1992

Estimation Of The Scale Matrix Of A Multivariate T-model

Anwarul Haque Joarder

ESTIMATION OF THE SCALE MATRIX OF A MULTIVARIATE T-MODEL

by

ANWARUL HAQUE JOARDER

Department of Statistical and Actuarial Sciences

Submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Faculty of Graduate Studies

The University of Western Ontario

London, Ontario, Canada

August 1992

National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

ISBN 0-315-75376-5

ABSTRACT

The classical theory of Multivariate Statistical Analysis is primarily based on the multivariate normal model. However, in the recent literature several authors have made studies as to how the conclusions will be affected if the population model departs from normality. The class of elliptical models shares some intrinsic properties of the multivariate normal model and has been getting increasing attention by the researchers in the recent literature.

In the present thesis we restrict the model to a suitable multivariate t-model which belongs to the class of elliptical models and at the same time accommodates the multivariate normal model. This model has found applications in the context of stock market problems. The main results of the thesis are outlined below.

Improved estimators of the scale matrix of the multivariate t-model have been obtained under a squared error loss function. Similar improved estimators for the characteristic roots of the scale matrix, trace of the scale matrix and also for the inverse of the scale matrix have been obtained. Some Improved estimators of the scale matrix of multivariate t-model have been obtained under the entropy loss function. Some other related new results are as follows.

An elegant expression has been obtained for the characteristic function of the multivariate t-distribution in terms of the well-known Macdonald function. Also a limit theorem for the Macdonald function has been obtained. Some identities involving expectations of the sum of product matrix, based on the multivariate t-model, have been derived.

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to my thesis adviser Professor Mir M. Ali for his expert guidance, active cooperation and stimulating discussions at all stages of my research work.

I thank Professors D.R. Bellhouse and D.A. Stanford for going through an earlier draft of the thesis and making constructive comments which have improved the thesis considerably.

The comments and suggestions by Professor M.S. Srivastava on the scale mixture representation of the multivariate t-model have led to substantial reduction in algebra for calculating many results on expectations.

I take this opportunity to thank Professor V.M. Joshi and Dr. S.E. Ahmed for their sincere academic help and care during my student life at the University of Western Ontario.

Finally, I wish to thank Ms. Maria Lavdas, Mr. Leslie Kwarciak and my friends Edmund Rudiuk, Philip Ragama, Sohail Khan, Mohammed Shamsun and Lubafu Nelson for their help related to computer system and computing.

CONTENTS

LIST OF TABLES

Minimum Relative Risk (MRR)

LIST OF FIGURES

Risk Difference= $f(c_2)$

Relative Risk= $RR(\hat{\Lambda}_1 : \tilde{\Lambda}_1; c_2)$

$p = 3$, $\Lambda = ((\lambda_{ik}))$ where $\lambda_{11} = 94, \lambda_{12} = 41, \lambda_{13} = 23, \lambda_{22} = 26, \lambda_{23} = 11$, and $\lambda_{33} = 6$ (see section 6.4.1)

CHAPTER 1

INTRODUCTION

The classical theory of Multivariate Statistical Analysis is based on the assumption that the underlying observation vectors arise from independent multivariate normal distributions. The multivariate normal models have indeed played a predominant role in the historical development of statistical theory, and found applications in almost all branches of science and technology.

There are two main reasons for using the multivariate normal distribution. Firstly, functions of multivariate observations are, usually, approximately normally distributed due to the central limit effect. This is especially true of sample means and covariance matrices, used extensively in formal inferential procedures. Secondly, the statistical analysis of multivariate observations is mathematically easily tractable under the assumption of normality.

Still the question arises as to what happens to the inferences when the normality assumption is violated. There could be several situations:

(i) the functional form of the underlying distribution is not known,

(ii) the underlying distribution is somewhat different from normality,

(iii) the underlying distribution does share some features of normality.

The statistical analysis for the first situation has led to the growth of distribution-free methods better known as Nonparametric Methods. The second situation has been considered by numerous statisticians and this area is usually termed as Robustness Studies. However, the situation when the underlying distribution shares some intrinsic properties of the multivariate normal distribution has been getting increasing attention in the recent literature. One such property is that the probability density function has equiprobable surfaces on concentric hyperspheres. This property is usually known as **spherical symmetry** and the related distributions are known as **spherical distributions**. A simple linear transformation (location-scale) leads to equiprobable surfaces on hyper-ellipses. This property is known as elliptical symmetry and the corresponding distributions are known as **elliptical distributions**.

Some notable subclasses of multivariate elliptical distributions are: multivariate stable distributions, contaminated normal distributions and the multivariate t-distribution. One of the most important multivariate elliptical distribution is the multivariate t-distribution. Since the multivariate t-distribution accommodates the multivariate normal distribution when the so-called degrees of freedom approaches infinity, it is a good candidate to check robustness of classical statistical inference.

The present thesis deals with the generalizations of some classical results developed under normality assumption when in fact the observations have a multivariate t-distribution. Some of the classical statistical theories have been recently generalized in the general set-up of multivariate elliptical models. It is, therefore, felt necessary to give a brief account of the related works on multivariate elliptical distributions.

A recent paper by Chmielewski (1981) presents an excellent review of the historical development of elliptical distributions as an attractive generalization of the multivariate normal distribution.

1.1 Distribution Theory

Earlier papers dealing with spherical distributions are by Bartlett (1934), Hartman and Wintner (1940), and Lord (1954).

Box (1952) notes that the usual F-statistic has the same null distribution for all spherically symmetric distributions. Bennet (1961) also notes that certain statistics have unchanged null distributions under a multivariate t-distribution. Efron (1969) shows that the Student t-distribution remains unchanged if we assume that the underlying observations to be spherically symmetric instead of normal. The robustness property of F-statistic has been proved by Thomas (1970) and Kariya and Eaton (1977) in connection with the study of linear models.

The first organized presentation of elliptical distribution theory is by Kelker (1970) who studies conditions for the existence of a probability density function, the marginal distribution, the conditional distribution, distributions of some functions of norm and quadratic forms etc.

Anderson and Fang (1990) also study the distribution of the sample covariance matrix, the generalized variance, characteristic roots of the covariance matrix and null distribution of Hotelling's T^2-statistic for a class of elliptical distributions. Sutradhar and Ali (1989) find the distribution of the covariance matrix for a class of elliptical distributions. They also specialize the result to the multivariate t-distribution and calculate the moments of the sum of products matrix.

For studies relating to the distribution of correlation coefficient for elliptical models we refer to Muirhead (1980), Muirhead and Waternaux (1980), Ali and Joarder (1991) and Joarder and Ali (1992).

Fan (1990) derives the distribution of generalized noncentral t, F and T^2-statistics for a class of spherical distributions. Fang and Wu (1990) discuss some properties of the distribution of quadratic forms based on a class of elliptical distributions. Anderson and Fang (1987) extend Cochran's theorem from normal distribution to elliptical distributions.

Fang, Fan and Xu (1990) study the distributions of quadratic forms of random idempotent matrices with applications to Hotelling's T^2-statistic and Tukey's test of nonadditivity. Cacoullos and Koutras (1982) find the distribution of generalized non-central chi-square statistic. Fan (1990b) finds the distributions of generalized non-central chi-square statistic and proves the matrix form of the generalized non-central Cochran's theorem.

1.2 Characterizations

Several authors such as Bartlett (1934), Hartman and Wintner (1940), Kelker (1970), Thomas (1970) and Nash and Klamkin (1976) have discussed the following important characterization of the normal law. If $Z_1, Z_2, ..., Z_N$ are independently and identically distributed random variables, then Z $=(Z_1, Z_2, ..., Z_N)'$ is spherically distributed if and only if Z has a normal distribution.

Kelker (1970) proves several beautiful characterization theorems of characteristic functions, norms and quadratic forms of a class of elliptical distributions.

Ali (1980) proves that in order that the sample mean $\overline{Z}$ and sample variance S^2 be independent when Z is spherically distributed, it is necessary and sufficient that $Z_1, Z_2, ..., Z_N$ be mutually independently normally distributed.

Cambanis, Huang and Simons (1981) present many interesting characterizations of elliptical distributions through a stochastic representation which follows from the work of Schoenberg (1938). This representation has been ever increasingly applied to study many tougher problems (see for example, Fang and Anderson, 1990, Fang, Kotz and Ng, 1990 and Fang and Zhang, 1990).

Mitchell (1988) as well as Mitchell (1989) studies the geometrical properties of elliptical distributions both from the viewpoint of robustness and manifolds of Amari (1985) and Lauritzen (1984).

1.3 Inferential Problems

Chmielewski (1980) as well as Chmielewski (1981a) considers some testing problems relating to scale matrices, such as, the equality of several matrices, sphericity, block diagonality and equicorreletedness. He proves that null distributions in these cases are robust in a class of elliptical distributions. The non-null distributions for tests of equality of scale matrices and sphericity of scale matrix are also shown to be robust.

Kariya (1981) shows that the Hotelling's T^2-test for testing that the location parameter equals the null vector in the one sample problem is robust against departures from normality. It remains Uniformly Most Powerful Invariant also for a class of elliptical distributions, and the null distribution is the same as that under normality. Sutradhar (1988) deals with the linear hypothesis with the multivariate t-error variable.

Kariya (1981a) gives necessary and sufficient conditions for which the null distributions of test statistics for most multivariate hypothesis testing problems remain the same in the class of elliptical distributions. He also shows that in certain special cases, the usual Multivariate Analysis of Variance tests remain Uniformly Most powerful Invariant also in a class of elliptical distributions.

Srivastava and Bilodeau (1989) proves that in a subclass of elliptical distributions Stein estimators are robust in estimating the mean vector and the regression parameters in a linear regression model. Singh (1991) studies Stein estimators in a regression model with errors having a multivariate t-model.

Anderson and Fang (1990a) generalize the theory of maximum likelihood estimation and likelihood ratio criteria from normal distributions to elliptical distributions. They also find that many usual likelihood ratio criteria and their null distributions are the same in the elliptical distributions.

Hsu (1990) develops an invariant test for testing the equality between the mean vector and a specified vector and its properties for a class of elliptical distributions. For a class of elliptical distributions Hsu (1990 a) also considers the following problems: MANOVA, lack of correlation among sets of variables, the equality of covariance matrices, the equality of the correlation coefficient to a given number and the equality of the multiple correlation coefficient to zero. He develops invariant tests and their properties for each of the above hypotheses.

Fan (1990a) finds shrinkage estimators and ridge regression estimators of regression parameters in a linear model when the sample comes from a class of elliptical distributions. He presents a class of shrinkage estimators and ridge regression estimators which dominate the ordinary least squares estimators under quadratic loss function. The results are also extended to the case when the loss function is a nondecreasing convex function of quadratic loss function.

Fan and Fang (1990) prove that the sample mean is a minimax estimator of the location parameter when the loss function is a decreasing function of the quadratic loss function and the underlying observations have a class of elliptical distributions. They also consider some sequential minimax properties of the sample mean and the Stien's two stage estimators.

Fan and Fang (1990a) again find some minimax estimators of the location parameter under a loss function when the underlying observations have a class of elliptical distributions. They also consider inadmissibility of sample regression coefficients.

Fan and Fang (1990b) also find conditions for minimaxity of an estimator for a location parameter under a general quadratic loss function when the underlying observations have a class of spherical distributions.

Quan (1990) proves that many likelihood ratio tests are Uniformly Most Powerful Invariant and unbiased when the observations belong to a class of elliptical distributions. Quan and Fang (1990) also study unbiasedness of likelihood ratio tests of some hypotheses regarding location and scale parameters when the underlying observations belong to a class of elliptical distributions.

Dey (1988) has developed estimators for the characteristic roots of the covariance matrix of the multivariate normal distribution by shrinking sample characteristic roots towards their geometric mean under a squared error loss function. Similar techniques have been adopted by Leung (1992) for estimating the characteristic roots of the scale matrix of a multivariate F-distribution. Dey (1990) estimates parameters of a scale mixture of normal distributions.

1.4 Moments and Identities

Kelker (1970) finds the mean and covariance of a class of elliptical distributions. Muirhead (1982) defines the kurtosis parameter of elliptical distributions.

Berkane and Bentler (1986) present an inductive method for computing the moments of a class of elliptical distributions. They introduce a sequence of new parameters relating centered higher order to second order moments.

Berkane and Bentler (1987) define new parameters characterizing a class of elliptical distributions. They also show that Mardia's (1970) coefficient of multivariate kurtosis is essentially one of these parameters. They establish a simple relation between centered multivariate product moments and second moments of the variables. Berkane and Bentler (1987a) find asymptotic distribution of marginal sample kurtoses and uses it to derive new estimators of the kurtosis parameter of multivariate elliptical distributions as well as tests for homogeneity of kurtoses. Gang (1990) expresses moments of a random vector in terms of some operators. He also shows some applications to a class of elliptical distributions.

Joarder and Ali (1992) generalize the identities of expectations of sum of product matrix due to Muirhead (1986) when the observations have a multivariate t-distribution instead of normal.

1.5 Some Applications of Elliptical Distributions

The class of elliptical distributions has found many applications. Dunnet and Sobel (1954) encounter the multivariate t-distribution in the context of certain multiple decision problems while Cornish (1954) discusses the multivariate t-distribution in connection with a set of normal sample deviates. Bechhofer, Dunnet and Sobel (1954) encounter the multivariate t-distribution for ranking means of normal populations with a common unknown variance.

There has been a great deal of work about the use of the t-distribution in financial studies. It has been observed by some authors that the empirical distribution of rates of return of common stock have somewhat thicker tails than that of the normal distribution. The multivariate t-distribution has fatter tails than that of the normal distribution and is, therefore, suitable to describe stock market data. Under the assumption that the errors follow the multivariate t-distribution Zellner (1976) considers a regression model to study a stock market data for a single stock. Sutradhar and Ali (1986) generalizes Zellner's regression model to study the performance of stocks of some selected firms relative to overall performance of all stocks trading on several stock exchanges under the assumption that the errors have the multivariate t-distribution.

Chib, Tiwari and Jammalamadaka (1988) study the prediction problem in linear regression models with elliptical errors. They also extend the results of Zellner (1976).

Sutradhar (1990) examines the behaviour of Fisher's linear discrimination criterion for classifying an observation into one of two t-populations.

Browne and Shapiro (1987) derive a test for factor analysis structure of the covariance matrix under the assumption that the error variates have a class of elliptical distribution.

The problems studied in the present thesis are presented in the next chapter.

CHAPTER 2

PROBLEMS STUDIED IN THE PRESENT THESIS

2.1 A Proposed Model

In multivariate statistical analysis we usually draw independent observations from a multivariate normal distribution. Suppose that the p-dimensional random variable X_j has the multivariate normal distribution given by $N_p(\theta,\ \Lambda)$; then X_j has p-components $X_{1j}, X_{2j}, \ldots, X_{pj}$, usually known as characteristics. Now if we draw N independent observations $X_j,\ j = 1, 2, \ldots, N$, then the observations $(X_1, X_2, \ldots, X_N)$ constitute a random sample of size N. In this case the joint probability density function of the sample is given by

$$f(x_1, x_2, \ldots, x_N) = (2\pi)^{-Np/2}\ |\Lambda|^{-N/2} exp\left(-\frac{1}{2}\sum_{j=1}^{N}(x_j - \theta)'\Lambda^{-1}(x_j - \theta)\right). \quad (2.1)$$

We will call the above joint probability density function the **multivariate normal model** . A natural generalization of the above multivariate normal model is to replace the exponential function by some suitable function g(.) such that

$$f(x_1, x_2, \ldots, x_N) = K_g(p, N)|\Lambda|^{-N/2} g\left(\sum_{j=1}^{N}(x_j - \theta)'\Lambda^{-1}(x_j - \theta)\right), \quad (2.2)$$

where $K_g(p, N)$ is the normalizing constant, is the joint p.d.f. of $(X_1, X_2, \ldots, X_N)$.

It is easily seen that the probability density functions given by (2.1) and (2.2) are constant on the surface of

$$\sum_{j=1}^{N}(x_j-\theta)'\Lambda^{-1}(x_j-\theta)=c^2$$

for every constant c.

We will call the probability density function given by (2.2) the **multivariate elliptical model.** The observations $(X_1, X_2, \ldots, X_N)$ are identically distributed but not necessarily independent, although they are uncorreleted. It is well-known (cf. Kelker, 1970) that the observations of the multivariate elliptical model are independently distributed with mean θ and scale matrix Λ if and only if $X_j(j = 1, 2\ldots, N)$ is distributed according to the p-variate normal distribution. This, of course, is a limitation to the generalization of the multivariate normal model by the multivariate elliptical model. However, the multivariate elliptical model given by (2.2) has found application in stock market problems (see e.g. Zellner, 1976 and Sutradhar and Ali, 1986).

The multivariate elliptical model given by (2.2) is too general to obtain specific results unless the functional form of $g(.)$ is specified. In the recent literature some subclasses of the multivariate elliptical model considered by statisticians are multivariate stable models, contaminated normal models and the multivariate t-model.

In the present thesis we will concentrate on the **multivariate t-model** given by

$$f(x_1, x_2, \ldots, x_N) = K(\nu, Np)\ |\Lambda|^{-N/2}\left(\nu+\sum_{j=1}^{N}(x_j-\theta)'\Lambda^{-1}(x_j-\theta)\right)^{-(\nu+Np)/2} \tag{2.3}$$

where the normalizing constant $K(\nu, Np)$ is given by

$$K(\nu, Np) = \frac{\nu^{\nu/2}\Gamma((\nu+Np)/2)}{\pi^{Np/2}\Gamma(\nu/2)}. \tag{2.4}$$

It is well known that

$$\left(1+\frac{1}{\nu}\sum_{j=1}^{N}(x_j-\theta)'\Lambda^{-1}(x_j-\theta)\right)^{-(\nu+Np)/2} \rightarrow exp\left(\frac{-1}{2}\sum_{j=1}^{N}(x_j-\theta)'\Lambda^{-1}(x_j-\theta)\right)$$

as $\nu \rightarrow \infty$. Thus the **multivariate t-model** accommodates the multivariate normal model by letting $\nu \rightarrow \infty$.

It can be easily verified by direct integration that the joint density function of multivariate t-model given by (2.3) can be rewritten as a mixture of density functions given by:

$$f(x_1,x_2,\ldots,x_N) = \int_0^{\infty} \frac{|\tau^2\Lambda|^{-N/2}}{(2\pi)^{Np/2}} exp\left(\frac{-1}{2}\sum_{j=1}^{N}(x_j-\theta)'(\tau^2\Lambda)^{-1}(x_j-\theta)\right) h(\tau)\, d\tau \tag{2.5}$$

where

$$h(\tau) = \frac{2\tau^{-(\nu+1)}}{\Gamma(\nu/2)(2/\nu)^{\nu/2}}\, exp\left(\frac{-\nu}{2\tau^2}\right). \tag{2.6}$$

This means that the multivariate t-model is a scale mixture of multivariate normal model with location parameter θ and scale parameter $\tau^2\Lambda$ with the mixing scale parameter τ^2 where τ^{-2} has a gamma distribution $G(\nu/2,\ 2/\nu)$. These representations have been exploited by Dey (1990) for some estimation problems, and by Singh (1991) for a linear regression problem with errors having a multivariate t-model.

The multivariate t-distribution has been encountered by several authors in many practical contexts (see e.g. Dunnet and Sobel, 1954, Cornish, 1954, Bechhofer, Dunnet and Sobel, 1954, Fama, 1965, Blattberg and Gonedes, 1974, Zellner, 1976, Sutradhar and Ali, 1986 and Sutradhar, 1990)

2.2 Motivation of the Present Thesis

As mentioned earlier, the theory of multivariate analysis has been developed primarily under the assumption that the observation vectors are independently normally distributed. Since the **multivariate t-model** converges to the **multivariate normal model** when the degrees of freedom approaches infinity, it provides a good basis for checking the sensitivity of classical statistical procedures for departures from normality. In the present thesis we study some of those procedures when the observations have actually a multivariate t-model. An outline of the problems is sketched below.

Fisher and Healy (1956) derive the characteristic function of the univariate t-distribution when ν is odd. Sutradhar (1986) derives a few series representations of the characteristic function of the multivariate t-distribution. We derive a neater form of the characteristic function of the multivariate t-distribution given by (4.1).

Muirhead (1986) derives some useful identities involving expectations of the sum of product matrix based on the multivariate normal distribution. We have generalized those results to the case when the observations follow a multivariate t-model given by (2.3).

Dey (1988) considers the problem of estimation of some functions of the scale matrix of the multivariate normal distribution. Leung (1992) considers the estimation of the characteristic roots of the scale matrix of multivariate F distribution. We have considered the problem of estimation of the scale matrix of the multivariate t-model given by (4.1) and of some functions of it when the observations follow the multivariate t-model given by (2.3) under the squared error loss function (cf. Leung, 1992) given by

$$L(u(A),\ w(\Lambda)) = tr(u(A) - w(\Lambda))^2$$

where $u(A)$ is any suitable estimator of $w(\Lambda)$ and $tr(A)$ means the trace of the square matrix A.

Dey and Srinivasan (1985) consider the problem of estimation of the scale matrix of the multivariate normal distribution under the entropy loss function. Dey (1990) also considers estimation of the scale parameters of the spherical t-model. We have considered the problem of estimation of the scale matrix of the multivariate t-model given by (2.3) under the entropy loss function.

2.3 Some Examples Where the Multivariate t-Model Arises

The multivariate t-model, as mentioned earlier, is a scale mixture of multivariate normal model with location parameter θ and scale parameter $\tau^2\Lambda$ with the mixing scale parameter τ^2 where τ^{-2} has a gamma distribution $G(\nu/2,\ 2/\nu)$. Thus the multivariate t-model fits well in a Bayesian Inference set-up when sampling from multivariate normal distribution $N_p(\theta,\ \tau^2\Lambda)$ with τ having p.d.f. given by (2.6).

Dunnet and Sobel (1954) encounter this model in the context of certain multiple decision problems while Cornish (1954) discusses the model in connection with a set of normal sample deviates. Bechhofer, Dunnet and Sobel (1954) encounter multivariate t-model when ranking means of normal populations with a common unknown variance.

The multivariate t-model has appeared in financial investigations. It was believed earlier that the rates of return on common stocks were adequately characterized by the multivariate normal model. It has been observed by several authors that the empirical distribution of rates of return of common stocks have somewhat thicker tails (more kurtosis) than that of the normal distribution. The evidence provided by Mandelbrot (1963) and Fama (1965) suggests that one could explicitly account for the observed fat tails by using the symmetric stable models.

The multivariate t-model has also fatter tails and can, therefore, characterize rates of return on common stocks. Under the assumption of independence of daily returns, Blattberg and Goenedes (1974) assess the suitability of the multivariate t-model as compared to the symmetric stable models. Under the assumption that the errors follow a multivariate t-model Zellner (1976) considers a regression model to study stock market data for a single stock.

Consider the following regression model:

$$Y_{ij} = \alpha_i + \sum_{r=1}^{m} \beta_{ir} x_{rj} + \epsilon_{ij}, \tag{2.7}$$

for $i = 1, 2, \ldots, p; j = 1, 2, \ldots, N$, the regression parameters of being α_i and $\beta_{i1}, \beta_{i2}, \ldots, \beta_{im}$, for $i = 1, 2, \ldots, p$. Then under the assumptions that

$$E(\epsilon_{ij}) = 0, \quad \text{for all } i,\ j$$

$$E(\epsilon_{ij}^2) = \tau^2 \lambda_{ii} \quad \text{for all } i,\ j$$

$$E(\epsilon_{ij}\epsilon_{kj}) = \tau^2 \lambda_{ik} \quad \text{for } k = 1, 2 \ldots, p, \text{ and for all } i,\ j$$

$$E(\epsilon_{ij}\epsilon_{kj'}) = 0 \quad \text{for all } i, j (\neq j'), k; j' = 1, 2, \ldots, N.$$

and that for a given τ, the errors $(\epsilon_1, \epsilon_2, \ldots, \epsilon_N)$ where $\epsilon_j = (\epsilon_{1j}, \epsilon_{2j}, \ldots, \epsilon_{pj})'$ $(j = 1, 2, \ldots, N)$ are independently and normally distributed as $N_p(0,\ \tau^2 \Lambda)$ while the parameter τ has the p.d.f. given by (2.6), it may be proved that

$$\begin{aligned} f(\epsilon_1, \epsilon_2, \ldots, \epsilon_N) = & \int_0^\infty \frac{|\tau^2 \Lambda|^{-N/2}}{(2\pi)^{Np/2}} exp\left(\frac{-1}{2} \sum_{j=1}^{N} \epsilon_j' (\tau^2 \Lambda)^{-1} \epsilon_j \right) h(\tau)\ d\tau \\ = & K(\nu, Np) |\Lambda|^{-N/2} \left(\nu + \sum_{j=1}^{N} \epsilon_j' \Lambda^{-1} \epsilon_j \right)^{-(\nu + Np)/2} \end{aligned} \tag{2.8}$$

where $K(\nu, Np)$ is given by (2.4). Thus we see that the errors have a multivariate t-model of the form (2.3).

Sutradhar and Ali (1986) consider the above set-up to study the performance of stocks of some selected firms relative to overall performance of all stocks trading on several stock exchanges. They consider the price change data for the stocks of four selected firms : 1. General Electric, 2. Standard Oil, 3. I.B.M. and 4. Sears, trading on the New York Stock Exchange in relation to the performance of New York Stock Exchange as a whole (or perhaps in conjunction with several other stock exchanges).

Let Y_{ij} denote the monthly return on a capital of \$100 , invested on the i-th ($i = 1, 2, 3, 4$) stock during the j-th ($j = 1, 2, \ldots, 20$) month. More specifically,

$$Y_{ij} = 100[(Q_{ij} - P_{ij}) + D_{ij}]/P_{ij},$$

where P_{ij} is the price of the i-th stock at the beginning of the j-th month, Q_{ij} the price at the end of the j-th month and D_{ij} the dividends earned during the j-th month. Let x_j denote the weighted average of these returns during the j-th month for the aggregate of all stocks trading on the New York Stock Exchange, called 'market' for short. To study the linear regression of the joint monthly returns of the selected stocks on the corresponding monthly returns of the 'market' as a whole, Sutradhar and Ali (1986) consider the following regression model:

$$Y_{ij} = \alpha_i + \beta_i x_j + \epsilon_{ij} \tag{2.9}$$

$i = 1, 2, \ldots, p(= 4); j = 1, 2, \ldots, N(= 20)$ and estimate the parameters

$$(\alpha_1,\ \beta_1), \ldots, (\alpha_4,\ \beta_4)$$

on the basis of a stock return data. They also consider the case of r ($r = 1, 2, \ldots, m$) markets in which case the regression model is given by (2.7) and develop theories for testing regression parameters.

For further details the reader is referred to Sutradhar and Ali (1986). Another application is due to Sutradhar (1990) where he examines the behaviour of Fisher's linear discrimination criterion for classifying an observation into one of two t-populations.

2.4 Contributions of the Present Thesis

In the present thesis some characteristic functions, some identities involving expectations, and some estimation problems have been considered.

Major Contributions of the Present Thesis:

The major original contributions of the thesis are contained in Chapter 6 and Chapter 7.

1. The scale matrix of the multivariate t-model has been estimated under a squared error loss function (see sections 6.2, 6.3). The risk functions of the estimators have been calculated (see Section 6.3). The proposed estimator has been compared to maximum likelihood estimator analytically and computationally (see Section 6.4). The result has been specialized to the case of the multivariate normal model (see Sections 6.6, 6.7, 6.8 and 6.9).

Some functions of the scale matrix of the multivariate t-model e.g. characteristic roots, the trace of the scale matrix and the inverted scale have also been estimated (see Sections 6.10, 6.11 and 6.12).

2. Some improved estimators of the scale matrix of the multivariate t-model have also been obtained under the entropy loss function (see Sections 7.2 and 7.3).

Some original but minor contributions of the thesis are as follows.

1. The characteristic function of the multivariate t-distribution has been derived in terms of the Macdonald function (see Section 4.3.1). A limit theorem for

the Macdonald function has also been obtained (see Section 4.8).

2. Some important identities involving expectations of the sum of products matrix based on the multivariate t-model has been obtained (see Chapter 5).

3. A Pearson Type II distribution has been proposed which is closed under marginal and conditional distributions (see Section 3.4).

CHAPTER 3

ELLIPTICAL DISTRIBUTIONS

3.1 Spherical and Elliptical Distributions

A p-dimensional random variable $Z = (Z_1, Z_2, \ldots, Z_p)'$ is said to have a spherical distribution if the probability density function (p.d.f.) is of the form

$$f(z) = g(z'z). \tag{3.1}$$

The density is constant on every concentric spherical surface $z'z = c^2$ centered at the fixed point $0 = (0, 0, \ldots, 0)'$ and hence Z is said to be spherically distributed.

A spherical random variable can also be defined in the following way:

A p-dimensional random variable Z is said to have a spherical distribution if for every C belonging to the class of orthogonal matrices

$$CZ \overset{d}{=} Z \tag{3.2}$$

where the symbol $\overset{d}{=}$ means that the two sides have the same distribution.

A p-dimensional random variable $X = (X_1, X_2, \ldots X_p)'$ is said to have an elliptical distribution if the p.d.f. of X is of the form

$$f(x; \theta, \Lambda) = |\Lambda|^{-1/2} g((x-\theta)'\Lambda^{-1}(x-\theta)). \tag{3.3}$$

where θ is a p-component location vector, Λ is a $p \times p$ positive definite scale matrix.

The p.d.f. in (3.3) is constant on every concentric ellipsoidal surface

$$(x-\theta)'\Lambda^{-1}(x-\theta)=c^2$$

centered at the fixed point $\theta=(\theta_1,\theta_2,...,\theta_p)'$ and hence X is said to be elliptically distributed. A simple linear transformation

$$Z=\Lambda^{-1/2}(X-\theta)$$

in (3.3) results in the above p.d.f. in (3.1).

A list of spherical and elliptical distributions are given below:

(i) **The Multivariate Normal Distribution $N_p(\theta,\Lambda)$**

$$f(x;\theta,\Lambda)=(2\pi)^{-p/2}|\Lambda|^{-1/2}exp(\frac{-1}{2}(x-\theta)'\Lambda^{-1}(x-\theta)). \qquad (3.4)$$

(ii) **The Multivariate t-Distribution $t_p(\theta,\nu\Lambda;\nu)$**

$$f(x)=K(\nu,\ p)(\nu+(x-\theta)'\Lambda^{-1}(x-\theta))^{-(\nu+p)/2}$$

where $K(\nu,\ p)=\frac{\nu^{\nu/2}\Gamma((\nu+p)/2)}{\pi^{p/2}\Gamma(\nu/2)}$.

(iii) **Contaminated Normal Distributions**

$$f(z)=(1-c)(2\pi)^{p/2}\exp(-z'z/2)+c(2\pi\lambda^2)^{-p/2}\exp(-z'z/2\lambda^2), \qquad (3.5)$$

where $0\leq c\leq 1$.

(iv) **Scale Mixture of Normal Distributions**

$$f(x)=\int_0^\infty\frac{|\tau^2\Lambda|^{-1/2}}{(2\pi)^{p/2}}\exp(-\frac{1}{2}(x-\theta)'(\tau^2\Lambda)^{-1}(x-\theta))\ h(\tau)d\tau \qquad (3.6)$$

where τ has a p.d.f. on $[0,\infty)$.

These distributions form a subclass of elliptical distributions with p.d.f. given by (3.3).

(v) **The Uniform Distribution inside Ellipsoid**

$$f(x) = \frac{\Gamma(p/2+1)}{((p+2)\pi)^{p/2}}\, I_T(x) \tag{3.7}$$

where $I_T(x)$ is the indicator function of the set

$$T = \{x : (x-\theta)'\Lambda^{-1}(x-\theta) \leq p+2\}.$$

(vi) **The Pearson Type II Distribution**

$$f(x) = k(b,p)|\Omega|^{-1/2}(1-(x-\theta)'\Omega^{-1}(x-\theta))^{(\nu-p)/2}$$

where $\nu > p$, $(x-\theta)'\Omega^{-}1(x-\theta) \leq 1$ and $k(b,p)$ is the normalizing constant given by

$$k(b,p) = \frac{\Gamma(\nu/2+1)}{\pi^{p/2}\Gamma(\frac{\nu-p}{2}+1)}.$$

(vii) **The Double-exponential Distribution**

$$f(x) = \frac{\Gamma(p/2)}{2\Gamma(p)}(\pi\lambda^2)^{-p/2} \exp\left(-|\sum_{j=1}^{n} \frac{(x_j-\theta)^2}{\lambda^2}|^{1/2}\right).$$

(cf. Bravo and MacGibbon, 1988, p 244)

3.2 Characterizations of Spherical and Elliptical Distributions

In the following theorem we present geometrical arguments to characterize spherical distributions. This theorem has been proved by Schoenberg (1938) in the context of radial symmetry in Hilbert space. The following geometrical proof is due to Ali (1989).

Theorem 3.1(Ali, 1989) In order that $(Z_1, Z_2, \ldots, Z_p)'$ belong to the class of spherical distributions, it is necessary and sufficient that joint distribution of $(Z_1, Z_2, \ldots, Z_p)'$ can be represented as a mixture of distributions of the following form:

$$F(Z_1, Z_2, \ldots, Z_p) = \int_0^\infty G(z_1/r, z_2/r, \ldots, z_p/r) dH(r) \tag{3.8}$$

where $H(r)$ is any arbitrary distribution function of a non-negative random variable and $G(u_1, u_2, \ldots, u_p)$ is the distribution function of a random variable $(U_1, U_2, \ldots, U_p)$ whose unit probability mass is uniformly distributed over the surface (shell) of unit sphere of $\Re^p$.

Proof. Consider any arbitrary probability distribution function $H(r)$ of a non-negative one dimensional random variable R i.e., $H(r) = 0$ for $r \leq 0$. Place this distribution in a Euclidean space $\Re^p$ with Cartesian co-ordinates $(Z_1, Z_2, \ldots, Z_p)'$ along the radial line emanating from the origin $0 = (0, 0, \ldots, 0)'$, coinciding this origin with the origin of R i.e. with $R = 0$. With R=0 fixed, rotate this radial line in the whole space $\Re^p$, and spread the mass $dH(r)$ uniformly over the surface of the sphere of radius r i.e. over

$$z_1^2 + z_2^2 + \ldots + z_p^2 = r^2,$$

for every value of r. It is clear that $H(r)$ would generate a spherical distribution.

On the other hand, given any spherical distribution of $(Z_1, Z_2, \ldots, Z_p)'$, collect the mass of constant density on each surface of radius r and place the mass at a distance r from the origin along any arbitrary line, for every value of r, and denote the resulting distribution function along the radial line by $H(r)$. Indeed,

$$H(r) = P(|Z| \leq r), \ |Z| = (Z_1^2 + Z_2^2 + \ldots + Z_p^2)^{1/2}.$$

Hence, $H(r)$ may be thought of as the generator of the spherical distribution and $H(r)$ completely specifies the spherical distribution.

Thus the members of the class of spherical distributions in $\Re^p$ are in one-to-one correspondence with the members of generator distributions, which are the one dimensional probability distributions of non-negative random variables.

The spreading of mass $dH(r)$ at a distance r along the radial line, uniformly over the sphere of radius r in $\Re^p$, can be easily accomplished by considering a unit mass spread over the sphere of radius r in $\Re^p$ and weighing each point by $dH(r)$.

Now consider a general point $(Z_1, Z_2, \dots, Z_p)'$ in $\Re^p$. The mass of the spherical surface, of a sphere of radius r in R^p, with the uniform distribution of unit mass over the sphere of radius r in $\Re^p$, cut off by

$$(Z_1 \leq z_1, Z_2 \leq z_2, \dots, Z_P \leq z_p)$$

is, by a simple dilation argument of bringing the sphere of radius r in $\Re^p$ to that of radius unity, and by the similarity principle, easily seen to be

$$G(z_1/r, z_2/r, \dots, z_p/r)$$

where G has been defined in the statement of the theorem; hence, weighting by $dH(r)$, the portion of the mass contained on the sphere of radius r in $\Re^p$ of the spherical distribution contained in

$$(Z_1 \leq z_1, Z_2 \leq z_2, \dots, Z_p \leq z_p)$$

is

$$G(z_1/r, z_2/r, \dots, z_p/r)dH(r)$$

so that collecting the mass for every sphere of radius r, we have the distribution function of the spherical distribution

$$F(z_1, z_2, \dots, z_p) = \int_0^\infty G(z_1/r, z_2/r, \dots, z_p/r)dH(r),$$

which is the mixture of distribution functions of the uniform distribution over a sphere of radius r in $\Re^p$ with the mixing distribution function $H(r)$, which is the distribution function of a non-negative random variable. Hence the theorem is proved.

By taking the Fourier-Stieltjes transform (characteristic function) of both sides of (2.9), it readily follows that the characteristic function $\phi(t_1, t_2, \ldots, t_p)$ of the spherical class must have the representation

$$\phi_Z(t_1, t_2, \ldots, t_p) = \int_0^\infty \psi(rt_1, rt_2, \ldots, rt_p) dH(r) \tag{3.9}$$

where $\psi(t_1, t_2, \ldots, t_p)$ is the characteristic function of a distribution whose unit mass is uniformly distributed over the surface of a unit sphere in $\Re^p$ with center at the origin and $H(r)$ is a distribution function with $H(r) = 0$ for $r \leq 0$. This characterization is due to Schoenberg (1938) who has proved that the above characteristic function is in one-to-one correspondence with the class of positive definite functions of $(t_1, t_2, \ldots, t_p)$ under radial symmetry in the Hilbert space and this class is completely characterized by the class of distribution functions $H(r)$.

The characteristic function of a spherical distribution can also be found by direct integration and is given in the following theorem.

Theorem 3.2 A random variable Z is said to have a spherical distribution if and only if its characteristic function $\phi_Z(t)$ is a function of $||t|| = (t't)^{1/2}$ in which case it must be of the form

$$\phi_Z(t) = \psi(||t||) = \int_0^\infty {}_0F_1\left(\frac{p}{2}; \frac{-(r||t||)^2}{4}\right) dH(r) \tag{3.10}$$

where $H(r)$ is the distribution function of norm $R = (Z'Z)^{1/2}$ and ${}_pF_q(a; b; z)$ is the generalized hypergeometric function defined by

$$ {}_pF_q(a; b; z) = \sum_{k=0}^\infty \frac{(a)_1(a)_2 \ldots (a)_p}{(b)_1(b)_2 \ldots (b)_q} \frac{z^k}{k!} \tag{3.11}$$

with $(a)_j$ the Pochammer polynomial defined by

$$(a)_j = \frac{\Gamma(a+j)}{\Gamma(a)} = a(a+1)\dots(a+j-1). \tag{3.12}$$

Proof. Let the spherical random variable Z have the p.d.f. given by

$$f(z) = g(z'z).$$

Then the characteristic function of Z is given by

$$\phi_Z(t) = \int_{\mathbb{R}^p} exp(it'z)\, g(z'z)\, dz.$$

By making the following orthogonal transformation

$$Y = (y_1, y_2, \dots, y_p)' = CZ$$

with $y_1 = (t'z)/||t||$, we have

$$\phi_Z(t) = \int_{\mathbb{R}^p} exp(i||t||y_1)\, g(y'y)\, dy.$$

Now let us make the p-dimensional polar transformation from $(y_1, y_2, \dots, y_p)$ to $(r, \theta_1, \dots, \theta_{p-1})$

$$\begin{aligned} y_k &= r\left(\prod_{l=1}^{k-1} sin\theta_l\right) cos\theta_k, \quad k = 1, 2, \dots, p-1, \\ y_p &= r\left(\prod_{k=1}^{p-2} sin\theta_k\right) sin\theta_{p-1} \end{aligned}$$

where $r\epsilon[0,\infty)$, $\theta_k \epsilon[0,\pi)$ for $k = 1, 2, \dots, p-2$, and $\theta_{p-1}\epsilon[0, 2\pi)$ so that the Jacobian of transformation is given by

$$r^{p-1} \prod_{k=1}^{p-2} (sin\theta_k)^{p-k-1}.$$

Then

$$\phi_Z(t) = \int_0^{\infty} \int_{\theta_1=0}^{\pi} \cdots \int_{\theta_{p-2}=0}^{\pi} \int_{\theta_{p-1}=0}^{2\pi} r^{p-1} exp(ir||t||cos\theta_1)$$
$$\times \prod_{k=1}^{p-2} (sin\theta_k)^{k-p-1}\ dr \prod_{k=1}^{p-1} d\theta_k.$$

Integration over $\theta_2, \ldots\ \theta_{p-1}$ yields

$$\phi_Z(t) = \frac{\Gamma(p/2)}{\sqrt{\pi}\Gamma((p-1)/2)} \int_{r=0}^{\infty} \int_{\theta_1=0}^{\pi} (sin\theta_1)^{p-2} exp(ir||t||cos\theta_1) d\theta_1\ dH(r)$$

where h(r) is the p.d.f. of norm $R = (Z'Z)^{1/2}$ given by

$$h(r) = \frac{2\pi^{p/2}}{\Gamma(p/2)}\ r^{p-1}\ g(r^2).$$

Now by the use of the following integral (see e.g.,Prudnikov et al., vol.1, 1986, p 457, formula no. 6)

$$\int_0^{\pi} (sin\ u)^{\alpha-1}\ exp(ix\ cos\ u)\ du = \sqrt{\pi}\ (2/x)^{(\alpha-1)/2}\ \Gamma(\alpha/2)\ J_{\alpha/2-1}(x), \quad \alpha > 0$$

with $\alpha = p-1, x = r||t||$ and $u = \theta_1$, we have

$$\phi_Z(t) = \Gamma(p/2) \int_0^{\infty} (\frac{2}{r||t||})^{p/2-1} J_{p/2-1}(r||t||)\ dH(r)$$

where $J_\alpha(t)$ is the Bessel function of first kind of arbitrary order α and argument x (see e.g. Lebedev, 1965, p 102). Then the expression in (3.10) follows by the use of the following relation

$${}_0F_1\left(b; \frac{-z^2}{4}\right) = \Gamma(b)(2/z)^{b-1} J_{b-1}(z)$$

with $b = p/2$ and $z = r||t||$.

Conversely let the characteristic function of any random variable Z be denoted by $\psi(||t||)$. Then the characteristic function of a random variable $Y = CZ$, where C is any orthogonal matrix, 's given by

$$E[exp(it'Y)] = E[exp(it'CZ)] = \psi(||C't||) = \psi(||t||)$$

which shows that the distribution of Z is invariant under orthogonal transformation and hence by (3.2) Z must have a spherical distribution. Hence the proof is complete.

The following stochastic representation of a spherical random variable follows easily from the previous theorem.

Theorem 3.3 Let Z have a spherical distribution with p.d.f. given by (2.1). Then Z has the stochastic representation given by

$$Z = RU$$

where $R = (Z'Z)^{1/2}$ is independent of U and the random variable U is uniformly distributed on the surface of unit sphere in $\Re^p$.

3.3 Uniform Distributions on or inside Unit Hyper-sphere

The uniform distributions on or inside hyper-spheres plays a very important role in statistical analysis of directional data analysis (see e.g. Watson, 1984). They have been powerful tools in studying spherical and elliptical distributions as well (cf. Fang and Anderson, 1990). In this section we find characteristic functions and moments of uniform distributions on or inside hyper-spheres. Most of these results follow from Cambanis, Huang and Simons (1981) and Fang, Kotz and Ng (1990).

In what follows we will require the following formula of volume $V(p,\ r)$ and surface area $S(p,\ r)$ of a p-dimensional sphere of radius r.

$$V(p,\ r) = \frac{(\pi r^2)^{p/2}}{(p/2)\Gamma(p/2)} \tag{3.13}$$

$$S(p,\ r) = \frac{2\pi^{p/2}}{\Gamma(p/2)}\, r^{p-1}. \tag{3.14}$$

3.3.1 The Uniform Distribution on the Surface of Unit Hyper-Sphere

Let U be uniformly distributed on the surface of unit sphere in $\Re^p$ i.e. the p.d.f. of U is given

$$f(u) = \frac{I_T(u)}{S(p,\ 1)}$$

where $S(p,\ 1)$ is defined in (3.14) and $I_T(u)$ is the indicator function of the set

$$T = \{u : u'u = 1\}.$$

In the next theorem we derive that the characteristic function of the uniform distribution on the surface of unit sphere as an immediate consequence of Theorem 3.2.

Theorem 3.4 Let U be uniformly distributed on the surface of unit sphere in $\Re^P$. Then

$$\phi_U(t) = {}_0F_1\left(\frac{p}{2}; \frac{-||t||^2}{4}\right) \tag{3.15}$$

$$E(U) = 0 \tag{3.16}$$

$$Cov(U) = p^{-1} I_p \tag{3.17}$$

and for any integers $m_1, m_2, \ldots m_k$ with $m = \sum_{k=1}^{p} m_k$, the product moment is given by

$$E\left(\prod_{k=1}^{p} U_k{}^{m_k}\right) = \begin{cases} 0 \text{ if at least one } m_k (k = 1, 2, \ldots, p) \text{ is odd,} \\ \frac{\Gamma(p/2)}{2^m \Gamma((m+p)/2)} \prod_{k=1}^{p} \frac{(m_k)!}{(m_k/2)!} \\ \text{if all } m'_k s (k = 1, 2, \ldots, p) \text{ are even.} \end{cases} \tag{3.18}$$

Proof. Since the distribution of R, as defin a Theorem 3.2, on the surface of unit sphere is degenerate at $r = 1$ i.e. the p.d.f. of R is given by

$$f(r) = 1, \ for\ r = 1,$$

it readily follows from Theorem 3.2 that the characteristic function of U is given by (3.15).

It follows from Theorem 3.3 that for any spherical random variable Z

$$E(Z) = E(||Z||)E(U) \tag{3.19}$$

and

$$Cov(Z) = E(||Z||^2)Cov(U). \tag{3.20}$$

Since the above relations are true for any spherical random variable Z, let Z have spherical normal distribution $N_p(0, I)$ so that $||Z||^2$ is distributed as ${\chi_p}^2$ and consequently

$$E(Z) = 0$$

$$Cov(Z) = I$$

and

$$E(||Z||^2) = p.$$

Then the expressions in (3.16) and (3.17) follows from (3.19) and (3.20). It also follows from Theorem 3.3 that *for any spherical random variable* Z

$$E\left(\prod_{k=1}^{p} Z_k^{m_k}\right) = E(R^m)\, E\left(\prod_{k=1}^{p} U_k^{m_k}\right) \tag{3.21}$$

or,

$$E\left(\prod_{k=1}^{p} U_k^{m_k}\right) = (E(R^m))^{-1} E\left(\prod_{k=1}^{p} Z_k^{m_k}\right).$$

Since this is also true for any random variable Z, let Z have a spherical normal distribution $N_p(0, I)$ so that for any integer l we have

$$E(R^l) = \frac{2^{l/2}\Gamma((l+p)/2)}{\Gamma(p/2)}$$

and

$$E(Z_1^l) = \begin{cases} 0 \text{ if at least one } l \text{ is odd}, \\ \frac{l!}{2^{l/2}(l/2)!} \text{ if all } l\text{'s are even}. \end{cases}$$

Hence the expression in (3.18) follows immediately.

We summarize these results on spherical distributions in the form of the following theorem.

Theorem 3.5 Let Z have spherical distribution given by (3.1). Then

$$E(Z) = 0, \tag{3.22}$$

$$Cov(Z) = (1/p)E(R^2)I_p \tag{3.23}$$

and for any integers $m_1, m_2, \ldots, m_k$ with $m = \sum_{k=1}^{p} m_k$, the product moment is given by

$$E\left(\prod_{k=1}^{p} Z_k^{m_k}\right) = \begin{cases} 0 \text{ if at least one } m_k (k = 1, 2, \ldots, p) \text{ is odd}, \\ E(R^m)\frac{\Gamma(p/2)}{2^m \Gamma((m+p)/2)} \prod_{k=1}^{p} \frac{(m_k)!}{(m_k/2)!} \\ \text{if all } m_k' s (k = 1, 2, \ldots, p) \text{ are even}. \end{cases} \tag{3.24}$$

where $R = (Z'Z)^{1/2}$ is the norm of Z with p.d.f.

$$h(r) = S(p,\ 1)\ r^{p-1} g(r^2). \tag{3.25}$$

The results of the above theorem can be easily extended to the case of elliptical distributions. Let X have the elliptical distribution given by (3.3). Then X can be written as

$$X = \theta + \Lambda^{1/2} Z$$

where Z is spherically distributed. It then follows that the characteristic function of X is given by

$$\begin{aligned}\phi_X(t) =& E(exp(it'X)) \\ =& E\left[exp\left(it'(\theta + \Lambda^{1/2}Z)\right)\right] \\ =& exp(it'\theta)\ E\left[exp(it'\Lambda^{1/2}Z)\right] \\ =& exp(it'\theta)\ \phi_Z(\Lambda^{1/2}t) \\ =& exp(it'\theta)\ \psi(||\Lambda^{1/2}t||)\end{aligned}$$

where $\psi(\cdot)$ is given by (3.10).

It follows from Theorem 3.3 that when Λ is of full rank then we have

$$\Lambda^{-1/2}(X - \theta) = RU,$$

so that

$$X = \theta + R\Lambda^{1/2}U.$$

It then follows from Theorem 3.5 that

$$E(X) = E(\theta + \Lambda^{1/2}Z) = \theta$$

and

$$\begin{aligned}Cov(X) =& E(X - \theta)(X - \theta)' \\ =& E(R\Lambda^{1/2}UU'\Lambda^{1/2}R) \\ =& E(R^2)\Lambda^{1/2}E(UU')\Lambda^{1/2} \\ =& E(R^2)\Lambda^{1/2}(p^{-1}I_p)\Lambda^{1/2} \\ =& (1/p)E(R^2)\Lambda\end{aligned}$$

The corresponding results for elliptical distributions are summarized in the form of the following theorem.

Theorem 3.6 Let X have the elliptical distribution with density given by (3.2). Then

(i) the characteristic function of X is given by

$$\phi_X(t) = exp(it'\theta)\psi(||\Lambda^{1/2}t||) = \int_0^\infty exp(it'\theta)\ {}_0F_1\left(\frac{p}{2}; \frac{-||r\Lambda^{1/2}t||^2}{4}\right) h(r)dr,$$

(ii)

$$X = \theta + R\Lambda^{1/2}U$$

where Λ is of full rank,

(iii)

$$E(X) = \theta$$

and

$$Cov(X) = (1/p)E(R^2)\Lambda = -2\psi'_X(0)\Lambda$$

where R has the p.d.f. given by (3.25) and ${}_pF_q(a; z)$ is defined in (3.11).

3.3.2 The Uniform Distribution inside Unit Hyper-sphere

Let W have a uniform distribution inside a unit sphere having p.d.f.

$$f(w) = \frac{I_T(w)}{V(p,\ 1)}$$

where $I_T(w)$is the indicator function of the set

$$T = \{w : w'w \leq 1\}.$$

By making a p-dimensional polar transformation from $(w_1, w_2, \ldots, w_p)$ to $(r, \theta_1, \theta_2, \ldots, \theta_{p-1})$ where $r\epsilon[0, 1)$, $\theta_k\epsilon[0, \pi)$, $k = 1, 2, \ldots, p-1$, and $\theta_{p-2}\epsilon[0, 2\pi)$ followed by integration over $\theta_1, \theta_2, \ldots, \theta_{p-1}$ we immediately have the p.d.f. of R given by

$$h(r) = \frac{r^{p-1}}{V(p,\ 1)}\ S(p,\ 1) = \frac{S(p,\ r)}{V(p,\ 1)} = p\ r^{p-1},\ r\epsilon[0, 1)$$

i.e. R is distributed as a beta random variable $Beta(p, 1)$.

