Portmanteau Tests For Univariate And Multivariate Time Series Models

by

Furat Barakat Hayran Dassi

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the

School of Computing and Mathematics,
Faculty of Computing, Engineering and Science,
University of South Wales

Director of Studies

Mark Griffiths

2022

Certificate of Research

I certify that, except where specific reference is made, the work described in this thesis is the result of the candidate's research. Neither this thesis, nor any part of it, has been presented, or is currently submitted, in candidature for any degree at any other University.

Candidate: ______________________

Director of Studies: ______________________

Date: ______________________

Abstract

This thesis examines how the number of available observations of a time series can influence its apparent stationarity as measured by two standard tests, namely the standard Dickey-Fuller (DF) test and the Augmented Dickey-Fuller (ADF) test. The univariate time series case is examined. A stationary time series generated from a first-order autoregressive process with positive or negative values of the parameter ϕ. Parameters were chosen that ensured that the series were theoretically stationary. The resulting time series produced were examined by, the DF, ADF, DF drift, ADF drift, DF trend and ADF trend tests. Monte Carlo experiments were undertaken using the R program for various values of parameters and different lengths of data, with each simulation repeated 10,000 times. The simulation studies show that the length of time series data affects the stationarity as identified by standard tests. For given values of the parameter ϕ of the first-order autoregressive model the minimum length of time series required to ensure the correct identification of the series' stationarity is presented.

Two new portmanteau tests were developed, bases on exponential weights of the residual autocorrelation function and the residual partial autocorrelation function. The asymptotic distributions of the new univariate portmanteau tests were derived. Monte Carlo experiments were used to compare the performance of the two new tests to existing tests. The simulation studies show that one of the new portmanteau tests, which is based on the partial autocorrelation function, is statistically more powerful than previous tests.

A new portmanteau test was developed for vector autoregressive moving average models, which is based on exponential weights of the residual covariance matrix. For this new multivariate portmanteau test the asymptotic distribution was derived. This new test was compared with previous tests using Monte Carlo experiments. The simulation study shows that the new multivariate portmanteau test is statistically more powerful than previous tests.

ACKNOWLEDGEMENTS

I would like to express very great appreciation to my director of studies Mark Griffiths for his valuable and constructive suggestions during the planning and development of this thesis. His willingness to give his time so generously has been very much appreciated.

I am also grateful to staff in the School of Computing and Mathematics, Faculty of Computing, Engineering and Science of the University of South Wales for sharing expertise, and sincere and valuable guidance and encouragement extended to me.

I take this opportunity to express gratitude to my parents, my brothers and sisters, and my friends for the unceasing encouragement, support and attention.

Finally, but by no means least thanks to my wife, Madleen, and my sons, Bakar and Oday: your love and understanding helped me through the dark times. Without you believing in me, I never would have made it. It is time to celebrate; you earned this degree right along with me.

The researcher

Furat Barakat Hayran Dassi

Contents

List of Figures

List of Tables

Glossary / Notation

$\mathrm{z_t}$	– is a value of a time series at time t.
$\tilde{\mathrm{z}}_{\mathrm{t}}$	– is a value of a non-stationary time series at time t.
$\bar{z}$	– is the sample mean of a time series.
μ	– is the mean of a time series.
Z	– is a random variable of a discrete sample space.
S	– is the sample variance of a time series.
σ^2	– is the variance of a time series.
$cov(z_{t_1}, z_{t_2})$	– is the covariance between two random variables z_{t_1} and z_{t_2}.
SE	– is the standard error.
k	– is a lag in a time series.
c_k	– is the autocovariance coefficient at lag k.
ϕ	– is a parameter of an autoregressive process.
$\hat{\phi}$	– is an estimation of a parameter of an autoregressive process.
e_t	– is a white noise at time t.
$\hat{e}_t$	– is a white noise estimation at time t.
θ	– is a parameter of moving average process.
$\hat{\theta}$	– is a parameter estimation of moving average process.
n	– is number of observations.
m	– is the maximum number of lags examined in a portmanteau test.
B	– is the backshift operator.
∇	– is the difference operator.
∇^s	– is the difference operator of order s.
$\rho(z_{t_1}, z_{t_2})$	– is the correlation function between z_{t_1} and z_{t_2}.
ρ_k	– is an autocorrelation function (ACF) at lag k.
$\hat{\rho}_k$	– is an estimation of the autocorrelation function (ACF) at lag k.
γ_k	– is an autocovariance function at lag k

$\boldsymbol{\Gamma}_n$	$-$ is an autocovariance matrix of order n.
$\boldsymbol{P}_n$	$-$ is an autocorrelation matrix of order n.
$\hat{R}_m$	$-$ is a residual correlation matrix of order m.
ϕ_{kk}	$-$ is a partial autocorrelation function (PACF) at lag k.
$\hat{\phi}_{kk}$	$-$ is an estimation of partial autocorrelation function (PACF) at lag k.
ψ_{j}	$-$ is a sequence of constants.
$\mathrm{AR}(p)$	$-$ is an autoregressive process of order p.
$\mathrm{MA}(q)$	$-$ is a moving average process of order q.
$\mathrm{ARMA}(p,q)$	$-$ is an autoregressive moving-average process of order (p,q).
χ^2	$-$ is the Chi-squared.
$\Rightarrow$	$-$ is the convergence in the distribution.
$\tilde{Q}_{BP}$	$-$ is the Box and Pierce portmanteau test.
$\tilde{Q}_{LB}$	$-$ is the Ljung and Box portmanteau test.
$\tilde{Q}_{M}$	$-$ is the Monti portmanteau test.
D_m	$-$ is the Peña and Rodríguez (2002) portmanteau test statistic.
D_m^*	$-$ is the Peña and Rodríguez (2006) portmanteau test statistic.
$\tilde{Q}_{MM}$	$-$ is the Mahdi and McLeod portmanteau test.
$\tilde{Q}_{FGLB}$	$-$ is the Fisher and Gallaher portmanteau test.
$\tilde{Q}_{GFK}$	$-$ is the Gallaher and Fisher portmanteau test with Kernel-based weights.
$\tilde{Q}_{GFD}$	$-$ is the Gallaher and Fisher portmanteau test with data adaptive weights.
$\tilde{Q}_{EXLB}$	$-$ is the Ljung and Box test portmanteau test with exponential weights.
$\tilde{Q}_{EXM}$	$-$ is the Monti test with exponential weights portmanteau test.
$\mathcal{K}(\cdot)$	$-$ is the Daniell Kernel function.
$\mathrm{VAR}(p)$	$-$ is a vector autoregressive process of order p.
$\mathrm{VMA}(q)$	$-$ is a vector moving-average process of order q.

VARMA(p, q)	− is a vector autoregressive moving-average process of order (p, q).
d	− is the number of components in a vector.
$\mathbf{z}_{\mathrm{t}}$	− is a $d \times 1$ vector of variables observed at t.
$\boldsymbol{\Psi}_0$	− is a $d \times d$ identity matrix.
I	− is a general $d \times d$ identity matrix.
D	− is the diagonal matrix.
$\boldsymbol{\Psi}_i$	− is a $d \times d$ coefficients matrices of a vector autoregressive process.
$\boldsymbol{\Psi}(B)$	− is a $d \times d$ matrix polynomial of the backshift operator B.
Π_i	− is a $d \times d$ coefficients matrices of a vector moving-average process.
$\mathbf{\Phi}$	− is a $d \times d$ parameter matrices of a vector autoregressive process.
$\boldsymbol{e}_t$	− is a zero mean vector white noise process of dimension d.
$\mathbf{\Theta}$	− is a $d \times d$ parameter matrices of vector moving average process.
$\widehat{\mathbf{\Phi}}$	− is an estimation of $d \times d$ parameter matrices of a vector autoregressive process.
$\hat{\boldsymbol{e}}_t$	− is a zero mean vector white noise process estimation of dimension d.
$\widehat{\mathbf{\Theta}}$	− is an estimation of $d \times d$ parameter matrices of vector moving average process.
$\widehat{\boldsymbol{R}}_k$	− is the sample autocorrelation matrix at lag k.
$\boldsymbol{L}$	− is a lower triangular matrix.
tr	− is the sum of the diagonal matrix.
$\mathbf{\Sigma}$	− is a covariance matrix of a vector white noise process.
$\mathbf{\Phi}(B)$	− is the matrix polynomial of the backshift operator B of an autoregressive process of order p.
$\mathbf{\Theta}(B)$	− is the matrix polynomial of the backshift operator B of a moving average process of order q.
$\mathbf{\Sigma}(k)$	− is a covariance matrix of a vector white noise process at lag k.
$\boldsymbol{\Gamma}(k)$	− is a covariance matrix at lag k.
$\boldsymbol{\rho}(k)$	− is a correlation matrix at lag k.

$\widehat{\boldsymbol{\mu}}$	− is a sample mean vector
$\widehat{\boldsymbol{\Gamma}}(0)$	− is a sample covariance matrix.
$\widehat{\boldsymbol{\Gamma}}(k)$	− is a sample covariance matrix at lag k.
$\widehat{\boldsymbol{\rho}}_{ij}(k)$	− is a sample correlation matrix at lag k.
vec	− is a vector operator.
$\otimes$	− is a Kronecker product.
$LR(k)$	− is a likelihood ratio test at lag k.
AIC	− is the Akaike information criterion.
BIC	− is the Bayesian information criterion.
HQ	− is the Hanna and Quinn information criterion.
$\tilde{Q}_H$	− is the Hosking portmanteau test of a VARMA model.
$\tilde{Q}_H^*$	− is the modified Li and Mcleod portmanteau test of a VARMA model.
$\tilde{Q}_{LM}$	− is the Li and McLeod portmanteau test of a VARMA model.
$\tilde{Q}_{LM}^*$	− is the modified Li and McLeod portmanteau test of a VARMA model.
$\tilde{Q}_{MMV}$	− is the Mahdi and McLeod portmanteau test of a VARMA model.
$\widehat{\mathfrak{R}}_m$	− is the residual autocorrelation matrix.
$\tilde{Q}_{EXCO}$	− is the exponential weights portmanteau test of a VARMA model and based on covariance matrix.
$\tilde{Q}_{EXAU}$	− is the exponential weights portmanteau test of a VARMA model and based on autocorrelation matrix.
$\boldsymbol{Q}$	− is an idempotent matrix.
†	− is a transpose operator of a vector or matrix.
$\lvert a \rvert$	− is the absolute value of real valued constant a.
$\lVert \boldsymbol{A} \rVert$	− is the determinant of a matrix, $\boldsymbol{A}$.
DF test	− is the Dickey-Fuller test.
ADF test	− is the Augmented Dickey-Fuller test.
μ_0	− is the drift.

$\mu_0 + \mu_1 t$	– is the deterministic linear trend.
τ	– is the sum of autoregressive coefficients.
$\hat{\tau}$	– is the estimation of the sum of autoregressive coefficients.
$l12$	– is the lag length.

Chapter 1 - Introduction

1.1 Introduction

Box and Jenkins published a classic book on time series analysis in 1970, which set the foundation for developments in time series analysis for the next 50 years. In this work they described the model building process as consisting of three main stages, namely: model identification, parameter estimation and diagnostic checking.

This thesis discusses the basic ideas of the Box-Jenkins methodology. It concentrates on diagnostic checking; in particular, it focuses on univariate portmanteau testing and multivariate portmanteau testing of models, with the aim of developing new and better portmanteau tests. The thesis also examines how the length of data can affect how standard unit root tests identify the stationarity of a time series.

This chapter introduces the aims of the thesis and also provides a brief history of time series analysis.

1.1.1 Aims of research

Portmanteau tests were introduced for the first time in 1970 by Box and Pierce (1970). Later other portmanteau tests were introduced by researchers, such as, Ljung and Box (1978), Monti (1994), and Gallagher and Fisher (2012, 2015). The aim of this thesis is to develop and evaluate new portmanteau tests that are more powerful than previously published portmanteau tests. The approach taken was be to conduct Monte Carlo experiments to explain the behaviour of the portmanteau tests and evaluate their performance compared to existing portmanteau tests. New portmanteau tests were developed for univariate autoregressive moving average models and for vector autoregressive moving average models, with the aim of improving on existing portmanteau tests.

Another aim of this thesis is to examine how the length of data of a time series can influence its apparent stationarity as measured by two standard tests. The univariate time series case is examined. To explore this issue, time series were generated from a known statistical model, a first-order autoregressive process and a second-order autoregressive process. Parameters were chosen that ensure the series were theoretically stationary. The standard Dickey-Fuller test and the Augmented Dickey- Fuller test were used to determine whether the series of observations produced were stationary or non-stationary. Monte Carlo experiments were undertaken using the R program for various model parameters and lengths of series and each simulation was repeated 10,000 times.

1.2 A history of time series

1.2.1 Definition of time series

A time series is a sequence of discrete observations arranged in chronological order. These data come from repeated observations and may be available, for instance, hourly, daily, weekly, monthly or yearly. Examples of time series abound in such fields as economics, business, the natural sciences, engineering, and the social sciences**.** For example, Figure 1.1 shows the Consumer Price Index (CPI) data of the UK inflation rate, monthly from January 2005 to January 2015 (CPI, 2015). The aim of time series analysis is to find the relationship between data over a period of time, and use this to forecast future measurements.

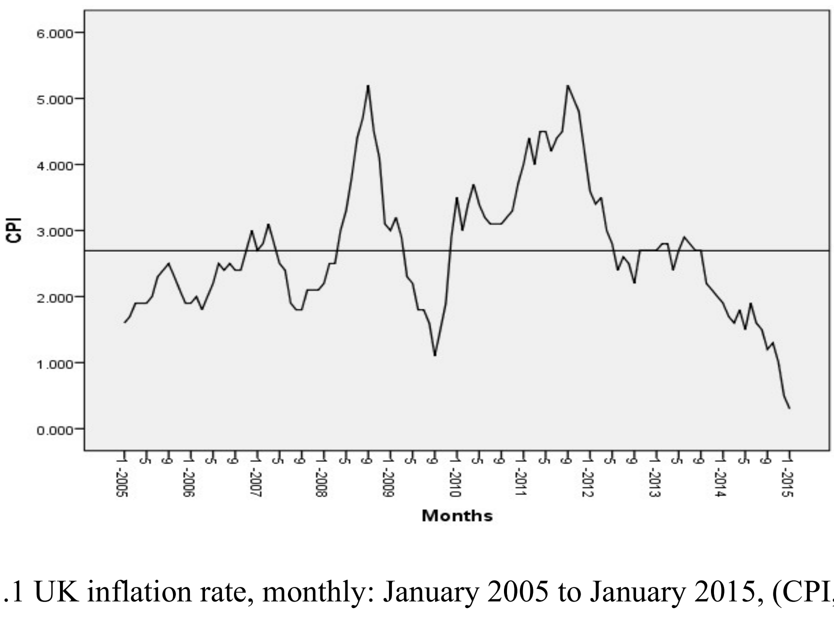

Figure 1.1 UK inflation rate, monthly: January 2005 to January 2015, (CPI, 2015).

1.2.2 The graphical representation of time series data

The plotting of data in time series analysis started with William Playfair, who was the first researcher to draw a chart of data against a time axis. For example, Playfair's first chart showed the sum total of England's imports and exports from the year 1700 to 1782 displayed as a line plot, this is reproduced in Figure 1.2, (Playfair, 1801). Several years later, the plotting of data over a period of time was used by the medical researcher Wunderlich (1870). He drew charts such as fever curves, which plotted a patient's temperature over the course of time (see Figure 1.3). Brinton (1914) used the time plot to represent data in many different subjects, for example, the average yearly earnings of Princeton graduates, (see Figure 1.4).

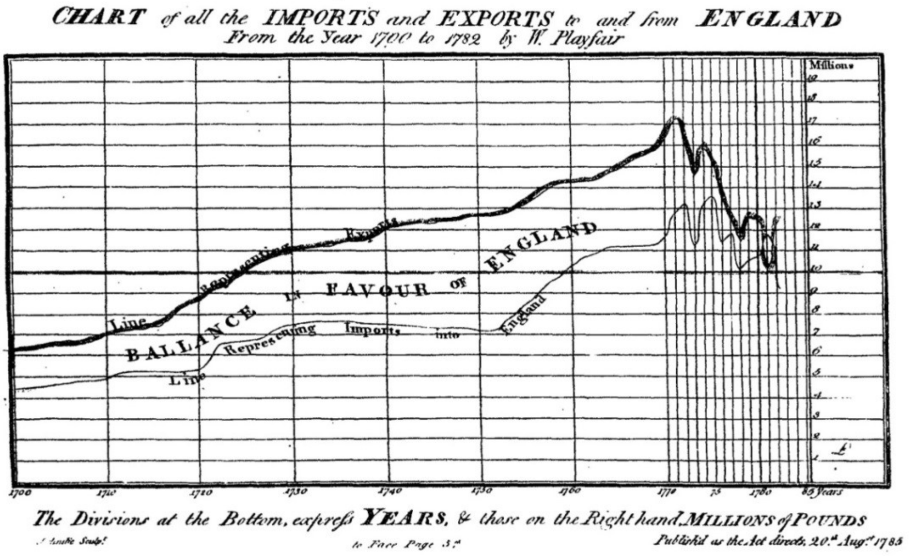

Figure 1.2 Imports and Exports to and from England from the year 1700 to 1782, reproduced from Playfair (1801).

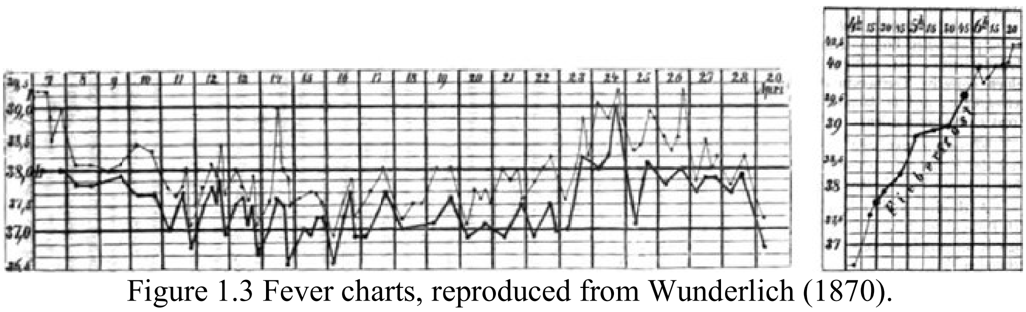

Figure 1.3 Fever charts, reproduced from Wunderlich (1870).

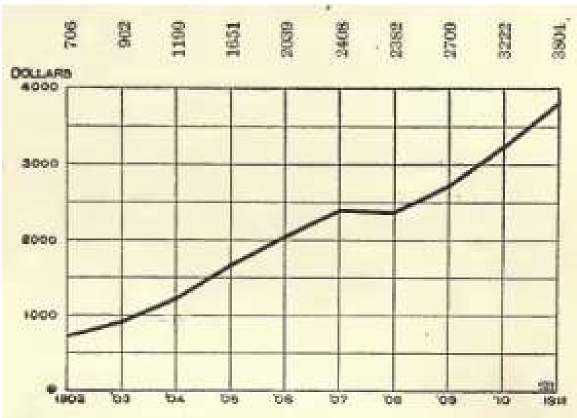

Figure 1.4 Average income of 155 Princeton graduates of the class 1901 for ten years after graduating, reproduced from Brinton (1914).

1.2.3 Thiele, Yule and Hooker concept of time series

Statistical analysis of time series data was first undertaken by Thiele (1880a; 1880b), when he published his first paper to analyse a model of a time series consisting of a regression component, a Brownian motion component, and white noise component. He derived Brownian motion with independent and normal distributions by using the method of least squares, and estimated variances proportional to the Brownian motion. He was also the first person to give a recursive computational methodology for filtering and predicting, which is now known as the Kalman filter (Kalman and Bucy, 1961). However, this work was unrecognized at the time. A more detailed description of Thiele's paper is presented by Lauritzen (1981, 2002).

The mean, variance, standard deviation, and the theory of correlation coefficients and partial correlation coefficients between data in time series were developed by the British statistician Yule (1895, 1896, 1897a, 1897b, 1907), and the statistician Hooker (1901). Yule (1895, 1896) used the correlation coefficients to examine the relationship between welfare and poverty in the field of economics statistics. The next study by Yule (1897a, 1897b) used the partial correlation coefficients. Hooker applied correlation to find the relationship between Britain's marriage rate and export trade over the period 1857-1899, (Hooker, 1901).

1.2.4 The periodogram method

The periodogram is used to identify and calculate the significance of different frequencies of a time series. The physicist Sir Arthur Schuster investigated periodicities in time series, such as, the periodicity of earthquakes, terrestrial magnetism and sunspot numbers. Schuster determined the periodogram by

$$A = \sum_{s=0}^{(n-1)\alpha} a_s \, cos \, \frac{2\pi}{n} s$$

$$B = \sum_{s=0}^{(n-1)\alpha} a_s \, sin \, \frac{2\pi}{n} s$$

where a_s takes the values $a_1, a_2, a_3, \ldots$, at equidistant values of time $t_0, t_0 + \alpha, t_0 + 2\alpha, \ldots$, where n is a number of the observations, and s is sunspots numbers. The calculated periodogram is plotted in Figure 1.5 and shows that a maximum amplitude occurs at a period of 10 and 11 years, (Schuster, 1897, 1906). Meanwhile, in 1922 Sir William Beveridge gave a periodogram analysis of wheat-price indices, extended over approximately 300 years from 1545 to 1845, (Beveridge, 1922). The theory of Schuster was modified by Whittaker and Robinson (1924), they constructed their periodogram from values of correlation ratios as they relate to each value of the arithmetic sequence. Meanwhile, the researcher Walker (1931) applied the correlation periodogram on the Port Darwin air pressure data. Later, the periodogram method was further developed by the researchers Davis (1941) and Kendall (1945).

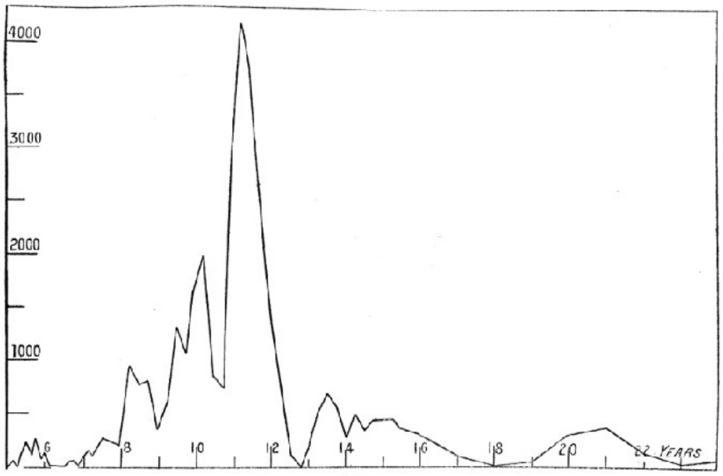

Figure 1.5 The periodogram of Wolfer's sunspot number for each month of the years 1749 to 1901, reproduced from Schuster (1906).

1.2.5 Autoregressive, moving average and mixed models

The next developments in time series analysis were by the British statistician Yule and the Russian statistician Slutsky. Yule (1927) introduced the scheme of the autoregressive model, which is defined as

$$\mathrm{z_t} = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} + e_t, \tag{1.1}$$

where $\mathrm{z_t}$ is a value of a time series at time t. The current value of the process $\mathrm{z_t}$ is expressed as a weighted sum of the previous p values (with weights $\phi_1, \phi_2, \ldots, \phi_p$) plus the current shock e_t (white noise). Yule applied the scheme of the autoregressive model to Wolfer's sunspot data that had been used by Schuster in his periodogram method (1897, 1906). Yule obtained better results with the autoregressive model than Schuster's periodogram method. An autoregressive model of order p is denoted as $\mathrm{AR}(p)$. Following Yule's work, Slutsky (1937) introduced the scheme of the moving average model, which is defined as

$$\mathrm{z_t} = e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} \tag{1.2}$$

where z_t is assumed to be generated as a finite moving average of a sequence of independent and identically distributed random variables e_t and $\theta_1, \theta_2, \dots, \theta_q$ are weights. A moving average model of order q is denoted as MA(q).

Furthermore, in 1950 Walker described a mixed autoregressive-moving average process for the first time, (Walker, 1950). Walker achieved this by adding together the moving average and autoregressive schemes, as defined by

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} \qquad (1.3)$$

Box and Jenkins in 1970 referred to the mixed autoregressive-moving average process, which had been described by Walker, and they named it the general mixed autoregressive-moving average process of order (p, q). They also gave this model the acronym ARMA(p, q), and showed that it can be put in the following form (Box and Jenkins, 1970).

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) z_t = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q) e_t$$

or

$$\phi(B) z_t = \theta(B) e_t \qquad (1.4)$$

where $\phi(B)$ and $\theta(B)$ are polynomials in B of degree p and q respectively, and B is the backshift operator, which is defined as

$$B^j z_t = z_{t-j} \qquad j = 0, 1, 2, \dots .$$

1.2.6 Stationary random processes

Many other researchers have contributed to the development of the mathematical foundations of stationary stochastic processes in time series. Kolmogorov (1931, 1941) developed the theory of Markov Processes, the theory of stochastic processes and introduced a general formula for the mean squared error of a linear extrapolation of a stationary random sequence. Khinchin (1933, 1934) developed the theory of stationary processes and correlation theory. Wold (1938) developed the probabilistic theory of stationary time series, which is based on the development of ergodic theory and prediction theory.

1.2.7 The methods of parameter estimation of the autoregressive moving average model

The methods of parameter estimation have an important role in time series analysis. There are many methods to estimate the parameters of the autoregressive, moving average and mixed models, such as, the maximum likelihood method and the least squares method.

The estimation of autoregressive processes.

Mann and Wald (1943) estimated an AR(p) process by using the method of maximum likelihood, while, Hurwicz (1945) estimated the AR(p) process by using the least-squares method. Meanwhile, Guy and Donald (1949) investigated the merits of autoregressive transformations and the reduced form transformation, with the main result of estimating the parameter structure of the AR(1) model. Kendall (1949) investigated the second-order autoregressive process.

The estimation of moving-average processes.

Durbin (1959) estimated the moving-average process by using the least squares method and the maximum likelihood estimators. This method was extended by Walker (1961).

The estimation of autoregressive and moving-average processes.

Durbin (1960) introduced the least squares method to estimate the parameters of the autoregressive moving average model.

1.2.8 The analysis of residuals and forecasting

Anscombe and Tukey (Anscombe 1961, Anscombe and Tukey, 1963) examined the analysis of residuals as a means of detecting departures from the model's assumptions, and they indicated how transformations might be constructed from certain functions of the residuals. Economic forecasting problems were investigated by Persons (1924) and Margret (1929). Later on work by Cowles (1933, 1944) predicted the future movements of stock price in his investigation of the Stock Market, which followed on from the earlier work of Bachelier (1900). Generally, forecasting and filtering methods have been developed by researchers, such as, Wiener, (1949), Kalman (1960), and Yaglom (1962).

1.2.9 Non-stationary time series

Most time series such as those found in economics and business are not stationary and they will exhibit deterministic trends, random walk and other non-stationary behaviour. Moving average trends were considered for the first time by Hooker in 1901, when he examined the correlation between the marriage rate and trade over the period 1857-1899, Hooker (1901). Hooker, in 1905, considered another method to remove the trend, which he called a differencing method. Hooker (1905) used differencing to remove the trend before he estimated the correlation between the variables, and he applied this method to corn price and marriage rate data. Later, the researcher Student further developed the differencing method, he estimated the first, second, ... nth differences between variables to get the correlation required, and he also determined the correlation between residual variation (Student, 1914). The differencing method was employed by the American economist and statistician Irving Fisher. He used the differencing method to transform data from non-stationary to stationary, (Fisher, 1925). Other methods of transforming time series from non-stationary to stationary, such as trends and random walk have been developed by researchers, such as, Grenander and Rosenblatt (1957), Box and Jenkins (1962), Box and Tiao (1965) and Priestley and Rao (1969).

In the case of a non-stationary time series, $\tilde{z}_t$ represents a value of the time series at time t. The differencing method is used to transform data from non-stationary to stationary.

$$\mathrm{z_t} = \nabla \tilde{z}_t$$

where $\nabla = (1 - B)$ is the differencing operator.

Thus, for the first difference $\nabla \tilde{z}_t = \tilde{z}_t - \tilde{z}_{t-1}$

The second difference $\nabla^2 \tilde{z}_t = \nabla(\tilde{z}_t - \tilde{z}_{t-1})$

$$= \nabla \tilde{z}_t - \nabla \tilde{z}_{t-1_t}$$

$$= (\tilde{z}_t - \tilde{z}_{t-1}) - (\tilde{z}_{t-1} - \tilde{z}_{t-2}), \quad t = 3, \dots, n$$

and for the sth difference $\nabla^s \tilde{z}_t = \nabla^{s-1} \tilde{z}_t - \nabla^{s-1} \tilde{z}_{t-1}$

The incorporation of the differencing method in to ARMA modelling was undertaken by Box and Jenkins (1970). This resulted in the general autoregressive integrated moving average process ARIMA(p, d, q), which can be defined as

$$\phi(B)\nabla^s \tilde{z}_t = \theta(B)e_t$$

1.2.10 Box and Jenkins methodology

Box and Jenkins (1970) developed a three-stage methodology to model time series data, namely:

1. **Identification**, which is the determination of a specific model on the basis of certain statistical figures by using the sample of autocorrelation function and the sample of partial autocorrelation function.
2. **Estimation**, which is estimation of the parameters of the model estimated by using either the maximum likelihood function, least squares method or Bayes' theorem.
3. **Diagnostic checking**, which involves the checking specification of the model by statistical tests.

1.2.11 Vector ARMA models

The extension of the univariate time series models to multivariate ARMA time series models was first proposed by Quenouille (1957). He studied a problem with five variables with 82 observations of each variable from the year 1867 to 1948. He fitted a vector first-order autoregressive process to the data. Following this, he discussed the identification method and the estimation method of vector autoregressive, vector moving average and vector autoregressive moving average processes.

Later, Whittle developed a method of fitting a model to a vector autoregressive process by using the autocovariance matrices together with the Yule-Walker equations, (Whittle, 1963). Meanwhile, Hannan (1970) discussed a multivariate time series process with theories of estimation methods for vector autoregressive, vector moving average and mixed vector autoregressive moving average processes. The researchers Zellner and Palm in 1974 applied simultaneous equation models within the context of the general linear multiple time series process, (Zellner and Palm, 1974). Furthermore, vector ARMA models have been discussed by many researchers, such as, Wallis (1977), Tiao and Box (1981), Hannan and Kavalieris (1984), Tiao and Tsay (1989), and Wei (2006).

Later other researchers extended the Box and Jenkins univariate model building methodology to vector ARMA models.

Model identification

The researcher Akaike in 1973 introduced the information criterion to identify the vector AR process, (Akaike, 1973). Later, Schwarz in 1978 introduced the Bayesian information criterion to identify the vector AR process. Hanna and Quinn (1979) and Quinn (1980) introduced another form to identify the vector AR process. Tiao and Box (1981) applied the likelihood ratio test to identify the vector AR process. Tiao and Box (1981) suggested the cross-correlation matrices to identify vector MA processes. Procedures of model identification for vector ARMA models have been developed by Tiao and Box (1981).

Model estimation

Tunnicliffe (1973), Reinsel (1979) and Anderson (1980) derived the conditional likelihood method to estimate VARMA models. The exact likelihood function of a vector moving average process was derived by Osborn (1977), and Phadke and Kedem (1978). Later, the exact likelihood function for a stationary vector ARMA was derived by Hillmer and Tiao (1979), Nicholls and Hall (1979) and Anderson (1980).

1.2.12 Portmanteau testing

Portmanteau testing has been developed to select the best fitted model after the ARMA models have been identified and estimated. Box and Pierce (1970) introduced the first portmanteau test, which is based on the residual of the autocorrelation function; this test is approximately distributed as a chi-squared distribution. Ljung and Box (1978) introduced a new portmanteau test and they showed this test to be more powerful than the Box and Pierce (1970) test. In addition, Monti (1994) introduced a test that is based on the residual of partial autocorrelation, and she showed that this test is at least as powerful as the Ljung and Box (1978) test. Later, other portmanteau tests, based on the determinant of the autocorrelation matrix, were introduced by researchers, such as, Peña and Rodríguez (2002, 2006). Mahdi and McLeod (2011) introduced a new test, which is based on the result of

Peña and Rodríguez (2002, 2006), the test is based on the log of the sample autocorrelation matrix. Later, Fisher and Gallagher (2012) presented a new portmanteau test that is a weighted sum of the squares of the residual autocorrelation coefficients. In 2015, Gallagher and Fisher introduced further portmanteau tests that are based on the weighted sums of the squared residual autocorrelations in three different cases, namely, the Kernel-based Weights test, the Geometrically Decaying Weights test and the Data Adaptive Weights test.

1.2.13 Multivariate portmanteau test

The first application of a portmanteau test to multivariate autoregressive models was by Chitturi in 1974, (Chitturi, 1974). Later, a portmanteau test was developed for vector ARMA models by Hosking (1980), which is based on the residual autocorrelation matrix. Hosking (1980) modified the multivariate portmanteau test, which is based on the residual autocovariance matrix. Li and McLeod (1981) gave another multivariate portmanteau test, which is based on the autocorrelation matrix. Mahdi and McLeod (2011) gave another multivariate portmanteau test, which is based on the residual autocorrelation matrix.

1.3 Thesis overview

Chapter 2 presents the notation used in the thesis and defines the key statistical terms employed in time series analysis, namely: the definition of the mean, variance, autocovariance, autocorrelation and partial autocorrelation functions. The same chapter also outlines the Box-Jenkins model building methodology, autoregressive process, moving average process and mixed autoregressive-moving average processes.

Chapter 3 offers a brief outline of the vector autoregressive moving average process, namely, the vector autoregressive process, the vector moving average process and the mixed vector autoregressive moving average process. Chapter 3 also introduces an outline of model building of vector autoregressive moving average models and associated portmanteau tests.

Chapter 4 details simulation studies that explore how the length of data in a time series affects the identification of its stationarity, as determined by standard tests (the DF test, the DF drift test and the DF trend test, the ADF test, the ADF drift test and the ADF trend test).

The aim of the thesis is to investigate whether the length of data in a time series influences its apparent stationarity.

Chapter 5 introduces two new portmanteau tests, the first is based on the exponential weighted sums of the squared sample autocorrelations function and the second is based on the exponential weighted sums of the squared sample partial autocorrelation function. The aim of the thesis is to investigate whether the new portmanteau tests are more powerful than previous portmanteau tests found in the literature.

Chapter 6 introduces a new multivariate portmanteau test, which is based on the residual covariance and autocorrelation matrices with exponential weights. The aim is to investigate whether the new multivariate portmanteau test is more powerful than the previous multivariate portmanteau tests that have been published.

Chapter 7 provides a summary of the thesis findings together with recommendations for future work.

1.4 Summary

A brief history of time series analysis has been provided, from the first chart drawn by William Playfair (1801) to the Box & Jenkins methodology in 1970, (Box & Jenkins, 1970). The development of portmanteau tests has been outlined, starting with Box and Pierce’s introduction of the first portmanteau test in 1970 (Box and Pierce, 1970), to latest developments introduced by Gallagher and Fisher (Gallagher and Fisher, 2015). The extension of ARMA time series models to the multivariate vector case has been considered, including the work on model building of VARMA time series models. Finally, existing vector portmanteau tests have been introduced, such as, Chitturi (1974), Hosking (1980), Li and McLeod (1981) and Mahdi and McLeod (2011).

Chapter 2 – Box And Jenkins Methodology

This chapter provides an introduction to the Box and Jenkins methodology of modelling time series. It provides the standard definitions used in time series analysis and gives the notation that is used throughout the thesis. It starts with definitions of the mean and the variance of a time series, the differencing of a time series, and then progresses on to the autocovariance, autocorrelation and partial autocorrelation functions. It also provides some definitions of time series such as stochastic process, stationary, Gaussian process, weak stationary, white noise, backshift operator, linear process and invertibility. Finally, the Chapter discusses the autoregressive, the moving average and the mixed autoregressive-moving average processes.

2.1 Standard definitions in time series analysis

2.1.1 The mean, the variance and the covariance functions

The mean of a stationary time series $\{z_t\}$ indicates the overall level of the series. The sample mean $\bar{z}$ provides an estimate of the true mean μ of the time series.

Definition 2.1.1 The sample mean of a time series is the sum of the observations for each time period z_t divided by the total number of observations n, then

$$\bar{z} = \frac{1}{n}\sum_{t=1}^{n} z_t \tag{2.1}$$

The sample variance of a time series is calculated using the normal approach, that is, determine the deviation of each observation from the mean, square each deviation, sum the deviations and divide by the total number of observations n, then

$$S_t^2 = \frac{1}{n}\sum_{t=1}^{n}(z_t - \bar{z})^2 \tag{2.2}$$

Definition 2.1.2 If Z is a random variable of a discrete sample space taking the values z_t, t = 0, +1, +2, …, then the expectation of Z is denoted as $E[Z]$ and is defined as

$$\mu = E[Z] = \sum_{t\geq 1} z_t \, p[Z = z_t], \tag{2.3}$$

where $p[Z = z_t]$ is the probability of the occurrence of the value of z_t.

Definition 2.1.3 The variance of Z, denoted by $var(Z)$, is defined by

$$\sigma^2 = var(Z) = E[(Z - \mu)^2]$$

where $\mu = E[Z]$ is the expectation of Z, then

$$= E[Z^2] - \mu^2$$

Definition 2.1.4 The covariance between two random variables z_{t_1} and z_{t_2} with expected values μ_{t_1} and μ_{t_2} is defined as

$$cov(z_{t_1}, z_{t_2}) = E\left[(z_{t_1} - \mu_{t_1})(z_{t_2} - \mu_{t_2})\right] \tag{2.4}$$

$$= E\left[z_{t_1} z_{t_2}\right] - \mu_{t_1}\mu_{t_2}.$$

Definition 2.1.5 The correlation function between z_{t_1} and z_{t_2} is defined as

$$\rho(z_{t_1}, z_{t_2}) = \frac{cov(z_{t_1}, z_{t_2})}{\sigma_{t1}\sigma_{t2}}. \tag{2.5}$$

2.1.2 Stochastic process

A stochastic process is a family or sequence of random variables $\{z_t\}_{-\infty}^{\infty}$ indexed by time *t*. In most applications, the time index is regularly spaced, and represents calendar time, for example, days, months, or years. A realization of a stochastic process with n observations is the sequence of observed data $\{z_t\}_1^n$. If the probability distribution associated with any set of times is a Normal distribution, the process is called a Normal or Gaussian process (Box and Jenkins, 1970).

2.1.3 Strict stationarity and weak stationarity

A stochastic process is said to be strictly stationary if its properties are unaffected by a change of time origin; that is, if the joint probability distribution function associated with n observations $z_{t_1}, z_{t_2}, \dots, z_{t_n}$, made at any set of times $t_1, t_2, \dots, t_n$, is the same as that associated with n observations, (Tsay, 2005).

A stochastic process is said to be weakly stationary if all its joint moments up to order n exist and are time invariant. Therefore, a second-order weak stationary process will have a mean, variance and covariance are time invariant. Sometimes, the term covariance stationary is used to describe a second-order weak stationary process, (Tsay, 2005).

2.1.4 Differencing

The aim of differencing is to transform a non-stationary time series to a stationary one. Differencing is a simple operation that involves calculating successive changes in the values of a time series. It is used when the mean of the series is changing over time.

For a non-stationary time series $\tilde{z}_t$ the first difference is calculated by Equation 2.6 which gives the series z_t,

$$z_t = \tilde{z}_t - \tilde{z}_{t-1}, \qquad t = 2,3, \dots, n \tag{2.6}$$

where z_t is called the first difference of $\tilde{z}_t$. If the first difference does not have a constant mean, the series can be differenced again to give the second difference of z_t ((Pankratz,1983).

The second difference is

$$\nabla^2 \tilde{z}_t = (\tilde{z}_t - \tilde{z}_{t-1}) - (\tilde{z}_{t-1} - \tilde{z}_{t-2}), \quad t = 3, \dots, n$$

2.1.5 The autocovariance function

If z_t is stationary then the covariance between z_t and its value z_{t+k}, separated by k intervals of time, is called the autocovariance at lag k and defined by

$$\gamma_k = cov[z_t, z_{t+k}] = E[(z_t - \mu)(z_{t+k} - \mu)] \tag{2.7}$$

Since z_t is stationary, this gives

$$\begin{aligned}\gamma_k &= cov[z_t, z_{t+k}] \\ &= cov[z_{t-k}, z_t] \\ &= \gamma_{-k}\end{aligned}$$

where, μ is the mean of observations and γ_k is the autocovariance function at lag k.

2.1.6 The autocorrelation function (ACF)

The standard formula for calculating an autocorrelation function is

$$\rho_k = \frac{E[(z_t - \mu)(z_{t+k} - \mu)]}{\sqrt{E[(z_t - \mu)^2]E[(z_{t+k} - \mu)^2]}}$$

$$= \frac{E[(z_t - \mu)(z_{t+k} - \mu)]}{\sigma_z^2} \tag{2.8}$$

where, σ_z^2 is the variance, μ is the mean of observations and ρ_k is the autocorrelation function at lag k.

Note that when $\sigma_z^2 = \gamma_0$, then

$$\rho_k = \frac{\gamma_k}{\gamma_0} \tag{2.9}$$

2.1.7 Autocovariance matrix

The covariance matrix associated with a stationary process for observations (z_1 , z_2 , ..., z_n) made at n successive times is

$$\boldsymbol{\Gamma}_n = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{n-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots & \gamma_{n-3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \gamma_{n-1} & \gamma_{n-2} & \gamma_{n-3} & \cdots & \gamma_0 \end{bmatrix}. \tag{2.10}$$

As with autocorrelations, autocovariance can be conveniently represented in matrix form. Start with matrix (2.10) and divide each element by γ_0. All elements on the main diagonal become one, indicating that each z_t is perfectly correlated with itself. All other γ_k values become ρ_k values as indicated by Equation (2.9):

$$= \sigma_z^2 \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{n-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{n-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{n-3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \cdots & 1 \end{bmatrix}$$

$$= \sigma_z^2 \boldsymbol{P}_n \tag{2.11}$$

where, $\boldsymbol{\Gamma}_n$ is the autocovariance matrix and $\boldsymbol{P}_n$ is the autocorrelation matrix (Box and Jenkins, 1970).

2.1.8 Conditions satisfied by the autocorrelations of a stationary process

The positive definiteness of the autocorrelation matrix (2.10) implies that its determinant and all principal minors are greater than zero, Box, Jenkins, Reinsel and Ljung, (2015).

For $n = 2$

$$\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix} > 0$$

so that

$$1 - \rho_1^2 > 0$$

and hence

$$-1 < \rho_1 < 1$$

Similarly, for $n = 3$, this requires

$$\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix} > 0, \quad \begin{vmatrix} 1 & \rho_2 \\ \rho_2 & 1 \end{vmatrix} > 0$$

$$\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix} > 0$$

which implies

$$-1 < \rho_1 < 1$$

$$-1 < \rho_2 < 1$$

$$-1 < \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} < 1$$

The determinant of the second matrix must be positive as well as the determinants of its principal minors, which implies $|\rho_1| \leq 1$ and $|\rho_2| \leq 1$, so

$$1 + 2\rho_1^2\rho_2 - 2\rho_1^2 - \rho_2^2 \geq 0 \quad \Rightarrow \left(\rho_2 - (2\rho_1^2 - 1)\right)(\rho_2 - 1) \leq 0$$

Since $|\rho_2| \leq 1$,

$$\rho_2 - (2\rho_1^2 - 1) \geq 0 \Rightarrow \rho_2 \geq 2\rho_1^2 - 1$$

Which lead to

$$-1 < \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} < 1$$

2.1.9 Estimation of autocovariance and autocorrelation functions

An estimate for the autocorrelation function can be obtained by,

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k}(z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=1}^{n}(z_t - \bar{z})^2}. \tag{2.12}$$

The variance is the average squared difference from the mean, by analogy the autocovariance of a time series is defined as the average product of differences at time t and $t + k$

$$c_k = \frac{1}{n}\sum_{t=1}^{n-k}(z_t - \bar{z})(z_{t+k} - \bar{z}), \qquad k = 0,1,2,\ldots,K \tag{2.13}$$

where c_k is the autocovariance coefficient at lag k, and c_0 is the variance. By combining Equations 2.12 and 2.13, the autocorrelation at lag k can be written in terms of the autocovariance:

$$\hat{\rho}_k = \frac{c_k}{c_0} \tag{2.14}$$

So that the estimator is asymptotically unbiased, where $\hat{\rho}_k$ is the estimation autocorrelation function.

The autocovariance function is sometimes computed with the alternative equation

$$c_k = \frac{1}{(n-k)} \sum_{t=1}^{n-k} (z_t - \bar{z})(z_{t+k} - \bar{z}), \qquad k = 0,1,2,\dots,K \tag{2.15}$$

The autocovariance function given by Equation 2.15 has a lower bias than the autocovariance function given by Equation 2.13 (Jenkins and Watts, 1968).

2.1.10 Standard errors of autocorrelation estimates

Bartlett, in 1946, derived an approximation expression for the variance of the estimated autocorrelation coefficient of a stationary normal process (Bartlett, 1946)

$$var\,[\rho_k] \simeq \frac{1}{n} \sum_{i=-\infty}^{\infty} \{\rho_i^2 + \rho_{i+k}\rho_{i-k} - 4\rho_k\rho_i\rho_{i-k} + 2\rho_i^2\rho_k^2\} \tag{2.16}$$

For a process with $\rho_i = 0$ this approximation simplifies to Equation 2.17, since, for $i > k-1$, all terms except the first appearing in the right-hand side of Equation 2.16 are zero. Then the variance of autocorrelation $var\,[\rho_k]$ is calculated as follows

$$var\,[\rho_k] \simeq \frac{1}{n}\left\{1 + 2\sum_{i=1}^{k-1} \rho_i^2\right\}. \tag{2.17}$$

A similar approximate expression for the covariance between the estimated correlation ρ_k and ρ_{k+t} at two different lags k and $k+t$ have been given by Bartlett (1946).

$$cov[\rho_k, \rho_{k+t}] \simeq \frac{1}{n} \sum_{i=-\infty}^{\infty} \rho_i \rho_{i+t} \tag{2.18}$$

Equation 2.18 is required in the interpretation of individual autocorrelations because large covariances can exist between neighbouring values.

The standard errors from Equation 2.17 for estimated autocorrelations ρ_k given by

$$SE[\rho_k] \simeq \sqrt{var\,[\rho_k]} \qquad k > 0 \tag{2.19}$$

2.1.11 The partial autocorrelation function

The partial autocorrelation measures the correlation between z_t and z_{t-k} that remains when the influences of $z_{t-1}, z_{t-2}, \dots, z_{t-k+1}$ on z_t and z_{t-k} have been eliminated**.** Consider the ith order of the correlation between z_t and z_{t-k},

$$z_t = \phi_{11} z_{t-1} + e_1$$

$$z_t = \phi_{21} z_{t-1} + \phi_{22} z_{t-2} + e_2$$

$$z_t = \phi_{31} z_{t-1} + \phi_{32} z_{t-2} + \phi_{33} z_{t-2} + e_3$$

$$z_t = \phi_{k1} z_{t-1} + \phi_{k2} z_{t-2} + \dots + \phi_{kk} z_{t-k} + e_k \tag{2.20}$$

where the sequence $\phi_{11}, \phi_{22}, \phi_{33}, \dots, \phi_{kk}$ denotes partial autocorrelations and e_t is an error term with mean zero and uncorrelated with z_{t-j} for $j = 1, 2, 3, \dots, k$.

Multiplying Equation 2.20 by z_{t-k} and then taking expected values, gives the autocovariance function

$$\gamma_j = \phi_{k1}\gamma_{j-1} + \dots + \phi_{k(k-1)}\gamma_{j-k+1} + \phi_{kk}\gamma_{j-k}, \tag{2.21}$$

hence,

$$\rho_j = \phi_{k1}\rho_{j-1} + \dots + \phi_{k(k-1)}\rho_{j-k+1} + \phi_{kk}\rho_{j-k} \qquad j = 1,2,\dots,k \tag{2.22}$$

where ρ_j is an autocorrelation function.

Substituting $j = 1,2,\dots,k$ into Equation 2.23 gives a set of linear equations for $\phi_1, \phi_2, \dots, \phi_k$ in terms of $\rho_1, \rho_2, \dots, \rho_k$, these are

$$\begin{array}{llll}
\rho_1 = \phi_{k1} & + \phi_{k2}\rho_1 & + \dots + \phi_{kk}\rho_{k-1} \\
\rho_2 = \phi_{k1}\rho_1 & + \phi_{k2} & + \dots + \phi_{kk}\rho_{k-2} \\
\quad\vdots \qquad \vdots & \quad \vdots & \quad\ \dots\ \vdots \\
\rho_k = \phi_{k1}\rho_{k-1} & + \phi_{k2}\rho_{k-2} & + \dots + \phi_{kk}
\end{array} \tag{2.23}$$

These are usually called the Yule-Walker equations (Yule, 1927, Walker, 1931).

The Yule-Walker Equations 2.23, may be written

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix} \tag{2.24}$$

Using Cramer's rule successively for $k = 1, 2, 3, \ldots,$ gives

$$\phi_{11} = \rho_1$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} = \frac{\rho_3 - (\phi_{21}\rho_2 + \phi_{22}\rho_1)}{1 - (\phi_{21}\rho_1 + \phi_{22}\rho_2)}$$

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{vmatrix}} \tag{2.25}$$

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_j}, \qquad k = 3,4, \ldots, \tag{2.26}$$

where

$$\phi_{k,j} = \phi_{k-1,j} - \phi_{kk}\phi_{k-1,k-j}\,, \quad k = 2, \ldots\,, \qquad j = 1, 2, \ldots, k-1$$

For an autoregressive process of order p, the partial autocorrelation function ϕ_{kk} will be nonzero for k less than or equal to p, and zero for k greater than p.

2.1.12 Estimation of the partial autocorrelation function

An estimated partial autocorrelation function (PACF) is similar to an estimated autocorrelation function (ACF). An estimated PACF is also a graphical representation of the statistical relationship between sets of order pairs (z_t, z_{t-k}) drawn from a single time series (Pankratz, 1983). The main idea of the PACF is to measure how z_t and z_{t+k} are related. To estimate the PACF, consider the regression relationship between z_{t+1} and the preceding value z_t:

$$z_{t+1} = \phi_{11} z_t + a_{t+1}$$

where ϕ_{11} is a partial autocorrelation coefficient to be estimated, and $k = 1$ and a_{t+1} is the error term.

When $k = 2$, then

$$z_{t+2} = \phi_{21} z_{t+1} + \phi_{22} z_t + a_{t+2}$$

When $k = 3$, then

$$z_{t+3} = \phi_{31} z_{t+2} + \phi_{32} z_{t+1} + \phi_{33} z_t + a_{t+3}$$

Thus, the partial autocorrelation function $\hat{\phi}_{kk}$ can be obtained by substituting ρ_i by $\hat{\rho}_i$ in Equation 2.25. Instead of calculating the complicated determinants for lag k in Equation 2.25, a recursive method starting with $\hat{\phi}_{11} = \hat{\rho}_1$ for computing $\hat{\phi}_{kk}$ has been given by Durbin (1960) as follows:

$$\hat{\phi}_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^{k} \hat{\phi}_{kj} \hat{\rho}_{k+1-j}}{1 - \sum_{j=1}^{k} \hat{\phi}_{kj} \hat{\rho}_j}, \tag{2.27}$$

where

$$\hat{\phi}_{k+1,j} = \hat{\phi}_{kj} - \hat{\phi}_{k+1,k+1} \hat{\phi}_{k,k+1-j}, \qquad j = 1,2,\dots,k \tag{2.28}$$

The method holds also for calculating the theoretical PACF ϕ_{kk}.

2.1.13 Standard errors of partial autocorrelation estimates

It was shown by Quenouille (1949), Jenkins (1954, 1956), and Daniels (1956) that given the hypothesis that the process is autoregressive of order p, the estimated partial autocorrelation of order $p+1$, and higher are approximately independently distributed with variance.

$$var[\hat{\phi}_{kk}] \cong \frac{1}{n} \qquad k \geq p+1$$

The standard error (S.E.) of the estimated partial autocorrelation $\hat{\phi}_{kk}$ is

$$S.E.[\hat{\phi}_{kk}] \simeq \frac{1}{\sqrt{n}} \qquad k \geq p+1 \tag{2.29}$$

2.1.14 White noise

A time series z_t is called a white noise process if $\{z_t\}$ is a sequence of independent and identically distributed random variables with normal distribution with constant mean $E(e_t) = \mu_e$ (it is usually assumed that $E(e_t) = 0$), constant variance $var(e_t) = \sigma_e^2$, and $y_k = cov(e_t, e_{t+k}) = 0$ for all $k \neq 0$. A white noise process $\{e_t\}$ is stationary with

the autocovariance function

$$\gamma_k = \begin{cases} \sigma_e^2, & k = 0, \\ 0, & k \neq 0, \end{cases} \tag{2.30}$$

the autocorrelation function

$$\rho_k = \begin{cases} 1, & k = 0, \\ 0, & k \neq 0, \end{cases} \tag{2.31}$$

and the partial autocorrelation function

$$\phi_{kk} = \begin{cases} 1, & k = 0, \\ 0, & k \neq 0, \end{cases} \tag{2.32}$$

2.1.15 Backshift operator

The backshift operator is a useful notion in time series analysis. For a time series $\{z_t\}_{t=0}^{n}$, the backshift operator can be used to model the data series and investigate the characteristics of $\{z_t\}_{t=0}^{n}$. The basic rules of the backshift operator are

1. $B\mathrm{z_t} = \mathrm{z_{t-1}}$

Example:

$$\nabla \mathrm{z_t} = \mathrm{z_t} - \mathrm{z_{t-1}} = \mathrm{z_t} - B\mathrm{z_t} = \mathrm{z_t} - B\mathrm{z_t} = (1 - B)\mathrm{z_t}.$$

2.

$$\frac{1}{1 - aB} = 1 + aB + a^2B^2 + a^3B^3 + \cdots, \qquad \text{if } |a| < 1$$

2.2 Linear process and invertibility

2.2.1 The linear process

The class of linear time series models, which includes the class of autoregressive moving-average (ARMA) models, provides a general framework for studying stationary processes. The time series $\{\mathrm{z_t}\}$ is a linear process if it has the representation

$$\mathrm{z_t} = e_\mathrm{t} + \sum_{j=0}^{\infty} \psi_j \, e_{t-j} \tag{2.33}$$

for all t, where $\{e_\mathrm{t}\}$ is a white noise of mean zero and constant variance σ_e^2 and $\{\psi_\mathrm{j}\}$ is a sequence of constants $\sum_{j=-\infty}^{\infty} |\psi_\mathrm{j}| < \infty$. The linear process $\mathrm{z_t}$ can be represented as a weighted sum of present and past values of the white noise process. The linear process was developed by several researchers, such as, Walker (1931), Slutsky (1937), Bartlett (1946), Doob (1953), Grenander and Rosenblatt (1957), and Hannan (1970).

For a linear time series defined by Equation 2.33, the dynamic structure of $\mathrm{z_t}$ is governed by the coefficients ψ_j, which are called the ψ-weights of $\mathrm{z_t}$ in the time series. If $\mathrm{z_t}$ is weakly stationary, it is possible to obtain its mean and variance easily by using the independence of $\{e_t\}$ as

$$E[\mathrm{z_t}] = 0, \tag{2.34}$$

$$var(\mathrm{z_t}) = \sigma_a^2 \sum_{j=1}^{\infty} \psi_j^2, \tag{2.35}$$

where σ_e^2 is the variance of e_t. Because $var(\mathrm{z_t}) < \infty$, $\{\psi_j^2\}$ must be a convergent sequence, that is $\psi_j^2 \to 0$ as $j \to \infty$. Consequently, for a stationary series, the impact of the remote white noise e_{t-j} on the return $\mathrm{z_t}$vanishes as j increases.

The lag-k autocovariance of $\mathrm{z_t}$ is

$$\gamma_k = cov(\mathrm{z_t}, z_{t-k}) = E\left[\left(\sum_{i=0}^{\infty} \psi_i e_{t-i}\right)\left(\sum_{j=0}^{\infty} \psi_j e_{t-k-j}\right)\right]$$

$$= E\left[\sum_{i,j=0}^{\infty} \psi_i \psi_j e_{t-i} e_{t-k-j}\right]$$

$$= \sum_{j=0}^{\infty} \psi_{j+k} \psi_j E\left[e_{t-k-j}^2\right]$$

where $E\left[e_{t-k-j}^2\right] = \sigma_e^2$, this gives

$$= \sigma_e^2 \sum_{j=0}^{\infty} \psi_{j+k} \psi_j. \tag{2.36}$$

Consequently, ψ-weights are related to the autocorrelations of $\mathrm{z_t}$ as follows

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\sum_{j=0}^{\infty} \psi_{j+k}\psi_j}{1 + \sum_{j=1}^{\infty} \psi_j^2}, k \geq 0, \tag{2.37}$$

where $\psi_0 = 1$. For a weakly stationary time series, $\psi_j \to 0$ as $j \to \infty$ and, hence, ρ_k converges to zero as k increases (Box, Jenkins, Reinsel and Ljung, 2015).

The Equation 2.33 implies that z_t can be written alternatively as a weighted sum of past values of z_t, plus an added white noise e_t, that is

$$\begin{aligned} z_t &= \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} + e_t \\ &= \sum_{j=1}^{p} \phi_j z_{t-j} + e_t \end{aligned} \tag{2.38}$$

The alternative form Equation 2.38 may be thought of as one where the current deviation z_t, from the level μ, is regressed on past deviations $z_{t-1}, z_{t-2}, \ldots$ of the process, where ϕ are weights of z_t.

2.2.2 Invertibility

If an MA(q) process can be represented by an AR(∞) process, then the process is said to be invertible (Box and Jenkins, 1970). Invertibility of a MA(q) process requires that all the roots of the polynomial $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q = 0$, lie outside the unit circle. The invertibility condition was investigated by Granger and Andersen in 1978.

The invertibility of an MA(1) process

To illustrate the invertibility condition, consider the first-order moving average process MA(1) if $|\theta| < 1$

$$z_t = (1 - \theta B)e_t \tag{2.39}$$

or

$$e_t = (1 - \theta B)^{-1} z_t \tag{2.40}$$

By pre-multiplying both sides of (2.39) by $(1 + \theta B + \theta^2 B^2 + \cdots + \theta^k B^k)$, gives

$$(1+\theta B+\theta^2B^2+\cdots+\theta^kB^k)z_t=(1+\theta B+\theta^2B^2+\cdots+\theta^kB^k)(1-\theta B)e_t$$

$$(1+\theta B+\cdots+\theta^kB^k)z_t=(1+\theta B+\cdots+\theta^kB^k-\theta B-\cdots-\theta^kB^k-\theta^{k+1}B^{k+1})e_t$$

then

$$(1+\theta B+\theta^2B^2+\cdots+\theta^kB^k)z_t=(1-\theta^{k+1}B^{k+1})e_t$$

By using the backshift operator rule

$$B^k(\mathrm{z_t})=\mathrm{z}_{\mathrm{t}-k}$$

hence,

$$z_t+\theta z_{t-1}+\theta^2z_{t-2}+\cdots+\theta^kz_{t-k}=e_t-\theta^{k+1}e_{t-1-k}$$

Thus,

$$z_t=-\theta z_{t-1}-\theta^2z_{t-2}-\cdots-\theta^kz_{t-k}+e_t-\theta^{k+1}e_{t-1-k} \tag{2.41}$$

If $|\theta|<1$, then the last term in this expression tends to zero as $k\to\infty$, and the infinite series can be written as

$$z_t=e_t+\sum_{i=1}^{\infty}(-\theta)^iz_{t-i} \tag{2.42}$$

So, $|\theta|<1$ is a sufficient condition for an MA(1) model to be invertible.

