Qüestiió - V. 8, n.º 4 (desembre 1984) pp. 155-163

SOME ASPECTS OF PARAMETER INFERENCE FOR NEARLY NONSTATIONARY AND NEARLY NONIN VERTIBLE ARMA MODELS II

S. AH TOLA, G. C. TIAO**
* UNIVERSITY OF HELSINKI
** UNIVERSITY OF CHICAGO

This article will extend the discussion in Ahtola and Tiao (1984a) of the finite sample distribution of the score function in nearly nonstationary first order autoregressions to nearly noninvertible first order moving average models. The distribution theory can be used to appreciate the behavior of the score function in situations where the asymptotic normal theory is known to give poor approximations in finite samples.

The approximate distributions suggested here can be used to test for the value of the moving average parameter when it is close to unity. In particular, a test for noninvertibility can be obtained with an exact finite sample distribution of the test statistic under the null hypothesis.

Keywords: APPROXIMATE DISTRIBUTION, CHI-SQUARE, FINITE SAMPLE, INVERTIBILITY, MINIMUM EIGENVALUE, MOVING AVERAGE, NONNORMALITY.

1. INTRODUCTION.

Parameter inference of ARMA models relies heavily on asymptotic distribution theory for the estimators of the parameters. To obtain asymptotic normality for the least squares estimators, say, both stationarity and invertibility are assumed. This asymptotic theory is known to break down if the autoregressive operator has roots on the unit circle. Fuller /10/, Dickey and Fuller /7/, Rao /18/, Evans and Savin /8/, Hasza and Fuller /11/, and Ahtola and Tiao /2/ provide theory to cover some cases when the model is purely autoregressive. The asymptotic theory for noninvertible models is yet to be developed. The presence of the moving average part in general, invertible or not, brings in a substantial complication for the estimation of the parameters, compared to the purely autoregressive models, where ordinary least squares can be used. Some sort of iterative procedures are called for and closed form expressions for the estimators are consequently not obtainable.

A serious problem arises when asymptotic normal distributions are used for finite sample inference, even when the true model is stationary and invertible. Evans and Savin /9/, report calculations with a first order autoregression, which clearly indicate that normality is a poor approximation for the least squares estimator of the autoregressive parameter even for large sample sizes, when the true parameter is near but below unity. Similar problems, but possibly somewhat less pronounced, are reported in Fuller /10/ (p.357) in connection with a nearly noninvertible first order moving average model. A common feature in both cases is the fact that the distributions become skewed. An additional difficulty with the finite sample distribution of an estimator of the moving average parameter is that the usually applied estimation techniques give estimates exactly on the noninvertibility boundary with positive probability; this probability being the bigger the closer to the noninvertibility boundary the true parameter is. Therefore the finite sample distribution of the estimator is of a mixed type and not absolutely continuous

- J. Ahtola* - Department of Economics - University of Helsinki - Helsinki - Finland
- G.C. Tiao** - Graduate School of Business - University of Chicago - Chicago - U.S.A.

- Article rebut el octubre del 1984.

everywhere. See Cryer and Ledolter /6/

1.1. THE CASE OF A FIRST ORDER AUTOREGRESSION.

To obtain some useful insights into the finite sample inference problem we suggest in Ahtola and Tiao /1/ the use of the score function and discuss the distribution theory associated with it in a first order autoregression in nearly nonstationary situations. More specifically, we consider the model

$$y_t = \phi y_{t-1} + a_t \,, \qquad (1.1)$$

where $y_o = 0$, $a_t \sim NID(0,1)$, $t = 1,2,\ldots,n$

The general case of $var(a_t) = \sigma^2$ poses no difficulties and is dealt with in Section 6. The score function, that is, the derivative of the loglikelihood function with respect to ϕ is readily written as

$$S_\phi(y) = \sum_{t=1}^{n} y_{t-1}a_t = \sum_{t=1}^{n} y_{t-1}^2(\hat{\phi} - \phi) \qquad (1.2)$$

where $\hat{\phi}$ is the least squares estimator of ϕ. $S_\phi(y)$ has a convenient quadratic form representation which can further be decomposed into a linear combination of independent and identically distributed chi-square variables with one degree of freedom. Specifically,

$$S_\phi(y) = a'Ca = \sum_{j=1}^{n} \eta_{jn}\chi_j^2(1) \,,$$

where $a = (a_1,\ldots,a_n)'$,

$$C = \frac{1}{2}\begin{bmatrix} 0 & 1 & \phi & \cdots & \phi^{n-2} \\ 1 & \cdot & & & \vdots \\ \phi & & \cdot & & \phi \\ \vdots & & & \cdot & 1 \\ \phi^{n-2} & \cdots & \phi & 1 & 0 \end{bmatrix}$$

and η_{jn}, $j=1,\ldots,n$ are the eigenvalues of C.

