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

David A. Dickey, Wayne A. Fuller


Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28197906%2974%3A366%3C427%3ADOTEDA%3E2.0.CO%3B2-3

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

Journal of the American Statistical Association is published by American Statistical Association. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/aastata.html.

Journal of the American Statistical Association
©1979 American Statistical Association

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2002 JSTOR
Let \( n \) observations \( Y_1, Y_2, \ldots, Y_n \) be generated by the model
\[
Y_t = \rho Y_{t-1} + \varepsilon_t, \quad t = 1, 2, \ldots
\]
where \( Y_0 = 0 \), \( \rho \) is a real number, and \( \{\varepsilon_t\} \) is a sequence of independent normal random variables with mean 0 and variance \( \sigma^2 \). Properties of the regression estimator of \( \rho \) are obtained under the assumption that \( \rho = \pm 1 \). Representations for the limit distributions of the estimator of \( \rho \) and of the regression \( t \) test are derived. The estimator of \( \rho \) 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
\[
Y_t = \rho Y_{t-1} + \varepsilon_t, \quad t = 1, 2, \ldots
\]  
where \( Y_0 = 0 \), \( \rho \) is a real number, and \( \{\varepsilon_t\} \) is a sequence of independent normal random variables with mean zero and variance \( \sigma^2 \) [i.e., \( \varepsilon_t \) 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 \( \sigma^2 \). 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 \( \rho \) is the least squares estimator
\[
\hat{\rho} = \left( \sum_{t=1}^{n} (1 - \rho^2)^{-1} \right) \sum_{t=1}^{n} Y_t Y_{t-1}.
\]

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
\[
Q_K = n \sum_{k=1}^{K} r_k^2,
\]
where
\[
r_k = \left( \sum_{l=1}^{n} \hat{\varepsilon}_l \right)^{-1} \sum_{l=1}^{n} \hat{\varepsilon}_l \hat{\varepsilon}_{l-k},
\]
and the \( \hat{\varepsilon}_l \)'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 \( \hat{\varepsilon}_t = Y_t - Y_{t-1} \).

The likelihood ratio test of the hypothesis \( H_0: \rho = 1 \) is a function of
\[
\hat{\tau} = (\hat{\rho} - 1) S^{-1} (\sum_{l=1}^{n} Y_{l-1}) \hat{\rho},
\]
where
\[
S^2 = (n-2)^{-1} \sum_{l=1}^{n} (Y_l - \hat{\rho} Y_{l-1})^2.
\]

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 \( \hat{\tau} \)
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 \( \hat{\beta} \) 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

\[
Y_t = \mu + \rho Y_{t-1} + e_t, \quad t = 1, 2, \ldots
\]

\[
Y_0 = 0
\]

and (c) the model

\[
Y_t = \mu + \beta t + \rho Y_{t-1} + e_t, \quad t = 1, 2, \ldots
\]

\[
Y_0 = 0.
\]

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

\[
Y' = (1, 1, 1, \ldots, 1),
\]

\[
t' = (1 - (n/2), 2 - (n/2), \ldots, (n/2), \ldots, n - 1 - (n/2)),
\]

\[
Y_{t-1} = (Y_1, Y_2, \ldots, Y_{n-1}).
\]

Let \( U_1 = Y_{t-1}, U_2 = (1, Y_{t-1}), \) and \( U_3 = (1, t, Y_{t-1}) \). We define \( \hat{\beta}_p \) as the last entry in the vector

\[
(U_2'U_2)^{-1}U_2'Y_t,
\]

(2.3) and define \( \hat{\beta}_r \), as the last entry in the vector

\[
(U_3'U_3)^{-1}U_3'Y_t.
\]

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

\[
\hat{\tau} = (\hat{\beta} - 1)(S_{\hat{\beta}}c_\rho)^{-1},
\]

(2.5)

\[
\hat{\tau}_p = (\hat{\beta}_p - 1)(S_{\hat{\beta}}c_\rho)^{-1},
\]

(2.6)

\[
\hat{\tau}_r = (\hat{\beta}_r - 1)(S_{\hat{\beta}}c_\rho)^{-1},
\]

(2.7) where \( S_{\hat{\beta}} \) is the approximation residual mean square

\[
S_{\hat{\beta}} = (n - k - 1)^{-1}[(Y_t' - U_3(U_3'U_3)^{-1}U_3')Y_t]
\]

(2.8) and \( c_\rho \) is the lower-right element of \( (U_3'U_3)^{-1} \).

