The Annals of Statistics
1977, Vol. 5, No. 4, 681–699

BEHAVIOR OF ROBUST ESTIMATORS IN THE REGRESSION MODEL WITH DEPENDENT ERRORS

BY HIRA L. KOUL

Michigan State University

This paper proves the asymptotic linearity in the regression parameter of a class of linear rank statistics when errors in the regression model are strictly stationary and strongly mixing. Besides this, several other weak convergence results are proved which yield the asymptotic normality of L and M estimators of the regression parameter under the above dependent structure. All these results are useful in studying the effect of the above dependence on the asymptotic behavior of R, M and L estimators vis-à-vis the least squares estimator. An example of linear model with Gaussian errors is given where it is shown that the asymptotic efficiency of certain classes of R, M and L estimators relative to the least squared estimator is greater than or equal to its value under the usual independent errors model.

0. Introduction. For each $n \geqq 1$, let $\{Z_{ni}, 1 \leqq i \leqq n\}$ be a sequence of random variables such that

$$Z_{ni} = \Delta x_{ni} + \varepsilon_{ni}, \qquad 1 \leqq i \leqq n \tag{0.1}$$

where $\{x_{ni}\}$ are some real numbers, Δ is the parameter of interest and where, for each $n \geqq 1$, $\{\varepsilon_{ni}, 1 \leqq i \leqq n\}$ is a sequence of strictly stationary strongly mixing (s.s.s.m.) random variables with mixing number α_n and has continuous joint distribution with a continuous common marginal cdf F. (Refer to Phillip [1967, Definition III] for the definition of strongly mixing triangular arrays.)

Here we prove the asymptotic linearity of $S(\Delta)$ (see (2.1) below) in Δ and the asymptotic normality of R, M and L estimators of Δ under the model (0.1). When errors are i.i.d. F, the asymptotic linearity of $S(\Delta)$ has been proved by Koul (1969, 1970) and Jurečková (1969). Theorem 2.1 below therefore extends these results to the above dependent model for bounded scores. A consequence of this result is the asymptotic normality of R-estimators.

Section 1 contains some basic results about certain basic processes which in turn are derived from the results in the appendix about weighted empirical processes. The results of Section 1 automatically yield the asymptotic normality of Huber's (1973) M estimators and Bickel's (1973) L estimator under the above model. This is shown in Section 3. It is noted that all these three estimators continue to have asymptotically normal distributions with the right means and variances that reflect the dependent structure of errors. Thus one needs to compare only the asymptotic variances in order to see how asymptotically robust

Received August 1974; revised November 1976.
AMS 1970 *subject classifications*. Primary 62G05; Secondary 62G35.
Key words and phrases. Regression parameter, R, M, L estimators, strongly mixing, robustness.

the above estimators of Δ are when the assumption of independent errors is weakened to the strongly mixing errors.

Recently in (1975b) Gastwirth and Rubin (G–R) have studied the effect of Δ-mixing, a notion which they introduced in (1975a), on various specified members of R, M and L estimators for location parameter. Our results here could be considered as a generalization of their theoretical results to include more estimators, to regression model and to strongly mixing dependence. Their Δ-mixing class is smaller than strongly mixing class.

In Section 4 we consider an application of our results of Sections 2 and 3 to the model $Z_i = i\Delta + \varepsilon_i$, $\{\varepsilon_i\}$ s.s.s.m. Gaussian, $E(\varepsilon_i) = 0$, $\mathrm{Var}(\varepsilon_i) = 1$ and $\mathrm{Cov}(\varepsilon_i, \varepsilon_{i+j}) = \rho^j$ for all i, j. It is concluded that the asymptotic efficiency of a class of R estimators (see Section 2 for the class of R-estimators) relative to the least square estimator, when computed under the above model, is greater than or equal to its value when computed under the model $Z_i = i\Delta + \varepsilon_i$, $\{\varepsilon_i\}$ i.i.d. $N(0, 1)$ provided $0 < \rho < 1$. Similar conclusions hold for M and L estimators of Section 3. These results are similar to the one given by G–R (1975b) for the case of L estimators of location parameter.

1. Assumptions and preliminaries. To allow us some flexibility for applications we give ourselves triangular arrays of real constants $\{c_{ni}, 1 \leqq i \leqq n\}$ and $\{d_{ni}, 1 \leqq i \leqq n\}$ and introduce the following basic processes. For $0 \leqq t \leqq 1$ and $|\Delta| \leqq b$, $0 < b < \infty$ fixed, define

$$(1.1)\qquad \begin{aligned} V_d(t, \Delta) &= \textstyle\sum_i d_i[I(Z_i \leqq H^{-1}(t) + c_i\Delta) - L_i(t, \Delta)]\,, \\ U_d(t, \Delta) &= \textstyle\sum_i d_i[I(Z_i \leqq F^{-1}(t) + c_i\Delta) - F(F^{-1}(t) + \Delta c_i)] \end{aligned}$$

where

$$(1.2)\qquad \begin{aligned} L_i(t, \Delta) &= F(H_\Delta^{-1}(t) + \Delta c_i)\,; \\ H_\Delta(y) &= n^{-1} \textstyle\sum_i F(y + \Delta c_i)\,, \qquad 1 \leqq i \leqq n\,. \end{aligned}$$

In this paper i in $\sum_i$ and $\max_i$ varies from 1 to n, y is a real number and the subscript n in triangular arrays and other entities is not exhibited for the sake of convenience.

About $\{c_i\}$ and $\{d_i\}$ we assume that

$$(1.3)\qquad \limsup_{n\to\infty} n \max_i d_i^2 < \infty\,, \qquad \tau_d^2 \equiv \textstyle\sum d_i^2 = 1$$

$$(1.4a)\qquad \max_i c_i^2 \to 0$$

$$(1.4b)\qquad \limsup_{n\to\infty} \textstyle\sum_i c_i^2 \leqq K < \infty\,.$$

About $\{\alpha_n\}$ we assume that for some $1 < q < 2$

$$(1.5)\qquad \limsup_{n\to\infty} \textstyle\sum_{j=0}^{n-1} (j+1)^2 \alpha_n^{1-1/q}(j) < \infty\,.$$

We begin by stating and proving

LEMMA 1.1. *Let* $\{Z_i\}$ *be* s.s.s.m. *with mixing numbers* $\{\alpha_n\}$ *and continuous marginal* cdf F. *Assume* $\{d_i\}$, $\{c_i\}$ *and* $\{\alpha_n\}$ *satisfy* (1.3), (1.4a) *and* (1.5) *respectively. Then* *for each fixed* $|\Delta| \leqq b$

$$\sup_t |V_d(t, \Delta) - U_d(t, \Delta)| = o_p(1)\,, \tag{1.6}$$

$$\sup_t |U_d(t, \Delta) - U_d(t, 0)| = o_p(1)\,. \tag{1.7}$$

Hence

$$\sup_t |V_d(t, \Delta) - V_d(t, 0)| = o_p(1)\,. \tag{1.8}$$

PROOF. For the sake of convenience write U, V for U_d, V_d. Note that $U(t, \Delta) = V(H_\Delta(F^{-1}(t)), \Delta)$. But F continuous and $|\Delta| \leqq b$ implies that

$$\sup_{t,\Delta} |H_\Delta(F^{-1}(t)) - t| = \sup_{y,\Delta} |H_\Delta(y) - F(y)| = o(1)\,. \tag{1.9}$$

Hence (1.6) follows from (1.9) and (16) of the appendix. Consequently for every $\varepsilon > 0$ and for Δ fixed

$$\lim_{\delta\to 0} \limsup_{n\to\infty} P(\sup_{|t-s|\leqq\delta} |U(t, \Delta) - U(s, \Delta)| > \varepsilon) = 0\,. \tag{1.10}$$

To prove (1.7) it is enough to prove that for each fixed Δ

(a) $|U(t, \Delta) - U(t, 0)| = o_p(1)$ for each fixed $0 \leqq t \leqq 1$ and
(b) for every $\varepsilon > 0$

$$\lim_{\delta\to\infty} \limsup_{n\to\infty} P(\sup_{|t-s|<\delta} |U(t, \Delta) - U(t, 0) - U(s, \Delta) + U(s, 0)| > \varepsilon) = 0\,.$$

But (b) follows from (1.10) when applied twice, once with fixed Δ and once with $\Delta = 0$. To prove (a) write

$$U(t, \Delta) - U(t, 0) = \textstyle\sum_i d_i \xi_i$$

where

$$\xi_i = [I(Z_i \leqq F^{-1}(t) + \Delta c_i) - I(Z_i \leqq F^{-1}(t)) - F(F^{-1}(t) + \Delta c_i) + t]\,, \quad 1 \leqq i \leqq n\,.$$

Now note that $\max_i |\xi_i| \leqq 1$ and

$$E|\xi_i|^2 \leqq |F(F^{-1}(t) + \Delta c_i) - t|\,. \tag{1.11}$$

Using the continuity of F and (1.4a) we have $\max_i ||\xi_i||_2 \to 0$. Also note that (1.5) implies (19) of Lemma A5 in the appendix. Thus Lemma A5 is applicable with the above $\{\xi_i\}$ and we have

$$\operatorname{Var}(U(t, \Delta) - U(t, 0)) \to 0$$

which entails (a) above. Hence (1.7) is proved. □

The following lemma is useful in studying the properties of Bickel's L-estimators and Huber's M-estimators.

