ON ALIGNED M-TESTS IN LINEAR MODELS

by

Pranab Kumar Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1346

ON ALIGNED M-TESTS IN LINEAR MODELS*

by

Pranab Kumar Sen

University of North Carolina, Chapel Hill

For a simple regression model, some M-test based on an alignment procedure are considered and their asymptotic properties are studied. The theory is extended to subhypothesis testing in some linear models.

1. Introduction. Consider the simple regression model

$$X_i = \theta + \beta c_i + e_i \ , \quad i=1,\ldots,n \ , \tag{1.1}$$

where the X_i are observable random variables (r.v.), θ (intercept) and β (slope) are unknown parameters, the c_i are known regression constants and the errors e_i are independent and identically distributed (i.i.d.)r.v. with continuous, unknown distribution function (d.f.) F, defined on the real line R. It is desired to test for

$$H_0: \ \beta = 0 \ \text{ against } \ H_1: \ \beta \neq (\text{or} > \text{or} <) \ 0, \tag{1.2}$$

treating θ as a nuisance parameter.

AMS Subject Classifications: 62F35, 62J99

Key Words & Phrases: Alignment, asymptotic normality, asymptotic stochastic linearity, linear models, M-tests, Pitman-efficiency, robustness

*Work partially supported by the National Heart, Lung and Blood Institute, Contract NIH-NHLBI-71-2243-L from the National Institutes of Health.

A variety of rank tests for (1.2) is available in the literature [viz., Hájek and Šidák (1967)]. Since the ranks are translation-invariant, the nuisance parameter θ is of no concern for these tests. The situation with the tests based on M-estimators is some what different. One way of testing for (1.2) is to estimate simultaneously θ, β and then to base the test on the estimate of β and its estimated standard error. Alternatively, as in Schreder and Hettmansperger (1980), one may employ some robust likelihood ratio type criteris. In either case, to apply the asymptotic theory, one needs to estimate $E\psi^2$ and $E\psi'$, where ψ is the score function (presumed to be absolutely continuous) generating the estimates. Actually, in Schrader and Hettmansperger (1980), following Huber (1973), it has been assumed that ψ has bounded derivatives (at least up to the second order). In addition, for testing H_0 against one-sided alternatives, these likelihood-ratio type tests may not be fully efficient. In recent developments on the asymptotic theory of M-estimators, Jurečková (1977, 1980), Yohai and Maronna (1979), Jurečková and Sen (1981) and others have eliminated the boundedness conditions on the derivatives of ψ and incorporated possible (finitely many) jump-discontinuities of ψ. This should also be the case for the dual testing problem.

The object of the present investigation is to focus on a simple alignment procedure which eliminates the need of estimating $E\psi'$, incorporates a larger class of score functions and provides a computationally simpler and asymptotically equi-efficient M-test for (1.2). Along with the preliminary notions, the proposed M-tests are considered in Section 2. Asymptotic properties of these tests are studied in Section 3. In the concluding section, the results are extended to some general linear models.

2. Aligned M-tests. For every $\underset{\sim}{t} = (t_1, t_2) \in R^2$, define $\underset{\sim}{M}_n(\underset{\sim}{t}) = (M_{n1}(\underset{\sim}{t}), M_{n2}(\underset{\sim}{t}))$ by letting

$$M_{n1}(\underset{\sim}{t}) = \sum_{i=1}^{n}\psi(X_i - t_1 - t_2c_i) , \qquad (2.1)$$

$$M_{n2}(\underset{\sim}{t}) = \sum_{i=1}^{n}c_i\psi(X_i - t_1 - t_2c_i) , \qquad (2.2)$$

where ψ is some score function. We assume that

$$\psi(x) = \psi_1(x) + \psi_2(x) , \quad x \in R , \qquad (2.3)$$

where ψ_1 is absolutely continuous on any bounded interval in R, and for the step-function ψ_2, we assume that for some $p(\geq 0)$, there exist open-intervals $J_r = (a_r, a_{r+1})$, $0 \leq r \leq p$, with $-\infty = a_0 < \ldots < a_p < a_{p+1} = +\infty$, such that

