Journal of the Korean
Statistical Society
Vol. 22, No. 1, 1993

# Test of Homogeneity for a Panel of Seasonal Autoregressive Processes†

Sung-Duck Lee [1]

**ABSTRACT**

Large sample test of homogeneity for a panel of more than two seasonal autoregressive processes is derived and its limiting distribution is found. Detailed results are shown for the important special case that the seasonal and nonseasonal autoregressive components are both of order one.

**KEYWORDS:** Seasonal autoregressive processes, Wald statistic, m-dependent processes, Limiting distribution

## 1. INTRODUCTION

It is often of interest to test whether or not two or more sequences of observations can be considered as coming from a common time series model. In practice, it is important that we employ the smallest possible number of parameters for adequate representation. If the hypothesis of homogeneity is

[1] Department of Statistics, Chungbuk National University, Cheongju, 360-763, Korea.
† This research was supported by Korea Research Foundation, 1990.

accepted, one can achieve parsimony in model specification and also obtain better estimates of the model parameters by pooling the data sets. For example, one has a panel of short time series, e.g., annual observations on some economic indicator from various regions of a country, or blood pressures of a number of patients on several days. Anderson(1978) studied statistical inference for a panel of autoregressive processes and a study of homogeneity tests for nonseasonal time series models was performed in Basawa,Billard and Srinivasan(1984). In this paper, we consider testing the homogeneity of a panel with $k$ multiplicative seasonal autoregressive processes which contain both seasonal and nonseasonal components. A multiplicative seasonal autoregressive process of order $p$ and $P$, SAR$(p, P)$, is given by

$$\phi(B)\Phi(B^s)X_t = e_t \tag{1.1}$$

where $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ and $\Phi(B^s) = 1 - \Phi_1 B^s - \cdots - \Phi_P B^{Ps}$ for the sequence of observations $X_t, t = 0, 1, 2, \cdots$ , and where $\{e_t\}$ is the error sequence, or white noise process. Since for general $(p, P)$ values, the detail becomes very messy algebraically, we derive our results for the particular case of the SAR$(1, 1)$ model. First, in Section 2, the asymptotic properties of the maximum likelihood estimates for the underlying parameters of SAR$(1, 1)$ model are discussed. In section 3, the *Wald statistic* as homogeneity test statistic is derived. Finally, we extend our results to the more general model in Section 4. A multiplicative seasonal first order autoregressive SAR$(1, 1)$ model has frequently proved useful in modeling a variety of real life time series data. Suppose we have $k$ independent SAR$(1, 1)$ models. Let $X_{tj}$ denote the observation at time $t$ from the $j^{th}$ multiplicative seasonal autoregressive process, SAR$(1, 1)$, that is,

$$(1 - \phi_j B)(1 - \Phi_j B^s)X_{t,j} = e_{t,j} \tag{1.2}$$

$t = 1, \cdots, n_j$, and $j = 1, \cdots, k$ ,and where $\{e_{t,j}\}$ is a sequence of independently and identically distributed normal random variables with mean zero and variance $\sigma^2$. Assume that $X_{0,j} = 0$ for each $j = 1, \cdots, k$. Let $n$ be the total number of observations from all samples, that is,$n = \sum_{j=1}^{k} n_j$ and assume that $n_j/n \to c_j$ as $n \to \infty$ where $0 < c_j < 1$ for all $j = 1, \cdots, k$. We assume that our model is stationary; so we have $|\phi_j| < 1$ and $|\Phi_j| < 1$. Since the process is stationary we can write

$$X_{t,j} = \{(1 - \phi_j B)(1 - \Phi_j B^s)\}^{-1} e_{t,j} = \sum_{r=0}^{\infty} \Psi_{r,j} e_{t-r,j} \tag{1.3}$$

with $\sum_{r=0}^{\infty} |\Psi_{r,j}| < \infty$. Hence, for the $j$th series, we can show that

