조사연구(2016. 2) 17권 1호: 71-83

특집논문

구조방정식모형의 모형평가 오·남용과 교정

Wrong Applications with Overall Model Evaluations and Its Corrections in Structural Equation Modeling

이기종[a)]
Kijong Rhee

수집된 자료에 설정된 모형이 적합한가의 여부를 따지는 구조방정식모형의 모형평가에서 널리 사용되는 정량적 지표가 χ^2와 RMSEA이다. 그러나 구조방정식모형에서 사용되는 통계검증의 논리와 유의수준 설정의 성격이 여타의 일반통계검증과는 다름에도 이를 일반통계검증처럼 취급하는 오남용이 빈번하게 일어난다. 이 글은 이런 차이점을 밝혀 구조방정식모형의 타당한 사용을 돕는 데 일조하고자 한다.

주제어: 구조방정식모형, 모형평가, χ^2, RMSEA

Some wrong applications in the area of model evaluation with the use of structural equation modeling are mainly associated with the uses of chi-square test and RMSEA. The reason why those wrong applications happen is that even though the logic of hypothesis testing with structural equation modeling is different from that of general hypothesis testing, many researchers do not recognize the differences between them. This paper focuses on the differences of the two different hypothesis testings and in consequence prevents

a) 국민대학교 교육학과 교수 이기종.
E-mail: rhee0408@kookmin.ac.kr

wrong applications with overall model evaluation in structural equation modeling.

Key words: structural equation modeling, model evaluation, χ^2, RMSEA

Ⅰ. 들어가며

사회과학 연구에서 주요 관심은 어떤 현상을 구성하고 있는 변수 간의 인과구조를 밝히는 것이다. 전통적으로 실험을 통해 인과성을 확보하였으나 실험이 여의치 않은 비실험적 상황에서 변수 간의 인과구조를 밝히는 수단으로 널리 사용되는 것이 구조방정식모형이다.

그러나 구조방정식모형을 이용한 연구가 확산되면서 올바른 이해 없이 사용된 구조방정식모형의 오용과 남용에 따른 부작용도 심각해지고 있다. 이런 오·남용은 현상을 왜곡하여 다른 그릇된 결론을 내리게 만들기도 하지만, 더욱 심각한 부작용은 오·남용에서 빚어진 그릇된 결과가 많은 후속연구를 잘못된 방향으로 인도하는 길잡이 역할을 한다는 것이다. 올바른 이해 없이 쓰인 잘못된 연구결과가 전문 학술지에 게재되는 탓에 이런 그릇된 연구를 후속연구가 인용하게 되고. 그 결과 잘못된 것이 제대로 된 것처럼 포장되어 진리처럼 여겨지게 되는 심각한 부작용을 빚는 것이다.

특히 구조방정식모형의 사용에서 빈번히 발생하는 오류는 모형평가에 관한 것이다. 설정된 모형이 수집된 자료에 적합한가를 따지는 모형평가에서 오·남용이 빈번한 것은 모형평가의 논리를 곡해하거나 자의적으로 받아들이고 있기 때문이다. 이런 오·남용은 크게 두 가지로 구분될 수 있다. 하나는 모형평가의 논리에 관한 곡해이며, 또 다른 하나는 모형평가 관련 적합도 지표에 관한 곡해이다. 이 글은 모형평가의 논리와 적합도 지표에 대한 직관적 접근을 통해 곡해를 일소하여 구조방정식모형의 올바른 사용에 일조하고자 한다.

Ⅱ. 모형적합도 판정의 틀: 자료인가 또는 모형인가?

모형적합도를 논의함에 있어 무엇보다 우선해야 할 것은 모형적합 여부를 결정짓는 판정의 중심이 무엇인가에 관한 질문이다. 판정의 중심을 어디에 두어야 하는가에 관한 질문으로 다음처럼 표현될 수 있다. 설정된 모형이 수집된 자료에 적합한 것인가? 아니면 수집된 자료가 설정된 모형에 적합한 것인가? 이 두 질문은 물론 부적합의 관점에서 표현될 수도 있다: 설정된 모형이 수집된 자료에 들어맞지 않는 것인가? 아니면 수집된 자료가 설정된 모형에 들어맞지 않는 것인가?

