Statistica Sinica 1(1991), 203-228

ASYMPTOTICALLY EFFICIENT ADAPTIVE L-ESTIMATORS IN LINEAR MODELS

A. H. Welsh

Australian National University

Abstract: An asymptotically efficient adaptive L-estimator of the slope in a linear model is proposed and investigated. The estimator is a one-step L-estimator of the type discussed by Welsh (1987a,b) with an estimate of the optimal "score" function. The optimal "score" function is related to the integral of (and hence should be easier to estimate than) the usual optimal L-estimator weight function. In constructing the estimator, the data is convolved with a vanishingly small Cauchy contaminant and then the conditional expectation given the data is taken. The "score" function can be treated as constant with respect to the conditional expectation. This means that the conditional expectation can be evaluated explicitly so that calculation of the estimator does not involve the numerical evaluation of an integral. A particular kernel based estimator of the optimal "score" function is examined.

Key words and phrases: Adaptive estimation, asymptotic efficiency, kernel estimators, L-estimator, linear model.

1. Introduction

The problem of constructing asymptotically efficient estimators of a finite dimensional Euclidian parameter θ in a semi-parametric model with an infinite dimensional nuisance shape parameter F has attracted considerable recent attention. Stein (1956) investigated the possibility of constructing estimators of θ (called adaptive estimators) which are asymptotically efficient in the sense that asymptotically they perform as well when F is unknown as when F is known. He showed that adaption is possible when F is a symmetric distribution and θ is the center of symmetry. This problem was treated in increasing generality by a number of authors; results for adaptive location R-, M- and L-estimators are contained in articles by Beran (1974), Stone (1975) and Sacks (1975) respectively. Bickel (1982) extended the work of Stein (1956) and showed in particular that adaption is possible when θ is the slope of a linear model with an intercept even if the underlying distribution F is asymmetric. Dionne (1981) (R-estimators), Bickel (1982) (M-estimators), Koul and Susarla (1983) (M-estimators), Newey (1988) (method of moment type estimators) and Portnoy and Koenker (1989)

(L-estimators based on regression quantiles) have recently proposed adaptive estimators for this problem. See also the general approaches of Begun et al. (1983) and Schick (1987). The results of extensive simulations are reported in Hsieh and Manski (1987); the simulation results in Portnoy and Koenker (1989) indicate that in finite samples, their adaptive estimator has good small sample properties.

In this paper, we propose and investigate a simple, fully adaptive, one-step L-estimator of the slope parameter in a linear model. The way in which the "score" function enters the estimator means that, in contrast to the estimators of Dionne (1981) and Bickel (1982), the estimator does not require sample splitting and is simple to compute. In particular, in contrast to the estimator of Koul and Susarla (1983), computation of the estimator does not require the evaluation of an integral. Both these properties are important in applications. The estimator is derived in the same way as the L-estimators of Welsh (1987a,b), using ideas from Sacks (1975), Stone (1975) and Portnoy and Koenker (1989). In contrast to the estimator of Portnoy and Koenker (1989), however, the present estimator is non-randomised and depends on the "score" function rather than its derivative. A simple easy-to-use kernel based method for estimating the "score" function is proposed. Finally, we show that the estimator is adaptive for any F under simple, general conditions on the design.

Suppose that we observe $Y_1, \ldots, Y_n$ satisfying

$$Y_j = \tau + x_j'\theta_0 + e_j, \qquad 1 \leq j \leq n,$$

where $\{x_j\}$ is a sequence of known p-vectors ($p \geq 1$), $\tau \in \mathbf{R}$ is an unknown intercept, $\theta_0 \in \mathbf{R}^p$ is the unknown slope parameter of interest and $\{e_j\}$ is a sequence of independent and identically distributed random variables with a common distribution function F. We will assume that the error distribution has a density function f with respect to Lebesgue measure. Since the components of $x_1, \ldots, x_n$ can be centered, we assume without loss of generality that $\sum_{j=1}^n x_j = 0$. The consequent dependence of $\{x_j\}$ on n is suppressed for notational simplicity.

The proposed estimator and a statement of the main result will now be presented. Thereafter, the derivation of the estimator and a sketch of the main steps in the proof of its asymtotic efficiency will be presented.

1.1. The estimator

Let θ_n be an initial estimator of the slope θ_0 and construct the (uncentered) residuals $r_j = Y_j - x_j'\theta_n$, $1 \leq j \leq n$. Let

$$G(z) = \frac{1}{\pi}\arctan(z) + \frac{1}{2}, \qquad -\infty \leq z \leq \infty,$$

denote the standard Cauchy distribution function. Then with

$$\gamma \equiv \gamma_n = (\log n)^{-1/2},$$

define the kernel distribution function estimator

$$\widehat{G}_n(y) = \frac{1}{n}\sum_{j=1}^{n} G\{(y - r_j)/\gamma\}, \qquad -\infty \le y \le \infty,$$

and set

$$\widehat{\alpha} \equiv \widehat{\alpha}_n = \widehat{G}_n(-(\log n)^{1/4}) + n^{-1/4} \text{ and } \widehat{\beta} \equiv \widehat{\beta}_n = \widehat{G}_n((\log n)^{1/4}) - n^{-1/4}.$$

Then the adaptive one-step L-estimator λ_n of θ_0 is

$$\lambda_n = \theta_n - \widehat{I}_n^{-1}(X'X)^{-1} \sum_{i=[n\widehat{\alpha}]+1}^{[n\widehat{\beta}]} \widehat{\eta}_n(i/n) \sum_{j=1}^{n} x_j \left\{ G\{(\widehat{Q}_n(i/n) - r_j)/\gamma\} - G\{(\widehat{Q}_n((i-1)/n) - r_j)/\gamma\} \right\}, \quad (1.1)$$

where $X'X = \sum_{j=1}^n x_j x_j'$, $\widehat{I}_n = n^{-1}\sum_{i=[n\widehat{\alpha}]+1}^{[n\widehat{\beta}]} \widehat{\eta}_n(i/n)^2$, $\widehat{Q}_n(u) = \widehat{G}_n^{-1}(u)$ and $\widehat{\eta}_n(u)$ is an estimator of the "score" function $\eta(u) = f'(F^{-1}(u))/f(F^{-1}(u))$, $0 \le u \le 1$.

There is inevitable arbitrariness in the choice of the estimator $\widehat{\eta}_n$. We will consider explicitly a simple kernel-based estimator of η. Let $k_1(x)$ and $k_2(x)$ be kernel functions with support on $[-1,1]$ which have bounded derivatives on $[-1,1]$ and which satisfy

$$\frac{1}{\ell!}\int_{-1}^{1} x^\ell k_i(x)dx = \begin{cases} 0 & \ell < i, \\ (-1)^i & \ell = i, \quad 0 \le \ell \le i, \quad i = 1,2. \end{cases} \quad (1.2)$$

For example, two simple polynomial kernels which satisfy these conditions are

$$k_1(x) = -\frac{3}{2}xI(|x| \le 1) \text{ and } k_2(x) = \left(-\frac{45}{2} - \frac{3}{5}x + \frac{45}{4}x^2 + x^3\right)I(|x| \le 1). \quad (1.3)$$

(Note that k_1 is the derivative of the Epanechnikov (1969) density kernel.) For $\delta_i \downarrow 0$ sufficiently small, let

$$\widehat{Q}_n^{(i)}(u) = \delta_i^{-1-i}\int_0^1 \widehat{Q}_n(w)k_i\{\delta_i^{-1}(u-w)\}\,dw \sim \delta_i^{-i}\int_{-1}^{1} \widehat{Q}_n(u - \delta_i x)k_i(x)dx,$$

$0 < u < 1$, $i = 1,2$. Now construct the estimator

$$\widehat{\eta}_n(u) = -\widehat{Q}_n^{(2)}(u)/\left\{\widehat{Q}_n^{(1)}(u)\right\}^2, \quad 0 < u < 1. \tag{1.4}$$

The results of Section 3 indicate that we should choose $\delta_i \sim n^{-\varepsilon_i}$ for $0 \le \varepsilon_i \le (4+2i)^{-1}$, $i = 1,2$. Since η is a nuisance parameter rather than a parameter of interest, classical theory provides little guidance in the precise choice of δ_1 and δ_2 and further research on this problem is required.

The following result will be proved in Sections 2–4.

Theorem. *Suppose that* $\{x_j\}$ *satisfies*

i) $$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} x_j x_j' = \Gamma, \quad \Gamma \textit{ non-singular,}$$

and

ii) $$\max_{1\le j\le n} |x_j| \le n^{1/4}C \quad \textit{for some } C < \infty.$$

Then if $n^{1/2}(\theta_n - \theta_0)$ *is bounded in probability as* $n \to \infty$, *it follows that*

$$n^{1/2}(\lambda_n - \theta_0) \xrightarrow{\mathcal{D}} N(0, I^{-1}\Gamma^{-1}),$$

where λ_n *is defined by* (1.1), $\widehat{\eta}$ *is defined by* (1.4) *with* k_1 *and* k_2 *satisfying* (1.2) *and* $I = \int_0^1 \eta(u)^2 du$. *Moreover*

$$\widehat{I}_n(\gamma)^{-1} n(X'X)^{-1} \xrightarrow{\mathrm{P}} I^{-1}\Gamma^{-1}.$$

Note that there are no conditions on F; the Fisher Information I can be finite or infinite. The conditions on $\{x_j\}$ are nearly minimal; in the second condition, we could replace $n^{1/4}$ by n^v, $0 < v < 1/2$, but the resulting increase in generality is unimportant. Nonetheless, these conditions are slightly weaker than those of Portnoy and Koenker (1989). If the density f is continuous and positive at $F^{-1}(1/2)$ then the least absolute deviations estimator can be used as the initial estimator θ_n (see Bassett and Koenker (1978)). Finally, it is clear from the arguments in Section 4 that the Theorem will hold for any estimator of η which satisfies the conclusions of Lemmas 3.5 and 3.6 and not just for the kernel estimator (1.4).

