ON ASYMPTOTIC DISTRIBUTIONS OF ESTIMATES OF PARAMETERS OF STOCHASTIC DIFFERENCE EQUATIONS

T. W. ANDERSON

Columbia University and Center for Advanced Study in the Behavioral Sciences

1. Summary and introduction. Let x_t $(t = 1, 2, \cdots)$ be defined recursively by

$$x_t = \alpha x_{t-1} + u_t, \qquad t = 1, 2, \cdots, \tag{1.1}$$

where x_0 is a constant, $\mathcal{E}u_t = 0$, $\mathcal{E}u_t^2 = \sigma^2$ and $\mathcal{E}u_t u_s = 0$, $t \neq s$. ($\mathcal{E}$ denotes mathematical expectation.) An estimate of α based on $x_1, \cdots, x_T$ (which is the maximum likelihood estimate of α if the u's are normally distributed) is

$$\hat{\alpha} = \left(\sum_{t=1}^{T} x_t x_{t-1}\right) \Big/ \left(\sum_{t=1}^{T} x_{t-1}^2\right). \tag{1.2}$$

If $|\alpha| < 1$, $\sqrt{T}(\hat{\alpha} - \alpha)$ has a limiting normal distribution with mean 0 under fairly general conditions such as independence of the u's and uniformly bounded moments of the u's of order $4 + \epsilon$, for some $\epsilon > 0$. (See [2], Chapter II, for example.) If $|\alpha| > 1$, White [3] has shown $(\hat{\alpha} - \alpha)|\alpha|^T / (\alpha^2 - 1)$ has a limiting Cauchy distribution under the assumption that $x_0 = 0$ and the u's are normally distributed; he has also found the distribution when $x_0 \neq 0$. His results can be easily modified and restated in the following form $(\sum_{t=1}^T x_{t-1}^2)^{\frac{1}{2}}(\hat{\alpha} - \alpha)$ has a limiting normal distribution if the u's are normally distributed and if $|\alpha| \neq 1$. Peculiarly, for $|\alpha| = 1$ this statistic has a limiting distribution which is not normal (and is not even symmetric for $x_0 = 0$). One purpose of this paper is to characterize the limiting distributions for $|\alpha| > 1$ when the u's are not necessarily normally distributed; it will be shown that for $|\alpha| > 1$ the results depend on the distribution of the u's. Central limit theorems are not applicable.

Secondly, the limiting distribution for $|\alpha| < 1$ will be shown to hold under the assumption that the u's are independently, identically distributed with finite variance. This was conjectured by White.

2. Asymptotic distributions in the unstable case. Here $|\alpha| > 1$. Let

$$A_T = \sum_1^T x_t x_{t-1} - \alpha \sum_1^T x_{t-1}^2 = \sum_1^T u_t x_{t-1}, \tag{2.1}$$

$$B_T = \sum_1^T x_{t-1}^2. \tag{2.2}$$

Then $\hat{\alpha} - \alpha = A_T/B_T$. Note that

$$\begin{aligned} x_t &= \alpha x_{t-1} + u_t = \alpha(\alpha x_{t-2} + u_{t-1}) + u_t = \cdots \\ &= u_t + \alpha u_{t-1} + \cdots + \alpha^{t-1} u_1 + \alpha^t x_0. \end{aligned} \tag{2.3}$$

Received August 25, 1958; revised February 17, 1959.

Let $\beta = 1/\alpha$ and let

$$z_t = \beta^{t-2} x_{t-1} = u_1 + \beta u_2 + \cdots + \beta^{t-2} u_{t-1} + \alpha x_0 . \tag{2.4}$$

It is easily verified that $\mathcal{E} z_t = \alpha x_0$ and $\operatorname{Var} z_t \to \sigma^2/(1 - \beta^2)$ as $T \to \infty$.

THEOREM 2.1.

$$\operatorname*{plim}_{T \to \infty} \left(\beta^{2(T-2)} B_T - \frac{1}{1 - \beta^2} z_T^2 \right) = 0. \tag{2.5}$$

PROOF. We shall show that

$$\begin{aligned} \beta^{2(T-2)} B_T - \frac{1}{1-\beta^2} z_T^2 & \\ &= \left[\beta^{2(T-2)} B_T - \frac{1 - \beta^{2T}}{1 - \beta^2} z_T^2 \right] + \left[\frac{1 - \beta^{2T}}{1 - \beta^2} z_T^2 - \frac{1}{1 - \beta^2} z_T^2 \right] \\ &= \sum_{s=1}^{T-1} \beta^{2s} (z_{T-s}^2 - z_T^2) - \frac{\beta^{2T}}{1 - \beta^2} z_T^2 \end{aligned} \tag{2.6}$$

converges stochastically to 0. From (1.1) and (2.4) we find

$$z_t = z_{t-1} + \beta^{t-2} u_{t-1} , \tag{2.7}$$

and hence

$$z_T = \beta^{T-2} u_{T-1} + \beta^{T-3} u_{T-2} + \cdots + \beta^{T-s-1} u_{T-s} + z_{T-s} . \tag{2.8}$$

