The Annals of Statistics
1975, Vol. 3, No. 4, 809–824

THE ASYMPTOTIC DISTRIBUTION THEORY OF THE EMPIRIC CDF FOR MIXING STOCHASTIC PROCESSES

BY JOSEPH L. GASTWIRTH[1] AND HERMAN RUBIN[2]

The George Washington University and Purdue University

This paper introduces a new mixing condition for stationary processes which is weaker than ϕ-mixing but stronger than strong mixing. Many processes arising in applications, e.g., first order autoregressive processes, obey the conditions. The main result is that the empiric cdf of a sample from such processes converges to a Gaussian process.

1. Introduction and summary. The purpose of this paper is to lay the probabilistic foundations for a study of the large sample theory of robust estimators for dependent processes. While there is substantial literature on the convergence of the empiric cdf to a Brownian bridge for i.i.d. observations, only recently [1] have these results been extended to dependent processes. In order to derive the results, we introduce a new mixing condition which is stronger than Rosenblatt's strong mixing [2], [9] yet weaker than the concept of mixing introduced by Doob [3] and used by Billingsley [1].

The new mixing condition is presented in Section 2 and several first order autoregressive processes are shown to satisfy it. The third section is concerned with proving the convergence of the empiric cdf to a Gaussian process. In the companion paper [4] the results are illustrated by deriving the asymptotic distribution of general linear combinations of order statistics.

2. Mixing conditions and their applications to statistics of the form $\sum f(X_i)$. In order for functions of a stochastic process to obey a central limit theorem some sort of asymptotic independence is usually required. Therefore, we develop analogs of Rosenblatt's mixing number and introduce another measure of dependence, between Rosenblatt's and Doob's, which is readily computable.

If $\{X_i\}$, $i \in I$, and $\{X_j\}$, $j \in J$, are two indexed families of rv's the mixing number measuring the dependence between them is

$$\alpha(I; J) = \sup_{A,B} |P(AB) - P(A)P(B)| , \tag{2.1}$$

where the range of A is the Borel field generated by the $\{X_i\}$, $i \in I$, and the range

Received September 1970; revised October 1974.

[1] This research was supported in part by the National Science Foundation Research Grants No. GP-7118 and GP-20527. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[2] This research was supported in part by the Office of Naval Research Contract N00014-67-A-0226-0008, project number NRO 42-216 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

AMS 1970 *subject classifications.* Primary 62E20; Secondary 60F05, 60G10.

Key words and phrases. Asymptotic distribution, empiric distribution function, strong mixing, stationary processes.

of B is the Borel field generated by the $\{X_j\}$, $j \in J$. Rosenblatt's mixing number for stationary processes will be denoted by α_n and is defined as

$$\alpha_n = \alpha(I; J) , \tag{2.2}$$

where $I = \{i; i \leqq 0\}$ and $J = \{j; j \geqq n\}$. For finite sets of rv's $I = \{i_1, \cdots, i_q\}$ and $J = \{j_1, \cdots, j_r\}$ the mixing number (2.1) will be denoted by

$$\alpha(i_1, \cdots, i_q; j_1, \cdots, j_r) . \tag{2.3}$$

Our measure of dependence is defined in terms of a function measuring the conditional deviation from independence. Let $W = \{X_i, i \in I\}$ have distribution R and $Z = \{X_j, j \in J\}$ have distribution Q, and (W, Z) have distribution P; then

$$\Delta(I; J; y) = \int |P(dz \mid y) - Q(dz)| . \tag{2.4}$$

In the case where ordinary conditional probability distributions exist $\Delta(I; J; y)$ is the total variation of the difference between the conditional distribution of z given $W = y$ and the marginal distribution of z. In the general situation $\Delta(I; J; \cdot)$ is defined to satisfy

$$\begin{aligned} &\int \Delta(I; J; y)g(y)R(dy) \\ &\quad = \sup_{|h|\leqq 1} \left[\iint h(y, z)g(y)P(dy, dz) - \iint h(y, z)g(y)R(dy)Q(dz)\right] \end{aligned} \tag{2.5}$$

for all positive R-integrable g. The analog of (2.3) is

$$\Delta(i_1, \cdots, i_q; j_1, \cdots, j_r; y) = \Delta(I; J; y) , \tag{2.6}$$

where $I = \{i_1, \cdots, i_y\}$ and $J = \{j_1, \cdots, j_r\}$. We also define

$$\Delta(n; y) = \Delta(I; J; y) , \tag{2.7}$$

where $I = \{i : i \leqq 0\}$ and $J = \{j : j \geqq n\}$, and

$$\Delta_n = \int \Delta(n; y)\, dR(y) = ||\Delta(n)||_1 . \tag{2.8}$$

For stationary processes, if $i_1 \leqq \cdots \leqq i_q < j_1 \leqq \cdots \leqq j_r$, it is obvious that

$$||\Delta(i_1, \cdots, i_q; j_1, \cdots, j_m)||_s \leqq ||\Delta(n)||_s , \tag{2.9}$$

where $n = j_1 - i_q$.

REMARKS. The quantity $||\Delta(I; J)||_1$ equals the total variation of $P - Q \times R$. Moreover,

$$\alpha(I; J) \leqq ||\Delta(I, J)||_1/4 . \tag{2.10}$$

On the other hand, conditions using the s-norms of the function (2.3) are weaker than those depending on the concept of ϕ mixing introduced by Doob [3] and developed by Billingsley [1] and Serfling [12]. Indeed, for stationary processes Billingsley's ϕ_n is $||\Delta(n)||_\infty/2$.

At this stage it seems appropriate to illustrate the computability of reasonable bounds for the s-norms of $\Delta(k)$ for first order autoregressive processes. For

Markov processes $\Delta(k) = \Delta(0; k)$ and $\Delta_k = ||\Delta(0; k)||_1$. Using the representation

$$X_k = \rho^k X_0 + U_k \tag{2.11}$$

and, denoting the stationary density of X_k by f, the mass of the non-absolutely continuous component of the distribution of U_k by M_k, and the density of the absolutely continuous component by f_k, we have

$$\Delta(0; k; y) = M_k + \int |f(x) - f_k(x - \rho^k y)|\, dx\,. \tag{2.12}$$

If the density f is not supported on a bounded interval of the real line, then

$$\sup_y \Delta(0; k; y) = 2\,. \tag{2.13}$$

This shows that autoregressive processes are not ϕ mixing. We now discuss three examples. Since all the examples have symmetric marginals the results are independent of the sign of ρ. Hence, we assume $\rho > 0$.

EXAMPLE 1. Our first example is the double-exponential process. Using the characterization due to Gastwirth and Wolff [5] it is seen that $M_k = \rho^{2k}$ and $f_k(x) = (\frac{1}{2})(1 - \rho^{2k}) \exp(-|x - \rho^k y|)$. Thus,

$$\begin{aligned} \Delta(0, k, y) &= \rho^{2k} + \int (\tfrac{1}{2})|\exp[-|x|] - (1 - \rho^{2k}) \exp[-|x - \rho^k y|]|\, dx \\ &\leqq \rho^{2k} + \tfrac{1}{2} \int \rho^{2k} \exp(-|x|)\, dx \\ &\qquad + \int (\tfrac{1}{2})(1 - \rho^{2k})|\exp(-|x|) - \exp(-|x - \rho^k y|)|\, dx \\ &\leqq 2\rho^{2k} + \tfrac{3}{2}|\rho^k y|(1 - \rho^{2k})\,. \end{aligned} \tag{2.14}$$

Moreover, the Minkowski inequality implies that

$$||\Delta(k)||_s \leqq 2\rho^{2k} + |\rho^k|(\tfrac{3}{2})(s!)^{1/s}\,, \tag{2.15}$$

so that $\Delta_k \leqq C|\rho^k|$.