1.2. The derivation of the estimator

To motivate the construction of λ_n suppose first that both η and F are known. Then the efficient $\dot{L}$-estimator has weight function $-\eta'(u)f(F^{-1}(u))$. Following Welsh (1987a,b), we seek an estimator λ_n satisfying

$$\lambda_n = \theta_0+I^{-1}(X'X)^{-1}\sum_{j=1}^{n} x_j\int_{-\infty}^{\infty}\{I(e_j \le y) - F(y)\}\eta'(F(y))f(y)dy + o_p(n^{-1/2})$$

$$= \theta_0+I^{-1}(X'X)^{-1}\sum_{j=1}^{n} x_j\int_{0}^{1}\{I\{e_j \le F^{-1}(u)\}-u\}\eta'(u)du + o_p(n^{-1/2}). \quad (1.5)$$

If the integral is approximated by a sum, (1.5) becomes

$$\lambda_n \approx \theta_0+I^{-1}(X'X)^{-1}\sum_{j=1}^{n} x_j n^{-1}\sum_{i=1}^{n}\{I\{e_j \le F^{-1}(i/n)\}-i/n\}\eta'(i/n)+o_p(n^{-1/2})$$

$$= \theta_0 + I^{-1}(X'X)^{-1}n^{-1}\sum_{i=1}^{n}\eta'(i/n)\sum_{j=1}^{n} x_j I\{e_j \le F^{-1}(i/n)\} + o_p(n^{-1/2})$$

as $\sum_{j=1}^{n} x_j = 0$. We could proceed to base the estimator on this expression but, for adaption, η' would have to be estimated as in Portnoy and Koenker (1989). However, we can avoid having to estimate η' by proceeding as in Sacks (1975): replace $\eta'(i/n)$ by the normalised difference $n\{\eta((i+1)/n) - \eta(i/n)\}$, sum by parts and ignore the end terms to obtain

$$\lambda_n \approx \theta_0+I^{-1}(X'X)^{-1}\sum_{i=1}^{n}\eta(i/n)\sum_{j=1}^{n} x_j I\{F^{-1}((i-1)/n) < e_j \le F^{-1}(i/n)\} + o_p(n^{-1/2}). \quad (1.6)$$

Now to construct an estimator of θ_0 satisfying (1.6), proceed as in Welsh (1987a,b) to modify the right hand side of (1.6) by replacing θ_0 by θ_n, e_j by r_j, $1 \le j \le n$, F by the empirical distribution function of the residuals and η by an appropriate estimator. To control the extreme residuals, we then trim a decreasingly small portion of the residuals. The analysis of the resulting estimator would require f to be bounded, to have three bounded derivatives and to vanish only at $\pm\infty$. Following Stone (1975), Koul and Susarla (1983) and Portnoy and Koenker (1989), these smoothness assumptions are avoided by first convolving the residuals with a vanishingly small smooth contaminant. Let $\{Z_j\}$ be a sequence of independent and identically distributed standard Cauchy random variables such that $\{Z_j\}$ is independent of $\{e_j\}$. With γ defined as above, let $r_j^* = r_j + \gamma Z_j$, $1 \le j \le n$, denote the contaminated residuals and then define

$$G_n(y) = \frac{1}{n}\sum_{j=1}^{n} I(r_j^* \le y), \qquad -\infty \le y \le \infty,$$

to be the empirical distribution function of the contaminated residuals. Let

$$\alpha^* \equiv \alpha_n^* = G_n(-(\log n)^{1/4}) + n^{-1/4} \text{ and } \beta^* \equiv \beta_n^* = G_n((\log n)^{1/4}) - n^{-1/4}.$$

Then, the randomised estimator is

$$\lambda_n^* = \theta_n - I_n^{*-1}(X'X)^{-1} \sum_{i=[n\alpha^*]+1}^{[n\beta^*]} \eta_n^*(i/n) \sum_{j=1}^{n} x_j I\left\{G_n^{-1}((i-1)/n) < r_j^* \le G_n^{-1}(i/n)\right\}.$$

On defining the ranks $\{D(j)\}$ of the contaminated residuals by

$$r_{nj}^* = r_{D(j)}^*, \qquad 1 \le j \le n,$$

where $r_{n1}^* \le r_{n2}^* \le \cdots \le r_{nn}^*$, we can rewrite the randomised adaptive one-step L-estimator λ_n^* of θ_0 in the computationally convenient form

$$\lambda_n^* = \theta_n - I_n^{*-1}(X'X)^{-1} \sum_{i=[n\alpha^*]+1}^{[n\beta^*]} x_{D(i)} \eta_n^*(i/n). \tag{1.7}$$

It is fairly straightforward to modify the arguments of Sections 2–4 to show that the conclusion of the Theorem also applies to the randomised estimator λ_n^* so that it is also an adaptive estimator of θ_0. To obtain a nonrandomised estimator λ_n, replace G_n in λ_n^* by $\widehat{G}_n$, its conditional expectation given $r_1, \ldots, r_n$, and then take the conditional expectation of λ_n^* given $r_1, \ldots, r_n$. This results in the estimator (1.1).

The contaminating distribution should be chosen to have a bounded density with three bounded derivatives. The thickness of the tails of this distribution will affect the rate at which the contaminant vanishes and the choice of the trimming parameters. We have chosen the Cauchy distribution for definiteness and simplicity but other distributions such as those used in Portnoy and Koenker (1989) could equally be used. After specifying a contaminating distribution, there is still some flexibility in the choice of $\widehat{\alpha}$, $\widehat{\beta}$ and γ. Other possible choices of these sequences are described in Section 2.

In constructing an estimator of η, we can regard either $\Psi(y) = \eta(F(y))$, $-\infty < y < \infty$, or, as is more natural for L-estimation, $\eta(u)$, $0 < u < 1$, as the basic nuisance parameter. If Ψ is regarded as the basic nuisance parameter, we can construct an estimator of $\Psi(y)$ and evaluate it at $y = Q_n(u)$ to obtain the randomised slope estimator

$$\lambda_n = \theta_n - I_n^{*-1}(X'X)^{-1} \sum_{i=[n\alpha^*]+1}^{[n\beta^*]} x_{D(i)} \Psi_n^*(r_{ni}^*), \quad \delta > 0, \tag{1.8}$$

or the nonrandomised slope estimator

$$\lambda_n = \theta_n - \widehat{I}_n^{-1}(X'X)^{-1} \sum_{i=[n\widehat{\alpha}]+1}^{[n\widehat{\beta}]} \widehat{\Psi}_n(\widehat{Q}_n(i/n)) \sum_{j=1}^{n} x_j \Big\{ G\{\widehat{Q}_n(i/n) - r_j)/\gamma\} - G\{(\widehat{Q}_n((i-1)/n - r_j)/\gamma\} \Big\}. \tag{1.9}$$

This approach to estimating the optimal weight function is essentially the approach adopted by Portnoy and Koenker (1989), except that they estimate $\Psi'(y) = \eta'(F(y))f(y)$ rather than Ψ. Note that in (1.8) and (1.9) the trimming depends on the ordering of the contaminated residuals rather than on their magnitude, so that (1.8) and (1.9) are not adaptive M-estimators. To convert (1.8) into an adaptive M-estimator, we would need to center the residuals $\{r_j\}$, to introduce a scale estimator $\widehat{\sigma}$ and to arrange to trim (r_j^*, x_j) when $|r_j^*|$ is large relative to $\widehat{\sigma}$, $1 \le j \le n$. To obtain a nonrandomised adaptive M-estimator, then take the conditional expectation given $r_1, \ldots, r_n$. The resulting estimators are similar to the estimators in Koul and Susarla (1983). The distinction between adaptive L- and M-estimators is quite subtle, particularly since integrating (1.5) by parts as opposed to summing by parts leads naturally to the construction of adaptive M-estimators.

An attractive advantage of the L-estimation approach over the M-estimation approach is that the score function can be treated as constant with respect to the conditional expectation. This means that the conditional expectation can be evaluated explicitly so that the non-randomised estimator does not involve an integral. The approach followed here is to regard η as the basic nuisance parameter. Since $\widehat{Q}_n(u) = \widehat{G}_n^{-1}(u)$ is an estimator of $Q(u) = F^{-1}(u)$, $0 \le u \le 1$, we can use differences (as in Sacks (1975) or Welsh (1987b)) or kernels (as in Welsh (1987c)) to estimate the derivatives of Q. As $Q^{(1)}(u) = 1/f(Q(u))$ and $Q^{(2)}(u) = -f'(Q(u))/f(Q(u))^3$, write

$$\eta(u) = -Q^{(2)}(u)/\{Q^{(1)}(u)\}^2, \qquad 0 \le u \le 1.$$

Following Welsh (1987c), construct kernel estimators of $Q^{(1)}$ and $Q^{(2)}$ which are then combined to obtain (1.4).

Provided the initial estimator θ_n is regression and scale equivariant and transformation invariant, so is the adaptive estimator (1.1).

1.3. The main steps in the proof of the Theorem

The proof of the Theorem is rather long and technical but the main steps are easily described. In nearly all the lemmas we prove, we obtain bounds which involve the product of stochastic terms, which are essentially differences between estimators and estimands, and a function of non-stochastic terms like F_n, Q_n and η_n etc. We ensure that these bounds tend to zero by establishing that the stochastic terms decrease at an algebraic rate and ensuring that the non-stochastic terms increase at most at a logarithmic rate. Indeed, α, β and γ are chosen to ensure that this requirement is satisfied. That the appropriate non-stochastic terms increase at most at logarithmic rates is established in Section 2. That the stochastic terms decrease at algebraic rates is established in Section 3. In particular, the rate of decrease to zero of $\widehat{Q}_n(u) - Q_n(u)$ is established in Lemma 3.2, using a modulus of continuity result (Lemma 3.1) which is proved using Bernstein's inequality. Lemma 3.2 is then used to obtain the rate of decrease to zero of $\widehat{\eta}_n(u) - \eta_n(u)$ in Lemmas 3.5 and 3.6. (Lemmas 3.3 and 3.4 establish intermediate steps in the argument by treating the first two derivatives of $\widehat{Q}_n(u) - Q_n(u)$.)