$$\psi_2(x) = d_r , \quad \forall x \in J_r , \quad r=0,1,\ldots,p , \qquad (2.4)$$

where the d_r are real numbers, not all equal. For $p \geq 1$, conventionally, we let $\psi_2(a_r) = \frac{1}{2}(d_r + d_{r-1})$, $1 \leq r \leq p$. Further, we assume that both ψ_1 and ψ_2 are non-decreasing and skew-symmetric, so that $\psi_j(-x) = -\psi_j(x)$, $\forall x$ and $j=1,2$. The d.f. F in (1.1) is assumed to be symmetric about 0, so that $\int_R \psi(x)dF(x) = 0$.

Let $(\hat{\theta}_n, \hat{\beta}_n)$ be the M-estimator of (θ, β), i.e., $(\hat{\theta}_n, \hat{\beta}_n)$ is a solution of the system of equations $\underset{\sim}{M}_n(\underset{\sim}{t}) = \underset{\sim}{0}$, or

$$\underset{\sim}{M}_n(\hat{\theta}_n, \hat{\beta}_n) = \underset{\sim}{0} . \qquad (2.5)$$

Also, let $\tilde{\theta}_n$ be the M-estimator of θ when $\beta = 0$ i.e., $\tilde{\theta}_n$ is a solution of $M_{n1}(t, 0) = 0$. Since $M_{n1}(t, 0)$ is $\searrow$ in t, $\tilde{\theta}_n$ may conveniently be written as

$$\tfrac{1}{2}[\sup\{t: M_{n1}(t, 0) > 0\} + \inf\{t: M_{n1}(t, 0) < 0\}) . \qquad (2.6)$$

The solution in (2.5), however, may have to be obtained by the trial and error method. If $Q(x)$ be defined by $Q'(x) \equiv \psi(x)$, then the Schrader and

Hettmansperger (1980) procedure rests on

$$2\gamma\sigma_0^{-2}\{\sum_{i=1}^n (Q(X_i - \tilde{\theta}_n) - Q(X_i - \hat{\theta}_n - \hat{\beta}_n c_i))\} \tag{2.7}$$

where

$$\sigma_0^2 = \int_R \psi^2(x)dF(x) \ , \quad \gamma = -\int_R \psi(x)f'(x)dx \tag{2.8}$$

and it is assumed that $0 < \sigma_0, \ \gamma < \infty$. In practice, their testing procedure needs to substitute suitable estimators of σ_0 and γ in (2.7) and rests on the asymptotic chi aquare distribution of (2.7), with one degree of freedom (D.F.), when H_0 holds. Also, for the solution in (2.5), for large n, we have [viz., Yohai and Maronna (1979)]

$$C_n^*(\hat{\beta}_n - \beta) \sim N(0, \sigma_0^2/\gamma^2) \ , \tag{2.9}$$

where

$$C_n^{*2} = \sum_{i=1}^n c_i^2 - n^{-1}(\sum_{i=1}^n c_i)^2 \ , \tag{2.10}$$

and hence, using suitable estimators of σ_0 and γ, an asymptotic test for (1.2) can be based on (2.9).

The procedure to be considered here is quite simple and does not require the estimation of γ or $(\hat{\theta}_n, \hat{\beta}_n)$. Let

$$\hat{M}_n = M_{n2}(\tilde{\theta}_n, 0) \ , \quad S_n^2 = n^{-1}\sum_{i=1}^n \psi^2(X_i - \tilde{\theta}_n) \ , \tag{2.11}$$

$$T_n = \hat{M}_n(C_n^* S_n)^{-1} \ . \tag{2.12}$$

We propose to use T_n as a test statistic. In the next section, we study the properties of T_n and show that for local (contiguous) alternatives, the asymptotic relative efficiency (A.R.E.) of T_n with respect to either of the earlier test is equal to 1. Thus the A.R.E. of T_n is not affected by its computational simplicity. Further, like the test based on (2.9), T_n may be employed for both one and two-sided alternatives.

