Simultaneous Estimation of Several Cumulative Distribution Function (CDF)s: Homogeneity Constraint

by

A. K. Md. Ehsanes Saleh [1] and B. M. Golam Kibria & Florence George [2]

Abstract

Let $\{x_{ij}(1 \le j \le n_i)|i = 1, 2, \ldots, k\}$ be k independent samples of size n_j from respective distributions of functions $F_j(x)(1 \le j \le k)$. A classical statistical problem is to test whether these k samples came form a common populations with distributions function, $F_0(x)$ whose form may or may not be known. In this paper, we consider the complementary problem of estimating the distribution functions suspected to be homogeneous in order to improve the basic estimator known as "empirical distribution function" (edf) in an asymptotic setup. To illustarte the findings of the paper some tables and graphs are given.

Keywords: Distribution Function; Dominance; Efficiency; Empirical Distribution; Estimation; Pre-test; Risk Function; Stein Type Estimator.

AMS 2010 Subject Classifications: 62G07, 62G05.

1 Introduction

Let $\{x_{ij}(1 \le j \le n_i)|i = 1, 2, \ldots, k\}$ be k $(k \ge 2)$ independent samples with respective distributions of functions $F_j(x)(1 \le j \le k)$. A classical statistical problems is the test of homogeneity or test of goodness of fit, namely,

$$\begin{aligned} H_0: &\quad F_1(x) = F_2(x) = \ldots F_k(x) = F(x) \quad (unknown) \\ and & \\ H_0: &\quad F_1(x) = F_2(x) = \ldots F_k(x) = F_0(x)\ (known) \end{aligned} \tag{1.1}$$

This paper deals with the estimation of

$$\mathbf{F}(x) = (F_1(x), F_2(x), \ldots, F_k(x))'$$

[1] A. K. Md. Ehsanes Saleh is a Distinguished Research Professor in School of Mathematics and Statistics, Carleton University, Ottawa, Canada K1S 5B6, Email: esaleh@math.carleton.ca

[2] B. M. Golam Kibria is a Professor and Florence George is an Associate Professor in the Department of Mathematics and Statistics, Florida International University, Modesto A. Maidique Campus Miami, FL 33199, USA, Email: kibriag@fiu.edu

when it is *suspected* that either of the hypothesis may hold. To begin with, we consider the basic estimator of $\mathbf{F}(x)$, namely, the vector of the empirical distribution function, called unrestricted edf (Uedf), as

$$\mathbf{F}(x) = (F_{n_1}(x), F_{n_2}(x), \ldots, F_{n_k}(x))',$$

where $x = F_0^{-1}(\lambda), \quad (0 < \lambda < 1)$ under goodness of fit hypothesis. Otherwise, we choose

$$\bar{\xi}_\lambda \approx F_n^{-1}(\lambda),$$

where the average of Uedf is given by

$$\bar{F}_n = \bar{F}_n(\bar{\xi}_\lambda) = \frac{1}{n}[n_1 F_{n_1}(\bar{\xi}_\lambda) + n_2 F_{n_2}(\bar{\xi}_\lambda) + \ldots + n_k F_{n_k}(\bar{\xi}_\lambda)] \tag{1.2}$$

is the empirical distribution function under homogeneity constraint. We designate $\bar{F}_n(\bar{\xi}_\lambda)$ as the restricted edf (Redf) of $\mathbf{F}(x)$, The Redf generally performs better than the Uedf when $\mathbf{F}(x)$ is close to $F(x)1_k$. But, when $\mathbf{F}(x)$ deviate from the $F(x)1_k$, $\bar{\mathbf{F}}_n(\bar{\xi}_\lambda)$ may be considerably biased, inefficient and even inconsistent. For this reason, often we incorporate the uncertain prior information, namely homogeneity or goodness of fit restriction in the estimation of $\mathbf{F}(x)$. As a first step, we consider the preliminary test edf (PTedf), $\mathbf{F}_n^{PT}(x)$ as the estimator of $\mathbf{F}(x)$ as suggested by Bancroft (1944), Han and Bancroft (1968), Judge and Bock (1978), and expanded by Saleh (2006) among others.

Note that $\mathbf{F}_n^{PT}$ chooses $\mathbf{F}_n$ or $F_n 1_k$ according to the preliminary test leading to rejection or acceptance of the hypothesis, $H_0 : F_1(x) = F_2(x) = \ldots = F_k(x)$. The PTE of $\mathbf{F}(x)$ is a discrete process and not stable as such we consider the continuous version of the $\mathbf{F}_n^{PT}$ known as the Stein-type estimator (SE), $\mathbf{F}_n^S$ (for details see Saleh 2006). This estimator is aimed at minimum quadratic risk. Further, improvement of $\mathbf{F}_n^S$, may be made by combining $\mathbf{F}_n^S$ and $\bar{F}_n 1_k$ to obtain the positive -rule Stein type estimator (PRSE), $\mathbf{F}_n^{S+}$. This paper is a follow up of the paper by Saleh and Ghania (2014a): "New estimators of a CDF" based on a single sample.

The proposed PTE, SE and PRSE, along with preliminary notions are presented in Section 2. The notion of asymptotic distributional bias (ADB), MSE (ADMSE) and quadratic risk (ADQR) are considered in section 3 and, in this light, ADQR results for various estimates are presented as well. The main result on asymptotic risk efficient (ARE) and relative efficieny will be discussed in Section 4. Finally some concluding remars are presented in section 5.

