

구성타당도 평가에 있어서 요인분석의 활용

강현철

호서대학교 정보통계학과

A Guide on the Use of Factor Analysis in the Assessment of Construct Validity

Kang, Hyuncheol

Department of Informational Statistics, Hoseo University, Asan, Korea

Purpose: The purpose of this study is to provide researchers with a simplified approach to undertaking exploratory factor analysis for the assessment of construct validity. **Methods:** All articles published in 2010, 2011, and 2012 in Journal of Korean Academy of Nursing were reviewed and other relevant books and articles were chosen for the review. **Results:** In this paper, the following were discussed: preliminary analysis process of exploratory factor analysis to examine the sample size, distribution of measured variables, correlation coefficient, and results of KMO measure and Bartlett's test of sphericity. In addition, other areas to be considered in using factor analysis are discussed, including determination of the number of factors, the choice of rotation method or extraction method of the factor structure, and the interpretation of the factor loadings and explained variance. **Conclusion:** Content validity is the degree to which elements of an assessment instrument are relevant to and representative of the targeted construct for a particular assessment purpose. This measurement is difficult and challenging and takes a lot of time. Factor analysis is considered one of the strongest approaches to establishing construct validity and is the most commonly used method for establishing construct validity measured by an instrument.

Key words: Factor analysis, Construct validity, Sample size, Communality, Factor loading

서론

본 논문의 목적은 탐색적 요인분석(Exploratory Factor Analysis [EFA])의 수행과 관련된 여러 가지 사항에 대하여 간호학을 포함한 다양한 분야의 연구자들에게 통합된 관점의 지침을 제공하는 것이다. 특히, 구성타당도(construct validity) 평가와 관련하여 요인분석을 적용할 때 고려해야 하는 주요 사항들에 대하여 여러 측면에서 포괄적으로 논의하였다.

1. 측정도구의 구성타당도

타당하고 신뢰성 있는 측정도구는 모든 연구에 있어서 필수적인

구성요소이며, 연구 결과에 대한 해석은 측정도구의 타당성에 의존하므로 연구자는 측정도구의 타당도를 검증해야 한다(Tafreshi & Yaghmaei, 2006). 타당도(validity)란 그 측정도구(검사)가 원래 측정하려 했던 것을 실제로 잘 측정하는가를 의미한다. 일반적으로 타당도 분석 방법은 크게 내용타당도(content validity), 준거타당도(criterion-related validity), 구성타당도(construct validity)로 구분된다. 측정도구의 특성과 측정도구가 사용되는 상황에 따라 더 중요한 타당도 분석 방법이 존재할 수 있으며, 어떤 경우에는 이들 중 일부의 타당도만 의미를 가질 수도 있다. 이와 같이 타당도는 절대적 개념이 아니며, 이러한 분석 방법들은 다른 형태의 타당도 분석이라기보다는 단지 타당화하는 데 있어서 방법상의 차이를 나타낼 뿐이다 (Tak, 2007).

주요어: 요인분석, 구성타당도, 표본크기, 공통성, 요인적재

Address reprint requests to : Kang, Hyuncheol

Department of Informational Statistics, Hoseo University, 20 Hoseo-ro 79 beon-gil, Asan 336-795, Korea
Tel: +82-41-540-5902 Fax: +82-41-540-5908 E-mail: hychkang@hoseo.edu

투고일: 2013년 8월 16일 심사완료일: 2013년 9월 11일 게재확정일: 2013년 9월 17일

Tafreshi와 Yaghmaei (2006)는 간호학 연구에서 사용된 많은 측정 도구들은 다른 연구를 위해 개발되고 타당화된 것이므로 간호학 연구에서 그것을 사용하고자 한다면 다시 타당화가 수행되어야 한다고 지적하고 있다. 또한, Shin 등(2010)은 2007년과 2008년의 JKAN (Journal of Korean Academy of Nursing)에 게재된 논문들을 분석하여 다음과 같이 논의하고 있다. - “도구의 타당도에 대해서 보고한 논문은 36.2%로 신뢰도 보고에 비해 상당히 낮은 것을 알 수 있다. 과학적인 연구를 하기 위해서는 도구의 신뢰도 및 타당도가 입증된 도구를 사용하는 것이 중요하므로, 논문의 타당도에 대한 기술이 증가되어야겠다. 특히, 본 분석에 따르면, 연구자가 기존에 개발된 도구를 수정 보완하여 그대로 사용한 경우가 많이 있었는데 수정 보완한 도구는 최소한 전문가로부터 내용타당도를 검증받는 절차가 필요하다. 그러므로 도구에 대한 심사 시 타당도에 대한 심사가 강화되어야겠다”

측정도구의 구성타당도 분석은 검사도구가 심리적 구성개념을 제대로 측정하고 있는지를 평가하는 방법이다(Tak, 2007). 구성타당도 분석은 쉽지 않고 시간소모적인 여러 절차들이 포함될 수 있는데 그룹간의 비교, 상관계수를 이용한 수렴타당도(convergent validity)와 변별타당도(discriminant validity)의 검토, 요인분석 등이 구성타당도 분석을 위해 사용될 수 있다. 이들 중 요인분석은 하나의 개념을 측정하는 도구가 여러 하위개념을 구분하여 측정한다고 가정하는 경우 구성타당도 분석을 위해 가장 널리 사용되는 통계적 분석기법이다.

