*Journal of Statistical Research* *ISSN 0256 - 422 X*
*2006, Vol. 40, No. 1, pp. 59-72*
*Bangladesh*

# A SHORT REVIEW OF MULTIVARIATE $t$-DISTRIBUTION

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

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

summary

This paper reviews most important properties of a location-scale multivariate $t$-distribution. A conditional representation of the distribution is exploited to outline moments, characteristic function, marginal and conditional distributions, distribution of linear combinations and quadratic forms. Stochastic representation is also used to determine the covariance matrix of the distribution. It also makes an attempt to justify an uncorrelated $t$- model and overviews distribution of the sum of products matrix and correlation matrix. Estimation strategies for parameters of the model is briefly discussed. Finally the recent trend of linear regression with the uncorrelated $t$- model is discussed.

*Keywords and phrases:* Elliptically Contoured Distributions; Marginal likelihood; Multivariate model; Predictive distribution; Robustness

*AMS 2000 Classification:* Primary 62J05; Secondary 62H99.

## 1 Introduction

The multivariate $t$-distribution is a natural generalization of the univariate Student $t$-distribution. Cornish (1954) derived it first in connection with a set of normal sample deviates.

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. Many 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

© Institute of Statistical Research and Training (ISRT), University of Dhaka, Dhaka 1000, Bangladesh.

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 a correlated $t$-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). Interested readers may go through Kelker (1970), Cambanis, Huang and Hsu (1981), Fang and Anderson (1990), Kotz and Nadarajah (2004) 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 correlated $t$-model.

## 2 The Probability Model

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

$$f(\boldsymbol{x}) = \frac{|\boldsymbol{\Sigma}|^{-1/2}}{C(\nu,p)\pi^{p/2}}\left[1+\frac{1}{\nu}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]^{-(\nu+p)/2}, \tag{2.1}$$

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

$$\Gamma((\nu+p)/2)C(\nu,p) = \nu^{p/2}\ \Gamma(\nu/2). \tag{2.2}$$

The $p$-variate random variable $\boldsymbol{X}$ has a mean vector $\boldsymbol{\mu}$ and a covariance matrix $\boldsymbol{\Sigma}^* = \nu^*\boldsymbol{\Sigma}$, where $\nu^* = \nu/(\nu-2)$ and can well be represented by $T_p(\boldsymbol{\mu},\ \nu^*\boldsymbol{\Sigma})$ where the shape parameter $\nu(>2)$ is assumed to be known. If $p=1$, $\boldsymbol{\mu}=0, \boldsymbol{\Sigma}=1$, then the p.d.f. in (2.1) defines the univariate $t$-distribution. When $p=2$, $\boldsymbol{\mu}=0, \boldsymbol{\Sigma}=1$, then the p.d.f. in (2.1) 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

$$f(\boldsymbol{x}) = \int_0^\infty \frac{|\omega^2\boldsymbol{\Sigma}|^{-1/2}}{(2\pi)^{p/2}}\exp(-(\boldsymbol{x}-\boldsymbol{\mu})'(\omega^2\boldsymbol{\Sigma})^{-1}(\boldsymbol{x}-\boldsymbol{\mu})/2)\ h(\omega)d\omega \tag{2.3}$$

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

$$h(\omega) = \frac{2(\nu/2)^{\nu/2}}{\Gamma(\nu/2)}\omega^{-(\nu+1)} \exp\left(\frac{-\nu}{2\omega^2}\right), \tag{2.4}$$

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

$$\boldsymbol{X}|\Omega = \omega \sim N_p(\boldsymbol{\mu}, \omega^2\boldsymbol{\Sigma}). \tag{2.5}$$

As $\nu \to \infty$, the random variable $\omega$ becomes a degenerate random variable with all the non-zero mass at the point unity and, consequently, the multivariate $t$-distribution in (2.1) converges to the multivariate normal distribution $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. 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 $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_N$ does not imply their independence unless $\nu \to \infty$.

