Ann. Inst. Statist. Math.
Vol. 47, No. 2, 267–271 (1995)

ESTIMATING THE ASYMPTOTIC DISPERSION OF THE L_1 MEDIAN

ARUP BOSE*

Theoretical Statistics and Mathematics Unit, Indian Statistical Institute, 203, B.T. Road, Calcutta-700 035, India

(Received February 21, 1994; revised July 27, 1994)

Abstract. A simple estimate of the asymptotic dispersion matrix of the L_1 median is proposed and its rate of convergence is studied.

Key words and phrases: L_1 median, asymptotic dispersion, consistent estimate, rate of convergence, strong law of large numbers, central limit theorem, law of iterated logarithm.

1. Introduction

Let $X_1, X_2, \ldots, X_n$ be n i.i.d. observations on a distribution F in R^d, $d \geq 2$. Medians for multivariate distributions have been defined in many ways in the literature. See Small (1990) for a survey in this area. The L_1 median of F is the value of θ which satisfies

$$E_F(|X-\theta| - |X|) = \inf_{\phi \in R_d} E_F(|X-\phi| - |X|).$$

Let F_n be the empirical distribution function of $(X_1, \ldots, X_n)$. Then a natural estimate of θ is the corresponding sample analogue θ_n which satisfies

$$\sum_{\alpha=1}^{n} |X_\alpha - \theta_n| = \inf_{\phi \in R^d} \sum_{\alpha=1}^{n} |X_\alpha - \phi|. \tag{1.1}$$

Under the assumption of boundedness of the density (Assumption B of Section 2), $n^{1/2}(\theta_n - \theta)$ has an asymptotic $N(0, D(\theta))$ distribution. The exact form of $D(\theta)$ is given in Section 2.

The problem of estimation of $D(\theta)$ was first discussed in Bose and Chaudhuri (1993). They exhibited an estimate D_n which, under the assumption of boundedness of the density of X_1 on every compact set (see Assumption B later) satisfies

$$D_n - D(\theta) = \begin{cases} O_p(n^{-1/2}) & \text{if } d \geq 3 \\ O_p(n^{-\delta}) & \text{for any } \delta < 1/2 \text{ if } d = 2. \end{cases} \tag{1.2}$$

* Research supported in part by a grant of the Indian Statistical Institute.

Note that these bounds are probability bounds and no almost sure results were given. This estimate is computed by splitting the data into two groups which are used to estimate two different components of $D(\theta)$ and is thus likely to be inefficient. It was claimed in Bose and Chaudhuri (1993) that there are serious difficulties in establishing asymptotic bounds, even in probability, for the simple plug in estimator.

In the next section we introduce the natural plug in estimators and show that they have excellent asymptotic properties provided certain conditions are satisfied. In particular, we show that irrespective of the dimension $d \geq 2$,

(a) $D_n - D(\theta) = O_P(n^{-1/2})$ if $E|X_1 - \theta|^{-2} < \infty$.

(b) $D_n - D(\theta) = O(n^{-1/2}(\log\log n)^{1/2})$ almost surely if

$$\sum_{i=1}^{\infty} P\left\{|X_1 - \theta|^{-2} \geq \frac{\epsilon i}{\log\log i}\right\} < \infty \qquad \text{for any small } \epsilon.$$

The condition of (b) (and hence of (a)) holds for dimension $d \geq 3$ under Assumption B. Thus in this case we have the sharpest possible rates. For $d = 2$, these conditions do not hold solely under Assumption B. Our method of proof also shows that under finiteness of higher inverse moments, D_n is asymptotically normal. Unless the above conditions are satisfied the plug in estimator perhaps does not have any good asymptotic properties. This is suggested by the method of proof that we have employed. These issues are discussed in more details in the Remarks following the proof of the main results.

Assumption B has been used by Chaudhuri (1992) for establishing various properties of θ_n. Our investigation seems to indicate that this assumption has a bearing on such results for the median only through the fact that it guarantees the existence of these inverse moments. It is plausible that such properties of the L_1 median hold true solely under such assumptions. This issue will be explored in a separate paper.

2. Results and discussion

For any vector $x \in R^d$, define the vector U and the matrix Q as,

$$U(x) = \begin{cases} \dfrac{x}{|x|} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0, \end{cases}$$

$$Q(x) = \begin{cases} \dfrac{I}{|x|} - \dfrac{xx^T}{|x|^3} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

$$= Q_1(x) - Q_2(x), \qquad \text{say.}$$

Then the matrix $D(\theta)$ (whenever it exists) may be written as

$$D(\theta) = A^{-1}BA^{-1}$$

where

$$A = E_F Q(X_1 - \theta), \qquad B = E_F[U(X_1 - \theta)U(X_1 - \theta)^T].$$

Note that B is always defined and A is defined under the following

ASSUMPTION A. $E_F[|X_1 - \theta|^{-1}] < \infty$.

