*The Annals of Mathematical Statistics*
1972, Vol. 43, No. 6, 2058-2066

# ON LIMIT THEOREMS FOR QUADRATIC FUNCTIONS OF DISCRETE TIME SERIES

BY E. J. HANNAN AND C. C. HEYDE

*Australian National University*

In this paper it is shown how martingale theorems can be used to appreciably widen the scope of classical inferential results concerning autocorrelations in time series analysis. The object of study is a process which is basically the second-order stationary purely non-deterministic process and contains, in particular, the mixed autoregressive and moving average process. We obtain a strong law and a central limit theorem for the autocorrelations of this process under very general conditions. These results show in particular that, subject to mild regularity conditions, the classical theory of inference for the process in question goes through if the best linear predictor is the best predictor (both in the least squares sense).

**1. Introduction.** A great deal of time series analysis is based upon quadratic functions of the data. In particular, many inferential results relate to theorems concerning the autocorrelations

$$r(j) = \frac{\sum_{n=1}^{N-j} \{x(n) - \bar{x}\}\{x(n+j) - \bar{x}\}}{\sum_{n=1}^{n} \{x(n) - \bar{x}\}^2}, \qquad j \geqq 0, \tag{1}$$
$$r(-j) = r(j),$$

$x(1), x(2), \cdots, x(N)$ being a sample of $N$ consecutive observations on some process $\{x(n)\}$. It is well known that, under certain conditions on the process $\{x(n)\}$, a strong law of large numbers and a central limit theorem hold for $r(j)$ (see, for example, Hannan [6], Chapter IV, VI). In this paper it is our object to show, using limit theorems for martingales, that the scope of the classical inferential theory can be appreciably widened in a natural way.

We shall be concerned with a process of the form

$$x(n) - \mu = \sum_{j=0}^{\infty} \alpha(j)\varepsilon(n-j), \qquad \sum_{j=0}^{\infty} \alpha^2(j) < \infty, \qquad \alpha(0) = 1; \tag{2}$$
$$E\varepsilon(n) = 0, \qquad E\{\varepsilon(m)\varepsilon(n)\} = 0, \qquad m \neq n.$$

If $x(n) - \mu$ is a second-order stationary, purely non-deterministic, process ([6], Chapter III) then it may be represented in this form with the $\varepsilon(n)$ as the linear prediction errors, having variance $\sigma^2 > 0$. As is well known, there will be many representations of such a stationary process in the form (2) but for only one of these will the $\varepsilon(n)$ be the prediction errors. However, our results extend beyond the stationary case so that we do not assume stationarity for the $\{\varepsilon(n)\}$ but only that (2) holds for $n \geqq 0$, together with other conditions to be discussed shortly. A process of the kind (2) arises from a wide variety of contexts; for example from a mixed autoregressive and moving averge process ([6], Chapter I) and as

Received August 7, 1971; revised April 7, 1972.

the response of a physically realizable filter to an uncorrelated sequence (Gikhman and Skorokhod [5], Chapter 5).

Now the classical theory of inference for the process (2) usually requires that the $\varepsilon(n)$ be independent and identically distributed with zero mean and variance $\sigma^2$ (which we paraphrase as i.i.d. $(0, \sigma^2)$). The essential feature of this paper is that, subject to some reasonable additional conditions, the classical theory goes through if the independence assumption is replaced by the weaker condition

$$E(\varepsilon(n) \mid \mathscr{F}_{n-1}) = 0 \quad \text{a.s.}, \qquad \text{all} \quad n, \tag{3}$$

where $\mathscr{F}_n$ is the $\sigma$-field generated by the $\varepsilon(m)$, $m \leqq n$. This requirement has a simple and natural interpretation in the case where $\{x(n)\}$ is stationary and thus purely non-deterministic and the $\varepsilon(n)$ are the linear prediction errors, for then $\mathscr{F}_n$ is also the $\sigma$-field generated by the $x(m)$, $m \leqq n$ so that, because of (3),

$$\varepsilon(n) = x(n) - E(x(n) \mid \mathscr{F}_{n-1}) . \tag{4}$$

To see this, write $\mathscr{G}_n$ for the $\sigma$-field generated by $x(m)$, $m \leqq n$. Clearly $\mathscr{F}_n \supseteqq \mathscr{G}_n$ and, when the $\varepsilon$'s are the prediction errors