We shall use the results that

$$\begin{aligned} \mathcal{E}(z_{T-s} - z_T)^2 &= \sigma^2[\beta^{2(T-2)} + \cdots + \beta^{2(T-s-1)}] \\ &\leqq \sigma^2 \frac{\beta^{2(T-s-1)}}{1 - \beta^2} , \end{aligned} \tag{2.9}$$

$$\begin{aligned} \mathcal{E}(z_{T-s} + z_T)^2 &\leqq 2\mathcal{E} z_{T-s}^2 + 2\mathcal{E} z_T^2 \\ &\leqq 4\mathcal{E} z_T^2 \leqq 4 \left(\frac{\sigma^2}{1 - \beta^2} + \alpha^2 x_0^2 \right). \end{aligned} \tag{2.10}$$

Then

$$\begin{aligned} \mathcal{E} \left| \beta^{2(T-2)} B_T - \frac{1 - \beta^{2T}}{1 - \beta^2} z_T^2 \right| &\leqq \sum_{s=1}^{T-1} \beta^{2s} \mathcal{E} \mid z_{T-s}^2 - z_T^2 \mid \\ &= \sum_{s=1}^{T-1} \beta^{2s} \mathcal{E} \mid (z_{T-s} + z_T)(z_{T-s} - z_T) \mid \\ &\leqq \sum_{s=1}^{T-1} \beta^{2s} [\mathcal{E}(z_{T-s} + z_T)^2 \mathcal{E}(z_{T-s} - z_T)^2]^{\frac{1}{2}} \\ &\leqq 2 \left(\frac{\sigma^2}{1 - \beta^2} + \alpha^2 x_0^2 \right)^{\frac{1}{2}} \frac{\sigma}{(1 - \beta^2)^{\frac{1}{2}}} \sum_{s=1}^{T-1} |\beta|^{T+s-1} \\ &\leqq 2 \left(\frac{\sigma^2}{1 - \beta^2} + \alpha^2 x_0^2 \right)^{\frac{1}{2}} \frac{\sigma}{(1 - \beta^2)^{\frac{1}{2}}} \frac{|\beta|^T}{1 - |\beta|} . \end{aligned} \tag{2.11}$$

By Tchebycheff's inequality

$$\Pr\left\{\left|\beta^{2(T-2)}B_T - \frac{1-\beta^{2T}}{1-\beta^2} z_T^2\right| > \epsilon\right\} \leqq \frac{K}{\epsilon} |\beta|^T, \tag{2.12}$$

where K is a constant, and for T sufficiently large this is arbitrarily small. Since

$$\mathcal{E} \frac{\beta^{2T}}{1-\beta^2} z_T^2 \leqq \beta^{2T} C \tag{2.13}$$

for C a suitable positive constant, the term in (2.6) converges in probability to 0 and the theorem follows.

The convergence in (2.5) is also with probability 1. The sum of (2.12) for $T = 1, 2, \cdots$ converges and similarly for (2.13). Hence, by the Borel-Cantelli Lemma (2.6) converges to 0 with probability 1.

It should be observed that z_T will have a limiting distribution. (In fact $\sum_1^\infty \beta^{(t-1)} u_t$ converges in the mean and with probability 1.) It will also be noted that $\beta^{2(T-2)}B_T$ is in the limit a nondegenerate random variable; if $\beta^{2(T-2)}$ is replaced by a function of T that decreases faster, then the resulting random variable converges stochastically to 0.

Let

$$y_T = u_T + \beta u_{T-1} + \cdots + \beta^{T-2} u_2 + \beta^{T-1} u_1. \tag{2.14}$$

THEOREM 2.2.

$$\operatorname*{plim}_{T\to\infty} (\beta^{T-2} A_T - y_T z_T) = 0. \tag{2.15}$$

PROOF. We have

$$\beta^{T-2} A_T - y_T z_T = \sum_{s=1}^{T-1} \beta^s u_{T-s}(z_{T-s} - z_T).$$

Then

$$\begin{aligned} \mathcal{E}\, |\beta^{T-2} A_T - y_T z_T| &\leqq \sum_{s=1}^{T-1} |\beta|^s \mathcal{E}\, |u_{T-s}(z_{T-s} - z_T)| \\ &\leqq \sum_{s=1}^{T-1} |\beta|^s [\mathcal{E} u_{T-s}^2 \mathcal{E}(z_{T-s} - z_T)^2]^{\frac{1}{2}} \\ &\leqq \frac{\sigma^2}{\sqrt{1-\beta^2}} \sum_{s=1}^{T-1} |\beta|^{T-1} \\ &= \frac{\sigma^2}{\sqrt{1-\beta^2}} (T-1)\, |\beta|^{T-1}. \end{aligned} \tag{2.16}$$

Since (2.16) converges to 0, the Tchebycheff inequality implies the theorem.

