King Fahd University of Petroleum & Minerals

DEPARTMENT OF MATHEMATICAL SCIENCES

Technical Report Series

TR 359

Oct 2006

On the Multivariate *T*-Distribution and Some of its Applications

B. M. Golam Kibria and Anwar H. Joarder

DHAHRAN 31261 ● SAUDI ARABIA ● www.kfupm.edu.sa/math/ ● E-mail: mathdept@kfupm.edu.sa

On the Multivariate *T*-Distribution and Some of its Applications

B. M. Golam Kibria
Department of Statistics
Florida International University
University Park
Miami, FL 33199, USA
Email: kibriag@fiu.edu

and

Anwar H. Joarder
Department of Mathematical Sciences
King Fahd University of Petroleum & Minerals
Dhahran 31261, Saudi Arabia
Email: anwarj@kfupm.edu.sa

Summary

This paper makes an attempt to justify a multivariate t-model and provides a modest review of most important results of this model developed in recent years. Essential properties and applications of the model in various fields are discussed. Special attention is given to pre-test and shrinkage estimation for regression parameters under certain restrictions. The predictive distributions under the multivariate t-distribution are also discussed. It is observed that the multivariate t-distribition is more realistic to model multivariate data than multivariate normal distribution because of its fat tail.

AMS 1991 Subject Classification: Primary 62J02, Secondary 62H12 and 62J07

Keywords and Phrases: Multivariate t-model; scale mixture of multivariate normal distribution; robustness of correlation; estimation of parameters; pretest and shrinkage estimation; ridge regression; predictive distribution; tests of hypotheses.

1. Introduction

The classical theory of statistical analysis is primarily based on the assumption that errors are normally distributed. Recently many authors have investigated as to how inferences are affected if the population model departs from normality. Most of economic and business data e.g. stock return data exhibit fat tailed distributions. The suitability of independent t-distributions for stock return data was assessed by Blattberg and Gonedes (1974). Soon after that Zellner (1976) considered analyzing stock return data by a simple regression model under the assumption that errors have a multivariate t-distribution. However, errors in this model are uncorrelated but not independent. Prucha and Kelejian (1984) discussed the inadequacy of normal distribution and suggested an uncorrelated t-model for many real world problems as a better alternative distribution. After a thorough investigation, Kelejian and Prucha (1985) proved that uncorrelated t-distributions are better able to capture heavy-tailed behavior than independent t-distributions.

The multivariate t-distribution is a viable alternative to the usual multivariate normal distribution and on the other hand results obtained under normality can be checked for robustness. For example, the distribution of product moment correlation coefficient obtained by Ali and Joarder (1991) is the same as that obtained by Fisher (1915) showing distribution robustness. Thus the t-test for testing significance of correlation is also robust (Joarder 2006). For more explanation of the background of this area of research interested readers may go through Cornish (1954), Kelker (1970), Cambanis, Huang and Hsu (1981), Fang and Anderson (1990), Kotz and Nadarajah (2004) , Nadarajah and Kotz (2006) and the references therein. In this paper we justify an uncorrelated multivariate t-model as the model for sample and present a modest review of the most important theories developed recently for statistical analysis with this model. This paper is expected to attract young researchers to develop an organized and solid foundation for the statistical analysis with an uncorrelated t-model.

2. The Multivariate *T*-Distribution

Different forms of multivariate t-distribution exist in literature. We will discuss some of them in this section. The probability density function (p.d.f.) of a p-variate t-distribution is given by

$$f(x) = \frac{|\Sigma|^{-1/2}}{C(\nu, p)\pi^{p/2}}\left[1+\frac{1}{\nu}(x-\mu)'\Sigma^{-1}(x-\mu)\right]^{-(\nu+p)/2}, \tag{1}$$

where x is the realized value of a $p\times 1$ random vector X, μ is a $p\times 1$ unknown mean vector and Σ is a $p\times p$ positive definite matrix of scale parameters while the normalizing constant $C(\nu, p)$ is given by

$$\Gamma((\nu+p)/2)C(\nu, p) = \nu^{p/2}\,\Gamma(\nu/2). \tag{2}$$

The p-variate random variable X has a mean vector μ and a covariance matrix $\Sigma^* = \nu^*\Sigma$, where $\nu^* = \nu/(\nu-2)$ and can well be represented by $T_p(\mu, \nu^*\Sigma)$ where the shape parameter $\nu(>2)$ is assumed to be known. If $p=1$, $\mu=0, \Sigma=1$, then the p.d.f. in (1) defines the univariate t-distribution. When $p=2$, $\mu=0, \Sigma=I_2$, then the p.d.f. in (1) is a slight modification of the bivariate surface of Pearson distribution (Pearson, 1923). It is well-known that the multivariate t-distribution can be written as

$$f(x) = \int_0^\infty \frac{|\omega^2\Sigma|^{-1/2}}{(2\pi)^{p/2}}\exp(-(x-\mu)'(\omega^2\Sigma)^{-1}(x-\mu)/2)\,h(\omega)d\omega \tag{3}$$

which is the mixture of the multivariate normal distribution $N_p(\mu, \omega^2\Sigma)$ and ω has the inverted gamma distribution with p.d.f.

$$h(\omega) = \frac{2(\nu/2)^{\nu/2}}{\Gamma(\nu/2)}\omega^{-(\nu+1)}\exp\left(\frac{-\nu}{2\omega^2}\right), \tag{4}$$

where ν is the degrees of freedom of inverted gamma distribution. Equivalently, $\nu\omega^{-2}$ has a chi-square distribution with ν degrees of freedom. Thus for given ω, the random vector X has a multivariate normal distribution i.e.

$$(X \mid \Omega = \omega) \sim N_p(\mu, \omega^2\Sigma). \tag{5}$$

As $\nu \to \infty$, the random variable Ω becomes a degenerate random variable with all the non-zero mass at the point unity and, consequently, the pdf of the multivariate t-distribution in (1) converges to that of the multivariate normal distribution $N_p(\mu, \Sigma)$. This also follows from the fact that as $\nu \to \infty$, we have $C(\nu, p) \to 2^{p/2}$ and $(1+u/\nu)^{-\nu} \to e^{-u}$. It is also worth mentioning that the uncorrelatedness of the components $X_1, X_2, \ldots, X_N$ does not imply their independence unless $\nu \to \infty$. In the following section, we will discuss some properties of the multivariate t-distribution.

3. Some Properties of the Multivariate *T*- Distribution

3.1 Moments and Characteristic Function

By the use of the mixture representation in (5), it can be easily proved that

$$E(X) = E(E(X \mid \omega)) = \mu$$

and

$$Cov\,(X\,) = E\,[Cov\,(X \mid \omega)] + Cov\,[E\,(X \mid \omega)] = E\,(\Omega^2\Sigma) = \nu^*\Sigma,$$

where $\nu^* = \nu/(\nu-2)$. The characteristic functions of the univariate and the multivariate t-distributions have been considered by many authors. The characteristic function of X following a multivariate t-distribution with p.d.f in (1) is given by

$$\phi_X(t) = E(e^{it'X}) = e^{it'\mu}\ \frac{\|(\nu\Sigma)^{1/2}t\|^{\nu/2}}{2^{\nu/2-1}\,\Gamma(\nu/2)}\,K_{\nu/2}(\|(\nu\Sigma)^{1/2}t\|) \tag{6}$$

(Joarder and Ali, 1996), where $K_{\nu/2}(\|(\nu\Sigma)^{1/2}t\|)$ is the Macdonald function with order $\nu/2$ and argument $\|(\nu\Sigma)^{1/2}t\|$.

The Macdonald function $K_\alpha(t)$ with order α and argument t admits by the following integral representation (see e.g. Watson, 1958, p. 172):

$$K_\alpha(t) = \left(\frac{2}{t}\right)^\alpha \frac{\Gamma(\alpha+1/2)}{\sqrt{\pi}}\int_0^\infty (1+u^2)^{-(\alpha+1)}\cos tu\; du, t>0, \alpha > -1/2. \tag{7}$$

A series representation of the Macdonald function $K_\alpha(r)$ where $r>0$ and α a nonnegative integer is well known (cf. Joarder and Ali, 1996). The quantity $Z = \Sigma^{-1/2}(X-\mu)$ has a spherical t -distribution $T_p(0, \nu^* I)$ whose product moment is given by

$$E\left(\prod_{i=1}^p Z_i^{k_i}\right) = \begin{cases} 0 \text{ if at least one } k_i\ (i=1,2,\ldots,p) \text{ is odd,} \\ \nu^{k/2}\ \dfrac{\Gamma((\nu-k)/2)}{2^k\ \Gamma(\nu/2)}\displaystyle\prod_{i=1}^p \frac{k_i!}{(k_i/2)!},\ \nu > k \\ \quad \text{if all } k_i\text{'s } (i=1,2,\ldots,p) \text{ are even} \end{cases} \tag{8}$$

where $k = \sum_{i=1}^p k_i$. The product moment can also be derived by using the stochastic representation $Z = RU$, where R^2/p has an $F(p,\nu)$ distribution, $R = (Z'Z)^{1/2}$ and U has a uniform distribution on the surface of unit hypersphere in $\Re^p$. Then

$$E\left(R^k\right) = \nu^{k/2}\,\frac{\Gamma((p+k)/2)\Gamma((\nu-k)/2)}{\Gamma(p/2)\Gamma(\nu/2)}, \nu > k \tag{9}$$

(See Theorem 2.8 of Fang, Kotz and Ng (1990) for details).