When $\phi=1$, the eigenvalues are

$$\eta_{jn} = -\frac{1}{2}\,, \quad j = 1,2,\ldots,n-1$$
$$\eta_{nn} = \frac{n-1}{2}$$

For $|\phi| < 1$ closed form expressions for η_{jn} are difficult to obtain. When normalized by the inverse of the standard deviation of $S_\phi(y)$,

$$c(\phi,n) = \begin{cases} (1-\phi^2)^{1/2}[n-(1-\phi^2)^{-1}(1-\phi^{2n})]^{-1/2}, & |\phi| < 1 \\ [\frac{1}{2}n(n-1)]^{-1/2}, & \phi = 1 \end{cases}$$

such that the score has variance equal to 1, $c(\phi,n)S_\phi(y)$ can be decomposed into two terms

$$c(\phi,n)S_\phi(y) = (Y - \beta X) + (1+\beta)X \qquad (1.3)$$

where

$$Y = -m\sum_{j=1}^{n}\chi_j^2(1), \qquad X = \sum_{j=1}^{n}(\lambda_{jn}+m)\chi_j^2(1)$$

$$\lambda_{jn} = c(\phi,n)\eta_{jn}, \qquad -m = \min_j(\lambda_{jn})$$

and

$\beta = \frac{-2m^2n}{1+2m^2n}$ is the regression coefficient of Y on X.

Note that $Y - \beta X$ and $(1+\beta)X$ are uncorrelated. $Y - \beta X$ is aproximately distributed as $-m\chi^2(nd^{-1})$, and $(1+\beta)X$ is approximately distributed as $(2mn)^{-1}\chi^2(2(mn)^2d^{-1})$, where $d = 1+2m^2n$.

It can also be shown that an asymptotic approximation to $-m$ is

$$-m = -(1+\phi)^{-1}c(\phi,n).$$

This approximation is exact when $\phi = 1$, for any n, and it performs very well for ϕ values dawn to .6 with sample sizes of, say, 25 and bigger. Therefore the normalized score function has a readily obtainable approximate expression as a sum of two uncorrelated weighted χ^2 variables. Note that this expression is a function of only two parameters m and n. The approximate expression is very convenient in the interpretation of the exact behavior of the score function, when ϕ and n are varied. In particular, the finite sample behavior of the score when ϕ approaches unity becomes very clear, and in fact, when ϕ is unity the approximate finite sample distribution is exact.

The plan of the rest of this article is as follows. In Section 2 we show how the results concerning the behavior of the score function in the first order autoregression can be extended to characterize the distribution of the score function in the first order moving average model, when the moving average parameter is close to unity. In Section 3 we give an interpretation for the score as a function of an estimator minus the true parameter. This interpretation gives additional intuition to why the score function is a natural test statistic to look at. Section 4 contains the interpretation of the transition of the distribution of the score when the moving average parameter approaches unity. In Section 5 we present tables for inference in practice, and Section 6 contains some concluding comments.

2. DISTRIBUTION OF THE SCORE IN A FIRST ORDER MOVING AVERAGE MODEL

Consider the model

$$y_t = a_t - \theta a_{t-1}, \qquad (2.1)$$

where $a_o = 0$, $a_t \sim NID(0,1)$, $t = 1,2,\dots,n$.

