# INVARIANT TESTS FOR SCALE PARAMETERS UNDER ELLIPTICAL SYMMETRY

by

Margaret A. Chmielewski

Dissertation submitted to the Graduate Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Statistics

APPROVED:

Donald R. Jensen

Donald R. Jensen, Chairman

Raymond H. Myers

Raymond H. Myers

Jesse C. Arnold

Jesse C. Arnold

Marion R. Reynolds, Jr.

Marion R. Reynolds, Jr.

Robert S. Schulman

Robert S. Schulman

November, 1978

Blacksburg, Virginia

ACKNOWLEDGMENTS

The author expresses her thanks and appreciation to the following people.

To my husband Joseph for his love, patience, and encouragement.

To Dr. Donald R. Jensen for suggesting the problem and for offering the needed help and encouragement.

To Drs. Robert V. Foutz and Robert S. Schulman for co-reading the manuscript.

To the Statistics Department for four years of financial support.

TABLE OF CONTENTS

TABLE OF CONTENTS (cont.)

# I. INTRODUCTION

In parametric statistical inference it often is assumed that observations are independent and Gaussian. The Gaussian assumption sometimes is justified on appeal to central limit theory or on the grounds that certain normal-theory procedures are robust. The independence assumption, usually unjustified, routinely facilitates the derivation of needed distribution theory.

In many situations it is known that the data are non-Gaussian. In economics, for example, price change data are known to have distributions with heavier tails than the Gaussian law [cf. Mandelbrot (1963)]. In other circumstances it is known that observations are dependent; for example, randomizations often used in designed experiments induce dependencies among the observations.

In this thesis, we consider tests for hypotheses pertaining to a class of distributions in which the observations are not necessarily either Gaussian or independent. These distributions are the spherically symmetric vector laws, with natural extensions of these to laws of random matrices. A random vector $\underline{x}$ of order $(n \times 1)$ is said to have a <u>spherically symmetric distribution</u> if and only if the distribution of $\underline{x}$ is the same as that of $\underline{Px}$ for all $(n \times n)$ orthogonal matrices $\underline{P}$. Such distributions have been proposed for use in several areas of application. Some models in communications systems can be represented in terms of independent spherically symmetric vectors [cf. Goldman (1976)]. In bombing problems the probability of a hit inside a circle centered at the origin is of some interest. Gilliland (1968) discusses spherically symmetric impact distributions for these problems. Other distributions have been modeled as scale mixtures of spherical Gaussian laws; Zellner (1976) discusses a stock market model in this context.

Further motivation for considering spherically symmetric laws stems from central limit theory. The limiting distribution of a normalized sum of independent random vectors, if it exists, is a multivariate stable law [cf. Butzer and Hahn (1978), for example]. Results established for spherically symmetric laws thus apply approximately in large samples to all distributions attracted to the spherically symmetric stable laws, this being a much larger class than those attracted to spherical Gaussian distributions under finite second moments.

In this thesis, we investigate various tests for scale parameters in the case of spherically symmetric vector and matrix distributions. Known tests for location parameters are discussed in the literature review. By combining the natural symmetries of the testing problems together with the symmetries of the underlying distributions, we are able to characterize all invariant tests for all distributions belonging to these classes. In contrast to some other methods of data reduction, the principle of invariance here obviates the need for an individual treatment of each distribution in a class and, instead, requires only that membership in the class be asserted.

The class of spherically symmetric distributions contains such heavy-tailed distributions as the spherical Cauchy law. As such laws need not have moments, the emphasis here is on tests for scale parameters which, if second-order moments exist, may be interpreted as tests for dispersion parameters.

A synopsis of this thesis follows. Chapter II contains a review of the literature, while Chapter III discusses elliptically symmetric distributions for both random vectors and random matrices. In Chapter IV we review the principle of invariance as it applies to certain hypothesis testing problems. In Chapters V through IX many standard normal-theory testing problems are revisited within the larger class of distributions. These problems include tests for the equality of k scale parameters, tests for the equality of k scale matrices, tests for sphericity, tests for block diagonal structure, tests for the uncorrelatedness of two variables within a set of m variables, and tests for the hypothesis of equi-correlatedness.

In all cases except the last three the null and non-null distributions of certain invariant test statistics are shown to be unique for all elliptically symmetric laws. Normal-theory procedures usually associated with these specific testing problems thus are exactly robust, and many known properties of these procedures extend directly to the larger class of distributions.

In the last three cases, the null distributions of certain invariant statistics are unique but the non-null distributions may depend on the underlying elliptically symmetric law. In tests for block diagonal structure, in the case of two blocks, we establish a monotone power property for the subclass of all elliptically symmetric unimodal distributions.

# II. LITERATURE REVIEW

Three of the earliest papers pertaining to spherically symmetric distributions are those of Maxwell (1860), Bartlett (1934) and Hartman and Wintner (1940). They all give the following characterization of the Gaussian law. Let $\underline{x}' = (x_1, \ldots, x_n)$ be a sample of n independently and identically distributed random variables. Then $\underline{x}$ has a spherically symmetric distribution if and only if $\underline{x}$ has a Gaussian law. More recently this characterization is noted by Kelker (1970), Thomas (1970), and Nash and Klamkin (1976).

One of the early modern papers is that of Lord (1954). He derives the characteristic function, $\phi_{\underline{x}}(\underline{t})$, when $\underline{x}$ has a spherically symmetric distribution. He also discusses the convolution of two independent spherically symmetric vectors and derives the probability density function (pdf) of $(\underline{x}'\underline{x})^{\frac{1}{2}}$.

Kelker (1970) gives an extensive discussion of spherically symmetric distributions with pdf's of the form $f(\underline{x}) = \psi(\underline{x}'\underline{x})$. He also extends his discussion to include elliptically symmetric distributions with pdf's of the type $f(\underline{x}) = |\underline{\Sigma}|^{-\frac{1}{2}} \psi[(\underline{x}-\underline{\theta})'\underline{\Sigma}^{-1}(\underline{x}-\underline{\theta})]$. He gives a condition for the existence of a probability density function, derives marginal distributions, shows that the conditional means have the same linear structure for all elliptically symmetric distributions, and gives several characterizations of the Gaussian distribution. He also proves that, if $\underline{\theta} = 0$, $\underline{\Sigma} = \underline{I}$ and the vector $\underline{x}$ is partitioned into $[\underline{x}_1'(1\times n_1), \underline{x}_2'(1\times n_2)]$, then the distribution of $(\underline{x}_1'\underline{x}_1/\underline{x}_2'\underline{x}_2)(n_2/n_1)$ has the usual F distribution derived under Gaussian assumptions. This fact, that certain statistics have unchanged null distributions under a wider class of distributions, is also noted earlier by Bennett (1961) and Efron (1969).

Bennett (1961) examines the following distribution as an alternative to the Gaussian distribution in tests of significance in multivariate analysis. Let $\underline{x}(p\times 1)$ have density $f(\underline{x}, \underline{\theta}, \underline{\Lambda}) = K[1 + \frac{1}{n} Q(\underline{x}, \underline{\theta}, \underline{\Lambda})]^{-\frac{1}{2}(n+p)}$ with $Q(\underline{x}, \underline{\theta}, \underline{\Lambda}) = (\underline{x}-\underline{\theta})'\underline{\Lambda}(\underline{x}-\underline{\theta})$. This is a multivariate generalization of Student's t distribution because, if $p = 1$, we get the univariate Student's t distribution with n degrees of freedom. When $\underline{\theta} = \underline{0}$ and $\underline{\Lambda} = \underline{I}$, Bennett notes that $\underline{x}'\underline{x}/(1 + \underline{x}'\underline{x})$ has a Beta distribution with parameters $p/2$ and $n/2$.

Efron (1969) uses Fisher's (1925) essentially geometric derivation of Student's t distribution to show that this distribution remains unchanged if we assume an underlying spherically symmetric distribution rather than a Gaussian distribution.

This uniqueness of certain distributions under the weaker assumption of spherical symmetry is also noted by Ghosh and Pollack (1975) and Thomas (1970).

Ghosh and Pollack (1975) prove that if $x_1, \ldots, x_n$ are random variables with unbounded support and with joint density spherically symmetric, then the joint distribution of $y_i = x_i/x_n$, $1 \leq i \leq n - 1$, is a multivariate Cauchy distribution.

Thomas (1970) looks at the univariate general linear model $\underline{y} = X\underline{\beta} + \underline{\varepsilon}$ when $\underline{\varepsilon}$ has a spherically symmetric distribution. He shows that the usual t and F test statistics used for tests on $\underline{\beta}$ have unchanged null distributions for this wider class of spherically symmetric laws. He also gives expressions for the non-null distributions for these two cases and for the statistic $\underline{x}'\underline{x}$. These distributions are not independent of the underlying distribution. Laurent (1974), under less stringent conditions than Thomas, gives several other expressions for the non-central distribution of $\underline{x}'\underline{x}$.

Thomas' (1970) unpublished work apparently is overlooked within the statistical community. A case in point is Zellner's (1976) consideration of the linear model $\underline{y} = \underline{X\beta} + \underline{\varepsilon}$ for $\underline{\varepsilon}$ distributed as a multivariate t. This is a specific example of a spherically symmetric distribution. Zellner shows that $\hat{\underline{\beta}} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{y}$ is the maximum likelihood estimate for $\underline{\beta}$ and, furthermore, that $\hat{\underline{\beta}}$ is a maximum likelihood estimate for $\underline{\beta}$ for all likelihood functions which are monotonically decreasing functions of $(\underline{y}-\underline{X\beta})'(\underline{y}-\underline{X\beta})$. He further adds that if second moments exist then $\hat{\underline{\beta}}$ is a minimum variance unbiased estimator. All this is noted by Thomas in his thesis. Zellner does not duplicate Thomas entirely, as he also considers problems of inference within a Bayesian framework.

Thomas, however, is not the first to use a spherically symmetric distribution instead of a Gaussian distribution in the analysis of linear models. His work was anticipated to some extent by Box (1952, 1953).

Box (1952) notes that the usual F statistics have the same null distribution for all $f(x) = \psi(\underline{x}'\underline{x})$. Box (1953) makes the statement, in an abstract, that those tests which are uniformly most powerful (UMP) under Gaussian assumptions are also UMP for any spherical distribution in which $\psi$ is a decreasing function. However, there is no published verification of these assertions, the validity of which accordingly remains an open question. Kariya and Eaton (1977) and Kariya (1977) also investigate power properties for several testing situations.

Kariya and Eaton (1977) assume that $f(\underline{x}) = \psi[(\underline{x}-\mu\underline{1})'(\underline{x}-\mu\underline{1})]$ for $\underline{1}' = (1,1,\ldots,1)$ and that $\psi$ is a non-increasing function. The test for H: $\mu = 0$ versus $K_1$: $\mu > 0$ or $K_2$: $\mu < 0$ is UMP and the null distribution is the ordinary Student's t distribution. They also show that the test for H: $\mu = 0$ versus K: $\mu \neq 0$ is uniformly most powerful unbiased (UMPU) when $\psi$ is also convex.

Kariya and Eaton (1977) and Kariya (1977) consider tests for serial correlation. Let the $(n\times 1)$ vector $\underline{x}$ have an elliptically symmetric distribution with $f(\underline{x}) = \psi[\underline{x}'[\gamma\underline{\Sigma}(\lambda)]^{-1}\underline{x}]$, $\gamma > 0$ and $\underline{\Sigma}(\lambda)^{-1} = \underline{I}_n + \lambda\underline{A}$ for $\underline{A}$ a known $(n\times n)$ matrix. They test H: $\lambda = 0$ versus $K_1$: $\lambda > 0$ and $K_2$: $\lambda \neq 0$. For these two tests the null distributions are identical to those derived under a Gaussian assumption. The one-sided test is UMP if we assume that $\psi$ is non-increasing. The two-sided test is UMPU if we assume also that $\psi$ is convex.

So far we have cited research pertaining only to vector distributions. Very little has been done for matrices $\underline{X}(n\times m)$ with spherically symmetric distributions.

Hsu (1940) finds the pdf of $\underline{S} = \underline{X}'\underline{X}$ when $\underline{X}(n\times m)$ has a pdf of the form $f(\underline{X}) = \psi(\underline{X}'\underline{X})$. Ahmad (1972) investigates the multivariate general linear model $\underline{Y} = \underline{X}\,\underline{\Theta} + \underline{E}$ with the assumption that $\underline{E}(n\times m)$ contains n independent and identically distributed spherically symmetric m vectors. He then finds the distribution of $\underline{E}'\underline{E}$. Unfortunately he seems to be unaware of Hsu's work. Furthermore, in his proof for the distribution of $\underline{E}'\underline{E}$ he assumes that $f(\underline{E}) = \psi(\underline{E}'\underline{E})$. Therefore, his proof reduces to a proof for the Wishart distribution because independence together with spherical symmetry imply that $f(\underline{E})$ is a Gaussian distribution.

We have discussed the role of spherically symmetric distributions in distribution theory and statistical inference. These distributions, however, are employed in several other areas such as minimax estimation, stochastic processes, fiducial inference, and probability inequalities. We discuss each of these areas in turn and give a brief summary of results.

Minimax Estimation

Strawderman (1974) finds families of minimax estimators for the location parameters of a p-variate ($p \geq 3$) distribution of the form

$$\int_0^\infty \frac{1}{(2\pi\sigma^2)^{p/2}} \exp[-\frac{1}{2\sigma^2}(\underline{x}-\underline{\theta})'(\underline{x}-\underline{\theta})]dG(\sigma)$$

where $G(\cdot)$ is a cumulative distribution function on $(0, \infty)$ and the loss is sum of squared errors.

Berger (1975) extends this by considering a density of the form

$$f(\underline{x}) = \int_0^\infty \frac{|\Sigma|^{-1/2}}{(2\pi)^{p/2}\,\sigma^p} \exp[-\frac{1}{2\sigma^2}(\underline{x}-\underline{\theta})'\underline{\Sigma}^{-1}(\underline{x}-\underline{\theta})]dG(\sigma)$$

and develops families of minimax estimators for $\underline{\theta}$ under a known quadratic loss.

Brandwein and Strawderman (1978) find families of minimax estimators for the location parameters of a p-variate ($p \geq 3$) spherically symmetric unimodal distribution with respect to general quadratic loss. They use the following definition of multivariate unimodality. A ($p\times 1$) random vector $\underline{x}$ is said to have a spherically symmetric unimodal distribution about $\underline{\theta}$ if the pdf of $\underline{x}$ with respect to Lebesgue measure is a non-increasing function of $(\underline{x}-\underline{\theta})'(\underline{x}-\underline{\theta})$.

## Stochastic Processes

Vershik (1964) defines a spherically invariant random process and shows that certain well-known results for the Gaussian process extend to this larger class, for example, closure under linear operators and the property of linear conditional expectations. McGraw and Wagner (1968) investigate bivariate elliptically symmetric distributions. They use these distributions to describe the second order moments of the transformation of a random signal by an instantaneous non-linear device.

Blake and Thomas (1968) derive characteristic functions and discuss the relationship of spherically invariant processes to linear estimation theory and Gaussian processes. Picinbono (1970) compares the class of spherically invariant processes with a particular class of Gaussian compound processes.

Yao (1973) considers various basic statistical properties of multivariate characteristic functions and probability density functions associated with spherically invariant random processes. He then applies this to estimation and detection problems in communication theory.

Chu (1973) shows that elliptically symmetric distributions sometimes can be expressed as integrals of a set of Gaussian densities, that is,

$$p(\underline{x}) = \int_0^\infty w(t)\, N_{\underline{x}}\, (t^{-1}\underline{C})dt$$

where $N_{\underline{x}}(t^{-1}\underline{C}) = (2\pi)^{-n/2} \exp[-\frac{1}{2}\,\underline{x}'(t^{-1}\underline{C})\underline{x}]$. The function $w(t)$ is defined to be $(2\pi)^{n/2}|\underline{C}|^{1/2}\, t^{-n/2}\, L^{-1}[f(s)]$ where $L(\cdot)$ is the Laplace transform operator, $L^{-1}(\cdot)$ is the inverse operator, and $f(s) = f(\frac{1}{2}\,\underline{x}'\underline{C}^{-1}\underline{x})$. Using this representation he derives marginal and conditional distributions. He then considers problems of computing optimal estimation, filtering, stochastic control, and team decisions in various linear systems.

Gualtierotti (1974) studies the equivalence properties of spherically invariant measures on a real and separable Hilbert space. He then uses this to give a likelihood ratio formula for equivalent spherically invariant measures [cf. Gualtierotti (1976)].

Goldman (1976) finds the pdf of $\underline{x} + \underline{n}$ when the random vector $\underline{x}$ has an elliptically symmetric distribution and $\underline{n}$ has a Gaussian distribution with $\underline{x}$ and $\underline{n}$ independent. He uses this to examine the problems of detecting a known signal vector in the presence of $\underline{x} + \underline{n}$, for $\underline{x}$ spherically symmetric.

## Fiducial Inference

All references cited here are for the multivariate structural model $\underline{Y}(n\times m) = X\underline{\beta} + \underline{E}\ \underline{\Gamma}$. Fraser and Haq (1969) derive error and structural distributions for this model. As a special case, they assume $\underline{E}$ has a spherically symmetric distribution with $f(\underline{E}) = \psi(\underline{E}'\underline{E})$.

Fraser and Haq (1970) derive the conditional distribution for the error given the response observations. Again, a special case considered is for $\underline{E}$ spherically symmetric.

Dawid (1977) considers inferences about $\underline{\beta}$, hypothesis testing and confidence interval estimation, under the assumption that $\underline{E}$ has a spherically symmetric distribution. Using studentization he gets a test statistic which has a matric t distribution. The null distribution of this test statistic is identical to the one derived under a Gaussian assumption.

## Probability Inequalities

Das Gupta et al (1971) discuss inequalities for probabilities of the type

$$\int_{-\underline{h}}^{\underline{h}} |\underline{\Sigma}|^{-1/2}\ \psi(\underline{x}'\underline{\Sigma}^{-1}\underline{x})\,d\underline{x}$$

and

$$\int_{-\infty}^{\underline{\ell}} |\Sigma|^{-1/2}\ \ \psi(\underline{x}'\underline{\Sigma}^{-1}\underline{x})\,d\underline{x}$$

where $\underline{h}' = (h_1, \ldots, h_n)$, $\underline{\ell}' = (\ell_1, \ldots, \ell_n)$ are constant vectors,

$h_i \geq 0$, $i=1,\ldots,n$, and $\underline{\Sigma}$ is positive definite. Their main result is the following. Let

$$\underline{\Sigma} = \begin{bmatrix} \underline{\Sigma}_{11} & \underline{\Sigma}_{12} \\ \underline{\Sigma}_{21} & \sigma_{pp} \end{bmatrix}$$

be a $(p\times p)$ positive definite matrix with $\underline{\Sigma}_{11}$: $(p-1)\times(p-1)$ and let $\underline{x}' = (x_1, \ldots, x_p)$ be a random vector with density function $|\underline{\Sigma}_\lambda|^{-1/2} \times \psi(\underline{x}'\underline{\Sigma}_\lambda^{-1}\underline{x})$, where

$$\Sigma_\lambda = \begin{bmatrix} \underline{\Sigma}_{11} & \lambda\underline{\Sigma}_{12} \\ \lambda\underline{\Sigma}_{21} & \sigma_{pp} \end{bmatrix}, \quad 0 \leq \lambda \leq 1 \ .$$

If C is a convex symmetric set in the Euclidean space $R^{p-1}$, then $P_\lambda\{(x_1,\ldots,x_{p-1}) \in C, |x_p| \leq h\}$ is non-decreasing in $\lambda$. Using this result they show that certain normal theory simultaneous confidence bounds extend to the class of elliptically symmetric distributions.

Wynn (1977) considers a spherically symmetric bivariate probability distribution centered at zero. He gives an inequality comparing the probability content of two different polygons circumscribing the same fixed circle with center at zero. These results can be used to compare the widths of certain simultaneous confidence intervals.

We now summarize a few papers on spherically symmetric distributions which do not fit into any designated category. These citings complete the review.

Kingman (1963) considers the problem of random walks with spherical symmetry. Higgins (1975) shows that $f(\underline{x}) = \psi(\underline{x}'\underline{x})$ has a functional form which is uniquely determined by the distribution of $\underline{x}'\underline{x}$. Using this characterization he gives a geometrical method for constructing multivariate distributions which is based upon surface integral techniques.

Devlin, Gnanadesikan, and Kettenring (1976) give a summary of distributional results associated with elliptically symmetric distributions. As the correlation matrices for all elliptically symmetric distributions having a given matrix $\Sigma$ are identical, they discuss using a subclass of these distributions, the variance mixtures of Gaussian distributions, for simulating non-Gaussian alternatives in studies of measures of correlation.

In contrast, Johnson and Ramberg (1977) actually derive computer algorithms for generating elliptically symmetric distributions. Convenient algorithms are given for generating bivariate elliptically symmetric variates.

Smith (1977) gives a non-parametric test for bivariate spherical symmetry using the empirical cumulative distribution function. He finds an asymptotic expansion for the null and non-null distributions under simple alternatives.

Box and Hunter (1957) derive moments of a rotatable design of order d. These are also the moments up to order 2d of a spherical distribution.

