The Annals of Statistics
1977, Vol. 5, No. 6, 1250-1257

ESTIMATION IN THE FIRST ORDER MOVING AVERAGE MODEL BASED ON SAMPLE AUTOCORRELATIONS[1]

BY RAÚL PEDRO MENTZ

University of Tucumán

For the first order moving average we consider a proposal by Walker (*Biometrika*, 1961) to use k sample autocorrelations ($1 < k < T$, T sample size), to estimate the first autocorrelation of the model, and hence its basic parameter. When $k = k_T \to \infty$ as $T \to \infty$, the estimator is proved consistent and asymptotically normal and efficient, the latter provided k_T dominates $\log T$ and is dominated by $T^{\frac{1}{2}}$. An alternative form of the estimator facilitates the calculations and the analysis of the role of k, without changing the asymptotic properties.

1. Introduction. We consider the moving average time series model

$$y_t = \varepsilon_t + \alpha\varepsilon_{t-1}\,, \qquad t = \cdots, -1, 0, 1, \cdots, \tag{1.1}$$

where the ε_t are i.i.d. (independent identically distributed) normal $(0, \sigma^2)$, $0 < \sigma^2 < \infty$. Then (1.1) defines a stationary stochastic process with covariance sequence $\sigma_0 = \sigma^2(1 + \alpha^2)$, $\sigma_1 = \sigma_{-1} = \sigma^2\alpha$, $\sigma_j = 0$, $|j| > 1$, and autocorrelation sequence $\rho \equiv \rho_1 = \rho_{-1} = \alpha/(1 + \alpha^2)$, $\rho_j = 0$, $|j| > 1$. We further assume that $|\alpha| < 1$, which makes $|\rho| < \frac{1}{2}$. Only $|\alpha| \neq 1$ is important, since for $|\alpha^*| > 1$ the parameters $1/\alpha^*$ and σ^2/α^{*2} provide an equivalent parameterization. See, for example, Anderson (1971), Chapter 7.

To estimate α or ρ, we consider a sample $y_1, \cdots, y_T$ from (1.1). Consistent estimators of the σ_j and ρ_j are, respectively, the sample autocovariances $c_j = T^{-1} \sum_{t=1}^{T-|j|} y_t y_{t+|j|}$, and the sample autocorrelations $r_j = c_j/c_0$, ($r_1 \equiv r$). Since we are interested in convergence in probability and in distribution, and $r \to \rho$ in probability as $T \to \infty$, we shall take $|r| < \frac{1}{2}$ throughout. In fact $r \to \rho$ a.s. as $T \to \infty$ (see, for example, Hannan (1970), Chapter IV), but we do not use this result here.

The moment estimator obtained by solving for $\hat{\alpha}$ the equation $r = \hat{\alpha}/(1 + \hat{\alpha}^2)$, namely $\hat{\alpha} = [1 - (1 - 4r^2)^{\frac{1}{2}}]/(2r)$, is consistent for α but Whittle (1953) proved it is inefficient compared with the maximum likelihood estimator. Its inefficiency can be ascribed to that of r as an estimator of ρ, and Walker (1961) proposed to improve the asymptotic efficiency by correcting it in terms of $r_2, \cdots, r_k$, for some k sufficiently large.

Let $\mathbf{W}(\boldsymbol{\rho})$ be the $k \times k$ covariance matrix of the limiting normal distribution

Received April 1976; revised February 1977.

[1] Work at Stanford was supported by a research contract with the Office of Naval Research (T. W. Anderson, Project Director), Contract N00014-75-C-0442 (NR-042-034).

AMS 1970 *subject classifications*. 62M10, 62M99.

