Journal of Statistical Research *ISSN 0256 - 422 X*
2005, Vol. 39, No. 2, pp. 71–86
Bangladesh

ESTIMATION OF PARAMETERS OF THE SIMPLE MULTIVARIATE LINEAR MODEL WITH STUDENT-t ERRORS

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

SUMMARY

This paper considers estimation of the intercept and slope vector parameters of the simple multivariate linear regression model with Student-t errors in the presence of uncertain prior information on the value of the unknown slope vector. The unrestricted, restricted, preliminary test, shrinkage, and positive-rule shrinkage estimators are defined together with the expressions for the bias, quadratic bias, quadratic risk and mean squared errors (mse) functions of the estimators are derived. Comparison of the estimators is made using quadratic risk criterion. Based on the study we conclude that for $p \geq 3$ shrinkage estimators are recommended, and for $p \leq 2$, the preliminary test estimators are preferable.

Keywords: Multivariate regression model; Student-t errors; unrestricted, restricted, preliminary test, shrinkage and positive-rule shrinkage estimators; bias, quadratic risk, mean squared error and relative efficiency.

AMS 2000 Classification: Primary 62F30, Secondary 62H12 and 62F10.

1 Introduction

The simple multivariate regression model is a more general model than the commonly used linear regression model where there is only one value of the response variable corresponding to one value of the explanatory variable. It is used to analyse data from studies where there are more than one value of the response variable for a particular value of the explanatory variable. For example, if several patients are given the same dose of a medicine to observe any response of the subjects, then, for one particular value of the explanatory variable, there will be several values of the response variable from different subjects. The model can be applied to any other experimental or observational studies where multiple responses are generated for one value of the independent variable.

Estimation of parameters is a core procedure in the statistical inference. Traditional estimators, such as the the maximum likelihood estimator (MLE) under the normal model or least squared estimator (LSE), are based exclusively on the observed data and are known to be best linear unbiased estimators. Improved estimators are based on non-sample uncertain prior information and the sample data, and often have better statistical properties than the classical estimators. We define several improved estimators and derive their various properties. Also, we compare the performances of the estimators under different criteria.

In the classical approach, estimators of unknown parameters are defined exclusively on the sample data. Bancroft (1944) first used non-sample prior information in estimating the parameters. The inclusion of non-sample information to the estimation of parameters is likely to 'improve' the quality of the estimators. The natural expectation is that the inclusion of additional information would result in a better estimator. In some cases this may be true, but in many other cases the risk of worse consequences cannot be ruled out. A number of estimators have been introduced in the literature that, under particular situation, over performs the traditional exclusive sample data based unbiased estimators when judged by criteria such as the mean squared error and squared error loss function.

The use of the normal distribution to model the errors of linear model is under increasing criticism for its inability to model fat or heavier tailed distributions as well as being non-robust. Fisher (1956, p. 133) warned against the consequences of inappropriate use of the traditional normal model. Fisher (1960, p. 46) analyzed Darwin's data (cf. Box and Taio, 1992, p. 133) by using a non-normal model. Fraser and Fick (1975) analyzed the same data by the Student-t model. Zellner (1976) provided both Bayesian and frequentist analyses of the multiple regression model with Student-t errors. Fraser (1979) illustrated the robustness of the Student-t model. Prucha and Kelegian (1984) proposed an estimating equation for the simultaneous equation model with the Student-t errors. Ullah and Walsh (1984) investigated the optimality of different types of tests used in econometric studies for the multivariate Student-t model. The interested readers may refer to the more recent work of Singh (1988), Lange et al. (1989), Giles (1991), Anderson (1993), Spanos (1994), Khan (1997), and Khan (2006) for different applications of the Student-t models. For a wide range of applications of the Student-t models refer to Lange et al. (1989). Indeed, Zellner (1976) noted that the Student t model is more wider than the normal model, as the later model is a special case of the former.

Bancroft (1944) and later Han and Bancroft (1968) developed the idea of improved estimation. They introduced the preliminary test estimator that uses uncertain non-sample prior information (not in the form of prior distributions), in addition to the sample information. Stein (1956) introduced the Stein-rule (shrinkage) estimator for multivariate normal population that dominates the usual maximum likelihood estimators under the quadratic error loss function. In a series of papers Saleh and Sen (1978, 1985) explored the preliminary test approach to Stein-rule estimation. Many authors have contributed to this area, notably Sclove et al. (1972), Judge and Bock (1978), Stein (1981), and Maatta and Casella (1990), to mention a few. Khan and Saleh (1995, 1997), and Khan (2000) investigated the problem for a family of Student-t populations. However, the relative performance of the preliminary test and shrinkage estimators of the intercept and slope parameters of multivariate regression model with Student-t error has not been investigated.

It is well known that the LSE of the intercept and slope parameters is unbiased. We suggest alternative improved estimators of the parameters that may be biased but would have some superior statistical properties in terms of another more popular statistical criterion, namely the quadratic risk based on quadratic error loss function. In this process, we define four biased estimators: the restricted estimator (RE), the preliminary test estimator (PTE) as a linear combination of the LSE and the RE, and the shrinkage estimator (SE) using the preliminary test approach, and the positive-rule shrinkage estimator (PRSE). We investigate the bias, quadratic risk and the mean squared error functions analytically to compare the performance of the estimators. The relative efficiency of the estimators are also studied to determine which estimator dominates other estimators under any specific condition. The analysis reveals the fact that althought there is no uniformly superior estimator that beats the others, the SE dominates the other two biased estimator if the non-sample information regarding the value of β is not too far from its true value. In practice, the non-sample information is usually available from past experience or expert knowledge, and hence it is expected that such information will not be too far from the true value.

The next Section specifies the simple multivariate regression model. Some preliminaries and four alternative 'improved' estimators of the slope parameter are provided in Section 3. The expressions of bias functions of the estimators are obtained in Section 4. The quadratic risk and mean squared error functions are derived in Section 5. Analysis of quadratic risk functions is discuses in Section 6. Some concluding remarks are given in Section 7.

2 Simple Multivariate Regression Model with Student–t Errors

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 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 Chmielewiski (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). Zellner (1976) first introduced the regression model with Student-t errors.

The j^{th} sample responses vector from a multivariate linear regression model can be expressed in the following convenient form

$$\boldsymbol{Y}_j = \boldsymbol{\theta} + \boldsymbol{\beta} x_j + \boldsymbol{e}_j \quad \text{for } j = 1, 2, \cdots, n; \tag{2.1}$$

where $\boldsymbol{Y}_j = (y_{1j}, \ldots, y_{pj})'$ is a p-dimensional column vector of responses produced by or associated with a single value of the explanatory variable x_j, $\boldsymbol{\theta}$ is a column vector of p-dimensional intercept parameters, $\boldsymbol{\beta}$ is a p-dimensional column vector of the slope parameters, and $\boldsymbol{e}_j = (e_{1j}, \ldots, e_{pj})'$ is a p-dimensional column vector of errors. Assume the errors jointly follow a multivariate Student-t distribution. Such a distribution can be viewed as a mixture distribution of normal and inverted gamma distributions. To be more specific, let the above errors be independently and identically distributed as normal variables so that $\boldsymbol{e}_j \sim N_p(\boldsymbol{0},\ \tau^2\Sigma)$ for any given value of τ. Assuming that τ follows an inverted gamma distribution with parameters ν and scale $\sigma = 1$, having density function

$$f(\tau;\nu) = \frac{2}{\Gamma\left(\frac{\nu}{2}\right)} \left(\frac{\nu}{2}\right)^{\nu/2} (\tau)^{-(\nu+1)} e^{-\frac{\nu}{2\tau^2}}, \qquad \tau > 0 \tag{2.2}$$

where ν is the shape parameter. It is well known (cf. Khan, 1997) that the mixture distribution of the errors, $\boldsymbol{e}_j$ and τ is a p-dimensional Student-t distribution with shape ν, and location $\boldsymbol{0}$. We write $[\boldsymbol{e}_j|\tau] \sim N_p(\nu, \boldsymbol{0}, \tau^2\Sigma)$ and $[\boldsymbol{e}_j] \sim t_p(\nu, \boldsymbol{0}, \frac{\nu}{\nu-2}\Sigma)$ where $\Sigma_\tau = \tau^2\Sigma$ is the covariance matrix of the normal errors, $\boldsymbol{e}_j$, when the value of τ is fixed.