This concludes the literature review. To summarize we note that elliptically symmetric distributions have been employed in a wide range of settings for a wide range of purposes. However, very little research pertaining to the scale parameters of such distributions has been done.

# III. SPHERICALLY SYMMETRIC DISTRIBUTIONS

## Section 3.1 Introduction

We introduce spherically symmetric distributions for a $(n\times 1)$ random vector $\underline{x}$ and for a random matrix $\underline{X}$ of order $(n\times m)$. Some properties of these distributions are cited as needed for a discussion of hypothesis tests.

## Section 3.2 Vector Distributions

We first define a spherically symmetric distribution for the $(n\times 1)$ random vector $\underline{x}$.

Definition 3.2.1: Let $\underline{x}(n\times 1)$ be a random vector with distribution $L(\underline{x})$. Then $\underline{x}$ is said to have a spherically symmetric distribution if, for each $\underline{P}$ belonging to the orthogonal group $\mathcal{O}(n)$, $L(\underline{x}) = L(\underline{Px})$. We write

$$L(\underline{x}) \in S_n[\underline{0},\underline{I}_n]$$

with $\underline{0}$ the vector of location parameters and $\underline{I}_n$ the matrix of scale parameters.

If $L(\underline{x})$ does not have an atom of weight at the origin, then $\underline{x}$ has a pdf with respect to Lebesgue measure [cf. Kelker (1970)]. The pdf takes the form $f(\underline{x}) = \psi(\underline{x}'\underline{x})$ for some function $\psi$. For all hypothesis testing problems considered in later chapters, we assume that a pdf exists.

Lord (1954) shows that, as the pdf is a function of $\underline{x}'\underline{x}$, the characteristic function $\phi_{\underline{x}}(\underline{t})$ is a function of $\underline{t}'\underline{t}$, i.e.

$$\phi_{\underline{x}}(\underline{t}) = \delta(\underline{t}'\underline{t}) \tag{3.2.1}$$

for some function $\delta$.

Kelker (1970) extends the definition of a spherically symmetric distribution by including a location vector $\underline{\mu}$ and a matrix $\underline{\Sigma}$ of scale parameters. This extension follows from the transformation $\underline{x} \to \underline{Ax} + \underline{\mu}$ where $\underline{\Sigma} = \underline{AA}'$ and the rank of $\underline{\Sigma}$ is n. We now write $L(\underline{x}) \in S_n[\underline{\mu},\underline{\Sigma}]$ with pdf

$$f(\underline{x}) = |\underline{\Sigma}|^{-1/2}\ \psi[(\underline{x}-\underline{\mu})'\underline{\Sigma}^{-1}(\underline{x}-\underline{\mu})] \tag{3.2.2}$$

and characteristic function

$$\phi_{\underline{x}}(\underline{t}) = \exp(i\underline{t}'\underline{\mu})\ \delta(\underline{t}'\underline{\Sigma t}) . \tag{3.2.3}$$

Suppose $L(\underline{x})$ has k moments. If $k = 1$, the mean is $\underline{\mu}$ and if $k = 2$ the covariance matrix is $\alpha\underline{\Sigma}$ for $\alpha \in (0,\infty)$. If $\underline{\Sigma} = \underline{I}_n$ then $\alpha$ is the variance of any univariate marginal distribution [cf. Kelker (1970)]. If $\underline{x}$ has a Gaussian distribution then $\alpha = 1$ and the covariance matrix equals the matrix $\underline{\Sigma}$ of scale parameters.

We discuss tests for various hypotheses concerning $\underline{\Sigma}$ in later chapters when the underlying distribution of $\underline{X}(n\times m)$ is elliptically symmetric as defined in Section 3.5. The moment results imply that these are tests for scale and not necessarily for variances and/or covariances. Therefore, the underlying distribution can be one for which second moments do not exist, e.g. the multivariate Cauchy distribution. In such situations we might still be interested in testing hypotheses regarding the scale parameters.

## Section 3.3 Representation in Polar Coordinates

We can represent $f(\underline{x}) = \psi(\underline{x}'\underline{x})$ in a system $(r,\underline{\theta})$ of polar coordinates because $f(\underline{x})$ is a function of $\underline{x}'\underline{x}$. There is no unique transformation from $\underline{x}$ to $(r,\underline{\theta})$ and we give one commonly used in the literature. Define

$$
\begin{aligned}
x_1 &= r\cos\theta_1 \\
x_2 &= r\sin\theta_1\cos\theta_2 \\
&\vdots \\
x_{n-1} &= r\sin\theta_1\sin\theta_2\ldots\cos\theta_{n-1} \\
x_n &= r\sin\theta_1\sin\theta_2\ldots\sin\theta_{n-1}
\end{aligned}
\tag{3.3.1}
$$

for $0 \le \theta_i \le \pi$, $i=1,\ldots,n-2$ and $0 \le \theta_{n-1} \le 2\pi$. Then $f(\underline{x})$ can be expressed as the product of two pdf's [cf. Thomas (1970)], i.e.

$$f(\underline{x}) = g(r)h(\underline{\theta}) \tag{3.3.2}$$

where

$$g(r) = \frac{2\pi^{n/2}}{\Gamma(n/2)}\, r^{n-1}\ \psi(r^2) \tag{3.3.3}$$

is the pdf of the radius and where

$$h(\underline{\theta}) = \frac{\Gamma(\frac{n}{2})}{2\pi^{n/2}} \sin^{n-2}\theta_1\ \sin^{n-3}\theta_2\ \ldots\ \sin\theta_{n-2}$$

is the pdf of the angles.

## Section 3.4 Some Examples of Spherical Vector Distributions

This section gives examples of distributions which are members of $S_n[\underline{0},\sigma^2\underline{I}_n]$. Examples (i)-(v) are discussed in Johnson and Ramberg (1977) while example (vi) is discussed by Kelker (1970) and Chu (1973).

(i) Gaussian pdf

$$(3.4.1) \qquad f(\underline{x}) = \frac{\sigma^{-n}}{(2\pi)^{n/2}} \exp[-(\underline{x}-\underline{\mu})'(\underline{x}-\underline{\mu})/2\sigma^2]$$

(ii) Cauchy pdf

$$(3.4.2) \qquad f(\underline{x}) = \frac{\sigma^{-n}\Gamma[\frac{1}{2}(n+1)]}{\pi^{\frac{1}{2}(n+1)}} \quad [1+(\underline{x}-\underline{\mu})'(\underline{x}-\underline{\mu})/\sigma^2]^{-\frac{1}{2}(n+1)}$$

(iii) Student's t pdf

$$(3.4.3) \qquad f(\underline{x}) = \frac{\sigma^{-n}\Gamma[\frac{1}{2}(n+\gamma)]}{(\pi\gamma)^{\frac{1}{2}n}\ \Gamma(\frac{1}{2}\gamma)} \quad [1+(\underline{x}-\underline{\mu})'(\underline{x}-\underline{\mu})/\sigma^2]^{-\frac{1}{2}(\gamma+n)}$$

For $\gamma = 1$, (3.4.3) becomes (3.4.2).

(iv) Pearson Type II pdf

$$(3.4.4) \qquad f(\underline{x}) = \frac{\sigma^{-n}\Gamma(\gamma+\frac{n}{2})}{\Gamma(\gamma)\pi^{n/2}} \quad [1-(\underline{x}-\underline{\mu})'(\underline{x}-\underline{\mu})/\sigma^2]^{\gamma-1},\ \gamma > 1$$

(v) Pearson Type VII pdf

$$(3.4.5) \qquad f(\underline{x}) = \frac{\sigma^{-n}\Gamma(\gamma)}{\pi^{n/2}\ \Gamma(\gamma-n/2)} \quad [1+(\underline{x}-\underline{\mu})'(\underline{x}-\underline{\mu})/\sigma^2]^{-\gamma},\ \gamma > n/2$$

For $\gamma = \frac{n+1}{2}$, (3.4.5) becomes (3.4.2).

(vi) Variance mixtures of Gaussian distributions

$$(3.4.6) \quad f(\underline{x}) = \int (2\pi)^{-n/2} (t\sigma^2)^{-n/2} \exp[-(\underline{x}-\underline{\mu})'(\underline{x}-\underline{\mu})/2t\sigma^2] \cdot dW(t)$$

with $W(t)$ a distribution function on $(0,\infty)$. This subclass contains the multivariate Gaussian law, the contaminated Gaussian law, the multivariate Student's t distributions, the multivariate Cauchy law, and the symmetric stable laws [cf. Kelker (1970) and Chu (1973)].

## Section 3.5 Matrix Distributions

We now define a spherically symmetric distribution for the $(n\times m)$ random matrix $\underline{X}$.

Definition 3.5.1: Let $\underline{X}(n\times m)$ be a random matrix with distribution $L(\underline{X})$. Then $L(\underline{X})$ is said to be left spherically invariant if, for each $\underline{P} \in \mathcal{O}(n)$, $L(\underline{X}) = L(\underline{PX})$.

If $\underline{X}$ has a pdf with respect to Lebesgue measure then $f(\underline{X}) = \psi(\underline{X}'\underline{X})$. In this thesis we further restrict our class of densities by assuming that $L(\underline{X})$ is both left and right invariant. If a pdf exists we assume it takes the form

$$(3.5.1) \qquad f(\underline{X}) = \psi(\mathrm{tr}\ \underline{X}'\underline{X})$$

and we write $L(\underline{X}) \in S_{nm}[\underline{\phi}, \underline{I}_n \otimes \underline{I}_m]$ where $\underline{\phi}$ is a zero matrix and $\underline{I}_n \otimes \underline{I}_m$ is the Kronecker product of two identity matrices. This dependence on the trace facilitates the study of distributions of test statistics in later chapters.

One reason for assuming a Kronecker product structure for the scale matrix is to achieve an extension of the multivariate Gaussian distribution. However, there is a more important reason. Let $\underline{X}(n\times m)$ be imbedded in $R^{nm}$, that is write $\underline{X}(n\times m)$ as

$$\text{(3.5.2)} \qquad \underline{X} = \begin{bmatrix} \underline{x}_1' \\ \vdots \\ \underline{x}_n' \end{bmatrix} \rightarrow \begin{bmatrix} \underline{x}_1 \\ \vdots \\ \underline{x}_n \end{bmatrix} = \underline{x}(nm\times 1) .$$

Thomas (1970) shows that, under the assumption of left invariance, the $(nm\times nm)$ scale matrix for $\underline{X}$ must take the form $\underline{I}_n \otimes \underline{\Sigma}$. This structure assumes a central role in later chapters.

Since the class of pdf's given in (3.5.1) is a subclass of the pdf's of the form $f(\underline{X}) = \psi(\underline{X}'\underline{X})$, we assume the Kronecker product structure for this subclass.

As in the vector case we can let $\underline{X} \rightarrow \underline{XA} + \underline{M}$ for $\underline{A}(m\times m)$ of full rank and extend the class of spherically symmetric distributions to the class of elliptically symmetric distributions $S_{nm}[\underline{M},\underline{I}_n \otimes \underline{\Sigma}]$ with characteristic function

$$\text{(3.5.3)} \qquad \phi_{\underline{X}}(\underline{T}) = \exp[i \text{ tr } \underline{T}'\underline{M}) \; \delta(\text{tr } \underline{T}'\underline{T\Sigma}) \; .$$

Next, we express our matrix distribution by imbedding $\underline{X}$ in $R^{mn}$. Let

$$f(\underline{X}) = |\underline{\Sigma}|^{-n/2} \psi(\text{tr } \underline{X}'\underline{X\Sigma}^{-1})$$

with

$$\underline{X}(n\times m) = \begin{bmatrix} \underline{x}_1' \\ \vdots \\ \underline{x}_n' \end{bmatrix} .$$

Then

$$\begin{aligned} \text{tr } \underline{X}'\underline{X}\underline{\Sigma}^{-1} &= \text{tr } \underline{\Sigma}^{-1} \sum_{j=1}^{n} \underline{x}_j \underline{x}_j' \\ &= \sum_{j=1}^{n} \underline{x}_j' \, \underline{\Sigma}^{-1} \underline{x}_j \\ &= \underline{x}'(\underline{I}_n \otimes \underline{\Sigma})^{-1} \underline{x} \end{aligned}$$

with $\underline{x}(nm\times 1)$ a random vector and

$$(3.5.4) \qquad f(\underline{x}) = |\underline{I}_n \otimes \underline{\Sigma}|^{-1/2} \, \psi[\underline{x}'(\underline{I}_n \otimes \underline{\Sigma})^{-1}\underline{x}] \quad .$$

This representation of $\underline{X}$ in $R^{nm}$ is useful in later chapters where we consider the distributional properties of various test statistics.

## Section 3.6 Singular Distributions

In later chapters we deal with spherically symmetric distributions which are singular. In anticipation of this we discuss singular distributions within the framework of elliptically symmetric distributions and prove a lemma regarding the characteristic functions of such distributions.

A singular distribution in Euclidean p space is a distribution which is concentrated on a set of lower dimension. That is, with unit probability the random element is concentrated on a r-dimensional subspace for $r < p$.

We now identify the subspace on which the singular measure is concentrated.

Lemma 3.6.1: Let $\phi_{\underline{x}}(\underline{t}) = \exp(i\ \underline{t}'\underline{\mu})\cdot\delta(\underline{t}'\underline{\Sigma t})$ with $\underline{\Sigma} = \sum_{j=1}^{r} \lambda_j\ \underline{q}_j\underline{q}_j'$. Then $\{\underline{q}_1,\ldots,\underline{q}_r\}$ forms a basis for the r-dimensional subspace on which the measure is concentrated and $\{\underline{q}_{r+1},\ldots,\underline{q}_p\}$ spans the space, complementary to the above r-dimensional subspace, with all mass concentrated at a point.

Proof: Since the rank of $\underline{\Sigma}$ is r there exists a $\underline{Q} \in \mathcal{O}(p)$ such that $\underline{\Gamma} = \underline{Q\Sigma Q}' = \mathrm{Diag}(\gamma_1,\ldots,\gamma_r,\ 0,\ldots,0)$. Let

$$\underline{y} = \underline{Qx}$$

and

$$\underline{\theta} = \underline{Q\mu}\ .$$

Then

$$\phi_{\underline{y}}(\underline{t}) = \exp(i\ \underline{t}'\underline{\theta})\ \delta(\underline{t}'\underline{\Gamma t}).$$

Partition $\underline{t}$, $\underline{\theta}$, and $\underline{\Gamma}$ as follows:

$$\underline{t}' = [\underline{t}_1'(1\times r),\ \underline{t}_2'(1\times(p-r))]$$

$$\underline{\theta}' = [\underline{\theta}_1',\ \underline{\theta}_2']$$

$$\underline{\Gamma} = \begin{bmatrix} \underline{\Gamma}_1 & \underline{\phi} \\ \underline{\phi} & \underline{\phi} \end{bmatrix}$$

with

$$\underline{\Gamma}_1 = \mathrm{Diag}(\gamma_1, \ldots, \gamma_r) \ .$$

Then (3.6.1) becomes

(3.6.2) $$\phi_{\underline{y}}(\underline{t}) = \exp(i\ \underline{t}_1'\underline{\theta}_1)\ \delta(\underline{t}_1'\underline{\Gamma}_1\underline{t}_1) \cdot \exp(i\ \underline{t}_2'\underline{\theta}_2) \ .$$

This is the product of a characteristic function for a nondegenerate elliptically symmetric distribution on $R^r$ and the characteristic function of a distribution on $R^{p-r}$ degenerate at $\underline{\theta}_2$. This completes the proof.

In later chapters we prove that certain distributions are scale invariant. These distributions will be singular with $\underline{\mu} = \underline{0}$ and hence $\underline{\theta}_2 = \underline{0}$. The lemma assures that we need prove only the scale invariance for the nondegenerate distribution defined on $R^r$.

# IV. INVARIANCE AND HYPOTHESIS TESTING

## Section 4.1 Introduction

The principle of invariance states that if a problem is invariant with respect to a group of transformations then the solution should be invariant with respect to the same group of transformations. In statistical inference the use of this principle is motivated by the fact that certain testing problems exhibit symmetries which place natural restrictions on these problems. In the present study there are three concomitant reasons for performing an invariant reduction for the hypothesis tests considered.

First, as the entire class possesses the required symmetries, we do not need to specify an underlying distribution in order to do an invariant reduction. Second, each invariant reduction produces a class of test statistics all of which are functions of the maximal invariants. We develop properties of the joint distributions of these maximal invariants which carry over to the distributions of certain functions of the maximal invariants. These functions serve as test statistics. Third, and most important, using the invariance structure of the underlying class of elliptically symmetric distributions, we prove distributional properties for a class of test statistics which apply to all underlying elliptically symmetric laws.

This is in contrast to other principles of data reduction. Likelihood ratio test procedures often require that we specify the underlying distribution. For a further study of individual cases, this method may be preferable. However, in this thesis we invoke the principle of invariance to characterize all invariant tests for all members of an underlying class of distributions.

## Section 4.2 Invariant Procedures

We review invariant procedures as they apply to hypothesis tests. All the material in this section is from Lehmann (1959).

Let $(\mathcal{X}, \mathcal{B})$ be a measurable space and let the random element X, with values in $\mathcal{X}$, have the probability distribution $P_\theta$, for $\theta$ an element of a parameter space $\Omega$. Then $(\mathcal{X}, \mathcal{B}, P_\theta)$ is a probability space.

Let G be a collection of one-to-one transformations of the sample space $\mathcal{X}$ onto itself such that gX denotes the random element which takes on the value gx for X = x and $g \varepsilon G$. Suppose that when $L(X) = P_\theta$, $\theta \varepsilon \Omega$, then $L(gX) = P_{\theta'}$, $\theta' \varepsilon \Omega$. We denote $\theta'$ by $\bar{g}\theta$ with

$$P_\theta\{gX \varepsilon B\} = P_{\bar{g}\theta}\{X \varepsilon B\} \tag{4.2.1}$$

and say that the transformation g on $\mathcal{X}$ induces a transformation $\bar{g}$ on $\Omega$.

The parameter space $\Omega$ is said to be preserved by G if

$$\bar{g}\ \Omega = \Omega \tag{4.2.2}$$

for each $g \varepsilon G$.

<u>Theorem 4.2.1</u>: Let g, g' be two transformations preserving $\Omega$. Then the transformations g'g and $g^{-1}$ defined by

$$(g'g)x = g'(gx)$$

and

$$g(g^{-1}x) = x$$

for all $x \varepsilon \mathcal{X}$ also preserve $\Omega$ and satisfy

(4.2.3) $$\overline{g'g} = \overline{g'}\cdot\overline{g} \quad \text{and} \quad (\overline{g^{-1}}) = (\overline{g})^{-1} \ .$$

Consider now the problem of testing H: $\theta \varepsilon \Omega_H$ versus K: $\theta \varepsilon \Omega_K$, where $\Omega_H$ and $\Omega_K$ are disjoint subsets of $\Omega$. We say that the testing problem remains invariant under $g \varepsilon G$ if in addition to (4.2.2)

(4.2.4) $$\overline{g}\ \Omega_H = \Omega_H$$

also holds.

Let $\mathcal{G}$ be a class of transformations satisfying (4.2.2) and (4.2.4). Also, let G be the smallest class of transformations containing $\mathcal{G}$ such that g, g' $\varepsilon$ G implies that $g'g \ \varepsilon$ G and $g^{-1} \ \varepsilon$ G. Then G is a group of transformations. By Theorem 4.2.1 G preserves both $\Omega$ and $\Omega_H$. From (4.2.3) we also have that the class of induced transformations $\overline{g}$ forms a group $\overline{G}$. Because any class $\mathcal{G}$ of transformations leaving the problem invariant can be extended to a group G, in all applications we consider G to be a group of transformations.

The principle of invariance restricts our attention to tests, $\phi(x)$, which are also invariant, that is

$$\phi(x) = \phi(gx), \quad x\varepsilon\mathcal{X},\ g\varepsilon G \ .$$

Definition 4.2.1: A function $T(x)$ defined on $X$ is maximal invariant under G if (i) $T(x) = T(gx)$, for each $x \varepsilon X$ and $g \varepsilon G$, and (ii) $T(x_1) = T(x_2)$, $x_1, x_2 \varepsilon X$ implies that there is some $g \varepsilon G$ such that $x_2 = gx_1$.

Theorem 4.2.2: A test $\phi(x)$ is invariant under G if and only if there exists a function h such that $\phi(x) = h(T(x))$.

The theorem states that an invariant test statistic will be a function of the sample maximal invariants produced by the group G. The induced group $\bar{G}$ produces corresponding maximal invariants on the parameter set $\Omega$.

Theorem 4.2.3: The distribution of $T(x)$ depends on $\Omega$ only through a maximal invariant function of the parameters.

It follows that an invariant reduction reduces both the sample space and the parameter space.

## Section 4.3 The Polar Representation of a Test Statistic

Let $L(\underline{x}) \varepsilon S_n[\underline{0}, \underline{I}_n]$. We discuss a method useful for showing that certain test statistics $T(\underline{x})$ have unique distributions. This method is extended to the case where the matrix $\underline{X}$ has a distribution belonging to $S_{nm}[\underline{\phi}, \underline{I}_n \otimes \underline{I}_m]$.

