Pakistan Journal of Information and Technology 1(2): 208-212, 2002

A Periodogram-Based Test Method for Comparing Stationary Stochastic Signals

Zazli Chik
Pusat Asasi Sains, University of Malaya, 50603 Kuala Lumpur, Malaysia

Abstract: A problem in gas industry and commodity prices describes a number of periodogram-based tests of the hypothesis that two independent time series are realizations of the same stationary signal. This paper describes the use of a periodogram-based test method called the randomisation test method for comparing stationary stochastic signals. The paper consider the case of comparing two signals which can be generalised to compare more than two signals. A number of test statistics are considered and the maximum distance between periodograms is recommended. It is important to standardise the signals before calculating the periodograms otherwise the test is considerably weakened.

Key Words: Kolmogorov-Smirnov, Randomisation Test, Detection of Changes, Cumulative Periodograms

Introduction

A number of methods of comparing stationary signals have been proposed, see for example Basseville (1988, 1989), Coates and Diggle (1986), Souza and Thomson (1982), Diggle and Fisher (1991) and Stoica (1990) and the references therein. The problem can be briefly described as follows. Let $\{x_t : t = 1, \ldots, n\}$ and $\{y_t : t = 1, \ldots, n\}$ denote two independent observed sets of data generated by the stationary processes $\{X_t\}$ and $\{Y_t\}$ respectively. Then the null hypothesis of interest is that the sets of data were generated by the same stationary process. In general, it is assumed that $\{X_t\}$ and $\{Y_t\}$ are stationary general linear processes with independent identically distributed innovations. It is often futher assumed that the innovations are Gaussian.

One approach is to assume a model, such as an autoregressive model. Stoica (1990) suggested using the Euclidean distance between the estimated parameters obtained by fitting autoregressive models to the two sets of data. Basseville (1989) commented that the Euclidean distance between the cepstral coefficients is a better measure than the Euclidean distance between the estimated parameters. Coates (1991), commented on the problems associated with the proposed test statistic in particular that the test is not symmetric and there is a lack of fit to the χ^2 distribution.

Another approach is to use the periodogram. Coates and Diggle (1986) described a number of periodogram-based tests of the hypothesis that two independent time series were generated by the same stationary process. A test based on the range of the periodogram ratios was found to be extremely weak. A test based on the cumulative sums of transformed periodogram ratios, like the test suggested by Stoica (1990), is not symmetric but depends on the arbitrary labelling of the two series $\{x_t\}$ and $\{y_t\}$. They recommended a test based on the ratio of the log of the periodograms. Their alternative hypothesis was that the ratio of the log of the spectra could reasonably be a quadratic model. Diggle and Fisher (1991) also approached the problem of comparing stationary signals using periodograms. In particular they suggested using the cumulative periodogram, leading to an informative graphical procedure.

For testing whether an observed set of data were generated by a white noise process, Jenkins and Watts (1968) recommended the use of the Kolmogorov-Smirnov test statistic. If the observed set of data were generated by a white noise process, then under the null hypothesis, the observed periodogram ordinates are independent and identically distributed. However, for the general problem of comparing stationary signals, the periodogram ordinates are not identically distributed although they are (asymptotically) independent. Hence the standard sampling distribution for the Kolmogorov-Smirnov test statistic cannot be used. The Kolmogorov-Smirnov test statistic is, however, an obvious measure of the distance between two cumulative periodograms and an approximate sampling distribution for the test statistic can be generated using a randomization test method. It is also important to standardise the observed set of data before being used, otherwise the test is considerably weakened.

Description of the Randomization Test Method: The randomization test is a method for determining the significance of experimental results by permuting a set of data in order to obtained repeated values of a particular test statistic (for example the t statistic). A randomization test is not a statistical test in the usual sense but it is a way of generating an approximate sampling distribution and hence of determining significance. There are two basic methods of permuting data. One is called systematic data permutation where all possible data permutations in a set are used in generating an approximate sampling distribution and the other is called random data permutation. Random data permutation uses only a random sample of all possible data permutations and in practice random

data permutation is more useful. It serves the same function as systematic data permutation but there is a substantial reduction in the number of permutations from typically many millions to as few as 100.
The validity of a randomization test using random data permutation depends on using random assignment to create the data permutation. That is, the allocation of observations to each sample is carried out at random so that each observation has an equal chance of being in each sample. Random assignment is the only random element necessary for determining the significance of experimental results by the randomization test method. Assumptions regarding normality, homogeneity of variances and so on are unnecessary although they may influence the power of the test. Any statistical test is transformed into a distribution-free test when the significance is determined by the randomization test method. Further details on the randomization test and the random assignment can be found in Edgington (1980), Hooton (1991) and Nelson (1992).

