

## 경로분석

아동의 학업성취 수준은 어떠한 변수들에 의해서 영향을 받는가? 학업성취에 영향을 미칠 수 있는 변수로는 지능이나 동기, 가족의 사회경제적 지위, 학교 및 교사의 특성 등 여러 가지 변인들을 생각해 볼 수 있다. 앞 장에서 다룬 중다회귀분석은 하나의 종속변수와 다수의 독립변수로 구성된 회귀 모형을 통하여 종속변수와 독립변수간의 관계를 밝히고자 하는 통계적 방법이었다. 따라서 회귀분석에서 관심의 대상이 되는 것은 종속변수에 영향을 미치는 독립변수는 무엇이며, 모형에 투입된 독립변수들을 가지고 종속변수의 변화를 얼마나 설명할 수 있는가의 문제이다. 회귀분석에서는 독립변수들간의 관계는 다루어지지 않는다. 위에서 언급한 연구 문제를 회귀분석이라는 통계기법을 적용하여 살펴본다고 가정해 보자. 종속변수는 아동의 학업성취 수준이 될 것이며, 독립변수는 지능, 동기, 가족의 사회경제적 지위 등이 될 것이다. 회귀분석은 학업성취라는 종속변수의 변화를 지능, 동기, 가족의 사회경제적 지위 등 독립변수들의 선형조합으로 설명하려는 분석기법이다. 따라서 지능이나 동기, 가족의 사회경제적 지위간의 인과 관계는 밝혀지지 않는다.

다시 말하면, 회귀분석에서는 독립변수들이 하나의 그룹으로 다루어져 그들간의 함수적인 관계가 있을 경우에도 그냥 하나의 회귀방정식 속에 묻혀서 종속변수의 값을 예측하는 데에만 공헌할 뿐이다(이순목, 1990). 즉, 변수들간의 보다 자세한 연결관계는 밝히지 못한다는 것이다. 예를 들어 지능은 동기와 어떤 연관관계에 있고, 지능과 가족의 사회경제적 지위는 어떤 연관관계에 있는지에 대해서는 알 수 없는 것이다. 그러나 우리는 이러한 변수들간의 연결 관계를 보다 자세하게 구명해야 할 필요가 있을 수 있다.

경로분석(path analysis) 및 구조 방정식 모형(structural equation models)은 변수들간의 연결 관계(network of relationship) 및 인과관계(casual relationship)가 가정되고 검증됨으로써 변수들이 어떻게 서로 영향을 미치는가에 대하여 설명해 준다. 본 장에서는 경로분석과 구조 방정식 모형의 개념 및 원리를 소개하고자 한다.

