SOME ALTERNATIVE MEASURES OF ASYMPTOTIC RELATIVE EFFICIENCY FOR THE MULTIPARAMETER TESTING PROBLEM WITH APPLICATION TO THE GROWTH CURVE PROBLEM

By

Robert Francis Woolson

Department of Biostatistics
University of North Carolina at Chapel Hill, N. C.

Institute of Statistics Mimeo Series No. 903

JANUARY 1974

SOME ALTERNATIVE MEASURES OF ASYMPTOTIC RELATIVE EFFICIENCY FOR THE MULTIPARAMETER TESTING PROBLEM WITH APPLICATION TO THE GROWTH CURVE PROBLEM

by

Robert Francis Woolson

A dissertation submitted to the faculty
of the University of North Carolina at
Chapel Hill in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy in the Department
of Biostatistics

Chapel Hill

1972

Approved by:

Adviser

Reader

Dana Quade
Reader

ABSTRACT

WOOLSON, ROBERT FRANCIS. Some Alternative Measures of Asymptotic Relative Efficiency for the Multiparameter Testing Problem with Application to the Growth Curve Problem. (Under the direction of PRANAB KUMAR SEN.)

Criteria for evaluating the loss in underfitting or overfitting a growth curve model are proposed. This problem has been formulated as one of comparing two test statistic sequences which have limiting chi-square distributions. As the degrees of freedom of the two competitive chi-square distributions may be different, alternative measures to the standard Pitman asymptotic relative efficiency (ARE) are considered. Two such criteria are the trace asymptotic relative efficiency (TARE) and the curvature asymptotic relative efficiency (CARE). Each of these quantities is shown to be the product of two factors: the first factor reflecting the degrees of freedom of the two tests and the common significance level, while the second factor is a function of the two noncentrality parameters. This second factor for the TARE is the ratio of the traces of the matrices in the noncentrality parameters while the second factor for the CARE is the q^{th} root of the ratio of the determinants of the matrices in the noncentrality parameters, where q is the number of parameters in the common hypothesis of interest. The CARE selects the test whose power function has the greater generalized Gaussian curvature at the null point, and the TARE selects the test with greater average local power, where the average is taken as the average over the family of spheres. The CARE and TARE are applied to the one-sample and multi-sample growth curve problems. Bounds for the TARE similar to the bounds which exist for the CARE are derived. It is shown that the c-sample efficiencies are multiples of the one-sample results. Numerical illustrations of the TARE and the CARE are presented for specific covariance matrices.

ACKNOWLEDGMENTS

The author appreciates the assistance of his advisor, Dr. P. K. Sen, and of the members of his advisory committee, Drs. S. K. Chatterjee, J. C. Cornoni, J. E. Grizzle, and D. E. Quade. He wishes to especially thank Dr. P. K. Sen for suggesting the topic of this dissertation and for thoughtful discussion at any time during the research. Special thanks also go to Dr. James E. Grizzle for serving as the author's course work advisor and for providing the author with the invaluable experience of working on numerous consulting projects with him.

The financial assistance of the Department of Biostatistics during the course of the author's graduate studies is gratefully acknowledged.

The author also wishes to thank his family for their encouragement during the course of this work. In particular he wishes to thank his wife, Linda, his son, Rob, and his mother, Mrs. Sallie Hagey.

The author also expresses sincere appreciation to Mrs. Gay Hinnant for the typing of the dissertation.

TABLE OF CONTENTS

LIST OF TABLES

Table | Page

NOTATION

$\underset{\sim}{B}$	is a matrix of real elements.
$\lvert\underset{\sim}{B}\rvert$	is the determinant of $\underset{\sim}{B}$.
$\lVert\underset{\sim}{\lambda}\rVert$	is the Euclidean norm of a vector $\underset{\sim}{\lambda}$.
$\mathcal{L}(\underset{\sim}{X})$	is the law or distribution of a random vector $\underset{\sim}{X}$.
$\xrightarrow{a.s.}$, $\xrightarrow{p}$	denote convergence almost surely and in probability respectively.
$tr(\underset{\sim}{B})$	is the trace of the matrix $\underset{\sim}{B}$.
$\underset{\sim}{A} \otimes \underset{\sim}{B}$	is the Kronecker or Direct Product of $\underset{\sim}{A}$ and $\underset{\sim}{B}$.
δ_{ij}	is the Kronecker Delta.
$ch(\underset{\sim}{B})$	denotes the set of characteristic roots of $\underset{\sim}{B}$.
R^t	denotes Euclidean t-space.
N_t	denotes the t dimensional normal distribution.
$\underset{\sim}{I}$	is the identity matrix.
$\underset{\sim}{J}$	is the matrix of ones.

CHAPTER I

INTRODUCTION AND REVIEW OF THE LITERATURE

1.1 Introduction

When two or more test procedures are available to test the same hypothesis we are faced with the problem of deciding which one to use. While the null hypothesis may be parametric or nonparametric, in most situations one is interested in a specific family of alternative hypotheses. Hence we assume that this family is parameterized by a parameter θ and, furthermore, that the value of θ when the null hypothesis is true is denoted by θ_0. For this reason we shall notationally represent the null hypothesis in this study as:

$$H_0: \ \theta = \theta_0$$

In this thesis the specific problem of comparing two test statistic sequences which have limiting chi-square distributions with possibly different degrees of freedom is to be studied. The main purpose of this work is to suggest and justify some measures of asymptotic relative efficiency (ARE) which may be used in comparing the two test sequences. Obviously, the measures of efficiency proposed should be such that, if the efficiency of test 2 relative to that of test 1 is greater than one, then test 2 possesses a limiting power function which in some sense is better than the limiting power function of test 1. In addition, it is desirable for any measure of ARE proposed to be free of arbitrary or unknown quantities and, furthermore, it should take on a single numerical value when sampling is performed from a completely specified distribution. Another purpose of this work is to apply the proposed criteria to the comparison of tests used for growth curve models. Of particular interest is the loss in efficiency when we underfit or overfit the true model. We shall see, however, that the efficiency criteria proposed are applicable to a larger set of problems and are not restricted to the study of growth curve problems alone.

1.2 Asymptotic Relative Efficiency

For comparing two α-level tests, ϕ_1 and ϕ_2, a reasonable measure of relative efficiency to use is:

$$RE(\alpha,\beta,\theta_a) = N_1/N_2$$

where N_1 and N_2 are the minimum sample sizes for which ϕ_1 and ϕ_2 at level α have power β against the alternative that $\theta = \theta_a$, where for simplicity we take θ to be scalar. To study this ratio for all values of (α,β,θ_a) is quite an involved study and, as an alternative, we may consider asymptotic comparisons of the two tests. One asymptotic approach is to fix $\theta_a, \theta_a \neq \theta_0$ and compare the limiting powers of ϕ_1 and ϕ_2 as sample size increases without bound for this fixed alternative. An obvious shortcoming of this approach is that if ϕ_1 and ϕ_2 are both consistent tests then both of the limiting powers are one and this method of comparison does not discriminate between the two tests. On the other hand, if the tests are inconsistent they are of little interest. Bahadur (1960) considers the comparison of the inverse ratio of the sample sizes in the limit for the two tests to achieve the same significance level; we shall see in Chapter II that this efficiency has serious deficiencies for our problem.

Pitman (1948) [see e.g. Noether (1955)] proposed that rather than considering a fixed alternative hypothesis, a sequence of alternative hypothesis depending on the sample size, N, be chosen such that the limit of this sequence approaches the null point and simultaneously the power is bounded away from one. In the Pitman sense we consider testing

$$H_0: \ \theta = \theta_0$$

against the sequence

$$H_N: \ \theta_N = \theta_0 + N^{-\delta} \lambda$$

where λ is a fixed but arbitrary nonzero real number and $\delta > 0$. The Pitman procedure for obtaining a measure of ARE is to consider two sequences of sample sizes, $\{N_1\}$ and $\{N_2\}$, chosen such that the limiting powers of the two tests through the sequence $\{\theta_N\}$, are the same. Then according to Pitman we have:

Definition: The asymptotic relative efficiency (ARE) of test ϕ_2 relative to test ϕ_1 is the limiting value of N_1/N_2 where N_1 is the number of observations required by test ϕ_1 for the power of test ϕ_1 to equal the power of ϕ_2 based on N_2 observations while simultaneously $N_2 \to \infty$ and $\theta_N \to \theta_0$.

If ϕ_1 and ϕ_2 are tests based on the statistics t_{N1} and t_{N2}, respectively, and if the limiting distributions of t_{Ni} through H_N converge to the normal distribution as $N \to \infty$, then with suitable restrictions on the limiting behavior of the first two moments of t_{Ni}

we may find the Pitman ARE quite easily. Denote the mean of t_{Ni} when $\theta = \theta_N$ by $\mu_{Ni}(\theta_N)$ and its variance by $\sigma^2_{Ni}(\theta_N)$. Consider the hypothesis H_0: $\theta = \theta_0$ against the one-sided sequence H_N: $\theta_N = \theta_0 + N^{-\delta}\lambda$ where $\lambda > 0$ (<0) and, suppose we have the following three conditions in addition to the limiting normality of t_{Ni}:

$$\lim_{N \to \infty} \frac{d\,\mu_{Ni}(\theta_0)}{d\theta} > 0 \quad (<0) \tag{1.2.1}$$

$$\lim_{N \to \infty} \frac{d\,\mu_{Ni}(\theta_0)}{d\theta} \Bigg/ \frac{d\,\mu_{Ni}(\theta_N)}{d\theta} = \lim_{N \to \infty} \frac{\sigma^2_{Ni}(\theta_N)}{\sigma^2_{Ni}(\theta_0)} = 1 \tag{1.2.2}$$

$$\lim_{N \to \infty} \frac{\dfrac{d\,\mu_{Ni}(\theta_0)}{d\theta}}{N^{\delta}\,\sigma_{Ni}(\theta_0)} = c_i > 0 \quad (<0) \tag{1.2.3}$$

then the ARE of ϕ_2 relative to ϕ_1 can be seen to be $\left(\dfrac{c_2}{c_1}\right)^2$ when $\delta = 1/2$ and in general it is $\left(\dfrac{c_2}{c_1}\right)^{1/\delta}$. The quantity c_i is called the efficacy of test ϕ_i.

Noether (1955) extended Pitman's definition of ARE to the case where the first $(m_i - 1)$ derivatives of $\mu_{Ni}(\theta)$ at θ_0 are zero and the $M_i\underline{th}$ derivative is nonzero at θ_0. With $M_1 = M_2 = M$, he found that the ARE of ϕ_2 relative to ϕ_1 is $\left(\dfrac{c_2}{c_1}\right)^{1/M\delta}$ where

$$c_i = \lim_{N \to \infty} \frac{\dfrac{d^M\,\mu_{Ni}(\theta_0)}{d\,\theta^M}}{N^{M\delta}\,\sigma_{Ni}(\theta_0)} .$$

Although not explicitly pointed out in Noether's article, it should be observed that if one does not require $M_1 = M_2$ then the ARE is indeterminate, depending as it does on the unknown λ_1 (or λ_2) raised to a power of $(M_2 - M_1)$. This point causes no practical limitation since if $M_1 \neq M_2$ then we are comparing tests which behave quite differently.

For the one sided alternative hypothesis, Blomqvist (1950) proposed an asymptotic local relative efficiency defined as the limit of the ratio of the sample sizes chosen so that the two limiting power functions have the same slope at θ_0. Under the conditions $M = 1$, $\delta = 1/2$ and

$$\frac{d\,\sigma_{Ni}(\theta_0)}{d\theta} \Big/ \sigma_{Ni}(\theta_0) = o(\sqrt{N}) \ ,$$

Noether (1955) established the equivalence of this definition and the Pitman definition. Kendall and Stuart (1963) discuss other cases and show, for example, for the two sided test with $M = 1$ that under mild restrictions the ARE is equal to the limiting ratio of the second derivatives of the power functions of θ_0.

It has been observed [see e.g. Puri & Sen (1971)] that the theory for comparing the two limiting distributions does not require asymptotic normality. It is sufficient that the two power functions can be made analytically the same by an appropriate choice of the sample sizes. They present the requisite theory for the comparison of two test statistics which have noncentral chi-square distributions both with the same degrees of freedom, p. In this case, since the noncentral chi-square is a monotonically increasing function of its noncentrality parameter, they found that when $M = 1$ and $\delta = 1/2$ that the ARE is simply the ratio of the two respective noncentrality parameters. They point out that this definition is entirely equivalent to the Pitman-Noether definition for comparison of tests based on one degree of freedom.

Turning now to the multiparameter testing problem we consider H_0: $\underset{q\times 1}{\underset{\sim}{\theta}} = \underset{\sim}{\theta}_0$ and the sequence of alternative hypotheses,

H_N: $\underset{\sim}{\theta}_N = \underset{\sim}{\theta}_0 + N^{-\delta} \underset{\sim}{\lambda}$ where $\underset{\sim}{\lambda}$ is a fixed but arbitrary non-null vector and $\delta > 0$. Puri and Sen study the case $p = q$ in their text and when $M = 1$, $\delta = 1/2$ observe that the Pitman-Noether efficiency is the ratio of two positive definite quadratic forms in $\underset{\sim}{\lambda}$. Clearly, in this case, there is no unique answer regarding the ARE since it depends, in general, on the arbitrary vector $\underset{\sim}{\lambda}$. By application of a theorem due to Courant on the extrema of the ratio of two positive definite quadratic forms [Puri & Sen (1971, p. 122)], bounds may be placed on the ARE over all non-null $\underset{\sim}{\lambda}$. This is, in fact, done for a number of cases in their text. The approach of placing bounds on the ARE provides some information on the ARE. However, some tests may be placed in a misleading position in the spectrum of tests for a given problem since the minimum bound on the ARE could be quite near zero, while the typical or average performance of the test may be better than its competitors. To develop appropriate average measures free from $\underset{\sim}{\lambda}$ is one of the goals of this investigation. In our study we shall consider a Pitman type of sequence of alternative hypotheses; however, a comparison in the sense of Bahadur (1960) is a possibility. Hence, we shall summarize the necessary points for the Bahadur criterion of efficiency.

In the Bahadur definition of efficiency it is assumed that we have a family of probability measures parameterized by $\underset{\sim}{\theta} \in \Omega$, defined on the same probability space. The parameter set is denoted by Ω and is partitioned into Ω_0 and $\Omega - \Omega_0$ where Ω_0 is the single point set containing $\underset{\sim}{\theta}_0$. Let us suppose that the two competing test statistic sequences are $\{T_{N1}\}$ and $\{T_{N2}\}$.

Definition: The sequence of test statistics $\{T_{Ni}\}$ is called a <u>standard</u> <u>sequence</u> for testing H_0 if, and only if, $\{T_{Ni}\}$ satisfies the following conditions

I. there exists a continuous distribution function $F_i(x)$ such that $\lim_{N \to \infty} \Pr_{\theta_0} [T_{Ni} \leq x] = F_i(x)$ for all real x,

II. there exists a constant $a_i \varepsilon (0,\infty)$ such that

$$-\log_e [1 - F_i(x)] = \frac{a_i x^2}{2} [1 + o(1)] \text{ as } x \to \infty,$$

III. there exists a function $b_i(\underset{\sim}{\theta})$ from $\Omega - \Omega_0$ to the positive real line such that for every $\varepsilon > 0$ we have

$$\lim_{N \to \infty} \Pr_\theta [\bigcup_{K \geq N} [|K^{-1/2} T_{Ki} - b_i(\theta)| > \varepsilon]] = 0$$

for all $\theta \varepsilon \Omega - \Omega_0$.

If $\{T_{Ni}\}$ is a standard sequence for testing H_0 then the quantity $a_i b_i^2(\theta)$ is called the asymptotic slope of the sequence $\{T_{Ni}\}$. With this in mind, if both $\{T_{N1}\}$ and $\{T_{N2}\}$ are standard sequences for testing H_0, then the Bahadur efficiency of test 2 to test 1 is $\frac{a_2 b_2^2(\theta)}{a_1 b_1^2(\theta)}$ for all $\theta \varepsilon \Omega - \Omega_0$. It can be shown that the Bahadur efficiency is the limit of the inverse ratio of the sample sizes needed for the two tests to have the same level of significance in large samples. In the later chapters we will show that under mild assumptions the Bahadur efficiency of ϕ_2 relative to ϕ_1 is

$\underset{\sim}{\eta}_2'(\underset{\sim}{\theta})\ \underset{\sim}{\Sigma}_2^{-1}(\underset{\sim}{\theta})\ \underset{\sim}{\eta}_2(\underset{\sim}{\theta})\ /\ \underset{\sim}{\eta}_1'(\underset{\sim}{\theta})\ \underset{\sim}{\Sigma}_1^{-1}(\underset{\sim}{\theta})\ \underset{\sim}{\eta}_1(\underset{\sim}{\theta})$ for $\underset{\sim}{\theta} \varepsilon \Omega - \Omega_0$, where both $\underset{\sim}{\eta}_i(\underset{\sim}{\theta})$ and $\underset{\sim}{\Sigma}_i(\underset{\sim}{\theta})$ depend on the alternative value of $\underset{\sim}{\theta}$. The Bahadur efficiency will also be shown to be insensitive to differences in degrees of freedom of the two tests.

The measures of ARE we shall propose will be defined using various notions of test optimality. Hence, we shall review the relevant areas of the theory of hypothesis testing in the following section.

1.3 Review of Optimal Parametric Tests

Uniformly most powerful tests (or critical regions) are known to exist in only the rarest situations. In order to derive tests with uniformly most powerful properties when restricting attention to a subset of all available tests, Neyman and Pearson (1936) generalized their fundamental lemma (for testing a simple hypothesis against a simple alternative) to power functions subject to more than one side condition. This generalized lemma is stated and proven in several places (e.g. Lehmann (1959)). For testing the hypothesis H_0: $\theta = \theta_0$ against the two-sided alternative H_a: $\theta \neq \theta_0$, where θ is a one-dimensional parameter, Neyman & Pearson (1936) propose the type A critical region which may be described as the locally best unbiased critical region. This test is the one which maximizes the second derivative of the power function evaluated at θ_0 subject to size and unbiasedness restrictions on the power functions. For the two parameter testing problem, Neyman & Pearson (1938) proposed type C critical regions; these regions can be constructed if one knows the relative importance locally of type II errors, since this region is defined to be the one with best local power along a given family of concentric ellipses with the same shape and direction of principal axes. If the family of ellipses are concentric circles then we say the type C region is <u>regular</u> otherwise it is said to be <u>nonregular</u>. Two main objections to type C regions can be raised; in the first place, one may not be able to state the relative importance of type II errors and, secondly, regular regions are not invariant under one to one, twice differentiable transformations of the parameter space. The last point simply means that regular regions can become nonregular regions even under some elementary transformations of the parameter space. To overcome these problems Isaacson (1951) proposed a type D critical region which does not require knowledge of the relative importance of type II errors. To motivate the type D test Isaacson observes that the type A power function satisfied an attractive geometrical property. Namely, if one considers a horizontal chord drawn at a fixed infinitesimal distance above θ_0 that the length of this chord, for the type A power function is a minimum when compared to the length of the chord for any other of the power functions satisfying the stated conditions of size and unbiasedness.

A type D region can be defined in the q parameter testing problem as that region which maximizes the generalized Gaussian curvature of the power function at $\underset{\sim}{\theta}_0$ subject to size and unbiasedness conditions. In order to be more specific let us denote by $\beta(\underset{\sim}{\theta}|\omega)$, the power function of a test with critical region ω, and let $\beta^{(i)}(\underset{\sim}{\theta}_0|\omega)$ denote the first partial derivative of $\beta(\underset{\sim}{\theta}|\omega)$ with respect to θ_i at $\underset{\sim}{\theta}_0$, $i = 1,2,\ldots,q$. Further let $\beta^{(i,j)}(\underset{\sim}{\theta}_0|\omega)$ denote the second partial derivative of $\beta(\underset{\sim}{\theta}|\omega)$ with respect to $\theta_i\theta_j$ evaluated at

$\underset{\sim}{\theta} = \underset{\sim}{\theta}_0$, $i, j = 1,2,\ldots,q$. Letting the determinant of a matrix A be denoted by $|A|$, then the following is Isaacson's definition of a type D region for testing H_0: $\underset{\sim}{\theta} = \underset{\sim}{\theta}_0$:

Definition: A region ω_0 is said to be an <u>unbiased critical region of type</u> D for testing H_0 if:

a. $\beta(\underset{\sim}{\theta}_0|\omega_0) = \alpha$,

b. $\beta^{(i)}(\underset{\sim}{\theta}_0|\omega_0) = 0 \quad i = 1,2,\ldots,q$,

c. $((\beta^{(i,j)}(\underset{\sim}{\theta}_0|\omega_0)))$ is positive definite,

d. $|((\beta^{(i,j)}(\underset{\sim}{\theta}_0|\omega_0)))| \geq |((\beta^{(i,j)}(\underset{\sim}{\theta}_0|\omega)))|$ for any other region ω satisfying conditions a, b and c.

In the two-dimensional case it follows that the type D region minimizes the area of an ellipse at an infinitesimal distance above the point $\underset{\sim}{\theta}_0$, and in the general q parameter case the type D region would be the one whose power function minimized the volume of a certain ellipsoid at a given cross-section of the power function. The type D region is then a generalization of the type A region to more than one dimension. Type D regions are characterized by Isaacson by use of the generalized Neyman-Pearson lemma. We know the generalized Gaussian curvature of $\beta(\underset{\sim}{\theta}|\omega)$ at $\underset{\sim}{\theta}_0$ is:

Definition: The generalized Gaussian curvature of $\beta(\underset{\sim}{\theta}|\omega)$ at $\underset{\sim}{\theta}_0$ is

$$K = |((\beta^{(i,j)}(\underset{\sim}{\theta}_0|\omega)))|/(1 + \sum_{j=1}^{q} \beta^{(j)^2}(\underset{\sim}{\theta}_0|\omega))^2$$

Thus, if a test is unbiased then $K = |((\beta^{(i,j)}(\underset{\sim}{\theta}_0|\omega)))|$. If we compare two critical regions (tests) then the test with the larger K has a power function which encloses an ellipsoid of smaller volume than the other power function along any of a family of infinitesimal contours. This presumes, of course, that both of the $((\beta^{(i,j)}(\underset{\sim}{\theta}_0|\omega))$'s are positive definite since a lack of definiteness of this matrix would mean that the intersection of the power function with the fixed hyperplane would not be an ellipsoid.

Wald (1943) defines a critical region ω_0 to have uniformly best average power with respect to a family of surfaces, K_c, and a weight function, $g(\underset{\sim}{\theta})$, if for any other region ω the surface integral of the power function of ω_0 times $g(\underset{\sim}{\theta})$ over K_c is greater than the surface integral of the power function of ω times $g(\underset{\sim}{\theta})$ over the surface K_c. If we define the surface as the unit sphere, i.e. $||\underset{\sim}{\lambda}|| = 1$, then we can show that the surface integral of a quadratic form $\underset{\sim}{\lambda}' \underset{\sim}{B} \underset{\sim}{\lambda}$ over $||\underset{\sim}{\lambda}|| = 1$ is proportional to the trace of the matrix $\underset{\sim}{B}$.

That is, $\int_{||\underset{\sim}{\lambda}||=1} \underset{\sim}{\lambda}' \underset{\sim}{B} \underset{\sim}{\lambda} \, d A = k \operatorname{tr}(\underset{\sim}{B})$. Using this fact, it will be possible to deduce that of two tests the one with a larger trace of $((\beta^{(i,j)}(\underset{\sim}{\theta}_0|\omega)))$ has greater average power locally over the family of spheres than the other.

1.4 Organization of the Study

Since many of the standard parametric and nonparametric multivariate test procedures are based on test statistics which are quadratic forms, we shall consider in the general development in Chapter II, two sequences of statistics which are quadratic forms in two sequences of random vectors. When the null hypothesis, $H_0: \underset{q\times 1}{\underset{\sim}{\theta}} = \underset{\sim}{\theta}_0$, is true it is assumed that each of the sequences of test statistics have limiting central chi-square distributions; one with t_1 degrees of freedom, the other with t_2 degrees of freedom. Sufficient conditions are presented under which we may compute the limiting power of our statistics through the sequence of alternative hypotheses H_N. Simplifications of the power functions are obtained in terms of the parameters. Three new definitions of ARE are proposed in Chapter II. These are: a) local asymptotic relative efficiency (LARE), b) curvature asymptotic relative efficiency (CARE) and c) trace asymptotic relative efficience (TARE). The precise definitions are presented in Chapter II. All three criteria depend on the level of significance of the tests, α, and the degrees of freedom, t_1 and t_2. Tabulations in Chapter II show the dependence on α is slight while the dependence on $t_2 - t_1$ is strong. In addition, all three criteria depend on the noncentrality parameters and the latter two criteria are shown to be "average" measures of efficiency, independent of the direction of approach of $\underset{\sim}{\theta}_N$ to $\underset{\sim}{\theta}_0$. The second criterion of ARE is a function of the ratio of determinants of the noncentrality parameters, while the trace criterion is a function of the ratio of the traces of the noncentrality parameters. There is, it seems, an interesting connection between these criteria and type D and type E optimality in the field of experimental design. Chapter II is concluded with a brief study of the Bahadur efficiency and its relationship to the LARE in a limiting sense.

Chapters III and IV are applications of the measures of ARE to the one-sample and multisample growth curve problems, respectively. Polynomial models are studied and the efficiency results are presented for the cases of underfitting and overfitting the correct model. These efficiency results are presented for the least squares procedures and the rank scores tests. The least squares and nonparametric procedures are also compared using the curvature and trace criteria. Bounds using the trace criterion are obtained similar to those known for the curvature criterion. The use of the higher order polynomial terms as covariables is also studied and the ARE is evaluated in this case.

Chapter V contains numerical computations of the results obtained in Chapters III and IV for some special cases.

CHAPTER II

MEASURES OF ASYMPTOTIC RELATIVE EFFICIENCY FOR THE MULTIPARAMETER TESTING PROBLEM

2.1 Introduction and Summary

The purpose of this chapter is to propose and study several competitive measures of asymptotic relative efficiency (ARE) for the multiparameter testing problem. We shall assume throughout that we have two sequences of test statistics available for testing the same hypothesis. It is customary in both parametric and nonparametric inference to consider a specified type of alternative hypothesis, e.g. translation alternatives, scale alternatives, etc. We label this family of alternatives by a parameter θ and we let θ_0 be the value of θ when the hypothesis we wish to test is true. In the text of this chapter we shall occasionally refer to the null hypothesis as H_0: $\theta = \theta_0$; however, we should keep in mind that the hypothesis may in fact be much more general. To compare the two test statistic sequences in large samples we shall consider a Pitman type sequence of alternative hypotheses and we shall present in the form of a theorem (Theorem 2.2.1), sufficient conditions under which these test statistics have limiting chi-square distributions through the sequence of alternative hypotheses. Under the assumptions of Theorem 2.2.1 we shall study in detail the limiting power functions of the statistics and obtain several simplifications. To derive these results we shall find it useful to prove some results for homogeneous polynomials in general and then apply these results to the expanded power functions. In order to compare the two test statistics we consider a common sequence of alternative hypotheses and derive the Pitman ARE of test 2 to test 1 when they both have limiting chi-square distributions with equal degrees of freedom. When the degrees of freedom are unequal we propose a local asymptotic relative efficiency (LARE) for the comparison of the two tests. This LARE is found to be equal to a scalar function, R, multiplied by the ratio of the noncentrality parameters of the limiting power functions. The function, R, has as its arguments the common significance level of the tests, α, and the two respective degrees of freedom of the tests. In the general setting $\underset{\sim}{\theta}$ is a vector of q parameters and the i^{th} test statistic has a limiting power function with t_i degrees of freedom for $i = 1,2$. The power depends also on a positive integer, M, (defined by the conditions of Theorem 2.2.1). In considering other measures of ARE we briefly discuss the problems when $M > 1$ and then restrict ourselves to the case $M = 1$. When at least one $t_i \geq q$ for $i = 1$ or 2 we define a measure of ARE based on the generalized Gaussian curvature of the limiting power of functions at the null value of $\underset{\sim}{\theta}$. We shall show that this criterion of ARE produces a quantity which is independent of the sequence of alternative hypotheses and furthermore when $t_1 = t_2$ this criterion reduces to the ratio of the geometric means of the characteristic roots of the matrices in the noncentrality parameters. When $t_1 < q$ and $t_2 < q$, the generalized Gaussian curvatures of the power functions are zero at $\underset{\sim}{\theta}_0$ and this method of comparison is useless. In this situation we propose a criterion of comparison based on the average power of the local power functions over a given family of surfaces. This approach leads to comparison of the trace of the matrices in the noncentrality parameters multiplied by the function, R.

We consider at the end of this chapter the efficiency criterion proposed by Bahadur (1960). We present sufficient conditions under which the statistics we consider form a standard sequence for testing the null hypothesis, and obtain the Bahadur efficiency of test 2 to test 1. It is shown that this efficiency (Bahadur) depends on the unknown alternative $\underset{\sim}{\theta}$ in both the covariance matrix and the location vector. The difference in degrees of freedom of the two tests is also not reflected in this efficiency criterion. We discuss briefly the limiting form of the Bahadur efficiency as $\underset{\sim}{\theta} \to \underset{\sim}{\theta}_0$ and under certain assumptions show that this limiting situation is equivalent to the LARE without the adjustment for degrees of freedom of the tests. The insensitivity of the Bahadur efficiency to differences in degrees of freedom of the tests and the complicated manner in which this efficiency depends on the unknown alternative value of $\underset{\sim}{\theta}$ lead us to dismiss this criterion as a possible mode of comparison of the two tests.

2.2 The Test Statistics and Their Limiting Distributions

For testing a specified null hypothesis against a parametric family of alternatives we have two tests of size α available, ϕ_1 and ϕ_2. The family of alternative hypotheses is parameterized by the vector $\underset{\sim}{\theta}$, which has q elements. When the null hypothesis is true the value of $\underset{\sim}{\theta}$ is denoted by $\underset{\sim}{\theta}_0$; for this reason we sometimes denote the null hypothesis as:

$$(2.2.1) \qquad H_0\colon\ \underset{\sim}{\theta} = \underset{\sim}{\theta}_0$$

and the alternative by

$$H_a\colon\ \underset{\sim}{\theta} = \underset{\sim}{\theta}_a \neq \underset{\sim}{\theta}_0.$$

The tests, ϕ_1 and ϕ_2, are functions of the statistics $\{Q_N^{(1)}\}$ and $\{Q_N^{(2)}\}$, respectively. These statistic sequences are further assumed to be quadratic forms in random vector sequences $\{T_N^{(1)}\}$ and $\{T_N^{(2)}\}$, respectively. Each vector $\{T_N^{(i)}\}$ is composed of t_i elements and the test statistics are written as:

$$(2.2.2) \qquad Q_N^{(i)} = N(\underset{\sim}{T}_N^{(i)} - \underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta}_0))' \, \hat{\underset{\sim}{\Sigma}}{}_N^{(i)^{-1}} \, (\underset{\sim}{T}_N^{(i)} - \underset{\sim}{\mu}_N^{(i)}(\theta_0));\ i = 1,2$$

where $\sqrt{N}\ \underset{\sim}{\mu}_N^{(i)}\ (\underset{\sim}{\theta}_0)$ is the mean vector of $\sqrt{N}\ T_N^{(i)}$ when $\underset{\sim}{\theta} = \underset{\sim}{\theta}_0$ and $\hat{\underset{\sim}{\Sigma}}{}_N^{(i)}$ is some consistent estimator of the covariance matrix of $\sqrt{N}\ \underset{\sim}{T}_N^{(i)}$. The tests would reject the null hypothesis for large $Q_N^{(i)}$ and accept the null hypothesis for small $Q_N^{(i)}$. We may assume that the discriminant of $Q_N^{(i)}$ is of full rank, since if it is not it is always possible to express $Q_N^{(i)}$ as a quadratic form in fewer variables whose discriminant would be nonsingular. To illustrate this point more clearly let us consider the multivariate multisample location problem. One statistic in common use is the Hotelling-Lawley Trace which is defined as:

$$(2.2.3) \qquad T_N^2 = \sum_{K=1}^{c} n_K \sum_{i=1}^{p} \sum_{j=1}^{p} s^{ij} (\bar{x}_K^{(i)} - \bar{\bar{x}}^{(i)})(\bar{x}_K^{(j)} - \bar{\bar{x}}^{(j)}) \text{ where } ((s^{ij}))$$

$$= (((N-c)^{-1} \sum_{K=1}^{c} \sum_{\ell=1}^{n_K} (x_{K\ell}^{(i)} - \bar{x}_K^{(i)})(x_{K\ell}^{(j)} - \bar{x}_K^{(j)}))^{-1} \qquad i,j = 1,\ldots,p$$

$$N = \sum_{K=1}^{c} n_K$$

T_N^2 is used to test the null hypothesis:

(2.2.4) $$H_0: \underset{p\times 1}{\underset{\sim}{\mu}_1} = \underset{p\times 1}{\underset{\sim}{\mu}_2} = \cdots = \underset{p\times 1}{\underset{\sim}{\mu}_c}$$

where $\underset{\sim}{\mu}_K$ is the location vector of an observation from population K. We can rewrite $\underset{\sim}{\mu}_K$ as

$$\underset{\sim}{\mu}_K = \underset{\sim}{\mu} + \underset{\sim}{\gamma}_K$$

where

$$\sum_{K=1}^{c} = \frac{n_K}{N} \underset{\sim}{\gamma}_K = \underset{\sim}{0}.$$

With this restriction the hypothesis (2.2.4) can be written equivalently as

(2.2.5) $$H_0: \underset{\sim}{\gamma}_1 = \underset{\sim}{\gamma}_2 = \cdots = \underset{\sim}{\gamma}_{c-1} = \underset{\sim}{0}.$$

If $\underset{\sim}{x}_{K1}, \ldots, \underset{\sim}{x}_{K,n_K}$ are independent and identically distributed p-variate normal vectors with mean $\underset{\sim}{\mu}_K$ and covariance matrix $\underset{\sim}{\Sigma}$, symmetric and positive definite, then

$$\bar{\underset{\sim}{x}}_K = n_K^{-1} \sum_{\ell=1}^{n_K} \underset{\sim}{x}_{K_\ell} \text{ is distributed } N_p(\underset{\sim}{\mu}_K, n_K^{-1} \underset{\sim}{\Sigma})$$

and

$$\bar{\bar{\underset{\sim}{x}}} = N^{-1} \sum_{K=1}^{c} n_K \bar{\underset{\sim}{x}}_K \text{ is distributed } N_p\Big(\sum_{K=1}^{c} \frac{n_K}{N} \underset{\sim}{\mu}_K, N^{-1} \underset{\sim}{\Sigma}\Big).$$

As a result we see that for each K = 1,2,...,c we have

$$(\bar{\underset{\sim}{x}}_K - \bar{\bar{\underset{\sim}{x}}}) \text{ is distributed } N_p(\underset{\sim}{\mu}_K - \underset{\sim}{\mu}, (n_K^{-1} - N^{-1}) \underset{\sim}{\Sigma}).$$

The covariance of $(\bar{\underset{\sim}{x}}_K - \bar{\bar{\underset{\sim}{x}}})(\bar{\underset{\sim}{x}}_{K'} - \bar{\bar{\underset{\sim}{x}}})'$ is $- N^{-1}\underset{\sim}{\Sigma}$ and consequently the joint distribution of $[(\underset{\sim}{\bar{x}}_K-\underset{\sim}{\bar{\bar{x}}});\ K = 1,2,\ldots,c]$ is pc-variate normal with mean $[(\underset{\sim}{\mu}_K-\underset{\sim}{\mu}),\ K = 1,2,\ldots,c]$ and covariance matrix defined by:

$$\underset{\sim}{\Sigma} \otimes \left(\frac{\delta_{Kq}}{n_K} - \frac{1}{N}\right)_{K,q = 1,2,\ldots,c} \tag{2.2.6}$$

with the restriction $\sum_{K=1}^{c} n_K(\underset{\sim}{\bar{x}}_K-\underset{\sim}{\bar{\bar{x}}}) = \underset{\sim}{0}$, we see that the rank of this distribution is p(c-1), which means the quadratic form in (2.2.3), if viewed as a quadratic form in $[(\underset{\sim}{\bar{x}}_K-\underset{\sim}{\bar{\bar{x}}});\ K = 1,2,\ldots,c]$, has a discriminant of rank p(c-1). On the other hand if we consider the p(c-1) vector;

$$[(\underset{\sim}{\bar{x}}_K-\underset{\sim}{\bar{\bar{x}}});\ K = 1,2,\ldots,c-1], \tag{2.2.7}$$

this is p(c-1) variate normal with mean

$$[(\underset{\sim}{\mu}_K-\underset{\sim}{\mu});\ K = 1,2,\ldots,c-1]$$

and covariance matrix given by

$$\underset{\sim}{\Sigma} \otimes \left(\frac{\delta_{Kq}}{n_K} - \frac{1}{N}\right)_{K,q = 1,2,\ldots,c-1} \tag{2.2.8}$$

The inverse of (2.2.8) is

$$\underset{\sim}{\Sigma} \otimes \left(\frac{n_K\ \delta_{Kq}}{N} + \frac{n_K\ n_q}{n_c}\right)_{K,q = 1,2,\ldots,c-1} \tag{2.2.9}$$

Since we have sampled from normal populations we know from the strong laws of large numbers that $\underset{\sim}{S} = ((s_{ij})) \overset{a.s.}{\rightarrow} \underset{\sim}{\Sigma}$ as $N \to \infty$. So a consistent estimate of (2.2.9) is provided when we substitute $\underset{\sim}{S}^{-1}$ in place of $\underset{\sim}{\Sigma}^{-1}$ in (2.2.9). With some algebra, and using the restriction

$\sum_{K=1}^{c} n_K(\bar{\underset{\sim}{x}}_K - \bar{\bar{\underset{\sim}{x}}}) = \underset{\sim}{0}$, we see that T_N^2, defined by (2.2.3), can be written equivalently as:

$$(2.2.10)\quad [\sqrt{N}(\bar{\underset{\sim}{x}}_K - \bar{\bar{\underset{\sim}{x}}}), K=1,2,\ldots,c-1]'(\underset{\sim}{S} \otimes (\frac{N\delta_{Kq}}{n_K} - 1))^{-1}[\sqrt{N}(\bar{\underset{\sim}{x}}_K - \bar{\bar{\underset{\sim}{x}}}), K = 1,2,\ldots,c-1]$$

But (2.2.10) is a quadratic form whose discriminant is of full rank, so we see that quadratic forms can be written in a "full rank" form. In the theory presented in this chapter, we may, therefore, consider without loss of generality that the quadratic forms, which we use as test statistics, to have nonsingular distributions. We now present a theorem which allows us to compute the limiting power of our statistics through a sequence of alternative hypotheses.

