of basis values:

- TWIN: Engineering basis values generated with the current methodology. This is expected to have a producer's risk of between 70 and 30 percent.
- Grade A: Engineering basis values generated with the current methodology is valid for this category. However, Grade A material may fall outside the "TWIN" category but does so without adversely affecting the strength characteristics.
- Grade B: Engineering basis values generated to accompany acceptance limits set with a producer's risk of approximately 5 percent.
- Grade C: Engineering basis values generated to accompany acceptance limits set with a producer's risk of 1 percent or less.

As more producers come on line with a material, a product that qualifies as "TWIN" can be added to the database of test results from which the basis values for "TWIN" are computed. Any materials that qualify as "Grade A" can be added to the database of test results from which the basis values for “Grade A” are computed; likewise for “Grade B” and “Grade C.” Materials that do not qualify as Grade C would require a larger set of test results in order to recommend basis values.

While a producer might be disappointed to have its material rated as Grade B or Grade C rather than Grade A, this may be preferable to the expense and delay of running additional tests to determine engineering basis values for their materials.

REFERENCES

REFERENCES

1. MIL-STD-105E: *Sampling Procedures and Tables for Inspection by Attributes*, Department of Defense, Washington, DC 20301, 1989.

2. S. Wellek, *Testing Statistical Hypotheses of Equivalence*, Chapman & Hall/CRC Boca Raton, Fl, 33431, 2003.

3. *CMH-HDBK-1G : The Composite Materials Handbook*, ASTM International, West Conshohocken, PA, 2010 .

4. *"DOT/FAA/AR-03/19: Material Qualification and Equivalency for Polymer Matrix Composite Material Systems: Updated Procedure"*, U.S. Department of Transportation, Federal Aviation Administration, Washington, D.C., 20591, Sept. 2003.

5. T. W. Anderson, "*An Introduction to Multivariate Statistical Analysis,*" 3rd edition, John Wiley & Sons, Inc., Hoboken, NJ, 2006.

6. R. A. Johnson and D. W. Wichern, "*Applied Multivariate Statistical Analysis,"* 5th edition, Pearson Education, Upper Saddle River, NJ, 2002.

7. J. R. Schott, *Matrix Analysis for Statistics*, John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158-0012, 1997

8. R. V. Hoag and A. T. Craig, *Introduction to Mathematical Statistics* 5th edition, Macmillian Publishing Co., Inc., New York, New York, 10022.

9. G. R. Shorack, *Probability for Statisticians,* Springer-Verlag New York, Inc., New York, New York, 2000.

10. X. Hu and F. T. Wright, "*Monotonicity Properties of the Power Functions of Likelihood Ratio Tests for Normal Mean Hypotheses Constrained by a Linear Space and a Cone*", Annals of Statistics, Vol 22, No. 3, 1547-1554, 1994.

11. T. W. Anderson, "*The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities"*, Proc Amer. Math. Soc. Vol 6, 170-176, 1955.

12. X. Hu, "*Multivariate Analysis Without vec and xO"*, The American Statistician, Vol. 61, No. 1. , Feb. 2007

13. *Advanced Composites Group ACG MTM45-1 6781 S-2 Glass 35% RC Qualification Material Property Data Report*, Test Report Number: CAM-RP-2009-001 Rev. A, National Institute for Aviation Research, Wichita, Kansas, 67218, Feb 2010.

14. *Advanced Composites Group MTM45-1/ Style 6781 S2 Glass Qualification Statistical Analysis Report*, NCAMP Report # NCP-RP-2009-001 N/C, National Institute for Aviation Research, Wichita, Kansas, 67218 , Mar 2010.