The core of the proof is given in Section 4. From (1.2), on a set whose probability can be made arbitrarily close to one, we have the decomposition

$$\begin{aligned} n^{-1/2}X'X(\lambda_n - \theta_0) &= n^{-1/2}X'X(\theta_n - \theta_0) - n^{-1/2}\widehat{I}_n^{-1} \sum_{i=[n\widehat{\alpha}]+1}^{[n\widehat{\beta}]} \widehat{\eta}_n(i/n) \\ &\quad \times \sum_{j=1}^{n} x_j \left\{G\{(\widehat{Q}_n(i/n) - r_j)/\gamma\} - G\{(\widehat{Q}_n((i-1)/n) - r_j)/\gamma\}\right\} \\ &= \Gamma T + R_1 + R_2 + R_3 + R_4, \end{aligned} \tag{1.10}$$

where

$$T = -n^{-1/2}\widehat{I}_n^{-1} \sum_{i=[n\alpha]+1}^{[n\beta]} \eta_n\Big(\frac{i}{n}\Big) \sum_{j=1}^{n} \Gamma^{-1} x_j \Big\{G\Big\{\Big(Q_n\Big(\frac{i}{n}\Big) - e_j\Big)/\gamma\Big\} - G\Big\{\Big(Q_n\Big(\frac{i-1}{n}\Big) - e_j\Big)/\gamma\Big\}\Big\},$$

$$\begin{aligned} R_1 &= n^{-1/2}\widehat{I}_n^{-1} \sum_{i=[n\widehat{\beta}]+1}^{[n\beta]} \eta_n\Big(\frac{i}{n}\Big) \sum_{j=1}^{n} x_j \Big\{G\Big\{\Big(Q_n\Big(\frac{i}{n}\Big) - e_j\Big)/\gamma\Big\} - G\Big\{\Big(Q_n\Big(\frac{i-1}{n}\Big) - e_j\Big)/\gamma\Big\}\Big\} \\ &\quad - n^{-1/2}\widehat{I}_n^{-1} \sum_{i=[n\alpha]+1}^{[n\widehat{\alpha}]} \eta_n\Big(\frac{i}{n}\Big) \sum_{j=1}^{n} x_j \Big\{G\Big\{\Big(Q_n\Big(\frac{i}{n}\Big) - e_j\Big)/\gamma\Big\} - G\Big\{\Big(Q_n\Big(\frac{i-1}{n}\Big) - e_j\Big)/\gamma\Big\}\Big\}, \end{aligned}$$

$$R_2 = -n^{-1/2}\widehat{I}_n^{-1} \sum_{i=[n\widehat{\alpha}]+1}^{[n\widehat{\beta}]} \Big\{\widehat{\eta}_n\Big(\frac{i}{n}\Big) - \eta_n\Big(\frac{i}{n}\Big)\Big\}$$

$$\times \sum_{j=1}^{n} x_j \Big\{G\Big\{\Big(Q_n\Big(\frac{i}{n}\Big) - e_j\Big)/\gamma\Big\} - G\Big\{\Big(Q_n\Big(\frac{i-1}{n}\Big) - e_j\Big)/\gamma\Big\}\Big\},$$

$$R_3 = n^{-1/2}\sum_{j=1}^{n} x_j \Big\{x_j'(\theta_n - \theta_0) - \widehat{I}_n^{-1} \sum_{i=[n\widehat{\alpha}]+1}^{[n\widehat{\beta}]} \widehat{\eta}_n\Big(\frac{i}{n}\Big)\Big\{F_n\Big(\widehat{Q}_n\Big(\frac{i}{n}\Big) + x_j'(\theta_n - \theta_0)\Big)$$

$$-F_n\Big(\widehat{Q}_n\Big(\frac{i-1}{n}\Big) + x_j'(\theta_n - \theta_0)\Big)\Big\}\Big\}$$

and

$$R_4 = -n^{-1/2}\widehat{I}_n^{-1} \sum_{i=[n\widehat{\alpha}]+1}^{[n\widehat{\beta}]} \widehat{\eta}_n\Big(\frac{i}{n}\Big)\sum_{j=1}^{n} x_j \Big\{G\Big\{\Big(\widehat{Q}_n\Big(\frac{i}{n}\Big) - r_j\Big)/\gamma\Big\} - G\Big\{\Big(\widehat{Q}_n\Big(\frac{i-1}{n}\Big) - r_j\Big)/\gamma\Big\}$$
$$- F_n\Big(\widehat{Q}_n\Big(\frac{i}{n}\Big) + x_j'(\theta_n - \theta_0)\Big) + F_n\Big(\widehat{Q}_n\Big(\frac{i-1}{n}\Big) + x_j'(\theta_n - \theta_0)\Big)$$
$$- G\Big\{\Big(Q_n\Big(\frac{i}{n}\Big) - e_j\Big)/\gamma\Big\} + G\Big\{\Big(Q_n\Big(\frac{i-1}{n}\Big) - e_j\Big)/\gamma\Big\}\Big\}.$$

The Theorem follows from Slutsky's Theorem if we prove a sequence of lemmas which show that with $I_n = \int_{[n\alpha]/n}^{[n\beta]/n} \eta_n(u)^2 du$,

$\widehat{I}_n = I_n + o_p(1)$ and $\widehat{I}_n = I + o_p(1)$,

$R_m \xrightarrow{P} 0$, $m = 1, 2, 3, 4$,

and

$T \xrightarrow{\mathfrak{D}} N(0, I^{-1}\Gamma^{-1})$.

The estimated Fisher information $\widehat{I}_n$ appears in all of R_1–R_4 and T so we treat it first. The required result is Lemma 4.1 which is obtained by straightforward approximation arguments and the application of a result of Stone (1975).

Now consider the remainder terms R_1–R_4. The first remainder term is treated by showing that its variance tends to zero (Lemma 4.2). The remaining terms R_2–R_4 are first summed by parts. The resulting terms in R_2 are straightforward to control once it can be shown that a particular weighted kernel distribution function estimator is bounded in probability. This is shown by a weak convergence argument (Lemma 4.3). The behaviour of R_3 depends on the behaviour of $\widehat{Q}_n(u) - Q_n(u)$ (given by Lemma 3.2) which we isolate by use of Taylor series expansions (Lemma 4.4). The treatment of R_4 avoids making an expansion by applying the modulus of continuity result (Lemma 3.1) in conjunction with Lemma 3.2 (Lemma 4.5).

Finally, the leading term T involves a sum over the whole sample and a trimmed sum. In the proof of Lemma 4.6, the trimmed sum is first approximated by an integral so that T can be written as the row sum of a triangular array of rowwise independent random variables. This sum is then shown to converge as required by Lyapounov's Central Limit Theorem and an argument of Koul and Susarla (1983).

2. Properties of the Smoothed Error Distribution

In order to avoid having to place strong conditions on the underlying error distribution F, the error distribution is convolved with a vanishingly small Cauchy contaminant. In this section, we derive some useful properties of the smoothed error distribution. The results and proofs are similar to those given by Stone (1975) and Portnoy and Koenker (1989). Throughout this section and in the sequel, K with or without subscripts denotes a generic positive constant which is not necessarily the same at each appearance.

For fixed $0 < c < 1$ let $\gamma \equiv \gamma_n = (\log n)^{-c}$. We specified the value $c = 1/4$ in the previous section for definiteness but the Theorem will be proved for the more general formulation. The convoluted errors $\{e_j + \gamma Z_j\}$ are independent and identically distributed with common distribution function

$$F_n(y) = \int_{-\infty}^{\infty} G((y-z)/\gamma)dF(z), \quad -\infty < y < \infty,$$

where $G(y) = \frac{1}{\pi}\arctan(y) + \frac{1}{2}$, $-\infty < y < \infty$. Note that $F_n(y)$ is absolutely continuous and has density

$$f_n(y) = \gamma^{-1}\int_{-\infty}^{\infty} g((y-z)/\gamma)dF(z), \quad -\infty < y < \infty,$$

where $g(y) = G'(y) = [\pi(1+y^2)]^{-1}$, $-\infty < y < \infty$. Moreover, $f_n(y)$ is three times differentiable and, because the derivatives of g are uniformly bounded,

$$\sup_{-\infty \le y \le \infty} |f_n^{(v)}(y)| \le K_v \gamma^{-(v+1)} = K_v (\log n)^{c(v+1)}, \quad v = 0,1,2,3. \tag{2.1}$$

Also note that if $|x_n| \to \infty$,

$$\begin{aligned} |f_n^{(v)}(x_n)| &\le K\left\{\gamma^{-(v+1)}|g^{(v)}(x_n/2\gamma)| + F(-|x_n|/2) + 1 - F(|x_n|/2)\right\} \\ &\longrightarrow 0, \qquad v = 0,1. \end{aligned} \tag{2.2}$$

We shall also need a lower bound for the smoothed density. Choose K_1, K_2

> 0 so that $F(K_2) - F(-K_1) > 0$. Then

$$\begin{aligned} f_n(y) &\geq \gamma^{-1} \int_{-K_1}^{K_2} g((y-z)/\gamma) dF(z) \\ &\geq \int_{-K_1}^{K_2} \frac{1}{\gamma\pi\,(1 + 2(y^2 + K_3^2)/\gamma^2)} dF(z), \quad K_3 = \max\{K_1, K_2\}, \\ &\geq \gamma K I(|y| \leq 1) + y^{-2}\gamma K I(|y| > 1), \quad K > 0, \end{aligned}$$

so that for any $a > 0$ and $M < \infty$

$$\inf_{|y| \leq M(\log n)^a} f_n(y) \geq K(\log n)^{-(2a+c)}. \tag{2.3}$$

(In the sequel, we will require $a + c < 1$.) It is convenient to express (2.3) in a slightly different form. For any $\rho \downarrow 0$ such that $0 < \rho < K(\log n)^{-(a+c)}$, put

$$U_n^*(\rho) = \\ \{u : F_n(-(\log n)^a - K^{-1}\rho(\log n)^{2a+c}) \leq u \leq F_n((\log n)^a + K^{-1}\rho(\log n)^{2a+c})\}$$

and note that with $\widetilde{\rho} \leq \rho$, by (2.3)

$$\begin{aligned} &F_n((\log n)^a + K^{-1}\rho(\log n)^{2a+c}) \\ &\quad = F_n((\log n)^a) + K^{-1}\rho(\log n)^{2a+c} f_n((\log n)^a + K^{-1}\widetilde{\rho}(\log n)^{2a+c}) \\ &\quad \geq F_n((\log n)^a) + \rho \end{aligned}$$

and, similarly,

$$F_n(-(\log n)^a - K^{-1}\rho(\log n)^{2a+c}) \leq F_n(-(\log n)^a) - \rho$$

so that for any $\rho \downarrow 0$ such that $0 < \rho < K(\log n)^{-(a+c)}$,

$$U_n(\rho) = \{u : F_n(-(\log n)^a) - \rho \leq u \leq F_n((\log n)^a) + \rho\} \subseteq U_n^*(\rho). \tag{2.4}$$

With

$$\alpha \equiv \alpha_n = F_n(-(\log n)^a) \quad \text{and} \quad \beta \equiv \beta_n = F_n((\log n)^a),$$

we can write

$$U_n(\rho) = [\alpha - \rho, \beta + \rho].$$

Throughout this section, suppose that $\rho \downarrow 0$ such that $0 < \rho < K(\log n)^{-(a+c)}$. We can then write (2.3) as

$$\inf_{u \in U_n(\rho)} f_n(Q_n(u)) \geq K(\log n)^{-(2a+c)}, \tag{2.5}$$

where $Q_n(u) = F_n^{-1}(u)$, $0 \le u \le 1$.

