Bahadur Representations for the Median Absolute Deviation and Its Modifications

Satyaki Mazumder[1] and Robert Serfling[2]

University of Texas at Dallas

May 2009
(Revision for *Statistics and Probability Letters*)

[1]Department of Mathematical Sciences, University of Texas at Dallas, Richardson, Texas 75083-0688, USA.

[2]Department of Mathematical Sciences, University of Texas at Dallas, Richardson, Texas 75083-0688, USA. Email: serfling@utdallas.edu. Website: www.utdallas.edu/~serfling. Support by NSF Grants DMS-0103698 and DMS-0805786 and NSA Grant H98230-08-1-0106 is gratefully acknowledged.

Abstract

The median absolute deviation about the median (MAD) is an important univariate spread measure having wide appeal due to its highly robust sample version. A powerful tool in treating the asymptotics of a statistic is a linearization, i.e., a Bahadur representation. Here we establish both strong and weak Bahadur representations for the sample MAD. The strong version is the first in the literature, while the weak version improves upon previous treatments by reducing regularity conditions. Our results also apply to a modified version of sample MAD (Tyler, 1994, and Gather and Hilker, 1997) introduced to obtain improved robustness for statistical procedures using sample median and MAD combinations over the univariate projections of multivariate data. The strong version yields the law of iterated logarithm for the sample MAD and supports study of almost sure properties of randomly trimmed means based on the MAD and development of robust sequential nonparametric confidence intervals for the MAD. The weak version is needed to simplify derivations of the asymptotic joint distributions of vectors of dependent sample median and sample MAD combinations, which arise in constructing nonparametric multivariate outlyingness functions via projection pursuit.

AMS 2000 Subject Classification: Primary 60F15 Secondary 62G20

Key words and phrases: Median absolute deviation; Bahadur representation; Law of iterated logarithm.

1 Introduction

The MAD (median absolute deviation about the median) is an important nonparametric spread measure with a highly robust sample version, of which a modified form is used in certain contexts. Past studies provide for the sample MAD and its modifications almost sure convergence, an exponential probability inequality, and joint asymptotic normality with the sample median. For treating the asymptotics of a statistic, a powerful tool is linearization, i.e., a Bahadur representation. Here we establish strong and weak Bahadur representations for both the sample MAD and its modifications. The strong version is entirely new and yields the law of iterated logarithm for the sample MAD, supports study of almost sure properties of randomly trimmed means based on the MAD, and facilitates development of robust sequential nonparametric confidence intervals for the MAD. The weak version relaxes regularity conditions assumed in previous treatments and is useful to simplify derivations of asymptotic joint distributions of vectors of dependent sample median and sample MAD combinations, as arise for example in constructing nonparametric multivariate outlyingness functions via projection pursuit.

Let us now make this precise. Let X have univariate distribution F. The *median* of F, or $\mathrm{Med}(F)$, is defined by $\nu = F^{-1}(1/2) = \inf\{x : F(x) \geq 1/2\}$ and satisfies

$$F(\nu-) \leq 1/2 \leq F(\nu). \tag{1}$$

The distribution G of $|X - \nu|$, i.e.,

$$G(y) = P(|X - \nu| \leq y) = F(\nu + y) - F(\nu - y-),\ y \in \mathbb{R}, \tag{2}$$

has median $\zeta = G^{-1}(1/2)$ satisfying

$$G(\zeta-) \leq 1/2 \leq G(\zeta). \tag{3}$$

The median ζ of G defines a scale parameter of F, the *median absolute deviation about the median (MAD)*, i.e., $\mathrm{Med}(G) = \mathrm{MAD}(F)$ (*not* the *mean* absolute deviation about the *mean*, sometimes also abbreviated by "MAD").

Sample versions Med_n and MAD_n for a random sample $\mathbb{X}_n = \{X_1, \ldots, X_n\}$ from F are defined as follows. With $X_{1:n} \leq \ldots \leq X_{n:n}$ the ordered sample values,

$$\mathrm{Med}_n = \frac{1}{2}\left(X_{\lfloor \frac{n+1}{2} \rfloor:n} + X_{\lfloor \frac{n+2}{2} \rfloor:n}\right).$$

Also, with $W^*_{1:n} \leq \ldots \leq W^*_{n:n}$ the ordered values of $W^*_i = |X_i - \mathrm{Med}_n|$, $1 \leq i \leq n$,

$$\mathrm{MAD}_n = \frac{1}{2}\left(W^*_{\lfloor \frac{n+1}{2} \rfloor:n} + W^*_{\lfloor \frac{n+2}{2} \rfloor:n}\right).$$

A *modified sample MAD* is defined, for any choice of $k = 1, \ldots, n-1$, as

$$\mathrm{MAD}^{(k)}_n = \frac{1}{2}\left(W^*_{\lfloor \frac{n+k}{2} \rfloor:n} + W^*_{\lfloor \frac{n+k+1}{2} \rfloor:n}\right),$$

thus including MAD_n for $k = 1$.

The advantages of using $\mathrm{MAD}_n^{(k)}$ with $k > 1$ arise in a variety of settings involving data $\mathbb{X}_n$ in $\mathbb{R}^d$. For example, for $\mathbb{X}_n$ in "general position" (no more than d points of $\mathbb{X}_n$ in any $(d-1)$-dimensional subspace) with $n \geq d+1$ and with either $k = d$ or $k = d-1$, the uniform breakdown point of $(\mathrm{Med}_n, \mathrm{MAD}_n^{(k)})$ over all univariate projections attains an optimal value (Tyler, 1994, Gather and Hilker, 1997). Further, for data as sparse as $n \leq 2d$, the usual MAD_n is not even defined and the modification $\mathrm{MAD}_n^{(k)}$ for some $k > 1$ becomes essential, not merely an option for improving breakdown points. Also, again for $\mathbb{X}_n$ in general position, and with $n \geq 2(d-1)^2+d$ and $k = d-1$, the projection median based on $(\mathrm{Med}_n, \mathrm{MAD}_n^{(k)})$ attains the optimal breakdown point possible for any translation equivariant location estimator (Zuo, 2003). See Serfling and Mazumder (2009) for a treatment of $\mathrm{MAD}_n^{(k)}$ providing an exponential probability inequality (new even for $k = 1$) and almost sure convergence to ζ under minimal regularity conditions.