Now we can express a_t as

$$a_t = y_t + \theta y_{t-1} + \dots + \theta^{t-1} y_1 \qquad (2.2)$$

The loglikelihood function is

$$l(\theta|y) = $$
$$= -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{n}(y_t + \theta y_{t-1} + \dots + \theta^{t-1}y_1)^2$$

The score function, $S_\theta(y)$, is now

$$S_\theta(y) = -\sum_{t=1}^{n}(y_t + \theta y_{t-1} + \dots + \theta^{t-1}y_1)(y_{t-1} + 2\theta y_{t-2} + \dots + (t-1)\theta^{t-2}y_1) \qquad (2.3)$$

Using (2.2) it is readily seen that

$$S_\theta(y) = -\sum_{t=1}^{n} a_t(a_{t-1} + \theta a_{t-2} + \dots + \theta^{t-2}a_1) \qquad (2.4)$$

Recall that the score function in the first order autoregression is $\Sigma a_t y_{t-1}$. However, y_{t-1} then has an expression

$$y_{t-1} = a_{t-1} + \phi a_{t-2} + \dots + \phi^{t-2}a_1$$

therefore we immediately see that $S_\theta(y)$ is the negative of the score function of the first order autoregression (1.1) with parameter θ. This fact implies that we get (see (1.3))

$$c(\theta,n)S_\theta(y) = -(1+\beta)X + (\beta X - Y)$$
$$\doteq -a_2\chi^2(b_2) + a_1\chi^2(b_1) \qquad (2.5)$$

where

$$a_1 = m, \qquad b_1 = nd^{-1}$$
$$a_2 = (2mn)^{-1}, \qquad b_2 = 2(mn)^2 d^{-1},$$
$$d = 1 + 2m^2 n, \qquad \text{and}$$
$$-m \doteq -(1+\theta)^{-1}c(\theta,n), \qquad (2.6)$$

and the two χ^2's are uncorrelated. In particular, when $\theta = 1$, all the approximations in (2.5) and (2.6) become exact and furthermore the two χ^2's are independent. We therefore have an exact result for $\theta = 1$.

$$c(\theta,n)S_\theta(y) = -[(n-1)(2n)^{-1}]^{1/2}\chi^2(1) + [2n(n-1)]^{1/2}\chi^2(n-1) \qquad (2.7)$$

where the two chi-squares are independent.

3. AN INTERPRETATION OF $S_\theta(y)$ AS A FUNCTION OF $\hat{\theta} - \theta$

In the autoregressive case it is straightforward to see that

$$S_\phi(y) = \sum_{t=1}^{n} y_{t-1}^2(\hat{\phi} - \phi) \qquad (3.1)$$

where $\hat{\phi}$ is the least squares estimator of ϕ. Therefore $S_\phi(y)$ offers a natural test statistic from the estimation point of view.

Since the least squares estimation in the moving average case is much less straightforward, expressions like (3.1) are not readily available. However, if we define an iterative least squares estimation procedure as in Fuller /10/ (p.344) (see also Macpherson and Fuller /13/, we can obtain an analogue to (3.1), but naturally only in an approximative sense. To this effect, let us denote (2.2) as

$$a_t(y,\theta) = y_t + \theta y_{t-1} + \ldots + \theta^{t-1} y_1 \qquad (3.2)$$

Define the least squares estimator of θ as the value which minimizes

$$Q_n(\theta) = \sum_{t=1}^{n} \{a_t(y,\theta)\}^2 \qquad (3.3)$$

The minimization of (3.3) is necessarily iterative, and let us denote by $\tilde{\theta}$ some iterate which can also be a suitable initial estimate. Expand now $a_t(y,\theta)$ as

$$a_t(y,\theta) = a_t(y,\tilde{\theta}) + W_t(y,\tilde{\theta})(\theta-\tilde{\theta}) + r_t(y,\tilde{\theta}) \qquad (3.4)$$

where $W_t(y,\tilde{\theta}) = \left.\frac{\partial a_t(y,\theta)}{\partial\theta}\right|_{\theta=\tilde{\theta}}$, and $r_t(y,\tilde{\theta})$ is the remainder term.

