

구조방정식모형(SEM)을 활용한 중국 이혼율 영향 요인 분석 연구*

후사정¹, 김유정², 김준홍³, 최용석⁴

요 약

본 연구는 세계 최대 인구 보유국인 중국이 최근 이혼 문제로 고심하고 있는 점을 고려해 급격하게 증가하고 있는 이혼율에 영향을 미치는 구체적인 요소를 파악하고자 하였다. 1987년 0.5(%, 1000명당 이혼 건수)였던 중국의 조이혼율(이하 이혼율로 통칭)은 2019년 3.4로 17년 연속 상승하고 있는 추세를 보인다. 중국의 이혼율에 관한 연구는 다양하게 진행되어 왔으나 구조방정식모형을 활용한 연구는 다소 미흡한 상태이다. 이에 본 연구에서는 중국 전체 지역을 기준으로 이혼율에 미치는 영향을 파악하고 난 후 지역별로 이혼율 분포를 파악하여 이혼율이 높은 지역, 평균인 지역, 낮은 지역 간 이혼율 상승 원인과 지리적인 차이를 비교 분석하고자 한다. 이를 위해 National Bureau of Statistics of China, Global Data Lab 그리고 CEI Data의 오픈 자료로부터 중국의 2001년부터 2019년까지의 이혼율과 상관이 높은 GDP, 소비수준, 여성 수입 및 교육수준, 출산율 등의 총 12개 변수를 선택하였고, 기술통계분석, 다변량 분산분석(MANOVA)과 프로파일 분석(profile analysis, PA)을 통해 이혼율이 높은 지역(헤이룽장성:黑龍江), 평균 지역(산시성:陝西), 낮은 지역(광둥성:廣東) 총 3개 집단 간에 차이가 있는지 확인하였다. 그리고 탐색적 요인 분석을 통해 이혼율에 미치는 요인을 분석했다. 끝으로 구조방정식모형의 분석에서 이혼율이 높은 헤이룽장성과 평균인 산시성, 낮은 광둥성 간에 경제사회 개발화 특성 및 현대가족 특성이 이혼율과 출산율에 미치는 영향을 검증하여 결론을 내리고자 한다. 따라서 본 연구에서 언급된 영향 외에도 이혼율에 영향을 미치는 이유는 여러 가지 있을 것으로 여겨진다.

주요용어 : 구조방정식모형, 이혼율, 잠재변수, 프로파일분석, 다변량분산분석.

1. 서론

본 연구는 세계 최대 인구 보유국인 중국이 최근 이혼 문제로 고심하고 있는 점을 고려해 급격하게 증가하고 있는 이혼율에 영향을 미치는 구체적인 요소를 파악하고자 하였다. Wang(2001)은 중국의 이혼율이 경제 전환기에 따라 증가할 가능성이 높지만 그 증가율은 곧 결혼 및 가족법을 포함하여 많은 요인에 의해 결정될 것을 시사하였고, Wang, Qin(2010)은 중국의 급속한 경제 성장과 이혼 및 재혼율의 현저한 증가는 경제 전환기의 이혼 및 재혼 패턴을 연구하였는데, 이혼율 및

*본 논문은 제1저자 후사정의 석사학위논문의 축약본입니다.

¹46241 부산광역시 금정구 부산대학교 63번길 2, 부산대학교 통계학과 석사과정. E-mail: huxiajingwyx@163.com

²46241 부산광역시 금정구 부산대학교 63번길 2, 부산대학교 통계학과 석사과정. E-mail: yj_kim13@naver.com

³46241 부산광역시 금정구 부산대학교 63번길 2, 부산대학교 통계학과 석사과정. E-mail: wnsghd4685@naver.com

⁴(교신저자)46241 부산광역시 금정구 부산대학교 63번길 2, 부산대학교 통계학과 전임교수.

E-mail: yschoi@pusan.ac.kr

[접수 2022년 11월 4일; 수정 2022년 12월 2일; 게재확정 2022년 12월 5일]

재혼율의 변화가 지역적 요인, 1인당 소득 및 교육 수준과 관련이 있음을 시사하고 있다. Chung(2007) 및 Baek, Ryu(2012), Kim, Choe, Jeon(2015)은 여성의 경제적 지위 변화 및 결혼 후 일과 가족을 조화하는 데서 오는 어려움, 결혼에 대한 의미와 태도변화, 가족의 기능변화로 인해 결혼으로 얻게 되는 이익감소와 결혼에 대한 사회적 부족현상을 혼인율 변화 양성의 원인으로 보았다. 또한, Kim, Choe, Jeon(2015)은 인구구성 요인의 예측은 일반적인 예측의 관점에서 장기예측을 하게 되는데, 약 20년간의 예측결과를 보면 연도별 예측결과에서 초혼율은 감소하고, 재혼율, 이혼율은 증가하는 추세로 나타나 혼인관련 세태의 변화는 사회적 변화 및 경제 상황 등 외부 요인의 변화에 영향을 기인했을 수 있으나, 사회 및 경제관련 외생변수는 다루고 있지 않다는 한계가 나타났다. 이어서 Chen(2020)은 중국의 급증하는 이혼율에 많은 관심을 기울이고 있음에도 불구하고, 지방 차원의 이혼 패턴과 추세에 대한 지식은 여전히 제한적이라는 한계를 시사하고 있다. 본 연구에서는 중국 전체 지역을 기준으로 이혼율이 높은 지역, 평균인 지역, 낮은 지역 간 이혼율 상승 원인과 지리적인 차이를 비교 분석하고자 하며 탐색적 요인분석(exploratory factor analysis, EFA)을 통해 이혼율에 미치는 요인을 분석한다. 마지막으로 구조방정식모형의 분석을 통해 이혼율이 높은 헤이룽장성(Heilongjiang)과 평균인 산시성(Shanxi), 낮은 광둥성(Guangdong) 간에 경제사회 개발화 특성 및 현대가족 특성이 이혼율과 출산율에 미치는 영향을 검증하여 결론을 내리고자 한다.

구조방정식모형(structural equation model)은 연구자의 관심 영역에 있는 변수들 간의 인과관계 및 연관관계를 방정식 체계로 정식화한 것으로 상관분석, 인자분석, 회귀분석 등 매우 넓은 영역의 통계적 분석기법들을 포함하고 있다(Hwang, Kang, 2022; Lee, Kang, 2021; Yang et al., 2019; Kim, 2019; Lee, 2017).

