Journal of the Korean
Statistical Society
Vol. 26, No. 4, 1997

A Robust Wald-Type Test in Linear Regression [†]

Ho Soo Nam [1]

Abstract

In this paper we propose a robust Wald-type test which is based on an efficient Mallows-type one-step GM-estimator. The proposed estimator based on the weight function of Song, Park and Nam (1996) has a bounded influence function and a high breakdown point. Under some regularity conditions, we compute the finite-sample breakdown point, and drive asymptotic normality of the proposed estimator. The level and power breakdown points, influence function and asymptotic distribution of the proposed test statistic are main points of this paper. To compare the performance of the proposed test with other tests, we perform some Monte Carlo simulations.

Key Words : Influence function; One-step GM-estimator; Power and level breakdown point; LTS estimator; MVE estimator.

[†] This paper was supproted by NON DIRECTED RESEARCH FUND, Korean Research Foundation, 1996.

[1] Department of Industrial Engineering, Dongseo University, Pusan, 617-716, Korea.

1. INTRODUCTION

In the linear regression model, *robust* in robust testing procedure means the distributional robustness, that is, robustness of level and power against heavy tailed error distribution or carrier's distribution in case of random carrier model. According to Markatou, Stahel and Ronchetti(1991) two fundamental goals in robust testing are (1) the level of a test should be stable under small and arbitrary departures from the null hypothesis (*robustness of validity*), and (2) the test should retain good power under small and arbitrary departures from specified alternatives (*robustness of efficiency*). The classical parametric tests can not be robust in this sense. If a test procedure uses the test statistic based on an estimator concerning the hypothesis, then any non-robust estimator can not lead the test to a robust one.

In this paper, we consider the linear regression model

$$y_i = \mathbf{x}_i^T \beta + \epsilon_i \ , \quad i = 1, 2, \cdots, n \ , \tag{1.1}$$

where $\{(\mathbf{x}_i, y_i) : i = 1, 2, \cdots, n\}$ is a sequence of independent and identically distributed (iid) random variables with distribution function $F(\mathbf{x}, y)$, $\mathbf{x}_i$ is a $p \times 1$ random vector, and β is a $p \times 1$ vector of unknown parameters. While, ϵ_i's are iid, independent of $\mathbf{x}_i$ and symmetric about 0 with scale parameter σ.

Let $\beta^T = (\beta_1^T, \beta_2^T)$, where β_1 is $(p-q) \times 1$ and β_2 is $q \times 1$ subvector of β. We want to test the null hypothesis

$$H_0 : \beta_2 = \mathbf{0} \ , \quad \beta_1 \text{ unspecified}$$

versus the alternative hypothesis

$$H_1 : \beta_2 \neq \mathbf{0} \ , \quad \beta_1 \text{ unspecified.} \tag{1.2}$$

Frequently, the null hypothesis is represented as $H\beta = \mathbf{0}$ for some $q \times p$ matrix H. As an alternative to the parametric F-test for the hypothesis (1.2), robust tests are discussed by many researchers: Markatou, Stahel and Ronchetti (1991), Markatou and He (1994), Heritier and Ronchetti (1994), etc.

We present a robust Wald-type test for the hypothesis (1.2). The test statistic is composed of the Mallows-type one-step generalized M(GM) estimator based on the weight function of the estimator proposed by Song, Park and Nam(1996). We also derive the influence function of the test statistic under the null hypothesis, and prove that the asymptotic level and power breakdown points of the test statistic is 0.5. The proposed test is asymptotically distribution-free and distributional robust on both size and power. To compare the performance of the proposed test with other tests, we perform some Monte Carlo simulations.

