

# Test of homogeneity for transition probabilities in panel Markov chains

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

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

(Received December 13, 2016; Revised December 27, 2016; Accepted December 27, 2016)

---

## Abstract

The test of transition probabilities in panel Markov chains are introduced. We deal with the hypotheses whether panel Markov chains have the same transition probabilities or not for all times. We suggest a LR test statistic for the test and its limit distribution is derived. We perform a simulation study to examine the limit distribution of test statistics when the number of the individuals are large.

Keywords: panel Markov chains, test of homogeneity, transition probability, LR statistic, simulation

---

## 1. 서론

하나의 현상에 대해 여러 개의 독립적인 마코프 체인으로 구성된 자료는 실제 현상에서 자주 나타나게 되는데 이를 패널 마코프 자료라 한다. 이때 개의 독립적인 마코프 체인이 생기는데, 이러한 자료를 분석할 때 먼저 마코프 체인들이 동일한 패턴을 갖고 있는지 확인한다. 만약 동일한 패턴을 가지고 있다면 각각의 마코프 체인들이 하나의 모집단에서 온 것으로 간주할 수 있으므로 각각의 독립적인 마코프 체인을 분석하는 것보다 자료를 종합하여 분석하는 것이 모수절약 원칙(the principle of parsimony)에 의해 더욱 효율적이다.

독립적인 마코프 체인들이 동일한 패턴을 가지고 있다면 하나의 모집단에서 온 것으로 간주할 수 있는데 Anderson과 Goodman (1957)은 이러한 동일한 패턴을 가지고 있는 패널 마코프 체인에 대하여 통계적으로 추론하였다. 이들은 패널 마코프 체인의 전이확률이 시간에 대하여 동일한지를 검정하기 위한 검정통계량이 패널의 수가 클 때 카이제곱분포를 따른다는 것을 증명하였다.

경제학의 다양한 분야에서 마코프 체인을 적용하여 통계적 추론을 거듭해 왔다. 효율적 시장가설에 대한 초반의 연구는 일일 또는 주간 주식 자료를 이용하여 단기 주식 수익률이 확률보행과정을 따른다고 주장하였다. 그러나 이러한 연구는 확률보행과정을 바탕으로 하기 때문에 과거와 현재의 주식 수익률 사이의 선형 종속성을 가정하는 문제점이 있다. 이를 보완하여 Niederhoffer와 Osborne (1966)은 1964년 10월, 22일의 거래일 동안 Dow Jones 산업 평균 지수에 포함된 7개 주식의 거래 가격의 변화를 마코프 체인에 적용하였다. 7개의 상태로 구성된 1차 마코프 체인을 고려하였으며 이들은 확률보행

---

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2016R1D1A3B03932557).

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

**Table 2.1.** Data structure for panel time finite Markov chains

$t$	0	1	2	...	$T$	
1 <sup>th</sup> individual $X_t(1)$	$X_0(1)$	$X_1(1)$	$X_2(1)$	...	$X_T(1)$	$\underline{X}(1)$
2 <sup>th</sup> individual $X_t(2)$	$X_0(2)$	$X_1(2)$	$X_2(2)$	...	$X_T(2)$	$\underline{X}(2)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$n^{\text{th}}$ individual $X_t(n)$	$X_0(n)$	$X_1(n)$	$X_2(n)$	...	$X_T(n)$	$\underline{X}(n)$

과정을 기각하고 전이확률이 정상적임을 보였다. Fielitz와 Bhargava (1973)는 1963년부터 1968년까지 200개 패널의 일일 주간 수익률을 사용하여 3가지 상태를 반영한 마코프 체인을 고려하였다. 이들은 200개의 패널을 하나의 패널 자료가 아닌 개별적으로 마코프 체인으로 적용하여야 함을 보였고, 시간에 따라 주식 수익률의 전이확률이 정상적이지 않음을 확인하였다. McQueen과 Thorley (1991)는 1947년부터 1987년동안 뉴욕증권거래소의 모든 주식의 연간 실질수익률과 초과수익률, 그리고 가치가증된 포트폴리오를 마코프 체인에 적용하였으나 검정 절차를 거치지 않고 암묵적으로 전이확률의 정상성을 가정하였다.

사전 연구들은 패널 자료를 개별적인 마코프 체인으로 전이확률의 정상성을 검정하거나 전이확률의 정상성에 대한 검정 없이 패널 마코프 체인을 적용하였다. 이에 본 논문은 패널 마코프 체인을 소개하고, 패널 마코프 체인의 전이확률의 동질성 검정통계량의 성질에 대하여 모의실험을 하고자 한다. 2장에서 패널 마코프 체인을 소개하고 우도함수로 전이확률을 추정한다. 3장에서는 전이확률의 동질성 검정 가설을 소개하고 검정통계량의 극한분포를 유도한다. 4장에서 모의실험을 통해 실제 패널의 수에 따른 검정통계량의 분포를 살펴보고 5장에서 결론을 맺는다.