It follows from (6) that the characteristic function of the univariate Student t-distribution with p.d.f.

$$f(x) = \frac{1}{C(\nu,1)\sqrt{\pi}}\left(1 + x^2/\nu\right)^{-(\nu+1)/2}, \nu > 0$$

is given by

$$\phi_X(t) = \frac{\nu^{\nu/4}\,|t|^{\nu/2}}{2^{\nu/2-1}\Gamma(\nu-2)} K_{\nu/2}(\sqrt{\nu}\,|t|), \quad \nu > 2,$$

where $C(\nu, p)$ is defined in (2) and $K_{\alpha}(t)$ is defined in (7). Based on this characteristic function Rahman and Saleh (1975) derived the exact distribution of the Behrens-Fisher statistic.

It may be remarked that the characteristic function of X in (6) can also be written as

$$\phi_X(t) = E(e^{it'X}) = e^{it'\mu}\ \psi(t'\Sigma t). \tag{10}$$

The covariance matrix and the kurtosis parameter can then be written as

$$Cov(X) = -2\psi'(0)\Sigma \text{ and } \kappa = \frac{\psi''(0)}{\{\psi'(0)\}^2} - 1 \tag{11}$$

respectively (Seo and Toyama, 1996).

3.2 Marginal and Conditional Distributions

It is well-known that marginal and conditional distributions of the components of X follow the multivariate t-distribution (see e.g. Sutradhar, 1984). Let X, μ, t and Σ be partitioned as

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \ t = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} \ \textit{and} \ \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

where $X_2, \mu_2, t_2\ \varepsilon\ \Re^q$ $(q < p)$ and Σ_2 is a $q \times q$ positive definite matrix. By the use of the characteristic function of X given by (6), it may be easily checked that $X_2 \sim T_q\ (\mu_2, \nu^*\Sigma_{22})$ where $\nu^* = \nu/(\nu-2)$. The conditional distribution of X_1 given $X_2 = x_2$ is $T_{p-q}\left(\mu_{1\cdot2}, \nu_{1\cdot2}\ \Sigma^*_{11.2}\right)$ where

$$\begin{aligned}
\mu_{1.2} &= \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2), \\
\nu_{1.2} &= \nu/(\nu + q - 2), \ \textit{and} \\
\Sigma^*_{11.2} &= \left(1 + (x_2-\mu_2)'(\nu\Sigma_{22})^{-1}(x_2-\mu_2)\right)\Sigma_{11.2}
\end{aligned}$$

with $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. The derivation of conditional covariance matrix discussed among others by Cambanis et. al. (1981) is detailed in the next section.

3.3 **Determination of Covariance Matrix by Stochastic Representation**

A p-dimensional random variable Z is said to have a spherical distribution if its probability density function (pdf) can be written as

$$f(z) = g(z'z) \tag{12}$$

Muirhead (1982) is the first to discuss spherical and elliptical distributions in a text book of multivariate analysis. Much of the theoretical development are available in Fang and Anderson (1990) and Fang, Kotz and Ng (1990). For applications of such distributions we refer to Lange, Little and Taylor (1989), Kotz and Nadarajah (2004) and the references therein. Let Z have the multivariate t-distribution with p.d.f.

$$f(z) = g(z'z) = \frac{1}{C(\nu,p)\pi^{p/2}}\left(1+z'z/\nu\right)^{-(\nu+p)/2} \tag{13}$$

where $C(\nu,p)$ is given by $C(\nu,p)\Gamma((\nu+p)/2) = \nu^{p/2}\Gamma(\nu/2)$. Then

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

and that

$$E\left(R^k\right) = \nu^{k/2}\,\frac{\Gamma((p+k)/2)\Gamma((\nu-k)/2)}{\Gamma(p/2)\Gamma(\nu/2)},\ \ \nu > k \tag{14}$$

(cf. Fang, Kotz and Ng, 1990, 22). In particular,

$$E\left(R^2\right) = \frac{\nu p}{\nu-2},\ \ \nu > 2 \text{ and } V\left(R^2\right) = \frac{2p(p+\nu-2)\nu^2}{(\nu-2)(\nu-4)^2},\ \ \nu > 4. \tag{15}$$

Consider the elliptical random variable $X = \mu + \Sigma^{1/2}Z$, where Z has the p.d.f. given by (12). It is well known (Cambanis, Hunag and Simons, 1981) that the covariance matrix of X is given by $Cov(X) = -2\psi'_X(0)\Sigma$, where $\phi_X(t) = E(e^{it'X}) = e^{it'\mu}\,\psi(t'\Sigma t)$ is the characteristic function of *X*.

Since most elliptical distributions do not have closed form for characteristic functions, an easy way out to determine covariance matrix is to exploit stochastic decomposition $\Sigma^{-1/2}(X-\mu) = Z = RU$ where $R = (Z'Z)^{1/2}$ is independent of U and the random variable U is uniformly distributed on the surface of unit sphere is R^p.

For any elliptical random variable X , where $X = \mu + \Sigma^{1/2}Z$ with Z having the p.d.f. (12), it is well known that $Cov(X) = p^{-1}E(R^2)\Sigma$ (Cambanis, Huang and Hsu, 1981, or Joarder, 1992). It follows from (15) that

$$Cov(X) = \frac{1}{p}\left(\frac{\nu p}{\nu-2}\right)\Sigma = \frac{\nu}{\nu-2}\Sigma,$$

where $X = \mu + \Sigma^{1/2}Z$ with Z having the p.d.f. given by (13).

3.4 Distribution of a Linear Function

Suppose the random variable X has a multivariate t-distribution with degrees of freedom ν, mean vector μ and covariance matrix Σ. Assume that A is a non-singular matrix and b is a constant vector, then $AX+b$ has the p-variate t distribution with mean vector $A\mu+b$, degrees of freedom ν and covariance matrix ARA'. The degrees of freedom for the distribution of the linear combination remain same. This result is similar to that for the multivariate normal distribution.

3.5 Distribution of Quadratic Forms

Suppose the random variable $\boldsymbol{X}$ has a multivariate t-distribution with mean vector μ, covariance matrix Σ and degrees of freedom ν, then $X'\Sigma^{-1}X/p$ has the F distribution with ν and p degrees of freedoms and non-centrality parameter $\mu'\Sigma^{-1}\mu/p$. For details we refer Mathai and Provost (1992).

4 Uncorrelated *T*- Model for Sample

The joint p.d.f. of N independent observations each having a p-variate t-distribution is given by

$$f_1(x_1, x_2, \ldots, x_N) = f(x_1) f(x_2) \ldots f(x_N) \tag{16}$$

which may be referred to as the *independent t-model*. However, recent interest is noticed in uncorrelated t-distributions. The joint p.d.f. of N uncorrelated random variables each having a multivariate t-distribution is given by

$$f_2(x_1, x_2, \ldots, x_N) = \frac{|\Sigma|^{-N/2}}{C(\nu, Np)\pi^{Np/2}} \left(1 + Q/\nu\right)^{-(\nu+Np)/2} \tag{17}$$

where $Q = \sum_{j=1}^{N} (x_j - \mu)' \Sigma^{-1} (x_j - \mu)$ and x_j $(j = 1,2,\ldots,N)$ is the realized value of a p-component random vector X_j $(j = 1,2,\ldots,N)$ having the t-distribution $T_p(\mu, \nu^* \Sigma)$ where $\nu^* = \nu/(\nu-2)$. The p.d.f. in (17) will hereinafter be called the *uncorrelated t-model*.

Kelejian and Prucha (1985) proved that the tails of the uncorrelated t -model is relatively thicker than those of the independent t-model given by (16). It may be remarked that observations in (17) are independent if and only $\nu \to \infty$, in which case the p.d.f. in (17) will be the product of that of N independent p-dimensional random variables each having normal distribution $N_p\ (\mu, \Sigma)$.

A more general case would be to consider k-samples, say, $(X_{g_1}, X_{g_2}, \ldots, X_{gN_g}), g = 1,2,\ldots,k$ is a sample of size N_g from $T_p(\mu_g, \nu^* \Sigma)$, $g = 1,2,\ldots,k$. The joint p.d.f. of observations of k-samples would be

$$f_3(x_{11}, x_{12}, \ldots, x_{kN_k}) = \frac{|\Sigma|^{-N/2}}{C(\nu, Np)\pi^{Np/2}} \left(1 + \frac{1}{\nu} Q\right)^{-(\nu+Np)/2} \tag{18}$$

where $Q = Q_1 + Q_2 + \cdots + Q_k = \sum_{g=1}^{k} Q_g$, $Q_g = \sum_{j=1}^{N_g} (x_{g_j} - \mu_g)' \Sigma^{-1} (x_{g_j} - \mu_g)$ and

$N = N_1 + N_2 + \cdots + N_k = \sum_{g=1}^{k} N_g$.

4.1 Distribution of the Sum of Products Matrix Based on the Uncorrelated T-Model

The sum of product matrix based on the uncorrelated t- model (17) is given by

$$A = \sum_{j=1}^{N} \left(X_j - \overline{X}\right)\left(X_j - \overline{X}\right) = (a_{ik}),$$

where $\overline{X} = \sum_{j=1}^{N} X_j / N$. It follows from (5) that for a given ω, the random matrix A has the usual Wishart distribution

$$(A \mid \Omega = \omega) \sim W_p(m, \omega^2 \Sigma),\ m = N - 1 \tag{19}$$

i.e. the p.d.f. of A is given by

$$\int_0^\infty \frac{|\omega^2 \Sigma|^{-1/2}}{2^{mp/2} \Gamma_p(m/2)} |A|^{(m-p-1)/2} \exp\left(-\frac{1}{2} tr(\omega^2 \Sigma)^{-1} A\right) h(\omega) d\omega, \tag{20}$$

where $A > 0$, $m = N - 1 \geq p$ and the generalized gamma function $\Gamma_p(\alpha)$ is defined by

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

with $\nu\omega^{-2} \sim \chi_\nu^2$. The completion of integration in (20) results in the p.d.f. of A given by

$$\frac{|\Sigma|^{-m/2}}{C(\nu, mp) 2^{mp/2} \Gamma_p(m/2)} |A|^{(m-p-1)/2} \left(1 + \frac{1}{\nu} tr \Sigma^{-1} A\right)^{-(\nu+mp)/2} \tag{21}$$

(cf. Sutradhar and Ali, 1989).