EXAMPLE 2. For Gaussian Markov processes we must bound

$$\begin{aligned} &\Delta(0; k; y) \\ &\quad = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \left| \frac{\exp[-(x - \rho^k y)^2/2(1 - \rho^{2k})]}{(1 - \rho^{2k})^{\frac{1}{2}}} - \exp[-x^2/2] \right| dx\,. \end{aligned} \tag{2.16}$$

In order that the reader can more easily follow the computations, we will use ρ instead of ρ^k until formula (2.22). We see that $\Delta(0; k; y)$ is bounded by

$$\begin{aligned} &(2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \left| \frac{\exp[-(x - \rho y)^2/2(1 - \rho^2)]}{(1 - \rho^2)^{\frac{1}{2}}} - \frac{\exp[-x^2/2(1 - \rho^2)]}{(1 - \rho^2)^{\frac{1}{2}}} \right| dx \\ &\qquad + (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \left| \frac{\exp[-x^2/2(1 - \rho^2)]}{(1 - \rho^2)^{\frac{1}{2}}} - \exp[-x^2/2] \right| dx\,. \end{aligned} \tag{2.17}$$

The first term of (2.17) is treated by noting that both functions are probability densities which are equal when $x = \rho y/2$. If $y > 0$ this integrand is positive when $x > \rho y/2$ (when $y < 0$, the reverse is true but the same bound occurs by symmetry) so that

$$(2\pi)^{-\frac{1}{2}} \int \left| \frac{\exp[-(x-\rho y)^2/2(1-\rho^2)]}{(1-\rho^2)^{\frac{1}{2}}} - \frac{\exp[-x^2/2(1-\rho^2)]}{(1-\rho^2)^{\frac{1}{2}}} \right| dx$$

$$= \frac{2}{[(2\pi)(1-\rho^2)]^{\frac{1}{2}}} \int_{x>\rho y/2} [\exp[-(x-\rho y^2)/2(1-\rho^2)] - \exp[-x^2/2(1-\rho^2)]]\, dx \tag{2.18}$$

$$= \frac{2}{[(2\pi)(1-\rho^2)]^{\frac{1}{2}}} \int_{-\rho y/2}^{\rho y/2} \exp[-t^2/2(1-\rho^2)]\, dt$$

$$= \frac{4}{(1-\rho^2)^{\frac{1}{2}}} \frac{1}{(2\pi)^{\frac{1}{2}}} \int_0^{\rho y/2} \exp[-t^2/2(1-\rho^2)]\, dt$$

$$\leqq \frac{K|\rho y|}{(1-\rho^2)^{\frac{1}{2}}} .$$

The right side of (2.17) is bounded by observing that the integrand is bounded by $Cx^2\rho^2\exp[-x^2/2]$ when $|\exp[-x^2\rho^2/2(1-\rho^2)] - (1-\rho^2)^{\frac{1}{2}}| < 0$. When $|\exp[-x^2\rho^2/2(1-\rho^2)] - (1-\rho^2)^{\frac{1}{2}}| > 0$, an application of Taylor's theorem yields the bound

$$|(1-\rho^2)^{-\frac{1}{2}} - 1| \leqq B\rho^2/(1-\rho^2)^{\frac{1}{2}} \tag{2.19}$$

which implies that the integrand is bounded by $B\rho^2\exp[-x^2/2]$. In any case the integrand in the right side of (2.17) is bounded by

$$(Ax^2 + B)\rho^2 \exp[-x^2/2]/(1-\rho^2)^{\frac{1}{2}} \tag{2.20}$$

so the second term of (2.17) is bounded by $K\rho^2$. When $|\rho|$ is bounded away from 1, expression (2.17) is bounded by

$$\frac{K\rho^2}{(1-\rho^2)^{\frac{1}{2}}} + \frac{K^1|\rho y|}{(1-\rho^2)^{\frac{1}{2}}} = C_1\rho^2 + C_2|\rho y| , \tag{2.21}$$

where $(1-\rho^2)^{-\frac{1}{2}}$ is absorbed in the constants C_1, C_2. Thus

$$||\Delta(k)|| = C_1\rho^{2k} + \rho^k C_2 \int |y| \frac{\exp[-y^2/2]\, dy}{(2\pi)^{\frac{1}{2}}} = C_1\rho^{2k} + C_2\rho^k . \tag{2.22}$$

More generally, applying the Minkowski inequality yields

$$||\Delta(k)||_s \leqq C_1\rho^{2k} + C_2|\rho|^k[\mu_s]^{1/s} , \tag{2.23}$$

where μ_s is the sth absolute moment of a unit normal rv.

Notice that (2.23) is very similar to the result (2.15) obtained for the double-exponential process and one might be tempted to conjecture that for Markov processes $\Delta_k = \Delta(0, k) \leqq C|\rho^k|$. Our final example shows that this is not true.

EXAMPLE 3. For the Cauchy process, U_k is a Cauchy rv with scale parameter $(1-\rho^k)$, so that

$$\Delta(0; k; y) = \pi^{-1} \int_{-\infty}^{\infty} |(1+x^2)^{-1}(1-\gamma)^{-1}(1+(x-\gamma y)^2/(1-\gamma)^2)^{-1}|\, dx , \tag{2.24}$$

where $\gamma = \rho^k$. Expression (2.24) is less than or equal to

$$\begin{aligned}(2.25)\qquad &\pi^{-1}\int |(1+x^2)^{-1} - (1-\gamma)^{-1}(1+x^2/(1-\gamma)^2)^{-1}|\,dx \\ &\quad + \pi^{-1}\int |(1-\gamma)^{-1}(1+x^2/(1-\gamma)^2)^{-1} \\ &\qquad - (1-\gamma)^{-1}(1+(x-\gamma y)^2/(1-\gamma)^2)^{-1}|\,dx\,.\end{aligned}$$

We discuss the case where $\gamma > 0$. The left side of (2.25) is a difference of two probability densities which are equal when $x = \pm(1-\gamma)^{\frac{1}{2}}$ and the factor $(1+x^2)^{-1}$ is larger when $x > (1-\gamma)^{\frac{1}{2}}$. Hence the left side of (2.25) is

$$\begin{aligned}(2.26)\qquad &\leqq 2\pi^{-1}\int_{|x|>(1-\gamma)^{\frac{1}{2}}} |(1+x^2)^{-1} - (1-\gamma)^{-1}(1+x^2/(1-\gamma)^2)^{-1}|\,dx \\ &\leqq 4\pi^{-1}\int_{x>(1-\gamma)^{\frac{1}{2}}} [(1+x^2)^{-1} - (1-\gamma)^{-1}(1+x^2/(1-\gamma)^2)^{-1}]\,dx \\ &= 4\pi^{-1}\arctan(1-\gamma)^{-\frac{1}{2}} - \arctan(1-\gamma)^{\frac{1}{2}}] \leqq 2\gamma\,.\end{aligned}$$

The second integral in (2.25) is handled by substituting $w = x(1-\gamma)^{-1}$ and $z = \gamma y(1-\gamma)^{-1}$, yielding

$$(2.27)\qquad \pi^{-1}\int \left|\frac{1}{(1+(w-z)^2)} - \frac{1}{1+w^2}\right| dw\,,$$

which is the difference of two Cauchy densities with locations z and 0 respectively and the densities are equal when $w = z/2$. The same type of argument used above shows that expression (2.27) is

$$(2.28)\qquad \leqq 4\pi^{-1}\arctan\left|\frac{\gamma y}{2(1-\gamma)}\right| \leqq \frac{4}{\pi(1-\gamma)}\arctan(\gamma|y|) \leqq C\arctan\gamma|y|$$

if $\gamma < \frac{1}{2}$. Thus

$$(2.29)\qquad \Delta(0;k;y) \leqq 2\gamma + C\arctan\gamma|y|$$

and

$$(2.30)\qquad ||\Delta(k)||_s \leqq 2\gamma + C\left[\int |\arctan(\gamma y)|^s \frac{1}{\pi(1+y^2)}\,dy\right]^{1/s}.$$