2. THE PROPOSED GM-ESTIMATOR

Song, Park and Nam (1996) proposed a bounded influence and high breakdown regression estimator, which is a solution of the simultaneous equations

$$\sum_{i=1}^{n} \eta(\mathbf{x}_i, r_i(\beta)/\sigma)\mathbf{x}_i = \mathbf{0} \ , \tag{2.1}$$

where

$$\eta\left(\mathbf{x}, \frac{r}{\sigma}\right) = w\left(\mathbf{x}, \frac{r}{\sigma}\right)\psi\left(\frac{r}{\sigma w\left(\mathbf{x}, r/\sigma\right)}\right), \tag{2.2}$$

$r = y - \mathbf{x}^T\beta$ and ψ is an odd and bounded function. Here, the weight function is given by

$$w\left(\mathbf{x}, \frac{r}{\sigma}\right) = \min\left(1,\ \frac{a \cdot \sigma v(\mathbf{x})}{|r|}\right), \tag{2.3}$$

where $v(\mathbf{x})$ is a measure of leverageness, and a is a tuning constant (see Song, Park and Nam (1996)). The equation (2.2) is a form of Schweppe-type, which is designed to overcome the defect that the Mallows-type downweights the leverage points regardless of the contribution to the model of these points. However, the weight (2.3) depends on both residuals and leverage points simultaneously. The Schweppe-type by using the weight (2.3) doubly downweights the outliers or leverage points, so it causes some deficiency. If this weight is used in the Mallows-type estimator, the defect of the Mallows-type may be overcomed. So, we propose a Mallows-type one-step GM-estimator based on the weight (2.3). Some simulation results, not listed in this paper, show that the following estimator in (2.6) has better performance than Song, Park and Nam (1996). The estimator is given by a solution of

$$\sum_{i=1}^{n} \eta(\mathbf{x}_i, r_i(\beta)/\sigma)\mathbf{x}_i = \mathbf{0} \ , \tag{2.4}$$

where

$$\eta\left(\mathbf{x}, \frac{r}{\sigma}\right) = w\left(\mathbf{x}, \frac{r}{\sigma}\right)\psi\left(\frac{r}{\sigma}\right). \tag{2.5}$$

Thus, the proposed one-step estimator by scoring method is given as follows:

$$\widehat{\beta} = \widehat{\beta}_0 + \widehat{\sigma}_0 H_0^{-1} g_0, \tag{2.6}$$

where

$$g_0 = \sum_{i=1}^{n} \eta(\mathbf{x}_i, r_i(\widehat{\beta}_0)/\widehat{\sigma}_0)\mathbf{x}_i$$

and

$$H_0 = X^T \widehat{A} X \ , \quad \widehat{A} = diag\left(\frac{1}{n}\sum_{j=1}^{n} \eta'(\mathbf{x}_i, \frac{r_j(\widehat{\beta}_0)}{\widehat{\sigma}_0})\right).$$

Here, X is the $n \times p$ matrix having rows $\mathbf{x}_i^T$, $\eta'(\mathbf{x}, r) = \partial\eta(\mathbf{x}, r)/\partial r$, $\widehat{\sigma}_0$ is a high breakdown scale estimator such as the median absolute deviation(MAD) estimator or Hill and Holland (1977)'s estimator, and $\widehat{\beta}_0$ is an initial estimator of β such as the least trimmed squares(LTS) estimator. We use the LTS estimator throughout this paper.

Now consider the properties of the proposed estimator (2.6) under the following assumptions.

(C1) The initial estimator $\widehat{\beta}_0$ has an influence function $IF(\mathbf{x}, y; \widehat{\beta}_0)$.

(C2) $\psi(\cdot)$ is bounded, $\eta(\mathbf{x}, r)$ is an odd function of r, and $\eta(\mathbf{x}, r) \geq 0$ for all $\mathbf{x} \in \Re^p$ and $r \in \Re^+$.

(C3) Without loss of generality, the first p observations are uncontaminated and $\mathbf{x}_1, \cdots, \mathbf{x}_p$ are linearly independent.

(C4) Assume that, for all $i = 1, \cdots, n$,

$$\sum_{j=1}^{n} \eta'(\mathbf{x}_i, r_j(\widehat{\beta}_0))$$

are positive at least for the first p observations, and nonnegative for the other observations.

(C5) $(1/n)H_0 \xrightarrow{P} D$, where D is a positive definite $p \times p$ matrix.