2. 요인분석

요인분석(factor analysis, 인자분석)은 여러 관찰가능한 특성(측정 변수)들로부터 소수의 일반적인 잠재변수(latent variable)를 발견해 내거나 만들어 내는 통계적 과정이라고 할 수 있다. 여기서 잠재변수는 측정변수들 간의 상관관계를 생성시키는 이론적인, 관찰할 수 없는, 저변에 깔려 있는 요인(factor)을 의미한다. 즉, 요인분석은 상관(혹은 공분산)행렬의 구조에 관한 통계적 모형을 구축하고, 그와 같은 구조를 생성시키는 소수 몇 개의 요인을 유도하여, 측정변수들 간의 구조적 관계를 해석하는 자료 분석 기법이라고 할 수 있다(Kang, Han, Kim, & Jhun, 2005).

여기서는 이 논문의 전개를 위해 필요한 범위 내에서 요인분석의 통계모형에 대한 수리적 특성을 간단히 서술한다. 먼저, 측정변수 $\chi_1, \chi_2, \dots, \chi_p$ 에 대하여 m ($\ll p$)개의 요인을 고려하는 통계적 모형을 수리적으로 표현하면 다음과 같다(단, 모든 측정변수들은 평균이 0이고 표준편차가 1이 되도록 표준화되었다고 가정하자).

$$\begin{cases} \chi_1 = \lambda_{11}f_1 + \lambda_{12}f_2 + \dots + \lambda_{1m}f_m + e_1 \\ \chi_2 = \lambda_{21}f_1 + \lambda_{22}f_2 + \dots + \lambda_{2m}f_m + e_2 \\ \dots \\ \chi_p = \lambda_{p1}f_1 + \lambda_{p2}f_2 + \dots + \lambda_{pm}f_m + e_p \end{cases}$$

여기서, 확률변수 f_1, f_2, \dots, f_m 은 모든 측정변수들의 변이에 영향을 주는 공통요인(common factor)이고, 반면에 e_i 는 단 한 개의 측정변수 χ_i 의 변이에만 공헌하는 확률변수로서 유일성변수(uniqueness variate)라고 한다. 위의 식에서 선형결합에 사용된 가중계수 λ_{ij} 를 요인적재(factor loading)라고 부르는데, 이는 모형에서 고려된 i 번째 측정변수 χ_i 에 관한 j 번째 요인 f_j 의 중요성을 나타낸다.

몇 가지 기본적인 가정 하에서 측정변수 χ_i 의 분산을 다음과 같이 분할하여 표현할 수 있다.

$$\text{Var}(\chi_i) = c_i + \psi_i = \{\lambda_{i1}^2 + \lambda_{i2}^2 + \dots + \lambda_{im}^2\} + \psi_i$$

여기서 $c_i (= \lambda_{i1}^2 + \lambda_{i2}^2 + \dots + \lambda_{im}^2)$ 를 측정변수 χ_i 의 공통성(communality)이라고 부르는데, 이는 χ_i 의 분산 중 m 개의 요인들에 의하여 설명되는 부분으로 해석될 수 있다. 따라서, 만약 어떤 측정변수의 공통성이 상대적으로 크다면 그 측정변수의 유일성(특수성) ψ_i 는 반대로 작게 될 것이다.

요인분석은 크게 탐색적 요인분석과 확증적 요인분석으로 구분될 수 있는데, 이 논문에서는 탐색적 요인분석의 적용에 대하여 다룰 것이다. 두 분석 방법에 대한 논의는 본 논문의 후반부에서 조금 더 자세히 다룰 것이다.

사전 자료 분석

1. 표본크기

요인분석에서 표본크기의 문제에 대하여 매우 다양한 견해와 경험법칙이 존재하는데, 요인분석의 최소 표본크기에 대한 전통적인 기준은 크게 두 가지의 범주로 구분할 수 있다. 첫 번째는 절대적 표본크기에 대한 기준이다. Hair, Anderson, Tatham과 Black (1995)은 표본크기가 100보다 커야 한다고 제안하였고, Guilford (1956)는 적어도 200은 되어야 한다고 주장하였으며, Tabachnick과 Fidell (2007)은 적어도 300개의 표본이 필요하다고 제안하였다. Comrey와 Lee (1992)를 인용한 여러 문헌들에서는 표본크기가 100이면 나쁜 편이고, 200 정도면 괜찮은 편이며, 300 정도면 좋은 편이고, 500 정도면 매우 좋은 편이고, 1,000 이상이면 훌륭하다고 서술하고 있다(Tak, 2007; Williams, Brown, & Onsmann, 2012). 두 번째 기준은 $n:p$ 비(사

례 수 대 측정변수의 비율)에 대한 기준이다. 이에 대하여도 20:1 (Hair et al.), 10:1 (Nunnally, 1978), 5:1 (Tabachnick & Fidell) 등 여러 기준들이 제안되어 있으며, Guilford는 2:1 이상이면 된다고 제안한 바 있다.

한편, Arrindel과 van der Ende (1985) 등은 $n:p$ 비는 그다지 중요하지 않으며 중요한 것은 사례 수대 요인 수의 비율이라고 주장하였다. 그들은 안정된 요인을 얻기 위해서는 사례 수가 추출되는 요인 수의 20배 이상은 되어야 한다고 주장하였다(Tak, 2007). 반면에 Hogarty, Hines, Kromrey, Ferron과 Mumford (2005)은 좋은 해(solution)를 얻기 위하여 적용될 수 있는 최소 표본크기 또는 $n:p$ 비에 대한 일반적인 기준은 존재하지 않는다고 언급하였다. 한편, Tak은 대략 사례 수가 200 이상이거나 사례 수와 측정변수의 비율이 5 대 1 이상이면 안정권이라 볼 수 있다고 하였다.