By decomposing the range of integration and using elementary inequalities one can show that

$$(2.31)\qquad ||\Delta(k)||_s \leqq 2\gamma + K_s\gamma^{\frac{1}{2}}$$

where K_s is a constant depending on s. More interesting is the case where $s = 1$. Here

$$\begin{aligned}(2.32)\qquad &\int_{-\infty}^{\infty} |\arctan\gamma y|(1+y^2)^{-1}\,dy \\ &\leqq 2\left[\int_0^{\gamma^{-1}} \frac{2\gamma y}{(1+y^2)}\,dy + \frac{\pi}{2}\int_{\gamma^{-1}}^{\infty}(1+y^2)^{-1}\,dy\right],\end{aligned}$$

where we use the bound $\arctan x \leqq x$ for $y \leqq (\gamma)^{-1}$. The right side of (2.32) is

$$(2.33)\qquad \leqq 2\left[(\tfrac{1}{2})\gamma\log(1+(\gamma^2)^{-1}) + \frac{\pi y}{2}\right] = \gamma + \gamma\log(1+(\gamma^2)^{-1})\,,$$

so that

$$||\Delta(k)||_1 \leqq C_1\rho^k + C_2\rho^k \log(1+\rho^{-2k}) . \tag{2.34}$$

For large x, $\log(1+x)$ is essentially $\log x$ so that $\log(1+\rho^{-2k})$ is nearly $2k\log(1/\rho)$, which implies that

$$||\Delta(k)||_1 \leqq A\rho^k + Bk\rho^k . \tag{2.35}$$

REMARK. If one refines the argument to get lower bounds on $\Delta(k)$, one can show that the k term in (2.35) cannot be eliminated.

In order to prove the asymptotic normality of rv's of the form $S_n = \sum f(X_i)$, where f is a bounded function, we present the result in terms of the mixing numbers. First we recall the following lemma due to Ibragimov [6].

LEMMA 2.1. *If U and V are bounded by C_1 and C_2 respectively and U is measurable* w.r.t. *the Borel field generated by X_i for $i \in I$ and V is measurable* w.r.t. *the Borel field generated by $\{X_j; j \in J\}$, then*

$$\operatorname{Cov}(U, V) \leqq 4C_1C_2\alpha(I; J) . \tag{2.36}$$

We now formally state

THEOREM 2.1. *Whenever $\{X_i\}$ is a strongly mixing* S.S.P. *such that*

$$\textstyle\sum_{k=1}^{\infty} \alpha(0, k) < \infty , \tag{2.37}$$

$$\textstyle\sum\sum_{j\neq k, 1\leqq j,k\leqq n} \min(\alpha(0, j; k), \alpha(0; j, k), \alpha(0, k; j)) = O(n) , \tag{2.38}$$

and

$$\begin{aligned}\textstyle\sum\sum\sum_{i+j+k\leqq n} \min(&\alpha(0; i, i+j, i+j+k),\\ &\alpha(0, i; i+j; i+j+k), \alpha(0, i, i+j; i+j+k)) = O(n) ,\end{aligned} \tag{2.39}$$

then any statistic of the form $S_n = \sum_{i=1}^n f(X_i)$, where f is a bounded function, is asymptotically normally distributed, i.e.,

$$n^{-\frac{1}{2}}[S_n - E(S_n)] \to_{\mathscr{L}} N(0, \sigma^2) , \tag{2.40}$$

where $\sigma^2 = \lim_{n\to\infty} n^{-1}V(S_n)$.

REMARK. These conditions show that asymptotic normality is determined by the strength of dependence of various subsets of four rv's, which suggests that strong mixing should not be necessary.

PROOF. The result will follow once the fourth moment condition of the Blum–Rosenblatt Theorem is verified, i.e., letting $A_i = f(X_i) - E[f(X_i)]$, $E|\sum A_i|^4$ must be $O(n^2)$. Expanding $E|\sum A_i|^4 = \sum\sum\sum\sum E(A_iA_jA_kA_l)$ one obtains

$$\begin{aligned}&\textstyle\sum_i EA_i^4 + 4\sum_i\sum_j E(A_i^3A_j) + 6\sum\sum_{i<j} E(A_i^2A_j^2)\\ &\quad\textstyle + 12\sum\sum_{i<j}\sum_k E(A_iA_jA_k^2)\\ &\quad\quad\textstyle + 24\sum\sum\sum\sum_{i<j<k<l} E(A_iA_jA_kA_l) .\end{aligned} \tag{2.41}$$

As the A_i are bounded with mean 0, $\sum E(A_i^4) \leqq Kn$, $\sum\sum E(A_i^3A_j) \leqq Kn^2$ and

$\sum\sum E(A_i^2 A_j^2) \leqq Kn^2$, when K is an appropriate constant. Now

$$(2.42)\qquad \begin{aligned} E(A_i A_j A_k^2) &= E(A_i A_j)E(A_k^2) + \operatorname{Cov}(A_i A_j;\, A_k) \\ &= \operatorname{Cov}(A_i;\, A_j A_k^2) \\ &= \operatorname{Cov}(A_i A_k^2;\, A_j)\,. \end{aligned}$$

Bounding the covariances by Lemma 2.1 yields

$$(2.43)\qquad \begin{aligned} |E(A_i A_j A_k^2)| &\leqq |E(A_i A_i)|E(A_k^2)| + K\alpha(i, j;\, k) \\ &\leqq K\alpha(i;\, j, k) \\ &\leqq K\alpha(i, k;\, j)\,. \end{aligned}$$

Since $E(A_i A_j) = E(A_i)E(A_j) + \operatorname{Cov}(A_i, A_j)$, $|E(A_i A_j)| \leqq K\alpha(0, j - i)$,

$$(2.44)\qquad \sum\sum\nolimits_{i<j} \sum\nolimits_k |E(A_i A_j)|E(A_k^2) \leqq n^2 K \sum\nolimits_{k=1}^{\infty} \alpha(0, k)\,.$$

As $E(A_i A_j A_k^2) \leqq |E(A_i A_j)|E(A_k)^2 + K \min(\alpha(i, j;\, k), \alpha(i;\, j, k), \alpha(i, k;\, j))$,

$$(2.45)\qquad \begin{aligned} &\sum\sum\sum\nolimits_{i<j<k} E(A_i A_j A_k^2) \\ &\qquad \leqq n^2 K \sum\nolimits_{k=1}^{\infty} \alpha(0, k) \\ &\qquad\qquad + nK \sum\sum\nolimits_{j,k} \min(\alpha(0, j;\, k), \alpha(0;\, j, k), \alpha(0, k;\, j))\,, \end{aligned}$$

where $j - i$ and $k - j$ are now denoted by j and k. The fourth order terms are handled by noting that if $i < j < k < l$, then

$$(2.46)\qquad \begin{aligned} |E(A_i A_j A_k A_l)| &\leqq |E(A_i A_j)||E(A_k A_l)| + K\alpha(i, j;\, k, l) \\ &\leqq K\alpha(i;\, j, k, l) \\ &\leqq K\alpha(i, j, k;\, l)\,. \end{aligned}$$

