⟨Research Paper⟩

Journal of the Korean
Statistical Society
Vol. 18, No.1, 1989

A Nonparametric Method for Nonlinear Regression Parameterers[+]

Hae Kyung Kim*

ABSTRACT

This paper is concerned with the development of a nonparametric procedure for the statistical inference about the nonlinear regression parameters. A confidence region and a hypothesis testing procedure based on a class of signed linear rank statistics are proposed and the asymptotic distributions of the test statistic both under the null hypothesis and under a sequence of local alternatives are investigated. Some desirable asymptotic properties including the asymptotic relative efficiency are discussed for various score functions.

1. Introduction and Assumptions

Consider the following nonlinear regression model for a univariate response Y,

$$Y_t = f(x_t, \theta) + \varepsilon_t, \quad t = 1, \cdots, n \tag{1.1}$$

where $x_t \in \Xi \subset R^m$ denotes the tth fixed known input vector, $\theta \in R^p$ is the parameter vector from a compact parameter space $\Theta \subset R^p$, f is a continuous function $f : R^{m+p} \rightarrow R^1$ and ε_t are independent identically distributed (i.i.d.) random errors with the common distribution function G. Suppose G possesses an absolutely continuous and symmetric (about zero) probability density function (pdf) g with a finite Fisher information

+ This research was supported by the Korean Ministry of Education(1988).

* Deparment of Mathematics, Yonsei University, Seoul 120, Korea.

$$\int_{-\infty}^{\infty} \left\{ \frac{g'(x)}{g(x)} \right\}^2 dG(x) \ (< \infty) \tag{1.2}$$

where $g'(x) = dg(x)/dx$ exists a.e..

Frequently the effect of the input vector x_t on the response Y_t is adequately approximated by a linear function

$$f(x_t, \theta) = X_t' \theta \tag{1.3}$$

where $X_t = (x_{t1}, \cdots, x_{tp})'$ and $x_{t1} = 1$.

The parameter θ is unknown and the regression problem is to make inference about θ in some optimal way on the basis of observations on Y_t, and x_t, $t = 1, \cdots, n$. Considerable attention has been devoted in the literature to the parameter estimation of the model (1.1), especially to the least squares estimation under the assumption that the errors ε_t are i.i.d.; Asymptotic results for nonlinear least squares estimation are given by Jennrich (1969), Malinvaud (1970) and Wu (1981).

The purpose of this paper is twofold. First, we develop a nonparametric procedure for the estimation and testing hypotheses about the nonlinear regression parameter θ based on signed linear rank statistics. Secondly, we evaluate the asymptotic efficiency of the proposed nonparametric procedure in comparison to an optimum parameter procedure in the sense of the least squares principle. For these, in Section 2 we construct a class of signed linear rank statistics suitable for deriving a confidence region and test statistics and investigate the asymptotic behavior of the statistics. An asymptotic confidence region and test procedure for the hypothesis about θ are proposed in Section 3. In order to study the asymptotic power properties of the proposed test, the limiting distribution of the test statistic is derived in Section 4 under a sequence of local alternatives tending to the null hypothesis at a suitable rate. Finally the asymptotic relative efficiency (ARE) of the proposed test with respect to the classical chi-square test based on the least squares estimators is derived in Section 5. It is also verified that the ARE is independent of the nature of the regression functions but depends on the error structure of the regression model.

Throughout the paper we make the following assumptions: for the model (1.1),

A1. $\dfrac{\partial f(x,\theta)}{\partial \theta_u}$, $u = 1, \cdots, p$, are continuous on $\Xi \times \Theta$, where θ_u is the uth element of θ.

A2. For all $\theta \in \Theta$, $c_{uv} = \lim\limits_{n \to \infty} \dfrac{1}{n} \sum\limits_{t=1}^{n} \dfrac{\partial f(x_t, \theta)}{\partial \theta_u} \dfrac{\partial f(x_t, \theta)}{\partial \theta_v}$ exists for $u, v = 1, \cdots, p$, and the matrix $C(\theta)=(c_{uv})_{u,v=1,\cdots,p}$ is nonsingular.

A3. For all $\theta\in\Theta$,

$$\lim_{n\to\infty}\frac{\max\limits_{1\leq t\leq n}\left[\dfrac{\partial f(x_t,\theta)}{\partial\theta_u}\dfrac{\partial f(x_t,\theta)}{\partial\theta_v}\right]}{\sum\limits_{t=1}^{n}\dfrac{\partial f(x_t,\theta)}{\partial\theta_u}\dfrac{\partial f(x_t,\theta)}{\partial\theta_v}}=0$$

for u, v= $1,\cdots,p$.

We now make some remarks on the assumptions:

Remark 1.1. In assumptions A1 through A3 the existence of the limits is essential and these conditions may be cumbersome to verify in practice. One simple way of checking the conditions would be to regard the input vector $\{x_n\}$ as a random samples from some distribution function $H(x)$ defined on Ξ. Then

$$c_{uv}=\int_{\Xi}\frac{\partial f(x,\theta)}{\partial\theta_u}\frac{\partial f(x,\theta)}{\partial\theta_v}\,dH(x)$$

provided the integral exists.