3. Properties of the aligned M-test. We assume that the d.f. F posesses an absolutely continuous probability density function (p.d.f.) f having a finite Fisher information

$$I(f) = \int_R \{f'(x)/f(x)\}^2 dF(x) . \tag{3.1}$$

Note that (3.1) insures that $\int_R |f'(x)|dx < \infty$. Concerning the constants c_i in (1.1), we assume that

$$\lim_{n\to\infty}\{\max_{1\le i\le n}(c_i - \bar{c}_n)^2/C_n^{*2}\} = 0 , \tag{3.2}$$

$$\lim_{n\to\infty} n^{-1}C_n^{*2} = C^{*2}: \quad 0 < C^* < \infty , \tag{3.3}$$

$$\lim_{n\to\infty} \bar{c}_n = \bar{c} \text{ exists, where } |\bar{c}| < \infty . \tag{3.4}$$

(Note that $\bar{c}_n = n^{-1}\sum_{i=1}^n c_i$.) Then, as a basis of our subsequent analysis, we quote the following result due to Jurečková (1977):

Under H_0: $\beta = 0$, for every K: $0 < K < \infty$,

$$\sup\{n^{-\frac{1}{2}}||\underset{\sim}{M}_n((\theta, 0) + \underset{\sim}{t}) - \underset{\sim}{M}_n((\theta, 0) + \gamma n \underset{\sim}{t}\underset{\sim}{V}||: \; ||\underset{\sim}{t}|| \le n^{-\frac{1}{2}}K\} \xrightarrow[p]{} 0 , \text{ as } n \to \infty , \tag{3.5}$$

where

$$\underset{\sim}{V} = \begin{pmatrix} 1 & \bar{c} \\ \bar{c} & C^{*2} + \bar{c}^2 \end{pmatrix} . \tag{3.6}$$

Further, from Yohai (1974), Jurečková (1977) and Jurečková and Sen (1981), ir follows that under the assumed regularity conditions,

$$|n^{\frac{1}{2}}(\tilde{\theta}_n - \theta)| = 0_p(1) , \text{ when } H_0 \text{ holds} , \tag{3.7}$$

$$|n^{\frac{1}{2}}(\hat{\theta}_n - \theta, \; \hat{\beta}_n - \beta)| = 0_p(1) . \tag{3.8}$$

From (2.6) and (3.5) through (3.8), we obtain that under H_0: $\beta = 0$, as $n \to \infty$,

$$n^{-\frac{1}{2}}\hat{M}_n = n^{-\frac{1}{2}}M_{n2}(\tilde{\theta}_n, 0)$$

$$= n^{-\frac{1}{2}}M_{n2}(\theta + n^{-\frac{1}{2}}\{n^{\frac{1}{2}}(\tilde{\theta}_n - \theta)\}, 0)$$

$$= n^{-\frac{1}{2}}\{M_{n2}(\theta, 0) - \bar{c}M_{n1}(\theta, 0)\} + o_p(1), \quad (3.9)$$

where under (θ, β), $\underset{\sim}{M}_n(\theta, \beta)$ has the same (joint) distribution as $\underset{\sim}{M}_n(\underset{\sim}{0})$ under H_0^*: $\theta = \beta = 0$, and further under H_0^*,

$$n^{-\frac{1}{2}}\underset{\sim}{M}_n(\underset{\sim}{0}) \xrightarrow[\mathcal{D}]{} N_2(\underset{\sim}{0}, \sigma_0^2\underset{\sim}{V}) . \quad (3.10)$$

Thus, by (3.9) and (3.10), under H_0: $\beta = 0$,

$$n^{-\frac{1}{2}}\hat{M}_n \xrightarrow[\mathcal{D}]{} N_1(0, \sigma_0^2 C^{*2}) , \quad (3.11)$$

so that by (3.3) and (3.11),

$$\hat{M}_n/C_n^* \xrightarrow[\mathcal{D}]{} N_1(0, \sigma_0^2), \text{ when } H_0 \text{ holds.} \quad (3.12)$$