It follows from Theorem 3.3 that $W = RU$ where R is distributed as a beta random variable $Beta(p,1)$ while U has a uniform distribution on the surface of a unit hyper-sphere and R and U are independently distributed.

Hence from Theorem 3.2 it follows that the characteristic function of W is given by

$$\phi_W(t) = \int_0^1 {}_0F_1\left(\frac{p}{2}; \frac{-||rt||^2}{4}\right) p\, r^{p-1} dr.$$

By the use of series representation of the generalized hypergeometric function, given by (3.11), followed by integrating over r, it is readily seen that

$$\phi_W(t) = {}_0F_1\left(p/2+1; \frac{-||t||^2}{4}\right).$$

The product moments of the components of $W = (W_1, W_2, \ldots, W_p)$, for any integers $m_1, m_2, \ldots, m_p$ with $m = \sum_{k=1}^{p} m_k$, is given by

$$E\left(\prod_{k=1}^{p} W_k^{m_k}\right) = \begin{cases} 0 \text{ if at least one } m_k(k=1,2,\ldots,p \text{ is odd,} \\ \frac{p}{m+p} \frac{\Gamma(p/2)}{2^m \Gamma((m+p)/2)} \prod_{k=1}^{p} \frac{(m_k)!}{(m_k/2)!} \\ \quad \text{if all } m_k's(k=1,2,\ldots,p) \text{ are even} \end{cases}$$

which follows from (3.24) by noting the fact that R is distributed according to $Beta(p,1)$ so that

$$E(R^m) = \frac{p}{m+p}.$$

3.4 The Pearson Type II Distribution

Kotz (1975) proposes the multivariate Pearson Type II Distribution as an extension to the univariate Pearson Type II Distribution. We make a slight modification to the form of the density so that the marginal and conditional distributions are well identified. We consider the following p.d.f.

$$f(x) = k(b,p)\ |\Omega|^{-1/2}\left(1-(x-\theta)'\Omega^{-1}(x-\theta)\right)^{(b-p)/2} \tag{3.26}$$

where

$$(x-\theta)'\Omega^{-1}(x-\theta) \le 1,$$

$b > p$ and the normalizing constant $k(b,p)$ is given by

$$k(b,p) = \frac{\Gamma(b/2+1)}{\pi^{p/2}\Gamma((b-p)/2+1)}.$$

We will denote the above p.d.f. by $PII(\theta,\ \Omega_{22}; b)$. The p.d.f. of

$$Z = \Omega^{-1/2}(X-\theta)$$

is given by

$$f(z) = k(b,p)(1-z'z)^{(b-p)/2},\ z'z \le 1. \tag{3.27}$$

When p=1, the p.d.f. of Z_1 is given by

$$f(z_1) = B^{-1}\left(\frac{1}{2},\frac{b+1}{2}\right)(1-z_1^2)^{(b-1)/2},\ \ z_1^2 \le 1,\ b>1$$

(cf. Stuart and Ord 1987, vol.1, p 248).

The Distribution of Norm

The distribution of norm $R = (Z'Z)^{1/2}$ is given by

$$h(r) = k(b,p)\ S(p,\ 1)\ r^{p-1}\ (1-r^2)^{(b-p)/2}.$$

The m-th moment of R is given by

$$E(R^m) = \frac{1}{2}B\left(\frac{2m+p}{2}, \frac{b-p+2}{2}\right). \tag{3.28}$$

It then follows from Theorem 3.6 that

$$E(X) = \theta$$

and

$$Cov(X) = \frac{1}{2p}B\left(\frac{p+4}{2}, \frac{b-p+2}{2}\right)\Lambda. \tag{3.29}$$

The product moments of the components of Z having p.d.f. (3.27) follows from (3.24) and (3.28).

The Characteristic Function

The characteristic function of the univariate Pearson Type II distribution is given by

$$\begin{aligned}\phi_{Z_1}(t) &= \int_{z_1=-1}^{1} \frac{\sqrt{\pi}\Gamma((b+1)/2)}{\Gamma((b+2)/2)} \exp(itz_1)\,(1-{z_1}^2)^{(b-1)/2} dz_1 \\ &= 2\frac{\sqrt{\pi}\Gamma((b+1)/2)}{\Gamma((b+2)/2)} \int_{z_1=0}^{1} \cos(tz_1)(1-z_1^2)^{(b-1)/2} dz_1.\end{aligned}$$

Now by the use of the following integral (see e.g. Prudnikov, et al., 1986, vol.1, p 389)

$$\int_0^a (a^2-x^2)^{c-1} costx\; dx = \frac{\sqrt{\pi}}{2}(2a/t)^{c-1/2}\;\Gamma(c)\;J_{c-1/2}(at)$$

with $a = 1$ and $c = (b+1)/2$, we have

$$\phi_{Z_1}(t) = 2^{b/2}\;\Gamma(b/2+1)\;|t|^{-b/2}\;J_{b/2}(|t|)$$

where $J_b(t)$ is the Bessel function of first kind.

Finally by using the relation between generalized hypergeometric function and the Bessel function of first kind (see Theorem 3.2) we obtain

$$\phi_{Z_1}(t) = \ _0F_1\left(\frac{b+2}{2}; \frac{-||t||^2}{4}\right).$$

Hence the characteristic function of the multivariate Pearson Type II distribution having p.d.f. f(z) given by (3.27) is given by

$$\phi_Z(t) = \int_{z'z\leq 1} cos\ t'z\ f(z)dz.$$

By an orthogonal transformation

$$Y = (y_1, y_2 \ldots, y_p)' = CZ$$

with $y_1 = t'Z/||t||$ followed by integration over $y_2, y_3, \ldots, y_p$, we immediately have

$$\phi_Z(t) = \phi_{Y_1}(||t||) = \phi_{Z_1}(||t||) = \ _0F_1\left(\frac{b+2}{2}; \frac{-||t||^2}{4}\right).$$

Hence the characteristic function of $X = \theta + \Omega^{1/2}Z$ is given by

$$\phi_X(t) = exp(it'\theta)\ _0F_1\left(\frac{b+2}{2}; \frac{-||\Lambda^{1/2}t||^2}{4}\right). \tag{3.30}$$

The above characteristic function of $X = \theta + \Omega^{1/2}Z$ can also be easily found by the use of Theorem 3.6. It follows that

$$\phi_X(t) = \int_0^\infty exp(it'\theta)\ _0F_1\left(\frac{p}{2}; \frac{-||r\Lambda^{1/2}t||^2}{4}\right)\ h(r)dr$$

where $h(r)$ is the p.d.f. of norm $r = (z'z)^{1/2}$. Then the result (3.30) follows by integration over r.

Marginal and Conditional Distributions

Let us partition X, θ, t and Ω as follows:

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}, t = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}, \Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}$$

where $\theta_2,\ t_2 \epsilon \Re^q$ $(q < p)$ and Ω_{22} is a $q \times q$ positive definite matrix.

Setting $t_1 = 0$ in the c.f. of X in (3.30) we immediately have

$$\phi_{X_2}(t) = exp(it_2'\theta)\ {}_0F_1\left(\frac{b+2}{2}; \frac{-||t_2||^2}{4}\right).$$

Hence $X_2 \sim PII(\theta_2,\ \Omega_{22};\ b)$ i.e. X_2 has the p.d.f.

$$f(x_2) = k(b,q)\ |\Omega_{22}|^{-1/2}\left(1 - (x_2 - \theta_2)'\Omega^{-1}(x_2 - \theta_2)\right)^{(b-q)/2} \tag{3.31}$$

where

$$(x_2 - \theta_2)'\Omega_{22}^{-1}(x_2 - \theta_2) \leq 1,\ b > q.$$

In what follows we will need the following two well-known identities:

$$|\Omega| = |\Omega_{22}|\ |\Omega_{11.2}| \tag{3.32}$$

and

$$y'\Omega y = y_2'\Omega_{22}^{-1}y_2 + (y_1 - \Omega_{12}\Omega_{22}^{-1}y_2)'\Omega_{11.2}^{-1}(y_1 - \Omega_{12}\Omega_{22}^{-1}y_2) \tag{3.33}$$

where $Y = X - \theta$ and

$$\Omega_{11.2} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}.$$

It follows from the identity in (3.33) that

$$\begin{aligned} 1 - y'\Omega\mathrm{y} &= 1 - y_2'\Omega_{22}^{-1}y_2 - (y_1 - \Omega_{12}\Omega_{22}^{-1}y_2)'\Omega_{11.2}^{-1}(y_1 - \Omega_{12}\Omega_{22}^{-1}y_2) \\ &= (1 - y_2'\Omega_{22}^{-1}y_2) \\ &+ \left[1 - (y_1 - \Omega_{12}\Omega_{22}^{-1}y_2)'V_{11.2}^{-1}(y_1 - \Omega_{12}\Omega_{22}^{-1}y_2)\right] \end{aligned}$$

where

$$V_{11.2} = (1 - y_2'\Omega_{22}^{-1}y_2)\Omega_{11.2}.$$

Then the joint density of Y_1 and Y_2 can be written as

$$\begin{aligned} f(y_1, y_2) =& k(b,p)|\Omega_{22}|^{-1/2}|\Omega_{11.2}|^{1/2}(1 - y_2'\Omega_{22}^{-1}y_2)^{(b-p)/2} \\ & \times \left[1 - (y_1 - \Omega_{12}\Omega_{22}^{-1}y_2)'V_{11.2}^{-1}(y_1 - \Omega_{12}\Omega_{22}^{-1}y_2)\right]^{(b-p)/2} \end{aligned} \tag{3.34}$$

so that from (3.31) we immediately have

$$\begin{aligned} f(y_1|y_2) =& k(b-q, p-q)|V_{11.2}|^{-1/2} \\ & \times \left[1 - (y_1 - \Omega_{12}\Omega_{22}^{-1}y_2)'(1 + y_2'\Omega_{22}^{-1}y_2)^{-1}V_{11.2}^{-1}(y_1 - \Omega_{12}\Omega_{22}^{-1}y_2)\right]^{(b-p)/2} \end{aligned}$$

From the relation $Y = X - \theta$ we finally have

$$(X_1|X_2 = x_2) \sim PII_{p-q}(\theta_{1.2},\ \Omega_{11.2}^*; b - q).$$

where

$$\theta_{1.2} = \theta_1 + \Omega_{12}\Omega_{22}^{-1}(x_2 - \theta_2)$$

and

$$\Omega_{11.2}^* = (1 - (x_2 - \theta_2)'\Omega_{22}^{-1}(x_2 - \theta_2))\Omega_{11.2}.$$

Then by the use (3.29) it follows from Theorem 3.6 that

$$E(X_1|X_2 = x_2) = \theta_{1.2}$$

and

$$Cov(X_1|X_2 = x_2) = \frac{1}{2p}B\left(\frac{b-q+4}{2}, \frac{b-p+2}{2}\right)\Omega_{11.2}^*.$$

A Pearson Type II Model

A Pearson Type II model for N p-dimensional observations $X_1, X_2, \ldots, X_N$ could be given by

$$f(x_1, x_2, \ldots, x_N) = k(b, Np)|\Omega|^{-N/2} \left(1 - \sum_{j=1}^{N}(x_j - \theta)'\Omega^{-1}(x_j - \theta)\right)^{(b-Np)/2}$$

where $b > Np$ and $k(b, Np)$ is the normalizing constant. The above model is a generalization of the Pearson Type II distribution given by (3.26) along the line (2.2).

CHAPTER 4

THE MULTIVARIATE T-DISTRIBUTION

The multivariate t-distribution is a natural generalization of the univariate Student t-distribution. The multivariate t-distribution converges to the multivariate normal distribution when the degrees of freedom of the multivariate t-distribution tends to infinity. As mentioned earlier the multivariate t-distributio... is a suitable candidate to check the robustness of statistical techniques developed under normality against alternatives with fat tails (see e.g. Sutradhar and Ali, 1986)

A p-variate random variable $X = (X_1, X_2, \ldots, X_p)'$ is said to have a multivariate t-distribution if the p.d.f. of X is given by

$$f(x) = K(\nu,p)|\Lambda|^{-1/2}\left(\nu + (x-\theta)'\Lambda^{-1}(x-\theta)\right)^{-(\nu+p)/2} \tag{4.1}$$

where $x \in \Re^p$, $\nu > 0$ and the normalizing constant $K(\nu,p)$ is given by

$$K(\nu,p) = \frac{\nu^{\nu/2}\Gamma((\nu+p)/2)}{\pi^{p/2}\Gamma(\nu/2)}.$$

The parameter θ is a p-dimensional location parameter, Λ a $p \times p$ positive definite scale matrix and ν is a scalar known as degrees of freedom.

The multivariate t-distribution has been studied in various contexts by several authors. Among them we mention Bechhofer, Dunnet and Sobel (1954), Cornish (1954), Siotani (1976), Sutradhar and Ali (1986), Sutradhar and Ali (1989) and Sutradhar (1990).

In this chapter we derive an elegant expression for the characteristic function (see section 4.3.1) of the multivariate t-distribution in terms of the Macdonald function. It is shown in section 4.6 that this representation allows us to demonstrate with great ease the well-known property that the multivariate t-distribution can be written as a scale mixture of a suitable multivariate normal distribution with the mixing scale parameter having a suitable chi-square distribution. The characteristic function leads to a limit theorem (see Section 4.8) for the Macdonald function. Some applications of these results are also demonstrated.

4.1 The Standard t-Distribution

We first define what we term as a **Multivariate Standard t-Distribution.**

Definition 4.1 A p-component random vector $Z = (Z_1, Z_2, \ldots, Z_p)'$ will be said to have a **Multivariate Standard t-Distribution** with location parameter 0 and scale parameter I if Z has the p.d.f. (probability density function) given by

$$f(z) = \frac{\Gamma((\nu+p)/2)}{\pi^{p/2}\Gamma(\nu/2)}(1+z'z)^{-(\nu+p)/2}, \ z\epsilon\Re^p, \ \nu > 0 \ . \qquad (4.2)$$

For brevity we will denote it by $Z \sim t_p(0, I; \ \nu)$. In particular, for the univariate case we omit the dimension p and refer to (4.2) simply as $Z \sim t(0,1; \ \nu)$ and in this case Z has the p.d.f.

$$f(z) = \frac{\Gamma(\nu+1/2)}{\sqrt{\pi}\,\Gamma(\nu/2)}(1+z^2)^{-(\nu+1)/2}, \ z\epsilon\Re, \ \nu > 0. \qquad (4.3)$$

The p.d.f. in (4.2) plays a similar role in multivariate t-distribution theory as that played by the standard normal distribution in multivariate normal theory. In particular, we note that the distribution of Z is spherical. Other t-type distributions spherical or elliptical with general location and scale parameters can be obtained by suitable linear transformation of the standard t-distribution. Our aim is to find the characteristic function for the multivariate t-distribution. Since the linear transform of a characteristic function is mathematically easily handled it is enough to derive the same for the standard case.

We now introduce location and scale parameters θ and Ω (a $p \times p$ positive definite matrix) respectively in (4.2).

Let $X = \theta + \Omega^{1/2} Z$ where Z is spherically distributed according to (4.2). Then the density function of X is given by

$$f(x) = K\ |\Omega|^{-1/2} \left(1 + (x-\theta)'\Omega^{-1}(x-\theta)\right)^{-(\nu+p)/2}, \tag{4.4}$$

where

$$K = \frac{\Gamma((\nu+p)/2)}{\pi^{p/2}\Gamma(\nu/2)},\ \nu > 0,\ x \in \Re^p$$

(cf. Johnson 1987, p 117). The p.d.f. in (4.4) can be denoted by $X \sim t_p(\theta, \Omega;\ \nu)$.

Several forms of the multivariate t-distribution have appeared in the literature. They can be obtained by suitable reparametrization of the model in (4.4). In particular the case when $\Omega = \ell(\nu)\Lambda$ where $\ell(\nu)$ is a suitable function of ν in (4.4) is extensively found in the literature. This is because the multivariate t-distribution is primarily useful as a generalization of the multivariate normal distribution in the context of robustness studies. Hence it is desirable that the model (4.4) include the multivariate normal distribution. But the model (4.4) does not in general include a multivariate normal model (even not necessarily when $\nu \to \infty$). However when $\Omega = \ell(\nu)\Lambda$ where the scalar $\ell(\nu) = a\nu + b > 0$, and $a > 0$, it is easily verified by direct computation that the model (4.4) does indeed tend to a multivariate normal model as $\nu \to \infty$, and hence this parametrization does accommodate the multivariate normal distribution as a limiting case.

4.2 Stochastic Representation of the Multivariate t-Distribution

Let $Z \sim t_p(0, I;\ \nu)$. Then it follows from Theorem 3.3 and Theorem 3.5 that Z has the following stochastic representation $Z = LU$ where $L = (Z'Z)^{1/2}$ has the p.d.f.

$$h(l) = \frac{2}{\Gamma(p/2, \nu/2)}\ l^{p-1}(1+l)^{-(\nu+p)/2}$$

i.e. $\nu L^2/p \sim F(p,\ \nu)$.

Again let $Y \sim t_p(0, \nu I;\ \nu)$ so that $R = (Y'Y)^{1/2}$ has the p.d.f. given by

$$h(r) = \frac{2\nu^{\nu/2}}{B(p/2, \nu/2)}\ r^{p-1}(\nu + r^2)^{-(\nu+p)/2} \tag{4.5}$$

i.e. $(R^2/p) \sim F(p/2, \nu/2)$.

It follows from Theorem 3.6 that the multivariate t-distribution $t_p(\theta, \nu\Lambda;\ \nu)$ has the following stochastic representation

$$X = \theta + R\Lambda^{1/2}U$$

where R has the p.d.f. given by (4.5).

4.3 The Characteristic Function

The characteristic function of the univariate Student t-distribution for odd degrees of freedom has been derived by Fisher and Healy (1956). Ifram (1970) gives a general result for all degrees of freedom but Pestana (1977) points out that this result is incorrect. A series representation of the characteristic function of the multivariate Student t-distribution has been given by Sutradhar (1986).

4.3.1 The Characteristic Function in Terms of the Macdonald Function

In this section we derive the characteristic function of the multivariate t-distribution for all val‿es of degrees of freedom in terms of the well-known special function, namely, the Macdonald function (see e.g. Lebedev, 1965). We list below some standard results on the integral and series representations of the Macdonald function which will be required in the sequel.

The Macdonald Function

The Macdonald function $K_\alpha(x)$, $x > 0$, $\alpha \epsilon \Re$ admits the following integral representation (see, for example, Lebedev 1965, p 119):

$$K_\alpha(x) = \frac{1}{2}\int_{-\infty}^{\infty} e^{-x\cosh w - \alpha w} dw. \tag{4.6}$$

By the substitution $e^w = 2u/x$ in the above integral it readily follows that

$$K_\alpha(x) = \frac{1}{2}\left(\frac{x}{2}\right)^\alpha \int_0^\infty u^{-\alpha-1} exp\left(-u - \frac{x^2}{4u}\right) du, \qquad x > 0,\ \alpha\epsilon\Re. \tag{4.7}$$

The integral given by (4.6) can also be rewritten as follows:

$$\begin{aligned} K_\alpha(x) =& \frac{1}{2}\int_0^\infty e^{-\frac{x}{2}(e^w+e^{-w})} e^{\alpha w} dw \\ &+ \frac{1}{2}\int_0^\infty e^{-\frac{x}{2}(e^w+e^{-w})} e^{-\alpha w} dw \\ =& \frac{1}{2}\int_0^\infty \left(e^{\alpha w} + e^{-\alpha w}\right) e^{-\frac{x}{2}(e^w+e^{-w})} dw. \end{aligned}$$

Hence it follows that

$$K_\alpha(x) = K_{-\alpha}(x), \quad \text{forall} \;\; x > 0,\ \alpha\epsilon\Re.$$

It then follows from (4.7) that for $x > 0, \alpha\epsilon\Re$

$$K_\alpha(x) = K_{-\alpha}(x) = \frac{1}{2}\left(\frac{x}{2}\right)^{-\alpha} \int_0^\infty u^{\alpha-1} \exp\left(-u - \frac{x^2}{4u}\right) du, \qquad x > 0,\ \alpha\epsilon\Re \tag{4.8}$$

(This result will be used in Section 4.5.2.)

Another integral representation of Macdonald function for $x > 0$ and $\alpha > -1/2$ (see, for example, Lebedev, 1965, p 140 and Watson, 1958, p 172) is given by

$$K_\alpha(x) = \left(\frac{2}{x}\right)^\alpha \frac{\Gamma(\alpha+1/2)}{\sqrt{\pi}} \int_0^\infty (1+u^2)^{-(\alpha+1/2)} \cos xu \; du \tag{4.9}$$

(This result will be used in course of proving Theorem 4.1).

In particular, we remark that (see, for example, Tranter 1968, p 19)

$$K_{1/2}(x) = \sqrt{\pi}\,(2x)^{-1/2} \quad \exp(-x) \tag{4.10}$$

(This result will be used in Section 4.3.2).

A series representation of the Macdonald function $K_\alpha(x)$ where $x > 0$ and α a nonnegative integer is as follows (see, for example, Lebedev, 1965, pp 107, 110):

$$K_\alpha(x) = 2^{\alpha-1} \sum_{k=0}^{\alpha-1} \frac{(-1)^k (\alpha-1-k)!}{k!\;4^k}\; x^{2k-\alpha}$$

$$+(-1)^\alpha\, 2^{-(\alpha+1)} \sum_{k=0}^{\infty} \frac{\{\xi(1+k)+\xi(\alpha+1+k)-\ell n(x^2/4)\}}{k!\;(\alpha+k)!\;4^k}\; x^{2k+\alpha} \tag{4.11}$$

where $\xi(x)$ is the digamma function defined by

$$\xi(x) = \Gamma'(x)/\Gamma(x)\;. \tag{4.12}$$

For non-integral positive values of α, a series representation of $K_\alpha(x)$ for $x > 0$ is given by (see, for example, Spainer and Oldham, 1987, chapter 51):

$$K_\alpha(x) = 2^{\alpha-1} \sum_{j=0}^{\infty} \frac{x^{2k-\alpha}}{k!\;(1-\alpha)_k\;4^k} + 2^{-(\alpha+1)} \sum_{k=0}^{\infty} \frac{x^{2k+\alpha}}{k!\;(1+\alpha)_k\;4^k} \tag{4.13}$$

where $(\alpha)_k$ is the Pochammer polynomial defined by (3.12).

Finally we remark that the Macdonald function has been referred to by different authors by various names, such as the Basset function, the Modified Hankel function, the Modified Bessel function of the third kind and for imaginary values of the argument as the Bessel function with imaginary argument. Further details on this topic is to be found in Lebedev (1965, chapter 5) and Spainer and Oldham (1987, chapter 51).

Theorem 4.1 Let Z be a univariate random variable having p.d.f. in (4.3), i.e. $Z \sim t(0, I;\ \nu)$. Then the c.f. of Z is given by

$$\phi_Z(t) = E(\exp(itZ)) = \psi_\nu(|t|) \tag{4.14}$$

with

$$\psi_\nu(|t|) = \frac{|t|^{\nu/2}}{2^{\nu/2-1}\Gamma(\nu/2)}\, K_{\nu/2}(|t|) \tag{4.15}$$

where $K_{\nu/2}(|t|)$ is the Macdonald function with order $\nu/2$ and argument $|t|$.

Proof. The c.f. of Z is given by

$$E(\exp(itZ)) = \frac{\Gamma((\nu+1)/2)}{\sqrt{\pi}\,\Gamma(\nu/2)} \int_{-\infty}^{\infty} (cos\ tz + i\ sin\ tz)(1+z^2)^{-(\nu+1)/2} dz\ .$$

Clearly, from the symmetry of the p.d.f. of Z, the imaginary part of the above integral is zero so that by virtue of (4.9) we immediately have (4.15) and the theorem is proved.

Theorem 4.2 Let the random vector Z have the standard p-variate t-distribution as defined in (4.2) i.e. $Z \sim t_p(0, I;\ \nu)$. Then the c.f. of Z is given by

$$\phi_Z(t) = E(\exp(it'Z)) = \psi_\nu(||t||) \tag{4.16}$$

where $\psi_\nu(\cdot)$ is the characteristic function of a univariate t-distribution as defined in (4.15).

Proof. The characteristic function of Z is given by

$$E(exp(it'Z)) = \int_{\Re^p} exp(it'z)\ f(z)\ dz$$

where $f(z)$ is given by (4.2).

An orthogonal transformation

$$Y = CZ$$

with $y_1 = \frac{t'z}{||t||}$ where C is a $p \times p$ orthogonal matrix, yields

$$E(exp(it'Z)) = K(\nu, p) \int_{\Re^p} exp(i||t||y_1)(1 + y'y)^{-(\nu+p)/2} \, dy.$$

Thus the characteristic function is a function of the scalar $||t|| = (t't)^{1/2}$ i.e.

$$E(exp(it'Z)) = H(||t||) \tag{4.17}$$

for some suitable function $H(\cdot)$.

It follows that the c.f. of the marginal distribution of the first component of Z i.e. Z_1 is then given by

$$E(\exp(it_1 Z_1)) = H(|t_1|) \ . \tag{4.18}$$

On the other hand, successive integration over $z_p, z_{p-1}, \ldots, z_2$ of $f(z_1, z_2, \ldots, z_p)$, given by (4.2), shows that Z_1 is distributed according to the p.d.f. given by (4.3) and hence from Theorem 4.1 we immediately have

$$E(\exp(it_1 Z_1)) = \psi_\nu(|t_1|). \tag{4.19}$$

Therefore from (4.18) and (4.19) we must have

$$H(|t_1|) = \psi_\nu(|t_1|) \text{ for all } t_1.$$

Hence it follows from (4.17) that

$$E(\exp(it'z)) = H(||t||) = \psi_\nu(||t||)$$

and the theorem is proved.

4.3.2 A List of Characteristic Functions of t-Type Distributions

Let X have p.d.f. given by (4.4). Then it follows from Theorem 4.2 that the c.f. of X is given by

$$\phi_X(t) = \exp(it'\theta)\ \psi_\nu(||\Omega^{1/2}t||). \tag{4.20}$$

The c.f's of other t-type distributions follow from it. We now give a list of the various forms of the p.d.f.'s of multivariate t-type distributions along with their characteristic functions expressed in terms of the Macdonald function. We recall that the function $\psi_\nu(\cdot)$ is defined in (4.15).

(i) The Univariate Student t-Distribution

$$\begin{aligned} f(x) =& K(1 + x^2/\nu)^{-(\nu+1)/2},\ x \in \Re,\ \nu > 0\ , \\ K =& \frac{\Gamma((\nu+1)/2)}{\sqrt{(\nu\pi)}\Gamma(\nu/2)}, \\ \phi_X(t) =& \psi_\nu(\sqrt{\nu}|t|)\ . \end{aligned}$$

(ii) The Spherical Student t-Distribution

$$\begin{aligned} f(x) =& K(1 + x'x/\nu)^{-(\nu+p)/2},\ x \in \Re^p\ \nu > 0\ , \\ K =& \frac{\Gamma((\nu+p)/2)}{(\nu\pi)^{-p/2}\Gamma(\nu/2)}\ , \\ \phi_X(t) =& \psi_\nu(\sqrt{\nu}||t||)\ . \end{aligned} \tag{4.21}$$

(iii) The Multivariate Elliptical Student t-Distribution

$$\begin{aligned} f(x) =& K|\Lambda|^{-1/2}(1 + (x-\theta)'(\nu\Lambda)^{-1}(x-\theta))^{-(\nu+p)/2},\ x \in \Re^p,\ \nu > 0\ , \\ K =& \frac{\Gamma((\nu+p)/2)}{(\nu\pi)^{-p/2}\Gamma(\nu/2)} \\ \phi_X(t) =& exp(it'\theta)\ \psi_\nu(||(\nu\Lambda)^{1/2}t||). \end{aligned} \tag{4.22}$$

(iv) The Multivariate Cauchy Distribution

$$\begin{aligned} f(x) =& K|\Lambda|^{-1/2}(1+(x-\theta)'\Lambda^{-1}(x-\theta))^{-(p+1)/2}, \; x \in \Re^p , \\ K =& \frac{\Gamma((p+1)/2)}{\pi^{(p+1)}} , \\ \phi_X(t) =& \exp(it'\theta)\psi_1(||\Lambda^{1/2}t||) \\ =& \exp(it'\theta)\exp(-||\Lambda^{1/2}t||) \qquad \text{by (4.10)} \end{aligned}$$

(cf. Press, 1982, p 175).

(v) Mean Variance Representation of the Multivariate t-Distribution (Sutradhar, 1990).

$$\begin{aligned} f(x) =& K|\Lambda|^{-1/2}(\nu-2+(x-\theta)'\Lambda^{-1}(x-\theta))^{-(\nu+p)/2}, \; x \in \Re^p, \nu > 2, \\ K =& \frac{(\nu-2)^{\nu/2}\Gamma((\nu+p)/2)}{\pi^{-p/2}\Gamma(\nu/2)} \\ \phi_X(t) =& \exp(it'\theta) \; \psi_\nu(\sqrt{\nu-2}||\Lambda^{1/2}t||). \end{aligned}$$

(vi) The Pearson Type VII Distribution (cf. Johnson and Kotz, 1970, p 114)

$$\begin{aligned} f(x) =& K|\Lambda|^{-1/2}(\beta+(x-\theta)'\Lambda^{-1}(x-\theta))^{-m}, x \in \Re^p, m \geq p/2, \beta > 0 , \\ K =& \frac{\beta^{m-p/2}\Gamma(m/2)}{\pi^{-p/2}\Gamma(m-p/2)}, \\ \phi_X(t) =& \exp(it'\theta) \; \psi_{2m-p}(\sqrt{\beta}||\Lambda^{1/2}t||). \end{aligned}$$

(vii) Multivariate t-Type Distributions

$$\begin{aligned} f(x) =& K|\beta\Lambda|^{-1/2}(1+(x-\theta)'(\beta\Lambda)^{-1}(x-\theta))^{-(\nu+p)/2}, \; x \in \Re^p, \; \nu, \beta > 0 , \\ K =& \frac{\Gamma((\nu+p)/2)}{\pi^{-p/2}\Gamma(\nu/2)}, \\ \phi_X(t) =& \exp(it'\theta) \; \psi_\nu(\sqrt{\beta}||\Lambda^{1/2}t||). \end{aligned}$$

The parameter β is not necessarily a function of ν (cf. Rao, 1973, pp. 169-170).

4.4 Series Representation of the Characteristic Function

We now indicate series representation, based on the Macdonald function, of the c.f. of the model given by (4.4).

Case (i) positive even ν

From (4.20) and (4.11) we have

$$\phi_X(t) = exp(it'\theta)\left[\sum_{k=0}^{\nu/2-1} C_1(k)\ ||\Omega^{1/2}t||^{2k} + \sum_{k=0}^{\infty} C_2(k)\ ||\Omega^{1/2}t||^{\nu+2k} + \sum_{j=0}^{\infty} C_3(k)(-ln(||\Omega^{1/2}t||))\ ||\Omega^{1/2}t||^{\nu+2k}\right]$$

where

$$C_1(k) = \frac{(-1)^k(\nu/2-1-k)!}{(\nu/2-1)!\ k!\ 4^k},$$
$$C_2(k) = \frac{(\xi(1+k)+\xi(\nu/2+1+k)+ln4)}{2^\nu(\nu/2-1)!\ k!\ (\nu/2+k)!\ 4^k},$$
$$C_3(k) = (2^{\nu-1}(\nu/2-1)!\ (\nu/2+k)!\ 4^k)^{-1},$$

and

$$\xi(x) = \Gamma'(x)/\Gamma(x)\ .$$

Case (ii) positive non-even ν

From (4.20) and (4.13) it is readily seen that

$$\phi_X(t) = \exp(it'\theta)\sum_{k=0}^{\infty}\left[D_1(k)\ ||\Omega^{1/2}t||^{2k} + D_2(k)\ ||\Omega^{1/2}t||^{\nu+2k}\right]$$

where

$$D_1(k) = (k!\ 4^k\ (1-\nu/2)_k)^{-1}$$

and

$$D_2(k) = \Gamma(-\nu/2)(2^\nu\Gamma(\nu/2)\ k!\ 4^k(1+\nu/2)_k)^{-1}$$

where $(a)_k$ is the Pochammer polynomial defined in (3.12).

4.5 On Moments of the Multivariate t-Distribution

Let X have the multivariate t distribution $t_p(\theta, \nu\Lambda;\ \nu)$ given by (4.1). Then it follows from Theorem 3.6 that

$$E(X) = \theta,$$
$$Cov(X) = \frac{\nu\Lambda}{\nu - 2}.$$

We now calculate product moments of the spherical t-distribution $t_p(\theta, \nu I;\ \nu)$ given by (4.21) by three different methods, namely, Series Representation, Integral Representation of the Macdonald Function and Stochastic Representation.

4.5.1 Calculation of Product Moments by Series Representation

Let $Y \sim t_p(0, \nu I;\ \nu)$ so that Y has the spherical t-distribution with density given by (4.21). Then it follows from section 4.4 that the characteristic function of Y is given by the following series representations:

Case (i) positive even ν

$$\phi_Y(t) = \sum_{k=0}^{\nu/2-1} C_1(k)\ \nu^k (t't)^k + \sum_{k=0}^{\infty} C_2(k)\ \nu^{\nu/2+k}(t't)^{\nu/2+k} + C_3(k)(-ln\sqrt{\nu}(t't)^{1/2})\ \nu^{\nu/2+k}(t't)^{\nu/2+k} \qquad (4.23)$$

where $C_i(k)'s(i = 1,2,3)$ are defined in section 4.4

Case (ii) positive noneven ν

$$\phi_Y(t) = \sum_{k=0}^{\infty} D_1(k)\ \nu^k (t't)^k + \sum_{k=0}^{\infty} D_2(k)\ \nu^{\nu/2+k}(t't)^{\nu/2+k} \qquad (4.24)$$

where $D_i(k)'s(i = 1,2)$ are defined in section 4.4.

To calculate $E(Y_1^2)$ for positive even ν, set $t_2 = t_3 = \ldots = t_p = 0$ in (4.23) so that we have

$$\phi_Y(t) = \sum_{k=0}^{\nu/2-1} C_1(k)\nu^k t_1^{2k} + \sum_{k=0}^{\infty} C_2(k)\ \nu^{\nu/2+k} t_1^{\nu+2k}$$
$$+C_3(k)(-ln\sqrt{\nu}t_1)\nu^{\nu/2+k} t_1^{\nu+2k}$$

Therefore

$$i^2 E(Y_1^2) = \frac{d^2}{dt_1^2}\ \phi_Y(t_1, 0, \ldots, 0)|_{t_1=0}$$
$$= \nu\left[2C_1(\nu) + 2\ \nu t_1^2\ C_1(2) + \ldots + \nu^{\nu/2}(\nu-3)_2\ t_1^{\nu-4} C_1(\nu/2-1)\right]$$
$$+\nu^{\nu/2}\left[(\nu-1)_2\ t_1^{\nu-2} C_2(0) + (\nu)_3\ t_1^{\nu} C_2(1) + \ldots\right]$$
$$+\nu^{\nu/2}\left[\{(\nu-1)_2\ t_1^{\nu-1} C_3(0) + (\nu)_3\ t_1^{\nu} C_3(1) + \ldots\}\right.$$
$$+\{(2\nu-1)\ t_1^{\nu-2} C_3(0) + \nu(2\nu+3)\ t_1^{\nu} C_3(1) + \ldots\}$$
$$\left. +\ \{(\nu-1)_2\ t_1^{\nu-2} ln(t_1) C_3(0) + (\nu)_3\ t_1^{\nu}\ ln(t_1) C_3(1) + \ldots\}\right]|_{t_1=0}.$$

Now by the use of the fact that, for $c > 0$

$$\lim_{t_1 \downarrow 0} t^{c_1}\ ln(t_1) = \lim_{t_1 \downarrow 0} \frac{ln(t_1)}{t_1^c}$$
$$= \lim_{t_1 \downarrow 0} \frac{t_1^{-1}}{-c\ t_1^{-(c+1)}}$$
$$= \lim_{t_1 \downarrow 0} \frac{t^{c_1}}{c}$$
$$= 0,$$

we have

$$-E(Y_1^2) = 2\nu C_1(1) \quad \text{for} \quad \nu > 2.$$

Hence

$$E(Y_1^2) = \frac{\nu}{\nu-2} \quad \text{for } \nu > 2.$$

By similar routine calculation it follows from (4.23) that for positive even values of ν

$$E(Y_1^4) = 24\nu^2 C_1(2) = \frac{3\nu^2}{(\nu-2)(\nu-4)} \quad \text{for } \nu > 4,$$

$$E(Y_1^2Y_2^2) = 8\nu^2 C_1(2) = \frac{\nu^2}{(\nu-2)(\nu-4)} \quad \text{for } \nu > 4$$

and

$$E(Y_1^2Y_2^2Y_3^2) = -48\nu^3 C_1(3) = \frac{\nu^3}{(\nu-2)(\nu-4)(\nu-6)} \quad \text{for } \nu > 6.$$

When ν is positive non-even we have, by differentiating (4.24), the same formulae for product moments as above except that C_1 is replaced by D_1.

In fact, it may be proved that for at least one odd m_k $(k = 1, 2, \dots, p)$,

$$E\left(\prod_{k=1}^{p} Y_k^{m_k}\right) = 0$$

and for all even m_k $(k = 1, 2, \dots, p)$ with $m = \sum_{k=1}^{p} m_k$,

$$E\left(\prod_{k=1}^{p} Y_k^{m_k}\right) = \begin{cases} \frac{(m/2)!(-4\nu)^{m/2}C_1(m/2)}{\pi^{p/2}}\{\prod_{k=1}^{p}(\frac{m_k+1}{2})\} & \text{for positive even } \nu(> m), \\ \frac{(m/2)!(-4\nu)^{m/2}D_1(m/2)}{\pi^{p/2}}\{\prod_{k=1}^{p}(\frac{m_k+1}{2})\} & \text{for positive non-even } \nu(> m), \end{cases}$$

where C_1 and D_1 are given in section 4.4.

Finally by setting the values of C_1 and D_1, we obtain the well-known result for the product moment of Spherical t-Distribution is given by

$$E\left(\prod_{k=1}^{p} Y_k^{m_k}\right) = \begin{cases} 0 \text{ if at least one } m_k (k = 1, 2, \dots, p) \text{ is odd}, \\ \nu^{m/2}\frac{\Gamma((\nu-m)/2)}{2^m\Gamma(\nu/2)}\prod_{k=1}^{p}\frac{m_k!}{(m_k/2)!}, \ \nu > m, \\ \quad \text{if all } m_k\text{'s } (k = 1, 2, \dots, p) \text{ are even} \end{cases} \tag{4.25}$$

(cf. Johnson and Kotz, 1972, p 136, Sutradhar, 1986).

4.5.2 Calculation of Product Moments by an Integral Representation of the Macdonald Function

The moments of the t-distribution are easily computed from the integral representation of the Macdonald function given by (4.8). Let the p-dimensional random variable Y have the p.d.f. given by (4.21) i.e. $Y \sim t_p(0, \nu I;\ \nu)$. Then it is easily checked by virtue of Theorem 4.1 and Theorem 4.2 that

$$\phi_Y(t) = \frac{\sqrt{\nu}||t||^{\nu/2}}{2^{\nu/2-1}\Gamma(\nu/2)}K_{\nu/2}(\sqrt{\nu}||t||)\ .$$

From (4.8) we have

$$K_{-\alpha}(x) = \frac{1}{2}(x/2)^{-\alpha} \int_0^\infty u^{\alpha-1} \exp(-u - x^2/(4u))du \ . \qquad (4.26)$$

If then follows from (4.15), (4.21) and (4.26) that the characteristic function of spherical t-distribution can be written as

$$\phi_Y(t) = \frac{(\sqrt{\nu}||t||)^{\nu/2}}{2^{\nu/2-1}\Gamma(\nu/2)} \left[\frac{1}{2} \left(\frac{\sqrt{\nu}||t||}{2} \right)^{-\nu/2} \int_0^\infty u^{\nu/2-1} \exp\left(-u - \frac{\nu||t||^2}{4u}\right) du \right]$$
$$= \int_0^\infty \frac{u^{\nu/2-1}\exp(-u)}{\Gamma(\nu/2)} \left[\prod_{k=1}^{p} \exp\left(\frac{-\nu t_k^2}{4u}\right) \right] du.$$

Thus

$$E\left[\prod_{k=1}^{p} Y_k^{m_k}\right] = i^{-m} \left[\prod_{k=1}^{p} \left(\frac{\partial}{\partial t_k}\right)^{m_k}\right] \phi_Y(t)|_{t=0}$$
$$= \int_0^\infty \frac{u^{\nu/2-1} exp(-u)}{i^m \Gamma(\nu/2)} \prod_{k=1}^{p} \left[\left(\frac{\partial}{\partial t_k}\right)^{m_k} exp\left(\frac{-\nu t_k^2}{4u}\right)\right] |_{t=0} \ du$$

where $m = m_1 + m_2 + \ldots + m_p$.

But $exp(-t_k^2/(4u))$ is the c.f. of $N(0, \ \nu/(2u))$ whose m_k-th moment is well-known to be given by

$$\left(\frac{\partial}{\partial t_k}\right)^{m_k} exp\left(\frac{-\nu t_k^2}{4u}\right) |_{t_k=0}$$

$$= \begin{cases} 0 \text{ if at least one } m_k (k = 1, 2, \ldots, p) \text{ is odd,} \\ \frac{(\nu/(2u))^{m_k/2} m_k!}{2^{m_k/2}(m_k/2)!}, \text{ if all } m_k\text{'s are even } (k = 1, 2, \ldots, p). \end{cases}$$

Hence it follows for at least one odd m_k $(k = 1, 2, \ldots, p)$ that

$$E(Y_1^{m_1} \ Y_2^{m_2} \ \ldots \ Y_p^{m_p}) = 0$$

and for all even m_k $(k = 1, 2, \ldots, p)$ that

$$E\left(\prod_{k=1}^{p} Y_k^{m_k}\right) = \nu^{m/2} \int_0^\infty \frac{u^{(\nu-m)/2-1} exp(-u)}{2^m \Gamma(\nu/2)} \prod_{k=1}^{p} \frac{m_k!}{(m_k/2)!} du \quad \nu > m.$$

Hence it follows that the product moment of the Spherical t-Distribution i.e. $Y \sim t_p(0,\ \nu I;\ \nu)$ is exactly the same as has been obtained in (4.25).

4.5.3 Calculation of Moments through Stochastic Representation

Let Y has spherical t-distribution $t_p(0, \nu I;\ \nu)$. Then we have from section 4.2 that $Y = RU$ where $R = (Y'Y)^{1/2}$ has the p.d.f. given by (4.5) and U has a uniform distribution on the surface of unit sphere in $\Re^p$.

Since the m-th moment of R is given by

$$E(R^m) = \frac{\nu^{m/2}\Gamma((m+p)/2)\Gamma((\nu-m)/2)}{\Gamma(p/2)\Gamma(\nu/2)},$$

it follows from (3.24) that the product moment $E(\prod_{k=1}^{p} Y_k^{m_k})$ is exactly the same as has been obtained in (4.25).

4.6 On Mixture of Normal Representation

We now demonstrate, with the help of the characteristic function, the well-known representation, already mentioned in (2.5), of the multivariate t-distribution as a suitable scale mixture of the multivariate normal distribution with the mixing scale parameter having a chi-square or an inverted Gamma distribution.

Let $X = (X_1, X_2, \ldots, X_p)'$ have the multivariate t-distribution given by (4.22). It then follows from (4.16), (4.22) and (4.26) that the characteristic function of the multivariate t-distribution can be written as

$$\begin{aligned}
\phi_X(t) =& exp(it'\theta)\ \frac{\|(\nu\Lambda)^{1/2}t\|^{\nu/2}}{2^{\nu/2-1}\Gamma(\nu/2)}\ K_{\nu/2}\left(\|(\nu\Lambda)^{1/2}t\|\right) \qquad (4.27)\\
=& exp(it'\theta)\ \frac{\|(\nu\Lambda)^{1/2}t\|^{\nu/2}}{2^{\nu/2-1}\Gamma(\nu/2)}\\
&\times \left[\frac{1}{2}\|(\nu\Lambda)^{1/2}t\|^{-\nu/2} \int_{u=0}^{\infty} u^{\nu/2-1} exp\left(-u - \frac{\|(\nu\Lambda)^{1/2}t\|^2}{4u}\right)\right] du
\end{aligned}$$

$$=exp(it'\theta)\int_{u=0}^{\infty}\frac{u^{\nu/2-1}exp(-u)}{\Gamma(\nu/2)}exp\left(-\frac{||(\nu\Lambda)^{1/2}t||^2}{4u}\right)du$$

$$=exp(it'\theta)\int_{u=0}^{\infty}\frac{u^{\nu/2-1}exp(-u)}{\Gamma(\nu/2)}\int_{\Re^p}exp(it'x)\ (2\pi)^{-p/2}\ |\frac{\nu\Lambda}{2u}|^{-1/2}$$

$$\times exp\left(\frac{-1}{2}(x-\theta)'\left(\frac{\nu\Lambda}{2u}\right)^{-1}(x-\theta)\right)dudx.$$

The last step follows from the identity

$$exp\left(-\frac{||(\nu\Lambda)^{1/2}t||^2}{4u}\right)=\int_{\Re^p}exp(it'x)\ (2\pi)^{-p/2}\ |\frac{\nu\Lambda}{2u}|^{-1/2}$$

$$\times exp\left(\frac{-1}{2}(x-\theta)'\left(\frac{\nu\Lambda}{2u}\right)^{-1}(x-\theta)\right)dx$$

which is the c.f. of $N_p(\theta, \frac{\nu}{2u}\Lambda)$.

Then the substitution $u=\frac{\nu}{2}\tau^{-2}$ yields

$$\phi_X(t)=\int_{\Re^p}exp(it'x)f_\nu(x)dx \tag{4.28}$$

where

$$f_\nu(x)=\int_{\tau^{-2}=0}^{\infty}h_\nu(\tau^{-2})(2\pi)^{-p/2}|\tau^2\Lambda|^{-1/2}$$

$$\times exp\left(\frac{-1}{2}(x-\theta)'(\tau^2\Lambda)^{-1}(x-\theta)\right)d\tau^{-2} \tag{4.29}$$

with

$$h_\nu(\tau^{-2})=\frac{(\tau^{-2})^{\nu/2-1}\ exp(\frac{-\tau^{-2}}{2/\nu})}{\Gamma(\nu/2)(2/\nu)^{\nu/2}}. \tag{4.30}$$

Hence the multivariate t-distribution is the mixture of the multivariate normal distribution $N_p(\theta,\ \tau^2\Lambda)$ with mixing scale parameter τ^2 where τ^{-2} has a gamma distribution given by $G(\nu/2,\ 2/\nu)$.

4.7 On the Limiting Distribution of the Multivariate t-Distribution

As is well-known, the multivariate t-distribution approaches the multivariate normal distribution as ν tends to infinity. It follows from (4.29) that

$$\lim_{\nu\to\infty} f_\nu(x) = \int_{\tau^{-2}=0}^{\infty} \lim_{\nu\to\infty} h_\nu(\tau^{-2})(2\pi)^{-p/2}|\tau^2\Lambda|^{-1/2} \times exp\left(\frac{-1}{2}(x-\theta)'(\tau^2\Lambda)^{-1}(x-\theta)\right) d\tau^{-2} \tag{4.31}$$

But when $\nu \to \infty$ the characteristic function of τ^{-2} is given by

$$\begin{aligned}\lim_{\nu\to\infty} \phi_{\tau^{-2}}(r) &= \lim_{\nu\to\infty} E(exp(ir\tau^{-2}))\\ &= \lim_{\nu\to\infty}(1-\frac{2ir}{\nu})^{-\nu/2}\\ &=\{\lim_{\nu\to\infty}(1-\frac{2ir}{\nu})^{-\nu}\}^{1/2}\\ &=(exp(-2ir))^{1/2}\\ &=e^{-ir}\end{aligned}$$

which is the characteristic function of a degenerate random variable with all the non-zero mass at the point unity.

Hence by virtue of one to one correspondence between c.f. and and p.d.f., if the later exists, it follows that

$$\lim_{\nu\to\infty} h_\nu(\tau^{-2}) = 1 \quad \text{for } \tau^{-2} = 1.$$

Then it follows from (4.29) and (4.31) that

$$\lim_{\nu\to\infty} f_\nu(x) = \frac{|\Lambda|^{-1/2}}{(2\pi)^{p/2}}\, exp\left(\frac{-1}{2}(x-\theta)'\Lambda^{-1}(x-\theta)\right). \tag{4.32}$$

4.8 A Limit Theorem for the Macdonald Function

Ismail (1977) establishes a few limit theorems of Macdonald functions. Some limit theorems of Macdonald functions have found application in Barndorff-Nielsen and Blaesild (1981) in the context of hyperbolic distributions. We prove a limit theorem which involves the Macdonald function.

Theorem 4.3 The ratio of $K_{\nu/2}(\sqrt{\nu}|c|)$ to

$$\frac{2^{\nu/2-1}\Gamma(\nu/2)}{(\sqrt{\nu}|c|)^{\nu/2}}\, exp(-c^2/2)$$

approaches the limit 1 as $\nu \to \infty$, i.e.

$$\lim_{\nu\to\infty}\frac{(\sqrt{\nu}|c|)^{\nu/2}}{2^{\nu/2}\Gamma(\nu/2)}K_{\nu/2}(\sqrt{\nu}|c|) = \frac{1}{2}exp(-c^2/2). \tag{4.33}$$

Proof. It follows from (4.28) and (4.32) that

$$\begin{aligned}\lim_{\nu\to\infty}\phi_X(t) &= \int_{\Re^p} exp(it'x)\lim_{\nu\to\infty} f_\nu(x)dx \\ &= \int_{\Re^p} exp(it'x)\frac{|\Lambda|^{-1/2}}{(2\pi)^{p/2}} \\ &\times exp\left(\frac{-1}{2}(x-\theta)'\Lambda^{-1}(x-\theta)\right)dx \\ &= exp(it'\theta)\; exp\left(-\frac{||\Lambda^{1/2}t||^2}{2}\right).\end{aligned}$$

Then it follows from (4.27) that

$$\lim_{\nu\to\infty}\frac{||(\nu\Lambda)^{1/2}t||^{\nu/2}}{2^{\nu/2-1}\Gamma(\nu/2)}\, K_{\nu/2}(||(\nu\Lambda)^{1/2}t||) = exp\left(-\frac{||\Lambda^{1/2}t||^2}{2}\right).$$

Since this is an identity in $||\Lambda^{1/2}t||$, the theorem is proved.

4.9 On Marginal and Conditional Distributions

The following well-known results follow easily from the representation of the characteristic function of the multivariate t-distribution as given in Theorem 4.2. Let X have the p.d.f. given by (4.4) i.e. $X \sim t_p(\theta, \Omega;\ \nu)$. Then for any non-singular matrix M, the c.f. of the linear combination $V = MX$ is given by

$$E(\exp(it'V)) = \exp(it'M\theta)\ \psi_\nu((t'M\Omega M't)^{1/2}).$$

Hence

$$V = MX \sim t_{p,\nu}(M\theta, M\Omega M').$$

The marginal and conditional distributions are easily seen to have multivariate t-distributions.

The Marginal Distribution

Let us partition X, θ, t and Ω as

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix},\ \theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix},\ t = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix},\ \Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}$$

where $\theta_2, t_2\ \epsilon \Re^q$ $(q < p)$ and Ω_{22} is a $q \times q$ positive definite matrix.

