

## Comparison between homogeneity test statistics for panel AR(1) model

Sung Duck Lee<sup>a,1</sup> · Sun Woo Kim<sup>a</sup> · Na Rae Jo<sup>a</sup>

<sup>a</sup>Department of Information and Statistics, Chungbuk National University

(Received December 15, 2015; Revised December 28, 2015; Accepted December 28, 2015)

---

### Abstract

We can achieve the principle of parsimony and efficiency if homogeneity for panel time series model is satisfied. We suggest a Rao test statistic and a Wald test statistic for the test of homogeneity for panel AR(1) and derived the limit distribution. We performed a simulation to examine statistics with the same chi-square distribution when number of the individual is small and in common with large. We also simulated to compare the empirical power of the statistics in a small panel. In application, we fit panel AR(1) model using regional monthly economical active population data and test homogeneity for panel AR(1). It is satisfied homogeneity, so it could be fitted AR(1) using the sample mean at the time point. We also compare the power of prediction between each individual and pooled model.

Keywords: panel first order autoregressive model, test of homogeneity, Rao test statistic, Wald test statistic, regional economical active population data

---

### 1. 서론

하나의 현상에 대해 여러 개의 독립적인 시계열로 구성된 자료는 실제 현상에서 자주 나타나게 되는 데 이를 패널 시계열 자료라 한다. 이때  $m$ 개의 독립적인 시계열 모형이 생기는데, 이러한 자료를 분석할 때 먼저  $m$ 개의 시계열들이 동일한 패턴을 갖고 있는지 확인한다. 만약 동일한 패턴을 가지고 있다면 각 시계열이 하나의 모집단에서 온 것으로 간주할 수 있을 것이므로 각각의 독립적인 시계열들을 개별적으로 분석하는 것보다 자료를 종합(pooling)하여 분석하는 것이 모수절약 원칙(the principle of parsimony)에 의해 더욱 효율적이다.

Anderson (1978)은 의사가 환자들의 혈압을 여러 날 동안 측정하여 이 환자들의 혈압 간에 동질성이 있는지를 알아보았다. Basawa 등 (1984)은 ARMA( $p, q$ ) 모형에 대하여 시계열 패널 자료에 대한 동질성 검정을, Lee (1993)는 계절성 시계열 패널 자료에 대한 동질성 검정을 연구하였다. 이런 패널 자료의 동질성 검정은 선형 모형 뿐만 아니라 비선형 시계열 모형에 대해서도 연구되었는데 Kim 등 (2000)은 일차 비선형 시계열의 정상성 조건을 유도하고, 여러개의 독립적인 일차 비선형 시계열로 구성된 패널 자료의 동질성 검정법과 그 극한분포를 제시하였다. 또한 Lee 등 (2013)은 패널 중선형 시계열 모형의 모수에 대한 동질성 검정과 검정 통계량의 극한 분포를 유도하고, 실증분석에서 실제 자료에 적합하였다.

This work was supported by the research grant of the Chungbuk National University in 2014.

<sup>1</sup>Corresponding author: Department of Information and Statistics, Chungbuk National University, 1 Chungdae-ro, Seowon-gu, Cheongju, Chungbuk 28644, Korea. E-mail: [sdlee@chungbuk.ac.kr](mailto:sdlee@chungbuk.ac.kr)

자료가 모든 시점에서 관측되지 않고 간헐적인 자료를 갖는 패널 시계열 모형의 동질성 검정에 관한 연구로는 Basawa와 Billard (1989)의 연구가 있는데 간헐적인 자료에 대한 자기상관 오차항을 갖는 회귀 모형의 동질성 검정과 검정 통계량의 극한분포를 소개하였고, 시계열의 시점 별 표본 평균을 이용하여 자료를 집약(aggreated)한 후 모수 추정의 문제를 다루었다. Basawa와 Billard (1989)의 연구는 상관이 있는 오차항을 갖는 비선형 회귀모형으로 확장 가능한데 Glasbey (1979)는 그러한 일반화가 필요한 실제 응용 사례를 제공했다. Lee (2014)는 간헐적인 자료를 갖는 패널 시계열 모형의 동질성 검정과 검정 통계량의 극한분포를 소개하고 모의실험을 통해 확인하였으며, 실제 자료에 적합하였다.