2 Proposed Estimators of $F(x)$: Homogeneity Hypothesis

First, we consider the unrestricted estimator (UE) of $\mathbf{F}$ as

$$\mathbf{F}_n(\bar{\xi}_\lambda) = (F_{n_1}(\bar{\xi}_\lambda), F_{n_2}(\bar{\xi}_\lambda), \ldots F_{n_k}(\bar{\xi}_\lambda))'.$$

It is well known that for fixed $\bar{\xi}_\lambda \in R^1$, $E(\mathbf{F}_n(x)) = \mathbf{F}(x)$ and

$$\begin{aligned} Cov(\mathbf{F}_n) &= Diag\left(\frac{1}{n_i}F_i(x)[1-F_i(x)], |i=1,2,\ldots,k\right) \\ &\leq Diag\left(\frac{1}{4n_i}|i=1,2,\ldots,k\right) \\ &= \frac{1}{4n}Diag\left(\frac{n}{n_i}|i=1,2,\ldots,k\right) = \frac{1}{4n}\Lambda_n^{-1} \end{aligned} \tag{2.1}$$

where $\Lambda_n = Diag(\frac{n_i}{n}|i=1,2,..,\ldots,k)$. Also, $\mathbf{F}_n \xrightarrow{\text{a.s.}} \mathbf{F}_x$ for fixed $\bar{\xi}_\lambda \in R^1$.

There has been an increasing amount of research work in the area of PTE, SE and PRSE since 2000. This is referenced by the recent work of Arashi (2012), Arashi and Tabatabaey (2011), Arashi et al. (2014a, b), Kibria and Saleh (2012), Saleh (2006), Saleh (2013), Saleh and Kibria (2011), Saleh and Ghania (2014), Shalabh et al. (2009), Shalabh and Wan (2000) among others. In this paper, we introduce for the first time, the estimation of CDF in a multi-sample situation. Next, the restricted estimator of $\mathbf{F}_x$ is the RE, $F_n(\bar{\xi}_\lambda)1_k$. Since the hypothesis $H_0 : F_1(x) = F_2(x) = \ldots = F_k(x)$ may hold, we consider the test of homogeneity given as

$$\mathcal{L}_n = 4n\mathbf{F}_n J_n' \Lambda_n J_n \mathbf{F}_n, \quad \Lambda_n = Diag\left(\frac{n_i}{n}|i=1,2,..,\ldots,k\right). \tag{2.2}$$

As $n \to \infty$ such that $\frac{n_i}{n} \to \lambda_j$, the distribution of $\mathcal{L}_n$ under the H_0 approximates the central chi-square distribution with $(k-1)$ degrees of freedom (D. F.). Let $\chi^2_{k-1}(\alpha)$ be the approximate upper α level critical value of the chi-square distribution with $(k-1)$ D.F. Then we define the PTE of $\mathbf{F}_X$ as PTedf given by

$$\mathbf{F}_n^{PT} = \mathbf{F}_n - (\mathbf{F}_n - \bar{F}_n 1_k) I(\mathcal{L}_n \leq \chi^2_{k-1}(\alpha)), \tag{2.3}$$

where $I(A)$ is the indicator function of the set A.

Next, we define the Stein-type estimators of $\mathbf{F}(x)$ as

$$\mathbf{F}_n^S = \mathbf{F}_n - (k-3)(\mathbf{F}_n - \bar{F}_n 1_k)\mathcal{L}_n^{-1} \tag{2.4}$$

and

$$\mathbf{F}_n^{S+} = \mathbf{F}_n^S - (\mathbf{F}_n^S - \bar{F}_n 1_k) I(\mathcal{L}_n \leq k-3) \tag{2.5}$$

which are known as James Stein type (JSE) and positive rule Styne type (PRS) estimator respectively.

3 Distributional Properties of the Five Estimators

In this section, we consider the notion of asymptotic distributional biases, MSE matrices and quadratic risk expressions for these estimators.

For the k dimensional unit cube, Ω, let $\omega(\subset \Omega)$ be a sub space for which $\mathbf{F}_X$ satisfies homogeneity constraint. Then by consistency of $\mathcal{L}_n$ test (2.2), we note that for fixed $\mathbf{F}_X :\not\in \omega$, $\mathcal{L}_n \to \infty$ as $n \to \infty$ and as $n \to \infty$, the risks of $\mathbf{F}_n^{PT}$, $\mathbf{F}_n^S$ and $\mathbf{F}_n^{S+}$ are equivalent which will be shown in the sequel. However, this situation is altered, when $\mathbf{F}_X \in \omega$, ie. when it belongs to a shrinkage neighborhood of ω. Thus we define a sequence of Pitman-alternative.

$$K_{(n)} : \mathbf{F}_x = \bar{F}_X 1_k + n^{-1/2}\xi, \quad \xi = (\xi_1, \xi_2, \ldots, \xi_k) \tag{3.1}$$

to find the asymptotic distribution of the $\mathbf{F}_n^{PT}$, $\mathbf{F}_n^S$ and $\mathbf{F}_n^{S+}$ respectively.

Let $G(x)$ be the cdf of $\sqrt{n}(\mathbf{F}_n^* - \mathbf{F}_x)$ under $\{K_{(n)}\}$ if it exists then

$$G(x) = \lim P_{K_{(n)}} \left(\sqrt{n}(\mathbf{F}_n^* - \mathbf{F}_x) \leq x\right) \tag{3.2}$$

We then define the asymptotic distributional bias (ADB), MSE matrices (AMSE) and quadratic risk (ADQR) as

$$\begin{aligned} b^*(\mathbf{F}_n^*) &= \int x dG^*(x) \\ M^*(\mathbf{F}_n^*) &= \int xx' dG^*(x) \\ R^*(\mathbf{F}_n^*; Q) &= tr(QM^*). \end{aligned} \tag{3.3}$$

Now, we have the following theorem on the asymptotic equivalence of $\mathbf{F}_n^{PT}$, $\mathbf{F}_n^S$ and $\mathbf{F}_n^{S+}$ to $\mathbf{F}_n(x)$ in ADQR while $\bar{F}_n(x)1_k$ has unbounded risk under fixed alternatives.