Theorem 2.2.1. For testing a specified hypothesis H_0 against a sequence of alternative hypothesis H_N, defined by:

$$H_N\colon \underset{q\times 1}{\underset{\sim}{\theta}_N} = \underset{q\times 1}{\underset{\sim}{\theta}_0} + N^{-\delta} \underset{q\times 1}{\underset{\sim}{\lambda}},$$

where $\delta > 0$, $\underset{\sim}{\lambda}$ is a fixed non-null vector and $\underset{\sim}{\theta}_0$ is the value of $\underset{\sim}{\theta}$ when H_0 is true; we have a t- vector $N^{M\delta}\underset{\sim}{T}_N$ with mean $N^{M\delta}\underset{\sim}{\mu}_N(\underset{\sim}{\theta})$ and positive definite covariance matrix $\underset{\sim}{\Sigma}(\underset{\sim}{\theta})$. Suppose the following five conditions are true:

$$\text{a)}\quad \frac{1}{r!}\left(\sum_{\ell=1}^{q} N^{-\delta}\lambda_\ell \frac{\partial}{\partial\theta_\ell}\right)^r \mu_{N,j}(\underset{\sim}{\theta}_0) = 0 \qquad \text{for } r = 1,2,\ldots,M-1 \text{ for } j = 1,2,\ldots,t$$

(2.2.11)

$$\text{b)}\quad \frac{1}{M!}\left(\sum_{\ell=1}^{q} N^{-\delta}\lambda_\ell \frac{\partial}{\partial\theta_\ell}\right)^M \mu_{N,j}(\underset{\sim}{\theta}_0) \neq 0 \qquad \text{for some } M \geq 1 \text{ and at least one } j \in (1,2,\ldots,t)$$

(2.2.12)

a) $\text{ch}\ \hat{\underset{\sim}{\Sigma}}_N\ \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_N) \xrightarrow{P} 1$ as $N \to \infty$

b) $\text{ch}\ \underset{\sim}{\Sigma}(\underset{\sim}{\theta})\ \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_0) \to 1$ as $\underset{\sim}{\theta} \to \underset{\sim}{\theta}_0$

(2.2.13) the M^{th} partial derivatives of $\mu_{Nj}(\underset{\sim}{\theta})$ are continuous at $\underset{\sim}{\theta}_0$ for each $j = 1,2,\dots,t$

(2.2.14) $\lim\limits_{N \to \infty} \left(\sum_{\ell=1}^{q} \lambda_\ell \frac{\partial}{\partial\theta_\ell}\right)^M \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_0) = \underset{t\times 1}{\underset{\sim}{c}}$

(2.2.15) $\mathcal{L}[N^{M\delta}(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_N))] \to N_t(\underset{\sim}{0},\underset{\sim}{\Sigma}(\underset{\sim}{\theta}_0))$ as $N \to \infty$ uniformly in $\underset{\sim}{\lambda}$, and the distribution is non-degenerate, then

(2.2.16) $\mathcal{L}[N^{2M\delta}(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_0))'\ \hat{\underset{\sim}{\Sigma}}_N^{-1}\ (\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_0))] \to \chi^2(t,\Delta)$ as $N \to \infty$,

where $\chi^2(t,\Delta)$ is a noncentral chi-square with t degrees of freedom and noncentrality parameter Δ and

(2.2.17)
$$\Delta = \frac{\underset{\sim}{c}'\ \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_0)\ \underset{\sim}{c}}{(M!)^2} .$$

Proof:

For each $j = 1,2,\dots,t$ by (2.2.11) and (2.2.13) we can write by Taylor's theorem and the mean-value theorem:

$$\mu_{Nj}(\underset{\sim}{\theta}_N) = \mu_{Nj}(\underset{\sim}{\theta}_0) + \frac{1}{M!}\left(\sum_{\ell=1}^{q} N^{-\delta} \lambda_\ell \frac{\partial}{\partial\theta_\ell}\right)^M \mu_{Nj}\ (\underset{\sim}{\theta}^*)$$

where

$$\underset{\sim}{\theta}^* = \underset{\sim}{\theta}_0 + h\, N^{-\delta}\, \underset{\sim}{\lambda};\ h \in (0,1),$$

therefore

$$N^{M\delta}\left[\mu_{Nj}(\underset{\sim}{\theta}_N) - \mu_{Nj}(\underset{\sim}{\theta}_0)\right] = \frac{1}{M!}\left(\sum_{\ell=1}^{q} \lambda_\ell \frac{\partial}{\partial\theta_\ell}\right)^M \mu_{Nj}(\underset{\sim}{\theta}_0) + \varepsilon_{Nj}(\underset{\sim}{\theta}^*,\underset{\sim}{\theta}_0)$$

where

$$\varepsilon_{Nj}(\underset{\sim}{\theta}^*,\underset{\sim}{\theta}_0) = \frac{1}{M!}\left(\sum_{\ell=1}^{q} \lambda_\ell \frac{\partial}{\partial\theta_\ell}\right)^M \mu_{Nj}(\underset{\sim}{\theta}^*) - \frac{1}{M!}\left(\sum_{\ell=1}^{q} \lambda_\ell \frac{\partial}{\partial\theta_\ell}\right)^M \mu_{Nj}(\underset{\sim}{\theta}_0)$$

But $\varepsilon_{Nj}(\underset{\sim}{\theta}^*,\underset{\sim}{\theta}_0) = o(1)$ as $N \to \infty$ because $\underset{\sim}{\theta}^* \to \underset{\sim}{\theta}_0$ as $N \to \infty$; hence by (2.2.13) $\varepsilon_{Nj}(\underset{\sim}{\theta}^*,\underset{\sim}{\theta}_0) = o(1)$. So

$$N^{M\delta}\left[\underset{\sim}{\mu}_N(\underset{\sim}{\theta}_N) - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_0)\right] = \frac{1}{M!}\left(\sum_{\ell=1}^{q} \lambda_\ell \frac{\partial}{\partial\theta_\ell}\right)^M \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_0) + \underset{\sim}{o}(1) \tag{2.2.18}$$

Let us define Q_N and Q_N^* by the following

$$Q_N = N^{2M\delta}\left(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_0)\right)' \hat{\underset{\sim}{\Sigma}}_N^{-1}\left(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_0)\right)$$

$$Q_N^* = N^{2M\delta}\left(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_0)\right)' \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_N)\left(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\theta_0)\right).$$

We first shall show that $\mathcal{L}(Q_N^*) \to \chi^2(t,\Delta)$ as $N \to \infty$, hence it would follow that $\mathcal{L}(Q_N) \to \chi^2(t,\Delta)$ as $N \to \infty$, as we shall show $|Q_N - Q_N^*| \overset{P}{\to} 0$.

Now $N^{M\delta}\left(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_0)\right) = N^{M\delta}\left(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_N)\right) + N^{M\delta}\left(\underset{\sim}{\mu}_N(\underset{\sim}{\theta}_N) - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_0)\right)$

and by (2.2.18) we have as $N \to \infty$.

$$N^{M\delta}\left(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_0)\right) = N^{M\delta}\left(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_N)\right) + \frac{1}{M!}\left(\sum_{\ell=1}^{q} \lambda_\ell \frac{\partial}{\partial\theta_\ell}\right)^M \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_0) + o(1) \tag{2.2.19}$$

By (2.2.15) and since $N^{2M\delta}\left(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_N)\right)' \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_0)\left(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_N)\right)$ is a continuous function of $N^{M\delta}\left(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_N)\right)$ we have that

$$\mathcal{L}(N^{2M\delta}(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_N))' \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_0)(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_N))) \rightarrow \chi^2(t) \text{ as } N \rightarrow \infty$$

but by (2.2.12b) and the fact that $\underset{\sim}{\theta}_N \rightarrow \underset{\sim}{\theta}_0$ as $N \rightarrow \infty$ we also have that

$$\mathcal{L}[N^{2M\delta}(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_N))' \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_N)(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_N))] \rightarrow \chi^2(t) \text{ as } N \rightarrow \infty.$$

Let

$$P_N = N^{2M\delta}(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_N))' \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_N)(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_N))$$

by (2.2.19) we can write Q_N^* as

$$Q_N^* = P_N + [\frac{1}{M!}(\sum_{\ell=1}^{q} \lambda_\ell \frac{\partial}{\partial\theta_\ell})^M \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_0)]' \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_N)[\frac{1}{M!}(\sum_{\ell=1}^{q} \lambda_\ell \frac{\partial}{\partial\theta_\ell})^M \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_0)] + o(1)$$

By (2.2.12b) we can replace $\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_N)$ by $\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_0) + \underset{\sim}{\Gamma}$ where $\underset{\sim}{\Gamma} = \underset{\sim}{o}(1)$ as $N \rightarrow \infty$ hence

$$Q_N^* = P_N + [\frac{1}{M!}(\sum_{\ell=1}^{q} N^{-\delta}\lambda_\ell \frac{\partial}{\partial\theta_\ell})^M \underset{\sim}{\mu}_N(\theta_0)]'$$

$$\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_0)[\frac{1}{M!}(\sum_{\ell=1}^{q} N^{-\delta}\lambda_\ell \frac{\partial}{\partial\theta_\ell})^M \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_0)] + o(1)$$

But by (2.2.14) the 2nd term becomes Δ as $N \rightarrow \infty$ hence $\mathcal{L}(Q_N^*) \rightarrow \chi^2(t,\Delta)$ as $N \rightarrow \infty$.

To show the asymptotic equivalence of Q_N and Q_N^* we first note that Q_N^* is bounded in probability since it has a χ^2 distribution; to be more explicit for any $\varepsilon > 0$ there exists a $K(N,\varepsilon)$, depending on N and ε such that

$$P[Q_N^* \leq K(N,\varepsilon)] \geq 1 - \varepsilon \text{ for every } N > N_\varepsilon.$$

We denote this property by writing

$$Q_N^* = O_p(1) \text{ as } N \rightarrow \infty.$$

Consider $|Q_N - Q_N^*|$. There exists an N_0 such that for all $N > N_0$ we have $Q_N^* > 0$. So we may write for $N > N_0$

$$|Q_N - Q_N^*| = \left|\frac{Q_N}{Q_N^*} - 1\right| \; Q_N^*$$

Now by Courant's theorem on the ratio of two positive definite quadratic forms we know that

$$\inf_{(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_0)) \neq 0} \frac{Q_N}{Q_N^*} = \hat{\gamma}_{1,N} \quad \text{where}$$

$\hat{\gamma}_{1,N}$ = smallest root of $\hat{\underset{\sim}{\Sigma}}_N \, \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_N)$, and

$$\sup_{(\underset{\sim}{T}_N - \underset{\sim}{\mu}_N(\underset{\sim}{\theta}_0)) \neq 0} \frac{Q_N}{Q_N^*} = \hat{\gamma}_{t,N} \quad \text{where}$$

$\hat{\gamma}_{t,N}$ = largest root of $\hat{\underset{\sim}{\Sigma}}_N \, \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_N)$. So we obtain

$$|Q_N - Q_N^*| \leq \max \left[|\hat{\gamma}_{1,N} - 1| \; Q_N^*, \; |\hat{\gamma}_{t,N} - 1| \; Q_N^* \right]$$

but by (2.2.12a) we know $\hat{\gamma}_{1,N}$ and $\hat{\gamma}_{t,N}$ converge stochastically to one $|Q_N - Q_N^*| \leq o_p(1) \; O_p(1) = o_p(1)$ as $N \to \infty$ and we see that Q_N and Q_N^* are asymptotically equivalent and the theorem is proven. Q.E.D.

We observe that if we have 2 sequences of test statistics $\{Q_{N_1}^{(1)}\}$ and $\{Q_{N_2}^{(2)}\}$ and a common sequence of alternative hypotheses:

$$H_N: \; \underset{\sim}{\theta}_N = \underset{\sim}{\theta}_0 + N^{-\delta} \underset{\sim}{\lambda}$$

where $\{N_1\}$ and $\{N_2\}$ both depend on N and furthermore if there exist constants ρ_1 and ρ_2 defined by:

$$\rho_i = \lim_{N \to \infty} \frac{N_i}{N} \qquad \rho_i \in (0,1] \quad i = 1,2 \tag{2.2.20}$$

then Theorem 2.2.1 could be seen to yield the conclusion that

$$(2.2.21) \qquad \mathcal{L}\Big(N_i^{2M_i\delta}\big(\underset{\sim}{T}_N^{(i)} - \underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta}_0)\big)' \hat{\underset{\sim}{\Sigma}}_N^{(i)^{-1}}\big(\underset{\sim}{T}_N^{(i)} - \underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta}_0)\big)\Big) \to \chi^2(t,\rho_i^{2M_i\delta}\Delta).$$

This fact will be used in later sections to obtain some efficiency results.

The only value of M which we consider in the applications in the later chapters is $M = 1$; no practical situations are known to us in which $M > 1$. With regard to δ we note that for the multivariate location problems and the growth curve application that $\delta = 1/2$; for tests of independence one would require $\delta = 1/4$ in order to keep the limiting power bounded away from zero and one.

Since $M = 1$ for the cases we study we present in the form of a theorem, an observation concerning the noncentrality parameter when $t < q$.

If t, q, $\underset{\sim}{\lambda}$, $\underset{\sim}{c}$, M and $\underset{\sim}{\Sigma}(\underset{\sim}{\theta}_0)$ are as defined in Theorem 2.2.1 we have the following theorem:

Theorem 2.2.2. If $t < q$ and $M = 1$ then there exists at least one $\underset{\sim}{\lambda} \neq \underset{\sim}{0}$ such that $\underset{\sim}{c} = \underset{\sim}{0}$.

Proof:

If $M = 1$ then $\Delta = \underset{1\times q}{\underset{\sim}{\lambda}'}\ \underset{q\times t}{\underset{\sim}{D}'}\ \underset{t\times t}{\underset{\sim}{\Sigma}^{-1}}(\underset{\sim}{\theta}_0)\ \underset{t\times q}{\underset{\sim}{D}}\ \underset{q\times 1}{\underset{\sim}{\lambda}}$ where $\underset{\sim}{D}$ is defined by $\underset{\sim}{c} = \underset{\sim}{D}\,\underset{\sim}{\lambda}$. Since $t < q$ we know that rank $(\underset{\sim}{D}) \leq t < q$ and further rank $(\underset{\sim}{D}'\,\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_0)\,\underset{\sim}{D}) \leq t < q$. So the matrix $\underset{\sim}{D}'\,\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_0)\,\underset{\sim}{D}$ of size $q \times q$ has rank s where $s < q$.

By the symmetry of $\underset{\sim}{D}'\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_0)\,\underset{\sim}{D}$ there exists an orthogonal matrix $\underset{\sim}{P}$ such that

$$\underset{\sim}{P}'(\underset{\sim}{D}'\,\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_0)\underset{\sim}{D})\,\underset{\sim}{P} = \mathrm{diag}(r_1,r_2,\ldots,r_s,0,0,\ldots,0) = R$$

where $\underset{\sim}{R}$ has the roots of $(\underset{\sim}{D}'\,\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_0)\,\underset{\sim}{D})$ on the diagonal.

Define the full rank transformation

$$\underset{\sim}{P}'\underset{\sim}{\lambda} = \underset{\sim}{\lambda}^*$$

therefore $$\underset{\sim}{\lambda} = \underset{\sim}{P}\,\underset{\sim}{\lambda}^* \quad \text{and}$$

$$\underset{\sim}{\lambda}'\,\underset{\sim}{D}'\,\underset{\sim}{\Sigma}^{-1}\,(\underset{\sim}{\theta}_0)\,\underset{\sim}{D}\,\underset{\sim}{\lambda} = \underset{\sim}{\lambda}^{*'}\,\underset{\sim}{R}\,\underset{\sim}{\lambda}^* .$$

Let the first s elements of $\underset{\sim}{\lambda}^*$ be zero and at least one of the remaining (q-s) elements be nonzero, then

$$\underset{\sim}{\lambda}^{*'}\,\underset{\sim}{R}\,\underset{\sim}{\lambda}^* = 0$$

but $\underset{\sim}{P}\,\underset{\sim}{\lambda}^* = \underset{\sim}{\lambda}$ and $|\underset{\sim}{P}| \neq 0$ therefore $\underset{\sim}{\lambda} \neq \underset{\sim}{0}$ and the result is proven.

Q.E.D.

We see that if $t < q$ there will exist directions for which Δ will be zero and hence our test will have power α in those directions. If $t \geq q$ and if $\underset{\sim}{D}$ has rank q then it follows that $\underset{\sim}{D}'\,\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_0)\,\underset{\sim}{D}$ is positive definite and for any $\underset{\sim}{\lambda} \neq \underset{\sim}{0}$ we would have $\Delta > 0$.

We now study the asymptotic power function of our statistics under the conditions of Theorem 2.2.1.

2.3 The Asymptotic Power Function of $\{Q_N\}$

2.3.1 Taylor Series Expansion of the Power Function

Under the assumptions of Theorem 2.2.1 and the uniform convergence in $\underset{\sim}{\lambda}$ we may write the limiting power function $P(\underset{\sim}{\lambda})$ as:

$$(2.3.1)\quad P(\underset{\sim}{\lambda}) = \sum_{r=0}^{\infty} \frac{1}{r!}\Big(\sum_{\ell=1}^{q} \lambda_\ell \frac{\partial}{\partial\lambda_\ell}\Big)^{r} P(\underset{\sim}{\lambda})\Big|_{\underset{\sim}{\lambda}=0} \quad \forall\, \underset{\sim}{\lambda} \ni ||\underset{\sim}{\lambda}|| < K$$

If we are going to use a Taylor series representation of $P(\underset{\sim}{\lambda})$ we need to evaluate terms of the form $\left.\dfrac{\partial^{r} P(\underset{\sim}{\lambda})}{\partial\lambda_{\ell_1}\cdots\partial\lambda_{\ell_r}}\right|_{\underset{\sim}{\lambda} = \underset{\sim}{0}}$.

Since $P(\underset{\sim}{\lambda})$ is a function of Δ one could evaluate the partial derivatives by application of the chain rule of differential calculus. Proceeding with this idea we get the following:

a)
$$\frac{\partial P(\underset{\sim}{\lambda})}{\partial\lambda_{K_1}} = \frac{\partial P(\underset{\sim}{\lambda})}{\partial\Delta}\cdot\frac{\partial\Delta}{\partial\lambda_{K_1}} \qquad \text{for } K_1 = 1,2,\ldots,q$$

as the first partial derivatives, and

b)
$$\frac{\partial^2 P(\underset{\sim}{\lambda})}{\partial\lambda_{K_1}\partial\lambda_{K_2}} = \frac{\partial^2 P(\underset{\sim}{\lambda})}{\partial\Delta^2}\,\frac{\partial\Delta}{\partial\lambda_{K_1}}\,\frac{\partial\Delta}{\partial\lambda_{K_2}} + \frac{\partial P(\underset{\sim}{\lambda})}{\partial\Delta}\,\frac{\partial^2\Delta}{\partial\lambda_{K_1}\partial\lambda_{K_2}}$$

for $K_1, K_2 = 1,2,\ldots,q$ as the second partial derivatives,

c)
$$\frac{\partial^3 P(\underset{\sim}{\lambda})}{\partial\lambda_{K_1}\partial\lambda_{K_2}\partial\lambda_{K_3}} = \frac{\partial^3 P(\underset{\sim}{\lambda})}{\partial\Delta^3}\,\frac{\partial\Delta}{\partial\lambda_{K_1}}\,\frac{\partial\Delta}{\partial\lambda_{K_2}}\,\frac{\partial\Delta}{\partial\lambda_{K_3}} + \frac{\partial^2 P(\underset{\sim}{\lambda})}{\partial\Delta^2}\,\frac{\partial^2\Delta}{\partial\lambda_{K_1}\partial\lambda_{K_2}}\,\frac{\partial\Delta}{\partial\lambda_{K_3}}$$
$$+ \frac{\partial^2 P(\underset{\sim}{\lambda})}{\partial\Delta^2}\,\frac{\partial^2\Delta}{\partial\lambda_{K_1}\partial\lambda_{K_3}}\,\frac{\partial\lambda}{\partial\lambda_{K_2}} + \frac{\partial^2 P(\underset{\sim}{\lambda})}{\partial\Delta^2}\,\frac{\partial^2\Delta}{\partial\lambda_{K_2}\partial\lambda_{K_3}}\,\frac{\partial\Delta}{\partial\lambda_{K_1}}$$
$$+ \frac{\partial P(\underset{\sim}{\lambda})}{\partial\Delta}\,\frac{\partial^3\Delta}{\partial\lambda_{K_1}\partial\lambda_{K_2}\partial\lambda_{K_3}}$$

for K_1, K_2, $K_3 = 1,2,\ldots,q$ as the third partial derivatives. Continuing in this fashion will lead to complicated expressions to evaluate so we seek alternative methods of computing these derivatives. Since we see that each term in the series representation of $P(\underset{\sim}{\lambda})$ in (2.3.1) is a polynomial of degree r in the λ_i, we will derive some results for homogeneous polynomials in general and with these results obtain a reduction of (2.3.1).

2.3.2 Results on Homogeneous Polynomials

Definition: A function $h(z_1, \ldots, z_q)$ is said to be homogeneous of degree n in a region $E \subset R^q$ if and only if for every positive number β and $(z_1, \ldots, z_q)$ with both $(z_1, \ldots, z_q)$ and $(\beta z_1, \ldots, \beta z_q)$ in E we have $h(\beta z_1, \ldots, \beta z_q) = \beta^n h(z_1, \ldots, z_q)$.

We remark that if h is a polynomial in $(z_1, \ldots, z_q)$ then we would say h is a homogeneous polynomial of degree n if it satisfied the condition of the above definition. Denoting a homogeneous polynomial of degree K as $h_K = h_K(z_1, \ldots, z_q)$ we now state and prove some results concerning h_K.

Lemma 2.3.1 $\frac{\partial h_K}{\partial z_s}$ is a homogeneous polynomial of degree

$$K - 1 \ (K = 1,2,\ldots, \ ), \ s = 1, \ldots, q.$$

Proof:

Fix an arbitrary $s \in (1,2,\ldots,q)$. We can represent h_K as

$$h_K = \sum_{(i_1,\ldots,i_q) \in I_K} a_{i_1,\ldots,i_q} \prod_{j=1}^{q} z_j^{i_j}$$

where I_K is the set defined as:

$$I_K = \{(i_1,\ldots,i_q): \ i_j = 0,1,\ldots,K; \ j = 1,2,\ldots,q \ \sum_{j=1}^{q} i_j = K\}$$

and $a_{i_1,\ldots,i_q}$ are real numbers (could possibly be zero). Now h_K is a finite sum so there is no problem in exchanging the derivative and summation operators, so we get that

$$\frac{\partial h_K}{\partial z_s} = \sum_{(i_1,\ldots,i_q) \varepsilon I_{K-1}} i_s \, a_{i_1,\ldots,i_q} \prod_{\substack{j=1 \\ j\neq s}}^{q} z_j^{i_j} \, z_s^{i_s - 1} = h_K^*$$

$$I_{K-1} = \{(i_1,\ldots,i_q)\colon \ (i_1,\ldots,i_q) \ \varepsilon \ I_K \text{ and } i_s \geq 1\}$$

let $i_s' = i_s - 1$ then we see that

$$I_{K-1} = \{(i_1,\ldots,i_s')\colon \ i_j = 0,\ldots,K;\ j = 1,2,\ldots,q;\ j \neq s;$$
$$i_s' = 0,\ldots,K-1;\ \sum_{\substack{j=1 \\ j\neq s}}^{q} i_j + i_s = K-1\}$$

So we see that h_K^* is a polynomial of degree K-1 in addition note that

$$h_K^*(\beta z_1,\ldots,\beta z_q) = \beta^{K-1} \, h_K^*(z_1,\ldots,z_q)$$

so it is a homogeneous polynomial of degree K-1. Since s was arbitrarily the result holds for s = 1,2,...,q.

Q.E.D.

It is worth noting that if h is a homogeneous polynomial of degree 1 then each of its partial derivatives are themselves constants; hence the derivatives would be identically zero.

Lemma 2.3.2 Any t^{th} partial derivative of $h_K(z_1,\ldots,z_q)$, K = 1,2,... is:

i) a homogeneous polynomial of degree K-t for $t \leq K$

ii) zero for all $t > K$.

Proof:

i) follows immediately from repeated application of Lemma 2.3.1 at the K^{th} step the homogeneous polynomial of degree 0 would also be a constant so all further derivatives would be zero so ii) follows.

Q.E.D.

Lemma 2.3.3 For any two positive integers a and b we have:

i) $h_a h_b = h_{a+b}$,

ii) $h_a + g_a = h_a$,

iii) $ch_a = h_a$ where c is a nonzero real number.

Proof:

i) The product of 2 polynomials is a polynomial of degree equal to the sum of the degrees of the individual polynomials; the homogeneity follows since:

$$\begin{aligned} h_{a+b}(\beta z_1,\ldots,\beta z_q) &= h_a(\beta z_1,\ldots,\beta z_q)\ h_b(\beta z_1,\ldots,\beta z_q) \\ &= \beta^{a+b}\ h_{a+b} \end{aligned}$$

ii) The sum of 2 polynomials is a polynomial of degree less than or equal to the maximum degree of the individual polynomials, hence $h_a + g_a$ is a polynomial of degree $\leq a$ and

$$h_a(\beta z_1,\ldots,\beta z_q) + g_a(\beta z_1,\ldots,\beta z_q) = \beta^a h'_a = \beta^a h_a$$

iii) follows trivially.

Q.E.D.

Lemma 2.3.4 Let $f(h_K(z_1,\dots,z_q))$ be a function of h_K, a homogeneous polynomial of degree K, possessing K^{th} partial derivatives with respect to z and continuous in z, then for any $s = 1,2,\dots,K$ we have

$$\frac{\partial^s f}{\partial x_{i_s}\cdots\partial x_{i_1}} = \sum_{j=1}^{s} \frac{\partial^j f}{\partial h_K^j}\, h_{jK-s}^{(s,j)}$$

, where the superscripts on h_{jK-s} are just to indicate that these homogeneous polynomials are in general distinct for $i_1, \dots, i_s \in (1, 2, \dots, q)$.

Proof:

i) $$\frac{\partial f}{\partial x_{i_1}} = \frac{\partial f}{\partial h_K}\,\frac{\partial h_K}{\partial x_{i_1}} = \frac{\partial f}{\partial h_K}\, h_{K-1}^{(1,1)}$$ by Lemma 2.3.1 for $i_1 = 1,2,\dots,q$

ii) $$\frac{\partial^2 f}{\partial x_{i_2}\partial x_{i_1}} = \frac{\partial^2 f}{\partial h_K^2}\, h_{K-1}^{(1,1)}\,\frac{\partial h_K}{\partial x_{i_2}} + \frac{\partial f}{\partial h_K}\,\frac{\partial h_{K-1}^{(1,1)}}{\partial x_{i_2}}$$

$$= \frac{\partial^2 f}{\partial h_K^2}\, h_{2K-2}^{(2,2)} + \frac{\partial f}{\partial h_K}\, h_{K-2}^{(2,1)}$$ by Lemmas 2.3.2 and 2.3.3

$$= \sum_{j=1}^{2} \frac{\partial^j f}{\partial h_K^j}\, h_{jK-s}^{(2,j)}$$ for $i_1, i_2 = 1, 2, \dots, q$

Similarly we find

iii) $$\frac{\partial^3 f}{\partial x_{i_3}\partial x_{i_2}\partial x_{i_1}} = \frac{\partial^3 f}{\partial h_K^3}\, h_{3x-3} + \frac{\partial^2 f}{\partial h_K^2}\, h_{2x-3} + \frac{\partial^2 f}{\partial h_K^2}\, h_{2K-3}$$

$$+ \frac{\partial f}{\partial h_K}\, h_{K-3} = \sum_{j=1}^{3} \frac{\partial^j f}{\partial h_K^j}\, h_{jK-3}$$ by 2.3.2 and 2.3.3

for $i_1, i_2, i_3 = 1, 2, \ldots, q$ and in general we see the result follows with repeated application of Lemmas 2.3.2 and 2.3.3.

Q.E.D.

If we define $h_r = 0$ for all $r < 0$ then the same formula Lemma 2.3.4 will hold if f has more than K^{th} continuous partial derivatives. It should be noted that Lemma 2.3.4 gives us no idea as to the precise homogeneous polynomials on the right hand side but for the reduction of the power function this is no limitation.

Lemma 2.3.5 $h_K(0, \ldots, 0) = 0 \;\forall\; K \geq 1$.

Proof:

$$h_K = \sum_{I_K} a_{i,\ldots,i_K} \prod_{j=1}^{q} z_j^{i_j}$$

Now if $a_{i_1,\ldots,i_K} = 0 \;\forall (i_1,\ldots,i_K)$ then of course $h_K \equiv 0$. On the other hand if $a_{i_1,\ldots,i_K} \neq 0$ for at least one K-triple $(i_1,\ldots,i_K)$ then $i_j > 0$ for at least one j and hence the terms with nonzero $a_{i_1,\ldots,i_K}$ are multiplied by zeroes hence

$$h_K(0,\ldots,0) = 0$$

Q.E.D.

2.3.3 A Simplification of the Power Function

In this section we shall substantially simplify (2.3.1) in terms of the parameter M in Theorem 2.2.1.

Lemma 2.3.6 Under the assumptions of Theorem 2.2.1 the noncentrality parameter Δ is a homogeneous polynomial of degree 2M.

Proof:

$$\Delta = \frac{1}{(M!)^2} \underset{\sim}{c}'\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_0)\underset{\sim}{c} \quad \text{where}$$

$$\underset{\sim}{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_t \end{bmatrix} \quad \& \quad c_j = \sum_{I_M} \lambda_{\ell_1} \cdots \lambda_{\ell_M} d_{j\ \ell_1,\ldots,\ell_M} \qquad j = 1,\ldots,t$$

$$I_M = \{(\ell_1,\ldots,\ell_M):\ \ell_s = 0,1,\ldots,M;\ s = 1,\ldots,M;\ \sum_{s=1}^{M} \ell_s = M\}$$

and
$$d_{j:\ell_1,\ldots,\ell_M} = \lim_{N \to \infty} \frac{\partial^M \mu_{N,j}(\underset{\sim}{\theta}_0)}{\partial\theta_{\ell_1} \cdots \partial\theta_{\ell_M}}$$

clearly for each $j = 1, 2, \ldots, t$ we see that c_j is a homogeneous polynomial of degree m. Let $\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\theta}_0) = ((\sigma^{jj'}))$.

So $\Delta = \frac{1}{(M!)^2} \sum_{j=1}^{t} \sum_{j'=1}^{t} \sigma^{jj'} c_j c_{j'}$, by Lemma 2.3.3 we see that Δ is a homogeneous polynomial of degree 2M.

Q.E.D.

Lemma 2.3.7 Under the assumptions of Theorem 2.2.1 all of the first (2M-1) partial derivatives of the limiting power function with respect to the λ_i's are zero at $\underset{\sim}{\lambda} = \underset{\sim}{0}$. That is

$$\left. \frac{\partial^s P(\underset{\sim}{\lambda})}{\partial\lambda_{\ell_1} \cdots \partial\lambda_{\ell_s}} \right|_{\underset{\sim}{\lambda} = \underset{\sim}{0}} = 0; \ s = 1, 2, \ldots, 2M-1.$$

Proof:

$P(\underset{\sim}{\lambda}) = P(\Delta(\lambda_1,\ldots,\lambda_q))$ and by Lemma 2.3.6 Δ is a homogeneous polynomial of degree 2M, furthermore P is continuous in its partial derivatives of all order so by Lemma 2.3.4 we can represent the partial derivatives as:

$$\frac{\partial^s P(\underset{\sim}{\lambda})}{\partial\lambda_{\ell_1} \cdots \partial\lambda_{\ell_s}} = \sum_{j=1}^{s} \frac{\partial^j P(\underset{\sim}{\lambda})}{\partial\Delta^j} h^{(s,j)}_{2jM-s}, \ \forall\ \underset{\sim}{\lambda} \ni ||\underset{\sim}{\lambda}|| < K \text{ and } s = 1,2,\ldots,2M-1$$

Now for $2jM-s \geq 1$ we know by Lemma 2.3.5 that $h^{(s,j)}_{2jM-s}(0,\ldots,0) = 0$ but $j \geq 1$ and $s < 2M$ so $2Mj > s$ therefore we see that $2Mj - s \geq 1$ for all $s = 1, 2, \ldots, 2M-1$. So we obtain the required result.

Q.E.D.

The preceding theorem yields a substantial reduction of the power function. We see that (2.3.1) can now be written as:

$$(2.3.2) \quad P(\underset{\sim}{\lambda}) = \alpha + \sum_{r=2M}^{\infty} \frac{1}{r!}\left(\sum_{\ell=1}^{q} \lambda_\ell \frac{\partial}{\partial\lambda_\ell}\right)^r P(\underset{\sim}{\lambda}) \Bigg|_{\underset{\sim}{\lambda}=\underset{\sim}{0}}$$

where α is the significance level of the test. We now reduce the expansion by the following theorem.

Lemma 2.3.8 $\left.\dfrac{\partial^s P(\underset{\sim}{\lambda})}{\partial\lambda_{\ell_1}\ldots\partial\lambda_{\ell_s}}\right|_{\underset{\sim}{\lambda}=\underset{\sim}{0}} = 0$ if s is not an integer multiple of $2M$; i.e. if there does not exist an integer b such that $s = 2Mb$.

Proof:

By Lemma 2.3.4 and the comment after it we can take derivatives higher than the 2M and use the representation

$$\left.\frac{\partial^s P(\underset{\sim}{\lambda})}{\partial\lambda_{\ell_1}\ldots\partial\lambda_{\ell_s}}\right|_{\underset{\sim}{\lambda}=\underset{\sim}{0}} = \sum_{j=1}^{s} \frac{\partial^j P}{\partial\Delta^j} \; h_{2M_j-s}\Bigg|_{\underset{\sim}{\lambda}=\underset{\sim}{0}}$$

where it is understood that $h_K = 0$ if $K < 0$.

But $h_{2M_j-s}\Big|_{\underset{\sim}{\lambda}=\underset{\sim}{0}} = 0$ if $2M_j - r \neq 0$ by Lemma 2.3.5 so the result follows.

Q.E.D.

So we can write the power function in its simplified form as

$$(2.3.3)\quad P(\underset{\sim}{\lambda}) = \alpha + \sum_{I_{2M}} \frac{1}{(2M)!} \lambda_{\ell_1} \cdots \lambda_{\ell_{2M}} \left. \frac{\partial^{2M} P(\underset{\sim}{\lambda})}{\partial\lambda_{\ell_1} \cdots \partial\lambda_{\ell_{2M}}} \right|_{\underset{\sim}{\lambda}=\underset{\sim}{0}} + \sum_{r=2}^{\infty}$$

$$\sum_{I_{2Mr}} \frac{1}{(2Mr)!} \lambda_{\ell_1} \cdots \lambda_{\ell_{2Mr}} \left. \frac{\partial^{2Mr} P(\underset{\sim}{\lambda})}{\partial\lambda_{\ell_1} \cdots \partial\lambda_{\ell_{2Mr}}} \right|_{\underset{\sim}{\lambda}=\underset{\sim}{0}}$$

We now give an alternative representation for the power function and then proceed to use these simplifications in discussing the proposed efficiency criteria.

2.3.4 An Alternative Representation of the Power Function

If we denote the i^{th} derivative of $P(\underset{\sim}{\lambda})$ with respect to Δ by $P^{(i)}(\Delta)$ then we can write $P(\underset{\sim}{\lambda})$ in a power series about $\Delta = 0$ as:

$$P(\Delta) = \sum_{r=0}^{\infty} \frac{\Delta^r}{r!} P^{(r)}(0)$$

where

$$P^{(r)}(0) = P^{(r)}(\Delta) \Big|_{\Delta=0}$$

but Δ itself is a function of $\lambda_1, \ldots \lambda_q$ and can be written in a power series in $\underset{\sim}{\lambda}$ about $\underset{\sim}{\lambda} = 0$ as

$$\Delta(\underset{\sim}{\lambda}) = \sum_{s=0}^{\infty} \frac{1}{s!} \left(\sum_{\ell=1}^{q} \lambda_\ell \frac{\partial}{\partial\lambda_\ell} \right)^s \Delta \Bigg|_{\underset{\sim}{\lambda}=\underset{\sim}{0}}$$

So

$$P(\Delta) = \sum_{r=0}^{\infty} \frac{\left\{ \sum_{s=0}^{\infty} \frac{1}{s!} \left(\sum_{\ell=1}^{q} \lambda_\ell \frac{\partial}{\partial\lambda_\ell} \right)^s \Delta \Big|_{\underset{\sim}{\lambda}=\underset{\sim}{0}} \right\}^r}{r!} P^{(r)}(0)$$

Lemma 2.3.9

$$P(\Delta) = \sum_{r=0}^{\infty} \frac{\left\{\frac{1}{(2M)!}\left(\sum_{\ell=1}^{q} \lambda_\ell \frac{\partial}{\partial\lambda_\ell}\right)^{2M} \Delta \Big|_{\underset{\sim}{\lambda}=\underset{\sim}{0}}\right\}^r}{r!} P^{(r)}(0)$$

under the conditions of Theorem 2.2.1.

Proof:

Δ is a homogeneous polynomial of degree 2M therefore, all of its derivatives less than the 2Mth with respect to the λ's must vanish at $\underset{\sim}{\lambda} = \underset{\sim}{0}$. On the other hand all derivatives greater than the 2Mth are zero since the 2Mth derivative itself is a constant in $\underset{\sim}{\lambda}$ so the result follows.

Q.E.D.

We now give a result which provides a representation of $P^{(r)}(\Delta)$ in general.

Lemma 2.3.10

$$P^{(r)}(\Delta) = \left(\frac{1}{2}\right)^r \sum_{s=0}^{r} \binom{r}{s}(-1)^s \, P(t + 2(r - s),\Delta),$$

$$r = 1, 2, \ldots$$

where t is as defined in Theorem 2.2.1 and

$$P(t + 2j,\Delta) = \Pr[\chi^2(t + 2j,\Delta) \geq X_{t,\alpha}]$$

i.e. the probability that a noncentral chi-square random variable with noncentrality parameter Δ and t + 2j degrees of freedom is greater than or equal to the (1-α) 100% point of a central chi-square random variable with t degrees of freedom.

Proof:

We know that

$$P(t,\Delta) = e^{-\frac{\Delta}{2}} \sum_{j=0}^{\infty} \frac{(\frac{\Delta}{2})^j}{j!} p_j,\ p_j = P(t + 2_j, 0).$$

Now the infinite series converges absolutely so there is no problem in exchanging the derivative operator and the infinite summation operator so we get

$$P^{(1)}(\Delta) = \frac{\partial P}{\partial \Delta} = -\frac{1}{2} e^{-\frac{\Delta}{2}} \sum_{j=0}^{\infty} \frac{(\frac{\Delta}{2})^j}{j!} p_j + \frac{1}{2} e^{-\frac{\Delta}{2}} \sum_{j=1}^{\infty} \frac{(\frac{\Delta}{2})^{j-1}}{(j-1)!} p_j$$

Therefore

$$P^{(1)}(\Delta) = \frac{1}{2}[P(t+2,\Delta) - P(t,\Delta)] = (\frac{1}{2})^1 \sum_{s=0}^{1} \binom{1}{s} (-1)^s P(t+2(1-s),\Delta)$$

We shall prove the result by induction:

Assume the result of the lemma is true for r = n, i.e.

$$P^{(n)}(\Delta) = (\frac{1}{2})^n \sum_{s=0}^{n} \binom{n}{s} (-1)^s P(t+2(n-s),\Delta)$$

$$P^{(n+1)}(\Delta) = \frac{\partial P^{(n)}(\Delta)}{\partial \Delta} = (\frac{1}{2})^n \sum_{s=0}^{n} \binom{n}{s} (-1)^s \frac{\partial P(t+2(n-s),\Delta)}{\partial \Delta}$$

$$= (\frac{1}{2})^n \sum_{s=0}^{n} \binom{n}{s} (-1)^s \{\frac{1}{2}(P(t+2(n-s)+2,\Delta) - P(t+2(n-s),\Delta))\}$$

$$= (\frac{1}{2})^{n+1} \{ \binom{n}{0} P(t+2n+2,\Delta) - \binom{n}{0} P(t+2n,\Delta) - \binom{n}{1} P(t+2n,\Delta)$$

$$+ \binom{n}{1} P(t+2n-2,\Delta) + \binom{n}{2} P(t+2n-2,\Delta) - \binom{n}{2} P(t+2n-4,\Delta)$$

$$- \binom{n}{3} P(t+2n-4,\Delta) + \dots + \binom{n}{n} (-1)^n P(t+2,\Delta) + \binom{n}{n} (-1)^{n+1} P(t,\Delta)\}$$

$$= \left(\tfrac{1}{2}\right)^{n+1} \binom{n}{0} P(t+2(n+1),\Delta) + \left(\tfrac{1}{2}\right)^{n+1} \sum_{s=0}^{n-1} (-1)^{s+1} \left[\binom{n}{s} + \binom{n}{s+1}\right]$$

$$P(t+2(n-s),\Delta) + \left(\tfrac{1}{2}\right)^{n+1} (-1)^{n+1} P(t+2(n-n),\Delta)$$

but $$\binom{n}{s} + \binom{n}{s+1} = \frac{[s+1+n-s]\ n!}{(n-s)!\ (s+1)!} = \binom{n+1}{s+1}$$

therefore we have

$$= \left(\tfrac{1}{2}\right)^{n+1} \binom{n+1}{0} P(t+2(n+1-0),\Delta) + \left(\tfrac{1}{2}\right)^{n+1} \sum_{s=0}^{n-1} (-1)^{s+1} \binom{n+1}{s+1}$$

$$P(t+2(n-s),\Delta) + \left(\tfrac{1}{2}\right)^{n+1} (-1)^{n+1} P(t+2(n+1-(n+1),\Delta)$$

let s' = s + 1

$$= \left(\tfrac{1}{2}\right)^{n+1} \binom{n+1}{0} P(t+2(n+1-0),\Delta) + \left(\tfrac{1}{2}\right)^{n+1} \sum_{s'=1}^{n} (-1)^{s'} \binom{n+1}{s'}$$

$$P(t+2(n+1-s'),\Delta) + \left(\tfrac{1}{2}\right)^{n+1} P(t+s(n+1)-(n+1),\Delta)$$

$$= \left(\tfrac{1}{2}\right)^{n+1} \sum_{s=0}^{n+1} \binom{n+1}{s} (-1)^{s} P(t+2(n+1-s),\Delta).$$

So by the principle of mathematical induction the result holds for all r = 1, 2, ...

Q.E.D.

In the next section we compute the Pitman ARE of test 2 to test 1 when $t_1 = t_2$; in the following sections we make use of the reductions we have obtained in this section.