APPENDICES

APPENDIX A

TABLES OF CRITICAL VALUES

TABLE A-1

CRITICAL VALUES FOR CASE 1 TEST STATISTIC

Critical Values for $n_1 = 6$, $n_2 = 2$, and $p = 6$					
δ	**α = 0.01**	**α = 0.02**	**α = 0.05**	**α = 0.10**	**α = 0.20**
0.1	0.8743	1.1373	1.6395	2.2096	3.0778
0.2	0.8809	1.1458	1.6518	2.2263	3.1009
0.3	0.8919	1.1602	1.6725	2.2541	3.1396
0.4	0.9076	1.1806	1.7018	2.2936	3.1944
0.5	0.9281	1.2072	1.7401	2.3450	3.2656
0.6	0.9538	1.2405	1.7879	2.4089	3.3539
0.7	0.9849	1.2808	1.8455	2.4860	3.4599
0.8	1.0219	1.3286	1.9138	2.5770	3.5845
0.9	1.0652	1.3846	1.9934	2.6826	3.7285
1.0	1.1154	1.4494	2.0851	2.8039	3.8928
1.1	1.1733	1.5237	2.1899	2.9418	4.0784
1.2	1.2394	1.6084	2.3087	3.0972	4.2860
1.3	1.3147	1.7045	2.4425	3.2713	4.5167
1.4	1.3999	1.8129	2.5925	3.4650	4.7712
1.5	1.4962	1.9347	2.7597	3.6793	5.0503
1.6	1.6045	2.0709	2.9451	3.9152	5.3547
1.7	1.7259	2.2228	3.1499	4.1736	5.6850
1.8	1.8615	2.3914	3.3751	4.4553	6.0417
1.9	2.0125	2.5778	3.6215	4.7610	6.4252
2.0	2.1801	2.7832	3.8902	5.0914	6.8359
2.1	2.3653	3.0085	4.1818	5.4470	7.2742
2.2	2.5692	3.2546	4.4971	5.8285	7.7403
2.3	2.7930	3.5226	4.8368	6.2362	8.2345
2.4	3.0374	3.8132	5.2014	6.6705	8.7570
2.5	3.3035	4.1271	5.5914	7.1317	9.3079
2.6	3.5919	4.4649	6.0074	7.6202	9.8874
2.7	3.9036	4.8273	6.4496	8.1362	10.4956
2.8	4.2390	5.2148	6.9184	8.6800	11.1326
2.9	4.5988	5.6278	7.4142	9.2517	11.7986
3.0	4.9835	6.0668	7.9373	9.8516	12.4936
3.1	5.3936	6.5321	8.4878	10.4797	13.2176
3.2	5.8295	7.0240	9.0659	11.1363	13.9709
3.3	6.2916	7.5429	9.6720	11.8215	14.7534

Critical Values for $n_1 = 6$, $n_2 = 2$, and $p = 6$					
δ	**α = 0.01**	**α = 0.02**	**α = 0.05**	**α = 0.10**	**α = 0.20**
3.4	6.7801	8.0889	10.3061	12.5353	15.5652
3.5	7.2955	8.6624	10.9684	13.2779	16.4064
3.6	7.8379	9.2635	11.6591	14.0494	17.2769
3.7	8.4076	9.8925	12.3783	14.8498	18.1769
3.8	9.0048	10.5494	13.1260	15.6792	19.1064
3.9	9.6298	11.2345	13.9025	16.5378	20.0653
4.0	10.2826	11.9479	14.7077	17.4255	21.0538
4.1	10.9635	12.6897	15.5418	18.3425	22.0719
4.2	11.6725	13.4600	16.4049	19.2887	23.1196
4.3	12.4099	14.2589	17.2969	20.2642	24.1969
4.4	13.1757	15.0866	18.2181	21.2691	25.3039
4.5	13.9701	15.9430	19.1685	22.3035	26.4405
4.6	14.7931	16.8284	20.1480	23.3672	27.6068
4.7	15.6449	17.7427	21.1568	24.4604	28.8029
4.8	16.5254	18.6861	22.1948	25.5832	30.0286
4.9	17.4349	19.6585	23.2623	26.7354	31.2841
5.0	18.3733	20.6601	24.3591	27.9173	32.5693
5.1	19.3407	21.6909	25.4853	29.1287	33.8843
5.2	20.3373	22.7509	26.6410	30.3697	35.2291
5.3	21.3629	23.8402	27.8261	31.6404	36.6037
5.4	22.4178	24.9588	29.0408	32.9407	38.0081
5.5	23.5019	26.1068	30.2850	34.2707	39.4423
5.6	24.6153	27.2842	31.5588	35.6304	40.9063
5.7	25.7580	28.4910	32.8621	37.0197	42.4001
5.8	26.9300	29.7273	34.1951	38.4388	43.9238
5.9	28.1314	30.9931	35.5576	39.8876	45.4773
6.0	29.3623	32.2884	36.9499	41.3662	47.0607

In Table 2, a value of "0" indicates that for that combination of ε and α, the simulation produced no values for T above 0 at the $1-\alpha^{th}$ percentile, while a value of "0.000" indicates that the simulation produced a value for the $1-\alpha^{th}$ percentile between 0 and 0.0005.