Here we develop both *strong* (i.e., almost sure) and *weak* (i.e., in probability) Bahadur representations for $\mathrm{MAD}_n^{(k)}$. These provide rates of convergence to zero for the error in approximating the estimation error by a simple weighted sum Y_n of $\widehat{F}_n(\nu)$, $\widehat{F}_n(\nu + \zeta)$, and $\widehat{F}_n(\nu - \zeta)$, with $\widehat{F}_n$ the usual empirical df. The rates establish negligibility in the senses needed for using Y_n to characterize the asymptotic behavior of $\mathrm{MAD}_n^{(k)} - \zeta$ for purposes of practical application.

Our strong Bahadur representation for $\mathrm{MAD}_n^{(k)}$ (Theorem 1), the first such result in the literature even for $k = 1$, yields the law of the iterated logarithm for that statistic (Corollary 4). It also makes possible developments such as robust sequential nonparametric confidence intervals for the MAD and studies of the almost sure properties of randomly trimmed means based on the MAD.

The weak version (Theorem 2) provides, among other applications, simplification of the derivations of asymptotic joint distributions of $\mathrm{MAD}_n^{(k)}$ with Med_n or other statistics of interest. The regularity conditions imposed in all previous weak versions (Hall and Welsh, 1985, Welsh, 1986, van der Vaart, 1998, and Chen and Giné, 2004), which also are all confined to the usual MAD_n, are reduced substantially. In particular, continuous differentiability and symmetry-type assumptions on F are avoided. Keeping to *minimal assumptions* is especially important, of course, in nonparametric applications.

The scope of application of our results is diverse. For example, the construction of highly robust quadratic form type outlyingness functions for multivariate data using projection pursuit (Pan, Fung, and Fang, 2000) involves vectors of (median, MAD) combinations, and in turn vectors of ratios of the form

$$\left(\frac{\boldsymbol{u}_1'\boldsymbol{x} - \mathrm{Med}\{\boldsymbol{u}_1'\boldsymbol{X}_i\}}{\mathrm{MAD}\{\boldsymbol{u}_1'\boldsymbol{X}_i\}}, \ldots, \frac{\boldsymbol{u}_J'\boldsymbol{x} - \mathrm{Med}\{\boldsymbol{u}_J'\boldsymbol{X}_i\}}{\mathrm{MAD}\{\boldsymbol{u}_J'\boldsymbol{X}_i\}}\right)$$

for some finite J. For establishing asymptotic multivariate normality of sample versions of such vectors under minimal assumptions on F, our Bahadur representations provide a straightforward approach that nicely handles the mutual dependence of the components. We also mention the metrically trimmed means based on observations within intervals of form

$\text{Med}_n \pm c\,\text{MAD}_n$, which yield high-breakdown analogues of the usual quantile-based trimmed means (see Hampel, 1985, Olive, 2001, and Chen and Giné, 2004). As another example, the MAD plays a role in designing robust screening methods in genomics. The ratio of the sample standard deviation (SD) and the sample MAD provides a measure of information content for each gene in a data set involving several DNA microarray experiments for operon prediction (Sabatti *et al.*, 2002). For high-throughput screening of large-scale RNA (ribonucleic acid) interference libraries, MAD-based hit selection methods are more resistant to outliers and rescue physiologically relevant false negatives that would have been missed using SD-based methods (Chung *et al.*, 2008). In comparing models for reliable gene selection in microarray data, not only the frequency of accurate classification but also the MAD of classification accuracies is used (Davis *et al.*, 2006).

As discussed in Serfling and Mazumder (2009) for almost sure convergence and asymptotic normality, similarly a Bahadur representation for $\text{MAD}_n^{(k)}$ for arbitrary n and k is exactly the same as that derived more conveniently for an arbitrary single order statistic version and can be obtained by minor modifications of our proofs. Thus, for any fixed integers $\ell \geq 1$ and $m \geq 1$, put

$$\widehat{\nu}_n = X_{\lfloor \frac{n+\ell}{2} \rfloor : n} \tag{4}$$

and

$$\widehat{\zeta}_n = W_{\lfloor \frac{n+m}{2} \rfloor : n}, \tag{5}$$

with $W_{1:n} \leq \ldots \leq W_{n:n}$ the ordered values of $W_i = |X_i - \widehat{\nu}_n|$, $1 \leq i \leq n$. For later use we note that, corresponding to $\hat{\nu}_n$, a sample analogue estimator for the distribution G is induced via (2):

$$\widehat{G}_n(y) = \widehat{F}_n(\widehat{\nu}_n + y) - \widehat{F}_n(\widehat{\nu}_n - y-), \; y \in \mathbb{R}. \tag{6}$$

For MAD_n as given by $\widehat{\zeta}_n$, we establish strong and weak Bahadur representations, which are stated and discussed in Section 2. The proof for the strong version is developed in Section 3 and that for the weak version in Section 4.

2 Strong and Weak Bahadur Representations

For the strong and weak Bahadur representations, we will assume, respectively, the conditions (**S**) or (**W**), defined as follows.

(**S**) F is continuous in neighborhoods of $\nu \pm \zeta$ and twice differentiable at ν and $\nu \pm \zeta$, with $F'(\nu) > 0$ and $G'(\zeta) = F'(\nu - \zeta) + F'(\nu + \zeta) > 0$.

(**W**) F is continuous in neighborhoods of $\nu \pm \zeta$ and differentiable at ν and $\nu \pm \zeta$, with $F'(\nu) > 0$ and $G'(\zeta) = F'(\nu - \zeta) + F'(\nu + \zeta) > 0$.

In a wide scope of typical application settings, it is not very restrictive to assume simply

> "F is continuous in neighborhoods of $\nu \pm \zeta$ and once (resp., twice) differentiable at ν and $\nu \pm \zeta$, with positive first derivatives at these points",

which implies (**W**) (resp., (**S**)). An advantage of our approach is that we are able to avoid imposing continous differentiability assumptions on F.

As in Serfling and Mazumder (2009), we use the notation

$$\begin{aligned}\alpha &= F(\nu-\zeta)+F(\nu+\zeta),\\ \beta &= F'(\nu-\zeta)-F'(\nu+\zeta),\\ \gamma &= \beta^2-4(1-\alpha)F'(\nu).\end{aligned}$$

Our Bahadur representations for $\widehat{\zeta}_n$ each approximate the estimation error $\widehat{\zeta}_n-\zeta$ by the same weighted sum of $\widehat{F}_n(\nu+\zeta)$, $\widehat{F}_n(\nu-\zeta)$, and $\widehat{F}_n(\nu)$, plus a remainder term Δ_n, as follows:

$$\widehat{\zeta}_n-\zeta=\frac{1/2-[\widehat{F}_n(\nu+\zeta)-\widehat{F}_n(\nu-\zeta)]}{G'(\zeta)}-\frac{\beta}{G'(\zeta)}\frac{1/2-\widehat{F}_n(\nu)}{F'(\nu)}+\Delta_n=:Y_n+\Delta_n. \tag{7}$$

The second term in the approximating random variable Y_n defined by (7) disappears when $\beta=0$ (implied by symmetry of F about ν). The negligibility of Δ_n is established in the following two results.