### 1. 경로분석

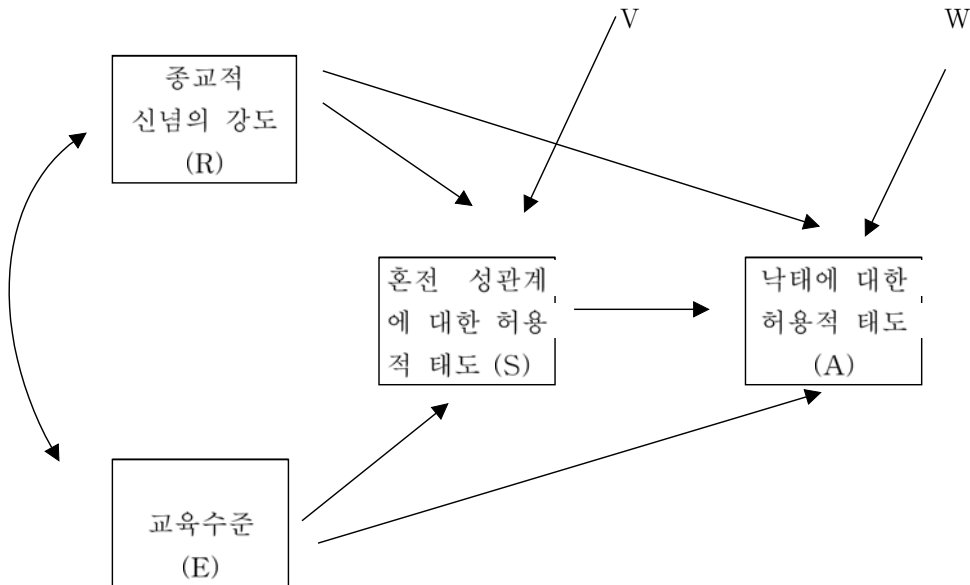
#### 1). 경로분석의 개념 및 기본 가정

경로분석은 기본적으로 일단의 변수들간의 선형관계(linear relationship)를 분해하고 이를 해석하는 방법이라 할 수 있다. 회귀분석에서 하나의 그룹으로 다루어지던 독립변수들 및 종속변수간에 인과적 관계가 설정되고, 설정된 인과적 가설을 검증함으로써 여러 변수들간의 인과 관계를 특정화하여 논리적으로 설명하는 방법이다(김두섭, 1994; Bohnstedt & Knoke, 1994).

경로분석은 이론적 배경에 입각하여 변수들간의 인과 관계를 나타내는 경로모형을 설정한 후,

설정된 경로모형에 따라 회귀분석을 여러번 행함으로써 이루어진다. 다음과 같은 예를 통해 경로 분석에 대해 자세히 살펴보고자 한다. [그림 1-1]은 4변수로 구성된 낙태 태도에 대한 경로모형이다.

[그림 1-1] 낙태 태도에 대한 경로모형



출처 : Bohmstedt & Knoke(1994), pp. 415

위에 설정된 경로모형에 따르면, 종교적 신념의 강도(R)와 교육수준(E)은 이미 주어진 것이며, 혼전 성관계에 대한 허용적 태도(S)는 종교적 신념의 강도와 교육수준에 의해서 영향을 받는 변수이고, 낙태에 대한 허용적 태도(A)는 교육수준과 종교적 신념의 강도에 의해서 직접 영향을 받는 동시에 혼전 성관계에 대한 허용적 태도에 의해서도 영향을 받는 변수이다. 결국, 낙태에 대한 허용적 태도(A)는 종교적 신념의 강도(R) 및 교육(E)에 의해서 영향을 받는 종속변수로 설정되어 있으며, 혼전 성관계에 대한 허용적 태도(S)는 종교적 신념의 강도(R)와 교육(E)의 영향을 받아 이를 낙태에 대한 허용적 태도(A)에 전달하는 매개변수로 설정되어 있다.

이 모형에서 종교적 신념의 강도(R) 및 교육수준(E)은 다른 변수의 영향을 받는 것이 아니라 이미 주어진 것으로 간주되는데, 이러한 변수를 외생변수(exogeneous variables)라 한다. 위의 모형에서 두 외생변수는 서로 상관관계에 있다. 그리고, 낙태에 대한 허용적 태도(A) 및 혼전 성관계에 대한 허용적 태도(S)처럼 모형 안에서 다른 변수들에 의해 설명되어지는 변수들을 내생변수(endogeneous variables)라고 한다. V와 W는 각각 혼전 성관계에 대한 허용적 태도(S)와 낙태에 대한 허용적 태도(A)의 모든 잔여원인변수(residual casual variables)라고 한다.

이상과 같은 경로모형을 정립하기 위해서는 기본적으로 인과적 기본가정(casual assumption)이 전제되어야 한다. 그러한 기본 가정은 다음과 같은 세가지로 설명할 수 있다(Bohmstedt &

