

Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: https://www.tandfonline.com/loi/uasa20

Distribution of the Estimators for Autoregressive Time Series with a Unit Root

David A. Dickey & Wayne A. Fuller

To cite this article: David A. Dickey & Wayne A. Fuller (1979) Distribution of the Estimators for Autoregressive Time Series with a Unit Root, Journal of the American Statistical Association, 74:366a, 427-431, DOI: 10.1080/01621459.1979.10482531

To link to this article: https://doi.org/10.1080/01621459.1979.10482531

Published online: 05 Apr 2012.

Submit your article to this journal

Article views: 950

View related articles

Citing articles: 2959 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=uasa20

Distribution of the Estimators for Autoregressive Time Series With a Unit Root

DAVID A. DICKEY and WAYNE A. FULLER*

Let n observations $Y_1, Y_2, \ldots, Y_n$ be generated by the model $Y_t = \rho Y_{t-1} + e_t$, where Y_0 is a fixed constant and $\{e_t\}_{t=1}^n$ is a sequence of independent normal random variables with mean 0 and variance σ^2. Properties of the regression estimator of ρ are obtained under the assumption that $\rho = \pm 1$. Representations for the limit distributions of the estimator of ρ and of the regression t test are derived. The estimator of ρ and the regression t test furnish methods of testing the hypothesis that $\rho = 1$.

KEY WORDS: Time series; Autoregressive; Nonstationary; Random walk; Differencing.

1. INTRODUCTION

Consider the autoregressive model

$$Y_t = \rho Y_{t-1} + e_t \ , \quad t = 1, 2, \ldots, \tag{1.1}$$

where $Y_0 = 0$, ρ is a real number, and $\{e_t\}$ is a sequence of independent normal random variables with mean zero and variance σ^2 [i.e., e_t NID$(0, \sigma^2)$].

The time series Y_t converges (as $t \to \infty$) to a stationary time series if $|\rho| < 1$. If $|\rho| = 1$, the time series is not stationary and the variance of Y_t is $t\sigma^2$. The time series with $\rho = 1$ is sometimes called a random walk. If $|\rho| > 1$, the time series is not stationary and the variance of the time series grows exponentially as t increases.

Given n observations $Y_1, Y_2, \ldots, Y_n$, the maximum likelihood estimator of ρ is the least squares estimator

$$\hat{\rho} = (\sum_{t=1}^{n} Y_{t-1}^2)^{-1} \sum_{t=1}^{n} Y_t Y_{t-1} \ . \tag{1.2}$$

Rubin (1950) showed that $\hat{\rho}$ is a consistent estimator for all values of ρ. White (1958) obtained the limiting joint-moment generating function for the properly normalized numerator and denominator of $\hat{\rho} - \rho$. For $|\rho| \neq 1$ he was able to invert the joint-moment generating function to obtain the limiting distribution of $\hat{\rho} - \rho$. For $|\rho| < 1$ the limiting distribution of $n^{\frac{1}{2}}(\hat{\rho} - \rho)$ is normal. For $|\rho| > 1$ the limiting distribution of $|\rho|^n(\rho^2 - 1)^{-1}(\hat{\rho} - \rho)$ is Cauchy. For $\rho = 1$, White was able to represent the limiting distribution of $n(\hat{\rho} - 1)$ as that of the ratio of two integrals defined on the Wiener process.

Rao (1961) extended White's results to higher-order autoregressive time series whose characteristic equations have a single root exceeding one and remaining roots less than one in absolute value. Anderson (1959) obtained the limiting distributions of estimators for higher-order processes with more than one root exceeding one in absolute value.

The hypothesis that $\rho = 1$ is of some interest in applications because it corresponds to the hypothesis that it is appropriate to transform the time series by differencing. Currently, practitioners may decide to difference a time series on the basis of visual inspection of the autocorrelation function. For example, see Box and Jenkins (1970, p. 174). The autocorrelation function of the deviations from the fitted model is then investigated as a test of the appropriateness of the model. Box and Jenkins (1970, p. 291) suggested the Box and Pierce (1970) test statistic

$$Q_K = n \sum_{k=1}^{K} r_k^2 \ , \tag{1.3}$$

where

$$r_k = (\sum_{t=1}^{n} \hat{e}_t^2)^{-1} \sum_{t=k+1}^{n} \hat{e}_t \hat{e}_{t-k}$$

and the $\hat{e}_t$'s are the residuals from the fitted model. Under the null hypothesis, the statistic Q_K is approximately distributed as a chi-squared random variable with $K - p$ degrees of freedom, where p is the number of parameters estimated. If $\{Y_t\}$ satisfies (1.1) then $p = 0$ under the null hypothesis and $\hat{e}_t = Y_t - Y_{t-1}$.

