

ON THE DERIVATIVES OF TRIMMED MEAN

Subhra Sankar Dhar and Probal Chaudhuri

Theoretical Statistics and Mathematics Unit, Indian Statistical Institute, Calcutta

Abstract: Trimmed mean, introduced by Tukey (1948), has been extensively studied in the literature as an estimate of the location parameter. It is well-known for being more robust and for having better mean square error than the usual mean when data arise from non-Gaussian distributions with heavy tails. In this paper, we consider the derivatives of trimmed mean with respect to the trimming proportion and investigate some statistical applications of those derivatives. We develop a diagnostic tool based on the first derivative of trimmed mean to determine whether the data is generated from a symmetric distribution or not. We also propose a test of symmetry of the distribution based on the first derivative and demonstrate that it performs better than several other well-known tests of symmetry studied by earlier authors. Further, we introduce an estimate of the contamination proportion $\beta \in (0, 1/2)$ in the contamination model $F(x) = (1 - \beta)H(x) + \beta G(x)$, where H and G are two distributions such that G is stochastically larger than H , based on the second derivative of the trimmed mean. In addition to some theoretical studies, we carry out a detailed numerical study to show that, in many situations, our proposed estimate of the contamination proportion outperforms other estimates, which are based on the idea of maximum likelihood estimation in mixture models.

Key words and phrases: Contamination model, Karhunen-Loeve expansion, maximum likelihood estimation, proportion of contamination, symmetric and skewed distributions, weak convergence of processes.

1 Introduction

It is well-known that while sample mean has smaller mean square error than sample median as an estimate of the location parameter for normally distributed data, the same is not true when data are generated from other distributions with heavier tails. Further, sample mean has the lowest breakdown point (0%), and sample median has the highest breakdown point (50%). For a random sample x_1, x_2, \dots, x_n , Tukey (1948) introduced the sample α -trimmed

mean \bar{x}_α , which is defined as

$$\bar{x}_\alpha = \frac{1}{n - 2[n\alpha]} \sum_{i=[n\alpha]+1}^{n-[n\alpha]} x_{(i)},$$

where $\alpha \in (0, 1/2)$, and $x_{(i)}$ is the i -th order statistic of the sample. Clearly, for specific choices of α , the α -trimmed mean will coincide with the sample mean and the sample median. It is well-known that the asymptotic breakdown point of the α -trimmed mean is α (see, e.g., Huber (1981)). Tukey and Mclaughlin (1963) proposed a trimmed version of the t -statistic using the α -trimmed mean and showed that the resulting test performs better than the usual t -test in different examples considered by them. In a review paper on adaptive robust procedures, Hogg (1967) discussed some practical reasons for using the α -trimmed mean. Bickel (1965) and Stigler (1973) investigated the asymptotic distribution of the α -trimmed mean under appropriate regularity conditions. Jaeckel (1971) proposed an estimate $\hat{\sigma}_\alpha^2$ for the asymptotic variance of the α -trimmed mean and used the value of α that minimizes $\hat{\sigma}_\alpha^2$ to construct an adaptive version of trimmed mean, which was subsequently studied by Hall (1981). Recently, Dhar and Chaudhuri (2009) has shown that, for a large class of symmetric distributions with exponential and polynomial tails, trimmed mean is more efficient than the least trimmed squares estimate, which is another robust estimate of location based on an alternative trimming procedure.

If we assume that the random sample consists of i.i.d. observations from an absolutely continuous distribution F having density f , the population analogue of the α -trimmed mean is defined as

$$\theta(\alpha) = \frac{1}{(1 - 2\alpha)} \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} xf(x)dx$$

(see, e.g., Serfling(1980, p.236)). Note that $\theta(0) = E_F(x)$ if it is finite, and as $\alpha \rightarrow (1/2+)$, $\theta(\alpha) \rightarrow F^{-1}(1/2)$ under appropriate conditions. If F happens to be a symmetric distribution with the center of symmetry μ , it is obvious that $\theta(\alpha) = \mu$ for all α . When $\theta(\alpha)$ is a continuously differentiable function of $\alpha \in (0, 1/2)$ (which is true, for instance, when f is a continuous probability density function), this implies that

$$\begin{aligned} \frac{d}{d\alpha}\theta(\alpha) = \theta'(\alpha) &= \frac{2}{(1 - 2\alpha)^2} \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} xf(x)dx - \frac{1}{(1 - 2\alpha)}\{F^{-1}(\alpha) + F^{-1}(1 - \alpha)\} = 0 \\ \Leftrightarrow \theta'(\alpha) &= \frac{1}{(1 - 2\alpha)} \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} xf(x)dx - \frac{\{F^{-1}(\alpha) + F^{-1}(1 - \alpha)\}}{2} = 0 \end{aligned}$$

for all $\alpha \in (0, 1/2)$. Therefore, one can develop diagnostic tools and tests for the presence or the absence of symmetry in the distribution generating the data based on estimates of $\theta'(\alpha)$. Such estimates are expected to be close to zero when the assumption of symmetry holds.

Several tests of symmetry have been proposed and studied in the literature. Mira (1999) proposed and investigated a test for $(Mean_F - Median_F) = 0$ based on sample mean and sample median. Though Mira's test is easy to implement, it has poor power for asymmetric distributions with mean and median close to each other. Besides, derivation of the asymptotic distribution of Mira's test statistic requires finiteness of the second moment of the underlying distribution. An extensive study comparing Mira's test with some other tests of symmetry proposed by earlier authors (e.g., Tajaiddin (1994), Modarres and Gastwirth (1996), Gupta (1967), Cabilio and Masaro (1996)) can be found in Mira (1999). It has been demonstrated there that Mira's test can outperform many of those earlier tests of symmetry in a wide variety of situations. In an earlier paper, Gastwirth (1971) also proposed a version of sign test, when the center of symmetry for the distribution is unspecified and used sample mean to estimate the center of symmetry. This test also tries to determine whether the locations of the mean and the median are close to each other. Csorgo and Heathcote (1987) proposed another test based on the estimate of the characteristic function of F . On the other hand, Ahmad and Li (1997) proposed a test that tries to compare kernel based estimates of $f(x)$ and $f(-x)$. Since nonparametric density estimates converge at rates slower than the usual $n^{-1/2}$ rate, the test statistics used by them have large variances leading to poor powers of their proposed tests. Schuster and Barker (1987) developed a test based on the Kolmogorov-Smirnov distance between the empirical distribution function F_n and its symmetrized version, and implemented the test using bootstrap techniques. This test is subsequently studied by Arcones and Gine (1991). In the next section, we have developed a graphical tool based on the derivative of α -trimmed mean to determine whether the data is generated from a symmetric distribution or not. We have also proposed a new test of symmetry and compared the performance of the test with that of some other well-known tests available in the literature.

The estimation of the contamination parameter β in the contamination model $F(x) = (1 - \beta)H(x) + \beta G(x)$, where $\beta \in (0, 1/2)$, and H and G are two distributions with G is stochastically larger than H , is important for assessing the extent of contamination in the data. Huber (1981) studied the performance of α -trimmed mean for data following

β -contaminated asymmetric normal distribution (see Huber (1981, p.2 and p.104)). He showed that the maximal asymptotic bias and variance of α -trimmed mean will be finite when $\alpha \geq \beta$, and it will be infinite when $\alpha < \beta$ (see Huber (1981, p. 104–105, Tables 4.9.1 and 4.9.2)). In Section 3, we have proposed an estimate of the contamination proportion β based on the second derivative of α -trimmed mean for data following the above contamination model. Since any such contamination model can be viewed as a special case of mixture models, one can also try to estimate the contamination proportion using maximum likelihood and related techniques (see, e.g., Everitt and Hand (1981)). However, such estimates of the mixing proportion will be based on specific choices of the parametric models for H and G , while our approach based on the second derivative of the α -trimmed mean is nonparametric in nature and not constrained by any parametric model for H and G . Besides, the implementation of maximum likelihood techniques in parametric mixture models involves several approximations and iterative computations. As we will see in Section 3, our estimate outperforms maximum likelihood type estimates for the contamination proportion in several contamination models.

2 Detection of asymmetry based on α -trimmed mean

A popular way of visually investigating the presence or the absence of symmetry in data is based on the histogram of the data. Alternatively, one may use some other density estimates too. However, the use of histogram or any other density estimate is justified only when one has a reasonably large sample. It will not be possible to form the class intervals and the frequency distribution in a sensible way in a small sample, which are basic ingredients needed to construct a histogram. Density estimates are often statistically very unreliable in small samples due to their high variability and slow convergence rates. In addition, histogram or any other density estimate involves the choice of a smoothing parameter.

As we have already mentioned in the preceding section, an assessment of symmetry or asymmetry in the data can be done using the derivative of the α -trimmed mean, i.e, $\theta'(\alpha)$. Note that a natural estimate of $\theta'(\alpha)$ is $T_n(\alpha) := \bar{x}_\alpha - \frac{\{F_n^{-1}(\alpha) + F_n^{-1}(1-\alpha)\}}{2}$, where F_n is the empirical distribution function of F . It is clear from this expression that $T_n(\alpha)$ is an L-estimate for $\theta'(\alpha)$, which is \sqrt{n} -consistent and has asymptotically normal distribution (see,

e.g., Serfling (1980)). In fact, we have the following theorem that describes the asymptotic behavior of the process $T_n(\alpha)$ as α varies over the interval $(0, 1/2)$.

Theorem 2.1: *Suppose that the x_i 's are i.i.d. observations with a common continuous and positive density on the entire real line. Then, for any $0 < b_1 < b_2 < 1/2$, the process $\sqrt{n}\{T_n(\alpha) - \theta'(\alpha)\}$, where α varies over the interval $[b_1, b_2]$, converges weakly to a Gaussian process with zero mean and covariance function $\lim_{n \rightarrow \infty} \sqrt{n}E\{T_n(\alpha_1) - \theta'(\alpha_1)\}\{T_n(\alpha_2) - \theta'(\alpha_2)\} = k(\alpha_1, \alpha_2)$, where the form of $k(\alpha_1, \alpha_2)$ is given in Lemma 4.1 in the Appendix.*

The following corollary follows from Theorem 2.1.

Corollary 2.2: *Under the conditions of Theorem 2.1, for any $0 < b_1 < b_2 < 1/2$, we have $\sup_{\alpha \in [b_1, b_2]} |T_n(\alpha) - \theta'(\alpha)| = O_p(n^{-1/2})$.*

Corollary 2.2 asserts that the function $T_n(\alpha)$ converges uniformly in probability to the function $\theta'(\alpha)$ at $n^{-1/2}$ rate on any compact subinterval of $(0, 1/2)$. So, if one plots the graph of $T_n(\alpha)$ to determine the presence or the absence of symmetry in the data, as the sample size grows, the graph will converge to the graph of $\theta'(\alpha)$ at a rapid rate ensuring correct determination with high probability. If the graph of $T_n(\alpha)$ is roughly constant around the value zero, one may conclude that the data is obtained from a symmetric distribution. Otherwise, one would conclude asymmetry in the distribution. Also, using the asymptotic normality of $\sqrt{n}\{T_n(\alpha) - \theta'(\alpha)\}$, for each fixed $\alpha \in (0, 1/2)$, one can calculate the asymptotic p -value for the testing problem with the null hypothesis $H_{0,\alpha} : \theta'(\alpha) = 0$ against the alternative hypothesis $H_{1,\alpha} : \theta'(\alpha) \neq 0$ for each $0 < \alpha < 1/2$. This p -value measures the extent of the deviation of $T_n(\alpha)$ from its null value (i.e., zero) for each α . In Figure 2.1, for different values of the trimming proportion α , we have plotted (the black curve) the average of $T_n(\alpha)$ and that of the corresponding p -values over 50 Monte-carlo replications, where each replication consists of 25 i.i.d. observations from standard normal distribution. The red and the blue curves plotted in the same plot with $T_n(\alpha)$ correspond to average \pm (std. dev.) and average ± 2 (std. dev.) limits, respectively. Here the std. dev. is computed from the values of $T_n(\alpha)$ obtained in different Monte-carlo replications.