$$\Psi_{r,j} = \phi_j^h[\phi_j^{(u+1)s} - \Phi_j^{(u+1)}]/(\phi_j^s - \Phi_j), \quad j = 1, \cdots, k, \tag{1.4}$$

where $r = us + h, \quad h = 0, 1, \cdots, s-1$, and $u = 0, 1, \cdots$. Hence $E(X_{t,j}) = 0$, and the autocovariance function at lag $h, \gamma_j(h)$, is given by

$$\gamma_j(h) = \sigma^2 \sum_{r=0}^{\infty} \Psi_{r,j}\Psi_{r+h,j}, \quad h = 0, 1, \cdots. \tag{1.5}$$

In particular,for each $j = 1, \cdots, k$, the variance of the SAR(1,1) model is

$$\gamma_j(0) = \sigma^2 \sum_{r=0}^{\infty} \Psi_{r,j}^2 = \sigma^2(1 + \phi_j^s\Phi_j)/\{(1 - \phi_j^2)(1 - \Phi_j^2)(1 - \phi_j^s\Phi_j)\}. \tag{1.6}$$

When the stationary condition holds, we see easily that $\gamma_j(0)$ is convergent.

## 2. STATISTICAL INFERRENCE FOR THE SAR(1,1) MODEL

The likelihood function based on all $n$ samples is given by

$$L = (2\pi\sigma^2)^{-\frac{n}{2}} \exp(-\frac{1}{2\sigma^2}) \sum_{j=1}^{k} \sum_{t=1}^{n_j} e_{t,j}^2, \tag{2.1}$$

where $e_{t,j} = X_{t,j} - \phi_j X_{t-1,j} - \Phi_j X_{t-s,j} + \phi_j \Phi_j X_{t-s-1,j}$. We set the derivatives $\partial \log L/\partial\phi_j$ and $\partial \log L/\partial\Phi_j$ to zero. The solutions that maximize (2.1) on $\phi_j$ and $\Phi_j$ give the unrestricted maximum likelihood estimates $\hat{\phi}_j$ and $\hat{\Phi}_j$ of the parameters $\phi_j$ and $\Phi_j$,respectively. Likewise, the maximum likelihood estimate $\hat{\sigma}^2$ of the variance $\sigma^2$ is

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^{k} \sum_{t=1}^{n_j} (X_{t,j} - \hat{\phi}_j X_{t-1,j} - \hat{\Phi}_j X_{t-s,j} + \hat{\phi}_j \hat{\Phi}_j X_{t-s-1,j})^2.$$

Let $\theta_j^T = (\phi_j, \Phi_j)$ and consider the Taylor expansion $e_{i,j}(\theta)$ at a suitable preliminary estimate $\theta_{0,j}$. Then,$e_{t,j}(\theta_j) \cong e_{t,j}(\theta_{0,j}) - D_{t,j}^T(\theta_j - \theta_{0,j}), \quad j = 1, \cdots, k$, where $D_{t,j} = (-\partial e_{t,j}/\partial\phi_j, \quad -\partial e_{t,j}/\partial\Phi_j)^T|_{\theta_j=\theta_{0,j}}$. For each $j = 1, \cdots, k$ ,

$$\sum_{t=1}^{n_j} e_{t,j}^2(\theta_j) \cong \sum_{t=1}^{n_j} \{e_{t,j}(\theta_{0,j}) - D_{t,j}^T(\theta_j - \theta_{0,j})\}^2. \tag{2.2}$$

When we minimize the right-hand side of (2.2) with respect to $\theta_j$, we obtain

$$\tilde{\theta}_j = \theta_{0,j} + (\sum_{t=1}^{n_j} D_{t,j} D_{t,j}^T)^{-1} \{\sum_{t=1}^{n_j} D_{t,j} e_{t,j}(\theta_{0,j})\}. \tag{2.3}$$