Theorem 3.1 For fixed $\mathbf{F}_X \not\in \omega$, $\bar{F}_n 1_k$ has unbounded ADRQ as $n \to \infty$, while $\mathbf{F}_n$, $\mathbf{F}_n^{PT}$, $\mathbf{F}_n^S$ and $\mathbf{F}_n^{S+}$ are asymptotically ADQR equivalent.

Proof:

$$\begin{aligned} n(\mathbf{F}_n^{PT}-\mathbf{F}_n)'Q(\mathbf{F}_n^{PT}-\mathbf{F}_n) &= n(\mathbf{F}_n - F_n 1_k)'Q(\mathbf{F}_n - F_n 1_k) I(\mathcal{L}_n < \chi^2_{k-1}(\alpha)) \\ &= \mathcal{L}_n I(\mathcal{L}_n < \chi^2_{k-1}(\alpha))\{n(\mathbf{F}_n - F_n 1_k)'Q(\mathbf{F}_n - F_n 1_k)\mathcal{L}_n^{-1}\} \\ &\leq \chi^2_{k-1}(\alpha)) I(\mathcal{L}_n < \chi^2_{k-1}(\alpha)) Ch_{max}(Q\Lambda^{-1}), \end{aligned}$$

where $\Lambda = Diag(\lambda_1, \lambda_2, \ldots, \lambda_k)$ and $\lim_{n\to\infty} \frac{n_i}{n} = \lambda_i$, $Ch_{max}(A)$ is the largest eigen value of A.

Also, note that

$$E\left[I(\mathcal{L}_n < \chi^2_{k-1}(\alpha))|F_x \notin \omega\right] \leq P(\mathcal{L}_n < \chi^2_{k-1}(\alpha))|F_x \notin \omega)$$

Thus for fixed $\mathbf{F}_X(x) \notin \omega$, $\mathbf{F}_n^{PT}(x)$ and $\mathbf{F}_n(x)$ are asymptotically risk equivalent.

For $\mathbf{F}_n^S(x)$, we note that on the set $\{\mathcal{L}_n\}$

$$\begin{aligned} & n(\mathbf{F}_n^S(x) - \mathbf{F}_n(x))'Q(\mathbf{F}_n^S(x) - \mathbf{F}_n(x)) \\ = \;& (k-3)^2 \mathcal{L}_n^{-2}\{n(\mathbf{F}_n(x) - \bar{F}_n(x)1_k)'Q(\mathbf{F}_n(x) - \bar{F}_n(x)1_k)\} \\ \leq \;& (k-3)^2 \mathcal{L}_n^{-1} Ch_{max}(Q\Lambda^{-1}) \end{aligned} \tag{3.4}$$

Thus, if we can show that

$$E\left[\mathcal{L}_n^{-1} I(\mathcal{L}_n > 0|\mathbf{F}_n(x) \notin \omega)\right] \to 0 \; as \; n \to \infty.$$

then $\mathbf{F}_n^S(x)$ and $\mathbf{F}_n(x)$ becomes asymptotically risk-equivalent for every $\bar{F}_n(x)1_k \notin \omega$. Now, $\mathcal{L}_n$ is non-negative and for every $\bar{F}_n(x) \notin \omega$, $n^{-1}\mathcal{L}_n \xrightarrow{\mathcal{P}} \Delta^2 = \delta'\Lambda\delta$ as $n \to \infty$ which is equivalent to $\mathcal{L}_n = O_p(n^{-1})$. Thus we have for every $\epsilon > 0$,

$$E\left[\mathcal{L}_n^{-1} I(0 < \mathcal{L}_n < \epsilon|\mathbf{F}_n(x) \notin \omega)\right] \to 0 \; as \; n \to \infty$$

Hence, $\mathbf{F}_n^S(x)$ is risk-equivalent to $\mathbf{F}_n(x)$. Similarly we can show that

$$\begin{aligned} & n(\mathbf{F}_n^{S+}(x) - \mathbf{F}_n(x))'Q(\mathbf{F}_n^{S+}(x) - \mathbf{F}_n(x)) \\ = \;& 4(k-3)^2 \mathcal{L}_n\{\mathcal{L}_n^{-1} I(\mathcal{L}_n < k-3) + I(\mathcal{L}_n < k-3)\}^2 \\ = \;& 4(k-3)^2\{\mathcal{L}_n^{-1} I(\mathcal{L}_n < k-3) + 2I(\mathcal{L}_n < k-3) + \mathcal{L}_n I(\mathcal{L}_n < k-3)\} \end{aligned} \tag{3.5}$$

Now

$$E\left\{\mathcal{L}_n^{-1} I(\mathcal{L}_n < k-3) + 2I(\mathcal{L}_n < k-3) + \mathcal{L}_n I(\mathcal{L}_n < k-3)|\mathbf{F}_n(x) \notin 0\right\} \to 0 \; as \; n \to \infty$$

Based on the above results we can say that $\mathbf{F}_n^{S+}(x)$ and $\mathbf{F}_n(x)$ are asymptotically risk equivalent. Next, we consider the case with Pitman alternatives.

Note that $F_n(x)1_k$ and $\mathbf{F}_n(x)$ both have non-negative elements bounded by 1, so the finite moment of $\sqrt{n}(\mathbf{F}_n(x) - \mathbf{F}(x))$ and $\sqrt{n}(\bar{F}_n(x)1_k - E[\bar{F}_n(x)1_k])$ of all finite orders exist and gave finite limit as $n \to \infty$. Also under $\{K_{(n)}\}$,

$$E[\bar{F}_n(x)1_k|K_{(n)}] - \mathbf{F}(x) = O(n^{-1/2})$$

so that in this case, $n||\mathbf{F}(x) - \bar{F}_n(x)1_k||^2 = O(1)$ in L-norm.

Thus, for (3.3) - (3.5), the convergence in distribution will ensure the convergence in second moment and hence (3.5) will ensure asymptotic risks under $\{K_{(n)}\}$. Note that Theorem 3.1 results applies as well for $\{K_{(n)}\}$. For our purpose it suffices to consider ADRQ through asymptotic distribution in (3.5). We present the asymptotic distributional theory and the ADB, ADMSE and ADQR results. Thus we present the following theorem.