## 2. 패널 마코프 체인

$\{X_t(k); t = 1, 2, \dots, T, k = 1, 2, \dots, n\}$ 는 상태공간  $S = \{1, 2, \dots, m\}$ 와 전이확률 행렬  $P = ((p_{ij}))$ ,  $i, j = 1, 2, \dots, m$ 를 가지는 정규 마코프 체인(Markov chains)이라 한다.  $n$ 개의 개체(individuals)를 반영한 패널 마코프 체인 자료의 구조는 Table 2.1과 같다.

$\{X_t(k)\}$ 는  $k$ 번째 개체의 시점  $t$ 에서 관측값으로  $n$ 개의 개체와  $T$ 개의 시점이 반영되었다.  $n$ 개의 개체  $\underline{X}(1), \underline{X}(2), \dots, \underline{X}(n)$ 는 모두 독립이고 전이확률  $P = ((p_{ij}))$ 를 가지는  $n$ -독립 연쇄( $n$ -independent chain)이다. 각 개체  $\underline{X}(1), \underline{X}(2), \dots, \underline{X}(n)$ 는 i.i.d.로 공통 밀도  $\prod_{t=1}^T p_{x_{t-1}x_t}$ 를 가진다고 가정한다. 전이확률  $p_{ij} = P(X_t = j | X_{t-1} = i)$ 는 마코프 성질에 따라 다음과 같이 정리 할 수 있다.

$$\begin{aligned}
 & P(X_t = x_t, X_{t-1} = x_{t-1}, \dots, X_0 = x_0) \\
 &= P(X_0 = x_0) \times P(X_1 = x_1 | X_0 = x_0) \times P(X_2 = x_2 | X_1 = x_1, X_0 = x_0) \\
 &\quad \times \dots \times P(X_t = x_t | X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2}, \dots, X_1 = x_1, X_0 = x_0) \\
 &= P(X_0 = x_0) p_{x_0 x_1} p_{x_1 x_2} \dots p_{x_{t-1} x_t} \\
 &= P(X_0 = x_0) \prod_{t=1}^T p_{x_{t-1} x_t}.
 \end{aligned}$$

전이확률의 마코프 성질을  $n$ 개의 개체에 대해 고려한 우도함수는 다음과 같다.

$$\begin{aligned} L_H &= \prod_{k=1}^n \prod_{t=1}^T p_{x_{t-1}(k) x_t(k)} \\ &= \prod_{t=1}^T \prod_{i,j=1}^m p_{ij}^{n_{ij}(t)} \\ &= \prod_{i,j=1}^m p_{ij}^{\sum_{t=1}^T n_{ij}(t)}, \end{aligned} \quad (2.1)$$

여기서  $n_{ij}(t)$ 는  $X_{t-1}(k) = i$ 이고  $X_t(k) = j$ 인 개체의 개수이다. 우도함수를 살펴보면 제한조건  $\sum_{j=1}^m p_{ij} = 1$ ,  $i = 1, 2, \dots, m$  하에서  $\ln L$ 을 최대로 하는 전이확률 행렬  $P = ((p_{ij}))$ 의 충분통계량은 가산행렬  $N = ((n_{ij}))$ 임을 알 수 있다. 라그랑지 승수 형태  $\phi = \ln L - \sum_{i=1}^m \lambda_i (\sum_{j=1}^m p_{ij} - 1)$ 와  $\partial\phi/\partial p_{ij} = 0$ ,  $\partial\phi/\partial \lambda_i = 0$ 을 동시에 풀어 다음과 같이  $p_{ij}$ 의 최대가능도추정량(MLE)을 구할 수 있다.

$$\widehat{p}_{ij} = \begin{cases} \frac{n_{ij}}{n_{i\cdot}}, & n_{i\cdot} > 0, \\ \delta_{ij}, & n_{i\cdot} = 0, \end{cases} \quad (2.2)$$

여기서

$$n_{i\cdot} = \sum_{j=1}^m n_{ij}, \quad \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j \end{cases}$$

이다.

실현치  $\underline{a}$ 를  $\underline{a} = (a_0, a_1, a_2, \dots, a_T)$ ,  $a_i \in S$ 라 하면  $k$ 번째 개체는

$$y_{\underline{a}}(k) = \begin{cases} 1, & X_0(k) = a_0, X_1 = a_1, \dots, X_T = a_T, \\ 0, & \text{otherwise} \end{cases}$$

