

IMPROVED ESTIMATION OF THE MEAN VECTOR FOR STUDENT-t MODEL

Shahjahan Khan

Department of Mathematics and Computing

University of Southern Queensland

Toowoomba, Q 4350, Australia

email: khans@usq.edu.au

Key Words: uncertain prior information; maximum likelihood estimator; likelihood ratio test; James-Stein estimator; preliminary test and shrinkage estimator; bias; quadratic risk; multivariate normal, Student-t and inverted gamma distributions; dominance; and relative efficiency.

ABSTRACT

Improved James-Stein type estimation of the mean vector $\boldsymbol{\mu}$ of a multivariate Student-t population of dimension p with ν degrees of freedom is considered. In addition to the sample data, *uncertain prior information* on the value of the mean vector, in the form of a null hypothesis, is used for the estimation. The usual maximum likelihood estimator (mle) of $\boldsymbol{\mu}$ is obtained and a test statistic for testing $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ is derived. Based on the mle of $\boldsymbol{\mu}$ and the test statistic the *preliminary test estimator* (PTE), Stein-type *shrinkage estimator* (SE) and *positive-rule shrinkage estimator* (PRSE) are defined. The *bias* and the *quadratic risk* of the estimators are evaluated. The relative performances of the estimators are investigated by analyzing the risks under different conditions. It is observed that the PRSE dominates over the other three estimators, regardless of the validity of the null hypothesis and the value ν .

1 INTRODUCTION

Ever since the publication of seminal papers of Stein (1956), and James and Stein (1961) there has been growing interest in the search for ‘improved’ estimator of the mean vector, $\boldsymbol{\mu}$ for the multivariate normal population. Earlier, Bancroft (1944) and later Han and Bancroft (1968) developed the preliminary test estimator that uses uncertain prior information, in addition to the sample information. In a series of papers Saleh and Sen (1978, 1985) explored the preliminary test approach to James-Stein type estimation. Many authors have contributed to this area, notably Sclove et al. (1972), Judge and Bock (1978), Stein (1981), Chou and Stewardman (1990), and Maatta and Casella (1990). All the above developments are based on the normal model. Investigations on improved estimation for Student-t model have been rather a recent development. Zellner (1976), Singh (1991), Giles (1991), Anderson (1993), Tabatabaey (1995), Celler et al (1995) and Khan and Saleh (1997) studied the preliminary test and Stein-type estimation for linear models with multivariate Student-t errors. However, Khan and Saleh (1998) investigated the problem from the sampling theory approach. The current work also pursues the same approach from a multivariate perspective and provides comprehensive results on ‘improved’ estimation of the mean for the multivariate Student-t model in the presence of uncertain prior information on $\boldsymbol{\mu}$.

Consider a random sample of size n from a multivariate Student-t population with unknown mean $\boldsymbol{\mu}$, common variance σ^2 and arbitrary shape parameter ν . Also assume that *uncertain prior information* on the value of $\boldsymbol{\mu}$ is available. This can be expressed in the form of a null hypothesis; $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ which may be true, but not sure. We wish to incorporate both the sample data and the uncertain prior information in estimating the mean vector, $\boldsymbol{\mu}$. First we obtain the unrestricted maximum likelihood estimator (mle) of the mean vector and the common variance from the sample. Based on the unrestricted and restricted mle of σ^2 , we derive the likelihood ratio test for testing $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$. This test is robust and is applicable to the entire class of elliptical models (cf. Anderson, 1993). Following Bancroft (1944), the *preliminary test estimator* (PTE), $\hat{\boldsymbol{\mu}}^{pt}$ of $\boldsymbol{\mu}$ has been defined by using the test statistic

and the mle of $\boldsymbol{\mu}$. Then apply the preliminary test approach (see Tabatabaey, 1995, for instance) to define James-Stein estimators, namely, the *shrinkage estimator* (SE), $\hat{\boldsymbol{\mu}}^s$ and the *positive-rule shrinkage estimator* (PRSE), $\hat{\boldsymbol{\mu}}^{s+}$ for the mean vector. The bias, quadratic bias and squared error loss functions are obtained for the above four estimators. The relative performance of the estimators is investigated by analyzing the risks under different conditions. It is observed that the PRSE dominates the other estimators when H_0 holds good. It also over performs the other estimators under the alternative hypothesis if the squared distance between $\boldsymbol{\mu}$ and $\boldsymbol{\mu}_0$ is not too large. The dominance picture of the estimators under the null and alternative hypotheses is provided. For the computations and derivations of the main results of the paper the Student-t distribution is viewed as a mixture of the normal distribution and Inverted Gamma (IG) distribution.

A brief justification and short review of the application of the Student-t model have been provided in the next section. This section also specifies the model as a mixture of normal and Inverted Gamma distributions. In Section 3, the unrestricted, preliminary test, shrinkage and positive rule shrinkage estimators are defined. Section 4 provides some useful results for the computation of the bias and quadratic risks of the estimators. The bias of the estimators are computed and analyzed in Section 5. In Section 6, we evaluate the expressions of quadratic risks of the estimators. The performances of the estimators are studied in Section 7.

2 THE STUDENT-t MODEL

Fisher (1956) discarded the normal distribution as a sole model for the distribution of errors. Fraser (1979) showed that the results based on the Student-t models for linear models are applicable to those of normal models, but not the vice-versa. Prucha and Kelejian (1984) critically analyzed the problems of normal distribution and recommended the Student-t distribution as a better alternative for many problems. The failure of the normal distribution to model the fat-tailed distributions has led to the use of the Student-t

model in such a situation. In addition to being robust, the Student-t distribution is a ‘more typical’ member of the elliptical class of distributions. Moreover, the normal distribution is a special (limiting) case of the Student-t distribution. It also covers the Cauchy distribution on the other extreme. Extensive work on this area of non-normal models has been done in recent years. A brief summary of such literature has been given by Chmielewski (1981), and other notable references include Fang and Zhang (1980), Khan and Haq (1990), Fang and Anderson (1990), Gupta and Vargava (1993) and Celler et al (1995).

Let \mathbf{X} be a p -dimensional random vector having a normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix, $\tau^2 I_p$. Assume τ follows an Inverted Gamma distribution with shape parameter ν and density function

$$p(\tau; \nu, \sigma^2) = \left(\frac{2}{\Gamma(\nu/2)} \right) \left(\frac{\nu\sigma^2}{2} \right)^{\nu/2} \tau^{-(\nu+1)} e^{-\frac{\nu\sigma^2}{2\tau^2}}. \quad (2.1)$$

Then the distribution of \mathbf{X} , conditional on τ , is denoted by $\mathbf{X}|\tau \sim N_p(\boldsymbol{\mu}, \tau^2 I_p)$, and that of τ by $\tau \sim IG(\nu, \sigma)$. In the literature, it is well known that the mixture distribution of \mathbf{X} and τ is a multivariate Student-t, and is obtained by completing the following integral

$$p(\mathbf{x}; \boldsymbol{\mu}, \Sigma, \nu) = \int_{\tau=0}^{\infty} N_p(\boldsymbol{\mu}, \tau^2 I_p) IG(\nu, \sigma) d\tau \quad (2.2)$$

where $\Sigma = \sigma^2 I_p$ and ν is the number of degrees of freedom. The integration yields the density of \mathbf{X} as

$$p(\mathbf{x}; \boldsymbol{\mu}, \Sigma, \nu) = k_1(\nu, p) |\Sigma|^{-1/2} \left[\nu + (\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right]^{-\frac{\nu+p}{2}} \quad (2.3)$$

where $k_1(\nu, p) = \left\{ \Gamma\left(\frac{\nu+p}{2}\right) \nu^{\nu/2} \right\} \left\{ \pi^{p/2} \Gamma\left(\frac{\nu}{2}\right) \right\}^{-1}$ is the normalizing constant. The above density is the p -dimensional multivariate Student-t density with an arbitrary unknown shape parameter ν . In notation we write $\mathbf{X} \sim t_p(\boldsymbol{\mu}, \Sigma, \nu)$. A method of moment estimator for ν is given by Singh (1988). But here we are interested in the estimation of the mean vector.