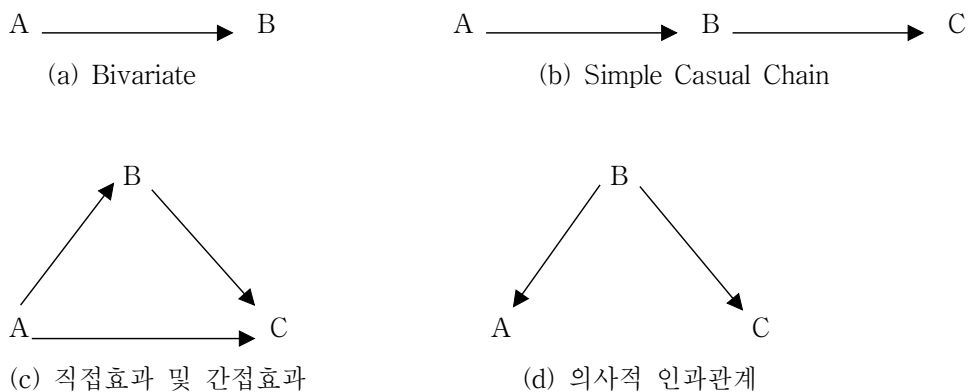
Knoke, 1994).

첫째, 변수들 사이에 인과 관계가 존재하기 위해서는 독립변수와 종속변수 사이에 공변량이 존재해야만 한다. 즉, 한 변수에 체계적인 변화나 차이가 발생하면, 다른 한 변수에도 체계적인 변화나 차이가 동반되어야 한다는 것이다.

둘째, 인과적 흐름의 순서(time order)가 전제되어야 한다. 독립변수로 설정된 변수의 변화가 종속변수로 설정된 변수의 변화보다 선행되어야 한다는 것이다. 가령, 교육수준(E)이 낙태에 대한 허용적 태도(A)에 영향을 미치는 것으로 인과적 흐름의 가정이 설정되었다면, E는 A에 영향을 미칠 수 있지만, 반대로 A가 E에 미치는 영향력의 가능성은 배제된다. 사회과학의 분석 대상이 되는 변수들은 실제의 상황에서는 그 인과적 흐름의 순서가 명백하지 않은 경우가 많다. 닭과 달걀의 관계처럼 원인과 결과를 분간하기 힘든 모호한 관계도 있으며, 인과적 흐름이 일방적이지 아니라 양방적인 경우도 있다. 그러나 경로분석에서는 변수들간의 인과적 흐름의 방향이 일방적인 경우에 국한되며, 변수들간에 인과적인 영향을 서로 주고 받거나 반복되는 경우는 제외된다.

셋째, 인과관계가 비의사적인 관계(nonspuriousness relationship)에 있어야 한다. 이는 변수들간의 공변량이 외부의 영향력으로부터 폐쇄되어 있어야 한다는 점에서 인과적 폐쇄성(casual closure)의 가정이라고도 할 수 있다(김두섭, 1994). 두 변수가 의사관계(spurious relationship)에 있을 때, 두 변수간 공변량은 외부의 영향력에 의한 것이므로 인과적으로 폐쇄되었다고 할 수 없다. 몇몇 인과구조의 형태가 아래의 [그림 1-2]에 제시되어 있다.

[그림 1-2] 인과구조의 형태



출처 : Bohrnstedt & Knoke(1994), pp. 412

위의 그림에서 (a)는 A를 독립변수로, B를 종속변수로 한 단순한 직접 효과의 인과관계를 보여 주고 있다. (b)에서는 A가 B에 영향을 미치고, 다시 B가 C에 영향을 미치고 있는데, 이는 B를 매개로 하여 A가 C에 영향을 미치는 간접효과의 인과관계 구조를 보여주는 것이다. (c)는 간접효과 및 직접효과를 모두 보여주는 인과관계 구조이다. 즉, C는 A로부터 직접적 영향을 받을

뿐만 아니라, B를 매개로 하여 간접적 영향도 받는다. 마지막으로 (d)는 A와 C간의 의사적 인과관계를 보여주고 있다. B와 C 사이에는 아무런 인과관계도 존재하지 않는다. 두 변수간에 관찰되는 공변량은 이들 모두가 외부의 변수 A에 인과적으로 종속되어 있기 때문이다. 따라서 이러한 인과구조에서는 B와 C라는 두 변수의 공변량이 외부의 영향력으로부터 폐쇄된 것으로 볼 수 없다.