로 정의할 수 있다.  $\underline{a}$ 인 개체수는  $n_{\underline{a}} = \sum_{k=1}^n y_{\underline{a}}(k)$ 로 나타낼 수 있고,  $\{n_{\underline{a}}\}$ 는 index  $n_{a_0}(0)$ 와 확률  $p_{a_0 a_1} p_{a_1 a_2} p_{a_2 a_3} \dots p_{a_{t-1} a_t}$ 를 가지는 다항분포이다.  $n_{u;ij}(t)$ 는  $a_0 = u$ 이고  $a_{t-1} = i$ ,  $a_t = j$ 인 개체수로  $n_{u;ij}(t) = \sum_{\underline{a}} n_{\underline{a}}$ 이다. 그리고  $\{n_{u;ij}(t)\}$ 는 다항분포를 따르는  $n_{\underline{a}}$ 의 합으로 index  $n_u(0)$ 와 확률  $p_{ui}^{(t-1)} p_{ij}$ 를 갖는 다항분포이다. 여기서  $p_{ui}^{(t-1)} p_{ij}$ 는  $X_0 = u$ 이고  $X_{t-1} = i$ ,  $X_t = j$ 일 확률이다.  $n_{u;i\cdot}(t)$ 는  $\sum_{j=1}^m n_{u;ij}(t)$ 로 다항분포의 성질에 의하여  $\{n_{u;ij}(t)|n_{u;i\cdot}(t)\}$  역시 index  $n_{u;i\cdot}(t) = \sum_{j=1}^m n_{u;ij}(t)$ 와 확률  $p_{ij}$ 를 갖는 다항분포이다. 따라서 새로운 확률변수  $Z_{u;ij}(t)$ 를 다음과 같이 고려하였다.

$$\begin{aligned} Z_{u;ij}(t) &= n_{u;ij}(t) - E(n_{u;ij}(t)|n_{u;i\cdot}(t)) \\ &= n_{u;ij}(t) - n_{u;i\cdot}(t)p_{ij}. \end{aligned}$$

새로운 확률변수  $\{Z_{u;ij}(t)\}$ 의 평균과 공분산 구조는 다음과 같다.

$$\begin{aligned} E(Z_{u;ij}(t)) &= 0, \\ \text{Var}(Z_{u;ij}(t)) &= n_u(0)p_{ui}^{(t-1)}p_{ij}(1-p_{ij}), \\ \text{Cov}(Z_{u;ij}(t), Z_{u;ij'}(t)) &= -n_u(0)p_{ui}^{(t-1)}p_{ij}p_{ij'}, \quad j \neq j', \\ \text{Cov}(Z_{ij}(t), Z_{i'j'}(s)) &= 0, \quad i \neq i' \text{ or } t \neq s. \end{aligned}$$

$Z_{u;ij}(t)$ 에서 가능한 모든 초기값  $u$ 를 고려하여 다음과 같은 새로운 확률변수  $Z_{ij}(t)$ 를 고려하였다.

$$\begin{aligned} Z_{ij}(t) &= \sum_{u=1}^m Z_{u;ij}(t) \\ &= n_{ij}(t) - n_{i \cdot}(t)p_{ij}. \end{aligned}$$

$Z_{ij}(t)$ 는 다음과 같은 성질을 갖는다.

$$\begin{aligned} E(Z_{ij}(t)) &= 0, \\ \text{Var}(Z_{ij}(t)) &= \left\{ \sum_{u=1}^m n_u(0)p_{ui}^{(t-1)} \right\} p_{ij}(1 - p_{ij}), \\ \text{Cov}(Z_{ij}(t), Z_{i'j'}(t)) &= - \left\{ \sum_{u=1}^m n_u(0)p_{ui}^{(t-1)} \right\} p_{ij}p_{i'j'}, \quad j \neq j', \\ \text{Cov}(Z_{ij}(t), Z_{i'j'}(s)) &= 0, \quad i \neq i' \text{ or } t \neq s. \end{aligned} \tag{2.3}$$

### 3. 패널 마코프 체인의 동질성 검정을 위한 통계량과 극한분포

동질성 검정을 위하여  $n \rightarrow \infty$ 일 때  $\sqrt{n}(\widehat{p}_{ij} - p_{ij})$ 의 극한분포를 구하기 위하여 다음과 같이 정리하였다.

$$\begin{aligned} \sqrt{n}(\widehat{p}_{ij} - p_{ij}) &= \sqrt{n} \left( \frac{n_{ij}}{n_{i \cdot}} - p_{ij} \right) \\ &= \sqrt{n} \left( \frac{\sum_{t=1}^T (n_{ij}(t) - n_{i \cdot}(t)p_{ij})}{\sum_{t=1}^T n_{i \cdot}(t)} \right) \\ &= \frac{\frac{1}{\sqrt{n}} \sum_{t=1}^T Z_{ij}(t)}{\frac{1}{n} \sum_{t=1}^T n_{i \cdot}(t)}. \end{aligned} \tag{3.1}$$

먼저 식 (3.1)의 분모는 다음과 같이 상수  $\phi_i$ 에 확률적으로 수렴한다.

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^T n_{i \cdot}(t) &= \frac{1}{n} \sum_{t=1}^T \sum_{u=1}^m \sum_{j=1}^m n_{u;ij}(t) \\ &= \sum_{t=1}^T \sum_{u=1}^m \sum_{j=1}^m \left( \frac{n_u(0)}{n} \right) \left( \frac{n_{u;ij}(t)}{n_u(0)} \right) \\ &\xrightarrow{p} \sum_{t=1}^T \sum_{u=1}^m \sum_{j=1}^m (\eta_u) (p_{ui}^{(t-1)} p_{ij}) \\ &= \sum_{t=1}^T \sum_{u=1}^m \eta_u p_{ui}^{(t-1)} = \phi_i, \end{aligned}$$

여기서  $0 < \eta_u < 1$ 이고  $\sum_{u=1}^m \eta_u = 1$ 이다.