The likelihood ratio test of the hypothesis H_0: $\rho = 1$ is a function of

$$\hat{\tau} = (\hat{\rho} - 1) S_e^{-1} (\sum_{t=2}^{n} Y_{t-1}^2)^{\frac{1}{2}} \ ,$$

where

$$S_e^2 = (n - 2)^{-1} \sum_{t=2}^{n} (Y_t - \hat{\rho} Y_{t-1})^2 \ .$$

In this article we derive representations for the limiting distributions of $\hat{\rho}$ and of $\hat{\tau}$, given that $|\rho| = 1$. The representations permit construction of tables of the percentage points for the statistics. The statistics $\hat{\rho}$ and $\hat{\tau}$

* Wayne A. Fuller is Professor of Statistics at Iowa State University, Ames, IA 50011. David A. Dickey is Assistant Professor of Statistics at North Carolina State University, Raleigh, NC 27650. This research was partially supported by Joint Statistical Agreement No. 76-66 with the Bureau of the Census.

© Journal of the American Statistical Association
June 1979, Volume 74, Number 366
Theory and Methods Section

are also generalized to models containing intercept and time terms.

In Section 4 the Monte Carlo method is used to compare the power of the statistics $\hat{\tau}$ and $\hat{\rho}$ with that of Q_K. Examples are given in Section 5.

2. MODELS AND ESTIMATORS

The class of models we investigate consists of (a) the model (1.1), (b) the model

$$Y_t = \mu + \rho Y_{t-1} + e_t \ , \quad t = 1, 2, \ldots \qquad (2.1)$$
$$Y_0 = 0$$

and (c) the model

$$Y_t = \mu + \beta t + \rho Y_{t-1} + e_t \ , \quad t = 1, 2, \ldots \qquad (2.2)$$
$$Y_0 = 0 \ .$$

Assume n observations $Y_1, Y_2, \ldots, Y_n$ are available for analysis and define the $(n-1)$ dimensional vectors,

$$\mathbf{1}' = (1, 1, 1, \ldots, 1) \ ,$$
$$\mathbf{t}' = (1 - (n/2), 2 - (n/2), 3 - (n/2), \ldots, n - 1 - (n/2)) \ ,$$
$$\mathbf{Y}_t' = (Y_2, Y_3, Y_4, \ldots, Y_n) \ ,$$
$$\mathbf{Y}_{t-1}' = (Y_1, Y_2, Y_3, \ldots, Y_{n-1}) \ .$$

Let $\mathbf{U}_1 = \mathbf{Y}_{t-1}$, $\mathbf{U}_2 = (\mathbf{1}, \mathbf{Y}_{t-1})$, and $\mathbf{U}_3 = (\mathbf{1}, \mathbf{t}, \mathbf{Y}_{t-1})$. We define $\hat{\rho}_\mu$ as the last entry in the vector

$$(\mathbf{U}_2'\mathbf{U}_2)^{-1}\mathbf{U}_2'\mathbf{Y}_t \ , \qquad (2.3)$$

and define $\hat{\rho}_\tau$ as the last entry in the vector

$$(\mathbf{U}_3'\mathbf{U}_3)^{-1}\mathbf{U}_3'\mathbf{Y}_t \ . \qquad (2.4)$$

The statistics analogous to the regression t statistics for the test of the hypothesis that $\rho = 1$ are

$$\hat{\tau} = (\hat{\rho} - 1)(S_{e1}^2 c_1)^{-\frac{1}{2}} \ , \qquad (2.5)$$

$$\hat{\tau}_\mu = (\hat{\rho}_\mu - 1)(S_{e2}^2 c_2)^{-\frac{1}{2}} \ , \qquad (2.6)$$

$$\hat{\tau}_\tau = (\hat{\rho}_\tau - 1)(S_{e3}^2 c_3)^{-\frac{1}{2}} \ , \qquad (2.7)$$

where S_{ek}^2 is the appropriate regression residual mean square

$$S_{ek}^2 = (n - k - 1)^{-1}[\mathbf{Y}_t'(\mathbf{I} - \mathbf{U}_k(\mathbf{U}_k'\mathbf{U}_k)^{-1}\mathbf{U}_k')\mathbf{Y}_t] \qquad (2.8)$$

and c_k is the lower-right element of $(\mathbf{U}_k'\mathbf{U}_k)^{-1}$.