Theorem 3.2: Under $\{K_{(n)}\}$ and the assumed regularity conditions, the following holds as $n \to \infty$

$$(i) \qquad \sqrt{n}(\mathbf{F}_n(x) - \mathbf{F}(x)) \overset{\mathcal{D}}{\Rightarrow} N_k\left(\xi, \frac{1}{4}\Lambda^{-1}\right)$$

$$(ii) \qquad \sqrt{n}(\mathbf{F}_n(x) - F_n(x)1_k) \overset{\mathcal{D}}{\Rightarrow} N_k\left(J\delta, \frac{1}{4}\Lambda^{-1}J'\right)$$

$$(iii) \qquad \sqrt{n}(F_n(x)1_k - F(\xi_\lambda))1_k \overset{\mathcal{D}}{\Rightarrow} N_k\left(0, \frac{1}{4}1_k1_k'\right)$$

$$(iv) \qquad \begin{pmatrix} \sqrt{n}(\mathbf{F}_n(x) - \mathbf{F}(x)) \\ \sqrt{n}(\mathbf{F}_n(x) - F_n(x)1_k) \end{pmatrix} \overset{\mathcal{D}}{\Rightarrow} \mathcal{N}_{2k}\left\{\begin{pmatrix} \xi \\ \delta \end{pmatrix}; \frac{1}{4}\begin{pmatrix} \Lambda^{-1} & \Lambda^{-1}J' \\ J\Lambda^{-1} & \Lambda^{-1}J' \end{pmatrix}\right\},$$

$$(v) \qquad \begin{pmatrix} \sqrt{n}(\mathbf{F}_n(x) - F(x)1_k) \\ \sqrt{n}(\mathbf{F}_n(x) - F_n(x)1_k) \end{pmatrix} \overset{\mathcal{D}}{\Rightarrow} \mathcal{N}_{2k}\left\{\begin{pmatrix} 0 \\ \delta \end{pmatrix}; \frac{1}{4}\begin{pmatrix} 1_k1_k' & O \\ O & \Lambda^{-1}J' \end{pmatrix}\right\},$$

$$(vi) \qquad \lim_{n\to\infty} P_{K_{(n)}}(\mathcal{L}_n \le x) = H_{k-1}(\chi^2_{k-1}(\alpha); \Delta^2), \quad \Delta^2 = 4\delta'\Lambda\delta, \quad \delta = \xi J \qquad (3.6)$$

where $H_\nu(.;\Delta^2)$ is the cdf of a non-central chi-square distribution with ν D.F. and non-centrality parameter Δ^2.

Now we proceed to compute the ADB, ADMSE and ADQR of the estimators based on the theorem 3.2. First, the ADB, ADMSE and ADQR of the UE is given by

$$b_1(\mathbf{F}_n(x)) \;=\; 0$$

$$\begin{aligned} M_1(\mathbf{F}_n(x)) &= \frac{1}{4}\Lambda^{-1} \quad \text{and} \\ R_1(\mathbf{F}_n(x);Q) &= \frac{1}{4}tr(Q\Lambda^{-1}) \end{aligned} \tag{3.7}$$

The ADB, ADMSE and ADQR of the RE is given by

$$\begin{aligned} b_2(\mathbf{F}_n(x)1_k) &= -\delta \\ M_2(\mathbf{F}_n(x)1_k) &= \frac{1}{4}1_k1_k' + \delta\delta' \quad \text{and} \\ R_2(\mathbf{F}_n(x)1_k;Q) &= \frac{1}{4}tr(1_k'Q1_k) + \delta'Q\delta \end{aligned} \tag{3.8}$$

The ADB, ADMSE and ADQR of the PTE is given by

$$\begin{aligned} b_3(\mathbf{F}_n^{PT}(x)) &= -\delta H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) \\ M_3(\mathbf{F}_n^{PT}(x)) &= \frac{1}{4}\Lambda^{-1} - \frac{1}{4}(I_k - 1_k1_k'\Lambda)\Lambda^{-1}H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) \\ &+ \delta\delta'\{2H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) - H_{k+3}(\chi^2_{k-1}(\alpha);\Delta^2)\}\text{and} \\ R_3(\mathbf{F}_n^{PT}(x);Q) &= \frac{1}{4}tr(Q\Lambda^{-1}) - \frac{1}{4}\{tr(Q\Lambda^{-1}) - tr(1_k'Q1_k')\}H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) \\ &+ \delta'Q\delta\{2H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) - H_{k+3}(\chi^2_{k-1}(\alpha);\Delta^2)\} \end{aligned} \tag{3.9}$$

The ADB, ADMSE and ADQR of the SE is given by

$$\begin{aligned} b_4(\mathbf{F}_n^{S}(x)) &= -(k-3)\delta E\left[\chi^{-2}_{k+1}(\Delta^2)\right] \\ M_4(\mathbf{F}_n^{S}(x)) &= \frac{1}{4}\Lambda^{-1} - \frac{1}{4}(k-3)\Lambda^{-1}\left\{2E\left[\chi^{-2}_{k+1}(\Delta^2)\right] - (k-3)E\left[\chi^{-4}_{k+1}(\Delta^2)\right]\right\} \\ &+ (k-3)(k+1)\delta\delta'E\left[\chi^{-4}_{k+3}(\Delta^2)\right] \\ R_4(\mathbf{F}_n^{S}(x);Q) &= \frac{1}{4}tr(Q\Lambda^{-1}) - \frac{1}{4}(k-3)tr(Q\Lambda^{-1})\left\{2E\left[\chi^{-2}_{k+1}(\Delta^2)\right] - (k-3)E\left[\chi^{-4}_{k+1}(\Delta^2)\right]\right\} \\ &+ (k-3)(k+1)\delta'Q\delta E\left[\chi^{-4}_{k+3}(\Delta^2)\right] \end{aligned} \tag{3.10}$$