Now, consider a random sample of size n from the above multivariate

Student-t population. The likelihood function of the sample is given by

$$L(\boldsymbol{\mu}, \Sigma, \nu; \mathbf{x}_1, \dots, \mathbf{x}_n) = k_n(\nu, p) |\Sigma|^{-\frac{n}{2}} \left[\nu + \sum_{j=1}^n (\mathbf{x}_j - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x}_j - \boldsymbol{\mu}) \right]^{-\frac{\nu+np}{2}} \quad (2.4)$$

where $k_n(\nu, p) = \left\{ \Gamma\left(\frac{\nu+np}{2}\right) \nu^{\nu/2} \right\} \left\{ \pi^{np/2} \Gamma\left(\frac{\nu}{2}\right) \right\}^{-1}$. Please refer to Khan (1997) for details on sampling from multivariate Student-t population by using the mixture of multivariate normal and inverted gamma distributions.

From the properties of the Student-t distribution, the marginal distribution of \mathbf{X}_j is p -dimensional Student-t with $E(\mathbf{X}_j) = \boldsymbol{\mu}$ and $\text{Cov}(\mathbf{X}_j) = \frac{\nu}{\nu-2} \Sigma$ for $j = 1, 2, \dots, n$. It may be noted that the \mathbf{X}_j 's are uncorrelated but not independent (cf. Anderson, 1993). However, for a given value of τ , each \mathbf{X}_j is independently normally distributed, that is, $[\mathbf{X}_j | \tau] \sim N_p(\boldsymbol{\mu}, \tau^2 I_p)$ for all j .

In this paper we wish to estimate the mean vector $\boldsymbol{\mu}$ based on the random sample $\mathbf{X}_1, \dots, \mathbf{X}_n$ when an *uncertain prior information* on $\boldsymbol{\mu}$ is available which can be presented by the null hypothesis, $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$. The researcher suspects that the H_0 may hold, but is not sure about it. To incorporate such prior information in to the estimation process Bancroft (1944) proposed the preliminary test estimator which removes the uncertainty in the null hypothesis through an appropriate statistical test on the H_0 . Towards the appropriate test, we derive the maximum likelihood estimator of the parameters and define the likelihood ratio test. Then based on the mle and the likelihood ratio test statistic, we define the preliminary test, shrinkage and positive rule shrinkage estimators for $\boldsymbol{\mu}$. To study the relative performance of the estimators, we compute the bias, quadratic bias and quadratic risk of the estimators. A detailed study of the risk analysis with respect to the departure of $\boldsymbol{\mu}$ from the H_0 has been provided.

3 THE ESTIMATORS

In this section, we derive the mle of the location and scale parameters of the Student-t model as well as the likelihood ratio test for the suspected null hypothesis. From (2.4) the log-likelihood function of the sample can be

written as

$$\ln L(\boldsymbol{\mu}, \sigma^2, \nu) = \ln k_n(\cdot) - \frac{np}{2} \ln \sigma^2 - \frac{\nu + np}{2} \ln \left[1 + \frac{\sum_{j=1}^n (\mathbf{x}_j - \boldsymbol{\mu})'(\mathbf{x}_j - \boldsymbol{\mu})}{\nu \sigma^2} \right]. \quad (3.1)$$

Then the maximum likelihood estimator (mle) of $\boldsymbol{\mu}$ is obtained as

$$\tilde{\boldsymbol{\mu}} = \frac{1}{n} \sum_{j=1}^n \mathbf{X}_j = \bar{\mathbf{X}} \quad (3.2)$$

and that of σ^2 as

$$\tilde{\sigma}^2 = \frac{1}{np} \sum_{j=1}^n (\mathbf{x}_j - \tilde{\boldsymbol{\mu}})'(\mathbf{x}_j - \tilde{\boldsymbol{\mu}}) = \frac{s^2}{np}. \quad (3.3)$$

The above mle's are termed as *unrestricted estimators* (UE) of $\boldsymbol{\mu}$ and σ^2 respectively.

Under the H_0 , the restricted log-likelihood function will replace $\boldsymbol{\mu}$ of the likelihood function by $\boldsymbol{\mu}_0$, and hence the restricted estimator of σ^2 becomes

$$\hat{\sigma}^2 = \frac{1}{np} \sum_{j=1}^n (\mathbf{x}_j - \boldsymbol{\mu}_0)'(\mathbf{x}_j - \boldsymbol{\mu}_0). \quad (3.4)$$

In the estimation of $\boldsymbol{\mu}$, when uncertain prior information on $\boldsymbol{\mu}$, in the form of the H_0 , $\boldsymbol{\mu} = \boldsymbol{\mu}_0$ is available, we need to first test the credibility of the H_0 .

An appropriate test statistic to test the null hypothesis is derived by using the likelihood ratio test procedure. The likelihood ratio statistic for the multivariate Student-t model is given by

$$\lambda = \left[\frac{\hat{\sigma}^2}{\tilde{\sigma}^2} \right]^{-np/2} = \left[\frac{\sum_{j=1}^n (\mathbf{x}_j - \boldsymbol{\mu}_0)'(\mathbf{x}_j - \boldsymbol{\mu}_0)}{s^2} \right]^{-np/2}. \quad (3.5)$$

Then it can be easily shown that under H_0 the statistic

$$T^2 = \lambda^{-2/np} = \frac{\sum_{j=1}^n (\mathbf{x}_j - \boldsymbol{\mu}_0)'(\mathbf{x}_j - \boldsymbol{\mu}_0)}{s^2} \quad (3.6)$$

has a scaled F -distribution with p and $m = (n - p)$ degrees of freedom (cf. Zellner, 1976), and can be used to test the H_0 . As discussed by Anderson (1993), the F -statistic stated above is robust, and it is valid for all the members of the elliptical class of distributions, not just for the normal or Student- t distributions. Thus to test the $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$, we perform the F -test based on the following monoton function of the T^2 -statistic:

$$F = \frac{p}{m} T^2 = \frac{\chi_n^2(\psi)}{\chi_m^2} \quad (3.7)$$

where $\psi = \frac{(\boldsymbol{\mu} - \boldsymbol{\mu}_0)'(\boldsymbol{\mu} - \boldsymbol{\mu}_0)}{2\sigma^2}$ is the non-centrality parameter when the H_0 is not true.

Following Bancroft (1944) and using Fisher's "recipe" of removing the uncertainty in the H_0 through performing an appropriate test, we define the *preliminary test estimator* (PTE) of $\boldsymbol{\mu}$ as a function of T^2 and $\tilde{\boldsymbol{\mu}}$ as follows:

$$\hat{\boldsymbol{\mu}}^{pt} = \hat{\boldsymbol{\mu}} I(T^2 \leq T_\alpha^2) + \tilde{\boldsymbol{\mu}} I(T^2 > T_\alpha^2) \quad (3.8)$$

where $\hat{\boldsymbol{\mu}} = \boldsymbol{\mu}_0$ is the *restricted estimator* (RE) of $\boldsymbol{\mu}$ under the H_0 ; T_α^2 is a value of the T^2 -statistic such that $P_r\{T^2 \leq T_\alpha^2\} = \alpha$ when the H_0 is true, for $0 \leq \alpha \leq 1$; and $I(A)$ is an indicator function of the set A .