2.4 Pitman ARE when $t_1 = t_2$

We consider now the problem of comparing the performance of $\{Q_N^{(2)}\}$ to $\{Q_N^{(1)}\}$ for large samples. We assume that both sequences of test statistics satisfy the conditions of Theorem 2.2.1 with parameters t_i, q, M_i, δ, $\underset{\sim}{\mu}_N^{(i)}$, Δ, $\underset{\sim}{\Sigma}^{(i)} = \underset{\sim}{\Sigma}^{(i)}(\underset{\sim}{\theta}_0)$ and $\underset{\sim}{c}^{(i)}$ for $i = 1, 2$. We denote the limiting power functions by $P_i(\underset{\sim}{\lambda})$ for $i = 1, 2$. We first notice that if $M_1 \neq M_2$ then one sequence of test statistics is behaving differently from the other. For example if $t_i \geq q$ and $M_1 = 1$, $M_2 > 1$ then test 1 has a limiting power function which attains a relative minimum at $\underset{\sim}{\theta}_0$ while test 2 is changing so rapidly at $\underset{\sim}{\theta}_0$ that we cannot guarantee that it has an extremum at $\underset{\sim}{\theta}_0$. For this reason we restrict consideration to $M_1 = M_2 = M$ in all the sections to follow.

Since it involves a slight extension of existing work (since q, t_1 and t_2 are general) we now derive the Pitman ARE of ϕ_2 with respect to ϕ_1 in the situation $t_1 = t_2$. For convenience we denote the Pitman ARE of ϕ_2 with respect to ϕ_1 as $e_{2,1}^P$.

Theorem 2.4.1. If $\{Q_N^{(1)}\}$ and $\{Q_N^{(2)}\}$ satisfy the conditions of Theorem 2.2.1 with parameters t_i and Δ_i respectively and if $t_1 = t_2$ then

$$e_{2,1}^P = \{ \frac{\Delta_2}{\Delta_1} \}^{1/2M\delta} .$$

Proof:

Recall that we require the $\lim_{N \to \infty} \frac{N_1(N)}{N_2(N)}$ where $\{N_1(N)\}$ and $\{N_2(N)\}$ are chosen such that the two tests $\{Q_{N_1}^{(1)}\}$ and $\{Q_{N_2}^{(2)}\}$ have the same limiting power against the same H_N,

$$H_N: \ \underset{\sim}{\theta}_N = \underset{\sim}{\theta}_0 + N^{-\delta} \underset{\sim}{\lambda}$$

Since $t_1 = t_2$ we know that the limiting powers are the same when the noncentralities are the same so if ρ_i are defined in (2.2.20) we see by (2.2.21) that

$$\mathcal{L}(Q^{(1)}_{N_1} \big| H_N) \rightarrow X^2(t, \rho_1^{2M\delta} \Delta_1) \text{ as } N \rightarrow \infty$$

and

$$\mathcal{L}(Q^{(2)}_{N_2} \big| H_N) \rightarrow X^2(t, \rho_2^{2M\delta} \Delta_2) \text{ as } N \rightarrow \infty$$

therefore we require

$$\rho_1^{2M\delta} \Delta_1 = \rho_2^{2M\delta} \Delta_2$$

$$\frac{\rho_1}{\rho_2} = \left\{ \frac{\Delta_2}{\Delta_1} \right\}^{1/2M\delta}$$

hence,

$$\lim_{N \rightarrow \infty} \frac{N_{1/N}}{N_{2/N}} = \lim_{N_1 \rightarrow \infty} \frac{N_1}{N_2} = \left\{ \frac{\Delta_2}{\Delta_1} \right\}^{1/2M\delta}$$

Q.E.D.

We remark that if exactly one Δ_i is zero the other test is clearly superior; if both Δ_1 and Δ_2 are zero then both are useless for that particular direction $\underset{\sim}{\lambda}$.

A common set of values for the parameters M, δ are 1, 1/2 respectively, and then

$$e^P_{2,1} = \frac{\underset{\sim}{\lambda}' \, \underset{\sim}{D}^{(2)'} \underset{\sim}{\Sigma}^{(2)^{-1}} \underset{\sim}{D}^{(2)} \underset{\sim}{\lambda}}{\underset{\sim}{\lambda}' \, \underset{\sim}{D}^{(1)'} \underset{\sim}{\Sigma}^{(1)^{-1}} \underset{\sim}{D}^{(1)} \underset{\sim}{\lambda}}$$

where $\underset{\sim}{D}^{(i)}$ is defined by $\underset{\sim}{C}^{(i)} = \underset{\sim}{D}^{(i)} \underset{\sim}{\lambda}$. In this form we have a general representation of the Pitman ARE for q and t arbitrary. If q > t then $e^P_{2,1}$ will be indeterminate for some $\underset{\sim}{\lambda}$ and care must be taken in the interpretation of $e^P_{2,1}$. If $q \leq t$ then one can apply Courant's theorem and place bounds on $e^P_{2,1}$ for all $\underset{\sim}{\lambda} \neq \underset{\sim}{0}$. If q = 1 no problem arises in the evaluation of $e^P_{2,1}$ since it would not depend on $\underset{\sim}{\lambda}$ in this case.

2.5 Local Asymptotic Relative Efficiency (LARE)

In general $t_1 \neq t_2$ then the preceding section's approach of computing the Pitman ARE does not work since for $t_1 \neq t_2$ the powers are not the same when $\Delta_1 = \Delta_2$. In this section we propose an alternative criterion of comparison in which we compare the limiting powers locally, i.e. when $\underset{\sim}{\lambda}$ is in some arbitrarily small neighborhood of the origin. To be precise we adopt the following:

Definition: The local asymptotic relative efficiency (LARE) of ϕ_2 with respect to ϕ_1 is defined to be the $\lim_{N \to \infty} \frac{N_1}{N_2}$ where $\{N_1\}$ and $\{N_2\}$ are chosen such that the two tests have the same limiting power locally through the same sequence of alternative hypotheses. Local power is the power function expansion up to the $(2M)^{th}$ derivative terms.

Theorem 2.5.1. The LARE of ϕ_2 with respect to ϕ_1 when $\{Q_N^{(1)}\}$ and $\{Q_N^{(2)}\}$ satisfy the conditions of Theorem 2.2.1 each with $\chi^2(t_i, \Delta_i)$ (i = 1,2) as $N \to \infty$ through H_N where

$$H_N: \quad \underset{\sim}{\theta}_N = \underset{\sim}{\theta}_0 + N^{-\delta} \underset{\sim}{\lambda}$$

is

$$\text{LARE} = \left\{ \left(\frac{P_2(t_2+2,0) - \alpha}{P_1(t_1+2,0) - \alpha} \right) \frac{\Delta_2}{\Delta_1} \right\}^{1/2M\delta} .$$

Proof: From section 2.3.4 we know that each power function, $P_i(\underset{\sim}{\lambda})$ can be written as

$$P_i(\underset{\sim}{\lambda}) = \alpha + \sum_{r=1}^{\infty} \frac{\Delta_i^{\,r}}{r!} P_i^{(r)}(\Delta_i)\Big|_{\Delta_i=0}$$

$$= \alpha + \Delta_i P_i^{(1)}(0) + \sum_{r=2}^{\infty} \frac{\Delta_i^r}{r!} P_i^{(r)}(\Delta_i)\Big|_{\Delta_i=0}$$

but $\sum_{r=2}^{\infty} \frac{\Delta_i^r}{r!} P_i^{(r)}(\Delta_i)\Big|_{\Delta_i=0}$ is the tail of a convergent series so we know that it is bounded by a finite number K.

In particular if $\underset{\sim}{\lambda}$ is sufficiently small then

$$\sum_{r=2}^{\infty} \frac{\Delta_i^r}{r!} P_i^{(r)}(\Delta_i)\Big|_{\Delta_i=0} = o(\lambda_{\max}^{4M}) \text{ as } \lambda_{\max} \to 0$$

where $\lambda_{\max} = \max_{\ell=1,2,\ldots,q} (\lambda_1,\ldots,\lambda_q)$.

So for $||\underset{\sim}{\lambda}|| < \varepsilon$, a small number, we see that

$$P_i(\underset{\sim}{\lambda}) \doteq \alpha + \Delta_i\, P_i^{(1)}(\Delta_i)\Big|_{\Delta_i=0}$$

So if we define local power as $\alpha + \Delta_i\, P_i^{(1)}(\Delta_i)\Big|_{\Delta_i=0}$

then we require the equality of terms like $\Delta_1\, P_1^{(1)}(\Delta_1)\Big|_{\Delta_1=0}$ and $\Delta_2\, P_2^{(1)}(\Delta_2)\Big|_{\Delta_2=0}$ in the limit. Now choose $\{N_1(N)\}$ and $\{N_2(N)\}$ such that

$$\rho_1^{2M\delta}\Delta_1\, P_1^{(1)}(\rho_1^{2M\delta}\Delta_1)\Big|_{\rho_1^{2M\delta}\Delta_1=0} = \rho_2^{2M\delta}\Delta_2\, P_2^{(1)}(\rho_2^{2M\delta}\Delta_2)\Big|_{\rho_2^{2M\delta}\Delta_2=0}$$

but by Lemma 2.3.10 we know

$$P_i^{(1)}(0) = (\tfrac{1}{2})[P_i(t_i + 2,0) - \alpha]$$

and we see

$$\frac{\rho_1}{\rho_2} = \left\{ \frac{[P_2(t_2 + 2,0) - \alpha]}{[P_1(t_1 + 2,0) - \alpha]} \; \frac{\Delta_2}{\Delta_1} \right\}^{1/2M\delta}$$

and hence

$$\frac{\rho_1}{\rho_2} = \lim_{N \to \infty} \frac{N_1(N)}{N_2(N)} = \lim_{N_1 \to \infty} \frac{N_1}{N_2} = \left\{ \frac{[P_2(t_2 + 2,0) - \alpha]}{[P_1(t_1 + 2,0) - \alpha]} \; \frac{\Delta_2}{\Delta_1} \right\}^{1/2M\delta}$$

Q.E.D.

Let us define

$$R(t_1,t_2,\alpha) = \frac{P_2(t_2 + 2,0) - \alpha}{P_1(t_1 + 2,0) - \alpha} \quad .$$

In the case where $M = 1$, $\delta = 1/2$ we get

$$\text{(2.5.1)} \qquad \text{LARE} = R(t_1,t_2,\alpha) \; \frac{\underset{\sim}{\lambda}'\underset{\sim}{D}^{(2)'}\underset{\sim}{\Sigma}^{(2)^{-1}}\underset{\sim}{D}^{(2)}\underset{\sim}{\lambda}}{\underset{\sim}{\lambda}'\underset{\sim}{D}^{(1)'}\underset{\sim}{\Sigma}^{(1)^{-1}}\underset{\sim}{D}^{(1)}\underset{\sim}{\lambda}} \quad .$$

It can be seen that we now have a measure of ARE which takes into account (a) the difference in degrees of freedom and (b) the difference in the noncentrality parameters for $\underset{\sim}{\lambda}$ near the origin. Furthermore, these two components are factored so that their product is the efficiency. The second quantity is the same as we get when we compute the Pitman ARE for $t_1 = t_2$, and we could proceed to place bounds on this ratio by Courant's theorem. The comments that were made in the previous section on the relationship of q and t_i should be kept in mind since either of these quadratic forms in (2.5.1) may be positive semidefinite. The scalar factor $R(t_1,t_2,\alpha)$ is given in Tables 2.5.1 to 2.5.5 for $t_1,t_2 = 1(1)10$ and $\alpha = .10, .05, .01, .005$ and $.001$. We note that the scalar adjustment can be quite large for large values of $|t_2-t_1|$, but this factor varies little with the value of α unless $|t_2-t_1|$ is large.

One major drawback of this LARE is that it depends on the arbitrary vector $\underset{\sim}{\lambda}$, even though we have assumed the vector to have arbitrarily small elements. This is an undesirable feature of this measure of ARE. We now propose a measure of ARE, based on the generalized curvature of the power function at the null point, which avoids this drawback.

TABLE 2.5.1

VALUES OF THE ADJUSTMENT FACTOR, $R(t_1,t_2,\alpha)$, FOR $\alpha = .10$

t_2	t_1 = 1	2	3	4	5	6	7	8	9	10
1	1.00	1.47	1.86	2.19	2.49	2.77	3.02	3.26	3.48	3.69
2	.68	1.00	1.26	1.49	1.69	1.88	2.05	2.21	2.36	2.51
3	.54	.79	1.00	1.18	1.34	1.49	1.62	1.75	1.87	1.99
4	.46	.67	.85	1.00	1.14	1.26	1.38	1.49	1.59	1.68
5	.40	.59	.75	.88	1.00	1.11	1.21	1.31	1.40	1.48
6	.36	.53	.67	.79	.90	1.00	1.09	1.18	1.26	1.34
7	.33	.49	.62	.73	.83	.92	1.00	1.08	1.15	1.22
8	.31	.45	.57	.67	.77	.85	.93	1.00	1.07	1.13
9	.29	.42	.53	.63	.72	.79	.87	.94	1.00	1.06
10	.27	.40	.50	.59	.67	.75	.82	.88	.94	1.00

Note: The entries in the table are

$$R(t_1,t_2,\alpha) = \frac{\Pr[\chi^2_{t_2+2} \geq X_{t_2,\alpha}] - \alpha}{\Pr[\chi^2_{t_1+2} \geq X_{t_1,\alpha}] - \alpha}.$$

TABLE 2.5.2

VALUES OF THE ADJUSTMENT FACTOR, $R(t_1,t_2,\alpha)$ FOR $\alpha = .05$

t_2	t_1									
	1	2	3	4	5	6	7	8	9	10
1	1.00	1.53	1.96	2.34	2.68	2.99	3.27	3.54	3.80	4.04
2	.65	1.00	1.28	1.53	1.75	1.95	2.14	2.32	2.48	2.64
3	.51	.78	1.00	1.19	1.36	1.52	1.67	1.81	1.94	2.06
4	.43	.65	.84	1.00	1.14	1.28	1.40	1.52	1.62	1.73
5	.37	.57	.73	.87	1.00	1.12	1.22	1.32	1.42	1.51
6	.33	.51	.66	.78	.90	1.00	1.10	1.19	1.27	1.35
7	.31	.47	.60	.71	.82	.91	1.00	1.08	1.16	1.23
8	.28	.43	.55	.66	.76	.84	.92	1.00	1.07	1.14
9	.26	.40	.52	.62	.70	.79	.86	.93	1.00	1.06
10	.25	.38	.49	.58	.66	.74	.81	.88	.94	1.00

Note: See note for Table 2.5.1.

TABLE 2.5.3

VALUES OF THE ADJUSTMENT FACTOR, $R(t_1,t_2,\alpha)$ FOR $\alpha = .01$

t_2	t_1									
	1	2	3	4	5	6	7	8	9	10
1	1.00	1.62	2.13	2.58	2.99	3.37	3.72	4.05	4.36	4.66
2	.62	1.00	1.32	1.60	1.85	2.08	2.30	2.50	2.69	2.88
3	.47	.76	1.00	1.21	1.40	1.58	1.74	1.90	2.05	2.18
4	.39	.63	.83	1.00	1.16	1.30	1.44	1.57	1.69	1.80
5	.33	.54	.71	.86	1.00	1.13	1.24	1.35	1.46	1.56
6	.30	.48	.63	.77	.89	1.00	1.10	1.20	1.29	1.38
7	.27	.44	.57	.69	.80	.91	1.00	1.09	1.17	1.25
8	.25	.40	.53	.64	.74	.83	.92	1.00	1.08	1.15
9	.23	.37	.49	.59	.69	.77	.85	.93	1.00	1.07
10	.21	.35	.46	.55	.64	.72	.80	.87	.94	1.00

Note: See note for Table 2.5.1.

TABLE 2.5.4

VALUES OF THE ADJUSTMENT FACTOR, $R(t_1,t_2,\alpha)$ FOR $\alpha = .005$

t_2	t_1									
	1	2	3	4	5	6	7	8	9	10
1	1.00	1.64	2.18	2.66	3.09	3.49	3.86	4.22	4.55	4.87
2	.61	1.00	1.33	1.62	1.88	2.12	2.35	2.56	2.77	2.96
3	.46	.75	1.00	1.22	1.42	1.60	1.77	1.93	2.08	2.23
4	.38	.62	.82	1.00	1.16	1.31	1.45	1.58	1.71	1.83
5	.32	.53	.71	.86	1.00	1.13	1.25	1.36	1.47	1.57
6	.29	.47	.63	.76	.89	1.00	1.11	1.21	1.30	1.39
7	.26	.43	.57	.69	.80	.90	1.00	1.09	1.18	1.26
8	.24	.39	.52	.63	.73	.83	.92	1.00	1.08	1.15
9	.22	.36	.48	.59	.68	.77	.85	.93	1.00	1.07
10	.21	.34	.45	.55	.64	.72	.79	.87	.93	1.00

Note: See note for Table 2.5.1.

TABLE 2.5.5

VALUES OF THE ADJUSTMENT FACTOR, $R(t_1,t_2,\alpha)$ FOR $\alpha = .001$

t_2	t_1									
	1	2	3	4	5	6	7	8	9	10
1	1.00	1.69	2.28	2.81	3.29	3.73	4.15	4.53	4.91	5.27
2	.59	1.00	1.35	1.66	1.94	2.20	2.45	2.68	2.90	3.11
3	.44	.74	1.00	1.23	1.44	1.63	1.82	1.99	2.15	2.31
4	.36	.60	.81	1.00	1.17	1.33	1.48	1.62	1.75	1.88
5	.30	.52	.69	.85	1.00	1.14	1.26	1.38	1.49	1.60
6	.27	.45	.61	.75	.88	1.00	1.11	1.22	1.32	1.41
7	.24	.41	.55	.68	.79	.90	1.00	1.10	1.18	1.27
8	.22	.37	.50	.62	.72	.82	.91	1.00	1.08	1.16
9	.20	.34	.46	.57	.67	.76	.84	.92	1.00	1.07
10	.19	.32	.43	.53	.62	.71	.79	.86	.93	1.00

Note: See note for Table 2.5.1.

2.6 Curvature Asymptotic Relative Efficiency (CARE)

We consider now another new measure of ARE for the special case $M_1 = M_2 = 1$. At the end of this section we discuss briefly the case where $M_i > 1$ and point out problems in the interpretation of this case.

We recall from Chapter I the definition of generalized Gaussian curvature of a function in several variables. We propose now as a measure of efficiency a criterion based on the generalized Gaussian curvature of the two limiting power functions, $P_i(\underset{\sim}{\lambda})$ at the origin.

Definition: The curvature asymptotic relative efficiency (CARE) of test ϕ_2 with respect to ϕ_1 is the $\lim\limits_{N_1 \to \infty} \frac{N_1}{N_2}$ where $\{N_1\}$ and $\{N_2\}$ are chosen such that the two tests have limiting power functions with the same generalized Gaussian curvature at $\underset{\sim}{\lambda} = \underset{\sim}{0}$ through the same sequence of alternative hypotheses.

Of course this definition of efficiency, denoted by (CARE), is not meaningful if M_1 and M_2 are both greater than one since one can then show by application of Lemma 2.3.7 that the two generalized Gaussian curvatures must be zero at $\underset{\sim}{\lambda} = \underset{\sim}{0}$. Let us now derive the generalized curvatures (we omit the adjective Gaussian for convenience) of the limiting power functions at $\underset{\sim}{\lambda} = \underset{\sim}{0}$.

Theorem 2.6.1. Under the conditions of Theorem 2.2.1, when $M_i = 1$ the generalized curvature of the limiting power function of $\{Q_N^{(i)}\}$ at $\underset{\sim}{\lambda} = \underset{\sim}{0}$ is:

$$G_i = \frac{[P_i(t_i + 2,0) - \alpha]^q}{2^q} \left| \underset{\sim}{D}^{(i)\prime} \underset{\sim}{\Sigma}^{(i)^{-1}} \underset{\sim}{D}^{(i)} \right|$$

where $\underset{\sim}{D}^{(i)}$ is defined by $\underset{\sim}{C}^{(i)} = \underset{\sim}{D}^{(i)} \underset{\sim}{\lambda}$.

<u>Proof</u>: By definition of G_i in Chapter I we have

$$G_i = \left| \left(\left(\frac{\partial^2 P_i(\underset{\sim}{\lambda})}{\partial\lambda_\ell \partial\lambda_K} \Big|_{\underset{\sim}{\lambda}=\underset{\sim}{0}} \right) \right) \right| \Bigg/ \left(1 + \sum_{\ell=1}^{q} \left[\frac{\partial P_i(\underset{\sim}{\lambda})}{\partial\lambda_\ell} \Big|_{\underset{\sim}{\lambda}=\underset{\sim}{0}} \right]^2 \right)^2$$

by Lemma 2.3.7 the denominator is one so

$$G_i = \left| \left(\left(\left[\frac{\partial^2 P_i(\underset{\sim}{\lambda})}{\partial\lambda_\ell \partial\lambda_K} \Big|_{\underset{\sim}{\lambda}=\underset{\sim}{0}} \right] \right) \right) \right|$$

and we know that

$$\frac{\partial^2 P_i(\underset{\sim}{\lambda})}{\partial\lambda_\ell \partial\lambda_K} \Big|_{\underset{\sim}{\lambda}=\underset{\sim}{0}} = \left[\frac{\partial^2 P_i(\underset{\sim}{\lambda})}{\partial\Delta_i^2} \frac{\partial\Delta_i}{\partial\Delta_\ell} \frac{\partial\Delta_i}{\partial\Delta_K} + \frac{\partial P_i(\underset{\sim}{\lambda})}{\partial\Delta_i} \frac{\partial^2\Delta_i}{\partial\lambda_\ell \partial\lambda_K} \right] \Bigg|_{\underset{\sim}{\lambda}=\underset{\sim}{0}}$$

But by Lemmas 2.3.9 and 2.3.10 we get

$$\frac{\partial^2 P_i(\underset{\sim}{\lambda})}{\partial\lambda_\ell \partial\lambda_K} \Big|_{\underset{\sim}{\lambda}=\underset{\sim}{0}} = \frac{[P_i(t_i + 2, 0) - \alpha]}{2} \left(\underset{\sim}{D}^{(i)'} \underset{\sim}{\Sigma}^{(i)^{-1}} \underset{\sim}{D}^{(i)} \right)_{\ell K}$$

where the second term on the right is the ℓK element of the matrix $\underset{\sim}{D}^{(i)'} \underset{\sim}{\Sigma}^{(i)^{-1}} \underset{\sim}{D}^{(i)}$. Therefore we get

$$G_i = \left(\frac{[P_i(t_i + 2, 0) - \alpha]}{2} \right)^q \left| \underset{\sim}{D}^{(i)'} \underset{\sim}{\Sigma}^{(i)^{-1}} \underset{\sim}{D}^{(i)} \right|$$

Q.E.D.

Now we observe that if $q > t_i$ with $M_i = 1$ then $G_i = 0$ by Theorem 2.2.2; on the other hand if $q \leq t$, then $G_i > 0$ if we assume that $\underset{\sim}{D}^{(i)}$ is of rank q which we do throughout. We are now in a position to find the CARE.

Theorem 2.6.2. If $\{Q_N^{(1)}\}$ and $\{Q_N^{(2)}\}$ satisfy the conditions of Theorem 2.2.1 and if $q \leq t_i (i = 1,2)$ with $M_1 = M_2 = 1$, then

$$\text{CARE} = R(t_1,t_2,\alpha)^{1/2\delta} \left\{ \frac{|\underset{\sim}{D}^{(2)'}\underset{\sim}{\Sigma}^{(2)^{-1}}\underset{\sim}{D}^{(2)}|}{|\underset{\sim}{D}^{(1)'}\underset{\sim}{\Sigma}^{(1)^{-1}}\underset{\sim}{D}^{(1)}|} \right\}^{1/2q\delta}$$

Proof: We must determine $\{N_1\}$ and $\{N_2\}$ such that $G_1 = G_2$ where G_i is the generalized curvature of the limiting power function of $\{Q_{N_i}^{(i)}\}$ through H_N as $N \to \infty$ where

$$H_N: \ \underset{\sim}{\theta}_N = \underset{\sim}{\theta}_0 + N^{-\delta} \underset{\sim}{\lambda}$$

By (2.2.21) we have

$$\mathcal{L}(Q_{N_i}^{(i)} \mid H_N) \to \chi^2(t_i, \rho_i^{2\delta} \Delta_i) \text{ as } N \to \infty$$

and ρ_i is defined in (2.2.20) and since M = 1

$$\Delta_i = \underset{\sim}{\lambda}' \underset{\sim}{D}^{(i)'} \underset{\sim}{\Sigma}^{(i)^{-1}} \underset{\sim}{D}^{(i)} \underset{\sim}{\lambda}.$$

By Theorem 2.6.1

$$G_i = \frac{[P_i(t_i+2,0) - \alpha]^q}{2^q} \rho_i^{2\delta q} \left| \underset{\sim}{D}^{(i)'} \underset{\sim}{\Sigma}^{(i)^{-1}} \underset{\sim}{D}^{(i)} \right|$$

Setting $G_1 = G_2$ we obtain

$$\frac{\rho_1}{\rho_2} = \lim_{N \to \infty} \frac{N_1(N)}{N_2(N)} = \lim_{N_1 \to \infty} \frac{N_1}{N_2} = R(t_1,t_2,\alpha)^{1/2\delta} \left\{ \frac{|\underset{\sim}{D}^{(2)'}\underset{\sim}{\Sigma}^{(2)^{-1}}\underset{\sim}{D}^{(2)}|}{|\underset{\sim}{D}^{(1)'}\underset{\sim}{\Sigma}^{(1)^{-1}}\underset{\sim}{D}^{(1)}|} \right\}^{1/2\delta q}$$

and the proof is complete. Q.E.D.

We now give a result for the special case of $t_1 = t_2$.

Theorem 2.6.3. Under the conditions of Theorem 2.6.2 with $t_1 = t_2$ the CARE is the geometric mean of the roots of the product matrix, $(\underset{\sim}{D}^{(2)'}\underset{\sim}{\Sigma}^{(2)^{-1}}\underset{\sim}{D}^{(2)})(\underset{\sim}{D}^{(1)'}\underset{\sim}{\Sigma}^{(1)^{-1}}\underset{\sim}{D}^{(1)})^{-1}$, of the Pitman ARE.

Proof: From section 2.5 we know that when $q \leq t_1 = t_2$, $M_1 = M_2 = 1$ we have

$$e^P_{2,1} = \left[\frac{\underset{\sim}{\lambda}'D^{(2)'}\Sigma^{(2)^{-1}}D^{(2)}\underset{\sim}{\lambda}}{\underset{\sim}{\lambda}'D^{(1)'}\Sigma^{(1)^{-1}}D^{(1)}\underset{\sim}{\lambda}}\right]^{1/2\delta} .$$

Since both $D^{(2)'}\Sigma^{(2)^{-1}}D^{(2)}$ and $D^{(1)'}\Sigma^{(1)^{-1}}D^{(1)}$ are positive definite we can find a nonsingular F such that $F'D^{(2)'}\Sigma^{(2)^{-1}}D^{(2)}F =$ diag $(r_1,\ldots,r_q)$, $F'D^{(1)'}\Sigma^{(1)^{-1}}D^{(1)}F = I$ where $r_1,\ldots,r_q$ are the roots of $(\underset{\sim}{D}^{(2)'}\Sigma^{(2)^{-1}}\underset{\sim}{D}^{(2)})(\underset{\sim}{D}^{(1)'}\Sigma^{(1)^{-1}}\underset{\sim}{D}^{(1)})^{-1}$. If we let $\lambda^* = F\underset{\sim}{\lambda}$ then we can represent $e^P_{2,1}$ in $\underset{\sim}{\lambda}^*$ space as

$$e^{*P}_{2,1} = \frac{\underset{\sim}{\lambda}^{*'}F'D^{(2)'}\Sigma^{(2)^{-1}}D^{(2)}F\,\underset{\sim}{\lambda}^*}{\underset{\sim}{\lambda}^{*'}F'D^{(1)'}\Sigma^{(1)^{-1}}D^{(1)}F\,\underset{\sim}{\lambda}^*} = \sum_{\ell=1}^{q} r_\ell\lambda_\ell^{*2} \Big/ \sum_{\ell=1}^{q} \lambda_\ell^{*2}$$

Now the geometric mean of $r_1,\ldots,r_q$ is given by

$$\left(\prod_{\ell=1}^{q} r_\ell\right)^{1/q} = \{|F'D^{(2)'}\Sigma^{(2)^{-1}}D^{(2)}F|\,|F'D^{(1)'}\Sigma^{(1)^{-1}}D^{(1)}F|^{-1}\}^{1/q}$$

$$= \{|D^{(2)'}\Sigma^{(2)^{-1}}D^{(2)}| \,/\, |D^{(1)'}\Sigma^{(1)^{-1}}D^{(1)}|\}^{1/q} .$$

So the CARE is the geometric mean of the roots of the product matrix. We also note that the CARE is equal to the ratio of the two geometric means of the roots of the individual matrices

$$\underset{\sim}{D}^{(i)'}\underset{\sim}{\Sigma}^{(i)^{-1}}\underset{\sim}{D}^{(i)}.$$

Q.E.D.

Thus the efficiency using the curvature provides us with a typical or average efficiency of our tests over all $\underset{\sim}{\lambda} \neq \underset{\sim}{0}$ and in the case $t_1 = t_2$ considerably simplifies the Pitman ARE since we have an alternative to simply placing bounds on the efficiency. Theorem 2.6.3 is a generalization of Bickel's (1965) result when he considered the multivariate one sample location problem with $q = t_1 = t_2$.

We can see that if $q \leq t_i$ for one i but $q > t_i$ for the other that the latter power function has zero curvature at the origin while the former has positive curvature so we could conclude that the former test is unquestionably superior to the other when using this efficiency criterion. If on the other hand both $t_i < q$ $(i = 1,2)$ then both curvatures are zero and this method of comparison is not able to discriminate between the tests. In this latter case we consider an alternative measure of ARE in a later section. Returning to the conditions of Theorem 2.6.2 let us define an <u>equi-power contour</u> as the set of $\underset{\sim}{\lambda}$ such that $P_i(\underset{\sim}{\lambda}) = \alpha + c$ $(0 < c \leq 1 - \alpha)$ where c is arbitrary. Consider the power function as written in equation (2.3.3) at the end of section 2.3.3; with $M_i = 1$ we find that an equi-power contour is equal to:

$$\alpha + \underset{\sim}{\lambda}'\left(\left(\frac{\partial^2 P_i(\underset{\sim}{\lambda})}{\partial\lambda_\ell \partial\lambda_K}\right)\Bigg|_{\underset{\sim}{\lambda}=\underset{\sim}{0}}\right)\underset{\sim}{\lambda}$$

$$+ \sum_{r=2}^{\infty}\left(\sum_{I_{2r}} \frac{\lambda_{\ell_1}\cdots\lambda_{\ell_{2r}}}{(2Mr)!} \frac{\partial^{2r} \quad P_i(\underset{\sim}{\lambda})}{\partial\lambda_{\ell_1}\cdots\partial\lambda_{\ell_{2r}}}\Bigg|_{\underset{\sim}{\lambda}=\underset{\sim}{0}}\right) = \alpha + c .$$

Now if we choose $\underset{\sim}{\lambda}$ sufficiently small so that we can ignore infinitesimals in λ_ℓ of order 4 and higher we see that the equi-power contour reduces to

$$\underset{\sim}{\lambda}'\left(\left(\frac{\partial^2 P_i(\underset{\sim}{\lambda})}{\partial\lambda_\ell\partial\lambda_K}\Bigg|_{\underset{\sim}{\lambda}=\underset{\sim}{0}}\right)\right)\underset{\sim}{\lambda} = c\ .$$

The matrix in this equation is positive definite when $q \leq t_i$ and hence this contour is an ellipsoid. Now the volume of this ellipsoid is given by

$$V = \int_{\underset{\sim}{\lambda}'\underset{\sim}{A}_i\underset{\sim}{\lambda}=c} \prod_{\ell=1}^{q} d\lambda_\ell \text{ where } A_i = \left(\left(\left(\frac{\partial^2 P_i(\underset{\sim}{\lambda})}{\partial\lambda_\ell\partial\lambda_K}\Bigg|_{\underset{\sim}{\lambda}=\underset{\sim}{o}}\right)\right)_{\ell,K=1,\ldots,q}\right)$$

Now $\underset{\sim}{A}_i$ is symmetric positive definite so there exists an orthogonal matrix $\underset{\sim}{B}$ such that $\underset{\sim}{B}'\underset{\sim}{A}_i\underset{\sim}{B} = R = \text{diag}\ (r_{i1},\ldots,r_{iq})$, r_{ij} = roots of $\underset{\sim}{A}_i$. Let $\underset{\sim}{C} = \underset{\sim}{B}\underset{\sim}{R}^{-1/2}$ and let $\underset{\sim}{\lambda} = \underset{\sim}{C}\ \underset{\sim}{\lambda}^*$

therefore $V = \displaystyle\int_{\underset{\sim}{\lambda}^{*\prime}\underset{\sim}{\lambda}^{*}=c} |\underset{\sim}{B}\underset{\sim}{R}^{-1/2}| \prod_{\ell=1}^{q} \lambda_\ell^* = |\underset{\sim}{B}\underset{\sim}{R}^{-1/2}|\ S^q(c)$ where $S^q(c)$ is the volume of the q-dimensional sphere of radius $\sqrt{c}$. Since $\underset{\sim}{R}^{-1/2}\ \underset{\sim}{B}'\ \underset{\sim}{A}_i\ \underset{\sim}{B}\ \underset{\sim}{R}^{-1/2} = \underset{\sim}{I}$ we notice that $|\underset{\sim}{A}_i|^{-1/2} = |\underset{\sim}{B}\underset{\sim}{R}^{-1/2}|$ and consequently $V = |\underset{\sim}{A}_i|^{-1/2}\ S_q(c) = \dfrac{S_q(c)}{|\underset{\sim}{A}_i|^{1/2}}$. The volume of the equi-power contour is thus inversely proportional to the square root of the generalized curvature of the power function at the origin. We see that to increase the curvature is to decrease the volume of a certain infinitesimal ellipsoid. In comparing the power functions of test 2 and test 1 along an equi-power contour, we see that if CARE > 1 then the power function of test 2 encloses an ellipsoid of smaller volume than test 1 along the same contour; intuitively the second test seems to satisfy an attractive property of <u>faster average growth locally</u> than test 1.

Under the assumptions of Theorem 2.6.2 both tests are unbiased, and we see that if CARE >1 then test 2 is more nearly optimum in the type D sense described by Isaacson (1951). If CARE > 1 we may say that test 2 has faster <u>average</u> growth locally when compared to test 1 but we should keep in mind that there may be some directions for which test 2 has <u>lower</u> power than test 1. This type of deficiency is almost certain to exist in any <u>single</u> quantity which attempts to measure multivariate efficiency.

If we consider the case where $M_1 = M_2 = M$ with $M > 1$, neglecting terms in $\underset{\sim}{\lambda}$ of power greater than the $2M^{th}$ and higher than the equi-power contour of each power function, reduces to

$$\sum_{I_{2M}} \lambda_{\ell_1} \cdots \lambda_{\ell_{2M}} \frac{1}{(2M)!} \frac{\partial^{2M} P_i(\underset{\sim}{\lambda})_{\underset{\sim}{\lambda}=\underset{\sim}{0}}}{\partial\lambda_{\ell_1} \cdots \partial\lambda_{\ell_{2M}}} = c \text{ for } c \in (0,1-\alpha] .$$

For general values of M it is not clear to me what type of contour this is or what conditions (similar to the positive definiteness assumed when M = 1), are needed on the

$$\frac{\partial^{2M} P_i(\underset{\sim}{\lambda})}{\partial\lambda_{\ell_1} \cdots \partial\lambda_{\ell_{2M}}}$$

to ensure a region which encloses an interior. Consequently attempts generalize the arguments on minimizing the volume of ellipsoids fail. One should anticipate problems in handling this for general M since even sufficient conditions to ensure a relative minimum of the power function at $\underset{\sim}{\lambda} = \underset{\sim}{0}$ are not generally available. A sufficient condition for M = 1 is, of course, the positive definiteness of the 2^{nd} partial derivative matrix. The tests considered in this work have M = 1 so we discuss the case M > 1 no longer.

2.7 Trace Asymptotic Relative Efficiency (TARE)

Again confining ourselves to the case $M_1 = M_2$ we propose another measure of ARE whose range of applicability extends to a wider class of problems than does the CARE. The criterion we propose is as follows: Definition: The trace asymptotic relative efficiency (TARE) of test ϕ_2 with respect to ϕ_1 is the $\lim\limits_{N_1 \to \infty} \frac{N_1}{N_2}$ where $\{N_1\}$ and $\{N_2\}$ are chosen such that the two tests have the same limiting average power locally over the unit sphere through the same sequence of alternative hypotheses.

Again as in section 2.6 we define locally to involve the terms up to the $2M^{th}$ derivatives in the limiting power function, and we assume negligible terms beyond the $2M^{th}$ derivatives. We expect this ARE to be an average efficiency since we are taking the average local power over all possible directions with respect to the surface area element on the q-dimensional sphere. When $M_1 = M_2 = 1$ we can write

$$P_i(\underset{\sim}{\lambda}) = \alpha + \frac{1}{2} \Delta_i [P_i(t_i + 2,0) - \alpha] + o(\lambda_{max}^4)$$

$$= \alpha + \frac{1}{2}[P_i(t_i + 2,0) - \alpha]\, \underset{\sim}{\lambda}'\, \underset{\sim}{D}^{(i)'} \underset{\sim}{\Sigma}^{(i)^{-1}} \underset{\sim}{D}^{(i)} \underset{\sim}{\lambda} + o(\lambda_{max}^4).$$

So ignoring higher order terms we obtain a local power of

(2.7.1) $P_i(\underset{\sim}{\lambda}) \doteq \alpha + \frac{1}{2}[P_i(t_i+2.0) - \alpha]\ \underset{\sim}{\lambda}'\underset{\sim}{D}^{(i)'}\underset{\sim}{\Sigma}^{(i)^{-1}}\underset{\sim}{D}^{(i)}\underset{\sim}{\lambda}.$

Let $\int_{K^q_{(R)}} f(\underset{\sim}{\lambda})dA$ denote the surface integral of the function $f(\underset{\sim}{\lambda})$ over the surface $K^q(R)$. Then we obtain the following theorem for the TARE of test 2 to test 1.

<u>Theorem 2.7.1</u>. If $\{Q_N^{(1)}\}$ and $\{Q_N^{(2)}\}$ satisfy the conditions of Theorem 2.2.1 and $M_1 = M_2 = 1$ then

$$\text{TARE} = \left\{ R(t_1,t_2,\alpha)\ \frac{\text{tr}\ \underset{\sim}{D}^{(2)'}\underset{\sim}{\Sigma}^{(2)^{-1}}\underset{\sim}{D}^{(2)}}{\text{tr}\ \underset{\sim}{D}^{(1)'}\underset{\sim}{\Sigma}^{(1)^{-1}}\underset{\sim}{D}^{(1)}} \right\}^{1/2\delta} .$$

<u>Proof</u>: We are required to choose $\{N_1\}$ and $\{N_2\}$ such that the two tests have the same average local power over the surface $||\lambda|| = 1$ and we must choose the sample sizes to guarantee this condition through the sequence H_N: $\underset{\sim}{\theta}_N = \underset{\sim}{\theta}_0 + N^{-\delta}\underset{\sim}{\lambda}$. By (2.2.21)

$$\mathcal{L}(Q_{N_i}^{(1)}|H_N) \to \chi^2(t_i,\rho_i^{2\delta}\Delta_i) \text{ as } N \to \infty$$

where ρ_i is defined in (2.2.20) and $\Delta_i = \underset{\sim}{\lambda}'\underset{\sim}{D}^{(i)'}\underset{\sim}{\Sigma}^{(i)^{-1}}\underset{\sim}{D}^{(i)}\underset{\sim}{\lambda}$.