Key words and phrases. First order moving average model, approximate maximum likelihood estimation, consistent and asymptotically efficient estimators.

of $T^{\frac{1}{2}}(\mathbf{r}-\boldsymbol{\rho})$, where $\mathbf{r}=(r_1,\cdots,r_k)'$, $\boldsymbol{\rho}=(\rho_1,\cdots,\rho_k)'=(\rho,0,\cdots,0)'$, and let $\mathbf{W}(\mathbf{r})$ be the same with ρ_j replaced by r_j. If we partition $\mathbf{r}=(r,\mathbf{r}^{(2)})$, and

$$\mathbf{W}(\mathbf{r})=\begin{pmatrix} w_{11} & \mathbf{w}_{12}' \\ \mathbf{w}_{12} & \mathbf{W}_{22}\end{pmatrix}, \tag{1.2}$$

then $w_{11}=1-3r^2+4r^4$, $\mathbf{w}_{12}=(2r(1-r^2),r^2,0,\cdots,0)'$, and $\mathbf{W}_{22}$ has $1+2r^2$, $2r$ and r^2 as components with indices i,j for $|i-j|=0,1,2$, respectively, and 0 elsewhere.

Walker's estimator of ρ for model (1.1) is

$$\begin{aligned}\hat\rho &= \textstyle\sum_{j=0}^{k-1} m(j)r_{j+1} = r+\sum_{j=1}^{k-1} m(j)r_{j+1} = r-\mathbf{w}_{12}'\mathbf{W}_{22}^{-1}\mathbf{r}^{(2)} \\ &= r-2r(1-r^2)\textstyle\sum_{j=1}^{k-1} w^{1j}r_{j+1} - r^2\sum_{j=1}^{k-1} w^{2j}r_{j+1},\end{aligned} \tag{1.3}$$

where w^{ij} are the components of the inverse matrix $\mathbf{W}_{22}^{-1}$. Note that the $m(j)$ are random variables, functions of r.

Walker (1961) developed the asymptotic theory for his proposal when k is treated as fixed. In the following sections we present the corresponding theory when $k=k_T$, a function of the series length T, such that $\lim_{T\to\infty}k_T=\infty$. It was conjectured by Walker [(1961), page 353] that such a theory could be developed, essentially by means of the tools we use below, except that the w^{ij} will be evaluated explicitly.

Proofs will be simplified below; for full details see Mentz (1975). General comments are collected in Section 5.

2. The components in two rows of $\mathbf{W}_{22}^{-1}$.

From Mentz (1976) we have that

$$w^{ij}=[C_1(j)+iC_2(j)]x^i+[C_3(j)+iC_4(j)]x^{-i}, \qquad i,j=1,2,\cdots,T, \tag{2.1}$$

where $x=[-1+(1-4r^2)^{\frac{1}{2}}]/(2r)$, and for rows $i=1,2$,

$$\begin{aligned} C_1(1)&=a_{22}/\Delta+O_p(x^{2k}), & C_2(1)&=-a_{21}/\Delta+O_p(x^{2k}), \\ C_1(2)&=-a_{12}/\Delta+O_p(x^{2k}), & C_2(2)&=a_{11}/\Delta+O_p(x^{2k}),\end{aligned} \tag{2.2}$$

while $C_s(t)=O_p(x^{2k})$ for all other s and t, where $a_{11}=(r/2)[(1-4r^2)^{\frac{1}{2}}-3]$, $a_{12}=-(r/2)[(1-4r^2)^{\frac{1}{2}}+1]$, $a_{21}=-r^2$, $a_{22}=0$, $\Delta=h_1+O_p(x^{2k})$, $h_1=a_{11}a_{22}-a_{12}a_{21}\neq 0$. Note that $|x|<1$ for $|r|<\frac{1}{2}$.

3. Consistency.

We now consider the following.

THEOREM 1. *Let $\{y_t\}$ satisfy equation (1.1), where $|\alpha|<1$ and the ε_t are* i.i.d. *normal, $\mathscr{E}\varepsilon_t=0$, $\mathscr{E}\varepsilon_t^2=\sigma^2(0<\sigma^2<\infty)$ for all t. Suppose that a set of observations of $\{y_t\}$ at times $t=1,2,\cdots,T$ is available, and that $k=k_T$ is a function of T $(T\geqq k+1)$ satisfying $\lim_{T\to\infty}k_T=\infty$. Then if $\hat\rho$ is defined by (1.3), $p\lim_{T\to\infty}\hat\rho=\rho$.*

