Journal of Applied Probability and Statistics
2020, Vol. 15, No. 1, pp. 99-109

PRELIMINARY TEST METHODOLOGY IN VIEW OF ASYMMETRY WITH ELLIPTICALLY CONTOURED DISTRIBUTION

M. Arashi
Faculty of Mathematical Sciences, Shahrood University of Technology, Iran
Email: m-arashi-stat@yahoo.com

Z. Hoque
Department of Analytics and Digital Era, College of Business and Economics, UAE University, P.O. Box 15551, Al Ain, UAE
Email: Zahirul.Hoque@uaeu.ac.ae

SUMMARY

This paper studies the preliminary test estimator (PTE) of the parameters of elliptically contoured distribution. The risk functions of the PTE and the sample information based unrestricted estimator have been derived under asymmetric linex loss function. These risk functions have been analysed both analytically and graphically to compare with each other. It has been revealed that under certain conditions PTE outperforms the sample information based estimator.

Keywords and phrases: Elliptically contoured distribution, Preliminary test estimator, Linex loss and risk function.

2010 Mathematics Subject Classification: Primary 62F03, secondary 62J05.

1 Introduction

The class of elliptical distributions contains a wide variety of distributions. It includes the multivariate normal (multinormal, henceforth) distribution as a special case, as well as many non-normal distributions including Cauchy, multivariate exponential, multivariate elliptical analogue of Student's t and non-normal variance mixtures of multinormal distributions. Elliptical distributions are symmetric and unimodal but are not constrained regarding kurtosis. This class of distributions has recently gained a lot of attention to the practitioners of statistics in different areas, especially in financial statistics. Particular use of this distribution is visible in risk management. Therefore, the estimation of parameters of elliptical distribution is of interest to its users in many areas.

In statistical literature different estimation strategies are suggested. Sample information based maximum likelihood estimator (MLE) is the best linear unbiased estimator and one of the most popular estimators. However, if the objective of any study is to minimize the risk of the estimator, MLE may not be an ideal choice as there are other estimators that outperform MLE. Khan *et al.* (2005) compared five different estimators of the intercept parameter of simple linear regression model under squared error loss and suggested that the least squares estimator (LSE) is not admissible as there are other estimators that outperform the LSE. Khan (2005) studied the different estimation strategies of parameters of the simple multivariate linear model with Student-t error with squared error loss. Later Khan (2008) investigated the estimation strategies of the intercept parameters of two linear regression models with normal errors, when it is a priori suspected that the two regression lines are parallel. An extensive investigation on the properties of alternative estimators of the parameters of various models is available in Saleh (2006).

Risk of the estimator depends on the choice of the loss function. In statistical literature there are various loss functions. Some loss functions are symmetric in nature and some are asymmetric. Symmetric losses, e.g. squared error loss, assigns equal weight to both under- and over-estimation. On the other hand, asymmetric losses such as linex loss assigns appropriate wight to under- and over-estimations. In real life there are situations where over-estimation is more severe than under-estimation and vice-versa. For example, in a dam construction underestimation of the peak water level is more serious than its over-estimation. In such situation the choice of an appropriate asymmetric loss is necessary for the study of the risk.

In an applied study of real estate assessment, Varian (1975) proposed the linex loss function. It includes the symmetric squared error loss as a special case. Hoque *et al.* (2009) investigated the performance of preliminary test estimator of simple linear regression model under asymmetric linex loss. In a recent study, Hoque *et al.* (2018) studied the performance of several estimators, including the shrinkage estimator, of the simple linear regression model under the linex loss function. For more accounts in the area readers may see Zellner (1986). In this paper, we investigate the performance of the preliminary test estimator of the elliptically contoured distribution under linex loss. The layout of the paper is as follows.

Preliminaries of the linex loss function, the elliptically contoured distribution (ECD) and the proposed estimator are discussed in Section 2. Risk functions of the estimators are derived and analyzed in Section 3. Finally, some concluding remarks are presented in Section 4.

2 Preliminaries

In this section we briefly describe the linex loss function and its properties along with the elliptically contoured distribution and the proposed estimator of it's parameter.

2.1 Linex Loss Function

In many real life situations over-estimation and under-estimation of the same magnitude of a parameter often have different economic and physical implications. As the symmetric loss functions, such as the squared error loss, fail to differentiate between over- and under-estimation, it is inappropriate to use such loss function, and the appropriate loss function is an asymmetric loss such as the linex loss. For the purpose of this study, we consider the following multivariate linear exponential (mlinex) loss function with the scale parameter b and the p-vector shape parameter $\boldsymbol{a}$

$$L(\hat{\boldsymbol{\theta}}, \boldsymbol{t}) = b\{e^{\boldsymbol{a}'(\hat{\boldsymbol{\theta}}-\boldsymbol{t})} - \boldsymbol{a}'(\hat{\boldsymbol{\theta}} - \boldsymbol{t}) - 1\}, \tag{2.1}$$

where $\hat{\boldsymbol{\theta}}$ is any natural estimator of vector parameter $\boldsymbol{t}$. A positive $\boldsymbol{a}$ indicates that over-estimation is more serious than under-estimation and a negative $\boldsymbol{a}$ represents the reverse situation. The magnitude of $\boldsymbol{a}$ reflects the degree of asymmetry about $(\hat{\boldsymbol{\theta}} - \boldsymbol{t}) = \boldsymbol{0}$. If $\boldsymbol{a} \to \boldsymbol{0}$, then the linex loss reduces to the symmetric squared error loss. For simplicity, without any loss of generality we assume $b = 1$ throughout this paper. For more details on the properties of the linex loss function readers may see Zellner (1986), Varian (1975), Parsian (1990), Zou (1997), Parsian and Kirimani (2002), Arashi *et al.* (2008), Hoque *et al.* (2009), Hoque *et al.* (2018) and Zou *et al.* (2009).

