International Journal of Statistical Sciences *ISSN 1683–5603*
Vol. 3 (Special Issue), 2004, pp 69–89

The Role of the Shape Parameter for the Shrinkage Estimators of the Mean Vector of Multivariate Student-t Population

Shahjahan Khan
Department of Mathematics and Computing
University of Southern Queensland
Toowoomba, Queensland 4350, Australia
E-mail: khans@usq.edu.au

[Received April 22, 2004; Accepted August 18, 2004]

Abstract

The role of the shape parameter in determining the properties of the shrinkage and positive-rule shrinkage estimators of the mean vector of multivariate populations is of main interest in this paper. The preliminary test approach to shrinkage estimation is used to define two Stein-type estimators based on the sample information and uncertain prior non-sample information. The impact of the change in the value of the shape parameter on the performances of the estimators with respect to the criteria of unbiasedness and quadratic risk is investigated. Graphical analysis of the effect of the shape and dimension of the population on the above properties is also provided.

Keywords and Phrases: Shrinkage and positive-rule shrinkage estimators; Shape parameter; Quadratic bias and risk; Multivariate normal and Student-t, Inverted gamma, Non-central chi-square and F distributions; and Incomplete beta ratio.

AMS Classification: Primary 62F30, Secondary 62H12 and 62F10.

1 Introduction

There has been many studies in the area of 'improved' estimation following the seminal work of Stein (1956), and James and Stein (1961). Earlier, Bancroft (1944) and later Han and Bancroft (1968) developed the preliminary test estimator that uses uncertain non-sample 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), Ahmad and Saleh (1989), Maatta and Casella (1990), and Ahmad (1992) to mention a few. All the above investigations are based on the normal model. Hence the impact of the shape parameter on the properties of the estimators has not been explored by any of the above studies. Investigations on improved estimation for the Student-t model have rather been a fairly recent development. Zellner (1976), Ullah and Walsh (1984), Khan and Haq (1990), Giles (1991), Anderson (1993), Tabatabaey (1995), and Khan and Saleh (1997) studied various linear models with multivariate Student-t errors. However, Khan and Saleh (1995) investigated the problem from the sampling theory approach. Lange et al. (1989) discussed a wide range of applications of the Student-t distribution as a model for robust statistical procedures. None of the above studies investigated the impact of the change in the value of the shape parameter on the properties of the James-Stein estimator. It is well known that the flat or heavier tailed distributions are often modelled by the Student-t distribution, rather than the normal distribution. The main difference between the normal distribution and the Student-t distribution is the involvement of the shape parameter (ν) in the latter distribution. For a sufficiently large value of the shape parameter, the Student-t distribution is not significantly different from the normal distribution. However, in the absence of a large number of degrees of freedom, the properties of the Stein-type estimators depend very much on the value of the shape parameter.

In this paper we investigate some of the important properties of two well known Stein-type estimators, namely, the shrinkage and the positive-rule shrinkage estimators, as the value of the shape parameter changes. Also, we study the relative performances of the two estimators under different conditions. Moreover, the effect of the change in the dimension of the multivariate population on the quadratic bias and risk functions are discussed. Some remarks on the effect of the sample size on the behavior of the estimators are also included. In addition to the analytical comparison, extensive computations have been used to produce graphs to critically check the impact of the change in the value of the shape parameter on the properties of the estimators.

Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n$ be a random sample of size n from a p-dimensional multivariate Student-t population with unknown mean $\boldsymbol{\mu}$, common scaled covariance matrix $\sigma^2 I_p$ and arbitrary shape parameter ν. Note that the covariance of $\boldsymbol{X}$ is $\frac{\nu}{\nu-2}\sigma^2 I_p$. Also assume that *uncertain non-sample 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 or may not be true. As suggested by R.A. Fisher, this uncertain prior information should be treated as a *nuisance parameter*, and its uncertainty can be removed by performing an appropriate test on the null hypothesis. We wish to incorporate both the sample data and the uncertain non-sample prior information in estimating the mean vector $\boldsymbol{\mu}$. Inclusion of such a non-sample information in the definition of the estimators is likely to improve the performance of the estimators with respect to some well known statistical criteria. First we obtain the unrestricted maximum likelihood estimator (mle) of the mean vector $\boldsymbol{\mu}$ and the common scaled variance σ^2 from the likelihood function of the parameters. Based on the unrestricted and restricted (by H_0) mle of σ^2, we derive the likelihood ratio test for testing $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ against the alternative, $H_1 : \ \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$. As discussed by Anderson (1993), this test is robust and is applicable to the entire class of elliptical models. Following Bancroft (1944), the *preliminary test estimator* (PTE), $\hat{\boldsymbol{\mu}}^{pt}$ of $\boldsymbol{\mu}$ has been defined by using an appropriate test statistic and the unrestricted as well as the restricted mle's of $\boldsymbol{\mu}$. Then the preliminary test approach has been applied 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 unknown mean vector, $\boldsymbol{\mu}$. The bias, quadratic bias and quadratic risk functions are obtained for the aforesaid estimators. The relative performance of the estimators is investigated by analyzing the risks under different conditions. The effect of the shape parameter on the performance of the estimators is of special interest. For the computations and derivations of the main results of the paper the multivariate Student-t distribution is viewed as a mixture of the multivariate normal distribution and Inverted Gamma (IG) distributions.

The multivariate Student-t model along with some of its applications have been introduced in the next section. Section 3 provides the unrestricted, restricted, preliminary test, shrinkage and positive rule shrinkage estimators of the mean vector. Some useful results for the computation of the bias and quadratic risk of the estimators are included in section 4. The bias of the estimators, as function of the shape parameter, are computed and analyzed in Section 5. Section 6 evaluates the expressions of quadratic risks of the estimators that directly depend on the shape parameter. The performances of the estimators, including the graphical comparisons, 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 distribution for linear models are applicable to those of normal models, but not vice-versa. Prucha and Kelejian (1984) critically analyzed the problems of normal distribution and advocated for the Student-t distribution as a better alternative for many econometric 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. Apart from 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. 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 Chmielewiski (1981), and other notable references include Zellner (1976), Fang and Zhang (1980), Ullah and Walsh

(1984), Khan and Haq (1990), Fang and Anderson (1990), Gupta and Vargava (1993), Khan (1997), and Khan and Saleh (1997).

Let $\boldsymbol{X}$ be a p-dimensional random vector that follows a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix, $\tau^2 I_p$. Assume that τ follows an Inverted Gamma distribution with a scale parameter σ and a shape parameter ν. Then the density function of τ is given by

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

The distribution of $\boldsymbol{X}$, conditional on τ, can be denoted by $[\boldsymbol{X}|\tau] \sim N_p(\boldsymbol{\mu}, \tau^2 I_p)$, and the unconditional distribution of τ can be denoted by $\tau \sim IG(\nu, \sigma^2)$.