Next, we investigate the behavior of $T_n(\alpha)$ when data are generated from some asymmetric distributions. Here, we have considered two different skewed distributions, namely, the gamma distribution with shape parameter 0.15 and scale parameter 1 and the mixture normal distribution $0.7N(0, 1) + 0.3N(5, 1)$. In Figures 2.2 and 2.3, we have displayed the graphs of average values of $T_n(\alpha)$ along with the average \pm (std. dev.) and average ± 2 (std. dev.) limits and the averages of corresponding p -values obtained from 50 Monte-

carlo replications, where each replication consists of a set of 25 i.i.d. observations as in Figure 2.1. Figure 2.2 corresponds to gamma distribution with shape parameter = 0.15 and scale parameter = 1 while Figure 2.3 corresponds to the mixture normal distribution $0.7N(0, 1) + 0.3N(5, 1)$.

Figure 2.1: The graphs of the averages of $T_n(\alpha)$ and corresponding p -values obtained from 50 Monte-carlo replications of 25 i.i.d. standard normal observations. The red and the blue curves represent average \pm (std. dev.) and average ± 2 (std. dev.) limits, respectively.

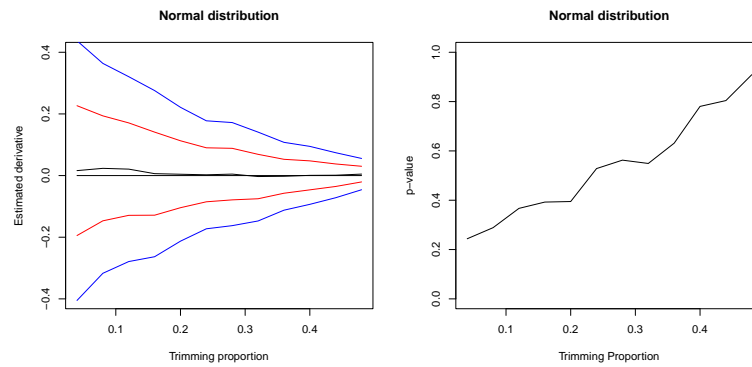
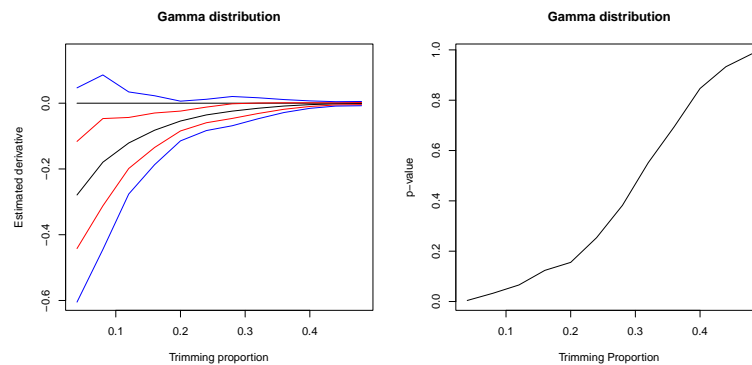
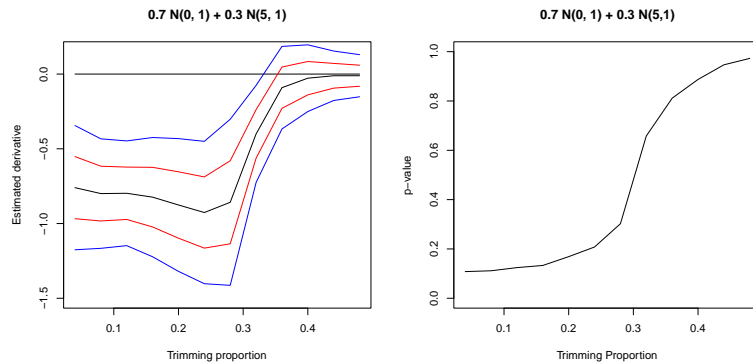


Figure 2.2: The graphs of the averages of $T_n(\alpha)$ and corresponding p -values obtained from 50 Monte-carlo replications of 25 i.i.d. observations from gamma distribution with shape parameter = 0.15 and scale parameter = 1. The red and the blue curves represent average \pm (std. dev.) and average ± 2 (std. dev.) limits, respectively.



The symmetry of the generated data is quite visible in Figure 2.1 while the asymmetry in the data is clearly indicated in each of Figures 2.2 and 2.3. In the next subsection, the detection of the nature of asymmetry in the data will be discussed.

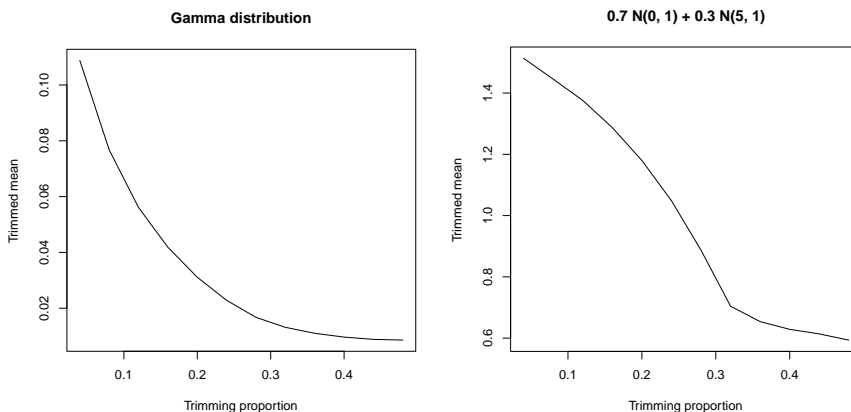
Figure 2.3: The graphs of the averages of $T_n(\alpha)$ and corresponding p -values obtained from 50 Monte-carlo replications of 25 i.i.d. observations from $0.7N(0, 1) + 0.3N(5, 1)$ distribution. The red and the blue curves represent average \pm (std. dev.) and average ± 2 (std. dev.) limits, respectively.



2.1 Positive and negative skewness and their detection

Note that in both of Figures 2.2 and 2.3, the graphs of the averages of the derivative of α -trimmed mean, i.e., $T_n(\alpha)$ lie below the x-axis for almost all α . This fact indicates a monotonically decreasing nature of the α -trimmed mean, when data are generated from the gamma and the mixture normal distributions (see Figure 2.4).

Figure 2.4: The graphs of the averages of \bar{x}_α obtained from 50 Monte-carlo replications of 25 i.i.d. observations from gamma and $0.7N(0, 1) + 0.3N(5, 1)$ distributions.



We will now state a result that establishes monotonically decreasing nature of α -trimmed means (i.e., $\theta'(\alpha) \leq 0$ for all $\alpha \in (0, 1/2)$) for distributions that are asymmetric in some appropriate sense.

Proposition 2.3: *Suppose that F is a distribution function having continuous density f and satisfying the condition $f(F^{-1}(\alpha)) \geq f(F^{-1}(1-\alpha))$ for all α . Then, the α -trimmed mean $\theta(\alpha) = (1-2\alpha)^{-1} \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} xf(x)dx$ associated with F will be a decreasing function of α , and $\theta'(\alpha) \leq 0$ for all α .*

The condition $f(F^{-1}(\alpha)) \geq f(F^{-1}(1-\alpha))$ for all α mentioned in Proposition 2.3 implies that the distribution has a long tail towards its right side, and hence it can be used as a definition of positively skewed distributions (see, e.g., Van Zwet (1979)). On the other hand, the condition $f(F^{-1}(\alpha)) \leq f(F^{-1}(1-\alpha))$ for all α implies that the distribution has a long tail towards its left side, which can be taken as a definition of negatively skewed distributions, and it will imply that $\theta'(\alpha) \geq 0$ for all $\alpha \in (0, 1/2)$ for such distributions.

Recall that Corollary 2.2 asserts that as the sample size grows, the graph of $T_n(\alpha)$ will converge to the graph of $\theta'(\alpha)$ at a rapid rate. If the graph of $T_n(\alpha)$ lies below the x-axis, one may conclude that the data are obtained from a positively skewed distribution. Similarly, if the graph of $T_n(\alpha)$ lies above the x-axis, it implies that the data are generated from a negatively skewed distribution. Motivated by these results, we propose a test of symmetry based on the derivative of α -trimmed mean in the next subsection.

2.2 A Statistical test for symmetry

We have already seen that $\theta'(\alpha) = 0$ for all $\alpha \in (0, 1/2)$ for a symmetric distribution, and hence one can formulate the problem of testing the hypothesis of symmetry as follows.

$H_0 : \int_{b_1}^{b_2} \{\theta'(\alpha)\}^2 d\alpha = 0$ for all $0 < b_1 < b_2 < 1/2$ against $H_1 : \int_{b_1}^{b_2} \{\theta'(\alpha)\}^2 d\alpha \neq 0$ for some $0 < b_1 < b_2 < 1/2$.

Note that a natural estimate of $\int_{b_1}^{b_2} \{\theta'(\alpha)\}^2 d\alpha$ is $V_n = (1/n) \sum_{i=[nb_1]+1}^{[nb_2]-1} \{T_n(i/n)\}^2$, and we propose to use it as a test statistic for testing H_0 against H_1 .

For $\alpha \in [b_1, b_2]$, the asymptotic distribution of the stochastic process $\sqrt{n}\{T_n(\alpha) - \theta'(\alpha)\}$ is Gaussian, which is stated in Theorem 2.1. Under H_0 , the limiting distribution of $V_n = \sum_{i=[nb_1]+1}^{[nb_2]-1} \{T_n(i/n)\}^2$ can also be derived, and it is stated in Theorem 2.4.

Theorem 2.4: *Under H_0 and the assumptions in Theorem 2.1, $V_n = \sum_{i=[nb_1]+1}^{[nb_2]-1} \{T_n(i/n)\}^2$ converges weakly to $\sum_{i=1}^{\infty} Z_i^2$, where the Z_i 's are independent $N(0, \lambda_i)$ random variables. Here the λ_i 's are the eigen values of the covariance function $k(\alpha_1, \alpha_2)$ ($\alpha_1, \alpha_2 \in [b_1, b_2]$), which is defined in Theorem 2.1. Further, a test with asymptotic size $0 < \rho < 1$, which rejects the null hypothesis $\int_{b_1}^{b_2} \{\theta'(\alpha)\}^2 d\alpha = 0$ when $V_n > \xi_\rho$, where ξ_ρ is the $(1-\rho)$ -th quantile of*

the distribution of $\sum_{i=1}^{\infty} Z_i^2$, will be a consistent test in the sense that under the alternative hypothesis, i.e., when data are generated from an asymmetric distribution, the asymptotic power of the test will be 1.

One can carry out the test of H_0 against H_1 using the asymptotic distribution of $V_n = \sum_{i=[nb_1]+1}^{[nb_2]-1} \{T_n(i/n)\}^2$ mentioned in Theorem 2.4. In order to implement the test, we first estimate the eigen values λ_i 's of the covariance kernel $k(\alpha_1, \alpha_2)$ ($\alpha_1, \alpha_2 \in [b_1, b_2]$), an expression for which is given in the statement of Lemma 4.1 in the Appendix. Note that the expression for $k(\alpha_1, \alpha_2)$ involves quantiles and the density function of the underlying distribution. One can estimate quantiles by usual sample quantiles and use some standard estimate of the density. We have used the kernel density estimate based on standard Gaussian kernel and used an adaptive choice of the bandwidth (namely, the default bandwidth in "kde" package in Statistical software *R*) in our numerical studies. Alternatively, one can estimate $k(\alpha_1, \alpha_2)$ by resampling methods such as the bootstrap. However, that will be computationally more expensive, and in our empirical study, we have not seen any difference in the performance of the test in terms of its level and power whether the covariance function is estimated by the first approach or using the bootstrap. In practice, one would estimate $k(\alpha_1, \alpha_2)$ for finitely many (r , say) values of α_1 and α_2 in the interval $[b_1, b_2]$, and form the $r \times r$ variance-covariance matrix having (i, j) -th element $k([i]/2r, [j]/2r)$; the eigen values of which will be computed subsequently. Note that in any finite sample situation, it will not be meaningful or necessary to consider a value of r larger than half of the sample size, i.e., $n/2$. Using estimated eigen values $\hat{\lambda}_i$'s, one can generate Gaussian random variables Z_i 's with zero mean and variances $\hat{\lambda}_i$'s, and then a finite approximation to the sum $\sum_{i=1}^{\infty} Z_i^2$, namely, $\sum_{i=1}^r Z_i^2$ can be computed. Finally, one can generate several Monte-carlo replications of that finite sum and, depending on the specified level of the test, choose the cut off point of the test as a specific quantile of the empirical distribution of that sum. In our implementation, we have considered $b_1 = 0.001$, $b_2 = 0.499$, $r = n/2$ and used 1000 Monte-carlo replications to obtain the empirical distribution of the finite sum $\sum_{i=1}^r Z_i^2$. We have made those choices based on our empirical experience keeping in mind the computational cost involved and the performance of the test. In the next two subsections, we carry out a detailed numerical study, where our proposed test is implemented on simulated and real data sets, and it is compared with a few other well-known tests of symmetry available in the literature.