Remark 1.2. It is not difficult to see that the assumptions are sufficiently general to cover a wide class of nonlinear regression functions including most of the commonly interesting ones such as:

(i) The p-compartment model

$$f(x,\theta)=\sum_{i=1}^{p}\alpha_i e^{\beta_i x}$$

with $\theta=(\alpha_1,\cdots,\alpha_p,\beta_1,\cdots,\beta_p)\in\Theta\subset R^{2p}$ where $\alpha_i,\beta_i\geq 0$, $i=1,\cdots,p$, and $x\geq 0$.

(ii) The sinusoidal models

$$f(x,\theta)=\alpha e^{-\beta x}\cos(\omega x+\phi)$$

with $\theta=(\alpha,\beta,\omega,\phi)$ where $\alpha,\beta\geq 0$, $x\geq 0$.

2. Signed Linear Rank Statistics and their Asymptotic Normality

In this section we shall propose a class of signed linear rank statistics and prove that under appropriate conditions the statistic has a normal distribuion in the limit.

Let $D_i(\theta)= Y_i - f(x_i,\theta)$, $i= 1,\cdots,n$, and $R_i(\theta)$ be the rank of $|D_i(\theta)|$ among $\{|D_t(\theta)|, 1\leq t\leq n\}$. Let $\phi(u)= \phi_1(u)-\phi_2(u)$, $0 < u < 1$, where $\phi_i(u)$, $i= 1,2$, are both nondecreasing, square integrable functions on $(0,1)$. Let $y= (Y_1,\cdots,Y_n)$. Define the statistic

$$S(y,\theta)= (S_1(y,\theta),\cdots,S_p(y,\theta))' \tag{2.1}$$

where

$$S_u(y,\theta)= \frac{1}{n} \sum_{t=1}^{n} \frac{\partial f(x_t,\theta)}{\partial\theta_u} \operatorname{Sgn}(D_t(\theta))\ \phi\left[\frac{R_t(\theta)}{n+1}\right], \tag{2.2}$$

$u= 1,\cdots,p$, where $\operatorname{Sgn}(x)= 1$ or -1 according as $x \geq 0$ or $x < 0$.

The choice of the regression constants, $\frac{\partial f(x_t,\theta)}{\partial\theta_u}$, $u= 1,\cdots,p$, is inspired by the normal equations of the method of least squares in the model (1.1).

The following theorem gives us the asymptotic normality of $S(y, \theta)$.

Theorem 2.1. Consider the model (1.1) and suppose that assumptions A1 through A3 are satisfied. Then $\sqrt{n}\,S(y,\theta)$ has asymptotically a p-variate normal distribution with mean vector zero and variance-covariance matrix $V(\theta)$ where

$$V(\theta)=\left[\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \frac{\partial f(x_t,\theta)}{\partial\theta_u} \frac{\partial f(x_t,\theta)}{\partial\theta_v} \int_0^1 \phi^2(w)\, dw\right]_{u,v=1,\cdots,p} \tag{2.3}$$

Proof. Let $\lambda= (\lambda_1,\cdots,\lambda_p)'$ be any arbitrary vector. Define a statistic which is slight different from $S_u(y,\theta)$ as

$$S_u^{(0)}(y,\theta)= \frac{1}{n} \sum_{t=1}^{n} \frac{\partial f(x_t,\theta)}{\partial\theta_u} \operatorname{Sgn}(D_t(\theta))\ \phi[G^*(|D_t(\theta)|)] \tag{2.4}$$

where G^* is the distribution function of $|D_t(\theta)|$. Let

$$S^{(0)}(y,\theta)= (S_1^{(0)}(y,\theta),\cdots,S_p^{(0)}(y,\theta))'.$$

For simplicity write $S = S(y,\theta)$, $S^{(0)} = S^{(0)}(y,\theta)$ and $V = V(\theta)$. We first prove that $\sqrt{n}\,\lambda' S$ and $\sqrt{n}\,\lambda' S^{(0)}$ are asymptotically equivalent random variables and then $\sqrt{n}\,\lambda' S^{(0)}$ converges in law to a normal distribution with mean zero and variance $\lambda' V \lambda$. Note that the vectors $(R_1,\cdots,R_n)$, $(|D_1(\theta)|,\cdots,|D_n(\theta)|)$ and $(\mathrm{Sgn}(D_1(\theta)),\cdots,\mathrm{Sgn}(D_n(\theta)))$ are mutually independent and $E_\theta\,\mathrm{Sgn}(D_t(\theta)) = 0$ for all t. Hence we obtain

$$E_\theta\{\sqrt{n}\,\lambda'(S - S^{(0)})\}^2 = \mathrm{Var}_\theta\{\sqrt{n}\,\lambda'(S - S^{(0)})\}$$
$$= \frac{1}{n}\sum_{t=1}^{n}\left[\sum_{u=1}^{p}\lambda_u \frac{\partial f(x_t,\theta)}{\partial \theta_u}\right]^2 E_\theta\left\{\phi\left[\frac{R_t(\theta)}{n+1}\right] - \phi\left[G^*(|D_t(\theta)|)\right]\right\}^2 .$$

Note that

$$E_\theta\left\{\phi\left[\frac{R_t(\theta)}{n+1}\right] - \phi\left[G^*(|D_t(\theta)|)\right]\right\}^2 = E_\theta\left\{\phi\left[\frac{R_1^*}{n+1}\right] - \phi(U_1)\right\}^2$$

which converges to zero, where $U_t = G^*(|D_t(\theta)|)$ are independent uniform (0,1) random variables and R_t^* are the ranks of U_t among $\{U_t,\ 1 \le t \le n\}$. Combining A2 and the above result yields the first part of the proof.