As in the introduction, assume that $\{X_t\}$ and $\{Y_t\}$ are stationary general linear processes with independent, identically distributed (not necessarily Gaussian innovations), the periodogram ordinates, $I_x(\omega_i)$ and $I_x(\omega_j)$ are asymptotically independent for i ≠ j, as are $I_y(\omega_i)$ and $I_y(\omega_j)$ (see Priestley, 1981). Also, each $I_x(\omega_i)$ is independent of all $I_y(\omega_j)$ and vice versa since the two series $\{x_t\}$ and $\{y_t\}$ are assumed to be independent. Also under the null hypothesis that the sets of data were generated by the same stationary process, $I_y(\omega_j)$ and $I_y(\omega_j)$ are identically distributed for each frequency $\omega_j = 2\pi j/n$, $j=1,\ldots,m$ and $m=[(n-1)/2]$.
Let d be some measure of the distance between the cumulative periodograms, $F_x(\omega\ )$ and $F_y(\omega\ )$. Then, under H_0, the distribution of d will be invariant under all 2^m possible interchanges of $I_x(\omega_j)$ and $I_y(\omega_j)$. The Kolmogorov-Smirnov test statistic, D_m is an obvious measure of the distance between the cumulative periodograms and is defined by

$$D_m = \sup_j \left| F_x(\omega_j) - F_y(\omega_j) \right|$$

where

$$F_x(\omega_j) = \sum_{i=1}^{j} I_x(\omega_i) \Big/ \sum_{i=1}^{m} I_x(\omega_i)$$

and similarly for $F_y(\omega_j)$.

The estimation of the power of the Kolmogorov-Smirnov test statistic, D_m when comparing two stationary signals, is of interest. This can be obtained by simulation.
Finally, the algorithm for the randomization test method for comparing two observed stationary signals is as follows:

Steps

- Obtain two independent observed sets of data $\{x_t$: t=1, ... , $n\}$ and $\{y_t$: t=1,..., $n\}$. Calculate the periodogram for each observed sets of data and call the periodogram ordinates for $\{x_t\}$ and $\{y_t\}$, $I_x(\omega_j)$ and $I_y(\omega_j)$, {j=1, ... ,m} respectively.
- Calculate the Kolmogorov-Smirnov test statistic, D_{m_0} for these original samples.
- Randomly permute these original samples. Calculate the Kolmogorov-Smirnov test statistics, D_{m_i} for these randomly permuted samples. Do this N-1 times.
- Combine D_{m_0} with the N-1 values of D_{m_i} to give a sample size of N, comprising $D_{m_0}, D_{m_1}, \ldots, D_{m_{N-1}}$. Count the number of test statistics that are equal to or larger than D_{m_0} (include D_{m_0} itself). If this number is S, then the significant level reached by the original sample is equal to S/N.
- To estimate the power of the Kolmogorov-Smirnov test statistic for particular stationary processes, repeat steps 1-4 a large number of times (for example 100 times). The power of the Kolmogorov-Smirnov test statistic at a given significance level is then the proportion of times a significant level is obtained.

The randomisation test method can be generalised to compare more than two signals using test statistics below

$$U = \sum_{q,r} \sup_j \left| F_q(\omega_j) - F_r(\omega_j) \right|$$

$$V = \sup_{q,r,j} \left| F_q(\omega_j) - F_r(\omega_j) \right|$$

$$T = \sup_j \sum_i (F_i(\omega_j) - \overline{F}(\omega_j))^2$$

where $\overline{F}(\omega_j) = (\sum_i F_i(\omega_j))/k$ is the mean of

the cumulative periodograms of the pooled k observed stationary signals at the jth frequency (see Kiefer, 1959).
The test statistic V is the obvious analogue of the Kolmogorov-Smirnov test statistic since it calculates the largest difference between any of the cumulative periodograms. In comparison, the test statistic U sums the Kolmogorov-Smirnov test statistic over all pairs of cumulative periodograms. By contrast, the test statistic T calculates the largest of the sum of squared differencs between a single cumulative periodogram and the average of all of the cumulative periodograms.

Results and Discussion

Simulation: The power of the proposed randomization test method based on the test statistic, D_{m} has been estimated by applying the method to pairs of simulated autoregressive (AR(p)) processes. The processes used were all of the form

$$X_t=\sum_{j=1}^{p}\alpha_j X_{t-j}+Z_t$$

where in each case $\{Z_t\}$ is a sequence of mutually independent $N(0,\sigma^2)$ random variables i.e. Gaussian white noise. The results were compared to those reported by Coates and Diggle (1986), Diggle and Fisher (1991).
Only 100 replicates of each pair of processes were performed for each simulation since good coverage of a range of cases is more important than precise estimation of the power for any particular case. For each replicate of each simulation, the randomization test method was applied using random data permutation. In each case, a sample size of n =64 was used and the number of random data permutations was N-1=99.
The following pairs of stationary processes were used.

