Biometrika (1986), **73**, 2, *pp.* 461-6
Printed in Great Britain

Nonlinearity tests for time series

BY RUEY S. TSAY

Department of Statistics, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213, U.S.A.

SUMMARY

This paper considers two nonlinearity tests for stationary time series. The idea of Tukey's one degree of freedom for nonadditivity test is generalized to the time series setting. The case of concurrent nonlinearity is discussed in detail. Simulation results show that the proposed tests are more powerful than that of Keenan (1985).

Some key words: Concurrent nonlinearity; Nonlinear time series; Tukey's nonadditivity test; Volterra expansion.

1. INTRODUCTION

Recently there has been a growing interest in studying nonlinear time series. In particular, various tests for testing linearity have been proposed to illustrate the nonlinear nature of certain well known processes (Subba Rao & Gabr, 1980; Hinich, 1982; Maravell, 1983; Hinich & Patterson, 1985) and to support the need for nonlinear time series models (Granger & Andersen, 1978). Recently Keenan (1985) adopted the idea of Tukey's (1949) one degree of freedom test for nonadditivity to derive a time-domain statistic, as an alternative of the frequency-domain statistics, e.g. bispectrum, for discriminating between nonlinear and linear models. Keenan's test is motivated by the similarity of Volterra expansions to polynomials, and is extremely simple both conceptually and computationally. However, as shown by Keenan's simulation in Table 1 of his paper, the power of his test could be very low.

In the present paper, we propose a modified test that retains the simplicity of Keenan's test yet is considerably more powerful. The new test statistic is given in § 2, and § 3 reviews numerical comparisons via simulation. Section 4 is devoted to the concurrent nonlinearity model for which Keenan's test fails, and § 5 briefly discusses the autoregressive-moving average models.

2. A TEST OF LINEARITY

A stationary time series Y_t can be written, in its very general form, as

$$Y_t = \mu + \sum_{i=-\infty}^{\infty} b_i e_{t-i} + \sum_{i,j=-\infty}^{\infty} b_{ij} e_{t-i} e_{t-j} + \sum_{i,j,k=-\infty}^{\infty} b_{ijk} e_{t-i} e_{t-j} e_{t-k} + \dots,$$

where μ is the mean level of Y_t and $\{e_t, -\infty < t < \infty\}$ is a strictly stationary process of independent and identically distributed random variables. Obviously, Y_t is nonlinear if any of the higher order coefficients, $\{b_{ij}\}$, $\{b_{ijk}\}, \dots$ is nonzero. For such a series, the proposed tests of this paper and that of Keenan (1985) are based on the following argument. Suppose, for the illustrative purpose, that b_{12} is nonzero. Then this nonlinearity will be distributionally reflected in the diagnostics of a fitted linear model if the residuals of the linear model are correlated with $Y_{t-1}Y_{t-2}$, a quadratic nonlinear term. In practice, since the orders of the higher order coefficients, if any, are unknown, Tukey's (1949) nonadditivity test simply uses the aggregated quantity $\hat{Y}_t^2$, the square of the fitted value of Y_t based on the entertained linear model, to obtain quadratic terms upon which the residuals can be correlated. This ingenious idea is extremely valuable when the sample size is small, because it only requires one degree of freedom. One disadvantage of using aggregated quantities, however, is that aggregation often loses potentially useful information. Thus, it seems preferable to employ disaggregated variables when the sample size is large or moderate, as is the case for most time series analyses.

The test proposed here is motivated by the above consideration and it consists of the following steps.

(i) Regress Y_t on $\{1, Y_{t-1}, \ldots, Y_{t-M}\}$ by least squares and obtain the residuals $\{\hat{e}_t\}$, for $t = M+1, \ldots, n$. The regression model will be denoted by

$$Y_t = W_t\Phi + e_t, \qquad (2\cdot1)$$

where $W_t = (1, Y_{t-1}, \ldots, Y_{t-M})$ and $\Phi = (\Phi_0, \Phi_1, \ldots, \Phi_M)^{\mathrm{T}}$ with M being a prespecified positive integer, n the sample size, and the superscript T denoting the matrix transpose.

(ii) Regress the vector Z_t on $\{1, Y_{t-1}, \ldots, Y_{t-M}\}$ and obtain the residual vector $\{\hat{X}_t\}$, for $t = M+1, \ldots, n$. Here the multivariate regression model is

$$Z_t = W_tH + X_t,$$

where Z_t is an $m = \frac{1}{2}M(M+1)$ dimensional vector defined by $Z_t^{\mathrm{T}} = \text{vech}\,(U_t^{\mathrm{T}}U_t)$ with $U_t = (Y_{t-1}, \ldots, Y_{t-M})$ and vech denoting the half stacking vector. In other words, Z_t^{T} is obtained from the symmetric matrix $U_i^{\mathrm{T}}U_t$ by the usual column stacking operator but using only those elements on or below the main diagonal of each column.