To prove the asymptotic normality of $\sqrt{n}\,\lambda' S^{(0)}$, note that $\sqrt{n}\,\lambda' S^{(0)} = \sum_{t=1}^{n} Z_t^{(0)}$

where

$$Z_t^{(0)} = \frac{1}{\sqrt{n}}\sum_{u=1}^{p}\lambda_u \frac{\partial f(x_t,\theta)}{\partial \theta_u}\,\mathrm{Sgn}(D_t(\theta))\,\phi(U_t).$$

Then, since $Z_t^{(0)}$ are independent, we can easily check that the Lindeberg condition is satisfied under our assumptions. Furthermore, we can also show that $E_\theta[\sqrt{n}\,\lambda' S^{(0)}]$ converges to zero, and $\mathrm{Var}_\theta[\sqrt{n}\,\lambda' S^{(0)}]$ converges to $\lambda' V \lambda$ as $n \to \infty$. This completes the proof. ■

3. Confidence Region and Test Statistic

In this section we shall consider asymptotic confidence region for the parameter θ in the model (1.1), and test procedure for the hypothesis about θ based on $S(y,\theta)$defined in Section 2. The asymptotic power properties of the test will be considered in later sections.

The asymptotic normality of $\sqrt{n}\ (S_1(y,\theta),\cdots,S_p(y,\theta))'$, derived in Theorem 2.1, under the regularity conditions suggests the use of the pivotal quantity of the form

$$Q_n(y,\theta) = n\, S'(y,\theta)\, V_n^{-1}(\theta)\, S(y,\theta) \tag{3.1}$$

where $S(y,\theta) = (S_1(y,\theta),\cdots,S_p(y,\theta))'$ and $V_n(\theta)$ is the p x p matrix with (u,v)th elements

$$\frac{1}{n}\sum_{t=1}^{n}\frac{\partial f(x_t,\theta)}{\partial\theta_u}\;\frac{\partial f(x_t,\theta)}{\partial\theta_v}\;\int_0^1 \phi^2(w)dw.$$

The following theorem gives the large sample distribution of $Q_n(y,\theta)$.

Theorem 3.1. Under the conditions of Theorem 2.1, $Q_n(y,\theta)$ has asymptotically a central chi-square distribution with p degrees of freedom.

Proof. Theorem 3.1 follows immediately from Theorem 2.1. ■

By reference to the null limiting distribution of $Q_n(y,\theta)$, we define $C_{1-\alpha}(\theta)$ as the set of θ such that

$$S'(y,\theta)\, V_n^{-1}(\theta)\, S(y,\theta) \leq \frac{1}{n}\,\chi^2_{1-\alpha}(p)$$

where $\chi^2_{1-\alpha}(p)$ is the $(1-\alpha)$th the quantile of the chi-square distribution. Then, for n large $C_{1-\alpha}(\theta)$ provides a $100(1-\alpha)$ percent confidence region for θ.

We also propose the following large sample α-size test procedure for hypothesis

$$H_0:\ \theta = \theta_0 \tag{3.2}$$

against H_1: $\theta \neq \theta_0$, where θ_0 is a specified vector: Reject or accept H_0 according as $Q_n(y,\theta_0) \geq$ or $< \chi^2_{1-\alpha}(p)$.

4. Asymptotic Distribution of Q_n under Contiguous Alternatives

In order to study the asymptotic power properties of the test considered in Section 3, it is necessary to study the limiting distribution of the test statistic under a sequence of local alternatives tending to the null hypothesis at a certain rate. In this section we shall specifically consider the asymptotic distribution of $Q_n(y,\theta)$, using the contiguity principle(LeCam, 1960), under the following sequence of local alternative hypotheses H_n defined by

$$H_n: \theta = \theta_n \quad \text{where } \theta_n = \theta_0 + \frac{h}{\sqrt{n}} \tag{4.1}$$

where $h \in R^p$ with each component of h's is positive.

Let (P_0, P_n) be a sequence of two probability distributions of $(Y_1,\cdots,Y_n)$ under H_0: $\theta = \theta_0$ and H_n respectively. It is well known that the contiguity of $\{P_n\}$ to $\{P_0\}$ provides the asymptotic normality of a statistic $\sqrt{n}\lambda' S(y,\theta_0)$ under the given sequence of probability distributions $\{P_n\}$ if the statistic has an asymptotic normal distribution under the sequence of probability distributions $\{P_0\}$.

