Confidence sets for the ratio of variance components in a mixed linear model with two variance components

Tsung-Hua Lin

Iowa State University
INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the original text directly from the copy submitted. Thus, some dissertation copies are in typewriter face, while others may be from a computer printer.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyrighted material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is available as one exposure on a standard 35 mm slide or as a 17" × 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. 35 mm slides or 6" × 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Confidence sets for the ratio of variance components in a mixed linear model with two variance components

Lin, Tsung-Hua, Ph.D.

Iowa State University, 1987
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark √.

1. Glossy photographs or pages
2. Colored illustrations, paper or print
3. Photographs with dark background
4. Illustrations are poor copy
5. Pages with black marks, not original copy
6. Print shows through as there is text on both sides of page
7. Indistinct, broken or small print on several pages
8. Print exceeds margin requirements
9. Tightly bound copy with print lost in spine
10. Computer printout pages with indistinct print
11. Page(s) lacking when material received, and not available from school or author.
12. Page(s) seem to be missing in numbering only as text follows.
13. Two pages numbered. Text follows.
14. Curling and wrinkled pages
15. Dissertation contains pages with print at a slant, filmed as received
16. Other

UMI
Confidence sets for the ratio of variance components in a mixed linear model with two variance components

by

Tsung-Hua Lin

A Dissertation Submitted to the Graduate Faculty in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1987
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. PRELIMINARY DISCUSSION AND LITERATURE REVIEW</td>
<td>1</td>
</tr>
<tr>
<td>1.1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2. Relationship between Confidence Sets and Tests of Hypotheses</td>
<td>9</td>
</tr>
<tr>
<td>1.3. Invariance and Similarity</td>
<td>11</td>
</tr>
<tr>
<td>1.4. Literature Review</td>
<td>18</td>
</tr>
<tr>
<td>2. APPLYING THE PRINCIPLES OF INVARIANCE AND SUFFICIENCY</td>
<td>24</td>
</tr>
<tr>
<td>2.1. A Reduction of $y$ by Translation Invariance</td>
<td>24</td>
</tr>
<tr>
<td>2.2. A Reduction of $(t,u)$ by Orthogonal Invariance</td>
<td>27</td>
</tr>
<tr>
<td>2.3. A Reduction of $(t,SS_g)$ by Scale Invariance</td>
<td>29</td>
</tr>
<tr>
<td>2.4. A Reduction of $y$ by Sufficiency and Translation-and-scale Invariance</td>
<td>31</td>
</tr>
<tr>
<td>3. INVARIANT ONE-SIDED TESTS OF $H_0 : \lambda = \lambda_0$ AND THE CORRESPONDING CONFIDENCE SETS</td>
<td>35</td>
</tr>
<tr>
<td>3.1. Most Powerful Invariant Test of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda = \lambda_1 (\lambda_1 \neq \lambda_0)$</td>
<td>36</td>
</tr>
<tr>
<td>3.2. Uniformly Most Powerful Invariant Test of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda \geq \lambda_0$</td>
<td>39</td>
</tr>
<tr>
<td>3.3. Locally Most Powerful Invariant Test of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda &gt; \lambda_0$</td>
<td>44</td>
</tr>
<tr>
<td>3.4. Wald's Test of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda &gt; \lambda_0$ as the MPI Test of $H_0$ against $H_A : \lambda = \infty$</td>
<td>52</td>
</tr>
<tr>
<td>4. INVARIANT TWO-SIDED TESTS OF $H_0 : \lambda = \lambda_0$ AND THE CORRESPONDING CONFIDENCE SETS</td>
<td>55</td>
</tr>
<tr>
<td>4.1. Likelihood Ratio Invariant Test of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda \neq \lambda_0$</td>
<td>55</td>
</tr>
</tbody>
</table>
4.1.1. Likelihood ratio invariant test of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda \neq \lambda_0 \) when \( \Lambda = \Lambda_3 \) and \( \Lambda_1 = \Delta \) for all \( i \) 58

4.1.2. Likelihood ratio invariant test of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda \neq \lambda_0 \) when \( \Lambda = \Lambda_1 \) and \( \Lambda_1 = \Delta \) for all \( i \) 60

4.1.3. Likelihood ratio invariant test of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda \neq \lambda_0 \) when \( \Lambda = \Lambda_2 \) and \( \Lambda_1 = \Delta \) for all \( i \) 62

4.2. Most Powerful Invariant Test of \( H_0 : \lambda = \lambda_0 \) against a Mixed Alternative 63

4.3. Other Two-sided Tests Obtained by Combining Two One-sided Tests 65

4.3.1. Combining most powerful invariant tests of \( H_0 : \lambda = \lambda_0 \) against \( H_1 : \lambda = \lambda_1 \) and of \( H_0 : \lambda = \lambda_0 \) against \( H_2 : \lambda = \lambda_2 \) 67

4.3.2. Combining uniformly most powerful invariant tests of \( H_0 : \lambda = \lambda_0 \) against \( H_1 : \lambda < \lambda_0 \) and of \( H_0 : \lambda = \lambda_0 \) against \( H_2 : \lambda > \lambda_0 \) 68

4.3.3. Combining locally most powerful invariant tests of \( H_0 : \lambda = \lambda_0 \) against \( H_1 : \lambda < \lambda_0 \) and of \( H_0 : \lambda = \lambda_0 \) against \( H_2 : \lambda > \lambda_0 \) 69

4.3.4. Combining Wald's tests of \( H_0 : \lambda = \lambda_0 \) against \( H_1 : \lambda < \lambda_0 \) and of \( H_0 : \lambda = \lambda_0 \) against \( H_2 : \lambda > \lambda_0 \) 70

5. OPTIMALITY OF SIMILARITY OF INVARIANT TESTS 73

5.1. Generalized Likelihood Ratio Translation-invariant Similar Test of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda \neq \lambda_0 \) 73

5.2. Most Powerful Translation-invariant Similar Test of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda = \lambda_1 \) (\( \lambda_1 \neq \lambda_0 \)) 76

5.3. A Locally Most Powerful Translation-Invariant Similar Test of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda > \lambda_0 \) 79
6. COMPUTING CRITICAL POINTS AND POWER FUNCTIONS OF INVARIANT TESTS

6.1. Monotonic Power Functions of Invariant Tests of $H_0 : \lambda = \lambda_0$ against One-sided Alternatives

6.1.1. The power function of the most powerful invariant test of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda = \lambda_1 (\lambda_1 > \lambda_0)$

6.1.2. The power function of the locally most powerful invariant test of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda > \lambda_0$

6.1.3. The power function of the Wald's test of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda > \lambda_0$

6.2. Power Functions of Invariant Tests of $H_0 : \lambda = \lambda_0$ against Two-sided Alternatives

6.3. Imhof's Technique for Evaluating the Distribution Function of a Linear Combination of Independent Chi-square Random Variables

6.4. Finding Critical Points of MPI and LMPI Tests

6.4.1. Bracketing the critical point of an MPI test

6.4.2. Bracketing the critical point of an LMPI test

6.5. Approximating Critical Points of LRI and MIXI Tests

6.6. Technique of Computing and Nature of the Confidence Sets

7. NUMERICAL STUDIES

7.1. Data Structures Used in Comparisons of Procedures

7.1.1. One-way classification

7.1.2. Lamb-weight data

7.2. Numerical Results

7.3. Concluding Remarks
1. PRELIMINARY DISCUSSION AND LITERATURE REVIEW

1.1. Introduction

Throughout this dissertation, we denote an n x 1 unity vector by \( \mathbf{1}_n \), an n x n identity matrix by \( \mathbf{I}_n \), an n x 1 null vector by \( \mathbf{0}_n \), and an m x n null matrix by \( \mathbf{0}_{mn} \). The subscripts m and n will be omitted whenever no confusion arises.

Suppose that \( \mathbf{y} \) is an n x 1 observable random vector. We consider the mixed linear model

\[
\mathbf{y} = \mathbf{X}\alpha + \mathbf{Z}\beta + \mathbf{e}
\]  

(1.1)

where

- \( \alpha \) is an a x 1 vector of unknown parameters,
- \( \beta \) is a b x 1 vector of unobservable random variables,
- \( \mathbf{e} \) is an n x 1 vector of unobservable random variables,
- \( \mathbf{X} \) is an n x a given non-random matrix,

and

- \( \mathbf{Z} \) is an n x b given non-random matrix.

It is assumed that \( \beta \) and \( \mathbf{e} \) are independently distributed as MVN\( _b(\mathbf{0}, \sigma^2_\beta \mathbf{I}) \) (b-variate normal with mean vector \( \mathbf{0} \) and variance-covariance matrix \( \sigma^2_\beta \mathbf{I} \)) and MVN\( _n(\mathbf{0}, \sigma^2_\mathbf{e} \mathbf{I}) \), respectively. The variance components \( \sigma^2_\mathbf{e} \) and \( \sigma^2_\beta \) are taken to be unknown parameters satisfying \( \sigma^2_\mathbf{e} \geq 0 \) and \( \sigma^2_\beta \geq 0 \).

The elements of \( \alpha \), \( \beta \), and \( \mathbf{e} \) are referred to as fixed effects, random effects, and random errors, respectively. Two special cases of the mixed linear model are
(1) the fixed-effects model in which \( Z \) is a "degenerate matrix", that is, \( Z \) has only "zero columns", and

(2) the random-effects model in which \( X \) is a unity vector and \( \alpha \) is a scalar.

The parameter space of model (1.1) is

\[
\Omega_1 = \{ (\alpha, \sigma^2, \sigma^2_\beta) : \alpha \in \mathbb{R}^3, \sigma^2 > 0, \sigma^2_\beta \geq 0 \}.
\]

Let \( \lambda = \sigma^2_\beta / \sigma^2_e \). Under model (1.1), it is clear that \( \lambda \geq 0 \) and that

\[
V(y) = \sigma^2_e (I + \lambda ZZ').
\]

In terms of \( \alpha, \sigma^2_e, \) and \( \lambda \), the parameter space of model (1.1) is

\[
\Omega'_1 = \{ (\alpha, \sigma^2_e, \lambda) : \alpha \in \mathbb{R}^3, \sigma^2_e > 0, \lambda \in \Lambda_1 \}
\]

where \( \Lambda_1 = (0, \infty) \).

Let \( \omega \) represent the largest eigenvalue of the matrix \( ZZ' \). The variance-covariance matrix of \( y \) is positive definite if and only if \( \lambda > -1/\omega \) (e.g., Harville and Fenech, 1985). This observation suggests a variation on model (1.1).

Let

\[
y = X\alpha + e^* \tag{1.2}
\]

where \( X \) and \( \alpha \) are as defined in model (1.1) and \( e^* \) is an \( n \times 1 \) unobservable random vector distributed as \( \text{MVN}_n(0, \sigma^2_e(I + \lambda ZZ')) \).

The parameter space of model (1.2) is taken to be

\[
\Omega_2 = \{ (\alpha, \sigma^2_e, \lambda) : \alpha \in \mathbb{R}^3, \sigma^2_e > 0, \lambda \in \Lambda_2 \}
\]

where \( \Lambda_2 = (-1/\omega, \infty) \).

Since \( ZZ' \) is a non-negative definite matrix, \( \omega \geq 0 \). Thus, \( \Omega_1' \)
Accordingly, \( \Omega_2 \) is referred to as an extended parameter space.

In model (1.2), \( \lambda \sigma^2_e \) is not necessarily interpretable as a variance as it is in model (1.1). Model (1.2) is sometimes more appropriate than model (1.1). For example, consider the one-way random model

\[
y_{ij} = \mu + \beta_i + e_{ij}, \quad j = 1, \ldots, n_i; \quad i = 1, \ldots, b \quad (1.3)
\]

where \( \beta_i \sim N(0, \sigma^2_\beta) \) (a normal distribution with mean zero and variance \( \sigma^2_\beta \)), \( e_{ij} \sim N(0, \sigma^2_e) \), the \( \beta \)'s are independent and identically distributed, the \( e \)'s are independent and identically distributed, and the \( \beta \)'s and \( e \)'s are statistically independent.

Model (1.3) is the special case of (1.1) obtained by taking \( n = \sum n_i, \alpha = (\mu), \mathbf{X} = \mathbf{1}_n \), and \( \mathbf{Z} = \text{diag}(1) \), an \( n \times b \) block diagonal matrix with \( \mathbf{1}_{n_i}, i = 1, \ldots, b \) as its diagonal blocks. In the special case where \( n_1 = \ldots = n_b = n_0 \) for some \( n_0 \), model (1.3) is referred to as a balanced one-way random model.

Define the intraclass correlation coefficient \( \rho \) to be the common correlation coefficient between any two members of the same class. That is,

\[
\rho = \frac{\text{Cov}(Y_{ij}, Y_{i'j'})}{\sqrt{\text{V}(Y_{ij}) \cdot \text{V}(Y_{i'j'})}} = \frac{\sigma^2}{\sigma^2_e + \sigma^2_\beta} = \frac{\lambda}{1 + \lambda} \quad (1.4)
\]

where \( j \neq j'; \quad i = 1, \ldots, b \).

It is implicit in model (1.3) that the intraclass correlation coefficient is non-negative. In some applications, a
negative intraclass correlation coefficient is conceivable. If, for example, the classes consist of pens of animals and there is competition for a limited supply of feed, the stronger animals may drive away the weaker and may regularly get more feed. Then, the animals in any particular pen may tend to be less alike than those in different pens, implying that the intraclass correlation coefficient is negative.

In this example, a model more appropriate than (1.3) is

\[ y_{ij} = \mu + e^*_ij, \quad j = 1, \ldots, n_i; \quad i = 1, \ldots, b \quad (1.5) \]

where \( e^*_ij \sim N(0, \sigma_e^2(1+\lambda)) \); \( \text{Cov}(e^*_ij, e^*_ij') = \lambda \sigma_e^2, \quad j \neq j' \) and \( \text{Cov}(e^*_ij, e^*_i'i') = 0, \quad i \neq i' \). It can be checked that the largest eigenvalue of \( ZZ' \) is \( \max\{n_i\} \) so that the parameter space of model (1.5) is \[ \{(\mu, \sigma_e^2, \lambda) : \mu \in \mathcal{R}, \sigma_e^2 > 0, \lambda > -1/\max\{n_i\} \} \, \text{if} \, 1 \leq i \leq b \]

Clearly, model (1.5) is a special case of model (1.2). The intraclass correlation coefficient \( \rho \) is now allowed to assume negative values. From (1.4), it is clear that \( \rho \) is a monotonically increasing function of \( \lambda \).

The concept of the intraclass correlation coefficient, first introduced by Fisher (1925), plays an important role in many areas of application. For example, \( \rho \) is used in investigating the reliability of measurements in psychometric studies (e.g., Winer, 1971), in assessing the repeatability and heritability of various traits in animals (e.g., Kempthorne, 1957), and in evaluating the
extent of familial resemblance in epidemiologic studies (e.g., Donner and Koval, 1980). In animal studies involving paternal half sibs, \( 4p \) may be interpretable as a heritability (e.g., Kempthorne, 1957). Since a heritability is inherently less than or equal to one, there is an implicit assumption that \( 0 \leq 4p \leq 1 \), or, in terms of \( \lambda \), that \( 0 \leq \lambda \leq 1/3 \).

The various models and parameter spaces may be consolidated into the single model

\[
y \sim MVN_n \{X \alpha, \sigma^2_e (I + \lambda ZZ')\} \tag{1.6}
\]

with the parameter space

\[
\Omega = \{ (\alpha, \sigma^2_e, \lambda) : \alpha \in \mathbb{R}^3, \sigma^2_e > 0, \lambda \in \Lambda \} \tag{1.7}
\]

where \( y, X, Z, \alpha, \) and \( \sigma^2_e \) are as defined in model (1.1) and \( \Lambda \) represents a closed, open, or half-open interval (the upper end point of \( \Lambda \) can be infinity). In particular, applications of this model were already discussed for the cases where \( \Lambda \) is restricted to \( \Lambda^1, \Lambda^2 \), or \([0, 1/3]\).

The subject of this dissertation is likelihood-based inference about the parameter \( \lambda \). In particular, we are interested in constructing a confidence set for \( \lambda \). To illustrate the nature of the problem, we consider the "ANOVA method" for testing the null hypothesis \( H_0 : \lambda = 0 \). We first introduce some notation for later reference. We denote the chi-square distribution with degrees of freedom \( a \) by \( \chi^2(a) \), the upper-\( \gamma \) point of \( \chi^2(a) \) by \( \chi^2_{\gamma}(a) \), the F distribution with degrees of freedom \( a \) and \( b \) by \( F(a,b) \), and the upper-\( \gamma \) point of \( F(a,b) \) by \( F_{\gamma}(a,b) \).
For the purpose of testing $H_0 : \lambda = 0$, the following ANOVA table, associated with model (1.1), is relevant.

**Table 1. Analysis of variance for testing $H_0 : \lambda = 0$**

<table>
<thead>
<tr>
<th>Source</th>
<th>D.F.</th>
<th>SS</th>
<th>MS</th>
<th>$E(\text{MS})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>rank($X$)</td>
<td>$SS_\alpha = y'P_x y$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta</td>
<td>\alpha$</td>
<td>$r = \text{rank}(X,Z) - \text{rank}(X)$</td>
<td>$SS_\beta = y'(P_x z - P_x)y$</td>
<td>$MS_\beta = \frac{SS_\beta}{r}$</td>
</tr>
<tr>
<td>Error</td>
<td>$n - \text{rank}(X,Z)$</td>
<td>$SS_e = y'y - SS_\alpha - SS_\beta$</td>
<td>$MS_e = \frac{SS_e}{v}$</td>
<td>$\sigma_e^2$</td>
</tr>
</tbody>
</table>

Here, $P_x = X(X'X)^{-1}X'$, $P_x z = (X,Z)[(X,Z)'(X,Z)]^{-1}(X,Z)'$ (for any matrix $A$, $A^-$ represents an arbitrary generalized inverse of $A$), and $K = \text{tr}(C)/\alpha$ where $C = Z'(I-P_x)Z$. By Lemma 1.(ii) of the Appendix, $\text{rank}(C) = \alpha$. It is well-known that $SS_\beta$ is statistically independent of $SS_\alpha$ and $SS_\alpha/\sigma_e^2 \sim \chi^2(\alpha)$.

Define $F = MS_\beta/MS_e$. Under $H_0 : \lambda = 0$, $SS_\beta/\sigma_e^2(1+\lambda K) \sim \chi^2(\alpha)$ and $F \sim F(\alpha,\nu)$. Thus, a size-$\gamma$ significance test of $H_0$ against the alternative $H_A : \lambda > 0$ is to reject $H_0$ if the observed $F$ is larger than $F_{\alpha}(\alpha,\nu)$ and to accept $H_0$ otherwise. If $\lambda \neq 0$, the distribution of $SS_\beta/\sigma_e^2(1+\lambda K)$ is not, in general, $\chi^2(\alpha)$ but rather is that of the sum of $(1+\lambda \Delta_i)z_i/(1+\lambda K)$, $i = 1, \ldots, \alpha$, where the $z_i$'s are independent $\chi^2(1)$ random variables and the $\Delta_i$'s are the non-zero eigenvalues of the matrix $C$, as we now show.

Let $D_\alpha = \text{diag}(\Delta_1, \ldots, \Delta_\alpha)$, a diagonal matrix with the $\Delta_i$'s as its diagonal elements. By Lemma 2 of the Appendix, there is an
orthogonal matrix \( P_1 = (R_1, U_1) \) such that \( P_1^T C P_1 = \begin{pmatrix} D_r & 0 \\ 0 & 0 \end{pmatrix} \), in which case \( R_1^T C R_1 = D_r \) and \( U_1^T C = 0 \). Therefore,

\[
R_1^T C^2 R_1 = R_1^T C (R_1^T C R_1 + U_1^T U_1^T) C R_1
= D_r D_r
= D_r^2
\]

implying that

\[
V[R_1^T Z'(I - P_x) y] = \sigma_e^2 R_1^T Z'(I - P_x)(I + \lambda ZZ')(I - P_x) Z R_1
= \sigma_e^2 (C + \lambda C^2) R_1
= \sigma_e^2 (D_r + \lambda D_r^2).
\]

Let

\[
t = D_r^{-1/2} R_1^T Z'(I - P_x) y
\]

(1.8)

where \( D_r^{-1/2} = \text{diag} \{ \Delta_1^{-1/2}, \ldots, \Delta_r^{-1/2} \} \). Then, it is clear that

\[
t \sim \text{MVN}_r [0, \sigma_e^2 (I + \lambda D_r)].
\]

(1.9)

Consider now the sum of squares \( t'^T t \). Let \( b \) represent any solution to the equation \( C b = Z'(I - P_x) y \). Then,

\[
t'^T t = y'^T Z R_1 D_r^{-1} R_1^T Z'(I - P_x) y
= b'^T C R_1 D_r^{-1} R_1^T C b.
\]

Note that \( R_1 D_r = R_1 R_1^T C R_1 = (I - U_1 U_1^T) C R_1 = C R_1 \). Thus, we find that

\[
t'^T t = b'^T R_1 D_r D_r^{-1} R_1^T C b
= b'^T R_1^T R_1 C b.
\]
= \mathbf{b}'(I - \mathbf{U}_1'\mathbf{U}_1)\mathbf{C}\mathbf{b} \\
= \mathbf{b}'\mathbf{C}\mathbf{b} \\
= \mathbf{b}'\mathbf{CC}'\mathbf{C}\mathbf{b} \\
= \mathbf{y}'(I - \mathbf{P}_x)\mathbf{Z}[\mathbf{Z}'(I - \mathbf{P}_x)\mathbf{Z}]^{-1}\mathbf{Z}'(I - \mathbf{P}_x)\mathbf{y}.

Clearly, \((I - \mathbf{P}_x)\mathbf{Z}[\mathbf{Z}'(I - \mathbf{P}_x)\mathbf{Z}]^{-1}\mathbf{Z}'(I - \mathbf{P}_x)\) is the unique projection matrix onto the column space \(\mathbf{C}[(I - \mathbf{P}_x)\mathbf{Z}]\). By Lemma 1.(iii) of the Appendix, \((I - \mathbf{P}_x)\mathbf{Z}[\mathbf{Z}'(I - \mathbf{P}_x)\mathbf{Z}]^{-1}\mathbf{Z}'(I - \mathbf{P}_x)\) is the unique projection matrix onto the column space \(\mathbf{C}(\mathbf{P}_x'\mathbf{z} - \mathbf{P}_x)\).

Therefore, \((I - \mathbf{P}_x)\mathbf{Z}[\mathbf{Z}'(I - \mathbf{P}_x)\mathbf{Z}]^{-1}\mathbf{Z}'(I - \mathbf{P}_x) = \mathbf{P}_x'\mathbf{z} - \mathbf{P}_x\). It follows that

\[ t' = \mathbf{y}'(\mathbf{P}_x'\mathbf{z} - \mathbf{P}_x)\mathbf{y} = \mathbf{S}\mathbf{S}'/\mathbf{c}_e^2. \]

As a consequence of (1.9), \(\mathbf{S}\mathbf{S}'/\mathbf{c}_e^2(1+\lambda K)\) is then indeed

\[ \chi^2(r) \]

distributed as \(\sum_{i=1}^{r}(1+\lambda\Delta_i)z_i/(1+\lambda K)\), a linear combination of \(r\) independent one-degree-of-freedom chi-square random variables.

Therefore, the F statistic is not always appropriate for testing the null hypothesis \(H_0: \lambda = \lambda_0\) when \(\lambda_0\) is not zero. One exception is the case where the matrix \(\mathbf{C}\) has only one distinct non-zero eigenvalue.

When \(\Delta_1 = \ldots = \Delta_\lambda = \Delta\) for some \(\Delta, K = \Delta\). Consequently,

\[ \mathbf{S}\mathbf{S}'/\mathbf{c}_e^2(1+\lambda K) \sim \chi^2(r) \quad \text{and} \quad F/(1+\lambda K) \sim F(r,f) \]

for \(\lambda \geq 0\). Thus, a size-\(\gamma\) significance test of \(H_0: \lambda = \lambda_0\) against \(H_A: \lambda > \lambda_0\) for any particular \(\lambda_0 \geq 0\) is to reject \(H_0\) if \(F/(1+\lambda_0\Delta) \geq F(r,f)\) and to accept \(H_0\) otherwise. In Section 3.2, this test will be shown to be the size-\(\gamma\) uniformly most powerful invariant test of \(H_0\).
1.2. Relationship between Confidence Sets and Test of Hypotheses

**Definition 1.1**

$C(y)$ is a $100(1-\gamma)\%$ confidence set for $\lambda$ if, for every $\alpha$, $\sigma^2$, and $\lambda$, $Pr\{ y : \lambda \in C(y); \alpha, \sigma^2, \lambda \} = 1 - \gamma$, that is, $C(y)$ covers the true value $\lambda$ with probability $1 - \gamma$, whatever the unknown true values of $\alpha$ and $\sigma^2$ are.

**Definition 1.2**

A measurable function $\phi(y)$ defined on the sample space of $y$ is a test function if $0 \leq \phi(y) \leq 1$ for every $y$.

A test function $\phi(y)$ is said to be non-randomized if $\phi(y)$ is either zero or one. In this dissertation, any test function will be meant to be a non-randomized test function unless it is stated otherwise.

Corresponding to a $100(1-\gamma)\%$ confidence set $C(y)$, define $A(\lambda) = \{ y : \lambda \in C(y) \}$ for every $\lambda$. It is well-known (e.g., Lehmann, 1959, p. 174) that for every $y$ and $\lambda$,

$$\lambda \in C(y) \text{ if and only if } y \in A(\lambda)$$

and that for each $\lambda_0$, $A(\lambda_0)$ defines an acceptance region of a size-$\gamma$ test of the composite null hypothesis $H_0 : \lambda = \lambda_0$. In other words, let $\phi(y)$ be a test function such that

$$\phi(y) = \begin{cases} 1, & \text{if } y \in A(\lambda_0), \\ 0, & \text{otherwise.} \end{cases}$$
Then, \( \Pr\{ y : \phi(y) = 1; \alpha, \sigma_e^2, \lambda \} = \gamma \) for every \( \alpha \) and \( \sigma_e^2 \).

Thus, a \( 100(1-\gamma)\% \) confidence set \( C(y) \) for \( \lambda \) is a collection of values of \( \lambda \), each of which is not rejected as a hypothesized value of \( \lambda \) by a size-\( \gamma \) test [with the acceptance region \( A(\lambda_0) \)] when \( y \) is observed.

The above argument is reversible. For each \( \lambda_0 \), let \( A(\lambda_0) \) represent the acceptance region of a size-\( \gamma \) test of the null hypothesis \( H_0 : \lambda = \lambda_0 \). Define \( C(y) = \{ \lambda \in \Lambda : y \in A(\lambda) \} \) for every \( y \). Then, (1.10) still holds and \( C(y) \) constitutes a \( 100(1-\gamma)\% \) confidence set for \( \lambda \).

The relationship (1.10) is sometimes referred to as the duality between a confidence set and a family of hypothesis tests.