Bounds will now be obtained for the derivatives of the quantile function $Q_n(u) = F_n^{-1}(u)$, $0 \le u \le 1$, and for the "score" function

$$\eta_n(u) = f_n^{(1)}(Q_n(u))/f_n(Q_n(u)), \qquad 0 \le u \le 1.$$

It follows immediately from (2.1) that

$$\sup_{0 \le u \le 1} \frac{1}{Q_n^{(1)}(u)} \le K(\log n)^c \tag{2.6}$$

and it follows from (2.1) and (2.5) that

$$\sup_{u \in U_n(\rho)} |Q_n^{(v)}(u)| \le K_v(\log n)^{2a+c+4(v-1)(a+c)}, \qquad v = 1,2,3,4. \tag{2.7}$$

By (2.1) and (2.5)

$$\sup_{u \in U_n(\rho)} |\eta_n(u)| \le K(\log n)^{2a+3c} \tag{2.8}$$

and

$$\sup_{u \in U_n(\rho)} |\eta_n^{(v)}(u)| \le K_v(\log n)^{2a+3c+4^v(a+c)}, \qquad v = 1,2. \tag{2.9}$$

Also, since $q - n^{-1} \le n^{-1}[nq] \le q$ for any $0 < q < 1$,

$$\sum_{i=[n\alpha]+1}^{[n\beta]} \left|\eta_n\Big(\frac{i+1}{n}\Big) - \eta_n\Big(\frac{i}{n}\Big)\right| \le \sup_{u \in U_n(1/n)} |\eta_n^{(1)}(u)| \le K(\log n)^{6a+7c} \tag{2.10}$$

by (2.9).

3. Properties of $\hat{\eta}_n$

In this section, we investigate the properties of $\hat{\eta}_n$ defined in (1.4). The main results depend on two preliminary results which will also be used in Section 4.

Lemma 3.1. *Let $\{c_j \equiv c_{jn}\}$ be any sequence of constants such that* $\lim_{n\to\infty} n^{-1}\sum_{j=1}^n c_j^2 < \infty$. *Then if $\{x_j\}$ satisfies* (i) *and* (ii) *of the Theorem, for any $a > 0$, $a + c < 1$, $d \ge 0$,*

$$\sup_{|y| \le D(\log n)^a} \ \sup_{|t| \le n^{-1/2}B_1} \ \sup_{|\tau| \le n^{-1/2}(\log n)^d B_2} \Big| \frac{1}{n}\sum_{j=1}^n c_j\{G\{(y - e_j + x_j't + \tau)/\gamma\}$$
$$- F_n(y + x_j't + \tau) - G\{(y - e_j)/\gamma\} + F_n(y)\}\Big| = O_p(n^{-3/4}(\log n)^{c+d+1/2}),$$

for any D, B_1 and $B_2 < \infty$.

Proof. Put $v_j = (1, x_j')' \in \mathbf{R}^{p+1}$ and let

$$H_n(y,t) = \frac{1}{n}\sum_{j=1}^{n} c_j \left\{ G\Big(\frac{y - e_j + v_j't}{\gamma}\Big) - F_n(y + v_j't) - G\Big(\frac{y - e_j}{\gamma}\Big) + F_n(y) \right\}, \tag{3.1}$$

$y \in \mathbf{R}$, $t \in \mathbf{R}^{p+1}$. Then the result will follow if it can be shown that

$$V_n = \sup_{|y| \le D(\log n)^a} \ \sup_{|t| \le n^{-1/2}(\log n)^d B} |H_n(y,t)| = O_p\big(n^{-3/4}(\log n)^{c+d+1/2}\big),$$

for any $B < \infty$.

Put

$$\xi_\ell \equiv \xi_{\ell n} = -D(\log n)^a + 2D(\log n)^a \ell/[n^{1/2}], \qquad \ell = 0, 1, \ldots, [n^{1/2}],$$

so $|\xi_\ell| \le C(\log n)^a$ and $\xi_{\ell+1} - \xi_\ell = 2D(\log n)^a/[n^{1/2}]$, $0 \le \ell \le [n^{1/2}]$. Then arguing as in the proof of Lemma 2.1 of Welsh (1987c),

$$V_n \le 2 \max_{1 \le \ell \le [n^{1/2}]} \ \sup_{|t| \le n^{-1/2}(\log n)^d B} |H_n(\xi_\ell, t)|.$$

Next cover the ball $\{|t| \le n^{-1/2}(\log n)^d B\}$ with cubes $\{b(t_k)\}$ about $\{t_k \equiv t_{kn}\}$ with sides of length $n^{-5/4}(\log n)^d B$. Note that $N = (n^{3/4} 2B)^{p+1}$ such cubes are required and that if $t \in b(t_k)$, then $|v_j'(t - t_k)| \le \tau_j$ where $\tau_j \equiv \tau_{jn} = n^{-5/4}(\log n)^d (p+1)^{1/2} B |v_j|$. Now,

$$V_n \le 2 \max_{1 \le \ell \le [n^{1/2}]} \max_{1 \le k \le N} \big\{ |H_n(\xi_\ell, t_k)| + \sup_{t \in b(t_k)} |H_n(\xi_\ell, t) - H_n(\xi_\ell, t_k)| \big\}. \tag{3.2}$$

Now g is bounded so

$$\begin{aligned} \sup_{t \in b(t_k)} |H_n(\xi_\ell, t) - H_n(\xi_\ell, t_k)| &\le n^{-1} \sum_{j=1}^{n} |c_j||v_j| n^{-5/4}(\log n)^{c+d} \\ &\le n^{-5/4}(\log n)^{c+d} K, \end{aligned}$$

by the conditions on $\{c_j\}$ and $\{x_j\}$. Similarly,

$$\begin{aligned} &\sum_{j=1}^{n} c_j^2 E\{G\{(\xi_\ell - e + x_j't_k)/\gamma\} - F_n(\xi_\ell + x_j't_k) - G\{(\xi_\ell - e)/\gamma\} + F_n(\xi_\ell)\}^2 \\ &\qquad \le \sum_{j=1}^{n} c_j^2 E\{G\{(\xi_\ell - e + x_j't_k)/\gamma\} - G\{(\xi_\ell - e)/\gamma\}\}^2 \\ &\qquad \le \sum_{j=1}^{n} c_j^2 |x_j't_k|^2 (\log n)^{2c} K_1 \\ &\qquad \le n^{1/2}(\log n)^{2(c+d)} K, \end{aligned}$$

so that by Bernstein's Inequality, for any $M < \infty$,

$$\begin{aligned}
P\Big\{ \max_{1\le \ell\le [n^{1/2}]} \max_{1\le k\le N} |H_n(\xi_\ell, t_k)| \ge n^{-3/4}(\log n)^{c+d+1/2}M \Big\} \\
\le 2n^{1/2}N \exp\{-M^2 \log n/(K + n^{-1/4}b(\log n)^{1/2-c-d}2M)\} \\
\le 2^{p+2}B^{p+1} \exp\Big\{\frac{p}{4} + \frac{5}{4} - \frac{M^2}{4K}\log n\Big\} \quad \text{for } n \text{ large} \\
\to 0 \quad \text{for } M > \sqrt{(p+5)K},
\end{aligned}$$

and the result obtains.

Note that when $d = 0$, the rate of convergence is $n^{-3/4}(\log n)^{c+1/2}$ rather than $n^{-1}(\log n)^{c+1/2}$ as one might expect. We get the slower rate because, under the conditions of the Lemma, $n^{-1}\sum_{j=1}^{n} c_{jn}^2|x_j|^2 = O(n^{1/2})$. If we make the stronger assumption that $n^{-1}\sum_{j=1}^{n} c_{jn}^2|x_j|^2 = O(1)$, the above argument yields the faster rate.

We are now able to establish the second preliminary result.

Lemma 3.2. *Suppose that* $\{x_j\}$ *satisfies* (i) *and* (ii) *of the Theorem and that* $n^{1/2}(\theta_n - \theta_0)$ *is bounded in probability. For any* $a > 0$, $a + c < 1$, *and* $0 \le \rho < K(\log n)^{-(a+c)}$, *define* $U_n(\rho)$ *by* (2.4). *Then*

$$\sup_{u\in U_n(\rho)} |\widehat{Q}_n(u) - Q_n(u)| = O_p(n^{-1/2}(\log n)^{2a+c}).$$

Proof. It follows from Lemma 3.1 with $\tau = 0$, $d = 0$ and $c_{jn} = 1$, $1 \le j \le n$, that

$$\begin{aligned}
\sup_{|y|\le D(\log n)^a} \big|\widehat{G}_n(y) - F_n(y)\big| &= \sup_{|y|\le D(\log n)^a} \Big|\frac{1}{n}\sum_{j=1}^{n} F_n(y + x_j'(\theta_n - \theta_0)) - F_n(y)\Big| \\
&\quad + O_p(n^{-3/4}(\log n)^{c+1/2} + n^{-1/2}) \\
&= O_p(n^{-1/2})
\end{aligned}$$

using a two-term Taylor series expansion. Hence, uniformly in $u \in U_n(\rho)$, on a set whose probability can be made arbitrarily close to one uniformly in n by taking K_2 sufficiently large,

$$\begin{aligned}
\widehat{G}_n\{Q_n(u) + n^{-1/2}(\log n)^{2a+c}K_1\} &\ge F_n\{Q_n(u) + n^{-1/2}(\log n)^{2a+c}K_1\} - n^{-1/2}K_2 \\
&\ge u + n^{-1/2}KK_1 - n^{-1/2}K_2 \\
&\ge u \qquad \text{for} \quad KK_1 > K_2
\end{aligned}$$

by (2.5), so uniformly in $u \in U_n(\rho)$, on a set whose probability can be made arbitrarily close to one uniformly in n,

$$\widehat{Q}_n(u) - Q_n(u) \leq n^{-1/2}(\log n)^{2a+c}K_1.$$

Similarly, uniformly in $u \in U_n(\rho)$, on a set whose probability can be made arbitrarily close to one uniformly in n,

$$\begin{aligned}\widehat{G}_n\{Q_n(u)+n^{-1/2}(\log n)^{2a+c}K_1\} &\leq F_n\{Q_n(u)+n^{-1/2}(\log n)^{2a+c}K_1\}+n^{-1/2}K_2\\ &\leq u - n^{-1/2}(\log n)^{2(a+c)}K_1K_2 + n^{-1/2}K_2\\ &\leq u\end{aligned}$$

by (2.1) and the result obtains.

Now consider the kernel estimators used in the estimator (1.4).

Lemma 3.3. *Suppose that k_1 and k_2 are kernel functions which satisfy* (1.2), *that $\{x_j\}$ satisfies conditions* (i) *and* (ii) *of the Theorem, and $n^{1/2}(\theta_n - \theta_0)$ is bounded in probability. Then for any $a > 0$, $a + c < 1$, and $\delta_i \sim n^{-\varepsilon_i}$, $0 < \varepsilon_i < 1/(2i)$,*

$$\sup_{u\in U_n(2/n)} |\widehat{Q}_n^{(i)}(u) - Q_n^{(i)}(u)| = O_p(n^{-\varepsilon_i}(\log n)^{2+4i} + n^{-1/2+i\varepsilon_i}(\log n)^2), \quad i = 1, 2,$$

where $U_n(\cdot)$ is defined by (2.4).