We can write (3.4) as

$$a_t(y,\tilde{\theta}) = -W_t(y,\tilde{\theta})(\theta-\tilde{\theta}) + r_t(y,\tilde{\theta}) + a_t \qquad (3.5)$$

which suggests regression of $a_t(y,\tilde{\theta})$ on $-W_t(y,\tilde{\theta})$ to obtain an estimator of $\theta-\tilde{\theta}$. Denote the new estimator obtained from this regression by $\hat{\theta} = \tilde{\theta} + \Delta\hat{\theta}$, where from above

$$\Delta\hat{\theta} = \frac{-\sum W_t(y,\tilde{\theta}) a_t(y,\tilde{\theta})}{\sum\{W_t(y,\tilde{\theta})\}^2} \qquad (3.6)$$

After iterating until convergence is achieved and arguing asymptotically, we obtain from (3.6) and (3.5), (see also Fuller /10/ (pp.347-348))

$$\hat{\theta} - \theta \doteq \frac{-\sum W_t(y,\theta) a_t}{\sum\{W_t(y,\theta)\}^2} \qquad (3.7)$$

where $\doteq$ denotes asymptotic equivalence. Therefore,

$$\sum_{t=1}^{n} \{W_t(y,\theta)\}^2 (\hat{\theta}-\theta) \doteq -\sum_{t=1}^{n} W_t(y,\theta) a_t \qquad (3.8)$$

Since the right hand side of (3.8) equals $S_\theta(y)$, (3.8) is now the analogue of (3.1). The two formulas are easily seen to be in agreement with each other, since in the autoregressive case

$$-W_t(y,\phi) = y_{t-1} \qquad (3.9)$$

which effectively allows the simple, exact formula (3.1).

To avoid possible confusion, we might note that the above iterative least squares procedure is not needed to obtain a test statistic for testing, say, $H_o: \theta = \theta_o$ vs. $H_1: \theta \neq \theta_o$. The test statistic can simply be obtained from (2.4) by calculating $c(\theta_o,n) S_{\theta_o}(y)$. Therefore no "estimation under the alternative hypothesis" is needed. This is a very tractable feature of testing procedures based on the idea of efficient score or Lagrange multiplier in general.

4. INTERPRETATION OF THE FINITE SAMPLE DISTRIBUTION OF $c(\theta,n)S_\theta(y)$

As a point of reference we have computed percentiles of the exact distribution of $c(\theta,n)S_\theta(y)$ by numerical methods for various θ and n. Numerical evaluation of the eigenvalues η_{jn} and numerical integrations to obtain $pr(\sum_{j=1}^{n} \lambda_{jn}\, \chi_j^2(1) \le x_o)$ for appropriate x_o were required. Table 1 summarizes these computations.

The distribution is seen to be very skewed to the left when θ is close to unity for all sample sizes. For large samples the skewness diminishes faster than for small samples when θ moves away from unity. This is the well known "central limit effect", which is known to work with invertible models.

The approximation (2.5) yields an intuitively appealing device to interpret the behavior of the distribution of $c(\theta,n)S_\theta(y)$. Recall that the exact distribution is that of a linear combination of n independently distributed chi-square random variables each with one degree of freedom. Therefore, in some sense, we have a total of n degrees of freedom. Now the degrees of freedom of the two approximating component chi-squares in (2.5) add up to n, and hence the approximation can be thought of as allocating the total nomber of degrees of freedom between these two components.

To see how the negative component of the approximation dominates the sum when θ is close to unity, let us denote

$$\mathrm{Var}(\beta X - Y) = w \quad \text{and} \quad \mathrm{Var}(-(1+\beta)X) = 1 - w,$$

where $w = (1+2m^2n)^{-1}2m^2n$. Then the degrees of freedom b_1 and b_2 in (2.5) can be written as $b_1 = n(1-w)$ and $b_2 = nw$. Thus, 1-w measures the portion of the total variability accounted for by the negative component of (2.5) and nw measures its skewness.

Tables 2a and 2b give, respectively, values of 1-w and nw for various values of θ.

From $\theta = .8$ and up the dominance of $-(1+\beta)X$ is very marked and also the skewness is seen to increase when θ increases and/or n decreases. For fixed n, we see a smooth transition in the behavior of the approximation (2.5) as θ decreases from unity. This corresponds directly with the exact behavior of $c(\theta,n)S_\theta(y)$ in Table 1.

In summary, the behavior of the distribution of $c(\theta,n)S_\theta(y)$ can be conveniently summarized as a linear combination of two chi-square random variables. The characteristics of this linear combination are determied by m and n, since (a_1,b_1) and (a_2,b_2) are functions of only these two quantities. In particular, the two functions 1-w and nw together illustrate in very simple terms the nature of the distribution in the transitional situation when θ is close to unity as well as the inadequacy of using asymptotic normality as an approximation for finite n in such a situation.

