Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root
Author(s): David A. Dickey and Wayne A. Fuller
Source: *Econometrica*, Vol. 49, No. 4 (Jul., 1981), pp. 1057-1072
Published by: The Econometric Society
Stable URL: http://www.jstor.org/stable/1912517

Econometrica, Vol. 49, No. 4 (July, 1981)

LIKELIHOOD RATIO STATISTICS FOR AUTOREGRESSIVE TIME SERIES WITH A UNIT ROOT

BY DAVID A. DICKEY AND WAYNE A. FULLER

Let the time series Y_t satisfy $Y_t = \alpha + \rho Y_{t-1} + e_t$, where Y_1 is fixed and the e_t are normal independent $(0, \sigma^2)$ random variables. The likelihood ratio test of the hypothesis that $(\alpha, \rho) = (0, 1)$ is investigated and a limit representation for the test statistic is presented. Percentage points for the limiting distribution and for finite sample distributions are estimated. The distribution of the least squares estimator of α is also discussed. A similar investigation is conducted for the model containing a time trend.

1. INTRODUCTION

LET Y_t SATISFY THE MODEL

$$Y_t = \alpha + \rho Y_{t-1} + e_t \qquad (t = 2, 3, \ldots, n), \tag{1.1}$$

where Y_1 is fixed and $\{e_t\}$ is a sequence of normal independent random variables with mean 0 and variance σ^2, $[e_t \sim \mathrm{NI}(0, \sigma^2)]$. The maximum likelihood estimators of ρ and α, conditional on Y_1, are the least squares estimators

$$\hat{\rho}_\mu = \left[\sum_{t=2}^{n} (Y_{t-1} - \bar{y}_{(-1)})^2\right]^{-1} \sum_{t=2}^{n} (Y_t - \bar{y}_{(0)})(Y_{t-1} - \bar{y}_{(-1)}), \tag{1.2}$$

$$\hat{\alpha}_\mu = \bar{y}_{(0)} - \hat{\rho}_\mu \bar{y}_{(-1)},$$

where $\bar{y}_{(i)} = (n-1)^{-1}\sum_{t=2}^{n} Y_{t+i}$ for $i = -1, 0$.

The statistic constructed by analogy to the regression "t statistic" for the estimated α is

$$\hat{\tau}_{\alpha\mu} = S_{\alpha\mu}^{-1}\hat{\alpha}_\mu,$$

where

$$S_{\alpha\mu}^2 = S_{e\mu}^2\left[(n-1)^{-1} + \bar{y}_{(-1)}^2\left\{\sum_{t=2}^{n}(Y_{t-1} - \bar{y}_{(-1)})^2\right\}^{-1}\right],$$

$$S_{e\mu}^2 = (n-3)^{-1}\sum_{t=2}^{n}(Y_t - \hat{\alpha}_\mu - \hat{\rho}_\mu Y_{t-1})^2.$$

An alternative model for Y_t is

$$Y_t = \alpha + \beta(t - 1 - \tfrac{1}{2}n) + \rho Y_{t-1} + e_t \qquad (t = 2, 3, \ldots, n), \tag{1.3}$$

where Y_1 is fixed and $e_t \sim \mathrm{NI}(0, \sigma^2)$. Let $\boldsymbol{X}$ denote the $(n-1) \times 3$ matrix whose ith row is $(1, i - \frac{1}{2}n, Y_i)$ and let $\boldsymbol{Y}' = (Y_2, Y_3, \ldots, Y_n)$. Then the least squares estimator of $\boldsymbol{\theta} = (\alpha, \beta, \rho)'$ is

$$\hat{\boldsymbol{\theta}}_\tau = (\hat{\alpha}_\tau, \hat{\beta}_\tau, \hat{\rho}_\tau)' = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}. \tag{1.4}$$

Let C_{ij} denote the ijth element of $(\boldsymbol{X}'\boldsymbol{X})^{-1}$. Then the "regression t statistics" are

$$\hat{\tau}_{\alpha\tau} = \left(C_{11}S_{e\tau}^2\right)^{-\frac{1}{2}}\hat{\alpha}_\tau, \tag{1.5}$$

$$\hat{\tau}_{\beta\tau} = \left(C_{22}S_{e\tau}^2\right)^{-\frac{1}{2}}\hat{\beta}_\tau, \tag{1.6}$$

where

$$S_{e\tau}^2 = (n-4)^{-1}\boldsymbol{Y}'\left[\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\right]\boldsymbol{Y}. \tag{1.7}$$

We shall study the likelihood ratio test of the hypothesis that $(\alpha, \rho) = (0, 1)$ for model (1.2), the likelihood ratio test of the hypothesis that $(\alpha, \beta, \rho) = (0, 0, 1)$ for model (1.3), and the likelihood ratio test of the hypothesis that $(\alpha, \beta, \rho) = (\alpha, 0, 1)$ for model (1.3). We also investigate the distributions of $\hat{\alpha}_\mu, \hat{\tau}_{\alpha\mu}, \hat{\alpha}_\tau, \hat{\beta}_\tau, \hat{\tau}_{\alpha\tau}$, and $\hat{\tau}_{\beta\tau}$ under the null model.

The likelihood ratio statistics are derived in Section 2 and the limiting distributions presented in Section 3. Percentage points for the distributions obtained by Monte Carlo methods are given in Section 4. In Section 5, it is shown that the limit distributions of the test statistics are unchanged when $\{e_t\}$ in (1.1) and (1.3) is replaced by a stationary pth order autoregressive process whose coefficients must be estimated. In Section 6 the powers of the likelihood ratio tests Φ_1, Φ_2, and Φ_3 are compared to the powers of other test statistics. Section 7 contains an illustration of the use of the test statistics.