모형은 측정모형(measurement model)과 구조모형(structural model)으로 분류되며 이는 잠재변수가 있는 구조방정식모형과 잠재변수가 없는 구조방정식모형인 경로분석모형(path analysis model)으로 분류된다. 잠재변수가 있는 구조방정식모형의 구축에서는 실제 잠재변수 간의 구조모형의 해석에 관심을 두게 되며 이들 간에 미치는 효과를 평가하게 된다. 인과모형분석은 사회과학과 자연과학에서 여러 변수들 사이의 인과관계를 밝히는 다양한 모형에 대한 분석으로서 구조방정식모형 외에 회귀모형, 확증적 인자분석모형(confirmatory factor analysis model), 공분산분석모형(ANCOVA model) 등이 있다.

2장에서는 MANOVA, 프로파일 분석 등을 이용하여 중국 전체와 3개 성(높은 이혼율 지역, 평균 이혼율 지역, 낮은 이혼율 지역)으로 이루어진 총 4개 집단 간의 평균을 비교하고자 한다. 프로파일 분석은 다변량 통계 기법으로 두 집단 이상에서 p 개의 처리에 의해 구성된 평균벡터에 대해 다양한 관계를 살펴보는 데 관심이 있는 분석방법으로서 반복 측정에 대한 MANOVA에 해당한다. 이 기술은 관찰된 점수를 수준과 패턴 효과로서 비직교적으로 분해하기 위해 교육, 심리학 및 의학 연구자들에 의해 널리 사용되고 있다.

3장에서는 탐색적 요인분석을 통해 12개 관측변수를 3개 변수군으로 나누었으며 이에 바탕으로 3개 지역 간의 구조방정식모형의 분석을 통해 지역 간에 경제사회 개발화 특성 및 현대가족 특성이 출산율과 이혼율에 미치는 영향을 분석하였다. 끝으로 4장에서는 본 연구의 결론을 정리하려 한다.

2. 분석자료 구축

2.1. 자료 설명

조이혼율(Crude Divorce rate, %)은 이혼에 관한 가장 기본적인 지표로서 1년간에 발생한 총 이혼건수를 당해연도의 총 인구로 나눈 수치를 1,000분비로 나타낸 것으로 인구 1000명 당 이혼건수를 의미한다. 본 연구에서는 조이혼율을 이혼율로 표기하기로 한다. Su et al.(2018)에 의하면 이혼율의 주요 영향 요인은 경제, 교육, 인구 구성의 세 가지 측면에서 분석할 수 있고, Mo(2020)에 의해 GDP, 문맹률 등 경제, 교육 등의 위주로 이혼율 영향 요인을 선택하였다. 이를 위해 본 연구는 National Bureau of Statistics of China(www.stats.gov.cn), Global Data Lab (globaldatalab.org), CEI data(ceidata.cei.cn)로부터 중국의 2001년부터 2019년까지의 1인당 국내 총생산(GDP, Gross domestic product, X_1), 소비수준(Consumption level, X_2), 여성 수입(Women's income, X_3), 전문대학 이상 졸업률(Graduation rate above junior college, X_4), 도시 인구비율(Urban population ratio, X_5), 인간개발지수(HDI, Human Development Index, X_6), 여성 교육지수(Women's Education Index, X_7), 출산율(Birthrate, X_8), 노인부양비율(Old-age dependency ratio, X_9), 성 개발지수(Gender-related Development Index, X_{10}), 문맹률(Illiteracy rate, X_{11}), 이혼율(Divorce rate, X_{12}) 등의 총 12개 변수를 사용하고자 한다. 특히 여기서 이혼율과 출산율의 측정단위는 천분율(% : per mille)로서 1000분의 1을 나타내고 있으며 인간개발지수(HDI) 및 소득지수 등의 변수 설명은 다음과 같이 정의된다.

$$HDI = \frac{1}{3}(\text{소득지수} + \text{기대수명지수} + \text{교육지수}),$$

$$\text{소득지수} = \frac{\text{구하려는 나라의 소득} - \text{최적값}}{\text{최고값} - \text{최적값}} \times \text{각 수의 log값},$$

$$\text{기대수명지수} = \frac{\text{구하려는 나라의 평균수명} - \text{최적값}}{\text{최고값} - \text{최적값}},$$

$$\text{교육지수} = \frac{2}{3} \times \text{성인문자 해독율} + \frac{1}{3} \times \text{학교등교률}.$$

GDI는 건강, 교육, 경제 자원에 대한 세 가지 기본 차원에서 남성과 여성의 성취도를 측정한 변수로서 남성 HDI에 대한 여성 HDI의 비율로 설명된다. 본 연구는 2019년 중국 대륙별 이혼율 분포를 살펴본 결과를 토대로 이혼율이 중간 집단으로 나타난 산시성을 기준으로 이혼율이 낮게 나타난 광둥성, 이혼율이 높게 나타난 헤이룽장성을 선택하였다.

변수 간의 관계를 추정, 기술하기 위해 2010년부터 2019년까지의 31개 성 데이터를 추출하여 기술통계 분석을 시행한다. 다음으로 MANOVA를 사용하여 평균값의 차이가 있는지 판정하고자 하는데 이때 현저한 차이가 발생한다면, 프로파일 분석을 이용하여 추가 분석을 진행할 수 있다. 여기서는 그룹 간 차이를 직관적으로 비교하기 위해서는 프로파일 분석을 적용한다.

실제 자료들은 시계열적 특성이 있으나 서로 상관성이 높은 것과 낮은 것이 혼합되어 있는 여러 변수에 의해서 추정된 관계를 서로 독립적인 것으로 가정한다.

2.2. 기술통계

전체적인 자료의 양상을 살펴보기 위해 연도별 이혼율 추이와 각 데이터의 분포를 파악하기 위해 기초 기술통계분석을 시행하였다. Table 1의 2001년부터 2019년까지 4개 집단의 이혼율 변화추이 그래프는 Figure 1과 같고, 특히 이혼율 증가가 뚜렷한 2개 연도 2010년 및 2019년 중국 전체의 12개 변수의 기술통계분석은 Table 2와 같다.