Setting $t_1 = 0$ in the c.f. of X in (4.20) we immediately have

$$E(\exp(it_2'X_2)) = \exp(it_2'\theta_2)\ \psi_\nu((t_2'\Omega_{22}t_2)^{1/2})\ .$$

Hence $X_2 \sim t_q(\theta_2, \Omega_{22};\ \nu)$ with p.d.f. given by

$$f(x_2) = K|\Omega_{22}|^{-1/2}(1 + (x_2 - \theta_2)'(\Omega_{22})^{-1}(x_2 - \theta_2))^{-(\nu+q)/2}, \qquad (4.34)$$

where $x_2\ \epsilon\ \Re^q$, $\nu > 0$ and the normalizing constant K is given by

$$K = \frac{\Gamma((\nu + q)/2)}{\pi^{q/2}\Gamma(\nu/2)}.$$

The Conditional Distribution

In what follows we will need two well-known identities given by (3.32) and (3.33).

It follows from the identity given by (3.33) that

$$\begin{aligned} y'\Omega y =& y_2'\Omega_{22}^{-1}y_2 \\ &+(y_1-\Omega_{12}\Omega_{22}^{-1}y_2)'\Omega_{11.2}^{-1}(y_1-\Omega_{12}\Omega_{22}^{-1}y_2) \\ =&(1+y_2'\Omega_{22}^{-1}y_2) \\ &+\left[1+(y_1-\Omega_{12}\Omega_{22}^{-1}y_2)'V_{11.2}^{-1}(y_1-\Omega_{12}\Omega_{22}^{-1}y_2)\right] \end{aligned}$$

where $Y = X - \theta$ and

$$V_{11.2} = (1 + y_2'\Omega_{22}^{-1}y_2)\Omega_{11.2}$$

with

$$\Omega_{11.2} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}.$$

Then, by the use of the identity in (3.32), the joint density of Y_1 and Y_2 can be written as

$$\begin{aligned} f(y_1, y_2) =& \frac{\Gamma((\nu+p)/2)}{\pi^{p/2}\Gamma(\nu/2)} |\Omega_{22}|^{-1/2}|\Omega_{11.2}|^{-1/2}(1+y_2'\Omega_{22}^{-1}y_2)^{-(\nu+p)/2} \\ &\times\left[1+(y_1-\Omega_{12}\Omega_{22}^{-1}y_2)'V_{11.2}^{-1}(y_1-\Omega_{12}\Omega_{22}^{-1}y_2\right]^{-(\nu+p)/2}. \end{aligned} \qquad (4.35)$$

so that from (4.34) we immediately have

$$\begin{aligned} f(y_1|y_2) =& \frac{\Gamma((\nu+p)/2)}{(\pi)^{(p-q)/2}\Gamma((\nu+q)/2)} |V_{11.2}|^{-1/2} \\ &\times[1+(y_1-\Omega_{12}\Omega_{22}^{-1}y_2)'V_{11.2}^{-1}(y_1-\Omega_{12}\Omega_{22}^{-1}y_2)]^{-(\nu+p)/2}. \end{aligned} \qquad (4.36)$$

Then from the relation $Y = X - \theta$ it is easily verified that

$$(X_1|X_2 = x_2) \sim t_{p-q}(\theta_{1.2}, \Omega_{11.2}^{*};\ \nu+q)$$

where

$$\theta_{1.2} = \theta_1 + \Omega_{12}\Omega_{22}^{-1}(x_2 - \theta_2)$$

$$\Omega^*_{11.2} = (1 + (x_2 - \theta_2)'\Omega_{22}^{-1}(x_2 - \theta_2))\Omega_{11.2}.$$

When $\Omega = \nu\Lambda$ in (4.4), we have the p.d.f. of the elliptical t-distribution given by (4.1). In this case

$$\Omega_{ik} = \nu\Lambda_{ik} \text{ for } i, k = 1, 2,$$

so that

$$\theta_{1.2} = \theta_1 + \nu\Lambda_{12}(\nu\Lambda_{22})^{-1}(x_2 - \theta_2)$$

and

$$\Omega^*_{11.2} = \nu\Lambda^*_{11.2}$$

where

$$\Lambda^*_{11.2} = (1 + (x_2 - \theta_2)'(\nu\Lambda_{22})^{-1}(x_2 - \theta_2))\Lambda_{11.2}$$

with

$$\Lambda_{11.2} = \Lambda_{11} - \Lambda_{12}\Lambda_{22}^{-1}\Lambda_{21}.$$

Then it follows from (4.36) that

$$(X_1|X_2 = x_2) \sim t_{p-q}(\theta_{1.2}, \nu\Lambda^*_{11};\ \nu + q)$$

so that

$$E(X_1|X_2 = x_2) = \theta_{1.2}$$

and

$$Cov(X_1|X_2 = x_2) = \nu\Lambda^*_{11.2}/(\nu + q - 2).$$

CHAPTER 5

IDENTITIES FOR EXPECTATIONS OF GENERALIZED WISHART MATRIX

In a recent paper Muirhead (1986) derives certain useful identities involving expectations taken with respect to the usual Wishart distribution. The present work generalizes the above results by taking expectations with respect to a generalized version of the Wishart distribution, considered by Sutradhar and Ali (1989), based on the multivariate t-model given by (2.3). We also generalize the identities to the case of sum of product matrix based on the multivariate elliptical model given by (2.2).

Let $X_1, X_2, \ldots, X_N$ be independently and identically distributed as $N_p(\mu, \Lambda)$, $\Lambda > 0$, $N > p$ each having p-components. Then it is well-known that the $p \times p$ matrix A of sample sum of squares and sum of products given by

$$A = \sum_{j=1}^{N} (X_j - \bar{X})(X_j - \bar{X})'$$

has the Wishart distribution $\mathcal{W}(n, \Lambda)$ with the probability density function given by

$$w(A) = \frac{|\Lambda|^{n/2}|A|^{(n-p-1)/2}}{2^{np/2}\Gamma_p(n/2)} \exp\left(-\frac{1}{2}tr(\Lambda^{-1}A)\right), \ A > 0, \ n = N - 1 \geq p \qquad (5.1)$$

where

$$\Gamma_p(\alpha) = \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma\left(\alpha - \frac{1}{2}(i-1)\right), \ \alpha > \frac{p-1}{2}. \qquad (5.2)$$

Sharma and Krishnamoorthy (1986), for the case $\Lambda = I_p$, proves that the identity

$$E\{(trA)^2 trA^\alpha\} = (np + 2 + 2\alpha)E\{trA\ tr(A^\alpha)\}$$

holds for all α for which the expectation exists. Efron and Morris (1976) obtains the following identity in the context of decision-theoretic estimation of Σ:

$$E_w\{tr(\Lambda^{-1}A)/trA\} = (np - 2)E(1/trA) .$$

Muirhead (1986) gives a generalized version of the above mentioned identities obtained by Sharma and Krishnamoorthy (1986) and Efron and Morris (1976).

5.1 Identities for Expectations of the Wishart Matrix Based on the Multivariate t-Model

In this section we consider a generalization of Muirhead's result by taking expectation with respect to a generalized Wishart distribution based on the multivariate t-model rather than the usual normality assumption. More precisely, in what follows, we assume that the p-dimensional random vectors $X_1, X_2, \ldots, X_N$ are identically distributed having the multivariate t-model given by (2.3). It may be noted that they are independently distributed only when $\nu \to \infty$ and in this special case we have the usual multivariate normal model given by (2.1) where $X_1, X_2, \ldots, X_N$ are independently and identically distributed according to $N_p(\theta, \Lambda)$.

Sutradhar and Ali (1989) show that when $X_1, X_2, \ldots, X_N$ have a joint distribution given by (2.3), the sample sum of products matrix $A = \sum_{j=1}^{N}(X_j - \overline{X})(X_j - \overline{X})'$ has the density

$$g(A) = C(\nu, p, n)|\Lambda|^{-n/2}|A|^{(n-p-1)/2}(\nu + tr(\Lambda^{-1}A))^{-(\nu+np)/2} \qquad (5.3)$$

for $A > 0, \Lambda > 0, n \geq p$, and

$$C(\nu, p, n) = \frac{\nu^{\nu/2}\Gamma\{(\nu + np)/2\}}{\Gamma(\nu/2)\Gamma_p(n/2)} \tag{5.4}$$

where $\Gamma_p(n/2)$ is the generalized gamma function defined by (5.2).

In view of the mixture representation of the multivariate t-model (see (2.5)) it is immediate that

$$A|\tau \sim W(n, \tau^2\Lambda)$$

where $\tau^{-2} \sim G(\nu/2, 2/\nu)$.

In this chapter we derive some identities involving expectations with respect to the density given by (5.3). The main result of this chapter is presented in Theorem 5.1 . But first we have Lemma 5.1 which was originally proved by direct integration. The following elegant version of the proof of Lemma 5.1 is based on the mixture representation of multivariate t-model and has been suggested by Professor M.S. Srivastava.

Lemma 5.1 Let the sum of products matrix (Wishart matrix) A have the p.d.f. given by (5.3). Then the rth moment of $|A|$ is given by

$$E_s(|A|^r) = \frac{\nu^{pr}\Gamma(\nu/2 - rp)}{\Gamma(\nu/2)} \frac{\Gamma_p(n/2 + r)}{\Gamma_p(n/2)}|\Lambda|^r$$

for $\nu > 2rp$.

Proof. If $A \sim W(n, \Lambda)$, it is well known that

$$E_w(|A|^r) = 2^{pr}\frac{\Gamma_p(n/2 + r)}{\Gamma_p(n/2)}\ |\Lambda|^r$$

(Muirhead, 1986).

Hence it follows that for any integer r

$$\begin{aligned} E_s(|A|^r) =& E[E(|A|^r|\tau)] \\ =& E\left[2^{pr}\ \frac{\Gamma_p(n/2 + r)}{\Gamma_p(n/2)}\ |\tau^2\Lambda|^r\right] \\ =& 2^{pr}\frac{\Gamma_p(n/2 + r)}{\Gamma_p(n/2)}\ |\Lambda|^r E(\tau^{2pr}). \end{aligned}$$

The proof is then completed by noting that for any integer r

$$E(\tau^r) = \left(\frac{\nu}{2}\right)^{r/2} \frac{\Gamma(\nu/2 - r/2)}{\Gamma(\nu/2)}, \quad \nu > r. \tag{5.5}$$

Theorem 5.1 Suppose that A has the density given by (5.3). Let $h(A)$ be a real-valued measurable function of A such that the function $\psi(t,\ A) = h(t^{-1}A)$ for $t > 0$ and $\psi'(t,\ A) = \frac{\partial}{\partial t}\psi(t,\ A)$ exists at $t = 1$. Then

$$E_g\left[\frac{h(A)\ tr(\Lambda^{-1}A)}{(\nu + tr(\Lambda^{-1}A))/(\nu + np)}\right] = np\ E_g\{h(A)\} - 2E_g\{\psi'(1,\ A)\} \tag{5.6}$$

where E_g stands for the expectation over the density given by (5.3) provided the expectations involved exist.

Proof. For $t > 0$, consider the function $\xi(t)$ given by

$$\xi(t) = C(\nu,p,n)|\Lambda|^{-n/2} \int_{A>0} h(A)|A|^{(n-p-1)/2}\ t^{np/2}(\nu + t\ tr(\Lambda^{-1}A))^{-(\nu+np)/2} dA\ . \tag{5.7}$$

It is readily seen that $\xi(1) = E_g\{h(A)\}$. Clearly

$$\begin{aligned}
\xi'(t) &= \frac{d}{dt}\xi(t) \\
&= C(\nu,p,n)|\Lambda|^{-n/2} \int_{A>0} h(A)|A|^{(n-p-1)/2} \\
&\times \Big[\frac{np}{2} t^{(np-2)/2}(\nu + t\ tr(\Lambda^{-1}A))^{-(\nu+np)/2} \\
&+ t^{np/2}\{-(\nu+np)/2\}(\nu + t\ tr(\Lambda^{-1}A))^{-(\nu+np+2)/2} tr(\Lambda^{-1}A)\Big]\ dA\ .
\end{aligned}$$

Differentiation of (5.7) with respect to t is justified by virtue of dominated convergence theorem. It then follows that

$$\begin{aligned}
\xi'(1) &= (np/2)E_g\{h(A)\} - \{(\nu+np)/2\}C(\nu,p,n)|\Lambda|^{-n/2} \\
&\int_{A>0} \frac{h(A)tr\Lambda^{-1}A}{\nu + tr(\Lambda^{-1}A)}(\nu + tr\Lambda^{-1}A)^{-(\nu+np)/2} dA \\
&= (np/2)E_g\{h(A)\} - \frac{1}{2}\ E_g\left[\frac{h(A)\ tr(\Lambda^{-1}A)}{(\nu + tr(\Lambda^{-1}A))/(\nu+np)}\right]\ .
\end{aligned} \tag{5.8}$$

The transformation

$$B = tA$$

in (5.7) yields

$$\begin{aligned}\xi(t) &= C(\nu,p,n)|\Lambda|^{-n/2}\int_{B>0} h(t^{-1}B)|t^{-1}B|^{(n-p-1)/2}t^{np/2} \\ &\times (\nu + tr\Lambda^{-1}B)^{-(\nu+np)/2}t^{-p(p+1)/2}dB \\ &= C(\nu,p,n)|\Lambda|^{-n/2}\int_{B>0} \psi(t,\ B)|B|^{(n-p-1)/2} \\ &\times (\nu + tr(\Lambda^{-1}B))^{-(\nu+np)/2}dB\ .\end{aligned}$$

Differentiation on both sides with respect to t gives us

$$\xi'(t) = C(\nu,p,n)|\Lambda|^{-n/2}\int_{B>0} \psi'(t,\ B)|B|^{(n-p-1)/2}(\nu + tr(\Lambda^{-1}B))^{-(\nu+np)/2}dB\ . \tag{5.9}$$

Since $B = A$ for $t = 1$, it follows from (5.9) that

$$\xi'(1) = E_s\{\psi'(1,\ A)\}\ . \tag{5.10}$$

Finally from (5.8) and (5.10) it is readily seen that

$$E_s\{\psi'(1,\ A)\} = \frac{np}{2}E_s\{h(A)\} - \frac{1}{2}E_s\left[\frac{h(A)tr(\Lambda^{-1}A)}{(\nu + tr(\Lambda^{-1}A))/(\nu+np)}\right]$$

so that

$$E_s\left[\frac{h(A)tr(\Lambda^{-1}A)}{(\nu + tr(\Lambda^{-1}A))/(\nu+np)}\right] = np\ E_s\{h(A)\} - 2E_s\{\psi'(1,\ A)\}$$

and the theorem is proved.

It has been stated earlier that by letting $\nu \to \infty$, the multivariate t-model given by (2.3) reduces to the multivariate normal model given by (2.1) so that the generalized density of A based on the multivariate t-model reduces to the usual

Wishart density in (5.1). Thus, by letting $\nu \to \infty$, in (5.6) we have the following theorem originally due to Muirhead (1986).

Theorem 5.2 Suppose that A has a density function given by (5.1). Let $h(A)$ be a real valued measurable function of A such that the function $\psi(t,\ A) = h(t^{-1}A)$ for $t > 0$ and suppose that $\psi'(t,\ A) = \frac{\partial}{\partial t}\{\psi(t,\ A)\}$ exists at $t = 1$. Then

$$E_w\{h(A)tr(\Lambda^{-1}A)\} = npE_w\{h(A)\} - 2E_w\{\psi'(1,\ A)\}$$

where E_w is the expectation over the usual Wishart density given by (5.1).

Another special case of Theorem 5.1 is presented below in the form of a corollary.

Corollary 5.1 Let W have the p.d.f. given by (5.3). Further for $x > 0$, let $h(xA) = x^l h(A)$ for some real l. Then

$$E_g\left[\frac{h(A)tr(\Lambda^{-1}A)}{(\nu + tr(\Lambda^{-1}A))/(\nu + np)}\right] = (np + 2l)E_g\{h(A)\}\ . \tag{5.11}$$

Proof. Using $h(xA) = x^l\ h(A)$ in (5.6), we have $\psi(t; A) = t^{-l}h(A)$. Now differentiating with respect to t and setting $t = 1$ we get

$$\psi'(1; A) = -lh(A)\ . \tag{5.12}$$

The identity in (5.11) readily follows from (5.6) and (5.12).

We remark that when $\nu \to \infty$ in (5.11) we obtain the following result originally due to Muirhead (1986):

Corollary 5.2 Let A has the usual Wishart density given by (5.1). Further for $x > 0$, let $h(xA) = x^l h(A)$ for some real l. Then

$$E_w\{h(A)tr(\Lambda^{-1}A)\} = (np + 2l)E_w\{h(A)\}\ .$$

The identity obtained by Sharma and Krishnamoorthy (1986) follows from Corollary 5.2 by setting $\Lambda = I$ and $h(A) = tr(A)tr(A^{\alpha})$ so that $l = \alpha + 1$. Also the identity used by Efron and Morris (1976) follows from Corollary 5.2 by setting $h(A) = (trA)^{-1}$ so that $l = -1$.

5.2 Identities for Expectations of the Wishart Matrix Based on the Elliptical Model

The identities developed so far involves the sum of product matrix (Wishart matrix) based on the multivariate t-model given by (2.3). The multivariate t-model is a member of the multivariate elliptical model given by (2.2). In the following theorem we generalize Theorem 5.1 for the multivariate elliptical model.

Theorem 5.3 (Sutradhar and Ali, 1989) Suppose that X_j's $j = 1, 2, \ldots, N$ have the elliptical model given by (2.2). Then the p.d.f of the sum of product matrix $A = \sum_{j=1}^{N}(X_j - \overline{X})(X_j - \overline{X})'$, is given by

$$f(A) = C_g(p,n)|\Lambda|^{-n/2}|A|^{(n-p-1)/2}g_{p,n}(tr(\Lambda^{-1}A)), \quad n = N-1 \geq p, \quad A > 0$$

where

$$g_{p,n}\left(\sum_{j=1}^{N} z_j' z_j\right) \propto \int_{\Re^p} g\left(\sum_{j=1}^{N} z_j' z_j\right) dz_N$$

and the normalizing constant $C_g(p,n)$ is given by

$$C_g(p,n) = \pi^{np/2}K_g(p,\ n)/\Gamma_p(n/2)$$

with

$$\Gamma_p(n/2) = \pi^{p(p-1)/4}\prod_{i=1}^{p}\Gamma\{(n+1-i)/2\}.$$

Theorem 5.4 Consider the p.d.f. of A given by Theorem 5.3. Let $h(A)$ be a real-valued measurable function of A such that the function $\psi(t,\ A) = h(t^{-1}A)$ for

$t > 0$ and $\psi'(t,\ A) = \frac{\partial}{\partial t}\{\psi(t,\ A)\}$ exists at $t = 1$. Then

$$
\begin{aligned}
&- 2C_g(p,n)|\Lambda|^{-n/2} \int_{A>0} h(A)|A|^{(n-p-1)/2} \frac{\partial}{\partial t} g_{p,n}(t\ tr(\Lambda^{-1}A))|_{t=1} dA \\
&= npE_f\{h(A)\} - 2E_f\{\psi'(1,\ A)\}
\end{aligned}
$$

where E_f is the expectation over the generalized Wishart density given by Theorem 5.3 provided the expectations involved exist.

Proof. For $t > 0$, consider the function $\xi(t)$ given by

$$
\xi(t) = C_g(p,n)|\Lambda|^{-n/2} \int_{A>0} h(A)|A|^{(n-p-1)/2}\ g_{p,n}(t\ tr(\Lambda^{-1}A))dA\ . \qquad (5.13)
$$

It is readily seen that $\xi(1) = E_g\{h(A)\}$. Clearly

$$
\begin{aligned}
\xi'(t) &= \frac{d}{dt}\xi(t) \\
&= C_g(p,n)|\Lambda|^{-n/2} \int_{A>0} h(A)|A|^{(n-p-1)/2} \\
&\times \left[\frac{np}{2} t^{(np-2)/2} g_{p,n}(t\ tr(\Lambda^{-1}A)) + t^{np/2} \frac{\partial}{\partial t} g_{p,n}(t\ tr(\Lambda^{-1}A))\right] dA\ .
\end{aligned}
$$

Differentiation of (5.13) with respect to t is justified by virtue of dominated convergence theorem. It follows that

$$
\begin{aligned}
\xi'(1) = &\frac{np}{2} E_f\{h(A)\} \\
&+ C_g(p,n)|\Lambda|^{-n/2} \int_{A>0} h(A)|A|^{(n-p-1)/2} \\
&\times \{\frac{\partial}{\partial t} g_{p,n}(tr(\Lambda^{-1}A))\}|_{t=1} dA\ .
\end{aligned}
$$

The transformation

$$
B = tA
$$

in (5.13) yields

$$
\xi(t) = C_g(p,n)|\Lambda|^{-n/2} \int_{B>0} h(t^{-1}B)|B|^{(n-p-1)/2} g_{p,n}(tr(\Lambda^{-1}A))dB.
$$

Differentiation on both sides with respect to t and by the use of the fact that $B = A$ for $t = 1$, we obtain

$$\xi'(1) = E_f\{\psi'(1,\ A)\}\ .$$

The proof is completed by equating two expressions for $\xi'(1)$.

We remark that when

$$g_{p,n}(tr(\Lambda^{-1}A)) = (\nu + tr(\Lambda^{-1}A))^{-(\nu+np)/2}$$

we have Theorem 5.1.

5.3 Some Special Cases

In what follows we assume that α stands for a real number while k and r stand for positive integers. The identities in this section involves the Pochammer polynomial symbol defined by (3.12). We also refer to the following result due to Sutradhar and Ali (1989):

$$E_g(A) = \frac{\nu n}{\nu - 2}\Lambda\ . \tag{5.14}$$

Case (i). Set $h(A) = \{tr(\Lambda^{-1}A)\}^{\alpha-1}$ in (5.11), so that l in (5.11) is given by $l = \alpha - 1$. We then have

$$E_g\left[\frac{(tr(\Lambda^{-1}A))^{\alpha}}{(\nu + tr(\Lambda^{-1}A))/(\nu + np)}\right] = \{np + 2(\alpha - 1)\}\, E_g\left[\{tr(\Lambda^{-1}A)\}^{\alpha-1}\right]. \tag{5.15}$$

Set $\alpha = 1$ in (5.15). It then follows that

$$E_g\left[\frac{tr(\Lambda^{-1}A)}{(\nu + tr(\Lambda^{-1}A))/(\nu + np)}\right] = np\ . \tag{5.16}$$

When we set $\alpha = 2$ in (5.15) and use (5.14) we readily obtain for $\nu > 2$,

$$E_g\left[\frac{\{tr(\Lambda^{-1}A)\}^2}{(\nu + tr(\Lambda^{-1}A))/(\nu + np)}\right] = \frac{\nu np}{\nu - 2}(np + 2)\ . \tag{5.17}$$

Letting $\nu \to \infty$ in (5.15), we obtain the recurrence relation

$$E_w\left[(tr(\Lambda^{-1}A))^\alpha\right] = \{2(np/2+\alpha-1)\}\, E_w\left[\{tr(\Lambda^{-1}A)\}^{\alpha-1}\right] . \tag{5.18}$$

Hence for any integer $k > 1$, setting $\alpha = k, k-1, \ldots, 2$ successively in (5.18), we obtain

$$E_w\left[(tr\Lambda^{-1}A)^k\right] = 2^{k-1}\left\{\prod_{i=k-1}^{1}(\frac{np}{2}+i)\right\} E_w[tr(\Lambda^{-1}A)] .$$

Next setting $\alpha = 1$ in (5.18) we finally obtain for $k \geq 1$,

$$E_w\left[\{tr(\Lambda^{-1}A)\}^k\right] = 2^k(np/2)_k \tag{5.19}$$

(cf. identity (7) of Muirhead, 1986).

The recurrence relation given by (5.18) can be rewritten as

$$E_w\left[\{tr(\Lambda^{-1}A)\}^{\alpha-1}\right] = \frac{-1/2}{(-np/2+1-\alpha)}\, E_w\left[\{tr(\Lambda^{-1}A)\}^\alpha\right] . \tag{5.20}$$

Setting $\alpha = -k+1, -k+2, \ldots, -1$ successively in (5.20) it is readily seen that

$$E_w\left[\{tr(\Lambda^{-1}A)\}^{-k}\right] = \frac{(-1/2)^{k-1}}{\prod_{i=k}^{2}(-np/2+i)} E_w\left[\{tr(\Lambda^{-1}A)\}^{-1}\right]$$

for all $k > 2$.

Finally setting $\alpha = 0$ in (5.20), we readily obtain for all $k < np/2$,

$$E_w\left[\{tr(\Lambda^{-1}A)\}^{-k}\right] = \frac{(-1/2)^k}{(-np/2+1)_k} \tag{5.21}$$

(cf. identity (8) of Muirhead, 1986).

Case (ii). Set $h(A) = \{tr(\Lambda^{-1}A)\}^{\alpha-1}trA$ in (5.11) so that l in (5.11) is given by $l = \alpha$. We then have

$$E_g\left[\frac{\{tr(\Lambda^{-1}A)\}^\alpha trA}{(\nu+tr(\Lambda^{-1}A))/(\nu+np)}\right] = (np+2\alpha)E_g\left[\{tr(\Lambda^{-1}A)\}^{\alpha-1}tr(A)\right] . \tag{5.22}$$

Putting $\alpha = 1$ in (5.22) and using (5.14), we obtain for $\nu > 2$,

$$E_g\left[\frac{\{tr(\Lambda^{-1}A)\}tr(A)}{(\nu + tr(\Lambda^{-1}A))/(\nu + np)}\right] = (np+2)\frac{\nu n}{\nu - 2}tr(\Lambda). \tag{5.23}$$

Letting $\nu \to \infty$ in (5.22), we obtain the recurrence relation

$$E_w\left[(tr(\Lambda^{-1}A))^{\alpha} trA\right] = 2\,(np/2+\alpha)\,E_w\left[\{tr(\Lambda^{-1}A)\}^{\alpha-1} trA\right] \ . \tag{5.24}$$

Setting $\alpha = k, k-1, \ldots, 2$ successively in (5.24) we obtain for $k \geq 2$,

$$E_w\left[\{tr(\Lambda^{-1}A)\}^k trA\right] = 2^{k-1}\left\{\prod_{i=k}^{2}(np/2+i)\right\}E_w\left[(tr\Lambda^{-1}A))trA\right] \ .$$

Setting $\alpha = 1$ in (5.24) we finally obtain for $k \geq 1$,

$$E_w\left[\{tr(\Lambda^{-1}A)\}^k trA\right] = n\ 2^k\ (np/2+1)_k \quad tr\Lambda \tag{5.25}$$

(cf. identity (9) of Muirhead, 1986).

The recurrence relation in (5.24) can also be written as

$$E_w\left[\{tr(\Lambda^{-1}A)\}^{\alpha-1} trA\right] = \frac{-1/2}{-np/2-\alpha}E_w\left[\{tr(\Lambda^{-1}A)\}^{\alpha}\ tr(A)\right] \ . \tag{5.26}$$

Setting $\alpha = -k+1, -k+2, \ldots, -1$ successively in (5.26) we obtain for $k > 1$,

$$E_w\left[\{tr(\Lambda^{-1}A)\}^{-k} trA\right] = \frac{(-1/2)^{k-1}}{\prod_{i=k-1}^{1}(-np/2+i)}E_w\left[\{tr\Lambda^{-1}A)\}^{-1} trA\right] \ .$$

Setting $\alpha = 0$ in (5.26) we finally have for $k < (np+2)/2$,

$$E_w\left[\{tr\Lambda^{-1}A)\}^{-k} trA\right] = \frac{(-1/2)^{k-1}}{p\ (-np/2+1)_{k-1}}\ \Lambda. \tag{5.27}$$

Case (iii). Set $h(A) = \{tr(\Lambda^{-1}A)\}^{\alpha-1}tr(A^{-1})$ in (5.11) so that l in (5.11) is given by $l = \alpha - 2$. We then have

$$E_g\left[\frac{\{tr(\Lambda^{-1}A)\}^{\alpha}tr(A^{-1})}{(\nu + tr(\Lambda^{-1}A))/(\nu + np)}\right] = 2\,(np/2+\alpha-1)\,E_g[\{tr(\Lambda^{-1}A)\}^{\alpha-1}tr(A^{-1})] \ . \tag{5.28}$$

Letting $\nu \to \infty$ in (5.28), we have the recurrence relation

$$E_w[\{tr(\Lambda^{-1}A)\}^{\alpha} trA^{-1}] = 2\,(np/2 + \alpha - 1)\,E_w[\{tr(\Lambda^{-1}A)\}^{\alpha-1} tr(A^{-1})] \ . \quad (5.29)$$

Setting $\alpha = k, k-1, \ldots, 1$ successively in the above relation we readily obtain for values of $n > p+1$ and all $k \geq 1$,

$$E_w[\{tr(\Lambda^{-1}A)\}^{k} tr(A^{-1})] = \frac{2^k \ (np/2 - 1)_k}{n - 1 - p} \ tr(\Lambda^{-1}) \quad (5.30)$$

(cf. identity (10) of Muirhead, 1986).

The identity in (5.29) can be written as

$$E_w[\{tr(\Lambda^{-1}A)\}^{\alpha-1} tr(A^{-1})] = \frac{-1/2}{(-np/2 + 2 - \alpha)} E_w[\{tr(\Lambda^{-1}A)\}^{\alpha} \ tr(A^{-1})] \ . \quad (5.31)$$

Setting $\alpha = -k+1, -k+2, \ldots, 0$ successively in the above relation we obtain for $n > p+1$ and $k < (np-2)/2$,

$$E_w[\{tr(\Lambda^{-1}A)\}^{-k} trA^{-1}] = \frac{(-1/2)^k}{(n-p-1)(-np/2+2)_k} . \quad (5.32)$$

Case (iv). Let $h(A) = \{tr(\Lambda^{-1}A)\}^{\alpha-1} tr(\Lambda A^{-1})$ in (5.11) so that l in (5.11) is given by $l = \alpha - 2$. We then have

$$E_g\left[\frac{\{tr(\Lambda^{-1}A)\}^{\alpha} tr(\Lambda A^{-1})}{(\nu + tr(\Lambda^{-1}A))/(\nu + np)}\right] = \{np + 2(\alpha - 2)\}\, E_g[\{tr(\Lambda^{-1}A)\}^{\alpha-1} tr(\Lambda A^{-1})] \ . \quad (5.33)$$

Letting $\nu \to \infty$ in (5.33), we have the recurrence relation

$$E_w[\{tr(\Lambda^{-1}A)\}^{\alpha} tr(\Lambda A^{-1})] = 2(np/2 + \alpha - 2) E_w[\{tr(\Lambda^{-1}A)\}^{\alpha-1} tr(\Lambda A^{-1})] \ . \quad (5.34)$$

Setting $\alpha = k, k-1, \ldots, 1$ successively in the above relation we obtain for $n > p+1$ and $k \geq 1$,

$$E_w[\{tr(\Lambda^{-1}A)\}^{k} tr(\Lambda A^{-1})] = \frac{2^k \ p(np/2 - 1)_k}{n - p - 1} \quad (5.35)$$

(cf. identity (11) of Muirhead, 1986).

The identity in (5.34) can be written as

$$E_w[\{tr(\Lambda^{-1}A)\}^{\alpha-1}tr(\Lambda A^{-1})] = \frac{-1/2}{-np/2+2-\alpha}E_w[\{tr(\Lambda^{-1}A)\}^{\alpha}\ tr(\Lambda^{-1}A)] \ . \tag{5.36}$$

Setting $\alpha = -k+1, -k+2, \ldots, 0$ successively in the above relation we obtain for $n > p+1$ and $k < (np-2)/2$,

$$E_w[\{tr(\Lambda^{-1}A)\}^{-k}tr(\Lambda A^{-1})] = \frac{p\ (-1/2)^k}{(n-p-1)(-np/2+2)_k}. \tag{5.37}$$

Case (v). Set $h(A) = \{tr(\Lambda^{-1}A)\}^{\alpha-1}|A|^r$ in (5.11) so that l in (5.11) is given by $l = \alpha + pr - 1$. We then have

$$E_g\left[\frac{\{tr(\Lambda^{-1}A)\}^{\alpha}|A|^r}{(\nu+tr(\Lambda^{-1}A))/(\nu+np)}\right] = \{np+2(\alpha+pr-1)\}\ E_g[\{tr(\Lambda^{-1}A)\}^{\alpha-1}|A|^r]. \tag{5.38}$$

Setting $\alpha = 1$ in (5.38) and using Lemma 5.1 we obtain for $\nu \geq 2pr$,

$$E_g\left[\frac{(tr(\Lambda^{-1}A))|A|^r}{(\nu+tr(\Lambda^{-1}A))/(\nu+np)}\right] = \frac{(np+2pr)\nu^{pr}\Gamma(\frac{\nu}{2}-pr)\Gamma_p\{(n+2r)/2\}}{\Gamma(\nu/2)\Gamma_p(n/2)}|\Lambda|^r \ . \tag{5.39}$$

Letting $\nu \to \infty$ in (5.38) we have

$$E_w[\{tr(\Lambda^{-1}A)\}^{\alpha}|A|^r] = 2(np/2+\alpha+pr-1)E_w[\{tr(\Lambda^{-1}A)\}^{\alpha-1}|A|^r]. \tag{5.40}$$

Setting $\alpha = k, k-1, \ldots, 1$ successively in the above relation we obtain for $k \geq 1$,

$$E_w[\{tr(\Lambda^{-1}A)\}^{k}|A|^r] = 2^{pr+k}\ (np/2+pr)_k\ \frac{\Gamma_p\{(n+2r)/2\}}{\Gamma_p(n/2)}\ |\Lambda|^r \tag{5.41}$$

(cf. identity (12) of Muirhead, 1986).

The identity in (5.40) can also be written as

$$E_w[\{tr(\Lambda^{-1}A)\}^{\alpha-1}|A|^r] = \frac{-1/2}{-np/2+1-pr-\alpha}E_w[\{tr(\Lambda^{-1}A)\}^{\alpha}|A|^r] \ . \tag{5.42}$$

Setting $\alpha = -k+1, -k+2, \ldots, 0$ successively in the above identity we obtain for $k - pr < np/2$,

$$E_w[\{tr(\Lambda^{-1}A)\}^{-k}|A|^r] = \frac{(-1)^k \; 2^{pr-k}\Gamma_p(n/2+r)}{\Gamma_p(n/2)(-np/2+1-pr)_k} \quad |\Lambda|^r. \tag{5.43}$$

Case (vi). Set $h(A) = A$ in (5.11), so that l in (5.11) is given by $l = 1$. Then it follows from (5.11) that

$$E_g\left[\frac{h(A)tr(\Lambda^{-1}A)}{(\nu + tr(\Lambda^{-1}A))/(\nu + np)}\right] = (np+2)E_g(A).$$

Then by the use of (5.14) we have for $\nu > 2$,

$$E_g\left[\frac{Atr(\Lambda^{-1}A)}{(\nu + tr(\Lambda^{-1}A))/(\nu + np)}\right] = \frac{\nu}{\nu - 2}n(np+2)\Lambda.$$

As $\nu \to \infty$ we have

$$E_w[Atr(\Lambda^{-1}A)] = n(np+2)\Lambda.$$

Case (vii). Let $h(A) = A\{tr(\Lambda^{-1}A)\}^{k-1}(\nu + tr(\Lambda^{-1}A))$. Then it follows from Theorem 5.1 by trivial induction that for $\nu > 2(k+1)$,

$$E_g\left[A\{tr(\Lambda^{-1}A)\}^k\right] = \frac{\nu^{k+1}(np/2)_k}{p\;(\nu/2 - k - 1)_{k+1}}\;\Lambda.$$

As $\nu \to \infty$, we have

$$E_w\left[A\{tr(\Lambda^{-1}A)\}^k\right] = p^{-1}(np/2)_k\;\Lambda.$$

Case (viii). Let $h(A) = A\{tr(\Lambda^{-1}A)\}^{-k}(\nu + tr(\Lambda^{-1}))$. Then it follows from Theorem 5.1 that for $np > 2(k-1)$ and $\nu > k$,

$$E_g\left[A\{tr(\Lambda^{-1}A)\}^k\right] = \{c_1(\nu) + c_2(\nu)\}\;\frac{n\Lambda}{2^{k-1}(np/2)_k}$$

where

$$c_1(\nu) = \nu^{-k/2} \prod_{i=1}^{k/2} (\nu + 2i)$$

and

$$c_2(\nu) = \left(\prod_{i=0}^{k} (\nu - 2i) \right)^{-1} \prod_{i=1}^{k/2} (\nu + 2i).$$

It follows that $c_1(\nu) \to 1$ and $c_2(\nu) \to 0$ as $\nu \to \infty$ and then

$$E_w \left[A\{tr\Lambda^{-1}A)\}^{-k} \right] = \frac{n\Lambda}{2^{k-1}(np/2)_k}.$$

Case (ix). Let $h(A) = \{tr(\Lambda^{-1}A)\}^{k-1} tr(A^2)(\nu + tr(\Lambda^{-1}A))$. Then it follows from Theorem 5.1 by trivial induction that for $\nu > 2(k+2)$,

$$E_g \left[\{tr(\Lambda^{-1}A)\}^k tr(A^2) \right] = \frac{\nu^k (np/2 + 2)_k}{(\nu/2 - k - 2)_k} E_g[tr(A^2)].$$

where $E(trA^2)$ is given by (6.20).

As $\nu \to \infty$ we have

$$E_w \left[\{tr(\Lambda^{-1}A)\}^k tr(A^2) \right] = (np/2 + 2)_k \; E_w[tr(A^2)]$$

where

$$E_w[tr(A^2)] = n[(n+1)\; tr(\Lambda^2) + (tr\; \Lambda)^2]$$

(see Srivastava and Khatri, 1979, p 99).

CHAPTER 6

ESTIMATION OF THE SCALE MATRIX OF THE MULTIVARIATE T-MODEL UNDER A SQUARED ERROR LOSS FUNCTION

6.1 Introduction and Summary

Consider N p-dimensional ($p \geq 2$) random vectors (not necessarily independent) $X_1, X_2, \ldots, X_N$ having a j · ıt probability density function (p.d.f.) given by

$$f(x_1, x_2 \ldots, x_N) = \frac{K(\nu,\ Np)}{|\Lambda|^{N/2}} \left(\nu + \sum_{j=1}^{N} (x_j - \theta)' \Lambda^{-1} (x_j - \theta) \right)^{-(\nu+Np)/2} \tag{6.1}$$

where $x_j = (x_{1j}, x_{2j}, \ldots, x_{pj})'$, θ an unknown vector of location parameters and Λ is an unknown positive definite matrix of scale parameters while the scalar ν is assumed to be a known positive constant. The normalizing constant $K(\nu, Np)$ is given by

$$K(\nu,\ Np) = \frac{\nu^{\nu/2}\Gamma((\nu+Np)/2)}{\pi^{Np/2}\Gamma(\nu/2)}.$$

As mentioned earlier, the model in (6.1) represents the multivariate t-model; it has been considered, among others, by Sutradhar and Ali (1986) in the context of a stock market problem and also in other contexts by Sutradhar and Ali (1989), Dey (1990) and Singh (1991).

We outline below some notations that will be used in this chapter.

Notations

1. Sample sum of products matrix based on the multivariate t-model:

$$A = \sum_{j=1}^{N} (X_j - \bar{X})(X_j - \bar{X})'.$$

2. Usual estimator of the scale matrix Λ of the multivariate t-model:

$$\tilde{\Lambda} = c_1 A,$$

c_1 is any fixed positive constant.

3. Maximum likelihood estimator of Λ :

$$\tilde{\Lambda}_1 = A/(n+1), \quad n = N - 1.$$

4. Unbiased estimator of Λ :

$$\tilde{\Lambda}_2 = (\nu - 2)A/(\nu n).$$

5. Proposed estimator of Λ:

$$\hat{\Lambda} = c_1 A - c_2 |A|^{1/p} I,$$

c_1 is a fixed positive constant ,

$$-\infty < c_2 \le d_0 = \frac{c_1 min\{a_{11}, \ldots, a_{pp}\}}{|A|^{1/p}}.$$

The above condition on c_2 guarantees that the diagonal elements of the proposed estimator $\hat{\Lambda}$ are nonnegative.

6. Sample sum of products matrix based on the multivariate normal model given by (2.1):

$$W = \sum_{j=1}^{N} (X_j - \bar{X})(X_j - \bar{X})'.$$

7. **Usual estimator of the scale matrix Σ of the multivariate normal model:**

$$\tilde{\Sigma} = c_1 W,$$

c_1 **is any fixed positive constant.**

8. **Maximum likelihood estimator of Σ:**

$$\tilde{\Sigma}_1 = W/(n+1), \quad n = N - 1.$$

9. **Unbiased estimator of Σ:**

$$\tilde{\Sigma}_2 = W/n.$$

10. **Proposed estimator of Σ:**

$$\hat{\Sigma} = c_1 W - c_2 |W|^{1/p} I,$$

$$c_1 > 0, \quad -\infty < c_2 \leq d_0 = \frac{c_1 min\{w_{11}, \ldots, w_{pp}\}}{|W|^{1/p}}.$$

The above condition on c_2 guarantees that the diagonal elements of the proposed estimator are nonnegative.

11. **Characteristic roots of A:**

$$m_1, m_2, \ldots, m_p, \quad (m_1 \geq m_2 \geq \ldots \geq m_p).$$

12. **Characteristic roots of Λ :**

$$\xi_1, \xi_2, \ldots, \xi_p, \quad (\xi_1 \geq \xi_2 \geq \cdots \geq \xi_p).$$

13. **Characteristic roots of W:**

$$l_1, l_2, \ldots, l_p, \quad (l_1 \geq l_2 \geq \cdots \geq l_p).$$

14. Characteristic roots of Σ:

$$\alpha_1, \alpha_2, \ldots, \alpha_p, \quad (\alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_p).$$

In this chapter we consider the problem of the estimation of the scale matrix Λ for the multivariate t-model given by (6.1). The present work is primarily motivated by the work of Dey (1988) dealing with the estimation of some functions of the scale matrix of the multivariate normal distribution under a squared error loss function and of Leung (1992) dealing with the scale matrix of the multivariate F distribution under a squared error loss function.

The scale matrix Λ is usually estimated, especially in the multivariate normal case (a special case of the model in 6.1), by multiples of the sum of products matrix A. For example, an unbiased estimator of Λ for the multivariate t-model in (6.1) is given by $\tilde{\Lambda} = (\nu - 2)A/(\nu n)$ where $n = N - 1$ (Anderson and Fang, 1990a, p 208).

The maximum likelihood estimation of Λ has been studied by Anderson, Fang and Hsu (1986) when $(X_1, X_2, \ldots, X_N)$ belongs to a class of elliptical distributions. The maximum likelihood estimator of Λ for the present case is given by $\tilde{\Lambda} = A/N$ (Anderson and Fang, 1990a, p 208). However, the most desirable optimum properties of the maximum likelihood estimator are based on the usual assumption of independence of the component variables $X_1, X_2, \ldots, X_N$. But in the model in (6.1) the components $X_1, X_2, \ldots, X_N$ are independently and identically distributed only when $\nu \to \infty$ and in that case the components are distributed according to $N_p(\theta, \Lambda)$. For finite values of ν the random vectors $X_1, X_2, \ldots, X_N$ are not independently distributed (although uncorrelated); therefore the usual properties of maximum likelihood estimator may not hold for the m.l e. of the scale matrix of the multivariate t-model.

We develop the estimators of Λ in the spirit of Dey (1988) and Leung (1992). As Dey (1988) points out, sample characteristic roots of A tend to be more spread than those of Λ. This suggests that one should take care of the sample characteristic roots by shrinking or expanding them depending on their magnitudes. Dey(1988) has developed estimators of population characteristic roots $\alpha_1, \alpha_2, \ldots, \alpha_p$ of the covariance matrix Σ of the multivariate normal distribution by shrinking sample characteristic roots $l_1, l_2, \ldots, l_p$ towards their geometric mean. He considers estimators of the form

$$\hat{\alpha}_i = c_1 l_i - c_2 (l_1 l_2 \ldots l_p)^{1/p}, \quad i = 1, 2, \ldots, p$$

under the loss function

$$L(\hat{\alpha}, \alpha) = \sum_{i=1}^{p} (\hat{\alpha}_i - \alpha_i)^2.$$

This technique has been exploited throughout this chapter in order to estimate different functions of the scale matrix Λ of the multivariate t-model.

We now consider an estimator of Λ of the form

$$\begin{aligned} \hat{\Lambda} = & c_1 A - c_2 (m_1 m_2 \ldots m_p)^{1/p} I \\ = & c_1 A - c_2 |A|^{1/p} \end{aligned}$$

(cf. Dey, 1988 and Leung, 1992) in order to improve upon the usual estimator $\tilde{\Lambda}$ of Λ of the form $\tilde{\Lambda} = c_1 A$. In section 6.2 we prove that the estimator $\hat{\Lambda}$ dominates the usual estimator $\tilde{\Lambda}$ under certain conditions in the sense of smaller risk as described below.

Consider the squared error loss function (cf. Dey, 1988 and Leung, 1992) given by

$$L(u(A),\ \Lambda) = tr[(u(A) - \Lambda)^2] \tag{6.2}$$

where $u(A)$ is an estimator of Λ. In estimating Λ by $u(A)$, we consider the risk function

$$R(u(A),\ \Lambda) = E[L(u(A),\ \Lambda)]. \tag{6.3}$$

An estimator $u_2(A)$ of Λ will be said to dominate another estimator $u_1(A)$ of Λ if, for all Λ belonging to the class of positive definite matrices, the following inequality holds

$$R(u_2(A),\ \Lambda) \leq R(u_1(A),\ \Lambda),$$

and the inequality

$$R(u_2(A),\ \Lambda) < R(u_1(A),\ \Lambda)$$

holds for at least one Λ.

Outline of the Results Obtained in This Chapter

1. In Theorem 6.1 and Theorem 6.2 (Section 6.2) we prove that the proposed estimator

$$\hat{\Lambda} = c_1 A - c_2 |A|^{1/p} I$$

of the scale matrix of the multivariate t-model dominates the usual estimator $\tilde{\Lambda} = c_1 A$; the risk functions of the estimators have been computed in Theorem 6.3 (Section 6.3). An expression for relative risk function of the estimators has been found in Theorem 6.4.

2. In Theorem 6.5 (section 6.4) we specialize Theorems 6.2, 6.3 and 6.4 by choosing $c_1 = (n+1)^{-1}$. In this case the estimator $\tilde{\Lambda} = c_1 A$ is the maximum likelihood estimator $\tilde{\Lambda}_1$ of Λ and the proposed estimator $\hat{\Lambda}$ is given by

$$\hat{\Lambda}_1 = \frac{1}{n+1} A - c_2 |A|^{1/p} I.$$

Some numerical computations have been performed to compare the proposed estimator $\hat{\Lambda}_1$ with the maximum likelihood estimator $\tilde{\Lambda}_1$ (see Section 6.4.1). Tables showing numerical computations have been given at the end of Chapter 6.

3. In Theorem 6.6 (Section 6.5), we specialize Theorems 6.2, 6.3 and 6.4 by choosing $c_1 = (\nu - 2)/(\nu n)$. In this case the proposed estimator is given by

$$\hat{\Lambda}_2 = \frac{\nu - 2}{\nu n} A - c_2 |A|^{1/p} I.$$

4. In Theorem 6.7 (Section 6.6) we specialize Theorem 6.2 to the case of multivariate normal distribution $N_p(\theta, \Sigma)$ by letting $\nu \to \infty$. In this case the proposed estimator is given by

$$\hat{\Sigma} = c_1 W - c_2 |W|^{1/p} I;$$

the risk functions of the estimators $\tilde{\Sigma}$ and $\hat{\Sigma}$ have been calculated in Theorem 6.8 (Section 6.7). The relative risk function of the estimators has been given in Theorem 6.9.

5. In Theorem 6.10 (Section 6.8), we specialize Theorems 6.7, 6.8 and 6.9 by choosing $c_1 = 1/(n+1)$. In this case the proposed estimator of Σ is given by

$$\hat{\Sigma}_1 = \frac{1}{n+1} W - c_2 |W|^{1/p} I.$$

6. In Theorem 6.11 (Section 6.9), we specialize Theorems 6.7, 6.8 and 6.9 by choosing ($c_1 = 1/n$). In this case the proposed estimator of Σ is given by

$$\hat{\Sigma}_2 = \frac{1}{n} W - c_2 |W|^{1/p} I.$$

7. Consider the vector of the characteristic roots of the scale matrix of the multivariate t-model:

$$\xi = (\xi_1, \xi_2, \ldots, \xi_p).$$

We propose the estimator

$$\hat{\xi} = (\hat{\xi}_1, \hat{\xi}_2, \ldots, \hat{\xi}_p)$$

where

$$\hat{\xi}_i = c_1 m_i - c_2(m_1 m_2 \ldots m_p)^{1/p}, \ (i = 1, 2, \ldots, p)$$

and

$$c_1 > 0, \quad -\infty < c_2 \le d_0 = \frac{c_1 min\{m_1, m_2, \ldots, m_p\}}{(m_1 m_2 \ldots m_p)^{1/p}}.$$

In Section 6.10.1, we prove a theorem that the proposed estimator $\hat{\xi}$ of ξ dominates the usual estimator $\tilde{\xi}$ where

$$\tilde{\xi} = (\tilde{\xi}_1, \tilde{\xi}_2, \ldots, \tilde{\xi}_p)$$

and

$$\tilde{\xi}_i = c_1 m_i, \quad c_1 > 0, \quad (i = 1, 2, \ldots, p).$$

The results of Section 6.10.1 are specialized in Section 6.10.2 to estimate the characteristic roots of the scale matrix Σ of the multivariate normal distribution.

8. In Section 6.11, we prove a theorem that the estimator $\hat{\delta}$ of $\delta = tr(\Lambda)$ given by

$$\hat{\delta} = c_1 tr(A) - c_2 p |A|^{1/p}$$

where

$$-\infty < c_2 \le d_0 = \frac{c_1 tr(A/p)}{|A|^{1/p}},$$

dominates the usual estimator $\tilde{\delta} = c_1 A, \quad c_1 > 0$ follows from Theorem 6.2.

9. In Section 6.12, we prove that the estimator $\hat{\Psi}$ of $\Psi = \Lambda^{-1}$ given by

$$\hat{\Psi} = c_1 A^{-1} - c_2 |A|^{-1} I$$

where

$$-\infty < c_2 \le d_0 = \frac{c_1 A^{-1}}{|A|^{-1/p}},$$

dominates the usual estimator $\tilde{\Psi} = c_1 tr(A^{-1})$.

Some Useful Results on the Sample Sum of Products Matrix

We now state some important results related to the distribution of the sum of products matrix A with p.d.f. given by (5.3).

We recall that the expectation of A is given by

$$E_g(A) = \frac{\nu n\ \Lambda}{\nu - 2}, \quad \nu > 2 \tag{6.4}$$

(Anderson and Fang, 1990a, p 208).

