TEST OF LINEARITY IN GENERAL REGRESSION MODELS (U)
PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS
X R CHEN ET AL. DEC 86 TR-86-49 AFSR-TR-87-1093
UNCLASSIFIED F49620-85-C-0008
F/G 12/3
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 196- A
Let \((X_i, Y_i), i = 1, \ldots, n\), be i.i.d. samples of \((X, Y)\). This paper proposes a method for testing the linearity of the regression function \(E(Y|X = x)\). The asymptotic distribution (under null hypothesis) and the
asymptotic power of the test are determined. Also, consistency of the test is proved under mild conditions.
TEST OF LINEARITY IN
GENERAL REGRESSION MODELS *

X.R. Chen and P.R. Krishnaiah
Center for Multivariate Analysis
University of Pittsburgh

Center for Multivariate Analysis
University of Pittsburgh

87 9 24 281
TEST OF LINEARITY IN
GENERAL REGRESSION MODELS *

X.R. Chen and P.R. Krishnaiah
Center for Multivariate Analysis
University of Pittsburgh

December 1986
Technical Report No. 86-49

Center for Multivariate Analysis
Fifth Floor Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

* Research sponsored by the Air Force Office of Scientific Research (AFSC), under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.
TEST OF LINEARITY IN GENERAL REGRESSION MODELS *

X.R. Chen and P.R. Krishnaiah
Center for Multivariate Analysis
University of Pittsburgh

ABSTRACT

Let \((X_i, Y_i), i = 1, \ldots, n\), be iid. samples of \((X, Y)\). This paper proposes a method for testing the linearity of the regression function \(E(Y|X = x)\). The asymptotic distribution (under null hypothesis) and the asymptotic power of the test are determined. Also, consistency of the test is proved under mild conditions.

AMS 1980 Subject Classification: Primary 62J05; Secondary 62C35.

Key words and phrases: Linear regression, consistency, strong approximation.

* Research sponsored by the Air Force Office of Scientific Research (AFSC), under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.
1. INTRODUCTION

Linear regression models are widely used in statistical analysis of experimental and observational data. Usually the linearity of the model is merely an assumption and cannot be taken for granted. In some planned experiments, repeated measurements on the dependent variable $Y$ can be taken while the independent variable $X$ is held fixed. In such cases standard analysis-of-variance technique can be employed to generate a test for linearity. In many applications, however, the independent variable is observed simultaneously with $Y$. That is to say, $X$, as well as $Y$, is a random variable. Under such circumstance the usual method for testing linearity cannot apply.

In this paper we shall study this problem in the large-sample context. We propose a method to test the linearity hypothesis based on a grouping of the data. The critical value of test-statistic is determined so that the test has a prescribed level of significance $\alpha$ asymptotically as the sample size tends to infinity. The consistency of the test is established, and the asymptotic power is calculated when the distance (in some sense) between the true regression function and the space of linear functions tends to zero in some specific rate.
2. ASSUMPTIONS OF THE MODEL

In this section we shall give a detailed account of our assumptions. We shall adhere to these assumptions in the sequel.

A basic assumption is that the conditional distribution of \( Y \) given \( \overline{X} = x \) has a form \( (F(y - m(x)), i.e., P(Y < y|\overline{X} = x) = F(y - m(x)) \), where \( F \) is a fixed distribution function. We do not assume \( F \) to be known. Let \((\overline{X}_1, Y_1), \ldots, (\overline{X}_n, Y_n)\) be iid. observations of \((\overline{X}, Y)\). Then, under this assumption, we can give \( Y_i \) a more convenient expression, as follows:

\[
Y_i = m(\overline{X}_i) + e_i, \quad i = 1, \ldots, n
\]  

where \( e_1, e_2, \ldots \) are iid. with common distribution function \( F \), and \( \overline{X}_1, \ldots, \overline{X}_n, e_1, \ldots, e_n \) are mutually independent. Also, \( \mathbb{E}e_1 = 0 \). We further assume that

1°. The moment generating function of \( e_1 \) exists in some neighborhood of zero.

2°. The variance \( \sigma^2 \) of \( e_1 \) (whose existence follows from 1°) is positive. Denote by \( \mu \) the probability distribution of \( \overline{X} \). We assume that

3°. \( \mathbb{E}\|\overline{X}\|^2 < \infty \), and \( \text{COV}(\overline{X}) > 0 \) (\( \|\cdot\| \) is the Euclidian norm).

4°. \( \mu \) has no singular component, and if \( \mu \) has an absolute component with density \( f \), then for sufficiently small \( a > 0 \) there exists an open set \( G_a \), such that

\[
|G_a \Delta \{x: f(x) > a\}| = 0
\]

where \( |A| \) denotes the Lebesgue measure of \( A \), and \( \Delta \) is symmetric difference. This condition is a rather weak one since it is satisfied by such \( f \) whose discontinuity points all lie in a closed set with Lebesgue measure zero.

For brevity, in the sequel we shall use "model (1)" in the sense that all of the above assumptions are met.
3. CONSTRUCTION OF THE TEST

Choose two numbers $\epsilon_1$ and $\epsilon_2$, with the condition

$$0 < \epsilon_1 < \epsilon_2 < 1/3.$$  \hspace{1cm} (2)

Some further restrictions on the choice of $\epsilon_1, \epsilon_2$ will be needed, which will be stated in Section 5.

Put $\xi_n = n^{-(1/3-\epsilon_2)/d}$. Decompose $\mathbb{R}^d$ into a set $J^*_n$ of supercube having the form

$$\{ (x(1), ..., x(d)) : a_i \xi_n \leq x(i) < (a_i + 1) \xi_n, \ i = 1, ..., d \},$$  \hspace{1cm} (3)

$$a_i = 0, \pm 1, \pm 2, \ i = 1, ..., d.$$  

For $J \in J^*_n$, use $\#(J)$ to denote the number of elements in the set $J \cap \{ \overline{x}_1, ..., \overline{x}_n \}$. Write

$$\{ J : J \in J^*_n, \ \#(J) \geq n^{2/3+\epsilon_1} \} = \{ J_{n1}, J_{n2}, ..., J_{nc_n} \}.$$  \hspace{1cm} (4)

Obviously,

$$c_n \leq n^{1/3-\epsilon_1}.$$  \hspace{1cm} (5)

Further, let

$$J_{ni} \cap \{ \overline{x}_1, ..., \overline{x}_n \} = \{ \overline{x}_{ni}(1), \overline{x}_{ni}(2), ..., \overline{x}_{ni}(n_i) \}.$$  

We have by definition

$$n_i \geq n^{(2/3)+\epsilon_1}, \ i = 1, ..., c_n.$$  \hspace{1cm} (6)

We shall write $Y_{ni}(j)$ and $e_{ni}(j)$ for $Y_k$ and $e_k$, when $\overline{x}_{ni}(j) = \overline{x}_k$.