Note that $\theta_{0,j}$ in (2.3) needs to be replaced by a preliminary estimate $\tilde{\theta}_{0,j}$. Let $\tilde{\theta}_j^*$ be $\tilde{\theta}_j$ with $\theta_{0,j}$ replaced by $\tilde{\theta}_{0,j}$. Then $\tilde{\theta}_j^*$ is an approximate least squares estimate of $\theta_j$. It can be shown that $\tilde{\theta}_j^*$ has the same limiting distribution as does $\tilde{\theta}_j$ and that this is also the limiting distribution of the maximum likelihood estimates $\hat{\theta}_j = (\hat{\phi}_j, \hat{\Phi}_j)^T$ (See Fuller,1976,p211-220). Thus from (2.3), we have that the limiting distribution of $\hat{\theta}_j$ can be found from

$$\sqrt{n_j}(\hat{\theta}_j - \theta_j) = (n_j^{-1} \sum_{t=1}^{n_j} D_{t,j} D_{t,j}^T)^{-1} \{n_j^{-\frac{1}{2}} \sum_{t=1}^{n_j} D_{t,j} e_{t,j}(\theta_j)\}. \tag{2.4}$$

**Lemma 2.1.** Let $\{X_{t,j}\}, t = 1, \cdots, n_j$, and $j = 1, \cdots, k$ be a stationary seasonal autoregressive process of order (1,1) satisfying equation (1.2). Then for each $j = 1, \cdots, k$,

$$n_j^{-1} \sum_{t=1}^{n_j} D_{t,j} D_{t,j}^T \xrightarrow{p} g_j(\theta_j) \qquad \text{as} \quad n_j \to \infty \tag{2.5}$$

where

$$g_j(\theta_j) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{2.6}$$

with $a_{11} = \gamma_j(0) - 2\Phi_j\gamma_j(s) + \Phi_j^2\gamma_j(0)$, $a_{22} = \gamma_j(0) - 2\phi_j\gamma_j(1) + \phi_j^2\gamma_j(0)$,
$a_{12} = a_{21} = \gamma_j(s-1) - \Phi_j\gamma_j(1) - \phi_j\gamma_j(s) + \phi_j\Phi_j\gamma_j(0)$,
and $\gamma_j(h)$, $h = 0, 1, \cdots$, is defined in (1.5).

**Proof.** For each $j = 1, \cdots, k$ ,we have

$$n_j^{-1} \sum_{t=1}^{n_j} D_{t,j} D_{t,j}^T = n_j^{-1} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

where $b_{11} = \sum_{t=1}^{n_j} (X_{t-1,j} - \Phi_j X_{t-s-1,j})^2$, $b_{22} = \sum_{t=1}^{n_j} (X_{t-s,j} - \phi_j X_{t-s-1,j})^2$,
$b_{12} = b_{21} = \sum_{t=1}^{n_j} (X_{t-1,j} - \Phi_j X_{t-s-1,j})(X_{t-s,j} - \phi_j X_{t-s-1,j})$.

By the Law of Large Numbers for $m$-dependent processes of Hoeffding and Robbins(1948), $n_j^{-1}\sum_{t=1}^{n_j} X_{t,j}X_{t+h,j} \xrightarrow{p} \gamma_j(h)$. Hence our result is followed.

**Lemma 2.2.** Let $\{X_{t,j}\}$ be a stationary seasonal autoregressive process of order (1,1) satisfying equation (1.2). Then for each $j = 1, \cdots, k$,

$$n_j^{-\frac{1}{2}} \sum_{t=1}^{n_j} D_{t,j} e_{t,j}(\theta_j) \xrightarrow{d} N_2(0, \sigma^2 g_j(\theta_j)) \quad \text{as } n_j \to \infty \tag{2.7}$$

where $g_j(\theta_j)$ is defined in (2.6).