Since the sum of (2.16) for $T = 1, 2, \cdots$ converges, the Borel-Cantelli lemma implies convergence with probability 1.

It will be noticed that y_T has the same form as z_T except for x_0 and the order of the u's is reversed and there is one more term. Under the assumptions we have made y_T does not necessarily have a limiting distribution. For example, u_T makes a not negligible contribution to y_T; if the u's are independent and if the sequence of distributions of u_T is wildly fluctuating, y_T will not have a limiting distribution. However, if the u's are independent and identically distributed, y_T has the same limiting distribution as z_T for $x_0 = 0$. The covariance between y_T and z_T is $(T - 1)\sigma^2\beta^{T-1}$, which converges to 0.

THEOREM 2.3. *If the u's are independently distributed, and if y_T has a limiting distribution, then (y_T, z_T) has a limiting distribution, say the distribution of (y, z), and y and z are independent.*[1]

PROOF. Let

$$z_T^* = \alpha x_0 + \sum_{t=1}^{[\frac{1}{2}T]} \beta^{t-1} u_t, \tag{2.17}$$

$$\tilde{z}_T = \sum_{t=[\frac{1}{2}T]+1}^{T-1} \beta^{t-1} u_t, \tag{2.18}$$

$$y_T^* = \sum_{t=[\frac{1}{2}T]+1}^{T} \beta^{T-t} u_t, \tag{2.19}$$

$$\tilde{y}_T = \sum_{t=1}^{[\frac{1}{2}T]} \beta^{T-t} u_t, \tag{2.20}$$

where $[\frac{1}{2}T]$ is the largest integer not greater than $\frac{1}{2}T$. Then z_T^* and y_T^* are independently distributed because they involve disjoint sets of u's. We have

$$\begin{aligned} \mathcal{E}(z_T - z_T^*)^2 &= \mathcal{E}\tilde{z}_T^2 \\ &= \sigma^2 \sum_{t=[\frac{1}{2}T]+1}^{T-1} \beta^{2(t-1)} \\ &\leqq \frac{\sigma^2\beta^{2[\frac{1}{2}T]}}{1-\beta^2} \leqq \frac{\sigma^2 |\beta|^{T-1}}{1-\beta^2}, \end{aligned} \tag{2.21}$$

$$\begin{aligned} \mathcal{E}(y_T - y_T^*)^2 &= \mathcal{E}\tilde{y}_T^2 \\ &= \sigma^2 \sum_{t=1}^{[\frac{1}{2}T]} \beta^{2(T-t)} \\ &\leqq \frac{\sigma^2\beta^{2(T-[\frac{1}{2}T])}}{1-\beta^2} \leqq \frac{\sigma^2 |\beta|^{T}}{1-\beta^2}. \end{aligned} \tag{2.22}$$

Then $z_T - z_T^*$ and $y_T - y_T^*$ converge stochastically and with probability 1 to 0 and the theorem follows.

THEOREM 2.4. *If (y_T, z_T) has a limiting distribution, say the distribution of (y, z) then $(\beta^{T-2}A_T, (1 - \beta^2)\beta^{2(T-2)}B_T)$ has a limiting distribution, the distribution of (yz, z^2).*

[1] This theorem as well as several other points, was suggested by Julius Blum.

THEOREM 2.5. *If* (y_T, z_T) *has a limiting distribution, the distribution of* (y, z), *and if* $\Pr\{z = 0\} = 0$, *then* $[\alpha^T/(\alpha^2 - 1)](\hat{\alpha} - \alpha)$ *has as a limiting distribution the distribution of* y/z.

THEOREM 2.6. *If the u's are independently normally distributed, the limiting distribution of* (y_T, z_T) *is normal with variances* $\sigma^2/(1 - \beta^2)$, *correlation* 0, $\mathcal{E}y = 0$ *and* $\mathcal{E}z = \alpha x_0$.

It will be observed that if the u's are independent and not all normally distributed, then z_T does not have a limiting normal distribution. For example, u_1 is not negligible; if it is not normal, z is not normal (since a convolution is normal only if the two component distributions are normal).

In the case of y_T, if all the u's beyond some t are normal and independent, then y_T will have a limiting normal distribution. If the u's are independently and identically distributed, then y_T has a limiting normal distribution if and only if the u's are normally distributed. If α is an integer and if the u's are independently distributed according to a rectangular discrete distribution over $0, 1, \cdots, \alpha - 1$, then the limiting distribution of y_T is (continuous) uniform on $(0, \alpha)$. It will be noted that central limit theorems are not applicable here.

THEOREM 2.7. *If the u's are independently normally distributed and if* $x_0 = 0$, $[\alpha^T/(\alpha^2 - 1)](\hat{\alpha} - \alpha)$ *has a Cauchy distribution as a limiting distribution.*

THEOREM 2.8. *If the u's are independently normally distributed* $(\sum x_{t-1}^2)^{\frac{1}{2}}(\hat{\alpha} - \alpha)$ *has a limiting normal distribution with mean* 0 *and variance* σ^2.