Now, consider a random sample of size n, $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n$, from the above p-variate normal population. Conditional on a given τ, the joint density function of the sample is given by

$$p(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n;\boldsymbol{\mu},\tau) = (2\pi\tau^2)^{-np/2} \exp\left\{-\frac{1}{2\tau^2}\sum_{j=1}^{n}(\boldsymbol{x}_j-\boldsymbol{\mu})'(\boldsymbol{x}_j-\boldsymbol{\mu})\right\}. \tag{2.2}$$

Then, it is well known that the mixture distribution of the random sample and τ is a multivariate Student-t distribution with p.d.f.

$$p(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n;\boldsymbol{\mu},\sigma,\nu) = k_n(\nu,p)(\sigma^2)^{-np/2}\left[1+\frac{1}{\nu\sigma^2}\sum_{j=1}^{n}(\boldsymbol{x}_j-\boldsymbol{\mu})'(\boldsymbol{x}_j-\boldsymbol{\mu})\right]^{-\frac{\nu+np}{2}} \tag{2.3}$$

where $k_n(\nu,p) = \{\Gamma\left((\nu+np)/2\right)\}\left\{(\pi\nu)^{\frac{np}{2}}\Gamma\left(\nu/2\right)\right\}^{-1}$ is the normalizing constant.

The above density is a multivariate Student-t density with shape parameter ν, location parameter $\boldsymbol{\mu}$ and scaled covariance matrix $\sigma^2 I_p$. Note that $\sigma^2 I_p$ is the scaled covariance matrix of $\boldsymbol{X}_j$ up to a multiplication factor of $\frac{\nu}{\nu-2}$. In the conventional notation we write $\boldsymbol{X}_j \sim t_p\left(\boldsymbol{\mu}, \sigma^2 I_p, \nu\right)$ for all j. A method of moment estimator of ν can be found in Singh (1988).

From the properties of the Student-t distribution, it may be noted that the $\boldsymbol{X}_j's$ are uncorrelated but not independent (cf. Anderson, 1993). However, for a given value of τ, each $\boldsymbol{X}_j$ is independently normally distributed, that is, $[\boldsymbol{X}_j|\tau] \sim N_p(\boldsymbol{\mu}, \tau^2 I_p)$ for all j.

3 The Estimators

From the expression (2.3), the log-likelihood function of the parameters based on the given sample can be written as

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

Then, following Khan (1997), the *unrestricted maximum likelihood estimator* (UE) of $\boldsymbol{\mu}$ and σ^2 is obtained as

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

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

respectively, and the corresponding *restricted maximum likelihood estimators* (RE) become

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

The likelihood ratio statistic for testing $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$, as specified in section 1, is given by

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

Under the null hypothesis, the statistic

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

follows a scaled F-distribution with p and $m = (n - p - 1)$ degrees of freedom (cf. Zellner, 1976). As discussed by Anderson (1993), the F-test 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. Therefore, the test of $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$, can be based on

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

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

The *preliminary test estimator* (PTE) of $\boldsymbol{\mu}$ as a function of the UE, RE and T^2 is defined by

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

where $I(A)$ is an indicator function of the set A; and T^2_α is the $(1-\alpha)^{th}$ quantile of the distribution of the T^2-statistic. Ahmed and Saleh (1989) defined a similar preliminary test estimator for the two-sample multivariate normal problem.

The PTE is an extreme choice between the UE and RE, as it does not allow any smooth transition between the two extremes. 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

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

where k^* is a *shrinkage constant* which is optimal in the sense of minimizing the quadratic risk, and an optimal value of k^* for the current problem is found to be $k = \dfrac{p-2}{m+2}$.

The SE becomes unstable and unreliable when the value of the T^2-statistic is too close to zero. To avoid this difficulty the *positive-rule shrinkage estimator* (PRSE) for $\boldsymbol{\mu}$ is defined as

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

The above two shrinkage estimators (SE and PRSE) are of the same form as for the multivariate normal model. However, as the forthcoming analysis reveals, the properties of the above Student-t based estimators are different from that of the normal model. The main objective of this paper is to study the effect of the shape parameter on the relative performance of the two shrinkage estimators based on the criteria of bias, quadratic bias and quadratic risk. A detailed study of the risk analysis with respect to the change in the shape parameter as well as the dimension of the multivariate Student-t model is also provided.

4 Some Useful Results

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

Lemma 4.1. *If $\boldsymbol{U}$ is an $n \times 1$ vector of Student-t variables with ν d.f., mean $\boldsymbol{\beta}$ and scaled covariance matrix $\sigma^2 I_n$, in which I_n is an identity matrix of order n, then*

$$E[\phi(\boldsymbol{U}'\boldsymbol{U})\boldsymbol{U}] = \boldsymbol{\beta} E[\phi(\chi^2_{n+2}, \lambda^*)] \tag{4.1}$$

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

Lemma 4.2. *If* $\boldsymbol{U}$ *is an* $n \times 1$ *vector of Student-t variables with* ν *d.f., mean* $\boldsymbol{\beta}$ *and scaled covariance matrix* $\sigma^2 I_n$, *and* $\boldsymbol{V}$ *is a positive definite matrix of order* n, *then*

$$E[\phi(\boldsymbol{U}'\boldsymbol{U})\boldsymbol{U}'\boldsymbol{V}\boldsymbol{U}] = E[\phi(\chi^2_{n+2}, \lambda^*)]tr(\boldsymbol{V}) + \boldsymbol{\beta}'\boldsymbol{V}\boldsymbol{\beta}E[\phi(\chi^2_{n+4}, \lambda^*)] \tag{4.2}$$

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

Lemma 4.3. *If* $\boldsymbol{U}$ *is an* $n \times 1$ *vector of Student-t variables with* ν *d.f., mean* $\boldsymbol{\beta}$ *and scaled covariance matrix* $\sigma^2 I_n$, *then*

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

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

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 $IG(\nu, \sigma^2)$ distribution.

5 Study of Bias

The bias of the UE $\tilde{\boldsymbol{\mu}}$ is $B_1(\tilde{\boldsymbol{\mu}}; \boldsymbol{\mu}) = 0$. However, the bias of the RE $\hat{\boldsymbol{\mu}}$ depends on the difference between the true and the suspected value of the mean vector $\boldsymbol{\mu}$, and is given by $B_2(\hat{\boldsymbol{\mu}}; \boldsymbol{\mu}) = (\boldsymbol{\mu}_0 - \boldsymbol{\mu}) = -\boldsymbol{\delta}$. The bias of the RE is non-zero under the alternative hypothesis. However, as it is in the following theorems, the bias of both the SE and PRSE depends on the value of the shape parameter of the Student-t model.