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

$$f_1(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N) = f(\boldsymbol{x}_1)f(\boldsymbol{x}_2)\ldots f(\boldsymbol{x}_N) \tag{2.6}$$

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 multivariate $t$-distributions is given by

$$f_2(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N) = \frac{|\boldsymbol{\Sigma}|^{-N/2}}{C(\nu, Np)\pi^{Np/2}}\left(1 + \frac{Q}{\nu}\right)^{-(\nu+Np)/2} \tag{2.7}$$

where $Q = \sum_{j=1}^{N}(\boldsymbol{x}_j - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_j - \boldsymbol{\mu})$ and $\boldsymbol{x}_j$ $(j = 1, 2, \ldots, N)$ is the realized value of a $p$-component random vector $\boldsymbol{X}_j$ $(j = 1, 2, \ldots, N)$ having the $t$ -distribution $T_p(\boldsymbol{\mu}, \nu^*\boldsymbol{\Sigma})$ where $\nu^* = \nu/(\nu - 2)$. The p.d.f. in (2.7) will hereinafter be called the *uncorrelated $t$-distribution.*

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 (2.6). It may be remarked that observations in (2.7) are independent if and only $\nu \to \infty$, in which case the p.d.f. in (2.7) will be the product that of $N$ independent $p$-dimensional normal distributions $N_p$ $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

A more general case would be to consider $k$-samples each of which

$$(\boldsymbol{X}_{g_1}, \boldsymbol{X}_{g_2}, \ldots, \boldsymbol{X}_{gN_g}), \qquad g = 1, 2, \ldots, k$$

is a sample of size $N_g$ from $T_p(\boldsymbol{\mu}_g, \nu^*\boldsymbol{\Sigma})$, $g = 1, 2, \ldots, k$. The joint p.d.f. of observations of $k$-samples would be

$$f_3(\boldsymbol{x}_{11}, \boldsymbol{x}_{12}, \ldots, \boldsymbol{x}_{kN_k}) = \frac{|\boldsymbol{\Sigma}|^{-N/2}}{C(\nu, Np)\pi^{Np/2}} \left(1 + \frac{1}{\nu}\, Q\right)^{-(\nu+Np)/2} \tag{2.8}$$

where $Q = Q_1 + Q_2 + \cdots + Q_k = \sum_{g=1}^{k} Q_g$, $Q_g = \sum_{j=1}^{N_g} (\boldsymbol{x}_{g_j} - \boldsymbol{\mu}_g)' \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}_{g_j} - \boldsymbol{\mu}_g)$ and $N = N_1 + N_2 + \cdots + N_k = \sum_{g=1}^{k} N_g$.

# 3 Some Properties of the Multivariate $t$- Distribution

## 3.1 Moments and Characteristic Function

By the use of the mixture representation in (2.5), it can be easily proved that

$$E(\mathbf{X}) = E(E(\mathbf{X}|\omega)) = \boldsymbol{\mu}$$

and

$$Cov(\mathbf{X}) = E[Cov(\mathbf{X}|\omega)] + \mathrm{Cov}[E(\mathbf{X}|\omega)] = E(\omega^2 \boldsymbol{\Sigma}) = \nu^* \boldsymbol{\Sigma},$$

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 $\mathbf{X}$ following a multivariate $t$-distribution with p.d.f in (2.1) is given by

$$\phi_{\boldsymbol{X}}(\boldsymbol{t}) = E(e^{i\boldsymbol{t}'\boldsymbol{X}}) = e^{i\boldsymbol{t}'\boldsymbol{\mu}} \; \frac{\|(\nu\boldsymbol{\Sigma})^{1/2}\boldsymbol{t}\|^{\nu/2}}{2^{\nu/2-1}\;\Gamma(\nu/2)} \; K_{\nu/2}(\|(\nu\boldsymbol{\Sigma})^{1/2}\boldsymbol{t}\|) \tag{3.1}$$