Denote by $\overline{x}_{n1}$ the arithmetic mean of $\overline{x}_{ni}(j)$, $j = 1, ..., n_i$. Similarly
we define $Y_{ni}$ and $e_{ni}$.

If the null hypothesis $H_0$:

$$m(x) = \alpha + \beta'x$$

is true, we have

$$Y_{ni} = \alpha + \bar{X}_{ni}\beta + e_{ni}, \quad i = 1, \ldots, c_n$$

where $Ee_{ni} = 0$. $\text{Var} e_{ni} = \sigma_n^2/n_i$, and $e_{n1}$, $\ldots$, $e_{nc_n}$ are independent.

Applying weighted least squares method to this model with weight matrix

$$W(n) = \text{diag}(n_1, n_2, \ldots, n_c)$$

one obtains the weighted residual sum of squares as

$$RSS_n = Y^t(n)W(n)Y(n) - Y^t(n)W(n)\bar{X}(n)(\bar{X}(n)W(n)\bar{X}(n))^{-1}\bar{X}(n)W(n)Y(n)$$

where

$$\bar{X}(n) = (\bar{x}_{n1}, \ldots, \bar{x}_{nc_n})', \quad Y(n) = (\bar{y}_{n1}, \ldots, \bar{y}_{nc_n})', \quad e(n) = (e_{n1}, \ldots, e_{nc_n})'$$

$$\bar{X}_{ni} = \bar{X}_n - \bar{X}_n, \quad \bar{y}_{ni} = Y_{ni} - \bar{y}_n, \quad \bar{e}_{ni} = e_{ni} - \bar{e}_n, \quad i = 1, \ldots, c_n$$

$$\bar{X}_n = \sum_{i=1}^{c_n} n_i \bar{x}_{ni}/N_n, \quad \bar{y}_n = \sum_{i=1}^{c_n} n_i \bar{y}_{ni}/N_n, \quad \bar{e}_n = \sum_{i=1}^{c_n} n_i e_{ni}/N_n$$

and

$$N_n = c_1 + n_2 + \ldots + n_{c_n}.$$ 

The definition of $RSS_n$ is meaningful even when (7) is false. In case

$$RSS_n = e^t(n)W(n)e(n) - e^t(n)W(n)\bar{X}(n)(\bar{X}(n)W(n)\bar{X}(n))^{-1}\bar{X}(n)W(n)e(n).$$

(11)
RSS_n tends to be large when the null hypothesis $H_0$ is not true, and this suggests a test for $H_0$. Reject $H_0$ when

$$\frac{RSS_n}{\sigma_n^2} > C$$

and accept $H_0$ when (12) does not hold. $\sigma_n^2$ is a suitable estimate of $\sigma^2$, to be defined in Section 5. In order to determine $C$ in (12), we must study the asymptotic distribution of the statistic $RSS_n/\sigma_n^2$ under $H_0$. This will be done in the following two sections.
4. ASYMPTOTIC DISTRIBUTION OF THE RESIDUAL SUM OF SQUARES

The purpose of this section is to prove the following:

**THEOREM 1.** Let \( p^* = p^*(\underline{x}_1, \underline{x}_2, \ldots) \) be the conditional distribution given \( \{\underline{x}_1, \underline{x}_2, \ldots\} \). Under model (1) and when \( H_0 \) is true, there exists random variable \( \xi_n \) whose conditional distribution given \( \{\underline{x}_1, \underline{x}_2, \ldots\} \) is \( \chi^2_{c_n - d} \), such that with probability one we have

\[
\frac{RSS/\sigma^2 - \xi_n}{P^*} \xrightarrow{n} 0, \quad \text{as } n \to \infty. \tag{13}
\]

**Proof:** First we proceed to show that

\[
\frac{\underline{x}'(n)W(n)\underline{x}(n)}{n} \to \text{COV}(\underline{x}), \quad \text{a.s.}, \quad \text{as } n \to \infty. \tag{14}
\]

For this purpose, denote the atoms of \( \underline{x} \) by \( a_1, a_2, \ldots \), and put \( p_i = P(\underline{x} = a_i) > 0, i = 1, 2, \ldots \). Given \( n > 0 \), find \( N \) so that \( p_{N+1} + p_{N+2} + \ldots < n \). Find \( b > 0 \) small enough, such that

\[
\int_{\{x = f(x) \leq b\}} f dx < n. \tag{15}
\]

According to condition 4°, there exist an open set \( G \), such that

\[
|G \Delta \{x: f(x) > b\}| = 0.
\]

Choose a bounded closed set \( F_0 \subset G \), so that \( \int_{G-F} f dx < n \). Denote by \( \epsilon_1 \) the distance between \( F_0 \) and the boundary of \( G \), then \( \epsilon_1 > 0 \).

By the inequality of Bennett (1962), we have

\[
P(|\#(J)/n - \mu(J)| \geq \epsilon) \leq 2\exp(- n\epsilon^2/(2 + \epsilon)) \tag{16}
\]

\[
P(|\#(J)/n - \mu(J)| \geq \epsilon\mu(J)) \leq 2\exp(- c\epsilon\mu(J)) \tag{17}
\]

where \( \epsilon > 0 \), and \( c > 0 \) is a constant depending solely on \( \epsilon \). Define two
subsets of $J^*_n$

$H_1 = \{J: J \in J^*_n, a_i \in J \text{ for some } i \leq N\}$

$H_2 = \{J: J \in J^*_n, J \cap F_0 \neq \emptyset\}$

and an event

$E_n = \{\#(J) \geq n^{2/3+\epsilon_1}, \text{ for every } J \in H_1 \cup H_2\}.$

Since the number of $J$'s contained in $H_1 \cup H_2$ is of the order $O(n^{-1/3+\epsilon_1})$ and

$\mu(J) \geq bn$ \text{ for } J \in H_1 \cap H_2 \text{ and } n \text{ sufficiently large, it is easy to prove by using (16) that}$

$$P(E_n) \geq 1 - O(e^{-cn^{1/3}}) \tag{18}$$

where $c$ does not depend on $n$.