THEOREM 2.9. *If the u's are independently normally distributed, the limiting moment generating function of* $\beta^{T-2}(1 - \beta^2)A_T / \sigma^2$ *and* $\beta^{2(T-2)}(1 - \beta^2)^2 B_T / \sigma^2$ *is*

$$(1 - U^2 - 2V)^{-\frac{1}{2}} \exp\left[\frac{\frac{1}{2}(\alpha^2 - 1)x_0^2}{\sigma^2} \frac{U^2 + 2V}{1 - U^2 - 2V}\right]. \tag{2.23}$$

PROOF. We have

$$\begin{aligned} &\mathcal{E}e^{(Uyz+Vz^2)(1-\beta^2)/\sigma^2} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1 - \beta^2}{2\pi\sigma^2} e^{-\frac{1}{2}(1-\beta^2)[y^2+(z-\alpha x_0)^2]/\sigma^2+(Uyz+Vz^2)(1-\beta^2)/\sigma^2}\, dy\, dz \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1 - \beta^2}{2\pi\sigma^2} e^{-\frac{1}{2}(1-\beta^2)\{y^2-2Uyz+(1-2V)z^2-2z\alpha x_0+\alpha^2x_0^2\}/\sigma^2}\, dy\, dz \end{aligned} \tag{2.24}$$

which is (2.23). This was given by White [3].

Theorem 2.8 permits setting up tests of hypotheses about α and forming confidence intervals for α if the u's are independently normally distributed. In the case of $|\alpha| < 1$, the result holds without the assumption of normality (see Section 4). It should be emphasized that statistical procedures based on the asymptotic normal distribution of $(\sum x_{t-1}^2)^{\frac{1}{2}}(\hat{\alpha} - \alpha)$ have wide scope when $|\alpha| < 1$ but when $|\alpha| > 1$ are justified only if the u's are normal.

If $\alpha = 1$, $x_t = \sum_1^t u_s + x_0$ and the numerator of $\hat{\alpha} - \alpha$ is

$$
\begin{aligned}
A_T &= \sum u_t x_{t-1} \\
&= \sum_{s<t} u_t u_s + \sum u_t x_0 \\
&= \tfrac{1}{2}[(\sum u_t)^2 - \sum u_t^2] + x_0 \sum u_t .
\end{aligned} \tag{2.25}
$$

In this case the normalization factor is T^2. Then

$$
\frac{1}{T} A_T = \tfrac{1}{2} (\sum u_t/\sqrt{T})^2 - \tfrac{1}{2} \sum u_t^2/T + x_0 \sum u_t/T. \tag{2.26}
$$

If the u's are independently and identically distributed, $\sum u_t / T$ converges stochastically to $\mathcal{E}u_t = 0$ and $\sum u_t^2 / T$ converges stochastically to $\mathcal{E}u_t^2 = \sigma^2$, and $\sum u_t / \sqrt{T}$ has a limiting normal distribution with mean 0 and variance σ^2. Thus the limiting distribution of A_T/T is that of $\frac{1}{2}x^2 - \frac{1}{2}\sigma^2$, where x has a normal distribution with mean 0 and variance σ^2. From this it is clear that A_T multiplied by any nonnegative function of the observations cannot have a limiting normal distribution since

$$
\lim_{t\to\infty} \Pr\{A_T \leqq 0\} = \Pr\{x^2 \leqq \sigma^2\} \tag{2.27}
$$

which is not $\frac{1}{2}$. White has observed that if the u's are independently normally distributed and if $x_0 = 0$, the limiting distribution of

$$
T(\hat{\alpha} - \alpha) = \frac{A_T/T}{B_T/T^2} \tag{2.28}
$$

is that of

$$
\frac{1}{2} \frac{x^2(1) - 1}{\int_0^1 x^2(t)\, dt} \tag{2.29}
$$

where $x(t)$ is the Wiener stochastic process with $\mathcal{E}x(t) = 0$ and $\mathcal{E}x^2(t) = t$, and he has given the limiting characteristic function of $(A_T/T, B_T/T^2)$.

It might be noted that in the case of $|\alpha| > 1$, the condition $\mathcal{E}u_t^2 = \sigma^2$ could be replaced by the condition $\mathcal{E}u_t^2 = \sigma_t^2 < M$ for some M. The results would involve such modifications as replacing $\sigma^2/(1 - \beta^2)$ by

$$
\sum_0^\infty \beta^{2s}\sigma_{s+1}^2 \; [<M/(1 - \beta^2)].
$$

3. Asymptotic distributions in the unstable vector case. Let x_t and u_t be p-component column vectors and α a $p \times p$ matrix. Let the process be defined by (1.1), where x_0 is a vector of constants, $\mathcal{E}u_t = 0$, $\mathcal{E}u_t u_t' = \Sigma$ and $\mathcal{E}u_t u_s' = 0$, $t \neq s$. The estimate of α is

$$
\hat{\alpha} = \sum x_t x_{t-1}' (\sum x_{t-1} x_{t-1}')^{-1}. \tag{3.1}
$$

The process is stable if all the characteristic roots of α are less than 1 in absolute value; we shall consider in this section the case that all p characteristic roots are greater than 1 in absolute value. The methods for the scalar case can be used here, but the results are more complicated. A more general case would include matrices α with some roots less and some roots greater than 1 in absolute value, but this would be much more involved.