3. LIMIT DISTRIBUTIONS

As the first step in obtaining the limit distributions we investigate the quadratic forms appearing in the statistics. Because the estimators are ratios of quadratic forms we lose no generality by assuming $\sigma^2 = 1$ in the sequel.

3.1 Canonical Representation of the Statistics

Given that $\rho = 1$, the quadratic form $\sum_{t=2}^{n} Y_{t-1}^2$ can be expressed as $\mathbf{e}_n'\mathbf{A}_n\mathbf{e}_n$, where $\mathbf{e}_n' = (e_1, e_2, \ldots, e_{n-1})$, the elements α_{ij} of $\mathbf{A}_n^{-1}$ satisfy $\alpha_{11} = 1$, $\alpha_{jj} = 2$ for $j > 1$, $\alpha_{j,j-1} = \alpha_{j,j+1} = -1$ for all j, and $\alpha_{ij} = 0$ otherwise. By a result of Rutherford (1946), the roots of $\mathbf{A}_n$ are

$$\lambda_{in} = (\tfrac{1}{4}) \sec^2((n - i)\pi/(2n - 1)) \ , \quad i = 1, 2, \ldots, n - 1 \ .$$

Let $\mathbf{M}$ be the $n - 1$ by $n - 1$ orthonormal matrix whose ith row is the eigenvector of $\mathbf{A}_n$ corresponding to λ_{in}. The itth element of $\mathbf{M}$ is

$$m_{it} = 2(2n - 1)^{-\frac{1}{2}} \cdot \cos[(4n - 2)^{-1}(2t - 1)(2i - 1)\pi] \ , \qquad (3.1)$$

and we can express the normalized denominator sum of squares appearing in $\hat{\rho}$ as

$$\Gamma_n = n^{-2} \sum_{t=2}^{n} Y_{t-1}^2 = n^{-2} \sum_{i=1}^{n-1} \lambda_{in} Z_{in}^2 \ , \qquad (3.2)$$

where $\mathbf{Z} = (Z_{1n}, Z_{2n}, \ldots, Z_{n-1,n})' = \mathbf{M}\mathbf{e}_n$.

Let

$$\mathbf{H}_n = n^{-\frac{1}{2}} \cdot \begin{bmatrix} n^2 & n^2 & n^2 & & n^2 \\ n(n-1) & n(n-2) & n(n-3) & \ldots & n \\ 0 & n-2 & 2(n-3) & \ldots & n-2 \end{bmatrix}$$

and

$$(T_n, W_n, V_n)' = n^{-\frac{1}{2}}(Y_{n-1}, n^{-1} \sum_{t=2}^{n} Y_{t-1}, n^{-2} \sum_{j=1}^{n-1} (n - j)(j - 1)e_j)'$$
$$= \mathbf{H}_n\mathbf{e}_n = \mathbf{H}_n\mathbf{M}^{-1}\mathbf{Z} \ . \qquad (3.3)$$

Then

$$n(\hat{\rho} - 1) = (2\Gamma_n)^{-1}(T_n^2 - 1) + O_p(n^{-\frac{1}{2}}) \ , \qquad (3.4)$$

$$n(\hat{\rho}_\mu - 1) = (2\Gamma_n - 2W_n^2)^{-1}(T_n^2 - 1 - 2T_nW_n) + O_p(n^{-\frac{1}{2}}) \ , \qquad (3.5)$$

$$n(\hat{\rho}_\tau - 1) = [2(\Gamma_n - W_n^2 - 3V_n^2)]^{-1} \cdot [(T_n - 2W_n)(T_n - 6V_n) - 1] + O_p(n^{-\frac{1}{2}}) \ . \qquad (3.6)$$

3.2 Representations for the Limit Distributions

Having expressed $n(\hat{\rho} - 1)$, $n(\hat{\rho}_\mu - 1)$, and $n(\hat{\rho}_\tau - 1)$ in terms of $(\Gamma_n, T_n, W_n, V_n)$ we obtain the limiting distribution of the vector random variable. The following lemma will be used in our derivation of the limit distribution.