최근의 주목할 만한 연구는 모집단 요인구조(population factor structure)의 재생성을 기준으로 한 것으로서(Arrindel & van der Ende, 1985; Barrett & Kline, 1981; MacCallum, Widaman, Zhang, & Hong, 1999), MacCallum 등은 20개의 측정변수에 대하여 60개의 사례로도 모집단 요인구조를 훌륭하게 재생해 낼 수 있음을 보였다. 그들은 표본크기보다는 자료의 질이 중요하다는 점을 강조하여 공통성과 요인적재값이 높은 측정변수들로 구성된 경우 표본크기가 작아도 좋은 결과를 얻을 수 있음을 밝혔다. 한편, MacCallum, Browne과 Sugawara (1996)은 적합도의 측도인 RMSEA(root mean square error of approximation; Kang, 2013)와 측정변수의 수 및 요인 수를 고려하여 구조방정식모형에서 검정력분석(power analysis)과 표본크기 결정에 대한 절차를 제시하였다. 그들은 SAS 프로그램을 제공하고 있으며(<http://www.unc.edu/~rcm/power/power.htm>), STATISTICA와 같은 일부 통계 프로그램에도 이러한 기능이 제공되고 있다. 아직까지는 탐색적 요인분석에 이 절차를 적용하기 위해서 상당한 수준의 수치연산이 필요하지만, 조만간 일반 연구자들이 쉽게 사용할 수 있는 프로그램 모듈이 제공될 것으로 기대된다.

요약하면, 요인분석에서의 표본크기 문제는 측정변수의 수뿐만 아니라 요인 수 및 공통성과 요인적재값 등이 종합적으로 고려되어야 하며, 너무 작은 표본크기는 당연히 바람직하지 않지만 요인분석이 탐색적 방법론이라는 점을 고려할 때 표본크기 문제에 지나치게 엄격한 기준을 적용할 필요는 없다고 생각된다. 다만, 표본크기가 작은 경우에는 요인분석 결과의 수용 및 기존의 개념-항목 관계의 변경 시 내용타당도 평가와 신뢰도 분석 등을 포함하여 매우 신중하게 결론을 유도할 필요가 있다. 참고로 Table 1에 최근 3년간(2010-2012) JKAN에 게재된 논문 중 구성타당도 검증을 위해 요인분석을 적용한 11편의 사례 수, 측정변수의 수, 요인 수를 제시하였다.

Table 1. The Number of Cases, Variables, and Factors in the Papers of JKAN (2010-2012)

No.	Vol. (No.)	Authors	Cases	Variables	Factors
1	42 (5)	Kim & Park	307	76	17
2	42 (5)	Seo et al.	188	54	5
3	42 (5)	Kim & Lee	137	35	7
4	42 (4)	Jo	388	35	7
5	42 (3)	Lee et al.	303	21	5
6	41 (4)	Lee & Ji	151	30	5
7	41 (3)	Jo	428	34	5
8	41 (3)	Kwon	120	20	5
9	41 (1)	Kim & Jang	960	29, 21, 26	5, 4, 4
10	40 (6)	Shin	1,687	72	16
11	40 (4)	Lee et al.	873	35	7

2. 측정변수들의 분포

측정변수들의 분포를 검토하는 것은 모든 통계분석의 기본이 된다. 먼저, 평균과 표준편차를 통해 측정변수들의 기본적 특징을 살펴보아야 한다. 평균을 통해 측정값들이 전체적으로 어디에 몰려 있는지를 살펴볼 필요가 있다. 만약 어떤 측정변수의 평균이 너무 높거나 너무 낮다면, 해당 문항의 내용 및 자료가 수집된 대상 등에 대한 종합적 검토를 수행할 필요가 있다. 표준편차는 자료의 변이(variation)를 나타낸다. 표준편차가 너무 크게 나타나면 특이값이 존재할 수 있으므로 주의해야 하고, 표준편차가 너무 작게 나타나면 해당 측정변수가 가지고 있는 정보가 많지 않다는 것을 의미하므로 그러한 이유를 파악해 보아야 한다. 이러한 과정에서 각 변수별 빈도표나 히스토그램을 이용하여 분포를 살펴보는 것은 매우 추천된다.

탐색적 요인분석이 정규성 가정 자체를 필요로 하지는 않지만 왜도(skewness, 치우친 정도)와 첨도(kurtosis, 뽕족한 정도)를 통해 측정변수들의 다른 특성들을 살펴보는 것도 바람직하다. 특히, 요인 추정방법으로 최대우도법을 사용할 경우 그 근거가 되므로 왜도와 첨도를 통해 정규성 여부를 평가할 필요가 있다. 왜도 및 첨도가 정규분포와 유의하게 다른가를 검토하기 위해서는 표준화값(standardized score, z-score)을 계산하여야 한다. 표준화값에 대한 기준은 유의수준 10% 하에서 1.65, 유의수준 5% 하에서 1.96, 유의수준 1% 하에서 2.58이다(표준화값이 아닌 왜도와 첨도의 원래 값에 대하여 잘못된 기준을 적용한 사례가 JKAN에서도 다수 발견되었다). SPSS를 사용하는 경우 왜도(첨도)를 그의 표준오차(standardized error)로 나누어 산출할 수 있으며, LISREL의 PRELIS 소프트웨어에서는 Z-Score라는 이름으로, AMOS 소프트웨어에서는 CR (critical ratio)이라는 이름으로 출력하여 주므로 이들 소프트웨어를 이용하

면 쉽게 왜도와 첨도의 표준화값을 얻을 수 있다. Table 2는 예시로 서 12개의 측정변수에 대하여 기술통계량을 제시한 예이다.