By definition, the PTE is an extreme choice between the UE and RE. Thus it ignores all the potential values of $\boldsymbol{\mu}$ in the interval $(\hat{\boldsymbol{\mu}}, \tilde{\boldsymbol{\mu}})$ or $(\tilde{\boldsymbol{\mu}}, \hat{\boldsymbol{\mu}})$, as the case may be, as an estimate of $\boldsymbol{\mu}$. This problem, along with the dependency of the PTE on the choice of the level of significance (α) of the test, reduces the attractiveness of the estimator in many practical applications. The Stein-type *shrinkage estimator* (SE) can be defined by using the T^2 statistic that allows a smooth transition between $\tilde{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\mu}}$. Moreover, the SE does not depend on the choice of the level of significance. The SE for $\boldsymbol{\mu}$, by using the preliminary test approach, is defined as

$$\hat{\boldsymbol{\mu}}^s = \hat{\boldsymbol{\mu}} + (1 - k^* T^{-2})(\tilde{\boldsymbol{\mu}} - \hat{\boldsymbol{\mu}}) \quad (3.9)$$

where k^* is a shrinkage constant. An optimal value of k^* that minimizes the value of the risk function is found to be $k = \frac{p-2}{m+2}$.

The SE becomes unstable and unreliable if T^2 is too close to zero. To avoid this difficulty a modified Stein-type estimator, namely the *positive-rule shrinkage estimator* (PRSE) is defined as follows:

$$\hat{\boldsymbol{\mu}}^{s+} = \hat{\boldsymbol{\mu}} + (1 - kT^{-2})(\tilde{\boldsymbol{\mu}} - \hat{\boldsymbol{\mu}})I(T^2 > k). \quad (3.10)$$

In the forthcoming sections we study the properties of the above defined four estimators. For the computation of the bias and the quadratic risk of these estimators we apply the results in the next section.

4 SOME USEFUL RESULTS

Stein (1966) investigated the multivariate normal model and proved a number of useful results regarding the expectation of non-central chi-squared variables. Since those results are necessary for the evaluation of bias and risk expressions, we extend those results for the multivariate Student-t model as follows:

Lemma 4.1. *If \mathbf{U} is an $n \times 1$ vector of Student-t variables with ν d.f., mean $\boldsymbol{\beta}$ and covariance matrix I_n , an identity matrix of order n , then*

$$E[\phi(\mathbf{U}'\mathbf{U})\mathbf{U}] = \boldsymbol{\beta}E_\tau[\phi(\chi_{n+2}^2, \lambda)] = \boldsymbol{\beta}E[\phi(\chi_{n+2}^2, \lambda^*)] \quad (4.1)$$

where $\phi(\cdot)$ is a Borel measurable function, and $\lambda^* = \frac{\nu - 2}{\nu} \lambda$ in which $\lambda = \frac{\boldsymbol{\beta}'\boldsymbol{\beta}}{2}$.

Lemma 4.2. *If \mathbf{U} is an $n \times 1$ vector of Student-t variables with ν d.f., mean $\boldsymbol{\beta}$ and covariance matrix I_n , and \mathbf{V} is a positive definite matrix of order n , then*

$$\begin{aligned} E[\phi(\mathbf{U}'\mathbf{U})\mathbf{U}'\mathbf{V}\mathbf{U}] &= E_\tau[\phi(\chi_{n+2}^2, \lambda)]\text{tr}(\mathbf{V}) + \boldsymbol{\beta}'\mathbf{V}\boldsymbol{\beta}E_\tau[\phi(\chi_{n+4}^2, \lambda)] \\ &= E[\phi(\chi_{n+2}^2, \lambda^*)]\text{tr}(\mathbf{V}) + \boldsymbol{\beta}'\mathbf{V}\boldsymbol{\beta}E[\phi(\chi_{n+4}^2, \lambda^*)] \end{aligned} \quad (4.2)$$

where $\phi(\cdot)$ is a Borel measurable function and $\lambda^* = \frac{\nu - 2}{\nu} \lambda$ in which $\lambda = \frac{\boldsymbol{\beta}'\boldsymbol{\beta}}{2}$.

Lemma 4.3. *If \mathbf{U} is an $n \times 1$ vector of Student-t variables with ν d.f., mean $\boldsymbol{\beta}$ and covariance matrix I_n , then*

$$E[\phi(\mathbf{U}'\mathbf{U})\mathbf{U}'] = E_\tau[\phi(\chi_{n+2}^2, \lambda)]I_n + \boldsymbol{\beta}\boldsymbol{\beta}'E_\tau[\phi(\chi_{n+4}^2, \lambda)]$$

$$= E[\phi(\chi_{n+2}^2, \lambda^*)]I_n + \boldsymbol{\beta}\boldsymbol{\beta}'E[\phi(\chi_{n+4}^2, \lambda^*)]. \quad (4.3)$$

where $\phi(\cdot)$ is a Borel measurable function and $\lambda^* = \frac{\nu - 2}{\nu} \lambda$ in which $\lambda = \frac{\boldsymbol{\beta}'\boldsymbol{\beta}}{2}$.

The proof of the above lemmas follow directly from the proof of Judge and Bock (1978) in Appendix B2 by taking expectations of the respective final expressions with respect to the $IG(\nu, \sigma)$ distribution in the appropriate manner.

5 STUDY OF BIAS

The expressions for the bias of the estimators are evaluated to investigate its nature and implications. It is well known that the bias of the UE, $\tilde{\boldsymbol{\mu}}$ is

$$B_1(\tilde{\boldsymbol{\mu}}; \boldsymbol{\mu}) = E[\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}] = \mathbf{O}. \quad (5.1)$$

Thus the UE, $\tilde{\boldsymbol{\mu}}$ is an unbiased estimator for $\boldsymbol{\mu}$.

Theorem 5.1. The bias of the PTE is given by

$$B_2(\tilde{\boldsymbol{\mu}}^{pt}; \boldsymbol{\mu}) = E[\hat{\boldsymbol{\mu}}^{pt} - \boldsymbol{\mu}] = -\delta G_{p+2,m}^{(2)}(l_\alpha; \Delta^*) \quad (5.2)$$

where

$$G_{p+2,m}^{(2)}(l_\alpha; \Delta^*) = \sum_{r=0}^{\infty} \frac{\Gamma\left(\frac{p+2+m}{2} + r\right)}{\Gamma\left(\frac{p+2}{2} + r\right) \Gamma\left(\frac{m}{2}\right)} \xi_r(l_\alpha) \xi_r(\nu) \xi_r(\nu, \Delta^*) \quad (5.3)$$

in which $l_\alpha = \frac{p}{p+2} F_{p,m}(\alpha)$ with $F_{p,m}(\alpha)$ being the $(1-\alpha)$ -th quantile of a central F -distribution with p and $m = n-p$ d.f.; $h_\alpha = \frac{1}{1+l_\alpha} = \frac{p+2}{p+2+pF_{p,m}(\alpha)}$;

$$\xi_r(l_\alpha) = I_{h_\alpha} \left(\frac{m}{2}; \frac{p+2+2r}{2} \right) \quad (5.4)$$

is the incomplete beta function evaluated at h_α ;

$$\xi_r(\nu) = \frac{\Gamma\left(\frac{\nu}{2} + r\right)}{r! \Gamma\left(\frac{\nu}{2}\right)} \quad (5.5)$$

$$\xi_r(\nu, \Delta^*) = \frac{(\Delta^*/\nu - 2)^r}{[1 + \Delta^*/(\nu - 2)]^{\nu/2+r}} \quad (5.6)$$

with $\Delta^* = \frac{\nu - 2}{\nu} \Delta$ such that $\Delta = \frac{\boldsymbol{\delta}'\boldsymbol{\delta}}{\tau^2}$; and $\boldsymbol{\delta} = (\boldsymbol{\mu} - \boldsymbol{\mu}_0)$.