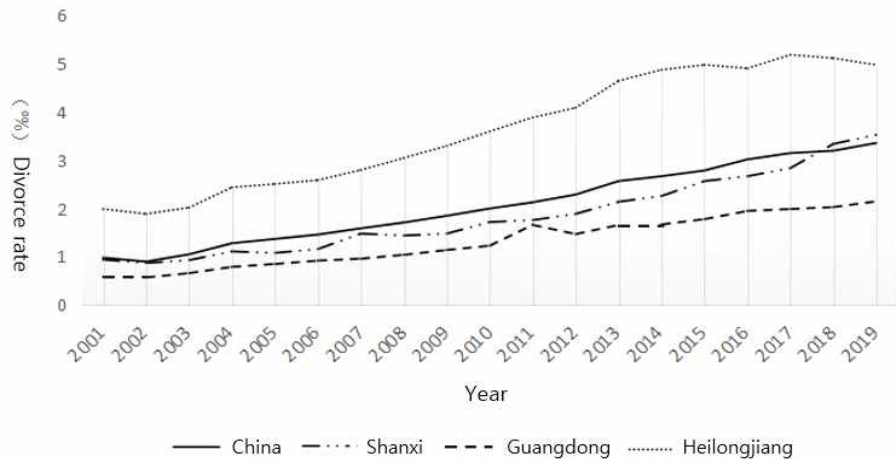


Figure 1. Changes in divorce rate from 2001 to 2019

Table 1. Changes in divorce rate from 2001 to 2019 (%)

Year	China	Shanxi	Guangdong	Heilongjiang
2019	3.36	3.53	2.15	4.98
2018	3.20	3.34	2.03	5.12
2017	3.15	2.84	1.99	5.19
2016	3.02	2.67	1.95	4.91
2015	2.79	2.57	1.78	4.98
2014	2.67	2.26	1.67	4.88
2013	2.57	2.14	1.66	4.65
2012	2.29	1.89	1.47	4.09
2011	2.13	1.76	1.66	3.89
2010	2.00	1.72	1.23	3.60
2009	1.85	1.48	1.14	3.30
2008	1.71	1.44	1.04	3.05
2007	1.59	1.48	0.96	2.8
2006	1.46	1.16	0.92	2.59
2005	1.37	1.08	0.85	2.51
2004	1.28	1.11	0.79	2.44
2003	1.05	0.93	0.66	2.02
2002	0.90	0.87	0.57	1.89
2001	0.98	0.94	0.58	1.99

중국 전체의 이혼율 수치는 2001년 0.98%에서 2019년 3.37%로 연평균 12.94%의 높은 증가율을 보였으며 산시성은 중국 전체와 비슷한 증가 곡선을 그리는 성으로 2001년에 비해 평균 14.55%의 높은 증가율을 볼 수 있었다. 헤이룽장성은 2017년에 5.19%까지 이르는 반면 광둥성은 이혼율은 가장 낮은 성으로 그해 1.99%에 불과하였다.

2010년과 2019년 중국 각 성의 경제, 교육, 인구 구성지표의 특징 비교를 통해 2010년 대비 2019년 중국 전체 평균 이혼율은 2.06%에서 3.37%로 높아진 것을 알 수 있었다. 동시에 경제 수준과 여성 교육수준이 크게 향상된 반면 출산율은 올라가지 않고 오히려 감소하였다.

Table 2. Characteristics of indicators such as economy, education, and population composition in each province of China in 2010 (2019)

Parameter	Mean	Standard Error	Minimum Value	Maximum Value
X_1	3.12(6.91)	1.66(3.26)	1.29(3.30)	7.94(16.46)
X_2	1.10(2.68)	0.59(2.96)	0.45(1.30)	3.23(18.08)
X_3	0.63(0.72)	0.07(0.58)	0.51(0.62)	0.78(0.86)
X_4	9.76(15.76)	5.25(8.18)	5.29(8.23)	31.50(51.37)
X_5	7.24(60.81)	1.81(11.59)	2.46(31.54)	10.08(88.3)
X_6	0.69(0.75)	0.06(0.06)	0.53(0.61)	0.84(0.90)
X_7	0.58(0.64)	0.08(0.08)	0.32(0.38)	0.79(0.86)
X_8	11.29(10.56)	2.72(2.52)	6.68(5.73)	15.99(14.60)
X_9	11.41(17.05)	2.08(4.08)	7.22(8.86)	16.17(23.26)
X_{10}	0.94(0.96)	0.02(0.02)	0.90(0.91)	0.96(0.98)
X_{11}	4.84(5.58)	4.30(5.76)	1.70(1.21)	24.42(33.11)
X_{12}	2.06(3.37)	0.87(0.89)	0.67(1.50)	4.50(5.00)

2.3. 다변량 분산분석

4개 집단(중국 전체, 산시성, 광둥성, 헤이룽장성) 간의 귀무가설 $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ 인 다변량 평균을 검정하고자 일원 MANOVA(Choi, 2019)를 적용하였다. 이는 4개 처리(지역) $j=1,2,3,4$ 에 대해 각 처리 별 자료 수는 동일하게 $n_j=19$ 이다. 또한 각 처리별 반응변수에서 언급한 바와 같이 변수의 수는 12개이고 따라서 각 처리별 평균과 총 평균벡터는 다음과 같다.

$$\begin{aligned} \text{총 평균: } \bar{\mathbf{y}} &= (2.27, 3.34, 1.31, \dots, 5.13)^T \\ \text{처리 1(산시성): } \bar{\mathbf{y}}_1 &= (2.07, 3.38, 1.27, \dots, 7.07)^T, \\ \text{처리 2(광둥성): } \bar{\mathbf{y}}_2 &= (1.85, 3.02, 1.02, \dots, 7.36)^T, \\ \text{처리 3(헤이룽장성): } \bar{\mathbf{y}}_3 &= (1.32, 4.80, 1.84, \dots, 4.02)^T. \end{aligned}$$

여기서 MANOVA를 위한 F -검정을 근사적으로 따르는 윌크스(Wilks) 검정통계량 $\Lambda = 1.115 \times 10^{-4}$ 에 대응하는 F -검정 통계량 $F_0 = 2.218 \times 10^3$ 의 유의확률이 $p\text{-value} < 0.001$ 으로 유의수준 0.05하에 귀무가설 $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ 를 기각하며 이로부터 네 처리 사이에 유의한 차이가 있음을 알 수 있었고, 이는 각 처리 별 평균벡터의 결과로부터 짐작 가능함을 볼 수 있다.