적합의 관점에서 진술된 것이든 부적합의 관점에서 진술된 것이든, 두 질문이 묻는 것은 무엇이 적합도 판정의 중심이어야 하는가이다. 수집된 자료에 중심을 두고 설정된 모형의 적합 여부를 판단해야 할 것인지, 아니면 설정된 모형에 중심을 두고 자료의 적합 여부를 판단해야 할 것인지에 관한 사려 깊은 판단이 요구되는 것이다.

구조방정식모형에서뿐만 아니라 표본을 대상으로 하는 연구에서는 언제나 수집된 표본 자료는 모집단을 잘 대표하도록 추출된 것임을 전제로 한다(이종성 외 2007; 정대연 1997) .이런 전제가 타당한 것이라면 모형적합도 판정은 자연스럽게 설정된 모형이 아니라 수집된 자료에 중심이 놓이게 된다. 따라서 설정된 모형이 수집된 자료에 들어맞지 않을 때, 모형을 재설정하여 적합도를 개선해야 하는 것이지 표본 자료가 모형에 적합하지 않아 자료를 재수집하거나 교정해야 한다는 표현은 잘못된 것이다. 모형적합도 판정에서 중심은 어디까지나 표본 자료이지 설정된 모형이 아닌 것이다.

Ⅲ. 정량적 판정과 정성적 판정

구조방정식모형에서 모형적합도를 판정하는 방법은 크게 두 가지로 분류될 수 있다. 하나는 도출된 통계량을 사용해 통계검증이 가능한 정량적 판정이고,

또 다른 하나는 통계검증이 불가능한 정성적 판정이다. 영가설이 맞다고 할 때 추출된 통계량의 분포를 이용해 가설검증이 가능한 것이 정량적 판정인 반면에, 추출된 통계량이 있으나 이 통계량의 분포를 알 수가 없어 통계검증이 불가능한 것이 정성적 판정이다(이기종 2012).

여기에서는 특히 오·남용이 심각한 정량적 판정에 초점을 둔다. 정성적 판정도 중요하나 기본적으로 정성적 판정은 판정을 위한 잣대 설정이 임의적이기 때문에 어떤 일반적 기준을 마련할 수 없기 때문이다. 그 결과 정량적 판정의 잣대로 보면 모형적합도가 낮은 경우라도 연구자가 적합하다고 강변할 소지가 있어, 정량적 판정이 우선시되고 정성적 판정은 보조역할을 하는 데 그쳐야 한다.

1. 정확적합도 지표 χ^2

설정된 모형이 수집된 자료에 잘 들어맞는지의 여부를 따지는 모형적합도 판정에서 가장 널리 사용되는 것이 중심 χ^2분포(central χ^2 distribution)를 이용한 χ^2검증이다. χ^2검증은 설정된 모형이 모집단에서 변수 간의 구조와 같다는 영가설이 맞다고 할 때, $(N-1)\times F$로 정의된 χ^2검증통계량이 중심 χ^2분포를 이룬다는 것에 기초한다(Jöreskog & Sörbom 2015). 여기서 F는 적합함수(fitting function)의 머리글자로 구조방정식모형에서 연구자가 모수추정에 사용한 적합함수의 값이다. 예를 들어, 최대가능법이 적합함수로 사용되었다면 아래 수식이 구해낸 값이 F이다.

$$FML = \log|\Sigma(\theta)| + tr[S\Sigma(\theta)-1] - \log|S| - (p+q)$$

$(N-1)\times F$로 정의된 χ^2 검증통계량은 $[\frac{1}{2}(p+q)(p+q+1)-t]$의 자유도를 갖는 χ^2분포이다. 여기서 p는 설정된 모형에서 y변수의 숫자, q는 x변수의 숫자, t는 자유모수의 숫자이다. χ^2검증이 널리 사용되는 것은 영가설이 맞다고 가정할 때 검증통계량의 분포를 알 수 있기 때문이다.