Proof. By the definition, the bias of the PTE is

$$B_2(\hat{\boldsymbol{\mu}}^{pt}; \boldsymbol{\mu}) = E_\tau \{E[\hat{\boldsymbol{\mu}}^{pt} - \boldsymbol{\mu} | \tau]\} \quad (5.7)$$

where

$$E[(\hat{\boldsymbol{\mu}}^{pt} - \boldsymbol{\mu}) | \tau] = -E[(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)I(T^2 \leq T_\alpha^2) | \tau]. \quad (5.8)$$

To evaluate the last expectation, let

$$\mathbf{y} = \frac{\sqrt{n}}{\tau}(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0). \quad (5.9)$$

Then for a given value of τ , $\mathbf{y} \sim N_p\left(\frac{\sqrt{n}}{\tau}\boldsymbol{\delta}; I_p\right)$, and the statistic $\mathbf{y}'\mathbf{y}$ follows a noncentral chi-squared distribution with p d.f. and non-centrality parameter $\Delta_\tau = \tau^{-2}\boldsymbol{\delta}'\boldsymbol{\delta}$. Therefore, the T^2 statistic defined in Section 3 can be expressed as

$$T^2 = \frac{\mathbf{y}'\mathbf{y}}{\chi_m^2} \quad (5.10)$$

and hence the statistic $\frac{p}{m}T^2$ follows a non-central F distribution with p and m d.f. and non-centrality parameter Δ_τ . So, conditionally on τ , we get

$$\begin{aligned} B_2(\hat{\boldsymbol{\mu}}^{pt}; \boldsymbol{\mu} | \tau) &= -E \left[\frac{\tau}{\sqrt{n}} \mathbf{y}' I \left(\frac{\mathbf{y}'\mathbf{y}}{\chi_m^2} \leq \frac{p}{m} F_{p,m}(\alpha) \right) \right] \\ &= -\boldsymbol{\delta} G_{p+2,m}(l_\alpha; \Delta_\tau) \end{aligned} \quad (5.11)$$

where $l_\alpha = \frac{p}{p+2} F_{p,m}(\alpha)$ and $G_{p+2,m}(l_\alpha; \Delta_\tau)$ is the distribution function of a non-central $F_{p+2,m}(\Delta_\tau)$ variable evaluated at l_α . Finally, taking expectation on the r.h.s. of (5.11) with respect to the distribution of τ , the bias of $\hat{\boldsymbol{\mu}}^{pt}$ is obtained as

$$B_2(\hat{\boldsymbol{\mu}}^{pt}; \boldsymbol{\mu}) = -\boldsymbol{\delta} G_{p+2,m}^{(2)}(l_\alpha; \Delta^*). \quad (5.12)$$

Under the H_0 , $\boldsymbol{\delta}$ is \mathbf{O} , and hence the PTE is an unbiased estimator of $\boldsymbol{\mu}$. However, under the alternative hypothesis the PTE is biased. Moreover, the size of the bias depends on $\boldsymbol{\delta}$, the departure of $\boldsymbol{\mu}_0$ from the true value of $\boldsymbol{\mu}$. It is observed that the bias of the RE, $\hat{\boldsymbol{\mu}}$ is $\boldsymbol{\delta}$, which is unbounded. But the bias of the PTE is less than RE since $0 \leq G_{p+2,m}^{(2)}(l_\alpha; \Delta^*) \leq 1$. Also note that as $\nu \rightarrow \infty$, $G_{p+2,m}^{(2)}(l_\alpha; \Delta^*)$ approaches to the c.d.f. of the central F variable,

$F_{p+2,m}(l_\alpha; 0)$ when the H_0 is true. The behavior of the bias of the PTE with respect to the change in the departure of $\boldsymbol{\mu}_0$ from $\boldsymbol{\mu}$ (i.e. the value of $\boldsymbol{\delta}$) can be investigated by studying the quadratic bias function,

$$QB_2 = \boldsymbol{\delta}'\boldsymbol{\delta}[G_{p+2,m}^{(2)}(l_\alpha; \Delta^*)]^2 \quad (5.13)$$

against different values of $\boldsymbol{\delta}'\boldsymbol{\delta} = \Delta$ for different given values of ν and α .

Theorem 5.2. The bias of the SE, $\hat{\boldsymbol{\mu}}^s$ is given by

$$B_3(\hat{\boldsymbol{\mu}}^s; \boldsymbol{\mu}) = -\boldsymbol{\delta}kmE^{(2)}[\chi_{p+2}^{-2}(\Delta^*)] \quad (5.14)$$

where

$$E^{(2)}[\chi_{p+2}^{-2}(\Delta^*)] = \sum_{r=0}^{\infty} \frac{1}{p+2r} \xi_r(\nu) \xi_r(\nu, \Delta^*) \quad (5.15)$$

in which the last two factors are as stated in equations (5.5) and (5.6) respectively.

Proof. By the definition, the bias of the SE is

$$B_3(\hat{\boldsymbol{\mu}}^s; \boldsymbol{\mu}) = E_\tau\{E[\hat{\boldsymbol{\mu}}^s - \boldsymbol{\mu}|\tau]\} \quad (5.16)$$

where

$$E[(\hat{\boldsymbol{\mu}}^s - \boldsymbol{\mu})|\tau] = -E[kT^{-2}(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)|\tau]. \quad (5.17)$$

Applying the substitution in (5.9), the right hand side of (5.17) becomes

$$-E\left[k\frac{\chi_m^2}{\mathbf{y}'\mathbf{y}}\mathbf{y}\frac{\tau}{\sqrt{n}}\right] = -\boldsymbol{\delta}kmE[\chi_{p+2}^{-2}(\Delta_\tau)]. \quad (5.18)$$

Then taking expectation on the right hand side of (5.18) with respect to the $IG(\nu, \sigma)$ distribution, we get the required expression in the notation of (5.14) and (5.15).

The bias of the Stein-type SE also depends on $\boldsymbol{\delta}$. Note that for large values of n and reasonably small p , the value of $km = \frac{(p-2)m}{m+2}$ is close to $(p-2)$. The quadratic bias of the SE,

$$QB_3 = \Delta k^2 m^2 \{E^{(2)}[\chi_{p+2}^{-2}(\Delta^*)]\}^2 \quad (5.19)$$

can be plotted against Δ to study the changes in QB_3 with the change in the value of Δ .

Theorem 5.3. The bias of the PRSE, $\hat{\mu}^{s+}$ is given by

$$B_4(\hat{\mu}^{s+}; \mu) = -\delta km E^{(2)}[\chi_{p+2}^{-2}(\Delta^*)] - \delta G_{p+2,m}^{(2)}(q_2; \Delta^*) \\ + \delta km E^{(2)}[\chi_{p+2}^{-2}(\Delta^*) I(F_{p+2,m}(\Delta^*) \leq q_2)] \quad (5.20)$$

where

$$E^{(2)}[\chi_{p+2}^{-2}(\Delta^*) I(F_{p+2,m}(\Delta^*) \leq q_2)] = \sum_{r=0}^{\infty} \frac{\Gamma\left(\frac{p+2+m}{2}\right) \xi_r(u_2) \xi_r(\nu) \xi_r(\nu, \Delta^*)}{\Gamma\left(\frac{p+2r}{2}\right) \Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{p+2+2r}{2}\right)} \quad (5.21)$$

in which $q_2 = \frac{m(p-2)}{(p+2)(m+2)}$, $u_2 = \frac{1}{1+q_2} = \frac{(p+2)(m+2)}{(p+2)(m+2) + m(p-2)}$ and $\xi_r(u_2)$ has the same form as $\xi_r(\ell_\alpha)$ in (5.4) but does not depend on α .