Theorem 1 *Strong Bahadur representation. Under Assumption* (**S**) *on* F,

$$\Delta_n \overset{a.s.}{=} O(n^{-3/4}(\log n)^{3/4}),\ \ n\to\infty. \tag{8}$$

Theorem 2 *Weak Bahadur representation. Under Assumption* (**W**) *on* F,

$$\Delta_n=o_p(n^{-1/2}),\ \ n\to\infty. \tag{9}$$

In view of the structure of Y_n as an average of i.i.d. random variables, the following two results are immediate using the classical central limit theorem and law of iterated logarithm, respectively. It is easily checked that the variance of Y_n is σ^2/n, where

$$\sigma^2=\frac{1}{4[G'(\zeta)]^2}\left(1+\frac{\gamma}{[F'(\nu)]^2}\right).$$

Corollary 3 *(Limit normal distribution for MAD) Under Assumption* (**W**) *on* F,

$$n^{1/2}(\widehat{\zeta}_n-\zeta)\xrightarrow{d}N(0,\sigma^2),\ \ n\to\infty.$$

Corollary 4 *(Law of iterated logarithm for MAD) Under Assumption* (**S**) *on* F,

$$\limsup_{n\to\infty}\frac{n^{1/2}(\widehat{\zeta}_n-\zeta)}{(2\sigma^2\log\log n)^{1/2}}=1\ \ a.s.$$

Remark 5 *Extensions and further developments.* (i) The steps of our proofs of Theorems 1 and 2 lead similarly to Bahadur representations for an arbitrary pth quantile ζ_p of G. Let $\widehat{\zeta}_{pn}$ denote the sample pth quantile of the deviations $W_i = |X_i - \widehat{\nu}_n|$, $1 \leq i \leq n$. Then, for F continuous in neighborhoods of $\nu \pm \zeta_p$ and once or twice differentiable at ν and $\nu \pm \zeta_p$, with $F'(\nu) > 0$ and $G'(\zeta_p) = F'(\nu - \zeta_p) + F'(\nu + \zeta_p) > 0$, and putting $\beta_p = F'(\nu - \zeta_p) - F'(\nu + \zeta_p)$, we have

$$\widehat{\zeta}_{pn} - \zeta_p = \frac{p - [\widehat{F}_n(\nu + \zeta_p) - \widehat{F}_n(\nu - \zeta_p)]}{G'(\zeta_p)} - \frac{\beta_p}{G'(\zeta_p)} \frac{1/2 - \widehat{F}_n(\nu)}{F'(\nu)} + \Delta_{pn}, \tag{10}$$

with Δ_{pn} satisfying the same convergences as above. When $\beta_p = 0$ (implied by symmetry of F about ν), the 2nd term in the approximation disappears.

(ii) Analogous to the classical robust nonparametric confidence interval for the median ν of F based on two order statistics $X_{j_1:n}$ and $X_{j_2:n}$, we can use selected $W_{j_1:n}$ and $W_{j_2:n}$ to form a robust nonparametric confidence interval for the MAD ζ of F. Foundational asymptotic theory for such procedures stems conveniently from the approximation (10). Also, the strong version opens up the possibility of developing *robust sequential fixed-width nonparametric confidence interval procedures* for ζ, paralleling well-established procedures for ν. This will be explored in a future study.

(iii) Motivated by the treatment of Rousseeuw and Croux (1993) showing that certain pth quantiles of the ordered $|X_i - X_j|$ provide higher Gaussian efficiency than the case $p = 1/2$ without sacrifice of breakdown point, the comparative performances of $\widehat{\zeta}_{pn}$ as spread measures may be investigated for various p using (10). This topic will be explored in a future study.

(iv) Extension of the above strong and weak Bahadur representations to the regression setting and the MAD of regression residuals will be developed in a separate paper. Previously, a weak version has been given by Welsh (1986) under certain regularity conditions.

(v) Randomly trimmed means based on the MAD date to Hampel (1971, 1985) and have been studied by Chen and Giné (2004) using a weak version of (7). Our strong version opens up the possibility of developing almost sure properties of such trimmed means. □

Remark 6 *Comparisons with previous results in the literature.* (i) For the case $\widehat{\nu}_n$ = the usual Med_n and $\widehat{\zeta}_n$ = the usual MAD_n, the asymptotic normality in Corollary 3 is long known. See Huber (1981) for implicit discussion considering MAD_n as a special case of M-estimator for scale. For a more comprehensive and explicit treatment, see Koul (2002). Under (**W**), one can derive not only Corollary 3 as stated but also joint asymptotic normality of $\widehat{\nu}_n$ and $\widehat{\zeta}_n$, using methods more elementary than those of the present paper (see Serfling and Mazumder, 2009, extending Falk, 1997). Of course, Corollary 3 does not imply the assertion of Theorem 2 regarding the error term Δ_n.

(ii) The earliest result establishing (9) is due to Hall and Welsh (1985), who assume (**W**) plus additional regularity assumptions on $F'(x)$ in neighborhoods of $\nu \pm \zeta$ and also symmetry-type assumptions $F(\nu + \zeta) = 1 - F(\nu - \zeta)$ and $F'(\nu + \zeta) = F'(\nu - \zeta)$. These conditions are relaxed somewhat in Welsh (1986).

(iii) van der Vaart (1998, Lemma 21.9), assuming (**W**) plus continuous differentiability of F in neighborhoods of $\nu \pm \zeta$, establishes that the functional $T(F) = \mathrm{MAD}(F)$ is *Hadamard differentiable* at F, tangentially to the set of functions h continuous at ν and in neighborhoods of $\nu \pm \zeta$, with derivative

$$T'_F(h) = \frac{h(\nu)}{F'(\nu)} \frac{F'(\nu+\zeta) - F'(\nu-\zeta)}{F'(\nu+\zeta) + F'(\nu-\zeta)} + \frac{h(\nu+\zeta) - h(\nu-\zeta)}{F'(\nu+\zeta) + F'(\nu-\zeta)}. \tag{11}$$

The formal substitution $h(\cdot) = n^{1/2}(\widehat{F}_n(\cdot) - F(\cdot))$ in (11) yields the random variable Y_n in (7), and then asymptotic normality of the sample analogue estimator $T(\widehat{F}_n)$ of $T(F)$ follows by a standard delta method argument,

$$n^{1/2}(T(\widehat{F}_n) - T(F)) \approx T'_F(n^{1/2}(\widehat{F}_n - F)),$$

applying weak convergence of $n^{1/2}(\widehat{F}_n(\cdot) - F(\cdot))$ to the corresponding F-Brownian bridge process with sample paths continuous everywhere that F is continuous. While elegantly connecting with functional derivatives, this approach requires, however, stronger regularity assumptions than Corollary 3. Also, under conditions even weaker than (**W**), the right hand side of (11) is a *Gâteaux* differential of $T(F)$, which facilitates the same practical applications.