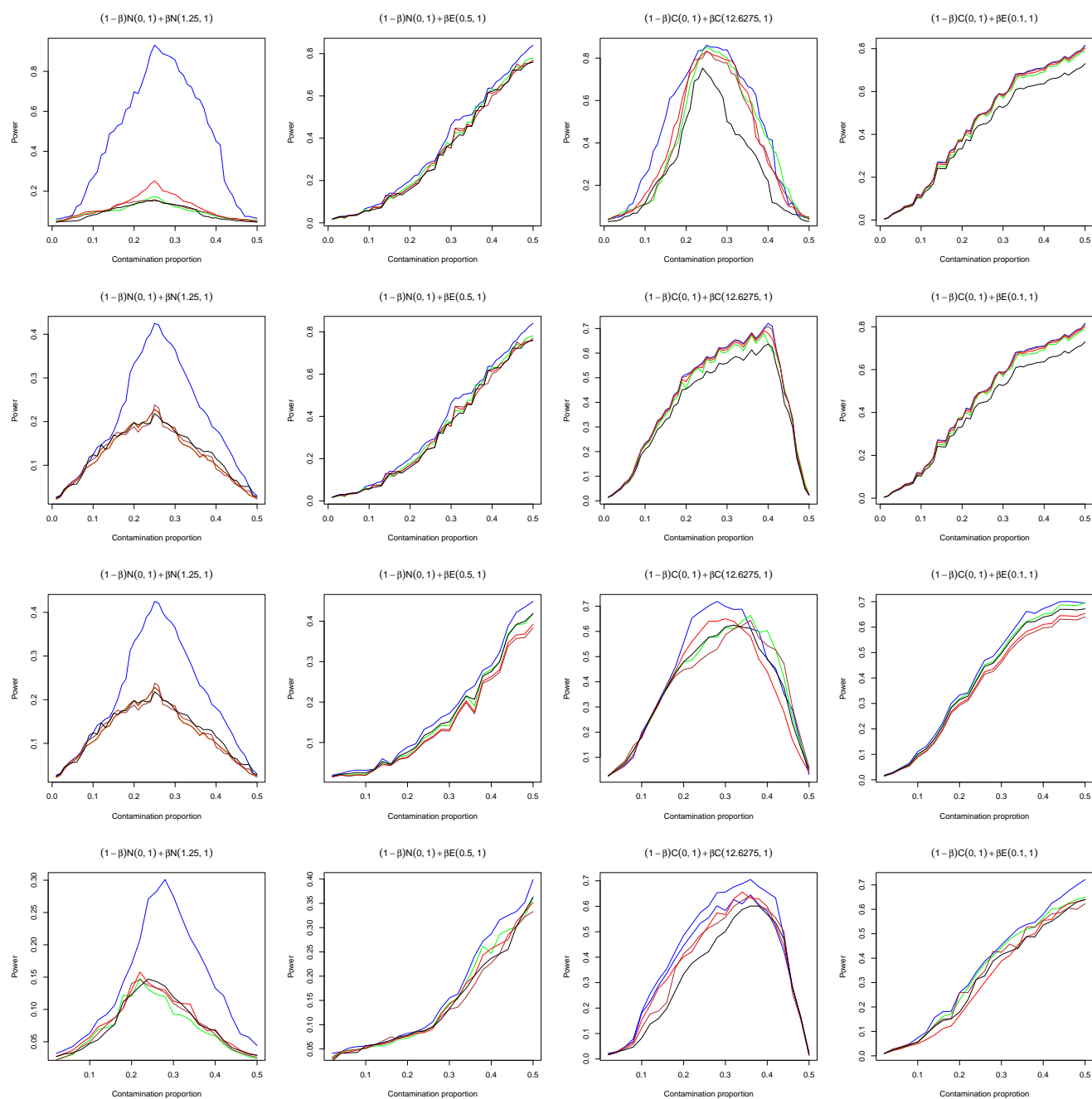
2.3 Comparison of the proposed test with other tests of symmetry

We have already presented a brief review of some well-known tests of symmetry in the Introduction. We now carry out a simulation study to compare our test proposed in this section with some of those earlier tests studied in the literature in terms of their powers and sizes. We have considered tests with nominal levels 5% and 1%, and in order to estimate the power and the size of a test, we have used the proportion of rejection of the null hypothesis of symmetry in several Monte-carlo replications, where each replication consists of i.i.d. samples from a specified distribution. In our study, we have used 1000 Monte-carlo replications with each replication consisting of a sample of size 50 or 100. For instance, in the previous examples of asymmetric distributions considered in Figures 2.2 and 2.3 (i.e., gamma distribution with shape parameter 0.15 and scale parameter 1 and the mixture normal distribution $0.7N(0, 1) + 0.3N(5, 1)$), the power of all of the tests mentioned here are found to be very close to 1 because those distributions are highly skewed. It is important to investigate the power of these tests when data are generated from less skewed distributions. Here, we have considered different mixture distributions of the form $(1 - \beta)F + \beta G$, where F and G are two probability distribution functions, and β lies in the closed interval $[0, 1/2]$ ensuring the inclusion of a wide variety of symmetric and asymmetric distributions in our study. In our investigation, we have considered F and G as normal $N(\mu, \sigma)$, Cauchy $C(\mu, \sigma)$ and exponential $E(\mu, \sigma)$ distributions with different location (μ) and scale (σ) parameters.

In the diagrams in Figure 2.5, different continuous curves having different colors denote the power curves of different tests. Here the *blue* curve is the power curve of our test (using adaptive choice of bandwidth), the *green* curve is the power curve of Mira's (1999) test, the *red* curve is the power curve of Csorgo and Heathcote's (1987) test, the *brown* curve is the power curve of J_{2n} test (see Ahmad and Li's (1997)), and the *black* curve is the power curve of Schuster and Barker's (1997) test.

It is amply indicated in the diagrams in Figure 2.5 that our test is more powerful than other tests in all the cases considered here, and all tests have approximately same observed levels (namely, 0.037 – 0.052 for nominally 5% tests and 0.009 – 0.016 for nominally 1% tests).

Figure 2.5: Empirical power of different tests at 5% and 1% nominal levels for different values of contamination proportion with sample sizes 100 and 50 when data follow different contamination models. The first row (5% level) and the second row (1% level) of curves are based on sample size 100. The third row (5% level) and the fourth row (1% level) of curves are based on sample size 50.



2.4 Analysis of real data

In this section, we have investigated the performance of our proposed test and other tests considered by Mira (1999), Csorgo and Heathcote (1987), Ahmad and Li (1997) and Schuster and Barker (1987), when these tests are applied on three real data sets, namely, the well-known Fisher's iris data, the diabetes data and the yeast data. The iris data and the yeast data are available in <http://archive.ics.uci.edu/ml>, and the diabetes data can be obtained from the "mclust" package in the software *R*. The first data set consists of three multivariate data sets corresponding to three different varieties of iris, namely, *Iris setosa*, *Iris virginica* and *Iris versicolor*, each of size 50. In each data point, there are four measurements, namely, sepal length, sepal width, petal length and petal width. The diabetes data was originally analyzed by Reven and Miller (1979). This data set contains measurements on three variables, namely, glucose level, insulin area and steady state plasma glucose (s.s.p.g.) level for 145 individuals. These individuals are classified into three classes according to some clinical criteria. These classes consist of normal individuals, chemical diabetes cases and overt diabetic individuals. There are 76, 36 and 33 individuals in these three classes, respectively. The yeast data set contains nine attributes, namely, sequence name, mcg, gvh, alm, mit, erl, pox, vac and nuc for 1484 instances.

At 5% level, all of the tests rejected the hypothesis of symmetry for petal widths of *Iris setosa* and *Iris versicolor*, for the variable s.s.p.g. in the case of normal individuals, and for all three variables in the case of overt and chemical diabetic individuals. Also, at 5% level, all those tests accepted the hypothesis of symmetry for sepal length, sepal width and petal length in the case of *Iris setosa*, for sepal width in the case of *Iris virginica*, for petal length and petal width in the case of *Iris virginica* and for sepal length and petal length in the case of *Iris versicolor*, and for glucose level and insulin area in the case of normal individuals. In the case of the yeast data, all tests lead to the same conclusion at 5% level for all the variables except gvh.

However, it is interesting to note that at 5% level, our test rejected the null hypothesis of symmetry for sepal length in the case of *Iris virginica*, for the sepal width in the case of *Iris versicolor* and for the gvh in the case of the yeast data while all of the other four tests accepted the null hypothesis of symmetry for these three variables. Following the graphical techniques and ideas introduced at the beginning of this section, we have drawn the graphs of $T_n(\alpha)$ and the corresponding p-values for $0 < \alpha < 1/2$ for the sepal length

of *Iris virginica*, the sepal width of *Iris versicolor* and the variable gvh in the yeast data in Figures 2.6 and 2.7. It is evident from Figures 2.6 and 2.7 that the distributions of these three variables are asymmetric, and consequently, our test rejects the null hypothesis for these three variables, while all other tests fail to detect the asymmetry in these three cases.

Figure 2.6: The graphs of $T_n(\alpha)$ for the sepal length of *Iris virginica*, the sepal width of *Iris versicolor* and the gvh of the yeast data.

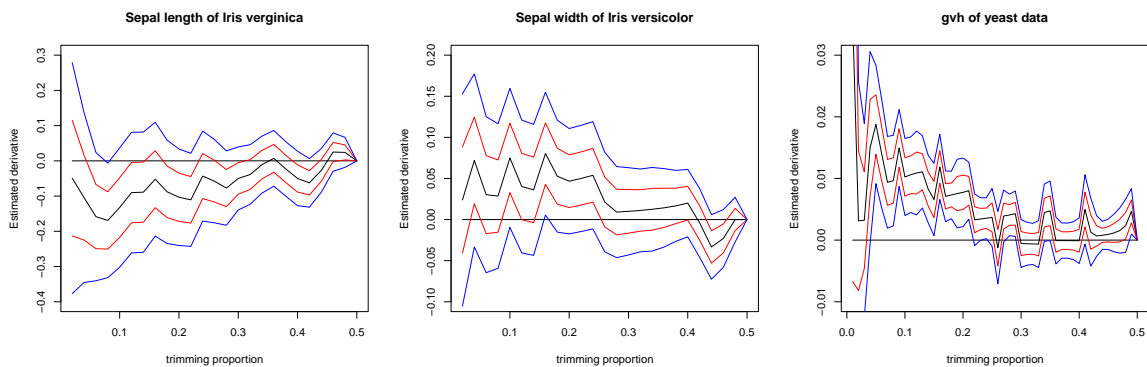
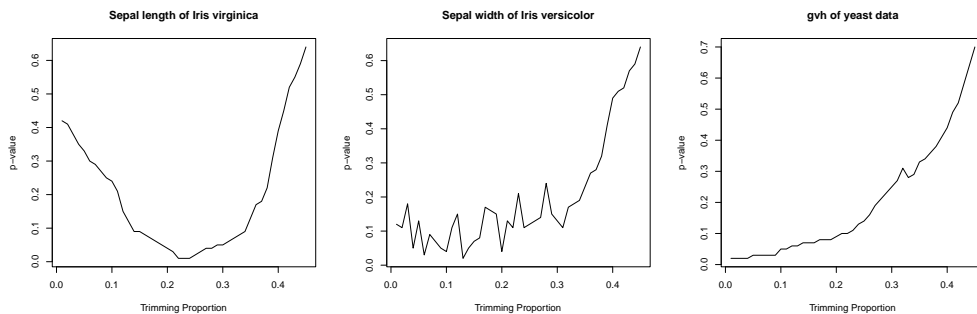


Figure 2.7: The graphs of p-values for different values of α of the sepal length of *Iris virginica*, the sepal width of *Iris versicolor* and the gvh of the yeast data.