In general, an MA(q) model is invertible if all the roots of MA(q) polynomial $\theta(B)=1-\theta_1B-\cdots-\theta_qB^q=0$, lie outside the unit circle (Box and Jenkins, 1970).

2.3 Autoregressive process

The general form for an autoregressive process of order p, an AR(p) process is

$$z_t=\phi_1z_{t-1}+\phi_2z_{t-2}+\cdots+\phi_pz_{t-p}+e_t \tag{2.43}$$

where the current value of the process is expressed as a weighted sum of previous values plus a white noise term. Equation 2.43 can be written as

$$\left(1-\phi_1B-\cdots-\phi_pB^p\right)z_t=\phi(B)z_t=e_t$$

The Equation 2.43 must satisfy certain conditions for the process to be stationary.

2.3.1 The first-order autoregressive process

The first-order autoregressive process AR(1), can be written as (Guy and Donald 1949),

$$(1 - \phi_1 B)z_t = e_t$$

which may also be written as

$$z_t = (1 - \phi_1 B)^{-1} e_t, \qquad |\phi_1| < 1$$

Use of lag operator rule 2 gives

$$z_t = \sum_{j=0}^{\infty} \phi_1^j e_{t-j}$$

providing that the infinite series on the right converges in an appropriate sense.

Hence,

$$\psi(B) = (1 - \phi_1 B)^{-1} = \sum_{j=0}^{\infty} \phi_1^j B^j, \tag{2.44}$$

or equivalently that

$$\sum_{j=0}^{\infty} |\phi_1|^j < \infty.$$

From Equation 2.44 an AR(1) process must satisfy the condition $|\phi_1| < 1$ to ensure stationarity. Since the root of $1 - \phi_1 B = 0$, this condition is equivalent to saying that the root of $1 - \phi_1 B = 0$ must lie outside the unit circle.

The autocorrelation of an AR(1) process

$$z_t = \phi_1 z_{t-1} + e_t$$

Multiplying by z_{t-k} on both sides gives

$$z_{t-k}\, z_t = \phi_1 z_{t-k} z_{t-1} + z_{t-k} e_t$$

and taking the expectation on both sides, where ϕ_1 is constant, gives

$$E[z_{t-k}\, z_t] = \phi_1 E[z_{t-k} z_{t-1}] + E[z_{t-k} e_t]$$

, because e_t and z_{t-k} are independent, it follows that for $k \geq 1$

$$cov\ (z_{t-k}, z_t) = \phi_1 cov(z_{t-k}, z_{t-1})$$

$$\gamma_k = \phi_1 \gamma_{k-1}, \qquad \text{for } k \geq 1 \tag{2.45}$$

The variance of an AR(1) process is given by

$$\gamma_0 = \sigma_z^2 = \frac{\sigma_e^2}{1 - \phi_1^2}, \qquad |\phi_1| < 1 \tag{2.46}$$

and the autocorrelation function is given by

$$\rho_k = \phi_1 \rho_{k-1} = \phi_1^k, \qquad \text{for } k \geq 1 \tag{2.47}$$

where $\rho_0 = 1$. Hence when $|\phi_1| < 1$ and the process is stationary, the autocorrelation exponentially decays in one of two forms depending on the sign of ϕ_1. If $0 < \phi_1 < 1$, then all autocorrelations are positive; if $-1 < \phi_1 < 0$, then the sign of the autocorrelations shows an alternating pattern, beginning with a negative value.

The partial autocorrelation of an AR(1) process

$$\phi_{kk} = \begin{cases} \rho_1 = \phi_1, & k = 1, \\ 0, & k \geq 2. \end{cases} \tag{2.48}$$

Hence, the partial autocorrelation of the AR(1) process shows a positive or negative spike at lag 1 depending on the sign of ϕ_1 and then cuts off, shown in Figure 2.1.

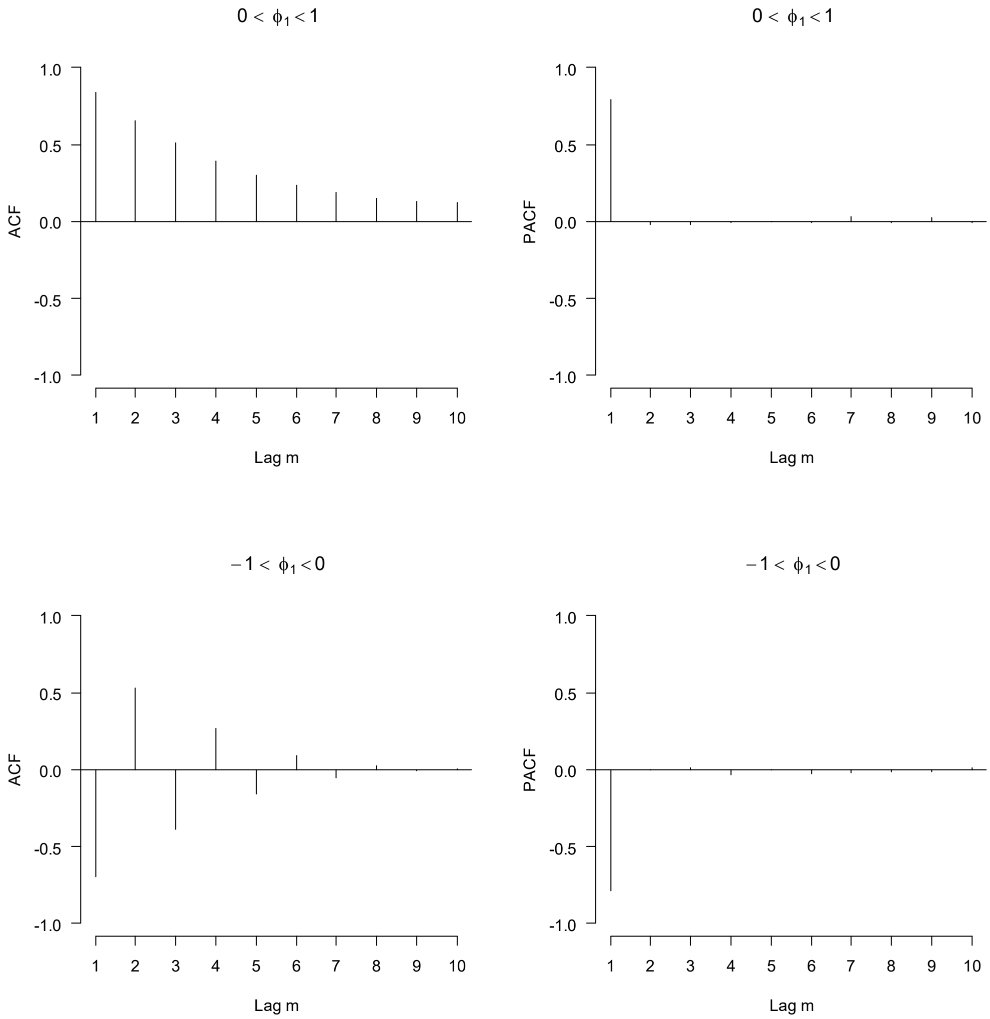

Figure 2.1 ACF and PACF of the AR(1) process.

2.3.2 The second-order autoregressive process

The second-order autoregressive process AR(2) may be written

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + e_t \tag{2.49}$$

The AR(2) process, as a finite autoregressive process, is always invertible if stationary. To be stationary the roots of $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 = 0$ must lie outside the unit circle. Consider the second-order polynomial equation (Wei, 2006)

$$1 - \phi_1 x - \phi_2 x^2 = 0$$

where the solution of this equation is

$$A_1 = \frac{-\phi_1 + \sqrt{\phi_1^2 + 4\,\phi_2}}{2\phi_2}, \qquad A_2 = \frac{-\phi_1 - \sqrt{\phi_1^2 + 4\,\phi_2}}{2\phi_2}$$

Taking the reciprocal both sides,

$$\frac{1}{A_1} = \frac{2\phi_2}{-\phi_1 + \sqrt{\phi_1^2 + 4\,\phi_2}} = \frac{2\phi_2}{-\phi_1 + \sqrt{\phi_1^2 + 4\,\phi_2}}\left[\frac{-\phi_1 - \sqrt{\phi_1^2 + 4\,\phi_2}}{-\phi_1 - \sqrt{\phi_1^2 + 4\,\phi_2}}\right]$$

$$= \frac{\phi_1 + \sqrt{\phi_1^2 + 4\,\phi_2}}{2}$$

Similarly

$$\frac{1}{A_2} = \frac{\phi_1 - \sqrt{\phi_1^2 + 4\,\phi_2}}{2}$$

For real roots, it is required that $\phi_1^2 + 4\,\phi_2 \geq 0$, which is

$$-1 < \frac{\phi_1 - \sqrt{\phi_1^2 + 4\,\phi_2}}{2} \leq \frac{\phi_1 + \sqrt{\phi_1^2 + 4\,\phi_2}}{2} < 1$$

Consider the left hand side

$$-1 < \frac{\phi_1 - \sqrt{\phi_1^2 + 4\,\phi_2}}{2}$$

$$\Leftrightarrow -2 < \phi_1 - \sqrt{\phi_1^2 + 4\,\phi_2}$$

$$\Leftrightarrow \sqrt{\phi_1^2 + 4\,\phi_2} < 2 + \phi_1$$

$$\Leftrightarrow \phi_1^2 + 4\,\phi_2 < \phi_1^2 + 4\phi_1 + 4$$

$$\Leftrightarrow \ \phi_2 < \phi_1 + 1$$

or

$$\phi_2 - \phi_1 < 1$$

Now consider the right hand side

$$\phi_2 + \phi_1 < 1$$

For complex roots, $\phi_1^2 + 4\phi_2 < 0$. Here A_1 and A_2 will be complex conjugates and $\left|\frac{1}{A_1}\right| = \left|\frac{1}{A_2}\right| < 1$ if and only if $\left|\frac{1}{A_1}\right|^2 < 1$. But

$$\left|\frac{1}{A_1}\right|^2 = \frac{[\phi_1^2 + (-\phi_1^2 - 4\phi_2)]}{4} = -\phi_2$$

so that $\phi_2 > -1$.

Thus, the stationarity condition of the AR(2) model is given by the following triangular region,

$$\begin{aligned} \phi_2 + \phi_1 &< 1 \\ \phi_2 - \phi_1 &< 1 \\ -1 < \phi_2 &< 1 \end{aligned} \tag{2.50}$$

The autocorrelation function of an AR(2) process

The autocorrelation function of an AR(2) process can be obtained by multiplying z_{t-k} on both sides of the AR(2) process in Equation 2.49

$$z_{t-k}z_t = \phi_1 z_{t-k}z_{t-1} + \phi_2 z_{t-k}z_{t-2} + z_{t-k}e_t$$

and then taking the expectation on both sides, where ϕ_1, ϕ_2 are constants gives

$$E[z_{t-k}z_t] = \phi_1 E[z_{t-k}z_{t-1}] + \phi_2 E[z_{t-k}z_{t-2}] + E[z_{t-k}e_t]$$

because e_t and z_{t-k} are independent, it follows that for $k \geq 1$

$$cov\ (z_{t-k}, z_t) = \phi_1 cov(z_{t-k}, z_{t-1}) + \phi_2 cov(z_{t-k}, z_{t-2})$$

$$\gamma_k = \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2}, \qquad k \geq 1 \tag{2.51}$$

The variance of AR(2) is

$$\gamma_0 = \sigma_{\tilde{z}}^2 = \frac{(1-\phi_2)\sigma_e^2}{(1+\phi_2)(1-\phi_2)^2 - \phi_1^2} \tag{2.52}$$

The autocorrelation function satisfies the second-order difference equation

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \qquad k \geq 1 \tag{2.53}$$

The stationary conditions for an AR(2) process are, in the cases of $k = 1$ and 2, given by

$$\rho_1 = \phi_1 + \phi_2 \rho_1$$

$$\rho_2 = \phi_1 \rho_1 + \phi_2,$$

which implies that

$$\left.\begin{aligned} \rho_1 &= \frac{\phi_1}{1-\phi_2} \\ \rho_2 &= \phi_2 + \frac{\phi_1^2}{1-\phi_2} \end{aligned}\right\} \tag{2.54}$$

Thus, ρ_1 and ρ_2 must lie in the region,

$$-1 < \rho_1 < 1$$

$$-1 < \rho_2 < 1$$

$$\rho_1^2 < \frac{1}{2}(\rho_2 + 1)$$

Thus, the ACF of the second-order autoregressive process will decay exponentially if the roots of $(1 - \phi_1 B - \phi_2 B^2) = 0$ are real, and will follow a damped sine wave if the roots of $1 - \phi_1 B - \phi_2 B^2 = 0$ are complex.

The partial autocorrelation of an AR(2) process

For the AR(2) process, because

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}$$

For $k \geq 1$, then

$$\phi_{11} = \rho_1 = \frac{\phi_1}{1 - \phi_2}$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

$$= \frac{\left(\frac{\phi_1^2 + \phi_2 - \phi_2^2}{1 - \phi_2}\right) - \left(\frac{\phi_1}{1 - \phi_2}\right)^2}{1 - \left(\frac{\phi_1}{1 - \phi_2}\right)^2}$$

$$= \frac{\phi_2[(1 - \phi_2)^2 - \phi_1^2]}{(1 - \phi_2)^2 - \phi_1^2} = \phi_2.$$

Hence, the partial autocorrelation function of an AR(2) process cuts off after lag 2.

2.3.3 The general pth-order autoregressive AR(p) process

The general AR(p) process may be written as

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} + e_t$$

or

$$\left(1 - \phi_1 B - \phi_2 B - \cdots - \phi_p B^p\right) z_t = \phi(B) z_t = e_t$$

This can then put in the form

$$\phi(B) = (1 - G_1 B)(1 - G_2 B) \cdots \left(1 - G_p B\right)$$

where $G_1^{-1}, G_2^{-1}, \cdots, G_p^{-1}$ are the root of $\phi(B) = 0$.

Now consider $\phi^{-1}(B)$ and using partial fractions, gives

$$\phi^{-1}(B) = \frac{1}{(1 - G_1 B)(1 - G_2 B) \cdots \left(1 - G_p B\right)}$$

$$= \frac{A_1}{(1 - G_1 B)} + \frac{A_2}{(1 - G_2 B)} + \cdots + \frac{A_p}{\left(1 - G_p B\right)}$$

$$= \sum_{j=1}^{p} \frac{A_j}{(1 - G_j B)}.$$

hence,

$$z_t = \phi^{-1}(B) e_t = \sum_{j=1}^{p} \frac{A_j}{(1 - G_j B)} e_t.$$

Thus, if $\psi(B) = \phi^{-1}(B)$ is to be a convergent series for $|\phi| < 1$, that is, if the weights

$$\psi_j = \sum_{i=1}^{p} A_i\, G_i^j,$$

are to be absolutely summable so that the AR(p) will represent a stationary process, it is required that $G_i < 1$, for $i = 1, 2, \dots, p$. Equivalently, the roots of $\phi(B) = 0$ must lie outside the unit circle, Box and Jenkins (1970).

The autocovariance function of the general pth-order autoregressive AR(p) process is given by Equation 2.43

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} + e_t$$

By multiplying z_{t-k} on both sides of the equation, gives

$$z_{t-k}\, z_t = \phi_1 z_{t-k}\, z_{t-1} + \phi_2 z_{t-k}\, z_{t-2} + \cdots + \phi_p z_{t-k}\, z_{t-p} + e_t$$

and taking the expectation on both sides, gives

$$E[z_{t-k}\, z_t] = \phi_1 E[z_{t-k}\, z_{t-1}] + \phi_2 E[z_{t-k}\, z_{t-2}] + \cdots + \phi_p E[z_{t-k}\, z_{t-p}] + E[z_{t-k} e_t]$$

because e_t and z_{t-k} are independent, it follows that

$$cov(z_{t-k}, z_t) = \phi_1 cov(z_{t-k}, z_{t-1}) + \phi_2 cov(z_{t-k}, z_{t-2}) + \cdots + \phi_p cov(z_{t-k}, z_{t-p})$$

Hence, the autocovariance function of the general pth- order autoregressive AR(p) process is

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \cdots + \phi_p \gamma_{k-p}, \qquad k > 0 \qquad (2.55)$$

The autocorrelation function of pth-order autoregressive AR(p)

The autocorrelation function of the general pth-order autoregressive AR(p) process will consist of a mixture of exponential decays and damped sine or cosine waves, (Box and Jenkins, 1970). Damped sine or cosine waves appear if some of the roots are complex. The autocorrelation function of autoregressive AR(p) can be found by solving a set of difference equations called the Yule-Walker equations given by

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p} \qquad k > 0 \tag{2.56}$$

The partial autocorrelation function of pth-order autoregressive AR(p)

For the partial autocorrelation function of the general pth-order autoregressive AR(p), the PACF ϕ_{kk} will vanish after lag p, Box and Jenkins (1970)

2.4 Moving average processes

The general form for a moving average process of order q, a MA(q) process, is

$$\begin{aligned} z_t &= e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} \\ &= \left(1 - \theta_1 B - \cdots - \theta_q B^q\right) e_t \\ &= \theta(B) e_t \end{aligned} \tag{2.57}$$

where $\theta(B) = \left(1 - \theta_1 B - \cdots - \theta_q B^q\right)$.

This moving average process is invertible if the roots of $\theta(B) = 0$ lie outside of the unit circle.

Thus, a moving-average model is expressed as the current value of the series against current and previous white noise. Moving average processes are useful in describing phenomena in which events produce an immediate effect that only lasts for short periods of time. This was first studied by Slutzky (1937).

2.4.1 The first-order moving average process

When $\theta(B) = (1 - \theta_1 B)$, then the first-order moving average MA(1) process is

$$z_t = e_t - \theta_1 e_{t-1} \tag{2.58}$$

$$= (1 - \theta_1 B)e_t$$

where $\{e_t\}$ is a zero mean white noise process with constant variance σ_e^2.

Autocorrelation function of the MA(1) process

To obtain autocorrelation function of the MA(1) process, it is necessary to find for $k = 1$ the mean, variance and autocovariance of the MA(1) process, from Equation 2.58.

The mean of the process

$$z_t = e_t - \theta_1 e_{t-1}$$

Taking the expectations on both sides, where θ_1 is constant

$$E[z_t] = E[e_t - \theta_1 e_{t-1}]$$

$$= 0$$

The variance of the MA(1) process

$$z_t = e_t - \theta_1 e_{t-1}$$

Taking the variance on both sides

$$var(z_t) = var(e_t - \theta_1 e_{t-1})$$

$$= \sigma_e^2 + \theta_1^2 \sigma_e^2 - 0$$

$$= (1 + \theta_1^2)\sigma_e^2$$

The first autocovariance of the MA(1) process

$$z_t = e_t - \theta_1 e_{t-1}$$

$$z_{t-k} = e_{t-1} - \theta_1 e_{t-1-k}$$

and taking the expectations, where θ_1 is constant

$$E[z_t z_{t-k}] = E[(e_t - \theta_1 e_{t-1})(e_{t-1} - \theta_1 e_{t-1-k})]$$
$$= -\theta_1 \sigma_e^2$$

Hence

$$\gamma_1 = E[z_t z_{t-k}] = -\theta_1 \sigma_e^2$$

Thus, the autocorrelation function of MA(1) process is given by

$$\rho_1 = \frac{-\theta_1}{1+\theta_1^2} \tag{2.59}$$

$$\rho_k = 0, \quad k > 1$$

which is cut off after lag 1, shown in Figure 2.2. Given the mean, variance and autocovariance of the MA(1) process are constants as show, this means the MA(1) process is always stationary. For the process to be invertible, the roots $1 - \theta_1 B = 0$ must lie outside the unit circle.

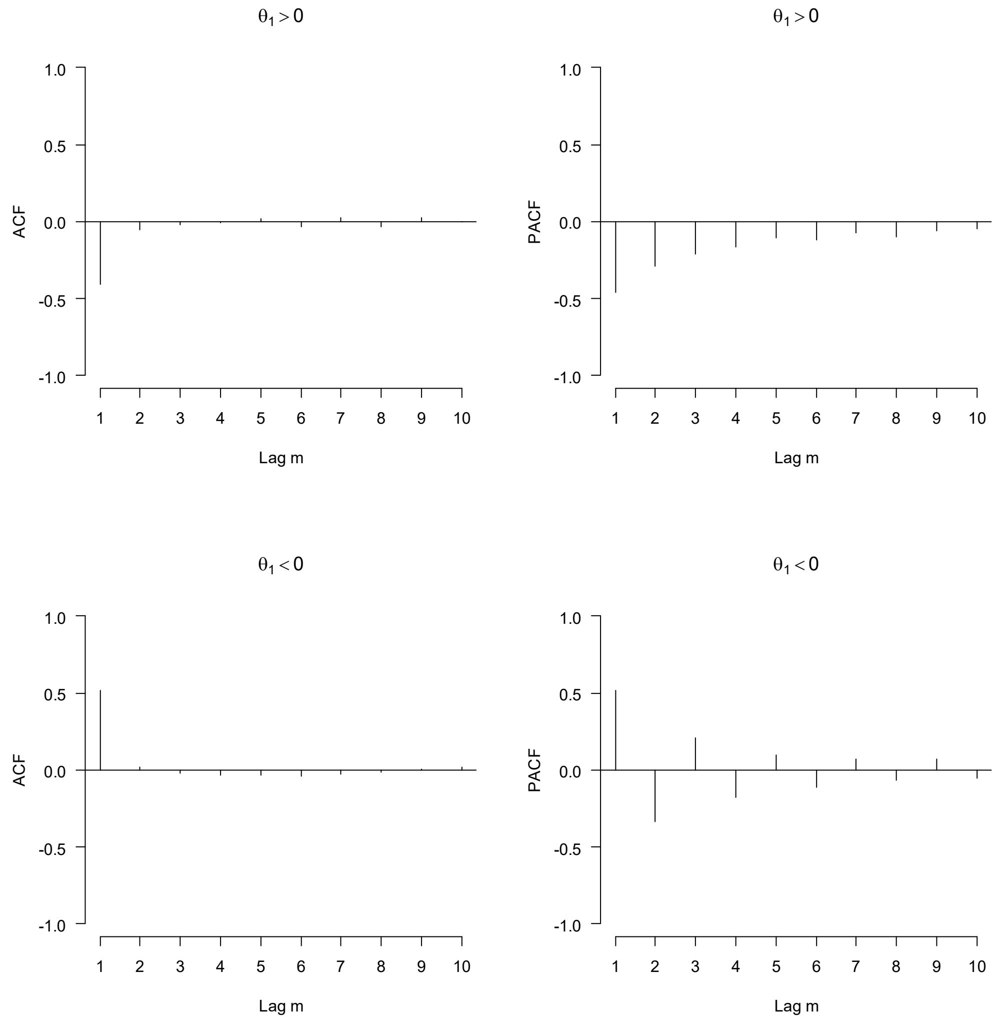

Figure 2.2 ACF and PACF of MA(1) process.

Partial autocorrelation function of the MA(1) process

The partial autocorrelation function has no cut-off, it can be shown to decay geometrically to zero. From Equations 2.25 and 2.59

$$\phi_{11} = \rho_1 = \frac{-\theta_1}{1+\theta_1^2} = \frac{-\theta_1(1-\theta_1^2)}{1-\theta_1^4}$$

$$\phi_{22} = \frac{\rho_1^2}{1-\rho_1^2} = \frac{-\theta_1^2}{1+\theta_1^2+\theta_1^4} = \frac{-\theta_1^2(1-\theta_1^2)}{1-\theta_1^6}$$

$$\phi_{33} = \frac{\rho_1^3}{1 - 2\rho_1^2} = \frac{-\theta_1^3}{1 + \theta_1^2 + \theta_1^4 + \theta_1^6} = \frac{-\theta_1^3(1 - \theta_1^2)}{1 - \theta_1^8}$$

In general

$$\phi_{kk} = \frac{(-1)^{k-1}\theta^k(1 - \theta^2)\theta_1^k(1 - \theta_1^2)}{1 - \theta_1^{2(k+1)}}. \tag{2.60}$$

Thus, $|\phi_{kk}| < \theta_1^k$, and the partial autocorrelation function is dominated by a damped exponential. If ρ_1 is positive, so that θ_1 is negative, the partial autocorrelations alternate in sign. However, if ρ_1 is negative, so that θ_1 is positive, the partial autocorrelations are negative.

2.4.2 The second-order moving average process

Invertibility conditions

The second-order moving average process is defined by

$$z_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} \tag{2.61}$$

and is stationary for all values of θ_1 and θ_2. However, it is invertible only if the roots of the characteristic equation

$$1 - \theta_1 B - \theta_2 B^2 = 0 \tag{2.62}$$

lie outside the unit circle, that is

$$\begin{gathered} \theta_2 + \theta_1 < 1 \\ \theta_2 - \theta_1 < 1 \\ -1 < \theta_2 < 1 \end{gathered} \tag{2.63}$$

Compare with the stationary conditions of the AR(2) process in Equation 2.50.

Autocorrelation function of the MA(2) process

The autocorrelation function of MA(2) process, from Equation 2.61,

$$z_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

So, the mean of the process, taking the expectations on both sides, where θ_1 and θ_2 are constants

$$E[z_t] = E[e_t] - \theta_1 E[e_{t-1}] - \theta_2 E[e_{t-2}]$$

$$= 0$$

The variance of the MA(2) process

$$z_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

Taking the variance on both sides

$$Var(z_t) = var(e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2})$$

where θ_1 and θ_2 are constants

$$= \sigma_e^2 + \theta_1^2 \sigma_e^2 + \theta_2^2 \sigma_e^2$$

$$= (1 + \theta_1^2 + \theta_2^2)\sigma_e^2$$

The autocovariance of the MA(2) process, from Equation 2.61

$$z_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

hence

$$\gamma_1 = (\theta_1 + \theta_2 \theta_1)\sigma_e^2$$

In general,

$$z_t z_{t-k} = (e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2})(e_{t-k} - \theta_1 e_{t-1-k} - \theta_2 e_{t-2-k})$$

Taking the expectation for both sides, then

$$E[z_t z_{t-k}] = E[(e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2})(e_{t-k} - \theta_1 e_{t-1-k} - \theta_2 e_{t-2-k})]$$

$$= E\left[\left(\sum_{i=0}^{2} \theta_i e_{t-i}\right)\left(\sum_{j=0}^{2} \theta_j e_{t-k-j}\right)\right]$$

$$= E\left[\sum_{i,j=0}^{2} \theta_i \theta_j e_{t-i}\, e_{t-k-j}\right]$$

hence,

$$\gamma_k = \sigma_e^2 \sum_{k,j=0}^{2} \theta_j \theta_{j+k} \qquad (2.64)$$

Partial autocorrelation function of the MA(2) process

The exact expression for the partial autocorrelation function of an MA(2) process is complicated, but it is dominated by the sum of two exponentials, if the roots of the characteristic Equation 2.62 are real; and by a damped sine wave, if the roots of Equation 2.62 are complex. Thus, it behaves like the autocorrelation of an AR(2) process. The autocorrelation functions (left-hand curves) and partial autocorrelation functions (right-hand curves).

2.4.3 The general qth-order moving average MA(q) process

The moving average model of order q, the MA(q) process, is given by

$$z_t = e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}$$
$$= \left(1 - \theta_1 B - \cdots - \theta_q B^q\right) e_t$$

where, as usual, $\{e_t\}$ is a zero mean white noise process with constant variance σ_e^2. This can be written

$$z_t = \theta(B) e_t \qquad (2.65)$$

where $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$ is the MA(q) operator.

The invertibility condition for higher order MA processes may be obtained by writing Equation 2.65 as

$$e_t = \theta^{-1}(B) z_t$$

Hence, if

$$\theta(B) = \prod_{j=1}^{q} \left(1 - H_j B\right)$$

where $1/H_j$, $j = 1,2, \dots, q$, are the roots of $\theta(B)$,

then

$$\theta(B) = (1 - H_1 B)(1 - H_2 B) \dots \left(1 - H_q B\right)$$

$$\theta^{-1}(B) = \frac{1}{(1 - H_1 B)(1 - H_2 B) \dots \left(1 - H_q B\right)}$$

Using partial fractions, such that

$$\pi(B) = \theta^{-1}(B) = \sum_{j=1}^{q} \frac{M_j}{\left(1 - H_j B\right)}$$

If H_j are all distinct, there exist M_j, which converges, if $\left|H_j\right| < 1$, when $j = 1,2, \dots, q$. Since the root of $\theta(B) = 0$ are H_j^{-1}, it follows that the invertibility condition for a MA(q) process is that the roots of the characteristic equation

$$\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q = 0 \tag{2.66}$$

lie outside the unit circle.

Autocorrelation function

The MA(q) process is given by

$$z_t = e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}$$

Taking the variance on both sides

$$var(z_t) = var(e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q})$$

where $\theta_1, \theta_2, \dots, \theta_q$ are constants

$$\gamma_0 = \left(1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2\right)\sigma_e^2 \tag{2.67}$$

Hence, the autocovariance function of a MA(q) process

$$z_t = e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}$$

$$z_{t-k} = e_{t-q} - \theta_1 e_{t-k-1} - \cdots - \theta_q e_{t-k-q}$$

By multiplying z_t and z_{t-k}, and taking the expectation, then

$$\gamma_k = \begin{cases} (-\theta_k + \theta_1\theta_{k+1} + \theta_2\theta_{k+2} + \cdots + \theta_{q-k}\theta_q)\sigma_e^2 & k = 1,2,\dots,q \\ 0 & k > q \end{cases} \tag{2.68}$$

Thus, the autocorrelation function is

$$\rho_k = \begin{cases} \dfrac{-\theta_k + \theta_1\theta_{k+1} + \theta_2\theta_{k+2} + \cdots + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2} & k = 1,2,\dots,q \\ 0, & k > q \end{cases} \tag{2.69}$$

The autocorrelation function of a moving average process has a cut-off at lag q.

Partial autocorrelation function

The partial autocorrelation function of the MA(q) process tails off as a mixture of exponential decays and/or damped sine waves depending on the nature of the roots of $(1 - \theta_1 B - \cdots - \theta_q B^q) = 0$. The partial autocorrelation function will contain damped sine waves if the roots of characteristic $(1 - \theta_1 B - \cdots - \theta_q B^q)$ are complex.

2.5 Mixed autoregressive-moving average processes

A large number of parameters reduce efficiency in estimation. Thus, in model building, it may be necessary to include both autoregressive and moving average terms in a model, which leads to the following useful mixed autoregressive-moving average (ARMA) model (Box and Jenkins, 1970)

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} \tag{2.70}$$

that is

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) z_t = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q) e_t$$

or

$$\phi(B)z_t = \theta(B)e_t \tag{2.71}$$

where $\phi(B)$ and $\theta(B)$ are polynomials of degrees p and q, in B.

Autocorrelation function of the ARMA(p, q) process

The autocorrelation function of the mixed process may be derived by a similar method to that used for the autoregressive process. On multiplying throughout in (2.70) by z_{t-k}

$$z_{t-k}\, z_t = \phi_1 z_{t-k}\, z_{t-1} + \phi_2 z_{t-k}\, z_{t-2} + \cdots + \phi_p z_{t-k}\, z_{t-p} + z_{t-k}\, e_t - \theta_1 z_{t-k}\, e_{t-1} - \cdots - \theta_q z_{t-k}\, e_{t-q}$$

and taking the expectation on both sides, where $\phi_1, \phi_2, \dots, \phi_p$ and $\theta_1, \theta_2, \dots, \theta_q$ are constants

$$E[z_{t-k}\, z_t] = \phi_1 E[z_{t-k}\, z_{t-1}] + \phi_2 E[z_{t-k}\, z_{t-2}] + \cdots + \phi_p E[z_{t-k}\, z_{t-p}] + E[z_{t-k}\, e_t] - \theta_1 E[z_{t-k}\, e_{t-1}] - \cdots - \theta_q E[z_{t-k}\, e_{t-q}]$$

because z_{t-k} and e_t are independent, then

$$E[z_{t-k}\, e_{t-i}] = 0, E[\, e_{t-i}] = 0 \quad \text{for k > i,}$$

hence$cov(z_{t-k}\, , z_t) = \phi_1 cov(z_{t-k}, z_{t-1}) + \phi_2 cov(z_{t-k}\, , z_{t-2}) + \cdots + \phi_p cov(z_{t-k}\, , z_{t-p}) + E[z_{t-k} e_t] - \theta_1 E[z_{t-k}\, e_{t-1}] - \cdots - \theta_q E[z_{t-k}\, e_{t-q}]$

this gives

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \cdots + \phi_p \gamma_{k-p} \tag{2.72}$$

The variance of the process, when $k = 0$, is given by

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \cdots + \phi_p \gamma_p + \sigma_e^2 - \theta_1 \gamma_1(-1) - \cdots - \theta_q \gamma_q(-q), \tag{2.73}$$

which has to be solved along with the p Equation 2.72 for $k = 1,2, \dots, p$ to obtain $\gamma_0, \gamma_1, \dots, \gamma_p$.

Hence, the autocorrelation function is

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \cdots + \phi_p\rho_{k-p} \qquad k \geq (q+1). \tag{2.74}$$

Partial autocorrelation function of the ARMA(p, q) process

The process given by Equation 2.71 may be written

$$e_t = \theta^{-1}(B)\phi(B)\tilde{z}_t$$

where $\theta^{-1}(B)$ is an infinite series in B. Hence, the partial autocorrelation function of a mixed process is infinite in extent. It behaves eventually like the partial autocorrelation function of a pure moving average process, being dominated by a mixture of damped exponentials and/or damped sine waves, depending on the order of the moving average and the values of the parameters it contains.

2.5.1 The first-order autoregressive first-order moving average process

A mixed process of considerable practical importance is the first-order autoregressive first-order moving average ARMA(1,1) process

$$z_t - \phi_1 z_{t-1} = e_t - \theta_1 e_{t-1} \tag{2.75}$$

that is

$$(1 - \phi_1 B)z_t = (1 - \theta_1 B)e_t$$

For stationarity, assuming that $|\phi_1| < 1$, and for invertibility, it requires that $|\theta_1| < 1$. When $\phi_1 = 0$, Equation 2.75 is reduced to an MA(1) process, and when $\theta_1 = 0$, it is reduced to AR(1) process. Thus, the AR(1) and MA(1) processes are special cases of the ARMA(1,1) process.

Autocorrelation function of the ARMA(1,1) process

To obtain the autocovariance for $\{z_t\}$, multiply z_{t-k} on both sides of the Equation 2.74

$$z_t - \phi_1 z_{t-1} = e_t - \theta_1 e_{t-1}$$

$$z_t = \phi_1 z_{t-1} + e_t - \theta_1 e_{t-1}$$

$$z_{t-k}\, z_t = \phi_1 z_{t-k}\, z_{t-1} + z_{t-k}\, e_t - \theta_1 z_{t-k}\, e_{t-1}$$

And take the expected value, where ϕ_1 and θ_1 are constants

$$E[z_{t-k}\, z_t] = \phi_1 E[z_{t-k}\, z_{t-1}] + E[z_{t-k}\, e_t] - \theta_1 E[z_{t-k}\, e_{t-1}]$$

because e_t and z_{t-k} are independent, it follows that

$$cov(z_{t-k}, z_t) = \phi_1 cov(z_{t-k}\, z_{t-1}) + E[z_{t-k}\, e_t] - \theta_1 E[z_{t-k}\, e_{t-1}]$$

$$\gamma_k = \phi_1 \gamma_{k-1} + E[z_{t-k}\, e_t] - \theta_1 E[z_{t-k}\, e_{t-1}] \tag{2.76}$$

when $k = 0$

$$\gamma_0 = \phi_1 \gamma_1 + E[z_t\, e_t] - \theta_1 E[z_t\, e_{t-1}]$$

Given that $E[z_t e_t] = \sigma_e^2$, and that the term $E[z_t e_{t-1}]$ can be written as

$$E[z_t e_{t-1}] = E[(\phi_1 z_{t-1} + e_t - \theta_1 e_{t-1}) e_{t-1}] = (\phi_1 - \theta_1)\sigma_e^2$$

then

$$\gamma_0 = \phi_1 \gamma_1 + \sigma_e^2 - \theta_1(\phi_1 - \theta_1)\sigma_e^2 \tag{2.77}$$

When $k = 1$, from Equation 2.76

$$\gamma_1 = \phi_1 \gamma_0 + E[z_{t-1}\, e_t] - \theta_1 E[z_{t-1}\, e_{t-1}]$$

$$= \phi_1 \gamma_0 - \theta_1 \sigma_e^2 \tag{2.78}$$

Substituting Equation 2.78 in (2.77), then

$$\gamma_0 = \phi_1(\phi_1 \gamma_0 - \theta_1 \sigma_e^2) + \sigma_e^2 - \theta_1(\phi_1 - \theta_1)\sigma_e^2$$

$$= \frac{(1 + \theta_1^2 - 2\phi_1\theta_1)}{(1 - \phi_1^2)}\sigma_e^2$$

Thus, from Equation (2.78) then

$$\gamma_1 = \frac{\phi_1(1 + \theta_1^2 - 2\phi_1\theta_1)}{(1 - \phi_1^2)}\sigma_e^2 - \theta_1 \sigma_e^2$$

$$= \frac{(\phi_1 - \theta_1)(1 - \phi_1\theta_1)}{(1 - \phi_1^2)}\sigma_e^2$$

For $k \geq 2$, from Equation 2.76

$$\gamma_k = \phi_1 \gamma_{k-1}, \quad k \geq 2$$

Hence, the ARMA(1,1) model has the following autocorrelation function:

$$\rho_k = \begin{cases} 1 & k = 0 \\ \dfrac{(\phi_1 - \theta_1)(1 - \phi_1 \theta_1)}{1 + \theta_1^2 - 2\phi_1 \theta_1}, & k = 1 \\ \phi_1 \rho_{k-1}, & k \geq 2 \end{cases} \tag{2.79}$$

The autocorrelation function of an ARMA(1,1) model combines characteristics of both AR(1) and MA(1) processes. The moving average parameter θ_1 enters into the calculation of ρ_1. Beyond ρ_1, the autocorrelation function of an ARMA(1,1) model follows the same pattern as the autocorrelation function of an AR(1) process.

Partial autocorrelation function of the ARMA(1,1) process

The partial autocorrelation function of the mixed ARMA(1,1) process Equation 2.75 consists of a single initial value $\phi_{11} = \rho_1$. Thereafter it behaves like the partial autocorrelation function of a pure MA(1) process, and is dominated by a damped exponential. Thus, as shown in Figure 2.3, when θ_1 is positive, it is dominated by a smoothly damped exponential which decays from a value of ρ_1, with sign determined by the sign of $(\phi_1 - \theta_1)$. Similarly, when θ_1 is negative, it is dominated by an exponential which oscillates as it decays from a value of ρ_1, with sign determined by the sign of $(\phi_1 - \theta_1)$.

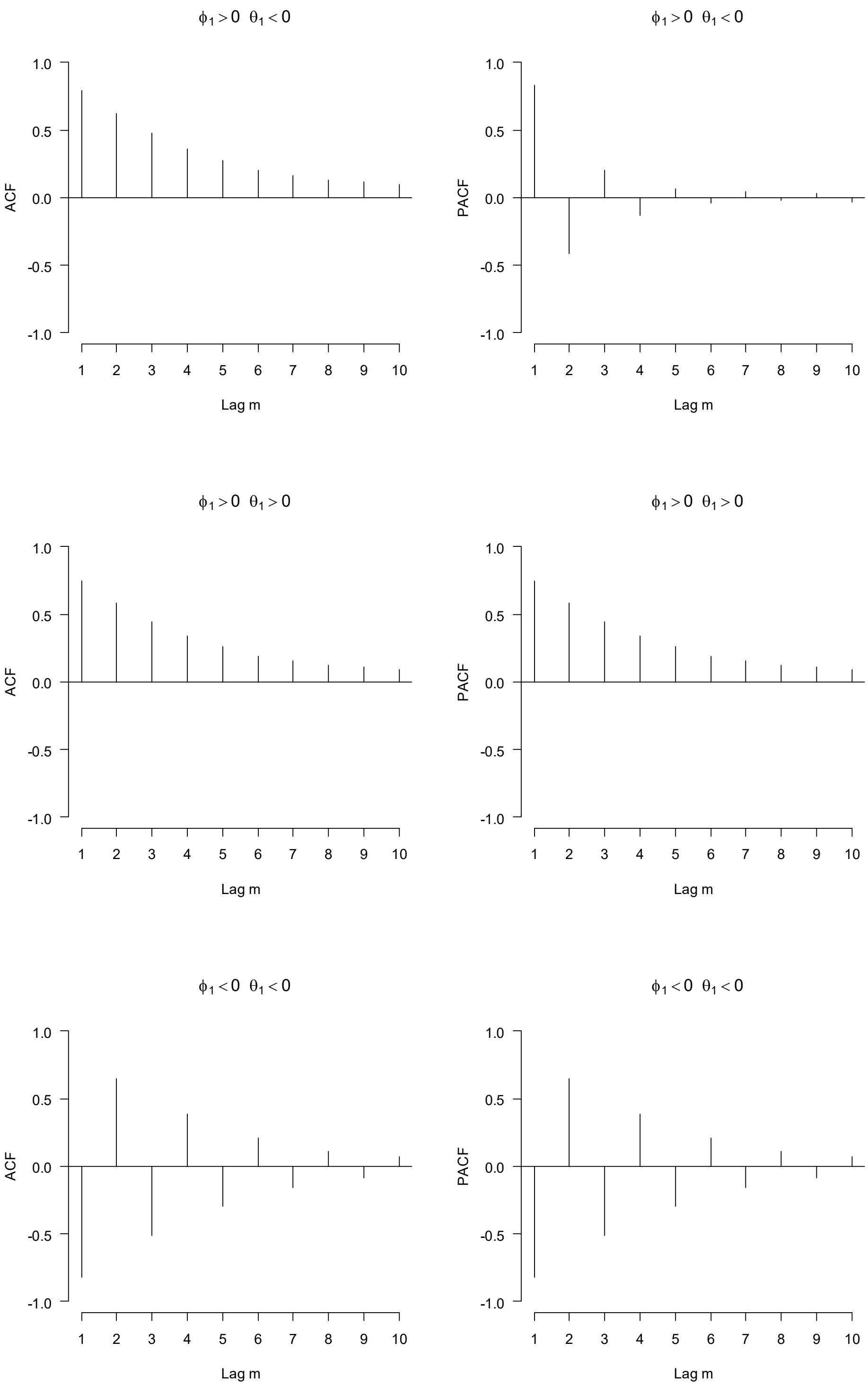

Figure 2.3 Autocorrelation and partial autocorrelation functions ρ_k and ϕ_{kk} for various ARMA (1, 1) models.

2.6 Model building

In 1970, Box and Jenkins proposed an important three-stages procedure in time series analysis for analysing an appropriated ARMA (p, q) process to forecast the observations. These three stages are namely, Identification, Estimation and Diagnostic Checking. Figure 2.4 shows the procedure of model building.

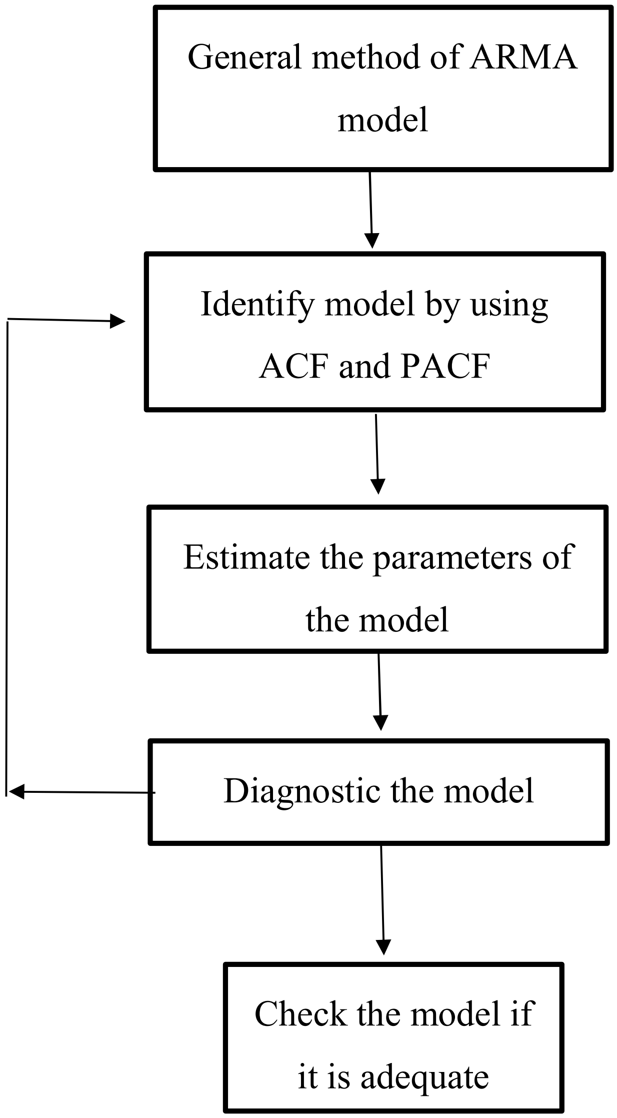

Figure 2.4 Stages of the Box and Jenkins methodology.

2.6.1 Model identifications

In time series analysis the most significant steps are to identify and build a model based on the available data. These steps require a good understanding of the process, particularly the characteristics of the process in terms of their autocorrelation function ρ_k and partial autocorrelation function ϕ_{kk} (Akaike,1974). In practice, the ACF and PACF are unknown, and for a given observed time series $z_1, z_2, \dots, z_n$, they have to be estimated by the sample ACF $\hat{\rho}_k$ and sample PACF $\hat{\phi}_{kk}$. Thus, in model identification, the goal is to match patterns in the sample ACF $\hat{\rho}_k$, and sample PACF $\hat{\phi}_{kk}$, with the theoretical patterns of the ACF ρ_k and the PACF ϕ_{kk}, for the ARMA processes. Table 2.1 shows the theoretical behaviour of the ACF ρ_k and the PACF ϕ_{kk} for processes known to be of type AR, MA and ARMA.

Process	**ACF**	**PACF**
AR$(\boldsymbol{p})$	Tails off exponential decay or damped sine-cosine wave	Cuts off after lag p
MA$(\boldsymbol{q})$	Cuts off after lag q	Tails off exponential decay or damped sine-cosine wave
ARMA$(\boldsymbol{p}, \boldsymbol{q})$	Tails off after $(q - p)$ lags, $q > p$	Tails off after $(p - q)$ lags, $p > q$

Table 2.1 Characteristics of theoretical ACF and PACF for stationary processes.

2.6.2 Model estimation

The main reason for model estimation is to determine an appropriate ARMA (p, q) model of a stationary time series. This involves calculating estimates of the mean, white noise variance and the coefficients of ϕ from an autoregressive process and θ from a moving average process. There are several methods of estimating the parameters of a time series model, namely, the maximum likelihood function, ordinary least squares and Bayes' theory. The maximum likelihood function approach has been taken in this thesis (see Equations (5.20) and (5.21)). These methods of estimation have been derived and developed by researchers, such as, Barnard (1949), Birnbaum (1962), Rao (1965), Kendall and Stuart (1966), and Hannan (1970).

2.6.3 Model diagnostic checking

The final stage of the Box and Jenkins methodology is Diagnostic Checking. The adequacy of a statistical model is examined - the residual autocorrelations and partial autocorrelations are used as a diagnostic check to test the goodness of fit of the model. A portmanteau test is used to test the goodness of fit of an ARMA model. Diagnostic checking will be discussed in Chapter 5.

2.7 Summary

This chapter has presented the definitions and formulas commonly used in time series analysis. This chapter has also presented the Box and Jenkins methodology of model building. In addition, the characteristics of autoregressive, moving average, and mixed autoregressive-moving average processes have also been provided.

Chapter 3 - Multivariate Vector ARMA Time Series

A major extension to the Box and Jenkins methodology has been the development of vector autoregressive moving average (VARMA) models. A VARMA model incorporates several time series at the same time, and takes into account interactions between one time series and another. VARMA models offer the potential for greater parsimony and an increase in forecasting accuracy. They are widely used in economics, in particular, in macroeconomic modelling, Reinsel (1993) and Lütkepohl (2005).

This chapter will discuss the mean, covariance and correlation matrix functions for the multivariate case, the vector white noise process, the linear process of vector time series, the vector autoregressive moving average process, the vector AR(1), the vector AR(p), the vector MA(1), the vector MA(q), the vector ARMA(1,1) and the model building of vector ARMA time series.

3.1 Vector time series

The basic idea of vector time series is that at each point in time there are a number of quantises that can be measured, and these can be regarded as the components of a vector. For example, consider three weather measurements made on a daily basis: rainfall, maximum temperature and minimum temperature. These can be represented by the three variables $z_{1,t}$, $z_{2,t}$ and $z_{3,t}$ respectively. These three variables can be regarded as the three components of a vector variable $\boldsymbol{z}_t = \left(z_{1,t}, z_{2,t}, z_{3,t}\right)^{\dagger}$.

In general, a vector time series can be represented by

$$\boldsymbol{z}_t = \left(z_{1,t}, z_{2,t}, \dots, z_{d,t}\right)^{\dagger}, t = 0, 1, 2, \dots, \tag{3.1}$$

where d is the number of components in the vector.

There are two main reasons for studying a vector time series. The first reason is to understand the relationship among the component series. The second reason is to enable forecasting to be made. Vector time series have been discussed by many researchers, such as, Quenouille (1957), Whittle (1963), Hannan (1970), Zellner and Palm (1974), Wallis (1977), Tiao and Box (1981), Hannan and Kavalieris (1984), Tiao and Tsay (1989) and Box, Jenkins, Reinsel and Ljung, (2015).

3.2 Mean, covariance and correlation matrix functions

Similar to the univariate case, terms can be derived for the mean, the covariance matrix function, and the correlation matrix function in a multivariate case.

3.2.1 The mean vector

Suppose that $\boldsymbol{z}_t = \left(z_{1,t}, z_{2,t}, \dots, z_{d,t}\right)^{\dagger}$, $t = 0,1,2,\dots$, denotes a d-dimensional jointly stationary real-valued vector process, so that, the mean $E\left[\boldsymbol{z}_{i,t}\right] = \mu_i$ is constant for each component $i = 1, 2, \dots, d$, then the mean vector can be written as

$$E\left[\boldsymbol{z}_{i,t}\right] = \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_d \end{pmatrix} \tag{3.2}$$

3.2.2 Covariance matrix function

A covariance matrix function of a vector time series is used to measure the strength of the linear dependence between the components of a stationary time series $\boldsymbol{z}_t$. Suppose that $\boldsymbol{z}_t = \left(z_{1,t}, z_{2,t}, \dots, z_{d,t}\right)^{\dagger}$, $t = 0, 1, 2, \dots$, denotes a d-dimensional jointly stationary real-valued vector process. The cross-covariance between $\boldsymbol{z}_{i,t}$ and $\boldsymbol{z}_{j,s}$ for all $i = 1, 2, \dots, d$ and $j = 1, 2, \dots, d$, are functions only of the time difference $(s - t)$.

Hence the lag-k covariance matrix for the vector time series $\boldsymbol{z}_t$ can be written as

$$\boldsymbol{\Gamma}(k) = Cov\{\boldsymbol{z}_t, \boldsymbol{z}_{t-k}\} = E[(\boldsymbol{z}_t - \boldsymbol{\mu})(\boldsymbol{z}_{t-k} - \boldsymbol{\mu})^{\dagger}]$$

and the lag-k cross-covariance matrix between two vectors $\boldsymbol{z}_t$ and $\boldsymbol{z}_{t-k}$, both $d \times 1$ vectors, can be written as a $d \times d$ matrix

$$\boldsymbol{\Gamma}(k) = Cov\{\boldsymbol{z}_t, \boldsymbol{z}_{t-k}\} = E[(\boldsymbol{z}_t - \boldsymbol{\mu})(\boldsymbol{z}_{t-k} - \boldsymbol{\mu})^{\dagger}]$$

$$= E\left[\begin{bmatrix}(z_{1,t}-\mu_1)\\(z_{2,t}-\mu_2)\\\vdots\\(z_{d,t}-\mu_d)\end{bmatrix}[(z_{1,t-k}-\mu_1),(z_{2,t-k}-\mu_2)\cdots(z_{d,t-k}-\mu_d)]\right]$$

$$= E\begin{bmatrix}(z_{1,t}-\mu_1)(z_{1,t-k}-\mu_1) & (z_{1,t}-\mu_1)(z_{2,t-k}-\mu_2) & \cdots & (z_{1,t}-\mu_1)(z_{d,t-k}-\mu_d)\\(z_{2,t}-\mu_2)(z_{1,t-k}-\mu_1) & (z_{2,t}-\mu_2)(z_{2,t-k}-\mu_2) & \cdots & (z_{2,t}-\mu_2)(z_{d,t-k}-\mu_d)\\\vdots & \vdots & \cdots & \vdots\\(z_{d,t}-\mu_d)(z_{1,t-k}-\mu_1) & (z_{d,t}-\mu_d)(z_{2,t-k}-\mu_2) & \cdots & (z_{d,t}-\mu_K)(z_{d,t-k}-\mu_d)\end{bmatrix}$$

$$= \begin{pmatrix}\gamma_{11}(k) & \gamma_{12}(k) & \cdots & \gamma_{1d}(k)\\\gamma_{21}(k) & \gamma_{22}(k) & \cdots & \gamma_{2d}(k)\\\vdots & \vdots & \cdots & \vdots\\\gamma_{d1}(k) & y_{d2}(k) & \cdots & \gamma_{dd}(k)\end{pmatrix} \tag{3.3}$$

where

$$\gamma_{ij}(k) = E\left[(z_{i,t}-\mu_i)(z_{j,t-k}-\mu_j)\right] = E\left[(z_{i,t-k}-\mu_i)(z_{j,t}-\mu_j)\right]$$

for $k = 0, 1, 2, \dots, i, j = 1, 2, \dots, d$ as a function of k, $\boldsymbol{\Gamma}(k)$ is called the covariance matrix function for the vector process $\boldsymbol{z}_t$.

For a stationary process $\{\boldsymbol{z}_t\}$ the covariance between $z_{i,t}$ and $z_{j,t+k}$ must depend only on the lag k, not on time t for all $i = 1, 2, \dots, d$ and $j = 1, 2, \dots, d$. For $i = j$, $\gamma_{ij}(k)$ is the autocovariance function for the ith component process $\boldsymbol{z}_{i,t}$. For $i \neq j$, $y_{ij}(k)$ is the cross-covariance function between $\boldsymbol{z}_{i,t}$ and $\boldsymbol{z}_{j,t}$.

From the definition in Equation 3.3, for negative lag k

$$\boldsymbol{\Gamma}(k) = E[(\boldsymbol{z}_t - \boldsymbol{\mu})(\boldsymbol{z}_{t-k} - \boldsymbol{\mu})^{\dagger}]$$

Because of stationarity

$$= E[(\boldsymbol{z}_{t+k} - \boldsymbol{\mu})(\boldsymbol{z}_t - \boldsymbol{\mu})^{\dagger}]$$

Applying the rule of matrix transpose, that is, $A = (A^{\dagger})^{\dagger}$ and $(AB)^{\dagger} = B^{\dagger}A^{\dagger}$, gives

$$= \{\{E[(\boldsymbol{z}_{t+k} - \boldsymbol{\mu})(\boldsymbol{z}_t - \boldsymbol{\mu})^{\dagger}]\}^{\dagger}\}^{\dagger}$$

$$= \{E[(\boldsymbol{z}_t - \boldsymbol{\mu})(\boldsymbol{z}_{t+k} - \boldsymbol{\mu})^{\dagger}]\}^{\dagger}$$

$$= \left\{E\left[(\boldsymbol{z}_t - \boldsymbol{\mu})\left(\boldsymbol{z}_{t-(-k)} - \boldsymbol{\mu}\right)^{\dagger}\right]\right\}^{\dagger}$$

From the definition of the covariance matrix

$$= \{\boldsymbol{\Gamma}(-k)\}^{\dagger}$$

3.2.3 Correlation matrix functions

A correlation matrix function is used to investigate the dependence between the multiple variables at the same time. The correlation matrix function for the vector process can be defined as

$$\boldsymbol{\rho}(k) = D^{-1/2}\boldsymbol{\Gamma}(k)D^{-1/2} = \rho_{ij}(k) \tag{3.4}$$

for $i, j = 1, 2, \dots, d$, where D is the diagonal matrix of Equation 3.3 in which the ith diagonal element is the variance of the ith process. Hence

$$D = \mathrm{diag}\left(\gamma_{11}(0), \gamma_{22}(0), \cdots, \gamma_{dd}(0)\right).$$

Thus, the ith diagonal element of $\boldsymbol{\rho}(k)$ is the autocorrelation function for the ith component series $\boldsymbol{z}_{i,t}$, whereas the $(i,j)th$ diagonal element of $\boldsymbol{\rho}(k)$ is

$$\boldsymbol{\rho}_{ij}(k) = \frac{\gamma_{ij}(k)}{\left(\gamma_{ii}(0)\gamma_{jj}(0)\right)^{1/2}} \tag{3.5}$$

where $\boldsymbol{\rho}_{ij}(k)$ is the cross-correlation function between component series $\boldsymbol{z}_{i,t}$ and $\boldsymbol{z}_{j,t}$.