By the use of the mixture representation in (19), it is easy to derive the expected values of $|A|^k, |A|^k A, |A|^k A^{-1}, (trA)^2, tr(A^2)$ etc. which are important in developing estimation strategies for functions based on the covariance matrix. See e.g. Joarder and Ali (1992a) and Joarder (1995a, 1995b).

4.2 Robustness of Correlation for Uncorrelated T -Model

Fisher (1915) derived the exact sampling distribution of Pearsonian correlation coefficient R for a random sample drawn from a bivariate normal population $N_2(\mu, \Sigma)$. Since then many statisticians have tried to investigate the behavior of R for non-normal situations. Ali and Joarder (1991) proved that both null and non-null distribution of R remain robust in a class of elliptical distributions which accommodates the uncorrelated t-model as a special case. The result has been generalized for the multivariate uncorrelated elliptical model by Joarder and Ali (1992b) for the correlation matrix R . The p.d.f. of R is given by

$$f(R) = \frac{|R|^{(m-p-1)/2}}{\left(|\rho| \prod_{i=1}^{p} \rho^{ii}\right)^{m/2} \Gamma_p(m/2)} H_{m,p}(\Gamma)$$

where

$$H_{m,p}(\Gamma) = 2^{-(m-2)p/2}\int_0^\infty \cdots \int_0^\infty (v_1 v_2 \ldots v_p)^{m-1} e^{-v'\Gamma v/2} \prod_{i=1}^{p} dv_i$$

with $v = (v_1, v_2, \ldots v_p)'$ and $\Gamma = (\rho_{ik}^*, r_{ik})$, $\rho_{ik}^* = (\rho^{ii}\rho^{kk})^{-1/2}\rho^{ik}$, ρ^{ik} denoting (i,k) th element of ρ^{-1} th element for all $i,k = 1,2,\ldots,p$. Note that Joarder and Ali (1992b) reported e^{-u} instead of $e^{-u/2}$ in the above integral. For more on the robustness of Pearsonian correlation coefficient, see Kotz, Balakrishnan and Johnson (2000) and Joarder (2006) among others.

5. Estimation of Parameters

5.1 Estimation of Parameters for One population

The maximum likelihood estimators of the parameters μ and Σ of the uncorrelated t-model in (17) are given by $\mu = \overline{X}$ and $\Sigma = A/N$ respectively (see Fang and Anderson, 1990, pp. 201--216). But maximum likelihood estimators in this case are not appealing because most important properties of maximum likelihood estimators, follow from the independence of the observations which is not the case for the model in (17) for finite value of the shape parameter ν. The sample mean $\overline{X}$ is obviously an unbiased and consistent estimator of μ. The unbiased estimator of Σ is given by $\Sigma = A/(\nu^* m)$, where $\nu^* = \nu/(\nu-2)$ and $m = N-1$ (see Fang and Anderson 1990, pp. 208).

Joarder (1995a) considered the estimation of the scale matrix Σ of the uncorrelated t-model under a squared error loss function. It may be remarked that the scale matrix Σ determines the covariance matrix up to a known constant ν^*. Joarder and Ahmed (1996) developed estimation strategy for eigenvalues of Σ of the uncorrelated t-model given by (17). The estimation of the trace of the scale matrix Σ under a squared error loss was considered by Joarder and Beg (1999). The estimation of Σ under a entropy loss function was considered by Joarder and Ali (1997).

5.2 Estimation of Parameters for Two Populations

Consider a two-sample problem i.e. the case of $k = 2$ in the situation discussed in (18). The equality of mean vectors μ_1 and μ_2 can then be tested by

$$T^2 = \left(\overline{X}_1 - \overline{X}_2\right)\left(\frac{S_p}{N_1} + \frac{S_p}{N_2}\right)^{-1}\left(\overline{X}_1 - \overline{X}_2\right)$$

where $(m_1 + m_2)S_p = m_1 S_1 + m_2 S_2$ with $m_1 = N_1 - 1$ and $m_2 = N_2 - 1$. The above result was derived by Sutradhar (1988a) for a scaled uncorrelated t-model obtained by reparametrizing $\nu^*\Sigma$ by Σ in (7). The following derivation of T^2-statistic is based on the mixture representation of multivariate uncorrelated t-model (see e.g. Khan 1997).

By virtue of the mixture representation of (5), it follows that conditional on ω,

$$\frac{m}{p}T^2 \sim F_{p,m}(\delta_\omega), \tag{22}$$

where $F_{p,m}(\delta_\omega)$ denotes a noncentral F -distribution with parameters $p, m = m_1 + m_2 - p + 1$ and

$\delta_\omega = (\mu_1 - \mu_2)'(\omega^2\Sigma)^{-1}(\mu_1 - \mu_2)$. The unconditional distribution of $\frac{m}{p}T^2$ can be obtained by completing the following integral

$$\int_0^\infty u_{p,m}(\delta_\omega)\, h(\omega) d\omega,$$

where $u_{p,m}(\delta_\omega)$ is the p.d.f. of $F_{p,m}(\delta_\omega)$. It follows from (22) that under $H_0 : \mu_1 = \mu_2$

$$T^2 \sim \frac{p}{m} F_{p,m}.$$

The power function of the test $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 \neq \mu_2$ was discussed by Sutradhar (1988a) and Sutradhar (1990).

Khan (1997) considered the estimation of the mean vector of the multivariate t-distribution in the presence of uncertain prior information. The usual MLE, restricted estimator and preliminary test estimators were considered; he compared their performances under the unbiasedness and minimum risk criterion. Several recommendations were made based on the condition on the departure parameter Δ. Khan (2004) also investigated the effect of shape parameters for the shrinkage estimators of the mean vector of multivariate t-distribution. Some properties of shrinkage and the positive-rule shrinkage estimators were discussed by changing the value of the shape parameter. He also studied the relative performance of these estimators under different conditions.

6. Linear Regression Models

Zellner (1976) considered univariate linear regression model to analyze stock return data with errors having a univariate uncorrelated t-model. It is King (1980) who laid the rigorous mathematical foundation of linear regression analysis under broader distributional assumptions of spherical symmetry which includes uncorrelated t-model as a special case. Prompted by the works of Zellner (1976) and King (1980), many authors used uncorrelated t-model for modeling real world data. Sutradhar and Ali (1986) generalized Zellner's model with errors having uncorrelated t-model given by (17). Lange, Little and Taylor (1989) applied uncorrelated t-model to a variety of situations.

The null distribution of the usual F-statistic in a linear regression model under uncorrelated t-model in (17) is robust but the power function depends on the form of (17); see e.g. Sutradhar (1988a) and Sutradhar (1990) for a detailed proof. For the linear regression model with errors having an uncorrelated t-model, it is known (Singh, 1987) that the usual least square estimator of the vector of regression coefficients is not only the maximum likelihood estimator but also the unique minimum variance estimator. Singh (1988) also developed methods of estimation of error variance in linear regression models with errors having an uncorrelated t-model with unknown degrees of freedom.

6.1 Pretest and Shrinkage Estimation under Multivariate t-Error

Consider the following linear regression model

$$y = X\beta + e, \tag{23}$$

where $y = (y_1, y_2, \ldots, y_N)'$ is an $N \times 1$ vector of observations on the dependent variable, X is an

$N \times p$ matrix of full rank p, $\beta = (\beta_1, \beta_2, \ldots, \beta_p)'$ is an $p \times 1$ vector of parameters and $e = (e_1, e_2, \ldots, e_N)'$ is an $N \times 1$ vector of errors, which are distributed according to the laws belonging to the class of spherical compound normal distributions (or equivalently scale mixture of normal distributions) with $E(e) = 0$ and $E(ee') = \sigma_e^2 I_N$, where I_N is the N-dimensional identity matrix and σ_e^2 the common variance of e_j ($j = 1, 2, \ldots, N$). This class is a subclass of the family of spherically symmetric distributions (SSD) and can be expressed as

$$f(e) = \int_0^{\infty} f(e \mid \omega) g(\omega) d\omega, \tag{24}$$

where $f(e)$ is the p.d.f. of e, $f(e \mid \omega)$ is the p.d.f. of normal with mean vector 0 and variance-covariance matrix $\omega^2 I_N$ and $g(\omega)$ is the p.d.f. of ω with support $[0, \infty)$. In this case, $E(\Omega^2) = \sigma_e^2$. The well-known members of the pdf in (24) are spherical normal distribution, spherical t-distribution and spherical Cauchy distribution. Interested readers may go through Joarder (2006) for an introduction to the bivariate t-distribution.