We shall define some auxiliary statistics. Suppose p_0 and p_n are product densities of $(Y_1,\cdots,Y_n)$ corresponding to P_0 and P_n respectively. Then the likelihood ratio statistic L_n for P_0 vs. P_n is

$$\log L_n = \sum_{t=1}^{n} \log \frac{g(Y_t - f(x_t,\theta_n))}{g(Y_t - f(x_t,\theta_0))} \tag{4.2}$$

where $g(y_t - f(x_t, \theta_0)) > 0$. Denote

$$W_n = 2\sum_{t=1}^{n} \left\{ \left[\frac{g(Y_t - f(x_t,\theta_n))}{g(Y_t - f(x_t,\theta_0))} \right]^{\frac{1}{2}} - 1 \right\}, \tag{4.3}$$

and

$$\Psi(u) = -\frac{g'(G^{-1}(u))}{g(G^{-1}(u))}, \quad 0<u<1$$

where $G(x) = \int_{-\infty}^{x} g(y)dy$.

For the contiguity of $\{P_n\}$ to $\{P_0\}$, we need the following additional assumptions: for the nonlinear model (1.1),

A4. $\lim_{n\to\infty} \sum_{t=1}^{n} [f(x_t,\theta_0) - f(x_t,\theta_n)]^2 < \infty.$

A5. $\lim_{n\to\infty} \dfrac{\max_{1\le t\le n} [f(x_t,\theta_0) - f(x_t,\theta_n)]^2}{\sum_{t=1}^{n} [f(x_t,\theta_0) - f(x_t,\theta_n)]^2} = 0$

Before we proceed to consider the main theorem of this section, we present the following two lemmas required subsequently.

Let E_0 and Var_0 denote the expected value and variance under P_0 respectively.

Lemma 4.1. Suppose g(x) satisfies condition (1.2). Then under P_0

$$\lim_{n\to\infty} \max_{1\le t\le n} P\left\{ \left| \frac{g(Y_t - f(x_t,\theta_n))}{g(Y_t - f(x_t,\theta_0))} - 1 \right| > \varepsilon \right\} = 0$$

for every $\varepsilon > 0$.

Proof. The proof easily follows from the fact that for every t

$$P_0 \left\{ \left| \frac{g(Y_t - f(x_t,\theta_n))}{g(Y_t - f(x_t,\theta_0))} - 1 \right| > \varepsilon \right\} \le \frac{1}{\varepsilon} E_0 \left| \frac{g(Y_t - f(x_t,\theta_n))}{g(Y_t - f(x_t,\theta_0))} - 1 \right|$$
$$= \frac{1}{\varepsilon} \left| f(x_t,\theta_0) - f(x_t,\theta_n) \right| \int_{-\infty}^{+\infty} \left| \frac{g(y - f(x_t,\theta_0)) - g(y - f(x_t,\theta_n))}{f(x_t,\theta_0) - f(x_t,\theta_n)} \right| dy$$
$$\le \frac{1}{\varepsilon} \left| f(x_t,\theta_0) - f(x_t,\theta_n) \right| \int_{-\infty}^{+\infty} \left| g'(y - f(x_t,\theta_0)) \right| dy \to 0, \text{ as } n \to \infty$$

which is guaranteed by the condition (1.2). ■

Lemma 4.2. Under assumptions A1, A4 and A5, $\log L_n$ has asymptotically $N(-\frac{1}{2}\sigma^2, \sigma^2)$ under P_0,

where

$$\sigma^2 = \lim_{n\to\infty} \sum_{t=1}^{n} [f(x_t,\theta_0) - f(x_t,\theta_n)]^2 \int_0^1 \Psi^2(w)dw.$$

Proof. It suffices to prove, by virtue of LeCam's second lemma and Lemma 4.1, that under P_0, W_n is asymptotically $N(-\frac{1}{4}\sigma^2, \sigma^2)$. Now define

$$V_n^* = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} [f(x_t,\theta_0) - f(x_t,\theta_n)]\, \mathrm{Sgn}(D_t(\theta_0))\, \Psi[G^*(|D_t(\theta_0)|)]. \tag{4.4}$$

First, we shall prove that $\sqrt{n}\ V_n^*$ has asymptotically $N(0,\sigma^2)$. Next, we will show that $\sqrt{n}\,V_n^*$ and $W_n - E_0W_n$ has the same limiting distribution. Finally we will show that

$$\lim_{n\to\infty} E_0W_n = -\frac{\sigma^2}{4}. \tag{4.5}$$

Denoting $U_t = G^*(|D_t(\theta_0)|)$, $1 \le t \le n$, $\sqrt{n}\,V_n^*$ can be rewritten as $\sqrt{n}\,V_n^* = \sum_{t=1}^{n} Z_t$,

where

$$Z_t = [f(x_t,\theta_0) - f(x_t,\theta_n)]\ \mathrm{Sgn}(D_t(\theta_0))\ \Psi(U_t). \tag{4.6}$$

From (4.6) and the fact that U_t are (i.i.d.) uniform (0,1) random variables, we obtain $E_n Z_t = 0$ and

$$\sigma_t^2 = \mathrm{Var}_0\ Z_t = [f(x_t,\theta_0) - f(x_t,\theta_n)]^2 \int_0^1 \Psi^2(w)dw.$$

Noting that Z_t are independent, for the normality of $\sqrt{n}\,V_n^*$ it is straightforward to verify the Lindeberg condition by using A5.