χ^2분포를 이용한 모형평가를 정확적합도라고 부르는 것은 구조방정식모형의 영가설과 관계가 있다. 구조방정식모형에서 영가설은 모집단에서의 변수간의 구조는 설정된 모형과 같다는 것이다. χ^2통계량은 바로 이 영가설, 즉 모집단공분산행렬은 추정된 자유모수의 함수인 모형공분산행렬과 같다는 것에 기초해 산출된다. 이 영가설을 기호를 사용해 수리적으로 나타내면 다음과 같다.

$$H_0 : \Sigma = \Sigma(\theta)$$

여기서 Σ는 모집단공분산행렬, θ는 추정된 자유모수를 담고 있는 벡터, 그리고 $\Sigma(\theta)$는 추정된 자유모수 θ의 함수로 표현된 모형공분산행렬이다. 이 영가설이 의미하는 바는, 모집단에서 여러 변수간의 구조를 나타내는 모집단공분산행렬은 선행이론에 의해 설정된 모형에 의해 나타내어지는 모형공분산행렬과 같다는 것이다. 따라서 영가설이 맞다면 모집단에서의 여러 변수 간의 관계를 나타내는 구조는 설정된 모형에 의해 생성되는 여러 변수 간의 관계를 나타나는 구조와 같다는 것이다. 즉, 모집단에서 변수간의 관계구조를 나타내는 모집단공분산행렬은 설정된 모형이 정확하다면 설정된 모형에 의해 생성되는 모형공분산행렬과 같다는 것이다.

따라서 영가설이 실재를 있는 그대로 반영하는 정확한 것이라면, 달리 말해 영가설이 타당한 것이라면, 들어맞고 남은 잔차(fitted residual)의 관점에서 영가설을 다시 표현하면 다음처럼 된다.

$$H_0 : \Sigma - \Sigma(\theta) = 0$$

이렇게 표현된 영가설이 의미하는 바는 모집단공분산행렬에서 모형공분산행렬을 뺀 잔차행렬은 모든 원소가 0인 영행렬이라는 것이다. 잔차행렬의 모든 원소가 0이라는 것은 모집단공분산행렬과 모형공분산행렬이 정확하게 일치한다는 것이며 이런 의미에서 이 영가설에 기초한 χ^2검증이 정확적합도라고 명명되는 것이다(Bollen 1989). 영가설이 맞다면 모집단공분산행렬에 모형공분산행렬이 정확하게 들어맞게 되고 그 차이는 0이 되어 두 개의 행렬은 서

로 같다고 판정되는 것이다.

이 영가설이 맞다는 전제 하에 모집단을 잘 대표하는 표본이 하나 추출되고, 이 표본에서 표본공분산행렬을 구한다. 그러면 이 표본공분산행렬은 모집단공분산행렬을 잘 대표하게 된다. 만약 영가설이 연구자가 관심을 갖는 현상의 실재를 잘 반영하고 있다면, 모수추정을 위해 어떤 적합함수가 사용되더라도 적합함수의 값은 0이 되며, 그 결과 $(N-1)\times F$로 정의된 χ^2값도 0이 된다. 따라서 χ^2가 표본크기에 민감해 모형적합도 지표로 바람직하지 않다는 주장은 논거가 취약하다. 영가설이 맞다면 적합함수의 값은 0이 되며 그 결과 표본크기가 어떠하든지 χ^2검증통계량도 0이 되기 때문이다.

위에서처럼 표본을 취하는 과정이 무한대로 시행되면, 표본이 아무리 모집단을 잘 대표한다고 해도 표본 그 자체가 모집단이 아님으로 해서 표본변산이 있게 마련이다. 그 결과 영가설이 맞는다고 해도 이런 표본변산으로 인해 χ^2값은 어떤 경우들에 있어서는 0보다 아주 큰 값이 생기게 되나 이런 경우는 영가설이 맞다면 드물게 일어나게 된다. 이렇게 무한대로 시행된 표본에서 나온 χ^2값들로 빈도분포를 구성하면, 대부분의 χ^2값은 0에 집중되고 0보다 아주 큰 χ^2값은 드물어 분포의 꼬리가 오른쪽으로 있는 정적편포를 이룬다.

이 분포가 영가설이 맞을 때 χ^2통계량이 이루는 이론적 분포인 것이다. 여기서 알 수 있는 것은 영가설이 맞다면 χ^2값은 작은 값을 갖게 되며, 그 결과 이런 작은 χ^2값이 관찰될 확률은 높아지게 된다는 것이다. 또한 영가설이 맞는다 해도 표본변산으로 인해 0보다 훨씬 큰 χ^2값이 발생할 수 있으나, 이런 큰 χ^2값은 영가설이 맞으면 드물게 발생하므로 큰 χ^2값이 관찰될 확률은 적어지게 된다는 것이다. 가설검증에서 0에 가까운 χ^2검증통계량이 도출되었다면 작은 그 χ^2검증통계량이 관찰될 확률이 높아지는 것이며 이는 영가설이 맞는 경우에서만 일어난다. 이와는 반대로 영가설이 맞지 않는다면 0에서 멀리 떨어진 큰 χ^2검증통계량이 도출되며 이 큰 χ^2검증통계량이 관찰될 확률은 높아지게 되는 것이다.