**Proof.** We can write $n_j^{-\frac{1}{2}} \sum_{t=1}^{n_j} D_{t,j} e_{t,j}(\theta_j) = n_j^{-\frac{1}{2}}[c_1 \ , c_2]^T$ where $c_1 = \sum_{t=1}^{n_j}(X_{t-1,j} - \Phi_j X_{t-s-1,j})e_{t,j}$, $c_2 = \sum_{t=1}^{n_j}(X_{t-s,j} - \phi_j X_{t-s-1,j})e_{t,j}$ Without loss of generality, consider the $j$ th sample and drop the subscript $j$, because of independence for different $j = 1, \cdots, k$. Let $X_{t,m} = \sum_{r=0}^{m} \psi_r e_{t-r}$ with $\psi_r$ defined (1.4). Let $Y_{t,m} = X_{t-1,m} - \Phi X_{t-s-1,m}$ and $Z_{t,m} = X_{t-s,m} - \phi X_{t-s-1,m}$. Then $(Y_{t,m}e_t)$ and $(Z_{t,m}e_t)$ are $(m+s+1)$-dependent processes, and $E(Y_{t,m}e_t) = 0$ and $E(Z_{t,m}e_t) = 0$. Under the stationary condition, it follows that

$$\lim_{m\to\infty} Var\,(Y_{t,m}e_t) = \sigma^2(1+\Phi^2)\gamma(0) - 2\Phi\sigma^2\gamma(s),$$
$$\lim_{m\to\infty} Var\,(Z_{t,m}e_t) = \sigma^2(1+\phi^2)\gamma(0) - 2\phi\sigma^2\gamma(1),$$
$$\lim_{m\to\infty} Cov\,(Y_{t,m}e_t, Z_{t,m}e_t) = \sigma^2[\gamma(s-1) - \phi\gamma(s) - \Phi\gamma(1) + \phi\Phi\gamma(0)].$$

Thus, by the central limit theorem for $m$-dependent processes of Hoeffding and Robbins(1948), we have equation (2.7), and hence our result is proved.

**Theorem 2.1.** Let $\{X_{t,j}\}, t = 1, \cdots, n_j$, and $j = 1, \cdots, k$, be a stationary seasonal autoregressive process of order (1,1) satisfying equation (1.2), and let $\theta^T = (\theta_1, \cdots, \theta_k)$ where $\theta_j^T = (\phi_j, \Phi_j), j = 1, \cdots, k$. Then,

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N_{2k}(0, G(\theta)) \quad \text{as } n \to \infty \tag{2.8}$$

where $G(\theta)$is the $(2k \times 2k)$ matrix,

$$G(\theta) = \sigma^2 Diag[\{c_1 g_1(\theta_1)\}^{-1}, \cdots, \{c_k g_k(\theta_k)\}^{-1}] \tag{2.9}$$

with $g_j(\theta_j)$ defined as in equation (2.6).

**Proof.** From Lemma 2.1., Lemma 2.2. and Slutsky's theorem, for each $j = 1, \cdots, k$, it follows that

$$\sqrt{n_j}(\hat{\theta}_j - \theta_j) \xrightarrow{d} N_2(0, \sigma^2 g_j^{-1}(\theta_j)) \quad \text{as } n_j \to \infty. \tag{2.10}$$

However, for different $j = 1, \cdots, k$, the $\hat{\theta}_j$ are independent and $n_j/n \to c_j$ as $n \to \infty$. Hence,

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N_{2k}(0, G(\theta)) \quad \text{as } n \to \infty.$$

## 3. LARGE SAMPLE TEST FOR THE SAR(1,1) MODEL

The null hypothesis to test the homogeneity of $k$ seasonal autoregressive processes of order (1,1) is given by $H : \phi_1 = \cdots = \phi_k$ and $\Phi_1 = \cdots = \Phi_k$ while the alternative hypothesis is $Kn : \phi_{j,n} = \phi_1 + h/\sqrt{n}$ and $\Phi_{j,n} = \Phi_1 + d/\sqrt{n}, \quad j = 1, \cdots, k$, where $h$ and $d$ are $(k \times 1)$ vectors of real numbers. Let us reparameterize $(\theta_1, \cdots, \theta_k)$ to $(\eta_2, \cdots, \eta_k)$ with $\eta_j = \theta_j - \theta_1, \quad j = 2, \cdots, k$. Then hypothesis $H$ becomes $H : \eta_j = 0, \quad j = 2, \cdots, k$. Let $\eta^T = (\eta_2, \cdots, \eta_k)$. Then, the *Wald statistic* $Q_{1n}$ for the hypothesis $H$ is given by