Define $B_p = \{\underline{x} \varepsilon R^p | (\underline{x}'\underline{x})^{1/2}=1\}$ and let U have a uniform distribution on $B_p$. Then $L(U)$ is the unique probability distribution on $B_p$ which is invariant under $O(p)$ [cf. Kariya and Eaton (1977)]. We now have the following theorem from Kariya and Eaton (1977).

Theorem 4.3.1: If $L(\underline{y}) = N_p[\underline{0},\underline{I}_p]$ and $L(\underline{x}) \varepsilon S_p[\underline{0},\underline{I}_p]$ then $L(\underline{x}/(\underline{x}'\underline{x})^{1/2}) = L(\underline{y}/(\underline{y}'\underline{y})^{1/2}) = L(U)$.

Upon taking $f(\underline{x}) = \psi(\underline{x}'\underline{x})$, we have the following corollary.

Corollary 4.3.1: Let a test statistic have the form $T(\underline{x}) = h(\underline{x}/(\underline{x}'\underline{x})^{1/2})$ for some function h. Then $T(\underline{x})$ has a unique distribution for all $L(\underline{x}) \varepsilon S_p[\underline{0},\underline{I}_p]$.

The corollary states that if $T(\underline{x})$ depends only on the angles in the polar representation given by (3.3.1), then $T(\underline{x})$ has a unique distribution. Therefore, to check whether a test statistic $T(\underline{x})$ has a unique distribution for every $L(\underline{x}) \varepsilon S_p[\underline{0},\underline{I}_p]$, we take $\underline{x} \to (r,\underline{\theta})$; we then transform $T(\underline{x})$ into $T(r,\underline{\theta})$; and finally we examine whether $T(r,\underline{\theta}) = T(\underline{\theta})$.

This discussion treats only vector distributions thus far, but the method works equally well for $L(\underline{X}) \varepsilon S_{nm}[\underline{\phi},\underline{I}_n \otimes \underline{I}_m]$.

Let $\underline{X}(n\times m)$ be represented in $\mathcal{R}^{nm}$ as in (3.5.2). Transform to $(r,\underline{\theta})$, then re-express the transformed variables $(r,\underline{\theta})$ in matrix form, i.e.

$$\underline{X} \to \underline{x} \to (r,\underline{\theta}) \to (r,\underline{\Theta}) \quad . \tag{4.3.1}$$

The $(n\times m)$ matrix $\underline{\Theta}$ contains the angular variables. Since $\underline{\theta}'\underline{\theta} = 1$ if and only if $\operatorname{tr} \underline{\Theta}'\underline{\Theta} = 1$, $(r,\underline{\Theta})$ is a polar transformation of $\underline{X}$ with $r^2 = \operatorname{tr} \underline{X}'\underline{X}$.

To check whether a test statistic $T(\underline{X})$ is a function of angular variables only, we transform $T(\underline{X})$ into $T[(r, \underline{\Theta})]$. If $T[(r, \underline{\Theta})] = T(\underline{\Theta})$, then we conclude that $T(\underline{X})$ has a unique distribution for all $f(\underline{X}) = \psi(\operatorname{tr} \underline{X}'\underline{X})$.

Suppose the null distribution of a test statistic is not dependent on the radius and that the non-null distribution is related to the null distribution by a simple change of scale. Then the non-null distribution of the test statistic also will not be a function of the radius and hence its distribution will be unique.

# V. TESTING EQUALITY OF k SCALE PARAMETERS

## Section 5.1 Introduction

For testing the equality of k scale parameters it often is assumed that (i) each sample scale statistic, $s_i^2$, is calculated from a random sample, $x_{i1},\ldots,x_{in_i}$, of $n_i$ observations, (ii) each $x_{ij}$, $j=1,\ldots,n_i$, has a Gaussian distribution with mean $\mu_i$ and variance $\sigma_i^2$, and (iii) the sample statistics, $s_i^2$, $i=1,\ldots,k$, are mutually independent.

In this chapter we develop invariant tests when the observations are not necessarily either Gaussian or independent. Before we present this theory we give an example for which an underlying joint spherical distribution might be appropriate. The example involves directional data, that is, angular data measured on a sphere of unit radius as discussed from a different viewpoint by Mardia (1972).

The example is a study of the navigational patterns of homing pigeons. Two groups of birds are released from a point and the bearings of their flight, called vanishing angles, are recorded just as they vanish in the distance. One group is a control while members of the other group have had their internal biological clocks altered. The random variable of interest is the location determined by vanishing angle and the distance traversed by each pigeon. Schmidt-Koenig (1965, pp. 243-246) offers empirical evidence which suggests that, under certain conditions, random patterns exist in the initial orientation of homeward-directed pigeons. For these situations, a spherical distribution seems to be appropriate for each pigeon. We further conjecture that the $n_1$

birds in the control group and the $n_2$ birds in the experimental group all have dependent flights governed by a joint spherical distribution. Under this assumption we then might test whether the two groups differ in terms of the variability of points determined by vanishing angles and distances traveled.

## Section 5.2 Statement of the Problem

Let $L(\underline{x}) \in S_n[\underline{\mu},\underline{\Sigma}]$ with

$$
\begin{aligned}
x'(1\times n) &= [\underline{x}_1'(1\times n_1),\ldots,\underline{x}'_k(1\times n_k)]\\
\underline{\mu}'(1\times n) &= [\underline{\mu}_1',\ldots,\underline{\mu}'_k]\\
\underline{\mu}_i'(1\times n_i) &= [\mu_i,\ldots,\mu_i] \qquad (5.2.1)\\
\underline{\Sigma} &= \mathrm{Diag}[\sigma_1^2\underline{I}_{n_1},\ldots,\sigma_k^2\underline{I}_{n_k}]\\
n &= n_1 + n_2 + \ldots + n_k
\end{aligned}
$$

and consider testing

$$H:\ \sigma_1^2 = \ldots = \sigma_k^2 = \sigma^2 \text{ (say)}$$

(5.2.2) versus

$$K:\ \sigma_i^2 \neq \sigma_j^2$$

for at least one pair $(i,j)$, $i \neq j$.

The unrestricted and restricted parameter spaces are respectively

$$\Omega:\ R_1^{n_1} \times \ldots \times R_1^{n_k} \times R_+^k$$

and

$$\Omega_H:\ R_1^{n_1} \times \ldots \times R_1^{n_k} \times R_+^1$$

with $\underline{\mu}_i \varepsilon R_1^{n_i}$ , $(\sigma_1^2,\ldots,\sigma_k^2) \varepsilon R_+^k$ , and $\sigma^2 \varepsilon R_+^1$ . In our notation, $\underline{a} \varepsilon R_1^{n_i}$ denotes a vector in Euclidean $n_i$-dimensional space with all components equal, while $\underline{b} \varepsilon R_+^k$ denotes a vector in Euclidean k space with all components positive.

The problem remains invariant under the group G acting on the sample space as

$$g[\underline{x}_1',\ldots,\underline{x}_k'] = [c(\underline{x}_1'+\underline{c}_1')\underline{P}_1',\ldots,c(\underline{x}_k'+\underline{c}_k')\underline{P}_k'] \tag{5.2.3}$$

where $P_i \varepsilon \mathcal{O}(n_i)$, $\underline{c}_i \varepsilon R_1^{n_i}$, i=1,...,k, and $c \varepsilon R_+^1$ .

Maximal invariant functions on the sample space and the parameter space are $\underline{z}_i'\underline{z}_i/\underline{z}_j'\underline{z}_j$, i,j=1,...,k, i≠j and $\sigma_i^2/\sigma_j^2$, i,j=1,...,k, i≠j, respectively with

$$\underline{z}_i' = [(x_{i1}-\bar{x}_i),\ldots,(x_{in_i}-\bar{x}_i)]$$

and

$$\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij} .$$

Theorem 5.2.1: For all $L(\underline{x})$ described in (5.2.1) the statistic $T(\underline{z}) = \{\underline{z}_i'\underline{z}_i/\underline{z}_j'\underline{z}_j$, i,j=1,...,k, i≠j} has a unique null and a unique non-null joint distribution.

Proof: (i) Null case

If H is true then $L(\underline{x}) \varepsilon S_n[\underline{\mu},\sigma^2\underline{I}_n]$ and $L(\underline{z}) \varepsilon S_n[\underline{0},\sigma^2\underline{A}]$ for $\underline{A} = \mathrm{Diag}(\underline{A}_1,\ldots,\underline{A}_k)$ and $\underline{A}_i = I_{n_i} - \frac{1}{n_i}\underline{1}\,\underline{1}'$ with $\underline{1}' = (1,\ldots,1)$. Since $\underline{A}_i$ is symmetric and idempotent of rank $n_i - 1$ there exists an orthogonal

$\underline{Q}_i$ such that $\underline{Q}_i\underline{A}_i\underline{Q}_i' = \text{Diag}(1,1,\ldots,1,0)$. The problem remains invariant under pre-multiplication by $\underline{P}_i$ so choose $\underline{P}_i = \underline{Q}_i$. Upon setting $c = \sigma^{-1}$ we may assume that

$$(5.2.4) \qquad L(\underline{z}) \in S_n[\underline{0}, \begin{bmatrix} \underline{I}_{n-k} & \underline{\phi} \\ \underline{\phi} & \underline{\phi} \end{bmatrix}].$$

However, $L(\underline{z})$ is singular and concentrated on a $n - k$ dimensional subspace of $R^n$ with unit probability. As we are interested in proving a property of $T(\underline{z})$ we need only examine the subspace on which the unit mass is nondegenerate. Therefore, from (3.6.1) we assume that $L(\underline{z}) \in S_N[\underline{0},\underline{I}_N]$ with $N = n - k$ and $N_i = n_i - 1$.

Next, we note that $T(\underline{z}) = \{\underline{z}_i'\underline{z}_i/\underline{z}_j'\underline{z}_j,\ i,j=1,\ldots,k,\ i\neq j\}$ is a function of $\underline{z}/(\underline{z}'\underline{z})^{1/2}$ because $\underline{z}'/(\underline{z}'\underline{z})^{1/2} = [\underline{z}_1',\ldots,\underline{z}_k']/(\underline{z}'\underline{z}')^{1/2}$ and $\underline{z}_i'\underline{z}_i/\underline{z}_j'\underline{z}_j = [\underline{z}_i'\underline{z}_i/(\underline{z}'\underline{z})^{1/2}]/[\underline{z}_j'\underline{z}_j/(\underline{z}'\underline{z})^{1/2}]$. Therefore, from Corollary 4.3.1 the distribution of $T(\underline{z})$ is unique for all $L(\underline{z})$.

(ii) Non-null case

The non-null distribution is related to the null distribution by a change of scale. Let $L(\underline{z}) \in S_N[\underline{0},\underline{I}_N]$ and define

$$\begin{gathered} \underline{z}_1 \to \sigma_1\underline{z}_1 \\ \vdots \\ \underline{z}_k \to \sigma_k\underline{z}_k \end{gathered} \quad .$$

Then $\underline{z}_i'\underline{z}_i/\underline{z}_j'\underline{z}_j$ is transformed into

$$(5.2.5) \qquad \left(\frac{\sigma_j^2}{\sigma_i^2}\right)\ \underline{z}_i'\underline{z}_i/\underline{z}_j'\underline{z}_j$$

and part (i) of the proof assures that (5.2.5) has a unique distribution for all $L(\underline{z})$ and hence for all $L(\underline{x})$.

The proof is complete because the method used here is equivalent to finding the distribution of $\{\underline{z}_i'\underline{z}_i/\underline{z}_j'\underline{z}_j,\ i,j=1,\ldots,k,\ i\neq j\}$ when $L(\underline{z}) \ \varepsilon\ S_N[\underline{0},\mathrm{Diag}(\sigma_1^2\underline{I}_{N_1},\ldots,\sigma_k^2\underline{I}_{N_k})]$.

If we let $s_i^2 = \underline{z}_i'\underline{z}_i/N_i$, then we have the following corollary.

Corollary 5.2.1: The usual normal theory test statistics have unique null and non-null distributions for the class of spherically symmetric laws.

Some examples of normal theory test statistics are (a) Hartley's (1950) maximum F test, (b) Cochran's (1941) test, and (c) Bartlett's modified likelihood ratio test [cf. Glaser (1976)] given respectively by

$$T_1 = \frac{\max\{s_i^2\}}{\min\{s_i^2\}} \tag{5.2.6}$$

$$T_2 = \frac{\max\{s_i^2\}}{\sum_{i=1}^{k} s_i^2} \tag{5.2.7}$$

and

$$T_3 = \frac{\prod_{i=1}^{k} (s_i^2)^{\gamma_i/\gamma}}{\frac{1}{\gamma}\sum_{i=1}^{k} \gamma_i s_i^2} \tag{5.2.8}$$

where $\gamma_i = n_i - 1$, and $\gamma = \sum_{i=1}^{k} \gamma_i$.

## Section 5.3 Extensions of Normal Theory Results

Because the statistics (5.2.6) through (5.2.8) have unique null and non-null distributions for spherically symmetric laws $L(\underline{x})$, we can choose a particular member of the class to (i) derive exact distributions of test statistics, (ii) show unbiasedness, and (iii) derive power properties. We choose the Gaussian distribution because so many results are available. The literature contains the following normal theory results which now are seen to hold when the underlying distribution is spherically symmetric.

Glaser (1976) tabulates exact critical values for the modified likelihood ratio test when the sample sizes are equal, while Chao and Glaser (1978) give an expression for computing exact critical values when the sample sizes are unequal. Carter and Srivastava (1977) show that the power of the modified likelihood ratio test is a monotone increasing function of $\delta_i = \sigma_i^2/\sigma_{i+1}^2$, $i=1,\ldots,k-1$, with $\sigma_1^2 \geq \ldots \geq \sigma_k^2$, when the remaining $k - 2$ parameters, $\delta_j = \sigma_j^2/\sigma_{j+1}^2$, $j=1,\ldots,k-1$, $j \neq i$, are held fixed.

Ramachandran (1956b) proves that Hartley's maximum F test is unbiased while Hartley (1950), David (1952), and Hartmann (1969) tabulate critical values. Cochran (1941) gives a brief table of critical values for his test.

To summarize, we first showed that Bartlett's modified likelihood ratio test, Hartley's maximum F test, and Cochran's test are invariant tests. Then we proved that these test statistics have unique null and non-null distributions for all underlying spherically symmetric laws.

## Section 5.4 Comparison of k-1 Scale Parameters to a Control

We make the same assumptions as in (5.2.1) except we now wish to test simultaneously the collection of hypotheses

(5.4.1) $$H = \{H_i: \sigma_i^2 = \sigma_k^2,\ i=1,\ldots,k-1\}.$$

It is not clear that this problem is amenable to reduction through invariance. Therefore, we proceed by considering each hypothesis in (5.4.1) individually. Next, we apply the union-intersection principle. Gnanadesikan (1959) uses this approach when $L(\underline{x})$ has a Gaussian distribution.

Invariance considerations applied marginally give the test ratios $\underline{z}_i'\underline{z}_i/\underline{z}_k'\underline{z}_k$, $i=1,\ldots,k-1$ . Using Theorem 5.2.1 we have that the statistics

$$T_i(\underline{z}) = \frac{\underline{z}_i'\underline{z}_i/N_i}{\underline{z}_k'\underline{z}_k/N_k},\ N_i=n_i-1,\ i=1,\ldots,k$$

have unique null and non-null joint distributions. Therefore, the test using $T_i(\underline{z})$ has acceptance region

(5.4.2) $$F_{i1} \leq T_i(\underline{z}) \leq F_{i2}$$

with $L(T_i(\underline{z}))$ a central F distribution having $N_i$ and $N_k$ degrees of freedom.

Next, we accept H if all the individual hypotheses, $H_i$, are accepted. That is, H can be expressed as

$$H:\ \bigcap_{i=1}^{k-1} H_i\ .$$

By the union-intersection principle [cf. Roy (1953)], the simultaneous test has acceptance region

$$(5.4.3) \qquad F_{11} \leq T_1(\underline{z}) \leq F_{12}, \ldots, F_{(k-1)1} \leq T_{k-1}(\underline{z}) \leq F_{(k-1)2} \; .$$

Known normal theory results for this problem automatically extend to the class of spherically symmetric distributions. We now give some of these results from the literature.

Ramachandran (1956a) gives an expression for evaluating one-sided probability statements when they are in the form

$$(5.4.4) \qquad P[T_i(\underline{z}) \leq a_i, \; i=1,\ldots,k-1] = 1 - \alpha.$$

A simplified expression emerges for the special case of $N_i = N$, $i=1,\ldots,k-1$. This case has $P[u = \max \underline{z}_i'\underline{z}_i/\underline{z}_k'\underline{z}_k \leq b] = 1 - \alpha$ with u called the studentized largest chi-squared statistic. Ramachandran (1956a) derives the pdf of u and gives upper 5% points for different values of N and $N_k$ for the case $k = 2$.

More extensive tables are available for this case. Gupta (1963) obtains critical values for $\alpha = .25; .10; .05; .01$, $N = 2(2)50$, and $k = 2(1)11$ under the additional assumption that $N = N_k$. Armitage and Krishnaiah (1964) obtain critical values for $\alpha = .10; .05; .025; .01$, $N = 1(1)19$, $k = 2(1)13$, and $N_k = 5(1)45$.

Krishnaiah (1965) shows that the power of the test associated with (5.4.4) is a monotone increasing function of each $\delta_{ik} = \sigma_i^2/\sigma_k^2$.

So far we have discussed only upper tail alternatives. Tables also exist for the studentized smallest chi-square [cf. Gupta and Sobel (1962), Ramachandran (1958), and Krishnaiah and Armitage (1964)].

Results for the two sided case are not available. However, we can apply Boole's inequality, $P(\bigcup_{i=1}^{k-1} E_i) \le \sum_{i=1}^{k-1} P(E_i)$, to obtain a conservative simultaneous testing procedure having all the properties under spherical laws which are enjoyed under Gaussian assumptions. From (5.4.3) we have that $E_i = F_{i1} \le T_i(\underline{z}) \le F_{i2}$, $i=1,\ldots,k-1$.

## VI. TESTING THE EQUALITY OF k SCALE MATRICES

### Section 6.1 Statement of the Problem

We undertake an extension of the testing problem considered in the preceding chapter. Instead of testing for the equality of k scale parameters, $\sigma_1^2,\ldots,\sigma_k^2$, we test for the equality of k scale matrices, $\underline{\Sigma}_1,\ldots,\underline{\Sigma}_k$.

Suppose $L(\underline{X}) \in S_{nm}[\underline{M},\underline{\Lambda}]$ with

(6.1.1)
$$\begin{aligned}
\underline{M}'(m\times n) &= [\underline{M}_1'(m\times n_1),\ldots,\underline{M}_k'(m\times n_k)]\\
\underline{M}_i' &= [\underline{\mu}_i',\ldots,\underline{\mu}_i']\\
\underline{\Lambda} &= \mathrm{Diag}[\underline{I}_{n_1} \otimes \underline{\Sigma}_1,\ldots,\underline{I}_{n_k} \otimes \underline{\Sigma}_k]\\
&\quad \underline{\Sigma}_i(m\times m),\ i=1,\ldots,k\\
n &= n_1 + \ldots + n_k\\
&\quad n_i > m,\ i=1,\ldots,k.
\end{aligned}$$

We wish to test

(6.1.2)
$$H\colon \underline{\Sigma}_1 = \underline{\Sigma}_2 = \ldots = \underline{\Sigma}_k = \underline{\Sigma} \text{ (say)}$$

agains the general alternative K: $\{\underline{\Sigma}_i \neq \underline{\Sigma}_j$, for at least one $i\neq j\}$.

Let $\mathcal{D}_m^+(k)$ denote the group of positive definite block diagonal matrices such that, if $\underline{A} \in \mathcal{D}_m^+(k)$ then $\underline{A} = \mathrm{Diag}(\underline{A}_1,\ldots,\underline{A}_k)$ with $\underline{A}(mk\times mk)$ and $\underline{A}_i(m\times m)$, $i=1,\ldots,k$.

The unrestricted and restricted parameter spaces for the problem are

$$\Omega: \quad \mathcal{F}^{1}_{n_1 \times m} \times \dots \times \mathcal{F}^{1}_{n_k \times m} \times \mathcal{D}^{+}_{m}(k)$$

and

$$\Omega_H: \quad \mathcal{F}^{1}_{n_1 \times m} \times \dots \times \mathcal{F}^{1}_{n_k \times m} \times \mathcal{S}^{+}_{m}$$

where $\underline{S} \in \mathcal{S}^{+}_{m}$ indicates that $\underline{S}(m \times m)$ is positive definite and $\underline{F}_i \in \mathcal{F}^{1}_{n_i \times m}$ indicates that $\underline{F}_i(n_i \times m)$ has $n_i$ equal row vectors each of dimension m.

The problem remains invariant under the group G such that, for $g \in G$,

$$(6.1.3) \qquad g[\underline{X}_1', \dots, \underline{X}_k'] = \underline{B}'[\underline{X}_1' + \underline{C}_1', \dots, \underline{X}_k' + \underline{C}_k'] \cdot \mathrm{Diag}(\underline{P}_1', \dots, \underline{P}_k')$$

where $\underline{P}_i \in \mathcal{O}(n_i)$, $\underline{C}_i \in \mathcal{F}^{1}_{n_i \times m}$, $i=1,\dots,k$, and $\underline{B} \in GL(m)$, with $GL(m)$ the group of all non-singular $(m \times m)$ matrices.