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 \( e_u'A_u e_u \), where \( e_u' = (e_1, e_2, \ldots, e_{n-1}) \), the elements \( \alpha_{ij} \) of \( A_u^{-1} \) satisfy \( \alpha_{11} = 1, \alpha_{ij} = 2 \) 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 \( A_u \) are

\[
\lambda_{ir} = \left( \frac{1}{2} \right) \text{sin}^2 \left( \frac{(n - i)\pi}{2(n - 1)} \right),
\]

(3.1)

\( i = 1, 2, \ldots, n - 1 \).

Let \( M \) be the \( n - 1 \) by \( n - 1 \) orthonormal matrix whose \( i \)th row is the eigenvector of \( A_u \) corresponding to \( \lambda_{ir} \). The \( i \)th element of \( M \) is

\[
m_{ir} = 2(2n - 1)^{-1} \cdot \cos \left[ \left( \frac{4n - 2}{2(2t - 1)} \right)(2i - 1)\pi \right],
\]

(3.1) and we can express the normalized denominator sum of squares appearing in \( \hat{\beta} \) as

\[
\Gamma_n = n^{-2} \sum_{t=2}^n Y_{t-1}^2 = n^{-2} \sum_{i=1}^{n-1} \lambda_{ir}Z_{ir}^2,
\]

(3.2) where \( Z = (Z_{1r}, Z_{2r}, \ldots, Z_{n-1,r})' = Me_r \).

Let

\[
H_n = n^{-1}
\]

\[
\begin{pmatrix}
\frac{n}{2} & \frac{n}{2} & \frac{n}{2} & \frac{n}{2} \\
0 & 0 & 0 & 0 \\
\end{pmatrix}
\]

\[
\begin{pmatrix}
n(n - 1) & n(n - 2) & n(n - 3) & \ldots & n \\
0 & 2(n - 3) & \ldots & n - 2 \\
\end{pmatrix}
\]

\[
\begin{pmatrix}
Y_{t-1} \sum_{t=2}^n Y_{t-1} = n^{-2} \sum_{i=1}^{n-1} \lambda_{ir}Z_{ir}^2
\end{pmatrix}
\]

(3.3)

Then

\[
n(\beta - 1) = (2\Gamma_n)^{-1}(T_s^2 - 1) + O_p(n^{-1})
\]

(3.4)

\[
n(\beta_p - 1) = (2\Gamma_n - 2W_r^2)^{-1}(T_s^2 - 1 - 2T_sW_r)
\]

\[
+ O_p(n^{-1})
\]

(3.5)

\[
n(\beta - 1) = 2[2\Gamma_n - W_s^2 - 3V_s^2]^{-1}
\]

\[
\cdot [(T_s - 2W_r)(T_s - 6V_s - 1)] + O_p(n^{-1})
\]

(3.6)

3.2 Representations for the Limit Distributions

Having expressed \( n(\beta - 1), n(\beta_p - 1), \) and \( n(\beta_r - 1) \) in terms of \( \Gamma_n, T_s, W_s, V_s \) 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 \( \sigma^2 \). Let \( \{w_i; i = 1, 2, \ldots\} \) be a sequence of real numbers and let \( \{w_i; i = 1, 2, \ldots, n - 1; n = 1, 2, \ldots\} \) be a triangular array of real numbers. If

\[
\sum_{i=1}^\infty w_i^2 < \infty,
\]

\[
\lim_{n \to \infty} \sum_{i=1}^{n-1} w_i^2 = \sum_{i=1}^\infty w_i^2,
\]

The following lemma will be used in our derivation of the limit distribution.
and
\[ \lim_{n \to \infty} w_i = w_i, \]
then \( \sum_{i=1}^{n} w_i Z_i \) is well defined as a limit in mean square and
\[ p \lim_{n \to \infty} \left( \sum_{i=1}^{n} w_i Z_i \right) = \sum_{i=1}^{\infty} w_i Z_i. \]

**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
\[ \sigma^2 \left| \sum_{i=1}^{n} w_i^2 \right| < \epsilon/9, \]
for all \( n > M \). Furthermore, given \( M \), we can choose \( N \) such that \( n > N \) implies
\[ \sigma^2 \sum_{i=1}^{M} (w_i - w_i)^2 < \epsilon/9 \]
and
\[ \sigma^2 \sum_{i=M+1}^{\infty} w_i^2 < 3\epsilon/9. \]
Hence, for all \( n > N \),
\[ \text{var} \left( \sum_{i=1}^{n} w_i Z_i - \sum_{i=1}^{\infty} w_i Z_i \right) < \epsilon, \]
and the result follows by Chebyshev’s inequality.