Lemma 1: Let $\{Z_i\}_{i=1}^{\infty}$ be a sequence of independent random variables with zero means and common variance σ^2. Let $\{w_i; i = 1, 2, \ldots\}$ be a sequence of real numbers and let $\{w_{in}; i = 1, 2, \ldots, n - 1; n = 1, 2, \ldots\}$ be a triangular array of real numbers. If

$$\sum_{i=1}^{\infty} w_i^2 < \infty \ ,$$
$$\lim_{n\to\infty} \sum_{i=1}^{n-1} w_{in}^2 = \sum_{i=1}^{\infty} w_i^2 \ ,$$

and

$$\lim_{n\to\infty} w_{in} = w_i ,$$

then $\sum_{i=1}^{\infty} w_i Z_i$ is well defined as a limit in mean square and

$$p\lim \{\sum_{i=1}^{n} w_{in} Z_i\} = \sum_{i=1}^{\infty} w_i Z_i .$$

Proof: Let $\epsilon > 0$ be given. Then we can choose an M such that

$$\sigma^2 \sum_{i=M+1}^{\infty} w_i^2 < \epsilon/9$$

and

$$\sigma^2 |\sum_{i=1}^{n} w_{in}^2 - \sum_{i=1}^{\infty} w_i^2| < \epsilon/9 ,$$

for all $n > M$. Furthermore, given M, we can choose $N_0 > M$ such that $n > N_0$ implies

$$\sigma^2 \sum_{i=1}^{M} (w_{in} - w_i)^2 < \epsilon/9$$

and

$$\sigma^2 \sum_{i=M+1}^{n} w_{in}^2 < 3\epsilon/9 .$$

Hence, for all $n > N_0$,

$$\operatorname{var}\{\sum_{i=1}^{n} w_{in} Z_i - \sum_{i=1}^{\infty} w_i Z_i\} < \epsilon ,$$

and the result follows by Chebyshev's inequality.

Theorem 1: Let $\{Z_i\}_{i=1}^{\infty}$ be a sequence of NID(0, 1) random variables. Let $\boldsymbol{\eta}_n' = (\Gamma_n, T_n, W_n, V_n)$, where the elements of the vector are defined in (3.2) and (3.3). Let

$$\boldsymbol{\eta}' = (\Gamma, T, W, V) ,$$

where

$$(\Gamma, T) = (\sum_{i=1}^{\infty} \gamma_i^2 Z_i^2, \sum_{i=1}^{\infty} 2^{\frac{1}{2}} \gamma_i Z_i) ,$$

$$(W, V) = (\sum_{i=1}^{\infty} 2^{\frac{1}{2}} \gamma_i^2 Z_i, \sum_{i=1}^{\infty} 2^{\frac{1}{2}} [2\gamma_i^3 - \gamma_i^2] Z_i) ,$$

$$\gamma_i = (-1)^{i+1} \sqrt{\gamma_i^2} ,$$

and

$$\gamma_i^2 = \lim_{n\to\infty} n^{-2}\lambda_{in} = 4[(2i-1)\pi]^{-2} .$$

Then $\boldsymbol{\eta}_n$ converges in distribution to $\boldsymbol{\eta}$, that is,

$$\boldsymbol{\eta}_n \xrightarrow{\mathcal{L}} \boldsymbol{\eta} .$$

Proof: Note that $\boldsymbol{\eta}$ is a well-defined random variable because $\sum_{i=1}^{\infty} \gamma_i^k < \infty$ for $k = 2, 3, \ldots, 6$. Let $\boldsymbol{\xi}_{in}$ be the ith column of $\mathbf{H}_n\mathbf{M}^{-1}$, where

$$\boldsymbol{\xi}_{in} = (a_{in}, b_{in}, g_{in})' = [\operatorname{cov}(T_n, Z_{in}), \operatorname{cov}(W_n, Z_{in}), \operatorname{cov}(V_n, Z_{in})]' .$$