Let

$$A_T = \sum x_t x'_{t-1} - \alpha \sum x_{t-1} x'_{t-1}, \tag{3.2}$$

$$B_T = \sum x_{t-1} x'_{t-1}, \tag{3.3}$$

$$\begin{aligned} z_T &= \alpha^{-(T-2)} x_{T-1} \\ &= u_1 + \alpha^{-1} u_2 + \cdots + \alpha^{-(T-2)} u_{T-1} + \alpha x_0, \end{aligned} \tag{3.4}$$

$$z = \sum_{t=1}^{\infty} \alpha^{-(t-1)} u_t + \alpha x_0, \tag{3.5}$$

$$F_T = z_T z'_T + \alpha^{-1} z_T z'_T \alpha^{-1\prime} + \cdots + \alpha^{-(T-1)} z_T z'_T \alpha^{-(T-1)\prime}, \tag{3.6}$$

$$G_T = u_T z'_T + u_{T-1} z'_T \alpha^{-1\prime} + \cdots + u_1 z'_T \alpha^{-(T-1)\prime}. \tag{3.7}$$

Then

$$\hat{\alpha} - \alpha = A_T B_T^{-1}. \tag{3.8}$$

THEOREM 3.1.

$$\operatorname*{plim}_{T\to\infty} (\alpha^{-(T-2)} B_T \alpha^{-(T-2)\prime} - F_T) = 0. \tag{3.9}$$

THEOREM 3.2.

$$\operatorname*{plim}_{T\to\infty} (A_T \alpha^{-(T-2)\prime} - G_T) = 0. \tag{3.10}$$

These theorems are proved by methods similar to those used for Theorems 2.1 and 2.2. The convergence in each case is also with probability 1.

Suppose that α is a matrix such that there exists a nonsingular matrix c such that

$$c\alpha c^{-1} = \lambda, \tag{3.11}$$

where λ is a diagonal matrix with the characteristic roots of α as the diagonal elements. Then $\alpha = c^{-1}\lambda c$, $\alpha^{-1} = c^{-1}\lambda^{-1}c$, and $\alpha^{-r} = c^{-1}\lambda^{-r}c$. Let $\lambda^{-1} = \gamma$. Then

$$\begin{aligned} F_T &= z_T z'_T + c^{-1}\gamma c z_T z'_T c' \gamma c^{-1\prime} + \cdots \\ &\quad + c^{-1}\gamma^{T-1} c z_T z'_T c' \gamma^{T-1} c^{-1\prime} \\ &= c^{-1}(c z_T z'_T c' + \cdots + \gamma^{T-1} c z_T z'_T c' \gamma^{T-1}) c^{-1\prime}. \end{aligned} \tag{3.12}$$

The i, jth element of the matrix in parentheses is the i, jth element of $c z_T z'_T c'$ multiplied by

$$1 + \gamma_i \gamma_j + \cdots + (\gamma_i \gamma_j)^{T-1} = \frac{1 - (\gamma_i \gamma_j)^T}{1 - \gamma_i \gamma_j}, \tag{3.13}$$

where γ_i is the ith diagonal element of γ. This converges to $1/(1 - \gamma_i\gamma_j)$. Then the i, jth element of cF_Tc' is asymptotically the i, jth element of $czz'c'$ divided by $1 - \gamma_i\gamma_j$. Let Γ be the matrix with $1/(1 - \gamma_i\gamma_j)$ as the i, jth element, and let Z_T be a diagonal matrix with ith diagonal element equal to the ith element of cz_T.

COROLLARY 3.1.

$$\operatorname*{plim}_{T\to\infty} (F_T - c^{-1}Z_T\Gamma Z_Tc^{-1\prime}) = 0. \tag{3.14}$$

Now consider

$$\begin{aligned} G_T &= u_Tz_T' + u_{T-1}z_T'c'\gamma c^{-1\prime} + \cdots + u_1z_T'c'\gamma^{T-1}c^{-1\prime} \\ &= (u_Tz_T'c' + u_{T-1}z_T'c'\gamma + \cdots + u_1z_T'c'\gamma^{T-1})c^{-1\prime}. \end{aligned} \tag{3.15}$$

The jth column of the matrix in parentheses is the jth element of $z'c'$ times

$$u_T + \gamma_ju_{T-1} + \cdots + \gamma_j^{T-1}u_1. \tag{3.16}$$

Let this be the jth element of a matrix Y_T.