$$\varepsilon(n) = x(n) - E(x(n) \mid \mathscr{G}_{n-1})$$

which is $\mathscr{G}_n$ measurable. Thus $\mathscr{G}_n \equiv \mathscr{F}_n$. Then, $E(x(n) \mid \mathscr{F}_{n-1})$ is the best linear predictor and the best linear predictor is the best predictor (both in the least squares sense). Conversely, if this is so, (4) must hold and hence (3). Thus (3) is equivalent to the condition that the best predictor is the best linear predictor, both in the least squares sense. In the stationary case our additional conditions are, for example, the regularity condition (7), below, together with the requirement that $E(\varepsilon^2(n) \mid \mathscr{F}_{n-1}) = \sigma^2$ a.s. Our results give that, subject to the mild regularity condition (7), the classical theory of inference for (2) goes through when the $\varepsilon(n)$ are the prediction errors provided the best linear predictor is the best predictor and the prediction variance, given the past, is a constant.

**2. Strong law for autocorrelations.** Here we consider the process (2) where $\varepsilon(n)$ satisfy the condition (3). We shall not require stationarity of $\{\varepsilon(n)\}$ but instead the condition

$$\lim_{N \to \infty} N^{-1} \sum_{n=1}^{N} E(\varepsilon^2(n) \mid \mathscr{F}_{n-1}) = \sigma^2 > 0 \quad \text{a.s.} \tag{5}$$

and the condition that there exists a random variable $X$ with $EX^2 < \infty$ such that

$$P(|\varepsilon(n)| > u) \leqq cP(|X|) > u) \tag{6}$$

for some $0 < c < \infty$ and all $n$, all $u \geqq 0$. If $x(n)$ is stationary we modify (5) to

$$E(\varepsilon^2(n) \mid \mathscr{F}_{n-1}) = \sigma^2 > 0 \quad \text{a.s.} \tag{5$'$}$$

and (6) is redundant. We shall, in Section 3, make use of the condition

$$\sum_{j=1}^{\infty} j^{\frac{1}{2}} \alpha(j)^2 < \infty . \tag{7}$$

Define

$$c(j) = c(-j) = N^{-1} \sum_{n=1}^{N-j} \{x(n) - \bar{x}\}\{x(n+j) - \bar{x}\} , \qquad j \geqq 0 ,$$

where $\bar{x}$ is the sample mean of the $x(n)$, $n = 1, \cdots, N$. If $x(n)$ is stationary its spectral density is

$$f(\lambda) = \frac{\sigma^2}{2\pi} |\sum_0^\infty \alpha(j) e^{ij\lambda}|^2$$

and the autocovariances $\gamma(j)$ satisfy

$$\gamma(j) = \int_{-\pi}^{\pi} e^{ij\lambda} f(\lambda)\, d\lambda = \sigma^2 \sum_{u=0}^\infty \alpha(u)\alpha(u+j) .$$

However we may define $f(\lambda)$ and $\gamma(j)$ by these formulae whether or not $x(n)$ is stationary. We now have the following theorem

THEOREM 1. *If* (3), (6) *and* $\sum |\alpha(k)| < \infty$ *hold,* $\bar{x}$ *converges* a.s. *to* $\mu$ *and if* (5) *holds also, then* $c(j)$ *converges in probabilty to* $\gamma(j)$. *If* $x(n)$ *is stationary,* $\bar{x}$ *converges* a.s. *to* $\mu$ *and if* (3), (5)′ *hold* $c(j)$ *converges* a.s. *to* $\gamma(j)$.

PROOF. If $\varepsilon(n)$ satisfies (3), (6) and $\sum |\alpha(k)| < \infty$, then it is easily seen that $\bar{x}$ has a variance which is $O(N^{-1})$ as $N \to \infty$. It consequently follows from the proof given in Doob [4], Theorem X 6.2, that $\bar{x}$ converges a.s. to $\mu$. It is then clear that $c(j)$ has the same a.s. behaviour as

$$\begin{aligned} c^*(j) &= N^{-1} \sum_{n=1}^{N-j} (x(n) - \mu)(x(n+j) - \mu) \\ &= \sum_{u=0}^\infty \sum_{v=0}^\infty \alpha(u)\alpha(v) N^{-1} \sum_{n=1}^{N-j} \varepsilon(n-u)\varepsilon(n+j-v) . \end{aligned}$$