3.2.4 The sample mean, sample covariance matrix and sample cross-correlation matrix

Suppose that $\boldsymbol{z}_t = \left(z_{1,t}, z_{2,t}, \dots, z_{d,t}\right)^{\dagger}$, $t = 0, 1, 2, \dots$, denotes a d-dimensional jointly stationary real-valued vector process, then sample mean vector can be defined as

$$\bar{Z} = \frac{1}{n}\sum_{t=1}^{n} \boldsymbol{z}_t \tag{3.6}$$

The sample covariance matrix of a time series can be written as

$$\widehat{\boldsymbol{\Gamma}}(0) = \frac{1}{n}\sum_{t=1}^{n}(\boldsymbol{z}_t - \bar{Z})\,(\boldsymbol{z}_t - \bar{Z})^{\dagger} \tag{3.7}$$

The sample cross-covariance matrix of a time series at lag k can be written as

$$\widehat{\boldsymbol{\Gamma}}(k) = \frac{1}{n}\sum_{t=1}^{n-k}(\boldsymbol{z}_t - \bar{Z})\,(\boldsymbol{z}_{t-k} - \bar{Z})^{\dagger} \tag{3.8}$$

The sample cross-correlations are defined

$$\widehat{\boldsymbol{\rho}}_{ij}(k) = \frac{\sum_{t=1}^{n-k}\left(z_{i,t} - \bar{Z}_i\right)\left(z_{j,t+k} - \bar{Z}_j\right)}{\left(\sum_{t=1}^{n}\left(z_{i,t} - \bar{Z}_i\right)^2 \sum_{t=1}^{n}\left(z_{j,t} - \widehat{\mu\bar{Z}_j}\right)^2\right)^{1/2}}, \qquad \text{for } i,j = 1,2,\dots,d \tag{3.9}$$

The cross-correlation $\boldsymbol{\rho}_{ij}(k)$ and the sample cross-correlation $\widehat{\boldsymbol{\rho}}_{ij}(k)$ are very useful in identifying a finite-order moving average model as $\boldsymbol{\rho}_{ij}(k) = 0$ for all $k > q$ for the vector MA(q) model. Unfortunately, the sample cross-correlations $\widehat{\boldsymbol{\rho}}_{ij}(k)$ may be difficult to estimate because of the large number of terms that may need to be estimated and examined for a multivariate time series.

3.3 Vector white noise process

A vector white noise process is defined as a sequence of independent random vectors, denoted as $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_t$ where $\boldsymbol{e}_t = (e_{1t}, e_{2t}, \dots, e_{dt})^{\dagger}$, is a zero mean white noise process with covariance matrix $\boldsymbol{\Sigma} = E[\boldsymbol{e}_t\boldsymbol{e}_t{}^{\dagger}]$, where $\boldsymbol{\Sigma}$ is a $d \times d$ symmetric positive definite matrix.

$$\boldsymbol{\Gamma}(k) = E[\boldsymbol{e}_t\boldsymbol{e}_{t+k}{}^{\dagger}] = \begin{cases}\boldsymbol{\Sigma}, & \text{if } k = 0\\ \boldsymbol{0}, & \text{if } k \neq 0\end{cases} \tag{3.10}$$

3.4 The linear process of vector time series

Suppose that $\mathbf{z}_{\mathrm{t}} = \left(\mathrm{z}_{1,\mathrm{t}}, \mathrm{z}_{2,\mathrm{t}}, \dots, \mathrm{z}_{d,t}\right)^{\dagger}$, $t = 0,1,2,\dots,$ denotes a d-dimensional jointly stationary real-valued vector process so that the mean $E\left[z_{i,t}\right] = \mu_i$ is constant for each $i = 1,2,\dots,d$. If the d-dimensional stationary vector process $\mathbf{z}_{\mathrm{t}}$ can be written as a combination of a sequence of d-dimensional white noise random vectors, then $\mathbf{z}_{\mathrm{t}}$ is a linear process

$$\mathbf{z}_\mathrm{t} = \boldsymbol{\mu} + \mathbf{e}_\mathrm{t} + \boldsymbol{\Psi}_1\mathbf{e}_{\mathrm{t}-1} + \boldsymbol{\Psi}_2\mathbf{e}_{\mathrm{t}-2} + \cdots$$

$$= \boldsymbol{\mu} + \sum_{i=0}^{\infty} \boldsymbol{\Psi}_i \mathbf{e}_{\mathrm{t-i}} \tag{3.11}$$

where $\boldsymbol{\mu}$ is a d-dimensional constant vector, $\boldsymbol{\Psi}_0 = I$ is a $d \times d$ identity matrix, the $\boldsymbol{\Psi}_i$'s are $d \times d$ coefficients matrices, and $\boldsymbol{e}_t = (e_{1t}, e_{2t}, \dots, e_{dt})^\dagger$, is a d-dimensional, zero mean, white noise process. For the stationary linear time series $\mathbf{z}_\mathrm{t}$ in Equation 3.11

$$E[\mathbf{z}_\mathrm{t}] = \boldsymbol{\mu}$$

and

$$\boldsymbol{\Gamma}(k) = E[(\boldsymbol{z}_t - \boldsymbol{\mu})(\boldsymbol{z}_{t-k} - \boldsymbol{\mu})^\dagger]$$

$$= E\left[\left(\sum_{i=0}^{\infty} \boldsymbol{\Psi}_i \mathbf{e}_{\mathrm{t-i}}\right)\left(\sum_{i=0}^{\infty} \boldsymbol{\Psi}_{i-k} \mathbf{e}_{\mathrm{t-i}}\right)^\dagger\right]$$

$$= E[(\mathbf{e}_\mathrm{t} + \boldsymbol{\Psi}_1\mathbf{e}_{\mathrm{t}-1} + \cdots)(\mathbf{e}_\mathrm{t} + \boldsymbol{\Psi}_{1-k}\mathbf{e}_{\mathrm{t}-1} + \cdots)^\dagger]$$

$$= E[(\mathbf{e}_\mathrm{t} + \boldsymbol{\Psi}_1\mathbf{e}_{\mathrm{t}-1} + \cdots)(\boldsymbol{e}_t{}^\dagger + (\boldsymbol{\Psi}_{1-k}\mathbf{e}_{\mathrm{t}-1})^\dagger + \cdots)]$$

$$= E[(\mathbf{e}_\mathrm{t}\boldsymbol{e}_t{}^\dagger + \mathbf{e}_\mathrm{t}(\boldsymbol{\Psi}_{1-k}\mathbf{e}_{\mathrm{t}-1})^\dagger + \cdots + \boldsymbol{\Psi}_1\mathbf{e}_{\mathrm{t}-1}\boldsymbol{e}_t{}^\dagger + \boldsymbol{\Psi}_1\mathbf{e}_{\mathrm{t}-1}(\boldsymbol{\Psi}_{1-k}\mathbf{e}_{\mathrm{t}-1})^\dagger + \cdots)]$$

From the vector white noise $\boldsymbol{e}_t = (e_{1t}, e_{2t}, \dots, e_{dt})^\dagger$ such that $E[\boldsymbol{e}_t] = 0$, $\boldsymbol{\Sigma} = E[\boldsymbol{e}_t\boldsymbol{e}_t{}^\dagger]$, $E[\boldsymbol{e}_t\boldsymbol{e}_{t+k}{}^\dagger] = 0$ for $k \neq 0$ and $\boldsymbol{\Psi}_0 = I$

$$= E\left[\left(\sum_{i=0}^{\infty} \boldsymbol{\Psi}_i \mathbf{e}_{\mathrm{t-i}}\, \mathbf{e}_{\mathrm{t-i}}{}^\dagger \boldsymbol{\Psi}_{i-k}^\dagger\right)\right]$$

$$= \sum_{i=0}^{\infty} \boldsymbol{\Psi}_i\, E[\mathbf{e}_{\mathrm{t-i}}\mathbf{e}_{\mathrm{t-i}}{}^\dagger]\boldsymbol{\Psi}_{i-k}^\dagger$$

$$= \sum_{i=0}^{\infty} \boldsymbol{\Psi}_i\, \boldsymbol{\Sigma}\, \boldsymbol{\Psi}_{i-k}^\dagger$$

3.5 The vector autoregressive process

The general form of the vector autoregressive process of order p, VAR(p), can be defined as

$$\boldsymbol{z}_t = \boldsymbol{\Phi}_1 \boldsymbol{z}_{t-1} + \boldsymbol{\Phi}_2 \boldsymbol{z}_{t-2} + \cdots + \boldsymbol{\Phi}_p \boldsymbol{z}_{t-p} + \boldsymbol{e}_t, \tag{3.12}$$

where $\boldsymbol{z}_t$ is a d-dimensional vector valued time series, $\boldsymbol{\Phi}_i$ $(i = 1,2, \dots, p)$ are $d \times d$ parameter matrices and $\boldsymbol{e}_t = (e_{1t}, e_{2t}, \dots, e_{dt})^{\dagger}$, is a d-dimensional zero mean white noise process with covariance matrix $\boldsymbol{\Sigma} = E[\boldsymbol{e}_t {\boldsymbol{e}_t}^{\dagger}]$. Using backshift operators, the vector autoregressive process of order p can be written as

$$\left(\boldsymbol{I} - \boldsymbol{\Phi}_1 B - \cdots - \boldsymbol{\Phi}_p B^p\right)\boldsymbol{z}_t = \boldsymbol{\Phi}(B)\boldsymbol{z}_t = \boldsymbol{e}_t$$

where $\boldsymbol{\Phi}(B)$ is a matrix polynomial of the backshift operator B of order p. The vector of autoregressive process VAR(p) was developed by researchers, such as, Sims (1980), Granger (1981), and Engle and Granger (1987).

The vector autoregressive process of order p will be stationary if this condition is satisfied that the zeros of $\left\|\boldsymbol{I} - \boldsymbol{\Phi}_1 B - \cdots - \boldsymbol{\Phi}_p B^p\right\|$ lie outside the unit circle, and if the roots of

$$\|\boldsymbol{\Phi}(B)\| = \left\|\boldsymbol{I} - \boldsymbol{\Phi}_1 B - \cdots - \boldsymbol{\Phi}_p B^p\right\| = 0,$$

are all greater than one in absolute value, where $\|\boldsymbol{A}\|$ is the determinant of a matrix $\boldsymbol{A}$, Reinsel (1993).

3.5.1 Covariance matrix function of vector autoregressive process of order p

The covariance matrix function of a vector AR(p) process can be obtained by multiplying Equation 3.12 by $\boldsymbol{z}_{t-k}^{\dagger}$ and taking the expectation, gives

$$E\left[\boldsymbol{z}_t \boldsymbol{z}_{t-k}^{\dagger}\right] = \boldsymbol{\Phi}_1 E\left[\boldsymbol{z}_{t-1} \boldsymbol{z}_{t-k}^{\dagger}\right] + \boldsymbol{\Phi}_2 E\left[\boldsymbol{z}_{t-2} \boldsymbol{z}_{t-k}^{\dagger}\right] + \cdots + E\left[\boldsymbol{e}_t \boldsymbol{z}_{t-k}^{\dagger}\right]$$

$$\boldsymbol{\Gamma}(k) = \boldsymbol{\Phi}_1 \boldsymbol{\Gamma}(k-1) + \boldsymbol{\Phi}_2 \boldsymbol{\Gamma}(k-2) + \cdots + \boldsymbol{\Phi}_p \boldsymbol{\Gamma}(k-p), \qquad k > p$$

For a vector AR(p) process, $\boldsymbol{z}_t$ is given by Equation 3.12, it follows from the infinite MA representation given in Equation 3.14 that

$$E[\boldsymbol{z}_{t-k}\boldsymbol{e}_t^{\dagger}] = \begin{cases} 0 & k > 0 \\ \boldsymbol{\Sigma} & k = 0 \end{cases}$$

By using the system of Yule-Walker matrix equations to obtain the AR coefficient matrices $\boldsymbol{\Phi}_i$ from the cross-covariance matrices $\boldsymbol{\Gamma}(0), \boldsymbol{\Gamma}(1), \dots, \boldsymbol{\Gamma}(p)$, it can be written in the form

$$\boldsymbol{\Gamma}(k) = \sum_{i=1}^{p} \boldsymbol{\Gamma}(k-i)\boldsymbol{\Phi}_i^{\dagger}, \qquad \text{for } k = 1,2,\dots,p \tag{3.13}$$

or

$$\begin{pmatrix} \boldsymbol{\Gamma}(0) \\ \boldsymbol{\Gamma}(1) \\ \vdots \\ \boldsymbol{\Gamma}(p) \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Gamma}(0) & \boldsymbol{\Gamma}(1)^{\dagger} & \boldsymbol{\Gamma}(2)^{\dagger} & \cdots & \boldsymbol{\Gamma}(p-1)^{\dagger} \\ \boldsymbol{\Gamma}(1) & \boldsymbol{\Gamma}(0) & \boldsymbol{\Gamma}(1)^{\dagger} & \cdots & \boldsymbol{\Gamma}(p-2)^{\dagger} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Gamma}(p-1) & \Gamma(p-2) & \boldsymbol{\Gamma}(p-3) & \cdots & \boldsymbol{\Gamma}(0) \end{pmatrix} \begin{pmatrix} \boldsymbol{\Phi}_1^{\dagger} \\ \boldsymbol{\Phi}_2^{\dagger} \\ \vdots \\ \boldsymbol{\Phi}_p^{\dagger} \end{pmatrix}$$

For the case $k = 0$

$$\boldsymbol{\Sigma} = \boldsymbol{\Gamma}(0) - \sum_{i=1}^{p} \boldsymbol{\Gamma}(-i)\boldsymbol{\Phi}_i^{\dagger} \tag{3.14}$$

3.5.2 The first-order vector autoregressive process

The first-order vector autoregressive can be written as

$$\boldsymbol{z}_t = \boldsymbol{\Phi}_1 \boldsymbol{z}_{t-1} + \boldsymbol{e}_t, \tag{3.15}$$

or

$$(I - \boldsymbol{\Phi}_1 B)\boldsymbol{z}_t = \boldsymbol{e}_t$$

For the case $k = 2$

$$\begin{pmatrix} \boldsymbol{z}_{1,t} \\ \boldsymbol{z}_{2,t} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Phi}_{1,11} & \boldsymbol{\Phi}_{1,12} \\ \boldsymbol{\Phi}_{1,21} & \boldsymbol{\Phi}_{1,22} \end{pmatrix} \begin{pmatrix} \boldsymbol{z}_{1,t-1} \\ \boldsymbol{z}_{2,t-1} \end{pmatrix} + \begin{pmatrix} \boldsymbol{e}_{1,t} \\ \boldsymbol{e}_{2,t} \end{pmatrix} \tag{3.16}$$

or

$$\boldsymbol{z}_{1,t} = \boldsymbol{\Phi}_{1,11}\boldsymbol{z}_{1,t-1} + \boldsymbol{\Phi}_{1,12}\boldsymbol{z}_{2,t-1} + \boldsymbol{e}_{1,t}$$

$$\boldsymbol{z}_{2,t} = \boldsymbol{\Phi}_{1,21}\boldsymbol{z}_{1,t-1} + \boldsymbol{\Phi}_{1,22}\boldsymbol{z}_{2,t-1} + \boldsymbol{e}_{2,t}$$

It is clear that each element of $\boldsymbol{z}_t$ is a function of each element of $\boldsymbol{z}_{t-1}$.

The first-order vector autoregressive process satisfies the stationarity condition if and only if $\|\boldsymbol{I} - \boldsymbol{\Phi}B\| = 0$, or equivalently $\|\lambda\boldsymbol{I} - \boldsymbol{\Phi}\| = 0$, where the eigenvalue $\lambda = B^{-1}$. It follows that the stationarity condition for the AR(1) model is equivalent to the condition that all eigenvalues of $\boldsymbol{\Phi}$, that is, all roots of $\|\boldsymbol{I} - \boldsymbol{\Phi}B\| = 0$, be less than one in absolute value.

For arbitrary $n > 0$, successive substitutions of $t + n$ in the right-hand side of Equation 3.15, gives

$$\mathbf{z}_{\mathrm{t}} = \sum_{i=0}^{t+n} \boldsymbol{\Phi}^{i}\mathbf{e}_{\mathrm{t-i}} + \boldsymbol{\Phi}^{t+n+1}\mathbf{z}_{\mathrm{-n-1}} \tag{3.17}$$

provided that all eigenvalues of $\boldsymbol{\Phi}$ are less than one in absolute value.

3.5.3 The covariance matrix function of the VAR(1) process

The covariance matrix for VAR(1) can be obtained by multiplying the Equation 3.15 by $\boldsymbol{z}_{t-k}^{\dagger}$ and taking the expectation, giving

$$E\left[\boldsymbol{z}_t\boldsymbol{z}_{t-k}^{\dagger}\right] = \boldsymbol{\Phi}_1 E\left[\boldsymbol{z}_{t-1}\boldsymbol{z}_{t-k}^{\dagger}\right] + E\left[\boldsymbol{e}_t\boldsymbol{z}_{t-k}^{\dagger}\right]$$

hence, when $k = 0$

$$\boldsymbol{\Gamma}(0) = \boldsymbol{\Phi}_1\boldsymbol{\Gamma}(-1) + \boldsymbol{\Sigma}$$

$$= \boldsymbol{\Phi}_1\boldsymbol{\Gamma}(1)^{\dagger} + \boldsymbol{\Sigma} \tag{3.18}$$

To compute $\Gamma(0)$, the values of $\boldsymbol{\Phi}_1$, $\boldsymbol{\Gamma}(1)^{\dagger}$ and $\boldsymbol{\Sigma}$ need to be known. The $\boldsymbol{\Phi}_1$, $\boldsymbol{\Gamma}(1)^{\dagger}$ and $\boldsymbol{\Sigma}$ can be obtained from

$$\boldsymbol{\Gamma}(k) = E\left[\boldsymbol{z}_{t-k}\boldsymbol{z}_t^{\dagger}\right]$$

Since $E\left[\boldsymbol{e}_t\boldsymbol{z}_{t-k}^{\dagger}\right] = 0$, for $k = 1, 2, \ldots$, then

$$\boldsymbol{\Gamma}(k) = \boldsymbol{\Phi}_1\boldsymbol{\Gamma}(k-1), \qquad \text{for } k = 1,2, \ldots,$$

so $k = 1$

$$\boldsymbol{\Gamma}(1) = \boldsymbol{\Phi}_1\boldsymbol{\Gamma}(0) \tag{3.19}$$

Hence,

$$\boldsymbol{\Gamma}(1)^{\dagger} = \boldsymbol{\Gamma}(0)\boldsymbol{\Phi}_1^{\dagger} \tag{3.20}$$

By substituting Equation 3.20 in Equation 3.18

$$\boldsymbol{\Gamma}(0) = \boldsymbol{\Phi}_1\boldsymbol{\Gamma}(0)\boldsymbol{\Phi}_1^{\dagger} + \boldsymbol{\Sigma} \tag{3.21}$$

Applying the vectorizing operation (vec) to both sides, then

$$vec\big(\boldsymbol{\Gamma}(0)\big) = vec\big(\boldsymbol{\Phi}_1\boldsymbol{\Gamma}(0)\boldsymbol{\Phi}_1^{\dagger}\big) + vec(\boldsymbol{\Sigma})$$

$$vec\big(\boldsymbol{\Gamma}(0)\big) = (\boldsymbol{\Phi}_1 \otimes \boldsymbol{\Phi}_1)vec\big(\boldsymbol{\Gamma}(0)\big) + vec(\boldsymbol{\Sigma})$$

$$vec\big(\boldsymbol{\Gamma}(0)\big) - (\boldsymbol{\Phi}_1 \otimes \boldsymbol{\Phi}_1)vec\big(\boldsymbol{\Gamma}(0)\big) = vec(\boldsymbol{\Sigma})$$

$$vec\big(\boldsymbol{\Gamma}(0)\big)\big(\boldsymbol{I}_{n^2} - (\boldsymbol{\Phi}_1 \otimes \boldsymbol{\Phi}_1)\big) = vec(\boldsymbol{\Sigma})$$

$$vec\big(\boldsymbol{\Gamma}(0)\big) = \big(\boldsymbol{I}_{n^2} - (\boldsymbol{\Phi}_1 \otimes \boldsymbol{\Phi}_1)\big)^{-1} vec(\boldsymbol{\Sigma}) \tag{3.22}$$

where $\otimes$ is a Kronecker product, which multiplies each element of matrix $\boldsymbol{C}^{\dagger}$ by the whole of matrix $\boldsymbol{A}$ to create a new matrix.

Note, vec is a linear vector operator that is used to transform a matrix to a vector, and it has the following vectorizing operation property, (Neudecker, 1969).

For matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$

$$vec(\boldsymbol{ABC}) = (\boldsymbol{C}^{\dagger} \otimes \boldsymbol{A})vec(\boldsymbol{B})$$

The Equation 3.22 can be used to find $\boldsymbol{\Gamma}(0)$ when $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Sigma}$ are known.

The $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Sigma}$ can be obtained from the parameters $\boldsymbol{\Gamma}(0)$ and $\boldsymbol{\Gamma}(1)$

$$\boldsymbol{\Gamma}(k) = \boldsymbol{\Gamma}(k-1)\boldsymbol{\Phi}_1^{\dagger}, \qquad \text{for } k = 1,2, \ldots$$

When $k = 1$

$$\boldsymbol{\Gamma}(1) = \boldsymbol{\Gamma}(0)\boldsymbol{\Phi}_1^{\dagger} \tag{3.23}$$

$$\boldsymbol{\Phi}_1^{\dagger} = \boldsymbol{\Gamma}(0)^{-1}\boldsymbol{\Gamma}(1) \tag{3.24}$$

By applying the matrix transpose rule to Equation 3.23

$$\boldsymbol{\Gamma}(1)^{\dagger} = \boldsymbol{\Phi}_1\boldsymbol{\Gamma}(0) \tag{3.25}$$

Equation 3.21 gives

$$\boldsymbol{\Gamma}(0) = \boldsymbol{\Phi}_1\boldsymbol{\Gamma}(0)\boldsymbol{\Phi}_1^{\dagger} + \boldsymbol{\Sigma}$$

By substituting Equations 3.24 and 3.25 in Equation 3.21

$$\boldsymbol{\Sigma} = \boldsymbol{\Gamma}(0) - \boldsymbol{\Gamma}(1)^{\dagger}\boldsymbol{\Gamma}(0)^{-1}\boldsymbol{\Gamma}(1) \tag{3.26}$$

Hence, $\boldsymbol{\Gamma}(k)$ is a covariance matrix function. The $\boldsymbol{\Phi}_1$ can be obtained from Equation 3.19 and the $\boldsymbol{\Sigma}$ can be obtained from Equation 3.26.

3.6 The vector moving average process

The general form of the vector moving average process of order q, VMA(q), can be written as

$$\boldsymbol{z}_t = \boldsymbol{\Theta}_1\boldsymbol{e}_{t-1} + \boldsymbol{\Theta}_2\boldsymbol{e}_{t-2} + \cdots + \boldsymbol{\Theta}_q\boldsymbol{e}_{t-q} + \boldsymbol{e}_t \tag{3.27}$$

where $\boldsymbol{z}_t = (z_{1t}, z_{2t}, \dots, z_{dt})^{\dagger}$, is a $d \times 1$ vector of time series observed at t, $\boldsymbol{e}_t = (e_{1t}, e_{2t}, \dots, e_{dt})^{\dagger}$, is a $d \times 1$ zero mean white noise process with a covariance matrix $\boldsymbol{\Sigma} = E[\boldsymbol{e}_t\boldsymbol{e}_t{}^{\dagger}]$ and $\boldsymbol{\Theta}_j$ is a $d \times d$ matrix of coefficients, for $(j = 1, 2, \dots, q)$. Using the backshift operator, the vector moving average process of order q can be written as

$$\left(\boldsymbol{I} - \boldsymbol{\Theta}_1 B - \cdots - \boldsymbol{\Theta}_q B^q\right)\boldsymbol{e}_t = \boldsymbol{\Theta}(B)\boldsymbol{e}_t = \boldsymbol{z}_t$$

where $\boldsymbol{\Theta}(B)$ is a matrix polynomial of the backshift operator B of order q.

The invertibility condition of VMA(q) process

The VMA(q) process is invertible if all roots of

$$\|\boldsymbol{\Theta}(B)\| = \left\|I - \boldsymbol{\Theta}_1 B - \cdots - \boldsymbol{\Theta}_q B^q\right\| = 0,$$

are greater than one in absolute value, where $\|.\|$ is the determinant of a matrix.

3.6.1 Covariance matrix function of vector moving average process of order q

The covariance matrix function of VMA(q) process can be obtained by

$$\begin{aligned}\boldsymbol{\Gamma}(k) &= \mathrm{cov}(\mathbf{z}_{\mathrm{t}}, \mathbf{z}_{\mathrm{t+k}}{}^{\dagger}) \\ &= E\left[\left(\boldsymbol{e}_t - \boldsymbol{\Theta}_1 \boldsymbol{e}_{t-1} - \cdots + \boldsymbol{\Theta}_q \boldsymbol{e}_{t-q}\right)\left(\boldsymbol{e}_{t+k} - \boldsymbol{\Theta}_1 \boldsymbol{e}_{t+k-1} - \cdots + \boldsymbol{\Theta}_q \boldsymbol{e}_{t+k-q}\right)^{\dagger}\right] \\ &= -\boldsymbol{\Sigma}\,\boldsymbol{\Theta}_k^{\dagger} + \boldsymbol{\Theta}_k \boldsymbol{\Sigma}\,\boldsymbol{\Theta}_{k+1}^{\dagger} + \cdots + \boldsymbol{\Theta}_{q-k} \boldsymbol{\Sigma}\,\boldsymbol{\Theta}_q^{\dagger}\end{aligned}$$

$$= \sum_{j=0}^{q-1} \boldsymbol{\Theta}_j \boldsymbol{\Sigma}\,\boldsymbol{\Theta}_{j+k}^{\dagger} \tag{3.28}$$

for $k = 1, 2, \dots, q$ with $\boldsymbol{\Theta}_0 = -I$, and $\boldsymbol{\Gamma}(k) = 0$ for $k > q$.
When $k = 0$

$$\boldsymbol{\Gamma}(0) = \boldsymbol{\Sigma} + \sum_{j=0}^{q-1} \boldsymbol{\Theta}_j \boldsymbol{\Sigma}\,\boldsymbol{\Theta}_{j+k}^{\dagger}$$

3.6.2 The first-order of the vector moving average process

The first-order of vector moving average can be written as

$$\boldsymbol{z}_t = \boldsymbol{e}_t - \boldsymbol{\Theta}_1 \boldsymbol{e}_{t-1} \tag{3.29}$$

or

$$(\boldsymbol{I} - \boldsymbol{\Theta}_1 B)\boldsymbol{e}_t = \boldsymbol{z}_t$$

For $k = 2$

$$\begin{pmatrix} \boldsymbol{z}_{1,t} \\ \boldsymbol{z}_{2,t} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{e}_{1,t} \\ \boldsymbol{e}_{2,t} \end{pmatrix} - \begin{pmatrix} \boldsymbol{\Theta}_{11} & \boldsymbol{\Theta}_{12} \\ \boldsymbol{\Theta}_{21} & \boldsymbol{\Theta}_{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{e}_{1,t-1} \\ \boldsymbol{e}_{2,t-2} \end{pmatrix} \tag{3.30}$$

or equivalently

$$\boldsymbol{z}_{1,t} = \boldsymbol{e}_{1,t} - \boldsymbol{\Theta}_{1,11} \boldsymbol{e}_{1,t-1} - \boldsymbol{\Theta}_{1,12} \boldsymbol{e}_{2,t-1}$$
$$\boldsymbol{z}_{2,t} = \boldsymbol{e}_{2,t} - \boldsymbol{\Theta}_{1,21} \boldsymbol{e}_{1,t-1} - \boldsymbol{\Theta}_{1,22} \boldsymbol{e}_{2,t-1}$$

It is clear that each element of $\boldsymbol{z}_t$ is a function of each element of $\boldsymbol{e}_{t-1}$.

The first-order of vector moving average process satisfies the invertible condition if all roots of $\|\boldsymbol{I} - \boldsymbol{\Theta}_1 B\| = 0$, are greater than one in absolute value. This is equivalent to the condition that all eigenvalues of $\boldsymbol{\Theta}_1$, that is, all roots λ of $\det\{\lambda \boldsymbol{I} - \boldsymbol{\Theta}_1\} = 0$, are less than one in absolute value.

The covariance matrix of the first-order of vector moving average process can be obtained by

$$\begin{aligned}\boldsymbol{\Gamma}(0) &= \mathrm{cov}(\mathbf{z}_\mathrm{t}, \mathbf{z}_\mathrm{t}{}^{\dagger}) \\ &= E[(\boldsymbol{e}_t - \boldsymbol{\Theta}_1 \boldsymbol{e}_{t-1})(\boldsymbol{e}_t - \boldsymbol{\Theta}_1 \boldsymbol{e}_{t-1})^{\dagger}] \\ &= \boldsymbol{\Sigma} + \boldsymbol{\Theta}_1 \boldsymbol{\Sigma}\, \boldsymbol{\Theta}_1^{\dagger},\end{aligned} \tag{3.31}$$

$$\boldsymbol{\Gamma}(1) = -\boldsymbol{\Sigma}\, \boldsymbol{\Theta}_1^{\dagger} \tag{3.32}$$

and

$$\boldsymbol{\Gamma}(k) = 0, \qquad \text{for } |k| > 1 \tag{3.33}$$

3.7 The vector autoregressive moving average (VARMA) process

The general form of the vector autoregressive moving average of order (p, q) can be defined as

$$\boldsymbol{z}_t = \boldsymbol{\Phi}_1 \boldsymbol{z}_{t-1} + \cdots + \boldsymbol{\Phi}_p \boldsymbol{z}_{t-p} + \boldsymbol{e}_t + \boldsymbol{\Theta}_1 \boldsymbol{e}_{t-1} + \boldsymbol{\Theta}_2 \boldsymbol{e}_{t-2} + \cdots + \boldsymbol{\Theta}_q \boldsymbol{e}_{t-q} \tag{3.34}$$

where $\boldsymbol{z}_t = (z_{1t}, z_{2t}, \dots, z_{dt})^{\dagger}$ is a $d \times 1$ vector of variables observed at t, $\boldsymbol{e}_t = (e_{1t}, e_{2t}, \dots, e_{dt})^{\dagger}$, is a $d \times 1$ zero mean white noise process with covariance matrix $\boldsymbol{\Sigma} = E[\boldsymbol{e}_t \boldsymbol{e}_t{}^{\dagger}]$, $\boldsymbol{\Phi}_i$ $(i = 1, 2, \dots, p)$ are $d \times d$ parameter matrices and $\boldsymbol{\Theta}_j$ is a $d \times d$ matrix of coefficients, for $(j = 1, 2, \dots, q)$. The vector autoregressive moving average process of order (p, q) can also be written using the backshift operator:

$$(\boldsymbol{I} - \boldsymbol{\Phi}_1 B - \cdots - \boldsymbol{\Phi}_p B^p)\boldsymbol{z}_t = (\boldsymbol{I} - \boldsymbol{\Theta}_1 B - \cdots - \boldsymbol{\Theta}_q B^q)\boldsymbol{e}_t$$

$$\boldsymbol{\Phi}(B)\boldsymbol{z}_t = \boldsymbol{\Theta}(B)\boldsymbol{e}_t$$

where $\boldsymbol{\Phi}(B)$ and $\boldsymbol{\Theta}(B)$ are matrix polynomials of the backshift operator B of order p and q. The vector autoregressive moving average process VARMA(p, q) has been discussed by researchers, such as, Hannan (1970, 1981), Reinsel (1993), Lütkepohl (2005), and Box et al., (2015).

The vector autoregressive moving average processes is stationary if the roots of $\left\|\boldsymbol{I} - \boldsymbol{\Phi}_1 B - \cdots - \boldsymbol{\Phi}_p B^p\right\|$ lie outside the unit circle, and if the roots of

$$\|\boldsymbol{\Phi}(B)\| = \left\|\boldsymbol{I} - \boldsymbol{\Phi}_1 B - \cdots - \boldsymbol{\Phi}_p B^p\right\| = 0,$$

are all greater than one in absolute value.

3.7.1 The covariance matrix function of the vector autoregressive moving average process

The covariance matrix function of the vector ARMA process can be obtained from the infinite MA representation

$$\mathbf{z}_{\mathrm{t}} = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \boldsymbol{\Psi}_i \mathbf{e}_{\mathrm{t-i}} = \boldsymbol{\mu} + \boldsymbol{\Psi}(B)\mathbf{e}_{\mathrm{t}}$$

Hence

$$E[\mathbf{z}_{\mathrm{t-k}} \boldsymbol{e}_{t-i}{}^{\dagger}] = \boldsymbol{\Psi}_{i-k}\boldsymbol{\Sigma}, \qquad \text{for } i \geq k$$

The covariance matrix function of the vector ARMA process is $\boldsymbol{\Gamma}(k) = Cov\{\boldsymbol{z}_{t-k}, \boldsymbol{z}_t\} = E[(\boldsymbol{z}_{t-k} - \boldsymbol{\mu})(\boldsymbol{z}_t - \boldsymbol{\mu})^{\dagger}]$ of $\boldsymbol{z}_t$, which satisfies the relation

$$\boldsymbol{\Gamma}(k) = Cov\{\boldsymbol{z}_{t-k}, \boldsymbol{z}_t\} = \sum_{i=1}^{p} \boldsymbol{\Gamma}(k-i)\, \boldsymbol{\Phi}_i^{\dagger} - \sum_{i=1}^{q} \boldsymbol{\Psi}_{i-k}\, \boldsymbol{\Sigma}\, \boldsymbol{\Theta}_i^{\dagger} \tag{3.35}$$

with $\boldsymbol{\Theta}_0 = -\boldsymbol{I}$ and $\boldsymbol{\Gamma}(k) = \sum_{i=1}^{p} \boldsymbol{\Gamma}(k-i)\, \boldsymbol{\Phi}_i^{\dagger}$ for $k = 1,2, \ldots, q$. The lag k cross-correlation matrix is given by

$$\boldsymbol{\rho}(k) = D^{-1}\boldsymbol{\Gamma}(k)D^{-1} = \rho_{ij}(k) \tag{3.36}$$

for $i, j = 1, 2, \dots, d$, where D is the diagonal matrix, in which the ith diagonal element is the variance of the ith process. Hence

$$D = \mathrm{diag}\big(\gamma_{11}(0), \gamma_{22}(0), \cdots, \gamma_{dd}(0)\big).$$

3.7.2 The first-order vector autoregressive moving average (VARMA) process

The first-order of vector autoregressive moving average process VARMA(1,1) can be written as

$$\boldsymbol{z}_t = \boldsymbol{\Phi}_1 \boldsymbol{z}_{t-1} + \boldsymbol{e}_t - \boldsymbol{\Theta}_1 \boldsymbol{e}_{t-1} \tag{3.37}$$

or

$$(\boldsymbol{I} - \boldsymbol{\Phi}_1 B)\boldsymbol{z}_t = (\boldsymbol{I} - \boldsymbol{\Theta}_1 B)\boldsymbol{e}_t$$

The VARMA(1,1) process satisfies the stationarity condition, if the solutions of the determinant equation $\|\boldsymbol{I} - \boldsymbol{\Phi}_1 B\| = 0$ are all greater than one in modulus and lie outside the unit circle, or if all the eigenvalues of $\boldsymbol{\Phi}_1$ are inside the unit circle. The MA representation of VARMA(1,1) process can be written as

$$(\boldsymbol{I} - \boldsymbol{\Phi}_1 B)(\boldsymbol{I} + \boldsymbol{\Phi}_1 B + \boldsymbol{\Phi}_1^2 B^2 + \cdots) = \boldsymbol{I}$$

hence $(\boldsymbol{I} - \boldsymbol{\Phi}_1 B)^{-1} = \boldsymbol{I} + \boldsymbol{\Phi}_1 B + \boldsymbol{\Phi}_1^2 B^2 + \cdots$, consequently, giving

$$\begin{aligned} \boldsymbol{z}_t &= (\boldsymbol{I} - \boldsymbol{\Phi}_1 B)^{-1}(\boldsymbol{I} - \boldsymbol{\Theta}_1 B)\boldsymbol{e}_t \\ &= \boldsymbol{e}_t + (\boldsymbol{\Phi}_1 - \boldsymbol{\Theta}_1)\boldsymbol{e}_{t-1} + \boldsymbol{\Phi}_1(\boldsymbol{\Phi}_1 - \boldsymbol{\Theta}_1)\boldsymbol{e}_{t-2} + \cdots \\ &= \sum_{i=0}^{\infty} \boldsymbol{\Psi}_i \, \boldsymbol{e}_{t-i} \end{aligned}$$

where $\boldsymbol{\Psi}_0 = \boldsymbol{I}$ and $\boldsymbol{\Psi}_i = \boldsymbol{\Phi}_1^{i-1}(\boldsymbol{\Phi}_1 - \boldsymbol{\Theta}_1)$ for $i \geq 1$.

The VARMA(1,1) process satisfies the invertible condition, if the solutions of the determinant equation $|\boldsymbol{I} - \boldsymbol{\Theta}_1 B| = 0$ are all outside the unit circle, or if all the eigenvalues of $\boldsymbol{\Theta}_1$ are inside the unit circle.

The covariance matrix function of the vector ARMA(1,1) process can be obtained by post-multiplying Equation 3.37 by $\boldsymbol{z}_{t-k}^{\dagger}$ and taking the expectation, giving

$$E\left[\boldsymbol{z}_t\boldsymbol{z}_{t-k}^{\dagger}\right] = \boldsymbol{\Phi}_1 E\left[\boldsymbol{z}_{t-1}\boldsymbol{z}_{t-k}^{\dagger}\right] + E\left[\boldsymbol{e}_t\boldsymbol{z}_{t-k}^{\dagger}\right] - \boldsymbol{\Theta}_1 E\left[\boldsymbol{e}_{t-1}\boldsymbol{z}_{t-k}^{\dagger}\right]$$

From the infinite MA representation

$$E[\mathbf{z}_{\mathrm{t-k}}\boldsymbol{e}_{t-i}{}^{\dagger}] = \boldsymbol{\Psi}_{i-k}\boldsymbol{\Sigma}, \qquad \text{for } i \geq k$$

Hence, for $k = 0$

$$\boldsymbol{\Gamma}(0) = \boldsymbol{\Phi}_1\boldsymbol{\Gamma}(-1) + \boldsymbol{\Sigma} - \boldsymbol{\Psi}_1\boldsymbol{\Sigma}\,\boldsymbol{\Theta}_1^{\dagger}$$

Then, $\boldsymbol{\Psi}_1 = \boldsymbol{\Phi}_1 - \boldsymbol{\Theta}_1$ and $\boldsymbol{\Gamma}(-1) = \boldsymbol{\Gamma}(1)^{\dagger}$

$$\boldsymbol{\Gamma}(0) = \boldsymbol{\Phi}_1\boldsymbol{\Gamma}(1)^{\dagger} + \boldsymbol{\Sigma} - (\boldsymbol{\Phi}_1 - \boldsymbol{\Theta}_1)\boldsymbol{\Sigma}\,\boldsymbol{\Theta}_1^{\dagger} \tag{3.38}$$

For $k = 1$

$$\boldsymbol{\Gamma}(1) = \boldsymbol{\Phi}_1\boldsymbol{\Gamma}(0) - \boldsymbol{\Theta}_1\boldsymbol{\Sigma}$$

Taking the transpose for both sides

$$\boldsymbol{\Gamma}(1)^{\dagger} = \left(\boldsymbol{\Phi}_1\boldsymbol{\Gamma}(0)\right)^{\dagger} - (\boldsymbol{\Theta}_1\boldsymbol{\Sigma})^{\dagger}$$

$$= \boldsymbol{\Gamma}(0)\boldsymbol{\Phi}_1^{\dagger} - \boldsymbol{\Sigma}\boldsymbol{\Theta}_1^{\dagger} \tag{3.39}$$

Substituting Equation 3.39 in Equation 3.38, gives

$$\boldsymbol{\Gamma}(0) = \boldsymbol{\Phi}_1\left(\boldsymbol{\Gamma}(0)\boldsymbol{\Phi}_1^{\dagger} - \boldsymbol{\Sigma}\boldsymbol{\Theta}_1^{\dagger}\right) + \boldsymbol{\Sigma} - (\boldsymbol{\Phi}_1 - \boldsymbol{\Theta}_1)\boldsymbol{\Sigma}\,\boldsymbol{\Theta}_1^{\dagger}$$

$$= \boldsymbol{\Phi}_1\boldsymbol{\Gamma}(0)\boldsymbol{\Phi}_1^{\dagger} - \boldsymbol{\Theta}_1\boldsymbol{\Sigma}\boldsymbol{\Phi}_1^{\dagger} + \boldsymbol{\Sigma} - (\boldsymbol{\Phi}_1 - \boldsymbol{\Theta}_1)\boldsymbol{\Sigma}\,\boldsymbol{\Theta}_1^{\dagger}$$

$$\boldsymbol{\Gamma}(0) - \boldsymbol{\Phi}_1\boldsymbol{\Gamma}(0)\boldsymbol{\Phi}_1^{\dagger} = \boldsymbol{\Sigma} - \boldsymbol{\Theta}_1\boldsymbol{\Sigma}\boldsymbol{\Phi}_1^{\dagger} - (\boldsymbol{\Phi}_1 - \boldsymbol{\Theta}_1)\boldsymbol{\Sigma}\,\boldsymbol{\Theta}_1^{\dagger} \tag{3.40}$$

Applying the vectorizing operation vec to both sides, gives

$$vec\left(\boldsymbol{\Gamma}(0)\right) = \left(\boldsymbol{I} - (\boldsymbol{\Phi}_1 \otimes \boldsymbol{\Phi}_1)\right)^{-1} vec\left(\boldsymbol{\Sigma} - \boldsymbol{\Theta}_1\boldsymbol{\Sigma}\boldsymbol{\Phi}_1^{\dagger} - (\boldsymbol{\Phi}_1 - \boldsymbol{\Theta}_1)\boldsymbol{\Sigma}\,\boldsymbol{\Theta}_1^{\dagger}\right) \tag{3.41}$$

For $k > 1$

$$\boldsymbol{\Gamma}(k) = \boldsymbol{\Phi}_1 \boldsymbol{\Gamma}(k-1), \qquad \text{for } k > 1 \tag{3.42}$$

The Equation 3.42 is useful to find $\boldsymbol{\Gamma}(k)$ when $k > 1$.

3.8 Model building VARMA models

The model building stages, such as, identification, estimation and diagnostic checking will be discussed for the vector of autoregressive moving average situation. For instance, identifying the VARMA process by using the sample covariance and correlation matrices. The maximum likelihood function will be used to estimate the VARMA processes.

3.8.1 Model identification of a vector time series

The procedure of model identification of a vector ARMA process follows the model building procedure in the univariate situation. In the univariate case, identification of a time series model of an ARMA process is based on the sample autocorrelation and the sample partial autocorrelation functions. In the case of a vector autoregressive process the identification is based on the covariance matrix function. Model identification can help to determine the order of a vector ARMA process; this method has been developed by researchers, such as, Zellner and Palm (1974), Wallis (1977) and Tiao, and Tsay (1989).

Identification of a vector autoregressive process of order $\boldsymbol{p}$

There are two methods to identify the vector autoregressive process of order p, which are the likelihood ratio test and the information criterion. The likelihood ratio test is used to determine the order of the VAR model, which is based on the estimates of the residual covariance matrices in the fitted models. The parameters of vector autoregressive process $\boldsymbol{\Phi}_k$ will be zero at lag k. This gives null hypothesis statistics at lag k, (Tiao and Box, 1981)

$$H_0: \boldsymbol{\Phi}_k = 0$$

$$H_1: \boldsymbol{\Phi}_k \neq 0$$

when the VAR has been fitted to the series. Then, the likelihood ratio test is given by

$$LR(k) = -(n - p - kk - 1.5)\ \ln\left(\frac{|S_{k-1}|}{|S_k|}\right) \tag{3.43}$$

where n is the number of observations of the vector of data and S_{k-1} is the residual sum of the square matrix, obtained by fitting the AR model of order $k-1$. The $LR(k)$ test asymptotically follows a chi square distribution with d^2 degree of freedom.

Another method to identify the VAR of order p is the information criterion, which can be defined in three different ways, namely

$$AIC(k) = \ln\left(\left|\widehat{\boldsymbol{\Sigma}}\right|\right) + 2r/n$$

$$BIC(k) = \ln\left(\left|\widehat{\boldsymbol{\Sigma}}\right|\right) + \ln(n)r/n$$

$$HQ(k) = \ln\left(\left|\widehat{\boldsymbol{\Sigma}}\right|\right) + 2\ln(\ln n)\, r/n$$

where r denotes the number of parameters estimated by maximum likelihood in the VARMA model and $\widehat{\boldsymbol{\Sigma}}$ is the maximum likelihood estimates of $\boldsymbol{\Sigma}$. AIC is the Akaike information criterion proposed by Akaike (1973), BIC stands for Bayesian information criterion (Schwarz 1978) and HQ is the criterion proposed by Hanna and Quinn (1979), (Quinn, 1980).

Identification of a vector moving average process of order q

The vector moving average process of order q can be identified by using the cross-correlation matrices, which was suggested by Tiao and Box (1981). The cross-correlation matrices satisfy $\boldsymbol{\rho}_i = 0$ for $i > q$. The elements of the sample cross correlation matrix at lag k is given by

$$\widehat{\boldsymbol{\rho}}_{ij}(k) = \frac{\sum_{t=1}^{n-k}\left(z_{i,t} - \bar{z}_i\right)\left(z_{j,t+k} - \bar{z}_j\right)}{\left(\sum_{t=1}^{n}\left(z_{i,t} - \bar{z}_i\right)^2 \sum_{t=1}^{n}\left(z_{j,t} - \bar{z}_j\right)^2\right)^{1/2}} \quad \text{for } i,j = 1,2,\dots,d \tag{3.44}$$

where n is the number of the observations, $\bar{z}_i$ is the sample mean and $\widehat{\boldsymbol{\rho}}_{ij}(k)$ is the lag k sample cross correlation matrix of $\boldsymbol{z}_t$. The cross-correlation matrix is zero if $\boldsymbol{z}_t$ follows vector moving average process of order q and $i > q$.

Identification of a vector autoregressive moving average process of order p and q

The parameters of vector autoregressive moving average of order p and q

$$\boldsymbol{\Phi}(B)\boldsymbol{z}_t = \boldsymbol{\Theta}(B)\boldsymbol{e}_t$$

can be identified by the patterns in the cross correlations of the residuals after a low order AR model has been fitted. For example, consider the case of a stationary ARMA(1,1) model

$$(\boldsymbol{I} - \boldsymbol{\Phi}B)\boldsymbol{z}_t = (\boldsymbol{I} - \boldsymbol{\Theta}B)\boldsymbol{e}_t \tag{3.45}$$

If an AR(1) model has been fitted to the series $\mathbf{z}_\mathrm{t}$, then the estimate is

$$\widehat{\boldsymbol{\Phi}}_{11} = \hat{\Gamma}(1)^{\dagger}\hat{\Gamma}(0)^{-1}$$

Thus, the residuals after the AR(1) model has been fitted will be

$$\begin{aligned}\hat{\mathbf{e}}_\mathrm{t} &= \mathbf{z}_\mathrm{t} - \widehat{\boldsymbol{\Phi}}_{11}\mathbf{z}_{\mathrm{t}-1} \\ &= \left(\boldsymbol{I} - \widehat{\boldsymbol{\Phi}}_{11}B\right)\end{aligned}$$

These will approximately follow the model, from Equation 3.45.

$$\hat{\mathbf{e}}_\mathrm{t} = \left(\boldsymbol{I} - \widehat{\boldsymbol{\Phi}}_{11}B\right)(\boldsymbol{I} - \boldsymbol{\Phi}B)^{-1}(\boldsymbol{I} - \boldsymbol{\Theta}B)\boldsymbol{e}_t$$

The residuals $\hat{\mathbf{e}}_\mathrm{t}$ of the sample correlations will behave approximately like a MA(1) model. Therefore, the correct identification of a vector autoregressive moving average process can be found by the examination of the residual cross correlation matrix after a AR(1) model has been fitted to the process (Tiao and Box, 1981).

3.8.2 Model estimation of a vector time series

The next step of model building is model estimation. The parameters of a vector ARMA process can be estimated by using the maximum likelihood function, the least square and Bayesian methods. These methods of estimation of a vector ARMA process of order p and q have been derived and developed by researchers, such as, Hillmer and Tiao (1979) and Nicholls and Hall (1979).

Estimate of a vector autoregressive process of order p

Consider the vector autoregressive process of order p VAR(p)

$$\boldsymbol{z}_t = \boldsymbol{\Phi}_1\boldsymbol{z}_{t-1} + \boldsymbol{\Phi}_2\boldsymbol{z}_{t-2} + \cdots + \boldsymbol{\Phi}_p\boldsymbol{z}_{t-p} + \boldsymbol{e}_t$$

where $\boldsymbol{z}_t$ is a d-dimensional vector valued time series and $\boldsymbol{e}_t$ is white noise. The log likelihood function of a VAR process of order p can be written as

$$l(\boldsymbol{\Phi}, \boldsymbol{\Sigma}) = -\frac{(n-p)}{2}\log(|\boldsymbol{\Sigma}|) - \frac{1}{2}\sum_{t=p+1}^{n} tr\left(\boldsymbol{e}_t^{\dagger}\boldsymbol{\Sigma}^{-1}\boldsymbol{e}_t\right)$$

By using the trace rules for a matrix, namely, $tr(AB) = tr(BA)$ and $tr(A+B) = tr(A) + tr(B)$, then

$$= -\frac{(n-p)}{2}\log(|\boldsymbol{\Sigma}|) - \frac{1}{2}\, tr\left(\boldsymbol{\Sigma}^{-1}\sum_{t=p+1}^{n}\left(\boldsymbol{e}_t\boldsymbol{e}_t^{\dagger}\right)\right) \tag{3.46}$$

where

$$\hat{\boldsymbol{e}}_t = \boldsymbol{z}_t - \boldsymbol{\Phi}_1\boldsymbol{z}_{t-1} - \boldsymbol{\Phi}_2\boldsymbol{z}_{t-2} - \cdots - \boldsymbol{\Phi}_p\boldsymbol{z}_{t-p}$$

and

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n-p}\sum_{t=p+1}^{n}\hat{\boldsymbol{e}}_t\hat{\boldsymbol{e}}_t^{\dagger}$$

Estimate of a vector moving average process of order q

Consider the vector moving average process of order q VMA(q)

$$\boldsymbol{z}_t = \boldsymbol{\Theta}_1\boldsymbol{e}_{t-1} + \boldsymbol{\Theta}_2\boldsymbol{e}_{t-2} + \cdots + \boldsymbol{\Theta}_q\boldsymbol{e}_{t-q} + \boldsymbol{e}_t$$

where $\boldsymbol{z}_t = (z_{1t}, z_{2t}, \ldots, z_{dt})^{\dagger}$, is a $d \times 1$ vector of variables observed at time t, $\boldsymbol{e}_t = (e_{1t}, e_{2t}, \ldots, e_{dt})^{\dagger}$, is a $d \times 1$ vector white noise process with zero mean, covariance matrix $\boldsymbol{\Sigma} = E[\boldsymbol{e}_t\boldsymbol{e}_t{}^{\dagger}]$ and $\boldsymbol{\Theta}_j$ is a $d \times d$ matrix of coefficients $(j = 1, 2, \ldots, q)$.

The exact likelihood estimation of the vector moving average can be written as

$$f(\boldsymbol{z}) = (2\pi)^{-dn/2}|\boldsymbol{\Sigma}|^{-1/2}\left|\boldsymbol{A}^{\dagger}\boldsymbol{\Sigma}^{-1}\boldsymbol{A}\right|^{-1/2}\exp\left\{\frac{-1}{2}(\boldsymbol{Bz}+\boldsymbol{A}\hat{\boldsymbol{e}}^{*})^{\dagger}\boldsymbol{\Sigma}^{-1}(\boldsymbol{Bz} + \boldsymbol{A}\hat{\boldsymbol{e}}^{*})\right\} \tag{3.47}$$

where

$$\boldsymbol{e}_t = \boldsymbol{Bz} + \boldsymbol{Ae}^{*},$$

$$\boldsymbol{e}^{*} = \left(\boldsymbol{e}_{1-q}^{\dagger}, \boldsymbol{e}_{2-q}^{\dagger}, \dots, \boldsymbol{e}_{0}^{\dagger}\right)^{\dagger},$$

$\boldsymbol{A}$ is a matrix of dimension $d(n+q) \times dq$ and $\boldsymbol{B}$ is a matrix of dimension $d(n+q) \times dn$, which are only determined by $\boldsymbol{\Theta}_1, \boldsymbol{\Theta}_2, \dots, \boldsymbol{\Theta}_q$, (Osborn, 1977).

Estimate vector autoregressive moving average process of order p and q

Consider the vector autoregressive moving average process of order p and q VARMA(p, q)

$$\boldsymbol{z}_t = \boldsymbol{\Phi}_1\boldsymbol{z}_{t-1} + \cdots + \boldsymbol{\Phi}_p\boldsymbol{z}_{t-p} + \boldsymbol{e}_t - \boldsymbol{\Theta}_1\boldsymbol{e}_{t-1} - \boldsymbol{\Theta}_2\boldsymbol{e}_{t-2} - \cdots - \boldsymbol{\Theta}_q\boldsymbol{e}_{t-q} \tag{3.48}$$

where $\boldsymbol{z}_t = (z_{1t}, z_{2t}, \dots, z_{dt})^{\dagger}$ is a $d \times 1$ vector of variables observed at time t, $\boldsymbol{e}_t$ is zero mean white noise process with covariance matrix $\boldsymbol{\Sigma}$, $\boldsymbol{\Phi}_i$ $(i = 1, 2, \dots, p)$ are $d \times d$ parameter matrices and $\boldsymbol{\Theta}_j$ is a $d \times d$ matrix of coefficients, for $(j = 1, 2, \dots, q)$. There are two methods to estimate the parameters of the VARMA model, which are the conditional likelihood method and the exact likelihood method.

The conditional likelihood method of vector ARMA process can be obtained by

$$l(\boldsymbol{\Phi}, \boldsymbol{\Theta}, \boldsymbol{\Sigma}) = -\frac{n}{2}\log(|\boldsymbol{\Sigma}|) - \frac{1}{2}\sum_{t=p+1}^{n}\left(\boldsymbol{e}_t^{\dagger}\boldsymbol{\Sigma}^{-1}\boldsymbol{e}_t\right) \tag{3.49}$$

From Equation 3.48

$$\boldsymbol{\Phi z}_t = \boldsymbol{\Theta e}_t$$

$$\boldsymbol{e}_t = \boldsymbol{\Theta}^{-1}\boldsymbol{\Phi z}_t$$

Therefore, the conditional likelihood method for vector ARMA can be written as

$$= -\frac{n}{2}\log(|\boldsymbol{\Sigma}|) - \frac{1}{2}\boldsymbol{e}_t^{\dagger}(I_n \otimes \boldsymbol{\Sigma}^{-1})\boldsymbol{e}_t \tag{3.50}$$

The conditional likelihood method has been derived by Tunnicliffe (1973), Reinsel (1979) and Anderson (1980).

The exact likelihood function of a stationary vector ARMA in Equation 3.48 has been derived by Hillmer and Tiao (1979) and Nicholls and Hall (1979). Further information for the exact likelihood function is detailed in Reinsel (1993).

3.9 Diagnostic checking of vector ARMA models

Diagnostic checking ensures the adequacy of a model in time series analysis, which can be conducted by using portmanteau tests. The portmanteau testing of VARMA models is based on the residual covariance matrices at several lags, this approach was developed by Hosking (1980). The existing portmanteau test used to examine VARMA models can be defined as (Hosking, 1980):

$$\tilde{Q}_H = n^2 \sum_{i=1}^{m} (n-i)^{-1} \mathrm{tr}\,(\boldsymbol{\Gamma}(i)\boldsymbol{\Sigma}^{-1}\boldsymbol{\Gamma}(i)^{\dagger}\boldsymbol{\Sigma}^{-1}), \tag{3.51}$$

where

$$\boldsymbol{\Gamma}(k) = n^{-1} \sum_{t=1}^{n-1} \hat{\boldsymbol{e}}_t \hat{\boldsymbol{e}}_{t-k} \quad k = 1,2,\ldots,m,$$

It has been shown by Hosking (1980) that the test statistic $\tilde{Q}_H$ is approximately distributed as chi-square with $d^2(m-p-q)$ degree of freedom.

The next chapter will investigate how the new portmanteau test $\tilde{Q}_{EXM}$ can be used with vector autoregressive moving average models. Monte Carlo experiments will be conducted to find the empirical size and power level to compare the new portmanteau test with the existing portmanteau test.

3.10 Summary

This chapter briefly discussed the extension of the Box and Jenkins methodology to a vector setting (VARMA models). An outline of the mean vector, the covariance and correlation matrix functions, the vector white noise process, the linear process of vector time series, the vector autoregressive, vector moving average, the vector autoregressive moving average processes and model building of vector ARMA time series have been provided.

Chapter 4 - The Influence Of Data Length On Testing Stationarity Of Univariate Time Series

The aim of this chapter is to examine how the number of available observations of a time series can influence its apparent stationarity (that is, its identification as being either stationary or non-stationary) as measured by two standard tests. The univariate time series case is examined. To explore this issue, time series are generated from a known statistical model, a first-order autoregressive process. Parameters are chosen that ensure that the series are theoretically stationary. The standard Dickey-Fuller test and the Augmented Dickey-Fuller test are used to determine whether the series of observations produced are stationary or non-stationary. Monte Carlo experiments are undertaken using the R program for various model parameters and lengths of series, and each simulation is repeated 10,000 times.

4.1 Introduction

A critical factor in fitting a model to a time series, using the Box and Jenkins methodology (1970), is identifying whether or not the time series is stationary or non-stationary. In the case of the data being non-stationary, extra steps need to be undertaken to make the data stationary, typically, by differencing the data or by some transformation of the data. The consequence of incorrectly identifying a stationary time series as being non-stationary is that it will lead to the data being altered inappropriately and the wrong model being fitted.

One factor that has a strong influence on the ability to identify a time series as being stationary or non-stationary is the length of data available. Note that, a process is stationary when its joint probability distribution does not change with time (Box and Jenkins methodology, 1970). A non-stationary process would be expected to exhibit deterministic trends, random walks and other non-stationary behaviour. The difficulty in identification arises from the fact that a stationary process may, for a short period, exhibit non-stationary behaviour, and conversely, a non-stationary process may exhibit stationary behaviour for a short period. Identification of stationarity can be achieved by the application of standard tests.

To examine the relationship between the length of a series of data and the reliability of the identification of its stationarity a simulation study was conducted. In outline, the procedure was as follows. First, stationary data were generated from a known statistical process of a univariate time series, that is, an autoregressive process. Next, different lengths of the simulated data were tested using the standard tests of stationarity to examine the apparent stationarity of the time series. The standard tests employed were; the Dickey-Fuller test (DF) and the Augmented Dickey-Fuller test (ADF).

When data are generated by a known statistical model (or collected from a real-world process), the length of data available may affect the apparent stationarity of the time series produced. For example, when considering a first-order autoregressive process, a short series of data with a value of parameter, such as 0.9, may produce a non-stationary time series, whereas a long series of data with the same parameter value may pass a test for stationarity.

4.2 Testing of non-stationary time series

There are many tests that have been developed to identify whether a time series is stationary or non-stationary. Unit root tests are widely used to test for stationarity in time series for different kinds of data, such as, stationary data, stationary data with a drift term, and stationary data with drift and trend terms. The null hypothesis of a unit root test for a first-order autoregressive process is, H_0: $|\phi_1|$ =1, the time series is non-stationary (has a unit root). The alternative hypothesis is, H_1: $|\phi_1| < 1$, the series is stationary (does not have a unit root).