본 논문에서는 패널의 수가 작은 패널 시계열 모형에서 동질성 검정을 살펴보았다. 패널의 수가 작을 때 패널 1차 자기회귀 모형의 동질성 검정 통계량의 분포를 살펴보고, 두 가지 동질성 검정 통계량의 검정력을 비교하였다. 2장에서는 패널 1차 자기회귀 모형을 소개하고, 3장에서는 패널 AR(1) 모형의 동질성 검정을 위한 세가지 검정 통계량을 구하고 그 극한분포를 유도한다. 4장에서는 패널의 수가 적은 패널 AR(1) 모형의 동질성 검정 통계량의 분포를 모의 실험하였으며, 두 검정 통계량의 검정력을 비교하였다. 5장에서는 실제 지역별 경제활동인구수 자료를 이용하여 동질성 검정을 수행하였으며, 개별적으로 적합한 모형과 시점별로 종합된 자료를 이용하여 적합한 모형의 예측력을 비교하고 6장에서 결론을 맺는다.

## 2. 패널 1차 자기회귀 모형

AR(1) 모형의 시계열  $\{X_t(j), t = 1, 2, \dots, n, j = 1, 2, \dots, m\}$ 가 시점  $t$ 에서의  $j$ 번째 패널 시계열일 때 이를 패널 1차 자기회귀 모형이라 하며, 패널 AR(1) 모형은 다음과 같은 식을 갖는다.

$$X_t(j) = \phi_j X_{t-1}(j) + e_t(j), \quad t = 1, 2, \dots, n, j = 1, 2, \dots, m, \quad (2.1)$$

여기서  $e_t(j)$ 는 독립이고 동일한 정규분포를 따르며 평균은 0이고, 분산은  $\sigma_e^2$ 인 확률변수이며, 모든 패널에서 1차 자기회귀계수  $\phi_j$ 가 정상조건  $|\phi_j| < 1$ 을 만족한다고 한다. 그리고 전체 관측치의 개수는 각 패널의 관측치의 개수의 합으로  $N = n_1 + n_2 + \dots + n_m$ 이다.

오차항  $e_t(j)$ 가 정규분포를 따른다는 가정을 통해서 총  $N$ 개의 표본에 대한 우도함수는 마코브 과정에 의해 다음과 같이 유도할 수 있다.

$$L_N = (2\pi\sigma_e^2)^{-\frac{1}{2}N} \exp \left[ -\frac{1}{2} \sum_{j=1}^m \sum_{t=1}^{n_j} \frac{\{X_t(j) - \phi_j X_{t-1}(j)\}^2}{\sigma_e^2} \right]. \quad (2.2)$$

이 우도함수를 최대로 하는  $\phi_j$ 의 최대우도추정량  $\hat{\phi}_j$ 는 다음과 같다.

$$\hat{\phi}_j = \frac{\sum_{t=1}^{n_j} X_t(j)X_{t-1}(j)}{\sum_{t=1}^{n_j} X_{t-1}^2(j)}, \quad j = 1, 2, \dots, m. \quad (2.3)$$

## 3. 패널 AR(1)의 동질성 검정을 위한 통계량과 극한분포

패널 AR(1)에서 각각의 시계열들의 자기회귀계수의 동질성 검정을 위한 가설은 다음과 같다.

$$\begin{aligned} H_0 : \phi_1 = \phi_2 = \dots = \phi_m, \\ H_1 : \text{적어도 하나의 } \phi_j \text{는 같지 않다.} \end{aligned} \quad (3.1)$$

패널 자기회귀 모형의 모수에 대한 동질성 검정에는 일반적으로 세 가지 통계량이 주로 사용되는 바 LR(likelihood ratio) 통계량, Rao 통계량, Wald 통계량이 있다. 먼저 LR 통계량은 다음과 같다.

$$LR = -2 \log \frac{L_H(\phi_H)}{L(\phi_j)}, \quad (3.2)$$

여기서  $L_H$ 는 귀무가설 하에서 우도함수이며  $\phi_H$ 는  $L_H$ 를 바탕으로 얻어진  $\phi_H$ 의 제한된 최우추정량으로  $L_H$ 와  $\hat{\phi}_H$ 은 아래와 같다.