식 (3.1)의 분자는 확률변수  $Z_{ij}(t)$ 의 선형결합으로  $R_{ij} = 1/\sqrt{n} \sum_{t=1}^T Z_{ij}(t)$ 라 한다.  $Z_{ij}(t)$ 는 다항분포  $\{n_{u;ij}(t)\}$ 의 선형변환으로 이루어져 있으므로 다항분포의 정규근사를 통해 결국  $R_{ij}$ 는  $m^2$ -다변량 정규분포로 분포 수렴하며 평균과 공분산 구조는 다음과 같다.

$$\begin{aligned} E(R_{ij}) &= 0, \\ \lim_{n \rightarrow \infty} \text{Var}(R_{ij}) &= p_{ij}(1 - p_{ij})\phi_i, \\ \lim_{n \rightarrow \infty} \text{Cov}(R_{ij}, R_{i'j'}) &= -\delta_{ii'}p_{ij}p_{i'j'}\phi_i, \quad j \neq j'. \end{aligned} \quad (3.2)$$

$\sqrt{n}(\widehat{p}_{ij} - p_{ij})$ 의 분자  $R_{ij}$ 는  $m^2$ -다변량 정규분포로 분포 수렴하고, 분모는 상수  $\phi_i$ 에 확률적으로 수렴하므로  $\sqrt{n}(\widehat{p}_{ij} - p_{ij})$ 는  $m^2$ -다변량 정규분포로 분포수렴한다.  $\sqrt{n\phi_i}(\widehat{p}_{ij} - p_{ij})$ 의 극한분포는  $R_{ij}/\sqrt{\phi_i}$ 의 극한분포와 같으므로  $\sqrt{n\phi_i}(\widehat{p}_{ij} - p_{ij})$ 는  $m^2$ -다변량 정규분포를 극한분포로 갖고 평균과 분산, 공분산은 다음과 같다.

$$\begin{aligned} E\left(\frac{R_{ij}}{\sqrt{\phi_i}}\right) &= 0, \\ \lim_{n \rightarrow \infty} \text{Var}\left(\frac{R_{ij}}{\sqrt{\phi_i}}\right) &= p_{ij}(1 - p_{ij}), \\ \lim_{n \rightarrow \infty} \text{Cov}\left(\frac{R_{ij}}{\sqrt{\phi_i}}, \frac{R_{i'j'}}{\sqrt{\phi_{i'}}}\right) &= -\delta_{ii'}p_{ij}p_{i'j'}, \quad j \neq j'. \end{aligned} \quad (3.3)$$

$i$ 가 고정되어 있다면  $\{\sqrt{n_{i\cdot}(t)}(\widehat{p}_{ij}(t) - p_{ij}(t)), j = 1, 2, \dots, m\}$ 은 평균이 0이고 분산이  $p_{ij}(t)(1 - p_{ij}(t))$ 이며 공분산이  $-\delta_{ii'}p_{ij}(t)p_{i'j'}(t)$ ,  $j \neq j'$ 인  $m$ -다변량 정규분포를 따름을 알 수 있다.

패널 마코프 체인의 시간에 대한 동질성 검정을 위한 가설은 다음과 같다.

$$H : p_{ij}(t) = p_{ij}, \quad \text{for all } t. \quad (3.4)$$

패널 마코프 체인의 전이확률이 시간에 따라 동일하다는 귀무가설을 검정하게 되는데, 이는 패널 마코프 체인이 정상적이라는 의미와 일맥상통한다. 가설 검정을 위한 비제한적(unrestricted) 우도함수와 최대우도추정량은 다음과 같다.

$$\begin{aligned} L &= \prod_t \prod_{i,j} \widehat{p}_{ij}(t)^{n_{ij}(t)}, \\ \widehat{p}_{ij}(t) &= \frac{n_{ij}(t)}{n_{i\cdot}(t)}. \end{aligned} \quad (3.5)$$

귀무가설 하에서의 제한된 우도함수 식 (2.1)로 구한 제한된 최대우도추정량 식 (2.2)를 이용하여 구한 LR 검정통계량은 다음과 같다.

$$\text{LR} = -2 \ln \lambda_n = -2 \ln \frac{\max L_H}{\max L} = -2 \ln \prod_t \prod_{i,j} \left( \frac{\widehat{p}_{ij}}{\widehat{p}_{ij}(t)} \right)^{n_{ij}(t)}. \quad (3.6)$$

귀무가설 하에서  $\text{LR} = -2 \ln \lambda_n$ 의 극한분포는 자유도가  $m(T - 1)(m - 1)$ 인  $\chi^2$  분포에 따른다.