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

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

$$K_\alpha(t) = \left(\frac{2}{t}\right)^{\alpha} \; \frac{\Gamma(\alpha + 1/2)}{\sqrt{\pi}} \int_0^\infty (1+u^2)^{-(\alpha+1)} \cos tu \, du, \qquad t > 0, \quad \alpha > -1/2. \tag{3.2}$$

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

$$E\left(\prod_{i=1}^{p} Z_i^{k_i}\right) = \begin{cases} 0 \quad \text{if at least one } \; k_i (i = 1, 2, \ldots, p) \; \text{ is odd} \\ \quad \nu^{k/2} \; \dfrac{\Gamma((\nu-k)/2)}{2^k \; \Gamma(\nu/2)} \; \displaystyle\prod_{i=1}^{p} \frac{k_i!}{(k_i/2)!}, \qquad \nu > k \\ \quad \text{if all } \; k_i\text{'s } \; (i = 1, 2, \ldots, p) \; \text{ are even} \end{cases} \tag{3.3}$$

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

$$E\left(R^k\right) = \nu^{k/2}\ \frac{\Gamma((p+k)/2)\Gamma((\nu-k)/2)}{\Gamma(p/2)\,\Gamma(\nu/2)}, \nu > k \tag{3.4}$$

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

It follows from (3.1) that the characteristic function of the univariate Student $t$-distribution with p.d.f.

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

is given by

$$\phi_{\mathbf{X}}(t) = \frac{\nu^{\nu/4}|t|^{\nu/2}}{2^{\nu/2-1}\ \Gamma(\nu-2)} \qquad K_{\nu/2}(\sqrt{\nu}|t|), \quad \nu > 2\ .$$

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 $\boldsymbol{X}$ in (3.1) can also be written as

$$\phi_{\boldsymbol{X}}(\boldsymbol{t}) = E(e^{i\boldsymbol{t}'\boldsymbol{X}}) = e^{i\boldsymbol{t}'\boldsymbol{\mu}}\ \psi(\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}). \tag{3.5}$$

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

$$Cov(\boldsymbol{X}) = -2\psi'(0)\boldsymbol{\Sigma} \text{ and } \kappa = \frac{\psi''(0)}{\{\psi'(0)\}^2} - 1 \tag{3.6}$$

respectively (Seo and Toyama, 1996).

## 3.2 Marginal and Conditional Distributions

It is well-known that linear combinations, marginal and conditional distributions of the components of $\boldsymbol{X}$ follow the multivariate $t$-distribution (see e.g. Sutradhar, 1984). Let $\boldsymbol{X}, \boldsymbol{\mu}, \boldsymbol{t}$ and $\boldsymbol{\Sigma}$ be partitioned as

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix}, \ \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \ \boldsymbol{t} = \begin{pmatrix} \boldsymbol{t}_1 \\ \boldsymbol{t}_2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$$

where $\boldsymbol{X}_2, \boldsymbol{\mu}_2, \boldsymbol{t}_2 \ \epsilon\ \Re^q \quad (q < p)$ and $\boldsymbol{\Sigma}_2$ is a $q \times q$ positive definite matrix. By the use of the characteristic function of $\boldsymbol{X}$ given by (3.1), it may be easily checked that $\boldsymbol{X}_2 \sim T_q\ (\boldsymbol{\mu}_2,\ \nu^*\boldsymbol{\Sigma}_{22})$ where $\nu^* = \nu/(\nu-2)$. The conditional distribution of $\boldsymbol{X}_1$ given $\boldsymbol{X}_2 = x_2$ is $T_{p-q}\ (\boldsymbol{\mu}_{1\cdot 2},\ \nu_{1\cdot 2}\ \boldsymbol{\Sigma}^*_{11.2})$ where

$$\begin{aligned} \boldsymbol{\mu}_{1.2} &= \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\ (\boldsymbol{x}_2 - \boldsymbol{\mu}_2), \\ \nu_{1.2} &= \nu/(\nu+q-2), \quad \text{and} \\ \boldsymbol{\Sigma}^*_{11.2} &= (1 + (\boldsymbol{x}_2-\boldsymbol{\mu}_2)'(\nu\boldsymbol{\Sigma}_{22})^{-1}\,(\boldsymbol{x}_2-\boldsymbol{\mu}_2))\,\boldsymbol{\Sigma}_{11.2} \end{aligned}$$

with $\boldsymbol{\Sigma}_{11.2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$. 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 $\mathbf{Z}$ is said to have a spherical distribution if its probability density function (*pdf*) can be written as

$$f(z) = g(z'z) \tag{3.7}$$

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), Kibria and Haq (1999a), Kibria and Saleh (2003), Kotz and Nadarajah (2004) and the references therein.

