IMS Collections
Volume Title
Vol. 0 (0000) 147–173

arXiv: math.PR/0000000

On the Estimation of Symmetric Distributions under Peakedness Order Constraints

Javier Rojo[1,*] **and José Batún-Cutz**

Rice University and Universidad Autónoma de Yucatán, Mérida, México

Abstract: Consider distribution functions F and G and suppose that F is *more peaked* about a than G is about b. The problem of estimating F or G, or both, when F and G are symmetric, arises quite naturally in applications. The empirical distribution functions F_n and G_m will not necessarily satisfy the order constraint imposed by the experimental conditions. Rojo and Batun-Cutz (2007) proposed some estimators that are strongly uniformly consistent when both m and n tend to infinity. However the estimators fail to be consistent when only either m or n tend to infinity. Here estimators are proposed that circumvent these problems and the asymptotic distribution of the estimators is delineated. A simulation study compares these estimators in terms of Mean Squared Error and Bias behavior with their competitors.

Contents

1. Introduction

The concept of stochastic order was pioneered by Lehmann (1955), and applications to hypotheses testing were discussed in Lehmann (1959), henceforth referred to as TSH-1. Lehmann and Rojo (1992) provided characterizations of stochastic

*Corresponding author
[1]Department of Statistics, MS-138; Rice University; 6100 Main Street; Houston, TX 77005; and School of Mathematics; Universidad Autónoma de Yucatán Mérida México e-mail: jrojo@rice.edu; jbatun@uady.mx
AMS 2000 subject classifications: Primary 62G30, 62G20; secondary 60F05
Keywords and phrases: Partial orders, Nonparametric statistics, Stochastic order, Empirical process, Weak convergence

ordering in terms of the maximal invariant with respect to the group of monotone transformations, and connections with other partial orderings were provided. Since the publication of TSH-1, there has been a large number of papers discussing various types of stochastic orders and their properties. Thus, one finds a large literature on stochastic orders in Economics (e.g. first-, second-, third-order stochastic dominance), reliability (e.g. IFR, IFRA, NBU, etc.), and applied probability (e.g. Laplace transform and dispersive orders). Marshall and Olkin (2007) and Shaked and Shantikumar (2007) are excellent references to the literature on stochastic orders.

The attention to this area of statistics and applied probability is well deserved. These concepts arise naturally in many applications in engineering, survival analysis, biology, economics, etc.

In corrosion engineering, for example, the times until pitting of metals immersed in a corrosive environment are measured under different solution corrosivities to discern the impact of the solution acidity on the pitting corrosion times. Shibata and Takeyama (1977) present data which strongly supports the belief that the times until pitting should be shorter in some sense, for the more corrosive environment. In toxicity studies, cells are grown in environments containing different levels of toxic materials (e.g. Arenaz *et al* (1992)). Invariably, the data supports the intuitive notion that the stronger the toxic solution is, the shorter the lifetimes of the organisms.

Another set of examples arises from clinical trials. This is illustrated by a clinical trial run to evaluate the efficiency of maintenance chemotherapy for acute myelongeneous leukemia (AML). The trial was conducted at Stanford University (Embury *et al* (1977)). After reaching a state of remission through treatment by chemotherapy, the patients who entered the study were randomized into two groups. The first group received maintenance chemotherapy; the second group did not. One would then expect that in this case, the survival times in the control group would be stochastically smaller than those in the first group.

Stochastic ordering, together with failure rate ordering, and monotone likelihood ratio ordering, are examples of *location* orderings. There are situations, however, when the interest lies in comparing distributions based on their *spread* rather than on their location.

Various concepts of spread, concentration, or dispersion have appeared in the literature. For example, Brown and Tukey (1946), Fraser (1957), Bickel and Lehmann (1979), Lehmann (1988), Doksum (1969), and Shaked (1980), define F to be more dispersive than G , denoted as $F >_d G$, if, for every $u > v$,

$$F^{-1}(u) - F^{-1}(v) \geq G^{-1}(u) - G^{-1}(v). \tag{1.1}$$

Shaked (1982), Bartoszewicz (1985a, 1985b, 1986), Oja (1981), and Rojo and He (1991), among others, have discussed various characterizations and properties of the dispersive order. Doksum (1969) utilized this concept to study power properties of rank tests, and showed that the power of certain rank tests is isotonic with respect to this order. Rojo (1995b, 1999) considered the problem of estimating the quantile function F^{-1} and the distribution function F when $F <_d G$, and the asymptotic theory of the resulting estimators was delineated. Rojo and Wang (1994) also showed that the power of tests based on L-statistics is isotonic with respect to the dispersive order. For other properties of the dispersive order, and connections with other partial orderings, see Bickel and Lehmann (1979), Proschan (1965), Karlin (1968), Shaked (1980, 1982), and Schweder (1982). When F and G

are assumed symmetric, (1.1) can be seen to be equivalent to

$$F^{-1}(u) - F^{-1}(1/2) \geq (\leq)\, G^{-1}(u) - G^{-1}(1/2)$$

depending on whether $u \geq (\leq)\ 1/2$.

Birnbaum (1948) proposed a different concept of dispersion based on the distribution functions rather than on the quantile functions. According to Brinbaum, the distribution function F is more peaked about the point a than the distribution function G is about the point b if, for all $x \geq 0$,

$$F((x+a)^-) - F(-x+a) \geq G((x+b)^-) - G(-x+b), \tag{1.2}$$

where $h(x^-) = \lim_{\epsilon \downarrow 0} h(x-\epsilon)$. We will write $F >_p G$ whenever (1.2) holds. It is easy to see that the condition (1.2) is equivalent to

$$\begin{array}{ll} F(x^-) \geq G(x^-) & \text{for } x \geq 0 \\ F(x) \leq G(x) & \text{for } x < 0. \end{array} \tag{1.3}$$

whenever F and G are symmetric about the point 0.

When F and G are continuous, it is easy to see that (1.2) is equivalent to requiring that $|X-a|$ be stochastically smaller than $|Y-b|$, and, although in general $F <_d G \not\Rightarrow F >_p G$ and $F >_p G \not\Rightarrow F <_d G$, it is easy to verify that $F <_d G \Rightarrow F >_p G$, when F and G are symmetric and continuous. When a and b in (1.2) are, respectively, the means of F and G, the condition (1.2) implies the obvious order on the variances of F and G.

An interesting example from statistical genetics, discussed in Rojo *et al* (2007), illustrates the importance of this concept in applications. Haseman-Elston (1972) proposed a regression model to assess the effect of a candidate gene on a phenotype when using sib-paired data. There have been some modifications of the initial proposal (see e.g. Elston *et al.* (2000)). The original model, Haseman-Elston (1972), represents the expected value of the squared phenotypic differences as a linear function of the proportion of alleles shared identical-by-descent (IBD) at the locus of interest. Let λ_i represent the proportion of alleles shared identical by descent ($\lambda_i = 0, \frac{1}{2}$, or 1). The Haseman and Elston (1972) regression model may then be written as follows: $E(X_i|\lambda_i) = \alpha + \beta\lambda_i$, where X_i represents the squared sib-pair difference for the i^{th} sib-pair conditional on λ_i. Writing $Z_{1i} = \theta + g_{1i} + \varepsilon_{1i}$ and $Z_{2i} = \theta + g_{2i} + \varepsilon_{2i}$ where Z_{1i} and Z_{2i} represent, respectively, the phenotype values for siblings one and two, and where θ is the population mean, and g_{ij} and ε_{ij} are the genetic and the residual effects, respectively, the model is then represented as

$$E(X_j|\lambda_j) \quad = \quad \eta_\varepsilon^2 + 2(1-\lambda_j)\eta_g^2$$

where, $\eta_\varepsilon^2 = E((\varepsilon_{1i} - \varepsilon_{2i})^2)$ and η_g^2 represents the variance in the trait due to allelic variation at the locus of interest. As a consequence of linkage between the candidate gene and the phenotype, siblings sharing two alleles IBD at the locus of interest will tend to be more similar than siblings sharing one allele IBD, and siblings sharing one allele IBD will in turn be more similar than siblings sharing no alleles IBD. It is then clear that phenotypical similarity of sibs within the same pair is being measured in terms of the spread of the distribution of the differences of the siblings' phenotypical measurements.

Existing sib-paired data illustrates very clearly that the distribution functions of sib-pair differences are symmetrically distributed. This will happen, for example, if

$(X-\mu_X, Y-\mu_Y)$ has the same distribution as $(\mu_X - X, \mu_Y - Y)$, as it happens under the assumption of a bivariate normal distribution, and if the means μ_X and μ_Y are equal, then the sib-pair differences are symmetrically distributed. When the candidate gene is linked to the phenotype of interest, the cumulative distributions of the differences within sib-pairs are ordered by peakedness. This is illustrated by sib-paired data on plasma Lipoprotein (a) data. Figure 1 shows the empirical

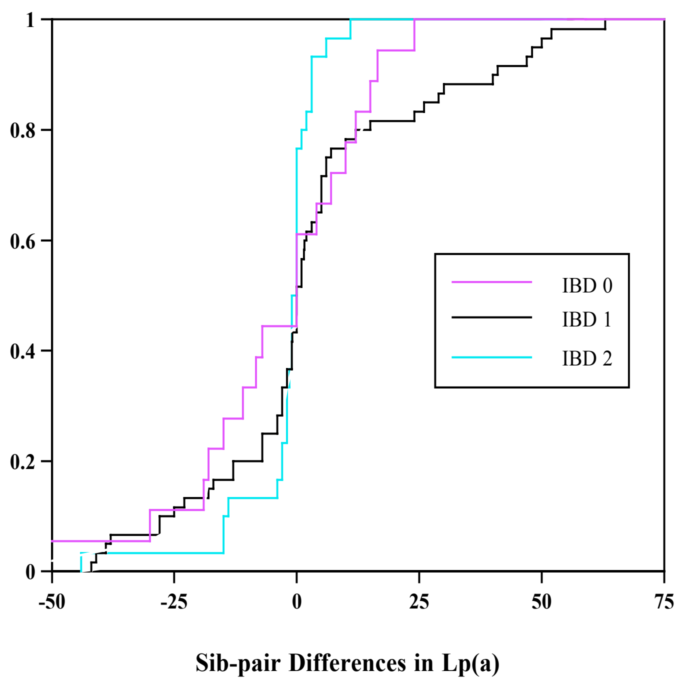

Fig 1. *Empirical distribution functions of phenotypic differences for the sib-pair data.*

distribution functions for plasma Lipoprotein (a) differences within sib-pairs for a sample of Caucasian individuals from the Dallas metroplex area. The pairs of siblings were classified into groups according to the number of shared alleles identical by descent.

Note that the assumptions of symmetry and peakedness are close to being satisfied, but the plots also show areas where these characteristics do not hold. We will illustrate our estimators later in section 4, by computing them for this example.

The points a and b about which peakedness of F and G will obtain, will be assumed known throughout this work. In the linkage example to be considered in section 4, the assumption of known a and b can be justified under the assumption of bivariate normality of the siblings' phenotypes with equal means. This is a common assumption in the literature. Thus, irrespective of whether a and b are known or unknown, the difference of the phenotypes is always symmetric about zero. Dropping the assumption of bivariate normality of the sib-pairs phenotypes, existing models, see *e.g.* Liu (1988) Table 15.7, yield a zero mean for the phenotypic differences. We, therefore, will assume that a and b are zero.

The goals of this paper are to develop estimators for symmetric F and G, which satisfy (1.2), and to delineate their asymptotic theory.

Under the assumption that F and G are discrete distributions satisfying (1.2), El Barmi and Rojo (1997) provided the nonparametric maximum likelihood estimators of F and G and tests were given to test the hypothesis of homogeneity of F and G against the alternative that F and G satisfy (1.2). Rojo, Batun, and Durazo (2007) proposed estimators for continuous F and G, when (1.2) holds and the case of censored data was also considered, but without the symmetry assumption. Rojo and Batun-Cutz (2007) proposed estimators for symmetric F and G when (1.2) holds using results from Schuster (1975), and the asymptotic theory was delineated for the case when both n and $m \to \infty$. El Barmi and Mukerjee (2008), following the ideas in Rojo (2004) and Rojo and Batun (2007), proposed estimators which are consistent for F (G) and their asymptotic theory was developed. Unfortunately, the proofs of their asymptotic results for the estimators of F and G depend on letting **both** n and m increase to infinity. The purpose of this paper is to consider modifications of the estimators proposed by Rojo and Batun-Cutz (2007) that yield consistent estimators for F (G) when only n (m) $\to \infty$. The asymptotic distribution theory is considered and a simulation study compares the estimators to the estimator of El Barmi and Mukerjee (2008).