Theorem 5.1.*For the p-dimensional Student-t model with* ν *as the shape parameter the bias of the SE* $\hat{\boldsymbol{\mu}}^s$ *is given by*

$$B_3(\hat{\boldsymbol{\mu}}^s; \boldsymbol{\mu}) = \boldsymbol{\delta}kmE^{(2)}[\chi^{-2}_{p+2}(\Delta^*)] \tag{5.1}$$

where

$$E^{(2)}[\chi^{-2}_{p+2}(\Delta^*)] = \sum_{r=0}^{\infty} \frac{1}{p+2r} h(r, \nu, \Delta^*) \tag{5.2}$$

in which

$$h(r, \nu, \Delta^*) = \frac{\Gamma\left(\frac{\nu}{2}+r\right)}{r!\Gamma\left(\frac{\nu}{2}\right)} \times \frac{[\frac{\Delta^*}{(\nu-2)}]^r}{\left[1+\frac{\Delta^*}{(\nu-2)}\right]^{\frac{\nu}{2}+r}} \tag{5.3}$$

with $\Delta^* = n\dfrac{\nu-2}{\nu\sigma^2}\Delta$ *such that* $\Delta = \boldsymbol{\delta}'\boldsymbol{\delta}$ *and* $\boldsymbol{\delta} = (\boldsymbol{\mu} - \boldsymbol{\mu}_0)$.

Proof: For the proof of the above theorem, and the forthcoming theorems, we use the following transformation:

$$\boldsymbol{y} = \frac{\sqrt{n}}{\tau}(\tilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0). \tag{5.4}$$

For a given value of τ, $\boldsymbol{y} \sim N_p\left(\frac{\sqrt{n}}{\tau}\boldsymbol{\delta}; I_p\right)$, and hence (under H_1) the statistic $\boldsymbol{y}'\boldsymbol{y}$ follows a non-central chi-square distribution with p degrees of freedom and non-centrality parameter $\Delta_\tau = n\tau^{-2}\boldsymbol{\delta}'\boldsymbol{\delta}$. Therefore, the T^2 statistic defined in Section 3 can be expressed as

$$T^2 = \frac{\boldsymbol{y}'\boldsymbol{y}}{\chi^2_m}. \tag{5.5}$$

Thus, under H_1 the statistic $\frac{p}{m}T^2$ follows a non-central F distribution with p and m degrees of freedom and non-centrality parameter $\Delta_\tau = n\tau^{-2}\boldsymbol{\delta}'\boldsymbol{\delta}$.

The expression of the bias, conditional on τ, is then obtained by working out the expected value of the difference between $\boldsymbol{\mu}^s$ and $\boldsymbol{\mu}$, and applying appropriate results from section 4. The theorem follows from the expectation of the conditional bias with respect to the $IG(\nu, \sigma^2)$ distribution.

The quadratic bias of the SE is obtained as

$$QB_3(\hat{\boldsymbol{\mu}}^s; \boldsymbol{\mu}) = \Delta\{kmE^{(2)}[\chi^{-2}_{p+2}(\Delta^*)]\}^2. \tag{5.6}$$

Theorem 5.2. *For the p-dimensional Student-t model with ν as the shape parameter the bias of the PRSE $\hat{\boldsymbol{\mu}}^{s+}$ is given by*

$$\begin{aligned} B_4(\hat{\boldsymbol{\mu}}^{s+}; \boldsymbol{\mu}) &= -\boldsymbol{\delta}kmE^{(2)}[\chi^{-2}_{p+2}(\Delta^*)] - \boldsymbol{\delta}G^{(2)}_{p+2,m}(q_2; \Delta^*) \\ &\quad +\boldsymbol{\delta}kmE^{(2)}[\chi^{-2}_{p+2}(\Delta^*)I(F_{p+2,m}(\Delta^*) \le q_2)] \end{aligned} \tag{5.7}$$

where

$$E[(\chi^{-2a}_{p+b}, \Delta^*)] = \sum_{r=0}^{\infty} h(r, \nu, \Delta^*)\frac{1}{\zeta(p, r, a)} \tag{5.8}$$

in which $\zeta(p, r, a) = \{p + b - 2 + 2r\}\{p + b - 4 + 2r\}\cdots\{p + b - 2(a - 1) + 2r\}$ for $a = 1, 2$ and $b = 0, 2, 4$;

$$G^{(a)}_{p+a,m}(q_a; \Delta^*) = \sum_{r=0}^{\infty} h(r, \nu, \Delta^*)I_{u_a}\left(\frac{p+a}{2} + r; \frac{m}{2}\right) \tag{5.9}$$

in which $I_{u_a}(\frac{p+a}{2} + r; \frac{m}{2})$ is the well known incomplete beta ratio, evaluated at $u_a = \frac{q_a}{1+q_a} = \frac{m(p-2)}{(p+a+2r)(m+2)+m(p-2)}$ with $q_a = \frac{m(p-2)}{(p+a+2r)(m+2)}$ for $a = 2, 4$; and

$$E^{(2)}\left[\chi^{-2}_{p+2}(\Delta^*)I(F_{p+2,m}(\Delta^*) \le q_2)\right] = \sum_{r=0}^{\infty} \frac{h(r, \nu, \Delta^*)}{\xi_2(p, r)}I_{u_2}\left(\frac{p+2}{2} + r; \frac{m}{2}\right) \tag{5.10}$$

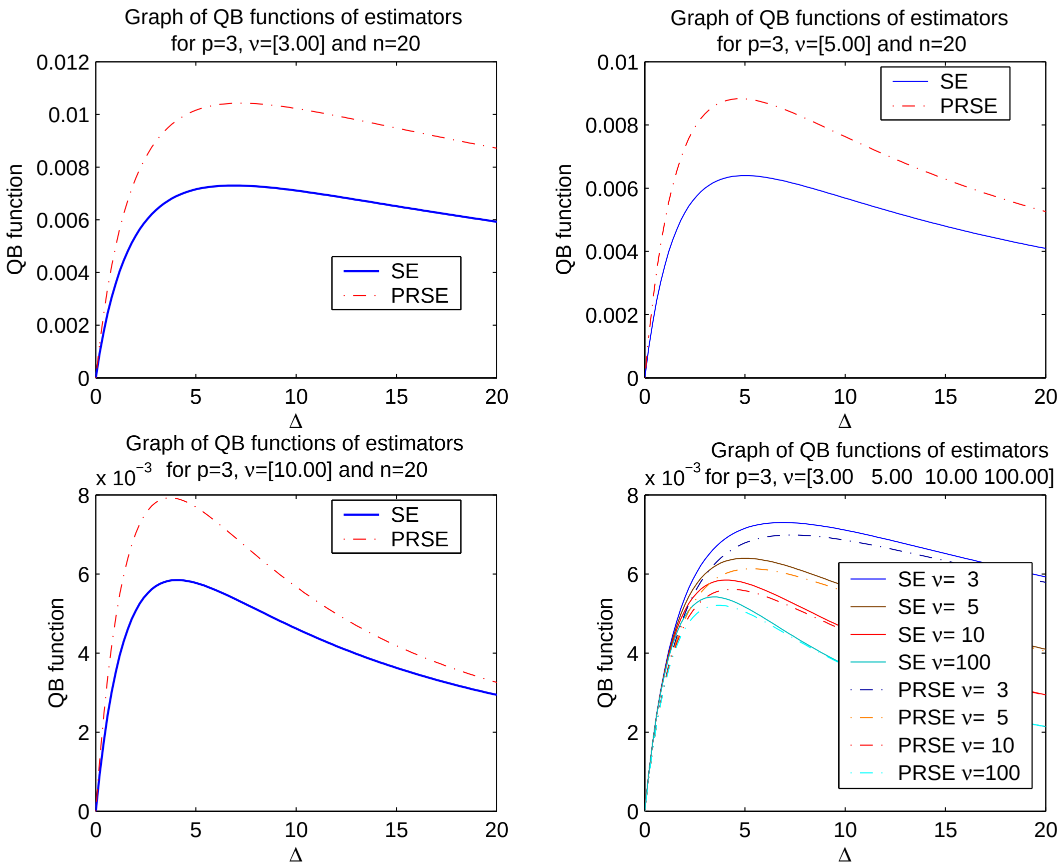

Figure 1: Quadratic Bias functions of the SE and PRSE for $p = 3$, and varying shape parameter ν

with $\xi_2(p,r) = (p+2r)$.