In most applied as well as theoretical research works, the error terms in linear models are assumed to be normally and independently distributed. However, such assumptions may not be appropriate in many practical situation (for example, see Gnanadesikan, 1977 and Zellner 1976). It happens particularly if the error distribution has heavier tails. One can tackle such situation by using the well known t-distribution as it has heavier tail than the normal distribution, specially for smaller degrees of freedom (e.g. Fama (1965), Blattberg and Gonedes (1974)). Because of the above reasons we are motivated to use the multivariate t-distribution as the error distribution. The multivariate Student's t-distribution can be obtained if $g(\omega)$ is assumed to have an inverted gamma density $IG(\nu, \sigma^2)$ with scale parameter σ^2, degrees of freedom ν and is given by

$$g(\omega) = \frac{2}{\Gamma(\nu/2)} \left(\frac{\nu\sigma^2}{2} \right)^{\nu/2} \omega^{-(\nu+1)} e^{-\frac{\nu\sigma^2}{2\omega^2}}, \quad 0 < \nu, \sigma, \omega < \infty. \tag{25}$$

Then, the multivariate t-distribution obtained from (24) is given by,

$$f(e) = \frac{\Gamma\left((N+\nu)/2\right)}{(\pi\nu)^{N/2}\Gamma(\nu/2)\sigma^N} \left(1 + \frac{e'e}{\nu\sigma^2}\right)^{-(N+\nu)/2}, \quad 0 < \nu, \sigma, < \infty, \quad -\infty < e_j < \infty. \tag{26}$$

The mean vector and variance-covariance matrix of e are respectively,

$$E(e) = 0, \quad and \quad E(ee') = \frac{\nu}{\nu - 2} \sigma^2 I_N = \sigma_e^2 I_N, \qquad \nu > 2.$$

The marginal distributions are univariate student t-distributions. For $\nu = 1$, the p.d.f. in (26) becomes that of Cauchy and as $\nu \to \infty$, the *pdf* approaches normal. For the full model the Unrestricted Estimator (UE) of β is given by

$$\hat{\beta}^{UE} = C^{-1} X'y, \tag{27}$$

where $C = X'X$ is the information matrix. The corresponding unbiased estimator of σ_e^2 is given by

$$\tilde{\sigma}_e^2 = \frac{(y - X\hat{\beta}^{UE})'(y - X\hat{\beta}^{UE})}{N - p}.$$

Our primary interest is to estimate the regression coefficients β when it is *apriori* suspected but not certain that β may be restricted to the subspace,

$$H_0 : H\beta \quad = \quad h, \tag{28}$$

where H is an $q \times p$ known matrix of full rank $q(< p)$ and h is an $q \times 1$ vector of known constants. The Restricted Estimator (RE) of β is given by

$$\hat{\beta}^{RE} = \hat{\beta}^{UE} - C^{-1}H'(HC^{-1}H')^{-1}(H\hat{\beta}^{UE} - h) \quad and \tag{29}$$

and the corresponding estimator of σ_e^2 is given by

$$\hat{\sigma}_e^2 = \frac{(y - X\hat{\beta}^{RE})'(y - X\hat{\beta}^{RE})}{N - p + q};$$

which is unbiased under the null hypothesis. Note that the restricted least squares estimator satisfies the condition $H\hat{\beta} = h$. The estimator of β in (27) is usually used in the case when there is no hypothesis information available on the vector of parameter of interest β. On the other hand, the estimator of β in (29) is useful in the presence of hypothesis (28). Furthermore, it is well known that the RE performs better than the UE, when the restrictions hold but as the parameters β moves away from the subspace $H\beta = h$, the RE becomes biased and inefficient while the performance of the UE remains stable. As a result, one may combine the UE and RE to obtain a better performance of the estimators in presence of the Uncertain Prior Information (UPI) $H\beta = h$, which leads to the Preliminary Test (PT) *estimator* and is defined as

$$\hat{\beta}^{PT} = \hat{\beta}^{RE} I(L_N \le L_{N,\alpha}) + \hat{\beta}^{UE} I(L_N > L_{N,\alpha}), \tag{30}$$

where,

$$L_N = \frac{(H\hat{\beta}^{UE} - h)'(HC^{-1}H')^{-1})(H\hat{\beta}^{UE} - h)}{q\tilde{\sigma}_e^{\,2}}$$

is the test-statistic for testing the null-hypothesis in (28), and $L_{N,\alpha}$ is the upper α-level critical value of L_N and $I(A)$ is the indicator function of the set A. Under the null hypothesis and normal theory, L_N follows a central F-distribution with $(q, N - p)$ degrees of freedom while under the alternative it follows the non-central F-distribution with $(q, N - p)$ degrees of freedom and non-centrality parameter $\Delta/2$, where

$$\Delta = \frac{(H\beta - h)'(HC^{-1}H')^{-1}(H\beta - h)}{\sigma_e^2} \tag{31}$$

is the departure parameter from the null hypothesis. It is important to remark that $\hat{\beta}^{PT}$ is bounded and performs better than $\hat{\beta}^{RE}$ in some part of the parameter space. For details see Judge and Bock (1978), Han and Bancroft (1968), Saleh and Sen (1978), Kibria and Saleh (1993) and Saleh (2006) among others. Note that, the Preliminary Test *(PT)* estimator has two characteristics: (i) it produces only two values, the unrestricted estimator and the restricted estimator, (ii) it depends heavily on the level of significance of the Preliminary Test *(PT)*. What about the intermediate value between $\hat{\beta}^{UE}$ and $\hat{\beta}^{RE}$? To overcome this shortcoming, we consider the Stein-type estimator. The Stein-type Shrinkage estimator *(SE)* of β is defined as

$$\hat{\beta}^{SE} = \hat{\beta}^{UE} - dL_N^{-1}(\hat{\beta}^{UE} - \hat{\beta}^{RE}), \tag{32}$$

where

$$d = \frac{(q-2)(N-p)}{q(N-p+2)}, \quad \text{and } q \geq 3.$$

The *SE* in (32) exhibits uniform improvement over $\hat{\beta}^{UE}$, however it is not a convex combination of $\hat{\beta}^{UE}$ and $\hat{\beta}^{RE}$. Both (30) and (32) involve the statistic L_N which adjusts the estimator for departure from H_0. For large value of L_n both (30) and (32) yield $\hat{\beta}^{UE}$, while for small value of L_n their performance is different. The SE has the disadvantage that the shrinkage factor $(1 - dL_N^{-1})$ becomes negative for $L_N < d$. This encourages one to find an alternative estimator. Hence, we define a better estimator, namely, the positive-rule shrinkage estimator (PR) of β as follows:

$$\hat{\beta}^{PR} = \hat{\beta}^{SE} - (1 - dL_N^{-1})I(L_N \leq d)(\hat{\beta}^{UE} - \hat{\beta}^{RE}). \tag{33}$$

The PR estimator in (33) provides uniform improvement over $\hat{\beta}^{UE}$ and it is a convex combination of $\hat{\beta}^{UE}$ and $\hat{\beta}^{RE}$. The properties of stein-type estimators have been analyzed under normality assumption by various researchers.

Tabatabaey et al. (2004a) considered aforementioned five well known estimators, namely, unrestricted estimator (*UE*), restricted estimator (*RE*), preliminary test (PT) estimator, shrinkage estimator (*SE*) and positive rule (*PR*) shrinkage estimator under the multivariate t-error assumption. The bias and risk functions of the proposed estimators are analyzed under both null and alternative hypotheses. Under the null hypothesis, the restricted estimator (*RE*) has the smallest risk followed by the pre-test or shrinkage estimators. However, the pre-test or shrinkage estimators perform the best followed by the unrestricted estimator (*UE*) and restricted estimator (*RE*) when the parameter moves away from the subspace of the restrictions. The conditions of superiority of the proposed estimator for departure parameter are provided in Tabatabaey et al. (2004a). It is demonstrated that the positive rule shrinkage estimator utilizes both sample and non-sample information and performs uniformly better than *UE* and ordinary shrinkage estimators.

Giles (1991) considered the pre-test estimator for the restricted linear model with spherically symmetric error disturbances. He has investigated some finite sample properties of the estimators numerically for the case of the multivariate t-error. Gilies (1992) also considered pre-test (*PT)* estimator for two sample linear regression model under spherically symmetric disturbances. Kibria (1996) considered SE for the multicollinear data and for the restricted linear model with Students t error. Using MSE criterion, he discussed the performance of the estimators with respect both to non-centrality and ridge parameters. Judge et al. (1985) discussed the finite sample properties of James-Stein type and its positive part of the location parameter vector under squared errors loss with errors having a multivariate t-distribution. They compared the risk functions of the estimators via a monte carlo experiments.

Singh (1991) discussed the properties of James-Stein rule estimators in a regression model with multivariate Student's t-error. In general, the risk characteristics are found to be the same under normal and non-normal errors. However, there is very limited literature on the analytical results relating to the finite samples properties of positive rule shrinkage estimator for the linear model with non-normal error distribution.

Khan (2004) discussed the role of shape parameter for the shrinkage estimators of the mean vector of multivariate student t-population. Stein-type estimators based on the sample information and uncertain non-sample information were defined in Khan (2004). The impact of shape parameter on the performance of the shrinkage and positive rules shrinkage estimators with respect to the criteria of unbiasedness and minimum quadratic risk were also investigated. For more on this topic the readers are referred to Khan (2004) and Tabatabaey et al. (2004a) and references therein.