The organization of this paper is as follows: Sections 2 proposes the estimators and finite sample properties are discussed. Section 3 delineates the asymptotic theory showing that the estimators are strongly and uniformly consistent and their asymptotic theory is developed. Section 4 illustrates the new estimators using the sib-pair data, and section 5 discusses the results of computer simulations which compare the bias and mean squared error of the new estimators with the bias and mean squared error of the estimators of Rojo and Batun-Cutz (2007) and El Barmi and Mukerjee (2008).

Although the estimators proposed in Rojo and Batun-Cutz (2007) have larger absolute bias than the estimators proposed here, the selection of the better estimators based on Mean Squared Error (MSE) behavior is not as clear. Whereas the new estimators have smaller MSE in a neighborhood of zero, the estimators of Rojo and Batun-Cutz have smaller MSE in the tails of the distributions, and the region of the support of the distribution where the latter estimators behave better seems to increase as the tail-heaviness of the distributions increase.

2. New Estimators and Their Finite Sample Properties

Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ be independent random samples from the symmetric distributions (about 0) F and G respectively, and let F_n and G_m be the empirical distribution functions based on $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$. Suppose than $F >_p G$. Rojo and Batun-Cutz (2007) considered the problem of the estimation of F and G under the peakedness restriction and proposed the following strongly uniformly consistent estimators

$$F^1_{n,m} = \Phi_1(\Phi_2(F_n, \Phi_1(G_m))) \tag{2.1}$$

$$F^2_{n,m} = \Phi_2(\Phi_1(F_n), \Phi_1(G_m)), \tag{2.2}$$

where Φ_1 and Φ_2 are operators defined by

$$\Phi_1(f)(x) = \frac{1}{2}(f(x) + 1 - f(-x^-)), \text{ and}$$

$$\Phi_2(f,g)(x) = \begin{cases} \max\{f(x), g(x)\} & \text{if} \quad x \geq 0 \\ \min\{f(x), g(x)\} & \text{if} \quad x < 0. \end{cases}$$

Note that the operator Φ_1 symmetrizes the function f, Schuster (1975), and the operator Φ_2 imposes the "stochastic order" restriction (see, e.g., Lo (1987), Rojo and Ma (1996), and Rojo (2004)). Unfortunately the estimators $F^i_{n,m}$, for $i = 1, 2$ do not converge to F when only $n \to \infty$. This follows since, for example, for $F^2_{n,m}$ when $x > 0$ and $\varepsilon > 0$,

$$\lim_{n\to\infty} P[F^2_{n,m}(x) - F(x) > \varepsilon] \quad \geq \quad P[\Phi_1(G_m(x)) - F(x) > \varepsilon] > 0.$$

This is a drawback of $F^2_{n,m}$ that is also shared by $F^1_{n,m}$ and $G^i_{n,m}$ for $i = 1, 2$, and the strong uniform consistency of these estimators requires that both m and n tend to infinity. To circumvent this problem, new estimators are proposed here.

2.1. Definition of the New Estimators

Let $\widehat{F}_n = \Phi_1(F_n)$ and $\widehat{G}_m = \Phi_1(G_m)$ be the symmetrized empirical distribution functions (Schuster, 1975). Then the empirical distribution function, and the symmetrized empirical distribution function of the combined samples are defined as follows:

$$C_{n,m} = \frac{n}{m+n} F_n + \frac{m}{n+m} G_m \text{ and}$$
$$\widehat{C}_{n,m} = \Phi_1(C_{n,m}) = \frac{n}{m+n} \widehat{F}_n + \frac{m}{n+m} \widehat{G}_m \tag{2.3}$$

respectively. Then our new estimators for F and G are

$$\widehat{F}^1_{n,m} = \Phi_1(\Phi_2(F_n, C_{n,m})), \tag{2.4}$$

$$\widehat{G}^1_{n,m} = \Phi_1(\Phi^*_2(G_m, C_{n,m})), \tag{2.5}$$

$$\widehat{F}^2_{n,m} = \Phi_2(\Phi_1(F_n), \Phi_1(C_{n,m})), \text{ and} \tag{2.6}$$

$$\widehat{G}^2_{n,m} = \Phi^*_2(\widehat{G}_m, \widehat{C}_{n,m}), \tag{2.7}$$

where

$$\Phi^*_2(f, g)(x) = \begin{cases} \min\{f(x), g(x)\} & \text{if} \quad x \geq 0 \\ \max\{f(x), g(x)\} & \text{if} \quad x < 0. \end{cases}$$

Note that the estimators $\widehat{F}^1_{n,m}$ and $\widehat{G}^1_{n,m}$ first impose the constraint of "stochastic order" by requiring that the estimator of F (G) be larger (smaller) than $C_{n,m}$ for $x \geq 0$ and smaller (larger) than $C_{n,m}$ for $x < 0$. The second requirement of symmetry is then imposed by the operator Φ_1. By contrast, the estimators $\widehat{F}^2_{n,m}$ and $\widehat{G}^2_{n,m}$, first impose the constraint of symmetry and then, through the operator Φ_2, the constraint of "stochastic order" is imposed.

El Barmi and Mukerjee (2008) proposed estimators for F and G when $F <_p G$. In our notation and making the appropriate change for the case $F >_p G$, their estimator for F is given, for $x \geq 0$, by

$$F^*_{nm}(x) \quad = \quad \frac{1}{2}(1 + \max\left\{F_n(x) - F_n(-x^-), C_{nm}(x) - C_{nm}(-x^-)\right\}).$$

This estimator is the same as our estimator $\widehat{F}^2_{n,m}$ since for $x \geq 0$,

$$\begin{aligned}
\widehat{F}^2_{n,m}(x) &= \max\left\{\frac{1}{2}(1+F_n(x)-F_n(-x^-)), \frac{1}{2}(1+C_{nm}(x)-C_{nm}(-x^-))\right\} \\
&= \frac{1}{2}+\frac{1}{2}\max\left\{F_n(x)-F_n(-x^-), C_{nm}(x)-C_{nm}(-x^-)\right\}) \\
&= F^*_{nm}(x).
\end{aligned}$$

Therefore, by symmetry, $\widehat{F}^2_{n,m} = F^*_{nm}$.

2.2. Bias Functions

The operator Φ_1 does not introduce any bias in the "symmetrization" procedure. In fact, it is well known that $\widehat{F}_n$ and $\widehat{G}_m$ are unbiased estimators for F and G, and have smaller variance than F_n and G_m respectively. However, the operators Φ_2 and Φ^*_2 do introduce bias when estimating F and G. The bias function of the estimators are discussed next and compared to the estimator provided by El Barmi and Mukerjee (2008).

For $x \geq 0$ define $F^+_n(x) = \frac{1}{n}\sum_{i=1}^n I_{[-x\leq X_i\leq x]}$, $F^{+*}_{nm} = \max\{F^+_n, \frac{nF^+_n+mG^+_m}{n+m}\}$ and finally, let $F^*_{nm} = \frac{1}{2}(1+F^{+*}_{nm})$; G^+_m, $G^{+*}_{n,m}$ and $G^*_{n,m}$ are defined similarly. The estimator F^*_{nm} is the estimator for F studied by El Barmi and Mukerjee (2008) following ideas of Rojo (2004). Note that for $x \geq 0$,

$$\begin{aligned}
E(F^*_{nm}(x)) &= \frac{1}{2}+\frac{1}{2}E(F^{+*}_{nm}(x)) \\
&= \frac{1}{2}+\frac{1}{2}E\{F^+_n(x)+\max\{0, \frac{m}{m+n}(G^+_m(x)-F^+_n(x))\}\} \\
&= \frac{1}{2}+\frac{1}{2}E(F^+_n(x))+\frac{m}{2(m+n)}E\{\max(0, G^+_m(x)-F^+_n(x))\}
\end{aligned}$$

and since $\frac{1}{2}+\frac{1}{2}E(F^+_n(x)) = F(x)$,

$$Bias(F^*_{nm}(x)) = \frac{m}{2(m+n)}E\{\max(0, G^+_m(x)-F^+_n(x))\}. \tag{2.8}$$

Note that $Bias(F^*_{nm}(x)) \to 0$ as $\frac{n}{m} \to \infty$. Since our estimator $\widehat{F}^2_{nm}$ defined by (2.6) turns out to be equal F^*_{nm}, then its bias function is also given by (2.8).

Now consider the estimator $\widehat{F}^1_{nm}$ given by (2.4). For $x \geq 0$,

$$\begin{aligned}
\widehat{F}^1_{n,m}(x) &= \Phi_1(\max(F_n(x), C_{nm}(x))) \\
&= \frac{1}{2}\{1+\max(F_n(x), C_{nm}(x))-\min(F_n(-x^-), C_{nm}(-x^-))\} \\
&= \frac{1}{2}\left\{1+F_n(x)-F_n(-x^-)+\max(0, C_{nm}(x)-F_n(x))\right. \\
&\quad \left.+\max(0, F_n(-x^-)-C_{nm}(-x^-))\right\}.
\end{aligned}$$

Thus, $E(\widehat{F}^1_{n,m}(x)) = F(x)+\frac{1}{2}E(\max(0, C_{nm}(x)-F_n(x)))+\frac{1}{2}E(\max(0, F_n(-x^-)-$

$C_{nm}(-x^-))$) and then, for $x \geq 0$

$$
\begin{aligned}
Bias(\widehat{F}^1_{n,m}(x)) &= \frac{1}{2}E(\max(0, \frac{m}{n+m}(G_m(x) - F_n(x)))) \\
&\quad + \frac{1}{2}E(\frac{m}{n+m}\max(0, F_n(-x^-) - G_m(-x^-))) \\
&= \frac{m}{2(m+n)}\{E(\max(0, G_m(x) - F_n(x))) \\
&\quad + E(\max(0, F_n(-x^-) - G_m(-x^-)))\} \\
&\geq \frac{m}{2(m+n)}E(\max(0, G^+_m(x) - F^+_n(x))) = Bias(F^*_{nm}).
\end{aligned}
$$

This result will also follow from the fact that $\widehat{F}^1_{nm} >_p \widehat{F}^2_{nm} = F^*_{nm}$.

Next consider the estimator F^2_{nm} defined in equation (2.2) and in Rojo and Batun-Cutz (2007):

$$
F^2_{nm}(x) = \max\left\{\frac{1}{2}(1 + F_n(x) - F_n((-x)^-)), \frac{1}{2}(1 + G_m(x) - G_m((-x)^-))\right\}.
$$

It follows easily that $E(F^2_{nm}(x)) = F(x) + \frac{1}{2}E(\max(0, G^+_m(x) - F^+_n(x)))$ and hence $Bias(F^2_{nm}(x)) = \frac{1}{2}E(\max(0, G^+_m(x) - F^+_n(x))) > Bias(F^*_{nm})$, for $x \geq 0$.