$$Q_{1n} = n(\hat{\eta} - \eta)^T J^{-1}(\theta)(\hat{\eta} - \eta) \tag{3.1}$$

where $\quad J(\theta) = (\partial\eta/\partial\theta)^T G(\theta)(\partial\eta/\partial\theta) \quad (3.2)$
with

$$(\partial\eta/\partial\theta)^T = \begin{pmatrix} -I & I & 0 & \cdots & 0 \\ -I & I & I & \cdots & 0 \\ \cdot & \cdot & \cdot & & \cdot \\ -I & 0 & 0 & \cdots & I \end{pmatrix}, \tag{3.3}$$

$I$ being the $2 \times 2$ dimensional identity matrix, and with $G(\theta)$ being defined as in (2.9).

**Theorem 3.1.** Under the null hypothesis $H$, the *Wald statistic* $Q_{1n}$ converges in distribution to that of a Chi-square distribution with $2(k-1)$ degrees of freedom.

**Proof.** By Theorem 2.1, we have $\sqrt{n}(\hat{\theta}-\theta)\xrightarrow{d}N_{2k}(0,G(\theta))$ as $n\to\infty$. Thus, under the null hypothesis $H$, $\sqrt{n}(\hat{\eta}-\eta)\xrightarrow{d}N_{2(k-1)}(0,J(\theta))$ as $n\to\infty$. Under $H$, the *Wald statistic*,$Q_{1n}$ becomes $Q_{1n}=n\hat{\eta}^T J^{-1}(\theta)\hat{\eta}$. Hence its limiting distribution follows a Chi-square distribution with $2(k-1)$ degrees of freedom.

# 4. EXTENSION TO THE SAR(p,P) PROCESS

Analogous results for the general multiplicative seasonal autoregressive process of order p and P, SAR(p,P), can likewise be obtained. From (1.1), the general form of a multiplicative seasonal autoregressive process of order (p,P) is specified by,$t=1,\cdots,n_j,\ j=1,\cdots,k$,

$$X_{t,j}=\phi_{1,j}X_{t-1,j}+\cdots+\phi_{p,j}X_{t-p,j}+\Phi_{1,j}X_{t-s,j}+\cdots+\Phi_{P,j}X_{t-Ps,j}$$
$$-\phi_{1,j}\Phi_{1,j}X_{t-s-1,j}-\cdots-\phi_{p,j}\Phi_{P,j}X_{t-Ps-p,j}+e_{t,j}. \tag{4.1}$$

Let $\Theta_j=(\phi_j,\Phi_j)^T$ be the $(p+P)\times 1$ vector with $\phi_j=(\phi_{1,j},\cdots,\phi_{p,j})^T$ and $\Phi_j=(\Phi_{1,j},\cdots,\Phi_{P,j})^T$ for each $j=1,\cdots,k$. Then, the null hypothesis for testing the homogeneity of $k$ SAR$(p,P)$ processes is given by $H:\Theta_1=\cdots=\Theta_k$, while the alternative hypothesis is given by $Kn:\Theta_{jn}=\Theta_1+h/\sqrt{n},\quad j=1,\cdots,k$, where $h$ is a $(p+P)\times 1$ vector of real numbers. Let $\eta^T=(\eta_2,\cdots,\eta_k)$ with $\eta_j=\Theta_j-\Theta_1,\quad j=1,\cdots,k$. The *Wald statistic*,$Q_{1n}$, is