**Theorem 1:** Let \( (Z_j)_{j=1}^{\infty} \) be a sequence of \( \text{NID}(0, 1) \) random variables. Let \( \eta' = (T, W, V) \), where the elements of the vector are defined in (3.2) and (3.3). Let
\[ \eta' = (\Gamma, T, W, V), \]

Then \( \eta' \) converges in distribution to \( \eta \), that is,
\[ \eta' \xrightarrow{d} \eta. \]

**Proof:** Note that \( \eta' \) is a well-defined random variable because \( \sum_{i=1}^{\infty} \gamma_i^k < \infty \) for \( k = 2, 3, \ldots, 6 \). Let \( \xi_i \) be the \( i \)-th column of \( \mathbf{H}, \mathbf{M}^{-1} \), where
\[ \xi_i = (a_i, b_i, g_i), \]
\[ = \left[ \text{cov}(T_i, Z_i), \text{cov}(W_i, Z_i), \text{cov}(V_i, Z_i) \right]'. \]

For fixed \( i \),
\[ \lim_{n \to \infty} \xi_i = (a_i, b_i, g_i)', \]
\[ = 2(\gamma_i \gamma_i', 2\gamma_i - \gamma_i'), \]
\[ \text{and} \]
\[ \lim_{n \to \infty} \left( \sum_{i=1}^{n} a_i^2, b_i^2, g_i^2 \right) = (1, 1, 1/30). \]
Let
\[ (T_{n*}, W_{n*}) = \left( \sum_{i=1}^{n-1} \gamma_i^2 Z_i, \sum_{i=1}^{n} a_i Z_i \right), \]
\[ (W_{n*}, V_{n*}) = \left( \sum_{i=1}^{n-1} b_i Z_i, \sum_{i=1}^{n} g_i Z_i \right). \]
Now, for example, by (3.3)
\[ \lim_{n \to \infty} \sum_{i=1}^{n} a_i^2 = \text{var} \left( T_{n*} \right) = \lim \text{var} \left( T_{n*} \right) = 1. \]
Therefore, by (3.7) and Lemma 1, \( T_{n*} \) converges in probability to \( T \). It follows by analogous arguments that \( (T_{n*}, W_{n*}, V_{n*}) \) converges in probability to \( (T, W, V) \). Because the distribution of \( (T_{n*}, T_{n*}, W_{n*}, V_{n*}) \) is the same as that of \( \eta_i \), we obtain the conclusion.

**Corollary 1:** Let \( Y_i \) satisfy (1.1) with \( \rho = 1 \). Then
\[ n(\beta - 1) \xrightarrow{d} \frac{1}{2}(T^3 - 1), \]
\[ n(\rho, 1) \xrightarrow{d} \frac{1}{2}(T - 3W)^3 - [(T^3 - 1) - 2TW], \]
and
\[ \tau_i \xrightarrow{d} \frac{1}{2}(T^3 - 1) \]
\[ \tau_r \xrightarrow{d} \frac{1}{2}(T^3 - 1) - 2TW. \]
Let \( Y_i \) satisfy (2.1) with \( \rho = 1 \). Then
\[ n(\beta, 1) \xrightarrow{d} \frac{1}{2}(T - 3W^3)^{-1} \]
\[ \cdot [(T - 2W)(T - 6W) - 1] \]
and
\[ \tau_i \xrightarrow{d} \frac{1}{2}(T - 3W)^3 - [(T - 2W)(T - 6W) - 1]. \]

**Proof:** The proof is an immediate consequence of Theorem 1 because the denominator quadratic forms in \( \beta, \rho, \beta, \tau \) are continuous functions of \( \eta \) that have probability 1 of being positive and the \( S_{10} \) of (2.8) converge in probability to \( \sigma^2 \).

The numerator and denominator of the limit representation of \( n(\beta - 1) \) are consistent with White’s (1958) limit joint-moment generating function.

Note that the limiting distributions of \( \beta, \rho, \beta, \tau \) are obtained under the assumption that the constant term \( \mu \) is zero. Likewise, the limiting distributions of \( \beta, \beta, \tau \) are derived under the assumption that the coefficient for time, \( \beta, \) is zero. The distributions of \( \beta, \beta, \tau \) are unaffected by the value of \( \mu \) in (2.2). If \( \mu \neq 0 \) for (2.1) or \( \beta \neq 0 \) for (2.2), the limiting distributions of \( \beta, \beta, \tau \) are normal. Thus if (2.1) is the maintained model and
the statistic \( \hat{\tau} \) 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 \( \hat{\rho} \), \( \hat{\beta} \), and \( \hat{\tau} \), 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 \( \hat{\tau}_L \) and \( \hat{\tau} \) 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 \( \hat{\beta} \) and \( \hat{\tau}_L \) do not depend on the value of \( Y_0 \). The limiting distribution of \( \hat{\beta} \) does not depend on \( Y_0 \), but the small-sample distribution of \( \hat{\beta} \) will be influenced by \( Y_0 \).