2. LIKELIHOOD RATIO STATISTICS

We construct the likelihood ratio statistics for the null hypothesis that the true model is a random walk with zero drift. We consider first the test that $(\alpha, \beta, \rho) = (0, 0, 1)$ in model (1.3) against the alternative that the null is not true. The logarithm of the likelihood function for a sample of n observations from model (1.3), conditional on Y_1, is

$$\log L = -\tfrac{1}{2}(n-1)\log(2\pi) - (n-1)\log\sigma - (2\sigma^2)^{-1}\sum_{t=2}^{n}\left[Y_t - \alpha - \beta(t - 1 - \tfrac{1}{2}n) - \rho Y_{t-1}\right]^2.$$

Under the null hypothesis, $H_0: (\alpha, \beta, \rho) = (0, 0, 1)$, the likelihood is maximized with respect to σ^2 to obtain

$$\hat{\sigma}_0^2 = (n-1)^{-1}\sum_{t=2}^{n}(Y_t - Y_{t-1})^2.$$

Under the alternative hypothesis the maximum of the likelihood occurs at

$(\hat{\sigma}_1^2, \hat{\boldsymbol{\theta}}_\tau')$, where $\hat{\boldsymbol{\theta}}_\tau$ was defined in (1.4) and $\hat{\sigma}_1^2 = (n-4)(n-1)^{-1}S_{e\tau}^2$. Therefore the likelihood ratio is

$$\left[\hat{\sigma}_0^{-1}\hat{\sigma}_1\right]^{n-1} = \left[1 + 3(n-4)^{-1}\Phi_2\right]^{\frac{1}{2}(n-1)},$$

where

$$\Phi_2 = \left(3S_{e\tau}^2\right)^{-1}\left[(n-1)\hat{\sigma}_0^2 - (n-4)S_{e\tau}^2\right].$$

Thus, the likelihood ratio test rejects the null hypothesis for large values of Φ_2, where Φ_2 is the usual regression "F test" of the hypothesis $H_0 : (\alpha, \beta, \rho) = (0, 0, 1)$.

In a similar manner, it can be shown that the likelihood ratio statistic for testing $H_0 : (\alpha, \rho) = (0, 1)$ against H_A : Not H_0, for the model (1.1) is

$$\left[1 + 2(n-3)^{-1}\Phi_1\right]^{\frac{1}{2}(n-1)},$$

where

$$\Phi_1 = \left(2S_{e\mu}^2\right)^{-1}\left[(n-1)\hat{\sigma}_0^2 - (n-3)S_{e\mu}^2\right].$$

The likelihood ratio test of the hypothesis $H_0 : (\alpha, \beta, \rho) = (\alpha, 0, 1)$ in model (1.3) is a monotone function of

$$\Phi_3 = \left(2S_{e\tau}^2\right)^{-1}\left[(n-1)\left\{\hat{\sigma}_0^2 - \left(\bar{y}_{(0)} - \bar{y}_{(-1)}\right)^2\right\} - (n-4)S_{e\tau}^2\right].$$

The statistics Φ_2 and Φ_3 are the common regression "F tests" one would construct for the hypotheses. The null hypothesis for test Φ_3 is that the time series is a random walk with drift α. It is easily demonstrated that the distribution of the test statistic Φ_3 does not depend upon α.

3. LIMITING DISTRIBUTIONS

Under the null hypotheses, the statistics introduced in Sections 1 and 2 can be expressed as functions of a few sample statistics. Let

$$\text{(3.1)} \qquad \Gamma_n = (n-1)^{-2}\sum_{t=2}^{n}\left(\sum_{j=1}^{t-1} e_j\right)^2,$$

$$T_n = (n-1)^{-\frac{1}{2}}\sum_{t=2}^{n} e_t = (n-1)^{\frac{1}{2}}\bar{e}_{(0)},$$

$$W_n = (n-1)^{-\frac{3}{2}}\sum_{t=1}^{n-1}(n-t)e_t = (n-1)^{-\frac{1}{2}}\bar{y}_{(-1)},$$

$$V_n = (n-1)^{-5/2}\sum_{t=1}^{n-1}(n-t)(t-1)e_t.$$

Then, for example,

$$(n-1)^{\frac{1}{2}}\hat{\alpha}_\mu = T_n - (n-1)(\hat{\rho}_\mu - 1)W_n$$

and

$$(n-1)(\hat{\rho}_\mu - 1) = \left(\Gamma_n - W_n^2\right)^{-1}\left\{\frac{1}{2}\left[\left(T_n + (n-1)^{-\frac{1}{2}}e_1\right)^2 - (n-1)^{-1}\sum_{t=1}^{n} e_t^2\right] - T_n W_n\right\}$$

$$= \left(\Gamma_n - W_n^2\right)^{-1}\left\{\tfrac{1}{2}\left(T_n^2 - \sigma^2\right) - T_n W_n\right\} + O_p(n^{-\frac{1}{2}}).$$

Because T_n and W_n are odd functions of $(e_1, e_2, \dots, e_n) = \boldsymbol{e}_n$ and Γ_n is an even function of $\boldsymbol{e}_n$, the distributions of $\hat{\alpha}_\mu$ and $\hat{\tau}_{\alpha\mu}$ are symmetric. Given that $\sigma^2 = 1$, Dickey and Fuller [7] have shown that $[\Gamma_n, T_n, W_n, V_n, n(\hat{\rho}_\mu - 1)]$ converges in distribution to $(\Gamma, T, W, V, \delta)$, where

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

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

$$\delta = \left(\Gamma - W^2\right)^{-1}\left[\tfrac{1}{2}(T^2 - 1) - TW\right],$$

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

and $\{Z_i\}_{i=1}^{\infty}$ is a sequence of normal independent (0, 1) random variables. Therefore, given that $(\alpha, \rho) = (0, 1)$,

$$n^{\frac{1}{2}}\sigma^{-1}\hat{\alpha}_\mu \xrightarrow{\mathcal{L}} T - \delta W,$$

and

$$\hat{\tau}_{\alpha\mu} \xrightarrow{\mathcal{L}} (T - \delta W)(\Gamma - W^2)^{\frac{1}{2}}\Gamma^{-\frac{1}{2}}, \tag{3.3}$$

because $S_{e\mu}^2$ converges in probability to σ^2.