$$L_H = \left( \frac{1}{2\pi\sigma_e^2} \right)^{-\frac{1}{2}N} \exp \left[ -\frac{1}{2} \sum_{j=1}^m \sum_{t=1}^{n_j} \frac{\{X_t(j) - \phi_H X_{t-1}(j)\}^2}{\sigma_e^2} \right], \quad (3.3)$$

$$\hat{\phi}_H = \frac{\sum_{j=1}^m \sum_{t=1}^{n_j} X_t(j) X_{t-1}(j)}{\sum_{j=1}^m \sum_{t=1}^{n_j} X_{t-1}^2(j)}. \quad (3.4)$$

다음으로 Rao (1973)가 제안한 패널 1차 자기회귀 모형의 모수에 대한 동질성 검정 Rao 통계량은 다음과 같으며, 여기서  $B_N(\phi)$ 는 Fisher information 행렬로  $m \times m$  행렬이다.

$$RAO = \left[ \frac{d \log L}{d\phi} \right]^T B_N^{-1}(\hat{\phi}_H) \left[ \frac{d \log L}{d\phi} \right]_{\phi=\hat{\phi}_H}, \quad (3.5)$$

$$B_N(\phi) = -\frac{\partial^2 \log L_N}{\partial \phi_r \partial \phi_s}, \quad r, s = 1, 2, \dots, m,$$

$$= \text{Diag} \left[ \sum_{t=1}^{n_1} X_{t-1}^2(1), \sum_{t=1}^{n_2} X_{t-1}^2(2), \dots, \sum_{t=1}^{n_m} X_{t-1}^2(m) \right].$$

식 (3.5)를 이용하여 패널 1차 자기회귀 모형에서의 Rao 통계량을 다음과 같이 계산할 수 있다.

$$RAO = \sum_{j=1}^m \sum_{t=1}^{n_j} \frac{1}{\sigma^2} X_{t-1}^2(j) (\hat{\phi}_j - \hat{\phi}_H)^2. \quad (3.6)$$

Lee (1993)는 패널 1차 자기회귀 모형에서는 LR 통계량과 Rao 통계량이 같음을 증명하였다. 마지막 Wald 통계량은 동질성 검정을 위한 가설, 식 (3.1)의 모수를  $\eta_{j-1} = \phi_j - \phi_1$ ,  $j = 2, \dots, m$ 로 이루어진 아래와 같은 가설로 재모수화하여 구할 수 있다.

$$H_0 : \eta_j = 0, \quad j = 2, 3, \dots, m, \quad (3.7)$$

$$H_1 : \text{적어도 하나의 } \eta_j \text{는 } 0 \text{이 아니다.}$$

식 (3.7)의 재모수화된 가설을 사용하여 구한 패널 1차 자기회귀 모형의 모수에 대한 동질성 검정을 위한 Wald 통계량은 다음과 같다.

$$W_N = \hat{\eta}^T \left\{ D B_N^{-1}(\hat{\phi}) D^T \right\}^{-1} \hat{\eta}, \quad (3.8)$$

여기서  $\hat{\phi}$ 는 다른 통계량에서와 같이  $\phi$ 의 제한되지 않은 최우추정량이고, 패널 1차 자기회귀 모형에서  $D$ 는 아래와 같다.

$$D = \frac{\partial \eta_r}{\partial \phi_s}, \quad r = 1, 2, \dots, m-1, \quad s = 1, 2, \dots, m$$