Proof. The proof of the above theorem follows from the same procedure and argument as in the proof of the previous theorem.

The first term in (5.7) is the same as the bias of the SE. The SE and PRSE are unbiased for $\boldsymbol{\mu}$ under the null hypothesis, but they are biased under the alternative hypothesis. 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 as well as the value of the shape parameter.

The quadratic bias of the PRSE is found to be

$$\begin{aligned} QB_4(\hat{\boldsymbol{\mu}}^s;\boldsymbol{\mu}) \quad = \quad & QB_3(\hat{\boldsymbol{\mu}}^s;\boldsymbol{\mu}) + \Delta\{G^{(2)}_{p+2,m}(q_2;\Delta^*)\}^2 \\ & +\Delta\{kmE^{(2)}[\chi^{-2}_{p+2}(\Delta^*)I(F_{p+2,m}(\Delta^*)\le q_2)]\}^2 \\ & +2\Delta kmE^{(2)}[\chi^{-2}_{p+2}(\Delta^*)]\times G^{(2)}_{p+2,m}(q_2;\Delta^*) \\ & -2(km)^2E^{(2)}[\chi^{-2}_{p+2}(\Delta^*)]\times E^{(2)}[\chi^{-2}_{p+2}(\Delta^*)I(F_{p+2,m}(\Delta^*)\le q_2)] \\ & -2kmG^{(2)}_{p+2,m}(q_2;\Delta^*)\times E^{(2)}[\chi^{-2}_{p+2}(\Delta^*)I(F_{p+2,m}(\Delta^*)\le q_2)]. \end{aligned} \tag{5.11}$$

5.1 Graphical Analysis of Quadratic Bias

The graphs in Figure 1 demonstrate the change in the value of the quadratic bias of the two shrinkage estimators for $p=3$, $n=20$, $\nu=3,\ 5,\ 10$ and 100. The form of the QB function of the SE and PRSE is quite similar. For both estimators the QB is 0 at $\Delta=0$. As the value of Δ grows larger the QB function very quickly moves upward and slowly declines after reaching the maximum for some moderate value of Δ around $\Delta=5$. The curve of the QB of the SE is always lower than or same as that of the PRSE for all values of p and ν. The maximum difference between the two QB curves of the two estimators is observed near the value of Δ for which the SE and PRSE attain their maximum values. However, for $\Delta=0$ as well as for very large values of Δ the difference between the QB functions of the SE and PRSE is negligible.

The minimum value of the QB of both the SE and PRSE occurs at $\Delta=0$ regardless of the choice of p and the value of the shape parameter. However, the maximum value of the BQ depends on the shape parameter ν. For $\nu=3$ the maximum value of the QB of the SE and PRSE is near 0.01 which occurs when Δ is near 5. Whereas, for $\nu=5$ the maximum value of the QB of the SE and PRSE is near 0.009 which occurs when Δ is near 5, and that for $\nu=10$ is 0.008. Hence the value of the QB function of both the SE and PRSE decreases as the value of the shape parameter of the model, ν increases. So, if a Student-t model is wrongly represented by a normal model then the value of the QB will be less than what it actually should have been under correct model. For large values of the shape parameter the QB of the PRSE approaches to that of the SE for smaller values of Δ. Thus in the case of misspecification of a Student-t model by a normal model would give a false impression that the QB curves of the PRSE converse

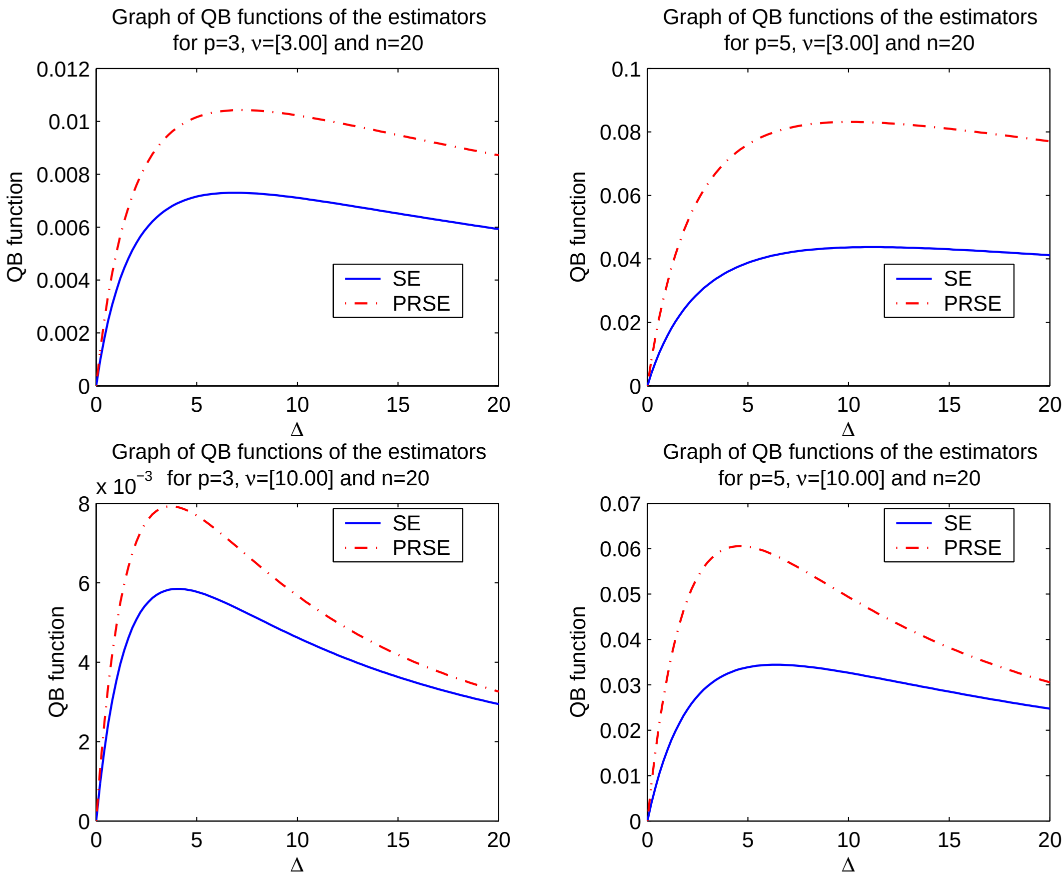

Figure 2: Quadratic Bias functions of the SE and PRSE for varying p and shape parameter ν

to that of the SE for smaller values of Δ and does so at a faster rate than it should be.