Let $\boldsymbol{Z}$ have the multivariate $t$-distribution with p.d.f.

$$f(\boldsymbol{z}) = g(\boldsymbol{z}'\boldsymbol{z}) = \frac{1}{c(\nu,p)\pi^{p/2}}\left(1+\frac{\boldsymbol{z}'z}{\nu}\right)^{(\nu+p)/2} \tag{3.8}$$

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

$$\frac{R^2}{p} \sim F(p,\nu)$$

and that

$$E\left(R^k\right) = \nu^{k/2}\ \frac{\Gamma((p+k)/2)\Gamma((\nu-k)/2)}{\Gamma(p/2)\,\Gamma(\nu/2)}, \nu > k \tag{3.9}$$

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

$$E\left(R^2\right) = \frac{\nu p}{\nu-2},\ \ \nu > 2 \tag{3.10}$$

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

Consider the elliptical random variable $\boldsymbol{X} = \boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2}\boldsymbol{Z}$, where $\boldsymbol{Z}$ has the p.d.f. given by (3.7). It is well known (Cambanis, Hunag and Simons, 1981) that the covariance matrix of $\boldsymbol{X}$ is given by $Cov(\boldsymbol{X}) = -2\psi'_{\boldsymbol{X}}(\mathbf{0})\boldsymbol{\Sigma}$, where $\psi_{\boldsymbol{X}}(t) =$ exp $(i\boldsymbol{t}'\boldsymbol{\mu})\psi(||\boldsymbol{\Sigma}^{1/2}t||)$ is the characteristic function of $\boldsymbol{X}$.

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

For any elliptical random variable $\boldsymbol{X}$, where $\boldsymbol{X} = \boldsymbol{\mu}+\boldsymbol{\Sigma}^{1/2}\boldsymbol{Z}$ with $\boldsymbol{Z}$ having the p.d.f.(3.7), it is well known that $Cov(\boldsymbol{X}) = \frac{1}{p}E\left(R^2\right)\boldsymbol{\Sigma}$ (Cambanis, Huang and Hsu, 1981, or Joarder, 1992). In this section we outline how the covariance matrix of multivariate $t$-distribution can be derived by the above result.

It follows from (3.10) that $Cov(\boldsymbol{X}) = \frac{1}{p}\left(\frac{\nu p}{\nu-2}\right)\boldsymbol{\Sigma} = \frac{\nu}{\nu-2}\boldsymbol{\Sigma}$, where $\boldsymbol{X} = \boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2}\boldsymbol{Z}$ with $\boldsymbol{Z}$ having the p.d.f. given by (3.8).

### 3.4 Distribution of a Linear Function

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

### 3.5 Distribution of Quadratic Forms

Suppose the random variable $\mathbf{X}$ has a multivariate $t$-distribution with mean vector $\boldsymbol{\mu}$, covariance matrix $\boldsymbol{\Sigma}$ and degrees of freedom $\nu$, then $\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X}/p$ has the $F$ distribution with $\nu$ and $p$ degrees of freedoms and non-centrality parameter $\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}/p$.