For model (1.3) with the assumption that $\boldsymbol{\theta}' = (0, 0, 1)$, we have

$$\mathbf{X}'\mathbf{X} = (n-1)\begin{bmatrix} 1 & 0 & (n-1)^{\frac{1}{2}}W_n \\ 0 & 12^{-1}n(n-2) & \frac{1}{2}(n-1)^{\frac{3}{2}}V_n \\ (n-1)^{\frac{1}{2}}W_n & \frac{1}{2}(n-1)^{\frac{3}{2}}V_n & (n-1)\Gamma_n \end{bmatrix}.$$

Letting $\boldsymbol{D}_n = \operatorname{diag}[(n-1)^{\frac{1}{2}}, (n-1)^{\frac{3}{2}}, (n-1)\sigma]$, we obtain

$$\boldsymbol{D}_n^{-1}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{D}_n^{-1} \xrightarrow{\mathcal{L}} \boldsymbol{A}, \tag{3.4}$$

$$\sigma^{-1}\boldsymbol{D}_n^{-1}(\boldsymbol{X}'\boldsymbol{Y} - \boldsymbol{X}'\boldsymbol{X}\boldsymbol{\theta}) \xrightarrow{\mathcal{L}} \boldsymbol{f}, \tag{3.5}$$

where

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & W \\ 0 & \frac{1}{12} & \frac{1}{2}V \\ W & \frac{1}{2}V & \Gamma \end{bmatrix}$$

and $\boldsymbol{f}' = [T, \frac{1}{2}T - W, \frac{1}{2}(T^2 - 1)]$. The matrix $\boldsymbol{A}$ is invertible with probability 1 and it is readily verified that

$$\boldsymbol{A}^{-1} = Q^{-1}\begin{bmatrix} Q + W^2 & 6VW & -W \\ 6VW & 12Q + 36V^2 & -6V \\ -W & -6V & 1 \end{bmatrix}, \tag{3.6}$$

where $Q = \Gamma - W^2 - 3V^2$. Thus

$$\sigma^{-1}\boldsymbol{D}_n(\hat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}) \xrightarrow{\mathcal{L}} \boldsymbol{A}^{-1}\boldsymbol{f}.$$

The third element of $\boldsymbol{A}^{-1}\boldsymbol{f}$ is the limit random variable for $n(\hat{\rho}_\tau - 1)$ given in Dickey and Fuller [7]. Using the fact that $S_{e\tau}^2$ converges in probability to σ^2, we obtain

$$\hat{\tau}_{\alpha\tau} \xrightarrow{\mathcal{L}} Q^{\frac{1}{2}}(Q + W^2)^{-\frac{1}{2}}(1,0,0)\boldsymbol{A}^{-1}\boldsymbol{f},$$

$$\hat{\tau}_{\beta\tau} \xrightarrow{\mathcal{L}} Q^{\frac{1}{2}}(12Q + 36V^2)^{-\frac{1}{2}}(0,1,0)\boldsymbol{A}^{-1}\boldsymbol{f},$$

$$\Phi_1 \xrightarrow{\mathcal{L}} 2^{-1}\{T^2 + \delta^2(\Gamma - W^2)\},$$

$$\Phi_2 \xrightarrow{\mathcal{L}} 3^{-1}\boldsymbol{f}'\boldsymbol{A}^{-1}\boldsymbol{f} = 3^{-1}\left[T^2 + 12(\tfrac{1}{2}T - W)^2 + \hat{\tau}_\tau^2\right],$$

$$\Phi_3 \xrightarrow{\mathcal{L}} 2^{-1}(\boldsymbol{f}'\boldsymbol{A}^{-1}\boldsymbol{f} - T^2) = 2^{-1}\left[12(\tfrac{1}{2}T - W)^2 + \hat{\tau}_\tau^2\right],$$

and

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

The limiting distributions hold for any fixed Y_1 and for e_t a sequence of independent identically distributed random variables.

TABLE I

EMPIRICAL DISTRIBUTION OF $\hat{\tau}_{\alpha\mu}$ FOR $(\alpha, \rho) = (0, 1)$ IN $Y_t = \alpha + \rho Y_{t-1} + e_t$.
(Symmetric Distribution)

Sample size n	Probability of a smaller value			
	0.90	0.95	0.975	0.99
25	2.20	2.61	2.97	3.41
50	2.18	2.56	2.89	3.28
100	2.17	2.54	2.86	3.22
250	2.16	2.53	2.84	3.19
500	2.16	2.52	2.83	3.18
∞	2.16	2.52	2.83	3.18
s.e.	0.003	0.004	0.006	0.008

TABLE II

EMPIRICAL DISTRIBUTION OF $\hat{\tau}_{\alpha\tau}$ FOR $(\alpha, \beta, \rho) = (0, 0, 1)$ IN $Y_t = \alpha + \beta t + \rho Y_{t-1} + e_t$.
(Symmetric Distribution)

Sample size n	Probability of a smaller value			
	0.90	0.95	0.975	0.99
25	2.77	3.20	3.59	4.05
50	2.75	3.14	3.47	3.87
100	2.73	3.11	3.42	3.78
250	2.73	3.09	3.39	3.74
500	2.72	3.08	3.38	3.72
∞	2.72	3.08	3.38	3.71
s.e.	0.004	0.005	0.007	0.008

TABLE III

EMPIRICAL DISTRIBUTION OF $\hat{\tau}_{\beta\tau}$ FOR $(\alpha, \beta, \rho) = (0, 0, 1)$ IN $Y_t = \alpha + \beta t + \rho Y_{t-1} + e_t$
(Symmetric Distribution)