2.2 The Model and Estimators

Let $\mathcal{S}(p)$ be the set of all positive definite matrices and $\Psi_p = \{\psi(\cdot) : \psi(t_1^2 + \cdots + t_p^2)\}$ be a p-dimensional characteristic function. Then $\Psi_1 \supset \Psi_2 \supset \cdots$. Define $\Psi_\infty = \bigcap_{i=1}^{\infty} \Psi_i$.

Definition 2.1. A $p \times 1$ random vector $\boldsymbol{X}$ is said to have an ECD with parameters $\boldsymbol{t}_{p\times 1}$ and $\boldsymbol{\Sigma}_{p\times p}$ and function ψ denoted by $\boldsymbol{X} \sim \mathcal{E}_p(\boldsymbol{t}, \boldsymbol{\Sigma}, \psi)$ if its characteristic function has the following form

$$\phi_{\boldsymbol{X}}(\boldsymbol{t}) = \exp(i\boldsymbol{t}'\boldsymbol{t})\psi(\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}),$$

where $\boldsymbol{t} \in \mathbb{R}^p$, $\boldsymbol{\Sigma} \in \mathcal{S}(p)$ and $\psi \in \Psi_p$.

Here ψ is called the characteristic generator. When $\boldsymbol{t} = 0$, for a positive-valued scalar σ^2 and $\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{I}_p$, $\boldsymbol{X}$ is said to have spherically symmetric distribution (SSD). Not necessarily $\boldsymbol{X}$ possess a density. However, if it does, the density is of the form .

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = k_n |\boldsymbol{\Sigma}|^{-1/2} h\left[(\boldsymbol{x} - \boldsymbol{t})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{t})\right]$$

for some function h (say), the density generator, where k_n is the normalizing constant. Note that h and ψ determine each other for each specified member of the family of distributions. In this case we may use the notation $\boldsymbol{X} \sim \mathcal{E}_n(\boldsymbol{t}, \boldsymbol{\Sigma}, h)$. It follows $\mathrm{E}(\boldsymbol{X}) = \boldsymbol{t}$ and $\mathrm{Cov}(\boldsymbol{X}) = -2\psi'(0)\boldsymbol{\Sigma}$ provided $|\psi'(0)| < \infty$. ? expressed the density of an ECD as an integral of a set of multivariate normal densities which is stated in the following lemma.

Lemma 2.1. Let $\boldsymbol{z}$ is a n-dimensional elliptically contoured random vector with mean equal to $\boldsymbol{t}$ and scale matrix $\boldsymbol{\Sigma}$ and density function $g(\cdot)$ with density generator $h(\cdot)$. If $h(t)$, $t \in [0, \infty)$ has the inverse Laplace transform then there exists a scalar function $\mathcal{W}(t)$ defined on $(0, \infty)$ such that

$$g(\boldsymbol{z}) \quad = \quad \int_0^\infty \mathcal{W}(t) f_{\mathcal{N}}(\boldsymbol{z}) dt,$$

where $f_{\mathcal{N}}(\cdot)$ denotes the density function of $\mathcal{N}_n(\boldsymbol{t}, t^{-1}\boldsymbol{\Sigma})$, and

$$\mathcal{W}(t) \quad = \quad (2\pi)^{\frac{n}{2}} \, |\boldsymbol{\Sigma}|^{\frac{1}{2}} t^{-\frac{n}{2}} \mathcal{L}^{-1}\{h(s)\},$$

$\mathcal{L}^{-1}\{h(s)\}$ denotes the inverse Laplace transform of $h(s)$.

For details on the properties of Laplace transform and its inverse see ?. On integrating $g(\boldsymbol{z})$ over $\mathbb{R}^n$, $\mathcal{W}(t)$ integrates to 1. Thus for non-negative function $\mathcal{W}(t)$, $g(z)$ is a density. Now precisely, let $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n$ be a random sample of size n from $\mathcal{E}_p(\boldsymbol{t}, \sigma^2 \boldsymbol{V}, h)$. Then the sample information based unrestricted estimator (UE) of $\boldsymbol{t}$ is given by

$$\hat{\boldsymbol{\theta}} = \bar{\boldsymbol{X}} = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{X}_i \sim \mathcal{E}_p(\boldsymbol{t}, \frac{\sigma^2}{n} \boldsymbol{V}, h). \tag{2.2}$$

Subsequently, the UE of σ^2 is given by

$$\hat{\sigma}^2 = \sum_{j=1}^{n} (\boldsymbol{X}_j - \hat{\boldsymbol{\theta}})' \boldsymbol{V}^{-1} (\boldsymbol{X}_j - \hat{\boldsymbol{\theta}}). \tag{2.3}$$