(iv) As a tool for treatment of randomly trimmed means, Chen and Giné (2004, Lemma 4.3) establish (9) for the usual MAD_n assuming (**W**) plus continuous differentiability of F in neighborhoods of ν and $\nu \pm \zeta$. Also, in their Lemma 5.3, they establish for MAD_n the *bootstrap analogue* of Theorem 2 above. This is given by substituting in (7) $\widehat{\zeta}_n^B - \widehat{\zeta}_n$ for $\widehat{\zeta}_n - \zeta$, $-(\widehat{F}_n^B - \widehat{F}_n)[\nu-\zeta, \nu+\zeta]$ for $1/2 - [\widehat{F}_n(\nu+\zeta) - \widehat{F}_n(\nu-\zeta)]$, and $-(\widehat{F}_n^B - \widehat{F}_n)(\nu)$ for $1/2 - \widehat{F}_n(\nu)$, where $\widehat{F}_n^B$ is the bootstrap empirical df based on sampling from $\widehat{F}_n$ and $\widehat{\zeta}_n^B$ is the corresponding estimate of $\widehat{\zeta}_n$.

(v) Writing Y_n in (7) as $Y_n = n^{-1} \sum_{i=1}^n g(X_i)$, with

$$g(x) = \frac{1/2 - \mathbf{1}\{\nu - \zeta < x \le \nu + \zeta\}}{A} - \frac{C}{A} \frac{1/2 - \mathbf{1}\{x \le \nu\}}{F'(\nu)},$$

we recognize $g(x)$ as the *influence function* of $T(F) = \mathrm{MAD}(F)$ (see Hampel *et al.*, 1986, and Rousseeuw and Croux, 1993).

(vi) Theorem 1, Corollary 4, and (10) are completely new and have no antecedents. □

3 Proof of the Strong Representation

Our approach applies a general paradigm for developing Bahadur representations given in Serfling (1980), p. 95.

Proof of Theorem 1. We first identify the structure of the remainder term in (7), by applying the above-mentioned paradigm, as follows. Define $\Delta_n^{(1)}$ by

$$\widehat{G}_n(\widehat{\zeta}_n) = \widehat{F}_n(\widehat{\nu}_n + \widehat{\zeta}_n) - \widehat{F}_n(\widehat{\nu}_n - \widehat{\zeta}_n -) = 1/2 + \Delta_n^{(1)}. \tag{12}$$

Using Taylor expansion and first-order differentiability of F at $\nu \pm \zeta$, define $\Delta_n^{(2)}$ and $\Delta_n^{(3)}$ by

$$F(\widehat{\nu}_n + \widehat{\zeta}_n) - F(\nu + \zeta) = F'(\nu + \zeta)(\widehat{\nu}_n + \widehat{\zeta}_n - \nu - \zeta) + \Delta_n^{(2)} \tag{13}$$

and

$$F(\widehat{\nu}_n - \widehat{\zeta}_n) - F(\nu - \zeta) = F'(\nu - \zeta)(\widehat{\nu}_n - \widehat{\zeta}_n - \nu + \zeta) + \Delta_n^{(3)}. \tag{14}$$

Further, define $\Delta_n^{(4)}$ and $\Delta_n^{(5)}$ by

$$F(\widehat{\nu}_n + \widehat{\zeta}_n) - F(\nu + \zeta) = \widehat{F}_n(\widehat{\nu}_n + \widehat{\zeta}_n) - \widehat{F}_n(\nu + \zeta) + \Delta_n^{(4)} \tag{15}$$

and

$$F(\widehat{\nu}_n - \widehat{\zeta}_n) - F(\nu - \zeta) = \widehat{F}_n(\widehat{\nu}_n - \widehat{\zeta}_n) - \widehat{F}_n(\nu - \zeta) + \Delta_n^{(5)}. \tag{16}$$

Combining these yields

$$\begin{aligned} & F'(\nu + \zeta)(\widehat{\nu}_n + \widehat{\zeta}_n - \nu - \zeta) - F'(\nu - \zeta)(\widehat{\nu}_n - \widehat{\zeta}_n - \nu + \zeta) \\ = \; & 1/2 - [\widehat{F}_n(\nu + \zeta) - \widehat{F}_n(\nu - \zeta)] + \Delta_n^{(1)} + \Delta_n^{(4)} - \Delta_n^{(2)} - \Delta_n^{(5)} + \Delta_n^{(3)}. \end{aligned} \tag{17}$$

On the other hand, the left hand side of (17) also may be expressed as

$$-(\widehat{\nu}_n - \nu)\beta + (\widehat{\zeta}_n - \zeta)G'(\zeta),$$

and it follows that

$$\widehat{\zeta}_n - \zeta = \frac{1/2 - [\widehat{F}_n(\nu + \zeta) - \widehat{F}_n(\nu - \zeta)]}{G'(\zeta)} + \frac{\beta}{G'(\zeta)}(\widehat{\nu}_n - \nu) + \frac{\Delta_n^*}{G'(\zeta)}, \tag{18}$$

with

$$\Delta_n^* = \Delta_n^{(1)} + \Delta_n^{(4)} - \Delta_n^{(2)} - \Delta_n^{(5)} + \Delta_n^{(3)}.$$

Finally, we use the classical Bahadur representation for $\widehat{\nu}_n$ to define $\Delta_n^{(6)}$:

$$\widehat{\nu}_n = \nu + \frac{1/2 - \widehat{F}_n(\nu)}{F'(\nu)} + \Delta_n^{(6)}. \tag{19}$$

Inserting this into (18), we arrive at (7) with

$$\Delta_n = \frac{\Delta_n^*}{G'(\zeta_p)} + \frac{\beta}{G'(\zeta_p)}\Delta_n^{(6)}.$$