Phillips and Perron in 1988 introduced a non-parametric modification to the standard Dickey-Fuller test of a unit root that they used to test for stationary behaviour in a time series analysis. The Phillips-Perron (PP) test deals with serial correlation. One advantage of the PP test is that the user does not have to specific the lag length, (Phillips and Perron,1988). Kwiatkowski, Phillips, Schmidt and Shin (KPSS) in 1992 derived another form of the null hypothesis test in time series analysis versus alternative of unit root. The series of observations of the KPSS test is expressed as the sum of the deterministic trend, a random walk and a stationary error term (Kwiatkowski, Phillips, Schmidt and Shin, 1992). Eliot, Rothenberg and Stock in 1996 introduced another modification of the Augmented Dickey-Fuller test in which the data are detrended so that explanatory variables are “taken out” of the data prior to running the test regression, it is known as the ERS test. One advantage of this test statistic is that it improves the power when the time series has unknown mean or trend, (Eliot, Rothenberg and Stock,1996). Perron and Ng in 1996 used the ADF-GLS detrended procedure data to create efficient versions of the modified form of Phillips and Perron, (Perron and Ng, 1996). Ng and Perron (2001) suggested the modified information criteria MIC for selecting the max lag. It is based on the Akaike information criteria AIC (Akaike, 1973) and the Schwarz information criteria BIC (Schwarz, 1978). Nason in 2013 introduces a new test for second-order stationarity that detects different kinds of departures from stationarity. The new test is also computationally fast, designed to work with Gaussian and a wide range of non-Gaussian time series, and can locate non-stationarities in time series, (Nason, 2013).

4.2.1 The Dickey-Fuller (DF) test

The Dickey-Fuller test (Dickey, Fuller, 1979) is a unit root test for non-stationarity of a time series. In 1979, Dickey and Fuller considered three different regression equations that are based on a first-order autoregressive process, which can be used to test a non-stationary time series, namely the test for a unit root, the test for a unit root with drift, and the test for a unit root with drift and a deterministic time trend. The first-order autoregressive process may be written as

$$z_t = \phi_1 z_{t-1} + e_t \tag{4.1}$$

where z_t is a value of a time series at time t, ϕ_1 is a real number and e_t is a white noise of mean zero and constant variance σ_e^2. Equation 4.1 can be transformed to

$$\nabla z_t = (\phi_1 - 1) z_{t-1} + e_t$$

where ∇ is the differencing operator.

Dickey and Fuller constructed a statistic by analogy to the t-ratio test for the estimate of ϕ_1, which is estimated by the least squared method,

$$\hat{\phi}_1 = \frac{\sum_{t=1}^{n} z_t z_{t-1}}{\sum_{t=1}^{n} z_{t-1}^2}$$

then,

$$DF = \frac{\hat{\phi}_1 - 1}{SE(\hat{\phi}_1)} \tag{4.2}$$

where $\mathrm{SE}(\hat{\phi}_1)$ is the standard error of the estimated parameter $\hat{\phi}_1$. The null and alternative hypotheses of the DF test are

$$\mathrm{H}_0: |\phi_1| = 1$$
$$\mathrm{H}_1: |\phi_1| < 1$$

The distribution of the DF test statistic was investigated by Dickey (1976), as an analytical description was not possible, he used simulations to calculate the critical values, see Appendix E for an outline of the method employed. He provided tables of the critical values of the DF test statistic's distribution for the three cases of: stationary, stationary with drift, and stationary with drift and trend time series.

The first-order autoregressive process with drift can be written as

$$z_t = \mu_0 + \phi_1 z_{t-1} + e_t \tag{4.3}$$

then, the unit root test with drift can be written as

$$\nabla z_t = \mu_0 + (\phi_1 - 1) z_{t-1} + e_t$$

The first-order autoregressive process with drift and a deterministic time trend can be written as

$$z_t = \mu_0 + \mu_1 t + \phi_1 z_{t-1} + e_t \tag{4.4}$$

So, the unit root test with drift and a deterministic time trend is

$$\nabla z_t = \mu_0 + \mu_1 t + (\phi_1 - 1) z_{t-1} + e_t$$

where $\mu_0 + \mu_1 t$ is a deterministic linear trend.

This provides three tests for data, for the cases where they are either stationary, stationary with drift, or stationary with a linear trend. In the rest of the chapter these three tests will be referred to as the DF test, the DF drift test and the DF trend test.

4.2.2 Augmented Dickey-Fuller test (ADF)

Dickey and Fuller (1981) generalized the DF test and applied it to the AR (p) process. This is named the Augmented Dickey-Fuller test (ADF). Equation 1.1 can be written using summation notation as

$$z_t = \sum_{i=1}^{p} \phi_i \, z_{t-i} + e_t$$

this can be transformed (Dickey and Fuller (1981) to

$$z_t = \tau z_{t-1} + \sum_{i=1}^{p} \phi_i \, \nabla z_{t-i} + e_t$$

hence,

$$\nabla z_t = (\tau - 1) z_{t-1} + \sum_{i=1}^{p} \phi_i \, \nabla z_{t-i} + e_t \tag{4.5}$$

where τ is the sum of the autoregressive coefficients, that is

$$\tau = \sum_{i=1}^{p} \phi_i$$

The test statistic for the Augmented Dickey-Fuller test (ADF) is

$$ADF = \frac{\hat{\tau} - 1}{SE(\hat{\tau})} \tag{4.6}$$

where $SE(\hat{\tau}.)$ is the standard error of the estimated parameter $\hat{\tau}$. The null and alternative hypotheses of the ADF test is

$$H_0 \colon |\hat{\tau}| = 1$$
$$H_1 \colon |\hat{\tau}| < 1$$

where H_0 is the null hypothesis (has unit root and non-stationary) and H_1 is the alternative hypothesis (stationary).

The distribution of the ADF test statistic was investigated by Dickey and Fuller (1981), they used simulations to calculate the critical values following the method of Dickey (1976), see Appendix E. They provided tables of the critical values of the ADF test statistic's distribution for the three cases of: stationary, stationary with drift, and stationary with drift and trend time series.

The autoregressive process of order p with drift μ_0 can be written as

$$z_t = \mu_0 + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} + e_t \tag{4.7}$$

then, the unit root test with drift can be written as

$$\nabla z_t = \mu_0 + (\tau - 1) z_{t-1} + \sum_{i=1}^{p} \phi_i \nabla z_{t-i} + e_t$$

The autoregressive process of order p with a deterministic time linear trend $\mu_0 + \mu_1 t$ can be written as

$$z_t = \mu_0 + \mu_1 t + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} + e_t \tag{4.8}$$

So, the unit root test with deterministic time linear trend is

$$\nabla z_t = \mu_0 + \mu_1 t + (\tau - 1) z_{t-1} + \sum_{i=1}^{p} \phi_i \nabla z_{t-i} + e_t$$

The critical values of this test statistic were calculated by simulation and can be obtained in Dickey (1976) and Fuller (1976). If the test statistic value is greater than the critical value, then the null hypothesis will be accepted, it means that the time series has a unit root and is non-stationary. If this is the case, the data will need to be transformed to obtain a stationary series.

An important issue of the ADF test is the specification of the lag length. If the lag length is too small then the remaining serial correlation in the error will bias the test, if the lag length is too large the power of the test will suffer (Schwert, 1989). A useful equation for determining the lag length that gives the test with the most power, as suggested by Schwert, is

$$l12 = \left[12\left(\frac{n}{100}\right)^{1/4}\right] \tag{4.9}$$

This gives three further tests for data that are either stationary, stationary with drift, or stationary with a linear trend. In the rest of this thesis these will be referred to as the ADF test, the ADF drift test, and ADF trend test respectively. In each case, the lag length employed is determined by Equation 4.9 as recommended by Schwert.

4.3 Monte Carlo experiment

The aim of this simulation is to show how the length of the series of observations will affect the apparent stationarity of the time series produced. A Monte Carlo experiment was conducted with 10,000 replications to simulate different lengths of series, namely, *n* = 25, 50, 75, 100, 250, 500, 750 and 1000 observations. The series were produced using the R language. Normally distributed N(0,1) pseudo random numbers were generated using the Mersenne-Twister generator (Matsumoto and Nishimura 1998). Then an AR(1) process $\mathrm{z_t} - \phi_1\, \mathrm{z_{t-1}} = \mathrm{e_t}$ was used to generate the data to be tested, using different positive and negative values of parameter ϕ_1= ±0.1, ±0.2, ±0.3, ±0.4, ±0.5, ±0.6, ±0.7, ±0.8, ±0.9, and ±0.99. Versions of the AR(1) process with drift $z_t = \mu_0 + \phi_1\, z_{t-1} + e_t$ and with drift and trend $z_t = \mu_0 + \mu_1 t + \phi_1\, z_{t-1} + e_t$ were also produced using the values $\mu_0 = \mu_1 = 0.5$ The DF and ADF tests were then used to determine whether the time series produced were stationary or non-stationary.

4.3.1 The steps of the Monte Carlo experiment to test the length of series observations

The steps of the Monte Carlo experiment using the DF and ADF tests to examine the stationarity of an AR(1) process, are:

1. Generate 1000 points of data from a Normal distribution (e_t white noise).
2. Use the e_t values to generate observations from an AR(1) process with parameter ϕ_1= ±0.1, ±0.2, ±0.3, ±0.4, ±0.5, ±0.6, ±0.7, ±0.8, ±0.9, and ±0.99.

3. Use the AR(1) data to produce three versions of the time series that are, stationary, stationary with drift $\mu_0 = 0.5$, and stationary with drift and a deterministic trend $\mu_0 = \mu_1 = 0.5$.
4. For each of the 3 time series in step 3, select the first *n* data points, where *n* = 25, 50, 75, 100, 250, 500, 750, 1000.
5. Test, at the 0.05 significance level, all the time series in step 4 using the DF and ADF tests, using the appropriate version of the tests (stationary, stationary with drift, or stationary with drift and trend).
6. Count the number of non-stationary series identified by both tests separately.
7. Repeat 10,000 times from 1-6.

Note, in the simulation result presented, values of μ_0 and $\mu_1 = 0.5$ were used. Simulations using values of μ_0 and $\mu_1 = 0.1, 0.3, 0.7$ and 0.9 were also conducted and gave the same results, but these not presented.

4.3.2 Generated data under an AR(1) process with positive values of parameters

Figure 4.1 gives the number of time series from an AR(1) process, identified as being non-stationary by the DF test, where ϕ_1 varies from 0.1 to 0.99 and different lengths of time series are examined. For the shortest time series (*n* = 25) the number identified as non-stationary is about 6% when $\phi_1 = 0.5$ and increases markedly for larger values of ϕ_1. For the medium lengths of time series (*n* = 50, 75, 100) the number identified as non-stationary increases when the values of ϕ_1 reaches 0.8, 0.9, and 0.9 respectively, with at least 20% identified as non-stationary in each case. For the longest lengths of data, such as *n* = 250, 500, 750, 1000 the number of non-stationary time series identified is near zero for all values of $\phi_1 < 0.99$.

For any given value of ϕ_1 the number of time series identified as being non-stationary decreases as the length of the time series increases. In the case of values of the parameter ϕ_1 very near to 1, such as, $\phi_1 = 0.99$, most of the time series are identified by the DF test as non-stationary, even for the longest series examined, which is to be expected, as at $\phi_1 = 1$ the AR(1) process (Equation 4.2) is no longer stationary

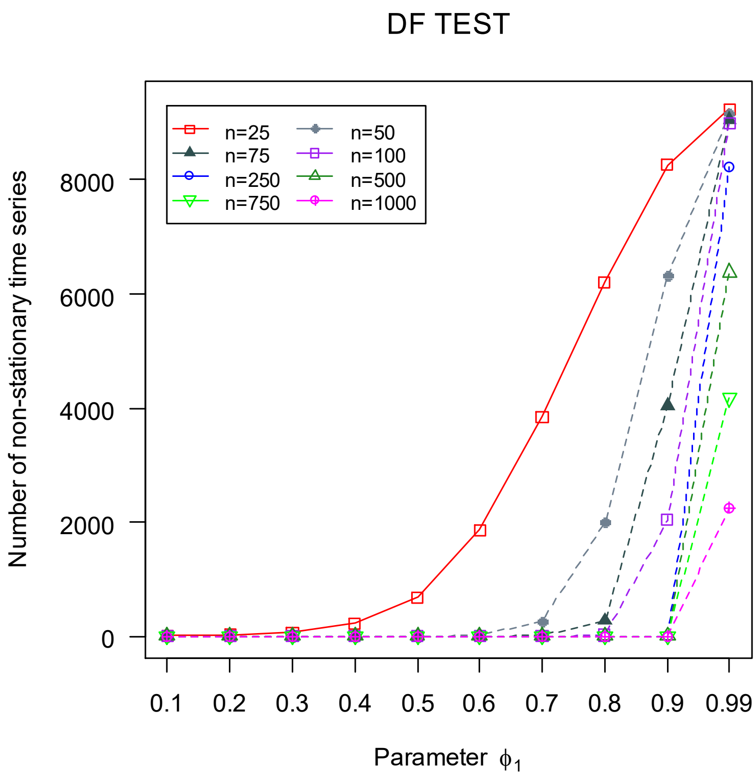

Figure 4.1 The number of series identified as non-stationary by the DF test, data generated under an AR(1) process, $z_t = \phi_1 z_{t-1} + e_t$, using a range of positive parameters ϕ_1 and different lengths of data.

Figure 4.2 gives the number of time series from an AR(1) process identified as being non-stationary by the ADF test. When the length of the times series is 100 or less, at least 20% of the series are identified as being non-stationary, irrespective of the value of ϕ_1. For the time series of length 250, the number of series identified is under 5% for all values of ϕ_1 below 0.9, at which point it is approximately 11%. When time series are generated with the longest lengths (n = 500, 750, and 1000) then all series produced are stationary except when the value of the parameter is very close to 1, i.e., $\phi_1 = 0.99$.

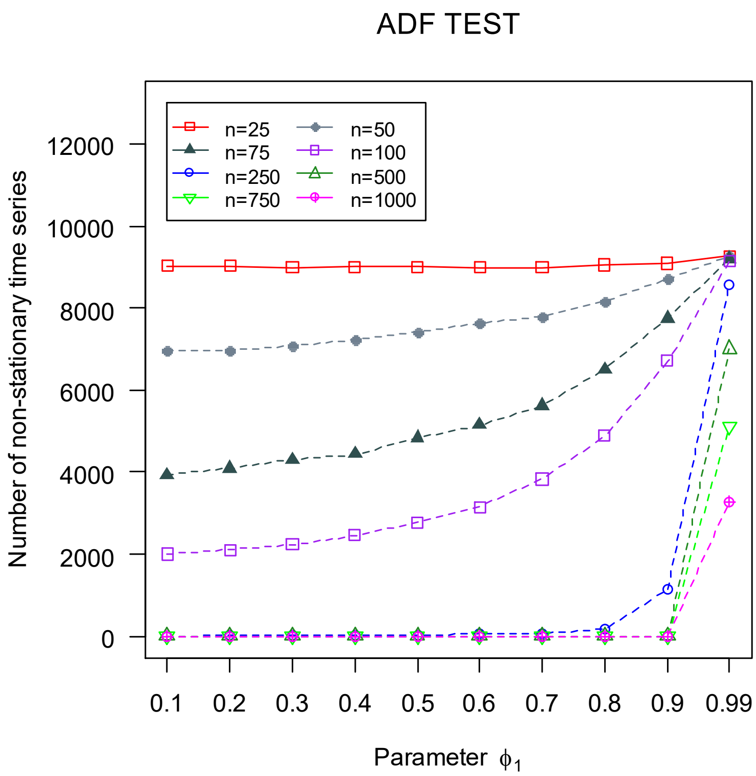

Figure 4.2 The number of series identified as non-stationary by the ADF test, data generated under an AR(1) process, $z_t = \phi_1 z_{t-1} + e_t$, using a range of positive parameters ϕ_1 and different lengths of data.

Figure 4.3 gives the number of time series from an AR(1) process with drift (drift parameter value $\mu_0 = 0.5$) identified as being non-stationary by the DF drift test. For the very shortest length of time series (n = 25) the number identified as non-stationary is above 5% for all values of ϕ_1 larger than 0.1. For medium lengths of time series (n = 50, 75, and 100) the number of non-stationary time series is above 5% for values of ϕ_1 above 0.6, 0.7 and 0.8 respectively. For the longest lengths of data (n = 250, 500, 750 and 1000) the number of non-stationary time series is near zero for all values of ϕ_1 below 0.99.

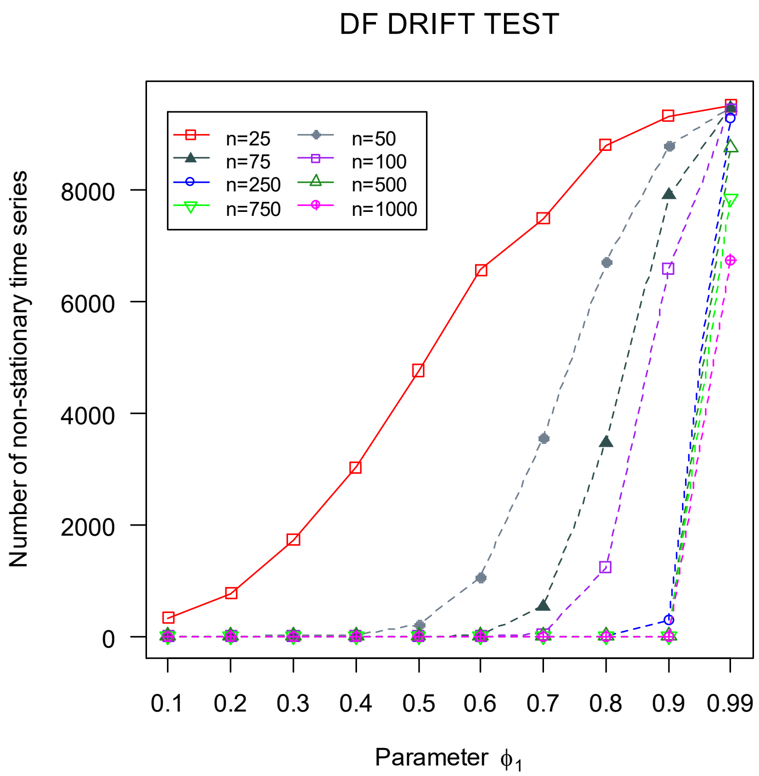

Figure 4.3 The number of series identified as non-stationary by the DF drift test, data generated under an AR(1) process with drift, $z_t = \mu_0 + \phi_1 z_{t-1} + e_t$, using a range of positive parameters ϕ_1 and different lengths of data.

Figure 4.4 gives the number of time series from an AR(1) process with drift (drift parameter value $\mu_0 = 0.5$) identified as being non-stationary by the ADF drift test. For the lengths of time series (n = 25, 50, 75, and 100) over 20% of the series are identified as being non-stationary for all vales of ϕ_1. For data lengths 250 and 500 the number of non-stationary time series rises above 5% for values of ϕ_1 above 0.1 and 0.8 respectively. For longer lengths of time series (n = 750, and 1000) the number of series identified as non-stationary is nearly zero for all vales of ϕ_1 below 0.9.

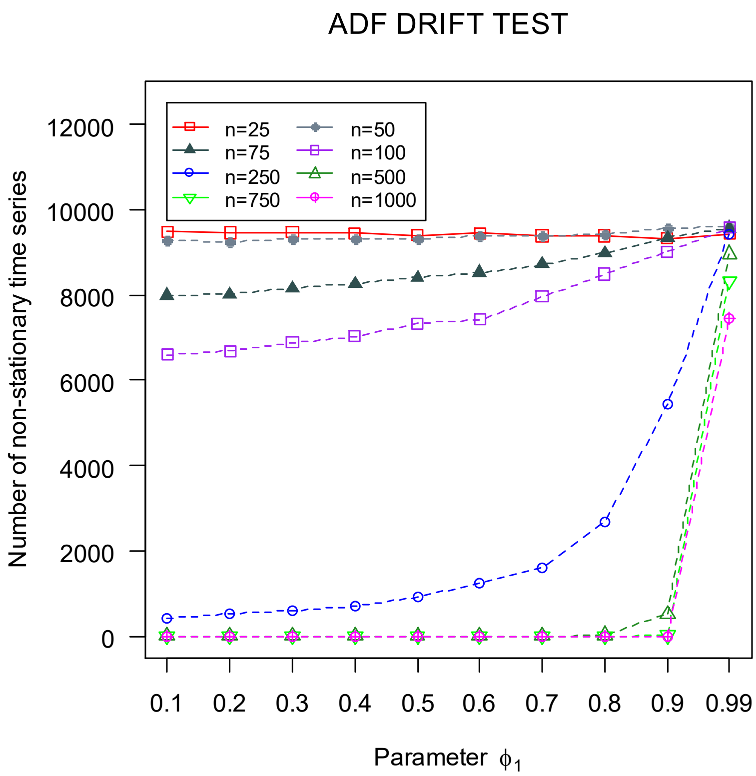

Figure 4.4 The number of series identified as non-stationary by the ADF drift test, data generated under an AR(1) process with drift, $z_t = \mu_0 + \phi_1 z_{t-1} + e_t$, using a range of positive parameters ϕ_1 and different lengths of data.

Figure 4.5 gives the number of time series from an AR(1) process with drift and trend (parameter values $\mu_0 = 0.5$ and $\mu_1 = 0.5$) identified as being non-stationary by the DF trend test. For the shortest length of time series ($n = 25$) the number of non-stationary time series is above 5%, even for the lowest value of the parameter examined $\phi_1 = 0.1$. For time series of lengths 50, 75 100 and 250 the number of series identified as being non-stationary is below 5% for all values of ϕ_1 up to 0.4, 0.6, 0.7, and 0.8 respectively. For the longest length series (n = 500, 750, and 1000) approximately zero percent of the series are identified as being non-stationary for all values of ϕ_1 below 0.99.

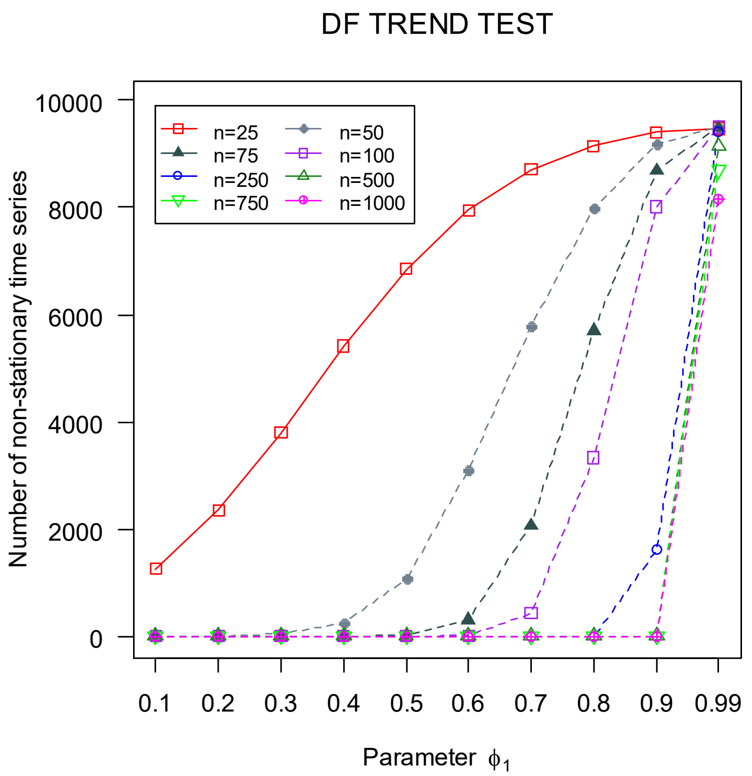

Figure 4.5 The number of series identified as non-stationary by the DF trend test, data generated under an AR(1) process with drift and trend, $z_t = \mu_0 + \mu_1 t + \phi_1 z_{t-1} + e_t$, using a range of positive parameters ϕ_1 and different lengths of data.

Figure 4.6 gives the number of time series from an AR(1) process with drift and trend (parameter values μ_0 = 0.5 and μ_1 = 0.5) identified as being non-stationary by the ADF trend test. For time series of lengths 25, 50, 75, 100 and 250 the number of non-stationary time series is above 5% for all values of ϕ_1. When the length of the time series is 500 the number identified as non-stationary is below 5% for value of ϕ_1 up to 0.9. For the series of longer lengths (n = 750, and 1000) the number of non-stationary time series is near zero for all values of ϕ_1 below 0.99.

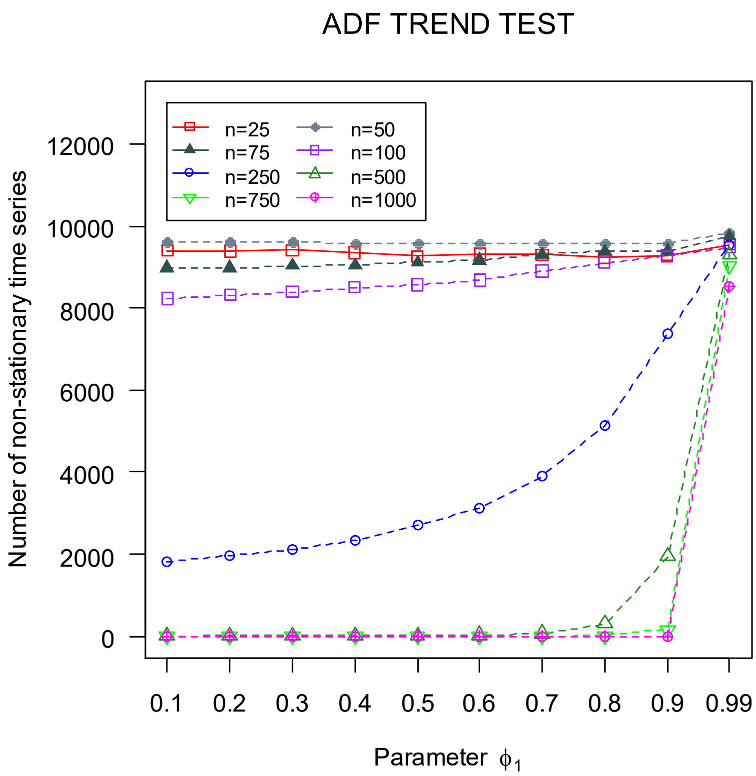

Figure 4.6 The number of series identified as non-stationary by the ADF trend test, data generated under an AR(1) process with drift and trend, $z_t = \mu_0 + \mu_1 t + \phi_1 z_{t-1} + e_t$, using a range of positive parameters ϕ_1 and different lengths of data.

4.3.3 Generated data under an AR(1) process with negative values of parameters

When data are generated under an AR(1) process with negative values of parameter ϕ_1, then the DF test identifies all the time series as being stationary irrespective of the length of the series. Similarly, for an AR(1) process with drift, all cases are identified by the DF drift test as stationary. Also, in the case of the AR(1) process with drift and trend, all cases are identified as stationary by the DF trend test.

Figure 4.7 gives the number of time series from an AR(1) process with negative values of parameter ϕ_1 that have been identified as being non-stationary by the ADF test. For time series of lengths 25, 50, 75 and 100 the number of non-stationary series are not affected by the values of parameters, but instead remains approximately constant and dependent on the length of the time series. The percentage of non-stationary series identified as non-stationary decreases with increasing length of series, and for series of length 25, 50, 75 and 100 it is approximately 87%, 63%, 33% and 15% respectively. For longer time series ($n >$ 100) all the series are identified as stationary.

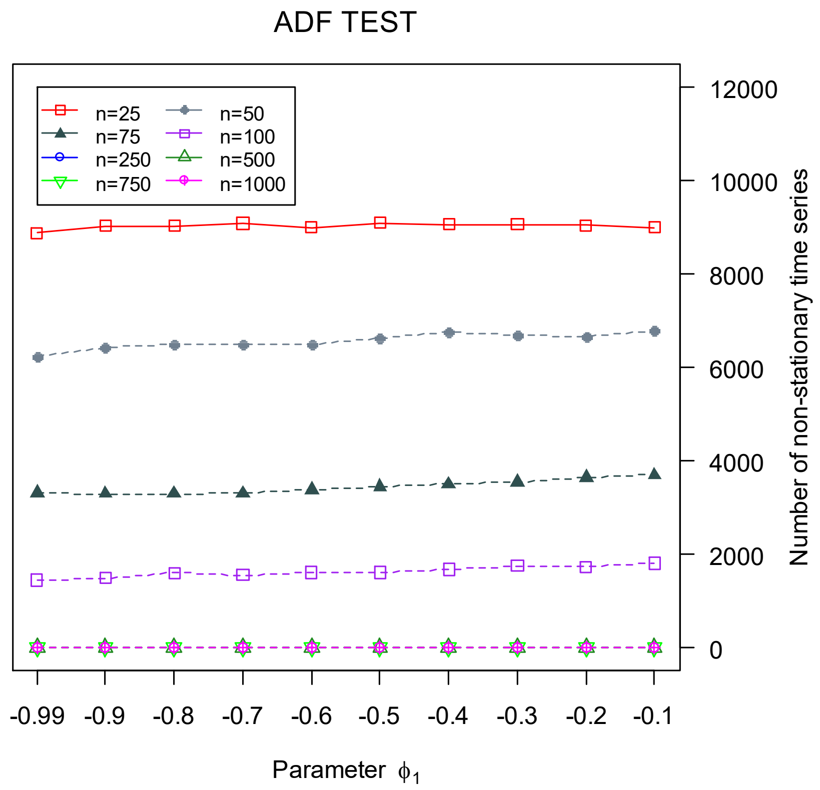

Figure 4.7 The number of series identified as non-stationary by the ADF test, data generated under an AR(1) process, $z_t = \phi_1 z_{t-1} + e_t$, using a range of negative parameters ϕ_1 and different lengths of data.

Figure 4.8 gives the number of time series from an AR(1) process with drift (drift parameter value $\mu_0 = 0.5$) and with negative values of parameter ϕ_1 that have been identified as being non-stationary by the ADF drift test. For time series of lengths 25, 50, 75 and 100 the number of non-stationary series are not affected by the values of the parameter, but instead remains approximately constant and dependent on the length of the time series. The percentage of non-stationary series identified as non-stationary decreases with increasing length of series, and for series of lengths 25, 50 75 and 100 it is approximately 95%, 90%, 78% and 60% respectively. For longer time series ($n > 100$) all the series are identified as stationary.

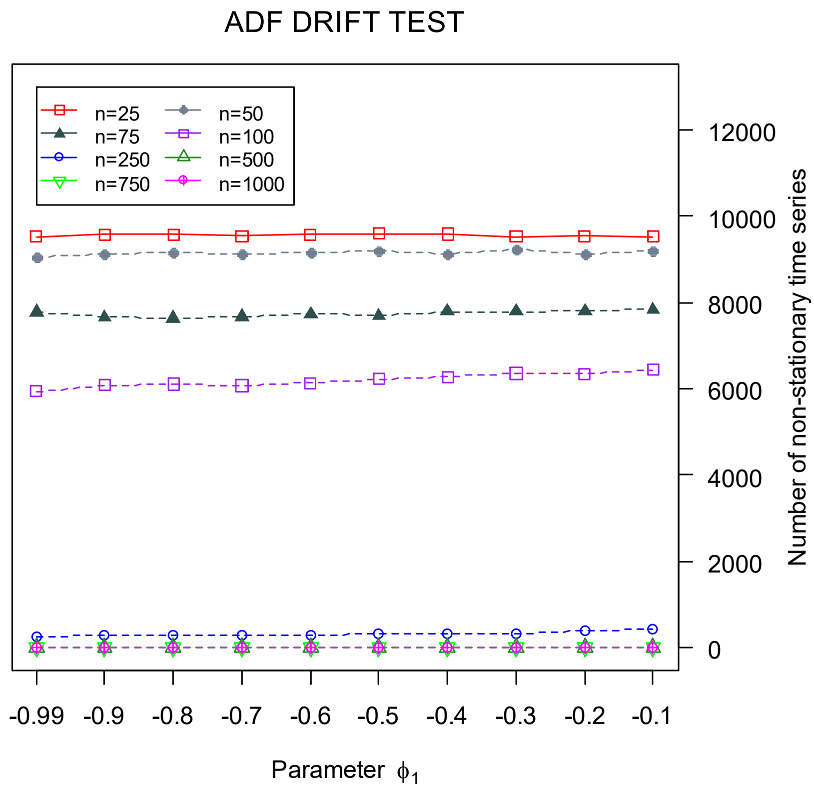

Figure 4.8 The number of series identified as non-stationary by the ADF drift test, data generated under an AR(1) process with drift, $z_t = \mu_0 + \phi_1 z_{t-1} + e_t$, using a range of negative parameters ϕ_1 and different lengths of data.

Figure 4.9 gives the number of time series from an AR(1) process with drift and trend (parameter values $\mu_0 = 0.5$ and $\mu_1 = 0.5$), and with negative values of parameter ϕ_1 that have been identified as being non-stationary by the ADF trend test. For time series of lengths 25, 50, 75, 100 and 250 the number of non-stationary series are not affected by the values of the parameter, but instead remains approximately constant and dependent on the length of the time series. The percentage of non-stationary series identified as non-stationary decreases with increasing length of series, and for series of lengths 25, 50 75, 100 and 250, it is approximately 95%, 95%, 88%, 79% and 13% respectively. For longer time series (n > 250) all the series are identified as stationary.

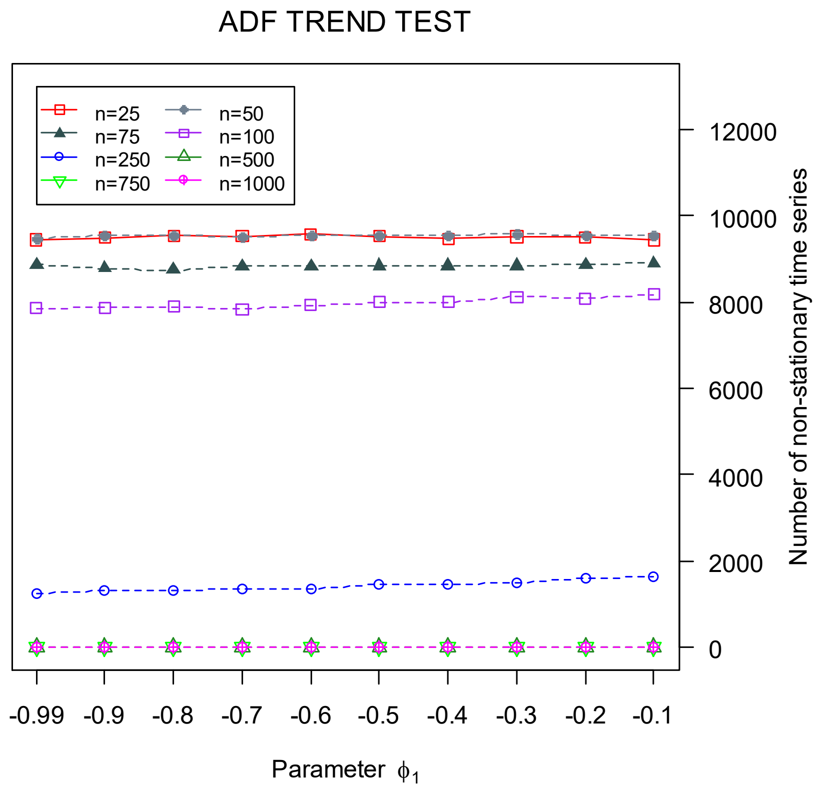

Figure 4.9 The number of series identified as non-stationary by the ADF trend test, data generated under an AR(1) process with drift and trend, $z_t = \mu_0 + \mu_1 t + \phi_1 z_{t-1} + e_t$, using a range of negative parameters ϕ_1 and different lengths of data.

4.4 Data generated under an AR(2) process with different values of parameters

A Monte Carlo experiment was conducted with 10,000 replications to simulate different lengths of series, namely, n = 25, 50, 75, 100, 250, 500, 750 and 1000 observations. Then an AR(2) process $z_t - \phi_1 z_{t-1} - \phi_2 z_{t-2} = e_t$ was used to generate the data to be tested, using different positive values of parameters ϕ_1 and ϕ_2. All combinations of parameters $\phi_1 = 0.1, 0.2, \ldots, 0.9$ and $\phi_2 = 0.1, 0.2, \ldots, 0.9$ were simulated, subject to the stationarity condition of an AR(2) process, namely, $\phi_1 + \phi_2 < 1$. Versions of the AR(2) process with drift $z_t = \mu_0 + \phi_1 z_{t-1} + \phi_2 z_{t-2} + e_t$ and with drift and trend $z_t = \mu_0 + \mu_1 t + \phi_1 z_{t-1} + \phi_2 z_{t-2} + e_t$ were also produced using the values $\mu_0 = \mu_1 = 0.5$, and the above combinations of parameters ϕ_1 and ϕ_2. All the time series were then tested by the DF tests and the ADF tests.

4.4.1 AR(2) processes tested for stationarity by the DF tests.

The DF tests was used to determine whether the time series produced were stationary or non-stationary. Figure 4.10 shows the simulation results for two representative values of ϕ_1, namely, 0.1 and 0.5, full results can be found in Appendix D.

Figure 4.10a gives the number of non-stationary time series when data are generated by the AR(2) process $z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + e_t$, with ϕ_1 = 0.1 and where ϕ_2 takes the values 0.1 to 0.9, tested by the DF test. For data lengths 25 most of the series are non-stationary when the values of ϕ_2 are greater than 0.6. For the data lengths equal or greater than 75, all series are non-stationary when the values of parameter ϕ_2 equal or greater than 0.8. Figure 4.10b, gives the number of non-stationary time series when $\phi_1 = 0.5$, and ϕ_2 varies. For the data lengths 50 to 75 most of the series are stationary when the values of ϕ_2 are less than or equal to 0.2. For data lengths equal or greater than 100 the number of non-stationary series start to increase rapidly when the values of parameter ϕ_2 reaches 0.5 or greater.

Figure 4.10c, shows data generated by the AR(2) process $z_t = \mu_0 + \phi_1 z_{t-1} + \phi_2 z_{t-2} + e_t$, with ϕ_1 = 0.1 and where ϕ_2 varies, tested by the DF drift test. For the data lengths 25 most of the series are stationary when the values of ϕ_2 are less than or equal to 0.2. For data lengths 50 to 100 the number of non-stationary series start to increase rapidly when the values of parameter ϕ_2 reaches 0.9. Figure 4.10d presents data generated by the AR(2) process with $\phi_1 = 0.5$, and where ϕ_2 varies. For the data lengths equal to 75 most of the series are non-stationary when the values of ϕ_2 greater than 0.2. For the data lengths equal to 100 majority of the series are non-stationary when the values of ϕ_2 greater than 0.3. For the data lengths equal or greater to 250 the number of non-stationary series start to increase rapidly for values of parameter ϕ_2 of 0.5 or greater.

Figure 4.10e, provides data generated by the AR(2) process $z_t = \mu_0 + \mu_1 t + \phi_1 z_{t-1} + \phi_2 z_{t-2} + e_t$, with ϕ_1 = 0.1 and where ϕ_2 varies, tested by the DF trend test. For the data lengths less than or equal 100, most of the series are non-stationary. For the data lengths 25 most of the series non-stationary with different values of ϕ_2. For the data lengths equal or greater than 50 most of the series are stationary with different values of ϕ_2. Figure 4.10f, presents data generated by the AR(2) process, with ϕ_1 = 0.5 and where ϕ_2 varies. For the data lengths equal to 75 most of the series are non-stationary when the values of ϕ_2 greater than 0.2. For the data lengths equal to 100 majority of the series are non-stationary when the values of ϕ_2 greater than 0.3. For the data lengths equal or greater to 250 the number of non-stationary series start to increase rapidly for values of parameter ϕ_2 0.5 or greater.

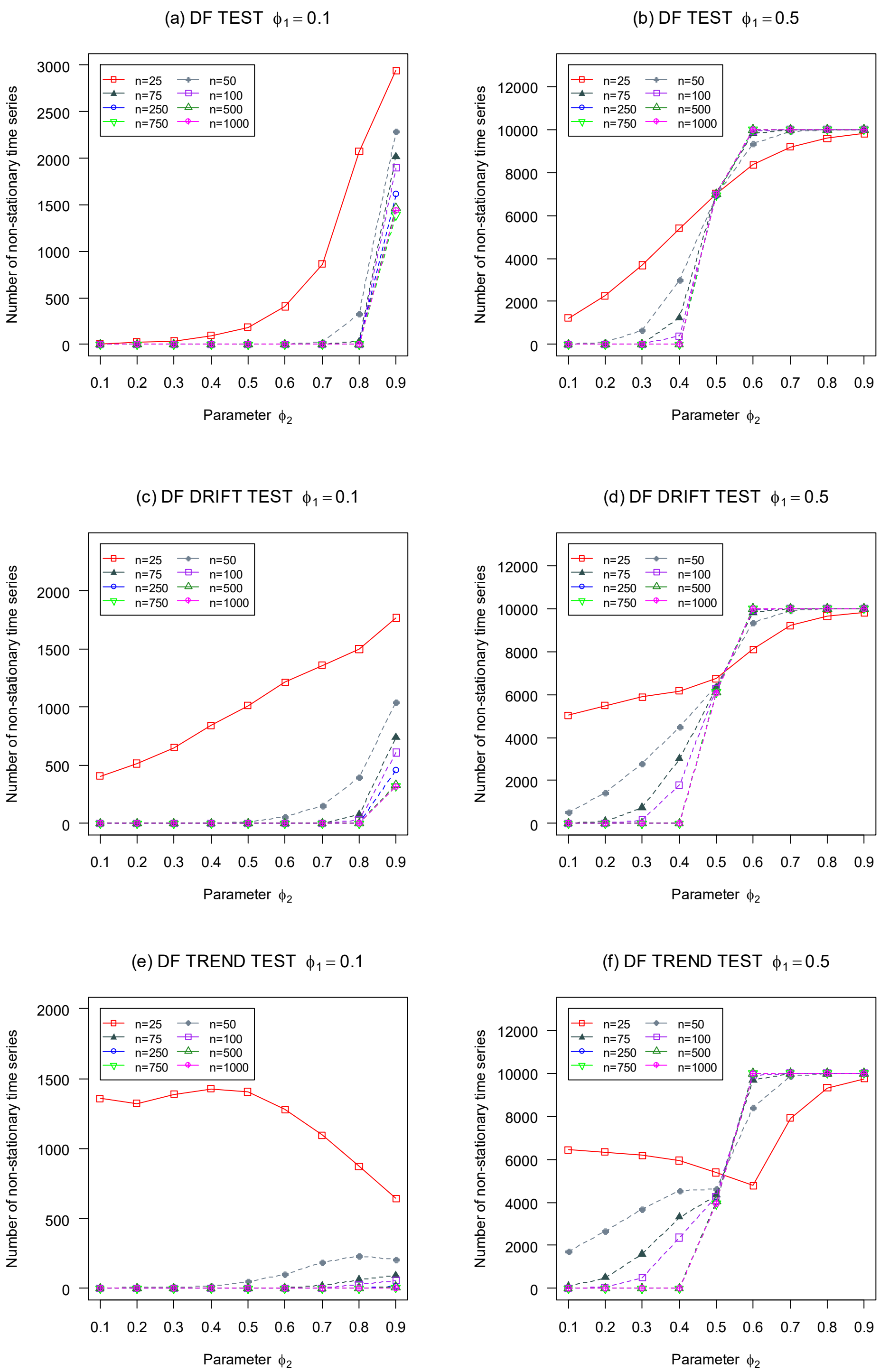

Figure 4.10 The number of series identified as non-stationary by the DF test, DF drift test and DF trend test, data generated under AR(2) process.

4.4.2 AR(2) processes tested for stationarity by the ADF tests.

The ADF tests was used to determine whether the time series produced were stationary or non-stationary. Figure 4.11 shows the simulation results for two representative values of ϕ_1, namely, 0.1 and 0.5, full results can be found in Appendix D.

Figure 4.11a gives the number of non-stationary time series when data generated by the AR(2) process $z_t - \phi_1 z_{t-1} - \phi_2 z_{t-2} = e_t$, with ϕ_1 = 0.1 and where ϕ_2 takes the values 0.1 to 0.9, tested by the ADF test. For data lengths less than or equal to 50 most of the series are non-stationary. For the data lengths equal or greater than 250, virtually all the series are stationary until the parameter ϕ_2 reaches to 0.6, with the number of non-stationary series increasing rapidly after this point. Figure 4.11b, gives the number of non-stationary time series when $\phi_1 = 0.5$, and ϕ_2 varies. For the data lengths less than or equal to 75 the majority of the series are non-stationary. For data lengths equal or greater than 250 the number of non-stationary series start to increase rapidly when the values of parameter ϕ_2 reaches 0.4 or greater.

Figure 4.11c, shows data generated by the AR(2) process $z_t = \mu_0 + \phi_1 z_{t-1} + \phi_2 z_{t-2} + e_t$, with ϕ_1 = 0.1 and where ϕ_2 varies, tested by the ADF drift test. For data lengths less than or equal to 100, most of the series are non-stationary. For data lengths equal or greater than 250, the number of non-stationary series start to increase for values of parameter ϕ_2 = 0.7 or greater. Figure 4.11d, presents data generated by the AR (2) process with $\phi_1 =$ 0.5, and where ϕ_2 varies. For the data lengths less than 250, most of the series are non-stationary. For the data lengths greater than 250, the number of non-stationary series start to increase rapidly for values of parameter ϕ_2 of 0.4 or greater.

Figure 4.11e, provides data generated by the AR(2) process $z_t = \mu_0 + \mu_1 t + \phi_1 z_{t-1} + \phi_2 z_{t-2} + e_t$, with ϕ_1 = 0.1 and where ϕ_2 varies, tested by the ADF trend test. For the data lengths less than or equal 100, most of the series are non-stationary. For data lengths greater than 500, the number of non-stationary series start to increase rapidly for values of parameter ϕ_2 = 0.6 or greater. Figure 4.11f, presents data generated by the AR(2) process, with ϕ_1 = 0.5 and where ϕ_2 varies. For data lengths less than or equal to 100, most of the series are non-stationary. For the data lengths greater then 250, the number of non-stationary series start to increase rapidly for values of parameter ϕ_2 = 0.3 or greater.

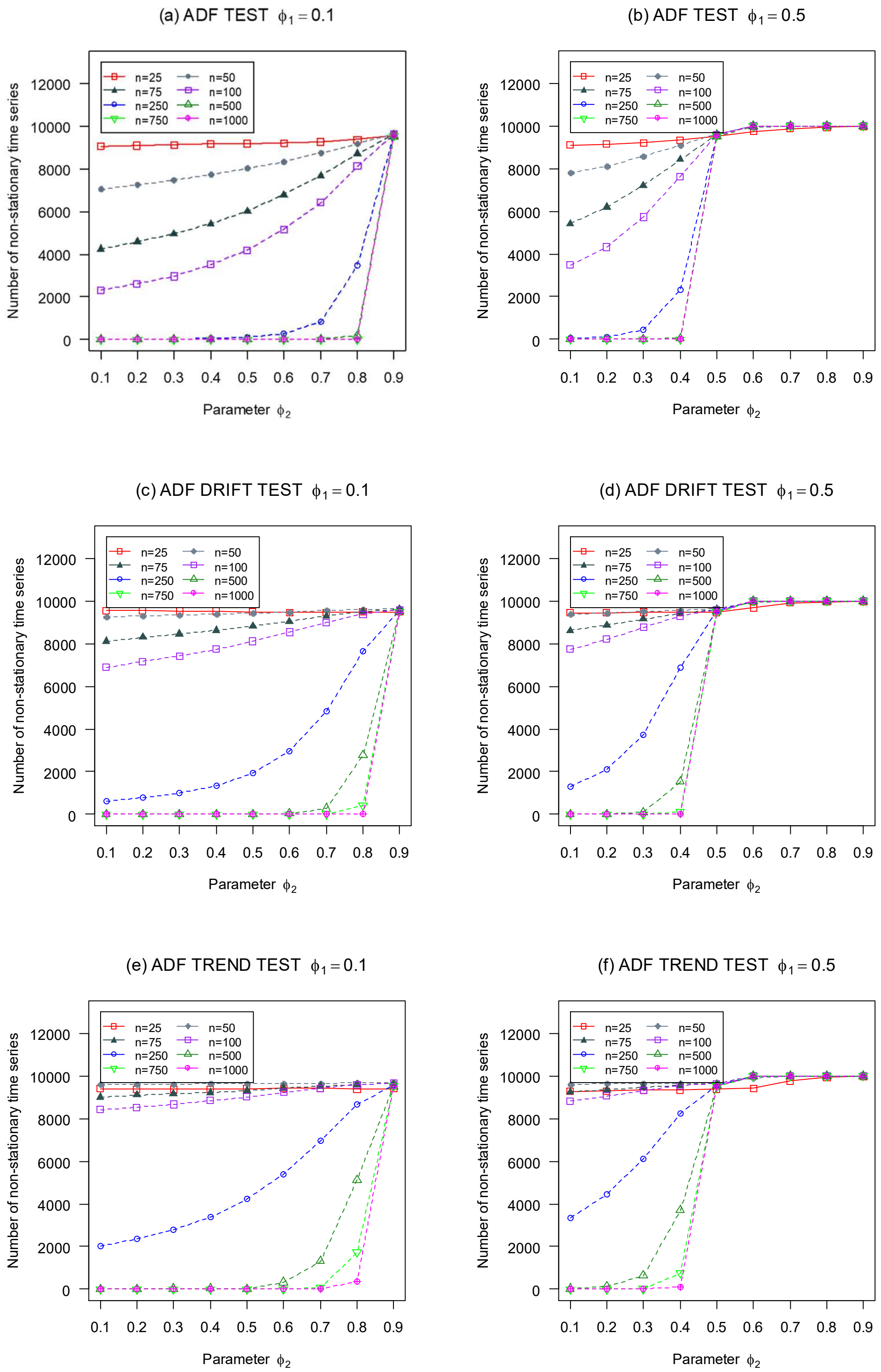

Figure 4.11 The number of series identified as non-stationary by the ADF test, ADF drift test and ADF trend test, data generated under AR(2) process.

4.5 Variability of DF tests and ADF tests

To discover the variability of the DF tests and the ADF tests a Monte-Carlo experiment was conducted with 1000 replications of the experiment carried out in Section 4.3.1. As this experiment involved 1000 replications of a simulation involving 10,000 replications this is only conducted ϕ_1 = 0.5. The aim was to calculate the mean and standard deviation for the DF test, the DF drift test, the DF trend test, the ADF test, the ADF drift test and the ADF trend test, when n = 100 observations under an AR(1) process. The DF tests and the ADF tests were calculated by using the steps of the previous Monte-Carlo experiment.

	n = 100					
ϕ_1 = 0.5	DF	DF drift	DF trend	ADF	ADF drift	ADF trend
Mean	0	0	0.15	2784.95	7256.425	8588.7
Standard deviation	0	0	0.36162	38.66155	40.26189	35.13943

Table 4.1 The mean and standard deviation for the test DF tests and ADF tests, data generated under an AR(1) process with ϕ_1 = 0.5 and n = 100.

As can be seen from Table 4.1 the results of this experiment show that for the parameter value ϕ_1 = 0.5 the DF and DF drift tests have the standard deviation equal to zero and the DF trend with standard deviation 0.36162. These results are as expected as the DF tests are the appropriate tests for this situation and almost all the time series are identified as stationary. In the case of the ADF test, the ADF drift test and the ADF trend test most time series are identified as non-stationary, which is to be expected as these are not the appropriate tests (this assumes prior knowledge that the process is an AR(1)).

The standard deviations for the ADF tests are around 1 percent of the associated mean.

4.6 Summary

The simulations undertaken show that the length of a time series critically affects the number of series identified as being non-stationary by standard tests (for example, the DF test and the ADF test).

A stationary time series was generated from an AR(1) process with positive values of the parameter and examined by the DF test, the DF drift test, and the DF trend test, see Figures 4.1, 4.3, and 4.5. It was found to be reliable to examine the series for stationarity by using the DF test, when n = 25 and the values of the parameter $\phi_1 < 0.5$, when n = 50 and for values of the parameter $\phi_1 < 0.8$, when n = 75 or 100 and for values of the parameter $\phi_1 < 0.9$, and when $n \geq 250$ with the values of the parameter $\phi_1 < 0.99$.

It is dependable to examine a time series for stationarity when data is examined by the DF test with drift, when n = 25 and for values of the parameter $\phi_1 < 0.2$, when $n = 50$ and for values of the parameter $\phi_1 < 0.6$, when n = 75 or 100 and for values of the parameter $\phi_1 < 0.8$, and when n > 250 and for values of the parameter $\phi_1 < 0.99$.

It is reliable to examine the series for stationarity by using the DF test with trend, when $n = 50$ and for values of the parameter $\phi_1 < 0.5$, when n = 75 and for values of the parameter $\phi_1 < 0.7$, when $n = 100$ and for values of the parameter $\phi_1 < 0.8$, and when n = 250 and for values of the parameter $\phi_1 < 0.9$, when n > 500 and for values of the parameter $\phi_1 < 0.99$.

A time series was generated from an AR(1) process with positive values of the parameter and examined by the ADF test, the ADF drift test, and the ADF trend test, see Figures 4.2, 4.4, and 4.6.

It is reliable to examine for stationarity by using the ADF test, when n = 250 and the values of parameter $\phi_1 < 0.9$, and when n > 250 with values of the parameter $\phi_1 < 0.99$, see Figures 4.2.

It was found to be reliable to examine for stationarity by using of the ADF test with drift, when $n = 250$ with values of the parameter $\phi_1 < 0.2$, see Figures 4.4.

It is dependable to examine for stationarity by the ADF test with trend, when $n = 500$ and values of the parameter $\phi_1 < 0.9$, and when $n > 500$ for values of the parameter $\phi_1 < 0.99$, see Figures 4.4.

Table 4.2 gives a summary of the findings from the Monte Carlo experiments conducted in Section 4.5, broken down by the DF, DF drift, DF trend, ADF, ADF drift and ADF trend tests. For each length of data examined in the experiments, the maximum value of ϕ_1 is given that ensures the series will be correctly identified as stationary (a cut-off of no more than 5% of the time series incorrectly identified is used, since the tests in the Monte Carlo experiments were conducted at the 0.05 significance level). The symbol (-) in Table 4.2 means all time series are non-stationary.

	Maximum positive values of ϕ_1							
Length n	25	50	75	100	250	500	750	1000
DF	0.4	0.7	0.8	0.8	0.9	0.9	0.9	0.9
DF Drift	0.2	0.5	0.6	0.7	0.9	0.9	0.9	0.9
DF Trend	-	0.4	0.6	0.7	0.9	0.9	0.9	0.9
ADF	-	-	-	-	0.8	0.9	0.9	0.9
ADF Drift	-	-	-	-	0.3	0.8	0.9	0.9
ADF Trend	-	-	-	-	-	0.8	0.9	0.9

Table 4.2 The maximum positive value of ϕ_1 that ensures that the series are correctly identified as stationary by the DF and ADF tests, the DF drift and ADF drift tests, and the DF trend and ADF trend tests, for positive values of ϕ_1 and different lengths of data.

For an AR(1) process generated using negative values of the parameter $\phi_1$and examined by the ADF tests, the ADF drift test, and the ADF trend test (see Figures 4.7, 4.8, and 4.9) the number of stationary series identified as non-stationary does not depend on the parameter ϕ_1, instead it is dependent only on the number of data points in the time series. As the number of data points increase the number of stationary time series identified as being non-stationary decreases. Table 4.3 gives the minimum number of data points required to ensure that the ADF, the ADF drift and the ADF trend tests correctly identify the time series as stationary for negative values of ϕ_1 (again using a 5% cut-off).

	Minimum number of data points
ADF	250
ADF Drift	250
ADF Trend	500

Table 4.3 The minimum number of data points in a time series generated from an AR(1) process (with negative ϕ_1) that ensures that the series is correctly identified as stationary by the ADF test, the ADF drift test, and the ADF trend test.

A time series was generated from an AR(2) process with a range of positives values of the parameters ϕ_1 and ϕ_2, and examined by the ADF test, the ADF drift test, and the ADF trend test, see Figure 4.10 and Appendix D.

Tables 4.4 and 4.5 give a summary of the findings from the Monte Carlo experiments conducted in Section 4.4, broken down by DF, DF drift, DF trend, ADF, ADF drift and ADF trend tests. For each length of data examined in the experiments, the maximum value of ϕ_2 is given that ensures the series will be correctly identified as stationary (a cut-off of no more than 5% of the time series incorrectly identified is used, since the tests in the Monte Carlo experiments were conducted at the 0.05 significance level). The symbol (-) in Tables 4.4 and 4.5 mean all time series are non-stationary.

$\phi_1 = 0.1$	Maximum positive values of ϕ_2							
Length n	25	50	75	100	250	500	750	1000
DF	0.6	0.8	0.8	0.8	0.8	0.8	0.8	0.8
DF Drift	0.2	0.8	0.8	0.8	0.8	0.8	0.8	0.8
DF Trend	-	0.9	0.9	0.9	0.9	0.9	0.9	0.9
ADF	-	-	-	-	0.6	0.8	0.8	0.8
ADF Drift	-	-	-	-	-	0.7	0.8	0.8
ADF Trend	-	-	-	-	-	0.6	0.7	0.8

Table 4.4 The maximum positive value of ϕ_2, when ϕ_1= 0.1 that ensures that the series are correctly identified as stationary by the DF tests, the DF drift tests, the DF trend tests, the ADF tests, the ADF drift tests, and the ADF trend tests, for positive values of ϕ_2 and different lengths of data.

$\phi_1 = 0.5$	Maximum positive values of ϕ_2							
Length n	25	50	75	100	250	500	750	1000
DF	-	0.2	0.3	0.4	0.4	0.4	0.4	0.4
DF Drift	-	-	0.2	0.3	0.4	0.4	0.4	0.4
DF Trend	-	-	0.2	0.3	0.4	0.4	0.4	0.4
ADF	-	-	-	-	0.3	0.4	0.4	0.4
ADF Drift	-	-	-	-	-	0.3	0.4	0.4
ADF Trend	-	-	-	-	-	0.2	0.3	0.4

Table 4.5 The maximum positive value of ϕ_2, when ϕ_1 = 0.5 that ensures that the series are correctly identified as stationary by the DF tests, the DF drift tests, the DF trend tests, ADF tests, the ADF drift tests, and the ADF trend tests, for positive values of ϕ_2 and different lengths of data

Table 4.4 shows that it is reliable to examine an AR(2) process for stationarity by using the DF test, the DF drift test, and the DF trend test, when n = 50 and $\phi_1 = 0.1$, with ϕ_2 less than or equal 0.8. It is dependable to examine an AR(2) process for stationarity by using the ADF test, the ADF drift test, and the ADF trend test, when n = 750 and $\phi_1 = 0.1$, with ϕ_2 less than or equal 0.8. It can be seen from Table 4.5 that it is reliable to examine an AR(2) process for stationarity by using the DF test, the DF drift test, and the DF trend test, when n = 50 and $\phi_1 = 0.5$, with ϕ_2 less than or equal 04. For the ADF test, the ADF drift test, and the ADF trend test, when n = 750 and $\phi_1 = 0.5$, then ϕ_2 must be less or equal 0.4 for stationarity to be correctly identified. From Appendix D it can be seen that for values of ϕ_1 larger than 0.5 all the ADF tests are unreliable even for large values of n (e.g., 1000) and for small values of ϕ_2 (i.e., less than or equal 0.4).

