Borrower: JHW

Lending String: JHE,AAA,*ALM,FDA,FTU

Patron: veeramani, Balaji

Journal Title: Journal of nonparametric statistics.

Volume: 6 **Issue:**
Month/Year: 1996 **Pages:** 205-221

Article Author:

Article Title: M. J. Hinich; Testing for Dependence in the Input to a Linear Time Series Model

ILLiad TN: 101615.0

ILL Number: 1412092

Call #: QA278.8 .J68 - Number of Items: 14 -

Location: Business Library Periodicals

ARIEL
Charge
Maxcost: 25.00IFM

Shipping Address:
Johns Hopkins University
Welch Medical Library, ILL
1900 E. Monument St.
Baltimore, MD 21205

Fax: 410-955-8020
Ariel: 162.129.60.21

Sorour, A. and Tong, H. (1993). A note on tests for threshold-type non-linearity in open loop systems. *Appl. Statist.*, **42**, 95–104.

Tiao, G. C. and Box, G. E. P. (1981). Modeling multiple time series with applications, *J. Amer. Statist. Asso.*, **76**, 802–816.

Tjøstheim, D. (1990). Nonlinear time series and Markov chains," *Adv. Appl. Prob.*, **22**, 587–611.

Tjøstheim, D. and Auestad, B (1994). Nonparametric identification of non-linear time series: selecting significant lags. *J. Amer. Statist. Asso.*, **89**, 1410–1419.

Tong, H. (1990). *Nonlinear time series: a dynamical system approach*. London: Oxford University Press.

Tsay, R. S. (1985). Model identification in dynamic regression (distributed lag) models, *J. Business Economic Statist.*, **3**, 228–237.

Tsay, R. S. (1986). Nonlinear tests for time series, *Biometrika*, **73**, 461–466.

Nonparametric Statistics, Vol. 6, pp. 205–221
Reprints available directly from the publisher
Photocopying permitted by license only

TESTING FOR DEPENDENCE IN THE INPUT TO A LINEAR TIME SERIES MODEL

MELVIN J. HINICH

University of Texas at Austin

This paper presents a simple test for dependence in the residuals of a linear parametric time series model fitted to non gaussian data. The test statistic is a third order extension of the standard correlation test for whiteness, but the number of lags used in this test is a function of the sample size. The power of this test goes to one as the sample size goes to infinity for any alternative which has non zero bicovariances $c_{e3}(r, s) = E[e(t)e(t+r)e(t+s)]$ for a zero mean stationary random time series. The asymptotic properties of the test statistic are rigorously determined. This test is important for the validation of the sampling properties of the parameter estimates for standard finite parameter linear models when the unobserved input (innovations) process is white but not gaussian. The sizes and power derived from the asymptotic results are checked using artificial data for a number of sample sizes. Theoretical and simulation results presented in this paper support the proposition that the test will detect third order dependencies in the input.

KEY WORDS: non Gaussian data, bicovariances, asymptotic properties, consistent test.

INTRODUCTION

This paper presents a test for dependence in the innovations process of a finite parameter time series model fitted to non gaussian data. The test statistic is based on a normalized sum of squared sample third order bicorrelations of the residuals. It can be considered as an extension of the portmanteau χ^2 test for independence of the residuals (Box and Pierce, 1970). One important difference between this test and others is that the number of lags is a function of the sample size. The user does not have to specify the dependence structure of the alternative to independence. The test is consistent as the sample size goes to infinity for any alternative which has third order dependence. The asymptotic properties of the test statistic are derived for a wide class of distributions for innovations which are independent, identically distributed random variables with zero mean (called a stationary pure noise process).

This test is of importance for time series analysis in two related ways. The first deals with detection of nonlinear structure in the observed data, which is the motivation of the bispectrum based test in Hinich (1982) and its extension to residuals (Ashley, Patterson and Hinich, 1985). But the test is also useful for the validation of the sampling properties of the parameter estimates of parameters of standard linear time series models when the unobserved input (innovations) process is white but not gaussian. The gaussian assumption is often made for statistical convenience, rather than out of conviction.

Although whiteness of the input process is sufficient to identify the structure of a finite parameter linear time series model from the sample covariances without assuming that the time series is gaussian, the finite sample distribution of the parameter estimates is unknown if the input is not gaussian. If the input is not gaussian, one usually assumes the input variables are independent and identically distributed in order to establish that ordinary least square (OLS) estimates are asymptotically gaussian (Mann and Wald, 1943; Hannan, p.329, 1970).

Least squares estimates are strongly consistent under the weaker assumption that the input is a *martingale difference* process (Lai and Wei, 1983). But if the input has higher order serial dependence, which includes martingle differences, then the form of the asymptotic distribution of ordinary least squares estimates of the parameters of a stable finite AR model is an open question. The asymptotic distribution will depend on the dependence structure of the process. More importantly, the finite sample distribution of these estimates can be very different from the gaussian limit since least squares estimates are ratios of weighted averages of the data.

A simple and reliable method for detecting dependence in the unobserved input in a linear time series model is an important component of a procedure for determining the statistical reliability of a linear model estimated data using a standard method. The U-statistic type of nonparametric test of Brock, Dechert, and Scheinkman (1987), called BDS, is not aimed at any specific type of alternative model or cumulant structure. The asymptotic normality of the BDS has been established but as with most asymptotic results, the large sample variance derived from the asymptotic may not yield a useful approximation to the variance of the test statistic for a given sample size. Mizrach (1991) uses a simpler form of a U-statistic to test the hypothesis that a time series is pth order Markov.

There are other approaches to testing for dependence. Chan and Tran (1992) present a bootstrap test based on estimates of the bivariate density at two time points and the marginal density function. Skaug and Tjøstheim (1993) use a test based on kernel estimates of the bivariate and marginal densities. Robinson (1991) presents a nonparametric entropy test for the bivariate case. These tests should have higher power than mine when the distribution of the innovations is symmetric since my test has zero power if the innovations have a symmetric joint density.

The paper begins with a brief review of linear time series models. The importance of the independence assumption is discussed in Section 2. Section 3 presents the test statistic for the residuals of a fitted linear model. Section 4 presents estimates of the power of the test using residuals from an ordinary least squares (OLS) fit of an AR(2) model using pseudo-random white noise innovations which have third order dependency.

1. LINEAR TIME SERIES MODELS WITH SKEWED INPUT

Let $\{x(t_k)\}$ denote a time series sampled at equally spaced times $t_k = k\tau$ from a stationary process $\{x(t_k)\}$ which has been low passed filtered so that the spectrum of the filtered signal is zero for all frequencies greater than $f = 1/2\tau$ (Papoulis, 1962). This sampling method will avoid aliasing, a problem which is often ignored in time series analysis of economic data (Hinich and Patterson, 1989 and 1993).

Suppose that $\{x(t_k)\} = \mathrm{H}[\{e(t_k)\}]$ is the output of an operator which acts on an input time series denoted $\{e(t_k)\}$. At this level of generality, the value of $x(t_k)$ can depend on future values of the input such as $e(t_k + t)$ for some $t > 0$. The operator is called *causal* if $x(t_k)$ only depends on the $e(t_j)$ for $j \leqslant k$. The operator is *linear* if

$$\mathbf{H}[\{c_1 e_1(t_k)\} + \{c_2 e_2(t_k)\}] = c_1 \mathbf{H}[\{e_1(t_k)\}] + c_2 \mathbf{H}[\{e_2(t_k)\}] \tag{1.1}$$

for any two time series $\{e_1(t_k)\}$ and $\{e_2(t_k)\}$ and constants c_1 and c_2. Such an operator is called a linear filter in the engineering and science literature (Priestley, 1981, Sect. 4.12).