**Table 4.1.** Frequency of empirical distribution ( $m = 5$ ) (%)

Mid Point	$\chi^2(4)$	$\phi = 0.1$		$\phi = 0.5$	
		Rao	Wald	Rao	Wald
0.5	9.0	8.7	8.8	9.2	10.0
1.5	17.4	20.8	21.3	19.5	19.4
2.5	17.8	17.9	18.4	17.3	17.3
3.5	15.2	14.8	14.1	15.0	15.0
4.5	11.9	10.6	10.8	11.4	12.3
5.5	8.8	7.4	7.6	9.1	7.8
6.5	6.3	6.2	6.6	6.8	4.8
7.5	4.4	4.6	4.0	3.3	4.5
8.5	3.0	2.9	2.6	2.7	2.9
9.5	2.1	1.7	1.7	2.0	1.8
10.5	1.4	1.8	1.8	1.3	1.9
11.5	0.9	0.9	0.7	0.7	1.0
	1.7	1.7	1.6	1.7	1.3
합계	100	100	100	100	100

$$= \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Basawa 등 (1984)은  $N \rightarrow \infty$ 일 때  $B_N(\phi) \rightarrow B(\phi)$ 으로 확률적 수렴한다고 하면 동질성 검정의 Wald 통계량은 정규조건 하에서  $W_N \rightarrow \chi^2(m-1, \lambda^2)$ 와 같이 비중심 카이제곱분포로 분포 수렴한다는 것을 증명하였다. 여기서  $B(\phi)$ 는 비확률, 정칙인 극한 Fisher information 행렬이고, 비중심모수는  $\lambda^2 = \delta^T \{DB^{-1}(\phi)D^T\}^{-1}$ 이지만 동질성 검정을 위한 귀무가설 하에서는 0이 되어 결국 Wald 통계량은 카이제곱분포를 따르게 된다.

또한 Basawa 등 (1984)은 시계열 모형에서 동질성 검정의 LR 통계량, Rao 통계량, Wald 통계량이 표본의 크기가 충분히 크면 점근적으로 같은 극한분포를 가짐을 증명하였다.

#### 4. 모의실험

앞에서 소개한 3가지 검정 통계량 중 주로 많이 사용되는 Wald 통계량과 Rao 통계량이 패널의 수가 적을 때 모의실험을 통해서  $\chi^2$ 분포로 접근하는지 알아보았다. 통계패키지 R을 이용하여 AR(1)을 따르는 난수를 발생시켜 패널 AR(1)을 따르는 데이터를 생성하였다. 이때 패널의 개수가 각각 5개, 10개, 15개인 경우 동질성 검정을 위한 Rao 통계량과 Wald 통계량이 각각  $\chi^2$ 분포로 근접하는지 알아보았다.

Table 4.1은 패널의 수가 5개이고 패널 1차 자기회귀 계수  $\phi_j$ 가 각각 0.1, 0.5인 패널 AR(1) 모형을 따르는 난수로 Rao 통계량과 Wald 통계량을 구한 후 이를 1,000번 반복하여 유의확률과 비교한 결과이다. Table 4.1의 결과를 보면 패널의 수가 5개인 경우  $\phi_j = 0.1, 0.5$ 인 경우 Rao 통계량과 Wald 통계량 모두  $\chi^2(4)$ 의 유의확률과 유사함을 확인할 수 있으며, Figure 4.1과 Figure 4.2에서도 두 통계량의 히스토그램이 모두  $\chi^2(4)$ 에 근접함을 확인할 수 있다.

같은 방법으로 Table 4.2는 패널의 수가 10개이고 패널 1차 자기회귀 계수  $\phi_j$ 가 각각 0.1, 0.5인 패널

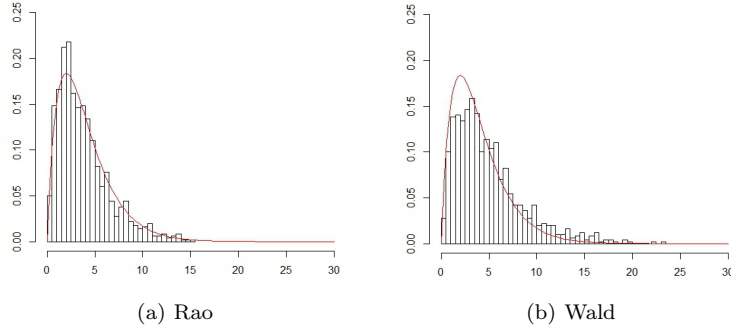


Figure 4.1. Histogram of two statistics ( $m = 5, \phi = 0.1$ ).