#### 4. 모의실험

통계프로그램 R을 사용하여 패널 마코프 체인의 패널의 수가 많을 때 동질성 검정을 위한 검정통계량인 카이제곱분포를 따르는지 모의실험 하였다. Tan과 Yilmaz (2002)가 제안한 방법으로 패널 마코프 체인을 생성한다. 먼저 AR(1)을 따르는 패널의 수가  $n$ 이고 시점이  $T$ 인  $\{Y_t(k); t = 0, 1, 2, \dots, T; k = 1, 2, \dots, n\}$ 를 난수생성한다.

$$Y_t(k) = \phi Y_{t-1}(k) + e_t, \quad e_t \sim N(0, 1).$$

이때  $\{Y_t(k)\}$ 가 평균보다 작으면 Down, 평균보다 크다면 Up이라 표시하는 새로운 확률변수  $\{X_t(k); t = 0, 1, 2, \dots, T; k = 1, 2, \dots, n\}$ 를 고려한다.

$$X_t(k) = \begin{cases} D, & Y_t(k) \geq \bar{Y}_k, \\ U, & Y_t(k) < \bar{Y}_k, \end{cases} \quad \bar{Y}_k = \frac{1}{T+1} \sum_{t=0}^T Y_t(k),$$

여기서  $\{X_t(k)\}$ 는 상태공간  $S = \{U, D\}$ 의 원소를 값으로 가지는 패널 마코프 체인이 되며 다음과 같은 전이확률을 갖는다.

$$\underline{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}, \quad \text{for } 1 = D, 2 = U.$$

시간에 대한 전이확률의 동질성 검정 가설은 식 (3.3)과 같다.

모의실험을 위하여 1차 자기회귀계수  $\phi$ , 시점  $T$ , 패널의 수  $n$ 을 각각 달리하여 패널 마코프 체인을 1,000번 반복하여 생성하고, 시간에 대한 전이확률의 동질성 검정을 위한 LR 검정통계량을 각각 1,000개씩 계산하였다.

Figure 4.1, Figure 4.2, Figure 4.3은 각 조건에서 LR 검정통계량의 분포와 카이제곱분포 곡선을 함께 그린 그림이다. 검정색 곡선은 난수발생으로 생성한 패널 마코프 체인으로 계산한 LR 검정통계량의 분포이고 빨간색 곡선은 자유도가  $2 \times (T - 1) \times (2 - 1)$ 인 카이제곱분포 곡선이며 파란색 선은  $\chi_{0.95}^2$  값을 나타낸다. 1차 자기회귀계수  $\phi$ 가 각각 0.1과 0.5로 패널 마코프 체인이 정상적인 경우, 패널의 수가 50개 이상일 때 LR 검정통계량의 분포 곡선과 이론적인 분포 곡선이 매우 유사함을 확인할 수 있다. 그러나 패널의 수가 10개, 25개인 경우, 비록 패널 마코프 체인이 정상적이라 하더라도 LR 검정통계량의 분포 곡선이 이론적인 분포 곡선보다 왼쪽으로 치우쳐 있음을 확인하였다. 따라서 정상적인 패널 마코프 체인이라도 패널의 수가 50개 이상일 때 동질성 검정을 위한 검정통계량의 분포가 자유도가  $2 \times (T - 1) \times (2 - 1)$ 인  $\chi^2$  분포를 따른다고 할 수 있다. 패널의 수가 작을 때에는 패널 마코프 체인에 대하여 동질성 검정통계량이  $\chi^2$  분포를 따른다고 할 수 없으므로 동질성 검정의 수행은 바람직하지 않다. 또한 1차 자기회귀계수  $\phi$ 가 0.9로 패널 마코프 체인이 비정상에 가까운 경우 실험한 모든 패널의 수에서 LR 검정통계량의 분포 곡선이  $\chi^2$  분포 곡선보다 왼쪽으로 치우쳐 있음을 확인할 수 있었다.

Table 4.1은 1차 자기회귀계수  $\phi$ 와 패널의 수  $n$ 을 달리하여 발생시킨 패널 마코프 체인으로 각각 1,000개의 LR 통계량 계산하여 귀무가설을 채택하는 통계량의 비율을 나타낸 표이다.  $\phi$ 가 0.1과 0.5로 정상적인 경우 패널의 수  $n$ 이 커질수록 귀무가설을 채택하는 비율이 1에 가까워지는 것을 확인할 수 있다. 반대로 1차 자기회귀계수  $\phi$ 가 0.9로 패널 마코프 체인이 비정상에 가깝거나 1차 자기회귀계수  $\phi$ 가 1로 확률보행과정인 경우 패널 마코프 체인이 비정상적인 경우 패널의 수가 커질수록 0에 가까워진다. 패널 마코프 체인에서 전이확률의 동질성 검정에 대하여 체인이 정상적인 경우 귀무가설을 채택하고 비정상적인 경우 귀무가설을 기각하여 동질성 검정이 정상적으로 작동함을 확인하였다.