경로분석은 위에서 설명한 변수간 공변량, 인과적 흐름의 순서, 비의사적 관계의 기본가정을 만족시키는 경우에만 사용될 수 있다. [그림 1-1]에서 종교적 신념의 강도는 V 및 W와 아무런 연결이 이루어지지 않고 있으며, 교육수준 역시 V 및 W와 연결되어 있지 않다. 그리고 낙태에 대한 허용적 태도는 V와 연결되어 있지 않으며, 혼전 성관계에 대한 허용적 태도는 W와 연결되어 있지 않다. 이는 경로모형에 나타난 변수들의 공변량이 의사관계에 의한 것이 아니며, 경로모형이 폐쇄된 인과구조를 지녀야 한다는 가정을 충족시키기 위한 것이다.

경로분석에서는 이상에서 언급된 인과적 기본가정(casual assumption) 이외에도 회귀분석에 적용되는 가정들을 포함하여 아래의 사항들이 엄격하게 충족되어야 한다(김두섭, 1994; 이순묵, 1990).

첫째, 경로모형의 설정에 오류가 없어야 한다. 이는 매우 기본적인 가정으로서, 설정된 모형에 이론적 결함으로 인하여 필요한 변수가 누락되거나, 반대로 부적절한 변수가 포함되었다면 인과구조를 정확히 파악할 수 없게 된다. 그리고, 경로분석을 위해서는 모형에 포함되지 않은 잔여원인 변수들이 선행변수들과 또는 잔여원인변수들끼리 상호 연관되어 있지 않다는 가정이 충족되어야 한다. 또한 경로모형에서 외생변수는 주어진 것으로 간주되며, 외생변수들간의 상관관계 역시 주어진 것으로 간주되어 분석대상에서 제외된다.

둘째, 경로모형에 포함된 변수들에 측정오차(measurement error)가 없다. 즉, 변수의 측정이 완벽해야 한다는 것이다.

셋째, 경로모형에 포함된 변수간의 구조관계는 선형(linear)이고, 가법적(additive)이어야 한다.

넷째, 경로모형에 포함되는 모든 변수들은 적어도 등간척도(interval scale)의 수준에서 측정되어야 한다. 회귀분석에서는 서열변수나 명목변수도 가변수(dummy variable)라는 편법을 사용하여 모형에 투입할 수 있으나, 경로모형에서는 이러한 방법을 사용할 수 없다.

다섯째, 독립변수들간에 다중공선성이 존재하지 않는다.

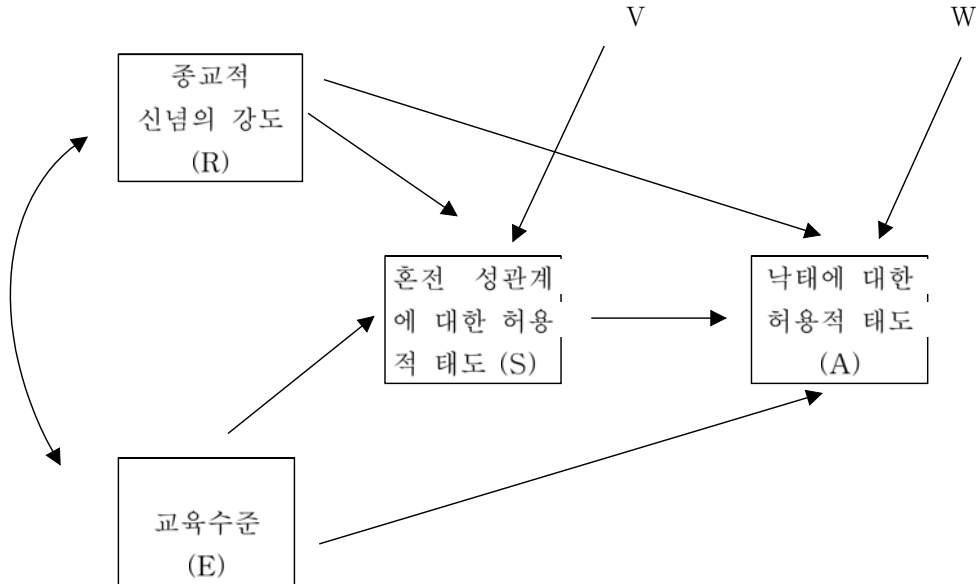
여섯째, 잔차(residual)에 대한 다음의 가정들이 충족된다. 즉, 잔차의 기대값은 0이고, 잔차들은 정규분포를 이루며, 동분산성을 지니고, 잔차들끼리 자기상관이 없어야 한다는 것이다.