Thus the (unconditional) density of $\boldsymbol{y}_j$ becomes

$$p(\boldsymbol{y}_j|\boldsymbol{\theta},\boldsymbol{\beta},\Sigma) = \frac{\Gamma\left(\frac{p+\nu}{2}\right)}{[\pi\nu\Sigma]^{\frac{p}{2}}\,\Gamma\left(\frac{\nu}{2}\right)} \left[1 + \frac{1}{\nu}\sum_{j=1}^{n}\left(\boldsymbol{y}_j - \boldsymbol{\theta} - \boldsymbol{\beta}x_j\right)' \Sigma^{-1}\left(\boldsymbol{y}_j - \boldsymbol{\theta} - \boldsymbol{\beta}x_j\right)\right]^{-\frac{\nu+p}{2}}. \tag{2.3}$$

Note that $\mathrm{E}[\boldsymbol{y}_j] = \boldsymbol{\theta} + \boldsymbol{\beta}x_j$, $\mathrm{Var}[\boldsymbol{y}_j|\tau] = \tau^2\Sigma$, and $\mathrm{Var}[\boldsymbol{y}_j] = \frac{\nu}{\nu-2}\Sigma$.

3 Some Preliminaries and Alternative Estimators

Following Zellner (1976), the *unrestricted estimator* (UE) of $(\boldsymbol{\theta}'\boldsymbol{\beta}')'$ is given by

$$\begin{pmatrix} \tilde{\boldsymbol{\theta}}_n \\ \tilde{\boldsymbol{\beta}}_n \end{pmatrix} = \begin{pmatrix} \bar{\boldsymbol{Y}} - \tilde{\boldsymbol{\beta}}_n \bar{x} \\ \frac{1}{Q}\sum_{j=1}^{n}(x_j - \bar{x})(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}}) \end{pmatrix} \tag{3.1}$$

where $Q = \sum_{j=1}^{n}(x_j - \bar{x})^2$ in which $\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$ and $\bar{\boldsymbol{Y}} = \frac{1}{n}\sum_{j=1}^{n}\boldsymbol{Y}_j$. The above UE is either the least squares or maximum likelihood estimator.

It is easily verified that $(\tilde{\boldsymbol{\theta}}_n', \tilde{\boldsymbol{\beta}}_n')'$ is *unbiased* with covariance matrix

$$\Sigma^* \otimes \begin{pmatrix} \frac{1}{n} + \frac{\bar{x}^2}{Q} & -\frac{\bar{x}}{Q} \\ -\frac{\bar{x}}{Q} & \frac{1}{Q} \end{pmatrix} \quad \text{in which } \Sigma^* = \frac{\nu}{\nu - 2}\boldsymbol{\Sigma} \tag{3.2}$$

where $\otimes$ is the Kronecker product of the two matrices. Further, one can verify that the matrix $\boldsymbol{S}$ defined by

$$\boldsymbol{S} = (n-2)^{-1}\sum_{j=1}^{n}\Big[(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}}) - \tilde{\boldsymbol{\beta}}_n(x_j - \bar{x})\Big]'\Big[(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}}) - \tilde{\boldsymbol{\beta}}_n(x_j - \bar{x})\Big] \tag{3.3}$$

is unbiased for Σ^*. In addition, one can test the null hypothesis $H_0 : \boldsymbol{\beta} = \boldsymbol{0}$ against $H_A : \boldsymbol{\beta} \neq \boldsymbol{0}$ based on the test statistic

$$\mathcal{L}_n = \frac{m}{p} Q \tilde{\boldsymbol{\beta}}_n' \boldsymbol{S}^{-1} \tilde{\boldsymbol{\beta}}_n \tag{3.4}$$

which follows a central F–distribution with (p, m) d.f. in which $m = n - p - 1$, under H_0, and under H_A it follows the mixed distribution given by

$$G_{p,m}(C_\alpha; \Delta^*) = \sum_{r\geq 0} \frac{\Gamma\big(\frac{\nu}{2} + r\big)}{\Gamma(r+1)\Gamma(\nu/2)} \; \frac{\big(\frac{\Delta^*}{\nu-2}\big)^r I_{C_\alpha}\big(\frac{1}{2}(q+r); \frac{1}{2}m\big)}{\big(1 + \frac{\Delta^2}{\nu-2}\big)^{\nu/2+r}} \tag{3.5}$$

where $\Delta^* = \big(\frac{\nu}{\nu-2}\big) Q \boldsymbol{\beta}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\beta}$, $C_\alpha = \frac{pF_{p,m}(\alpha)}{m + pF_{p,m}(\alpha)}$ and $I_x(a; b)$ is the usual incomplete beta function. Further, let

$$G^{(j)}_{p+2i,m}(\Delta^*) = \sum_{r=0}^{\infty} \frac{\Gamma\big(\frac{\nu}{2} + r + j - 2\big)}{\Gamma(r+1)\Gamma\big(\frac{\nu}{2} + j - 2\big)} \; \frac{(\Delta^*/(\nu-2))^r I\big[\frac{1}{2}(p+2i) + r, \frac{1}{2}m\big]}{(1 + \Delta^*/(\nu-2))^{\nu/2+r+j-2}} \tag{3.6}$$

for $j = 1, 2$ and $i = 1, 2$. Also,

$$\begin{aligned} E^{(j)}\left[\chi^{-2}_{p+2i}(\Delta^*)\right] &= \sum_{r=0}^{\infty} \frac{\Gamma\big(\frac{\nu}{2} + r + j - 2\big)}{\Gamma(r+1)\Gamma\big(\frac{\nu}{2} + j - 2\big)} \frac{(\Delta^*/(\nu-2))^r (p + 2i - 2 + 2r)^{-1}}{(1 + \Delta^*/(\nu-2))^{\nu/2+r+j-2}} \\ E^{(j)}\big[\chi^{-2}_{p+2i}(0)\big] &= (p + 2i - 2)^{-1}, \quad j = 1, 2 \end{aligned} \tag{3.7}$$

and

$$\begin{aligned} & E^{(j)}\Big[F^{-1}_{p+2i,m}(\Delta^*) I\Big(F_{p+2i,m}(\Delta^*) \leq \frac{qd}{p+2i}\Big)\Big] \\ & \quad = \sum_{r=0}^{\infty} \frac{\Gamma\big(\frac{\nu}{2} + r + j - 2\big)}{\Gamma(r+1)\Gamma\big(\frac{\nu}{2} + j - 2\big)} \frac{(\frac{\Delta^*}{\nu-2})^r (p+2i)p + 2i - 2 + 2r)^{-1}}{(1 + \Delta^*/(\nu-2))^{\nu/2+r+j-2}} \\ & \qquad \times I_x\Big[\frac{1}{2}(p + 2i - 2 + 2r), \frac{1}{2}(m+2)\Big], \\ & E^{(j)}\Big[F^{-1}_{p+2i,m}(0) I\Big(F_{p+2i,m}(0) \leq \frac{qd}{p+2i}\Big)\Big] \\ & \quad = (p+2i)(p+2i-2)^{-1} I_x\Big[\frac{1}{2}(p+2i-2), \frac{1}{2}(m+2)\Big], \end{aligned} \tag{3.8}$$

where $x = \frac{qd}{m+qd}$ in which $d == \frac{(p-2)m}{p(m+2)}$ is a positive real number.