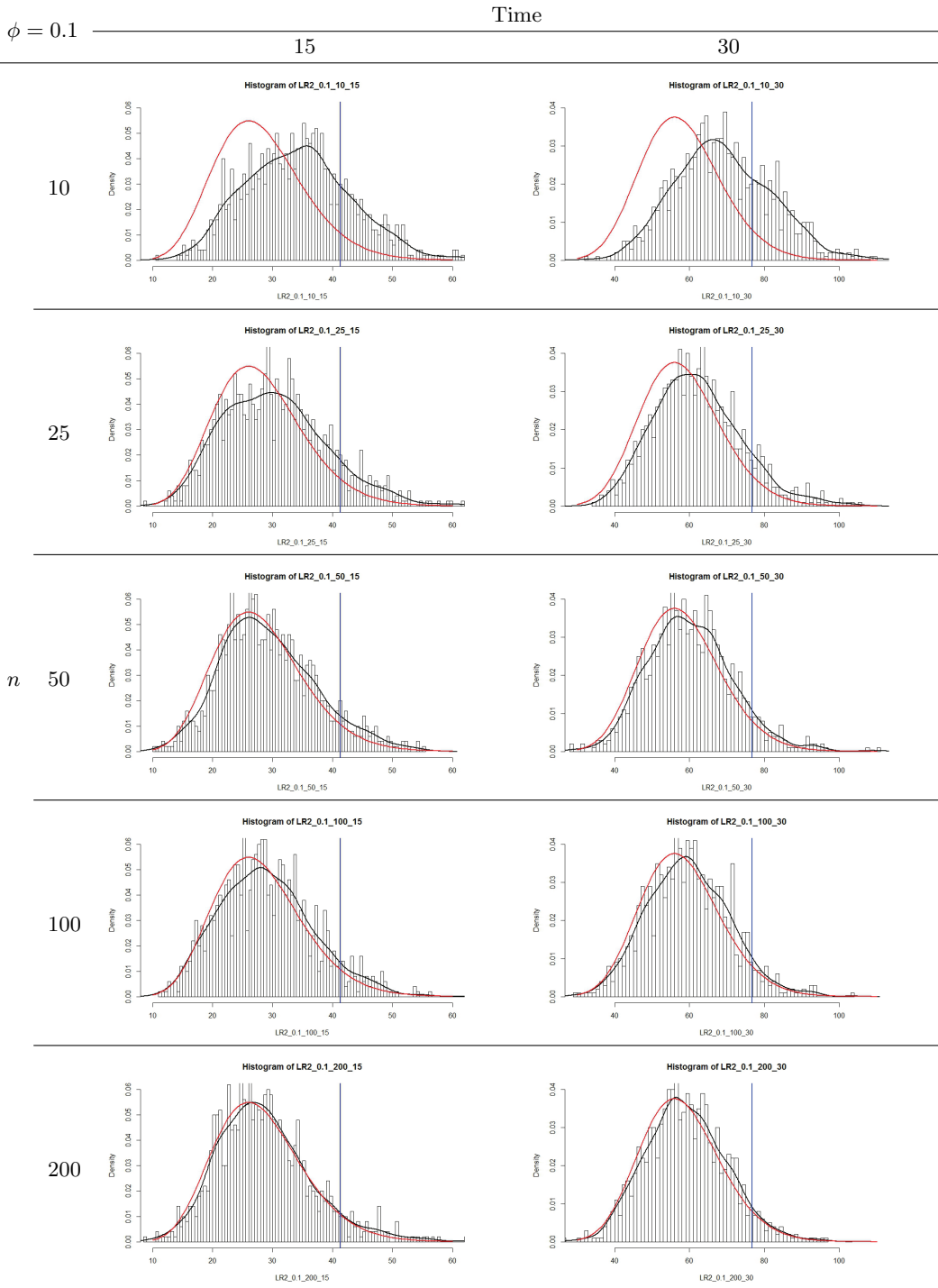


Figure 4.1. Distributions of LR statistics ( $\phi = 0.1$ ).