Now

$$E|N^{-1} \sum_{n=1}^{N-j} \varepsilon(n-u)\varepsilon(n+j-v)| < K < \infty$$

by virtue of (6) and if $\sum |\alpha(k)| < \infty$ then

$$\lim_{p\to\infty} E|\sum_{u=p+1}^\infty \sum_{v=0}^\infty \alpha(u)\alpha(v) N^{-1} \sum_{n=1}^{N-j} \varepsilon(n-u)\varepsilon(n+j-v)| = 0 . \tag{8}$$

The same is true if in the left term in (8) the first two sums are over $0 \leqq u \leqq p$, $p < v < \infty$. On the other hand,

$$N^{-1} \sum_{n=1}^{N} E\{\varepsilon(n)\varepsilon(n-k) \mid \mathscr{F}_{n-1}\} = 0 \quad \text{a.s.} , \qquad k > 0 ,$$

and it follows from (6) and a law of large numbers for martingales due to Heyde (Theorem 1 of Heyde and Seneta [8]) that $N^{-1} \sum_{n=1}^{N} \varepsilon(n)\varepsilon(n-k)$, $k > 0$, and hence $N^{-1} \sum_{n=1}^{N-j} \varepsilon(n-u)\varepsilon(n+j-v)$, $u \neq v - j$ converges in probability to zero. (In order to obtain the uniform bound on the distribution of $\varepsilon(n)\varepsilon(n-k)$ required to justify the application of Theorem 1 of [8] we note that

$$\begin{aligned} P(|\varepsilon(n)\varepsilon(n-k)| > u) &\leqq P(\varepsilon^2(n) + \varepsilon^2(n-k) > 2u) \\ &\leqq P(\varepsilon^2(n) > u) + P(\varepsilon^2(n-k) > u) \\ &\leqq 2cP(X^2 > u) \end{aligned} \tag{9}$$

using (6).) Furthermore, by the same theorem together with (5), $N^{-1} \sum_{n=1}^{N-j} \varepsilon^2(n)$ converges in probability to $\sigma^2$. Thus

$$\sum_{u=0}^{p} \sum_{v=0}^{p} \alpha(u)\alpha(v) N^{-1} \sum_{n=1}^{N-j} \varepsilon(n-u)\varepsilon(n+j-v) \to \sigma^2 \sum{}' \alpha(u)\alpha(u+j)$$

in probability as $N \to \infty$ where $\sum'$ is a sum over $0 \leqq u$, $u + j \leqq p$. It follows from this together with (8) and Markov's inequality that $c^*(j)$, and hence $c(j)$, converges in probability to $\gamma(j)$.

If $x(n)$ is stationary with a.c. spectrum it is well known that $\bar{x}$ converges a.s. to $\mu$. In place of $c^*(j)$ consider

$$\tilde{c}(j) = c^*(j) - \sum_{u=0}^{\infty} \alpha(u)\alpha(u+j)N^{-1} \sum_{n=1}^{N-j} \varepsilon^2(n-u) .$$

Assuming $x(n)$ stationary and (3), (5)′ we may show that the mean of $\tilde{c}(j)$ is zero and its variance converges to zero. The proof of the first is obvious and we prove the second, for simplicity, in case $j = 0$. The variance is

$$\sum\sum_{p\neq q,=0}^{\infty} \alpha(p)\alpha(q) \sum\sum_{r\neq s,=0}^{\infty} \alpha(r)\alpha(s) \times [N^{-2} \sum\sum_{m,n=1}^{N} E\{\varepsilon(m-p)\varepsilon(m-q)\varepsilon(n-r)\varepsilon(n-s)\}] .$$

We evaluate the expectation using (3) and (5)′. The only contribution comes when $m - p = n - r$ and $s = q - p + r$ or $m - p = n - s$ and $s = p - q + r$. Both sets of identifications give the same result and we take the first. After evaluating the expectation let us add back

$$N^{-1}\sigma^4 \sum\sum_{|p-r|<N,=0}^{\infty} \alpha(p)^2\alpha(r)^2 \left(1 - \frac{|p-r|}{N}\right)$$

which clearly converges to zero as $N$ increases. Then we obtain

$$N^{-1}\sigma^4 \sum_{j=-N+1}^{N-1} \left(1 - \frac{|j|}{N}\right) \{\sum_0^{\infty} \alpha(p)\alpha(p+|j|)\}^2 .$$