2.4. 프로파일 분석

Choi(2019)에 의하면 프로파일 분석은 g 개 집단 간의 p 개 처리에 의해 구성된 평균벡터 $\mu_j = (\mu_{j1}, \dots, \mu_{jp})^T$, $j=1, \dots, g$ 의 다양한 관계를 살펴볼 수 있다. 여기서 프로파일 분석을 사용하여 집단 간의 평행, 일치, 수직 관계를 존재 여부를 확인할 수 있다. Srivastava(2002)에 따르면 일반화 프로파일 분석(generalized PA)은 세 집단 이상의 PA로 MANOVA를 활용하여 g 개 처리와 p 개의 반응변수로 구성된 y_{ij} , $i=1, \dots, n_j$, $j=1, \dots, g$ 는 다변량 정규분포 $N_p(\mu_j, \Sigma)$ 를 만족하며 식 (1)과 같은 일원 MANOVA 모형을 따른다.

$$y_{ij} = \mu_j + e_{ij} \quad (1)$$

일반화 프로파일 분석은 다음과 같이 3단계 별 귀무가설 H_{01} , H_{02} , H_{03} 에 대해 검정을 하게 된다.

<1단계> H_{01} : Parallel Profiles (평행성 프로파일)

$$H_{01}: C(\mu_1 - \mu_g) = \cdots = C(\mu_{g-1} - \mu_g) = 0$$

$$\Leftrightarrow H_{01}: \mu_1 - \mu_g = \gamma_1 \mathbf{1}_p, \cdots, \mu_{g-1} - \mu_g = \gamma_{g-1} \mathbf{1}_p$$

<2단계> H_{02} : Coincident Profiles (일치성 프로파일)

$$H_{02}: \gamma = (\gamma_1, \cdots, \gamma_{g-1})^T = 0$$

<3단계> H_{03} : Level Profiles (수준성 프로파일)

$$H_{03}: \mu_g = \delta \mathbf{1}_p \Rightarrow C\mu = 0$$

여기서 C 는 $(p-1) \times p$ 인 대비행렬, $\mathbf{1}_p$ 는 1로 구성된 행렬이고 γ_j 는 j 번째 집단의 수준 차이를 나타내며, δ 는 상수이다. 본 연구에서는 일반화 프로파일 분석을 적용하여 중국 전체, 산시성, 광둥성, 헤이룽장성 4개 집단 간의 총 12개 관측변수를 비교한 결과는 Table 3과 Figure 2와 같다.

검정 결과를 보면, H_{01} : 평행성이 기각되어 4개 집단 간의 프로파일이 유사하지 않음을 보여주고 있다. H_{02} : 일치성도 기각되어 4개 집단이 다름을 나타내고 있고 H_{03} : 수준성도 기각이 되어 차이가 있음을 보여준다. 따라서 4개 집단 간에 현저한 차이가 있음을 나타낸다.

다음 부분은 프로파일 분석을 적용하여 각 성과 중국 전체의 차이를 검토하고자 한다. 첫째는 중국 전체와 이혼율의 분포가 평균 집단으로 나타난 산시성 간의 비교이고, 둘째는 중국 전체와 이혼율이 높은 헤이룽장성 간의 비교이다. 마지막으로 중국 전체와 이혼율이 낮은 광둥성 간의 비교이다. 비교한 결과를 Table 4에 요약하였다.

Figure 3을 보면, 형태는 비슷한 면이 있지만, 결과 Table 4에 따르면 H_{01} : 평행성이 기각되어 중국 전체와 3개 성의 프로파일이 유사하지 않음을 보여주고 있다. 일반화 프로파일 분석의 경우 평행성이 기각되어 다음 가설검정을 진행하지 않아도 되지만, 추가적으로 다음 가설검정을 함께 살펴보면 H_{02} : 일치성의 경우 중국 전체와 산시성, 중국 전체와 헤이룽장성의 귀무가설이 기각되어 다름을 나타내고 있는 한편 중국 전체와 광둥성 간의 비교 결과는 채택되어 일치성이 유사함을 나타낸다. H_{03} : 수준성은 중국 전체와 3개 성 모두 기각이 되어 차이가 있음을 보여준다. 이는 전체적으로 프로파일에 차이가 있음을 나타내고 있다.

Table 3. Results of 4 groups of PA

Stage	H_0 (null hypothesis)	Test Statistics	p -value	Result
Step 1	H_{01} : Parallel	$\chi^2_{01} = 1.839 \times 10^{-4}$	< 0.001	reject
Step 2	H_{02} : Coincident	$F_{02} = 13.207$	< 0.001	reject
Step 3	H_{03} : Level	$F_{03} = 16125.075$	< 0.001	reject

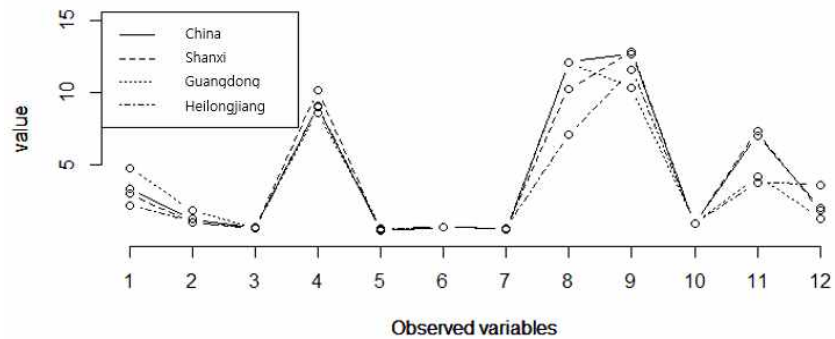


Figure 2. PA plots of four groups

Table 4. Results of PA in China and three provinces

	H_{01} : Parallel			H_{02} : Coincident			H_{03} : Level		
	F_{01}	p -value	Result	F_{02}	p -value	Result	F_{03}	p -value	Result
China vs Shanxi	58.423	< 0.001	reject	11.479	< 0.001	reject	44477.75	< 0.001	reject
China vs Heilongjiang	423.234	< 0.001	reject	4.284	< 0.001	reject	27966.05	< 0.001	reject
China vs Guangdong	363.042	< 0.001	reject	0.903	0.550	accept	9661.360	< 0.001	reject