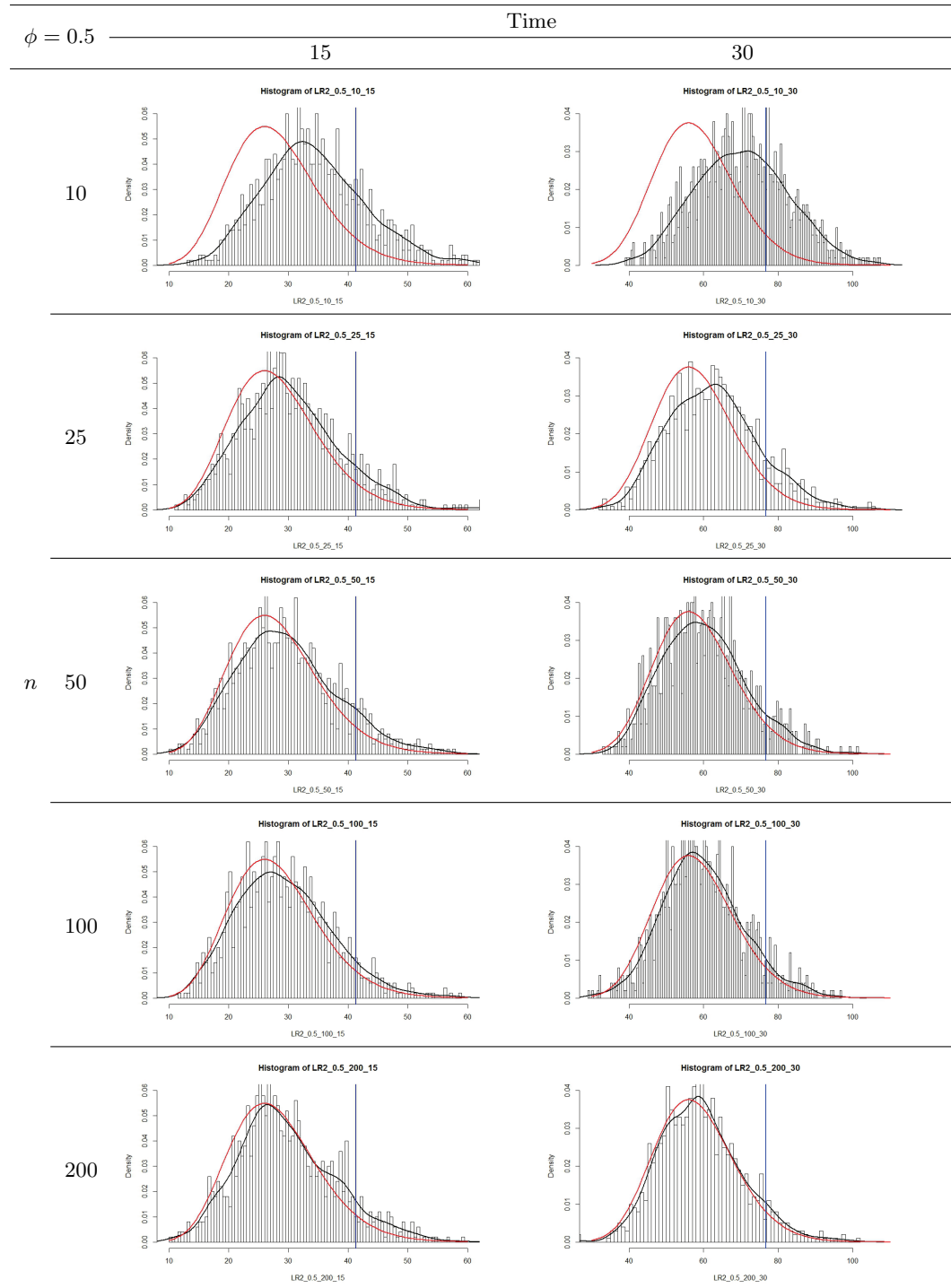


Figure 4.2. Distributions of LR statistics ( $\phi = 0.5$ ).