LEMMA 1.2. *In addition to the assumptions of Lemma* 1.1 *assume that F has a bounded density f and* (1.4b) *is satisfied. Then*

$$\sup_{\Delta,t} |U_d(t, \Delta) - U_d(t, 0)| \to 0 \quad \text{in probability as} \quad n \to \infty\,. \tag{1.12}$$

PROOF. Write $J(t, \Delta) = \sum d_i I(Z_i \leqq F^{-1}(t) + \Delta c_i)$ and $\bar{J}(t, \Delta) = EJ(t, \Delta) = \sum_i d_i F(F^{-1}(t) + \Delta c_i)$. Fix an $\varepsilon > 0$. Choose a partition $-b = \Delta_0 < \Delta_1 < \cdots < \Delta_r = b$ of $[-b, b]$ such that

$$\max_{1 \leqq j \leqq r} (\Delta_j - \Delta_{j-1}) \|f\|_\infty \leqq \varepsilon / 2K^{-\frac{1}{2}} \tag{1.13}$$

where K is as in (1.4). The proof of (1.12) will be divided into four cases;

(i) $c_i \geqq 0$, $d_i \geqq 0$, $1 \leqq i \leqq n$,
(ii) $c_i \geqq 0$, $d_i \leqq 0$, $1 \leqq i \leqq n$,
(iii) $c_i \leqq 0$, $d_i \geqq 0$, $1 \leqq i \leqq n$,
(iv) $c_i \leqq 0$, $d_i \leqq 0$, $1 \leqq i \leqq n$.

The details will be given only for case (i), the other cases being handled similarly. Once the proof is given for the above cases, the proof of (1.12) for general $\{d_i\}$ and $\{c_i\}$ will follow from these by decomposing $U_d(t, \Delta)$ and $U_d(t, 0)$ into the above four cases. We proceed to prove (1.12) under case (i).

Assume $c_i \geqq 0$, $d_i \geqq 0$, $1 \leqq i \leqq n$. Suppose $\Delta_{j-1} \leqq \Delta \leqq \Delta_j$ for some $1 \leqq j \leqq r$. Then for such a Δ we have, writing U for U_d,

$$\begin{aligned} &U(t, \Delta_{j-1}) - U(t, 0) + \bar{J}(t, \Delta_{j-1}) - \bar{J}(t, \Delta_j) \\ &\quad \leqq U(t, \Delta) - U(t, 0) \\ &\quad \leqq U(t, \Delta_j) - U(t, 0) + \bar{J}(t, \Delta_j) - \bar{J}(t, \Delta_{j-1}) \quad 0 \leqq t \leqq 1 . \end{aligned} \tag{1.14}$$

Now since F has a bounded density f and $\max_i |c_i| \to 0$ we have, by the mean value theorem, Cauchy–Schwarz inequality and the fact that $\sum d_i^2 = 1$, that

$$\begin{aligned} &|\bar{J}(t, \Delta_j) - \bar{J}(t, \Delta_{j-1})| \\ &\quad \leqq (\Delta_j - \Delta_{j-1}) (\textstyle\sum c_i^2)^{\frac{1}{2}} \|f\|_\infty \\ &\quad \leqq \varepsilon/2 , \qquad 0 \leqq t \leqq 1 ; \quad 1 \leqq j \leqq r ; \quad \text{by (1.4b) and (1.13).} \end{aligned} \tag{1.15}$$

Hence from (1.14) and (1.15) we have

$$\sup_{\Delta, t} |U(t, \Delta) - U(t, 0)| \leqq \max_{1 \leqq j \leqq r} \sup_t |U(t, \Delta_j) - U(t, 0)| + \varepsilon/2 . \tag{1.16}$$

Therefore (1.12) follows in this case from (1.16) and (1.7). □

2. Linearity of $S_d(\Delta)$.

Define

$$S_d(\Delta) = \textstyle\sum_i d_i \varphi(R_{i\Delta}/(n + 1)) \tag{2.1}$$

where $R_{i\Delta}$ is the rank of $Z_i - \Delta c_i$ among $\{Z_j - \Delta c_j, 1 \leqq j \leqq n\}$, and φ is a nondecreasing bounded and right continuous function on $(0, 1)$.

Here we state and prove the following

THEOREM 2.1. *For each $n \geqq 1$, let $\{Z_i, 1 \leqq i \leqq n\}$ be* s.s.s.m. *with mixing number α satisfying* (1.5). *Also assume that the joint distribution of $\{Z_i, 1 \leqq i \leqq n\}$ is continuous for each $n \geqq 1$ with marginal F which has an absolutely continuous density f such that*

$$0 < I(f) = \int (f'/f)^2 \, dF < \infty . \tag{2.2}$$

Furthermore let $\{c_i\}$ *and* $\{d_i\}$ *satisfy* (1.3), (1.4) *and*

$$(c_i - c_j)(d_i - d_j) \geqq 0 \qquad (2.3)$$
$$(\text{or } (c_i - c_j)(d_i - d_j) \leqq 0)\,, \qquad 1 \leqq i \leqq j \leqq n\,.$$

Then

$$\sup_\Delta |S_d(\Delta) - S_d(0) + \Delta \textstyle\sum_i c_i d_i b(\varphi, f)| = o_p(1) \qquad \text{as} \quad n \to \infty\,, \qquad (2.4)$$

where $b(\varphi, t) = -\int_0^1 \varphi(t)[f'(F^{-1}(t))/f(F^{-1}(t))]\,dt$.

Proof. Introduce, for $0 \leqq t \leqq 1$,

$$S_d(t, \Delta) = \textstyle\sum_i d_i I(R_{i\Delta} \leqq tn) \qquad (2.5)$$
$$T_d(t, \Delta) = S_d(t, \Delta) - \mu_d(t, \Delta)\,; \qquad \mu_d(t, \Delta) = \textstyle\sum_i d_i L_i(t, \Delta)$$

where from (1.3) it should be recalled that $\sum d_i^2 = 1$. Note that $S_d(\Delta) = \int \varphi(nt/(n+1))\,dS_d(t, \Delta)$. We split the proof in several lemmas. For all the following lemmas the assumptions of Theorem 2.1 are holding. Thus the statements will include only the results and no assumptions. The proofs should make clear as to which assumptions are being used.

Lemma 2.1. *For any fixed* $|\Delta| \leqq b$ *we have*

$$\sup_t |T_d(t, \Delta) - T_d(t, 0)| = o_p(1)\,. \qquad (2.6)$$

Proof. Writing T, V for T_d, V_d, etc. we have

$$T(t, \Delta) = V(H_{n\Delta}(H_\Delta^{-1}(t)), \Delta) + \textstyle\sum_i d_i[L_i(H_{n\Delta}(H_\Delta^{-1}(t)), \Delta) - L_i(t, \Delta)] \qquad (2.7)$$

where $H_{n\Delta}$ is the empirical cumulative of $\{Z_i - \Delta c_i, 1 \leqq i \leqq n\}$.