The consistency of S_n^2, as an estimator of σ_0^2, follows as in Jurečková and Sen (1981), and hence, by (3.12), under H_0,

$$T_n \xrightarrow[\mathcal{D}]{} N_1(0, 1) . \quad (3.13)$$

By (3.13), the asymptotic critical values of T_n can be taken as the appropriate percentile points of the standard normal d.f. To study the asymptotic power properties of T_n, we confine ourselves to a class of local alternatives $\{H_n^0\}$, specified by

$$H_n^0: \beta = n^{-\frac{1}{2}}\lambda , \text{ where } \lambda(\neq 0) \text{ is fixed.} \quad (3.14)$$

Note that under (1.1), (3.1), (3.2)-(3.4) and (3.14), the contiguity of the (sequence of) probability measures under $\{H_n^0\}$ to those under H_0 follows as in Hájek and Šidák (1967, ch. 6). (3.9) and this contiguity imply that as $n \to \infty$,

$$n^{-\frac{1}{2}}|\hat{M}_n - M_{n2}(\theta, 0) + \bar{c}M_{n1}(\theta, 0)| \overset{p}{\to} 0 \text{ , under } \{H_n^0\} \text{ .} \qquad (3.15)$$

Also, by (2.1), (2.2), (3.10) and the contiguity established before, under $\{H_n^0\}$,

$$n^{-\frac{1}{2}}\underset{\sim}{M}_n(\theta, 0) \underset{\mathcal{D}}{\to} \mathcal{N}_2(\underset{\sim}{\xi}, \sigma_0^2\underset{\sim}{V}) \text{ ,} \qquad (3.16)$$

where

$$\underset{\sim}{\xi} = \lambda\gamma(\bar{c}, C^{*2} + \bar{c}^2) \qquad (3.17)$$

and γ is defined by (2.8). Thus, by (3.15), (3.16) and (3.17), under $\{H_n^0\}$,

$$\hat{M}_n/C_n^* \underset{\mathcal{D}}{\to} \mathcal{N}_1(\lambda\gamma C^*, \sigma_0^2) \text{ .} \qquad (3.18)$$

Finally, the stochastic convergence of S_n^2 (to σ_0^2) under H_0 and the aforesaid contiguity insue that $S_n^2 \underset{p}{\to} \sigma_0^2$, under $\{H_n^0\}$ as well. Thus, by (3.18) and the above, under $\{H_n^0\}$,

$$T_n \underset{\mathcal{D}}{\to} \mathcal{N}_1(\lambda\gamma C^*/\sigma_0, 1). \qquad (3.19)$$

The efficacy of $\{T_n\}$ (in the Pitman-sense) is therefore equal to

$$\gamma^2 C^{*2}/\sigma_0^2 \qquad (3.20)$$

and this agrees with the efficacy of the test based on $\hat{\beta}_n$ or the likelihood ratio type test, considered by Schreder and Hettmansperger (1980). Hence, the aligned M-test based on T_n is asymptotically equi-efficient (for Pitman-type alternatives) with respect to either of the other tests.

4. Aligned M-tests in general linear model. Consider now the general linear model

$$X_i = \beta' c_i + e_i \ , \quad 1 \le i \le n \ , \tag{4.1}$$

where $\beta' = (\beta_1, \ldots, \beta_q)$ for some $q(\ge 1)$ and the c_i are specified q-vectors. Let $\beta' = (\beta_1', \beta_2')$, where β_1 and β_2 are r and s vectors, with $r + s = q$. By a canonical reduction, we may consider the following linear hypothesis

$$H_0: \ \beta_2 = 0 \text{ against } H_1: \ \beta_2 \ne 0 \ , \tag{4.2}$$

where β_1 is treated as a nuisance parameter (vector).