So by (2.7.1) we require that

(2.7.2) $$\frac{1}{2}[P_1(t_1+2,0) - \alpha] \int_{||\underset{\sim}{\lambda}||=1} \rho_1^{2\delta}\underset{\sim}{\lambda}'\underset{\sim}{D}^{(1)'}\underset{\sim}{\Sigma}^{(1)^{-1}}\underset{\sim}{D}^{(1)}\underset{\sim}{\lambda}\ dA$$

$$= \frac{1}{2}[P_2(t_2+2,0) - \alpha] \int_{||\underset{\sim}{\lambda}||=1} \rho_2^{2\delta}\underset{\sim}{\lambda}'\underset{\sim}{D}^{(2)'}\underset{\sim}{\Sigma}^{(2)^{-1}}\underset{\sim}{D}^{(2)}\underset{\sim}{\lambda}dA.$$

But by a result in differential geometry [Weyl (1939)]

$$\int_{||\underset{\sim}{\lambda}||=1} \underset{\sim}{\lambda}'\underset{\sim}{B}\,\underset{\sim}{\lambda}\, dA = \frac{\mathrm{tr}\ \underset{\sim}{B}}{q} \text{ Area } (S^{q-1}_{(1)}) \tag{2.7.3}$$

so (2.7.2) becomes

$$\frac{\rho_1}{\rho_2} = \left\{ R(t_1,t_2,\alpha)\ \frac{\mathrm{tr}(\underset{\sim}{D}^{(2)'}\underset{\sim}{\Sigma}^{(2)^{-1}}\underset{\sim}{D}^{(2)})}{\mathrm{tr}(\underset{\sim}{D}^{(1)'}\underset{\sim}{\Sigma}^{(1)^{-1}}\underset{\sim}{D}^{(1)})} \right\}^{1/2\delta} .$$

Recalling how the ρ_i are defined we see the result follows.

Q.E.D.

The trace of a matrix is equal to the sum of its roots so using this definition of ARE we find that we get a quantity which is equal to the arithmetic mean of the roots. Hence, if we had defined the TARE as the limiting ratio of sample sizes such that the tests have the same arithmetic mean of the roots of the 2nd partial derivative matrix at the origin then we would have arrived at the same result. We see that if TARE > 1 then test 2 is more nearly optimum in the sense of Wald (1943) for sufficiently small $\underset{\sim}{\lambda}$, since it has greater average local power over the unit sphere.

2.8 Bahadur Efficiency

To be consistent with the notation in Chapter I we let Ω be the parameter set consisting of the values of $\underset{\sim}{\theta}$ and let Ω_0 be that subset of the parameter set consisting of the single point $\underset{\sim}{\theta}_0$. We consider in the Bahadur method of comparison, a fixed alternative hypothesis, H_a: $\underset{\sim}{\theta} = \underset{\sim}{\theta}_a \in \Omega - \Omega_0$; the statistics we compare are

$$Q_N^{(i)} = N(\underset{\sim}{T}_N^{(i)} - \underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta}_0))'\hat{\underset{\sim}{\Sigma}}^{(i)^{-1}}(\underset{\sim}{T}_N^{(i)} - \underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta}_0)); \quad i = 1,2.$$

We give sufficient conditions under which $\{\overset{+}{\sqrt{Q_N^{(i)}}}\}$ is a standard sequence for testing H_0:

Theorem 2.8.1. Suppose that to test the hypothesis $\underset{\sim}{\theta} = \underset{\sim}{\theta}_0$ we have the two sequences of test statistics, $\{Q_N^{(1)}\}$, and $\{Q_N^{(2)}\}$, satisfying the following three conditions:

(2.8.1) a) $Q_N^{(i)}$ has a central χ^2 distribution with t_i degrees of freedom as $N \to \infty$ when $\underset{\sim}{\theta} = \underset{\sim}{\theta}_0$.

(2.8.2) b) $(\underset{\sim}{T}_N^{(i)} - \underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta}_0)) \overset{a.s.}{\to} \underset{\sim}{\eta}^{(i)}(\underset{\sim}{\theta})$ as $N \to \infty$ for all $\underset{\sim}{\theta} \in \Omega - \Omega_0$ where $\underset{\sim}{\eta}^{(i)}(\underset{\sim}{\theta})$ is a fixed non-null vector of t_i-components.

(2.8.3) c) $\hat{\underset{\sim}{\Sigma}}_N^{(i)} \overset{a.s.}{\to} \underset{\sim}{\Sigma}^{(i)}(\underset{\sim}{\theta})$ as $N \to \infty$ for $\underset{\sim}{\theta} \in \Omega - \Omega_0$ where $\underset{\sim}{\Sigma}^{(i)}(\underset{\sim}{\theta})$ is positive definite.

Then the sequence $\{\overset{+}{\sqrt{Q_N^{(i)}}}\}$ is a standard sequence for testing H_0 with $a_i = 1$, $b_i^2(\underset{\sim}{\theta}) = \underset{\sim}{\eta}^{(i)'}(\underset{\sim}{\theta})\, \underset{\sim}{\Sigma}^{(i)^{-1}}(\underset{\sim}{\theta})\, \underset{\sim}{\eta}^{(i)}(\underset{\sim}{\theta})$.

Proof: Let $R_N^{(i)} = \overset{+}{\sqrt{Q_N^{(i)}}}$. We must now verify the three conditions for a standard sequence as indicated in the definition in Chapter I.

I. $\mathcal{L}(Q_N^{(i)}) \to \chi^2(t_i)$ implies

$\mathcal{L}(R_N^{(i)}) \to \chi\ (t_i)$, where $\chi(t_i)$ is a chi-distribution with t_i degrees of freedom, for $\underset{\sim}{\theta} \in \Omega_0$ since $R_N^{(i)}$ is a continuous function of $Q_N^{(i)}$. Hence we see there exists a continuous distribution function $F_i(X)$ such that

$$\lim_{N \to \infty} \Pr[R_N^{(i)} \leq X] = F_i(X) \ \forall \ X \text{ and } \underset{\sim}{\theta} \in \Omega_0.$$

II. For each X and $\underset{\sim}{\theta} \in \Omega_0$ we notice that

$\Pr_{\underset{\sim}{\theta}}[R_N^{(i)} \leq X] = \Pr_{\underset{\sim}{\theta}}[Q_N^{(i)} \leq X^2]$ therefore from I we see that

since $F_i(X) = \Pr[\chi^2(t_i) \leq X^2]$ that

$$1 - F_i(X) = \int_{X^2}^{\infty} (2^{\frac{t_i}{2}} \Gamma(\tfrac{t_i}{2}))^{-1} e^{-z/2} z^{\frac{t_i-2}{2}} dz$$

let $u = z^{\frac{t_i-2}{2}}$ $\qquad dv = e^{-z/2} dz$

$du = \frac{t_i-2}{2} z^{\frac{t_i-4}{2}} dz \qquad v = -2 e^{-z/2}$

So for each X

$$(2.8.4) \quad 1 - F_i(X) = (2^{t_i/2} \Gamma(t_i/2))^{-1} [2e^{-X^2/2} X^{(t_i-2)} + 2 \int_{X^2}^{\infty} e^{-z/2} (\frac{t_i-2}{2}) z^{\frac{t_i-2}{2} -1} dz]$$

let $\omega = z/2$ then (2.8.4) becomes

$$(2.8.5) \quad 1 - F_i(X) = (2^{t_i/2} \Gamma(t_i/2))^{-1} 2[e^{-X^2/2} X^{(t_i-2)} + 2^{\frac{t_i-2}{2}} (t_i-2) \int_{\frac{X^2}{2}}^{\infty} e^{-\omega} \omega^{\frac{t_i-2}{2} -1} d\omega].$$

Now we know that

$$\int_{\frac{X^2}{2}}^{\infty} e^{-\omega}\,\omega^{\frac{t_i-2}{2}-1}\,d\omega = O\left(e^{-\frac{X^2}{2}}\left(\frac{X^2}{2}\right)^{\frac{t_i-2}{2}-1}\right) \text{ as } X \to +\infty$$

therefore as $X \to \infty$ (2.8.5) is

$$1 - F_i(X) = (2^{t_i-2}\,\Gamma(t_i/2))^{-1}\, 2[e^{-X^{t_i-2}} + 2^{t_i-2}(t_i-2)\, O(e^{-X^2/2}X^{t_i-4})]$$

So as $X \to \infty$

$$1 - F_i(X) = e^{-X^2/2}\,X^{t_i-2}[2\,(2^{t_i/2}\,\Gamma(t_i/2))^{-1} + O(X^{-2})]$$

$$= e^{-X^2/2}\,X^{t_i-2}\,2(2^{t_i/2}\,\Gamma(t_i/2))^{-1}\,[1 + o(1)] \text{ as } X \to \infty .$$

Hence,

$$\log_e(1 - F_i(X) = -\frac{X^2}{2} + (t_i-2)\log_e X + \log c_{t_i} + o(1) \text{ as } X \to \infty$$

where $c_{t_i} = 2(2^{t_i/2}\,\Gamma(t_i/2))^{-1}$ and we see

(2.8.6) $$\log_e(1-F_i(X)) = -\frac{X^2}{2}\left[1 - \frac{2(t_i-2)\log_e X}{X^2} + o(1)\right] \text{ as } X \to \infty$$

$$\log_e(1-F_i(X)) = -\frac{X^2}{2}[1 + o(1)] \text{ as } X \to \infty.$$

So condition II required for a standard sequence is satisfied for $R_N^{(i)}$ with $a_i = 1$.

III. Since $(T_{\underset{\sim}{N}}^{(i)} - \mu_{\underset{\sim}{N}}^{(i)}(\underset{\sim}{\theta}_0)) \overset{a.s.}{\to} \underset{\sim}{\eta}^{(i)}(\underset{\sim}{\theta})\ \forall\ \underset{\sim}{\theta} \in \Omega - \Omega_0$ as $N \to \infty$

and

$$\hat{\underset{\sim}{\Sigma}}{}_N^{(i)} \overset{a.s.}{\rightarrow} \underset{\sim}{\Sigma}^{(i)}(\underset{\sim}{\theta}) \text{ as } N \rightarrow \infty, \ \forall\ \theta \in \Omega - \Omega_0$$

therefore we see that

$$N^{-1}Q_N^{(i)} \overset{a.s.}{\rightarrow} \underset{\sim}{\eta}^{(i)'}(\underset{\sim}{\theta})\ \underset{\sim}{\Sigma}^{(i)^{-1}}(\underset{\sim}{\theta})\underset{\sim}{\eta}^{(i)}(\underset{\sim}{\theta}) \text{ as } N \rightarrow \infty .$$

As a result, we have

$$N^{-1/2}R_N^{(i)} \overset{a.s.}{\rightarrow} \sqrt{\underset{\sim}{\eta}^{(i)'}(\underset{\sim}{\theta})\underset{\sim}{\Sigma}^{(i)^{-1}}(\underset{\sim}{\theta})\underset{\sim}{\eta}^{(i)}(\underset{\sim}{\theta})}\ \forall\ \underset{\sim}{\theta} \in \Omega - \Omega_0.$$

So if $\underset{\sim}{\eta}^{(i)}(\underset{\sim}{\theta}) \neq \underset{\sim}{0}$ for $\underset{\sim}{\theta} \in \Omega - \Omega_0$ then by the assumed positive definitiveness of $\underset{\sim}{\Sigma}^{(i)}(\underset{\sim}{\theta})\ \forall\ \underset{\sim}{\theta}$; the result of the theorem follows.

Q.E.D.

As a result if we assume that our statistics satisfied the conditions of Theorem 2.8.1, which they do for all the applications considered, then the Bahadur efficiency of ϕ_2 with respect to ϕ_1 is simply

$$(2.8.7) \qquad e^B_{2,1}(\underset{\sim}{\theta}) = \frac{a_2 b_2^2(\underset{\sim}{\theta})}{a_1 b_1^2(\underset{\sim}{\theta})} = \frac{\underset{\sim}{\eta}^{(2)'}(\underset{\sim}{\theta})\underset{\sim}{\Sigma}^{(2)^{-1}}(\underset{\sim}{\theta})\underset{\sim}{\eta}^{(2)}(\underset{\sim}{\theta})}{\underset{\sim}{\eta}^{(1)}(\underset{\sim}{\theta})\underset{\sim}{\Sigma}^{(1)^{-1}}(\underset{\sim}{\theta})\underset{\sim}{\eta}^{(1)}(\underset{\sim}{\theta})} \text{ for } \underset{\sim}{\theta} \in \Omega - \Omega_0.$$

We notice that (2.8.7) depends on the parameter $\underset{\sim}{\theta}$ in a most complicated fashion; primarily we notice that the covariance $\underset{\sim}{\Sigma}^{(i)}(\underset{\sim}{\theta})$ depends on the fixed alternative value chosen. For some problems $\underset{\sim}{\Sigma}^{(i)}(\underset{\sim}{\theta})$ would not of course depend on $\underset{\sim}{\theta}$, for example, in the standard multivariate one sample problem when the data represent observations from a normal population with specified location $\underset{\sim}{\theta}_0$ under the null hypothesis and location $\underset{\sim}{\theta}$ for the alternatives. For other problems this dependence of the covariance on $\underset{\sim}{\theta}$ is real, for example, if in the above problems we consider nonparametric tests constructed under an appropriate invariance structure we would find that $\underset{\sim}{\Sigma}^{(i)}(\underset{\sim}{\theta})$ would depend on $\underset{\sim}{\theta}$. So we see this dependence of the matrix $\underset{\sim}{\Sigma}^{(i)}(\underset{\sim}{\theta})$ on $\underset{\sim}{\theta}$ creates a problem in the evaluation of this efficiency. We should also notice by (2.8.6) that the measure of efficiency, which was constructed for large X, is not sensitive to the degrees of freedom of the test, as the dominant terms in (2.8.6) do not involve t_i as $X \to \infty$. For these reasons we find the Bahadur criterion an unattractive measure of efficiency for our problem.

Bahadur (1960) established sufficient conditions under which the Pitman ARE is equal to the limiting value of the Bahadur efficiency when the <u>univariate</u> parameter θ approaches the null value $\underset{\sim}{\theta}_0$. We now establish a relationship between the limiting Bahadur efficiency as $\underset{\sim}{\theta} \to \underset{\sim}{\theta}_0$ and the LARE which we proposed earlier

<u>Theorem 2.8.2</u>. If for testing H_0: $\underset{\sim}{\theta} = \underset{\sim}{\theta}_0$ we have the two sequences of test statistics $\{Q_N^{(i)}\}$ $i = 1,2$, each sequence satisfying the conditions of Theorem 2.8.1, and

(2.8.8) a) $\underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta})$ has first partial derivatives with respect to the elements of $\underset{\sim}{\theta}$ and uniformly continuous in a small open neighborhood of $\underset{\sim}{\theta}_0$; for every N,

(2.8.9) b) $$\lim_{N \to \infty} \frac{\partial\, \underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta}_0)}{\partial\, \underset{\sim}{\theta}} = \underset{\sim}{D}^{(i)} \text{ exists}$$

(2.8.10) c) $$\underset{\sim}{\Sigma}^{(i)}(\underset{\sim}{\theta})\underset{\sim}{\Sigma}^{(i)^{-1}}(\underset{\sim}{\theta}_0) \to \underset{\sim}{I} \text{ as } \underset{\sim}{\theta} \to \underset{\sim}{\theta}_0$$

(2.8.11) d) $$\underset{\sim}{\eta}^{(i)}(\underset{\sim}{\theta}) = \lim_{N \to \infty} [\underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta}) - \underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta}_0)]$$

then

$$(2.8.12)\quad e^{B}_{2,1}(\underset{\sim}{\theta}) = \frac{(\underset{\sim}{\theta}-\underset{\sim}{\theta}_0)'\underset{\sim}{D}^{(2)'}\underset{\sim}{\Sigma}^{(2)^{-1}}(\underset{\sim}{\theta}_0)\underset{\sim}{D}^{(2)}(\underset{\sim}{\theta}-\underset{\sim}{\theta}_0)[1+o(1)]}{(\underset{\sim}{\theta}-\underset{\sim}{\theta}_0)'\underset{\sim}{D}^{(1)'}\underset{\sim}{\Sigma}^{(1)^{-1}}(\underset{\sim}{\theta}_0)\underset{\sim}{D}^{(1)}(\underset{\sim}{\theta}-\underset{\sim}{\theta}_0)[1+o(1)]}$$

as $\underset{\sim}{\theta} \to \underset{\sim}{\theta}_0$

where $\underset{\sim}{D}^{(i)}$ is defined by (2.8.9).

<u>Proof</u>: By Theorem 2.8.1

$$e^{B}_{2,1}(\underset{\sim}{\theta}) = \frac{\underset{\sim}{\eta}^{(2)'}(\underset{\sim}{\theta})\underset{\sim}{\Sigma}^{(2)^{-1}}(\underset{\sim}{\theta})\underset{\sim}{\eta}^{(2)}(\underset{\sim}{\theta})}{\underset{\sim}{\eta}^{(1)'}(\underset{\sim}{\theta})\underset{\sim}{\Sigma}^{(1)^{-1}}(\underset{\sim}{\theta})\underset{\sim}{\eta}^{(1)'}(\underset{\sim}{\theta})} \quad \text{for every } \underset{\sim}{\theta} \in \Omega - \Omega_0.$$

But by (2.8.11) we have for every $\underset{\sim}{\theta} \in \Omega - \Omega_0$

$$(2.8.13)\quad e^{B}_{2,1}(\underset{\sim}{\theta}) = \lim_{N\to\infty} \frac{[\underset{\sim}{\mu}_N^{(2)}(\underset{\sim}{\theta})-\underset{\sim}{\mu}_N^{(2)}(\underset{\sim}{\theta}_0)]'\underset{\sim}{\Sigma}^{(2)^{-1}}(\underset{\sim}{\theta})[\underset{\sim}{\mu}_N^{(2)}(\underset{\sim}{\theta})-\underset{\sim}{\mu}_N^{(2)}(\underset{\sim}{\theta}_0)]}{[\underset{\sim}{\mu}_N^{(1)}(\underset{\sim}{\theta})-\underset{\sim}{\mu}_N^{(1)}(\underset{\sim}{\theta}_0)]'\underset{\sim}{\Sigma}^{(1)^{-1}}(\underset{\sim}{\theta})[\underset{\sim}{\mu}_N^{(1)}(\underset{\sim}{\theta})-\underset{\sim}{\mu}_N^{(1)}(\underset{\sim}{\theta}_0)]}.$$

But by (2.8.10) there exists matrices $\underset{\sim}{\Gamma}^{(i)}(\underset{\sim}{\theta})$ defined by

$$\underset{\sim}{\Sigma}^{(i)^{-1}}(\underset{\sim}{\theta}) = \underset{\sim}{\Sigma}^{(i)^{-1}}(\underset{\sim}{\theta}_0) + \underset{\sim}{\Gamma}^{(i)}(\underset{\sim}{\theta})$$

where

$$\underset{\sim}{\Gamma}^{(i)}(\underset{\sim}{\theta}) = \underset{\sim}{o}(1) \text{ as } \underset{\sim}{\theta} \to \underset{\sim}{\theta}_0 .$$

So for each i we see that for $\underset{\sim}{\theta}$ close to $\underset{\sim}{\theta}_0$ we have for each N:

$$(2.8.14)\quad [\underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta})-\underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta}_0)]'\underset{\sim}{\Sigma}^{(i)^{-1}}(\underset{\sim}{\theta}_0)[\underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta})-\underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta}_0)] + o(1).$$

Now choose $\hat{\underset{\sim}{\theta}}$ such that

$$\frac{\partial \underset{\sim}{\mu}_N^{(i)}(\hat{\underset{\sim}{\theta}})}{\partial\, \underset{\sim}{\theta}}(\underset{\sim}{\theta} - \underset{\sim}{\theta}_0) = \underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta}) - \underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta}_0)$$

but by (2.8.8)

$$\frac{\partial \underset{\sim}{\mu}_N^{(i)}(\hat{\underset{\sim}{\theta}})}{\partial \underset{\sim}{\theta}} = \frac{\partial \underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta}_0)}{\partial \underset{\sim}{\theta}} [1 + o(1)] \text{ as } \underset{\sim}{\theta} \to \underset{\sim}{\theta}_0 .$$

So we see that (2.8.14) becomes for each i

$$(2.8.15) \quad (\underset{\sim}{\theta}-\underset{\sim}{\theta}_0)' \left[\frac{\partial \underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta}_0)}{\partial \underset{\sim}{\theta}}\right]' \underset{\sim}{\Sigma}^{(i)^{-1}}(\underset{\sim}{\theta}_0) \left[\frac{\partial \underset{\sim}{\mu}_N^{(i)}(\underset{\sim}{\theta}_0)}{\partial \underset{\sim}{\theta}}\right] (\underset{\sim}{\theta}-\underset{\sim}{\theta}_0) + o(1)$$

$$\text{as } \underset{\sim}{\theta} \to \underset{\sim}{\theta}_0.$$

Now this is true for each N hence we have that as $\underset{\sim}{\theta} \to \underset{\sim}{\theta}_0$ that

$$\underset{\sim}{\eta}^{(i)'}(\underset{\sim}{\theta})\underset{\sim}{\Sigma}^{(i)^{-1}}(\underset{\sim}{\theta})\underset{\sim}{\eta}^{(i)}(\underset{\sim}{\theta}) = (\underset{\sim}{\theta}-\underset{\sim}{\theta}_0)'\underset{\sim}{D}^{(i)'}\underset{\sim}{\Sigma}^{(i)^{-1}}(\underset{\sim}{\theta}_0)D^{(i)}(\underset{\sim}{\theta}-\underset{\sim}{\theta}_0) + o(1).$$

Hence the result (2.8.12) follows.

Q.E.D.

We notice that if we let $\underset{\sim}{\lambda} = (\underset{\sim}{\theta} - \underset{\sim}{\theta}_0)$ then $\underset{\sim}{\lambda} \to \underset{\sim}{0}$ as $\underset{\sim}{\theta} \to \underset{\sim}{\theta}_0$ and we see that the quantity (2.8.12) is similar to the LARE. In short if the $\{Q_N^{(i)}\}$ satisfied the conditions of Theorem 2.2.1 with M = 1 we see that the LARE and the limiting Bahadur we have considered here would be nearly identical. The limiting Bahadur efficiency in (2.8.12) even though simple to compute does not account for differences in degrees of freedom.

CHAPTER III

APPLICATION OF TARE AND CARE TO THE ONE SAMPLE GROWTH CURVE PROBLEM

3.1 Introduction

In this chapter we apply the measures of ARE proposed in Chapter II to evaluate the efficiencies of some common procedures used in the study of polynomial growth curve models. The general statistical model along with its reduction for the one sample growth curve model is given in Section 3.2. While the results of this chapter are easily extended for other hypotheses, the hypothesis we study is the hypothesis of a constant growth curve over time. In Section 3.4 we discuss the reduction of the basic data to estimates of the assumed model. Once a given model has been decided upon a common procedure to follow in practice is to estimate the growth curve parameters by unweighted least squares. This basic reduction is assumed throughout this chapter. After obtaining this set of summary statistics for each observation vector we apply the Hotelling T^2 test to test the null hypothesis. We then evaluate the TARE and CARE of an incorrect specification of the model to the correct specification of the number of parameters. Similar results are obtained in Section 3.5 for the corresponding nonparametric rank scores procedure. A brief summary of the one sample rank scores procedure is presented in Section 3.5.2. Section 3.6 is devoted to a comparison of the parametric and nonparametric procedures based on the trace and curvature criteria of ARE. Bounds for the TARE are derived similar to those available for the CARE. The chapter is concluded with a brief section on the use of covariance adjustment of the statistics with the higher order polynomial coefficients.

3.2 The Statistical Model

The model given in this section is sufficiently general to encompass the c-sample problem and more complicated designs and will be used in a later chapter on the c-sample problem. We consider two index sets, I and T where

$$(3.2.1) \quad I = \{\underset{\sim}{i} = (i_1, \ldots, i_M) \colon 1 \le i_j \le K_j \ (\ge 1) \text{ for } j = 1, \ldots, M\}$$

and

$$(3.2.2) \quad T = \{t_1. \ \ldots \ t_h\} \text{ where } t_1 < \ldots < t_h.$$

We notice that I contains $K_1 K_2 \ldots K_M$ distinct points and we may think of I as specifying the design across individuals. T is the set of distinct time points. Corresponding to each $\underset{\sim}{i} \in I$ and $t_\ell \in T$ we have a set of $n(\underset{\sim}{i})$ random vectors of b elements, namely

$$(3.2.3) \quad \underset{1\times b}{\underset{\sim}{X}_S(\underset{\sim}{i}, t_\ell)} = \begin{pmatrix} X_{S1}(\underset{\sim}{i}, t_\ell) \\ \vdots \\ X_{Sb}(\underset{\sim}{i}, t_\ell) \end{pmatrix}' \qquad S = 1, 2, \ldots, n(\underset{\sim}{i}).$$

In addition, let the h×b matrix

$$(3.2.4) \quad \underset{\sim}{X}_S(\underset{\sim}{i}) = \begin{pmatrix} \underset{\sim}{X}_S(\underset{\sim}{i}, t_1) \\ \vdots \\ \underset{\sim}{X}_S(\underset{\sim}{i}, t_h) \end{pmatrix} \qquad S = 1, 2, \ldots, n(\underset{\sim}{i}).$$

We assume that $\underset{\sim}{X}_S(\underset{\sim}{i})$ is a collection of independent stochastic matrices from a bh-variate continuous distribution function $G(\underset{\sim}{X};\underset{\sim}{i})$ for $\underset{\sim}{X} \varepsilon R^{bh}$.

For the normal theory, we would assume

$$G(\underset{\sim}{X};\underset{\sim}{i}) = N_{bh}(\underset{\sim}{M}(i), \underset{bh\times bh}{\underset{\sim}{\Sigma}})$$

where the location vector

$$\underset{\sim}{M}(\underset{\sim}{i}) = \begin{pmatrix} \underset{\sim}{\mu}(\underset{\sim}{i},t_1) \\ \vdots \\ \underset{\sim}{\mu}(\underset{\sim}{i},t_h) \end{pmatrix} \qquad \underset{\sim}{i} \varepsilon I \text{ and}$$

$\underset{\sim}{\Sigma}$ does not depend on $\underset{\sim}{i}$. We note that with this specification

$$\underset{\sim}{X}_S(\underset{\sim}{i}) - \underset{\sim}{M}(\underset{\sim}{i}) \sim N_{bh}(\underset{\sim}{0},\underset{\sim}{\Sigma}) \qquad S = 1,\ldots,n(\underset{\sim}{i})$$

that is, the distribution of $\underset{\sim}{X}_S(\underset{\sim}{i}) - \underset{\sim}{M}(\underset{\sim}{i})$ does not depend on $\underset{\sim}{i}$. To eliminate the normality assumption, in the nonparametric approach we assume that

$$G(\underset{\sim}{X};\underset{\sim}{i}) = G(\underset{\sim}{X} + \underset{\sim}{M}(\underset{\sim}{i})) \qquad \underset{\sim}{i} \varepsilon I.$$

In this way $\underset{\sim}{X}_S(\underset{\sim}{i}) - \underset{\sim}{M}(\underset{\sim}{i})$ is distributed independently of $\underset{\sim}{i}$.

For the growth curve model one assumes that the $\underset{\sim}{\mu}(\underset{\sim}{i},t_\ell)$ can be written as a function of certain parameters $\underset{\sim}{\theta}(\underset{\sim}{i})$ and the time point, t_ℓ. For example we may assume that

$$\mu_j(\underset{\sim}{i},t_\ell) = \gamma_j(\underset{\sim}{\theta}_j(\underset{\sim}{i}), t_\ell) \qquad \begin{array}{l} \underset{\sim}{i} \varepsilon I \\ 1 \le j \le b \\ 1 \le \ell \le h \end{array} \tag{3.2.7}$$

where $\underset{\sim}{\theta}_j(\underset{\sim}{i})$ is a vector of r elements ($r \le h$). Since the $\underset{\sim}{\theta}_j(\underset{\sim}{i})$ are assumed to be unknown we can consider problems in testing hypotheses concerning the $\underset{\sim}{\theta}_j(\underset{\sim}{i})$ and the estimation of $\underset{\sim}{\theta}_j(\underset{\sim}{i})$. If we let

$$(3.2.8) \qquad \underset{r\times b}{\underset{\sim}{\theta}(\underset{\sim}{i})} = (\underset{\sim}{\theta}_1(\underset{\sim}{i}),\dots,\underset{\sim}{\theta}_b(\underset{\sim}{i})) \qquad \underset{\sim}{i} \in I$$

then we wee that we have a reparameterization from $\underset{\sim}{M}(\underset{\sim}{i})$ to $\underset{\sim}{\theta}(\underset{\sim}{i})$ which is dimension-reducing. In this chapter we are interested in evaluating the loss in the incorrect specification of $\underset{\sim}{\theta}(\underset{\sim}{i})$ for the one sample problem.

We shall consider only polynomial functions of t_ℓ where $\underset{\sim}{\theta}_j(\underset{\sim}{i})$ are the coefficients of these polynomials. To be explicit, we require the functions γ_j defined by (3.2.7) to be polynomials in t_ℓ. The set T corresponds to h distinct abscissae and it is well known that corresponding to this set T there exists a unique set of h orthogonal polynomials $P_0(t)$, ..., $P_{h-1}(t)$ satisfying the conditions:

$$(3.2.9) \qquad \sum_{j=1}^{h} P_i(t_j)P_k(t_j) = \begin{cases} 1 & i = K \\ 0 & i \neq K \end{cases}$$

and $P_i(t)$ is a polynomial of exactly degree i in t. We denote by the vector $\underset{\sim}{b}_i$ the value of $P_i(t)$ for each t_ℓ, $\ell = 1, \dots, h$; that is

$$(3.2.10) \qquad \underset{h\times 1}{\underset{\sim}{b}_i} = \begin{bmatrix} P_i(t_1) \\ \vdots \\ P_i(t_h) \end{bmatrix} \qquad i = 0,1,\dots,h-1.$$

We shall consider in this study only the case b = 1. We observe that for the one sample problem I consists of only one element so we omit the subscripts $\underset{\sim}{i}$ in the remaining discussion of the one sample case. In the one sample problem with b = 1 we have:

$$\underset{h\times 1}{\underset{\sim}{X}_S} = \begin{pmatrix} X_{S1}(t_1) \\ \vdots \\ X_{Sh}(t_h) \end{pmatrix} \qquad S = 1,2,\ldots,n$$

which are independent and identically distributed as $G(\underset{\sim}{X})$.

3.3 Comments on Data Reduction

Since we consider only polynomial models we see that $\underset{\sim}{\mu}(t)$ can be any of the following:

(3.3.1) $\mu(t_\ell) = \alpha_0 P_0(t_\ell) + \alpha_1 P_1(t_\ell) + \ldots + \alpha_{h-1} P_{h-1}(t_\ell)$

(3.3.2) $\mu(t_\ell) = \alpha_0 P_0(t_\ell) + \alpha_1 P_1(t_\ell) + \ldots + \alpha_{h-2} P_{h-2}(t_\ell)$

(3.3.3) $\mu(t_\ell) = \alpha_0 P_0(t_\ell) + \alpha_1 P_1(t_\ell)$ for each $\ell = 1,2,\ldots,h$.

With each of these models one could construct a test of the hypothesis that the growth curve is constant with respect to t_ℓ. For example, for the model (3.3.1) the appropriate hypothesis to test is

$$(3.3.4) \qquad H_0: \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{h-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} ; \quad \alpha_0 \text{ unspecified.}$$

While under the assumption of model (3.3.3) the appropriate hypothesis to test is:

$$(3.3.5) \qquad H_0: \ \alpha_1 = 0, \quad \alpha_0 \text{ unspecified}$$

Given what maximum degree of the model is assumed then we can determine which parameters to test for equality to zero. Once it has been decided how many parameters we wish to estimate we then have the problem of reducing each observation vector of h-elements to a vector of t-elements which represent the estimates for the individual's growth curve. One reduction technique is the method of unweighted least squares, which amounts to choosing for each individual the vector $\underset{\sim}{\alpha}$ which minimizes the quantity

$$(3.3.6) \qquad (\underset{\sim}{X}_S - \underset{\sim}{B}\underset{\sim}{\alpha})' \; (\underset{\sim}{X}_S - \underset{\sim}{B}\underset{\sim}{\alpha})$$

where $\underset{\sim}{B}\underset{\sim}{\alpha}$ is the assumed growth curve model. The estimate for each $\underset{\sim}{X}_S$ is denoted by $\underset{\sim}{Y}_S$ and is equal to $(\underset{\sim}{B}'\underset{\sim}{B})^{-1}\underset{\sim}{B}'\underset{\sim}{X}_S$, i.e.

$$\underset{\sim}{Y}_S = (\underset{\sim}{B}'\underset{\sim}{B})^{-1} \underset{\sim}{B}'\underset{\sim}{X}_S .$$

Now if $\underset{\sim}{B}$ is the matrix of orthogonal polynomials $\underset{\sim}{P}$ satisfying (3.2.9) then,

$$(3.3.7) \qquad \underset{\sim}{Y}_S = \underset{\sim}{B}'\underset{\sim}{X}_S .$$

Under the assumption of the Gauss-Markov theorem this is the minimum variance linear unbiased estimator for $\underset{\sim}{\alpha}$ if the covariance matrix of $\underset{\sim}{X}_S$ assumed to exist) is equal to $\sigma^2 \underset{h\times h}{I}$. When this covariance pattern is incorrect then the estimator $\underset{\sim}{Y}_S$ is not the best linear unbiased estimator (BLUE), which in this case is given by the method of generalized least squares.

In the absence of information on the true covariance $\underset{\sim}{\Sigma}$, the estimator (3.3.7) is the estimator suggested by several authors, [e.g. Potthoff and Roy, 1964]. Other estimates have also been proposed which introduce stochastic weights [e.g. Grizzle and Allen, 1969]; however, we explore estimators of the form (3.3.7) since these are also used frequently in practice.

The estimators needed to test the null hypothesis would not involve the first element of $\underset{\sim}{Y}_S$ since this estimates α_0. The estimators for testing the null hypothesis are:

$$(3.3.8) \qquad \underset{\sim}{Y}_{Si} = \underset{\sim}{B}_i' \underset{\sim}{X}_S \qquad S = 1,\ldots,n$$

where $\underset{\sim}{B}_i$ correspond to the orthogonal polynomials $\underset{\sim}{b}_1,\ldots,\underset{\sim}{b}_i$ to be used to test that $\alpha_1,\alpha_2,\ldots,\alpha_i$ are zero. We now consider the standard parametric analysis which is applied to these reduced data.

3.4 Parametric Procedures

In the standard parametric analysis we assume that $G(\underset{\sim}{X})$ is $N(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$; however, if one assumes that $G(\underset{\sim}{X})$ has finite moments up to order $2 + \delta$ for some $\delta > 0$ then the parametric procedure in large samples will still yield the same probability distributions. To be more explicit, we suppose that $G(\underset{\sim}{X})$ is an h-variate continuous distribution function with location vector

$$\underset{\sim}{\mu} = \begin{pmatrix} \mu(t_1) \\ \vdots \\ \mu(t_h) \end{pmatrix}$$

and dispersion $\underset{h\times h}{\underset{\sim}{\Sigma}}$, positive definite. Consider now the random sample $\underset{\sim}{X}_S$, $S = 1,\ldots,n$ and define

$$\underset{h\times 1}{\bar{\underset{\sim}{X}}_n} = n^{-1} \sum_{S=1}^{n} \underset{\sim}{X}_S$$

$$\underset{\sim}{S}_n = \frac{n}{n-1}\{n^{-1} \sum_{S=1}^{n} \underset{\sim}{X}_S \underset{\sim}{X}_S' - \bar{\underset{\sim}{X}}_n \bar{\underset{\sim}{X}}_n\},$$

by the arguments presented in Puri and Sen (1971, p. 173) we see that

$$\mathcal{L}(n^{1/2}\,\bar{\underset{\sim}{X}}_n) \to N(\underset{\sim}{0},\underset{\sim}{\Sigma}) \quad \text{if } \underset{\sim}{\mu} = 0 \quad \text{and}$$

$\underset{\sim}{S}_n \overset{P}{\to} \underset{\sim}{\Sigma}$ as $n \to \infty$. In addition for a sequence of alternative hypotheses $n^{-1/2}\,\underset{\sim}{\lambda}$ where $\underset{\sim}{\lambda}$ is fixed and non-null

$$\mathcal{L}(n^{1/2}\,\bar{\underset{\sim}{X}}\,|H_n) \to N(\underset{\sim}{\lambda},\underset{\sim}{\Sigma}) \text{ as } n \to \infty$$

therefore,

$$\mathcal{L}(n\bar{\underset{\sim}{X}}'_n\, S_n^{-1}\, \bar{\underset{\sim}{X}}_n) \to \chi^2(h,\underset{\sim}{\lambda}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{\lambda}) \text{ as } n \to \infty$$

Clearly, if we now consider the function defined by

$$\underset{\sim}{Y}_{Si} = \underset{\sim}{B}_i\, \underset{\sim}{X}_S; \text{ for each } S = 1,\dots,n$$

then we can apply the same logic to the random variables $\bar{\underset{\sim}{Y}}_n$ to find the distribution of $n^{1/2}\,\bar{\underset{\sim}{Y}}_n$ for large n. In particular consider now the special problem of comparing a linear to a quadratic model when the true model is linear.

3.4.1 Parametric Procedure - True Model Linear

If $\mu(t_\ell)$ is linear in t_ℓ, then to test the null hypothesis we consider the following:

(3.4.1) $$\underset{\sim}{Y}_{S1} = \underset{\sim}{b}'_1\, \underset{\sim}{X}_S; \quad S = 1,\dots, n$$

and if we (incorrectly) tested for the quadratic also we consider the transformation:

(3.4.2) $$\underset{\sim}{Y}_{S2} = \begin{pmatrix} b'_1 \\ b'_2 \end{pmatrix} X_S; \quad S = 1,\dots, n.$$

Since we know

$$\mu(t_\ell) = \alpha_0 P_0(t_\ell) + \alpha_1 P_1(t_\ell),$$

it follows that $E(\underset{\sim}{Y}_{S1}) = \alpha_1$ and its covariance matrix is given by $\underset{\sim}{b}_1' \underset{\sim}{\Sigma} \underset{\sim}{b}_1$. Similarly $E(\underset{\sim}{Y}_{S2}) = \binom{\alpha_1}{0}$ and the covariance matrix of $\underset{\sim}{Y}_{S2}$ is given by

$$\begin{bmatrix} \underset{\sim}{b}_1' \underset{\sim}{\Sigma} \underset{\sim}{b}_1 & \underset{\sim}{b}_1' \underset{\sim}{\Sigma} \underset{\sim}{b}_2 \\ \underset{\sim}{b}_2' \underset{\sim}{\Sigma} \underset{\sim}{b}_1 & \underset{\sim}{b}_2' \underset{\sim}{\Sigma} \underset{\sim}{b}_2 \end{bmatrix} .$$

Consider the sequence of alternative hypotheses,

$$\text{(3.4.3)} \qquad H_N: \ \alpha_{1n} = n^{-1/2} \lambda, \ \lambda \neq 0 \ .$$

Then the usual parametric one sample estimates are

$$\bar{\underset{\sim}{Y}}_i = n^{-1} \sum_{S=1}^{n} \underset{\sim}{Y}_{Si} \ ; \quad i = 1,2.$$

Since G has moments of order $2 + \delta$ ($\delta > 0$), the conditions of Theorem 2.2.1 are satisfied; the first four conditions are obvious and the fifth follows from the Berry-Esseén Theorem. Therefore

$$\mathcal{L}(n\bar{\underset{\sim}{Y}}_1'(\underset{\sim}{b}_1'\underset{\sim}{S}_n\underset{\sim}{b}_1)^{-1} \bar{\underset{\sim}{Y}}_1) \to \chi^2(1,\Delta_1) \text{ as } n \to \infty$$

where

$$\Delta_1 = \lambda_1[\underset{\sim}{b}_1' \underset{\sim}{\Sigma} \underset{\sim}{b}_1]^{-1} \lambda_1.$$

In a similar fashion the test based on

$$\sqrt{n}\, \bar{\underset{\sim}{Y}}_2' \begin{bmatrix} \underset{\sim}{b}_1'\underset{\sim}{S}_n\underset{\sim}{b}' & \underset{\sim}{b}_1'\underset{\sim}{S}_n\underset{\sim}{b}_2 \\ \underset{\sim}{b}_2'\underset{\sim}{S}_n\underset{\sim}{b}_1 & \underset{\sim}{b}_2'\underset{\sim}{S}_n\underset{\sim}{b}_2 \end{bmatrix}^{-1} \sqrt{n}\, \bar{\underset{\sim}{Y}}_2$$

has a limiting $\chi^2(2,\Delta_2)$ as n → ∞, where

$$\Delta_2 = \lambda_1(1,0)\begin{bmatrix} \underset{\sim}{b}_1' \underset{\sim}{\Sigma} \underset{\sim}{b}_1 & \underset{\sim}{b}_1' \underset{\sim}{\Sigma} \underset{\sim}{b}_2 \\ \underset{\sim}{b}_2' \underset{\sim}{\Sigma} \underset{\sim}{b}_1 & \underset{\sim}{b}_2' \underset{\sim}{\Sigma} \underset{\sim}{b}_2 \end{bmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \lambda_1.$$

This reduces to

$$\Delta_2 = \lambda_1(\underset{\sim}{b}_2' \underset{\sim}{\Sigma} \underset{\sim}{b}_2)[(\underset{\sim}{b}_1' \underset{\sim}{\Sigma} \underset{\sim}{b}_1)(\underset{\sim}{b}_2' \underset{\sim}{\Sigma} \underset{\sim}{b}_2) - (\underset{\sim}{b}_1' \underset{\sim}{\Sigma} \underset{\sim}{b}_2)(\underset{\sim}{b}_2' \underset{\sim}{\Sigma} \underset{\sim}{b}_1)]^{-1} \lambda_1.$$

If we let ϕ_1 be the test based on the correct model and ϕ_2 the test on the quadratic model; we see that the ARE of ϕ_2 with respect to ϕ_1 using trace, curvature or local criteria becomes:

$$(3.4.4)\quad \text{ARE} = R(1,2,\alpha)\,\frac{(\underset{\sim}{b}_2' \underset{\sim}{\Sigma} \underset{\sim}{b}_2)(\underset{\sim}{b}_1' \underset{\sim}{\Sigma} \underset{\sim}{b}_1)}{[(\underset{\sim}{b}_1' \underset{\sim}{\Sigma} \underset{\sim}{b}_1)(\underset{\sim}{b}_2' \underset{\sim}{\Sigma} \underset{\sim}{b}_2) - (\underset{\sim}{b}_1' \underset{\sim}{\Sigma} \underset{\sim}{b}_2)(\underset{\sim}{b}_2' \underset{\sim}{\Sigma} \underset{\sim}{b}_1)]}$$

Continuing with the parametric procedure let us derive the efficiency for the general case of overfitting the model.