3 The second derivative of α -trimmed mean and estimation of contamination proportion

Consider the contamination model $F(x) = (1 - \beta)H(x) + \beta G(x)$, where $\beta \in (0, 1/2)$, and H and G are such that G is stochastically larger than H . As we have already pointed out in the Introduction, in this model, the contamination proportion β is a parameter of statistical importance. It is appropriate to note that there is a sharp curvature change in

the graph of the average of α -trimmed mean obtained from 50 Monte-carlo simulations, when α is close to the contamination proportion 0.3 in Figure 2.4. Also, in Figure 2.3, for α close to 0.3, we have observed a sharp change in the graph that plots the average over several Monte-carlo simulations of $T_n(\alpha)$, the derivative of α -trimmed mean. This behavior of α -trimmed mean and its derivative motivated us to investigate the behavior of the second derivative of α -trimmed mean when data are generated from a contamination model. The second derivative of $\theta(\alpha)$ is given by

$$\begin{aligned} \theta''(\alpha) &= \frac{2}{1-2\alpha} \left[\frac{1}{1-2\alpha} \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x f(x) dx - \frac{1}{2} \{F^{-1}(\alpha) + F^{-1}(1-\alpha)\} \right. \\ &\quad \left. - \frac{1}{2} \left\{ \frac{1}{f(F^{-1}(\alpha))} - \frac{1}{f(F^{-1}(1-\alpha))} \right\} \right]. \end{aligned}$$

One can use $S_n(\alpha) := \frac{2}{1-2\alpha} \left[\bar{x}_\alpha - \frac{1}{2} \{F_n^{-1}(\alpha) + F_n^{-1}(1-\alpha)\} - \frac{1}{2} \left\{ \frac{1}{f_n(F_n^{-1}(\alpha))} - \frac{1}{f_n(F_n^{-1}(1-\alpha))} \right\} \right]$ as a natural estimate of $\theta''(\alpha)$. Here \bar{x}_α is the sample α -trimmed mean, F_n is the empirical distribution function, and f_n is some suitable estimate of the density f . In our numerical work, we have used the kernel density estimate with an adaptive choice of bandwidth, namely, the default bandwidth in the “kde” package in the statistical software *R* as mentioned in Section 2.2. In Figure 3.1, we have plotted the average of the values of $S_n(\alpha)$ over 50 Monte-carlo replications, where each replication consists of 25 i.i.d. observations from the mixture normal distribution $0.7N(0, 1) + 0.3N(5, 1)$.

It is observed in Figure 3.1 that the maxima of $S_n(\alpha)$ is close to $\beta = 0.3$. This fact motivated us to investigate the behavior of the maxima of $\theta''(\alpha)$. In the discussion following Proposition 3.1, we will see that the behavior of the maxima of $\theta''(\alpha)$ depends on the overlap between the distributions H and G , where G is a stochastically larger than H . As an illustration, the overlap between two normal distributions with the same standard deviation but different means is shown as the black region in the next figure. A quantitative measure of the overlap between H and G is the common probability mass $\Delta_{H,G}$ between H and G , which can be defined as

$$\Delta_{H,G} = \int_{-\infty}^{\gamma} g(x) dx + \int_{\gamma}^{\infty} h(x) dx$$

Here γ is the unique intersecting point between h and g , the density functions of H and G , respectively, if we assume that such a unique point of intersection exists. On the other hand, if H and G have disjoint supports, we set $\Delta_{H,G} = 0$.

Figure 3.1 The graph of the average of $S_n(\alpha)$ obtained from 50 Monte-carlo replications each consisting of 25 i.i.d. observations from $0.7N(0, 1) + 0.3N(5, 1)$ distribution.

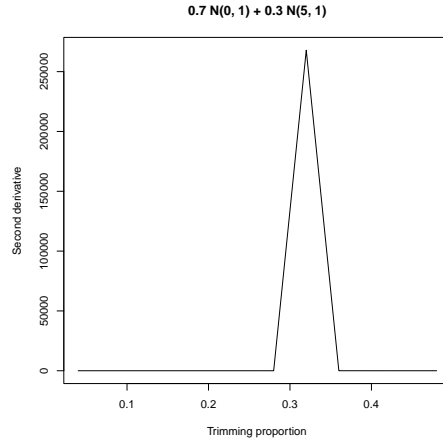
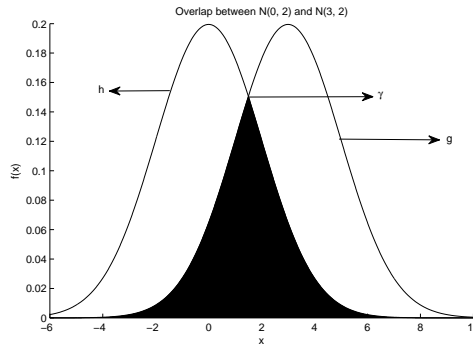


Figure 3.2 The common probability mass between two normal distributions with the same standard deviation but different means.



One can find a closed form expression for Δ_{H,H_ϕ} in the case of location contamination model, i.e., $F(x) = (1 - \beta)H(x) + \beta H_\phi(x)$, where H has a symmetric unimodal density, ϕ is a location shift, and $H_\phi(x) = H(x - \phi)$. It follows from the definition that

$$\begin{aligned} \Delta_{H,H_\phi} &= \int_{-\infty}^{\phi/2} h(x - \phi)dx + \int_{\phi/2}^{\infty} h(x)dx \\ &\quad (\phi/2 \text{ is the intersecting point of the densities of } H(x) \text{ and } H(x - \phi)) \\ &= H(-\phi/2) + 1 - H(\phi/2) = 1 - H(\phi/2) + 1 - H(\phi/2) \\ &= 2\{1 - H(\phi/2)\}. \end{aligned}$$

The relation $\Delta_{H,H_\phi} = 2\{1 - H(\phi/2)\}$ indicates that Δ_{H,H_ϕ} depends on ϕ , i.e., the amount of separation between the centers of the distributions of $H(x)$ and $H_\phi(x)$. We now state a proposition related to the behavior of the maxima of $\theta''(\alpha)$.

Proposition 3.1 *Consider the model $F(x) = (1 - \beta)H(x) + \beta G(x)$, where $\beta \in (0, 1/2)$, H is any distribution function having a continuous density h , and G is a stochastically larger than H with continuous density g . Further, suppose that h and g are positive on any compact subinterval strictly inside the supports of h and g , and they tend to zero toward the end points of their supports, i.e., $\lim_{\alpha \rightarrow 0^+} h(H^{-1}(\alpha)) = \lim_{\alpha \rightarrow 1^-} h(H^{-1}(\alpha)) = \lim_{\alpha \rightarrow 0^+} g(G^{-1}(\alpha)) = \lim_{\alpha \rightarrow 1^-} g(G^{-1}(\alpha)) = 0$. Also, assume that $f(F^{-1}(\alpha)) > f(F^{-1}(1 - \alpha))$ for all α , where $f = (1 - \beta)h + \beta g$ is the density function of F . In the case when h is supported on a compact interval, and its support is disjoint from that of g , we have $\sup_{\alpha \in [\beta - \delta, \beta + \delta], \alpha \neq \beta} \theta''(\alpha) = \infty$ and $\sup_{\alpha \notin [\beta - \delta, \beta + \delta]} \theta''(\alpha) < \infty$ for any $\delta > 0$. On the other hand, for any given H with its density h supported on the entire real line, if G , which is stochastically larger than H , varies in such a way that $\inf_{q \in (0,1)} \{G^{-1}(q) - H^{-1}(q)\} \rightarrow \infty$, we have $\theta''(\beta) \rightarrow \infty$ while $\sup_{\alpha \notin [\beta - \delta, \beta + \delta]} \theta''(\alpha)$ remains bounded above for any $\delta > 0$.*

Note that the preceding proposition implies that when $\Delta_{H,G}$ is zero or close to zero, and α lies in a small neighborhood of β , $\theta''(\alpha)$ assumes very large values; and $\theta''(\alpha)$ takes relatively smaller values for α lying outside that neighborhood. Hence, if we obtain a maximizer of an appropriate estimate of $\theta''(\alpha)$ with respect to α , that maximizer will be close to β for small values of $\Delta_{H,G}$. This behavior of the maxima of $\theta''(\alpha)$ motivated us to propose an estimate $\hat{\beta}$ of the parameter β as defined below.

$$\hat{\beta} = \arg \max_{\alpha \in [b_1, b_2]} S_n(\alpha), 0 < b_1 < b_2 < 1/2.$$

For location contamination model $F(x) = (1 - \beta)H(x) + \beta H_\phi(x)$ with H symmetric, unimodal and supported on the entire real line, which we have used in different examples in preceding and subsequent sections, it follows that $\inf_{q \in (0,1)} \{G^{-1}(q) - H^{-1}(q)\} = \phi \rightarrow \infty$ if and only if $\Delta_{H,H_\phi} \rightarrow 0$ using the relation $\Delta_{H,H_\phi} = 2\{1 - H(\phi/2)\}$ derived earlier. In view of the preceding discussion, it is expected that the performance of $\hat{\beta}$ will be good when $\Delta_{H,G}$ is small. In the following subsection, we investigate the behavior of $\hat{\beta}$ in different location contamination models with varying choices of $\Delta_{H,G}$.

3.1 A comparison with maximum likelihood estimation

Recall that the contamination model described in this section can be viewed as a special case of mixture models. The estimation of mixing proportion in mixture models is thoroughly discussed in Everitt and Hand (1981) using maximum likelihood and related techniques. We have compared the performance of our proposed estimate with some other estimates

of β based on the idea of maximum likelihood, when $\Delta_{H,G} = 10\%$, 15% and 20% for location contamination models involving normal, Cauchy and Laplace distributions. Using the relation $\Delta_{H,H_\phi} = 2\{1 - H(\phi/2)\}$, we have considered appropriate values of the location shift ϕ and varying choices of $\beta \in (0, 1/2)$ in the simulation study.

We have simulated $m = 1000$ samples from each distribution with sample sizes 100 and 1000, and calculated mean square error (m.s.e.) $= \frac{1}{m} \sum_{i=1}^m (b_i - \beta)^2$ for different estimates. Here b_i is the estimate of β for the i -th sample. We have computed the efficiency of m.l.e. relative to our estimate, where the efficiency of an estimate E_1 relative to another estimate E_2 is defined as $m.s.e.(E_2)/m.s.e.(E_1)$. Results are summarized in Figures 3.3, 3.4 and 3.5.

For mixtures of normal distributions, we have computed the efficiency of m.l.e. based on E.M. algorithm and Newton-Raphson (NR) method (see, e.g., Everitt and Hand (1981)) relative to our estimate with sample sizes 100 and 1000, and $\Delta_{H,G} = 10\%$, 15% and 20% . From the following figures, we observe that our estimate outperforms m.l.e. in all the cases considered here. The *green*, *blue* and *red* curves denote the efficiencies corresponding to $\Delta_{H,G} = 10\%$, 15% and 20% , respectively. Throughout the paper, we will use same colors to correspond to those values of $\Delta_{H,G}$.

In the case of mixtures of Cauchy distributions, the EM algorithm did not converge in our numerical studies. It is indicated by Figure 3.4 that $\hat{\beta}$ is more efficient than m.l.e. based on Newton-Raphson method in almost all cases of mixtures of Cauchy distributions considered by us.