## 4 Uncorrelated $t$- Model

### 4.1 Distribution of the Sum of Products Matrix Based on the Uncorrelated $t$-Model

The sum of product matrix based on the uncorrelated $t$- distribution (2.7) is given by

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

where $\overline{\boldsymbol{X}} = \sum_{j=1}^{N}\boldsymbol{X}_j/N$. It follows from (2.5) that for a given $\omega$, the random matrix $\boldsymbol{A}$ has the usual Wishart distribution

$$\mathbf{A}|\Omega = \omega \sim W_p(m, \omega^2\Sigma), \qquad m = N-1 \tag{4.1}$$

i.e. the p.d.f. of $\boldsymbol{A}$ is given by

$$\int_0^\infty \frac{|\omega^2\boldsymbol{\Sigma}|^{-1/2}}{2^{mp/2}\Gamma_p(m/2)}\ |\boldsymbol{A}|^{(m-p-1)/2}\ \exp\left(-\frac{1}{2}\mathrm{tr}(\omega^2\boldsymbol{\Sigma})^{-1}\boldsymbol{A}\right) \qquad h(\omega)d\omega, \tag{4.2}$$

where $\boldsymbol{A} > 0$, $m = N-1 \geq p$ 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 $\nu\omega^{-2} \sim \chi^2_\nu$. The completion of integration in (4.2) results in the p.d.f. of $\boldsymbol{A}$ given by

$$\frac{|\boldsymbol{\Sigma}|^{-m/2}}{C(\nu, mp)2^{mp/2}\ \Gamma_p(m/2)}\ |\boldsymbol{A}|^{(m-p-1)/2}\left(1+\frac{1}{\nu}\ \mathrm{tr}\ \boldsymbol{\Sigma}^{-1}\boldsymbol{A}\right)^{-(\nu+mp)/2} \tag{4.3}$$

(cf. Sutradhar and Ali, 1989).

By the use of the mixture representation in (4.1), it is easy to derive the expected values of $|\boldsymbol{A}|^k, |\boldsymbol{A}|^k\boldsymbol{A},\ |\boldsymbol{A}|^k\boldsymbol{A}^{-1},\ (\mathrm{tr}\ \boldsymbol{A})^2, \mathrm{tr}(\boldsymbol{A}^2)$ 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 $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. 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 the entire class of elliptical distributions which accommodates the correlated $t$-distribution as a special case. The result has been generalized for multivariate elliptical distribution by Joarder and Ali (1992b) for the correlation matrix $\boldsymbol{R}$. The *pdf* of $\boldsymbol{R}$ is given by

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

where

$$H_{n,p}(\Gamma) = 2^{-\frac{(n-3)p}{2}} \int_0^\infty \cdots \int_0^\infty (v_1 v_2 \ldots v_p)^{n-2} e^{-v'\Gamma v/2} \prod_{i=1}^p dv_i$$

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

## 5 Estimation of Parameters

### 5.1 Estimation of Parameters for One population

The maximum likelihood estimators of the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ of the correlated $t$-distribution in (2.7) are given by $\hat{\boldsymbol{\mu}} = \overline{\boldsymbol{X}}$ and $\hat{\boldsymbol{\Sigma}} = \boldsymbol{A}/N$ 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 (2.7) for finite value of the shape parameter $\nu$. The sample mean $\overline{\boldsymbol{X}}$ is obviously an unbiased and consistent estimator of $\boldsymbol{\mu}$. The unbiased estimator of $\boldsymbol{\Sigma}$ is given by $\hat{\boldsymbol{\Sigma}} = \boldsymbol{A}/(\nu^* m)$, where $\nu^* = \nu/(\nu-2)$ and $m = N-1$ (see Fang and Anderson, pp. 208).

Joarder (1995a) considered the estimation of the scale matrix $\boldsymbol{\Sigma}$ of the uncorrelated $t$-distribution under a squared error loss function. It may be remarked that the scale matrix $\boldsymbol{\Sigma}$ determines the covariance matrix up to a known constant $\nu^*$. Joarder and Ahmed (1996) developed estimation strategy for eigenvalues of $\boldsymbol{\Sigma}$ of the correlated $t$-distribution given by (2.7). The estimation of the trace of the scale matrix $\boldsymbol{\Sigma}$ under a squared error loss was considered by Joarder and Beg (1999). The estimation of $\boldsymbol{\Sigma}$ 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 (2.8). The equality of mean vectors $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ can then be tested by

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

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 correlated $t$-distribution obtained by reparametrizing $\nu^*\boldsymbol{\Sigma}$ by $\boldsymbol{\Sigma}$ in (2.7). The following derivation of $T^2$-statistic is based on the mixture representation of multivariate $t$-distribution (see e.g. Khan 1997).