χ^2검증도 다른 통계검증과 마찬가지로 영가설이 맞다고 할 때 우연하게 일

어날 수 있는 경우를 유의수준으로 정하게 되며, 표본에서 계산된 χ^2통계량이 설정된 유의수준의 경계값보다 작으면 영가설을 수용하게 되고, 반대로 경계값보다 크면 영가설을 기각하게 된다. 구조방정식모형에서 설정되는 유의수준은 .05 이상이며 많은 경우에서 .05보다 훨씬 큰 값이 설정된다.

구조방정식모형에서는 영가설이 맞으면 그 결과 작은 χ^2값을 갖게 되고 따라서 영가설을 수용하게 된다. 이는 수집된 자료인 표본공분산행렬과 모형에서 도출된 모형공분산행렬이 유의수준이 정한 오차한계 내에서 서로 같음을 나타낸다. 이를 확률의 관점에서 표현하면, 표본에서 계산된 χ^2값이 나타내는 확률수준은 영가설이 맞다고 할 때 그 표본 χ^2값이 나타날 확률이다. 그러므로 계산된 χ^2값이 나타내는 확률이 클수록 설정된 모형은 수집된 자료인 표본공분산행렬에 더 잘 들어맞는 것이다.

또한 자유도의 관점에서 보면 χ^2검증은 자유도를 검증하는 것이다(Bollen 1989). 다른 조건이 모두 동일하다고 하면, 자유도가 작은 모형이 자유도가 큰 모형에 비해 상대적으로 양호한 적합도를 보이게 된다. 자유도가 작은 모형은 추정되어야 할 자유모수의 숫자가 많아 설정된 모형이 수집된 자료에 잘 들어맞게 되나, 자유도가 큰 모형은 자유모수의 숫자가 적어 모형이 자료에 잘 들어맞지 않게 된다. 모집단 변수구조를 나타낼 때 가능한 경로가 모두 모형에 설정되면 그 숫자만큼 자유모수가 필요하게 되어 모형이 포화된다. 포화모형은 언제나 완벽한 적합도를 보이기는 하나 자유도가 0이라는 것에서 알 수 있듯이 χ^2검증은 자유도를 검증하는 것이다. 자유도가 작아질수록 모형이 자료에 가까워져 적합도가 개선되는 이점이 있어 연구자는 자유도를 희생해 적합도를 개선하는 선택을 고려할 수 있다. 그러나 자유도가 작은 모형은 큰 모형에 비해 최소한의 변수로 현상을 설명하려는 간명성이 떨어지는 단점이 있어, 적합도 개선을 선택할 것인지 또는 간명성의 수준을 올릴 것인지 동시적으로 고려해야 한다.

그러나 χ^2검증도 모든 면에서 완벽한 것은 아니어서 흠결이 없지는 않다. 흠결로 많이 지적되는 것은 χ^2검증이 연구모형, 표본크기, 그리고 변수분포에 영향을 받는다는 점이다. 그러나 주지하는 바와 같이 이러한 지적은 χ^2검증에

만 해당되는 것은 아니며, 어떤 종류에 관계없이 통계량을 사용하는 통계검증에는 예외 없이 적용되는 흠결인 것이다.

2. 근사적합도 지표 RMSEA

구조방정식모형에서의 영가설은 설정된 모형이 모집단에서의 변수 간의 구조와 같다는 것이다. 그러나 이 영가설은 정확적합도의 경우에서 보듯 잔차행렬의 모든 원소가 0인 영행렬이라는 것이며 이는 현실적으로 가능하지 않은 무리한 것이라는 주장이 있다(Kline 1998). 모집단의 관점에서 보면 영가설이 타당할 수 있으나 실제 연구는 모집단을 잘 대표하는 표본을 대상으로 하는 것임에도 이런 영가설을 가정하는 것은 현실을 도외시한 것이라는 지적이다. 그러므로 완벽히 정확하게 들어맞는 모형이 아닌 적당하게 들어맞는 모형은 실재를 제대로 반영하고 있어도 결국 정확하게 들어맞지 않음으로 인해 표본이 큰 연구에서는 영가설이 기각되는 결과를 초래하게 된다는 지적이다.