Any time invariant linear filter can be expressed as the following convolution of a sequence of coefficients here denoted $\{h(m)\}$, which is called the *impulse response* of the filter:

$$x(t_k) = \sum_{m=-\infty}^{\infty} h(m) e(t_{k-m}). \tag{1.2}$$

If the filter is casual, then $h(m) = 0$ for m < 0. The filter is called *stable* if $\{h(m)\}$ is absolutely summable. A causal filter is *invertible* if the roots of the polynomical $\sum_m h(m) z^m$ are outside the unit circle $|z| = 1$. If so, then there exists a causal and stable impulse response $\{b(n)\}$ which inverts the time series as follows: $e(t_k) = \sum_n b(n) x(t_{k-n})$ for each k.

No randomness has been addressed as yet. Consider a sample of the output $\{x(t_k)\}$ of a linear filter where the input values $e(t_k)$ are unobserved random variables with a stationary joint density. Assume for convenience that $E[e(t_k)] = 0$ and that one wants to estimate the impulse response in order to forecast future values of the $x(t_k)$.

The most popular approach to this problem is to assume that the filter is a finite parameter ARMA(p, q) model

$$\sum_{n=0}^{p} b(n) x(t_{k-n}) = \sum_{m=0}^{q} h(m) e(t_{k-m}). \tag{1.3}$$

Note that the MA part is a linear filter acting on the input e's while the AR part is a linear filter acting on the x's. Four basic assumptions are usually made for this model: 1) the AR(p) filter is stable, 2) the MA(q) filter is invertible, 3) the stationary input process, $\{e(t_k)\}$ is a gaussian time series, and 4) $\{e(t_k)\}$ is a zero mean white process, i.e., $E[e(t_k)e(t_k + r)] = 0$ for all nonzero r. If these assumptions hold, then the output is a stationary gaussian time series. Also note that since the input is both white and gaussian, then the $e(t_k)$ variates are mutually independent, which makes the model identified.

Any stable finite parameter ARMA model whose MA part is invertible can be represented by an infinite order one-sided autoregressive difference equation of the output. The parameters of such infinite order representation has a geometric descending pattern (Fuller, Sect. 2.4, 1976). Many investigators use finite AR models rather than ARMA models in order to use least squares to estimate the parameters. The following discussion applies to finite parameter AR models.

A stationary gaussian process $\{x(t_k)\}$ is completely characterized by its mean, which has been assumed to be zero, its variance σ^2 and its autocorrelations

(Rosenblatt, Chapter 1, 1985). If the process is not gaussian, then the bicorrelation $c_{x3}=(r,s)=E[x(t_k)x(t_k+r)x(t_k+s)]/\sigma^3$ may be non-zero for many r and s values.

If the input is white but is not gaussian and not pure noise, then there can exist dependence in the input which will cause the sampling distribution of OLS estimates of the AR filter parameters to be non-gaussian even for large samples.

2. DEPENDENCE MATTERS

The class of dependent processes are so vast that a discussion of testing for dependence should focus on a parametric model for dependence which can generate a rich class of time series which are dependent. The following nonlinear finite parameter model provides a simple way to generate nonlinear time series which have non-zero bicorrelations as well as significant skewness and kurtosis. Let $\{\varepsilon(t_k)\}$ be a zero mean pure noise process with unit variance:

$$e(t_k)=\sum_{m=0}^{p}h(m)\,\varepsilon(t-m)+\sum_{r=1}^{q}\sum_{s=0}^{q}a(r,s)\varepsilon(t_k-r)\varepsilon(t_k-r-s). \tag{2.1}$$

This model is a linear function of the $\varepsilon(t_k)$ and cross products $\varepsilon(t_k-r)\varepsilon(t_k-r-s)$. Nonlinear operations with memory (delayed responses) produce dependent processes.

The specific model used in this work to generate dependent variates is similar but not identical to one of the two quadratic models used in Hinich and Patterson (1992). The following model for a white process with dependence was used to generate the artificial data:

$$e(t_k)=\varepsilon(t_k)+(\beta/4)\sum_{m=2}^{17}\varepsilon(t_k-1)\varepsilon(t_k-m). \tag{2.2}$$

The non-zero bicorrelations of $\{e(t)\}$ are $c_{e3}(1,m)=(\beta/4)(1+\beta^2)^3]^{-1/2}$ for $m=2,\ldots,17$.

In order to demonstrate that dependence matters in estimating an AR model, the following AR(2) model was used to generate pseudo-random AR(2) variates whose innovations were generated using model (2.2):

$$x(t_k)=a(1)x(t_k-1)+a(2)x(t_k-2)+e(t_k)\quad\text{for } a(1)=\rho(8/3)^{1/2},$$
$$a(2)=-\rho^2,\rho<1. \tag{2.3}$$

The difference equation in (2.3) has a pair of complex roots on the circle of radius ρ, and it is stable is $\rho<1$.

The pure noise input variates $\varepsilon(t_k)$ in (2.2) were either 1) gaussian, 2) exponential, or 3) uniform (0, 1) pseudo-random generated using either 1) the IMSL program RNNOA (standard normal), 2) RNEXP (standard exponential) and 3) RNUN (uniform 0 – 1). A fixed seed values were passed to the random number generator for the first replication in each run. For each replication, the ε's were normalized by subtracting $E[\varepsilon(t_k)]$ from $\varepsilon(t_k)$ and dividing by its standard deviation.

For each of the three distributions of the ε's in (2.2), three values were used for β:$\beta=0$, $\beta=1$, $\beta=5$, and two values for N:$N=50$ and $N=200$. The runs for $\beta=0$ were designed to benchmark the distribution of the estimated $a(1)$ and $a(2)$ for pseudo-random pure noise series.

Each run had 1000 replications. For each replication, the AR(2) innovations $\{e(n)\}$ were standardized by subtracting the sample mean and dividing by the standard deviation of the sample. The innovations were used to generate $x(n)$ values using the AR(2) model (2.3) with $\rho=0.7$, which yields the coefficients $a(1)=1.143$ and $a(2)=-0.49$.

For each run of 1000 replications, OLS estimates and their standard errors were computed for $a(1)$ and $a(2)$ for each of 1000 replications. Tables 1a(for $N=50$) and 1b(for $N=200$) present the sample mean values over the 1000 replications of $a(1)$ and $a(2)$, the sample standard deviation of the estimates, the 25%– 50%– 75% quantiles of the estimates, and the calculated standard errors of the estimates. The standard errors of the estimates were calculated using the sum of squares of the residuals from each replication. Although not shown in the tables, the average standard deviation of the residuals were nearly one for all the runs. The worst deviation from one was 0.96 for the gaussian quadratic model with $N=50$ and $\beta=1$.

The average standard deviation of the estimates are larger than their average standard errors. The gaussian quadratic model gave the largest deviation from the average standard error. For example when $N=50$ and $\beta=1$, the average standard deviation of the estimate of $a(1)$ is 0.21 whereas the average of the calculated standard error of the estimate is 0.12. The average of the calculated standard error of the estimate of $a(2)$ is 0.17 for this case. These simulations results and others which I have run indicate that dependence in the input series may cause an investigator to have a false sense of confidence in the quality of the fit.

The estimates of $a(1)$ and $a(2)$ are only biased for the exponential quadratic model and $N=200$. For $N=50$, the small sample negative bias of OLS is unaffected by the dependence in the input.

3. A THIRD ORDER PORTMANTEAU TEST OF THE RESIDUALS

To specify the null hypothesis for our test, suppose that the data is a sample of size N from a time series $\{x(t_k)\}$ which obeys a stable AR linear model of order p whose unobserved input is a pure noise process $\{e(t_k)\}$ with finite moments up to order twelve. Assume that the roots of the z polynomial are well inside the unit circle. Then if the lag length of the AR is known, the model is correctly specified and thus OLS estimates of the parameters of the model will be very close to the true values if $N-p$ is large (Kreiss and Franke, 1992).