For fixed i,

$$\lim_{n\to\infty} \boldsymbol{\xi}_{in} = \boldsymbol{\xi}_i = (a_i, b_i, g_i)' = 2^{\frac{1}{2}}(\gamma_i, \gamma_i^2, 2\gamma_i^3 - \gamma_i^2)' . \tag{3.7}$$

By Jolley (1961, p. 56, #307,308) we have

$$\sum_{i=1}^{\infty} (a_i^2, b_i^2, g_i^2) = (1, 1/3, 1/30) .$$

Let

$$(\Gamma_n^*, T_n^*) = (\sum_{i=1}^{n-1} n^{-2}\lambda_{in} Z_i^2, \sum_{i=1}^{n-1} a_{in} Z_i) ,$$

$$(W_n^*, V_n^*) = (\sum_{i=1}^{n-1} b_{in} Z_i, \sum_{i=1}^{n-1} g_{in} Z_i) .$$

Now, for example, by (3.3)

$$\lim_{n\to\infty} \sum_{i=1}^{n-1} a_{in}^2 = \lim_{n\to\infty} \operatorname{var}\{T_n^*\} = \lim_{n\to\infty} \operatorname{var}\{T_n\} = 1 .$$

Therefore, by (3.7) and Lemma 1, T_n^* converges in probability to T. It follows by analogous arguments that $(\Gamma_n^*, T_n^*, W_n^*, V_n^*)$ converges in probability to (Γ, T, W, V). Because the distribution of $(\Gamma_n^*, T_n^*, W_n^*, V_n^*)$ is the same as that of $\boldsymbol{\eta}_n$ we obtain the conclusion.

Corollary 1: Let Y_t satisfy (1.1) with $\rho = 1$. Then

$$n(\hat{\rho} - 1) \xrightarrow{\mathcal{L}} \tfrac{1}{2}(\Gamma^{-1}(T^2 - 1) ,$$

$$n(\hat{\rho}_\mu - 1) \xrightarrow{\mathcal{L}} \tfrac{1}{2}(\Gamma - {}^2W)^{-1}[(T^2 - 1) - 2TW] ,$$

$$\hat{\tau} \xrightarrow{\mathcal{L}} \tfrac{1}{2}\Gamma^{-\frac{1}{2}}(T^2 - 1) ,$$

and

$$\hat{\tau}_\mu \xrightarrow{\mathcal{L}} \tfrac{1}{2}(\Gamma - W^2)^{-\frac{1}{2}}[(T^2 - 1) - 2TW] .$$

Let Y_t satisfy (2.1) with $\rho = 1$. Then

$$n(\hat{\rho}_\tau - 1) \xrightarrow{\mathcal{L}} \tfrac{1}{2}(\Gamma - W^2 - 3V^2)^{-1} \cdot [(T - 2W)(T - 6V) - 1]$$

and

$$\hat{\tau}_\tau \xrightarrow{\mathcal{L}} \tfrac{1}{2}(\Gamma - W^2 - 3V^2)^{-\frac{1}{2}}[(T - 2W)(T - 6V) - 1] .$$

Proof: The proof is an immediate consequence of Theorem 1 because the denominator quadratic forms in $\hat{\rho}$, $\hat{\rho}_\mu$, $\hat{\rho}_\tau$ are continuous functions of $\boldsymbol{\eta}$ that have probability 1 of being positive and the S_{ek}^2 of (2.8) converge in probability to σ^2.

The numerator and denominator of the limit representation of $n(\hat{\rho} - 1)$ are consistent with White's (1958) limit joint-moment generating function.

Note that the limiting distributions of $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ are obtained under the assumption that the constant term μ is zero. Likewise, the limiting distributions of $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$ are derived under the assumption that the coefficient for time, β, is zero. The distributions of $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$ are unaffected by the value of μ in (2.2). If $\mu \neq 0$ for (2.1) or $\beta \neq 0$ for (2.2), the limiting distributions of $\hat{\tau}_\mu$ and $\hat{\tau}_\tau$ are normal. Thus if (2.1) is the maintained model and

the statistic $\hat{\tau}_\mu$ is used to test the hypothesis $\rho = 1$, the hypothesis will be accepted with probability greater than the nominal level where $\mu \neq 0$.