Finally, consider the estimator F^1_{nm} given in Rojo and Batun-Cutz (2007). For $x \geq 0$

$$
\begin{aligned}
F^1_{nm}(x) &= \frac{1}{2}\left\{1 + \max(F_n(x), \frac{1}{2}(1 + G_m(x) - G_m((-x)^-)))\right. \\
&\quad \left. - \min(F_n(-x), \frac{1}{2}(1 + G_m((-x)^-) - G_m((x))))\right\} \\
&= \frac{1}{2}(1 + F_n(x) - F_n((-x)^-)) + \frac{1}{2}\max(0, \frac{1}{2}(1 - 2F_n(x) + G^+_m(x))) \\
&\quad + \frac{1}{2}\max(0, \frac{1}{2}(-1 + 2F_n((-x)^-) + G^+_m(x))).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
E(F^1_{nm}(x)) = F(x) &+ \frac{1}{4}E(\max(0, (1 - 2F_n(x) + G^+_m(x)))) \\
&+ \frac{1}{4}E(\max(0, (-1 + 2F_n((-x)^-) + G^+_m(x)))).
\end{aligned}
$$

Then, for $x \geq 0$,

$$
\begin{aligned}
Bias(F^1_{nm}(x)) &= \frac{1}{4}E\left(\max\{0, G^+_m(x) - F^+_n(x) - F_n((-x)^-) - F_n(x) + 1\}\right) \\
&\quad + \frac{1}{4}E\left(\max\{0, G^+_m(x) - F^+_n(x) - 1 + F_n(x) + F_n((-x)^-)\}\right).
\end{aligned}
$$

The last expression is then seen to be equal to

$$
\begin{aligned}
&\frac{1}{4}E(\max\{\max(0, G^+_m(x) - F^+_n(x) - F_n((-x)^-) - F_n(x) + 1), \\
&\qquad \max(G^+_m(x) - F^+_n(x) + F_n((-x)^-) + F_n(x) - 1, 2(G^+_m - F^+_n)\}) \\
&\qquad \geq \frac{1}{4}E(\max(0, 2(G^+_m(x) - F^+_n(x)))) = Bias(F^*_{n,m}).
\end{aligned}
$$

The corresponding inequalities for the case of $x < 0$ follow by symmetry. Similar results may be obtained for the estimators $G^1_{n,m} = \Phi_1(\Phi^*_2(\Phi_1(F_n), G_m))$, $G^2_{n,m} = \Phi^*_2(\Phi_1(F_n), \Phi_1(G_m))$, and $\widehat{G}^1_{n,m}$ and $\widehat{G}^2_{n,m}$. It is easy to see that all the estimators for F have positive (negative) bias for $x > 0$ ($x < 0$), while the estimators for G have negative (positive) bias for $x > 0$ ($x < 0$). The following proposition summarizes the results about the bias functions.

Proposition 1. *Let $F >_p G$ be symmetric distribution functions, and let $X_1, \dots, X_n$ and $Y_1, \dots, Y_m$ be independent random samples from F and G respectively. The bias functions of the estimators for F and G given by (2.1), (2.2), (2.4), (2.5), (2.6), and (2.7), satisfy the following properties. For all x,*

$$
\begin{aligned}
(i)|Bias(\widehat{F}^1_{n,m}(x))| &\geq |Bias(\widehat{F}^2_{n,m}(x))| \\
&= \frac{m}{2(m+n)} E\{\max(0, G^+_m(|x|) - F^+_n(|x|)\} = |Bias(F^*_{n,m}(x))|
\end{aligned}
$$

(ii) $|Bias(F^1_{n,m}(x))| \geq |Bias(F^2_{n,m}(x))| \geq |Bias(\widehat{F}^2_{n,m}(x))|$

(iii) $|Bias(\widehat{G}^1_{n,m}(x))| \geq |Bias(\widehat{G}^2_{n,m}(x))| = -\frac{m}{2(m+n)} E\{\min(0, F^+_n(|x|) - G^+_m(|x|))\}$

(iv) $|Bias(G^1_{n,m}(x))| \geq |Bias(G^2_{n,m}(x))| \geq |Bias(\widehat{G}^2_{n,m}(x))|$.

2.3. Estimators as Projections onto Appropriate Convex Spaces

Recall the definitions of the new estimators given by (2.4) - (2.7). Schuster (1975) showed that the operator Φ_1 projects its argument to its closest symmetric distribution. That is, letting $\mathcal{F}$ be the convex set of symmetric distributions about zero, then for an arbitrary distribution H, $\| \Phi_1(H) - H \|_\infty = \inf_{F \in \mathcal{F}} \| H - F \|_\infty$. Rojo and Ma (1996), and Rojo and Batun-Cutz (2007) have shown that the operator Φ_2 has the property that for arbitrary distributions H and G, $|\Phi_2(H(x), G(x)) - H(x)| = \inf_{F \in \mathcal{F}^*} |F(x) - G(x)|$, where $\mathcal{F}^*$ is the convex set of distributions F satisfying (1.3). Thus, for F and G distribution functions let

$$\mathcal{F}_1 = \{\text{distribution functions } F \text{ satisfying (1.3) with } G \text{ replaced by } C_{n,m}\}$$

$$\mathcal{F}^*_1 = \{\text{symmetric distributions } F \text{ satisfying (1.3) with } G \text{ replaced by } \Phi_1(C_{n,m})\}$$

$$\text{and} \qquad \mathcal{F}^*_2 = \{\text{all symmetric at 0 distribution functions}\}.$$

Thus the estimator $\widehat{F}^2_{n,m}$ first projects F_n onto $\mathcal{F}^*_2$ and then projects $\Phi_1(F_n)$ onto $\mathcal{F}^*_1$. By contrast, the estimator $\widehat{F}^1_{n,m}$ first projects F_n onto $\mathcal{F}_1$ to obtain $\Phi_2(F_n, C_{n,m})$ and then projects the latter onto $\mathcal{F}^*_1$. With appropriate changes in the above notation, similar comments hold for the estimators $\widehat{G}^i_{n,m}$ for $i = 1, 2$.

2.4. Peakedness Order of New and Previous Estimators

By construction, the estimators $F^i_{n,m}$ and $\widehat{F}^i_{n,m}$, for $i = 1, 2$ are more peaked than the estimators $G^i_{n,m}$ and $\widehat{G}^i_{n,m}$, respectively. Rojo and Batun-Cutz (2007) showed that $F^1_{n,m} >_p F^2_{n,m}$. The next theorem provides comparisons in terms of peakedness for several of the estimators and provides a simple relationship between $F^2_{n,m}$ and $\widehat{F}^2_{n,m}$.

Lemma 1. *Let $F >_p G$ be symmetric distribution functions, and let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ be independent random samples from F and G respectively. Consider the estimators for F and G given by (2.1), (2.2), (2.4), (2.5), (2.6), (2.7). Then*

(i) $\widehat{F}^2_{n,m} = \dfrac{n}{n+m}\widehat{F}_n + \dfrac{m}{n+m}F^2_{n,m}$

(ii) $\widehat{F}^1_{n,m} >_p \widehat{F}^2_{n,m} >_p \widehat{G}^2_{n,m} >_p \widehat{G}^1_{n,m}$

(iii) $F^1_{n,m} >_p F^2_{n,m} >_p \widehat{F}^2_{n,m}$, *and* $G^1_{n,m} <_p G^2_{n,m} <_p \widehat{G}^2_{n,m}$.

Proof. (i) For $x \geq 0$,

$$\begin{aligned}\widehat{F}^2_{n,m}(x) = \max\{\widehat{F}_n(x), \widehat{C}_{n,m}(x)\} &= \frac{n}{n+m}\widehat{F}_n(x) + \frac{m}{n+m}\max\{\widehat{F}_n(x), \widehat{G}_m(x)\} \\ &= \frac{n}{n+m}\widehat{F}_n(x) + \frac{m}{n+m}F^2_{n,m}(x).\end{aligned}$$

The result then follows by symmetry.

(ii) First we prove that $\widehat{F}^1_{n,m} >_p \widehat{F}^2_{n,m}$. Let $x \geq 0$, then

$$\begin{aligned}\widehat{F}^1_{n,m}(x) &= \frac{1}{2}\left[\max\{F_n(x), C_{n,m}(x)\} + 1 - \min\{F_n((-x)^-), C_{n,m}((-x)^-)\}\right] \\ &\geq \frac{1}{2}\left[C_{n,m}(x) + 1 - C_{n,m}((-x)^-)\right] = \widehat{C}_{n,m}(x). \qquad (2.9)\end{aligned}$$

Using similar arguments it can be shown that $\widehat{F}^1_{n,m}(x) \geq \widehat{F}_n(x)$. Therefore, combining the last inequality and (2.9) we obtain $\widehat{F}^1_{n,m}(x) \geq \widehat{F}^2_{n,m}(x)$. The result follows from symmetry.

We now prove that $\widehat{F}^2_{n,m} >_p \widehat{G}^2_{n,m}$. For $x \geq 0$, $\widehat{F}^2_{n,m}(x) = \max\{\widehat{F}_n(x), \widehat{C}_{n,m}(x)\} \geq \widehat{C}_{n,m}(x) \geq \widehat{G}^2_{n,m}(x)$. The result follows by symmetry.

Since for $x \geq 0$, $\widehat{G}^1_{n,m}(x) \leq \widehat{C}_{n,m}(x)$ and $\widehat{G}^1_{n,m}(x) \leq \widehat{G}_m(x)$. Then $\widehat{G}^2_{n,m} >_p \widehat{G}^1_{n,m}$ by symmetry.

Finally consider (iii). The result that $F^1_{n,m} >_p F^2_{n,m}$ follows from Rojo and Batun-Cutz (2007). The result that $F^2_{n,m} >_p \widehat{F}^2_{n,m}$ follows from the arguments used to prove that Bias$(F^2_{n,m}) \geq$Bias$(\widehat{F}^2_{n,m})$.

Note that (i) implies that for $x \geq 0$, $Bias(F^2_{n,m}(x)) = \frac{m+n}{m}Bias(\widehat{F}^2_{n,m}(x))$, so that $|Bias(F^2_{n,m}(x))| = \frac{m+n}{m}|Bias(\widehat{F}^2_{n,m}(x))|$ for all x, thus providing a more accurate description of the result about bias given in proposition 1.

3. Asymptotics

This section discusses the strong uniform convergence of the estimators and their asymptotic distribution theory. One important aspect of the asymptotic results for the estimators $\hat{F}^i_{n,m}$ $(\hat{G}^i_{n,m})$, $i = 1, 2$ discussed here is that they hold even when only n (m) tends to infinity. This is in sharp contrast with the results of Rojo and Batun-Cutz (2007) and those of El Barmi and Mukerjee (2008). We discuss the strong uniform convergence first.

3.1. Strong Uniform Convergence

The following theorem provides the strong uniform convergence of the estimators $\widehat{F}^i_{n,m}$ ($\widehat{G}^i_{n,m}$), $i=1,2$. The results use the strong uniform convergence of the symmetrized $\hat{F}_n$ ($\hat{G}_m$) to F (G) as $n\to\infty$ ($m\to\infty$), Schuster (1975).

Theorem 3.1. *Let F and G be symmetric distribution functions with $F >_p G$, and let $X_1,\dots,X_n$ and $Y_1,\dots,Y_m$ be independent random samples from F and G respectively. Then,*

(i) *$\widehat{F}^i_{n,m}$, for $i=1,2$, converge uniformly with probability one to F as $n\to\infty$.*
(ii) *$\widehat{G}^i_{n,m}$ for $i=1,2$ converge uniformly with probability one to G as $m\to\infty$.*

Proof. **(i)** Consider $\widehat{F}^2_{n,m}$ first. Then, for $x\geq 0$,

$$(3.1)\quad \widehat{F}^2_{n,m}(x)-F(x)=\widehat{F}_n(x)-F(x)+\frac{m}{n+m}\max\{0,\widehat{G}_m(x)-\widehat{F}_n(x)\}.$$

But, since $F(x)\geq G(x)$,

$$\widehat{G}_m(x)-\widehat{F}_n(x)\leq \widehat{G}_m(x)-G(x)+F(x)-\widehat{F}_n(x).$$

Hence

$$\begin{aligned}\max\{0,\widehat{G}_m(x)-\widehat{F}_n(x)\} &\leq \max\{0,\widehat{G}_m(x)-G(x)+F(x)-\widehat{F}_n(x)\}\\ (3.2)\qquad &\leq |\widehat{G}_m(x)-G(x)|+|\widehat{F}_n(x)-F(x)|\end{aligned}$$

and therefore, the left side of (3.1) is bounded above by

$$|\widehat{F}_n(x)-F(x)|+(\frac{m}{m+n})\{|\widehat{G}_m(x)-G(x)|+|\widehat{F}_n(x)-F(x)|.\}$$

Since $\widehat{F}_n$, and $\widehat{G}_m$ are strongly and uniformly consistent for F and G, then as $n\to\infty$, with probability one,

$$\sup_{x\geq 0}|\widehat{F}^2_{n,m}(x)-F(x)|\to 0,$$

regardless of whether $m\to\infty$ or not. When $x<0$ the result follows by symmetry.