Table 2. Descriptive Statistics from an Example

Variables	Mean	Standard deviation		Skewness		Kurtosis	
		SD	CV(%)	Raw	Z-score	Raw	Z-score
x1	3.86	0.97	25.0	-0.76	-3.83	0.36	1.05
x2	4.29	0.88	20.6	-1.35	-5.94	2.01	3.32
x3	4.25	0.84	19.8	-1.26	-5.67	1.96	3.28
x4	4.32	0.74	17.0	-0.93	-4.52	0.65	1.61
x5	3.99	0.97	24.2	-1.02	-4.86	0.87	1.97
x6	3.38	1.20	35.6	-0.17	-0.99	-0.97	-5.16
x7	3.56	1.17	32.9	-0.53	-2.82	-0.54	-1.92
x8	3.60	1.05	29.2	-0.40	-2.17	-0.36	-1.07
x9	3.52	1.02	28.9	-0.22	-1.21	-0.41	-1.27
x10	2.38	0.97	40.7	0.27	1.52	-0.54	-1.93
x11	3.24	1.09	33.7	-0.30	-1.65	-0.62	-2.34
x12	2.33	1.03	44.2	0.54	2.84	-0.11	-0.18

CV=(SD/Mean) × 100(%)

3. 상관계수

상관계수는 요인분석의 직접적인 대상이 되므로 주의 깊게 살펴볼 필요가 있다. 너무 낮은 상관성은 해당 변수가 요인분석에 적합하지 않음을 나타낼 수 있으므로 이에 대한 검토가 필요하다. 즉, 각 측정변수는 같은 요인으로 묶일 것으로 기대되는 변수 그룹들과 상대적으로 높은 상관계수를 가져야 한다. Tabachnick과 Fidell (2007)은 0.3을 기준으로 상관계수의 크기를 살펴볼 것을 추천하였고, Hair 등(1995)은 상관계수가 ±0.3일 때 최소한도의, ±0.4일 때 중요한, ±0.5일 때 실제적으로 유의한 정도라고 분류하였다. 한편, 상관계수가 너무 커서 절대값이 1에 가까운지도 살펴보아야 한다. 어떤 두 측정변수의 상관계수가 1에 가까우면 이는 두 변수가 거의 동일한 의미이거나 같은 정보를 가지고 있다는 것을 나타낼 수 있기 때문이다.

분석대상이 되는 측정변수들이 매우 많을 경우에는 사전에 상관계수를 살펴보는 작업이 현실적으로 쉽지 않을 수 있다. 이런 경우는 먼저 요인분석을 수행한 후 그 결과를 살펴보는 과정에서 이상한 현상이 발견되는 측정변수에 대하여 평균, 표준편차, 상관계수 등을 살펴보는 것도 좋은 전략이라 할 수 있다.

4. KMO 표본적합성 측도와 Bartlett의 구형성 검정

Kaiser-Meyer-Olkin의 표본적합성 측도(measure of sampling ade-

quacy)는 관측된 상관계수들의 값과 편상관계수들의 값을 비교하는 지수로서, 이 값이 클수록 측정변수들 저변에 공통적인 잠재요인이 존재함을 나타낸다. Kaiser (1974)에 의거하면 KMO 측도의 값이 0.90보다 큰 경우를 훌륭한(marvelous), 0.80-0.89인 경우를 가치 있는(meritorious), 0.70-0.79인 경우를 중급의(middling), 0.60-0.69인 경우를 평범한(mediocre), 0.50-0.59인 경우를 빈약한(miserable), 0.5 이하를 받아들이기 힘든(unacceptable) 경우로 분류하였다. 즉, 이 측도의 값이 적으면 요인분석을 위한 변수들의 선정이 좋지 못함을 나타낸다.

Bartlett의 구형성(sphericity) 검정은 “상관계수 행렬이 단위행렬이다.”라는 귀무가설을 기각할 수 있는지를 검정하는 것이다. 즉, “귀무가설: 공통요인이 존재하지 않는다.”와 “대립가설: 공통요인이 존재한다.”를 대상으로 검정을 실시하는 것이다. 따라서, Bartlett 검정의 p -값이 유의수준(예컨대, 0.05)보다 작아서 귀무가설이 기각되어야 그 자료에 요인분석을 실시할 가치가 있음을 나타낸다.

KMO 측도와 Bartlett 검정은 요인분석을 수행하기 전에 자료가 최소한의 조건을 가지고 있는지를 검토하기 위한 것이다. 여러 문헌에서 KMO 측도와 Bartlett 검정을 통해 요인분석의 결과가 타당한 것으로 나타났다는 식으로 결론을 짓는 경우가 종종 있는데(JKAN에서도 이러한 서술이 일부 발견되었다), 요인분석 결과의 타당성은 요인적재값, 공통성 등 여러 가지 요인분석의 결과를 바탕으로 종합하여 판정할 문제이다.

요인분석의 수행과 결과 해석

1. 요인의 개수

요인적재값을 추정하기 위한 여러 방법들은 대부분의 경우 자료에 부합하는 요인의 개수를 미리 지정하도록 요구하고 있다. 대부분의 소프트웨어에서 Kaiser의 기준을 기본적으로 사용하고 있지만, 요인의 개수를 결정하는 데 있어 유일한 기준은 없으며, 여러 기준들이 주는 결과들을 종합적으로 검토하여 연구분야의 이론적인 바탕 위에서 합리적으로 판단하여야 한다. 연구자가 잠정적으로 상정하고 있는 요인 수를 포함하여 서로 다른 요인 수를 지정하여 요인분석을 수행하고 그 결과를 비교하여 결론을 짓는 것은 실제 연구에서 매우 추천된다.