Proof. Note that by a Taylor expansion,

$$\begin{aligned}&\sup_{u\in U_n(2/n)}\left|\delta_i^{-1}\int_{-1}^{1} Q_n(u-\delta_i x)k_i(x)dx - Q_n^{(i)}(u)\right|\\ &\qquad \leq \frac{\delta_i}{(i+1)!}\sup_{u\in U_n(\delta_i+2/n)}|Q_n^{(i+1)}(u)|\int_{-1}^{1}|x^{i+1}k_i(x)|dx\\ &\qquad \leq \delta_i(\log n)^{2a+c+4i(a+c)}K\end{aligned}$$

by (2.7) as $\delta_i + 2/n \leq K(\log n)^{-a-c}$. Also

$$\begin{aligned}&\sup_{u\in U_n(2/n)}\left|\delta_i^{-1-i}\int_0^1\{\widehat{Q}_n(w) - Q_n(w)\}k_i(\delta_i^{-1}(u-w))dw\right|\\ &\qquad = \sup_{u\in U_n(2/n)}\left|\delta_i^{-i}\int_{-1}^{1}\{\widehat{Q}_n(u-\delta_i x) - Q_n(u-\delta_i x)\}k_i(x)dx\right|\\ &\qquad \leq \delta_i^{-i}\sup_{u\in U_n(\delta_i+2/n)}|\widehat{Q}_n(u) - Q_n(u)|\int_{-1}^{1}|k_i(x)|dx\\ &\qquad = O_p(n^{-1/2+i\varepsilon_i}(\log n)^{2a+c})\end{aligned}$$

by Lemma 3.2. The Lemma obtains.

We shall also require the following result.

Lemma 3.4. *Suppose that k_1 and k_2 are kernel functions which satisfy* (1.2), *that $\{x_j\}$ satisfies conditions* (i) *and* (ii) *of the Theorem, and $n^{1/2}(\theta_n - \theta_0)$ is bounded in probability. Then for any $a > 0$, $a + c < 1$, and $\delta_i \sim n^{-\varepsilon_i}$, $0 < \varepsilon_i < 1/(4+2i)$,*

$$n \sup_{u\in U_n(1/n)} \left|\widehat{Q}_n^{(i)}(u+n^{-1}) - \widehat{Q}_n^{(i)}(u) - Q_n^{(i)}(u+n^{-1}) + Q_n^{(i)}(u)\right|$$
$$= O_p\left(n^{-\varepsilon_i}(\log n)^{2+4^{i+1}} + n^{-1/2+(2+i)\varepsilon_i}(\log n)^2\right),$$

$i = 1, 2$, where $U_n(\cdot)$ is defined by (2.4).

Proof. Note that by (2.7)

$$\sup_{u\in U_n(1/n)} \left|\delta_i^{-i}\int_{-1}^{1}\{Q_n(u+n^{-1}-\delta_i x) - Q_n(u-\delta_i x)\}k_i(x)dx - Q_n^{(i)}(u+n^{-1}) + Q_n^{(i)}(u)\right|$$
$$\le \sup_{u\in U_n(1/n)} \left|\delta_i^{-i}\int_{-1}^{1} n^{-1}Q_n^{(1)}(u-\delta_i x)k_i(x)dx - n^{-1}Q_n^{(i+1)}(u)\right|$$
$$+ n^{-2}\{\delta_i^{-i}(\log n)^{6a+5c} + (\log n)^{2a+c+4(i+1)(a+c)}\}K$$
$$\le n^{-1}\sup_{u\in U_n(1/n)}\left|\frac{\delta_i}{(i+1)!}\int_{-1}^{1} Q_n^{(i+2)}(u-\widetilde{\delta}_i x)k_i(x)dx\right| + n^{-2+i\varepsilon_i}(\log n)^{6a+5c}K$$
$$\le n^{-1}\delta_i(\log n)^{2a+c+4(i+1)(a+c)}K, \quad 0 \le \widetilde{\delta}_i \le \delta_i.$$

Also,

$$\sup_{u\in U_n(1/n)}\left|\delta_i^{-1-i}\int_0^1 \widehat{Q}_n(w)\{k_i\{\delta_i^{-1}(u+n^{-1}-w)\} - k_i\{\delta_i^{-1}(u-w)\}\}dw\right.$$
$$\left. - \delta_i^{-i}\int_{-1}^{1}\{Q_n(u+n^{-1}-\delta_i x) - Q_n(u-\delta_i x)\}k_i(x)dx\right|$$
$$= \sup_{u\in U_n(1/n)}\left|\delta_i^{-1-i}\int_0^1\{\widehat{Q}_n(w)-Q_n(w)\}\{k_i\{\delta_i^{-1}(u+n^{-1}-w)\}-k_i\{\delta_i^{-1}(u-w)\}\}dw\right|$$
$$\le \sup_{u\in U_n(1/n)}\left|\delta_i^{-1-i}\int_{u-\delta_i+n^{-1}}^{u+\delta_i}\{\widehat{Q}_n(w)-Q_n(w)\}\{k_i\{\delta_i^{-1}(u+n^{-1}-w)\}-k_i\{\delta_i^{-1}(u-w)\}\}dw\right|$$
$$+ \sup_{u\in U_n(1/n)}\left|\delta_i^{-1-i}\int_{u-\delta_i}^{u-\delta_i+n^{-1}}\{\widehat{Q}_n(w) - Q_n(w)\}k_i\{\delta_i^{-1}(u-w)\}dw\right|$$

$$+ \sup_{u \in U_n(1/n)} \left| \delta_i^{-1-i} \int_{u+\delta_i}^{u+\delta_i+n^{-1}} \{\widehat{Q}_n(w) - Q_n(w)\} k_i\{\delta_i^{-1}(u+n^{-1}-w)\} dw \right|$$
$$\leq \sup_{u \in U_n(\delta_i + 2n^{-1})} |\widehat{Q}_n(u) - Q_n(u)| \{n^{-1+(2+i)\varepsilon_i} K + 2n^{-1+(1+i)\varepsilon_i} K\}$$
$$\leq n^{-3/2+(2+i)\varepsilon_i} (\log n)^{2a+c} K$$

by Lemma 3.2. The lemma obtains.

For future reference, note that the conditions of Lemma 3.4 imply those of Lemma 3.3.

We now apply Lemmas 3.3 and 3.4 to obtain the following results for $\widehat{\eta}_n$.

Lemma 3.5. *Suppose that the conditions of Lemma 3.3 hold. Then*

$$\sup_{u \in U_n(1/n)} |\widehat{\eta}_n(u) - \eta_n(u)| = O_p(n^{-\sigma}),$$

where $0 < \sigma < \min\{\varepsilon_1, \frac{1}{2} - \varepsilon_1, \varepsilon_2, \frac{1}{2} - 2\varepsilon_2\}$ *and* $U_n(\cdot)$ *is defined by* (2.4).

Proof. Write

$$\sup_{u \in U_n(1/n)} |\widehat{\eta}_n(u) - \eta_n(u)|$$
$$\leq \sup_{u \in U_n(1/n)} |\widehat{Q}_n^{(2)}(u) - Q_n^{(2)}(u)| / Q_n^{(1)}(u)^2$$
$$+ \sup_{u \in U_n(1/n)} |\widehat{Q}_n^{(1)}(u) - Q_n^{(1)}(u)| \, |\widehat{Q}_n^{(2)}(u)| \, |Q_n^{(1)}(u) + \widehat{Q}_n^{(1)}(u)| / \widehat{Q}_n^{(1)}(u)^2 Q_n^{(1)}(u)^2$$
$$= O_p\big(n^{-\varepsilon_2}(\log n)^{12} + n^{-1/2+2\varepsilon_2}(\log n)^4 + n^{-\varepsilon_1}(\log n)^{16} + n^{-1/2+\varepsilon_1}(\log n)^{12}\big)$$

by Lemma 3.3, (2.6) and (2.7).

Lemma 3.6. *Suppose that the conditions of Lemma 3.4 hold. Then with* $\alpha = F_n(-(\log n)^a)$, $\beta = F_n((\log n)^a)$ *for* $a > 0$, $a + c < 1$,

$$\sum_{i=[n\alpha]+1}^{[n\beta]} \left| \widehat{\eta}_n\Big(\frac{i+1}{n}\Big) - \widehat{\eta}_n\Big(\frac{i}{n}\Big) - \eta_n\Big(\frac{i+1}{n}\Big) + \eta_n\Big(\frac{i}{n}\Big) \right| = O_p(n^{-\kappa}),$$

where $0 < \kappa < \min\{\varepsilon_1, \frac{1}{2} - 3\varepsilon_1, \varepsilon_2, \frac{1}{2} - 4\varepsilon_2\}$.

Proof. Let

$$\Delta\widehat{Q}_n^{(i)}(u) = \widehat{Q}_n^{(i)}(u + n^{-1}) - \widehat{Q}_n^{(i)}(u)$$

and

$$\Delta Q_n^{(i)}(u) = Q_n^{(i)}(u + n^{-1}) - Q_n^{(i)}(u), \quad 0 < u < 1, \quad i = 1, 2.$$

Then

$$\sum_{i=[n\alpha]+1}^{[n\beta]} \left|\widehat{\eta}_n\Big(\frac{i+1}{n}\Big) - \widehat{\eta}_n\Big(\frac{i}{n}\Big) - \eta_n\Big(\frac{i+1}{n}\Big) + \eta_n\Big(\frac{i}{n}\Big)\right|$$

$$\le n \sup_{u\in U_n(1/n)} |\widehat{\eta}_n(u+n^{-1}) - \widehat{\eta}_n(u) - \eta_n(u+n^{-1}) + \eta_n(u)|$$

$$\le n \sup_{u\in U_n(1/n)} \left|\Delta \widehat{Q}_n^{(2)}(u) - \Delta Q_n^{(2)}(u)\right| / Q_n^{(1)}(u)^2$$

$$+ n \sup_{u\in U_n(1/n)} \left|\Delta \widehat{Q}_n^{(2)}(u)\right| \sup_{u\in U_n(2/n)} \left|\widehat{Q}_n^{(1)}(u) - Q_n^{(1)}(u)\right| \frac{|\widehat{Q}_n^{(1)}(u) + Q_n^{(1)}(u)|}{\{\widehat{Q}_n^{(1)}(u) Q_n^{(1)}(u)\}^2}$$

$$+ \sup_{u\in U_n(1/n)} \left|\widehat{Q}_n^{(2)}(u) - Q_n^{(2)}(u)\right| \, |D_n(u)| \, n \left|\Delta Q_n^{(1)}(u)\right|$$

$$+ \sup_{u\in U_n(1/n)} \left|\widehat{Q}_n^{(2)}(u)\right| \, |\widehat{D}_n(u) - D_n(u)| \, n \left|\Delta Q_n^{(1)}(u)\right|$$

$$+ \sup_{u\in U_n(1/n)} \left|\widehat{Q}_n^{(2)}(u)\right| \, |\widehat{D}_n(u)| \, n \left|\Delta \widehat{Q}_n^{(1)}(u) - \Delta Q_n^{(1)}(u)\right|,$$

where $\widehat{D}_n(u) = \{\widehat{Q}_n^{(1)}(u) + \widehat{Q}_n^{(1)}(u+n^{-1})\}/\{\widehat{Q}_n^{(1)}(u)\widehat{Q}_n^{(1)}(u+n^{-1})\}^2$ and$D_n(u) = \{Q_n^{(1)}(u) + Q_n^{(1)}(u+n^{-1})\}/\{Q_n^{(1)}(u)Q_n^{(1)}(u+n^{-1})\}^2$, $0 < u < 1$. The result obtains from Lemmas 3.3 and 3.4 and the bounds (2.6) and (2.7).