(C6) $(1/n)(X^T \widehat{V} X) \xrightarrow{P} E$, where $\widehat{V}$ is the diagonal matrix with diagonal elements $\sum_{j=1}^{n} \eta^2(\mathbf{x}_i, r_j(\widehat{\beta}_0)/\widehat{\sigma}_0)/n$, $i = 1, \cdots, n$ and E is a positive definite $p \times p$ matrix.

(C7) $\max_{i,j} |x_{ij}| = o_p(n^{1/2})$, for all $i = 1, \cdots, n$; $j = 1, \cdots, p$.

(C8) $\widehat{\beta}_0 = \beta + O_p(n^{-1/2})$.

(C9) $\widehat{\sigma}_0 = \sigma + O_p(n^{-1/2})$.

Theorem 2.1. (1) Under the assumptions (C1) and (C2), the proposed estimator has a bounded influence function.
(2) Under the assumptions (C2) through (C4), the proposed one-step estimator has a breakdown point of $([n/2] - p + 1)/n$.
(3) Under the assumptions (C2) and (C5) through (C9),

$$\sqrt{n}(\widehat{\beta} - \beta) \xrightarrow{d} N(\mathbf{0}, \sigma^2\Sigma),$$

where $\Sigma = D^{-1}ED^{-1}$ with

$$D = \int \eta' \left(\mathbf{x}, \frac{y - \mathbf{x}^T\beta}{\sigma}\right) \mathbf{x}\mathbf{x}^T dF(\mathbf{x}, y)$$

and

$$E = \int \eta^2 \left(\mathbf{x}, \frac{y - \mathbf{x}^T\beta}{\sigma}\right) \mathbf{x}\mathbf{x}^T dF(\mathbf{x}, y).$$

For inferences about β, the asymptotic variance of $\sqrt{n}\widehat{\beta}$, $nVar(\widehat{\beta}) = \sigma^2 D^{-1}ED^{-1}$, can be estimated as follows:

$$n\widehat{Var}(\widehat{\beta}) = \widehat{\sigma}_0^2\ \widehat{D}^{-1}\widehat{E}\,\widehat{D}^{-1}. \tag{2.7}$$

Here, $\widehat{D}$ and $\widehat{E}$ are computed as follows (Song, Park and Nam (1996)):

$$\widehat{D} = \frac{1}{n}(X^T\widehat{A}X)\ , \quad \widehat{E} = \frac{1}{n}(X^T\widehat{V}X).$$

3. TEST STATISTIC AND ITS PROPERTIES

In this section, we construct a test statistic for testing the hypothesis (1.2). A quadratic form of the second part $\widehat{\beta}_2$ in (2.6) can be used for a Wald-type test statistic. It is given as follows:

$$W_n^2 = \frac{\sqrt{n}\widehat{\beta}_2^T\ \widehat{\Sigma}_{22}^{-1}\ \sqrt{n}\widehat{\beta}_2}{\widehat{\sigma}_0^2}, \tag{3.1}$$

where $\widehat{\sigma}_0^2\,\widehat{\Sigma}_{22}$ is an estimator of asymptotic variance of $\sqrt{n}\widehat{\beta}_2$. Practically, the asymptotic variance of $\sqrt{n}\widehat{\beta}$ can be estimated by the same method as in (2.7), and $\widehat{\sigma}_0^2\widehat{\Sigma}_{22}$ is its $q \times q$ submatrix.

Now, consider the properties of the test based on the statistic W_n^2. Our concerns about the test statistic are focused on: (1) boundedness of the influence function, (2) level and power breakdown point, and (3) asymptotic distribution.