**Definition 1.3**

The power, as a function of \( \alpha, \sigma_e^2, \) and \( \lambda \), of a test function \( \phi(y) \) of \( H_0 : \lambda = \lambda_0 \) is

\[
\beta_\phi(\alpha, \sigma_e^2, \lambda ; \lambda_0) = \Pr\{ y : \phi(y) = 1; \alpha, \sigma_e^2, \lambda \}.
\]

According to (1.10), for any value \( \lambda_0 \in \Lambda \), we have

\[
\Pr\{ y : \lambda_0 \in C(y); \alpha, \sigma_e^2, \lambda \} = \Pr\{ y : y \in A(\lambda_0); \alpha, \sigma_e^2, \lambda \} = 1 - \beta_\phi(\alpha, \sigma_e^2, \lambda ; \lambda_0)
\]

for every \( \alpha, \sigma_e^2, \) and \( \lambda \). Suppose that \( \lambda_0 \) is a false value. Since it is desirable that the confidence set \( C(y) \) cover the false value with "small" probability for every \( \alpha, \sigma_e^2, \) and \( \lambda \) (Neyman, 1937), the test \( \phi(y) \) of \( H_0 : \lambda = \lambda_0 \) with the acceptance region \( A(\lambda_0) \) is desirable when it rejects the false null hypothesis with "large"
probability for every \( \alpha, \sigma^2_e, \) and \( \lambda, \) that is, when \( \phi(y) \) has large power \( \beta(y; \alpha, \sigma^2_e, \lambda; \lambda_0) \) for every \( \alpha, \sigma^2_e, \) and \( \lambda. \)

It is well known that a size-\( y \) test of \( H_0: \lambda = \lambda_0 \) with uniformly largest probability of rejecting \( H_0 \) for every \( \alpha, \sigma^2_e, \) and \( \lambda \) does not exist in general. Therefore, a \( 100(1 - y) \) confidence set with uniformly smallest probability of covering \( \lambda_0 \) for every \( \alpha, \sigma^2_e, \) and \( \lambda \) is not in general attainable. Alternatively, within a class of invariant tests (see Section 1.3) of \( H_0, \) we consider the most powerful test against \( H_A: \lambda = \lambda_1 (\lambda_1 \neq \lambda_0), \) the locally most powerful test against \( H_A: \lambda > \lambda_0, \) Wald's test against \( H_A: \lambda > \lambda_0, \) and others in Chapters 3 and 4.

1.3. Invariance and Similarity

For purposes of making inferences about \( \lambda \) or functions of \( \lambda \) under model (1.6), the elements of \( \alpha \) and \( \sigma^2_e \) are nuisance parameters. In the presence of nuisance parameters, it is customary to restrict attention to statistical procedures having certain properties. Invariance and (in conjunction with testing hypotheses) similarity are two such properties (Cox and Hinkley, 1974). In particular, most statistical procedures for making inferences about the variance components in mixed linear models are translation invariant (e.g., Harville, 1985).

Let \( (\mathcal{X}, \mathcal{A}, \mathcal{P}_\Theta) \) be any probability space. Here \( \mathcal{X} \) is the sample space, \( \mathcal{A} \) the \( \sigma \)-algebra of subsets of \( \mathcal{X}, \) and \( \mathcal{P}_\Theta \) a probability distribution defined on \( \mathcal{A} \) and parametrized by \( \Theta. \) Let \( x \) and \( \Lambda \)
represent arbitrary elements of $X$ and $\mathcal{A}$, respectively. Both $x$ and $\theta$ could be vector-valued and $\theta$ belongs to a parameter space $\Theta$.

Let $\mathcal{P} = \{ P_\theta : \theta \in \Theta \}$. The notation $E[\phi(x) ; \Theta]$ represent the expectation of a statistic $\phi(x)$ when the distribution function of $x$ is parametrized by $\theta$.

Let $g$ be a one-to-one transformation from $X$ onto itself, and let $G$ be a collection of such transformations. Define a transformed $\sigma$-algebra by $\mathcal{A}_g = \{ gA : A \in \mathcal{A} \}$. Corresponding to a distribution $P_\theta$, an induced distribution $P_{g_\theta}$ on $\mathcal{A}_g$ may be defined by $P_{g_\theta}(gA) = P_\theta(A)$ for every $A \in \mathcal{A}$. $g_\theta$ is regarded as a transformation of $\theta$ induced by $g$. The collection of such induced transformations $g$ on $\Theta$ is denoted by $\mathcal{G}$.

**Definition 1.4**

The parameter space $\Theta$ is said to be invariant with respect to $\mathcal{G}$ if $\mathcal{G} \Theta \subseteq \Theta$ for all $\theta \in \Theta$ and if for every $\theta' \in \Theta$, there is a $\theta \in \Theta$ such that $\theta' = g_\theta$.

Note that $\Theta$ is invariant with respect to $\mathcal{G}$ if and only if $\mathcal{G} \Theta = \Theta$. We also say that $g$ leaves $\Theta$ invariant if $g_\Theta = \Theta$.

**Definition 1.5**

The parameter space $\Theta$ is said to be invariant with respect to $\mathcal{G}$ if $\mathcal{G} \Theta = \Theta$ for every $g \in \mathcal{G}$.

**Lemma 1.1**

Let $g_1$ and $g_2$ be two one-to-one transformations from $X$ onto itself which leave $\Theta$ invariant. Then, the transformations $g_1 \circ g_2$ and
\(g^{-1}\) defined by \((g_1 \circ g_2)(x) = g_1(g_2(x))\) and \(g_1(g^{-1}_1(x)) = x\) for all \(x \in X\) also leave \(\Theta\) invariant. The induced transformations \(g_1, g_2, g_1^{-1} \circ g_2,\) \(g_1 \circ g_2,\) and \(g_1^{-1}\) satisfy \(g_1 \circ g_2 = g_1 \circ g_2\) and \(g_1^{-1} \circ g_1^{-1} = (g_1^{-1})^{-1}\).

The proof of Lemma 1.1 is given, e.g., by Lehmann (1959). According to Lemma 1.1, any set of transformations which leaves \(\Theta\) invariant can always be extended to a group of transformations which leaves \(\Theta\) invariant. Subsequently, \(G\) represents a group of one-to-one transformations from \(X\) onto itself; and \(\tilde{G}\) represents the group of transformations from \(\Theta\) onto \(\Theta\) induced by the elements of \(G\).

**Definition 1.6**

Let \(\Theta_0\) and \(\Theta_A\) be two disjoint subsets of \(\Theta\). The problem of testing the null hypothesis \(H_0: \Theta \subseteq \Theta_0\) against the alternative hypothesis \(H_A: \Theta \subseteq \Theta_A\) is said to be invariant with respect to \(G\) if \(\tilde{g}\Theta_0 = \Theta_0\) and \(\tilde{g}\Theta_A = \Theta_A\) for all \(\tilde{g} \in \tilde{G}\).

If a problem of testing is invariant with respect to \(G\), then it may be reasonable to restrict attention to test functions that are invariant with respect to \(G\) in the following sense:

**Definition 1.7**

A function \(T(x)\) defined on \(X\) is said to be invariant with respect to \(G\) if \(T(g(x)) = T(x)\) for all \(x \in X\) and all \(g \in G\).

We now introduce the concept of a maximal invariant and use the concept to characterize the totality of tests which are invariant with respect to \(G\).
Definition 1.8

With respect to $G$, an invariant function $T(x)$ defined on $X$ is said to be a maximal invariant if $T(x_1) = T(x_2)$ implies that $x_1 = g(x_2)$ for some $g \in G$.

Lemma 1.2

Let $T(x)$ be a maximal invariant with respect to $G$. A necessary and sufficient condition for a statistic $\phi(x)$ to be invariant with respect to $G$ is that $\phi(x)$ depends on $x$ only through $T(x)$.

For a proof of Lemma 1.2, refer, for example, to Lehmann (1959).

Definition 1.9

A function $\nu(\theta)$ defined on $\Theta$ is said to be invariant with respect to $\bar{G}$ if $\nu(\theta) = \nu(g\theta)$ for all $\theta \in \Theta$ and all $g \in \bar{G}$.

Definition 1.10

With respect to $\bar{G}$, an invariant function $\nu(\theta)$ defined on $\Theta$ is said to be a maximal invariant if $\nu(\theta_1) = \nu(\theta_2)$ implies that $\theta_1 = g\theta_2$ for some $g \in \bar{G}$.

Lemma 1.3

Let $\nu(\theta)$ be a maximal invariant with respect to $\bar{G}$. A necessary and sufficient condition for a parametric function $\mu(\theta)$ to be invariant with respect to $\bar{G}$ is that $\mu(\theta)$ depends on $\theta$ only through $\nu(\theta)$.
The proof of Lemma 1.3 is analogous to that of Lemma 1.2.

**Corollary 1.1**

Suppose that the elements of $\mathcal{P}$ are distinct in the sense that $\theta_1 \neq \theta_2$ implies that $P_{\theta_1} \neq P_{\theta_2}$. Also suppose that the function $T(x)$ defined on $\mathcal{X}$ is invariant with respect to $G$ and that the function $\nu(\theta)$ defined on $\mathfrak{G}$ is a maximal invariant with respect to $\tilde{G}$. Then, the distribution of $T(x)$ depends on $\theta$ only through $\nu(\theta)$.

**Proof**

Let $(\mathcal{S}, \mathcal{B}, P^T_{\mu(\theta)})$ be the probability space induced by the transformation $T$ on $(\mathcal{X}, \mathcal{S}, P_{\theta})$ satisfying $P^T_{\mu(\theta)}(B) = P_{\theta}(T^{-1}(B))$ for any $B \in \mathcal{B}$, where $T^{-1}$ denotes the inverse image of $T$.

Then, for any $B \in \mathcal{B}$ and $g \in G$,

\[ P_{\theta}(T^{-1}(B)) = P_{\tilde{g}\theta}(g(T^{-1}(B))) \]

\[ = P^T_{\mu(\tilde{g}\theta)}(T(g(T^{-1}(B)))) \]

\[ = P^T_{\mu(\tilde{g}\theta)}(T(T^{-1}(B))) \]

\[ = P^T_{\mu(\tilde{g}\theta)}(B) \]

\[ = P_{\tilde{g}\theta}(T^{-1}(B)). \]

Since the elements of $\mathcal{P}$ are distinct, $\theta = \tilde{g}\theta$. Thus, $\mu(\theta) = \mu(\tilde{g}\theta)$. By Lemma 1.3, $\mu(\theta)$ depends on $\theta$ only through $\nu(\theta)$.

Q.E.D.

We now review some results on similar tests.

Let $\theta' = (\theta'_s, \theta'_t)$ where the elements of $\theta'_s$ consist of the
parameters of interest while those of $\Theta_t$ are nuisance parameters.

A null hypothesis of the form $H_0: \Theta_s = \Theta_{so}$ for some $\Theta_{so}$ hypothesizes that the vector $\Theta_s$ is equal to the vector $\Theta_{so}$ elementwise.

**Definition 1.11**

In testing $H_0: \Theta_s = \Theta_{so}$, $W_\gamma \in \mathfrak{A}$ is said to be a size-$\gamma$ similar critical region if $Pr\{ x: x \in W_\gamma; \Theta_{so}, \Theta_t \} = \gamma$ for every value of $\Theta_t$.

**Definition 1.12**

In testing $H_0: \Theta_s = \Theta_{so}$, a test $\phi(x)$ is said to be a size-$\gamma$ similar test if it is based on a size-$\gamma$ critical region $W_\gamma$, that is, if

$$
\phi(x) = \begin{cases} 
1, & \text{if } x \in W_\gamma, \\
0, & \text{otherwise}.
\end{cases}
$$

Consider now the problem of characterizing the class of size-$\gamma$ similar tests of $H_0: \Theta_s = \Theta_{so}$ in the special case where there exists a set of sufficient statistics for $\mathcal{P}$ under $H_0$ such that the family of distributions of sufficient statistics is boundedly complete. Let $S(x)$ be a set of sufficient statistics for $\mathcal{P}$ when $\Theta_s = \Theta_{so}$.

**Definition 1.13**

In testing $H_0: \Theta_s = \Theta_{so}$, a critical region $W_\gamma \in \mathfrak{A}$ is said to have Neyman structure of size $\gamma$ with respect to $S(x)$ if the conditional probability $Pr\{ x: x \in W_\gamma | S(x); \Theta_{so} \}$ equals $\gamma$ for
every value of $S(x)$ except for a set of zero probability.

The following lemma is an almost immediate consequence of Definitions 1.11 - 1.13.

Lemma 1.4

If a critical region for testing $H_0 : \theta = \theta_0$ has Neyman structure of size $\gamma$ with respect to $S(x)$, then the test based on the critical region is a size-$\gamma$ similar test.

The following lemma gives a condition under which the converse of Lemma 1.4 is true, that is, under which Neyman structure with respect to $S(x)$ is a necessary condition for a test to be similar.

Lemma 1.5

If the family of distributions of $S(x)$ is boundedly complete under $H_0$, then the critical region of a size-$\gamma$ similar test of $H_0$ has Neyman structure of size $\gamma$ with respect to $S(x)$.

Proof

Since $\gamma = E[\phi(x);\theta_0,\theta]$ = $E[E[\phi(x)|S(x)];\theta_0]$, the bounded completeness of $S(x)$ implies that $E[\phi(x)|S(x)] = \gamma$ for every value of $S(x)$ except for a set of zero probability.

Q.E.D.

Therefore, when $S(x)$ is sufficient and boundedly complete under $H_0$, we may concentrate on the class of tests with Neyman structure of size $\gamma$ with respect to $S(x)$ in order to construct "optimal" size-$\gamma$ similar tests. In particular, we form tests of $H_0$ based on the conditional distribution of $x$ given $S(x)$ in such a
way that the probability of rejecting $H_0$ when $H_0$ is true is $\gamma$ for almost every value of $S(x)$. Notice that since $S(x)$ is sufficient for $\theta_t$ under $H_0$, the original problem of testing a composite null hypothesis $H_0 : \theta_S = \theta_{so}$ with $\theta_t$ unspecified now reduces to a problem of testing a single null hypothesis $H'_0 : \theta_S = \theta_{so}$ with no nuisance parameter $\theta_t$.

There is an important difference between invariant tests and similar tests of $H : \theta = \theta_0$. Suppose that the problem of testing is invariant with respect to a group of transformations $G$ and a maximal invariant with respect to $G$, whose distribution does not depend on $\theta_t$, exists. Then, an invariant test, depending on the data only through the maximal invariant, is free of nuisance parameters in the sense that its power function does not depend on $\theta_t$ under either $H_0$ or $H_A$. On the other hand, the power function of a similar test may depend on $\theta_t$ under $H_A$. Nevertheless, similar tests are often found to be invariant.

1.4. Literature Review

Wald (1940) gave a confidence interval for $\lambda$ in a one-way random model (1.3) with unequal class frequencies. Subsequently, Wald (1947) generalized the idea in his previous paper to the special case of linear model (1.1) where the model matrix $(X,Z)$ has full column rank.

Corresponding to his confidence interval are significance tests of null hypotheses of the form $H_0 : \lambda = \lambda_0$ (where $\lambda_0 \geq 0$),
including the null hypothesis \( H_0 : \lambda = 0 \), or equivalently, \( H_0 : \sigma^2_{\beta} = 0 \).

Wald's procedure was extended by Seely and El-Bassiouni (1983) to cover the cases in which the model matrix \((X,Z)\) might not have full rank. The extended version of Wald's procedure is applicable whenever \( r = \text{rank}(X,Z) - \text{rank}(X) \) and \( f = n - \text{rank}(X,Z) \) are positive. Later, Harville and Fenech (1985) gave a more computationally oriented description of the extended Wald's procedure. We now describe this procedure.

For any matrix \( A \), let \( C(A) \) represent the column space of \( A \) and let \( C(A)\) represent the orthogonal complement of \( C(A) \). Let \( L \) and \( F \) be \( n \times r \) and \( n \times f \) matrices whose columns form orthonormal bases for \( C[(I - P_x)Z] \) and \( C(X,Z)\) respectively. The choice of \( L \) and \( F \) is not unique. To see this, let \( O_1 \) and \( O_2 \) be \( r \times r \) and \( f \times f \) orthogonal matrices. Then, \( LO_1 \) and \( FO_2 \) are \( n \times r \) and \( n \times f \) matrices whose columns form orthonormal bases for \( C[(I - P_x)Z] \) and \( C(X,Z)\) respectively.

As indicated by Seely and El-Bassiouni,

\[
L' y \sim \text{MVN}_r[0, \sigma^2_e(I + \lambda L' ZZ' L)] \tag{1.12}
\]

and

\[
F' y \sim \text{MVN}_f[0, \sigma^2_e I]. \tag{1.13}
\]

Further, since both \( Z \) and \( L \) are orthogonal to \( F \),

\[
\text{Cov}(L' y, F' y) = \sigma^2_e L' (I + \lambda ZZ') F
\]
implying that $L'\ y$ and $F'\ y$ are statistically independent.

Define

$$Q(\lambda) = \frac{y' L (I + \lambda L' Z Z' L)^{-1} L' y}{y' F F' y}.$$  \hspace{1cm} (1.14)

By (1.12), (1.13), and (1.14), $Q(\lambda)$ is distributed as $F(r, f)$. Since its distribution does not depend on $\lambda$, $Q(\lambda)$ can serve as a pivotal quantity. It is easy to verify that $Q(\lambda)$ is invariant to the choice of $L$ and $F$.

It can be seen that $t = \frac{D^{-1/2}_r R'_1 Z' (I - P^_x)y}{P^_y}$ of (1.8) is one choice of $L'y$, in which case $L' Z Z' L$ is equal to $D^_r$. Furthermore, there is an orthogonal matrix $P_2 = (R_2, U_2)$ such that $P'_2(I - P^'_x z)P_2 = \begin{bmatrix} I_f & 0 \\ 0 & 0 \end{bmatrix}$, in which case $R'_2(I - P^'_x z)R_2 = I_f$ and $U'_2(I - P^'_x z) = 0$, as can be seen by applying Lemma 2 of the Appendix.

Let

$$u = R'_2(I - P^'_x z)y.$$  \hspace{1cm} (1.15)

It is easy to see that $u$ is one choice for $F'y$. Note that $u \sim \text{MVN}_f [0, \sigma^2_e I]$. In light of (1.15), $u'u$ can be expressed in terms of $y$ as follows:

$$u'u = y'(I - P^'_x z)R'_2 R_2(I - P^'_x z)y$$

$$= y'(I - P^'_x z)(I - U_2 U'_2)(I - P^'_x z)y$$

$$= y'(I - P^'_x z)y.$$  

Note that $y'(I - P^'_x z)y$ is the residual sum of squares obtained by regressing $y$ on $(X, Z)$. Thus, $SS_e = u'u$. Subsequently we take $S_e$
to be the positive square root of \( u'u \).

Let \( t_1, \ldots, t_r \) represent the elements of \( t \). Then, the pivotal quantity \( Q(\lambda) \) can be written as

\[
Q(\lambda) = \sum_{1}^{r} \frac{t_i^2}{1 + \lambda \Delta_i} \cdot \frac{1}{SS_e} \cdot \frac{f}{x}.
\]  

Clearly, \( Q(\lambda) \) is strictly decreasing in \( \lambda \).

Therefore, a significance test of \( H_0 : \lambda = \lambda_0 \) is to reject \( H_0 \) if the observed \( Q(\lambda_0) \) is (1) too large, (2) too small, (3) either too large or too small when testing against (1) \( H_A : \lambda > \lambda_0 \), (2) \( H_A : \lambda < \lambda_0 \) and (3) \( H_A : \lambda \neq \lambda_0 \), respectively.

Accordingly, a confidence interval for \( \lambda \) can be constructed from any of these three types of tests -- refer to the discussion of Section 1.2.

Thompson (1955b) discussed Wald's method as applied to linked incomplete block designs and partially linked block designs. More detailed information about Wald's procedure will be given in Section 3.4 and Subsection 4.3.4.

It is shown in Chapter 2 that \((1/Se)t\) is a set of sufficient statistics for the distribution of a maximal invariant with respect to a group of locational and scale transformations. The statistic is also shown to be a maximal invariant with respect to a group of orthogonal and scale transformations defined on the sample space of \((t,u)\), a maximal invariant with respect to a group of locational transformations on the sample space of \(y\).

Even though the test statistic \( Q(\lambda) \) of Wald's test is a
function of such maximal invariant statistics, the properties of Wald's test relative to other invariant tests did not seem to have received much attention. In this dissertation, Wald's test is compared with the uniformly most powerful invariant (UMPI) test, the most powerful invariant (MPI) test, the locally most powerful invariant (LMPI) test, and the likelihood ratio invariant (LRI) test.

We now summarize briefly previous work on optimal invariant tests. For the one-way random model (1.3) with balanced data, Herbach (1959, Section 9) showed that the uniformly most powerful translation and scale invariant test for testing $H_0 : \sigma^2 = 0$ against $H_A : \sigma^2 > 0$, or equivalently, for testing $H_0 : \lambda = 0$ against $H_A : \lambda > 0$, exists and is the "usual" F-test. Lehmann (1959, Section 7.7) extended this result to the problem of testing $H_0 : \lambda = \lambda_0$ against $H_A : \lambda > \lambda_0$ (where $\lambda_0 \geq 0$). It is shown in Section 3.2 that the UMPI test for testing $H_0 : \lambda = \lambda_0$ against $H_A : \lambda > \lambda_0$ exists when the model matrix $(X,Z)$ is such that $Z'(I - P_X)Z$ has only one distinct non-zero eigenvalue. Herbach (1959, Section 7) derived the likelihood ratio test of $H_0 : \sigma^2 = 0$ against $H_A : \sigma^2 > 0$ (or equivalently, of $H_0 : \lambda = 0$ against $H_A : \lambda > 0$) for the balanced one-way random model.

Thompson (1955a) derived a maximin translation and scale invariant test of $H_0 : \lambda \leq \lambda_0$ against $H_A : \lambda > \lambda_1$ (where $\lambda_1 > \lambda_0 > 0$) for a general incomplete block model, i.e., a model of (1.1) with every entry of $(X,Z)$ being either one or zero. Spjøtvoll
(1967) derived the most powerful translation-invariant similar test of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda = \lambda_1 \) for model (1.3). This test was shown to be the most powerful translation, orthogonal, and scale invariant test. Spjøtvoll pointed out that Wald's test of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda > \lambda_0 \) for model (1.3) is the most powerful translation-invariant similar test of \( H_0 : \lambda = \lambda_0 \) against the alternative "\( \lambda = \infty \)."

At the final stage of writing this dissertation, it came to the author's attention that LaMotte, McWhorter, and Prasad (1987) compared the most powerful invariant test with the so-called LM test (LaMotte and McWhorter, 1978), of which Wald's test is a special case. The comparison was based on several unbalanced one-way random models and made in terms of (a) the power function of the tests and (b) the average length of confidence intervals corresponding to the tests in a simulation study. It was suggested that, in a very unbalanced model, the LM procedure can perform poorly.
2. APPLYING THE PRINCIPLES OF INVARIANCE AND SUFFICIENCY

2.1. A Reduction of y by Translation Invariance

The subsequent discussion applies to the general model (1.6) with the parameter space \( \Omega \) of (1.7).

Define a group of locational transformations on the sample space of \( y \) by

\[
G_1 = \{ g_\delta : y \rightarrow y + x\delta, \delta \in \mathcal{R}^3 \}.
\]

The induced group of transformations on the parameter space \( \Omega \) is

\[
\overline{G}_1 = \{ \overline{g}_\delta : (\alpha, \sigma^2, \lambda) \rightarrow (\alpha + \delta, \sigma^2, \lambda), \delta \in \mathcal{R}^3 \}.
\]

Clearly, the parameter space \( \Omega \) is invariant with respect to \( \overline{G}_1 \).

Let \( \Lambda_0 \) and \( \Lambda_A \) represent two disjoint subsets of \( \Lambda \). Consider the problem of testing a composite null hypothesis of the form \( H_0 : \lambda \in \Lambda_0 \) against a composite alternative hypothesis \( H_A : \lambda \in \Lambda_A \).

Let \( \Omega_0 \) and \( \Omega_A \) represent the subspaces of \( \Omega \) under the null and the alternative hypotheses, respectively.

Since the transformation \( \overline{g}_\delta \) does not affect \( \lambda \), both \( \Omega_0 \) and \( \Omega_A \) are invariant with respect to \( \overline{G}_1 \). In other words, the problem is invariant with respect to \( G_1 \).

For purpose of characterizing the class of translation invariant tests, we introduce the concept of an error contrast.

**Definition 2.1**

Under the model (1.6), a linear combination \( a'y \) of the elements of \( y \) is said to be an error contrast if \( E(a'y) = 0 \) for all \( \alpha, \sigma^2, \text{ and } \lambda \), i.e., if \( a'X = 0 \).
It can be shown that $a'y$ is an error contrast if and only if $a$ belongs to the space $C(I - P_x)$ which is known as the error space. Clearly, the elements of the vector $(I - P_x)y$ are error contrasts. The elements of $(I - P_x)y$ form a maximal invariant with respect to $G_1$. To see this, observe that:

1. $(I - P_x)g_d(y) = (I - P_x)(y + X_6) = (I - P_x)y$ for every $y$ and every $\delta \in \mathbb{R}^a$; and
2. if $(I - P_x)y_1 = (I - P_x)y_2$ for two vectors $y_1$ and $y_2$, then $(I - P_x)(y_1 - y_2) = 0$, implying that $y_1 - y_2 \in C(P_x)$, i.e., that $y_1 = y_2 + X_6$ for some $\delta \in \mathbb{R}^a$.

More generally, we have the following result.

**Lemma 2.1**

Let $Ay$ be a vector of error contrasts, $r + f$ of which are linearly independent. Then, $Ay$ is a maximal invariant with respect to $G_1$.

**Proof**

1. Invariance: According to the Definition 2.1, $AX = 0$, implying that $A(y + X_6) = Ay$ for every $y$ and every $\delta \in \mathbb{R}^a$.

2. Maximal invariance: Since $\text{rank}(X) = n - (r + f)$, $C(A')$ is the orthogonal complement of $C(X)$ with respect to $\mathbb{R}^n$. Thus, for any vectors $y_1$ and $y_2$ that satisfying $Ay_1 = Ay_2$, or equivalently, $A(y_1 - y_2) = 0$, $y_1 - y_2 \in C(X)$. That is, $y_1 = y_2 + X_6$ for some $\delta \in \mathbb{R}^a$.

Q.E.D.

The following lemma will prove useful.
Lemma 2.2

$C(I - P_x, z)$ and $C((I - P_x)Z)$ are orthogonal complements with respect to $C(I - P_x)$.

Proof

Since $(I - P_x, z) = (I - P_x)(I - P_x, z)$ and $P_x, z - P_x = (I - P_x)(P_x, z - P_x)$, $C(I - P_x, z) \subseteq C(I - P_x)$ and $C(P_x, z - P_x) \subseteq C(I - P_x)$. Moreover, since $(I - P_x, z)(P_x, z - P_x) = 0$, $C(I - P_x, z)$ is orthogonal to $C(P_x, z - P_x)$. Now making use of Lemma 1.iii) of the Appendix,

$$\text{rank}(I - P_x) = \text{rank}(I - P_x, z + P_x, z - P_x)$$
$$= \text{rank}(I - P_x) + \text{rank}(P_x, z - P_x)$$
$$= \text{rank}(I - P_x) + \text{rank}((I - P_x)Z).$$

Q.E.D.