Now, we go to the expressions of restricted, preliminary test and shrinkage estimators of $(\boldsymbol{\theta}', \boldsymbol{\beta}')'$.

The restricted estimator (RE) of $(\boldsymbol{\theta}', \boldsymbol{\beta}')$ is given by

$$\begin{pmatrix} \hat{\boldsymbol{\theta}}_n \\ \hat{\boldsymbol{\beta}}_n \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{\theta}}_n + \tilde{\boldsymbol{\beta}}_n \bar{x} \\ \mathbf{0} \end{pmatrix}. \tag{3.9}$$

Following Saleh (2006) we define the preliminary test estimators (PTE) of $(\boldsymbol{\theta}', \boldsymbol{\beta}')'$ as

$$\begin{pmatrix} \hat{\boldsymbol{\theta}}_n^{\mathrm{PT}} \\ \hat{\boldsymbol{\beta}}_n^{\mathrm{PT}} \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{\theta}}_n + \tilde{\boldsymbol{\beta}}_n \bar{x}\, I(\mathcal{L}_n < F_{p,m}(\alpha) \\ \tilde{\boldsymbol{\beta}}_n I(\mathcal{L}_n \geq F_{p,m}(\alpha)) \end{pmatrix} \tag{3.10}$$

where $F_{p,m}(\alpha)$ is the α-level upper quantile of an F-distribution with p and m degrees of freedom. Similarly, the Stein-type shrinkage estimator (SE) of $(\boldsymbol{\theta}', \boldsymbol{\beta}')'$ is given by

$$\begin{pmatrix} \hat{\boldsymbol{\theta}}_n^{S} \\ \hat{\boldsymbol{\beta}}_n^{S} \end{pmatrix} = \begin{pmatrix} \tilde{\boldsymbol{\theta}}_n + \tilde{\boldsymbol{\beta}}_n \bar{x}(1 - d\mathcal{L}_n^{-1}) \\ \tilde{\boldsymbol{\beta}}_n(1 - d\mathcal{L}_n^{-1}) \end{pmatrix}, \text{ with } d = \frac{(p-2)m}{p(m+2)} \tag{3.11}$$

and the positive-rule shrinkage estimator (PRSE) of $(\boldsymbol{\theta}', \boldsymbol{\beta}')'$ is as follows:

$$\begin{pmatrix} \hat{\boldsymbol{\theta}}_n^{S+} \\ \hat{\boldsymbol{\beta}}_n^{S+} \end{pmatrix} = \begin{pmatrix} \tilde{\boldsymbol{\theta}}_n + \tilde{\boldsymbol{\beta}}_n \bar{x}(1 - d\mathcal{L}_n^{-1}) I(\mathcal{L}_n > d) \\ \tilde{\boldsymbol{\beta}}_n(1 - d\mathcal{L}_n^{-1}) I(\mathcal{L}_n > d) \end{pmatrix}. \tag{3.12}$$

In the next section we present the bias, quadratic bias functions, MSE matrices and quadratic risk functions of these five sets of estimators. The estimators belong to the class of quasi-empirical Bayes estimators of the form

$$\begin{pmatrix} \boldsymbol{\theta}_n^{*} \\ \boldsymbol{\beta}_n^{*} \end{pmatrix} = \begin{pmatrix} \tilde{\boldsymbol{\theta}}_n + \bar{x}\tilde{\boldsymbol{\beta}}_n g(\mathcal{L}_n) \\ \tilde{\boldsymbol{\beta}}_n g(\mathcal{L}_n) \end{pmatrix} \tag{3.13}$$

where $g(\mathcal{L}_n)$ takes the values 0, 1, $I(\mathcal{L}_n < F_{p,m}(\alpha))$ $(1 - d\mathcal{L}_n^{-1})$ and $(1 - d\mathcal{L}_n^{-1}) I(\mathcal{L}_n > d)$ respectively to yield the UE, RE, PTE, SE and PRSE of $(\boldsymbol{\theta}', \boldsymbol{\beta}')'$ respectively (cf Saleh, 2006).

4 The Bias and Quadratic Bias

In this section, the bias and quadratic bias expressions are given in the following theorem.

Theorem 4.1: For the simple multivariate linear model with Student-t errors having ν degrees of freedom the bias and quadratic bias functions of the UE, RE, PTE, SE and PRSE of the intercept and slope vectors are given by

$$\begin{pmatrix} \boldsymbol{b}_1(\tilde{\theta}_n) \\ \boldsymbol{b}_1(\tilde{\boldsymbol{\beta}}_n) \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} B_1(\tilde{\theta}_n) \\ B_1(\tilde{\boldsymbol{\beta}}_n) \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix} \tag{4.1}$$

$$\begin{pmatrix} \boldsymbol{b}_2(\hat{\theta}_n) \\ \boldsymbol{b}_2(\hat{\boldsymbol{\beta}}_n) \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta}\bar{x} \\ -\boldsymbol{\beta} \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} B_2(\hat{\theta}_n) \\ B_2(\hat{\boldsymbol{\beta}}_n) \end{pmatrix} = \begin{pmatrix} \bar{x}^2\Delta^* \\ \Delta^* \end{pmatrix} \tag{4.2}$$

$$\begin{pmatrix} \boldsymbol{b}_3(\hat{\theta}_n^{\mathrm{PT}}) \\ \boldsymbol{b}_3(\hat{\boldsymbol{\beta}}_n^{\mathrm{PT}}) \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta}\bar{x}G^{(2)}_{p+2,m}(\ell_\alpha;\Delta^*) \\ -\boldsymbol{\beta}\bar{x}G^{(2)}_{p+2,m}(\ell_\alpha;\Delta^*) \end{pmatrix}, \quad \ell_\alpha = \frac{p}{p+2}\,F_{p,m}(\alpha) \tag{4.3}$$

and

$$\begin{pmatrix} B_3(\hat{\theta}_n^{\mathrm{PT}}) \\ B_3(\hat{\boldsymbol{\beta}}_n^{\mathrm{PT}}) \end{pmatrix} = \begin{pmatrix} \bar{x}^2\Delta^*\{G^{(2)}_{p+2,m}(\ell_\alpha;\Delta^*)\}^2 \\ \Delta^*\{G^{(2)}_{p+2,m}(\ell_\alpha;\Delta^*)\}^2 \end{pmatrix} \tag{4.4}$$

$$\begin{pmatrix} \boldsymbol{b}_4(\hat{\theta}_n^{S}) \\ \boldsymbol{b}_4(\hat{\boldsymbol{\beta}}_n^{S}) \end{pmatrix} = \begin{pmatrix} dp\boldsymbol{\beta}\bar{x}E^{(2)}[\chi^{-2}_{p+2}(\Delta^*)] \\ -dp\boldsymbol{\beta}E^{(2)}[\chi^{-2}_{p+2}(\Delta^*)] \end{pmatrix}, \tag{4.5}$$

and

$$\begin{pmatrix} B_4(\hat{\theta}_n^{S}) \\ B_4(\hat{\boldsymbol{\beta}}_n^{S}) \end{pmatrix} = \begin{pmatrix} p^2d^2\bar{x}^2\Delta^*\{E^{(2)}[\chi^{-2}_{p+2}(\Delta^*)]\}^2 \\ p^2d^2\Delta^*\{E^{(2)}[\chi^{-2}_{p+2}(\Delta^*)]\}^2 \end{pmatrix}$$

$$\begin{pmatrix} \boldsymbol{b}_5(\hat{\theta}_n^{S+}) \\ \boldsymbol{b}_5(\hat{\boldsymbol{\beta}}_n^{S+}) \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta}\bar{x}\big\{E^{(2)}\big[(1-d_1F^{-1}_{p+2,m}(\Delta^*))I(F_{p+2,m}(\Delta^*)<d_1)\big] + E^{(2)}[F^{-1}_{p+2,m}(\Delta^*)]\big\} \\ -\boldsymbol{\beta}\big\{E^{(2)}\big[(1-d_1F^{-1}_{p+2,m}(\Delta^*))I(F_{p+2,m}(\Delta^*)<d_1)\big] + E^{(2)}[F^{-1}_{p+2,m}(\Delta^*)]\big\} \end{pmatrix},$$