3.4.2 Parametric Procedures - General Case of Overfitting

Suppose that the correct model is:

$$(3.4.5)\qquad \mu(t_\ell) = \alpha_0 P_0(t_\ell) + \sum_{j=1}^{q} \alpha_j P_j(t_\ell).$$

Make the transformation

$$(3.4.6)\qquad \underset{q\times 1}{\underset{\sim}{Y}_{S1}} = \begin{pmatrix} \underset{\sim}{b}_1' \\ \vdots \\ \underset{\sim}{b}_q' \end{pmatrix} \underset{\sim}{X}_S = \underset{\sim}{B}_1' \underset{\sim}{X}_S \qquad S = 1,\ldots,n$$

Now we see that under the model (3.4.5)

$$E(\underset{\sim}{Y}_{S1}) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_q \end{pmatrix} ;\ D(\underset{\sim}{Y}_{S1}) = \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 .$$

Consider now the sequence of alternative hypotheses

$$H_N: \begin{pmatrix} \alpha_{1n} \\ \vdots \\ \alpha_{qn} \end{pmatrix} = n^{-1/2} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_q \end{pmatrix}, \underset{\sim}{\lambda} \neq 0 .$$

Then similarly to the previous section we see that

$$Q_n^{(1)} = \sqrt{n}\, \bar{\underset{\sim}{Y}}'_{S1} (\underset{\sim}{B}'_1 \underset{\sim}{S}_n \underset{\sim}{B}_1)^{-1} \sqrt{n}\, \bar{\underset{\sim}{Y}}_{S1}$$

will have $\chi^2(q,\Delta_1)$ distribution as $n \to \infty$ where

$$\Delta_1 = \underset{\sim}{\lambda}'(\underset{\sim}{B}'_1 \underset{\sim}{\Sigma} \underset{\sim}{B}_1)^{-1} \underset{\sim}{\lambda} .$$

If we incorrectly assumed that

$$\mu(t_\ell) = \alpha_0 P_0(t_\ell) + \sum_{j=1}^{q} \alpha_j P_j(t_\ell) + \sum_{j=q+1}^{p} \alpha_j P_j(t_\ell); \qquad p \leq (h-1).$$

Denote

$$\underset{\sim}{B}'_2 = \begin{pmatrix} \underset{\sim}{b}'_{q+1} \\ \vdots \\ \underset{\sim}{b}'_p \end{pmatrix}$$

and the accompanying transformation to $\underset{\sim}{Y}_{S2}$ by

$$\underset{\sim}{Y}_{S2} = \begin{pmatrix} \underset{\sim}{B}'_1 \\ \\ \underset{\sim}{B}'_2 \end{pmatrix} \underset{\sim}{X}_S \qquad S = 1,2,\ldots,n.$$

We see that

$$E(\underset{\sim}{Y}_{S2}) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_q \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

and the covariance of $\underset{\sim}{Y}_{S2}$ is

$$\begin{pmatrix} \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 & \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2 \\ \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 & \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2 \end{pmatrix} .$$

We observe that in the notation of Theorem 2.2.1

$$\underset{\sim}{\mu}_n(\alpha_1 \dots \alpha_q) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_q \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

thus

$$\underset{\sim}{D}^{(2)} = \begin{bmatrix} \underset{\sim}{I} \\ q\times q \\ \underset{\sim}{0} \\ (p-q)\times q \end{bmatrix}$$

in the notation of Theorem 2.2.1. Clearly, the limiting distribution of $Q_n^{(2)}$ is $\chi^2(p,\Delta_2)$ as n → ∞, where

$$Q_n^{(2)} = n \, \bar{\underset{\sim}{Y}}_2' \begin{pmatrix} \underset{\sim}{B}_1' \underset{\sim}{S}_n \underset{\sim}{B}_1 & \underset{\sim}{B}_1' \underset{\sim}{S}_n \underset{\sim}{B}_2 \\ \underset{\sim}{B}_2' \underset{\sim}{S}_n \underset{\sim}{B}_1 & \underset{\sim}{B}_2' \underset{\sim}{S}_n \underset{\sim}{B}_2 \end{pmatrix}^{-1} \bar{\underset{\sim}{Y}}_2$$

and

$$\Delta_2 = \underset{\sim}{\lambda}' \, \underset{\sim}{D}^{(2)'} \begin{bmatrix} \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 & \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2 \\ \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 & \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2 \end{bmatrix}^{-1} \underset{\sim}{D}^{(2)} \underset{\sim}{\lambda} .$$

By Theorem 8.2.1 of Graybill (1969) we have:

$$(3.4.7)\quad \begin{bmatrix} \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 & \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2 \\ \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2 & \underset{\sim}{B}' \underset{\sim}{\Sigma} \underset{\sim}{B}_2 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} [\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 - \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1} \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1]^{-1}, & * \\ *, & * \end{bmatrix}$$

where * indicates other terms in $\underset{\sim}{B}_i' \underset{\sim}{\Sigma} \underset{\sim}{B}_j$ which are eliminated by multiplication by $\underset{\sim}{D}^{(2)'}$. Hence using (3.4.7) we obtain

$$\Delta_2 = \underset{\sim}{\lambda}'[\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 - \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1} \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1]^{-1} \underset{\sim}{\lambda} .$$

Again designating the test with the correct number of parameters as ϕ_1 and letting ϕ_2 be the overfit we obtain the following efficiencies of ϕ_2 relative to ϕ_1:

$$(3.4.8)\quad \text{CARE} = R(q,p,\alpha) \left\{ \frac{|\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1|}{|\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 - \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1}\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)|} \right\}^{1/q} ;$$

$$q \leq p$$

and

$$(3.4.9)\quad \text{TARE} = R(q,p,\alpha) \frac{\operatorname{tr}[\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 - \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1}\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1]^{-1}}{\operatorname{tr}(\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)^{-1}} ;$$

$$q \leq p.$$

3.4.3 Parametric Procedures - General Case of Underfitting

For this situation we consider the same model 3.4.5 and suppose ϕ_1 is again based on the test $Q_n^{(1)}$ of Section 3.4.2 but ϕ_2 is based on $Q_n^{(2)}$, a quadratic form in p variables, where p < q. We define

$\underset{\sim}{Y}_{S2}$ by

$$(3.4.10) \qquad \underset{p\times 1}{\underset{\sim}{Y}_{S2}} = \begin{bmatrix} \underset{\sim}{b}_1' \\ \vdots \\ \underset{\sim}{b}_p' \end{bmatrix} \underset{\sim}{X}_S = \underset{\sim}{B}' \underset{\sim}{X}_S; \quad \text{each } S = 1,\dots,n.$$

Notice that

$$B_1' = \begin{pmatrix} B' \\ b_{p+1}' \\ \vdots \\ b_q' \end{pmatrix} .$$

In the context of Theorem 2.2.1 we have $\mu_{nj}(\alpha_1,\dots,\alpha_q) = \alpha_j$, $j = 1,\dots,p$, and

$$\underset{p\times q}{\underset{\sim}{D}^{(2)}} = [\underset{p\times p}{\underset{\sim}{I}} \; : \; \underset{p\times(q-p)}{\underset{\sim}{0}}] .$$

The conditions of Theorem 2.2.1 are satisfied again and the limiting distribution of $Q_n^{(2)}$ where

$$Q_n^{(2)} = \sqrt{n}\, \bar{\underset{\sim}{Y}}_{S2}' (\underset{\sim}{B}' \underset{\sim}{S}_n \underset{\sim}{B})^{-1} \sqrt{n}\, \bar{\underset{\sim}{Y}}_{S2}$$

is $X^2(p,\Delta_2)$ where $\Delta_2 = \underset{\sim}{\lambda}' \underset{\sim}{D}^{(2)'} (\underset{\sim}{B}' \underset{\sim}{\Sigma} \underset{\sim}{B})^{-1} \underset{\sim}{D}^{(2)} \underset{\sim}{\lambda}$. As pointed out in Chapter II this is a situation in which the test has power α in (q-p) principal directions and hence the generalized Gaussian curvature is zero. We can consider the trace criterion which yields TARE of ϕ_2 with respect to ϕ_1:

$$\text{TARE} = R(q,p,\alpha) = \frac{\text{tr}\, \underset{\sim}{D}^{(2)'} (\underset{\sim}{B}' \underset{\sim}{\Sigma} \underset{\sim}{B})^{-1} \underset{\sim}{D}^{(2)}}{\text{tr}(\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)^{-1}}$$

which reduces to

$$(3.4.11) \qquad \text{TARE} = R(q,p,\alpha)\ \frac{\operatorname{tr}(\underset{\sim}{B}'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B})^{-1}}{\operatorname{tr}(\underset{\sim}{B}_1'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_1)}\ ;\ \text{for } q > p.$$

This quantity will also be numerically evaluated for some specific covariance structures in Section 5.3.

3.5 Nonparametric Procedures

When using nonparametric methods for growth curve analysis two approaches suggest themselves. The first is to estimate each individual's growth curve parameters using some robust procedure and then apply a nonparametric test to these estimates.

The second approach is to use a least squares reduction of each observation as in the previous section and then apply a standard nonparametric test to these least squares estimates. The latter approach is taken in this section.

3.5.1 Data Reduction

The observations $\underset{h\times 1}{\underset{\sim}{X}_S} - \underset{\sim}{\mu}$, $S = 1,\ldots,n$ are assumed to have been selected from an h-variate continuous distribution function $G(\underset{\sim}{X})$, diagonally symmetric about $\underset{\sim}{0}$ where $\mu(t_\ell)$ is represented by one of the models (3.3.1), (3.3.2) or (3.3.3). As we did in the parametric case we shall first consider the special case where the model is linear and evaluate the efficiency by overfitting this linear model. The tests we compare will be rank order tests applied to $\underset{\sim}{Y}_{S1}$ and $\underset{\sim}{Y}_{S2}$ defined by (3.4.1) and (3.4.2) respectively.

Since $\underset{\sim}{X}_S - \underset{\sim}{\mu}$ is assumed to be from G which is continuous and $\underset{\sim}{X}_S - \underset{\sim}{\mu}$ is diagonally symmetric about zero it follows from the symmetry alone that the characteristic function of $\underset{\sim}{X}_S - \underset{\sim}{\mu}$ is real. If B is any h×q matrix of rank $q(q \leq h)$ and we define

$$\underset{\sim}{Z}_S = \underset{\sim}{B}'(\underset{\sim}{X}_S - \underset{\sim}{\mu})$$

then the characteristic function of $\underset{\sim}{Z}_S$, $\psi(\underset{\sim}{t})$, is also real since

$$\psi(\underset{\sim}{t}) = E(e^{i\underset{\sim}{t}'\underset{\sim}{Z}_S}) = E(e^{i\underset{\sim}{t}'\underset{\sim}{B}'(\underset{\sim}{X}_S - \mu)}) = \phi(\underset{\sim}{B}\underset{\sim}{t})$$

where ϕ is the characteristic function of $(\underset{\sim}{X}_S - \underset{\sim}{\mu})$. But ϕ is real for all $\underset{\sim}{t}$ therefore ϕt is real for all $\underset{\sim}{B}\underset{\sim}{t}$, hence ψ is real and $\underset{\sim}{Z}_S$ is diagonally symmetric about $\underset{\sim}{0}$. If we let H be the distribution function of $\underset{\sim}{Z}_S$, then we can think of $\underset{\sim}{Z}_S$ $S = 1,\ldots,n$ as a sample from H and hence letting $\underset{\sim}{Y}_S = \underset{\sim}{B}' \underset{\sim}{X}_S$; $\underset{\sim}{Y}_S$ can be assumed to be from a q-variate continuous distribution function, F, diagonally symmetric about $\underset{\sim}{B}'\underset{\sim}{\mu}$. We are now in a position to apply the multivariate one sample tests as outlined in Puri and Sen (1971). We conclude from this section that we have a reduction of the data to a set of data to which we may apply the nonparametric multivariate one sample test procedures. We shall discuss, as in the parametric case, the problem of overfitting a linear growth curve model but first we briefly summarize the multivariate nonparametric results which we shall use.

3.5.2 One Sample Procedures

Let $\underset{q\times 1}{\underset{\sim}{V}_S}$, $S = 1,\ldots,n$ be a random sample from a diagonally symmetric continuous distribution function F. Under the null hypothesis we assume the location vector of $\underset{\sim}{V}_S$ is $\underset{\sim}{0}$ and for large sample study we have H_n: $\underset{\sim}{\theta}_n = n^{-1/2}\,\underset{\sim}{\Gamma}$. We denote by $\underset{\sim}{R}_n$, the rank matrix of the absolute values of V_{Sk}; i.e.

$$\underset{\sim}{R}_n = \begin{pmatrix} R_{11} & \cdots & R_{1n} \\ \vdots & & \\ R_{q1} & \cdots & R_{qn} \end{pmatrix}$$

where R_{jk} = Rank of $|V_{Sk}|$ among $|V_{1k}|, \ldots, |V_{nk}|$.

Ties can be ignored, at least in theory, by the assumed continuity of F. A transformation of the ranks is made to a matrix of rank scores, $\underset{\sim}{E}_n$, where

$$\underset{\sim}{E}_n = \begin{pmatrix} E_n\, R_{11}, & \cdots & E_n\, R_{1n} \\ \vdots & & \\ E_n\, R_{q1}, & \cdots & E_n\, R_{qn} \end{pmatrix} .$$

The transformation is defined by the score function, J_n, defined by

$$E_{n,\alpha} = J_n(\alpha/(n+1)) \qquad 1 \leq \alpha \leq n$$

where J_n is required to satisfy the following conditions:

I. $\lim_{n \to \infty} J_n(u) = J(u)$ exists for $0 < u < 1$ and $J(u)$ is not constant

II. $$\int_{-\infty}^{\infty} [J_n[\frac{n}{n+1} H_{nj}(X)] - J[\frac{n}{n+1} H_{nj}(X)]]\, dF_{nj}(X) = 0_p(n^{-1/2})$$ $j = 1,2,\ldots,q$

where for each $j = 1,2,\ldots,q$

$$H_{nj}(X) = n^{-1}[\# \text{ of } |V_{Sj}| \leq X;\ S = 1,2,\ldots,n]$$

III. $$\left|\frac{d^i J(u)}{d\, u^i}\right| \leq K[u(1-u)]^{-i - 1/2 + \delta} \qquad i = 0,1 \text{ for some } \delta > 0$$

$$\text{IV.}\quad \lim_{n \to \infty} \int_0^1 [J_n(u) - J(u)]^2 \, du = 0 .$$

Also define $C_{S\alpha}$ to be

$$C_{S\alpha} = \begin{cases} +1 & \quad V_{S\alpha} > 0 \\ -1 \quad \text{if} & \quad V_{S\alpha} < 0 \end{cases} .$$

Under the basic sign invariance structure a class of conditionally distribution free tests for the null hypothesis exists and is characterized in Puri and Sen (1971). We shall use from their text, corollary 4.4.31, for the limiting distribution of $n^{1/2}(\underset{\sim}{T}_n - \gamma)$ through a sequence of alternatives H_n. To summarize their result:

If

$$T_{nj} = \sum_{\alpha=1}^{n} E_{n\, R_{j\alpha}} \, C_{j\alpha} \quad \text{for } j = 1,2,\ldots,q$$

and

$$\underset{\sim}{\gamma} = \frac{1}{2} \begin{pmatrix} \int_0^1 J(u)du \\ \vdots \\ \int_0^1 J(u)du \end{pmatrix}$$

then through H_n the vector $n^{1/2}(\underset{\sim}{T}_n - \underset{\sim}{\gamma})$ has a q-variate normal distribution as n → ∞ with mean

$$\underset{\sim}{C} = \frac{1}{2} \begin{bmatrix} \Gamma_1 \, A_1(F_1) \\ \vdots \\ \Gamma_q \, A_q(F_q) \end{bmatrix} \quad \text{where } A_j(F_j) = \int_{-\infty}^{\infty} \frac{d}{dX} J^*(F_j(X)) dF_j(X)$$

$$j = 1,\ldots,q$$

and covariance matrix

$$\underset{\sim}{\Sigma} = \frac{1}{4}((\sigma_{jj'})) \qquad j,j' = 1,\ldots,q.$$

$$\sigma_{jj'} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J^*(F_j(X))\ J^*(F_{j'}(y))\quad dF_{j'}(X,y)$$

The additional assumptions have been made that $E_{n,\alpha}$ is the expectation of the α^{th} smallest observation from a sample of size n drawn from a distribution $\psi_j(X)$ with

$$\psi_j(X) = \begin{cases} 2\psi_j^*(X) - 1 & X \geq 0 \\ 0 & X < 0 \end{cases}$$

$$\psi_j^*(X) + \psi_j^*(-X) = 1 \qquad \forall \quad X$$

So $J(0) = 0$ and

$$J(u) = \psi_j^{-1}(u) = \psi_j^{*^{-1}}\left(\frac{1+u}{2}\right) = J^*\left(\frac{1+u}{2}\right) \qquad \begin{array}{l} 0 < u < 1 \\ \text{each } j = 1,\ldots,q \end{array}$$

and we assume that

$$f_j(X)\ J^{*'}(F_j(X)) \text{ is bounded as } X \rightarrow \pm\infty.$$

We see that we may construct quadratic forms in $n^{1/2}(\underset{\sim}{T}_n - \underset{\sim}{\gamma})$, with $\hat{\underset{\sim}{\Sigma}}_n$ a consistent estimate for $\underset{\sim}{\Sigma}$ and test H_0. Alternatively, one may use the test statistic (quadratic form) constructed under the permutation group generated by the sign invariance structure and use these test statistics. In large samples these two statistics are power equivalent through the sequence H_n.

3.5.3 Application of the One Sample Test

For our problem we started with an original h-variate continuous distribution function, diagonally symmetric about $\underset{\sim}{B}\ \underset{\sim}{\alpha}$ where $\underset{\sim}{B}$ is the portion of the orthogonal polynomial matrix needed for the model. To test constancy of the growth curve model we must test

$$H_0: \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_q \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

hence we transform from $\underset{\sim}{X}_S$ to $\underset{\sim}{Y}_S$ by (3.3.8).

By our previous discussion $\underset{\sim}{Y}_S$ is diagonally symmetric about $\underset{\sim}{B}_1'\underset{\sim}{B}\begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_q \end{pmatrix}$, $\left(\text{which in the orthogonal polynomial representation is } \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_q \end{pmatrix} \right)$.

We observe that if $\underset{\sim}{B}_1$ and $\underset{\sim}{B}$ are orthogonal polynomial matrices that

$$\underset{q\times h}{\underset{\sim}{B}_1'} \; \underset{h\times(q+1)}{\underset{\sim}{B}} = [\underset{q\times 1}{\underset{\sim}{0}} \; : \; \underset{q\times q}{\underset{\sim}{I}}] \; .$$

We note that a sufficient condition for $\underset{\sim}{Y}_S$ to have a distribution which does not depend on α_0 is that the first column of $\underset{\sim}{B}_1'\underset{\sim}{B}$ is the zero vector. If in this more general situation we let

$$\underset{\sim}{B}_1'\underset{\sim}{B} = [\underset{q\times 1}{\underset{\sim}{0}} \; : \; \underset{\sim}{B}_1'\underset{\sim}{B}_2]$$

then

$$\underset{\sim}{B}_1'\underset{\sim}{B}\,\underset{\sim}{\alpha} = \underset{\sim}{B}_1'\underset{\sim}{B}_2 \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_q \end{pmatrix}.$$

Denoting

$$\underset{\sim}{\theta} = \underset{\sim}{B}_1'\underset{\sim}{B}_2 \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_q \end{pmatrix}$$

we complete our connection between this problem and the multivariate nonparametric one-sample problem by noting that

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_q \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

implies $\underset{\sim}{\theta} = \underset{\sim}{0}$ and the sequence of alternative values

$$\begin{pmatrix} \alpha_{1n} \\ \vdots \\ \alpha_{qn} \end{pmatrix} = n^{-1/2} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_q \end{pmatrix}$$

in terms of $\underset{\sim}{\theta}$ is

$$\theta_n = n^{-1/2} (\underset{\sim}{B}_1'\underset{\sim}{B}_2) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_q \end{pmatrix} = n^{-1/2} \begin{pmatrix} \Gamma_1 \\ \vdots \\ \Gamma_q \end{pmatrix}$$

where

$$\Gamma_j = (\underset{\sim}{B}_1'\underset{\sim}{B}_2)_j \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_q \end{pmatrix} \qquad j = 1,2,\ldots,q.$$

Hence we may test H_0 by testing the hypothesis $\underset{\sim}{\theta} = \underset{\sim}{0}$ and the sequence $\begin{pmatrix} \alpha_{1n} \\ \vdots \\ \alpha_{qn} \end{pmatrix}$ in terms of the test on $\underset{\sim}{\theta}$ is equal to $\underset{\sim}{\theta}_n$. Hence in the notation of Section 3.5.2 if we consider general rank scores applied to the $\underset{\sim}{Y}_S$ then the test statistics constructed would be noncentral X^2 with q degrees of freedom and noncentrality Δ defined by

$$\Delta = \underset{\sim}{c}' \, \underset{\sim}{\Sigma}^{-1} \, \underset{\sim}{c}$$

where

$$\underset{\sim}{c} = \begin{bmatrix} (\underset{\sim}{B}_1'\underset{\sim}{B}_2)_1 \underset{\sim}{\lambda} A_1(F_1) \\ \\ (\underset{\sim}{B}_1'\underset{\sim}{B}_2)_q \underset{\sim}{\lambda} A_q(F_q) \end{bmatrix},$$

$\underset{\sim}{\Sigma}$ is as defined in Section 3.5.2, and F_j is the marginal distribution function. Thus if we let

$$\underset{\sim}{\tau} = ((\tau_{jK})) \qquad j,K = 1,\ldots,q$$

$$\tau_{jK} = \frac{\sigma_{jK}}{A_j(F_j)A_K(F_K)}$$

where σ_{jK} is defined in Section 3.5.2.

Then we see

$$\Delta = \underset{\sim}{\lambda}'(\underset{\sim}{B}_1'\underset{\sim}{B}_2)'\,\underset{\sim}{\tau}^{-1}(\underset{\sim}{B}_1'\underset{\sim}{B}_2)\underset{\sim}{\lambda}$$

and we see that $(\underset{\sim}{B}_1'\underset{\sim}{B}_2)$ corresponds therefore to the $\underset{\sim}{D}^{(i)}$ of Theorem 2.2.1.

The effect of testing the hypothesis with too few or too many parameters i.e., underfitting or overfitting respectively, simply changes $\underset{\sim}{B}_1'\underset{\sim}{B}_2$ to a non-square matrix. For example, if we had too few statistics, say p, $(p < q)$ then $\underset{\sim}{B}_1'\underset{\sim}{B}_2$ is $p\times q$ and hence Δ is positive semidefinite. We now discuss the linear growth curve model and compute the efficiency of a quadratic assumed model to the test with the correct number of parameters.

3.5.4 True Model Linear

If the model (3.3.3) is the true one then the hypothesis to be tested is (3.3.5) against the sequence of alternatives (3.4.3). Letting $\underset{\sim}{Y}_{S1}$ be as in (3.4.1) and $\underset{\sim}{Y}_{S2}$ as in (3.4.2) then $\underset{\sim}{Y}_{S1}$ can be viewed as a sample from F_1 (a univariate continuous distribution function symmetric about α_1) and $\underset{\sim}{Y}_{S2}$ can be viewed as a sample from F_2 (a bivariate continuous distribution function, diagonally symmetric about $\binom{\alpha_1}{0}$). On applying rank scores to $\underset{\sim}{Y}_{S1}$ and $\underset{\sim}{Y}_{S2}$ as described in Section 3.5.2 and letting $Q_n^{(i)}$ be the quadratic form in the rank statistics of $\underset{\sim}{Y}_{Si}$ $(i = 1,2)$, we see that

$$\left.\begin{aligned} \mathcal{L}(Q_n^{(1)}) &\to \chi^2(1,\Delta_1) \\ \mathcal{L}(Q_n^{(2)}) &\to \chi^2(2,\Delta_2) \end{aligned}\right. \qquad \text{as } n \to \infty$$

where

$$\Delta_1 = \lambda_1 \left[\frac{\nu_{1,11}}{A_1^2(F_1)} \right]^{-1} \lambda_1$$

and

$$\Delta_2 = \lambda_1 (1,0)\, \underset{\sim}{\tau}_2^{-1} \binom{1}{0} \lambda_1$$

where

$$\nu_{1,11} = \Big[\int_{-\infty}^{\infty} J^*(F_1(X))\, dF_1(X) \Big]^2 ; \quad A_1(F_1) = \int_{-\infty}^{\infty} \frac{d}{dX} J^{*\prime}(F_1(X))\, dF_1(X)$$

$$\underset{\sim}{\tau}_2 = \left(\left(\frac{\nu_{2,jK}}{A_j(F_{2j})\, A_K(F_{2\,K})} \right) \right) = (\tau_{2,jK}) \qquad j,K = 1,2$$

$F_{2,j}$ is the j^{th} marginal distribution function of F_2.

$$\nu_{2,jK} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J^*(F_{2,j}(X))\, J^*(F_{2,K}(Y))\, dF_{2,jK}(X,Y); \qquad j,K = 1,2.$$

We observe that we have used the same score function for each test procedure and each variate. Obviously we want to use the same score function for the common variates in the competitive tests;

however one may be interested in using a different score function for the extraneous variates (in the case of overfitting). These points are other possibilities which could be explored; however, we continue in our discussion under the assumption of a common score function.

We note that $F_1 = F_{2,1}$ and hence

$$\nu_{1,11} = \nu_{2,11} ,$$

and

$$A_1(F_1) = A_1(F_{2,11}).$$

The ARE of ϕ_2 with respect to ϕ_1 using any of the three criteria, curvature, trace, or local is:

$$\text{(3.5.1)} \qquad \text{ARE} = R(1,2, \ ) \frac{\tau_{2,11}\,\tau_{2,22}}{(\tau_{2,11}\,\tau_{2,22} - \tau_{2,12}\,\tau_{2,21})} .$$

We now consider the general case of overfitting in order to obtain results comparable to the parametric results.

3.5.5 Nonparametric Procedures - General Case of Overfitting

For the general case of overfitting we let ϕ_1 be the test which is based on the rank scores applied to $\underset{q\times 1}{\underset{\sim}{Y}_{S1}}$ defined by (3.4.6), i.e. $\underset{\sim}{Y}_{S1} = \underset{q\times h}{\underset{\sim}{B}_1'}\ \underset{\sim}{X}_S$. In the context of Section 3.5.3 we see that $\underset{\sim}{B}_2 = \underset{\sim}{B}_1$, hence the test based on $Q_n^{(1)}$ (the quadratic form in $T_n^{(1)}$) through the sequence

$$\text{(3.5.5.1)} \qquad H_n: \begin{pmatrix} \alpha_{1n} \\ \vdots \\ \alpha_{qn} \end{pmatrix} = n^{-1/2} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_q \end{pmatrix}$$

would be $\chi^2(q,\Delta_1)$ as n → ∞

with

$$\Delta_1 = \underset{\sim}{\lambda}' \, \underset{\sim}{\tau}_1^{-1} \, \underset{\sim}{\lambda}$$

where $\underset{\sim}{\tau}_1$ is defined as $\underset{\sim}{\tau}$ in Section 3.5.3. On the other hand if we tested for an additional (p-q) parameters $\begin{pmatrix} \alpha_{q+1} \\ \vdots \\ \alpha_p \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ by reducing $\underset{\sim}{X}_S$ to $\underset{\sim}{Y}_{S2}$ by $\underset{\sim}{Y}_{S2} = \begin{bmatrix} \underset{\sim}{B}_1' \\ \\ \underset{\sim}{B}_3' \end{bmatrix} \underset{\sim}{X}_S \quad S = 1,\ldots,n$ where $\underset{\sim}{B}_3' = \begin{pmatrix} \underset{\sim}{b}_{q+1} \\ \vdots \\ \underset{\sim}{b}_p' \end{pmatrix}$.

We then see that $Q_n^{(2)}$, the quadratic form in the rank scores of $\underset{\sim}{Y}_{S2}$, has through the same sequence of alternative hypotheses, a limiting $\chi^2(p,\Delta_2)$ as $n \to \infty$, with

$$\Delta_2 = \underset{1\times q}{\underset{\sim}{\lambda}} \; [\underset{q\times q}{\underset{\sim}{I}} \; : \; \underset{q\times(p-q)}{\underset{\sim}{0}}] \; \underset{\sim}{\tau}_2^{-1} \begin{bmatrix} \underset{\sim}{I} \\ \vdots \\ \underset{\sim}{0} \end{bmatrix} \underset{\sim}{\lambda} \, .$$

We denote the null distribution functions of $\underset{\sim}{Y}_{S1}$ and $\underset{\sim}{Y}_{S2}$ by F_1 and F_2 respectively. The corresponding marginals (bivariate and univariate) are denoted by $F_{1,j}$; $F_{1,jK}$ and $F_{2,\ell}$; $F_{2,\ell m}$, where $j,K = 1,\ldots,q$ while $\ell,m = 1,\ldots,q, q+1,\ldots,p$. It is apparent by the construction of $\underset{\sim}{Y}_{S1}$ and $\underset{\sim}{Y}_{S2}$ that

$$F_{1,j} = F_{2,j}$$

$$\text{for } j,K = 1,2,\ldots,q$$

$$F_{1,jK} = F_{2,jK}$$

and by equality of the score functions we therefore have that $\underset{\sim}{\tau}_1 = \underset{\sim}{\tau}_{2,11}$ where

$$\underset{\sim}{\tau}_2 = \begin{bmatrix} \underset{q\times q}{\underset{\sim}{\tau}_{2,11}} & \underset{q\times(p-q)}{\underset{\sim}{\tau}_{2,12}} \\ \underset{(p-q)\times q}{\underset{\sim}{\tau}_{2,21}} & \underset{(p-q)\times(p-q)}{\underset{\sim}{\tau}_{2,22}} \end{bmatrix}$$

and

$$\underset{\sim}{\tau}_2^{-1} = \begin{bmatrix} (\underset{\sim}{\tau}_{2,11} - \underset{\sim}{\tau}_{2,12}\ \underset{\sim}{\tau}_{2,22}^{-1}\ \underset{\sim}{\tau}_{2,21})^{-1} & , & * \\ * & , & * \end{bmatrix} .$$

Thus the ARE of ϕ_2 with respect to ϕ_1 using the curvature criterion is:

$$(3.5.5.1)\ \text{CARE} = R(q,p,\alpha)\left\{\frac{|\underset{\sim}{\tau}_1|}{|\underset{\sim}{\tau}_{2,11} - \underset{\sim}{\tau}_{2,12}\ \underset{\sim}{\tau}_{2,22}^{-1}\ \underset{\sim}{\tau}_{2,21}|}\right\}^{1/q} \qquad q \leq p$$

Using the trace criterion the ARE is:

$$(3.5.5.2)\ \text{TARE} = R(q,p,\alpha)\ \frac{\text{tr}[\underset{\sim}{\tau}_{2,11} - \underset{\sim}{\tau}_{2,12}\ \underset{\sim}{\tau}_{2,22}^{-1}\ \underset{\sim}{\tau}_{2,21}]^{-1}}{\text{tr}\ \underset{\sim}{\tau}_1^{-1}} \qquad q \leq p$$

3.5.6 Nonparametric Procedures - General Case of Underfitting

We let ϕ_1 be the test in the previous section for the same hypothesis and we let ϕ_2 be the corresponding rank test when we have based the test on only the first p of the $(\alpha_1,\ldots,\alpha_q)'$ where $p < q$. Letting $\underset{p\times 1}{\underset{\sim}{Y}_{S2}} = \underset{p\times h}{\underset{\sim}{B}_4'}\ \underset{h\times 1}{\underset{\sim}{X}_S}$ $\quad S = 1,\ldots,n$. We note $\underset{\sim}{B}_1 = [\underset{\sim}{B}_4, \underset{\sim}{b}_{p+1},\ldots,\underset{\sim}{b}_q]$

From the previous discussions the quadratic forms constructed in the rank order statistics of the $\underset{\sim}{Y}_{S2}$ denoted by $Q_n^{(2)}$ have a limiting

$$\chi^2(p,\Delta_2) \text{ as } n \to \infty ,$$

where

$$\Delta_2 = \underset{\sim}{\lambda}' \begin{bmatrix} \underset{\sim}{I} \\ p\times p \\ \cdots \\ \underset{\sim}{0} \\ (q-p)\times p \end{bmatrix} \underset{\sim 2}{\tau}^{-1} [\underset{p\times p}{\underset{\sim}{I}} : \underset{p\times(q-p)}{\underset{\sim}{0}}] \underset{\sim}{\lambda}$$

The ARE of ϕ_2 with respect to ϕ_1 using the trace criterion is therefore:

$$\text{(3.5.6.1)} \qquad \text{TARE} = R(q,p,\alpha) \frac{\text{tr } \underset{\sim 2}{\tau}^{-1}}{\text{tr } \underset{\sim 1}{\tau}^{-1}} ; \qquad p < q$$

3.5.7 Reduction for Wilcoxon Signed Rank Test

In the notation of Section 3.5.2 the Wilcoxon score or signed rank defines $E_{n\alpha} = J_n(\frac{\alpha}{n+1}) = \frac{\alpha}{n+1}$; $\alpha = 1,\ldots,n$. Then $\underset{\sim}{\tau} = ((\tau_{jK}))$ $j,K = 1,\ldots,q$ is equal to

$$\text{(3.5.7.1)} \qquad \tau_{jK} = \begin{cases} 1/12 \ [\int_{-\infty}^{\infty} f_j(X) \ dF_j(X)]^2 & K = j = 1,\ldots,q \\ \rho_{jK}^g \Big/ \left\{ 12[\int_{-\infty}^{\infty} f_j(X) \ dF_j(X)][\int_{-\infty}^{\infty} f_K(X) dF_K(X)] \right\} & j \neq K = 1,\ldots,q \end{cases}$$

where F is the distribution function of $\underset{\sim S}{Y}$ and ρ_{jK}^g is the grade correlation, i.e.

$$\rho_{jK}^g = 12 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [F_j(X) - \tfrac{1}{2}] \ [F_K(y) - \tfrac{1}{2}] \ dF_{jK}(X,y) \qquad j,K = 1,\ldots,q .$$

Further if G is multivariate normal then F is multivariate normal and

$$\text{(3.5.7.2)} \qquad \int_{-\infty}^{\infty} f_j(X) \ dF_j(X) = [2 \sqrt{\pi} \ \gamma_{jj}^{1/2} \ ]^{-1} \qquad j = 1,\ldots,q$$

where γ_{jj} is the variance of the j^{th} component of $\underset{\sim}{Y}_S$.

Denoting ρ_{jK} as the correlation of the j^{th} and K^{th} variates of $\underset{\sim}{Y}_S$ we have

$$\rho^g_{jK} = \frac{6}{\pi} \sin^{-1} \left(\frac{\rho_{jK}}{2}\right) \qquad j,K = 1,\ldots,q.$$

We note that if we assume that G has a covariance matrix $\underset{\sim}{\Sigma}$ then $\gamma_{jj} = \underset{\sim}{b}_j' \underset{\sim}{\Sigma} \underset{\sim}{b}_j$. We see that

$$(3.5.7.3) \quad \tau_{jK} = \begin{cases} 1/12\{2\sqrt{\pi}\,[\underset{\sim}{b}_j' \underset{\sim}{\Sigma} \underset{\sim}{b}_j]^{1/2}\}^{-2} & j = K = 1,\ldots,q \\ \frac{1}{12}[\frac{6}{\pi} \sin^{-1}(\frac{\rho_{jK}}{2})][2\sqrt{\pi}(\underset{\sim}{b}_j' \underset{\sim}{\Sigma} \underset{\sim}{b}_2)^{1/2}][2\sqrt{\pi}(\underset{\sim}{b}_K' \underset{\sim}{\Sigma} \underset{\sim}{b}_K)^{1/2}] & \\ & j \neq K = 1,\ldots,q. \end{cases}$$

In general

$$(3.5.7.4) \quad \tau_{jK} = \begin{cases} 1/12(4\pi\, \underset{\sim}{b}_j' \underset{\sim}{\Sigma} \underset{\sim}{b}_j)^{-1}; & j = K = 1,\ldots,q \\ \frac{1}{12}(4\pi(\underset{\sim}{b}_j' \underset{\sim}{\Sigma} \underset{\sim}{b}_j)^{1/2}(\underset{\sim}{b}_K' \underset{\sim}{\Sigma} \underset{\sim}{b}_K)^{1/2})[\frac{6}{\pi} \sin^{-1}(\frac{\rho_{jK}}{2})]; & \\ & j \neq K = 1,2,\ldots,q \end{cases}$$

where

$$\rho_{jK} = \underset{\sim}{b}_j' \underset{\sim}{\Sigma} \underset{\sim}{b}_K / (\underset{\sim}{b}_j' \underset{\sim}{\Sigma} \underset{\sim}{b}_j)^{1/2} (\underset{\sim}{b}_K' \underset{\sim}{\Sigma} \underset{\sim}{b}_K)^{1/2} \qquad j,K = 1,2,\ldots,q.$$

Obviously, the additional subscript (1 or 2) needs to be added for the comparison of the two tests.

3.6 Comparison of Nonparametric to Parametric Procedures

In this section we consider a comparison of the nonparametric procedures available for the one sample problem to the standard parametric procedure. The results in this section have applications to the general one sample shift problem so we shall first present results for the general one sample problem. We then discuss the extension of these results to the comparison of the two procedures when applied to the growth curve problem when

a) the model is correctly specified;

b) the model is overspecified; and

c) the model is underspecified.

3.6.1 <u>One Sample Location Problem</u>

In this case we shall let ϕ_1 be the test based on the parametric (T_n^2) test and ϕ_2 be the nonparametric (rank scores) procedure. We can have either of the two measures of ARE in this case, the curvature or the trace; i.e.

$$\text{CARE} = \{|\underset{\sim}{\Sigma}|/|\underset{\sim}{\tau}|\}^{1/q} \tag{3.6.1}$$

or

$$\text{TARE} = \text{tr}\ \underset{\sim}{\tau}^{-1}/\text{tr}\ \underset{\sim}{\Sigma}^{-1}. \tag{3.6.2}$$

We shall want to place bounds on (3.6.1) and (3.6.2) over suitable subclasses of the entire class of distribution functions, $\mathcal{F}$. The class, $\mathcal{F}$, for the one sample location problem, is the class of q-variate absolutely continuous distribution functions, diagonally symmetric about its median, $\underset{\sim}{\theta}$, and possessing moments of order $2 + \delta$ for some $\delta > 0$.

Bounds for (3.6.1) are well known since (3.6.1) is nothing more than the efficiency of the rank scores <u>estimator</u> to the least squares <u>estimator</u> using the Wilk's criterion of asymptotic generalized variance; this fact was noted by Bickel (1965). Bounds on (3.6.1) over subclasses in $\mathcal{F}$ are presented in Puri and Sen (1971) and Bickel (1964).

We summarize these results briefly:

If ϕ_2 is the normal scores estimator and

a) F is q-variate normal then CARE = 1,

b) $F(\underset{\sim}{X}) = \prod_{j=1}^{q} F_j(X_j)$ then CARE $\geq$ 1.