Lemma 6.1 Consider the p.d.f. of A given by (5.3). Then for any real number k and any positive number ν satisfying the conditions $n + 2k \ge 0$ and $\nu \ge 2(kp+1)$, the following result holds:

$$E_g(|A|^k A) = \nu^{kp+1}\ (n/2 + k)\ \frac{\Gamma(\nu/2 - kp - 1)}{\Gamma(\nu/2)}\ \frac{\Gamma_p(n/2+k)}{\Gamma_p(n/2)}\ |\Lambda|^k \Lambda.$$

Proof. It is readily verified that for any real number r

$$\begin{aligned} E_g(|A|^k A^r) &= \int_{A>0} |A|^k A^r\ g(A)dA \\ &= \frac{C(\nu,p,n)}{C(\nu^*,p,n^*)} \left(\frac{\nu^*}{\nu}\right)^{\nu^*/2} |\Lambda|^k \int_{A>0} A^r\ g^*(A)dA, \end{aligned} \tag{6.5}$$

where the p.d.f. $g(A)$ is given by (5.3) while the p.d.f. $g^*(A)$ is given by

$$g^*(A) = C(\nu^*,p,n^*)|\Lambda^*|^{-n^*/2}|A|^{(n^*-p-1)/2}\left(\nu^* + tr((\Lambda^*)^{-1}A)\right)^{-(\nu^*+n^*p/2}, $$

where $A > 0$, $n^* = n + 2k \ge p$, $\nu^* = \nu - 2kp > 0$ and $\Lambda^* = \nu\Lambda/\nu^*$.

Set n, ν and Λ equal to n^*, ν^* and Λ^* respectively in $g(A)$ given by (5.3). Then it is readily verified from (6.4) that

$$E_{g*}(A) = \int_{A>0} A\, g^*(A)dA = \frac{\nu^* n^*}{\nu^* - 2}\Lambda^*$$

so that from (6.5) we immediately have

$$E_g(|A|^k A) = \frac{C(\nu,p,n)}{C(n^*,p,n^*)}\left(\frac{\nu^*}{\nu}\right)^{\nu^*/2}|\Lambda|^k\left(\frac{\nu^* n^*}{\nu^* - 2}\Lambda^*\right).$$

On substitution of $C(\nu,p,n)$ from (5.4), We finally have

$$\begin{aligned} E_g(|A|^k A) =& \frac{\nu^{\nu/2}\Gamma((\nu+np)/2)}{\Gamma(\nu/2)\Gamma_p(n/2)}\frac{\Gamma(\nu^*/2)\Gamma_p(n^*/2)}{(\nu^*)^{\nu^*/2}\Gamma((\nu^*+n^*p)/2)} \\ & \times\left[\left(\frac{\nu^*}{\nu}\right)^{\nu^*/2}|\Lambda|^k\left(\frac{\nu^* n^*}{\nu^* - 2}\right)\frac{\nu\Lambda}{\nu^*}\right] \\ =& \frac{\nu^{kp+1}\Gamma(\nu^*/2-1)}{\Gamma(\nu/2)}\frac{(n^*/2)\Gamma_p(n^*/2)}{\Gamma_p(n/2)}|\Lambda|^k\,\Lambda. \end{aligned}$$

The proof of the lemma is completed by setting $n^* = n + 2k$ and $\nu^* = \nu - 2kp$ in the last expression.

Lemma 6.2 Let A have the p.d.f. given by (5.3). Then for $n > p + 1$,

$$E(A^{-1}) = \frac{\Lambda^{-1}}{n-p-1}.$$

Proof. From (5.3) a straightforward computation shows that the p.d.f. of A^{-1} is given by

$$g(A^{-1}) = C(\nu,p,n)|\Lambda|^{-n/2}|A|^{(n+p+1)/2}\left(\nu + tr(\Lambda^{-1}A)\right)^{-(\nu+np)/2}$$

where $C(\nu,p,n)$ is given by

$$C(\nu,p,n) = \frac{\nu^{\nu/2}\Gamma((\nu+np)/2)}{\Gamma(\nu/2)\Gamma_p(n/2)}.$$

Hence we have

$$
\begin{aligned}
E(A^{-1}) &= \int_{A^{-1}>0} A^{-1} g(A^{-1}) dA^{-1} \\
&= \int_{A^{-1}>0} A^{-1} \frac{\nu^{-np/2}\Gamma((\nu+np)/2)}{\Gamma(\nu/2)\Gamma_p(n/2)} |\Lambda|^{-n/2}|A|^{(n+p+1)/2} \\
&\times \left(1+tr(\Lambda^{-1}A/\nu)\right)^{-(\nu+np)/2} dA^{-1} \\
&= \int_{A^{-1}>0} A^{-1} \frac{\nu^{-np/2}\Gamma((\nu+np)/2)}{\Gamma(\nu/2)\Gamma_p(n/2)} |\Lambda|^{-n/2}|A|^{(n+p+1)/2} \\
&\times \left[\frac{2^{-(\nu+np)/2}}{\Gamma((\nu+np)/2)} \int_{u=0}^{\infty} u^{(\nu+np)/2-1} exp\left(-\frac{u}{2}(1+tr(\Lambda^{-1}A/\nu))\right) du\right] dA^{-1}.
\end{aligned}
$$

where U is a random variable having a gamma distribution

$$
G\left(\frac{\nu+np}{2},\ \frac{2}{1+tr(\Lambda^{-1}A/\nu)}\right).
$$

Then by the use of the transformation $A = \nu U^{-1} W$ with Jacobian

$$
J(A^{-1} \to W^{-1}) = (\nu^{-1}U)^{p(p+1)/2}
$$

it is immediately seen that

$$
\begin{aligned}
E(A^{-1}) =& \nu^{-1} \int_{u=0}^{\infty} \frac{u^{\nu/2} exp(-u/2)}{2^{\nu/2}\Gamma(\nu/2)} \\
&\times \int_{W^{-1}>0} W^{-1} \frac{|\Lambda|^{-n/2}|W|^{(n+p+1)/2}}{2^{np/2}\Gamma_p(n/2)} exp\left(\frac{-1}{2} tr(\Lambda^{-1}W)\right) dW^{-1} \\
=& \nu^{-1} \int_{v=0}^{\infty} v f(v) \int_{W^{-1}>0} W^{-1} f(W^{-1}) dW^{-1} \\
=& \nu^{-1}\ E(V)\ E(W^{-1}) \\
=& E(W^{-1}) \\
=& \Lambda^{-1}/(n-p-1)
\end{aligned}
$$

where $f(v)$ is the p.d.f. of a random variable V having a chi-square distribution with ν degrees of freedom, W is a random matrix having usual Wishart distribution $\mathcal{W}_p(n,\ \Lambda)$ with p.d.f. $f(W)$, while $f(W^{-1})$ is the p.d.f. of W^{-1}.

The last step is given in most standard texts; see e.g. Muirhead (1982, p97).

Lemma 6.3 Suppose that A has the p.d.f. given by (5.3). Then for any real number k, $n+2k > p+1$, $\nu \geq 2(kp+1)$, and $n = N-1$ the following result hold:

$$E(|A|^k A^{-1}) = \frac{2\,\nu^{kp-1}}{n+2k-p-1}\,\frac{\Gamma(\nu/2-kp+1)}{\Gamma(\nu/2)}\,\frac{\Gamma_p(n/2+k)}{\Gamma_p(n/2)}\,|\Lambda|^k\Lambda^{-1}.$$

Proof. The proof follows by the use of Lemma 6.2 in (6.5) with $r = -1$.

A shorter version of the proofs of Lemma 6.1, Lemma 6.2 and Lemma 6.3 based on the mixture representation of the multivariate t-model has been suggested by Professor M.S. Srivastava.

Since $A|\tau \sim W_p(n,\ \tau^2\Lambda)$ where $\tau^{-2} \sim G(\nu/2, 2/\nu)$, it follows from Dey (1988, p 140) that for any any real number k

$$\begin{aligned}
E[|A|^k A] =& E\left[E\left(|A|^k A|\tau\right)\right]\\
=& E\left[2^{kp}\ (n+2k)\ \frac{\Gamma_p(n/2+k)}{\Gamma_p(n/2)}\ |\tau^2\Lambda|^k\ (\tau^2\Lambda)\right]\\
=& 2^{kp}\ (n+2k)\ \frac{\Gamma_p(n/2+k)}{\Gamma_p(n/2)}\ |\Lambda|^k\Lambda\ E\left(\tau^{2kp+2}\right).
\end{aligned}$$

Hence the proof of Lemma 6.1 follows by noting (5.5).

Similarly it also follows from Muirhead (1982, p 97) that

$$\begin{aligned}
E[A^{-1}] =& E\left[E\left(A^{-1}|\tau\right)\right]\\
=& E\left[\frac{(\tau^2\Lambda)^{-1}}{n-p-1}\right]\\
=& \frac{(\Lambda)^{-1}}{n-p-1}\ E(\tau^{-2}).
\end{aligned}$$

Thus the proof of Lemma 6.2 is completed by noting that $E(\tau^{-2}) = 1$.

It also follows from Dey (1988, p 141) that

$$\begin{aligned}
E[|A|^k A^{-1}] =& E\left[E\left(|A|^k A^{-1}|\tau\right)\right]\\
=& E\left[\frac{2^{kp}}{n+2k-p-1}\ \frac{\Gamma_p(n/2+k)}{\Gamma_p(n/2)}\ |\tau^2\Lambda|^k\ (\tau^2\Lambda)^{-1}\right]\\
=& \frac{2^{kp}}{n+2k-p-1}\ \frac{\Gamma_p(n/2+k)}{\Gamma_p(n/2)}\ |\Lambda|^k\ (\Lambda)^{-1}\ \ E\left(\tau^{2kp-2}\right).
\end{aligned}$$

Hence the proof of Lemma 6.3 follows by noting (5.5).

6.2 Estimation of the Scale Matrix of the Multivariate t-Model

Consider the multivariate t-model given by (6.1) with $\nu > 4$. Also consider the following two estimators of Λ:

$$\text{usual estimator } \tilde{\Lambda} = c_1 A, \tag{6.6}$$

$$\text{proposed estimator } \hat{\Lambda} = c_1 A - c_2 |A|^{1/p} I, \tag{6.7}$$

where

$$c_1 > 0, \quad -\infty < c_2 \leq d_0 = \frac{c_1 min\{a_{11}, \ldots, a_{pp}\}}{|A|^{1/p}}. \tag{6.8}$$

In the following theorem we prove that the proposed estimator of the scale matrix of the multivariate t-model dominates the usual estimator under certain conditions.

Theorem 6.1 Consider the multivariate t-model given by (6.1) for $\nu > 4$. Then the proposed estimator

$$\hat{\Lambda} = c_1 A - c_2 |A|^{1/p} I,$$

where c_1 is *a given positive number* while c_2 satisfies the condition

$$-\infty < c_2 \leq d_0 = \frac{c_1 min\{a_{11}, \ldots, a_{pp}\}}{|A|^{1/p}},$$

dominates the usual estimator $\tilde{\Lambda} = c_1 A$ under the squared error loss function given by (6.2) for any c_2 satisfying the conditions stated below:

1. For a given c_1 satisfying

$$c_1 < \frac{\nu - 4}{\nu} \frac{p}{np + 2} \quad (d^* < 0),$$

c_2 satisfies

$$d^* < c_2 < 0. \tag{6.9}$$

2. For a given c_1 satisfying

$$c_1 > \frac{\nu-4}{\nu}\frac{p}{np+2} \quad (d^* > 0),$$

c_2 satisfies the following scheme:

$$(i)\ 0 < c_2 \leq d_0 \quad \text{if } 0 < d_0 < d^*, \tag{6.10a}$$

$$\text{and} \quad (ii)\ 0 < c_2 < d^* \quad \text{if } d^* \leq d_0 < \infty \tag{6.10b}$$

where

$$d^* = \left(c_1\frac{np+2}{p} - \frac{\nu-4}{\nu}\right)\frac{\Gamma_p(n/2+1/p)}{\Gamma_p(n/2+2/p)}. \tag{6.11}$$

Proof. Let us consider the following risk functions:

$$R(\tilde{\Lambda},\ \Lambda) = E[tr(\tilde{\Lambda} - \Lambda)^2]$$

and

$$R(\hat{\Lambda}, \Lambda; c_2) = E[tr(\hat{\Lambda} - \Lambda)^2]$$

with the Risk Difference

$$D(\Lambda,\ c_2) = R(\hat{\Lambda},\ \Lambda; c_2) - R(\tilde{\Lambda},\ \Lambda).$$

Then in order that the proposed estimator $\hat{\Lambda}$ dominates the usual estimator $\tilde{\Lambda}$ it is necessary and sufficient that $D(\Lambda,\ c_2) < 0$.

A simple calculation shows that

$$
\begin{aligned}
D(\Lambda,\ c_2) =& Etr(\hat{\Lambda}-\Lambda)^2 - Etr(\tilde{\Lambda}-\Lambda)^2 \\
=& Etr\left[\hat{\Lambda}^2 - \tilde{\Lambda}^2 - (\hat{\Lambda}-\tilde{\Lambda})\Lambda - \Lambda(\hat{\Lambda}-\tilde{\Lambda})\right] \\
=& Etr\left[(\hat{\Lambda}^2 - \tilde{\Lambda}^2) - 2(\hat{\Lambda}-\tilde{\Lambda})\Lambda\right] \\
=& Etr\left[(-2c_2|A|^{1/p}\tilde{\Lambda} + c_2^2|A|^{2/p}I) - 2\{-c_2|A|^{1/p}I\}\Lambda\right] \\
=& -2c_1c_2\ Etr(|A|^{1/p}A) + 2c_2\ E(|A|^{1/p})tr(\Lambda) + c_2^2\ E(|A|^{2/p})tr(I).
\end{aligned}
$$

It then follows from Lemma 5.1 and Lemma 6.1 that

$$
\begin{aligned}
D(\Lambda,\ c_2) =& -2c_1c_2\ tr\left[\frac{2\nu^2(np+2)}{(\nu-2)(\nu-4)p}\frac{\Gamma_p(n/2+1/p)}{\Gamma_p(n/2)}|\Lambda|^{1/p}\Lambda\right] \\
&+2c_2\ \left[\frac{2\nu}{\nu-2}\frac{\Gamma_p(n/2+2/p)}{\Gamma_p(n/2)}|\Lambda|^{1/p}\right] tr(\Lambda) \\
&+c_2^2 p\left[\frac{4\nu^2}{(\nu-2)(\nu-4)}\frac{\Gamma_p(n/2+2/p)}{\Gamma_p(n/2)}|\Lambda|^{2/p}\right] \\
=& \frac{4\nu^2 c_2 p}{(\nu-2)(\nu-4)}\ \frac{|\Lambda|^{2/p}}{\Gamma_p(n/2)} \\
&\times\left[\left(-c_1\frac{np+2}{p}+\frac{\nu-4}{\nu}\right)\Gamma_p(n/2+1/p)\frac{tr(\Lambda/p)}{|\Lambda|^{1/p}} + c_2\Gamma_p(n/2+2/p)\right].
\end{aligned}
$$

Hence

$$
\begin{aligned}
D(\Lambda,\ c_2) =& \frac{4\nu^2 p}{(\nu-2)(\nu-4)}\frac{\Gamma_p(n/2+2/p)}{\Gamma_p(n/2)}\ |\Lambda|^{2/p} \\
&\times c_2\left(c_2 - \frac{tr(\Lambda/p)}{|\Lambda|^{1/p}}d^*\right) \qquad (6.12)
\end{aligned}
$$

where d^* is given by (6.11).

As stated earlier, in order that $\hat{\Lambda}$ dominates $\tilde{\Lambda}$ it is necessary and sufficient that $D(\Lambda,\ c_2) < 0$. However, $D(\Lambda,\ c_2) < 0$ if and only if

$$
\frac{tr(\Lambda/p)}{|\Lambda|^{1/p}}d^* < c_2 < 0, \quad \text{or} \quad 0 < c_2 < \frac{tr(\Lambda/p)}{|\Lambda|^{1/p}}d^*.
$$

The above conditions involve $tr(\Lambda)$ and $|\Lambda|$ which are unknown quantities. Now let $\xi_1, \xi_2, \ldots, \xi_p$ be the characteristic roots of Λ. Then

$$
tr(\Lambda/p) = (\xi_1 + \xi_2 + \ldots + \xi_p)/p
$$

and

$$|\Lambda|^{1/p} = (\xi_1 \xi_2 \dots \xi_p)^{1/p}.$$

It then follows from the well-known inequality between arithmetic mean and geometric mean that

$$\frac{tr(\Lambda/p)}{|\Lambda|^{1/p}} \geq 1.$$

Hence $\hat{\Lambda}$ dominates $\tilde{\Lambda}$ if

$$d^* < c_2 < 0, \quad \text{or} \quad 0 < c_2 < d^*.$$

Moreover, in order that the proposed estimator $\hat{\Lambda}$ have all non-negative diagonal elements we must also have $c_2 \leq d_0$ where d_0 is given by (6.8). We also note that

$$d^* < 0 \ \text{ if and only if } \ c_1 < \frac{\nu - 4}{\nu} \frac{p}{np + 2},$$

$$\text{while } \ d^* > 0 \ \text{ if and only if } \ c_1 > \frac{\nu - 4}{\nu} \frac{p}{np + 2}.$$

Hence the proposed estimator $\hat{\Lambda}$ dominates the usual estimator $\tilde{\Lambda}$ if c_2 satisfies the following conditions:

1. For a given c_1 satisfying

$$c_1 < \frac{\nu - 4}{\nu} \frac{p}{np + 2},$$

c_2 satisfies

$$d^* < c_2 < 0.$$

2. For a given c_1 satisfying

$$c_1 > \frac{\nu - 4}{\nu} \frac{p}{np + 2},$$

c_2 satisfies the following scheme:

$$(i) \ \ 0 < c_2 \leq d_0 \ \text{ if } 0 < d_0 < d^*,$$

$$\text{and } (ii) \;\; 0 < c_2 < d^* \;\; \text{if } d^* \le d_0 < \infty.$$

Hence the theorem is proved.

Remark 6.1 We note that for the case when

$$c_1 = \frac{\nu - 4}{\nu} \frac{p}{np + 2},$$

we have $d^* = 0$. In this case it is seen from (6.12) that $D(\Lambda, c_2) \ge 0$ so that there exists no proposed estimator $\hat{\Lambda}$ dominating the usual estimator $\tilde{\Lambda}$. However, $D(\Lambda, c_2) = 0$ only if $c_2 = 0$ in which case the two estimators coincide.

Optimal Value of c_2

We now look for narrower range of c_2 (compared to the range of c_2 obtained in Theorem 6.1) in which the risk of the proposed estimator will be much lower than that of the usual estimator. It is seen from (6.12) that for given Λ, the Risk Difference $D(\Lambda, c_2)$ is a polynomial of degree 2 in c_2, and it is minimized at

$$c_2 = \frac{tr(\Lambda/p)}{|\Lambda|^{1/p}} \frac{d^*}{2}.$$

Now by virtue of the inequality between arithmetic mean and geometric mean of the characteristic roots of Λ, it follows that the optimal value of c_2 satisfies the following conditions:

$$(i) \qquad c_2 \le \frac{d^*}{2} \text{ if } d^* < 0.$$

$$\text{and } (ii) \qquad c_2 \ge \frac{d^*}{2} \text{ if } d^* > 0$$

$$i.e. (i) \qquad -\infty < c_2 \le \frac{d^*}{2} \text{ if } d^* < 0 \qquad (6.13)$$

$$\text{and } (ii) \qquad \frac{d^*}{2} \le c_2 \le d_0 \text{ if } d^* > 0 \qquad (6.14).$$

In the light of the above results we have the following version of Theorem 6.1 narrowing down the range of c_2.

Theorem 6.2 Consider the multivariate t-model given by (6.1) for $\nu > 4$. Then the proposed estimator

$$\hat{\Lambda} = c_1 A - c_2 |A|^{1/p} I,$$

where c_1 is a given positive number while c_2 satisfies the condition

$$-\infty < c_2 \leq d_0 = \frac{c_1 min\{a_{11}, \ldots, a_{pp}\}}{|A|^{1/p}},$$

dominates the usual estimator $\tilde{\Lambda} = c_1 A$ under the squared error loss function given by (6.2) for any c_2 satisfying the conditions stated below:

1. For a given c_1 satisfying

$$c_1 < \frac{\nu - 4}{\nu} \frac{p}{np + 2} \quad (i.e.\ d^* < 0),$$

c_2 satisfies

$$d^* < c_2 \leq \frac{d^*}{2}$$

where d^* is given by (6.11).

2. For a given c_1 satisfying

$$c_1 > \frac{\nu - 4}{\nu} \frac{p}{np + 2} \quad (i.e.\ d^* > 0),$$

c_2 satisfies the following scheme:

$$\begin{aligned}
&\text{(i)} \quad c_2 = d_0 \ \text{ for } \ 0 < d_0 \leq \frac{d^*}{2}, \\
&\text{(ii)} \quad \frac{d^*}{2} \leq c_2 \leq d_0 \ \text{ for } \ \frac{d^*}{2} < d_0 < d^*, \\
&\text{(iii)} \quad \frac{d^*}{2} \leq c_2 < d^* \ \text{ for } \ d^* \leq d_0,
\end{aligned}$$

where d^* is defined by (6.11).

Proof. For a given c_1 satisfying

$$c_1 < \frac{\nu - 4}{\nu} \frac{p}{np + 2} \quad (i.e.\ d^* < 0),$$

it follows from (6.9) and (6.13) that the proposed estimator $\hat{\Lambda}$ dominates the usual estimator $\tilde{\Lambda}$ if c_2 satisfies

$$d^* < c_2 < 0 \text{ and } -\infty < c_2 \leq \frac{d^*}{2}$$

$$\text{i.e. } d^* < c_2 \leq \frac{d^*}{2}$$

where d^* is given by (6.11).

It also follows from (6.10a), (6.10b) and (6.14) that for a given c_1 satisfying

$$c_1 > \frac{\nu - 4}{\nu} \frac{p}{np + 2} \quad (i.e.\ d^* > 0),$$

the proposed estimator $\hat{\Lambda}$ dominates the usual estimator $\tilde{\Lambda}$ if c_2 satisfies the following scheme:

$$(i)\ 0 < c_2 \leq d_0 \text{ when } 0 < d_0 \leq \frac{d^*}{2},$$

$$\text{and } (ii)\ 0 < c_2 \leq d_0 \text{ and } \frac{d^*}{2} \leq c_2 \leq d_0 \text{ when } \frac{d^*}{2} < d_0 < d^*,$$

$$\text{i.e. } \frac{d^*}{2} \leq c_2 \leq d_0, \text{ when } \frac{d^*}{2} < d_0 < d^*,$$

$$\text{while } (iii)\ 0 < c_2 < d^* \text{ and } \frac{d^*}{2} \leq c_2 \leq d_0 \text{ when } d^* \leq d_0 < \infty,$$

$$\text{i.e. } \frac{d^*}{2} \leq c_2 < d^* \text{ when } d^* \leq d_0 < \infty$$

where d_0 and d^* are given by (6.8) and (6.11) respectively.

In the case (i) above

$$0 < c_2 \leq d_0 \leq \frac{tr(\Lambda/p)}{|\Lambda|^{1/p}}$$

when $0 < d_0 \leq d^*/2$. Thus the Risk Difference $D(\Lambda, c_2)$ given by (6.12) is monotone decreasing over $0 < c_2 \leq d_0$ when $0 < d_0 \leq d^*/2$ so that the best value of c_2 in the case (i) of the above scheme is given by $c_2 = d_0$.

Hence the theorem is proved.

Finally, we remark that based on the numerical computations of Risk Difference and Relative Risk of the estimators given at the end of this chapter it appears that a general rule of thumb may be to take $c_2 = .75\ d^*$ in all cases except when $0 < d_0 \leq d^*/2$ in which case $c_2 = d_0$ is of course the best choice for c_2.

6.3 Comparison of Risks of the Estimators

In this section we find explicit expressions for the risk functions of the usual and the proposed estimators of the scale matrix of the multivariate t-model and define a measure of comparing the risk functions.

Risk Function of Usual Estimator

The risk function of the usual estimator $\tilde{\Lambda}$ defined by (6.6) is given by

$$\begin{aligned} R(\tilde{\Lambda},\ \Lambda) =& Etr(c_1 A - \Lambda)^2 \\ =& c_1^2 Etr(A^2) + \left(1 - \frac{2\nu c_1 n}{\nu - 2}\right)\ tr(\Lambda^2) \\ =& c_1^2 \varphi(\Lambda^2) + \left(1 - \frac{2\nu c_1 n}{\nu - 2}\right)\ tr(\Lambda^2),\quad \nu > 4. \end{aligned} \tag{6.15}$$

where

$$\varphi(\Lambda^2) = Etr(A^2) = E\left[\sum_{i=1}^{p}\sum_{k=1}^{p} a_{ik}^2\right] = \sum_{i=1}^{p}\sum_{k=1}^{p} E(a_{ik}^2).$$

It follows from Sutradhar and Ali (1989) that for $\nu > 4$

$$\begin{aligned} E(a_{ik}^2) =& \frac{\nu^2 n}{(\nu-2)(\nu-4)}\left[n\ \left(\sum_{l=1}^{p}\delta_{il}\delta_{kl}\right)^2\right. \\ &\left. + 2\sum_{l=1}^{p}\delta_{il}^2\delta_{kl}^2 + \sum_{l<m}(\delta_{il}\delta_{km} + \delta_{im}\delta_{kl})^2\right] \end{aligned} \tag{6.16}$$

where $\Lambda^{1/2} = \Delta = ((\delta_{ik}))$, a positive definite square root matrix of Λ.

It is easily seen that

$$\begin{aligned}
&2\sum_{l=1}^{p}\delta_{il}^2\delta_{kl}^2 + \sum_{l<m}(\delta_{il}\delta_{km} + \delta_{im}\delta_{kl})^2 \\
&= \sum_{l=1}^{p}\delta_{il}^2\delta_{kl}^2 + \sum_{l=1}^{p}\delta_{il}^2\delta_{kl}^2 \\
&+ \sum_{l<m}\delta_{il}^2\delta_{km}^2 + \sum_{l<m}\delta_{im}^2\delta_{kl}^2 \\
&+2\sum_{l<m}\delta_{il}\delta_{km}\delta_{im}\delta_{kl} \\
&= \left(\sum_{l=1}^{p}\delta_{il}^2\delta_{kl}^2 + 2\sum_{l<m}\delta_{il}\delta_{km}\delta_{im}\delta_{kl}\right) \\
&+ \sum_{l=1}^{p}\delta_{il}^2\delta_{kl}^2 + \sum_{l<m}\delta_{il}^2\delta_{km}^2 + \sum_{l<m}\delta_{im}^2\delta_{kl}^2 \\
&= \left(\sum_{l=1}^{p}\delta_{il}\delta_{kl}\right)^2 + \left(\sum_{l=1}^{p}\delta_{il}^2\right)\sum_{l=1}^{p}\delta_{kl}^2
\end{aligned} \tag{6.17}$$

since

$$\begin{aligned}
\left(\sum_{l=1}^{p}\delta_i^2\right)\sum_{l=1}^{p}\delta_{kl}^2 &= (\delta_{i1}^2 + \delta_{i2}^2 + \ldots + \delta_{ip}^2)(\delta_{k1}^2 + \delta_{k2}^2 + \ldots + \delta_{kp}^2) \\
&= \delta_{i1}^2\delta_{k1}^2 + \delta_{i1}^2\delta_{k2}^2 + \ldots + \delta_{i1}^2\delta_{kp}^2 \\
&= \delta_{i2}^2\delta_{k1}^2 + \delta_{i2}^2\delta_{k2}^2 + \ldots + \delta_{i2}^2\delta_{kp}^2 \\
&+ \ldots \quad \ldots \quad \ldots \\
&+ \delta_{ip}^2\delta_{k1}^2 + \delta_{ip}^2\delta_{k2}^2 + \ldots + \delta_{ip}^2\delta_{kp}^2 \\
&= \sum_{l=1}^{p}\delta_{il}^2\delta_{kl}^2 + \sum_{l<m}\delta_{il}^2\delta_{km}^2 + \sum_{l<m}\delta_{im}^2\delta_{kl}^2.
\end{aligned}$$

Again by virtue of $\Lambda = \Delta^2$ we have

$$((\lambda_{ik})) = \left(\left(\sum_{l=1}^{p}\delta_{il}\delta_{kl}\right)\right) \tag{6.18)}$$

It then follows from (6.17) and (6.18) that

$$2\sum_{l=1}^{p}\delta_{il}^2\delta_{kl}^2+\sum_{l<m}(\delta_{il}\delta_{km}+\delta_{im}\delta_{kl})^2=\lambda_{ik}^2+\lambda_{ii}\lambda_{kk},\quad \nu>4. \tag{6.19}$$

Hence we have from (6.16)

$$E(a_{ik}^2)=\frac{\nu^2 n}{(\nu-2)(\nu-4)}\left[(n+1)\lambda_{ik}^2+\lambda_{ii}\lambda_{kk}\right].$$

As $\nu\to\infty$, the result matches with the corresponding result under normality i.e. with $E(w_{ik}^2)$ where W has the usual Wishart distribution $\mathcal{W}_p(n,\Lambda)$ (see e.g. Anderson, 1958, p 161).

Finally we have for $\nu>4$

$$\begin{aligned}\varphi(\Lambda^2)=&Etr(A^2)\\=&E\left(\sum_{i=1}^{p}\sum_{k=1}^{p}a_{ik}^2\right)\\=&\sum_{i=1}^{p}\sum_{k=1}^{p}E(a_{ik}^2)\\=&\frac{\nu^2 n}{(\nu-2)(\nu-4)}\left[(n+1)\sum_{i=1}^{p}\sum_{k=1}^{p}\lambda_{ik}^2+\sum_{i=1}^{p}\sum_{k=1}^{p}\lambda_{ii}\lambda_{kk}\right]\\=&\frac{\nu^2 n}{(\nu-2)(\nu-4)}\left[(n+1)\,tr\,(\Lambda^2)+(tr\Lambda)^2\right].\end{aligned} \tag{6.20}$$

The result in (6.20) can also be proved quickly exploiting the mixture representation of the multivariate t-model. Since $A|\tau\sim\mathcal{W}_p(n,\tau^2\Lambda)$ where $\tau^{-2}\sim G(\nu/2,2/\nu)$, it follows from Srivastava and Khatri (1979, p 99) that

$$\begin{aligned}Etr(A^2)=&E\left[E\left[tr(A^2)|\tau\right]\right]\\=&n\,E\left[(n+1)\,tr(\tau^2\Lambda)^2+(tr\,\tau^2\Lambda)^2\right]\\=&E(\tau^4)\,n\left[(n+1)\,tr(\Lambda^2)+(tr\,\Lambda)^2\right]\\=&\frac{\nu^2 n}{(\nu-2)(\nu-4)}\left[(n+1)\,tr(\Lambda^2)+(tr\,\Lambda)^2\right].\end{aligned}$$

Risk Function of Proposed Estimator

In course of proving Theorem 6.1 we have defined

$$D(\Lambda,\ c_2) = R(\hat{\Lambda},\ \Lambda; c_2) - R(\tilde{\Lambda},\ \Lambda)$$

and found an explicit expression for $D(\Lambda,\ c_2)$ in (6.12). The risk function of the usual estimator $\tilde{\Lambda}$ is also calculated in (6.15). Hence a computable form of the risk function of the proposed estimator $\hat{\Lambda}$ defined by (6.7) follows from

$$R(\hat{\Lambda},\ \Lambda; c_2) = D(\Lambda,\ c_2) + R(\tilde{\Lambda},\ \Lambda).$$

Thus we have proved the following theorem.

Theorem 6.3 The risk functions of the proposed estimator

$$\hat{\Lambda} = c_1 A - c_2 |A|^{1/p} I,$$

where c_1 is a given positive number while c_2 satisfies the condition

$$-\infty < c_2 \leq d_0 = \frac{c_1 min\{a_{11}, \ldots, a_{pp}\}}{|A|^{1/p}},$$

and the usual estimator $\hat{\Lambda} = c_1 A$ are given by

$$R(\tilde{\Lambda},\ \Lambda) = \left[1 + \frac{\nu c_1 n}{\nu - 2}\left(\frac{\nu c_1 (n+1)}{\nu - 4} - 2\right)\right]\ tr(\Lambda^2) + \frac{\nu^2 c_1^2 n}{(\nu-2)(\nu-4)}(tr\Lambda)^2 \quad (6.21)$$

and

$$\begin{aligned} R(\hat{\Lambda},\ \Lambda; c_2) =& \frac{4\nu^2 p}{(\nu-2)(\nu-4)} \frac{\Gamma_p(n/2 + 2/p)}{\Gamma_p(n/2)}\ |\Lambda|^{2/p} c_2 \left(c_2 - \frac{tr(\Lambda/p)}{|\Lambda|^{1/p}} d^*\right) \\ &+ \left[1 + \frac{\nu c_1 n}{\nu - 2}\left(\frac{\nu c_1 (n+1)}{\nu - 4} - 2\right)\right]\ tr(\Lambda^2) + \frac{\nu^2 c_1^2 n}{(\nu-2)(\nu-4)}(tr\Lambda)^2. \end{aligned}$$

We remark that the risk functions of the estimators depend on Λ only through

$$\prod_{i=1}^{p} \xi_i \quad \text{and} \quad \sum_{i=1}^{p} \xi_i^2$$

where $\xi_1, \xi_2, \ldots, \xi_p$ are the characteristic roots of Λ.

Measures of Relative Risk

To compare the risk of the two estimators $\hat{\Lambda}$ and $\tilde{\Lambda}$, the usual way is to use the measure relative risk. Let the relative risk of the two estimators $\hat{\Lambda}$ and $\tilde{\Lambda}$ be given by

$$RR(\hat{\Lambda} : \tilde{\Lambda}; c_2) = \frac{R(\hat{\Lambda},\ \Lambda; c_2)}{R(\tilde{\Lambda},\ \Lambda)}, \tag{6.22}$$

where $0 < RR(\hat{\Lambda} : \tilde{\Lambda}; c_2) \leq 1$ for the choices of c_2 given by Theorem 6.1.

Following Dey (1988) we define the measure Percentage Improvement in Risk (PIR) of $\hat{\Lambda}$ relative to $\tilde{\Lambda}$ as

$$PIR(\hat{\Lambda} : \tilde{\Lambda}; c_2) = 100\left(1 - \frac{R(\hat{\Lambda}, \Lambda; c_2)}{R(\tilde{\Lambda}, \Lambda)}\right) = \frac{-100D(\Lambda,\ c_2)}{R(\tilde{\Lambda},\ \Lambda)} \tag{6.23}$$

where $D(\Lambda,\ c_2)$ is given by (6.12) while $R(\tilde{\Lambda},\ \Lambda)$ is given by Theorem 6.3. Clearly

$$0 < PIR(\hat{\Lambda} : \tilde{\Lambda}; c_2) \leq 100$$

and

$$PIR(\hat{\Lambda} : \tilde{\Lambda}; c_2) = 100\{1 - RR(\hat{\Lambda} : \tilde{\Lambda}; c_2)\}. \tag{6.24}$$

Minimum Value of Relative Risk Function (MRR)

Now we find the minimum value of the relative risk function $RR(\hat{\Lambda} : \tilde{\Lambda}; c_2)$. Let

$$b_1 = b_2 \frac{tr(\Lambda/p)}{|\Lambda|^{1/p}} d^* \tag{6.25}$$

$$b_2 = \frac{4\nu^2 p}{(\nu-2)(\nu-4)} \frac{\Gamma_p(n/2 + 2/p)}{\Gamma_p(n/2)} \frac{|\Lambda|^{2/p}}{R(\tilde{\Lambda},\ \Lambda)} \tag{6.26}$$

where $R(\tilde{\Lambda},\ \Lambda)$ and d^* are given by Theorem 6.3 and (6.11) respectively. The following theorem deals with the minimum value of the relative risk function with respect to c_2 for given Λ and we denote it by

$$MRR(\hat{\Lambda} : \tilde{\Lambda}) = \min_{c_2} RR(\hat{\Lambda} : \tilde{\Lambda}; c_2) = \min_{c_2} \frac{R(\hat{\Lambda},\ \Lambda; c_2)}{R(\tilde{\Lambda},\ \Lambda)}.$$

Theorem 6.4 Consider the relative risk functions

$$RR(\hat{\Lambda} : \tilde{\Lambda}; c_2) = \frac{R(\hat{\Lambda},\ \Lambda; c_2)}{R(\tilde{\Lambda},\ \Lambda)}$$

where the estimators $\hat{\Lambda}$ and $\tilde{\Lambda}$ are the proposed and usual estimators defined by

$$\hat{\Lambda} = c_1 A - c_2 |A|^{1/p} I,$$

where c_1 is a given positive number while c_2 satisfies the condition

$$-\infty < c_2 < d_0 = \frac{c_1 min\{a_{11}, \ldots, a_{pp}\}}{|A|^{1/p}}$$

and $\hat{\Lambda} = c_1 A$ respectively.

Then

$$RR(\hat{\Lambda} : \tilde{\Lambda}; c_2) = 1 - b_1 c_2 + b_2 c_2^2, \quad -\infty < c_2 \le d_0 \tag{6.27}$$

and for given Λ, the relative risk function is minimized at

$$c_2(opt) = \frac{b_1}{2b_2} = \frac{tr(\Lambda/p)}{|\Lambda|^{1/p}} \frac{d^*}{2} \tag{6.28}$$

provided this is admissible, and

$$MRR(\hat{\Lambda} : \tilde{\Lambda}) = \min_{c_2} RR(\hat{\Lambda} : \tilde{\Lambda}; c_2) = 1 - \frac{b_1^2}{4b_2}$$

where d_0 and d^* are given by (6.8) and (6.11) respectively, and b_1 and b_2 are defined by (6.25) and (6.26) respectively.

Proof. It is readily verified that

$$RR(\hat{\Lambda} : \tilde{\Lambda}; c_2) = 1 + \frac{D(\Lambda,\ c_2)}{R(\tilde{\Lambda},\ \Lambda)} \tag{6.29}$$

where $D(\Lambda,\ c_2)$ is given by (6.12).

Now it follows from (6.25) and (6.26) that

$$D(\Lambda,\ c_2) = (b_2 c_2^2 - b_1 c_2) R(\tilde{\Lambda},\ \Lambda).$$

Then from (6.29) we immediately have

$$RR(\hat{\Lambda} : \tilde{\Lambda}; c_2) = 1 - b_1 c_2 + b_2 c_2^2$$

which is a polynomial of degree 2 in c_2 and for given Λ the relative risk function is minimized at

$$c_2 = \frac{b_1}{2b_2} = \frac{tr(\Lambda/p)}{|\Lambda|^{1/p}} \frac{d^*}{2}$$

and the minimum value is given by

$$\begin{aligned} \min_{c_2}\ RR(\hat{\Lambda} : \tilde{\Lambda}; c_2) =& 1 - b_1 \left(\frac{b_1}{2b_2}\right) + b_2 \left(\frac{b_1}{2b_2}\right)^2 \\ =& 1 - \frac{b_1^2}{4b_2}. \end{aligned}$$

Remark 6.2 It may be proved similarly for the estimators $\tilde{\Lambda}$ and $\hat{\Lambda}$ defined by (6.6) and (6.7) respectively that the Percentage Improvement in Risk (PIR) is given by

$$PIR(\hat{\Lambda} : \tilde{\Lambda}; c_2) = 100(b_1 c_2 - b_2 c_2^2), \quad -\infty < c_2 \leq d_0$$

and for given Λ the PIR maximizes at

$$c_2(opt) = \frac{b_1}{2b_2} = \frac{tr(\Lambda/p)}{|\Lambda|^{1/p}} \frac{d^*}{2}$$

with

$$\max_{c_2} PIR(\hat{\Lambda} : \tilde{\Lambda}; c_2) = 25\frac{b_1^2}{b_2}.$$

Some Special Cases

(1). For the bivariate case i.e for $p = 2$, the usual and the proposed estimators of Λ are given by

$$\tilde{\Lambda} = c_1 A,$$

and

$$\hat{\Lambda} = c_1 A - c_2 |A|^{1/2}$$

respectively, where $c_1 > 0$ and

$$-\infty < c_2 \leq d_0 = \frac{c_1 min\{a_{11}, a_{22}\}}{|A|^{1/2}}.$$

It follows from Theorem 6.2 that the proposed estimator dominates the usual estimator under the squared error loss function given by (6.2) for $\nu > 4$ and any c_2 satisfying the following conditions:

1. For a given c_1 satisfying

$$c_1 < \frac{1 - 4/\nu}{n + 1} \quad (i.e.\ d^* < 0),$$

c_2 satisfies

$$d^* < c_2 \leq \frac{d^*}{2}.$$

2. For a given c_1 satisfying

$$c_1 > \frac{1 - 4/\nu}{n + 1} \quad (i.e.\ d^* > 0),$$

c_2 satisfies the following scheme:

$$c_2 = d_0, \quad \text{for} \quad 0 < d_0 \le \frac{d^*}{2},$$
$$\text{and} \quad \frac{d^*}{2} \le c_2 \le d_0 \quad \text{for} \quad \frac{d^*}{2} < d_0 < d^*,$$
$$\text{while} \quad \frac{d^*}{2} \le c_2 < d^* \quad \text{for} \quad d^* \le d_0$$

where

$$d^* = \frac{2}{n}\{c_1(n+1) - 1 + 4/\nu\}.$$

It follows from Theorem 6.3 and Theorem 6.4 that the relative risk of the proposed estimator to usual estimator is given by

$$RR(\hat{\Lambda} : \tilde{\Lambda}; c_2) = 1 - b_1 c_2 + b_2 c_2^2, \quad -\infty < c_2 \le d_0$$

and for given Λ, the minimum value of the relative risk function is given by

$$\min_{c_2} \; RR(\hat{\Lambda} : \tilde{\Lambda}; c_2) = 1 - \frac{b_1^2}{4b_2}$$

where

$$b_1 = b_2 \frac{(\lambda_{11} + \lambda_{22})/2}{(\lambda_{11}\lambda_{22})^{1/2}} d^*$$
$$b_2 = \frac{2\nu^2 n(n-1)}{(\nu-2)(\nu-4)} \frac{\lambda_{11}\lambda_{22}}{R(\tilde{\Lambda},\ \Lambda)}$$

and

$$R(\tilde{\Lambda},\ \Lambda) = \left[1 + \frac{\nu c_1 n}{\nu - 2}\left(\frac{\nu c_1(n+1)}{\nu - 4} - 2\right)\right](\lambda_{11}^2 + \lambda_{22}^2) + \frac{\nu^2 c_1^2 n}{(\nu-2)(\nu-4)}(\lambda_{11} + \lambda_{22})^2.$$

We remark that in this case the relative risk depends on the characteristic roots of Λ only through their ratio.

(2). For the trivariate case i.e for $p = 3$, the usual and proposed estimators of Λ are given by

$$\tilde{\Lambda} = c_1 A,$$

and

$$\hat{\Lambda} = c_1 A - c_2 |A|^{1/3} I$$

respectively, where c_1 is a fixed positive number and

$$-\infty < c_2 \le d_0 = \frac{c_1 min\{a_{11}, a_{22}, a_{33}\}}{|A|^{1/2}}.$$

It follows from Theorem 6.2 that the proposed estimator dominates the usual estimator for $\nu > 4$ and any c_2 satisfying the following conditions:

1. For a given c_1 satisfying

$$c_1 < \frac{\nu - 4}{\nu} \frac{3}{3n+2} \quad (i.e.\ d^* < 0),$$

c_2 satisfies

$$d^* < c_2 \le \frac{d^*}{2}.$$

2. For a given c_1 satisfying

$$c_1 > \frac{\nu - 4}{\nu} \frac{3}{3n+2} \quad (i.e.\ d^* > 0),$$

c_2 satisfies the following scheme:

$$\begin{aligned} c_2 = d_0 \quad &\text{for} \quad 0 < d_0 \le \frac{d^*}{2}, \\ \text{and} \quad \frac{d^*}{2} \le c_2 \le d_0 \quad &\text{for} \quad \frac{d^*}{2} < d_0 < d^* \\ \text{while} \quad \frac{d^*}{2} \le c_2 < d^* \quad &\text{for} \quad d^* \le d_0 \end{aligned}$$

where

$$d^* = \left(c_1 \frac{3n+2}{3} - \frac{\nu - 4}{\nu}\right) \frac{\Gamma(n/2 - 1/6)\Gamma(n/2 + 1/3)\Gamma(n/2 - 2/3)}{\Gamma(n/2 + 1/6)\Gamma(n/2 - 1/3)\Gamma(n/2 + 2/3)}.$$

It also follows from Theorem 6.4 that the relative risk $RR(\hat{\Lambda} : \tilde{\Lambda}; c_2)$ of our proposed estimator $\hat{\Lambda} = c_1 A - c_2 |A|^{1/3} I$ relative to usual estimator $\tilde{\Lambda} = c_1 A$ is given by

$$RR(\hat{\Lambda} : \tilde{\Lambda}; c_2) = 1 - b_1 c_2 + b_2 c_2^2, \quad -\infty < c_2 \leq d_0$$

and for given Λ, the minimum value of relative risk function is given by

$$\min_{c_2} \; RR(\hat{\Lambda} : \tilde{\Lambda}; c_2) = 1 - \frac{b_1^2}{4b_2}$$

where

$$b_1 = b_2 \frac{(\lambda_{11} + \lambda_{22} + \lambda_{33})/3}{(\lambda_{11}\lambda_{22}\lambda_{33})^{1/3})} d^*, \tag{6.30}$$

$$b_2 = \frac{12\nu^2}{(\nu-2)(\nu-4)} \frac{(\lambda_{11}\lambda_{22}\lambda_{33})^{2/3}}{R(\tilde{\Lambda},\ \Lambda)} \times \frac{\Gamma(n/2+2/3)\Gamma(n/2+1/6)\Gamma(n/2-1/3)}{\Gamma(n/2)\Gamma(n/2-1/2)\Gamma(n/2-1)} \tag{6.31}$$

and

$$R(\tilde{\Lambda},\ \Lambda) = \left[1 + \frac{\nu c_1 n}{\nu - 2}\left(\frac{\nu c_1 (n+1)}{\nu - 4} - 2\right)\right](\lambda_{11}^2 + \lambda_{22}^2 + \lambda_{33}^3) + \frac{\nu^2 c_1^2 n}{(\nu-2)(\nu-4)}(\lambda_{11} + \lambda_{22} + \lambda_{33})^2.$$

So far we have discussed proposed estimator in contrast to usual estimator (multiples of sample sum of products matrix A). Among the usual estimators two estimators of greater interest are the maximum likelihood estimator and the unbiased estimator. The following section specializes the results of sections 6.2 and 6.3 to compare the proposed estimator with the maximum likelihood estimator.

6.4 Proposed Estimator Dominating the Maximum Likelihood Estimator

The maximum likelihood estimator of Λ is given by $\tilde{\Lambda}_1 = A/(n+1)$ (Anderson and Fang, 1990a, p 208) so that $c_1 = (n+1)^{-1}$.

The following theorem follows from Theorem 6.2, Theorem 6.3 and Theorem 6.4 by setting $c_1 = (n+1)^{-1}$.

Theorem 6.5 Consider the model given by (6.1) for $\nu > 4$. The proposed estimator

$$\hat{\Lambda}_1 = A/(n+1) - c_2|A|^{1/p}I$$

where

$$-\infty < c_2 \leq d_0 = \frac{min\{a_{11}, \ldots, a_{pp}\}}{(n+1)|A|^{1/p}},$$

dominates m.l.e $\tilde{\Lambda}_1 = A/(n+1)$ for any c_2 satisfying the following conditions :

1. For $p \geq 3$ and

$$\frac{1}{n+1} < \frac{\nu-4}{\nu}\frac{p}{np+2} \quad (i.e.\ d^* < 0)$$

c_2 satisfies

$$d^* < c_2 \leq \frac{d^*}{2}.$$

2. For $p \geq 2$ and

$$\frac{1}{n+1} > \frac{\nu-4}{\nu}\frac{p}{np+2} \quad (i.e.\ d^* > 0)$$

c_2 satisfies the following scheme:

$$\begin{aligned} c_2 = d_0 \ \text{ for } \ 0 < d_0 \leq \frac{d^*}{2}, \\ \text{and} \ \frac{d^*}{2} \leq c_2 \leq d_0 \ \text{ for } \ \frac{d^*}{2} < d_0 < d^*, \\ \text{while} \ \frac{d^*}{2} \leq c_2 < d^* \ \text{ for } \ d^* \leq d_0 \end{aligned}$$

where

$$d^* = \left(\frac{1}{n+1}\frac{np+2}{p} - \frac{\nu-4}{\nu}\right)\frac{\Gamma_p(n/2+1/p)}{\Gamma_p(n/2+2/p)}.$$

Also the relative risk function is given by

$$RR(\hat{\Lambda}_1 : \tilde{\Lambda}_1; c_2) = 1 - b_1 c_2 + b_2 c_2^2, \quad -\infty < c_2 \leq d_0$$

and for given Λ,

$$MRR(\hat{\Lambda}_1 : \tilde{\Lambda}_1) = \min_{c_2} RR(\hat{\Lambda} : \tilde{\Lambda}; c_2) = 1 - \frac{b_1^2}{4b_2}$$

where

$$b_1 = b_2 \frac{tr(\Lambda/p)}{|\Lambda|^{1/p}} d^*$$

$$b_2 = \frac{4\nu^2 p}{(\nu-2)(\nu-4)} \frac{\Gamma_p(n/2 + 2/p)}{\Gamma_p(n/2)} \frac{|\Lambda|^{2/p}}{R(\tilde{\Lambda}_1, \Lambda)}$$

and

$$R(\tilde{\Lambda}_1, \Lambda) = \left[1 + \frac{\nu}{\nu-2}\frac{n}{n+1}\left(\frac{\nu}{\nu-4} - 2\right)\right] tr(\Lambda^2) + \frac{\nu^2}{(\nu-2)(\nu-4)}\frac{n}{(n+1)^2}(tr\Lambda)^2.$$

Some Special Cases

(1). For the bivariate case i.e for $p = 2$, the maximum likelihood and the proposed estimators of Λ are given by

$$\tilde{\Lambda}_1 = A/(n+1),$$

and

$$\hat{\Lambda}_1 = A/(n+1) - c_2|A|^{1/2} I$$

respectively, where

$$-\infty < c_2 \leq d_0 = \frac{min\{a_{11}, a_{22}\}}{(n+1)|A|^{1/2}}.$$

It follows from Theorem 6.2 by putting $c_1 = (n+1)^{-1}$ that in this case the proposed estimator dominates the m.l.e. under the squared error loss function given by (6.2) for $\nu > 4$ and any c_2 satisfying the following conditions:

For a given c_1 satisfying

$$c_1 > \frac{\nu-4}{\nu}\frac{p}{np+2} \quad (i.e.\ d^* > 0),$$

c_2 satisfies the following scheme:

$$\begin{aligned} c_2 = d_0 &\quad \text{for} \quad 0 < d_0 \le \frac{d^*}{2}, \\ \text{and} \quad \frac{d^*}{2} \le c_2 \le d_0 &\quad \text{for} \quad \frac{d^*}{2} < d_0 < d^*, \\ \text{while} \quad \frac{d^*}{2} \le c_2 < d^* &\quad \text{for} \quad d^* \le d_0, \end{aligned}$$

where $d^* = 8/(\nu n)$.