This is $N^{-1}$ by the Cesàro sum of the Fourier series, evaluated at the origin, of the convolution of $f(\lambda)$ with itself. It thus converges to zero. Thus $\tilde{c}(0)$, and in the same way $\tilde{c}(j)$, converges in probability to zero. However

$$\sum_{u=0}^{\infty} \alpha(u)\alpha(u+j)N^{-1} \sum_{n=1}^{N-j} \varepsilon^2(n-u)$$

converges in probability to $\gamma(j)$ by the same kind of argument as was used earlier in the proof (the convergence of $\sum |\alpha(u)|$ not now being needed). Thus $c^*(j)$ and hence $c(j)$ converges in probability to $\gamma(j)$ and since, by the ergodic theorem, $c(j)$ converges almost surely it must converge almost surely to $\gamma(j)$. This completes the proof.

**3. Central limit theorems for autocorrelations.** From Theorem 1 we note that, depending on the conditions, $r(j) = c(j)/c(0)$ converges either in probability or almost surely to $\rho(j) = \gamma(j)/\gamma(0)$. Here we offer two theorems, the first of which is for the non-stationary case.

THEOREM 2. *Suppose* (7) *holds,* $\sum |\alpha(j)| < \infty$ *and* $\{\varepsilon(n)\}$ *is a stochastic sequence, with* $E\varepsilon^2(n) = \sigma^2$, *all* $n$, *satisfying* (3), (5), *and* (6) *with bounding random variable* $X$ *having finite fourth moment. Suppose also that* $E\{\varepsilon^2(n)\varepsilon(n-r)\varepsilon(n-s)\} = \sigma^4\tau_{rs}$ *is finite and uniformly bounded for every* $n$, $r \geqq 1$, $s \geqq 1$, *and*

$$n^{-1} \sum_{t=1}^{n} \varepsilon(t-r)\varepsilon(t-s)E(\varepsilon^2(t) \mid \mathscr{F}_{t-1}) \rightarrow_{\text{a.s.}} \sigma^4\tau_{rs} \tag{10}$$

*as* $n \rightarrow \infty$ *for any* $r \geqq 1$, $s \geqq 1$. *Then, the joint distribution of* $N^{\frac{1}{2}}(r(j) - \rho(j))$,

$1 \leqq j \leqq s$, *converges to the s-variate normal distribution with zero mean and non-singular covariance matrix* $W = [w_{ij}]$ *where*

$$w_{ij} = \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} \tau_{rs}\{\rho(r+i) + \rho(r-i) - 2\rho(r)\rho(i)\} \times \{\rho(r+j) + \rho(r-j) - 2\rho(r)\rho(j)\}.$$

PROOF. This theorem may be established by adapting the proof of the theorem of Anderson and Walker [1], where the case $\{\varepsilon(n)\}$ i.i.d. $(0, \sigma^2)$ is considered. The only parts which cannot immediately be adapted involve the replacement of their condition $\sum j\alpha^2(j) < \infty$ by our (7), their application of a central limit theorem of Diananda in the proof of their Lemma 1 and their demontstration that $N^{-1} \sum_{n=1}^{N} \{x(n) - \mu\}^2$ converges in probability to $E\{x(1) - \mu\}^2$ as $N \to \infty$. This last result, however, is covered by our Theorem 1.

The replacement of $\sum j\alpha^2(j) < \infty$ by our (7) is easy to justify. The only point at issue concerns the proof of their Lemma 3 where it is necessary to note that the bound on the expectation of their (2.23) can be suitably sharpened for $|i| > n$. Full details are given in the proof that we give for Theorem 3 below.

It then remains to consider the central limit part. What is required is just that, for any sequence of constants $c_1, \cdots, c_m$, $n^{-\frac{1}{2}} \sum_{r=1}^{m} c_r \sum_{t=1}^{n} \varepsilon(t)\varepsilon(t+r)$ converges in distribution to a certain normal law. In order to obtain this under our conditions we first note that it suffices to establish the convergence result for $n^{-\frac{1}{2}} \sum_{r=1}^{m} c_r \sum_{t=1}^{n} \varepsilon(t)\varepsilon(t-r)$ which differs from the former in a fixed finite number of terms (and hence the difference goes in probability to zero as $n \to \infty$). We shall obtain this last result with the aid of a central limit theorem for martingales due to Brown [3].