It is easy to show that $S^2 = \frac{1}{n-1} \hat{\sigma}^2$ is unbiased for σ_h^2, where

$$\sigma_h^2 = -2\sigma^2 \psi'(0) = \sigma^2 \kappa^{(1)}, \qquad \kappa^{(i)} = \int_0^\infty \left(\frac{1}{t}\right)^i \mathcal{W}(t) dt. \tag{2.4}$$

Under a restricted setting, assume the sub-space restriction $\boldsymbol{t} = \boldsymbol{t}_0$ holds. Then the restricted estimator (RE) of $\boldsymbol{t}$ is $\boldsymbol{t}_0$. However, this restriction is under suspicion as we do not know the outcome of testing the hypothesis $H_0 : \boldsymbol{t} = \boldsymbol{t}_0$. A remedy is to use the Bancroft (1944) approach of the preliminary test estimation. A simple form of the preliminary test estimator (PTE) is given by

$$\begin{aligned} \hat{\boldsymbol{\theta}}^{\mathrm{PT}} &= \hat{\boldsymbol{\theta}} \, I(\mathcal{L}_n \geq \mathcal{L}_{n,\alpha}) + \boldsymbol{t}_0 \, I(\mathcal{L}_n < \mathcal{L}_{n,\alpha}) \\ &= \hat{\boldsymbol{\theta}} - (\hat{\boldsymbol{\theta}} - \boldsymbol{t}_0) \, I(\mathcal{L}_n < \mathcal{L}_{n,\alpha}), \end{aligned} \tag{2.5}$$

where $\mathcal{L}$ is the test-statistic for testing $H_0 : \boldsymbol{t} = \boldsymbol{t}_0$, $\mathcal{L}_{n,\alpha}$ is the α-level critical value of $\mathcal{L}$ and $I(A)$ is the indicator function of the set A. The following theorem reveals the test statistic and its non-null distribution.

Theorem 2.1. Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n$ be a random sample from $\mathcal{E}_p(\boldsymbol{t}, \sigma^2\boldsymbol{V}, h)$. Also let $\Omega = \{(\boldsymbol{t}, \sigma, \boldsymbol{V}) : \boldsymbol{t} \in \mathbb{R}, \sigma \in \mathbb{R}^+, \boldsymbol{V} > 0\}$, and, $\omega = \{(\boldsymbol{t}, \sigma, \boldsymbol{V}) : \boldsymbol{t} = \boldsymbol{t}_0, \boldsymbol{t}_0 \in \mathbb{R}, \sigma \in \mathbb{R}^+, \boldsymbol{V} > 0\}$. Moreover, suppose $y^{\frac{n}{2}}h(y)$ has a finite positive maximum y_h. Then the likelihood ratio criterion for testing hypothesis $H_0 : \boldsymbol{t} = \boldsymbol{t}_0$ is given by

$$\mathcal{L}_n = \frac{n(\hat{\boldsymbol{\theta}} - \boldsymbol{t}_0)'\boldsymbol{V}^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{t}_0)}{S^2},$$

and $\mathcal{L}_n$ has the following modified generalized non-central F distribution given by

$$g^*_{1,m}(\mathcal{L}) = \sum_{r\geq 0} \frac{\left(\frac{1}{m}\right)^{\frac{1}{2}(1+2r)} \mathcal{L}_n^{\frac{1}{2}(2r-1)} K_r^0(\Delta^2)}{r!\; B\left(\frac{2r+1}{2}, \frac{m}{2}\right)\left(1 + \frac{1}{m}\mathcal{L}_n\right)^{(1+2r)/2}}, \tag{2.6}$$

where $m = n - 1$, $\Delta^2 = \xi/\sigma_e^2$ for $\xi = (\boldsymbol{t} - \boldsymbol{t}_0)'\boldsymbol{V}^{-1}(\boldsymbol{t} - \boldsymbol{t}_0)$, and

$$K_r^h(\Delta^2) = \int_0^\infty \frac{e^{-\Delta^2/2}}{r!}\left(-\Delta^2/2\right)^r t^{-h}\mathcal{W}(t)\,dt. \tag{2.7}$$

Proof: For testing $H_0 : \boldsymbol{t} = \boldsymbol{t}_0$ against $H_1 : \boldsymbol{t} \neq \boldsymbol{t}_0$, let

$$\tilde{\sigma}^2 = \sum_{i=1}^n (\boldsymbol{X}_i - \boldsymbol{t}_0)'\boldsymbol{V}^{-1}(\boldsymbol{X}_i - \boldsymbol{t}_0). \tag{2.8}$$