The ADB, ADMSE and ADQR of the PR is given by

$$\begin{aligned} b_5(\mathbf{F}_n^{S+}(x)) &= -\delta\left\{(k-3)E\left[\chi^{-2}_{k+1}(\Delta^2)\right]\right. \\ &+ \left. E\left[\left(1-(k-3)\chi^{-2}_{k+1}(\Delta^2)\right)^2 I\left(\chi^2_{k+1}(\Delta^2 < k-3\right)\right]\right\} \\ M_5(\mathbf{F}_n^{S+}(x)) &= M_4(\mathbf{F}_n^{S}(x)) - \frac{1}{4}(k-3)\Lambda^{-1}E\left[\left(1-(k-3)\chi^{-2}_{k+1}(\Delta^2)\right)^2\right. \\ &\times \left. I\left(\chi^2_{k+1}(\Delta^2) < k-3\right)\right] \\ &+ (\delta\delta')\left\{2E\left[\left(1-(k-3)\chi^{-2}_{k+1}(\Delta^2)\right) I\left(\chi^2_{k+1}(\Delta^2 < k-3\right)\right]\right. \\ &- \left. E\left[\left(1-(k-3)\chi^{-2}_{k+3}(\Delta^2)\right)^2 I\left(\chi^2_{k+3}(\Delta^2) < k-3\right)\right]\right\} \end{aligned}$$

$$
\begin{aligned}
R_5(\mathbf{F}_n^{S+}(x);Q) &= R_4(\mathbf{F}_n^S(x)) - \frac{1}{4}(k-3)tr(\Lambda^{-1}Q)E\left[\left(1-(k-3)\chi_{k+1}^{-2}(\Delta^2)\right)^2\right. \\
&\times \left. I\left(\chi_{k+1}^2(\Delta^2) < k-3\right)\right] \\
&+ (\delta' Q\delta)\left\{2E\left[\left(1-(k-3)\chi_{k+1}^{-2}(\Delta^2)\right) I\left(\chi_{k+1}^2(\Delta^2) < k-3\right)\right]\right. \\
&- \left. E\left[\left(1-(k-3)\chi_{k+3}^{-2}(\Delta^2)\right)^2 I\left(\chi_{k+3}^2(\Delta^2) < k-3\right)\right]\right\} \qquad (3.11)
\end{aligned}
$$

4 Analysis of ADQR Properties

In this section, we will compare the performance of the estimators via ADQR expressions.

4.1 Comparison of Uedf and Redf

Consider the ADQR-difference between Uedf and Redf as follows

$$R_1(\mathbf{F}_n(x);Q) - R_2(\mathbf{F}_n(x)1_k;Q)) = \frac{1}{4}tr[Q(\Lambda^{-1} - 1_k 1_k')] - \delta' Q\delta \qquad (4.1)$$

By Courant Theorem

$$\Delta^2 Ch_{min}(Q\Lambda^{-1}) \leq \Delta^2 \frac{\delta' Q\delta}{\delta'\Lambda\delta} \leq \Delta^2 Ch_{max}(Q\Lambda^{-1})$$

where $Ch_{min}(A)$ and $Ch_{max}(A)$ are the minimum and maximum eigen values of A respectively. Thus we obtain that Redf, $F_n(x)1_k$ performs better than Uedf, $\mathbf{F}_n(x)$, whenever,

$$\Delta^2 \leq \frac{tr[Q(\Lambda^{-1} - 1_k 1_k')]}{4Ch_{max}(Q\Lambda^{-1})}$$

and Uedf, $\mathbf{F}_n(x)$ performs better than Redf, $F_n(x)1_k$, whenever,

$$\Delta^2 > \frac{tr[Q(\Lambda^{-1} - 1_k 1_k')]}{4Ch_{max}(Q\Lambda^{-1})}$$

If $Q = \Lambda$, then Redf performs better than Uedf whenever,

$$\Delta^2 < \frac{k-1}{4}$$

4.2 Comparison of $\mathbf{F}_n^{PT}(x)$, $\mathbf{F}_n(x)$ and $F_n(x)1_k$

Consider the ADQR expression for PTefd, $\mathbf{F}_n^{PT}(x)$ given by

$$\frac{1}{4}tr(Q\Lambda^{-1}) - \frac{1}{4}tr(Q\Lambda^{-1})H_{k+1}(\chi_{k-1}^2(\alpha);\Delta^2)$$

$$+ \quad \delta' Q \delta \{2H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) - H_{k+3}(\chi^2_{k-1}(\alpha);\Delta^2)\} \tag{4.2}$$

If $\alpha = 0$, then $\chi^2_{k-1}(0) = 1$, hence ADQR expression reduces to $\frac{1}{4}tr(1'_k Q 1_k) + \delta' Q \delta$, the ADQR of Redf, $\mathbf{F}_n(x)$. Now, by Courant theorem we have

$$\begin{aligned}
&\left[1 - H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) + 4\Delta^2 \frac{Ch_{max}(Q\Lambda^{-1})}{tr(Q\Lambda^{-1})} K(\alpha,\Delta^2)\right]^{-1} \\
&\le ARE(\mathbf{F}_n^{PT}(x) : \mathbf{F}_n(x)) \le \\
&\left[1 - H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) + 4\Delta^2 \frac{Ch_{min}(Q\Lambda^{-1})}{tr(Q\Lambda^{-1})} K(\alpha,\Delta^2)\right]^{-1}
\end{aligned} \tag{4.3}$$

$$K(\alpha,\Delta^2) = \left\{2H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) - H_{k+3}(\chi^2_{k-1}(\alpha);\Delta^2)\right\}$$