First denote $Q_n^2 = W_n^2/n$. The functional version of the statistic Q_n^2 at F is given by

$$Q^2(F) = \frac{\beta_2^T(F)\Sigma_{22}^{-1}\beta_2(F)}{\sigma^2}.$$

Note that if S is a test statistic with its functional $S(F)$, the influence function of S at F is defined by

$$IF(\mathbf{x}, y; S, F) = \lim_{t\to 0}\frac{S((1-t)F + t\Delta_{(\mathbf{x},y)}) - S(F)}{t},$$

where $\Delta_{(\mathbf{x},y)}$ denotes the pointmass 1 at $(\mathbf{x}, y)$. Since the statistic Q_n^2 can be represented as

$$Q_n^2 = (\widehat{\beta}_2^T\widehat{\Sigma}_{22}^{-1/2}/\widehat{\sigma}_0)\cdot(\widehat{\Sigma}_{22}^{-1/2}\widehat{\beta}_2/\widehat{\sigma}_0) = \widehat{U}^T\widehat{U},$$

where $\widehat{U} = \widehat{\Sigma}_{22}^{-1/2}\widehat{\beta}_2/\widehat{\sigma}_0$, the functional $Q^2(F)$ is divided as follows with $U(F) = \Sigma_{22}^{-1/2}\beta_2(F)/\sigma$:

$$Q^2(F) = U(F)^T U(F).$$

Thus, to prove the boundedness of the influence function of W_n^2, it is enough to show that the influence function of $Q(F)(=\sqrt{Q^2(F)})$ is bounded.

Theorem 3.1. If the assumption (C1) and (C2) hold, then under $H_0 : \beta_2 = 0$, the influence function of the test statistic W_n^2 is bounded.

Proof. To prove the theorem, it suffices that the influence function of $Q(F)(=\sqrt{Q^2(F)})$ is bounded. Let $F_t = (1-t)F + t\Delta_{(\mathbf{x},y)}$ and $\sigma = 1$ for simplicity. Then we have

$$U(F_t) = U(F) + tIF(\mathbf{x}, y; U, F) + O(t^2),$$

where $IF(\mathbf{x}, y; U, F)$ is the influence function of $U(F)$. Therefore,

$$Q^2(F_t) = [U(F) + tIF(\mathbf{x}, y; U, F)]^T\cdot[U(F) + tIF(\mathbf{x}, y; U, F)] + O(t^2).$$

To derive the influence function of $Q(F)$, we use the technique of Markatou, Stahel and Ronchetti (1991). Note that

$$\begin{aligned} IF(\mathbf{x},y;Q,F) &= \lim_{t\to 0}\left\{\frac{\partial}{\partial t}Q(F_t)\right\} \\ &= \lim_{t\to 0}\frac{1}{2}\left\{\frac{\partial}{\partial t}Q^2(F_t)\right\}\ [Q^2(F_t)]^{-1/2} \\ &= \lim_{t\to 0}\left\{\Big[2IF(\mathbf{x},y;U,F_t)^T U(F_t) \right. \\ &\quad \left. +2tIF(\mathbf{x},y;U,F)^T IF(\mathbf{x},y;U,F)+O(t^2)\Big]/2Q(F_t)\right\}. \qquad (3.2) \end{aligned}$$

If $U(F)=\mathbf{0}$, which is equivalent to '$Q^2(F)=0$', then we can apply L'Hopital's rule to (3.2) to obtain non-zero influence function. Thus, we have

$$IF(\mathbf{x},y;Q,F)=\begin{cases} IF(\mathbf{x},y;U,F)^T U(F)/Q(F), & if \quad U(F)\neq \mathbf{0} \\ ||IF(\mathbf{x},y;U,F)||, & if \quad U(F)=\mathbf{0}. \end{cases}$$