Sample size n	Probability of a smaller value			
	0.90	0.95	0.975	0.99
25	2.39	2.85	3.25	3.74
50	2.38	2.81	3.18	3.60
100	2.38	2.79	3.14	3.53
250	2.38	2.79	3.12	3.49
500	2.38	2.78	3.11	3.48
∞	2.38	2.78	3.11	3.46
s.e.	0.004	0.005	0.006	0.009

TABLE IV

EMPIRICAL DISTRIBUTION OF Φ_1 FOR $(\alpha, \rho) = (0, 1)$ IN $Y_t = \alpha + \rho Y_{t-1} + e_t$

Sample size n	Probability of a smaller value 0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
25	0.29	0.38	0.49	0.65	4.12	5.18	6.30	7.88
50	0.29	0.39	0.50	0.66	3.94	4.86	5.80	7.06
100	0.29	0.39	0.50	0.67	3.86	4.71	5.57	6.70
250	0.30	0.39	0.51	0.67	3.81	4.63	5.45	6.52
500	0.30	0.39	0.51	0.67	3.79	4.61	5.41	6.47
∞	0.30	0.40	0.51	0.67	3.78	4.59	5.38	6.43
s.e.	0.002	0.002	0.002	0.002	0.01	0.02	0.03	0.05

TABLE V

EMPIRICAL DISTRIBUTION OF Φ_2 FOR $(\alpha, \beta, \rho) = (0, 0, 1)$ IN $Y_t = \alpha + \beta t + \rho Y_{t-1} + e_t$

Sample size n	Probability of a smaller value 0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
25	0.61	0.75	0.89	1.10	4.67	5.68	6.75	8.21
50	0.62	0.77	0.91	1.12	4.31	5.13	5.94	7.02
100	0.63	0.77	0.92	1.12	4.16	4.88	5.59	6.50
250	0.63	0.77	0.92	1.13	4.07	4.75	5.40	6.22
500	0.63	0.77	0.92	1.13	4.05	4.71	5.35	6.15
∞	0.63	0.77	0.92	1.13	4.03	4.68	5.31	6.09
s.e.	0.003	0.003	0.003	0.003	0.01	0.02	0.03	0.05

TABLE VI

EMPIRICAL DISTRIBUTION OF Φ_3 FOR $(\alpha, \beta, \rho) = (\alpha, 0, 1)$ IN $Y_t = \alpha + \beta t + \rho Y_{t-1} + e_t$

Sample size n	Probability of a smaller value 0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
25	0.74	0.90	1.08	1.33	5.91	7.24	8.65	10.61
50	0.76	0.93	1.11	1.37	5.61	6.73	7.81	9.31
100	0.76	0.94	1.12	1.38	5.47	6.49	7.44	8.73
250	0.76	0.94	1.13	1.39	5.39	6.34	7.25	8.43
500	0.76	0.94	1.13	1.39	5.36	6.30	7.20	8.34
∞	0.77	0.94	1.13	1.39	5.34	6.25	7.16	8.27
s.e.	0.004	0.004	0.003	0.004	0.015	0.020	0.032	0.058

4. SIMULATION

Tables I–VI contain percentiles for the null distributions of the "regression t statistics" ($\hat{\tau}_{\alpha\mu}, \hat{\tau}_{\alpha\tau}, \hat{\tau}_{\beta\tau}$), and "regression F tests" (Φ_1, Φ_2, Φ_3). The null model is given in each table.

The empirical distributions of the statistics for finite samples were created from statistics for samples generated by the model with $Y_1 = 0$ and $Y_t = Y_{t-1} + e_t$, $t = 2, 3, \ldots, n$, for $n = 25$, 50, 100, 250, and 500. Three replicates of 50,000 samples were generated for $n = 25$, two for $n = 50$, 100, and 250, and one for $n = 500$. The simulation of the limit case was conducted using the procedure given in Dickey [**6**]. Three replicates of 50,000 were generated for the limit case. For symmetric distributions, cells equidistant from zero were pooled to create a symmetric histogram.

For each of the six estimators and for each sample size, the 0.01, 0.025, 0.05, 0.10, 0.90, 0.95, 0.975, and 0.99 percentage points of the distributions were calculated. These empirical percentiles were then plotted against n. Based on the plots, regression functions of the form $P = \alpha + \beta n^{\gamma}$ were fitted to the percentiles of the empirical distributions. Because several observations on each percentile were available for $n = 25$, 50, 100, 250, and for the limit case, regression F tests for lack of fit for the smoothing regressions were computed. Of the 36 lack of fit statistics computed, 7 were significant at the 0.25 level, 2 at the 0.05 level, and none at the 0.01 level. The regression smoothed percentiles are given in Tables I through VI.

David [**5**, Section 2.5] gives a method for constructing distribution free confidence intervals for the percentiles of a distribution based on empirical percentiles. In Tables I through VI the number in the row labeled "s.e" is the largest of the two half lengths of the 68.26% confidence intervals constructed for $n = 25$ and for the limit case. These entries provide an upper bound for the estimated standard errors of the regression smoothed percentiles.

The histogram $\hat{\tau}_{\alpha\mu}$ for 50,000 samples with $n = 25$ is shown in Figure 1. Figure 2 contains the histogram for $\hat{\tau}_{\beta\tau}$ constructed from 50,000 samples of size $n = 25$

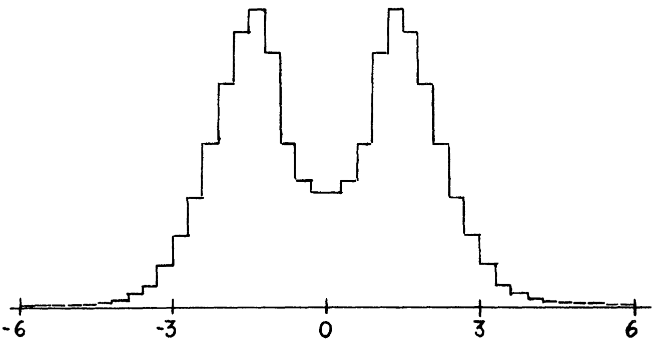

FIGURE 1.—Histogram for 50,000 values of $\hat{\tau}_{\alpha\mu}$ constructed with $n = 25$.