Define $X_t = \sum_{r=1}^{m} c_r \varepsilon(t)\varepsilon(t-r)$, noting that $\{S_n = \sum_{t=1}^{n} X_t, \mathscr{F}_n, n \geqq 1\}$ is a martingale. Let

$$V_n^2 = \sum_{t=1}^{n} E(X_t^2 \mid \mathscr{F}_{t-1}) = \sum_{t=1}^{n} \sum_{r=1}^{m} \sum_{s=1}^{m} c_r c_s \varepsilon(t-r)\varepsilon(t-s)E(\varepsilon^2(t) \mid \mathscr{F}_{t-1})$$

and

$$s_n^2 = EV_n^2 = \sum_{t=1}^{n} EX_t^2 .$$

In order to apply Theorem 2 of [3] we need to show that

$$\text{(i)} \quad s_n^{-2} V_n^2 \to_p 1 \qquad \text{and} \qquad \text{(ii)} \quad s_n^{-2} \sum_{t=1}^{n} E(X_t^2 I(|X_t| \geqq \varepsilon s_n)) \to 0$$

for any $\varepsilon > 0$, $I(\cdot)$ being the indicator function.

The condition (10) clearly ensures that (i) holds upon noting that

$$s_n^2 = n\sigma^4 \sum_{r=1}^{m} \sum_{s=1}^{m} c_r c_s \tau_{rs} . \tag{11}$$

To obtain (ii) we first note that for any $u \geqq 0$,

$$\begin{aligned} P(|X_n| > u) &\leqq P(\textstyle\sum_{r=1}^{m} |c_r \varepsilon(n)\varepsilon(n-r)| > u) \\ &\leqq P(\textstyle\bigcup_{r=1}^{m} \{|c_r \varepsilon(n)\varepsilon(n-r)| > u/m\}) \\ &\leqq \textstyle\sum_{r=1}^{m} P(|\varepsilon(n)\varepsilon(n-r)| > u/m|c_r|) \qquad \text{(using (9))} \\ &\leqq 2c \textstyle\sum_{r=1}^{m} P(X^2 < u/m|c_r|) \\ &\leqq 2cmP(X^2 < u/mc^*) \end{aligned} \tag{12}$$

where $c^* = \max_{1 \leq r \leq m} |c_r|$. Thus, using integration by parts and (12),

$$EX_n^2 I(|X_n| \geqq \varepsilon s_n) \leqq 2 \int_{\varepsilon s_n}^\infty x P(|X_n| > x)\, dx$$
$$\leqq 2cm \int_{\varepsilon s_n}^\infty x P(X^2 > x/mc^*)\, dx \to 0$$

as $n \to \infty$ and (ii) follows. Brown's theorem then gives that $s_n^{-1} \sum_{t=1}^n X_t$ converges in distribution to $N(0, 1)$. That is, $n^{-\frac{1}{2}} \sum_{r=1}^m c_r \sum_{t=1}^n \varepsilon(t)\varepsilon(t-r)$ converges in distribution to $N(0, \sigma^4 \sum_{r=1}^m \sum_{s=1}^m c_r c_s \tau_{rs})$. The proof of Theorem 2 is then completed along the lines of Anderson and Walker [1]. Our second theorem relates to the stationary case.

THEOREM 3. *Let $x(n)$ be stationary and satisfy* (3), (5)′ *and* (7). *Then the joint distribution of $N^{\frac{1}{2}}(r(j) - \rho(j))$, $1 \leqq j \leqq s$, converges to the s-variate multivariate normal distribution with zero mean vector and covariance matrix $W = [w_{ij}]$ where*

$$w_{ij} = \sum_{r=1}^\infty \{\rho(r+i) + \rho(r-i) - 2\rho(r)\rho(i)\}$$
$$\times \{\rho(r+j) + \rho(r-j) - 2\rho(r)\rho(j)\}$$
$$= \sum_{r=-\infty}^\infty \{\rho(r)\rho(r+i-j) + \rho(r)\rho(r+i+j) + 2\rho^2(r)\rho(i)\rho(j)$$
$$- 2\rho(r)\rho(i)\rho(r+j) - 2\rho(r)\rho(j)\rho(r+i)\}\,.$$

This proof also follows [1] and improves the theorem of that paper in three ways. Firstly the independence of the $\varepsilon(n)$ is replaced by (3), (5)′. Secondly $\sum j\alpha^2(j) < \infty$ is replaced by $\sum j^{\frac{1}{2}}\alpha^2(j) < \infty$. Thirdly $\sum |\alpha(j)| < \infty$ is avoided. Because we use the conditions (3), (5)′ we must confine ourselves to a one-sided sum, $\sum_0^\infty \alpha(j)\varepsilon(n-j)$ (i.e., to the purely non-deterministic case) while [1] treats the case of a two-sided moving average. The second and third relaxations of the conditions of the theorem in [1] may be made under the other conditions of that theorem.