For the mixtures of Laplace distributions, Newton-Raphson method is not feasible as the density functions involved are not differentiable. However, the EM algorithm can be carried out in this case, and it is indicated by Figure 3.5 that $\hat{\beta}$ is more efficient than m.l.e. based on EM algorithm in almost all cases considered by us when data are generated from mixtures of Laplace distributions.

The computation of maximum likelihood estimates based on Newton-Raphson method has been done by the “micEcon” package in the Statistical software *R*, and we have used the *R* codes given in Horton, Brown and Qian (2004, p.353) to compute the m.l.e. based on E.M. algorithm. Overall, our empirical study motivates us to recommend the use of $\hat{\beta}$ as an estimate of contamination proportion in asymmetric contamination model instead of estimates based on maximum likelihood techniques computed using EM algorithm or Newton-Raphson iteration.

Figure 3.3: The graphs of the efficiency of m.l.e. based on E.M. algorithm and Newton-Raphson method with respect to our estimate when data are generated from mixtures of normal distributions with sample sizes 100 and 1000.

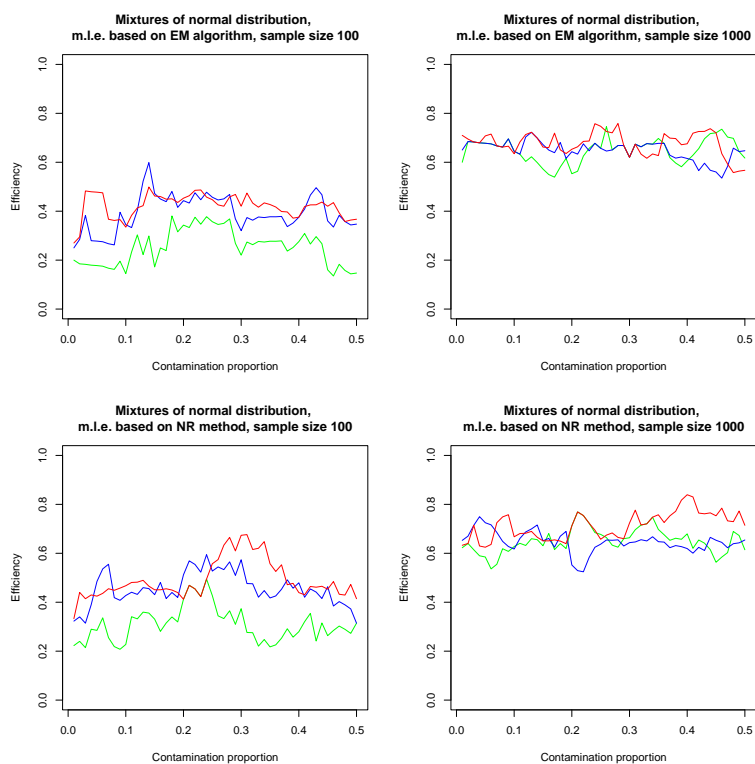


Figure 3.4 The graphs of the efficiency of m.l.e. with respect to our estimate when data are generated from mixtures of Cauchy distributions with sample sizes 100 and 1000.

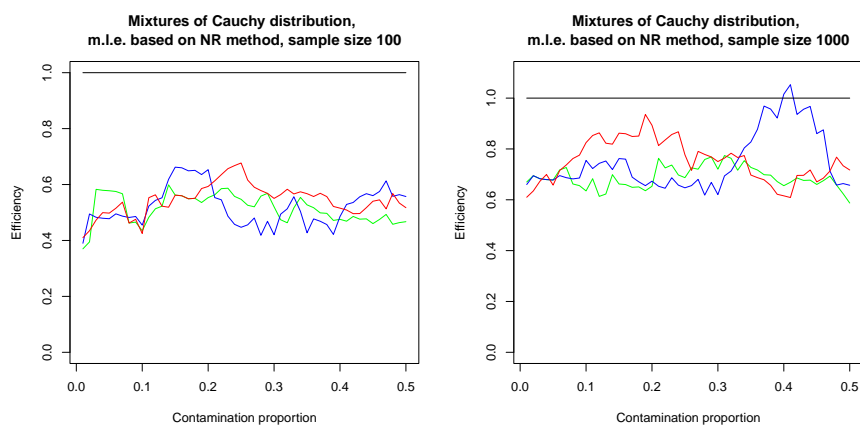
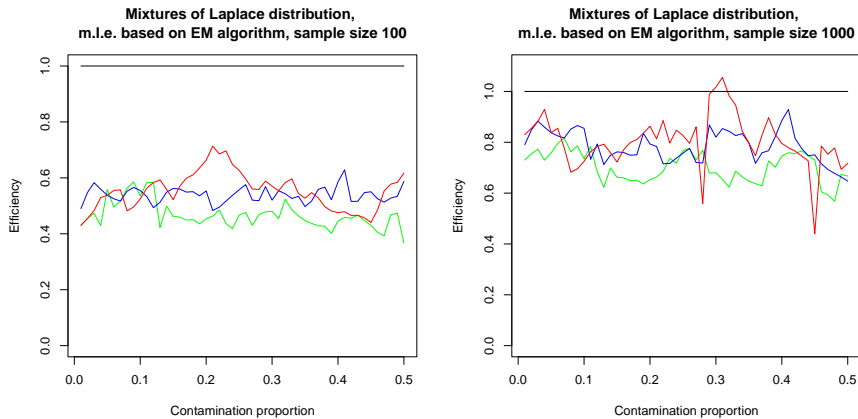


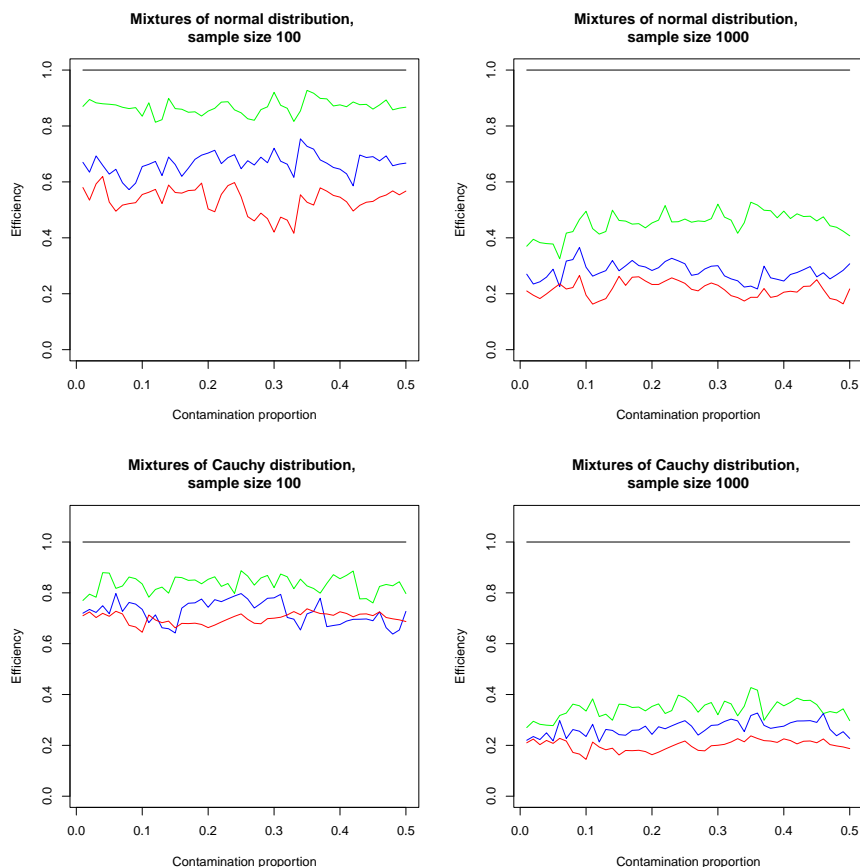
Figure 3.5 The graphs of the efficiency of m.l.e. based on EM algorithm with respect to our estimate when data are generated from mixtures of Laplace distributions with sample sizes 100 and 1000.



3.2 A comparison with Cramer-Rao lower bound

From the Figures 3.3, 3.4 and 3.5, it is clear that our proposed estimator performs better than the computed versions of m.l.e., i.e., m.l.e. based on EM algorithm or Newton-Raphson method, when sample sizes are 100 and 1000 in almost all the cases considered here. As we have already pointed out in the Introduction, $\hat{\beta}$ does not require any iterative computation while computed versions of m.l.e. are based on the iterative procedures, and the performance of those versions of m.l.e. depend on the initial values of the estimate required by those iterative computations. In all our numerical computations in the simulation study, we have used the true values of the parameters to start the iterations. Moreover, the second order differentiability of the model is required for computing m.l.e. by Newton-Raphson method whereas one needs only the existence of the density function, which may not even be differentiable for obtaining $\hat{\beta}$. However, the true m.l.e. is \sqrt{n} -consistent, asymptotically normal, and its asymptotic variance coincides with the Cramer-Rao lower bound (C.R.L.B.). We have already seen that $\hat{\beta}$ outperforms computed versions of m.l.e. in finite sample simulations for different examples considered here. So, a natural question is how does the m.s.e. of $\hat{\beta}$ compare with the true C.R.L.B., i.e., the asymptotic variance of the true m.l.e. Here we have computed the efficiency of $\hat{\beta}$ with respect to true C.R.L.B. when data follow the mixtures of normal and Cauchy distributions.

Figure 3.6 The graphs of the efficiency of $\hat{\beta}$ with respect to true C.R.L.B. when data follow mixtures of normal and Cauchy distributions with sample sizes 100 and 1000.



It is evident from Figure 3.6 that the m.s.e. of $\hat{\beta}$ is close to C.R.L.B. for sample size 100 but in the case of sample size 1000, the low efficiency of $\hat{\beta}$ is reflected in the figure.

Acknowledgment The research of the first author is partially supported by a grant from Council of Scientific and Industrial Research (CSIR), Government of India.

4 Appendix:Proofs

In order to prove Theorem 2.1, we first need to prove following two lemmas.

Lemma 4.1: *If the observations are generated from an absolutely continuous distribution function F having a positive and continuous density f on the entire real line, then, for any $\alpha_1, \alpha_2, \dots, \alpha_k \in (0, 1/2)$, where $k > 1$, the asymptotic distribution of $\sqrt{n}\{T_n(\alpha_1) - \theta'(\alpha_1), T_n(\alpha_2) - \theta'(\alpha_2), \dots, T_n(\alpha_k) - \theta'(\alpha_k)\}$ is k -variate normal with zero*

mean and a variance-covariance matrix for which the (i, j) -th entry ($1 \leq i \leq j \leq k$) is

$$\begin{aligned}
 k(\alpha_i, \alpha_j) &= \frac{2 \int_0^{F^{-1}(1-\alpha_i)} x^2 f(x) dx + 2\alpha_i F^{-1}(1-\alpha_i) F^{-1}(1-\alpha_j)}{(1-2\alpha_i)(1-2\alpha_j)} \\
 &+ \frac{\alpha_j F^{-1}(\alpha_i)}{2(1-2\alpha_i) f(F^{-1}(\alpha_j))} + \frac{\alpha_i F^{-1}(\alpha_i) + F^{-1}(1-\alpha_i) \int_{F^{-1}(1-\alpha_i)}^{F^{-1}(1-\alpha_j)} x f(x) dx}{2(1-2\alpha_i) f(F^{-1}(1-\alpha_j))} \\
 &+ \frac{\alpha_j F^{-1}(\alpha_j) + \int_{F^{-1}(\alpha_j)}^{F^{-1}(\alpha_i)} x f(x) dx}{2(1-2\alpha_j) f(F^{-1}(\alpha_i))} + \frac{\alpha_j(1-\alpha_i)}{4f(F^{-1}(\alpha_i)) f(F^{-1}(\alpha_j))} \\
 &+ \frac{\alpha_i \alpha_j}{4f(F^{-1}(\alpha_i)) f(F^{-1}(1-\alpha_j))} + \frac{\alpha_i F^{-1}(\alpha_j) + \int_{F^{-1}(\alpha_j)}^{F^{-1}(1-\alpha_i)} x f(x) dx}{2(1-2\alpha_j) f(F^{-1}(1-\alpha_i))} \\
 &+ \frac{\alpha_i \alpha_j}{4f(F^{-1}(1-\alpha_i)) f(F^{-1}(\alpha_j))} + \frac{\alpha_j(1-\alpha_i)}{4f(F^{-1}(1-\alpha_i)) f(F^{-1}(1-\alpha_j))}.
 \end{aligned}$$