COROLLARY 3.2.

$$\operatorname*{plim}_{T\to\infty} (G_T - Y_TZ_Tc^{-1\prime}) = 0. \tag{3.17}$$

It should be noted that γ and c do not need to be real, but $c^{-1}Z_T\Gamma Z_Tc^{-1\prime}$ and $Y_TZ_Tc^{-1\prime}$ will be real. In fact the diagonal elements of Z_T are the elements of cz_T, where c is complex and z_T consists of real random variables. The elements of Y_T are complex linear combinations of real random variables. When we speak of Y_T having a limiting distribution we mean the set of real random variables has a limiting distribution. (The coefficients of the linear combinations remain fixed.)

THEOREM 3.3. *If the u's are independently distributed and if Y_T has a limiting distribution, then (Y_T, Z_T) has a limiting distribution, say the distribution of (Y, Z), and Y and Z are independent.*

THEOREM 3.4. *If (Y_T, Z_T) has a limiting distribution, say the distribution of (Y, Z), then $(A_T\alpha^{-(T-2)\prime}, \alpha^{-(T-2)}B_T\alpha^{-(T-2)\prime})$ has a limiting distribution, the distribution of $(YZc^{-1\prime}, c^{-1}Z\Gamma Zc^{-1\prime})$.*

THEOREM 3.5. *If (Y_T, Z_T) has a limiting distribution, the distribution of (Y, Z), if the probability is 1 that each diagonal component of Z is different from 0, and if Γ is nonsingular, then $(\hat{\alpha} - \alpha)\alpha^{(T-2)}$ has as a limiting distribution the distribution of $Y\Gamma^{-1}Z^{-1}c$.*

It may be noted that Γ is nonsingular if and only if the characteristic roots of α are all different. Since a diagonal component of Z is a linear combination of the components of z, all the diagonal components will be different from 0 with probability 1 if the probability is 0 that the components of z satisfy a linear relation.

THEOREM 3.6. *If the u's are independently normally distributed the limiting distribution of (Y_T, Z_T) is that of (Y, Z) where Y and Z are composed of linear combinations of two sets of independent normal variables.*

The mean of z is αx_0 and the covariance matrix is

$$c^{-1}\left(\frac{\sum_{k,l} c_{ik}\,\sigma_{kl}\,c_{jl}}{1-\gamma_i\gamma_j}\right)c^{-1\prime}. \tag{3.18}$$

The mean of Y is 0. The covariances are harder to describe. Let w be an arbitrary real p-component vector and let W be the diagonal matrix with the elements of cw as the diagonal elements. Then the covariance matrix of the ith and jth rows of YWc^{-1} is $\sigma_{ij}c^{-1}W\Gamma Wc^{-1\prime}$.

We can give a kind of analogue of Theorem 2.8. In the scalar case, if the u's are independently normally distributed, $\alpha^{(T-2)}(\hat{\alpha}-\alpha)$ and $B_T\alpha^{-2(T-2)}$ have as a limiting distribution, the distribution of y/z and z^2; this limiting distribution has the property that the conditional distribution of y/z given z is normal with mean 0 and variance σ^2/z^2. In the vector case if the u's are independently normally distributed and the characteristic roots of α are all different, $(\hat{\alpha}-\alpha)\alpha^{(T-2)}$ and $\alpha^{-(T-2)}B_T\alpha^{-(T-2)\prime}$ have as a limiting distribution the distribution of $Y\Gamma^{-1}Z^{-1}c$ and $c^{-1}Z\Gamma Zc^{-1\prime}$; this has the property that the conditional distribution of $Y\Gamma^{-1}Z^{-1}c$ given Z is normal with mean 0 and covariances $\Sigma \times x(c^{-1}Z\Gamma Zc^{-1\prime})^{-1}$. This result can be used to justify the usual procedures of testing hypotheses and confidence intervals when the above conditions are satisfied.

The mth order scalar difference equation can be treated by writing it as a special first order vector equation by letting the vector x_t' be made up of the scalars $(x_t, x_{t-1}, \cdots, x_{t-m+1})$, and the mth order vector case can be treated similarly.

4. Asymptotic distributions in the stable case. In this section we assume that the u's are independently and identically distributed and that $|\alpha| < 1$. Then we show that $\sqrt{T}\,(\hat{\alpha}-\alpha)$ has a limiting normal distribution. The important feature here is that the variance of u_t is assumed finite, but nothing is assumed about moments of higher order. Diananda [1] proved a result similar to this when $\alpha = 0$.