Given the above assumptions on the model and N, let $\{u(t_k)\}$ denote the residuals from the fitted model scaled by its sample standard deviation so that the sample variance of the u's is one. The difference between $u(t_k)$ and $e(t_k)$ is of order $O_p(N^{-1/2})$, and the difference between a given joint cumulant of the residuals and its equivalent joint cumulant for the innovations is $O(N^{-1/2})$. In the theorem which establishes the asymptotic distribution of the test statistic, this error is negligible.

Table 1a. Statistics for estimates of AR(2) parameters N = 50, Replications = 1000 White noise innovations from quadratic model (2.2) with 16 lags

Table 1b. Statistics for estimates of AR(2) parameters N = 200, Replications = 1000 White noise innovations from quadratic model (2.2) with 16 lags

G, E, or U β Amplitude	Mean est(a)	Stand. Dev.	Std. Error of Estimates	.25 Quant	.50 Quant	.75 Quant
	a(1) = 1.14					
G β=0	1.07	0.13	0.12	0.99	1.08	1.16
G β=1	1.03	0.24	0.12	0.86	1.06	1.22
G β=5	1.04	0.19	0.12	0.92	1.06	1.17
E β=0	1.06	0.12	0.12	0.98	1.07	1.14
E β=1	0.96	0.25	0.12	0.77	0.97	1.16
E β=5	1.02	0.18	0.12	0.90	1.02	1.15
U β=0	1.06	0.13	0.12	0.98	1.07	1.15
U β=1	1.07	0.14	0.12	0.97	1.07	1.16
U β=5	1.04	0.21	0.12	0.90	1.07	1.21
	a(2) = −0.49					
G β=0	−0.46	0.11	0.12	−0.54	−0.47	−0.39
G β=1	−0.46	0.18	0.12	−0.59	−0.48	−0.33
G β=5	−0.46	0.16	0.12	−0.57	−0.47	−0.36
E β=0	−0.45	0.12	0.12	−0.54	−0.46	−0.38
E β=1	−0.42	0.19	0.12	−0.56	−0.42	−0.28
E β=5	−0.45	0.16	0.12	−0.56	−0.45	−0.34
U β=0	−0.45	0.12	0.12	−0.54	−0.46	−0.37
U β=1	−0.46	0.12	0.12	−0.55	−0.47	−0.38
U β=5	−0.46	0.17	0.12	−0.59	−0.48	−0.35
	a(1) = 1.14					
G β=0	1.13	0.06	0.06	1.09	1.13	1.17
G β=1	1.11	0.10	0.06	1.03	1.12	1.21
G β=5	1.11	0.10	0.06	1.06	1.12	1.18
E β=0	1.12	0.06	0.06	1.08	1.13	1.16
E β=1	1.07	0.14	0.06	0.97	1.08	1.18
E β=5	1.10	0.10	0.06	1.04	1.10	1.16
U β=0	1.12	0.06	0.06	1.08	1.12	1.17
U β=1	1.12	0.07	0.06	1.08	1.12	1.17
U β=5	1.12	0.12	0.06	1.04	1.13	1.20
	a(2) = −0.49					
G β=0	−0.48	0.06	0.06	−0.52	−0.49	−0.44
G β=1	−0.48	0.10	0.06	−0.55	−0.49	−0.42
G β=5	−0.48	0.09	0.06	−0.55	−0.49	−0.43
E β=0	−0.48	0.06	0.06	−0.53	−0.49	−0.44
E β=1	−0.46	0.11	0.06	−0.54	−0.45	−0.38
E β=5	−0.47	0.09	0.06	−0.53	−0.47	−0.41
U β=0	−0.48	0.06	0.06	−0.53	−0.48	−0.44
U β=1	−0.48	0.06	0.06	−0.53	−0.48	−0.44
U β=5	−0.48	0.09	0.06	−0.55	−0.49	−0.42

We will now treat the residuals $u(t_k)$ as if they have the same joint cumulants as the $e(t_k)$ in order to establish the large N sampling results for our test statistics.

Since $E[u(t_1)u(t_2)u(t_3)]$ is invariant to permutations of (t_1, t_2, t_3), the triangle $\{0 < r \leqslant s\}$ in the first quadrant can be chosen as the *principal domain* for the

bicorrelations $c_{u3}(r,s) = E[u(t_k)u(t_k+r)u(t_k+s)]$. The input $\{e(t_k)\}$ is assumed to be pure noise under the null hypothesis, and thus treating the residuals as the input, it follows that $c_{u3}(r,s) = 0$ for **all** r and s except when $r = s = 0$. If the input has third order dependence then $c_{u3}(r,s) \neq 0$ for at least one pair of r and s values.

Recall that Lai and Wei (1983) have shown that OLS is strongly consistent if the input $\{e(t_k)\}$ is a martingale difference, that is if the conditional expectation of $e(t_k)$ given its past is zero. If the input is a martingale difference, then $c_{e3}(r,s) = 0$ for all r and s except when $r = s$. The test statistic is a function of the sample bicorrelations for $r < s$ in order to improve the rate of convergence of our test statistic's distribution to normality, at the expense of power to discriminate between a martingale difference input versus pure noise (the null hypothesis). Ignoring sample bicorrelations for $r = s$ greatly simplifies the analysis of the sampling properties of the test statistic. The Hinich-Patterson (1992) bispectum based test for martingale differences can be used to detect such alternatives.

In order to take advantage of this zero bicorrelation property of pure noise, I propose a third order portmanteau statistic computed from the estimated values of the bicorrelation of the residuals (treated as the true input for the asymptotic analysis).

The following statistic is the r,s sample bicorrelation multiplied by $(N-s)^{1/2}$ to standardize its variance:

$$G(r,s) = (N-s)^{-1/2} \sum_{k=1}^{N-s} u(t_k)u(t_k+r)u(t_k+s) \tag{3.1}$$

Then $E[G(r,s)] = 0$ and $E[G^2(r,s)] = 1$ under the null hypothesis. The following theorem is proven in the Appendix.

THEOREM 1. Set $L = N^c$ where $0 < c < 0.5$. Our test statistic H_N is the following normalized sum of the $(L-1)L/2$ values $G^2(r,s)$ for $1 < r < s = L$:

$$H_N = L^{-1} \sum_{s=2}^{L} \sum_{r=1}^{s-1} [G^2(r,s) - 1].$$

H_N is asymptotically $N(0,1)$ as $N \to \infty$. For large N, $\mathrm{Var}(H_N) = 1 + O(v(\kappa)N^{2C-1})$ where $v(\kappa) = \kappa^2 + 6\kappa + 6$. See the Appendix for a proof.

The value of a test has to be judged by the types of alternatives which it has power against. The orientation for this test is to detect nonlinear structure in the innovations if it has a number of nonzero bicorrelations. This test should complement the portmanteau test of McLeod and Li (1983) which is based on a normalized sum of squared autocorrelations of squared residuals, and thus uses a subset of the sample normalized fourth order cumulants.

4. THE SIZE AND POWER OF THE TEST

To provide some insight into the power of test for large samples, assume that the test is applied to data which is white. The null hypothesis is that the data are realizations of a pure noise process. The alternative hypothesis is that the process has M nonzero bicorrelations in the set $0<r<s\leqslant L$. Thus there is third-order dependence in the process.

To simplify the discussion, suppose that the process has finite dependence, which is a very strong form of mixing. Applying the central limit theorem for dependent variates to (3.1) to get a large sample error, $G(r,s)=N^{1/2}(c_{u3}(r,s)+O_p(N^{-1/2})$. Thus $G^2(r,s)=N(c_{u3}^2+O_p(N^{-1/2})$, and consequently

$$H_N=(N/L)B^2[1+O_p(1/LN^{1/2})] \quad \text{where} \quad B^2=\sum_{s=2}^{L}\sum_{r=1}^{s-1}c^2(r,s). \tag{4.1}$$

The power of the test will be near one if $N^{1-c}B^2$ is large, and clearly the test is consistent as $N\to\infty$ for any alternative with at least one nonzero bicorrelation.