Proof. By the definition, the bias of the PRSE, $\hat{\mu}^{s+}$ is

$$B_4(\hat{\mu}^{s+}; \mu) = E_\tau\{E[\hat{\mu}^{s+} - \mu]|\tau\}. \quad (5.22)$$

where

$$E[(\hat{\mu}^{s+} - \mu)|\tau] = E[\{-(\tilde{\mu} - \mu_0)kT^{-2} - (\tilde{\mu} - \mu_0)I(T^2 \leq k) \\ + (\tilde{\mu} - \mu_0)kT^{-2}I(T^2 \leq k)\}|\tau]. \quad (5.23)$$

Now applying the transformation in (5.9), we get

$$E[\{(\tilde{\mu} - \mu_0)kT^{-2}\}|\tau] = \delta km E[\chi_{p+2}^{-2}(\Delta_\tau)]; \quad (5.24)$$

$$E[\{(\tilde{\mu} - \mu_0)I(T^2 \leq k)\}|\tau] = \left[\left\{ \frac{\tau}{\sqrt{n}} I\left(\frac{\mathbf{y}'\mathbf{y}}{\chi_m^2} \leq k\right) \right\} |\tau \right] \\ = \delta E[\{I(F_{p+2,m}(\Delta_\tau) \leq q_2)\}|\tau] \\ = \delta G_{p+2,m}(q_2; \Delta_\tau); \quad (5.25)$$

$$E[\{(\tilde{\mu} - \mu_0)kT^{-2}I(T^2 \leq k)\}|\tau] = \delta km E[\chi_{p+2}^{-2}(\Delta_\tau) I(F_{p+2,m}(\Delta_\tau) \leq q_2)]. \quad (5.26)$$

Thus, for a given τ , we have

$$E[(\hat{\mu}^{s+} - \mu)|\tau] = -\delta km E[\chi_{p+2}^{-2}(\Delta_\tau)] - \delta G_{p+2,m}(q_2; \Delta_\tau) \\ + \delta km E[\chi_{p+2}^{-2}(\Delta_\tau) I(F_{p+2,m}(\Delta_\tau) \leq q_2)]. \quad (5.27)$$

The final expression in (5.20) is obtained by taking expectation on the r.h.s. of (5.27) with respect to the distribution of τ .

The first term in (5.27) is the same as the bias of the SE. Like PTE and SE, the PRSE is unbiased for $\boldsymbol{\mu}$ under the null hypothesis but biased under the alternative hypothesis. Similarly, the size of the bias is a function of $\boldsymbol{\delta}$, the magnitude of the difference between the value of the mean vector, $\boldsymbol{\mu}$ specified by the H_0 and its true value.

6 EVALUATION OF RISKS

Consider the *quadratic error loss function* of the form:

$$L(\boldsymbol{\theta}^*; \boldsymbol{\theta}) = n(\boldsymbol{\theta}^* - \boldsymbol{\theta})'(\boldsymbol{\theta}^* - \boldsymbol{\theta}) \quad (6.1)$$

where $\boldsymbol{\theta}^*$ is any estimator of $\boldsymbol{\theta}$. Then the quadratic risk function associated with (6.1) is defined as

$$R(\boldsymbol{\theta}^*; \boldsymbol{\theta}) = E[n(\boldsymbol{\theta}^* - \boldsymbol{\theta})'(\boldsymbol{\theta}^* - \boldsymbol{\theta})]. \quad (6.2)$$

Using the above definition we evaluate the quadratic risk of the four different estimators under study.

Theorem 6.1. The quadratic risk of the UE, $\tilde{\boldsymbol{\mu}}$ in estimating the mean vector $\boldsymbol{\mu}$ of a p -dimensional Student-t model with ν d.f. is

$$R_1(\tilde{\boldsymbol{\mu}}; \boldsymbol{\mu}) = \frac{\nu\sigma^2}{\nu-2} p. \quad (6.3)$$

Proof. The risk of $\tilde{\boldsymbol{\mu}}$ with respect to the loss function in (6.1) is

$$R_1(\tilde{\boldsymbol{\mu}}; \boldsymbol{\mu}) = E_\tau[E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})|\tau\}], \quad (6.4)$$

since for a given τ , $\frac{\sqrt{n}}{\tau}(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}) \sim N_p(\mathbf{O}; I_p)$,

$$E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})|\tau\} = E\{\tau^2 \chi_p^2\} = \tau^2 p.$$

Then $R_1(\tilde{\boldsymbol{\mu}}; \boldsymbol{\mu}) = E_\tau[\tau^2 p] = p \frac{\nu\sigma^2}{\nu-2}$. Note that $E_\tau(\tau^2) = \int_{\tau=0}^{\infty} \tau^2 IG(\nu, \sigma) d\tau = \frac{\nu\sigma^2}{\nu-2}$. As $\nu \rightarrow \infty$, the risk of $\tilde{\boldsymbol{\mu}}$ tends to p , a well-known result for the normal model.

Theorem 6.2. The quadratic risk of the PTE, $\hat{\boldsymbol{\mu}}^{pt}$ in estimating the mean vector $\boldsymbol{\mu}$ of a p -dimensional Student-t model with ν d.f. is

$$R_2(\hat{\boldsymbol{\mu}}^{pt}; \boldsymbol{\mu}) = \frac{\nu\sigma^2}{\nu-2}p \left[1 - G_{p+2,m}^{(2)}(l_\alpha; \Delta^*) \right] + \Delta \left[2nG_{p+2,m}^{(2)}(l_\alpha; \Delta^*) - G_{p+4,m}^{(4)}(l_\alpha^*; \Delta^*) \right], \quad (6.5)$$

where

$$G_{p+2,m}^{(4)}(l_\alpha^*; \Delta^*) = \sum_{r=0}^{\infty} \frac{\Gamma\left(\frac{p+4+m}{2} + r\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{p+4}{2} + r\right)} \xi_r(h_\alpha^*) \xi_r(\nu) \xi_r(\nu, \Delta^*) \quad (6.6)$$

in which $l_\alpha^* = \frac{p}{p+4}F_{p,m}(\alpha)$ and $h_\alpha^* = \frac{1}{1+l_\alpha^*} = \frac{p+4}{p+4+pF_{p,m}(\alpha)}$.

Recall that $\xi_r(\cdot)$'s have been defined in (5.4) – (5.6).

Proof. The risk of $\hat{\boldsymbol{\mu}}^{pt}$ with respect to the loss function in (6.1) is

$$R_2(\hat{\boldsymbol{\mu}}^{pt}; \boldsymbol{\mu}) = E_\tau[E\{n(\hat{\boldsymbol{\mu}}^{pt} - \boldsymbol{\mu})'(\hat{\boldsymbol{\mu}}^{pt} - \boldsymbol{\mu})\}|\tau] \quad (6.7)$$

where

$$\begin{aligned} E\{n(\hat{\boldsymbol{\mu}}^{pt} - \boldsymbol{\mu})'(\hat{\boldsymbol{\mu}}^{pt} - \boldsymbol{\mu})|\tau\} &= E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})|\tau\} \\ &+ E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})I(T^2 \leq T_\alpha^2)|\tau\} \\ &- 2E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})I(T^2 \leq T_\alpha^2)|\tau\}. \end{aligned} \quad (6.8)$$

The first term in (6.8) is $\tau^2 p$, the risk of $\tilde{\boldsymbol{\mu}}$, for a given τ . To evaluate the second and third terms, note that $\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu} = (\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0) - (\boldsymbol{\mu} - \boldsymbol{\mu}_0)$. Then applying the transformation in (5.9) we obtain

$$\begin{aligned} E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)I(T^2 \leq T_\alpha^2)|\tau\} &= \tau^2 E\left\{\mathbf{y}'\mathbf{y}I\left(\frac{\mathbf{y}'\mathbf{y}}{\chi_m^2} \leq l_\alpha\right)|\tau\right\} \\ &= \tau^2 \left[pG_{p+2,m}(l_\alpha; \Delta_\tau) + \frac{\Delta}{\tau^2}G_{p+4,m}(l_\alpha^*; \Delta_\tau) \right] \end{aligned} \quad (6.9)$$

where $\Delta_\tau = \frac{\boldsymbol{\delta}'\boldsymbol{\delta}}{\tau^2}$; and

$$E\{n(\boldsymbol{\mu} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)I(T^2 \leq T_\alpha^2)|\tau\} = n\Delta G_{p+2,m}(l_\alpha; \Delta_\tau). \quad (6.10)$$

Now collecting the terms, and simplifying, the right hand side of (6.8) becomes

$$\tau^2 p - \tau^2 p G_{p+2,m}(l_\alpha; \Delta_\tau) + \Delta [2n G_{p+2,m}(l_\alpha; \Delta_\tau) - G_{p+4,m}(l_\alpha^*; \Delta_\tau)]. \quad (6.11)$$

Finally taking expectation on (6.11) with respect to the $IG(\nu, \sigma)$ distribution we obtain the risk expression in (6.5).