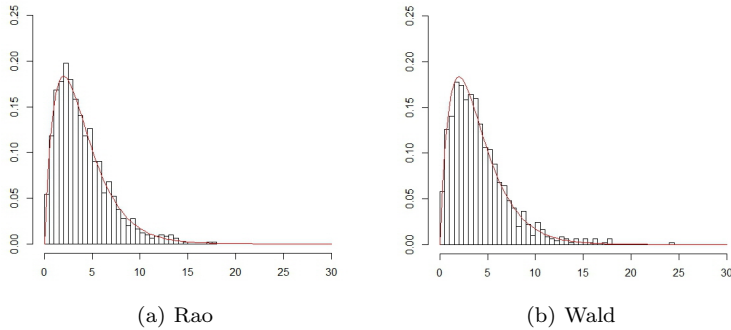


Figure 4.2. Histogram of two statistics ( $m = 5, \phi = 0.5$ ).

Table 4.2. Frequency of empirical distribution ( $m = 10$ ) (%)

Mid Point	$\chi^2(9)$	$\phi = 0.1$		$\phi = 0.5$	
		Rao	Wald	Rao	Wald
1	0.9	0.8	0.8	1.2	1.0
3	8.0	10.4	10.3	9.3	8.3
5	17.1	17.1	16.7	17.3	17.1
7	20.6	18.8	18.8	20.5	17.4
9	18.4	19.7	20.6	14.7	17.6
11	13.7	13.3	13.3	15.7	13.4
13	9.1	8.2	8.1	10.0	9.5
15	5.5	5.7	4.4	5.4	7.7
17	3.2	2.6	3.1	3.5	3.6
19	1.7	1.7	1.8	1.5	2.5
21	0.9	1.0	1.1	0.7	1.0
23	0.5	0.4	0.8	0.1	0.2
	0.4	0.3	0.2	0.1	0.7
합계	100	100	100	100	100

AR(1) 모형을 따르는 난수로 Rao 통계량과 Wald 통계량을 구한 후 이를 1,000번 반복하여 유의확률과 비교한 결과이다. Table 4.2의 결과를 보면 패널의 수가 10개이고  $\phi = 0.1, 0.5$ 인 경우 역시 Rao 통계량과 Wald 통계량 모두  $\chi^2(9)$ 의 유의확률과 유사함을 확인할 수 있다. Figure 4.3과 Figure 4.4를 보아도

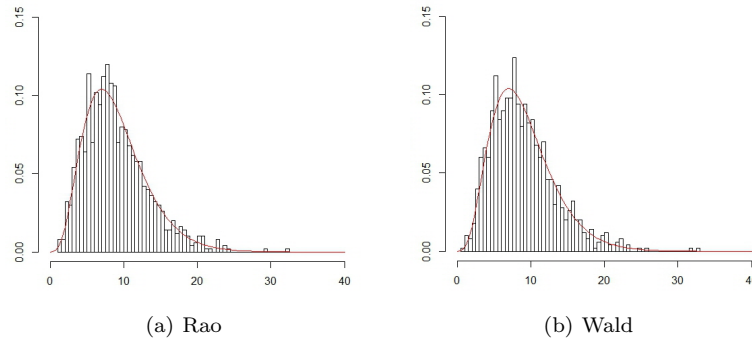


Figure 4.3. Histogram of two statistics ( $m = 10, \phi = 0.1$ ).

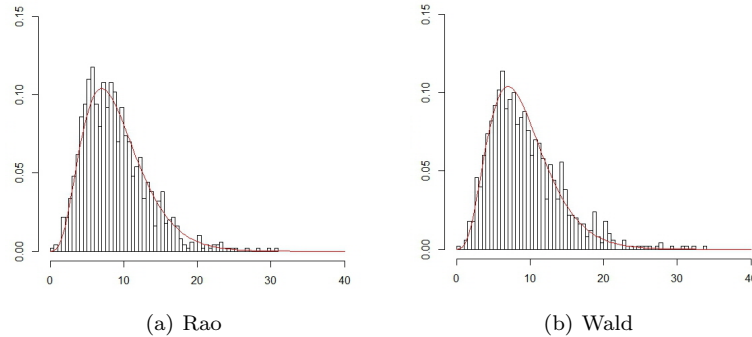


Figure 4.4. Histogram of two statistics ( $m = 10, \phi = 0.5$ ).