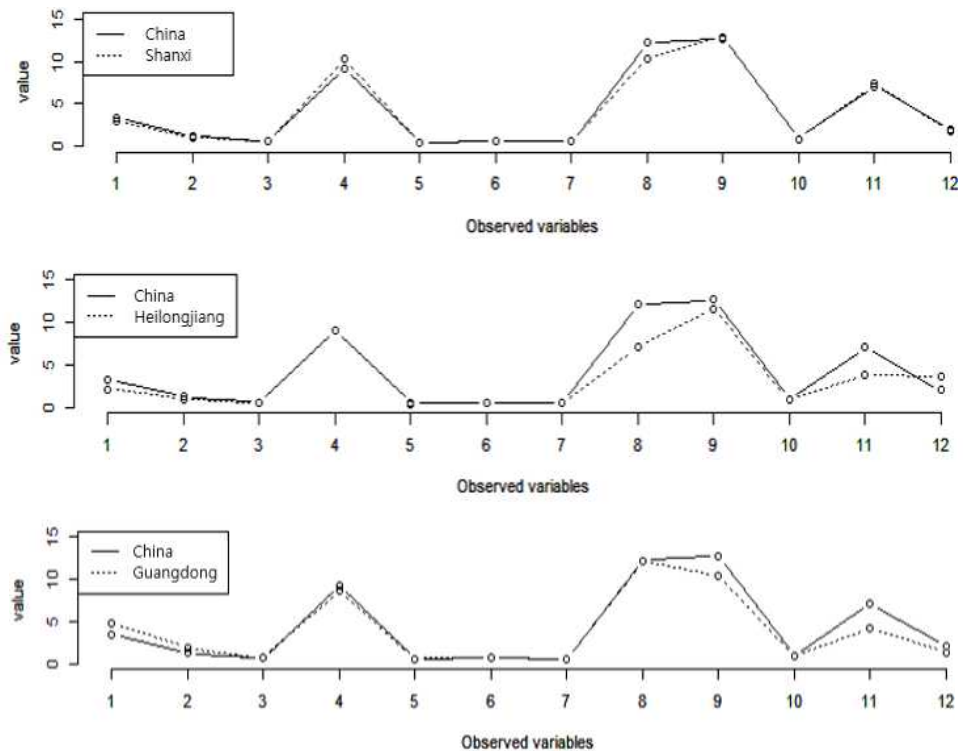


Figure 3. PA plots of China and three provinces

3. 분석 사례

3.1. 탐색적 요인분석

탐색적 요인분석은 다중 관측변수의 구조를 찾아내고 변수 수를 줄이는 차원축소를 처리하는 방법이다. 여기서 사용된 변수는 X_1 (1인당 국내 총생산), X_2 (소비수준), X_3 (여성 수입), X_4 (전문대학 이상 졸업률), X_5 (도시 인구비율), X_6 (인간개발지수), X_7 (여성 교육지수), X_8 (출산율), X_9 (노인부양비율), X_{10} (성 개발지수), X_{11} (문맹률), X_{12} (이혼율)이고, 총 12개 변수의 탐색적 요인분석의 결과는 Table 5와 같다. 특히 변수의 성질을 일치시키기 위하여 영향계수가 음인 변수 X_8 (출산율)과 X_{11} (문맹률)에 대하여 역수를 사용했다.

결과에 따르면 12개 변수는 3개 그룹으로 나누어진다. F_1 은 경제사회 개발화 특성으로 정의되고 F_3 는 현대가족 특성으로 정의되며 F_2 는 혼인 특성이 된다. 특히 이들 F_1 , F_2 , F_3 는 서로 독립으로 정하여 구조방정식모형의 분석을 실시할 때 이들 인자를 3개 주요 변수군으로 진행할 수 있다.

Table 5. Results of the EFA of 12 observation variables

Variables	Factor loading		
	F_1	F_2	F_3
X_1	0.888	-0.168	0.304
X_2	0.880	0.005	0.290
X_3	0.969	-0.137	0.194
X_4	0.560	0.178	0.806
X_5	0.840	0.285	-0.097
X_6	0.955	0.106	0.269
X_7	0.786	0.437	0.417
X_8	-0.199	0.946	0.117
X_9	0.070	0.226	0.671
X_{10}	0.690	0.651	0.205
X_{11}	0.666	0.508	0.007
X_{12}	0.178	0.883	0.400
Eigenvalue	6.074	2.770	1.785
Contribution rate	50.4%	23.1%	14.9%
Total contribution rate	88.3%		

3.2 구조방정식모형의 분석

1) 이론 고찰

Choi(2021)에 의하면 구조방정식모형은 측정모형과 구조모형을 통해서 모형 간의 인과관계를 파악하는 통계적 이론모형을 의미한다. 이는 인과분석을 위해서 요인분석과 회귀분석을 결합한 형태로 구조화되며 잠재변수가 있는 구조방정식모형의 구축은 측정모형과 구조모형에 의해서 이루어진다. 관측변수 X 형에 의한 $\mathbf{x} = (x_1, \dots, x_q)^T$ 에 관한 측정모형은 식 (2)와 같이 오차변수 $\boldsymbol{\delta} = (\delta_1, \dots, \delta_q)^T$ 와 잠재변수 $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)^T$ 에 의해서 구축된다.

$$\mathbf{x} = \mathbf{A}_x \boldsymbol{\xi} + \boldsymbol{\delta} \quad (2)$$

유사하게 관측변수 $\mathbf{Y}$ 형에 의한 $\mathbf{y} = (y_1, \dots, y_p)^T$ 에 대해선 오차변수 $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_p)^T$ 와 잠재변수 $\boldsymbol{\eta} = (\eta_1, \dots, \eta_m)^T$ 에 의해서 식 (3)과 같은 측정모형이 구축되어 진다.

$$\mathbf{y} = \mathbf{A}_y \boldsymbol{\eta} + \boldsymbol{\epsilon} \quad (3)$$

구조모형은 $\mathbf{x}$ 에 관한 측정모형과 $\mathbf{y}$ 에 관한 측정모형을 통한 회귀분석적 또는 경로분석적으로 결합하여 관계화한 모형을 의미하며, 이는 식 (4)와 같이 표현된다.

$$\boldsymbol{\eta} = \Gamma \boldsymbol{\xi} + \beta \boldsymbol{\eta} + \boldsymbol{\zeta} \quad (4)$$

실제 관측되지 않는 $\boldsymbol{\eta} = (\eta_1, \dots, \eta_m)^T$ 는 내생잠재변수이고, $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)^T$ 는 외생잠재변수이며, β 및 Γ 는 계수행렬이고, $\boldsymbol{\eta}_1$ 는 랜덤오차를 표시한 벡터이다.