Theorem 6.3. The quadratic risk of the SE, $\hat{\boldsymbol{\mu}}^s$ in estimating the mean vector $\boldsymbol{\mu}$ of a p -dimensional Student-t model with ν d.f. is

$$R_3(\hat{\boldsymbol{\mu}}^s; \boldsymbol{\mu}) = \frac{\nu\sigma^2}{\nu-2} \left[p - 2km + k^2m(m+2)E^{(0)}\{\chi_p^{-2}(\Delta^*)\} \right] + 2\Delta kmnE^{(2)}\{\chi_{p+2}^{-2}(\Delta^*)\} \quad (6.12)$$

where $E^{(0)}\{\chi_p^{-2}(\Delta^*)\} = \sum_{r=0}^{\infty} \frac{1}{p+2r} \xi_r(\nu) \xi_r(\nu, \Delta^*)$.

Proof. The risk of $\hat{\boldsymbol{\mu}}^s$ with respect to the loss function in (6.1) is

$$R_3(\hat{\boldsymbol{\mu}}^s; \boldsymbol{\mu}) = E_\tau[E\{n(\hat{\boldsymbol{\mu}}^s - \boldsymbol{\mu})'(\hat{\boldsymbol{\mu}}^s - \boldsymbol{\mu})\}|\tau] \quad (6.13)$$

where

$$\begin{aligned} E\{n(\hat{\boldsymbol{\mu}}^s - \boldsymbol{\mu})'(\hat{\boldsymbol{\mu}}^s - \boldsymbol{\mu})|\tau\} &= E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})|\tau\} \\ &+ k^2 E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)T^{-4}|\tau\} \\ &- 2k E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)T^{-2}|\tau\}. \end{aligned} \quad (6.14)$$

The first term in (6.14) is $\tau^2 p$, the risk of $\tilde{\boldsymbol{\mu}}$ for a given value of τ . For the second term, by applying the transformation in (5.9) we have,

$$\begin{aligned} E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)T^{-4}|\tau\} &= E\left\{\frac{\chi_m^4}{(\mathbf{y}'\mathbf{y})^2}(\mathbf{y}'\mathbf{y})\tau^2|\tau\right\} \\ &= \tau^2 m(m+2)E\{\chi_p^{-2}(\Delta_\tau)\}. \end{aligned} \quad (6.15)$$

In the third term, first we note that $\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu} = (\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0) - (\boldsymbol{\mu} - \boldsymbol{\mu}_0)$, and then using (5.9) we get,

$$E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)T^{-2}|\tau^2\} = E\left\{\frac{\chi_m^2}{(\mathbf{y}'\mathbf{y})} \times (\mathbf{y}'\mathbf{y})\tau^2\right\} = \tau^2 m \quad (6.16)$$

and

$$E\{n(\boldsymbol{\mu} - \boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)T^{-2}|\tau\} = \Delta mnE\{\chi_{p+2}^{-2}(\Delta_\tau)\}. \quad (6.17)$$

Now collecting all the terms and simplifying, we have

$$\begin{aligned}
E\{n(\hat{\boldsymbol{\mu}}^s - \boldsymbol{\mu})' (\hat{\boldsymbol{\mu}}^s - \boldsymbol{\mu})|\tau\} &= \tau^2 p + k^2 \tau^2 m(m+2)E\{\chi_p^{-2}(\Delta_\tau)\} \\
&\quad - 2k \left[\tau^2 m - nm \Delta E\{\chi_{p+2}^{-2}(\Delta_\tau)\} \right] \\
&= \tau^2 \left[p - 2km + k^2 m(m+2)E\{\chi_p^{-2}(\Delta_\tau)\} \right] \\
&\quad + 2knm \Delta E\{\chi_{p+2}^{-2}(\Delta_\tau)\}. \tag{6.18}
\end{aligned}$$

The expectation on the r.h.s. of (6.18) with respect to the $IG(\nu, \sigma)$ distribution yields the risk expression in (6.12).

Theorem 6.4. The quadratic risk of the PRSE, $\hat{\boldsymbol{\mu}}^{s+}$ in estimating the mean vector $\boldsymbol{\mu}$ of a p -dimensional Student-t model with ν d.f. is

$$\begin{aligned}
R_4(\hat{\boldsymbol{\mu}}^{s+}; \boldsymbol{\mu}) &= \frac{\nu\sigma^2}{\nu-2} p \left[1 - G_{p+2,m}^{(2)}(q_2; \Delta^*) \right] - 2 \frac{\nu\sigma^2}{\nu-2} km \left[1 - G_{p,m}^{(0)}(q; \Delta^*) \right] \\
&\quad + \frac{\nu\sigma^2}{\nu-2} k^2 m(m+2) \left[E^{(0)}\{\chi_p^{-2}(\Delta^*)\} \right. \\
&\quad \quad \left. - E^{(2)}\{\chi_{p+2}^{-2}(\Delta^*) I(F_{p+2,m}(\Delta^*) \leq q_2)\} \right] \\
&\quad - \Delta n \left[G_{p+4,m}^{(4)}(q_4; \Delta^*) - 2G_{p+2,m}^{(2)}(q_2; \Delta^*) \right] \\
&\quad + \Delta 2kmn \left[E^{(2)}\{\chi_{p+2}^{-2}(\Delta^*)\} - E^{(2)}\{\chi_{p+2}^{-2}(\Delta^*) I(F_{p+2,m}(\Delta^*))\} \right] \tag{6.19}
\end{aligned}$$

where $q_4 = \frac{m(p-2)}{(p+4)(m+2)}$ and $q_2 = \frac{m(p-2)}{(p+2)(m+2)}$.

Proof. The risk of $\hat{\boldsymbol{\mu}}^{s+}$ with respect to the loss function in (6.1) is

$$R_4(\hat{\boldsymbol{\mu}}^{s+}; \boldsymbol{\mu}) = E_\tau[E\{\hat{\boldsymbol{\mu}}^{s+} - \boldsymbol{\mu}\}'(\hat{\boldsymbol{\mu}}^{s+} - \boldsymbol{\mu})|\tau] \tag{6.20}$$

where

$$\begin{aligned}
E\{n(\hat{\boldsymbol{\mu}}^{s+} - \boldsymbol{\mu})'(\hat{\boldsymbol{\mu}}^{s+} - \boldsymbol{\mu})|\tau\} &= E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})|\tau\} \\
&\quad + E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)k^2 T^{-4}|\tau\} + E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)I(T^2 \leq k)|\tau\} \\
&\quad + E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)k^2 T^{-4}I(T^2 \leq k)|\tau\} \\
&\quad - 2E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)kT^{-2}|\tau\} - 2E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)I(T^2 \leq k)|\tau\} \\
&\quad + 2E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)kT^{-2}I(T^2 \leq k)|\tau\}
\end{aligned}$$

$$\begin{aligned}
& +2E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)kT^{-2}I(T^2 \leq k)|\tau\} \\
& -2E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)k^2T^{-4}I(T^2 \leq k)|\tau\} \\
& -2E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)kT^{-2}I(T^2 \leq k)|\tau\}. \tag{6.21}
\end{aligned}$$