The graphs in Figure 2 demonstrate the change in the value of the quadratic bias of the two shrinkage estimators for fixed $n = 20$, and varying values of $p = 3,\ 5$ and $\nu = 3,\ 10$. The top two graphs vary only in the value of p, the dimension of the population of interest. For smaller value of p the QB of both the SE and PRSE are smaller than that for larger value of p. For example, the maximum value of the PRSE is near 0.01 for $p = 3$ compared to that of 0.08 for $p = 5$. Moreover, the maximum value of the QB of the SE is about 0.007 when $p = 3$ and just over 0.04 for $p = 5$. The maximum difference between the QB of the two estimators increases with the increase in the value of p. The bottom two graphs of Figure 2 illustrate very similar picture. They also reveal that the QB curve of the PRSE converges to that of the SE for larger values of Δ when the value of p is larger. Furthermore, it is observed that the value of the QB for both the estimators goes up as the dimension of the population increases.

The above graphical analyses of the QB of the SE and PRSE reveals that both estimators are biased but the SE has smaller QB than the PRSE except for Δ near 0 or very large values of Δ. However, the magnitude of this large value of Δ depends on the choice of the shape parameter ν, and or the dimension of the population p.

6 Evaluation of Risks

The quadratic risk of an estimator $\boldsymbol{\theta}^*$, based on a random sample of size n, for estimating a parameter $\boldsymbol{\theta}$ under the quadratic error loss function of the form

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

is defined as

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

of $\boldsymbol{\mu}$, $\tilde{\boldsymbol{\mu}}$, conditional on τ^2, is $p\tau^2$. The expectation of the conditional risk with respect to the distribution of τ yields

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

Note that as $\nu \to \infty$ the risk of the UE reduces to $p\sigma^2$, the risk of the mle under the multivariate normal model. However, as it is in the following theorems, the quadratic risks of both the SE and PRSE depend on the value of the shape parameter.

Theorem 6.1.*For the p-dimensional Student-t model with ν as the shape parameter the quadratic risk of the SE $\hat{\boldsymbol{\mu}}^s$ in estimating the mean vector $\boldsymbol{\mu}$ is*

$$\begin{aligned} R_3(\hat{\boldsymbol{\mu}}^s;\boldsymbol{\mu}) &= \frac{\nu\sigma^2}{\nu-2}\left[p - 2km + k^2m(m+2)E^{(0)}\{\chi_p^{-2}(\Delta^*)\}\right] \\ &\quad +2\Delta kmnE^{(2)}\{\chi_{p+2}^{-2}(\Delta^*)\}. \end{aligned} \tag{6.4}$$

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

$$R_3(\hat{\boldsymbol{\mu}}^s;\boldsymbol{\mu}) = E_\tau[E\{n(\hat{\boldsymbol{\mu}}^s-\boldsymbol{\mu})'(\hat{\boldsymbol{\mu}}^s-\boldsymbol{\mu})\}|\tau] \tag{6.5}$$

where

$$\begin{aligned} E\{n(\hat{\boldsymbol{\mu}}^s-\boldsymbol{\mu})'(\hat{\boldsymbol{\mu}}^s-\boldsymbol{\mu})|\tau\} &= E\{n(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu})|\tau\} \\ &\quad + k^2E\{n(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)T^{-4}|\tau\} \\ &\quad - 2kE\{n(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)T^{-2}|\tau\}. \end{aligned} \tag{6.6}$$

The first term in (6.6) 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 (5.4) we have

$$\begin{aligned} E\{n(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)T^{-4}|\tau\} &= E\left\{\frac{\chi_m^4}{(\boldsymbol{y}'\boldsymbol{y})^2}(\boldsymbol{y}'\boldsymbol{y})\tau^2|\tau\right\} \\ &= \tau^2 m(m+2)E\{\chi_p^{-2}(\Delta_\tau)\}. \end{aligned} \tag{6.7}$$

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 (5.4) we get

$$E\{n(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)T^{-2}|\tau\} = E\left\{\frac{\chi_m^2}{(\boldsymbol{y}'\boldsymbol{y})}\times(\boldsymbol{y}'\boldsymbol{y}))\tau^2\right\} = \tau^2 m \tag{6.8}$$

and

$$E\{n(\boldsymbol{\mu}-\boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)T^{-2}|\tau\} = \Delta mnE\{\chi_{p+2}^{-2}(\Delta_\tau)\}. \tag{6.9}$$

Now collecting all the terms in equation (6.6) and simplifying, we have

$$\begin{aligned} E\{n(\hat{\boldsymbol{\mu}}^s-\boldsymbol{\mu})'(\hat{\boldsymbol{\mu}}^s-\boldsymbol{\mu})|\tau\} &= \tau^2 p + k^2\tau^2 m(m+2)E\{\chi_p^{-2}(\Delta_\tau)\} \\ &\quad - 2k\left[\tau^2 m - nm\Delta E\{\chi_{p+2}^{-2}(\Delta_\tau)\}\right] \\ &= \tau^2\left[p - 2km + k^2 m(m+2)E\{\chi_p^{-2}(\Delta_\tau)\}\right] \\ &\quad + 2knm\Delta E\{\chi_{p+2}^{-2}(\Delta_\tau)\}. \end{aligned} \tag{6.10}$$

The expectation on the above expression with respect to the $IG(\nu,\sigma^2)$ distribution yields the risk expression in (6.4).

Using the following relationship of the expectation of the noncentral chi-square variables,

$$\Delta_\tau E\{\chi_{p+2}^{-2}(\Delta_\tau)\} = 1-(p-2)E\{\chi_p^{-2}(\Delta_\tau)\} \tag{6.11}$$

the expression in (6.10) can be written as

$$R_3(\hat{\boldsymbol{\mu}}^s;\boldsymbol{\mu}) = \tau^2\left[p-(p-2)^2\frac{m}{m+2}E\{\chi_p^{-2}(\Delta_\tau)\}\right]. \tag{6.12}$$