The sample maximal invariants under the action of the group G, and the parametric maximal invariants under the induced group $\bar{G}$, are given in the next theorem.

Theorem 6.1.1: The respective maximal invariants for the sample and parameter spaces are the roots of

$$(6.1.4) \qquad |\underline{Z}_i'\underline{Z}_i - d\underline{Z}_j'\underline{Z}_j| = 0$$

and

$$|\underline{\Sigma}_i - \lambda \underline{\Sigma}_j| = 0, \quad i,j=1,\dots,k, \quad i \neq j$$

with

$$\underline{Z}_i' = [(\underline{x}_{i1}-\underline{\bar{x}}_{i\cdot}),\ldots,(\underline{x}_{in_i}-\underline{\bar{x}}_{i\cdot})]$$

$$\underline{\bar{x}}_{i\cdot} = \frac{1}{n_i}\sum_{j=1}^{n_i} \underline{x}_{ij} \quad .$$

Proof: Let $G_1$ and $G_2$ be two groups such that

$$g_1[\underline{X}_1',\ldots,\underline{X}_k'] = [\underline{X}_1'+\underline{C}_1',\ldots,\underline{X}_k'+\underline{C}_k']\cdot\text{Diag}(\underline{P}_1',\ldots,\underline{P}_k'),$$

$$\underline{C}_i \in F_{n_i\times m},\ \underline{P}_i \in \mathcal{O}(n_i)$$

and

$$g_2\underline{X} = \underline{X}\ \underline{B},\ \underline{B} \in GL(m).$$

Maximal invariant statistics under $G_1$ are $\underline{Z}_1'\underline{Z}_1,\ldots,\underline{Z}_k'\underline{Z}_k$ [cf. Lehmann (1959), p. 296]. Theorem 4.2.1 now assures us that we need only show, under $G_2$, that (6.1.4) is maximal invariant within the space of $\underline{Z}_1'\underline{Z}_1,\ldots,\underline{Z}_k'\underline{Z}_k$.

Under $G_2$, $\underline{Z}_i'\underline{Z}_i \to \underline{B}'\underline{Z}_i'\underline{Z}_i\underline{B}$, $i=1,\ldots,k$. The roots of $|\underline{Z}_i'\underline{Z}_i-d\underline{Z}_j'\underline{Z}_j| = 0$, $i,j=1,\ldots,k$, $i\neq j$ are invariant. We show these roots are maximal invariant by assuming that there are matrices $\underline{U}_i$, $\underline{U}_j$ such that

$$|\underline{Z}_i'\underline{Z}_i-d\underline{Z}_j'\underline{Z}_j| = 0 = |\underline{U}_i'\underline{U}_i-d\underline{U}_j'\underline{U}_j| \quad .$$

Lehmann (1959, p. 278) proves that there is some $\underline{B}$ which transforms $\underline{Z}_i'\underline{Z}_i$ into $\underline{U}_i'\underline{U}_i$ and $\underline{Z}_j'\underline{Z}_j$ into $\underline{U}_j'\underline{U}_j$. Thus (6.1.4) is maximal invariant. The proof for the parameter space is similar.

Let $D^{ij} = \{d_1^{ij}, \ldots, d_m^{ij}\}$ be the roots of $|\underline{Z}_i'\underline{Z}_i - \underline{Z}_j'\underline{Z}_j| = 0$ for $i,j=1,\ldots,k$, $i \neq j$. Then we have the following theorem.

Theorem 6.1.2: For both the null and non-null cases the joint distribution of $\{D^{ij},\ i,j=1,\ldots,k,\ i\neq j\}$ does not depend on any particular underlying elliptically symmetric distribution. That is, the null and non-null distributions are unique for all $L(\underline{X})$ defined in (6.1.1).

Proof: (i) Null Case

If H is true then $L(\underline{X}) \in S_{nm}[[\underline{M}_1', \ldots, \underline{M}_k'], \mathrm{Diag}(\underline{I}_{n_1} \otimes \underline{\Sigma}, \ldots, \underline{I}_{n_k} \otimes \underline{\Sigma})]$ and $L(\underline{Z}) \in S_{nm}[\underline{\phi}, \mathrm{Diag}(\underline{F}_1 \otimes \underline{\Sigma}, \ldots, \underline{F}_k \otimes \underline{\Sigma})]$ with $\underline{F}_i = \underline{I}_{n_i} - \frac{1}{n_i}\underline{1}\,\underline{1}'$. Using arguments similar to those in the proof of Theorem 5.2.1 we assume that

$$(6.1.5) \qquad L(\underline{Z}) \in S_{nm}\left[\underline{\phi}, \begin{bmatrix} \underline{I}_{n-k} \otimes \underline{I}_m & \underline{\phi} \\ \underline{\phi} & \underline{\phi} \end{bmatrix}\right].$$

That is, choose $\underline{P}_i$ and $\underline{B}$ in (6.1.3) so that $\underline{P}_i\underline{F}_i\underline{P}_i' = \mathrm{Diag}(1,1,\ldots,1,0)$ and $\underline{B}'\Sigma\underline{B} = \underline{I}_m$.

Now (6.1.5) can be expressed as a distribution in $R^{nm}$ with unit probability concentrated on a $r = (n-k)m = Nm$-dimensional subspace of $R^{nm}$. We assume that

$$(6.1.6) \qquad L(\underline{Z}) \in S_{Nm}[\underline{\phi}, \underline{I}_N \otimes \underline{I}_m]$$

as we need to concern ourselves only with this subspace. This follows from Lemma 3.6.1.

Next, we embed $\underline{Z}$ in $R^{Nm}$ and change to polar coordinates, i.e.

(6.1.7) $$\underline{Z} \to \underline{z} \to (r, \underline{\theta})$$

Now put the extreme right hand side of (6.1.7) back into matrix form, i.e.

$$(r, \underline{\theta}) \to (r, \underline{\Theta})$$

where $\underline{\Theta}(N \times m)$ is a matrix containing angular variables.

Since $\underline{\theta}'\underline{\theta} = 1$ if and only if tr $\underline{\Theta}'\underline{\Theta} = 1$ we have that $(r, \underline{\Theta})$ is a polar transformation of the matrix $\underline{Z}$. Under this transformation, $|\underline{Z}_i'\underline{Z}_i - d\underline{Z}_j'\underline{Z}_j| = 0$ becomes $|r^2\underline{\Theta}_i'\underline{\Theta}_i - dr^2\underline{\Theta}_j'\underline{\Theta}_j| = 0 = |\underline{\Theta}_i'\underline{\Theta}_i - d\underline{\Theta}_j'\underline{\Theta}_j|$, and the $D^{ij}$ do not depend on $r^2 = \text{tr } \underline{Z}'\underline{Z}$. Therefore, the null distribution must be unique for all $L(\underline{Z})$ and also for all $L(\underline{X})$.

(ii) Non-null case

Suppose $L(\underline{Z}) \in S_{Nm}[\underline{\phi}, \text{Diag}(\underline{I}_{N_1} \otimes \underline{\Sigma}_1, \ldots, \underline{I}_{N_k} \otimes \underline{\Sigma}_k)]$. Let

$$\begin{aligned} \underline{Y}_1 &= \underline{Z}_1\underline{\Sigma}_1^{-1/2} \\ Y_2 &= \underline{Z}_2\underline{\Sigma}_2^{-1/2} \\ &\vdots \\ \underline{Y}_k &= \underline{Z}_k\underline{\Sigma}_k^{-1/2} \end{aligned} \quad .$$

Then $L(\underline{Y}_1, \underline{Y}_2, \ldots, \underline{Y}_k) \in S_{Nm}[\underline{\phi}, \text{Diag}(\underline{I}_{N_1} \otimes \underline{I}_m, \ldots, \underline{I}_{N_k} \otimes \underline{I}_m)]$. As in part (i) let

$$\underline{Y} \to \underline{y} \to (r, \underline{\theta}) \to (r, \underline{\Theta})$$

and note that

$$|\underline{Z}_i'\underline{Z}_i - d\underline{Z}_j'\underline{Z}_j| = 0 \rightarrow$$

$$|\underline{\Sigma}_i^{1/2}\underline{Y}_i'\underline{Y}_i\underline{\Sigma}_i^{1/2} - d\underline{\Sigma}_j^{1/2}\underline{Y}_j'\underline{Y}_j\underline{\Sigma}_j^{1/2}| = 0$$

$$\rightarrow |\underline{\Sigma}_i^{1/2}\underline{\Theta}_i'\underline{\Theta}_i\underline{\Sigma}_i^{1/2} - d\underline{\Sigma}_j^{1/2}\underline{\Theta}_j'\underline{\Theta}_j\underline{\Sigma}_j^{1/2}| = 0 \quad .$$

It follows that the non-null distribution of $\{D^{ij}, i,j=1,\ldots,k, i\neq j\}$ does not depend on tr $\underline{X}'\underline{X}$ and thus is unique for all $L(\underline{X})$.

One example of a test statistic with unique null and non-null distributions for $L(\underline{X})$ defined in (6.1.1) is the normal-theory modified likelihood ratio statistic

$$\text{(6.1.7)} \qquad T_1 = \frac{\prod_{i=1}^{k} |\underline{Z}_i'\underline{Z}_i|^{\frac{1}{2}\gamma_i}}{|\underline{Z}'\underline{Z}|^{\frac{1}{2}\gamma}}$$

with $\gamma_i = N_i$ and $\gamma = \sum_{i=1}^{k} \gamma_i$.

Another example is a test statistic proposed by Pillai and Young (1974). Let

$$\text{(6.1.8)} \qquad T_2 = \frac{\max_{1\le i\le k} U_i}{\min_{1\le j\le k} U_j}$$

where

$$U_i = N_{1,i+1} \text{ tr } \underline{S}_{2,i}\, \underline{S}_{1,i+1}^{-1}/N_{2,i}, \quad i=1,\ldots,k-1$$

$$U_k = \begin{cases} N_{2,k} \text{ tr } \underline{S}_{1,1}\, \underline{S}_{2,k}^{-1}/N_{1,1}, & k > 2 \\ N_{1,1} \text{ tr } \underline{S}_{2,2}\, \underline{S}_{1,1}^{-1}/N_{2,2}, & k = 2 \end{cases}$$

and

$$\underline{S}_{\ell,i} = \sum_{j=1}^{n_{\ell,i}} (\underline{x}_{ij}^{(\ell)} - \underline{\bar{x}}_{i\cdot}^{(\ell)})(\underline{x}_{ij}^{(\ell)} - \underline{\bar{x}}_{i\cdot}^{(\ell)})', \quad \ell = 1,2$$

is calculated after the original sample of $n_i$ observations is split into two sub-samples of $n_{1,i}$ and $n_{2,i}$ observations. The statistic $T_2$ is called the maximum U-ratio.

We also investigate the special case of testing H: $\underline{\Sigma}_1 = \underline{\Sigma}_2$ versus K: $\underline{\Sigma}_1 \neq \underline{\Sigma}_2$ before we mention extensions of normal-theory results to the larger class of distributions. The respective maximal invariant statistics and parameters are the roots of $|\underline{Z}_1'\underline{Z}_1 - d\underline{Z}_2'\underline{Z}_2| = 0$ and the roots of $|\underline{\Sigma}_1 - \lambda\underline{\Sigma}_2| = 0$. Examples of test statistics with unique null and non-null distributions are:

(6.1.9)
$$T_3 = \frac{|\underline{Z}_1'\underline{Z}_1|^{\frac{1}{2}\gamma_1} \; |\underline{Z}_2'\underline{Z}_2|^{\frac{1}{2}\gamma_2}}{|\underline{Z}_1'\underline{Z}_1 + \underline{Z}_2'\underline{Z}_2|^{\frac{1}{2}\gamma}}$$

This is the special case of (6.1.7) with k = 2.

(6.1.10)
$$T_4 = ch_1[(\underline{Z}_1'\underline{Z}_1)(\underline{Z}_2'\underline{Z}_2)^{-1}]$$

(Roy's largest root)

(6.1.11)
$$T_5 = tr(\underline{Z}_1'\underline{Z}_1)(\underline{Z}_2'\underline{Z}_2)^{-1}$$

(Lawley-Hotelling trace)

(6.1.12)
$$T_6 = tr(\underline{Z}_1'\underline{Z}_1)(\underline{Z}_1'\underline{Z}_1 + \underline{Z}_2'\underline{Z}_2)^{-1}$$

(Pillai trace)

Section 6.2 Extensions of Normal-Theory Results

Because the usual normal theory tests have unique distributions with respect to the underlying class of elliptically symmetric laws, we present some normal theory results which extend to this larger class. These results pertain mostly to tests for H: $\underline{\Sigma}_1 = \underline{\Sigma}_2$ versus K: $\underline{\Sigma}_1 \neq \underline{\Sigma}_2$.

Sugiura and Nagao (1968) show that the modified likelihood ratio test using $T_3$ is unbiased. Das Gupta and Giri (1973) investigate admissible critical regions for $T_3$ and give conditions for unbiasedness.

Anderson and Das Gupta (1964b) consider tests dependent on the roots of $(\underline{Z}_1'\underline{Z}_1)(\underline{Z}_2'\underline{Z}_2)^{-1}$ and investigate tests for the alternatives

$$K_1\colon \lambda_i \geq 1,\ i=1,\ldots,m,\ \sum_{i=1}^{m} \lambda_i > m$$

and

$$K_2\colon \lambda_i \leq 1,\ i=1,\ldots,m,\ \sum_{i=1}^{m} \lambda_i < m$$

where $\{\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m\}$ are the ordered roots of $\underline{\Sigma}_1\underline{\Sigma}_2^{-1}$. They show that the power is a monotonically increasing function of each $\lambda_i$ for a class of test statistics which includes $T_4$ and $T_5$.

Mikhail (1962) investigates the test statistics $T_4$ and $ch_m[(\underline{Z}_1'\underline{Z}_1)(\underline{Z}_2'\underline{Z}_2)^{-1}]$, where $ch_m(\underline{A})$ denotes the smallest root of the matrix $\underline{A}$. He studies the four alternatives $\lambda_1 > 1$, $\lambda_1 < 1$, $\lambda_m > 1$, and $\lambda_m < 1$ and shows in all cases that the power is a monotonically increasing function of each $\lambda_i$.

Pillai and Jayachandran (1968) compare the power for $T_3$, $T_4$, $T_5$, and $T_6$. They conclude that none of these four test statistics is uniformly best. They also derive the exact null distributions of $T_3$, $T_5$, and $T_6$ for the case $m = 2$.

Krishnaiah and Chang (1973) express the exact null distributions of $T_3$ and $T_4$ as linear combinations of inverse Laplace transformations of the products of certain double integrals, while Giri (1968) shows that $T_6$ is a locally best invariant test with power monotonically increasing in each $\lambda_i$.

Pillai and Young (1974) give the exact null distribution of the maximum U ratio for the case $k = 2$ and $m = 2$ and tabulate percentage points for the values $\alpha = .05$, $N_1 = 3(2)7$, and $N_2 = 10(5)40$; 50; 60; 80; 100.

In summary we note that the joint null and non-null distributions of the maximal invariant statistics are unique for all underlying elliptically symmetric distributions. Specifically the usual normal-theory test statistics have distributions which are exactly robust. Therefore, all results derived under a Gaussian assumption automatically extend to this larger class.

# VII. TESTS FOR SPHERICITY

## Section 7.1 Statement of the Problem

The hypothesis of sphericity and its alternatives are

$$H: \underline{\Sigma} = \sigma^2 \underline{V}$$

(7.1.1) versus

$$K: \underline{\Sigma} \neq \sigma^2 \underline{V}$$

with $\sigma^2$ an unknown parameter and $\underline{V}$ a fixed matrix.

This hypothesis is usually tested under Gaussian assumptions, that is, we assume $L(\underline{x}) = N_m[\underline{\mu}, \underline{\Sigma}]$ and take a random sample of size n from this population. If $\underline{V} = \underline{I}_n$ then this is a test for the homoscedasticity and independence of the components $x_1, \ldots, x_m$ of $\underline{x}$.

We test this hypothesis for the case $L(\underline{X}) \varepsilon S_{nm}[\underline{M}, \underline{I}_n \otimes \underline{\Sigma}]$ with $\underline{M}' = [\underline{\gamma}, \ldots, \underline{\gamma}] \varepsilon \mathcal{F}^1_{n\times m}$ and $\underline{\Sigma} \varepsilon \mathcal{S}^+_m$. Under this larger class of distributions (7.1.1) is still the hypothesis of sphericity as we now are testing that $\underline{X}$ has a spherical distribution against the alternative that $\underline{X}$ has an elliptically symmetric distribution.

We now put the problem in canonical form before undertaking an invariant reduction. Let $\underline{Y} = \underline{XV}^{-1/2}$, then

$$L(\underline{Y}) \ \varepsilon \ S_{nm}[\underline{N}, \underline{I}_n \otimes \underline{\Lambda}] \tag{7.1.2}$$

with

$$\underline{N}' = \underline{V}'^{-1/2}\, \underline{M}' = [\underline{\mu},\ldots\underline{\mu}] \in F^1_{n\times m}$$

$$\underline{\Lambda} = \underline{V}^{-1/2}\, \underline{\Sigma}\, \underline{V}^{-1/2} \in S^+_m .$$

An equivalent form of (7.1.1) is

$$H: \underline{\Lambda} = \sigma^2 \underline{I}_m$$

(7.1.3) versus

$$K: \underline{\Lambda} \neq \sigma^2 \underline{I}_m$$

with unrestricted and restricted parameter spaces

$$\Omega: \quad F^1_{n\times m} \times S^+_m$$

and

$$\Omega_H: \quad F^1_{n\times m} \times \underline{I}_m \times R^1_+ .$$

The group G for which the testing problem remains invariant is given in the following lemma.

Lemma 7.1.1: The testing problem is invariant under the group G, such that for $g \in G$,

(7.1.4) $$g\underline{Y} = c\underline{P}(\underline{Y}+\underline{B})\underline{Q}$$

with $\underline{P} \in O(n)$, $\underline{Q} \in O(m)$, $\underline{B} \in F^1_{n\times m}$ and $c \in R^1_+$ .

Proof: Observe that

$$\bar{g}[\underline{N},\underline{\Lambda}] = [c\underline{P}(\underline{N}+\underline{B})\underline{Q}, c^2\underline{Q}'\underline{\Lambda}\underline{Q}] \in \Omega$$

and

$$\bar{g}[\underline{N},\sigma^2\underline{I}_n] = [c\underline{P}(\underline{N}+\underline{B})\underline{Q},c^2\sigma^2\underline{I}_n] \; \varepsilon \; \Omega_H \; .$$

It is further true that $L(g\underline{Y}) \; \varepsilon \; S_{nm}[c\underline{P}(\underline{N}+\underline{B})\underline{Q},\underline{I}_n \otimes c^2\underline{Q}'\underline{\Lambda}\underline{Q}]$ and $P_\theta \to P_{\theta'}$ implies that $\theta' = \bar{g}\theta$ for $\theta' = [c\underline{P}(\underline{N}+\underline{B})\underline{Q},c^2\underline{Q}'\underline{\Lambda}\underline{Q}]$. Therefore, the problem remains invariant under G.

In the next theorem we give sample and parametric maximal invariants under the action of the group G and the action of the induced group $\bar{G}$.

Theorem 7.1.1: Maximal invariant statistics and parameters are respectively

$$\frac{d_1}{\sum_{i=1}^{m} d_i},\ldots,\frac{d_m}{\sum_{i=1}^{m} d_i} \tag{7.1.5}$$

and

$$\frac{\lambda_1}{\sum_{i=1}^{m} \lambda_i},\ldots,\frac{\lambda_m}{\sum_{i=1}^{m} \lambda_i} \; . \tag{7.1.6}$$

Here $\{d_1,\ldots,d_m\}$ contains the roots of $|\underline{Z}'\underline{Z}-d\underline{I}_m| = 0$ and $\{\lambda_1,\ldots,\lambda_m\}$ contains the roots of $|\underline{\Lambda}-\lambda\underline{I}_m| = 0$ with

$$\underline{Z}' = [(\underline{y}_1-\bar{\underline{y}}),\ldots,(\underline{y}_n-\bar{\underline{y}})]$$

and

$$\bar{\underline{y}} = \frac{1}{n}\sum_{j=1}^{n} \underline{y}_j \; .$$

<u>Proof</u>: Let $G_1$ and $G_2$ be two groups such that

$$g_1\underline{Y} = \underline{P}(\underline{Y}+\underline{B})\underline{Q},\ \underline{P} \in \mathcal{O}(n),\ \underline{Q} \in \mathcal{O}(m),\ \underline{B} \in \mathcal{F}^1_{n\times m}$$

and

$$g_2\underline{Y} = c\underline{Y},\ c \in \mathcal{R}^1_+ .$$

We can show that the maximal invariant statistics under $G_1$ are the roots of $|\underline{Z}'\underline{Z}-d\underline{I}_m| = 0$ by using an argument similar to the proof of Theorem 6.1.1. Theorem 4.2.1 shows that we need only consider the action of $G_2$ on the space of eigenvalues $\{d_1,\ldots,d_m\}$ in order to prove maximal invariance. Suppose $\{d_1,\ldots,d_m\}$ and $\{e_1,\ldots,e_m\}$ are such that

$$d_j/\sum_{i=1}^m d_i = e_j/\sum_{i=1}^m e_i,\ j=1,\ldots,m.$$

It follows that

$$e_j = d_j \sum_{i=1}^m e_i/\sum_{i=1}^m d_i = cd_j$$

and thus (7.1.5) is maximal invariant. A similar proof holds for the action of $\bar{G}$ on the parameter space.