PROOF. Let us take the case where $\mu$ is known to be zero since mean correction makes no difference to the truth of the theorem. Instead of $N^{\frac{1}{2}}\{r((j) - \rho(j)\}$ we may consider the limit distribution of $N^{\frac{1}{2}}\{c^*(j) - \rho(j)c^*(0)\}/c^*(0)$ and since $c^*(0)$ converges a.s. to $\gamma(0)$, by Theorem 1, we thus consider

$$N^{\frac{1}{2}}\{c^*(j) - \rho(j)c^*(0)\}\,. \tag{13}$$

We first show that we may omit all terms involving an $\varepsilon^2(n)$. These terms are (see [1] page 1301)

$$N^{-\frac{1}{2}}\{\sum_0^\infty \alpha(k)\alpha(k+j)T_{N,k}^{(j)} - \rho(j) \sum_0^\infty \alpha^2(k)T_{N,k}^{(0)}\} \tag{14}$$

where

$$T_{N,k}^{(j)} = \sum_{1-k}^{N-j-k} \varepsilon^2(n) - \sum_1^N \varepsilon^2(n)$$

and

$$E|T_{N,k}^{(j)}| = \sigma^2 \min(2k+j, 2N-j)\,.$$

(In [1] a bound by $(2k+j)\sigma^2$ is given for $E|T_{N,k}^{(j)}|$ and this accounts for the replacement of our $\sum j^{\frac{1}{2}}\alpha^2(j) < \infty$ by $\sum j\alpha^2(j) < \infty$ in [1].) Thus, taking the first term in (14), for example,

$$N^{-\frac{1}{2}}E\{|\sum_0^\infty \alpha(k)\alpha(k+j)T_{N,k}^{(j)}|\}$$
$$\leqq N^{-\frac{1}{2}}\sigma^2 \sum_0^\infty |\alpha(k)\alpha(k+j)| \min(2k+j, 2N-j)$$
$$\leqq \sigma^2[\{N^{-\frac{1}{2}} \sum_0^\infty \alpha^2(k) \min(2k+j, 2N-j)\}^{\frac{1}{2}}$$
$$\times \{N^{-\frac{1}{2}} \sum_0^\infty \alpha^2(k+j) \min(2k+j, 2N-j)\}^{\frac{1}{2}}] .$$

However,

$$N^{-\frac{1}{2}} \sum_0^\infty \alpha^2(k) \min(2k+j, 2N-j)$$
$$\leqq 2 \sum_0^{N-j} \alpha^2(k)k^{\frac{1}{2}}\{k/N\}^{\frac{1}{2}} + jN^{-\frac{1}{2}} \sum_0^{N-j} \alpha^2(k) + 2 \sum_{N-j}^\infty \alpha^2(k)k^{\frac{1}{2}} ,$$

which converges to zero. The same is true of the second factor and the second term in (14) and thus (14) converges in probability to zero.

Let us put $x(n) = x_1(n) + x_2(n)$ where

$$x_2(n) = \sum_{K+1}^\infty \alpha(j)\varepsilon(n-j) .$$

We also put

$$c_{ij}(k) = N^{-1} \sum_1^N x_i(n)x_j(n+k) ; \qquad i, j = 1, 2 .$$

If $\gamma_{ij}(k) = Ec_{ij}(k)$, then

$$\gamma_{ij}(k) = \int_{-\pi}^{\pi} e^{ik\lambda} f_{ij}(\lambda)\, d\lambda$$

and

$$f_{11}(\lambda) = \frac{\sigma^2}{2\pi} |\sum_0^K \alpha(j)e^{ij\lambda}|^2 , \qquad f_{22}(\lambda) = \frac{\sigma^2}{2\pi} |\sum_{K+1}^\infty \alpha(j)e^{ij\lambda}|^2 ,$$

$$f_{12}(\lambda) = \overline{f_{21}(\lambda)} = \frac{\sigma^2}{2\pi} \sum_0^K \alpha(j)e^{ij\lambda} \sum_{K+1}^\infty \alpha(j)e^{-ij\lambda} .$$