5. APPROXIMATE FINITE SAMPLE DISTRIBUTION FOR $c(\theta,n)S_\theta(y)$

We have used (2.5) to interpret and appreciate the distribution of $c(\theta,n)S_\theta(y)$ when θ and n vary rather than suggesting it as an approximate distribution for practical use.

A somewhat better overall fit to the exact distribution can be obtained by using approximations

$$-(1+\beta)X \doteq -bF(\nu_1,\nu_2) \qquad (5.1)$$

$$\beta X - Y \doteq N(\mu,w) \qquad (5.2)$$

where the parameters of the scaled F-distribution are obtained by equating the first three moments on both sides of (5.1), and μ is the approximate mode of βX-Y. For details of the evaluation of these parameters and the motivation behind these approximations the reader is referred to Ahtola and Tiao /1/.

We now present percentiles of the approximation

$$-(1+\beta)X + (\beta X-Y) \doteq -b\,F(\nu_1,\nu_2) + N(\mu,w) \qquad (5.3)$$

where the F part and the normal part are treated as independent. We call this approximation the F-Normal approximation. We also compare this approximation to the exact distribution by giving the exact cumulative probabilities at the percentiles of the approximation.

Table 3 summarizes the performance of the approximation for various n and θ. The lower tail is generally extremely well approximated. Although the upper tails show some discrepancies particularly when n = 100 and θ is around .99, the F-Normal approximation seems to work well enough for practical use and performs far better than the asymptotic normal approximation.

6. CONCLUSIONS.

We have demonstrated how the distributions of the score functions of the first order autoregression and the first order moving average model are in close connection with each other. The discussion was carried through by assuming the disturbance variance, $\mathrm{var}(a_t)$ equal to 1. It is straightforward to see that if $\mathrm{var}(a_t) = \sigma^2$, the test statistic $c(\theta,n)S_\theta(y)$ need to be divided by σ^2 and all the previous results apply. Usually it is the case that σ^2 is unknown, in which case σ^2 must be substituted for by some estimate for practical applications. Simulation studies indicate that using e.g. the MLE of σ^2 has no practically significant

TABLE 1 Exact Percentiles of $c(\theta,n)S_{\theta}(y)$

(a) n = 25

θ	.010	.025	.050	.100	.900	.950	.975	.990
.6	-3.01	-2.34	-1.81	-1.28	1.12	1.41	1.66	1.95
.8	-3.36	-2.52	-1.88	-1.27	1.01	1.24	1.43	1.66
.9	-3.66	-2.65	-1.92	-1.22	.92	1.09	1.24	1.41
.95	-3.84	-2.75	-1.96	-1.20	.84	.98	1.10	1.25
.975	-3.90	-2.79	-1.98	-1.20	.80	.92	1.03	1.15
.99	-3.92	-2.80	-1.98	-1.20	.78	.89	.99	1.10
.995	-3.92	-2.80	-1.99	-1.20	.78	.89	.98	1.10

(b) n = 50

θ	.010	.025	.050	.100	.900	.950	.975	.990
.6	-2.83	-2.25	-1.79	-1.30	1.17	1.46	1.71	2.00
.8	-3.12	-2.41	-1.86	-1.30	1.09	1.32	1.51	1.74
.9	-3.43	-2.56	-1.91	-1.28	1.00	1.17	1.32	1.48
.95	-3.70	-2.63	-1.94	-1.23	.89	1.02	1.14	1.26
.975	-3.87	-2.77	-1.97	-1.21	.81	.92	1.01	1.11
.99	-3.94	-2.81	-1.99	-1.20	.76	.85	.92	1.00
.995	-3.95	-2.82	-2.00	-1.20	.75	.83	.89	.97

(c) n = 100

θ	.010	.025	.050	.100	.900	.950	.975	.990
.6	-2.68	-2.17	-1.75	-1.30	1.21	1.51	1.77	2.06
.8	-2.91	-2.30	-1.82	-1.31	1.16	1.40	1.61	1.84
.9	-3.17	-2.44	-1.88	-1.31	1.08	1.28	1.44	1.62
.95	-3.46	-2.58	-1.93	-1.28	.98	1.13	1.25	1.38
.975	-3.72	-2.70	-1.95	-1.24	.88	.99	1.08	1.17
.99	-3.91	-2.80	-1.99	-1.20	.78	.86	.92	.99
.995	-3.96	-2.83	-2.00	-1.20	.74	.81	.86	.92