It follows from Theorem 6.4 that for the bivariate case i.e for $p = 2$, the relative risk $RR(\hat{\Lambda}_1 : \tilde{\Lambda}_1; c_2)$ of our proposed estimator

$$\hat{\Lambda}_1 = A/(n+1) - c_2|A|^{1/2}I$$

relative to the maximum likelihood estimator $\tilde{\Lambda}_1 = A/(n+1)$ is given by

$$RR(\hat{\Lambda}_1 : \tilde{\Lambda}_1; c_2) = 1 - b_1c_2 + b_2c_2^2, \quad -\infty < c_2 \le d_0$$

and for given Λ,

$$\min_{c_2} RR(\hat{\Lambda}_1 : \tilde{\Lambda}_1; c_2) = 1 - \frac{b_1^2}{4b_2}$$

where

$$\begin{aligned} b_1 &= b_2 \frac{(\lambda_{11} + \lambda_{22})/2}{(\lambda_{11}\lambda_{22})^{1/2}} d^*, \\ b_2 &= \frac{2\nu^2 n(n-1)}{(\nu-2)(\nu-4)} \frac{\lambda_{11}\lambda_{22}}{R(\tilde{\Lambda}_1,\ \Lambda)}, \end{aligned}$$

and

$$\begin{aligned} R(\tilde{\Lambda}_1,\ \Lambda) = &\left[1 + \frac{\nu}{\nu-2}\frac{n}{n+1}\left(\frac{\nu}{\nu-4} - 2\right)\right](\lambda_{11}^2 + \lambda_{22}^2) \\ &+ \frac{\nu^2}{(\nu-2)(\nu-4)}\frac{n}{(n+1)^2}(\lambda_{11} + \lambda_{22})^2. \end{aligned}$$

(2). For the trivariate case namely for $p = 3$ the maximum likelihood and the proposed estimators are given by

$$\tilde{\Lambda}_1 = A/(n+1)$$

and

$$\hat{\Lambda}_1 = A/(n+1) - c_2|A|^{1/3}I$$

respectively, where

$$-\infty < c_2 \leq d_0 = \frac{min\{a_{11}, a_{22}, a_{33}\}}{(n+1)|A|^{1/3}}.$$

It follows from Theorem 6.2 by putting $c_1 = (n+1)^{-1}$ that in this case the proposed estimator dominates the usual estimator for $\nu > 4$ and any c_2 satisfying the following conditions:

1. For $\nu > 12(n+1)$ (*i.e.* $d^* < 0$), the value of c_2 satisfies

$$d^* < c_2 \leq \frac{d^*}{2}.$$

2. For $\nu < 12(n+1)$ (*i.e.* $d^* > 0$), c_2 satisfies the following scheme:

$$c_2 = d_0 \quad \text{for} \quad 0 < d_0 \leq \frac{d^*}{2},$$
$$\text{and} \quad \frac{d^*}{2} \leq c_2 \leq d_0 \quad \text{for} \quad \frac{d^*}{2} < d_0 < d^*$$
$$\text{while} \quad \frac{d^*}{2} \leq c_2 < d^* \quad \text{for} \quad d^* \leq d_0$$

where

$$d^* = \left(\frac{-1}{3(n+1)} + \frac{4}{\nu}\right) \frac{\Gamma(n/2 - 1/6)\Gamma(n/2 + 1/3)\Gamma(n/2 - 2/3)}{\Gamma(n/2 + 1/6)\Gamma(n/2 - 1/3)\Gamma(n/2 + 2/3)}.$$

It also follows from Theorem 6.4, by putting $c_1 = (n+1)^{-1}$, that for the trivariate case namely for $p = 3$, the relative risk $RR(\hat{\Lambda}_1 : \tilde{\Lambda}_1; c_2)$ of our proposed estimator

$$\hat{\Lambda}_1 = A/(n+1) - c_2|A|^{1/3}I$$

relative to the maximum likelihood estimator $\tilde{\Lambda}_1 = A/(n+1)$ is given by

$$RR(\hat{\Lambda}_1 : \tilde{\Lambda}_1; c_2) = 1 - b_1 c_2 + b_2 c_2^2, \quad -\infty < c_2 \leq d_0$$

and for given Λ,

$$\min_{c_2} RR(\hat{\Lambda}_1 : \tilde{\Lambda}_1; c_2) = 1 - \frac{b_1^2}{4b_2}$$

where b_1 and b_2 are given by (6.30) and (6.31) with d^* as given above and

$$R(\tilde{\Lambda}_1, \Lambda) = \left[1 + \frac{\nu}{\nu-2}\frac{n}{n+1}\left(\frac{\nu}{\nu-4} - 2\right)\right] (\lambda_{11}^2 + \lambda_{22}^2 + \lambda_{33}^2)$$
$$+ \frac{\nu^2}{(\nu-2)(\nu-4)}\frac{n}{(n+1)^2}(\lambda_{11} + \lambda_{22} + \lambda_{33})^2.$$

6.4.1 Numerical Computation of Relative Risk Function

Some numerical computations have been performed to compare the proposed estimator

$$\hat{\Lambda}_1 = A/(n+1) - c_2|A|^{1/p} I$$

with the maximum likelihood estimator (m.l.e.)

$$\tilde{\Lambda}_1 = A/(n+1).$$

where

$$-\infty < c_2 \leq d_0 = \frac{min\{a_{11}, a_{22}, \ldots, a_{pp}\}}{(n+1)|A|^{1/p}}.$$

We compute the Minimum Relative Risk (MRR)

$$MRR(\hat{\Lambda}_1 : \tilde{\Lambda}_1) = \min_{c_2} RR(\hat{\Lambda}_1 : \tilde{\Lambda}_1; c_2) = \min_{c_2} \frac{R(\hat{\Lambda}_1, \Lambda; c_2)}{R(\tilde{\Lambda}_1, \Lambda)} = 1 - \frac{b_1^2}{4b_2}$$

where b_1 and b_2 are defined in Theorem 6.5.

For computational purposes we consider mostly the case when the scale matrix Λ is diagonal (in which case the diagonal elements are also characteristic roots).

However the diagonal elements of Λ are so chosen that they represent a broad spectrum of characteristic roots of Λ. Minimum Relative Risk has also been computed for a correleted scale matrix Λ (see Table 6.13 on pages 160–161).

The numerical computations have been summarized in thirteen tables. The Tables 6.1 to 6.12 are of the same pattern; the first column shows the values of ν and n, the second column shows the values of Minimum Relative Risk (MRR), the third column shows the values of d^* while the fourth column shows the optimum value of c_2 for different values of ν and n. Tables 6.1 to 6.12 on pages 142–159 show numerical computations of Minimum Relative Risk when Λ is diagonal.

Graphs Showing Risk Difference

Theorem 6.1 is based on the Risk Difference $D(\Lambda, c_2)$ of the estimators. We consider an example to calculate the Risk Difference $D(\Lambda, c_2)$ given by (6.12) for $p = 3, c_1 = (n+1)^{-1}$ with the following scale matrix Λ

$$\Lambda = \begin{pmatrix} 94 & 41 & 23 \\ 41 & 26 & 11 \\ 23 & 11 & 6 \end{pmatrix}.$$

Our aim is to check the behaviour of the Risk Difference between the proposed estimator $\hat{\Lambda}_1$ and the maximum likelihood estimator $\tilde{\Lambda}_1 = (n+1)^{-1}A$ (see Section 6.4). We plot the Risk Difference

$$\begin{aligned} f(c_2) =& (404.90376)^{-1} D(\Lambda, c_2) \\ =& \frac{\nu^2 G}{(\nu-2)(\nu-4)}\, c_2(c_2 - 7.2304281 d^*) \end{aligned}$$

against different values of c_2 where

$$d^* = \left(\frac{-1}{3(n+1)} + \frac{4}{\nu}\right) \frac{\Gamma(n/2 - 1/6)\Gamma(n/2 + 1/3)\Gamma(n/2 - 2/3)}{\Gamma(n/2 + 1/6)\Gamma(n/2 - 1/3)\Gamma(n/2 + 2/3)},$$

and

$$G = \frac{\Gamma(n/2 + 2/3)\Gamma(n/2 + 1/6)\Gamma(n/2 - 1/3)}{\Gamma(n/2)\Gamma(n/2 - 1/2)\Gamma(n/2 - 1)}.$$

Graphs showing $f(c_2)$ against c_2 for different n and ν are shown in Figures 6.1 to 6.7 on pages 162–168.

Graphs Showing Relative Risk

Graphs showing the Relative Risk Function

$$RR(\hat{\Lambda}_1 : \tilde{\Lambda}_1; c_2) = \frac{R(\hat{\Lambda}_1,\ \Lambda; c_2)}{R(\tilde{\Lambda}_1,\ \Lambda)}$$

where

$$R(\tilde{\Lambda}_1, \Lambda) = (n+1)^{-2}\varphi(\Lambda^2) + 14210\left(1 - \frac{\nu}{\nu-2}\frac{2n}{n+1}\right),$$

$$R(\hat{\Lambda}_1, \Lambda) = D(\Lambda, c_2) + R(\tilde{\Lambda}_1, \Lambda)$$

and

$$\begin{aligned}\varphi(\Lambda^2) &= \frac{\nu^2 n}{(\nu-2)(\nu-4)}\left[(n+1)\ tr(\Lambda^2) + (tr\Lambda)^2\right] \\ &= \frac{\nu^2 n}{(\nu-2)(\nu-4)}\left[14210\ (n+1) + (126)^2\right] \\ &= \frac{\nu^2 n}{(\nu-2)(\nu-4)}\left[14210\ n + 30086\right]\end{aligned}$$

against different values of c_2 are shown in Figures 6.8 to 6.14 on pages 169–175.

Summary of Numerical Results

We note that the lower the Minimum Relative Risk (MRR), the better the proposed estimator as compared to the maximum likelihood estimator. Based on the numerical computation we have the following comments:

Although the proposed estimator always dominates the maximum likelihood estimator,

1. the higher the value of n, the lower the Minimum Relative Risk,
2. the higher the value of ν, the higher the Minimum Relative Risk ,
3. the higher the value of p, the higher the Minimum Relative Risk.

6.4.2 An Example

Let us have the following observed sum of products matrix

$$A = \begin{pmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{pmatrix}$$

corresponding to $n = 10$, $\nu = 5$ and $p = 3$. We calculate the maximum likelihood estimator and the proposed estimator ($\tilde{\Lambda}_1$ and $\hat{\Lambda}_1$ respectively) defined in section 6.4.

Here we have $d^* = 0.17143446$ (see Table 6.4 on page 148). From the Wishart matrix we have

$$d_0 = \frac{min\{a_{11}, a_{22}, a_{33}\}}{(n+1)|A|^{1/3}} = 0.08017177$$

so that d_0 satisfies $0 < d_0 \leq d^*/2$ and consequently

$$c_2 = d_0 = 0.08017177$$

(see Condition 2 of Theorem 6.5). Then it is readily seen that the maximum likelihood estimator $\tilde{\Lambda}_1 = A/(n+1)$ is given by

$$\tilde{\Lambda}_1 = \begin{pmatrix} 1.18181818 & -0.36363636 & 0.18181818 \\ -0.36363636 & 1.18181818 & -0.18181818 \\ 0.18181818 & -0.18181818 & 0.90909091 \end{pmatrix}$$

while the proposed estimator $\hat{\Lambda}_1$ is given by

$$\hat{\Lambda}_1 = \begin{pmatrix} 0.27272728 & -0.36363636 & 0.18181818 \\ -0.36363636 & 0.27272728 & -0.18181818 \\ 0.18181818 & -0.18181818 & 0.00000000 \end{pmatrix}.$$

The following section specializes the results of Section 6.2 and Section 6.3 to compare the unbiased estimator with the proposed estimator.

6.5 Proposed Estimator Dominating Unbiased Estimator

An unbiased estimator of Λ is given by $\tilde{\Lambda}_2 = (\nu - 2)A/(\nu n)$ (Anderson and Fang, 1990a, p 208) so that $c_1 = (\nu - 2)/(\nu n)$.

The following theorem follows from Theorem 6.2, Theorem 6.3 and Theorem 6.4 by setting $c_1 = (\nu - 2)/(\nu n)$.

Theorem 6.6 The proposed estimator

$$\hat{\Lambda}_2 = \frac{\nu - 2}{\nu n} A - c_2 |A|^{1/p} I$$

where

$$-\infty < c_2 \leq d_0 = \frac{(\nu - 2) min\{a_{11}, \ldots, a_{pp}\}}{\nu n |A|^{1/p}},$$

dominates the unbiased estimator $\tilde{\Lambda}_2 = (\nu - 2)A/(\nu n)$ under the squared error loss function given by (6.2) for $\nu > 4$ and any c_2 satisfying the following conditions:

For a given

$$\frac{\nu - 2}{\nu - 4} > \frac{np}{np + 2} \quad (i.e.\ d^* > 0),$$

c_2 satisfies the following scheme:

$$c_2 = d_0 \quad \text{for} \quad 0 < d_0 \leq \frac{d^*}{2},$$
$$\text{and} \quad \frac{d^*}{2} \leq c_2 \leq d_0 \quad \text{for} \quad \frac{d^*}{2} < d_0 < d^*,$$
$$\text{while} \quad \frac{d^*}{2} \leq c_2 < d^* \quad \text{for} \quad d^* \leq d_0$$

where

$$d^* = \left(\frac{\nu - 2}{\nu n} \frac{np + 2}{p} - \frac{\nu - 4}{\nu} \right) \frac{\Gamma_p(n/2 + 1/p)}{\Gamma_p(n/2 + 2/p)}.$$

Also the relative risk function is given by

$$RR(\hat{\Lambda}_2 : \tilde{\Lambda}_2; c_2) = 1 - b_1 c_2 + b_2 c_2^2, \quad -\infty < c_2 \leq d_0$$

and for given Λ

$$MRR(\hat{\Lambda}_2 : \tilde{\Lambda}_2) = \min_{c_2} \; RR(\hat{\Lambda}_2 : \tilde{\Lambda}_2; c_2) = 1 - \frac{b_1^2}{4b_2}$$

where

$$\begin{aligned} b_1 &= b_2 \frac{tr(\Lambda/p)}{|\Lambda|^{1/p}} d^* \\ b_2 &= \frac{4\nu^2 p}{(\nu-2)(\nu-4)} \frac{\Gamma_p(n/2+2/p)}{\Gamma_p(n/2)} \frac{|\Lambda|^{2/p}}{R(\tilde{\Lambda}_2,\ \Lambda)} \end{aligned}$$

and

$$R(\tilde{\Lambda}_2,\ \Lambda) = \frac{\nu + 2(n-1)}{(\nu-4)n} \, tr(\Lambda^2) + \frac{\nu-2}{(\nu-4)n}(tr\Lambda)^2.$$

Some Special Cases

(1). For the bivariate case i.e for $p = 2$, the unbiased and the proposed estimators of Λ are given by

$$\tilde{\Lambda}_2 = \frac{\nu-2}{\nu n} A,$$

and

$$\hat{\Lambda}_2 = \frac{\nu-2}{\nu n} A - c_2 |A|^{1/2} I$$

respectively, where

$$-\infty < c_2 \le d_0 = \frac{(\nu-2) min\{a_{11}, a_{22}\}}{\nu n |A|^{1/2}}.$$

It follows from Theorem 6.6 by putting $c_1 = (\nu-2)/(\nu n)$ that the proposed estimator dominates the unbiased estimator under the squared error loss function given by (6.2) for $\nu > 4$ and any c_2 satisfying the following conditions:

$$\begin{aligned} & c_2 = d_0 \;\text{ for }\; 0 < d_0 \le \frac{d^*}{2}, \\ \text{and} \;\; & \frac{d^*}{2} \le c_2 \le d_0 \;\text{ for }\; \frac{d^*}{2} < d_0 < d^*, \\ \text{while} \;\; & \frac{d^*}{2} \le c_2 < d^* \;\text{ for }\; d^* \le d_0, \end{aligned}$$

where d^* is given by

$$d^* = \frac{2}{\nu n^2}\{\nu + 2(n-1)\}.$$

It also follows from Theorem 6.6 that when $p = 2$, the relative risk of the proposed estimator to the unbiased estimator is given by

$$RR(\hat{\Lambda}_2 : \tilde{\Lambda}_2; c_2) = 1 - b_1 c_2 + b_2 c_2^2, \quad -\infty < c_2 \leq d_0$$

and for given Λ, the minimum value of the relative risk function is given by

$$\min_{c_2} \; RR(\hat{\Lambda}_2 : \tilde{\Lambda}_2; c_2) = 1 - \frac{b_1^2}{4b_2}$$

where

$$b_1 = b_2 \frac{(\lambda_{11} + \lambda_{22})/2}{(\lambda_{11}\lambda_{22})^{1/2}} d^*,$$

$$b_2 = \frac{2\nu^2 n(n-1)}{(\nu-2)(\nu-4)} \frac{\lambda_{11}\lambda_{22}}{R(\tilde{\Lambda}_2,\ \Lambda)},$$

with

$$R(\tilde{\Lambda}_2,\ \Lambda) = \frac{\nu + 2(n-1)}{(\nu-4)n}(\lambda_{11}^2 + \lambda_{22}^2) + \frac{\nu - 2}{(\nu-4)n}(\lambda_{11} + \lambda_{22})^2.$$

We remark that in this case the relative risk depends on the characteristic roots of Λ only through their ratio.

(2). For the trivariate case i.e for $p = 3$, the unbiased and proposed estimators of Λ are given by

$$\tilde{\Lambda}_2 = \frac{\nu - 2}{\nu n} A,$$

and

$$\hat{\Lambda}_2 = \frac{\nu - 2}{\nu n} A - c_2 |A|^{1/3} I$$

respectively, where

$$-\infty < c_2 \leq d_0 = \frac{(\nu - 2) min\{a_{11}, a_{22}, a_{33}\}}{\nu n |A|^{1/2}}.$$

It follows from Theorem 6.6 that for the trivariate case the proposed estimator dominates the unbiased estimator under the squared error loss function given by (6.2) for $\nu > 4$ and for any c_2 satisfying the following conditions:

$$c_2 = d_0, \text{ for } 0 < d_0 \leq \frac{d^*}{2},$$
$$\text{and } \frac{d^*}{2} \leq c_2 \leq d_0 \text{ for } \frac{d^*}{2} < d_0 < d^*,$$
$$\text{while } \frac{d^*}{2} \leq c_2 < d^* \text{ for } d^* \leq d_0$$

where

$$d^* = \frac{2(\nu + 3n - 2)}{3\nu n} \frac{\Gamma(n/2 - 1/6)\Gamma(n/2 + 1/3)\Gamma(n/2 - 2/3)}{\Gamma(n/2 + 1/6)\Gamma(n/2 - 1/3)\Gamma(n/2 + 2/3)}.$$

It also follows from Theorem 6.6 that for $p = 3$, the relative risk of the proposed estimator relative to the unbiased estimator is given by

$$RR(\hat{\Lambda}_2 : \tilde{\Lambda}_2; c_2) = 1 - b_1 c_2 + b_2 c_2^2, \quad -\infty < c_2 \leq d_0$$

and for given Λ, the minimum value of relative risk function is given by

$$\min_{c_2} RR(\hat{\Lambda}_2 : \tilde{\Lambda}_2; c_2) = 1 - \frac{b_1^2}{4b_2}$$

where b_1 and b_2 are given by (6.30) and (6.31) with d^* as given above and

$$R(\tilde{\Lambda}_2, \Lambda) = \frac{\nu + 2(n-1)}{(\nu - 4)n}(\lambda_{11}^2 + \lambda_{22}^2 + \lambda_{33}^3) + \frac{\nu - 2}{(\nu - 4)n}(\lambda_{11} + \lambda_{22} + \lambda_{33})^2.$$

6.6 Estimation of the Scale Matrix (Normal Case)

As mentioned earlier in Chapter 5 that the multivariate t-model converges to the multivariate normal model $N_p(\theta, \Sigma)$ as $\nu \to \infty$ and then the distribution of the sum of products matrix,

$$A = \sum_{j=1}^{N}(X_j - \bar{X})(X_j - \bar{X})'$$

based on the multivariate t-model given by (5.3), approaches the usual Wishart distribution $\mathcal{W}_p(n, \Lambda)$ given by (5.1). To avoid confusion, we will denote the sum of products matrix A by W and the scale matrix Λ by Σ. In this section we want to estimate the scale matrix Σ of the multivariate normal distribution $N_p(\theta, \Sigma)$. In doing so we consider the following squared error loss function

$$L(u(W), \Sigma) = tr[(u(W) - \Sigma)^2] \tag{6.32}$$

where $u(W)$ is any suitable estimator of Σ. The risk function is defined as usual

$$R(u(W), \Sigma) = E[L(u(W), \Sigma)].$$

In the following theorem we specialize the results of Sections 6.2 and 6.3 by letting $\nu \to \infty$ for the case of the multivariate normal distribution. *From the numerical computation we observe that the proposed estimator does not have substantial gain over the maximum likelihood estimator in the normal case .*

Theorem 6.7 Suppose that the p-dimensional ($p \geq 2$) random vectors $X_1, X_2, \ldots,$ X_N are independently and identically distributed according to $N_p(\theta, \Sigma)$. Consider the following two estimators of Λ:

$$\text{usual estimator } \tilde{\Sigma} = c_1 W, \tag{6.33}$$

$$\text{proposed estimator } \hat{\Sigma} = c_1 W - c_2 |W|^{1/p} I, \tag{6.34}$$

where c_1 is a fixed positive number and

$$-\infty < c_2 \leq d_0 = \frac{c_1 min\{w_{11}, \ldots, w_{pp}\}}{|W|^{1/p}}.$$

Then the proposed estimator dominates the usual estimator under the squared error loss function given by (6.32) for any c_2 satisfying the following conditions:

1. For a given c_1 satisfying

$$c_1 < p/(np+2) \quad (i.e.\ d^* < 0),$$

c_2 satisfies

$$d^* < c_2 \le \frac{d^*}{2},$$

2. For a given c_1 satisfying

$$c_1 > p/(np+2) \quad (i.e.\ d^* > 0),$$

c_2 satisfies the following scheme:

$$\begin{aligned} & c_2 = d_0 \text{ for } 0 < d_0 \le \frac{d^*}{2}, \\ \text{and} \quad & \frac{d^*}{2} \le c_2 \le d_0 \quad \text{for} \quad \frac{d^*}{2} < d_0 < d^*, \\ \text{while} \quad & \frac{d^*}{2} \le c_2 < d^* \quad \text{for} \quad d^* \le d_0 \end{aligned}$$

where

$$d^* = \left(c_1 \frac{np+2}{p} - 1\right) \frac{\Gamma_p(n/2+1/p)}{\Gamma_p(n/2+2/p)}. \tag{6.35}$$

Remark 6.3 In this case it follows from (6.12), by letting $\nu \to \infty$ that

$$D(\Sigma,\ c_2) = 4p \frac{\Gamma_p(n/2+2/p)}{\Gamma_p(n/2)} |\Sigma|^{2/p} c_2 \left(c_2 - \frac{tr(\Sigma/p)}{|\Sigma|^{1/p}} d^*\right) \tag{6.36}$$

where d^* is given by (6.35).

We also note that when

$$c_1 = \frac{p}{np+2}$$

we have $d^* = 0$. It then follows from (6.36) that $\hat{\Sigma}$ dominates $\tilde{\Sigma}$ unless $c_2 = 0$. When $c_2 = 0$ the two estimators $\hat{\Sigma}$ and $\tilde{\Sigma}$ coincide.

6.7 Comparison of Risks of the Estimators (Normal Case)

In this section we calculate the risk functions of the usual and proposed estimators of the scale matrix of the multivariate normal distribution. The estimators are defined by (6.33) and (6.34) respectively.

Computable forms of the risk function of the usual and proposed estimators follow from Theorem 6.3 by letting $\nu \to \infty$. The following theorem summarizes the results of the risk functions of the usual and proposed estimators.

Theorem 6.8 The risk function of the usual estimator $\tilde{\Sigma} = c_1 W$ and the proposed estimator

$$\hat{\Sigma} = c_1 W - c_2 |W|^{1/p} I,$$

where c_1 is a fixed positive number and

$$-\infty < c_2 \leq d_0 = \frac{c_1 min\{w_{11}, \ldots, w_{pp}\}}{|W|^{1/p}},$$

are given by

$$R(\tilde{\Sigma}, \Sigma) = [1 + c_1 n(c_1 n + c_1 - 2)]\ tr(\Sigma^2) + c_1^2\ n(tr\Sigma)^2 \tag{6.37}$$

and

$$\begin{aligned} R(\hat{\Sigma},\ \Sigma; c_2) =& 4p\ \frac{\Gamma_p(n/2 + 2/p)}{\Gamma_p(n/2)}\ |\Sigma|^{2/p} c_2 \left(c_2 - \frac{tr(\Sigma/p)}{|\Sigma|^{1/p}} d^*\right) \\ &+ [1 + c_1 n(c_1 n + c_1 - 2)]\ tr(\Sigma^2) + c_1^2\ (tr\Sigma)^2. \end{aligned} \tag{6.38}$$

where d^* is given by (6.35).

The measure of relative risk of the estimators defined by (6.33) and (6.34) respectively can be defined exactly the same way as have discussed in (6.22) and (6.23). Now we have the following theorem by letting $\nu \to \infty$ in Theorem 6.4.

Theorem 6.9 Consider the relative risk function

$$RR(\hat{\Sigma} : \tilde{\Sigma}; c_2) = \frac{R(\hat{\Sigma},\ \Sigma; c_2)}{R(\tilde{\Sigma},\ \Sigma)}$$

where the estimators $\hat{\Sigma}$ and $\tilde{\Sigma}$ are defined by

$$\hat{\Sigma} = c_1 W - c_2 |W|^{1/p} I,$$

where c_1 is a fixed positive number and

$$-\infty < c_2 \leq d_0 = \frac{c_1 min\{w_{11}, \ldots, w_{pp}\}}{|W|^{1/p}},$$

and $\tilde{\Sigma} = c_1 W$ respectively .

Then

$$RR(\hat{\Sigma} : \tilde{\Sigma}; c_2) = 1 - b_1 c_2 + b_2 c_2^2, \quad -\infty < c_2 \leq d_0$$

where

$$\begin{aligned} b_1 &= b_2 \frac{tr(\Sigma/p)}{|\Sigma|^{1/p}} d^* \\ b_2 &= \frac{4\nu^2 p}{(\nu-2)(\nu-4)} \frac{\Gamma_p(n/2+2/p)}{\Gamma_p(n/2)} \frac{|\Sigma|^{2/p}}{R(\tilde{\Sigma},\ \Sigma)} \end{aligned}$$

and for given Σ the value of c_2 which minimizes the relative risk function is given by

$$c_2 = \frac{b_1}{2b_2} = \frac{tr(\Sigma/p)}{|\Sigma|^{1/p}} \frac{d^*}{2}$$

with

$$\min_{c_2} \ RR(\hat{\Sigma} : \tilde{\Sigma}; c_2) = 1 - \frac{b_1^2}{4b_2}.$$

where d^* is given by (6.35).

Remark 6.4 It may be proved similarly that the Percentage Improvement in Risk (PIR) of the estimator $\hat{\Sigma}$ relative to the estimator $\tilde{\Sigma}$ is given by

$$PIR(\hat{\Sigma} : \tilde{\Sigma}; c_2) = 100(b_1 c_2 - b_2 c_2^2), \quad -\infty < c_2 \leq d_0$$

and for given Σ, we have

$$\max_{c_2} PIR(\hat{\Sigma} : \tilde{\Sigma}; c_2) = 25 \frac{b_1^2}{b_2}.$$

6.8 Proposed Estimator Dominating the Maximum Likelihood Estimator (Normal Case)

The maximum likelihood estimator of Σ is given by $\tilde{\Sigma}_1 = W/(n+1)$ (Anderson, 1984, p 63) where W has usual Wishart distribution given by (5.1). We want to compare the proposed estimator with the m.l.e.

When $c_1 = (n+1)^{-1}$ the conditions given by (6.9), (6.10a) and (6.10b) simplifies to

$$(i)\ \ d^* < c_2 < 0, \quad p > 2,$$

$$(ii)\ \ 0 < c_2 \leq d_0, \text{ if } 0 < d_0 < d^*, \quad p < 2,$$

$$(iii)\ \ 0 < c_2 < d^*, \text{ if } d^* \leq d_0 < \infty, \quad p < 2.$$

But we have considered the model (6.1) for $p \geq 2$. Thus by putting $c_1 = (n+1)^{-1}$ and then letting $\nu \to \infty$ in Theorems 6.1 , 6.2, 6.3 and 6.4 we have the following theorem:

Theorem 6.10 Consider the multivariate normal model for $p > 2$. Then the proposed estimator

$$\hat{\Sigma}_1 = W/(n+1) - c_2|W|^{1/p}I$$

where

$$-\infty < c_2 \leq d_0 = \frac{min\{w_{11}, \ldots, w_{pp}\}}{(n+1)|W|^{1/p}},$$

dominates the m.l.e. $\tilde{\Sigma}_1 = W/(n+1)$ under the squared error loss function given by (6.32) if c_2 satisfies the following conditions:

$$d^* < c_2 \leq \frac{d^*}{2}$$

where

$$d^{*} = -\frac{p-2}{(n+1)p}\frac{\Gamma_p(n/2+1/p)}{\Gamma_p(n/2+2/p)}.$$

The relative risk of the proposed estimator

$$\hat{\Sigma}_1 = W/(n+1) - c_2|W|^{1/p}I$$

relative to the m.l.e $\tilde{\Sigma}_1 = W/(n+1)$ is given by

$$RR(\hat{\Sigma}_1 : \tilde{\Sigma}_1; c_2) = 1 - b_1c_2 + b_2c_2^2$$

and for a given Σ, it minimizes at

$$c_2 = \frac{b_1}{2b_2} = \frac{tr(\Sigma/p)}{|\Sigma|^{1/p}}\frac{d^{*}}{2}$$

with

$$\min_{c_2} \; RR(\hat{\Sigma}_1 : \tilde{\Sigma}_1; c_2) = 1 - \frac{b_1^2}{4b_2},$$

where

$$\begin{aligned} b_1 &= b_2\frac{tr(\Sigma/p)}{|\Sigma|^{1/p}}d^{*}, \\ b_2 &= 4p\;\frac{\Gamma_p(n/2+2/p)}{\Gamma_p(n/2)}\frac{|\Sigma|^{2/p}}{R(\tilde{\Sigma}_1,\;\Sigma)}, \end{aligned}$$

and

$$R(\tilde{\Sigma}_1,\;\Sigma) = \frac{tr(\Sigma^2)}{n+1} + \frac{n(tr\Sigma)^2}{(n+1)^2} \qquad (6.39).$$

Some Special Cases

(1). We note that the m.l.e. $\tilde{\Sigma}_1$ and the proposed estimator $\hat{\Sigma}_1$ are identical in the bivariate case.

(2). For the trivariate case i.e for $p = 3$, the estimators are given by

$$\tilde{\Sigma}_1 = W/(n+1)$$

and

$$\hat{\Sigma}_1 = W/(n+1) - c_2|W|^{1/3}I$$

where

$$-\infty < c_2 \leq d_0 = \frac{min\{w_{11}, w_{22}, w_{33}\}}{(n+1)|A|^{1/3}}.$$

In this case it follows from Theorem 6.10 that the proposed estimator dominates the m.l.e if

$$d^* < c_2 \leq \frac{d^*}{2}$$

where d^* is given by

$$d^* = \frac{-1}{3(n+1)} \frac{\Gamma(n/2 - 1/6)\Gamma(n/2 + 1/3)\Gamma(n/2 - 2/3)}{\Gamma(n/2 + 1/6)\Gamma(n/2 - 1/3)\Gamma(n/2 + 2/3)}.$$

In this case the relative risk $RR(\hat{\Sigma}_1 : \tilde{\Sigma}_1, c_2)$ of our proposed estimator relative to usual estimator is given by

$$RR(\hat{\Sigma}_1 : \tilde{\Sigma}_1; c_2) = 1 - b_1 c_2 + b_2 c_2^2, \quad -\infty < c_2 \leq d_0$$

where

$$\begin{aligned} b_1 =& b_2 \, \frac{(\sigma_{11} + \sigma_{22} + \sigma_{33})/3}{(\sigma_{11}\sigma_{22}\sigma_{33})^{1/3}} \, d^*, \\ b_2 =& 12 \, \frac{\Gamma(n/2 + 2/3)\Gamma(n/2 + 1/6)\Gamma(n/2 - 1/3)}{\Gamma(n/2)\Gamma(n/2 - 1/2)\Gamma(n/2 - 1)} \, \frac{(\sigma_{11}\sigma_{22}\sigma_{33})^{2/3}}{R(\tilde{\Sigma}_1, \, \Sigma)} \end{aligned}$$

and

$$R(\tilde{\Sigma}_1, \, \Sigma) = \frac{\sigma_{11}^2 + \sigma_{22}^2 + \sigma_{33}^2}{n+1} + \frac{n(\sigma_{11} + \sigma_{22} + \sigma_{33})^2}{(n+1)^2}.$$

6.9 Proposed Estimator Dominating the Unbiased Estimator (Normal Case)

An unbiased estimator of Λ is given by $\tilde{\Sigma}_2 = W/n$ (Anderson, 1984, p 71) so that $c_1 = n^{-1}$.

The following theorem follows from Theorem 6.7, Theorem 6.8 and Theorem 6.9 by setting $c_1 = n^{-1}$.

Theorem 6.11 The proposed estimator

$$\hat{\Sigma}_2 = n^{-1}W - c_2|W|^{1/p}I$$

where

$$-\infty < c_2 \leq d_0 = \frac{min\{w_{11}, \ldots, w_{pp}\}}{n|W|^{1/p}}$$

dominates the unbiased estimator $\tilde{\Sigma}_2 = W/n$ for any c_2 satisfying the following conditions :

$$c_2 = d_0 \ \text{ for } \ 0 < d_0 \leq \frac{d^*}{2},$$

$$\text{and} \ \ \frac{d^*}{2} \leq c_2 \leq d_0 \ \text{ for } \ \frac{d^*}{2} < d_0 < d^*,$$

$$\text{while} \ \ \frac{d^*}{2} \leq c_2 < d^* \ \text{ for } \ d^* \leq d_0,$$

where

$$d^* = \frac{2}{np}\frac{\Gamma_p(n/2 + 1/p)}{\Gamma_p(n/2 + 2/p)}.$$

The relative risk function of the proposed estimator relative to the unbiased estimator is given by

$$RR(\hat{\Sigma}_2 : \tilde{\Sigma}_2; c_2) = 1 - b_1c_2 + b_2c_2^2, \quad -\infty < c_2 \leq d_0$$

and for given Σ

$$MRR(\hat{\Sigma}_2 : \tilde{\Sigma}_2) = \min_{c_2} \ RR(\hat{\Sigma}_2 : \tilde{\Sigma}_2; c_2) = 1 - \frac{b_1^2}{4b_2}$$

where

$$b_1 = b_2\frac{tr(\Sigma/p)}{|\Sigma|^{1/p}}d^*$$

$$b_2 = 4p\,\frac{\Gamma_p(n/2 + 2/p)}{\Gamma_p(n/2)}\frac{|\Sigma|^{2/p}}{R(\tilde{\Sigma}_2, \ \Sigma)}$$

and

$$R(\tilde{\Sigma}_2,\ \Sigma) = n^{-1}[tr(\Sigma^2) + (tr\Sigma)^2] \qquad (6.40).$$

Some Special Cases

(1). For the bivariate case i.e for $p = 2$, the unbiased and the proposed estimators of Σ are given by

$$\tilde{\Sigma}_2 = \frac{W}{n}$$

and

$$\hat{\Sigma}_2 = \frac{W}{n} - c_2|W|^{1/2}$$

respectively,

$$-\infty < c_2 \leq d_0 = \frac{min\{w_{11}, w_{22}\}}{n|W|^{1/2}}.$$

It follows from Theorem 6.11 that the proposed estimator dominates the unbiased estimator under the squared error loss function given by (6.32) if c_2 satisfies the following conditions on c_2 :

$$c_2 = d_0 \ \text{ for } \ 0 < d_0 \leq \frac{d^*}{2},$$
$$\text{and } \ \frac{d^*}{2} \leq c_2 \leq d_0 \ \text{ for } \ \frac{d^*}{2} < d_0 < d^*,$$
$$\text{while } \ \frac{d^*}{2} \leq c_2 < d^* \ \text{ for } \ d^* \leq d_0,$$

where d^* is given by $d^* = 2n^{-2}$.

It follows from Theorem 6.11 that when $p = 2$ the relative risk of the proposed estimator to the unbiased estimator is given by

$$RR(\hat{\Sigma}_2 : \tilde{\Sigma}_2; c_2) = 1 - b_1c_2 + b_2c_2^2, \quad -\infty < c_2 \leq d_0$$

and for given Σ, the minimum value of the relative risk function is given by

$$\min_{c_2} \ RR(\hat{\Sigma}_2 : \tilde{\Sigma}_2; c_2) = 1 - \frac{b_1^2}{4b_2}$$

where

$$b_1 = b_2 \frac{(\sigma_{11}+\sigma_{22})/2}{(\sigma_{11}\sigma_{22})^{1/2}} d^*$$

$$b_2 = 2n(n-1) \frac{\sigma_{11}\sigma_{22}}{R(\tilde{\Sigma}_2, \Sigma)}$$

with

$$R(\tilde{\Sigma}_2, \Sigma_2) = 2n^{-1}(\sigma_{11}^2 + \sigma_{22}^2 + \sigma_{11}\sigma_{22}).$$

We remark that in this case the relative risk depends on the characteristic roots of Σ only through their ratio.

(2). For the trivariate case i.e for $p = 3$, the unbiased and proposed estimators of Σ are given by

$$\tilde{\Sigma}_2 = \frac{W}{n}$$

and

$$\hat{\Sigma}_2 = \frac{W}{n} - c_2|A|^{1/3} I$$

respectively, where

$$-\infty < c_2 \leq d_0 = \frac{min\{w_{11}, w_{22}, w_{33}\}}{n|W|^{1/2}}.$$

It follows from Theorem 6.11 that for the trivariate case the proposed estimator dominates the unbiased estimator under the squared error loss function given by (6.32) for any c_2 satisfying the following conditions :

$$c_2 = d_0 \quad \text{for} \quad 0 < d_0 \leq \frac{d^*}{2},$$

$$\text{and} \quad \frac{d^*}{2} \leq c_2 \leq d_0 \quad \text{for} \quad \frac{d^*}{2} < d_0 < d^*,$$

$$\text{while} \quad \frac{d^*}{2} \leq c_2 < d^* \quad \text{for} \quad d^* \leq d_0$$

where

$$d^* = \frac{2}{3n} \frac{\Gamma(n/2-1/6)\Gamma(n/2+1/3)\Gamma(n/2-2/3)}{\Gamma(n/2+1/6)\Gamma(n/2-1/3)\Gamma(n/2+2/3)}.$$

It also follows from Theorem 6.11 that for $p = 3$ the relative risk of the proposed estimator relative to the unbiased estimator is given by

$$RR(\hat{\Sigma}_2 : \tilde{\Sigma}_2; c_2) = 1 - b_1 c_2 + b_2 c_2^2, \quad -\infty < c_2 \le d_0$$

and for given Σ, the minimum value of relative risk function is given by

$$\min_{c_2} \; RR(\hat{\Sigma}_2 : \tilde{\Sigma}_2; c_2) = 1 - \frac{b_1^2}{4b_2}$$

where

$$\begin{aligned} b_1 =& b_2 \frac{(\sigma_{11} + \sigma_{22} + \sigma_{33})/3}{(\sigma_{11}\sigma_{22}\sigma_{33})^{1/3})} d^*, \\ b_2 =& \frac{12(\sigma_{11}\sigma_{22}\sigma_{33})^{2/3}}{R(\tilde{\Sigma}_2, \; \Sigma)} \; \frac{\Gamma(n/2 + 2/3)\Gamma(n/2 + 1/6)\Gamma(n/2 - 1/3)}{\Gamma(n/2)\Gamma(n/2 - 1/2)\Gamma(n/2 - 1)}, \end{aligned}$$

with

$$R(\tilde{\Sigma}_2, \; \Sigma) = n^{-1}[(\sigma_{11}^2 + \sigma_{22}^2 + \sigma_{33}^2) + (\sigma_{11} + \sigma_{22} + \sigma_{33})^2].$$

6.10 Simultaneous Estimation of Characteristic Roots of Scale Matrix

In this section we estimate the characteristic roots of the scale matrix Λ of the multivariate t-model and of the scale matrix Σ of the multivariate normal model.

6.10.1 Simultaneous Estimation of Characteristic Roots of the Scale Matrix of the Multivariate t-Model

We now prove that the simultaneous estimation of the characteristic roots $\xi_1, \xi_2, \ldots, \xi_p$ of Λ under a squared error loss function is similar to the estimation of the scale matrix under the loss function given by (6.2) discussed in Theorem 6.1 and 6.2. Let

$$m = (m_1, m_2, \ldots, m_p)'$$

and

$$\xi = (\xi_1, \xi_2, \ldots, \xi_p)'.$$

In estimating the characteristic roots $\xi_i, (i = 1, 2, \ldots, p)$ we consider the following loss function:

$$L(u(m),\ \xi) = \sum_{i=1}^{p} (u(m_i) - \xi_i)^2 \tag{6.41}$$

where $u(m_i), (i = 1, 2, \ldots, p)$ is any estimator of $\xi_i, (i = 1, 2, \ldots, p)$, $u(m) = (u(m_1), u(m_2), \ldots, u(m_p))$, and the risk function is defined as usual by taking expectation of the loss function i.e.

$$R(u(m),\ \xi) = E[L(u(m),\ \xi)].$$

Further consider the following estimators of

$$\xi = (\xi_1, \xi_2, \ldots, \xi_p)' :$$

$$\text{usual estimator}\ \ \tilde{\xi} = (\tilde{\xi}_1, \tilde{\xi}_2, \ldots, \tilde{\xi}_p)', \tag{6.42}$$

$$\text{proposed estimator}\ \ \hat{\xi} = (\hat{\xi}_1, \hat{\xi}_2, \ldots, \hat{\xi}_p)', \tag{6.43}$$

where

$$\tilde{\xi}_i = c_1\ m_i,\ (i = 1, 2, \ldots, p),$$

$$\hat{\xi}_i = c_1\ m_i - c_2\ (m_1 m_2 \ldots m_p)^{1/p},\ (i = 1, 2, \ldots, p),$$

c_1 is a fixed positive number and

$$-\infty < c_2 \le d_0 = \frac{c_1 m_p}{(m_1 m_2 \ldots m_p)^{1/p}}.$$

We now prove a theorem that the proposed estimator of the characteristic roots of Λ dominates its usual estimator.

Theorem 6.12 Consider the multivariate t-model given by (6.1) for $\nu > 4$. Then the proposed estimator $\hat{\xi}$ (defined by (6.43)) dominates the usual estimator $\tilde{\xi}$ (defined by (6.42)) under the squared error loss function given by (6.41) if c_2 satisfies the following conditions (exactly the same as 6.9, 6.10a and 6.10b):

1. For a given c_1 satisfying

$$c_1 < \frac{\nu - 4}{\nu} \frac{p}{np+2} \quad (d^* < 0),$$

c_2 satisfies $d^* < c_2 < 0$.

2. For a given c_1 satisfying

$$c_1 > \frac{\nu - 4}{\nu} \frac{p}{np+2} \quad (d^* > 0),$$

c_2 satisfies the following scheme:

$$(i)\ 0 < c_2 \leq d_0 \quad \text{if } 0 < d_0 < d^*,$$

$$(ii)\ 0 < c_2 < d^* \quad \text{if } d^* \leq d_0 < \infty$$

where d^* is given by (6.11).

where

$$d^* = \left(c_1 \frac{np+2}{p} - \frac{\nu-4}{\nu}\right) \frac{\Gamma_p(n/2+1/p)}{\Gamma_p(n/2+2/p)}.$$

Proof. Let us suppose that

$$D(\xi,\ c_2) = R(\hat{\xi},\ \xi) - R(\tilde{\xi},\ \xi).$$

Then

$$\begin{aligned} D(\xi,\ c_2) =& E \sum_{i=1}^{p} \left[(\hat{\xi}_i - \xi_i)^2 - (\tilde{\xi}_i - \xi_i)^2 \right] \\ =& E \sum_{i=1}^{p} \left[\hat{\xi}_i^2 - \tilde{\xi}_i^2 - 2(\hat{\xi}_i - \tilde{\xi}_i)\xi_i \right]. \end{aligned}$$

Hence by virtue of

$$\hat{\xi}_i^2 = \{\tilde{\xi}_i^2 - c_2|A|^{1/p}\}^2$$
$$= \tilde{\xi}_i^2 - 2c_2\tilde{\xi}_i^2|A|^{1/p} + c_2^2|A|^{2/p},$$

for $i = 1, 2, \dots, p$, we have

$$D(\xi,\ c_2) = E\sum_{i=1}^{p}\left[-2c_2\tilde{\xi}_i^2|A|^{1/p} + c_2^2|A|^{1/p} + 2c_2|A|^{1/p}\xi_i\right]$$
$$= -2c_1c_2\ Etr(|A|^{1/p}A) + 2c_2\ E(|A|^{1/p})tr(\Lambda) + pc_2^2\ E(|A|^{2/p})$$

which is exactly the same as that we have found in Theorem 6.1. Hence the rest of the proof is omitted because of the similarity to that of Theorem 6.1.

In the light of Theorem 6.2 we have the following version of Theorem 6.12.

Theorem 6.13 Consider the multivariate t-model given by (6.1) for $\nu > 4$. Then the proposed estimator $\hat{\xi}$ (defined by (6.43)) dominates the usual estimator $\tilde{\xi}$ (defined by (6.42)) under the squared error loss function given by (6.41) for any c_2 satisfying the following conditions :

1. For a given c_1 satisfying

$$c_1 < \frac{\nu-4}{\nu}\frac{p}{np+2} \quad (i.e.\ d^* < 0),$$

c_2 must satisfy

$$d^* < c_2 \leq \frac{d^*}{2}.$$

2. For a given c_1 satisfying

$$c_1 > \frac{\nu-4}{\nu}\frac{p}{np+2} \quad (i.e.\ d^* > 0),$$

c_2 satisfies the following scheme:

$$c_2 = d_0 \ \text{ for } \ 0 < d_0 \leq \frac{d^*}{2},$$
$$\text{and } \ \frac{d^*}{2} \leq c_2 \leq d_0 \ \text{ for } \ \frac{d^*}{2} < d_0 < d^*,$$
$$\text{while } \ \frac{d^*}{2} \leq c_2 < d^* \ \text{ for } \ d^* \leq d_0,$$

where d^* is defined by (6.11).

The following section specializes the above theorems to the case of the estimation of characteristic roots of the multivariate normal distribution.

6.10.2 Simultaneous Estimation of Characteristic Roots of the Scale Matrix of the Multivariate Normal Distribution

Suppose that we want to estimate the characteristic roots of the scale matrix Σ of the multivariate normal distribution $N_p(\theta, \Sigma)$. We also suppose that the p-dimensional vectors $X_1, X_2, \ldots, X_N$ are independently and identically distributed according to $N_p(\theta, \Sigma)$.

Let

$$l = (l_1, l_2, \ldots, l_p)'$$

and

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_p)'.$$

In estimating the characteristic roots $\alpha_i, (i = 1, 2, \ldots, p)$ we consider the following loss function:

$$L(u(l),\ \alpha) = \sum_{i=1}^{p} (u(l_i) - \alpha_i)^2 \tag{6.44}$$

where $u(l_i), (i = 1, 2, \ldots, p)$ is any estimator of $\alpha_i, (i = 1, 2, \ldots, p)$, $u(l) = (u(l_1), u(l_2), \ldots, u(l_p))'$, and the risk function is defined as usual by taking expectation of the loss function i.e.

$$R(u(l),\ \alpha) = E[L(u(l),\ \alpha)].$$

Further consider the following estimators of $\xi = (\alpha_1, \alpha_2, \ldots, \alpha_p)'$:

$$\text{usual estimator } \ \tilde{\alpha} = (\tilde{\alpha}_1, \tilde{\alpha}_2, \ldots, \tilde{\alpha}_p)', \tag{6.45}$$

$$\text{proposed estimator } \hat{\alpha} = (\hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_p)', \tag{6.46}$$

where

$$\tilde{\alpha}_i = c_1 \; l_i, \; (i = 1, 2, \ldots, p),$$

$$\hat{\alpha}_i = c_1 \; l_i - c_2 \; (l_1 l_2 \ldots l_p)^{1/p}, \; (i = 1, 2, \ldots, p),$$

c_1 is a fixed positive number and

$$-\infty < c_2 \leq d_0 = \frac{c_1 l_p}{(l_1 l_2 \ldots l_p)^{1/p}}.$$

We now prove a dominance theorem that the proposed estimator $\hat{\xi}$ dominates the usual estimator $\tilde{\xi}$ of ξ.

Theorem 6.14 Suppose that the p-dimensional vectors $X_1, X_2, \ldots, X_N$ are independently and identically distributed according to $N_p(\theta, \Sigma)$. Then the proposed estimator $\hat{\alpha}$ (defined by (6.46)) dominates the usual estimator $\tilde{\xi}$ (defined by (6.45)) under the squared error loss function given by (6.44) for any c_2 satisfying the following conditions :

1. For a given c_1 satisfying

$$c_1 < \frac{p}{np+2} \;\; (i.e. \; d^* < 0),$$

c_2 satisfies $d^* < c_2 < 0$.

2. For a given c_1 satisfying

$$c_1 > \frac{p}{np+2} \;\; (i.e. \; d^* > 0),$$

c_2 satisfies the following scheme:

$$(i) \;\; 0 < c_2 \leq d_0 \;\; \text{for} \;\; 0 < d_0 < d^*,$$

$$(ii)\ \ 0 < c_2 < d^* \ \text{ for } \ d^* \leq d_0$$

where d^* is defined by (6.35).

Remark 6.5 The above theorem becomes similar to Theorem 2.1 of Dey (1988) if ξ, c_1, c_2, n and d^* are replaced by λ, c, b, k and

$$2d\ \frac{c_{p,k+2/p}}{c_{p,k+4/p}}$$

respectively.

By the arguments as used in establishing Theorem 6.2 we have the following stronger version of Theorem 6.14.

Theorem 6.15 Suppose that the p-dimensional vectors $X_1, X_2, \ldots, X_N$ are independently and identically distributed according to $N_p(\theta, \Sigma)$. Then the proposed estimator $\hat{\alpha}$ (defined by (6.46)) dominates the usual estimator $\tilde{\xi}$ (defined by (6.45)) under the squared error loss function given by (6.44) for any c_2 satisfying the following conditions :

1. For a given c_1 satisfying

$$c_1 < p/(np+2) \ \ (i.e.\ d^* < 0),$$

c_2 satisfies

$$d^* < c_2 \leq \frac{d^*}{2}.$$

2. For a given c_1 satisfying

$$c_1 > p/(np+2) \ \ (i.e.\ d^* > 0),$$

c_2 satisfies the following scheme:

$$c_2 = d_0 \ \text{ for } \ 0 < d_0 \leq \frac{d^*}{2},$$

$$\text{and} \ \ \frac{d^*}{2} \leq c_2 \leq d_0 \ \text{ for } \ \frac{d^*}{2} < d_0 < d^*,$$

$$\text{while} \ \ \frac{d^*}{2} \leq c_2 < d^* \ \text{ and for } \ d^* \leq d_0,$$

where d^* is defined by (6.35).

6.11 Estimation of the Trace of the Scale Matrix of the Multivariate t-Model

In this section we propose estimator for the trace of the scale matrix of the multivariate t-model. Let $\delta = tr(\Lambda)$. In estimating δ by an estimator $u(A)$ we consider the following loss function

$$L(u(A),\ \delta) = (u(A) - \delta)^2 \tag{6.47}$$

where $u(A)$ is any estimator of tr(Λ). The risk function is defined as usual by taking expectation over the loss function i.e.

$$R(u(A),\ \delta) = E[L(u(A),\ \delta)].$$

Furthermore consider the following two estimators of $\delta = tr\Lambda$:

$$\text{usual estimator}\ \ \tilde{\delta} = c_1 tr(A), \tag{6.48}$$

$$\text{proposed estimator}\ \ \hat{\delta} = c_1 tr(A) - c_2\ p|A|^{1/p}, \tag{6.49}$$

where c_1 is a fixed positive constant and

$$-\infty < c_2 \le d_0 = \frac{c_1 tr(A/p)}{|A|^{1/p}}.$$

We now prove a dominance theorem that the proposed estimator $\hat{\delta}$ dominates the usual estimator $\tilde{\delta}$ of $\delta = tr(\Lambda)$.