Note that the influence function of the proposed estimator is $D^{-1}w(\mathbf{x},y-\mathbf{x}^T\beta)\psi(y-\mathbf{x}^T\beta)\mathbf{x}$. By using this, we can easily obtain the influence function of $Q(F)$ under H_0. If F_{β_0} is the distribution function under the null hypothesis with $\beta_0^T=(\beta_1^T,\mathbf{0}^T)$ then we have

$$IF^2(\mathbf{x},y;Q,F_{\beta_0})=w^2(\mathbf{x},y-\mathbf{x}^T\beta)\psi^2(y-\mathbf{x}^T\beta)(\mathbf{x}^T D^{-1})_2\Sigma_{22}^{-1}(D^{-1}\mathbf{x})_2,$$

where $(D^{-1}\mathbf{x})_2$ is the subvector of q components of $D^{-1}\mathbf{x}$ that corresponds to β_2. Since D and Σ are positive definite matrix and $||w(\mathbf{x},y-\mathbf{x}^T\beta)\psi(y-\mathbf{x}^T\beta)\mathbf{x}||<\infty$ by (C2), $IF^2(\mathbf{x},y;Q,F_{\beta_0})$ is bounded as a function of $\mathbf{x}$ and $|y-\mathbf{x}^T\beta|$. Hence the theorem follows.

If the influence of a contaminated point under the null hypothesis is too excessive, then the level of the test becomes unstable. So, the boundedness of its influence has an important meaning. However, it is meaningless to argue whether the influence function is bounded or not under the alternative. Now, we divert our concern to breakdown analysis. Markatou and He (1994) defined the level and power breakdown points of a test as follows: the level breakdown point is the least amount of contamination of the null that is serious enough to drive the test statistic to arbitrary point in its domain, while the power breakdown point is the least amount of contamination that can possibly drive the test statistic to its null value regardless of the true alternative value. To define the level and power breakdown points of a test, we define the level breakdown function $\varepsilon^{**}_{\beta_2}$ and power breakdown function $\varepsilon^{*}_{\beta_2}$ as follows:

$$\varepsilon^{**}_{\beta_2} = \inf\{\varepsilon > 0 : S((1-\varepsilon)F + \varepsilon G; \beta_1, \mathbf{0}) = S(F; \beta_1, \beta_2), \text{ for some } G \text{ and } \beta_1\},$$

$$\varepsilon^{*}_{\beta_2} = \inf\{\varepsilon > 0 : S((1-\varepsilon)F + \varepsilon G; \beta_1, \beta_2) = S(F; \beta_1, \mathbf{0}), \text{ for some } G \text{ and } \beta_1\},$$

where $S(F; \beta_1, \beta_2)$ is the functional version of a test statistic S. Then the level breakdown point ε^{**} and the power breakdown point ε^{*} of the test are defined by

$$\epsilon^{**} = \sup_{\beta_2}\{\epsilon^{**}_{\beta_2}\}, \qquad \epsilon^{*} = \sup_{\beta_2}\{\epsilon^{*}_{\beta_2}\}.$$

With these definitions, we can define the finite-sample power breakdown point ε^{*}_{n} and level breakdown point ε^{**}_{n} of the test statistic W^2_n as follows:

$$\varepsilon^{*}_{n} = \min\left\{\frac{m}{n}\ ;\ W^2_{n(m)} = 0, \ \text{ regardless of the true alternative}\right\}$$

and

$$\varepsilon^{**}_{n} = \min\left\{\frac{m}{n}\ ;\ W^2_{n(m)} \text{ is infinite, under the null}\right\},$$

where $W^2_{n(m)}$ is the test statistic which is obtained by replacing m points among n with arbitrary points. Now, we can derive the following theorem according to these definitions of the finite sample power and level breakdown points. And Theorem 3.2 shows that the asymptotic power and level breakdown points of the test statistic are 0.5.

Theorem 3.2. Under the assumptions (C2) through (C4), the Wald-type test based on W^2_n has both level and power breakdown points of $([n/2]-p+1)/n$.