일반적으로 가설적 모형에 대한 평가는 합도 지수(fit index)들에 의해서 이루어지는데, 적합도 지수에 대한 적절한 이해와 타당한 모형을 선택하기 위한 절차는 중요하게 다루어져야 할 내용이라고 할 수 있다. (Song, Kang, 2021; Kang, 2013). 구조방정식모형을 평가하기 위하여 여러 가지 적합도 지수들이 제안되어 있는데 가장 기본적인 카이제곱(χ^2)값이 있으며, 카이제곱값을 변형한 것으로 Normed χ^2 , GFI, AGFI, RMSEA, RMR, SRMR, CN 등의 절대 적합도 지수가 있고, CFI, TLI, NFI 등 중분 적합도 지수와 PGFI, PNFI, PCFI 등 간명 적합도 지수, AIC, BCC, BIC, CAIC, ECVI 등 정보기준에 근거한 적합도 지수 등 다양한 적합도 지수들이 존재하는데 이들이 적합도의 서로 다른 측면을 평가하기 때문이며, 연구자는 여러 가지 적합도 지수들을 이용하여 모형을 종합적으로 평가하는 것이 추천된다(Song, Kang, 2021).

모형이 주어진 경험 자료에 잘 맞는지를 평가하는 적합성 평가는 부합지수를 활용하나 부합지수 중 어느 하나라도 무조건 최고라고 내세울 수 있는 지수는 없으므로 가설화 된 모형에 대하여 가장 큰 장점과 작은 문제점을 지닌 부합지수를 선택하여 모형의 적합성 검토에 사용해야 한다. 본 연구에서는 Rosseel(2012)이 개발한 R의 lavaan 패키지를 활용하여 분석을 진행한다.

2) 광둥성, 산시성, 헤이룽장성 간 비교

3개 지역 간 비교의 모형을 설계하기 위해 요인분석을 통해 혼인 관련 특성 X_{12} (이혼율)과 X_8 (출산율)을 Y_1 과 Y_2 로 표시하고 잠재변수를 η_1 으로 설정하여 경제사회 개발화 특성(ξ_1) 및 여성 현대화 특성(ξ_2)이 출산율과 이혼율(η_1)에 미치는 영향을 제시하고자 한다. 관측변수와 잠재변수를 Table 6과 같이 설정하며 다음과 같은 연구가설을 설계하였다.

H_{01} : 경제사회 개발화 특성(ξ_1)은 혼인 특성(η_1)에 영향을 미친다.

H_{02} : 현대가족 특성(ξ_2)은 혼인 특성(η_1)에 영향을 미친다.

Figure 4의 경로도를 바탕으로 잠재변수가 있는 구조방정식모형을 구축하면 다음과 같다. 여기서 Table 7과 같은 측정모형과 구조모형을 정의했다.

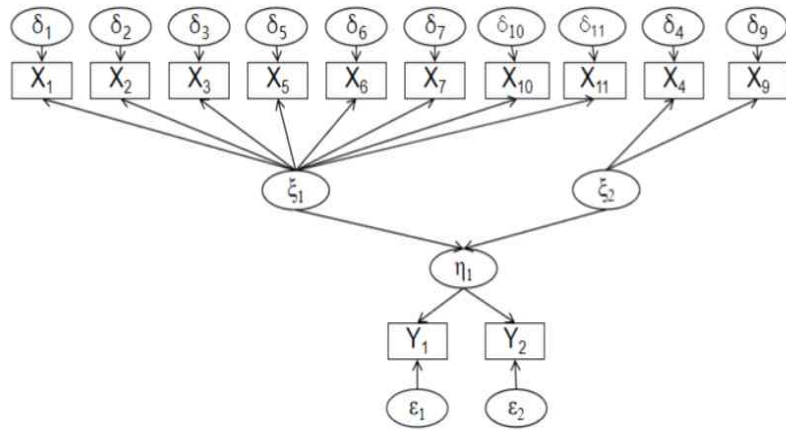


Figure 4. Pathways of comparisons between three regions

Table 6. Observation variable and Latent variable

Latent Variables	Variable
ξ_1 : economic and social development characteristics	$X_1, X_2, X_3, X_5, X_6, X_7, X_{10}, X_{11}$
ξ_2 : Characteristics of Modern Family	X_4, X_9
η_1 : fertility on divorce	Y_1, Y_2

Table 7. Measurement model and Structure model

MM, Measurement Model	SM, Structure Model
$x_1 = \lambda_1^x \xi_1 + \delta_1$ $x_2 = \lambda_2^x \xi_1 + \delta_2$ $x_3 = \lambda_3^x \xi_1 + \delta_3$ $x_5 = \lambda_5^x \xi_1 + \delta_5$ $x_6 = \lambda_6^x \xi_1 + \delta_6$ $x_7 = \lambda_7^x \xi_1 + \delta_7$ $x_{10} = \lambda_{10}^x \xi_1 + \delta_{10}$ $x_{11} = \lambda_{11}^x \xi_1 + \delta_{11}$	$\eta_1 = (\gamma_1 \gamma_2) \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \zeta_1$
$x_4 = \lambda_4^x \xi_2 + \delta_4$ $x_9 = \lambda_9^x \xi_2 + \delta_9$	
$y_1 = \lambda_1^y \eta_1 + \epsilon_1$ $y_2 = \lambda_2^y \eta_1 + \epsilon_2$	

표본의 수가 작은 관계로 재표집(resampling) 기법인 부트스트랩(Bootstrap)을 적용하고, 그룹 간의 비교를 위해 group 변수로 세 지역을 추가하면서 하나의 모델로 구조방정식모형 적합결과를 살펴본 결과, 적합지수들 중 표본의 크기에 무관한 모형적합 지수 GFI(goodness-of-fit index) 값이 0.955 이고 SRMR(standardized root mean square residual)은 0.074로 각각 모형 적합 기준 0.9 이상과 0.08 이하를 만족하고 있다.

각 모수들에 대한 추정치와 모수의 유의성 검정을 위한 검정통계량 z -값과 유의확률 p -value 의 결과를 요약하자면 Table 8, Table 9, Table 10과 같다.