(iii) Regress $\hat{e}_t$ on $\hat{X}_t$ and let $\hat{F}$ be the F ratio of the mean square of regression to the mean square of error. That is, fit

$$\hat{e}_t = \hat{X}_t\beta + \varepsilon_t \quad (t = M+1, \ldots, n) \qquad (2\cdot2)$$

and define

$$\hat{F} = \{(\Sigma\, \hat{X}_t\hat{e}_t)(\Sigma\, \hat{X}_t^{\mathrm{T}}\hat{X}_t)^{-1}(\Sigma\, \hat{X}_t^{\mathrm{T}}\hat{e}_t)/m\}/\{\Sigma\, \hat{\varepsilon}_t^2/(n-M-m-1)\}, \qquad (2\cdot3)$$

where the summations are over t from $M+1$ to n and $\hat{\varepsilon}_t$ is the least squares residual for (2·2).

Obviously, this procedure reduces to Keenan's if one aggregates Z_t, with weights determined by the least squares estimate of (2·1), to become a scalar variable. Notice that a by-product of this disaggregated approach is that from the regression (2·2) one can easily identify the significant nonlinear terms to be incorporated in the model. Thus, the proposed testing procedure can be used as a diagnostic tool for building linear or nonlinear time series models.

THEOREM 1. *Let Y_t be a stationary autoregressive process of order M satisfying the model*

$$(Y_t - \mu) = \sum_{i=1}^{M} \Phi_i(Y_{t-i} - \mu) + e_t,$$

where the e_t's are independent and identically distributed random variables with mean zero, variance σ_e^2, and finite fourth moment. Then, for large n, the statistic $\hat{F}$ defined in (2·3) *follows approximately a F distribution with degrees of freedom*

$$\tfrac{1}{2}M(M+1), \quad n-\tfrac{1}{2}M(M+3)-1.$$

Proof. Let $\Phi_0=(1-\Phi_1-\ldots-\Phi_M)\mu$ and $\Phi=(\Phi_0,\Phi_1,\ldots,\Phi_M)^{\mathrm{T}}$. Then the least squares estimate $\hat{\Phi}$ of (2·1) converges to Φ almost surely under the conditions of the theorem (Lai & Wei, 1983). Using this result and Slutsky's Theorem (Bickel & Doksum, 1977, p. 461), and adopting an argument similar to that of Lemma 3.1 of Keenan (1985), it is clear that to prove the theorem it is sufficient to show that

$$n^{-\frac{1}{2}}\sum_{t=M+1}^{n} X_t^{\mathrm{T}} e_t. \tag{2·4}$$

converges in distribution to a multivariate normal random variable, where X_t is given in (2·2). Since X_t depends only on $\{Y_{t-j}, j>0\}$ which is independent of e_t, $X_t^{\mathrm{T}}e_t$ forms a stationary and ergodic martingale difference process. The asymptotic normality of (2·4) then follows from a multivariate version of a martingale central limit theorem (Billingsley, 1961). The associated covariance matrix is the limit of $n^{-1}\Sigma X_t^{\mathrm{T}}X_t$. Finally, the large-sample F distribution of the testing statistic $\hat{F}$ of (2·3) follows from an argument similar to that of the usual analysis of variance. □

Note that the limit of $\{\tfrac{1}{2}M(M+1)\}\hat{F}$ is a chi-squared random variable with degrees of freedom $\tfrac{1}{2}M(M+1)$. This is a straightforward generalization of Corollary 3·1 of Keenan (1985). In practice, we prefer to use the approximate F distribution.

3. Comparison

In this section we use simulation to compare the F test of § 2 with Keenan's test. For simplicity, the six models of Keenan (1985) are used in this study. They are:

Model 1, $Y_t=e_t-0·4e_{t-1}+0·3e_{t-2}$;
Model 2, $Y_t=e_t-0·4e_{t-1}+0·3e_{t-2}+0·5e_te_{t-2}$;
Model 3, $Y_t=e_t-0·3e_{t-1}+0·2e_{t-2}+0·4e_{t-1}e_{t-2}-0·25e_{t-2}^2$;
Model 4, $Y_t=0·4Y_{t-1}-0·3Y_{t-2}+e_t$;
Model 5, $Y_t=0·4Y_{t-1}-0·3Y_{t-2}+0·5Y_{t-1}e_{t-1}+e_t$;
Model 6, $Y_t=0·4Y_{t-1}-0·3Y_{t-2}+0·5Y_{t-1}e_{t-1}+0·8e_{t-1}+e_t$.