All of these functions are square integrable over $[-\pi, \pi]$. In fact to see that, for example, $f(\lambda)$ is square integrable under (7) note that this is equivalent to

$$\sum_{k=0}^\infty \{\sum_{j=1}^\infty \alpha(j)\alpha(j+k)\}^2 < \infty$$

and

$$\sum_{k=0}^\infty \{\sum_{j=1}^\infty j^{\frac{1}{4}}\alpha(j) \cdot j^{-\frac{1}{4}}\alpha(j+k)\}^2 \leqq \sum_{k=0}^\infty \{\sum_{j=1}^\infty j^{\frac{1}{2}}\alpha^2(j)\}\{\sum_{j=1}^\infty j^{-\frac{1}{2}}\alpha^2(j+k)\}$$
$$\leqq \{\sum_{j=1}^\infty j^{\frac{1}{2}}\alpha^2(j)\} \sum_{k=0}^\infty \sum_{j=1}^\infty j^{-\frac{1}{2}}\alpha^2(j+k)$$
$$\leqq \{\sum_{j=1}^\infty j^{\frac{1}{2}}\alpha^2(j)\} \sum_{k=1}^\infty \alpha^2(k) \sum_{j=1}^k j^{-\frac{1}{2}} < \infty$$

under (7).

Now (13) becomes

$$(15) \qquad N^{\frac{1}{2}}[\{c_{11}(j) + c_{22}(j) + c_{12}(j) + c_{21}(j)\} - \rho(j)\{c_{11}(0) + c_{22}(0) + 2c_{12}(0)\}] .$$

We call $c'_{ij}(k)$ the expression $c_{ij}(k)$ with all terms involving an $\varepsilon^2(n)$ omitted. We wish to show that the contribution to the primed form of (15) from the $c'_{ij}(k)$ for $i$, $j$ not both equal to unity has a variance which, for all sufficiently large $N$, may be made arbitrarily small by taking $K$ large. To this end we put

$$N^{\frac{1}{2}}\{c_{ij}(k) - Ec_{ij}(k)\} = N^{\frac{1}{2}}c'_{ij}(k) + N^{\frac{1}{2}}\{c''_{ij}(k) - Ec''_{ij}(k)\} ,$$

wherein $c''_{ij}(k)$ contains all terms in $c_{ij}(k)$ involving an $\varepsilon^2(n)$. If we evaluate the variance of this as if the $\varepsilon(n)$ and $x(n)$ were Gaussian we shall not affect the variance of $N^{\frac{1}{2}}c'_{ij}(k)$. Since on this Gaussian assumption the two terms on the right are uncorrelated, we obtain the variance of the left-hand term as an upper bound to the variance of $N^{\frac{1}{2}}c'_{ij}(k)$. The variance of the left-hand term is, on the Gaussian assumption, ([6] page 210)

$$\sum_{-N+1}^{N-1} (1 - |u|/N)\{\gamma_{ii}(u)\gamma_{jj}(u) + \gamma_{ij}(u+k)\gamma_{ji}(u-k)\}$$

which converges to

$$2\pi \int_{-\pi}^{\pi} \{f_{ii}(\lambda)f_{jj}(\lambda) + |f_{ij}(\lambda)|^2 e^{2ik\lambda}\}\, d\lambda \tag{16}$$

because of the square integrability of the $f_{ij}(\lambda)$ and Parseval's theorem. However, as $K \to \infty$,

$$\int_{-\pi}^{\pi} f_{22}^2(\lambda)\, d\lambda\,, \qquad \int_{-\pi}^{\pi} f_{22}(\lambda)\, d\lambda$$

converge to zero because the Fourier series of a function in $L_p(-\pi, \pi)$, $1 < p < \infty$, converges in the $L_p$ norm to the function ([9] page 50). (In our case the function is $\sum_{K+1}^{\infty} \alpha(j) \exp ij\lambda \in L_4$.) Thus taking $K$ sufficiently large we may, if $i, j$ are not both unity, make (16) arbitrarily small and hence the variance of $N^{\frac{1}{2}}c'_{ij}(k)$, for all sufficiently large $N$, arbitrarily small.