Asymptotic theory does not give much insight into whether the asymptotic gaussian distribution for H_N applies for a given sample size. In order to provide such insight, I turned to artificial data analysis.

The size (Type I error) of the test was estimated for residuals to an OLS fit of an AR(2) model using sample sizes for $N=50$ and $N=200$, and the three types of distributions for pure noise $\varepsilon(n)$'s used in the simulations of the OLS estimates presented in Section 2. Six thousand replications were made for each set of parameters. These pseudo-random variates where generated by the same routine as was used to generate the $\varepsilon(n)$ variates in model (2.2) as discussed in Section 2 using $c=0.4$, and thus rounding to the nearest integer, the L's for these two sample sizes are $L=5$ (10 bicorrelations) and $L=8$ (28 bicorrelations). Table 2 shows the estimated sizes at the 1% and 5% nominal sizes for $N=50$ assuming that the test statistic was $N(0,1)$.

There is no logical reason why the large sample approximation is valid for small samples, yet the simulation results show that the sizes hold. It is not surprising that the sizes are above their nominal level for exponential pure noise and $N=200(L=8)$. The kurtosis of an exponential variate is $\kappa=6$. Recall that the major error term in Theorem 1 is $O(\nu(\kappa)N^{2c-1})$ where $\nu(\kappa)=\kappa^2+6\kappa+6$. For $c=0.4$, this term is $\nu(\kappa)N^{-1/5}$. For a gaussian pure noise process, the error term is $O(6N^{-1/5})$. To achieve the same magnitude for the theoritical error of exponentials as compared with gaussian variates, the sample size must be $13^5=371{,}293$ times the sample size for the gaussian model. This enormous sample size penalty can be reduced by using a smaller c, but then the test would have reduced power since the number of lags L is reduced. For the example used in the simulations, the two L's are smaller than the span of dependence.

A simple way to reduce the kurtosis of the residuals without seriously reducing the power of the test is to clip the residuals by trimming the observations in the tails of the empirical cdf. It is not necessary to do this for the examples shown since the inflation of the sizes for the exponential model is modest.

The power of the test is shown in Table 3 for the same set of parameters as are used in Table 2. The test has modest power for the gaussian model and $N=50$ and little power for the uniform. It is somewhat surprising that the test has any power for $N=50$ since $L=50$ and thus only four of the non-zero bicorrelation are included in the sum of squares. The sum of squared bicorrelations for $\beta=1$ is $B^2=4/16=(1/4)$. The power of the test increases to acceptable levels for $N=200$

Table 2. Estimated 5% and 1% sizes from 6000 replications 6000 replications of gaussian, exponential, uniform innovations

N–Sample Size *Model*		*5%* *%> 1.64*	*1%* *%> 2.33*
$N=50$			
	G	3.5%	1.3%
	E	4.6%	2.8%
	U	2.8%	0.9%
$N=200$			
	G	5.2%	1.9%
	E	8.7%	5.3%
	U	5.2%	1.6%

Table 3. Estimated power of the test on the residuals of an AR(2) fit Innovations generated using mode (2.2) for 1000 replications

N = 50	*5%* *%> 1.64*	*1%* *%> 2.33*	*N–200*	*5%* *%> 1.64*	*1%* *%> 2.33*
$\beta=1$			$\beta=1$		
G	20.0%	13.5%	G	72.8%	63.2%
E	18.8%	13.7%	E	69.5%	60.3%
U	2.4%	1.1%	U	5.4%	2.0%
$\beta=5$			$\beta=5$		
G	41.5%	33.5%	G	90.0%	85.6%
E	29.2%	23.4%	E	76.3%	71.4%
U	5.0%	2.1%	U	37.4%	22.8%

and the gaussian and exponential distributions for the ε's. In this case $B^2=7/16$ since there are seven nonzero bicorrelations in the sum of squares. The power for the uniform for $N=200$ is low, which may be due to subtle departures for statistical independence for the uniform generator which show up in statistics which are functions of the sample bicorrelations. The congruential uniform generators are a type of chaotic process and are not random. Other simulations I have run point in that direction but the question is still open why the uniform model is hard to detect using the statistic presented in this paper.

5. CONCLUSION

A simple test for third order dependence in a random time series is presented in this paper. The asymptotic sampling properties are presented and checked by simulations using artificial data. The test is applied to the residuals of a stable AR(2) model whose parameters are estimated by least squares. The innovations of the AR(2) model are not gaussian and have dependence. The power of the test depends on the sum of squared nonzero bicorrelations, the distribution of the innovations, and of course the sample size.

A Fortran 77 program to compute these statistics is available from the author upon request. Please e-mail requests to hinich@gov.utexas.edu. The Fortran program used in the simulations is also available for anyone who wishes to replicate the simulations.

Acknowledgments

This work was supported by the Office of Naval Research. I wish to thank the referee for providing me with many very helpful comments and for catching my many errors and slips. The reviewing process was one of the best I have ever experienced.

References

Ashley, R., Patterson, D. M., and Hinich, M. J. (1986). A diagnostic test for nonlinear serial dependence in time series fitting errors. *J. Time Series Anal.* 7(3), 165–178.

Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J. Amer. Statist. Assoc.* 65, 1509–1526.

Brillinger, D. (1975). *Time Series, Data Analysis and Theory.* New York: Holt, Rinehart and Winston.

Brock, W. A., Dechert, W. D., and Scheinkman, J. (1987). A test for independence based on the correlation dimension, University. of Wisconsin-Madison, W. P. 8702.

Chan, N. H. and Tran, L. T. (1992). Nonparametric tests for serial dependence. *J. Time Ser. Anal.* 13(1), 19–28.

Fuller, W. A. (1976). *Introduction to Statistical Time Series.* New York: Wiley.

Hannan, E. J. (1970). *Multiple Time Series.* New York: Wiley.

Hinich, M. J. (1982). Testing for gaussianity and linearity of a stationary time series. *J. Time Ser. Anal.* 3(3), 169–176.

Hinich, M. J. and Patterson, D. M. (1989). Evidence of linearity in the trade-by-trade stock market return generating process. Chapter 16, *Economics Complexity: Chaos, Sunspots, Bubbles,* and Nonlinearity, ed. by W. A. Barnett, J. Geweke, and K. Shell. New York: Cambridge University Press, 383–409.

Hinich, M. J. and Patterson, D. M. (1992). A new diagnostic test of model inadequacy which uses the martingale difference criterion. *J. Time Ser. Anal.* 13(3), 233–252.

Hinich, M. J. and Patterson, D. M. (1993). Intra-day nonlinear behavior of stock prices. ed. by R. Day, and P. Chen, *Nonlinear Dynamics and Evolutionary Economics.* Oxford: Oxford University Press.

Kreiss, J.-P. and Franke, J. (1992). Bootstrapping stationary autoregressive moving-average models. *J. Time Ser. Anal* 13(4), 297–318.

Lai, T. L. and Wei, C. Z. (1983). Asymptotic properties of general autoregressive models and strong consistency of least-squares estimates of their parameters. *J. Multiv.* 13, 1–23.

Leonov, V. P. and Shiryaev, A.N. (1959). On a method of calculation of semi-invariants. *Theor. Prob. Appl.* 4, 319–329.

Mann, H. B. and Wald, A. (1943a). On stochastic limit and order relationships. *Ann. Math. Statistic.* 14, 217–226.

Mann, H. B. and Wald. A. (1943b). On the statistical treatment of linear stochastic difference equations. *Econometrica* 11, 173–220.

McLeod, A. I. and Li, W. K. (1983). Diagnostic checking ARMA Time Series Models using squared-residual autocorrelation, *J. Time Ser. Anal.* 4(4), 269–273.