6.2 Pre-test and Shrinkage Estimation under Stochastic Constraint and Multivariate *T*- Error

The pre-test or shrinkage estimation under the general linear hypothesis (exact or non-stochastic) are available in literature. In rare cases we have exact prior information on the linear combination of parameters while estimating economic relations. Some uncertainty about the prior information are stochastic for many practical situation (see Theil and Golberger, 1961 and Theil 1963). In that case non-stochastic constraint does not work. Here we will discuss about the estimation of β when the error distribution belongs to (24) and it is suspected that β may be restricted to the stochastic subspace defined by

$$h = H\beta + \delta, \tag{34}$$

where h is an $q \times 1$ vector of observations, H is an $q \times p$ matrix of known constants of full rank q, and δ is an $q \times 1$ vector of errors, which is distributed according to the laws belonging to the class of compound normal distributions. That is,

$$f(\delta) = \int_0^\infty f_q(\delta \mid \omega) g(\omega) d\omega, \tag{35}$$

where $f(\delta)$ is the p.d.f. of δ, $f_q(v \mid \omega)$ is the normal p.d.f. with mean vector ψ and variance-covariance matrix $\omega^2 \Lambda \, (\omega > 0)$ and $g(\omega)$ is the inverted Gamma density with scale σ^2 and degrees of freedom ν, denoted by $IG(\nu, \sigma^2)$. Thus

$$f(\delta) = \frac{\Gamma\left((\nu+q)/2\right) |\Lambda|^{-1/2}}{(\nu\pi)^{q/2} \Gamma(\nu/2) \sigma^q} \left(1 + \frac{(\delta - \psi)' \Lambda^{-1} (\delta - \psi)'}{\nu\sigma^2}\right)^{-(\nu+q)/2}, \quad 0 < \nu, \sigma, < \infty, \quad -\infty < \delta < \infty. \tag{36}$$

with

$$E(\delta) = \psi, \; E(\delta\delta') = \sigma_e^2 \Lambda + \psi\psi', \; \sigma_e^2 = \frac{\nu}{\nu - 2}\sigma^2, \; \nu > 2.$$

We assume that $\delta \mid \gamma$ and $e \mid \gamma$ are independent. We combine the sample and stochastic prior information to get the following linear statistical model

$$\begin{bmatrix} y_1 \\ h \end{bmatrix} = \begin{bmatrix} X \\ H \end{bmatrix} \beta + \begin{bmatrix} e \\ \delta \end{bmatrix}, \tag{37}$$

where

$$\begin{bmatrix} e \\ \delta \end{bmatrix} | \, \omega \sim N_{N+q} \left\{ \begin{bmatrix} 0 \\ \psi \end{bmatrix}, \omega^2 \begin{bmatrix} I_N & 0 \\ 0 & \Lambda \end{bmatrix} \right\},$$

subject to condition

$$[-H, I_q] \begin{bmatrix} \beta \\ H\beta + \psi \end{bmatrix} = \psi = 0.$$

Rewrite the model as

$$y = Z\phi + u,$$

subject to exact restriction

$$R\phi = \psi = 0, \tag{38}$$

where

$$y = \begin{bmatrix} y_1 \\ \Lambda^{-1/2} h \end{bmatrix}, Z = \begin{bmatrix} X & 0 \\ 0 & \Lambda^{-1/2} \end{bmatrix}, \phi = \begin{bmatrix} \beta \\ H\beta + \psi \end{bmatrix}, R = [-H, I_q]$$

and

$$u = \begin{bmatrix} e \\ \Lambda^{-1/2} (\delta - \psi) \end{bmatrix} | \, \omega \sim N_{N+q} (0, \omega^2 I_{n+q}).$$

Then using (35), the p.d.f. of u is obtained as,

$$f(u) = \frac{\Gamma\left((N+q+\nu)/2\right)}{(\pi\nu)^{(N+q+\nu)/2} \Gamma(\nu/2) \sigma^{(N+q)}} \left(1 + \frac{u'u}{\nu\sigma^2}\right)^{-(N+q+\nu)/2}, \quad 0 < \nu, \sigma, < \infty, \quad -\infty < u_i < \infty.$$

For the full model the Unrestricted Least Squares Estimator (UE) of ϕ is given by

$$\hat{\phi}^{UE} = (Z'Z)^{-1} Z'y = \begin{bmatrix} \hat{\phi}_1^{UE} \\ \hat{\phi}_2^{UE} \end{bmatrix} = \begin{bmatrix} \hat{\beta}^{UE} \\ h \end{bmatrix}, \tag{39}$$

where $\hat{\beta}^{UE} = (X'X)^{-1} X'y$ is the unrestricted estimator of β. The Restricted Least Squares Estimator (RE) of ϕ is given by

$$\hat{\phi}^{RE} = \hat{\phi}^{UE} - (Z'Z)R'[R(Z'Z)^{-1}R']^{-1}\hat{\phi}^{UE} = \begin{bmatrix} \hat{\phi}_1^{RE} \\ \hat{\phi}_2^{RE} \end{bmatrix} = \begin{bmatrix} \hat{\beta}^{RE} \\ H\hat{\beta}^{RE} \end{bmatrix}, \tag{40}$$

where $\hat{\beta}^{RE}$ is the stochastic hypothesis restricted estimator of β and is given by

$$\hat{\beta}^{RE} = \hat{\beta}^{UE} - S^{-1}H'(HS^{-1}H'+\Omega)^{-1}(H\hat{\beta}^{UE} - h), \tag{41}$$

where $S = X'X$ is the information matrix. Note that unlike usual restricted estimator, the stochastic hypothesis restricted least squares estimator does not satisfy the condition $H\hat{\beta}^{RE} = h$. The Preliminary Test Least Squares Estimator (PT) of ϕ is given by

$$\begin{aligned} \hat{\phi}^{PT} &= \hat{\phi}^{RE} I(L_N \leq L_{N,\alpha}) + \hat{\phi}^{UE} I(L_N > L_{N,\alpha}) \\ &= \begin{pmatrix} \hat{\beta}^{PT} \\ H\hat{\beta}^{RE} I(L_N \leq L_{N,\alpha}) + hI(L_N > L_{N,\alpha}) \end{pmatrix}, \end{aligned} \tag{42}$$

where

$$\hat{\beta}^{PT} = \hat{\beta}^{RE} I(L_N \leq L_{N,\alpha}) + \hat{\beta}^{UE} I(L_N > L_{N,\alpha})$$

is the stochastic preliminary test least squares estimator,

$$L_N = \frac{(H\hat{\beta}^{UE} - r)'(HS^{-1}H'+\Omega)^{-1}(H\hat{\beta}^{UE} - h)}{qS_e^2},$$

with

$$S_e^2 = \frac{(y - Z\hat{\phi}^{UE})'(y - Z\hat{\phi}^{UE})}{N-p} = \frac{(y_1 - X\hat{\beta}^{UE})'(y_1 - X\hat{\beta}^{UE})}{N-p},$$

$L_{N,\alpha}$ is the upper α-level critical value of L_N and $I(A)$ is the indicator function of the set A. Under the null hypothesis and normal theory, L_N follows a central F-distribution with $(q, N-p)$ degrees of freedom while under the alternative hypothesis, $H\phi \neq 0$, the pdf of L_N is given by

$$g_{q,N-p}(L_N, \Delta, \nu) =$$

$$\sum_{r=0}^{\infty} \left(\frac{q}{N-p}\right)^{q/2+r} \frac{\Gamma\left(\frac{\nu}{2}+r\right)\Gamma\left(\frac{N-p+q}{2}+r\right)}{\Gamma(r+1)\Gamma\left(\frac{\nu}{2}\right)\Gamma\left(\frac{q}{2}+r\right)\Gamma\left(\frac{N-p}{2}\right)} \frac{\left(\frac{\Delta}{\nu-2}\right)^r}{\left(1+\frac{\Delta}{\nu-2}\right)^{\nu/2+r}} \frac{L_N^{q/2+r-1}}{\left(1+\frac{q}{N-p}L_N\right)^{\frac{N-p+q}{2}+r}}, \tag{43}$$

where

$$\Delta = \frac{\psi'(HC^{-1}H'+\Lambda)^{-1}\psi}{\sigma_e^2},$$

is the departure parameter from the null hypothesis. It is important to remark that $\hat{\phi}^{PT}$ is bounded and performs better than $\hat{\phi}^{RE}$ in some parameter space. The Stein-type shrinkage estimator (SE) of ϕ is defined as

$$\hat{\phi}^{SE} = \hat{\phi}^{UE} - dL_N^{-1}(\hat{\phi}^{UE} - \hat{\phi}^{RE}) = \begin{pmatrix} \hat{\beta}^{SE} \\ h - dL_N^{-1}(h - H\hat{\beta}^{RE}) \end{pmatrix}, \tag{44}$$

where,

$$\hat{\beta}^{SE} = \hat{\beta}^{UE} - dL_N^{-1}(\hat{\beta}^{UE} - \hat{\beta}^{RE})$$

is the stochastic shrinkage estimator and

$$d = \frac{(q-2)(N-p)}{q(N-p+2)}, \quad and\ q \geq 3.$$

The SE in (44) will provide uniform improvement over $\hat{\phi}^{UE}$, however it is not a convex combination of $\hat{\phi}^{UE}$ and $\hat{\phi}^{RE}$. This encourage one to find an alternative estimator. Hence, we define a better estimator, namely, the positive-rule shrinkage estimator (PR) of β as follows:

$$\hat{\phi}^{PR} = \hat{\phi}^{SE} - (1 - dL_N^{-1}) I(L_N \leq d)(\hat{\phi}^{UE} - \hat{\phi}^{RE}) = \begin{pmatrix} \hat{\beta}^{PR} \\ h - dL_N^{-1}(h - H\hat{\beta}^{RE}) - (1 - dL_N^{-1}) I(L_N \leq d)(h - H\hat{\beta}^{RE}) \end{pmatrix}, \tag{45}$$

where

$$\hat{\phi}^{PR} = \hat{\beta}^{SE} - (1 - dL_N^{-1}) I(L_N \leq d)(\hat{\beta}^{UE} - \hat{\beta}^{RE})$$

is the stochastic positive rule shrinkage estimator. Tabatabaey et al. (2004b) analyzed above five well known possible stochastic restricted estimators namely, unrestricted estimator (*UE*), restricted estimator (*RE*), preliminary test (*PT*) estimator, shrinkage estimator (*SE*) and positive rule (*PR*) shrinkage estimators for the multivariate *t* regression model. The bias and risk functions of the proposed estimators are analyzed under both the null and alternative hypotheses. Under the null hypothesis, the restricted estimator (*RE*) has the smallest risk followed by the pre-test or shrinkage estimators. However, the pre-test or shrinkage estimator performs the best followed by the unrestricted estimator (*UE*) and restricted estimator (*RE*) when the parameter moves away from the subspace of the restrictions. It has been evident that the positive rule shrinkage estimator utilizes both sample and non-sample information and performs uniformly better than *UE* and ordinary shrinkage estimator.