Note that (2.2) implies f is bounded and that for any $-\infty < y, z < +\infty$

$$|f(y) - f(z)|^2 \leqq |y - z| I(f) \|f\|_\infty\,. \qquad (2.8)$$

Therefore we have that (2.2) implies f is uniformly continuous. This in turn implies that

$$\max_i \sup_t l_i(t, \Delta) \to 1 \qquad \text{as} \quad n \to \infty \qquad (2.9)$$

where

$$l_i(t, \Delta) = \frac{\partial}{\partial t} L_i(t, \Delta) \equiv f(H_\Delta^{-1}(t) + \Delta x_i)/n^{-1} \textstyle\sum_i f(H_\Delta^{-1}(t) + \Delta x_i)\,, \qquad 1 \leqq i \leqq n\,.$$

Then $\forall\, 0 < K < \infty$

$$\max_i \sup_{|t-s| \leqq kn^{-\frac{1}{2}}} n^{\frac{1}{2}} |L_i(t, \Delta) - L_i(s, \Delta) - (t - s) \cdot 1| \equiv o(1) \qquad \text{as} \quad n \to \infty\,. \qquad (2.10)$$

Writing V_1 for V_d with $d_i \equiv n^{-\frac{1}{2}}$ and using (16) of the appendix with this modification we conclude that the random variables $\{\sup_t |V_1(t, \Delta)|\}$ are bounded in probability. This follows from Theorems 15.5 and 15.3 (i) of Billingsley (1968). Using this fact and (16) of the appendix one can conclude that

$$\sup_t |n^{\frac{1}{2}}[H_\Delta(H_{n\Delta}^{-1}(t)) - t] + V_1(t, \Delta)| = o_p(1)\,. \qquad (2.11)$$

A proof of a statement like (2.11) is given in Koul–Staudte (1972) in connection with the independent observations. It is easy to see that the same proof carries over to prove (2.11). From here on one uses an argument similar to one used in Koul–Staudte (1972) (which requires applying (2.9), (2.10), (2.11) and the fact that $\sup_t |V_1(t, \Delta)| = O_p(1)$ to (2.7)) to conclude that

$$\sup_t |T_d(t, \Delta) - K_d(t, \Delta)| = o_p(1) \qquad \text{for each} \quad |\Delta| \leqq b \tag{2.12}$$

where

$$K_d(t, \Delta) \equiv V_d(t, \Delta) - n^{\frac{1}{2}} \bar{d} V_1(t, \Delta) .$$

Hence one has (2.6) holding in view of (1.3) and (1.8) once applied to V_d and once to V_1. □

LEMMA 2.2. *Assume* (2.2) *holds. Then* (1.4) *and* (1.3) *imply that for each fixed* $|\Delta| < b$

$$\sup_t |\mu_d(t, \Delta) - \mu_d(t, 0) - \Delta(\textstyle\sum_i c_i d_i) b(t, f)| \to 0 \qquad as \quad n \to \infty \tag{2.13}$$

with

$$b(t, f) = f(F^{-1}(t)) , \qquad 0 \leqq t \leqq 1 .$$

PROOF. Follows easily using uniform continuity of $f(F^{-1})$. □

LEMMA 2.3.

$$\sup_{\Delta, t} |S_d(t, \Delta) - S_d(t, 0) - \Delta b_n(t, f)| \to_p 0 \qquad as \quad n \to \infty \tag{2.14}$$

with

$$b_n(t, f) = (\textstyle\sum_i c_i d_i) f(F^{-1}(t)) .$$

PROOF. From the above Lemmas 2.1, 2.2 we have that for each fixed Δ,

$$\sup |S_d(t, \Delta) - S_d(t, 0) - \Delta b_n(t, f)| \to_p 0 \qquad \text{as} \quad n \to \infty . \tag{2.15}$$

Theorem 2.1 of Jurečková (1969) thows that (2.3) implies that $S_d(t, \Delta)$ is a monotone step function of Δ for each fixed t; this is just a functional statement and not a distributional statement. The remainder of the proof is just analogous to that of Lemma 1.2.

PROOF OF THEOREM 2.1. Writing

$$S_d(\Delta) = \int \varphi(t)\, dS_d(t_n, \Delta) ; \qquad t_n = (n+1)n^{-1}t ,$$

and integrating by parts one has

$$S_d(\Delta) - S_d(0) = -\int [S_d(t_n, \Delta) - S_d(t_n, 0)]\, d\varphi(t) .$$

Also (2.2) implies that $f(x) \to 0$ as $x \to \pm\infty$ and hence

$$\begin{aligned} \int f(F^{-1}(t))\, d\varphi(t) &= -\int \varphi(t)[f'(F^{-1}(t))/f(F^{-1}(t))]\, dt \\ &= b(\varphi, f) . \end{aligned}$$

Therefore

$$S_d(\Delta) - S_d(0) + \Delta \sum_i c_i d_i b(\varphi, f) \\ = -\int [S_d(t_n, \Delta) - S_d(t_n, 0) - \Delta \sum_i d_i c_i f(F^{-1}(t_n))] \, d\varphi(t)$$

which completes the proof in view of (2.14) and the fact that $\sup_t |t_n - t| \to 0$.

Application to R-estimator. Now assume the model (0.1) holds. Consider the R-estimator $\hat{\Delta}_R$ of Δ defined as the solution of the equation $S_x(\Delta) = 0$ where $S_x(\Delta) = \sum_i (x_i - \bar{x})\varphi(R_{i\Delta}/(n+1))$ with $R_{i\Delta}$ now as the rank of $Z_i - \Delta x_i$. We have

COROLLARY 2.1. *Assume*

$$\limsup_{n\to\infty} n^{\frac{1}{2}} \max_i |x_i|/\sigma_x < \infty \qquad \textit{where} \quad \sigma_x^2 = \sum_i (x_i - \bar{x})^2 . \tag{2.17}$$

Let $\{\varepsilon_i, 1 \leqq i \leqq n\}$ *and* $\{Z_i, 1 \leqq i \leqq n\}$ *be as in model* (0.1). *Let* F *satisfy the conditions of Theorem* 2.1. *Then*

$$\sigma_x|(\hat{\Delta}_R - \Delta_0) - \hat{S}_x(\Delta_0)\sigma_x^{-2}b^{-1}(\varphi, f)| \to_p 0 \qquad \textit{as} \quad n \to \infty \tag{2.18}$$

where probability is computed under model (0.1) *with* Δ_0 *being the true value of* Δ *and where* $\hat{S}_x(\Delta_0) = \sum_i (x_i - \bar{x})\varphi(F(Z_i - \Delta_0 x_i \sigma_x^{-1}))$.

PROOF. Since $\hat{\Delta}_R$ is translation invariant, assume without loss of generality that $\Delta_0 = 0$. Then $\{Z_i\}$ of model (0.1) satisfies the conditions of Theorem 2.1.

Put in Theorem 2.1, $d_i \equiv (x_i - \bar{x})/\sigma_x$ and $c_i \equiv x_i/\sigma_x$. Now (2.17) obviously implies that (1.3) and (1.4) are satisfied for these $\{d_i\}$ and $\{c_i\}$. Hence Theorem 2.1 is applicable and one concludes in a straightforward fashion that

$$\sigma_x|\hat{\Delta}_R - \sigma_x^{-2}S_x(0)b^{-1}| \to_p 0 \qquad \text{as} \quad n \to \infty . \tag{2.19}$$

Now use (2.12) above with $\Delta = 0$, the above $\{d_i\}$ and the fact that φ is bounded nondecreasing to conclude that

$$\sigma_x^{-1}|S_x(0) - \hat{S}_x(0)| \to_p 0 \qquad \text{as} \quad n \to \infty ; \tag{2.20}$$

this is proven by writing

$$S_x(0) = \int \varphi(t) \, dT_d(t_n, 0) , \quad t_n = (n+1)^{-1}nt ; \qquad d_i = (x_i - \bar{x})/\sigma_x .$$

On combining (2.20) with (2.21) one gets (2.18). □

THEOREM 2.2. *Let* $\{\varepsilon_i, 1 \leqq i \leqq n\}$ *be* s.s.s.m. *with mixing numbers* $\{\alpha_n\}$ *satisfying* (1.5). *Furthermore assume that for each* $n \geqq 1$, $\{\varepsilon_i, 1 \leqq i \leqq n\}$ *have continuous joint distribution with the absolutely continuous marginal* F *having an absolutely continuous density* f *satisfying* (2.2). *Assume* $\{x_i\}$ *satisfy* (2.17). *Finally assume that*

$$\liminf_{n\to\infty} \sigma_x^{-2}\tau_n^2 > 0 \tag{2.21}$$

where

$$\tau_n^2 = \operatorname{Var}_0 \left(\sum_{i=1}^n (x_i - \bar{x})\varphi(F(Z_i))\right) .$$

Then

$$\mathscr{L}_\Delta(\sigma_x^2\tau_n^{-1}(\hat{\Delta}_R - \Delta)) \to N(0, b^{-2}) \tag{2.22}$$

with $b \equiv b(\varphi, f)$.