Mizrach, B. (1991). A simple nonparametric test for independence of order (p) Wharton School, University. of Pennsylvania.

Papoulis, A. (1962). *The Fourier Integral and Its Applications,* New York: McGraw Hill.

Priestley, M. B. (1981). *Spectral Analysis and Time Series,* New York: Academic Press.

Robinson, P. M. (1991). Consistent nonparametric entropy-based testing. *Rev. Econ. Studies* 58, 437–453.

Rosenblatt, M. (1985). *Stationary Sequences and Random Fields,* Boston: Birkhauser.

Skaug, H. J. and Tjøstheim, D. (1994). Nonparametric tests of serial independence. *Biometrika.*

APPENDIX

Let X denote a random variable whose density function has finite moments of all order, i.e., $\mu_n = E(X^n)$ exits for all integers n. The function $g_x(s) = E[\exp(sX)]$ is called the *moment generating function* (mgf) of X since the coefficient of $s^n/n!$ in the Taylor series expansion of $g_x(s)$ around $s = 0$ is $\mu_n = E(X^n)$. Although the generating function is subscripted by x, the notation in this appendix is simplied if the moments are only subscripted by the order of the moment. The moments (and cumulants defined below) will only be subscripted by their random variables when it is necessary to avoid ambiguity.

The natural log of $g_x(s)$ plays an important role in proofs of central limit theorems. The coefficient of $s^n/n!$ in the Taylor series expansion of $\log[g_x(s)]$ is called the nth cumulant of X, and is usually denoted κ_n. An equivalent definition is that κ_n is the nth derivative of $\log[g_x(s)]$ at $s = 0$. The Taylor series expansion is:

$$\log[g_x(s)] = \sum_{n=1}^{\nu} \kappa_n(s^n/n!) + o(|s|^\nu). \tag{1}$$

From the first two derivatives of $\log[g_x(s)]$ evaluated at $s = 0$, it can be shown that $\kappa_1 = \mu_1 = E(X)$ and $\kappa_2 = \mu_2 - \mu_1{}^2$.

Note that the nth order cumulant of $X + c$ is also κ_n for $n > 1$. We then can simplify the exposition of cumulants if we set $E(X) = 0$ from now on. Then $\kappa_2 = \mu^2 = \sigma^2$ (the variance of X), $\kappa_3 = \mu_3$, and $\kappa_4 = \mu_4 - 3\sigma^4$. Thus the variance of X^2 is

$$\mathrm{Var}(X^2) = E(X^4) - \sigma^4 = \kappa_4 + 2\sigma^4. \tag{2}$$

Now consider a vector of l zero mean random variables $\mathbf{X} = (X_1, \ldots, X_l)$. The generating function of $\mathbf{X}$ is $g_x(s) = E[\exp(\mathbf{s}'\mathbf{X})]$ where $\mathbf{s} = (s_1, \ldots, s_l)'$. For any set of non-negative integers $n_1, \ldots, n_l$, let $\kappa(x_1^{n_1} x_2^{n_2} \ldots x_l^{n_l})$ denote the $(n_1, \ldots, n_l)$ *joint cumulant* of $(X_1^{n_1}, X_2^{n_2}, \ldots, X_l^{n_l})$. This cumulant is the coefficient of $(s_1^{n_1} \ldots s_l^{n_l})/(n_1! \ldots n_l!)$ in the Taylor series expansion of $\log[g_x(s)]$. This definition implies that for $n = n_1 + \ldots + n_l$,

$$\kappa(x_1^{n_1} x_2^{n_2} \ldots x_l^{n_l}) = \frac{\partial^n \log[g_x(\mathbf{s})]}{(\partial s_1)^{n_1}(\partial s_2)^{n_2} \ldots (\partial s_l)^{n_l}} \tag{3}$$

where the joint partial derivative is evaluated $\mathbf{s} = \mathbf{0}$.

Note that if $n_l = 0$, $s_l^{n_l} = 1$ and thus in our notation, $\kappa(x_1^{n_1} \ldots x_{l-1}^{n_{l-1}})$ is the $(n_1, \ldots, n_{l-1})$ joint cumulant of $(X_1^{n_1}, \ldots, X_{l-1}{}^{n_{l-1}})$. This implies that the nth cumulant of X_1 is $\kappa_x(x_1{}^n)$ in our joint cumulant notation. If the X_k's are identically distributed, then $\kappa_x(x^n) = \kappa_n$.

To simplify notation, we will now let $\kappa(x_1 \ldots x_k)$ denote the $(1, \ldots, 1)$ cumulant of the subset $(X_1, \ldots, X_k)$ of $\mathbf{X}$. When all the exponents of X's are equal to one, the joint cumulant of the X's is called *simple*.

It is helpful now to consider some examples of joint cumulants for zero mean variates. The joint cumulant $\kappa(x_1, x_2) = E(X_1 X_2)$ is defined to be the *covariance* of (X_1, X_2), or in notational shorthand, $\mathrm{Cov}(X_1, X_2)$. The simple joint cumulant of (X_1, X_2, X_3) is $\kappa(x_1 x_2 x_3) = E(X_1 X_2 X_3)$.

In general, $E(X_1^{n_1} X_2^{n_2} \dots X_l^{n_l})$ is the sum over all partitions of $n_1, n_2, \dots, n_l$ of the products of the cumulants of each set in the partition (see Brillinger, Sect. 2.3, 1975). For example the simple 4th order joint moment has the following expansion:

$$E(X_1 X_2 X_3 X_4) = \kappa(x_1 x_2 x_3 x_4) + \kappa(x_1 x_2)\kappa(x_3 x_4) + \kappa(x_1 x_3)\kappa(x_2 x_4) + \kappa(x_1 x_4)\kappa(x_2 x_3). \quad (4)$$

which implies that the simple joint cumulant of (X_1, X_2, X_3, X_4) is

$$\kappa(x_1 \dots x_4) = E(X_1 X_2 X_3 X_4) - E(X_1 X_2)E(X_3, X_4) - E(X_1 X_3)E(X_2 X_4) - E(X_1 X_4)E(X_2 X_3). \quad (5)$$

Another implication of (4) is as follows. Let $Y_1 = X_1 X_2$ and $Y_2 = X_3 X_4$. Since $E(Y_1) = \kappa(x_1 x_2) = \mathrm{Cov}(X_1, X_2)$ and $E(Y_2) = \kappa(x_3 x_4) = \mathrm{Cov}(X_3, X_4)$, it follows from (4) that

$$\mathrm{Cov}(Y_1, Y_2) = \kappa(x_1 x_2 x_3 x_4) + \kappa(x_1 x_3)\kappa(x_2 x_4) + \kappa(x_1 x_4)\kappa(x_2 x_3). \quad (6)$$

This implies that $\mathrm{Cov}(X_1^2, X_2^2) = \kappa(x_1^2, x_2^2) + 2\kappa(x_1 x_2)$.

The following six results for cumulants are modifications of the conditions presented on page 19 in Brillinger (1975). These results are obtained by a methodical application of the definition of the joint cumulant.

P1) The $n_1, \dots, n_l$ joint cumulant of the permutation $(X_2^{n_2}, X_1^{n_1}, X_3^{n_3}, \dots, X_l^{n_l})$ is $\kappa(x_2^{n_2} x_1^{n_1} x_3^{n_3} \dots x_l^{n_l})$. Thus the simple cumulant of any permutation of $(X_1, \dots, X_l)$ is $\kappa(x_1 \dots x_l)$.

P2) For any constants $a_1, \dots, a_l$, $\kappa(x_1 + a_1 \dots x_l + a_l) = \kappa(x_1 \dots x_l)$.