4. Properties of λ_n

Throughout this section, we assume that $\{x_j\}$ satisfies (i) and (ii) of the Theorem, $n^{1/2}(\theta_n - \theta_0)$ is bounded in probability and $\widehat{\eta}$ satisfies the conclusion of Lemmas 3.5 and 3.6. For $a, c > 0$, $a + c < 1$, and $0 < b < 1/2$, let

$$\widehat{\alpha} = \widehat{G}_n(-(\log n)^a) + n^{-b} \qquad \alpha = F_n(-(\log n)^a)$$
$$\widehat{\beta} = \widehat{G}_n((\log n)^a) - n^{-b} \qquad \beta = F_n((\log n)^a)$$
$$\gamma = (\log n)^{-c}.$$

Note that by Lemma 3.1, $\widehat{G}_n(\pm(\log n)^a) - F_n(\pm(\log n)^a) = O_p(n^{-1/2})$; see the proof of Lemma 3.2. It follows then that for n large enough, on a set whose probability can be made arbitrarily close to one,

$$\alpha \le \widehat{\alpha} \le \alpha + n^{-b} \text{ and } \beta - n^{-b} \le \widehat{\beta} \le \beta. \tag{4.1}$$

It will be established that the Theorem holds under these conditions and for these choices of $\widehat{\alpha}$, $\widehat{\beta}$ and γ.

We prove the theorem by proving a sequence of lemmas which establish the behaviour of the terms in the decomposition (1.10) as described in Section 1.3.

Lemma 4.1. *$\widehat{I}_n = I_n + o_p(1)$ and $\widehat{I}_n = I + o_p(1)$.*

Proof. Let $I_n^* = n^{-1}\sum_{i=[n\alpha]+1}^{[n\beta]} \eta_n(i/n)^2$. Then by (4.1), with probability which can be made arbitrarily close to one,

$$\begin{aligned}|\widehat{I}_n - I_n^*| &\le \frac{1}{n}\sum_{i=[n\alpha]+1}^{[n\beta]} \left|\widehat{\eta}_n\Big(\frac{i}{n}\Big)^2 - \eta_n\Big(\frac{i}{n}\Big)^2\right| + \frac{1}{n}\sum_{i=[n\alpha]+1}^{[n\widehat{\alpha}]} \eta_n\Big(\frac{i}{n}\Big)^2 + \frac{1}{n}\sum_{i=[n\widehat{\beta}]+1}^{[n\beta]} \eta_n\Big(\frac{i}{n}\Big)^2\\ &\le \sup_{u\in U_n(1/n)} |\widehat{\eta}_n(u) - \eta_n(u)|\{|\widehat{\eta}_n(u)| + |\eta_n(u)|\} + 2n^{-b}\sup_{u\in U_n(1/n)} \eta_n(u)^2\\ &\xrightarrow{P} 0\end{aligned}$$

by Lemma 3.5, (2.8) and (4.1). Also, by (2.7) and (2.8),

$$|I_n - I_n^*| \le \Big|\sum_{i=[n\alpha]+1}^{[n\beta]} \Big\{\int_{(i-1)/n}^{i/n} \eta_n(u)^2 du - n^{-1}\eta_n\Big(\frac{i}{n}\Big)^2\Big\}\Big| \le Kn^{-1}(\log n)^{8a+10c}.$$

The result then obtains from Theorem 4.1 of Stone (1975).

Lemma 4.2. *$R_1 \xrightarrow{P} 0$.*

Proof. Note that

$$\begin{aligned}&\mathrm{Var}\Big[n^{-1/2}\sum_{i=[n\widehat{\beta}]+1}^{[n\beta]} \eta_n\Big(\frac{i}{n}\Big)\sum_{j=1}^{n} x_j\Big\{G\Big\{\Big(Q_n\Big(\frac{i}{n}\Big)-e_j\Big)/\gamma\Big\} - G\Big\{\Big(Q_n\Big(\frac{i-1}{n}\Big)-e_j\Big)/\gamma\Big\}\Big\}\Big]\\ &= n^{-1}\sum_{j=1}^{n} x_j^2 E\Big[\sum_{i=1}^{n} I([n\widehat{\beta}] < i \le [n\beta])\eta_n(i/n)\{G\{(Q_n(i/n) - e)/\gamma\}\\ &\qquad\qquad - G\{(Q_n((i-1)/n) - e)/\gamma\}\}\Big]^2\\ &\le KE\Big[\sum_{i=1}^{n} I([n\widehat{\beta}] < i \le [n\beta])\eta_n(i/n)\frac{g\{(Q_n(i/n) - e)/\gamma\}}{n\gamma f_n(Q_n((i-\theta)/n))}\Big]^2, \qquad |\theta| \le 1,\\ &\le K(\log n)^{8a+10c} n^{-1}([n\beta] - [n\beta - n^{1-b}])\\ &\to 0,\end{aligned}$$

as $n \to \infty$ by (2.6) and (2.8), so the first term in R_1 converges in probability to zero.

The remaining term in R_1 may be treated via a similar argument and the Lemma obtains.

Lemma 4.3. $R_2 \xrightarrow{P} 0$.

Proof. Let $v \in \mathbf{R}^p$ be any fixed vector and let $c_j = \max\{v'x_j, 0\}$ and $d_j = c_j - v'x_j$ so that $v'x_j = c_j - d_j$, $1 \le j \le n$. Then write

$$\begin{aligned}
P_n(u) &= n^{-1/2}\sum_{j=1}^{n} v'x_j\{G\{(F_n^{-1}(u) - e_j)/\gamma\} - u\}, \qquad 0 \le u \le 1. \\
&= n^{1/2}\bar{c}(n\bar{c})^{-1}\sum_{j=1}^{n} c_j\Big\{G\{(F_n^{-1}(u) - e_j)/\gamma\} - u\} \\
&\qquad\qquad - n^{-1/2}\bar{d}(n\bar{d})^{-1}\sum_{j=1}^{n} d_j\{G\{(F_n^{-1}(u) - e_j)/\gamma\} - u\Big\} \\
&= n^{1/2}\bar{c}\int_{-\infty}^{\infty} G\{(F_n^{-1}(u) - y)/\gamma\}d\{P_n^c(y) - F(y)\} \\
&\qquad\qquad - n^{1/2}\bar{d}\int_{-\infty}^{\infty} G\{(F_n^{-1}(u) - y)/\gamma\}d\{P_n^d(y) - F(y)\}, \quad (4.2)
\end{aligned}$$

where $P_n^c(y) = (n\bar{c})^{-1}\sum_{j=1}^{n} c_j I(e_j \le y)$ and $P_n^d(y) = (n\bar{d})^{-1}\sum_{j=1}^{n} d_j I(e_j \le y)$. Now by a straightforward application of Theorem 15.6 of Billingsley (1968), it follows that $\bar{c}n^{1/2}\{P_n^c(\cdot) - F(\cdot)\}$ and $\bar{d}n^{1/2}\{P_n^d(\cdot) - F(\cdot)\}$ converge weakly in $D[-\infty,\infty]$ to mean zero Gaussian processes and hence that

$$\sup_{-\infty\le y\le\infty} \bar{c}n^{1/2}|P_n^c(y) - F(y)| = O_p(1) \quad\text{and}$$
$$\sup_{-\infty\le y\le\infty} \bar{d}n^{1/2}|P_n^d(y) - F(y)| = O_p(1).$$

Then, integrating by parts,

$$\begin{aligned}
&\sup_{0\le u\le 1}\left|n^{1/2}\bar{c}\int_{-\infty}^{\infty} G\{(F_n^{-1}(u) - y)/\gamma\}d\{P_n^c(y) - F(y)\}\right| \\
&= \sup_{0\le u\le 1}\left|n^{1/2}\bar{c}\int_{-\infty}^{\infty}\{P_n^c(y) - F(y)\}dG\{(F_n^{-1}(u) - y)/\gamma\}\right| \\
&\le \sup_{-\infty\le y\le\infty}\left|\bar{c}n^{1/2}\{P_n^c(y) - F(y)\}\right| \sup_{0\le u\le 1}\left|\int_{-\infty}^{\infty} dG\{(F_n^{-1}(u) - y)/\gamma\}\right| \\
&= O_p(1).
\end{aligned}$$

A similar result holds for the second term in (4.2) so that

$$\sup_{0\le u\le 1} |P_n(u)| = O_p(1). \tag{4.3}$$

Now, summing by parts, it follows that with probability which can be made arbitrarily close to one

$$\begin{aligned}
&|R_2|\\
&\le \Big| n^{-1/2} \sum_{i=[n\widehat{\alpha}]+1}^{[n\widehat{\beta}]} \Big\{\widehat{\eta}_n\Big(\frac{i+1}{n}\Big) - \widehat{\eta}_n\Big(\frac{i}{n}\Big) - \eta_n\Big(\frac{i+1}{n}\Big) + \eta_n\Big(\frac{i}{n}\Big)\Big\} \sum_{j=1}^n x_j G\Big\{\Big(Q_n\Big(\frac{i}{n}\Big) - e_j\Big)/\gamma\Big\}\Big|\\
&\quad + \Big|\{\widehat{\eta}_n(([n\widehat{\beta}]+1)/n) - \eta_n(([n\widehat{\beta}]+1)/n)\} n^{-1/2} \sum_{j=1}^n x_j G\{(Q_n([n\widehat{\beta}]/n) - e_j)/\gamma\}\Big|\\
&\quad + \Big|\{\widehat{\eta}_n(([n\widehat{\alpha}]+1)/n) - \eta_n(([n\widehat{\alpha}]+1)/n)\} n^{-1/2} \sum_{j=1}^n x_j G\{(Q_n([n\widehat{\alpha}]/n) - e_j)/\gamma\}\Big|\\
&\le \sum_{[n\alpha]+1}^{[n\beta]} \Big|\widehat{\eta}_n\Big(\frac{i+1}{n}\Big) - \widehat{\eta}_n\Big(\frac{i}{n}\Big) - \eta_n\Big(\frac{i+1}{n}\Big) + \eta_n\Big(\frac{i}{n}\Big)\Big| \sup_{|y|\le M(\log n)^a} \Big| n^{-1/2} \sum_{j=1}^n x_j G\Big(\frac{y-e_j}{\gamma}\Big)\Big|\\
&\quad + \sup_{u\in U_n(1/n)} |\widehat{\eta}_n(u) - \eta_n(u)| \Big| n^{-1/2} \sum_{j=1}^n x_j \Big\{G\Big\{\Big(Q_n\Big(\frac{[n\widehat{\beta}]}{n}\Big) - e_j\Big)/\gamma\Big\} + G\Big\{\Big(Q_n\Big(\frac{[n\widehat{\alpha}]}{n}\Big) - e_j\Big)/\gamma\Big\}\Big\}\Big|\\
&\xrightarrow{\ \mathrm{P}\ } 0
\end{aligned}$$

by (4.1), (4.3) and Lemmas 3.5 and 3.6.