Kaiser의 규칙은 표본상관행렬의 고유값 중 1보다 큰 개수만큼 요인을 보유하는 것이다(Table 3). 이 방법은 대부분의 통계분석 소프트웨어에서 초기설정(디폴트) 방법으로 사용되고 있으며, 일반적으로 연구자가 기대하는 요인의 개수와 일치하는 경우가 많을 뿐 아니라 요인의 개수를 미리 정해 놓고 인위적으로 만든 확률모형으로부터 취한 표본에 이 기준을 적용했을 때에도 매우 잘 작동되는 것

Table 3. Eigenvalues and Explained Variance (SPSS output)

Component	Initial eigenvalue			Extracted SS loadings			Rotated SS loadings		
	Sum	%Var	%Cum	Sum	%Var	%Cum	Sum	%Var	%Cum
1	2.93	24.4	24.4	2.93	24.4	24.4	2.35	19.6	19.6
2	1.71	14.2	38.7	1.71	14.2	38.7	1.81	15.0	34.7
3	1.27	10.5	49.3	1.27	10.5	49.3	1.75	14.6	49.3
4	.98	8.1	57.4						
5	.88	7.3	64.8						
6	.85	7.1	71.9						
7	.77	6.4	78.4						
8	.66	5.5	83.9						
9	.60	5.0	89.0						
10	.50	4.1	93.1						
11	.45	3.7	96.9						
12	.36	3.0	100						

으로 알려져 있다. 반면에, 이 방법에 의한 결과는 종종 보유되어야 할 것보다 적은 개수의 요인을 결정하게 할 수 있다는 점에 유의할 필요가 있다(Kang et al., 2005; Tabachnick & Fidell, 2007).

각 요인의 설명분산(explained variance) 비율 및 누적 비율에 대한 기준은 하나의 요인이 실제적으로 중요한 의미를 가지기 위해서는 그 요인이 전체 변이에 대해 가지는 공헌도가 최소한 얼마 이상이 되어야 한다는 것을 규정하는 것이다. Hair 등(1995)은 사회과학 분야에서 누적 설명분산이 보통 50-60% 정도는 되어야 한다고 제시하였다. 주축요인추출법이나 최대우도법의 경우 이 기준을 사용하기 위해서는 요인의 수를 변경하며 여러 번 요인분석을 수행해 보아야 한다.

Cartell과 Harman (1966)에 의해 제안된 스크리 검사(scree test) 방법은 고유값의 크기를 그래프에 나타낸 도형(scree graph)에 근거하여 적절한 요인의 수를 결정하는 것이다. 이 방법은 고유값의 크기를 단순히 그래프화한 것이어서, 이 기법에 의해 요인의 개수를 결정할 때에는 주관성이 개입되나 그림이 가지는 시각적 효과를 얻을 수 있다는 장점이 있다. 이 기법은 도형의 모양이 스크리 형태를 보여주지 않을 경우에는 적용하기가 상당히 어려우나, 만약 사소한 요인들이 상당히 존재하지만 흥미의 대상이 추가 되는 요인의 수를 찾는 데 있어서는 다른 방법들보다 우수하다는 사실이 여러 가지 모의실험의 결과를 통해 밝혀져 있다(Kang et al., 2005).

요인의 수를 0으로부터 시작하여 상정한 요인의 수에 대한 적합도 검정의 결과가 기각되지 않을 때까지 축차적으로 카이제곱 적합도검정을 수행하는 것을 사용할 수 있다. 이 방법은 정규성과 같은 몇 가지 가정이 충족되는 경우 통상적으로 만족스런 결과를 제공하는 것으로 알려져 있다(Lawley & Maxwell, 1971). 다만 이 방법은

Table 4. χ^2 Test for the Number of Factors (SAS output)

Tests	DF	χ^2	p
H0: No common factors H1: At least one common factor	66	356.0	<.001
H0: 1 Factor is sufficient H1: More factors are needed	54	141.7	<.001
H0: 2 Factor is sufficient H1: More factors are needed	43	88.1	<.001
H0: 3 Factor is sufficient H1: More factors are needed	33	50.5	.026
H0: 4 Factor is sufficient H1: More factors are needed	24	35.3	.063

일부 소프트웨어에서는 아직 제공하지 않는 경우도 있다. Table 4는 SAS 소프트웨어의 결과를 요약한 것으로, 3개 또는 4개의 요인을 적절한 요인 수로 고려할 수 있음을 보여주고 있다.

2. 요인 추출방법

요인을 추출한다는 것은 결국 자료(상관행렬)로부터 요인적재값 λ_i 와 유일성(특수성) ψ_i 를 추정하는 것이다. 대부분의 소프트웨어에서 제공하는 요인 추출방법은 매우 다양하지만 문헌에서 주로 언급되는 것은 주성분분석법(principal component method), 주축요인법(principal axis factor method), 최대우도법(maximum likelihood method) 등이다.

주성분분석은 서로 연관되어 있는 측정변수들을 서로 독립적인 소수의 인공변수인 주성분으로 차원축소를 꾀하는 것이며, 요인분석에서의 주성분분석법이란 단순히 첫 몇 개의 주성분을 요인으로 취하는 것이다. 이러한 주성분분석법은 측정변수들의 전체 분산이 최대한 설명될 수 있도록 주성분(요인)을 추출한다. 주성분분석법은 요인 추출과정이 단순하기 때문에 대부분의 소프트웨어에서 초기설정 방법으로 사용되며, 최종 결과를 얻기 전에 요인분석을 시험적으로 수행하고자 할 때 효율적이다. Snock과 Gorsuch (1989)는 주성분분석법의 경우 요인적재값의 크기가 실제보다 크게 나타나는 경향이 있음을 보였는데, 따라서, 주성분분석법을 실시했을 때 요인구조를 해석하기 쉬운 경향이 있으나 이는 잘못된 결과일 수 있으므로 주의해야 한다(Tak, 2007). 그러나 구성타당도 평가가 요인구조의 정확성보다는 변수들의 군집성에 관심을 둔다는 점을 고려할 때, 주성분분석법의 이러한 성질은 구성타당도 평가를 위해 도움이 되는 측면도 있다고 하겠다.