P3) Let $Y_k = a_k X_k$ for $k = 1, \dots, l$. Then $(a_1^{n_1} a_2^{n_2} \dots a_l^{n_l})\kappa(x_1^{n_1} x_2^{n_2} \dots x_l^{n_l})$ is the $(n_1, \dots, n_l)$ joint cumulant of $(Y_1, \dots, Y_l)$. Thus the simple cumulant of $(aY_1, \dots, aY_l)$ is $a^l \kappa(x_1 \dots x_l)$.

P4) The simple joint cumulant of $(X_1 + X_{l+1}, X_2 \dots, X_l)$ for $l+1$ variates is $\kappa(x_1 x_2 \dots x_l) + \kappa(x_{l+1} x_2 \dots x_l)$.

P5) If $(X_1, \dots, X_l)$ and $(Y_1, \dots, Y_l)$ are independent, then the simple cumulant of $(X_1 + Y_1, \dots, X_l + Y_l)$ is $\kappa(x_1 \dots x_l) + \kappa(y_1 \dots y_l)$.

P6) If any subset of X_k's are independent of the other X's, then $\kappa(x_1^{n_1} \cdots x_l^{n_l}) = 0$.

We now turn to a theorem which we use to prove the asymptotic normality of our test statistic.

THEOREM A. Let S denote the sum of the l variates $X_1 \dots, X_l$ i.e., $S = \mathbf{1}'\mathbf{X}$ where $\mathbf{1}' = (1, \dots 1)$. Then $\kappa(s^n)$, the nth cumulant of S, is

$$\sum_{k_1=1}^{l} \sum_{k_2=1}^{l} \cdots \sum_{k_n=1}^{l} \kappa(x_{k_1} \dots x_{k_n})$$

When $k_1 = k_2 = k$, then $\kappa(x_{k_1} x_{k_2} \dots) = \kappa(x_k^2 x_{k_3} \dots)$.

Proof. The mgf of S is $g_s(s) = E[\exp(sS)] = E[\exp(s\mathbf{1}'\mathbf{X})] = g_x(s, s, \dots, s)$. Let $h_s(s) = \log[g_s(s)]$ and $h_x(s_1, \dots, s_l) = \log[g_x(s)]$. Thus $\partial h_s / \partial_s = \partial h_x / \partial s_1 + \cdots + \partial h_x / \partial s_l$ by the chain rule. The second partial is $\partial^2 h_s / (\partial s)^2 = \sum_{k_1} \sum_{k_2} \partial^2 h_x / \partial s_{k_1} \partial s_{k_2}$. The result follows by continuing the chain rule for partial differentiation, setting $s = 0$, and appling (3). ■

Proof of Theorem 1 in the paper. Under the null hypothesis that the $u(t_k)$ are i.i.d. zero mean random variables where $\sigma_u = 1$. The *skewness* of $u(t_k)$ is denoted $\kappa_3 = \gamma$ and its *kurtosis* is denoted κ_4.

Redefine the three time points in the triple product of the u's for a given (r, s) as follows: $t_{k_1} = t_k, t_{k_2} = t_k + r, t_{k_3} = t_k + s$ $(k = 1, l)$. Since $r < s, t_{k_2} < t_{k_3}$. Let $Y(t_{k_1}, t_{k_2}, t_{k_3}) = u(t_{k_1})u(t_{k_2})u(t_{k_3})$ for $0 < r < s$. To apply these cumulant results to obtain the joint cumulants of $Y(t_{k_1}, t_{k_2}, t_{k_3})$, replace k by t_k and X_k by $u(t_k)$ in the above results. Then from P6, $\kappa[u(t_{k_1}), \dots, u(t_{k_l})] = 0$ unless $t_{k_1} = \cdots t_{k_l} = t$, when the l th order joint cumulant equals κ_l, the l th order cumulant of $u(t_k)$. Then a) $e[Y(t_{k_1}, t_{k_2}, t_{k_3}) = 0$ unless $t_{k_1} = t_{k_2} = t_{k_3}$, b) $E[Y(t_{k_1}, t_{k_2}, t_{k_3})]^2 = 1$, and c) $E[Y(t_{j_1}, t_{j_2}, t_{j_3}) Y(t_{k_1}, t_{k_2}, t_{k_3})] = 0$ if $j \neq k$.

The nth order cumulant of a product of variates can be related to the joint cumulants of the variates, but the relationship is more complicated than the one between moments and cumulants stated above. There is no simple approach to deal with the combinatorial relationships between the nth order joint cumulants of $Y(t_{k_1}, t_{k_2}, t_{k_3})$ for various values of t_k, r, and s and the cumulants of $u(t)$ even though the $u(t_k)$'s are independent. The relationships rest on a definition of *indecomposable partitions* of two dimensional tables of subscripts of the t's (see Leonov and Shiryaev, 1959 and Sect. 2.3 of Brillinger, 1975). We display the table of the t's next to the table of their subscripts which Brillinger uses in his exposition.

Consider the following $l \times 3$ tables of $t_{k_1} = t_k, t_{k_2} = t_k + r_k, t_{k_3} = t_k + s_k (k = 1, l)$:

Times			Using Delay Notation		
t_{11}	t_{12}	t_{13}	t_1	$t_1 + r_1$	$t_1 + s_1$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
t_{l_1}	t_{l_2}	t_{l_3}	t_l	$t_l + r_l$	$t_l + s_l$

(7)

Let $v = v_1 \cup \dots \cup v\mathbf{P}$ denote a partition of the t_{k_i} in this table into P sets, where $k = 1, \dots l$ and $i = 1, 2, 3$. There are many partitions of the $l \times 3$ times from the single set of all the elements to $l \times 3$ sets of one element.

Let $I(p)$ denote the number of elements is set v_p of the partition v. These $I(p)$ elements are denoted $(t_{k_{1(p)} i_{1(p)}}, \dots, t_{k_{I(p)} i_{I(p)}})$. The cumulant of $[Y_{k_{1(p)} i_{1(p)}}, \dots, Y_{k_{I(p)} i_{I(p)}}]$ is $\kappa[y(t_{k_{1(p)} i_{1(p)}}, \dots, y(t_{k_{I(p)} i_{I(p)}})]$. The symbol $\kappa[v_p]$ will be used for this joint cumulant.

A partition is *decomposable* if one set or a union of some sets in v equals a subset of the rows of table 7. A partition is called *indecomposable* if it is not decomposable.

Consider for example the following 2×3 table:

t_{11}	t_{12}	t_{13}	t_1	$t_1 + r_r$	$t_1 + s_1$
t_{21}	t_{22}	t_{23}	t_2	$t_2 + r_2$	$t_2 + s_2$

(8)

The decomposable partitions are: $(t_{11}, t_{12}, t_{13}) \cup (t_{21}, t_{22}, t_{23})$, which is the union of the two rows, and all its sub-partitions. Two indecomposable partitions of this table are $(t_{11}, t_{21}) \cup (t_{12}, t_{22}) \cup (t_{13}, t_{23})$ and $(t_{11}, t_{22}) \cup (t_{12}, t_{21}) \cup (t_{13}, t_{23})$, which are drawn in display 8.

The nth order joint cumulant of $[Y(t_{11}, t_{12}, t_{13}), \dots, Y(t_{l_1}, t_{l_2}, t_{l_3})]$ is the sum of products of cumulants taken over all indecomposable partitions of the table. For an indecomposable partition v the pth cumulant in the product is $\kappa[v_p] = \kappa[y(t_{k_{1(p)} i_{1(p)}}, \dots, y(t_{k_{I(p)} i_{I(p)}})]$. This is the cumulant of the Y's whose indices are in set v_p.

From now on this product will be called PIP(v) for *product of the cumulants of the sets in an indecomposable partition* v.

Recall that $\kappa[u(t_{k_1}),\ldots,u(t_{k_l})]=0$ unless $t_{k_1}=\cdots t_{k_l}=t$. It then follows by an enumeration of each of the cumulants of the sets in the indecomposable partitions $v=v_1\cup\cdots\cup v_p$ of table (8) that most of the products of cumulants are zero for a given partition.