So that $\mathbf{F}_n^{PT}(x)$ performs better than the $\mathbf{F}_n(x)$ whenever

$$\Delta^2 \le \frac{\frac{1}{4}tr(Q\Lambda^{-1})H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2)}{4Ch_{max}(Q\Lambda^{-1})\left\{2H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) - H_{k+3}(\chi^2_{k-1}(\alpha);\Delta^2)\right\}} = \Delta_1^2(\alpha,k) \tag{4.4}$$

and $\mathbf{F}_n(x)$ performs better than the $\mathbf{F}_n^{PT}(x)$ whenever

$$\Delta^2 \ge \frac{\frac{1}{4}tr(Q\Lambda^{-1})H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2)}{4Ch_{max}(Q\Lambda^{-1})\left\{2H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) - H_{k+3}(\chi^2_{k-1}(\alpha);\Delta^2)\right\}} = \Delta_2^2(\alpha,k) \tag{4.5}$$

4.3 Comparison of $\mathbf{F}_n^S(x)$ and $\mathbf{F}_n(x)$

The ADQR of $\mathbf{F}_n^S(x)$ can be written as

$$\begin{aligned}
R_4(\mathbf{F}_n^S(x);Q) &= \frac{1}{4}tr(Q\Lambda^{-1}) - \frac{1}{4}(k-3)tr(Q\Lambda^{-1})\left\{2E\left[\chi^{-2}_{k+1}(\Delta^2)\right] - (k-3)E\left[\chi^{-4}_{k+1}(\Delta^2)\right]\right\} \\
&+ (k-3)(k+1)\delta' Q \delta E\left[\chi^{-4}_{k+3}(\Delta^2)\right] \\
&= \frac{1}{4}tr(Q\Lambda^{-1}) - \frac{1}{4}(k-3)tr(Q\Lambda^{-1})\left\{(k-3)E\left[\chi^{-4}_{k+1}(\Delta^2)\right]\right. \\
&+ \left. 2\Delta^2 E\left[\chi^{-4}_{k+3}(\Delta^2)\right]\left[1 - \frac{(k+3)\delta' Q\delta}{2\Delta^2 tr(Q\Lambda^{-1})}\right]\right\} \ge 0
\end{aligned} \tag{4.6}$$

for $\frac{tr(Q\Lambda^{-1})}{Ch_{max}(Q\Lambda^{-1})} \ge \frac{p+2}{2}$. Hence,

$$R_4(\mathbf{F}_n^S(x);Q) \le R_1(\mathbf{F}_n(x);Q)$$

$\forall(\Delta^2, Q)$ satisfy

$$\left\{Q : \frac{tr(Q\Lambda^{-1})}{Ch_{max}(Q\Lambda^{-1})} \ge \frac{p+2}{2}\right\}$$

4.4 Comparison of $\mathbf{F}_n^S(x)$ and $\mathbf{F}_n^{PT}(x)$

The AQR difference is

$$\begin{aligned} R_4(\mathbf{F}_n^{SE}(x);Q \quad &- \quad R_3(\mathbf{F}_n^{PT}(x);Q) = \frac{1}{4}\{tr(Q\Lambda^{-1}) - tr(1_k'Q1_k')\}H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) \\ &- \quad \delta'Q\delta\{2H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) - H_{k+3}(\chi^2_{k-1}(\alpha);\Delta^2)\} \\ &- \quad \frac{1}{4}(k-3)tr(Q\Lambda^{-1})\left\{2E\left[\chi^{-2}_{k+1}(\Delta^2)\right] - (k-3)E\left[\chi^{-4}_{k+1}(\Delta^2)\right]\right\} \\ &+ \quad (k-3)(k+1)\delta'Q\delta E\left[\chi^{-4}_{k+3}(\Delta^2)\right] \end{aligned} \tag{4.7}$$

The above difference will be non-negative when

$$\begin{aligned} &\delta Q\delta\left\{(2H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) - H_{k+3}(\chi^2_{k-1}(\alpha);\Delta^2)) - (k-3)(k+1)E\left[\chi^{-4}_{k+3}(\Delta^2)\right]\right\} \geq \\ &\frac{1}{4}\{tr(Q\Lambda^{-1}) - tr(1_k'Q1_k')\}H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) \\ - \quad &\frac{1}{4}(k-3)tr(Q\Lambda^{-1})\left\{2E\left[\chi^{-2}_{k+1}(\Delta^2)\right] - (k-3)E\left[\chi^{-4}_{k+1}(\Delta^2)\right]\right\} \end{aligned} \tag{4.8}$$

Using Courant theorem, $\mathbf{F}_n^S(x)$ will dominate $\mathbf{F}_n^{PT}(x)$ when

$$\Delta^2 \quad \geq \quad \frac{\left\{(2H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) - H_{k+3}(\chi^2_{k-1}(\alpha);\Delta^2)) - (k-3)(k+1)E\left[\chi^{-4}_{k+3}(\Delta^2)\right]\right\}}{Ch_{max}(Q\Lambda^{-1})f_1(\alpha,k)} \tag{4.9}$$

where

$$\begin{aligned} f_1(\alpha,k) \quad &= \quad \{tr(Q\Lambda^{-1}) - tr(1_k'Q1_k')\}H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) \\ &- \quad (k-3)tr(Q\Lambda^{-1})\left\{2E\left[\chi^{-2}_{k+1}(\Delta^2)\right] - (k-3)E\left[\chi^{-4}_{k+1}(\Delta^2)\right]\right\}. \end{aligned} \tag{4.10}$$

However, $\mathbf{F}_n^{PT}(x)$ will dominate $\mathbf{F}_n^S(x)$ when

$$\Delta^2 \quad \leq \quad \frac{\left\{(2H_{k+1}(\chi^2_{k-1}(\alpha);\Delta^2) - H_{k+3}(\chi^2_{k-1}(\alpha);\Delta^2)) - (k-3)(k+1)E\left[\chi^{-4}_{k+3}(\Delta^2)\right]\right\}}{Ch_{min}(Q\Lambda^{-1})f_1(\alpha,k)}. \tag{4.11}$$