주축요인법과 최대우도법은 앞에서 설명한 요인분석 모형을 상정하여 측정변수들의 공분산이 최대한 설명될 수 있도록 요인을 추출한다. 따라서, 주성분분석법에 비하여 전체 분산에 대한 설명

비율은 작아질 수 있다. 일반적으로 요인분석의 목적이 많은 측정 변수들이 공통적으로 갖고 있는 의미 있는 구조를 추출하는 것이 목적이라고 할 때 주축요인법이나 최대우도법이 더 타당하다고 할 수 있다(Tak, 2007). 주축요인법이나 최대우도법에서는 최종 해(solution)를 얻기 위해 반복적인 수치적 방법을 이용하는데, 이 과정에서 해가 타당한 결과로 수렴하지 않거나 공통성의 추정값이 1을 넘는 소위 Heywood 상황이 발생할 수 있다. 이러한 상황은 여러 가지 복잡한 원인에 의해 발생할 수 있는데(Kang et al., 2005), 아무튼 Heywood와 같은 상황에서 얻어지는 해는 타당하지 못하다고 할 수 있다. 또한, 최대우도법은 다변량 정규분포의 가정 하에서 타당한 방법으로 측정변수들의 정규성이 대체로 만족될 경우 추천되는 방법이라고 할 수 있다.

많은 경험적 결과에 의하면 요인 추출방법에 따라 그 결과가 크게 달라지지는 않는다. 그러나 추정방법에 따라 요인적재값의 패턴이 달라질 수 있으므로 연구자는 여러 가지 추출방법으로 결과를 얻고 연구분야의 이론적 근거에 기초하여 타당한 요인구조를 선택하여야 한다.

3. 요인구조의 회전

구성타당도를 검증하기 위한 요인분석이 동일한 구성개념을 측정하는 변수들 간에는 상대적으로 높은 상관성(수렴성)을, 반면에 서로 다른 구성개념을 측정하는 변수들 간에는 상대적으로 낮은 상관성(변별성)을 가지는지를 알아보고자 하는 것이므로, 요인구조를 단순화하기 위한 회전은 실제로 필수적이라 할 수 있다.

연구자는 varimax 또는 quartimax와 같은 직교회전(orthogonal rotation)이나 oblimin 또는 promax와 같은 사각회전(oblique rotation)을 사용할 수 있다.

직교회전은 요인의 직교성(독립성)을 유지하면서 요인구조를 회전하는 방법이며, 직교회전 방법 중 varimax는 많은 연구자들에 의하여 추천되며 문헌에서 가장 많이 나타나는 방법이다. 요인분석의 실제 목적이 서로 상대적으로 독립적인 소수 몇 개의 요인을 유도하여 이를 통해 측정변수들 간의 구조를 파악하는 데 있다는 점을 중시할 때, 통상적으로는 직교회전을 이용하는 것이 더 바람직하다고 판단된다.

사각회전은 요인의 직교성을 포기하고 단순성이라는 측면에서 요인구조의 해석 그 자체에 더 큰 비중을 둔 개념이라고 할 수 있다. 그러나 사각회전에 의해 회전된 요인구조를 얻었을 경우, 요인의 해석은 쉬워질 수 있으나 그 대가로 요인들 간의 상관관계가 도입되어 관찰될 수 없는 요인들 사이의 상관성을 또 다시 설명해야 하는 문제가 제기된다. 따라서, 사각회전은 보조적인 도구로 조심스럽게 사

용될 필요가 있으며, 사각회전을 사용할 때에는 확증적 요인분석을 사용하는 것이 더 바람직하다는 주장도 있다.

4. 요인적재값과 설명분산

요인분석의 결과 해석 시에 필요한 사항은 요인구조(factor structure)를 통해 어느 측정변수가 어느 요인에 속하는지를 결정하고, 높은 적재값을 갖는 측정변수들의 공통적인 의미가 무엇인지를 파악하여 요인의 의미를 부여하는 것이다. 높은 적재값에 대한 기준도 다양하며 보통 절대값 0.3, 0.4, 0.5 정도가 사용된다.

앞에서 요인구조의 각 행(row)의 원소들의 제곱합 $c_i = \lambda_{i1}^2 + \lambda_{i2}^2 + \dots + \lambda_{im}^2$ 를 측정변수 λ_{ij} 의 공통성이라고 하며 이는 λ_{ij} 의 분산 중 m 개의 요인들에 의하여 설명되는 부분으로 해석될 수 있다는 것을 설명한 바 있다. 반면에 요인구조의 각 열(column)의 원소들의 제곱합 $\lambda_{1j}^2 + \lambda_{2j}^2 + \dots + \lambda_{mj}^2$ 는 측정변수들의 전체 분산(변이) 중 요인에 의해 설명되는 부분을 의미한다(SPSS의 경우 이것을 '회전 제곱합 적재값'이라는 이름으로 출력한다).

여러 문헌에서 고유값과 설명분산(explained variance)을 혼동하여 제시하는 경우가 많이 있다(JKAN에서도 이러한 서술이 다수 발견되었다). 사실 고유값은 요인분석의 이전 단계에서 요인분석과는 독립적으로 계산되는 것이며, 회전된 요인과 1 대 1로 대응되지 않는다. 따라서, Table 5에서와 같이 요인구조와 함께 제시되어야 할 결과는 각 요인에 의하여 설명되는 분산(회전 제곱합 적재값)이다.