Theorem 6.2.*For the p-dimensional Student-t model with ν as the shape parameter the quadratic risk of the PRSE $\hat{\boldsymbol{\mu}}^{s+}$ in estimating the mean vector $\boldsymbol{\mu}$ is*

$$\begin{aligned} R_4(\hat{\boldsymbol{\mu}}^{s+};\boldsymbol{\mu}) &= \frac{\nu\sigma^2}{\nu-2}p\left[1-G^{(2)}_{p+2,m}(q_2;\Delta^*)\right]-2\frac{\nu}{\nu-2}km\left[1-G^{(0)}_{p,m}(q;\Delta^*)\right]\\ &+\frac{\nu\sigma^2}{\nu-2}k^2m(m+2)\left[E^{(0)}\left\{\chi_p^{-2}(\Delta^*)\right\}-E^{(2)}\left\{\chi_{p+2}^{-2}(\Delta^*)I\left(F_{p+2,m}(\Delta^*)\le q_2\right)\right\}\right]\\ &-\Delta n\left[G^{(4)}_{p+4,m}(q_4;\Delta^*)-2G^{(2)}_{p+2,m}(q_2;\Delta^*)\right]\\ &+\Delta 2kmn\left[E^{(2)}\left\{\chi_{p+2}^{-2}(\Delta^*)\right\}-E^{(2)}\left\{\chi_{p+2}^{-2}(\Delta^*)I\left(F_{p+2,m}(\Delta^*)\le q_2\right\}\right]. \end{aligned}\tag{6.13}$$

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

$$R_4(\hat{\boldsymbol{\mu}}^{s+};\boldsymbol{\mu}) = E_\tau[E\{n(\hat{\boldsymbol{\mu}}^{s+}-\boldsymbol{\mu})'(\hat{\boldsymbol{\mu}}^{s+}-\boldsymbol{\mu})\}|\tau] \tag{6.14}$$

where

$$\begin{aligned} E\{n(\hat{\boldsymbol{\mu}}^{s+}-\boldsymbol{\mu})'(\hat{\boldsymbol{\mu}}^{s+}-\boldsymbol{\mu})|\tau\} &= E\{n(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu})'\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu})|\tau\}\\ &+E\{n(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)k^2T^{-4}|\tau\}+\ E\{n(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)I(T^2\le k)|\tau\}\\ &+E\{n(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)k^2T^{-4}kI(T^2\le K)|\tau\}\\ &-2E\{n(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)kT^{-2}|\tau^2\}-2E\{n(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu})'(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)I(T^2\le k)|\tau\}\\ &+2E\{(n(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)kT^{-2}I(T^2\le k)|\tau\}\\ &+2E\{n(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)kT^{-2}I(T^2\le k)|\tau\}\\ &-2E\{n(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)k^2T^{-4}I(T^2\le k)|\tau\}\\ &-2E\{n(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)'(\tilde{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)kT^{-2}I(T^2\le k)|\tau\}. \end{aligned}\tag{6.15}$$

Now following the same procedure and argument as in the proof of the previous theorem this theorem can be proved after completing appropriate algebraic computations and simplifications.

The risk of the PRSE in (6.13) can alternatively be expressed as

$$
\begin{aligned}
R_4(\hat{\boldsymbol{\mu}}^{s+};\boldsymbol{\mu}) = R_3(\hat{\boldsymbol{\mu}}^{s};\boldsymbol{\mu}) & \\
+\frac{\nu\sigma^2}{\nu-2}\Big[& 2kmG^{(0)}_{p,m}(q;\Delta^*) - pG^{(2)}_{p+2,m}(q_2;\Delta^*) \\
& -k^2m(m+2)E^{(2)}\Big\{\chi^{-2}_{p+2}(\Delta^*)I\left(F_{p+2,m}(\Delta^*)\le q_2\right)\Big\}\Big] \\
-\Delta n\Big[& G^{(4)}_{p+4,m}(q_4;\Delta^*) \\
& -2G^{(2)}_{p+2,m}(q_2;\Delta^*) + 2kmE^{(2)}\Big\{\chi^{-2}_{p+2}(\Delta^*)I\left(F_{p+2,m}(\Delta^*)\le q_2\right)\Big\}\Big]
\end{aligned} \tag{6.16}
$$

7 Analysis of Risks

The UE $\tilde{\boldsymbol{\mu}}$ has a constant risk that depends on the shape and the dimension of the Student-t model. As expected, this risk approaches the risk 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 actually should be, if a Student-t model is mis-specified as a normal model. It is well known that the quadratic risk of the Stein-type *shrinkage estimator* $\hat{\boldsymbol{\mu}}^s$ is smaller than that of $\tilde{\boldsymbol{\mu}}$. Near $\Delta^*=0$ the risk of $\hat{\boldsymbol{\mu}}^s$ is the smallest with compared to that of $\tilde{\boldsymbol{\mu}}$. But $R_3(\hat{\boldsymbol{\mu}}^s;\boldsymbol{\mu})$ approaches to $R_1(\tilde{\boldsymbol{\mu}};\boldsymbol{\mu})$ as Δ^* grows larger. For smaller values of ν the risk curve of the SE exceeds the risk curve of $\tilde{\boldsymbol{\mu}}$. Thus $\hat{\boldsymbol{\mu}}^s$ dominates $\tilde{\boldsymbol{\mu}}$ for large values of ν. Hence there is no uniform dominance of the SE over the UE for all values of ν. However, for larger values of Δ^* the risk of $\hat{\boldsymbol{\mu}}^s$ approaches that of $\tilde{\boldsymbol{\mu}}$ for all ν. Nevertheless the rate is faster for a larger value of ν than a smaller ν.

Under H_0, the *positive-rule shrinkage estimator* has a smaller risk than the shrinkage estimator for any given ν. It dominates $\hat{\boldsymbol{\mu}}^s$, and $\tilde{\boldsymbol{\mu}}$ when the null hypothesis is true. $R_4(\hat{\boldsymbol{\mu}}^{s+};\boldsymbol{\mu})$ approaches $R_1(\tilde{\boldsymbol{\mu}};\boldsymbol{\mu})$ as Δ^* increases, and both risk functions coincide as $\Delta^*\to\infty$ after meeting at some value of Δ^*. Both $R_3(\hat{\boldsymbol{\mu}}^s;\boldsymbol{\mu})$ and $R_4(\hat{\boldsymbol{\mu}}^{s+};\boldsymbol{\mu})$ approaches to $R_1(\tilde{\boldsymbol{\mu}};\boldsymbol{\mu})$; but the former tends to meet it for a smaller value of Δ^* than the later for any fixed ν. 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 and for any given ν, the following dominance picture of the estimators emerges:

$$\hat{\boldsymbol{\mu}}^{s+}\succ\hat{\boldsymbol{\mu}}^{s}\succ\tilde{\boldsymbol{\mu}} \tag{7.1}$$

where '$\succ$' stands for domination. However, the picture is not so clear cut under the alternative hypothesis, as there is no uniform domination of each other for every value of ν when H_0 is not true.