Proof. First, note that the breakdown point of $\widehat{\beta}_2$ is $([n/2]-p+1)/n$ by (C2), (C3) and (C4). The test statistic (3.1) can be rewritten as

$$W^2_n = \widehat{\beta}_2^T(n\widehat{D})(n\widehat{E})^{-1}(n\widehat{D})\widehat{\beta}_2/\widehat{\sigma}_0^2.$$

Assume that m points are contaminated $(p \le m \le [n/2]-p+1)$. Let $a_i = \sum_{j=1}^{n} \eta'(\mathbf{x}_i, r_j(\widehat{\beta}_0)/\widehat{\sigma}_0)$ and $v_i^2 = \sum_{j=1}^{n} \eta^2(\mathbf{x}_i, r_j(\widehat{\beta}_0)/\widehat{\sigma}_0)$. Then, we have

$$\lambda_{min}(n\widehat{E}) = \lambda_{min}(\sum_{i=1}^{n} v_i^2 \mathbf{x}_i\mathbf{x}_i^T) \ge \lambda_{min}(\sum_{i=1}^{p} v_i^2 \mathbf{x}_i\mathbf{x}_i^T) > 0 \qquad (3.3)$$

and

$$\lambda_{min}(n\widehat{D}) = \lambda_{min}(\sum_{i=1}^{n} a_i \mathbf{x}_i \mathbf{x}_i^T) \geq \lambda_{min}(\sum_{i=1}^{p} a_i \mathbf{x}_i \mathbf{x}_i^T) > 0, \tag{3.4}$$

because a_i is positive at least for p good points. Note that for a positive semidefinite $p \times p$ matrix P

$$\lambda_{max}(P) \leq \sum_{k=1}^{p} \lambda_k(P) = tr(P),$$

where $\lambda_{min}(P)$, $\lambda_{max}(P)$ and $tr(P)$ mean "minimum", "maximum" eigen values of P and "trace" of P, respectively. From this, we have

$$\lambda_{max}(n\widehat{E}) = \lambda_{max}(\sum_{i=1}^{n} v_i^2 \mathbf{x}_i \mathbf{x}_i^T) \leq \sum_{i=1}^{n} v_i^2 ||\mathbf{x}_i||^2 < \infty, \tag{3.5}$$

since $||\eta(\mathbf{x}_i, r_i)\mathbf{x}_i|| < \infty$ for all i. Finally,

$$\lambda_{max}(n\widehat{D}) \leq \sum_{i=1}^{n} a_i ||\mathbf{x}_i||^2 < \infty. \tag{3.6}$$

Now, consider the power and level breakdown points of the test statistic. First, note that, for the two positive definite matrix P and Q, the following inequalities hold

$$\frac{||\mathbf{z}||^2 \lambda_{min}(P)}{\lambda_{max}(Q)} \leq \mathbf{z}^T P Q^{-1} \mathbf{z} \leq \frac{||\mathbf{z}||^2 \lambda_{max}(P)}{\lambda_{min}(Q)}.$$

The power breakdown occurs when the test statistic is driven to the null value 0. But we have $W_n^2 > 0$, because

$$W_n^2 \geq \frac{||\widehat{\beta}_2||^2 \lambda_{min}^2(n\widehat{D})}{\lambda_{max}(n\widehat{E})},$$

by (3.4) and (3.5). Note that $\widehat{\beta}$ has a breakdown point of $([n/2] - p + 1)/n$, so $||\widehat{\beta}_2||$ remains positive and bounded.

Next, if the test statistic is driven to an arbitrary large value in its domain $[0, \infty)$, then the level breakdown occurs. By (3.3) and(3.6) we also have

$$W_n^2 \leq \frac{||\widehat{\beta}_2||^2 \lambda_{max}^2(n\widehat{D})}{\lambda_{min}(n\widehat{E})} < \infty.$$

Thus, both the level and power breakdown points are of $([n/2] - p + 1)/n$. Hence the proof is completed.

Since the LS estimator has the asymptotic breakdown point of 0, the classical F-test based on the LS estimator has both power and level breakdown points of 0. The Wald-type test based on the one-step GM-estimator with high breakdown initial fit is considered by Simpson *et al.* (1992) and Coakley and Hettmansperger (1993), Markatou, Stahel and Ronchetti (1991), Markatou and He (1994), among others. These tests have bounded influence and high breakdown points for both level and power.

Now, we derive the asymptotic null and alternative distributions of the test statistic. The following theorem stems directly from Theorem of Stroud (1971).