Now if H is true then $L(\underline{Y}) \in S_{nm}[N,\underline{I}_n \otimes \sigma^2\underline{I}_n]$ and $L(\underline{Z}) \in S_{nm}[\underline{\phi},\underline{F} \otimes \sigma^2\underline{I}_m]$ with $\underline{F} = \underline{I}_n - \frac{1}{n}\underline{1}\,\underline{1}'$. As before, we assume that

(7.1.7) $$L(\underline{Z}) \in S_{Nm}[\underline{\phi},\underline{I}_N \otimes \underline{I}_m]$$

with $N = n - 1$.

All invariant tests necessarily depend on the maximal invariant statistics. Therefore, we conclude that an invariant test statistic must be scale invariant, i.e. it cannot depend on the tr $\underline{Y}'\underline{Y}$. This implies that all invariant test statistics have unique null distributions for $L(\underline{Z})$ defined in (7.1.7).

The main result pertaining to invariant tests for sphericity in the class $S_{Nm}[\underline{M},\underline{I}_N \otimes \underline{\Sigma}]$ is the following invariance property of the non-null distribution.

Theorem 7.1.2: All invariant statistics for testing the sphericity hypothesis have unique non-null distributions for $L(\underline{Z}) \varepsilon S_{Nm}[\underline{\phi},\underline{I}_N \otimes \underline{\Sigma}]$.

Proof: Let $\underline{W} = \underline{Z}\ \underline{\Sigma}^{-1/2}$. Then $L(\underline{W}) \varepsilon S_{Nm}[\underline{\phi},\underline{I}_N \otimes \underline{I}_m]$. Imbed W in $\mathcal{R}^{Nm}$ and change to polar coordinates. Then $(r,\underline{\Theta})$, where $\underline{\Theta}$ is of order $(N\times m)$, is a polar transformation of $\underline{W}$. Under this transformation

$$|\underline{Z}'\underline{Z}-d\underline{I}_m| = 0 \rightarrow$$

$$|\underline{W}'\underline{W}-d\underline{\Sigma}^{-1}| = 0 \rightarrow$$

$$|\underline{\Theta}'\underline{\Theta} - \frac{d}{r^2}\underline{\Sigma}^{-1}| = 0 \quad . \tag{7.1.8}$$

From (7.1.8) we see that $\frac{d_1}{\sum_{i=1}^{m} d_i}, \ldots, \frac{d_m}{\sum_{i=1}^{m} d_i}$ is not a function of $r^2 = \text{tr}\ \underline{Z}'\underline{Z}$ and the proof is complete.

Some examples of invariant test statistics follow. The first is the modified likelihood ratio statistic under normal theory. The other three examples have been discussed briefly in the literature. We discuss all four examples in more detail in Section 7.3.

(7.1.9) (i) $$T_1 = \left[\frac{|\underline{Z}'\underline{Z}|^{1/m}}{\frac{1}{m}\,\mathrm{tr}\,\underline{Z}'\underline{Z}}\right]^{\frac{1}{2}mN}$$

(7.1.10) (ii) $$T_2 = \frac{\mathrm{tr}\ (\underline{Z}'\underline{Z})^2}{(\mathrm{tr}\ \underline{Z}'\underline{Z})^2}$$

(7.1.11) (iii) $$T_3 = \frac{ch_1\ (\underline{Z}'\underline{Z})}{\mathrm{tr}\ \underline{Z}'\underline{Z}}$$

(7.1.12) (iv) $$T_4 = \frac{ch_m\ (\underline{Z}'\underline{Z})}{\mathrm{tr}\ \underline{Z}'\underline{Z}}$$

Before we conclude, we stress the significance of the scale invariance in (7.1.4). This invariance gives the property of scale invariance for the maximal invariant statistics, and, in turn, implies that all invariant test statistics have unique null and non-null distributions for $L(\underline{X}) \in S_{nm}[\underline{M},\underline{I}_n \otimes \underline{\Sigma}]$. The importance of this scale invariance is demonstrated further in the next section when we consider a hypothesis similar to sphericity, but involving test statistics which are not invariant under the group of scale changes.

## Section 7.2 The case $\sigma^2 = 1$

We assume the model (7.1.2) and we wish to test

$$H: \underline{\Lambda} = \underline{I}_m$$

(7.2.1) versus

$$K: \underline{\Lambda} \neq \underline{I}_m$$

with unrestricted and restricted parameter spaces

$$\Omega: \quad \mathcal{F}^1_{n\times m} \times \mathcal{S}^+_m$$

and

$$\Omega_H: \mathcal{F}^1_{n\times m} \times \underline{I}_m \quad .$$

The testing problem remains invariant under the group G such that, for $g \in G$,

(7.2.2) $$g\underline{Y} = \underline{P}(\underline{Y}+\underline{B})\underline{Q}$$

with $\underline{P} \in \mathcal{O}(n)$, $\underline{Q} \in \mathcal{O}(m)$, and $\underline{B} \in \mathcal{F}^1_{n\times m}$.

Section 7.1 shows that the maximal invariant statistics and parameters are the roots of $|\underline{Z}'\underline{Z}-d\underline{I}_m| = 0$ and the roots of $|\underline{\Lambda}-\lambda\underline{I}_m| = 0$, respectively.

Some examples of invariant test statistics are

(7.2.3) $$T_5 = ch_1\ (\underline{Z}'\underline{Z})$$

(7.2.4) $$T_6 = tr\ \underline{Z}'\underline{Z}$$

and

(7.2.5) $$T_7 = |\underline{Z}'\underline{Z}|^{\frac{1}{2}N} \exp[tr\ (-\tfrac{1}{2}\ \underline{Z}'\underline{Z})].$$

Other examples of invariant test statistics for this case are given by $T_1$, $T_2$, $T_3$, and $T_4$ from the preceding section. However, $T_5$, $T_6$, and $T_7$ do not have null and non-null distributions independent of a specific underlying elliptically symmetric law.

To show this, let $\underline{Z} \to (r, \underline{\Theta})$. Then

$$|\underline{Z}'\underline{Z} - d\underline{I}_m| = 0 \to$$

$$|r^2\underline{\Theta}'\underline{\Theta} - d\underline{I}_m| = 0$$

and the null and non-null distributions of the maximal invariant statistics are seen to depend on r and thus on the underlying elliptically symmetric law. Since $T_5$, $T_6$, and $T_7$ depend directly on these maximal invariant statistics, they also have non-unique null and non-null distributions.

By completely specifying $\sigma^2$ we eliminate a restriction on the sample and parameter spaces. This produces a larger group of invariant test statistics, some of which do not have distributions independent of the underlying elliptically symmetric law.

## Section 7.3 Extensions of Normal-Theory Results

The test statistics $T_1$, $T_2$, $T_3$, and $T_4$ in Section 7.1 have unique null and non-null distributions. Therefore, the following normal-theory results extend immediately to the class of elliptically symmetric distributions.

John (1971) shows that the test based upon $T_2$ is a locally best invariant test. He later obtains the exact null distribution of $T_2$ for $m > 2$ [cf. John (1972)]. Finally, John (1976) tabulates critical values of $V = (1-m^{-1})^{-1}(T_2-m^{-1})$ for the cases $N = 3(1)16; 18(2)24; 180$, $\alpha = .1; .05; .01; .001$, and $m = 3$. He also gives a table for $\alpha = .05$, $m = 4(1)10$, and the N values given previously.

Gleser (1966) and Das Gupta (1969) show that $|\underline{Z}'\underline{Z}|/\text{tr } (\underline{Z}'\underline{Z})^m$ is an unbiased test statistic while Nagarsenker and Pillai (1973) find the exact null distribution in series form of $|\underline{Z}'\underline{Z}|/\text{tr } (\underline{Z}'\underline{Z}/m)^m$ and calculate percentage points for the cases $m = 4(1)10$, $\alpha = .01; .05$, and for the values $n = 5(1)20; 22(2)30; 34(8)50; 60(20)100; 140; 200; 300$.

Schuurmann, Krishnaiah, and Chattopadhyay (1973) derive the exact null distributions of $T_3$ and $T_4$. They tabulate critical values for $m = 3(1)16$, $r = (N-m-1)/2 = 0(1)16; 18(2)22; 25$, and for $\alpha = .05; .01(T_3)$, and $\alpha = .95; .99(T_4)$.

In summary, we have extended the test for sphericity by extending to a class of elliptically symmetric distributions. Furthermore, under the assumption that $\sigma^2$ is unspecified in H: $\underline{\Lambda} = \sigma^2\underline{I}_m$, the joint null and non-null distributions of the maximal invariant statistics are unique. In fact <u>all</u> invariant tests are exactly robust. If $\sigma^2$ is completely specified then the null and non-null distributions of the maximal invariant statistics do depend on the underlying elliptically symmetric distribution.

# VIII. TESTS FOR BLOCK DIAGONAL STRUCTURE

## Section 8.1 Introduction

Let $L(\underline{X}) \in S_{nm}[\underline{M}, \underline{I}_n \otimes \underline{\Sigma}]$ with $\underline{M} \in F^1_{n\times m}$ and $\underline{\Sigma} \in S^+_m$. We investigate properties of invariant tests for two cases.

We first test the hypothesis H: $\underline{\Sigma} = \mathrm{Diag}(\sigma_{11},\ldots,\sigma_{mm})$. If second moments exist this is a test for the mutual uncorrelatedness of the m variables $x_1,\ldots,x_m$. This becomes a test for the mutual independence of $x_1,\ldots,x_m$ if $L(\underline{X})$ is Gaussian.

To define the second testing situation partition $\underline{\Sigma}$ into

$$\begin{bmatrix} \underline{\Sigma}_{11} & \underline{\Sigma}_{12} \\ \underline{\Sigma}_{21} & \underline{\Sigma}_{22} \end{bmatrix}$$

with $\underline{\Sigma}_{11}(p\times p)$, $\underline{\Sigma}_{22}(q\times q)$, and $p + q = m$. We then test H: $\underline{\Sigma}_{12} = 0$. For this case we prove that the power functions associated with certain invariant tests are monotonically increasing functions of each non-centrality parameter for the subclass of elliptically symmetric unimodal distributions.

## Section 8.2 Testing $\underline{\Sigma} = \mathrm{Diag}(\sigma_{11},\ldots,\sigma_{mm})$

Given that $L(\underline{X}) \in S_{nm}[\underline{M}, \underline{I}_n \otimes \underline{\Sigma}]$ we wish to test

$$H: \underline{\Sigma} = \mathrm{Diag}(\sigma_{11},\ldots,\sigma_{mm})$$

(8.2.1) versus

$$K: \underline{\Sigma} \neq \mathrm{Diag}(\sigma_{11},\ldots,\sigma_{mm})$$

with $\underline{M} \in \mathcal{F}^1_{n\times m}$ and $\underline{\Sigma} \in \mathcal{S}^+_m$.

The unrestricted and restricted parameter spaces for the problem are

$$\Omega\colon\ \mathcal{F}^1_{n\times m} \times \mathcal{S}^+_m$$

and

$$\Omega_H\colon\ \mathcal{F}^1_{n\times m} \times \mathcal{D}^+(m)$$

where $\underline{D} \in \mathcal{D}^+(m)$ implies $D = \mathrm{Diag}(d_{11},\ldots,d\ \ )$ with $d_{ii} > 0$. The problem remains invariant under the group G such that, for $g \in G$,

(8.2.2) $$g\underline{X} = \underline{P}(\underline{X}+\underline{B})\underline{D}$$

with $\underline{P} \in \mathcal{O}(n)$, $\underline{B} \in \mathcal{F}^1_{n\times m}$, and $\underline{D} \in \mathcal{D}^+(m)$.

<u>Theorem 8.2.1</u>: Write the sample matrix $\underline{S} = [s_{ij}]$, where

$$\underline{S} = \sum_{j=1}^{n} (\underline{x}_j-\underline{\bar{x}})(\underline{x}_j-\underline{\bar{x}})'$$

$$= (\underline{X}-\underline{\bar{X}})'(\underline{X}-\underline{\bar{X}}) = \underline{Z}'\underline{Z}$$

with $\underline{\bar{x}} = \frac{1}{n}\sum_{j=1}^{n} \underline{x}_j$. Then the sample and parametric maximal invariants are $\underline{R}$ and $\underline{\mathcal{R}}$ with

(8.2.3) $$\underline{R} = [r_{ij}]$$

and

$$\underline{\mathcal{R}} = [\rho_{ij}]$$

where $r_{ij} = s_{ii}^{-1/2}\ s_{ij}\ s_{jj}^{-1/2}$ and $\rho_{ij} = \sigma_{ii}^{-1/2}\ \sigma_{ij}\ \sigma_{jj}^{-1/2}$.

Proof: Let $G_1$ and $G_2$ be two groups such that $g_1\underline{X} = \underline{P}(\underline{X}+\underline{B})$ and $g_2\underline{X} = \underline{X}\ \underline{D}$ with $\underline{P}$, $\underline{B}$, and $\underline{D}$ defined in (8.2.2). The action of $G_1$ gives $\underline{S}$ and $\underline{\Sigma}$ as maximal invariants. We need to prove that (8.2.3) is maximal invariant within the space of the matrix $\underline{S}$. The group $G_2$ takes $\underline{S}$ into

$$\underline{D}'\underline{S}\ \underline{D} = \begin{bmatrix} d_{11}s_{11}d_{11} & \cdots & d_{11}s_{1m}d_{mm} \\ \vdots & & \vdots \\ d_{mm}s_{m1}d_{11} & \cdots & d_{mm}s_{mm}d_{mm} \end{bmatrix}$$

Now, $\underline{R} = [r_{ij}] = [s_{ii}^{-1/2}\ s_{ij}\ s_{jj}^{-1/2}]$ and $\underline{R}$ is clearly invariant. To show maximal invariance let $\underline{D}_1'\underline{S}\ \underline{D}_1 = \underline{D}_2'\underline{T}\ \underline{D}_2$. Then $\underline{T} = \underline{D}_2'^{-1}\underline{D}_1\underline{S}\ \underline{D}_1\underline{D}_2^{-1} = \underline{D}'\underline{S}\ \underline{D}$ for $\underline{D} \in \mathcal{D}^+(m)$ and (8.2.3) is maximal invariant. A similar proof holds for the parameter space.

Definition 8.2.1: If second moments are defined for $L(\underline{X})$, then $\underline{R}$ is the sample correlation matrix and $\underline{\mathcal{R}}$ is the population correlation matrix.

The next theorem characterizes the joint null distribution of the components of $\underline{R}$.

Theorem 8.2.2: The joint null distribution of $T = \{r_{ij} = s_{ii}^{-1/2}\ s_{ij}\ s_{jj}^{-1/2}$, $i,j=1,\ldots,m,\ i\neq j\}$ is unique for all elliptically symmetric $L(\underline{X})$.

Proof: If H is true then $L(\underline{X}) \in S_{nm}[\underline{M},\underline{I}_n \otimes \mathrm{Diag}(\sigma_{11},\ldots,\sigma_{mm})]$ and

$L(\underline{Z}) \in S_{nm}[\underline{\phi}, (\underline{I}_n - \frac{1}{n}\,\underline{1}\,\underline{1}') \otimes \mathrm{Diag}(\sigma_{11},\ldots,\sigma_{mm})]$. Premultiplication by an appropriate $\underline{P} \in \mathcal{O}(n)$ and post-multiplication by an appropriate $\underline{D} \in \mathcal{D}^+(m)$ gives

$$(8.2.5) \qquad L(\underline{Z}) \in S_{nm}[\underline{\phi}, \begin{bmatrix} \underline{I}_{n-1} & \underline{\phi} \\ \underline{\phi} & \underline{0} \end{bmatrix} \otimes \underline{I}_m].$$

This is a singular distribution with unit probability concentrated on a (n-1)m-dimensional subspace of $\mathcal{R}^{nm}$. So we assume

$$L(\underline{Z}) \in S_{Nm}[\underline{\phi}, \underline{I}_N \otimes \underline{I}_m]$$

for N = n - 1.

Next, let $\underline{Z} \to \underline{z} \to (r,\underline{\theta}) \to (r,\underline{\Theta})$. Then $\underline{S} \to \underline{Z}'\underline{Z} \to r^2\underline{\Theta}'\underline{\Theta}$. Write $\underline{\Theta}'\underline{\Theta}$ in the same manner as $\underline{S}$, i.e.

$$\underline{\Theta}'\underline{\Theta} = \begin{bmatrix} \underline{\theta}_1'\underline{\theta}_1 & \cdots & \underline{\theta}_1'\underline{\theta}_m \\ \vdots & & \vdots \\ \underline{\theta}_m'\underline{\theta}_1 & & \underline{\theta}_m'\underline{\theta}_m \end{bmatrix} .$$

Under this polar transformation,

$$s_{ii}^{-\frac{1}{2}}\, s_{ij}\, s_{jj}^{-\frac{1}{2}} \to$$

$$(r^2\underline{\theta}_i'\underline{\theta}_i)^{-\frac{1}{2}}\,(r^2\underline{\theta}_i'\underline{\theta}_j)(r^2\underline{\theta}_j'\underline{\theta}_j)^{-\frac{1}{2}}$$

$$= (\underline{\theta}_i'\underline{\theta}_i)^{-\frac{1}{2}}\,(\underline{\theta}_i'\underline{\theta}_j)(\underline{\theta}_j'\underline{\theta}_j)^{-\frac{1}{2}} .$$

Therefore, the joint distribution of the $r_{ij}$ is not a function of $r^2 = \text{tr}\ \underline{Z}'\underline{Z}$ and is unique for all $L(\underline{Z})$ and also for all $L(X) \in S_{nm}[\underline{M}, \underline{I}_n \otimes \text{Diag}(\sigma_{11}, \ldots, \sigma_{mm})]$.

The non-null distribution is not related to the null distribution by a simple change of scale. Therefore, the technique used thus far to show uniqueness of the non-null distribution is no longer applicable. We accordingly leave open the question as to whether or not the non-null distribution depends on the underlying elliptically symmetric law. Since power properties for an underlying Gaussian distribution are still unsolved, we make no statement about non-null distributions and power properties for this larger class of distributions.

The modified likelihood ratio test under normal theory gives an example of an invariant test statistic with a unique null distribution. This test statistic is given by

(8.2.6) $$T_1 = \frac{|\underline{S}|^{N/2}}{\prod_{j=1}^{m} s_{jj}^{N/2}} \quad .$$

Anderson (1958) gives moments of $T_1^{2/N}$ and derives an explicit form of its pdf.

## Section 8.3 Tests for H: $\underline{\Sigma}_{12} = \underline{0}$

Assume $L(\underline{X}) \in S_{nm}[\underline{M}, \underline{I}_n \otimes \underline{\Sigma}]$ with

$$\underline{X}(n\times m) = [\underline{X}_1(n\times p), \underline{X}_2(n\times q)]$$

$$\underline{M} \in \mathcal{F}^1_{n\times m}$$

(8.3.1)
$$\underline{\Sigma}(m\times m) = \begin{bmatrix} \underline{\Sigma}_{11} & \underline{\Sigma}_{12} \\ \underline{\Sigma}_{21} & \underline{\Sigma}_{22} \end{bmatrix} \in \mathcal{S}^+_m$$

$$\underline{\Sigma}_{11}(p\times p),\ \underline{\Sigma}_{22}(q\times q),\ p + q = m,\ p \le q.$$

We wish to test

$$H: \underline{\Sigma}_{12} = 0$$

versus

$$K: \underline{\Sigma}_{12} \neq 0\ .$$

Let $\mathcal{D}^+(m(2))$ denote the group of block diagonal positive definite matrices such that, if $\underline{A} \in \mathcal{D}^+(m(2))$, then $\underline{A}(m\times m) = \mathrm{Diag}(\underline{A}_{11},\underline{A}_{22})$ with $\underline{A}_{11}(p\times p)$ and $\underline{A}_{22}(q\times q)$ both symmetric positive definite. Let $\mathcal{D}(m(2))$ denote the group of block diagonal non-singular matrices. Then the unrestricted and restricted parameter spaces for the problem are

$$\Omega:\ \mathcal{F}^1_{n\times m} \times \mathcal{S}^+_m$$

and

$$\Omega_H:\ \mathcal{F}^1_{n\times m} \times \mathcal{D}^+(m(2))\ .$$

The problem also remains invariant under the group G such that, for $g \in G$, we have

(8.3.2)
$$g\underline{X} = \underline{P}(\underline{X}+\underline{B})\underline{A}$$

where $\underline{P} \in \mathcal{O}(n)$, $\underline{B} \in F^1_{n\times m}$, and $\underline{A} \in \mathcal{D}(m(2))$.

The next theorem gives the sample and parametric maximal invariants under the action of the group G. The proof of this theorem is given in Eaton (1972, pp. 10.21-10.22).