A summary of the 2nd order joint cumulants for $0<r_1<s_1$ and $0<r_2<s_2$ is as follows: $\kappa[y(t_1,t_1+r_1,t_1+s_1)\ y(t_2,t_2+r_2,t_2+s_2)]=0$ unless $t_1=t_2=t, r_1=r_2=r$, and $s_1=s_2=s$. If so, then $\kappa[y^2(t,t+r,t+s)]=1$.

Recall that the statistic in the test sum is

$$G(r,s)=(N-s)^{-1/2}\sum_{t_k=1}^{N-s} Y(t_k,t_k+r,t_k+s),\quad 0<r<s.$$

By Theorem A the covariance of $G(r_1,s_1)$ and $G(r_2,s_2)$ is $[(N-s_1)(N-s_2)]^{-1/2}$ times a double sum of covariances of the Y's. There are $N-s$ non-zero terms (all equal to one) in the double sum of covariances. Then from Theorem A and P3, $\mathrm{Var}[G(r,s)]=(N-s)/(N-s)=1$ and $\mathrm{Cov}[G(r_1,s_2),G(r_2,s_2)]=0$.

To obtain the 3rd order joint cumulants, consider the following 3×3 table for $0<r_k<s_k$:

t_{11}	t_{12}	t_{13}		t_1	t_1+r_r	t_1+s_1
t_{21}	t_{22}	t_{23}		t_2	t_2+r_2	t_2+s_2
t_{31}	t_{32}	t_{33}		t_3	t_3+r_3	t_3+s_3

Using the delay notation for the indices, first consider the following indecomposable partition: $v_1=(t_1,t_2,t_3)\cup(t_1+r_1,t_2+r_2,t_3+r_3)\cup(t_1+s_1,t_2+s_2,t_3+s_3)$. If 1) $t_1=t_2=t_3$, 2) $r_1=r_2=r_3$, and 2) $s_1=s_2=s_3$, then the 3rd order cumulants of the three columns equal γ, and thus $\mathrm{PIP}(v_1)=\gamma^3$. If any one of the equalities in 1, 2, or 3 do not hold, $\mathrm{PIP}(v_1)=0$.

Now consider the indecomposable partition

$v_2=(t_1,t_2,t_3)\cup(t_1+r_1,t_2+r_1,t_3+r_3,t_1+s_1)\cup(t_2+s_2,t_3+s_3)$. $t_1=t_2=t_3=t$, $r_1=r_2=r_3=s_1$, and $s_2=s_3$, $\mathrm{PIP}(v_2)=\gamma\kappa\sigma_u^2=\gamma\kappa$. But $r_1<s_1$ and thus the PIP of such a partition is zero. Other such three set partition will have a zero PIP since $0<r<s$ for each row. It then follows that $\kappa[y^3(t,t+r_1,t+s_1)]=\gamma^3$

Suppose that $s_3\neq r_1$. Consider the indecomposable partition (a sub-partition of v_2) $v_3=(t_1,t_2,t_3)\cup(t_1+r_1,t_2+r_1)\cup(t_3+r_3,t_1+s_1)\cup(t_2+s_2,t_3+s_3)$. Then if $t_1=t_2=t_3=t$, and 4) $r_1=r_2, s_2=s_3, r_3=s_1$, then $\mathrm{PIP}(v_3)=\gamma$. Since $s_3\neq r_1$, this is the only partition of the table into one triple and three sets of two indices which has a non-zero PIP. If these equalities hold then all partitions into one triple and four pairs of indices have zero PIP's. Thus $\kappa[y(t,t+r_1,t+s_1)y(t,t+r_1,t+s_2)\ y(t,t+s_1,t+s_2)]=\gamma$. If one the inequalities in 4) are broken, then $\kappa\ [y(t,t+r_1,t+\ s_1)\ y(t,t+r_2,t+s_2)\ y(t,t+r_3,t+s_3)]=0$

Equality 1 constrains the three t's to lie on a line in the three dimensional space (t_1,t_2,t_3). This one dimensional constraint applies to all indecomposable partitions which have non-zero PIP's. Thus for each (r,s) in the three dimensional subset of the six dimensional space of $(r_1,s_1,r_2,s_2,r_3,s_3)$, the number of non-zero 3rd order joint cumulants of the Y's is at most of the order $O(N)$.

From Theorem A the 3rd order joint cumulant of the G's is $[(N-s_1)(N-s_2)(N-s_3)]^{-1/2}=O(N^{-3/2})$ times the three fold sum of 3rd order cumulants of the Y's. It then follows from above that the 3rd order joint cumulant $\kappa[g^2(r,s)]=O(\gamma^3N^{-1/2})$. The 3rd order joint cumulant $\kappa[g(r_1,s_1)g(r_1,s_2)g(r_1,s_2)]$ is $O(\gamma N^{-1/2})$ if $r_3=s_1\neq s_2$. If the r's or the s's are different, then $\kappa[g(r_1,s_1)g(r_2,s_2)g(r_3,s_2)]=0$.

We also require an understanding of the higher order joint cumulants to prove the asymptotic properties of our test statistic. The general form can be deduced from the 4th order case by enumerating the sets in the indecomposable partitions of the 4×3 table:

t_{11}	t_{12}	t_{13}		t_1	t_1+r_1	t_1+s_1
t_{21}	t_{22}	t_{23}		t_2	t_2+r_2	t_2+s_2
t_{31}	t_{32}	t_{33}		t_3	t_3+r_3	t_3+s_3
t_{41}	t_{42}	t_{43}		t_4	t_4+r_4	t_4+s_4

Consider the following indecomposable partition of this table into three sets of four indices: $v_4=(t_1,t_2,t_3,t_4)\cup(t_1+r_1,t_2+r_2,t_3+r_3,t_4+r_4)\cup(t_1+s_1,t_2+s_2,t_3+s_3,t_4+s_4)$. The three 4th order joint cumulants of the sets of four are not zero if the following equalities hold: 5) $t_4=t_3=t_2=t_1=t$, 6) $r_4=r_3=r_2=r_1=r$, 7) $s_4=s_3=s_2=s_1=s$. If all these equalities hold, $\mathrm{PIP}(v)=\kappa^3$.

Now consider the following indecomposable partition of v_4 into two pairs and two sets of four indices (col's 1 and 2).

$v_5=$

$$(t_1,t_2,t_3,t_4)\cup(t_1+r_1,t_2+r_2,t_3+r_3,t_4+r_4)\cup(t_1+s_1,t_3+s_3)\cup(t_2+s_2,t_4+s_4).$$

The two covariances for the pairs and the two 4th order cumulants of the sets of four are not zero if equalities 5 and 6 hold, and 8) $s_3=s_1, s_4=s_2$. If so then $\mathrm{PIP}(v_5)=\kappa^2$.

Suppose that $s_1=s_2$. Then equality constraint 8 implies that 7 also holds. If so there are nine indecomposable partitions of the table which are similar to v_5 in the sense that the partition consists of two columns and two pairs of indices whose cumulants are not zero, and thus have non-zero PIP's .One example is

$$(t_1,t_2,t_3,t_4)\cup(t_1+r_1,t_2+r_2,t_3+r_3,t_4+r_4)\cup(t_1+s_1,t_3+s_3)\cup(t_2+s_2,t_4+s_4).$$

Each of these partitions can be partitioned into one column and four sets of pairs, for example

$$(t_1,t_2,t_3,t_4)\cup(t_1+r_1,t_2+r_2)\cup(t_3+r_3,t_4+r_4)\cup(t_1+s_1,t_3+s_3)\cup(t_2+s_2,t_4+s_4).$$

There are 27 such indecomposable partitions whose PIP is κ.