For the second part of the proof, note that

$$\mathrm{Sgn}(D_t(\theta_0))\ \Psi[G^*(|D_t(\theta_0)|)] = \mathrm{Sgn}(D_t(\theta_0))\frac{g^{*\prime}(|D_t(\theta_0)|)}{g^*(|D_t(\theta_0)|)} = \frac{g'(D_t(\theta_0))}{g(D_t(\theta_0))} \tag{4.7}$$

where $g^*(x) = G^{*\prime}(x)$. Let $h(y) = [g(y)]^{1/2}$ so that $h'/h = g'/2g$. Using (4.3), (4.4) and (4.7), it follows that

$$\begin{aligned}
&\mathrm{Var}_0(W_n - E_0W_n - \sqrt{n}\,V_n^*) = \mathrm{Var}_0(W_n - \sqrt{n}\,V_n^*)\\
&= 4\sum_{t=1}^{n} \mathrm{Var}_0\left\{\frac{h(D_t(\theta_n)) - h(D_t(\theta_0))}{h(D_t(\theta_0))} - \frac{1}{2}[f(x_t,\theta_0) - f(x_t,\theta_n)]\frac{g'(D_t(\theta_0))}{g(D_t(\theta_0))}\right\}\\
&< 4\sum_{t=1}^{n} E_0\left\{\frac{h(D_t(\theta_n)) - h(D_t(\theta_0))}{h(D_t(\theta_0))} - [f(x_t,\theta_0) - f(x_t,\theta_n)]\frac{h'(D_t(\theta_0))}{h(D_t(\theta_t))}\right\}^2\\
&= 4\sum_{t=1}^{n}[f(x_t,\theta_0) - f(x_t,\theta_n)]^2\\
&\quad \times \int_{-\infty}^{\infty}\left[\frac{h(y-f(x_t,\theta_n)) - h(y-f(x_t,\theta_n))}{f(x_t,\theta_n) - f(x_t,\theta_0)} - h'(y - f(x_t,\theta_0))\right]^2 dy.
\end{aligned} \tag{4.8}$$

Since the integral in (4.8) converges uniformly in t to zero as $n \to \infty$, it follows from A4 that $\mathrm{Var}_0(W_n - E_0W_n - \sqrt{n}\ V_n^*)$ converges to zero as $n \to \infty$.

Finally, to prove (4.5) note that

$$2\,E_0\left[\frac{h(D_t(\theta_n))}{h(D_t(\theta_0))}-1\right] = -\,E_0\left[\frac{h(D_t(\theta_n))}{h(D_t(\theta_0))}-1\right]^2 \tag{4.9}$$

Using (4.9), it follows that E_0W_n become

$$-\sum_{t=1}^{n}\,[f(x_t,\theta_0)-f(x_t,\theta_n)]^2\int_{-\infty}^{\infty}\left[\frac{h(y-f(x_t,\theta_n))-h(y-f(x_t,\theta_0))}{f(x_t,\theta_0)-f(x_t,\theta_n)}\right]^2 dy. \tag{4.10}$$

The proof is complete from the fact that the integral in (4.10) converges to $\frac{1}{4}\int_{-\infty}^{\infty}\Psi^2(w)dw$ by using (1.2). The lemma follows. ■

We have the following theorem from Lemmas 4.1 and 4.2.

We shall write $\frac{\partial f(x_t,\theta_0)}{\partial\theta_u}$ for $\frac{\partial f(x_t,\theta)}{\partial\theta_u}$ when it is evaluated at $\theta=\theta_0$.

Theorem 4.3. Let the model (1.1) and condition (1.2) hold. Suppose that assumptions A1 through A5 are astisfied. Then, under P_n, $\sqrt{n}\,\lambda' S(y,\theta_0)$ has asymptotically a normal distribution with mean $\lambda'\mu$ and variance $\lambda' V(\theta_0)\,\lambda$ where $\mu=(\mu_1,\cdots,\mu_p)'$,

$$\mu_u=\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\frac{\partial f(x_t,\theta_0)}{\partial\theta_u}\;[f(x_t,\theta_0)-f(x_t,\theta_n)]\int_0^1\phi(w)\,\Psi(w)dw \tag{4.11}$$

for $u=1,\cdots,p$.

Proof. By virte of LeCam's first lemma and Lemma 4.2, $\{P_n\}$ is contiguous to $\{P_0\}$. Therefore it suffices to prove, by LeCam's third lemma, that under P_0, $(\log L_n, \sqrt{n}\,\lambda' S(y,\theta_0))$ has asymptotically a bivariate normal distribution with mean vector $\tilde{\mu}=(-\frac{1}{2}\sigma^2, 0)$ and variance-covariance matrix $\Sigma=[\sigma_{ij}]$ where $\sigma_{11}=\sigma^2$, $\sigma_{12}=\sigma_{21}=\lambda'\mu$ and $\sigma_{22}=\lambda' V(\theta_0)\,\lambda$.