TABLE 2a

Values of $1-w = Var(1+\beta)X$

θ \ n	100	50	25
.6	.663	.660	.652
.8	.814	.810	.800
.9	.900	.895	.882
.95	.950	.939	.823
.975	.969	.961	.943
.99	.983	.973	.953
.995	.986	.977	.957

TABLE 2b

Values of $nw = b_2$

θ \ n	100	50	25
.6	33.684	17.021	8.696
.8	18.605	9.524	5.000
.9	10.000	5.263	2.938
.95	5.405	3.026	1.906
.975	3.072	1.940	1.427
.99	1.748	1.348	1.163
.995	1.353	1.168	1.080

TABLE 3   Percentiles of the F-Normal Approximation

(a) n = 25

θ	.010	.025	.050	.100	.900	.950	.975	.990
.6	-2.99	-2.40	-1.94	-1.47	.98	1.26	1.50	1.77
.8	-3.29	-2.52	-1.96	-1.40	.95	1.19	1.37	1.58
.9	-3.53	-2.61	-1.95	-1.33	.89	1.07	1.22	1.38
.95	-3.69	-2.74	-1.96	-1.28	.84	.98	1.09	1.22
.975	-3.80	-2.74	-1.98	-1.26	.78	.90	1.00	1.11
.99	-3.89	-2.80	-2.01	-1.26	.74	.84	.93	1.02
.995	-3.92	-2.82	-2.02	-1.26	.72	.82	.90	.99
	Exact cumulative probabilities at the above percentiles							
.6	.010	.023	.042	.078	.863	.928	.961	.982
.8	.011	.025	.046	.086	.878	.941	.969	.986
.9	.011	.026	.049	.090	.889	.946	.973	.988
.95	.011	.027	.050	.093	.900	.951	.974	.988
.975	.011	.026	.050	.095	.890	.944	.970	.986
.99	.010	.025	.049	.095	.875	.930	.962	.980
.995	.010	.025	.049	.095	.863	.923	.955	.977

(b) n = 50

θ	.010	.025	.050	.100	.900	.950	.975	.990
.6	-2.84	-2.31	-1.89	-1.44	1.07	1.37	1.62	1.90
.8	-3.09	-2.42	-1.91	-1.39	1.05	1.29	1.50	1.73
.9	-3.35	-2.52	-1.92	-1.33	.98	1.18	1.34	1.51
.95	-3.55	-2.60	-1.92	-1.28	.91	1.06	1.18	1.31
.975	-3.70	-2.66	-1.93	-1.24	.84	.96	1.05	1.14
.99	-3.83	-1.74	-1.97	-1.23	.77	.86	.93	1.01
.995	-3.89	-2.79	-1.99	-1.23	.73	.81	.88	.95
	Exact cumulative probabilities at the above percentiles							
.6	.010	.023	.043	.082	.875	.937	.968	.986
.8	.010	.025	.047	.090	.887	.945	.974	.990
.9	.011	.026	.050	.094	.894	.952	.978	.991
.95	.011	.027	.051	.095	.908	.959	.982	.993
.975	.011	.028	.052	.097	.915	.964	.983	.993
.99	.011	.027	.051	.098	.907	.956	.978	.991
.995	.010	.026	.050	.098	.887	.942	.971	.987

(c) n = 100

θ	.010	.025	.050	.100	.900	.950	.975	.990
.6	-2.69	-2.22	-1.83	-1.40	1.13	1.45	1.71	2.01
.8	-2.89	-2.31	-1.85	-1.37	1.12	1.39	1.61	1.86
.9	-3.13	-2.42	-1.88	-1.34	1.07	1.29	1.47	1.66
.95	-3.37	-2.52	-1.90	-1.30	1.00	1.17	1.30	1.45
.975	-3.56	-2.59	-1.90	-1.25	.92	1.04	1.14	1.25
.99	-3.74	-2.68	-1.93	-1.21	.82	.91	.97	1.04
.995	-3.83	-2.74	-1.95	-1.22	.77	.84	.89	.95
	Exact cumulative probabilites at the above percentiles							
.6	.010	.023	.044	.087	.882	.942	.971	.998
.8	.010	.025	.048	.093	.891	.948	.975	.991
.9	.011	.026	.050	.096	.897	.953	.979	.992
.95	.011	.027	.051	.098	.907	.960	.982	.994
.975	.011	.028	.052	.099	.921	.966	.986	.996
.99	.011	.028	.052	.100	.929	.972	.987	.995
.995	.011	.027	.052	.099	.925	.967	.984	.994

influence on the distribution of the score compared to the known σ^2 case.