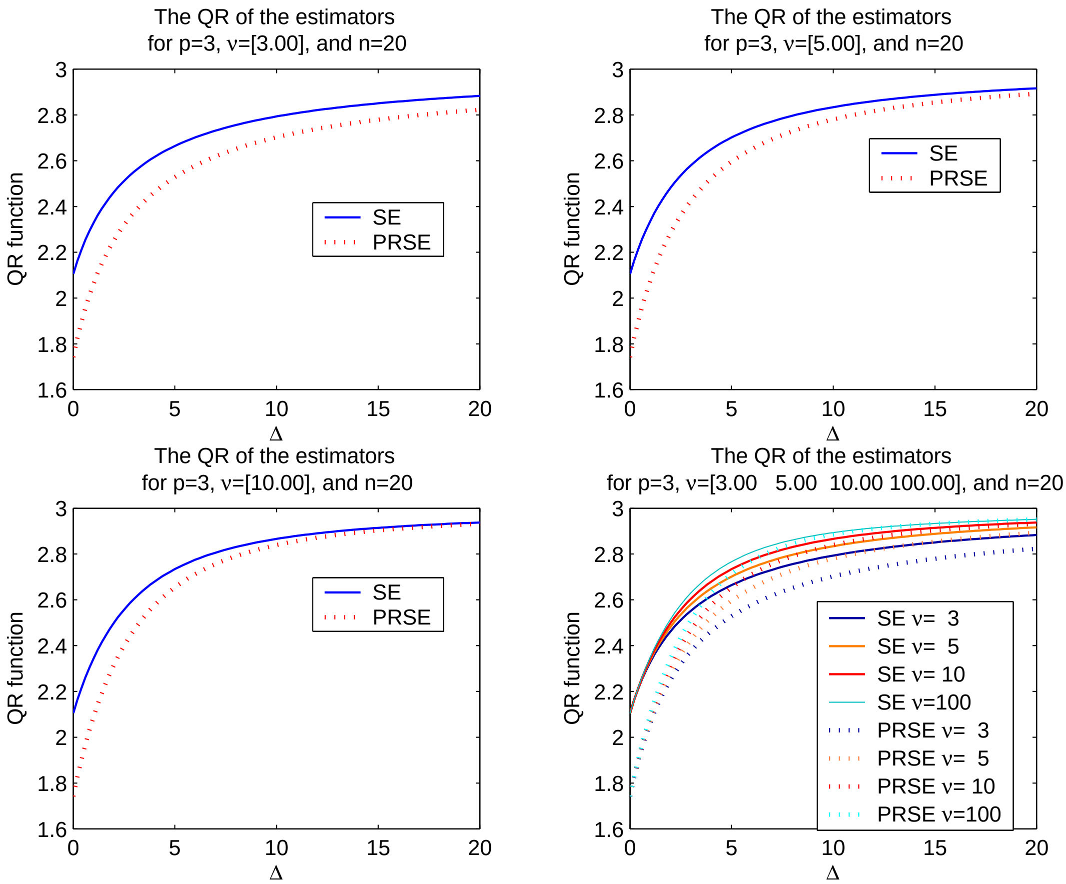

Figure 3: Quadratic Risk functions of the SE and PRSE for $p = 3$, and varying shape parameter ν

7.1 Graphical Analysis of Quadratic Risk

Without loss of generality, the graphs for the quadratic risk (QR) of the estimators are produced for some arbitrary value of σ^2 and n. The graphs of Figure 3 are produced for fixed $n = 20$, $p = 3$ but varying values of the shape parameter, $\nu = 3, \ , 5, \ , 10, \ 100$. The form of the QR curves for both the SE and PRSE is very similar.

A common feature of the quadratic risk functions of the SE and the PRSE is that they start at the minimum (zero) when $\Delta = 0$. Then it climbs up as Δ grows larger and larger at a faster rate near $\Delta = 0$. The QR curve of the PRSE approaches to that of the SE from below as the value of Δ grows larger. Finally, as $\Delta \to \infty$, the difference between the two curves reduces gradually with the decrease in the rate of growth in the QR curves.

The general shape of the QR curve is upward increasing as the value of Δ increases. But the rate of growth of the curve diminishes as Δ grows larger and larger. At $\Delta = 0$, both the SE and PRSE has the minimum QR. This QR is always smaller than that of the maximum likelihood estimator (which equals the dimension of the population) regardless of the value of the shape parameter and the dimension of the population $p > 3$. It is observed that the PRSE always has a smaller QR than the SE if Δ is not too large, for the same value of the shape parameter as well as for the same p and n.

The difference between the QR of the two estimators decreases as the value of Δ increases, and finally the two curves meet each other for some large value of Δ. As the value of the shape parameter ν increases the value of Δ at which the QR curves of the two estimators meet decreases. So if a Student-t model is misspecified by a normal model then the meeting (and converging) of the two curves will occur at a much more smaller value of Δ and it should be. Also the rate of convergence of the QR of the PRSE to the SE is faster for larger values of ν. From the graphs in Figure 3, the two curves meet a lost faster when $\nu = 10$ than for $\nu = 3, $ or 5.

The maximum difference between the QR of the two shrinkage estimators occurs at $\Delta = 0$ and grows higher as the dimension of the population increases. That is, the PRSE performs relatively better than the SE on a higher dimension than on a lower dimension. It is noted that the difference between the risk curves of the two estimators is higher when the value of the shape parameter is smaller. Thus for a smaller value of the shape parameter the QR function of the SE and PRSE are closer to one another than for a larger value of the shape parameter. Moreover, this difference increases as the dimension of the population increases.

Finally, the cost of mis-specification of a multivariate Student-t model with smaller degrees of freedom by a normal model has the potential to be highly significant. However, the use of an appropriate multivariate Student-t model eliminates the chance of under stating the actual risk of the shrinkage estimators.

The graphs of Figure 4 are produced for fixed $n = 20$ but varying values of the dimension of the population $p = 3, \ 5$ and the shape parameter $\nu = 3, \ 10$. From the graphs it is clear that the QR of the SE and PRSE is minimum at $\Delta = 0$, but the minimum value depends on the value of p. Both the minimum and maximum QR of