Recall that under P_0, Log L_n and $\sqrt{n}\,V_n^*-\frac{1}{2}u_n^2$

where

$$u_n^2=\sum_{t=1}^{n}\,[f(x_t,\theta_0)-f(x_t,\theta_n)]^2\int_0^1\phi^2(w)dw,$$

$\sqrt{n}\lambda' S(y,\theta_0)$ and $\sqrt{n}\lambda' S^{(0)}(y,\theta_0)$ have the same limiting normal distribution respectively. Therefore it suffices to show the asymptotic normality of $(\sqrt{n}V_n^* - \frac{1}{2}u_n^2, \sqrt{n}\lambda' S^{(0)}(y,\theta_0))$. First consider $\sqrt{n}(V_n^*, \lambda' S^{(0)}(y,\theta_0))$. Now, from (2.4) and (4.4) we have, under P_0

$$\sqrt{n}(V_n^*, \lambda' S^{(0)}(y,\theta_0))$$

$$= \left[\sum_{t=1}^{n}[f(x_t,\theta_0) - f(x_t,\theta_n)]\,\Psi(U_t), \frac{1}{n}\sum_{t=1}^{n}\sum_{u=1}^{p}\lambda_u\frac{\partial f(x_t,\theta_0)}{\partial\theta_u}\,\phi(U_t)\right]\mathrm{Sgn}(D_t(\theta_0)).$$

Therefore,

$$s_{11}(n) = \mathrm{Var}_0[\sqrt{n}V_n^*] = \sum_{t=1}^{n}[f(x_t,\theta_0) - f(x_t,\theta_n)]^2\int_0^1\Psi^2(w)dw,$$
$$s_{22}(n) = \mathrm{Var}_0[\sqrt{n}\lambda' S^{(0)}(y,\theta_0)] = \frac{1}{n}\sum_{t=1}^{n}\left[\sum_{u=1}^{p}\lambda_u\frac{\partial f(x_t,\theta_0)}{\partial\theta_u}\right]^2\int_0^1\phi^2(w)dw,$$
$$s_{12}(n) = s_{21}(n) = \mathrm{Cov}_0\sqrt{n}[V_n^*, \lambda' S^{(0)}(y,\theta_0)]$$
$$= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left[\sum_{u=1}^{p}\lambda_u\frac{\partial f(x_t,\theta_0)}{\partial\theta_u}\right][f(x_t,\theta_0) - f(x_t,\theta_n)]\int_0^1\phi(w)\Psi(w)dw$$

which, by A2 and A4, tend to finite. From the above results, as $n\to\infty$ we have $\tilde{\mu} = (-\frac{1}{2}\sigma^2, 0)$, $\sigma_{11} = \lim_{n\to\infty} s_{11}(n) = \sigma^2$, $\sigma_{12} = \sigma_{21} = \lim_{n\to\infty} s_{21}(n) = \lambda'\mu$ and $\sigma_{22} = \lim_{n\to\infty} s_{22}(n) = \lambda' V(\theta_0)\lambda$.

Now, we have to show that under P_0, $\sqrt{n}(V_n^*, \lambda' S^{(0)}(y,\theta_0))$ has a limiting bivariate normal distribution. By the same argument as in the proof of Theorem 2.1, it suffices to establish the asymptotic normality of an arbitrary linear combination of the components. For any reals β_1 and β_2, let

$$J_n = \beta_1(\sqrt{n}V_n^*) + \beta_2(\sqrt{n}\lambda' S^{(0)}(y,\theta_0)).$$

Then $J_n = \sum_{t=1}^{n} T_t$ where

$$T_t = \mathrm{Sgn}(D_t(\theta_0))\left[\frac{\beta_1}{n}\sum_{u=1}^{p}\lambda_u\frac{\partial f(x_t,\theta_0)}{\partial\theta_u}\,\phi(U_t) + \beta_2[f(x_t,\theta_0) - f(x_t,\theta_n)]\,\Psi(U_t)\right],$$

which are independent, so that $E_0T_t = 0$ and $\sigma_t^2 = \mathrm{Var}_0T_t$ is equal

$$\sigma_t^2 = \frac{\beta_1^2}{n}\left[\sum_{u=1}^{p}\lambda_u \frac{\partial f(x_t,\theta_0)}{\partial\theta_u}\right]^2 \int_0^1 \phi^2(w)dw + \beta_2^2\,[f(x_t,\theta_0)-f(x_t,\theta_n)]^2 \int_0^1 \Psi^2(w)dw$$
$$+2\frac{\beta_1\beta_2}{\sqrt{n}}\left[\sum_{u=1}^{p}\lambda_u \frac{\partial f(x_t,\theta_0)}{\partial\theta_u}\right][f(x_t,\theta_0)-f(x_t,\theta_n)]\int_0^1 \phi(w)\,\Psi(w)dw.$$

Finally, also we can easily check the Lindeberg condition. Hence the theorem follows. ■

Theorem 4.4. Under the assumptions of Theorem 4.3, $\sqrt{n}\ S(y,\theta_0)$ has asymptotically a p-variate normal distribution with mean μ and variance-covariance matrix $V(\theta_0)$ under $\{H_n\}$.