The proof is now completed by establishing (8), which follows from proving

$$\Delta_n^{(i)} \overset{a.s.}{=} O(n^{-3/4}(\log n)^{3/4}), \; n \to \infty, \tag{20}$$

individually for each $i = 1, \ldots, 6$, under conditions that are included in $\mathbf{S}$. For $i = 6$, this is given by the classical result for the Bahadur representation for the median (Bahadur, 1966, or Serfling, 1980, Theorem 2.5.1). The remaining cases are established via Lemmas 7–11 below. $\square$

The first lemma, to which we will appeal several times, gives almost sure convergence of $\widehat{\nu}_n$ to ν and $\widehat{\zeta}_n$ to ζ, along with exponential probability inequalities for $\widehat{\nu}_n$ and $\widehat{\zeta}_n$. Statement (i) is classical (Serfling, 1980) and statements (ii), (iii), and (iv) are established in Serfling and Mazumder (2009), Corollary 4 and Theorem 1. Note that the uniqueness assumptions on ν and ζ imposed in the lemma are implied by the conditions $F'(\nu) > 0$ and $G'(\zeta) = F'(\nu - \zeta) + F'(\nu + \zeta) > 0$, respectively, included in both (**S**) and (**W**).

Lemma 7 *Let $\nu = F^{-1}(1/2) =$ Med(F) be the unique solution of (1) and $\zeta = G^{-1}(1/2) =$ MAD(F) the unique solution of (3). Define $\widehat{\nu}_n$ and $\widehat{\zeta}_n$ by (4) and (5), for fixed positive integers ℓ and m. Then*

(i) $\widehat{\nu}_n \xrightarrow{a.s.} \nu$, $n \to \infty$;

(ii) For every $\varepsilon > 0$,

$$P(|\widehat{\nu}_n - \nu| > \varepsilon) \leq 2e^{-2n\delta^2_{\varepsilon,n}}, \tag{21}$$

where $\delta_{\varepsilon,n}$ $(= \Delta_{\varepsilon,n}(\ell, m))$ $= \min\{a_0(\varepsilon), b_0(\varepsilon)\}$, with

$$\begin{aligned} a_0(\varepsilon) &= (F(\nu + \varepsilon/2) - (\lfloor (n+\ell)/2 \rfloor - 1)/n)^+ , \\ b_0(\varepsilon) &= \lfloor (n+\ell)/2 \rfloor / n - F(\nu - \varepsilon/2). \end{aligned}$$

Also,

(iii) $\widehat{\zeta}_n \xrightarrow{a.s.} \zeta$, $n \to \infty$;

(iv) For every $\varepsilon > 0$,

$$P(|\widehat{\zeta}_n - \zeta| > \varepsilon) \leq 6e^{-2n\Delta^2_{\varepsilon,n}}, \tag{22}$$

where $\Delta_{\varepsilon,n}$ $(= \Delta_{\varepsilon,n}(\ell, m))$ $= \min\{a_0(\varepsilon), b_0(\varepsilon), c_0(\varepsilon), d_0(\varepsilon)\}$, with a_0 and b_0 as above and

$$\begin{aligned} c_0(\varepsilon) &= (F(\nu + \zeta + \varepsilon/2) - F(\nu - \zeta - \varepsilon/2) - \lfloor (n+m)/2 \rfloor / n)^+ , \\ d_0(\varepsilon) &= \lfloor (n+m)/2 \rfloor / n - [F(\nu + \zeta - \varepsilon/2) - F(\nu - \zeta + \varepsilon/2)]. \end{aligned}$$

The following result yields (20) for $i = 1$.

Lemma 8 *Let $\nu = F^{-1}(1/2) =$ Med(F) be the unique solution of (1) and $\zeta = G^{-1}(1/2) =$ MAD(F) the unique solution of (3), and let F be continuous in neighborhoods of $\nu \pm \zeta$. Then almost surely*

$$\widehat{G}_n(\widehat{\zeta}_n) = 1/2 + O(n^{-1}),\ n \to \infty.$$

Proof. For convenience we take $\ell = m = 1$. The argument for other cases is similar. Let N_0 be the union of some neighborhoods of $\nu \pm \zeta$ in which F is continuous. Since $\widehat{\zeta}_n = W_{\lfloor (n+1)/2 \rfloor : n}$, we have

$$\widehat{G}_n(\widehat{\zeta}_n) = \lfloor (n+1)/2 \rfloor / n,\ n \to \infty,$$

unless there is a tie $W_{Q:n} = W_{\lfloor (n+1)/2 \rfloor:n}$ for some $Q > \lfloor (n+1)/2 \rfloor$. If such a tie exists, then for some $i \neq j$ we have both $X_i = \widehat{\nu}_n \pm \widehat{\zeta}_n$ for some choice of sign, and $X_j = \widehat{\nu}_n \pm \widehat{\zeta}_n$ for some choice of sign. It is easily seen that in this case we have either a tie between X_i and X_j or else a tie between $(X_i + X_j)/2$ and $\widehat{\nu}_n$ (the median observation). Now, by the strong convergences in Lemma 7 (i, iii), $\widehat{\nu}_n \pm \widehat{\zeta}_n$ belong to N_0 for all sufficiently large n, in which case the above-mentioned possible ties are precluded almost surely by the continuity of F in N_0. To complete the proof, we use $\lfloor (n+1)/2 \rfloor / n = 1/2 + O(n^{-1})$, $n \to \infty$. □

To obtain (20) for $i = 2, 3, 4$, and 5, we will need statement (i) of the following result.

Lemma 9 *Let F be differentiable at ν and $\nu \pm \zeta$, with $F'(\nu) > 0$ and $G'(\zeta) = F'(\nu - \zeta) + F'(\nu + \zeta) > 0$. Then*

(i) Almost surely

$$|(\widehat{\nu}_n + \widehat{\zeta}_n) - (\nu + \zeta)| \leq D_1 \frac{(\log n)^{1/2}}{n^{1/2}}, \text{ for all } n \text{ sufficiently large,} \tag{23}$$

and

$$|(\widehat{\nu}_n - \widehat{\zeta}_n) - (\nu - \zeta)| \leq D_1 \frac{(\log n)^{1/2}}{n^{1/2}}, \text{ for all } n \text{ sufficiently large,} \tag{24}$$

where

$$D_1 = \min\{8/F'(\nu),\ 8/G'(\zeta)\}.$$

(ii) Also,

$$|(\widehat{\nu}_n + \widehat{\zeta}_n) - (\nu + \zeta)| = O_p(n^{-1/2}),\ n \to \infty, \tag{25}$$

and

$$|(\widehat{\nu}_n - \widehat{\zeta}_n) - (\nu - \zeta)| = O_p(n^{-1/2}),\ n \to \infty. \tag{26}$$