## 2) 경로분석의 절차

경로분석은 이론적 배경에 입각하여 변수들간의 인과 관계를 나타내는 경로모형을 설정한 후, 설정된 경로모형에 따라 회귀분석을 여러번 행함으로써 이루어진다. 본 절에서는 [그림 1-1]의

경로모형을 통해 경로분석의 절차 및 방법에 대하여 살펴보고자 한다.

[그림 1-3] 낙태 태도에 대한 경로모형



출처 : Bohrnstedt & Knoke(1994), pp. 415

위와 같이 경로모형에는 각 경로의 화살표 위에 경로계수(path coefficient)를 표시하게 된다. 이 경로계수는 회귀식에서 얻은 독립변수의 비표준화회귀계수(unstandardized regression coefficient) 또는 표준화회귀계수(standardized regression coefficient: Beta)를 사용할 수 있다. 그러나 경로분석을 최초로 제안한 Wright를 비롯하여 대부분의 학자들이 관례적으로 표준화회귀계수를 사용하고 있다. 표준화회귀계수를 경로계수로 사용하면 각 변수들이 다른 후속 변수에 미치는 영향력의 상대적인 크기를 나타내는 데 보다 유리하고, 따라서 분석 결과의 해석이 보다 편리해지는 장점이 있다(김두섭, 1994).

낙태 태도에 대한 위의 경로모형에서 화살표로 표시된 각 경로의 경로계수를 구하기 위해서는 두 번의 중다회귀분석을 실시해야 한다. 첫 번째 중다회귀분석은 혼전 성관계에 대한 허용적 태도(S)를 종속변수로 하고, 종교적 신념의 강도(R) 및 교육수준(E)을 독립변수로 한 회귀모형에 의해 실시한다. 두 번째 중다회귀분석은 낙태에 대한 허용적 태도(A)를 종속변수로 하고, 종교적 신념의 강도(R), 교육수준(E) 및 혼전 성관계에 대한 허용적 태도(S)를 독립변수로 한 회귀모형에 의해 실시한다. 이를 방정식으로 표기하면 다음과 같다.

$$S = p_{sr}R + p_{se}E + p_{sv}V \dots\dots\dots (1)$$

$$A = p_{ar}R + p_{ae}E + p_{as}S + p_{aw}W \dots\dots\dots (2)$$

여기에서,  $p_{sr}$ ,  $p_{se}$ ,  $p_{sv}$ ,  $p_{ar}$ ,  $p_{ae}$ ,  $p_{as}$ ,  $p_{aw}$  등은 각각의 경로계수를 의미한다. 이중  $p_{sr}$ ,  $p_{se}$  는 혼전 성관계에 대한 허용적 태도(S)를 종속변수로 하고, 종교적 신념의 강도(R) 및 교육수준(E)을 독립변수로 한 중다회귀분석의 표준화회귀계수와 같다. 그리고  $p_{ar}$ ,  $p_{ae}$ ,  $p_{as}$  는 낙태에 대한 허용적 태도(A)를 종속변수로 하고, 종교적 신념의 강도(R), 교육수준(E) 및 혼전 성관계에 대한 허용적 태도(S)를 독립변수로 한 중다회귀분석의 표준화회귀계수와 같다. 한편, 잔여원인변수로부터의 경로계수인  $p_{sv}$ ,  $p_{aw}$  는 다음과 같은 방식에 의해 계산된다.