By the results of Fuller (1976, p. 370), the limiting distributions of $\hat{\rho}$, $\hat{\rho}_\mu$, and $\hat{\rho}_\tau$, given that $\rho = -1$, are identical and equal to the mirror image of the limiting distribution of $\hat{\rho}$ given that $\rho = 1$. Likewise, the limiting distributions of $\hat{\tau}$, $\hat{\tau}_\mu$, and $\hat{\tau}_\tau$ for $\rho = -1$ are identical and equal to the mirror image of the limiting distribution of $\hat{\tau}$ for $\rho = 1$.

In our derivations Y_0 is fixed. The distributions of $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ do not depend on the value of Y_0. The limiting distribution of $\hat{\rho}$ does not depend on Y_0, but the small-sample distribution of $\hat{\rho}$ will be influenced by Y_0.

In the derivations we assumed the e_t to be NID$(0, \sigma^2)$. The limiting distributions also hold for e_t that are independent and identically distributed nonnormal random variables with mean zero and variance σ^2. White (1958) and Hasza (1977) have discussed this generalization.

The statistic $\hat{\tau}$ is a monotone function of the likelihood ratio when Y_0 is fixed under the null model of $\rho = 1$ and under the alternative model of $\rho \neq 1$. Tests based on the τ statistics are not likelihood ratios and not necessarily the most powerful that can be constructed if, for example, the alternative model is that $(Y_0, Y_1, \ldots, Y_n)$ is a portion of a realization from a stationary autoregressive process.

A set of tables of the percentiles of the distributions is given in Fuller (1976, pp. 371,373) and a slightly more accurate set in Dickey (1976). Dickey also presents details of the table construction and gives estimates of the sampling error of the estimated percentiles.

4. POWER COMPARISONS

The powers of the statistics studied in this article were compared with that of the Box–Pierce Q statistic in a Monte Carlo study using the model

$$Y_t = \rho Y_{t-1} + e_t \ , \quad t = 1, 2, \ldots, n \ ,$$

where the $e_t \sim$ NID$(0, \sigma^2)$ and $Y_0 = 0$. Four thousand samples of size $n = 50$, 100, 250 were generated for $\rho = .80, .90, .95, .99, 1.00, 1.02, 1.05$. The random-number generator SUPER DUPER from McGill University was used to create the pseudonormal variables.

Eight two-sided size .05 tests of the hypothesis $\rho = 1$ were applied to each sample. The tests were $\hat{\rho}$, $\hat{\tau}$, $\hat{\rho}_\mu$, $\hat{\tau}_\mu$, Q_1, Q_5, Q_{10}, Q_{20}, where Q_K is the Box–Pierce Q statistic defined in (1.3) with $\hat{e}_t = Y_t - Y_{t-1}$.

There are several conclusions to be drawn from the results presented in the table. First, the Q statistics are less powerful than the statistics introduced in this article. For example, when $n = 250$ and $\rho = .8$ the worst of the statistics introduced in this article rejected the null hypothesis 100 percent of the time, while the best of the Q statistics rejected the null hypothesis in only 45 percent of the samples.

Second, the performances of $\hat{\rho}$ and $\hat{\tau}$ were similar, and they were uniformly more powerful than the other test statistics. It is not surprising that $\hat{\rho}$ and $\hat{\tau}$ are superior to $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ because $\hat{\rho}$ and $\hat{\tau}$ use the knowledge that the true value of the intercept in the regression is zero.