TABLE A-2

CRITICAL VALUES FOR CASE 2 TEST STATISTIC

Critical Values for $n_1 = 6$, $n_2 = 2$, **p = 3**, $\delta = \varepsilon\sqrt{\sigma}$					
ε	**α = 0.01**	**α = 0.02**	**α = 0.05**	**α = 0.10**	**α = 0.20**
0.1	0	0	0	0	0
0.2	0	0	0	0	0
0.3	0.000	0.000	0	0	0
0.4	0.002	0.001	0	0	0
0.5	0.003	0.002	0.001	0	0
0.6	0.006	0.004	0.002	0.000	0
0.7	0.009	0.007	0.004	0.001	0
0.8	0.013	0.010	0.006	0.003	0.000
0.9	0.017	0.013	0.009	0.005	0.001
1.0	0.022	0.017	0.011	0.007	0.002
1.1	0.029	0.022	0.015	0.009	0.003
1.2	0.035	0.027	0.018	0.012	0.005
1.3	0.042	0.032	0.022	0.015	0.006
1.4	0.049	0.038	0.026	0.018	0.008
1.5	0.057	0.044	0.031	0.021	0.011
1.6	0.065	0.051	0.035	0.025	0.013
1.7	0.073	0.057	0.040	0.029	0.015
1.8	0.081	0.064	0.045	0.032	0.018
1.9	0.090	0.071	0.051	0.037	0.021
2.0	0.096	0.077	0.056	0.041	0.024
2.1	0.104	0.084	0.061	0.045	0.027
2.2	0.112	0.091	0.067	0.050	0.030
2.3	0.121	0.099	0.073	0.054	0.033
2.4	0.130	0.107	0.080	0.059	0.037
2.5	0.139	0.115	0.086	0.065	0.040
2.6	0.148	0.123	0.093	0.070	0.044
2.7	0.157	0.132	0.100	0.076	0.048
2.8	0.169	0.141	0.107	0.082	0.052
2.9	0.179	0.150	0.115	0.088	0.056
3.0	0.190	0.160	0.123	0.094	0.061
3.1	0.201	0.170	0.131	0.101	0.066
3.2	0.214	0.181	0.140	0.107	0.071

Critical Values for $n_1 = 6$, $n_2 = 2$, **p = 3**, $\delta = \varepsilon\sqrt{\sigma}$					
ε	**α = 0.01**	**α = 0.02**	**α = 0.05**	**α = 0.10**	**α = 0.20**
3.3	0.226	0.192	0.148	0.114	0.076
3.4	0.240	0.204	0.158	0.122	0.081
3.5	0.253	0.215	0.167	0.129	0.087
3.6	0.267	0.228	0.177	0.137	0.092
3.7	0.281	0.240	0.187	0.145	0.098
3.8	0.296	0.253	0.197	0.153	0.104
3.9	0.312	0.267	0.207	0.161	0.110
4.0	0.328	0.280	0.218	0.170	0.116
4.1	0.343	0.294	0.229	0.179	0.122
4.2	0.360	0.308	0.241	0.188	0.129
4.3	0.377	0.322	0.252	0.197	0.136
4.4	0.394	0.337	0.263	0.206	0.142
4.5	0.413	0.353	0.276	0.216	0.149
4.6	0.430	0.369	0.288	0.225	0.156
4.7	0.449	0.384	0.300	0.235	0.163
4.8	0.467	0.400	0.313	0.245	0.171
4.9	0.486	0.417	0.326	0.256	0.178
5.0	0.506	0.433	0.339	0.266	0.186

Critical Values for $n_1 = 6$, $n_2 = 2$, **p = 4**, $\delta = \varepsilon\sqrt{\sigma}$					
ε	**α = 0.01**	**α = 0.02**	**α = 0.05**	**α = 0.10**	**α = 0.20**
0.1	0	0	0	0	0
0.2	0	0	0	0	0
0.3	0	0	0	0	0
0.4	0.000	0	0	0	0
0.5	0.001	0.000	0	0	0
0.6	0.003	0.002	0	0	0
0.7	0.005	0.003	0.001	0	0
0.8	0.007	0.005	0.002	0.000	0
0.9	0.009	0.007	0.004	0.001	0
1.0	0.012	0.010	0.006	0.003	0
1.1	0.015	0.012	0.008	0.004	0.000
1.2	0.019	0.015	0.011	0.006	0.001
1.3	0.023	0.019	0.013	0.009	0.002
1.4	0.027	0.022	0.016	0.011	0.004
1.5	0.031	0.026	0.019	0.013	0.005
1.6	0.036	0.030	0.023	0.016	0.007
1.7	0.041	0.035	0.026	0.019	0.009
1.8	0.047	0.039	0.030	0.022	0.011
1.9	0.052	0.044	0.033	0.025	0.013
2.0	0.058	0.049	0.037	0.028	0.015