Lemma 4.4. $R_3 \xrightarrow{\mathrm{P}} 0$.

Proof. Without loss of generality, let $\theta_0 = 0$. Then summing by parts,

$$\begin{aligned}
R_3 &= n^{-1/2} \sum_{j=1}^n x_j \Big[x_j'\theta_n + \widehat{I}_n^{-1} \sum_{i=[n\widehat{\alpha}]+1}^{[n\widehat{\beta}]} \Big\{\widehat{\eta}_n\Big(\frac{i+1}{n}\Big) - \widehat{\eta}_n\Big(\frac{i}{n}\Big)\Big\} F_n\Big(\widehat{Q}_n\Big(\frac{i}{n}\Big) + x_j'\theta_n\Big)\\
&\qquad - \widehat{I}_n^{-1} x_j'\theta_n \eta_n\Big(\frac{[n\widehat{\beta}]+1}{n}\Big) f_n\Big(Q_n\Big(\frac{[n\widehat{\beta}]}{n}\Big)\Big) + \widehat{I}_n^{-1} x_j'\theta_n \eta_n\Big(\frac{[n\widehat{\alpha}]+1}{n}\Big) f_n\Big(Q_n\Big(\frac{[n\widehat{\alpha}]}{n}\Big)\Big)\Big]\\
&\quad + n^{-1/2} \widehat{I}_n^{-1} \sum_{j=1}^n x_j \Big\{\widehat{\eta}_n\Big(\frac{[n\widehat{\alpha}]+1}{n}\Big) F_n\Big(\widehat{Q}_n\Big(\frac{[n\widehat{\alpha}]}{n}\Big) + x_j'\theta_n\Big) - x_j'\theta_n \eta_n\Big(\frac{[n\widehat{\alpha}]+1}{n}\Big) f_n\Big(Q_n\Big(\frac{[n\widehat{\alpha}]}{n}\Big)\Big)\Big\}\\
&\quad - n^{-1/2} \widehat{I}_n^{-1} \sum_{j=1}^n x_j \Big\{\widehat{\eta}_n\Big(\frac{[n\widehat{\beta}]+1}{n}\Big) F_n\Big(\widehat{Q}_n\Big(\frac{[n\widehat{\beta}]}{n}\Big) + x_j'\theta_n\Big) - x_j'\theta_n \eta_n\Big(\frac{[n\widehat{\beta}]+1}{n}\Big) f_n\Big(Q_n\Big(\frac{[n\widehat{\beta}]}{n}\Big)\Big)\Big\}\\
&= R_{31} + R_{32} + R_{33},
\end{aligned}$$

say. Now, for $1 \le j \le n$ and $[n\alpha] \le i \le [n\beta]$,

$$F_n\Big(\widehat{Q}_n\Big(\frac{i}{n}\Big)+x_j'\theta_n\Big)=\frac{i}{n}+\Big\{x_j'\theta_n+\widehat{Q}_n\Big(\frac{i}{n}\Big)-Q_n\Big(\frac{i}{n}\Big)\Big\}f_n\Big(Q_n\Big(\frac{i}{n}\Big)\Big)+L_j\Big(\frac{i}{n}\Big), \quad (4.4)$$

where

$$\max_{[n\alpha]\le i\le[n\beta]} L_j\Big(\frac{i}{n}\Big) \le O_p(1)\{n^{-1}|x_j|^2+n^{-1}(\log n)^4\}(\log n)^2,$$

by Lemma 3.2, (2.1) and the assumption that $|\theta_n| = O_p(n^{-1/2})$. Hence, using the fact that $\sum_{j=1}^n x_j = 0$,

$$\begin{aligned}|R_{32}| &= \widehat{I}_n^{-1}\Big|n^{-1/2}\sum_{j=1}^n x_jx_j'\theta_n\{\widehat{\eta}_n(([n\widehat{\alpha}]+1)/n)-\eta_n(([n\widehat{\alpha}]+1)/n)\}\\ &\qquad\times f_n(Q_n([n\widehat{\alpha}]/n))+\frac{1}{2}n^{-1/2}\sum_{j=1}^n x_jL_j([n\widehat{\alpha}]/n)\widehat{\eta}_n(([n\widehat{\alpha}]+1)/n)\Big|\\ &\le O_p(1)\{n^{-\sigma}(\log n)+n^{-1/4}(\log n)^3+n^{-1/2}(\log n)^6\}\\ &\xrightarrow{P} 0\end{aligned}$$

by (4.1), (2.1), Lemma 3.5, (2.8), (2.4) and conditions (i) and (ii). Similarly, $R_{33} \xrightarrow{P} 0$ and it remains to show that $R_{31} \xrightarrow{P} 0$.

By the expansion (4.4) and the fact that $\sum_{j=1}^n x_j = 0$, with probability which can be made arbitrarily close to one

$$\begin{aligned}&|R_{31}|\\ &=\Big|n^{-1/2}\sum_{j=1}^n x_jx_j'\theta_n\Big[1+\widehat{I}_n^{-1}\sum_{i=[n\widehat{\alpha}]+1}^{[n\widehat{\beta}]}\Big\{\widehat{\eta}_n\Big(\frac{i+1}{n}\Big)-\widehat{\eta}_n\Big(\frac{i}{n}\Big)\Big\}f_n\Big(Q_n\Big(\frac{i}{n}\Big)\Big)\\ &\quad-\widehat{I}_n^{-1}\eta_n(([n\widehat{\beta}]+1)/n)f_n(Q_n([n\widehat{\beta}]/n))+\widehat{I}_n^{-1}\eta_n(([n\widehat{\alpha}]+1)/n)f_n(Q_n([n\widehat{\alpha}]/n))\Big]\\ &\qquad+n^{-1/2}\widehat{I}_n^{-1}\sum_{j=1}^n x_j\sum_{i=[n\widehat{\alpha}]+1}^{[n\widehat{\beta}]}\Big\{\widehat{\eta}_n\Big(\frac{i+1}{n}\Big)-\widehat{\eta}_n\Big(\frac{i}{n}\Big)\Big\}L_j\Big(\frac{i}{n}\Big)\Big|\\ &\le O_p(1)n^{-1}\widehat{I}_n^{-1}\sum_{i=[n\alpha]+1}^{[n\widehat{\alpha}]}\eta_n\Big(\frac{i}{n}\Big)^2+O_p(1)n^{-1}\widehat{I}_n^{-1}\sum_{i=[n\widehat{\beta}]+1}^{[n\beta]}\eta_n\Big(\frac{i}{n}\Big)^2\end{aligned}$$

$$+ O_p(1)(\log n)^2 \widehat{I}_n^{-1} \sum_{i=[n\alpha]+1}^{[n\beta]} \left|\widehat{\eta}_n\Big(\frac{i+1}{n}\Big) - \widehat{\eta}_n\Big(\frac{i}{n}\Big) - \eta_n\Big(\frac{i+1}{n}\Big) + \eta_n\Big(\frac{i}{n}\Big)\right|$$

$$+ O_p(1)\widehat{I}_n^{-1}\Big| \sum_{i=[n\widehat{\alpha}]+1}^{[n\widehat{\beta}]} \Big\{\eta_n\Big(\frac{i+1}{n}\Big) - \eta_n\Big(\frac{i}{n}\Big)\Big\} f_n\Big(Q_n\Big(\frac{i}{n}\Big)\Big) + \frac{1}{n} \sum_{i=[n\widehat{\alpha}]+1}^{[n\widehat{\beta}]} \eta_n\Big(\frac{i}{n}\Big)^2$$

$$- \eta_n(([n\widehat{\beta}]+1)/n) f_n(Q_n([n\widehat{\beta}]/n)) + \eta_n(([n\widehat{\alpha}]+1)/n) f_n(Q_n([n\widehat{\alpha}]/n))\Big|$$

$$+ O_p(1)\Big\{n^{-1/4}(\log n)^2 + n^{-1/2}(\log n)^6\Big\} \widehat{I}_n^{-1} \sum_{i=[n\alpha]+1}^{[n\beta]} \left|\widehat{\eta}_n\Big(\frac{i+1}{n}\Big) - \widehat{\eta}_n\Big(\frac{i}{n}\Big)\right|$$

$$\le O_p(1)\widehat{I}_n^{-1}\Big|\frac{1}{n} \sum_{i=[n\widehat{\alpha}]+1}^{[n\widehat{\beta}]} \eta_n\Big(\frac{i}{n}\Big)^2 - \sum_{i=[n\widehat{\alpha}]+1}^{[n\widehat{\beta}]} \eta_n\Big(\frac{i}{n}\Big)\Big\{f_n\Big(Q_n\Big(\frac{i}{n}\Big)\Big) - f_n\Big(Q_n\Big(\frac{i-1}{n}\Big)\Big)\Big\}\Big| + o_p(1)$$

by Lemmas 4.1 and 4.2, (4.1), (2.8), Lemma 3.6 and (2.9). The result follows on expanding $f_n(Q_n((i-1)/n))$ to two terms and then applying (2.8) and (2.9).

Lemma 4.5. $R_4 \xrightarrow{\text{P}} 0$.