Monte Carlo Power of Two-Sided Size .05 Tests of $\rho = 1$

		ρ						
n	Test	.80	.90	.95	.99	1.00	1.02	1.05
50	Q_1	.09	.05	.05	.04	.04	.07	.47
	Q_5	.07	.04	.03	.03	.04	.08	.53
	Q_{10}	.05	.04	.03	.03	.03	.09	.54
	Q_{20}	.03	.02	.02	.02	.02	.08	.52
	$\hat{\rho}$	.57	.18	.08	.05	.05	.14	.71
	$\hat{\tau}$	.57	.18	.08	.04	.05	.23	.70
	$\hat{\rho}_\mu$	.28	.10	.06	.05	.06	.11	.67
	$\hat{\tau}_\mu$	.18	.06	.04	.04	.05	.13	.68
100	Q_1	.15	.07	.05	.04	.05	.26	.94
	Q_5	.13	.08	.05	.04	.04	.34	.95
	Q_{10}	.11	.06	.05	.03	.04	.37	.95
	Q_{20}	.08	.05	.04	.03	.03	.38	.95
	$\hat{\rho}$	.99	.55	.17	.05	.05	.54	.98
	$\hat{\tau}$	.99	.55	.17	.04	.05	.59	.97
	$\hat{\rho}_\mu$	.86	.30	.10	.05	.05	.49	.98
	$\hat{\tau}_\mu$	.73	.18	.06	.04	.05	.51	.98
250	Q_1	.34	.12	.06	.05	.06	.94	1.00
	Q_5	.45	.13	.07	.04	.05	.95	1.00
	Q_{10}	.34	.12	.06	.04	.05	.95	1.00
	Q_{20}	.24	.10	.05	.04	.04	.95	1.00
	$\hat{\rho}$	1.00	1.00	.74	.08	.05	.98	1.00
	$\hat{\tau}$	1.00	1.00	.74	.08	.05	.97	1.00
	$\hat{\rho}_\mu$	1.00	.96	.43	.06	.05	.98	1.00
	$\hat{\tau}_\mu$	1.00	.89	.28	.04	.05	.98	1.00

Third, for $\rho < 1$ the statistic $\hat{\rho}_\mu$ yielded a more powerful test than the statistic $\hat{\tau}_\mu$. For $\rho > 1$ the ranking was reversed and the $\hat{\tau}_\mu$ statistic was more powerful.

For sample sizes of 50 and 100, and $\rho < 1$, Q_1 was the most powerful of the Q statistics studied. For sample size 250, Q_5 was the most powerful Q statistic. The size of the Q tests for $K \geq 5$ was considerably less than .05 for $n = 50$.

There is evidence that $\hat{\tau}$ and $\hat{\tau}_\mu$ are biased tests, accepting the null hypothesis more than 95 percent of the time for ρ close to, but less than, one. Because the tests are consistent, the minimum point of the power function is moving toward one as the sample size increases.

5. EXAMPLES

Gould and Nelson (1974) investigated the stochastic structure of the velocity of money using the yearly observations from 1869 through 1960 given in Friedman and Schwartz (1963). Gould and Nelson concluded that the logarithm of velocity is consistent with the model $X_t = X_{t-1} + e_t$, where $e_t \sim N(0, \sigma^2)$ and X_t is the velocity of money.

Two models,

$$X_t - X_1 = \rho(X_{t-1} - X_1) + e_t \qquad (5.1)$$

and

$$X_t = \mu + \rho X_{t-1} + e_t \ , \qquad (5.2)$$

were fit to the data. For (5.1) the estimates were

$$\hat{X}_t - X_1 = \underset{(.0094)}{1.0044}(X_{t-1} - X_1) \ , \quad \hat{\sigma}^2 = .0052$$

and for (5.2),

$$\hat{X}_t = \underset{(.0176)}{.0141} + \underset{(.0199)}{.9702}X_{t-1} \ , \quad \hat{\sigma}^2 = .0050 \ .$$

Model (5.1) assumes that it is known that no intercept enters the model if X_1 is subtracted from all observations. Model (5.2) permits an intercept in the model. The numbers in parentheses are the "standard errors" output by the regression program. For (5.1) we compute

$$n(\hat{\rho} - 1) = 91(.0044) = .4004$$

and

$$\hat{\tau} = (.0094)^{-1}(.0044) = .4681 \ .$$

Using either Table 8.5.1 or 8.5.2 of Fuller (1976), the hypothesis that $\rho = 1$ is accepted at the .10 level.

For (5.2) we obtain the statistics

$$n(\hat{\rho}_\mu - 1) = 92(.9702 - 1) = -2.742$$

and

$$\hat{\tau}_\mu = (.0199)^{-1}(.9702 - 1.0) = -1.50 \ .$$

Again the hypothesis is accepted at the .10 level.

As a second example we study the logarithm of the quarterly Federal Reserve Board Production Index for the period 1950–1 through 1977–4. We assume that the time series is adequately represented by the model

$$Y_t = \beta_0 + \beta_1 t + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + e_t \ ,$$

where e_t are NID$(0, \sigma^2)$ random variables.