Let us now consider the case of $\widehat{F}^1_{n,m}$. For $x\geq 0$

$$\begin{aligned}(3.3)\qquad \widehat{F}^1_{n,m}(x)-F(x) &= \widehat{F}_n(x)-F(x)+\frac{1}{2}\frac{m}{n+m}\left[\max\{0,G_m(x)-F_n(x)\}\right.\\ &\quad \left.-\min\{0,G_m(-x^-)-F_n(-x^-)\}\right].\end{aligned}$$

Since $F(x)\geq G(x)$ and $F(-x)\leq G(-x)$, then it follows that

$$\max\{0,G_m(x)-F_n(x)\}-\min\{0,G_m(-x^-)-F_n(-x^-)\}$$

is bounded above by

$$\max\{0,G_m(x)-G(x)+F(x)-F_n(x)\}-\min\{0,G_m(-x^-)-G(-x)+F(-x)-F_n(-x^-)\}$$

and, therefore, the left side of (3.2) is bounded above by

$$
\begin{aligned}
|\widehat{F}_n(x) - F(x)| \quad &+ \quad \frac{1}{2}\frac{m}{m+n}(|G_m(x) - G(x)| + |F(x) - F_n(x)| \\
&+ \quad |G_m(-x^-) - G(-x)| + |F(-x) - F_n(-x^-)|). \qquad (3.4)
\end{aligned}
$$

Taking the supremum over x in (3.4), and then letting $n \to \infty$, the result follows, whether $m \to \infty$ or not, from the strong uniform convergence of $\widehat{F}_n$, G_m, and F_n to F, G, and F respectively. The result for $x < 0$ follows by symmetry.

(iii) The proof for the strong uniform convergence of $\widehat{G}^2_{n,m}$ to G, when only $m \to \infty$ is similar. We sketch the proof. For $x < 0$

$$\widehat{G}^2_{n,m}(x) - G(x) \quad = \quad \widehat{G}_m(x) - G(x)) + \frac{n}{n+m} \max\{0, \widehat{F}_n(x) - \widehat{G}_m(x)\}.$$

Therefore, since $F(x) < G(x)$ for $x < 0$, $\max\{0, \widehat{F}_n(x) - \widehat{G}_m(x)$ is bounded above by

$$\max\{0, \widehat{F}_n(x) - F(x) + G(x) - \widehat{G}_m(x)\} \leq |\widehat{F}_n(x) - F(x)| + |G(x) - \widehat{G}_m(x)|.$$

When $m \to \infty$, the result follows, regardless of whether $n \to \infty$ or not, from the strong uniform convergence of $\widehat{F}_n$ and $\widehat{G}_m$ and using a symmetry argument to handle the case of $x > 0$.
(iv) This case is omitted as it follows from similar arguments.

3.2. Weak Convergence

Consider first the point-wise asymptotic distribution for $\widehat{F}^i_{n,m}$, $i = 1, 2$. Recall that

$$\sqrt{n}(\widehat{F}_n(x) - F(x)) \xrightarrow{W} N\left(0, \frac{F(-|x|)(2F(|x|) - 1)}{2}\right).$$

Therefore, when $n/m \to \infty$, and using (3.1)-(3.4), Slutsky's theorem and the central limit theorem for $\widehat{F}_n$, we get the following result:

$$\sqrt{n}(\widehat{F}^i_{nm}(x) - F(x)) \quad \xrightarrow{W} \quad N\left(0, \frac{F(-|x|)(2F(|x|) - 1)}{2}\right). \qquad (3.5)$$

Thus, under these conditions, $\widehat{F}^i_{n,m}$, $i = 1, 2$, are $\sqrt{n}$–equivalent and have the same asymptotic distribution as the symmetrized $\widehat{F}_n$ which happens to have the same asymptotic limit as in the case when G is completely known. Note that this result assumes only that $n/m \to \infty$ and hence the result holds if m is fixed and $n \to \infty$. This is in sharp contrast with the results of El Barmi and Mukerjee (2008) that require that both n and m tend to infinity. Similar results hold for the estimators $\widehat{G}^i_{n,m}$, $i = 1, 2$. These are summarized in the following theorem.

Theorem 3.2. *Suppose that $F >_p G$ and let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ be random samples from F and G respectively. Then for $i = 1, 2$,*

(i) If $n/m \to \infty$ then

$$\sqrt{n}(\widehat{F}^i_{nm}(x) - F(x)) \quad \xrightarrow{\mathcal{D}} \quad N\left(0, \frac{F(-|x|)(2F(|x|) - 1)}{2}\right).$$

(ii) If $m/n \to \infty$ then

$$\sqrt{n}(\widehat{G}^i_{nm}(x) - G(x)) \overset{\mathcal{D}}{\to} N\left(0, \frac{G(-|x|)(2G(|x|)-1)}{2}\right).$$

We now turn our attention to the weak convergence of the processes

$$\left\{\sqrt{n}\left(\widehat{F}^i_{nm}(x) - F(x)\right) : -\infty < x < \infty\right\},$$

and

$$\left\{\sqrt{n}\left(\widehat{G}^i_{nm}(x) - G(x)\right) : -\infty < x < \infty\right\},$$

for $i = 1, 2$. Only the results for $\widehat{F}^i_{n,m}$, $i = 1, 2$ will be discussed in detail as the results for $\widehat{G}^i_{n,m}$, $i = 1, 2$ can be obtained by similar arguments. Although the processes $\left\{\sqrt{n}\left(\widehat{F}^i_{nm}(x) - F(x)\right) : -\infty < x < \infty\right\}$ for $i = 1, 2$ are correlated, we are only interested in their marginal behavior. For that purpose let $\{W_1(x) : -\infty < x < \infty\}$ denote a mean zero Gaussian process with covariance function

$$(3.6) \quad E(W_1(x)W_1(y)) = \begin{cases} \frac{1}{2}(1-F(y))(F(x)-F(-x)) & \text{if } |y| > |x| \\ \frac{1}{2}F(x)(F(-y)-F(y)) & \text{if } |y| < |x|, \end{cases}$$

and let $\{W_2(x) : -\infty < x < \infty\}$ denote a mean zero Gaussian process with covariance function

$$(3.7) \quad E(W_2(x)W_2(y)) = \begin{cases} \frac{1}{2}(1-G(y))(G(x)-G(-x)) & \text{if } |y| > |x| \\ \frac{1}{2}G(x)(G(-y)-G(y)) & \text{if } |y| < |x|. \end{cases}$$

We have the following result:

Theorem 3.3. *Under the conditions of the previous Theorem,*

(i) If $n/m \to \infty$, then

$$\{\sqrt{n}(\widehat{F}^i_{nm}(x) - F(x)), -\infty < x < \infty\} \overset{W}{\to} \{W_1(x) : -\infty < x < \infty\}, \text{ and}$$

(ii) If $m/n \to \infty$, then

$$\{\sqrt{n}(\widehat{G}^i_{nm}(x) - G(x)), -\infty < x < \infty\} \overset{W}{\to} \{W_2(x) : -\infty < x < \infty\}.$$

Proof. The proof follows easily by the continuous mapping Theorem after observing that the weak limit of $\{\sqrt{n}(\widehat{F}_n(x) - F(x))), -\infty < x < \infty\}$ is the process $\{W_1(x) : -\infty < x < \infty\}$, together with the fact that, using (3.1),

$$(3.8) \ \widehat{F}^2_{n,m}(x) - F(x) = \widehat{F}_n(x) - F(x) + \frac{m}{n+m}\max\{0, \widehat{G}_m(x) - \widehat{F}_n(x)\},$$

with $\| \sqrt{n}\frac{m}{n+m}\{\max 0, \widehat{G}_m - \widehat{F}_n\} \|_\infty \to 0$ with probability one, where $\| \cdot \|_\infty$ denotes the sup norm. Similar arguments yield the results for the other processes.

The asymptotic theory for $\widehat{F}^2_{n,m}$ was discussed by El Barmi *et al* (2008) for the case that both n and m go to infinity and hence their result does not include our result here when m is bounded and $n \to \infty$. When $n/m \to c$ with $0 \le c < \infty$, and $F(x) > G(x)$ for all $x > 0$ the weak limit of $\{\sqrt{n}(\widehat{F}^i_{nm}(x) - F(x)), -\infty < x < \infty\}$ is $\{W_1(x) : -\infty < x < \infty\}$, for $i = 1, 2$, which is the weak limit of the process $\{\sqrt{n}(F_{n,2}(x) - F(x)), -\infty < x < \infty\}$ discussed in Rojo and Batun-Cutz (2007). Let $\{Z(x), -\infty < x < \infty\}$ represent the weak limit of the empirical process $\{\sqrt{n}(F_n(x) - F(x)), -\infty < x < \infty\}$. That is $\{Z(x), -\infty < x < \infty\}$ is a mean zero Gaussian process with covariance function $E(Z(t)Z(s)) = F(s)(1 - F(t))$ for $s \leq t$. When $n/m \to c$ with $0 \leq c < \infty$, and $F(x) = G(x)$ for all x the weak limits of $\{\sqrt{n}(\widehat{F}^i_{nm}(x) - F(x)), -\infty < x < \infty\}$ for $i = 1, 2$ follow from the results in Rojo (2004) as follows:

Theorem 3.4. *Let $F(x) = G(x)$ for all x and let $n/m \to c$ for $0 \leq c < \infty$. Let $\{W_i(x), -\infty < x < \infty\}$, for $i = 1, 2$ be the mean zero Gaussian processes with covariance functions given by (3.6) and (3.7), respectively. Let $W^*_i(x) = W_i(|x|)sgn(x)$, for $i = 1, 2$. Then*

(i) The process $\sqrt{n}(\widehat{F}^2_{n,m} - F(x)), -\infty < x < \infty\}$ converges weakly to the process $\{\max(W^*_1(x), \frac{\sqrt{c}W^*_2(x)+cW^*_1(x)}{1+c}), -\infty < x < \infty\}$ with $W^*_1 \overset{\mathcal{D}}{=} W^*_2$ and independent.

(ii) The process $\sqrt{n}(\widehat{F}^1_{n,m} - F(x)), -\infty < x < \infty\}$ converges weakly to the process $\{H(|x|)sgn(x), -\infty < x < \infty]\}$ where $H(x) = \frac{1}{2}\{\max\{Z_1(x), \frac{c}{1+c}Z_1(x) + \frac{\sqrt{c}}{1+c}Z_2(x)\} - \min\{Z_1(-x), \frac{c}{1+c}Z_1(-x) + \frac{\sqrt{c}}{1+c}Z_2(-x)\}$, and $\{Z_i(x), -\infty < x < \infty\}$, $i = 1, 2$ are independent copies of the process $\{Z(x), -\infty < x < \infty\}$.

Proof. (i) Consider $\widehat{F}^2_{n,m}$ first. When $F(x) = G(x)$ for all x, it follows from (3.8) that, for $x \geq 0$,

$$\begin{aligned}\sqrt{n}(\widehat{F}^2_{n,m}(x) - F(x)) = \max\{&\sqrt{n}(\widehat{F}_n(x) - F(x)), \sqrt{n/m}\frac{m}{n+m}\sqrt{m}(\widehat{G}_m(x) - G(x)) \\ &+ \frac{n}{n+m}\sqrt{n}(F(x) - \widehat{F}_n(x))\}. \end{aligned} \tag{3.9}$$

By the independence of $\widehat{F}_n$ and $\widehat{G}_m$ and their weak convergence to W_1 and W_2, it follows that the bivariate process

$$\{\sqrt{n/m}\frac{m}{n+m}\sqrt{m}(\widehat{G}_m(x) - G(x)), \frac{n}{n+m}\sqrt{n}(F(x) - \widehat{F}_n(x)), -\infty < x < \infty\}$$

converges weakly to the process $\{\frac{\sqrt{c}}{1+c}W_2(x), \frac{c}{1+c}W_1(x), -\infty < x < \infty\}$. Since for $x < 0$, $\widehat{F}^2_{n,m}(x) - F(x) \overset{\mathcal{D}}{=} F(-x) - \widehat{F}^2_{n,m}(-x)$, the result then follows for $0 < c < \infty$ from the continuous mapping theorem after observing that the mapping $h(x, y) = (\frac{1+c}{c}y, x + y)$ is continuous, and then applying it to (3.9) to get the result. The case of $c = 0$ follows immediately since it then follows that the second term on the right side of (3.9) converges to zero in probability.