As

$$(2.47)\qquad \sum\sum\sum\sum\nolimits_{i<j<k<l} |E(A_i A_j)||E(A_k A_l)| \leqq n^2 K^2 \Big(\sum\nolimits_{r=1}^{\infty} \alpha(0, r)\Big)^2\,,$$

$$(2.48)\qquad \begin{aligned} &\sum\sum\sum\sum\nolimits_{i<j<k<l} |E(A_i A_j A_k A_l)| \\ &\qquad \leqq n^2 K' \Big[\sum \alpha(0, r)\Big]^2 + K \sum\sum\sum\sum\nolimits_{i<j<k<l} \min(\alpha(i, j;\, k, l), \\ &\qquad\qquad \alpha(i;\, j, k, l), \alpha(i, j, k;\, l))\,. \end{aligned}$$

The second term on the right side of (2.48) is

$$\begin{aligned} \leqq Kn \sum\sum\sum\nolimits_{j+k+l\leqq n} \min\{&\alpha(0, j;\, j + k, j + k + l), \\ &\alpha(0;\, j, j + k, j + k + l), \alpha(0, j, j + k;\, j + k + l)\} \end{aligned}$$

when j, k and l now denote $j - i$, $k - j$ and $l - k$. Thus, assumptions (2.37), (2.38), and (2.39) imply that (2.45) and (2.48) are $O(n^2)$, so that the conditions of the Blum-Rosenblatt theorem hold.

An important Corollary of Theorem 2.1 is

COROLLARY 2.1. *Let $\{X_i\}$ obey the conditions of Theorem 2.1, then the finite dimensional marginal distributions of the empiric process $n^{\frac{1}{2}}[F_n(t) - F(t)]$ converges to a multivariate normal distribution.*

PROOF. For any t, $F_n(t) = n^{-1} \sum_1^n I_i(t)$, where $I_i(t)$ is 1 if $X_i < t$ and 0 otherwise. For any set $t_1, \cdots, t_k$ of k values of t, any linear combination $\sum_{j=1}^k a_j F_n(t_j)$ is a function of the form $n^{-1} \sum_{i=1}^n f(X_i)$ and Theorem 2.1 applies.

REMARKS.

1) Corollary 2.1 includes the asymptotic normality of the sign test statistic.
2) The conditions of Theorem 2.1 hold whenever $\sum k^2 \alpha_k = O(n)$.
3) Since $\alpha_k \leqq \Delta_k$, Theorem 2.1 is valid whenever $\sum k^2 \Delta_k = O(n)$. This condition is easily verified for the three examples discussed as $\sum k^2 \Delta_k$ converges in all three examples.

We conclude this section by showing that for Gaussian processes the sign test statistic is asymptotically normally distributed if $\sum |\rho_k| < \infty$. Since bounded functions f have finite variance and the normal distribution is determined by its moments we can approximate f in L^2 (w.r.t. the normal distribution) by a polynomial P_ε so that $||f - P_\varepsilon|| = ||f_\varepsilon|| < \varepsilon$. The statistic $n^{-\frac{1}{2}} \sum_{i=1}^n P_\varepsilon(X_i)$ is asymptotically normally distributed by Sun's Theorem [13]. The variance of $n^{-\frac{1}{2}} \sum_{i=1}^n f_\varepsilon(X_i) < \varepsilon \sum_{-n}^n |\rho_k|$, where we have bounded $\text{Cov}\,(f_\varepsilon(X_i), f_\varepsilon(X_j))$ by $\varepsilon|\rho_{i-j}|$ by Sarmanov's Lemma [10]. By the Mann–Wald Theorem [7], the result follows.

3. The convergence of the empric cdf to a Gaussian process. In this section we show that the empiric cdf of a strong mixing S.S.P. converges to a Gaussian process provided that the Δ (or α) functions defined in Section 2 obey some regularity conditions. These conditions are weaker than Doob's concept of ϕ mixing so that our result is stronger than Billingsley's Theorem 22.1. In particular, the Gaussian, Cauchy and double-exponential first order autoregressive processes are not ϕ mixing but satisfy our conditions. For Gaussian processes, a special result is derived showing that $\sum |\rho_k| < \infty$ suffices to guarantee the convergence of the empiric cdf to a Gaussian process.

The first step in the proof is an application of a lemma of Rubin [11] which gives verifiable conditions which imply Prokhorov's necessary and sufficient condition for processes to converge to a limiting process with a.s. continuous sample paths. The next step is to apply Theorem 2.1 to prove that the finite dimensional marginal distributions converge to the appropriate multivariate Gaussian distribution.

Before stating our main result we prove a generalization of Doob's Lemma 7.1 [3] which applies to our Δ functions. Specifically we have

LEMMA 3.1. *If f and g are functions such that $E(|f(X)|^q) < \infty$, $E(|g(Y)|^r) < \infty$ and if $1/r + 1/q + 1/s = 1$ and $1 \leqq q, r, s \leqq \infty$, then*

$$|\text{Cov}\,(f(X), g(Y))| \leqq 2^{1/q} ||f||_q ||g||_r ||\Delta^{1/r+1/s}||_s, \tag{3.1}$$

where Δ is defined in Section 2.

PROOF. Writing

$$\text{Cov}\,[f(X)g(Y)] = \int g(y)\{\int f(x)[dP(x\,|\,y) - dQ(x)]\}\,dR(y) \tag{3.2}$$

and applying the Hölder inequality to $g(y)$ and $h(y) = \int f(x)[dP(x\,|\,y) - dQ(x)]$ yields

$$|\text{Cov}\,f(X)g(Y)| \leqq [\int |g(y)|^r\,dR(y)]^{1/r} \cdot [\int |h(y)|^t\,dR(y)]^{1/t}\,, \tag{3.3}$$

where $r^{-1} + t^{-1} = 1$. Since

$$h(y) = \int f(x)\,\text{sgn}\,(dP(x\,|\,y) - dQ(x))|dP(x\,|\,y) - dQ(x)|\,,$$

applying Hölder's inequality yields

$$|h(y)|^t \leqq (\int |f(x)|^t|dP(x\,|\,y) - dQ(x)|)\Delta^{t/r}(y)\,. \tag{3.4}$$

Setting $1/t = 1/q + 1/s$ applying the Hölder inequality again yields

$$\int |f(x)|^t|dP(x\,|\,y) - dQ(x)| \leqq (\int |f(x)|^q|dP(x\,|\,y) - dQ(x)|)^{t/q} \cdot [\Delta(y)]^{t/s} \tag{3.5}$$

so that the second factor on the right side of (3.3) is bounded by the tth root of

$$\int (\int |f(x)|^q|dP(x\,|\,y) - dQ(x)|)^{t/q} \cdot [\Delta(y)]^{t(1/s+1/r)}\,dR(y)\,. \tag{3.6}$$

Applying the Hölder inequality once more shows that (3.6) is

$$\leqq [\int (\int |f(x)|^q|dP(x\,|\,y) - dQ(x)|)\,dR(y)]^{t/q} \cdot [\int \Delta^{1+s/r}(y)\,dR(y)]^{t/s} \tag{3.7}$$

and (3.1) follows by taking tth roots and noting that

$$\begin{aligned} &\int\int |f(x)|^q|dP(x\,|\,y) - dQ(x)|\,dR(y) \\ &\qquad \leqq \int |f(x)|^q \int [|dP(x\,|\,y) + dQ(x)|]\,dR(y) \\ &\qquad \leqq 2 \int |f(x)|^q\,dQ(x)\,. \end{aligned} \tag{3.8}$$

A further useful generalization is

LEMMA 3.2. *If f_i has support F_i, g_i has support G_i, where the sets F_i are pairwise disjoint and the sets G_i are pairwise disjoint, then*

$$\sum |\text{Cov}\,(f_i, g_i)| \leqq 2^{1/q}||f||_q||g||_r||\Delta^{1/r+1/s}||_s \tag{3.9}$$

where $f = \sum f_i$, $g = \sum g_i$ and g, r, s are as in Lemma 3.1.