On the basis of the results of Fuller (1976, p. 379) the coefficient of Y_{t-1} in the regression equation

$$Y_t - Y_{t-1} = \beta_0 + \beta_1 t + (\alpha_1 + \alpha_2 - 1)Y_{t-1} - \alpha_2(Y_{t-1} - Y_{t-2}) + e_t$$

can be used to test the hypothesis that $\rho = \alpha_1 + \alpha_2 = 1$. This hypothesis is equivalent to the hypothesis that one of the roots of the characteristic equation of the process is one. The least squares estimate of the equation is

$$\hat{Y}_t - Y_{t-1} = \underset{(.15)}{.52} + \underset{(.00034)}{.00120}t - \underset{(.033)}{.119}Y_{t-1} + \underset{(.081)}{.498}(Y_{t-1} - Y_{t-2}), \quad \hat{\sigma}^2 = .033 \ .$$

There are 110 observations in the regression. The numbers in parentheses are the quantities output as "standard errors" by the regression program. On the basis of the results of Fuller, the statistic $(n - p)(\hat{\rho} - 1)(1 + \hat{\alpha}_2)^{-1}$, where $\hat{\rho}$ is the coefficient of Y_{t-1} and p is the number of parameters estimated, is approximately distributed as $n(\hat{\rho}_\tau - 1)$. Also the "t statistic" constructed by dividing the coefficient of Y_{t-1} by the regression standard error is approximately distributed as $\hat{\tau}_\tau$. For this example we have

$$(n - p)(\hat{\rho} - 1)(1 + \hat{\alpha}_2)^{-1} = 106(-.119)(.502)^{-1} = -25.1$$

and

$$\hat{\tau}_\tau = (.033)^{-1}(-.119) = -3.61 \ .$$

Both statistics lead to rejection of the null hypothesis of a unit root at the 5 percent level if the alternative hypothesis is that both roots are less than one in absolute value. The Monte Carlo study of Section 4 indicated that tests based on the estimated ρ were more powerful for tests against stationarity than the $\hat{\tau}$ statistics. In this example the test based on $\hat{\rho}$ rejects the hypothesis at a smaller size (.025) than that of the $\hat{\tau}$ statistic (.05).

[*Received November 1976. Revised November 1978.*]

REFERENCES

Anderson, Theodore W. (1959), "On Asymptotic Distributions of Estimates of Parameters of Stochastic Difference Equations," *Annals of Mathematical Statistics*, 30, 676–687.

Box, George E.P., and Jenkins, Gwilym M. (1970), *Time Series Analysis Forecasting and Control*, San Francisco: Holden-Day.

Box, George E.P., and Pierce, David A. (1970), "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models," *Journal of the American Statistical Association*, 65, 1509–1526.

David, Herbert A. (1970), *Order Statistics*, New York: John Wiley & Sons.

Dickey, David A. (1976), "Estimation and Hypothesis Testing in Nonstationary Time Series," Ph.D. dissertation, Iowa State University.

Friedman, Milton, and Schwartz, A.J. (1963), *A Monetary History of the United States 1867–1960*, Princeton, N.J.: Princeton University Press.

Fuller, Wayne A. (1976), *Introduction to Statistical Time Series*, New York: John Wiley & Sons.

Gould, John P., and Nelson, Charles R. (1974), "The Stochastic Structure of the Velocity of Money," *American Economic Review*, 64, 405–417.

Hasza, David P. (1977), "Estimation in Nonstationary Time Series," Ph.D. dissertation, Iowa State University.

Jolley, L.B.W. (1961), *Summation of Series* (2nd ed.), New York: Dover Press.

Rao, M.M. (1961), "Consistency and Limit Distributions of Estimators of Parameters in Explosive Stochastic Difference Equations," *Annals of Mathematical Statistics*, 32, 195–218.

Rao, M.M. (1978), "Asymptotic Distribution of an Estimator of the Boundary Parameter of an Unstable Process," *Annals of Statistics*, 6, 185–190.

Rubin, Herman (1950), "Consistency of Maximum-Likelihood Estimates in the Explosive Case," in *Statistical Inference in Dynamic Economic Models*, ed. T.C. Koopmans, New York: John Wiley & Sons.

Rutherford, D.E. (1946), "Some Continuant Determinants Arising in Physics and Chemistry," *Proceedings of the Royal Society of Edinburgh*, Sect. A, 62, 229–236.

White, John S. (1958), "The Limiting Distribution of the Serial Correlation Coefficient in the Explosive Case," *Annals of Mathematical Statistics*, 29, 1188–1197.