THEOREM 4.1. *The limiting distribution of $A_T/\sqrt{T}$ is normal with mean 0 and variance $\sigma^4/(1-\alpha^2)$.*

PROOF.

$$A_T = \sum_{t=2}^{T} u_t u_{t-1} + \alpha \sum_{t=3}^{T} u_t u_{t-2} + \cdots + \alpha^{T-2}u_T u_1 + x_0 \sum_{t=1}^{T} \alpha^{t-1}u_t. \tag{4.1}$$

The last term has mean 0 and variance $x_0^2\sigma^2(1-\alpha^{2T})/(1-\alpha^2)$; this divided by T converges to 0, the random term converges stochastically to 0 and can be neglected. Let

$$A_T^* = A_T - x_0 \sum_{1}^{T} \alpha^{t-1}u_t. \tag{4.2}$$

Then A_T^* is a linear combination of terms $u_t u_s$, $t \neq s$. Each term has mean $\mathcal{E}u_t u_s = 0$ and variance $\mathcal{E}(u_t u_s)^2 = \mathcal{E}u_s^2 u_t^2 = \sigma^4$. Each term is uncorrelated with each other term.

Let

$$C_{T,S} = \sum_{2}^{T} u_t u_{t-1} + \alpha \sum_{3}^{T} u_t u_{t-2} + \cdots + \alpha^S \sum_{t=S+2}^{T} u_t u_{t-S-1} \tag{4.3}$$

for $S \leqq T - 2$ and let $C_{T,S} = A_T^*$ for $S > T - 2$. Then $A_T^* - C_{T,S}$ has mean 0 and variance bounded by

$$\frac{\sigma^4 \alpha^{2(S+1)}}{1 - \alpha^2} [T - (S + 2)]. \tag{4.4}$$

Then $A^*/\sqrt{T} - C_{T,S}/\sqrt{T}$ has mean 0 and a variance bounded (uniformly in T) by $\sigma^4\alpha^{2(S+1)}/(1 - \alpha^2)$. This can be made arbitrarily small by making S sufficiently large. Now let

$$C_{T,S}^* = \sum_{S+2}^{T} [u_t u_{t-1} + \alpha u_t u_{t-2} + \cdots + \alpha^S u_t u_{t-S-1}]. \tag{4.5}$$

The limiting distribution of $C_{T,S}^*/\sqrt{T}$ is the same as of $C_{T,S}/\sqrt{T}$. Let

$$y_t = u_t u_{t-1} + \alpha u_t u_{t-2} + \cdots + \alpha^S u_t u_{t-S-1}. \tag{4.6}$$

Then

$$\mathcal{E} y_t^2 = \frac{1 - \alpha^{2(S+1)}}{1 - \alpha^2} \sigma^4, \tag{4.7}$$

$$\mathcal{E} y_t y_s = 0, \qquad t \neq s, \tag{4.8}$$

and y_t is an $(S + 1)$-dependent sequence. Theorem 4.4 below applies, and hence $C_{T,S}^*/\sqrt{T}$ has a limiting normal distribution with mean 0 and variance (4.7). Theorem 4.5 below completes the proof.

THEOREM 4.2.

$$\operatorname*{plim}_{T \to \infty} B_T/T = \sigma^2/(1 - \alpha^2). \tag{4.9}$$

PROOF.

$$\begin{aligned} B_T = \sum_{1}^{T} x_{t-1}^2 &= x_0^2 + (u_1 + \alpha x_0)^2 + (u_2 + \alpha u_1 + \alpha^2 x_0)^2 \\ &\quad + \cdots + (u_{T-1} + \alpha u_{T-2} + \cdots + \alpha^{T-1} x_0)^2 \\ &= [u_1^2(1 + \alpha^2 + \cdots + \alpha^{2(T-2)}) + \cdots + u_{T-1}^2] \\ &\quad + 2[\alpha(u_2 u_1 + \cdots + u_{T-1} u_{T-2}) + \cdots + \alpha^{T-2} u_{T-1} u_1] \\ &\quad + 2x_0[u_1(\alpha + \alpha^3 + \cdots + \alpha^{2T-3}) + \cdots + \alpha^{T-1} u_{T-1}] \\ &\quad + x_0^2[1 + \alpha^2 + \cdots + \alpha^{2(T-1)}]. \end{aligned} \tag{4.10}$$

The last term divided by T converges to 0. The next to last term has mean 0 and variance bounded by a constant times T; when this term is divided by T it converges stochastically to 0. The second bracket has mean 0 and variance

$$(4.11)\qquad [\alpha^2(T-2)+\alpha^4(T-3)+\cdots+\alpha^{2(T-2)}]\sigma^4 < T\sigma^4[1+\alpha^2+\cdots] = T\sigma^4/(1-\alpha^2).$$

This term divided by T converges stochastically to 0. Thus B_T/T has the probability limit of the first bracket divided by T. But

$$(4.12)\qquad \begin{aligned} \frac{1}{1-\alpha^2}\sum_1^{T-1} u_t^2 - [u_1^2(1+\alpha^2+\cdots+\alpha^{2(T-2)})+\cdots+u_{T-1}^2] &= u_1^2(\alpha^{2(T-1)}+\alpha^{2T}+\cdots)+\cdots+u_{T-1}^2(\alpha^2+\alpha^4+\cdots) \\ &= u_1^2\frac{\alpha^{2(T-1)}}{1-\alpha^2}+\cdots+\frac{\alpha^2}{1-\alpha^2}u_{T-1}^2. \end{aligned}$$