Table 5. Factor Loadings, Communalities, and Explained Variances

Variables	Principal component method			Principal axis factor method				
	Factor structure	Communalities		Factor structure	Communalities			
x1	.72	-.06	-.10	.53	.57	.01	-.13	.34
x2	.56	.29	-.26	.47	.48	.26	-.26	.37
x3	.77	.07	-.03	.61	.68	.10	-.07	.48
x4	.78	.13	-.12	.64	.71	.16	-.14	.56
x5	.27	.31	-.54	.46	.25	.25	-.45	.33
x6	.13	.49	-.09	.27	.13	.33	-.08	.13
x7	.10	.76	.05	.59	.08	.71	.03	.51
x8	-.10	.65	-.07	.45	-.01	.41	-.03	.17
x9	.33	.49	.32	.46	.24	.42	.19	.28
x10	-.14	.28	.59	.46	-.11	.19	.44	.25
x11	.07	-.05	.72	.53	.00	-.01	.54	.29
x12	-.20	-.06	.59	.39	-.18	-.05	.43	.22
Explained variance	2.35	1.81	1.75		1.74	1.18	1.05	
Explained (%)	19.6	15.1	14.6		14.6	9.9	8.8	
Cumulative (%)	19.6	34.7	49.3		14.6	24.4	33.2	

Orthogonal varimax rotation method is used.

추가적인 이슈

1. 확증적 요인분석

탐색적 요인분석(Exploratory Factor Analysis [EFA])은 요인의 개수나 구조에 대한 특별한 가정 없이 자료에 내재되어 있는 특성을 탐색하여 관심 있는 가설이나 모형 및 구조 등을 생성함으로써 추가적인 연구를 수행하는데 기여한다. 그러나 탐색적 요인분석을 사용하는 데 있어 몇 가지 제한 사항이 존재한다. 첫째, 연구자는 모든 요인들이 서로 상관되어 있지 않거나(직교회전의 경우) 모두 상관되어 있는(사각회전의 경우) 것으로 가정해야 한다. 즉, 특정한 요인들 사이에만 상관성을 두거나 없앨 수 없다. 둘째, 모든 측정변수들은 모든 요인에 적재된다고 가정해야 한다. 즉, 특정 측정변수를 특정 요인에 배정하거나 다른 요인에 배정할 수 없다.

확증적 요인분석(Confirmatory Factor Analysis [CFA])은 구조방정식모형의 특수한 경우로서 탐색적 요인분석의 위와 같은 단점을 해소할 수 있을 뿐만 아니라, 요인구조에 대한 연구자의 가설을 모형화하고 실제 자료에 의한 가설의 지지여부를 모형의 부합도라는 관점에서 검증할 수 있다(Kang, 2013).

확증적 요인분석이 정확한 가설검정에 필요한 통계적 도구를 제공하지만, 어떠한 과정을 어디까지 사용할 것인지는 궁극적으로 연구자에게 달려 있다. 구체적으로 계획된 예측이 있어도 연구자는 탐색적 요인분석을 사용할 수 있으며, 반면 확증적 요인분석도 상당히 탐색적 방식으로 사용될 수 있다(Munro, 2005). 일반적으로 측정변수의 수가 많을수록 타당도는 떨어지게 된다(Seong, 2002). 또한, 확증적 요인분석은 제2종 오류의 위험성이 높기 때문에(Munro), 측정변수의 수가 많을 경우 받아들일만한 적합도를 얻기 위해 비상식적인 모수(parameter)를 추가해야 하는 마찰이 일어날 수 있다. 따라서, 요인 수에 비하여 측정변수가 매우 많을 경우에는 확증적 요인분석의 남용을 자제할 필요가 있다.

2. 반영적 지표와 형성적 지표

반영적 지표(reflective indicator)는 측정변수의 값이 잠재변수의 영향을 받아 결정되는 구조로 잠재변수가 측정변수의 원인임을 가정한다. 행동과학 분야에서 보편적으로 사용하는 지능이나 성격과 같이 이론적이며 추상적인 개념들이 반영적 지표에 해당한다. 반영적 지표는 측정변수의 값이 잠재변수의 영향으로 인하여 변화될 수 있기 때문에 측정변수들 사이에 상관관계가 높을 수 있으며 측정변수의 변화가 잠재변수인 개념을 변화시키지 않는다는 특징이 있다(Cho, 2011; Jarvis, MacKenzie, & Podsakoff, 2003). 반영적 지표에

서 측정변수들은 해당 구성개념의 함수이므로 동일한 구성개념의 측정변수 값들 간에는 내적일관성이 기대된다(Lee & Kim, 2013). 따라서, 내적일관성 신뢰도 및 요인분석을 통해 반영적 지표의 타당성을 검토할 필요가 있다.

형성적 지표(formative indicators)는 측정변수가 잠재변수의 값을 결정하는 구조로 측정변수가 잠재변수를 형성하는 원인임을 가정한다. 사회경제적 지위(SES)와 같은 지표(index)가 대표적인 형성적 지표에 해당한다. 또한, 특정 부분에 대한 지식의 경우도 형성적 지표에 해당될 수 있다(예를 들어, 골다공증이나 유방암에 대한 지식을 측정하는 20개의 문항). 사회경제적 지위는 교육수준이나 직업, 수입 등에 의해 영향을 받으므로, 교육수준, 직업, 수입 등 측정변수의 변화에 의해 잠재변수인 사회경제적 지위가 변화되지만, 잠재변수의 변화가 측정변수의 변화를 수반하지는 않는다(Chin, 1998). 이러한 형성적 지표는 잠재변수에 의해 공통적으로 영향을 받아 측정변수의 값이 결정되는 것이 아니기 때문에 측정변수들 간의 높은 상관관계와 내적일치도를 필요로 하지 않는다(Bollen & Lennox, 1991). 따라서, 형성적 지표를 이루는 측정변수들은 내적일관성 신뢰도 또는 요인분석의 대상이 되지 않는다. 그리고 측정변수의 추가 또는 제거가 잠재변수에 해당하는 개념의 변화로 이어질 수 있다는 특징을 가지므로(Jarvis et al., 2003), 주로 주관적 판정에 의해 측정변수의 선정이 결정된다.