The natural plug in estimators A_n and B_n are then defined by

$$A_n = n^{-1} \sum_{\alpha=1}^{n} Q(X_\alpha - \theta_n), \tag{2.1}$$

$$B_n = n^{-1} \sum_{\alpha=1}^{n} U[(X_\alpha - \theta_n]U[(X_\alpha - \theta_n)^T]. \tag{2.2}$$

We now introduce an assumption which has been used by Chaudhuri (1992) to study the properties of θ_n.

ASSUMPTION B. X_1 has a density f which is bounded on every compact subset of R^d.

By using the fact that $d \geq 2$, it is easy to see that Assumption B implies Assumption A. Under Assumption B, the following representation for θ_n follows from Theorem 3.2 of Chaudhuri (1992).

$$\theta_n - \theta = A^{-1} n^{-1} \sum_{\alpha=1}^{n} U(X_\alpha - \theta) + R_n \tag{2.3}$$

where

$$R_n = O(n^{-1+\epsilon}) \qquad \text{a.s. for any} \quad \epsilon > 0.$$

From this representation and the CLT and LIL we have (under Assumption B),

$$\theta_n - \theta = O_P(n^{-1/2}) \tag{2.4}$$

and

$$\theta_n - \theta = O(n^{-1/2}(\log\log n)^{1/2}) \qquad \text{a.s.} \tag{2.5}$$

The relations (2.4) and (2.5) motivate the following results on the asymptotic properties of the plug in estimators. In the following theorems the estimator θ_n which is used to define A_n and B_n need not necessarily be that defined by (1.1).

THEOREM 2.1. *Suppose Assumption* A *holds. For any estimator* θ_n,
(a) *If* (2.4) *holds then* $B_n - B = O_P(n^{-1/2})$.
(b) *If* (2.5) *holds then* $B_n - B = n^{-1/2}(\log\log n)^{1/2}$ *almost surely.*

The behaviour of the function Q is quite different from that of U. The above rates are achievable by A_n only under further restrictions. We introduce the following assumptions.

ASSUMPTION C. $E_F[|X_1 - \theta|^{-2}] < \infty$.

ASSUMPTION D. $\sum_{i=1}^{\infty} P\{|X_1 - \theta|^{-2} \geq \frac{\epsilon i}{\log\log i}\} < \infty$ for any small ϵ.

THEOREM 2.2. (a) *If Assumption* C *and* (2.4) *hold then* $A_n - A = O_P(n^{-1/2})$.

(b) *If Assumption* D *and* (2.5) *hold then* $A_n - A = O(n^{-1/2}(\log\log n)^{1/2})$ *a.s.*

We defer a discussion of our results till the end of the proofs.

PROOF OF THEOREM 2.1. Define $B_n(\theta)$ as in (2.2) with θ_n replaced by θ. Note that $B_n(\theta) - B$ is a mean of i.i.d. *bounded* random variables. Hence

$$B_n(\theta) - B = O(n^{-1/2}(\log\log n)^{-1/2}) \quad \text{a.s.,} \tag{2.6}$$

$$B_n(\theta) - B = O_P(n^{-1/2}). \tag{2.7}$$

On the other hand, from the inequality $|U(x-\phi_1) - U(x-\phi_2)| \leq |\phi_1 - \phi_2| \min\{|x-\phi_1|^{-1}, |x-\phi_2|^{-1}\}$, it follows that

$$|B_n(\theta) - B_n| \leq K|\theta_n - \theta| n^{-1} \sum_{\alpha=1}^{n} |X_\alpha - \theta|^{-1}. \tag{2.8}$$

Using Assumption A, the second factor is bounded by the strong law of large numbers. The theorem then follows from the relations (2.6), (2.7) and (2.8).