Chapter 5 - Portmanteau Tests

The aim of this chapter is to review previous studies in the area of portmanteau tests, which have been developed by several researchers. The chapter also presents a new portmanteau test, which is based on exponential weights of the residual partial autocorrelation function. Monte Carlo experiments are used to compare the performance of the new portmanteau test to existing tests.

5.1 Introduction

Diagnostic checking is the third stage of the Box and Jenkins methodology. The adequacy of a statistical model is examined, by considering the model's residuals, then the autocorrelation and partial autocorrelation functions are used as diagnostic tools to test the goodness of fit of the model. A portmanteau test is an important method of diagnostic checking, which is used to test the goodness of fit of an ARMA model of a time series, which has been studied by both Box and Pierce (1970) and Ljung and Box (1978).

A portmanteau test is calculated by summing the residuals of the autocorrelation or partial autocorrelation function of the fitted model. Then the value of the portmanteau test is compared with a critical value. If the value of the portmanteau test is less than the critical value, it means the model is appropriate for the data. Alternatively, if the value of the portmanteau test is bigger than the critical value, it means the model is inappropriate for the data.

Suppose that a time series $\{z_t\}$ is generated by a stationary and invertible ARMA(p, q) process

$$\phi(B)z_t = \theta(B)e_t$$

where $\{e_{\mathrm{t}}\}$ is a white noise process of mean zero and constant variance σ_e^2, and $\phi(B)$ and $\theta(B)$ are polynomials given by $\phi(B) = 1 - \phi_1 B - \cdots - \phi_2 B^p$ and $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$, has been fitted by maximum likelihood estimates $\left(\hat{\phi}, \hat{\theta}\right)$ obtained for the parameters, then it is possible to identify the residuals as

$$\hat{e}_t = \hat{\theta}^{-1}(B)\hat{\phi}(B)z_t, \tag{5.1}$$

The residuals are computed recursively using Equation 1.3 in the following form

$$\hat{e}_t = z_t - \sum_{j=1}^{p} \hat{\phi}_j \, z_{t-j} + \sum_{j=1}^{q} \hat{\theta}_j \, \hat{e}_{t-j} \qquad t = 1,2,\cdots,n \tag{5.2}$$

These residuals $\hat{e}_t$ from the ARMA model will be random if the model is correct, this means that the autocorrelation of the residuals ρ_k will be zero at all lags k. This gives the null hypothesis for all lags k

$$H_0: \rho_k = 0$$

$$H_1: \rho_k \neq 0$$

When considering the partial autocorrelation ϕ_{kk} of the residuals at lags k the hypothesis test can be given in the equivalent form

$$H_0: \phi_{kk} = 0$$

$$H_1: \phi_{kk} \neq 0$$

All the following tests will use one of these forms of the hypothesis, depending on whether the statistic relies on ρ_k or ϕ_{kk}.

5.1.1 Box and Pierce test

Box and Pierce (1970) showed that if the fitted model is appropriate then the portmanteau test statistic

$$\tilde{Q}_{BP} = n \sum_{k=1}^{m} \hat{\rho}_k^2 \, (\hat{e}) \tag{5.3}$$

is approximately distributed as $\chi^2(m - p - q)$, where $\hat{\rho}_k$ is the sample autocorrelation function, and n is the number of observations and m is the maximum lag taken into account.

5.1.2 Ljung and Box test

Ljung and Box (1978) showed that the chi-squared distribution does not provide a sufficiently accurate approximation to the distribution of the statistic $\tilde{Q}_{BP}$ under the null hypothesis, with $\tilde{Q}_{BP}$ tending to the smaller values than expected under the chi-squared distribution. Empirical evidence to support this was presented by Davies, Triggs and Newbold (1977). Consequently, Ljung and Box (1978) proposed a modified form of the portmanteau test statistic given by

$$\tilde{Q}_{LB} = n(n+2)\sum_{k=1}^{m}(n-k)^{-1}\,\hat{\rho}_k^2(\hat{e}) \tag{5.4}$$

where

$$\hat{\rho}_k(\hat{e}) = \sum_{t=k+1}^{n} e_t e_{t-k} \Big/ \sum_{t=1}^{n} e_t^2$$

The modified statistic has, approximately, a mean of $E(\tilde{Q}) \approx m-p-q$ of the $\chi^2(m-p-q)$ distribution, where n is the number of observations and m is the maximum lag taken into account.

5.1.3 Monti test

Monti (1994) showed that the statistic $\tilde{Q}_M$ in Equation 5.5 is asymptotically distributed as $\chi^2(m-p-q)$, analogous to the asymptotic distribution of the statistic $\tilde{Q}_{LB}$ in Equation 5.4. The portmanteau test $\tilde{Q}_{LB}$ is based on the autocorrelation functions. Monti (1994) suggested a portmanteau test statistic

$$\tilde{Q}_M = n(n+2)\sum_{k=1}^{m}(n-k)^{-1}\,\hat{\phi}_{kk}^2(\hat{e}) \tag{5.5}$$

where $\hat{\phi}_{kk}(\hat{e})$ is the residual partial autocorrelation at lag k, n is the number of observations and m is the maximum lag taken into account. Hence a test of model adequacy can be based on referring the value of $\tilde{Q}_M$ to the upper critical value of the $\chi^2(m-p-q)$ distribution. If the model is correct, $\hat{\phi}_{kk}^2(\hat{e})$ is approximately distributed as a normal random variable with mean zero and variance $(n-k)/\big(n(n+2)\big)$. The test based on $\tilde{Q}_M$ has been found to be typically at least as powerful as $\tilde{Q}_{LB}$ (Monti, 1994).

5.1.4 Peña and Rodríguez 2002 test

Peña and Rodríguez (2002) showed that the portmanteau goodness-of-fit test statistic is based on a general measure of multivariate dependence. Denote the correlation matrix up to order lag m of residual $\hat{e}_t$ from the fitted ARMA(p,q) model by

$$\hat{R}_m(\hat{e}) = \begin{pmatrix} 1 & \hat{\rho}_1(\hat{e}) & \hat{\rho}_2(\hat{e}) & \cdots & \hat{\rho}_k(\hat{e}) \\ \hat{\rho}_1(\hat{e}) & 1 & \hat{\rho}_2(\hat{e}) & \cdots & \hat{\rho}_{k-1}(\hat{e}) \\ \hat{\rho}_2(\hat{e}) & \hat{\rho}_1(\hat{e}) & 1 & \cdots & \hat{\rho}_{k-2}(\hat{e}) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \hat{\rho}_k(\hat{e}) & \hat{\rho}_{k-1}(\hat{e}) & \hat{\rho}_{k-2}(\hat{e}) & \cdots & 1 \end{pmatrix}$$

The proposed portmanteau test statistic is based on the determinant of this correlation matrix, a general measure of a dependence in multivariate analysis, and is given by

$$\hat{D}_m = n\left(1 - \left|\hat{R}_m(\hat{e})\right|^{1/m}\right) \tag{5.6}$$

where n is the length of the time series. If the model is correctly identified, $\hat{D}_m$ is asymptotically distributed as a linear combination of Chi-squared random variables and is approximately a Gamma distribution random variable for large values of m with parameter α and β.

The distribution of $\hat{D}_m$ can be approximated by the Gamma distribution (Peña and Rodríguez, 2002), where the parameters are defined by

$$\alpha = \frac{3m[(m+1) - 2(p+q)]^2}{2[2(m+1)(2m+1) - 12m(p+q)]}$$

and

$$\beta = \frac{3m[(m+1) - 2(p+q)]}{2(m+1)(2m+1) - 12m(p+q)}$$

and the distribution has a mean of

$$\frac{\alpha}{\beta} = \frac{m+1}{2} - (p+q)$$

and a variance of

$$\frac{\alpha}{\beta^2} = \frac{(m+1)(2m+1)}{3m} - 2(p+q)$$

5.1.5 Peña and Rodríguez 2006 test

Peña and Rodríguez (2006) gave a new portmanteau test statistic, which is the log of the determinant of the mth autocorrelation matrix

$$D_m^* = -\frac{n}{m+1}\log\left|\left\|\hat{R}_m\right\|\right| \quad (5.7)$$

where $\hat{R}_m$ is the residual correlation matrix of order m and $|.|$ is the absolute value of the constant. There are two approximations to the asymptotic distribution, which are based on the Gamma and Normal distribution. The Gamma distribution is the approximation distribution of D_m^*, where the parameters are defined by

$$\alpha = \frac{3(m+1)[m-2(p+q)]^2}{2[2m(2m+1)-12(m+1)(p+q)]} \quad (5.8)$$

and

$$\beta = \frac{3(m+1)[m-2(p+q)]}{2m(2m+1)-12(m+1)(p+q)} \quad (5.9)$$

and the distribution has a mean of

$$\frac{\alpha}{\beta} = \frac{m}{2} - (p+q)$$

and a variance of

$$\frac{\alpha}{\beta^2} = \frac{m(2m+1)}{3(m+1)} - 2(p+q)$$

Peña and Rodríguez (2006) suggested a power transformation which reduces the skewness in order to improve the normal approximation. The test statistic is

$$ND_m^* = \left(\frac{\alpha}{\beta}\right)^{-1/\lambda}\left(\frac{\lambda}{\sqrt{\alpha}}\right)\left((D_m^*)^{1/\lambda} - \left(\frac{\alpha}{\beta}\right)^{1/\lambda}\left(1 - \frac{1}{2\alpha}\left(\frac{\lambda - 1}{\lambda^2}\right)\right)\right)$$

and

$$\lambda = \left\{1 - \frac{2\big(m/2 - (p+q)\big)\Big(m^2/\big(4(m+1)\big) - (p+q)\Big)}{3\big(m(2m+1)/\big(6(m+1)\big) - (p+q)\big)^2}\right\}^{-1}$$

where m is moderately large $\lambda \simeq 4$, α and β are the values obtained in Equations (5.8) and (5.9). The statistic ND_m^* is distributed as N(0,1).

5.1.6 Mahdi and McLeod test

Mahdi and McLeod (2011) showed that for large n, the portmanteau test statistic

$$\tilde{Q}_{MM} = -\frac{3n}{2m+1}\log\left|\hat{R}_m\right| \tag{5.10}$$

is approximately distributed as a chi squared random variable with

$$\frac{1.5m(m+1)}{2m+1} - (p+q),$$

and $\chi^2(m-p-q)$ degrees of freedom. In their paper the simulation study compared performance of their test with the $\tilde{Q}_{LB}$ and $\tilde{Q}_M$ tests.

5.1.7 Fisher and Gallagher test

Fisher and Gallagher (2012) introduced a new portmanteau test statistic $\tilde{Q}_{FGLB}$ that is based on the square of the mth-order autocorrelation matrix

$$\tilde{Q}_{FGLB} = n(n+2)\sum_{k=1}^{m}\frac{(m-k+1)}{m}\frac{\hat{\rho}_k^2}{n-k} \tag{5.11}$$

where $\hat{\rho}_k^2$ is the autocorrelation at lag k.The statistic is a weighted sum of the squares of the sample autocorrelation coefficients, where the weights consist of a convolution of the Ljung-Box standardizing weights with the sequence

$$\left\{\frac{1}{m}, \frac{2}{m}, \dots, \frac{m-1}{m}, 1\right\}.$$

The $\tilde{Q}_{FGLB}$ is approximately distributed as a Gamma distribution with shape

$$\alpha = \frac{3m(m+1)}{8m+4}$$

and scale

$$\beta = \frac{2(2m+1)}{3m}$$

The simulation study indicates that $\tilde{Q}_{FGLB}$ is more powerful than $\tilde{Q}_{BP}$, $\tilde{Q}_{LB}$ and $\tilde{Q}_M$ (Fisher and Gallagher, 2012).

5.1.8 Gallagher and Fisher tests

Gallagher and Fisher (2015) introduced three portmanteau test statistics, created by taking general weighted sums of the first $m = m(n)$ squared sample autocorrelations:

$$Q_w = n \sum_{k=1}^{m} w_k \hat{\rho}_k^2 \tag{5.12}$$

where n is the number of observations, k is the number of lags taken into account and m is the maximum lag. In addition to the weight given in Equation 5.12 they consider three additional weighting schemes:

1 - Kernel-based weights: The weights are based on the square of a kernel function and blended with the Ljung-Box standardizing terms to construct a sequence of weights

$$w_k = \big((n+2)/(n-k)\big)[\mathcal{K}(k/m)]^2$$

where $\mathcal{K}(\cdot)$ is the Daniell Kernel function, which is defined as

$$\mathcal{K}(k/m) = \begin{Bmatrix} \frac{\sin\left(\sqrt{3}\pi(k/m)\right)}{\sqrt{3}\pi(k/m)} & : & |k/m| < 1 \\ 0 & : & |k/m| \geq 1 \end{Bmatrix}.$$

The theoretical asymptotic distributions of the weighting scheme is given by

$$\frac{Q_w - \sum w_k}{\sqrt{2w_k^2}} \Rightarrow N(0,1),$$

where ⇒ denotes convergence in the distribution. The asymptotic distribution of this test is the normal distribution (Gallagher and Fisher, 2015).

2 - Geometrically Decaying Weights: For ARMA models autocorrelations decay exponentially with respect to the lag. It seems intuitive then that the weights in Equation 5.12 be selected to decay quickly as well, since even under the alternative hypothesis of an underfitted model, the correlations at large lags should still be relatively small. They consider weights of the sum of the form

$$w_k = (p+q)a^{k-1},$$

for some user-specified ratio $0 < a < 1$. This weighting will be

$$\sum w_k = \frac{(p+q)(1-a^m)}{(1-a)}$$

and

$$\sum w_k^2 = \frac{(p+q)^2(1-a^{2m})}{1-a^2}$$

The simulation studies used a value of $a = 0.9$ (Gallagher and Fisher, 2015).

3 - Data Adaptive Weights: In this portmanteau test they used the sample autocorrelation $\hat{\rho}$ and the sample partial autocorrelation $\hat{\phi}_{kk}$. The test is defined as

$$\tilde{Q}_{GFD} = n\sum_{k=1}^{m_0} \frac{n+2}{n-k}\hat{\rho}^2 + n\sum_{k=m_0+1}^{m} w_k\,\hat{\rho}^2,$$

The first m_0 terms use the standardizing weight $(n+2)/(n-k)$ from the Ljung-Box statistics, and the remaining terms use the weights

$$w_k = -\log\left(1 - |\hat{\phi}_{kk}|\right)$$

where $m_0 = \min\,(\log(n), M)$, and M is a finite bound. The simulation studies indicate that using data adaptive weights is more powerful that all the previous portmanteau tests (Gallagher and Fisher, 2015).

4 - Asymptotic distribution of Gallagher and Fisher (2015)

Gallagher and Fisher (2015) considered the asymptotic behaviour of general weighted portmanteau statistics satisfying Equation 5.12. The gamma approximation is used for geometrically decaying weights and data adaptive weights similar to Peña and Rodríguez (2002-2006) and Fisher and Gallagher (2012), which is based on the work of Satterthwaite (1941-1946) and Box (1954). That is $Q_w \sim \Gamma(\alpha, \beta)$ with shape and scale

$$\alpha = \frac{(\sum w_k)^2}{2(\sum w_k^2 - p - q)} \tag{5.13}$$

and

$$\beta = \frac{2(\sum w_k^2 - p - q)}{\sum w_k} \tag{5.14}$$

respectively (Gallagher and Fisher, 2015).

5.2 A new weighted portmanteau test

5.2.1 Exponential weighted portmanteau test

This thesis introduces two new portmanteau tests that are based on exponential weights. The first new test is a development of Ljung and Box's test (1978) and the second test is a development of Monti's test (1994). These new portmanteau test statistics are defined as

$$\tilde{Q}_{EXLB} = n(n+2)\sum_{k=1}^{m} w_k \frac{\hat{\rho}_k^2}{n-k}. \tag{5.15}$$

$$\tilde{Q}_{EXM} = n(n+2)\sum_{k=1}^{m} w_k \frac{\hat{\phi}_{kk}^2}{n-k}, \tag{5.16}$$

where $\hat{\rho}_k^2$ is the sample autocorrelation and $\hat{\phi}_{kk}^2$ is the sample partial autocorrelation at lag m, and w_k is an exponential weight.

5.2.2 Derivation of the exponential weight w_k

Consider an exponential function of the form

$$f(x) = a^x, \quad 0 < a < 1, \quad 0 \leq x < 1$$

where a is the base and x is the exponent.

Constrain x to the values $(k-1)/m$, that is, terms from $\{0, 1/m, 2/m, \dots, (m-1)/\mathrm{m}\}$, where k is the length of lag used in the autocorrelation function and the partial autocorrelation function, and m is the maximum lag.

Also, constrain a to take the value $\frac{1}{m}$.

The exponential function now takes the form

$$f\left(\frac{k-1}{m}\right) = \left(\frac{1}{m}\right)^{\left(\frac{k-1}{m}\right)}$$

Then, this can be rearranged as,

$$= e^{\ln\left(\frac{1}{m}\right)^{\left(\frac{k-1}{m}\right)}}$$

$$= e^{\left(\frac{k-1}{m}\right)\ln\left(\frac{1}{m}\right)}$$

$$= e^{-\left(\frac{k-1}{m}\right)\ln}$$

Since, $f\left(\frac{k-1}{m}\right)$ is now only a function of the variable k, and m is a constant it can be rewritten as a function of the lag k, $w(k)$, which can be written as the exponential weight w_k.

So that

$$w_k = e^{-\left(\frac{k-1}{m}\right)\ln m} \quad k = 1,2,\dots,m \tag{5.17}$$

This exponential weight has similar distribution behaviour to the weights employed by Fisher and Gallagher (2012), and Gallagher and Fisher (2015).

5.2.3 Asymptotic distribution of the new univariate portmanteau test

Theorem 5.1 Suppose that a univariate time series $\{z_t\}$ is generated by a stationary and invertible ARMA(p, q) process with mean zero and constant variance (see Equation 5.1). Then, the new univariate portmanteau test statistic is asymptotically distributed as

$$\sum_{k=1}^{m} \lambda_k \chi_{1,k}^2$$

where m is the maximum lag, $\chi_{1,k}^2$ $(k = 1, 2, \dots, m)$ are independent χ_1^2 random variables and λ_k $(\lambda_1, \lambda_2, \dots, \lambda_m)$ are the eigenvalues of $(\boldsymbol{I}_m - \boldsymbol{Q}^*)\boldsymbol{G}$, where $\boldsymbol{G}$ is a $m \times m$ diagonal matrix

$$\boldsymbol{G} = \begin{pmatrix} w_1 I & 0 & \cdots & 0 \\ 0 & w_2 I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_m I \end{pmatrix}$$

where w_k $(k = 1, 2, \cdots, m)$ are weights that satisfy $0 < w_k \leq 1$, and $\boldsymbol{Q}^*$ is an idempotent matrix of univariate, which is define as $\boldsymbol{Q}^* = \boldsymbol{Y}\boldsymbol{V}^{-1}\boldsymbol{Y}^\dagger$, $\boldsymbol{V}$ is the information matrix for the parameters ϕ and θ and $\boldsymbol{Y}$ is a $m \times (p + q)$ matrix with elements $\phi^\dagger$ and $\theta^\dagger$define d by

$$\frac{\mathbf{1}}{\phi(B)} = \sum_{i=0}^{\infty} \phi_i^\dagger B^i$$

and

$$\frac{\mathbf{1}}{\theta(B)} = \sum_{i=0}^{\infty} \theta_i^\dagger B^i$$

The form of the idempotent matrix $\boldsymbol{Q}^*$ was first derived by Box and Pierce (1970) in the development of their univariate portmanteau tests.

Proof of Theorem 5.1:

Let, $\chi_{1,k}^2$ $(k = 1, 2, \dots, m)$ be independent χ_1^2 random variables. By using the idempotent matrix form Box and Pierce (1970) and multiplying the exponential weight with $(\boldsymbol{I}_m - \boldsymbol{Q}^*)$, then

$$(\boldsymbol{I}_m - \boldsymbol{Q}^*)\boldsymbol{G}$$

the λ_k $(\lambda_1, \lambda_2, \dots, \lambda_m)$ are the eigenvalues of $(\boldsymbol{I}_m - \boldsymbol{Q}^*)\boldsymbol{G}$.

Summing the eigenvalues and applying the tr matrix to the idempotent matrix, it gives

$$\sum_{k=1}^{m} \lambda_k = tr\big((\boldsymbol{I}_m - \boldsymbol{Q}^*)\boldsymbol{G}\big)$$
$$= tr(\boldsymbol{G}) - tr(\boldsymbol{Q}^*) + (1/m)tr(\boldsymbol{Q}^*\boldsymbol{F})$$

where $\boldsymbol{F}$ is a diagonal matrix with elements $f_k = k$, where $k = 0,1, \dots, (m-1)$, and

$$\sum_{k=1}^{m} \lambda_k^2 = tr\big((\boldsymbol{I}_m - \boldsymbol{Q}^*)\boldsymbol{G}(\boldsymbol{I}_m - \boldsymbol{Q}^*)\boldsymbol{G}\big)$$
$$= tr(\boldsymbol{G})^2 - tr(\boldsymbol{Q}^*) + (2/m)tr(\boldsymbol{Q}^*\boldsymbol{F}) - (2/(m^2)^2)tr(\boldsymbol{Q}^*\boldsymbol{F^2})$$
$$+ (1/(m^2))tr(\boldsymbol{Q}^*\boldsymbol{F}\boldsymbol{Q}^*\boldsymbol{F})$$

As $\boldsymbol{Q}^*$ is the idempotent matrix with rank $(p+q)$, then

$$\sum_{k=1}^{m} \lambda_k = \sum_{k=1}^{m} w_k - (p+q) + (1/m)\sum_{k=2}^{m}(k-1)\, q_{kk}$$

$$\sum_{k=1}^{m} \lambda_k^2 = \sum_{k=1}^{m} w_k^2 - (p+q) + (2/m)\sum_{k=2}^{m}(k-1)\, q_{kk} - (2/(m^2)^2)\sum_{k=2}^{m}(k-1)^2\, q_{kk}$$
$$+ (1/m)^2 \sum_{i=2}^{m}\sum_{j=2}^{m}(i-1)(j-1)q_{ij}^2$$

where q_{ij} are the elements of $\boldsymbol{Q}^*$.

Using Kronecker's lemma (Davidson, 1994), as $m \to \infty$, then

$$(1/m)\sum_{k=2}^{m}(k-1)\, q_{kk} \to 0$$

$$(2/m)\sum_{k=2}^{m}(k-1)\, q_{kk} \to 0$$

$$(2/(m^2)^2)\sum_{k=2}^{m}(i-1)^2\,q_{kk}\ \rightarrow 0$$

and

$$(1/m)^2\sum_{i=2}^{m}\sum_{j=2}^{m}(i-1)(j-1)q_{ij}^2 \rightarrow 0$$

Thus,

$$\sum_{k=1}^{m}\lambda_k = \sum_{i=1}^{m} w_i \ \text{ as } m \rightarrow \infty \tag{5.18}$$

$$\sum_{k=1}^{m}\lambda_k^2 = \sum_{i=1}^{m} w_i^2 - (p+q) \ \text{ as } m \rightarrow \infty \tag{5.19}$$

From the result of Box (1954), as a vector $\boldsymbol{z}_t$ having mean zero and constant covariance matrix, then the asymptotic distribution of portmanteau test statistic in Equations 5.18 and 5.19 distributed as

$$\sum_{k=1}^{m}\lambda_k \chi_{1,k}^2$$

5.2.4 Approximation distribution of the new univariate portmanteau test

The two new portmanteau tests are based on exponential weights of the autocorrelation function or the partial autocorrelation function. By using the results of the Hong (1996 a, b), the Hong test statistic can be defined as

$$Q_n^* = \frac{n\sum_{j=n-1}^{n-1}\kappa^2\left(\frac{j}{M}\right)\hat{\rho}(j) - MS(\kappa)}{(2MD(\kappa))^{1/2}}$$

M is a sequence of truncation values. If the smoothing parameter $M = n(M) \rightarrow \infty$ and $M/n \rightarrow 0$, then

$$S(\kappa) = \int_{-\infty}^{\infty}\kappa^2\,dz,\ D(\kappa) = \int_{-\infty}^{\infty}\kappa^4\,dz$$

where $\hat{\rho}(j)$ is the correlation function, n is number of observations and κ is the Kernel function. From Equation 5.18 and 5.19 the normalize terms are $\sum_{k=1}^{m} w_k$ and $\sum_{k=1}^{m} w_k^2$, then these two normalize terms could be replaced by $S(\kappa)$ and $D(\kappa)$. The approximation distribution of the new portmanteau test statistics of ARMA models can be written as $\tilde{Q}_{EXLB} \sim \Gamma(\alpha, \beta)$ and $\tilde{Q}_{EXM} \sim \Gamma(\alpha, \beta)$ with shape

$$\alpha = \frac{(\sum w_k)^2}{2(\sum w_k^2 - p - q)}$$

and scale,

$$\beta = \frac{2(\sum w_k^2 - p - q)}{\sum w_k}$$

5.3 Monte Carlo experiment

5.3.1 Simulation studies

The aim of the simulation study is to compare the new exponential portmanteau test with the portmanteau tests used in previous studies. The new test is compared with the other tests, which were developed by Ljung and Box (1978) $\tilde{Q}_{LB}$, Monti (1994) $\tilde{Q}_{M}$, Mahdi and McLeod (2011) $\tilde{Q}_{MM}$, Fisher and Gallagher (2012) $\tilde{Q}_{FGLB}$, Gallagher and Fisher (2015) Kernel-based weights $\tilde{Q}_{GFK}$ and Data Adaptive Weights $\tilde{Q}_{GFD}$. The empirical size and the power level of the tests were investigated by conducting simulations studies using the R program.

5.3.2 Empirical size

A Monte-Carlo experiment was conducted with 10,000 replications. The aim was to simulate n = 100 observations under an AR(1) process $z_t - \phi\, z_{t-1} = e_t$ with different parameters $\phi = 0.1, 0.3, 0.5, 0.7$ and 0.9. Next, an AR(1) model was fitted to the generated data producing an estimate $\hat{\phi}$ of the underlying parameter ϕ. The method employed to achieve the fitted model uses the maximum likelihood function, using approximation 2 from Box, Jenkins and Reinsel, (2008, p. 321).

$$(n-2)(n-1)^{-1}\sum_{t=2}^{n} z_t z_{t-1} \Bigg/ \sum_{t=2}^{n-1} z_t^2 \qquad (5.20)$$

Next, the autocorrelations of the fitted model were calculated using the residuals $\hat{e}_t = z_t - \hat{\phi}\, z_{t-1}$ $(t = 2, \dots, n)$. The test statistics $\tilde{Q}_{LB}$, $\tilde{Q}_{M}$, $\tilde{Q}_{MM}$, $\tilde{Q}_{FGLB}$, $\tilde{Q}_{GFK}$, $\tilde{Q}_{GFD}$, $\tilde{Q}_{EXLB}$ and $\tilde{Q}_{EXM}$ were calculated. This was repeated for lags of autocorrelations and partial autocorrelations for maximum lags m = 10, 20 and 30.

Method of a Monte-Carlo experiment to calculate the empirical size of a range of portmanteau tests.

Below are the steps of a Monte-Carlo experiment where data are generated by an AR(1) process, $z_t = \phi\, z_{t-1} + e_t$, then fitted under an AR(1) model to find the empirical size of the following portmanteau tests $\tilde{Q}_{LB}$, $\tilde{Q}_{M}$, $\tilde{Q}_{MM}$, $\tilde{Q}_{FGLB}$, $\tilde{Q}_{GFK}$, $\tilde{Q}_{GFD}$, $\tilde{Q}_{EXLB}$ and $\tilde{Q}_{EXM}$.

1. Select the value of the process parameter ϕ and maximum lag m. In this example, $\phi = 0.1$ and m = 10.
2. Generate n = 100 values from a Normal distribution (e_t white noise).
3. Use the e_t values to generate observations z_t from an AR(1) process with parameter ϕ.
4. Fit an AR(1) model to the observations by estimating its parameters using the maximum likelihood function.
5. Find the residuals $\hat{e}_t$.
6. Find the residual autocorrelation and partial autocorrelation functions for the model.
7. Calculate the various portmanteau test statistics. For example, $\phi = 0.1$, gives

ϕ	$\tilde{Q}_{LB}$	$\tilde{Q}_{M}$	$\tilde{Q}_{MM}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
0.1	16.6580	17.0012	11.890	8.120	0.917	8.976	6.744	7.058

8. Look up the 5 percentage point of the χ^2_{m-1} distribution and the gamma distribution.

Distributions	χ^2_{m-1}			Gamma				
Tests	$\tilde{Q}_{LB}$	$\tilde{Q}_{M}$	$\tilde{Q}_{MM}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
$m = 10$	16.9		14.1	9.92	1.6	11.55	7.764	

9. Reject the fitted AR(1) model if the value of portmanteau test is bigger than the critical value in step 7 (using the appropriate distribution for each portmanteau test).

ϕ	$\tilde{Q}_{LB}$	$\tilde{Q}_{M}$	$\tilde{Q}_{MM}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
0.1	Accept	Reject	Accept	Accept	Accept	Accept	Accept	Accept

10. Repeat 10,000 times for steps 1-8.

11. For each portmanteau test use the number of rejected AR(1) models (out of 10,000) to find the percentage rejected.

Tables 5.1, 5.2 and 5.3 give the results of the Monte-Carlo experiment and show the proportion of the test statistics $\tilde{Q}_{LB}$, $\tilde{Q}_{M}$, $\tilde{Q}_{MM}$, $\tilde{Q}_{FGLB}$, $\tilde{Q}_{GFK}$, $\tilde{Q}_{GFD}$, $\tilde{Q}_{EXLB}$ and $\tilde{Q}_{EXM}$ that are above the upper 5 percentage point of the χ^2_{m-1} distribution or gamma distribution. The tables show data fitted under the AR(1) with different parameters $\phi = 0.1, 0.3, 0.5, 0.7$ and 0.9 with n = 100, and lags of autocorrelations and partial autocorrelations are m = 10, 20 and 30.

ϕ	$\tilde{Q}_{LB}$	$\tilde{Q}_{M}$	$\tilde{Q}_{MM}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
0.1	**0.0531**	0.0548	0.0294	0.0311	0.0122	0.0301	0.0292	0.0287
0.3	**0.0516**	0.0541	0.0347	0.0352	0.0165	0.0336	0.0335	0.0355
0.5	0.0574	**0.0531**	0.0312	0.0369	0.0167	0.0357	0.0369	0.0378
0.7	0.0518	**0.0514**	0.0308	0.0343	0.0185	0.0317	0.0347	0.0349
0.9	0.0605	0.0568	0.0429	0.0474	0.0364	0.0475	0.0535	**0.0504**

Table 5.1 Empirical size of the test statistics $\tilde{Q}_{\mathrm{LB}}$, $\tilde{Q}_{\mathrm{M}}$, $\tilde{Q}_{\mathrm{MM}}$, $\tilde{Q}_{\mathrm{FGLB}}$, $\tilde{Q}_{\mathrm{GFK}}$, $\tilde{Q}_{\mathrm{GFD}}$, $\tilde{Q}_{\mathrm{EXLB}}$ and $\tilde{Q}_{\mathrm{EXM}}$ at 5% significance level for fitted AR(1) models, n = 100 and m = 10.

Table 5.1 shows the values of significant level when $\alpha = 0.05$, n = 100 and m = 10. The value of the $\tilde{Q}_{LB}$ test is closer to the 0.05 significance level in two cases, i.e., when $\phi =$ 0.1 and 0.3. The value of the $\tilde{Q}_M$ test is closer to 0.05 in two cases, i.e., when $\phi = 0.5$ and 0.7. The value of the $\tilde{Q}_{EXM}$ test is closer to 0.05 in one case, i.e., when $\phi = 0.9$. Overall, the $\tilde{Q}_M$ test is better than the other tests in most cases.

ϕ	$\tilde{Q}_{LB}$	$\tilde{Q}_M$	$\tilde{Q}_{MM}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
0.1	0.0616	**0.0557**	0.025	0.0412	0.0223	0.0385	0.0353	0.0316
0.3	0.0616	**0.0545**	0.0255	0.0398	0.0232	0.0397	0.0344	0.0336
0.5	0.0622	**0.0513**	0.0249	0.0438	0.0255	0.044	0.0404	0.0337
0.7	0.0671	0.0555	0.0281	**0.0504**	0.0284	0.0465	0.0467	0.0404
0.9	0.0706	**0.0500**	0.0284	0.0514	0.0318	0.0456	0.0491	0.0438

Table 5.2 Empirical size of the test statistics $\tilde{Q}_{\mathrm{LB}}$, $\tilde{Q}_{\mathrm{M}}$, $\tilde{Q}_{\mathrm{MM}}$, $\tilde{Q}_{\mathrm{FGLB}}$, $\tilde{Q}_{\mathrm{GFK}}$, $\tilde{Q}_{\mathrm{GFD}}$, $\tilde{Q}_{\mathrm{EXLB}}$ and $\tilde{Q}_{\mathrm{EXM}}$ at 5% significance level for fitted AR(1) models, $n = 100$ and $m = 20$.

Table 5.2 shows the values of significance level when $\alpha = 0.05$, n = 100 and m = 20. The value of the $\tilde{Q}_M$ test is closest to 0.05 in four cases, i.e., when $\phi = 0.1, 0.3, 0.5$ and 0.9. The values of the $\tilde{Q}_{FGLB}$ test is closest to 0.05 in one case, i.e., when $\phi = 0.7$. Overall, the $\tilde{Q}_M$ test is better than the other tests in most cases.

ϕ	$\tilde{Q}_{LB}$	$\tilde{Q}_M$	$\tilde{Q}_{MM}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
0.1	0.0731	0.0450	0.0163	**0.0507**	0.0264	0.0451	0.0392	0.0348
0.3	0.0721	0.0475	0.0202	**0.0518**	0.0262	0.0440	0.0424	0.0380
0.5	0.0684	0.0419	0.0158	**0.0504**	0.0263	0.0458	0.0402	0.0327
0.7	0.0768	0.0448	0.0165	0.0573	0.0300	0.0474	**0.0481**	0.0377
0.9	0.0749	0.0405	0.0204	0.0569	0.0325	0.0557	**0.0544**	0.0444

Table 5.3 Empirical size of the test statistics $\tilde{Q}_{\mathrm{LB}}$, $\tilde{Q}_{\mathrm{M}}$, $\tilde{Q}_{\mathrm{MM}}$, $\tilde{Q}_{\mathrm{FGLB}}$, $\tilde{Q}_{\mathrm{GFK}}$, $\tilde{Q}_{\mathrm{GFD}}$, $\tilde{Q}_{\mathrm{EXLB}}$ and $\tilde{Q}_{\mathrm{EXM}}$ at 5% significance level for fitted AR(1) models, $n = 100$ and $m = 30$.

Table 5.3 shows the values of significance level when $\alpha = 0.05$, $n = 100$ and $m = 30$. The value of the $\tilde{Q}_{FGLB}$ test is closest to 0.05 in three cases, i.e., when $\phi = 0.1, 0.3$ and 0.5. The values of the $\tilde{Q}_{EXLB}$ test is closest to 0.05 in two cases, i.e., when $\phi = 0.7$ and 0.9. No test stands out as the best over the range of ϕ values.

As can be seen from the previous tables of the empirical size, the $\tilde{Q}_{\mathrm{M}}$ test is the best test in most cases, when $m = 20$. However, this is not the situation when $m = 10$ or $m = 30$.

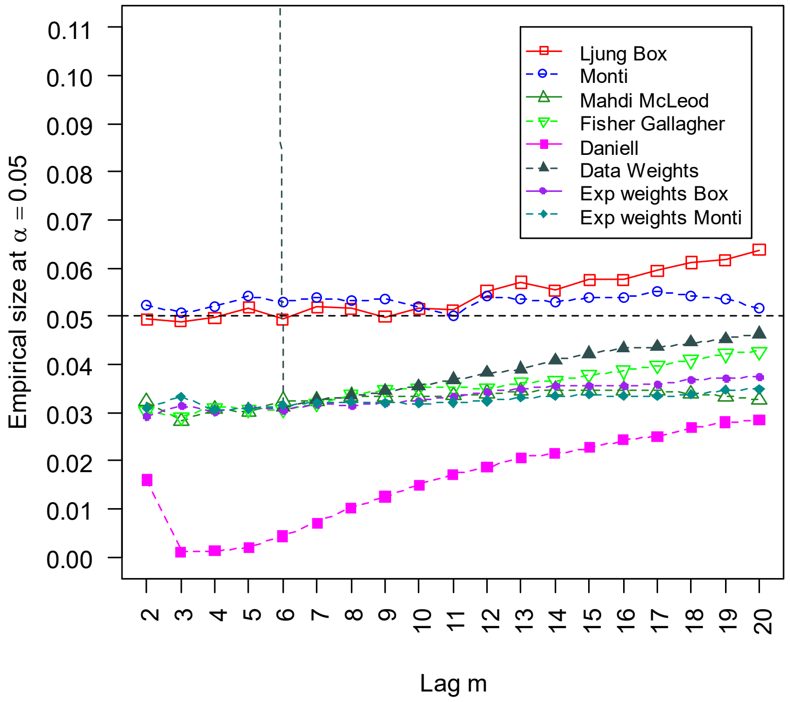

Figure 5.1 Empirical size for lags from 2 to 20 for a correctly fitted AR(1) model, data generated by an AR(1) process with $\phi_1 = 0.5$, at 5% significance level, series of length $n = 150$.

Figure 5.1 shows the empirical size of lags from 2 to 20 of the $\tilde{Q}_{LB}$, $\tilde{Q}_{M}$, $\tilde{Q}_{MM}$, $\tilde{Q}_{FGLB}$, $\tilde{Q}_{GFK}$, $\tilde{Q}_{GFD}$, $\tilde{Q}_{EXLB}$ and $\tilde{Q}_{EXM}$ tests based on a 5% significance level, when data are generated by an AR(1) process $\phi = 0.5$ and fitted under an AR(1) model with $n = 150$ and 10,000 replications. The $\tilde{Q}_{LB}$, $\tilde{Q}_{FGLB}$, and $\tilde{Q}_{GFD}$ tests increase as the lag increases. In addition, the

$\tilde{Q}_{MM}$, $\tilde{Q}_{EXLB}$ and $\tilde{Q}_{EXM}$ tests slowly increase as the lag increases. The $\tilde{Q}_{GFK}$ test rapidly decreases at lag 3, then increases as the lag increases. The $\tilde{Q}_{M}$ test is not affected as the lag increases. The $\tilde{Q}_{GFD}$ test always rejects the correct models at lags 2, 3, 4 and 5 (these points are off the scale in Figure 5.1), then from lag 6 the $\tilde{Q}_{GFD}$ increases as the lag increases.

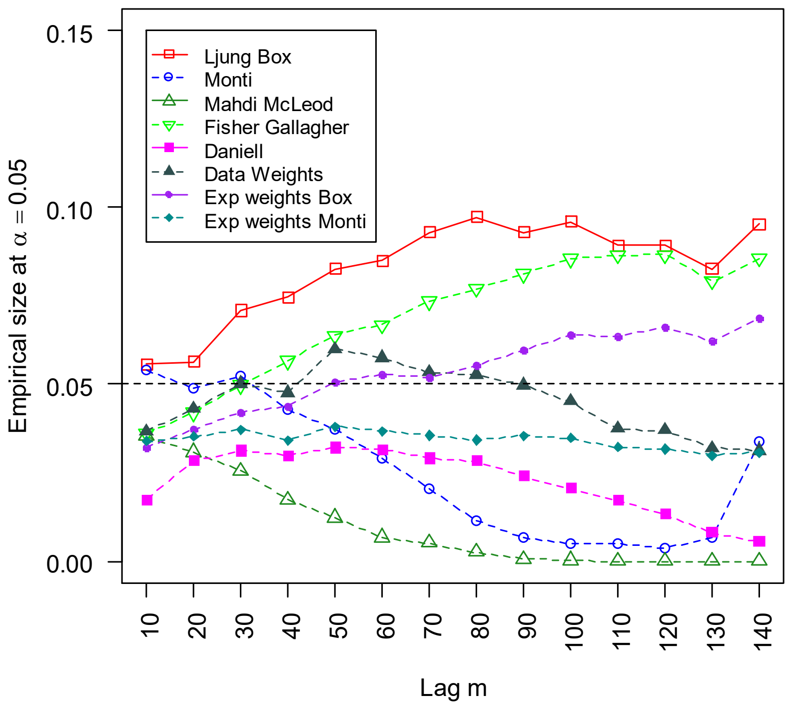

Figure 5.2 Empirical size for maximum lags for a properly fitted AR(1) model, data generated by an AR(1) process with $\phi_1 = 0.5$, at 5% significance level, series of length $n = 150$.

Figure 5.2 shows the empirical size of large lags based on a 5% significance level when data are generated by an AR(1) process with $\phi = 0.5$ and fitted by an AR(1) model with n = 100 and 10,000 replications. The $\tilde{Q}_{LB}$, $\tilde{Q}_{FGLB}$ and $\tilde{Q}_{EXLB}$ tests increase as the lag increases. Other tests such as, $\tilde{Q}_{M}$, $\tilde{Q}_{MM}$ and $\tilde{Q}_{GFK}$ decrease when the lag increases and the $\tilde{Q}_{GFD}$ test initially increases with increasing lag but then decreases for larger lags. The $\tilde{Q}_{EXM}$ test remains approximately constant as the lag increases.

5.3.3 Power studies

The aim of the power studies is to show which tests are the most powerful. The same processes and parameters were employed as those used in Monti (1994) to compare the portmanteau tests for the statistics $\tilde{Q}_{LB}$, $\tilde{Q}_{M}$, $\tilde{Q}_{MM}$, $\tilde{Q}_{FGLB}$, $\tilde{Q}_{GFK}$, $\tilde{Q}_{GFD}$, $\tilde{Q}_{EXLB}$ and $\tilde{Q}_{EXM}$.

The approach taken was to generate data by a number of alternative ARMA(2,2) processes,

$$z_t = e_t + \phi_1 z_{t-1} + \phi_2 z_{t-2} - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

and to fit an AR(1) model and a MA(1) model to the data. Next the residual of the data was obtained, and the ACF and PACF were calculated. For each alternative set of parameters for the ARMA(2,2) process, 10,000 replications of 100 observations were generated. For each test the power was computed with lags m = 10, 20 and 30.

In these experiments the AR(1) model parameter was estimated by using Equation 5.18. The MA(1) model parameter was estimated by using the maximum likelihood function, which is

$$\ln(\theta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \sum_{t=1}^{n}\frac{e_t^2}{2\sigma^2}. \tag{5.21}$$

Method of a Monte-Carlo experiment to calculate the power level of a range of portmanteau tests.

Below are the steps of a Monte-Carlo experiment (with an integrated example) where data are generated by an ARMA(2, 2) process $z_t = e_t + \phi_1 z_{t-1} + \phi_2 z_{t-2} - \theta_1 e_{t-1} - \theta_2 e_{t-2}$ then fitted under an AR(1) model to find the power level of the following portmanteau tests $\tilde{Q}_{LB}$, $\tilde{Q}_{M}$, $\tilde{Q}_{MM}$, $\tilde{Q}_{FGLB}$, $\tilde{Q}_{GFK}$, $\tilde{Q}_{GFD}$, $\tilde{Q}_{EXLB}$ and $\tilde{Q}_{EXM}$.

1. Select the values of the process parameters ϕ_1, $\phi_2, \theta_1, \theta_2$ and maximum lag m. In this example, $\theta_1 = -0.5, \theta_2 = 0,\ \phi_1 = 0,\ \phi_2 = 0$ and m = 10.
2. Generate 100 values from a Normal distribution (e_t white noise).
3. Use the e_t values to generate observations from an ARMA(2,2) process with parameters ϕ and θ.

4. Fit an AR(1) model to the observations by estimating its parameters using the maximum likelihood function.
5. Find the residual $\hat{e}_t$.
6. Find the residual autocorrelation and partial autocorrelation functions for the model.
7. Calculate the portmanteau test(s). For example, with an ARMA(2, 2) process with parameter $\theta_1 = -0.5, \theta_2 = 0$, $\phi_1 = 0$, $\phi_2 = 0$, one randomly generated series gave

θ_1	$\tilde{Q}_{LB}$	$\tilde{Q}_M$	$\tilde{Q}_{MM}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
-0.5	18.3804	13.8646	12.220	10.310	0.489	10.550	7.705	6.782

8. Look up the 5 percentage point of the χ^2_{m-1} distribution and gamma distribution.

Distributions	χ^2_{m-1}			Gamma				
Tests	$\tilde{Q}_{LB}$	$\tilde{Q}_M$	$\tilde{Q}_{MM}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
$m = 10$	16.9		14.1	9.92	1.6	11.37	7.764	

9. Reject the fitted AR(1) model if the value of portmanteau test is bigger than the critical value in step 7 (using the appropriate distribution for each portmanteau test). That is, with the example results from step 7

θ_1	$\tilde{Q}_{LB}$	$\tilde{Q}_M$	$\tilde{Q}_{MM}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
-0.5	Reject	Accept	Accept	Reject	Accept	Accept	Accept	Accept

10. Repeat 10,000 times for steps 1-8.
11. For each portmanteau test use the number of rejected AR(1) models (out of 10,000) to find the proportion rejected.

These steps can be used for data generated by ARMA(2,2) models and fitted under MA(1) model.

					$m = 10$							
Model	ϕ_1	ϕ_2	θ_1	θ_2	$\tilde{Q}_{LB}$	$\tilde{Q}_{M}$	$\tilde{Q}_{MM}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
1	**---**	**---**	**-0.50**	**---**	0.2634	0.3081	0.3886	0.3314	0.3395	0.3266	0.3702	**0.4326**
2	**---**	**---**	**-0.80**	**---**	0.7444	0.9659	0.9872	0.8995	0.9392	0.9034	0.9336	**0.9901**
3	**---**	**---**	**-0.60**	**0.30**	0.7792	0.9880	0.9935	0.9187	0.9300	0.9275	0.9431	**0.9953**
4	**0.10**	**0.30**	**---**	**---**	0.4283	0.4269	0.5239	0.5290	0.5424	0.5295	0.5612	**0.5619**
5	**1.30**	**-0.35**	**---**	**---**	0.7211	0.7088	0.8467	0.8454	**0.9089**	0.8238	0.8929	0.8972
6	**0.70**	**---**	**-0.40**	**---**	0.5541	0.6179	0.7605	0.6958	0.7821	0.6519	0.7713	**0.8263**
7	**0.70**	**---**	**-0.90**	**---**	0.9872	**1**	**1**	0.9997	**1**	0.9996	**1**	**1**
8	**0.40**	**---**	**-0.60**	**0.30**	0.8414	0.9975	**0.9992**	0.9649	0.9813	0.9669	0.983	**0.9992**
9	**0.70**	**---**	**0.70**	**-0.15**	0.1742	0.1630	0.1822	0.1929	0.1395	**0.2024**	0.1999	0.1928
10	**0.70**	**0.20**	**0.50**	**---**	0.7506	0.7456	0.8150	0.8121	0.7543	0.8066	0.8258	**0.8322**
11	**0.70**	**0.20**	**-0.50**	**---**	0.3915	0.4801	0.6468	0.5482	0.6764	0.5012	0.6489	**0.7268**
12	**0.90**	**-0.40**	**1.20**	**-0.30**	0.7201	0.9735	0.9800	0.8529	0.7698	0.8746	0.8713	**0.9813**
	Average				0.6130	0.6979	0.7603	0.7159	0.7303	0.7095	0.7501	**0.7863**

Table 5.4 Power level of the test statistics $\tilde{Q}_{\mathrm{LB}}$, $\tilde{Q}_{\mathrm{M}}$, $\tilde{Q}_{\mathrm{MM}}$, $\tilde{Q}_{\mathrm{FGLB}}$, $\tilde{Q}_{\mathrm{GFK}}$, $\tilde{Q}_{\mathrm{GFD}}$, $\tilde{Q}_{\mathrm{EXLB}}$ and $\tilde{Q}_{\mathrm{EXM}}$ where data are generated under various alternative ARMA(2,2) processes and fitted by an AR(1) model.

Table 5.4 shows the power levels based on a 5% significance level when data are generated from an ARMA(2, 2) process and an AR(1) model is fitted, with $n = 100$ and $m = 10$. Table 5.4 shows that the $\tilde{Q}_{EXM}$ test is more powerful than other portmanteau tests in 8 cases, but the $\tilde{Q}_{GFK}$ and $\tilde{Q}_{GFD}$ tests are better than the other tests in one case, that is, for models 5 and 9 respectively. The $\tilde{Q}_{M}$, $\tilde{Q}_{MM}$, $\tilde{Q}_{GFK}$, $\tilde{Q}_{EXLB}$ and $\tilde{Q}_{EXM}$ tests are jointly the most powerful in one case, that is, for model 7. The $\tilde{Q}_{MM}$ and $\tilde{Q}_{EXM}$ tests are jointly the most powerful in one case, that is, model 8. The average value has been taken for each test when data are fitted under an AR(1) model with $m = 10$, it illustrates that overall the $\tilde{Q}_{EXM}$ test is better than the other tests. It means that the new $\tilde{Q}_{EXM}$ test, which is based on the partial autocorrelation function, is in general more powerful than other tests when data are fitted by an AR(1) model with $m = 10$.

					$m = 10$							
Model	ϕ_1	ϕ_2	θ_1	θ_2	$\tilde{Q}_{LB}$	$\tilde{Q}_{M}$	$\tilde{Q}_{MM}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
13	**0.50**	---	---	---	0.2836	0.2689	0.3215	0.3438	0.3552	0.3369	**0.3765**	0.3631
14	**0.80**	---	---	---	0.9823	0.9758	0.9903	0.9926	**0.9942**	0.9921	0.9940	0.9933
15	**1.10**	**-0.35**	---	---	0.9961	0.9957	0.9989	0.9989	**0.9997**	0.9989	0.9993	0.9995
16	---	---	**0.80**	**-0.50**	0.8389	0.9375	0.9734	0.9415	0.9487	0.9481	0.9584	**0.9791**
17	---	---	**-0.60**	**0.30**	0.3868	0.4626	0.5940	0.5209	0.6002	0.4784	0.6001	**0.6727**
18	**0.50**	---	**-0.70**	---	0.8773	0.8606	0.9365	0.9405	**0.9648**	0.9338	0.9613	0.9575
19	**-0.50**	---	**0.70**	---	0.8933	0.8763	0.9452	0.9516	**0.9697**	0.9458	0.9660	0.9633
20	**0.30**	---	**0.80**	**-0.50**	0.6265	0.7602	0.8378	0.7518	0.7323	0.7807	0.7857	**0.8579**
21	**0.80**	---	**-0.50**	**0.30**	0.9786	0.9626	0.9847	0.9897	**0.9931**	0.9886	0.9928	0.9898
22	**1.20**	**-0.50**	**0.90**	---	0.4685	**0.7108**	0.6157	0.4761	0.1232	0.4932	0.4300	0.5735
23	**0.30**	**-0.20**	**-0.70**	---	0.2649	0.2852	0.3491	0.3268	0.3088	0.3237	0.3721	**0.3962**
24	**0.90**	**-0.40**	**1.20**	**-0.30**	0.7888	0.9335	0.9571	0.8958	0.8121	0.9085	0.9076	**0.9615**
	Average				0.6988	0.7525	0.7920	0.7608	0.7335	0.7607	0.7787	**0.8090**

Table 5.5 Power level of the test statistics $\tilde{Q}_{\mathrm{LB}}$, $\tilde{Q}_{\mathrm{M}}$, $\tilde{Q}_{\mathrm{MM}}$, $\tilde{Q}_{\mathrm{FGLB}}$, $\tilde{Q}_{\mathrm{GFK}}$, $\tilde{Q}_{\mathrm{GFD}}$, $\tilde{Q}_{\mathrm{EXLB}}$ and $\tilde{Q}_{\mathrm{EXM}}$ where data are generated under various alternative ARMA(2,2) processes and fitted by a MA (1) model.

Table 5.5 shows the power levels based on a 5% significance level when data are generated from an ARMA(2, 2) process and a MA(1) model is fitted, with $n = 100$ and $m = 10$. As is apparent in Table 5.5 the $\tilde{Q}_{EXM}$ test is more powerful than other portmanteau tests in 5 cases, but the $\tilde{Q}_{M}$ and $\tilde{Q}_{EXLB}$ tests are better than other tests in one case, that is, for models 22 and 13 respectively. The $\tilde{Q}_{GFK}$ is better in 5 cases. The average across all models has been calculated for each test, it illustrates that the $\tilde{Q}_{EXM}$ test is better than other tests. This means that the new $\tilde{Q}_{EXM}$ test is more powerful than other tests when data are fitted under a MA(1) model with $m = 10$.

					$m = 20$							
Model	ϕ_1	ϕ_2	θ_1	θ_2	$\tilde{Q}_{LB}$	$\tilde{Q}_{M}$	$\tilde{Q}_{MM}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
1	**---**	**---**	**-0.50**	**---**	0.2197	0.2038	0.2688	0.2614	0.3021	0.3434	0.3292	**0.3859**
2	**---**	**---**	**-0.80**	**---**	0.5932	0.8651	0.9636	0.7804	0.8928	0.9128	0.8919	**0.9879**
3	**---**	**---**	**-0.60**	**0.30**	0.6239	0.9461	0.9838	0.8140	0.9138	0.9363	0.9109	**0.9931**
4	**0.10**	**0.30**	**---**	**---**	0.3624	0.2908	0.3951	0.4487	0.5105	**0.5610**	0.5272	0.5175
5	**1.30**	**-0.35**	**---**	**---**	0.6289	0.5488	0.7288	0.7604	0.8337	0.8392	0.8534	**0.8537**
6	**0.70**	**---**	**-0.40**	**---**	0.4765	0.4828	0.6347	0.6030	0.6810	0.6977	0.7204	**0.7861**
7	**0.70**	**---**	**-0.90**	**---**	0.9302	0.9992	**1**	0.9951	0.9998	0.9998	0.9998	**1**
8	**0.40**	**---**	**-0.60**	**0.30**	0.6874	0.9764	0.9960	0.8748	0.9596	0.9682	0.9577	**0.9988**
9	**0.70**	**---**	**0.70**	**-0.15**	0.1596	0.1192	0.1248	0.1762	0.1770	**0.2262**	0.1955	0.1806
10	**0.70**	**0.20**	**0.50**	**---**	0.6378	0.5977	0.7210	0.7500	0.7931	**0.8247**	0.8038	0.8107
11	**0.70**	**0.20**	**-0.50**	**---**	0.3114	0.2987	0.4392	0.4045	0.4870	0.5102	0.5344	**0.6194**
12	**0.90**	**-0.40**	**1.20**	**-0.30**	0.5661	0.8960	0.9604	0.7313	0.8379	0.8851	0.8320	**0.9817**
	Average				0.5164	0.6021	0.6847	0.6333	0.6990	0.7254	0.7130	**0.7596**

Table 5.6 Power level of the test statistics $\tilde{Q}_{\mathrm{LB}}$, $\tilde{Q}_{\mathrm{M}}$, $\tilde{Q}_{\mathrm{MM}}$, $\tilde{Q}_{\mathrm{FGLB}}$, $\tilde{Q}_{\mathrm{GFK}}$, $\tilde{Q}_{\mathrm{GFD}}$, $\tilde{Q}_{\mathrm{EXLB}}$ and $\tilde{Q}_{\mathrm{EXM}}$ where data are generated under various alternative ARMA(2,2) processes and fitted by an AR(1) model.

Table 5.6 shows power levels based on a 5% significance level when data are generated from an ARMA(2, 2) process and an AR(1) model is fitted, with $n = 100$ and $m = 20$. Table 5.6 shows that the $\tilde{Q}_{EXM}$ test is more powerful than other portmanteau tests in 8 cases, but the $\tilde{Q}_{GFD}$ test is better in three cases. In addition, the $\tilde{Q}_{EXM}$ and $\tilde{Q}_{MM}$ tests are jointly better in 1 case, that is, for model 7. The average value obtained by each test shows that the $\tilde{Q}_{EXM}$ test is better than other tests. This means that the new $\tilde{Q}_{EXM}$ test is, in general, more powerful than other tests when data are fitted under an AR(1) model with $m = 20$.

					$m = 20$							
Model	ϕ_1	ϕ_2	θ_1	θ_2	$\tilde{Q}_{LB}$	$\tilde{Q}_{M}$	$\tilde{Q}_{MM}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
13	**---**	**---**	**-0.50**	**---**	0.2468	0.1816	0.2192	0.2882	0.3194	**0.3593**	0.3465	0.3197
14	**---**	**---**	**-0.80**	**---**	0.9632	0.9310	0.9745	0.9841	0.9911	**0.9925**	0.9912	0.9891
15	**---**	**---**	**-0.60**	**0.30**	0.9879	0.9840	0.9981	0.9979	0.9994	0.9993	**0.9996**	0.9994
16	**0.10**	**0.30**	**---**	**---**	0.7042	0.8146	0.9260	0.8611	0.9313	0.9522	0.9313	**0.9678**
17	**1.30**	**-0.35**	**---**	**---**	0.3168	0.3237	0.4425	0.4110	0.4883	0.5105	0.5340	**0.6110**
18	**0.70**	**---**	**-0.40**	**---**	0.7930	0.7223	0.8643	0.8933	0.9355	0.9405	**0.9425**	0.9363
19	**0.70**	**---**	**-0.90**	**---**	0.8187	0.7520	0.8831	0.9113	0.9463	0.9497	**0.9512**	0.9451
20	**0.40**	**---**	**-0.60**	**0.30**	0.4959	0.5731	0.7214	0.6417	0.7273	0.7895	0.7385	**0.8259**
21	**0.70**	**---**	**0.70**	**-0.15**	0.9624	0.9124	0.9685	0.9826	0.9899	0.9906	**0.9911**	0.9864
22	**0.70**	**0.20**	**0.50**	**---**	0.3866	0.5803	0.6031	0.4463	0.4256	0.5129	0.4570	**0.6371**
23	**0.70**	**0.20**	**-0.50**	**---**	0.2260	0.1939	0.2366	0.2707	0.2964	**0.3516**	0.3307	0.3466
24	**0.90**	**-0.40**	**1.20**	**-0.30**	0.6292	0.8283	0.9189	0.8007	0.8795	0.9141	0.8805	**0.9582**
	Average				0.6276	0.6498	0.7297	0.7074	0.7442	0.7719	0.7578	**0.7936**

Table 5.7 Power level of the test statistics $\tilde{Q}_{\mathrm{LB}}$, $\tilde{Q}_{\mathrm{M}}$, $\tilde{Q}_{\mathrm{MM}}$, $\tilde{Q}_{\mathrm{FGLB}}$, $\tilde{Q}_{\mathrm{GFK}}$, $\tilde{Q}_{\mathrm{GFD}}$, $\tilde{Q}_{\mathrm{EXLB}}$ and $\tilde{Q}_{\mathrm{EXM}}$ where data are generated under various alternative ARMA(2,2) processes and fitted by a MA(1) model.