Proof. Without loss of generality, let $\theta_0 = 0$. Let $v \in \mathbf{R}^p$ be any fixed vector and let

$$H_{ni}(\tau, t) = n^{-1/2} \sum_{j=1}^{n} v' x_j \{G\{(Q_n(i/n) - e_j + x_j' t + \tau)/\gamma\}$$
$$- F_n(Q_n(i/n) + x_j' t + \tau) - G\{(Q_n(i/n) - e_j)/\gamma\}\}.$$

Then, summing by parts, it follows that with probability which can be made arbitrarily close to one

$$|R_4| = \Big|\widehat{I}_n^{-1} \sum_{i=[n\widehat{\alpha}]+1}^{[n\widehat{\beta}]} \{\widehat{\eta}_n((i+1)/n) - \widehat{\eta}_n(i/n)\} H_{ni}\{\widehat{Q}_n(i/n) - Q_n(i/n), \theta_n\}$$
$$+ \widehat{I}_n^{-1} \widehat{\eta}_n(([n\widehat{\alpha}]+1)/n) H_{n,[n\widehat{\alpha}]/n}\{\widehat{Q}_n([n\widehat{\alpha}]/n) - Q_n([n\widehat{\alpha}]/n), \theta_n\}$$
$$- \widehat{I}_n^{-1} \widehat{\eta}_n(([n\widehat{\beta}]+1)/n) H_{n,[n\widehat{\beta}]/n}\{\widehat{Q}_n([n\widehat{\beta}]/n) - Q_n([n\widehat{\beta}]/n), \theta_n\}\Big|$$

$$\le \widehat{I}_n^{-1} \sum_{i=[n\alpha]+1}^{[n\beta]} \left|\widehat{\eta}_n\Big(\frac{i+1}{n}\Big) - \widehat{\eta}_n\Big(\frac{i}{n}\Big)\right| \max_{[n\alpha]\le i\le[n\beta]} \left|H_{ni}\Big\{\widehat{Q}_n\Big(\frac{i}{n}\Big) - Q_n\Big(\frac{i}{n}\Big), \theta_n\Big\}\right|$$
$$+ O_p(1)(\log n)^3 |H_{n,[n\widehat{\alpha}]/n}\{\widehat{Q}_n([n\widehat{\alpha}]/n) - Q_n([n\widehat{\alpha}]/n), \theta_n\}|$$
$$+ O_p(1)(\log n)^3 |H_{n,[n\widehat{\beta}]/n}\{\widehat{Q}_n([n\widehat{\beta}]/n) - Q_n([n\widehat{\beta}]/n), \theta_n\}|$$

by (4.1), Lemma 3.5 and (2.8). Now

$$\max_{[n\alpha]\le i\le[n\beta]} |H_{ni}\{\widehat{Q}_n(i/n) - Q_n(i/n), \theta_n\}| = O_p(n^{-3/4}(\log n)^4)$$

by Lemmas 3.2 and 3.1 so the result follows from Lemma 3.6 and (2.11).

Lemma 4.6. $T \xrightarrow{\mathfrak{D}} \mathrm{N}(0, I^{-1}\Gamma^{-1})$.

Proof. Let $v \in \mathbf{R}^p$ be any fixed vector and write

$$T = -\widehat{I}_n^{-1} n^{-1/2} \sum_{j=1}^{n} v'\Gamma^{-1} x_j \sum_{i=[n\alpha]+1}^{[n\beta]} \eta_n\Big(\frac{i}{n}\Big)\Big\{G\Big\{\Big(Q_n\Big(\frac{i}{n}\Big) - e_j\Big)/\gamma\Big\} - G\Big\{\Big(Q_n\Big(\frac{i-1}{n}\Big) - e_j\Big)/\gamma\Big\}\Big\}$$

$$= \widehat{I}_n^{-1} n^{-1/2} \sum_{j=1}^{n} v'\Gamma^{-1} x_j n^{-1} \sum_{i=[n\alpha]+1}^{[n\beta]} \eta_n(i/n) \frac{g\{(Q_n(i/n) - e_j)/\gamma\}}{\gamma f_n(Q_n(i/n))} + o_p(1),$$

using the fact that g and g' are bounded and applying (2.1), (2.5) and (4.4). Then, note that

$$\Big| n^{-1/2} \sum_{j=1}^{n} v'\Gamma^{-1} x_j \Big\{ \int_{[n\alpha]/n}^{[n\beta]/n} \eta_n(u)\, dG\{(Q_n(u) - e_j)/\gamma\} - n^{-1} \sum_{i=[n\alpha]+1}^{[n\beta]} \eta_n(i/n) \frac{g\{(Q_n(i/n) - e_j)/\gamma\}}{\gamma f_n(Q_n(i/n))} \Big\} \Big|$$

$$= \Big| n^{-1/2} \sum_{j=1}^{n} v'\Gamma^{-1} x_j \sum_{i=[n\alpha]+1}^{[n\beta]} \Big\{ \int_{(i-1)/n}^{i/n} \eta_n(u)\, dG\{(Q_n(u) - e_j)/\gamma\} - n^{-1} \eta_n(i/n) \frac{g\{(Q_n(i/n) - e_j)/\gamma\}}{\gamma f_n(Q_n(i/n))} \Big\} \Big|$$

$$\le n^{-1/2} \sum_{j=1}^{n} |v'\Gamma^{-1} x_j| n^{-1} \sup_{u \in U_n(1/n)} \{ |\eta_n^{(1)}(u)/\gamma f_n(Q_n(u))| + |\eta_n(u)/\gamma^2 f_n(Q_n(u))^2| + |\eta_n(u) f_n^{(1)}(Q_n(u))/\gamma f_n(Q_n(u))^3| \}$$

$$\xrightarrow{\mathrm{P}} 0$$

by (2.2), (2.5), (2.8), (2.9) and condition (i) of the Theorem, so that using

$\sum_{j=1}^{n} v'\Gamma^{-1}x_j = 0$,

$$\begin{aligned}
T &= \widehat{I}_n^{-1} n^{-1/2} \sum_{j=1}^{n} v'\Gamma^{-1} x_j \int_{[n\alpha]/n}^{[n\beta]/n} \eta_n(u) dG\{(Q_n(u) - e_j)/\gamma\} + o_p(1) \\
&= (\widehat{I}_n^{-1} I_n) I_n^{-1} n^{-1/2} \sum_{j=1}^{n} v'\Gamma^{-1} x_j (\Lambda_{nj} - E\Lambda_{nj}) + o_p(1) \\
&= (\widehat{I}_n^{-1} I_n) T^* + o_p(1),
\end{aligned}$$

say, where $\Lambda_{nj} = \int_{[n\alpha]/n}^{[n\beta]/n} \eta_n(u) dG\{(Q_n(u) - e_j)/\gamma\}$, $1 \le j \le n$, $n \ge 1$, and $T^* = I_n^{-1} n^{-1/2} \sum_{j=1}^{n} v'\Gamma^{-1} x_j (\Lambda_{nj} - E\Lambda_{nj})$. Clearly, T^* has mean zero, and variance

$$\begin{aligned}
\text{Var}(T^*) &= I_n^{-2} n^{-1} \sum_{j=1}^{n} (v'\Gamma^{-1} x_j)^2 \{E\Lambda_{n1}^2 - (E\Lambda_{n1})^2\} \\
&\le I_n^{-2} n^{-1} \sum_{j=1}^{n} (v'\Gamma^{-1} x_j)^2 E\Lambda_{n1}^2 \\
&\le I_n^{-2} n^{-1} \sum_{j=1}^{n} (v'\Gamma^{-1} x_j)^2 \int_{[n\alpha]/n}^{[n\beta]/n} \eta_n(u)^2 du \\
&= I_n^{-1} n^{-1} \sum_{j=1}^{n} (v'\Gamma^{-1} x_j)^2
\end{aligned}$$

so that

$$\limsup \text{Var } T^* \le I^{-1} v'\Gamma^{-1} v$$

by Theorem 4.1 of Stone (1975). Thus, if $I = \infty$, $T^* \xrightarrow{P} 0$ and the result obtains. Finally, if $I < \infty$, a straightforward computation using condition (ii) of the Theorem and the bound (2.8) shows that $E|T^*|^3 \to 0$ so the result follows from Lyapounov's Central Limit Theorem and the argument at the end of the proof of Theorem 2.1 of Koul and Susarla (1983).

References

Bassett, G. and Koenker, R. (1978). Asymptotic theory of least absolute error regression. *J. Amer. Statist. Assoc.* **73**, 618–622.

Beran, R. (1974). Asymptotically efficient adaptive rank estimates in location models. *Ann. Statist.* **2**, 63–74.

Begun, J. M., Hall, W. J., Huang, W. M. and Wellner, J. A. (1983). Information and asymptotic efficiency in parametric-nonparametric models. *Ann. Statist.* **11**, 432–452.

Bickel, P. J. (1982). On adaptive estimation. *Ann. Statist.* **10**, 647–671.

Billingsley, P. (1968). *Convergence of Probability Measures.* John Wiley, New York.

Dionne, L. (1981). Efficient nonparametric estimators of parameters in the general linear hypothesis. *Ann. Statist.* **9**, 457–460.

Epanechnikov, V. A. (1969). Non-parametric estimation of a multivariate probability density. *Theor. Probab. Appl.* **14**, 153–158.

Hsieh, D. A. and Manski, C. F. (1987). Monte Carlo evidence on adaptive maximum likelihood estimation of a regression. *Ann. Statist.* **15**, 541–551.

Koul, H. L. and Susarla, V. (1983). Adaptive estimation in linear regression. *Statistics and Decisions* **1**, 379–400.

Newey, W. (1988). Adaptive estimation of regression models via moment restrictions. *J. Econmtx.* **9**, 301–339.

Portnoy, S. and Koenker, R. (1989). Adaptive L-estimation for linear models. *Ann. Statist.* **17**, 362–381.

Royden, H. L. (1968). *Real Analysis*, 2nd edition. Macmillan, New York.

Sacks, J. (1975). An asymptotically efficient sequence of estimates of a location parameter. *Ann. Statist.* **3**, 285–298.

Schick, A. (1987). A note on the construction of asymptotically linear estimators. *J. Statist. Plann. Inferences* **16**, 89–105.

Stein, C. (1956). Efficient nonparametric testing and estimation. *Proc. 3rd Berkeley Symp. Math. Statist. Prob.* **1**, 187–196.

Stone, C. (1975). Adaptive maximum likelihood estimators of a location parameter. *Ann. Statist.* **3**, 267–284.

Welsh, A. H. (1987a). The trimmed mean in the linear model (with discussion). *Ann. Statist.* **15**, 20–36. Correction ibid (1988) **16**, 480.

Welsh, A. H. (1987b). One-step L-estimators for the linear model. *Ann. Statist.* **15**, 626–641. Correction ibid (1988) **16**, 481.

Welsh, A. H. (1987c). Kernel estimates of the sparsity function. *Statistical Data Analysis Based on the L_1-norm and Related Methods* (Edited by Y. Dodge), 369–378. Elsevier Science Publishers B. V., Amsterdam.

Department of Statistics, Australian National University, Australia.

(Received June 1989; accepted June 1990)