Browne & Cudeck(1993)은 정확적합도 대신에 적당하게 들어맞는 근사적합도를 검증하는 방법인 $RMSEA$를 제안하고 있다. $RMSEA$는 다음과 같이 정의된다.

$$RMSEA = \sqrt{\frac{F - \frac{df}{N-1}}{df}}$$

위 $RMSEA$의 정의에서 분자는 모집단공분산행렬과 모형공분산행렬의 차이를 나타내는 적합함수 값에서 모형의 자유도를 표본크기로 교정한 값을 뺀 것이다. 분자는 수집된 자료와 설정된 모형 간의 차이를 나타내는 것으로, 설정된 모형이 얼마나 잘 들어맞는가에 초점을 맞춘다. 설정된 모형이 복잡하면 자료에 잘 들어맞게 되며 그 결과 적합함수 값은 작아진다. 분자는 설정된 모형의 복잡성을 고려하지 않는다. 이런 점에 제재를 가한 것이 자유도로 분자를 나눈 것이다. 분자에 대한 분모의 비율에다 제곱근을 취한 것은 분자의 성질이 제곱값이기 때문에 해석이 쉽지 않아 보다 해석을 용이하게 하기 위해서이다.

편차점수제곱합을 표본크기로 나눈 분산에 제곱근을 취해 표준편차를 만든 것과 같은 이치라고 할 수 있다.

$RMSEA$에서 한 가지 제한점은 영가설이 맞다면 적합함수의 값은 0이 되어 분자는 음수값을 갖게 되며, 그 결과 제곱근 안의 값이 0보다 작은 음수가 나온다는 점이다. 이럴 경우를 대비해 $RMSEA$의 값이 음수가 나오면 0을 택하게 하는 max($RMSEA$, 0)가 사용된다.

일반적으로 $RMSEA$의 값이 작을수록 설정된 모형이 자료에 잘 들어맞는 것을 나타낸다. $RMSEA$가 .05보다 작으면 양호한 상태를 나타낸다고 하나 Browne & Cudeck(1993)이 지적하는 바와 같이 이 기준은 본질적으로 주관적이다. 그럼에도 Hu & Bentler(1999)는 .10 이상이면 나쁨, .10보다 작으면 보통, .08보다 작으면 수용가능, .05보다 작으면 근사, .00이면 정확적합이라는 견해를 표명하고 있다.

근사적합도 $RMSEA$의 검증은 이것이 .05와 같거나 작다는 영가설을 설정하고 이를 검증하는 것이다. 영가설을 기호로 나타내면 다음과 같다.

$$H_0 : RMSEA \leq .05$$

그러나 이때 통계량이 관찰될 확률은 정확적합도에서와 같은 방법으로 구해지는 것이 아니고, 비중심모수(noncentrality parameter) δ에 의해 결정되는 비중심 χ^2분포(noncentral χ^2distribution)의 누적분포함수를 1에서 뺀 것으로 결정된다. 위에 진술된 영가설에서 비중심모수 δ는 $N \times df \times .05^2$이다. 만약 $RMSEA$가 .05가 아닌 다른 값으로 설정된다면 비중심모수 δ는 $N \times df \times$다른 값의 제곱이다. 근사적합도에서 이 영가설은 검증통계량이 관찰될 확률이 미리 정해진 유의수준보다 작다면 기각된다.

여기서 하나 덧붙일 것은 위의 영가설은 등호가 아닌 부등호의 형태로 진술되었다는 점이다. 다시 말해 $H_0 : RMSEA = .05$가 아닌 $H_0 : RMSEA \leq .05$로 진술된 것이다. 그러나 Browne & Cudeck(1993)이 진술한 것과 같은 부등호와 등호가 결합된 형태는 영가설이 맞다고 가정할 때 표본분포의 모습을 알

수 없다는 취약점이 있다. 예컨대, $RMSEA$가 .05보다 작다면 그것이 어떤 값을 지칭하는 것인지 불명료하게 되고, 그 결과 영가설이 맞다고 할 때 표본분포를 알 수가 없어서 영가설 수용역과 기각역을 구분하는 경계값을 설정할 수 없게 되는 것이다. 따라서 영가설이 부등호와 등호가 결합된 형태로 진술될 수 있어도 이는 부정확한 표현이며, 실제 가설검증에서는 등호로 진술된 영가설을 사용하게 된다.