(ii) Note that for $x > 0$

$$\begin{aligned}\widehat{F}^1_{nm}(x) - F(x) &= \frac{1}{2}\max\{F_n(x) - F(x), \\ &\qquad \frac{n}{m+n}(F_n(x) - F(x)) + \frac{m}{n+m}(G_m(x) - F(x))\} \\ &+ \frac{1}{2}\min\{F_n(-x) - F(-x), \\ &\qquad \frac{n}{m+n}(F_n(-x) - F(-x)) + \frac{m}{n+m}(G_m(-x) - F(-x))\}.\end{aligned}$$

Since the function $h(x,y,z,w) = \frac{1}{2}\left[\max\{x, x+z\} - min\{y, y+w\}\right]$ is continuous, by the continuous mapping theorem we obtain

$$\sqrt{n}(\widehat{F}^1_{nm}(x) - F(x)) \overset{W}{\to} \frac{1}{2}\Big[\max\{Z_1(x), \frac{c}{1+c}Z_1(x) + \frac{\sqrt{c}}{1+c}Z_2(x)\} \tag{3.10}$$
$$- \min\{Z_1(-x), \frac{c}{1+c}Z_1(x) + \frac{\sqrt{c}}{1+c}Z_2(-x)\}\Big] = H(x).$$

The result then follows after considering the case $x < 0$ and following a similar argument.

It has been observed, e.g. Rojo (1995a), Rojo (2004), and Rojo and Batun-Cutz (2007), that weak convergence of the processes of interest fails to hold when the underlying distributions F and G coincide at some point x_0 and are unequal in some neighborhood to the right of x_0. That is the case here as well. Suppose that $F(x_0) = G(x_0)$ for $x_0 > 0$ and $F(x) > G(x)$ for $x \in (x_0, x_0+\nu)$, $\nu > 0$. If $\frac{m}{n} \to c$, $0 < c \leq \infty$, as $m, n \to \infty$, it follows that

$$\sqrt{n}(\widehat{F}^1_{nm}(x_0) - F(x_0)) \overset{\mathcal{D}}{\to} H(|x_0|)sgn(x_0), \tag{3.11}$$

with $H(x)$ defined as in *(ii)* of theorem 3.4 with $(Z_1(x_0), Z_2(x_0))$ a zero-mean bivariate normal distribution vector with covariance $(1-F(x_0))F(x_0)$.

However, for $x \in (x_0, x_0+\nu)$ the sequence $\sqrt{n}(\widehat{F}^1_{nm}(x) - F(x))$ converges in distribution to the distribution given in (3.5). Then it can be seen, using arguments as in Rojo (1995a), that the process $\{\sqrt{n}(\widehat{F}^1_{nm}(x) - F(x)) : -\infty < x < \infty\}$ is not tight and hence cannot converge weakly.

We finish this section with results that provide the weak convergence of the processes $\{\sqrt{n}(F^i_{n,m}(x) - F(x)), -\infty < x < \infty\}$ for $i = 1, 2$, in the case that $F(x) = G(x)$ for all x.

Theorem 3.5. *Let $n/m \to c$ with $0 \leq c < \infty$, and $F(x) = G(x)$ for all x.*

(i) The process $\{\sqrt{n}(F^2_{n,m}(x) - F(x)), -\infty < x < \infty\}$ converges weakly to

$$\{sgn(x)\max\{sgn(x)W_1(x), sgn(x)\sqrt{c}W_2(x), -\infty < x < \infty\},$$

where W_1 and W_2 are independent copies of the mean zero Gaussian process with covariance function defined by (3.6).

(ii) The process $\{\sqrt{n}(F^1_{n,m}(x) - F(x)), -\infty < x < \infty\}$ converges weakly to

$$\frac{1}{2}\{\max\{Z(xsgn(x)), \sqrt{c}W(xsgn(x)\}$$
$$-sgn(x)\min\{Z(-xsgn(x)), \sqrt{c}W(-xsgn(x))\}; -\infty < x < \infty\},$$

where Z and W are independent mean zero Gaussian process with covariance functions defined by $E(Z(s)Z(t)) = F(s)(1-F(t))$ for $s < t$, and (3.6) respectively.

Proof. *(i)* The result follows from the independence of $\{\sqrt{n}(\widehat{F}^*_n(x) - F(x)), -\infty < x < \infty\}$ and $\{\sqrt{m}(\widehat{G}^*_m(x) - G(x)), -\infty < x < \infty\}$, their weak convergence to W_1 and W_2, and the continuous mapping theorem after observing that

$$\sqrt{n}(F^2_{n,m}(x) - F(x)) = sgn(x)\max\{sgn(x)\sqrt{n}(\widehat{F}^*_n(x) - F(x)),$$
$$sgn(x)(\sqrt{\frac{n}{m}}\sqrt{m}(\widehat{G}^*_m(x) - G(x))\}.$$

(ii) Consider the case of $x > 0$ and write

$$\begin{aligned}\sqrt{n}(F^1_{n,m}(x) - F(x)) &= \frac{\sqrt{n}}{2}\{1 - 2F(x) + \max(F_n(x), \widehat{G}_m(x)) - \min(F_n(-x), \widehat{G}_m(-x))\} \\ &= \frac{1}{2}\{\max\{\sqrt{n}(F_n(x) - F(x)), \sqrt{\frac{n}{m}}\sqrt{m}(\widehat{G}_m(x) - G(x))\} \\ &\quad - \min\{\sqrt{n}(F_n(-x) - F(-x)), \sqrt{\frac{n}{m}}\sqrt{m}(\widehat{G}_m(-x) - G(-x))\}\}.\end{aligned}$$

For $x < 0$, a similar argument leads to

$$\begin{aligned}\sqrt{n}(F^1_{n,m}(x) - F(x)) &= \frac{1}{2}\{\min\{\sqrt{n}(F_n(x) - F(x)), \sqrt{\frac{n}{m}}\sqrt{m}(\widehat{G}_m(x) - G(x))\} \\ &\quad - \max\{\sqrt{n}(F_n(-x) - F(-x)), \sqrt{\frac{n}{m}}\sqrt{m}(\widehat{G}_m(-x) - G(-x))\}\},\end{aligned}$$

The result then follows by the continuous mapping theorem after letting $n/m \to c$ with $\sqrt{n}(F_n(x) - F(x))$ and $\sqrt{m}(\widehat{G}_m(x) - G(x))$ independent and converging weakly to Z and W respectively.

4. Example with Sib-pair Data: An Illustration

In this section the estimator $\widehat{F}^2_{n,m}$ is illustrated by using the sib-paired data for the Caucasian population in the Dallas metroplex area. As can be observed from Figure 2, the new estimated distribution functions now satisfy both the constraint of symmetry and the constraint of peakedness. Thus, since siblings with two alleles identical by descent are more similar than those siblings sharing only one allele identical by descent, the distribution function denoted by IBD2 is more peaked about zero than the other two distribution functions. Similar comments apply to the other comparisons.

5. Simulation Work

Monte Carlo simulations were performed to study the finite-sample properties of the estimators $\widehat{F}^1_{nm}$ and $\widehat{F}^2_{nm}$ defined by (2.4) and (2.6) respectively. We consider various examples of underlying distributions (Normal, Cauchy, Laplace, mixtures of normals, and T), and sample sizes ($n = 10, 20, 30$ for F and $m = 10, 20, 30$ for G). Each simulation consisted of 10,000 replications.

Figures 3 and 4 show the bias functions for the four estimators considered here. Figure 3 considers $F \sim Cauchy(0, 1)$ and $G \sim Cauchy(0, 2)$, and Figure 4 considers the case with $F \sim Laplace(0, 1)$ and $G \sim Laplace(0, 1.5)$. As shown in Proposition 1, the estimator $\widehat{F}^2_{n,m}$ has uniformly the smallest absolute bias. These Figures are representative of the results that we obtained. One result that holds in all of our simulations is that $|Bias(F^1_{n,m}(x))| \geq |Bias(\widehat{F}^1_{n,m}(x))|$ for all x. Unfortunately, we are unable to prove this conjecture.

Turning our attention to comparing the estimators in terms of the Mean Squared Error (MSE) Figures 5 - 10 show the ratio of the MSE of the empirical distribution to the MSE of each of the four estimators considered here. These plots are representative of all the examples considered. As it can be seen from the plots, the

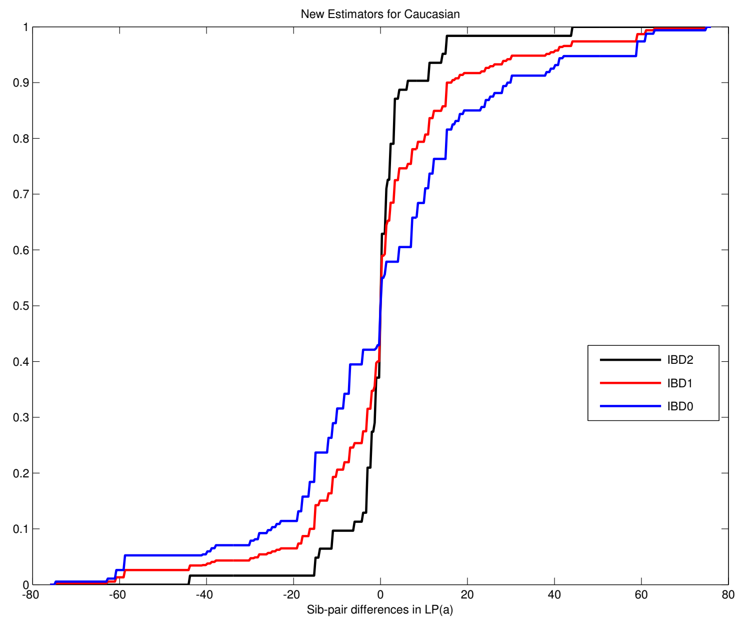

FIG 2. *Order restricted estimators for the sib-pair data incorporating peakedness.*

empirical distribution function is dominated by the estimators in every case and for all x. Whereas the estimators $\widehat{F}^i_{n,m}$ behave better than the estimators $F^i_{n,m}$, $i = 1, 2$ in a neighborhood of zero, the roles are reversed on the tails of the underlying distribution. What is observed is that the region of the support of F where $\widehat{F}^i_{n,m}$ dominate $F^i_{n,m}$, $i = 1, 2$ shrinks as the tails of the distributions get heavier, and when the distribution G is far from F. Thus, there is no clear choice among the four estimators, unless the tail is of special interest, in which case the estimator $F^2_{n,m}$ seems to be the clear choice.

6. Conclusions

Estimators were proposed for the distribution functions F and G when it is known that $F >_p G$, and F and G are symmetric about zero. The estimator for F (G) was seen to be strongly uniformly consistent when only n (m) goes to infinity and the asymptotic theory of the estimators was delineated without requiring that both n and m go to infinity. Finite sample properties of the estimators were considered and it was shown that the estimator $\widehat{F}^2_{n,m}$ has the uniformly smaller absolute bias of the four estimators considered here. The choice of which estimator is best in terms of mean squared error (mse), however, is not clear. Although the estimators $\widehat{F}^i_{n,m}$ for $i = 1, 2$ have smaller mse than the estimators $F^i_{n,m}$, $i = 1, 2$ in a neighborhood of zero, the tails are problematic for $\widehat{F}^i_{n,m}$ and the estimators $F^i_{n,m}$ tend to have smaller mse as demonstrated by the simulation study.