Table 8. Estimation of Parameters and Significance of Structural Equations Model in Heilongjiang

Parameter	Estimation	Standard Error	z-value	p-value
λ_1^x	1.000	-	-	-
λ_2^x	1.599	0.104	15.315	<0.001
λ_3^x	1.294	0.054	23.753	<0.001
λ_5^x	1.922	0.214	8.969	<0.001
λ_6^x	1.690	0.078	21.632	<0.001
λ_7^x	2.012	0.058	34.842	<0.001
λ_{10}^x	1.832	0.108	16.902	<0.001
λ_{11}^x	1.796	0.196	9.175	<0.001
λ_4^x	1.000	-	-	-
λ_9^x	1.197	0.154	7.772	<0.001
λ_1^y	1.000	-	-	-
λ_2^y	0.579	6.886	6.886	<0.001
γ_1	1.089	0.502	2.171	0.030
γ_2	0.474	0.261	1.812	0.070

Table 9. Estimation of Parameters and Significance of Structural Equation Model in Shaanxi

Parameter	Estimation	Standard Error	z-value	p-value
λ_1^x	1.000	-	-	-
λ_2^x	0.860	0.025	33.930	<0.001
λ_3^x	1.187	0.096	12.405	<0.001
λ_5^x	0.926	0.032	29.277	<0.001
λ_6^x	1.200	0.119	10.088	<0.001
λ_7^x	1.165	0.113	10.311	<0.001
λ_{10}^x	1.073	0.136	7.881	<0.001
λ_{11}^x	0.481	0.057	8.465	<0.001
λ_4^x	1.000	-	-	-
λ_9^x	0.597	0.062	9.613	<0.001
λ_1^y	1.000	-	-	-
λ_2^y	-0.035	0.034	-1.023	0.036
γ_1	0.011	0.354	0.030	0.976
γ_2	0.545	0.309	1.763	0.078

Table 8, Table 9, Table 10에 따르면 헤이룽장성과 광둥성의 경제사회 개발화 특성(ξ_1)의 영향계수는 1.089와 0.470으로 높게 나타난 반면 산시성의 영향계수는 0.011로 낮게 나타난다. 그리고 현대가족 특성(ξ_2)의 관점에서 보았을 때, 헤이룽장성과 산시성의 영향계수는 0.474와 0.545로 서로 유사하게 나타난 반면 광둥성은 -0.171로 음의 영향을 미치고 있으나 통계적으로는 영향을 주지 못하고 있다($p=0.276$). 잠재변수 간의 효과분할표 Table 11에 따르면 혼인 특성(η_1)에 미치는 경제사회 개발화 특성(ξ_1)과 현대가족 특성(ξ_2)의 적합효과들이 존재한다. 다만 광둥성의 경우 음의 영향을 미침을 보여준다.

Table 10. Estimation of Parameters and Significance of Structural Equation Model in Guangdong

Parameter	Estimation	Standard Error	z-value	p-value
λ_1^x	1.000	-	-	-
λ_2^x	0.987	0.008	118.207	<0.001
λ_3^x	0.660	0.028	23.287	<0.001
λ_5^x	0.208	0.031	6.764	<0.001
λ_6^x	0.776	0.044	17.550	<0.001
λ_7^x	0.848	0.053	15.967	<0.001
λ_{10}^x	0.577	0.044	12.979	<0.001
λ_{11}^x	0.709	0.077	9.255	<0.001
λ_4^x	1.000	-	-	-
λ_9^x	-0.209	0.114	-1.827	0.068
λ_1^y	1.000	-	-	-
λ_2^y	0.232	0.123	1.885	0.059
γ_1	0.470	0.116	4.047	<0.001
γ_2	-0.171	0.156	-1.090	0.276

Table 11. Division of Effects

Provinces	Endogenous latent variable	Preliminary latent variable	Correlation coefficient	Division of Effects		
				Indirect effect	direct effect	total effect
Heilongjiang	η_1	ξ_1	0.950	-	1.089	1.089
		ξ_2	0.812	-	0.474	0.474
Shanxi		ξ_1	0.803	-	0.011	0.011
		ξ_2	-0.131	-	0.545	0.545
Guangdong		ξ_1	0.742	-	0.470	0.470
		ξ_2	-0.047	-	-0.171	-0.171

상관계수는 전체적인 현대가족 특성(ξ_2)보다 경제사회 개발화 특성(ξ_1)에서 더 높게 나타난다. 그리고 산시성과 광둥성의 현대가족 특성(ξ_2)과 혼인 특성(η_1)의 상관계수는 -0.131과 -0.047로 음의 수치를 나타낸 반면 헤이룽장성은 0.812로 높은 상관성이 있는 것을 보여준다. 그 이유는 2020년 중국 인구 조사의 결과(National Bureau of Statistics of China, 2021)를 통해 헤이룽장성은 고령화 문제가 심각한 지역인 것으로 추측되기 때문이다.

종합적으로 구조방정식모형의 결과를 통해 경제, 교육, 여성의 지위가 높아짐에 따라 이혼율은 높아지는 한편 출산율과 이혼율은 반비례한다는 것을 도출하였다.

4. 결론

본 연구는 중국의 이혼율이 꾸준히 상승하고 있어, 심각한 사회 문제로 대두되고 있기에 이혼율에 영향을 미치는 원인을 파악하고 지역별 이혼율 수준에 따라 세 집단(헤이룽장성, 산시성, 광둥성)을 분류한 후 인구 구성적 특성인 이혼율, GDI, 문맹률에 미치는 요인을 구체적으로 파악하고자 하였다. National Bureau of Statistics of China, Global Data Lab, CEI data로부터 데이터를 활용하였으며 중국의 2001년부터 2019년까지의 이혼율, GDP, 소비수준, 여성 수입, 전문대학 이상 졸업률, 도

시 인구비율, 인간개발지수, 여성 교육수준, 출산율, 노인부양비율, 성 개발지수, 문맹률 등의 변수를 사용하여 분석을 진행하였다. 분석결과 중국 정부의 이혼율 및 인구 구성적 특성에 미치는 요소가 파악되었기에 이는 몇 가지 정책적 방향 설정에 도움이 될 것으로 생각된다.

첫째, MANOVA, 프로파일 분석 등을 이용하여 3개 지역 간의 평균을 비교한 결과를 따르면 지역 간 차이가 있는 것으로 나타났다. 이에 3장의 연구가 의의가 있음을 설명한다. 둘째, 요인분석 후에 3개 변수 군으로 나누어지고 구조방정식모형의 분석을 진행한 결과를 살펴보면 경제사회 개발화 특성이 현대가족 특성보다 더 높게 나타내며 지역 간 차이도 볼 수 있다. 전체적으로 경제, 교육, 여성의 지위가 높아짐에 따라 이혼율도 높아지는 한편, 출산율과 이혼율은 반비례한다는 결론이 도출되었다.

이 연구에서 언급된 영향 외에도 이혼율에 영향을 미치는 이유는 여러 가지 있을 것으로 파악된다. 예를 들면 사람의 성격, 부부간 애정의 돈독함, 그리고 관념의 차이 등 다양한 영향이 있을 것으로 추정되기에 이혼율 상승의 영향을 철저히 파악하기에는 역부족이었다. 또한 통계청의 자료 집계 특성상 2019년도의 데이터까지만 존재했기에 전년도 및 당해년도 자료 파악은 진행하지 못하였다. 2021년 1월 중국 정부의 이혼 숙려제도의 도입 후 이혼율에 또 다른 요인변수가 영향을 미쳤을 수도 있어 최근 년도 자료를 활용하지 못한 부분에 있어 다소 아쉬움이 존재한다. 따라서 이혼 숙려제도(2021)의 도입 전후로 이혼율의 영향에 대해서 살펴보는 것은 추후의 과제들에서 다루어지게 되기를 기대한다.