PROOF. Observe that in deriving Lemma 3.1 only that part of $\Delta(y)$ for y in the support of g is used. Letting X_{G_i} denote the indicator function of the set G_i and $\Delta_i = \Delta X_{G_i}$, Lemma 3.1 implies that

$$\text{Cov}\,(f_i, g_i) \leqq 2^{1/q}||f_i||_q||g_i||_s||\Delta_i^{1/r+1/s}||_s\,. \tag{3.10}$$

The conclusion follows by applying Hölder's inequality to the series and using the fact that the supporting sets of each function are disjoint.

The main technical result of the section is given by

THEOREM 3.1. *Whenever* $\{X_i\}$ *is a* S.S.P. *such that*

$$\sum_{k=1}^{\infty} ||\Delta(0, k)||_1 < \infty , \tag{3.11}$$

$$\sum_{j+k\leq n} \min (||\Delta(0, j; j+k)||_1, ||\Delta(0; j, j+k)||_1) = (n(\log n)^{-5}) , \tag{3.12}$$

$$\begin{aligned}\sum_{i+j+k\leq n} \min (&||\Delta(0; i, i+j, i+j+k)||_1, ||\Delta(0, i; i+j, i+j+k)||_1,\\ &||\Delta(0, i, i+j; i+j+k)||_1) = o(n(\log n)^{-5})\end{aligned} \tag{3.13}$$

and either

$$\sum_{k=1}^{\infty} ||\Delta(0, k)||_s < \infty \qquad \textit{for some} \quad s > 1 \tag{3.14a}$$

or

$$\alpha(0, k) = O(k^{-1}(\log k)^{-6-\varepsilon}) , \tag{3.14b}$$

then the empiric process $n^{\frac{1}{2}}[F_n(t) - F(t)]$ *obeys the conditions of Prokhorov's continuity theorem*, i.e., *as* $n \to \infty$ *it converges to a process with* a.s. *continuous paths if the finite dimensional marginals converge.*

Because of the importance of Theorem 3.1 we make the

DEFINITION. A process $\{X_i\}$ obeying the conditions of Theorem 3.1 is a strongly mixing Δ_s process.

Before proceeding to the proof of the theorem we recall some useful results due to Rubin [11].

LEMMA 3.3 (Rubin). *Let* X_n *be a separable process defined on* [0, 1] *such that*

$$X_n(t + u) - X_n(t) > -\psi_n(u) \qquad \textit{for} \quad u > 0 , \tag{3.15}$$

where ψ_n *is increasing on* (0, 1), *and for some* $\lambda > 0$

$$\sum_{j=1}^{2^i} E\left|X_n(j/2^i) - X_n\left(\frac{j-1}{2^i}\right)\right|^{\lambda} \leq \gamma_{in} < \infty . \tag{3.16}$$

For any $\varepsilon > 0$, *let* $R_n(\varepsilon)$ *be the smallest integer such that*

$$\psi_n(2^{-R_n(\varepsilon)}) < \varepsilon . \tag{3.17}$$

The Prokhorov continuity condition is satisfied if for every $\varepsilon > 0$ *and* $\eta > 0$

$$\limsup_n \sum_{i=l}^{R_n(\varepsilon)} \gamma_{in}^{1/\lambda+1} < \eta . \tag{3.18}$$

A frequently useful corollary is

LEMMA 3.4. *If* X_n *is a separable process satisfying* (3.15) *and*

$$E|X_n(t + u) - X_n(t)|^{\lambda} < \phi_n(u) \tag{3.19}$$

then Prokhorov's condition is satisfied if for every $\varepsilon > 0$ *and* $\eta > 0$ *there is an* l *such that*

$$\limsup_n \sum_{i=l}^{R_n(\varepsilon)} (2^i\phi_n(2^{-i}))^{1/\lambda+1} < \eta , \tag{3.20}$$

where R *is as in Lemma* 3.3.

REMARK. If γ_{in} in Lemma 3.3 or ϕ_n in Lemma 3.4 can be written as a sum of a fixed finite number of functions satisfying (3.18) or (3.20) respectively, the conclusion follows.

PROOF OF THEOREM 3.1. Letting $Y_i = F(X_i)$ one can transform the empiric process to the unit interval and we shall assume that this has been done. We shall verify that the conditions of Lemma 3.3 are satisfied where $\lambda = 4$, $\psi_n(u) = n^{\frac{1}{2}}u$ and $R_n(\varepsilon)$ is the smallest integer greater than $\log(1/\varepsilon) + \frac{1}{2}\log n$, where logarithms are taken to the base 2. Let $B_i(t, u) = 1$ if $Y_i \in [t, t+u)$ and 0 otherwise and let $A_i(t, u) = B_i(t, u) - u$. We shall omit the arguments u and t where no confusion will arise. The computation of the bounds on the fourth moments required to verify condition (3.18) is similar to the derivation of Theorem 2.1. First we note that if $V(t, u) = \prod_{k=1}^{K} A_{j_k}(t, u)$, $W(t, u) = \prod_{g=1}^{G} A_{i_g}(t, u)$, then $|V| \leqq 1$, $E|V|^q \leqq 2u$ for all $q \geqq 1$, $E|W|^r < 2u$ for $r \geqq 1$, and Lemma 3.1 implies that

$$|\mathrm{Cov}\,(V, W)| \leqq \min\,(2^{1/q}(2u)^{1/q+1/r}||\Delta^{1-1/q}||_s, Cu)\,, \tag{3.21}$$

where C is a constant.

Note that any product of A_i's is a constant plus a linear combination of indicator functions (of sets whose probabilities are $\leqq u$), and if the intervals $[t_i, t_i + u)$ are disjoint

$$\sum |\mathrm{Cov}\,(V(t_i, t_i + u), W(t_i, t_i + u))| \leqq M2^{1/q}||\Delta^{1-1/q}||_s\,. \tag{3.22}$$

Clearly (3.22) is minimized when $q = \infty$ and $s = 1$. (For the duration of this proof $||\cdot||$ without a subscript will denote $||\cdot||_1$.)

We now bound the terms in the expansion of $n^{-2}\sum_{j=1}^{2^i} E\{\sum_{k=1}^{n} A_k((j-1)/2^i, j/2^i)\}^4$. Arguing as in Section 2, but using more powerful bounds, we obtain

$$n^{-2}\sum_{j=1}^{2^i}\sum_{k=1}^{n} E(A_k^4) \leqq 2/n\,, \tag{3.23}$$

$$\begin{aligned} &n^{-2}\sum_{j=1}^{2^i}\sum_{k=1}^{n}\sum_{l=1,k\neq l}^{n} E(A_k^3 A_l) \\ &\quad = n^{-2}\sum_{j=1}^{2^i}\sum_{k=1}^{n}\sum_{l=1,k\neq l}^{n}\mathrm{Cov}\,(A_k^3, A_l) \\ &\quad \leqq n^{-2}\sum_{k=1}^{n}\sum_{l=1}^{n} M||\Delta(k, l)|| \leqq 2Mn^{-2}\sum_{k=1}^{n}||\Delta(0, k)||\,, \end{aligned} \tag{3.24}$$

and

$$\begin{aligned} &n^{-2}\sum_{j=1}^{2^i}\sum_{k=1}^{n}\sum_{l=1,k\neq l}^{n} E(A_k^2 A_l^2) \\ &\quad = n^{-2}\sum_{j=1}^{2^i}\sum_{k\neq l}\sum E(A_k^2)E(A_l^2) + n^{-2}\sum_j\sum_{k\neq l}\sum \mathrm{Cov}\,(A_k^2, A_l^2) \\ &\quad \leqq 4\cdot 2^{-i} + 2Mn^{-1}\sum_{k=1}^{n}||\Delta(0, k)||\,. \end{aligned} \tag{3.25}$$