6.3 Ridge Regression under Multivariate *T*- Error

It is observed from (27) that the properties of the usual Least Squares Estimator (LSE) of β depends heavily on the characteristics of the information matrix $C = X'X$. If the columns of C matrix are linearly dependent, then the least squares estimator (LSE) produces unduly large sampling variances. Moreover, some of the regression coefficients may be statistically insignificant with wrong sign and the deduction of meaningful statistical inference become difficult for the researcher. Hoerl and Kennard (1970 a,b) found that multicollinearity is a common problem in many fields of applications. To resolve this problem, they suggested to use $C(k) = X'X + kI_p, \quad (k \geq 0)$ rather than C in the estimation of β. The resulting estimator of β is known as the Ridge Regression Estimator (RRE). Hoerl and Kennard (1970) considered the unrestricted ridge regression estimator ($URRE$) of β, as

$$\tilde{\beta}(k) = R(k)\tilde{\beta}, \tag{46}$$

where $R(k) = [I_p + kC^{-1}]^{-1}$ is the ridge or biasing parameter and $k \geq 0$ is the shrinkage parameter. Based on the $RLSE$, Sarkar (1992) proposed the following restricted ridge regression estimator ($RRRE$),

$$\hat{\beta}^{RE}(k) = R(k)\hat{\beta}. \tag{47}$$

Similarly, Saleh and Kibria (1993) considered the Preliminary Test Ridge Regression Estimator ($PTRRE$) of β as,

$$\hat{\beta}^{PT}(k) = \hat{\beta}(k)I(L_N \leq L_{N,\alpha}) + \tilde{\beta}(k)I(L_N > L_{N,\alpha}) = R(k)\hat{\beta}^{PT}. \tag{48}$$

Under MSE criterion Saleh and Kibria (1993) compared *URRE*, *RRRE* and *PTRRE* and discussed their relative merits and demerits. Tabatabey et al. (2004a) considered the *SE* and *PRE* under the ridge regression and obtained the corresponding Shrinkage Ridge Regression Estimator (*SRRE*) and Positive rule ridge regression estimator (*PRRRE*). Under the both quadratic bias and risk criterion he compared all five ridge regression estimators, namely, *URRE*, *RRRE*, *PTRRE*, *SRRE* and *PRRRE*. They also pinpointed some insights of these five estimators.

6.4 Ridge Regression under the Conflicting Statistics and Multivariate T -Error

In order to define the preliminary test estimators of β, we first consider the three well-known test-statistics for testing $H_0: H\beta = h$ against $H_A: H\beta \neq h$ with Students' t-error, namely (i) the Wald (W) test (ii) the likelihood ratio (LR) test and (iii) the Lagrangian Multiplier (LM) test and they are respectively given by

$$\begin{aligned} L_W &= \lambda(N)\frac{(H\tilde{\beta}-h)'(HC^{-1}H')^{-1}(H\tilde{\beta}-h)}{\tilde{\sigma}_N^2}, \\ L_{LR} &= N[\ln(\hat{\sigma}_N^2) - ln(\tilde{\sigma}_N^2)], \\ L_{LM} &= \frac{(H\tilde{\beta}-h)'(HC^{-1}H')^{-1}(H\tilde{\beta}-h)}{\lambda(N)\hat{\sigma}_N^2}, \end{aligned} \tag{49}$$

where $\tilde{\sigma}_N^2 = \frac{1}{N}(y - X\tilde{\beta})'(y - X\tilde{\beta})$ and $\hat{\sigma}_N^2 = \frac{1}{N}(y - X\hat{\beta})'(y - X\hat{\beta})$ are the unrestricted and restricted maximum likelihood estimator of σ^2 and

$$\lambda(N) = \frac{N+\nu}{N+\nu+2}, \quad 0 < \lambda(N) < 1.$$

For details, we refer to Ullah and Walsh (1984). Note that if N is large then $\lambda(N)$ is close to 1, then the results in (49) also hold for normal regression model.

The test statistics in (49) can also be written as follows

$$
\begin{aligned}
L_W &= \lambda(N)\frac{Nq}{N-p}L_N, \\
L_{LR} &= N\ ln\left(1+\frac{q}{N-p}L_N\right), \\
L_{LM} &= (\lambda(N))^{-1}\left(\frac{NqL_N}{N-p+qL_N}\right),
\end{aligned} \tag{50}
$$

where L_N is the test statistic for testing the null hypothesis (28). The test statistic L_N follows a central F-distribution with $(q, N-p)$ degrees of freedoms under H_0 (see Zellner 1976, and King 1980). Ullah and Zinde-Walsh (1984) showed that for these test statistics in (50) the following inequalities hold:

$$
\begin{aligned}
&L_W < L_{LR} < L_{LM} && if \quad \omega < \omega_1(\lambda(N)), \\
&L_{LR} \le L_W < L_{LM} && if \quad \omega_1(\lambda(N)) \le \omega < \omega_2(\lambda(N)), \\
&L_W \ge L_{LM} > L_{LR} && if \quad \omega_2(\lambda(N)) \le \omega < \omega_3(\lambda(N)), \\
&L_W > L_{LR} \ge L_{LM} && if \quad \omega > \omega_3(\lambda(N)),
\end{aligned} \tag{51}
$$

where $\omega = L_W/N$, and $\omega_1(\lambda(N)), \omega_2(\lambda(N))$ and $\omega_3(\lambda(N))$ are respectively the unique positive roots (depending on $\lambda(N)$) of the equations

$$
\begin{aligned}
\lambda(N)\omega - log\,(1+\omega) &= 0, \\
(\lambda(N))^2\omega - \omega(1+\omega)^{-1} &= 0, \\
\lambda(N) log\,(1+\omega) - \omega(1+\omega)^{-1} &= 0.
\end{aligned} \tag{52}
$$

Clearly under $H_0 : H\beta = h$, we have

$$
P_0\left(L_W \ge \lambda(N)\frac{Nq}{N-p}F_{q,N-p}(\alpha)\right) = P_0\left(F_{q,N-p} \ge F_{q,N-p}(\alpha)\right) = \alpha. \tag{53}
$$

It follows from (53) that the size of the Wald test can be greater or less than the LR-test depending on the solution for ω and the value of $\lambda(N)$. Similar comments apply to the size of LR and LM tests. The exact sampling distribution of the three test statistics are complicated. Thus in practice the critical regions of the tests are commonly based on asymptotic approaches (see Evans and Savin, 1982). It is known that the asymptotic distributions of the three tests is approximated by the chi-squared distribution with q degrees of freedom. Let the α level critical value of the distribution be $\chi_q^2(\alpha)$ as a first approximation. This choice of critical value for three tests leads to conflicts as in the case of finite sample inference. The inequalities of statistics given in (51) will occur if

$$
\begin{aligned}
&\text{Either } L_{LR} < \chi_q^2(\alpha) < L_{LM}, \text{ or, } L_W < \chi_q^2(\alpha) < L_{LR}, \\
&\text{Either } L_W < \chi_q^2(\alpha) < L_{LM}, \text{ or, } L_{LR} < \chi_q^2(\alpha) < L_W, \\
&\text{Either } L_{LM} < \chi_q^2(\alpha) < L_W, \text{ or, } L_{LR} < \chi_q^2(\alpha) < L_{LM}, \\
&\text{Either } L_{LPR} < \chi_q^2(\alpha) < L_W, \text{ or, } L_{LM} < \chi_q^2(\alpha) < L_{LR}
\end{aligned} \tag{54}
$$

respectively. For the normal error case, Evan and Savin (1982) showed that on using the $\chi_q^2(\alpha)$ critical value there are two characteristics: First, they will differ with respect to their sizes and powers in small samples and there may be conflict between their conclusions. Second: when the sizes of the tests are corrected to be the same, the power are approximately the same and there may no any confliction. For the Students' t error case, Ullah and Walsh (1984) showed that the inequalities in (51) is complicated, the relationship among the sizes of these test and the possibility of conflict quite different than the normal case. For more details, please see Ullah and Walsh (1984). For excellent references and for various researches on W, LR and LM tests, readers are refereed to Savin (1976), Berndt and Savin (1977), Rao and Mukerjee (1977), Evans and Savin (1982), Ullah and Zinde-Walsh (1984), Billah and Saleh (2000 a,b), Kibria and Saleh (2003 a,b) and Kibria (2004) among others. Based on the above considerations, Kibria and Saleh (2003b) considered the following preliminary test ridge regression estimators (*PTRRE*) based on W, LR and LM tests, which are given below,

$$\hat{\beta}_*^{PT}(k) \quad = \quad \hat{\beta}(k)I(L_* \leq \chi_q^2(\alpha)) + \tilde{\beta}(k)I(L_* > \chi_q^2(\alpha)), \tag{55}$$

where L_* stands for either of L_W, L_{LR} and L_{LM} tests. They studied the finite sample properties of the *PTRRE* based on the Wald, the Likelihood Ratio and the Lagrangian Multiplier tests. They have effectively determined some conditions on the departure parameter and the shrinkage parameter for the superiority of the proposed estimators. They have also discussed the method of choosing optimum level of significance to obtain minimum guaranteed efficient estimators. The *PTRRE* based on W test is found to perform the best with the choice of smallest level of significance to yield the best estimator in the sense of highest minimum guaranteed efficiency. It is worth noting that the analysis of Kibria and Saleh (2003b) under the assumption of *t*-distribution coincides with the traditional results developed under normality assumption for large degrees of freedom. The analysis of Kibria and Saleh (2003b) also coincide with that of Billah and Saleh (2000b), where they considered performance of Preliminary Test Least Squares Estimator (*PTLSE*) based on W, LR and LM tests.