Since $\text{rank}(I - P_x, z) = f$ and $\text{rank}((I - P_x)Z) = r$, we may extract a set of $f$ linearly independent vectors from $C(I - P_x, z)$ and $r$ linearly independent vectors from $C((I - P_x)Z)$ to form a set of $r + f$ linearly independent error contrasts. In fact, one sees easily that the $f$ components of $u$ of (1.15) and the $r$ components of $t$ of (1.8) together form such a set of error contrasts.

The location parameters are eliminated by the principle of translation invariance as the distribution of the maximal invariant $(t, u)$ is free of $\alpha$.

Note that the normal distribution of $t$ has a positive definite variance-covariance matrix for all $\lambda$ such that $\lambda > -1/\Delta^*$ where $\Delta^* = \max\{\Delta_1, \ldots, \Delta_r\}$. According to Callanan (1985), $-1/\Delta^*$
Let $\Lambda_3 = (-1/\Delta^*,\infty)$.

It is straightforward to verify that the probability density function (p.d.f.) of $(t,u)$ is

$$g(t,u;\sigma_e^2,\lambda) = \frac{\exp\left(\frac{1}{2\sigma_e^2} \sum_1^r \frac{t_1^2}{1 + \lambda \Delta_1} \right) \exp\left(\frac{1}{2\sigma_e^2} \sum_1^f u_j^2\right)}{(2\pi\sigma_e^2)^{r+f/2} \prod_1^r (1 + \lambda \Delta_1)^{s/2} (2\pi\sigma_e^2)^{f/2}} \frac{\exp\left(-\frac{1}{2\sigma_e^2} \left[\sum_1^r \frac{t_1^2}{1 + \lambda \Delta_1} + \sum_1^f u_j^2\right]\right)}{(2\pi\sigma_e^2)^{(r+f)/2} \prod_1^r (1 + \lambda \Delta_1)^{s/2}}$$

(2.1)

The function $g(\cdot;\sigma_e^2,\lambda)$ is a legitimate p.d.f. for all values of $\sigma_e^2$ and $\lambda$ in the set $\{(\sigma_e^2,\lambda) : \sigma_e^2 > 0, \lambda \in \Lambda_3\}$. Subsequently, let

$\mathcal{W} = \{(\sigma_e^2,\lambda) : \sigma_e^2 > 0, \lambda \in \Lambda\}$

(2.2)

where $\Lambda$ stands for any closed, open, or half-open interval with the upper point being possibly infinity.

2.2. A Reduction of $(t,u)$ by Orthogonal Invariance

Define a group of orthogonal transformations on the sample space of $(t,u)$ by

$$G_2 = \{ g_{A} : (t,u) \rightarrow (t,Au), \text{ where } A \text{ is orthogonal}\}.$$

Recall that $u \sim \text{MVN}(O,\sigma_e^2 I)$. Thus, $Au \sim \text{MVN}(O,\sigma_e^2 I)$.

Furthermore, since $t$ and $u$ are statistically independent, $(t,Au)$ is distributed the same as $(t,u)$. In other words, the induced
group of transformations on the parameter space \( \bar{\mathcal{G}} \) is

\[
\bar{G}_2 = \{ \bar{g}_A : (\sigma^2_\varepsilon, \lambda) \rightarrow (\sigma^2_\varepsilon, \lambda) \text{ where } A \text{ is orthogonal} \}.
\]

That is, \( \bar{g} \) is invariant with respect to \( \bar{G}_2 \). Moreover, it is clear that the problem of testing the composite null hypothesis \( H_0 : \lambda \in \Lambda_0 \) against \( H_A : \lambda \in \Lambda_A \) is invariant with respect to \( G_2 \).

The set of statistics \( S_2(t, u) = (t, uu' \mid u) = (t, SS_\varepsilon) \) is a maximal invariant with respect to \( G_2 \). To prove this, we require the following lemma.

**Lemma 2.3**

Let \( a \) and \( b \) be any two \( n \times 1 \) vectors. Then \( a'a = b'b \) implies that \( a = Ab \) for some orthogonal matrix \( A \).

**Proof**

The proof is trivial when one of \( a \) and \( b \) is null.

Suppose that \( a \) and \( b \) are non-null. Let \( c = a'a = b'b \). Note that \( aa' \) and \( bb' \) are \( n \times n \) symmetric matrices and that \( \text{rank}(aa') = \text{rank}(bb') = 1 \). Therefore, the only non-zero eigenvalue of \( aa' \) is \( a'a \).

By Lemma 2 of the Appendix, there is an orthogonal matrix \( P \) such that \( P'aa'P = \begin{bmatrix} c & 0' \\ 0 & 0 \end{bmatrix} \). Similarly, there is an orthogonal matrix \( Q \) such that \( Q'bb'Q = \begin{bmatrix} c & 0' \\ 0 & 0 \end{bmatrix} \). Clearly, \( P'aa'P = Q'bb'Q \), implying that

\[
aa' = PQ'bb'QP'
\]

\[= dd'\]
where \( d = PQ'b \).

Thus, either \( a = d \) or \( a = -d \). In either case, it is clear that \( a = Ab \) for some orthogonal matrix \( A \).

Q.E.D.

Now, \( S_2(t,u) \) is a maximal invariant with respect to \( G_2 \) because (1) \( S_2(g(t,u)) = S_2(t,Au) = (t,u'A' Au) = (t,u'u) = S_2(t,u) \) for every \( (t,u) \) and every \( g \in G_2 \); and (2) if \( S_2(t_1,u_1) = S_2(t_2,u_2) \), then \( t_1 = t_2 \) and \( u_1'A' u_1 = u_2'A' u_2 \), implying that \( (t_1,u_1) = (t_2,Au_2) \) for some orthogonal matrix \( A \).

Since \( t \) and \( u \) are statistically independent, so are \( t \) and \( SS_e \). Note that \( SS_e/\sigma_e^2 \sim \chi^2(f) \) since \( u \sim \text{MVN}_f(0,\sigma_e^2I) \). Hence, the p.d.f. of \( (t,SS_e) \) is

\[
h(t,SS_e;\sigma_e^2,\lambda) = \exp\left\{ -\frac{1}{2} t' [\sigma_e^2(I + \lambda D_\tau)]^{-1} t \right\} \frac{\exp\left\{ -\frac{SS_e}{2\sigma_e^2} \right\}}{(2\pi)^{f/2} |\sigma_e^2(I + \lambda D_\tau)|^{f/2} \Gamma(f/2)(2\sigma_e^2)^{f/2}}
\]

(2.3)

2.3. A Reduction of \( (t,SS_e) \) by Scale Invariance

To eliminate the nuisance parameter \( \sigma_e^2 \), we define a group of scale transformations \( G_3 \) on the sample space of \( (t,SS_e) \) by
Let \( S(t, SS_e) = x \) where \( x \) is the \( r \times 1 \) vector whose \( i \)th element is \( x_i = t_i/S_e \), \( i = 1, \ldots, r \). The function \( S_3(t, SS_e) \) is a maximal invariant with respect to \( G_3 \) since (1) \( S_3(g_c(t, SS_e)) = S_3(ct, c^2 SS_e) = S_3(t, SS_e) \); and (2) if \( S_3(t_{1i}, SS_{e,1}) = S_3(t_{2i}, SS_{e,2}) \) for some \( t_{1i}, t_{2i}, SS_{e,1}, \) and \( SS_{e,2} \) then \( t_{1i}/S_{e,1} = t_{2i}/S_{e,2} \) for \( i = 1, \ldots, r \), where \( t_{ki} \) is the \( i \)th element of \( t_k \) and \( S_{e,k} \) is the positive square root of \( SS_{e,k} \), \( k = 1, 2 \), implying that \((t_{1i}, SS_{e,1}) = g_c(t_{2i}, SS_{e,2}) \) for \( c = S_{e,1}/S_{e,2} \). Consequently, it follows from Theorem 2 of Lehmann (1959, Chapter 6) that \( x \), when regarded as a function of \( t \) and \( u \), is a maximal invariant with respect to the following group of transformations

\[
G^* = \{ g_{c,A} : (t, u) \rightarrow (ct, cAu), \text{ where } c > 0 \text{ and } A \text{ is orthogonal} \}.
\]
the joint distribution of \( q_i = \frac{t_i}{(1 + \lambda \Delta_i)^{1/2}} \) for \( i = 1, \ldots, r \), is that of a multivariate-t distribution with p.d.f.

\[
h(q_1, \ldots, q_r) = \frac{\Gamma((r+f)/2)}{(\pi f)^{r/2} \Gamma(f/2)} \left(1 + \frac{\sum_{i=1}^{r} q_i^2}{f}\right)^{-\frac{r + f}{2}}
\]

(Johnson and Kotz, 1972) for \((q_1, \ldots, q_r)' \in \mathbb{R}^r\). Thus, the p.d.f. of \( x \) is

\[
k(x; \lambda) = \frac{\Gamma((r+f)/2)}{(\pi f)^{r/2} \Gamma(f/2)} \left(1 + \frac{x^2}{\sum_{i=1}^{r} (1 + \lambda \Delta_i)}\right)^{-\frac{r + f}{2}} \prod_{i=1}^{r} (1 + \lambda \Delta_i)^{1/2}
\]

for \( x \in \mathbb{R}^r \). The function \( k(\cdot; \lambda) \) is a legitimate p.d.f. for all values of \( \lambda \) in the set \( \Lambda_3 = (-1/\Delta^*, \infty) \).

2.4. A Reduction of \( y \) by Sufficiency and Translation-and-scale Invariance

In Sections 2.1 - 2.3, \( x \) was shown to be a maximal invariant with respect to a group of orthogonal and scale transformations on the sample space of \((t,u)\). In this section, \( x \) is first shown to be invariant with respect to a group of locational and scale transformations on the sample space of a set of sufficient statistics for the distribution of \( y \). Then, as a result of applying the Stein Theorem (Hall, Wijisman, and Ghosh, 1965, p. 607), \( x \) is a set of sufficient statistics for the distribution of
a maximal invariant with respect to a group of locational and scale transformations on the sample space of $y$.

Partition the vector space $\mathbb{R}^n$ into three orthogonal subspaces $C(X)$, $C((I - P_X)Z)$, and $C(I - P_{X,z})$. These three subspaces have ranks $n - (r + f)$, $r$, and $f$, respectively. Also, the columns of $(I - P_X)ZR_1$ form a basis for $C((I - P_X)Z)$, while the columns of $(I - P_{X,z})R_2$ form a basis for $C(I - P_{X,z})$ (Sections 1.1 and 1.4).

Now, let $R_3$ represent an $n \times (n - r - f)$ matrix such that the columns of $XR_3$ form a basis for $C(X)$. Define the following non-singular matrix

$$Q = \begin{bmatrix} R_3'X' \\ R_1'Z'(I - P_X) \\ R_2'(I - P_{X,z}) \end{bmatrix}.$$ 

Since it is a linear transformation of $y$, the distribution of $Qy$ is multivariate normal. Its mean vector and variance-covariance matrix are

$$E(Qy) = \begin{bmatrix} R_3'X'X\alpha \\ 0 \\ 0 \end{bmatrix},$$

and

$$V(Qy) = \sigma_e^2 Q(I + \lambda ZZ')Q'.$$

$$= \sigma_e^2 \begin{bmatrix} R_3'X'(I + \lambda ZZ')XR_3 & \lambda R_3'X'ZCR_1 & 0 \\ \lambda R_1'CXZ'XR_3 & D_x + \lambda D^2_x & 0 \\ 0 & 0 & I_f \end{bmatrix}.$$
respectively. Therefore, \((R^X'y, t, SS_e)\) is a set of sufficient statistics for the distribution of \(y\). Subsequently, let \(S(y) = (\phi_1, \ldots, \phi_{n-2}) = (y'R_3, t, SS_e)\).

Consider the following group of locational and scale transformations on the sample space of \(y\):

\[ G_4 = \{ g_{c, \delta} : y \rightarrow cy + X\delta, \text{ where } c > 0 \text{ and } \delta \in \mathbb{R}^3 \}. \]

This group induces the following group of transformations on the sample space of \((R^X'y, t, SS_e)\):

\[ G'_4 = \{ g'_{c, \delta} : (R^X'y, t, SS_e) \rightarrow (cR^X'y + R^X'X\delta, ct, c^2SS_e), \text{ where } c > 0 \text{ and } \delta \in \mathbb{R}^3 \}. \]

The group of transformations on the parameter space \(\Omega\) induced by \(G'_4\) is

\[ G'_4 = \{ g'_{c, \delta} : (\alpha, \sigma^2, \lambda) \rightarrow (c\alpha + \delta, c^2\sigma^2, \lambda), \text{ where } c > 0 \text{ and } \delta \in \mathbb{R}^3 \}. \]

Clearly, the parameter space \(\Omega\) is invariant with respect to \(G'_4\). The problem of testing \(H_0 : \lambda \leq \Lambda_0\) against \(H_A : \lambda \leq \Lambda_A\) is invariant with respect to \(G'_4\). Furthermore, it is straightforward to verify that \(x\) is a maximal invariant with respect to \(G'_4\).

The path of the above process was sufficiency followed by translation-and-scale invariance. The order of applying sufficiency and invariance may be reversed. Many statisticians (Cox and Hinkley, 1974; Ferguson, 1957) tend to recommend the sufficiency-then-invariance strategy because it is believed to be easier than the invariance-then-sufficiency strategy. The two
methods sometimes give equivalent answers. This is the case in
the present application. This claim can be justified by verifying
that the Assumption C of the Stein Theorem is satisfied.

By the Factorization Theorem (e.g., Ferguson, 1967, p. 115),
the p.d.f. of \(y\) is equal to \(f_1(S(y); \alpha, \sigma^2_0, \lambda) \cdot f_2(y)\), where \(f_1\) and \(f_2\)
are non-negative measurable functions defined on the sample spaces
of \(S(y)\) and \(y\), respectively; and \(f_2(y)\) does not depend on
\((\alpha, \sigma^2_0, \lambda)\). Then, in our application, the Assumption C of the Stein
Theorem requires the establishment of the following four
conditions:

1. for each \(g_{c, \delta} \in G_4\), the transformation \(g_{c, \delta}(y)\) of \(y\) is
   continuously differentiable, and the Jacobian of transformation
depends only on \(S(y)\);
2. for each \(g_{c, \delta} \in G_4\), \(S(y_1) = S(y_2)\) implies that
\(S(g_{c, \delta}(y_1)) = S(g_{c, \delta}(y_2))\);
3. the transformation \(S(y)\) on \(y\) is continuously
differentiable and the matrix \(\frac{\partial g_{c, \delta}}{\partial y_j} \) has rank \(n-f+1\);
and
4. for each \(g_{c, \delta} \in G_4\), \(f_2(g_{c, \delta}(y))/f_2(y)\) depends only on \(S(y)\).

It is straightforward to verify that the conditions (1)-(4)
are satisfied in our application. Consequently, \(x\) is a set of
sufficient statistics for the distribution of a maximal invariant
with respect to \(G_4\).
3. INVARIANT ONE-SIDED TESTS OF $H_0 : \lambda = \lambda_0$ AND THE CORRESPONDING CONFIDENCE SETS

In what follows, attention is restricted to tests of hypotheses that depend on $y$ only through the value of $x$. According to the discussion of Chapter 2, these tests are interpretable as either translation and scale invariant (when we base our decisions upon sufficient statistics) or orthogonal and scale invariant (when we base our decisions upon translation invariant statistics). We refer to any such test simply as an invariant test.

Various optimal invariant tests of $H_0 : \lambda = \lambda_0$ will be discussed. In particular, we construct the size-$\gamma$ most powerful invariant (MPI) test of $H_0$ against $H_A : \lambda = \lambda_1$ ($\lambda_1 \neq \lambda_0$), the size-$\gamma$ uniformly most powerful invariant (UMPI) test (when it exists) of $H_0$ against $H_A : \lambda > \lambda_0$, and the size-$\gamma$ locally most powerful invariant (LMPI) test of $H_0$ against $H_A : \lambda > \lambda_0$. Since the pivotal quantity $Q(\lambda)$ (1.16) depends on $y$ only through the maximal invariant $x$, Wald’s test of $H_0 : \lambda = \lambda_0$ is also an invariant test. In fact, Wald’s test of $H_0$ against $H_A : \lambda > \lambda_0$ is the most powerful invariant test of $H_0$ against $H_A' : \lambda = \omega$. Both the locally most powerful invariant test and Wald’s test reduce to the uniformly most powerful invariant test when $\Delta_1 = \ldots = \Delta_\tau$.

Recall that $x$ is sufficient for the distribution of a maximal invariant with respect to $G_4$. Thus, the optimal invariant tests
mentioned above are optimal not only in the class of tests which depend on \( y \) only through \( x \), but also in the class of tests which depend on the data only through a maximal invariant with respect to \( G_4 \) (see, e.g., Mood, Graybill, and Boes, 1974, p. 408).

For each size-\( \gamma \) optimal invariant test of \( H_0 \) with an acceptance region \( \Lambda(\lambda_0) \), a 100(1-\( \gamma \))% confidence set may be constructed as \( C(x) = \{ \lambda \in \Lambda_\lambda : x \in \Lambda(\lambda) \} \) by applying the duality (1.10). The optimality of each test is easily translated into that of the confidence set in terms of the probability of false coverage by applying the equality (1.11).

For restricted parameter spaces like \( \Lambda_1 \) and \( \Lambda_2 \), \( C(x) \) may be modified in the obvious way, giving the confidence sets \( C_1(x) = C(x) \cap \Lambda_1 \) and \( C_2(x) = C(x) \cap \Lambda_2 \), respectively. Modified confidence sets are "conservative" in the sense that their confidence levels may be greater than 1 - \( \gamma \) for values of \( \lambda \) which are on the boundary of the parameter space.

Optimal invariant tests of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda < \lambda_0 \) may be constructed in the same fashion as those of \( H_0 \) against \( H_A : \lambda > \lambda_0 \). The details are omitted.

3.1. Most Powerful Invariant Test of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda = \lambda_1 \) (\( \lambda_1 \neq \lambda_0 \))

Let \( j(\lambda;x) \) represent the likelihood function of \( \lambda \) when \( x \) is regarded as the data vector. By (2.4), we have
\[
\zeta(\lambda;\mathbf{x}) = \frac{\Gamma((r+f)/2)}{(\pi)^{r/2} \Gamma(f/2)} \cdot \frac{\left(1+\sum_1^r \left(x_1^2/(1+\lambda \Delta_1)\right)\right)^{-\frac{r+f}{2}}}{\prod_1^r (1+\lambda \Delta_1)^{1/2}} \tag{3.1}
\]

Consider testing the null hypothesis \( H_0 : \lambda = \lambda_0 \) against the alternative \( H_A : \lambda = \lambda_1 \) where \( \lambda_0, \lambda_1 \in \Lambda \) and \( \lambda_0 \neq \lambda_1 \). According to the Neyman-Pearson Lemma, the critical region of a most powerful invariant test consists of the points \( \mathbf{x} \) such that the likelihood ratio \( \zeta(\lambda_1;\mathbf{x})/\zeta(\lambda_0;\mathbf{x}) \) is sufficiently large. Hence, the size-\( \gamma \) MPI test function is

\[
\phi_{\text{MPI}}(\mathbf{x}) = \begin{cases} 
1+\sum_1^r \left(x_1^2/(1+\lambda_0 \Delta_1)\right) \\
1+\sum_1^r \left(x_1^2/(1+\lambda_1 \Delta_1)\right)
\end{cases}
\]

\[
> K_\gamma(\lambda_0, \lambda_1), \\
0, \quad \text{otherwise},
\tag{3.2}
\]

where \( K_\gamma(\lambda_0, \lambda_1) \) is chosen so that

\[
\Pr\{ \mathbf{x} : \phi_{\text{MPI}}(\mathbf{x}) = 1 ; \lambda_0 \} = \gamma.
\]

The acceptance region of \( \phi_{\text{MPI}}(\mathbf{x}) \) is

\[
\lambda_{\text{MPI}}(\lambda_0) = \{ \mathbf{x} : \frac{1+\sum_1^r \left(x_1^2/(1+\lambda_0 \Delta_1)\right)}{1+\sum_1^r \left(x_1^2/(1+\lambda_1 \Delta_1)\right)} \leq K_\gamma(\lambda_0, \lambda_1) \}. \tag{3.3}
\]
Then, according to the duality (1.10) between a confidence set and a family of hypothesis tests, a $100(1-\gamma)\%$ confidence set for $\lambda$ is

$$C_{\text{MPI}}(x) = \{ \lambda \in \Lambda_3 : x \in A_{\text{MPI}}(\lambda) \}$$

$$\begin{align*}
&= \{ \lambda \in \Lambda_3 : \frac{1}{1 + \sum_{i=1}^{r} \left( \frac{x_i^2}{1 + \lambda A_i} \right)} \leq K_\gamma(\lambda, \lambda_1) \}.
\end{align*}$$

(3.4)

Since $\phi_{\text{MPI}}(x)$ is the size-$\gamma$ MPI test of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda = \lambda_1$, we have that

$$\int_{\mathcal{R}^r} \phi(x)k(x; \lambda_1)dx \leq \int_{\mathcal{R}^r} \phi_{\text{MPI}}(x)k(x; \lambda_1)dx \quad (3.5)$$

for any size-$\gamma$ test function $\phi(x)$.

Corresponding to the test function $\phi(x)$, a $100(1-\gamma)\%$ invariant confidence set $C(x)$ of $\lambda$ can be obtained by invoking the duality relationship (1.10). Consequently, (3.5) and (1.11) together imply that

$$\Pr\{ x : \lambda_0 \in C(x); \lambda_1 \} \geq \Pr\{ x : \lambda_0 \in C_{\text{MPI}}(x); \lambda_1 \}.$$

In other words, $C_{\text{MPI}}(x)$ has the smallest probability of covering any false value of $\lambda$ when $\lambda_1$ is the true value among the class of $100(1-\gamma)\%$ invariant confidence sets of $\lambda$. 

3.2. Uniformly Most Powerful Invariant Test of $H_0 : \lambda = \lambda_0$
against $H_A : \lambda > \lambda_0$

Since the test function (3.2) depends on the value $\lambda_1$ of the
alternative hypothesis, a uniformly most powerful invariant (UMPI)
test of $H_0$ does not exist in general. We now show that, in the
special case where $\Delta_1 = \ldots = \Delta_r$, there exists a UMPI test of $H_0$
against the one-sided alternative $H_A : \lambda > \lambda_0$.

Let $\zeta^*(\lambda; x)$ denote the likelihood function associated with $x$
in the special case where $\Delta_1 = \Delta, i = 1, \ldots, r$, for some $\Delta > 0$. Then,

\[ \zeta^*(\lambda; x) = \frac{\Gamma((r+f)/2)}{(\pi)^{r/2} \Gamma(f/2)} \left\{ \frac{1+\sum_{i=1}^{r} x_i^2/(1+\lambda\Delta)}{1+\lambda\Delta} \right\}^{-(r+f)/2} \]  

(3.6)

It is a straightforward exercise to show that, for $\lambda_1 > \lambda_0$, the
likelihood ratio $\zeta^*(\lambda_1; x)/\zeta^*(\lambda_0; x)$ is a monotonically increasing
function of $\sum_{i=1}^{r} x_i^2$. Thus, the distribution function of $x$ has
monotone likelihood ratio in $\sum_{i=1}^{r} x_i^2$ (see, e.g., Lehmann, 1959, p.
68). By Lehmann's Theorem 2 (Lehmann, 1959, p. 68), the size-$\gamma$
UMPI test of $H_0 : \lambda \leq \lambda_0$ against $H_A : \lambda > \lambda_0$ is

\[
\phi_{UMPX}(x) = \begin{cases} 
1, & \text{if } \sum_{i=1}^{r} x_i^2 > K_\gamma(\lambda_0), \\
0, & \text{otherwise}
\end{cases} 
\]  

(3.7)
where \( K_\gamma(\lambda_0) \) is determined so that

\[
Pr \{ x : \phi_{\text{UMP}}(x) = 1; \lambda = \lambda_0 \} = \gamma.
\]

In particular, (3.7) defines the size-\( \gamma \) UMPI test of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda > \lambda_0 \).

Expressed in terms of \( t \) and \( SS_e \), the test statistic \( \sum_{1}^{r} x_1^2 \)
becomes \( \sum_{1}^{r} t_1^2/SS_e \). Recall, from Chapter 1, that \( \sum_{1}^{r} t_1^2/\sigma_e^2(1 + \lambda \Delta) \sim \chi^2(r) \), that \( SS_e/\sigma_e^2 \sim \chi^2(\ell) \), and that \( t \) is statistically independent of \( SS_e \). Thus, under \( H_0, \)

\[
\frac{\sum_{1}^{r} x_1^2}{1 + \lambda_0 \Delta} \sim F(r, \ell)
\]

the critical point of the test is

\[
K_\gamma(\lambda_0) = \frac{r}{f} \cdot (1 + \lambda_0 \Delta) \cdot F_\gamma(r, \ell).
\]

(3.8)

It is clear that (3.7) defines the test using the regular ANOVA table method in Section 1.1.

Converting from (3.7), a 100(1-\( \gamma \))% confidence interval for \( \lambda \)
is given as

\[
C_{\text{UMP}}(x) = \{ \lambda \in \Lambda_3 : \frac{\sum_{1}^{r} x_1^2}{\sum_{1}^{r} x_1^2} \leq \frac{r}{f} (1 + \lambda \Delta) \cdot F_\gamma(r, \ell) \}
\]

\[
= \left\{ \lambda \in \Lambda_3 : \left[ \frac{\sum_{1}^{r} x_1^2}{\sum_{1}^{r} x_1^2} \right] - 1/\Delta \leq \lambda \right\}.
\]

(3.9)

Consequently, we have obtained a 100(1-\( \gamma \))% confidence lower limit for \( \lambda \).

Since \( \phi_{\text{UMP}}(x) \) is the size-\( \gamma \) uniformly most powerful invariant test of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda > \lambda_0 \), we have, for any size-\( \gamma \)
test $\phi(x)$, that

$$
\int_{\mathbb{R}^2} \phi(x) k(x; \lambda_1) \, dx \leq \int_{\mathbb{R}^2} \phi_{\text{UMPI}}(x) k(x; \lambda_1) \, dx
$$

for every $\lambda_1$ such that $\lambda_1 \geq \lambda_0$. Therefore, for any $100(1 - \gamma)$% invariant confidence set $C(x)$, we have that

$$
\Pr \{ x : \lambda_0 \in C(x); \lambda_1 \} \geq \Pr \{ x : \lambda_0 \in C_{\text{UMPI}}(x); \lambda_1 \}
$$

for every $\lambda_1$ such that $\lambda_1 \geq \lambda_0$. That is, among the class of $100(1 - \gamma)$% invariant confidence sets, $C_{\text{UMPI}}(x)$ covers any false value $\lambda_0$ with the smallest probability for any true value of $\lambda$ such that $\lambda_0 < \lambda$. $C_{\text{UMPI}}(x)$ defines the so-called uniformly most accurate invariant lower confidence bound for $\lambda$ (Ferguson, 1967, section 5.8).