PROOF. Substituting (2.1) in (1.3) we delete terms having x^{2k} as a factor. Hence to prove that $\sum_{j=1}^{k-1}m(j)r_{j+1}\to_P 0$, we replace the $m(j)$ by $-\{2r(1-r^2)[C_1(1)+jC_2(1)]+r^2[C_1(2)+jC_2(2)]\}x^j$. From (2.2) it suffices to prove that

$\sum_{j=1}^{k-1} x^j r_{j+1} \to_P 0$ as $T \to \infty$. Now

$$P\{|\sum_{j=1}^{k-1} r_{j+1}x^j| > \varepsilon\} \leqq \sum_{j=2}^{n} P\{|r_j| > \varepsilon/2^n\} + P\left\{\frac{|x|^n}{1-|x|} > \varepsilon/2\right\}. \tag{3.1}$$

In the second term $|x| \to_P |\alpha| < 1$, and by proper choice of n (independently of k and T) the term can be made arbitrarily small. The first term tends to 0 since $r_j \to_P 0$ for $j = 2, \cdots, n$. A similar argument can be used with $\sum_{j=1}^{k-1} jx^j r_{j+1}$.

4. Asymptotic normality.

We now prove the following.

THEOREM 2. *Let the conditions of Theorem* 1 *hold together with*

$$\lim_{T\to\infty} k_T^{-1}\log T = 0, \qquad \lim_{T\to\infty} T^{-1}k_T^2 = 0. \tag{4.1}$$

Then as $T \to \infty$, $T^{\frac{1}{2}}(\hat\rho - \rho)$ has a limiting normal distribution with parameters 0 *and* $(1-\alpha^2)^3/(1+\alpha^2)^4$.

PROOF. The proof is done in three parts.

Part 1. (Replacement of r_j by c_j and simplification.)

$$\begin{aligned}\hat\rho - \rho &= \sum_{j=1}^{k} m(j-1)r_j - \rho \\ &\equiv c_0^{-1}\sum_{j=0}^{k} m(j-1)(c_j - \mathscr{E}c_j) - T^{-1}c_0^{-1}\sigma_1,\end{aligned} \tag{4.2}$$

where we introduced $m(-1) \equiv -\sigma_0^{-1}\sigma_1 = -\rho = -\alpha/(1+\alpha^2)$. Since $c_0 \to \sigma_0$ and $T^{\frac{1}{2}}T^{-1}c_0^{-1}\sigma_1 \to 0$ in probability as $T \to \infty$, we see that $T^{\frac{1}{2}}(\hat\rho - \rho)$ has the same limiting distribution as $T^{\frac{1}{2}}\sigma_0^{-1}\sum_{j=0}^{k} m(j-1)(c_j - \mathscr{E}c_j)$.

From the argument in Section 2 we have that $m(j) = m_1(j) + x^{\lambda k}m_2(j)$, $j = 1, 2, \cdots, k-1$, for some $0 < \lambda \leqq 2$ and functions m_1 and m_2. Hence

$$\begin{aligned}\hat\rho - \rho = \sigma_0^{-1}(&\textstyle\sum_{j=0}^{k} m_1(j-1)(c_j - \mathscr{E}c_j) \\ &+ x^{\lambda k}\textstyle\sum_{j=0}^{k} m_2(j-1)(c_j - \mathscr{E}c_j) - x^{\lambda k}m_2(-1)(c_0 - \mathscr{E}c_0)),\end{aligned} \tag{4.3}$$

where $m_2(-1) = m_1(-1)$. The first two terms have sums of the same nature, and it will be shown below that the first term normalized by $T^{\frac{1}{2}}$ has a limiting normal distribution. Since the second term has a factor $x^{\lambda k} \to 0$ in probability, this part will be completed if we show that $T^{\frac{1}{2}}|x|^{\lambda k} \to_P 0$. But $\log T/k_T \to 0$ assumed in (4.1) implies this result.