Now, $\mathbf{F}_n^S(x)$ will dominate $\mathbf{F}_n^{RE}(x)$ when

$$\Delta^2 \geq \frac{1-(k-3)(k+1)E\left[\chi^{-4}_{k+3}(\Delta^2)\right]}{Ch_{max}(Q\Lambda^{-1})f_2(k,\Delta^2)} \tag{4.12}$$

where

$$f_2(k,\Delta^2) = \{tr(Q\Lambda^{-1}) - tr(1_k' Q 1_k')\} - (k-3)tr(Q\Lambda^{-1}) \\ \left\{2E\left[\chi_{k+1}^{-2}(\Delta^2)\right] - (k-3)E\left[\chi_{k+1}^{-4}(\Delta^2)\right]\right\} \tag{4.13}$$

However, $\mathbf{F}_n^{RE}$ will dominate $\mathbf{F}_n^S(x)$ when

$$\Delta^2 \le \frac{1-(k-3)(k+1)E\left[\chi_{k+3}^{-4}(\Delta^2)\right]}{Ch_{min}(Q\Lambda^{-1})f_2(k,\Delta^2)}. \tag{4.14}$$

4.5 Comparison of $\mathbf{F}_n^S(x)$ and $\mathbf{F}_n^{S+}(x)$

Now, consider the comparison of $\mathbf{F}_n^S(x);Q)$ and $\mathbf{F}_n^{S+}(x);Q)$. Since, $1-(k-3)\chi_{k+1}^{-2}(\Delta^2) \le 0$, the ADQR difference is

$$R_5(\mathbf{F}_n^{S+}(x);Q) - R_4(\mathbf{F}_n^S(x);Q) \le 0$$

Hence

$$R_5(\mathbf{F}_n^{S+}(x);Q) \le R_4(\mathbf{F}_n^S(x);Q) \le R_4(\mathbf{F}_n(x);Q) \quad \forall \Delta^2$$

We could compare all possible pairs of estimators using the corresponding risk functions. Since the comparison teachniques are similar, for the brevity of the paper, we compare them via table and figures. For $Q=\Lambda$, the values of risk for different Δ^2 are presneted in Table 4.1 and the graph in Figure 4.1. The values of relative efficiency for different Δ^2 are presneted in Table 4.2 and graph in Figure 4.2.

From the Tables 4.1 and 4.2 and Figures 4.1 and 4.2, we can see that near the null hypothesis both restricted and PTE perform better than the rest. However, for $k>3$, the positive rule estimators uniformly domonates both LSE and JSE.

5 Concluding Remarks

We have considered several estimators for estimating the population CDF, $F_X(x)$. The asymptotic bias, MSE and quadratuc risk of the estimators are provided. The perfromance of the estimators are discussed in the smaller quadratic risk sense. Under the null hypotheie, the RE performed the best. For $k>3$, we showed that the Stein-type positive rule estimator dominates both shrinkage and the unrestricted estimator uniformly in an asymptotic set-up,

Table 4.1: Risk of the Estimators for $\alpha = 0.05$ and different values of k and Δ^2

	Δ^2	UE	RE	PT	SE	PR
K=4						
	0.000	1.000	0.250	0.287	0.667	0.534
	1.000	1.000	0.500	0.556	0.737	0.646
	2.000	1.000	0.750	0.783	0.788	0.726
	5.000	1.000	1.500	1.199	0.875	0.856
	10.000	1.000	2.750	1.285	0.930	0.928
	20.000	1.000	5.250	1.059	0.964	0.964
	30.000	1.000	7.750	1.005	0.976	0.976
K=6						
	0.000	1.500	0.250	0.312	0.600	0.060
	1.000	1.500	0.500	0.597	0.733	0.361
	2.000	1.500	0.750	0.858	0.837	0.581
	5.000	1.500	1.500	1.436	1.042	0.958
	10.000	1.500	2.750	1.769	1.208	1.194
	20.000	1.500	5.250	1.607	1.334	1.334
	30.000	1.500	7.750	1.514	1.385	1.385

while preliminary test estimators works reasonably for $k = 3$. Tables and graphs are given for practical use of the methodology.

References

Arashi, M. (2012). Preliminary test and Stein-estimators in simultaneous equation. *Linear Alegebra and its Application.* 435(5), 1195-1211.

Arashi, M. and Tabatabaey, S.M.M. (2011). A note on Stein type estimators in ellitically contoured models. *Journal of Statistical Planning and Inference*, 140, 1206-1213.

Arashi, M., Kibria, B. M. G., Norouzirad, M. and Nadarajah, S. (2014a). Improved Preliminary test and Stein-Rule Liu Estimators for the Ill-Conditioned Elliptical Linear Regression Model. *Journal of Multivariate Analysis.* 126, 53-74.

Arashi, M, Kibria, B.M. G. and Tajadod, A. (2014b). On Shrinkage Estimators in Matrix Variate Elliptical Models. To appear in *Metrika*.

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

Table 4.2: Relative Efficiencies of the Estimators for $\alpha = 0.05$ and different values of k and Δ^2

	Delta2	RE	PT	SE	PR
K=4					
	0.000	4.000	3.478	1.500	1.872
	1.000	2.000	1.799	1.356	1.547
	2.000	1.333	1.277	1.268	1.377
	5.000	0.667	0.834	1.143	1.169
	10.000	0.364	0.778	1.075	1.078
	20.000	0.190	0.944	1.038	1.038
	30.000	0.129	0.995	1.025	1.025
K=6					
	0.000	6.000	4.800	2.500	25.010
	1.000	3.000	2.511	2.047	4.157
	2.000	2.000	1.749	1.792	2.583
	5.000	1.000	1.045	1.440	1.566
	10.000	0.545	0.848	1.242	1.256
	20.000	0.286	0.933	1.124	1.125
	30.000	0.194	0.991	1.083	1.083

Han, C.P. and Bancroft, T.A. (1968). On pooling means when variance is unknown. *Jour. Amer. Statist. Assoc.* 63. 1312-1333.