There are, at least, four cases where $\Delta_1 = \ldots = \Delta_r$. These four cases are (1) the one-way random model with only two levels of the random factor, (2) the balanced one-way random model, (3) the randomized complete block design with equal block sizes, and (4) the linked incomplete block design. For both designs (3) and (4), the mixed linear model is assumed with block effects being treated as the random effects. Thompson (1955b) showed that $\Delta_1 = \ldots = \Delta_r$ in case (4). In the following, cases (1)-(3) are briefly discussed.

First, consider model (1.3) when $b = 2$. This model is the special case of model (1.1) where $X = 1_n$ and $Z = \begin{bmatrix} 1_n & 0 \\ 0 & 1_{n_2} \end{bmatrix}$.
For this case, \( \text{rank}(C) = \text{rank}(X,Z) - \text{rank}(X) = 1 \). Therefore, \( C \) has only one non-zero eigenvalue. (It is easy to show that the non-zero eigenvalue is \( 2n_1n_2/n \).) Second, consider the balanced one-way random model, i.e., model (1.3) with \( n_1 = \ldots = n_b = n_0 \) for some \( n_0 \). This is the special case of model (1.1) where \( X = 1_n \) and \( Z = \text{diag}(1, \ldots, 1) \). It is straightforward to show that \( C = n_0(I_b - (1/b)1_b1_b') \), that \( \text{rank}(C) = b - 1 \), and that the unique non-zero eigenvalue of \( C \) is \( n_0 \).

Finally, we show that, in the case of a complete block design with equal block size, all of the non-zero eigenvalues of \( C \) are equal. A complete block design for \( b \) blocks of size \( k \) is one in which \( t \) treatments are replicated \( r_1, \ldots, r_t \) times, respectively, in such a way that \( n_i = kr_i/n \) units in each block are assigned to the \( i \)th treatments \( (i = 1, \ldots, t) \). Here, \( n = \sum_{i=1}^{t} r_i = bk \) represents the total number of units. A possible model for the observations from a complete block design is

\[
y_{ijl} = \mu + \tau_i + \beta_j + e_{ijl},
\]

\( l = 1, \ldots, n_i; i = 1, \ldots, t; j = 1, \ldots, b \) \hspace{1cm} (3.10)

where \( y_{ijl} \) is the \( l \)th observation of the \( j \)th block to which the \( i \)th treatment is assigned, \( \mu \) and \( (\tau_1, \tau_2, \ldots, \tau_t) \) are unknown constants, \( (\beta_1, \beta_2, \ldots, \beta_b)' \) is a \( b \times 1 \) random vector distributed as \( \text{MVN}_b(0, \sigma^2_\beta I) \), \( (e_{111}, e_{112}, \ldots, e_{tbn})' \) is an \( n \times 1 \) random vector distributed as \( \text{MVN}_n(0, \sigma^2_e I) \), and the \( \beta \)'s are statistically
independent of the e's. The variance components \( \sigma_\beta^2 \) and \( \sigma_e^2 \) are taken to be unknown parameters satisfying \( \sigma_\beta^2 \geq 0 \) and \( \sigma_e^2 > 0 \).

Model (3.10) can be formulated as a special case of model (1.1) by taking \( y = (y_{111}, y_{112}, \ldots, y_{tbn_t})' \), \( \alpha = (\mu, \tau_1, \ldots, \tau_t)' \), \( \beta = (\beta_1, \ldots, \beta_b)' \), \( X = (1, W) \), and

\[
Z = \begin{bmatrix}
I_b \otimes 1_{n_1} \\
I_b \otimes 1_{n_2} \\
\vdots \\
I_b \otimes 1_{n_t}
\end{bmatrix},
\]

where \( W = \text{diag}(1_{n_1}) \) and \( I_b \otimes 1_{n_i} \) represents the Kronecker product of \( I_1 \) and \( 1_{n_i} \), \( i = 1, \ldots, t \) (e.g., Anderson, 1984, p. 599).

It is a straightforward exercise to show that, in this special case, the matrix \( C \) equals \( k(I_b - \frac{1}{b}1_b 1_b' \), \( \text{rank}(C) = b - 1 \), and the non-zero eigenvalues of \( C \) are all equal to \( k \).

It is easy to see that the size-\( \gamma \) UMPI test of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda < \lambda_0 \) is

\[
\phi_{\text{UMPI}}(x) = \begin{cases} 
1, & \text{if } \sum \frac{x_i^2}{1} < \frac{x}{f} \cdot (1 + \lambda_0 \Delta) \cdot F_{1-\gamma}(x,f) \\
0, & \text{otherwise}
\end{cases}
\]

(3.11)

and the corresponding 100(1-\( \gamma \))% confidence interval for \( \lambda \) is

\[
C_{\text{UMPI}}(x) = \{ \lambda \in \Lambda_3 : \sum \frac{x_i^2}{1} \geq \frac{x}{f} \cdot (1 + \lambda \Delta) \cdot F_{1-\gamma}(x,f) \}.
\]
3.3. Locally Most Powerful Invariant Test of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda > \lambda_0$

It was pointed out in Section 3.2 that UMPI test of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda > \lambda_0$ does not exist in general. We now consider the locally most powerful invariant (LMPI) test. To test $H_0$ against $H_A$, an LMPI test, when it exists, is optimal in a neighborhood of $\lambda_0$ in the sense that the slope of the power function of the test at $\lambda = \lambda_0$ is maximized.

The existence of an LMPI test is established by the following lemma (see, e.g., Ferguson, 1967, p. 235).

**Lemma 3.1**

Let $\phi(x)$ be an invariant test and take $\beta_\phi(\lambda)$ to be the power function of the test defined by $\beta_\phi(\lambda) = \int_{\mathcal{X}} \phi(x) k(x;\lambda) dx$ ($\lambda \in \Lambda_3$).

Then,

(i) $\beta_\phi(\lambda)$ is continuous in $\lambda$;

(ii) $\frac{d}{d\lambda} \beta_\phi(\lambda) = \int_{\mathcal{X}} \phi(x) \frac{\partial}{\partial \lambda} k(x;\lambda) dx$.

**Proof**

(i) Let $\lambda$ represent an arbitrary point in $\Lambda_3$. Since $\phi(x)$ is a test function, $\phi(x)$ is measurable and $0 \leq \phi(x) \leq 1$. Thus, since

$$= \left\{ \lambda \in \Lambda_3 : \left[-\frac{f}{r} \cdot \left( \sum_{i=1}^{r} \frac{x_i^2}{F_{1-\gamma}(r, \ell)} \right) - 1\right] / \Delta \geq \lambda \right\}$$

which defines a $100(1-\gamma)$% confidence upper limit.
k(x; λ) is integrable, φ(x)k(x; λ) is integrable for every λ ∈ Λ_3.

By definition,

\[ \beta_\phi(\lambda) = \int_{\mathcal{R}^r} \phi(x) \frac{\Gamma(-r/2) (1+\sum_{1}^{r} \frac{x_1^2}{1+\lambda \Delta_1})^{-(r+f)/2}}{\Gamma(-f/2) (\pi)^{r/2} \prod_{1}^{r} (1+\lambda \Delta_1)^{1/2}} \, dx \]

= \frac{\beta(\lambda)}{c p_1(\lambda)} \int_{\mathcal{R}^r} \phi(x) p_2(x; \lambda) \, dx

where

\[ c = \frac{\Gamma(-r/2)}{\Gamma(-f/2) (\pi)^{r/2}}, \]

\[ p_1(\lambda) = \prod_{1}^{r} (1+\lambda \Delta_1)^{-1/2}, \]

and

\[ p_2(x; \lambda) = (1+\sum_{1}^{r} \frac{x_1^2}{1+\lambda \Delta_1})^{-(r+f)/2}. \]

Let \( \{ \lambda_n \}_{1}^{\infty} \) represent a sequence of numbers in Λ_3 such that \( \lambda_n \leq B \) for every \( n \) and some \( B \) and \( \lambda_n \to \lambda \) as \( n \to \infty \). Then, for each \( n \),

\[ p_2(x; \lambda_n) \leq (1+\sum_{1}^{r} \frac{x_1^2}{1+B \Delta_1})^{-(r+f)/2} = p_2(x; B). \]

That is, the sequence of functions \( p_2(x; \lambda_n) \) is dominated by an integrable function. By the Lebesgue Dominated Convergence Theorem (e.g., Bartle, 1966, Corollary 5.7), we have

\[ \lim_{n \to \infty} \beta_\phi(\lambda_n) = \lim_{n \to \infty} c p_1(\lambda_n) \int_{\mathcal{R}^r} \phi(x) p_2(x; \lambda_n) \, dx \]
\[= \lim_{n \to \infty} p_1(\lambda_n) \int_{\mathbb{R}} \phi(x) \lim_{n \to \infty} p_2(x; \lambda_n) \, dx\]

\[= c \left( \int_{\mathbb{R}} \phi(x) p_2(x; \lambda) \, dx \right)\]

\[= \beta \phi(\lambda).\]

(11) Clearly,

\[
\frac{\partial}{\partial \lambda} \phi(x) k(x; \lambda) = c \phi(x) \cdot \frac{\partial}{\partial \lambda} \left\{ \frac{x^2}{\prod_{i=1}^{\infty} (1 + \lambda \Delta_i)^{1/2}} \right\}^{-(r+f)/2} \left[ 1 + \sum_{i=1}^{r} \frac{x_i}{1 + \lambda \Delta_i} \right]^{-1/2} \]

Furthermore,

\[
\frac{\partial}{\partial \lambda} \prod_{i=1}^{\infty} (1 + \lambda \Delta_i)^{-1/2} = \prod_{i=1}^{\infty} \left\{ \prod_{j \neq i}^{\infty} (1 + \lambda \Delta_j)^{-1/2} \cdot \frac{\partial}{\partial \lambda} (1 + \lambda \Delta_i)^{-1/2} \right\}^{-1}\]

\[= \prod_{i=1}^{\infty} \left\{ \prod_{j \neq i}^{\infty} (1 + \lambda \Delta_j)^{-1/2} \cdot \left( -\frac{\Delta_i}{2} \right) \cdot (1 + \lambda \Delta_i)^{-3/2} \right\}\]

\[= -\frac{1}{2} \sum_{i=1}^{\infty} \prod_{j=1}^{\infty} (1 + \lambda \Delta_j)^{-1/2} \cdot \left\{ \frac{\Delta_i}{1 + \lambda \Delta_i} \right\}\]

\[= -\frac{1}{2} \prod_{i=1}^{\infty} (1 + \lambda \Delta_i)^{-1/2} \sum_{i=1}^{\infty} \frac{\Delta_i}{1 + \lambda \Delta_i}\]

and

\[
\frac{\partial}{\partial \lambda} \left\{ 1 + \sum_{i=1}^{r} \frac{x_i^2}{1 + \lambda \Delta_i} \right\}^{-\frac{r+f}{2}} = -\frac{r+f}{2} \left\{ 1 + \sum_{i=1}^{r} \frac{x_i^2}{1 + \lambda \Delta_i} \right\}^{-\frac{r+f-1}{2}} \cdot \frac{\partial}{\partial \lambda} \left\{ \sum_{i=1}^{r} \frac{x_i^2}{1 + \lambda \Delta_i} \right\}\]
Hence,

\[
\frac{\partial}{\partial \lambda} \phi(x) k(x; \lambda) =
\]

\[
\phi(x) \cdot c \left\{ \frac{-1}{2} \left[ \frac{\Delta_1}{1+\lambda \Delta_1} \right]^{-1/2} \cdot \left[ \frac{\Delta_1}{1+\lambda \Delta_1} \right] \cdot \left[ \frac{\Delta_1}{(1+\lambda \Delta_1)^2} \right] \cdot \left[ \frac{1}{\lambda} \right] \cdot \left[ \frac{1}{1+\lambda \Delta_1} \right] \right\}
\]

\[
\leq \frac{1}{2} k(x; \lambda) \cdot \left\{ \frac{\Delta_1}{1+\lambda \Delta_1} + (x+f) \cdot \frac{\Delta_1}{1+\lambda \Delta_1} \right\}
\]

Let \( \Delta^* = \max \{ \Delta_i \} \). It is clear that

\[
\frac{\Delta_1}{1+\lambda \Delta_1} \leq \frac{\Delta^*}{1+\lambda \Delta^*}, \quad 1 \leq i \leq r
\]

Thus,

\[
\frac{\Delta_1}{1+\lambda \Delta_1} + (x+f) \cdot \frac{\Delta^*}{1+\lambda \Delta^*} \leq \frac{\Delta^*}{1+\lambda \Delta^*} + (x+f) \cdot \frac{\Delta^*}{1+\lambda \Delta^*}
\]

\[
\leq (2r+f) \frac{\Delta^*}{1+\lambda \Delta^*}
\]

Therefore,
\[ \left| \frac{\partial}{\partial \lambda} \phi(x) k(x; \lambda) \right| \leq \frac{2x+f}{2} \frac{\Delta^*}{1+\lambda \Delta^*} k(x; \lambda) \]

\[ = \frac{2x+f}{2} \frac{\Delta^*}{1+\lambda \Delta^*} c \ p_1(\lambda) \ p_2(x; \lambda). \quad (3.13) \]

Since \( \Lambda^* \) is an open set on the real line, there exist \( \underline{\lambda} \) and \( \overline{\lambda} \) in \( \Lambda^* \) such that \( \underline{\lambda} < \lambda < \overline{\lambda} \). Observe that

\[ \frac{\Delta^*}{1+\lambda \Delta^*} p_1(\lambda) < \frac{\Delta^*}{1+\lambda \Delta^*} p_1(\lambda) \quad (3.14)\]

and

\[ p_2(x; \lambda) < p_2(x; \overline{\lambda}). \quad (3.15) \]

Thus, \[ \left| \frac{\partial}{\partial \lambda} \phi(x) k(x; \lambda) \right| \] is dominated by \[ \frac{2x+f}{2} \frac{\Delta^*}{1+\lambda \Delta^*} c \ p_1(\lambda) \ p_2(x; \overline{\lambda}) \]

which is free of \( \lambda \) and integrable. By Corollary 5.9 of Bartle (1966), the proof is complete.

Q.E.D.

By definition, a size-\( \gamma \) LMP test of \( H_0 \) against \( H_1 \) is obtained by maximizing the slope of the power function at \( H_0 \) subject to the size requirement, that is, by maximizing

\[ \frac{d}{d\lambda} \beta(\lambda) \bigg|_{\lambda_0} = \int_{\mathbb{R}} \phi(x) \frac{\partial}{\partial \lambda} k(x; \lambda) \bigg|_{\lambda=\lambda_0} \ dx \]

subject to

\[ \beta(\lambda_0) = \int_{\mathbb{R}} \phi(x) k(x; \lambda_0) \ dx = \gamma. \quad (3.16) \]
By the generalized Neyman-Pearson Lemma (see, e.g., Lehmann, 1959, p. 83), this maximization problem is solved by taking the test function to be

\[
\phi(x) = \begin{cases} 
1, & \text{if } \frac{\partial}{\partial \lambda} k(x; \lambda)\bigg|_{\lambda = \lambda_0} > c_0 k(x; \lambda_0), \\
0, & \text{otherwise}
\end{cases}
\]

\[
= \begin{cases} 
1, & \text{if } \log k(x; \lambda)\bigg|_{\lambda = \lambda_0} > c_0, \\
0, & \text{otherwise}
\end{cases}
\]

where the constant \( c_0 \) is chosen so that (3.16) is satisfied.

Thus, the LMP test of \( H_0 \) against \( H_A \) consists of rejecting \( H_0 \) for values of \( x \) such that \( \frac{\partial}{\partial \lambda} \log k(x; \lambda)\bigg|_{\lambda = \lambda_0} \) is sufficiently large. Since

\[
\frac{\partial}{\partial \lambda} \log k(x; \lambda) = \frac{\partial}{\partial \lambda} \left\{ -\frac{1}{2} \sum_{i=1}^{r} \log(1+\lambda \Delta_i) - \frac{r+f}{2} \log(1 + \sum_{i=1}^{r} \frac{x_i^2}{1+\lambda \Delta_i}) \right\}
\]

\[
= -\frac{1}{2} \sum_{i=1}^{r} \frac{\Delta_i x_i^2}{1+\lambda \Delta_i} + \frac{r+f}{2} \cdot \frac{r \Delta_i x_i^2}{1+\sum_{i=1}^{r} \frac{x_i^2}{1+\lambda \Delta_i}},
\]

the size-\( \gamma \) LMP test function is
\[
\phi^{(\text{LMPX})}(x) = \begin{cases} 
1, & \text{if } \sum_{1}^{r} \frac{\Delta_{1} x_{1}^{2}}{(1+\lambda_{1} \Delta_{1})^{2}} > K_{\gamma}(\lambda_{0}), \\
1 + \sum_{1}^{r} \frac{x_{1}^{2}}{1+\lambda_{1} \Delta_{1}}, & \text{otherwise} 
\end{cases}
\]  

(3.17)

where \(K_{\gamma}(\lambda_{0})\) is determined to satisfy

\[
\Pr \{ x : \phi^{(\text{LMPX})}(x) = 1; \lambda_{0} \} = \gamma. 
\]  

(3.18)

A \((1-\gamma)\)% confidence set of \(\lambda\) is therefore

\[
C^{(\text{LMPX})}(x) = \left\{ \lambda \in \Lambda_{3} : \frac{\sum_{1}^{r} \Delta_{1} x_{1}^{2}}{(1+\lambda \Delta_{1})^{2}} \leq K_{\gamma}(\lambda) \right\}. 
\]  

(3.19)

An "optimal" property of \(C^{(\text{LMPX})}(x)\) may be deduced as follows.

Since \(\phi^{(\text{LMPX})}(x)\) is the size-\(\gamma\) locally most powerful invariant test of \(H_{0} : \lambda = \lambda_{0}\) against \(H_{A} : \lambda > \lambda_{0}\), for any size-\(\gamma\) test \(\phi(x)\) of \(H_{0}\), it is clear that

\[
\frac{d}{d\lambda} \int_{\mathcal{X}} \phi(x) k(x;\lambda) dx \big|_{\lambda_{0}} \leq \frac{d}{d\lambda} \int_{\mathcal{X}} \phi^{(\text{LMPX})}(x) k(x;\lambda) dx \big|_{\lambda_{0}}.
\]

Consequently, for any \(100(1-\gamma)\)% confidence set \(C(x)\), we have

\[
\frac{d}{d\lambda} \Pr \{ x : \lambda_{0} \in C(x); \lambda \} \big|_{\lambda_{0}} \geq \ldots
\]
\[
\frac{d}{d\lambda} \Pr\{ x : \lambda_0 \in C_{\text{LMPI}}(x) ; \lambda \} \bigg| _{\lambda = \lambda_0}.
\]

That is, for true values of \( \lambda \) in a sufficiently small neighborhood of a false value \( \lambda_0 \), the probability of false coverage of the confidence set \( C_{\text{LMPI}}(x) \) is smaller than that of any other \( 100(1-\gamma)\% \) invariant confidence set.

It is straightforward to verify that the size-\( \gamma \) LMPI test of \( H_0 : \lambda < \lambda_0 \) is

\[
\phi_{\text{LMPI}}(x) = \begin{cases} 
1, & \text{if } \frac{\sum_{i=1}^{r} \frac{\Delta_i x_i^2}{(1+\lambda_0 \Delta_i)^2}}{1+\sum_{i=1}^{r} \frac{x_i}{1+\lambda_0 \Delta_i}} < k_{1-\gamma}(\lambda_0), \\
0, & \text{otherwise}
\end{cases}
\]

(3.20)

where \( k_{1-\gamma}(\lambda_0) \) is chosen so that

\[
\Pr\{ x : \phi_{\text{LMPI}}(x) = 1; \lambda_0 \} = \gamma.
\]

The corresponding \( 100(1-\gamma)\% \) confidence set is

\[
C_{\text{LMPI}}(x) = \left\{ \lambda \in \Lambda_3 : \frac{\sum_{i=1}^{r} \frac{\Delta_i x_i^2}{(1+\lambda \Delta_i)^2}}{1+\sum_{i=1}^{r} \frac{x_i}{1+\lambda \Delta_i}} \geq k_{1-\gamma}(\lambda) \right\}.
\]

(3.21)

When \( \Delta_1 = \ldots = \Delta_r \), the test statistic of the LMPI test reduces to \( \sum_{i=1}^{r} x_i^2 \), which is the test statistic of the UMPI test.
That is, the LMPI test reduces to the UMPI test and $C_{\text{LMPI}}(x)$ of (3.19) and (3.21) reduces to $C_{\text{UMPI}}(x)$ of (3.9) and (3.12), respectively, when $\Delta_i = \Delta$, $i = 1, \ldots, r$.

3.4. Wald's Test of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda > \lambda_0$ as the MPI Test of $H_0$ against $H_A' : \lambda = \infty$

Recall that the pivotal quantity $Q(\lambda)$ of (1.16) is distributed as $F(r, f)$ and is monotonically decreasing in $\lambda$.

Therefore, a size-$\gamma$ Wald's test of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda > \lambda_0$ is

$$
\phi_w(x) = \begin{cases} 
1, & \text{if } \frac{1}{1+\lambda_0} \sum_{i=1}^{r} \frac{x_i^2}{\Delta_i} > F_\gamma(r, f), \\
0, & \text{otherwise.}
\end{cases}
$$

(3.22)

We now consider the nature of the MPI test (3.2) as $\lambda_1 = \infty$.

Observe that

$$
\lim_{\lambda_1 \to \infty} \frac{1 + \sum_{i=1}^{r} \frac{x_i^2}{1+\lambda_0 \Delta_i}}{1 + \sum_{i=1}^{r} \frac{x_i^2}{1+\lambda_1 \Delta_i}} = 1 + \sum_{i=1}^{r} \frac{x_i^2}{1+\lambda_0 \Delta_i}.
$$

Thus, in "testing $H_0$ against $H_A' : \lambda = \infty$", the MPI test rejects $H_0$ when $1 + \sum_{i=1}^{r} \frac{x_i^2}{1+\lambda_0 \Delta_i}$ is sufficiently large, or equivalently, when $\sum_{i=1}^{r} \frac{x_i^2}{1+\lambda_0 \Delta_i}$ is sufficiently large. This test is identical to Wald's test (3.22) of $H_0$ against $H_A : \lambda > \lambda_0$. Therefore, Wald's test can
be expected to perform well when the true value of $\lambda$ is large. As a matter of fact, it is shown in Subsection 6.1.3 that the power of Wald's test of $H_0$ against $H_A$ approaches one when the true value of $\lambda$ approaches $\infty$. The locally most powerful invariant test of $H_0$ against $H_A : \lambda > \lambda_0$ given by (3.17) does not necessarily have this property.

A 100(1 - $\gamma$)% confidence set for $\lambda$ is

$$C_w(x) = \left\{ \lambda \in \Lambda_3 : \sum_{i=1}^{r} \frac{x_i^2}{1 + \lambda \Delta_i} \cdot \frac{f}{r} \leq F_\gamma(r, f) \right\}. \quad (3.23)$$

Since $\sum_{i=1}^{r} \frac{x_i^2}{1 + \lambda \Delta_i}$ is strictly decreasing in $\lambda$, (3.23) defines a 100(1 - $\gamma$)% confidence interval $[\lambda_x, \infty)$ where the lower bound $\lambda_x$ satisfies the equation $\sum_{i=1}^{r} \frac{x_i^2}{1 + \lambda \Delta_i} \cdot \frac{f}{r} = F_\gamma(r, f)$. The confidence interval $C_w(x)$ has the property that, for sufficiently large true values of $\lambda$, the probability of covering a false value of $\lambda$ is smaller for $C_w(x)$ than for any other 100(1 - $\gamma$)% invariant confidence set.

Clearly, the size-$\gamma$ Wald's test of $H_0$ against $H_A : \lambda < \lambda_0$ is

$$\phi_w(x) = \begin{cases} 1, & \text{if } \sum_{i=1}^{r} \frac{x_i^2}{1 + \lambda_0 \Delta_i} \cdot \frac{f}{r} < F_{1-\gamma}(r, f), \\ 0, & \text{otherwise}, \end{cases} \quad (3.24)$$

and the corresponding 100(1 - $\gamma$)% confidence interval for $\lambda$ is
\[ C_w(x) = \left\{ \lambda \in \Lambda_3 : \frac{\sum_{1}^{r} x_i^2}{1 + \lambda \Delta_1} \cdot \frac{f}{r} \geq F_{1-\gamma}(r, f) \right\} \] (3.25)

which apparently gives a confidence upper limit.

Note that when \( \Delta_i = \Delta \) for all \( i \), the test statistic \( \phi_w(x) \) reduces to \( \sum_{1}^{r} x_i^2 \), which is the test statistic of the UMPI test.

Accordingly, Wald's test reduces to the UMPI test whenever the latter exists and the two confidence intervals \( C_w(x) \) of (3.23) and (3.25) are identical to \( C_{\text{UMPI}}(x) \) of (3.9) and (3.12), respectively.
4. INVARIANT TWO-SIDED TESTS OF $H_o : \lambda = \lambda_0$ AND THE CORRESPONDING CONFIDENCE SETS

In the Chapter 3, several confidence sets for $\lambda$ were derived from families of optimal invariant tests of $H_o : \lambda = \lambda_0$ against one-sided alternatives. In these tests, only one direction of departure from the null hypothesis $H_o$ was emphasized. Consider now the case where both directions of departure from $H_o$ are to be emphasized, that is, consider the problem of testing $H_o : \lambda = \lambda_0$ against $H_A : \lambda \neq \lambda_0$ for some $\lambda_0 \in \Lambda$. Confidence sets will be constructed from families of optimal invariant two-sided tests.

Like the optimal invariant tests discussed in the Chapter 3, the optimal invariant two-sided tests are optimal not only in the class of tests depending on the data only through $x$ but also in that depending on the data only through a maximal invariant with respect to $G_A$.