Critical Values for $n_1 = 6$, $n_2 = 2$, **p = 4**, $\delta = \varepsilon\sqrt{\sigma}$					
ε	**α = 0.01**	**α = 0.02**	**α = 0.05**	**α = 0.10**	**α = 0.20**
2.1	0.064	0.054	0.042	0.031	0.018
2.2	0.070	0.060	0.046	0.035	0.020
2.3	0.077	0.065	0.051	0.039	0.023
2.4	0.083	0.071	0.055	0.042	0.026
2.5	0.090	0.077	0.060	0.046	0.029
2.6	0.097	0.084	0.066	0.051	0.032
2.7	0.105	0.090	0.071	0.055	0.035
2.8	0.112	0.097	0.076	0.060	0.039
2.9	0.120	0.104	0.082	0.064	0.042
3.0	0.129	0.112	0.088	0.069	0.046
3.1	0.136	0.118	0.094	0.074	0.049
3.2	0.145	0.126	0.101	0.080	0.053
3.3	0.155	0.135	0.107	0.085	0.057
3.4	0.164	0.143	0.114	0.090	0.061
3.5	0.174	0.151	0.121	0.096	0.066
3.6	0.184	0.160	0.128	0.102	0.070
3.7	0.194	0.169	0.136	0.108	0.075
3.8	0.205	0.179	0.144	0.114	0.079
3.9	0.216	0.188	0.151	0.121	0.084
4.0	0.227	0.198	0.159	0.127	0.089
4.1	0.238	0.207	0.167	0.134	0.094
4.2	0.250	0.218	0.175	0.141	0.099
4.3	0.261	0.229	0.184	0.148	0.104
4.4	0.274	0.239	0.193	0.154	0.109
4.5	0.286	0.250	0.201	0.162	0.115
4.6	0.298	0.261	0.210	0.169	0.120
4.7	0.311	0.272	0.219	0.176	0.126
4.8	0.324	0.284	0.229	0.184	0.131
4.9	0.338	0.295	0.238	0.191	0.137
5.0	0.350	0.306	0.247	0.199	0.143

Critical Values for $n_1 = 6$, $n_2 = 2$, **p = 5**, $\delta = \varepsilon\sqrt{\sigma}$					
ε	**α = 0.01**	**α = 0.02**	**α = 0.05**	**α = 0.10**	**α = 0.20**
0.1	0	0	0	0	0
0.2	0	0	0	0	0
0.3	0	0	0	0	0
0.4	0	0	0	0	0
0.5	0	0	0	0	0
0.6	0.001	0	0	0	0
0.7	0.003	0.001	0	0	0
0.8	0.004	0.003	0.000	0	0