A major usefulness of the decomposition of the score function into two approximately chi-square distributed, uncorrelated random variables is the simplicity of the ensuing characterization of the finite sample distribution, when the parameter approaches the boundary value. Also exactly on the boundary, characterization can be sharpened, since it can be shown that all the approximations become exact and uncorrelatedness becomes independence. Therefore, the exact distribution (2.7) could be used as such for testing for noninvertibility of the model.

7. REFERENCES.

/1/ AHTOLA, J. and TIAO, G.C.: "Parameter inference for a Nearly Nonstationary First Order Autoregressive Model". Biometrika, 71, 263-272 (1984a).

/2/ AHTOLA, J. and TIAO, G.C.: "Distributions of Autoregressive Parameters with Complex Roots on the Unit Circle", Research Report No.50, Dept. of Statistics, University of Helsinki. (1984b).

/3/ ANDERSON, T.W.: "On Asymptotic Distributions of Estimates of Parameters of Stochastic Difference Equations". Ann. Math. Statist. 30, 676-687. (1959).

/4/ BOX, G.E.P.: "A General Distribution Theory for a Class of Likelihood Criteria". Biometrika, 36, 317-346. (1949).

/5/ BUSE, A.: "The Likelihood Ratio, Wald and Lagrange Multiplier Tests: An Expository Note". The American Statistician 36, 153-157. (1982).

/6/ CRYER, J.D. and LEDOLTER, J.: "Small-Sample Properties of the Maximum Likelihood Estimator in the First Order Moving Average Model". Biometrika, 68, 691-694 (1981).

/7/ DICKEY, D.A. and FULLER, W.A.: "Distribution of the Estimators for Autoregressive Time Series with a Unit Root". J. Amer. Statist. Assoc., 74, 427-431. (1979)

/8/ EVANS, G.B.A. and SAVIN, N.E.: "The Calculation of the Limiting Distribution of the Least Squares Estimator of the Parameter in a Random Walk Model". Ann. Statist. 9, 1114-1118. (1981a)

/9/ EVANS, G.B.A. and SAVIN, N.E.: "Testing for Unit Roots: 1". Econometrica, 49, 753-799 (1981b)

/10/ FULLER, W.A.: "Introduction to Statistical Time Series", New York: Wiley (1976).

/11/ HASZA, D.P. and FULLER, W.A.: "Estimation for Autoregressive Processes with Unit Roots". Ann. Statist. 7, 1106-1120 (1979).

/12/ IMHOF. J.P.: "Computing the Distribution of Quadratic Forms in Normal Variables". Biometrika, 48, 419-426. (1961).

/13/ MACPHERSON, B.D. and FULLER, W.A.:"Consistency of the Least Squares Estimator of the First Order Moving Average Parameter". Ann. Statist., 11, 326-329. (1983).

/14/ MANN, H.B. and WALD, A.: "On the Statistical Treatment of Linear Stochastic Difference Equations". Econometrica, 11, 173-220. (1943).

/15/ PHILLIPS, P.C.B.: "Approximations to Some Finite Sample Distributions Associated with a First-Order Stochastic Difference Equation". Econometrica 45, 463-485. (1977).

/16/ PHILIPS, P.C.B.: "Edgeworth and Saddlepoint Approximations in the First-Order Noncircular Autoregression". Biometrika, 65, 91-98. (1978).

/17/ RAO, C.R.: "Linear Statistical Inference and Its Applications". New York: Wiley. (1973).

/18/ RAO, M.M.: "Asymptotic Distribution of an Estimator of the Boundary Parameter of an Unstable Process". Ann. Statist. 6, 185-190. (1978).

/19/ SATTERTHWAITE, F.E.: "Synthesis of Variance". Psychometrika, 6, 309-316. (1941)