4.1. Likelihood Ratio Invariant Test of $H_o : \lambda = \lambda_0$ against $H_A : \lambda \neq \lambda_0$

A size-$\gamma$ likelihood ratio invariant test of $H_o$ against $H_A$, in the class of invariant tests which depend on $y$ only through $x$, is (e.g., Bickel and Docksum, 1977, p. 209)

$$\phi_{LR}(x) = \begin{cases} 
1, & \text{if } \frac{\ell(\lambda_0; x)}{\sup_{\lambda' \in \Lambda} \ell(\lambda'; x)} < K_{\gamma}(\lambda_0), \\
0, & \text{otherwise},
\end{cases} \quad (4.1)$$
where \( \sup_{\lambda' \in \Lambda} \ell(\lambda'; \mathbf{x}) \) is the supremum, as a function of \( \lambda', \) of \( \ell(\lambda; \mathbf{x}) \)

over \( \Lambda \) and \( K_\gamma(\lambda_0) \) is such that

\[
Pr\{ \mathbf{x} : \phi_{\text{LRI}}(\mathbf{x}) = 1; \lambda_0 \} = \gamma.
\]

The acceptance region of the test (4.1) is

\[
A_{\text{LRI}}(\lambda_0) = \{ \mathbf{x} : \ell(\lambda_0; \mathbf{x}) / \sup_{\lambda' \in \Lambda} \ell(\lambda'; \mathbf{x}) \geq K_\gamma(\lambda_0) \}. \quad (4.2)
\]

A 100(1-\( \gamma \))% confidence set is

\[
C_{\text{LRI}}(\mathbf{x}) = \{ \lambda : \ell(\lambda; \mathbf{x}) / \sup_{\lambda' \in \Lambda} \ell(\lambda'; \mathbf{x}) \geq K_\gamma(\lambda) \}. \quad (4.3)
\]

It is easy to verify that \( \lim \log \ell(\lambda; \mathbf{x}) = \lim \log \ell(\lambda; \mathbf{x}) = \lambda \rightarrow -1/\Delta \)

\(-\infty \) for almost every \( \mathbf{x} \) and that \( \log \ell(\lambda; \mathbf{x}) \) is continuous over

\((-1/\Delta, \infty)\). Therefore, there exists at least one value of \( \lambda \) which

maximizes \( \ell(\lambda; \mathbf{x}) \) for \( \lambda \in \Lambda \). Such a maximum value must satisfy

the likelihood equation \( \frac{\partial}{\partial \lambda} \log \ell(\lambda; \mathbf{x}) = 0. \) Since

\[
\frac{\partial}{\partial \lambda} \log \ell(\lambda; \mathbf{x}) = -\frac{1}{2} \sum_{i=1}^{r} \frac{\Delta_i x_i^2}{1 + \lambda \Delta_i} + \frac{r + \ell}{1 + \sum_{i=1}^{r} \frac{x_i^2}{1 + \lambda \Delta_i}},
\]

the likelihood equation is

\[
-\sum_{i=1}^{r} \frac{\Delta_i}{1 + \lambda \Delta_i} + (r + \ell) \frac{\sum_{i=1}^{r} x_i^2}{1 + \sum_{i=1}^{r} \frac{x_i^2}{1 + \lambda \Delta_i}} = 0. \quad (4.4)
\]
A solution to the equation (4.4) may be outside the range \( \Lambda_1 = (0, \infty) \). Thus, when \( \Lambda \) is taken to be \( \Lambda_1 \), \( \xi(\lambda; x) \) may attain its maximum at the boundary point zero. Similarly, when \( \Lambda \) is taken to be \( \Lambda_2 \), \( \xi(\lambda; x) \) may attain its supremum at the boundary point \(-1/\omega\) which is not included in the parameter space. Although a maximum value of \( \xi(\lambda; x) \) for \( \lambda \in \Lambda_2 \) might not exist, \( \sup_{\lambda \in \Lambda_2} \log \xi(\lambda; x) \) is well-defined.

Let \( \tilde{\lambda} \) represent any value of \( \lambda \in \Lambda \) such that \( \xi(\tilde{\lambda}; x) = \sup_{\lambda' \in \Lambda} \xi(\lambda'; x) \). That is, \( \tilde{\lambda} \) is a value of \( \lambda \) at which \( \xi(\lambda; x) \) attains its maximum over \( \Lambda \) when \( \Lambda \) is taken as \( \Lambda_1 \) or \( \Lambda_3 \), and its supremum when \( \Lambda \) is \( \Lambda_2 \). Then, the size-\( \gamma \) likelihood ratio invariant (LRI) test of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda \neq \lambda_0 \) is

\[
\phi_{LRI}(x) = \begin{cases} 
1, & \text{if } 1 + \frac{x^2}{(1 + \tilde{\lambda} \Delta_1)} < K_\gamma(\lambda_0) \\
1 + \frac{x}{\sum_1^r (x^2/(1 + \tilde{\lambda} \Delta_1))}, & \text{otherwise,}
\end{cases}
\]

(4.5)

where \( K_\gamma(\lambda_0) \) is determined so that

\[
\Pr \{ x : \phi_{LRI}(x) = 1; \lambda_0 \} = \gamma.
\]

The 100(1-\( \gamma \))% confidence set \( C_{LRI}(x) \) can be re-expressed as

\[
C_{LRI}(x) =
\]
4.1.1. Likelihood ratio invariant test of $H_0 : \lambda = \lambda_0$ against

$H_A : \lambda \neq \lambda_0$ when $\Lambda = \Lambda_3$ and $\Delta_i = \Delta$ for all $i$

When $\Lambda = \Lambda_3$ and $\Delta_i = \Delta$ for all $i$, $\mathcal{L}(\lambda; x)$ is maximized uniquely at $\tilde{\lambda}_3 = \frac{1}{\Delta} - \frac{1}{\Delta} \sum x_i^2 - 1 = \frac{1}{\Delta} (r - 1)$ as can be easily verified from (4.4). Let $T = \sum x_i^2 / (1 + \lambda_0 \Delta)$. Then, the LRI test of $H_0$ against $H_A$ is to reject $H_0$ when the observed $S_3(T)$ is sufficiently small, where

$$S_3(T) = \frac{r^2}{(1 + T)^{r/2}}.$$

It is easy to see that $\lim_{T \to \infty} S_3(T) = \lim_{T \to 0} S_3(T) = 0$ and that

$$\frac{d}{dT} S_3(T) = -\frac{r}{2} T^{r/2-1} (1 + T)^{-r/2}$$

$$- \frac{r + f}{2} (1 + T)^{-(r+f)/2-1} T^{r/2}$$

$$= \frac{1}{2} T^{r/2-1} (1 + T)^{-(r+f)/2-1} (r - fT)$$
\[
\phi_{LRI}(T) = \begin{cases} 
1, & \text{if } T < T_1(\lambda_0) \text{ or } > T_2(\lambda_0), \\
0, & \text{otherwise},
\end{cases}
\]

(4.6)

where \( T_1(\lambda_0) \) and \( T_2(\lambda_0) \) are determined so that \( T_1(\lambda_0) \leq T_2(\lambda_0) \),
\( S_3(T_1(\lambda_0)) = S_3(T_2(\lambda_0)) \), and

\[\Pr\{ T : \phi_{LRI}(T) = 1 ; \lambda_0 \} = \gamma.\]

Clearly, \( \frac{\hat{f}}{\ell} - T \sim F(\gamma, \ell) \) under \( H_0 \). Thus, (4.6) defines a two-sided \( F \)-test

\[
\phi_{LRI}(x) = \begin{cases} 
1, & \frac{\sum x_i^2}{\ell} < F_{1-\gamma_1}(\gamma, \ell) \text{ or } > F_{\gamma_2}(\gamma, \ell), \\
0, & \text{otherwise},
\end{cases}
\]

(4.7)

where \( \gamma_1 \) and \( \gamma_2 \) satisfy \( \gamma_1 + \gamma_2 = \gamma \) and

\[
\begin{bmatrix} 
\frac{F_{1-\gamma_1}(\gamma, \ell)}{F_{\gamma_2}(\gamma, \ell)} \\
\frac{F_{1-\gamma_1}(\gamma, \ell)}{F_{\gamma_2}(\gamma, \ell)} 
\end{bmatrix} \equiv \begin{bmatrix} 
\frac{\ell + \gamma \cdot F_{1-\gamma_1}(\gamma, \ell)}{\ell + \gamma \cdot F_{\gamma_2}(\gamma, \ell)} \\
\frac{\ell + \gamma \cdot F_{1-\gamma_1}(\gamma, \ell)}{\ell + \gamma \cdot F_{\gamma_2}(\gamma, \ell)} 
\end{bmatrix} \quad (4.8)
\]
4.1.2. Likelihood ratio invariant test of $H_0: \lambda = \lambda_0$ against $H_A: \lambda \neq \lambda_0$ when $A = A_1$ and $A_1 = \Delta$ for all $i$

In the special case when $A = A_1$ and $A_1 = \Delta$ for all $i$, the maximum likelihood estimator of $\lambda$ is $\lambda = \max\{0, \lambda_3\} = \max\{0, (\frac{f}{r} \sum x_i^2 - 1)/\Delta\}$. Thus, the null hypothesis is rejected when the observed value of $S_1(T)$ is sufficiently small, where

$$S_1(T) = \begin{cases} 
  S_3(T), & \text{if } \frac{f}{r} \sum x_i^2 > 1, \\
  \left(1 + \lambda_0 \Delta\right)^{-x/2} \cdot \left[ \frac{1 + \sum x_i^2}{1 + \sum x_i^2/(1 + \lambda_0 \Delta)} \right]^{-x/2}, & \text{otherwise}
\end{cases}$$

(4.9)

$S_1(T)$ is a continuous function of $T$. Note that $\lim_{T \to \infty} S_1(T) = 0$, $\lim_{T \to 0} S_1(T) = (1 + \lambda_0 \Delta)^{-x/2} > 0$, and $S_1(T) > S_3(T)$ for $\frac{f}{r} (1 + \lambda_0 \Delta)T < 1$. Since the null hypothesis $H_0$ is rejected when the observed $S_1(T)$ is sufficiently small, the test is either a one-sided or a two-sided $P$-test, depending on the values of $r$, $f$, $\Delta$, $\lambda_0$, and $\gamma$. When $\gamma$ is less than $\Pr\{ T : S_1(T) < (1 + \lambda_0 \Delta)^{-x/2}; \lambda_0 \}$, the size-$\gamma$ LRI test becomes
\[
\Phi_{LRI}(T) = \begin{cases} 
1, & \text{if } T > T(\lambda_0), \\
0, & \text{otherwise},
\end{cases} \quad (4.10)
\]

where \( T(\lambda_0) \) is chosen so that
\[
\Pr\{ T : T > T(\lambda_0) \ ; \lambda_0 \} = \gamma.
\]

Since \( \frac{f}{\xi} T \sim \chi^2(r,f) \) under \( H_0 \), we get \( T(\lambda_0) = \frac{f}{\xi} \chi^2(r,f) \). Hence the test (4.10) defines a one-sided F-test which is identical to (3.7), the size-\( \gamma \) UMPI test of \( H_0 \) against \( H_A : \lambda > \lambda_0 \).

On the other hand, when \( \gamma \) is greater than or equal to
\[
\Pr\{ T : S_1(T) < (1+\lambda_0 \Delta)^{-r/2}; \lambda_0 \},
\]

the size-\( \gamma \) LRI test becomes
\[
\Phi_{LRI}(T) = \begin{cases} 
1, & \text{if } T > T_2(\lambda_0) \text{ or } < T_1(\lambda_0), \\
0, & \text{otherwise},
\end{cases} \quad (4.11)
\]

where \( T_1(\lambda_0) \) and \( T_2(\lambda_0) \) are chosen so that \( T_1(\lambda_0) < T_2(\lambda_0) \),

\[
S_1(T_1(\lambda_0)) = S_1(T_2(\lambda_0)),
\]

and
\[
\Pr\{ T : T < T_1(\lambda_0) \text{ or } > T_2(\lambda_0) ; \lambda_0 \} = \gamma.
\]

Clearly, \( \Phi_{LRI}(T) \) of (4.11) defines a two-sided F-test.

A size-\( \gamma \) non-randomized LRI test of \( H_0 : \lambda = 0 \) based on \( S_1(T) \) may not exist for some values of \( \gamma \). When \( \lambda_0 = 0 \), (4.9) reduces to
\[
S_1(T) = \begin{cases} 
S_3(T), & \text{if } \frac{f}{\xi} T > 1, \\
1, & \text{otherwise}.
\end{cases} \quad (4.12)
\]

Thus, \( H_0 \) will be accepted with probability at least
\[
\Pr\{ T : S_1(T) = 1; \lambda_0 \} = \Pr\{ \chi^2(r,f) \leq 1 \}
\]

under \( H_0 \). That is, the size of the test must be at most
\[
\Pr\{ \chi^2(r,f) \geq 1 \}.
4.1.3. Likelihood ratio invariant test of $H_0: \lambda = \lambda_0$ against $H_A: \lambda \neq \lambda_0$ when $A = A_2$ and $\Delta_i = \Delta$ for all $i$

The above discussion may be modified to cover the case when $A = A_2$ and $\Delta_i = \Delta$ for all $i$. Let $\lambda = \max\{-1/\omega, \lambda_3\}$. Then, the null hypothesis is rejected when the observed value of $S_2(T)$ is sufficiently small, where

$$S_2(T) = \begin{cases} S_3(T), & \text{if } \frac{f}{r} \sum_1^r x_i^2 > 1 - \frac{\Delta}{\omega}, \\ \frac{1 - \frac{\Delta}{\omega}}{1 + \lambda_0 \Delta} \left( \frac{1 + \frac{f}{r} \sum_1^r x_i^2}{1 - \frac{\Delta}{\omega}} \right)^{\frac{r}{2}}, & \text{otherwise} \end{cases}$$

$$S_3(T) = \begin{cases} S_3(T), & \text{if } \frac{f}{r} (1 + \lambda_0 \Delta) T > 1 - \frac{\Delta}{\omega}, \\ \frac{1 - \frac{\Delta}{\omega}}{1 + \lambda_0 \Delta} \left( \frac{1 + \frac{1 + \lambda_0 \Delta}{1 - \frac{\Delta}{\omega}} \frac{r}{T}}{1 + T} \right)^{\frac{r + f}{2}}, & \text{otherwise}. \end{cases}$$

We have that $\lim_{T \to \infty} S_2(T) = 0$, $S_2(T) > S_3(T)$ for $\frac{f}{r} (1 + \lambda_0 \Delta) T < 1 - \frac{\Delta}{\omega}$, and $\lim_{T \to 0} S_2(T) = \left( \frac{1 - \frac{\Delta}{\omega}}{1 + \lambda_0 \Delta} \right)^{\frac{r}{2}} > 0$. Therefore, the LRI test of $H_0$ based on $S_2(T)$ could be either a one-sided or a two-sided
F-test and a one-sided F-test is identical to the UMPI test of $H_0$ against $H_A: \lambda > \lambda_0$.

4.2. Most Powerful Invariant Test of $H_0: \lambda = \lambda_0$ against a Mixed Alternative

The LRI test function given in (4.5) is also applicable when the parameter space is restricted to $[\lambda_1, \lambda_2]$ for some $\lambda_1$ and $\lambda_2$ such that $-1/A^* < \lambda_1 < \lambda_2 < \infty$. When the parameter space is so restricted, there is an interesting two-sided test of $H_0: \lambda = \lambda_0$ where $\lambda_0 \in [\lambda_1, \lambda_2]$. This test is the one which, subject to $\beta_\phi(\lambda_0) = \gamma$, maximizes a weighted average $c_1 \beta_\phi(\lambda_1) + c_2 \beta_\phi(\lambda_2)$ of $\beta_\phi(\lambda_1)$ and $\beta_\phi(\lambda_2)$ for pre-assigned non-negative coefficients $c_1$ and $c_2$ such that $c_1 + c_2 = 1$. Some special cases are (i) $(c_1, c_2) = (0, 1)$, in which case the test reduces to the MPI test (3.2); (ii) $(c_1, c_2) = (0.5, 0.5)$, in which case there is equal emphasis on the departures from $H_0$ in either direction.

Note that since $\beta_\phi(\lambda) = \int_{\mathcal{X}} \phi(x) k(x; \lambda) dx$, $c_1 \beta_\phi(\lambda_1) + c_2 \beta_\phi(\lambda_2)$ is equal to $\int_{\mathcal{X}} \phi(x) [c_1 k(x; \lambda_1) + c_2 k(x; \lambda_2)] dx$. Clearly, $c_1 k(x; \lambda_1) + c_2 k(x; \lambda_2)$ is a mixture p.d.f. (e.g., Robbins and Pitman, 1949).

It follows from the generalized Neyman-Pearson Lemma that a size-$\gamma$ invariant test maximizing $c_1 \beta_\phi(\lambda_1) + c_2 \beta_\phi(\lambda_2)$ is

$$\phi_{\text{mixx}}(x) = \cdots$$
where $K_{Y}(\lambda_0, \lambda_1, \lambda_2)$ is determined so that

$$\Pr\{ X : \phi_{\text{max}}(X) = 1; \lambda_0 \} = \gamma.$$  

The test (4.13) could be extended to the case where we wish to emphasize the power of detecting the deviation from the null hypothesis at more than two alternative values of $\lambda$. To do so, it suffices to find the most powerful invariant test of $H_0 : \lambda = \lambda_0$ against the alternative hypothesis that the true p.d.f. is a certain mixture of more than two p.d.f.'s. However, the computations of the critical points for these tests would present a difficult, if not intractable, problem.

As $\lambda_2$ approaches infinity, the test (4.13) reduces to the MPI test of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda = \lambda_1$. Hence, letting $\lambda_2 \to \infty$ is equivalent to taking $(c_1, c_2) = (1, 0)$.

Based on (4.13), a $100(1 - \gamma)$% confidence set is

$$\left\{ \left[ 1 + \sum_{i=1}^{r+1} \frac{x_i^2}{1 + \lambda_i \Delta_1} \right] \frac{c_1 + c_2}{\prod_{i=1}^{r}(1 + \lambda_i \Delta_1)} \right\} > K_{Y}(\lambda_0, \lambda_1, \lambda_2).$$

(4.13)
4.3. Other Two-sided Tests Obtained by Combining
Two One-sided Tests

A two-sided test may be constructed by combining two
one-sided tests (Cox and Hinkley, 1974, p. 105). As applied to
the problem of testing $H_0 : \lambda = \lambda_0$ against $H_A : \lambda = \lambda_1$ or $\lambda = \lambda_2$, this would consist of combining tests of (i) $H_0 : \lambda = \lambda_0$ against $H_1 : \lambda = \lambda_1$ and (ii) $H_0 : \lambda = \lambda_0$ against $H_2 : \lambda = \lambda_2$. Rejection
of $H_0$ in testing either (i) or (ii) is a cause for rejection of $H_0$
in the two-sided problem. In Subsection 4.3.1, we will combine
two MPI tests of (i) and (ii).

Similarly, the problem of testing $H_0 : \lambda = \lambda_0$ against
$H_A : \lambda \neq \lambda_0$ consists of two one-sided problems: (i) $H_0 : \lambda = \lambda_0$
against $H_1 : \lambda < \lambda_0$ and (ii) $H_0 : \lambda = \lambda_0$ against $H_2 : \lambda > \lambda_0$.
Rejecting $H_0$ in testing either (i) or (ii) results in rejection of
$H_0$ in the two-sided problem. We show how to combine two one-sided
UMPI, LMPI, or Wald's tests of $H_0$ in Subsections 4.3.2, 4.3.3, and 4.3.4, respectively.

Let $\phi_j(x)$ represent a size-$\gamma_j$ test of $H_0$ against the alternative $H_j$, $j = 1, 2$, where $\gamma_1$ and $\gamma_2$ are non-negative and $\gamma_1 + \gamma_2 \leq 1$. Correspondingly, let $A_j(\lambda_0)$ represent the acceptance region of the test $\phi_j(x)$, $j = 1, 2$. Using the duality relationship (1.10), a $100(1 - \gamma_j)$% confidence set for $\lambda$ is

$$C_j(x) = \{ \lambda \in \Lambda_3 : x \in A_j(\lambda) \}, \ j = 1, 2.$$ 

Since the acceptance region of the two-sided test of $H_0$ constructed by combining $\phi_1(x)$ and $\phi_2(x)$ is clearly $A_c(\lambda_0) = A_1(\lambda_0) \cap A_2(\lambda_0)$, a confidence set for $\lambda$ is given by

$$C_c(x) = \{ \lambda \in \Lambda_3 : x \in A_c(\lambda) \} = \{ \lambda \in \Lambda_3 : x \in A_1(\lambda) \} \cap \{ \lambda \in \Lambda_3 : x \in A_2(\lambda) \} = C_1(x) \cap C_2(x).$$

It is important to note that, in general, $C_c(x)$ is "conservative" in the sense that its confidence level is greater than or equal to $100(1 - \gamma_1 - \gamma_2)$%. This is easily seen by applying the Bonferroni inequality (see, e.g., Bickel and Docksum, 1977, p. 439). $C_c(x)$ is an exact $100(1 - \gamma_1 - \gamma_2)$% confidence set for $\lambda$ if and only if the complements of $C_1(x)$ and $C_2(x)$ are disjoint, that is, if the rejection regions of $\phi_1(x)$ and $\phi_2(x)$ are disjoint for every $\lambda_0 \in \Lambda_3$. 
4.3.1. Combining most powerful invariant tests of $H_0 : \lambda = \lambda_0$

against $H_1 : \lambda = \lambda_1$ and of $H_0 : \lambda = \lambda_0$ against $H_2 : \lambda = \lambda_2$

Let $\phi_{\text{MPI}, j}(x)$ denote the size-$\gamma_j$ MPI test of $H_0 : \lambda = \lambda_0$

against $H_j : \lambda = \lambda_j$ as defined by (3.2), $j = 1, 2$. A combined
test function of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda = \lambda_1$ or $\lambda = \lambda_2$ is

$$
\phi_{\text{MPI}, C}(x) = \begin{cases} 
1, & \text{if } \phi_{\text{MPI}, 1}(x) = 1 \text{ or } \phi_{\text{MPI}, 2}(x) = 1, \\
0, & \text{otherwise}
\end{cases}
$$

$$
\begin{align*}
&= \begin{cases} 
1, & \text{if } \frac{1 + \sum_{1}^{x_1^2}}{1 + \lambda_1 \Delta_1} > K_1^* (\lambda_0, \lambda_1) \text{ or } \frac{1 + \sum_{1}^{x_1^2}}{1 + \lambda_2 \Delta_1} > K_2^* (\lambda_0, \lambda_2), \\
0, & \text{otherwise}
\end{cases}
\end{align*}
$$

(4.14)

where $K_j^* (\lambda_0, \lambda_j)$ is determined so that

$$
\Pr \{ x : \phi_{\text{MPI}, j}(x) = 1 ; \lambda_0 \} = \gamma_j, \quad j = 1, 2.
$$

(4.15)

Note that $\phi_{\text{MPI}, C}(x)$ has a size less than or equal to $\gamma_1 + \gamma_2$
as

$$
\Pr \{ x : \phi_{\text{MPI}, C}(x) = 1 ; \lambda_0 \} \leq \sum_{1}^{2} \Pr \{ x : \phi_{\text{MPI}, j}(x) = 1 ; \lambda_0 \}.
$$

The corresponding confidence set for $\lambda$ is
\[ C_{FMPI,c}(x) = \begin{cases} \lambda \in \Lambda_3 : & \frac{1}{1+\sum_1^r \frac{x_1^2}{1+\lambda_1\Delta}} \leq K_{\gamma_1}(\lambda,\lambda_1) \\ \frac{1}{1+\sum_1^r \frac{x_1^2}{1+\lambda_2\Delta}} \leq K_{\gamma_2}(\lambda,\lambda_2) \end{cases} \] (4.16)

whose confidence level is no less than \(100(1 - \gamma_1 - \gamma_2)\%\).

4.3.2. Combining uniformly most powerful invariant tests of \(H_0 : \lambda = \lambda_0\) against one-sided alternative exists when \(\Delta_i = \Delta\) for all \(i\). Let \(\phi_{UMP\_1}(x)\) denote the size-\(\gamma_1\) UMPI test of \(H_0 : \lambda < \lambda_0\) and \(\phi_{UMP\_2}(x)\) the size-\(\gamma_2\) UMPI test of \(H_0 : \lambda > \lambda_0\), respectively. A combined test function of \(H_0 : \lambda = \lambda_0\) against \(H_A : \lambda \neq \lambda_0\) is

\[
\phi_{UMP\_C}(x) = \begin{cases} 1, & \text{if } \phi_{UMP\_1}(x) = 1 \text{ or } \phi_{UMP\_2}(x) = 1, \\ 0, & \text{otherwise} \end{cases}
\]
When \( \gamma_1 + \gamma_2 \leq 1 \), \( F_{\gamma_2}(r, f) \geq F_{1-\gamma_1}(r, f) \) and the size of the test \( \phi_{\text{UMPI}, c}(x) \) is \( \gamma_1 + \gamma_2 \).

Consequently, a \( 100(1 - \gamma_1 - \gamma_2) \)% confidence interval is

\[
C_{\text{UMPI}, c}(x) = \left\{ \lambda \leq \lambda_0 : F_{1-\gamma_1}(r, f) \leq \frac{\sum x_i^2}{1 + \lambda_0 A} \leq F_{\gamma_2}(r, f) \right\}.
\]

Note that in the special case where \( \gamma_1 \) and \( \gamma_2 \) are chosen to satisfy the equation (4.8), the test (4.17) is identical to the LRI test (4.7).

### 4.3.3. Combining locally most powerful invariant tests of \( H_0 : \lambda = \lambda_0 \) against \( H_1 : \lambda < \lambda_0 \) and of \( H_0 : \lambda = \lambda_0 \) against \( H_2 : \lambda > \lambda_0 \)

Let \( \phi_{\text{LMP}, i}(x) \) represent the size-\( \gamma_1 \) LMP test of \( H_0 \) against \( H_1 : \lambda < \lambda_0 \) and \( \phi_{\text{LMP}, i}(x) \) the size-\( \gamma_2 \) LMP test of \( H_0 \) against \( H_2 : \lambda > \lambda_0 \), respectively. A combined test of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda \neq \lambda_0 \) is

\[
\phi_{\text{LMP}, c}(x) = \begin{cases} 
1, & \text{if } \phi_{\text{LMP}, i}(x) = 1 \text{ or } \phi_{\text{LMP}, i}(x) = 1, \\
0, & \text{otherwise}
\end{cases}
\]
\[ G, \text{ otherwise} \]

where \( K_{1-\gamma_1}(\lambda_0) \) and \( K_{\gamma_2}(\lambda_0) \) satisfy (3.18) for \( \gamma = 1 - \gamma_1 \) and \( \gamma = \gamma_2 \), respectively. Clearly, the size of the test \( \phi^{LMPI_c}(x) \) is exactly \( \gamma_1 + \gamma_2 \).

The corresponding 100(1 - \( \gamma_1 - \gamma_2 \))% confidence set is

\[
C_{LMPI,c}(x) = \left\{ \lambda \in \Lambda_3 : K_{1-\gamma_1}(\lambda) \leq \frac{r}{1 + \sum_{i=1}^{r} \frac{x_i^2}{1 + \lambda_0 A_i}} \leq K_{\gamma_2}(\lambda) \right\}.
\]

As noted earlier, the one-sided LMPI test of \( H_0 : \lambda = \lambda_0 \) is a UMPI test of \( H_0 \) when \( A_i = \Delta \) for all \( i \). That is, the size-\( (\gamma_1 + \gamma_2) \) test (4.18) is identical to the size-\( (\gamma_1 + \gamma_2) \) test (4.17) when \( \Delta_i = \Delta \) for all \( i \).

4.3.4. Combining Wald’s tests of \( H_0 : \lambda = \lambda_0 \) against \( H_1 : \lambda < \lambda_0 \) and of \( H_0 : \lambda = \lambda_0 \) against \( H_2 : \lambda > \lambda_0 \)

Let \( \phi_{w_1}(x) \) represent the size-\( \gamma_1 \) Wald’s test of \( H_0 \) against \( H_1 : \lambda < \lambda_0 \) and \( \phi_{w_2}(x) \) the size-\( \gamma_2 \) Wald’s test of \( H_0 \) against \( H_2 : \lambda > \lambda_0 \), respectively. A combined test of \( H_0 : \lambda = \lambda_0 \) against \( H_\Lambda : \lambda \neq \lambda_0 \)
\( \lambda \neq \lambda_0 \) is

\[
\phi_{w,c}(x) = \begin{cases} 
1, & \text{if } \phi_{w,1}(x) = 1 \text{ or } \phi_{w,2}(x) = 1, \\
0, & \text{otherwise}
\end{cases}
\]

\[= \begin{cases} 
1, & \text{if } Q(\lambda_0) < F_{1-\gamma_1}(r,f) \text{ or } > F_{\gamma_2}(r,f) \\
0, & \text{otherwise.}
\end{cases}
\] (4.20)

Clearly, the size of the test \( \phi_{w,c}(x) \) is \( \gamma_1 + \gamma_2 \). The size-\( (\gamma_1 + \gamma_2) \) test (4.20) is identical to the size-\( (\gamma_1 + \gamma_2) \) test (4.17) when \( \Delta_1 = \Delta \) for all i.