This is a nonnegative random variable with expected value

$$(4.13)\qquad \frac{1}{1-\alpha^2}[\alpha^2+\cdots+\alpha^{2(T-1)}]\sigma^2 = \frac{\alpha^2}{1-\alpha^2}\,\frac{1-\alpha^{2(T-1)}}{1-\alpha^2}\,\sigma^2,$$

and divided by T converges to 0. Thus

$$(4.14)\qquad \operatorname{plim}\frac{B_T}{T} = \operatorname{plim}\frac{\sum_1^{T-1} u_t^2}{(1-\alpha^2)T} = \frac{\sigma^2}{1-\alpha^2}$$

by the law of large numbers.

THEOREM 4.3. *The limiting distribution of* $\sqrt{T}(\hat\alpha-\alpha)$ *is normal with mean* 0 *and variance* $1-\alpha^2$.

PROOF.

$$(4.15)\qquad \sqrt{T}\,(\hat\alpha-\alpha) = \sqrt{T}\,\frac{A_T}{B_T} = \frac{A_T/\sqrt{T}}{B_T/T}.$$

This proof exploits the fact that the second-order moments of A_T involve only the second-order moments of the u's (because A_T only involves products of independent u's) and that a special central limit theorem applies. The result can easily be extended to the vector case, where the characteristic roots of the matrix α are less than 1 in absolute value. In turn this permits extension to the general-order difference equation (scalar or vector) in the stable case. The case of

$$(4.16)\qquad x_t = \alpha x_{t-1} + \gamma + u_t$$

again can be treated this way. However, the case of

$$(4.17)\qquad x_t = \alpha x_{t-s} + \gamma z_t + u_t,$$

where z_t is a sequence of fixed variates, will not in general yield to this treatment (unless restrictions are made so that asymptotically z_t washes out); the reason is that in addition to terms like $u_t u_{t-i}$ there will be terms $u_t z_{t-i}$ and these will not be identically distributed.

The following central limit theorem was given by Diananda:

THEOREM 4.4. *Let* $y_1, y_2, \cdots,$ *be a sequence of random variables such that the distribution of* $(y_{t+t_1}, y_{t+t_2}, \cdots, y_{t+t_n})$ *is independent of* t *for every* $t_1 < t_2 < \cdots < t_n (t_1 \geqq 0)$ *and* n *and such that this collection is independent of* $(y_{s+s_1}, y_{s+s_2}, \cdots, y_{s+s_p})$ *for every* $s_1 < s_2 < \cdots < s_p (s_1 \geqq 0)$ *and* p *if* $s > t + t_n + m$. *Assume* $\mathcal{E} y_t = 0$, $\mathcal{E} y_t^2 < \infty$. *Then* $\sum_1^T y_t/\sqrt{T}$ *has a limiting normal distribution with mean* 0 *and variance*

$$\mathcal{E} y_1^2 + 2\mathcal{E} y_1 y_2 + \cdots + 2\mathcal{E} y_1 y_{m+1} . \tag{4.18}$$

The sequence y_t is called m-dependent. The proof depends upon a theorem proved by Diananda, which is similar to the following:[2]

THEOREM 4.5. *Let*

$$S_T = Z_{kT} + X_{kT}, \qquad T = 1, 2, \cdots, \quad k = 1, 2, \cdots, \tag{4.19}$$

such that

$$\mathcal{E} X_{kT}^2 \leqq M_k, \tag{4.20}$$

$$\lim_{k \to \infty} M_k = 0, \tag{4.21}$$

$$\Pr\{Z_{kT} \leqq z\} = F_{kT}(z) \to F_k(z), \quad \textit{as} \quad T \to \infty, \tag{4.22}$$

$$\lim_{k \to \infty} F_k(z) = F(z) \tag{4.23}$$

at every continuity point. Then

$$\lim_{T \to \infty} \Pr\{S_T \leqq z\} = F(z) \tag{4.24}$$

at every continuity point of $F(z)$.

The condition on X_{kT} is essentially that it converge stochastically to 0 uniformly in T.

REFERENCES

[1] P. H. DIANANDA, "Some probability limit theorems with statistical applications," *Proc. Cambridge Philos. Soc.*, Vol. 49 (1953), pp. 239–246.

[2] T. KOOPMANS, Ed., *Statistical Inference in Dynamic Economic Models—Cowles Commission Monograph 10*, John Wiley and Sons, New York, 1950.

[3] JOHN S. WHITE, "The limiting distribution of the serial correlation coefficient in the explosive case," *Ann. Math. Stat.*, Vol. 29 (1958), pp. 1188–1197.

---

[2] Theorem 4.4 and 4.5 were proved for the present paper before the author was aware of Diananda's results.