Proof of Lemma 4.1: For simplicity of notation, we assume that $k = 2$. For other values of k , the proof will be similar. Recall that $T_n(\alpha) = \bar{x}_\alpha - \frac{\{F_n^{-1}(\alpha) + F_n^{-1}(1-\alpha)\}}{2}$, where $\alpha \in (0, 1/2)$. The asymptotic linear expansions of α -trimmed mean and quantiles are given in Dasgupta (2008) and Serfling (1980), respectively. As given in Dasgupta (2008), we have

$$\begin{aligned}
 &\bar{x}_\alpha - \frac{1}{(1-2\alpha)} \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x f(x) dx \\
 &= \frac{1}{n} \sum_{i=1}^n \frac{F^{-1}(\alpha) 1_{\{x_i \leq F^{-1}(\alpha)\}} + x_i 1_{\{F^{-1}(\alpha) \leq x_i \leq F^{-1}(1-\alpha)\}} + F^{-1}(1-\alpha) 1_{\{x_i \geq F^{-1}(1-\alpha)\}}}{(1-2\alpha)} + o_p\left(\frac{1}{\sqrt{n}}\right).
 \end{aligned}$$

Also, from Serfling (1980), we have

$$F_n^{-1}(\alpha) - F^{-1}(\alpha) = \frac{1}{n} \sum_{i=1}^n \frac{\alpha - 1_{\{x_i \leq F^{-1}(\alpha)\}}}{f(F^{-1}(\alpha))} + R_n,$$

where, as $n \rightarrow \infty$, $R_n = O(n^{-3/4}(\log n)^{3/4})$ with probability 1. Using these linear expansions of α -trimmed mean and quantiles, it is straightforward to see that $\sqrt{n}\{T_n(\alpha_1) - \theta'(\alpha_1), T_n(\alpha_2) - \theta'(\alpha_2)\}$ can be written as a sum of i.i.d. bivariate random vectors along with a remainder term that goes to zero in probability. By an application of C.L.T. and Slutsky's theorem, we have the asymptotic normality of $\sqrt{n}\{T_n(\alpha_1) - \theta'(\alpha_1), T_n(\alpha_2) - \theta'(\alpha_2)\}$ with zero mean. The asymptotic variance-covariance matrix as given in the statement of the lemma can be obtained by a direct algebraic computation.

Lemma 4.2: *Under the assumptions stated in Lemma 4.1 and for $\alpha \in [b_1, b_2]$, the stochastic process $\sqrt{n}\{T_n(\alpha) - \theta'(\alpha)\}$ is tight, where $0 < b_1 < b_2 < 1/2$.*

Proof of Lemma 4.2: Note that the process $T_n(\alpha)$ for $\alpha \in [b_1, b_2]$ lies in $D[b_1, b_2]$ for each $n \geq 1$, where $D[b_1, b_2]$ is the space of real functions on $[b_1, b_2]$ that are right continuous and have left-hand limits (see Billingsley (1999, p. 121)). In order to prove the tightness of $T_n(\alpha)$, one needs to verify the two conditions stated in Theorem 13.2 in Billingsley (1980, p. 139). It follows from Theorem 3.1 in Bickel (1967) that for $\alpha \in [b_1, b_2]$, where $0 < b_1 < b_2 < 1/2$, and under the conditions stated in Lemma 4.1, the standardized quantile process will converge weakly to a Gaussian process. This implies that the processes $\sqrt{n}\{F_n^{-1}(\alpha) - F^{-1}(\alpha)\}$ and $\sqrt{n}\{F_n^{-1}(1 - \alpha) - F^{-1}(1 - \alpha)\}$ are tight for $\alpha \in [b_1, b_2]$, and consequently, the quantile processes that we need to deal with here satisfy Conditions 1 and 2 in Theorem 13.2 in Billingsley (1999, p. 139). Next, we try to establish the tightness of the α -trimmed mean process. For the α -trimmed mean process, the Condition 2 in Theorem 13.2 in Billingsley (1999) related to the oscillation of the stochastic process follows from Theorem A.1 in Leger and Romano (1990, p. 311–312) considering F_n and F instead of \hat{G}_n and F_n , respectively. The Condition 1 in Theorem 13.2 in Billingsley (1999) related to the uniform boundedness of the process holds for the α -trimmed mean process because the α -trimmed mean is the average of certain quantiles. Note that here we are using the fact that quantile processes satisfy the condition of uniform boundedness for $\alpha \in [b_1, b_2]$. Consequently, the α -trimmed mean process is tight, and the process $T_n(\alpha)$ is also tight for $\alpha \in [b_1, b_2]$ and $0 < b_1 < b_2 < 1/2$. This completes the proof of the lemma.

Proof of Theorem 2.1: It follows from Lemmas 4.1 and 4.2 that any finite dimensional distribution of the stochastic process $\sqrt{n}\{T_n(\alpha) - \theta'(\alpha)\}$ is multivariate normal, and the process satisfies the tightness condition. Therefore, $\sqrt{n}\{T_n(\alpha) - \theta'(\alpha)\}$ converges weakly to a Gaussian process in view of Theorem 13.1 in Billingsley (1999, p. 139).

Proof of Corollary 2.2: We have already seen in the proof of Lemma 4.2 that the stochastic process $\sqrt{n}\{T_n(\alpha) - \theta'(\alpha)\}$ is tight in $D[b_1, b_2]$ equipped with the supremum norm. Now, it follows from Condition 1 in Theorem 13.2 in Billingsley (1999, p.139) that, for every $0 < \eta < 1$ and for any $0 < b_1 < b_2 < 1/2$, there exists a constant $M(\eta, b_1, b_2) > 0$ such that $P[\sup_{\alpha \in [b_1, b_2]} \sqrt{n}|T_n(\alpha) - \theta'(\alpha)| \leq M(\eta, b_1, b_2)] > 1 - \eta$, which implies that $\sup_{\alpha \in [b_1, b_2]} \sqrt{n}|T_n(\alpha) - \theta'(\alpha)| = O_p(n^{-1/2})$. This completes the proof.

Proof of Proposition 2.3: The α -trimmed mean $\theta(\alpha) = \frac{1}{(1-2\alpha)} \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} xf(x)dx$ will

be a decreasing function of α if

$$\begin{aligned} \frac{d}{d\alpha}\theta(\alpha) \leq 0 &\Leftrightarrow \frac{1}{(1-2\alpha)} \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} xf(x)dx \leq \frac{F^{-1}(\alpha) + F^{-1}(1-\alpha)}{2} \\ &\Leftrightarrow \lim_{N \rightarrow \infty} \sum_{j=1}^{j=N} \frac{1}{N} \frac{F^{-1}(\alpha_j) + F^{-1}(1-\alpha_j)}{2} \leq \frac{F^{-1}(\alpha) + F^{-1}(1-\alpha)}{2}, \quad (1) \end{aligned}$$

where $(\alpha_1, \alpha_2, \dots, \alpha_N)$ is an equally spaced partition of $(\alpha, 1-\alpha)$. The last implication follows from the convergence of Riemann sum to the Riemann integral. In order to prove (1), it is enough to show that for any $j = 1, \dots, N$, $\frac{F^{-1}(\alpha_j) + F^{-1}(1-\alpha_j)}{2}$ is smaller than $\frac{F^{-1}(\alpha) + F^{-1}(1-\alpha)}{2}$. So, it is sufficient to prove that $\frac{F^{-1}(\alpha) + F^{-1}(1-\alpha)}{2}$ is a decreasing function of α , i.e.,

$$\frac{d}{d\alpha} \frac{F^{-1}(\alpha) + F^{-1}(1-\alpha)}{2} \leq 0 \Leftrightarrow f(F^{-1}(\alpha)) \geq f(F^{-1}(1-\alpha)).$$

This completes the proof of the proposition.

Proof of Theorem 2.4: Recall from Theorem 2.1 that for $\alpha \in [b_1, b_2]$, $\sqrt{n}\{T_n(\alpha) - \theta'(\alpha)\}$ converges weakly to a Gaussian process with zero mean and covariance function $k(\alpha_1, \alpha_2)$. Here, $k(\alpha_1, \alpha_2)$ is a non-negative kernel, and it follows from its expression in Lemma 4.1 that it is a continuous function of α_1 and α_2 . Let the $e_i(\alpha)$'s and the λ_i 's be the eigen functions and eigen values of the kernel $k(\alpha_1, \alpha_2)$, respectively. Now, it follows from the Karhunen-Loeve expansion (see, e.g., Loeve (1978)) of the weak limit of the process $\sqrt{n}\{T_n(\alpha) - \theta'(\alpha)\}$ that this process converges in distribution to the process $\sum_{i=1}^{\infty} Z_i e_i(\alpha)$, where the Z_i 's are independent random variables such that Z_i has $N(0, \lambda_i)$ distribution, and $\alpha \in [b_1, b_2]$. It now follows using the continuity of the integral functional on $D[b_1, b_2]$ equipped with the supremum norm and the orthonormality of the $e_i(\alpha)$'s that, under H_0 , when $\theta'(\alpha) = 0$ for all $\alpha \in [b_1, b_2]$, $\int_{b_1}^{b_2} n\{T_n(\alpha)\}^2 d\alpha$ converges weakly to $\sum_{i=1}^{\infty} Z_i^2$.

We now show that $\int_{b_1}^{b_2} n\{T_n(\alpha)\}^2 d\alpha - \sum_{i=[nb_1]+1}^{[nb_2]-1} \{T_n(i/n)\}^2$ converges to zero in probability. Note that $\int_{b_1}^{b_2} n\{T_n(\alpha)\}^2 d\alpha = \int_{b_1}^{b_2} n \left\{ \bar{x}_\alpha - \frac{F_n^{-1}(\alpha) + F_n^{-1}(1-\alpha)}{2} \right\}^2 d\alpha$. In other words, it is enough to show that $n \int_{b_1}^{b_2} \{F_n^{-1}(\alpha)\}^2 d\alpha - \sum_{i=[nb_1]+1}^{[nb_2]-1} \{F_n^{-1}(i/n)\}^2 \xrightarrow{P} 0$, $n \int_{b_1}^{b_2} \{F_n^{-1}(1-\alpha)\}^2 d\alpha - \sum_{i=[nb_1]+1}^{[nb_2]-1} \{F_n^{-1}(1-i/n)\}^2 \xrightarrow{P} 0$ and $\int_{b_1}^{b_2} \{\bar{x}_\alpha\}^2 d\alpha - \sum_{i=[nb_1]+1}^{[nb_2]-1} \bar{x}_{i/n}^2 \xrightarrow{P} 0$.

Note that

$$\sum_{i=[nb_1]+1}^{[nb_2]-1} \{F_n^{-1}(i/n)\}^2 = \left[x_{([nb_1]+1)}^2 + \dots + x_{([nb_2]-1)}^2 \right],$$

and

$$\begin{aligned}
n \int_{b_1}^{b_2} \{F_n^{-1}(\alpha)\}^2 d\alpha &= n \left[\int_{b_1}^{([nb_1]+1)/n} \{F_n^{-1}(\alpha)\}^2 d\alpha + \int_{([nb_1]+1)/n}^{([nb_1]+2)/n} \{F_n^{-1}(\alpha)\}^2 d\alpha \right. \\
&\quad \left. + \dots + \int_{([nb_2]-1)/n}^{[nb_2]/n} \{F_n^{-1}(\alpha)\}^2 d\alpha + \int_{([nb_2]-1)/n}^{b_2} \{F_n^{-1}(\alpha)\}^2 d\alpha \right] \\
&= \left[x_{[nb_1]+1}^2 + \dots + x_{([nb_2]-1)}^2 \right] + x_{([nb_2])}^2 \{b_2 - [nb_2]/n\}.
\end{aligned}$$

Hence, $n \int_{b_1}^{b_2} \{F_n^{-1}(\alpha)\}^2 d\alpha - \sum_{i=[nb_1]+1}^{[nb_2]-1} \{F_n^{-1}(i/n)\}^2 = x_{([nb_2])}^2 (b_2 - [nb_2]/n) \xrightarrow{P} 0$ using the facts that $(b_2 - [nb_2]/n) \rightarrow 0$ and $x_{([nb_2])} = O_p(1)$ as $n \rightarrow \infty$. The proofs of $n \int_{b_1}^{b_2} \{F_n^{-1}(1 - \alpha)\}^2 d\alpha - \sum_{i=[nb_1]+1}^{[nb_2]-1} \{F_n^{-1}(1 - i/n)\}^2 \xrightarrow{P} 0$ and $\int_{b_1}^{b_2} \{\bar{x}_\alpha\}^2 d\alpha - \sum_{i=[nb_1]+1}^{[nb_2]-1} \bar{x}_{i/n}^2 \xrightarrow{P} 0$ follow after similar algebraic computations.