Finally each of the 27 can be partitioned into six sets of pairs, but not all will be indecomposable. There are 12 indecomposable partitions into six pairs whose covariances are not zero. one example is

$$(t_1,t_4)\cup(t_2,t_3)\cup(t_1+r_1,t_2+r_2)\cup(t_3+r_3,t_4+r_4)\cup(t_1+s_1,t_3+s_3)\cup(t_2+s_2,t_4+s_4).$$

Since the 4th order joint cumulant of the Y's is the sum of the PIP's for all indecomposable partitions, it follows from the above enumeration that

$$\kappa[y^4(t,t+r,t+s)y(t,t+r)] = \kappa^3 + 9\kappa^2 + 27\kappa + 12. \tag{9}$$

Recall partition v_5 for $s_1 \neq s_2$ and equalities 5, 6, and 8. An enumeration of the indecomposable partitions which have non-zero PIP's are as follows. There are three indecomposable partitions of this partition of the form $(t_1+s_1, t_3+s_3) \cup (t_2+s_2, t_4+s_4) \cup$ (col. 2) $\cup$ (two pairs from col. 1) which have non-zero PIPs. Similarly, there are there composed of col. 2 and four sets of pairs from col. 1 and three with non-zero PIP's. Finally there are six partitions with six pairs such that their PIP's are positive. Thus

$$\kappa[y^2(t,t+r,t+s_1)y^2(t,t+r,t+s_2)] = \kappa^2 + 6\kappa + 6. \tag{10}$$

A similar pattern of partitions holds for the partition made up of the first two columns with the pair $(t_1+s_1, t_4+s_4) \cup (t_2+s_2, t_3+s_3)$. The PIP for this partition is the same as for 5), and $s_4 = s_1$, $s_3 = s_2$. But then the rows are a permutation of v_5 for $s_1 \neq s_2$, and thus this partition is redundant.

The partition which yields the major error term for the theorem is the partition of the 2nd and 3rd columns and the pairs (t_1, t_4) and (t_2, t_3). The cumulants for these four sets will be non-zero if 9) $t_4 = t_1, t_3 = t_2 = t - d$, 10) $d = r_2 - r_1 = s_2 - s_1$, and 11) $r_4 = r_1, r_3 = r_2, s_4 = s_1, s_3 = s_2$. This partition has not been counted if $d \neq 0$ and thus $t_2 \neq t_1$. The enumeration of the partitions of subsets of this partition is the same as for v_5. Thus

$$\kappa[y^2(t,t+r_1,t+s_1)y^2(t,t+r_1+d,t+s_1+d)] = \kappa^2 + 6\kappa + 6. \tag{11}$$

Suppose that $r_2 - r_1 \neq s_2 - s_1$, and $r_4 = r_1$, $r_3 = r_2$, $s_4 = s_1$, $s_3 = s_2$ (equality 11 holds). This is the least restrictive set of equalities for the (r_k, s_k), and is a four dimensional subset of the eight dimensional space of the indices for the 4th order cumulants. Then the PIP $= 1$ for the following indecomposable partition of four pairs:

$(t_1, t_4) \cup (t_2, t_3) \cup (t_1+r_1, t_2+r_2) \cup (t_3+r_3, t_4+r_4) \cup (t_1+s_1, t_4+s_4) \cup (t_2+s_2, t_3+s_3)$ if $t_4 = t_1, t_3 = t_2 = t_1 - (r_2 - r_1)$. Thus $\kappa[y^2(t,t+r,t+s_1)y^2(t,t+r_1-r_2,t+s_1-s_2)] = 1$.

As in the 3rd order case, equalities 5 and 9 constrain the four t's to lie on a line in the four dimensional space (t_1, t_2, t_3, t_4). A one dimensional constraint applies to all indecomposable partitions which have non-zero PIP's. The equality constraint 11 on the four (r_k, s_k) determine the number of non-zero PIP's in the four fold sum of the 4th order joint cumulant of the Y's in the calculation of the 4th order joint cumulants of the G's. For each (r_1, s_1) and (r_2, s_2) there are $O(N)$ non-zero 4th order cumulants of the Y's. Thus from Theorem A for (r_k, s_k) in the four dimensional space defined by equality 11), the 4th order joint cumulant of the $G(r,s)$ is $[(N-s_1)(N-s_2)(N-s_3)(N-s_4)]^{-1/2}O(N) = O(N^{-1})$. The same result holds for non-zero 4th order cumulants for indices in four dimensional spaces defined by other partitions of four pairs.

The same pattern holds of partitions of the general $l \times 3$ table of subscripts into pairs with identical indices. The non-zero cumulants are those where the t's lie on a line, and the (r_k, s_k) lie in an l dimensional subset of the $2l$ dimensional space of the indices. The lth joint cumulant of the G's is of order $O(N^{1-l/2})$.

These results will be now applied to prove that the test statistic $H_N = L^{-1} \sum_m [G^2(r_m, s_m) - 1]$ is asymptotically normal. It has already been shown that $\mathrm{Var}[G(r_m, s_m)] = E(G^2(r_m, s_m)) = 1$ and thus $E(H_N) = 0$ under the null hypothesis of independence. To obtain the variances and covariances of the G^2s, recall the result implied by expression (6) for $X_1 = G(r_1, s_1)$ and $X_2 = G(r_2, s_2)$. Then

$$\mathrm{Var}[G^2(r_1,s_1)G^2(r_2,s_2)] = \kappa[g^2(r_1,s_1)g^2(r_2,s_2)] + 2,$$

$$\mathrm{Cov}[G^2(r_1,s_1)G^2(r_2,s_2)] = \kappa[g^2(r_1,s_1)g^2(r_2,s_2)].$$

Suppose that $r_4 = r_3 = r_2 = r_1 = r$ and $s_4 = s_3 = s_2 = s_1 = s$. Then from (9) Theorem A, $\kappa[g^2(r,s)] = O(\mu(\kappa)/N)$ where $\mu(\kappa) = \kappa^3 + 9\kappa^2 + 27\kappa + 12$. Thus $\mathrm{Var}[G^2(r,s)] = 2 + O(\mu(\kappa)/N)$.

If $d = r_2 - r_1 = s_2 - s_1$ and $r_4 = r_1, r_3 = r_2, s_4 = s_1, s_3 = s_2$ (equalities 9 and 10) then from (11) it follows that $\kappa[g^2(r_1,s_1), g^2(r_1+d), s_1+d)] = O(\nu(\kappa))$ where $\nu(\kappa) = \kappa^2 + 6\kappa + 6$. Thus $\mathrm{Cov}[G^2(r_1,s_1)G^2(r_2,s_2)] = O(\mu(\kappa)/N)$, and otherwise the covariance error is $O(1/N)$. Since the number of $G^2(r_m, s_m) - 1$ terms in the sum is $L^2/2 (L = N^c)$, $\mathrm{Var}(H_N) = 1 + O(L^2/N) \to 1$ as $N \to \infty$ since $0 < c < 1/2$.

There are approximately $L^2/2$ such $(r_1, s_1), (r_2, s_2), (r_3, s_3)$ and (r_4, s_4) in the double sum which satisfy equalities 9 and 10. Thus the error in the variance of H_N is $O(\nu(\kappa)N^{2c-1})$.

To complete the proof we will now demonstrate that the cumulants of H_N of order $l \geqslant 3$ go to zero as $N \to \infty$. The lth cumulant of H_N depends on the $2l$ order joint cumulant of the $G^2(r_k, s_k)$ for $k = 1, l$. From above these cumulants are of order $O(N^{1-2l/2}) = O(N^{1-l})$. The (r_k, s_k) lie in an $2l$ dimensional subset since they are paired. Thus the lth cumulant of H_N is of order $O(L^{-l}L^{2l}N^{1-l}) = O(N^{l(1-c)+1})$ which goes to zero as $N \to \infty$ since $l(1-c)+1 < 0$. ■