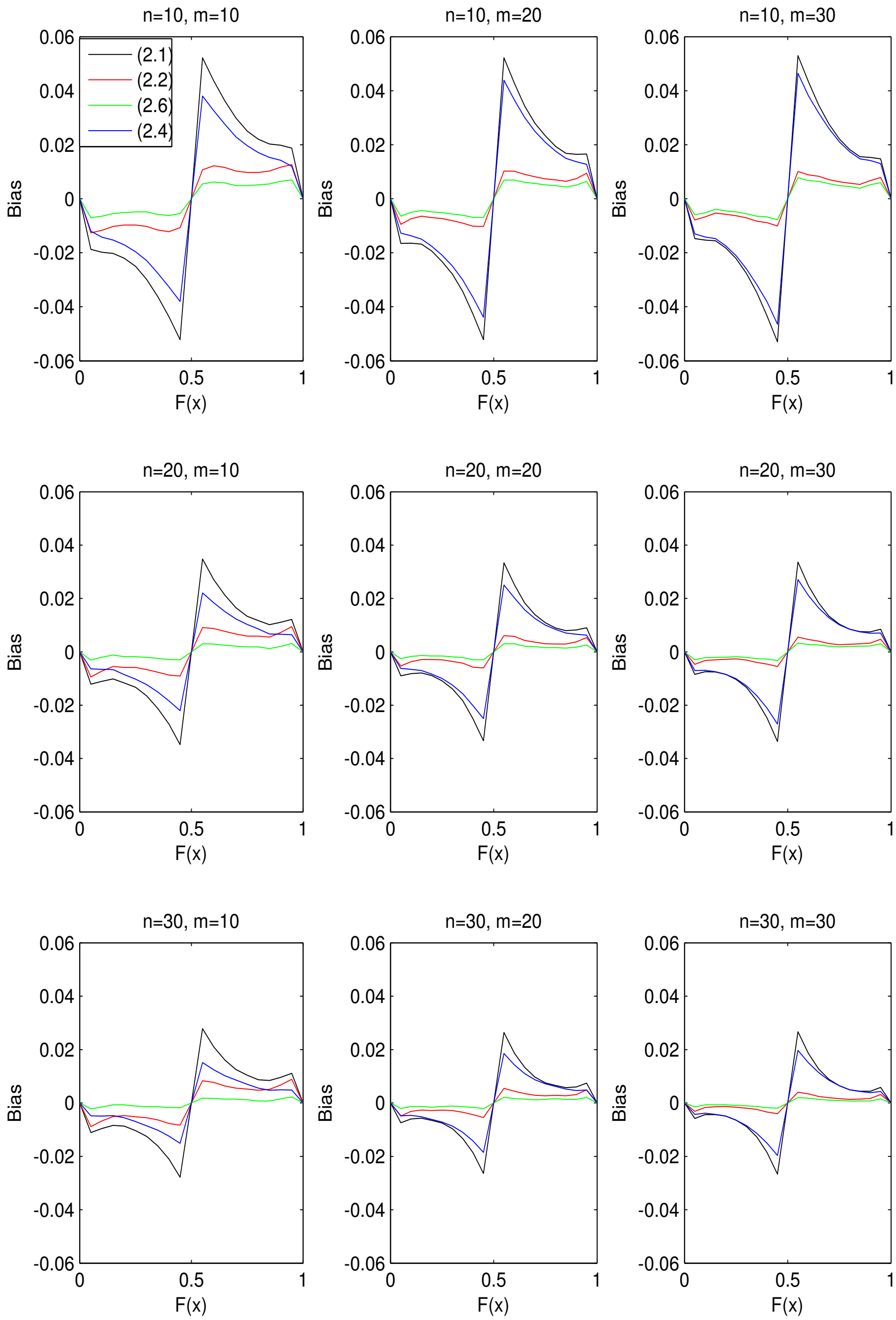

Fig 3. *Bias of the estimators when estimating F∼Cauchy(0,1) with G∼Cauchy(0,2)*

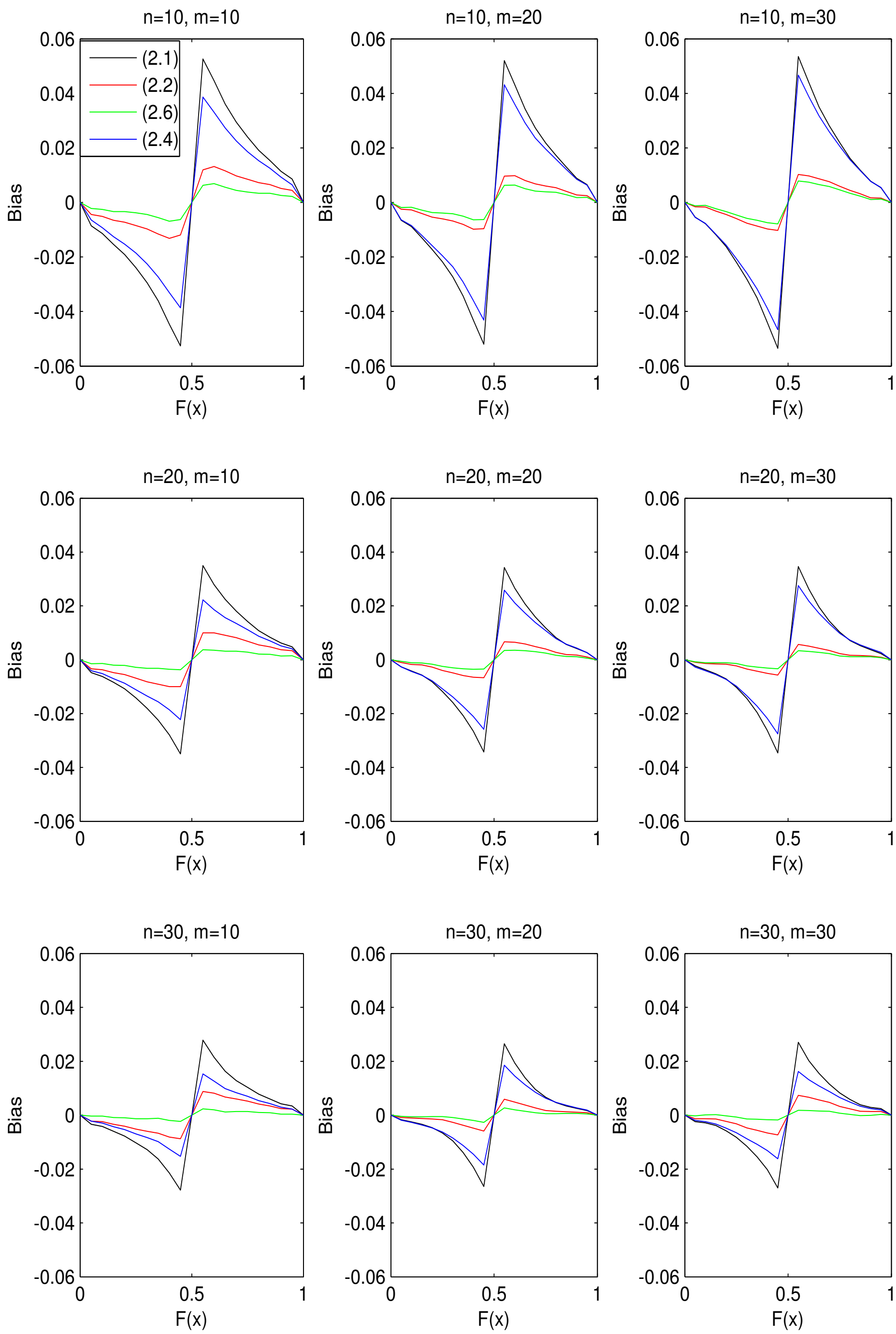

FIG 4. *Bias of the estimators when estimating F∼Laplace(0,1) with G∼Laplace(0,2).*

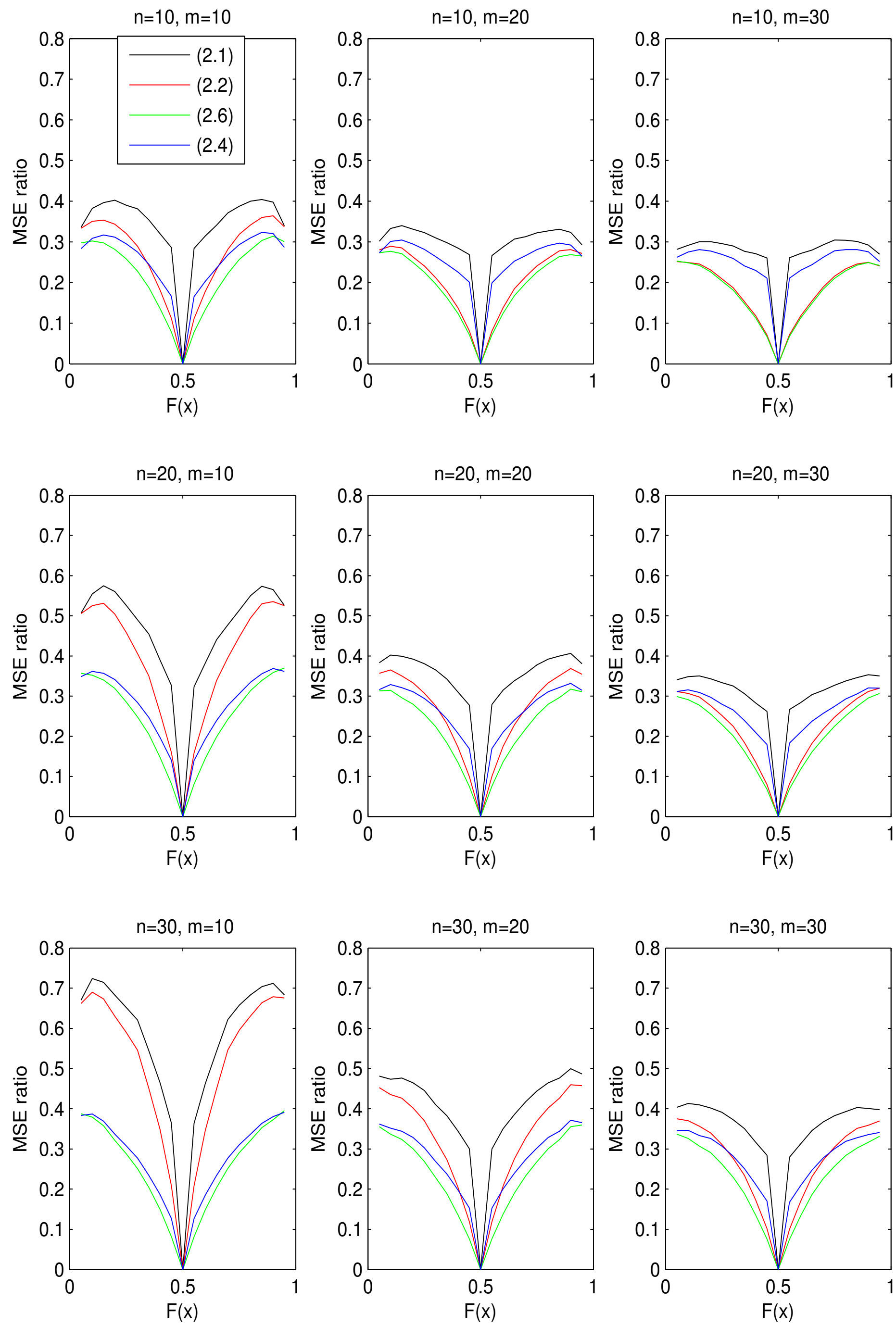

Fig 5. *Mean Squared Error of the estimators when estimating F∼Normal(0,1) with G∼Normal(0,1.1).*