Next

$$\begin{aligned} &n^{-2}|\sum_{j=1}^{2^i}\sum\sum\sum_{k\neq l\neq m} E(A_k^2 A_l A_m) \\ &\quad \leqq n^{-2}\sum\sum\sum\sum E(A_k^2)|E(A_l A_m)| \\ &\qquad + n^{-2}\sum\sum\sum\sum \min\,(|\mathrm{Cov}\,(A_k^2, A_l A_m)|, \\ &\qquad\qquad |\mathrm{Cov}\,(A_k^2 A_l, A_m)|, |\mathrm{Cov}\,(A_k^2 A_m, A_l)|)\,. \end{aligned} \tag{3.26}$$

The first term is bounded by $4\cdot 2^{-i}\sum_l ||\Delta(0, l)||$ and the second is bounded by

$$\begin{aligned} &Mn^{-2}\sum\sum\sum \min\,(||\Delta(k; l, m)||, \Delta(k, l; m)||, ||\Delta(k, m; l)||) \\ &\qquad \leqq 6Mn^{-1}\sum\sum_{k<l}\min\,(||\Delta(0; k, l)||, ||\Delta(0, k; l)||)\,. \end{aligned} \tag{3.27}$$

The most complicated term is

$$\begin{aligned}(3.28)\qquad & 24n^{-2}|\textstyle\sum_j \sum\sum\sum\sum_{k<l<m<h} E(A_k A_l A_m A_h)| \\ & \leqq 24n^{-2} \textstyle\sum |E(A_k A_l)E(A_m A_h)| \\ & \quad + 24n^{-2} \textstyle\sum\sum\sum\sum\sum \min(|\operatorname{Cov}(A_k A_l, A_m A_h)|, \\ & \quad |\operatorname{Cov}(A_k, A_l A_m A_h)|, |\operatorname{Cov}(A_k A_l A_m, A_h)|) .\end{aligned}$$

Treating the second term in a manner similar to that used for the second term of (3.26) shows that it is

$$\begin{aligned}(3.29)\qquad & \leqq 24Mn^{-1} \textstyle\sum_{i+j+k\leqq n} \min(||\Delta(0; i, i+j, i+j+k)||, \\ & ||\Delta(0, i; i+j, i+j+k)||, ||\Delta(0, i, i+j; i+j+k)||) .\end{aligned}$$

There are two alternative bounds for the first term in (3.28). Using Lemma 3.1 or the factor $E(A_k A_l)$ gives the bound

$$48n^{-2}(2^{-i})^{1-1/s} \textstyle\sum\sum\sum\sum\sum ||\Delta(k; l)||_s |E(A_m, A_h)| .$$

Applying Lemma 3.2 (with $q = 0$) to the factor $E(A_m A_h)$ shows that $\sum_j |E(A_m A_h)| \leqq ||\Delta(m_1 h)||_1$. Thus, the final bound is

$$(3.30)\qquad K(2^{-i})^{1-1/s} \textstyle\sum_k ||\Delta(0, k)||_s \sum_m ||\Delta(0, m)||_1 .$$

Alternatively we can apply Ibragimov's lemma to $E(A_k A_l)$ and treat $E(A_m A_h)$ as before. The resulting bound is

$$(3.31)\qquad M \textstyle\sum_{l=1}^{n} \min(2^{-i}, \alpha(0, l)) \sum_k ||\Delta(0, k)|| .$$

Putting terms together we see that

$$(3.32)\qquad \textstyle\sum_{j=1}^{2^i} E|X(j/2^i) - X((j-1)/2^i)|^4 \leqq \sum_{r=1}^{8} \zeta_{rin} ,$$

where the ζ's are the various bounds derived above. To check that (3.18) is satisfied note that there are $O(\log n)$ terms and each summand ζ_{rin} in the bound for γ_{in} is either $o((\log n)^{-5})$ uniformly in i or the series $\sum_{i=1}^{\infty} (\zeta_{rin})^{1/5}$ converges uniformly in n.

REMARK. If the original $\{X_i\}$ are strongly mixing, the conditions of Theorem 3.1 are stronger than the conditions of Theorem 2.1, so that the finite dimensional marginals of the empiric process converge to a multivariate normal distribution and the process converges to a Gaussian process.

A useful corollary is

COROLLARY 3.1. *Whenever* $\{X_i\}$ *is a strongly mixing* s.s.p. *such that* $\Delta_k = o(k^{-2}(\log k)^{-5})$, $n^{\frac{1}{2}}[F_n(t) - F(t)]$ *converges to a Gaussian process with* a.s. *continuous paths.*

PROOF. Since $\alpha(0, k)$ and $||\Delta(0, k)|| \leqq \Delta_k$, (3.11) and (3.14b) are satisfied. The left side of (3.12) is less than $2 \sum_{k=1}^{\infty} k\Delta_k$ and the left side of (3.13) is less than $3 \sum k^2\Delta_k$ so the result will follow once $\sum k^2\Delta_k$ is shown to be $o(n(\log n)^{-5})$. As Δ_k is $o(k^{-2}(\log k)^{-5})$ and as the logarithm function is slowly varying we are done.

REMARK. By the monotonicity of Δ_i no better result of this type can be obtained.

REMARK. Theorem 3.1 applies to the three Markov processes discussed in Section 2.

An alternate theorem using only the mixing numbers is

THEOREM 3.2. *Whenever X_i is a strongly mixing* s.s.p. *such that $\alpha_k = o(k^{-\frac{5}{2}})$, then the empiric process $n^{\frac{1}{2}}[F_n(t) - F(t)]$ converges to a Gaussian process.*

PROOF. The finite dimensional marginals converge to Gaussian marginals by Theorem 2.1. The verification of Prokhorov's conditions proceeds as before except that Lemma 3.4 is used and the covariances are bounded by $K \min(u, \alpha_k)$. In the expansion of $n^{-2}E(\sum A_i)^4$ the worst terms, as in Theorem 3.1 are those in the expansion of

$$\sum\sum\sum\sum_{k<l<m<h} E(A_k A_l A_m A_h)\,,$$

where the bound is

$$M(\textstyle\sum_k \min(u, \alpha_k))^2 + Mn^{-1} \sum k^2 \min(u, \alpha_k)\,. \tag{3.33}$$

It can be shown that in the second term the u does not improve the bound appreciably, therefore to satisfy (3.20) we require that

$$\textstyle\sum_{i=l}^{R_n(\varepsilon)} (2^i M n^{-1} \sum_k k^2 \alpha_k)^{\frac{1}{5}} = (2^{R_n(\varepsilon)} M n^{-1} \sum k^2 \alpha_k)^{\frac{1}{5}} \dfrac{(1 - 2^{\frac{1}{5}(l-1-R_n(\varepsilon))})}{1 - 2^{-\frac{1}{5}}} \tag{3.34}$$

is $< \eta$. The second factor approaches a limit and $2^{R_n(\varepsilon)} < 2n^{\frac{1}{2}}/\varepsilon$. This is equivalent to the first factor being $o(1)$ or $\sum k^2\alpha_k = o(n^{\frac{1}{2}})$. As α_k is a monotonically decreasing sequence this reduces to $\alpha_k = o(k^{-\frac{5}{2}})$.