The corresponding 100(1 - \( \gamma_1 - \gamma_2 \))% confidence set is

\[
C_{w,c}(x) = \{ \lambda \in \Lambda_1 : F_{1-\gamma_1}(r,f) \leq Q(\lambda) \leq F_{\gamma_2}(r,f) \}. \] (4.21)

Since \( Q(\lambda) \) is strictly decreasing in \( \lambda \), (4.21) defines an interval on the real line. More specifically, the solutions to the equations

\[
Q(\lambda) = F_{1-\gamma_1}(r,f) \quad \text{and} \quad Q(\lambda) = F_{\gamma_2}(r,f)
\] (4.22)

are the upper and lower limits of the confidence interval, respectively. The equations (4.22) can be solved iteratively as discussed in Section 6.4.

In addition to the exact confidence interval (4.21), Harville and Fenech (1985) considered the approximate confidence intervals that result from approximating the pivotal quantity \( Q(\lambda) \) by either

\[
Q_1(\lambda) = \frac{f}{r} \cdot \frac{1}{1+\lambda\hat{\lambda}} \cdot \frac{r}{1} \sum x_i^2
\] (4.23)
where \( \bar{\Delta} = \frac{\sum I}{1} \Delta_1 / r \), or

\[
Q_2(\lambda) = \frac{1}{x} \cdot \frac{\sum I}{\Delta^{-1} + \lambda} \frac{x^2}{\Delta_1}
\]  

(4.24)

where \( \Delta^{-1} = \frac{1}{r} \sum \frac{1}{\Delta_1} \), respectively. Approximation (4.23) leads to the approximate 100(1 - \( \gamma_1 \) - \( \gamma_2 \))% confidence interval

\[
C_{w,c}^1(\lambda) = \{ \lambda \in A_3 : F_{1-\gamma_1}(r,f) \leq Q_1(\lambda) \leq F_{\gamma_2}(r,f) \}, \quad (4.25)
\]

and approximation (4.24) to the approximate 100(1 - \( \gamma_1 \) - \( \gamma_2 \))% confidence interval

\[
C_{w,c}^2(\lambda) = \{ \lambda \in A_3 : F_{1-\gamma_1}(r,f) \leq Q_2(\lambda) \leq F_{\gamma_2}(r,f) \}. \quad (4.26)
\]

As discussed by Harville and Fenech (1985), the computations required to form the approximate confidence intervals (4.25) and (4.26) are much less extensive than those required to form the exact confidence interval (4.21).

The second approximation was proposed by Thomas and Hultquist (1978) in the case of the unbalanced one-way random model. Six methods of approximating the upper and lower confidence limits of (4.21), including (4.25) and (4.26), were considered (in the context of the unbalanced one-way random model) by Burdick, Masood, and Graybill (1986) in a simulation study.
5. OPTIMALITY OF SIMILARITY OF INVARIANT TESTS

In Chapters 3 and 4, attention was restricted to hypothesis tests that are invariant. In this chapter, we use the principle of similarity, together with that of translation invariance, to "reduce" the class of hypothesis tests. The discussion involves the likelihood function of \((\sigma_e^2, \lambda)\) associated with, or equivalently, the p.d.f. of, the maximal invariant \((t, u)\).

5.1. Generalized Likelihood Ratio Translation-invariant Similar Test of \(H_0 : \lambda = \lambda_0\) against \(H_A : \lambda \neq \lambda_0\)

Let \(L(\sigma_e^2, \lambda; t, u)\) represent the likelihood function of \((\sigma_e^2, \lambda)\) associated with the p.d.f. \(g(t, u; \sigma_e^2, \lambda)\) of \((t, u)\). Then,

\[
\log L(\sigma_e^2, \lambda; t, u) = \frac{C}{2\sigma_e^2} \left[ \sum_{i=1}^{r} \frac{t_i^2}{1 + \lambda_1} + SS_e \right] - \frac{x_f + f}{2} \log \sigma_e^2
\]

\[
- \frac{1}{2} \sum_{i=1}^{r} \log(1 + \lambda_1) \tag{5.1}
\]

where

\[
C = -\frac{x_f + f}{2} \log 2\pi.
\]

In this section, we show that the generalized likelihood ratio test of \(H_0 : \lambda = \lambda_0\) against \(H_A : \lambda \neq \lambda_0\), based on \(L(\sigma_e^2, \lambda; t, u)\), turns out to be the generalized likelihood ratio translation-invariant similar test.

The partial derivative of (5.1) with respect to \(\sigma_e^2\) is
The solution to the equation (in $\sigma_e^2$) obtained by equating (5.2) to zero is

$$
\hat{\sigma}_e^2(\lambda) = \frac{SS_e + \sum_{i=1}^{n} \frac{t_i^2}{1+\lambda \Delta_i}}{r + f}.
$$

(5.3)

It is straightforward to show that

$$
\frac{\partial^2}{\partial \sigma_e^2} \log \mathcal{L}(\sigma_e^2, \lambda;t,u) \bigg|_{\sigma_e^2 = \hat{\sigma}_e^2(\lambda)} < 0.
$$

Therefore, for any particular value of $\lambda$, $\mathcal{L}(\sigma_e^2, \lambda;t,u)$ attains its maximum at $\sigma_e^2 = \hat{\sigma}_e^2(\lambda)$.

Substituting $\hat{\sigma}_e^2(\lambda)$ for $\sigma_e^2$ in $\mathcal{L}(\sigma_e^2, \lambda;t,u)$, we obtain the so-called concentrated likelihood function

$$
\mathcal{L}_c(\lambda;t,u) = \frac{\exp\left\{ -\frac{1}{2} (r + f) \right\} }{[2\pi \hat{\sigma}_e^2(\lambda)]^{(r+f)/2} \prod_{i=1}^{r} \left( 1 + \lambda \Delta_i \right)^{1/2}}
$$

$$
= q(SS_e).\mathcal{L}(\lambda;x),
$$

where

$$
q(SS_e) = \frac{\Gamma\left( \frac{f}{2} \right)}{\Gamma\left( \frac{r+f}{2} \right)} \frac{\exp\left\{ -\frac{1}{2} (r + f) \right\} }{\pi^{f/2}} \left( \frac{r+f}{2SS_e} \right)^{(r+f)/2}.
$$

and $\mathcal{L}(\lambda;x)$ is the likelihood function of $\lambda$ associated with $x$ given
by (3.1). Note that, when regarded as a function of \( \lambda \), \( \mathcal{L}_c(\lambda; t, u) \) is proportional to \( \xi(\lambda; x) \), implying in particular that any value of \( \lambda \) that maximizes one of these functions maximizes the other as well.

Let \( \Phi_0 = \{ (\sigma_e^2, \lambda) : \sigma_e^2 > 0, \lambda = \lambda_0 \} \) where \( \lambda_0 \in \Lambda \). Then, the test statistic of the generalized likelihood ratio translation-invariant similar (GLRS) test of \( \mathcal{H}_0 : \lambda = \lambda_0 \) against \( \mathcal{H}_A : \lambda \neq \lambda_0 \) is

\[
\sup_{\Phi_0} \frac{\mathcal{L}(\sigma_e^2, \lambda; t, u)}{\mathcal{L}(\sigma_e^2, \lambda; t, u)} = \sup_{\sigma_e^2 > 0} \frac{\mathcal{L}(\sigma_e^2, \lambda_0; t, u)}{\sup_{\lambda \in \Lambda} \sup_{\sigma_e^2 > 0} \mathcal{L}(\sigma_e^2, \lambda; t, u)} = \frac{\mathcal{L}_c(\lambda_0; t, u)}{\sup_{\lambda \in \Lambda} \mathcal{L}_c(\lambda; t, u)} = \frac{\xi(\lambda_0; x)}{\xi(\tilde{\lambda}; x)}
\]

where \( \tilde{\lambda} \) is such that \( \xi(\tilde{\lambda}; x) = \sup_{\lambda' \in \Lambda} \xi(\lambda'; x) \).

The statistic (5.4) is identical to the test statistic of the LRI test of \( \mathcal{H}_0 : \lambda = \lambda_0 \) against \( \mathcal{H}_A : \lambda \neq \lambda_0 \) given in Section 4.1. Thus, the GLRS test of \( \mathcal{H}_0 \) against \( \mathcal{H}_A \) coincides with the LRI test.
5.2. Most Powerful Translation-invariant Similar Test of

\[ H_0 : \lambda = \lambda_0 \] against \[ H_A : \lambda = \lambda_1 \ (\lambda_1 \neq \lambda_0) \]

For the purpose of characterizing the class of translation-invariant similar tests of the composite null hypothesis \( H_0 : \lambda = \lambda_0 \), we seek a sufficient statistic for \{ \( g(t, u; \sigma^2_e, \lambda) \) \} under \( H_0 \). It is clear from (2.1) that a sufficient statistic for \{ \( g(t, u; \sigma^2_e, \lambda_0) \) \} is

\[
\hat{\sigma}^2_e(\lambda_0) = SS_e + \sum_1^r \frac{t_1^2}{1 + \lambda_0 \Delta_1}.
\]  

(5.5)

Moreover, since \{ \( g(t, u; \sigma^2_e, \lambda_0) \) \} constitutes an exponential family whose parameter space \{ \( \sigma^2_e : \sigma^2_e > 0 \) \} contains an open set on the real line, the family of distributions of \( \hat{\sigma}^2_e(\lambda_0) \) is complete, and hence boundedly complete (see, e.g., Lehmann, 1959, p. 132). Therefore, according to Lemma 1.5, we may concentrate on the class of translation-invariant tests with Neyman structure with respect to \( \sigma^2_e(\lambda_0) \) in order to construct a most powerful translation-invariant similar (MPS) test.

To be specific, we regard the composite alternative hypothesis \( H_A : \lambda = \lambda_1 \) as a class of simple hypotheses of the form \( H'_A : \lambda = \lambda_1, \sigma^2_e = \sigma^2_{e,1} \). By the Neyman-Pearson Lemma, the test statistic of an MPS test of \( H_0 \) against \( H'_A \) for some \( \sigma^2_{e,1} \) and \( \lambda_1 \) is given by
\[
\frac{g(t,u;\sigma^2_{e,1},\lambda_1)}{g(t,u;\sigma^2_{e,0})} = \exp\left\{ -\frac{1}{2\sigma^2_{e,1}} \left[ SS_e + \sum_{i=1}^{r} \frac{t_i^2}{1+\lambda_1\Delta_1} \right] \right\} \left(\sigma^2_{e,1}\right)^{(r+f)/2} \frac{t_i}{\prod_{i=1}^{r}(1+\lambda_0\Delta_1)^{1/2}} \exp\left\{ -\frac{1}{2\sigma^2_{e}} \left[ SS_e + \sum_{i=1}^{r} \frac{t_i^2}{1+\lambda_0\Delta_1} \right] \right\}
\]

\[
= B(\bar{\sigma}^2_{e}(\lambda_0)).\exp\left\{ -\frac{\bar{\sigma}^2_{e}(\lambda)}{2\sigma^2_{e,1}} \right\},
\]

where

\[
B(\bar{\sigma}^2_{e}(\lambda_0)) = \left(\frac{\sigma^2_{e}}{\bar{\sigma}^2_{e,1}}\right)^{\frac{r+f}{2}} \prod_{i=1}^{r} \frac{1}{1+\lambda_0\Delta_1} \exp\left\{ -\frac{\bar{\sigma}^2_{e}(\lambda_0)}{2\sigma^2_{e}} \right\},
\]

which is a constant depending on \((t,u)\) only through the value of \(\bar{\sigma}^2_{e}(\lambda_0)\). Thus, the test is to reject \(H_0\) if the observed value of \(-\bar{\sigma}^2_{e}(\lambda_1)\) is sufficiently large, or equivalently, if the observed value of \(\bar{\sigma}^2_{e}(\lambda_0)/\bar{\sigma}^2_{e}(\lambda_1)\) is sufficiently large. That is, the size-\(\gamma\) MPS test of \(H_0\) against \(H_A\) is given by the function

\[
\phi(t,u) = \begin{cases} 
1, & \text{if } \bar{\sigma}^2_{e}(\lambda_0)/\bar{\sigma}^2_{e}(\lambda_1) > K_{\gamma}(\lambda_0,\lambda_1,\bar{\sigma}^2_{e}(\lambda_0)), \\
0, & \text{otherwise},
\end{cases}
\tag{5.6}
\]

where \(K_{\gamma}(\lambda_0,\lambda_1,\bar{\sigma}^2_{e}(\lambda_0))\) is chosen so that
for every value of $\sigma^2_e(\lambda_o)$ except for a set of zero probability.

Clearly, the critical region of this test is independent of the choice of $\sigma^2_{e,1}$. Thus, (5.6) and (5.7) define the size-$\gamma$ MPS test of $H_0$ against $H_A$.

Observe that the test statistic can be expressed in terms of $x$ as

$$\frac{\sigma^2_e(\lambda_o)}{\sigma^2_e(\lambda_1)} = \frac{SS_e + \sum \frac{t_i^2}{1+\lambda_o \Delta_i}}{SS_e + \sum \frac{t_i^2}{1+\lambda_1 \Delta_i}} = 1 + \frac{\sum \frac{x_i^2}{1+\lambda_0 \Delta_i}}{\sum \frac{x_i^2}{1+\lambda_1 \Delta_i}}.$$  

Since the distribution of $x$ is free of $\sigma^2$ and $\sigma^2_e(\lambda_o)$ is a boundedly complete and sufficient statistic for $\sigma^2$ under $H_0$, it follows from Basu's (1955, 1958) theorem that $\frac{\sigma^2_e(\lambda_o)}{\sigma^2_e(\lambda_1)}$ is statistically independent of $\sigma^2_e(\lambda_o)$ under $H_0$. Consequently, $K(\lambda_o, \lambda_1, \sigma^2_e(\lambda_o))$ does not depend on $\sigma^2_e(\lambda_o)$, and the size-$\gamma$ MPS test of $H_0$ against $H_A$ is given by

$$\phi_{MPS}(x) = \begin{cases} 
1, & \text{if } 1 + \sum \frac{x_i^2}{1+\lambda_0 \Delta_i} > K_\gamma(\lambda_o, \lambda_1), \\
1 + \sum \frac{x_i^2}{1+\lambda_1 \Delta_i}, & \text{otherwise},
\end{cases}$$
where $K_\gamma(\lambda_0, \lambda_1)$ is chosen so that

$$\Pr\{ \mathbf{x} : \phi_{\text{MPS}}(\mathbf{x}) = 1; \lambda_0 \} = \gamma.$$ 

The MPS test is identical to the MPI test given by (3.2). In the special case where $\Delta_i = \Delta, i = 1, \ldots, r$, there exists a uniformly most powerful translation-invariant similar test of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda > \lambda_0$, which is of course identical to the UMPI test given by (3.7).

The MPS test was given previously by Spjøtvoll (1967, Section 3) for the one-way random model (1.3).

### 5.3. A Locally Most Powerful Translation-invariant Similar Test of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda > \lambda_0$

The composite alternative hypothesis $H_A : \lambda > \lambda_0$ can be regarded as a class of composite hypotheses of the form $H'_A : \lambda > \lambda_0, \sigma_e^2 = \sigma_e^2$. Let $\beta_\phi(\lambda, \sigma_e^2)$ represent the power function of a translation-invariant similar test $\phi(t, u)$ of $H_0$, i.e., let

$$\beta_\phi(\lambda, \sigma_e^2) = \int \phi(t, u) \cdot g(t, u; \sigma_e^2, \lambda) dt du.$$ 

A size-$\gamma$ locally most powerful translation-invariant similar (LMPS) test of $H_0$ against $H'_A$ is obtained by maximizing $\frac{\partial}{\partial \lambda} \beta_\phi(\lambda, \sigma_e^2)$ subject to the restriction $\beta_\phi(\lambda_0, \sigma_e^2) = \gamma$. Using the same technique employed in proving Lemma 3.1, it can be shown that

$$\frac{\partial}{\partial \lambda} \beta_\phi(\lambda, \sigma_e^2) = \int \phi(t, u) \cdot \frac{\partial}{\partial \lambda} g(t, u; \sigma_e^2, \lambda) dt du.$$
Hence, according to the generalized Neyman-Pearson Lemma, the test statistic of an LMPS test for testing $H_0$ against $H_A'$ is

$$
\frac{\partial}{\partial \lambda} \frac{g(t,u;\sigma_e^2,1,\lambda)}{g(t,u;\sigma_e^2,\lambda_0)} \bigg|_{\lambda=\lambda_0}

= \left(\frac{\sigma_e^2}{\sigma_e^2,1} \right)^{\frac{r+f}{2}} \exp\left\{ -\frac{\sigma_e^2(\lambda_0)}{2\sigma_e^2} \right\} \cdot \exp\left\{ -\frac{\sigma_e^2(\lambda_0)}{2\sigma_e^2_1} \right\} \bigg( \frac{r}{2} \sum_1^r \frac{\Delta_1 t_1^2}{1+\lambda_0 \Delta_1^2} \bigg) - \frac{1}{2} \sum_1^r \frac{\Delta_1}{1+\lambda_0 \Delta_1} \bigg),

$$

where $\sigma_e^2(\lambda_0)$ is as given in (5.5) and

$$
B(\sigma_e^2(\lambda_0)) = \left(\frac{\sigma_e^2}{\sigma_e^2,1} \right)^{\frac{r+f}{2}} \exp\left\{ -\frac{\sigma_e^2(\lambda_0)}{2\sigma_e^2} \right\} \bigg[ \frac{1}{\sigma_e^2} - \frac{1}{\sigma_e^2_1} \bigg],

$$

a "constant" depending on $(t,u)$ only through the value of $\sigma_e^2(\lambda_0)$. Therefore, we reject $H_0$ if the observed value of

$$
\sum_1^r \frac{\Delta_1 t_1^2}{1+\lambda_0 \Delta_1^2}

$$

is sufficiently large, or equivalently, if the observed value of

$$
\sum_1^r \frac{\Delta_1 t_1^2}{1+\lambda_0 \Delta_1^2} / \sigma_e^2(\lambda_0)

$$

is sufficiently large. That is, the size-$\gamma$ LMPS test of $H_0$ against $H_A'$ is
\[ \phi(t,u) = \begin{cases} 1, & \text{if } \sum_{i=1}^{r} \frac{\Delta_i t_i^2}{1 + \lambda_0 \Delta_i} > K_\gamma(\lambda_0, \sigma_e^2(\lambda_0)) \\ 0, & \text{otherwise}, \end{cases} \quad (5.8) \]

where \( K_\gamma(\lambda_0, \sigma_e^2(\lambda_0)) \) is chosen so that

\[ \mathbb{E}[\phi(t,u) | \sigma_e^2(\lambda_0); \lambda_0] = \gamma \quad (5.9) \]

for every value of \( \sigma_e^2(\lambda_0) \) except for a set of zero probability.

Since the test defined by (5.8) and (5.9) does not depend on the choice of \( \sigma_e^2, \) the test is actually a size-\( \gamma \) LMPS test of \( H_0 \) against \( H_A. \)

Expressed in terms of \( x, \) the test statistic of (5.8) becomes

\[
\frac{r \sum_{i=1}^{r} \frac{\Delta_i t_i^2}{1 + \lambda_0 \Delta_i}}{\sigma_e^2(\lambda_0)} = \frac{r \sum_{i=1}^{r} \frac{\Delta_i t_i^2}{1 + \lambda_0 \Delta_i}}{SS_e + \sum_{i=1}^{r} \frac{t_i^2}{1 + \lambda_0 \Delta_i}} = \frac{r \sum_{i=1}^{r} \frac{\Delta_i x_i^2}{1 + \lambda_0 \Delta_i}}{1 + \sum_{i=1}^{r} \frac{x_i^2}{1 + \lambda_0 \Delta_i}}.
\]

Thus, by an argument similar to that employed in the previous section, \( K_\gamma(\lambda_0, \sigma_e^2(\lambda_0)) \) does not depend on \( \sigma_e^2(\lambda_0), \) and the size-\( \gamma \) LMPS test is
$\phi_{\text{LMPS}}(x) = \begin{cases} 
1, & \text{if } \sum_{i=1}^{r} \frac{\Delta_i x_i^2}{1 + \lambda_0 \Delta_i} > K_\gamma(\lambda_0), \\
1 + \sum_{i=1}^{r} \frac{x_i^2}{1 + \lambda_0 \Delta_i}, & \text{otherwise} 
\end{cases}$

where $K_\gamma(\lambda_0)$ is chosen so that

$\Pr\{ x : \phi_{\text{LMPS}}(x) = 1; \lambda_0 \} = \gamma$.

Clearly, the LMPS test is identical to the LMPI test given by (3.17). In the case where $\Delta_i = \Delta$, $i = 1, \ldots, r$, the uniformly most powerful translation-invariant similar test of $H_0$ against $H_A$ exists and is identical to the UMPI test of (3.7).
6. COMPUTING CRITICAL POINTS AND POWER FUNCTIONS OF IN Variant TESTS

The various invariant tests of \( H_0 : \lambda = \lambda_0 \) introduced in Chapters 3, 4, and 5 can be compared on the basis of their power functions. As discussed in Section 1.2, the power of a test to reject a false \( H_0 \) can be re-interpreted as the probability that the corresponding confidence set does not cover the false value \( \lambda_0 \).

Recall (from Lemma 3.1) that the power functions of invariant tests are continuous. We show in Section 6.1 that the power function of each of the following one-sided tests of \( H_0 \) is increasing in \( \lambda \): the MPI test against \( H_A : \lambda = \lambda_1 (\lambda_1 > \lambda_0) \), the LMPI test against \( H_A : \lambda > \lambda_0 \), and Wald's test against \( H_A : \lambda > \lambda_0 \). In Section 6.2, the power functions of invariant tests of \( H_0 \) against two-sided alternatives will be discussed.

In what follows, let \( z' = (z_1', \ldots, z_{r+1}) \) where the \( z_i' \)'s represent independent random variables such that \( z_i' \sim \chi^2(1), i = 1, \ldots, r \) and \( z_{r+1} \sim \chi^2(f) \). The power functions of the MPI, LMPI, and Wald's tests can be expressed in terms of the distribution of linear combinations of the \( r + 1 \) independent chi-square random variables \( z_1', \ldots, z_{r+1} \).

In Section 6.3, we present a result due to Imhof (1961), which is useful in numerically evaluating the distribution function of a linear combination of independent chi-square random
variables. Imhof’s procedure can be used to compute the power functions of invariant tests. In Section 6.4, we show that the bisection method may be used in conjunction with Imhof’s procedure to compute the critical points of the LMPI and MPI tests. For the two-sided tests \( \phi_{LPI}(x) \) and \( \phi_{MIXX}(x) \), computing the critical points presents a difficult problem. In Section 6.5, we propose that the critical points be approximated by conducting simulation studies. In Section 6.5, the nature of various confidence sets and the technique of computing are discussed.

6.1. Monotonic Power Functions of Invariant Tests of \( H_0 : \lambda = \lambda_0 \) against One-sided Alternatives

In what follows, the power functions of the MPI, LMPI, and Wald’s tests of \( H_0 \) against right-hand-side alternatives are shown to be increasing in \( \lambda \). A similar argument could be used to show that the MPI, LMPI, and Wald’s tests of \( H_0 \) against left-hand-side alternatives are decreasing in \( \lambda \).

6.1.1. The power function of the most powerful invariant test of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda = \lambda_1 (\lambda_1 > \lambda_0) \)

It follows from (3.2) that the power function of the size-\( \gamma \) MPI test of \( H_0 \) against \( H_A \) is
\[ \beta_{\text{MPI}}(\lambda) = \Pr \left\{ x : \frac{1 + \frac{x^2}{1 + \lambda_0 \Delta_1}}{1 + \frac{x^2}{1 + \lambda_1 \Delta_1}} > K_{\gamma}(\lambda_0, \lambda_1); \lambda \right\} \]

\[ = \Pr \left\{ z : \frac{r}{1 + \lambda \Delta_1} \frac{1 + \lambda_0 \Delta_1}{z_1} > K_{\gamma}(\lambda_0, \lambda_1) \right\} \quad (6.1) \]

and hence that

\[ \lim_{\lambda \to \infty} \beta_{\text{MPI}}(\lambda) = \Pr \left\{ z : \frac{r}{1 + \lambda \Delta_1} \frac{1 + \lambda_0 \Delta_1}{z_1} > K_{\gamma}(\lambda_0, \lambda_1) \right\}, \]

which is equal to one if and only if \((1 + \lambda_1 \Delta_1)/(1 + \lambda_0 \Delta_1) > K_{\gamma}(\lambda_0, \lambda_1)\) for all \(i\).

Let \(G(\lambda) = \left( z_{r+1} + \frac{r}{1 + \lambda_0 \Delta_1} \frac{1 + \lambda \Delta_1}{z_1} \right) \left( z_{r+1} + \frac{r}{1 + \lambda_1 \Delta_1} \frac{1 + \lambda \Delta_1}{z_1} \right) \).

We now show that \(G(\lambda)\) is increasing in \(\lambda\) and hence that \(\beta_{\text{MPI}}(\lambda)\) is increasing in \(\lambda\). The derivative of \(G(\lambda)\) is

\[ \frac{d}{d\lambda} G(\lambda) = \]

\[ \frac{r}{1 + \lambda \Delta_1} \frac{\Delta_1 z_1}{z_{r+1} + \frac{r}{1 + \lambda_1 \Delta_1} \frac{1 + \lambda \Delta_1}{z_1}} - \frac{r}{1 + \lambda \Delta_1} \frac{\Delta_1 z_1}{z_{r+1} + \frac{r}{1 + \lambda_0 \Delta_1} \frac{1 + \lambda \Delta_1}{z_1}} \]

\[ = \left( z_{r+1} + \frac{r}{1 + \lambda \Delta_1} \frac{1 + \lambda_1 \Delta_1}{z_1} \right)^2 \]
Since \( \lambda_1 > \lambda_0 \), the numerator of \( \frac{d}{d\lambda} G(\lambda) \) is clearly greater than

\[
\sum_{1}^{r} \frac{\Delta_1 z_i}{1+\lambda_0 \Delta_1} \cdot \sum_{1}^{r} \frac{1+\lambda \Delta_1}{1+\lambda \Delta_1} - \sum_{1}^{r} \frac{\Delta_1 z_i}{1+\lambda_0 \Delta_1} \cdot \sum_{1}^{r} \frac{1+\lambda \Delta_1}{1+\lambda \Delta_1} z_i
\]

\[
= \sum_{1}^{r} \frac{\Delta_1 z_i}{1+\lambda_0 \Delta_1} \cdot \sum_{1}^{r} \frac{z_i}{1+\lambda \Delta_1} - \sum_{1}^{r} \frac{\Delta_1 z_i}{1+\lambda_0 \Delta_1} \cdot \sum_{1}^{r} \frac{z_i}{1+\lambda \Delta_1}
\]

\[
= \sum_{i \neq j}^{r} H(i,j),
\]

where

\[
H(i,j) = \frac{z_i z_j}{(1+\lambda_0 \Delta_1) \cdot (1+\lambda \Delta_j)} \cdot (\Delta_1 - \Delta_j).
\]

Thus, to show that \( \frac{d}{d\lambda} G(\lambda) \) is positive, it suffices to show that

\( H(i,j) + H(j,i) \) is non-negative for each pair (\( i, j \)) with \( i > j \).