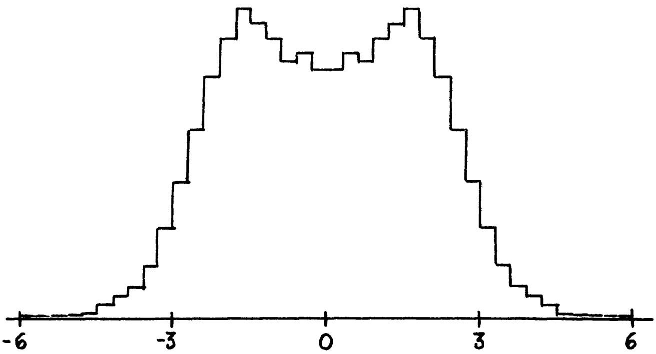

FIGURE 2.—Histogram for 50,000 values of $\hat{\tau}_{\beta\tau}$ constructed with $n = 25$.

generated with $\boldsymbol{\theta}' = (0, 0, 1)$. The distributions are symmetric and the histograms were constructed to be symmetric. The distributions of the τ statistics are distinctive in two respects; the distribution is bimodal and the "spread" of the distribution is much larger than that of Student's t distribution.

5. DISTRIBUTIONS FOR HIGHER ORDER PROCESSES

In this section we demonstrate that the test statistics investigated in the previous sections can be applied in higher order autoregressive processes. Consider data generated by the model

$$
\begin{aligned}
(5.1)\qquad & Y_1 = 0, \\
& Y_t = Y_{t-1} + Z_t \qquad (t = 2, 3, \ldots),
\end{aligned}
$$

where

$$Z_t = \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \ldots + \theta_p Z_{t-p} + e_t$$

is a stationary autoregressive process and the e_t are NID$(0, \sigma^2)$ random variables. The model can also be written

$$Y_t = \rho Y_{t-1} + \sum_{i=1}^{p} \theta_i (Y_{t-i} - Y_{t-1-i}) + e_t,$$

where $\rho = 1$ and $Z_t = Y_t - Y_{t-1}$. To simplify the presentation we assume, without loss of generality, $\sigma^2 = 1$.

Consider the regression equation

$$Y_{t+p} = \alpha + \beta\left[t - \tfrac{1}{2}(n - p + 1)\right] + \rho Y_{t+p-1} + \sum_{i=1}^{p} \theta_i Z_{t+p-i} + e_{t+p},$$

$t = 1, 2, \ldots, n - p$. Let $\boldsymbol{H}_n$ denote the $(p + 3) \times (p + 3)$ sums of squares and products matrix needed to compute the regression, let $\boldsymbol{M}_n$ denote the square roots of the diagonal elements of H_n, let $\gamma_n' = (\alpha, \beta, \rho, \theta_1, \theta_2, \ldots, \theta_p)$, and let $\hat{\gamma}_n'$ denote the least squares estimator of γ_n'. Then

$$M_n(\hat{\gamma}_n - \gamma_n) = \left[M_n^{-1} H_n M_n^{-1} \right]^{-1} M_n^{-1} g_n,$$

where

$$g_n' = \sum_{t=1}^{n-p} \left(1, \left[t - \tfrac{1}{2}(n-p+1)\right], Y_{t+p-1}, Z_{t+p-1}, \ldots, Z_t\right)' e_{t+p}.$$

Fuller [8, p. 374] has demonstrated that $n^{-\frac{1}{2}} Y_t$ is converging to $n^{-\frac{1}{2}} \cdot (1 - \sum_{i=1}^{p} \theta_i)^{-1} \sum_{j=1}^{t} e_j$ as t increases. By the results of Fuller, we have

$$n^{-1} \sum_{t=2}^{n} Z_t = O_p(n^{-\frac{1}{2}}),$$

$$\sum_{t=2}^{n} Y_{t-1}^2 = O_p(n^2),$$

$$n^{-2} \sum_{t=1}^{n} \left[t - \tfrac{1}{2}(n+p-1)\right] Z_{t+p-j} = O_p(n^{-\frac{1}{2}}),$$

$$\sum_{t=2}^{n} Y_{t-1} Z_{t-1} = O_p(n).$$

Therefore

$$\operatorname{plim} M_n^{-1} H_n M_n^{-1} = \text{block diag}(H_{11}, H_{22}),$$

$$H_{11} = \begin{bmatrix} 1 & 0 & \Gamma^{-\frac{1}{2}} W \\ 0 & 1 & \Gamma^{-\frac{1}{2}} 3^{\frac{1}{2}} V \\ \Gamma^{-\frac{1}{2}} W & \Gamma^{-\frac{1}{2}} 3^{\frac{1}{2}} V & 1 \end{bmatrix},$$

H_{22} is the $p \times p$ correlation matrix of the process Z_t, and Γ, W, and V are as defined in (3.2). It follows that the limiting distribution of the vector composed of the first three elements of $M_n(\hat{\gamma}_n - \gamma_n)$ is the same as the limiting distribution of

$$\left[n^{\frac{1}{2}} \hat{\alpha}_\tau, n^{\frac{3}{2}} \hat{\beta}_\tau, \left(\sum_{t=2}^{n} Y_{t-1}^2 \right)^{\frac{1}{2}} (\hat{\rho}_\tau - 1) \right]$$

discussed in Section 3. Similar results are easily obtained for the regression that does not contain the time trend.