Proof. The proof of this theorem follows directly from Theorem 4.3. ■

Then, we have the following main theorem.

Theorem 4.5. Under the assumptions of Theorem 4.3, and $\{H_n\}$ in (4.1), the statistic $Q_n(y,\theta_0)$ has asymptotically a noncentral chi-square distribution with p degrees of freedom and noncentrality parameter equal to

$$\triangle_Q = \frac{1}{2}\ \mu' V^{-1}(\theta_0)\ \mu \tag{4.12}$$

where μ and $V(\theta_0)$ are defined in Theorem 4.3.

Proof. The proof follows, by the same argument as in proof of Theorem 3.1, from Theorem 4.4. ■

5. Asymptotic Efficiency of the Proposed Tests

In this section we shall consider the relative asymptotic power efficiency of the proposed tests (Q-tests) based on $Q_n(y,\theta)$ with respect to a classical parametric chi-square test (χ^2-test) derived from the asymptotic normality of the least squares estimator (LSE).

It is known [e.g., Chanda (1976)] that under certain regularity conditions, the sequence of the LSE's $\hat{\theta}_n$ has asymptotically a normal distribution in the sense that

$$L\{\sqrt{n}\,(\hat{\theta}_n-\theta)\} \to N_p(0,\ \sigma^2 C^{-1}(\theta))$$

where σ^2 is the common variance of errors in the model (1.1), and $C(\theta)$ is a p x p matrix with (u,v)th element as

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \frac{\partial f(x_t,\theta)}{\partial \theta_u} \frac{\partial f(x_t,\theta)}{\partial \theta_v}.$$

For simplicity, we assume that σ^2 is known. The χ^2-test statistic is, therefore,

$$X^2(y,\theta)= \frac{n}{\sigma^2} (\hat{\theta}_n-\theta)' C_n(\theta)(\hat{\theta}_n-\theta)$$

where $C_n(\theta)$ is defined as $C(\theta)$ before taking limit in the elements and the large sample size α-test for H_0: $\theta=\theta_0$ in (3.2) has the form: Reject or accept H_0, according as $X^2(y, \theta_0) \geq$ or $<$ $\chi^2_{1-\alpha}(p)$. It is easily seen that under H_n this statistic has asymptotically a noncentral chi-square distribution with p degrees of freedom and the noncentrality parameter

$$\triangle_{\chi^2} = \frac{1}{2\sigma^2} h'C(\theta_0)h \tag{5.1}$$

where h is defined in (4.1).

It follows from Theorem 4.5 and the above result that the asymptotic Pitman relative efficiency (ARE) of the Q-tests with respect to the χ^2-test is [Hannan, 1956] equal to

$$e(Q,\chi^2) = \frac{\triangle_Q}{\triangle_{\chi^2}}$$

where $\triangle_Q$ and $\triangle_{\chi^2}$ are given by (4.12) and (5.1) respectively. Noting that

$V(\theta_0)= \left[\int_0^1 \phi^2(w)dw \right]^{-1} C(\theta_0)$, the ARE reduces to

$$\frac{\sigma^2 \left[\int_0^1 \phi(w) \Psi(w)dw \right]^2}{\int_0^1 \phi^2(w)dw} \frac{\mu^{*\prime} C^{-1}(\theta_0) \mu^*}{h'C(\theta_0) h} \tag{5.2}$$

where $\mu^*= (\mu_1^*,\cdots,\mu_p^*)'$ and μ_u^*, u= 1,⋯,p, are defined as μ_u's in (4.11), except the integral terms are deleted. Thus the expression of (5.2) depends not only on the distribution function G through the score functions but depends on the θ_0, h, and regression function f through μ^* and C. However, by virtue of the following lemma, the expression for ARE can be simplified.

Lemma 5.1. Under the model (1.1) and $\{H_n\}$,

$$\frac{\mu^{*\prime}C^{-1}(\theta_0)\mu^*}{h'C(\theta_0)h} = 1. \tag{5.3}$$

In particular, for the linear model (1.3), the equality in (5.3) holds for every n before taking limits in μ^* and $C(\theta_0)$.

Proof. Let $u_n(\theta_0)$ be a p x 1 vector with the uth element as

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial f(x_t,\theta_0)}{\partial\theta_u}\,[f(x_t,\theta_0)-f(x_t,\theta_n)],\quad u=1,\cdots,p.$$

Then we have

$$f(x_t,\theta_0)-f(x_t,\theta_n)=f(x_t,\theta_0)-f(x_t,\theta_0+\frac{h}{\sqrt{n}})=\frac{1}{\sqrt{n}}[\,\nabla_t f(\xi)\cdot h]$$

for some ξ on the segment joining θ_0 and $\theta_0+\frac{h}{\sqrt{n}}$, where $\nabla_t f(\theta)=(\frac{\partial f(x_t,\theta)}{\partial\theta_1},\cdots,\frac{\partial f(x_t,\theta)}{\partial\theta_p})$.