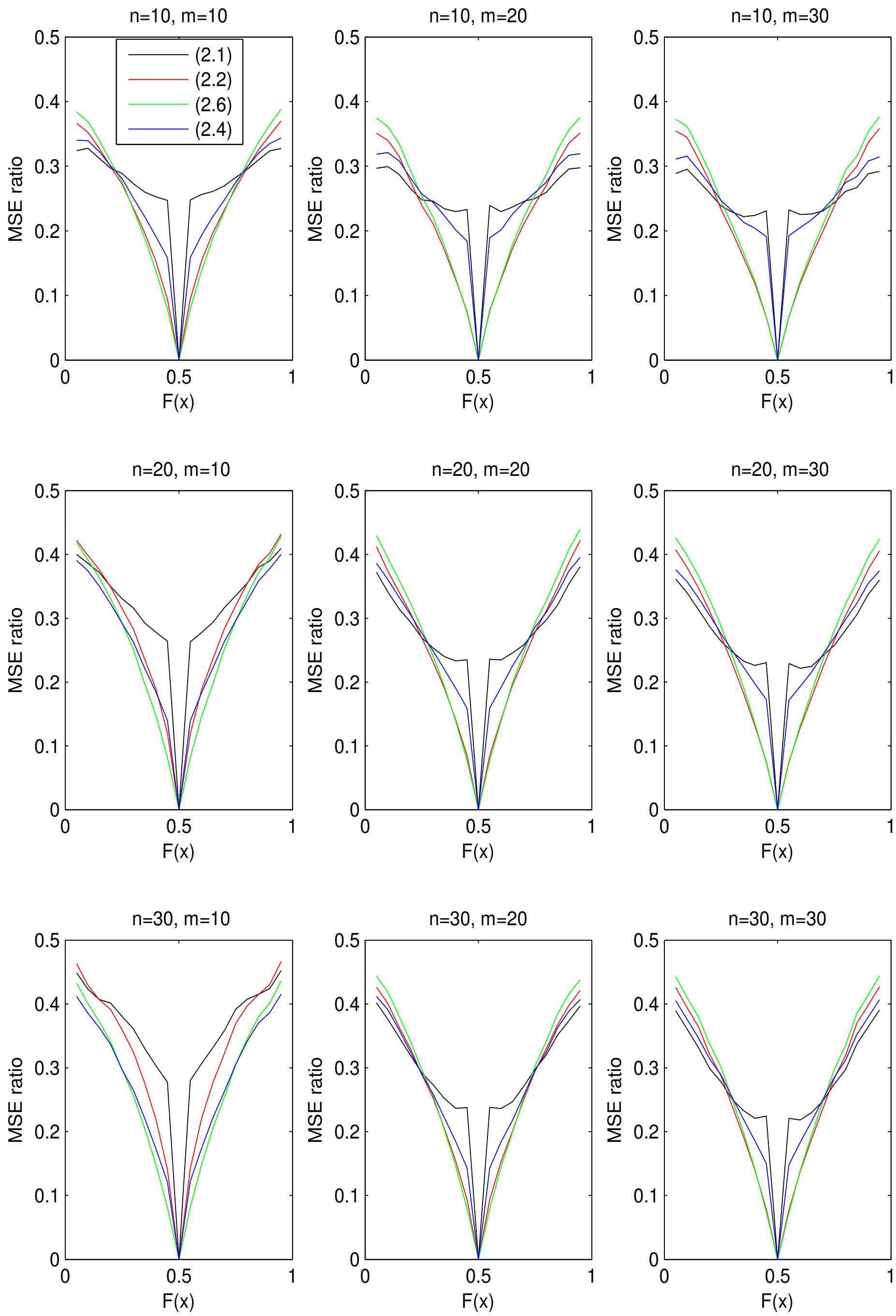

FIG 6. *Mean Squared Error of the estimators when estimating F∼Normal(0,1) with G∼Normal(0,2).*

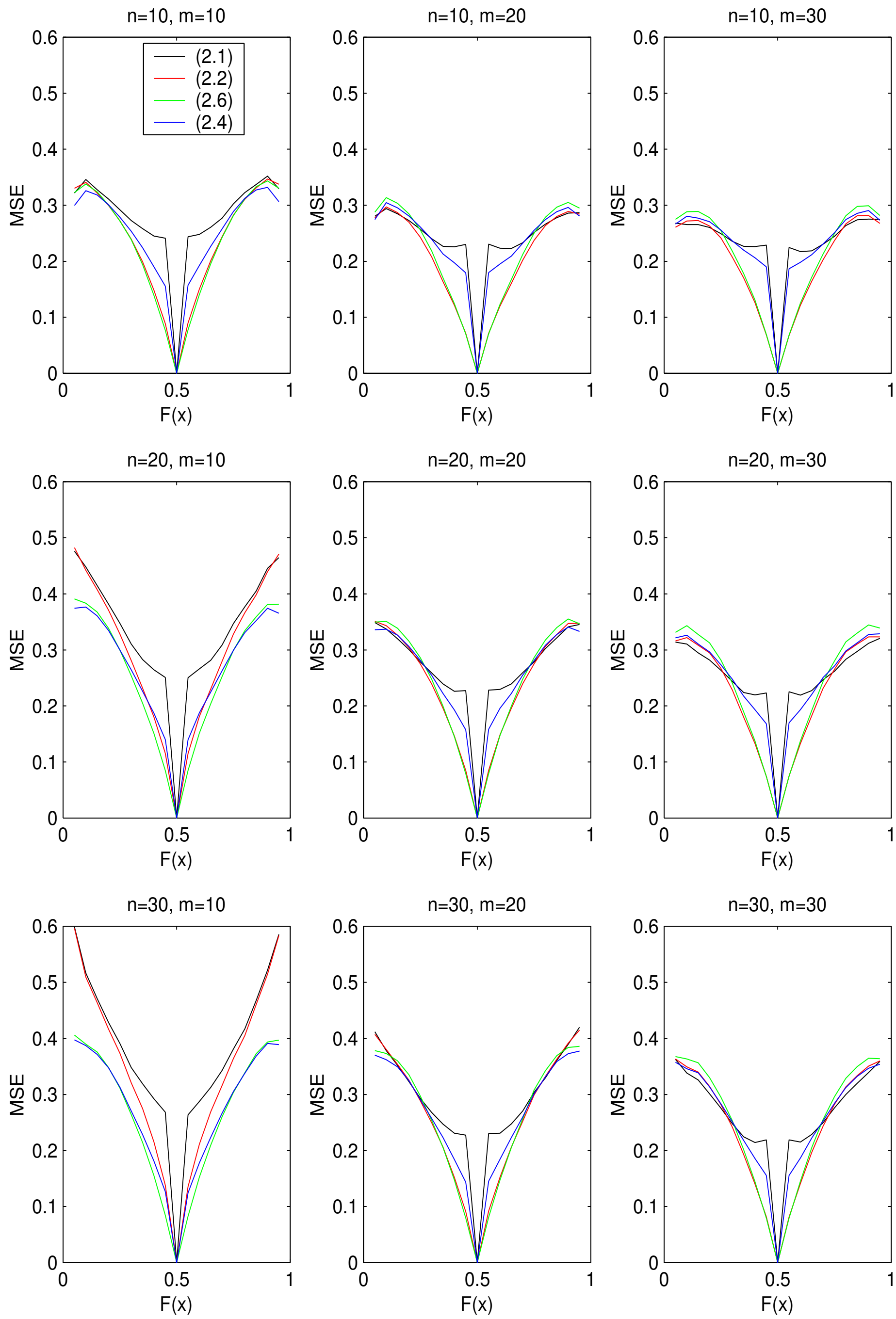

FIG 7. *Mean Squared Error of the estimators when estimating F∼Cauchy(0,1) with G∼Cauchy(0,1.5).*

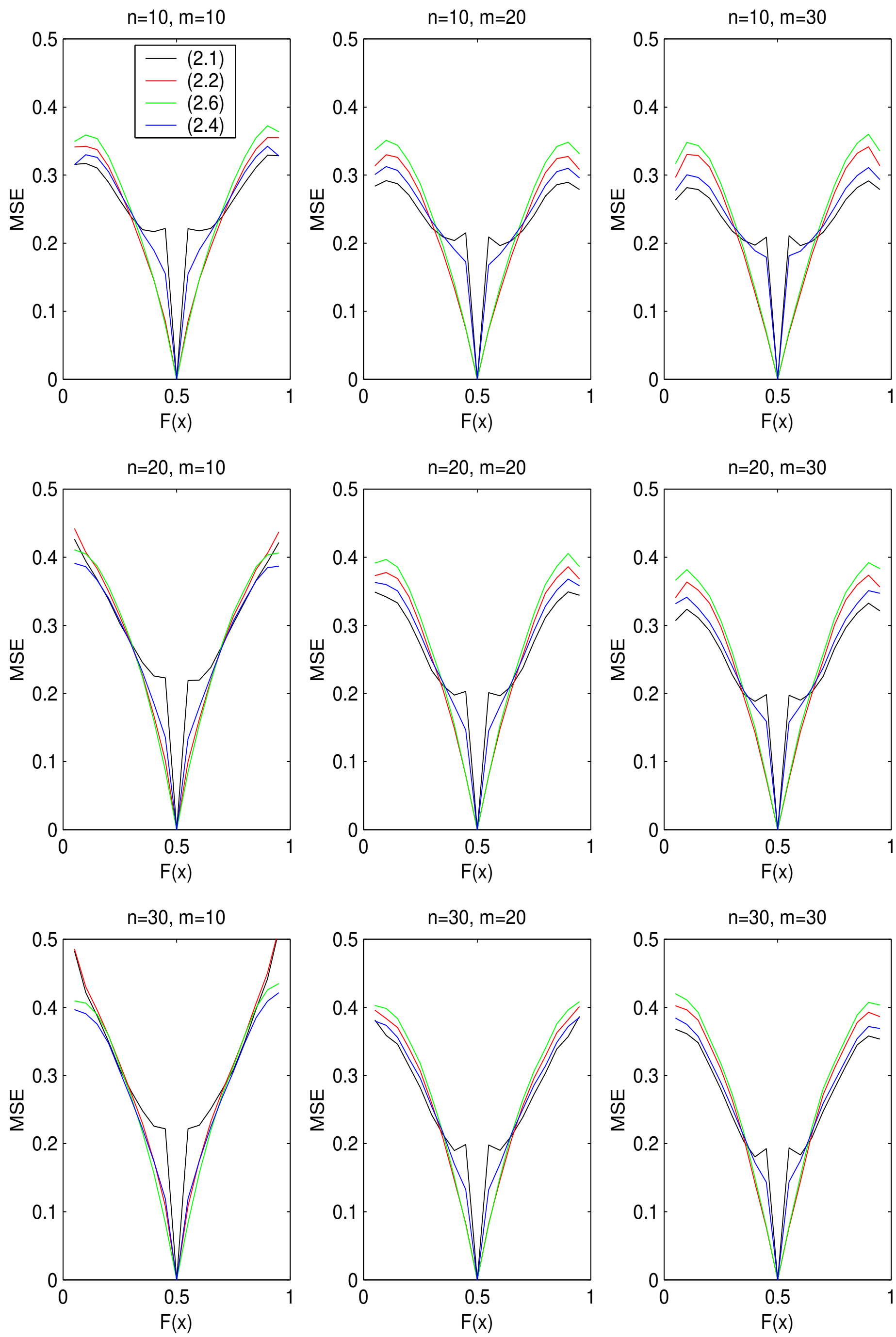

FIG 8. *Mean Squared Error of the estimators when estimating F∼Cauchy(0,1) with G∼Cauchy(0,2).*

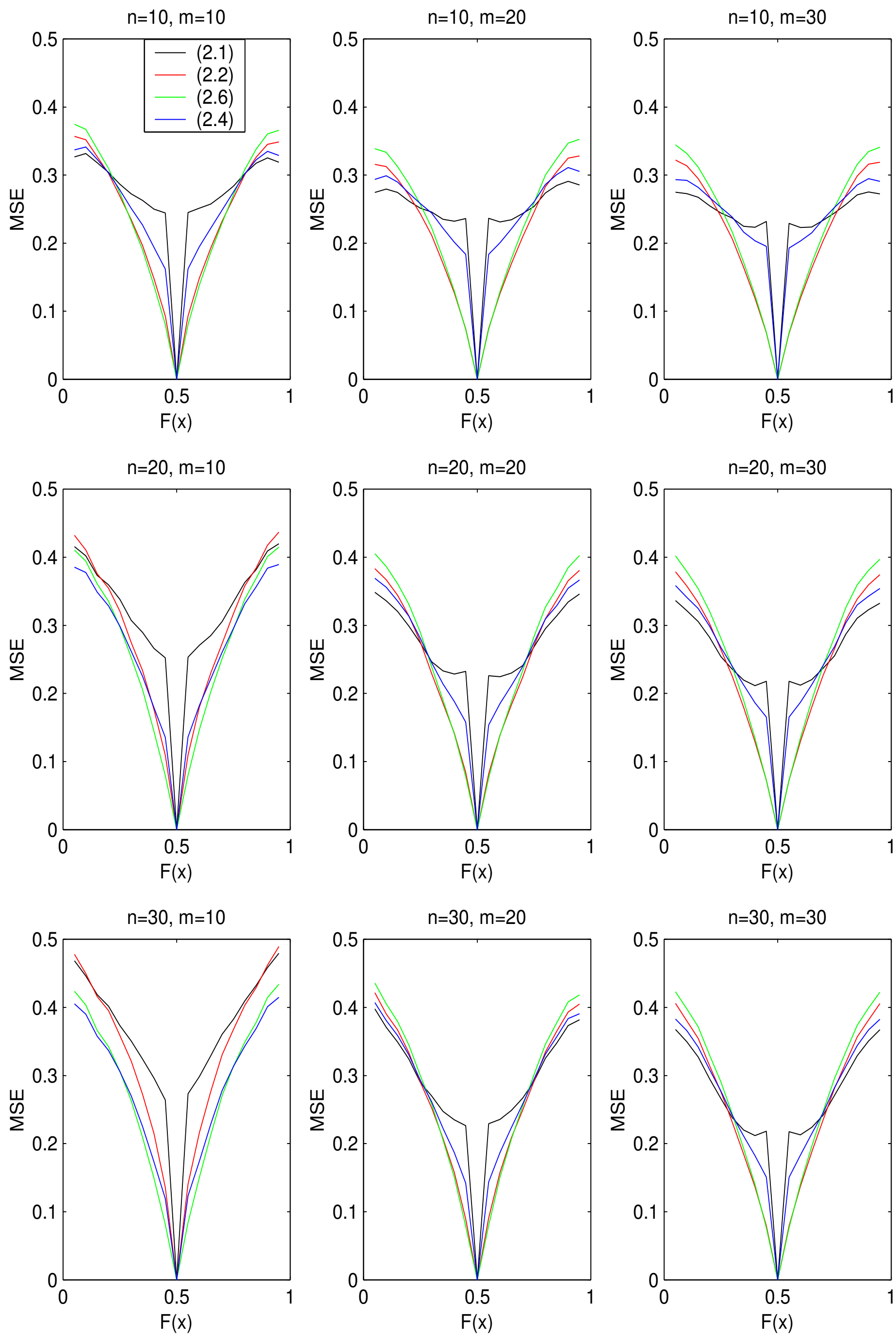

FIG 9. *Mean Squared Error of the estimators when estimating F∼Laplace(0,1) with G∼Laplace(0,1.5).*