6. EMPIRICAL POWER OF TESTS

Tables VII–IX were constructed to give information on the power of the tests. In Table VII the power was computed for samples of size $n = 100$ generated for model (1.1) with $\rho = 0.8, 0.9, 0.95, 1.00, 1.02$, and 1.05 and $\alpha = 0.0, 0.5$, and 1.0.

TABLE VII

EMPIRICAL POWER OF TWO SIDED SIZE 0.05 TESTS FOR SAMPLES OF SIZE 100(Y_1 FIXED)

	$\rho = 0.8$			$\rho = 0.9$			$\rho = 0.95$			$\rho = 0.99$			$\rho = 1.0$			$\rho = 1.02$			$\rho = 1.05$		
	α			α			α			α			α			α			α		
Statistic	0.00	0.50	1.00	0.00	0.50	1.00	0.00	0.50	1.00	0.00	0.50	1.00	0.00	0.50	1.00	0.00	0.50	1.00	0.00	0.50	1.00
Φ_1	0.78	0.83	0.93	0.22	0.36	0.73	0.08	0.21	0.91	0.04	0.76	1.00	0.05	0.98	1.00	0.43	1.00	1.00	0.95	1.00	1.00
Φ_2	0.41	0.47	0.65	0.09	0.13	0.39	0.04	0.09	0.63	0.03	0.67	1.00	0.05	0.97	1.00	0.45	1.00	1.00	0.95	1.00	1.00
Φ_3	0.57	0.57	0.72	0.15	0.19	0.43	0.07	0.11	0.37	0.05	0.06	0.13	0.05	0.05	0.04	0.08	1.00	1.00	0.94	1.00	1.00
$n(\hat{\rho}_\mu - 1)$	0.86	0.87	0.88	0.29	0.25	0.13	0.11	0.02	0.00	0.05	0.03	0.00	0.05	0.14	0.07	0.47	1.00	1.00	0.97	1.00	1.00
$\hat{\tau}_\mu$	0.71	0.77	0.89	0.18	0.27	0.56	0.06	0.12	0.53	0.04	0.07	0.14	0.06	0.26	0.35	0.49	1.00	1.00	0.97	1.00	1.00
$n(\hat{\rho}_\tau - 1)$	0.57	0.55	0.53	0.14	0.10	0.05	0.06	0.04	0.01	0.05	0.07	0.12	0.05	0.05	0.05	0.18	0.99	1.00	0.96	1.00	1.00
$\hat{\tau}_\tau$	0.46	0.50	0.62	0.10	0.12	0.22	0.04	0.05	0.09	0.05	0.06	0.11	0.05	0.05	0.05	0.17	0.97	1.00	0.96	1.00	1.00

For $\alpha = 0$ power is computed from 10,000 samples.
For $\alpha \neq 0$ power is computed from 3000 samples.

TABLE VIII

EMPIRICAL POWER OF TWO SIDED SIZE 0.05 TESTS AGAINST THE STATIONARY ALTERNATIVE FOR SAMPLES OF n = 50, 100, AND 250 (20,000 SAMPLES)

	$\rho = 0.8$			$\rho = 0.9$			$\rho = 0.95$			$\rho = 0.99$			$\rho = 1.0$		
	n			n			n			n			n		
Statistic	50	100	250	50	100	250	50	100	250	50	100	250	50	100	250
Φ_1	0.25	0.80	1.0	0.09	0.24	0.94	0.05	0.09	0.37	0.05	0.05	0.05	0.05	0.05	0.05
Φ_2	0.09	0.44	1.0	0.04	0.10	0.64	0.04	0.04	0.15	0.04	0.04	0.04	0.05	0.05	0.05
Φ_3	0.17	0.59	1.0	0.08	0.16	0.78	0.06	0.07	0.23	0.06	0.05	0.06	0.05	0.05	0.05
$n(\hat{\rho}_\mu - 1)$	0.30	0.87	1.0	0.10	0.29	0.97	0.05	0.09	0.43	0.05	0.04	0.05	0.05	0.05	0.05
τ_μ	0.20	0.74	1.0	0.07	0.19	0.91	0.04	0.06	0.31	0.04	0.04	0.04	0.05	0.05	0.05
$n(\hat{\rho}_\tau - 1)$	0.14	0.57	1.0	0.06	0.14	0.78	0.05	0.05	0.21	0.05	0.04	0.04	0.05	0.05	0.05
τ_τ	0.11	0.48	1.0	0.05	0.11	0.69	0.04	0.05	0.16	0.05	0.04	0.04	0.05	0.05	0.05

TABLE IX

EMPIRICAL POWER OF ONE SIDED SIZE 0.05 TESTS FOR SAMPLES OF SIZE 100(Y_1 FIXED)

	$\rho = 0.8$			$\rho = 0.9$			$\rho = 0.95$			$\rho = 0.99$			$\rho = 1.00$		
	α			α			α			α			α		
Statistic	0.00	0.50	1.00	0.00	0.50	1.00	0.00	0.50	1.00	0.00	0.50	1.00	0.00	0.50	1.00
d	0.97	0.95	0.89	0.53	0.31	0.03	0.23	0.01	0.00	0.08	0.00	0.00	0.05	0.00	0.00
$n(\hat{\rho}_\mu - 1)$	0.95	0.95	0.96	0.46	0.42	0.29	0.19	0.06	0.00	0.07	0.00	0.00	0.05	0.00	0.00
τ_μ	0.86	0.90	0.95	0.30	0.43	0.73	0.12	0.21	0.66	0.06	0.06	0.20	0.05	0.00	0.00
$n(\hat{\rho}_\tau - 1)$	0.73	0.72	0.72	0.24	0.20	0.12	0.10	0.06	0.01	0.05	0.04	0.02	0.05	0.05	0.06
τ_τ	0.64	0.67	0.78	0.18	0.22	0.34	0.08	0.09	0.15	0.05	0.04	0.03	0.05	0.05	0.05

For $\alpha = 0$ power is computed from 10,000 samples.
For $\alpha \neq 0$ power is computed from 3000 samples.