Hence the limiting distribution is unchanged if $m(j)$ is replaced by $m_1(j)$, where

$$m_1(-1) = -\alpha/(1+\alpha)^2, \qquad m_1(j) = x^j[1 + j(1-4r^2)^{\frac{1}{2}}], \quad j = 0, \cdots, k-1. \tag{4.4}$$

Part 2. (Substituting parameters for rv in the $m_1(j)$.) We now prove that if $\mu_1(j)$ denotes $m_1(j)$ with r replaced by ρ, then

$$p\lim_{T\to\infty} T^{\frac{1}{2}}\sum_{j=1}^{k-1}[m_1(j) - \mu_1(j)]c_{j+1} = 0, \tag{4.5}$$

where we used that $\mathscr{E}c_j = 0$ for $j > 1$. Let $\bar{x} \equiv x(\rho) = -\alpha$. Then

$$m_1(j) - \mu_1(j) = x^j[1 + j(1 - 4r^2)^{\frac{1}{2}}] - \bar{x}^j[1 + j(1 - 4\rho^2)^{\frac{1}{2}}]$$
$$(4.6) \qquad = [(1 - 4r^2)^{\frac{1}{2}} - (1 - 4\rho^2)^{\frac{1}{2}}]j\bar{x}^j + [1 + j(1 - 4r^2)^{\frac{1}{2}}](x^j - \bar{x}^j),$$

so that the rv in (4.5) will be taken to be formed by the corresponding two terms.

The sum on j of the first term in (4.6) gives a contribution of the form considered below, that is $T^{\frac{1}{2}} \sum_j \bar{x}_j c_{j+1}$. Since $\bar{x}^j = (-\alpha)^j$ is summable ($|\alpha| < 1$), such term converges in distribution to a normal rv with 0 expected value and finite variance. Further $(1 - 4r^2)^{\frac{1}{2}} \to (1 - 4\rho^2)^{\frac{1}{2}}$ in probability as $T \to \infty$, so that the contribution due to the first term of (4.6) converges stochastically to 0.

In the second term we are led to deal with $T^{\frac{1}{2}} \sum_{j=1}^{k-1} (x_j - \bar{x}_j)c_{j+1}$, or a similar expression with weights $j(x_j - \bar{x}_j)$. Using the same technique as in (3.1), together with the facts that $x \to_P \bar{x}$ and $T^{\frac{1}{2}}(c_2, \cdots, c_n)$ is asymptotically normally distributed with 0 expectations and finite variances, this expression is shown to converge to 0 in probability as $T \to \infty$.

Part 3. (The asymptotic distribution.) The conclusion of the preceding argument is that we have to find the limiting distribution of

$$(4.7) \qquad \Omega = T^{\frac{1}{2}}\sigma_0^{-1} \sum_{j=0}^{k-1} \mu_1(j-1)(c_j - \mathscr{E}c_j) \equiv T^{\frac{1}{2}} \sum_{t=1}^{T} W_t,$$

where $\mu_1(j) = \delta_j^*(-\alpha)^j[1 + j(1 - \alpha^2)/(1 + \alpha^2)]$, $j = 0, 1, \cdots, k-1$, $\delta_j^* = \frac{1}{2}$ for $j = -1$ and 1 otherwise, and we take $W_t = \sum_{j=0}^{k} \sigma_0^{-1}\mu_1(j-1)(y_t y_{t+j} - \mathscr{E}y_t y_{t+j})$, $t = 1, \cdots, T$. For $t = T - k + 1, \cdots, T$ the sums should range only up to $T - t$, but the simplification means adding a total of $k^2/2$ extra terms which is negligible compared with the existing Tk terms, because $k^2/T \to 0$ as $T \to \infty$. Taken as a stochastic process $\{W_t\}$ is stationary, finitely dependent of order $k + 1$, and finitely correlated of order 1.