and

$$\begin{pmatrix} B_5(\hat{\theta}_n^{S+}) \\ B_5(\hat{\boldsymbol{\beta}}_n^{S+}) \end{pmatrix} = \begin{pmatrix} \bar{x}^2\Delta^*\big\{E^{(2)}\big[(1-d_1F^{-1}_{p+2,m}(\Delta^*))I(F_{p+2,m}(\Delta^*)<d_1)\big] + E^{(2)}[F^{-1}_{p+2,m}(\Delta^*)]\big\}^2 \\ \Delta^*\big\{E^{(2)}\big[(1-d_1F^{-1}_{p+2,m}(\Delta^*))I(F_{p+2,m}(\Delta^*)<d_1)\big] + E^{(2)}[F^{-1}_{p+2,m}(\Delta^*)]\big\}^2 \end{pmatrix} \tag{4.6}$$

with $d_1 = \frac{dp}{p+2}$ where $\boldsymbol{b}_i$ and B_i for $i = 1, 2, \ldots, 5$ represent bias and quadratic bias functions of the estimators respectively.

We observe from the expressions of quadratic biases that

$$0 \leq B_3(\hat{\boldsymbol{\beta}}_n^{\mathrm{PT}}) \leq B_2(\hat{\boldsymbol{\beta}}_n) < \infty.$$

Moreover, the quadratic biases of $\hat{\boldsymbol{\beta}}_n^{S}$ and $\hat{\boldsymbol{\beta}}_n^{S+}$ satisfy the ordering

$$0 = B_1(\tilde{\boldsymbol{\beta}}_n) \leq B_4(\hat{\boldsymbol{\beta}}_n^{S}) \leq B_5(\hat{\boldsymbol{\beta}}_n^{S+}) < \infty.$$

For the comparison of the quadratic biases of $\hat{\boldsymbol{\beta}}_n^{\mathrm{PT}}$ and $\hat{\boldsymbol{\beta}}_n^{S}$ we note that

$$B_3(\hat{\boldsymbol{\beta}}_n^{\mathrm{PT}}) - B_4(\hat{\boldsymbol{\beta}}_n^{S}) = \Delta^* \left[\left\{ G_{p+2,m}^{(2)}(\ell_\alpha; \Delta^*) \right\}^2 - \left\{ E^{(2)} \left[\chi_{p+2}^{-2}(\Delta^*) \right] \right\}^2 \right] \geq 0$$

whenever $G_{p+2,m}^{(2)}(\ell_\alpha; \Delta^*) \geq E^{(2)} \left[\chi_{p+2}^{-2}(\Delta^*) \right]$ otherwise $B_4(\hat{\boldsymbol{\beta}}_n^{S}) > B_3(\hat{\boldsymbol{\beta}}_n^{\mathrm{PT}})$.

Similar conclusions hold for the biases of $\tilde{\boldsymbol{\theta}}_n$, $\hat{\boldsymbol{\theta}}_n$, $\hat{\boldsymbol{\theta}}_n^{\mathrm{PT}}$, $\hat{\boldsymbol{\theta}}_n^{S}$ and $\hat{\boldsymbol{\theta}}_n^{S+}$.

5 The MSE and Risk Expressions of the Estimators

The following theorem gives the expressions for the MSE matrices and quadratic risks of the estimators.

Theorem 5.1: *For the simple multivariate linear model with Student-t errors having ν degrees of freedom the MSE matrices and quadratic risk functions of the UE, RE, PTE, SE and PRSE of the intercept and slope vectors are given by*

(i) $\boldsymbol{M}_1(\tilde{\boldsymbol{\theta}}_n) = \left(\frac{1}{n} + \frac{\bar{x}^2}{Q} \right) \boldsymbol{\Sigma}^*$ and $R_1(\tilde{\boldsymbol{\theta}}_n; \boldsymbol{W}) = \left(\frac{1}{n} + \frac{\bar{x}^2}{Q} \right) \mathrm{tr}(\boldsymbol{W}\boldsymbol{\Sigma}^*)$

and

$$\boldsymbol{M}_1(\tilde{\boldsymbol{\beta}}_n) = \frac{1}{Q}\boldsymbol{\Sigma}^* \quad \text{and} \quad R_1(\tilde{\boldsymbol{\beta}}_n; \boldsymbol{W}) = \frac{1}{Q}\mathrm{tr}(\boldsymbol{W}\boldsymbol{\Sigma}^*).$$

(ii) $\boldsymbol{M}_2(\hat{\boldsymbol{\theta}}_n) = \left(\frac{1}{n} + \frac{\bar{x}^2}{Q}\Delta^* \right) \boldsymbol{\Sigma}^*$ and $R_2(\hat{\boldsymbol{\theta}}_n; \boldsymbol{W}) = \left(\frac{1}{n} + \frac{\bar{x}^2}{Q}\Delta^* \right) \mathrm{tr}(\boldsymbol{W}\boldsymbol{\Sigma}^*)$

and

$$\boldsymbol{M}_2(\hat{\boldsymbol{\beta}}_n) = \boldsymbol{\beta}\boldsymbol{\beta}' \quad \text{and} \quad R_2(\hat{\boldsymbol{\beta}}_n; \boldsymbol{W}) = \Delta^* = Q\boldsymbol{\beta}'\boldsymbol{\Sigma}*^{-1}\boldsymbol{\beta}.$$

(iii) $\boldsymbol{M}_3(\hat{\boldsymbol{\theta}}_n^{\mathrm{PT}}) = \left(\frac{1}{n} + \frac{\bar{x}^2}{Q} \right) \boldsymbol{\Sigma}^* - \frac{\bar{x}^2}{Q}\boldsymbol{\Sigma}^* G_{p+2,m}^{(2)}(\ell_\alpha; \Delta^*)$

$$+ \bar{x}^2 \boldsymbol{\beta}\boldsymbol{\beta}' \left\{ 2G_{p+2,m}^{(1)}(\ell_\alpha; \Delta^*) - G_{p+4,m}^{(1)}(\ell_\alpha^*; \Delta^*) \right\} \text{ where } \ell_\alpha^* = \tfrac{p}{p+4} F_{p,m}(\alpha).$$

$$R_3(\hat{\boldsymbol{\theta}}_n^{\mathrm{PT}};\boldsymbol{W}) = \left(\frac{1}{n}+\frac{\bar{x}^2}{Q}\right)\mathrm{tr}(\boldsymbol{W}\Sigma^2) - \frac{\bar{x}^2}{Q}\mathrm{tr}(\boldsymbol{W}\Sigma^*)G^{(2)}_{p+2,m}(\ell_\alpha;\Delta^*)$$
$$+\ \bar{x}^2(\boldsymbol{\beta}'\boldsymbol{W}\boldsymbol{\beta})\left\{2G^{(1)}_{p+2,m}(\ell_\alpha;\Delta^*) - G^{(1)}_{p+4,m}(\ell^*_\alpha;\Delta^*)\right\}.$$
$$\boldsymbol{M}_3(\hat{\boldsymbol{\beta}}_n^{\mathrm{PT}}) = \frac{1}{Q}\boldsymbol{\Sigma}^* - \frac{1}{Q}\boldsymbol{\Sigma}^* G^{(2)}_{p+2,m}(\ell_\alpha;\Delta^*) + \boldsymbol{\beta}\boldsymbol{\beta}'$$
$$\left\{2G^{(1)}_{p+2,m}(\ell_\alpha;\Delta^*) - G^{(1)}_{p+4,m}(\ell^*_\alpha;\Delta^*)\right\} \quad \text{and}$$
$$R_3(\hat{\boldsymbol{\beta}}_n^{\mathrm{PT}}) = \frac{1}{Q}\mathrm{tr}(\boldsymbol{W}\Sigma^*) - \frac{1}{Q}\mathrm{tr}(\boldsymbol{W}\Sigma^*)G^{(2)}_{p+2,m}(\ell_\alpha;\Delta^*)$$
$$+\ (\beta'\boldsymbol{W}\beta)\left\{2G^{(1)}_{p+2,m}(\ell_\alpha;\Delta^*) - G^{(1)}_{p+4,m}(\ell^*_\alpha;\Delta^*)\right\}.$$