Then using Theorem 1 and Corollary 1 of Anderson *et al.* (1986) we have

$$\begin{aligned}
\Lambda &= \frac{\max_\omega L(\boldsymbol{x})}{\max_\Omega L(\boldsymbol{x})} = \frac{d_n|\tilde{\sigma}^2\boldsymbol{V}|^{-\frac{1}{2}}\max_t f\left(\frac{\sum_{i=1}^n(\boldsymbol{x}_i-\boldsymbol{t})'\boldsymbol{V}^{-1}(\boldsymbol{x}_i-\boldsymbol{t})}{2\sigma^2}\right)}{d_n|\hat{\sigma}^2\boldsymbol{V}|^{-\frac{1}{2}}\max_t f\left(\frac{\sum_{i=1}^n(\boldsymbol{x}_i-\boldsymbol{t})'\boldsymbol{V}^{-1}(\boldsymbol{x}_i-\boldsymbol{t})}{2\sigma^2}\right)} \\
&= \left(\frac{\hat{\sigma}}{\bar{\sigma}}\right)^n \frac{h(y_h)}{h(y_h)} = \left(\frac{\sum_{j=1}^n(\boldsymbol{X}_i - \hat{\boldsymbol{\theta}})'\boldsymbol{V}^{-1}(\boldsymbol{X}_i - \hat{\boldsymbol{\theta}})}{\sum_{i=1}^n(\boldsymbol{X}_i - \boldsymbol{t}_0)'\boldsymbol{V}^{-1}(\boldsymbol{X}_i - \boldsymbol{t}_0)}\right)^n \\
&= \left(\frac{mS^2}{mS^2 + n(\hat{\boldsymbol{\theta}} - \boldsymbol{t}_0)'\boldsymbol{V}^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{t}_0)}\right)^n = (1 + \mathcal{L}_n/m)^{-n}.
\end{aligned}$$

Here $\mathcal{L}_n$ is the likelihood ratio test statistic for testing the underlying null hypothesis. For its non-null distribution, we note that under the assumption in which $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n$ is a random sample from $\mathcal{N}_n(\boldsymbol{t}, t^{-1}\sigma^2\boldsymbol{V})$ the test statistic $\mathcal{L}_n$ follows non-central F-distribution with $(1, m)$ d.f. and non-centrality parameter $\Delta_t^2 = \frac{\xi}{t^{-1}\sigma^2}$ Saleh (2006). Then applying Lemma 2.1, the result follows. Accordingly, we have

Corollary 2.1. Under H_0 the pdf of $\mathcal{L}_n$ is given by

$$g^*_{1,m}(\mathcal{L}_n) = \frac{\left(\frac{1}{m}\right)^{\frac{1}{2}} \mathcal{L}_n^{\frac{1}{2}-1}}{B\left(\frac{1}{2}, \frac{m}{2}\right)\left(1 + \frac{1}{m}\mathcal{L}_n\right)^{\frac{1}{2}(m+1)}},$$

which is the central F-distribution with $(1, m)$ d.f.

Corollary 2.2. The power function at γ-level of significance of $\mathcal{L}_n$, say, modified generalized non-central F cumulative distribution function of the statistic $\mathcal{L}_n$ is given by

$$\mathcal{G}_{p,m}(l_\gamma; \Delta^2) = \sum_{r\geq 0} \frac{1}{r!} K_r^0(\Delta^2) I_x \left[\frac{1}{2}(1+2r), \frac{m}{2}\right], \tag{2.9}$$

where $I_x(\cdot,\cdot)$ is the incomplete beta function, $x = \frac{l_\gamma}{m+l_\gamma}$ and $l_\gamma = F_{1,m}(\gamma)$.

The nature of the PTE depends on α $(0 < \alpha < 1)$, the level of significance of the test. Also it yields the extreme results, either the unrestricted estimator or the restricted estimator, depending on the outcome of the test.

3 Risk Functions

In this section we derive the risk functions of the estimators under the LINEX loss function given by (2.1). By definition, the risk function for any estimator $\boldsymbol{t}^*$ of $\boldsymbol{t}$ associated to any loss function is given by

$$\mathrm{R}(\boldsymbol{t}^*; \boldsymbol{t}) = \mathrm{E}[\mathrm{L}(\boldsymbol{t}^*, \boldsymbol{t})]. \tag{3.1}$$

To facilitate the derivation of the risk functions of our proposed estimators we need the following lemma.

Lemma 3.1. If $\boldsymbol{Z} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{I}_p)$ and $S^2 \sim \chi_k^2$ are independent, then for any Borel measurable function $\phi : \Re \times (0, \infty) \to \Re$ and for any $c \in \Re^p$,

$$\mathrm{E}[\exp(\boldsymbol{c}'\boldsymbol{Z})\, \phi(\boldsymbol{Z}, S^2)] = \exp(\boldsymbol{c}'\boldsymbol{\mu} + \boldsymbol{c}'\boldsymbol{c}/2)\, \mathrm{E}\left[\phi(\boldsymbol{Z} + \boldsymbol{c} + \boldsymbol{\mu}, S^2)\right]$$

provided $(\exp(\boldsymbol{c}'\boldsymbol{Z})\, \phi(\boldsymbol{Z}, S^2))$ is integrable.