Now define the event

$F_n = \{\min\{\mu(I_{1i}) : i = 1, \ldots, c_n\} \leq n^{-1/3+\epsilon_{1/2}}\}$

We are going to prove that for any given positive integer $k$, it is true that

$$P(F_n) = O(n^{-k}). \tag{19}$$

For this purpose, define $D_n = \{(x(1), \ldots, x(d)) : |x(i)| \leq n^{(k+1)/2}, i = 1, \ldots, d\}$. By condition $4^0$, it is readily shown

$$P(\bar{X}_i \in D_n, i = 1, \ldots, n) \geq 1 - O(n^{-k}).$$

Define

$H_3 = \{J: J \in J^*_n, J \cap D_n \neq \emptyset, \mu(J) \leq n^{-1/3+\epsilon_{1/2}}\}.$

There are at most $n^{(k+1)d}$ elements in $H_3$. Using (16), for $n$ sufficiently large, we have

$$P(F_n) \leq Cn^{-k} + \sum_{J \in H_3} P(\#(J) \geq n^{2/3+\epsilon_1})$$

$$\leq Cn^{-k} + \sum_{J \in H_3} P(\#(J)/n - \mu(J) \geq n^{-1/3+\epsilon_{1/2}})$$

$$\leq Cn^{-k} + n^{(k+1)d} O(e^{-n^{1/3}}) = O(n^{-k}) \tag{20}$$
which proves (19).

For each \( t \in \mathbb{R}^d \), denote by \( t^{(u)} \) the \( u \)-th coordinate of \( t \). Write \( E\bar{x}^{(u)} = m_u \). For each \( J \in J_n^* \), choose a point \( t_j \in J \), and if \( J \) is one of \( I_{n_1}, \ldots, I_{n_c} \), then choose \( t_j = \bar{x}_{ni} \). Put \( \tilde{J}_n = J_n^* - \{ I_{ni} : i = 1, \ldots, c_n \} \), we have

\[
| \sum_{i=1}^{c_n} \frac{n_i}{n} \bar{x}^{(u)}(u) - m_u | \leq | \sum_{J \in \tilde{J}_n} t_j^{(u)}(u)(J) - m_u | + | \sum_{i=1}^{c_n} \frac{n_i}{n} \bar{x}^{(u)}(u) - \sum_{J \in \tilde{J}_n^*} t_j^{(u)}(u)(J) |
\]

\[
\leq | \sum_{J \in \tilde{J}_n^*} t_j^{(u)}(u)(J) - m(u) | + | \sum_{i=1}^{c_n} \bigg( \frac{n_i}{n} - \mu(I_{ni}) \bigg) \bar{x}^{(u)}(u) | + | \sum_{J \in \tilde{J}_n} t_j^{(u)}(u)(J) |
\]

\[
\leq L_1 + L_2 + L_3.
\]  

(21)

Since \( E\|\bar{x}\| < \infty \), \( L_1 \to 0 \) as \( n \to \infty \). As for \( L_3 \), when event \( E_n \) occurs, we have

\[
\lim_{n \to \infty} \sup_{J \in \tilde{J}_n} \sum_{i=1}^{c_n} \bigg( \frac{n_i}{n} - \mu(I_{ni}) \bigg) \bar{x}^{(u)}(u) \leq E(\|\bar{x}^{(u)}\| I(F_0^C U(a_{N+1}, a_{N+2}, \ldots))).
\]  

(22)

By definition of \( N \) and \( F_0 \), we have \( P(\bar{x} \in F_0^C U(a_{N+1}, a_{N+2}, \ldots)) < 3n \). Letting \( n > 0 \) small enough, the right hand side of (22) can be made small enough.

This, combining with (18), shows that

\[
\lim_{n \to \infty} L_3 = 0, \quad a.s.
\]  

(23)

Finally we turn to \( L_2 \). When event \( F_n \) occurs, we have \( \mu(I_{ni}) \geq n^{1/3+\epsilon_1/2} \).

Also, by (17) we have

\[
P(\#(I_{ni})/n - \mu(I_{ni}) \geq \epsilon_\mu(I_{ni}) \text{ for at least one } i \text{ such that } (I_{ni}) > n^{-1/3+\epsilon_1/2} \leq n \cdot 2\exp(-\epsilon_\mu(I_{ni})^{-2} n^{-1/3+\epsilon_1/2}) = 0(n^{-2}).
\]  

(24)
This, combining with (19) (choosing \( k = 2 \) in (19), shows that with probability one \(|n_i/n - \mu(I_{ni})| \leq \varepsilon \mu(I_{nk}), \ i = 1, ..., c_n, \) for \( n \) sufficiently large. In this event we have

\[
L_2 \leq \epsilon \sum_{i=1}^{c_n} |\bar{X}_n(u)| \mu(I_{ni}) \leq \epsilon \sum_{J \subseteq \mathbb{J}^*_n} |\mathcal{T}_J(u)| \mu(J) + \epsilon E|\bar{X}(u)|.
\]

Therefore

\[
\lim_{n \to \infty} L_2 = 0, \ a.s.
\]

and we have proved that

\[
\lim_{n \to \infty} \sum_{i=1}^{c_n} \frac{n_i}{n} \bar{X}_{ni} = m_u, \ u = 1, ..., d, \ a.s. \tag{25}
\]

Similar arguments shows that

\[
\lim_{n \to \infty} \sum_{i=1}^{c_n} \frac{n_i}{n} \bar{X}_{ni} = E(\bar{X}(u)\bar{X}(v)), \ u, v = 1, ..., d, \ a.s. \tag{26}
\]

Now (14) follows from (25), (26).

According to an important theorem in the theory of strong approximation proved by Komlős and others (1975), one can find a random variable

\[
Z_{ni} \sim N(0, \sigma^2/n_i), \ Z_{ni}, ..., N_{nc_n} \text{ independent, and}
\]

\[
P(|e_{ni} - Z_{ni}| \geq (\log n_i)^2/n_i) \leq K \exp(-\lambda(\log n_i)^2), \ i = 1, ..., c_n \tag{27}
\]

where \( K > 0, \ \lambda > 0 \) are constants not depending on \( n, i \). Put

\[\begin{align*}
V_{ni} &= e_{ni} - Z_{ni}, \ i = 1, ..., c_n, \\
\tilde{V}_{ni} &= V_{ni} - \frac{1}{n} \sum_{i=1}^{c_n} n_i V_{ni}, \ \tilde{Z}_{ni} &= Z_{ni} - \frac{1}{n} \sum_{i=1}^{c_n} n_i Z_{ni}, \ i = 1, ..., c_n, \\
V(n) &= (\tilde{V}_{ni}, ..., \tilde{V}_{nc_n})', \ Z(n) &= (\tilde{Z}_{ni}, ..., \tilde{Z}_{nc_n})', \\
\xi_n &= Z(n)^*W(n)Z(n) - Z(n)^*W(n)\bar{X}(n)^*X(n)^*W(n)\bar{X}(n)^*X(n)^*W(n)Z(n)\ 
\end{align*}\]
Then \( \text{RSS}_n \), defined by (19), can be written as

\[
\text{RSS}_n = \xi_n + V'(n)W(n)V(n) + 2V'(n)W(n)Z(n)
\]

\[
- V'(n)W(n)\bar{x}(n)(\bar{x}(n)W(n)\bar{x}(n))^{-1}V'(n)W(n)V(n)
\]

\[
- 2V'(n)W(n)\bar{x}(n)(\bar{x}(n)W(n)\bar{x}(n))^{-1}V'(n)W(n)Z(n)
\]

\[
\bar{\xi}_n + Q_1 + 2Q_2 - Q_3 - 2Q_4. \quad (29)
\]