정확적합도 영가설인 $H_0 : \Sigma - \Sigma(\theta) = 0$을 근사적합도 관점에서 표현하면 다음과 같다.

$$H_0 : RMSEA = .05$$

정확적합도와 근사적합도의 영가설이 다른 점은 정확적합도에서는 $RMSEA$가 0이라는 것이며, 근사적합도에서는 이것이 .05보다 작거나 같다는 것이다.

두 영가설의 차이점을 구체적으로 비교해 보자. 정확적합도에서는 설정된 모형이 모집단에서 변수구조를 정확하게 반영한다는 것이며, 만약 정확하게 반영하지 못한다면 이때 차이가 통계적으로 받아들일 수 있는 차이인지에 관심을 두는 것이다. 근사적합도에서는 설정된 모형은 모집단에서의 변수구조를 어느 정도는—예를 들면, .05만큼—정확하게 반영하지 못하고 있다는 것이며, 이런 가정하에 차이가 받아들일 수 있는 것인지에 관심을 둔다. 근사적합도에서 어느 정도 덜 맞는 판단의 준거가 .05이어야 하는지, 또는 이보다 작은 값 혹은 큰 값이 되어야 하는지는 Browne & Cudeck(1993)이 진술한 것처럼 본질적으로 주관적이다.

Ⅳ. 구조방정식모형과 일반통계검증의 차이

구조방정식모형과 일반통계검증의 차이는 영가설을 보는 시각에서도 확연히 다르다. 구조방정식모형에서는 영가설이 수용되는가에 관심을 두지만, t검증 같은 일반통계검증에서는 영가설이 기각되는가에 관심을 둔다. 이런 차이

가 있는 것은 구조방정식모형에서는 영가설에 연구자가 관심을 갖는 현상이 투영되어 있지만, 일반통계검증에서는 대안가설에 연구자가 관심을 갖는 현상이 투영되어 있기 때문이다.

구조방정식모형에서 영가설이 수용된다는 것은 설정된 모형이 모집단에서의 변수 간의 구조를 잘 나타낸다는 것이며, 이는 모형으로 표현된, 연구자가 관심을 갖는 현상이 모집단에서도 그러하리라는 유추를 가능하게 한다. 이런 이유로 구조방정식모형에서는 영가설 수용에 관심을 두는 것이다. 반면에 영가설이 기각된다면, 모형으로 표현된, 연구자가 관심을 갖는 현상이 모집단에서의 그것과는 다른 것임을 시사하는 것이어서 판정의 중심이 되는 모집단에 적합한 다른 변수구조를 갖는 현상이 존재하고 있는 것이다. 이는 모형의 타당성이 훼손되는 것이어서 선행이론이나 연구자의 통찰에 의해 설정된 모형이 개선되어야 함을 뜻하는 것이다.

일반통계검증이 영가설 기각에 관심을 두는 것은 일종의 더미인 영가설이 기각되어야 연구자가 관심을 갖는 현상을 나타내고 있는 대안가설을 받아들일 수 있기 때문이다. 만약 영가설이 수용된다면 더미인 영가설을 극복할 수 있는 충분한 경험적 근거가 없는 것임을 나타낸다. 그 결과, 수집된 자료의 관점에서 보면 대안가설은 연구자가 관심을 갖는 현상이 실재와는 거리가 있는 타당하지 않은 것이다. 연구자가 관심을 갖는 현상이 실재를 타당하게 반영하는 것이기 위해서는 영가설이 기각되고 대안가설이 채택되어야 하기 때문에 일반통계검증에서는 영가설 기각에 관심을 두는 것이다.

일반통계검증에서는 영가설의 기각 여부가 관심이지만 구조방정식모형에서는 영가설의 수용 여부가 관심이다. 이 점이 여타의 일반통계검증으로부터 구조방정식모형을 구분짓는 중요한 특징이다.