Proof: Using the method of double expectation

$$\begin{aligned}
\mathrm{E}[\exp(\boldsymbol{c}'\boldsymbol{Z})\, \phi(\boldsymbol{Z}, S^2)] &= \mathrm{E}\left(\mathrm{E}(\exp(\boldsymbol{c}'\boldsymbol{Z})\, \phi(\boldsymbol{Z}, S)) | S^2 = s\right) \\
&= \mathrm{E}\left\{ \left(\frac{1}{\sqrt{2\pi}}\right)^p \int_{\Re^p} \phi(\boldsymbol{z}, s) \exp\left[\boldsymbol{c}'\boldsymbol{z} - \frac{(\boldsymbol{z}-\boldsymbol{\mu})'(\boldsymbol{z}-\boldsymbol{\mu})}{2}\right] d\boldsymbol{z} \middle| S^2 = s \right\} \\
&= \exp\left(\boldsymbol{c}'\boldsymbol{\mu} + \frac{\boldsymbol{c}'\boldsymbol{c}}{2}\right) \mathrm{E}\left\{ \left(\frac{1}{\sqrt{2\pi}}\right)^p \int_{\Re^p} \phi(\boldsymbol{z}, s) \right. \\
&\quad \left. \times \exp\left\{-\frac{1}{2}[\boldsymbol{z} - (\boldsymbol{c}+\boldsymbol{\mu})]'[\boldsymbol{z} - (\boldsymbol{c}+\boldsymbol{\mu})]\right\} d\boldsymbol{z} \middle| S^2 = s \right\} \\
&= \exp(\boldsymbol{c}'\boldsymbol{\mu} + \boldsymbol{c}'\boldsymbol{c}/2)\, \mathrm{E}\left[(2\pi)^{-\frac{p}{2}} \int_{\Re^p} \phi(\boldsymbol{u} + \boldsymbol{c} + \boldsymbol{\mu}, S^2) \exp(-\boldsymbol{u}'\boldsymbol{u}/2)\, d\boldsymbol{u}\right] \\
&= \exp(\boldsymbol{c}'\boldsymbol{\mu} + \boldsymbol{c}'\boldsymbol{c}/2)\, \mathrm{E}\left[\phi(\boldsymbol{Z} + \boldsymbol{c} + \boldsymbol{\mu}, S^2)\right]
\end{aligned}$$

which completes the proof of the lemma.

As $\bar{\boldsymbol{X}}$ is an unbiased estimator of $\boldsymbol{t}$ we have $\boldsymbol{a}'\mathrm{E}[\bar{\boldsymbol{X}} - \boldsymbol{t}] = \boldsymbol{0}$. Consider under normality $\boldsymbol{D} = (\frac{\sigma^2}{tn}\boldsymbol{V})^{-\frac{1}{2}}(\bar{\boldsymbol{X}} - \boldsymbol{t})|t \sim \mathcal{N}_p(\boldsymbol{0}, \boldsymbol{I}_p)$. Then applying Lemma 3.1 with ϕ as identity we get the risk function of the unrestricted estimator $\hat{\boldsymbol{\theta}}$ as follows.

$$\mathrm{R}(\hat{\boldsymbol{\theta}}; \boldsymbol{t}) = \int_0^\infty W(t)\, e^{\frac{\sigma^2 \boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}}{2tn}}\, dt - 1. \tag{3.2}$$

Theorem 3.1. The risk function of the preliminary test estimator of $\boldsymbol{t}$ under LINEX loss function is given by

$$\begin{aligned}\mathrm{R}(\hat{\boldsymbol{\theta}}^{\mathrm{PT}}; \boldsymbol{t}) &= e^{\boldsymbol{a}'(\boldsymbol{t}_0 - \boldsymbol{t})}\mathcal{G}_{p,m}(l_\gamma; \Delta^2) + \left[1 - \mathcal{G}_{p,m}(l_\gamma; \Delta_*^2)\right] \\ &\times \int_0^\infty \mathcal{W}(t) e^{\frac{\sigma^2 \boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}}{2tn}} - \boldsymbol{a}'(\boldsymbol{t} - \boldsymbol{t}_0)\mathcal{G}_{p+1,m}(l_\gamma; \Delta^2) - 1,\end{aligned}$$

where $\mathcal{G}_{i,j}(k; \Delta)$ is the cumulative distribution function of the non-central F distribution with (i, j) d.f., non-centrality parameter Δ and evaluated at k.

Proof: By definition we have

$$\mathrm{R}(\hat{\boldsymbol{\theta}}^{\mathrm{PT}}; \boldsymbol{t}) = \mathrm{E}(e^{\boldsymbol{a}'\boldsymbol{t}^*}) - \boldsymbol{a}'\mathrm{E}(\boldsymbol{t}^*) - 1, \tag{3.3}$$

where $\boldsymbol{t}^* = \hat{\boldsymbol{\theta}}^{\mathrm{PT}} - \boldsymbol{t}$.

Let $\boldsymbol{T} = \boldsymbol{V}^{-\frac{1}{2}}(\hat{\boldsymbol{\theta}} - \boldsymbol{t}_0)$. Using the fact that $\boldsymbol{T} \sim \mathcal{E}_p(\boldsymbol{\mu}, \frac{\sigma^2}{n}\boldsymbol{I}_p, h)$ where $\boldsymbol{\mu} = \boldsymbol{V}^{-\frac{1}{2}}(\boldsymbol{t} - \boldsymbol{t}_0)$, and Lemma 2.1 we get

$$\boldsymbol{u}|t = \tau\boldsymbol{T}|t \sim \mathcal{N}_p\left(\tau\boldsymbol{\mu}, \boldsymbol{I}_p\right), \qquad \tau = \sqrt{tn}/\sigma$$

is independent of $mS^2|t \sim \sigma^2 t^{-1}\chi_m^2$. Applying Lemma 3.1 we get