From the well-known normal theory and the assumptions made on model (1), it follows that the conditional distribution of \( \xi_n \) given \( \{\bar{x}_1, \ldots, \bar{x}_n\} \) is \( \chi^2_{c_n - d} \) (so the unconditional distribution of \( \xi_n \) is also \( \chi^2_{c_n - d} \)). Now we proceed to estimate \( Q_1 - Q_4 \). First take \( Q_1 \). Considering (6), we see that the event

\[
T_n = \bigcup_{i=1}^{c_n} \{|e_{n_i} - Z_{n_i}| \geq (\log n_i)^2/n_i\}
\]

has a probability \( O(n^{-2}) \). Hence, with probability one, we can assert that

\[
|V_{n_i}| \leq (\log n_i)^2/n_i, \quad i = 1, \ldots, c_n \quad (30)
\]

for \( n \) sufficiently large. Now, in case (30) is true,

\[
Q_1 = \sum_{i=1}^{c_n} n_i \bar{V}_{n_i} \leq \sum_{i=1}^{c_n} n_i V_{n_i} \leq \sum_{i=1}^{c_n} (\log n_i)^4/n_i.
\]

Considering (5) and (6), we have (again in case (30) is true)

\[
Q_1 \leq n^{1/3-\epsilon_1}(\log n)^4 n^{-2/3-\epsilon_1} = n^{-1/3-2\epsilon_1}(\log n)^4 \quad (31)
\]

which proves that

\[
\lim_{n \to \infty} Q_1 = 0, \quad \text{a.s.} \quad (32)
\]
Next come to $Q_2$. By Schwartz inequality,

$$Q_2^2 \leq Q_1 Z \sum_{i=1}^{C_n} n_i Z_{ni}^2$$

(33)

But $\sum_{i=1}^{C_n} n_i Z_{ni}^2 - x^2$. Hence $\sum_{i=1}^{C_n} n_i Z_{ni}^2 / c_n = O_p(1)$, as $n \to \infty$. From this, (5), and (31), we see that

$$Q_2 \overset{p}{\to} 0, \text{ as } n \to \infty.$$ (34)

In order to deal with $Q_3$, we shall in the following use the symbol $k_n$ to denote a quantity depending on $X_1$, ..., $X_n$, with the property that

$$\sup_n k_n < \infty \text{ with probability one.}$$

$k_n$ may assume different values in each of its appearance. We also use the symbol $I$ to denote an identity matrix with appropriate order. Under these conventions, we have by (14) that

$$\left(\frac{X_i}{n} W(n) X_i(n)\right)^{-1} \leq k_n I / n$$

(35)

for $n$ sufficiently large. Therefore,

$$Q_3 \leq V_i(n) W(n) X_i(n) (k_n I / n) \frac{X_i}{n} W(n) X_i(n) W(n) X_i(n) V(n)$$

$$\leq k_n V_i(n) W(n) \left(\sum_{i=1}^{C_n} n_i \frac{X_i}{n} W(n) X_i(n) V(n)\right).$$

By this, and using again (14), we have

$$Q_3 \leq k_n V_i(n) W(n) X_i(n) V(n) = k_n Q_1$$

(36)

with probability one for $n$ sufficiently large. By (32), we get

$$\lim_{n \to \infty} Q_3 = 0, \text{ a.s.}$$

(37)

Finally we come to $Q_4$. By Schwartz inequality,

$$Q_4^2 \leq Q_3 Z \frac{X_i}{n} W(n) X_i(n) (\frac{X_i}{n} W(n) X_i(n)^{-1} \frac{X_i}{n} W(n)) Z(n).$$

(38)
Arguing as in the case of $Q_3$, we have

$$Z'_n W(n) X(n)^{-1} X'_n W(n) Z(n) \leq k_n Z'_n W(n) Z(n)$$

with probability one for $n$ sufficiently large. From (36), (38), (39), we have (again with probability one for $n$ sufficiently large)

$$Q^2 \leq k_n Q_1 Z'_n W(n) Z(n).$$

(40)

In the process of dealing with $Q_2$, we have shown that the right hand side of (40) tends to zero in probability as $n \to \infty$. Hence, with probability one,

$$Q_4 \overset{p}{\to} 0, \text{ as } n \to \infty.$$  (41)

From (29), (32), (34), (37), (41), and the observation made on $\xi_n$, we finally reach (13). Theorem 1 is proved.
5. ESTIMATION OF THE VARIANCE OF ERROR

In this section we shall give an estimate of the variance $\sigma^2$. For our purpose, we need an estimator $\frac{\sigma^2}{n}$ which tends to $\sigma^2$ with a rate as $O(c_n^{-1/2})$.

We shall consider separately the two cases that $1^\circ, d \leq 3$, $2^\circ, d > 3$.

For the first case, we choose $c_1, c_2$ satisfying (2) and

$$c_2 = \frac{(9c_1 + 1)/12}$$

and define

$$\sigma_n^2 = \sum_{i=1}^{c_n} \sum_{j=1}^{n_i} (X_{ni}(j) - X_{nj})^2 / (N_n - C_n)$$

For the second case, we choose $c_1, c_3$, such that

$$0 < c_1 < 1/3, \quad 2(1 - c_3)/d > \frac{1}{6} - \frac{c_1}{2}$$

Define $q_n = n^{-(1-c_3)/d}$. Decompose $R^d$ into a set $\tilde{J}_n$ of supercubes having the form (3), but $\epsilon_n$ in (3) is changed to $q_n$. For each $J \in \tilde{J}_n$, write the elements in $\{X_1, \ldots, X_n\} \cap J$ as $\{X_{nj}(1), \ldots, X_{nj}(n_j)\}$. The observations of $Y$ corresponding to these $X$'s is written as $\{Y_{nj}(1), \ldots, Y_{nj}(n_j)\}$.

Finally we define

$$\sigma_n^2 = \sum_{J \in \tilde{J}_n} \sum_{j=1}^{n_j} (Y_{nj}(j) - Y_{nj})^2 / \sum_{J \in \tilde{J}_n} (n_j - 1)$$

where $Y_{nj} = \sum_{j=1}^{n_j} Y_{nj}(j)/n_j$.