Theorem 8.3.1: Maximal invariants for the sample and parameter spaces are $\{v_1,\ldots,v_p\}$ and $\{\rho_1,\ldots,\rho_p\}$ where $\{v_1,\ldots,v_p\}$ contains the square roots of the eigenvalues of $|\underline{S}_{11}^{-1}\ \underline{S}_{12}\ \underline{S}_{22}^{-1}\ \underline{S}_{21} - v^2\underline{I}_p| = 0$ and $\{\rho_1,\ldots,\rho_p\}$ contains the square roots of the eigenvalues of $|\underline{\Sigma}_{11}^{-1}\ \underline{\Sigma}_{12}\ \underline{\Sigma}_{22}^{-1}\ \underline{\Sigma}_{21} - \rho^2\underline{I}_p| = 0$.

Definition 8.3.1: The sample canonical statistics are $\{v_1 \geq \ldots \geq v_p \geq 0\}$ where the $v_i$'s are the singular values of $\underline{S}_{11}^{-1/2}\ \underline{S}_{12}\ \underline{S}_{22}^{-1/2}$. The population canonical parameters are $\{\rho_1 \geq \ldots \geq \rho_p \geq 0\}$ where the $\rho_i$'s are the singular values of $\underline{\Sigma}_{11}^{-1/2}\ \underline{\Sigma}_{12}\ \underline{\Sigma}_{22}^{-1/2}$.

We next characterize the joint null distribution of $\{v_1,\ldots,v_p\}$.

Theorem 8.3.2: Let $L(\underline{Z}) \in S_{Nm}[\phi, \underline{I}_N \otimes \underline{I}_m]$ for $N = n - 1$, and partition $\underline{S} = \underline{Z}'\underline{Z}$ as

$$\underline{S} = \begin{bmatrix} \underline{S}_{11} & \underline{S}_{12} \\ \underline{S}_{21} & \underline{S}_{22} \end{bmatrix}.$$

Then the joint null distribution of $T = \{v_1,\ldots,v_p\}$ is unique for all $L(\underline{Z})$.

Proof: Let $\underline{Z} \to \underline{z} \to (r,\underline{\theta}) \to (r,\underline{\Theta})$. Then the matrices $\underline{S}_{11}$, $\underline{S}_{12}$, and $\underline{S}_{22}$ can be expressed in polar form as

$$\underline{S}_{11} = r^2\underline{\Theta}_1'\underline{\Theta}_1$$

$$\underline{S}_{12} = r^2\underline{\Theta}_1'\underline{\Theta}_2$$

$$\underline{S}_{22} = r^2\underline{\Theta}_2'\underline{\Theta}_2 \ .$$

It follows that

$$\underline{S}_{11}^{-1/2}\ \underline{S}_{12}\ \underline{S}_{22}^{-1/2} \to (\underline{\Theta}_1'\underline{\Theta}_1)^{-1/2}\ (\underline{\Theta}_1'\underline{\Theta}_2)(\underline{\Theta}_2'\underline{\Theta}_2)^{-1/2}$$

and thus T does not depend on $r^2 = \text{tr}\ \underline{Z}'\underline{Z}$. The distribution of T therefore is unique for all $L(\underline{Z})$.

Corollary 8.3.1: The following invariant test statistics have unique null distributions.

(8.3.3) $$T_2 = \frac{|\underline{S}|}{|\underline{S}_{11}|\,|\underline{S}_{22}|}$$

(8.3.4) $$T_3 = \text{ch}_1\ [\underline{S}_{11}^{-1}\ \underline{S}_{12}\ \underline{S}_{22}^{-1}\ \underline{S}_{21}]$$

(8.3.5) $$T_4 = \text{tr}\ [\underline{S}_{11}^{-1}\ \underline{S}_{12}\ \underline{S}_{22}^{-1}\ \underline{S}_{21}]$$

As the joint pdf of $\{v_1,\ldots,v_p\}$ is unique for all $L(\underline{Z})$, the results of Krishnaiah and Waikar (1971) extend to the larger class. They give the joint pdf of $\{v_1,\ldots,v_p\}$ and then derive the following pdf's: (i) the joint pdf of $v_i^2/v_{i+1}^2$, $i=1,\ldots,p-1$, (ii) the joint pdf of

$v_i^2/v_p^2$, i=1,...,p-1, and (iii) the joint pdf of $v_i^2/\sum_{j=1}^{p} v_j^2$, i=1,...,p-1.

Section 8.4 Power Properties for H: $\underline{\Sigma}_{12} = 0$

Let $T(\underline{X})$ be an invariant test statistic depending on the maximal invariant statistics $\{v_1,\ldots,v_p\}$. Then its non-null distribution depends on the maximal invariant parameters $\{\rho_1,\ldots,\rho_p\}$. We show that the power is a monotonically increasing function of each $\rho_i$ for the subclass of elliptically symmetric unimodal distributions.

There are several different definitions of unimodality for multivariate distributions. We use the one given by Anderson (1955). For extensions of Anderson's definition, definitions of other types of unimodality, and comparisons of the different types see Sherman (1955), Mudholkar (1966), Olshen and Savage (1970), Fefferman, Jodeit, and Perlman (1972), Ghosh (1974), Wolfe (1975), Das Gupta (1976), Dharmadhikari and Jogdeo (1976), and Wells (1978). We now give Anderson's definition of unimodality.

Definition 8.4.1: Let $\underline{x}(n\times 1)$ be a random vector with pdf $f(\underline{x})$. Then $f(\underline{x})$ is said to have a unimodal distribution if $\{\underline{x}|f(\underline{x})\geq u\}$ is convex for all u $(0<u<\infty)$.

We use this definition together with the results of Anderson and Das Gupta (1964a) for the Gaussian distribution to prove that the power is a monotonically increasing function of each $\rho_i$. Fundamental to the proof is the following theorem by Anderson (1955).

<u>Theorem 8.4.1</u>: Let E be a convex set in $R^n$, symmetric about the origin. For $\underline{x}(n\times 1)$, let $f(\underline{x}) \geq 0$ be a function such that

(i) $f(\underline{x}) = f(-\underline{x})$

(ii) $\{\underline{x} | f(\underline{x}) \geq u\} = K_u$ is convex for all u $(0<u<\infty)$.

(iii) $\int_E f(\underline{x})d\underline{x} < \infty$ .

Then

$$\int_E f(\underline{x}+k\underline{y})\,d\underline{x} \geq \int_E f(\underline{x}+\underline{y})\,d\underline{x}$$

for each fixed $\underline{y}$ and for $0 \leq k \leq 1$.

Anderson and Das Gupta's (1964a) proof for the monotonicity of the power function uses several conditional distributions which, in the Gaussian case, are just the marginal distributions. The marginal distribution is a Gaussian distribution and also unimodal. We apply the method of Anderson and Das Gupta (1964a) to the subclass of unimodal elliptical distributions. Because the conditional distribution is no longer the marginal distribution and no longer clearly unimodal we must prove that the conditional distribution is unimodal. We first give a result for all marginal distributions.

<u>Theorem 8.4.2</u>: For $\underline{x}(n\times 1)$, let $f(\underline{x}) \geq 0$ be unimodal and symmetric about zero. Partition $\underline{x} \in R^n$ as $[\underline{x}_1', \underline{x}_2']'$ with $\underline{x}_1 \in R^p$, $\underline{x}_2 \in R^q$, and $p + q = n$. Then

$$f(\underline{x}_1) = \int_{R_q} f(\underline{x}_1,\underline{x}_2)d\underline{x}_2$$

is a symmetric unimodal function in $R^p$.

Proof: See Eaton and Perlman (1975).

In the next theorem we prove that conditional distributions are unimodal when a pdf $f(\underline{x})$ is unimodal and symmetric about zero.

Theorem 8.4.3: For $\underline{x}(n\times 1)$, let the pdf $f(\underline{x})$ be unimodal and symmetric about zero. Partition $\underline{x} \in R^n$ as $[\underline{x}_1',\underline{x}_2']'$ with $\underline{x}_1 \in R^p$, $\underline{x}_2 \in R^q$, and $p + q = n$. Then

$$f(\underline{x}_1|\underline{x}_2) = f(\underline{x}_1,\underline{x}_2)/f(\underline{x}_2)$$

is a symmetric unimodal pdf in $R^p$ for each fixed $\underline{x}_2$.

Proof: By assumption

$$\{(\underline{x}_1,\underline{x}_2)|f(\underline{x}_1,\underline{x}_2) \geq u\} \tag{8.4.1}$$

is convex for all $0 < u < \infty$. Fix $\underline{x}_2 = \underline{x}_2^*$. Then the convex set in (8.4.1) is intersected by a hyperplane, the intersection being a convex set in $R^p$. Therefore,

$$\{(\underline{x}_1,\underline{x}_2^*)|f(\underline{x}_1,\underline{x}_2^*) \geq u\} \tag{8.4.2}$$

is convex for all $0 < u < \infty$ and for each $\underline{x}_2^*$.

For the conditional distribution to be unimodal

(8.4.3) $$\{(\underline{x}_1,\underline{x}_2^*)\,|\,f(\underline{x}_1|\underline{x}_2^*) \geq v\}$$

must be convex for all $0 < v < \infty$ and for each $\underline{x}_2^*$. Rewrite (8.4.3) as

$$\{(\underline{x}_1,\underline{x}_2^*)\,|\,f(\underline{x}_1,\underline{x}_2^*) \geq vf(\underline{x}_2^*)\}$$

$$= \{(\underline{x}_1,\underline{x}^*{}_2)\,|\,f(\underline{x}_1,\underline{x}_2^*) \geq u\}$$

for $u = vf(\underline{x}_2^*)$. This expression is convex by (8.4.2). Therefore, the conditional distribution is unimodal.

We have now presented all the preliminary definitions and theorems needed to prove that the power functions associated with invariant tests are monotonically increasing functions of each $\rho_i$.

Let $f(\underline{Y}) = |\underline{\Lambda}|^{-n/2}\,\psi[\mathrm{tr}\ \underline{Y}'\underline{Y}\ \underline{\Lambda}^{-1}]$ be unimodal, where

$$\underline{\Lambda} = \begin{bmatrix} \underline{I}_p & [\underline{D}_\rho \underline{\phi}] \\ [\underline{D}_\rho \underline{\phi}]' & \underline{I}_q \end{bmatrix} .$$

We need only consider the pdf in this form because the problem is invariant under the group G defined in (8.3.2). Here $\underline{Y} = \underline{P}(\underline{X}+\underline{B})\underline{A}$ and $[\underline{D}_\rho\underline{\phi}]$ is a $(p\times q)$ matrix with

$$\underline{D}_\rho(p\times p) = \mathrm{Diag}(\rho_1,\ldots,\rho_p)$$

and $\underline{\phi}$ $(p\times(q-p))$ a zero matrix.

Next partition $\underline{Y}(n\times m)$ into $[\underline{U}(n\times p),\ \underline{V}(n\times q)]$ for $p + q = m$ and $p \leq q$. Using standard multivariate techniques (cf. Anderson (1958), p. 23) we have that

$$(8.4.4) \qquad f(\underline{U},\underline{V}) = |\underline{\Lambda}_{11\cdot 2}|^{-n/2}$$

$$\cdot\ \psi\{\mathrm{tr}\ [(\underline{U}-\underline{V}\underline{\Lambda}_{22}^{-1}\underline{\Lambda}_{21})'(\underline{U}-\underline{V}\underline{\Lambda}_{22}^{-1}\underline{\Lambda}_{21})\underline{\Lambda}_{11\cdot 2}^{-1}$$

$$+ \underline{V}'\underline{V}\ \underline{\Lambda}_{22}^{-1}]\}$$

with

$$\underline{\Lambda}_{11\cdot 2} = \underline{\Lambda}_{11}-\underline{\Lambda}_{12}\underline{\Lambda}_{22}^{-1}\underline{\Lambda}_{21}$$

$$\underline{\Lambda}_{11} = \underline{I}_p$$

$$\underline{\Lambda}_{22} = \underline{I}_q$$

$$\underline{\Lambda}_{12} = [\underline{D}_\rho \underline{\phi}] \quad .$$

After substituting into (8.4.4) we have

$$(8.4.5) \qquad f(\underline{U},\underline{V}) = \prod_{i=1}^{p} (1-\rho_i^2)^{-n/2}$$

$$\cdot\ \psi\ \{\mathrm{tr}\ [\underline{U}-\underline{V}[\underline{D}_\rho \underline{\phi}]')'(\underline{U}-\underline{V}[\underline{D}_\rho \underline{\phi}]')\underline{D}_{1-\rho^2}^{-1}$$

$$+ \underline{V}'\underline{V}]\}$$

where $D_{1-\rho^2}^{-1} = [\mathrm{Diag}(1-\rho_1^2,\ldots,1-\rho_p^2)]^{-1}$.

We next condition on $\underline{V} = \underline{V}^*$. We know from Theorem 8.4.3 that the conditional distribution is unimodal. The conditional distribution is given by

$$(8.4.6) \qquad f(\underline{U}|\underline{V}=\underline{V}^*) = \frac{f(\underline{U},\underline{V}^*)}{f(\underline{V}^*)}$$

where $f(\underline{V}) = \int f(\underline{U},\underline{V})d\underline{U} = \psi_1(\text{tr } \underline{V}'\underline{V})$ for some function $\psi_1$.

We temporarily set $f(\underline{V}^*)$ aside because it is a fixed positive quantity and consider only $f(\underline{U},\underline{V}^*)$. Eventually we will want to consider the unconditional distribution, at which time we bring $f(\underline{V})$ back into the discussion. Now,

$$(8.4.7) \qquad f(\underline{U},\underline{V}^*) = \prod_{i=1}^{p} (1-\rho_i^2)^{-n/2}$$

$$\cdot\ \psi\{\text{tr } [\underline{U}-\underline{V}^*[\underline{D}_\rho\underline{\phi}]')'(\underline{U}-\underline{V}^*[\underline{D}_\rho\underline{\phi}]')D^{-1}_{1-\rho^2}$$

$$+\ \underline{V}^*{}'\underline{V}^*]\}\ .$$

Define

$$\underline{S}_h = (\underline{U}'\underline{V}^*)(\underline{V}^*{}'\underline{V}^*)^{-1}(\underline{V}^*{}'\underline{U})$$

$$\underline{S}_e = \underline{U}'\underline{U} - (\underline{U}'\underline{V}^*)(\underline{V}^*{}'\underline{V}^*)^{-1}(\underline{V}^*{}'\underline{U}).$$

If $v_i^2$ is the $i^{th}$ largest root of $(\underline{U}'\underline{U})^{-1}(\underline{U}'\underline{V}^*)(\underline{V}^*{}'\underline{V}^*)^{-1}(\underline{V}^*{}'\underline{U})$ then $v_i^2/(1-v_i^2)$ is the $i^{th}$ largest root of $\underline{S}_h\ \underline{S}_e^{-1}$. Since all invariant tests depend only on the $v_i^2$'s, the class of test procedures based on the roots of $(\underline{U}'\underline{U})^{-1}(\underline{U}'\underline{V}^*)(\underline{V}^*{}'\underline{V}^*)^{-1}(\underline{V}^*{}'\underline{U})$ is the same as the class of test procedures based on the roots of $\underline{S}_h\ \underline{S}_e^{-1}$.

We next define

$$\underline{W}(q\times p) = \underline{F}\ \underline{U}\ \underline{D}_{1-\rho^2}^{-1/2}$$

(8.4.8)

$$\underline{Z}((n-q)\times p) = \underline{G}\ \underline{U}\ \underline{D}_{1-\rho^2}^{-1/2} .$$

The matrices $\underline{F}(q\times n)$ and $\underline{G}((n-q)\times n)$ satisfy

$$\underline{F}'\underline{F} = \underline{V}^*(\underline{V}^{*\prime}\underline{V}^*)^{-1}\underline{V}^{*\prime}$$

$$\underline{G}'\underline{G} = \underline{I}_n - \underline{V}^*(\underline{V}^{*\prime}\underline{V}^*)^{-1}\underline{V}^{*\prime}$$

and the roots of $\underline{S}_h\ \underline{S}_e^{-1}$ are identical to the roots of $(\underline{W}'\underline{W})(\underline{Z}'\underline{Z})^{-1}$.

The Jacobian of the transformation is $|\underline{D}_{1-\rho^2}|^{n/2} = \prod_{i=1}^{p} (1-\rho_i^2)^{n/2}$. The matrices $\underline{F}$ and $\underline{G}$ can be found [cf. Roy (1957), pp. 84-86] so that

(8.4.9)
$$f(\underline{W},\underline{Z},\underline{V}^*) = [\mathrm{tr}\ (\underline{Z}'\underline{Z}+\underline{V}^{*\prime}\underline{V}^*)$$
$$+ \sum_{j=1}^{p} (\underline{w}_j-\tau_j\underline{e}_j)'(\underline{w}_j-\tau_j\underline{e}_j)$$
$$+ \sum_{j=p+1}^{q} \underline{w}_j'\underline{w}_j]$$

with $\underline{e}_j$ a vector having as its $i^{th}$ element

$$\delta_{ij} = \begin{cases} 1, & i=j \\ 0, & i\neq j \end{cases}$$

and where $\tau_1^2 \geq \dots \geq \tau_p^2$ are the characteristic roots of

$$[\underline{D}_\rho\underline{\phi}]\underline{V}^{*\prime}\underline{V}^*[\underline{D}_\rho\underline{\phi}]'\underline{D}_{1-\rho^2}^{-1} .$$

We now give the main theorem of this section.

Theorem 8.4.4: Let $f(\underline{W},\underline{Z} \mid \underline{V}=\underline{V}^*)$ be unimodal and consider an invariant test for which the acceptance region is convex and symmetric in each row vector of $\underline{W}$ for each set of fixed $\underline{Z}$ and the other rows of $\underline{W}$. Then this test has a power function which is monotonically increasing in each $\rho_i$.

Proof: Still working with $f(\underline{W},\underline{Z},\underline{V}^*)$ we condition on $\underline{Z}$ and the last $q - 1$ row vectors of $\underline{W}$. This conditional distribution is unimodal because $f(\underline{W},\underline{Z}|\underline{V}=\underline{V}^*)$ is unimodal. Therefore, we can apply Theorem 8.4.1 to this conditional distribution. Observe that

$$f(\underline{w}_1,\underline{V}^*|\underline{w}_2,\ldots,\underline{w}_q,\underline{Z}) = \frac{f(\underline{W},\underline{Z},\underline{V}^*)}{f(\underline{w}_2,\ldots,\underline{w}_q,\underline{Z})}$$

and consider $f(\underline{W},\underline{Z},\underline{V}^*)$ for $\underline{w}_2,\ldots,\underline{w}_q$ and $\underline{Z}$ fixed. Then from Theorem 8.4.1,

$$\int_E f(\underline{w}_1,\underline{w}_2^*,\ldots,\underline{w}_q^*,\underline{Z}^*,\underline{V}^*)d\underline{w}_1$$

$$(8.4.10) \qquad = \int_E \psi[\mathrm{tr}\ (\underline{Z}^{*\prime}\underline{Z}^*+\underline{V}^{*\prime}\underline{V}^*)$$

$$+ \sum_{j=2}^{p} (\underline{w}_j^*-\tau_j\underline{e}_j)'(\underline{w}_j^*-\tau_j\underline{e}_j)$$

$$+ \sum_{j=p+1}^{q} \underline{w}_j^{*\prime}\underline{w}_j^*$$

$$+ (\underline{w}_1-\tau_1\underline{e}_1)'(\underline{w}_1-\tau_1\underline{e}_1)]d\underline{w}_1$$

$$\le \int_E \psi[\mathrm{tr}\ (\underline{Z}^{*\prime}\underline{Z}^{*}+\underline{V}^{*\prime}\underline{V}^{*})$$

$$+ \sum_{j=2}^{p} (\underline{w}_j^{*}-\tau_j\underline{e}_j)'(\underline{w}_j^{*}-\tau_j\underline{e}_j)$$

$$+ \sum_{j=p+1}^{q} \underline{w}_j^{*\prime}\underline{w}_j^{*}$$

$$+ (\underline{w}_1-k\tau_1\underline{e}_1)'(\underline{w}_1-k\tau_1\underline{e}_1)]d\underline{w}_1$$

for $0 \le k \le 1$ and for E a set, in the space of $\underline{W}$ and $\underline{Z}$, which is convex and symmetric in $\underline{w}_1$ given the other $\underline{w}_j$'s and $\underline{Z}$.

The inequality in (8.4.10) is not changed if we multiply both sides by the marginal density $f(\underline{w}_2,\ldots,\underline{w}_q,\underline{Z})$. Therefore, for $\underline{V} = \underline{V}^{*}$ the conditional probability of the acceptance region monotonically decreases in each $\tau_j$.

The $\tau_1^2,\ldots,\tau_p^2$ are the roots of $\underline{V}^{*\prime}\underline{V}^{*}[\underline{D}_\rho\underline{\phi}]'\underline{D}^{-1}_{1-\rho^2}[\underline{D}_\rho\underline{\phi}] = \underline{T}\ \underline{T}'\underline{\Gamma}$. The $(q\times q)$ non-singular matrix $\underline{T}$ is defined such that $\underline{T}\ \underline{T}' = \underline{V}^{*\prime}\underline{V}^{*}$ and

$$\underline{\Gamma} = [\underline{D}_\rho\underline{\phi}]'\underline{D}^{-1}_{1-\rho^2}[\underline{D}_\rho\underline{\phi}]$$

$$= \begin{bmatrix} \underline{D}_\theta & \underline{\phi} \\ \underline{\phi} & \underline{\phi} \end{bmatrix}$$

with $\underline{D}_\theta = \mathrm{Diag}(\theta_1,\ldots,\theta_p)$, $\theta_i = \dfrac{\rho_i}{1-\rho_i^2}$ .