Thus,

$$u_n(\theta_0)=\Sigma^*\,h$$

where Σ^* is a p x p matrix with the (u,v)th elements as

$$\frac{1}{n}\sum_{t=1}^{n}\frac{\partial f(x_t,\theta_0)}{\partial\theta_u}\;\frac{\partial f(x_t,\xi)}{\partial\theta_v}.$$

Note $n\to\infty$, $u_n(\theta_0)\to\mu^*$ and $\Sigma^*\to C(\theta_0)$. This implies that $\mu^*=C(\theta_0)\,h$, therefore $\mu^{*\prime}C^{-1}(\theta_0)\mu^*=h'\,C'(\theta_0)\,C^{-1}(\theta_0)\,C(\theta_0)\,h=h'C(\theta_0)\,h$.

For the second part of theorem, it suffices to show that $u_n'(\theta_0)\;C_n'(\theta_0)\;u_n(\theta_0)=h'C_n(\theta_0)h$ for all n, where $C_n(\theta_0)$ is defined above. It follows easily from the fact that, for linear model (1.3), $u_n(\theta_0) = C_n(\theta_0)\, h$ for every n. ■

From (5.2) and Lemma 5.1, the ARE of the Q-tests with respect to the χ^2-test is, under the conditions specified in Theorem 4.3, equal to

$$e(Q, \chi^2) = \sigma^2 \frac{\left[\int_0^1 \phi(w)\Psi(w)dw\right]^2}{\int_0^1 \phi^2(w)dw}, \tag{5.4}$$

and this implies that the ARE does not depend on the nature of the underlying regression function $f(x,\theta)$, but depend on the unknown distribution function $G(x)$ of errors, through the score function ϕ.

Note that the expression (5.4) is the same as the one obtained (for the simple linear model) by Adichie (1967).

Various interesting results allied to the expression of the ARE are given below for the specific cases: First, let $\phi(u) = u$ on (0,1) (Wilcoxon score). Then, in this case

$$e(Q, \chi^2) = 12\sigma^2 \left[\int_{-\infty}^{\infty} g^2(x)\, dx\right]^2.$$

It is known [Hodges and Lehmann (1956)] that $e(Q, \chi^2) \geq 0.864$ for all continuous G. Some particular values are $e(Q, \chi^2) = \frac{3}{\pi} = 0.955$ when g is a normal density, $e(Q, \chi^2) = 1$ for the case of a uniform, and $e(Q, \chi^2) = 81/64$ when $g(x) = x^2 e^{-x} / \Gamma(3)$ for $x \geq 0$. It is also known that $e(Q, \chi^2)$ exceeds one for distributions G with heavy tails (e.g., Cauchy, double-exponential, logistic distribution etc.). Second, let $\phi(u) = \Phi^{-1}(u)$ on (0,1) (Normal score), where Φ is the standard normal distribution function having the density Φ',

$$e(Q, \chi^2) = \sigma^2 \left\{\int_{-\infty}^{\infty} \frac{g^2(x)}{\Phi'(\Phi^{-1}[G(x)])}\, dx\right\}^2.$$

It has been shown [Chernoff and Savage (1958)] that $e(Q, \chi^2) \geq 1$ for all G. Mikulski (1963) has known that $e(Q, \chi^2) = 1$ only if G is normal. Finally let $\phi(u) = 1$ on (0,1) (sign score), then $e(Q, \chi^2) = 4\sigma^2 g^2(0)$.

Thus from the ARE point of view the Q-test using Wilcoxon or normal scores appears to be superior compared with the classical χ^2-test in many situations. In particular, the Q-test using Wilcoxon scores is preferable when the tails of the error distribution are heavy.

References

(1) Adichie, J.N. (1967) Asymptotic efficiency of a class of nonparametric tests for regression parameters. *Annals of Mathematical Statistics*, 38, 884-893.

(2) Chanda, K.C. (1976) Efficiency and robustness of least squares estimators. *Sankyā*, B, 38, 153-163.

(3) Chernoff, H. and Savage, I.R.(1958) Asymptotic normality and efficiency of certain non-parametric test statistics. *Annals of Mathematical Statistics*, 29, 972-994.

(4) Hannan, B.J. (1956) The asymptotic powers of certain tests based on multiple correlations. *Journal of the Royal Statistical Society*, B, 18, 227-233.

(5) Hodges, J.L. and Lehmann, E.L. (1956) The efficiency of some nonparametric competitions of the t-test. *Annals of Mathematical Statistics*, 27, 324-335.

(6) Jennrich, R.I. (1969) Asymptotic properties of non-linear least squares estimators. *Annals of Mathematical Statistics*, 40, No.2, 633-643.

(7) LeCam, L. (1960) Locally asymptotic normal families of distribution. *University of California Publications in Statistics*, 3, No. 2, 37-98.

(8) Mainvaud, E.(1970) The consistency of nonlinear regression. *Annals of Mathematical Statistics*, 41, 956-969.

(9) Mikulski, P.W. (1963) On the efficiency of optimal nonparametric procedures in the two sample case. *Annals of Mathematical Statistics*, 34, 22-32.

(10) Wu, C.F.(1981) Asymptotic theory of nonlinear least squares estimation. *Annals of Statistics*, 9, 501-513.