THEOREM 2. Under model (1) and suppose that $H_0$ is true, then with probability 1

$$\sqrt{c_n}(\sigma_n^2 - \sigma^2) \overset{p}{\to} 0, \quad \text{as } n \to \infty.$$
Proof. We shall only give the details for the case \( d \leq 3 \), as the case \( d > 3 \) can be dealt with in an entirely similar way. Put \( b_{ni}(j) = (\overline{x}_{ni}(j) - \overline{x}_{ni}) \beta \), then \( |b_{ni}(j)| \leq \sqrt{d} \| \beta \| \varepsilon_n \). From (18), we have

\[
\sum_{j=1}^{n_i} (Y_{ni}(j) - Y_{ni})^2 = \sum_{j=1}^{n_i} (e_{ni}(j) - e_{ni})^2 + \sum_{j=1}^{n_i} b_{ni}(j) + 2 \sum_{j=1}^{n_i} b_{ni}(j)(e_{ni}(j) - e_{ni}).
\]

(47)

Put \( N'_n = N_n - C_n \). In view of (5), in order to prove (46), it is enough to verify the following assertions (all with probability one).

1. \( n \to 0 \)

\[
\frac{1}{n} \frac{1}{n} \frac{1}{N'} \frac{p^*}{n} \text{ where } T_n = \sum_{i=1}^{c_n} \sum_{j=1}^{n_i} (e_{ni}(j) - \sigma^2).
\]

2. \( n \to 0 \)

\[
\frac{1}{n} \frac{1}{n} \frac{1}{N'} \frac{p^*}{n} \sum_{i=1}^{c_n} \sum_{j=1}^{n_i} b_{ni}(j) / N'_n \to 0.
\]

3. \( n \to 0 \)

\[
\frac{1}{n} \frac{1}{n} \frac{1}{N'} \frac{p^*}{n} \sum_{i=1}^{c_n} \sum_{j=1}^{n_i} b_{ni}(j) / N'_n \to 0.
\]

4. \( S_n \to 0 \)

\[
\frac{1}{n} \frac{1}{n} \frac{1}{N'} \frac{p^*}{n} \sum_{i=1}^{c_n} \sum_{j=1}^{n_i} b_{ni}(j)(e_{ni}(j) - e_{ni}) / N'_n.
\]

Under model (1), \( T_n \) is a sum of \( N_n \) iid. random variables with zero mean and finite variance \( \sigma^2 \). By (6), we have \( N'_n \geq 2 / 3 - \epsilon_1 \geq 1 - \frac{1}{2} / n \). Hence (with probability one)

\[
p^*(|n \to 0 | T_n / N'_n | > \epsilon) \leq p^*(|T_n / N'_n | > \epsilon) \leq \frac{4}{\epsilon^2 + 3 \epsilon_1} \to 0.
\]

This proves 1. For 2, notice that \( E(n_i e_{ni}^2) = \sigma^2 \), hence

\[
\frac{1}{n} \frac{1}{n} \frac{1}{N'} \frac{p^*}{n} \sum_{i=1}^{c_n} n_i e_{ni}^2 / N'_n \leq 1 - \frac{5 \epsilon_1}{2}
\]

\[
E^* \left( \sum_{i=1}^{c_n} n_i e_{ni}^2 / N'_n \right) \leq 1 - \frac{5 \epsilon_1}{2} \quad c_n / N'_n \leq 2n / n \to 0.
\]
This proves 1°. For 2°, notice that \( E(n_i e_{n_i}^2) = \sigma^2 \), hence

\[
\frac{1}{6-\varepsilon_1/2} c_n \left( \sum_{i=1}^{n} n_i e_{n_i}^2 / N_n' \right) \leq \frac{1}{6-\varepsilon_1/2} \frac{C_n / N_n'}{2 n_{1/2}} \to 0
\]

and 2° follows. 3° is proved by noticing that \( N_n' \equiv N + 1 \), and

\[
\frac{1}{6-\varepsilon_1/2} c_n n \sum_{i=1}^{n} b_{n_i}(j) / N_n' \leq \frac{N_n}{N_n'} d \| \beta \|_2^2 c_n \frac{1}{6-\varepsilon_1/2}
\]

under the condition \( 0 < \varepsilon_1 < 1/3 \), \( 0 < \varepsilon_2 < (9\varepsilon_1 + 1)/12 \) and \( d \leq 3 \), we have \( -(2/3 - 2\varepsilon_2)/d + (1/6 - \varepsilon_1/2) < 0 \). Therefore, the right hand side tends to zero as \( n \to \infty \).

Finally we come to 4°. Since the sum \( \sum_{i=1}^{c_n} \sum_{j=1}^{c_n} d_{n_i}(j)(e_{n_i}(j) - e_{n_i}) \) can be written in a form \( \sum_{i=1}^{c_n} \sum_{j=1}^{n_i} d_{n_i}(j)e_{n_i}(j) \) with \( |d_{n_i}(j)| \leq |b_{n_i}(j)| \), it follows that

\[
\text{Var}^*(S_n) \leq n \sum_{i=1}^{c_n} \sum_{j=1}^{n_i} b_{n_i}(j)\sigma^2 / N_n'.
\]

It is easy to see that the right hand side is less than a constant \( g_n \) which does not depend on \( \overline{x}_1, \ldots, \overline{x}_n \), and \( g_n \to 0 \) as \( n \to \infty \). Also, \( \text{E}^*(S_n) = 0 \). Therefore,

\[
E*S_n^2 = \text{Var} S_n + (E(S_n))^2
\]

\[\leq g_n \to 0,\]

from which 4° follows, and Theorem 2 is proved.

A more tedious argument allows us to prove that

\[
\lim_{n \to \infty} \sqrt{n}(\overline{C_n}(\sigma^2 - \sigma) = 0, \quad \text{as.}
\]
6. A LARGE SAMPLE DETERMINATION OF C IN TEST (12)

We are now in a position to prove the following theorem, which gives a large sample determination of C in test (12).

**THEOREM 3.** Under model (1) and suppose that the null hypothesis \( H_0 \) is true, the distribution of the statistic

\[
T_n^* = \sqrt{2RSS_n} \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} - \sqrt{2(c_n - d)}
\]

converges to the standard normal \( N(0,1) \) as \( n \to \infty \).

**Proof.** From Theorem 1 it follows that the conditional distribution of \( T_n \) given \( \{S_1, S_2, \ldots\} \) tends to \( N(0,1) \), with probability one. Since \( \sigma_{\bar{X}} < \sigma \), we have

\[
|\bar{X} - \mu| \leq |\sigma_{\bar{X}} - \sigma|/\sigma_{\bar{X}} \leq \sqrt{2 RSS_n |\sigma_{\bar{X}} - \sigma|/\sigma_{\bar{X}}^2}.
\]

By Theorem 1, with probability one \( RSS_n/c_n = O_p(1) \). Therefore, from (48) and Theorem 2, we have \( T_n - T_n^* \to 0 \), with probability one. Hence we have proved that with probability one, the conditional distribution of \( T_n^* \) given \( \{S_1, S_2, \ldots\} \) tends to \( N(0,1) \), as \( n \to \infty \), and the same holds true for the unconditional distribution of \( T_n^* \). Theorem 3 is proved.