We now proceed as in Anderson [(1971), pages 538–539]. Let $\{N\}$ be a sequence of integers such that $k/N \to 0$ as $T \to \infty$, and let $M = [T/N]$. Then (4.7) has the same limiting distribution as $M^{-\frac{1}{2}} \sum_{j=1}^{M} (Z_j + Y_j) + T^{-\frac{1}{2}}R$, where

$$Z_j = N^{-\frac{1}{2}} \sum_{i=1}^{N-k} W_{(j-1)N+i}, \qquad Y_j = N^{-\frac{1}{2}} \sum_{i=N-k+1}^{N} W_{(j-1)N+i}, \qquad j = 1, \cdots, M,$$

and $R = W_{NM+1} + \cdots + W_T$ ($R \equiv 0$ if $NM = T$). This random variable has the same limiting distribution as

$$(4.8) \qquad \Omega^* = M^{-\frac{1}{2}} \sum_{j=1}^{M} Z_j,$$

which is proved by showing that the other terms converge to 0 in mean square. Next we have

$$(4.9) \qquad \mathscr{E}Z_j^2 = N^{-1}\{(N-k)\mathscr{E}W_1^2 + (N-k-1)2\mathscr{E}W_1W_2\},$$

and let us write

$$\Omega^* = (\mathscr{E}Z_1^2) \sum_{j=1}^M \frac{Z_j}{(M\mathscr{E}Z_j^2)^{\frac{1}{2}}} \equiv (\mathscr{E}Z_1^2) \sum_{j=1}^M z_j \, . \tag{4.10}$$

Then $\mathscr{E} \sum z_j = 0$, $\mathscr{E} \sum z_j^2 = 1$. To use Liapounov's central limit theorem [Loève (1963), Chapter VI] it suffices to prove that for some $\delta > 0$, $\lim_{T\to\infty} \sum \mathscr{E}|z_j|^{2+\delta} = 0$. We choose $\delta = 2$. Then it suffices to prove that

$$\sum_{j=1}^M \mathscr{E}z_j^4 = \sum_{j=1}^M \mathscr{E} \frac{Z_j^4}{M^2(\mathscr{E}Z_j^2)^2} = \frac{\mathscr{E}Z_1^4}{M(\mathscr{E}Z_1^2)^2} \tag{4.11}$$

converges to 0 as $T \to \infty$, where $\mathscr{E}Z_1^4 = N^{-2} \sum_{s,t,q,v=1}^{N-k} \mathscr{E}W_s W_t W_q W_v$. Proceeding as in Anderson [(1971), page 539] or Berk [(1974), page 498], we conclude that we only need to show that those terms with $v = t$, $|t - s| \leqq k + 1$, $|s - q| \leqq k + 1$, $t < s < q$ provide contributions that tend to 0 as $T \to \infty$. There are at most $4(N - k)(k + 1)^2$ such terms, and the total contribution to (4.11) is bounded by a constant times $4\sigma^8(N - k)(k + 1)^2 M^{-1}(\mathscr{E}Z_1^2)^{-2}\{\sum_{j=0}^k \mu_1(j - 1)\}^4$. This completes the proof because this tends to 0 as $T \to \infty$, by (4.1), and means that $T^{\frac{1}{2}}(\hat{\rho} - \rho)$ is asymptotically normal with expectation 0 and variance

$$\lim_{T\to\infty} \mathscr{E}Z_1^2 = \lim_{T\to\infty} (\mathscr{E}W_1^2 + 2\mathscr{E}W_1 W_2) \, . \tag{4.12}$$

It remains to prove that (4.12) equals $(1 - \alpha^2)^3/(1 + \alpha^2)^4$. This is done by direct but rather laborious calculation, and the details are omitted here.

COROLLARY 1. *Under the conditions of Theorem* 2, *let $\hat{\alpha}'$ be the moment estimator with r replaced by $\hat{\rho}$. Then $T^{\frac{1}{2}}(\hat{\alpha}' - \alpha)$ has a limiting normal distribution with parameters* 0 *and* $1 - \alpha^2$.