By virtue of the mixture representation of (2.5 ), it follows that conditional on $\omega$,

$$\frac{m}{p}\, T^2 \sim F_{p,m}(\delta_\omega), \tag{5.1}$$

where $F_{p,m}(\delta_\omega)$ denotes a noncentral $F$-distribution with parameters $p,\ m = m_1 + m_2 - p + 1$ and $\delta_\omega = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)'\,(\omega^2\boldsymbol{\Sigma})^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$. 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 (5.1) that under $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$

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

The power function of the test $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ against $H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$ was discussed by Sutradhar (1988a) and Sutradhar (1990).

Khan (1997) considered the estimation of the mean vector of a 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 $\Delta$. 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 multivariate uncorrelated $t$-distribution. It is King (1980) who laid the rigorous mathematical foundation of linear regression analysis under broader distributional assumptions of spherical symmetry which includes uncorrelated $t$-distribution as a special case. Prompted by the works of Zellner (1976) and King (1980), many authors used uncorrelated $t$-distribution for modeling real world data. Sutradhar and Ali (1989) generalized Zellner's model with errors having a correlated $t$-distribution given by (1.7). Lange, Little and Taylor (1989) applied uncorrelated $t$ -distribution to a variety of situations.

The null distribution of the usual $F$-statistic in a linear regression model under correlated $t$-distribution in (1.7) is robust but the power function depends on the form of (1.7); see e.g. Sutradhar(1988b,1990) for a detailed proof. For the linear regression model with errors having an uncorrelated $t$-distribution, 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 $t$-distribution with unknown degrees of freedom.

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), Blatberg and Gonedes (1974)).

Because of the above reasons, Sutradhar and Ali (1989), Giles (1991, 1992), Singh (1988, 1991), Kibria (1996, 2004), Kibria and Haq (1999b), Judge et al. (1985), Kibria and Saleh (2003), and Tabatabaey et al. (2004a, 2004b), among others were motivated to use the multivariate $t$-distribution as the error distribution of the linear regression model. They have considered several pre-test and shrinkage type estimators for estimating the regression parameters and compared their performance with respect to various loss functions. The predictive distribution under the multivariate $t$ -distribution was studied by Kibria & Haq (1998, 1999b) among others. Since the application of the uncorrelated $t$-distribution is increasing in business and econometric studies especially through pre-test and shrinkage estimation and predictive inference, an updated and comprehensive review is necessary, which is under the current investigation.

# 7 Acknowledgments

The first author gratefully acknowledges the excellent research facilities provided by the King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia. The authors are thankful to the Chief Editor, Professor A. K. Md. E. Saleh for his constructive suggestions that have helped to improve the presentation of the paper.

# References

[1] 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.

[2] Blatberg, 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.

[3] Cambanis, S., S. Huang, and G. Simons (1981). On the theory of elliptically contoured distributions. *J. Mult. Anal.*, 11, 368– 385.

[4] Cornish, E.A. (1954). The multivariate $t$-distribution associated with a set of normal sample deviates. *Austr. J. Physics*, 7, 531–542.

[5] Fama, E. F. (1965). The behavior of stock market prices. *J. Bus.*, 38, 34-105.

[6] Fang, K.T. and Anderson, T.W. (1990). Statistical Inference in Elliptically Contoured and Related Distributions. Allerton Press, New York.

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

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

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

[10] 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.

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

[12] Joarder, A.H. (1992). Estimation of the Scale Matrix of a Multivariate $T$ -Model. Unpublished Ph.D. thesis. The University of Western Ontario, Canada.