(iv) $$\boldsymbol{M}_4(\hat{\boldsymbol{\theta}}_n^S) = \left(\frac{1}{n}+\frac{\bar{x}^2}{Q}\right)\boldsymbol{\Sigma}^* - dp\,\frac{\bar{x}^2}{Q}\boldsymbol{\Sigma}^*\Big\{2E^{(2)}\big[\chi^{-2}_{p+2}(\Delta^*)\big]$$
$$-\ (p-2)E^{(2)}\big[\chi^{-4}_{p+2}(\Delta^*)\big] + dp(p+2)\bar{x}^2\boldsymbol{\beta}\boldsymbol{\beta}'E^{(2)}\big[\chi^{-4}_{p+4}(\Delta^*)\big],$$
$$R_4(\hat{\boldsymbol{\theta}}_n^S;\boldsymbol{W}) = \left(\frac{1}{n}+\frac{\bar{x}^2}{Q}\right)\mathrm{tr}(\boldsymbol{W}\boldsymbol{\Sigma}^*) - dp\,\frac{\bar{x}^2}{Q}\mathrm{tr}(\boldsymbol{W}\boldsymbol{\Sigma}^*)\big\{2E^{(2)}\big[\chi^{-2}_{p+2}(\Delta^*)\big]$$
$$-\ (p-2)E^{(2)}\big[\chi^{-4}_{p+2}(\Delta^*)\big]\big\} + dp(p+2)\bar{x}^2(\boldsymbol{\beta}'\boldsymbol{W}\boldsymbol{\beta})E^{(2)}\big[\chi^{-4}_{p+4}(\Delta^*)\big]$$

and

$$\boldsymbol{M}_4(\hat{\boldsymbol{\theta}}_n^S) = \frac{1}{Q}\Sigma^* - \frac{1}{Q}dp\Sigma^*\big\{2E^{(2)}\big[\chi^{-2}_{p+2}(\Delta^*)\big] - (p-2)E^{(2)}\big[\chi^{-4}_{p+2}(\Delta^*)\big]\big\}$$
$$+\ dp(p+2)\beta\boldsymbol{\beta}'E^{(2)}\big[\chi^{-4}_{p+4}(\Delta^*)\big]$$

and

$$R_4(\hat{\boldsymbol{\beta}}_n^S;\boldsymbol{W}) = \frac{1}{Q}\mathrm{tr}(\boldsymbol{W}\boldsymbol{\Sigma}^*) - \frac{dp}{Q}\mathrm{tr}(\boldsymbol{W}\boldsymbol{\Sigma}^*)\big\{2E^{(2)}\big[\chi^{-2}_{p+2}(\Delta^*)\big]$$
$$-(p-2)E^{(2)}\big[\chi^{-4}_{p+2}(\Delta^*)\big]\big\} + dp(p+2)\bar{x}^2(\boldsymbol{\beta}'\boldsymbol{W}\boldsymbol{\beta})E^{(2)}\big[\chi^{-4}_{p+4}(\Delta^*)\big].$$

(v) $$\boldsymbol{M}_5(\hat{\boldsymbol{\theta}}_n^{S+}) = \boldsymbol{M}_4(\hat{\boldsymbol{\theta}}_n^S) - \frac{\bar{x}^2}{Q}\boldsymbol{\Sigma}^*E^{(1)}\Big[\big(1-d_1F^{-1}_{p+2,m}(\Delta^*)\big)^2 I\big(F_{p+2,m}(\Delta^*)<d_1\big)\Big]$$
$$+\ \bar{x}^2\boldsymbol{\beta}\boldsymbol{\beta}'\big\{2E^{(2)}\big[\big(1-d_1F^{-1}_{p+2,m}(\Delta^*)\big)I\big(F_{p+2,m}(\Delta^*)<d_1\big)\big]$$
$$-E^{(2)}\big[\big(1-d_2F^{-1}_{p+4,m}(\Delta^*)\big)^2 I\big(F_{p+4,m}(\Delta^*)<d_2\big)\big]\big\}$$

with $d_1 = \frac{dp}{p+2}$ and $d_2 = \frac{dp}{p+4}$,

$$R_5(\hat{\boldsymbol{\theta}}_n^{S+};\boldsymbol{W}) = R_4(\hat{\boldsymbol{\theta}}_n^S;\boldsymbol{W}) - \frac{\bar{x}^2}{Q}\mathrm{tr}(\boldsymbol{W}\Sigma^*)E^{(1)}\big[\big(1-d_1F^{-1}_{p+2,m}(\Delta^*)\big)^2 I\big(F_{p+2,m}(\Delta^*)<d_1\big)\big]$$
$$+(\boldsymbol{\beta}'\boldsymbol{W}\boldsymbol{\beta})\big\{2E^{(2)}\big[\big(1-d_1F^{-1}_{p+2,m}(\Delta^*)\big)I\big(F_{p+2,m}(\Delta^*)<d_1\big)\big]$$
$$-E^{(2)}\big[\big(1-d_2F^{-1}_{p+2,m}(\Delta^*)\big)^2 I\big(F_{p+4,m}(\Delta^*)<d_2\big)\big]\big\}.$$

Similarly,

$$\boldsymbol{M}_4(\hat{\boldsymbol{\beta}}_n^{S+}) = \boldsymbol{M}_4(\hat{\boldsymbol{\beta}}_n^{S}) - \frac{1}{Q}\Sigma^* E^{(1)}\left[(1-d_1F_{p+2,m}^{-1}(\Delta^*)\big)^2 I\big(F_{p+2,m}(\Delta^*) < d_1\big)\right]$$
$$+\bar{x}^2\boldsymbol{\beta}\boldsymbol{\beta}'\Big\{2E^{(2)}\left[(1-d_1F_{p+2,m}^{-1}(\Delta^*)\big)I\big(F_{p+2,m}(\Delta^*) < d_1\big)\right]$$
$$-E^{(2)}\left[\big(1-d_2F_{p+4,m}^{-1}(\Delta^*)\big)I\big(F_{p+4,m}(\Delta^*) < d_2\big)\right]\Big\}.$$

6 Analysis of Quadratic Risks

It is clear that the risks of $\tilde{\boldsymbol{\theta}}_n$ and $\tilde{\boldsymbol{\beta}}_n$ are constant while that of $\hat{\boldsymbol{\theta}}_n$ and $\hat{\boldsymbol{\beta}}_n$ depend on Δ^* and are monotonic functions of Δ^*. So, as $\Delta^* \to \infty$, they are unbounded. For $\boldsymbol{W} = \boldsymbol{\Sigma}*^{-1}$, we then have

$$R_2\big(\hat{\boldsymbol{\theta}}_n; \boldsymbol{\Sigma}^{*-1}\big) = p\Big(\frac{1}{n} + \frac{\bar{x}^2}{Q}\Delta^*\Big) \quad \text{and} \quad R_2\big(\hat{\boldsymbol{\beta}}_n; \Sigma^{*-1}\big) = p/Q. \tag{6.1}$$

Further,

$$\left[\Big(\frac{1}{n} + \frac{\bar{x}^2}{Q}\Delta^*\Big) - \Big(\frac{1}{n} + \frac{\bar{x}^2}{Q}\Big)\right]\mathrm{tr}(\boldsymbol{W}\boldsymbol{\Sigma}^*) = \frac{\bar{x}^2}{Q}(\Delta^* - 1)\mathrm{tr}(\boldsymbol{W}\boldsymbol{\Sigma}^*). \tag{6.2}$$