6.5 Predictive Inference with Multivariate Students *T*-Error

The distribution of future responses for given a set of observed data is known as predictive distribution. The predictive distribution for future responses for the regression model with errors having a multivarite *t*-distribution can be obtained by using structural, classical and Bayesian approaches and they give the same future predictive distribution (see Kibria 2006). Using the structural relation of the model $Y = \beta X + \sigma e$, where e has the density (1), Haq and Khan (1990) derived the predictive distribution of N_f future responses for the future regression model, $Y_f = \beta X_f + \sigma e_f$. The predictive distribution of Y_f has been obtained as N_f dimensional multivariate t-distribution with mean vector $b_y X_f$, variance covariance matrix ,

$$S_y \mid I_N - X_f[XX' + X_f X_f']^{-1} X_f \mid^{-1/2}$$

and degrees of freedom $N - p$. It is to noted that the predictive distribution does not depend on the degrees of freedom of the parent t-distribution. Since

$$(N-p)S_y^2(Y_f - b_y X_f)(I_n - X_f[XX' + X_f X_f']^{-1} X_f)(Y_f - b_y X_f)'$$

has an F-distribution with N_f and $N - p$ degrees of freedoms, the prediction region for a set of future responses can be determined for any desired coverage probability. The literature on predictive inference for the multivariate linear model with *ARMA(p,q)* processes are limited. However, most of the studies relating to predictive inference from the multivariate linear models have been considered with errors having a Gaussian *ARMA(p,q)* process. Kibria and Haq (1998) have considered the multivariate linear model with error having a multivariate t-distribution and a *ARMA(1,1)* process. They derived the marginal likelihood function of the parameters and predictive distribution of a set of future responses. The intra-class correlation coefficient ρ is often used to measure the degree of intra-family resemblance with respect to biomedical attributes such as blood pressure, weight, height etc. Intra-class correlation also arises in psychology, education and genetics, where the population may be divided into clusters. For instance, in sampling from a biological population, it is advantageous to select a sample of population clusters and then to select a sample of organisms within these clusters. The consequences of such a sampling procedure is that the sample observations within a cluster may exhibit a residual covariance of intra-class structure rather than diagonal form (Wiorkowski, 1975).

Using structural relation of the model, Kibria and Haq (1999) derived the predictive distribution for a set of future responses from a multivariate linear model with error following a multivariate t-distribution and intra-class covariance structure. The predictive distribution obtained as multivariate t with appropriate parameters. For $\rho = 0$, the results in Kibria and Haq (1999) reduces to Haq and Khan (1990).

6.6 Bayesian Regression Analysis under Multivariate *T*-Distribution

Zellenr (1976) analyzed the traditional multiple regression model

$$y = X\beta + u$$

under the assumption that error terms have a joint multivariate t-distribution with the following p.d.f.

$$p(u \mid \nu, \sigma) = \frac{\nu^{\nu/2}(\nu + N/2)}{\pi^{N/2}\Gamma(\nu/2)\sigma^N}\left(\nu + \frac{u'u}{\sigma^2}\right)^{-(N+\nu)/2}.$$

Thus he considered uncorrelated but dependent errors for the model for which it can be shown that maximum likelihood estimator for the regression coefficient vector, $\hat{\beta} = (X'X)^{-1}X'Y$ is simply the least square estimator and the maximum likelihood estimator is a minimum variance linear unbiased estimator when relevant moments exist. Zellner (1976) also considered the Bayesian analysis of the regression model with a diffuse prior p.d.f. for the regression coefficients as

$$p(\beta, \sigma) \propto \frac{1}{\sigma^2}$$

and found that the joint posterior distribution of the parameters as multivariate t-distribution as arises from multivariate normal model. The marginal posterior distribution for the regression parameter was also found to have a multivariate t-distribution, which does not depend on the unknown parameter ν. However, the posterior distribution of the scale parameter was found to be in the form of F-distribution. He has shown several important inference procedures about the regression coefficients developed under normality remain valid for the multivariate t-model. However, inference about the scale parameter will depend on the extent of the departure from normality, as measured by the value of degrees of freedoms parameters. He also presented a natural conjugate prior distribution for the regression model having a multivariate Student t-distribution.

7 Concluding Remarks

This paper has discussed some basic properties of the multivariate t-distribution with applications to various fields of science and business. Special attention has been made for the estimation for the parameters for the linear regression model under the multivarite t-distribution. Predictive distributions for future observation under the multivariate t-distribution are also discussed. Since the application of the multivariate t-distribution has been increasing day by day in business and econometrics, the paper will help and encourgae young researchers to stimulate further research.

Acknowledgments

The paper was completed while the first author was visiting (July 2006) Professor B. K. Sinha, Stat-Math Unit, Indian Statistical Institute (ISI), Kolkata, India. He is grateful to ISI for providing him with excellent research facilities. The second author gratefully acknowledges the excellent research facilities provided by the King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia.

References

Ali, M.M. and Joarder, A.H. (1991). Distribution of the correlation coefficient for the class of bivariate elliptical models. *Canadian Journal of Statisttics*, **19**, 447--452.

Berndt, E. and Savin, N. E. (1977). Conflict among criteria for testing Hypothesis in the multivariate linear regression model, *Econometrica*, 45, 1263-1278.

Billah, B. and Saleh, A. K. Md. E. (2000a). Performance of the PTE based on the conflicting W, LR and LM Tests in Regression model, *Advances on Methodological and Applied Aspects of Probability and Statistics* (N.Balakrishnan Editor. Gordon and Breach Science Publishers.

Billah, B. and Saleh, A. K. Md. E. (2000b). Performance of Three Large Sample Tests in the Actual Formation of the Pretest Estimators for Regression Models in the t -error. *Journal of Applied Statistical Science.* Vol 9,No .3 pp237-250

Blattberg, R.C. and Gonedes, N.J. (1974). A comparison of the stable and student distributions as statistical models for stock prices. *Journal of Business*, **47**, 224--280.

Cambanis, S., S. Huang, and G. Simons (1981). On the theory of elliptically contoured distributions. *Journal of Multivariate Analysis*, 11, 368-385.

Cornish, E.A. (1954). The multivariate t-distribution associated with a set of normal sample deviates. *Australian Journal of Physics*, **7**, 531--542.

Evans, G. B. A. and Savin, N. E. (1982). Conflict among the criteria revisited; The W, LR and LM Tests, *Econometrica*, 50, 737-748.

Fama, E. F. (1965). The behavior of stock market prices. *Journal of Business*, 38, 34-105.

Fang, K.T. and Anderson, T.W. (1990). *Statistical Inference in Elliptically Contoured and Related*

Distributions. Allerton Press, New York.

Fang, K.T.; Kotz, S. and Ng, K.W. (1990). Symmetric Multivariate and Related Distributions. *Chapman and Hall*, London.

Fisher, R.A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, **10**, 507--521.

Giles, A. J. (1991). Pretesting for Linear Restrictions in a Regression Model with Spherically Symmetric Distributions. *J. Econometrics*, **50**, 377-398.

Giles, A. J. (1992). Estimation of the error variance after a preliminary test of homogeneity in a regression model with spherically symmetric disturbances. *J. Econometrics,* **53**, 345-361.

Gnanadesikan, R. (1977). *Methods for Statistical Data Analysis of Multivariate Observations*, New York: Wiley.

Han, C-P, and Bancroft, T. A. (1968). On Pooling means when variance is unknown. *Journal of the American Statistical Association*, **63**, 1333-1342.

Haq, M.S. and *Khan, S.* (1990). Prediction Distribution for the Linear Model with Multivariate Student-t Error Distribution. *Communications In Statistics, Theory & Methods,* **19**, 4705-4712.

Hoerl, A. E. and Kennard, R. W. (1970a). Ridge Regression: Biased Estimation for Non-orthogonal Problems. *Technometrics,* **12**, 55-67.

Hoerl, A. E. and Kennard, R. W. (1970b). Ridge Regression: Application to Non-Orhtogonal Problem. *Technometrics*, **12**, 69-82.

James, W. and Stein, C. (1961). Estimation with quadratic loss. *Proceeding of the Fourth Barkeley Symposium on Mathematical Statistics and Probability*. University of California Press, Berkeley, CA, 361-379.

Joarder, A.H. (1992). Estimation of the scale matrix of a multivariate t-model. Unpublished Ph.D. thesis. The University of Western Ontario, Canada.

Joarder, A.H. (1995a). Estimation of the scale matrix of a multivariate t-model under a squared error loss function. *Joarder of Statisical Research*, **29**, 55--66.

Joarder, A.H. (1995b). Estimation of the trace of the scale matrix of a multivariate t-model. *Econometrics Conference at Monash.* Monash University, Clayton, Australia.

Joarder, A.H. (2006). An introduction to bivariate *t*-distribution. Technical Report No. 345, Department of Mathematical Sciences, King Fahd University of Petroleum & Minerals, Saudi Arabia.