5. Comments. We here collect several comments about Walker's proposal and our findings.

a. The estimator (1.3) can be interpreted as an improvement in the asymptotic efficiency of r, a consistent estimator of ρ, by approximating the maximum likelihood estimator by means of a linear combination of $r, r_2, \cdots, r_k$, with random coefficients. Since $T^{\frac{1}{2}}(r - \rho)$ tends in distribution to a normal with variance $1 - 3\rho^2 + 4\rho^4 = (1 - \alpha^2)^3(1 + \alpha^2)^{-4} + \{4\alpha^2 + \alpha^4(1 + \alpha^2)\}(1 + \alpha^2)^{-4}$, the improvement achieved by $\hat{\rho}$ is the second term, the first term being also the asymptotic variance of the maximum likelihood estimator of ρ. In terms of α, comparing the moment estimator with the same with r replaced by $\hat{\rho}$, the variance of the limiting distribution of the former is [cf. Whittle (1953)] $(1 + \alpha^2 + 4\alpha^4 + \alpha^6 + \alpha^8)(1 - \alpha^2)^{-2} = 1 - \alpha^2 + \alpha^2\{4 + \alpha^2(1 + \alpha^2)^2\}(1 - \alpha^2)^{-2}$, where $1 - \alpha^2$ is also the asymptotic variance of the maximum likelihood estimator of α.

Walker's approach to derive (1.3) is to work with the limiting normal distribution of $T^{\frac{1}{2}}(\mathbf{r} - \boldsymbol{\rho})$ and use maximum likelihood ideas. Since some approximations are introduced, the final estimator comes closer to approximating the least-squares estimator, the Jacobian being omitted. These approximations have no relevance for the asymptotic theory we developed, but may be important in small samples.

b. The estimator (1.3) is consistent and asymptotically efficient, proved biased for small samples and "a priori" appears as robust to departures from normality in the distribution of the ε_t, for moderate sample sizes.

The conclusions for large samples follow from the studies by Walker (1961) [see also Anderson, (1971), Section 5.7.2] and the present paper. Small sample studies by Monte Carlo trials were made by Walker (1961) and more extensively by McClave (1974).

In the small-sample studies ($T = 100$), (1.3) showed considerable efficiency (i.e., agreement with the large-sample variances), but also rather important biases. Walker (1961) proposed a correction for bias that has not been studied empirically in detail.

The robustness argument arises because only the limiting normal distribution of $T^{\frac{1}{2}}(\mathbf{r} - \boldsymbol{\rho})$ enters in the derivation of (1.3), and it is well known that the same limiting distribution holds for a wide class of distributions of the ε_t. Unfortunately no empirical results for small samples are available in this connection.

c. If in a practical situation $\mathscr{E}y_t$ were unknown, it would be estimated from the data. Then c_j would be replaced by $c_j^* = T^{-1} \sum_{t=1}^{T-|j|} (y_t - \bar{y})(y_{t+|j|} - \bar{y})$, where $\bar{y} = T^{-1} \sum_{t=1}^{T} y_t$; see, for example, Anderson (1971), for this and other types of mean corrections. It is easily proved that our results hold for the modified version of the estimators, since, for example $T^{\frac{1}{2}} \sum_{j=0}^{k} \{m(j-1)(c_j - \mathscr{E}c_j) - m^*(j-1)(c_j^* - \mathscr{E}c_j^*)\} \to_P 0$ as $T \to \infty$, where $m^*(j-1)$ is $m(j-1)$ with c_j replaced by c_j^*.

d. Our analysis has been restricted to the first-order moving average model. Extension of the approach to a model of order $q > 1$ seems quite feasible. The components of the $\mathbf{W}$ matrix in (1.2) are known for all q. $\mathbf{W}_{22}$ will be a Toeplitz matrix with equal elements along its central diagonals, and zeroes elsewhere; the components of the inverses of such matrices are given as functions of the roots of the associated polynomial equation in Mentz (1976). $\mathbf{W}_{22}$ is positive definite and can therefore be taken as the covariance matrix of a stationary moving average process; by the argument in Anderson [(1971), pages 224–225] half of the roots are larger and half are less than one in absolute value, as was the case in Section 2.

6. A modification to simplify the computations. From the argument in the proof of Theorem 2, it follows that

$$\hat{\rho}^* = \sum_{j=0}^{k-1} m_1(j) r_{j+1}, \tag{6.1}$$

where $m_1(j)$ was defined in (4.4), has the same large-sample properties of (1.3). It discards parts having x^k as dominating factor, and hence differs only slightly from $\hat{\rho}$ if k is moderately large.