Table 5.7 shows the power levels based on a 5% significance level, when data are generated from an ARMA(2, 2) process and a MA(1) model is fitted, with n = 100 and m = 20. It is evident from Table 5.7 that the $\tilde{Q}_{EXM}$ test is more powerful than other portmanteau tests in 5 cases, while the $\tilde{Q}_{EXLB}$ test is better in 4 cases. The $\tilde{Q}_{GFD}$ is better in three cases. The average value calculated for each test illustrates that the $\tilde{Q}_{EXM}$ test is better than the others. It means that the new $\tilde{Q}_{EXM}$ test is, in general, more powerful than other the tests when data are fitted under a MA(1) model with m = 20.

					$m = 30$							
Model	ϕ_1	ϕ_2	θ_1	θ_2	$\tilde{Q}_{LB}$	$\tilde{Q}_{M}$	$\tilde{Q}_{MM}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
1	**---**	**---**	**-0.50**	**---**	0.2167	0.1628	0.1869	0.2493	0.2514	**0.3668**	0.3132	0.3571
2	**---**	**---**	**-0.80**	**---**	0.5426	0.7478	0.9060	0.7081	0.8107	0.9206	0.8621	**0.9797**
3	**---**	**---**	**-0.60**	**0.30**	0.5747	0.8577	0.9522	0.7355	0.8307	0.9302	0.8736	**0.9910**
4	**0.10**	**0.30**	**---**	**---**	0.3442	0.2221	0.2939	0.4223	0.4510	**0.5787**	0.5066	0.4841
5	**1.30**	**-0.35**	**---**	**---**	0.5924	0.4335	0.6076	0.7145	0.7611	**0.8342**	0.8273	0.8228
6	**0.70**	**---**	**-0.40**	**---**	0.4215	0.3433	0.4760	0.5318	0.5804	0.6953	0.6685	**0.7250**
7	**0.70**	**---**	**-0.90**	**---**	0.8817	0.9957	0.9999	0.9816	0.9973	0.9994	0.9995	**1**
8	**0.40**	**---**	**-0.60**	**0.30**	0.6307	0.9292	0.9887	0.8078	0.9043	0.9677	0.9379	**0.9983**
9	**0.70**	**---**	**0.70**	**-0.15**	0.1697	0.0932	0.0836	0.1723	0.1533	**0.2367**	0.1895	0.1675
10	**0.70**	**0.20**	**0.50**	**---**	0.5986	0.4971	0.6353	0.7183	0.7633	**0.8336**	0.7916	0.7916
11	**0.70**	**0.20**	**-0.50**	**---**	0.3303	0.2613	0.3392	0.3890	0.4108	0.5420	0.5158	**0.5828**
12	**0.90**	**-0.40**	**1.20**	**-0.30**	0.5148	0.7824	0.9101	0.6634	0.7538	0.8861	0.7971	**0.9731**
	Average				0.4848	0.5272	0.6150	0.5912	0.6390	0.7326	0.6902	**0.7394**

Table 5.8 Power level of the test statistics $\tilde{Q}_{\mathrm{LB}}$, $\tilde{Q}_{\mathrm{M}}$, $\tilde{Q}_{\mathrm{MM}}$, $\tilde{Q}_{\mathrm{FGLB}}$, $\tilde{Q}_{\mathrm{GFK}}$, $\tilde{Q}_{\mathrm{GFD}}$, $\tilde{Q}_{\mathrm{EXLB}}$ and $\tilde{Q}_{\mathrm{EXM}}$ where data are generated under various alternative ARMA(2,2) processes and fitted by an AR(1) model.

Table 5.8 shows the power levels based on a 5% significance level, when data are generated from an ARMA(2, 2) process and with an AR(1) model is fitted, with n = 100 and m = 30. Table 5.8 shows that the $\tilde{Q}_{EXM}$ test is more powerful than other portmanteau tests in 7 cases, but the $\tilde{Q}_{GFD}$ test is better in 5 cases. The average value calculated for each test illustrates that the $\tilde{Q}_{EXM}$ test is better than other tests. It means that the new $\tilde{Q}_{EXM}$ test is more powerful than other tests when data are fitted under an AR(1) model with m = 30.

					$m = 30$							
Model	ϕ_1	ϕ_2	θ_1	θ_2	$\tilde{Q}_{LB}$	$\tilde{Q}_{M}$	$\tilde{Q}_{MM}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
13	**0.50**	**---**	**---**	**---**	0.2394	0.1400	0.1535	0.2753	0.2854	**0.3850**	0.3394	0.2998
14	**0.80**	**---**	**---**	**---**	0.9487	0.8841	0.9549	0.9785	0.9857	**0.9917**	0.9889	0.9845
15	**1.10**	**-0.35**	**---**	**---**	0.9760	0.9606	0.9908	0.9939	0.9974	**0.9989**	0.9980	0.9982
16	**---**	**---**	**0.80**	**-0.50**	0.6528	0.6949	0.8509	0.8017	0.8735	0.9467	0.9055	**0.9547**
17	**---**	**---**	**-0.60**	**0.30**	0.3097	0.2560	0.3349	0.3842	0.4072	0.5241	0.5007	**0.5672**
18	**0.50**	**---**	**-0.70**	**---**	0.7618	0.6188	0.7853	0.8670	0.9014	**0.9423**	0.9280	0.9154
19	**-0.50**	**---**	**0.70**	**---**	0.7847	0.6501	0.8149	0.8860	0.9209	**0.9571**	0.9432	0.9311
20	**0.30**	**---**	**0.80**	**-0.50**	0.4599	0.4460	0.6012	0.5847	0.6551	**0.7985**	0.7132	0.7963
21	**0.80**	**---**	**-0.50**	**0.30**	0.9487	0.8642	0.9418	0.9757	0.9824	**0.9900**	0.9875	0.9800
22	**1.20**	**-0.50**	**0.90**	**---**	0.3553	0.4514	0.5073	0.4200	0.4300	0.5257	0.4608	**0.6497**
23	**0.30**	**-0.20**	**-0.70**	**---**	0.2230	0.1515	0.1618	0.2523	0.2515	**0.3620**	0.3141	0.3184
24	**0.90**	**-0.40**	**1.20**	**-0.30**	0.5758	0.7016	0.8448	0.7261	0.8174	0.9150	0.8485	**0.9474**
	Average				0.6030	0.5683	0.6618	0.6788	0.7090	0.7781	0.7440	**0.7786**

Table 5.9 Power level of the test statistics $\tilde{Q}_{\mathrm{LB}}$, $\tilde{Q}_{\mathrm{M}}$, $\tilde{Q}_{\mathrm{MM}}$, $\tilde{Q}_{\mathrm{FGLB}}$, $\tilde{Q}_{\mathrm{GFK}}$, $\tilde{Q}_{\mathrm{GFD}}$, $\tilde{Q}_{\mathrm{EXLB}}$ and $\tilde{Q}_{\mathrm{EXM}}$ where data are generated under various alternative ARMA(2,2) processes and fitted by a MA(1) model.

Table 5.9 shows the power levels based on a 5% significance level, when data are generated from an ARMA(2, 2) process and with a MA(1) model fitted, with n = 100 and m = 30. From Table 5.9 it is clear that the $\tilde{Q}_{GFD}$ test is more powerful than other portmanteau tests in 8 cases, but the $\tilde{Q}_{EXM}$ test is better in 4 cases. The average value has been taken for each test, which illustrates that the $\tilde{Q}_{EXM}$ test is the best overall. This means that the new $\tilde{Q}_{GFD}$ test is more powerful than other tests when data are fitted under a MA(1) model with m = 30.

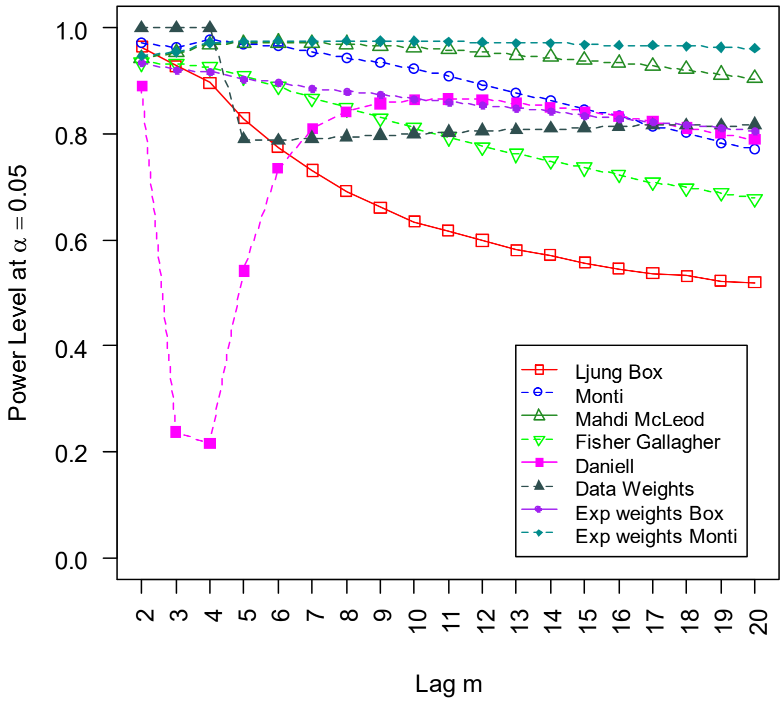

Figure 5.3 Power level for lags from 2 to 20 for a correctly fitted AR(1) model, data generated by a MA(1) process with $\theta_1 = -0.8$, at 5% significance level, series of length $n = 85$.

Figure 5.3 shows the power level for lags from 2 to 20 based on a 5% significance level, when data are generated by a MA(1) process with $\theta = -0.8$ and fitted under an AR(1) model with n = 85 (following the simulation of Gallagher and Fisher (2015)) and 10,000 replications. The power of the $\tilde{Q}_{LB}$, $\tilde{Q}_M$ and $\tilde{Q}_{FGLB}$ tests decreases as the lag increases. The $\tilde{Q}_{MM}$ and $\tilde{Q}_{EXLB}$ tests slowly decrease as the lag increases. The $\tilde{Q}_{GFK}$ test rapidly decreases at lags 3 and 4, then slowly increases as the lag increases. The $\tilde{Q}_{GFD}$ test rapidly decreases at lag 5, then slowly increases as the lag increases. From Figure 5.3 it is apparent that this test rejects all models, even correct ones, for the lags examined. The $\tilde{Q}_{EXM}$ test remains constant as the lag increases. In most cases, the $\tilde{Q}_{EXM}$ is the most powerful tests.

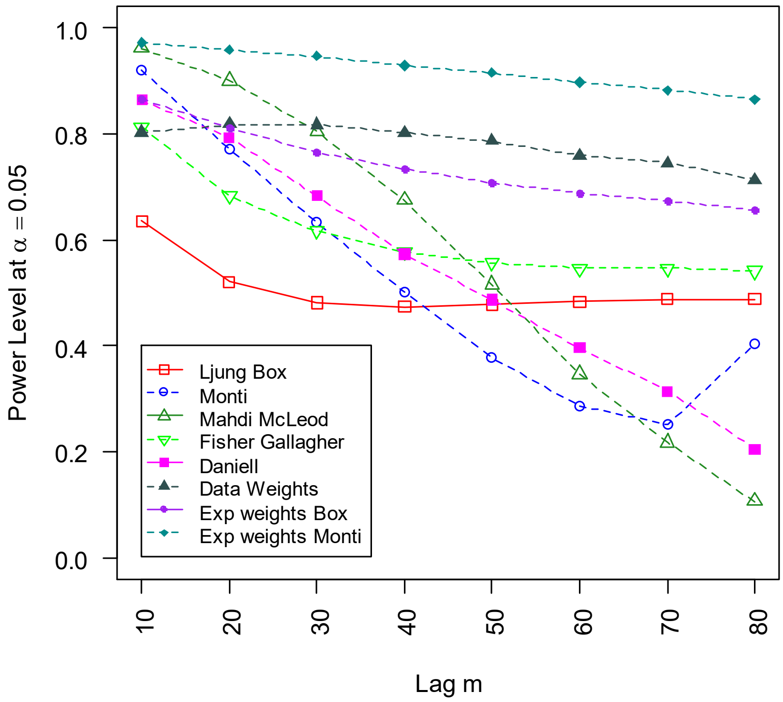

Figure 5.4 Power level for maximum lags for a correctly fitted AR(1) model, data generated by a MA(1) process with $\theta_1 = -0.8$, at 5% significance level, series of length $n = 85$.

Figure 5.4 shows the power level of large lags based on a 5% significance level, when data are generated by a MA(1) process with $\theta = -0.8$ and fitted under an AR(1) model with n = 85 and 10,000 replications. The power of the $\tilde{Q}_{GFD}$, $\tilde{Q}_{EXLB}$ and $\tilde{Q}_{EXM}$ tests slowly decreases as the lag increases. The $\tilde{Q}_{LB}$ and $\tilde{Q}_{FGLB}$ tests have similar behaviour to each other, initially slowly decreasing and then remaining constant as the lag increases further. The power of the $\tilde{Q}_M$, $\tilde{Q}_{MM}$ and $\tilde{Q}_{GFK}$ tests decreases as the lag increases. In all cases, the $\tilde{Q}_{EXM}$ is the most powerful test.

The next study is similar to Gallagher and Fisher (2015), where data are generated under an ARMA(2,2) process and are fitted by an ARMA(1,1) model, see Tables 5.10 and 5.11.

Following on from previously published research in this area, m was set at 10 and 20, and n was set at 100, and in each case, the critical value was determined from the corresponding asymptotic distribution.

					$m = 10$							
Model	ϕ_1	ϕ_2	θ_1	θ_2	$\tilde{Q}_{LB}$	$\tilde{Q}_{M}$	$\tilde{Q}_{MM}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
4	**0.10**	**0.30**	**---**	**---**	0.2988	0.2681	0.3252	0.3497	0.2951	0.3310	**0.4343**	0.4148
5	**1.30**	**-0.35**	**---**	**---**	0.1183	0.1292	0.123	0.1105	0.0659	0.1001	0.1552	**0.1651**
9	**0.70**	**---**	**0.70**	**-0.15**	0.1180	0.1121	0.1032	0.1107	0.0534	0.1068	**0.1483**	0.1446
10	**0.70**	**0.20**	**0.50**	**---**	0.1663	0.1538	0.1737	0.1863	0.1528	0.1715	**0.2490**	0.2407
12	**0.90**	**-0.40**	**1.20**	**-0.30**	0.3856	0.3902	0.4289	0.4338	0.2873	0.4325	**0.5132**	0.5032
15	**1.10**	**-0.35**	---	---	0.1428	**0.1460**	0.1193	0.1168	0.0201	0.097	0.1383	0.1459
16	**---**	**---**	**0.80**	**-0.50**	0.3664	0.4477	0.4691	0.3859	0.1234	0.4323	0.4412	**0.5085**
17	**---**	**---**	**-0.60**	**0.30**	0.1194	0.1179	0.1188	0.1151	0.0598	0.1033	0.1565	**0.1679**
20	**0.30**	**---**	**0.80**	**-0.50**	0.3986	0.4626	0.4941	0.4332	0.1756	0.4683	0.4867	**0.5383**
21	**0.80**	**---**	**-0.50**	**0.30**	0.1195	0.1329	0.1241	0.1062	0.0364	0.0955	0.1421	**0.1619**
22	**1.20**	**-0.50**	**0.90**	**---**	0.4549	**0.7459**	0.6585	0.4605	0.0500	0.4703	0.4895	0.6768
23	**0.30**	**-0.20**	**-0.70**	**---**	0.1836	0.1829	0.1896	0.1948	0.0894	0.2044	0.2391	**0.2412**
	Average				0.2339	0.2741	0.2773	0.2503	0.1174	0.2511	0.2995	**0.3257**

Table 5.10 Power level of the test statistics $\tilde{Q}_{\mathrm{LB}}$, $\tilde{Q}_{\mathrm{M}}$, $\tilde{Q}_{\mathrm{MM}}$, $\tilde{Q}_{\mathrm{FGLB}}$, $\tilde{Q}_{\mathrm{GFK}}$, $\tilde{Q}_{\mathrm{GFD}}$, $\tilde{Q}_{\mathrm{EXLB}}$ and $\tilde{Q}_{\mathrm{EXM}}$ where data are generated under various alternative ARMA(2,2) processes, and fitted by an ARMA(1,1) model $m = 10$.

Table 5.10 shows the power levels based on a 5% significance level, when data are generated from an ARMA(2, 2) process and an ARMA(1,1) model is fitted, with $n = 100$ and $m = 10$. Table 5.10 demonstrates that the $\tilde{Q}_{EXM}$ test is more powerful than other portmanteau tests in 7 cases, but the $\tilde{Q}_{EXLB}$ test is better in 4 cases, and the $\tilde{Q}_{EXLB}$ and $\tilde{Q}_{M}$ tests are best in 1 case each. The average value has been taken for each test, which illustrates that the $\tilde{Q}_{EXM}$ test is better than other tests. This means that the new $\tilde{Q}_{EXM}$ test is more powerful than other tests when data are fitted under an ARMA(1,1) model with $m = 10$.

					$m = 20$							
Model	ϕ_1	ϕ_2	θ_1	θ_2	$\tilde{Q}_{LB}$	$\tilde{Q}_{M}$	$\tilde{Q}_{MM}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
4	**0.10**	**0.30**	**---**	**---**	0.2440	0.1774	0.2106	0.2724	0.2720	**0.3615**	0.3525	0.3247
5	**1.30**	**-0.35**	**---**	**---**	0.1105	0.0885	0.0704	0.0841	0.0595	0.1131	0.1110	**0.1139**
9	**0.70**	**---**	**0.70**	**-0.15**	0.1194	0.0899	0.0706	0.1026	0.0742	**0.1294**	0.1247	0.1078
10	**0.70**	**0.20**	**0.50**	**---**	0.1449	0.1053	0.1052	0.1409	0.1310	**0.1954**	0.1936	0.1795
12	**0.90**	**-0.40**	**1.20**	**-0.30**	0.2996	0.3140	0.3320	0.3461	0.3472	**0.4625**	0.4367	0.4288
15	**1.10**	**-0.35**	---	---	0.1298	0.1141	0.0883	0.1119	0.0711	0.1185	0.1233	**0.1211**
16	**---**	**---**	**0.80**	**-0.50**	0.2940	0.3063	0.3569	0.3201	0.2812	**0.4711**	0.3888	0.4525
17	**---**	**---**	**-0.60**	**0.30**	0.1451	0.1062	0.0999	0.1168	0.0840	0.1375	0.1448	**0.1472**
20	**0.30**	**---**	**0.80**	**-0.50**	0.3143	0.3168	0.3699	0.3422	0.3120	**0.4997**	0.4138	0.4635
21	**0.80**	**---**	**-0.50**	**0.30**	0.1125	0.0999	0.0795	0.0907	0.0600	0.1115	0.1119	**0.1217**
22	**1.20**	**-0.50**	**0.90**	**---**	0.3466	0.6033	0.6279	0.3828	0.3069	0.4911	0.4428	**0.6838**
23	**0.30**	**-0.20**	**-0.70**	**---**	0.1641	0.1307	0.1295	0.1624	0.1384	**0.2365**	0.2005	0.1938
	Average				0.2021	0.2044	0.2117	0.2061	0.1781	0.2773	0.2537	**0.2782**

Table 5.11 Power level of the test statistics $\tilde{Q}_{\mathrm{LB}}$, $\tilde{Q}_{\mathrm{M}}$, $\tilde{Q}_{\mathrm{MM}}$, $\tilde{Q}_{\mathrm{FGLB}}$, $\tilde{Q}_{\mathrm{GFK}}$, $\tilde{Q}_{\mathrm{GFD}}$, $\tilde{Q}_{\mathrm{EXLB}}$ and $\tilde{Q}_{\mathrm{EXM}}$ when data are generated under various alternative ARMA(2,2) processes, and fitted by an ARMA(1,1) model $m = 20$.

Table 5.11 shows the power levels based on a 5% significance level, when data are generated from an ARMA(2, 2) process and an ARMA(1,1) model is fitted, with $n = 100$ and $m = 20$. Table 5.11 shows that the $\tilde{Q}_{EXM}$ test is more powerful than other portmanteau tests in 5 cases, but the $\tilde{Q}_{GFD}$ test is better in 7 cases. The average value calculated for each test illustrates that the $\tilde{Q}_{EXM}$ test is better than other tests. Generally, the new $\tilde{Q}_{EXM}$ test is more powerful than other tests in most cases when data fitted under an ARMA(1,1) model with $m = 10$ or 20.

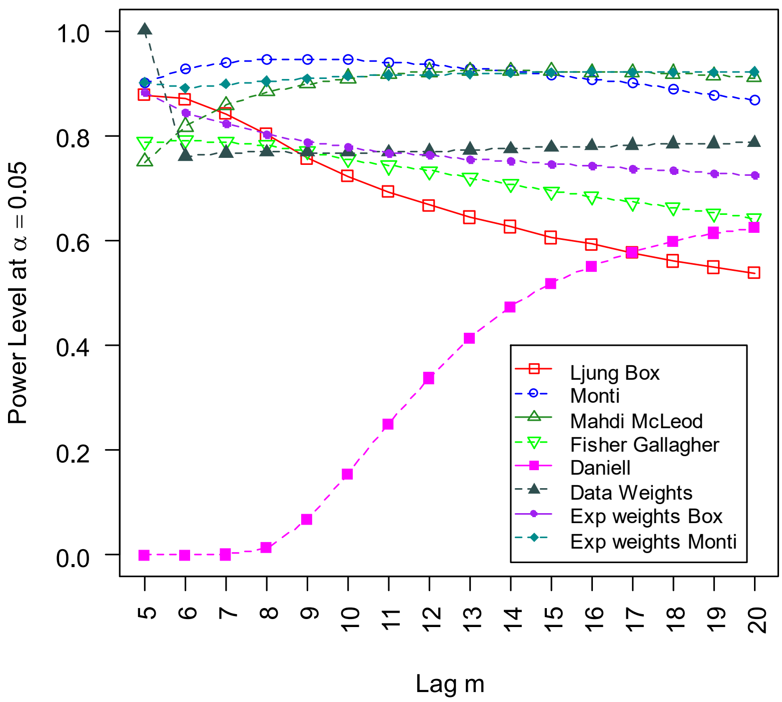

Figure 5.5 Power level for lags from 5 to 20 for a correctly fitted ARMA(1,1) model, data generated by an ARMA(2,1) process with ϕ_1=1.2, ϕ_2= -0.5 and θ_1= -0.9, at 5% significance level, series of length $n = 150$.

Figure 5.5 shows the power level for lags from 5 to 20 based on a 5% significance level, when data are generated by an ARMA(2,1) process with ϕ_1= 1.2, ϕ_2=-0.5 and $\theta_1 = -0.9$, and fitted under an ARMA(1,1) model with n = 150 and 10,000 replications. The power of the $\tilde{Q}_{LB}$, $\tilde{Q}_{EXLB}$ and $\tilde{Q}_{FGLB}$ tests decreases as the lag increases. The $\tilde{Q}_M$ test increases up to lag 10, then it slowly decreases as the lag increases. The $\tilde{Q}_{MM}$ test increases as the lag increases. The $\tilde{Q}_{GFK}$ test is stable at lags 5, 6 and 7, then rapidly increases as the lag increases. The $\tilde{Q}_{GFD}$ test rapidly decreases at lag 6, then slowly increases as the lag increases further. The power level of the $\tilde{Q}_{EXM}$ test is approximately constant as the lag increases. In general, the $\tilde{Q}_{EXM}$ is the most powerful test.

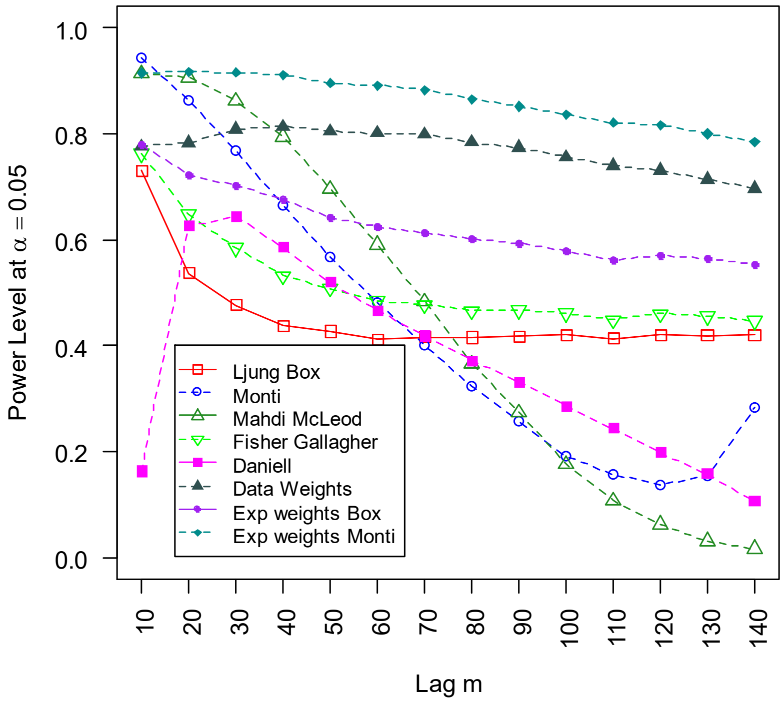

Figure 5.6 Power level for maximum lags for a correctly fitted ARMA(1,1) model, data generated by an ARMA(2,1) process with ϕ_1=1.2, ϕ_2= -0.5 and θ_1= -0.9, at 5% significance level, series of length $n = 150$.

Figure 5.6 shows the power level for large lags based on a 5% significance level, when data are generated by an ARMA(2,1) process with $\phi_1 = 1.2$, $\phi_2 = -0.5$ and $\theta_1 = -0.9$, and fitted under an ARMA(1,1) model with n = 150 and 10,000 replications. The results of Figure 5.6 are similar to the results of Figure 5.4, except in the case of the $\tilde{Q}_{GFK}$ test. The value of the $\tilde{Q}_{GFK}$ test is less than 0.2 at lag 10, rapidly increasing at lag 20 and then decreasing as lag increases further.

5.4 Variability of the univariate portmanteau test

To find out the variability of the new univariate portmanteau test statistics a Monte-Carlo experiment was conducted with 1000 replications of the experiment to determine the empirical size carried out in Section 5.3.3. As this experiment involved 1000 replications of a simulations involving 10,000 replications this is only conducted for one value of ϕ_1. The aim was to calculate the mean and standard deviation for the test statistics $\tilde{Q}_{LB}$, $\tilde{Q}_M$, $\tilde{Q}_{FGLB}$, $\tilde{Q}_{GFK}$, $\tilde{Q}_{GFD}$, $\tilde{Q}_{EXLB}$ and $\tilde{Q}_{EXM}$, when n = 100 observations under an AR(1) process $z_t - \phi\, z_{t-1} = e_t$ with parameter $\phi_1 = 0.5$ and maximum lags m = 10. The test statistics $\tilde{Q}_{LB}$, $\tilde{Q}_M$, $\tilde{Q}_{FGLB}$, $\tilde{Q}_{GFK}$, $\tilde{Q}_{GFD}$, $\tilde{Q}_{EXLB}$ and $\tilde{Q}_{EXM}$ were calculated by using the steps of the previous Monte-Carlo experiment.

	m =10						
$\phi_1 = 0.5$	$\tilde{Q}_{LB}$	$\tilde{Q}_M$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
Mean	0.052669	0.053590	0.034766	0.015218	0.031812	0.033215	0.032987
Standard deviation	0.00231	0.002312	0.001943	0.001163	0.001684	0.001818	0.001884

Table 5.12 The mean and standard deviation for the test statistics $\tilde{Q}_{LB}$, $\tilde{Q}_M$, $\tilde{Q}_{FGLB}$, $\tilde{Q}_{GFK}$, $\tilde{Q}_{GFD}$, $\tilde{Q}_{EXLB}$ and $\tilde{Q}_{EXM}$, data generated under an AR(1) process with parameter ϕ_1= 0.5, and fitted under AR(1) with n = 100 and maximum lags m = 10.

As can be seen from Table 5.12 the results of this experiment are consistent with those from Table 5.1. The standard deviations for the new tests and previous tests are around 5 percent of their associate mean, in most cases.

A second experiment was conducted to examine the variability of the new test when applied to the power study in Section 5.3.4. Again 1000 replications of the original experiment were undertaken for a single choice of θ_1 (i.e. -0.8).

	m =10						
$\theta_1 = -0.8$	$\tilde{Q}_{LB}$	$\tilde{Q}_{M}$	$\tilde{Q}_{FGLB}$	$\tilde{Q}_{GFK}$	$\tilde{Q}_{GFD}$	$\tilde{Q}_{EXLB}$	$\tilde{Q}_{EXM}$
Mean	0.75297	0.964627	0.905353	0.938861	0.850433	0.939353	0.990351
Standard deviation	0.002310	0.002312	0.001943	0.001163	0.001684	0.001818	0.001884

Table 5.13 The mean and standard deviation for the test statistics $\tilde{Q}_{LB}$, $\tilde{Q}_{M}$, $\tilde{Q}_{FGLB}$, $\tilde{Q}_{GFK}$, $\tilde{Q}_{GFD}$, $\tilde{Q}_{EXLB}$ and $\tilde{Q}_{EXM}$, data generated under an MA(1) process with parameter $\theta_1 = -0.8$, and fitted under AR(1) with n = 100 and maximum lags m = 10.

As can be seen in Table 5.13 the standard deviations are less than 1 percent of the associate mean for the new tests and almost all the previous tests.

5.5 Summary

The empirical size simulations (see Tables 5.1, 5.2 and 5.3) show that the Monti $\tilde{Q}_{MM}$ test is better than other tests when data are generated from an AR(1) process and fitted by an AR(1) model with n = 100 and m =10, 20 and 30. The empirical size Figure 5.1 shows that portmanteau tests from previous studies do not have significant levels that are stable with respect to lag length. Figure 5.2, the exponential weighted portmanteau test is not affected by lag length, which means the $\tilde{Q}_{EXM}$ test is stable with respect to lag length.

The power level simulations (see Tables 5.4, 5.5, 5.6, 5.7, 5.8 and 5.9) show that the new portmanteau $\tilde{Q}_{EXM}$ test is more powerful than previous tests, when data are generated from an ARMA(2,2) process and fitted under either an AR(1) model or a MA(1) model with n = 100 and m = 10, 20 and 30. The average power level for each portmanteau test given in these tables shows that the new test $\tilde{Q}_{EXM}$ is the best. When data are fitted under an AR(1) model and m = 30 the new test $\tilde{Q}_{EXM}$ is better than the previous tests in 7 cases, while data fitted under a MA(1) model with m = 30, the $\tilde{Q}_{GFD}$ test is better in 8 cases. The power level Figures 5.3 and 5.4 show that the $\tilde{Q}_{EXM}$ test is more powerful than those from the previous studies in both cases of small and large lags.

Furthermore, when data are generated from an ARMA(2,2) process and fitted under an ARMA(1,1) model with m = 10, the results of the power level simulations given in Tables 5.10 and 5.11 show that the new test $\tilde{Q}_{EXM}$ is better in 6 cases. When m = 20 the new test is better in 5 cases. So the new $\tilde{Q}_{EXM}$ test is more powerful in the power levels when m = 10 and the $\tilde{Q}_{GFD}$ test when m = 10. The average value of the power level tables for each portmanteau test shows that the new test $\tilde{Q}_{EXM}$ is better than the previous tests. The power level experiments given in Figures 5.5 and 5.6 show that the $\tilde{Q}_{EXM}$ test is more powerful than previous portmanteau tests with small and large lags.

Monte-Carlo studies of the variability of the new portmanteau tests show the standard deviations to be low in comparison to their associated means, and also low in comparison to the standard deviation of some of the previous tests.

Chapter 6 - Portmanteau Tests Of Vector Time Series

The aim of this chapter is to review previous studies in the area of portmanteau tests of vector autoregressive moving average models, which have been developed by several researchers. This chapter will focus on developing a new portmanteau test, which is based on exponential weights of the residual covariance matrix. The performance of the new portmanteau test is then compared with previous studies by the use of Monte Carlo experiments using the R program.

6.1 Introduction

A major extension of portmanteau tests has been the application to vector autoregressive moving average (VARMA) time series models. In the univariate time series case, portmanteau tests are based on the residual of the autocorrelation and partial autocorrelation functions. In the vector time series case, portmanteau tests are based on the residual of the covariance matrices and the cross correlation matrices. The first application of a portmanteau test to multivariate autoregressive models was by Chitturi (1974). Since then, Portmanteau tests for VARMA(p, q) models have been developed by many researchers, such as, Hosking (1980), Poskitt and Tremayne (1982) and Li and McLeod (1981).

A portmanteau test of a vector autoregressive moving average model is calculated by summing the residuals of the covariance or autocorrelation matrix of the fitted model. Then the value of the portmanteau test is compared with a critical value. If the value of the portmanteau test is less than the critical value, it means the model is an appropriate one for the data. Alternatively, if the value of the portmanteau test is bigger than the critical value, it means that the model is inappropriate for the data.

Consider a vector time series $\{\boldsymbol{z}_t\}$ generated by a stationary and invertible VARMA(p, q) process given by

$$\boldsymbol{z}_t = \boldsymbol{\Phi}_1 \boldsymbol{z}_{t-1} + \boldsymbol{\Phi}_2 \boldsymbol{z}_{t-2} + \cdots + \boldsymbol{\Phi}_p \boldsymbol{z}_{t-p} + \boldsymbol{e}_t - \boldsymbol{\Theta}_1 \boldsymbol{e}_{t-1} - \boldsymbol{\Theta}_2 \boldsymbol{e}_{t-2} - \cdots - \boldsymbol{\Theta}_q \boldsymbol{e}_{t-q}$$

$$\left(\boldsymbol{I} - \boldsymbol{\Phi}_1 B - \cdots - \boldsymbol{\Phi}_p B^p\right)\boldsymbol{z}_t = \left(\boldsymbol{I} - \boldsymbol{\Theta}_1 B - \cdots - \boldsymbol{\Theta}_q B^q\right)\boldsymbol{e}_t \qquad (6.1)$$

$$\mathbf{\Phi}(B)\boldsymbol{z}_t = \mathbf{\Theta}(B)\boldsymbol{e}_t \tag{6.2}$$

where $\boldsymbol{z}_t = (z_{1t}, z_{2t}, \dots, z_{dt})^\dagger$ is a $d \times 1$ vector of variables observed at time t, $\boldsymbol{e}_t = (e_{1t}, e_{2t}, \dots, e_{dt})^\dagger$, is a $d \times 1$ zero mean white noise process with covariance matrix $\boldsymbol{\Sigma} = E[\boldsymbol{e}_t\boldsymbol{e}_t{}^\dagger]$, $\mathbf{\Phi}_i$ $(i = 1, 2, \dots, p)$ are $d \times d$ parameter matrices and $\mathbf{\Theta}_j$ is a $d \times d$ matrix of coefficients, for $(j = 1, 2, \dots, q)$ and $\mathbf{\Phi}(B)$ and $\mathbf{\Theta}(B)$ are matrix polynomials of the backshift operator B of order p and q respectively.

The parameters matrices $\mathbf{\Phi}(B)$ and $\mathbf{\Theta}(B)$ can be estimated fitted by use of the conditional likelihood method (see Equation 3.56) to obtain the fitted models

$$\hat{\boldsymbol{e}}_t = \widehat{\mathbf{\Phi}}(B)\boldsymbol{z}_t\widehat{\mathbf{\Theta}}^{-1}(B) \tag{6.3}$$

The residuals $\hat{\boldsymbol{e}}_t$ are computed by

$$\hat{\boldsymbol{e}}_t = \boldsymbol{z}_t - \sum_{j=1}^{p} \widehat{\mathbf{\Phi}}_j \, z_{t-j} + \sum_{j=1}^{q} \widehat{\mathbf{\Theta}}_j \, \hat{\boldsymbol{e}}_{t-j} \qquad t = 1,2,\cdots,n \tag{6.4}$$

For the VARMA(p,q) model to be correct the residuals $\hat{\boldsymbol{e}}_t$ need to be approximately zero, this means the autocovariance matrix of the residuals $\boldsymbol{\Gamma}(\boldsymbol{k})$ will be zero at all lags k. This gives the null hypothesis for all lags k

$$H_0\colon \boldsymbol{\Gamma}(\boldsymbol{k}) = 0$$

$$H_1\colon \boldsymbol{\Gamma}(\boldsymbol{k}) \neq 0.$$

6.1.1 Hosking vector portmanteau tests

Hosking (1980, 1981) gave a general form of a multivariate portmanteau test statistic for VARMA(p,q) models, which is based on the residual autocorrelation matrix, it can be written as

$$\tilde{Q}_H = n \sum_{k=1}^{m} \left(\mathrm{vec}(\widehat{\boldsymbol{R}}_k)\right)^\dagger (\widehat{\boldsymbol{R}}_0^{-1} \otimes \widehat{\boldsymbol{R}}_0^{-1})\mathrm{vec}(\widehat{\boldsymbol{R}}_k) \tag{6.5}$$

This is asymptotically chi-squared distributed with degrees of freedom $d^2(m-p-q)$, where n is the series length, m is the maximum lag and $\widehat{\boldsymbol{R}}_k$ is the sample autocorrelation matrix at lag k. The residual of the autocorrelation matrix can be found by

$$\widehat{\boldsymbol{R}}_k = \widehat{\boldsymbol{L}}^{\dagger}\widehat{\boldsymbol{\Gamma}}_k\widehat{\boldsymbol{L}} \tag{6.6}$$

where

$$\widehat{\boldsymbol{\Gamma}(\boldsymbol{k})} = n^{-1}\sum_{t=k+1}^{n} \hat{\boldsymbol{e}}_t\,\hat{\boldsymbol{e}}_{t-k}^{\dagger}$$

is a sample autocovariance matrix and $\boldsymbol{L}$ is a lower triangular matrix such that $\boldsymbol{L}\boldsymbol{L}^{\dagger} = \boldsymbol{\Gamma}(\boldsymbol{0})^{-1}$.

Hosking (1980) gave the modified multivariate portmanteau test statistic, which is based on the residual autocovariance matrix, it can be written as

$$\tilde{Q}_H^{*} = n^2\sum_{k=1}^{m}(n-k)^{-1}\mathrm{tr}\,[\boldsymbol{\Gamma}(k)^{\dagger}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Gamma}(k)\boldsymbol{\Sigma}^{-1}] \tag{6.7}$$

where

$$\boldsymbol{\Gamma}(k) = n^{-1}\sum_{t=1}^{n-1}\hat{\boldsymbol{e}}_t\hat{\boldsymbol{e}}_{t-k}^{\dagger}$$

and tr (trace) is the sum of the diagonal matrix. This is asymptotically chi-squared distributed with degrees of freedom $d^2(m-p-q)$.

6.1.2 Li and McLeod vector portmanteau test

Li and McLeod (1981) provided another multivariate portmanteau test statistic, which is based on the autocorrelation matrix, it can be written as

$$\tilde{Q}_{LM} = n\sum_{k=1}^{m}\left(\mathrm{vec}\left(\widehat{\boldsymbol{R}}_k^{(*)}\right)\right)^{\dagger}\left(\widehat{\boldsymbol{R}}_0^{-1}\otimes\widehat{\boldsymbol{R}}_0^{-1}\right)\mathrm{vec}\left(\widehat{\boldsymbol{R}}_k^{(*)}\right) \tag{6.8}$$

which is asymptotically chi-squared distributed with degrees of freedom $d^2(m-p-q)$, where

$$\widehat{\boldsymbol{R}}_k^{(*)} = \left(\hat{r}_{i,j}(k)\right)_{d\times d}$$

and $\hat{r}_{i,j}(k) = \hat{\gamma}_{i,j}(k)/\sqrt{\hat{\gamma}_{i,i}(0)\hat{\gamma}_{i,i}(0)}$, $i,j = 1,2,\dots,d$, $\hat{\gamma}_{i,j}(k) = n^{-1}\sum_{t=k+1}^{n}\hat{\boldsymbol{e}}_{i,t}\,\hat{\boldsymbol{e}}_{j,t-k}^{\dagger}$

Li and McLeod (1981) recommended a multivariate modified portmanteau test statistic, which is defined as

$$\tilde{Q}_{LM}^{*} = \tilde{Q}_H + \frac{d^2 m(m+1)}{2n} \tag{6.9}$$

which is asymptotically chi-squared distributed with degrees of freedom $d^2(m-p-q)$.

6.1.3 Mahdi and McLeod vector portmanteau test

Mahdi and McLeod (2011) proposed another multivariate portmanteau test statistic, which is based on the residual autocorrelation matrix as

$$\tilde{Q}_{MMV} = -n\log\left|\widehat{\mathfrak{R}}_m\right| \tag{6.10}$$

where

$$\widehat{\mathfrak{R}}_m = \begin{pmatrix} I_d & \widehat{\boldsymbol{R}}_1 & \cdots & \widehat{\boldsymbol{R}}_m \\ \widehat{\boldsymbol{R}}_1^{\dagger} & I_d & \cdots & \widehat{\boldsymbol{R}}_{m-1} \\ \vdots & \vdots & \cdots & \vdots \\ \widehat{\boldsymbol{R}}_m^{\dagger} & \widehat{\boldsymbol{R}}_{m-1}^{\dagger} & \cdots & I_d \end{pmatrix}$$

The $\tilde{Q}_{MMV}$ test is approximately distributed as $a\chi_b^2$, where

$$a = \frac{2m+1}{3}$$

$$b = \frac{3d^2 m(m+1)}{2(2m+1)} - d^2(p+q).$$

6.2 A new weighted portmanteau test of vector ARMA models

The new weighted portmanteau test statistics of univariate time series models developed in Chapter 5 are based on the residual autocorrelation and partial autocorrelation functions with exponential weights. This chapter provides a new portmanteau test statistic for vector time series models, which is based on exponential weights of the residual covariance and residual autocorrelation matrices. The approximate distribution of the test is derived. The test can be written as

$$\tilde{Q}_{EXCO} = n^2 \sum_{k=1}^{m} \frac{w_k}{n-k} tr[\boldsymbol{\Gamma}(k)^{\dagger}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Gamma}(k)\boldsymbol{\Sigma}^{-1}] \tag{6.11}$$

the modified portmanteau test is

$$\tilde{Q}_{EXAU} = n^2 \sum_{k=1}^{m} \frac{w_k}{n-k} \left(\mathrm{vec}(\widehat{\boldsymbol{R}}_k)\right)^{\dagger} (\widehat{\boldsymbol{R}}_0^{-1} \otimes \widehat{\boldsymbol{R}}_0^{-1})\mathrm{vec}(\widehat{\boldsymbol{R}}_k) \tag{6.12}$$

where n is the number of observations, m is the maximum lag taken into account, $\boldsymbol{\Gamma}(k)$ is a covariance matrix at lag k, $\widehat{\boldsymbol{R}}_k$ is an autocorrelation matrix at lag k and w_k is an exponential weight, given by

$$w_k = e^{-\left(\frac{k-1}{m}\right)\ln\ ^2 m}$$

6.2.1 Asymptotic distribution of the new multivariate portmanteau test

Theorem 6.1. Suppose that a vector time series $\{\boldsymbol{z}_t\}$ is generated by a stationary and invertible VARMA(p,q) process with mean zero and constant covariance matrix (see Equation 6.2). Then, the new portmanteau test statistic is asymptotically distributed as

$$\sum_{k=1}^{d^2 m} \lambda_k \chi_{1,k}^2$$

where m is the maximum lag, d is the number of vector components, $\chi_{1,k}^2$ $(k = 1, 2, \dots, d^2 m)$ are independent χ_1^2 random variables and λ_k $(\lambda_1, \lambda_2, \dots, \lambda_{d^2 m})$ are the eigenvalues of $(\boldsymbol{I}_{d^2 m} - \boldsymbol{Q})\boldsymbol{W}$, where $\boldsymbol{W}$ is a $d^2 m \times d^2 m$ diagonal matrix

$$\boldsymbol{W} = \begin{pmatrix} w_1 I_{d^2} & 0 & \cdots & 0 \\ 0 & w_2 I_{d^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_m I_{d^2} \end{pmatrix}$$

where w_k $(k = 1, 2, \cdots, m)$ are weights that satisfy $0 < w_k \leq 1$, and $\boldsymbol{Q}$ is an idempotent matrix, which is define as

$$\boldsymbol{Q} = \boldsymbol{X}(\boldsymbol{X}^{\dagger}\boldsymbol{M}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^{\dagger}\boldsymbol{M}^{-1} \tag{6.13}$$

and $\boldsymbol{X}$ is a $d^2m \times d^2(p+q)$ matrix with elements $\boldsymbol{\Psi}$ and Π defined by

$$\boldsymbol{\Psi}(B) = \boldsymbol{\Phi}(B)^{-1} = \sum_{i=0}^{\infty} \boldsymbol{\Psi}_i B^i$$

and

$$\Pi(B) = \boldsymbol{I} - \sum_{i=0}^{\infty} \Pi_i B^i = \boldsymbol{\Theta}(B)^{-1}$$

and $\boldsymbol{M} = \boldsymbol{I}_m \otimes \boldsymbol{\Gamma}_0 \otimes \boldsymbol{\Gamma}_0$ is a positive-definite symmetric.

The form of the idempotent matrix $\boldsymbol{Q}$ was first derived by Box and Pierce (1970), and has been subsequently used by McLeod (1978), Hosking (1980), and Mahdi and McLeod (2011) in the development of their multivariate portmanteau tests.

Proof of Theorem 6.1:

Let, $\chi^2_{1,k}$ $(k = 1, 2, \dots, d^2m)$ be independent χ^2_1 random variables. By using the idempotent matrix form Hosking (1980) and multiplying the exponential weight with $(\boldsymbol{I}_{d^2m} - \boldsymbol{Q})$, then

$$(\boldsymbol{I}_{d^2m} - \boldsymbol{Q})\boldsymbol{W}$$

the λ_k $(\lambda_1, \lambda_2, \dots, \lambda_{d^2m})$ are the eigenvalues of $(\boldsymbol{I}_{d^2m} - \boldsymbol{Q})\boldsymbol{W}$.

Summing the eigenvalues and applying the tr matrix to the idempotent matrix, it gives

$$\sum_{k=1}^{d^2m} \lambda_k = tr\big((\boldsymbol{I}_{d^2m} - \boldsymbol{Q})\boldsymbol{W}\big)$$
$$= tr(\boldsymbol{W}) - tr(\boldsymbol{Q}) + (1/d^2m)tr(\boldsymbol{QC})$$

where $\boldsymbol{C}$ is a diagonal matrix with elements $c_k = k$, where $k = 0,1, \dots, (d^2m - 1)$, and

$$\sum_{k=1}^{d^2m} \lambda_k^2 = tr\big((\boldsymbol{I}_{d^2m} - \boldsymbol{Q})\boldsymbol{W}(\boldsymbol{I}_{d^2m} - \boldsymbol{Q})\boldsymbol{W}\big)$$
$$= tr(\boldsymbol{W})^2 - tr(\boldsymbol{Q}) + (2/d^2m)tr(\boldsymbol{QC}) - (2/(d^2m^2)^2)tr(\boldsymbol{QC^2})$$
$$+ (1/(d^2m^2))tr(\boldsymbol{QCQC})$$

As $\boldsymbol{Q}$ is the idempotent matrix with rank $d^2(p+q)$, then

$$\sum_{k=1}^{d^2m} \lambda_k = \sum_{k=1}^{d^2m} w_k - d^2(p+q) + (1/d^2m)\sum_{k=2}^{d^2m}(k-1)\, q_{kk}$$

$$\begin{aligned}\sum_{k=1}^{d^2m} \lambda_k^2 = \sum_{k=1}^{d^2m} w_k^2 - d^2(p+q) + (2/d^2m)\sum_{k=2}^{d^2m}(k-1)\, q_{kk} \\ - (2/(d^2m^2)^2)\sum_{k=2}^{d^2m}(k-1)^2\, q_{kk} \\ + (1/d^2m)^2\sum_{i=2}^{d^2m}\sum_{j=2}^{d^2m}(i-1)(j-1)q_{ij}^2\end{aligned}$$

where q_{ij} are the elements of $\boldsymbol{Q}$.

Using Kronecker's lemma (Davidson, 1994), as $d^2m \to \infty$, then

$$(1/d^2m)\sum_{k=2}^{d^2m}(k-1)\, q_{kk} \to 0$$

$$(2/d^2m)\sum_{k=2}^{d^2m}(k-1)\, q_{kk} \to 0$$

$$(2/(d^2m^2)^2)\sum_{k=2}^{d^2m}(i-1)^2\, q_{kk} \to 0$$

and

$$(1/d^2m)^2\sum_{i=2}^{d^2m}\sum_{j=2}^{d^2m}(i-1)(j-1)q_{ij}^2 \to 0$$

Thus,

$$\sum_{k=1}^{d^2m} \lambda_k = \sum_{i=1}^{d^2m} w_i \text{ as } d^2m \to \infty \tag{6.14}$$

$$\sum_{k=1}^{d^2m} \lambda_k^2 = \sum_{i=1}^{d^2m} d^2 w_i^2 - d^2(p+q) \text{ as } d^2m \to \infty \tag{6.15}$$

From the result of Box (1954), as a vector $\boldsymbol{z}_t$ having mean zero and constant covariance matrix, then the asymptotic distribution of portmanteau test statistic in Equation 6.11 distributed as

$$\sum_{k=1}^{d^2m} \lambda_k \chi_{1,k}^2$$

6.2.2 Approximation distribution of the new multivariate portmanteau test

Duchesne and Roy (2004) proposed a test statistic for checking the hypothesis of non-correlation or independence in the Gaussian case. The test statistic is

$$T_n = \frac{\tilde{Q}_H - d^2 M_n(\text{Ҡ})}{\left(2d^2 V_n(\text{Ҡ})\right)^{1/2}}$$

where $\tilde{Q}_H$ is the Hosking multivariate portmanteau test, Ҡ is a Kernel function,

$$M_n(\text{Ҡ}) = \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) \text{Ҡ}^2\left(\frac{k}{P_n}\right),$$

$$V_n(\text{Ҡ}) = \sum_{k=1}^{n-2} \left(1 - \frac{k}{n}\right)\left(1 - \frac{(k+1)}{n}\right) \text{Ҡ}^4\left(\frac{k}{P_n}\right),$$

and P_n is a sequence of truncation values. If $P_n \to \infty$ and $P_n/n \to 0$, then $M_n(\text{Ҡ}) = \int_0^\infty \text{Ҡ}^2$ and $V_n(\text{Ҡ}) = \int_0^\infty \text{Ҡ}^4$. From Equation 6.14 and 6.15 the normalize terms are $\sum_{k=1}^{d^2m} w_k$ and $\sum_{k=1}^{d^2m} w_k^2$, then these two normalize terms could be replaced by $M_n(\text{Ҡ})$ and $V_n(\text{Ҡ})$. Duchesne and Roy (2004) proved that $T_n \to N\,(0,1)$ and $\hat{T}_n \to T_n$ is $O_P(1)$, where $O_P(\,.\,)$ is "order in probability". By using the result of Duchesne and Roy (2004) then the approximation distribution of the new portmanteau test statistic of vector ARMA models can be written as $Q_w \sim \Gamma(\alpha, \beta)$ with shape and scale

$$\alpha = \frac{\left(\sum_{k=1}^{d^2m} w_k\right)^2}{2d^2\left(\sum_{k=1}^{d^2m} w_k^2 - p - q\right)} \tag{6.16}$$

and

$$\beta = \frac{2d^2\left(\sum_{k=1}^{d^2m} w_k^2 - p - q\right)}{\sum_{k=1}^{d^2m} w_k} \tag{6.17}$$

6.3 Monte Carlo experiment with vector time series

6.3.1 Simulation studies

The aim of this simulation study is to examine the effectiveness of the new exponential portmanteau test $\tilde{Q}_{EXCO}$ compared with the portmanteau tests used in previous studies. Specifically, the new test is compared with tests developed by Hosking (1980) $\tilde{Q}_H$, and Mahdi and McLeod (2011) $\tilde{Q}_{VMM}$. The empirical size and the power level of the tests were investigated by conducting simulations studies using the R program.

6.3.2 Empirical size

A Monte-Carlo experiment was conducted with 10,000 replications. The procedure used was the same as that outlined in Section 5.3.3, with the exception of step 5 in which the calculation of the autocorrelation and partial autocorrelation function is replaced by the calculated covariance matrices. The aim was to simulate 100 and 200 observations under the VAR(1) process.
One model was taken from Hosking (1980) ($j = 1$)

<u>Model 1</u>

$$\boldsymbol{\Phi}_1 = \begin{pmatrix} 0.9 & 0.1 \\ -0.6 & 0.4 \end{pmatrix}, \qquad \boldsymbol{\Sigma}_1 = \begin{pmatrix} 1.0 & 0.4 \\ 0.4 & 1.0 \end{pmatrix}$$

and three models were taken from Li and McLeod (1981) ($j = 2, 3, 4$).

<u>Model 2</u>

$$\boldsymbol{\Phi}_2 = \begin{pmatrix} -1.5 & 1.2 \\ -0.9 & 0.5 \end{pmatrix}, \qquad \boldsymbol{\Sigma}_2 = \begin{pmatrix} 1.0 & 0.5 \\ 0.5 & 1.0 \end{pmatrix}$$

Model 3

$$\boldsymbol{\Phi}_3 = \begin{pmatrix} 0.4 & 0.1 \\ -1.0 & 0.5 \end{pmatrix}, \qquad \boldsymbol{\Sigma}_3 = \begin{pmatrix} 1.0 & 0.5 \\ 0.5 & 1.0 \end{pmatrix}$$

Model 4

$$\boldsymbol{\Phi}_4 = \begin{pmatrix} -0.2 & 0.3 \\ -0.6 & 1.1 \end{pmatrix}, \qquad \boldsymbol{\Sigma}_4 = \begin{pmatrix} 1.0 & 0.5 \\ 0.5 & 1.0 \end{pmatrix}$$

All models are a Gaussian bivariate VAR(1) model of the form $\boldsymbol{z}_t = \boldsymbol{\Phi}_j \boldsymbol{z}_{t-1} + \boldsymbol{e}_t,\ j = 1, 2, \ldots, 4$. The coefficients matrices and the covariance matrices are taken from Hosking (1980) $(\boldsymbol{\Phi}_1, \boldsymbol{\Sigma}_1)$, and Li and McLeod (1981) $\{(\boldsymbol{\Phi}_2, \boldsymbol{\Sigma}_2),\ (\boldsymbol{\Phi}_3, \boldsymbol{\Sigma}_3), (\boldsymbol{\Phi}_4, \boldsymbol{\Sigma}_4)\}$

Tables 6.1 and 6.2 show the proportion of $\tilde{Q}_H, \tilde{Q}_{VMM}$ and $\tilde{Q}_{EXCO}$ values that are above the upper 5 percentage point of the appropriate distribution, $\chi^2_{d^2(m-1)}$ for the $\tilde{Q}_H$ and $\tilde{Q}_{VMM}$ tests, and the gamma distribution for the $\tilde{Q}_{EXCO}$ test. The data fitted under the VAR(1) with different coefficients matrices $\boldsymbol{\Phi}$ and covariance matrices $\boldsymbol{\Sigma}$ with n = 100 and 200, and lags of covariance matrices of m = 10, 20 and 30.

	$m = 10$			$m = 20$			$m = 30$		
Model	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$
1	0.0570	**0.0559**	0.0701	**0.0637**	0.0692	0.0823	**0.0706**	0.0968	0.0908
2	0.0522	**0.0518**	0.0544	0.0638	0.0688	**0.0633**	**0.0665**	0.0916	0.0786
3	0.0521	0.0448	**0.0503**	0.0615	0.0659	**0.0606**	**0.0677**	0.0921	0.0737
4	**0.0579**	0.0609	0.0655	**0.0595**	0.0702	0.0768	**0.0712**	0.1036	0.0928

Table 6.1 Empirical size of $\tilde{Q}_H, \tilde{Q}_{VMM}$ and $\tilde{Q}_{EXCO}$ at 5% significance level for fitted VAR(1) models, $n = 100$ and $m = 10, 20$ and 30.

Table 6.1 shows the values of significance level when $\alpha = 0.05$, n = 100 and m = 10, 20 and 30. When m = 10, the value of $\tilde{Q}_H$ test is also closer to 0.05 in one case with model 4, the $\tilde{Q}_{VMM}$ test is closer to 0.05 significance level than any other tests in two cases with models 1 and 2. The value of $\tilde{Q}_{EXCO}$ test is closer to 0.05 in one case with model 3. When m = 20, the value of $\tilde{Q}_{EXCO}$ test is close to 0.05 in two cases with models 2 and 3. The value of $\tilde{Q}_H$ test is also closer to 0.05 in two cases with models 1 and 4. When m = 30, the value of $\tilde{Q}_H$ test is close to 0.05 in four cases with models 1, 2, 3 and 4.

	$m = 10$			$m = 20$			$m = 30$		
Model	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$
1	**0.0509**	0.0474	0.0688	0.0592	**0.0533**	0.0754	**0.0599**	0.0646	0.0876
2	0.0532	0.0407	**0.0508**	0.0593	**0.052**	0.0633	**0.0606**	0.0611	0.0708
3	**0.0530**	0.0386	0.0453	0.0548	**0.0490**	0.0573	**0.0590**	0.0619	0.0733
4	0.0549	**0.0519**	0.0706	0.0562	**0.056**	0.0753	**0.0647**	0.0653	0.0900

Table 6.2 Empirical size of $\tilde{Q}_H$, $\tilde{Q}_{VMM}$ and $\tilde{Q}_{EXCO}$ at 5% significance level for fitted VAR(1) models, $n = 200$ and $m = 10, 20$ and 30.

Table 6.2 shows the values of significance level when $\alpha = 0.05$, $n = 200$ and $m = 10$, 20 and 30. When $m = 10$, the value of $\tilde{Q}_H$ and $\tilde{Q}_{EXCO}$ tests are closer to 0.05 significance level in two cases with models 1, 3 and 2, 4 respectively. When $m = 20$, the value of $\tilde{Q}_{VMM}$ test is closer to 0.05 significance level than any other test in all cases. When $m = 30$, the value of $\tilde{Q}_H$ test is closer to 0.05 in all cases.

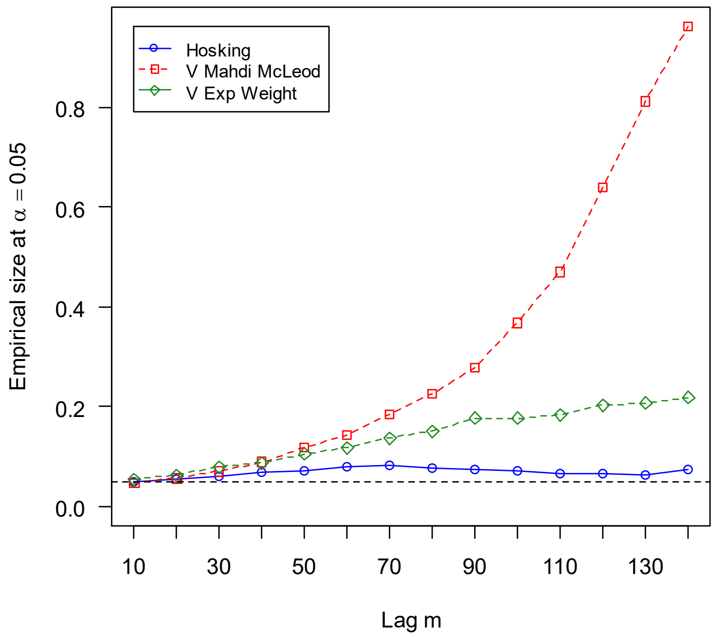

Figure 6.1 Empirical size for maximum lags of $\tilde{Q}_H$, $\tilde{Q}_{VMM}$ and $\tilde{Q}_{EXCO}$ tests at 5% significance level fitted VAR(1), and data generated by model 2, series of length $n = 150$.