Joarder, A.H. and Ahmed, S.E. (1996). Estimation of the characteristic roots of scale matrix. *Metrika*, **44**, 259--267.

Joarder, A.H. and Ali, M.M. (1992a). On some generalized Wishart expectations. *Communications in Statistics – Theory and Methods*, **21**, 283--294.

Joarder, A.H. and Ali, M.M. (1992b). Distribution of the correlation matrix for a class of elliptical models. *Communications in Statistics- Theory and Methods*, **21**, 1953--1964.

Joarder, A.H. and Ali, M.M. (1996). On the characteristic function of the multivariate *t* distribution. *Pakistan Journal of Statistics*, **12**, 55--62.

Joarder, A.H. and Ali, M.M. (1997). Estimation of the scale matrix of a multivariate t-model under entropy loss function. *Metrika*, **46**, 21--32.

Joarder, A.H. and Beg, G.K. (1999). Estimation of the trace of the scale matrix under a squared error loss. *Statistica*, 59(2), 181-192.

Judge, G.G. and Bock, M.E. (1978). *The Statistical Implications of Pre-test and Stein-rule Estimators in Econometrics*, North-Holland Publishing Company, Amsterdam.

Judge, G. C., Miyazaki, S. and Yancey, T. (1985). Minimax estimators for the location vectors of spherically symmetric densities. *Econometric Theory*, **1**, 409-417.

Kelejian, H.H. and Prucha, I.R. (1985). Independent or uncorrelated disturbances in linear regression. *Economic Letters*, **19**, 35--38.

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

Khan, S. (2004). The role of the shape parameter for the shrinkage estimators of the mean vector of multivariate student t population. *International Journal of Statistical Science*, 3, 69-89.

Khan, S. (1997). Inference on the mean vectors of two multivariate Student-*t* population. Technical Report No. 224, Department of Mathematical Sciences. King Fahd University of Petroleum and Minerals. Saudi Arabia.

Kibria, B. M. G. (1996). On shrinkage ridge regression estimators for restricted linear models with multivariate t disturbances. *Student*, **1** (3), 177-188.

Kibria, B. M. G. (2004). Performance of the Shrinkage Preliminary Tests Ridge Regression Estimators Based on the Conflicting of W, LR and LM Tests. *Journal of Statistical Computation And Simulation*, **74** (11). To appear.

Kibria, B. M. G. (2006). The Matric t Distribution and its Applications in Predictive Inference. *Journal of Multivariate Analysis,* **97**, (3), 785-795.

Kibria, B. M. G. and Haq, M. S. (1998). Marginal likelihood and prediction for the multivariate ARMA(1,1) linear models with multivariate t error distribution. *Journal of Statistical Research,* 32(1), 71-80.

Kibria, B. M. G. and Haq, M. S. (1999). The multivariate linear model with multivariate t and intra-class covariance structure. *Statistical Paper*, **40**, 263-276.

Kibria, B. M. G. and Saleh, A. K. Md. E. (1993). Performance of shrinkage preliminary test estimator in regression analysis. *Jahangirnagar Review*, **A 17**, 133-148.

Kibria, B. M. G. and Saleh, A. K. Md. E. (2003a). Effect of W, LR, and LM Tests on the

Performance of Preliminary Test Ridge Regression Estimators. *Journal of the Japan Statistical Society*, **33**(1), 119-136

Kibria, B. M. G. and Saleh, A. K. Md. E. (2003b). Preliminary test ridge regression estimators with Student's t errors and conflicting test-statistics, *Metrika*, **59**, 105-124.

Kibria, B. M. G. and Saleh, A. K. Md. E. (2004). Performance of Positive Rule Estimator in the Ill-Conditioned Gaussian Regression Model, *Calcutta Statistical Association Bulletin*. **55** (219-220), 211-241.

King, M.L. (1980). Robust tests for spherical symmetry and their applications to least squares regression. *Annals of Statistics*,, 8, 1265--1271.

Kotz, S.; Balakrishnan, N and Johnson, N.L. (2000). *Continuous Multivariate Distributions*, Volume 1. John Wiley, New York.

Kotz, S. and Nadarajah, S. (2004). *Multivariate t Distributions and their Applications*. Cambridge University Press, USA.

Lange, K.L.; Little, R.J.A. and Taylor, J.M.G. (1989). Robust statistical modeling using the t distribution. *J. Amer. Statist. Assoc.*, **84**, 881 --896.

Mosteller, F. (1948). On pooling data. *Journal of the American Statistical Association*, **43**, 231-242.

Mathai, A. M. and Provost, S. B. (1992). *Quadratic forms in Random Variables*, New York, Marcel Dekker..

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. John Wiley, New York.

Nadarajah, S. and Kotz, S. (2006). Percentage points of the multivariate t distribution. *Internatioal Statistical Review*, 74, 15-30.

Pearson, K. (1923). On non-skew frequency surface. *Biometrika*, 15, 231.

Prucha, I. R. , and Kelejian, H. H. (1984). The structure of simultaneous equation estimators: A generalization towards nonnormal disturbances, *Econometrica*, **52** , 721-736.

Rahman, M. and Saleh, A.K.M.E. (1974). Explicit form of the distribution of the Behrens-Fisher d-statistic. *Journal of the Royal Statistical Society*, Series B, **36**, 54-60.

Rao, C. R. and Mukerjee, R. (1977). Comparison of LR, Score and Wald test in a non-ID setting. *Journal of Multivariate Analysis*, 60, 99-110.

Saleh, A. K. Md. E. (2006). *Theory of Preliminary test and Stein- Type Estimation with Applications*. Wiley, New York.

Saleh, A. K. Md. E. and Kibria, B. M. G. (1993). Performances of some new preliminary test ridge regression estimators and their properties. *Communications in Statistics-Theory and Methods*, **22**, 2747-2764.

Saleh, A. K. Md. E. and Sen, P. K. (1978). Non-parametric estimation of location parametric after a preliminary test regression. *Annals of Statistics*, **6**, 154-168.

Sarkar, N. (1992). A new Estimator Combining the Ridge Regression and the Restricted Least Squares Method of Estimation. *Communications in Statistics-Theory and Methods*, 21(7), 1987-2000.

Savin, N. E. (1976). Conflict among testing procedures in a linear regression model with autoregressive disturbances. *Econometrica*, **44**, 1303-1315.

Seo, T. and Toyama, T. (1996). On the estimation of kurtosis parameter in elliptical distributions. *Journal of Japan Statisical Society*, **26**, 59--68.

Shalabh (1995). Performance of Stein-rule procedure for simultaneous prediction of actual and average values of study variable in linear regression models. *Bulletin of International Statistical Institute*, **56**, 1375-1390.

Singh, R.S. (1987). A family of improved estimators in linear regression models with errors having multivariate Student-t distribution. Working paper. Department of Econometrics and Operations Research, Monash University, Clayton, Australia.

Singh, R.S. (1988). Estimation of error variance in linear regression models with errors having multivariate Student-t distribution with unknown degrees of freedom. *Economic Letters*, **27**, 47--53.

Singh, R. S. (1991). James-Stein rule estimators in linear regression models with multivariate t distributed error. *Australia Journal of Statistics*, **33**, 145-158.

Sutradhar, B.C. (1984). Contributions to Multivariate Analysisbased on Elliptic T Model. Unpublished Ph.D. Thesis. The University of Western Ontario, Canada.

Sutradhar, B.C. (1988a). The non-null distribution of Hotelling's T^2 under t-population. *Recent Developments in Statistics and Actuarial Science*, University of Western Ontario, Canada.

Sutradhar, B.C. (1988b). Testing linear hypothesis with t error variables. *Sankhya*, **50**, 175--180.

Sutradhar, B.C. (1990). A note on the power of the F-test for testing linear hypothesis with elliptic t errors. *Communications in Statistics -- Theory and Methods*, **19**, 863--867.

Sutradhar, B. C. and Ali, M. M. (1986). Estimation of the parameter of a regression model with a multivariate t error. *Communications in Statistics - Theory and Methods*, **A 15**, 429-450.

Sutradhar, B.C. and Ali, M.M. (1989). Estimation of parameters of regression model with a multivariate-t error model. *Communications in Statistics - Theory and Methods*, **15**, 429--450.

Tabatabaey, S. M., Saleh, A. K. Md. E. and Kibria, B. M. G. (2004a). Estimation Strategies For Parameters of the Linear Regression Models With Spherically Symmetric Distributions. *Journal of Statistical Research*, **38** (1), 13-31.

Tabatabaey, S. M., Saleh, A. K. Md. E. and Kibria, B. M. G. (2004b). Simultaneous estimation of regression parameters with spherically symmetric errors under possible stochastic constraints. *International journal of Statistical Sciences*, **3**, 1-20.

Theil, H. (1963). On the use of incomplete prior information in regression analysis. *Journal of the American Statistical Association*, **58**, 401-414.

Theil, H. and Goldberger, A. S. (1961). On pure and mixed statistical estimation in Economics. *International Economics Review*, **2**, 65-78.

Ullah, A. and Zinde-Walsh, V. (1984). On the robustness of LM, LR and W tests in regression models. *Econometrics*, **52**, 1055-1066.

Watson, G. N. (1958). *A Treatise on the Theory of Bessel Functions*. Cambridge University Press, Cambridge.

Wiorkowski, J.J. (1975). Unbalanced regression analysis with residuals having a covariance structure of inta-class form. *Biometrics,* 31, 611-618.

Zellner, A. (1976). Bayesian and non-Bayesian analysis of the regression model with multivariate Student t error terms, *Journal of the American Statistical Association*, **71**, 400-405.