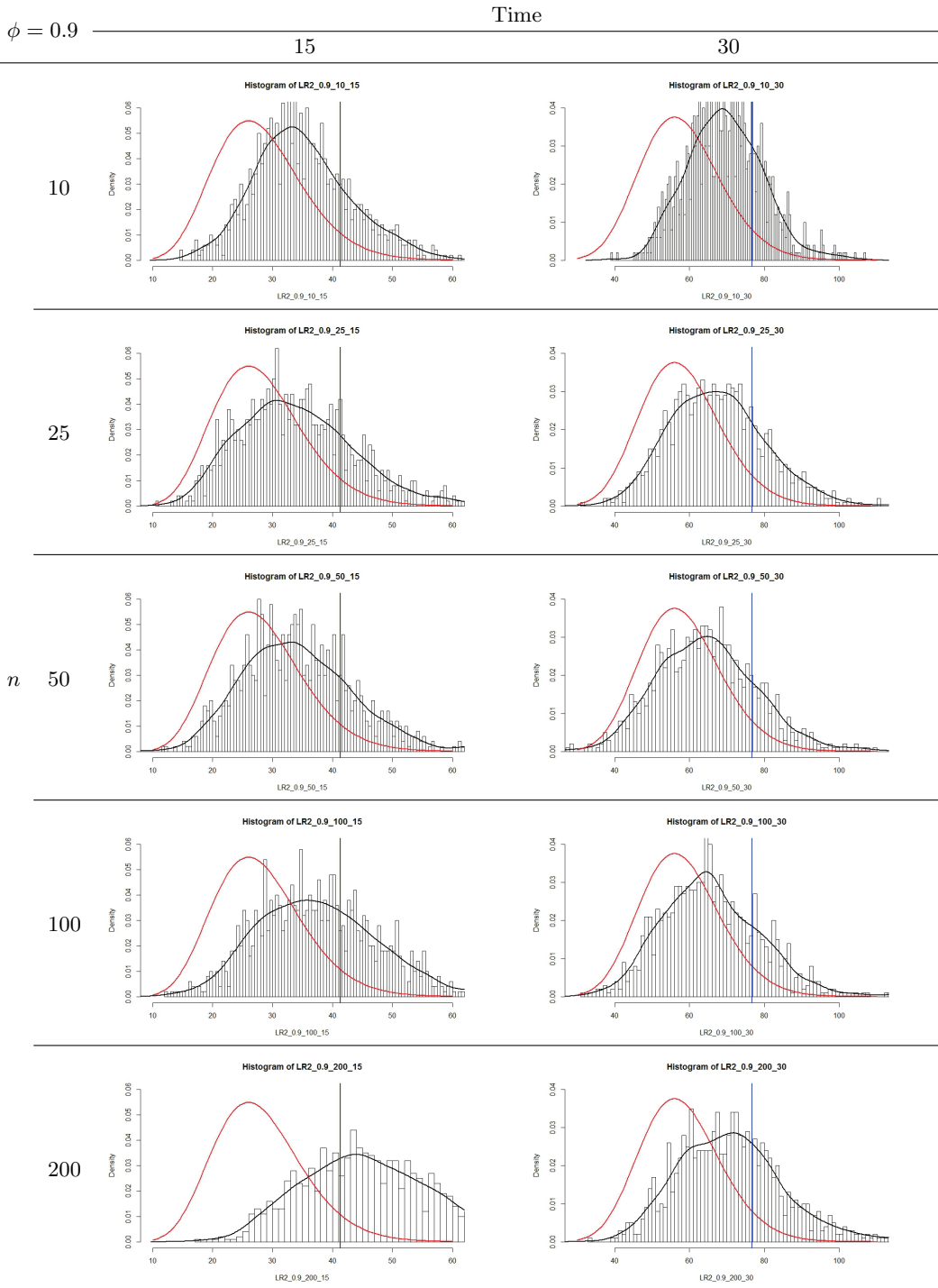


Figure 4.3. Distributions of LR statistics ( $\phi = 0.9$ ).

**Table 4.1.** Empirical acceptance of LR statistics

$\phi$	$n$							
	5	10	15	25	50	100	200	400
0.1	0.554	0.728	0.847	0.892	0.928	0.942	0.947	0.952
0.5	0.633	0.694	0.815	0.875	0.920	0.932	0.932	0.934
0.9	0.943	0.774	0.758	0.779	0.824	0.795	0.686	0.325
1.0	0.986	0.824	0.711	0.649	0.600	0.365	0.047	-

## 5. 결론

본 논문에서는 패널의 수에 따른 패널 마코프 체인의 전이확률의 동질성 검정을 수행하였다. 모의실험을 통해 시간에 대한 전이확률의 동질성 검정을 위한 LR 검정통계량의 분포를 살펴본 결과 패널의 수가 50개이상인 경우 검정통계량이 카이제곱분포를 따름을 보였다. 하지만 패널의 수가 작을 때에는 LR 검정통계량이 카이제곱분포를 따른다고 할 수 없었다. 또한 시간에 대한 전이확률의 동질성 검정은 패널 마코프 체인이 정상적일 때와 비정상적일 때 모두 바람직하게 작동함을 확인하였다.

## References

- Anderson, T. W. and Goodman, L. A. (1957). Statistical inference about Markov chains, *Annals of Mathematical Statistics*, **28**, 12–40.
- Fielitz, B. D. and Bhargava, T. N. (1973). The behavior of stock-price relatives: a Markovian analysis, *Operation Research*, **21**, 1183–1199.
- McQueen, G. and Thorley, S. (1991). Are stock returns predictable? A test using Markov chains, *Journal of Finance*, **46**, 239–263.
- Niederhoffer, V. and Osborne, M. (1966). Market making and reversal on the stock exchange, *Journal of the American Statistical Association*, **61**, 897–916.
- Tan, B. and Yilmaz, K. (2002). Markov Chain test for time dependence and homogeneity: an analytical and empirical evaluation, *European Journal of Operational Research*, **137**, 524–543.

# 패널 마코프 체인의 전이확률에 대한 동질성 검정

이성덕<sup>a,1</sup> · 조나래<sup>a</sup>

<sup>a</sup>충북대학교 정보통계학과

(2016년 12월 13일 접수, 2016년 12월 27일 수정, 2016년 12월 27일 채택)

---

## 요약

패널 마코프 체인의 구조를 소개하고 우도함수를 도출하여 전이확률을 추정하였다. 패널 마코프 체인의 전이확률의 동질성 검정통계량으로 LR 통계량을 제안하고 그 극한분포를 제시하였다. 동질성 검정통계량의 극한분포를 패널의 수를 달리하여 모의실험하였으며 패널의 수가 50개 이상인 경우 동질성 검정통계량의 분포가 카이제곱분포를 따르는 것을 확인하였다. 정상적인 경우 검정통계량이 우수한 검정력을 가지는 것을 보였고, 확률보행과정과 같이 비정상적인 경우 검정통계량이 전이확률의 비동질성을 잘 반영하는 것을 확인하였다.

주요용어: 패널 마코프 체인, 동질성 검정, 전이확률, LR 통계량, 모의실험

---

---

이 논문은 2016년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 이공학 개인기초연구지원사업 연구비 지원에 의하여 연구되었음 (NRF-2016R1D1A3B03932557).

<sup>1</sup>교신저자: (28644) 충북 청주시 서원구 충대로 1, 충북대학교 정보통계학과. E-mail: sdlee@chungbuk.ac.kr