PROOF. First of all the parameter Δ may be assumed to be 0. By (2.21) and Corollary 2.1 above we have

$$\lim_{n\to\infty} \mathscr{L}_0(\sigma_x^2\tau_n^{-1}\hat{\Delta}_R) = \lim_{n\to\infty} \mathscr{L}_0(\tau_n^{-1}b^{-1}\hat{S}_x(0)) . \tag{2.23}$$

The proof of the theorem is now completed by applying Lemma A6 to the random variables $\xi_{nj} = \varphi(F(Z_j))$ with $d_{nj} = (x_j - \bar{x})$, $1 \leqq j \leqq n$. Observe that (1.5), (2.17) and (2.21) above imply the conditions of the Lemma A6. □

3a. Huber's *M*-estimator. Recall from Huber (1973) that the *M*-estimator $\hat{\Delta}_M$ of Δ is defined as the solution of the equation

$$h(\Delta) \equiv \textstyle\sum_i x_i \psi(Z_i - \Delta x_i) = 0 \tag{3.1}$$

where

(3.1′) ψ is a right continuous $\nearrow$ bounded function on $(-\infty, +\infty) \ni \int_{-\infty}^{+\infty} \psi(y)\, dF(y) = 0$.

Define, for $-\infty < y < +\infty$,

$$G(y, \Delta) = \textstyle\sum_i x_i I(Z_i \leqq y + \Delta x_i) . \tag{3.2}$$

Then $\hat{\Delta}_M$ is the solution of

$$\int \psi(y)\, dG(y, \Delta) = 0 .$$

PROPOSITION 3.1. *Let* $\{Z_i, 1 \leqq i \leqq n\}$ *and* $\{\varepsilon_i, 1 \leqq i \leqq n\}$ *be as in model* (0.1) *and* $\{\alpha_n\}$ *satisfy* (1.5). *Assume*

$$\limsup_{n\to\infty} n \max x_i^2/\tau_x^2 < \infty ; \qquad \tau_x^2 = \textstyle\sum_i x_i^2 . \tag{3.3}$$

Let ψ *be as in* (3.1′) *and F satisfy* (2.2). *Then, if* 0 *is the true parameter*

$$\tau_x^{-1} \sup_\Delta |h(\Delta\tau_x^{-1}) - h(0) + \Delta\tau_x\, d(\psi, f)| \to_p 0 . \tag{3.4}$$

Hence, if Δ_0 *is the true parameter,*

$$\tau_x|(\hat{\Delta}_M - \Delta_0) - h(\Delta_0\tau_x^{-1})\, d^{-1}(\psi, f)\tau_x^{-2}| \to_p 0 \tag{3.5}$$
$$d(\psi, f) = -\int_{-\infty}^{\infty} \psi(y) f'(y)\, dy .$$

Moreover if

$$\liminf_{n\to\infty} \tau_x^{-2}\tau_n^2 > 0 \tag{3.6}$$

where now $\tau_n^2 = \mathrm{Var}_0\,(\sum_i x_i \psi(Z_i))$, *then*

$$\mathscr{L}_{\Delta_0}(\tau_x^2\tau_n^{-1}(\hat{\Delta}_M - \Delta_0)) \to N(0, d^{-2}(\psi, f)) . \tag{3.7}$$

PROOF. Let

$$W(y, \Delta) = \tau_x^{-1}[G(y, \Delta\tau_x^{-1}) - \nu(y, \Delta\tau_x^{-1})]$$

with

$$\nu(y, \Delta) = \sum_i x_i F(y + x_i \Delta) .$$

Put $c_i \equiv x_i/\tau_x$ and $d_i \equiv x_i/\tau_x$ in the definition of U as given in (1.1). Then

$$W(y, \Delta) = U_x(F(y), \Delta) , \qquad -\infty < y < \infty , \quad \forall |\Delta| \leqq b . \tag{3.8}$$

Note that (3.3) implies that (1.3) and (1.4) are satisfied for the above $\{c_i\}$ and $\{d_i\}$. Furthermore (2.2) implies F satisfies the assumptions of Lemma 1.5. Also if 0 is the true parameter then model (0.1) implies that $\{Z_i\}$ are s.s.s.m. Hence applying Lemma 1.2 one has, using uniform continuity of F,

$$\sup_{\Delta, y} |W(y, \Delta) - W(y, 0)| \to_p 0 \qquad \text{as} \quad n \to \infty . \tag{3.9}$$

But one also has

$$\sup_{y, \Delta} \tau_x^{-1} |\sum_i x_i [F(y + \Delta x_i \tau_x^{-1}) - F(y) - \Delta x_i \tau_x^{-1} f(y)]| \to 0 \tag{3.10}$$

because of $\max_i x_i^2/\tau_x^2 \to 0$.

Combining (3.9) with (3.10) one has that if 0 is the true parameter, then

$$\sup_{\Delta, y} \tau_x^{-1} |G(y, \Delta\tau_x^{-1}) - G(y, 0) - \Delta\tau_x f(y)| \to_p 0 \qquad \text{as} \quad n \to \infty . \tag{3.11}$$

Now since ψ is bounded and $\nearrow$ use (3.11) to conclude (3.4) in the same fashion as (2.13) is used to conclude (2.4). Thus we conclude the proof of (3.4).

Again $\hat{\Delta}_M$ being translation invariant, one may assume $\Delta_0 = 0$. Then (3.5) follows from (3.4) in the same fashion as does (2.18) from (2.4).

Finally (3.7) follows from (3.5) and Lemma A6 applied to the random variables $\xi_{nj} = \psi(Z_j)$, $d_{nj} = x_j$; $1 \leqq j \leqq n$. Boundedness of ψ implies that $\{\xi_{nj}\}$ are uniformly bounded. □

REMARK. There are two ways Proposition 3.1 differs from some of the results of Huber (1973). The first being that we have dependent data but the second difference is that we allow ψ which need not have derivatives. Of course this is balanced by having F satisfy stronger conditions. If indeed ψ did have a derivative then

$$\begin{aligned} d(\psi, f) &= -\int \psi(y) f'(y)\, dy = \int f(y)\psi'(y)\, dy \\ &= \int \psi'\, dF \end{aligned}$$

which agrees with the constant given in Huber.

3b. Bickel's L-estimator. In order to define these estimators we need to decompose $x_j = x_j^+ - x_j^-$, $1 \leqq j \leqq n$ and define

$$\begin{aligned} G^+(y, \Delta) &= \sum_i x_i^+ I(Z_i \leqq y + \Delta x_i) ; \\ G^-(y, \Delta) &= \sum_i x_i^- (Z_i \leqq y + \Delta x_i) . \end{aligned} \tag{3.12}$$

Also define

$$\begin{aligned} Q^+(y, \Delta) &= (\sum x_i^+)^{-1} G^+(y, \Delta) ; \\ Q^-(y, \Delta) &= (\sum x_i^-)^{-1} G^-(y, \Delta) , \quad -\infty < y < +\infty . \end{aligned} \tag{3.13}$$

Let Δ^* be a preliminary estimator of Δ. Let Λ be the distribution function of a finite signed measure on (0, 1) with $\Lambda(1) = 1$. Then Bickel (1973) defines L-estimator of type 1 of Δ as

$$\hat{\Delta}_L = \Delta^* + \tau_x^{-2}[(\textstyle\sum_j x_j^+) \int_0^1 (Q_n^+)^{-1}(\omega, \Delta^*)\, d\Lambda(\omega) + (\sum_j x_j^-) \int_0^1 (Q_n^-)^{-1}(\omega, \Delta^*)\, d\Lambda(\omega)] \tag{3.14}$$

[(3.14) above corresponds to (2.10) of [2] with $c_j \equiv x_j$.]

Our problem here is to see under what sufficient conditions does $\hat{\Delta}_L$ continue to have the asymptotic normality when model (0.1) above is true.