The e_t's are independent $N(0, 1)$ random variates generated from the GGNML subroutine of the IMSL package. Table 1 gives a summary in terms of empirical significance levels

Table 1. *Empirical frequencies of rejection of the null hypothesis of linearity*; $n=70$, 204; $M=4$, *and* 350 *replications. Nominal significance level*, 0·05.

True model	$n=70$ Keenan	$n=70$ F test	$n=204$ Keenan	$n=204$ F test
		(a) Linear		
Model 1	0·063	0·066	0·051	0·054
Model 4	0·060	0·066	0·051	0·046
		(b) Nonlinear		
Model 3	0·366	0·509	0·843	0·971
Model 5	0·534	0·760	0·811	0·986
Model 6	0·549	0·707	0·857	0·934
Model 2	0·100	0·166	0·097	0·217

and powers. For each model, the results are based on 350 replications of two combinations of $n=70$, $M=4$ and $n=204$, $M=4$. It is clear that the proposed F test is more powerful than Keenan's test in identifying the nonlinear models while the empirical significance levels of the two tests are reasonable and remain comparable for the linear models.

4. Concurrent nonlinearity

A striking result of Table 1 is that both Keenan's test and the F test are not powerful in handling Model 2, which contains a concurrent nonlinear term $e_t e_{t-2}$. In this section, we consider the problem of concurrent nonlinearity in further detail.

Obviously, the failure of Keenan's test and the F test is, to a large extent, because the constructed variable Z_t of step (ii) of § 2, hence X_t of step (iii), does not contain concurrent nonlinear terms. We therefore consider in step (ii) the variable $\hat{A}_t=(Y_{t-1}\hat{e}_t, \ldots, Y_{t-M}\hat{e}_t)$ and define in step (iii) the statistics

$$\hat{R}_t=\hat{A}_t\hat{e}_t-U_t\sigma_e^2, \tag{4·1}$$

$$\hat{C}=\frac{(\Sigma\, \hat{R}_t)(\Sigma\, \hat{R}_t^{\mathrm{T}}\hat{R}_t)^{-1}(\Sigma\, \hat{R}_t^{\mathrm{T}})/M}{\Sigma\, \hat{e}_t^2/(n-M-1)}, \tag{4·2}$$

where, again, $U_t=(Y_{t-1}, \ldots, Y_{t-M})$ and summations are over t from $M+1$ to n. The motivation for using $\hat{A}_t$ is obvious, and that for $\hat{R}_t$ will become clear in the following proof.

THEOREM 2. *Under the conditions of Theorem* 1, $\hat{C}$ *of* (4·2) *follows approximately a F distribution with degrees of freedom M and* $n-M-1$, *provided that n is sufficiently large.*

Proof. Using the consistency property of the least squares estimates in the same way as in Theorem 1, we need only show the asymptotic normality of the statistic

$$n^{-\frac{1}{2}}\sum_{t=M+1}^{n} R_t^{\mathrm{T}}=n^{-\frac{1}{2}}\sum_{t=M+1}^{n}\{Y_{t-1}(e_t^2-\sigma_e^2), \ldots, Y_{t-M}(e_t^2-\sigma_e^2)\}^{\mathrm{T}},$$

where R_t is the theoretical counterpart of $\hat{R}_t$. Let $\mathscr{F}_{t-1}=\{Y_{t-1}, Y_{t-2}, \ldots\}$ be the σ-field generated by the available information at time $t-1$. Then, it is clear that (i) $E(R_t^{\mathrm{T}}|\mathscr{F}_{t-1})=0$ and (ii) each component of R_t forms a stationary martingale difference process with constant variance $(\kappa-1)\sigma_e^4\{\mathrm{var}(Y_t)+\mu^2\}$, where κ is the fourth cumulant of e_t. The expected asymptotic normality follows, again, from Billingsley's (1961) martingale central limit theorem. Furthermore, since the process $\{Y_{t-j}e_t^2; j=1, \ldots, M\}$ is stationary and ergodic, $n^{-1}\Sigma\, R_t^{\mathrm{T}}R_t$ converges in probability to the variance-covariance matrix of R_t. □

Note that the limiting distribution of $M\hat{C}$ is chi-squared with M degrees of freedom. The critical value of a chi-squared table, therefore, can be used when n is large. Here we use the F distribution to match with that of § 2. In practice, σ_e^2 of (4·1) is unknown. However, under the null hypothesis of linear model, it can be estimated by the residual mean square of error of the regression (2·1).

Table 2 compares Keenan's test, the F test of § 2 and the concurrent C test of Theorem 2. Here Model 1 is used to show the significance level and Model 2 to illustrate the power of each test in recognizing the concurrent nonlinear model. Again, 350 replications of each model were used for a given sample size. It is clear from the table that the concurrent C test increases the power markedly over the other two in identifying concurrent nonlinear models, especially when the sample size is moderate or large. None of the tests, however, is very powerful when the sample size is small.