As before, noting that $\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu} = (\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0) - (\boldsymbol{\mu} - \boldsymbol{\mu}_0)$, and applying the transformation in (5.9) the terms on the r.h.s. of (6.21) can be evaluated by using the fact that

$$\begin{aligned}
& E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})|\tau\} = \tau^2 p \\
& E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)k^2T^{-4}|\tau\} = \tau^2 k^2 m(m+2)E\{\chi_p^{-2}(\Delta_\tau)\} \\
& E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)I(T^2 \leq k)|\tau\} = \tau^2 p G_{p+2,m}(q_2; \Delta_\tau) \\
& \quad + \Delta n G_{p+4,m}(q_4; \Delta_\tau) \\
& E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)k^2T^{-4}I(T^2 \leq k)|\tau\} \\
& \quad = \tau^2 k^2 m(m+2)E\{\chi_{p+2}^{-2}(\Delta_\tau)I(F_{p+2,m}(\Delta_\tau) \leq q_2)\} \\
& E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)kT^{-2}|\tau\} = E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)kT^{-2}|\tau\} \\
& E\{n(\boldsymbol{\mu} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)kT^{-2}|\tau\} = -\tau km + nkm\Delta E\{\chi_{p+2}^{-2}(\Delta_\tau)\} \\
& E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)I(T^2 \leq k)|\tau\} = E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)I(T^2 \leq k)|\tau\} \\
& \quad - E\{n(\boldsymbol{\mu} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)I(T^2 \leq k)|\tau\} \\
& \quad = [\tau^2 p G_{p+2,m}(q_2; \Delta_\tau) + n\Delta G_{p+4,m}(q_4; \Delta_\tau)] - n\Delta G_{p+2,m}(q_2; \Delta_\tau) \\
& E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)kT^{-2}I(T^2 \leq k)|\tau\} = E\{nk(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0) \\
& \quad \times T^{-2}I(T^2 \leq k)|\tau\} - E\{nk(\boldsymbol{\mu} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)T^{-2}I(T^2 \leq k)|\tau\} \\
& \quad = \tau^2 km G_{p,m}(q; \Delta_\tau) - nkm\Delta E\{\chi_{p+2}^{-2}(\Delta_\tau)I(F_{p+2,m}(\Delta_\tau) \leq q_2)\} \\
& E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)kT^{-2}I(T^2 \leq k)|\tau\} = \tau^2 km G_{p,m}(q; \Delta_\tau) \\
& E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)k^2T^{-4}I(T^2 \leq k)|\tau\} \\
& \quad = \tau^2 k^2 m(m+2)E\{\chi_{p+2}^{-2}(\Delta_\tau)I(F_{p+2,m}(\Delta_\tau) \leq q_2)\} \\
& E\{n(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)kT^{-2}I(T^2 \leq k)|\tau\} = \tau^2 km G_{p,m}(q; \Delta_\tau). \tag{6.22}
\end{aligned}$$

Now collecting all the terms on the r.h.s. of (6.22) and plugging into the right hand side of (6.21) we get, after simplification and rearrangement,

$$\begin{aligned}
& E\{n(\hat{\boldsymbol{\mu}}^{s+} - \boldsymbol{\mu})'(\hat{\boldsymbol{\mu}}^{s+} - \boldsymbol{\mu})|\tau\} = \\
& \quad \tau^2 p [1 - G_{p+2,m}(q_2; \Delta_\tau)] - 2\tau^2 km [1 - G_{p,m}(q; \Delta_\tau)] \\
& \quad + \tau^2 km(m+2) [E\{\chi_p^{-2}(\Delta_\tau)\} - E\{\chi_{p+2}^{-2}(\Delta_\tau)I(F_{p+2,m}(\Delta_\tau) \leq q_2)\}]
\end{aligned}$$

$$\begin{aligned}
& - \Delta n [G_{p+4,m}(q_4; \Delta_\tau) - 2G_{p+2,m}(q_2; \Delta_\tau)] \\
& + \Delta 2kmn \left[E\{\chi_{p+2}^{-2}(\Delta_\tau)\} - E\{\chi_{p+2}^{-2}(\Delta_\tau)I(F_{p+2,m}(\Delta_\tau) \leq q_2)\} \right]. \quad (6.23)
\end{aligned}$$

Then completing the expectations on the r.h.s. of (6.23) with respect to the $IG(\nu, \sigma)$ distribution the final risk expression in (6.19) is obtained.

For computations, note that

$$E \left[\chi_a^{-b}(\Delta^*) \right] = \sum_{r=0}^{\infty} \frac{\xi_r(\nu) \xi_r(\nu, \Delta^*)}{(a-2+2r)(a-4+2r) \dots \{a-2(b-1)+2r\}}$$

and

$$E^{(2)} \left[\chi_{p+2}^{-2}(\Delta^*) I(F_{p+2,m}(\Delta^*) \leq q_2) \right] = \sum_{r=0}^{\infty} \frac{\Gamma\left(\frac{p+2+m}{2}\right) \xi_r(u_2) \xi_r(\nu) \xi_r(\nu, \Delta^*)}{\Gamma\left(\frac{p+2r}{2}\right) \Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{p+2+2r}{2}\right)}.$$

7 ANALYSIS OF RISKS

The *unrestricted estimator*, $\tilde{\boldsymbol{\mu}}$ has a constant risk that depends on the d.f. and dimension of the Student-t model. As expected, this risk of the mle of $\boldsymbol{\mu}$ for the Student-t model approaches the risk of the mle of the normal model as ν grows large. However, for smaller values of ν , the risk of $\tilde{\boldsymbol{\mu}}$ for the normal model is smaller than that of the Student-t model. So, there could be a misleading risk figure, which may appear to be much smaller than it should be, if a Student-t model is misspecified as a normal model.

The risk of the *preliminary test estimator*, $\hat{\boldsymbol{\mu}}^{pt}$ is a function of the choice of α . For a given ν and a particular choice of α , the risk of $\hat{\boldsymbol{\mu}}^{pt}$ is smaller than that of $\tilde{\boldsymbol{\mu}}$ at $\Delta^* = 0$. But it grows larger as Δ^* increases, after crossing the risk curve of $\tilde{\boldsymbol{\mu}}$ from below and then slowly approaches to $R_1(\tilde{\boldsymbol{\mu}}; \boldsymbol{\mu})$ as $\Delta^* \rightarrow \infty$. The behavior of $R_2(\hat{\boldsymbol{\mu}}^{pt}; \boldsymbol{\mu})$ changes significantly as the value of α changes. The difference of risks of the UE and PTE is obtained as

$$D_{12} = \frac{\nu\sigma^2}{\nu-2} p G_{p+2,m}^{(2)}(\ell_\alpha; \Delta^*) - \Delta \left[2n G_{p+2,m}^{(2)}(\ell_\alpha; \Delta^*) - G_{p+4,m}^{(4)}(\ell_\alpha^*; \Delta^*) \right]. \quad (7.1)$$

Hence the PTE over performs the UE whenever $D_{12} \geq 0$, that is,

$$\Delta^* \leq \frac{G_{p+2,m}^{(2)}(\ell_\alpha; \Delta^*)}{2n G_{p+2,m}^{(2)}(\ell_\alpha; \Delta^*) - G_{p+4,m}^{(4)}(\ell_\alpha^*; \Delta^*)}. \quad (7.2)$$

If the inequality does not hold, then the UE dominates the PTE. This is true for all ν and α .

It is well known that the risk of the Stein-type *shrinkage estimator*, $\hat{\boldsymbol{\mu}}^s$ is smaller than that of $\tilde{\boldsymbol{\mu}}$ for the multivariate ($p > 3$) normal model. This is also true for the Student-t model with any value of ν . It also dominates the PTE for any choice of α under the H_0 and for all ν . Near $\Delta^* = 0$ the risk of $\hat{\boldsymbol{\mu}}^s$ is the smallest with compared to that of $\tilde{\boldsymbol{\mu}}$. But the risk curve of $R_3(\hat{\boldsymbol{\mu}}^s; \boldsymbol{\mu})$ approaches to that of $R_1(\tilde{\boldsymbol{\mu}}; \boldsymbol{\mu})$ from below as Δ^* grows larger, but does not exceed the risk curve of $\tilde{\boldsymbol{\mu}}$. Thus $\hat{\boldsymbol{\mu}}^s$ always dominates $\tilde{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\mu}}^{pt}$, regardless of the value of ν , α and Δ^* . However, for larger values of Δ^* the risk of $\hat{\boldsymbol{\mu}}^s$ equals that of $\tilde{\boldsymbol{\mu}}$, after meeting at some value of Δ^* , for all ν . The value of Δ^* at which $R_1(\tilde{\boldsymbol{\mu}}; \boldsymbol{\mu})$ and $R_3(\hat{\boldsymbol{\mu}}^s; \boldsymbol{\mu})$ meet decreases as the value of ν increases. Thus for a Student-t model with smaller value of the shape parameter they meet at a slower pace than for the normal model.