$$\begin{aligned}\mathrm{E}&\left[e^{\boldsymbol{a}'(\hat{\boldsymbol{\theta}} - \boldsymbol{t})} I(\mathcal{L}_n > \mathcal{L}_n(\alpha))\Big| t\right] = e^{\boldsymbol{a}'(\boldsymbol{t}_0 - \boldsymbol{t})}\mathrm{E}\left\{\exp\left[\boldsymbol{a}'(\hat{\boldsymbol{\theta}} - \boldsymbol{t}_0)\right]\right. \\ &\times I\left(\frac{n(\hat{\boldsymbol{\theta}} - \boldsymbol{t}_0)'\boldsymbol{V}^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{t}_0)}{S^2} > \mathcal{L}_n(\alpha)\right)\left| t\right\} \\ &= e^{\boldsymbol{a}'(\boldsymbol{t}_0 - \boldsymbol{t})}\mathrm{E}\left\{\exp\left[\tau^{-1}\boldsymbol{a}'\boldsymbol{V}^{\frac{1}{2}}\boldsymbol{u}\right] I\left(\frac{\boldsymbol{u}'\boldsymbol{u}}{S^2/(t^{-1}\sigma^2)} > \mathcal{L}_n(\alpha)\right)\left| t\right.\right\} \\ &= e^{\boldsymbol{a}'(\boldsymbol{t}_0 - \boldsymbol{t})} e^{\tau^{-1}\boldsymbol{a}'\boldsymbol{V}^{-\frac{1}{2}}\tau\boldsymbol{\mu}} e^{\frac{\sigma^2 \boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}}{2tn}} \\ &\times \mathrm{E}\left\{1 - I\left(\frac{(\boldsymbol{u} + \tau^{-1}\boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{a})'(\boldsymbol{u} + \tau^{-1}\boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{a})}{S^2/(t^{-1}\sigma^2)} \le \mathcal{L}_n(\alpha)\right)\left| t\right.\right\}.\end{aligned} \tag{3.4}$$

Thus, integrating with respect to t from (3.4) and using Corollary 2.2 we get

$$\mathrm{E}\left[e^{\boldsymbol{a}(\hat{\boldsymbol{\theta}} - \boldsymbol{t})} I(\mathcal{L}_n > \mathcal{L}_n(\alpha))\right] = \left[1 - \mathcal{G}_{p,m}(l_\gamma; \Delta_*^2)\right] \int_0^\infty \mathcal{W}(t) e^{\frac{\sigma^2}{2tn}\boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}} dt \tag{3.5}$$

where $\Delta_*^2 = \Delta^2 + 2\boldsymbol{a}'(\boldsymbol{t} - \boldsymbol{t}_0) + \frac{\boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}}{\sigma_e^2}$.
Now using (3.5) and Corollary 2.2 we get

$$\begin{aligned}
\mathrm{E}(e^{\boldsymbol{a}'\boldsymbol{t}^*}) &= \mathrm{E}\left\{\mathrm{E}\left[e^{\boldsymbol{a}'(\hat{\boldsymbol{\theta}}-\boldsymbol{t})-\boldsymbol{a}'(\hat{\boldsymbol{\theta}}-\boldsymbol{t}_0)I(\mathcal{L}_n\le\mathcal{L}_n(\alpha))}|\mathcal{L}_n \le \mathcal{L}_n(\alpha)\right]\right\} \\
&\quad + \mathrm{E}\left\{\mathrm{E}\left[e^{\boldsymbol{a}'(\hat{\boldsymbol{\theta}}-\boldsymbol{t})-\boldsymbol{a}'(\hat{\boldsymbol{\theta}}-\boldsymbol{t}_0)I(\mathcal{L}_n\le\mathcal{L}_n(\alpha))}|\mathcal{L}_n > \mathcal{L}_n(\alpha)\right]\right\} \\
&= e^{\boldsymbol{a}'(\boldsymbol{t}_0-\boldsymbol{t})}\mathrm{E}[I(\mathcal{L}_n \le \mathcal{L}_n(\alpha))] + \mathrm{E}\left[e^{\boldsymbol{a}'(\hat{\boldsymbol{\theta}}-\boldsymbol{t})}I(\mathcal{L}_n > \mathcal{L}_n(\alpha))\right] \\
&= e^{\boldsymbol{a}'(\boldsymbol{t}_0-\boldsymbol{t})}\mathcal{G}_{p,m}(l_\gamma;\Delta^2) + [1 - \mathcal{G}_{p,m}(l_\gamma;\Delta_*^2)] \\
&\quad \times \int_0^\infty \mathcal{W}(t)e^{\frac{\sigma^2}{2tn}\,\boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}}dt. \qquad (3.6)
\end{aligned}$$

Also we have

$$\mathrm{E}(\boldsymbol{t}^*) = \mathrm{E}(\hat{\boldsymbol{\theta}}^{\mathrm{PT}} - \boldsymbol{t}) = (\boldsymbol{t} - \boldsymbol{t}_0)\mathcal{G}_{p+1,m}(l_\gamma;\Delta^2) \qquad (3.7)$$

which is the bias function of the PTE of θ. Substituting the results from (3.6) and (3.7) in (3.3), we obtain the final expression of the risk function of PTE.