If ϕ_2 is the Wilcoxon scores estimator then

(3.6.3) a) CARE $\geq .864 \, [\, |\rho_{jK}| \, / \, |\rho^{g}_{jK}| \,]^{1/q}$ for all F in $\mathcal{F}$;

b) If F is q-variate normal then

(3.6.4) $$\text{CARE} = \frac{3}{\pi}[\, |\rho_{jK}| \, / \, |\tfrac{6}{\pi} \sin^{-1}(\rho_{jK}/2)| \,]^{1/q}$$

c) If F is bivariate normal then

$$\text{CARE} = \frac{3}{\pi}[(1-\rho^2) \mid (1 - \frac{36}{\pi^2} (\sin^{-1}(\rho/2))^2]^{1/2}$$

and if we find the maximum and minimum with respect to $\rho(-1 \leq \rho \leq 1)$ we find $.91 \leq$ CARE $\leq .95$ where the lower bound is reached when $|\rho| \to 1$ and .95 is the value at $\rho = 0$.

d) If F is pairwise independent then (3.6.3) has a lower bound of .864 and (3.6.4) has a lower bound of $3/\pi$.

We now shall find bounds similar to the above bounds when we use the trace criterion of efficiency.

Let us consider the normal scores test, i.e. ϕ_2 is the normal score procedure. We first note that if F is q-variate normal then $\underset{\sim}{\Sigma} = \underset{\sim}{\tau}$ and therefore TARE = 1.

Theorem 3.6.1. If $F(\underset{\sim}{X}) = \prod_{j=1}^{q} F_j(X_j)$ and if $F_j(X_j)$ has density $f_j(X_j)$ with finite variance σ_j^2 and if $\frac{d}{dX}\,\Phi^{-1}(F_j(X))$ is bounded as $X \to \pm\infty$ then TARE ≥ 1.

Proof:

For all $j,K = 1,2,\ldots,q$ we know that

$$\tau_{jK} = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Phi^{-1}(F_j(X))\ \Phi^{-1}(F_K(y))dF_{jK}(X,y)}{(\int_{-\infty}^{\infty} \frac{d}{dX}\,\Phi^{-1}(F_j(X))\ dF_j(X))\ (\int_{-\infty}^{\infty} \frac{d}{dy}\,\Phi^{-1}(F_K(y))\ dF_K(y))}$$

but for $j \neq K$ the numerator of τ_{jK} is zero by the diagonal symmetry of $F(\underset{\sim}{X})$ and the symmetric nature of the score function Φ^{-1}.

Let

$$A(F_j) = \int_{-\infty}^{\infty} \frac{d}{dX}\,\Phi^{-1}(F_j(X))\ dF_j(X); \qquad j = 1,2,\ldots,q.$$

Therefore

$$\tau_{jK} = \begin{cases} 0 & j \neq K \\ \int_{-\infty}^{\infty} [\Phi^{-1}(F_j(X))]^2\ dF_j(X)/A^2(F_j) & j = K = 1,2,\ldots,q. \end{cases}$$

But by the boundedness assumption and general properties of the normal distribution it has been shown [see e.g. Gastwirth and Wolf (1968)] that

$$A^2(F_j) \geq 1/\sigma_j{}^2 \qquad j = 1,2,\ldots,q.$$

We know that

$$\int_{-\infty}^{\infty} [\Phi^{-1}(F_j(X))]^2\ dF_j(X) = 1.$$

Therefore

$$\underset{\sim}{\tau}^{-1} = \text{diag}\ (A^2(F_1),\ \ldots,\ A^2(F_q))$$

$$\Sigma^{-1} = \text{diag}\ (1/\sigma_1^{\,2}, \ldots, 1/\sigma_q^{\,2})$$

and the result follows.

Q.E.D.

We note that if the distribution function, F, has a diagonal covariance matrix then the above theorem still follows since each diagonal element of $\underset{\sim}{\tau}^{-1}$ is greater than or equal to $1/A^2(F_i)$. Hence the result of the theorem is valid for a wider class of distributions. If ϕ_2 is the Wilcoxon score then from Puri and Sen (1971, p. 175) we have:

$$\tau_{jK} = \begin{cases} 1/12\left(\int_{-\infty}^{\infty} f_j(X)\ dF_j(X)\right)^2; & j = K = 1,2,\ldots,q \\ \rho_{jK}^{g}/12\left(\int_{-\infty}^{\infty} f_j(X)\ dF_j(X)\right)\left(\int_{-\infty}^{\infty} f_K(X)\ dF_K(X)\right), & j \neq K, \end{cases}$$

where

$$\rho_{jK}^{g} = 12 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [F_j(X) - 1/2]\ [F_K(y) - 1/2]\ dF_{jK}(X,y).$$

Let

$$A(F_j) = \int_{-\infty}^{\infty} f_j(X)\ dF_j(X).$$

In general tr $\underset{\sim}{\tau}^{-1}$/tr $\underset{\sim}{\Sigma}^{-1}$ is not easily evaluated for arbitrary $F(\underset{\sim}{X})$. We consider some special cases.

<u>Theorem 3.6.2</u>. If $\underset{\sim}{X}$ is an interchangeable random vector then

$$(3.6.5) \qquad \text{TARE} = \frac{12a^2\sigma^2(1-\rho)\ (1 + (q-2)\rho^g)\ (1 + (q-1)\rho)}{(1 - \rho^g)\ (1 + (q-2)\rho)\ (1 + (q-1)\rho^g)}$$

where $a = A(F_j) = \int_{-\infty}^{\infty} f_j(X)\ dF_j(X) = \int_{-\infty}^{\infty} f(X)\ dF(X).$

Proof:

Since $$\underset{\sim}{\Sigma} = \sigma^2 \begin{bmatrix} 1 & \rho & \cdots & \rho \\ & 1 & \cdots & \\ \vdots & & & \\ \rho & & & 1 \end{bmatrix}$$

and $$\rho^g_{jK} = \rho^g \qquad j \neq K = 1,2,\ldots,q$$

then

$$\tau_{jK} = \begin{matrix} \frac{1}{12a^2} & & j = K \\ \frac{\rho^g}{12a^2} & & j \neq K \end{matrix}$$

$$\underset{\sim}{\tau} = \frac{1}{12a^2} \begin{pmatrix} 1 & \rho^g & \cdots & \rho^g \\ & & & \vdots \\ & \rho^g & & 1 \end{pmatrix}$$

Denote $\underset{q\times q}{\underset{\sim}{J}} = \underset{q\times q}{\underset{\sim}{1}}$ then we represent $\underset{\sim}{\Sigma}$ & $\underset{\sim}{\tau}$ and their inverses are given by the following four equations:

$$\underset{\sim}{\Sigma} = [\sigma^2(1-\rho)\,\underset{\sim}{I} + \sigma^2\,\rho J]$$

$$\underset{\sim}{\tau} = [\frac{1}{12a^2}\,(1-\rho^g)\,\underset{\sim}{I} + \frac{\rho^g}{12a^2}\,\underset{\sim}{J}]$$

$$\underset{\sim}{\Sigma}^{-1} = (\sigma^2(1-\rho))^{-1}\,[\underset{\sim}{I} - \sigma^2\rho/(\sigma^2 + (q-1)\sigma^2\rho)\,\underset{\sim}{J}]$$

$$\underset{\sim}{\tau}^{-1} = (\frac{1}{12a^2}(1-\rho^g))^{-1}[\underset{\sim}{I} - \frac{\rho^g}{12a^2} \,/\, (\frac{1}{12a^2} + (q-1)\,\frac{\rho^g}{12a^2})\,\underset{\sim}{J}]$$

hence the result follows.

Q.E.D.

Theorem 3.6.3. If $F(\underset{\sim}{X})$ has covariance matrix $\underset{\sim}{\Sigma}$ which is diagonal then

$$\text{TARE} \geq .864.$$

Proof:

We know that

$$\text{tr } \underset{\sim}{\tau}^{-1} \geq 12 \sum_{j=1}^{q} \left(\int_{-\infty}^{\infty} f_j(X) \, dF_j(X) \right)^2 = 12 \sum_{j=1}^{q} A^2(F_j)$$

and

$$\text{tr } \underset{\sim}{\Sigma}^{-1} = \sum_{j=1}^{q} 1/\sigma_j^2$$

therefore

$$\text{TARE} \geq 12 \sum_{j=1}^{q} A^2(F_j) / \left(\sum_{j=1}^{q} 1/\sigma_j^2 \right) = \frac{\sum_{j=1}^{q} 12 \, \sigma_j^2 \, A^2(F_j)/\sigma_j^2}{\sum_{j=1}^{q} 1/\sigma_j^2}$$

but $12\sigma_j^2 A^2(F_j) \geq .864$ [Hodges and Lehmann (1956)].

Thus, $\text{TARE} \geq .864$.

Q.E.D.

Theorem 3.6.4. If $F(\underset{\sim}{X})$ is q-variate normal with pairwise independent coordinates

$$\text{TARE} \geq 3/\pi \; .$$

Proof:

When F is normal we have

$$A(F_j) = (2 \sqrt{\pi} \, \sigma_j)^{-1}$$

and

$$\rho_{jK}^{g} = \frac{6}{\pi} \sin^{-1} (\rho_{jK}/2)$$

hence we have the following

$$\underset{\sim}{\tau} = \frac{\pi}{3} \begin{pmatrix} \sigma_1^2 & \frac{6\sigma_1\sigma_2}{\pi} \sin^{-1}(\rho_{12/2}) & \cdots & \frac{6\sigma_1\sigma_2}{\pi} \sin^{-1}(\rho_{1q/2}) \\ & \cdot & & \\ & & \cdot & \\ \text{symmetric} & & \cdot & \\ & & \cdot & \sigma_q^2 \end{pmatrix} .$$

Pairwise independence implies $\rho_{jK} = 0$ thus $\rho^g_{jK} = 0$. Therefore,

$$\underset{\sim}{\tau}^{-1} = \frac{3}{\pi} \operatorname{diag}(1/\sigma_1^2, \ldots, 1/\sigma_q^2)$$

$$\underset{\sim}{\Sigma}^{-1} = \operatorname{diag}(1/\sigma_1^2, \ldots, 1/\sigma_q^2)$$

and the result follows.

Q.E.D.

We now establish a result for the trace criterion which parallels the curvature result when F is q-variate normal ($q \geq 3$).

Theorem 3.6.5. For $q > 2$ then

$$\inf_{\phi \in Q} \text{TARE} = 0$$

where Q is the class of nonsingular q-variate normal distribution.

Proof:

As in Bickel (1964) it is sufficient to establish the result for $q = 3$. Consider the covariance matrix:

$$\underset{\sim}{\Sigma}(\alpha) = \sigma^2 \begin{pmatrix} 1 & 0 & (1-\alpha)/\sqrt{2} \\ 0 & 1 & (1-\alpha)/\sqrt{2} \\ (1-\alpha)/\sqrt{2} & (1-\alpha)/\sqrt{2} & 1 \end{pmatrix} .$$

Hence

$$\underset{\sim}{\Sigma}^{-1}(\alpha) = (\sigma^2(1-(1-\alpha)^2))^{-1} \begin{pmatrix} 1-\frac{(1-\alpha)^2}{2}, & \frac{(1-\alpha)^2}{2}, & -\frac{(1-\alpha)}{\sqrt{2}} \\ \frac{(1-\alpha)^2}{2}, & 1-\frac{(1-\alpha)^2}{2}, & -\frac{(1-\alpha)}{\sqrt{2}} \\ -\frac{(1-\alpha)}{\sqrt{2}}, & -\frac{(1-\alpha)}{\sqrt{2}}, & 1 \end{pmatrix}$$

$\text{tr}\ \underset{\sim}{\Sigma}^{-1}(\alpha) = [\sigma^2(1-(1-\alpha)^2)]^{-1}\ [3-(1-\alpha)^2]$ and we see as $\alpha \to 0$ that $\text{tr}\ \underset{\sim}{\Sigma}^{-1}(\alpha) \to \infty$. On the other hand,

$$\underset{\sim}{\tau}(\alpha) = \frac{\pi}{6}\sigma^2 \begin{pmatrix} 1 & 0 & \frac{6}{\pi}\sin^{-1}\left(\frac{1-\alpha}{\sqrt{2}}\right) \\ 0 & 1 & \frac{6}{\pi}\sin^{-1}\left(\frac{1-\alpha}{\sqrt{2}}\right) \\ \frac{6}{\pi}\sin^{-1}\left(\frac{1-\alpha}{\sqrt{2}}\right) & \frac{6}{\pi}\sin^{-1}\left(\frac{1-\alpha}{\sqrt{2}}\right) & 1 \end{pmatrix}$$

$$\underset{\sim}{\tau}^{-1}(\alpha) = \frac{3}{\pi}\sigma^2(1-2b^2)^{-1} \begin{pmatrix} 1-b^2 & b^2 & -b \\ b^2 & 1-b^2 & -b \\ -b & -b & 1 \end{pmatrix}$$

where $$b = \frac{6}{\pi}\sin^{-1}\left(\frac{1-\alpha}{\sqrt{2}}\right).$$

Clearly we have

$$\text{tr}\ \underset{\sim}{\tau}^{-1}(\alpha) = \frac{3}{\pi}\sigma^2(1-2b^2)^{-1}\ (3-2b^2).$$

Notice that $\frac{3\sigma^2}{\pi}(1-2b^2)^{-1}\ (3-2b^2) \to \sigma^2\ .42$

since

$$b^2 = \left[\frac{6}{\pi}\sin^{-1}\left(\frac{1-\alpha}{\sqrt{2}}\right)\right]^2 = \frac{36}{\pi^2}\left(\sin^{-1}\left(\frac{1-\alpha}{\sqrt{2}}\right)\right)^2 \to 2.25 \text{ as } \alpha \to 0,$$

therefore $(1-2b^2)^{-1} \to -.29$

$$(3 - 2b^2) \to -1.50.$$

So we see that tr $\underset{\sim}{\tau}^{-1}(\alpha) \to .42\,\sigma^2 \neq \infty$ as $\alpha \to 0$. So TARE $\to 0$ as $\alpha \to 0$ and the theorem follows.

Q.E.D.

Hence the same type of degeneracy occurs using the trace criterion as occurs with the curvature and the smallest root criterion. Let us now study the bivariate normal distribution function and derive bounds on the TARE.

Theorem 3.6.6. If F is bivariate normal with mean $\binom{0}{0}$ and

$$\underset{\sim}{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\,\sigma_1\sigma_2 \\ \rho\,\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$ and positive definite then

a) TARE is independent of σ_1^2 and σ_2^2;

b) $.87 \leq \text{TARE} \leq .95$ where the lower bound is reached as $|\rho| \to 1$ and the upper bound is reached at $\rho = 0$.

Proof:

Since F is bivariate normal we know

$$\underset{\sim}{\tau} = \frac{\pi}{3}\begin{pmatrix} \sigma_1^2 & \frac{6\sigma_1\sigma_2}{\pi}\sin^{-1}(\rho/2) \\ \frac{6\sigma_1\sigma_2}{\pi}\sin^{-1}(\rho/2) & \sigma_2^2 \end{pmatrix}.$$

In the 2×2 case we know that for a nonsingular matrix $\underset{\sim}{A}$ that

$$\operatorname{tr} \underset{\sim}{A}^{-1} = \operatorname{tr} \underset{\sim}{A}/|A| \ .$$

Hence,

$$\operatorname{tr} \underset{\sim}{\tau}^{-1} = \frac{\pi}{3}(\sigma_1^2 + \sigma_2^2)/\sigma_1^2\ \sigma_2^2\ \frac{\pi^2}{9}\ (1 - \frac{36}{\pi^2}\ (\sin^{-1}(\rho/2))^2)$$

$$\operatorname{tr} \underset{\sim}{\Sigma}^{-1} = (\sigma_1^2 + \sigma_2^2)/\sigma_1^2\ \sigma_2^2\ (1 - \rho^2)$$

and

$$\text{TARE} = \operatorname{tr} \underset{\sim}{\tau}^{-1}/\operatorname{tr} \underset{\sim}{\Sigma}^{-1} = \frac{3}{\pi}(1 - \frac{36}{\pi^2}(\sin^{-1}\ (\rho/2))^2)^{-1}\ (1 - \rho^2)$$

which proves part a of the theorem. We first notice that the TARE is symmetric about $\rho = 0$ since if we denote $h(\rho)$ = TARE we see

$$h(-\rho) = \frac{3}{\pi}(1 - \frac{36}{\pi^2}(\sin^{-1}(-\rho/2))^2)^{-1}\ (1 - \rho^2) = h(\rho).$$

For $0 \leq \rho \leq 1, \sin^{-1}(\rho/2)$ is increasing, therefore, $[\sin^{-1}(\rho/2)]^2$ is increasing, similarly, for $1 \leq \rho \leq 0, \sin^{-1}(\rho/2)$ is decreasing, and therefore, $(\sin^{-1}(\rho/2))^2$ is increasing, thus $[\sin^{-1}(\rho/2)]^2$ is increasing as $|\rho|$ increases. Bickel (1964) has shown that

$$\rho^g = 3(1 - \frac{2}{\pi}\cos^{-1}\ \rho/2)$$

hence

$$\text{TARE} = \frac{3}{\pi}\ (1 - \rho^2)(1 - 9(1 - \frac{2}{\pi}\cos^{-1}\ \rho/2)^2)^{-1} .$$

Further Bickel (1964, p. 1087) demonstrated that this function has a maximum at $\rho = 0$ and is monotonically decreasing in $|\rho|$. Hence the maximum of the TARE is at $\rho = 0$ which is $3/\pi$ or .95. For the lower bound consider the monotonic transformation

$$u = \sin^{-1}\ (\rho/2)$$

$$2 \sin u = \rho,$$

thus

$$\text{TARE} = \frac{3}{\pi}\,\frac{(1 - 4\sin^2 u)}{(1 - \frac{36}{\pi^2}u^2)} ,$$

and

$$\lim_{|\rho| \to 1} \frac{3}{\pi}\,\frac{(1-\rho^2)}{(1 - \frac{36}{\pi^2}(\sin^{-1}(\rho/2)^2)} = \lim_{|\rho| \to \frac{\pi}{6}} \frac{3}{\pi}\,\frac{(1-\rho^2)}{(1 - \frac{36}{\pi^2}u^2)} .$$

Applying L'Hospital's rule we obtain

$$\lim_{|u| \to \frac{\pi}{6}} \frac{3}{\pi}\,\frac{(-8\sin u \cos u)}{-2\,\frac{(36)}{\pi^2}\,u} = \frac{3}{\pi}\,\frac{8\sin\frac{\pi}{6}\cos\frac{\pi}{6}}{2(\frac{6}{\pi})}$$

$$= \frac{\sqrt{3}}{2} = .866 = .87.$$

Q.E.D.

We see the trace leads to slightly wider bounds than the curvature criterion does; but we have the same basic result that the efficiency decreases as $|\rho|$ increases. In any case we see that using curvature or the trace is likely to produce similar results for the efficiency for an underlying bivariate normal distribution.

3.6.2 Growth Curve Model Correctly Specified

In the notation of Chapter II we have $t_1 = t_2 = q$ and if ϕ_2 is the nonparametric procedure and ϕ_1 the parametric procedure then the efficiencies of ϕ_2 relative to ϕ_1 are:

$$\text{CARE} = (|\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1| / |\underset{\sim}{\tau}|)^{1/q} \tag{3.6.6}$$

$$\text{TARE} = \text{tr}\,\underset{\sim}{\tau}^{-1}/\text{tr}(\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)^{-1} \text{ where } \underset{\sim}{B}_1' = [\underset{\sim}{b}_1,\ldots,\underset{\sim}{b}_q]. \tag{3.6.7}$$

$\underset{\sim}{\Sigma}$ is the covariance matrix of the original observations from the parent distribution function G, an h-variate absolutely continuous distribution function diagonally symmetric about zero with moments of order $2 + \delta$ for some $\delta > 0$. F is, in this problem, the distribution function of the reduced quantities $\underset{\sim}{Y}_S$, defined by $\underset{\sim}{B}_1' \underset{\sim}{X}_S$. The least squares procedure and the rank procedures are applied to the $\underset{\sim}{Y}_S$. Practically speaking, the bounds attained in the general one sample problem apply to this problem as well. The only point of interest is to determine what set of circumstances in the growth curve problem lead to these bounds. Consider the normal scores procedure. If F is q-variate normal then the efficiency using either criterion is 1; if F has pairwise independent components it follows that

$$\underset{\sim}{b}_j' \underset{\sim}{\Sigma} \underset{\sim}{b}_K = 0 \quad \forall \quad j \neq K.$$

One case where this happens is when $\underset{\sim}{\Sigma} = \sigma^2 \underset{\sim}{I}$, hence the CARE and TARE are bounded below by unity.

Turning now to the Wilcoxon scores we see that if F has pairwise independent components then the efficiency (trace and curvature) is bounded below by .864 and if F is normal the bound is $3/\pi$, again pairwise independence necessitates that $\underset{\sim}{b}_j' \underset{\sim}{\Sigma} \underset{\sim}{b}_K = 0 \ \forall \ j \neq K$. If F is bivariate normal then using curvature efficiency the lower bound of .91 is reached when

$$\left| (\underset{\sim}{b}_1' \underset{\sim}{\Sigma} \underset{\sim}{b}_2)/(\underset{\sim}{b}_1' \underset{\sim}{\Sigma} \underset{\sim}{b}_1)^{1/2}(\underset{\sim}{b}_2' \underset{\sim}{\Sigma} \underset{\sim}{b}_2)^{1/2} \right| \to 1$$

while the trace criterion achieves a lower bound of .87 for the same limit. The upper bound is attained for both when $\rho = 0$. One case where $\rho = 0$ is $\underset{\sim}{\Sigma} = \sigma^2 \underset{\sim}{I}$, so the upper bound is attained for this situation. The lower bound for the growth curve model cannot in general be reached for the bivariate case (i.e. linear and quadratic). To see this point more clearly, we know there exists a nonsingular $\underset{\sim}{D}$ such that

$$(\underset{\sim}{D}^{-1})' \underset{\sim}{\Sigma} \underset{\sim}{D}^{-1} = I$$

therefore

$$\underset{\sim}{\Sigma} = \underset{\sim}{D}'\underset{\sim}{D}$$

let

$$\underset{\sim}{a}_i = \underset{\sim}{D}\,\underset{\sim}{b}_i \qquad i = 1,2$$

thus

$$\rho = \underset{\sim}{a}_1'\underset{\sim}{a}_2/(\underset{\sim}{a}_1'\underset{\sim}{a}_1)^{1/2}\ (\underset{\sim}{a}_2'\underset{\sim}{a}_2)^{1/2}$$

and this can be one only if there is a nonzero constant c such that

$$\underset{\sim}{a}_1 = \underset{\sim}{c}\,\underset{\sim}{a}_2$$

which implies that

$$\underset{\sim}{b}_1 = \underset{\sim}{c}\,\underset{\sim}{b}_2$$

which we know cannot happen since for h time points, the elements of $\underset{\sim}{b}_1$ are monotonically increasing while the elements of $\underset{\sim}{b}_2$ are not monotonic. Thus the correlation coefficient of the linear and quadratic cannot be 1 so the lower bound in general cannot be reached.

3.6.3 Growth Curve Model Overspecified

In the notation of the previous sections we have t_1 and t_2 both equal to t which is greater than q. We denote

$$\underset{\sim}{B}_1 = [\underset{\sim}{b}_1, \ldots, \underset{\sim}{b}_q] \tag{3.6.8}$$

$$\underset{\sim}{B}_2 = [\underset{\sim}{b}_{q+1}, \ldots, \underset{\sim}{b}_t]$$

where the underlying model is given by $\underset{\sim}{B}_1 \underset{\sim}{\alpha}$ and $\underset{\sim}{B}_2$ corresponds to the orthonormal polynomials we have overfit. The rank scores and least scores procedures are applied to the

$$\underset{\sim}{Y}_S = \begin{pmatrix} \underset{\sim}{B}_1' \\ \underset{\sim}{B}_2' \end{pmatrix} \underset{\sim}{X}_S \; ; \qquad S = 1,2,\ldots,n.$$

Under the assumptions of Theorem 2.2.1 the quadratic forms constructed in the rank statistics have a chi-square distribution with t degrees of freedom and noncentrality parameter

$$\underset{1\times q}{\underset{\sim}{\lambda}'} \; [\underset{q\times q}{\underset{\sim}{I}} \; : \underset{q\times(t-q)}{\underset{\sim}{0}} \;] \; \underset{\sim}{\tau}^{-1} \begin{bmatrix} \underset{\sim}{I} \\ \underset{\sim}{0} \end{bmatrix} \underset{\sim}{\lambda}$$

through H_N where

$$H_N\colon \begin{pmatrix} \alpha_{1N} \\ \vdots \\ \alpha_{qN} \end{pmatrix} = N^{-1/2} \underset{\sim}{\lambda} \; .$$

Each term of $\underset{\sim}{\tau}$ is the corresponding covariance of the rank scores of the $\underset{\sim}{Y}_S$ divided by the $A(F_j)$ $A(F_K)$ defined earlier in this section. The corresponding parametric test would also have chi-square distribution with t degrees of freedom and noncentrality

$$\underset{\sim}{\lambda}'[\underset{\sim}{I}:\underset{\sim}{0}] \left[\begin{pmatrix} \underset{\sim}{B}_1' \\ \underset{\sim}{B}_2' \end{pmatrix} \underset{\sim}{\Sigma} \; (\underset{\sim}{B}_1, \underset{\sim}{B}_2) \right]^{-1} \begin{bmatrix} \underset{\sim}{I} \\ \underset{\sim}{0} \end{bmatrix} \underset{\sim}{\lambda}$$

through H_N. The efficiency of ϕ_2 relative to ϕ_1 is therefore given by:

$$\text{(3.6.9)} \quad \text{CARE} = \{ |\underset{\sim}{B}_1'\underset{\sim}{\Sigma}\underset{\sim}{B}_1 - \underset{\sim}{B}_1'\underset{\sim}{\Sigma}\underset{\sim}{B}_2(\underset{\sim}{B}_2'\underset{\sim}{\Sigma}\underset{\sim}{B}_2)^{-1}\underset{\sim}{B}_2'\underset{\sim}{\Sigma}\underset{\sim}{B}_1| / |\underset{\sim}{\tau}_{11} - \underset{\sim}{\tau}_{12}\underset{\sim}{\tau}_{22}^{-1}\underset{\sim}{\tau}_{21}| \}^{1/q}$$

and

$$\text{(3.6.10)} \quad \text{TARE} = \text{tr}(\underset{\sim}{\tau}_{11} - \underset{\sim}{\tau}_{12}\underset{\sim}{\tau}_{22}^{-1}\underset{\sim}{\tau}_{21})^{-1} / \text{tr}(\underset{\sim}{B}_1'\underset{\sim}{\Sigma}\underset{\sim}{B}_1 - \underset{\sim}{B}_1'\underset{\sim}{\Sigma}\underset{\sim}{B}_2(\underset{\sim}{B}_2'\underset{\sim}{\Sigma}\underset{\sim}{B}_2)^{-1}\underset{\sim}{B}_2'\underset{\sim}{\Sigma}\underset{\sim}{B}_1)^{-1}$$

where $\underset{\sim}{\tau}_{ij}$ is the portion of $\underset{\sim}{\tau}$ due to the covariance of the rank scores applied to $\underset{\sim}{B}_i' \underset{\sim}{X}_S$ and $\underset{\sim}{B}_j' \underset{\sim}{X}_S$ for $i,j = 1,2$.

3.6.4 Growth Curve Model Underspecified

In this case $t < q$ and we let $\underset{h \times t}{\underset{\sim}{B}}$ be the first t rows of $\underset{\sim}{B}_1$. The least squares procedure and nonparametric procedure are applied to

$$\underset{\sim}{Y}_S = \underset{\sim}{B}' \underset{\sim}{X}_S \qquad S = 1, 2, \ldots, n.$$

The curvature criterion cannot be applied in this case so we consider the trace criterion which yields the efficiency of ϕ_2 relative to ϕ_1 as:

$$\text{TARE} = \text{tr} \begin{bmatrix} \underset{\sim}{I}_{t \times t} \\ \underset{\sim}{0} \end{bmatrix} \underset{\sim}{\tau}^{-1} [\underset{\sim}{I} : \underset{\sim}{0}] / \text{tr} \begin{bmatrix} \underset{\sim}{I} \\ \underset{\sim}{0} \end{bmatrix} (\underset{\sim}{B}' \underset{\sim}{\Sigma} \underset{\sim}{B})^{-1} [\underset{\sim}{I} : \underset{\sim}{0}] .$$

This reduces to:

$$\text{TARE} = \text{tr}\, \underset{\sim}{\tau}^{-1} / \text{tr} (\underset{\sim}{B}' \underset{\sim}{\Sigma} \underset{\sim}{B})^{-1} . \tag{3.6.11}$$

3.7 Covariance Adjustment

In this section we investigate the effect of the use of the higher order polynomial terms as covariables on the TARE and CARE. We compute the efficiencies of the procedure with covariance adjustment with respect to the procedure without covariance adjustment when the model has been a) correctly specified, b) overspecified, and c) underspecified. In cases a and c we are able to show that you always gain by covariance adjustment.

3.7.1 Model Correctly Specified

The hypothesis of interest in this problem is

$$H_0\colon\ (\alpha_1,\ \ldots,\ \alpha_q)' = \underset{\sim}{0}'$$

(3.7.1)

$$\text{against} \qquad H_N(\alpha_{1N},\ \ldots,\ \alpha_{qN})' = N^{-1/2}(\lambda_1,\ \ldots,\ \lambda_q)' \ .$$

From the previous sections we know that the parametric test based on

$$\underset{q\times 1}{\underset{\sim}{Y}_{S1}} = \underset{\sim}{B}_1'\ \underset{\sim}{X}_S; \qquad S = 1,\ 2,\ \ldots,\ n$$

will have noncentral $\chi^2(q,\Delta_1)$ through H_N as $N \to \infty$ where $\Delta_1 = \underset{\sim}{\lambda}'(\underset{\sim}{B}_1'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_1)^{-1}\ \underset{\sim}{\lambda}$. While the nonparametric test using rank scores would also have limiting noncentral $\chi^2(q,\Delta_2)$ through H_N where $\Delta_2 = \underset{\sim}{\lambda}'\ \underset{\sim}{\tau}_1^{-1}\ \underset{\sim}{\lambda}$. $\underset{\sim}{B}_1$ is defined by (3.6.8). We assume that we shall use r of the higher degree terms as covariables in the analysis, i.e. we let

$$\underset{\sim}{Y}_{S2} = \underset{r\times h}{\underset{\sim}{B}_2'}\ \ \underset{\sim}{X}_S \qquad\qquad S = 1,\ 2,\ \ldots,\ n$$

where

$$\underset{\sim}{B}_2 = [\underset{\sim}{b}_{q+1},\ \ldots,\ \underset{\sim}{b}_{q+r}] \qquad r = 1,2,\ldots,(h-q)-1.$$

The expected value of $\bar{\underset{\sim}{Y}}_1$ given $\bar{\underset{\sim}{Y}}_2$ is

$$\underset{\sim}{\alpha} + (\underset{\sim}{B}_1'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_2)(\underset{\sim}{B}_2'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_2)^{-1}\ \bar{\underset{\sim}{Y}}_2\ ,$$

and its covariance matrix is

$$\underset{\sim}{B}_1'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_1 - \underset{\sim}{B}_1'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_2(\underset{\sim}{B}_2'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_2)^{-1}\ \underset{\sim}{B}_2'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_1\ .$$

The residuals are thus

$$\bar{\underset{\sim}{Y}}_1 - \underset{\sim}{\alpha} - (\underset{\sim}{B}_1'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_2)(\underset{\sim}{B}_2'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_2)^{-1}\ \bar{\underset{\sim}{Y}}_2,$$

and under H_0 are

$$\bar{\underset{\sim}{Y}}_1 - (\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1} \bar{\underset{\sim}{Y}}_2$$

with covariance matrix

$$[\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 - \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1} \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1] = \Sigma_{1.2}.$$

Hence,

$$\mathcal{L}(\sqrt{n}[\bar{\underset{\sim}{Y}}_1 - (\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1} \bar{\underset{\sim}{Y}}_2]) \to N(0,\underset{\sim}{\Sigma}_{1.2}) \text{ as } N \to \infty$$

and through H_N

$$\mathcal{L}(\sqrt{n}[\bar{Y}_1 - (\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1}\bar{\underset{\sim}{Y}}_2 - n^{-1/2}\underset{\sim}{\lambda}]) \to N(\underset{\sim}{\lambda},\underset{\sim}{\Sigma}_{1.2}) \text{ as } N \to \infty .$$

In a similar fashion [see e.g. Sen and Puri (1970)] the test based on the rank scores procedure would have noncentral $\chi^2(q,\Delta_3)$ through H_N as $N \to \infty$ with

$$\Delta_3 = \underset{\sim}{\lambda}' \underset{\sim}{\tau}_{1.2}^{-1} \underset{\sim}{\lambda}$$

where

$$\underset{\sim}{\tau}_{1.2} = \underset{\sim}{\tau}_{11} - \underset{\sim}{\tau}_{12} \underset{\sim}{\tau}_{22}^{-1} \underset{\sim}{\tau}_{21} .$$

We then see that the efficiencies of the test with covariables relative to the test without covariates for the parametric procedure are:

$$\text{CARE} = (|\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1|/|\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 - \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1} \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1|)^{1/q}$$

and

$$\text{TARE} = \text{tr}[\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 - \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1} \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1]^{-1}/\text{tr}[\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1]^{-1}.$$

The corresponding efficiencies for the nonparametric procedures are:

$$\text{CARE} = (|\underset{\sim}{\tau}_{11}|/|\underset{\sim}{\tau}_{11} - \underset{\sim}{\tau}_{12}\underset{\sim}{\tau}_{22}^{-1}\underset{\sim}{\tau}_{21}|)^{1/q}$$

and

$$\text{TARE} = \text{tr}[\underset{\sim}{\tau}_{11} - \underset{\sim}{\tau}_{12}\underset{\sim}{\tau}_{22}^{-1}\underset{\sim}{\tau}_{21}]^{-1}/\text{tr}\ \underset{\sim}{\tau}_{11}^{-1}\ .$$

We know $\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1} \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1$ is symmetric positive semi-definite while $\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1$ and $\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 - \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1} \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1$ are symmetric positive definite.

Hence by Theorem 1.44 of Graybill (1961) we see that

$$|\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1| \geq |\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 - \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1} \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1|.$$

A similar argument shows that

$$|\underset{\sim}{\tau}_{11}| \geq |\underset{\sim}{\tau}_{11} - \underset{\sim}{\tau}_{12}\underset{\sim}{\tau}_{22}^{-1}\underset{\sim}{\tau}_{21}|.$$

Therefore, using the curvature criterion of ARE we see that the efficiency of the procedures using covariance adjustment to the corresponding procedure without covariance adjustment is bounded below by unity.

Using the trace criterion we also find that the ARE is bounded below by unity. To demonstrate this we consider the following argument. There exists a nonsingular matrix C such that

$$(3.7.2) \qquad \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 = \underset{\sim}{C}\,\underset{\sim}{C}' \qquad \text{and}$$

$$\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 - \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1} \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 = \underset{\sim}{C}\,\underset{\sim}{\Lambda}\,\underset{\sim}{C}'$$

where $\underset{\sim}{\Lambda} = \text{diag}(\gamma_1, \ldots, \gamma_q)$, $\underset{\sim}{\gamma}_i$ is a characteristic root of

$$(\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 - \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1} \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)(\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)^{-1}.$$

Each γ_i is between zero and one since the roots of

$$(\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 - \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1} \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)(\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)^{-1}$$

are the same as the roots of

$$\underset{\sim}{I} - \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1} \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1(\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)^{-1}$$

i.e. the solution of

$$|\underset{\sim}{I} - \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1} \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1(\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)^{-1} - \gamma \underset{\sim}{I}| = 0$$

for γ or the solution of

$$(3.7.3) \qquad |\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1} \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 - \lambda(\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)| = 0$$

for λ where $\lambda = 1 - \gamma$ are the same. The roots of (3.7.3) are greater than or equal to zero, i.e. $1 - \gamma \geq 0$, therefore $\gamma \leq 1$. Clearly each γ_i is between zero and one. So we see that from (3.7.2)

$$\begin{aligned} \operatorname{tr}[\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 - \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1}\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1]^{-1} & \\ = \operatorname{tr} \underset{\sim}{C}'^{-1} \underset{\sim}{\Lambda}^{-1} \underset{\sim}{C}^{-1} &= \operatorname{tr} \underset{\sim}{\Lambda}^{-1}(\underset{\sim}{C} \underset{\sim}{C}')^{-1} \\ = \operatorname{tr} \underset{\sim}{\Lambda}^{-1}(\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)^{-1} &\geq \operatorname{tr}(\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)^{-1} \end{aligned}$$

since each element of the diagonal matrix $\underset{\sim}{\Lambda}^{-1}$ is greater than or equal to one. Therefore

$$\text{TARE} = \operatorname{tr}[\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1 - \underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_2(\underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_2)^{-1} \underset{\sim}{B}_2' \underset{\sim}{\Sigma} \underset{\sim}{B}_1]^{-1}/\operatorname{tr}[\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1]^{-1} \geq 1.$$

A similar argument can be applied to the nonparametric procedure. Covariance adjustment does improve both the CARE and TARE in this case.