Critical Values for $n_1 = 6$, $n_2 = 2$, **p = 5**, $\delta = \varepsilon\sqrt{\sigma}$					
ε	**α = 0.01**	**α = 0.02**	**α = 0.05**	**α = 0.10**	**α = 0.20**
0.9	0.006	0.004	0.001	0	0
1.0	0.008	0.006	0.003	0.000	0
1.1	0.010	0.008	0.005	0.001	0
1.2	0.013	0.011	0.007	0.003	0
1.3	0.015	0.013	0.009	0.005	0.000
1.4	0.018	0.016	0.011	0.007	0.001
1.5	0.022	0.019	0.014	0.009	0.002
1.6	0.025	0.022	0.016	0.011	0.003
1.7	0.029	0.025	0.019	0.013	0.005
1.8	0.033	0.028	0.022	0.016	0.007
1.9	0.037	0.032	0.025	0.018	0.008
2.0	0.042	0.036	0.028	0.021	0.010
2.1	0.046	0.040	0.031	0.024	0.012
2.2	0.051	0.044	0.035	0.027	0.015
2.3	0.056	0.049	0.039	0.030	0.017
2.4	0.061	0.053	0.043	0.033	0.020
2.5	0.067	0.058	0.047	0.036	0.022
2.6	0.072	0.063	0.051	0.040	0.025
2.7	0.078	0.068	0.055	0.043	0.028
2.8	0.084	0.074	0.060	0.047	0.030
2.9	0.090	0.079	0.064	0.051	0.033
3.0	0.097	0.085	0.069	0.055	0.037
3.1	0.103	0.091	0.074	0.059	0.040
3.2	0.110	0.097	0.079	0.064	0.043
3.3	0.117	0.104	0.084	0.068	0.046
3.4	0.124	0.110	0.090	0.072	0.050
3.5	0.132	0.117	0.096	0.077	0.054
3.6	0.140	0.124	0.101	0.082	0.057
3.7	0.148	0.131	0.107	0.087	0.061
3.8	0.156	0.138	0.113	0.092	0.065
3.9	0.164	0.146	0.120	0.097	0.069
4.0	0.173	0.153	0.126	0.102	0.073
4.1	0.181	0.161	0.132	0.108	0.077
4.2	0.190	0.169	0.139	0.113	0.081
4.3	0.200	0.177	0.146	0.119	0.086
4.4	0.209	0.185	0.152	0.124	0.090
4.5	0.219	0.194	0.159	0.130	0.094
4.6	0.228	0.202	0.167	0.136	0.099
4.7	0.238	0.211	0.174	0.142	0.103
4.8	0.247	0.220	0.181	0.148	0.108
4.9	0.257	0.228	0.188	0.154	0.112
5.0	0.268	0.238	0.196	0.160	0.117

Critical Values for $n_1 = 6$, $n_2 = 2$, **p = 6**, $\delta = \varepsilon\sqrt{\sigma}$					
ε	**α = 0.01**	**α = 0.02**	**α = 0.05**	**α = 0.10**	**α = 0.20**
0.1	0	0	0	0	0
0.2	0	0	0	0	0
0.3	0	0	0	0	0
0.4	0	0	0	0	0
0.5	0	0	0	0	0
0.6	0	0	0	0	0
0.7	0.001	0	0	0	0
0.8	0.002	0.000	0	0	0
0.9	0.004	0.002	0	0	0
1.0	0.006	0.004	0.000	0	0
1.1	0.008	0.006	0.002	0	0
1.2	0.010	0.008	0.004	0.000	0
1.3	0.012	0.010	0.006	0.002	0
1.4	0.014	0.012	0.008	0.003	0
1.5	0.017	0.014	0.010	0.005	0.000
1.6	0.019	0.017	0.013	0.007	0.001
1.7	0.022	0.019	0.015	0.010	0.002
1.8	0.026	0.022	0.017	0.012	0.003
1.9	0.029	0.025	0.020	0.014	0.005
2.0	0.032	0.028	0.022	0.016	0.007
2.1	0.036	0.032	0.025	0.019	0.009
2.2	0.040	0.035	0.028	0.021	0.011
2.3	0.044	0.039	0.031	0.024	0.013
2.4	0.048	0.043	0.035	0.027	0.015
2.5	0.053	0.047	0.038	0.030	0.017
2.6	0.057	0.051	0.041	0.033	0.020
2.7	0.062	0.055	0.045	0.036	0.022
2.8	0.067	0.059	0.049	0.039	0.025
2.9	0.072	0.064	0.053	0.042	0.028
3.0	0.077	0.069	0.057	0.046	0.030
3.1	0.083	0.074	0.061	0.049	0.033
3.2	0.088	0.079	0.065	0.053	0.036
3.3	0.094	0.084	0.070	0.057	0.039
3.4	0.100	0.089	0.074	0.061	0.042
3.5	0.106	0.095	0.079	0.065	0.045
3.6	0.113	0.101	0.084	0.069	0.048
3.7	0.119	0.106	0.089	0.073	0.052
3.8	0.126	0.112	0.094	0.077	0.055
3.9	0.132	0.119	0.099	0.081	0.059
4.0	0.139	0.125	0.104	0.086	0.062
4.1	0.146	0.131	0.109	0.090	0.066

Critical Values for $n_1 = 6$, $n_2 = 2$, **p = 6**, $\delta = \varepsilon\sqrt{\sigma}$					
ε	**α = 0.01**	**α = 0.02**	**α = 0.05**	**α = 0.10**	**α = 0.20**
4.2	0.154	0.138	0.115	0.095	0.069
4.3	0.161	0.145	0.121	0.100	0.073
4.4	0.168	0.151	0.126	0.105	0.077
4.5	0.177	0.158	0.132	0.109	0.081
4.6	0.184	0.165	0.138	0.114	0.084
4.7	0.191	0.172	0.144	0.119	0.088
4.8	0.200	0.180	0.150	0.124	0.092
4.9	0.208	0.187	0.156	0.130	0.096
5.0	0.217	0.194	0.163	0.135	0.100