References

- Chen, M. (2020). Divorce trends in China across time and space: an update, *Asian Population Studies*, 121-147, DOI: <https://doi.org/10.1080/17441730.2020.1858571>
- Choi Y.S., (2019). *Estimation and Testing for Multivariate Data Analysis with R*, Gyeongmunsa(in Korean).
- Choi Y.S., (2021). *Understanding and Application of Causal Model with R-SEM*, Gyeongmunsa(in Korean).
- Chung, H.-S. (2007). A critical analysis of the characteristics and causes of the changes in marriage rates and recommendations for family policy, *Journal of Korean Home Management Association*, 24(6), 177-193. (in Korean).
- Hwang, I.O., Kang, H.C. (2022). Comparative Study on Testing Methods of Path Coefficient in Structural Equation Model, *Journal of the Korean Data Analysis Society*, 24(3), 1007-1016. (in Korean). DOI : <https://scholar.kyobobook.co.kr/article/detail/4010031909724>
- Kang, H. (2013). Discussions on the suitable interpretation of model fit indices and the strategies to fit model in structural equation modeling, *Journal of the Korean Data Analysis Society*, 15(2), 653-668. (in Korean). DOI : <https://scholar.kyobobook.co.kr/article/detail/4010027882127>
- Kim, K.W., Choe, S.B., Joen, S.B. (2015). A Study on Modeling of First Marriage, Divorce and Remarriage, *Journal of the Korean Data Analysis Society*, 17(5), 2447-2459. (in Korean). DOI : <https://scholar.kyobobook.co.kr/article/detail/4010027873286>
- Lee, K.T., Kang, H. (2021). Evaluation of type I error of fit indices for structural equation model under normal and exponential distribution, *Journal of the Korean Data Analysis Society*, 23(1), 173-181. (in Korean). DOI : <https://scholar.kyobobook.co.kr/article/detail/4010028126024>
- Mo, L. (2020). Trends in the Divorce Rate and its Regional Disparity in China, *Journal of Comparative Family Studies*, 48(4), 383-394.
- National Bureau of Statistics of China, (2021). *Communiqué of the Seventh National Population Census* (No. 5).
- Rosseel, Y. (2012). lavaan: an R package for structural equation modeling, *Journal of Statistical Software*, 48(2), 1-36.

- Song, Y.-S., Kang, H. C. (2021). Evaluation of Type II Error of Fit Indices for Structural Equation Model under Normal and Exponential Distribution, *Journal of the Korean Data Analysis Society*, 23(2), 599-609.. (in Korean). DOI : <https://scholar.kyobobook.co.kr/article/detail/4010028244166>
- Srivastava, M.S. (2002). *Method of Multivariate Statistics*, John Wiley & Sons, Inc., New York.
- Su, L., Liang, C., Yang, X., Liu, Y., Allan, C.P. (2018). Influence Factors Analysis of Provincial Divorce Rate Spatial Distribution in China, *Discrete Dynamics in Nature and Society*. doi:10.1155/2018/6903845.
- Wang, Q. (2001). China's divorce trends in the transition toward a market economy. *Journal of Divorce & Remarriage*, 35(1 - 2): 173 - 189. [Taylor & Francis Online], Doi: https://www.tandfonline.com/doi/10.1300/J087v35n01_11
- Wang, Q., Qin Z. (2010). China's Divorce and Remarriage Rates: Trends and Regional Disparities, *Journal of Divorce & Remarriage*, 51:4, 257-267. [Taylor & Francis Online], Doi : <https://doi.org/10.1080/10502551003597949>

A Study on China's Divorce Rate using Structural Equation Model

Xiajing Hu¹, You Jeong Kim², Jun Hong Kim³, Yong-Seok Choi⁴

Abstract

In recent years, the divorce rate for China has been rising year by year. Based on the crude divorce rate from 2001 to 2019, this paper selects four groups of Heilongjiang, Shanxi and Guangdong, where the total divorce rate and the total divorce rate of China are high, medium and low respectively, and explores the reasons for the rising divorce rate of China from three aspects of economy, education and population composition. Compare the provinces with high divorce rate and low divorce rate at the same time, and analyze the reasons and geographical differences of rising divorce rate. This paper is mainly divided into two parts. The first part is the comparison of differences between four groups based on multivariate statistical analysis, and the second part is structural equation model analysis. The first part is mainly descriptive statistics and variance analysis of all groups in China, Shaanxi, Guangdong and Heilongjiang, to confirm whether there are differences among the four groups. Then, profile analysis was conducted to compare the whole in China and Shaanxi province with the median divorce rate, as well as heilongjiang Province with the high divorce rate and Guangdong Province with the low divorce rate. In the second part of structural equation model analysis, first of all, the factors affecting the divorce rate were analyzed. Secondly, through the structural equation model analysis, the effect of economic and social development characteristics and modernization characteristics and fertility on divorce rates between Heilongjiang Province, the average Shanxi Province, and low Guangdong Province is verified and concluded. Through a variety of analyses, this paper draws a conclusion that the divorce rate for China is greatly influenced by social development characteristics and modernization characteristics and fertility. However, there is little difference in the influence factors of the three provinces of China.

Keywords : Structural Equation Model, Divorce Late, Latent Variable, Profile Analysis, Multivariate Analysis of Variance (MANOVA).

¹Master, Department of Statistics, Pusan National University2, Busandaehak-ro 63beon-gil, Geumjeong-gu, Busan, Korea. E-mail: huxiajingwyx@163.com

²Master, Department of Statistics, Pusan National University2, Busandaehak-ro 63beon-gil, Geumjeong-gu, Busan, Korea. E-mail: yj_kim13@naver.com

³Master, Department of Statistics, Pusan National University2, Busandaehak-ro 63beon-gil, Geumjeong-gu, Busan, Korea. E-mail: wnsghd4685@naver.com

⁴(Corresponding Author) Professor, Department of Statistics, Pusan National University2, Busandaehak-ro 63beon-gil, Geumjeong-gu, Busan, Korea. E-mail: yschoi@pusan.ac.kr

[Received 4 November 2022; Revised 2 December; Accepted 5 December 2022]