The statistic Φ_1 is the likelihood ratio test of $(\alpha, \rho) = (0, 1)$ against the alternative $(\alpha, \rho) \neq (0, 1)$ for model (1.1). The statistic Φ_2 is the likelihood ratio test of $(\alpha, \beta, \rho) = (0, 0, 1)$ against the general alternative of model (1.3). Note that in models (1.1) and (1.3) the initial value Y_1 is fixed. Because the alternative is broader for Φ_2, Φ_2 displays smaller power than Φ_1 in Table VII where the parameter $\beta = 0$ for all examples of the table. Both Φ_1 and Φ_2 display bias, having power less than the size for $\rho = 0.99$. The power of both tests increases as α increases.

The statistic Φ_3 is the likelihood ratio test of $(\alpha, \beta, \rho) = (\alpha, 0, 1)$ against the general alternative of model (1.3). In Table VII the power of Φ_3 is between those of Φ_1 and Φ_2 for $\rho < .99$. At $\rho = 1.02$ and $\alpha = 0$, the power of Φ_3 is considerably less than the powers of Φ_1 and Φ_2. No bias is evident in Φ_3.

We have included in Table VII the statistics $\hat{\rho}_\mu, \hat{\rho}_\tau, \hat{\tau}_\mu$, and $\hat{\tau}_\tau$ discussed by Fuller [**8**, Section 8.5]. The null distributions of the statistics $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ are computed under model (1.1) with the assumption that $(\alpha, \rho) = (0, 1)$. The distributions of the statistics for $(\alpha, 1), \alpha \neq 0$ differ from those with $\alpha = 0$. The null distributions of the statistics $\Phi_3, \hat{\rho}_\tau$, and $\hat{\tau}_\tau$ are independent of α and therefore they maintain their size for $\rho = 1$ and $\alpha \neq 0$. The tests used for Table VII were constructed from $\hat{\rho}_\mu, \hat{\rho}_\tau, \hat{\tau}_\mu$, and $\hat{\tau}_\tau$ by removing equal areas from the two tails of the distribution. The tests $\hat{\tau}_\mu$ and $\hat{\tau}_\tau$ are generally less powerful than the corresponding tests Φ_1 and Φ_2 when $\rho < 1$.

Table VIII contains the estimated power for the test statistics against the stationary first order autoregressive time series. The tests are the two sided tests of Table VII. The observations for $\rho \neq 1$ for the samples of Table VIII were generated using the model

$$Y_t = \rho Y_{t-1} + e_t \qquad (t = 2, 3, \ldots, n),$$

$$Y_t = (1 - \rho^2)^{-\frac{1}{2}} e_1 \qquad (t = 1),$$

where the e_t are NID(0, 1) random variables. The power was computed for 20,000 samples of each of the three sample sizes. Generally speaking $\hat{\rho}_\mu$ is the most powerful of the tests considered. The test Φ_3 is the most powerful of the tests that permit the null model to contain drift.

Table IX contains the power for one sided tests when the true model is (1.1) with $\rho \leqslant 1$. Included in this table is the von Neumann ratio

$$d = n\left[\sum_{t=1}^{n} (Y_t - \bar{y}_n)^2\right]^{-1} \sum_{t=2}^{n} (Y_t - Y_{t-1})^2,$$

where

$$\bar{y}_n = n^{-1} \sum_{t=1}^{n} Y_t.$$

Sargan [15] gives percentiles for d when Y_t is generated by model (1.1) with $(\alpha, \rho) = (0, 1)$. For sample sizes 50 and 100 and significance level 0.05 we use the percentiles from Sargan's paper. The jth percentile of the limit distribution for d is the reciprocal of the $(100 - j)$th percentile in the table of Anderson and Darling [2, p. 203]. For finite sample sizes not considered by Sargan, we use the asymptotic percentiles as critical values for the power calculations of Table IX. Fuller [9] has constructed modifications of the statistic d that are applicable to higher order autoregressive processes and to model (1.3). The methods used to generate the samples of Table IX are those used to generate the samples of Table VII with $\rho \leqslant 1$. The statistic d is an appropriate test when the alternative is that Y_t is a stationary first order autoregressive time series. It displays good power for this alternative (that is, when $\alpha = 0$ and $\rho < 1$). The statistic $n(\hat{\rho}_\mu - 1)$ is only slightly less powerful than d for $\alpha = 0$ and maintains somewhat better power for $\alpha \neq 0$. For $\rho < 1$ and $\alpha \neq 0$ the estimator $\hat{\rho}_\mu$ is closer to one on the average than the corresponding estimator associated with $\alpha = 0$. Therefore, the tests $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ display poor power for values of ρ close to one and $\alpha \neq 0$. Because the estimator $\hat{\rho}_\mu$ converges to ρ for $\rho \leqslant 1$, there is some sample size for any $\rho < 1$ for which the

TABLE X

MODELS AND TEST STATISTICS

Null Model	Alternative Model	Test Statistic
$Y_t = Y_{t-1} + e_t$	$Y_t = \alpha + \rho Y_{t-1} + e_t$	Φ_1
$Y_t = Y_{t-1} + e_t$	$Y_t = \alpha + \beta t + \rho Y_{t-1} + e_t$	Φ_2
$Y_t = \alpha + Y_{t-1} + e_t$	$Y_t = \alpha + \beta t + \rho Y_{t-1} + e_t$	Φ_3
$Y_t = Y_{t-1} + e_t$	$Y_t = \alpha + \rho Y_{t-1} + e_t$ $\rho \neq 1$	$\hat{\rho}_\mu, \hat{\tau}_\mu^{a}$
$Y_t = \alpha + Y_{t-1} + e_t$	$Y_t = \alpha + \beta t + \rho Y_{t-1} + e_t$ $\rho \neq 1$	$\hat{\rho}_\tau, \hat{\tau}_\tau^{a}$
$Y_t = Y_{t-1} + e_t$	$Y_t = \alpha + \rho Y_{t-1} + e_t$ $\lvert\rho\rvert < 1$	d

[a]Poor power for Y_t not stationary, $\alpha \neq 0$, small n, and ρ less than, but close to one.

statistics will have power greater than the size. Because the null distributions are derived under the assumption that $(\alpha, \rho) = (0, 1)$, there is no sample size for which the tests $d, \hat{\rho}_\mu$, and $\hat{\tau}_\mu$ are appropriate if the alternative includes $\alpha \neq 0$ and $\rho = 1$.