Proof. (i) The assumptions on F imply the conditions of Lemma 7 and together the inequalities (ii) and (iv) of that lemma yield, for any $\varepsilon > 0$,

$$P(|(\widehat{\nu}_n + \widehat{\zeta}_n) - (\nu + \zeta)| > \varepsilon) \leq 8e^{-2n\Delta^2_{\varepsilon/2,n}}. \tag{27}$$

Put

$$\varepsilon_n = D_1 \frac{(\log n)^{1/2}}{n^{1/2}},$$

Now, since $F(\nu) = 1/2$ (implied by the above condition), we have

$$\begin{aligned} a_0(\varepsilon_n/2) &= F(\nu + \varepsilon_n/4) - (\lfloor (n+\ell)/2 \rfloor - 1)/n \\ &= F(\nu + \varepsilon_n/4) - 1/2 + O(n^{-1}) \\ &= F(\nu + \varepsilon_n/4) - F(\nu) + O(n^{-1}) \\ &= \frac{F'(\nu)}{4}\varepsilon_n + o(\varepsilon_n) + O(n^{-1}) \\ &> \frac{(\log n)^{1/2}}{n^{1/2}}, \text{ for all } n \text{ sufficiently large,} \end{aligned}$$

and similarly

$$b_0(\varepsilon_n/2) > \frac{(\log n)^{1/2}}{n^{1/2}}, \text{ for all } n \text{ sufficiently large.}$$

By similar arguments using $F(\nu+\zeta) - F(\nu-\zeta) = 1/2$ (implied by the above condition), we also obtain

$$c_0(\varepsilon_n/2) > \frac{(\log n)^{1/2}}{n^{1/2}}, \text{ for all } n \text{ sufficiently large,}$$

and

$$d_0(\varepsilon_n/2) > \frac{(\log n)^{1/2}}{n^{1/2}}, \text{ for all } n \text{ sufficiently large.}$$

Hence

$$2n\Delta^2_{\varepsilon_n/2,n} \geq 2\log n, \text{ for all } n \text{ sufficiently large.}$$

Now using (27) with a Borel-Cantelli argument we obtain (23). A similar argument yields (24).

(ii) Let M be any fixed (large) number, and put

$$\tilde{\varepsilon}_n = D_1 \frac{M}{n^{1/2}}.$$

By similar steps as above with $\tilde{\varepsilon}_n$ in place of ε_n, we obtain

$$2n\Delta^2_{\tilde{\varepsilon}_n/2,n} \geq 2M^2, \text{ for all } n \text{ sufficiently large,}$$

whence

$$P(n^{1/2}|(\widehat{\nu}_n + \widehat{\zeta}_n) - (\nu+\zeta)| > D_1 M) \leq e^{-2M^2}, \text{ for all } n \text{ sufficiently large.}$$

Since M may be arbitrarily large, (25) follows, and a similar argument yields (26). □

Lemma 10 *Let F be differentiable at ν and twice differentiable at $\nu \pm \zeta$, with $F'(\nu) > 0$ and $G'(\zeta) = F'(\nu-\zeta) + F'(\nu+\zeta) > 0$. Then, for $i = 2$ and 3, almost surely*

$$\Delta_n^{(i)} = O\left(\frac{\log n}{n}\right), \ n \to \infty.$$

Proof. By the second order differentiability of F at $\nu \pm \zeta$, and using Young's form of Taylor's Theorem (Serfling, 1980, Theorem 1.12.1C) and the almost sure convergence of $\widehat{\nu}_n + \widehat{\zeta}_n$ to $\nu + \zeta$ (as per Lemma 7 (i, iii)), we obtain almost surely

$$\Delta_n^{(2)} = O\left(\left((\widehat{\nu}_n + \widehat{\zeta}_n) - (\nu+\zeta)\right)^2\right), \ n \to \infty.$$

Applying Lemma 9 (i), (23) gives the result for $i = 2$. Similar steps with (24) yield the result for $i = 3$. □

Lemma 11 *Let F be differentiable at ν and twice differentiable at $\nu \pm \zeta$, with $F'(\nu) > 0$ and $G'(\zeta) = F'(\nu - \zeta) + F'(\nu + \zeta) > 0$. Then (20) holds for $i = 4$ and 5.*

PROOF. Let us express $\nu + \zeta$ as a pth quantile of F: $\nu + \zeta = F^{-1}(p) = \theta_p$, say. Put $x_n = (\widehat{\nu}_n + \widehat{\zeta}_n) - (\nu + \zeta) = (\widehat{\nu}_n + \widehat{\zeta}_n) - \theta_p$. Then $(\widehat{\nu}_n + \widehat{\zeta}_n) = \theta_p + x_n$. By Lemma 9 (i), we have almost surely

$$|x_n| \le \varepsilon_n = D_1(\log n)^{1/2}/n^{1/2},$$

for all sufficiently large n, whence

$$\Delta_n^{(4)} \le \sup_{|x| \le \varepsilon_n} \left| [\widehat{F}_n(\theta_p + x) - \widehat{F}_n(\xi_p)] - [F(\xi_p + x) - F(\xi_p)] \right| =: H_{pn}.$$

Now H_{pn} is the random variable shown in a key lemma of Bahadur (see Serfling, 1980, Lemma 2.5.4E) to satisfy almost surely

$$H_{pn} = O(n^{-3/4}(\log n)^{3/4}),\ n \to \infty,$$

using the twice differentiability of F at θ_p. It can be checked that we do not require $F'(\theta_p) > 0$ for this step. Thus (20) holds for $i = 4$, and similarly the case $i = 5$ is proved. □

4 Proof of the Weak Representation

In the present proof, we wish to show (9) instead of (8). We again treat the remainder term in (7) using the structure developed in the proof of Theorem 1. Thus, it will suffice to establish

$$\Delta_n^{(i)} = o_p(n^{-1/2}),\ n \to \infty, \tag{28}$$

individually for each $i = 1, \dots, 6$, under conditions that are included in **W**. The case $i = 1$ again follows immediately from Lemma 8. For $i = 6$, (28) is given by the classical weak Bahadur representation for the median (Ghosh, 1971). The remaining cases are established via Lemmas 12–14 below. □

The following analogue of Lemma 10 takes care of the cases $i = 2$ and 3.