To begin with we introduce the processes

$$Y_n(t) = \tau_x^{-1} \textstyle\sum_j |x_j|[I(Z_j \leqq F^{-1}(t)) - t]/q(t)\,; \quad q(t) \equiv b(t, f) \equiv f(F^{-1}(t))\,, \qquad \gamma \leqq t \leqq 1 - \gamma \tag{3.15}$$

where γ is fixed in $(0, \frac{1}{2})$.

LEMMA 3.1. *Let* $\{Z_i, 1 \leqq i \leqq n\}$ *be* s.s.s.m. *with mixing numbers* α_n *satisfying* (1.5). *Assume* F *has continuous positive density* f. *Furthermore assume that*

$$K(t, s) = \lim_{n\to\infty} EY_n(t)Y_n(s)\,, \qquad \gamma \leqq s, t \leqq 1 - \gamma \tag{3.16}$$

exists and that $\{x_j\}$ *satisfy* (3.3). *Then* $Y_n \Rightarrow Y$ *in* $D[\gamma, 1 - \gamma]$; Y *is an almost surely continuous Gaussian process with mean function* 0 *and covariance function* K *given by* (3.16).

REMARK. This is an analogue of Proposition 4.3 of [2].

PROOF. Putting $d_i \equiv |x_i|/\tau_x$ in $U_d(t, 0)$ of Section 1 above yields

$$Y_n(t) = U_d(t, 0)/q(t)\,, \qquad \gamma \leqq t \leqq 1 - \gamma\,. \tag{3.17}$$

Now because f is positive and continuous on R, we have that $1/q$, is continuous, bounded on $[\gamma, 1 - \gamma]$ and hence uniformly continuous.

Next observe that with the above $\{d_i\}$, (3.3) $\Leftrightarrow$ (1.3). Hence (1.10) is applicable and one concludes that $\forall\, \varepsilon > 0$

$$\lim_{\delta\to 0} \limsup_{n\to\infty} P(\sup_{|t-s|\leqq\delta} |Y_n(t) - Y_n(s)| > \varepsilon) = 0\,. \tag{3.18}$$

Next apply Lemma A6 with

$$\xi_{ni} = \textstyle\sum_{j=1}^k [I(Z_i \leqq F^{-1}(t_j)) - t_j]\,, \qquad d_i = x_i\,, \qquad 1 \leqq i \leqq n$$

to conclude, in view of (3.3) and (3.16), that

$$\mathscr{L}(Y_n(t_j), 1 \leqq j \leqq k) \to \mathscr{L}(Y(t_j), 1 \leqq j \leqq k) \quad \forall\, k < \infty\,.$$

This together with (3.18) concludes the lemma. □

We next state and indicate the proof of an analogue of Theorem 2.1, part I of [2].

THEOREM 3.2. *Suppose* $\{Z_i, 1 \leqq i \leqq n\}$ *are as in model* (0.1) *above. Suppose* $\{x_i\}$ *satisfy* (3.3) *and* F *has uniformly continuous positive bounded density* f. *Furthermore let* Δ^* *be translation invariant and* $\tau_x\Delta^*$ *be bounded in probability when* 0 *is* *the true parameter* $[= B$ *of* [2]]. *Also assume* Λ *to be concentrating on* $[\gamma, 1-\gamma]$ *for* $0 < \gamma < \frac{1}{2}$ *fixed such that* $\int_0^1 F^{-1}(t)\, d\Lambda(t) = 0$. *Finally let mixing numbers* $\{\alpha_n\}$ *satisfy* (1.5). *Then*

$$\begin{aligned}\lim_{n\to\infty} &\mathscr{L}_{\Delta_0}(\tau_x(\hat{\Delta}_L - \Delta_0)) \\ &= \lim_{n\to\infty} \mathscr{L}_0(-\textstyle\int Y_n(t)\, d\Lambda(t)) - \mathscr{L}(-\int Y(t)\, d\Lambda(t)) \\ &= N(0, K(\Lambda, F))\end{aligned} \tag{3.19}$$

where

$$K(\Lambda, F) = \lim_n \textstyle\iint_\gamma^{1-\gamma} K_n(t, s)\, d\Lambda(t)\, d\Lambda(s) = \iint_\gamma^{1-\gamma} K(t, s)\, d\Lambda(t)\, d\Lambda(s) \tag{3.20}$$

with $K(t, s)$ *defined by* (3.16) *above.*

PROOF. To begin with, without loss of generality, assume $\Delta_0 = 0$ so that $\{Z_i, 1 \leqq i \leqq n\}$ are s.s.s.m. Now under the above assumptions on F it may be seen that (3.11) above is still valid. But observe that (3.11) above is the same result as Lemma 4.1 of [2] after appropriate modifications. Also realize that Lemma 4.1 of [2] is basic to Propositions 4.1 and 4.2 of Bickel (1973). In other words once we have (3.11), the rest of the proof of the above theorem is similar to that of Theorem 2.1 of Bickel (1973), except where he uses his Proposition 4.3, we use our Lemma 3.1 above. All one has to do is carry out the details with $\{c_j\}$ of Bickel replaced by our $\{x_j\}$. (We remark here that in (4.23) of [2] $\sum_j c_j$ in the numerator of the right-hand side should be $\sum |c_j|$.)

As far as interchange of limit and integral signs in (3.20) is concerned one proceeds as follows. Let

$$g_s(Z_j) = [I(Z_j \leqq F^{-1}(t)) - t]\,. \tag{3.21}$$

Then, for $\gamma \leqq s \leqq t \leqq 1-\gamma$,

$$\begin{aligned}K_n(t, s) &= [q(t)q(s)]^{-1}[s(1-t) + 2\tau_x^{-2} \textstyle\sum\sum_{i<j=1}^n |x_i x_j| E g_s(Z_i) g_t(Z_j)] \\ &\leqq [q(t)q(s)]^{-1}[s(1-t) + 20n\tau_x^{-2} \max_i x_i^2 \cdot \textstyle\sum_{i=0}^{n-1} \alpha_n(i)]\end{aligned} \tag{3.22}$$

which follows from Lemma A1 applied with $q = \infty$ because our $\{g_s(Z_j)\}$ are bounded by 1.

Now apply dominated convergence theorem to conclude (3.20) which is justified in view of the assumptions (3.3), (1.5) and that f is positive on $[\gamma, 1-\gamma]$, $\Lambda[\gamma, 1-\gamma] = 1$. □

REMARK. Observe that all the above estimators continue to have asymptotically normal distributions with the right asymptotic means and variances that reflect the dependence structure of the errors. Thus in order to compare the effect of strongly mixing dependence on the asymptotic behavior of the above estimators one need only compare their asymptotic variance. In the following section we give one example of such a comparison.

4. An application. Consider the model

$$Z_i = i\Delta + \varepsilon_i\,, \qquad 1 \leqq i \leqq n \tag{4.1}$$

where $\{\varepsilon_i\}$ is a s.s.s.m. Gaussian process with a mixing number α (note that $\{\varepsilon_i\}$ and α do not depend on n) satisfying

(4.2) $$\sum_{j=0}^{\infty} (j+1)^2 \alpha^{1-1/q}(j) < \infty \qquad \text{for some} \quad 1 < q < 2$$

and where $E(\varepsilon_i) = 0$, $\text{Var}(\varepsilon_i) = 1$ for all i and $\text{Cov}(\varepsilon_i, \varepsilon_{i+j}) = \rho^j$ for all i, j. Clearly this model satisfies the conditions of model (0.1) when $x_{ni} = i$ in (0.1). [Note that in $\{\varepsilon_i]$, i may vary over all negative as well as nonnegative integers.]

Observe that the constants $x_{ni} \equiv i$ satisfy (3.3) and (2.17). Thus to apply the results of Sections 2 and 3 all we need to do is to verify (2.21) for R-estimators, (3.6) for M-estimators and (3.16) for L-estimators. We will actually derive explicit formulas for these limits in the above model (4.1).

Notice that in all three cases covariances are involved and the underlying functions φ, ψ and Λ are bounded and nondecreasing. We will be using the following lemma repeatedly.