Hence, $\hat{\boldsymbol{\theta}}_n$ is better than $\tilde{\theta}_n$ if $\Delta^* < 1$ and $\tilde{\theta}_n$ is better than $\hat{\theta}_n$ if $\Delta^* > 1$. Similar conclusions hold for $\hat{\boldsymbol{\beta}}_n$ and $\tilde{\boldsymbol{\beta}}_n$, i.e.

$$R_2(\hat{\boldsymbol{\beta}}_n; \boldsymbol{W}) - R_1(\tilde{\boldsymbol{\beta}}_n; \boldsymbol{W}) = \Delta^* - \frac{1}{Q}\mathrm{tr}(\boldsymbol{W}\boldsymbol{\Sigma}^*)$$

implies that $\hat{\boldsymbol{\beta}}_n$ is better than $\tilde{\boldsymbol{\beta}}_n$ if $\Delta^* < \frac{1}{Q}\mathrm{tr}(\boldsymbol{W}\boldsymbol{\Sigma}^*)$ and $\tilde{\boldsymbol{\beta}}_n$ is better than $\hat{\boldsymbol{\beta}}_n$ if $\Delta^* > \frac{1}{Q}\mathrm{tr}(\boldsymbol{W}\boldsymbol{\Sigma}^*)$.

Now, consider the comparison of $\tilde{\boldsymbol{\theta}}_n$ and $\hat{\boldsymbol{\theta}}_n^{\mathrm{PT}}$, and $\tilde{\boldsymbol{\beta}}_n$ and $\hat{\boldsymbol{\beta}}_n^{\mathrm{PT}}$. First, we consider $\tilde{\boldsymbol{\beta}}_n$ and $\hat{\boldsymbol{\beta}}_n^{\mathrm{PT}}$. The risk-difference $R_1(\tilde{\boldsymbol{\beta}}_n; W) - R_3(\hat{\boldsymbol{\beta}}_n^{\mathrm{PT}}; \boldsymbol{W})$ which is

$$\frac{1}{Q}\mathrm{tr}(\boldsymbol{W}\boldsymbol{\Sigma}^*)G_{p+2,m}^{(2)}(\ell_\alpha; \Delta^*) - (\boldsymbol{\beta}'\boldsymbol{W}\boldsymbol{\beta})\left\{2G_{p+2,m}^{(1)}(\ell_\alpha; \Delta^*) - G_{p+4,m}^{(1)}(\ell_\alpha^*; \Delta^*)\right\} \gtreqless 0$$

whenever

$$Q(\boldsymbol{\beta}'\boldsymbol{W}\boldsymbol{\beta}) \lesseqgtr \frac{\mathrm{tr}(\boldsymbol{W}\boldsymbol{\Sigma}^*)G_{p+2,m}^{(2)}(\ell_\alpha; \Delta^*)}{\left[2G_{p+2,m}^{(1)}(\ell_\alpha; \Delta^*) - G_{p+4,m}^{(1)}(\ell_\alpha^*; \Delta^*)\right]}.$$

If $\boldsymbol{W} = \boldsymbol{\Sigma}^{*-1}$, then the above reduces to

$$\Delta^* \lesseqgtr \frac{pG_{p+2,m}^{(2)}(\ell_\alpha; \Delta^*)}{\left[2G_{p+2,m}^{(1)}(\ell_\alpha; \Delta^*) - G_{p+4,m}^{(1)}(\ell_\alpha^*; \Delta^*)\right]}.$$

Hence, $\hat{\boldsymbol{\beta}}_n^{\mathrm{PT}}$ is better than $\tilde{\boldsymbol{\beta}}_n$ if

$$\Delta^* < \frac{pG^{(2)}_{p+2,m}(\ell_\alpha;\Delta^*)}{\big[2G^{(1)}_{p+2,m}(\ell_\alpha;\Delta^*) - G^{(1)}_{p+4,m}(\ell^*_\alpha;\Delta^*)\big]},$$

otherwise $\tilde{\boldsymbol{\beta}}_n$ is better.

Similar conclusions follow for $\tilde{\boldsymbol{\theta}}_n$ and $\hat{\boldsymbol{\theta}}_n^{\mathrm{PT}}$. Consider the risk-difference of the estimators

$$\begin{aligned}&R_1(\tilde{\boldsymbol{\theta}}_n;\boldsymbol{W}) - R_3\big(\hat{\boldsymbol{\theta}}_n^{\mathrm{PT}};\boldsymbol{W}\big)\\&= \frac{\bar{x}^2}{Q}\mathrm{tr}(\boldsymbol{W\Sigma}^*)G^{(2)}_{p+2,m}(\ell_\alpha;\Delta^*) - \bar{x}^2(\boldsymbol{\beta}'\boldsymbol{W}\boldsymbol{\beta})\big[2G^{(1)}_{p+2,m}(\ell_\alpha;\Delta^*) - G^{(1)}_{p+4,m}(\ell^*_\alpha;\Delta^*)\big].\end{aligned}$$

Hence, the risk-difference is $\gtreqless$ whenever

$$Q(\boldsymbol{\beta}'\boldsymbol{W}\boldsymbol{\beta}) \gtreqless \frac{\mathrm{tr}(\boldsymbol{W\Sigma}^*)G^{(2)}_{p+2,m}(\ell_\alpha;\Delta^*)}{\big[2G^{(1)}_{p+2,m}(\ell_\alpha;\Delta^*) - G^{(1)}_{p+4,m}(\ell^*_\alpha;\Delta^*)\big]}.$$

Hence, $\hat{\boldsymbol{\theta}}_n^{\mathrm{PT}}$ is better than $\tilde{\boldsymbol{\theta}}_n$ if

$$Q(\boldsymbol{\beta}'\boldsymbol{W}\boldsymbol{\beta}) \le \frac{\mathrm{tr}(\boldsymbol{W\Sigma}^*)G^{(2)}_{p+2,m}(\ell_\alpha;\Delta^*)}{\big[2G^{(1)}_{p+2,m}(\ell_\alpha;\Delta^*) - G^{(1)}_{p+4,m}(\ell^*_\alpha;\Delta^*)\big]}.$$

otherwise $\tilde{\boldsymbol{\theta}}_n$ is better than $\hat{\boldsymbol{\theta}}_n^{\mathrm{PT}}$.

Now, consider the risk-difference

$$\begin{aligned}R_1(\tilde{\boldsymbol{\beta}}_n;\boldsymbol{W}) - R_4(\hat{\boldsymbol{\beta}}_n^S;\boldsymbol{W}) &= \frac{dp}{Q}\mathrm{tr}(\boldsymbol{W\Sigma}^*)\Big\{2E^{(2)}\big[\chi^{-2}_{p+2}(\Delta^*)\big] - (p-2)E^{(2)}\big[\chi^{-4}_{p+2}(\Delta^*)\big]\Big\}\\&\quad -dp(p+2)(\boldsymbol{\beta}'\boldsymbol{W}\boldsymbol{\beta})E^{(2)}\big[\chi^{-4}_{p+4}(\Delta^*)\big]\\R_1(\tilde{\boldsymbol{\theta}}_n;\boldsymbol{W}) - R_4(\hat{\boldsymbol{\theta}}_n^S;\boldsymbol{W}) &= \frac{dp}{Q}\mathrm{tr}(W\Sigma^*)\Big\{(p-2)E^{(1)}\big[\chi^{-4}_{p+2}(\Delta^*)\big]\\&\quad +\Big[1 - \frac{(p+2)(\boldsymbol{\beta}'\boldsymbol{W}\boldsymbol{\beta})}{2\Delta^*\mathrm{tr}(\boldsymbol{W}\Sigma^*)}\Big]2\Delta^* E^{(2)}\big[\chi^{-4}_{p+4}(\Delta^*)\big]\Big\}.\end{aligned}$$

The risk-difference is positive for all $\mathcal{A}$ such that

$$\left\{\mathcal{A} : \frac{\mathrm{tr}(\boldsymbol{W}\Sigma^*)}{\mathrm{Ch}_{\max}(\boldsymbol{W}\Sigma^*)} \ge \frac{p+2}{2}\right\}$$

where $\mathrm{Ch}_{\max}$ is the maximum characteristic root of the matrix $(\boldsymbol{W}\Sigma^*)$. Thus, $\hat{\boldsymbol{\theta}}_n^S$ uniformly dominates $\tilde{\boldsymbol{\theta}}_n$ as well as $\hat{\boldsymbol{\beta}}_n^S$ uniformly dominates $\tilde{\boldsymbol{\beta}}_n$.