Figure 6.1 shows the empirical size for maximum lags of the $\tilde{Q}_{\mathrm{H}}$, $\tilde{Q}_{\mathrm{VMM}}$ and $\tilde{Q}_{\mathrm{EXCO}}$ tests at 5% significance level fitted under a VAR(1) model, with data generated by Model 2 with n = 150. The $\tilde{Q}_{\mathrm{H}}$ test is stable as the lag increases, and the $\tilde{Q}_{\mathrm{VMM}}$ test rapidly increases as the lag increases, whereas the $\tilde{Q}_{\mathrm{EXCO}}$ test slowly increases as the lag increases. This means the $\tilde{Q}_{\mathrm{H}}$ test is more stable with respect to lag length than other tests as the lag increases.

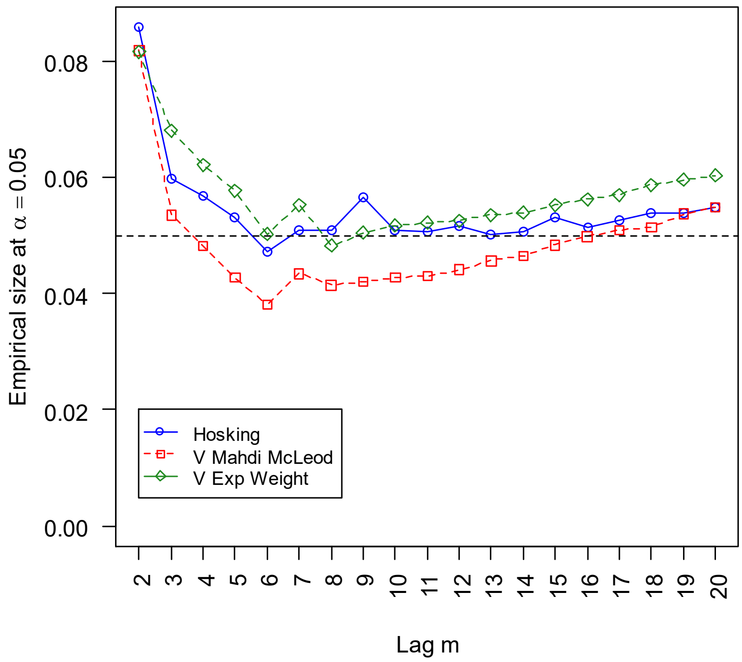

Figure 6.2 Empirical size for lags from 2 to 20 of the $\tilde{Q}_{H}$, $\tilde{Q}_{VMM}$ and $\tilde{Q}_{EXCO}$ tests at 5% significance level fitted VAR(1), and data generated by model 2, series of length n = 150.

Figure 6.2 shows the empirical size for lags from 2 to 20 of the $\tilde{Q}_{\mathrm{H}}$, $\tilde{Q}_{\mathrm{VMM}}$ and $\tilde{Q}_{\mathrm{EXCO}}$ tests at 5% significance level fitted under a VAR(1) model, and data generated by Model 2 with n = 150. The $\tilde{Q}_{EXCO}$ test is close to 0.05 at lags 6, 8, 9 and 10, then slowly increases as lag size increases, whereas the $\tilde{Q}_{\mathrm{VMM}}$ test is close to 0.05 at lag 4, then slowly decreases as the lag increases up to lag 6. From this point it becomes close to the significance level, thereafter it slowly increases as lag increase. The $\tilde{Q}_{EXCO}$ test decreases as the lag increases and becomes close to 0.05 at lag 6. Then the $\tilde{Q}_{EXCO}$ test increases at lag 7, after that it becomes very close 0.05 at lag 8, thereafter it slowly increases as the lag increases. The $\tilde{Q}_H$ test is close to 0.05 at lags $5 \leq m \leq 140$.

Figures 6.1 and 6.2 together show the effect of the choice of large lag size on the empirical size of the three tests. The $\tilde{Q}_{\mathrm{H}}$ test is effective when $5 \leq m \leq 140$, while the $\tilde{Q}_{\mathrm{VMM}}$ test is effective when $6 \leq m \leq 25$, and the $\tilde{Q}_{\mathrm{EXCO}}$ test when $5 \leq m \leq 25$.

6.3.3 Power studies

The aim of the power studies is to show which tests are the most powerful. The data were generated by a number of different VARMA(2,2) processes.

$$\boldsymbol{z}_t = \boldsymbol{\Phi}_1 \boldsymbol{z}_{t-1} + \boldsymbol{\Phi}_2 \boldsymbol{z}_{t-2} + \boldsymbol{e}_t - \boldsymbol{\Theta}_1 \boldsymbol{e}_{t-1} - \boldsymbol{\Theta}_2 \boldsymbol{e}_{t-2}$$

For each alternative, 10,000 replications of 100 and 200 observations were generated. The procedure used was the same as that outlined in Section 5.3.4, with the exception of step 5 in which the calculation of the autocorrelation and partial autocorrelation function is replaced by the calculated covariance matrices. For each test the power was computed with lags $m = 10, 20$ and 30. The residuals of the data were obtained.

The VARMA models below were used to compare the new method against the tests of Mahdi and McLeod and Hoskings. These models were selected from the literature by Mahdi and McLeod (2011) to give a representative sample of models on which to test their methods.

These Models have been fitted under the VAR(1) model to make comparisons among the three tests, the models are:

<u>Model 5</u>

Lütkepohl, (2005, p.17).

$$\begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} - \begin{pmatrix} 0.5 & 0.1 \\ 0.4 & 0.5 \end{pmatrix} \begin{pmatrix} z_{1,t-1} \\ z_{2,t-1} \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0.3 & 0 \end{pmatrix} \begin{pmatrix} z_{1,t-2} \\ z_{2,t-2} \end{pmatrix} = \begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix}$$
$$\boldsymbol{\Sigma} = \begin{pmatrix} 1.00 & 0.71 \\ 0.71 & 1.00 \end{pmatrix}$$

Model 6

Brockwell and Davis (1991, p. 428).

$$\begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} - \begin{pmatrix} 0.7 & 0.0 \\ 0.0 & 0.6 \end{pmatrix} \begin{pmatrix} z_{1,t-1} \\ z_{2,t-1} \end{pmatrix} = \begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix} + \begin{pmatrix} 0.5 & 0.6 \\ -0.7 & 0.8 \end{pmatrix} \begin{pmatrix} e_{1,t-1} \\ e_{2,t-1} \end{pmatrix}$$
$$\boldsymbol{\Sigma} = \begin{pmatrix} 1.00 & 0.71 \\ 0.71 & 2.00 \end{pmatrix}$$

Model 7

Reinsel (1993, p. 71).

$$\begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} - \begin{pmatrix} 1.2 & -0.5 \\ 0.6 & 0.3 \end{pmatrix} \begin{pmatrix} z_{1,t-1} \\ z_{2,t-1} \end{pmatrix} = \begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix} - \begin{pmatrix} -0.6 & 0.3 \\ 0.3 & 0.6 \end{pmatrix} \begin{pmatrix} e_{1,t-1} \\ e_{2,t-1} \end{pmatrix}$$
$$\boldsymbol{\Sigma} = \begin{pmatrix} 1.00 & 0.50 \\ 0.50 & 1.25 \end{pmatrix}$$

Model 8

Tsay (2005, 2nd, p. 371).

$$\begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} - \begin{pmatrix} 0.8 & -2.0 \\ 0.0 & 0.0 \end{pmatrix} \begin{pmatrix} z_{1,t-1} \\ z_{2,t-1} \end{pmatrix} = \begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix} - \begin{pmatrix} -0.5 & 0.0 \\ 0.0 & 0.0 \end{pmatrix} \begin{pmatrix} e_{1,t-1} \\ e_{2,t-1} \end{pmatrix}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1.00 & 0.71 \\ 0.71 & 1.00 \end{pmatrix}$$

Model 9

Reinsel (1993, p. 30).

$$\begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} = \begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix} - \begin{pmatrix} 0.8 & 0.7 \\ -0.4 & 0.6 \end{pmatrix} \begin{pmatrix} e_{1,t-1} \\ e_{2,t-1} \end{pmatrix}$$
$$\boldsymbol{\Sigma} = \begin{pmatrix} 4.00 & 1.00 \\ 1.00 & 2.00 \end{pmatrix}$$

Model 10

Tsay (2005, 2nd, p. 350).

$$\begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} = \begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix} - \begin{pmatrix} 0.2 & 0.3 \\ -0.6 & 1.1 \end{pmatrix} \begin{pmatrix} e_{1,t-1} \\ e_{2,t-1} \end{pmatrix}$$
$$\boldsymbol{\Sigma} = \begin{pmatrix} 2.00 & 1.00 \\ 1.00 & 1.00 \end{pmatrix}$$

Model 11

Lütkepohl, (2005, p. 445).

$$\begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} - \begin{pmatrix} 0.5 & 0.1 \\ 0.4 & 0.5 \end{pmatrix} \begin{pmatrix} z_{1,t-1} \\ z_{2,t-1} \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0.25 & 0 \end{pmatrix} \begin{pmatrix} z_{1,t-2} \\ z_{2,t-2} \end{pmatrix} = \begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix} + \begin{pmatrix} 0.6 & 0.2 \\ 0.0 & 0.3 \end{pmatrix} \begin{pmatrix} e_{1,t-1} \\ e_{2,t-1} \end{pmatrix}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1.00 & 0.30 \\ 0.30 & 1.00 \end{pmatrix}$$

Model 12

Reinsel et al. (1992, p. 141).

$$\begin{pmatrix} z_{1,t} \\ z_{2,t} \\ z_{3,t} \end{pmatrix} - \begin{pmatrix} 0.4 & 0.3 & -0.6 \\ 0.0 & 0.8 & 0.4 \\ 0.3 & 0.0 & 0.0 \end{pmatrix} \begin{pmatrix} z_{1,t-1} \\ z_{2,t-1} \\ z_{3,t-1} \end{pmatrix} = \begin{pmatrix} e_{1,t} \\ e_{2,t} \\ e_{3,t} \end{pmatrix} - \begin{pmatrix} 0.7 & 0.0 & 0.0 \\ 0.1 & 0.2 & 0.0 \\ -0.4 & 0.5 & -0.1 \end{pmatrix} \begin{pmatrix} e_{1,t-1} \\ e_{2,t-1} \\ e_{3,t-1} \end{pmatrix}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1.00 & 0.50 & 0.40 \\ 0.50 & 1.00 & 0.70 \\ 0.40 & 0.70 & 1.00 \end{pmatrix}$$

Tables 6.3 and 6.4 show the power level of $\tilde{Q}_H$, $\tilde{Q}_{VMM}$ and $\tilde{Q}_{EXCO}$ values that are above the upper 5 percentage point of $\chi^2_{d^2(m-1)}$ distribution and gamma distribution. The data was fitted under a VAR(1) model with different coefficients matrices and covariance matrices with $n = 100$ and 200, and lags of covariance matrices of $m = 10, 20$ and 30.

	$m = 10$			$m = 20$			$m = 30$		
Model	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$
5	0.3522	0.5569	**0.7001**	0.2518	0.4692	**0.6554**	0.2088	0.4428	**0.6309**
6	0.7827	**0.992**	0.9814	0.6528	**0.9748**	0.9709	0.6062	0.9485	**0.9625**
7	0.9921	**1**	**1**	0.8632	0.9998	**0.9999**	0.6952	0.9996	**1**
8	0.5363	0.8845	**0.933**	0.4133	0.7743	**0.8769**	0.3696	0.7179	**0.8433**
9	0.8892	**1**	0.9999	0.6629	**0.9998**	0.9994	0.5530	**0.9993**	0.9981
10	0.6783	**0.9972**	0.9742	0.5118	**0.9877**	0.9495	0.4549	**0.9703**	0.9306
11	0.3330	0.5725	**0.6491**	0.224	0.4725	**0.6204**	0.2073	0.4336	**0.6011**
12	0.768	**0.9975**	0.9896	0.5921	**0.9961**	0.9779	0.5095	**0.9958**	0.9657
Average	0.6664	0.8750	**0.9034**	0.5214	0.8342	**0.8812**	0.4505	0.8134	**0.8665**

Table 6.3 Power level of $\tilde{Q}_H$, $\tilde{Q}_{VMM}$ and $\tilde{Q}_{EXCO}$ at 5% significance level for fitted VAR(1) models, $n = 100$ and $m = 10, 20$ and 30.

Table 6.3 shows that the $\tilde{Q}_{VMM}$ test is more powerful than other portmanteau tests in 5 cases, and the $\tilde{Q}_{EXCO}$ test is more powerful than other portmanteau tests in 4 cases, when $m = 10$. In addition, when $m = 20$, the $\tilde{Q}_{VMM}$ test is the most powerful in 4 cases, and the $\tilde{Q}_{EXCO}$ test is more powerful than other portmanteau tests in 4 cases. When $m = 30$, the $\tilde{Q}_{EXCO}$ test is more powerful than other portmanteau tests in 5 cases, and the $\tilde{Q}_{VMM}$ test is more powerful in 3 cases. It means that the $\tilde{Q}_{EXCO}$ test is more powerful when $m = 30$ and $n = 100$.

	$m = 10$			$m = 20$			$m = 30$		
Model	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$
5	0.7204	0.8820	**0.9647**	0.5198	0.8032	**0.9457**	0.4213	0.7358	**0.9369**
6	0.9993	**1**	**1**	0.9798	**1**	**1**	0.9545	**1**	**1**
7	**1**	**1**	**1**	**1**	**1**	**1**	0.9992	**1**	**1**
8	0.9192	0.9987	**0.9998**	0.7709	0.9922	**0.9988**	0.6790	0.9791	**0.9982**
9	**1**	**1**	**1**	0.9939	**1**	**1**	0.9586	**1**	**1**
10	0.9916	**1**	**1**	0.9162	**1**	**1**	0.8302	**1**	**1**
11	0.7373	0.9050	**0.9562**	0.5119	0.8381	**0.9474**	0.4144	0.7765	**0.9363**
12	0.9967	**1**	**1**	0.9499	**1**	**1**	0.7586	**1**	**1**
Average	0.9205	0.9732	**0.9900**	0.8303	0.9541	**0.9864**	0.7519	0.9364	**0.9839**

Table 6.4 Power level of $\tilde{Q}_H$, $\tilde{Q}_{VMM}$ and $\tilde{Q}_{EXCO}$ at 5% significance level for fitted VAR(1) models, $n = 200$ and $m = 10, 20$ and 30.

Table 6.4 shows that the $\tilde{Q}_{EXCO}$ test is more powerful in all cases than other portmanteau tests, when $m = 10, 20$ and 30. The $\tilde{Q}_{VMM}$ test is more powerful than other tests in 5 cases, when $m = 10, 20$ and 30. The $\tilde{Q}_H$ test is more powerful in 2 cases when $m = 10$, and 1 case when $m = 20$. This means that the $\tilde{Q}_{EXCO}$ test is more powerful than other tests when $m = 30$ and $n = 200$.

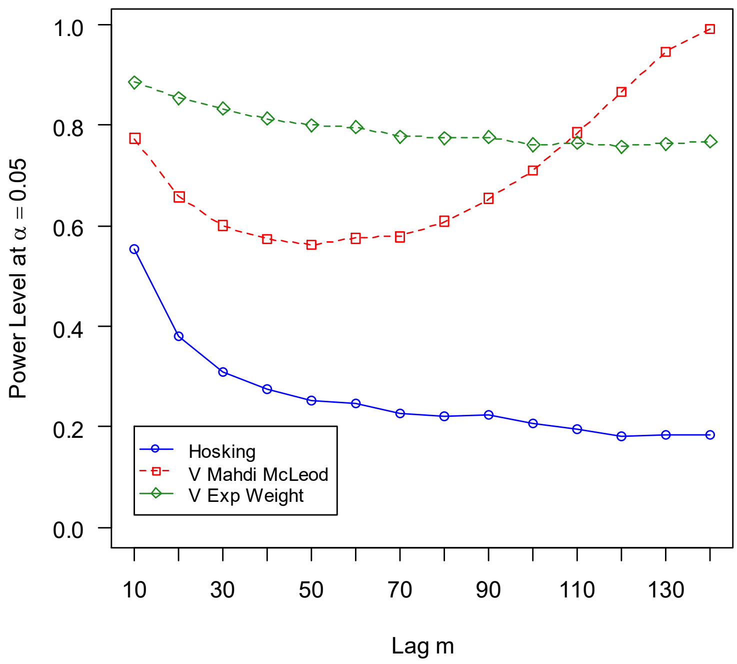

Figure 6.3 Power level for maximum lags of $\tilde{Q}_H$, $\tilde{Q}_{VMM}$ and $\tilde{Q}_{EXCO}$ at 5% significance level fitted VAR(1), and data generated by model 5, series of length $n = 150$.

Figure 6.3 shows the power level for maximum lags of the $\tilde{Q}_{\mathrm{H}}$, $\tilde{Q}_{\mathrm{VMM}}$ and $\tilde{Q}_{\mathrm{EXCO}}$ tests at 5% significance level fitted under a VAR(1) model, with data generated by model 5 with $n = 150$. The $\tilde{Q}_{\mathrm{H}}$ test decreases as the lag increases, and the $\tilde{Q}_{\mathrm{VMM}}$ test decreases till lag 50, then it increases as the lag increases. The $\tilde{Q}_{\mathrm{EXCO}}$ test decreases only slightly as the lag increases. This means the $\tilde{Q}_{\mathrm{EXCO}}$ test is more stable and powerful than other tests as the lag increases.

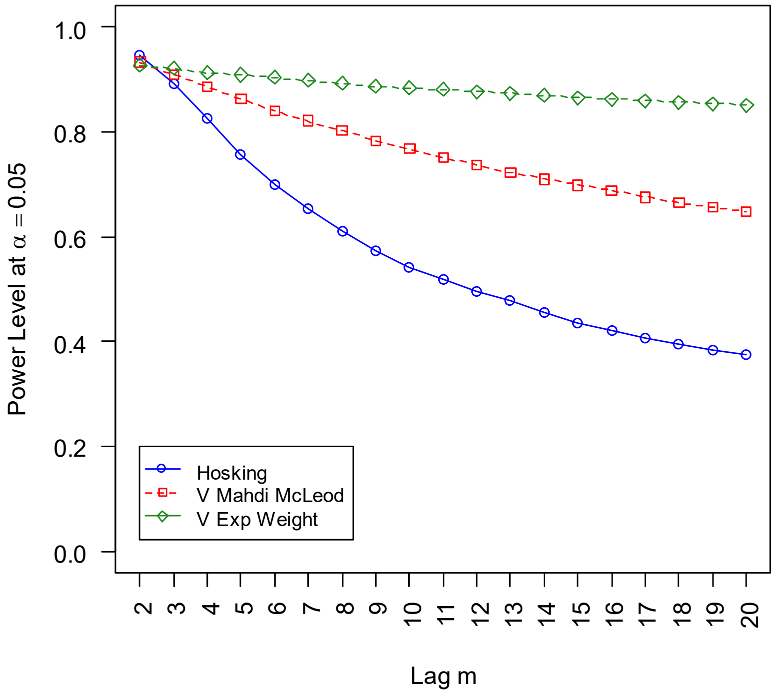

Figure 6.4 Power level for lags from 2 to 20 of $\tilde{Q}_{\mathrm{H}}$, $\tilde{Q}_{\mathrm{VMM}}$ and $\tilde{Q}_{\mathrm{EXCO}}$ at 5% significance level fitted VAR(1), and data generated by model 5, series of length $n = 150$..

Figure 6.4 shows the power level for lags from 2 to 20 of the $\tilde{Q}_{\mathrm{H}}$, $\tilde{Q}_{\mathrm{VMM}}$ and $\tilde{Q}_{\mathrm{EXCO}}$ tests at 5% significance level fitted under a VAR(1) model, with data generated by model 5 with $n = 150$. The $\tilde{Q}_{\mathrm{H}}$ and $\tilde{Q}_{\mathrm{VMM}}$ tests decrease as the lag increases, and the $\tilde{Q}_{\mathrm{EXCO}}$ test slowly decreases as the lag increases. As can be seen from Figure 6.4 the $\tilde{Q}_{\mathrm{EXCO}}$ test is more stable and powerful than other tests as the lag increases.

Figures 6.3 and 6.4 together show the effect of the choice of large lag size on the power level of all three tests. For a power level of 75%, the $\tilde{Q}_{\mathrm{H}}$ test is effective when $m \leq 5$, while the $\tilde{Q}_{\mathrm{VMM}}$ test is effective when $m \leq 11$, and the $\tilde{Q}_{\mathrm{EXCO}}$ test when $m \leq 140$.

Next, a VMA(1) model is fitted to the data to compare the portmanteau tests $\tilde{Q}_{H}$, $\tilde{Q}_{VMM}$ and $\tilde{Q}_{EXCO}$. The data are generated by a number of alternative VARMA(2,2) processes, which have been selected from the literature (Lütkepohl, (2005).

$$\boldsymbol{z}_t = \boldsymbol{\Phi}_1 \boldsymbol{z}_{t-1} + \boldsymbol{\Phi}_2 \boldsymbol{z}_{t-2} + \boldsymbol{e}_t - \boldsymbol{\Theta}_1 \boldsymbol{e}_{t-1} - \boldsymbol{\Theta}_2 \boldsymbol{e}_{t-2}$$

For each alternative, 10,000 replications of 100 and 200 observations were generated. The coefficients matrices were estimated by using the conditional maximum likelihood function. For each test the power was computed with lags $m = 10, 20$ and 30. The residual of the data was obtained.

There is no published empirical research on fitting a VAM(1) model to non-VAM(1) processes. Consequently, the VARMA processes below were selected from the literature to give a representative sample of processes on which to test the new method. As a comparison, the methods of Mahdi and McLeod, and Hoskings were also used to evaluated the fitted models

Model 13

Lütkepohl, (2005, p.17).

$$\begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} - \begin{pmatrix} 0.5 & 0.1 \\ 0.4 & 0.5 \end{pmatrix} \begin{pmatrix} z_{1,t-1} \\ z_{2,t-1} \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0.3 & 0 \end{pmatrix} \begin{pmatrix} z_{1,t-2} \\ z_{2,t-2} \end{pmatrix} = \begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1.00 & 0.71 \\ 0.71 & 1.00 \end{bmatrix}$$

Note that this is the same as Model 5 as used in the power studies experiments in Section 6.3.3.

Model 14

Brockwell and Davis (1991, p. 428).

$$\begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} - \begin{pmatrix} 0.7 & 0.0 \\ 0.0 & 0.6 \end{pmatrix} \begin{pmatrix} z_{1,t-1} \\ z_{2,t-1} \end{pmatrix} = \begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix} + \begin{pmatrix} 0.5 & 0.6 \\ -0.7 & 0.8 \end{pmatrix} \begin{pmatrix} e_{1,t-1} \\ e_{2,t-1} \end{pmatrix}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1.00 & 0.71 \\ 0.71 & 2.00 \end{pmatrix}$$

Note that this is the same as Model 6 as used in the power studies experiments in Section 6.3.3.

Model 15

Huong (2013, p. 113)

$$\begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} - \begin{pmatrix} 0.4588 & 0.4390 \\ -0.0299 & 0.5162 \end{pmatrix} \begin{pmatrix} z_{1,t-1} \\ z_{2,t-1} \end{pmatrix} = \begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix} - \begin{pmatrix} -0.0589 & -0.3047 \\ 0.0093 & -0.1212 \end{pmatrix} \begin{pmatrix} e_{1,t-1} \\ e_{2,t-1} \end{pmatrix}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1.00 & 0.20 \\ 0.20 & 1.00 \end{pmatrix}$$

<u>Model 16</u>

Huong (2013 p. 112)

$$\begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} - \begin{pmatrix} 0.5603 & 0.5361 \\ -0.0366 & 0.6303 \end{pmatrix} \begin{pmatrix} z_{1,t-1} \\ z_{2,t-1} \end{pmatrix} = \begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1.00 & 0.20 \\ 0.20 & 1.00 \end{pmatrix}$$

<u>Model 17</u>

Lütkepohl, (2005, p. 445).

$$\begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} - \begin{pmatrix} 0.5 & 0.1 \\ 0.4 & 0.5 \end{pmatrix} \begin{pmatrix} z_{1,t-1} \\ z_{2,t-1} \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0.25 & 0 \end{pmatrix} \begin{pmatrix} z_{1,t-2} \\ z_{2,t-2} \end{pmatrix} = \begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix} + \begin{pmatrix} 0.6 & 0.2 \\ 0.0 & 0.3 \end{pmatrix} \begin{pmatrix} e_{1,t-1} \\ e_{2,t-1} \end{pmatrix}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1.00 & 0.30 \\ 0.30 & 1.00 \end{pmatrix}$$

Note that this is the same as Model 11 as used in the power studies experiments in Section 6.3.3.

Tables 6.5 and 6.6 show the power level of $\tilde{Q}_H, \tilde{Q}_{VMM}$ and $\tilde{Q}_{EXCO}$ values that are above the upper 5 percentage point of $\chi^2_{d^2(m-1)}$ distribution and gamma distribution. The data were fitted under the VMA(1) model with different coefficients matrices and covariance matrices with $n = 100$ and 200, and with lags of covariance matrices of $m = 10, 20$ and 30.

	$m = 10$			$m = 20$			$m = 30$		
Model	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$
13	0.9999	0.9999	**1**	0.9898	0.9996	**1**	0.9722	0.9983	**1**
14	0.9228	0.9887	**0.9918**	0.8009	0.9775	**0.9897**	0.7314	0.9520	**0.9859**
15	0.3962	0.4290	**0.5781**	0.3144	0.3961	**0.5744**	0.3064	0.3923	**0.5748**
16	0.9766	0.9899	**0.9972**	0.9391	0.9791	**0.9967**	0.9001	0.9672	**0.9959**
17	0.5585	0.7317	**0.8371**	0.3642	0.6673	**0.8242**	0.2918	0.6093	**0.8056**
Average	0.7708	0.8278	**0.8808**	0.6816	0.8039	**0.877**	0.6403	0.7838	**0.8724**

Table 6.5 Power level of $\tilde{Q}_H, \tilde{Q}_{VMM}$ and $\tilde{Q}_{EXCO}$ at 5% significance level for fitted VMA(1) models, $n = 100$ and $m = 10, 20$ and 30.

Table 6.5 shows that the $\tilde{Q}_{EXCO}$ test is more powerful than other portmanteau tests in all cases, and the $\tilde{Q}_{VMM}$ test is also more powerful than the $\tilde{Q}_H$ test in all cases, when $n = 100$ and $m = 10, 20$ and 30.

	$m = 10$			$m = 20$			$m = 30$		
Model	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$	$\tilde{Q}_H$	$\tilde{Q}_{VMM}$	$\tilde{Q}_{EXCO}$
13	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
14	**1**	**1**	**1**	0.9992	**1**	**1**	0.9924	**1**	**1**
15	0.7680	0.8384	**0.9300**	0.6486	0.7740	**0.9286**	0.5784	0.7124	**0.9180**
16	**1**	**1**	**1**	0.9996	**1**	**1**	0.9994	**1**	**1**
17	0.9378	0.9800	**0.9934**	0.7874	0.9538	**0.9928**	0.6692	0.9212	**0.9890**
Average	0.9411	0.9636	**0.9846**	0.8869	0.9455	**0.9842**	0.8478	0.9267	**0.981**

Table 6.6 Power level of $\tilde{Q}_H, \tilde{Q}_{VMM}$ and $\tilde{Q}_{EXCO}$ at 5% significance level for fitted VMA(1) models, $n = 200$ and $m = 10, 20$ and 30.

Table 6.6 shows that the $\tilde{Q}_{EXCO}$ test is more powerful than other portmanteau tests in all cases, when $m = 10, 20$ and 30. The $\tilde{Q}_{VMM}$ and $\tilde{Q}_H$ tests are more powerful in three cases, when $m = 10$. When $m = 20$ and 30, the $\tilde{Q}_{VMM}$ test is more powerful in three cases, and the $\tilde{Q}_H$ is more powerful in one case.

6.4 Variability of the new multivariate portmanteau test

To explore the variability of the new multivariate portmanteau test statistics a Monte-Carlo experiment was conducted with 1000 replications of the experiment to determine the empirical size carried out in Section 6.3.2. As this experiment involved 1000 replications of a simulations involving 10,000 replications this is only conducted for Model 3. The aim was to calculate the mean and standard deviation for the test statistics $\tilde{Q}_H$ and $\tilde{Q}_{EXCO}$, when n = 100 observations under an VAR(1) process by using Model 3 and maximum lags m = 10. The test statistics $\tilde{Q}_H$ and $\tilde{Q}_{EXCO}$ were calculated by using the steps of the previous Monte-Carlo experiment.

	m = 10	
Model 3	$\tilde{Q}_H$	$\tilde{Q}_{EXCO}$
Mean	0.050165	0.054235
Standard deviation	0.002192	0.002337

Table 6.7 The mean and standard deviation for the test statistics $\tilde{Q}_H$ and $\tilde{Q}_{EXCO}$, data fitted under an VAR(1) and generated by model 3, with n = 100 and maximum lags m = 10.

As can be seen from Table 6.7 the results of this experiment are consistent with those from Table 6.1. The standard deviations for the new test is less than 5 percent of the associate mean.

A second experiment was conducted to examine the variability of the new test when applied to the power study in Section 6.3.3. Again 1000 replications of the original experiment were undertaken for a single choice of Model 8.

	m = 10	
Model 8	$\tilde{Q}_H$	$\tilde{Q}_{EXCO}$
Mean	0.470114	0.9331
Standard deviation	0.004848	0.002644

Table 6.8 The mean and standard deviation for the test statistics $\tilde{Q}_H$ and $\tilde{Q}_{EXCO}$, data fitted under an VAR(1) and generated by model 8, with n = 100 and maximum lags m = 10.

As can be seen in Table 6.8 the standard deviations of the new multivariate test is less than 1 percent of the associate mean for the new test.

6.5 Summary

The simulation of empirical size (see Tables 6.1 and 6.2) shows that none of the $\tilde{Q}_H$, $\tilde{Q}_{VMM}$ and $\tilde{Q}_{EXCO}$ tests are better than the others in all cases. The experiments have been conducted with 10,000 replications, and data generated by a range of VAR(1) processes, then fitted under the VAR(1) model.

The results of the simulation studies of power level (see Tables 6.3, 6.4, 6.5 and 6.6) show that the new portmanteau $\tilde{Q}_{EXCO}$ test is more powerful than other tests. The experiments have been conducted with 10,000 replications, and data generated by a range of VARMA(2,2) processes, then fitted under a VAR(1) model and a VMA(1) model.

The simulation of empirical size and power level (see Figures 6.1, 6.2, 6.3 and 6.4) show that the $\tilde{Q}_{EXCO}$, $\tilde{Q}_{VMM}$ $\tilde{Q}_H$ tests are close to 0.05 at lags 6, 3, and 7 respectively in the empirical size, whereas all three tests are the most powerful at lag 2 in the power level. Furthermore, the $\tilde{Q}_{EXCO}$ test slowly increases, while the $\tilde{Q}_{VMM}$ test rapidly increases, and the $\tilde{Q}_H$ test is stable as lag increases in the empirical size. In comparison, the power level of the $\tilde{Q}_{EXCO}$ is the most stable and most powerful as lag increases. The $\tilde{Q}_{VMM}$ test starts to increase at lag 70. Overall, the $\tilde{Q}_{EXCO}$ test is better than other tests for empirical size when the lag is small, and better for power level for all lags.

Chapter 7 - Conclusion

The thesis has explored how the length of data of a univariate time series can influence its apparent stationarity as measured by two standard tests, namely, the Dickey-Fuller test and the Augmented Dickey-Fuller test.

The research has examined the effectiveness of two new exponential portmanteau tests, for univariate time series. These new portmanteau test statistics developed in Chapter 5 are

$$\tilde{Q}_{EXLB} = n(n+2)\sum_{k=1}^{m} w_k \frac{\hat{\rho}_k^2}{n-k}$$

and

$$\tilde{Q}_{EXM} = n(n+2)\sum_{k=1}^{m} w_k \frac{\hat{\phi}_{kk}^2}{n-k}$$

where $\hat{\rho}_k^2$ is the autocorrelation and $\hat{\phi}_{kk}^2$ is the partial autocorrelation at lag k, and w_k is the exponential weight.

Next, the thesis explored the effectiveness of a new exponential portmanteau test for multivariate time series. The new portmanteau test statistics of multivariate time series that is given in Chapter 6 is

$$\tilde{Q}_{EXCO} = n^2 \sum_{k=1}^{m} \frac{w_k}{n-k} tr[\boldsymbol{\Gamma}(k)^{\dagger}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Gamma}(k)\boldsymbol{\Sigma}^{-1}]$$

or equivalently as

$$\tilde{Q}_{EXAU} = n^2 \sum_{k=1}^{m} \frac{w_k}{n-k}\left(\mathrm{vec}(\widehat{\boldsymbol{R}}_k)\right)^{\dagger}(\widehat{\boldsymbol{R}}_0^{-1} \otimes \widehat{\boldsymbol{R}}_0^{-1})\mathrm{vec}(\widehat{\boldsymbol{R}}_k)$$

where n is the number of observations, m is the maximum lag taken into account, $\boldsymbol{\Gamma}(k)$ is a covariance matrix at lag k, $\widehat{\boldsymbol{R}}_k$ is an autocorrelation matrix at lag k and w_k is an exponential weight.

7.1 The length of a time series

Chapter 4 examined how the length of data in a time series affects the identification of its stationarity as identified by the standard tests (the DF test, the DF drift test and the DF trend test, the ADF test, the ADF drift test and the ADF trend test).

A time series was generated from an AR(1) process with positive values of parameters with different lengths of series, namely, n = 25, 50, 75, 100, 250, 500, 750 and 1000 observations. This was then tested by the DF test, the DF drift test, the DF trend test, the ADF test, the ADF drift test and the ADF trend test. As is evident in Figures 4.1, 4.2, 4.3, 4.4, 4.5 and 4.6, the number of series identified as being non-stationary increases as the value of the AR(1) process parameter ϕ_1 increases, and the number identified as being non-stationary decreases as the length of data increases. The researcher proves that the length of data has a strong impact on the stationarity in time series when data are generated under an AR(1) process with positive values of parameter ϕ_1.

For an AR(1) process generated using negative values of parameter ϕ_1 and examined by the ADF test, the ADF drift test, and the ADF trend test, the number of series identified as non-stationary does not depend on the parameter ϕ_1, instead it depends only on the number of data points in the time series. As the number of data points increases the number of time series identified as non-stationary decreases, see Figures 4.7, 4.8, and 4.9. So, for negative values of parameter ϕ_1 the stationarity of the time series depends only on the length of data.

The minimum number of data points required to ensure that the ADF, the ADF drift and the ADF trend tests correctly identify the time series as being stationary, when data are generated from an AR(1) process using negative values of ϕ_1 (using a 5% cut-off), is given in Table 4.3.

A time series was generated from an AR(2) process with positive values of parameters ϕ_1 and ϕ_2, with different lengths of series, namely, n = 25, 50, 75, 100, 250, 500, 750 and 1000 observations. This was then tested by, the DF test, the DF drift test the DF trend test, the ADF test, the ADF drift test and the ADF trend test. As is evident in Figures 4.10 and 4.11, the number of series identified as being non-stationary increases as the value of the

AR(2) process parameters ϕ_1 and ϕ_2 increase, and the number identified as being non-stationary decreases as the length of data increases. The researcher proves that the length of data has a strong impact on the ADF tests' ability to correct identify stationarity in time series when data are generated under an AR(2) process.

7.2 New univariate portmanteau test

The aim of the simulation studies in Chapter 5 was to compare the effectiveness of the new autocorrelation and partial autocorrelation portmanteau tests in relation to those portmanteau tests proposed in previous studies. Monte-Carlo experiments were conducted to examine the empirical size. The empirical size simulations (Section 5.3.3) of different lags, based on a 5% significance level, show that the $\tilde{Q}_{EXM}$ test is not affected by the length of lag used, however, other portmanteau tests are affected. For example, the empirical size simulations based on a 5% significance level show that the values of the $\tilde{Q}_{LB}$, $\tilde{Q}_{FGLB}$, $\tilde{Q}_{GFD}$ and $\tilde{Q}_{EXLB}$ tests increase to 5% significance level as the lag increases and the values of the $\tilde{Q}_M$, $\tilde{Q}_{MM}$ and $\tilde{Q}_{GFK}$ tests decrease to 5% significance level when the lag increases, see Figures 5.1 and 5.2.

In Section 5.3.4 the power level simulation study identified the most powerful portmanteau tests. Monte-Carlo experiments were conducted for 12 different ARMA(2,2) processes. These models were based on those from Monti (1994). An AR(1) model and a MA(1) model were fitted to the data generated from ARMA(2,2) processes. For each test the power was computed with lags $m = 10, 20$ and 30. The residual of the fitted model was obtained. The power level simulations show that the new exponential portmanteau $\tilde{Q}_{EXM}$ test is more powerful than all previous tests used for this purpose, see Tables 5.4, 5.5, 5.6, 5.7, 5.8 and 5.9. The power level simulations for different lags at the 5% significance show that the $\tilde{Q}_{LB}$, $\tilde{Q}_M$, $\tilde{Q}_{MM}$, $\tilde{Q}_{FGLB}$, $\tilde{Q}_{GFK}$, $\tilde{Q}_{GFD}$ and $\tilde{Q}_{EXLB}$ tests decrease as the lag increases, while the $\tilde{Q}_{EXM}$ test slowly decreases as the lag increases, see Figures 5.3 and 5.4.

Following the methodology of the simulation study conducted by Gallagher and Fisher (2015), the data were generated from a subset of the ARMA(2,2) process and fitted under an ARMA(1,1) model. The new $\tilde{Q}_{EXM}$ test proved to be more powerful than all other tests proposed in previous studies. The power level simulations for different lags at the 5%

significance show that the $\tilde{Q}_{EXM}$ test is not affected as the lag increases, while the $\tilde{Q}_{LB}$, $\tilde{Q}_{M}$, $\tilde{Q}_{MM}$, $\tilde{Q}_{FGLB}$, $\tilde{Q}_{GFK}$, $\tilde{Q}_{GFD}$ and $\tilde{Q}_{EXLB}$ tests decrease or increase as the lag increases, see Figures 3.5 and 3.6.

7.3 New multivariate portmanteau test.

The aim of the VARMA simulation study in Chapter 6 was to compare the new exponential portmanteau test $\tilde{Q}_{EXCO}$ of vector autoregressive moving average models to those identified in previous studies of vector portmanteau tests, namely, Hosking's $\tilde{Q}_{H}$ test (1980), and Mahdi and McLeod's $\tilde{Q}_{VMM}$ test (2011). Monte-Carlo experiments were conducted for four different VAR(1) processes, then fitted under a VAR(1) model with n = 100 and 200, and with $m = 10, 20$ and 30 in the simulation of the empirical size. The power level simulation study, which was conducted in Section 6.3.3, involved the generation of 12 different VARMA(2,2) processes, each of which were fitted under either a VAR(1) or a VMA(1) model (with n = 100 and 200, and with $m = 10, 20$ and 30).

The simulation study of empirical size shows that none of the $\tilde{Q}_{H}$, $\tilde{Q}_{VMM}$ and $\tilde{Q}_{EXCO}$ tests are better than the others in all cases, see Tables 6.1 and 6.2. However, the power level simulation study shows that the new portmanteau $\tilde{Q}_{EXCO}$ test is more powerful compared with the $\tilde{Q}_{H}$ and $\tilde{Q}_{VMM}$ tests, for more details see Tables 6.3, 6.4, 6.5 and 6.6.

A Monte-Carlo experiment of empirical size was conducted with data generated by

$$\boldsymbol{\Phi}_2 = \begin{pmatrix} -1.5 & 1.2 \\ -0.9 & 0.5 \end{pmatrix}, \qquad \boldsymbol{\Sigma}_2 = \begin{pmatrix} 1.0 & 0.5 \\ 0.5 & 1.0 \end{pmatrix}$$

in Section 6.3.2, with different number of lags and $n = 150$, then fitted under a VAR(1) model.

The data were generated by

$$\begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} - \begin{pmatrix} 0.5 & 0.1 \\ 0.4 & 0.5 \end{pmatrix} \begin{pmatrix} z_{1,t-1} \\ z_{2,t-1} \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0.3 & 0 \end{pmatrix} \begin{pmatrix} z_{1,t-2} \\ z_{2,t-2} \end{pmatrix} = \begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1.00 & 0.71 \\ 0.71 & 1.00 \end{pmatrix}$$

in Section 6.3.3, with different number of lags and $n = 150$, then fitted under VAR(1) model in the simulation of power level. This simulation study of different lags shows that the $\tilde{Q}_{\mathrm{EXCO}}$ test is more stable than the $\tilde{Q}_{\mathrm{H}}$ and $\tilde{Q}_{\mathrm{VMM}}$ tests, in that, the $\tilde{Q}_{\mathrm{H}}$ and $\tilde{Q}_{\mathrm{VMM}}$ tests decrease or increase as different length of lag are employed in both empirical size and power level simulation studies, see Figures 6.1, 6.2, 6.3 and 6.4.

7.4 Summary and key findings

The thesis has shown that the length of data has a strong impact on the apparent stationarity of a time series when data generated by AR(1) process and AR(2) process, and it has demonstrated how to select the appropriate estimated value of parameter for a corresponding length of time series.

For a univariate time series the $\tilde{Q}_{EXM}$ test has been presented and shown to be more powerful than the $\tilde{Q}_{EXLB}$ test and all the previous tests (namely, the $\tilde{Q}_{LB}$, $\tilde{Q}_{M}$, $\tilde{Q}_{MM}$, $\tilde{Q}_{FGLB}$, $\tilde{Q}_{GFK}$ and $\tilde{Q}_{GFD}$ tests). A new portmanteau test of a multivariate time series has been presented and proven to be better that the $\tilde{Q}_{H}$ and $\tilde{Q}_{VMM}$ tests.

7.5 Future work

Chapter 4 shows how the length of a time series affects its apparent stationarity when data are generated from a known statistical process of a univariate autoregressive process. Future work will examine how the length of a time series affects the apparent stationarity when data are generated from a moving average process of a univariate time series with positive and negative values of parameters. Another area that will be investigated is how the length of a time series affects the apparent stationarity when data are generated from ARMA(1,1) process with positive and negative values of parameters. Other types of unit root tests will be used to examine how the length of a time series affects the apparent stationarity, such as the tests of Phillips and Perron (PP), Kwiatkowski, Phillips, Schmidt and Shin (KPSS), and Eliot, Rothenberg and Stock (ADF-GLS).

In Chapter 5 new portmanteau tests for univariate time series were developed. These new portmanteau test statistics assume the data is non-seasonal. Future work will aim to extend these tests to seasonal data and make a comparison with previous studies of seasonality portmanteau tests.

The Autoregressive Conditional Heteroscedastic (ARCH) model explicitly models the change in variance within a time series over time (Engle, 1982). Current practice is for the $\tilde{Q}_{LB}$ portmanteau test to be used to test for remaining ARCH model effects in the variance equation and to check the specification of the variance equation. The new portmanteau test statistics will be extended in future work and applied to the (ARCH) model and compared with the existing $\tilde{Q}_{LB}$ test.

Bollerslev (1986) extended the ARCH model to Generalized Autoregressive Conditional Heteroscedastic (GARCH) models, in which the current conditional variance equation also includes the past conditional variance. Bollerslev (1986) already suggested a Lagrange multiplier (LM) test for testing a GARCH model against a higher order GARCH model, Li and Mak (1994). The new portmanteau test statistics will be extended in future work and applied to the (GARCH) model and compared with the existing LM test.

In Chapter 6, the new multivariate portmanteau test statistic was applied to multivariate non-seasonal data. This equation will be extended in future work to cover seasonal data, and work will be undertaken to compare the new multivariate portmanteau test statistics with previous studies of seasonal portmanteau tests.

Many researchers have extended the Generalized Autoregressive Conditional Heteroscedastic (GARCH) model from the univariate to the multivariate case to create the Multivariate Generalized Autoregressive Conditional Heteroscedastic (MGARCH) model. For instance, Bollerslev (1990) studied the changing variance structure of the exchange rate regime in the European Monetary System, assuming the correlations to be time invariant. Ling and Li (1997) introduced a portmanteau test for testing the adequacy of the multivariate MGARCH model. The new multivariate portmanteau test statistic will be extended in future work to enable it to apply to the MGARCH model.

References

Akaike, H., (1973), *Information theory and extension of the maximum likelihood principle*, In B.N. Petrov and Csaki, F. (eds.), 2nd International Symposium on Information Theory, Akademi ai Kiado, Budapest, pp. 267-281.

Akaike, H., (1974), *A new look at the statistical model identification*, IEEE Transactions Automatic Control, Vol. AC-19, pp. 716-723.

Anderson, T. W., (1980) *Maximum likelihood estimation for vector autoregressive moving average models*, In Direction in Time series, eds., D. R. Brillinger and O. C. Tiao, Institute of Mathematical Statistics, pp. 49-59.

Anscombe, F. J., (1961), *Examination of residual*, Proc. 4th Berkeley Symp., 1, pp. 1-36.

Anscombe, F. J. and Tukey, J. W., (1963), *The examination and analysis of residual*, Technometrics, Vol. 5, pp. 141-160.

Bachelier, L., (1900), *Théorie de la of Spéculation*, Annales de L Ecole Normle Supérieure third series 17, pp. 22-86. Reprinted in (1964), *Theory of Speculation*, trans. James Boness, A., The Random Character of Stock Market Prices, edited by Cootner, P. H., pp. 17- 78, MIT Press, Cambridge, MA.

Barnard, G. A., (1949), *Statistical inference*, Journal of the Royal Statistical Society, B11, pp. 115-139.

Bartlett, M. S., (1946), *On the theoretical specification of sampling properties of auto-correlated time series*, Journal of the Royal Statistical Society, B8, 27.

Beveridge, Sir William H., (1922), *Crop yields and rainfall in western Europe*, Journal of the Royal Statistical Society, 85, 412.

Birnbaum, A., (1962), *On the foundations of statistical inference*, Journal American Statistical Association, Vol. 57, pp. 269-306.

Bollerslev, T., (1986), *Generalized autoregressive conditional heteroskedasticity*, Journal of Econometrics, 31: 307–327.

Bollerslev, T., (1990), *Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model*, Review of Economics and Statistics, 72, 498-505.

Box, G. P., (1954), *Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, I. Effect of Inequality of Variance in the One-Way Classification*, The Annals of Mathematical Statistics, Vol. 25, No. 2, pp. 290-302.

Box, G. E. P. and Jenkins, G. M., (1962), *Some statistical aspects of adaptive optimization and control*, Journal of the Royal Statistical Society, Series B, Vol. 24, No. 2, pp. 297-343.

Box, G. E. P. & Jenkins, G. M., (1970), *Time series analysis: Forecasting and control,* San Francisco: Holden-Day.

Box, G. E. P., Jenkins, G., M., Reinsel, G. C. and Ljung, G. M., (2015), *Time series analysis: Forecasting and control,* 5th ed., Hoboken, NJ: Wiley.

Box, G. P. and Pierce, D. A., (1970), *Distribution of residual autocorrelations in autoregressive integrated moving average time series models*, Journal of the American Statistical Association, Vol. 65, No. 332, pp. 1509-1526.

Box, G. E. P. and Muller, M. E., (1958), *A note on the generation of random normal deviates*, The Annals of Mathematics Statistics, 29, pp. 610-611.

Box, G. E. P. and Tiao, G. C., (1965), *A change in level of a non-stationary time series*, Biometrika, Vol. 52, No. 1/2, pp. 181-192.

Brinton, W. C., (1914), *Graphic methods for presenting facts*, The Engineering Magazine Company, New York, NY., USA.

Brockwell, P. J. and Davis, R. A., (1991), *Time Series Theory and Methods*, Springer-Verlag New York.

Chitturi, R. V., (1974), *Distribution of residual autocorrelations in multiple autoregressive schemes*, Journal of American Statistical Association, Vol. 69, pp. 928-934.

Cowles, A., (1933), *Can stock market forecasters forecast*? Econometrica, Vol. 1, pp. 309-324.

Cowles, A., (1944), *Stock Market forecasting*, Econometrica, Vol. 12, No. 3/4, pp. 206-214.

CPI, (2015), *The consumer price index12-month inflation rate for the last 10 years: January 2005 to January 2015*, Available at *http://www.ons.gov.uk/ons/rel/cpi/consumer-price-indices/january-2015/stbconsumer-price-indicesjanuary2015.html?format=print* (Accessed: 6 October 2013).

Daniels, H. E., (1956), *The approximate distribution of serial correlation coefficients*, Biometrika, 43, 169-185.

Davidson, J., (1994), *Stochastic limit theory*, New York, Oxford University Press.

Davis, H. T., (1941), *The analysis of the economic time series*, India.

Davies, N., Triggs, C. M. and Newbold P., (1977), *Significance levels of the Box-Pierce portmanteau statistic in infinite samples,* Biometrika, 64, pp. 517-522.

Dickey, D.A., (1976), *Estimation and hypothesis testing in non-stationary time series*, Iowa State University Ph.D. Thesis.

Dickey, D.A. and Fuller W.A., (1979), *Distribution of the estimates for autoregressive time series with unit root*, Journal of the American Statistical Association,74, pp. 427-431.

Dickey D.A. and Fuller W.A., (1981), *Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root*, Econometrica, 49, pp. 1057-1072.

Doob, J. L., (1953), *Stochastic Process*, Wiley, New York.

Duchesne, P. and Roy, R. (2004), O*n consistent testing for serial correlation of unknown form in vector time series models*, Journal of Multivariate Analysis, 89, pp. 148–180.

Durbin, J., (1959), *Efficient estimation of parameters in moving-average models*, Biometrika, Vol. 45, pp. 306-316.

Durbin, J., (1960), *The fitting of time series models*, Rev. International Statistical Institute, 28, pp. 233-244.

Elliott G., Rothenberg T. J., and Stock J. H., (1996), *Efficient tests for an autoregressive unit root*, Econometrica, 64(4): 813–836.

Engle, R., (1982), *Autoregressive Conditional Heteroskedasticity with Estimate of Variance of U.K Inflation*, Econometrica, 50: 987-1008.

Engle, R. F. and Granger, C. W. J., (1987), *Cointegration and error correction: Representation, Estimation and Testing*, Econometrica, Vol. 55, pp. 251-276.

Fisher, I., (1925), *Our unstable dollar and the so-called business cycle*, Journal of the American Statistical Association, Vol. 20, pp. 179- 202.

Fisher, T. J. and Gallagher, C. M., (2012), *New Weighted Portmanteau Statistics for time series goodness of fit testing,* Journal of the American Statistical Association, Vol. 107(498), pp. 777-787.

Fuller, W.A., (1976), *Introduction to statistical time series. New York*, John Wiley and Sons.

Gallagher, C. M. and Fisher, T. J., (2015), *On weighted of portmanteau tests for time-series goodness-of-fit*, Journal of Time Series Analysis, Vol. 36, pp. 67-83.

Granger, C. W. J. and Andersen, A., (1978), *On the invertibility of time series models*, Stochastic Processes and their Applications, Vo. 8, No. 1, pp. 87-92.

Granger, C. W. J., (1981), *Some properties of time series data and their use in econometric model specification*, Journal of Econometrics, Vol. 16, pp. 121-130.

Grenander, U. and Rosenblatt, M., (1957), *Statistical analysis of stationary time series*, New York Wiley.

Guy, H. and Donald, C., (1949), *A sampling study of the merits of autoregressive and reduced from transformations in regression analysis*, Journal of American Statistical Association, Vol. 44, No. 247, pp. 356-372.

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

Hannan, E. J., (1981), *Estimating the dimension of a linear system*, Journal of Multivariate Analysis, Vol. 11, pp. 459-473.

Hannan, E. J. and Kavalieris, L., (1984), *Multivariate linear time series models*, Advance in Applied Probability, 16, pp. 492-561.

Hannan, E. J. and Quinn, B. G., (1979), *The determination of the order of an autoregression*, Journal of the Royal Statistical Society, Series B 41, pp. 190-195.

Hillmer, S. C. and Tiao, G. C., (1979), *Likelihood function of stationary multiple autoregressive moving-average models*, Journal of the American Statistical Association, Vol. 74, pp. 652-660.

Hong, Y., (1996a), Consistent testing for serial correlation of unknown form, Econometrica, 64(4): 837–864.

Hong, Y., (1996b), Testing for independence between two covariance stationary time series, Biometrika, 83(3): 615–625.

Hosking, J. R. M., (1980), *The multivariate portmanteau statistics*, Journal of American Statistical Association, Vol. 75, pp. 602-608.

Hosking, J. R. M., (1981), *Equivalent forms of the multivariate portmanteau statistic*, Journal of the Royal Statistical Society, Series B 43, pp. 261-262.

Hooker, R. H., (1901), *Correlation of the marriage-rate with trade*, Journal of the Royal Statistical Society, Vol. 64, pp. 485-492.

Hooker, R. H., (1905), *On the correlation of successive observations*, Journal of the Royal Statistical Society, Vol. 68, pp. 696-703.

Huong, N. T., (2013), *Goodness-of-fit in multivariate time series*, Ph.D. Thesis, University Carlos III de Madrid.

Hurwicz, L., (1945), *Least-square bias in time series*, In Statistical Inference in Dynamic Economic Models, ed., T.C. Koopmans, New York, John Wiley, pp. 365-383.

Jenkins, G. M., (1954, 1956), *Tests of hypotheses in the linear autoregressive model*, I, Biometrika, 41, pp. 405-419, II, Biometrika, 43, pp. 186-199.

Jenkins, G. M. and Watts, D. G., (1968), *Spectral analysis and its applications*, Holden-Day, San Francisco.

Kalman, R. E., (1960), *A new approach to linear filtering and prediction problems*, Transactions of the ASME Journal of Basic Engineering, 82, Series D, pp. 35-45.

Kalman, R. E. and Bucy, R., (1961), *New results in linear filtering and prediction*, Journal of Basic Engineering, 83 D, pp. 95-108.

Kendall, M. G., (1945), *On the analysis of oscillatory time series*, Journal of the Royal Statistical Society, A108, pp. 93-129.

Kendall, M. G., (1949), *The estimation of parameters in linear autoregressive time series*, Econometrica, Vol. 17, Supplement: Report of the Washington Meeting, pp. 44-57.

Kendall, M. G. and Stuart, A., (1966), *The advanced theory of statistics*, Vol. 3, Charles Griffin, London.

Khinchin, A. Y., (1933), *Úber stationäre reihen zufalliger variablen*, Matematicheskii Sbornik, Vol. 40, pp. 124-128.

Khinchin, A. Y., (1934), *Korrelationstheorie der stationären stochastischen prozesse*, Mathematische Annalen, Vol. 109, pp. 604-615.

Kolmogorov, A., (1931), *Úber die analytischen methoden in der wahrscheinlichkeit-srechnung*, Mathematische Annalen, Vol. 104, pp. 415-458.

Kolmogorov, A., (1941), *Interpolation und Extrapolation von stationaren zufalligen folgen*, Bull. Acad. Sci. (Nauk) U.S.S.R., Ser. Math., Vol. 5, pp. 3-14.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y., (1992), *Testing the null hypothesis of stationarity against the alternative of a unit root*, Journal of Econometrics, 54: 159-178.

Lauritzen, S. L., (1981), *Time series analysis in 1880: A Discussion of Contributions Made by T.N. Thiele*, International Statistical Review, 49, pp. 319-331.

Lauritzen, S. L., (2002), *Thiele: Pioneer in Statistics*, Oxford, Oxford University Press.

Li, W. K. and McLeod, A. I., (1981), *Distribution of the residual autocorrelation in multivariate ARMA time series models*, Journal of the Royal Statistical Society, B 43, pp. 231-239.

Ljung, G. M. and Box, G. E. P., (1978), *On a measure of lack of fit in time series models*, Biometrika, Vol. 65, No. 2, pp. 297-303.

Li, W.K., Mak, T.K., (1994), *On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity*, Journal of Time Series Analysis, 15, 627–636.

Ling S. and Li, W. K., (1997), *Diagnostic checking of nonlinear multivariate time series with multivariate arch errors*, Journal of time series analysis, Vol. 18, No. 5, pp. 447-464.

Lütkepohl, H., (2005), *New Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.

Mahdi, E. and McLeod I. A., (2011), *Improved multivariate portmanteau test*, Journal of Time Series Analysis 33(2), pp. 211-222.

Mann, H. B. and Wald, A.,(1943), *On the statistical treatment of linear stochastic difference equations*, Econometrica, Vol. 11, No. ¾, pp. 173-220.

Margret, A. W., (1929), *Morgenstern on the methodology of economic forecasting*, Journal of Political Economy, Vol. 37, pp. 312-339.

Matsumoto, M. and Nishimura T., (1998), *Mersenne Twister: A 623-dimensionally equidistributed uniform pseudo-random number generator*, ACM Transactions on Modeling and Computer Simulation, 8, pp. 3-30.

McLeod, A. I. (1978), *On the Distribution of Residual Autocorrelations in Box-Jenkins Models*, Journal of the Royal Statistical Society, Ser. B, 40, 296–302.

Monti, A. C., (1994), *A proposal for residual autocorrelation test in linear models*, Biometrika, Vol. 81, No. 4, pp. 776-780.

Nason, G. P., (2013), *A test for second order stationarity and approximate confidence intervals for localized autocovariances for locally stationary time series*, Journal of the Royal Statistical Society Series B, 75, 879-904.

Neudecker, H., (1969), *Some theorems on matrix differentiation with special reference to Kronecker matrix products*, Journal of American Statistical, Vol. 64, pp. 953-962.

Nicholls, D. F. and Hall, A. D., (1979), *The exact likelihood function of multivariate autoregressive moving average models*, Biometrika, Vol. 66, pp. 259-264.

Osborn, D. R., (1977), *Exact and approximate maximum likelihood estimators for vector moving average process*, Journal of the Royal Statistical Society, B 39, pp. 114-118.

Ng, S. and Perron, P., (2001), *Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power*, Econometrica, 69: 1519-1554.

Pankratz, A., (1983*), Forecasting with univariate Box-Jenkins models concepts and cases*, John Wiley and Sons. Inc.

Peña, D., Rodríguez, J., (2002), *A powerful portmanteau test of lack of fit from time series*, Journal of the American Statistical Association, Vol. 97, No. 458, pp. 601-610.

Peña, D., Rodríguez, J., (2006), *The log of the autocorrelation matrix for testing goodness of fit in time series,* Journal of statistical planning and inference, Vol. 136 (8), pp. 2706-2718.