Under H_0 , the *positive-rule shrinkage estimator* has the smallest risk than any of the other three estimators. It dominates $\hat{\boldsymbol{\mu}}^s$, which dominates $\hat{\boldsymbol{\mu}}^{pt}$ and $\tilde{\boldsymbol{\mu}}$. $R_4(\hat{\boldsymbol{\mu}}^{s+}; \boldsymbol{\mu})$ approaches $R_1(\tilde{\boldsymbol{\mu}}; \boldsymbol{\mu})$ from below as Δ^* increases, and both the risk functions coincide as $\Delta^* \rightarrow \infty$ after meeting one another at some value of Δ^* . Both $R_3(\hat{\boldsymbol{\mu}}^s; \boldsymbol{\mu})$ and $R_4(\hat{\boldsymbol{\mu}}^{s+}; \boldsymbol{\mu})$ meet $R_1(\tilde{\boldsymbol{\mu}}; \boldsymbol{\mu})$ from below; but the former meets for a smaller value of Δ^* than the later. Therefore, $\hat{\boldsymbol{\mu}}^{s+}$ not only dominates $\hat{\boldsymbol{\mu}}^s$, but also dominates for a wider (range) of values of Δ^* .

Based on the foregoing discussion under H_0 , the following dominance picture of the estimators emerges:

$$\hat{\boldsymbol{\mu}}^{s+} \succ \hat{\boldsymbol{\mu}}^s \succ \hat{\boldsymbol{\mu}}^{pt} \succ \tilde{\boldsymbol{\mu}} \quad (7.3)$$

where the symbol ‘ \succ ’ stands for domination. However, under the alternative, the above dominance picture will change as $\hat{\boldsymbol{\mu}}^{pt}$ and $\tilde{\boldsymbol{\mu}}$ will interchange the positions starting from some value of Δ^* .

ACKNOWLEDGEMENTS

The author is grateful to two referees for suggesting some important corrections and improvements on an earlier version of the paper. He also extends sincere gratitude to Professor A.K.Md.E. Saleh for his kind advice and valuable discussions on the problem of improved estimation. The author thankfully acknowledges the excellent research facilities provided by the King Fahd University of Petroleum and Minerals, Saudi Arabia.

BIBLIOGRAPHY

- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *Second International Symposium on Information Theory*. B.N. Petrov and F. Csaki eds. Akademiai, Budapest, 267-181.
- Anderson, T.W. (1993). Nonnormal multivariate distributions: Inference based on elliptically contoured distributions. In *Multivariate Analysis: Future Directions*. ed by C.R. Rao. North-Holland, Amsterdam, 1-24.
- Bancroft, T.A. (1944). On biases in estimation due to the use of the preliminary tests of significance. *Ann. Math. Statist.*, Vol. 15, 190-204.
- Celler, D., Fourdrunier, D and Strawderman, W.E. (1995). Shrinkage positive rule estimators for spherically symmetric distributions. *Journal of Multivariate Analysis*, **53**, 194-209.
- Chmielewski, M.A. (1981). Elliptically symmetric distributions. *International Statistical Review*, Vol. 49, 67-74.
- Chou, J.P. and Strawderman, W.E. (1990). Minimax estimation of means of multivariate normal mixtures. *Journal of Multivariate Analysis*, **35**, 141-151.
- Fang, K.T. and Anderson, T.W. (1990). *Statistical Inference in Elliptically Contoured and Related Distributions*. Allerton Press Inc., New York.
- Fisher, R.A. (1956). *Statistical Methods in Scientific Inference*. Oliver and Boyd, London.
- Fraser, D.A.S. (1979). *Inference and Linear Models*, McGraw-Hill, New York.
- Giles, J.A. (1991). Pre-testing for linear restrictions model with spherically symmetric disturbances. *Journal of Econometrics*, Vol. 50, 377-398.
- Gupta, A.K. and Varga, T. (1993). *Elliptically Contoured Models in Statistics*. Kluwer Academic Publishers, London.
- Han, C.P. and Bancroft, T.A. (1968). On pooling means when variance is unknown. *Jou. Amer. Statist. Assoc.*, Vol. 63, 333-1342.

- James, W. and Stein, C. (1961). Estimation with quadratic loss. Proceedings of the Fourth Berkeley Symposium on Math. Statist. and Probability, University of California Press, Berkeley, Vol. 1, 361–379.
- Judge, G.G. and Bock, M.E. (1978). The Statistical Implications of Pre-test and Stein-rule Estimators in Econometrics. North-Holland, New York.
- Khan, S. (1997). Likelihood based inference on the mean vectors of two multivariate Student-t populations. *Far East Journal of Theoretical Statistics*. Vol. **1**, 77-91.
- Khan, S. and Haq, M.S. (1990). Prediction distribution for the linear model with multivariate Student-t error. *Communications in Statistics: Theory & Methods*, Vol. 19, 4705–4712.
- Khan, S. and Saleh, A.K.Md.E. (1998). Comparison of estimators of means based on p -samples from multivariate Student-t population. *Communications in Statistics: Theory & Methods*. **27**, 193-210.
- Khan, S. and Saleh, A.K.Md.E. (1997). Shrinkage pre-test estimator of the intercept parameter for a regression model with multivariate Student-t errors. *Biometrical Journal*. Vol.**39**, 131-147.
- Maatta, J.M. and Casella, G. (1990). Developments in decision-theoretic variance estimation, *Statistical Sciences*, Vol. 5, 90–101.
- Prucha I.R. and Kelegian, H.H. (1984). The structure of simultaneous equation estimators: A generalization towards non-normal disturbances, *Econometrica*, Vol. 52, 721–736.
- Saleh, A.K.Md.E. and Sen, P.K. (1985). Shrinkage least squares estimation in a general multivariate linear model. Proceedings of the Fifth Pannonian Symposium on Mathematical Statistics, 307–325.
- Saleh, A.K.Md.E. and Sen, P.K. (1978). Nonparametric estimation of location parameter after a preliminary test on regression. *Annals of Statistics*, Vol 6, 154–168.
- Sclove, S.L., Morris, C. and Radhakrishnan, R. (1972). Non-optimality of preliminary-test estimators for the mean of a multivariate normal distribution. *Ann. Math. Statist.*, Vol. 43, 1481–1490.
- Singh, R.S. (1988). Estimation of error variance in linear regression models with errors having multivariate Student-t distribution with unknown degrees of freedom. *Economics Letters*, Vol. 27, 47-53.
- Singh, R.S. (1991). James-Stein rule estimators in linear regression models with multivariate-t errors. *Australian Journal of Statistics*, Vol. 33, 145–158.
- Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution, Proceedings of the Third Berkeley Symposium on Math. Statist. and Probability, University of California Press, Berkeley, Vol. 1, 197–206.
- Stein, C. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.*, Vol. 9, 1135–1151.

- Tabatabaey, S.M.M. (1995). Preliminary Test Approach Estimation: Regression Model With Spherically Symmetric Errors. Unpublished Ph.D. Thesis, Carleton University, Canada.
- Ullah, A. and Walsh, V.Z. (1984). On the robustness of ML, LR and W test in regression. *Econometrica*. **52**, 1055-1066.
- Zellner, A. (1976). Bayesian and non-Bayesian analysis of the regression model with multivariate Student-t error term. *Jou. Amer. Statist. Assoc.*, Vol. 60, 601-616.