$$Q_{1n}=n(\hat{\Theta}-\Theta)^T[HG_n^{-1}(\hat{\Theta})H^T]^{-1}(\hat{\Theta}-\Theta) \tag{4.2}$$

where $H$ is the $(k-1)(p+P)\times k(p+P)$ matrix, $H=(\partial\eta/\partial\Theta)^T$ and $G_n(\hat{\Theta})$ is the $k(p+P)\times k(p+P)$ matrix, that is, $G_n(\hat{\Theta})=Diag[G_{n1}(\hat{\Theta}),\cdots,G_{nk}(\hat{\Theta})]$ with $G_{nj}(\hat{\Theta})$ being the $(p+P)\times(p+P)$ matrix

$$G_{nj}(\Theta)=n^{-1}\sigma^{-2}\begin{pmatrix}A & C\\ C^T & B\end{pmatrix} \tag{4.3}$$

(evaluated at $\Theta=\hat{\Theta}$) where the $(p\times p)$ matrix **A** has elements $(a_{t-i}^T a_{t-m}), i, m=1,\cdots,p$, the $(P\times P)$ matrix **B** has elements $(b_{t-is}^T b_{t-ms}), i,m=1,\cdots,,P$, and the $(p\times P)$ matrix **C** has elements $(a_{t-1}^T b_{t-ms}+c_l^T d_{im}), i=1,\cdots,p$ ,$m=1,\cdots,P$ with $a_{t-i}^T=(X_{1-i}-\Phi_1 X_{1-i-s}-\cdots-\Phi_P X_{1-Ps-i},\cdots,X_{n-i}-\Phi_1 X_{n-i-s}-\cdots-\Phi_P X_{n-Ps-i})$, $b_{t-is}^T=(X_{1-is}-\phi_1 X_{1-is}-\cdots-\phi_p X_{1-is-p},$

$\cdots, X_{n-is}-\phi_1 X_{n-1-is}-\cdots-\phi_p X_{n-is-p})$, $c_t^T = (X_1-\Phi_1 X_0-\cdots+\phi_p\Phi_P X_{1-Ps-i},$
$\cdots, X_n - \phi_1 X_{n-1} - \cdots + \phi_p X_{n-Ps-p})$ and $d_{im}^T = (X_{1-is-m}, \cdots, X_{n-is-m})$.
In (4.2), $\hat{\Theta}$ is the unrestricted maximum likelihood estimate of $\Theta$ and $G_n(\hat{\Theta}_H)$ is defined as in (4.3) but with $\hat{\Theta}_H$ in place of $\Theta$.

**Result.** Let $\{X_{i,j}\}$, $t = 1, \cdots, n_j$, $j = 1, \cdots, k$, be a stationary process satisfying (4.1). Then under the null hypothesis $H$, the *Wald statistic* $Q_{1n}$ converges in distribution to that of a Chi-square distribution with $(k-1)(p+P)$ degrees of freedom.

# REFERENCES


(1) Anderson. T. W.(1978). Repeated Measurements on Autoregressive Processes. *Journal of the American Statistical Association*, 73, 371-378.

(2) Basawa. I. V., Billard. L. and Srinivasan. R.(1984). Large Sample Tests of Homogeneity for Time Series Models. *Biometrika*, 71, 203-206.

(3) Box. G. E. P. and Jenkins. G. M.(1976). *Time Series Analysis; Forecasting and Control*, 2nd ed., Holden Day Inc., California.

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

(5) Hoeffding. W. and Robbins. H.(1948). The Central Limit Theorem for Dependent Random Variables. *Duke Mathematical Journal*, 15, 773-780.

(6) Lee. S. D.(1989). *Large Sample Tests for the Multiplicative Seasonal Time Series Models*, Ph.D. Dissertation, University of Geogia.

(7) Wald. A.(1943). Tests of Statistical Hypothesis Concerning Several Parameters When the Number of Observations is Large. *Transactions of the American Mathematical Society*, 54, 426-482.