또한 어떤 유형의 통계검증이든 통계검증에서 반드시 고려해야 할 것이 유의수준 설정이다. 유의수준 설정은 현상을 반영하는 실재가 어느 가설에 나타나는가에 직결되어 있다. 앞서도 진술했듯이 영가설을 기각해야만 실재를 반영하는 대안가설을 채택할 수 있는 경우는 영가설을 기각하는 데 초점을 두게 된다. 이렇게 영가설을 기각하는 것에 초점을 둔다면, 영가설이 맞다고 할 때 우연히 일어나는 경우를 높게 허용할 수 없게 된다. 그 결과 자연스럽게 영가

설이 맞을 때 그것이 우연히 일어나는 경우를 확률로 나타낸 유의수준은 .01이나 .05처럼 작은 값을 갖게 된다.

이와는 달리 구조방정식모형에서의 모형평가처럼 가설검증의 초점이 영가설을 수용하는 데 있다면, 영가설이 맞을 때 그것이 우연히 일어나는 경우를 높게 허용하는 것이 논리적 귀결이다. 그 결과 유의수준은 실험연구처럼 일반통계검증에서 관행적으로 사용되는 .05보다는 높은 수준이어야 한다. 많은 경우 .30이나 .50처럼 상대적으로 큰 값을 갖도록 설정된다. 그러나 전문학술지에 게재된 한 연구를 보면, χ^2검증통계량이 관찰될 확률이 설정된 유의수준보다 낮게 나온 것을 t나 F검증에서처럼 영가설을 기각하고 이를 설정된 모형이 수집된 자료에 적합하다는 증거로 해석하는 잘못을 범하고 있다(신숙재·정문자 1998).

덧붙여 구조방정식모형을 포함한 모든 통계검증에서 유의수준의 결정은 단순히 영가설의 수용여부에 관한 사항으로만 결정되는 것은 아니며 다른 요소도 유의수준의 결정에서 중요하게 고려해야 하는 요소이다. 예를 들면, $(1-\beta)$로 정의되는 통계검증력도 유의수준의 결정에서 함께 고려되어야 하는 중요한 요소이다(Satorra & Saris 1985; Glass & Hopkins 1984).

Ⅴ. 나오며

구조방정식모형에서 수집된 자료에 설정된 모형이 적합한가를 판단하는 모형평가는 매우 중요하다. 자료에 모형이 맞지 않은 상태에서 다른 무엇을 논하는 것은 아무 의미가 없기 때문에 모형평가는 구조방정식모형의 핵심부분이라고 할 수 있다. 이런 모형평가에 관한 판단이 학문적 토대 위에서 논의되지 않고 근거 없는 주장에 의지하거나 잘못된 선행연구를 본보기로 해서 논의되는 것은 부끄러운 일이다. 구조방정식모형에서 모형평가와 관련된 가설검증의 논리와 유의수준 설정이 다른 여타의 일반통계검증과 다른 점을 인지하여 구조방정식모형이 제공하는 여러 가지 강점을 최대한 이용해야 할 것이다.

참고문헌

신숙재·정문자. 1998. “어머니의 양육 스트레스, 사회적 지원과 부모효능감이 양육행동에 미치는 영향.” ≪아동학회지≫ 19(1): 27-42.

이기종. 2012. ≪구조방정식모형≫. 국민대 출판부.

이종성·강계남·김양분·강상진. 2007. ≪사회과학여구를 위한 통계방법≫. 박영사.

정대연. 1997. ≪사회과학방법론사전≫. 백의출판사.

Bollen, K.A. 1989. *Structural Equations with Latent Variables*. New York, NY: John Wiley & Sons.

Browne, M.W. and R. Cudeck. 1993. “Single Sample Cross Validation Indices for Covariance Structures.” *Multivariate Behavioral Research* 24: 445-455.

Glass, G.V. and K.D. Hopkins. 1984. *Statistical Methods in Educational and Psyhology*. Englewood Cliffs. NJ: Prentice Hall.

Hu, Li-tze and P. Bentler. 1999. “Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives.” *Structural Equation Modeling* 6(1):1-55. http://dx.doi.org/10.1080/10705519909540118

Jöreskog, K.G. and D. Sörbom. 2015. *LISREL 9.2: Structural Equation Modeling with the SIMPLIS Command Language*. Chicago, IL: Scientific Software International Inc.

Kline, R.B. 1998. *Principles and Practice of Structural Equation Modeling*. New York, NY: The Guilford Press.

Satorra, A. and W.E. Saris. 1985. “Power of the Likelihood Ratio Test in Covariance Structure Analysis.” P*sychometrika* 50: 83-90.

<접수 2016/1/31, 게재확정 2016/2/12>