Anderson and Das Gupta (1964a) show that for $\theta_i^* \geq \theta_i$, i=1,...,p

$$ch_i(\underline{T}'\underline{\Gamma}^*\underline{T}) \geq ch_i(\underline{T}'\underline{\Gamma}\ \underline{T}),\ i=1,\ldots,p$$

where $ch_i(\underline{A})$ denotes the $i^{th}$ largest characteristic root of the matrix $\underline{A}$.

So, for any given $\underline{V} = \underline{V}^*$, the conditional probability of the acceptance region monotonically decreases in each $\theta_i$ and hence each $\rho_i$. Our final step is to multiply by the marginal distribution $f(\underline{V})$. This does not change the conditional probability inequality and the theorem is proved.

Results of Anderson and Das Gupta (1964a), for specific invariant tests, extend to the subclass of elliptically symmetric unimodal distributions.

Let $e_1 \geq \ldots \geq e_p$ be the roots of $(\underline{W}'\underline{W})(\underline{Z}'\underline{Z})^{-1}$. Then $e_i = \frac{v_i^2}{1-v_i^2}$. The relation $v_1^2 \leq c$ is equivalent to the relation $e_1 \leq \frac{c}{1-c} = c^*$. We then have the following corollary.

Corollary 8.4.1: The test statistic $T = ch_1\ [\underline{S}_{11}^{-1}\underline{S}_{12}\underline{S}_{22}^{-1}\underline{S}_{21}]$ has a power function which is monotonically increasing in each $\rho_i$.

Let $f_i = 1 + e_i$, i=1,...,p, and let $W_k$ be the sum of all different products of $f_1,\ldots,f_p$ taken k at a time, k=1,...,p. In particular

$$W_p = \prod_{i=1}^{p} f_i = [\prod_{i=1}^{p} (1-v_i^2)]^{-1}\ .$$

We now have the following theorem and corollary.

<u>Theorem 8.4.5</u>: A test having acceptance region $\sum_{k=1}^{p} a_k W_k \le c$, $a_k \ge 0$ has a power function which is monotonically increasing in each $\rho_i$.

<u>Proof</u>: See Anderson and Das Gupta (1964a).

<u>Corollary 8.4.2</u>: The test statistic $T = \frac{|\underline{S}|}{|\underline{S}_{11}||\underline{S}_{22}|}$ has a power function which is monotonically increasing in each $\rho_i$.

We have seen that the null distributions of certain test statistics are unique, while leaving open the possibility that non-null distributions may depend on the underlying elliptically symmetric law. The main result of this chapter is the extension of a normal-theory power property for H: $\underline{\Sigma}_{12} = \underline{0}$ when the underlying distribution is an elliptically symmetric unimodal distribution.

# IX. OTHER INVARIANT TESTS

## Section 9.1 Introduction

The purpose here is to apply the principle of invariance to two nonstandard multivariate testing problems not usually treated in textbooks even in the Gaussian case. The first problem considered is a test for the uncorrelatedness of two variables within a set of m variables. In the second problem we test that the m variables are equi-correlated.

In the first instance we show how the principle of invariance reduces the problem in terms of the simple correlation coefficient; in the second instance connections are made with likelihood ratio tests as given in the literature.

We continue to use the term "correlated" with the understanding that elliptical laws need not have moments of any order. In the latter case the testing problems are to be interpreted in terms of the structure of the matrix $\underline{\Sigma}$ of scale parameters.

## Section 9.2 Tests for the Uncorrelatedness of Two Variables Within a Set of m Variables

Given $L(\underline{X}) \in S_{nm}[\underline{M}, \underline{I}_n \otimes \underline{\Sigma}]$ with $\underline{M} \in F^1_{n\times m}$, $\underline{\Sigma} \in S^+_m$, and $m = 3$, we wish to test

$$(9.2.1) \qquad H: \underline{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & 0 & \sigma_{33} \end{bmatrix} = \Sigma_0$$

versus

$$K: \underline{\Sigma} \neq \underline{\Sigma}_0 \ .$$

We consider the problem for m = 3 but later note that these developments can be extended for an arbitrary m.

The unrestricted and restricted parameter spaces are

$$\Omega: \ F^1_{n\times m} \times S^+_m$$

and

$$\Omega_H: \ F^1_{n\times m} \times T^+_m$$

where $\underline{U} \in T^+_m$ implies $\underline{U} = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & 0 \\ u_{31} & 0 & u_{33} \end{bmatrix}$ and U is positive definite.

The testing problem is invariant under a group G such that, for each $g \in G$, we have

$$(9.2.2) \qquad g\underline{X} = \underline{P}(\underline{X}+\underline{B})\underline{A}$$

with $\underline{P} \in O(n)$, $\underline{B} \in F^1_{n\times m}$, and $\underline{A} \in A$, where $\underline{A} \in A$ implies that $\underline{A}$ has the form

$$\underline{A} = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & 0 & a_{33} \end{bmatrix} .$$

The next theorem gives the sample and parametric maximal invariants under the group G.

Theorem 9.2.1: The respective sample and parametric maximal invariants under G are $\underline{V}$ and $\underline{\Gamma}$ with

$$(9.2.3) \qquad \underline{V} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & r_{23} \\ 0 & r_{32} & 1 \end{bmatrix}$$

and

$$(9.2.4) \qquad \underline{\Gamma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \rho_{23} \\ 0 & \rho_{32} & 1 \end{bmatrix}$$

where $r_{23} = \dfrac{s_{23}}{s_{22}^{1/2}\, s_{33}^{1/2}}$, $\rho_{23} = \dfrac{\sigma_{23}}{\sigma_{22}^{1/2}\, \sigma_{33}^{1/2}}$ and

$$\underline{S} = \sum_{j=1}^{n} (\underline{x}_j - \bar{\underline{x}})(\underline{x}_j - \bar{\underline{x}})'$$

$$= \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{bmatrix} .$$

Proof: Let $G_1$ and $G_2$ be two groups such that $g_1\underline{X} = \underline{P}(\underline{X}+\underline{B})$ and $g_2\underline{X} = \underline{XA}$ with $\underline{P}$, $\underline{B}$, and $\underline{A}$ as defined in (9.2.2). The sample matrix $\underline{S}$ is maximal invariant under the action of $G_1$. We examine the action of $G_2$ on $\underline{S}$ to show that (9.2.3) is maximal invariant.

Under $G_2$ we have $\underline{S} \to \underline{A}'\underline{S}\,\underline{A}$. To show that $\underline{V}$ is invariant choose $a_{22} = s_{22}^{-1/2}$, $a_{33} = s_{33}^{-1/2}$, and $a_{11}$, $a_{21}$, $a_{31}$ such that the (1,1) element of $\underline{A}'\underline{S}\,\underline{A}$ is one and the (2,1) and (3,1) elements are zero. Each $a_{i1}$, $i=1,2,3$ will be a function of the $s_{ij}$, $i,j=1,2,3$. Next, suppose $\underline{A}_1'\underline{S}\,\underline{A}_1 = \underline{A}_2'\underline{T}\,\underline{A}_2$, $\underline{A}_1$, $\underline{A}_2 \in \mathcal{A}$. Then $T = \underline{A}_2'^{-1}\underline{A}_1'\underline{S}\,\underline{A}_1\underline{A}_2^{-1}$. But $\underline{A}_1\underline{A}_2^{-1} \in \mathcal{A}$ and $\underline{V}$ is maximal invariant.

We now have the following corollary to Theorem 8.4.2.

Corollary 9.2.1: The test statistic $T = r_{23}$ has a unique null distribution for all $L(\underline{X}) \in S_{nm}(\underline{M},\underline{I}_n \otimes \underline{\Sigma})$ with $\underline{M} \in \mathcal{F}^1_{n\times m}$ .

The problem can be extended to $\underline{\Sigma}(m\times m)$ and

$$H:\ \sigma_{m-1,m} = 0$$

(9.2.5) versus

$$K:\ \sigma_{m-1,m} \neq 0 \ .$$

The $\underline{A}$ matrix in (9.2.2) would have $a_{11},\ldots,a_{m1}$ in the first column, $a_{(m-1)(m-1)}$ and $a_{mm}$ in the (m-1,m-1) and (m,m) positions respectively, and zeros elsewhere.

We have another extension if we partition $\underline{\Sigma}(m\times m)$ as

$$\begin{bmatrix} \underline{\Sigma}_{11} & \underline{\Sigma}_{12} & \underline{\Sigma}_{13} \\ \underline{\Sigma}_{21} & \underline{\Sigma}_{22} & \underline{\Sigma}_{23} \\ \underline{\Sigma}_{31} & \underline{\Sigma}_{32} & \underline{\Sigma}_{33} \end{bmatrix}$$

with $\underline{\Sigma}_{11}(p\times p)$, $\underline{\Sigma}_{22}(q\times q)$, $\underline{\Sigma}_{33}(r\times r)$, and $m = p + q + r$. The hypothesis to be tested is

$$H: \underline{\Sigma}_{23} = 0$$

versus

$$K: \underline{\Sigma}_{23} \neq 0 \quad .$$

The $\underline{A}$ matrix in (9.2.2) becomes

$$\begin{bmatrix} \underline{A}_{11} & \phi & \phi \\ \underline{A}_{21} & \underline{A}_{22} & \phi \\ \underline{A}_{31} & \phi & \underline{A}_{33} \end{bmatrix} .$$

Let $\underline{A}_{22} = \underline{S}_{22}^{-1/2}$, $\underline{A}_{33} = \underline{S}_{33}^{-1/2}$ and choose $\underline{A}_{11}$, $\underline{A}_{21}$, $\underline{A}_{31}$ such that the (1,1) element of $\underline{A}'\underline{S}\,\underline{A}$ is $\underline{I}_p$ and the (2,1) and (3,1) elements are zero matrices. Then

$$\underline{V} = \begin{bmatrix} \underline{I}_p & \phi & \phi \\ \underline{\phi} & \underline{I}_q & [\underline{D}_d\underline{\phi}] \\ \underline{\phi} & [\underline{D}_d\underline{\phi}]' & \underline{I}_r \end{bmatrix}$$

is maximal invariant. The $(q\times r)$ matrix $[\underline{D}_d\underline{\phi}]$ has $\underline{D}_d = \mathrm{Diag}(d_1,\ldots,d_q)$ and $\underline{\phi}(q\times(r-q))$ with the $d_i$'s the positive square roots of the eigenvalues of $|\underline{S}_{22}^{-1}\underline{S}_{23}\underline{S}_{33}^{-1}\underline{S}_{32}-d^2\underline{I}_q| = 0$.

We have the following corollary to Theorem 8.4.2.

Corollary 9.2.2: The joint null distribution of $T = \{d_i,\ i=1,\ldots,q\}$ is unique for all $L(\underline{X}) \in S_{nm}(\underline{M},\underline{I}_n \otimes \underline{\Sigma})$ with $\underline{M} \in \mathcal{F}^1_{n\times m}$.

As in Section 8.4, examples of some invariant test statistics are

(a) $$T_1 = ch_1\ (\underline{S}_{22}^{-1}\underline{S}_{23}\underline{S}_{33}^{-1}\underline{S}_{32})$$

(b) $$T_2 = tr\ (\underline{S}_{22}^{-1}\underline{S}_{23}\underline{S}_{33}^{-1}\underline{S}_{32})$$

and

(c) $$T_3 = \frac{|\underline{S}|}{|\underline{S}_{22}|\,|\underline{S}_{33}|}$$

with

$$\underline{S} = \begin{bmatrix} \underline{S}_{22} & \underline{S}_{23} \\ \underline{S}_{32} & \underline{S}_{33} \end{bmatrix}.$$

The power properties demonstrated in Chapter VIII for the subclass of elliptically symmetric unimodal distributions can be applied to the test statistics in Corollaries 9.2.1 and 9.2.2. Therefore, in Corollary 9.2.1 the power is a monotonically increasing function of $\rho^2_{23}$ and in Corollary 9.2.2 the power is a monotonically increasing function of each $\rho_i$ when $L(\underline{X})$ is unimodal. The $\rho_i$ are the square roots of the eigenvalues of $|\underline{\Sigma}_{22}^{-1}\underline{\Sigma}_{23}\underline{\Sigma}_{33}^{-1}\underline{\Sigma}_{32} - \rho^2\underline{I}_q| = 0$.

Section 9.3 Tests for Equi-Correlatedness

Given that $L(\underline{X}) \in S_{nm}[\underline{M},\underline{I}_n \otimes \underline{\Sigma}]$ with $\underline{M} \in F^1_{n\times m}$ and $\underline{\Sigma} \in S^+_m$, we wish to test

$$\text{H: } \underline{\Sigma} = [(1-\rho)\underline{I}_m+\rho\underline{1}\ \underline{1}']$$

(9.3.1) versus

$$\text{K: } \underline{\Sigma} \neq [(1-\rho)\underline{I}_m+\rho\underline{1}\ \underline{1}'].$$

We put the problem in canonical form before undertaking a reduction through invariance. There is an orthogonal $\underline{Q}$ with first column $m^{-1/2}\ \underline{1}$ such that $\underline{Q}'\underline{\Sigma}\ \underline{Q} = \underline{\Lambda}$ where

$$\underline{\Lambda} = \text{Diag}(\sigma^2(1-\rho)+\sigma^2\rho m,\ \sigma^2(1-\rho),\ldots,\sigma^2(1-\rho))$$

$$= \text{Diag}(\lambda_1,\lambda_2,\ldots,\lambda_2)$$

with $\lambda_1 = \sigma^2[(1-\rho)+\rho m]$ and $\lambda_2 = \sigma^2(1-\rho)$.

We now assume that $L(\underline{Y}) \in S_{nm}[\underline{N},\underline{I}_n \otimes \underline{\Lambda}]$ with $\underline{Y} = \underline{XQ}$, $\underline{N} = \underline{MQ} \in F^1_{n\times m}$, and $\underline{\Lambda} = \underline{Q}'\underline{\Sigma Q} \in S^+_m$. The hypothesis we test is

$$\text{H: } \underline{\Lambda} = \text{Diag}(\lambda_1,\lambda_2,\ldots,\lambda_2)$$

(9.3.2) versus

$$\text{K: } \underline{\Lambda} \neq \text{Diag}(\lambda_1,\lambda_2,\ldots,\lambda_2)$$

with unrestricted and restricted parameter spaces

$$\Omega: \quad \mathcal{F}^1_{n\times m} \times \mathcal{S}^+_m$$

and

$$\Omega_H: \quad \mathcal{F}^1_{n\times m} \times \mathcal{D}(2)$$

where $\underline{D} \in \mathcal{D}(2)$ implies $\underline{D} = \text{Diag}(d_1, d_2, \ldots, d_2)$.

The problem remains invariant under the group G, such that for $g \in G$,

(9.3.3) $$g\underline{Y} = \underline{P}(\underline{Y}+\underline{B})\underline{D}$$

with $\underline{P} \in \mathcal{O}(n)$, $\underline{B} \in \mathcal{F}^1_{n\times m}$, and $\underline{D} \in \mathcal{D}(2)$. The actions of G on the sample space and of $\bar{G}$ on the parameter space give the following theorem.

Theorem 9.3.1: The respective sample and parametric maximal invariants are

(9.3.4)
$$v_{1j}\, v_{11}^{-1/2} \left(\sum_{i=2}^{m} v_{ii}\right)^{-1/2}, \quad j=2,\ldots,m$$
$$v_{ij} \left(\sum_{i=2}^{m} v_{ii}\right)^{-1}, \quad i,j=2,\ldots,m$$

and

(9.3.5)
$$\lambda_{1j}\, \lambda_{11}^{-1/2} \left(\sum_{i=2}^{m} \lambda_{ii}\right)^{-1/2}, \quad j=2,\ldots,m$$
$$\lambda_{ij} \left(\sum_{i=2}^{m} \lambda_{ii}\right)^{-1}, \quad i,j=2,\ldots,m$$

for $\underline{V} = [v_{ij}] = \sum_{j=1}^{n} (\underline{y}_j - \bar{\underline{y}})(\underline{y}_j - \bar{\underline{y}})'$ and $\underline{\Lambda} = [\lambda_{ij}]$.

Proof: Let $G_1$ and $G_2$ be two groups such that $g_1\underline{Y} = \underline{P}(\underline{Y}+\underline{B})$ and $g_2\underline{Y} = \underline{YD}$ with $\underline{P}$, $\underline{B}$, and $\underline{D}$ defined in (9.3.3). The action of $G_1$ gives $\underline{V}$ and $\underline{\Lambda}$ as maximal invariants. The action of $G_2$ on $\underline{V}$ gives

$$\underline{D}'\underline{VD} = \begin{bmatrix} d_1^2 v_{11} & d_1 d_2 v_{12} & \cdots & d_1 d_2 v_{1m} \\ d_1 d_2 v_{21} & d_2^2 v_{22} & \cdots & d_2^2 v_{2m} \\ \vdots & \vdots & & \vdots \\ d_1 d_2 v_{m1} & d_2^2 v_{m2} & & d_2^2 v_{mm} \end{bmatrix}$$

and (9.3.4) is clearly invariant. To show maximal invariance, let $\underline{D}_1\underline{V}_1\underline{D}_1 = \underline{D}_2\underline{V}_2\underline{D}_2$. Then $\underline{V}_2 = \underline{D}_2^{-1}\underline{D}_1\underline{V}_1\underline{D}_1\underline{D}_2^{-1}$. But $\underline{D}_1\underline{D}_2^{-1} \in \mathcal{D}(2)$. Therefore, $\underline{V}_2 = \underline{DV}_1\underline{D}$ and (9.3.4) is maximal invariant. The proof for the parameter space is similar.

Theorem 9.3.2: The joint null distribution of $T = \{v_{1j}\, v_{11}^{-1/2} (\sum_{i=2}^{m} v_{ii})^{-1/2}$, $j=2,\ldots,m;\ v_{ij}(\sum_{i=2}^{m} v_{ii})^{-1}, i,j=2,\ldots,m\}$ is unique for all $L(\underline{Y}) \in S_{nm}(\underline{M}, \underline{I}_n \otimes \underline{\Sigma})$.

Proof: If H is true, then $L(\underline{Z}) \in S_{Nm}[\underline{\phi}, \underline{I}_N \otimes \underline{I}_m]$ with $N = n - 1$ and $\underline{Z}'\underline{Z} = \sum_{j=1}^{n} (\underline{y}_j - \bar{\underline{y}})(\underline{y}_j - \bar{\underline{y}})' = \underline{V}$. Next, transform $\underline{Z}$ into $(r, \underline{\Theta})$. Then $\underline{V} \to r^2\underline{\Theta}'\underline{\Theta}$, where

$$\underline{\Theta}'\underline{\Theta} = \begin{bmatrix} \underline{\theta}_1'\underline{\theta}_1 & \underline{\theta}_1'\underline{\theta}_2 & \cdots & \underline{\theta}_1'\underline{\theta}_m \\ \vdots & \vdots & & \vdots \\ \underline{\theta}_m'\underline{\theta}_1 & \underline{\theta}_m'\underline{\theta}_2 & \cdots & \underline{\theta}_m'\underline{\theta}_m \end{bmatrix} .$$

Therefore,

$$v_{1j}\, v_{11}^{-1/2} \left(\sum_{i=2}^{m} v_{ii}\right)^{-1/2} \to$$

$$r^2(\underline{\theta}_1'\underline{\theta}_j)(r^2\underline{\theta}_1'\underline{\theta}_1)^{-1/2}\left(r^2 \sum_{i=2}^{m} \underline{\theta}_i'\underline{\theta}_i\right)^{-1/2}$$

and

$$v_{ij} \left(\sum_{i=2}^{m} v_{ii}\right)^{-1} \to$$

$$(r^2\underline{\theta}_i'\underline{\theta}_j)\left(r^2 \sum_{i=2}^{m} \underline{\theta}_i'\underline{\theta}_i\right)^{-1} \quad .$$

The null distribution of T is unique for all $L(\underline{Z})$ and hence for all $L(\underline{Y})$ since the maximal invariants do not depend on $r^2 = \mathrm{tr}\ \underline{Z}'\underline{Z}$.

Corollary 9.3.2: The invariant test statistic

$$T_1 = \frac{|\underline{V}|}{v_{11}\left(\frac{\sum_{i=2}^{m} v_{ii}}{m-1}\right)^{m-1}} , \tag{9.3.6}$$

which is the likelihood ratio statistic under normal theory [Gleser and Olkin (1969)], has a unique null distribution for all $L(\underline{Y})$.

Wilks (1946) expresses $T_1$ in terms of the original sample $\underline{X}$. In this sample space we have

$$T_1 = \frac{|\underline{S}|}{\left(\frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{m} s_{ij}\right)\left[\frac{1}{m-1}\left(\sum_{i=1}^{m} s_{ii} - \frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{m} s_{ij}\right)\right]^{m-1}} \quad . \tag{9.3.7}$$

We next show that (9.3.6) and (9.3.7) are equivalent. Let $\underline{S} = \sum_{j=1}^{n} (\underline{x}_j - \bar{\underline{x}})(\underline{x}_j - \bar{\underline{x}})'$. Then

$$\underline{V} = \underline{Q}'\underline{S}\underline{Q}$$

$$v_{11} = \frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{m} s_{ij}$$

$$\sum_{i=2}^{m} v_{ii} = \sum_{i=1}^{m} s_{ii} - \frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{m} s_{ij}$$

and the two expressions are equivalent.