To compare $\hat{\boldsymbol{\theta}}_n$ and $\hat{\boldsymbol{\theta}}_n^S$ we may write

$$R_4(\hat{\boldsymbol{\theta}}_n^S; \boldsymbol{W}) - R_2(\hat{\theta}_n; \boldsymbol{W}) = \frac{\bar{x}^2}{Q}(1-\Delta^*)\mathrm{tr}(\boldsymbol{W}\Sigma^*)$$
$$+dp\frac{\bar{x}^2}{Q}\mathrm{tr}(\boldsymbol{W}\Sigma^*)\big\{(p-2)E^{(2)}\big[\chi_{p+2}^{-2}(\Delta^*)\big] + \Delta^* E^{(2)}\big[\chi_{p+4}^{-4}(\Delta^*)\big]\big\}$$
$$+dp(p+2)\bar{x}^2(\boldsymbol{\beta}'\boldsymbol{W}\boldsymbol{\beta})E^{(2)}\big[\chi_{p+4}^{-4}(\Delta^*)\big].$$

Under H_0, the above risk-difference becomes

$$\frac{\bar{x}^2}{Q}\mathrm{tr}(\boldsymbol{W}\boldsymbol{\Sigma}^*) + dp(p-2)\frac{\bar{x}^2}{Q}E^{(2)}\big[\chi_{p+2}^{-2}(0)\big] \geq R_2(\hat{\boldsymbol{\theta}}_n; \boldsymbol{W})$$

while under H_A

$$R_2(\hat{\boldsymbol{\theta}}_n; \boldsymbol{W}) = \frac{1}{n}\mathrm{tr}(\boldsymbol{W}\boldsymbol{\Sigma}^*) \leq R_1(\tilde{\theta}_n; \boldsymbol{W}).$$

Thus, $\hat{\boldsymbol{\theta}}_n$ performs better than $\tilde{\boldsymbol{\theta}}_n$ under H_0. However, as $\boldsymbol{\beta}$ moves away from the origin Δ^* increases and the risk of $\hat{\boldsymbol{\theta}}_n$ becomes unbounded while the risk of $\hat{\boldsymbol{\theta}}_n^S$ remains below the risk of $\tilde{\boldsymbol{\theta}}_n$ and the two merge as $\Delta^* \to \infty$. Thus $\hat{\boldsymbol{\theta}}_n^S$ dominates $\hat{\theta}_n$ outside an interval around the origin.

Similarly,

$$R_4(\hat{\boldsymbol{\beta}}_n^S; \boldsymbol{W}) - R_2(\hat{\boldsymbol{\beta}}_n; \boldsymbol{W})$$
$$= \frac{1}{Q}\mathrm{tr}(\boldsymbol{W}\Sigma^*)\Big[1 - dp\Big\{(p-2)E^{(2)}\big[\chi_{p+2}^{-2}(\Delta^*)\big] + 2\Delta^* E^{(2)}\big[\chi_{p+4}^{-4}(\Delta^*)\big]\Big\}\Big]$$
$$+\ dp(p+2)(\boldsymbol{\beta}'\boldsymbol{W}\boldsymbol{\beta})E^{(2)}\big[\chi_{p+4}^{-4}(\Delta^*)\big].$$

Under H_0, the above becomes

$$\frac{1}{Q}\mathrm{tr}(W\Sigma^*)\big[1 - dp(p+2)E^{(2)}\big[\chi_{p+2}^{-2}(\Delta^*)\big]\big]$$

which, under H_0

$$R_2(\hat{\boldsymbol{\beta}}_n; \boldsymbol{W}) = 0 \leq R_1(\tilde{\boldsymbol{\beta}}_n; \boldsymbol{W}).$$

Thus, $\boldsymbol{\beta}_n$ performs better than $\tilde{\boldsymbol{\beta}}_n^S$ under H_0.

However, as $\boldsymbol{\beta}$ moves away from the origin $\mathbf{0}$, $\boldsymbol{\beta}'\boldsymbol{W}\boldsymbol{\beta}$ as well as Δ^* increases and the risk of $\hat{\boldsymbol{\beta}}_n$ becomes unbounded while the risk of $\hat{\boldsymbol{\beta}}_n^S$ remains below the risk of $\tilde{\boldsymbol{\beta}}_n$ and the two merge as $\Delta^* \to \infty$. Thus, $\hat{\boldsymbol{\beta}}_n^S$ dominates $\hat{\boldsymbol{\beta}}_n$ outside an interval around the origin.

Now, we compare $\hat{\boldsymbol{\theta}}_n^S$ and $\hat{\boldsymbol{\theta}}_n^{S+}$ (also $\hat{\boldsymbol{\beta}}_n^S$ and $\hat{\boldsymbol{\beta}}_n^{S+}$). The risk-difference is

$$R_4(\hat{\theta}_n^S; \boldsymbol{W}) - R_5(\hat{\theta}_n^{S+}; \boldsymbol{W})$$
$$= \frac{\bar{x}^2}{Q}\mathrm{tr}(\boldsymbol{W}\Sigma^*)E^{(1)}\big[\big(1 - d_1 F_{p+2,m}^{-1}(\Delta^*)\big)^2 I\big(F_{p+2,m}(\Delta^*) < d_1\big)\big]$$
$$= \bar{x}^2\big[R_4(\hat{\boldsymbol{\beta}}_n^S; bW) - R_5(\hat{\boldsymbol{\beta}}_n^{S+}; \boldsymbol{W})\big] \geq 0.$$

Hence, the relative dominance of ($\hat{\boldsymbol{\theta}}_n^{S+}$, $\hat{\boldsymbol{\theta}}_n^{S}$ and $\tilde{\boldsymbol{\theta}}_n$) is given by

$$R_5(\hat{\boldsymbol{\theta}}_n^{S+};\boldsymbol{W}) \le R_4(\hat{\boldsymbol{\theta}}_n^{S};\boldsymbol{W}) \le R_1(\tilde{\theta}_n;\boldsymbol{W}) \;\; \forall \; \Delta^*.$$

Similarly, the relative dominance of $\hat{\boldsymbol{\beta}}_n^{S+}$, $\hat{\boldsymbol{\beta}}_n^{S}$ and $\tilde{\boldsymbol{\beta}}_n$ is given by

$$R_5(\hat{\boldsymbol{\beta}}_n^{S+};\boldsymbol{W}) \le R_4(\hat{\boldsymbol{\beta}}_n^{S};\boldsymbol{W}) \le R_1(\tilde{\boldsymbol{\beta}}_n;\boldsymbol{W}) \;\; \forall \; \Delta^*.$$