From Theorem 3 it follows that if we choose

\[
C = \left\{ \sqrt{2(c_n - d)} + u_{\alpha} \right\}^2/2
\]

in (12), where \( u_{\alpha} \) is defined by

\[
\int_{-\infty}^{u_{\alpha}} x^2/2dx/\sqrt{2\pi} = 1 - \alpha.
\]

then the test is asymptotically similar with size \( \alpha \).
7. CONSISTENCY OF THE TEST

Denote by \( Q_n(F) \) the power of the test defined by (12) and (49), where \( F \) is the joint distribution of \((\mathbf{x}, Y)\). The following theorem shows that this test is consistent over a wide class of alternatives.

**Theorem 4.** Suppose that the following conditions are satisfied:

1. \( E\|\mathbf{x}\|^2 < \infty, \ \text{COV}(\mathbf{x}) > 0, \ \text{E}Y^2 < \infty \).
2. The distribution of \( \mathbf{x} \) has no singular component.
3. The closure \( \overline{A}_1 \) of the set \( A \) of all atoms of \( \mathbf{x} \) possesses Lebesgue measure zero.
4. The closure \( \overline{A}_2 \) of all discontinuity points of \( m(x) = E(Y|\mathbf{x} = x) \) possesses Lebesgue measure zero.
5. \( P(Y < y|\mathbf{x} = x) \) depends only on \( y - m(x) \).
6. There exists no linear function \( \alpha + \beta'x \) such that \( P(m(\mathbf{x}) = \alpha + \beta'\mathbf{x}) = 1 \).

Then, denoting the distribution of \((\mathbf{x}, Y)\) by \( F \), we have

\[
\lim_{n \to \infty} Q_n(F) = 1.
\]

**Proof.** First choose \( d+2 \) points \( t_1, \ldots, t_{d+2} \), such that the points

\[
(t'_1, m(t_1)), \ldots, (t'_{d+2}, m(t_{d+2}))
\]

do not lie on a common hyperplane in \( \mathbb{R}^{d+1} \). If \( \mathbf{x} \) possesses no absolute continuous component, then, by conditions 1° and 6°, we can choose such points \( t_1, \ldots, t_{d+2} \) from the atom set \( A \). If \( \mathbf{x} \) has an absolute continuous component with density \( f \), then, according to conditions 1°, 3°, 4°, 6°, we can choose \( t_1, \ldots, t_{d+2} \) from the set \( \mathbb{R}^d - (\overline{A}_1 \cup \overline{A}_2) \). With the additional
property that there exists two constants $h_1 > 0$, $h_2 > 0$, such that for any $I = ((x^{(1)}), \ldots, x^{(d)})$: $b_1 \leq x^{(i)} < b_1 + h$, $i = 1, \ldots, d$ with $h \in (0, h_1)$, we have

$$\int_I f dx \geq h_2 h^d. \quad (52)$$

For large $n$, we can find $d+2$ distinct supercubes $J_{n1}, \ldots, J_{nd+2}$ in $J_n$ (see (3)), such that $t_i \in J_{ni}$, $i = 1, \ldots, d+2$. Denote the elements of $\{\bar{x}_1, \ldots, \bar{x}_n\} \cap J_{ni}$ by $\bar{x}_{ni}(j)$, $j = 1, \ldots, n_i$. In the same way as we did in Section 2, we define $Y_{ni}(j)$, $\bar{x}_{ni}$, $Y_{ni}$. Fit a hyperplane

$$y = \alpha_n + \beta_n' x$$

to the $d+2$ points $(\bar{x}_{ni}, Y_{ni})$, $i = 1, \ldots, d+2$, by the weighted least squares method with weights $n_1, \ldots, n_{d+2}$. The weighted residual sum of squares is

$$R_n = \sum_{i=1}^{d+2} n_i (Y_{ni} - \alpha_n - \beta_n' \bar{x}_{ni})^2.$$ 

For definiteness, we assume that the points $t_1, \ldots, t_{d+2}$ are chosen from the set $R_d - (\bar{A}_1 \cup \bar{A}_2)$. In this case, from (52) and the fact that $\gamma_n = n^{-1/3+\varepsilon_2}$, with probability one we have

$$n_i \geq h_2^{2/3+\varepsilon_2} / 2, \quad i = 1, \ldots, d+2 \quad (53)$$

for $n$ sufficiently large.

Given $\delta_1 > 0$, we can find $n_0$ such that for $n \geq n_0$

$$\sup \{m(x) : x \in J_{ni}\} - \inf \{m(x) : x \in J_{ni}\} < \delta_1 / 2, \quad i = 1, \ldots, d+2. \quad (54)$$

Use again $P* = P*(\bar{x}_1, \bar{x}_2, \ldots)$ the conditional probability distribution given $\bar{x}_1, \bar{x}_2, \ldots$. By (53), (54), it is easy to see that

$$p*(|Y_{ni} - m(t_i)| \geq \delta_1) \leq \delta_1^2 2h_2^{-1/2} \alpha^2 n^{-2/3+\varepsilon_2}, \quad i = 1, \ldots, d+2 \quad (55)$$
with probability one for \( n \) sufficiently large.

Since the points (51) do not lie simultaneously on any hyperplane, and since \( \overline{X}_{ni} \to t_i, i = 1, \ldots, d+2, \) as \( n \to \infty, \) it follows that there exists constant \( \delta_2 > 0 \) not depending on \( n \) and \( (\overline{X}_i, Y_i), i = 1, 2, \ldots, \) such that

\[
\sup \{|m(t_i) - a_n - \alpha_i\overline{X}_{n-ni}|: i = 1, \ldots, d+2\} \geq \delta_2
\]

for \( n \) sufficiently large. Choose \( \delta_1 = \delta_2/2 \) in (54). Then if the inequalities \( |Y_{ni} - m(t_i)| < \delta_1 \) all hold true, we shall have

\[
R_n \geq \delta_1^2 \min(n_1, \ldots, n_{d+2})
\]

for \( n \) sufficiently large. Summing up all the arguments above, we reach the following conclusion:

\[
\lim_{n \to \infty} P(R_n > n^{2/3}) = 1. \tag{57}
\]

Obviously, we have \( RSS_n \geq R_n. \) Hence by (57)

\[
\lim_{n \to \infty} P(RSS_n > n^{2/3}) = 1. \tag{58}
\]

Now consider \( \sigma_n^2 \) defined by (45). The equality (47) still holds if we redefine \( b_{ni}(j) = m(X_{ni}(j)) - m_{ni}, \) with \( m_{ni} = \frac{\sum_{j=1}^{n_i} m(X_{ni}(j))}{n_i}. \) We have

\[
\sum_{i=1}^{c_n} \sum_{j=1}^{n_i} b_{ni}(j)/N_n \leq \sum_{i=1}^{c_n} \sum_{j=1}^{n_i} m^2(X_{ni}(j))/N_n \leq \sum_{i=1}^{n} m^2(X_i)/N_n. \tag{59}
\]