형성적 지표로 구성되어야 하는 측정모형이 반영적 지표로 간주되거나 반영적 지표로 구성되어야 하는 측정모형이 형성적 지표로 간주되는 것은 편향된 모수추정 결과의 원인이 되며, 오류의 가능성을 높일 수 있다. 따라서, 연구자는 측정모형이 반영적 지표나 형성적 지표로 구성될 수 있다는 것을 인지할 필요가 있으며, 변수에 대한 이론적 배경 등을 고려하여 측정모형을 정의하고 타당한 통계 방법론을 적용해야 한다(Cho, 2011).

결론

탐색적 요인분석은 이름 그대로 탐색적인 자료 분석 방법이며, 많은 분석 단계를 포함하는 복잡한 통계분석 기법이다. 또한, 요인분석을 적용함에 있어 수많은 옵션과 경험적 기준이 존재한다. 따라서, 연구자는 다양한 옵션들과 기준들을 조합하여 가급적 다양한 형태의 결과를 얻고 그들을 세심하게 비교하여, 연구분야의 이론과 경험에 부합하는 합리적인 결론을 내릴 수 있도록 노력해야 한다. 본 논문에서는 탐색적 요인분석의 수행과 관련된 여러 가지 사항에 대하여 통합된 관점의 지침을 제공하고자 하였다. 본 논문에서 논의된 내용들이 요인분석을 수행하고자 하는 많은 연구자들에게 실제적 도움이 되기를 기대한다.

REFERENCES

- Arrindel, W. A., & van der Ende, J. (1985). An empirical test of the utility of the observations-to-variables ratio in factor and components analysis. *Applied Psychological Measurement, 9*(2), 165-178. <http://dx.doi.org/10.1177/014662168500900205>
- Barrett, P. T., & Kline, P. (1981). The observation to variable ratio in factor analysis. *Personality Study and Group Behavior, 1*, 23-33.
- Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin, 110*(2), 305-314.
- Cartell, L., & Harman, A. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245-276.
- Chin, W. W. (1998). Issue and opinion on structural equation modeling. *MIS Quarterly, 22*(1), 1-10.
- Cho, H. D. (2011). *A study on issues of using structure equation modeling in education study*. Unpublished master's thesis, Korea University, Seoul.
- Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Guilford, J. P. (1956). *Psychometric methods* (2nd ed.). New York, NY: McGraw-Hill.
- Hair, J. F. Jr., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis* (4th ed.). Upper Saddle River, NJ: Prentice Hall.
- Hogarty, K. Y., Hines, C. V., Kromrey, J. D., Ferron, J. M., & Mumford, K. R. (2005). The quality of factor solutions in exploratory factor analysis: The influence of sample size, communality, and overdetermination. *Educational and Psychological Measurement, 65*(2), 202-226. <http://dx.doi.org/10.1177/0013164404267287>
- Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research, 30*(2), 199-218. <http://dx.doi.org/10.1086/376806>
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31-36.
- 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.
- Kang, H., Han, S. T., Kim, K., & Jhun, M. (2005). *Multivariate data analysis using SAS by examples*. Paju: Freedom Academy.
- Lawley, D. N., & Maxwell, A. E. (1971). *Factor analysis as a statistical method* (2nd ed.). New York, NY: American Elsevier Pub. Co.
- Lee, H., & Kim, J. H. (2013). *Structural equation modeling and AMOS 20.0*. Seoul: JypHyunJae Publishing Co.
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*(2), 130-149.
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84-99.
- Munro, B. H. (2005). *Statistical methods for health care research* (5th ed.). Philadelphia, PA: Lippincott Williams & Wilkins.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.
- Seong, T. J. (2002). *Validity and reliability* (2nd ed.). Seoul: Hakjisa Publisher.
- Shin, H. S., Hyun, M. S., Ku, M. O., Cho, M. O., Kim, S. Y., Jeong, J. S., et al. (2010). Analysis of research papers published in the Journal of the Korean Academy of Nursing—focused on research trends, intervention studies, and level of evidence in the research. *Journal of Korean Academy of Nursing, 40*(1), 139-149. <http://dx.doi.org/10.4040/jkan.2010.40.1.139>
- Snook, S. C., & Gorsuch, R. L. (1989). Component analysis versus common factor analysis: A Monte Carlo study. *Psychological Bulletin, 106*(1), 148-154. <http://dx.doi.org/10.1037/0033-2909.106.1.148>
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson/Allyn & Bacon.
- Tafreshi, M. Z., & Yaghmaei, F. (2006). Factor analysis of construct validity: A review of nursing articles. *Journal of Medical Education, 10*(1), 19-26.
- Tak, J. K. (2007). *Psychological testing: An understanding of development and evaluation method* (2nd ed.). Seoul: Hakjisa Publisher.
- Williams, B., Brown, T., & Onsmann, A. (2012). Exploratory factor analysis: A five-step guide for novices. *Journal of Emergency Primary Health Care, 8*(3), Article 1.