LEMMA 4.1. *Let G be a nondecreasing bounded function on $(-\infty, +\infty)$. Let X, Y be any two random variables with a joint distribution on $R \times R$. Then*

(4.3) $$\text{Cov}(G(X), G(Y)) = \iint [P(X \leqq x, Y \leqq y) - P(X \leqq x)P(Y \leqq y)]\, dG(x)\, dG(y)\,.$$

PROOF. Write $G(X) = \int I(X \leqq x)\, dG(x) + a$ constant, do the similar thing with $G(Y)$ and apply Fubini's theorem to conclude the lemma. □

From G–R (1975a) we recall that (their Lemma 2.1)

(4.4) $$P_0[Z_1 \leqq x, Z_{j+1} \leqq y] - P_0(Z_1 \leqq x)P_0(Z_{j+1} \leqq y) = \mathbf{n}(x)\mathbf{n}(y) \sum_{k=1}^{\infty} H_{k-1}(x)H_{k-1}(y)\rho^{jk}(k!)^{-1}\,, \qquad -\infty < x, y < +\infty$$

where H_k is the kth Hermite polynomial, $\mathbf{n}$ is the unit normal density and P_0 is the probability measure given by the model (4.1) above when true $\Delta = 0$.

Now consider the case of *M-estimators*. Here, because $x_i = i$,

(4.5) $$\tau_n^2 = s_n^2\sigma_\psi^2 + 2\sum_{i=1}^{n-1}\sum_{j=1}^{n-i} i(j+i)\,\text{Cov}_0(\psi(Z_1), \psi(Z_{j+1})) = I_{1n} + 2I_{2n}\,, \quad \text{say,}$$

where $s_n^2 \equiv \tau_x^2 = \sum_{i=1}^{n} i^2$.

Now recall that ψ is nondecreasing and bounded and hence it is a priori square integrable with respect to $\mathbf{n}(x)\, dx$. Thus G–R conditions are satisfied and we have, using (4.3) with $G = \psi$ and (4.4), that

(4.6) $$\text{Cov}_0(\psi(Z_1), \psi(Z_{j+1})) = \sum_{k=1}^{\infty} \rho^{jk}(k!)^{-1}c_k$$

where

$$c_k = [\textstyle\int H_{k-1}(x)\mathbf{n}(x)\, d\psi(x)]^2\,, \qquad k \geqq 1\,.$$

Observe that

(4.7) $$\sigma_\psi^2 = \sum_{k=1}^{\infty} c_k/k! < \infty\,.$$

Let

$$B_{nk} = \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} i(j+i)\rho^{jk}, \qquad n \geqq 1, k \geqq 1. \tag{4.8}$$

Then (4.6) implies that

$$I_{2n} = \sum_{k=1}^{\infty} B_{nk} c_k (k!)^{-1}.$$

Summing of various geometric series in (4.8) yields (with $a = \rho^k$)

$$\begin{aligned} B_{nk} = a(1-a)^{-2}[n(n+1)/2 - (1+n(1-a))\{na(1-a)^{-1} \\ - a(1-a^n)(1-a)^{-2}\}] + a(1-a)^{-1}\tau_x^2, \qquad n \geqq 1, k \geqq 1. \end{aligned} \tag{4.9}$$

From (4.9) and (4.8) one has, because $|\rho| < 1$,

$$\begin{aligned} s_n^{-2}|B_{nk}| \leqq s_n^{-2}|B_{n1}| \leqq |\rho|(1-|\rho|)^{-2}[1 + 2|\rho|(1-|\rho|)^{-2}] \\ + |\rho|(1-|\rho|)^{-1}, \qquad k \geqq 1, n \geqq 1. \end{aligned} \tag{4.10}$$

From (4.9) we also get

$$\lim_{n\to\infty} s_n^{-2} B_{nk} = \rho^k(1-\rho^k)^{-1} \qquad \text{for each} \quad k \geqq 1. \tag{4.11}$$

Combining (4.5) through (4.11) we have proved that (using dominated convergence theorem)

$$\lim_{n\to\infty} s_n^{-2}\tau_n^2 = \sigma_\psi^2 + 2\sum_{k=1}^{\infty} \rho^k(1-\rho^k)^{-1}c_k(k!)^{-1}.$$

Now using expansion of $(1-\rho^k)^{-1}$ and interchanging the infinite sums one gets

$$\begin{aligned} \lim_{n\to\infty} s_n^{-2}\tau_n^2 &= \sigma_\psi^2 + 2\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} c_k(k!)^{-1}\rho^{jk} \\ &= \sum_{j=-\infty}^{\infty}\sum_{k=1}^{\infty} c_k(k!)^{-1}\rho^{|j|k}. \end{aligned} \tag{4.12}$$

Next consider *R-estimators*. Here $\sigma_x^2 \equiv \sum_{i=1}^{n}(x_i - \bar{x})^2 = n(n^2-1)/12$ and

$$\begin{aligned} \tau_n^2 = \sigma_x^2\sigma_\varphi^2 + 2\sum_{i=1}^{n-1}\sum_{j=1}^{n-i}\left(i - \left(\frac{n+1}{2}\right)\right) \\ \times \left(i + j - \frac{n+1}{2}\right)\operatorname{Cov}_0(\varphi(F(Z_1)), \varphi(F(Z_{j+1}))). \end{aligned} \tag{4.13}$$

Carrying out calculations similar to the ones that led us up to (4.12) above one gets (using Lemma 4.1 above with $G = \varphi \circ F$)

$$\begin{aligned} \lim_{n\to\infty} \sigma_x^{-2}\tau_n^2 &= \sigma_\varphi^2 + 2\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} c_k(k!)^{-1}(\rho^j)^k \\ &= \sum_{j=-\infty}^{\infty}\sum_{k=1}^{\infty} c_k(k!)^{-1}\rho^{|j|k} \quad \left(\text{again } \sigma_\varphi^2 = \sum_{k=1}^{\infty}(c_k/k!)\right) \end{aligned}$$

where now

$$c_k \equiv c_k(\varphi) = \left[\int H_{k-1}(x)\mathbf{n}(x)\, d\varphi(F(x))\right]^2, \quad (F = \Phi = N(0,1)cd). \tag{4.14}$$

Similarly for the *L-estimators* one gets

$$K(\Lambda, F) = \sum_{j=-\infty}^{\infty}\sum_{k=1}^{\infty}(c_k/k!)\rho^{|j|k} \tag{4.15}$$

where now

$$c_k = \left[\int H_{k-1}(x)\mathbf{n}(x)\, d\Lambda(F(x))\right]^2, \qquad k \geqq 1.$$

Now consider the case of

Least squares estimator. Here the least squares estimator is $\hat{\Delta}_{LS} = \sum_{i=1}^n iZ_i / \sum_i i^2$. Under the model (4.1) with the assumption (4.2) one can conclude (e.g., using Phillips (1967) results) that

$$\mathscr{L}_0(s_n \hat{\Delta}_{LS}) \to N(0, \sigma^2) \quad (s_n^2 = \textstyle\sum_i^n i^2)$$

where

$$\sigma^2 = \lim_{n\to\infty} s_n^{-2} \operatorname{Var}_0 (s_n \hat{\Delta}_{LS})$$
$$= \lim_{n\to\infty} s_n^{-2} \operatorname{Var}_0 (\textstyle\sum_{i=1}^n iZ_i) .$$

But since $\operatorname{Cov}_0 (Z_1, Z_{j+1}) = \rho^j$, we conclude using (4.9) and (4.11) with $k = 1$ that

$$\sigma^2 = 1 + 2\rho(1-\rho)^{-1} = \textstyle\sum_{j=-\infty}^{\infty} \rho^{|j|} . \tag{4.16}$$

We now state a

COROLLARY. *The asymptotic efficiency of any M estimator (within the class given by* (3.1′)) *relative to the least squares estimator under model* (4.1) *with* $0 < \rho < 1$ *is greater than or equal to its value under the model*

$$Z_i = i\Delta + \varepsilon_i , \qquad \varepsilon_i \text{ i.i.d. } N(0, 1) , \qquad i \geqq 1 . \tag{4.17}$$

PROOF. The proof follows from (4.12) and Lemma 4.1 of G–R (1975a).

Since the form of the asymptotic variances of the above R and L estimators is similar to that of the M-estimators, be it stated that the above corollary remains valid for R and L estimators also.