Theorem 6.16 Consider the multivariate t-model given by (6.1) for $\nu > 4$. Then the proposed estimator $\hat{\delta}$ (defined by (6.49)) dominates over the usual estimator $\tilde{\delta}$ (defined by (6.48)) under the squared error loss function given by (6.47) for any c_2 satisfying the conditions stated in Theorem 6.1.

Proof. Let $D(\Lambda,\ c_2) = R(\hat{\delta},\ \delta; c_2) - R(\tilde{\delta},\ \delta)$.

Then it is easy to show that

$$\begin{aligned} D(\Lambda,\ c_2) =& E\left[(\hat{\delta}-\delta)^2 - (\tilde{\delta}-\delta)^2\right] \\ =& E\left[\hat{\delta}^2 - \tilde{\delta}^2 - 2(\hat{\delta}-\tilde{\delta})\delta\right]. \end{aligned}$$

Hence by virtue of

$$\begin{aligned} \hat{\delta}^2 =& \{\tilde{\delta} - c_2 p|A|^{1/p}\}^2 \\ =& \tilde{\delta}^2 - 2c_2 p\tilde{\delta}|A|^{1/p} + c_2^2 p^2 |A|^{2/p} \end{aligned}$$

we have

$$\begin{aligned} D(\Lambda,\ c_2) =& E\left[-2c_2 p\tilde{\delta}|A|^{1/p} + c_2^2 p^2 |A|^{2/p} - 2(-c_2 p|A|^{1/p})tr(\Lambda)\right] \\ =& p\ E\left[-2c_1 c_2 |A|^{1/p} tr(A) + c_2^2 p|A|^{2/p} + 2c_2|A|^{1/p} tr(\Lambda)\right] \end{aligned}$$

which is p-times $D(\Lambda,\ c_2)$ obtained in course of proving Theorem 6.1. Hence the rest of the proof is omitted because of the similarity to that of Theorem 6.1.

By the argument as used in establishing Theorem 6.2 we have the following theorem from Theorem 6.16.

Theorem 6.17 Consider the multivariate t-model given by (6.1) for $\nu > 4$. Then the proposed estimator $\hat{\beta}$ (defined by (6.49)) dominates over the usual estimator $\tilde{\beta}$ (defined by (6.48)) under the squared error loss function given by (6.47) for any c_2 satisfying the following conditions:

1. For a given c_1 satisfying

$$c_1 < \frac{\nu-4}{\nu}\frac{p}{np+2} \quad (i.e.\ d^* < 0),$$

c_2 satisfies

$$d^* < c_2 \le \frac{d^*}{2}.$$

2. For a given c_1 satisfying

$$c_1 > \frac{\nu - 4}{\nu} \frac{p}{np+2} \quad (i.e.\ d^* > 0),$$

c_2 satisfies the following scheme:

$$c_2 = d_0 \quad \text{for} \quad 0 < d_0 \le \frac{d^*}{2},$$

$$\text{and} \quad \frac{d^*}{2} \le c_2 \le d_0 \quad \text{for} \quad \frac{d^*}{2} < d_0 < d^*,$$

$$\text{while} \quad \frac{d^*}{2} \le c_2 < d^* \quad \text{for} \quad d^* \le d_0$$

where d^* is defined by (6.11).

6.12 Estimation of the Inverted Scale Matrix of the Multivariate t-Model

In this section we develop estimator for inverted scale matrix i.e. for $\Psi = \Lambda^{-1}$. To estimate the inverted scale matrix Ψ, we consider the following loss function

$$L(u(A),\ \Psi) = tr[(u(A) - \Psi)^2] \tag{6.50}$$

where $u(A)$ is any suitable estimator of Ψ. The risk function is defined as usual by

$$R(u(A),\ \Psi) = E[tr\{(u(A) - \Psi)^2\}].$$

Further consider the following estimators of Ψ:

$$\text{usual estimator} \quad \tilde{\Psi} = c_1 A^{-1}, \tag{6.51}$$

$$\text{proposed estimator} \quad \hat{\Psi} = c_1 A^{-1} - c_2 |A|^{-1/p} I, \tag{6.52}$$

where

$$A = \sum_{j=1}^{N} (X_j - \overline{X})(X_j - \overline{X})',$$

c_1 is a fixed positive number and

$$-\infty < c_2 \le k_0 = \frac{c_1 A^{-1}}{|A|^{-1/p}}.$$

We now prove a dominance theorem that the proposed estimator $\hat{\Psi}$ dominates the usual estimator $\tilde{\Psi}$.

Theorem 6.18 Consider the multivariate t-model given by (6.1). Then under the loss function given by (6.50), the proposed estimator $\hat{\Psi}$ (defined by (6.52)) dominates the usual estimator $\tilde{\Psi}$ (defined by (6.51)) for $n \ge 4$ and any c_2 satisfying the following conditions:

1. For a given c_1 satisfying

$$c_1 < \frac{\nu}{\nu+2}(n-p-1-2/p),$$

c_2 satisfies $k^* < c_2 < 0$.

2. For a given c_1 satisfying

$$c_1 > \frac{\nu}{\nu+2}(n-p-1-2/p),$$

c_2 satisfies the following scheme:

$$(i)\ 0 < c_2 \le k_0 \quad \text{if } 0 < k_0 < k^*,$$

$$(ii)\ 0 < c_2 < k^* \quad \text{if } k^* \le k_0 < \infty$$

where

$$k^* = 4\left(\frac{c_1}{n-p-1-2/p} - \frac{\nu}{\nu+2}\right)\frac{\Gamma_p(n/2-1/p)}{\Gamma_p(n/2-2/p)}. \qquad (6.53).$$

Proof. Let

$$D(\Lambda,\ c_2) = R(\hat{\Psi},\ \Psi) - R(\tilde{\Psi},\ \Psi).$$

Then $\hat{\Psi}$ dominates $\tilde{\Psi}$ if and only if $D(\Lambda, c_2) < 0$. A simple calculation shows that

$$\begin{aligned} D(\Lambda,\ c_2) =& E\left[\hat{\Psi} - \Psi)^2 - (\tilde{\Psi} - \Psi)^2\right] \\ =& Etr\left[\hat{\Psi}^2 - \tilde{\Psi}^2 - 2(\hat{\Psi} - \hat{\Psi})\Psi\right] \end{aligned}.$$

Hence by virtue of

$$\begin{aligned} \hat{\Psi}^2 =& \{\tilde{\Psi} - c_2|A|^{-1/p}I\}^2 \\ =& \tilde{\Psi}^2 - 2c_2\tilde{\Psi}|A|^{-1/p}I + c_2^2|A|^{-2/p}I, \end{aligned}$$

we have

$$\begin{aligned} D(\Lambda,\ c_2) =& Etr\left[-2c_2\tilde{\Psi}|A|^{-1/p}I + c_2^2|A|^{-2/p}I - 2(-c_2|A|^{-1/p}I)\Psi\right] \\ =& -2c_1c_2Etr(|A|^{-1/p}A^{-1}) + c_2^2pE(|A|^{-2/p}) + 2c_2E(|A|^{-1/p})tr(\Psi). \end{aligned}$$

It then follows from Lemma 5.1 and Lemma 6.3 that

$$\begin{aligned} D(\Lambda,\ c_2) =& -2c_1c_2\ tr\left[\frac{\nu+2}{2\nu(n-p-1-2/p)}\frac{\Gamma_p(n/2-1/p)}{\Gamma_p(n/2)}|\Lambda|^{-1/p}\Psi\right] \\ &+c_2^2\ p\left[\frac{\nu+2}{4\nu}\frac{\Gamma(n/2-2/p)}{\Gamma(n/2)}|\Lambda|^{-2/p}\right] \\ &+2c_2\left[\frac{\Gamma_p(n/2-1/p)}{2\Gamma_p(n/2)}|\Lambda|^{-1/p}\right]tr(\Lambda^{-1}) \\ =& \frac{\nu+2}{\nu}\ \frac{c_2p}{4}\ \frac{\Gamma_p(n/2-2/p)}{\Gamma_p(n/2)}\ |\Lambda|^{-2/p} \\ &\times\left[\left(\frac{\nu}{\nu+2} - \frac{c_1}{n-p-1-2/p}\right)\frac{tr(\Lambda^{-1})}{|\Lambda|^{-1/p}} + c_2\right] \\ =& \frac{(\nu+2)p}{4\nu}\ \frac{\Gamma_p(n/2-2/p)}{\Gamma_p(n/2)}\ |\Lambda|^{-2/p}c_2\left[c_2 - \frac{tr(\Lambda^{-1})}{p\ |\Lambda|^{-1/p}}k^*\right] \\ =& \frac{(\nu+2)p}{4\nu}\ \frac{\Gamma_p(n/2-2/p)}{\Gamma_p(n/2)}\ |\Psi|^{2/p}c_2\left[c_2 - \frac{tr(\Psi/p)}{|\Psi|^{1/p}}k^*\right] \end{aligned}$$

where k^* is stated in the theorem.

Clearly $D(\Lambda,\ c_2) < 0$ so that $\hat{\Psi}$ dominates $\tilde{\Psi}$ if and only if

$$\frac{tr(\Psi/p)}{|\Psi|^{1/p}}k^* < c_2 < 0, \quad \text{or} \quad 0 < c_2 < \frac{tr(\Psi/p)}{|\Psi|^{1/p}}\ k^*.$$

The proof then follows by the use of the inequality between the arithmetic mean and the geometric mean of the characteristic roots of Ψ (cf. Theorem 6.1)

By the arguments as used in establishing Theorem 6.2 we have the following theorem from Theorem 6.18.

Theorem 6.19 Consider the multivariate t-model given by (6.1). Then under the loss function given by (6.52), the proposed estimator $\hat{\Psi}$ (defined by (6.54)) dominates the usual estimator $\tilde{\Psi}$ (defined by (6.53)) for $n \geq 4$ if c_2 satisfies the following conditions:

1. For a given c_1 satisfying

$$c_1 < \frac{\nu}{\nu+2}(n-p-1-2/p),$$

c_2 satisfies

$$k^* < c_2 \leq \frac{k^*}{2}.$$

2. For a given c_1 satisfying

$$c_1 > \frac{\nu}{\nu+2}(n-p-1-2/p),$$

c_2 satisfies the following scheme:

$$\begin{aligned} c_2 = k_0 \quad &\text{for } 0 < k_0 \leq \frac{k^*}{2}, \\ \text{and} \quad \frac{k^*}{2} \leq c_2 \leq k_0 \quad &\text{for } \frac{k^*}{2} < k_0 < k^*, \\ \text{while} \quad \frac{k^*}{2} \leq c_2 < k^* \quad &\text{for } k^* \leq k_0 \end{aligned}$$

where k^* is defined by (6.53).

Table 6.1: MRR (Minimum Relative Risk) of proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 2$ and $\Lambda = diag(1, 1)$.

n	ν	MRR	d^*	c_2(opt)
5	5	0.42970298	0.31999999	0.16000000
	10	0.78345865	0.16000000	0.08000000
	15	0.89152543	0.10666666	0.05333333
	20	0.93571429	0.08000000	0.04000000
	25	0.95767818	0.06400000	0.03200000
	30	0.97009346	0.05333333	0.02666667
	35	0.97776920	0.04571428	0.02285714
	40	0.98283671	0.04000000	0.02000000
	45	0.98635394	0.03555555	0.01777778
	50	0.98889317	0.03200000	0.01600000
10	5	0.30664549	0.16000000	0.08000000
	10	0.68930102	0.08000000	0.04000000
	15	0.83005951	0.05333333	0.02666667
	20	0.89406616	0.04000000	0.02000000
	25	0.92797917	0.03200000	0.01600000
	30	0.94796943	0.02666667	0.01333333
	35	0.96069745	0.02285714	0.01142857
	40	0.96928501	0.02000000	0.01000000
	45	0.97534561	0.01777778	0.00888889
	50	0.97977904	0.01600000	0.00800000
15	5	0.25304156	0.10666667	0.05333333
	10	0.63549454	0.05333333	0.02666667
	15	0.78946287	0.03555556	0.01777778
	20	0.86408798	0.02666667	0.01333333
	25	0.90536825	0.02133333	0.01066667
	30	0.93045166	0.01777778	0.00888889
	35	0.94678397	0.01523810	0.00761905
	40	0.95799344	0.01333333	0.00666667
	45	0.96601172	0.01185185	0.00592593
	50	0.97194136	0.01066667	0.00533333

Table 6.1 (continued): MRR (Minimum Relative Risk) of proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 2$ and $\Lambda = diag(1,1)$.

n	ν	MRR	d^*	c_2(opt)
20	5	0.22263713	0.08000001	0.04000000
	10	0.60023855	0.04000000	0.02000000
	15	0.76028323	0.02666667	0.01333333
	20	0.84118650	0.02000000	0.01000000
	25	0.88734686	0.01600000	0.00800000
	30	0.91605050	0.01333333	0.00666667
	35	0.93507356	0.01142857	0.00571429
	40	0.94831287	0.01000000	0.00500000
	45	0.95788998	0.00888889	0.00444444
	50	0.96503797	0.00800000	0.00400000
25	5	0.20299665	0.06400000	0.03200000
	10	0.57528797	0.03200000	0.01600000
	15	0.73824343	0.02133333	0.01066667
	20	0.82307525	0.01600000	0.00800000
	25	0.87260932	0.01280000	0.00640000
	30	0.90397159	0.01066667	0.00533333
	35	0.92505587	0.00914286	0.00457143
	40	0.93989998	0.00800000	0.00400000
	45	0.95074027	0.00711111	0.00355556
	50	0.95889537	0.00640000	0.00320000
30	5	0.18924950	0.05333334	0.02666667
	10	0.55668430	0.02666667	0.01333333
	15	0.72099396	0.01777778	0.00888889
	20	0.80838118	0.01333333	0.00666667
	25	0.86032262	0.01066667	0.00533333
	30	0.89368642	0.00888889	0.00444444
	35	0.91638122	0.00761905	0.00380952
	40	0.93251476	0.00666667	0.00333333
	45	0.94439252	0.00592593	0.00296296
	50	0.95338963	0.00533333	0.00266667

Table 6.2: MRR (Minimum Relative Risk) of the proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 2$ and $\Lambda = diag(2, 1)$.

n	ν	MRR	d^*	c_2(opt)
	5	0.47034484	0.31999999	0.16970562
	10	0.79550296	0.16000000	0.08485281
	15	0.89691275	0.10666666	0.05656854
	20	0.93872341	0.08000000	0.04242641
5	25	0.95959076	0.06400000	0.03394112
	30	0.97141439	0.05333333	0.02828427
	35	0.97873558	0.04571428	0.02424366
	40	0.98357414	0.04000000	0.02121320
	45	0.98693507	0.03555555	0.01885618
	50	0.98936288	0.03200000	0.01697056
	5	0.36331308	0.16000000	0.08485281
	10	0.71002961	0.08000000	0.04242641
	15	0.84003265	0.05333333	0.02828427
	20	0.89978528	0.04000000	0.02121320
10	25	0.93164975	0.03200000	0.01697056
	30	0.95051250	0.02666667	0.01414214
	35	0.96255856	0.02285714	0.01212183
	40	0.97070393	0.02000000	0.01060660
	45	0.97646218	0.01777778	0.00942809
	50	0.98068007	0.01600000	0.00848528
	5	0.31773945	0.10666667	0.05656854
	10	0.66228505	0.05333333	0.02828427
	15	0.80320083	0.03555556	0.01885618
	20	0.87222817	0.02666667	0.01414214
15	25	0.91068608	0.02133333	0.01131371
	30	0.93417311	0.01777778	0.00942809
	35	0.94952335	0.01523810	0.00808122
	40	0.96008909	0.01333333	0.00707107
	45	0.96766407	0.01185185	0.00628539
	50	0.97327617	0.01066667	0.00565685

Table 6.2 (continued): MRR (Minimum Relative Risk) of the proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 2$ and $\Lambda = diag(2, 1)$.

n	ν	MRR	d^*	c_2(opt)
20	5	0.29216473	0.08000001	0.04242641
	10	0.63142229	0.04000000	0.02121320
	15	0.77708015	0.02666667	0.01414214
	20	0.85143617	0.02000000	0.01060660
	25	0.89416557	0.01600000	0.00848528
	30	0.92087819	0.01333333	0.00707107
	35	0.93865471	0.01142857	0.00606092
	40	0.95106677	0.01000000	0.00530330
	45	0.96006910	0.00888889	0.00471405
	50	0.96680269	0.00800000	0.00424264
25	5	0.27574799	0.06400000	0.03394113
	10	0.60977854	0.03200000	0.01697056
	15	0.75754291	0.02133333	0.01131371
	20	0.83514988	0.01600000	0.00848528
	25	0.88077742	0.01280000	0.00678823
	30	0.90982144	0.01066667	0.00565685
	35	0.92943051	0.00914286	0.00484873
	40	0.94328375	0.00800000	0.00424264
	45	0.95342926	0.00711111	0.00377124
	50	0.96107992	0.00640000	0.00339411
30	5	0.26430512	0.05333334	0.02828427
	10	0.59374575	0.02666667	0.01414214
	15	0.74236581	0.01777778	0.00942809
	20	0.82203677	0.01333333	0.00707107
	25	0.86969853	0.01066667	0.00565685
	30	0.90047363	0.00888889	0.00471405
	35	0.92149704	0.00761905	0.00404061
	40	0.93649529	0.00666667	0.00353553
	45	0.94757003	0.00592593	0.00314270
	50	0.95598003	0.00533333	0.00282843

Table 6.3: MRR (Minimum Relative Risk) of proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 2$ and $\Lambda = diag(25, 1)$.

n	ν	MRR	d^*	c_2(opt)
5	5	0.64096928	0.31999999	0.41599999
	10	0.85083208	0.16000000	0.20800000
	15	0.92256032	0.10666666	0.13866666
	20	0.95330058	0.08000000	0.10400000
	25	0.96894859	0.06400000	0.08320000
	30	0.97791842	0.05333333	0.06933333
	35	0.98351451	0.04571428	0.05942857
	40	0.98723237	0.04000000	0.05200000
	45	0.98982474	0.03555555	0.04622222
	50	0.99170282	0.03200000	0.04160000
10	5	0.58791162	0.16000000	0.20800000
	10	0.79931522	0.08000000	0.10400000
	15	0.88502778	0.05333333	0.06933333
	20	0.92631405	0.04000000	0.05200000
	25	0.94898740	0.03200000	0.04160000
	30	0.96267742	0.02666667	0.03466667
	35	0.97154435	0.02285714	0.02971429
	40	0.97760354	0.02000000	0.02600000
	45	0.98192204	0.01777778	0.02311111
	50	0.98510587	0.01600000	0.02080000
15	5	0.56753396	0.10666667	0.13866666
	10	0.77336593	0.05333333	0.06933333
	15	0.86287127	0.03555556	0.04622222
	20	0.90869561	0.02666667	0.03466667
	25	0.93503234	0.02133333	0.02773333
	30	0.95148483	0.01777778	0.02311111
	35	0.96242281	0.01523810	0.01980952
	40	0.97005235	0.01333333	0.01733333
	45	0.97558073	0.01185185	0.01540741
	50	0.97971226	0.01066667	0.01386667

Table 6.3 (continued): MRR (Minimum Relative Risk) of proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 2$ and $\Lambda = diag(25, 1)$.

n	ν	MRR	d^*	c_2(opt)
20	5	0.55661791	0.08000001	0.10400001
	10	0.75754544	0.04000000	0.05200000
	15	0.84807174	0.02666667	0.03466667
	20	0.89613734	0.02000000	0.02600000
	25	0.92460013	0.01600000	0.02080000
	30	0.94281015	0.01333333	0.01733333
	35	0.95515111	0.01142857	0.01485714
	40	0.96389507	0.01000000	0.01300000
	45	0.97031361	0.00888889	0.01155556
	50	0.97516292	0.00800000	0.01040000
25	5	0.54979630	0.06400000	0.08320000
	10	0.74686511	0.03200000	0.04160000
	15	0.83746072	0.02133333	0.02773333
	20	0.88671000	0.01600000	0.02080000
	25	0.91648642	0.01280000	0.01664000
	30	0.93587255	0.01066667	0.01386667
	35	0.94920361	0.00914286	0.01188571
	40	0.95876558	0.00800000	0.01040000
	45	0.96585793	0.00711111	0.00924444
	50	0.97126422	0.00640000	0.00832000
30	5	0.54512397	0.05333334	0.06933334
	10	0.73916295	0.02666667	0.03466667
	15	0.82947366	0.01777778	0.02311111
	20	0.87936629	0.01333333	0.01733333
	25	0.90999013	0.01066667	0.01386667
	30	0.93019297	0.00888889	0.01155556
	35	0.94424462	0.00761905	0.00990476
	40	0.95442271	0.00666667	0.00866667
	45	0.96203635	0.00592593	0.00770370
	50	0.96788304	0.00533333	0.00693333

Table 6.4: MRR (Minimum Relative Risk) of proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 3$ and $\Lambda = diag(1, 1, 1)$.

n	ν	MRR	d^*	c_2(opt)
	5	0.55672033	0.37651530	0.18825765
	10	0.86523787	0.17420857	0.08710429
	15	0.94428877	0.10677300	0.05338650
	20	0.97282673	0.07305521	0.03652760
5	25	0.98547639	0.05282453	0.02641227
	30	0.99183642	0.03933742	0.01966871
	35	0.99530300	0.02970377	0.01485188
	40	0.99729273	0.02247853	0.01123926
	45	0.99846982	0.01685889	0.00842945
	50	0.99917404	0.01236319	0.00618159
	5	0.39966232	0.17143446	0.08571723
	10	0.76835727	0.08234253	0.04117127
	15	0.88740307	0.05264523	0.02632261
	20	0.93691091	0.03779657	0.01889829
10	25	0.96132332	0.02888738	0.01444369
	30	0.97481725	0.02294792	0.01147396
	35	0.98289880	0.01870545	0.00935272
	40	0.98803141	0.01552359	0.00776180
	45	0.99143882	0.01304882	0.00652441
	50	0.99377962	0.01106900	0.00553450
	5	0.32640359	0.11141464	0.05570732
	10	0.70763715	0.05421782	0.02710891
	15	0.84605410	0.03515221	0.01757611
	20	0.90827828	0.02561941	0.01280970
15	25	0.94071221	0.01989973	0.00994986
	30	0.95944275	0.01608661	0.00804330
	35	0.97108945	0.01336295	0.00668147
	40	0.97874382	0.01132020	0.00566010
	45	0.98399442	0.00973140	0.00486570
	50	0.98772043	0.00846036	0.00423018

Table 6.4 (continued): MRR (Minimum Relative Risk) of proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 3$ and $\Lambda = diag(1,1,1)$.

n	ν	MRR	d^*	c_2(opt)
20	5	0.28305635	0.08258202	0.04129101
	10	0.66546503	0.04045516	0.02022758
	15	0.81450303	0.02641287	0.01320644
	20	0.88507042	0.01939173	0.00969586
	25	0.92327623	0.01517904	0.00758952
	30	0.94600136	0.01237059	0.00618529
	35	0.96048053	0.01036454	0.00518227
	40	0.97019956	0.00886001	0.00443001
	45	0.97699467	0.00768982	0.00384491
	50	0.98190308	0.00675367	0.00337684
25	5	0.25430098	0.06561938	0.03280969
	10	0.63440331	0.03227533	0.01613766
	15	0.78962374	0.02116065	0.01058032
	20	0.86590896	0.01560330	0.00780165
	25	0.90839653	0.01226890	0.00613445
	30	0.93423772	0.01004596	0.00502298
	35	0.95100594	0.00845815	0.00422908
	40	0.96243820	0.00726729	0.00363365
	45	0.97054153	0.00634107	0.00317053
	50	0.97646804	0.00560009	0.00280005
30	5	0.23380615	0.05444283	0.02722142
	10	0.61055652	0.02685055	0.01342528
	15	0.76949947	0.01765313	0.00882656
	20	0.84982915	0.01305441	0.00652721
	25	0.89556738	0.01029518	0.00514759
	30	0.92388180	0.00845570	0.00422785
	35	0.94252492	0.00714178	0.00357089
	40	0.95539404	0.00615634	0.00307817
	45	0.96461510	0.00538989	0.00269494
	50	0.97142461	0.00477673	0.00238836

Table 6.5: MRR (Minimum Relative Risk) of proposed esti- ..or $\hat{\Lambda}_1$ relative to $\tilde{\Lambda}_1$ for $p = 3$ and $\Lambda = diag(4, 2, 1)$.

n	ν	MRR	d^*	c_2(opt)
5	5	0.62141238	0.37651530	0.21963392
	10	0.87995912	0.17420857	0.10162167
	15	0.94963936	0.10677300	0.06228425
	20	0.97526844	0.07305521	0.04261554
	25	0.98673117	0.05282453	0.03081431
	30	0.99252396	0.03933742	0.02294683
	35	0.99569164	0.02970377	0.01732720
	40	0.99751385	0.02247853	0.01311247
	45	0.99859358	0.01685889	0.00983435
	50	0.99924033	0.01236319	0.00721186
10	5	0.50309706	0.17143446	0.10000343
	10	0.79961726	0.08234253	0.04803315
	15	0.90048577	0.05264523	0.03070972
	20	0.94357179	0.03779657	0.02204800
	25	0.96514803	0.02888738	0.01685097
	30	0.97719265	0.02294792	0.01338629
	35	0.98445552	0.01870545	0.01091151
	40	0.98909115	0.01552359	0.00905543
	45	0.99218025	0.01304882	0.00761181
	50	0.99430866	0.01106900	0.00645691
15	5	0.45111036	0.11141464	0.06499187
	10	0.75191197	0.05421782	0.03162706
	15	0.86629316	0.03515221	0.02050546
	20	0.91918941	0.02561941	0.01494466
	25	0.94726538	0.01989973	0.01160817
	30	0.96368146	0.01608661	0.00938385
	35	0.97398089	0.01336295	0.00779505
	40	0.98079570	0.01132020	0.00660345
	45	0.98549508	0.00973140	0.00567665
	50	0.98884403	0.00846036	0.00493521

Table 6.5 (continued): MRR (Minimum Relative Risk) of proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 3$ and $\Lambda = diag(4, 2, 1)$.

n	ν	MRR	d^*	c_2(opt)
20	5	0.42123025	0.08258202	0.04817284
	10	0.71990755	0.04045516	0.02359884
	15	0.84103407	0.02641287	0.01540751
	20	0.89998640	0.01939173	0.01131184
	25	0.93251559	0.01517904	0.00885444
	30	0.95213037	0.01237059	0.00721618
	35	0.96475635	0.01036454	0.00604598
	40	0.97329890	0.00886001	0.00516834
	45	0.97930929	0.00768982	0.00448573
	50	0.98367322	0.00675367	0.00393964
25	5	0.40175442	0.06561938	0.03827797
	10	0.69690110	0.03227533	0.01882728
	15	0.82160385	0.02116065	0.01234371
	20	0.88449653	0.01560330	0.00910193
	25	0.92019441	0.01226890	0.00715686
	30	0.94221626	0.01004596	0.00586015
	35	0.95666343	0.00845815	0.00493392
	40	0.96659905	0.00726729	0.00423925
	45	0.97369123	0.00634107	0.00369896
	50	0.97890843	0.00560009	0.00326672
30	5	0.38803721	0.05444283	0.03175832
	10	0.67955431	0.02685055	0.01566282
	15	0.80619177	0.01765313	0.01029766
	20	0.87174502	0.01305441	0.00761507
	25	0.90976265	0.01029518	0.00600552
	30	0.93363568	0.00845570	0.00493249
	35	0.94953293	0.00714178	0.00416604
	40	0.96060721	0.00615634	0.00359120
	45	0.96860204	0.00538989	0.00314410
	50	0.97454322	0.00477673	0.00278642

Table 6.6: MRR (Minimum Relative Risk) of proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 3$ and $\Lambda = diag(25, 1, 1)$.

n	ν	MRR	d^*	c_2(opt)
	5	0.77210436	0.37651530	0.57944889
	10	0.91970221	0.17420857	0.26810322
	15	0.96490918	0.10677300	0.16432133
	20	0.98242411	0.07305521	0.11243038
5	25	0.99046399	0.05282453	0.08129582
	30	0.99458882	0.03933742	0.06053944
	35	0.99686640	0.02970377	0.04571345
	40	0.99818538	0.02247853	0.03459396
	45	0.99897074	0.01685889	0.02594547
	50	0.99944290	0.01236319	0.01902668
	5	0.72094168	0.17143446	0.26383392
	10	0.87564960	0.08234253	0.12672338
	15	0.93480648	0.05264523	0.08101986
	20	0.96183063	0.03779657	0.05816811
10	25	0.97593184	0.02888738	0.04445705
	30	0.98402152	0.02294792	0.03531635
	35	0.98899462	0.01870545	0.02878728
	40	0.99221451	0.01552359	0.02389047
	45	0.99438393	0.01304882	0.02008185
	50	0.99589177	0.01106900	0.01703495
	5	0.70149978	0.11141464	0.17146472
	10	0.85286671	0.05421782	0.08344005
	15	0.91620630	0.03515221	0.05409850
	20	0.94749107	0.02561941	0.03942772
15	25	0.96486241	0.01989973	0.03062525
	30	0.97535268	0.01608661	0.02475694
	35	0.98209464	0.01336295	0.02056529
	40	0.98663926	0.01132020	0.01742155
	45	0.98981974	0.00973140	0.01497642
	50	0.99211359	0.00846036	0.01302032

Table 6.6 (continued): MRR (Minimum Relative Risk) of proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 3$ and $\Lambda = diag(25, 1, 1)$.

n	ν	MRR	d^*	c_2(opt)
20	5	0.69102895	0.08258202	0.12709194
	10	0.83878244	0.04045516	0.06225961
	15	0.90354215	0.02641287	0.04064884
	20	0.93702043	0.01939173	0.02984345
	25	0.95633838	0.01517904	0.02336022
	30	0.96838758	0.01237059	0.01903806
	35	0.97635000	0.01036454	0.01595081
	40	0.98185110	0.00886001	0.01363537
	45	0.98578798	0.00768982	0.01183447
	50	0.98868697	0.00675367	0.01039375
25	5	0.68445740	0.06561938	0.10098680
	10	0.82919573	0.03227533	0.0496[illegible]
	15	0.89436227	0.02116065	0.03256578
	20	0.92905470	0.01560330	0.02401315
	25	0.94960314	0.01226890	0.01888157
	30	0.96271301	0.01004596	0.01546052
	35	0.97154885	0.00845815	0.01301691
	40	0.97776087	0.00726729	0.01118421
	45	0.98227699	0.00634107	0.00975877
	50	0.98565077	0.00560009	0.00861842
30	5	0.67994259	0.05444283	0.08378634
	10	0.82224463	0.02685055	0.04132242
	15	0.88740242	0.01765313	0.02716778
	20	0.92279528	0.01305441	0.02009046
	25	0.94415553	0.01029518	0.01584407
	30	0.95801296	0.00845570	0.01301314
	35	0.96749208	0.00714178	0.01099105
	40	0.97424515	0.00615634	0.00947448
	45	0.97921364	0.00538989	0.00829492
	50	0.98296615	0.00477673	0.00735128

Table 6.7: MRR (Minimum Relative Risk) of proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 5$ and $\Lambda = diag(1,1,1,1,1)$.

n	ν	MRR	d^*	c_2(opt)
10	5	0.51278498	0.19909764	0.09954882
	10	0.84135187	0.09226476	0.04613238
	15	0.93186338	0.05665380	0.02832690
	20	0.96595775	0.03884832	0.01942416
	25	0.98147861	0.02816503	0.01408252
	30	0.98943067	0.02104284	0.01052142
	35	0.99382985	0.01595556	0.00797778
	40	0.99638784	0.01214010	0.00607005
	45	0.99792032	0.00917252	0.00458626
	50	0.99884987	0.00679846	0.00339923
15	5	0.42310014	0.12167796	0.06083898
	10	0.78275670	0.05784690	0.02892345
	15	0.89647768	0.03656988	0.01828494
	20	0.94326172	0.02593137	0.01296568
	25	0.96608515	0.01954826	0.00977413
	30	0.97854681	0.01529286	0.00764643
	35	0.98590407	0.01225328	0.00612664
	40	0.99049953	0.00997360	0.00498680
	45	0.99349191	0.00820052	0.00410026
	50	0.99550187	0.00678205	0.00339103
20	5	0.36668366	0.08790116	0.04395058
	10	0.73888957	0.04232278	0.02116139
	15	0.86[illegible]719	0.02712999	0.01356499
	20	0.92349067	0.01953359	0.00976680
	25	0.95204069	0.01497575	0.00748788
	30	0.96819936	0.01193719	0.00596860
	35	0.97805686	0.00976680	0.00488340
	40	0.98441331	0.00813900	0.00406950
	45	0.98868920	0.00687293	0.00343647
	50	0.99166190	0.00586008	0.00293004

Table 6.8: MRR (Minimum Relative Risk) of proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 5$ and $\Lambda = diag(5, 4, 3, 2, 1)$.

n	ν	MRR	d^*	c_2(opt)
10	5	0.54403941	0.19909764	0.11463603
	10	0.84841159	0.09226476	0.05312401
	15	0.93433215	0.05665380	0.03262001
	20	0.96704513	0.03884832	0.02236801
	25	0.98202251	0.02816503	0.01621680
	30	0.98972305	0.02104284	0.01211600
	35	0.99399310	0.01595556	0.00918686
	40	0.99648018	0.01214010	0.00699000
	45	0.99797205	0.00917252	0.00528133
	50	0.99887785	0.00679846	0.00391440
15	5	0.46471353	0.12167796	0.07005948
	10	0.79418188	0.05784690	0.03330697
	15	0.90091821	0.03656988	0.02105613
	20	0.94538349	0.02593137	0.01493071
	25	0.96723528	0.01954826	0.01125546
	30	0.97922311	0.01529286	0.00880529
	35	0.98632402	0.01225328	0.00705517
	40	0.99077011	0.00997360	0.00574258
	45	0.99367059	0.00820052	0.00472168
	50	0.99562165	0.00678205	0.00390495
20	5	0.41558238	0.08790116	0.05061155
	10	0.75420498	0.04232278	0.02436852
	15	0.87384779	0.02712999	0.01562085
	20	0.92671398	0.01953359	0.01124701
	25	0.95386404	0.01497575	0.00862271
	30	0.96931623	0.01193719	0.00687317
	35	0.97878045	0.00976680	0.00562351
	40	0.98490159	0.00813900	0.00468625
	45	0.98902875	0.00687293	0.00395728
	50	0.99190339	0.00586008	0.00337410

Table 6.9: MRR (Minimum Relative Risk) of proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 5$ and $\Lambda = diag(6, 5, 3, 1, 1)$.

n	ν	MRR	d^*	c_2(opt)
10	5	0.61042200	0.19909764	0.12952023
	10	0.86443548	0.09226476	0.06002157
	15	0.94009734	0.05665380	0.03685535
	20	0.96962253	0.03884832	0.02527224
	25	0.98332335	0.02816503	0.01832237
	30	0.99042651	0.02104284	0.01368913
	35	0.99438755	0.01595556	0.01037967
	40	0.99670400	0.01214010	0.00789757
	45	0.99809774	0.00917252	0.00596706
	50	0.99894596	0.00679846	0.00442264
15	5	0.55078443	0.12167796	0.07915592
	10	0.81940973	0.05784690	0.03763150
	15	0.91105273	0.03656988	0.02379003
	20	0.95031846	0.02593137	0.01686929
	25	0.96994280	0.01954826	0.01271685
	30	0.98082848	0.01529286	0.00994856
	35	0.98732691	0.01225328	0.00797120
	40	0.99141925	0.00997360	0.00648819
	45	0.99410079	0.00820052	0.00533473
	50	0.99591090	0.00678205	0.00441197
20	5	0.51504258	0.08790116	0.05718289
	10	0.78733814	0.04232278	0.02753250
	15	0.88820788	0.02712999	0.01764904
	20	0.93408978	0.01953359	0.01270731
	25	0.95809581	0.01497575	0.00974227
	30	0.97193448	0.01193719	0.00776558
	35	0.98048943	0.00976680	0.00635365
	40	0.98606143	0.00813900	0.00529471
	45	0.98983900	0.00687293	0.00447109
	50	0.99248175	0.00586008	0.00381219

Table 6.10: MRR (Minimum Relative Risk) of proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 7$ and $\Lambda = diag(1,1,1,1,1,1,1)$.

n	ν	MRR	d^*	c_2(opt)
10	5	0.58699354	0.23987493	0.11993746
	10	0.87823455	0.10934228	0.05467114
	15	0.95100734	0.06583140	0.03291570
	20	0.97696590	0.04407596	0.02203798
	25	0.98827876	0.03102269	0.01551135
	30	0.99382795	0.02232052	0.01116026
	35	0.99674892	0.01610468	0.00805234
	40	0.99834458	0.01144280	0.00572140
	45	0.99922243	0.00781689	0.00390844
	50	0.99969087	0.00491616	0.00245808
15	5	0.49126240	0.13410122	0.06705061
	10	0.82491058	0.06308781	0.03154390
	15	0.92103821	0.03941667	0.01970834
	20	0.95857622	0.02758110	0.01379055
	25	0.97623150	0.02047976	0.01023988
	30	0.98558261	0.01574553	0.00787277
	35	0.99094854	0.01236394	0.00618197
	40	0.99420463	0.00982775	0.00491387
	45	0.99625974	0.00785515	0.00392758
	50	0.99759231	0.00627708	0.00313854
20	5	0.42869051	0.09389402	0.04694701
	10	0.78321700	0.04486233	0.02243116
	15	0.89544658	0.02851843	0.01425922
	20	0.94192327	0.02034648	0.01017324
	25	0.96477362	0.01544331	0.00772166
	30	0.97735541	0.01217454	0.00608727
	35	0.98485175	0.00983969	0.00491985
	40	0.98958140	0.00808856	0.00404428
	45	0.99269598	0.00672657	0.00336328
	50	0.99481492	0.00563698	0.00281849

Table 6.11: MRR (Minimum Relative Risk) of proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 7$ and $\Lambda = diag(4,3,3,2,2,1,1)$.

n	ν	MRR	d^*	c_2(opt)
10	5	0.62747266	0.23987493	0.13478430
	10	0.88610333	0.10934228	0.06143878
	15	0.95353261	0.06583140	0.03699027
	20	0.97800161	0.04407596	0.02476602
	25	0.98876026	0.03102269	0.01743147
	30	0.99406581	0.02232052	0.01254177
	35	0.99686841	0.01610468	0.00904912
	40	0.99840324	0.01144280	0.00642964
	45	0.99924919	0.00781689	0.00439226
	50	0.99970126	0.00491616	0.00276236
15	5	0.54854198	0.13410122	0.07535068
	10	0.83856126	0.06308781	0.03544867
	15	0.92593709	0.03941667	0.02214799
	20	0.96078624	0.02758110	0.01549766
	25	0.97737163	0.02047976	0.01150746
	30	0.98622175	0.01574553	0.00884732
	35	0.99132615	0.01236394	0.00694723
	40	0.99443503	0.00982775	0.00552215
	45	0.99640272	0.00785515	0.00441376
	50	0.99768141	0.00627708	0.00352705
20	5	0.49850336	0.09389402	0.05275849
	10	0.80236544	0.04486233	0.02520788
	15	0.90286122	0.02851843	0.01602434
	20	0.94545539	0.02034648	0.01143257
	25	0.96668726	0.01544331	0.00867751
	30	0.97848370	0.01217454	0.00684080
	35	0.98555670	0.00983969	0.00552887
	40	0.99004014	0.00808856	0.00454491
	45	0.99300320	0.00672657	0.00377962
	50	0.99502478	0.00563698	0.00316738

Table 6.12: MRR (Minimum Relative Risk) of proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 7$ and $\Lambda = diag(7,2,1,1,1,1,1)$.

n	ν	MRR	d^*	c_2(opt)
10	5	0.73746629	0.23987493	0.16453272
	10	0.91075629	0.10934228	0.07499901
	15	0.96192244	0.06583140	0.04515444
	20	0.98154819	0.04407596	0.03023216
	25	0.99043918	0.03102269	0.02127879
	30	0.99490513	0.02232052	0.01530988
	35	0.99729361	0.01610468	0.01104637
	40	0.99861329	0.01144280	0.00784873
	45	0.99934550	0.00781689	0.00536169
	50	0.99973880	0.00491616	0.00337205
15	5	0.69522813	0.13410122	0.09198143
	10	0.87891562	0.06308781	0.04327259
	15	0.94146190	0.03941667	0.02703631
	20	0.96806862	0.02758110	0.01891817
	25	0.98122206	0.02047976	0.01404728
	30	0.98841681	0.01574553	0.01080003
	35	0.99263890	0.01236394	0.00848056
	40	0.99524338	0.00982775	0.00674096
	45	0.99690799	0.00785515	0.00538793
	50	0.99799809	0.00627708	0.00430551
20	5	0.67058005	0.09389402	0.06440289
	10	0.85653562	0.04486233	0.03077154
	15	0.92545858	0.02851843	0.01956109
	20	0.95670996	0.02034648	0.01395587
	25	0.97296527	0.01544331	0.01059273
	30	0.98226180	0.01217454	0.00835064
	35	0.98795320	0.00983969	0.00674915
	40	0.99161785	0.00808856	0.00554803
	45	0.99406951	0.00672657	0.00461382
	50	0.99575859	0.00563698	0.00386646

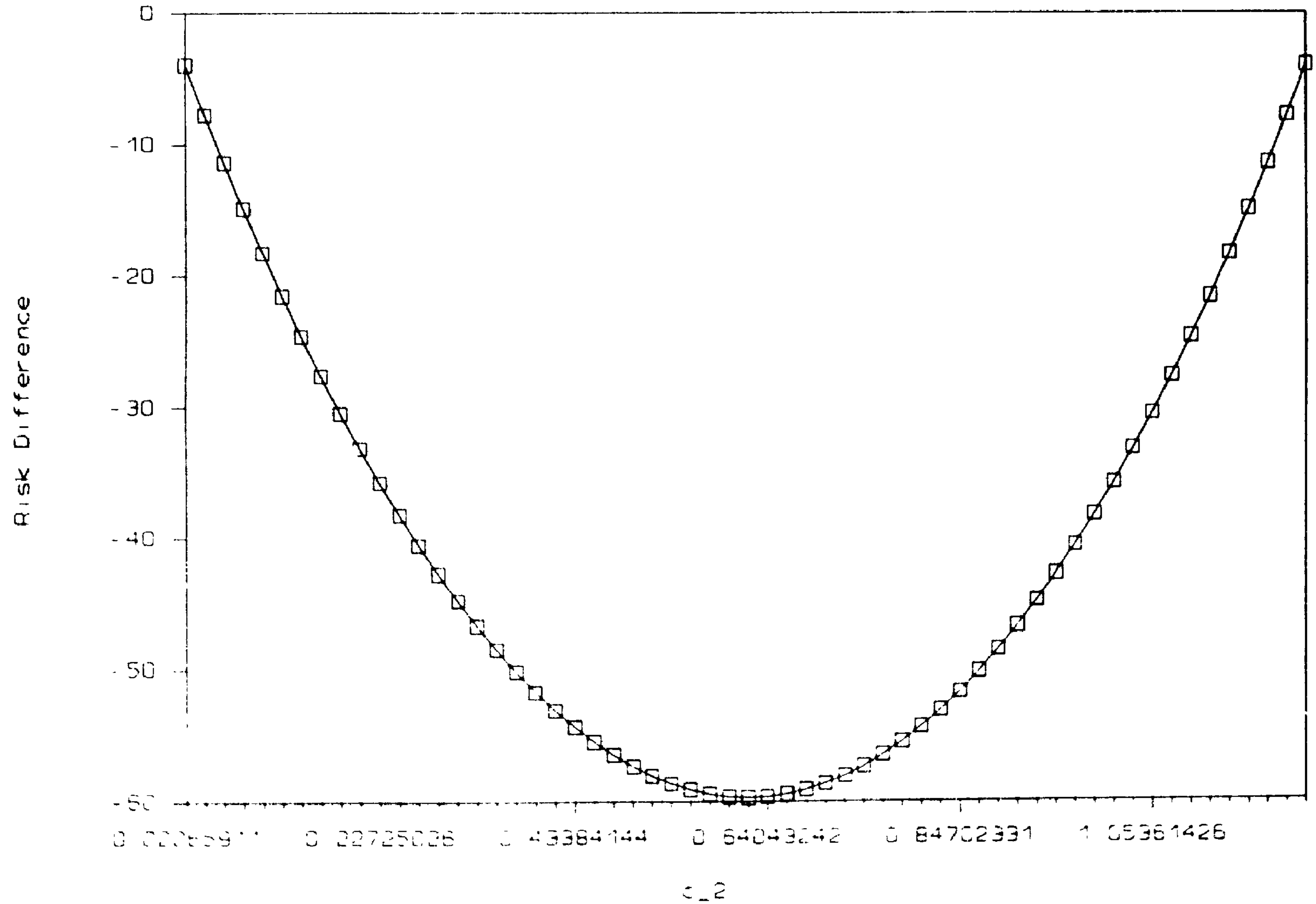
Table 6.13: MRR (Minimum Relative Risk) of proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 3$ and $\Lambda = ((\lambda_{ik}))$ where $\lambda_{11} = 94, \lambda_{12} = 41, \lambda_{13} = 23, \lambda_{22} = 26, \lambda_{23} = 11$ and $\lambda_{33} = 6$.

n	ν	MRR	d^*	c_2(opt)
5	5	0.77922797	0.37651515	1.36118281
	10	0.92180097	0.17420851	0.62980103
	15	0.96575099	0.10677297	0.38600713
	20	0.98282689	0.07305519	0.26411012
	25	0.99067664	0.05282452	0.19097194
	30	0.99470735	0.03933741	0.14221315
	35	0.99693418	0.02970375	0.10738543
	40	0.99822426	0.02247852	0.08126464
	45	0.99899262	0.01685889	0.06094850
	50	0.99945468	0.01236318	0.04469556
10	5	0.73051786	0.17143475	0.61977327
	10	0.87935990	0.08234267	0.29768637
	15	0.93657935	0.05264532	0.19032410
	20	0.96280611	0.03779664	0.13664293
	25	0.97652066	0.02888743	0.10443424
	30	0.98440009	0.02294796	0.08296178
	35	0.98924905	0.01870548	0.06762431
	40	0.99239111	0.01552362	0.05612121
	45	0.99450934	0.01304884	0.04717435
	50	0.99598235	0.01106901	0.04001686
15	5	0.71214372	0.11141477	0.40278825
	10	0.85756493	0.05421788	0.19600925
	15	0.91866732	0.03515226	0.12708294
	20	0.94894052	0.02561944	0.09261977
	25	0.96578777	0.01989975	0.07194186
	30	0.97597837	0.01608662	0.05815659
	35	0.98253608	0.01336296	0.04830997
	40	0.98696083	0.01132022	0.04092501
	45	0.99005997	0.00973141	0.03518115
	50	0.99229658	0.00846037	0.03058606

Table 6.13 (continued): MRR (Minimum Relative Risk) of proposed estimator $\hat{\Lambda}_1$ relative to m.l.e. $\tilde{\Lambda}_1$ for $p = 3$ and $\Lambda = ((\lambda_{ik}))$ where $\lambda_{11} = 94, \lambda_{12} = 41, \lambda_{13} = 23, \lambda_{22} = 26, \lambda_{23} = 11$ and $\lambda_{33} = 6$.