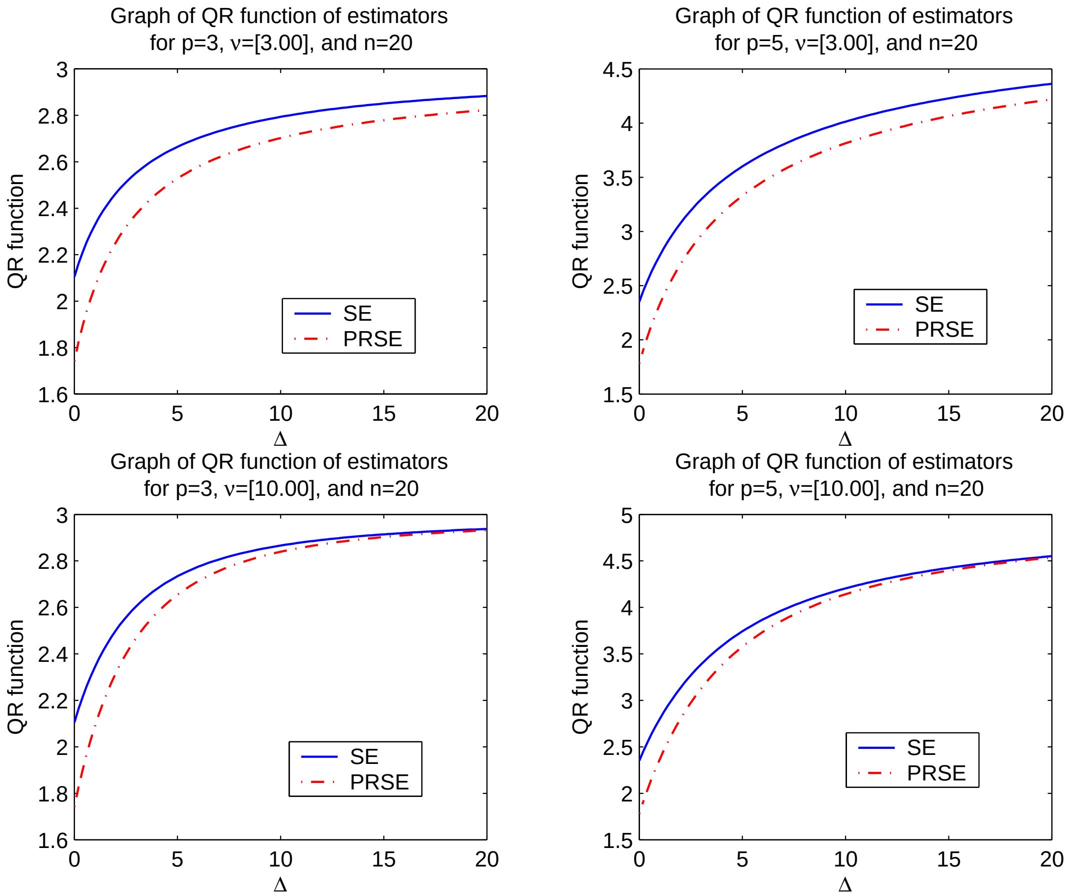

Figure 4: Quadratic Risk functions of the SE and PRSE for varying p and shape parameter ν

the estimators depend on the value of p.

It is clear from the above analyses of the risk function that the value of the shape parameter plays a key role in the determination of the value of the QR of the shrinkage estimators. Also, the risk function is quite sensitive to the change in the dimension of the population under study. In general, the risk of the shrinkage estimators increases with the increase in the dimension of the population.

8 Concluding Remarks

The PRSE uniformly dominates the SE for all ν and p in terms of the QR. The QR is higher for both the SE and PRSE for smaller values of ν. For very large values of Δ both estimators perform the same. The minimum QR of both the estimators increases as the dimension of the population increases. Since the value of Δ depends on the quality of the uncertain non-sample prior information, and it is often obtained from expert knowledge or previous studies it is more likely that the value of Δ would not be far away from 0. For any value of Δ near 0, the PRSE always dominates over the SE for all ν and p, but does more so for smaller ν.

Acknowledgements

The author sincerely thanks Professor A.K.Md.E. Saleh for his kind advice and valuable discussion on the problem of improved estimation during a visit of the author to Carleton University, Canada. The author is also very grateful to Professor Bikas K Sinha of the ISI-Calcutta for pointing out a specification error and suggesting other improvements on an earlier version of the paper.

References

[1] Ahmed, S.E. (1992). Shrinkage preliminary test estimation in multivariate normal distributions. Jou. of Statist. Comput. Simul., Vol. 43, 177–195.

[2] Ahmed, S.E. and Saleh, A.K.Md.E. (1989). Pooling multivariate data. Jou. of Statist. Comput. Simul., Vol. 35, 209–226.

[3] 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.

[4] 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.

[5] Chmielewski, M.A. (1981). Elliptically symmetric distributions. International Statistical Review, Vol. 49, 67–74.

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

[7] Fisher, R.A. (1956). Statistical Methods in Scientific Inference. Oliver and Boyd, London.

[8] Fraser, D.A.S. (1979). Inference and Linear Models, McGraw–Hill, New York.

[9] Fraser, D.A.S. (1979). Inference and Linear Models. McGraw-Hill, New York

[10] Giles, J.A. (1991). Pre-testing for linear restrictions model with spherically symmetric disturbances. Journal of Econometrics, Vol. 50, 377–398.

[11] Gupta, A.K. and Varga, T. (1993). Elliptically Contoured Models in Statistics. Kluwer Academic Publishers, London.

[12] Han, C.P. an Bancroft, T.A. (1968). On pooling means when variance is unknown. Jou. Amer. Statist. Assoc., Vol. 63, 1333–1342.

[13] 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.

[14] Judge, G.G. and Bock, M.E. (1978). The Statistical Implications of Pre-test and Stein-rule Estimators in Econometrics. North-Holland, New York.

[15] Khan, S. (1997). Likelihood based inference of the mean vectors of two multivariate Student-t populations. Far East Jou. Theo. Statist. Vol. 1, 77-91.

[16] Khan, S. and Haq, M.S. (1990). Prediction distribution for the linear model with multivariate Student-t error, Comm. in Statist. Vol. 19, 4705–4712.

[17] 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. 39, 1-17.

[18] Khan, S. and Saleh, A.K.Md.E. (1995). Preliminary test estimators of the mean based on p-samples from multivariate Student-t populations. Bulletin of the International Statistical Institute. 50th Session of the ISI, Beijing, 599–600.

[19] Lange, K.L., Little, R.J.A. and Taylor, J.M.G. (1989). Robust statistical modeling using the t distribution. *Journal of the American Statistical Association*, **84**, 881-896.

[20] Maatta, J.M. and Casella, G. (1990). Developments in decision-theoretic variance estimation, Statistical Sciences, Vol. 5, 90–101.

[21] Prucha I.R. and Kelejian, H.H. (1984). The structure of simultaneous equation estimators: A generalization towards non-normal disturbances, Econometrica, Vol. 52, 721–736.

[22] 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.

[23] 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.

[24] Sclove, S.L., Morris, C. and Rao, C.R. (1972). Non-optimality of preliminary-test estimators for the mean of a multivariate normal distribution. Ann. Math. Statist. Vol. 43, 1481–1490.

[25] Singh, R.S. (1991). James-Stein rule estimators in linear regression models with multivariate-t errors. Australian Journal of Statistics, Vol. 33, 145–158.

[26] 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.

[27] Stein, C. (1981). Estimation of the mean of a multivariate normal distribution. Ann. Statist. Vol. 9, 1135–1151.

[28] Tabatabaey, S.M.M. (1995). Preliminary Test Approach Estimation: Regression Model With Spherically Symmetric Errors. Unpublished Ph.D. Thesis, Carleton University, Canada.

[29] Ullah, A. and Walsh, V.Z. (1984). On the robustness of ML, LR and W test in regression models. Econometrica. Vol. 52, 1055–1066.

[30] Zellner, A. (1976). Bayesian and non-Bayesian analysis of regression model with multivariate Student-t error term. Jou. Amer. Statist. Assoc., vol. 66, 601–616.

BLANK PAGE