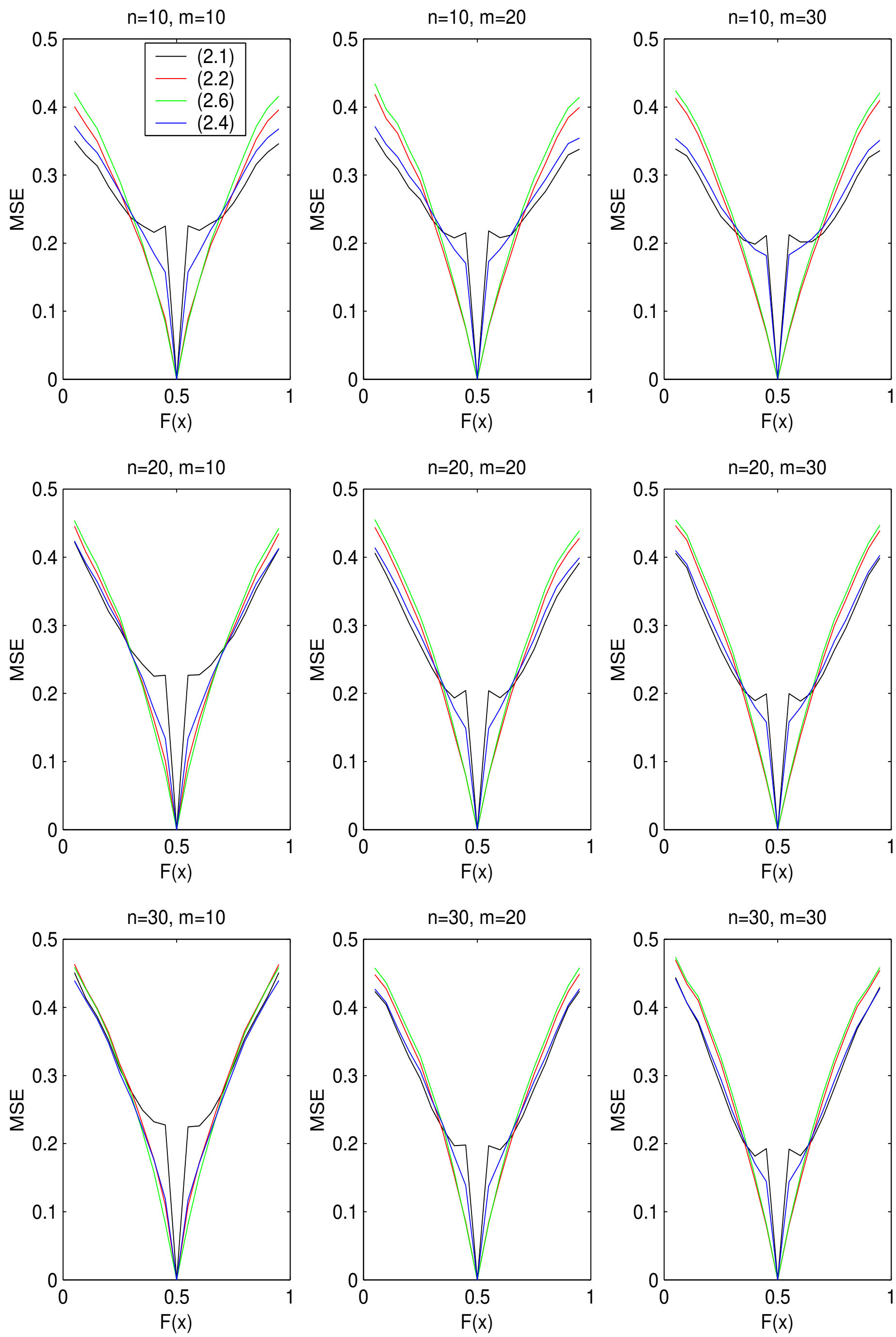

FIG 10. *Mean Squared Error of the estimators when estimating F∼Laplace(0,1) with G∼Laplace(0,2).*

Acknowledgements

The work of the first author was partially supported by NSF Grant DMS-053246, NSA Grant H98230-06-1-0099, and NSF REU Grant DMS-0552590. The second author acknowledges support by project PROMEP/103.5/08/2988. The authors are grateful to Drs Rudy Guerra of Rice University and Jonathan Cohen of University of Texas Southwestern Medical Center who made their data accessible. Their work was supported by NIH grant RO1 HL-53917.

References

[1] Arenaz, P., Bitticks L., Panell, K. H. and García, S. (1992). Genotoxic potential of crown ethers in mammalian cells; induction of sister chromatid exchanges. *Mutation Research*, **280**, 109–115.

[2] Bartoszewicz, J. (1985a). Moment inequalities for order statistics from ordered families of distributions. *Metrika*, **32**, 383–389.

[3] Bartoszewicz, J. (1985b). Dispersive ordering and monotone failure rate distributions. *Adv. Appl. Prob.*, **17**, 472–47.

[4] Bartoszewicz, J. (1986). Dispersive ordering and the total time on test transformation. *Statist. Prob. Lett.*, **4**, 285–288.

[5] Bickel, P. J. and Lehmann, E. L. (1979). Descriptive statistics for nonparametric models. IV. Spread. In *Contributions to Statistics, Jaroslav Hájek Memorial Volume*, ed. J. Jureckova. Dordrecht, Riedel, 33–40.

[6] Birnbaum, Z. W. (1948). On random variables with comparable peakedness. *Ann. Math. Statist.*, **19**, 6–81.

[7] Brown, G. and Tukey, J. W. (1946). Some distributions of sample means. *Ann. Math. Statist.*, **7**, 1–12.

[8] Doksum, K. A. (1969). Starshaped transformations and the power of rank tests. *The Annals of Mathematical Statistics*, **40**, 1167–1176.

[9] El Barmi, H. and Rojo, J. (1997). Likelihood ratio test for peakedness in multinomial populations. *J. Nonparam. Statist.*, **7**, 221–237.

[10] El Barmi, H. and Mukerjee, H. (2008). Peakedness and peakedness ordering in symmetric distributions. *Journal of Multivariate Analysis*, doi:10.1016/j.jmva2008.06.011 (In Press)

[11] Elston, R. C., Boxbaum, S. and Olson, M. (2000). Haseman and Elston Revisited. *Genetic Epidem.*, **19**, 1–17.

[12] Embury, S. H., Elias, L., Heller, P. H., Hood, C. E., Greenberg, P. L. and Schrier, S. L. (1977). Remission maintenance therapy in acute myelogenous lukemia. *Western Journal of Medicine*, **126**, 267–272.

[13] Fraser, D. A. S. (1957). *Nonparametric Methods in Statistics.* Wiley, New York.

[14] Haseman, J. K. and Elston, R. C. (1972). The investigation of linkage between a quantitative trait and a marker locus. *Behav. Genet.*, **2**, 2–19.

[15] Karlin, S. (1968). *Total Positivity.* Stanford University Press, CA.

[16] Lehmann, E. L. (1955). Ordered families of distributions. *Ann Statist.*, **26**, 399–419.

[17] Lehmann, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

[18] Lehmann, E. L. (1988). Comparing location experiments. *Ann. Statist.*, **16**, 521–533.

[19] Lehmann, E. L. and Rojo, J. (1992). Invariant directional orderings. *Ann Statist.*, **20**, 2100–2110.

[20] Liu, B. H. (1988). *Statistical Genomics, Linkage, Mappings, and QTL Analyisis.* CRC Press, New York.

[21] Lo, S. H. (1987). Estimation of distribution functions under order restrictions. *Statistics and Decisions*, **5**, 251–262.

[22] Marshall, A. W. and Olkin, I. (2007). *Life Distributions: Structure of Nonparametric, Semiparametric, and Parametric Families.* Springer Science+Bussines Media, LCC, New York.

[23] Oh, M. (2004). Inference for peakedness ordering between two distributions. *J. Kor. Statist. Soc.*, **33**, 303–312.

[24] Oja, H. (1981). On location, scale, skewness, and kurtosis of univariate distributions. *Scand. J. Statist.*, **8**, 154–168.

[25] Proschan, F. (1965). Peakedness of distributions of convex combinations. *Ann. Math. Statist.*, **36**, 1703–1706.

[26] Rojo, J. and He, G. Z. (1991). New properties and characterizations of the dispersive ordering. *Statist. & Prob. Lett.*, **11**, 365–372.

[27] Rojo, J. and Wang, J. (1994). Test based on L-statistics to test the equality in dispersion of two probability distributions. *Statistics and Probability Letters.* **21**, 107–113.

[28] Rojo, J. (1995a). On the weak convergence of certain estimators of stochastically ordered survival functions. *J. Nonparam. Statist.*, **4**, 349–363.

[29] Rojo, J. (1995b). Nonparametric quantile estimation under order constraints, *J. Nonparam. Statist.*, **5**, 185-200.

[30] Rojo, J. and Ma, Z. (1996). On the estimation of stochastically ordered survival functions. *J. Statist. Compu. and Simu.*, **55**, 1–21.

[31] Rojo, J. (1998). Estimation of the quantile function of an IFRA distribution. *Scand. J. Statist.*, **25.2**, 293–310.

[32] Rojo, J. (1999). On the estimation of a survival function under a dispersive order constraint. *J. Nonparam. Statist.*, **11**, 107–135.

[33] Rojo, J. (2004) On the estimation of survival functions under a stochastic order constraint. *The First Erich L. Lehmann Symposium - Optimality*, (J. Rojo, *ed*), IMS LNMS Vol 44, 37–61.

[34] Rojo, J. and Batun-Cutz, J. (2007). Estimation of symmetric distributions subjects to peakedness order. *Series in Biostatistics Vol 3, Advances in Statistical Modeling and Inference* Chapter 13, 649–670.

[35] Rojo, J., Batun-Cutz, J. and Durazo, R. (2007). Inference under peakedness restrictions. *Statistica Sinica* **17.3**, 1165–1189.

[36] Schuster, E. (1975). Estimating the distribution function of a symmetric distribution. *Biometrika,* 62, 3, 631–635.

[37] Schweder, T. (1982). On the dispersion of mixtures. *Scand. J. Statist.*, **9**, 165–169.

[38] Shaked, M. (1980). On mixtures from exponential families. *J. R. Statist. Soc. B.*, **42**, 192–198.

[39] Shaked, M. (1982). Dispersive orderings of distributions. *J. Appl. Prob.*, **19**, 310–320.

[40] Shaked, M. and Shantikumar, J. G. (2007) *Stochastic Orders*. Springer Science+Bussines Media, LCC, New York.

[41] Shibata, T. and Takeyama, T. (1977). Stochastic theory of pitting corrosion. *Corrosion*, **33.7**, 243.