두 통계량의 히스토그램이 모두  $\chi^2(9)$ 에 근접함을 확인할 수 있다.

한편 패널의 수가 작을 때 Rao 통계량과 Wald 통계량 중 어떤 통계량이 더 바람직한지 알아보기 위하여 통계프로그램 R을 이용하여  $\phi_1 = 0.1$ 로 고정하고  $\phi_2$ 와  $\phi_3$ 를 변화시켜 시점의 수는 100개이고 패널의 수가 3개인 패널 AR(1)을 따르는 시계열 난수를 각각 1,000개씩 발생시켰다. Table 4.3과 Table 4.4는 계산된 각각의 1,000개의 Rao 통계량과 Wald 통계량이 동질성 귀무가설을 기각할 빈도(%)를 계산한 결과이다.

Table 4.3과 Table 4.4를 보면 Rao 통계량과 Wald 통계량 모두  $\phi_1$ 과  $\phi_2, \phi_3$ 의 차이가 커질수록 빠르게 100으로 가까이 가는 것을 확인할 수 있다. 두 통계량을 비교해 본다면 Rao 통계량이 Wald 통계량보다 더 민감하게 패널 1 차 자기회귀 계수의 차이를 감지하는 것을 확인할 수 있다. 따라서 패널의 수가 작을 때, 패널 1차 자기회귀 모형의 동질성 검정을 위한 검정 통계량으로는 Rao 통계량이 Wald 통계량보다 더 우수하다고 할 수 있다.

## 5. 사례연구

통계청 국가통계포털(KOSIS)에서 제공하는 2000년 1월부터 2013년 12월까지 14년(168개월)동안 5개 시도(충청북도, 전라남도, 경상북도, 경상남도, 제주도)의 월별 경제활동인구수 자료를 사용하였다. KOBIS에서 제공하는 자료는 단위가 “명”이므로 포아송 분포를 따르는 변수로 이 자료를 정규분포로 근사시키기 위해 각 지역별로 표준화하였으며 계절차분을 통하여 정상성을 만족하는 자료로 변환하였다.

**Table 4.3.** Empirical power of Rao statistics (%)

Rao	$\phi_2$									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	
$\phi_3$	0.1	3.8	9.4	28.9	58.0	87.8	97.5	99.9	100.0	100.0
	0.2	9.2	9.4	20.7	49.9	79.0	95.2	99.5	100.0	100.0
	0.3	29.2	23.6	28.0	52.6	76.7	94.3	99.2	100.0	100.0
	0.4	59.3	49.7	49.8	57.2	75.9	93.1	98.9	100.0	100.0
	0.5	87.0	79.1	74.9	75.7	86.3	93.5	99.1	99.8	100.0
	0.6	97.9	95.7	93.8	91.6	95.1	97.8	99.2	99.9	100.0
	0.7	100.0	99.5	99.2	98.9	98.6	99.7	100.0	100.0	100.0
	0.8	100.0	100.0	99.8	99.9	99.9	99.9	100.0	100.0	100.0
	0.9	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

**Table 4.4.** Empirical power of Wald statistics (%)

Wald	$\phi_2$									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	
$\phi_3$	0.1	3.4	8.9	26.4	55.6	85.9	97.0	99.8	100.0	100.0
	0.2	8.5	8.4	18.3	45.3	75.3	93.6	99.4	99.9	100.0
	0.3	26.8	20.8	25.5	48.6	71.9	92.4	98.8	100.0	100.0
	0.4	54.9	45.2	45.5	54.6	73.3	92.3	98.7	100.0	100.0
	0.5	84.5	75.9	70.8	73.1	84.6	92.7	98.5	99.8	100.0
	0.6	96.9	94.3	91.7	90.5	93.7	97.2	99.0	99.9	100.0
	0.7	99.9	99.5	98.9	98.0	98.3	99.6	99.9	99.9	100.0
	0.8	100.0	99.9	99.7	99.9	99.9	99.9	99.8	100.0	100.0
	0.9	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

5개 지역의 변환한 월별 경제활동인구수는 아래와 같이 패널 AR(1)을 따른다.

$$X_t(j) = \phi_j X_{t-1}(j) + e_t(j), \quad t = 1, 2, \dots, 144, \quad j = 1, 2, \dots, 5, \quad (5.1)$$

( $j = 1$  : 충북,  $j = 2$  : 전남,  $j = 3$  : 경북,  $j = 4$  : 경남,  $j = 5$  : 제주).