(a). white noise *versus* AR(1), $\alpha_1>0$;

(b). white noise *versus* AR(2), $\alpha_1=0, \alpha_2>0$;

(c). white noise *versus* AR(3), $\alpha_2=0$, $\alpha_1=-\alpha_3>0$;

(d). AR(1), $\alpha_1=0.5$, *versus* AR(1) , $\alpha_1>0$.

A Fortran program was constructed for this simulation study. Pseudorandom numbers used were generated using the NAG (1984) random number generator. The periodograms ordinates were calculated using the NAG(1984) subroutine.
Three cases were studied. Firstly, pairs of processes were simulated with values of σ^2 adjusted so that the processes $\{X_t\}$ and $\{Y_t\}$ had the same variances. The results are shown in Table 1 and they are similar to the results obtained by Diggle and Fisher (see Table 2, 1991). The major drawback of this method is that it is only possible when the structure of the data is known, as in the case of a simulation study. In such a case, the theoretical variances of the processes can be calculated and the process variances standardised.

Table 1: Estimated Power of D_{m}-tests, Series Length n =64, with Different Noise Variances so that the Process Variances are the Same

α_1 or α_2	Results for the following sizes 0.10	0.05	0.01
white noise *versus* AR(1), $\alpha_1>0$			
0.2	.30	.19	.04
0.4	.64	.50	.18
0.6	.87	.78	.45
0.8	.99	.94	.80
white noise *versus* AR(2), $\alpha_1=0, \alpha_2>0$			
0.2	.22	.11	.02
0.4	.32	.20	.05
0.6	.46	.30	.13
0.8	.53	.43	.22
white noise *versus* AR(3), $\alpha_2=0$, $\alpha_1=-\alpha_3>0$			
0.2	.30	.14	.03
0.4	.67	.56	.26
0.6	.93	.87	.52
AR(1), $\alpha_1=0.5$, *versus* AR(1) , $\alpha_1>0$			
0.1	.64	.51	.26
0.3	.30	.20	.07
0.5	.10	.06	.01
0.7	.35	.26	.09
0.9	.86	.81	.63

A practical and easy alternative to standardising the variances of the processes $\{X_t\}$ and $\{Y_t\}$ is to calculate the variances for the observed signals and divide the original signals by their empirical process standard deviations. The results are shown in Table 2 and are similar to those in Table 1.
When comparing stationary processes, the processes should be standardised. If the processes are not standardised and the error variances for both of the processes are the same, say both are unity, then there is a substantial decrease in power as the data becomes highly correlated. The results are shown in Table 3. For example, the estimated power for white noise versus AR(1) , $\alpha_1=0.8$ for the three sizes should be more than for white noise versus AR(1) , $\alpha_1=0.6$.
Results in Table 2 can be compared to the results in Table 2 of Coates and Diggle (1986) where they have used the semiparametric approach which they recommend. The results from using the randomization test method show more power for almost all of the pairs of processes except for (b). In this case, the log spectral ratio is quadratic in form and the semiparametric approach would be expected to be more powerful as it assumes a quadratic form for the log spectral ratio. In general, the randomization test method is a better approach than the semiparametric

Table 2: Estimated Power of D_m-tests, Series Length n =64, Standardised so that the Observed Series have the Same Variances

α_1 or α_2	Results for the following sizes 0.10	0.05	0.01
white noise *versus* AR(1), $\alpha_1 >0$			
0.0	.15	.11	.02
0.2	.32	.24	.05
0.4	.67	.56	.21
0.6	.88	.80	.50
0.8	.98	.98	.83
white noise *versus* AR(2), $\alpha_1 =0$, $\alpha_2 >0$			
0.2	.24	.12	.02
0.4	.36	.22	.03
0.6	.46	.35	.11
0.8	.56	.42	.25
white noise *versus* AR(3), $\alpha_2 =0$, $\alpha_1 = - \alpha_3 >0$			
0.2	.31	.17	.03
0.4	.69	.59	.25
0.6	.94	.87	.50
AR(1), $\alpha_1 =0.5$, *versus* AR(1) , $\alpha_1 >0$			
0.1	.63	.55	.25
0.3	.38	.24	.11
0.5	.15	.10	.03
0.7	.34	.26	.13
0.9	.78	.74	.51

approach of Coates and Diggle since it does not depend on the form of the log spectral ratio to be powerful.
Stoica (1990) suggested an approach based on the Euclidean distance between the estimated parameters obtained by fitting autoregressive models to two sets of data. There are a number of problems associated with this method, Coates (1991) for details.
The Euclidean distance between the logarithm of spectral densities has been suggested as the basis for comparing stationary signals. One obvious estimate of the Euclidean distance between the logarithm of the spectral densities is the Euclidean distance between the logarithm of the periodograms. Two problems arise, firstly, the simple Euclidean distance between the logarithms of two periodogram gives a very weak test statistics since most frequencies will contribute very little (Coates, 1991). Secondly, the randomization test method based on this measure fails since the distance $|\log I_x(\omega_j)$ - $\log I_y(\omega_j)|$ will be the same for every random permutation. Bloomfield (1976) suggested smoothing the *cross periodogram* to estimate the coherency of two stationary processes. Using this idea, a simple three point moving average was used to smooth the permuted periodograms. A problem with smoothing the permuted periodograms is that there are problems with estimating values at the ends of the periodograms, but these are often where the differences between the periodograms are most