3.1 Performance of PTE

In this subsection we compare the risk of PTE with that of the UE. From equation (3.2) and Theorem 3.1 we get

$$\mathrm{R}(\hat{\boldsymbol{\theta}}^{\mathrm{PT}};\boldsymbol{t}) = \mathrm{R}(\hat{\boldsymbol{\theta}};\boldsymbol{t}) + \mathcal{F}(\Delta), \qquad (3.8)$$

where

$$\begin{aligned}
\mathcal{F}(\Delta) \quad &= \quad e^{\boldsymbol{a}'(\boldsymbol{t}_0-\boldsymbol{t})}\mathcal{G}_{p,m}(l_\gamma;\Delta^2) - \mathcal{G}_{p,m}(l_\gamma;\Delta_*^2)\left(\int_0^\infty \mathcal{W}(t)e^{\frac{\sigma^2}{2tn}\,\boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}}dt\right) \\
&\quad -\boldsymbol{a}'(\boldsymbol{t} - \boldsymbol{t}_0)\mathcal{G}_{p+1,m}(l_\gamma;\Delta^2). \qquad (3.9)
\end{aligned}$$

Under the null hypothesis $\theta = \theta_0$ and hence $\Delta = 0$. Then from (3.8) we obtain

$$\begin{aligned}
\mathrm{R}(\hat{\boldsymbol{\theta}}^{\mathrm{PT}};\boldsymbol{t}) \le \mathrm{R}(\hat{\boldsymbol{\theta}};\boldsymbol{t}) \quad &\Leftrightarrow \quad \left(\int_0^\infty \mathcal{W}(t)e^{\frac{\sigma^2}{2tn}\,\boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}}dt\right) \ge \frac{\mathcal{G}_{p,m}(l_\gamma;0)}{\mathcal{G}_{p,m}(l_\gamma;\frac{\boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}}{\sigma_e^2})} \\
&\Leftrightarrow \quad \left(\int_0^\infty \mathcal{W}(t)e^{\frac{\sigma^2}{2tn}\,\boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}}dt\right) \ge \frac{1-\gamma}{\mathcal{G}_{p,m}(l_\gamma;\frac{\boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}}{\sigma_e^2})},
\end{aligned}$$

since $\mathcal{G}_{p,m}(l_\gamma;.) \ge 0$ and $\mathcal{G}_{p,m}(l_\gamma;0) = 1 - \gamma$ (see Corollaries 2.1 and 2.2).

Since $\mathrm{R}(\hat{\boldsymbol{\theta}};\boldsymbol{t}) \ge 0$, using equation (3.2) we get

$$\int_0^\infty \mathcal{W}(t)e^{\frac{\sigma^2}{2tn}\,\boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}}dt \ge 1. \qquad (3.10)$$

If, $\mathcal{G}_{p,m}(l_\gamma; \frac{\boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}}{\sigma_e^2}) \geq 1 - \gamma$, we conclude that, $\hat{\theta}^{\mathrm{PT}}$ outperforms $\hat{\theta}$ under the null hypothesis. For the values $\boldsymbol{t} > \boldsymbol{t}_0$, from (3.10) and the fact that all $\mathcal{G}_{p,m}(l_\gamma; \Delta^2)$, $\mathcal{G}_{p,m}(l_\gamma; \Delta_*^2)$ and $\mathcal{G}_{p+1,m}(l_\gamma; \Delta^2)$ are ≤ 1, we get

$$\begin{aligned} -\mathcal{G}_{p,m}(l_\gamma; \Delta^2) \left(\int_0^\infty \mathcal{W}(t) e^{\frac{\sigma^2}{2tn} \boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}} dt \right) &\leq 1 \\ e^{\boldsymbol{a}'(\boldsymbol{t}_0 - \boldsymbol{t})} \mathcal{G}_{p,m}(l_\gamma; \Delta^2) &< 1 \\ -\boldsymbol{a}'(\boldsymbol{t} - \boldsymbol{t}_0) \mathcal{G}_{p+1,m}(l_\gamma; \Delta^2) &\leq 1. \end{aligned} \tag{3.11}$$

Collecting all the terms in (3.11) and replacing in (3.9) yields

$$\mathcal{F}(\Delta) < 1, \quad \text{whenever} \quad \boldsymbol{t} > \boldsymbol{t}_0. \tag{3.12}$$

Thus from (3.12) we conclude that $\hat{\theta}^{\mathrm{PT}}$ outperforms $\hat{\theta}$ for $\boldsymbol{t} > \boldsymbol{t}_0$. In a more general