The test statistics discussed in this section and the hypotheses for which they are appropriate are summarized in Table X.

7. EXAMPLE

To illustrate the use of the tables we study the logarithm of the quarterly Federal Reserve Board Production Index 1950–1 through 1977–4. We assume that the time series is adequately represented by the model

$$(7.1) \qquad Y_t = \beta_0 + \beta_1 t + \alpha_1 Y_{t-1} + \alpha_2 (Y_{t-1} - Y_{t-2}) + e_t,$$

where e_t are independent identically distributed $(0, \sigma^2)$ random variables. The ordinary least squares estimates are

$$\hat{Y}_t - Y_{t-1} = \underset{(0.15)}{0.52} + \underset{(0.00034)}{0.00120t} - \underset{(0.033)}{0.119}\, Y_{t-1} + \underset{(0.081)}{0.498}\,(Y_{t-1} - Y_{t-2}),$$

$$\text{R.S.S.} = 0.056448,$$

$$\hat{Y}_t - Y_{t-1} = \underset{(0.0025)}{0.0054} + \underset{(0.083)}{0.447}\,(Y_{t-1} - Y_{t-2}), \quad \text{R.S.S.} = 0.063211,$$

$$\hat{Y}_t - Y_{t-1} = \underset{(0.079)}{0.511}\,(Y_{t-1} - Y_{t-2}), \quad \text{R.S.S.} = 0.065966,$$

where R.S.S. denotes the residual sum of squares. The numbers in parentheses are the quantities output as "standard errors" by the regression program.

To test the hypothesis that $\beta_0 = \beta_1 = 0$ and $\alpha_1 = 1$ against the general alternative (7.1) we compute

$$\Phi_2 = \frac{0.065966 - 0.056448}{3(0.000533)} = 5.95,$$

where $0.000533 = 0.056448/106$ is the residual mean square for the full model regression. As there are 110 observations in the regression the 97.5 per cent point of the distribution of Φ_2, as given in Table V, is 5.59. Therefore the hypothesis $\beta_0 = \beta_1 = 0$ and $\alpha_1 = 1$ is rejected at the 2.5 per cent level.

To test the hypothesis that $\beta_1 = 0$ and $\alpha_1 = 1$ against the general alternative (7.1) we compute

$$\Phi_3 = \frac{0.063211 - 0.056448}{2(0.000533)} = 6.34.$$

The 95 per cent point of the distribution is given in Table VI as 6.49 and the 90 per cent point as 5.47. Therefore at the 5 per cent level one could accept the hypothesis that the second order autoregressive process has a unit root with possible drift under the maintained hypothesis that the process is second order. The null hypothesis would be rejected at the 10 per cent level. We note that on the basis of Table 8.5.2 of Fuller [**8**] the statistic

$$\hat{\tau}_\tau = \frac{-0.119}{0.033} = -3.61$$

would lead to rejection of the hypothesis of a unit root at the 10 per cent level if a two sided test is performed. If the alternative is that both roots are less than one in absolute value the hypothesis of a unit root is rejected at the 5 per cent level.

North Carolina State University
and
Iowa State University

Manuscript received June, 1978; final revision received April, 1980.

REFERENCES

[1] ANDERSON, T. W.: "On Asymptotic Distributions of Estimates of Parameters of Stochastic Difference Equations," *Annals of Mathematical Statistics*, 30(1959), 676–687.

[2] ANDERSON, T. W., AND D. A. DARLING: "Asymptotic Theory of Certain 'Goodness of Fit' Criteria Based on Stochastic Processes," *Annals of Mathematical Statistics*, 23(1952), 193–212.

[3] BOX, G. E. P., AND G. M. JENKINS: *Time Series Analysis Forecasting and Control*. San Francisco: Holden-Day, 1970.

[4] BULMER, M. G.: "The Statistical Analysis of Density Dependency," *Biometrics*, 31(1975), 901–911.

[5] DAVID, H. A.: *Order Statistics*. New York: John Wiley and Sons, 1970.

[6] DICKEY, D. A.: "Estimation and Hypothesis Testing in Nonstationary Time Series," Iowa State University Ph.D. thesis, 1976.

[7] DICKEY, D. A., AND W. A. FULLER: "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, 74(1979), 427–431.

[8] FULLER, W. A.: *Introduction to Statistical Time Series*. New York: John Wiley and Sons, 1976.

[9] ———: "Testing the Autoregressive Process for a Unit Root," Report to the U.S. Census Bureau, 1979.

[10] HASZA, D.: "Estimation in Nonstationary Time Series," Iowa State University Ph.D. thesis, 1977.

[11] MANN, H. B., AND A. WALD: "On Stochastic Limit and Order Relationships," *Annals of Mathematical Statistics*, 14(1943), 217–226.

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

[13] ———: "Asymptotic Distribution of An Estimator of the Boundary Parameter of an Unstable Process," *Annals of Statistics*, 6(1978), 185–190.

[14] RUBIN, H.: "Consistency of Maximum-Likelihood Estimates in the Explosive Case," in *Statistical Inference in Dynamic Economic Models*, ed. by T. C. Koopmans, 1950.

[15] SARGAN, J. D.: "The Durbin-Watson Ratio of the Gaussian Random Walk," London School of Economics manuscript, 1973.