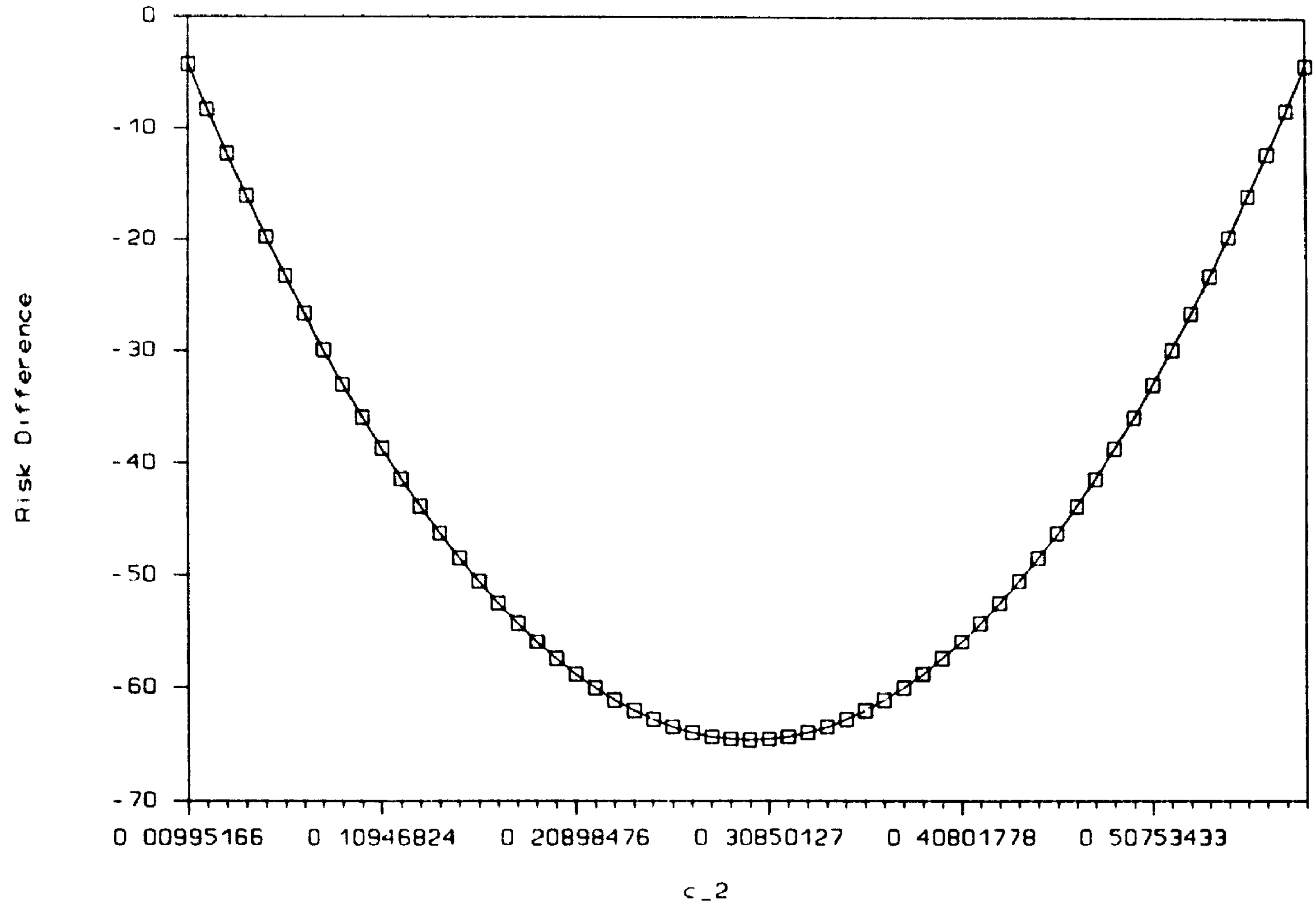
n	ν	MRR	d^*	c_2(opt)
20	5	0.70228201	0.08258148	0.29854974
	10	0.84414595	0.04045489	0.14625311
	15	0.90652263	0.02641270	0.09548757
	20	0.93885636	0.01939160	0.07010479
	25	0.95755345	0.01517894	0.05487513
	30	0.96923482	0.01237051	0.04472202
	35	0.97696453	0.01036448	0.03746980
	40	0.98231059	0.00885996	0.03203064
	45	0.98613989	0.00768977	0.02780018
	50	0.98896188	0.00675363	0.02441581
25	5	0.69610256	0.06561891	0.23722641
	10	0.83503467	0.03227510	0.11668140
	15	0.89774430	0.02116050	0.07649972
	20	0.93120688	0.01560319	0.05640888
	25	0.95106554	0.01226881	0.04435438
	30	0.96375567	0.01004589	0.03631805
	35	0.97231996	0.00845809	0.03057781
	40	0.97834772	0.00726724	0.02627263
	45	0.98273396	0.00634102	0.02292416
	50	0.98601341	0.00560005	0.02024538
30	5	0.69185865	0.05444259	0.19682162
	10	0.82843810	0.02685043	0.09707006
	15	0.89110184	0.01765305	0.06381955
	20	0.92520875	0.01305435	0.04719428
	25	0.94582945	0.01029514	0.03721912
	30	0.95922720	0.00845566	0.03056902
	35	0.96840376	0.00714175	0.02581895
	40	0.97494835	0.00615631	0.02225639
	45	0.97976804	0.00538987	0.01948552
	50	0.98341119	0.00477671	0.01726882

Figure 1 RISK DIFFERENCE (n=10, $\hat{y}$=5)



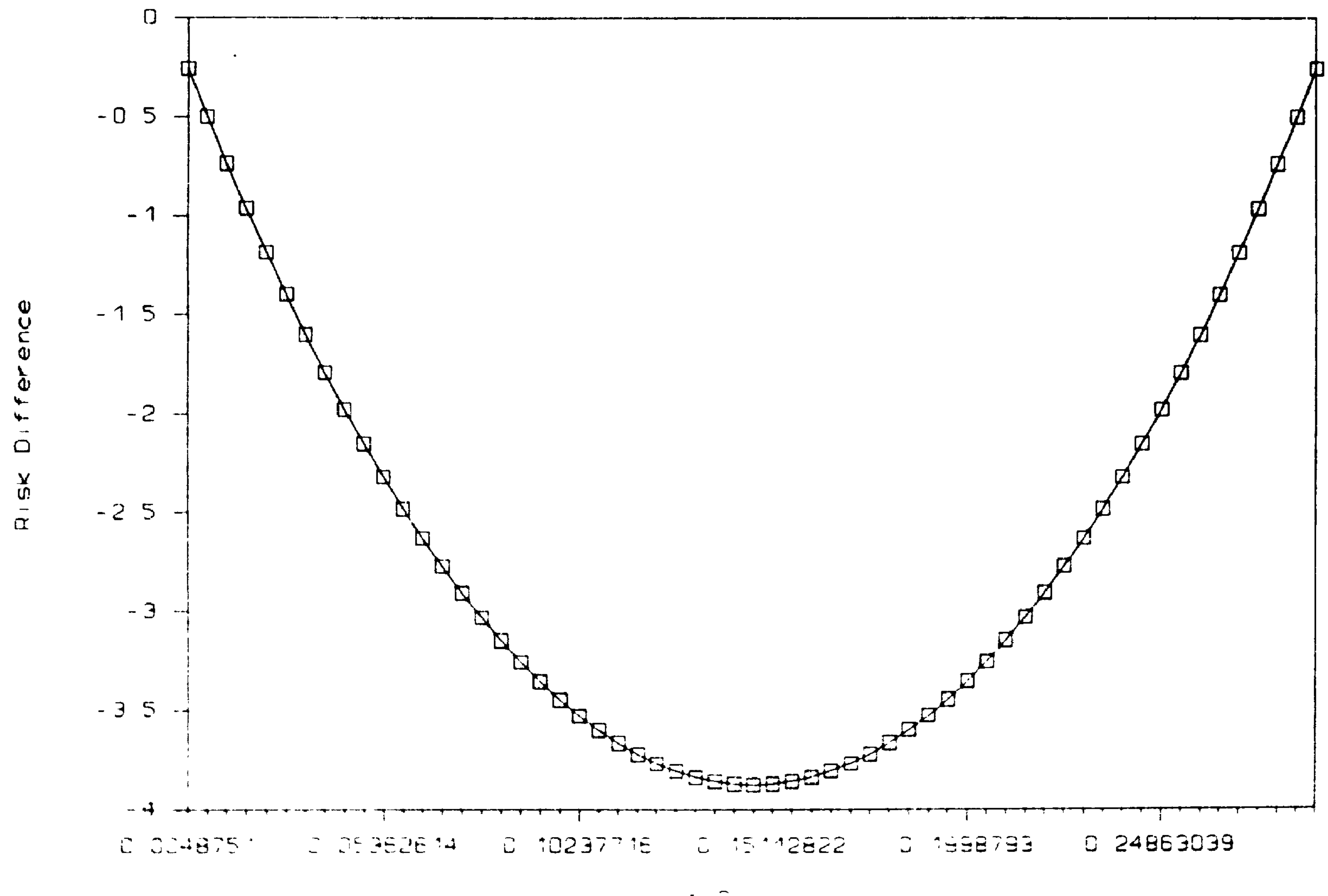
$d^* = 0.17143446$

Figure 2 RISK DIFFERENCE (n=20, ν=5)



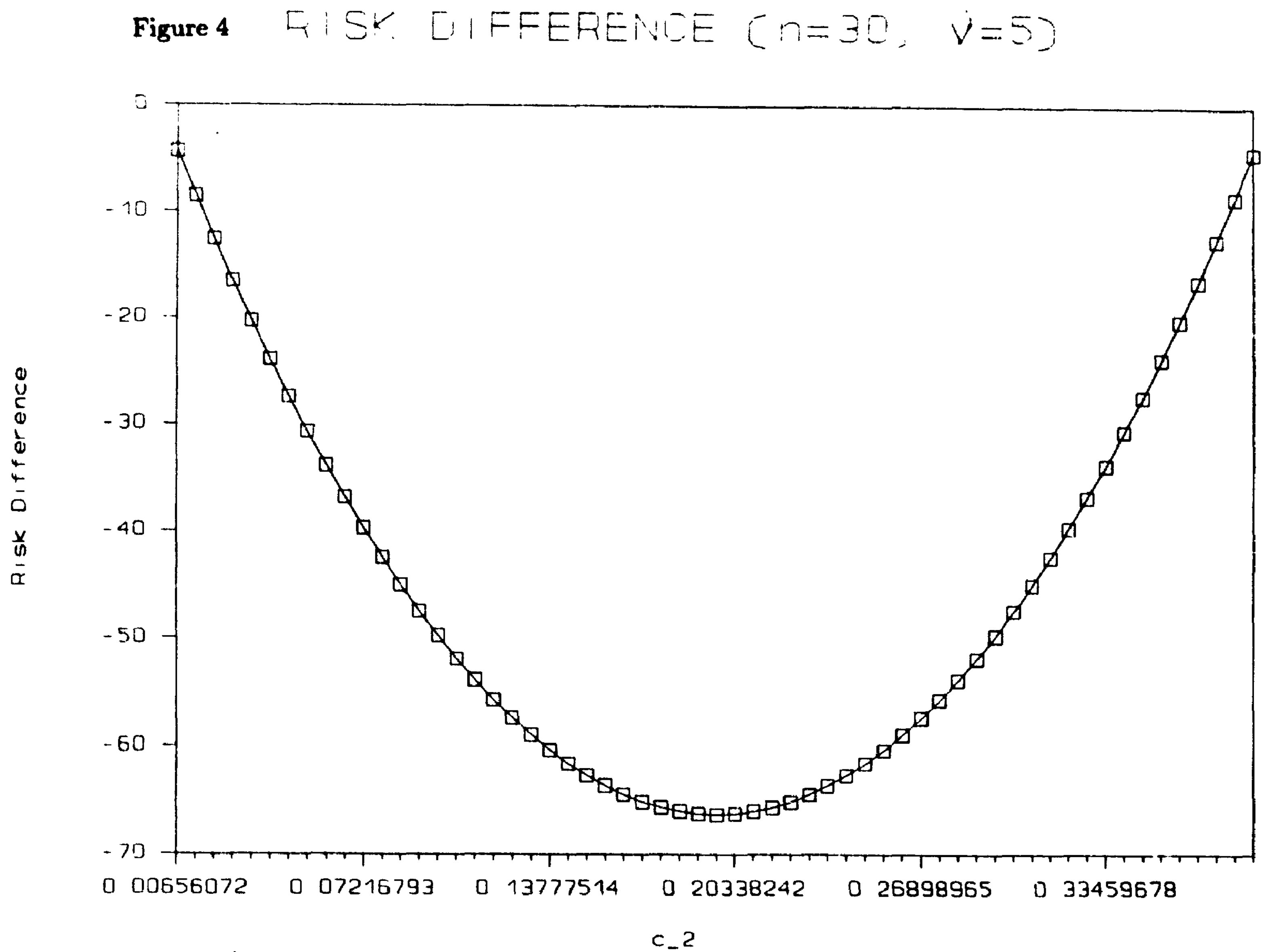
$d^* = 0.08258202$

Figure 3 RISK DIFFERENCE (n=20, ẏ=10)



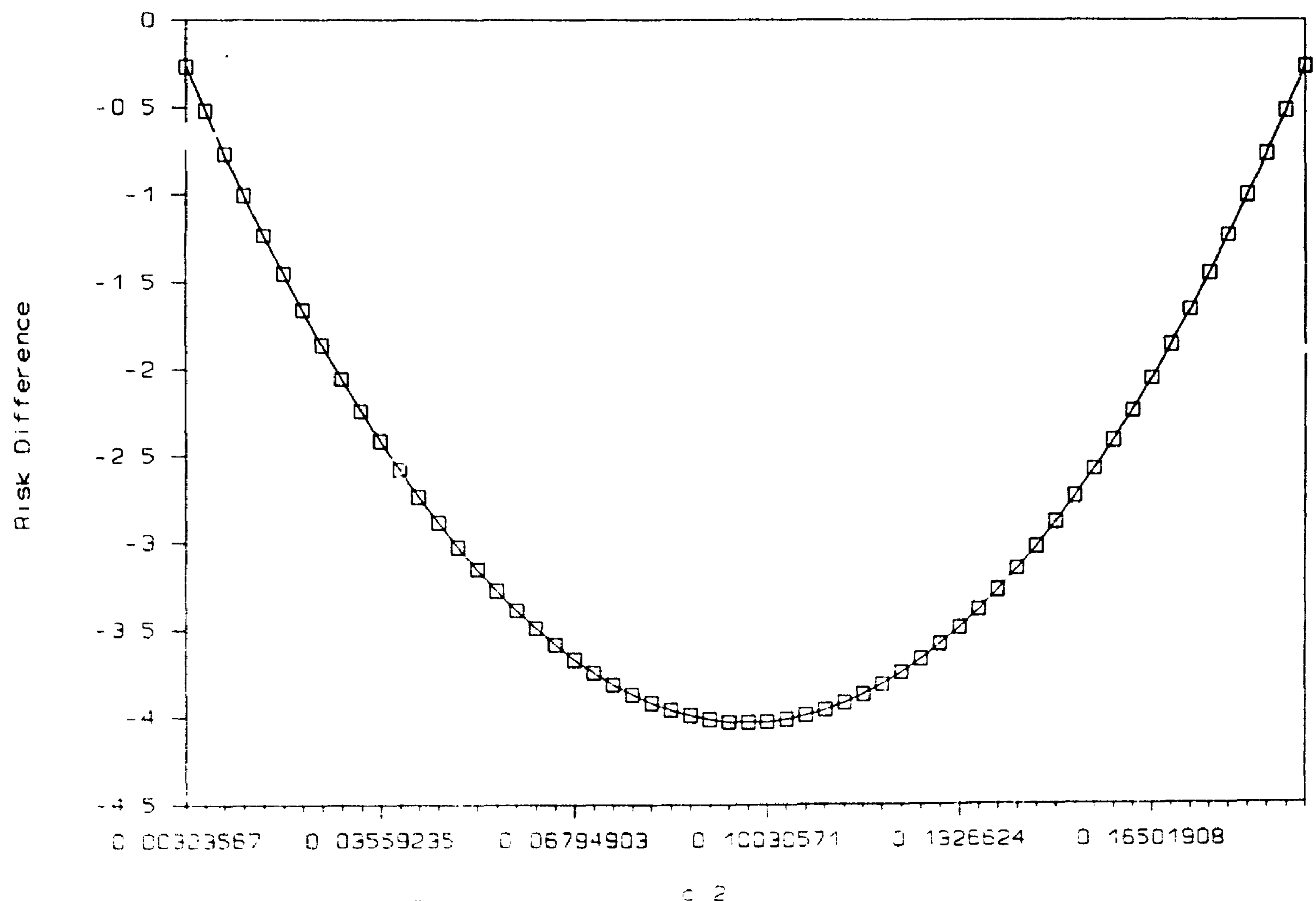
$d^* = 0.04045516$

Figure 4



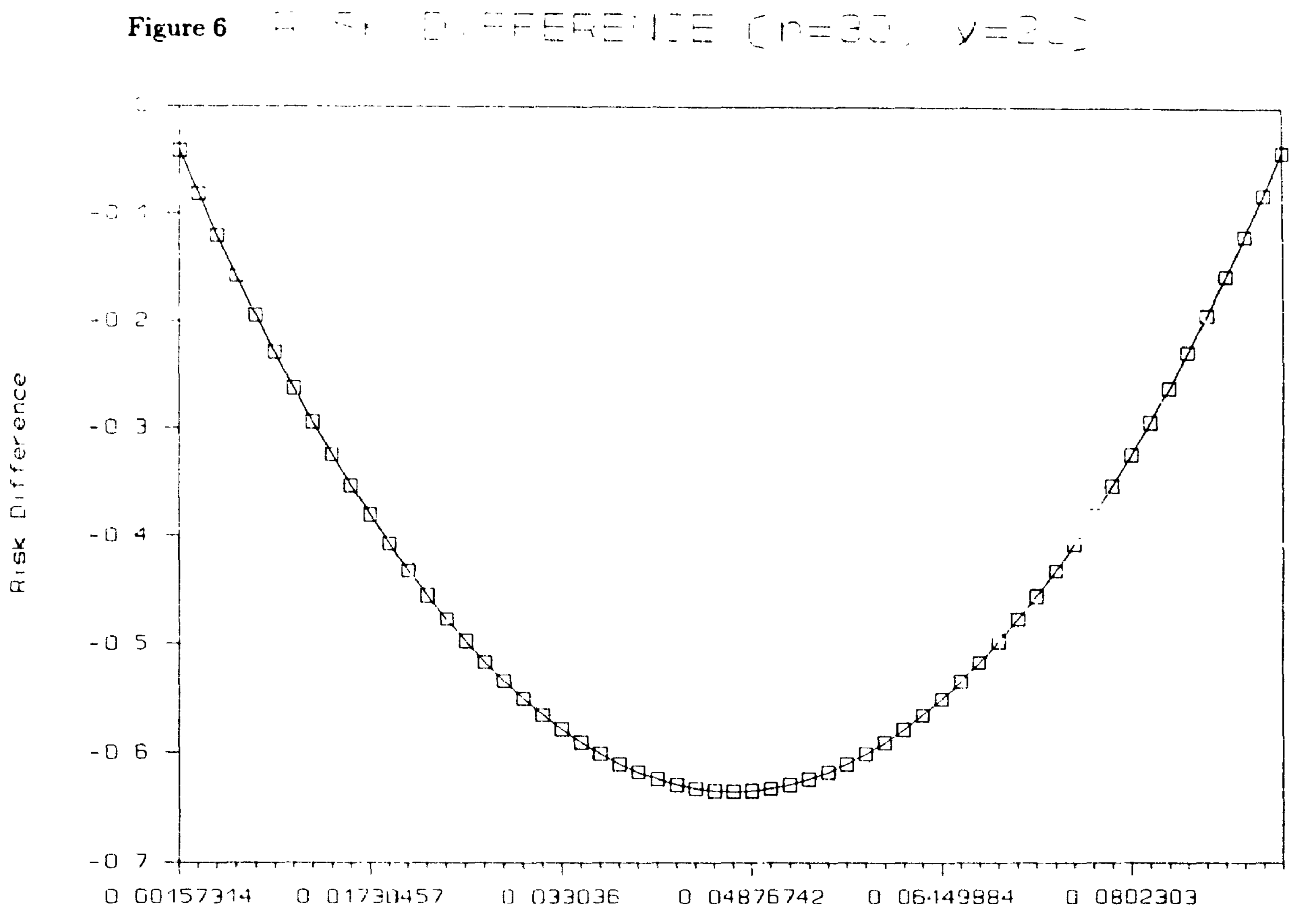
$d^* = 0.05444283$

Figure 5 RISK DIFFERENCE (n=30, ν=10)



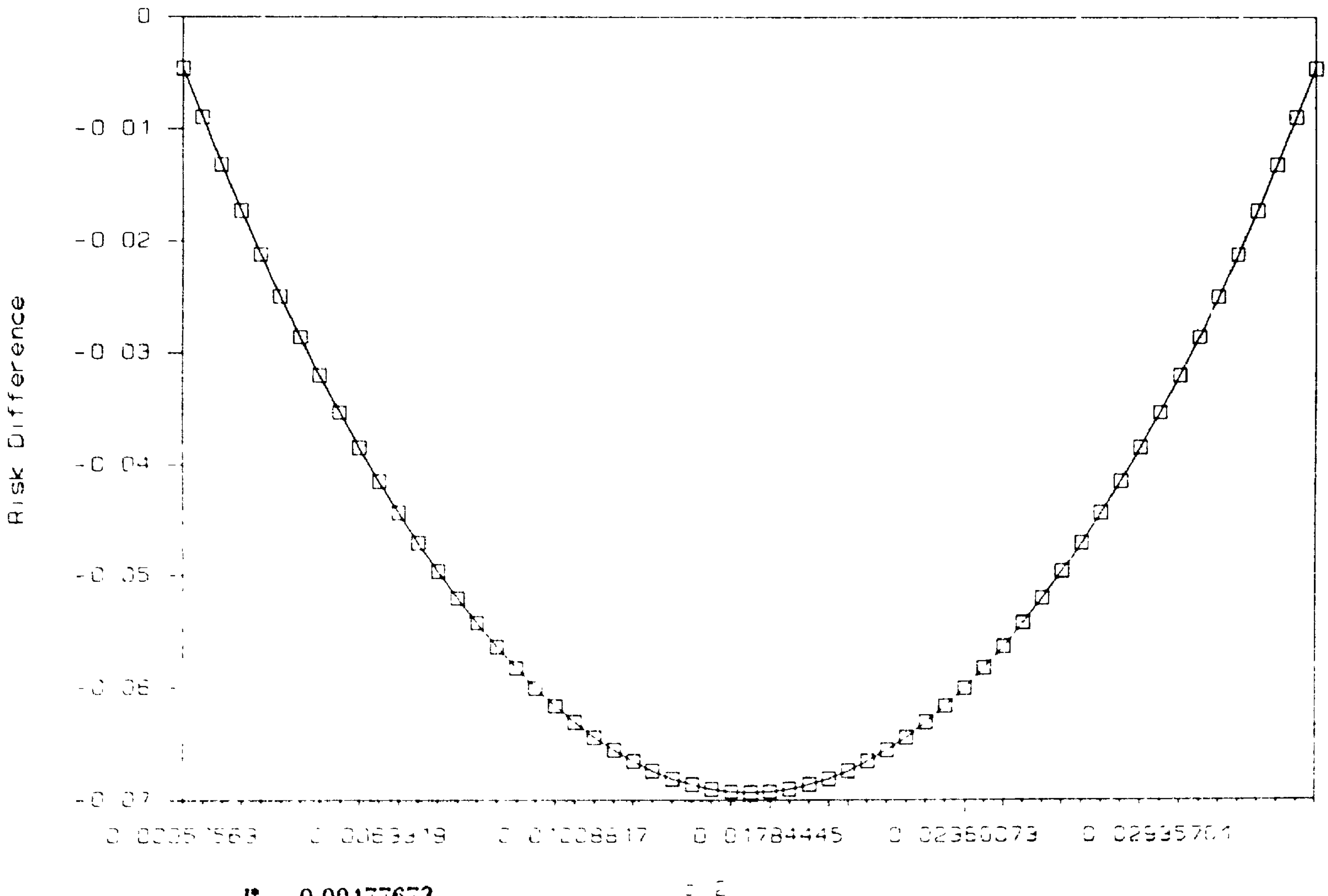
$d^* = 0.02685055$

Figure 6



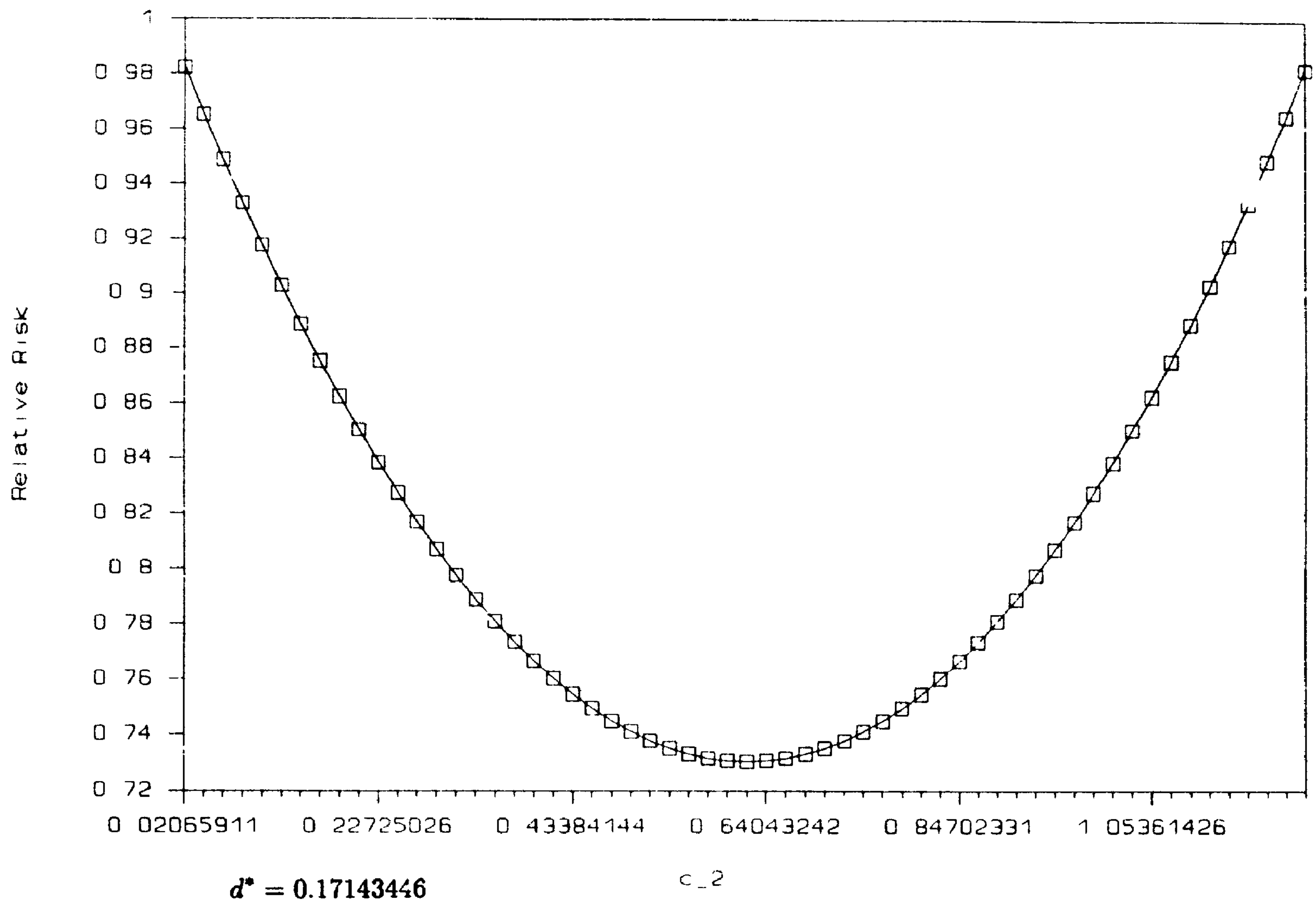
$d^* = 0.01305441$

Figure 7 RISK DIFFERENCE (n=30, ν=50)



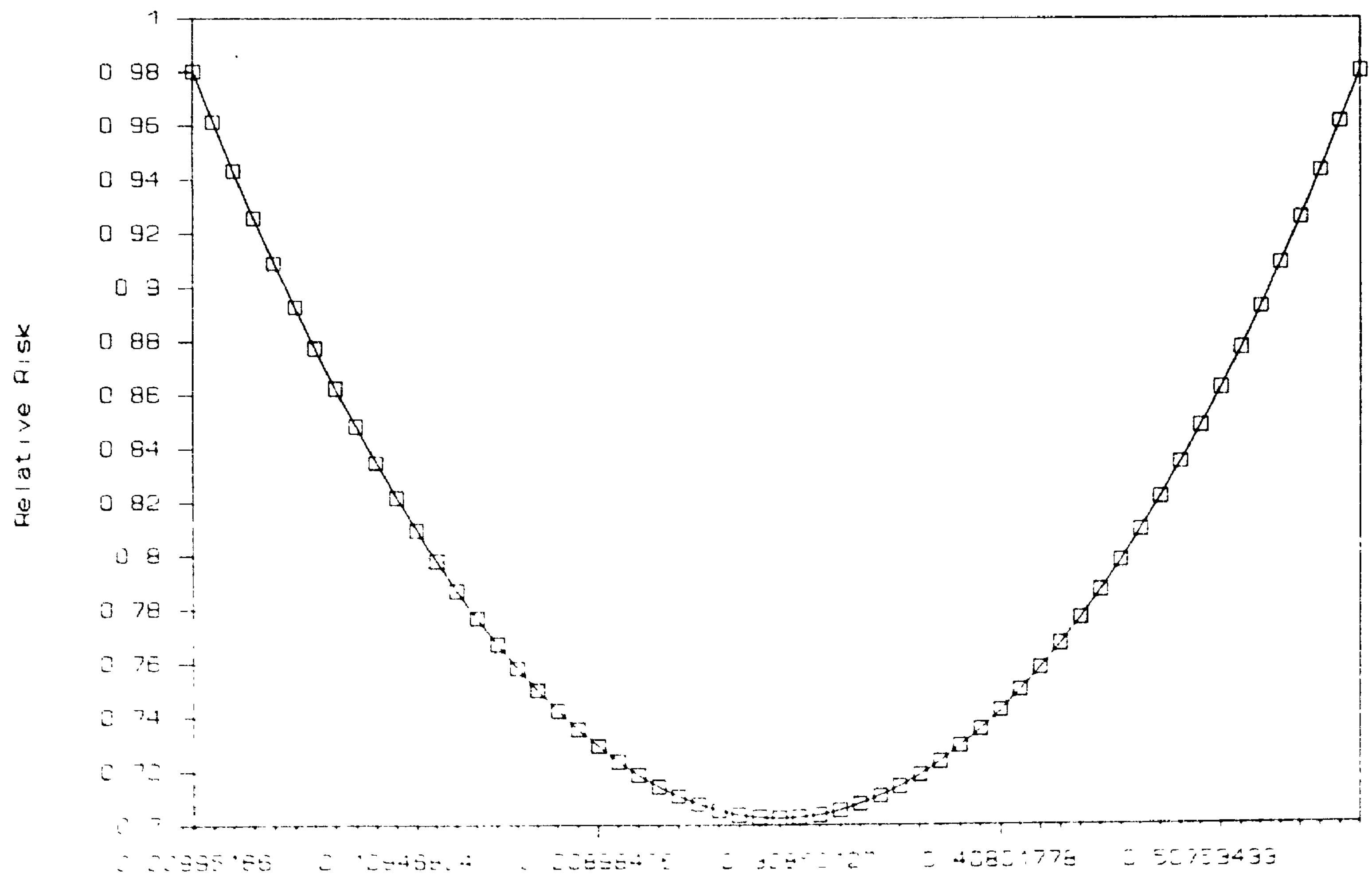
$d^* = 0.00477673$

Figure 8 RELATIVE RISK (n=10, ν=5)



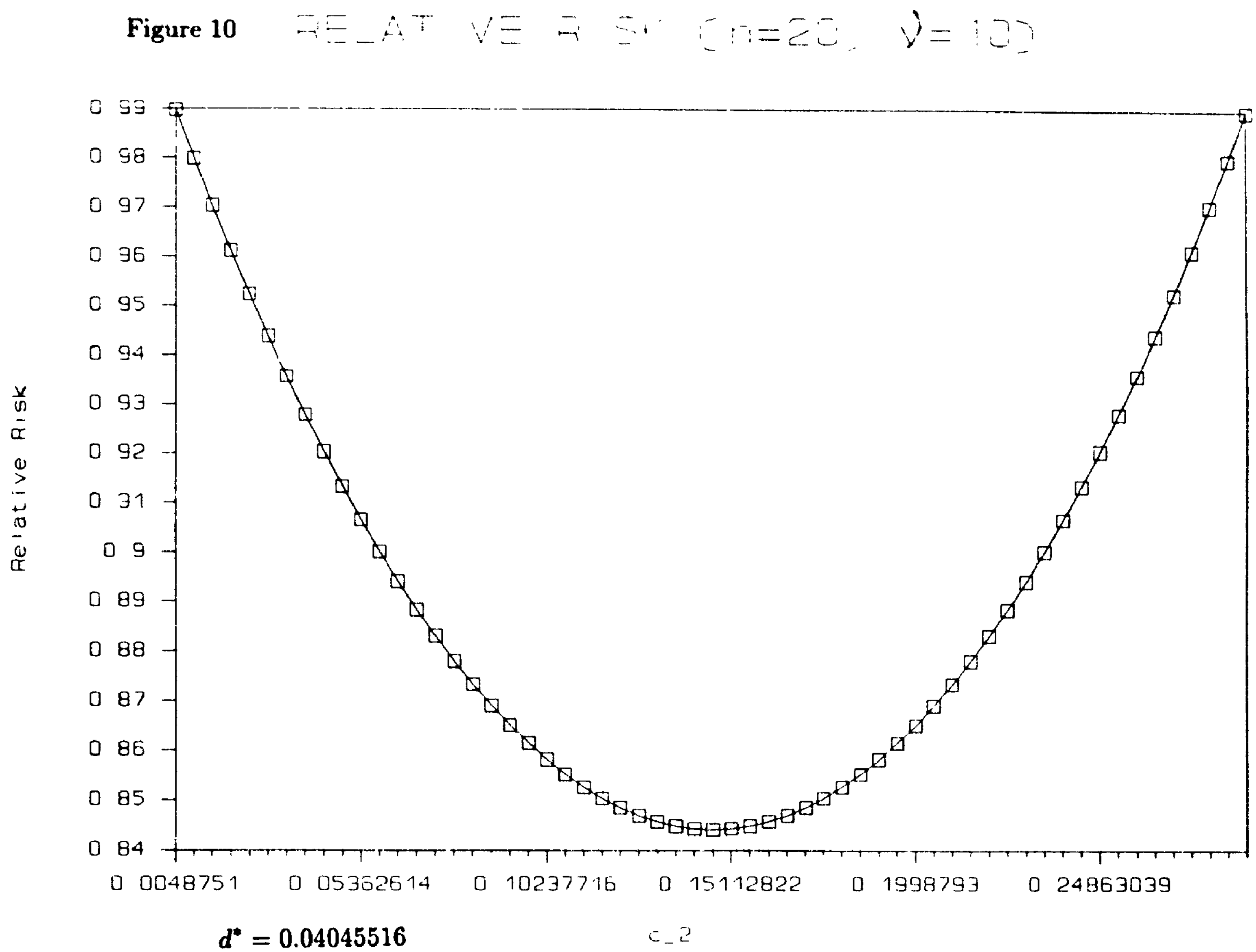
$d^* = 0.17143446$

Figure 9 RELATIVE RISK (n=20, ν=5)



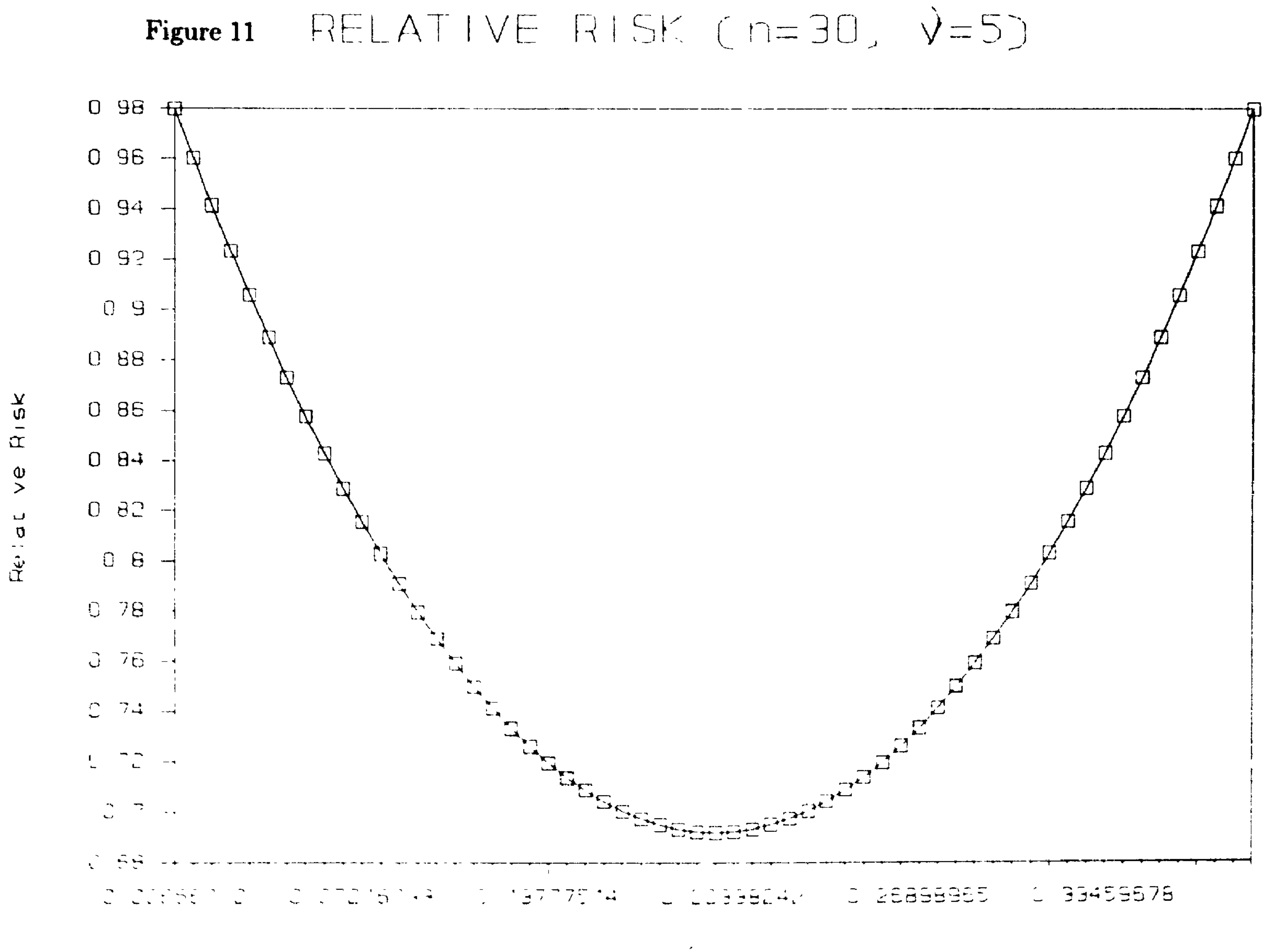
$d^* = 0.08258202$

Figure 10



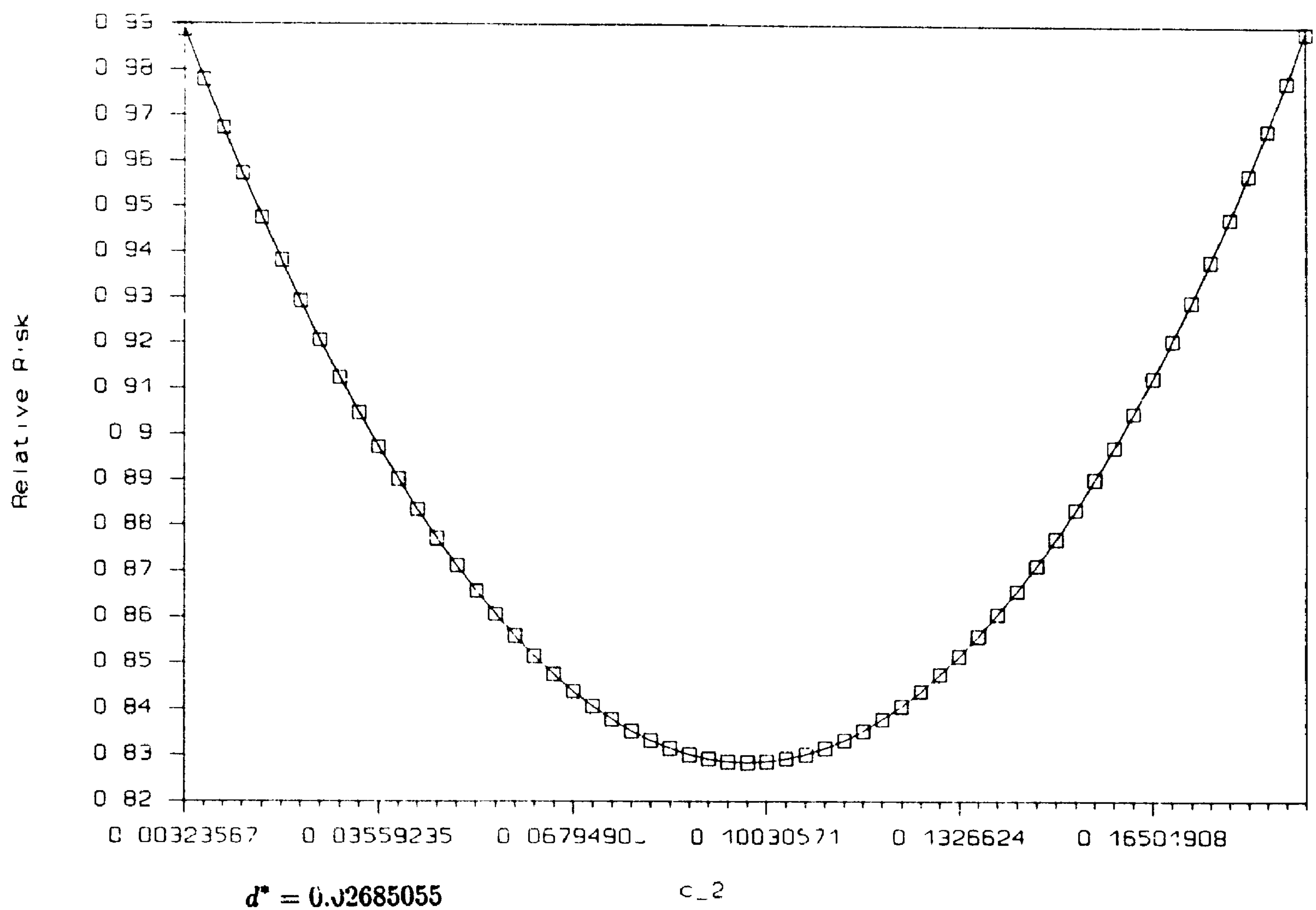
$d^* = 0.04045516$

Figure 11



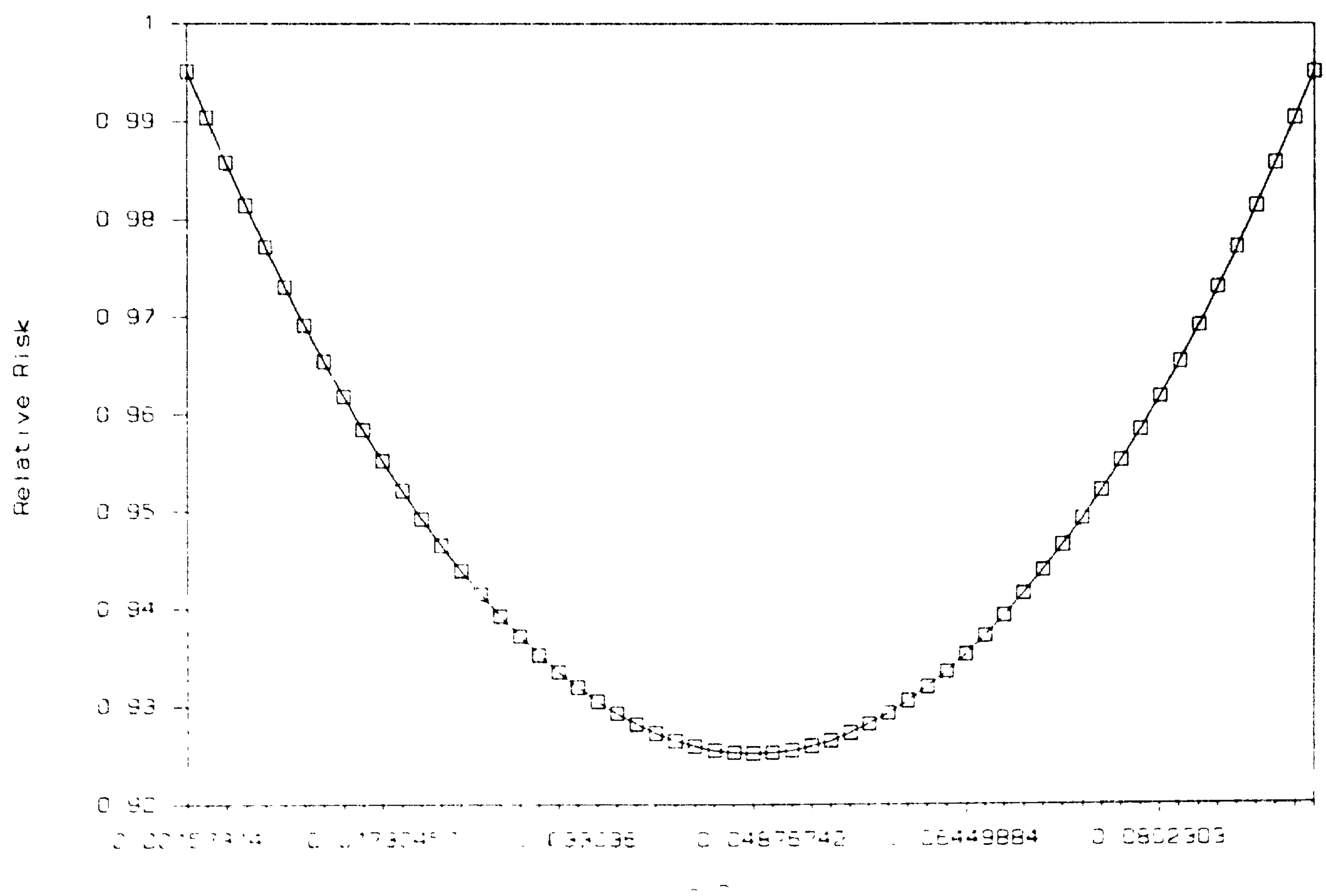
$d^* = 0.05444283$

Figure 12 RELATIVE RISK (n=30, y=10)



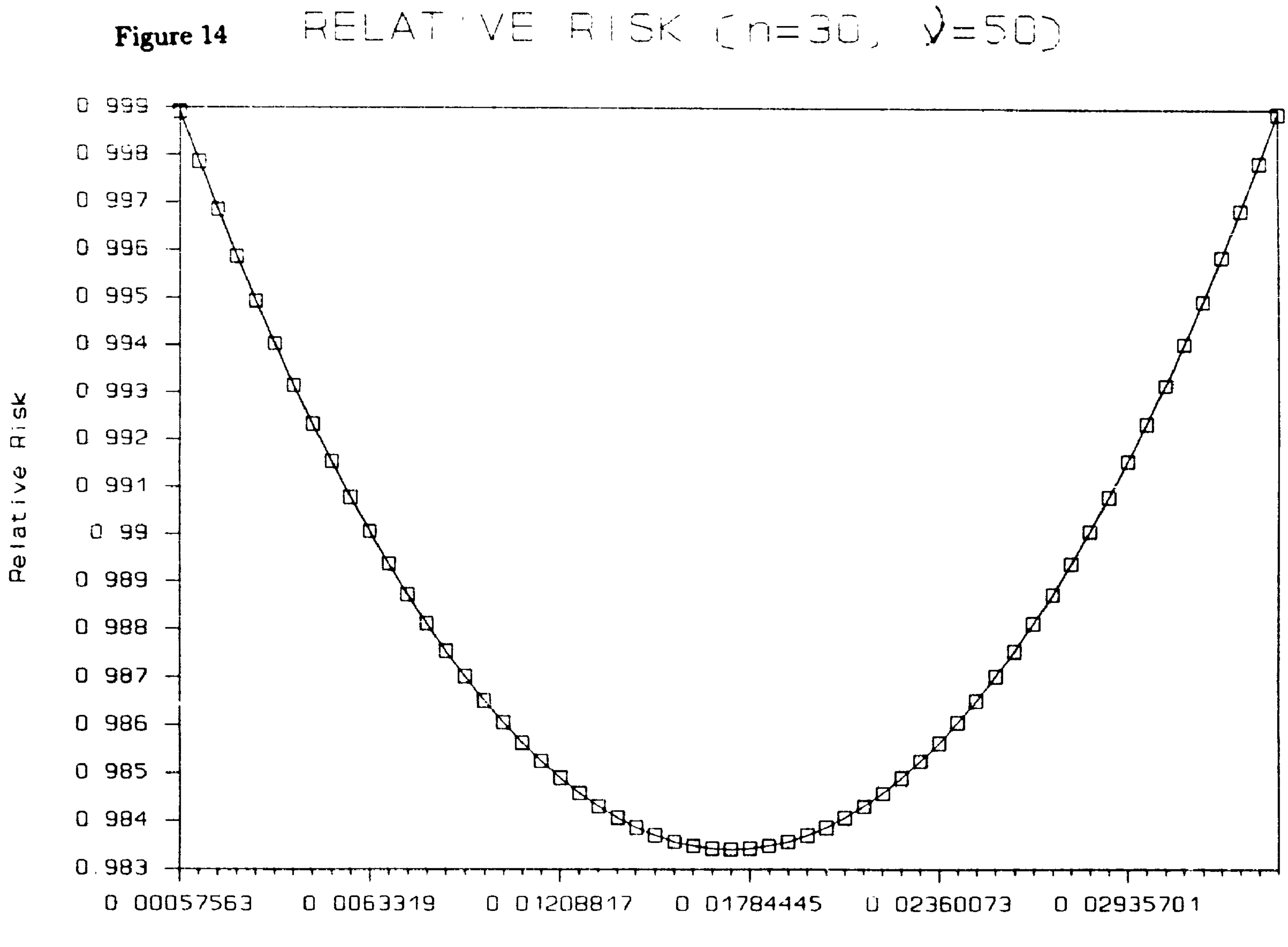
$d^* = 0.02685055$

Figure 13 RELATIVE RISK (n=30, ν=20)



$d^{*} = 0.01305441$

Figure 14



$d^* = 0.00477673$

CHAPTER 7

ESTIMATION OF THE SCALE MATRIX OF THE MULTIVARIATE T-MODEL UNDER THE ENTROPY LOSS FUNCTION

7.1 Estimation under Entropy Loss

Consider the multivariate t-model given by (2.3)

$$g(x_1, x_2, \ldots, x_N) = K(\nu, Np)|\Lambda|^{-N/2} \left(\nu + \sum_{j=1}^{N} (X_j - \theta)' \Lambda^{-1} (X_j - \theta) \right)^{-(\nu+Np)/2}$$

where

$$K(\nu, Np) = \frac{\nu^{\nu/2}\Gamma((\nu + Np)/2)}{\pi^{Np/2}\Gamma(\nu/2)}.$$

In Chapter 6 we have developed estimators for the scale matrix Λ which dominate the best multiple estimator based on the sum of products matrix A under a squared error loss function. In this chapter we develop estimators for the scale matrix Λ under the entropy loss function given by

$$L(u(A),\ \Lambda) = tr(\Lambda^{-1}u(A)) - ln|\Lambda^{-1}u(A)| - p \tag{7.1}$$

where $u(A)$ is any estimator of Λ based on the sample sum of products matrix A.

In estimating Λ by $u(A)$, we consider the risk function

$$R(u(A), \Lambda) = E[L(u(A), \Lambda)].$$

An estimator $u_2(A)$ of Λ will be said to dominate another estimator $u_1(A)$ of Λ if for all Λ belonging to the class of positive definite matrices the following inequality holds

$$R(u_2(A), \Lambda) \leq R(u_1(A), \Lambda),$$

and the inequality

$$R(u_2(A), \Lambda) < R(u_1(A), \Lambda),$$

holds for at least one Λ.

An estimator is said to be admissible if there exists no other estimator which dominates it. On the other hand an estimator is said to be inadmissible if it is dominated by another estimator.

Notations

In what follows we will use the following notations:

1. Sum of products matrix based on multivariate t-model:

$$A = \sum_{j=1}^{N} (X_j - \overline{X})(X_j - \overline{X})'.$$

2. Lower triangular decomposition of A:

$$A = TT',$$

where T a lower triangular matrix.

3. Characteristic roots of A:

$m_1, m_2, \ldots, m_p$ $(m_1 \geq m_2 \geq \ldots \geq m_p)$.

4. Spectral decomposition of A:

$$A = RMR',$$

where $M = Diag(m_1, m_2, \ldots, m_p)$ is the diagonal matrix with its characteristic roots as the diagonal elements and R the corresponding matrix of normalized characteristic vectors ($RR' = R'R = I$).

5. Unbiased estimator of Λ:

$$\Lambda_1^* = \frac{\nu - 2}{\nu n} A$$

6. James-Stein Type Estimator of Λ:

$$\Lambda_2^* = T\ Diag(d_1^*, d_2^*, \ldots, d_p^*) T'$$

where

$$d_i^* = \left(\frac{\nu}{\nu - 2} (n + 1 + p - 2i) \right)^{-1}, \quad i = 1, 2, \ldots, p.$$

7. Improved Estimator of Λ:

$$\Lambda_3^* = R \phi(M) R',$$

where

$$\phi(M) = Diag(d_1^* m_1, d_2^* m_2, \ldots, d_p^* m_p)$$

with d_i^* ($i = 1, 2, \ldots, p$) as given above.

8. Sum of products matrix based on N observations on $N_p(0, \Lambda)$:

$$W = \sum_{j=1}^{N} (Z_j - \overline{Z})(Z_j - \overline{Z})'.$$

9. Lower triangular decomposition of W:

$$W = UU',$$

where U a lower triangular matrix.

10. Spectral decomposition of W:

$$W = RLR',$$

where $L = Diag(l_1, l_2, \ldots, l_p)$ is the diagonal matrix with its characteristic roots as the diagonal elements and R the corresponding matrix of normalized characteristic vectors ($RR' = R'R = I$).

11. Characteristic roots of W:

$l_1, l_2, \ldots, l_p \quad (l_1 \geq l_2 \geq \ldots \geq l_p)$.