Since Z_t is a function of U_t that is orthogonal to $\hat{e}_t$, $\hat{C}$ of (4·2) in fact can be defined in terms of the residual $\hat{\varepsilon}_t$ of (2·2). That is, one may replace $\hat{e}_t$ by $\hat{e}_t$. This substitution allows us to use the F test of (2·3) first and then switch to $\hat{C}$, if necessary, for concurrent nonlinearity. The numbers in parentheses in Table 2 give the corresponding results based on $\hat{\varepsilon}_t$. For Models 1 and 2, the effect of the F test of (2·3) on the performance of $\hat{C}$ is negligible.

Table 2. *Empirical frequencies of rejection the null hypothesis of linearity*; $n = 70, 140, 204$; $M = 4$, *and* 350 *replications. Nominal significance level,* 0·05

	$n=70$			$n=140$			$n=204$		
Model	K	F	C	K	F	C	K	F	C
				(a) Linear					
Model 1	0·037	0·069	0·074 (0·091)	0·049	0·054	0·054 (0·066)	0·066	0·066	0·046 (0·069)
				(b) Nonlinear					
Model 2	0·100	0·166	0·326 (0·320)	0·103	0·166	0·663 (0·751)	0·097	0·217	0·820 (0·911)

K, Keenan's test; F, F test; C, concurrent C test. Values in brackets use $\hat{\varepsilon}_t$.

5. Autoregressive-moving average models

In practice autoregressive-moving average models are often used in time series analysis with parameters estimated by the maximum likelihood method. In this case the proposed F test, or the concurrent C test, can be employed as a diagnostic tool for checking the linearity assumption of the process. We summarize the result as follows. Proof uses the same techniques as for Theorem 1, together with consistency properties of the maximum likelihood estimates.

Theorem 3. *Suppose that Y_t is a stationary and invertible* ARMA (p, q) *process, say,*

$$(Y_t - \mu) = \sum_{i=1}^{p} \Phi_i (Y_{t-i} - \mu) + e_t - \sum_{i=1}^{q} \Theta_i e_{t-i}, \tag{5·1}$$

where $\{e_t\}$ is a sequence of independent and identically distributed $N(0, \sigma_e^2)$ random variables. Let $M = p + q$ and $Z_t = \mathrm{vech}\,\{U_t^{\mathrm{T}} U_t\}$, where $U_t = (Y_{t-1}, \ldots, Y_{t-p}, \hat{e}_{t-1}, \ldots, \hat{e}_{t-q})$ with $\hat{e}_{t-i}$'s being the residuals for model (5·1) *fitted by maximum likelihood. Then, for large n, $\hat{F}$ defined similarly to* (2·3) *follows approximately a F distribution with degrees of freedom $\frac{1}{2}M(M+1)$ and $n - \frac{1}{2}M(M+3) - 1$.*

Acknowledgements

The author wishes to thank the Editor for comments that improved the presentation of the paper, and the ALCOA Research Foundation for partial financial support.

References

Bickel, P. J. & Doksum, K. A. (1977). *Mathematical Statistics: Basic Ideas and Selected Topics.* San Francisco: Holden-Day.

BILLINGSLEY, P. (1961). The Lindeberg-Lévy theorem for martingales. *Proc. Am. Math. Soc.* **12**, 788-92.
GRANGER, C. W. J. & ANDERSEN, A. P. (1978). *Introduction to Bilinear Time Series Models.* Gottingen: Vandenhoeck and Ruprecht.
HINICH, M. J. (1982). Testing for Gaussianity and linearity of a stationary time series. *J. Time Series Anal.* **3**, 169-76.
HINICH, M. J. & PATTERSON, D. M. (1985). Evidence of nonlinearity in daily stock returns. *J. Bus. Econ. Statist.* **3**, 69-77.
KEENAN, D. M. (1985). A Tukey nonadditivity-type test for time series nonlinearity. *Biometrika* **72**, 39-44.
LAI, T. L. & WEI, C. Z. (1983). Asymptotic properties of general autoregressive models and strong consistency of least squares estimates of their parameters. *J. Mult. Anal.* **13**, 1-23.
MARAVELL, A. (1983). An application of nonlinear time series forecasting. *J. Bus. Econ. Statist.* **1**, 66-74.
SUBBA RAO, T. & GABR, M. M. (1980). A test for linearity of stationary time series. *J. Time Series Anal.* **1**, 145-58.
TUKEY, J. W. (1949). One degree of freedom for non-additivity. *Biometrics* **5**, 232-42.

[*Received May* 1985. *Revised October* 1985]