An example of $\{\varepsilon_i\}$ satisfying the conditions of model (4.1) would be $\{\varepsilon_i\}$ generated by the first order autoregressive Gaussian process where G–R (1975a) showed that $\alpha(j) \leqq (\frac{1}{4})(C_1\rho^{2j} + C_2|\rho|^j)$ for some constants C_1 and C_2.

APPENDIX

Before stating the main proposition of this section we recall a lemma from Deo (1973) and state it as

LEMMA A1. *Suppose for each* $n \geqq 1$, $\{\xi_{nj}, 1 \leqq j \leqq n\}$ *are strongly mixing random variables with mixing number* α_n. *Suppose X and Y are two random variables respectively measurable with respect to* $\sigma\{\xi_{n1}, \cdots, \xi_{nk}\}$ *and* $\sigma\{\xi_{nk+m}, \cdots, \xi_{nn}\}$, $1 \leqq m$, $m + k \leqq n$. *Assume p, q and r are* $\ni p^{-1} + q^{-1} + r^{-1} = 1$, *and* $||X||_p < \infty$, $||Y||_q < \infty$. *Then for each* $1 \leqq m$, $k + m \leqq n$

$$|E(XY) - E(X)E(Y)| \leqq 10 \cdot \alpha_n^{1/r}(m)||X||_p||Y||_q . \tag{1}$$

Consequently if $||X||_\infty = B < \infty$ *then for* $q > 1$ *and each* $1 \leqq m$, $k + m \leqq n$

$$|E(XY) - E(X)E(Y)| \leqq 10 \cdot \alpha_n^{1-1/q}(m)||Y||_q . \tag{2}$$

COMMENT. Deo (1973) has a proof of this lemma when $\{\xi_{ni}\}$ are s.s.s.m. The same proof goes through in view of Lemma 2.1 of Phillip, page 157 (1967) for any strongly mixing sequences $\{\xi_{ni}\}$. The stationarity assumption is not crucial to Deo's lemma.

Next we state and prove

PROPOSITION A1. *For each* $n \geqq 1$, *let* $\{Y_{nj}, 1 \leqq j \leqq n\}$ *be a sequence of strongly mixing* rv's *in* [0, 1], *with mixing number* α_n *and the marginal* cdf's $\{G_{nj}, 1 \leqq j \leqq n\}$ *satisfying*

(3) $$n^{-1} \sum_{j=1}^{n} G_{nj}(t) = t\,, \qquad 0 \leqq t \leqq 1\,.$$

Moreover assume $\{d_{ni}\}$ *satisfy* (1.3) *and* $\{\alpha_n\}$ *satisfy* (1.5) *above. Then for every* $\varepsilon > 0$

(4) $$\lim_{\delta\to 0} \limsup_{n\to\infty} P(\sup_{|t-s|\leqq \delta} |W_d(t) - W_d(s)| \geqq \varepsilon) = 0$$

where $W_d(t) = \sum_i d_{ni}[I(Y_{ni} \leqq t) - G_{ni}(t)]$, $0 \leqq t \leqq 1$. *Recall from* (1.3) *that* $\sum_i d_{ni}^2 = 1$.

PROOF. The proof consists of the following three lemmas. In each of these lemmas appropriate conditions of the proposition are in action.

LEMMA A2. *For* $1 < q < 2$

(5) $$E|W_d(t) - W_d(s)|^4 \leqq B_n\{3|t-s|^{2/q} + n^{-1}|t-s|^{1/q}\}$$

where

$$B_n = 10(4!)k_d^4 k_n^2(\alpha, q)\,; \qquad k_n(\alpha, q) = \sum_{j=0}^{n-1} (j+1)^2 \alpha_n^{1-1/q}(j)$$

and $k_d^2 = n \max_i d_i^2$.

PROOF. Without loss of generality assume $0 \leqq s \leqq t \leqq 1$ and write

(6) $$W_d(t) - W_d(s) = \sum_i d_i \xi_i$$

where

$$\xi_i = d_i I(s < Y_i \leqq t) - p_i\,, \qquad p_i = G_i(t) - G_i(s)\,, \qquad 1 \leqq i \leqq n\,.$$

One has

(7) $$E|W_d(t) - W_d(s)|^4 \leqq 4!\; k_d^4 n^{-2} \sum |E(\xi_i \xi_{i+j} \xi_{i+j+k} \xi_{i+j+k+l})|$$

where summation is over $1 \leqq i, j, k, l \leqq n$, $j + k + l \leqq n - i$. Next observe that $|\xi_i| \leqq 1$ and $||\xi_i||_q \leqq p_i^{1/q}$, $1 \leqq i \leqq n$. After applying Lemma A1 repeatedly with $B = 1$ one gets

(8) $$\begin{aligned} &|E(\xi_i \xi_{i+j} \xi_{i+j+k} \xi_{i+j+k+l})| \\ &\quad \leqq 10 \min\{\alpha^{1-1/q}(j)p_i^{1/q}, \alpha^{1-1/q}(k)p_i^{1/q} \\ &\qquad + 10[\alpha(j)\alpha(l)]^{1-1/q}(p_i p_{i+j+k})^{1/q}, \alpha^{1-1/q}(i)p_i^{1/q}\}\,. \end{aligned}$$

This yields that

(9) $$\text{LHS (7)} \leqq 10.4!k_d^4 n^{-2}\{I_1 + I_2\}$$

where

$$I_1 = 3 \sum_{i=1}^{n} \sum_{(1)} \alpha^{1-1/q}(j)p_i^{1/q}\,; \qquad \sum_{(1)} = \text{summation over } \; k, l \leqq j\,, \; k + l + j \leqq n - i$$

and

$$I_2 = 10 \sum_i \sum_k \sum_{j,l \le k} [\alpha(j)\alpha(l)]^{1-1/q}(p_i p_{i+j+k})^{1/q}$$

where the summation over k is from 1 to $n-i-j$. One has

$$(10) \qquad I_2 \leqq 10(\sum_{i=1}^n p_i^{1/q})^2(\sum_{j=0}^{n-1} \alpha^{1-1/q}(j))^2 \quad \text{and}$$
$$I_1 \leqq 3(\sum_{i=1}^n p_i^{1/q}) \sum_{j=0}^{n-1} (j+1)^2 \alpha^{1-1/q}(j) .$$

Now observe that (3) above and moment inequality implies

$$(11) \qquad n^{-1} \sum_{i=1}^n p_i^{1/q} \leqq |t-s|^{1/q} .$$

Also note that $\sum_{j=0}^{n-1} \alpha^{1-1/q}(j) \leqq \sum_{j=0}^{n-1} (j+1)^2 \alpha^{1-1/q}(j) \equiv k_n(\alpha, q)$. Combining this observation with (11), (10) and (9) and the fact that $k_n(\alpha, q) \geqq 1$, one concludes the proof of (5). □

Define for $(i-1)/n \leqq t \leqq i/n$, $1 \leqq i \leqq n$,

$$(12) \qquad Z_d(t) = W_d((i-1)/n) + [nt-(i-1)]\{W_d(i/n) - W_d((i-1)/n)\} .$$

LEMMA A3. *For any* $1 \leqq q < 2$

$$(13) \qquad E|Z_d(t) - Z_d(s)|^4 \leqq 144 B_n |t-s|^{2/q} , \qquad 0 \leqq s, t \leqq 1 .$$

PROOF. In the case where $d_i \equiv 1/n^{\frac{1}{2}}$ and $\alpha \equiv 0$ (independence) a proof of (13) is given by Shorack with $q = 1$. In view of the above inequality (5) it is clear that the same proof goes through for general $\{d_i\}$ and therefore is not reproduced here. □

LEMMA A4.

$$(14) \qquad \sup_{0 \le t \le 1} |W_d(t) - Z_d(t)| \to_p 0 .$$

PROOF. For the moment assume $d_j \geqq 0$, $1 \leqq j \leqq n$. Since Z_d is a linear interpolation of W_d, we have $\sup_t |Z_d(t) - W_d(t)| \leqq \mathscr{W}_{n1} + \mathscr{W}_{n2}$ where

$$\mathscr{W}_{n1} = \max_i \sup_{(i-1)/n \le t \le i/n} |W_d(t) - W_d(i-1)/n|$$

and

$$\mathscr{W}_{n2} = \max_i \sup_{(i-1)/n \le t \le i/n} |W_d(t) - W_d(i/n)| .$$

Now $d_j \geqq 0$, $1 \leqq j \leqq n$ and (3) imply that for $(i-1)/n \leqq t \leqq i/n$

$$(15) \qquad |W_d(t) - W_d((i-1)/n)| \leqq |W_d(i/n) - W_d((i-1)/n)| + 2 \max_i |d_i| .$$