Now we examine the first term in (3.33) which depends on

$$\textstyle\sum_{k=1}^{n} \min(2^{-i}, \alpha_k) \leqq \sum_{k=1}^{2^{.4i}} 2^{-i} + \sum_{k>2^{.4i}} \alpha_k = 2^{-.6i} + o(2^{-.6i}) \tag{3.35}$$

which is $O(2^{-.6i})$, where we have used the assumption that $\alpha_k = o(k^{-\frac{5}{2}})$. Hence

$$\textstyle\sum_{i=l}^{R_n(\varepsilon)} [2^i(\sum_{k=1}^{n} \min(2^{-i}, \alpha_k)^2]^{\frac{1}{5}} \leqq M \sum_{i=l}^{\infty} 2^{-.04i} \tag{3.36}$$

which can be made less than η if l is chosen sufficiently large.

As before a much better result can be obtained when the original rv's $\{X_i\}$ are a Gaussian process. We require

LEMMA 3.5. *For every $M > 0$, $\delta > 0$, $m > 0$ there exists a number $c > 0$ and polynomials S, R, R independent of M, such that whenever $X_1, X_2, Y_1, \cdots, Y_m$ are jointly normal with means* 0, *variances* 1, *$E(X_1X_2) = \rho$, $E(X_i Y_j) = \tau_{ij}$ and the covariance matrix of the (X, Y) vector has determinant exceeding δ, $0 \leqq a < b \leqq a + M$, $B_i = I_{(a,b)}(X_i)$, $\alpha = E(B_i)$, then for all y, $a \leqq y_i \leqq b$ for $i = 1, \cdots, m$,*

$$\begin{aligned} &|E(B_1 - \alpha)(B_2 - \alpha) \mid Y = y) - E((B_1 - \alpha)(B_2 - \alpha))| \\ &\qquad \leqq \alpha^c S(b)(|\rho| \textstyle\sum_{i=1}^{2} \sum_{j=1}^{m} |\tau_{ij}| + \prod_{i=1}^{2} \sum_{j=1}^{m} |\tau_{ij}|) \\ &\qquad\qquad \textit{if} \quad \min_i \textstyle\sum_j |\tau_{ij}| < \dfrac{\delta}{2} \quad \textit{or} \quad m = 1\,, \end{aligned} \tag{3.37}$$

$$(3.38)\quad \begin{aligned} &|E((B_i-\alpha)(B_j-\alpha)\mid Y=y)-E((B_i-\alpha)(B_j-\alpha))| \\ &\qquad \leqq M^2R(b)(|\rho|\textstyle\sum_{i=1}^{2}\sum_{j=1}^{m}|\tau_{ij}|+\prod_{i=1}^{2}\sum_{j=1}^{m}|\tau_{ij}|) \qquad \textit{otherwise.}\end{aligned}$$

PROOF. The conditional distribution of X_1 and X_2 given $Y_1, \cdots, Y_m$ is normal with means $\tau_i'Q^{-1}Y$ and covariance matrix

$$(3.39)\quad \Sigma(\rho,\tau_1,\tau_2)=\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}\begin{pmatrix}\tau_1'\\ \tau_2'\end{pmatrix}Q^{-1}(\tau_1\tau_2)\,,$$

where Q is the covariance matrix of the Y's. By the uniform nonsingularity, all elements of Q^{-1} are bounded by $1/\delta$ and all elements of $\Sigma^{-1}(\rho, \tau_1, \tau_2)$ as well as the determinant of that matrix are bounded by $1/\delta$.

Let

$$(3.40)\quad \begin{aligned} &P(\rho,\tau_1,\tau_2,y)\\ &\quad=\int_a^b\int_a^b\frac{1}{2\pi|\Sigma(\rho,\tau_1,\tau_2)|^{\frac12}}\\ &\qquad\times\exp\{-\tfrac12(x-\tau'Q^{-1}y)'\Sigma^{-1}(\rho,\tau_1,\tau_2)(x-\tau'Q^{-1}y)\}\,dx_1\,dx_2\,.\end{aligned}$$

Then expanding,

$$(3.41)\quad \begin{aligned} &E((B_1-\alpha)(B_2-\alpha)\mid Y=y)-E((B_1-\alpha)(B_2-\alpha))\\ &\quad=P(\rho,\tau_1,\tau_2,y)-P(0,0,\tau_2,y)-P(0,\tau_1,0,y)\\ &\qquad-P(\rho,0,0,y)+2P(0,0,0,y)\,,\end{aligned}$$

since $\alpha=\int_a^b(2\pi)^{-\frac12}\exp(-\frac12x^2)\,dx$.

Now the right-hand expression in (3.41) is

$$(3.42)\quad \begin{aligned} &\int_0^1[P_\rho(t\rho,t\tau_1,t\tau_2,y)-P_\rho(t\rho,0,0,y)]\rho\\ &\quad+[P_{\tau_1}(t\rho,t\tau_1,t\tau_2,y)-P_{\tau_1}(0,t\tau_1,0,y)]\tau_1\\ &\quad+[P_{\tau_2}(t\rho,t\tau_1,t\tau_2,y)-P_{\tau_2}(0,0,t\tau_2,y)]\tau_2\,dt\\ &\qquad=\int_0^1\int_0^t[P_{\rho\tau_1}(t\rho,s\tau_1,s\tau_2,y)+P_{\rho\tau_1}(s\rho,t\tau_1,s\tau_2,y)]\rho\tau_1\\ &\qquad\quad+[P_{\rho\tau_2}(t\rho,s\tau_1,s\tau_2,y)+P_{\rho\tau_2}(s\rho,s\tau_1,t\tau_2,y)]\rho\tau_2\\ &\qquad\quad+[P_{\tau_1\tau_2}(s\rho,t\tau_1,s\tau_2,y)+P_{\tau_1\tau_2}(s\rho,s\tau_1,t\tau_2,y)]\tau_1\tau_2\,ds\,dt\,.\end{aligned}$$

Since τ_1 and τ_2 are vectors, by $P_{\tau_1}\tau_1$ we mean $\sum_i P_{\tau_{1i}}\tau_{1i}$, etc.

Hence the expression is bounded by

$$(3.43)\quad \begin{aligned} &|\rho|\left(\sum|\tau_{1i}|\sup\left|\frac{\partial^2P(r,t_1,t_2)}{\partial r\,\partial t_{1i}}\right|+\sum|\tau_{2i}|\sup\left|\frac{\partial^2P(r,t_1,t_2)}{\partial r\,\partial t_{2i}}\right|\right)\\ &\qquad\qquad+\sum\sum|\tau_{1i}||\tau_{2j}|\sup\left|\frac{\partial^2P(r,t_1,t_2)}{\partial t_{1i}\,\partial t_{2j}}\right|\,,\end{aligned}$$

where the sup is over the smallest convex set generated by the covariance matrices. An elementary algebraic argument shows that all matrices involved in the argument are uniformly non-singular, so that we can differentiate under the integral sign and obtain the result that all second derivatives of the integrand of (3.40) are bounded by a polynomial $R(b)$, which proves (3.38).

If $\min_i \sum_j |\tau_{ij}| < 1/2\delta$ or $m = 1$, the sum of the regression coefficients will not be near 1 for both variables, and hence for all x, y with $a \leqq x_i \leqq b$, $a \leqq y_j \leqq b$

$$(3.44) \qquad (x - \tau' Q^{-1} y)' \Sigma(\rho, \tau_1, \tau_2)^{-1} (x - \tau' Q^{-1} y) \geqq rb^2 - s .$$

We may assume $r \leqq 2$. Now

$$(3.45) \qquad \begin{aligned} \textstyle\int_a^b \int_a^b e^{-\frac{1}{2} rb^2} dx_1\, dx_2 &= (\textstyle\int_a^b e^{-\frac{1}{2} rb^2} dx)^2 \\ &\leqq (\textstyle\int_a^b e^{-\frac{1}{2} rx^2} dx)^2 \leqq (\textstyle\int_a^b e^{-\frac{1}{2} x^2} dx)^r (b - a)^{2-r} \end{aligned}$$

from which (3.37) follows with $c = r$, $S = 2\pi M^{2-r} e^s R$.