We will consider estimators of Λ based on (i) multiple of the sample sum of products matrix (ii) lower triangular decomposition of the sample sum of products matrix and (iii) spectral decomposition of the sample sum of products matrix.

In other words we consider

(i) estimators of the type cA, $c > 0$,

(ii) estimators of the type $T\Delta T'$ where T is an upper triangular matrix such that $A = TT'$ and Δ is an arbitrary positive definite diagonal matrix in analogy with the work, in the context of the multivariate normal distribution, by James and Stein (1961).

(iii) estimators of the type $R\phi(M)R'$ where

$$\phi(M) = Diag(\delta_1 m_1, \delta_2 m_2, \ldots, \delta_p m_p)$$

with positive numbers δ_i ($i = 1, 2 \ldots, p$), once again in analogy with the works of Stein (1975) and Dey and Srinivasan (1985).

Since we will develop estimators along the lines of Stein Estimation it seems appropriate to describe briefly what is meant by Stein effect.

Stein Effect

Suppose that $X_i \sim N(\theta_i, 1)$ and it is desired to estimate θ_i under the squared error loss function. Then it is well known that the sample mean $\delta_i(X_i) = X_i$ is admissible if the expe- . loss is the criterion. Consider next the problem of estimating the mean vector θ of p-dimensional normal population $N_p(\theta, I)$. The usual estimator of $\theta = (\theta_1, \theta_2, \ldots, \theta_p)'$ is the sample mean

$$\delta(X) = (\delta_1(X_1, \delta_2(X_2), \delta_p(X_p)) = (X_1, X_2, \ldots, X_p)'.$$

Stein (1956) proves that the sample mean $\delta(X)$ is inadmissible if the following risk function is the criterion:

$$R(\delta(X), \theta) = \sum_{i=1}^{p} R_i(\delta(X_i), \theta_i)$$

where

$$R_i(\delta(X_i), \theta_i) = (\delta_i(X_i) - \theta_i)^2 \quad (i = 1, 2, \ldots, p), \; p \geq 3.$$

James and Stein (1961) prove that the sample mean $\delta(X)$ is inadmissible when $p \geq 3$ and expected loss is the criterion; They prove that the sample mean $\delta(X)$ is dominated by the estimator

$$\delta^{JS}(X) = \left(1 - \frac{p-2}{X'X}\right) X.$$

In the above problem X_i's are independent and the θ_i's need not be related in any way. This surprising fact that one could combine unrelated problems to obtain improved estimators is commonly called the Stein effect (see, for example, Berger, 1988).

The papers by Stein (1956) and James and Stein (1961) have had a profound influence on estimation problems in statistics. Stein's ideas and theories have been extended in two main directions, namely, to more general loss functions, and to other distributions with location parameters.

Another important area is the estimation of the scale matrix of the multivariate normal distribution. James and Stein (1961) consider the estimation of the scale matrix Λ of the multivariate normal distribution $N_p(0, \Lambda)$ under the entropy loss function given by (7.1). It has been shown that the estimator UDU', where $D = Diag(d_1, d_2, \ldots, d_p)$, with

$$d_i = (n + p + 1 - 2i)^{-1}, \quad for \quad i = 1, 2, ..., p,$$

and U is the lower triangular matrix such that the Wishart matrix W can be decomposed as $W = UU'$, is a minimax estimator for the scale matrix of the multivariate normal distribution among the class of estimators $U\Delta U'$ where Δ is an arbitrary positive definite diagonal matrix.

Later Stein (1975) considers the following class of estimators for the scale matrix of the multivariate normal distribution:

$$R\phi(L)R'$$

where

$$L = Diag(l_1, l_2, \ldots, l_p),$$

$\phi(L)$ is a diagonal matrix with diagonal elements as some functions of the characteristic roots of W and the Wishart matrix W has the spectral decomposition given by $W = RLR'$ with R the matrix of normalized characteristic vectors ($RR' = R'R = I$). Dey and Srinivasan (1985) develop improved as well as minimax estimators for the scale matrix Λ of the multivariate normal distribution along the line of James and Stein (1961) and Stein (1975).

Some of these results have been extended in the present work to the case when the underlying observations have the multivariate t-model rather than the multivariate normal model.

In order to avoid digressions we prove some lemmas that will be needed in what follows.

Lemma 7.1 For any integer $n \geq p \geq 1$,

$$\sum_{i=1}^{p} ln\left(\frac{n+p-(2i-1)}{n}\right) \leq 0.$$

Proof. Consider the arithmetic mean and geometric mean of the numbers $n+p-1, n+p-3, n+p-5, \ldots, n+p-(2p-1)$. Then from the arithmetic and geometric mean inequality we have

$$n \geq \left(\prod_{i=1}^{p}\{n+p-(2i-1)\}\right)^{1/p},$$

or

$$1 \geq \prod_{i=1}^{p}\left(\frac{n+p-(2i-1)}{n}\right).$$

The proof is completed by taking logarithm in both sides of the above inequality.

Lemma 7.2 Consider the multivariate t-model given by (2.3) for $\nu > 2$. Then the following result holds:

$$E(A) = \frac{\nu n}{\nu-2}\Lambda$$

(cf. Sutradhar and Ali, 1989).

Proof. (i) By definition

$$E(A) = \int_{X \epsilon \Re^{Np}} \sum_{j=1}^{N} (X_j - \bar{X})(X_j - \bar{X})' \int_0^\infty \frac{|\tau^2\Lambda|^{-N/2}}{(2\pi)^{Np/2}}$$
$$\times\, exp\left(\frac{-1}{2}\sum_{j=1}^{N}(x_j-\theta)'(\tau^2\Lambda)^{-1}(x_j-\theta)\right) h(\tau)\; d\tau dX$$

where $X = (X_1, X_2, \ldots, X_N)$ (see for example, (2.5)).

Tb: ı nder the transformation

$$X_j = \theta + \tau Z_j, \quad (j = 1, 2, \ldots, N)$$

we have

$$E(A) = \int_0^\infty \int_{Z\epsilon\Re^{Np}} \tau^2 \sum_{j=1}^{N} (Z_j - \bar{Z})(Z_j - \bar{Z})' \; \frac{\tau^{-Np}|\Lambda|^{-N/2}}{(2\pi)^{Np/2}}$$

$$\times\, exp\left(\frac{-1}{2}\sum_{j=1}^{N} z_j'\Lambda^{-1}z_j\right) h(\tau)\; \tau^{Np} dZ\; d\tau$$

$$= \left(\int_0^\infty \tau^2 h(\tau) d\tau\right) \int_{Z\epsilon\Re^{Np}} \sum_{j=1}^{N} (Z_j - \bar{Z})(Z_j - \bar{Z})'$$

$$\times\, \frac{|\Lambda|^{-N/2}}{(2\pi)^{Np/2}} exp\left(\frac{-1}{2}\sum_{j=1}^{N} z_j'\Lambda^{-1}z_j\right) \; dZ.$$

Hence by virtue of

$$E(\tau^2) = \frac{\nu}{\nu - 2}$$

and

$$E(W) = n\Lambda$$

we finally have

$$E(A) = \frac{\nu}{\nu - 2} \; (n\Lambda).$$

Lemma 7.3 Consider the multivariate t-model for $\nu > 2$. Then the following two identities hold:

(i)

$$E\left[ln(|A|)\right] = E\left[ln(|W|)\right] + 2pE(ln\tau) \qquad (7.2)$$

where

$$W = \sum_{j=1}^{N} (Z_j - \bar{Z})(Z_j - \bar{Z})'$$

with $Z_j \sim N(0,\Lambda)$, $(j = 1,2,\ldots,N)$.

(ii)

$$E\left[ln(|\Lambda^{-1}A|)\right] = \sum_{i=1}^{p} E\left[ln(\chi^2_{n+1-i})\right] + 2pE(ln\tau) \tag{7.3}$$

where τ^{-2} has a gamma distribution given by $G(\nu/2, 2/\nu)$.

Proof. (i) The proof is very analogous to Lemma 7.2.

(ii) It follows from (7.2) that

$$\begin{aligned} E[ln(\Lambda^{-1}A)] =& E[ln(A)] - ln(\Lambda) \\ =& E\left[ln(|W|)\right] + 2pE(ln\tau) - ln(\Lambda). \end{aligned}$$

Since

$$\frac{|W|}{|\Lambda|} \sim \prod_{i=1}^{p} \chi^2_{n+1-i},$$

(see e.g. Muirhead, 1982, pp 85, 100) it follows that

$$\begin{aligned} E[ln(\Lambda^{-1}A)] =& E\left[ln\left(\prod_{i=1}^{p} \chi^2_{n+1-i}\right)\right] + 2pE(ln\tau) \\ =& \sum_{i=1}^{p} E\left[ln\left(\chi^2_{n+1-i}\right)\right] + 2pE(ln\tau) \end{aligned}$$

where τ^{-2} has a gamma distribution given by $G(\nu/2, 2/\nu)$.

Lemma 7.4 Consider the triangular decomposition of the sum of products matrix

$$A = \sum_{j=1}^{N} (X_j - \bar{X})(X_j - \bar{X})' = TT'$$

where T is a lower triangular matrix and $X = (X_1, X_2, \ldots, X_N)$ has the multivariate t-model. Then for $\nu > 2$ the following identities hold:

(i)

$$E[tr(\Lambda^{-1}T\Delta T')] = \frac{\nu}{\nu - 2}\; E[tr(\Lambda^{-1}U\Delta U')], \tag{7.4}$$

(ii)

$$E[ln(|\Lambda^{-1}T\Delta T'|)] = E[ln(|\Lambda^{-1}U\Delta U'|)] + 2pE[ln(\tau)], \tag{7.5}$$

where U is a lower triangular matrix such that

$$W = \sum_{j=1}^{N}(Z_j - \bar{Z})(Z_j - \bar{Z})' = UU'$$

with $Z_j \sim N_p(0, \Lambda)$ and $\tau^{-2} \sim G(\nu/2, 2/\nu)$.

Proof. (i) As in Lemma 7.2 we have

$$E[tr(\Lambda^{-1}T\Delta T')] = \int_{X\epsilon\Re^{Np}} tr(\Lambda^{-1}T\Delta T') \int_0^\infty \frac{|\tau^2\Lambda|^{-N/2}}{(2\pi)^{Np/2}}$$
$$\times\, exp\left(\frac{-1}{2}\sum_{j=1}^{N}(x_j - \theta)'(\tau^2\Lambda)^{-1}(x_j - \theta)\right) h(\tau)\; d\tau dX.$$

Under the transformation $X_j = \theta + \tau Z_j$, $(j = 1, 2, \ldots, p)$, the sum of products matrix

$$\sum_{j=1}^{N}(X_j - \bar{X})(X_j - \bar{X})' = TT'$$

becomes

$$\tau^2\sum_{j=1}^{N}(Z_j - \bar{Z})(Z_j - \bar{Z})' = \tau^2 UU'$$

so that

$$E[tr(\Lambda^{-1}T\Delta T')] = \int_0^\infty \int_{Z\epsilon\Re^{Np}} E[tr(\Lambda^{-1}\tau U\Delta\tau U')]\;\; \frac{\tau^{-Np}|\Lambda|^{-N/2}}{(2\pi)^{Np/2}}$$
$$\times\, exp\left(\frac{-1}{2}\sum_{j=1}^{N} z_j'\Lambda^{-1}z_j\right) h(\tau)\; \tau^{Np} dZ\; d\tau$$
$$= \left(\int_0^\infty \tau^2 h(\tau)d\tau\right)\int_{Z\epsilon\Re^{Np}} tr(\Lambda^{-1}U\Delta U')$$
$$\times\, \frac{|\Lambda|^{-N/2}}{(2\pi)^{Np/2}} exp\left(\frac{-1}{2}\sum_{j=1}^{N} z_j'\Lambda^{-1}z_j\right)\; dZ$$
$$= \frac{\nu}{\nu - 2}\;\; E[tr(\Lambda^{-1}U\Delta U')].$$

Hence (7.4) is proved. (ii) The proof for (7.5) runs exactly along the same lines as above.

Lemma 7.5 Consider the spectral decomposition of the sum of products matrix

$$A = \sum_{j=1}^{N}(X_j - \bar{X})(X_j - \bar{X})' = RMR'$$

where $X = (X_1, X_2, \ldots, X_N)$ have the multivariate t-model. Then for $\nu > 2$ the following identities hold:

(i)

$$E[ln(\Lambda^{-1}R\phi(M)R')] = \sum_{i=1}^{p} E[ln(\chi^2 - n + 1 - i)] + \sum_{i=1}^{p} ln(d_i^*) + 2pE(ln\tau), \quad (7.6)$$

(ii)

$$E[tr(\Lambda^{-1}R\phi(M)R' = 2E\left[\sum_{i=1}^{p}\sum_{t=i+1}^{p}\frac{d_i l_i - d_t l_t}{l_i - l_t}\right] + (n - p + 1)\sum_{.=1}^{p} d_i \quad (7.7)$$

where l_i's and m_i's $(i = 1, 2, \ldots, p)$ are the characteristic roots of W and A respectively,

$$d_i^* = \left(\frac{\nu}{\nu - 2}(n + p + 1 - 2i)\right)^{-1} = (\frac{\nu}{\nu - 2})^{-1} d_i, \quad (i = 1, 2, \ldots, p),$$

and

$$\phi(M) = Diag(d_1^* m_1, d_2^* m_2, \ldots, d_p^* m_p).$$

Proof. (i) It is easily checked that

$$\begin{aligned} E[ln(\Lambda^{-1}R\phi(M)R')] =& ln(|\Lambda^{-1}|) + E[ln(|\phi(M)|)] + ln(|RR'|) \\ =& ln(|\Lambda^{-1}|) + ln(\prod_{i=1}^{p} d_i^* m_i) \\ =& ln(|\Lambda^{-1}|) + \sum_{i=1}^{p} ln(d_i^*) + E[ln(|A|)] \\ =& E[ln(|\Lambda^{-1}A|)] + \sum_{i=1}^{p} ln(d_i^*). \end{aligned}$$

Then by the use of (7.3) we have (7.6).

(ii) We have

$$\begin{aligned} & E[tr(\Lambda^{-1}R\phi(M)Diag(d_1^* m_1, d_2^* m_2, \ldots, d_p^* m_p)R')] \\ &= \int_{X\epsilon\Re^{Np}} tr(\Lambda^{-1}Diag(d_1^* m_1, d_2^* m_2, \ldots, d_p^* m_p)) \int_0^\infty \frac{|\tau^2\Lambda|^{-N/2}}{(2\pi)^{Np/2}} \\ &\times exp\left(\frac{-1}{2}\sum_{j=1}^{N}(x_j-\theta)'(\tau^2\Lambda)^{-1}(x_j-\theta)\right) h(\tau)\, d\tau dX. \end{aligned}$$

Under the transformation $X_j = \theta + \tau Z_j$, $(j = 1, 2, \ldots, p)$, the sum of products matrix

$$R\ Diag(m_1, m_2, \ldots, m_p)R' = \sum_{j=1}^{N}(X_j - \bar{X})(X_j - \bar{X})'$$

becomes

$$\tau^2 \sum_{j=1}^{N}(Z_j - \bar{Z})(Z_j - \bar{Z})' = R\tau^2 Diag(l_1, l_2, \ldots, l_p)R'$$

so that $m_i = \tau^2 l_i$, $(i = 1, 2, \ldots, p)$.

Then as in Lemma 7.2 it is easily verified that

$$\begin{aligned} & E[tr(\Lambda^{-1}RDiag(d_1^* m_1, d_2^* m_2, \ldots, d_p^* m_p)R')] \\ &= \left(\int_0^\infty \tau^2 h(\tau)\, d\tau\right) \int_{Z\epsilon\Re^{Np}} tr(\Lambda^{-1}RDiag(d_1^* l_1, d_2^* l_2, \ldots, d_p^* l_p)R') \\ &\times \frac{|\Lambda|^{-N/2}}{(2\pi)^{Np/2}} exp\left(\frac{-1}{2}\sum_{j=1}^{N} z_j'\Lambda z_j\right) dZ. \end{aligned}$$

Now by the use of

$$E(\tau^2) = \frac{\nu}{\nu - 2}$$

and

$$d_i^* = \left(\frac{\nu}{\nu - 2}\right)^{-1} d_i, \quad (i = 1, 2, \ldots, p)$$

we have

$$\begin{aligned} & E[tr(\Lambda^{-1} R Diag(d_1^* m_1, d_2^* m_2, \ldots, d_p^* m_p) R')] \\ &= \int_{Z \epsilon \Re^{Np}} tr(\Lambda^{-1} R Diag(d_1 l_1, d_2 l_2, \ldots, d_p l_p) R') \\ &\times \frac{|\Lambda|^{-N/2}}{(2\pi)^{Np/2}} \, exp\left(\frac{-1}{2}\sum_{j=1}^{N} z_j' \Lambda z_j\right) dZ. \end{aligned} \tag{7.8}$$

We then have from Dey and Srinivasan (1985, p 1583) that

$$\begin{aligned} & \int_{Z \epsilon \Re^{Np}} tr(\Lambda^{-1} R \phi(L) R') \, \frac{|\Lambda|^{-N/2}}{(2\pi)^{Np/2}} \, exp\left(\frac{-1}{2}\sum_{j=1}^{N} z_j' \Lambda z_j\right) dZ \\ &= 2E\left[\sum_{i=1}^{p}\sum_{t=i+1}^{p} \frac{\phi_i(L) - \phi_t(L)}{l_i - l_t} + 2\sum_{i=1}^{p} \frac{\partial \phi_i(L)}{\partial l_i} + (n - p - 1)\sum_{i=1}^{p} \frac{\phi_i(L)}{l_i}\right] \end{aligned} \tag{7.9}$$

where $\phi(L) = Diag(\phi_1(L), \phi_2(L), \ldots, \phi_p(L))$. Thus the Lemma 7.5 follows from (7.8) and (7.9) with $\phi_i(L) = d_i l_i$.

Finally we restate somewhat more precisely the results we shall prove in this chapter.

1. Among the class of estimators of Λ, of the form cA, the estimator Λ_1^* has the minimum risk under the entropy loss function given by (7.1)(see Theorem 7.1). This estimator happens to be the unbiased estimator of Λ.

2. The estimator Λ_2^* is the minimum risk estimator among the class of estimators of the form $T\Delta T'$ where Δ is any arbitrary positive definite diagonal matrix (see Theorem 7.2); Also the estimator Λ_2^* dominates the estimator Λ_1^* under the entropy loss function given by (7.1).

3. The estimator Λ_3^* dominates the abovementioned estimator Λ_2^* under the entropy loss function (see Theorem 7.3).

7.2 Estimators Based on Multiples of the Sample Sum of Products Matrix

As mentioned in Chapter 6, the scale matrix Λ of the multivariate normal distribution is usually estimated by cW where $c > 0$ and W is the usual Wishart matrix. It is well known (see e.g. Muirhead, 1982, p 129) that under the entropy loss function given by (7.1), the best estimator (smallest risk) of the scale matrix of the multivariate normal distribution, of the form cW, is given by W/n.

In this section we consider estimators of the form $\tilde{\Lambda} = cA$, where $c > 0$, for the scale matrix Λ of the multivariate t-model and find optimum value of c for which the risk function of the estimator under the entropy loss function given by (7.1) is minimized. We will prove that the optimum value of c is given by $c = (\nu - 2)/(\nu n)$ and resulting estimator of the scale matrix is given by

$$\Lambda_1^* = \frac{\nu - 2}{\nu n} A$$

which is also unbiased. We find that the best (smallest risk) estimator of the scale matrix of the multivariate t-model, of the form $\tilde{\Lambda} = cA$ is given by Λ_1^*.

It may be mentioned that the maximum likelihood estimator of Λ is given by

$$\tilde{\Lambda}_1 = \frac{1}{n+1} A$$

(Anderson and Fang, 1990a, p 208). Since the maximum likelihood estimator of Λ belongs to the class cA, it follows that the unbiased estimator Λ_2^* of Λ dominates the maximum likelihood estimator $\tilde{\Lambda}_1$.

Theorem 7.1 Consider the multivariate t-model given by (6.1) for $\nu > 2$. Then under the entropy loss function the unbiased estimator of Λ, namely

$$\Lambda_1^* = \frac{\nu - 2}{\nu n} A$$

has the smallest risk among the class of estimators $\tilde{\Lambda} = cA$, *for* $c > 0$ and the corresponding minimum risk is given by

$$\begin{aligned} R(\Lambda_1^*,\ \Lambda) =& p\ ln(n) - \sum_{i=1}^{p} E[ln(\chi^2_{n+1-i})] \\ &+ p\ ln\left(\frac{\nu}{\nu - 2}\right) - 2p\ E(ln\ \tau) \end{aligned} \tag{7.10}$$

where τ^{-2} has a gamma distribution $G(\nu/2, 2/\nu)$.

Proof. The risk function of the estimator cA is given by

$$\begin{aligned} R(cA,\ \Lambda) =& E[L(cA),\ \Lambda] \\ =& E[tr(\Lambda^{-1}cA) - ln(|\Lambda^{-1}cA|) - p] \\ =& c\ tr[\Lambda^{-1}E(A)] - p\ ln(c) - E[ln\ (|\Lambda^{-1}A|)] - p. \end{aligned}$$

Then it follows from Lemmas 7.2 and 7.3 that

$$\begin{aligned} R(cA,\ \Lambda) =& c\ tr\left[\Lambda^{-1}\frac{\nu n}{\nu - 2}\Lambda\right] - p\ ln(c) \\ &- \left(\sum_{i=1}^{p} E[ln(\chi^2_{n+1-i})] + 2p\ E(ln\tau)\right) - p \\ =& \frac{\nu np}{\nu - 2}\ c - \sum_{i=1}^{p} E[ln(\chi^2_{n+1-i})] - p\ ln(c) - p - 2p\ E(ln\ \tau). \end{aligned} \tag{7.11}$$

Now taking derivatives we have

$$\begin{aligned} \frac{\partial R(cA,\ \Lambda)}{\partial c} =& \frac{\nu np}{\nu - 2} - \frac{p}{c}, \\ \frac{\partial^2 R(cA,\ \Lambda)}{\partial c^2} =& \frac{p}{c^2} > 0. \end{aligned}$$

Hence $c = (\nu - 2)/(\nu n)$ minimizes the risk function given by (7.11) and the corresponding estimator is given by $\Lambda_1^* = (\nu - 2)A/(\nu n)$.

It follows from (7.11), by putting $c = (\nu - 2)/(\nu n)$, that the risk function of the estimator Λ_1^* is given by

$$R(\Lambda_1^*,\ \Lambda) = -\sum_{i=1}^{p} E[ln(\chi^2_{n+1-i})] - v\ ln\left(\frac{\nu - 2}{\nu n}\right) - 2p\ E(ln\ \tau),$$

or

$$R(\Lambda_1^*,\ \Lambda) = -\sum_{i=1}^{p} E[ln(\chi^2_{n+1-i})] + p\ ln(n)$$
$$+p\ ln\left(\frac{\nu}{\nu - 2}\right) - 2p\ E(ln\ \tau).$$

7.3 Estimators Based on a Triangular Decomposition of the Sample Sum of Products Matrix

Following James and Stein (1961) we propose estimators of the form $T\Delta T'$ where T is a lower triangular matrix such that the sample sum of products matrix A has the decomposition $A = TT'$ and Δ an arbitrary positive definite diagonal matrix. We find the optimum value of Δ for which the risk function of the estimator $T\Delta T'$ under the entropy loss function is minimized and denote it by D^*. We will call the resulting estimator as the James and Stein Type Estimator and denote it by $\Lambda_2^* = TD^*T'$.

Theorem 7.2 Under the entropy loss function given by (7.1), the estimator

$$\Lambda_2^* = TD^*T'$$

where T is a lower triangular matrix such that $A = TT'$ and

$$D^* = Diag(d_1^*, d_2^*, \ldots, d_p^*)$$

with

$$d_i^* = \left(\frac{\nu}{\nu-2}(n+1+p-2i)\right)^{-1} = \left(\frac{\nu}{\nu-2}\right)^{-1} d_i, \quad i = 1, 2, \ldots, p, \tag{7.12}$$

has the smallest risk among the class of estimators $T\Delta T'$ where Δ belongs to the class of all positive definite diagonal matrices, and the risk function of the estimator Λ_2^* is given by

$$\begin{aligned} R(\Lambda_2^*,\ \Lambda) = &\sum_{i=1}^{p} ln(n+1+p-2i) - \sum_{i=1}^{p} E[ln(\chi^2_{n+1-i})] \\ &+p\ ln\left(\frac{\nu}{\nu-2}\right) - 2p\ E(ln\tau) \end{aligned} \tag{7.13}$$

where τ^{-2} has a gamma distribution $G(\nu/2, 2/\nu)$.

Furthermore Λ_2^* dominates the unbiased estimator $\Lambda_1^* = (\nu-2)A/(\nu n)$.

Proof. The risk function of the estimator $T\Delta T'$ is given by

$$\begin{aligned} R(T\Delta T',\ \Lambda) =&E\left[tr(\Lambda^{-1}T\Delta T') - ln|\Lambda^{-1}T\Delta T'| - p\right] \\ =&E\left[tr(\Lambda^{-1}T\Delta T')\right] - E\left[ln(|\Lambda^{-1}T\Delta T'|)\right] - p. \end{aligned}$$

Then from Lemma 7.4 we have

$$\begin{aligned} R(T\Delta T',\ \Lambda) ==&\frac{\nu}{\nu-2}E\left[tr(\Lambda^{-1}U\Delta U')\right] - E\left[ln(|\Lambda^{-1}U\Delta U'|)\right] \\ &-2p\ E(ln\tau) - p. \end{aligned} \tag{7.14}$$

Then following Muirhead (1982, pp 130-132), it can be proved that the risk function given by (7.14) does not depend on Λ and that

$$\begin{aligned} R(T\Delta T',\ \Lambda) =&\frac{\nu}{\nu-2}\sum_{i=1}^{p}\delta_i(n+1+p-2i) - 2p\ E(ln\tau) \\ &-\left[\sum_{i=1}^{p} ln(\delta_i) + \sum_{i=1}^{p} E[ln(\chi^2_{n+1-i})]\right] - p. \end{aligned} \tag{7.15}$$

This attains its minimum value of

$$R(\Lambda_2^*,\ \Lambda) = -\,2p\ E(ln\tau) + \sum_{i=1}^{p} ln\left[\frac{\nu}{\nu-2}(n+1+p-2i)\right]$$
$$-\sum_{i=1}^{p} E[ln(\chi^2_{n+1-i})]$$

when

$$\delta_i = \left(\frac{\nu}{\nu-2}(n+1+p-2i)\right)^{-1} = d_i^*, (say)\ \ i = 1, 2, \ldots, p.$$

The risk function of the unbiased estimator Λ_1^* is given by (7.10) as

$$R(\Lambda_1^*,\ \Lambda) = p\ ln(n) - \sum_{i=1}^{p} E[ln(\chi^2_{n+1-i})]$$
$$+p\ ln\left(\frac{\nu}{\nu-2}\right) - 2p\ E(ln\ \tau).$$

while the risk of the James-Stien Type Estimator Λ_2^* is given by

$$R(\Lambda_2^*, \Lambda) = -\,2pE(ln\tau) + \sum_{i=1}^{p} ln\left[\frac{\nu}{\nu-2}(n+1+p-2i)\right]$$
$$-\sum_{i=1}^{p} E[ln(\chi^2_{n+1-i})].$$

so that

$$R(\Lambda_2^*, \Lambda) - R(\Lambda_1^*, \Lambda) = \sum_{i=1}^{p} ln\left[\frac{\nu}{\nu-2}(n+1+p-2i)\right] - p\ ln\left(\frac{\nu n}{\nu-2}\right)$$
$$= \sum_{i=1}^{p} ln\left(\frac{n+1+p-2i}{n}\right).$$

Hence it follows from Lemma 7.1 that the estimator Λ_2^* dominates the unbiased estimator Λ_1^*.

7.4 Estimators Based on the Spectral Decomposition of the Sample Sum of Products Matrix

As mentioned earlier, the sum of products matrix A can be decomposed as $A = RMR'$.

We consider, following Dey and Srinivasan (1985), estimators of the form $R\phi(M)R'$ where

$$\phi(M) = Diag(\phi_1(M), \phi_2(M), \ldots, \phi_p(M))$$

and $\phi_i(M)(i = 1, 2, \ldots, p)$ is a function of the characteristic roots $m_1, m_2, \ldots m_p$ satisfying $\phi_i(M) > 0$. We note that if

$$\phi_i(M) = \frac{1}{n+1} m_i, \quad i = 1, 2, \ldots, p,$$

then

$$\phi(M) = \frac{1}{n+1} M$$

and consequently

$$R\phi(M)R' = \frac{1}{n+1} RMR' = \frac{1}{n+1} A$$

which is the maximum likelihood estimator $\tilde{\Lambda}_1$ of Λ. We also note that if

$$\phi_i(M) = \frac{(\nu-2)m_i}{\nu n}, \quad i = 1, 2, \ldots, p.$$

then

$$\phi(M) = \frac{\nu-2}{\nu n} M$$

and consequently

$$R\phi(M)R' = \frac{\nu-2}{\nu n} RMR' = \frac{\nu-2}{\nu n} A$$

which is the unbiased estimator of Λ (Anderson and Fang, 1990a, p 208).

In the following theorem we consider an estimator Λ_3^* of Λ based on the above spectral decomposition of the sample sum of products matrix. We prove that this estimator dominates Λ_2^* considered in Theorem 7.2.

Theorem 7.3 Let $\Lambda_3^* = R\phi(M)R$ be an estimator for Λ where A has the spectral decomposition $A = RMR'$, with

$$M = Diag(m_1, m_2, \ldots, m_p), \quad m_1 \geq m_2 \geq \ldots, \geq m_p,$$

and

$$\phi(M) = Diag(d_1^* m_1, d_2^* m_2, \ldots, d_p^* m_p)$$

with d_i^*'s as given by (7.12). Then Λ_3^* dominates Λ_2^* .

Proof. The risk function of the estimator Λ_3^* is given by

$$\begin{aligned} R(\Lambda_3^*,\ \Lambda) =& E\left[tr(\Lambda^{-1}\Lambda_3^* - ln|\Lambda^{-1}\Lambda_3^*| - p\right] \\ =& E[tr(\Lambda^{-1}R\phi(M)R')] - E[ln(|\Lambda^{-1}R\phi(M)R'|)] - p. \end{aligned}$$

Then it follows from Lemma 7.5 that

$$\begin{aligned} R(\Lambda_3^*, \Lambda) =& 2E\left[\sum_{i=1}^{p}\sum_{t=i+1}^{p}\frac{d_i l_i - d_t l_t}{l_i - l_t}\right] + (n-p+1)\sum_{i=1}^{p} d_i \\ & - \left(\sum_{i=1}^{p} ln(d_i^*) + \sum_{i=1}^{p} E[ln(\chi^2_{n+1-i})] + 2pE[ln(\tau)]\right) - p, \end{aligned} \tag{7.16}$$

where l_i's are the characteristic roots of the Wishart matrix W and d_i's are given by (7.12). Now

$$\begin{aligned} & \sum_{i=1}^{p}\sum_{t=i+1}^{p}\frac{d_i l_i - d_t l_t}{l_i - l_t} \\ =& \sum_{i=1}^{p}\sum_{t=i+1}^{p}\frac{l_i(d_i - d_t)}{l_i - l_t} + \sum_{i=1}^{p}\sum_{t=i+1}^{p} d_t. \end{aligned}$$

But for $t = i+1, i+2, \ldots, p$; $i = 1, 2, \ldots, p$, we have $d_i \le d_t$, $l_i \ge l_t$ so that

$$\frac{l_i}{l_i - l_t} > 1$$

and

$$\frac{l_i(d_i - d_t)}{l_i - l_t} \le d_i - d_t \; ;$$

We then have

$$\begin{aligned}
&\sum_{i=1}^{p}\sum_{t=i+1}^{p}\frac{d_i l_i - d_t l_t}{l_i - l_t}\\
\le &\sum_{i=1}^{p}\sum_{t=i+1}^{p}(d_i - d_t) + \sum_{i=1}^{p}\sum_{t=i+1}^{p} d_t\\
= &\sum_{i=1}^{p}\sum_{t=i+1}^{p} d_i\\
= &\sum_{i=1}^{p}(p-i)d_i.
\end{aligned} \tag{7.17}$$

Finally by the use of (7.17) in (7.16) we have

$$\begin{aligned}
R(\Lambda_3^*, \Lambda) \le & 2 \sum_{i=1}^{p}(p-i)d_i + (n-p+1)\sum_{i=1}^{p} d_i\\
& - \left(\sum_{i=1}^{p} ln(d_i^*) + \sum_{i=1}^{p} E[ln(\chi^2_{n+1-i})] + 2pE[ln(\tau)]\right) - p\\
= & -\sum_{i=1}^{p} ln(d_i^*) - \sum_{i=1}^{p} E[ln(\chi^2_{n+1-i})] - 2p\, E[ln(\tau)]\\
= & R(\Lambda_2^*, \Lambda) \quad \text{(see (7.12) and (7.13))}
\end{aligned}$$

which means that Λ_3^* dominates Λ_2^*.

Finally we remark that the estimator Λ_3^* dominates the estimator Λ_2^*, and the estimator Λ_2^* dominates the estimator Λ_1^*. It happens that the estimator Λ_1^* is the unbiased estimator of Λ. However, the unbiased estimator of Λ dominates the maximum likelihood estimator under the entropy loss function (see Section 7.2).

7.4.1 An Example

Suppose we have the following observed sum of products matrix

$$A = \begin{pmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{pmatrix}$$

with $n = 10$, $\nu = 5$ and $p = 3$. We calculate the estimators

$$\Lambda_1^* = \frac{\nu - 2}{\nu n} A,$$

$$\Lambda_2^* = TD^*T'$$

and

$$\Lambda_3^* = R\phi(M)R'$$

where

$$D^* = Diag(d_1^*, d_2^*, d_3^*)$$

$$M = Diag(m_1, m_2, m_3)$$

and

$$\phi(M) = Diag(d_1^* m_1, d_2^* m_2, d_3^* m_3).$$

It is easy to verify that $D^* = Diag(.05, .06, .075)$,

$$T = \begin{pmatrix} 3.60555128 & 0.00000000 & 0.00000000 \\ -1.10940039 & 3.43063125 & 0.00000000 \\ 0.55470020 & -0.40360368 & 3.08697453 \end{pmatrix},$$

$m_1 = 18, m_2 = 9, m_3 = 9, M = Diag(18, 9, 9)$ and

$$R = \begin{pmatrix} 0.66666667 & 0.70710678 & 0.23570226 \\ -0.66666667 & 0.70710678 & -0.23570226 \\ 0.33333333 & 0.00000000 & -0.94280904 \end{pmatrix}$$

so that $A = TT'$, $A = RMR'$ and

$$\phi(M) = \begin{pmatrix} 0.90000000 & 0.00000000 & 0.00000000 \\ 0.00000000 & 0.54000000 & 0.00000000 \\ 0.00000000 & 0.00000000 & 0.67500000 \end{pmatrix}.$$

Hence the unbiased estimator $\Lambda_1^* = (\nu - 2)A/(\nu n)$ is given by

$$\Lambda_1^* = \begin{pmatrix} 0.78000000 & -0.24000000 & 0.12000000 \\ -0.24000000 & 0.78000000 & -0.12000000 \\ 0.12000000 & -0.12000000 & 0.60000000 \end{pmatrix}$$

and the estimator Λ_2^* is given by

$$\Lambda_2^* = \begin{pmatrix} 0.65000000 & -0.20000000 & 0.10000000 \\ -0.20000000 & 0.76769231 & -0.11384615 \\ 0.10000000 & -0.11384615 & 0.73986425 \end{pmatrix}$$

while the estimator Λ_3^* is given by

$$\Lambda_3^* = \begin{pmatrix} 0.70750000 & -0.16750000 & 0.05000000 \\ -0.16750000 & 0.70750000 & -0.05000000 \\ 0.05000000 & -0.05000000 & 0.70000000 \end{pmatrix}.$$

REFERENCES

Ali, M.M. (1980). Characterization of the normal distribution among the continuous symmetric spherical class. *J. Roy. Statist Soc.*, B, **42**, 162–164.

Ali, M.M. (1989). Characterizations of spherical distributions. *Proceedings of the First Islamic Countries Conference on Statistical Science.* Lahore, Pakistan.

Ali, M. M. and Joarder, A.H. (1991). Distribution of the correlation coefficient for the class of bivariate elliptical models. *Canad. J. Statist.*, **19**, 447–452.

Amari, S.I. (1985). Differential – Geometric Methods in Statistics , Lecture Notes in Statistics **28**. Springer-Verlag, Berlin.

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

Anderson, T.W. (1984). An Introduction to Multivariate Statistical Analysis. John Wiley, New York.

Anderson, T.W.; Fang, K.T. and Hsu, H. (1986) Maximum likelihood estimates and likelihood-ratio criteria for multivariate elliptically contoured distributions. *Canad. J. Statist.,* 14, 55–59.

Anderson T.W. and Fang, K.T. (1987). Cochran's theorem for elliptically contoured distributions. *Sankhya,* Series A, **49**, 303–315.

Anderson, T.W. and Fang, K.T. (1990). On the theory of multivariate elliptically contoured distributions and their applications. *Statistical Inference in Elliptically Contoured and Related Distributions* (K.T. Fang and T.W. Anderson, eds.), 1–23, Allerton Press, New York.

Anderson, T.W. and Fang, K.T. (1990a). Inference in multivariate elliptically contoured distribution based on maximum likelihood. *Statistical Inference in Elliptically Contoured and Related Distributions* (K.T. Fang and T.W. Anderson, eds.), 201–216, Allerton Press, New York.

Barndorff-Nielsen, O. and Blaesild, P. (1981). Hyperbolic distributions and ramifications: contributions to theory and applications. *Statistical Distributions in Scientific Work,* (C. Taillie et al. eds.), 4, 19-44. D. Reidel Publishing Co., Dordrecht, Holland.

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

Bechhofer, R.E.; Dunnet, C.W. and Sobel, M. (1954). A two sample multiple decision procedure for ranking means of normal populations with a common unknown variance. *Biometrika,* 41, 170–176.

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

Berger, J. (1990). The Stein effect. *Encyclopedia of Statistical Sciences* (S. Kotz and N.L. Johnson, eds.), 757-761, Wiley, New York.

Berkane, M. and Bentler, P.M. (1986). Moment of elliptically distributed random variables. *Statist. and Probab. Lett.,* 4, 333–335.

Berkane, M. and Bentler, P.M. (1987). Characterizing parameters of multivariate elliptical distributions. *Commun. Statist. Simul. and Compu.*, **16**, 193–198.

Berkane, M. and Bentler, P.M. (1987a). Distribution of kurtoses with estimators and tests of homogeneity of kurtosis. *Statist. and Probab. Lett.*, **5**, 201–207.

Blattberg, R.C. and Gonedes, N.J. (1974). A comparison of the stable and Student distributions as statistical models for stock prices. *J. Business*, **47**, 224-280.

Box, G.E.P. (1952). Spherical distributions. (Abstract) *Ann. Math. Statist.*, **24**, 687–688.

Bravo, G. and MacGibbon, B. (1988). Improved shrinkage estimators for the mean vector of a scale mixture of normals with unknown variance. *Canad. J. Statist.*, **16**, 237-245.

Browne, M.W. and Shapiro, A. (1987). Adjustments for kurtosis in factor analysis with elliptically distributed errors. *J. Roy. Statist. Soc.*, B, **49**, 346–352.

Cacoullos, T. and Koutras, M. (1982). Quadratic forms in spherical random variables: generalized noncentral χ^2-distribution, *Naval. Res. Logist. Quart.*, **31**, 447-461.

Cambanis, S.; Huang, S. and Simons, G. (1981). On the theory of elliptically contoured distributions. *J. Mult. Anal.*, **11**, 368–385.

Chib, S.; Tiwari, R.C. and Jammalamadaka, S.R. (1988). Bayes prediction in regression with elliptical errors. *J. Economet.*, **38**, 349–360.

Chmielewski, M.A. (1980). Invariant scale matrix hypothesis tests under elliptical symmetry. *J. Mult. Anal.*, **10**, 343–350.

Chmielewski, M.A. (1981). Elliptically symmetric distributions. A review and bibliography. *Internat. Statist. Rev.*, **49**, 67–74.

Chmielewski, M.A. (1981a). Invariant tests for the equality of k- scale parameters under spherical symmetry. *J. Statist. Plann. and Infer.*, **5**, 341–346.

Cornish, E.A. (1954). The multivariate t-distribution associated with a set of normal sample deviates. *Austr. J. Physics*, **7**, 531–542.

Dey, D.K. and Srinivasan, C. (1985). Estimation of a covariance matrix under Stein's loss. *Ann. Statist.*, **13**, 1581–1591.

Dey, D.K. (1988). Simultaneous estimation of eigenvalues. *Ann. Inst.. Statist. Math.*, **40**, 137–147.

Dey, D.K. (1990). Estimation of scale parameters in mixture distributions. *Canad. J. Statist.*, **18**, 171–178.

Dunnet, C.W and Sobel, M. (1954). A bivariate generalization of Student's t. *Biometrika,* 41, 153–169.

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

Efron, B. and Morris, C. (1976). Multivariate empirical Bayes and estimation of covariance matrices. *Ann. Statist.* **4**, 22–32.

Fama, E.F. (1965). The behaviour of stock market prices. *J. Business,* **38**, 34–105.

Fan, J.Q. (1990). Generalized non-central t, F, and T^2-distributions. ***Statistical Inference in Elliptically Contoured and Related Distributions*** (K.T. Fang and T.W. Anderson, eds.), 79–95, Allerton Press, New York.

Fan, J.Q. (1990 a). Shrinkage estimators and ridge regression estimators for elliptically contoured distributions. ***Statistical Inference in Elliptically Contoured and Related Distributions*** (K.T. Fang and T.W. Anderson, eds.), 313–326, Allerton Press, New York.

Fan, J.Q. (1990 b). Distributions of quadratic forms and noncentral Cochran's theorem. ***Statistical Inference in Elliptically Contoured and Related Distributions*** (K.T. Fang and T.W. Anderson, eds.), 177–191, Allerton Press, New York.

Fan, J.Q. and Fang, K.T. (1990). Minimax estimators and Stein's two-stage estimators of location parameters. ***Statistical Inference in Elliptically Contoured and Related Distributions*** (K.T. Fang and T.W. Anderson, eds.), 299–312, Allerton Press, New York.

Fan, J.Q. and Fang, K.T. (1990a). Inadmissibility of sample mean and sample regression coefficients for elliptically contoured distributions. ***Statistical Inference in Elliptically Contoured and Related Distributions*** (K.T. Fang and T.W. Anderson, eds.), 275–290, Allerton Press, New York.

Fan, J.Q. and Fang, K.T. (1990b). Inadmissibility of the usual estimator for the location parameters of spherically symmetric distributions. ***Statistical Inference in Elliptically Contoured and Related Distributions*** (K.T. Fang and T.W. Anderson, eds.), 291–298, Allerton Press, New York.

Fang, K.T. and Anderson, T.W. (1990). Statistical Inference in Elliptically

Contoured and Related Distributions. Allerton Press, New York.

Fang, K.T.; Fan, J.Q. and Xu, J.L. (1990). The distribution of quadratic forms of random idempotent matrices with their applications. *Statistical Inference in Elliptically Contoured and Related Distributions* (K.T. Fang and T.W. Anderson, eds.), 163–175, Allerton Press, New York.

Fang, K.T.; Kotz, S. and Ng, K.W. (1990). Symmetric Multivariate and Related Distributions. Chapman and Hall, New York.

Fang, K.T. and Wu, Y. (1990). Distribution of quadratic forms and Cochran's theorem. *Statistical Inference in Elliptically contoured and Related Distributions* (K.T. Fang and T.W. Anderson, eds.), 147–161, Allerton Press, New York.

Fang, K.T. and Zhang, Y. (1990). Generalized Multivariate Analysis. Springer-Verlag,

Fisher, R.A., and Healy, M.J.R. (1956). New tables of Behren's test of significance. *J. Roy. Statist. Soc.*, Ser. B, **18**, 212.

Gang L. (1990). Moments of a random vector and its quadratic forms. *Statistical Inference in Elliptically Contoured and Related Distributions* (K.T. Fang and T.W. Anderson, eds.), 433–440, Allerton Press, New York.

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

Hsu, H. (1990). Generalized T^2 test for multivariate elliptically contoured distributions. *Statistical Inference in Elliptically Contoured and Related Distributions* (K.T. Fang and T.W. Anderson, eds.), 243–256, Allerton Press, New York.

Hsu, H. (1990a). Invariant tests for multivariate elliptically contoured distributions. *Statistical Inference in Elliptically Contoured and Related Distributions* (K.T. Fang and T.W. Anderson, eds.), 257–274, Allerton Press, New York.

Ifram, A.F. (1970). On the characteristic function of F and t distributions. *Sankhya,* Ser.A, **32**, 350-352.

Ismail, M.E. (1977). Integral representations and complete monotonocity of various quotients of Bessel functions. *Canad. J. Math.,* **29**, 1198-1207.

James, W. and Stein's, C. (1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.,* 1, 361-379. University of California Press, Berkeley.

Joarder, A.H. and Ali, M. M. (1992). Distribution of the correlation matrix for a class of elliptical models. *Commun. Statist. - Theor. Meth.*, **21**, 1953-1964

Joarder, A.H. and Ali, M. M. (1992). On some generalized Wishart expectations. *Commun. Statist. - Theor. Meth.*, **21**, 283–294.

Johnson, M.E. (1987). Multivariate Statistical Simulation. Wiley, New York.

Johnson, N.L. and Kotz, S. (1970). Distributions in Statistics : Continuous Univariate Distributions – 2. Wiley, New York.

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

Kariya T. (1981). A robustness property of Hotelling's T^2-test. *Ann. Statist.,* **9**, 210-214.

Kariya, T. (1981a). Robustness of Multivariate Tests. *Ann. Statist.*, 9, 1267–1275.

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

Kotz, S. (1975). Multivariate distributions at a cross road. *Statistical Distributions in Scientific Work*, 1, 247-270. D. Reidel Publishing Co., Dordrecht, Holland.

Lauritzen, S.L. (1984). Statistical Manifolds. Preprint No. R 84-12, Institute of Electronic Systems, Aalborg University Centre, Denmark.

Leung, P.L. (1992). Estimation of eigenvalues of the scale matrix of the multivariate F-distribution . *Commun. Statist.-Theory Meth.*, 21, 1845-1856.

Lebedev, N.N. (1965). Special Functions and their Applications. Prentice-Hall Inc., New Jersey.

Lord, R.D. (1954). The use of the Hankel transform in statistics 1. general theory and examples. *Biometrika*, 41, 44-55.

Mandelbrot, B. (1963). The variation of certain speculative prices. *J. Business*, 36, 394–419.

Mardia, K.V. (1970). Measures of multivariate skewness and kurtosis with applications.*Biometrika*, 57, 519-530.

Mitchell, A.F.S. (1988). Statistical manifolds of univariate elliptic distributions. *Internat. Satist. Rev.*, 56, 1-16.

Mitchell, A.F.S. (1989). The information matrix, skewness tensor and α connections for the general multivariate elliptical distribution. *Ann. Inst.. Statist. Math.*, 14, 289-304.

Muirhead, R.J. (1980). The effects of elliptical distributions on some standard procedures involving correlation coefficients: a review. *In Multivariate Statistical Analysis* (R.P. Gupta ed.), 143-159, North-Holland, Amsterdam.

Muirhead, R.J. and Waternaux, C.M. (1980). Asymptotic distributions in canonical correlation analysis and other multivariate procedures for non-normal population. *Biometrika,* 1, 31-43.

Muirhead, R.J. (1982). Aspects of Multivariate Statistical Theory . Wiley, New York, chapter 5.

Muirhead, R.J. (1986): A note on some Wishart expectations. *Metrika* 33, 247–251.

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

Pestana, D. (1977). Note on a paper of Ifram. *Sankhya*, Ser. A, 39, 396–397.

Press, S.J. (1982). Applied Multivariate Analysis. Robert E. Krieger Publishing Co., Florida.

Prudnikov, A.P.; Brychkov, Y.A. and Marichev, O.I. (1986). Integrals and Series. v.1, Gordon and Breach Science Publishers, New York.

Quan, H. (1990). Some optimal properties of testing hypotheses of elliptically contoured distribution. *Statistical Inference in Elliptically contoured and Related Distributions* (K.T. Fang and T.W. Anderson, eds.), 327–341, Allerton Press, New York.

Quan, H. and Fang K.T. (1990). Unbiasedness of the parameter tests for generalized multivariate distributions. *Statistical Inference in Elliptically con-*

toured and Related Distributions (K.T. Fang and T.W. Anderson, eds.), 343–359, Allerton Press, New York.

Rao, C.R. (1973). Linear Statistical Inference and Its Applications. Wiley, New York.

Schoenberg, I.J. (1938). Metric spaces and completely monotone functions. *Ann. Math.*, **39**, 811-841.

Sharma, D. and Krishnamoorthy, K. (1986). An identity concerning a Wishart matrix. *Metrika*, **33**, 65–68.

Spainer, J. and Oldham, K.B. (1987). An Atlas of Functions. Hemisphere Publishing Corporation, Washington D.C.

Singh, R.S. (1991). James-Stein rule estimators in linear regression models with multivariate t-distributed error. *Austr. J. Statist.*, **33**, 145-158.

Siotani, M. (1976). Conditional and stepwise multivariate t-distribution. *Essays in Probability and Statistics* , (S. Ideka ed.), 365–395. Shinko Tausho, Tokyo.

Srivastava, M.S. and Bilodeau, M. (1989). Stein estimation under elliptical distribution. *J. Mult. Anal.* **28**, 247-259.

Srivastava, M.S. and Khatri, C.G. (1979). An Introduction to Multivariate Statistics. North-Holland, New York.

Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proc. Third Berkeley Symp. Math. Statist. Prob.*, 1, 197–206 . University of California Press, Berkeley.

Stein, C. (1975). Reitz Lecture, 39th Annual Meeting IMS. Atlanta, Georgia.

Stuart, A. and Ord, J.K. (1987). Kendal's Advanced Theory of Statistics. vol.1, Charles Griffin and Co. Ltd., London.

Sutradhar, B.C. (1986). On the characteristic function of multivariate student t-distribution. *Canad. J. Statist.*, **14**, 329–337.

Sutradhar, B.C. and Ali, M.M. (1986). Estimation of the parameters of a regression model with a multivariate t error variable. *Commun. Statist. Theor. Meth.*, **15**, 429–450.

Sutradhar, B.C.(1988). Testing linear hypothesis with t error variable. *Sankhya,* B, **50**, 175-180.

Sutradhar, B.C. and Ali, M.M. (1989). A generalization of the Wishart distribution for the elliptical model and its moments for the multivariate t model. *J. Mult. Anal.,* **29**, 155–162.

Sutradhar, B.C. (1990). Discrimination of observations into one of two t populations. *Biometrics,* **46**, 827–835.

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

Tranter, C.J. (1968). Bessel Functions with some Physical Applications. English Universities Press, Ltd., London.

Watson, G.S. (1984). Statistics on Spheres. Wiley-Interscience, New York.

Watson, G.N. (1958). A treatise on the theory of Bessel Functions. Cambridge University Press, London.

Zellner, A. (1976). Bayesian and non-Bayesian analysis of the regression model with multivariate Student-*t* error term. *J. Amer. Statist. Assoc.*, **71**, 400-405 (correction, **71**, 1000).