3.7.2 Model Underspecified

In this situation the hypothesis is (3.7.1); however, we have one test with t statistics ($t < q$) and zero covariables and the second test with t statistics and r covariables. We let

$$\underset{\sim}{B} = [\underset{\sim}{b}_1, \ldots, \underset{\sim}{b}_t : \underset{\sim}{b}_{t+1}, \ldots, \underset{\sim}{b}_q] = [\underset{\sim}{B}_1 : \underset{\sim}{B}_2]$$

and assume the least squares reduction

$$\underset{t\times 1}{\underset{\sim}{Y}_{S1}} = \underset{\sim}{B}_1' \underset{\sim}{X}_S; \quad S = 1, 2, \ldots, n.$$

From the previous sections the parametric procedure based on the $\underset{\sim}{Y}_{S1}$ will have noncentral $\chi^2(t,\Delta_1)$ through H_N as $N \to \infty$ where

$$\Delta_1 = \underset{\sim}{\lambda}' \begin{bmatrix} \underset{\sim}{I}_{t\times t} \\ \underset{\sim}{0} \end{bmatrix} (\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)^{-1} [\underset{t\times t}{\underset{\sim}{I}} : 0]\underset{\sim}{\lambda}$$

and

$$\Delta_1 = \underset{\sim}{\lambda}' \begin{bmatrix} \underset{t\times t}{(\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)^{-1}} & \underset{t\times(q-t)}{\underset{\sim}{0}} \\ \underset{(q-t)\times t}{\underset{\sim}{0}} & \underset{(q-t)\times(q-t)}{\underset{\sim}{0}} \end{bmatrix} \underset{\sim}{\lambda} .$$

The nonparametric procedure with rank scores applied to the $\underset{\sim}{Y}_{S1}$ would result in a limiting $\chi^2(t,\Delta_2)$ through H_N with

$$\Delta_2 = \underset{\sim}{\lambda}' \begin{bmatrix} \underset{t\times t}{\underset{\sim}{\tau}_{11}^{-1}} & \underset{t\times(q-t)}{\underset{\sim}{0}} \\ \underset{(q-t)\times t}{\underset{\sim}{0}} & \underset{(q-t)\times(q-t)}{\underset{\sim}{0}} \end{bmatrix} \underset{\sim}{\lambda}.$$

We wish to compare these test to the tests which use r of the higher order terms as covariables. We let

$$\underset{\sim}{Y}_{S2} = \underset{r\times h}{\underset{\sim}{B}_3'} \; \underset{\sim}{X}_S \qquad r = 1, 2, \ldots, (h-t) - 1$$

where

$$\underset{\sim}{B}_3 = [\underset{\sim}{b}_{t+1}, \underset{\sim}{b}_{t+2}, \ldots, \underset{\sim}{b}_{t+r}].$$

We notice that $\underset{\sim}{B}_3$ will contain terms from $\underset{\sim}{B}_2$ and in fact may actually equal $\mathbf{B}_2$. The expected value of $\bar{Y}_1$ given $\bar{Y}_2$ is given by

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_t \end{pmatrix} + (\mathbf{B}_1' \mathbf{\Sigma} \mathbf{B}_3)(\mathbf{B}_3' \mathbf{\Sigma} \mathbf{B}_3)^{-1} \bar{\mathbf{Y}}_2$$

with covariance matrix

$$\mathbf{\Sigma}_{1.3} = \mathbf{B}_1' \mathbf{\Sigma} \mathbf{B}_1 - \mathbf{B}_1' \mathbf{\Sigma} \mathbf{B}_3 (\mathbf{B}_3' \mathbf{\Sigma} \mathbf{B}_3)^{-1} \mathbf{B}_3' \mathbf{\Sigma} \mathbf{B}_1 .$$

The residuals are

$$(\bar{\mathbf{Y}}_1 - (\boldsymbol{\alpha} + (\mathbf{B}_1' \mathbf{\Sigma} \mathbf{B}_3)(\mathbf{B}_3' \mathbf{\Sigma} \mathbf{B}_3)^{-1} \bar{\mathbf{Y}}_2))$$

with the same covariance matrix. Under similar assumptions as before we have that

$$N(\bar{\mathbf{Y}}_1 - \mathbf{B}_1' \mathbf{\Sigma} \mathbf{B}_3 (\mathbf{B}_3' \mathbf{\Sigma} \mathbf{B}_3)^{-1} \bar{\mathbf{Y}}_2) \, \mathbf{\Sigma}_{1.3}^{-1} (\bar{\mathbf{Y}}_1 - \mathbf{B}_1' \mathbf{\Sigma} \mathbf{B}_3 (\mathbf{B}_3' \mathbf{\Sigma} \mathbf{B}_3)^{-1} \bar{Y}_2)$$

is noncentral $\chi^2(t,\Delta_3)$ through H_N as $N \rightarrow \infty$, where

$$\Delta_3 = \boldsymbol{\lambda}' \begin{bmatrix} [\mathbf{B}_1' \mathbf{\Sigma} \mathbf{B}_1 - \mathbf{B}_1' \mathbf{\Sigma} \mathbf{B}_3 (\mathbf{B}_3' \mathbf{\Sigma} \mathbf{B}_3)^{-1} \mathbf{B}_3 \mathbf{\Sigma} \mathbf{B}_1]^{-1} ; & \mathbf{0} \\ \mathbf{0} \qquad ; & \mathbf{0} \end{bmatrix} \boldsymbol{\lambda} .$$

In a like fashion the nonparametric procedure with covariables would have a noncentral $\chi^2(t,\Delta_4)$ through H_N as $N \rightarrow \infty$ where

$$\Delta_4 = \boldsymbol{\lambda}' \begin{bmatrix} [\boldsymbol{\tau}_{11} - \boldsymbol{\tau}_{13}' \boldsymbol{\tau}_{33}^{-1} \boldsymbol{\tau}_{31}]^{-1} , & \mathbf{0} \\ \mathbf{0} \qquad , & \mathbf{0} \end{bmatrix} \boldsymbol{\lambda} .$$

We thus have the following as the efficiency of the test using covariates to the test without the use of covariates for the parametric procedure:

$$\text{TARE} = \operatorname{tr}[\mathbf{B}_1' \mathbf{\Sigma} \mathbf{B}_1 - \mathbf{B}_1' \mathbf{\Sigma} \mathbf{B}_3 (\mathbf{B}_3' \mathbf{\Sigma} \mathbf{B}_3)^{-1} \mathbf{B}_3' \mathbf{\Sigma} \mathbf{B}_1]^{-1} / \operatorname{tr}[\mathbf{B}_1' \mathbf{\Sigma} \mathbf{B}_1]^{-1} ,$$

and for the corresponding nonparametric procedure is:

$$\mathrm{TARE} = \mathrm{tr}[\underset{\sim}{\tau}_{11} - \underset{\sim}{\tau}_{13}\, \underset{\sim}{\tau}_{33}^{-1}\, \underset{\sim}{\tau}_{31}]^{-1} / \mathrm{tr}\ \underset{\sim}{\tau}_{11}^{-1} .$$

The trace criterion is the only method of comparison which we can use in this case since the curvature of the power function of each test is zero. We also note these efficiencies are bounded below by one from the arguments in section 3.7.1. Even if we have underspecified the model we always do better to use covariance adjustment.

3.7.3 Model Overspecified

In this situation the hypothesis is again given in (3.7.1); however, we have one test with t statistics $(t > q)$ and no covariables and the second test with t statistics and r covariables. We let $\underset{\sim}{B} = [\underset{\sim}{b}_1, \ldots, \underset{\sim}{b}_q, \underset{\sim}{b}_{q+1}, \ldots, \underset{\sim}{b}_t]$ and assume the least squares reduction

$$\underset{t\times 1}{\underset{\sim}{Y}_{S1}} = \underset{\sim}{B}'\, \underset{\sim}{X}_S; \qquad S = 1, 2, \ldots, n.$$

From the previous sections the parametric procedure based on the $\underset{\sim}{Y}_{S1}$ will have noncentral $\chi^2(t,\Delta_1)$ through H_N as $N \to \infty$ where

$$\Delta_1 = \underset{\sim}{\lambda}'[\underset{\sim}{B}_1'\, \underset{\sim}{\Sigma}\, \underset{\sim}{B}_1 - \underset{\sim}{B}_1'\, \underset{\sim}{\Sigma}\, \underset{\sim}{B}_2(\underset{\sim}{B}_2'\, \underset{\sim}{\Sigma}\, \underset{\sim}{B}_2)^{-1}\, \underset{\sim}{B}_2'\, \underset{\sim}{\Sigma}\, \underset{\sim}{B}_1]^{-1}\, \underset{\sim}{\lambda}.$$

The nonparametric procedure with rank scores applied to the $\underset{\sim}{Y}_{S1}$ would result in a limiting $\chi^2(t,\Delta_2)$ through H_N as $N \to \infty$, where

$$\Delta_2 = \underset{\sim}{\lambda}'[\underset{\sim}{\tau}_{11} - \underset{\sim}{\tau}_{12}\, \underset{\sim}{\tau}_{22}^{-1}\, \underset{\sim}{\tau}_{21}]^{-1}\, \underset{\sim}{\lambda}.$$

The partitioned matrices are defined by $B = [\underset{h\times q}{\underset{\sim}{B}_1}, \ \underset{h\times(t-q)}{\underset{\sim}{B}_2}]$ and $\underset{\sim}{\tau}_{ij}$ are the covariances of the rank scores of $\underset{\sim}{B}_i'\, \underset{\sim}{X}_S$ and $\underset{\sim}{B}_j'\, \underset{\sim}{X}_S$. We wish to compare these tests to the tests which use r of the higher order terms as covariables we let

$$\underset{\sim}{Y}_{S2} = \underset{r\times h}{\underset{\sim}{B}_3'}\ \underset{\sim}{X}_S \qquad r = 1, 2, \ldots, (h-t)-1$$

where

$$\underset{\sim}{B}_3 = [\underset{\sim}{b}_{t+1}, \ldots, \underset{\sim}{b}_{t+r}].$$

The expected value of $\bar{\underset{\sim}{Y}}_1$ given $\bar{\underset{\sim}{Y}}_2$ is

$$\begin{pmatrix} \underset{\sim}{\alpha} \\ \underset{\sim}{0} \end{pmatrix} + (\underset{\sim}{B}'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_3)(\underset{\sim}{B}_3'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_3)^{-1}\ \bar{\underset{\sim}{Y}}_2$$

with covariance matrix

$$\underset{\sim}{\Sigma}_{1.3} = \underset{\sim}{B}'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B} - \underset{\sim}{B}'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_3(\underset{\sim}{B}_3'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_3)^{-1}\ \underset{\sim}{B}_3'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}$$

The residuals as

$$\left(\bar{\underset{\sim}{Y}}_1 - \left(\begin{pmatrix} \underset{\sim}{\alpha} \\ \underset{\sim}{0} \end{pmatrix} + (\underset{\sim}{B}'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_3)(\underset{\sim}{B}_3'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_3)^{-1}\ \bar{\underset{\sim}{Y}}_2 \right) \right)$$

with the same covariance matrix and through H_N we assume that

$$\left(\sqrt{n} \left(\bar{\underset{\sim}{Y}}_1 - (\underset{\sim}{B}'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_3)(\underset{\sim}{B}_3'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_3)^{-1}\bar{\underset{\sim}{Y}}_2 - n^{-1/2} \begin{pmatrix} \underset{\sim}{\lambda} \\ \underset{\sim}{0} \end{pmatrix} \right) \right) \to N_t\left(\begin{pmatrix} \underset{\sim}{\lambda} \\ \underset{\sim}{0} \end{pmatrix}, \Sigma_{1.3} \right)$$

as $n \to \infty$.

We denote

$$\underset{\sim}{A}_{ij} = \underset{\sim}{B}_i'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_j \qquad i,j = 1,2,3,$$

and

$$\underset{\sim}{A}_{ij\cdot 3} = \underset{\sim}{A}_{ij} - \underset{\sim}{A}_{i3}\ \underset{\sim}{A}_{33}^{-1}\ \underset{\sim}{A}_{3j} \qquad i,j = 1,2.$$

Also let

$$A^* = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \text{and}$$

$$A^{**} = \begin{pmatrix} A_{11.3} & A_{12.3} \\ A_{21.3} & A_{22.3} \end{pmatrix} .$$

We know that

$$|\underset{\sim}{A}_{11} - \underset{\sim}{A}_{12} \underset{\sim}{A}_{22}^{-1} \underset{\sim}{A}_{21}| = |\underset{\sim}{A}^{*}|/|\underset{\sim}{A}_{22}|$$

and

$$|\underset{\sim}{A}_{11.3} - \underset{\sim}{A}_{12.3} \underset{\sim}{A}_{22.3}^{-1} \underset{\sim}{A}_{21.3}| = |\underset{\sim}{A}^{**}|/|\underset{\sim}{A}_{22.3}| .$$

Thus using the curvature criterion we get

$$\text{(3.7.4)} \quad \text{CARE} = \frac{|\underset{\sim}{A}_{11} - \underset{\sim}{A}_{12} \underset{\sim}{A}_{22}^{-1} \underset{\sim}{A}_{21}|}{|\underset{\sim}{A}_{11.3} - \underset{\sim}{A}_{12.3} \underset{\sim}{A}_{22.3}^{-1} \underset{\sim}{A}_{21.3}|} = \frac{|\underset{\sim}{A}^{*}|}{|\underset{\sim}{A}^{**}|} \; \frac{|\underset{\sim}{A}_{22.3}|}{|\underset{\sim}{A}_{22}|}$$

for the efficiency of the test with covariates relative to the test without covariates. The first ratio in (3.7.4) is at least unity, however, the second term is less than or equal to one. It does not seem obvious whether the CARE is greater than or equal to one. Using the nonparametric rank scores procedures we have for the efficiency of the test with covariates to the test without covariates:

$$\text{CARE} = \frac{|\underset{\sim}{\tau}_{11} - \underset{\sim}{\tau}_{12} \underset{\sim}{\tau}_{22}^{-1} \underset{\sim}{\tau}_{21}|}{|\tau_{11.3} - \tau_{12.3} \tau_{22.3}^{-1} \tau_{21.3}|} .$$

As in the parametric case it is not clear if this quantity can be less than unity.

3.8 Comments

It should be observed that the hypothesis

$$H_0: \begin{pmatrix} \alpha_{\ell+1} \\ \vdots \\ \alpha_{h-1} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \text{ with } \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_\ell \end{pmatrix}$$

unspecified ($\ell < h$) could have been the hypothesis tested in this chapter in place of the hypothesis given in (3.3.4). To test this hypothesis we would apply the parametric or nonparametric one sample procedure to

$$\underset{\sim}{Y}_S = \underset{\sim}{B}' \underset{\sim}{X}_S, \quad S = 1, 2, \ldots, n$$

where

$$\underset{\sim}{B} = [\underset{\sim}{b}_{\ell+1}, \ldots, \underset{\sim}{b}_{h-1}].$$

The bounds attained in this chapter would not change; however, the actual efficiency formulae would change to reflect the $\underset{\sim}{B}$ matrix. In general the results of this chapter are applicable to tests for the one sample problem. In Chapter IV we find that the TARE and CARE reduce to scalar multiples of the corresponding one sample results for the hypothesis that the intercept is a fixed number and the curve is stationary.

CHAPTER IV

APPLICATION OF TARE AND CARE TO THE MULTI-SAMPLE GROWTH CURVE PROBLEM

4.1 Introduction

We shall present results in this chapter for the c-sample problem ($c > 1$) similar to the results obtained in Chapter III for the one-sample problem. As in Chapter III we shall restrict attention to the polynomial growth curve model. The hypothesis of interest will be the hypothesis of the equality of the c growth curves, under various assumptions of what the model is, we shall see that this test of homogeneity will give rise to different test statistics. The unweighted least squares reduction will be used for reduction of each observation vector to the estimates for the assumed model. We shall apply the general rank scores statistics to these reduced vectors and also shall apply the Hotelling-Lawley trace statistics to test the null hypothesis. Efficiency formulae are presented for each procedure.

4.2 The Statistical Model and Data Reduction

The model for the multi-sample (c-sample) problem is a special case of the general model presented in Section 3.2. The index set, I, is the set of c indices $1,2,\ldots,c$. The index set, T, is the same as in (3.2.2) and the data are characterized by (3.2.3) and (3.2.4). For notational ease, we write n_i in place of $n(i)$ since $i = 1,2,\ldots,c$.

Restricting attention to the case b = 1 we see that we have the following for the multi-sample problem.

$$(4.2.1) \qquad \underset{h\times 1}{\underset{\sim}{X}_{Si}} = \begin{bmatrix} X_{Si}(t_1) \\ \vdots \\ X_{Si}(t_h) \end{bmatrix} \qquad \begin{matrix} i = 1,2,\ldots,c \\ S = 1,2,\ldots,n_i \end{matrix}$$

where for each i the $\underset{h\times 1}{\underset{\sim}{X}_{Si}}$ are i.i.d. as $G(\underset{\sim}{X};i)$ an h-variate absolutely continuous distribution function. The location vector of $\underset{\sim}{X}_{Si}$ is $M(\underset{\sim}{i})$ where

$$\underset{\sim}{M}(i) = \begin{bmatrix} \mu(i,t_1) \\ \vdots \\ \mu(i,t_h) \end{bmatrix} \qquad i = 1,2,\ldots,c$$

and in the orthogonal polynomial models which we consider, we have that

$$(4.2.2) \qquad M(i) = \underset{h\times r}{\underset{\sim}{B}_1} \; \underset{r\times 1}{\underset{\sim}{\alpha}_i} \qquad i = 1,2,\ldots,c$$

where $\underset{\sim}{B}_1$ are the orthogonal polynomials of degree 0 to r-1 satisfying (3.2.9) and (3.2.10). Hence, we have again a dimension reducing transformation from $\underset{\sim}{M}(i)$ to $\underset{\sim}{\alpha}_i$. The assumption

$$(4.2.3) \qquad G(\underset{\sim}{X};i) = G(\underset{\sim}{X} + \underset{\sim}{M}(i)) \qquad i = 1,2,\ldots,c$$

is made so that $\underset{\sim}{X}_{Si} - \underset{\sim}{M}(i)$ is distributed independently of i. In the normal theory we assume

$$(4.2.4) \qquad G(\underset{\sim}{X};i) = N_h(\underset{\sim}{M}(i), \underset{\sim}{\Sigma}) \qquad i = 1,2,\ldots,c,$$

if M(i) is given by (4.2.2). The hypothesis of homogeneity of the growth curves is given by

$$(4.2.5) \qquad H_0\colon \underset{\sim}{\alpha}_1 = \underset{\sim}{\alpha}_2 = \ldots = \underset{\sim}{\alpha}_c = \underset{\sim}{\alpha} \quad \text{(unspecified)}.$$

Defining $\underset{\sim}{b}_j$ $j = 0,1,\ldots,h-1$ as in (3.2.10) we shall consider the unweighted least squares estimate of $\underset{\sim}{\alpha}_i$ for each observation; hence denoting $\underset{\sim}{B}_1$ by

$$\underset{\sim}{B}_1 = [\underset{\sim}{b}_0,\underset{\sim}{b}_1,\ldots,\underset{\sim}{b}_{r-1}]$$

then consider the transformation

$$\underset{\sim}{Y}_{Si} = \underset{\sim}{B}_1' \underset{\sim}{X}_{Si} \qquad i = 1,2,\ldots,c \quad S = 1,2,\ldots,n_i. \tag{4.2.6}$$

In the next sections we compute the efficiencies of the parametric and nonparametric procedures for the overspecified and underspecified growth curve models.

4.3 Parametric Procedure

Under the assumption that $\underset{\sim}{X}_{Si}$ are from a multivariate normal distribution defined by (4.2.4) it follows that $\underset{\sim}{Y}_{Si}$ defined by (4.2.6) has a multivariate normal distribution. Hence, to test the null hypothesis (4.2.5) we could follow any of several procedures, for example, the likelihood criterion which reduces to Wilks lambda, the Hotelling-Lawley trace, or Roy's largest root. Under the assumption of normality the Hotelling-Lawley trace criterion (denoted T_0^2) and the likelihood ratio criterion are asymptotically equivalent, in fact, both would lead to noncentral chi-square statistics for large samples. On the other hand if the parent distribution is not necessarily normal but has moments of order $2 + \delta$ for some $\delta > 0$ then T_0^2 still has a central chi-square under the null hypothesis and a noncentral chi-square through an appropriate sequence of alternatives. We shall use the T_0^2 statistic, but the equivalence of the two criteria in large samples for normal distribution functions is noteworthy. We consider the data characterized by (3.2.3) and (3.2.4) and define

$$\bar{\underset{\sim}{X}}_i = n_i^{-1} \sum_{S=1}^{n_i} \underset{\sim}{X}_{Si} \qquad i = 1,2,\ldots,c$$

$$\bar{\underset{\sim}{X}} = \sum_{i=1}^{c} \frac{n_i}{N} \bar{\underset{\sim}{X}}_i \qquad i = 1,2,\ldots,c$$

where
$$N = \sum_{i=1}^{c} n_i$$

$$\underset{\sim}{S} = (N-c)^{-1} \sum_{i=1}^{c} \Big(\sum_{S=1}^{n_i} (\underset{\sim}{X}_{Si} - \bar{\underset{\sim}{X}}_i)(\underset{\sim}{X}_{Si} - \bar{\underset{\sim}{X}}_i)' \Big).$$

Assume there exist c constants, $\gamma_1, \gamma_2, \ldots, \gamma_c$, each in the open interval from zero to one such that

$$\lim_{N \to \infty} \frac{n_i}{N} = \gamma_i; \quad i = 1,2,\ldots,c \text{ and } \sum_{i=1}^{c} \gamma_i = 1.$$

Since the parent distribution function of $\underset{\sim}{X}_{Si}$ has moments of order $2 + \delta$ for some $\delta > 0$, it follows from the central limit theorem that

$$(4.3.1) \qquad \mathcal{L}\big(n_i^{1/2} (\bar{\underset{\sim}{X}}_i - E\, \bar{\underset{\sim}{X}}_i)\big) \to N_h(\underset{\sim}{0},\underset{\sim}{\Sigma}) \text{ as } n_i \to \infty .$$

Furthermore we know that $\underset{\sim}{S} \overset{P}{\to} \underset{\sim}{\Sigma}$ by the laws of large numbers. From (4.3.1) it follows that

$$(4.3.2) \qquad \mathcal{L}\big(n_i^{1/2} (\bar{\underset{\sim}{Y}}_i - E\, \bar{\underset{\sim}{Y}}_i)\big) \to N_r(\underset{\sim}{0},\underset{\sim}{B}' \underset{\sim}{\Sigma} \underset{\sim}{B}) \text{ as } n_i \to \infty$$

as $\bar{Y}_i$ is a continuous function of $\bar{X}_i$. In addition it is obvious from (4.3.2) that

$$(4.3.3) \quad \mathcal{L}\big(N^{1/2}(\bar{\underset{\sim}{Y}}_i - E\, \bar{\underset{\sim}{Y}}_i);\ i = 1,2,\ldots,c\big) \to N_{rc}\Big(\underset{\sim}{0}, \underset{r\times r}{(\underset{\sim}{B}' \underset{\sim}{\Sigma} \underset{\sim}{B})} \otimes \underset{c\times c}{\underset{\sim}{\Gamma}}\Big)$$

where

$$\underset{\sim}{\Gamma} = \begin{bmatrix} \frac{1}{\gamma_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\gamma_c} \end{bmatrix} .$$

Writing $\underset{\sim}{\alpha}_i$ as $\underset{\sim}{B} + \underset{\sim}{\theta}_i$ where $\sum_{i=1}^{c} \gamma_i \underset{\sim}{\theta}_i = \underset{\sim}{0}$ we observe that only c - 1 of the $\underset{\sim}{\theta}_i$ are linearly independent; hence the null hypothesis implies that $\underset{\sim}{\theta}_1 = \underset{\sim}{\theta}_2 = \cdots = \underset{\sim}{\theta}_i = \underset{\sim}{0}$. The sequence of alternative hypotheses, H_N, is defined by

$$(4.3.4) \qquad H_N: \ \underset{\sim}{\theta}_{iN} = N^{-1/2} \underset{\sim}{\lambda}_i \qquad i = 1,2,\ldots,c$$

where $\underset{\sim}{\lambda}_i$ is non-null for at least one i.

To test H_0, we may define T_0^2 in its symmetric, less than full rank, form of

$$(4.3.5) \qquad T_0^2 = \sum_{i=1}^{n} n_i(\underset{\sim}{\bar{Y}}_i - \underset{\sim}{\bar{Y}})(\underset{\sim}{B}_1' \underset{\sim}{S} \underset{\sim}{B}_1)^{-1}(\underset{\sim}{\bar{Y}}_i - \underset{\sim}{\bar{Y}})'$$

or we may consider the statistic in one of its full rank forms, for example:

$$(4.3.6) \quad T_0^2 = N[\underset{\sim}{\bar{Y}}_i - \underset{\sim}{\bar{Y}}); \ i = 1,2,\ldots,c-1]' \left\{ (\underset{\sim}{B}_1' \underset{\sim}{S} \underset{\sim}{B}_1) \otimes \left[\frac{\delta_{ii'}N}{n_i} - 1 \right]_{i,i'=1,2,\ldots,c-1} \right\}^{-1} [(\underset{\sim}{\bar{Y}}_i - \underset{\sim}{\bar{Y}}); \ i=1,2,\ldots,c-1].$$

By (4.3.3) and computation of the covariance we know, under H_0, that

$$\mathcal{L}([\sqrt{N}(\bar{\underset{\sim}{Y}}_i - \bar{\underset{\sim}{Y}});\ i = 1,2,\dots,c-1] \to N_{r(c-1)}\left(\underset{\sim}{0}, (\underset{\sim}{B}_1'\underset{\sim}{\Sigma}\,\underset{\sim}{B}_1) \otimes \left(\frac{\delta_{ii'}}{\gamma_i} - 1\right)_{i,i'=1,2,\dots,c-1}\right) \text{ as } N \to \infty.$$

Furthermore through H_N we see that

$$(4.3.7)\quad \mathcal{L}(\sqrt{N}(\bar{\underset{\sim}{Y}}_i - \bar{\underset{\sim}{Y}});\ i = 1,2,\dots,c-1] \to N_{r(c-1)}\left([\underset{\sim}{\lambda}_i - \bar{\underset{\sim}{\lambda}}), i=1,2,\dots,c-1];\ (\underset{\sim}{B}_1'\underset{\sim}{\Sigma}\,\underset{\sim}{B}) \otimes \left(\frac{\delta_{ii'}}{\gamma_i} - 1\right)_{i,i'=1,2,\dots,c-1}\right) \text{ as } N \to \infty$$

where $\bar{\underset{\sim}{\lambda}} = \sum_{i=1}^{c} \gamma_i \underset{\sim}{\lambda}_i$. Without loss of generality we may assume $\bar{\underset{\sim}{\lambda}}$ is zero because of the r restrictions. Clearly, by virtue of (4.3.7) T_0^2 has a limiting $\chi^2(r(c-1), \Delta\ )$ through H_N as $N \to \infty$ with

$$(4.3.8)\quad \Delta_1 = [\underset{\sim}{\lambda}_i;\ i = 1,2,\dots,c-1]' \left[(\underset{\sim}{B}_1'\underset{\sim}{\Sigma}\,\underset{\sim}{B}_1) \otimes \left(\frac{\delta_{ii'}}{\gamma_i} - 1\right)_{i,i'=1,2,\dots c-1}\right]^{-1} [\underset{\sim}{\lambda}_i;\ i = 1,2,\dots,c-1].$$

Written in a symmetric form in the $\underset{\sim}{\lambda}_i$, $i = 1,2,\dots,c$ Δ_1 becomes

$$\Delta_1 = \sum_{i=1}^{c} \gamma_i\, \underset{\sim}{\lambda}_i' \,(\underset{\sim}{B}_1'\, \underset{\sim}{\Sigma}\, \underset{\sim}{B}_1)^{-1} \underset{\sim}{\lambda}_i.$$

We shall need to compute the trace and determinant of the discriminant in (4.3.8) so we reduce this discriminant to a simpler form. Noting that

$$(4.3.9)\quad \left(\frac{\delta_{ii'}}{\gamma_i} - 1\right)_{i,i'=1,2,\dots,c-1} = \begin{bmatrix} \frac{1}{\gamma_1} & 0 \dots & 0 \\ 0 & & \\ \vdots & & \\ 0 & & \frac{1}{\gamma_{c-1}} \end{bmatrix} + (-1)\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}(1,1,\dots,1)$$

thus by Theorem 8.3.3 in Graybill (1969) the inverse of (4.3.9) is

$$\begin{pmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \cdot & & \\ \vdots & & \cdot & \\ 0 & & & \gamma_{c-1} \end{pmatrix} + K \begin{pmatrix} \gamma_1 \\ \cdot \\ \cdot \\ \cdot \\ \gamma_{c-1} \end{pmatrix} (\gamma_1, \ldots, \gamma_{c-1})$$

where

$$K = -(-1)\Big(1 + (-1)\sum_{i=1}^{c-1} \gamma_i\Big)^{-1} = (1 - (1-\gamma_c))^{-1} = \gamma_c^{-1}$$

Consequently, we have

$$(4.3.10) \qquad \left[\frac{\delta_{ii'}}{\gamma_i} - 1\right]^{-1}_{i,i'=1,2,\ldots,c-1} = \left[\delta_{ii'}\gamma_i + \frac{\gamma_i\gamma_{i'}}{\gamma_c}\right]_{i,i'=1,2,\ldots,c-1} .$$

Furthermore, since the inverse of a direct product is the direct product of the inverses, it follows from (4.3.10) that

$$\left[(\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1) \otimes \left[\frac{\delta_{ii'}}{\gamma_i} - 1\right]_{i,i'=1,2,\ldots,c-1}\right]^{-1}$$

$$= (\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)^{-1} \otimes \left[\delta_{ii'}\gamma_i + \frac{\gamma_i\gamma_{i'}}{\gamma_c}\right]_{i,i'=1,2,\ldots,c-1} .$$

In addition, by Theorem 8.8.10 and 9.1.11 in Graybill (1969) it is apparent that:

$$(4.3.11) \qquad \left|(\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)^{-1} \otimes \left[\delta_{ii'}\gamma_i + \frac{\gamma_i\gamma_{i'}}{\gamma_c}\right]_{i,i'=1,2,\ldots,c-1}\right|$$

$$= \left|(\underset{\sim}{B}_1' \underset{\sim}{\Sigma} \underset{\sim}{B}_1)^{-1}\right|^{c-1} \left|\left[\delta_{ii'}\gamma_i + \frac{\gamma_i\gamma_{i'}}{\gamma_c}\right]_{i,i'=1,2,\ldots,c-1}\right|^{(r)}$$

and

$$(4.3.12)\qquad \mathrm{tr}\left[(\underset{\sim}{B}_1'\underset{\sim}{\Sigma}\,\underset{\sim}{B}_1)^{-1} \otimes \left[\delta_{ii'}\gamma_i + \frac{\gamma_i\gamma_{i'}}{\gamma_c}\right]_{i,i'=1,2,\ldots,c-1}\right]$$

$$= \mathrm{tr}(\underset{\sim}{B}_1'\underset{\sim}{\Sigma}\,\underset{\sim}{B}_1)^{-1}\ \mathrm{tr}\left[\left[\delta_{ii}\gamma_i + \frac{\gamma_i\gamma_{i'}}{\gamma_c}\right]_{i,i'}\right].$$

Let us define

$$(4.3.13)\qquad \underset{\sim}{A} = \left(\left(\delta_{ii'}\gamma_i + \frac{\gamma_i\gamma_{i'}}{\gamma_c}\right)\right)_{i,i'=1,2,\ldots,c-1}$$

4.3.1 Parametric Procedure - Overfitting

If we had incorrectly assumed that the polynomial model was of degree (t-1) in each group (t>r) then we would make the transformation:

$$(4.3.14)\qquad \underset{\sim}{Y}_{Si} = [\underset{\sim}{b}_0,\underset{\sim}{b}_1,\ldots,\underset{\sim}{b}_{r-1},\ldots,\underset{\sim}{b}_{t-1}]'\,\underset{\sim}{X}_{Si} \qquad \begin{array}{l} i = 1,2,\ldots,c \\ S = 1,2,\ldots,n_i \end{array}$$

and for notational convenience we let $\underset{\sim}{B} = [\underset{\sim}{B}_1 : \underset{\sim}{B}_2]$

$$\underset{\sim}{B}_1 = [\underset{\sim}{b}_0,\ldots,\underset{\sim}{b}_{r-1}]$$

$$\underset{\sim}{B}_2 = [\underset{\sim}{b}_r,\ldots,\underset{\sim}{b}_{t-1}].$$

Following, the same reasoning as in (4.3.6) and (4.3.7) we see that through H_N

$$(4.3.15)\qquad \mathcal{L}(\sqrt{N}(\bar{\underset{\sim}{Y}}_i-\bar{\underset{\sim}{Y}});\ i = 1,2,\ldots,c-1) \to N_{t(c-1)}\left[[\binom{\underset{\sim}{\lambda}_i}{\underset{\sim}{0}}, i = 1,2,\ldots,c-1];\ (\underset{\sim}{B}'\underset{\sim}{\Sigma}\,\underset{\sim}{B}) \otimes \left[\frac{\delta_{ii'}}{\gamma_i} - 1\right]_{i,i'=1,2,\ldots,c-1}\right] \text{ as } N \to \infty$$

where for each i; $\underset{\sim}{\lambda}_i$ is r×1 and $\underset{\sim}{0}$ is (t-r)×1. Therefore, the statistic

T_0^2, defined as the quadratic form in the t(c-1) vectors in (4.3.15) would have limiting noncentral $\chi^2(t(c-1),\Delta_2)$ as $N \to \infty$ where

$$\Delta_2 = [\underset{\sim}{\binom{\underset{\sim}{\lambda}_i}{0}};i=1,2,\ldots,c-1]' \left[(\underset{\sim}{B}'\underset{\sim}{\Sigma}\underset{\sim}{B})^{-1} \otimes \left(\delta_{ii'}\gamma_i + \frac{\gamma_i\gamma_{i'}}{\gamma_c} \right)_{i,i'=1,2,\ldots,c-1} \right]$$

$$[\underset{\sim}{\binom{\underset{\sim}{\lambda}_i}{0}};\ i=1,2,\ldots,c-1].$$

By (3.4.7) and the fact that the last (t-r) elements of each $\underset{\sim}{\binom{\underset{\sim}{\lambda}_i}{0}}$ vector are zeroes we have

$$\Delta_2 = [\underset{\sim}{\lambda}_i,\ i=1,2,\ldots,c-1]' ([\underset{\sim}{B}_1'\underset{\sim}{\Sigma}\underset{\sim}{B}_1 - \underset{\sim}{B}_1'\underset{\sim}{\Sigma}\underset{\sim}{B}_2(\underset{\sim}{B}_2'\underset{\sim}{\Sigma}\underset{\sim}{B}_2)^{-1} \underset{\sim}{B}_2'\underset{\sim}{\Sigma}\underset{\sim}{B}_1]^{-1} \otimes \underset{\sim}{A})$$

$$[\underset{\sim}{\lambda}_i;\ i = 1,2,\ldots,c-1].$$

As in (4.3.11) and (4.3.12) we have the determinant and trace of the discriminant of Δ_2 as

$$|[\underset{\sim}{B}_1'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_1 - \underset{\sim}{B}_1'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_2(\underset{\sim}{B}_2'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_2)^{-1}\ \underset{\sim}{B}_2'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}]^{-1}|^r\ |\underset{\sim}{A}|^{c-1}$$

and

$$\text{tr}[\underset{\sim}{B}_1'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_1 - \underset{\sim}{B}_1'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_2(\underset{\sim}{B}_2'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_2)^{-1}\ \underset{\sim}{B}_2'\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_1]^{-1}\ \text{tr}\ \underset{\sim}{A},$$

respectively.

Representing the test based on the correct number of parameters as ϕ_1 and the overfitted procedure as ϕ_2 we easily find the following efficiencies of ϕ_2 relative to ϕ_1:

$$\text{CARE} = R(r(c-1),t(c-1),\alpha)\left\{ \frac{|\underset{\sim}{B}_1\underset{\sim}{\Sigma}\underset{\sim}{B}_1|^{(c-1)}}{|\underset{\sim}{B}_1'\underset{\sim}{\Sigma}\underset{\sim}{B}_1 - \underset{\sim}{B}_1'\underset{\sim}{\Sigma}\underset{\sim}{B}_2(\underset{\sim}{B}_2'\underset{\sim}{\Sigma}\underset{\sim}{B}_2)^{-1}\underset{\sim}{B}_2'\underset{\sim}{\Sigma}\underset{\sim}{B}_1|^{(c-1)}} \right\}^{\frac{1}{r(c-1)}}$$

which becomes

$$(4.3.16)\quad \text{CARE} = R(r(c-1),t(c-1),\alpha)\left\{\frac{|\underset{\sim}{B}_1\underset{\sim}{\Sigma}\underset{\sim}{B}_1|}{|\underset{\sim}{B}_1'\underset{\sim}{\Sigma}\underset{\sim}{B}_1 - \underset{\sim}{B}_1'\underset{\sim}{\Sigma}\underset{\sim}{B}_2(\underset{\sim}{B}_2'\underset{\sim}{\Sigma}\underset{\sim}{B}_2)^{-1}\underset{\sim}{B}_2'\underset{\sim}{\Sigma}\underset{\sim}{B}_1|}\right\}^{1/r}$$

for $t \geq r$.

Using the trace criterion we obtain

$$(4.3.17)\quad \text{TARE} = R(r(c-1),t(c-1),\alpha)\left\{\frac{\text{tr}[\underset{\sim}{B}_1'\underset{\sim}{\Sigma}\underset{\sim}{B}_1 - \underset{\sim}{B}_1'\underset{\sim}{\Sigma}\underset{\sim}{B}_2(\underset{\sim}{B}_2'\underset{\sim}{\Sigma}\underset{\sim}{B}_2)^{-1}\underset{\sim}{B}_2'\underset{\sim}{\Sigma}\underset{\sim}{B}_1]^{-1}}{\text{tr}[\underset{\sim}{B}_1'\underset{\sim}{\Sigma}\underset{\sim}{B}_1]^{-1}}\right\}$$

for $t \geq r$.

We see therefore that the c-sample efficiencies for overfitting are the one sample efficiency, in both cases adjusted by a scalar function. The equations (4.3.16) and (4.3.17) do illustrate the fact that the loss in efficiency may be more severe for the c-sample problem than it was for the one-sample problem.

4.3.2 Parametric Procedure - Underfitting

If we had assumed that the polynomial growth curve was of degree $t-1(t<r)$ in each group then we would make the transformation

$$(4.3.18)\qquad \underset{\sim}{Y}_{Si} = [\underset{\sim}{b}_0,\underset{\sim}{b}_1,\ldots,\underset{\sim}{b}_{t-1}]'\ \underset{\sim}{X}_{Si} \qquad \begin{array}{l} i = 1,\ldots,c \\ S = 1,2,\ldots,n_i \end{array}.$$

Let

$$\underset{\sim}{B}^* = [\underset{\sim}{b}_0,\underset{\sim}{b}_2,\ldots,\underset{\sim}{b}_{t-1}]$$

and

$$\underset{\sim}{B}_1^* = [\underset{\sim}{B}^*;\underset{\sim}{b}_t,\ldots,\underset{\sim}{b}_{r-1}] = [\underset{\sim}{B}^*:\underset{\sim}{B}_2^*].$$

Then through H_N we see that

$$(4.3.19)\ \mathcal{L}(\sqrt{N}(\bar{\underset{\sim}{Y}}_i - \bar{\underset{\sim}{Y}});\ i=1,2,\ldots,c-1) \rightarrow N_{t(c-1)}([\underset{\sim}{\lambda}_i^*,\ i=1,2,\ldots,c-1];$$

$$(\underset{\sim}{B}^{*\prime}\underset{\sim}{\Sigma}\underset{\sim}{B}^*) \otimes \underset{\sim}{A}^{-1}) \text{ as } N \rightarrow \infty.$$

where $\underset{\sim}{\lambda}_i^*$ is defined by $\underset{\sim}{\lambda}_i = \begin{bmatrix} \underset{t\times 1}{\lambda_i^*} \\ \lambda_{i(t+1)} \\ \vdots \\ \lambda_{i(r)} \end{bmatrix}$. Hence, the quadratic form in the statistics $\sqrt{N}(\bar{\underset{\sim}{Y}}_i - \bar{\underset{\sim}{Y}})$ would have $\chi^2(t(c-1),\Delta_3)$ as $N \to \infty$ through H_N where

$$\Delta_3 = [\underset{\sim}{\lambda}_i^*;\ i=1,2,\ldots,c-1]'\ [(\underset{\sim}{B}^{*'}\underset{\sim}{\Sigma}\underset{\sim}{B}^*)^{-1} \otimes \underset{\sim}{A}]\ [\underset{\sim}{\lambda}_i^*, i=1,2,\ldots,c-1].$$

Using the trace criteria the ARE of the underfitting procedure, ϕ_2, to the procedure with the correct numbers of parameters, ϕ_1, is:

$$\text{TARE} = R(r(c-1),t(c-1),\alpha)\ \frac{\operatorname{tr}[\underset{\sim}{B}^{*'}\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}^*]^{-1}}{\operatorname{tr}[\underset{\sim}{B}_1^{*'}\ \underset{\sim}{\Sigma}\ \underset{\sim}{B}_1^*]^{-1}} \tag{4.3.20}$$

4.4 Nonparametric Procedures

We shall present efficiency formulae in this section for the general rank scores procedures for the multivariate multisample problem as outlined in Puri and Sen (1971). Since the $\underset{\sim}{X}_{Si}$ have been observed from a continuous h-variate distribution function $G(\underset{\sim}{X},i)$ with location vector $\underset{\sim}{M}(i)$ where $\underset{\sim}{M}(i) = \underset{h\times r}{\underset{\sim}{B}_1}\ \underset{r\times 1}{\underset{\sim}{\alpha}_i}$, it is obvious that the $\underset{\sim}{B}_1\,\underset{\sim}{X}_{Si}$ have a continuous r-variate distribution function with location vector $\underset{\sim}{\alpha}_i$. Hence, the multisample rank scores procedure may be applied to the $\underset{\sim}{B}_1\,\underset{\sim}{X}_{Si}$. The procedure is simply to rank each coordinate of $\underset{\sim}{Y}_{Si}$ in the set of all the $\underset{\sim}{Y}_{Si}$ and apply a score function to these ranks. Mean scores are computed for each sample and to test the null hypothesis of equality of location vectors; a set of r(c-1) contrasts in the mean scores is constructed. A quadratic form in these r(c-1) contrasts is constructed. A quadratic form in these r(c-1) contrasts is defined to be the test statistic; numerically large values of this statistic lead to rejection of the null hypothesis. Referring specifically to Section 5.6 of Puri and Sen (1971, we note through the sequence H_N defined in (4.3.4) that the quadratic form would be noncentral chi-square with r(c-1) degrees of freedom with noncentrality Δ_1 as $N \to \infty$. The noncentrality Δ_1 is defined by

$$(4.4.1) \qquad \Delta_1 = \sum_{i=1}^{c} \gamma_i \, \underset{\sim}{\lambda}_i' \, \underset{\sim}{\tau}_{11}^{-1}(F) \, \underset{\sim}{\lambda}_i$$

where γ_i are defined in the previous section $i = 1,2,\ldots,c$ and

$$\underset{\sim}{\tau}_{11}(F) = \left(\left(\frac{\sigma_{jj'}}{c(F_j)c(F_{j'})}\right)\right)_{j,j'=1,2,\ldots,r}$$

and $\sigma_{jj'}$ is $\nu_{jj'}(F)$ defined by (5.4.28) of Puri and Sen (1971) and

$$c(F_j) = \int_{-\infty}^{\infty} \frac{d}{dX} J(F_j(X)) \, dF_j(X); \text{ for } j = 1,2,\ldots,r.$$

We have again assumed that $\bar{\underset{\sim}{\lambda}} = \underset{\sim}{0}$, which is no loss of generality. We observe that (4.4.1) can be written in its full rank form as:

$$(4.4.2) \quad \Delta_1 = [\underset{\sim}{\lambda}_i;\ i=1,2,\ldots,c-1]'[\underset{\sim}{\tau}_{11}^{-1}(F) \otimes \underset{\sim}{A}][\underset{\sim}{\lambda}_i;\ i=1,2,\ldots,c-1]$$

where $\underset{\sim}{A}$ is defined by (4.3.13). Hence, we have

$$(4.4.3) \qquad |\underset{\sim}{\tau}_{11}^{-1}(F) \otimes \underset{\sim}{A}| = |\underset{\sim}{\tau}_{11}^{-1}(F)|^{c-1} \; |A|^{r}$$

and

$$(4.4.4) \qquad \mathrm{tr}(\underset{\sim}{\tau}_{11}^{-1}(F) \otimes \underset{\sim}{A}) = \mathrm{tr}\ \tau_{11}^{-1}(\underset{\sim}{F})\ \mathrm{tr}\ \underset{\sim}{A}.$$

4.4.1 Nonparametric Procedures - Overfit

If we had overfit the polynomial model in each group with polynomials of degree t-1 then we would apply the rank scores procedures to the variables defined by (4.3.14). The quadratic form in the t(c-1) contrasts in rank scores would have noncentral chi-square distribution with t(c-1) degrees of freedom and noncentrality Δ_2 through H_N, where

$$\Delta_2 = \left[\begin{pmatrix}\boldsymbol{\lambda}_i\\ \boldsymbol{0}\end{pmatrix};\ i = 1,2,\ldots,c-1\right]' \left[\boldsymbol{\tau}^{-1}(F) \otimes \boldsymbol{A}\right] \left[\begin{pmatrix}\boldsymbol{\lambda}_i\\ \boldsymbol{0}\end{pmatrix};\ i = 1,2,\ldots,c-1\right] \tag{4.4.5}$$

and

$$\boldsymbol{\tau}(F) = \begin{pmatrix} \boldsymbol{\tau}_{11}(F) & \boldsymbol{\tau}_{12}(F) \\ \boldsymbol{\tau}_{21}(F) & \boldsymbol{\tau}_{22}(F) \end{pmatrix},$$

the covariance matrix of the entire set of t scores. In an obvious way we have

$$\Delta_2 = [\boldsymbol{\lambda}_i,\ i = 1,2,\ldots,c-1]' \left[(\boldsymbol{\tau}_{11}(F) - \boldsymbol{\tau}_{12}(F)\boldsymbol{\tau}_{22}^{-1}(F)\boldsymbol{\tau}_{21}(F)\right]^{-1} \otimes \boldsymbol{A}[\boldsymbol{\lambda}_i, i=1,2,\ldots,c-1].$$

The efficiencies of the incorrect test ϕ_2 to ϕ_1, the test based on the correct number of parameters using curvature criterion is:

$$\text{CARE} = R(r(c-1),t(c-1),\alpha) \left\{ \frac{|\boldsymbol{\tau}_{11}(F)|}{|\boldsymbol{\tau}_{11}(F) - \boldsymbol{\tau}_{12}(F)\boldsymbol{\tau}_{22}^{-1}(F)\boldsymbol{\tau}_{21}(F)|} \right\}^{1/r};\quad r \le t \tag{4.4.5}$$

and using trace criterion we obtain:

$$\text{TARE} = R(r(c-1),t(c-1),\alpha)\ \frac{\operatorname{tr}[\boldsymbol{\tau}_{11}(F) - \boldsymbol{\tau}_{12}(F)\boldsymbol{\tau}_{22}^{-1}(F)\boldsymbol{\tau}_{21}(F)]^{-1}}{\operatorname{tr}\ \boldsymbol{\tau}_{11}(F)^{-1}}\quad r \le t. \tag{4.4.6}$$

We note the similarity to the one-sample results and observe that the scalar adjustment, $R(r(c-1),t(c-1),\alpha)$ may result in large reductions when comparing many samples.