But

\[
H(i,j) + H(j,i)
\]

\[
= z_i z_j (\Delta_1 - \Delta_j) \left\{ \frac{1}{(1+\lambda_0 \Delta_1) \cdot (1+\lambda \Delta_j)} - \frac{1}{(1+\lambda_0 \Delta_1) \cdot (1+\lambda \Delta_j)} \right\}
\]

\[
= \frac{z_i z_j (\Delta_1 - \Delta_j)^2}{(1+\lambda_0 \Delta_1) \cdot (1+\lambda_0 \Delta_j) \cdot (1+\lambda \Delta_j) \cdot (1+\lambda_0 \Delta_1) \cdot (\lambda_1 - \lambda_0)}
\]

\( \geq 0. \)

Hence, the MPI test has an increasing power function.

Expression (6.1) may be re-expressed as
\[
\beta_{\text{MPF}}(\lambda) = \Pr \left\{ z : \sum_{1}^{r} \left[ \frac{1}{1+\lambda_{0} \Delta_{1}} - \frac{K_{\gamma}(\lambda_{0}, \lambda_{1})}{1+\lambda_{1} \Delta_{1}} \right] (1+\lambda \Delta_{1}) z_{1} \right. \\
\left. + \left[ 1 - K_{\gamma}(\lambda_{0}, \lambda_{1}) \right] . z_{x+1} > 0 \right\}, \quad (6.2)
\]

which is the distribution function of a linear combination of independent chi-square random variables evaluated at zero.

6.1.2. The power function of the locally most powerful invariant test of \( H_{0} : \lambda = \lambda_{0} \) against \( H_{A} : \lambda > \lambda_{0} \)

It follows from (3.17) that the power function of the size-\( \gamma \) LMPI test of \( H_{0} \) against \( H_{A} \) is

\[
\beta_{\text{LMPF}}(\lambda) = \Pr \left\{ x : \sum_{1}^{r} \frac{\Delta_{1} x_{1}^{2}}{(1+\lambda_{0} \Delta_{1})^{2}} > K_{\gamma}(\lambda_{0}); \lambda \right\}
\]

\[
= \Pr \left\{ z : \sum_{1}^{r} \frac{\Delta_{1}(1+\lambda \Delta_{1})}{(1+\lambda_{0} \Delta_{1})^{2}} \cdot z_{1} \right. \\
\left. + \sum_{1}^{r} \frac{1+\lambda \Delta_{1}}{1+\lambda_{0} \Delta_{1}} \cdot z_{x+1} > K_{\gamma}(\lambda_{0}) \right\} \quad (6.3)
\]

and hence that
$$\lim_{\lambda \to \infty} \beta_{L_{\text{MPX}}} (\lambda) = \Pr \left\{ \frac{1}{1 + \lambda_0 \Delta_1} z_1 > \frac{\sum_{i=1}^{r} \frac{\Delta_i^2 z_i}{(1 + \lambda_0 \Delta_1)^2}}{\sum_{i=1}^{r} \frac{\Delta_i z_i}{1 + \lambda_0 \Delta_1}} > k_{\gamma} (\lambda_0) \right\},$$

which is equal to one if and only if $\Delta_i / (1 + \lambda_0 \Delta_1) > k_{\gamma} (\lambda_0)$ for all $i$.

Let $M(\lambda) = \sum_{i=1}^{r} \frac{\Delta_i (1 + \lambda \Delta_i)}{(1 + \lambda_0 \Delta_1)^2} z_i / (z_r + 1 + \sum_{i=1}^{r} \frac{1 + \lambda \Delta_i}{1 + \lambda_0 \Delta_1} z_i)$. The derivative of $M(\lambda)$ is

$$\frac{d}{d\lambda} M(\lambda) =$$

$$\frac{\sum_{i=1}^{r} \frac{\Delta_i^2 z_i}{(1 + \lambda_0 \Delta_1)^2} \left( z_r + 1 + \sum_{i=1}^{r} \frac{1 + \lambda \Delta_i}{1 + \lambda_0 \Delta_1} z_i \right) - \sum_{i=1}^{r} \frac{\Delta_i z_i}{1 + \lambda_0 \Delta_1} \sum_{i=1}^{r} \frac{\Delta_i (1 + \lambda \Delta_i)}{(1 + \lambda_0 \Delta_1)^2} z_i}{(z_r + 1 + \sum_{i=1}^{r} \frac{1 + \lambda \Delta_i}{1 + \lambda_0 \Delta_1} z_i)^2}.$$ 

The numerator of $\frac{d}{d\lambda} M(\lambda)$ is clearly greater than

$$\sum_{i=1}^{r} \frac{\Delta_i^2 z_i}{(1 + \lambda_0 \Delta_1)^2} \left( z_r + 1 + \sum_{i=1}^{r} \frac{1 + \lambda \Delta_i}{1 + \lambda_0 \Delta_1} z_i \right) - \sum_{i=1}^{r} \frac{\Delta_i z_i}{1 + \lambda_0 \Delta_1} \sum_{i=1}^{r} \frac{\Delta_i (1 + \lambda \Delta_i)}{(1 + \lambda_0 \Delta_1)^2} z_i$$

$$= \sum_{i=1}^{r} \frac{\Delta_i^2 z_i}{(1 + \lambda_0 \Delta_1)^2} \sum_{i=1}^{r} \frac{z_i}{1 + \lambda_0 \Delta_1} - \sum_{i=1}^{r} \frac{\Delta_i z_i}{1 + \lambda_0 \Delta_1} \sum_{i=1}^{r} \frac{\Delta_i (1 + \lambda \Delta_i)}{(1 + \lambda_0 \Delta_1)^2} z_i$$

$$= \sum_{i \neq j} N(i, j),$$
where

\[ N(i,j) = \frac{z_i z_j \Delta_i}{(1+\lambda_o \Delta_i)(1+\lambda_o \Delta_j)} \cdot \left( \frac{\Delta_i}{1+\lambda_o \Delta_i} - \frac{\Delta_j}{1+\lambda_o \Delta_j} \right). \]

\[ = \frac{z_i z_j \Delta_i}{(1+\lambda_o \Delta_i)^2(1+\lambda_o \Delta_j)^2} (\Delta_i - \Delta_j). \]

Since

\[ N(i,j) + N(j,i) = \frac{z_i z_j (\Delta_i - \Delta_j)^2}{(1+\lambda_o \Delta_i)^2(1+\lambda_o \Delta_j)^2} \geq 0 \]

for each pair of \((i,j)\) with \(i > j\), \(\frac{d}{d\lambda} M(\lambda) > 0\). Hence, the LMPI test has an increasing power function. Consequently, \(\phi_{\text{LMPI}}(\lambda)\) of (3.17) is also the size-\(\gamma\) LMPI test of \(H_0 : \lambda \leq \lambda_o\) against \(H_A : \lambda > \lambda_o\).

The power function \(\beta_{\text{LMPI}}(\lambda)\), like the power function \(\beta_{\text{MPI}}(\lambda)\), is also expressible as the distribution function of a linear combination of independent chi-square random variables evaluated at zero. Specifically,

\[ \beta_{\text{LMPI}}(\lambda) = \text{Pr}\{ z : \sum_{1}^{r} \left[ \frac{\Delta_i}{(1+\lambda_o \Delta_i)^2} - \frac{K_{\gamma}(\lambda_o)}{1+\lambda_o \Delta_1} \right](1+\lambda \Delta_i)z_i - K_{\gamma}(\lambda_o)z_{r+1} > 0 \}. \]

(6.4)
6.1.3. The power function of Wald's test of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda > \lambda_0$

In light of (3.22), the power function of the size-$\gamma$ Wald's test of $H_0$ against $H_A$ is

$$\beta_w(\lambda) = \Pr\{ x : f \sum_{i=1}^{r+1} 1 + \lambda \Delta_i \frac{x_i^2}{\lambda_0 \Delta_i} > F_{\gamma}(x,f) ; \lambda \}$$

$$= \Pr\{ z : f \sum_{i=1}^{r+1} 1 + \lambda \Delta_i \frac{z_i}{\lambda_0 \Delta_i} > F_{\gamma}(x,f) \}. \quad (6.5)$$

It follows from (6.5) that $\lim_{\lambda \to \infty} \beta(\lambda) = 1$. Moreover, it is clear that $\beta_w(\lambda)$ is increasing in $\lambda$. Thus, Wald's test of (3.22) also defines the size-$\gamma$ Wald's test of $H_0 : \lambda \leq \lambda_0$ against $H_A : \lambda > \lambda_0$.

The power function $\beta_w(\lambda)$ is also expressible as

$$\beta_w(\lambda) = \Pr\{ z : f \sum_{i=1}^{r+1} 1 + \lambda \Delta_i \frac{z_i}{\lambda_0 \Delta_i} - F_{\gamma}(x,f)z_{r+1} > 0 \}. \quad (6.6)$$

6.2. Power Functions of Invariant Tests of $H_0 : \lambda = \lambda_0$ against Two-sided Alternatives

In this section, we consider the power functions of the two-sided versions of the MPI, LMPI, and Wald's tests as introduced in Section 4.3.

The test (4.14) for testing $H_0$ against $H_A : \lambda = \lambda_1$ or $\lambda = \lambda_2$ where $\lambda_2 > \lambda_0 > \lambda_1$ was constructed by combining two MPI tests. The power function of the combined test is
\begin{align*}
\beta_{\text{MPI},c}(\lambda) &= \Pr \left\{ x : \frac{1 + \sum_{i=1}^{r} \frac{x_i^2}{1 + \lambda_0 \Delta_i}}{1 + \sum_{i=1}^{r} \frac{x_i^2}{1 + \lambda_1 \Delta_i}} > K_{\gamma_1}(\lambda_0, \lambda_1) \text{ or } \frac{1 + \sum_{i=1}^{r} \frac{x_i^2}{1 + \lambda_0 \Delta_i}}{1 + \sum_{i=1}^{r} \frac{x_i^2}{1 + \lambda_1 \Delta_i}} > K_{\gamma_2}(\lambda_0, \lambda_2) \right\} \\
&= \sum_{j=1}^{2} \Pr \left\{ z : \frac{z_{r+1} + \sum_{i=1}^{r} \frac{1 + \lambda_0 \Delta_i}{1 + \lambda_j \Delta_i} z_i}{z_{r+1} + \sum_{i=1}^{r} \frac{1 + \lambda_0 \Delta_i}{1 + \lambda_1 \Delta_i} z_i} > K_{\gamma_j}(\lambda_0, \lambda_j) \right\} \\
&= \Pr \left\{ z : \frac{z_{r+1} + \sum_{i=1}^{r} \frac{1 + \lambda_0 \Delta_i}{1 + \lambda_1 \Delta_i} z_i}{z_{r+1} + \sum_{i=1}^{r} \frac{1 + \lambda_0 \Delta_i}{1 + \lambda_2 \Delta_i} z_i} > K_{\gamma_1}(\lambda_0, \lambda_1), \right. \\
&\left. \frac{z_{r+1} + \sum_{i=1}^{r} \frac{1 + \lambda_0 \Delta_i}{1 + \lambda_1 \Delta_i} z_i}{z_{r+1} + \sum_{i=1}^{r} \frac{1 + \lambda_0 \Delta_i}{1 + \lambda_2 \Delta_i} z_i} > K_{\gamma_2}(\lambda_0, \lambda_2) \right\} \\
&= (6.7)
\end{align*}

where \( K_{\gamma_j}(\lambda_0, \lambda_j) \) satisfies (4.17) for \( j = 1, 2 \).

For testing \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda \neq \lambda_0 \), the test (4.18) was constructed by combining two LMPI tests. The power function of the combined test is
\[ \beta_{LMP,c}(\lambda) = \Pr \left\{ z : \frac{\sum_{i=1}^{r} \frac{\Delta_i (1+\lambda \Delta_i)}{(1+\lambda \Delta_i)^2} z_i}{z_{r+1} + \sum_{i=1}^{r} \frac{1+\lambda \Delta_i}{1+\lambda_0 \Delta_i} z_i} < k_{1-\gamma_1}(\lambda_0) \right\} \]

\[ + \Pr \left\{ z : \frac{\sum_{i=1}^{r} \frac{\Delta_i (1+\lambda \Delta_i)}{(1+\lambda \Delta_i)^2} z_i}{z_{r+1} + \sum_{i=1}^{r} \frac{1+\lambda \Delta_i}{1+\lambda_0 \Delta_i} z_i} > k_{\gamma_2}(\lambda_0) \right\} \] (6.8)

where \( k_{1-\gamma_1}(\lambda_0) \) and \( k_{\gamma_2}(\lambda_0) \) satisfy (3.18) for \( \gamma = 1 - \gamma_1 \) and \( \gamma = \gamma_2 \), respectively.

Also for testing \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda \neq \lambda_0 \), the two-sided Wald's test (4.20) has the power function

\[ \beta_{w,c}(\lambda) = \Pr \left\{ z : \frac{\sum_{i=1}^{r} \frac{1+\lambda \Delta_i}{1+\lambda_0 \Delta_i} z_i}{z_{r+1}} < f_{1-\gamma_1}(r,f) \right\} \]

\[ + \Pr \left\{ z : \frac{\sum_{i=1}^{r} \frac{1+\lambda \Delta_i}{1+\lambda_0 \Delta_i} z_i}{z_{r+1}} > f_{\gamma_2}(r,f) \right\} \] (6.9)
6.3. Imhof's Technique for Evaluating the Distribution Function of a Linear Combination of Independent Chi-square Random Variables

Let \( \chi^2(m, \eta) \) represent the distribution of a non-central chi-square random variable with degrees of freedom \( m \) and non-centrality parameter \( \eta \). Let \( s, m_1, \ldots, m_s \) represent positive integers and \( \eta_1, \ldots, \eta_s \) non-negative constants. Let \( T = \sum_{j=1}^{s} a_j v_j \), where \( a_j \) are arbitrary non-zero constants and \( v_j \) are independent random variables such that \( v_j \sim \chi^2(m_j, \eta_j) \), \( j = 1, \ldots, s \). The Inversion Theorem (e.g., Chung, 1974, pp. 152-154) may be used to evaluate the distribution functions of \( T \). Imhof (1961) pointed out that such an integral formula was implicit in the work of Gurland (1948). An explicit representation was first given by Gil-Pelaez (1951). In the following, we present Imhof's work for the special case where the \( \eta_j \)'s are all zero, i.e., where \( v_j \sim \chi^2(m_j), j = 1, \ldots, s \).

Since the \( v_j \)'s are independent, the characteristic function of \( T \) is \( f(u) = \prod_{j=1}^{s} (1-2\lambda_j u)^{-m_j/2} \). Consequently, by the inversion formula, the distribution function of \( T \) evaluated at \( t \) is

\[
\Pr\{ T \leq t \} = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty u^{-2} \mathcal{J}\{ e^{-\lambda u} f(u) \} \, du
\]

where \( \mathcal{J}(.) \) denotes the imaginary part of a complex number (i.e., \( \mathcal{J}(c + id) = d \) where \( c \) and \( d \) are real numbers).

Let \( \text{arg}(.) \) represent the principal argument of a non-zero complex number (e.g., Kaplan, 1984, p. 572). By using the relationships
arg[(1 - dbu)^\theta] = g \tan^{-1}(bu),

| (1 - dbu)^\theta | = (1 + b^2u^2)^{-g/2},

\arg \left( \exp\left(\frac{iau}{1-ibu}\right) \right) = \frac{au}{1 + b^2u^2},

and

\left| \exp\left(\frac{iau}{1-ibu}\right) \right| = \exp\left(\frac{-abu^2}{1 + b^2u^2}\right),

Imhof found that

\Pr\{ T \leq t \} = \frac{1}{2} \cdot \frac{1}{ \pi} \int_{0}^{\infty} \frac{\sin \Theta(u)}{u \rho(u)} \, du \quad (6.10)

where

\Theta(u) = \frac{1}{2} \sum_{1}^{s} m_j \tan^{-1}(a_ju) - \frac{1}{2} tu

and

\rho(u) = \prod_{1}^{s} (1 + a_j^2u^2)^{m_j/4}.

It can be verified that \( \lim_{u \to 0} \frac{\sin \Theta(u)}{u \rho(u)} = \frac{1}{2} \frac{s}{\pi} \sum_{1}^{s} a_j m_j - t \) and

that

\lim_{u \to \infty} \Theta(u) = \begin{cases} -\infty, & \text{if } t > 0, \\ \infty, & \text{if } t < 0, \\ \frac{\pi}{4} \sum_{1}^{s} m_j a_j |a_j|^{-2}, & \text{if } t = 0. \end{cases}

Imhof pointed out that the function \( u \rho(u) \) increases
monotonically as \( u \to \infty \) and suggested that, for computational purposes, (6.10) can be approximated by

\[
I_U = \frac{1}{2} - \frac{1}{\pi} \int_0^U \frac{\sin \theta(u)}{u \rho(u)} \, du, \tag{6.11}
\]

where \( U \) is a positive constant to be determined so that the difference between (6.10) and (6.11)

\[
t_U = \frac{1}{\pi} \int_U^\infty \frac{\sin \theta(u)}{u \rho(u)} \, du
\]

is small. An upper bound for \( |t_U| \) is

\[
T_U = \frac{1}{\pi k U} \prod_{j=1}^s |a_j|^{-m_j/2},
\]

where \( k = \sum_{j=1}^s m_j/2 \). By making use of this upper bound, \( U \) can be chosen to achieve any desired level of accuracy. Numerical integration can be used to approximate integral (6.11). Simpson's rule and the trapezoidal rule were considered by Imhof. It was suggested, on the basis of results for several special cases, that the trapezoidal rule is superior to Simpson's rule.

Imhof's method was reported by Solomon and Stephens (1977) to be very satisfactory, for purposes of computing the percentage points of the distribution of \( T \), when all the \( a_j \)'s are positive. Davies (1973, 1980) developed an algorithm for implementing Imhof's method. The subroutine DCADRE of IMSL is useful for computing (6.11) and thus may be used to implement Imhof's procedure as well.
6.4. Finding Critical Points of MPI and LMPI Tests

In this section, we indicate how the approximation (6.11) to (6.10) may be incorporated into an algorithm for computing the critical points of invariant tests. Note that the critical points of the one-sided or two-sided Wald's tests are simply percentage points of $F(r, f)$. Algorithms and software for computing these percentage points are readily available.

In general, the problem of finding a critical point of an MPI or an LMPI test is one of solving a non-linear equation

$$\beta(\lambda_0) = \gamma \tag{6.12}$$

where $\gamma$ is the size of test, $0 < \gamma < 1$, and $\beta(\lambda_0)$ is the power function $\beta(\lambda)$ evaluated at $\lambda = \lambda_0$.

Kennedy and Gentle (1980, section 5.2) describe three commonly used numerical methods for computing the root $x_p$ of the equation $F(x) - p = 0$ for some function $F$ of $x$. The three methods are all iterative. On the $(i + 1)$st iteration, a new approximation to $x_p$ -- called an iterate -- is computed from previous iterate(s) and the functional values of $F$ at these iterate(s). The three methods are Newton's method, the secant method, and the bisection method. Let $f(x) = F(x) - p$. The $(1 + 1)$st iterate given by Newton's method is

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}, \tag{6.13}$$

where $f'(x_i)$ is the derivative of $f(x)$ evaluated at $x = x_i$. The $(1 + 1)$st iterate of the secant method is
Algorithm (6.14) is the same as (6.13) except that the derivative \( f'(X) \) is replaced by an approximation.

The bisection method is begun with an interval \([x^0, x^1]\) which is known to cover a root. Thus, \( x_0 \) and \( x_1 \) are such that \( f(x_0) f(x_1) < 0 \). On the first step of the method, this interval is cut at its midpoint \( x_2 = \frac{1}{2} (x_0 + x_1) \), thereby splitting the interval in half. Unless the root happens to be \( x_2 \), it lies in one and only one of the two subintervals. Then, a new interval, shorter than the previous and containing the root, is formed. This new interval forms the starting point for another iteration.

In all three methods, the iterative process continues until some convergence criteria are satisfied. In particular, we could stop the bisection process at the \( i \)th step if \( |x_i - x_{i-1}| < \delta_1 \) and \( |f(x_{i+1})| < \delta_2 \), where \( x_i \) and \( x_{i-1} \) are the end points and \( x_{i+1} \) the midpoint of the \( i \)th interval and \( \delta_1 \) and \( \delta_2 \) are pre-chosen constants. The root is then approximated by \( x_{i+1} \). Although it is usually slower than the other two methods, the bisection method is generally more stable than its competitors.

### 6.4.1. Bracketing the critical point of an MPI test

For a size-\( \gamma \) MPI test of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda = \lambda_1 (\lambda_1 > \lambda_0) \), we now construct an interval \([K_0, K_1]\) which brackets the critical point \( k_{\gamma}(\lambda_0, \lambda_1) \). It follows from (6.1) that
\[ \beta_{\text{MPI}}(\lambda_0) = \Pr \left( z : \frac{z_{r+1} + \sum_{i=1}^{r} z_i}{z_{r+1} + \frac{\lambda_0}{1 + \lambda_0 \Delta_i} \sum_{i=1}^{r} z_i} > K_0(\lambda_0, \lambda_1) \right). \]

We wish to choose \( K_0 \) and \( K_1 \) so that \( \beta_{\text{MPI}}(\lambda_0) \geq \gamma \) for \( K_0(\lambda_0, \lambda_1) = K_0 \) and \( \beta_{\text{MPI}}(\lambda_0) \leq \gamma \) for \( K_1(\lambda_0, \lambda_1) = K_1 \). For the purposes of choosing \( K_0 \) and \( K_1 \), let

\[ q^* = \frac{1 + \lambda_0 \Delta_*}{1 + \lambda_1 \Delta_*} \quad \text{and} \quad q_* = \frac{1 + \lambda_0 \Delta_*}{1 + \lambda_1 \Delta_*}, \]

where \( \Delta_* = \max\{\Delta_i\} \) and \( \Delta_* = \min\{\Delta_i\} \). Observe that

\[ q^* \leq \frac{1 + \lambda_0 \Delta_1}{1 + \lambda_1 \Delta_1} \leq q_*, \quad i = 1, \ldots, r. \]

Therefore, \( K_0 \) and \( K_1 \) can be chosen so that

\[ \Pr \left( z : \frac{z_{r+1} + \sum_{i=1}^{r} z_i}{z_{r+1} + q_* \sum_{i=1}^{r} z_i} > K_0 \right) = \gamma \]

and

\[ \Pr \left( z : \frac{z_{r+1} + \sum_{i=1}^{r} z_i}{z_{r+1} + q^* \sum_{i=1}^{r} z_i} > K_1 \right) = \gamma', \]

respectively.

Recall that the \( z_i \)'s are independent and that \( z_1 \sim \chi^2(1), i = 1, \ldots, r, \) and \( z_{r+1} \sim \chi^2(r) \). It is straightforward to verify that
and

\[ K_0 = \frac{f + rF_y(r, f)}{f + rq*F_y(r, f)} \]

\[ K_1 = \frac{f + rF_y(r, f)}{f + rq*F_y(r, f)} \]

### 6.4.2. Bracketing the critical point of an LMPI test

Next, we consider computing the critical point of a size-\(\gamma\) LMPI test of \(H_0 : \lambda = \lambda_0\) against \(H_A : \lambda > \lambda_0\). It follows from (6.3) that

\[
\beta_{\text{LMPI}}(\lambda_0) = \text{Pr}\left\{ z : \frac{\sum_{i=1}^{r} \frac{\Delta_i z_i}{1 + \lambda_0 \Delta_i}}{\frac{r}{z_{r+1} + \sum_{i=1}^{r} z_i}} > K_y(\lambda_0) \right\}.
\]

Let

\[ q^* = \frac{\Delta^*}{1 + \lambda_0 \Delta^*} \quad \text{and} \quad q_* = \frac{\Delta_*}{1 + \lambda_0 \Delta_*}. \]

Observe that

\[ q_* \leq \frac{\Delta_1}{1 + \lambda_0 \Delta_1} \leq q^*, \quad i = 1, \ldots, r. \]

Therefore, \(\beta_{\text{LMPI}}(\lambda_0) \geq \gamma\) for \(K_y(\lambda_0) = K_0\) and \(\beta_{\text{LMPI}}(\lambda_0) \leq \gamma\) for \(K_y(\lambda_0) = K_1\), where \(K_0\) and \(K_1\) are defined by

\[
\text{Pr}\left\{ z : \frac{\sum_{i=1}^{r} \frac{z_i}{z_{r+1} + \sum_{i=1}^{r} z_i}}{\sum_{i=1}^{r} \frac{z_i}{z_{r+1} + \sum_{i=1}^{r} z_i}} > K_0 \right\} = \gamma.
\]
and

\[
\text{Pr}\left\{ z : \frac{q^* \sum_{l=1}^{r} z_1}{z_{r+1} + \sum_{l=1}^{r} z_1} > K_1 \right\} = \gamma,
\]

respectively.

Consequently, \([K_0, K_1]\) is an interval which brackets the critical point \(K_\gamma(\lambda_0)\). It is easy to verify that

\[
K_0 = q^* \cdot \frac{rF_\gamma(x, f)}{\int rF_\gamma(x, f) dx}
\]

and

\[
K_1 = q^* \cdot \frac{rF_\gamma(x, f)}{\int rF_\gamma(x, f) dx}.
\]

6.5. Approximating Critical Points of LRI and MIXI Tests

It is difficult to compute the exact critical points of the two-sided tests \(\phi_{\text{LRI}}(x)\) of (4.5) and \(\phi_{\text{MIXI}}(x)\) of (4.13). Practically, we may approximate the critical points by conducting simulation studies. In the following, we briefly outline a procedure which is useful in approximating the critical points of \(\phi_{\text{LRI}}(x)\). We omit the description of the approximation procedure for the critical points of \(\phi_{\text{MIXI}}(x)\) because it has the same manner as that of \(\phi_{\text{LRI}}(x)\).

Let \(S(x)\) represent the test statistic of the \(\phi_{\text{LRI}}(x)\), that
Let $\{x^{(j)}; j = 1, \ldots, N\}$ represent a r-variate random sample of size N, where $x^{(j)} = t^{(j)}/\sqrt{SS_e^{(j)}}$ is such that $t^{(j)}$ is distributed as $\text{MVN}_r(0, (I + \lambda D_x))$, $SS_e^{(j)}$ is distributed as $\chi^2(f)$, and $t^{(j)}$ is statistically independent of $SS_e^{(j)}$, $j = 1, \ldots, N$. The random sample $x^{(j)}$'s can be generated easily as algorithms and softwares (e.g., SAS and IMSL) for generating multivariate normal and chi-square random variates are readily available.