Table 3: Estimated Power of D_m-tests, Series Length n =64, Having the Same Noise Variances so the Process Variances are Different

α_1 or α_2	Results for the following sizes 0.10	0.05	0.01
white noise *versus* AR(1), $\alpha_1 >0$			
0.2	.30	.17	.05
0.4	.60	.48	.17
0.6	.74	.61	.31
0.8	.66	.58	.25
white noise *versus* AR(2), $\alpha_1 =0$, $\alpha_2 >0$			
0.2	.22	.10	.02
0.4	.30	.11	.03
0.6	.27	.15	.02
0.8	.16	.05	.03
white noise *versus* AR(3), $\alpha_2 =0$, $\alpha_1 = - \alpha_3 >0$			
0.2	.29	.16	.03
0.4	.63	.46	.12
0.6	.52	.35	.15
AR(1), $\alpha_1 =0.5$, *versus* AR(1) , $\alpha_1 >0$			
0.1	.58	.38	.17
0.3	.26	.14	.07
0.5	.10	.06	.01
0.7	.20	.11	.04
0.9	.20	.14	.04

important. This approach gave power estimates which were much weaker than those above.
An alternative to the Kolmogorov-Smirnov test statistic is the area between the cumulative periodograms, $F_x(\omega)$ and $F_y(\omega)$. Similar results to Table 1 or Table 2 can be obtained using the area. This is to be expected since most frequencies will contribute little (Coates, 1991).
The semiparametric approach by Coates and Diggle has limited applicability, in general, parametric assumptions even for the log spectral ratio are difficult to justify. An important advantage of the proposed non-parametric test is the simplicity of the assumptions required for its validity. In particular the validity of the randomization test method depends only on random assignment, in this case random data permutation.
The traditional method of testing whether a stationary series was generated by a white noise process is to use the Kolmogorov-Smirnov test (Jenkins and Watts, 1968). The standard sampling distribution cannot be used when comparing general stationary series, but the randomization test method can be used to create an approximate sampling distribution. The computation involved is not too great and the results obtained based on this method are equally competitive and sometimes much better than other methods. However, it is important to ensure that the stationary series has been standardised.

References

Basseville, M., 1988. Detecting Changes in Signals and Systems- A urvey, Automatica, 24: 309-326.

Basseville, M., 1989 Distance Measures for Signal Processing and Pattern Recognition, Signal Processing, 18: 349-369.

Bloomfield, P., 1976. Fourier Analysis of Time Series, John Wiley, London.

Coates, D. S., 1991. Comments on "Performance Evaluation Of Some Methods for Off-Line Detection of Changes in Autoregressive Signals" by P. Stoica, Signal Processing, 24: 353-357.

Coates, D. S. and P. J. Diggle, 1986. Tests for Comparing Two Estimated Spectral Densities, J. of Time Series Analysis. 7:7-20.

de Souza, P. and P. J. Thomson, 1982. LPC Distance Measures and Statistical Tests with Particular Reference to the Likelihood Ratio, IEEE Transaction on Acoustic Speech Signal Processing, ASSP-30: 304-315.

Diggle, P. J. and N. I. Fisher, 1991. Non-Parametric Comparison of Cumulative Periodograms, J. of Applied Statistics, 40: 423-434.

Edgington, E. S., 1980. Randomization Tests, New York, Marcel Dekker.

Hooton, J.W.L., 1991. Randomization tests: Statistics for Experimenters, Computer Methods and Programs in Biomedicine, 35: 43-51.

Jenkins, G. M. and D. G. Watts, 1968. Spectral Analysis and Its Application, Holden-Day, San Francisco.

Kiefer, J., 1959. K-sample Analogues of the Kolmogorov-Smirnov and Cramér-von Mises Tests, Annals of Mathematical Statististics, 30: 420-447.

Nelson, L.S., 1992. Technical Aids: A Randomization Test for Ordered Alternatives, J. of Quality Technology, 24: 51-53.

Priestley, M. B., 1981. Spectral Analysis and Time Series, Academic Press, London.

Stoica, P., 1990. Performance evaluation of Some Methods for Off-Line Detection of Changes in Autoregressive Signals, Signal Processing. 19: 301-310.