It can be shown that there exists constant \( q > 0, \) such that with probability one we have

\[
N_n \geq qn \tag{60}
\]
for \( n \) sufficiently large (see the end of this section). From (59), (60), using the strong law of large numbers, we get

\[
\limsup_{n \to \infty} \sum_{i=1}^{c_n} \sum_{j=1}^{n_i} b_{n_i}(j)/N_n \leq q^{-1}E(m^2(X)) \leq q^{-1}E(Y^2) < \infty, \text{ a.s.} \quad (61)
\]

Employing the method of handling \( S_n \) in the proof of Theorem 2, we get

\[
P^*(| \sum_{i=1}^{c_n} \sum_{j=1}^{n_i} b_{n_i}(j)(e_{n_i}(j) - e_{n_i})|/N_n \geq \epsilon) \leq \epsilon^{-2}N^{-2}a^2 \sum_{i=1}^{c_n} \sum_{j=1}^{n_i} b_{n_i}(j) \leq \epsilon^{-2}N^{-2}a^2 \sum_{i=1}^{n} m^2(\bar{X}_i) \to 0, \text{ a.s.} \quad (62)
\]

Finally, we have

\[
\sum_{i=1}^{c_n} \sum_{j=1}^{n_i} (e_{n_i}(j) - e_{n_i})^2/N_n \xrightarrow{p*} 0. \quad (63)
\]

Combining (61)-(63), we see that there exists constant \( M < \infty \), such that

\[
\lim_{n \to \infty} P(\sigma_n^2 < M) = 1. \quad (64)
\]

From (58), (64), it follows that

\[
\lim_{n \to \infty} P(RSS_n/\sigma_n^2 > n^{2/3}M^{-1}) = 1. \quad (65)
\]

But by (5), we have

\[
C < (\sqrt{2(c_n - d) + u_o})^2 < 8c_n < 8n^{1/3} < M^{-1}n^{2/3}
\]

for \( n \) sufficiently large. This means

\[
\lim_{n \to \infty} P(RSS_n/\sigma_n^2 > C) = 1
\]
which is no other than (50), and the theorem is proved.

Proof of (60). If \( \mathfrak{X} \) has at least one atom, then clearly (60) holds.

Otherwise we choose a sufficiently large such that
\[
\int_I f dx > 1/2, \quad I = \{(x(1), \ldots, x(d)) : |x(i)| \leq a, i = 1, \ldots, d\}.
\]
I contains \((2a)^{-1/3-\varepsilon_2} \) elements of \( J^* \). Put
\[
H_n = \{J : J \in J^*, J \subset I, \mu(J) \geq 4^{-1}(2a)^{-d_n} \}^{(1/3-\varepsilon_2)}
\]
Evidently we have
\[
\sum_{J \in H_n} \mu(J) \geq 1/4. \tag{66}
\]
Using (16), and noticing that \( H_n \) has at most \((2a)^{-d_n} \) members, we obtain
\[
P(|\#(J)/n - \mu(J)| \geq 8^{-1}(2a)^{-d_n} \tag{1/3-\varepsilon_2}, \text{ for at least one } J \in H_n)
\leq (2a)^{-d_n} \frac{1/3-\varepsilon_2}{2} \exp(-cn^{1/3+2\varepsilon_2}) = O(n^{-2}) \tag{67}
\]
where \( c > 0 \) does not depend on \( n \). From (67) we see that, with probability one, we have
\[
|\#(J)/n - \mu(J)| < 8^{-1}(2a)^{-d_n} \tag{1/3-\varepsilon_2}, \text{ for all } J \in H_n \tag{68}
\]
for \( n \) sufficiently large. But when (68) holds we shall have
\[
N_n \geq \sum_{J \in H_n} \#(J) \geq n \sum_{J \in H_n} [\mu(J) - 8^{-1}(2a)^{-d_n} \tag{1/3-\varepsilon_2}]
\geq n \sum_{J \in H_n} \frac{1}{2} \mu(J) \geq n/8
\]
on account of the fact that \( \varepsilon_2 > \varepsilon_1 \) and (66). This shows (60) holds with \( q = 1/8 \).
8. ASYMPTOTIC POWER OF THE TEST

Suppose that \((\overline{x}, y(1)), (\overline{x}, y(2)), \ldots\) is a sequence of alternatives, where \(m(n)(x) \triangleq E(y(n)|\overline{x} = x)\) approaches some linear function \(\alpha + \beta'x\) as the sample size \(n\) increases. Specifically we assume that

\[
m(n)(x) = \alpha + \beta'x + m_0(x)/g_n
\]

(69)

where \(g_n \to \infty\) as \(n \to \infty\). We are interested in finding such \(g_n\) for which \(Q_n(F(n))\) tends to some limit greater than the size \(\alpha_0\) and smaller than one (\(F(n)\) is the distribution of \((\overline{x}, y(n))\)).

THEOREM 5. Suppose that \((\overline{x}_1, y_1), \ldots, (\overline{x}_n, y_n)\) are iid. samples of \((\overline{x}, y(n))\), where \(y_i = m(n)(\overline{x}_i) + e_i, i = 1, \ldots, n, m(n)(x)\) is defined by (69) and \(\overline{x}_i, e_i, i = 1, \ldots, n\) satisfy the conditions specified in model (1). Also, assume that \(EM^2_0(\overline{x}) < \infty\), and

\[
\lim_{n \to \infty} g_n/(\sqrt{n} c_n^{-1/4}) = h^{-1} e (0, \infty).
\]

(70)

Define

\[
l = u_{\alpha_0} - \frac{h^2}{\sqrt{2}} \min_{\alpha, \beta} E[m_0(\overline{x}) - \alpha - \beta'\overline{x}]^2.
\]

(71)

Then we have

\[
\lim_{n \to \infty} Q_n(F(n)) = \frac{1}{\sqrt{2\pi}} \int_{-l}^{\infty} e^{-t^2/2} dt.
\]

(72)

Proof. As the detailed proof is tedious, we give only a sketch of the main points of the proof.

1. Since the linear hypothesis \(H_0\) is not assumed to be true, the residual sum of squares \(\text{RSS}_n (9)\) cannot be reduced to (11). Under this circumstance, instead of (27), we have to use a strong approximation to \(Y_{ni}:\)
where \( Z_{ni} = N(u_{ni}, \sigma^2/n_i) \), with
\[
u_{ni} = \alpha + \beta \frac{\bar{X}_{ni}}{\sigma^2} + \sum_{j=1}^{c_i} m_0(\bar{X}_{ni}(j))/(n_i g_n).
\]

2. Put \( Z_{ni}^* = Z_{ni}^* = \alpha - \beta \frac{\bar{X}_{ni}}{\sigma^2} \). It is easy to verify that the expression (29) still holds true with \( Z_{ni} \) so defined.