Lemma 12 *Let F be differentiable at ν and $\nu \pm \zeta$, with $F'(\nu) > 0$ and $G'(\zeta) = F'(\nu - \zeta) + F'(\nu + \zeta) > 0$. Then, for $i = 2$ and 3,*

$$\Delta_n^{(i)} = o_p(n^{-1/2}),\ n \to \infty.$$

PROOF. By the first order differentiability of F at $\nu \pm \zeta$, and using Young's form of Taylor's Theorem (Serfling, 1980, Theorem 1.12.1C) and the almost sure convergence of $\widehat{\nu}_n + \widehat{\zeta}_n$ to $\nu + \zeta$ (as per Lemma 7 (i, iii)), we obtain almost surely

$$\Delta_n^{(2)} = o\left(\left| (\widehat{\nu}_n + \widehat{\zeta}_n) - (\nu + \zeta) \right| \right),\ n \to \infty.$$

Applying Lemma 9 (ii), (25) gives the result for $i = 2$. Similar steps with (26) yield the result for $i = 3$. □

For the cases $i = 4$ and 5, we will apply the following basic lemma.

Lemma 13 *(Ghosh, 1971) Let* $\{U_n\}$ *and* $\{V_n\}$ *be sequences of random variables on some probability space* $(\Omega, \mathcal{A}, P)$*. Suppose that*

(a) $V_n = O_p(1)$*,* $n \to \infty$*,*

and

(b) For all t *and all* $\varepsilon > 0$*,*

$$\begin{aligned} \lim_{n\to\infty} P(U_n \geq t+\varepsilon, V_n \leq t) &= 0 \\ \lim_{n\to\infty} P(U_n \leq t, V_n \geq t+\varepsilon) &= 0. \end{aligned}$$

Then $U_n - V_n = o_p(1)$*,* $n \to \infty$*.*

Lemma 14 *Let* F *be continuous in neighborhoods of* $\nu \pm \zeta$ *and differentiable at* ν *and* $\nu \pm \zeta$*, with* $F'(\nu) > 0$ *and* $G'(\zeta) = F'(\nu-\zeta) + F'(\nu+\zeta) > 0$*. Then (28) holds for* $i = 4$ *and 5.*

Proof. Put

$$U_n = n^{1/2}[\widehat{F}_n(\widehat{\nu}_n + \widehat{\zeta}_n) - \widehat{F}_n(\nu+\zeta)]$$

and

$$V_n = n^{1/2}[F(\widehat{\nu}_n + \widehat{\zeta}_n) - F(\nu+\zeta)]$$

Then $n^{1/2}\Delta_n^{(4)} = V_n - U_n$ and for (28) it suffices to show that $U_n - V_n = o_p(1)$. Now, by the assumptions of the lemma, and using the same argument as in the proof of Lemma 12, we have

$$\begin{aligned} F(\widehat{\nu}_n + \widehat{\zeta}_n) - F(\nu+\zeta) &= O\left(\left|(\widehat{\nu}_n + \widehat{\zeta}_n) - (\nu+\zeta)\right|\right) \text{ a.s., } n \to \infty, \\ &= O_p(n^{-1/2}),\ n \to \infty, \end{aligned}$$

and thus V_n satisfies (a) of Lemma 13.

Now let $\varepsilon > 0$ and put

$$\alpha = \lim_{t\uparrow 0} F^{-1}(F(\nu+\zeta) + t/\sqrt{n})), \quad \beta = \lim_{t\downarrow 0} F^{-1}(F(\nu+\zeta) + t/\sqrt{n})).$$

Consider the case $t > 0$. If $F'(\nu+\zeta) > 0$, then $F^{-1}(F(\nu+\zeta)+t/\sqrt{n}) > \nu+\zeta$ and also $\beta = \nu+\zeta$, i.e., F^{-1} is continuous at $F(\nu+\zeta)$. Then, using $F(x) < p$ iff $x < F^{-1}(p)$, we have

$$\begin{aligned} \{V_n \leq t\} &= \left\{F(\widehat{\nu}_n + \widehat{\zeta}_n) - F(\nu+\zeta) \leq t/\sqrt{n}\right\} \\ &\subset \left\{F(\widehat{\nu}_n + \widehat{\zeta}_n) < F(\nu+\zeta) + t/\sqrt{n}\right\} \\ &= \left\{\widehat{\nu}_n + \widehat{\zeta}_n < F^{-1}\left(F(\nu+\zeta) + t/\sqrt{n}\right)\right\} \\ &\subset \left\{\widehat{F}_n(\widehat{\nu}_n + \widehat{\zeta}_n) \leq \widehat{F}_n\left(F^{-1}\left(F(\nu+\zeta) + t/\sqrt{n}\right)\right)\right\}. \end{aligned}$$

Then

$$P(U_n \geq t+\varepsilon, V_n \leq t) \leq P\left(Z_n \geq \frac{t+\varepsilon}{\sqrt{n}}\right), \tag{29}$$

where

$$Z_n = \widehat{F}_n\left(F^{-1}\left(F(\nu+\zeta)+t/\sqrt{n}\right)\right) - \widehat{F}_n(\nu+\zeta).$$

Since $F^{-1}(F(\nu+\zeta)+t/\sqrt{n}) > \nu+\zeta$, the random variable nZ_n is Binomial(n, p_n) with

$$p_n = F\left(F^{-1}\left(F(\nu+\zeta)+t/\sqrt{n}\right)\right) - F(\nu+\zeta).$$

Also, since $F^{-1}\left(F(\nu+\zeta)+t/\sqrt{n}\right) \to \nu+\zeta$ and F is continuous in a neighborhood of $\nu+\zeta$, we have, for n sufficiently large,

$$\begin{aligned} p_n &= F\left(F^{-1}\left(F(\nu+\zeta)+t/\sqrt{n}\right)\right) - F(\nu+\zeta) \\ &= t/\sqrt{n}, \end{aligned}$$

where we have used $F(F^{-1}(u)) = u$ if F is continuous at $F^{-1}(u)$. Hence, for all n sufficiently large,

$$\begin{aligned} P\left(Z_n \geq \frac{t+\varepsilon}{\sqrt{n}}\right) &\leq P\left(Z_n - p_n \geq \frac{t+\varepsilon}{\sqrt{n}} - p_n\right) \\ &\leq P\left(|Z_n - p_n| \geq \frac{\varepsilon}{\sqrt{n}}\right) \\ &\leq \frac{p_n(1-p_n)}{\varepsilon^2} \to 0,\ n \to \infty, \end{aligned}$$

and, returning to (29), we see that the first statement in (b) of Lemma 13 is established for $t > 0$ and $F'(\nu+\zeta) > 0$. The same steps carry through if $F'(\nu+\zeta) = 0$ with $\beta = \nu+\zeta$.