Next, we try to calculate the asymptotic power of the proposed test. The asymptotic power of the test is given by $\lim_{n \rightarrow \infty} P_{H_1} \left[\sum_{i=[nb_1]+1}^{[nb_2]-1} \{T_n(i/n)\}^2 > c \right]$, where c is the critical value determined from the level of the test based on the distribution of $\sum_{i=1}^{\infty} Z_i^2$. Now, we have

$$\begin{aligned}
&\lim_{n \rightarrow \infty} P_{H_1} \left[\sum_{i=[nb_1]+1}^{[nb_2]-1} \{T_n(i/n)\}^2 > c \right] \\
&= \lim_{n \rightarrow \infty} P_{H_1} \left[\int_{b_1}^{b_2} n \{T_n(\alpha)\}^2 d\alpha > c \right] \quad (\text{since } \int_{b_1}^{b_2} n \{T_n(\alpha)\}^2 - \sum_{i=[nb_1]+1}^{[nb_2]-1} \{T_n(i/n)\}^2 \xrightarrow{P} 0) \\
&= \lim_{n \rightarrow \infty} P_{H_1} \left[\int_{b_1}^{b_2} n \{T_n(\alpha) - \theta'(\alpha)\}^2 d\alpha > c - 2n \int_{b_1}^{b_2} T_n(\alpha) \theta'(\alpha) d\alpha + n \int_{b_1}^{b_2} \{\theta'(\alpha)\}^2 d\alpha \right] \\
&= \lim_{n \rightarrow \infty} P_{H_1} \left[\sum_{i=1}^{\infty} Z_i^2 > c - n \int_{b_1}^{b_2} \{\theta'(\alpha)\}^2 d\alpha \right] \\
&\quad (\text{since } 2 \int_{b_1}^{b_2} T_n(\alpha) \theta'(\alpha) d\alpha \xrightarrow{P} 2 \int_{b_1}^{b_2} \{\theta'(\alpha)\}^2 d\alpha \text{ using Corollary 2.2}) \\
&= 1 \quad (\text{since } n \int_{b_1}^{b_2} \{\theta'(\alpha)\}^2 d\alpha \rightarrow \infty \text{ as } n \rightarrow \infty \text{ since } \theta'(\alpha) \neq 0 \text{ for some } 0 < b_1 < b_2 < 1/2)
\end{aligned}$$

This completes the proof of the theorem.

Proof of Proposition 3.1: First, we consider the case, when H is supported on a compact interval, and H and G have disjoint supports. Recall that

$$\begin{aligned}
\theta''(\alpha) &= \frac{2}{1-2\alpha} \left[\frac{1}{1-2\alpha} \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x f(x) dx - \frac{1}{2} \{F^{-1}(\alpha) + F^{-1}(1-\alpha)\} \right. \\
&\quad \left. - \frac{1}{2} \left\{ \frac{1}{f(F^{-1}(\alpha))} - \frac{1}{f(F^{-1}(1-\alpha))} \right\} \right].
\end{aligned}$$

It follows from the proof of Proposition 2.3 that, under the skewness condition $f(F^{-1}(\alpha)) > f(F^{-1}(1 - \alpha))$, $\left[\frac{1}{1-2\alpha} \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} xf(x)dx - \frac{1}{2}\{F^{-1}(\alpha) + F^{-1}(1 - \alpha)\} \right]$ is bounded above for all $\alpha \in [b_1, b_2]$. So, in order to prove Proposition 3.1, it is enough to investigate the behavior of $M(\alpha) := -\frac{1}{2} \left[\frac{1}{f(F^{-1}(\alpha))} - \frac{1}{f(F^{-1}(1-\alpha))} \right]$, where $\alpha \in [b_1, b_2]$ and $0 < b_1 < b_2 < 1/2$.

Assume that $y_1 = F^{-1}(\alpha)$. If $\alpha < (1 - \beta)$, then y_1 will be located inside the support of H as G is stochastically larger than H , and they have disjoint supports. In other words,

$$(1 - \beta)H(y_1) = \alpha \Leftrightarrow y_1 = H^{-1}\left(\frac{\alpha}{(1 - \beta)}\right).$$

In the case $\alpha > (1 - \beta)$, y_1 will be located inside the support of G , and in that case,

$$\begin{aligned} F^{-1}(\alpha) = y_1 &\Leftrightarrow (1 - \beta)H(y_1) + \beta G(y_1) = \alpha \\ &\Leftrightarrow (1 - \beta) + \beta G(y_1) = \alpha \text{ (since } H(y_1) = 1) \Leftrightarrow y_1 = G^{-1}\left(\frac{\alpha - (1 - \beta)}{\beta}\right). \end{aligned}$$

If $\alpha = (1 - \beta)$, then $F^{-1}(\alpha)$ can be defined as any point that lies between $H^{-1}(1)$ and $G^{-1}(0)$.

Next, when H and G have disjoint supports, we show that $\sup_{\alpha \in [\beta - \delta, \beta + \delta], \alpha \neq \beta} M(\alpha) = \infty$, and $\sup_{\alpha \notin [\beta - \delta, \beta + \delta]} M(\alpha) < \infty$ for any $\delta > 0$. Now, we have

$$\begin{aligned} \lim_{\alpha \rightarrow \beta^+} M(\alpha) &= \lim_{\alpha \rightarrow \beta^+} -\frac{1}{2} \left[\frac{1}{(1 - \beta)h(F^{-1}(\alpha)) + \beta g(F^{-1}(\alpha))} - \frac{1}{(1 - \beta)h(F^{-1}(1 - \alpha)) + \beta g(F^{-1}(1 - \alpha))} \right] \\ &= \lim_{\alpha \rightarrow \beta^+} -\frac{1}{2} \left[\frac{1}{(1 - \beta)h(H^{-1}(\frac{\alpha}{1 - \beta})) + \beta g(H^{-1}(\frac{\alpha}{1 - \beta}))} - \frac{1}{(1 - \beta)h(H^{-1}(\frac{1 - \alpha}{1 - \beta})) + \beta g(H^{-1}(\frac{1 - \alpha}{1 - \beta}))} \right] \\ &= \infty \text{ (since } h(H^{-1}(x)) \rightarrow 0 \text{ as } x \rightarrow 1- \text{ and } g(H^{-1}(x)) \rightarrow 0 \text{ as } x \rightarrow 1- \text{).} \end{aligned}$$

In the same way,

$$\begin{aligned} \lim_{\alpha \rightarrow \beta^-} M(\alpha) &= \lim_{\alpha \rightarrow \beta^-} -\frac{1}{2} \left[\frac{1}{(1 - \beta)h(F^{-1}(\alpha)) + \beta g(F^{-1}(\alpha))} - \frac{1}{(1 - \beta)h(F^{-1}(1 - \alpha)) + \beta g(F^{-1}(1 - \alpha))} \right] \\ &= \lim_{\alpha \rightarrow \beta^-} -\frac{1}{2} \left[\frac{1}{(1 - \beta)h(H^{-1}(\frac{\alpha}{1 - \beta})) + \beta g(H^{-1}(\frac{\alpha}{1 - \beta}))} - \frac{1}{(1 - \beta)h(G^{-1}(\frac{(1 - \alpha) - (1 - \beta)}{\beta})) + \beta g(G^{-1}(\frac{(1 - \alpha) - (1 - \beta)}{\beta}))} \right] \\ &= \infty \text{ (since } h(G^{-1}(x)) \rightarrow 0 \text{ as } x \rightarrow 0+ \text{ and } g(G^{-1}(x)) = 0 \text{ as } x \rightarrow 0+ \text{).} \end{aligned}$$

Hence, $\lim_{\alpha \rightarrow \beta} M(\alpha) = \infty$. This implies that $\sup_{\beta - \delta \leq \alpha \leq \beta + \delta, \alpha \neq \beta} M(\alpha) = \infty$.

Next, we try to investigate $M(\alpha)$ when $\alpha < \beta - \delta$ or $\alpha > \beta + \delta$ for any $\delta > 0$. If $\alpha > \beta + \delta$, we have

$$\begin{aligned} M(\alpha) &= -\frac{1}{2} \left[\frac{1}{(1-\beta)h(F^{-1}(\alpha)) + \beta g(F^{-1}(\alpha))} - \frac{1}{(1-\beta)h(F^{-1}(1-\alpha)) + \beta g(F^{-1}(1-\alpha))} \right] \\ &= -\frac{1}{2} \left[\frac{1}{(1-\beta)h(H^{-1}(\frac{\alpha}{(1-\beta)})) + \beta g(H^{-1}(\frac{\alpha}{(1-\beta)}))} - \frac{1}{(1-\beta)h(H^{-1}(\frac{(1-\alpha)}{(1-\beta)})) + \beta g(H^{-1}(\frac{(1-\alpha)}{(1-\beta)}))} \right]. \end{aligned}$$

The last expression is bounded as h and g are continuous and positive on any compact subinterval strictly inside the supports of h and g . Next, we consider $\alpha < \beta - \delta$. In that case,

$$\begin{aligned} M(\alpha) &= -\frac{1}{2} \left[\frac{1}{(1-\beta)h(F^{-1}(\alpha)) + \beta g(F^{-1}(\alpha))} - \frac{1}{(1-\beta)h(F^{-1}(1-\alpha)) + \beta g(F^{-1}(1-\alpha))} \right] \\ &= -\frac{1}{2} \left[\frac{1}{(1-\beta)h(H^{-1}(\frac{\alpha}{(1-\beta)})) + \beta g(H^{-1}(\frac{\alpha}{(1-\beta)}))} - \frac{1}{(1-\beta)h(G^{-1}(\frac{(1-\alpha)-(1-\beta)}{\beta})) + \beta g(G^{-1}(\frac{(1-\alpha)-(1-\beta)}{\beta}))} \right]. \end{aligned}$$

Again, the last expression is bounded as h and g are continuous and positive on any compact subinterval strictly inside the supports of h and g . This completes the proof in the case when H is supported on compact interval, and H and G have disjoint supports.

Next, we consider the case when H is supported on the entire real line, and G varies in such a way that $G >_{st} H$ and $\inf_{q \in (0,1)} \{G^{-1}(q) - H^{-1}(q)\} \rightarrow \infty$. As in the case of disjoint supports for H and G , it is enough to investigate the term $M(\alpha) = -\frac{1}{2} \left[\frac{1}{f(F^{-1}(\alpha))} - \frac{1}{f(F^{-1}(1-\alpha))} \right]$ that appears in the expression of $\theta''(\alpha)$.