PROOF OF THEOREM 2.2. Define $A_n(\theta)$ as in (2.1) with θ_n replaced by θ. Then $A_n(\theta) - A$ is the average of zero mean i.i.d. variables with finite second moment. Hence by the CLT and LIL,

$$A_n(\theta) - A = O(n^{-1/2}(\log\log n)^{1/2}) \quad \text{a.s.} \tag{2.9}$$

and

$$A_n(\theta) - A = O_P(n^{-1/2}). \tag{2.10}$$

Note that $|Q_1(x-\phi_1) - Q_1(x-\phi_2)| \leq 2|\phi_1 - \phi_2| \max\{|x-\phi_1|^{-2}, |x-\phi_2|^{-2}\}$ and by a simple manipulation, the same bound holds for Q_2 with 2 replaced by a larger constant. Hence for a constant K,

$$\begin{aligned}|A_n - A_n(\theta)| &\leq K|\theta_n - \theta| n^{-1} \max\left\{\sum_{\alpha=1}^{n} |X_\alpha - \theta_n|^{-2}, \sum_{\alpha=1}^{n} |X_\alpha - \theta|^{-2}\right\} \\ &= K|\theta_n - \theta| \max(T_{1n}, T_{2n}) \quad \text{say.}\end{aligned}$$

By the strong law of large numbers, T_{2n} is bounded almost surely. To tackle the other term, define for sufficiently large K_1, $N = \{|\theta_n - \theta| < K_1 n^{-1/2}\}$ and for sufficiently small K_2, $Y_i - \theta = (X_i - \theta) I(|X_i - \theta|^{-2} \geq K_2 i)$. Observe that

$\sum_{i=1}^{\infty} P(Y_i - \theta \neq X_i - \theta) < \infty$ since $E_F|X-\theta|^{-2} < \infty$. Hence T_{1n} is bounded a.s. if $T_{1n}^* = n^{-1}\sum_{\alpha=1}^{n}|Y_\alpha - \theta_n|^{-2}$ is so.

On the other hand it is easy to see if $|u| \leq K_1 n^{-1/2}$ and K_2 is such that $4K_1K_2^{1/2} \leq 1$ then,

$$|Y_\alpha - \theta + u|^2 \geq |Y_\alpha - \theta|^2 - 2|Y_\alpha - \theta|K_1 n^{-1/2} \geq 2^{-1}|Y_\alpha - \theta|^2.$$

Hence on the set N, we have almost surely for large n,

$$T_{1n}^* \leq Kn^{-1}\sum_{\alpha=1}^{n}|X_\alpha - \theta|^{-2} \quad \text{which is bounded a.s.} \tag{2.11}$$

Part (a) of the theorem follows by combining (2.4), (2.10) and (2.11).

To prove the second part, redefine $N = \{|\theta_n - \theta| \leq K_1 n^{-1/2}(\log\log n)^{-1/2}\}$ and $Y_i - \theta = (X_i - \theta)I(|X_i - \theta|^{-2} \geq K_2 i(\log\log i)^{-1})$ where K_1 is sufficiently large and K_2 is sufficiently small and follow the above argument.

Remarks. (a) It is easily checked that if $d \geq 3$ Assumption B implies Assumption D (and hence Assumption C too). Hence for $d \geq 3$, the best possible probability and almost sure rates hold for the estimator defined in (1.1) under Chaudhuri's condition.

(b) For $d = 2$ Assumption B guarantees $E|X - \theta|^{-(2-\epsilon)}$ for any $\epsilon > 0$. This may tempt one to believe that even though Theorem 2.2 is not applicable, perhaps a slower rate is achievable. That this is not the case is clear from a careful scrutiny of the proof. One needs to "kill" the maximum fluctuation in θ_n and since this is of the order $n^{-1/2}$ in probability and $n^{-1/2}(\log\log n)^{1/2}$ almost surely, appropriate truncation levels are those given in the theorem. For $d = 2$, the estimates of Bose and Chaudhuri (1993) are available which achieves the rate in $O_p(n^{-1/2+\epsilon})$ for any $\epsilon > 0$.

(c) From the proof of the Theorems it is also clear that with extra (inverse) moment conditions our estimates will be asymptotic normal (for θ_n defined by (1.1)). One simply would use a Taylor expansion, and tackle the remainder as in the proofs of the Theorems.

(d) Our approach shows that the inverse moment conditions $E_F[|X_1 - \theta|^{-1}] < \infty$ and $E_F[|X_1 - \theta|^{-2}] < \infty$ are crucial for the plug in estimators to work. It is plausible that Chaudhuri's representation remains true solely under Assumption C and the boundedness of the density as such is not needed. However most common distributions do satisfy Assumption B.

Acknowledgements

The author is thankful to Dr. Ratan Dasgupta for his comments. He is grateful to the referee for his prompt and constructive comments.

References

Bose, A. and Chaudhuri, P. (1993). On the dispersion of multivariate median, *Ann. Inst. Statist. Math.*, **45**(3), 541–550.

Chaudhuri, P. (1992). Multivariate location estimation using extension of R-estimates through U-statistics type approach, *Ann. Statist.*, **20**, 897–916.

Small, C. G. (1990). A survey of multidimensional medians, *Internat. Statist. Rev.*, **58**, 263–277.