Heilers, S. and Willerst, R. (1988). Asymptotic normality of Restimators in linear models. *Statistics.* 19. 173-184.

Judge, G.G. and Bock, M.E. (1978). *The Statistical Implications of Pre-test and Stein-rule Estimators in Econometrics*, North-Holland Publishing Company, Amsterdam.

Kibria, B. M. G. and Saleh, K. Md. E. (2012). Improving the Estimators of the Parameters of a Probit Regression Model: A Ridge Regression Approach. *Journal of Statistical Planning and Inference.* 142 (2012) 14211435.

Saleh, A. K. Md. Ehsanes (2006). *Theory of Preliminary Test and Stein- Type Estimation with Applications.* John Wiley and Sons, Inc., publication.

Saleh, A. K. Md. Ehsanes (2013). Ridge Autoregression R-estimation: Subspace restriction. To appear in *Festhcrift in Honor of H.L. Koul.* Springer.

Saleh, K. Md. E. and Kibria, B. M. G. (2011). On Some Ridge Regression Estimators: A Nonparametric Approach. *Journal of Nonparametric Statistics.* 23(3), 819-851.

Saleh, A. K. Md. Ehsanes and Ghania, A. F. (2014a) New estimators of distribution functions. To appear in *Commun. Stat. Theo-method*.

Saleh, A. K. Md. Ehsanes and Ghania, A. F. (2014b) Ridge Autoregression Estimation: LS Method. To appear in *Commun. Stat. Theo-method*.

Shalabh, H. Toutenburg & C. Heumann (2009). Stein-rule estimation under an extended balanced loss function. *Journal of Statistical Computation and Simulation* 79 (10), 1259-1273.

Shalabh and Wan, A.T.K. (2000). Stein-rule estimation on mixed regression, *Biomedical Journal*, 42, pp203-214.

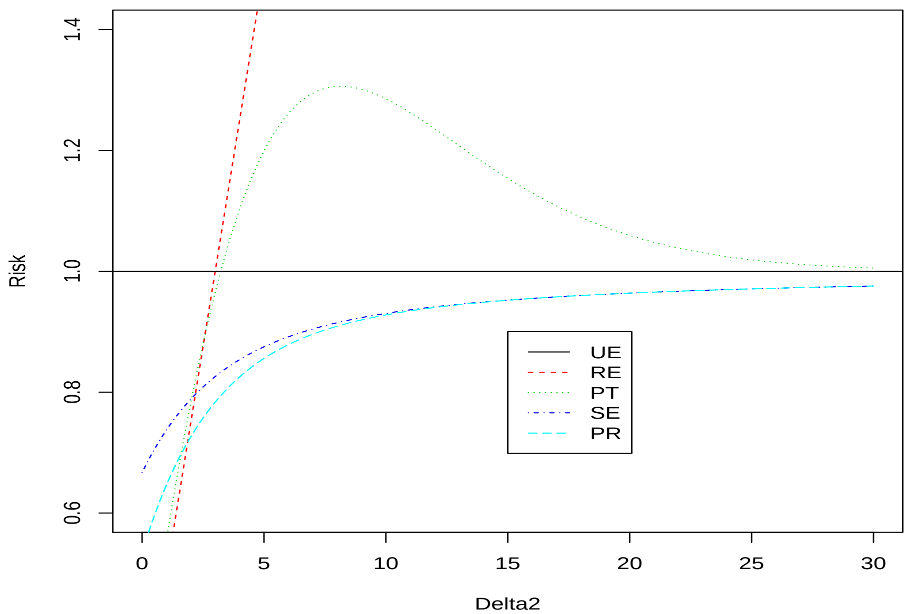

Figure 4.1: Risk behavior of the estimators for $\alpha = 0.05$ and $k = 4$

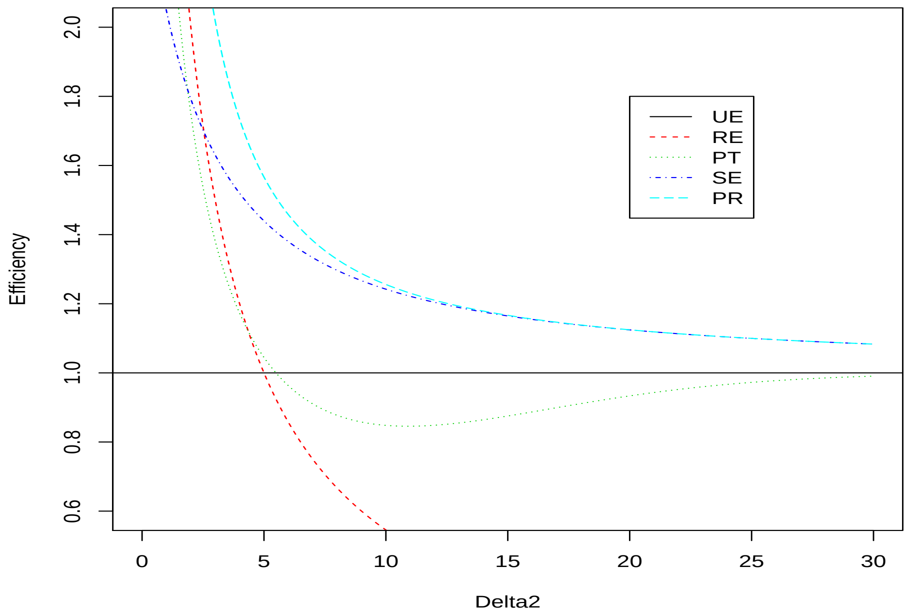

Figure 4.2: Efficiency of the estimators for $\alpha = 0.05$ and $k = 6$