**Definition 6.1**

Given the random sample $\{x^{(j)}; j = 1, \ldots, N\}$, the empirical distribution function of $S(x)$ is

$$F(\phi) = \frac{1}{N} \sum_{j=1}^{N} I\{S(x^{(j)}) \leq \phi\},$$

where $0 \leq \phi \leq \infty$ and for $1 \leq j \leq N$,

$$I\{S(x^{(j)}) \leq \phi\} = \begin{cases} 1, & \text{if } S(x^{(j)}) \leq \phi, \\ 0, & \text{otherwise}. \end{cases}$$

One needs to solve the equation (4.4) for the maximum likelihood estimate $\lambda$ in order to construct $\hat{F}(\phi)$. Callanan (1985) considered fourteen algorithms for computing the restricted maximum likelihood estimates of the variance components, hence of
For any \( r \) such that \( 0 \leq r \leq 1 \), let \( \phi_r \) represent any member of \( \{ S(x^{(j)}), j = 1, \ldots, N \} \) which satisfies

\[
(a) \quad \frac{1}{N} \sum_{j=1}^{N} I\{ S(x^{(j)}) < \phi_r \} \leq r
\]

and

\[
(b) \quad F(\phi_r) \geq r.
\]

The number of \( \phi_r \)'s satisfying (a) and (b) is at least one and at most two. In the latter case, we denote the smaller and the larger of the two \( \phi_r \)'s by \( \phi_{r,1} \) and \( \phi_{r,2} \) respectively. Define

\[
\hat{K}_r(\lambda_0) = \begin{cases} 
\phi_r & \text{when } \phi_r \text{ is the unique number satisfying (a) and (b)}, \\
(\phi_{r,1} + \phi_{r,2})/2 & \text{when both } \phi_{r,1} \text{ and } \phi_{r,2} \text{ satisfy (a) and (b)}. 
\end{cases}
\]

**Definition 6.2**

Given the random sample \( \{ x^{(j)}, j = 1, \ldots, N \} \), \( \hat{K}(\lambda_0) \) is the lower-\( r \) point of the empirical distribution of \( S(x) \).

Then, the critical point \( K_r(\lambda_0) \) of the test \( \phi_{\text{LMI}}(x) \) can be approximated by \( \hat{K}_r(\lambda_0) \).

Suggested by the standard asymptotic theory (e.g., Cox and Hinkley, 1974), Harville and Callanan (1987) proposed that the distribution of \( -2 \log S(x) \) be approximated by \( \chi^2(1) \) under the null hypothesis. Based on the approximation, the following approximate 100(1-\( r \))% confidence set for \( \lambda \) was given:
It is possible to construct pathological examples in which the confidence set (6.15) is not an interval.

6.6. Technique of Computing and Nature of the Confidence Sets

To find the \( 100(1-\gamma)\% \) confidence lower limit corresponding to the family of size-\( \gamma \) Wald's tests of \( H_0: \lambda = \lambda_0 \) against the alternative \( H_A: \lambda > \lambda_0 \), one may first plot the curve \( P_\gamma(x) = Q(\lambda) - F_{\gamma}(x; f) \) against \( \lambda \) and locate an interval covering the root to the non-linear equation \( P_\gamma(\lambda) = 0 \). The non-linear equation is solvable by applying the bisection method (see Section 6.4). The root to \( P_\gamma(\lambda) = 0 \) is the \( 100(1-\gamma)\% \) confidence lower limit for \( \lambda \). A confidence upper limit for \( \lambda \) corresponding to the family of Wald's tests of \( H_0 \) against the alternative \( H_A: \lambda < \lambda_0 \) can be obtained by using the same process.

The technique is also applicable for constructing confidence sets of the LMPI and MPI procedures. For example, to construct the \( 100(1-\gamma)\% \) confidence set \( C_{\text{LMPI}}(x) \) of (3.19), we plot the curve

\[
P_{\text{LMPI}}(\lambda) = \sum \frac{r \Delta_i x_1^2}{(1 + \lambda \Delta_i)^2} / \left( 1 + \sum \frac{r \Delta_i x_1^2}{1 + \lambda \Delta_i} \right) - K_\gamma(\lambda)
\]

against \( \lambda \) and locate interval(s) covering the root(s) to the non-linear equation.
\( P_{\text{LMP}}(\lambda) = 0 \). As soon as the root(s) to the non-linear equation are obtained by applying the bisection method, the confidence set \( C_{\text{LMP}}(\lambda) \) may be identified as the collection of values of \( \lambda \) such that \( P_{\text{LMP}}(\lambda) \leq 0 \).

Similarly, to construct the 100(1-\( \gamma \))% confidence set \( C_{\text{MPI}}(\lambda) \) of (3.4), we plot the curve \( P_{\text{MPI}}(\lambda) = \frac{r x_1^2}{1 - \lambda \Delta_1} / \frac{r x_1^2}{1 + \lambda \Delta_1} \) to locate the interval(s) covering the root(s) to the non-linear equation \( P_{\text{MPI}}(\lambda) = 0 \). Then, the bisection method is applied to obtain the root(s) to form the confidence set \( C_{\text{MPI}}(\lambda) = \{ \lambda \in \Lambda : P_{\text{MPI}}(\lambda) \leq 0 \} \).

It is computationally more intensive to construct the confidence sets corresponding to the LMP and MPI procedures than to the Wald's procedure. This is due to the fact that the critical points of the LMP and MPI tests, unlike those of Wald's test, depend on the value of null hypothesis. In addition, since it is difficult to investigate analytically the properties of the curves \( P_{\text{LMP}}(\lambda) \) and \( P_{\text{MPI}}(\lambda) \), it is inconclusive yet whether either \( C_{\text{LMP}}(\lambda) \) or \( C_{\text{MPI}}(\lambda) \) is necessarily an interval. We have found, in a numerical example, that \( C_{\text{LMP}}(\lambda) \) of (3.19) gives a confidence lower limit, that \( C_{\text{LMP}}(\lambda) \) of (3.21) gives a confidence upper limit, and that \( C_{\text{MPI}}(\lambda) \) of (3.4) gives bounded confidence intervals for different choices of \( \lambda_1 \). These results are included in the numerical study in the Chapter 7.

Since we plot the curves \( P_{\text{LMP}}(\lambda) \) and \( P_{\text{MPI}}(\lambda) \) only over a
finite range of values of $\lambda$, we would never be certain that all
the roots to the non-linear equations $P_{\text{LMPI}}(\lambda) = 0$ and $P_{\text{MPI}}(\lambda) = 0$
are obtained. Thus, it seems to be possible that the correct
confidence sets are not obtained for the LMPI and MPI procedures.
7. NUMERICAL STUDIES

In this chapter, numerical comparisons among the various confidence-set procedures are made for each of five examples: the example considered by Harville and Fenech (1985) and four examples in which the model is unbalanced one-way random model -- each with a different pattern of \( n_i \)'s.

The basis for the comparisons is the probability of covering a "false" value \( \lambda_0 \) of \( \lambda \). In Section 7.1, we describe the five examples in detail, and then in Section 7.2, present the results of the numerical comparisons. In Section 7.3, we summarize our results.

7.1. Data Structures Used in Comparisons of Procedures

7.1.1. One-way classification

Burdick, Magsood, and Graybill (1986) considered the construction of confidence intervals for \( \lambda \) in the unbalanced one-way random model (1.3). They compared six approximations of Wald's procedure to the "exact" Wald's procedure. In implementing the exact Wald's procedure, they used the bisection method to solve the non-linear equations (4.22). The basis of their comparisons was a simulation study of the confidence coefficients and of the average lengths produced by the different methods. Their study covered ten different patterns of \( n_i \)'s.

Four of their ten patterns were included in the present
study. These four patterns and the non-zero eigenvalues of the matrix C for each of the four are given in the Table 2. Note that there are \( b - 1 \) non-zero eigenvalues of C for a one-way model with \( b \) classes.

Table 2. Four patterns of \( n_i \)'s for the unbalanced one-way classification and the non-zero eigenvalues of the matrix C

<table>
<thead>
<tr>
<th>Pattern Number</th>
<th>Number of Classes</th>
<th>Class Sizes</th>
<th>Non-zero Eigenvalues of C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1(4)</td>
<td>6</td>
<td>1,1,1,1,1,100</td>
<td>1.0, 1.0, 1.0, 1.0, 5.7143</td>
</tr>
<tr>
<td>2(6)</td>
<td>6</td>
<td>1,100,100,100,100,100</td>
<td>1.1976, 100.0, 100.0, 100.0, 100.0</td>
</tr>
<tr>
<td>3(8)</td>
<td>10</td>
<td>1,1,4,5,6,6,8,8,10,10</td>
<td>1.0, 1.2025, 4.2004, 5.2293, 6.0, 6.7991, 8.0, 9.0603, 10.0</td>
</tr>
<tr>
<td>4(10)</td>
<td>10</td>
<td>3,3,4,5,6,6,8,8,10,10</td>
<td>3.0, 3.3065, 4.2444, 5.2485, 6.0, 6.8356, 8.0, 9.0793, 10.0</td>
</tr>
</tbody>
</table>

*The pattern numbers in parentheses are those assigned by Burdick, Maqsood, and Graybill (1986).*

7.1.2. Lamb-weight data

An example consisting of data of the weights at birth of 62 single-birth male lambs was introduced by Harville and Fenech (1985, Table 1). These lambs are the offspring of a total of 23 rams, belonging to five distinct population lines, and 62 dams, which are categorized by age into class 1 (1-2 years), 2 (2-3 years), or 3 (over 3 years).

Let \( y_{ijkl} \) represent the birth weight of the \( l \)th of the lambs
that are the offspring of the kth sire in the jth line and of a
dam belonging to the ith age category. The data may be modeled as

\[ Y_{ijkl} = \mu + \delta_i + \pi_j + \beta_{jk} + e_{ijkl} \]

where the age effects (\(\delta_1, \delta_2, \delta_3\)) and the line effects (\(\pi_1, ..., \pi_5\)) are fixed effects, the sire within line effects (\(\beta_{11}, ..., \beta_{58}\)) are random effects, and (\(e_{1111}, ..., e_{3582}\)) are random errors. It is assumed that the \(\beta's\) are independently distributed as \(N(0, \sigma^2_\beta)\), that the e's are independently distributed as \(N(0, \sigma^2_e)\), and that the \(\beta's\) and e's are statistically independent of each other.

An ANOVA table for testing \(H_0: \sigma^2_\beta = 0\) against \(H_A: \sigma^2_\beta > 0\), or equivalently, \(H_0: \lambda = 0\) against \(H_A: \lambda > 0\), is

Table 3. Analysis of variance for testing \(H_0: \lambda = 0\) for lamb weight data

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>E(MS)</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha)</td>
<td>7</td>
<td>SS(\alpha) = 7454.959</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\beta</td>
<td>\alpha)</td>
<td>18</td>
<td>SS(\beta) = 80.296</td>
<td>4.460</td>
<td>(\sigma^2_e + K\sigma^2_\beta)</td>
</tr>
<tr>
<td>Error</td>
<td>37</td>
<td>SS(e) = 102.235</td>
<td>2.763</td>
<td>(\sigma^2_e)</td>
<td></td>
</tr>
</tbody>
</table>

Here, \(K = 2.2118\) and the values of \(\Delta_i\) and \(t_i\), \(i = 1, ..., 18\) are

<table>
<thead>
<tr>
<th>(\Delta)</th>
<th>(t)</th>
<th>(\Delta)</th>
<th>(t)</th>
<th>(\Delta)</th>
<th>(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8400</td>
<td>-2.9062</td>
<td>1.4078</td>
<td>1.2882</td>
<td>2.7482</td>
<td>-2.5893</td>
</tr>
<tr>
<td>Value 1</td>
<td>Value 2</td>
<td>Value 3</td>
<td>Value 4</td>
<td>Value 5</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>0.9027</td>
<td>-0.4508</td>
<td>1.7077</td>
<td>-2.5007</td>
<td>3.1505</td>
<td>-2.4773</td>
</tr>
<tr>
<td>1.0000</td>
<td>-4.8083</td>
<td>1.9329</td>
<td>-2.1365</td>
<td>3.3236</td>
<td>3.0048</td>
</tr>
<tr>
<td>1.0750</td>
<td>0.7319</td>
<td>2.0000</td>
<td>-1.5313</td>
<td>3.5644</td>
<td>-1.5521</td>
</tr>
<tr>
<td>1.1644</td>
<td>-0.7361</td>
<td>2.0000</td>
<td>0.9594</td>
<td>4.2340</td>
<td>-1.8835</td>
</tr>
<tr>
<td>1.3456</td>
<td>1.2924</td>
<td>2.3293</td>
<td>-1.2735</td>
<td>5.0875</td>
<td>0.7676</td>
</tr>
</tbody>
</table>

Also, $\Lambda_1 = [0, \infty)$, $\Lambda_2 = (-1/\omega, \infty) = (-0.111, \infty)$, and $\Lambda_3 = (-1/\Delta^*, \infty) = (-0.197, \infty)$. An 80% confidence interval $[\lambda_*, \lambda^*]$ for $\lambda$ based on Wald's procedure is obtained by equating $Q(\lambda)$ to $F_{0.10}(18, 37)$ and $F_{0.90}(18, 37)$ and solving for $\lambda_*$ and $\lambda^*$, respectively. The interval, as given by Harville and Fenech (1985), is $[-0.008, 1.125]$. Harville and Callanan (1987) also constructed the approximate 80% confidence set (6.15), which consists of the interval $[-0.079, 0.729]$.

### 7.2. Numerical Results

For the five examples introduced in Section 7.1, we compare the confidence sets corresponding to various LMPI tests, MPI tests, and Wald's tests. Specifically, the following three confidence sets are investigated:

1. the confidence sets corresponding to the LMPI, MPI, and Wald's tests of $H_0 : \lambda = \lambda_0$ against the alternative $H_A : \lambda > \lambda_0$ for the LMPI and Wald's tests and $H_A : \lambda = 1/3$ for the MPI tests;

2. the confidence sets corresponding to the LMPI, MPI, and Wald's tests of $H_0 : \lambda = \lambda_0$ against the alternative $H_A : \lambda < \lambda_0$.
for the LMPI and Wald's tests and \( H_A : \lambda = 0.0 \) for the MPI tests; and

(3) the confidence sets corresponding to the combined LMPI and Wald's tests of \( H_0 : \lambda = \lambda_0 \) against \( H_A : \lambda \neq \lambda_0 \).

The size of each test is taken to be \( \gamma = 0.05 \), corresponding to a confidence coefficient of 0.95. In case of a two-sided alternative, the size is equally divided, i.e., we take \( \gamma_1 = \gamma_2 = 0.025 \). In comparing two-sided invariant tests of \( H_0 : \lambda = \lambda_0 \), the LRI test and the test combining two one-sided MPI tests were not included due to the extensive computations that would have been required.

By using the equality (1.11), the probability of false coverage of any of the confidence-set procedures can be determined from the power of its corresponding test function. The critical points of the various tests were computed in accordance with the discussion of Section 6.4 while the powers of the tests were computed in accordance with the discussion of Sections 6.1, 6.2, and 6.3.

For each of the five examples introduced in Section 7.1, the probability of falsely covering \( \lambda = 0.0 \), was plotted (as a function of the true value of \( \lambda \)) for each of the three confidence sets corresponding, respectively, to the LMPI, MPI, and Wald's tests of \( H_0 : \lambda = \lambda_0 \) against the alternative \( H_A : \lambda > \lambda_0 \) for the LMPI and Wald's test and \( H_A : \lambda = 1/3 \) for the MPI test. The results are presented in Figures 1 to 5. As expected, the
The confidence set corresponding to the LMPI procedure has the smallest probability of covering \( \lambda = 0.0 \) when the true value of \( \lambda \) is in the neighborhood of \( \lambda = 0.0 \), while the set corresponding to Wald's procedure has the smallest probability of false coverage when the true value is far from \( \lambda = 0.0 \). Not surprisingly, the confidence set corresponding to the MPI procedure has the smallest probability of false coverage when the true values of \( \lambda \) is "close" to the point \( \lambda = 1/3 \). There was no true value of \( \lambda \) (over the range considered) for which the probability of false coverage of the MPI procedure exceeded that of both of the other two procedures.

The probability of falsely covering \( \lambda = 1/3 \) is plotted in Figures 6 to 10 for each of the three confidence sets corresponding to the LMPI, MPI, and Wald's tests of \( H_0: \lambda = \lambda_0 \) against the alternative \( H_A: \lambda < \lambda_0 \) for the LMPI and Wald's tests and \( H_A: \lambda = 0.0 \) for the MPI test. Wald's procedure is clearly inferior to the other two procedures over the range of \( \lambda \)-values considered. Although the LMPI procedure has the smallest probability of falsely covering \( \lambda = 1/3 \) when true value of \( \lambda \) is close to 1/3, the difference among the three procedures are small for all values of \( \lambda \), around and to the right of the point \( \lambda = 1/3 \). The MPI procedure has the smallest probability of false coverage for true values of \( \lambda \) between \(-1/\Delta^\bullet\) and a point in the interval \((0, 1/3)\).

Figures 11 to 15 give plots of the probability of falsely covering \( \lambda = 0.15 \) of the confidence sets corresponding to the LMPI and Wald's tests of \( H_0: \lambda = \lambda_0 \) against \( H_A: \lambda = \lambda_0 \). For each of
the five examples, Wald's procedure is the best of the two procedures when the true value is not too far to the left of 0.0 and when it is sufficiently far to the right of 0.0.

For the lamb-weight data, confidence sets corresponding to various test procedures were constructed in accordance with the discussion in Section 6.6. The 90% confidence set corresponding to the LMPI tests of $H_0 : \lambda = \lambda_0$ against the alternative $H_A : \lambda > \lambda_0$ consists of the interval $[-0.114, \infty)$ while that against $H_A : \lambda < \lambda_0$ is $(-0.197, 0.900]$. Consequently, an 80% confidence interval corresponding to the LMPI tests of $H_0$ against $H_A : \lambda \neq \lambda_0$ is $[-0.114, 0.900]$. On the other hand, the 90% confidence set corresponding to the MPI tests of $H_0$ against $H_A : \lambda = 1/3$ consists of the interval $[-0.055, 0.809]$ while that against $H_A : \lambda = 0.0$ is $[-0.082, 0.707]$. Consequently, a confidence set corresponding to the MPI tests of $H_0$ against the alternative $H_A : \lambda = 0.0$ or 1/3 is $[-0.055, 0.707]$ with confidence level at least 80%.

7.3. Concluding Remarks

The main purpose of this dissertation has been the construction of confidence sets for $\lambda$, the ratio of the variance components in a mixed linear model with two variance components. To obtain the confidence sets, we have constructed families of "optimal" invariant tests of $H_0 : \lambda = \lambda_0$, where $\lambda_0 \in \Lambda$. Corresponding to each family of tests is a confidence set.

Figures 1 to 15 provide some information about the different
probability of false coverage for each of the confidence-set procedures. However, in choosing among the procedures, there is another consideration. The critical points of the LMPI and MPI tests depend on the value of the null hypothesis, while those of Wald's test do not. Thus, the LMPI and MPI procedures are much more computationally intensive than Wald's procedure. Moreover, when using the technique given in Section 6.6 to construct the confidence sets for \( \lambda \) corresponding to the LMPI and MPI procedures, it is possible that the correct confidence sets are not obtained when the parameter space \( \Lambda \) is unbounded.

Unless it is believed that the true value of \( \lambda \) is large and one is interested in obtaining a confidence lower limit for \( \lambda \), one should hesitate to adopt Wald's procedure. Especially when the parameter space for \( \lambda \) is restricted (for example, to the interval \([0, 1/3]\), as in the lamb-weight example), a confidence set corresponding to a most powerful invariant test should be preferred.
Figure 1. Probability of falsely covering $\lambda = 0$ by 95% confidence sets corresponding to tests of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda > \lambda_0$ for the LMP1 and Wald tests and $H_A : \lambda = 1/3$ for the MPI test.
Figure 2. Probability of falsely covering $\lambda = 0$ by 95% confidence sets corresponding to tests of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda > \lambda_0$ for the LMPI and Wald tests and $H_A : \lambda = 1/3$ for the MPI test.
Figure 3. Probability of falsely covering $\lambda = 0$ by 95% confidence sets corresponding to tests of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda > \lambda_0$ for the LMP1 and Wald tests and $H_A : \lambda = 1/3$ for the MPI test.
Figure 4. Probability of falsely covering $\lambda = 0$ by 95% confidence sets corresponding to tests of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda > \lambda_0$ for the LMPI and Wald tests and $H_A : \lambda = 1/3$ for the MPI test.
Figure 5. Probability of falsely covering $\lambda = 0$ by 95% confidence sets corresponding to tests of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda > \lambda_0$ for the LMPI and Wald tests and $H_A : \lambda = 1/3$ for the MPI test.
Figure 6. Probability of falsely covering $\lambda = 1/3$ by 95% confidence sets corresponding to tests of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda < \lambda_0$ for the LMPI and Wald tests and $H_A : \lambda = 0$ for the MPI test.
Figure 7. Probability of falsely covering $\lambda = 1/3$ by 95% confidence sets corresponding to tests of $H_0: \lambda = \lambda_0$ against $H_A: \lambda < \lambda_0$ for the LMP1 and Wald tests and $H_A: \lambda = 0$ for the MPI test.
Figure 8. Probability of falsely covering $\lambda = 1/3$ by 95% confidence sets corresponding to tests of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda < \lambda_0$ for the LMPI and Wald tests and $H_A : \lambda = 0$ for the MPI test.
Figure 9. Probability of falsely covering $\lambda = 1/3$ by 95% confidence sets corresponding to tests of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda < \lambda_0$ for the LMPI and Wald tests and $H_A : \lambda = 0$ for the MPI test.
Figure 10. Probability of falsely covering $\lambda = \frac{1}{3}$ by 95% confidence sets corresponding to tests of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda < \lambda_0$ for the LMPI and Wald tests and $H_A : \lambda = 0$ for the MPI test.
Figure 11. Probability of falsely covering $\lambda = .15$ by 95% confidence sets corresponding to tests of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda \neq \lambda_0$. 

---

**PATTERN 1**

- Probability of false covering $X = .15$ by 95% confidence sets corresponding to tests of $H_0 : X = X_0$ against $H_A : X \neq X_0$.
Figure 12. Probability of falsely covering $\lambda = .15$ by 95% confidence sets corresponding to tests of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda \neq \lambda_0$.
Figure 13. Probability of falsely covering $\lambda = .15$ by 95% confidence sets corresponding to tests of $H_0: \lambda = \lambda_0$ against $H_A: \lambda \neq \lambda_0$. 
Figure 14. Probability of falsely covering $\lambda = .15$ by 95% confidence sets corresponding to tests of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda \neq \lambda_0$. 
Figure 15. Probability of falsely covering $\lambda = .15$ by 95% confidence sets corresponding to tests of $H_0 : \lambda = \lambda_0$ against $H_A : \lambda \neq \lambda_0$
For an arbitrary matrix $A$, let $C(A)$ represent the column space of $A$, and let $P_A$ represent the unique projection matrix onto $C(A)$, i.e., let $P_A = A(A'A)^{-1}A'$ where $(A'A)^{-1}$ is a generalized inverse of $A'A$. It is well-known that $P_A$ is symmetric and idempotent, that $P_A A = A$ and $A' P_A = A'$, that $C(A) = C(P_A)$, and that $\text{rank}(A) = \text{rank}(P_A) = \text{tr}(P_A)$. For matrices $A$ and $B$ having the same number of rows, let $C(A,B)$ represent the column space of the partitioned matrix $(A,B)$, and let $P_{A,B}$ represent the projection matrix onto $C(A,B)$.

**Lemma 1**

(i) $C(A,B) = C[A,(I - P_A)B]$;

(ii) $\text{rank}((I - P_A)B) = \text{rank}(A,B) - \text{rank}(A)$;

(iii) $C(P_{A,B} - P_A) = C((I - P_A)B)$.

**Proof**

(i) Let $F = \begin{bmatrix} I & (A'A)^{-1}A' \\ 0 & I \end{bmatrix}$. Then, $(A,B) = [A,(I - P_A)B]F$.

Hence, $C(A,B) \subseteq C[A,(I - P_A)B]$. Let $G = \begin{bmatrix} I & -(A'A)^{-1}A' \\ 0 & I \end{bmatrix}$. Then, $[A,(I - P_A)B] = (A,B)G$. Hence, $C(A,B) \supseteq C[A,(I - P_A)B]$.

Therefore, $C(A,B) = C[A,(I - P_A)B]$.

(ii) From (i), $\text{rank}(A,B) = \text{rank}[A,(I - P_A)B]$. Since $A'(I - P_A)B = 0$, $\text{rank}((I - P_A)B) = \text{rank}(A,B) - \text{rank}(A)$.

(iii) It is clear that $(I - P_A)B = (P_{A,B} - P_A)B$. Hence,
$C((I - P_A)B) \subseteq C(P_{A,B} - P_A)$. Note that $P_{A,B} - P_A$ is idempotent. Thus, $\text{rank}(P_{A,B} - P_A) = \text{tr}(P_{A,B} - P_A) = \text{tr}(P_{A,B}) - \text{tr}(P_A) = \text{rank}(A,B) - \text{rank}(A)$. Using (ii), $\text{rank}(P_{A,B} - P_A) = \text{rank}[(I - P_A)B]$. Hence, $C(P_{A,B} - P_A) = C((I - P_A)B)$.

Q.E.D.

Lemma 2

Suppose that $A$ is a symmetric real matrix of dimensions $k \times k$, and let $r = \text{rank}(A)$. Then, there is a $k \times k$ orthogonal matrix $P$ such that $P'AP = \begin{bmatrix} D_r & 0 \\ 0 & 0 \end{bmatrix}$, where $D_r$ is an $r \times r$ diagonal matrix. Moreover, the diagonal elements of $D_r$ are the $r$ non-zero eigenvalues of $A$. $P$ may be partitioned into $(R, U)$ where $R$ has $r$ columns so that $R'AR = D_r$, $U'AU = 0$. In addition, $U'A = 0$ and $R'A^2R = D_r^2$.

The proof of Lemma 2 is given, e.g., by Searle (1982, Chapter 11A).
9. BIBLIOGRAPHY


Wald, A. 1940. A Note on the Analysis of Variance with Unequal Class Frequencies. Annals of Mathematical Statistics

10. ACKNOWLEDGEMENTS

I wish to express my appreciation to my major professor, Dr. David A. Harville, for his guidance and patience, without which this dissertation would have been impossible.

Dr. Yasuo Amemiya, Dr. James Cornette, Dr. Oscar Kempthorne, and Dr. Glen Meeden are thanked for their valuable discussions and comments in my final oral examination. I am grateful to the staffs of the Numerical Analysis group of the Department of Statistics for their assistance in computational work of this dissertation. Special thanks go to both Dr. David Cox and Dr. Paul Hinz for their constant encouragement and friendship.

I am indebted to my wife (and typist), Nai-Huei, who showed her understanding and support in so many ways in this long and sometimes frustrating process of obtaining my degree.

This research was partly supported by the Office of Naval Research Contract N0014-85-K-0418.