The aligned M-test for (4.2) is very similar to that in Section 2. We partition c_i' as (c_{i1}', c_{i2}'), $1 \le i \le n$, and note that under H_0, we have $X_i = \beta_1' c_{i1} + e_i$, $1 \le i \le n$. Let $\tilde{\beta}_{n,1}$ be the M-estimator of β_1 (under H_0) i.e., it is a solution of

$$\textstyle\sum_{i=1}^n c_{i1}\psi(X_i - t' c_{i1}) = 0 \ . \tag{4.3}$$

Let then

$$\hat{M}_n = \textstyle\sum_{i=1}^n c_{i2}\psi(X_i - \tilde{\beta}_{n,1}' c_{i1}) \ , \tag{4.4}$$

$$C_{nrs} = \textstyle\sum_{i=1}^n c_{ir} c_{is}' \ , \ r,s = 1,2 \ , \tag{4.5}$$

$$C_n^* = C_{n22} - C_{n21} C_{n11}^- C_{n12} \ , \tag{4.6}$$

$$s_n^2 = n^{-1}\textstyle\sum_{i=1}^n \psi^2(X_i - \tilde{\beta}_{n,1}' c_{i1}) \tag{4.7}$$

and

$$T_n^2 = s_n^{-2}(\hat{M}_n' (C_n^*)^- \hat{M}_n) \ . \tag{4.8}$$

If we assume that

$$\lim_{n\to\infty} n^{-1}\underset{\sim}{C}_n^* = \underset{\sim}{C}^* \text{ exists and is positive definite} \tag{4.9}$$

and (3.2) and (3.4) hold coordinate arise, then under the assumed regularity conditions on ψ and F and the additional concordance-discordance condition on the $\underset{\sim}{c}_i$, due to Jurečková (1977), we are able to make use of her Theorem 4.1 and proceeding as in Section 3 obtain the following:

(i) Under H_0, T_n^2 has asymptotically the central chi square distribution with s D.F.

(ii) Under $\{H_n^0\}$, where H_n^0: $\underset{\sim}{\beta}_2 = n^{-\frac{1}{2}}\underset{\sim}{\lambda}$, T_n^2 has asymptotically a non-central chi-square d.f. with s D.F. and noncentrality parameter

$$\Delta = \gamma^2\sigma_0^{-2}(\underset{\sim}{\lambda}'(\underset{\sim}{C}^*)^{-1}\underset{\sim}{\lambda}) \ . \tag{4.10}$$

By virtue of these results, it follows that T_n^2 has the same A.R.E. as the likelihood ratio type test considered by Schrader and Hettmansperger (1980). On the other hand, T_n^2 does not need the computation of the unrestricted M-estimator $\hat{\underset{\sim}{\beta}}_n$ of $\underset{\sim}{\beta}$ in (4.1) nor the estimator of γ .

REFERENCES

[1] Hájek, J. and Šidák, Z. (1967). *The Theory of Rank Tests.* New York: Academic Press.

[2] Huber, P.J. (1973). Robust regression: Asymptotics, conjectures and Monte Carlo. *Ann. Statist.* 1, 799-821.

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

[4] Jurečková, J. (1977). Asymptotic relations of M-estimates and R-estimates in linear regression model. *Ann. Statist. 5*, 464-472.

[5] Jurečková, J. (1980). Asymptotic representation of M-estimators of location. *Math. Oper. Statist., Ser. Statist.* 11, 61-73.

[6] Jurečková, J. and Sen, P.K. (1981). Sequential procedures based on M-estimators with discontinuous score functions. *Jour. Statist. Plan. Inf.* 5, in press.

[7] Schrader, R.M. and Hettmansperger, T.P. (1980). Robust analysis of variance based upon a likelihood ratio criterion. *Biometrika, 67*, 93-101.

[8] Yohai, V.J. (1974). Robust estimation in the linear model. *Ann. Statist.* **3**, 562-567.

[9] Yohai, V.J. and Maronna, R.A. (1979). Asymptotic behavior of M-estimators for the linear model. *Ann. Statist.* **7**, 258-268.