Figure 1: Risk of the UE and PTE against the estimation error for $a = 4$

condition, since for every x, $\exp(x) - x - 1 \geq 0$, from (3.9) we obtain

$$\begin{aligned} \mathcal{F}(\Delta) \geq\; & \boldsymbol{a}'(\boldsymbol{t}_0 - \boldsymbol{t}) \left[\mathcal{G}_{p,m}(l_\gamma; \Delta^2) + \mathcal{G}_{p+1,m}(l_\gamma; \Delta^2) \right] \\ & + \mathcal{G}_{p,m}(l_\gamma; \Delta^2) - \mathcal{G}_{p,m}(l_\gamma; \Delta_*^2) \left(\int_0^\infty \mathcal{W}(t) e^{\frac{\sigma^2}{2tn} \boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}} dt \right) \\ \geq\; & 0, \end{aligned}$$

whenever

$$\boldsymbol{a}'(\boldsymbol{t}_0 - \boldsymbol{t}) \geq \frac{\mathcal{G}_{p,m}(l_\gamma; \Delta_*^2) \left(\int_0^\infty \mathcal{W}(t) e^{\frac{\sigma^2}{2tn} \boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}} dt \right) - \mathcal{G}_{p,m}(l_\gamma; \Delta^2)}{\left[\mathcal{G}_{p,m}(l_\gamma; \Delta^2) + \mathcal{G}_{p+1,m}(l_\gamma; \Delta^2) \right]},$$

Therefore, UE outperforms PTE and vice versa whenever

$$a'(t_0 - t) \le \frac{\mathcal{G}_{p,m}(l_\gamma; \Delta_*^2)\left(\int_0^\infty \mathcal{W}(t) e^{\frac{\sigma^2}{2tn} a'Va} dt\right) - \mathcal{G}_{p,m}(l_\gamma; \Delta^2)}{[\mathcal{G}_{p,m}(l_\gamma; \Delta^2) + \mathcal{G}_{p+1,m}(l_\gamma; \Delta^2)]}.$$

To illustrate the superiority of the PTE over the UE, data has been simulated from the risk functions of both estimators and plotted against the estimation error in Figure 1. It reaffirms that if the estimation error Δ is not too high, the PTE outperforms the UE. As Δ starts departing from 0, the risk of the PTE starts increasing with a higher rate in the positive side of Δ than the negative side. This is because the value of a is taken as positive assuming that the over-estimation is more serious than the under-estimation. Similar to the shape of the linex loss function, the risk function of the PTE is also asymmetric in nature. As the UE is solely based on the sample information, the risk of this estimator is constant regardless of the value of Δ.

4 Concluding Remarks

In this paper, we have derived and compared the risk of the preliminary test estimator with that of the unrestricted estimator. It has been found that if the prior information is not too far from the true value of the parameter, PTE is a better choice than the unrestricted estimator. As the prior information is usually obtained from some previous study or expert knowledge, such information is expected to be close to the true value of the parameter and hence the value of Δ is expected to be near 0.

4.1 Acknowledgments

The authors thankfully acknowledge the comments and suggestions of the chief editor and the associate editor of the journal which has resulted in a significant improvement of the quality of the paper.

References

Anderson, T.W., Fang, K.T. and Hsu, H., (1986). Maximum-likelihood estimates and likelihood ration criterioa for multivariate elliptically contoured distributions, *Can. J. of Stat.*, **14**, 55–59.

Arashi, M., Tabatabaey, S.M.M. and Khan, S., (2008). Estimation in multiple regression model with elliptically contoured errors under MLINEX loss, *J. Appl. Prob. Statist.*, **3**, 23-35.

Bancroft, T.A., (1944) On biases in estimation due to the use of the preliminary test of significance. *Ann. Math. Statist.*, **15**, 190-204.

Hoque, Z., Khan, S. and Wesolowski, J., (2009). Performance of preliminary test estimator under LINEX loss function, *Commun. Statist. Theory Meth.*, **38**(2), 252-261.

Hoque, Z., Wesolowski, J. and Hossain, S., (2018). Shrinkage estimator of regression model under asymmetric loss, *Commun. Statist. Theory Meth.*, **48**(22), 5547–5557.

Khan, S., (2005). Estimation of parameters of the simple multivariate linear model with Student-t error, *J. Statist. Res.*, **39**(2), 79-94.

Khan, S., (2008). Shrinkage estimators of intercept parameters of two simple regression models with suspected equal slopes, *Commun. Statist. Theory Meth.*, **37**(2), 247–260.

Khan, S., Hoque, Z. and Saleh, A.K.Md.E., (2005). Estimation of the intercept parameter for linear regression model with uncertain prior information *Stat. Pap.*, **46**(3), 379–398.

Parsian, A., (1990). On the Admissibility of an Estimator of a Normal Mean Vector Under a LINEX Loss Function, *Annals of the Institute of Statistical Mathematics*, 42(4): 657–669.

Parsian, A., and Kirmani, S.N.U.A., (2002). Estimation under LINEX loss function. *Handbook of applied econometrics and statistical inference, Statistics Textbooks Mono-graphs*, 53–76.

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

Varian, H.R., (1975). Bayesian approach to real state assessment, *Studies in Bayesian econometrics and statistics in honor of Leonard J. Savage*, (eds. S. E. Fienberg and A. Zellner), 195–208, North Holland, Amsterdam.

Zellner, A., (1986). Bayesian estimation and prediction using asymmetric loss function, *J. Amer. Statist. Assoc.*, 81, 446–451.

Zou, G., (1997). Admissible estimation for finite population under the Linex loss function, *J. Statist. Plan. Inf.*, 61, 373–384.

Zou, G., Zeng, J., Wan, A.T.K., and Guan, Z., (2009). Stein-type improved estimation of standard error under asymmetric LINEX loss function, *Statistics*, **43**(2), 121–129.