Now let us verify the Prokhorov continuity condition for stationary Gaussian process with $\sum |\rho_i| < \infty$. Instead of using the usual probability integral transform, let us transform to density $6t(1 - t)$, i.e., $P(Y_i \leqq t) = 3t^2 - 2t^3$. We now have $n^{\frac{1}{2}}(\{F_{nY}(t + u) - F_Y(t + u)\} - \{F_{nY}(t) - F_Y(t)\}) > -6n^{\frac{1}{2}} u$, which is not essentially different from the case of the probability integral transform.

The proof of Rubin's lemma involves showing that

$$(3.46) \qquad \textstyle\sum_{i=k}^{R_n} \sum_{j=0}^{2i} P\left(\left|X_n\left(\frac{j+1}{2^i}\right) - X_n\left(\frac{j}{2^i}\right)\right| > \zeta_i\right) \to 0 ,$$

where $\sum \zeta_i < \varepsilon$. We shall bound all except the extreme terms for each i by a 4th moment Markov inequality, and the extreme ones by the Tchebychev inequality. We can even use a fixed number instead of ζ_i for the extreme terms. Now

$$(3.47) \qquad E[(X_n(2^{-i}) - X_n(0))^2] = \frac{1}{n} \sum\sum E(A_i A_j) ,$$

where $A_i = I_{(-\infty,\xi)}(X_i) - \alpha$, $\alpha = \int_{-\infty}^{\xi} (2\pi)^{-\frac{1}{2}} \exp(-\frac{1}{2}x^2)\, dx = 3 \cdot 2^{-2i} - 2^{1-3i}$. Now $E(A_i^2) = \alpha(1 - \alpha)$ and by Sarmonov's lemma [10], $|E(A_i A_j)| \leqq \alpha |\rho_{i-j}|$. Thus

$$(3.48) \qquad E[(X_n(2^{-i}) - X_n(0))^2] \leqq 3(1 + 2 \textstyle\sum_{i=1}^{\infty} |\rho_i|) 2^{-2i} ,$$

and hence $P(|X_n(2^{-i}) - X_n(0)| > \varepsilon) \leqq K\varepsilon^{-2} 2^{-2i}$, which is more than adequate for our purposes.

For the remaining terms, we proceed as before to verify Rubin's lemma. We will illustrate the use of Lemma 3.5 only on the fourth-order terms; the others are even easier.

Let $B_i = I_{t,t+u}(Y_i)$, $\alpha = E(B_i)$, $A_i = B_i - \alpha$, and $t \geqq \frac{1}{2}$. The terms for $t < \frac{1}{2}$ give an equal contribution. Let $i < j < k < l$ and consider

$$(3.49) \qquad \begin{aligned} &E(A_i A_j A_k A_l) - E(A_i A_j) E(A_k A_l) \\ &\quad = E([E(A_i A_j \mid X_k X_l) - E(A_i A_j)] B_k B_l) \\ &\qquad - \alpha E([E(A_i A_j \mid X_k) - E(A_i A_j)] B_k) \\ &\qquad - \alpha E([E(A_i A_j \mid X_l) - E(A_i A_j)] B_l) . \end{aligned}$$

If $|\rho_{ik}| + |\rho_{il}| < \delta/2$ or $|\rho_{jk}| + |\rho_{jl}| < \delta/2$, we can apply (3.37) to all terms on the right side of (3.49), obtaining

$$\begin{aligned}|E(A_iA_jA_kA_l) - E(A_iA_j)E(A_kA_l)| \\ \leqq 3\alpha^{1+c}Q(b)(|\rho_{ij}|(|\rho_{ik}| + |\rho_{il}| + |\rho_{jk}| + |\rho_{jl}|) \\ + (|\rho_{ik}| + |\rho_{il}|)(|\rho_{jk}| + |\rho_{jl}|))\,.\end{aligned} \tag{3.50}$$

If $|\rho_{ik}| + |\rho_{il}| \geqq \delta/2$ and $|\rho_{jk}| + |\rho_{jl}| \geqq \delta/2$, we just use $6M^2\alpha Q(b)(|\rho_{ik}| + |\rho_{il}|) \times (|\rho_{jk}| + |\rho_{jl}|)$ as a bound. Furthermore, $\alpha < 6(1-t)u$. Also, since the normal tail drops off rapidly, $(1-t)^{1+c}Q(b)$ and $(1-t)R(b)$ are uniformly bounded for all terms. Hence, except for a multiple of n terms, if $i < j < k < l$

$$\begin{aligned}|E(A_iA_jA_kA_l)| &\leqq |E(A_iA_j)||E(A_kA_l)| \\ &\quad + Ku^{1+c}(|\rho_{ij}|(|\rho_{ik}| + |\rho_{il}| + |\rho_{jk}| + |\rho_{jl}|) \\ &\quad + (|\rho_{ik}| + |\rho_{il}|)(|\rho_{jk}| + |\rho_{jl}|)) \\ &\leqq u^2|\rho_{ij}||\rho_{kl}| + Ku^{1+c}(|\rho_{ij}|(|\rho_{ik}| + |\rho_{jk}| + |\rho_{il}| + |\rho_{jl}|) \\ &\quad + (|\rho_{ik}| + |\rho_{il}|)(|\rho_{jk}| + |\rho_{jl}|))\,.\end{aligned} \tag{3.51}$$

The sum of these terms is bounded by $n^2M(u^2 + u^{1+c})$ since $\sum |\rho_{i-j}| < \infty$. The remaining terms are bounded by $K\alpha$, or $6KM^2(1-t)R(b)u$, so that their sum is bounded by Cnu. We obtain similar results for the other terms. Consequently, we can apply Rubin's lemma with $\phi_n(u) = M(u^2 + u^{1+c}) + Du/n$.

REFERECES

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

[2] BLUM, J. R. and ROSENBLATT, M. (1956). A class of stationary processes and a central limit theorem. *Proc. Nat. Acad. Sci.* **43** 412–423.

[3] DOOB, J. L. (1953). *Stochastic Processes.* Wiley, New York.

[4] GASTWIRTH, J. L. and RUBIN, H. (1975). The behavior of robust estimators on dependent data. To appear in *Ann. Statist.*

[5] GASTWIRTH, J. L. and WOLFF, S. S. (1965). A characterization of the Laplace distribution. Department of Statistics Report 28. The Johns Hopkins Univ.

[6] IBRAGIMOV, I. A. (1962). Some limit theorems for stationary processes. *Theor. Probability Appl.* **7** 349–382.

[7] MANN, H. B. and WALD, A. (1943). On stochastic limit and order relations. *Ann. Math. Statist.* **14** 217–226.

[8] PROHOROV, YU. V. (1956). Convergence of random processes and limit theorems in probability theory. *Theor. Probability Appl.* **1** 157–214.

[9] ROSENBLATT, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci.* **42** 43–47.

[10] ROZANOV, YU. A. (1967). *Stationary Random Processes.* Holden-Day, San Francisco.

[11] RUBIN, H. (1968). A useful device for continuity theorems, (Abstract 17). *Ann. Math. Statist.* **39** 1090.

[12] SERFLING, R. J. (1968). Contributions to central limit theory for dependent variables. *Ann. Math. Statist.* **39** 1158–1175.

[13] SUN, T. C. (1963). A central limit theorem for non-linear functions of a normal stationary process. *J. Math. Mech.* **12** 945–977.

DEPARTMENT OF STATISTICS
GEORGE WASHINGTON UNIVERSITY
2201 G ST. N.W.
WASHINGTON D.C. 20006

DEPARTMENT OF STATISTICS
PURDUE UNIVERSITY
W. LAFAYETTE, INDIANA 47907