Suppose, however, that $t > 0$ and $F'(\nu+\zeta) = 0$ with $\beta > \nu+\zeta$. Let θ be any point in the open interval $(\nu+\zeta, \beta)$. By the strong convergence of $\widehat{\nu}_n + \widehat{\zeta}_n$ to $\nu+\zeta$, we have $P(\widehat{\nu}_n + \widehat{\zeta}_n > \theta) \to 0$ and

$$P(U_n \geq t+\varepsilon, V_n \leq t) = P(U_n \geq t+\varepsilon, V_n \leq t, \widehat{\nu}_n + \widehat{\zeta}_n \leq \theta) + o(1),\ n \to \infty.$$

Now

$$\begin{aligned} \{V_n \leq t, \widehat{\nu}_n + \widehat{\zeta}_n \leq \theta\} &= \left\{F(\widehat{\nu}_n + \widehat{\zeta}_n) - F(\nu+\zeta) \leq t/\sqrt{n}, \widehat{\nu}_n + \widehat{\zeta}_n \leq \theta\right\} \\ &\subset \left\{\widehat{\nu}_n + \widehat{\zeta}_n < F^{-1}\left(F(\nu+\zeta)+t/\sqrt{n}\right), \widehat{\nu}_n + \widehat{\zeta}_n \leq \theta\right\} \\ &\subset \left\{\widehat{F}_n(\widehat{\nu}_n + \widehat{\zeta}_n) \leq \widehat{F}_n\left(\min\left\{\theta, F^{-1}\left(F(\nu+\zeta)+t/\sqrt{n}\right)\right\}\right)\right\}. \end{aligned}$$

Then

$$P(U_n \geq t+\varepsilon, V_n \leq t, \widehat{\nu}_n + \widehat{\zeta}_n \leq \theta) \leq P\left(Z_n \geq \frac{t+\varepsilon}{\sqrt{n}}\right),$$

where

$$Z_n = \widehat{F}_n\left(\min\left\{\theta, F^{-1}\left(F(\nu+\zeta)+t/\sqrt{n}\right)\right\}\right) - \widehat{F}_n(\nu+\zeta).$$

But, by definition of β and θ, almost surely there are no observations in the interval $[\nu+\zeta,\theta]$, so that almost surely $Z_n = 0$ and hence

$$P\left(Z_n \geq \frac{t+\varepsilon}{\sqrt{n}}\right) = 0$$

in the case under consideration. Thus we now have established the first statement in (b) of Lemma 13 for $t > 0$. A similar approach takes care of the case $t \leq 0$. Finally, similar steps establish the second statement in (b) of Lemma 13.

Thus we have proved $n^{1/2}\Delta_n^{(4)} = V_n - U_n = o_p(1)$ and so (28) holds for $i = 4$. The case $i = 5$ is obtained similarly. □

Acknowledgment

Very helpful comments and suggestions by G. L. Thompson and a reviewer have led to improvements in the paper and are greatly appreciated. Support under National Science Foundation Grants DMS-0103698 and DMS-0805786 and NSA Grant H98230-08-1-0106 is gratefully acknowledged.

References

[1] Bahadur, R. R. (1966). A note on quantiles in large samples. *Annals of Mathematical Statistics* **37** 577–580.

[2] Chen, Z. and Giné, E. (2004). Another approach to asymptotics and bootstrap of randomly trimmed means. *Annals of the Institute of Statistical Mathematics* **56** 771–790.

[3] Chung, N., Zhang, X. D., *et al.* (2008). Median absolute deviation to improve hit selection for genome-scale RNAi screens. *Journal of Biomolecular Screening* **13** 149–158.

[4] Davis, C., Gerick, F., *et al.* (2006). Reliable gene signatures for microarray classification: assessment of stability and performance. *Bioinformatics* **22** 2356–2363.

[5] Falk, M. (1997). Asymptotic independence of median and MAD. *Statistics and Probability Letters* **34** 341–345.

[6] Gather, U. and Hilker, T. (1997). A note on Tyler's modification of the MAD for the Stahel-Donoho estimator. *Annals of Statistics* **25** 2024-2026.

[7] Ghosh, J. K. (1971). A new proof of the Bahadur representation of sample quantiles and an application. *Annals of Mathematical Statistics* **42** 1957–1961.

[8] Hall, P. and Welsh, A. H. (1985). Limit theorems for the median deviation. *Annals of Institute of Statistical Mathematics* **37** 27–36.

[9] Hampel, F. R. (1971) A general qualitative definition of robustness. *Annals of Mathematical Statistics* **42** 1887–1896.

[10] Hampel, F. R. (1985) The breakdown points of the mean combined with some rejection rules. *Technometrics* **27** 95–107.

[11] Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986). *Robust Statistics: The Approach Based on Influence Functions*. Wiley, New York.

[12] Huber, P. J. (1981). *Robust Statistics*. Wiley, New York.

[13] Koul, Hira L. (2002). *Weighted Empirical Processes in Dynamic Nonlinear Models.* [Second edition of *Weighted Empiricals and Linear Models*, IMS Lecture Notes and Monograph Series, Hayward, CA, 1992] *Lecture Notes in Statistics* **166**, Springer.

[14] Olive, D. J. (2001). High breakdown analogs of the trimmed mean. *Statistics and Probability Letters* **51** 87–92.

[15] Pan, J.-X., Fung, W.-K., and Fang, K.-T. (2000). Multiple outlier detection in multivariate data using projection pursuit techniques. *Journal of Statistical Planning and Inference* **83** 153–167.

[16] Rousseeuw, P. J. and Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association* **88** 1273–1283.

[17] Sabatti, C., Rohlin, L., Oh, M.-K., and Liao, J. C. (2002). Co-expression pattern from DNA microarray experiments as a tool for operon prediction. *Nucleic Acids Research* **30** 2886–2893.

[18] Serfling, R. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons, New York.

[19] Serfling, R. and Mazumder, S. (2009). Exponential probability inequality and convergence results for the median absolute deviation and its modifications *Statistics and Probability Letters*, in press.

[20] Tyler, D. E. (1994). Finite sample breakdown points of projection based multivariate location and scatter statistics. *Annals of Statistics* **22** 1024–1044.

[21] van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press.

[22] Welsh, A. H. (1986). Bahadur representations for robust scale estimators based on regression residuals. *Annals of Statistics* **14** 1246–1251.

[23] Zuo, Y. (2003). Projection-based depth functions and associated medians. *Annals of Statistics* **31** 1460–1490.