Theorem 3.3. Under the assumptions (C2) and (C5) through (C9),

$$W_n^2 \xrightarrow{d} \chi^2(q), \quad \text{under } H_0 : \beta_2 = \mathbf{0},$$

$$W_n^2 \xrightarrow{d} \chi^2(q, \delta), \quad \text{under } H_1 : \beta_2 = n^{-1/2}\Delta,$$

where $\delta = \Delta^T(\Sigma_{22})^{-1}\Delta/\sigma^2$, the noncentrality parameter.

The above theorem gives us an approximate test for H_0. That is, for sufficiently large n, we reject H_0 at level α if

$$W_n^2 \;\geq\; \chi^2_{q,\alpha}\,, \tag{3.7}$$

where $\chi^2_{q,\alpha}$ is the upper α-quantile of chi-square distribution with q degrees of freedom.

4. SOME MONTE CALRO RESULTS

In this section, we want to compare the proposed test (P) with the classical F-test (F), the Wald-type tests based on the Huber's M-estimator (H), Mallows-type (M) and Schweppe-type (S) GM-estimators. For this, the following model is considered:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i, \quad i = 1, 2, \cdots, 50,$$

where $\beta_0 = 1,\ \beta_1 = 1$. The hypothesis is $H_0 : \beta_2 = 0$ versus $H_1 : \beta_2 \neq 0$. The random carriers x_{ij}'s are independently generated from the normal distribution $N(0, 1)$, for $i = 1, \cdots, 50$ and $j = 1, 2$. These generated values are fixed through simulation, and there are 15 leverage points in the aspect of robust Mahalanobis distance based on the minimum volume ellipsoid estimators. In this simulation, three situations are considered: (1) ϵ_i's are independently generated from the standard normal distribution, (2) ϵ_i's are independently generated from the contaminated normal $CN(0.2, 5)$, and (3) ϵ_i's are independently generated from $CN(0.2, 10)$. Here, the distribution function of an ε-contaminated normal distribution $CN(\varepsilon, \sigma)$ is given by

$$F(x) = (1 - \varepsilon)\Phi(x) + \varepsilon\Phi(x/\sigma).$$

The LTS estimate is used as an initial estimate of β. For the scale estimate, we use Hill and Holland(1977)'s estimator:

$$\widehat{\sigma}_0 = 2.1 * median\ \{\ largest\ n - p + 1\ of\ the\ |r_i(\widehat{\beta}_0)|\}. \tag{4.1}$$

The MAD can be used as a robust(high breakdown) estimate of the scale σ. However, the use of the MAD incurs very poor control for level, since the MAD usually underestimates the scale σ. For the well controlled level, we use the estimator (4.1) corrected to reduce bias.

The tuning constants used for the weight function and ψ-function are as follows:

$$a = 3.0, \quad c = 1.5,$$

where the tuning constant c is used for the Huber's ψ. All M- or GM-estimates considered in this simulation use the Huber's ψ.

The results in Table 4.1 are based on 500 replications each, and the comparison is performed in three types of significance levels(α): 0.01, 0.05, 0.10. The constant m indicates the change of β_2. If $m = 0$, it means the null. In this simulation, m is composed as follows:

$$\beta_2 = m \cdot \delta,$$

where δ is a positive constant, which is different in cases. By inspecting the results, the classical F-test is a most powerful test in Case 1. However, it is unsatisfactory in Case 2 and Case 3, which have many bad leverage points. While the proposed test shows better performance in Case 2 and Case 3 in either stability of levels and powers than other tests.