By what is sometimes called Bernstein's lemma ([6] page 242) the theorem will now result if it is shown that the $N^{\frac{1}{2}}\{c'_{11}(j) - \rho(j)c'_{11}(0)\}$ are jointly asymptotically normal with a covariance matrix which converges, as $K$ is increased, to $W$. The proof of the asymptotic normality is the same as that given in the course of proving Theorem 2. Putting

$$\rho'(j) = \sum_{k\leq K} \alpha(k)\alpha(k+j)/\sum_{k\leq K} \alpha^2(k)\,,$$

the covariance of the $N^{\frac{1}{2}}\{c'_{11}(j) - \rho'(j)c'_{11}(0)\}$ converges to $W'$ where $W'$ is obtained from $W$ by replacing $f(\lambda)$ by $f_{11}(\lambda)$. Since $N^{\frac{1}{2}}(\rho(j) - \rho'(j))c'_{11}(0)$ evidently converges in probability to zero and $W'$ converges to $W$ as $K \to \infty$ because of the theorem, quoted above, on $L_p$ convergence, the theorem is proved.

**4. Some applications and extensions.** The most obvious application of the results of Sections 2 and 3 is to the autoregression

$$\sum_{k=0}^{q} \beta(k)\{x(n-k) - \mu\} = \varepsilon(n)\,, \qquad \beta(0) = 1\,, \tag{17}$$

wherein we assume that

$$\sum_{k=0}^{q} \beta(k)z^k \neq 0\,, \qquad |z| \leqq 1\,,$$

so that $x(n)$ can be represented in the form (2), and (7) is satisfied (see, for example, [6], Chapter I). The $\beta(j)$ and $\sigma^2$ $(=E\varepsilon^2(n))$ are estimated through

$$\sum_{j=0}^{q} \hat{\beta}(j)c(k-j) = \delta_{0,k}\hat{\sigma}^2\,, \qquad k = 0, 1, \cdots, q\,,$$

([6], Chapter VI) and, remembering that $\beta(0) = 1$, we see that $\hat{\beta}(1), \cdots, \hat{\beta}(q)$ are functions only of the $r(j)$, $j = 1, 2, \cdots, q$ and thus Theorem 3 may be applied. Similar considerations apply to the mixed autoregressive and moving average process obtained when (17) is altered only by the replacement of the right-hand side by

$$\sum_{j=0}^{r} \delta(j)\varepsilon(n-j) .$$

It seems that all of the classical inferential theory for the $\beta(j)$, $\delta(j)$ (see [2]; [6], Chapter VI) will continue to apply under the appropriate conditions of the present paper.

There are other problems that yield to the same treatment as we have applied to the $r(j)$ in the present paper. For example, if $x(n)$ is generated by (2) and is stationary with finite fourth moment and (3) holds, then

$$\lim_{N\to\infty} \sup_{|\lambda|\leq\pi} |N^{-1} \sum_{n=1}^{N} (x(n) - \mu)e^{in\lambda}| = 0 \quad \text{a.s.}$$

The proof of this proposition, which is important in connection with the estimation of the frequencey of a sinusoidal signal rèceived together with noise, follows the same lines as that given in Hannan [7].

**Acknowledgment.** We are indebted to E. Seneta for some useful conversations during the initial stage of this work.

## REFERENCES

[1] ANDERSON, T. W. and WALKER, A. M. (1964). On the asymptotic distribution of the autocorrelations of a sample from a linear stochastic process. *Ann. Math. Statist.* **35** 1296–1303.

[2] BOX, G. E. P. and JENKINS, G. M. (1970). *Time Series Analysis Forecasting and Control.* Holden-Day, San Francisco.

[3] BROWN, B. M. (1971). Martingale central limit theorems. *Ann. Math. Statist.* **42** 59–66.

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

[5] GIKHMAN, I. I. and SKOROKHOD, A. V. (1969) *Introduction to the Theory of Random Processes.* Saunders, Philadelphia.

[6] HANNAN, E. J. (1970). *Multiple Time Series.* Wiley, New York.

[7] HANNAN, E. J. (1971). Non linear time series regression. *J. Appl. Probability* **8** 767–780.

[8] HEYDE, C. C. and SENETA, E. (1972). Estimation theory for growth and immigration rates in a muliplicative process. To appear in *J. Appl. Probability.*

[9] KATZNELSON, Y. (1968). *Harmonic Analysis.* Wiley, New York.

DEPARTMENT OF STATISTICS
AUSTRALIAN NATIONAL UNIVERSITY
BOX 4, POST OFFICE, CANBERRA A.C.T. 2600
AUSTRALIA