In the derivations we assumed the \( e_i \) to be NID(0, \( \sigma^2 \)). The limiting distributions also hold for \( e_i \) that are independent and identically distributed nonnormal random variables with mean zero and variance \( \sigma^2 \). 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 \( \tau \) 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 \( e_t \sim \text{NID}(0, \sigma^2) \) 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 .05 tests of the hypothesis \( \rho = 1 \) were applied to each sample. The tests were \( \hat{\beta}, \hat{\tau}, \hat{\beta}_L, \hat{\tau}_L, \hat{\beta}_R, \hat{\tau}_R, Q_1, Q_5, Q_{10}, Q_{25} \), where \( Q_k \) is the Box–Pierce Q statistic defined in (1.3) with \( e_t = Y_t - Y_{t-1} \).

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 \( \hat{\beta} \) and \( \hat{\tau} \) were similar, and they were uniformly more powerful than the other test statistics. It is not surprising that \( \hat{\beta} \) and \( \hat{\tau} \) are superior to \( \hat{\beta}_L \) and \( \hat{\tau}_L \) because \( \hat{\beta} \) and \( \hat{\tau} \) use the knowledge that the true value of the intercept in the regression is zero.

Third, for \( \rho < 1 \) the statistic \( \hat{\beta}_L \) yielded a more powerful test than the statistic \( \hat{\tau}_L \). For \( \rho > 1 \) the ranking was reversed and the \( \hat{\tau}_L \) 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 \( \hat{\tau}_L \) are biased tests, accepting the null hypothesis more than 95 percent of the time for \( \rho \) 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,

\[
X_t - X_1 = \rho(X_{t-1} - X_1) + e_t \quad (5.1)
\]

and

\[
X_t = \mu + \rho X_{t-1} + e_t, \quad (5.2)
\]
were fit to the data. For (5.1) the estimates were
\[ \hat{X}_1 - X_1 = 1.0044(X_{t-1} - X_1), \quad \hat{\sigma}^2 = .0052 \]
and for (5.2),
\[ \hat{X}_1 = .0141 + .9702X_{t-1}, \quad \hat{\sigma}^2 = .0050. \]
(0.0176) (0.0199)
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
\[ n(\beta - 1) = 91(.0044) = .4004 \]
and
\[ \hat{\tau} = (.0094)^{-1}(0.0044) = .4681. \]
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
\[ n(\rho_0 - 1) = 92(.9702 - 1) = .2742 \]
and
\[ \hat{\tau} = (.0199)^{-1}(.9702 - 1.0) = .150. \]
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
\[ Y_t = \beta_0 + \beta_1t + \alpha_1Y_{t-1} + \alpha_2Y_{t-2} + \epsilon_t, \]
where \( \epsilon_t \) 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
\[ Y_t - Y_{t-1} = \beta_0 + \beta_1t + (\alpha_1 + \alpha_2 - 1)Y_{t-1} - \alpha_3(Y_{t-1} - Y_{t-2}) + \epsilon_t, \]
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
\[ \hat{Y}_t - Y_{t-1} = .52 + .00120t - .119Y_{t-1} \]
(0.15) (0.0034) (.033)
+ .498(Y_{t-1} - Y_{t-2}), \quad \hat{\sigma}^2 = .033. \]
(.031)
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 \( (n - p)(\beta - 1)(1 + \alpha_2)^{-1} \)
where \( \beta \) is the coefficient of \( Y_{t-1} \) and \( p \) is the number of parameters estimated, is approximately distributed as
\[ \chi^2(n - p, \hat{\sigma}^2). \]
Also the “t statistic” constructed by dividing the coefficient of \( Y_{t-1} \) by the regression standard error is approximately distributed as \( \hat{\tau}. \)
For this example we have
\[ (n - p)(\beta - 1)(1 + \alpha_2)^{-1} = 166(-.119)(.502)^{-1} = -25.1 \]
and
\[ \hat{\tau} = (.033)^{-1}(-.119) = -3.61. \]
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 \( \rho \) were more powerful for tests against stationarity than the \( \hat{\tau} \) statistics. In this example the test based on \( \rho \) rejects the hypothesis at a smaller size (.025) than that of the \( \hat{\tau} \) statistic (.05).

[Received November 1976. Revised November 1978.]

REFERENCES