If possible, suppose now that $\sup_{\alpha > \beta + \delta} \theta''(\alpha)$ is not bounded above as G varies satisfying the conditions stated above. So, there must exist a sequence of distributions $G_n >_{st} H$ and a sequence of positive real numbers $1/2 > \alpha_n > \beta + \delta$ such that $\inf_{q \in (0,1)} \{G_n^{-1}(q) - H^{-1}(q)\} \rightarrow \infty$ and $\theta''_n(\alpha_n) \rightarrow \infty$ as $n \rightarrow \infty$. Let y_n be the α_n -th quantile of $F_n = (1-\beta)H + \beta G_n$, i.e., $y_n = F_n^{-1}(\alpha_n)$. So, we have

$$\begin{aligned} F_n^{-1}(\alpha_n) = y_n &\Leftrightarrow (1-\beta)H(y_n) + \beta G_n(y_n) = \alpha_n \\ &\Leftrightarrow (1-\beta)H(y_n) + \beta H(y_n - z_{y_n}) = \alpha_n \quad (z_{y_n} > 0 \text{ as } G_n >_{st} H). \end{aligned} \quad (2)$$

Note that $z_{y_n} \rightarrow \infty$ as $n \rightarrow \infty$ for any y_n since $\inf_{q \in (0,1)} \{G_n^{-1}(q) - H^{-1}(q)\} \rightarrow \infty$ as $n \rightarrow \infty$. If $\theta''_n(\alpha_n) \rightarrow \infty$ (and hence, $M_n(\alpha) \rightarrow \infty$) as $n \rightarrow \infty$, then we must have

either $y_n = F_n^{-1}(\alpha_n) \rightarrow \pm\infty$ and $y_n - z_{y_n} \rightarrow \pm\infty$, or $u_n := F_n^{-1}(1 - \alpha_n) \rightarrow \pm\infty$ and $u_n - z_{u_n} \rightarrow \pm\infty$ as $n \rightarrow \infty$. This follows from the expression

$$M_n(\alpha_n) = -\frac{1}{2} \left[\frac{1}{(1 - \beta)h(y_n) + \beta h(y_n - z_{y_n})} - \frac{1}{(1 - \beta)h(u_n) + \beta h(u_n - z_{u_n})} \right],$$

and the fact that h is positive and continuous on any compact subinterval within its support. If $y_n \rightarrow -\infty$ and $z_{y_n} \rightarrow \infty$, we have $y_n - z_{y_n} \rightarrow -\infty$ because $z_{y_n} \rightarrow \infty$. Now, using $y_n \rightarrow -\infty$ and $y_n - z_{y_n} \rightarrow -\infty$ in (2), we have $\alpha_n \rightarrow 0$ as $n \rightarrow \infty$, which contradicts the fact that $\alpha_n > \beta + \delta$ for all n . If $y_n \rightarrow \infty$ and $z_{y_n} \rightarrow \infty$, then $y_n - z_{y_n}$ may be bounded or unbounded. In that case, we again need to consider several distinct possibilities, as either $y_n - z_{y_n}$ remains bounded as $n \rightarrow \infty$, or *we can extract a subsequence along which* $y_n - z_{y_n} \rightarrow \pm\infty$ as $n \rightarrow \infty$. Using $y_n \rightarrow \infty$ and $y_n - z_{y_n} \rightarrow \infty$ in (2), we have $\alpha_n \rightarrow 1$, which is a contradiction to the fact that $\alpha_n < 1/2$ for all n . Lastly, if we consider the case $y_n \rightarrow \infty$ while $y_n - z_{y_n}$ either remains bounded or tend to $-\infty$ as $n \rightarrow \infty$, then it follows from (2) that for all n sufficiently large, $\alpha_n > 1/2$, in view of the fact that $(1 - \beta) > 1/2$, and this is again a contradiction. Combining all these, it now follows that the sequence y_n must remain bounded as $n \rightarrow \infty$. In a similar way, one can show that u_n remains bounded as $n \rightarrow \infty$ using the equation

$$(1 - \beta)H(u_n) + \beta H(u_n - z_{u_n}) = (1 - \alpha_n), \quad (3)$$

where z_{u_n} satisfies $G_n(u_n) = H(u_n - z_{u_n})$, and hence $z_{u_n} \rightarrow \infty$ as $n \rightarrow \infty$ in view of the condition $\inf_{q \in (0,1)} \{G_n^{-1}(q) - H^{-1}(q)\} \rightarrow \infty$ as $n \rightarrow \infty$. This completes the proof of boundedness of $\sup_{\alpha > \beta + \delta} M_n(\alpha)$ as $n \rightarrow \infty$.

Next, if possible, assume that $\sup_{\alpha < \beta - \delta} \theta_n''(\alpha)$ is not bounded above. Hence, as before, there must exist a sequence of distributions $G_n >_{st} H$ and a sequence of positive real numbers $1/2 > \alpha_n > \beta + \delta$ satisfying the conditions as stated in the preceding paragraph. Once again, if $\theta''(\alpha_n) \rightarrow \infty$ as $n \rightarrow \infty$, it follows from the preceding expression of $M_n(\alpha_n)$ that, as $n \rightarrow \infty$, either $y_n \rightarrow \pm\infty$ and $y_n - z_{y_n} \rightarrow \pm\infty$, or we have $u_n \rightarrow \pm\infty$ and $u_n - z_{u_n} \rightarrow \pm\infty$. Now, arguing in a similar way as in the preceding paragraph and using (3), one can show that $u_n - z_{u_n}$ remains bounded as $n \rightarrow \infty$. Since h is positive and continuous on any compact subinterval of its support, we can conclude that $\sup_{\alpha < \beta - \delta} M_n(\alpha)$ remains bounded as $n \rightarrow \infty$.

Finally, we try to investigate the case $\alpha = \beta$. Here also, let G_n be a sequence of distributions satisfying $G_n >_{st} H$ and $\inf_{q \in (0,1)} \{G_n^{-1}(q) - H^{-1}(q)\} \rightarrow \infty$ as $n \rightarrow \infty$. Let

$v_n := F_n^{-1}(1 - \beta)$ and $w_n := F_n^{-1}(\beta)$. Hence, we have

$$(1 - \beta)H(v_n) + \beta H(v_n - z_{v_n}) = (1 - \beta), \quad (4)$$

where z_{v_n} satisfies $G_n(v_n) = H(v_n - z_{v_n})$, and consequently, $z_{v_n} \rightarrow \infty$ as $n \rightarrow \infty$ in view of the condition $\inf_{q \in (0,1)} \{G_n^{-1}(q) - H^{-1}(q)\} \rightarrow \infty$ as $n \rightarrow \infty$. Now, we try to show that (4) will be satisfied only when $v_n \rightarrow \infty$ and $v_n - z_{v_n} \rightarrow -\infty$ as $n \rightarrow \infty$. If $v_n \rightarrow -\infty$ and $v_n - z_{v_n} \rightarrow -\infty$ as $n \rightarrow \infty$ in (4), then we have $\beta = 1$, which is a contradiction to the fact $\beta < 1/2$. On the other hand, if v_n is bounded and $v_n - z_{v_n} \rightarrow -\infty$ as $n \rightarrow \infty$, then (4) along with the fact that H is supported on the entire real line implies that the left hand side of (4) is strictly smaller than $(1 - \beta)$ for all n sufficiently large, which is also a contradiction. If both v_n and $v_n - z_{v_n}$ tend to ∞ as $n \rightarrow \infty$, it follows from (4) that $\beta = 0$, which is not possible. Further, if we consider the possibility that $v_n \rightarrow \infty$ and $v_n - z_{v_n}$ remains bounded as $n \rightarrow \infty$ in (4), then $\{(1 - \beta) - (1 - \beta)H(v_n)\}$ tends to zero as $n \rightarrow \infty$ while $\beta H(v_n - z_{v_n})$ remains bounded away from zero as $n \rightarrow \infty$, which leads to a contradiction in view of (4). Combining all these facts, one can conclude that (4) will be satisfied only if $v_n \rightarrow \infty$ and $v_n - z_{v_n} \rightarrow -\infty$ as $n \rightarrow \infty$. Consequently, we must have $M_n(\beta) = \left[\frac{1}{(1-\beta)h(w_n)+\beta h(w_n-z_{w_n})} - \frac{1}{(1-\beta)h(v_n)+\beta h(v_n-z_{v_n})} \right] \rightarrow \infty$ as $n \rightarrow \infty$ since $h(x) \rightarrow 0$ as $x \rightarrow \pm\infty$. This completes the proof of the theorem.

References

- Ahmad, I. A. and Li, Q. (1997) Testing symmetry of an unknown density function by kernel method. *Journal of Nonparametric Statistics* **7**, 279–293.
- Arcones, M. A. and Gine, E. (1991) Some Bootstrap Tests of Symmetry for Univariate Continuous Distributions. *The Annals of Statistics* **19**, 1496–1511.
- Bickel, P. J. (1965) On some robust estimates of location. *The Annals of Statistics* **36**, 847–858.
- Bickel, P. J. (1967) Some contributions to the theory of order statistics. *Proceedings of the Fifth Berkeley Symposium Mathematical Statistics and Probability* **1**, 575–591.
- Billingsley, P. (1968) *Convergence of Probability Measures*. Wiley Series in Probability and Statistics.
- Cabilio, P. and Masaro, J. (1996) A simple test of symmetry about an unknown median. *The Canadian Journal of Statistics* **24**, 349–361.
- Csorgo, S. and Heathcote, C. R. (1987) Testing for Symmetry. *Biometrika* **74**, 177–184.

- DasGupta, A. (2008) *Asymptotic Theory of Statistics and Probability*, Springer Verlag, New York, March 2008.
- Dhar, S. S. and Chaudhuri, P. (2009) A Comparison of Robust Estimators Based on Two Types of Trimming. *Advances in Statistical Analysis* **93**, 151–158.
- Everitt, B. S. and Hand, D. J. (1981) *Finite Mixture Distributions*. Chapman & Hall.
- Gastwirth, J. L. (1971) On the Sign Test for Symmetry. *Journal of the American Statistical Association* **66**, 821–823.
- Gupta, M. K. (1967) An Asymptotically Nonparametric Test of Symmetry. *The Annals of Mathematical Statistics* **38**, 849–866.
- Hall, P. (1981) Large sample properties of Jackel's adaptive trimmed mean. *The Annals of Institute of Statistical Mathematics* **33**, 449–462.
- Hogg, R. V. (1967) A Partial Review and Some Suggestions for Future Applications and Theory. *Journal of the American Statistical Association* **69**, 909–923.
- Horton, N. J., Brown, E. R. and Qian, L. (2004) Use of R as a Toolbox for Mathematical Statistics Exploration. *The American Statistician* **58**, 343–357.
- Huber, P. J. (1981) *Robust Statistics*. John Wiley & Sons, New York.
- Jaekel, L. A. (1971) Some Flexible Estimates of Location. *The Annals of Mathematical Statistics* **42**, 1540–1552.
- Leger, C. and Romano, J. P. (1990) Bootstrap Adaptive Estimation: The Trimmed-Mean Example. *The Canadian Journal of Statistics* **18**, 297–314.
- Loeve, M. (1978) *Probability Theory II*. Springer-Verlag.
- Mira, A. (1999) Distribution-free test for symmetry based on Bonferroni's measure. *Journal of Applied Statistics* **26**, 959–972.
- Modarres, R. and Gastwirth, J. L. (1996) A modified runs test for symmetry. *Statistics and Probability Letters* **31**, 107–112.
- Reven, G. M. and Miller, R. G. (1979) An attempt to define the nature of chemical diabetes using a multidimensional analysis. *Diabetologia* **16**, 17–24.
- Schuster, E. F. and Barker, R. C. (1987) Using the bootstrap in testing symmetry versus asymmetry. *Communication in Statistics: Simulation and Computation* **16**, 69–84.
- Serfling, R. J. (1980) *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons, New York.
- Stigler, S. M. (1973) The Asymptotic Distribution of The Trimmed Mean. *The Annals of Statistics* **1**, 472–477.

Tajauddin, I. H. (1994) Distribution-free test for symmetry based on Wilcoxon two-sample test. *Journal of Applied Statistics* **21**, 409–415.

Tukey, J. W. (1948) Some elementary problems of importance to small sample practice. *Human Biology* **20**, 205–214.

Tukey, J. W. and McLaughlin, D. H. (1963) Less Vulnerable Confidence and Significance Procedures for Location Based on a Single Sample:Trimming/Winsorization 1. *Sankhya, Series A* **25**, 331–352.

van Zwet, W. R. (1979) Mean, Median, Mode, II. *Statist. Neerl.* **33**, 1–5.

203 B. T. Road, Theoretical Statistics and Mathematics Unit

Indian Statistical Institute, Calcutta.

E-mail: subhra_r@isical.ac.in, probal@isical.ac.in.