동질성 검정을 위한 가설은 아래와 같다.

$$H_0 : \phi_1 = \phi_2 = \dots = \phi_5, \\ H_1 : \text{적어도 하나의 } \phi_j \text{는 같지 않다.} \quad (5.2)$$

통계프로그램 R을 이용하여 동질성 검정을 위한 Rao 통계량을 구한 결과  $Rao = 0.2008$ 으로  $\chi^2(4) = 9.4877$ 보다 작아 위 귀무가설을 기각할 수 없다. 따라서 변환한 월별 경제활동인구수 자료의 각 도별  $\phi_j$  값은 동질적이라고 결론지을 수 있다. 그러므로 시도 별로 각각의 모형으로 적합하여 분석하는 것보다 5개 시도의 자료를 하나로 집약한 모형을 이용하여 분석을 진행하는 것이 더 타당하다고 할 수 있다.

동질성 검정으로 각 패널들의 동질성을 보였다면 각 패널의 자기회귀계수  $\phi_j$  ( $j = 1, 2, \dots, m$ )를 하나의 공통 자기회귀계수로 표현할 수 있게 된다. 현실에서 실제 관측치보다 각 시점에서의 표본평균과 같은 집합체 데이터를 보다 편리하게 구할 수 있는 것이 사실이다. 오차항이 정규분포를 따른다면 표본평균은  $\phi$ 의 충분통계량이라는 점에 착안하여 각 시점에서의 표본평균을 이용하여 집합체 데이터를 형성한다. 각 시점에서의 표본평균을 이용하여 AT(1) 모형으로 적합시킨 모형은 다음과 같으며 오차항  $e_t$ 는

**Table 5.1.** MLE of  $\phi_j$ 

지역	충청북도	전라남도	경상북도	경상남도	제주도	집합체
$\phi_j$	0.7787	0.8099	0.7312	0.8088	0.7335	0.623

MLE = maximum likelihood estimators.

**Table 5.2.** Comparison of MSPE

	시계열 별 예측	평균 예측
MSPE	3.772959	2.085439

MSPE = mean square predicted error.

백색잡음을 따른다.

$$Y_t = \phi Y_{t-1} + e_t, \quad e_t \sim WN(0, \sigma^2), \quad (5.3)$$

$$Y_t = \bar{X}_t = \frac{1}{5} \sum_{j=1}^5 X_t(j), \quad t = 1, 2, \dots, 144.$$

각 지역별로 각각 AR(1)으로 적합하여 구한  $\phi_j$ 와 집합체 데이터로 구한  $\phi$ 의 최대우도추정량은 Table 5.1과 같다.

각 도의 패널 AR(1) 모형과 동질성 가정을 바탕으로 각 시점에서 표본평균으로 추정된 AR(1) 모형의 예측력을 비교하기 위하여 예측값과 실제값의 차를 제공하여 얻은 평균제곱예측오차(Mean Square Predicted Error; MSPE)를 구한 결과는 Table 5.2와 같다. Table 5.2의 결과를 보면 전체적으로 각 지역별로 AR(1)모형에 적합하여 예측한 MSPE보다 pooling한 데이터로 예측한 MSPE가 더 우수하다는 것을 알 수 있다.

## 6. 결론

패널 시계열 자료에서 독립적인 시계열들이 동일한 패턴을 가지고 있다면 각 시계열이 하나의 모집단에서 온 것으로 간주할 수 있다. 패널 시계열 자료의 동질성이 만족된다면 각각의 독립적인 시계열들을 개별적으로 분석하는 것보다 자료를 종합하여 분석하는 것이 모수절약 원칙에 의해 더욱 효율적이다.