Table 4.1 Empirical Levels and Powers

m	(α)	F	H	P	S	M
		Case 1 : $\epsilon \sim N(0,1)$				
0	(.01)	.008	.010	.012	.016	.008
	(.05)	.048	.042	.048	.052	.040
	(.10)	.102	.082	.080	.090	.084
1	(.01)	.170	.148	.144	.150	.136
	(.05)	.350	.318	.306	.330	.314
	(.10)	.482	.442	.440	.440	.428
2	(.01)	.280	.264	.264	.288	.248
	(.05)	.528	.466	.468	.470	.456
	(.10)	.658	.620	.612	.590	.592
3	(.01)	.478	.432	.428	.432	.400
	(.05)	.704	.678	.666	.676	.648
	(.10)	.816	.784	.788	.790	.784
		Case 2 : $\epsilon \sim CN(0.2,5)$				
0	(.01)	.008	.010	.012	.010	.010
	(.05)	.042	.050	.054	.046	.048
	(.10)	.088	.086	.094	.100	.086
1	(.01)	.056	.136	.164	.148	.116
	(.05)	.164	.284	.344	.330	.266
	(.10)	.260	.388	.456	.432	.390
2	(.01)	.132	.398	.472	.432	.372
	(.05)	.304	.612	.676	.660	.578
	(.10)	.420	.708	.772	.742	.688
3	(.01)	.190	.490	.568	.556	.476
	(.05)	.380	.700	.758	.740	.672
	(.10)	.510	.790	.844	.812	.776
		Case 3 : $\epsilon \sim CN(0.2,10)$				
0	(.01)	.008	.008	.010	.018	.008
	(.05)	.052	.054	.056	.060	.050
	(.10)	.108	.100	.106	.108	.110
1	(.01)	.018	.110	.208	.150	.102
	(.05)	.080	.282	.362	.302	.268
	(.10)	.126	.376	.488	.412	.356
2	(.01)	.050	.286	.444	.346	.270
	(.05)	.122	.528	.686	.586	.492
	(.10)	.224	.646	.784	.708	.634
3	(.01)	.056	.452	.644	.546	.424
	(.05)	.178	.694	.830	.734	.674
	(.10)	.276	.780	.886	.836	.758

5. CONCLUSION

The estimator proposed in this paper is one-step GM-estimator that uses Mallows-type weighting scheme. Under some regularity conditions the proposed estimator has good properties in the aspects of bounded influence function, high breakdown point, and efficiency.

The main point of this paper is a robust test based on the proposed estimator. The proposed test has the maximal breakdown points of level and power, and the influence function of the test statistic under the null is bounded.

Some Monte Carlo results show that the proposed test has better performance than other tests in either stability of levels and powers when there are some outliers.

ACKNOWLEDGEMENTS

The author thanks the anonymous referees and the associate editor for their helpful comments and suggestions.

REFERENCES

(1) Coakley, C.W. and Hettmansperger, T.P. (1993). A Bounded Influence, High Breakdown, Efficient Regression Estimator, *Journal of the American Statistical Association*, **88**, 872-880.

(2) Heritier, S. and Ronchetti, E. (1994). Robust Bounded-Influence Tests in General Parametric Models, *Journal of the American Statistical Association*, **89**, 897-904.

(3) Hill, R.W. and Holland, P.W. (1977). Two Robust Alternatives to Least-Squares Regression, *Journal of the American Statistical Association*, **72**, 828-833.

(4) Markatou, M. and He, X. (1994). Bounded Influence and High Breakdown Point Testing Procedures in Linear Models, *Journal of the American Statistical Association*, **89**, 543-549.

(5) Markatou, M., Stahel, W.A., and Ronchetti, E. (1991). Robust M-type Testing Procedures for Linear Models. In *Directions in Robust Statistics and Diagnostics*, edited by W. Stahel and S. Weisberg, part II, 201-220. New York: Springer-Verlag.

(6) Simpson, D.G., Ruppert, D. and Carroll, R.J. (1992). On One-Step GM Estimates and Stability of Inferences in Linear Regression, *Journal of the American Statistical Association*, **87**, 439-450.

(7) Song, M.S., Park, C.S. and Nam, H.S. (1996). A High Breakdown and Efficient GM-Estimator in Linear Models, *Journal of the Korean Statistical Society*, **25**(4).

(8) Stroud, T.W. (1971). On Obtaining Large-Sample Tests from Asymptotically Normal Estimators, *The Annals of Mathmatical Statistics,* **42**(4), 1412-1424.