[13] Joarder, A.H. (1995a). Estimation of the scale matrix of a multivariate $t$ -model under a squared error loss function. *J. Statist. Res.*, 29 , 55–66.

[14] 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.

[15] Joarder, A.H. (2006). Some useful integrals and their applications in correlation analysis. To appear.

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

[17] Joarder, A.H. and Ali, M.M. (1992a). On some generalized Wishart expectations. *Commun. Statist. – Theor. Meth.*, 21, 283–294.

[18] Joarder, A.H. and Ali, M.M. (1992b). Distribution of the correlation matrix for a class of elliptical models. *Commun. Statist. – Theor. Meth.*, 21, 1953–1964.

[19] Joarder, A.H. and Ali, M.M. (1996). On the characteristic function of the multivariate $t$ distribution. *Pak. J. Statist.*, 12, 55–62.

[20] 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.

[21] Joarder, A.H. and Beg, G.K. (1999). Estimation of the trace of the scale matrix of a multivariate t-model under a squared error loss, Statistica, 59(2), 181-189.

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

[23] Kelejian, H.H. and Prucha, I.R. (1985). Independent or uncorrelated disturbances in linear regression. *Econ. Lett.*, 19, 35–38.

[24] Kelker, D. (1970). Distribution theory of spherical distributions and locations scale parameters. *Sankhya*, A 32, 419-430.

[25] Khan, S. (2004). The role of the shape parameter for the shrinkage estimators of the mean vector of multivariate student $t$ population. *Int. J. Statist. Sci.*, 3, 69-89.

[26] 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.

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

[28] Kibria, B. M. G. (2004). Conflict among the shrinkage estimators induced by W, LR and LM tests under a Student's $t$ regression model. *J. Kor. Statist. Soc.*, 33 (4), 411-433.

[29] 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. *J. Statist. Res.*, 32(1) 71-80.

[30] Kibria, B. M. G. and Haq, M. S. (1999a). Predictive inference for the elliptical linear model. *J. of Mult. Anal.,*. 68, 235-249.

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

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

[33] King, M.L. (1980). Robust tests for spherical symmetry and their applications to least squares regression. *Ann. Statist.*, 1265–1271.

[34] Kotz, S. and Nadarajah, S. (2004). Multivariate $t$ Distributions and their Applications. Cambridge University Press, USA.

[35] 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.

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

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

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

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

[40] Rahman, M. and Saleh, A. K. Md. E. (1975). Explicit form of the distribution of Behrens-Fisher statistic. *J. Roy. Statist. Soc.*, 36 (1), 54-60. Corrigendum, 36 (3), 466.

[41] Seo, T. and Toyama, T. (1996). On the estimation of kurtosis parameter in elliptical distributions. *J. Japan Statist. Soc.*, 26, 59–68.

[42] 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.

[43] Singh, R.S. (1988). Estimation of error variance in linear regression models with errors having multivariate Student-$t$ distribution with unknown degrees of freedom. *Econ. Lett.*, 27, 47–53.

[44] Singh, R. S. (1991). James-Stein rule estimators in linear regression models with multivariate $t$ distributed error. *Aus. J. of Statist.*, 33, 145-158.

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

[46] 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.

[47] Sutradhar, B.C. (1988b). Testing linear hypothesis with $t$ error variables. *Sankhya*, 50, 175–180.

[48] Sutradhar, B.C. (1990). A note on the power of the $F$-test for testing linear hypothesis with elliptic $t$ errors. *Commun. Statist.– Thero. Meth.*, 19, 863–867.

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

[50] Sutradhar, B.C. and Ali, M.M. (1989). Estimation of parameters of regression model with a multivariate-$t$ error model. *Commun. Statist. - Theor. Meth,*, 15, 429–450.

[51] 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. *J. Statist. Res.*, 38 (1), 13-31.

[52] 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 stichastic constraints. *Int. J. Statist. Sci.*, 3, 1-20.

[53] Watson, G. N. (1958). A Treatise on the Theory of Bessel Functions. Cambridge University Press, Cambride.

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