APPENDIX B

SAS CODE

SAS Code to Generate Table A in Appendix A

```
Data TestStat2;
n1 = 6;
 n2 = 2;
p = 6;
do delta = .1 to 4 by 0.1;
        ncp = delta*delta*n1*n2/(n1+n2);
        do q = 0 to .99 by 0.01;
                x = cinv(q, p, ncp);
                y = cdf('CHISQ',x, p, ncp);
                output;
        end;
 end;

run;
```

SAS Code to Generate Table B in Appendix A

```
*---------------------------------------------+
| April 2, 2010                                          |
| Generate simulated random test statistics   |
+---------------------------------------------*;

/*  generate random values */

   data work.temp2;
/*  Code to allow computations of multiple values of n1 and n2 */
 /*   do n1 = 3 to 10;
         do n2 = 2 to 8;
   */
   n1 = 6; n2 = 2;
   p=6;
        m = n1 + n2;
             do p = 3 to m-2;
                  do _j_ = .1 to 5 by .1;
                       expR = (m-2)*p; /* the expected value for sigma is the degrees of
freedom of chi-square dist divided by m */
                        epsilon = _j_;
                             ncp = (n1*n2*epsilon*epsilon)/m;
                        retain _seed_ 0;
                             do _i_ = 1 to 1000000;
                                  R = RAND('CHISQUARE', (m-2)*p);
                                  T1 = RAND('UNIFORM');
                                       if(T1 = 0) then T1 = RAND('UNIFORM');
                                  T3 = quantile('CHISQ', T1, p, ncp);
                                  T4 = (T3*m)/(n1*n2);
                                  If sqrt(T4) < epsilon*sqrt(n1*n2/m) then T =
(epsilon - sqrt(T4*n1*n2/m))**2/R; Else T = 0;
                                  output;
                       end;
                       end;
                  end;
/*             end;
        end;
 end;
*/
Keep n1 n2 p epsilon ncp T4 R T;
run;
```

```
proc sort; by p epsilon;
run;
/*  Run univariate to determine quantiles and statistics for each set of test results */
proc univariate data = work.temp2 noprint;
by p epsilon;
 var T ;
output out=sasuser.six_two pctlpts = 80 98 pctlpre= T
pctlname pct80 pct98
mean = mean std = stdev p90 = pct90 p95=pct95 p99 = pct99 max = max
;
 run;

quit;

data work.temp2;
/*  Code to allow computations of multiple values of n1 and n2 */
 /*   do n1 = 3 to 10;
          do n2 = 2 to 8;
   */
   n1 = 6; n2 = 2;
   p=6;
              m = n1 + n2;
                    do p = 3 to m-2;
                         do _j_ = .1 to 5 by .1;
                              expR = (m-2)*p; /* the expected value for sigma is the degrees of
freedom of chi-square dist divided by m */
                               epsilon = _j_;
                                    ncp = (n1*n2*epsilon*epsilon)/m;
                               retain _seed_ 0;
                                    do _i_ = 1 to 1000000;
                                        R = RAND('CHISQUARE', (m-2)*p);
                                        T1 = RAND('UNIFORM');
                                             if(T1 = 0) then T1 = RAND('UNIFORM');
                                        T3 = quantile('CHISQ', T1, p, ncp);
                                        T4 = (T3*m)/(n1*n2);
                                        If sqrt(T4) < epsilon*sqrt(n1*n2/m) then T =
(epsilon - sqrt(T4*n1*n2/m))**2/R; Else T = 0;
                                        output;
                              end;
                              end;
                         end;
/*                  end;
          end;
 end;
*/
Keep n1 n2 p epsilon ncp T4 R T;
```

```
run;

proc sort; by p epsilon;
run;
/*  Run univariate to determine quantiles and statistics for each set of test results */
proc univariate data = work.temp2 noprint;
by p epsilon;
 var T ;
output out=sasuser.six_two2 pctlpts = 80 98 pctlpre= T
pctlname pct80 pct98
mean = mean std = stdev p90 = pct90 p95=pct95 p99 = pct99 max = max
;
 run;

quit;

data sasuser.sims;
set work.temp2;
run;
quit;
```