From a practical point of view (6.1) makes easy the choice of k, guided by the degree of numerical approximation that is desired. Consider Table 1 where for r negative the values of $m_1(j)$ are those in the table all taken with positive signs. Once the numerical value of r is available, the table can be used to see how many sample autocorrelations r_j, $2 \leqq j \leqq k$ to include in (6.1).

Using data generated in a computer, Walker (1961) studied in detail a set of $T = 100$ observations from (1.1) with $\alpha = 0.5$ ($\rho = 0.4$). He estimated $r = 0.35005$, $r_2 = -0.06174$, $r_3 = -0.08007$, $r_4 = -0.14116$ and $r_5 = -0.15629$. Then his estimates for $1 \leqq k \leqq 5$ are compared in Table 2 with $\hat{\rho}^*$. Only r_j up to $j = 5$ are provided by Walker. From Table 1 for $r = 0.35$ it is apparent that a larger k will be called for. However the behavior of $\hat{\rho}^*$ seems comparable to that of $\hat{\rho}$.

TABLE 1
Values of $m_1(j)$ for selected values of r

j	.05	.15	.25	.35	.45
1	−.1000000	−.3000000	−.5000000	−.7000000	−.9000000
2	.0075125	.0685482	.1961524	.4049504	.7353557
3	−.0005018	−.0139772	−.0692193	−.2140023	−.5682477
4	.0000314	.0026761	.0230114	.1072520	.4234477
5	−.0000018	−.0004922	−.0073620	−.0519085	−.3075804
6	.0000001	.0000880	.0022931	.0245097	.2192185
7	0.0000000	−.0000154	−.0007003	−.0113615	−.1539701
8	0.0000000	.0000026	.0002106	.0051919	.1068903
9	0.0000000	−.0000004	−.0000626	−.0023457	−.0735060
10	0.0000000	0.0000000	.0000184	.0010500	.0501521
11	0.0000000	0.0000000	−.0000053	−.0004664	−.0339916
12	0.0000000	0.0000000	.0000015	.0002058	.0229082
13	0.0000000	0.0000000	−.0000004	−.0000903	−.0153631
14	0.0000000	0.0000000	.0000001	.0000394	.0102590
15	0.0000000	0.0000000	0.0000000	−.0000171	−.0068249

TABLE 2
Estimates of ρ, $T = 100$, Walker's data

k	$\hat{\rho}$	$\hat{\rho}^*$
2	0.38051	0.39327
3	0.36879	0.36084
4	0.38498	0.39106
5	0.37934	0.37429

To compute (1.3) exactly, Walker (1961) proposed an iterative procedure; (6.1) is of course much simpler. In fact (2.2) could be written in detail to give the exact form of (1.3), and a table similar to Table 1 prepared for that case, but we omit these details here.

Acknowledgments. This study is a part of the author's Ph. D. dissertation in the Department of Statistics, Stanford University. Appreciation is expressed to Professor T. W. Anderson for generous guidance and help while directing this work. The author is also indebted to a referee, an associate editor and the editor for useful comments.

REFERENCES

ANDERSON, T. W. (1971). *The Statistical Analysis of Time Series.* Wiley, New York.

BERK, K. N. (1974). Consistent autoregressive spectral estimates. *Ann. Statist.* **2** 489–502.

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

LOÈVE, MICHEL (1963). *Probability Theory* (3rd ed.) Van Nostrand, New York.

MCCLAVE, JAMES T. (1974). A comparison of moving average estimation procedures. *Comm. Statist.* **3** 865–883.

MENTZ, RAÚL P. (1975). Asymptotic properties of some estimators in moving average models. Technical Report No. 21, Department of Statist., Stanford Univ.

MENTZ, RAÚL P. (1976). On the inverse of some covariance matrices of Toeplitz type. *SIAM J. Appl. Math.* **31** 426–437.

WALKER, A. M. (1961). Large sample estimation of parameters for moving average models. *Biometrika* **48** 343–357.

WHITTLE, P. (1953). Estimation and information in stationary time series. *Ark. Mat.* **2** 423–434.

INSTITUTE OF STATISTICS
UNIVERSITY OF TUCUMÁN
CASILLA DE CORREO 209
TUCUMÁN, ARGENTINA