3. We have to verify that (32), (34), (37) and (41) still holds true in the present case. For \( Q_1 \), nothing has been changed and (31) is true.

For \( Q_2 \), notice that \( Z_{ni}^*W_{ni}Z_{ni}^* \) obeys a non-central \( \chi^2 \) distribution with degree of freedom \( c_i \) and non-central parameter
\[
\delta_n^2 = g_n^{-2} \sum_{i=1}^{c_i} \left( \sum_{j=1}^{n_i} m_0(\bar{X}_{ni}(j))/n_i \right)^2.
\]

It is easily seen that as \( n \to \infty, \delta_n^2 = ng_n^{-2}[Em_0(\bar{X}) + o(1)] \). Since we have \( ng_n^{-2} = O(\sqrt{C}) \) in view of (70), by (5), we have
\[
E(Z_{ni}^*W_{ni}Z_{ni}^*) = o(C_n) = o(n^{1/2})
\]
and (34) follows from (31) and (73). \( Q_3 \) and \( Q_4 \) can be dealt with in the same manner.

4. Given \( \bar{X}_1, \bar{X}_2, \ldots \), the conditional distribution of \( \xi_n/\sigma^2 \) is
\[
\chi^2_{c_i, \delta_n^2},
\]
where
\[
\delta_n^2 = ng_n^{-2} \left( \sum_{i=1}^{c_i} \frac{n_i}{n} \nu_{ni} \right) - \left( \frac{c_i^{n_i} \bar{X}_{ni} \bar{X}_{ni} - \bar{X}_{ni} \bar{X}_{ni}}{\bar{X}_{ni} \bar{X}_{ni}} \right)
\]
where
\[
\nu_{ni} = \sum_{j=1}^{n_i} m_0(\bar{X}_{ni}(j))/n_i - m_0, \quad m_0 = Em_0(\bar{X}).
\]

From (14) and (70), it is easily seen
\[
\delta_n^2 = \sqrt{c_n} \{ \text{Var}(m_0(X)) - \mathbb{E}(m_0(X) - m_0)\mathbb{X}'\} \mathbb{X}^{-1} \mathbb{E}(m_0(X) - m_0)\mathbb{X} + o(1)h^2 \\
= \sqrt{c_n} \{ \min_{\alpha, \beta} \mathbb{E}(m_0(X) - \alpha - \beta \mathbb{X})^2 + o(1)\}h^2.
\] (74)

Given \( \{\overline{X}_1, \overline{X}_2, \ldots\} \), the conditional distribution of \( \xi_n/\sigma^2 \) is the same as the sum of iid. variables \( (\eta_i + \delta_\sigma/\sqrt{c_n})^2, i = 1, \ldots, c_n \), where \( \eta_1, \ldots, \eta_{c_n} \) are iid., \( \eta_i \sim N(0,1) \). The central limit theorem can be applied to this case, which gives

\[
\frac{\xi_n/\sigma^2 - (c_n + \delta_n^2)}{\sqrt{2c_n + 4\delta_n^2}} \rightarrow N(0,1).
\] (75)

5. Define \( \sigma_n^2 \) as in (45). We have to verify that (46) still holds true in the present case. Starting from (47), in which \( b_{n_1}(j) \) is redefined as

\[
b_{n_1}(j) = (\overline{X}_{n_1}(j) - \overline{X}_{n_1})'\beta + g_n^{-1}u_{n_1}(j)
\]

\[
u_{n_1}(j) = m_0(\overline{X}_{n_1}(j)) - \sum_{j=1}^{n_1} m_0(\overline{X}_{n_1}(j))/n_i
\]

it is easily seen that the first and third term in the right hand side of (47) can be handled in the same way as before, and it suffices to verify that

\[
c_n \sum_{i=1}^{n} \frac{n_i}{\sum_{i=1}^{n_1} b_{n_1}(j)/n} = o(\sqrt{c_n}).
\] (76)

Considering (46), (75) reduces to

\[
g_n^{-2} \sum_{i=1}^{c_n} \frac{n_i}{\sum_{i=1}^{n_1} u_{n_1}(j)/n} = o(\sqrt{c_n}).
\] (77)
Since
\[ \sum_{i=1}^{c_n} \sum_{j=1}^{n_l} u_{n_l}^2(j)/n \leq \sum_{i=1}^{c_n} \sum_{j=1}^{n_l} m_0^2(\bar{X}_{n_l}(j))/n \leq \sum_{i=1}^{n} m_0^2(\bar{X}_i)/n + E(m_0^2(\bar{X})), \quad \text{a.s.} \]

Further, by (5) and (70), we have
\[ g_n^{-2} \leq \frac{1}{2} n^{-1} \sqrt{c_n} \leq n^{-5/6} \]
for \( n \) sufficiently large. This proves (77).

6. From (45) and (75), we have
\[ \frac{\xi_n/\sigma_n^2 - (c_n + \delta_n^2)}{\sqrt{2c_n + 4\delta_n^2}} \rightarrow N(0,1) \]
as \( n \to \infty \), under the conditional distribution given \((\bar{X}_1, \bar{X}_2, \ldots)\). From (49), (74), we have
\[ \lim_{n \to \infty} \frac{((2(c_n - d)^2 + u_{\alpha_0})^2/2 - (c_n + \delta_n^2)}{\sqrt{2c_n + 4\delta_n^2}} = \xi. \]

This leads to (72), concluding the proof of the theorem.

The statistical meaning of the quantity \( \xi \) is clear, since
\[ \min_{\alpha, \beta} E[\delta_0(X) - \alpha - \beta'X]^2 \]
measures the distance of \( \delta_0(X) \) to linearity in the MSE sense. As this distance increases, \( \xi \) decreases, and the asymptotic power of the test (right hand side of (72)) increases.

The theorem indicates that, roughly speaking, our test has a discrimination power for those \( Y \), for which the distance (in the sense of MSE) of \( E(Y|X = x) \) to the linear function space is not smaller than \( O(n^{-1/2}c_n^{1/4}) \). Under quite general conditions, we can prove that \( c_n \to 0(1/3-\varepsilon_2) \), and \( O(n^{-1/2}c_n^{1/4}) = O(n^{-5/12-\varepsilon_2/4}) \). In principle, \( \varepsilon_2 \) can be chosen arbitrarily near 1/3. So the order for possible discrimination can be made arbitrarily
near $O(n^{-1/2})$, but this order cannot be reached, unless $\mathcal{X}$ is purely atomic with a finite number of atoms.
REFERENCES


END
12-87
DTIC