4.4.2 Nonparametric Procedure - Underfitting

Underfitting the model leads to applying the rank scores to the $\underset{\sim}{Y}_{Si}$ defined by (4.3.18). The quadratic form in the $t(c-1)$ contrasts of the mean rank scores would lead to a limiting noncentral chi-square distribution with $t(c-1)$ degrees of freedom and noncentrality parameter Δ_3 through H_N where,

$$\Delta_3 = [\underset{\sim}{\lambda}_i^*, \ i=1,2,\ldots,c-1]'[(\underset{\sim}{B}^{*'} \underset{\sim}{\Sigma} \underset{\sim}{B}^*)^{-1} \otimes \underset{\sim}{A}] \ [\underset{\sim}{\lambda}_i^*; \ i=1,2,\ldots,c-1],$$

$\underset{\sim}{\lambda}_i^*$ and $\underset{\sim}{B}^*$ are defined as in Section 4.3.2. The efficiency of the underfitted procedure to the correctly fitted procedure is:

$$\text{(4.4.7)} \qquad \text{TARE} = R(r(c-1),t(c-1),\alpha) \ \frac{\text{tr } \underset{\sim}{\tau}^{*-1}(F)}{\text{tr } \underset{\sim}{\tau}_{11}^{-1}(F)}$$

when $\underset{\sim}{\tau}^*(F)$ is the portion of $\underset{\sim}{\tau}_{11}(F)$, corresponding to the upper t×t portion of $\underset{\sim}{\tau}_{11}(F)$.

4.5 Comments

We conclude from Sections 4.3 and 4.4 that for large values of c, the adjustment of the one-sample efficiencies to obtain the c-sample efficiencies, may be quite large. The loss in efficiency in overfitting is greater for the c-sample problem than the corresponding one sample problem. On the other hand, the loss in efficiency in underfitting the correct model is not as great for the c-sample problem as it is for the one-sample problem.

A comparison of the parametric and nonparametric procedures would obviously lead to the same formulae as presented for the one-sample problem. This fact is deduced by simply observing that the function, $R(q,t,\alpha)$, is 1 when $q = t$ and the ratio of the determinants or traces of the noncentralities would be independent of the matrix $\underset{\sim}{A}$. Hence, the one-sample bounds obtained in Chapter III are applicable to the c-sample problem.

Adjustment of the statistics with the higher order terms as covariables would also lead to identical results as presented in Section 3.7 and the bounds attained in that section would apply for the c-sample problem as well.

CHAPTER V

NUMERICAL ILLUSTRATIONS OF THE CARE AND TARE

5.1 Introduction

The purpose of this chapter is to provide some numerical illustrations of the CARE and TARE in specific situations. We have seen in Chapters III and IV that the CARE and TARE for the polynomial growth curve model depend in general on: a) the covariance matrix, $\underset{\sim}{\Sigma}$, b) the number of parameters in the model, q, c) the number of time points, h, d) the use of covariates and e) the score function. We concern ourselves in this chapter with the parametric procedures only. We shall find, however, that the results in Section 5.3 for the parametric tests are similar to the results obtained for the Wilcoxon score. Sections 5.3 and 5.4 are devoted to the evaluation of the TARE and CARE for the problem of underfitting and overfitting, respectively. Particular attention is paid to the situation where the data vectors are observations at five equally spaced time points. Section 5.2 is included to show the redution of the TARE and CARE for the uniform correlation model, and a tabulation of the TARE in this case is given. The covariance matrix defined in (5.2.6) is the one which is used to provide numerical examples of the values of TARE and CARE.

5.2 Special Covariance Patterns

If the covariance matrix is uniform, for example,

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & & & \\ \vdots & & \ddots & & \\ \rho & & & & 1 \end{bmatrix} \tag{5.2.1}$$

then the TARE and CARE for the test of stationarity studied in detail in Chapter III do not depend on σ^2 or ρ. To demonstrate this we let $\underset{\sim}{B} = [\underset{\sim}{b}_1, \ldots, \underset{\sim}{b}_t]$, be the orthonormal polynomials of degrees 1 to t, then

$$\underset{\sim}{B}' \underset{\sim}{\Sigma} \underset{\sim}{B} = \sigma^2(1-\rho) \underset{\sim}{I}.$$

Consequently, the TARE and CARE for overfitting with t statistics relative to the test with q statistics are:

$$\left.\begin{aligned} \text{TARE} &= R(q,t,\alpha)\ t/q \\ \text{CARE} &= R(q,t,\alpha)\ . \end{aligned}\right. \qquad \text{for } t \geq q. \tag{5.2.2}$$

If $t < q$ then

$$\text{TARE} = R(q,t,\alpha)\ t/q. \tag{5.2.3}$$

The quantity (5.2.3) is tabulated in Table 5.2.1 with $q = t_1$, and $t = t_2$ in the notation used in that table. The entries in Table 5.2.1 are valid for all values of the number of time points, h, and are valid for both the parametric and Wilcoxon procedures discussed in Chapter III. Inspection of the table reveals that the efficiency is not greatly reduced for underfitting the model by only one or two parameters but is quite low for underfitting by a great deal.

As we pointed out in Chapter III in the concluding comments, the results of that chapter are applicable to other hypotheses of interest; e.g., we may test the hypothesis that the intercept is a fixed quantity, say α_{00} and that the curve is stationary. In this case the $\underset{\sim}{B}$ matrix is augmented by $\underset{\sim}{b}_0$. This hypothesis is of special interest since the c-sample efficiencies have been shown to be multiples of the corresponding one-sample efficiencies of that particular test. Defining $\underset{\sim}{B}$ as $\underset{\sim}{B} = [\underset{\sim}{b}_0, \underset{\sim}{b}_1, \ldots, \underset{\sim}{b}_q]$, we find that

$$(5.2.4) \qquad \underset{\sim}{B}' \underset{\sim}{\Sigma} \underset{\sim}{B} = \sigma^2 \begin{bmatrix} 1-\rho+h\rho & 0 & 0 & 0 \\ 0 & 1-\rho & 0 & 0 \\ \vdots & 0 & \vdots & \vdots \\ \vdots & \vdots & & 0 \\ 0 & 0 & & 1-\rho \end{bmatrix} .$$

Hence, the TARE and CARE are independent of σ^2; however, the TARE depends in general on ρ. Inspection of (5.2.4) and its determinant readily show that the CARE would be independent of ρ. In fact, the quantity (5.2.2) would define the CARE for this problem. We remark also that covariance adjustment by the higher order polynomial estimates does not alter the results discussed up to this point in this section since $\underset{\sim}{B}' \underset{\sim}{\Sigma} \underset{\sim}{B}$ is diagonal in both cases. While the CARE does not depend on ρ for the second hypothesis, the TARE does depend on ρ. With some algebra the TARE of the test with t statistics to the test with q statistics can be shown to be:

$$(5.2.5) \qquad \text{TARE} = R(q,t,\alpha)[t(1-\rho) + h(t-1)\rho]/[q(1-\rho) + h(q-1)\rho].$$

We have not tabulated values of (5.2.5) but it may be of interest to do so in future work. The point to be made is that we get different efficiency results by including the intercept in the hypothesis.

A model considered in time series analysis is the first order auto-regressive model, with covariance structure defined by (5.2.6).

$$(5.2.6) \qquad \underset{\sim}{\Sigma} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & \cdots & \rho^{h-1} \\ \rho & 1 & \rho & & & \vdots \\ \rho^2 & \rho & 1 & & & \vdots \\ \vdots & & & 1 & & \\ \vdots & & & & \ddots & \\ \rho^{h-1} & \cdots & \cdots & \cdots & \rho & 1 \end{pmatrix} .$$

In section 5.3 and 5.4 we shall evaluate the TARE and CARE for this correlation pattern for various values of ρ and h. The value of σ^2 is, of course, immaterial since the TARE and CARE are independent of σ^2. It is known that the inverse of $\underset{\sim}{\Sigma}$, defined by (5.2.6) is:

$$\underset{\sim}{\Sigma}^{-1} = [\sigma^2(1-\rho^2)]^{-1} \begin{pmatrix} 1 & -\rho & 0 & & \cdots & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \cdots & 0 \\ 0 & & & & & \\ \vdots & & & & & \\ 0 & & & & -\rho & 1 \end{pmatrix}$$

which is useful in computing the covariance matrix of the weighted least squares estimates. This covariance matrix is, of course, given by $(\underset{\sim}{B}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{B})^{-1}$. When the underlying distribution function is normal, the covariance matrix $(\underset{\sim}{B}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{B})^{-1}$ is identical to the covariance matrix for the maximum likelihood estimates. Hence, we see that $(\underset{\sim}{B}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{B})$ is the matrix in the noncentrality parameter of these statistics. In Section 5.3 we compare the underfitting procedure based on the unweighted least squares estimates to the trace of $(\underset{\sim}{B}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{B})$.

TABLE 5.2.1

VALUES OF TARE FOR UNDERFITTING FOR UNIFORM COVARIANCE PATTERNS

	t_2													
t_1	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	.76													
3	.65	.86												
4	.58	.76	.89											
5	.54	.70	.82	.92										
6	.50	.65	.76	.85	.93									
7	.47	.61	.72	.80	.87	.94								
8	.44	.58	.68	.76	.83	.89	.95							
9	.42	.55	.65	.72	.79	.85	.90	.95						
10	.40	.53	.62	.69	.75	.81	.86	.91	.96					
11	.39	.51	.59	.66	.73	.78	.83	.88	.92	.96				
12	.37	.49	.57	.64	.70	.75	.80	.85	.89	.93	.96			
13	.36	.47	.55	.62	.68	.73	.77	.82	.86	.90	.93	.97		
14	.35	.46	.54	.60	.66	.71	.75	.79	.83	.87	.90	.94	.97	

Note: Entries in the table are TARE of the incorrect test with t_2 degrees of freedom to the test specifying the correct number of statistics, t_1. Values in the table were computed with $\alpha = .05$.

5.3 Underfitting

Because of the relationship between the multi-sample problem and the one-sample problem for the hypothesis of a specified intercept and a stationary or constant response over time, we shall consider this hypothesis in this section. This section specifically considers the evaluation of the TARE of the unweighted least squares procedure which underfits the model with respect to the unweighted least squares procedure which specifies the correct number of parameters. We also consider the TARE when covariance adjustment has been used in the underfitting procedure. These two comparisons are analogous to two possible procedures one may follow in practice; the first procedure simply being an unweighted least squares with no covariables, while the second is the unweighted least squares estimates with all remaining higher order polynomial estimates used as covariables. In addition to comparing these two underfitting procedures to the unweighted least squares procedure based on the correct number of parameters, we also compare them to the weighted least squares procedure based on the correct number of parameters. The covariance matrix (5.2.6) was used for the covariance matrix of the original observations. Tables of the ratio of the trace of the noncentrality parameters were generated for $\rho = \pm.1, \pm.5, \pm.9$ and $\pm.95$ for five and also ten equally spaced time points. For five time points these figures are found in Tables 5.3.1 - 5.3.8, two illustrations for ten time points are given in Tables 5.3.9 and 5.3.10. To compute the TARE one only needs to obtain the value of R corresponding to the α-level and the number of samples and multiply by the appropriate number selected from the tables. Values of R are found in Chapter II. While the numerical results in Tables 5.3.1 - 5.3.10 are for the parametric test procedures, the corresponding tabulation for the Wilcoxon score yields nearly identical results for the figures below the diagonals. The maximum difference was, in fact, .01. Each figure below the diagonal is the ratio of the trace of the matrix in the noncentrality parameter of the test based on the unweighted least squares procedure with too few parameters to the corresponding trace for the unweighted least squares test with the correct number of parameters. For this problem t_1, in the tables is the number of parameters in the true model while t_2 is the number of parameters assumed. In the upper triangle of the tables, we compare the underfitted tests to the test using the weighted least squares reduction. The roles of t_1 and t_2 are thus reversed in this comparison. The entries above the diagonal in the 'a' tables are therefore:

$$\text{(5.3.1)} \qquad \operatorname{tr}(\underset{\sim}{B}'_{t_1} \underset{\sim}{\Sigma} \underset{\sim}{B}_{t_1})^{-1}/\operatorname{tr}(\underset{\sim}{B}'_{t_2} \underset{\sim}{\Sigma}^{-1} \underset{\sim}{B}_{t_2})$$

and below the diagonal are:

$$\text{(5.3.2)} \qquad \operatorname{tr}(\underset{\sim}{B}'_{t_1} \underset{\sim}{\Sigma} \underset{\sim}{B}_{t_1})^{-1}/\operatorname{tr}(\underset{\sim}{B}'_{t_2} \underset{\sim}{\Sigma} \underset{\sim}{B}_{t_2})^{-1}.$$

The entries in the 'b' tables have identical denominators, but the covariance matrices in the numerator have been changed to:

$$\text{(5.3.3)} \quad \operatorname{tr}(\underset{\sim}{B}'_{t_1} \underset{\sim}{\Sigma} \underset{\sim}{B}_{t_1} - \underset{\sim}{B}'_{t_1} \underset{\sim}{\Sigma} \underset{\sim}{B}_{h-t_1}(\underset{\sim}{B}'_{h-t_1} \underset{\sim}{\Sigma} \underset{\sim}{B}_{h-t_1})^{-1} \underset{\sim}{B}'_{h-t_1} \underset{\sim}{\Sigma} \underset{\sim}{B}_{t_1})^{-1}.$$

$\underset{\sim}{B}_{h-t_1}$ is used to denote the matrix of orthonormal polynomials of degree t_1 to (h-1).

Examining the tables we see that for $\rho = \pm.1$ covariance adjustment does not change the ratio of traces of the underfit to either procedure. For all t_1 and t_2 the ratio of traces is quite low when h = 5. When h = 10 the ratio is near .90 if the degree of the true model is high and the degree of the incorrect test is only one less. For $\rho = .5$ and h = 5 or 10 slight improvements are observed using covariance adjustment; these are small, however. The ratio of traces improves by as much as .07 with covariance adjustment for $\rho = -.5$. If ρ is .9 or .95 the ratio of traces is extremely small whether covariance adjustment is used or not. Nearly all of the ratios of traces are less than .50 in this case and many are near zero. Hence, for large positive correlations the loss in the TARE in underfitting the model is substantial. For large negative values of ρ we see that covariance adjustment greatly improves the efficiency results. For example, in comparing the test based on the unweighted least squares with two parameters to the test based on unweighted squares with three, we see that the ratio of traces is .44. Adjustment of the former test with the three higher order terms as covariables increases the ratio of traces to 1.06. This is not a surprising result since the test with covariance adjustment is providing information about the underlying covariance matrix which the unweighted least squares procedure does not obtain. Marked improvements are also noted for the adjusted test as compared to the weighted least squares procedure. When adjustment is made for the factor R, one will find that the TARE will in several cases be greater than one for the test based on fewer degrees of freedom relative to the test based on a larger number of degrees of freedom. This can be explained by one of two possibilities in the case of comparing the unweighted least squares tests: either the covariance is such that not weighting causes a reduction in the trace of the matrix in the expression for the noncentrality or the increase in the size of the trace of the test with more degrees of freedom is too small to account for the loss in sensitivity of a test with more degrees of freedom. The interpretation of this must be made in light of the fact that we have considered the local power functions which means the alternatives are close to the null point. No generalization to alternatives at a greater distance from the null point should be made since the truncated power function is not an adequate approximation to the entire power function for these alternatives. A fuller discussion of the relationship between the degrees of freedom and the noncentrality parameter is found in Krishnaiah (1966, pp. 91-92). The interpretation of the increase in the TARE of the test with fewer degrees of freedom relative to the test with more degrees of freedom based on the weighted least squares procedure lies simply in the fact that the gain in degrees of freedom of the latter test has not been accompanied by a sufficiently large increase in the trace of the matrix in its noncentrality. The weighted least squares procedure which is equivalent to the maximum likelihood procedure in the normal theory case is the uniformly most powerful test in the class of invariant tests for the one-sample problem. Underspecifying the true model results in choosing estimates which are singular transformations of the estimates for the correct number of parameters. Hence, tests constructed for the underfitted model do not belong to the same invariant class as the weighted least squares test statistics. According to a result of Wald (1943) the test based on the maximum likelihood statistics (weighted estimates) has best average power over the family of ellipsoids defined by equating its noncentrality parameter to a constant. Integration over the family of spheres for local alternatives weights all the parameters equally, while in using the ellipsoids as the surface of integration the weights are determined by the covariance matrix. Since the likelihood criterion provides the best test in the sense of best average power on a given family of ellipsoids, it may not perform as well when the surfaces of integration are spheres and the information matrix is not proportional to $\underset{\sim}{I}$.

Care must be used in the interpretation of the efficiency results presented in this section because we have compared the tests in a very restricted manner. First, we have allowed only local alternatives and, second, we have taken <u>average</u> power in this local sense. While the test based on fewer degrees of freedom may have better average local power, it is at the same time inferior in those directions which have not been fit.

In conclusion, we see the numerical results show that we never lose by using covariance adjustment of the primary variates with the higher order polynomial terms. Furthermore, there are covariance matrices that show great improvements in the TARE by use of the covariance technique. While underfitting the model is not desirable, the results in this section seem to suggest that the loss in the average local power is not large for certain covariance matrices. For others, the loss is severe and we see that the decision to include or exclude higher degree polynomial terms depends to some extent on the underlying covariance pattern.

TABLE 5.3.1

RATIO OF TRACES OF NONCENTRALITY PARAMETERS FOR UNDERFITTING: FIRST ORDER MODEL $\rho = .1$, $h = 5$

a. Without Covariance Adjustment

	t_2				
t_1	1	2	3	4	5
1		.48	.30	.22	.17
2	.48		.64	.46	.35
3	.31	.64		.72	.55
4	.22	.46	.72		.77
5	.17	.35	.55	.77	

b. With Covariance Adjustment

	t_2				
t_1	1	2	3	4	5
1		.48	.31	.22	.17
2	.48		.64	.46	.35
3	.31	.64		.72	.55
4	.22	.46	.72		.77
5	.17	.35	.55	.77	

Note: Entries are defined by (5.3.1), (5.3.2) and (5.3.3).

TABLE 5.3.2

RATIO OF TRACES OF NONCENTRALITY PARAMETERS FOR UNDERFITTING: FIRST ORDER MODEL $\rho = -.1$, $h = 5$

a. Without Covariance Adjustment

t_1	t_2 1	2	3	4	5
1		.52	.36	.28	.23
2	.52		.69	.54	.45
3	.36	.69		.78	.65
4	.28	.54	.78		.83
5	.23	.45	.65	.33	

b. With Covariance Adjustment

t_1	t_2 1	2	3	4	5
1		.52	.36	.28	.23
2	.52		.69	.54	.45
3	.36	.69		.78	.65
4	.28	.54	.78		.83
5	.23	.45	.65	.83	

Note: See note for Table 5.3.1.

TABLE 5.3.3

RATIO OF TRACES OF NONCENTRALITY PARAMETERS FOR UNDERFITTING: FIRST ORDER MODEL $\rho = .5$, $h = 5$

a. Without Covariance Adjustment

t_1	t_2 1	2	3	4	5
1		.34	.16	.09	.06
2	.35		.46	.26	.17
3	.16	.47		.56	.36
4	.09	.27	.56		.64
5	.06	.17	.36	.64	

b. With Covariance Adjustment

t_1	t_2 1	2	3	4	5
1		.35	.17	.09	.06
2	.36		.47	.27	.17
3	.17	.48		.57	.37
4	.09	.27	.57		.64
5	.06	.17	.37	.64	

Note: See note for Table 5.3.1.

TABLE 5.3.4

RATIO OF TRACES OF NONCENTRALITY PARAMETERS FOR UNDERFITTING: FIRST ORDER MODEL $\rho = -.5$, $h = 5$

a. Without Covariance Adjustment

	t_2				
t_1	1	2	3	4	5
1		.52	.39	.33	.31
2	.59		.67	.57	.52
3	.43	.73		.78	.72
4	.34	.59	.81		.89
5	.31	.52	.72	.89	

b. With Covariance Adjustment

	t_2				
t_1	1	2	3	4	5
1		.57	.43	.37	.34
2	.65		.75	.64	.59
3	.47	.82		.85	.78
4	.38	.66	.88		.92
5	.34	.59	.78	.92	

Note: See note for Table 5.3.1.

TABLE 5.3.5

RATIO OF TRACES OF NONCENTRALITY PARAMETERS FOR UNDERFITTING: FIRST ORDER MODEL $\rho = .9$, $h = 5$

a. Without Covariance Adjustment

t_1	t_2: 1	2	3	4	5
1		.09	.02	.01	.01
2	.10		.25	.11	.06
3	.03	.26		.43	.24
4	.01	.11	.43		.56
5	.01	.06	.24	.56	

b. With Covariance Adjustment

t_1	t_2: 1	2	3	4	5
1		.09	.03	.01	.01
2	.10		.27	.12	.07
3	.03	.27		.44	.25
4	.01	.12	.44		.57
5	.01	.07	.25	.57	

Note: See note for Table 5.3.1.

TABLE 5.3.6

RATIO OF TRACES OF NONCENTRALITY PARAMETERS FOR UNDERFITTING: FIRST ORDER MODEL $\rho = -.9$, $h = 5$

a. Without Covariance Adjustment

	t_2				
t_1	1	2	3	4	5
1		.19	.15	.13	.12
2	.45		.32	.28	.27
3	.20	.44		.64	.61
4	.14	.32	.72		.84
5	.12	.27	.61	.84	

b. With Covariance Adjustment

	t_2				
t_1	1	2	3	4	5
1		.61	.48	.41	.39
2	1.48		.78	.68	.65
3	.65	1.06		.87	.83
4	.47	.77	.99		.95
5	.39	.65	.83	.95	

Note: See note for Table 5.3.1.

TABLE 5.3.7

RATIO OF TRACES OF NONCENTRALITY PARAMETERS FOR UNDERFITTING: FIRST ORDER MODEL $\rho = -.95$, $h = 5$

a. Without Covariance Adjustment

	t_2				
t_1	1	2	3	4	5
1		.10	.08	.07	.06
2	.32		.24	.21	.20
3	.11	.33		.62	.59
4	.07	.24	.71		.83
5	.06	.20	.59	.83	

b. With Covariance Adjustment

	t_2				
t_1	1	2	3	4	5
1		.61	.48	.42	.40
2	2.01		.78	.68	.65
3	.67	1.10		.87	.83
4	.48	.78	.99		.95
5	.40	.65	.83	.95	

Note: See note for Table 5.3.1.

TABLE 5.3.8

RATIO OF TRACES OF NONCENTRALITY PARAMETERS FOR UNDERFITTING: FIRST ORDER MODEL $\rho = .95$, $h = 5$

a. Without Covariance Adjustment

t_1	t_2 1	2	3	4	5
1		.05	.01	.01	.00
2	.05		.23	.10	.06
3	.01	.24		.41	.23
4	.00	.10	.42		.56
5	.00	.06	.23	.56	

b. With Covariance Adjustment

t_1	t_2 1	2	3	4	5
1		.05	.01	.01	.00
2	.05		.24	.10	.06
3	.01	.25		.42	.24
4	.01	.10	.43		.56
5	.00	.06	.24	.56	

Note: See note for Table 5.3.1.

TABLE 5.3.9

RATIO OF TRACES OF NONCENTRALITY PARAMETERS FOR UNDERFITTING: FIRST ORDER MODEL ρ = .1, h = 10

a. Without Covariance Adjustment

	t_2									
t_1	1	2	3	4	5	6	7	8	9	10
1		.49	.32	.23	.18	.15	.12	.11	.09	.08
2	.49		.65	.48	.37	.30	.25	.22	.19	.17
3	.32	.65		.73	.57	.47	.39	.33	.29	.26
4	.29	.48	.73		.78	.64	.53	.46	.40	.35
5	.18	.37	.57	.78		.81	.68	.58	.51	.45
6	.15	.30	.47	.64	.81		.84	.72	.62	.55
7	.12	.25	.39	.53	.68	.84		.86	.75	.66
8	.11	.22	.33	.46	.58	.72	.86		.87	.77
9	.09	.19	.29	.40	.51	.62	.75	.87		.88
10	.08	.17	.26	.35	.45	.55	.66	.77	.88	

b. With Covariance Adjustment

	t_2									
t_1	1	2	3	4	5	6	7	8	9	10
1		.49	.32	.23	.18	.15	.12	.11	.09	.08
2	.49		.65	.48	.37	.30	.26	.22	.19	.17
3	.32	.65		.73	.57	.47	.39	.34	.29	.26
4	.23	.48	.74		.78	.64	.53	.46	.40	.35
5	.18	.37	.57	.78		.81	.68	.59	.51	.45
6	.15	.30	.47	.64	.82		.84	.72	.63	.55
7	.13	.26	.39	.53	.68	.84		.86	.75	.66
8	.11	.22	.34	.46	.59	.72	.86		.87	.77
9	.09	.19	.29	.40	.51	.63	.75	.87		.88
10	.08	.17	.26	.35	.45	.55	.66	.77	.88	

Note: See note for Table 5.3.1.

TABLE 5.3.10

RATIO OF TRACES OF NONCENTRALITY PARAMETERS FOR UNDERFITTING: FIRST ORDER MODEL $\rho = -.95$, $h = 10$

a. Without Covariance Adjustment

	t_2									
t_1	1	2	3	4	5	6	7	8	9	10
1		.34	.25	.20	.17	.15	.14	.13	.13	.12
2	.85		.29	.23	.20	.18	.16	.15	.15	.15
3	.45	.53		.44	.38	.33	.31	.29	.28	.28
4	.27	.32	.60		.63	.56	.51	.48	.46	.46
5	.20	.24	.45	.75		.74	.68	.64	.62	.61
6	.17	.19	.37	.62	.82		.83	.78	.75	.75
7	.15	.17	.33	.54	.72	.88		.88	.85	.84
8	.14	.16	.30	.50	.66	.81	.92		.93	.92
9	.13	.15	.28	.47	.63	.77	.87	.95		.97
10	.12	.15	.28	.46	.61	.75	.84	.92	.97	

b. With Covariance Adjustment

	t_2									
t_1	1	2	3	4	5	6	7	8	9	10
1		.55	.40	.32	.27	.24	.22	.21	.20	.20
2	1.37		.73	.58	.49	.44	.40	.38	.37	.36
3	.72	1.31		.80	.68	.60	.55	.52	.50	.50
4	.43	.79	1.08		.85	.75	.69	.65	.63	.62
5	.33	.59	.82	1.02		.89	.81	.76	.74	.73
6	.27	.48	.67	.83	.98		.92	.86	.83	.82
7	.24	.43	.59	.73	.86	.98		.94	.91	.90
8	.22	.39	.54	.68	.79	.90	.98		.97	.96
9	.21	.37	.51	.64	.75	.85	.93	.99		.99
10	.20	.36	.50	.62	.73	.82	.90	.96	.98	

Note: See note for Table 5.3.1.

5.4 Overfitting

We present numerical results for the CARE of the parametric test that overspecifies the model relative to the parametric test which correctly specifies the model for h equal to five. We consider the unweighted least squares procedure for both tests. We restrict attention to the first order auto-regressive model and the hypothesis discussed in Section 5.3. In Tables 5.4.1 - 5.4.4 we have tabulated the $t_1{}^{th}$ root of the ratio of the determinants above the diagonal, while below the diagonal we have the CARE of the overspecified test to the test based on the correct number of parameters for the one sample problem (or two sample problem) for $\alpha = .05$. Hence, above the diagonals are:

$$(5.4.1) \qquad D = \{|\underset{\sim}{B}'_{t_1} \underset{\sim}{\Sigma} \underset{\sim}{B}_{t_1}| / |\underset{\sim}{B}'_{t_1} \underset{\sim}{\Sigma} \underset{\sim}{B}_{t_1} - \underset{\sim}{B}'_{t_1} \underset{\sim}{\Sigma} \underset{\sim}{B}_{t_2-t_1} (\underset{\sim}{B}_{t_2-t_1} \underset{\sim}{\Sigma} \underset{\sim}{B}_{t_2-t_1})^{-1} \underset{\sim}{B}'_{t_2-t_1} \underset{\sim}{\Sigma} \underset{\sim}{B}_{t_1}|\}^{1/t_1},$$

and below the diagonal are:

$$(5.4.2) \qquad \text{CARE} = R(t_1, t_2, .05)D.$$

We have tabulated the results for $\rho = \pm.5, \pm.90$. The tables were generated for other values of ρ; however, the results for intermediate values of ρ are between the values listed in the tables. Small values of $|\rho|$ result in D, defined by (5.4.1), being very close to one. The comparison to the test based on the weighted least squares reduction yields similar values of the CARE for $|\rho| \leq .5$. For larger values of $|\rho|$ the CARE for comparison to the weighted least squares procedure is much lower (see Table 5.4.4). We note that for $\rho = \pm.5$ the CARE for overfitting is about .40 for fitting five terms when only one is needed. If we overfit a four parameter model by a five parameter model the CARE is about .90. When $\rho = +.9$ we obtain results identical with the case where $\rho = +.5$, while for $\rho = -.9$ we see that overfitting in most cases produces a CARE greater than one. This most probably is caused by the fact that the overfitting yields information about the covariance matrix not used in the unweighted least squares procedure which specifies the correct number of parameters. Comparison to the weighted least squares procedure based on the correct number of parameters for $\rho = -.9$ produces rather low values of the CARE. On the other hand, with underfitting in this same case we have a loss in the ratio of the traces of the noncentralities; however, a large portion of this loss is recovered by covariance adjustment (see Table 5.3.6). Furthermore, all the numbers in Table 5.3.6 will increase when the adjustment is made for R. This suggests that overfitting has more loss associated with it than does underfitting for certain covariance matrices.

TABLE 5.4.1

VALUES OF D AND CARE FOR OVERFITTING $\rho = .5$

t_1	t_2				
	1	2	3	4	5
1		1.00	1.04	1.04	1.04
2	.65		1.02	1.04	1.04
3	.53	.79		1.01	1.02
4	.44	.68	.85		1.00
5	.39	.59	.75	.88	

Note: Entries above diagonal are defined by (5.4.1).
Entries below diagonal are defined by (5.4.2).

TABLE 5.4.2

VALUES OF D AND CARE FOR OVERFITTING $\rho = -.5$

t_1	t_2				
	1	2	3	4	5
1		1.00	1.10	1.10	1.11
2	.65		1.05	1.13	1.13
3	.56	.82		1.05	1.12
4	.47	.74	.88		1.05
5	.41	.65	.82	.91	

Note: See note for Table 5.4.1.

TABLE 5.4.3

VALUES OF D AND CARE FOR OVERFITTING $\rho = +.9$

	t_2				
t_1	1	2	3	4	5
1		1.00	1.03	1.03	1.03
2	.65		1.02	1.04	1.04
3	.53	.79		1.02	1.02
4	.44	.68	.85		1.01
5	.39	.59	.75	.88	

Note: See note for Table 5.4.1.

TABLE 5.4.4

VALUES OF D AND CARE FOR OVERFITTING $\rho = -.9$

	t_2				
t_1	1	2	3	4	5
1		1.00	3.15	3.15	3.25
2	.65(.20)		1.77	2.34	2.38
3	1.60(.58)	1.38(.49)		1.20	2.19
4	1.35(.46)	1.53(.64)	1.01(.41)		1.56
5	1.22(.87)	1.36(.73)	1.60(.57)	1.37(.37)	

Note: See note for Table 5.4.1. Numbers in parenthesis are the CARE of the overfit with respect to the weighted least squares test.

CHAPTER VI

SUMMARY AND SUGGESTIONS FOR FURTHER RESEARCH

We have considered in this study two measures of ARE which may be used when comparing two statistics with limiting chi-square distributions. The two quantities, the CARE and TARE, may be used when the chi-square distributions have different degrees of freedom. The CARE selects that test of the two competing tests whose power function has the greater generalized Gaussian curvature at the null point. On the other hand, the TARE selects the test whose power function has the greater average local power over the family of spheres. Both the CARE and TARE depend on the degrees of freedom of the tests, the significance level and the noncentrality parameters. The CARE and TARE computing formulae have been derived for the one-sample and the multi-sample polynomial growth curve problems. Particular attention has been paid to the formulae for the underfitting and overfitting problems. Numerical examples of the CARE and the TARE have also been computed to provide some idea of the efficiency results for some special cases.

Several areas of future work may be suggested. In the area of applications it would be useful to study the CARE and TARE for examples other than the ones chosen in Chapter V. In addition, the work of Chapters III and IV should be generalized to include more complicated designs. Hopefully, these results would reduce to the one-sample problem as the multi-sample problem did so the same bounds on the TARE and

CARE would apply, excepting the factor R. It seems reasonable that the noncentrality parameters for the more complicated models would factor as the noncentrality parameters for the multi-sample problem did.

In addition to the growth curve problem it would be desirable to study other applied problems. The results of Chapter II are in no way restricted to the study of growth curve models, but may be applied to virtually any situation where a reduction to a set of summary statistics is made and we want to determine the efficiency of the reduction. For example, if the course of a disease is characterized by a stochastic process and it is of interest to compare different groups of persons in this disease process then maybe certain summary statistics define or describe the process. One way to compare them is to reduce each person's observations to the basic summary statistics by some suitable estimation procedure then compare the summary statistics of the different groups. A natural question arises as to whether the reduction is a 'good' reduction to use. The basic approach used in this study would be helpful in this problem.

Another area of possible future work is to consider other efficiency criteria. From the results obtained in Chapter V for the underfitting problem, we see that it may be meaningful to define the TARE as the ratio of the average local powers where the average is taken over the family of ellipsoids defined by the noncentrality of the likelihood ratio test.

In addition, the restriction of local alternatives makes the TARE a somewhat restricted measure of ARE.

REFERENCES

[1] Bahadur, R. R. (1960). "Stochastic comparison of tests," <u>Annals of Mathematical Statistics</u>, <u>31</u>, 276-295.

[2] Blomqvist, N. (1950). "On a measure of dependence between two random variables," <u>Annals of Mathematical Statistics</u>, <u>21</u>, 593-600.

[3] Bickel, P. J. (1964). "On some alternative estimates for shift in the p-variate one-sample problem," <u>Annals of Mathematical Statistics</u>, <u>35</u>, 1079-1090.

[4] Bickel, P. J. (1965). "On some asymptotically nonparametric competitors to Hotelling's T^2," <u>Annals of Mathematical Statistics</u>, <u>36</u>, 160-173.

[5] Gastwirth, J. L. and Wolff, S. S. (1968). "An elementary method for obtaining lower bounds on the asymptotic power of rank tests," <u>Annals of Mathematical Statistics</u>, <u>39</u>, 2128-2131.

[6] Graybill, F. A. (1961). <u>An Introduction to Linear Statistical Models, Volume 1</u>, McGraw-Hill Book Company, Inc., New York.

[7] Graybill, F. A. (1969). <u>Introduction to Matrices with Applications in Statistics</u>, Wadsworth Publishing Company, Inc., Belmont, California.

[8] Grizzle, J. E. and Allen, D. (1969). "Analysis of growth and dose response curves," <u>Biometrics</u>, <u>25</u>, 357-382.

[9] Hodges, J. L., Jr. and Lehmann, E. L. (1956). "The efficiency of some nonparametric competitors of the t-test," <u>Annals of Mathematical Statistics</u>, <u>27</u>, 324-335.

[10] Isaacson, S. L. (1951). "On the theory of unbiased tests of simple statistical hypotheses specifying the values of two or more parameters," <u>Annals of Mathematical Statistics</u>, <u>22</u>, 217-234.

[11] Kendall, M. and Stuart, A. (1963). <u>The Advanced Theory of Statistics, Volume II</u>, Second Edition, Hafner Publishing Company, New York.

[12] Krishnaiah, P. R. (1966). Multivariate Analysis, Academic Press, New York.

[13] Lehmann, E. L. (1959). Testing Statistical Hypotheses, John Wiley and Sons, Inc., New York.

[14] Neyman, J. and Pearson, E. S. (1936). "Contributions to the theory of testing statistical hypotheses I. Unbiased critical regions of type A and type A_1," Statistical Research Memoirs, Volume 1, 1-37.

[15] Neyman, J. and Pearson, E. S. (1938). "Contributions to the theory of testing statistical hypotheses," Statistical Research Memoirs, Volumes II and III, 25-57.

[16] Noether, G. E. (1955). "On a theorem of Pitman," Annals of Mathematical Statistics, 26, 64-68.

[17] Potthoff, R. F. and Roy, S. N. (1964). "A generalized multivariate analysis of variance model useful especially for growth curve problems," Biometrika, 51, 313-326.

[18] Puri, M. L. and Sen, P. K. (1971). Nonparametric Methods in Multivariate Analysis, John Wiley and Sons, Inc., New York.

[19] Sen, P. K. and Puri, M. L. (1970). "Asymptotic theory of likelihood ratio and rank order tests in some multivariate linear models," Annals of Mathematical Statistics, 41, 87-100.

[20] Wald, A. (1943). "Tests of statistical hypotheses concerning several parameters when the number of observations is large," Transactions of the American Mathematical Society, 54, 426-482.

[21] Weyl, H. (1939). "On the volume of tubes," American Journal of Mathematics, 461.