They hypothesis given in (9.3.2) can be regarded as a hypothesis of no relationship among $y_1, \ldots, y_m$. From the discussion in Section 8.2 we know that power properties for the Gaussian distribution are still unsolved. We therefore make no statements here about power properties for this larger class.

Finally, we note that Kariya and Eaton (1977) and Kariya (1977) examine a testing problem similar to the one defined in (9.3.1). However, they are working with a (n×1) random vector and are testing for serial correlation among the n components.

# X. SUMMARY

We have characterized all invariant tests for various hypotheses about scale parameters when the underlying distribution belongs to the class of elliptically symmetric laws. The only member of this class for which the sample observations are independent is the Gaussian distribution. Accordingly, we essentially have weakened the usual assumptions to include not only non-Gaussian laws, but also dependent observations.

Under the assumption of an underlying elliptically symmetric law we have shown that all normal-theory tests considered here are exactly robust and that non-null distributions often are exact as well. For one case in which it is not known whether the non-null distribution is exact, we have proved a monotone power property in the subclass of elliptically symmetric unimodal distributions.

We also have shown that the principle of invariance may be applied effectively to other than the standard hypothesis testing problems. In the context of this study, the clear advantage of this principle is that one need not specify a particular underlying distribution belonging to the standard class in order to derive properties of test procedures which hold for all members of the class.

In practice it is understood that normal-theory procedures are approximate procedures for distributions attracted to Gaussian laws. We have shown that various normal-theory tests for scale parameters are exact in the class of a-1 elliptical laws. As this class contains the elliptical stable laws, one practical consequence of this study is to extend the use of normal-theory procedures, as large-sample approximate procedures, to all distributions in the domains of attraction of the elliptically symmetric stable laws.

Finally, we note that the results in this thesis extend beyond a study of robustness and approximation theory since they apply directly to fields in which problems often exhibit an inherent elliptical symmetry. Examples of problems exhibiting spherical symmetry can be found in models for communications systems [cf. Goldman (1975)], bombing problems [cf. Gilliland (1968)], stock market models [cf. Zellner (1976)], and in problems involving angular data [cf. Mardia (1972)].

## BIBLIOGRAPHY


Ahmad, R. (1972). Extension of the normal family to spherical families. *Trabajos Estadist.* 23 51-60.

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

Anderson, T.W. (1958). *An Introduction to Multivariate Statistical Analysis.* John Wiley and Sons, Inc., New York.

Anderson, T.W. and Das Gupta, S. (1964a). Monotonicity of the power function of some tests of independence between two sets of variates. *Ann. Math. Statist.* 35 206-208.

Anderson, T.W. and Das Gupta, S. (1964b). A monotonicity property of the power functions of some tests of the equality of two covariance matrices. *Ann. Math. Statist.* 35 1059-1063.

Armitage, J.V. and Krishnaiah, P.R. (1964). *Tables for the Studentized Largest Chi-Square Distribution and Their Applications.* ARL64-188, Aerospace Research Laboratories, Wright Patterson Air Force Base, Ohio.

Bartlett, M.S. (1934). The vector representation of a sample. *Proc. Camb. Phil. Soc.* 30 327-340.

Bennett, B.M. (1961). On a certain multivariate non-normal distribution. *Proc. Camb. Phil. Soc.* 57 434-436.

Berger, J. (1975). Minimax estimation of location vectors for a wide class of densities. *Ann. Statist.* 3 1318-1328.

Blake, I.F. and Thomas, J.B. (1968). On a class of processes arising in linear estimation theory. *IEEE Trans. Inform. Thy.* IT-14 12-16.

Box, G.E.P. (1952). Multi-factor designs of first order. *Biometrika* 39 49-57.

Box, G.E.P. (1953). Spherical Distributions. *Ann. Math. Statist.* 24 687-688 (Abstract).

Box, G.E.P. and Hunter, J.S. (1957). Multi-factor experimental designs for exploring response surfaces. *Ann. Math. Statist.* 28 195-241.

Brandwein, A.R.C. and Strawderman, W.E. (1978). Minimax estimation of location parameters for spherica-ly symmetric unimodal distributions under quadratic loss. *Ann. Statist.* 6 377-416.

Butzer, P.L. and Hahn, L. (1978). General theorems on rates of convergence in distribution of random variables. II: Applications to the stable limit laws and weak law of large numbers. *J. Multivariate Anal.* 8 202-221.

Carter, E.M. and Srivastava, M.S. (1977). Monotonicity of the power function of modified likelihood ratio criterion for the homogeneity of variances and of the sphericity test. *J. Mult. Anal.* 7 229-233.

Chao, M. and Glaser, R.E. (1978). The exact distribution of Bartlett's test statistic for homogeneity of variances with unequal sample sizes. *J. Amer. Statist. Assoc.* 73 422-426.

Chu, K. (1973). Estimation and decision for linear systems with elliptical random processes. *IEEE Trans. Automat. Contr.* AC-18 499-505.

Cochran, W.G. (1941). The distribution of the largest of a set of estimated variances as a fraction of their total. *Ann. Eugenics* 11 47-52.

Consul, P.C. (1967). On the exact distributions of likelihood ratio criteria for testing independence of sets of variates under null hypotheses. *Ann. Math. Statist.* 38 1160-1169.

Das Gupta, S. (1969). Properties of power functions of some tests concerning dispersion matrices of multivariate normal distributions. *Ann. Math. Statist.* 40 697-701.

Das Gupta, S. (1976). A generalization of Anderson's theorem on unimodal functions. *Proc. Amer. Math. Soc.* 60 85-91.

Das Gupta, S., Anderson, T.W., and Mudholkar, G.S. (1964). Monotonicity of the power functions of some tests of the multivariate linear hypotheses. *Ann. Math. Statist.* 35 200-205.

Das Gupta, S., Eaton, M.L., Olkin, I., Perlman, M., Savage, L.J. and Sobel, M. (1971). Inequalities on the probability content of convex regions for elliptically contoured distributions. *Sixth Berkeley Symposium on Math. Statist. and Prob.*, Vol. II, 241-265.

Das Gupta, S. and Giri, N. (1973). Properties of tests concerning covariance matrices of normal distributions. *Ann. Statist.* 1 1222-1224.

David, H.A. (1952). Upper 5 and 1 % points of the maximum F-ratio. *Biometrika* 39 422-424.

Dawid, A.P. (1977). Spherical matrix distributions and a multivariate model. *J. Roy. Statist. Soc. B* 39 254-261.

Devlin, S.J., Gnanadesikan, R., and Kettenring, J.R. (1976). Some multivariate applications of elliptical distributions. *Essays in Probability and Statistics*, Chapter 24. Shinko Tsusho Co., Ltd., Tokyo, Japan.

Dharmadhikari, S.W. and Jogdeo, K. (1976). Multivariate unimodality. *Ann. Statist.* 4 607-613.

Eaton, M.L. (1972). *Multivariate Statistical Analysis.* Institute of Mathematical Statistics, University of Copenhagen, Denmark.

Eaton, M.L. and Perlman, M.D. (1975). Monotonicity of certain multivariate normal probabilities. *Canadian J. Statist.* 3 165-173.

Efron, B. (1969). Student's t-test under symmetry conditions. *J. Amer. Statist. Assoc.* 64 1278-1302.

Fefferman, C., Jodeit, J.R., and Perlman, M.D. (1972). A spherical surface measure inequality for convex sets. *Proc. Amer. Math. Soc.* 33 114-119.

Fisher, R.A. (1925). Applications of Student's distribution. *Metron* 5(3) 90-104.

Fraser, D.A.S. and Haq, M.S. (1969). Structural probability and prediction for the multivariate model. *J. Roy. Statist. Soc. B* 31 217-331.

Fraser, D.A.S. and Haq, M.S. (1970). Inference and prediction for the multilinear model. *J. Statist. Res.* 4 93-109.

Ghosh, P. (1974). On generalized unimodality. *Comm. Statist.* 3 567-580.

Ghosh, M. and Pollack, E. (1975). Some properties of multivariate distributions with pdf's constant on ellipsoids. *Comm. Statist.* 4 1157-1160.

Gilliland, D.C. (1968). Maximization of the integral of a bell-shaped function over a symmetric set. *Naval Res. Logistics Quart.* 15 507-515.

Giri, N. (1968). On tests of the equality of two covariance matrices. *Ann. Math. Statist.* 39 275-277.

Glaser, R.E. (1976). Exact critical values for Bartlett's test for homogeneity of variances. *J. Amer. Statist. Assoc.* 71 488-490.

Gleser, L.J. (1966). A note on the sphericity test. *Ann. Math. Statist.* 37 464-467.

Gleser, L.J. and Olkin, I. (1969). Testing for equality of means, equality of variances, and equality of covariances under restrictions upon the parameter space. *Ann. Inst. Statist. Math.* 21 33-48.

Gnanadesikan, R. (1959). Equality of more than two variances and of more than two dispersion matrices against certain alternatives. *Ann. Math. Statist.* 30 177-184.

Goldman, J. (1976). Detection in the presence of spherically symmetric random vectors. *IEEE Trans. Inform. Thy.* IT-22 52-59.

Gualtierotti, A.F. (1974). Some remarks on spherically invariant distributions. *J. Mult. Anal.* 4 347-349.

Gualtierotti, A.F. (1976). A likelihood ratio formula for spherically invariant processes. *IEEE Trans. Inform. Thy.* 22 610.

Gupta, S.S. (1963). On a selection and ranking procedure for gamma populations. *Ann. Inst. Statist. Math.* 14 199-216.

Gupta, S.S. and Sobel, M. (1962). On selecting a subset containing the population with the smallest variance. *Biometrika* 49 495-508.

Hartley, H.O. (1950). The maximum F-ratio as a short-cut test for heterogeneity of variance. *Biometrika* 37 308-312.

Hartman, P. and Wintner, A. (1940). On the spherical approach to the normal distribution law. *Amer. J. Math.* 62 759-779.

Hartmann, N.A., Jr. (1969). *F-max Tables for Mean Squares with Unequal Degrees of Freedom.* Technical Report 15, Department of Statistics, Oregon State University, Corvallis, Oregon.

Higgins, J.J. (1975). A geometrical method of constructing multivariate densities and some related inferential procedures. *Comm. Statist.* 4 955-966.

Hsu, P.L. (1940). An algebraic derivation of the distribution of rectangular coordinates. *Proc. Edinburgh Math. Soc.* 6 Ser. 2 185-189.

Jensen, D.R. (1978). Class notes. Mult. Statist. Anal. II, VPI&SU, Blacksburg, Virginia.

John, S. (1971). Some optimal multivariate tests. *Biometrika* 58 123-127.

John, S. (1972). The distribution of a statistic used for testing sphericity of normal distributions. *Biometrika* 59 169-173.

John, S. (1976). Fitting sampling distribution agreeing in support and moments and tables of critical values of sphericity criterion. *J. Mult. Anal.* 6 601-607.

Kariya, T. (1977). A robustness property of the tests for serial correlation. *Ann. Statist.* 5 1212-1220.

Kariya, T. and Eaton, M.L. (1977). Robust tests for spherical symmetry. *Ann. Statist.* 5 206-215.

Keifer, J. and Schwartz, R. (1965). Admissible character of $T^2$, $R^2$, and other fully invariant tests for classical multivariate normal problems. *Ann. Math. Statist.* 36 747-770.

Kelker, D. (1970). Distribution theory of spherical distributions and a location-scale parameter generalization. *Sankhya A* 32 419-430.

Kingman, J.F.C. (1963). Random walks with spherical symmetry. *Acta Mathematica* 109 11-53.

Krishnaiah, P.R. (1965). Simultaneous tests for the equality of variances against certain alternatives. *Austral. J. Statist.* 7 105-109.

Krishnaiah, P.R. and Armitage, J.V. (1964). *Distribution of the Studentized Smallest Chi-Square, with Tables and Applications.* ARL 64-218, Aerospace Research Laboratories, Wright Patterson Air Force Base, Ohio.

Krishnaiah, P.R. and Chang, T.C. (1972). On the exact distributions of the traces of $S_1(S_1+S_2)^{-1}$ and $S_1S_2^{-1}$. *Sankhya* 34 153-160.

Krishnaiah, P.R. and Schuurmann, F.J. (1974). On the evaluation of some distributions that arise in simultaneous tests for the equality of the latent roots of the covariance matrix. *J. Mult. Anal.* 265-282.

Krishnaiah, P.R. and Waikar, V.B. (1971). Simultaneous tests for equality of latent roots against certain alternatives--I. *Ann. Inst. Statist. Math.* 23 451-468.

Laurent, A.G. (1974). The intersection of random spheres and the non-central radial error distribution for spherical models. *Ann. Statist.* 2 182-189.

Lehmann, E.L. (1959). *Testing Statistical Hypotheses.* John Wiley and Sons, New York.

Lord, R.D. (1954). The use of the Hankel transform in statistics. I. General theory and examples. *Biometrika* 41 44-55.

Mardia, K.V. (1972). *Statistics of Directional Data.* Academic Press, New York.

Maxwell, J.C. (1860). Illustration of the dynamical theory of gases --Part I. On the motions and collisions of perfectly elastic bodies. *Taylor's Philosophical Magazine* 19 19-32.

McGraw, D.K. and Wagner, J.F. (1968). Elliptically symmetric distributions. *IEEE Trans. Inform. Thy.* IT-14 110-120.

Mikhail, W.F. (1962). On a property of a test for the equality of two normal dispersion matrices against one-sided alternatives. *Ann. Math. Statist.* 33 1463-1465.

Mudholkar, G.S. (1966). The integral of an invariant unimodal function over an invariant convex set - An inequality and applications. *Proc. Amer. Math. Soc.* 17 1327-1333.

Nagarsenker, B.N. and Pillai, K.C.S. (1973). The distribution of the sphericity test criterion. *J. Mult. Anal.* 3 226-235.

Narain, R.D. (1950). On the completely unbiased character of tests of independence in multivariate normal systems. *Ann. Math. Statist.* 21 293-298.

Nash, D. and Klamkin, M.S. (1976). A spherical characterization of the normal distribution. *J. Mult. Anal.* 55 156-158.

Olkin, I. and Rubin, H. (1964). Multivariate Beta distributions and independence properties of the Wishart distribution. *Ann. Math. Statist.* 35 261-269.

Olshen, R.A. and Savage, L.J. (1970). A generalized unimodality. *J. Appl. Prob.* 7 21-34.

Picinbono, B. (1970). Spherically invariant and compound Gaussian stochastic processes. *IEEE Trans. Inform. Thy.* IT-16 77-79.

Pillai, K.C.S. and Dotson, C.O. (1969). Power comparisons of tests of two multivariate hypotheses based on individual characteristic roots. *Ann. Inst. Statist. Math.* 21 49-66.

Pillai, K.C.S. and Jayachandran, K. (1967). Power comparisons of tests of two multivariate hypotheses based on four criteria. *Biometrika* 54 195-210.

Pillai, K.C.S. and Jayachandran, K. (1968). Power comparisons of tests of equality of two covariance matrices based on four criteria. *Biometrika* 55 335-342.

Pillai, K.C.S. and Young, D.L. (1974). On the max U-ratio and likelihood ratio tests of equality of several covariance matrices. *Comm. Statist.* 3 29-53.

Ramachandran, K.V. (1956a). On the simultaneous analysis of variance test. *Ann. Math. Statist.* 27 521-528.

Ramachandran, K.V. (1956b). On the Tukey test for the equality of means and the Hartley test for equality of variances. *Ann. Math. Statist.* 27 825-831.

Ramachandran, K.V. (1958). On the studentized smallest chi-square. *J. Amer. Statist. Assoc.* 53 868-872.

Robbins, H. (1948). Mixture of distributions. *Ann. Math. Statist.* 19 360-369.

Roy, S.N. (1953). On a heuristic method of test construction and its use in multivariate analysis. *Ann. Math. Statist.* 24 220-238.

Roy, S.N. (1957). *Some Aspects of Multivariate Analysis.* John Wiley and Sons, Inc., New York.

Roy, S.N. and Mikhail, W.F. (1961). On the monotonic character of the power functions of two multivariate tests. *Ann. Math. Statist.* 32 1145-1151.

Schmidt-Koenig, K. (1965). Current problems in bird orientation, 217-278. *Advances in the Study of Behavior*. Lehrman, D.S. et. al., eds. Academic Press, New York.

Schuurmann, F.J., Krishnaiah, P.R. and Chattopadhyay, A.K. (1973). On the distribution of the ratios of the extreme roots to the trace of the Wishart matrix. *J. Mult. Anal.* 3 445-453.

Sherman, S. (1955). A theorem on convex sets with applications. *Ann. Math. Statist.* 26 763-767.

Smith, P.J. (1977). A nonparametric test for bivariate circular symmetry based on the empirical cdf. *Comm. Statist.* A6 209-220.

Strawderman, W.E. (1974). Minimax estimation of location parameters for certain spherically symmetric distributions. *J. Mult. Anal.* 4 255-264.

Sugiura, N. and Nagao, H. (1968). Unbiasedness of some test criteria for the equality of one or two covariance matrices. *Ann. Math. Statist.* 39 1686-1692.

Thomas, D.H. (1970). Some contributions to radial probability distributions, statistics, and the operational calculi. Ph.D. dissertation, Wayne State University.

Vershik, A.M. (1964). Some characteristic properties of Gaussian stochastic processes. *Thy. Prob. Appl.* 9 353-356.

Wells, D.R. (1978). A monotone unimodal distribution which is not central convex unimodal. *Ann. Statist.* 6 926-931.

Wilks, S.S. (1946). Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution. *Ann. Math. Statist.* 17 257-281.

Wolfe, J.S. (1975). On the unimodality of spherically symmetric stable distribution functions. *J. Mult. Anal.* 236-242.

Wynn, H.P. (1977). An inequality for certain bivariate probability integrals. *Biometrika* 64 411-414.

Yao, K. (1973). A representation theorem and its applications to spherically invariant random processes. *IEEE Trans. Inform. Thy.* IT-19 600-608.

Zellner, A. (1976). Bayesian and non-Bayesian analysis of the regression model with multivariate student-t error terms. *J. Amer. Statist. Assoc.* 71 400-405.

# VITA

The author was born on July 2, 1946, in Cleveland, Ohio. After receiving a B.A. in Math and Psychology in 1970 and a M.S. in Math in 1971 from Cleveland State University, she obtained an instructorship in mathematics at Virginia State College in Petersburg, Virginia. After leaving Virginia State College she taught mathematics for Richard Bland College, also in Petersburg. She next decided to enroll in the Ph.D. program at Virginia Polytechnic Institute and State University.

During her four year stay at VPI & SU she taught statistics on a graduate teaching assistantship and, in conjunction with the biology department, performed the statistical analyses for a DOE grant entitled "Oil persistence in tundra and its impact on the belowground ecosystem."

The author is now an assistant professor in the Institute of Statistics, Texas A&M University, College Station, Texas 77843.

Margaret A Chmielewski

INVARIANT TESTS FOR SCALE PARAMETERS

UNDER ELLIPTICAL SYMMETRY

by

Margaret A. Chmielewski

(ABSTRACT)

In the parametric development of statistical inference it often is assumed that observations are independent and Gaussian. The Gaussian assumption sometimes is justified on appeal to central limit theory or on the grounds that certain normal theory procedures are robust. The independence assumption, usually unjustified, routinely facilitates the derivation of needed distribution theory.

In this thesis a variety of standard tests for scale parameters is considered when the observations are not necessarily either Gaussian or independent. The distributions considered are the spherically symmetric vector laws, i.e. laws for which $\underline{x}(n\times 1)$ and $\underline{Px}$ have the same distribution for every $(n\times n)$ orthogonal matrix $\underline{P}$, and natural extensions of these to laws of random matrices. If $\underline{x}$ has a spherical law, then the distribution of $\underline{Ax} + \underline{b}$ is said to be elliptically symmetric.

The class of spherically symmetric laws contains such heavy-tailed distributions as the spherical Cauchy law and other symmetric stable distributions. As such laws need not have moments, the emphasis here is on tests for scale parameters which become tests regarding dispersion parameters whenever second-order moments are defined.

Using the principle of invariance it is possible to characterize the invariant tests for certain hypotheses for all elliptically symmetric distributions. The particular problems treated are tests for the equality of k scale parameters, tests for the equality of k scale matrices, tests for sphericity, tests for block diagonal structure, tests for the uncorrelatedness of two variables within a set of m variables, and tests for the hypothesis of equi-correlatedness. In all cases except the last three the null and non-null distributions of invariant statistics are shown to be unique for all elliptically symmetric laws. The usual normal-theory procedures associated with these particular testing problems thus are exactly robust, and many of their known properties extend directly to this larger class.

In the last three cases, the null distributions of certain invariant statistics are unique but the non-null distributions depend on the underlying elliptically symmetric law. In testing for block diagonal structure in the case of two blocks, a monotone power property is established for the subclass of all elliptically symmetric unimodal distributions.