Finally we compare $\hat{\boldsymbol{\theta}}_n^{\mathrm{PT}}$ and $\hat{\boldsymbol{\theta}}_n^{S}$. Note that under H_0

$$R_4(\hat{\boldsymbol{\theta}}_n^{S};\boldsymbol{W}) = R_3(\hat{\boldsymbol{\theta}}_n^{\mathrm{PT}};\boldsymbol{W}) + \frac{\bar{x}^2}{Q}\left[G_{p+2,m}^{(1)}(\ell_\alpha;0) - d\right] \ge R_3(\hat{\theta}_n^{\mathrm{PT}};\boldsymbol{W})$$

where $G_{p+2,m}^{(1)}(\ell_\alpha;0) = 1-\alpha$ and $\ell_\alpha = pF_{p,m}(\alpha)/p+2$. Note ℓ_α satisfies the inequality

$$\left\{\alpha : \ell_\alpha \ge F_{p+2,m}^{-1}(\alpha;0)\right\}.$$

Thus, the risk of $\hat{\boldsymbol{\theta}}_n^{\mathrm{PT}}$ is smaller than that of $\hat{\boldsymbol{\theta}}_n^{S}$ when ℓ_α satisfies the above inequality. This implies that $\hat{\boldsymbol{\theta}}_n^{S}$ does not always dominate $\hat{\theta}_n^{\mathrm{PT}}$ under H_0. We may order the quadratic risks under H_0 as

$$R_2(\hat{\boldsymbol{\theta}}_n;\boldsymbol{W}) \le R_3(\hat{\boldsymbol{\theta}}_n^{\mathrm{PT}};\boldsymbol{W}) \le R(\hat{\boldsymbol{\theta}}_n^{S};\boldsymbol{W}) \le R_1(\tilde{\boldsymbol{\theta}}_n;\boldsymbol{W}).$$

The relative dominance picture changes as Δ^* diverts from 0. As $\Delta^* \to \infty$, the risks of $\hat{\boldsymbol{\theta}}_n^{S}$ and $\hat{\boldsymbol{\theta}}_n^{\mathrm{PT}}$ converges to the risk of $\tilde{\boldsymbol{\theta}}_n$ but for reasonably small values of Δ^* near 0, $\hat{\theta}_n^{\mathrm{PT}}$ dominates $\hat{\boldsymbol{\theta}}_n^{S}$ for a solution α from the set $\{\alpha : \ell_\alpha \ge F_{p+2,m}^{-1}(\alpha,0)\}$. Thus, none of the estimators uniformly dominate each other for $p \ge 3$.

Similarly, under H_0

$$R_4(\hat{\boldsymbol{\beta}}_n^{S};\boldsymbol{W}) = R_3(\hat{\boldsymbol{\beta}}_n^{\mathrm{PT}};\boldsymbol{W}) + \frac{1}{Q}\left[G_{p+2,m}^{(2)}(\ell_\alpha;0) - d\right]$$

and the analysis is similar to the risks of $\hat{\boldsymbol{\theta}}_n^{\mathrm{PT}}$ and $\hat{\boldsymbol{\theta}}_n^{S}$ follows and none of the estimators uniformly dominate each other for $p \ge 3$. Analysis of the MSE matrices is also possible, but not included in this paper.

7 Conclusions

In this paper we have studied five different estimators of $(\boldsymbol{\theta}', \boldsymbol{\beta}')'$ when the hypothesis $H_0 : \boldsymbol{\beta} = \mathbf{0}$ is suspected to hold. Note that the estimators $(\hat{\boldsymbol{\theta}}_n^{\prime S}, \hat{\boldsymbol{\beta}}_n^{\prime S})$ and $(\hat{\boldsymbol{\theta}}_n^{\prime S+}, \hat{\boldsymbol{\beta}}_n^{\prime S+})'$ are constrained by $p \ge 3$ while $(\hat{\boldsymbol{\theta}}_n^{\prime \mathrm{PT}}, \hat{\boldsymbol{\beta}}_n^{\prime \mathrm{PT}})'$ does not need to satisfy this constraint but needs the determination of the size of the test and depends on the diversion parameter Δ^*. Also maximal saving in risk for the shrinkage estimators in the normal case is $\frac{(p-2)m}{p(m+2)}$. In the case of Student-t models the maximal risk saving is $\frac{(p-2)m}{p(m+2)}\left(\frac{\nu-2}{\nu}\right)$. Thus, for moderate and small values of ν the risk-saving is significant compared to the normal case. The risk saving in the case of the PTE in the normal case is $G^{(2)}_{p+2,m}(\ell_\alpha; 0)$ and for the Student-t errors it is $\frac{\nu-2}{\nu}G^{(2)}_{p+1,m}(\ell_\alpha; 0)$ which converges to $G^{(2)}_{p+2,m}(\ell_\alpha; 0)$ as $\nu \to \infty$. Thus, the performance of the five estimators is robust in the case of the Student-t errors which is determined by the d.f. ν (≥ 4). For $p \geq 3$ one uses the $(\hat{\boldsymbol{\theta}}_n^{S+'}, \hat{\boldsymbol{\beta}}_n^{S+'})'$ while for $p < 3$ the PTE is used with optimum size of α. Therefore, from the point of robust efficiency both the PTE and the positive-rule Stein-type estimators may be advocated for application.

7.1 Acknowledgements

The author is very thankful to Professor A K Md E Saleh, Carleton University, Canada, for his valuable advice and comments that significantly improved the quality and presentation of the manuscript.

References

[1] 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.

[2] Bancroft, T.A. (1944). On biases in estimation due to the use of the preliminary tests of significance. Ann. Math. Statist., **15**, 190–204.

[3] Chiou, P. and Saleh, A.K.Md.E. (2002). Preliminary test confidence sets for the mean of a multivariate normal distribution. Journal of Propagation in Probability and Statistics, **2**, 177-189.

[4] Box, G.E.P. and Taio, G.C. (1990). Bayesian Inference in Statistical Analysis. Wiley, New York.

[5] Cornish, E.A. (1954). Multivariate t-distribution associated with a set of normal sample deviates. Australian Journal of Physics. Vol. 7, 531–542.

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

[7] Fisher, R.A. (1960). The Design of Experiments. 7th ed., Hafner, New York.

[8] Fraser, D.A.S. and Fick, G.H. (1975). Necessary analysis and its implementation. In Proc. Symposium on Statistics and Related Topics. ed. by A.K. Md. E. Saleh, Carleton University, Canada, 5.01–5.30.

[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] Han, C.P. and Bancroft, T.A. (1968). On pooling means when variance is unknown. Jou. Amer. Statist. Assoc., **63**, 1333-1342.

[12] James, W. and Stein, C. (1961). Estimation with quadratic loss. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability. Berkeley, California, 361–379.

[13] Judge, G.C. and Bock, M.E. (1978). Statistical Implications of Pre-test and Stein-rule Estimators in Econometrics. North Holland, Amsterdam.

[14] Khan, S. (1997). Likelihood based inference on the mean vectors of two multivariate Student-t populations. Far East Journal of Theoretical Statistics. 1, 1-17.

[15] Khan, S. (2000). Improved estimation of the mean vector for Student-t model. Communications In Statistics: Theory & Methods 29, (3), 507-527.

[16] Khan, S. (2006). Optimal Tolerance Regions for Some Functions of Multiple Regression Model with Student-t Errors. Journal of Statistics & Management Sciences, Forthcoming

[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. Vol. 39, 131–147.

[18] Khan, S. and Saleh, A.K.Md.E. (1998). Comparison of estimators of the mean based on p-samples from multivariate Student-t population. Communications in Statistics: Theory & Methods. Vol. 27(1), 193-210.

[19] Khan, S. and Saleh, A.K.Md.E. (2001). On the comparison of the pre-test and shrinkage estimators for the univariate normal mean. Statistical Papers, **42**(4), 451-473.

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

[21] 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.

[22] Saleh, A K Md E. (2006). Theory of Preliminary Test and Stein-type Estimation with Applications. Wiley, New York.

[23] 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.

[24] Spanos, A. (1994). On modeling heteroskedasticity: The Student's t and elliptical linear regression models. Econometric Theory, Vol. 10, 286–315.

[25] 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.

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

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

[28] Zellner, A. (1976). Bayesian and non-Bayesian analysis of the regression model with multivariate Student-t error term. Journal of American Statistical Association, Vol. 66, 601–616.