본 논문에서는 패널의 수가 작은 패널 1차 자기회귀 모형의 동질성 검정에 대해 살펴 보았다. 모의실험을 통하여 패널의 수가 작은 패널 1차 자기회귀 모형의 두 가지 동질성 검정 통계량과 그 분포를 살펴본 결과, 패널의 수가 작을 때에도 Rao 통계량과 Wald 통계량이 카이제곱 분포를 따름을 보였다. 두 검정 통계량의 검정력을 비교한 바, 패널의 수가 작을 때 Rao 통계량이 더 바람직함을 보였다. 실제 지역별 월별 경제활동인구 자료를 활용하여 동질성 검정을 수행하고 각 시계열별로 적합한 모형과 시점별 표본 평균을 이용해 적합한 모형의 예측력을 비교하였다.

## References

- Anderson, T. W. (1978). Repeated measurement on autoregressive process, *Journal of American Statistical Association*, **73**, 371–378.
- Basawa, I. V., Billard, L., and Srinivasan, R. (1984). Large-sample tests of homogeneity for time series models, *Biometrika*, **71**, 203–206.
- Basawa, I. V. and Billard, L. (1989). Large-sample inference for a regression model with autocorrelated errors, *Biometrika*, **76**, 283–288.



- Glasbey, C. A. (1979). Correlated residuals in non-linear regression applied to growth data, *Applied Statistics*, **28**, 251–259.
- Kim, I., Hwang, S., and Lee, S. (2000). Homogeneity test of random coefficient for the first order nonlinear time series panel data, *The Korean Journal of Applied Statistics*, **13**, 97–104.
- Lee, S. D. (1993). Test of homogeneity for a panel of seasonal autoregressive processes, *The Journal of Korean Statistical Society*, **22**, 125–132.
- Lee, S. D. (2014). Test of Homogeneity for intermittent panel AR(1) processes and application, *The Journal of Korean Statistical Society*, **27**, 1163–1170.
- Lee, S. Y., Kim, S. W., and Lee, S. D. (2013). Test of homogeneity for panel bilinear time series model, *The Korean Journal of Applied Statistics*, **26**, 521–529.
- Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed., John Wiley & Sons, New York.

# 패널 1차 자기회귀과정들의 동질성 검정 통계량 비교

이성덕<sup>a1</sup> · 김선우<sup>a</sup> · 조나래<sup>a</sup>

<sup>a</sup>충북대학교 정보통계학과

(2015년 12월 15일 접수, 2015년 12월 28일 수정, 2015년 12월 28일 채택)

---

## 요약

패널 시계열 자료를 소개하고 패널 1차 자기회귀 모형을 고려하였다. 패널 1차 자기회귀 모형의 동질성 검정을 위한 검정 통계량으로 Rao 통계량과 Wald 통계량을 제안하고, 그 극한분포를 제시하였다. 모의실험을 통해 패널의 수가 작을 때에도 패널의 수가 많을 때와 마찬가지로 두 검정 통계량의 분포가 카이제곱분포를 따르는 것을 확인하였으며, 패널의 수가 작을 때 Rao 통계량이 Wald 통계량 보다 더 우수한 검정력을 가짐을 모의실험을 통해 확인하였다. 시도별 월별 경제활동인구수 자료를 패널 1차 자기회귀 모형으로 적합하여 동질성 검정을 수행한 결과 동질성을 만족하였다. 동질성 검정을 만족한 자료를 시점별 평균을 이용하여 종합하고 이를 1차 자기회귀모형으로 적합하였다. 각각의 시도별로 적합한 모형과 시점별 평균을 이용하여 적합한 모형의 예측력을 비교한 결과 동질성 검정을 통과한 패널 1차 자기회귀모형의 경우 자료를 종합하여 적합한 모형의 예측력이 더 우수함을 확인하였다.

주요용어: 패널 1차 자기회귀 모형, 동질성 검정, Rao 검정 통계량, Wald 검정 통계량, 지역별 경제활동인구수

---

이 논문은 2014년도 충북대학교 학술연구지원사업의 연구비 지원에 의하여 연구되었음.

<sup>1</sup>교신저자: (28644) 충북 청주시 서원구 충대로 1(개신동), 충북대학교 정보통계학과.

E-mail: sdlee@chungbuk.ac.kr