From (5) above we have for $1 < q < 2$

$$\sum_{i=1}^n |EW_d(i/n) - W_d((i-1)/n)|^4 \leqq 4B_n n^{1-2/q}$$

and from (1.3) $\max_i |d_i| \to 0$ as $n \to \infty$. Combining these observations with (14) we observe that $\mathscr{W}_{n1} = o_p(1)$ as $n \to \infty$. Similarly one shows $\mathscr{W}_{n2} = o_p(1)$ as $n \to \infty$. This way then (14) is proved for nonnegative $\{d_i\}$. In general write $d_j = d_j^+ - d_j^-$, $1 \leqq j \leqq n$; $W_d = \tau_{d^+} W_{d^+} - \tau_{d^-} W_{d^-}$ and observe that because $\tau_d = 1$, we have $\tau_{d^+} \leqq 1$, $\tau_{d^-} \leqq 1$. One also has a similar decomposition for Z_d and using the above proof for each part $\{W_{d^+}\}$ and $\{W_{d^-}\}$ one has the proof of the lemma. □

PROOF OF PROPOSITION A1. From Lemma A3 above and assumptions (1.3) and (1.5) it follows that $\{Z_d\}$ satisfy (4) above. (See Billingsley (1968), page 95.) In view of (14) $\{W_d\}$ must then also satisfy (4). This completes the proof of Proposition A1. □

The following corollary is used in Section 2.

COROLLARY A1. *Let* $\{Z_i, 1 \leqq i \leqq n\}$ *be* s.s.s.m. *with mixing numbers* $\{\alpha_n\}$ *satisfying* (1.5). *Assume that F of model* (0.1) *is a continuous* cdf *and* $\{d_i\}$ *satisfy* (1.3). *Then for each fixed* $|\Delta| \leqq b$

$$\lim_{\delta\to 0} \limsup_{n\to\infty} P(\sup_{|t-s|<\delta} |V_d(t, \Delta) - V_d(s, \Delta)| > \varepsilon) = 0 \quad \forall \varepsilon > 0 \tag{16}$$

where V_d *is defined by* (1.1).

PROOF. In Proposition A1 above take $Y_i = H_\Delta(Z_i - c_i\Delta)$, $1 \leqq i \leqq n$. Then $\{Z_i\}$ s.s.s.m. implies $\{Y_i\}$ is strongly mixing. Also $0 \leqq Y_i \leqq 1$ and $G_i(t) = L_i(t, \Delta)$ and (3) is a priori satisfied. Hence (4) yields (16). □

We also need the following

LEMMA A5. *Let* $\{\xi_{n1}, \cdots, \xi_{nn}\}$ *be a sequence of* s.m. rv's *with mixing numbers* α_n. *Assume* (1.3) *holds for* $\{d_{ni}\}$,

$$\max_j |\xi_{nj}| \leqq 1\,, \qquad n \geqq 1 \tag{17}$$

$$\max_j \|\xi_{nj}\|_2 \to 0 \quad \textit{as} \quad n \to \infty \tag{18}$$

and

$$\limsup_{n\to\infty} \textstyle\sum_{j=0}^{n-1} \alpha_n^{1-1/q}(j) < \infty \quad \textit{for some} \quad 1 < q < \infty\,. \tag{19}$$

Then

$$\operatorname{Var}(\textstyle\sum_j d_{nj}\xi_{nj}) \to 0 \quad \textit{as} \quad n \to \infty\,. \tag{20}$$

PROOF. From Lemma A1 we have, in view of (17), that

$$|\textstyle\sum_{i<j} d_i d_j \operatorname{Cov}(\xi_i, \xi_j)| \leqq 10 k_d^2 \max_j \|\xi_{nj}\|_q \sum_{j=1}^{n-1} \alpha_n^{1-1/q}(j)\,. \tag{21}$$

Again (17) implies that for $q > 1$

$$\|\xi_{nj}\|_q \leqq \|\xi_{nj}\|_2^{1/q}\,, \qquad 1 \leqq j \leqq n \tag{22}$$

so that (18) implies

$$\max_j \|\xi_{nj}\|_q \to 0\,. \tag{23}$$

Moreover in view of (1.3) we have

$$\operatorname{Var}(\textstyle\sum_j d_j\xi_j) \leqq \max_j \|\xi_{nj}\|_2 + 2|\sum_{i<j} d_i d_j \operatorname{Cov}(\xi_i, \xi_j)|\,.$$

The first term on the right goes to zero by (18), whereas the convergence of the second term to zero follows from (21), (22) and (23). □

We end this section by stating the following

LEMMA A6. *For each* $n \geqq 1$, *let* $\{\xi_{nj}, 1 \leqq j \leqq n\}$ *be* s.m. *with mixing number* α_n. *Assume these* rv's *to be uniformly bounded and*

$$\limsup_{n\to\infty} \textstyle\sum_{j=1}^n j^2\alpha_n(j) < \infty . \tag{24}$$

Further assume

$$\max_i d_{ni}^2/\tau_d^2 \to 0 \tag{25}$$

and

$$\liminf_{n\to\infty} \tau_n^2\tau_d^{-2} > 0 \tag{26}$$

where $\tau_n^2 = \mathrm{Var}\,(\sum_{i=1}^n d_{ni}\xi_{ni})$. *Then*

$$\mathscr{L}(\tau_n^{-1} \textstyle\sum_1^n d_{ni}\xi_{ni}) \to N(0, 1) .$$

PROOF. This is essentially Theorem 3.1 of Mehra–Rao (1975). It is clear from their proof that their theorem remains valid under the above slightly general conditions. □

REMARK. It is important to note that our Proposition A1 above is not contained in Mehra–Rao (1975), Theorem 3.2 since in their case $\{Y_{nj}\}$ must have the same marginals. Being in the regression model we necessarily have rv's with different marginals. □

REFERENCES

[1] ADICHIE, J. N. (1967). Estimation of regression parameter based on ranks. *Ann. Math. Statist.* **38** 894–904.

[2] BICKEL, P. J. (1973). On some analogues to linear combination of order statistics and linear models. *Ann. Statist.* **1** 597–616.

[3] BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[4] DEO, C. M. (1973). A note on empirical processes of strong mixing sequences. *Ann. Probability* **1** 870–875.

[5] GASTWIRTH, J. L. and RUBIN, H. (1975a). Asymptotic distribution theory of empiric cdf for mixing processes. *Ann. Statist.* **3** 809–824.

[6] GASTWIRTH, J. L. and RUBIN, H. (1975b). Behavior of robust estimators on dependent data. *Ann. Statist* **3** 1070–1100.

[7] HUBER, P. J. (1973). Robust regression: Asymptotic, conjectures, Monte Carlo. *Ann. Statist.* **1** 799–821.

[8] JURECKOVA, J. (1969). Asymptotic linearity of rank statistics in regression parameter. *Ann. Math. Statist.* **40** 1889–1900.

[9] KOUL, H. L. (1969). Asymptotic behavior of Wilcoxon type confidence regions in multiple linear regression. *Ann. Math. Statist.* **40** 1950–1979.

[10] KOUL, H. L. (1970). A class of ADF tests for the subhypothesis in regression. *Ann. Math. Statist.* **41** 1273–1281.

[11] KOUL, H. L. and STAUDTE, R. G., JR. (1972). Weak convergence of weighted empirical cumulatives based on ranks. *Ann. Math. Statist.* **43** 832–841.

[12] MEHRA, K. L. and RAO, M. S. (1975). Weak convergence of generalized empirical processes relative to d_q under strong mixing. *Ann. Probability* **3** 979–991.

[13] PHILLIP, W. (1969). The central limit problem for mixing sequences of random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **12** 155–171.

[14] SHORACK, G. (1973). Convergence of reduced empirical and quantile processes with applications to functions of order statistics in non-i.i.d. case. *Ann. Statist.* **1** 146–152.

DEPARTMENT OF STATISTICS AND PROBABILITY
MICHIGAN STATE UNIVERSITY
WELLS HALL
EAST LANSING, MICHIGAN 48824