Perron, P and Ng. S., (1996), *Useful Modifications to Unit Root Tests with Dependent Errors and their Local Asymptotic Properties*, Review of Economic Studies, 63(3): 435-465.

Persons, W. M., (1924), *The problem of Business Forecasting*, Pollak Foundation for Economic Research Publications, No. 6, London, Pitman.

Phadke, M. S. and Kedem, G., (1978), *Computation of the exact likelihood function of multivariate moving average models*, Biometrika, 65, pp. 511-519.

Phillips, P.C.B., and Perron P., (1988), *Testing for a unit root in time series regression*, Biometrika, 75: 335–346.

Playfair, W., (1801), *The Commercial and Political*, 3rd ed. London.

Poskitt, D. S. and Tremayne, A. R., (1982), *Diagnostic tests for multiple time series models*, The Annals of Statistics, Vol. 10, pp. 114-120.

Priestley, M. B. and Rao, T. S., (1969), *A test for non-stationary of time series*, Journal of Royal Statistical Society, Series B, Vol. 31, No. 1, pp. 140-149.

Quenouille, M. H., (1949), *Approximate tests of correlation in time series*, Journal of Royal Statistical Society, B11, pp. 68-84.

Quenouille, M. H., (1957), *The analysis of multiple time series*, Griffin, London.

Quinn, B. G., (1980), *Order determination for a multivariate autoregression*, Journal of Royal Statistical Society, B42, No. 2, pp. 182-185.

Rao, C. R., (1965), *Linear statistical inference and its applications*, Wiley, New York.

Reinsel, G., (1979), *FIML estimation of the dynamic simultaneous equations model with ARMA disturbances*, Journal of Econometrics, Vol. 9, pp. 263-281.

Reinsel, G. C., (1993), *Elements of Multivariate Time Series Analysis*, Springer-Verlag, New York.

Reinsel, G.C., Basu, S. and Yap, S. F., (1992), *Maximum likelihood estimators in the multivariate autoregressive moving average model from a generalized least squares viewpoint*, Journal of Time Series Analysis, Vol. 13, pp. 133-145.

Satterthwaite, F. E., (1941), *Synthesis of variance,* Psychometrik*a*, 6, pp. 309-316.

Satterthwaite, F. E., (1946), *An approximate distribution of estimates of variance components*, Biometrics Bulletin, 2(6), pp. 110-114.

Schuster, A., (1897), *On lunar and solar periodicities of earthquakes*, Proceedings of the Royal Society of London, Vol. 61, pp. 455-465.

Schuster, A., (1906), *On the periodicities of sunspots*, Philosophical Translations of the Royal Society of London, Series A, Vol. 206, pp. 69-100.

Schwarz, G., (1978), *Estimating the dimension of a model*, Annals of Statistics, Vol. 6, pp. 461-464.

Schwert, G.W., (1989), *Test for unit roots: A Monte Carlo Investigation*, Journal of Business and Economic Statistics, 7(2), pp. 147-159.

Sims, C. A., (1980), *Macroeconomics and Realty*, Economitrica, Vol. 48, pp. 1-48.

Slutzky, E., (1937), *The summation of random causes as the source of cyclic processes, Econometrica*, 5, 105-146 (1937), Translated from the earlier paper of the same title in Problems of Economic conditions (Ed. Conjuncture Institute Moscow).

Stralkowski, C. M., (1968), *Lower order autoregressive-moving average stochastic models and their use for the characterization of abrasive cutting tools*, Ph.D. Thesis, University of Wisconsin.

Student (Gosset, W. S.), (1914), *The elimination of spurious correlation due to position in time and space*, Biometrika, Vol. 10, pp. 179-180.

Thiele, T. N., (1880a), *Om Anvendelse af mindset Kvadraters Methode I nogle Tilfalde, hvor en Komplikation af vise Slags uensartede tilfaldige Fejlkilder giver Fejlene en 'systematisk' Karakter*, Vidensk, Selsk. Skr. 5, RK., naturvid, og mat , Afd., vol. 12, pp. 381-408.

Thiele, T. N., (1880b), *Sur la compensation de quelques erreurs- quasl-systematiques par la méthode des moindres carrés*, Kobenhaven: Reitzel, (French translation of Thiele (1880).

Tiao, G. C. and Box, G. E. P., (1981), *Modeling multiple time series with applications*, Journal of the American Statistical Association, Vol. 76, pp. 802-816.

Tiao, G. C. and Tsay, R. S., (1989), *Model specification in multivariate time series*, Journal of the Royal Statistical Society, Series B, Vol. 51, pp. 157-213.

Tsay, R. S., (2005), *Analysis of Financial Time series*, 2nd ed., New York, Wiley.

Tunnicliffe, W. G., (1973), *The estimation of parameters in multivariate time series models*, Journal of the Royal Statistical Society, B, Vol. 35, pp. 76-85.

Walker, A. M., (1950), *Note on a generalization of the large sample goodness of fit test for linear autoregressive schemes,* Journal of the Royal Statistical Society, Series B, Vol. 12, No. 1, pp. 102-107.

Walker, A. M., (1961), *Large-sample estimation of parameters for moving-average models*, Biometrika, Vol. 48, pp. 343-357.

Walker, G. T., (1931), *On periodicity in series of related terms*, Proceedings of the Royal Society of London, Series A, Vol. 131, pp. 518-532.

Wallis, K. F., (1977), *Multiple time series analysis and the final form of econometric models*, Econometrica, Vol. 45, pp. 1481-1497.

Wei, W. W. S., (2006), *Time series analysis univariate and multivariate methods*, 2nd edition, United States America.

Whittaker, E. T. and Robinson, G., (1924), *The Calculus of Observations*, OLD Bailey, London.

Whittle, P., (1963), *On the fitting of multivariate autoregressions, and the approximate canonical factorization of a spectral density matrix*, Biometrika, Vol. 50, pp. 129-134.

Wiener, N., (1949), *Extrapolation Interpolation and Smoothing of Stationary Time Series*, Wiley, New York.

Wold, H., (1938), *A study in the analysis of stationary time series*, Uppsala, 2nd ed..

Wunderlich, C. A., (1870), *Dsa Verhalten der Eigenwärme in Krankheiten*, Otto Wigand, Leipzig, Germany, 2nd edition.

Yaglom, A. M., (1962), *An introduction to the theory of stationary random function*, Prentice Hall, Englewood Cliffs, NJ.

Yule, G. U., (1895), *On the correlation of total pauperism with proportion of out relief*, I: all ages, Economic Journal, Vol. 5, No. 20, pp. 603-611.

Yule, G. U., (1896), *On the correlation of total pauperism with proportion of out relief*, II: males, over sixty-five, Economic Journal, Vol. 6, No. 24, pp. 613-623.

Yule, G. U., (1897a), *On the significant of Bravais' formula for regression, &c, in the case of skew variable*, Proceedings of the Royal Society of London, Vol. 60, pp. 477-489.

Yule, G. U., (1897b), *On the theory of correlation*, Journal of the Royal Statistical Society, Vol. 60, No. 4, pp. 812-854.

Yule, G. U., (1907), *On the theory of correlation for any number of variables*, treated by a new system of notation, Proceedings of the Royal Society of London, serial A, Vol. 79, pp. 182-193.

Yule, G. U., (1927), *On the method of investigating periodicities in distributed series with special reference to Wolfer's sunspot numbers*, Phil. Trans., A 226, pp. 267- 295.

Zellner, A. and Palm, F., (1974), *Time series analysis and simultaneous equation econometric*, Journal of Econometrics, Vol. 2, pp. 17-54.

Appendix A

R program for data length simulation

```
ns<-10000    ### number of simulations
n<-25          ### number of data
a1<- list(mode="vector",length=ns) ## produce vector as same length of simulation
ad1<- list(mode="vector",length=ns)
a2<- list(mode="vector",length=ns)
ad2<- list(mode="vector",length=ns)
a3<- list(mode="vector",length=ns)
ad3<- list(mode="vector",length=ns)
for (i in 1:ns)
{
fun<-rnorm(n)        ### generate data by using normal distribution
y<- 0.5*seq(1,n)     ### the trend sequence
sim<-arima.sim(list(ar=0.1),n=n,innov=fun) ### simulate AR(1)
simd<-arima.sim(list(ar=0.1),n=n,innov=fun)+0.5  ### simulate AR(1) with drift
simt<-arima.sim(list(ar=0.1),n=n,innov=fun)+0.5+y   ### simulate AR(1) with drift and
trend
adf1<-ur.df(sim,type="none",lags=0) ### DF test
adf2<-ur.df(simd,type="drift",lags=0) ### DF test with drift
adf3<-ur.df(simt,type="trend",lags=0)  ### DF test with drift and trend
a1[[i]]<-adf1@teststat                        ### critical value of  DF test
ad1[[i]]<-a1[[i]]<adf1@cval[2]
a2[[i]]<-adf2@teststat[1,1]               ### critical value of  DF test with drift
ad2[[i]]<-a2[[i]]<adf2@cval[1,2]
a3[[i]]<-adf3@teststat[1,1]
ad3[[i]]<-a3[[i]]<adf3@cval[1,2]       ### critical value of  DF test with drift and trend
}
length(ad1[ad1 ==FALSE])   ### number of false for DF test
length(ad2[ad2 ==FALSE])   ### number of false for DF test with drift
length(ad3[ad3 ==FALSE])   ### number of false for DF test with drift and trend
```

Appendix B

R program of simulate empirical size and power level for univariate portmanteau tests.

```
n<-100                         ### number of observations
ns<-1000                       ### number of simulations
m<-10                          ### maximum lag
round <- round(log(n))     ### indicating the number of decimal places
round1<-round +1
mmsd<-((3*m*(m+1))/(2*(2*m+1)))-1
mmchi<-qchisq(.95, df=mmsd)  ### chi-squared critical value Mahdi and McLeod (2012)
boxchi<- qchisq(.95, df=(m-1)) ### chi-squared critical value Ljung and Box (1978)
w<- numeric(m) ### creates a real vector of the specified length
wfg<- numeric(m)
r<- numeric(m)
p<- numeric(m)
fg<- numeric(m)
exacf<- numeric(m)
expacf<- numeric(m)
kernel<- numeric(m)
ker <- numeric(m)
rdaw<- numeric(round)
ss<- numeric(round)
ss1<- numeric(round)
ln<- numeric(m)
lnlagacf<- numeric(m)
lnsquar<- numeric(m)
newsd<- numeric(m)
fgsd<- numeric(m)
Qlb <- list(mode="vector",length=ns) ## produce vector as same length of simulation
Qm <- list(mode="vector",length=ns)
```

```
Qlbe <- list(mode="vector",length=ns)
Qme <- list(mode="vector",length=ns)
Qfg<- list(mode="vector",length=ns)
Qke<- list(mode="vector",length=ns)
boxtest<- list(mode="vector",length=ns)
montitest<- list(mode="vector",length=ns)
boxnewtest<- list(mode="vector",length=ns)
montinewtest<- list(mode="vector",length=ns)
fgtest<- list(mode="vector",length=ns)
ketest<- list(mode="vector",length=ns)
MLtest<-list(mode="vector",length=ns)
Qmm<- list(mode="vector",length=ns)
matrixcolumn<- list(mode="vector",length=ns)
dawtest<- list(mode="vector",length=ns)
Qdaw<- list(mode="vector",length=ns)
gammadaw<- list(mode="vector",length=ns)
y<-c(1:m)
for (j in 1:m){
newsd[j]<-(1/m)^((y[j]-1)/m) ### loop for weight new test
fgsd[j]<- ((m- y[j]+1)/(m))     ### loop for weight Fisher and Gallagher (2012) test
}
squarnewsd<-(newsd^2)   ### square the weight of the new test
sum1<-(sum(newsd))^2    ### sum the square weight  of the new test
sum2<- 2*(sum(squarnewsd)-1)
divd1<- sum1/sum2
divd2<- sum2/sqrt(sum1)
gamma<-qgamma(0.95,divd1,1/divd2) ### gamma distribution for new test
squarfgsd<-(fgsd^2)       ##### square the weight of the FG (2012) test
sumfg1<-(sum(fgsd))^2    ### sum the square weight  of the FG (2012) test
sumfg2<- 2*(sum(squarfgsd)-1)
divdfg1<- sumfg1/sumfg2
divdfg2<- sumfg2/sqrt(sumfg1)
```

```
gammafg<-qgamma(0.95,divdfg1,1/divdfg2) ###gamma distribution for FG (2012) test
set.seed(1234)
for (i in 1:ns)
{
fun<-rnorm(n) ### generate random number
sim<-arima.sim(list(ar=(0.1)),n=n,innov=fun) ### simulate AR(1) process
fit <- arima(sim, order = c(1,0,0))  ### fitted AR(1) model
res<-residuals(fit) ### the residual of the fitted model
acf<- acf(res, lag=m, plot=F)$acf  ### autocorrelation function
pacf<- pacf(res, lag=m, plot=F)$acf  ### partial autocorrelation function
for(l in 1:m) ### loop for the weight with maximum lag
{
w[l]<-(1/m)^((y[l]-1)/m) ### for new test
wfg[l]<- ((m- y[l]+1)/(m)) ### for FG (2012) test
r[l]<-((acf[l+1])^2)/(n-l)
p[l]<-((pacf[l])^2)/(n-l)
fg[l]<- wfg[l]* r[l]
exacf[l]<- w[l]* r[l]
expacf[l]<- w[l]* p[l]
kernel[l]<-((n+2)/(n-l))*((sin(sqrt(3)*pi*(l/m))/(   sqrt   (3)*pi*(l/m)))^2)*(r[l]*(n-l))   ##
kernel test
sumkernel<-n*sum( kernel)
ker[l]<-(((sin(sqrt(3)*pi*(l/m)))/(sqrt(3)*pi*(l/m)))^2)*( (n+2)/(n-l))
sumk1<- sumkernel-sum(ker)
sumk2<- sqrt(2*sum( (ker)^2))
}
for (s in 1:round) ###critical value of data adaptive weight
{
rdaw[s]<-((acf[s+1])^2)/(n-s)
ss[s]<-((n+2)/(n-s))
ss1[s]<-((n+2)/(n-s))^2
}
for(t in round1:m)
```

```
{
ln[t]<- -(log(1-abs(pacf[t])))
lnlagacf[t]<-ln[t]*(acf[t+1])^2
lnsquar[t]<-(ln[t])^2
}
Qlb[[i]]<-n*(n+2)*sum(r)  ### Ljung and Box
Qm[[i]]<-(n*(n+2))*sum(p) ### Monti (1994)
Qfg[[i]]<-(n*(n+2))*sum(fg) ### Fisher and Gallagher (2012)
Qlbe[[i]]<-(n*(n+2))*sum(exacf) ### new test with autocorrelation function
Qme[[i]]<-(n*(n+2))*sum(expacf)  ### new test with partial autocorrelation function
Qmm[[i]]<-portest(res,lag=m,order=0,SquaredQ=FALSE,Kernel=FALSE)###Mahdi  and
McLeod (2012)
matrixcolumn[[i]]<-Qmm[[i]][2]
MLtest[[i]]<-matrixcolumn[[i]]<mmchi ###Look up the 5 percentage point
boxtest[[i]]<- Qlb[[i]]< boxchi ###Look up the 5 percentage point
montitest[[i]]<- Qm[[i]]< boxchi ###Look up the 5 percentage point
boxnewtest[[i]]<- Qlbe[[i]]< gamma ###Look up the 5 percentage point
montinewtest[[i]]<- Qme[[i]]< round(gamma,digits=2)
fgtest[[i]]<- Qfg[[i]]< gammafg
Qke[[i]]<- sumk1/sumk2
ketest[[i]]<- Qke[[i]]< 1.6
Qdaw[[i]]<-n*(n+2)*sum(rdaw)+n*sum(lnlagacf) ### data adaptive weight test
sumsd1<-(sum(ln)+sum(ss))^2
sumsd2<-2*((sum(lnsquar)+sum(ss1))-1)
alpha<- sumsd1/ sumsd2
beta<- sumsd2/sqrt(sumsd1)
gammadaw[[i]]<-qgamma(0.95,alpha,1/beta)
dawtest[[i]]<- Qdaw[[i]]< gammadaw[[i]]
}
length(boxtest[boxtest ==FALSE])/ns ### account number of rejection
length(montitest[montitest ==FALSE])/ns
length(MLtest[MLtest ==FALSE])/ns
length(fgtest[fgtest ==FALSE])/ns
```

```
length(ketest[ketest ==FALSE])/ns
length(dawtest[dawtest ==FALSE])/ns
length(boxnewtest[boxnewtest ==FALSE])/ns
length(montinewtest[montinewtest ==FALSE])/ns
```

Appendix C

R program of simulate empirical size and power level for multivariate portmanteau tests.

```
n<-150              ### number of observations
m<-20               ### maximum lag
k <- 2              ### vector dimension
ns<-10000           ### number of simulations
lag<-k^2*m
lag2<-k^4*m
tracelag<- numeric(m)   ### creates a real vector of the specified length
tracelagw<- numeric(m)
w<- numeric(m)
newsd<- numeric(lag)
Qvho<- list(mode="vector",length=ns)  ### produce vector as same length of simulation
Qve<- list(mode="vector",length=ns)
hostest<- list(mode="vector",length=ns)
newtest<- list(mode="vector",length=ns)
MLtest<-list(mode="vector",length=ns)
Qmm<- list(mode="vector",length=ns)
matrixcolumn<- list(mode="vector",length=ns)
mmsd<-(k^2)*((3*m*(m+1))/(2*(2*m+1)))-(k^2)
mmchi<-qchisq(.95, df=mmsd) ### critical value for vector Mahdi and McLeod (2011) test
hchi<- qchisq(.95, df=k^2*(m-1)) ### critical value for vector Hosking (1980) test
y<-c(1:lag)
for (j in 1:lag){
newsd[j]<-(1/(lag2))^((y[j]-1)/lag) ### loop for the weight new vector portmanteau test
}
squarnewsd<-(newsd^2)
sum1<-sum(newsd)^2
sum2<- k^2*(sum(squarnewsd)-1)
```

```
divd1<- sum1/sum2
divd2<- sum2/sqrt(sum1)
gamma<-qgamma(0.95,divd1,1/divd2) ## gamma distribution for new vector portmanteau test
set.seed(1234)
for (i in 1:ns)
{
phi <- array(c(0.5,0.4,0.1,0.5,0,0.3,0,0),dim=c(k,k,2))
theta <- NULL
d <- NA
sigma <- matrix(c(1,0.71,0.71,1),k,k)
sim1 <- varima.sim(phi, theta, d, sigma, n) ### generate data
fitVAR <- VAR(sim1, p=1) ### fitted data
res <- residuals(fitVAR)  ### residual of the fitted data
cov <- acf(res[,1:k],lag.max=m,type="covariance",plot=F)$acf ## covariance matrix
invclag0 <- solve(cov[1,,]) ### inverse matrix
x<-c(1:m)
for(l in 1:m)
{
w[l]<-(1/((k^2)*m))^((x[l]-1)/m) ### loop for the weight of new vector portmanteau test
clag<-cov[l+1,,]
tclag<-t(clag)  ### trace of the covariance matrix
mclag<- tclag%*% invclag0%*% clag%*% invclag0
tracelag[l]<- tr(mclag)/(n-l)
tracelagw[l]<- w[l]*tr(mclag)/(n-l)
Qvho[[i]]<-(n*n)*sum(tracelag)  ### vector Hosking test
Qve[[i]]<-(n*n)*sum(tracelagw) ### vector new test
Qmm[[i]]<-gvtest(res,lag=m,order=0,SquaredQ=FALSE,Kernel=FALSE)      ##      vector Mahdi and McLeod (2011)
matrixcolumn[[i]]<-Qmm[[i]][,2,drop=F]
MLtest[[i]]<- matrixcolumn [[i]]< mmchi
hostest[[i]]<-Qvho[[i]]<hchi
newtest[[i]]<-Qve[[i]]<gamma
```

```
}
}
length(hostest[hostest==FALSE])/ns  ### account number of rejection
length(MLtest [MLtest ==FALSE])/ns
length(newtest[newtest==FALSE])/ns
```

Appendix D

Simulation studies of data generated under an AR(2) process tested by the DF tests and the ADF tests.

ϕ_1 and ϕ_2 are given from 0.1 to 0.9 (subject to the stationarity condition of an AR(2) process, i.e. $\phi_1 + \phi_2 < 1$), and n = 25, 50, 75, 100, 250, 500, 750, and 1000.

	df							
	25	50	75	100	250	500	750	1000
0.1,0.1	6	0	0	0	0	0	0	0
0.1,0.2	23	0	0	0	0	0	0	0
0.1,0.3	36	0	0	0	0	0	0	0
0.1,0.4	96	0	0	0	0	0	0	0
0.1,0.5	183	0	0	0	0	0	0	0
0.1,0.6	407	1	0	0	0	0	0	0
0.1,0.7	862	27	0	0	0	0	0	0
0.1,0.8	2073	332	29	1	0	0	0	0
0.1,0.9	2941	2285	2019	1896	1614	1465	1389	1439

	df drift							
	25	50	75	100	250	500	750	1000
0.1,0.1	407	0	0	0	0	0	0	0
0.1,0.2	515	0	0	0	0	0	0	0
0.1,0.3	650	1	0	0	0	0	0	0
0.1,0.4	842	5	0	0	0	0	0	0
0.1,0.5	1014	11	0	0	0	0	0	0
0.1,0.6	1213	58	2	0	0	0	0	0
0.1,0.7	1359	154	5	0	0	0	0	0
0.1,0.8	1498	399	74	26	0	0	0	0
0.1,0.9	1766	1041	741	609	463	331	319	317

	df trend							
	25	50	75	100	250	500	750	1000
0.1,0.1	1359	0	0	0	0	0	0	0
0.1,0.2	1323	6	0	0	0	0	0	0
0.1,0.3	1387	12	0	0	0	0	0	0
0.1,0.4	1427	18	0	0	0	0	0	0
0.1,0.5	1406	50	1	0	0	0	0	0
0.1,0.6	1279	99	5	0	0	0	0	0
0.1,0.7	1097	184	19	0	0	0	0	0
0.1,0.8	874	234	63	26	0	0	0	0
0.1,0.9	644	208	92	56	16	10	5	5

	df							
	25	50	75	100	250	500	750	1000
0.2,0.1	31	0	0	0	0	0	0	0
0.2,0.2	60	0	0	0	0	0	0	0
0.2,0.3	151	0	0	0	0	0	0	0
0.2,0.4	342	0	0	0	0	0	0	0
0.2,0.5	705	4	0	0	0	0	0	0
0.2,0.6	1404	83	1	0	0	0	0	0
0.2,0.7	2619	738	137	15	0	0	0	0
0.2,0.8	4283	3885	3721	3625	3584	3440	3394	3446
0.2,0.9	6158	7990	9286	9753	9999	10000	10000	10000

	df drift							
	25	50	75	100	250	500	750	1000
0.2,0.1	907	2	0	0	0	0	0	0
0.2,0.2	1192	6	0	0	0	0	0	0
0.2,0.3	1470	15	0	0	0	0	0	0
0.2,0.4	1755	50	0	0	0	0	0	0
0.2,0.5	2023	159	0	0	0	0	0	0
0.2,0.6	2262	410	44	2	0	0	0	0
0.2,0.7	2462	928	331	109	0	0	0	0
0.2,0.8	3088	2430	2096	2052	1779	1677	1676	1694
0.2,0.9	5113	7672	9269	9763	9999	10000	10000	10000

	df trend							
	25	50	75	100	250	500	750	1000
0.2,0.1	2346	12	0	0	0	0	0	0
0.2,0.2	2405	33	0	0	0	0	0	0
0.2,0.3	2404	88	0	0	0	0	0	0
0.2,0.4	2388	158	3	0	0	0	0	0
0.2,0.5	2324	298	15	0	0	0	0	0
0.2,0.6	2182	499	87	10	0	0	0	0
0.2,0.7	1931	739	288	127	0	0	0	0
0.2,0.8	1475	764	498	435	290	268	212	213
0.2,0.9	998	4200	8411	9606	9999	10000	10000	10000

	df							
	25	50	75	100	250	500	750	1000
0.3,0.1	101	0	0	0	0	0	0	0
0.3,0.2	253	0	0	0	0	0	0	0
0.3,0.3	570	1	0	0	0	0	0	0
0.3,0.4	1119	20	0	0	0	0	0	0
0.3,0.5	2119	196	8	0	0	0	0	0
0.3,0.6	3576	1350	359	64	0	0	0	0
0.3,0.7	5406	5155	5110	4987	4933	4937	4894	4871
0.3,0.8	7121	8664	9542	9843	10000	10000	10000	10000
0.3,0.9	8431	9783	9980	9998	10000	10000	10000	10000

	df drift							
	25	50	75	100	250	500	750	1000
0.3,0.1	1963	11	0	0	0	0	0	0
0.3,0.2	2344	34	0	0	0	0	0	0
0.3,0.3	2759	131	0	0	0	0	0	0
0.3,0.4	3103	356	11	0	0	0	0	0
0.3,0.5	3427	922	141	10	0	0	0	0
0.3,0.6	3738	1908	880	389	0	0	0	0
0.3,0.7	4498	3923	3679	3577	3374	3258	3235	3290
0.3,0.8	6449	8466	9546	9863	10000	10000	10000	10000
0.3,0.9	8247	9799	9985	9998	10000	10000	10000	10000

	df trend							
	25	50	75	100	250	500	750	1000
0.3,0.1	3660	25	0	0	0	0	0	0
0.3,0.2	3690	208	0	0	0	0	0	0
0.3,0.3	3675	409	8	0	0	0	0	0
0.3,0.4	3608	782	64	0	0	0	0	0
0.3,0.5	3442	1239	292	68	0	0	0	0
0.3,0.6	3224	1752	925	481	1	0	0	0
0.3,0.7	2687	1747	1463	1340	1116	1074	1004	972
0.3,0.8	2021	6116	9097	9773	10000	10000	10000	10000
0.3,0.9	5367	9651	9976	9998	10000	10000	10000	10000

	df							
	25	50	75	100	250	500	750	1000
0.4,0.1	422	1	0	0	0	0	0	0
0.4,0.2	836	6	0	0	0	0	0	0
0.4,0.3	1649	44	0	0	0	0	0	0
0.4,0.4	2852	378	15	1	0	0	0	0
0.4,0.5	4523	2119	733	173	0	0	0	0
0.4,0.6	6309	6127	6186	6131	6049	6083	6027	6042
0.4,0.7	7861	9077	9708	9909	10000	10000	10000	10000
0.4,0.8	8876	9853	9987	9998	10000	10000	10000	10000
0.4,0.9	9451	9976	10000	10000	10000	10000	10000	10000

	df drift							
	25	50	75	100	250	500	750	1000
0.4,0.1	3420	86	0	0	0	0	0	0
0.4,0.2	3871	280	0	0	0	0	0	0
0.4,0.3	4280	753	37	0	0	0	0	0
0.4,0.4	4684	1774	356	44	0	0	0	0
0.4,0.5	4997	3146	1827	937	0	0	0	0
0.4,0.6	5727	5240	5062	5012	4809	4809	4737	4754
0.4,0.7	7387	9031	9716	9923	10000	10000	10000	10000
0.4,0.8	8820	9867	9991	9999	10000	10000	10000	10000
0.4,0.9	9455	9978	10000	10000	10000	10000	10000	10000

	df trend							
	25	50	75	100	250	500	750	1000
0.4,0.1	5136	478	3	0	0	0	0	0
0.4,0.2	5113	928	3	0	0	0	0	0
0.4,0.3	5026	1545	191	4	0	0	0	0
0.4,0.4	4872	2288	737	198	0	0	0	0
0.4,0.5	4617	3041	1959	1234	7	0	0	0
0.4,0.6	4048	3107	2844	2683	2473	2402	2326	2296
0.4,0.7	3328	7472	9458	9881	10000	10000	10000	10000
0.4,0.8	6868	9786	9987	9999	10000	10000	10000	10000
0.4,0.9	8938	9972	10000	10000	10000	10000	10000	10000

	df							
	25	50	75	100	250	500	750	1000
0.5,0.1	1227	12	0	0	0	0	0	0
0.5,0.2	2271	91	1	0	0	0	0	0
0.5,0.3	3685	651	32	1	0	0	0	0
0.5,0.4	5402	2997	1237	394	0	0	0	0
0.5,0.5	7028	6938	7033	6982	6923	6989	6932	6987
0.5,0.6	8355	9346	9821	9938	10000	10000	10000	10000
0.5,0.7	9188	9901	9993	9999	10000	10000	10000	10000
0.5,0.8	9603	9983	10000	10000	10000	10000	10000	10000
0.5,0.9	9803	9998	10000	10000	10000	10000	10000	10000

	df drift							
	25	50	75	100	250	500	750	1000
0.5,0.1	5040	540	2	0	0	0	0	0
0.5,0.2	5491	1408	107	4	0	0	0	0
0.5,0.3	5879	2774	752	141	0	0	0	0
0.5,0.4	6166	4487	3015	1798	4	0	0	0
0.5,0.5	6746	6418	6297	6256	6107	6115	6104	6083
0.5,0.6	8097	9352	9828	9952	10000	10000	10000	10000
0.5,0.7	9213	9919	9994	9999	10000	10000	10000	10000
0.5,0.8	9649	9986	10000	10000	10000	10000	10000	10000
0.5,0.9	9824	9998	10000	10000	10000	10000	10000	10000

	df trend							
	25	50	75	100	250	500	750	1000
0.5,0.1	6442	1722	107	3	0	0	0	0
0.5,0.2	6338	2644	498	50	0	0	0	0
0.5,0.3	6181	3678	1603	498	0	0	0	0
0.5,0.4	5942	4539	3319	2357	36	0	0	0
0.5,0.5	5390	4618	4336	4273	4025	4042	3941	3951
0.5,0.6	4774	8411	9703	9925	10000	10000	10000	10000
0.5,0.7	7912	9865	9993	9999	10000	10000	10000	10000
0.5,0.8	9319	9984	10000	10000	10000	10000	10000	10000
0.5,0.9	9740	9998	10000	10000	10000	10000	10000	10000

df	25	50	75	100	250	500	750	1000
0.6,0.1	2972	158	2	0	0	0	0	0
0.6,0.2	4568	997	79	3	0	0	0	0
0.6,0.3	6218	3887	1849	700	0	0	0	0
0.6,0.4	7668	7642	7688	7673	7687	7727	7723	7746
0.6,0.5	8777	9549	9874	9980	10000	10000	10000	10000
0.6,0.6	9419	9932	9998	10000	10000	10000	10000	10000
0.6,0.7	9719	9994	10000	10000	10000	10000	10000	10000
0.6,0.8	9873	9999	10000	10000	10000	10000	10000	10000
0.6,0.9	9944	9999	10000	10000	10000	10000	10000	10000

df drift	25	50	75	100	250	500	750	1000
0.6,0.1	6625	2237	242	9	0	0	0	0
0.6,0.2	6935	3970	1417	326	0	0	0	0
0.6,0.3	7169	5755	4278	2895	20	0	0	0
0.6,0.4	7624	7418	7338	7266	7220	7238	7195	7246
0.6,0.5	8677	9587	9893	9986	10000	10000	10000	10000
0.6,0.6	9466	9945	9998	10000	10000	10000	10000	10000
0.6,0.7	9758	9996	10000	10000	10000	10000	10000	10000
0.6,0.8	9898	9999	10000	10000	10000	10000	10000	10000
0.6,0.9	9949	9999	10000	10000	10000	10000	10000	10000

df trend	25	50	75	100	250	500	750	1000
0.6,0.1	7485	4004	1058	133	0	0	0	0
0.6,0.2	7332	5153	2726	1121	0	0	0	0
0.6,0.3	7091	6021	4760	3703	129	0	0	0
0.6,0.4	6547	6087	5779	5816	5646	5678	5609	5660
0.6,0.5	6109	9040	9830	9970	10000	10000	10000	10000
0.6,0.6	8676	9931	9998	10000	10000	10000	10000	10000
0.6,0.7	9580	9994	10000	10000	10000	10000	10000	10000
0.6,0.8	9863	9999	10000	10000	10000	10000	10000	10000
0.6,0.9	9924	9999	10000	10000	10000	10000	10000	10000

df	25	50	75	100	250	500	750	1000
0.7,0.1	5396	1455	141	7	0	0	0	0
0.7,0.2	6967	4738	2578	1099	0	0	0	0
0.7,0.3	8227	8204	8285	8294	8322	8316	8335	8377
0.7,0.4	9108	9712	9912	9983	10000	10000	10000	10000
0.7,0.5	9600	9968	9997	10000	10000	10000	10000	10000
0.7,0.6	9812	9996	10000	10000	10000	10000	10000	10000
0.7,0.7	9914	10000	10000	10000	10000	10000	10000	10000
0.7,0.8	9962	10000	10000	10000	10000	10000	10000	10000
0.7,0.9	9987	10000	10000	10000	10000	10000	10000	10000

df drift	25	50	75	100	250	500	750	1000
0.7,0.1	7793	5211	2341	648	0	0	0	0
0.7,0.2	7956	6894	5539	4161	51	0	0	0
0.7,0.3	8278	8169	8087	8105	8048	8053	8063	8108
0.7,0.4	9073	9777	9927	9991	10000	10000	10000	10000
0.7,0.5	9651	9976	9997	10000	10000	10000	10000	10000
0.7,0.6	9859	9997	10000	10000	10000	10000	10000	10000
0.7,0.7	9930	10000	10000	10000	10000	10000	10000	10000
0.7,0.8	9973	10000	10000	10000	10000	10000	10000	10000
0.7,0.9	9992	10000	10000	10000	10000	10000	10000	10000

df trend	25	50	75	100	250	500	750	1000
0.7,0.1	8170	6557	4132	2054	0	0	0	0
0.7,0.2	8000	7341	6245	5185	339	0	0	0
0.7,0.3	7544	7325	7115	7210	7078	7113	7064	7099
0.7,0.4	7271	9482	9899	9990	10000	10000	10000	10000
0.7,0.5	9200	9966	9997	10000	10000	10000	10000	10000
0.7,0.6	9749	9996	10000	10000	10000	10000	10000	10000
0.7,0.7	9901	10000	10000	10000	10000	10000	10000	10000
0.7,0.8	9968	10000	10000	10000	10000	10000	10000	10000
0.7,0.9	9986	10000	10000	10000	10000	10000	10000	10000

df	25	50	75	100	250	500	750	1000
0.8,0.1	7626	5641	3407	1615	1	0	0	0
0.8,0.2	8674	8713	8757	8793	8809	8797	8826	8886
0.8,0.3	9349	9800	9960	9992	10000	10000	10000	10000
0.8,0.4	9717	9981	9998	10000	10000	10000	10000	10000
0.8,0.5	9859	9998	10000	10000	10000	10000	10000	10000
0.8,0.6	9945	10000	10000	10000	10000	10000	10000	10000
0.8,0.7	9968	10000	10000	10000	10000	10000	10000	10000
0.8,0.8	9997	10000	10000	10000	10000	10000	10000	10000
0.8,0.9	9994	10000	10000	10000	10000	10000	10000	10000

df drift	25	50	75	100	250	500	750	1000
0.8,0.1	8581	7877	6712	5435	137	0	0	0
0.8,0.2	8738	8740	8695	8718	8709	8689	8723	8749
0.8,0.3	9374	9841	9967	9995	10000	10000	10000	10000
0.8,0.4	9780	9986	9999	10000	10000	10000	10000	10000
0.8,0.5	9900	9998	10000	10000	10000	10000	10000	10000
0.8,0.6	9952	10000	10000	10000	10000	10000	10000	10000
0.8,0.7	9979	10000	10000	10000	10000	10000	10000	10000
0.8,0.8	9998	10000	10000	10000	10000	10000	10000	10000
0.8,0.9	9994	10000	10000	10000	10000	10000	10000	10000

df trend	25	50	75	100	250	500	750	1000
0.8,0.1	8630	8317	7578	6706	812	0	0	0
0.8,0.2	8241	8247	8165	8285	8186	8225	8179	8221
0.8,0.3	8184	9743	9953	9993	10000	10000	10000	10000
0.8,0.4	9533	9981	9999	10000	10000	10000	10000	10000
0.8,0.5	9841	9997	10000	10000	10000	10000	10000	10000
0.8,0.6	9939	10000	10000	10000	10000	10000	10000	10000
0.8,0.7	9975	10000	10000	10000	10000	10000	10000	10000
0.8,0.8	9998	10000	10000	10000	10000	10000	10000	10000
0.8,0.9	9994	10000	10000	10000	10000	10000	10000	10000

df	25	50	75	100	250	500	750	1000
0.9,0.1	9036	9123	9127	9172	9210	9190	9189	9266
0.9,0.2	9546	9897	9974	9996	10000	10000	10000	10000
0.9,0.3	9810	9989	10000	10000	10000	10000	10000	10000
0.9,0.4	9901	10000	10000	10000	10000	10000	10000	10000
0.9,0.5	9960	10000	10000	10000	10000	10000	10000	10000
0.9,0.6	9980	10000	10000	10000	10000	10000	10000	10000
0.9,0.7	9991	10000	10000	10000	10000	10000	10000	10000
0.9,0.8	9993	10000	10000	10000	10000	10000	10000	10000
0.9,0.9	9999	10000	10000	10000	10000	10000	10000	10000

df drift	25	50	75	100	250	500	750	1000
0.9,0.1	9041	9137	9153	9161	9186	9145	9171	9213
0.9,0.2	9612	9934	9987	9999	10000	10000	10000	10000
0.9,0.3	9864	9993	10000	10000	10000	10000	10000	10000
0.9,0.4	9931	10000	10000	10000	10000	10000	10000	10000
0.9,0.5	9970	10000	10000	10000	10000	10000	10000	10000
0.9,0.6	9986	10000	10000	10000	10000	10000	10000	10000
0.9,0.7	9992	10000	10000	10000	10000	10000	10000	10000
0.9,0.8	9995	10000	10000	10000	10000	10000	10000	10000
0.9,0.9	9998	10000	10000	10000	10000	10000	10000	10000

df trend	25	50	75	100	250	500	750	1000
0.9,0.1	8813	8918	8928	8974	8965	8981	8963	9015
0.9,0.2	8868	9878	9979	9999	10000	10000	10000	10000
0.9,0.3	9734	9991	10000	10000	10000	10000	10000	10000
0.9,0.4	9896	10000	10000	10000	10000	10000	10000	10000
0.9,0.5	9958	10000	10000	10000	10000	10000	10000	10000
0.9,0.6	9986	10000	10000	10000	10000	10000	10000	10000
0.9,0.7	9991	10000	10000	10000	10000	10000	10000	10000
0.9,0.8	9992	10000	10000	10000	10000	10000	10000	10000
0.9,0.9	9998	10000	10000	10000	10000	10000	10000	10000

	adf	lag12		statinary				
	25	50	75	100	250	500	750	1000
0.1,0.1	9073	7050	4247	2309	11	0	0	0
0.1,0.2	9098	7260	4588	2602	16	0	0	0
0.1,0.3	9132	7478	4944	2959	28	0	0	0
0.1,0.4	9161	7727	5422	3517	45	0	0	0
0.1,0.5	9181	8019	6021	4187	103	0	0	0
0.1,0.6	9216	8343	6775	5155	261	0	0	0
0.1,0.7	9269	8756	7675	6430	823	4	0	0
0.1,0.8	9376	9194	8719	8132	3465	181	1	0
0.1,0.9	9551	9631	9608	9611	9576	9522	9549	9585

	adf drift	lag12						
	25	50	75	100	250	500	750	1000
0.1,0.1	9553	9276	8125	6897	605	0	0	0
0.1,0.2	9550	9316	8304	7157	771	0	0	0
0.1,0.3	9525	9342	8454	7413	988	0	0	0
0.1,0.4	9519	9387	8629	7748	1322	0	0	0
0.1,0.5	9500	9452	8826	8115	1942	4	0	0
0.1,0.6	9475	9501	9039	8546	2964	34	0	0
0.1,0.7	9473	9552	9291	9003	4816	276	4	0
0.1,0.8	9488	9606	9505	9379	7631	2766	418	37
0.1,0.9	9497	9654	9608	9598	9549	9494	9491	9498

	adf drift tren	lag12						
	25	50	75	100	250	500	750	1000
0.1,0.1	9399	9591	9023	8422	2039	5	0	0
0.1,0.2	9382	9616	9110	8531	2343	8	0	0
0.1,0.3	9378	9622	9170	8672	2770	13	0	0
0.1,0.4	9389	9639	9233	8861	3398	31	0	0
0.1,0.5	9411	9654	9321	9026	4216	11	0	0
0.1,0.6	9418	9659	9434	9235	5405	316	0	0
0.1,0.7	9423	9664	9533	9426	6953	1312	68	0
0.1,0.8	9393	9677	9588	9595	8667	5122	1738	370
0.1,0.9	9410	9692	9641	9647	9560	9529	9525	9516

	adf	lag12						
	25	50	75	100	250	500	750	1000
0.2,0.1	9089	7186	4451	2487	14	0	0	0
0.2,0.2	9095	7419	4819	2830	24	0	0	0
0.2,0.3	9131	7668	5300	3363	35	0	0	0
0.2,0.4	9162	7944	5884	4006	79	0	0	0
0.2,0.5	9195	8286	6656	4978	221	0	0	0
0.2,0.6	9266	8717	7578	6275	719	2	0	0
0.2,0.7	9376	9182	8652	8026	3198	139	1	0
0.2,0.8	9546	9634	9609	9609	9586	9528	9553	9584
0.2,0.9	9741	9922	9967	9985	9999	10000	10000	10000

	adf drift	lag12						
	25	50	75	100	250	500	750	1000
0.2,0.1	9541	9284	8232	7048	695	0	0	0
0.2,0.2	9531	9328	8397	7326	905	0	0	0
0.2,0.3	9517	9368	8583	7649	1227	0	0	0
0.2,0.4	9500	9439	8788	8042	1778	3	0	0
0.2,0.5	9479	9490	8999	8487	2742	29	0	0
0.2,0.6	9475	9549	9252	8947	4566	210	4	0
0.2,0.7	9476	9599	9492	9385	7491	2482	300	21
0.2,0.8	9504	9653	9605	9600	9546	9496	9492	9501
0.2,0.9	9641	9942	9985	9994	10000	10000	10000	10000

	adf trend	lag12						
	25	50	75	100	250	500	750	1000
0.2,0.1	9370	9611	9080	8486	2216	7	0	0
0.2,0.2	9367	9620	9150	8636	2613	10	0	0
0.2,0.3	9366	9625	9233	8806	3212	21	0	0
0.2,0.4	9400	9644	9295	8987	4010	82	0	0
0.2,0.5	9407	9654	9406	9188	5189	255	0	0
0.2,0.6	9417	9653	9525	9405	6776	1135	50	0
0.2,0.7	9409	9677	9591	9584	8583	4796	1496	271
0.2,0.8	9403	9683	9644	9640	9552	9529	9526	9512
0.2,0.9	9438	9864	9977	9997	10000	10000	10000	10000

	adf	lag12						
	25	50	75	100	250	500	750	1000
0.3,0.1	9081	7337	4687	2715	19	0	0	0
0.3,0.2	9133	7614	5150	3186	31	0	0	0
0.3,0.3	9138	7890	5749	3831	63	0	0	0
0.3,0.4	9193	8232	6505	4754	183	0	0	0
0.3,0.5	9255	8679	7449	6089	610	2	0	0
0.3,0.6	9358	9151	8596	7906	2904	112	1	0
0.3,0.7	9541	9637	9604	9609	9574	9522	9550	9578
0.3,0.8	9740	9947	9978	10000	10000	10000	10000	10000
0.3,0.9	9867	9989	10000	10000	10000	10000	10000	10000

	adf drift	lag12						
	25	50	75	100	250	500	750	1000
0.3,0.1	9506	9305	8344	7229	812	0	0	0
0.3,0.2	9497	9348	8517	7549	1120	0	0	0
0.3,0.3	9481	9420	8724	7967	1607	1	0	0
0.3,0.4	9485	9486	8950	8408	2529	19	0	0
0.3,0.5	9477	9541	9207	8893	4310	161	4	0
0.3,0.6	9474	9588	9478	9363	7299	2164	218	14
0.3,0.7	9495	9656	9597	9600	9552	9498	9493	9502
0.3,0.8	9660	9954	9989	10000	10000	10000	10000	10000
0.3,0.9	9878	9997	10000	10000	10000	10000	10000	10000

	adf trend	lag12						
	25	50	75	100	250	500	750	1000
0.3,0.1	9344	9605	9126	8569	2472	9	0	0
0.3,0.2	9345	9623	9205	8737	3027	13	0	0
0.3,0.3	9367	9633	9286	8938	3829	53	0	0
0.3,0.4	9380	9651	9390	9153	4952	218	0	0
0.3,0.5	9407	9654	9515	9386	6577	961	35	0
0.3,0.6	9405	9670	9588	9570	8489	4429	1249	197
0.3,0.7	9415	9682	9635	9632	9554	9533	9528	9512
0.3,0.8	9458	9886	9970	10000	10000	10000	10000	10000
0.3,0.9	9692	9993	10000	10000	10000	10000	10000	10000

	adf	lag12						
	25	50	75	100	250	500	750	1000
0.4,0.1	9077	7522	5009	3020	28	0	0	0
0.4,0.2	9122	7847	5581	3639	52	0	0	0
0.4,0.3	9165	8174	6362	4546	152	0	0	0
0.4,0.4	9243	8630	7331	5924	511	1	0	0
0.4,0.5	9355	9126	8522	7777	2632	50	1	0
0.4,0.6	9540	9640	9615	9608	9579	9522	9544	9580
0.4,0.7	9746	9942	9985	9993	10000	10000	10000	10000
0.4,0.8	9879	9990	9999	10000	10000	10000	10000	10000
0.4,0.9	9952	9999	10000	10000	10000	10000	10000	10000

	adf drift	lag12						
	25	50	75	100	250	500	750	1000
0.4,0.1	9476	9337	8464	7443	1014	0	0	0
0.4,0.2	9466	9385	8664	7840	1441	0	0	0
0.4,0.3	9469	9461	8917	8319	2310	10	0	0
0.4,0.4	9475	9538	9182	8827	4036	130	0	0
0.4,0.5	9463	9588	9452	9342	7104	1832	147	8
0.4,0.6	9488	9655	9605	9603	9550	9496	9494	9506
0.4,0.7	9669	9962	9990	9998	10000	10000	10000	10000
0.4,0.8	9900	9996	9999	10000	10000	10000	10000	10000
0.4,0.9	9959	10000	10000	10000	10000	10000	10000	10000

	adf trend	lag12						
	25	50	75	100	250	500	750	1000
0.4,0.1	9314	9611	9180	8689	2835	11	0	0
0.4,0.2	9331	9628	9266	8892	3587	41	0	0
0.4,0.3	9342	9638	9361	9114	4729	180	0	0
0.4,0.4	9371	9656	9507	9353	6347	785	21	1
0.4,0.5	9389	9666	9584	9556	8360	4078	1015	134
0.4,0.6	9409	9666	9630	9636	9551	9539	9526	9510
0.4,0.7	9439	9901	9982	9994	10000	10000	10000	10000
0.4,0.8	9738	9993	9999	10000	10000	10000	10000	10000
0.4,0.9	9924	10000	10000	10000	10000	10000	10000	10000

	adf	lag12						
	25	50	75	100	250	500	750	1000
0.5,0.1	9116	7805	5417	3478	44	0	0	0
0.5,0.2	9147	8113	6208	4324	121	0	0	0
0.5,0.3	9213	8570	7202	5725	434	0	0	0
0.5,0.4	9335	9100	8434	7640	2323	48	1	0
0.5,0.5	9528	9652	9624	9599	9573	9530	9548	9576
0.5,0.6	9755	9950	9991	9991	10000	10000	10000	10000
0.5,0.7	9884	9993	9997	9999	10000	10000	10000	10000
0.5,0.8	9946	9999	10000	10000	10000	10000	10000	10000
0.5,0.9	9978	10000	10000	10000	10000	10000	10000	10000

	adf drift	lag12						
	25	50	75	100	250	500	750	1000
0.5,0.1	9448	9374	8616	7738	1299	0	0	0
0.5,0.2	9458	9447	8867	8213	2094	5	0	0
0.5,0.3	9476	9531	9159	8773	3716	90	2	0
0.5,0.4	9466	9584	9441	9309	6862	1527	93	6
0.5,0.5	9475	9658	9613	9599	9555	9494	9493	9505
0.5,0.6	9679	9961	9991	9999	10000	10000	10000	10000
0.5,0.7	9911	9998	9999	10000	10000	10000	10000	10000
0.5,0.8	9966	10000	10000	10000	10000	10000	10000	10000
0.5,0.9	9986	10000	10000	10000	10000	10000	10000	10000

	adf trend	lag12						
	25	50	75	100	250	500	750	1000
0.5,0.1	9283	9620	9254	8834	3357	25	0	0
0.5,0.2	9304	9632	9344	9074	4462	136	0	0
0.5,0.3	9334	9644	9492	9320	6098	627	10	0
0.5,0.4	9356	9662	9575	9544	8247	3696	758	83
0.5,0.5	9395	9669	9632	9639	9552	9540	9526	9513
0.5,0.6	9419	9923	9987	9999	10000	10000	10000	10000
0.5,0.7	9763	9997	9999	9999	10000	10000	10000	10000
0.5,0.8	9930	10000	10000	10000	10000	10000	10000	10000
0.5,0.9	9974	10000	10000	10000	10000	10000	10000	10000

	adf	lag12						
	25	50	75	100	250	500	750	1000
0.6,0.1	9127	8056	6014	4127	90	0	0	0
0.6,0.2	9186	8492	7032	5527	360	0	0	0
0.6,0.3	9311	9067	8345	7478	2074	31	0	0
0.6,0.4	9525	9649	9617	9597	9578	9529	9552	9579
0.6,0.5	9761	9955	9980	9997	10000	10000	10000	10000
0.6,0.6	9906	9994	9999	10000	10000	10000	10000	10000
0.6,0.7	9951	9998	10000	10000	10000	10000	10000	10000
0.6,0.8	9978	10000	10000	10000	10000	10000	10000	10000
0.6,0.9	9986	10000	10000	10000	10000	10000	10000	10000

	adf drift	lag12						
	25	50	75	100	250	500	750	1000
0.6,0.1	9432	9429	8801	8089	1895	3	0	0
0.6,0.2	9455	9516	9116	8690	3429	66	1	0
0.6,0.3	9452	9572	9420	9277	6590	1267	53	3
0.6,0.4	9471	9656	9611	9604	9555	9492	9494	9504
0.6,0.5	9714	9967	9994	9999	10000	10000	10000	10000
0.6,0.6	9930	10000	10000	10000	10000	10000	10000	10000
0.6,0.7	9979	10000	10000	10000	10000	10000	10000	10000
0.6,0.8	9985	10000	10000	10000	10000	10000	10000	10000
0.6,0.9	9991	10000	10000	10000	10000	10000	10000	10000

	adf trend	lag12						
	25	50	75	100	250	500	750	1000
0.6,0.1	9272	9634	9321	9006	4204	100	2	0
0.6,0.2	9332	9645	9457	9287	5823	483	6	0
0.6,0.3	9328	9666	9565	9531	8084	3289	558	52
0.6,0.4	9384	9669	9623	9636	9554	9546	9527	9516
0.6,0.5	9432	9933	9992	10000	10000	10000	10000	10000
0.6,0.6	9797	9997	10000	10000	10000	10000	10000	10000
0.6,0.7	9957	10000	10000	10000	10000	10000	10000	10000
0.6,0.8	9979	10000	10000	10000	10000	10000	10000	10000
0.6,0.9	9992	10000	10000	10000	10000	10000	10000	10000

	adf	lag12						
	25	50	75	100	250	500	750	1000
0.7,0.1	9165	8436	6870	5265	294	0	0	0
0.7,0.2	9285	9033	8244	7280	1734	23	0	0
0.7,0.3	9515	9654	9619	9594	9578	9535	9548	9583
0.7,0.4	9774	9968	9987	9997	10000	10000	10000	10000
0.7,0.5	9921	9996	10000	10000	10000	10000	10000	10000
0.7,0.6	9959	10000	10000	10000	10000	10000	10000	10000
0.7,0.7	9979	10000	10000	10000	10000	10000	10000	10000
0.7,0.8	9987	10000	10000	10000	10000	10000	10000	10000
0.7,0.9	9990	10000	10000	10000	10000	10000	10000	10000

	adf drift	lag12						
	25	50	75	100	250	500	750	1000
0.7,0.1	9423	9516	9079	8582	3097	42	1	0
0.7,0.2	9417	9578	9402	9239	6281	987	33	2
0.7,0.3	9461	9648	9616	9602	9553	9498	9493	9500
0.7,0.4	9713	9981	9993	9999	10000	10000	10000	10000
0.7,0.5	9949	9999	10000	10000	10000	10000	10000	10000
0.7,0.6	9981	10000	10000	10000	10000	10000	10000	10000
0.7,0.7	9989	10000	10000	10000	10000	10000	10000	10000
0.7,0.8	9993	10000	10000	10000	10000	10000	10000	10000
0.7,0.9	9997	10000	10000	10000	10000	10000	10000	10000

	adf trend	lag12						
	25	50	75	100	250	500	750	1000
0.7,0.1	9293	9647	9432	9244	5514	362	4	1
0.7,0.2	9316	9660	9564	9514	7925	2912	373	31
0.7,0.3	9359	9663	9625	9637	9546	9544	9523	9516
0.7,0.4	9412	9950	9991	9999	10000	10000	10000	10000
0.7,0.5	9827	9997	10000	10000	10000	10000	10000	10000
0.7,0.6	9962	9999	10000	10000	10000	10000	10000	10000
0.7,0.7	9983	10000	10000	10000	10000	10000	10000	10000
0.7,0.8	9990	10000	10000	10000	10000	10000	10000	10000
0.7,0.9	9998	10000	10000	10000	10000	10000	10000	10000

	adf	lag12						
	25	50	75	100	250	500	750	1000
0.8,0.1	9238	8983	8124	7095	1427	17	0	0
0.8,0.2	9496	9651	9611	9593	9578	9540	9548	9583
0.8,0.3	9778	9958	9988	9997	10000	10000	10000	10000
0.8,0.4	9912	9997	10000	10000	10000	10000	10000	10000
0.8,0.5	9960	9999	10000	10000	10000	10000	10000	10000
0.8,0.6	9975	10000	10000	10000	10000	10000	10000	10000
0.8,0.7	9987	10000	10000	10000	10000	10000	10000	10000
0.8,0.8	9990	10000	10000	10000	10000	10000	10000	10000
0.8,0.9	9988	10000	10000	10000	10000	10000	10000	10000

	adf drift	lag12						
	25	50	75	100	250	500	750	1000
0.8,0.1	9401	9573	9387	9200	5936	743	19	1
0.8,0.2	9438	9640	9610	9604	9557	9501	9495	9505
0.8,0.3	9736	9982	9996	10000	10000	10000	10000	10000
0.8,0.4	9963	9999	10000	10000	10000	10000	10000	10000
0.8,0.5	9984	10000	10000	10000	10000	10000	10000	10000
0.8,0.6	9989	10000	10000	10000	10000	10000	10000	10000
0.8,0.7	9995	10000	10000	10000	10000	10000	10000	10000
0.8,0.8	9996	10000	10000	10000	10000	10000	10000	10000
0.8,0.9	9996	10000	10000	10000	10000	10000	10000	10000

	adf trend	lag12						
	25	50	75	100	250	500	750	1000
0.8,0.1	9284	9648	9554	9494	7704	2465	277	13
0.8,0.2	9332	9671	9615	9629	9550	9545	9524	9516
0.8,0.3	9444	9966	9997	9999	10000	10000	10000	10000
0.8,0.4	9862	10000	10000	10000	10000	10000	10000	10000
0.8,0.5	9960	10000	10000	10000	10000	10000	10000	10000
0.8,0.6	9989	10000	10000	10000	10000	10000	10000	10000
0.8,0.7	9996	10000	10000	10000	10000	10000	10000	10000
0.8,0.8	9997	10000	10000	10000	10000	10000	10000	10000
0.8,0.9	9996	10000	10000	10000	10000	10000	10000	10000

	adf	lag12						
	25	50	75	100	250	500	750	1000
0.9,0.1	9475	9651	9618	9596	9577	9541	9543	9583
0.9,0.2	9800	9976	9991	10000	10000	10000	10000	10000
0.9,0.3	9920	9996	10000	10000	10000	10000	10000	10000
0.9,0.4	9962	10000	10000	10000	10000	10000	10000	10000
0.9,0.5	9980	10000	10000	10000	10000	10000	10000	10000
0.9,0.6	9986	10000	10000	10000	10000	10000	10000	10000
0.9,0.7	9991	10000	10000	10000	10000	10000	10000	10000
0.9,0.8	9988	10000	10000	10000	10000	10000	10000	10000
0.9,0.9	9991	10000	10000	10000	10000	10000	10000	10000

	adf drift	lag12						
	25	50	75	100	250	500	750	1000
0.9,0.1	9411	9634	9609	9597	9546	9495	9493	9501
0.9,0.2	9767	9993	9999	10000	10000	10000	10000	10000
0.9,0.3	9960	9999	10000	10000	10000	10000	10000	10000
0.9,0.4	9980	10000	10000	10000	10000	10000	10000	10000
0.9,0.5	9990	10000	10000	10000	10000	10000	10000	10000
0.9,0.6	9995	10000	10000	10000	10000	10000	10000	10000
0.9,0.7	9996	10000	10000	10000	10000	10000	10000	10000
0.9,0.8	9995	10000	10000	10000	10000	10000	10000	10000
0.9,0.9	9996	10000	10000	10000	10000	10000	10000	10000

	adf trend	lag12						
	25	50	75	100	250	500	750	1000
0.9,0.1	9312	9660	9625	9634	9552	9546	9527	9513
0.9,0.2	9444	9979	9997	10000	10000	10000	10000	10000
0.9,0.3	9871	9999	10000	10000	10000	10000	10000	10000
0.9,0.4	9967	10000	10000	10000	10000	10000	10000	10000
0.9,0.5	9992	10000	10000	10000	10000	10000	10000	10000
0.9,0.6	9996	10000	10000	10000	10000	10000	10000	10000
0.9,0.7	9995	10000	10000	10000	10000	10000	10000	10000
0.9,0.8	9994	10000	10000	10000	10000	10000	10000	10000
0.9,0.9	9997	10000	10000	10000	10000	10000	10000	10000

Appendix E

Dickeys' method for calculating the critical values of the DF test and the ADF test.

For example, for the DF test (similar procedure for the ADF test), (Dickey, 1976)

1- Generate n points of data from a Normal distribution (e_t white noise).
2- Use the e_t values to generate observations from an AR(1) process with parameter ϕ.
3- Estimate all parameters by least squared method using the data you have, and compute the t-test statistic.
4- Fix all estimated parameters except τ which you set to zero.
5- Using this sample re-estimate τ and then the t-test statistic using least squared method, which is a random number drawn from the sampling distribution under the null, say t_1.
6- Repeat steps (3) and (4) above 10,000 times, and produce a set t_1, $t_2\cdots$.
7- Percentiles of this distribution give the critical values.