$$p_{sv} = \sqrt{1 - R_{SRE}^2} \dots\dots\dots (3)$$

$$p_{aw} = \sqrt{1 - R_{ARES}^2} \dots\dots\dots (4)$$

즉,  $p_{sv}$ 는 혼전 성관계에 대한 허용적 태도(S)를 종속변수로 해서 실시한 회귀분석에서 구해진 중다결정계수  $R^2$ 의 값을 1에서 뺀 후 제곱근 한 값이며,  $p_{aw}$ 는 낙태에 대한 허용적 태도(A)를 종속변수로 해서 실시한 회귀분석을 통해 구한 중다결정계수  $R^2$ 의 값을 1에서 뺀 후 제곱근 한 값이다.

위의 경로모형에서 설정된 두 차례의 중다회귀분석 결과가 <표 1-1>과 같이 제시되어 있다.

<표 1-1> 혼전 성경험에 대한 태도 및 낙태에 대한 태도의 중다회귀분석 결과

독립 변수	혼전 성경험에 대한 태도(S)	낙태에 대한 태도(A)
종교적 신념의 강도 (R)	-0.293***	-0.084*
교육수준 (E)	0.178***	0.227***
혼전 성경험에 대한 태도 (S)	--	0.338***
중다결정계수	.120***	.219***

이러한 결과로부터 경로모형에 제시된 각 경로의 경로계수를 구할 수 있다.

$$p_{sr} = -.293$$

$$p_{se} = .178$$

$$p_{sv} = \sqrt{1-.120} = .938$$

$$p_{ar} = -.084$$

$$p_{as} = .338$$

$$p_{ae} = .227$$

$$p_{aw} = \sqrt{1-.219} = .884$$

이렇게 산출한 경로계수들은 한 변수가 다른 한 변수에 미치는 직접적인 인과효과만을 측정

한 것이다. 따라서 경로계수만으로는 두 변수간의 인과적 효과를 완전히 파악할 수 없다. 하나의 독립변수가 하나의 종속변수에 미치는 효과를 측정하는 효과계수(effect coefficient)  $C_{ij}$ 는 경로계수와 같거나 보다 큰 값을 지니게 된다. 즉, 혼전 성경험에 대한 허용적 태도(S)와 낙태에 대한 허용적 태도(A)간 경로에 대한 경로계수  $p_{as}$ 는 효과계수와 같다. 그러나 교육수준(E)이 낙태에 대한 허용적 태도(A)에 미치는 효과를 측정하는 효과계수는 교육수준이 낙태에 대한 허용적 태도에 미치는 직접효과 뿐 아니라 혼전 성관계에 대한 허용적 태도를 매개로 미치는 간접효과도 포함되는 것이다. 즉, 직접효과와 간접효과를 합한 것이 효과계수가 되는 것이다. 효과계수는 결국 두 변수간의 인과적 효과라 할 수 있다. 그리고, 두 변수간의 공변량은 이러한 인과적 효과와 비인과적 효과로 구성된다. 혼전 성경험에 대한 허용적 태도(S)가 낙태에 대한 허용적 태도(A)에 미치는 효과 및 교육수준(E)이 낙태에 대한 허용적 태도(A)에 미치는 효과, 종교적 신념의 강도(R)가 낙태에 대한 허용적 태도(A)에 미치는 효과를 나타내는 효과계수는 다음과 같다.

$$C_{as} = P_{as} = .338$$

$$C_{ae} = P_{ae} + (P_{se} \times P_{as}) = .227 + (.178 \times .338) = .287$$

$$C_{ar} = P_{ar} + (P_{sr} \times P_{as}) = -.084 + (-.293 \times .338) = -.183$$

위의 결과에 따르면, 교육수준(E)이 낙태에 대한 허용적 태도(A)에 직접적으로 미치는 효과는 .227이며, 교육수준(E)이 혼전 성관계에 대한 허용적 태도(S)를 매개로 하여 낙태에 대한 허용적 태도(A)에 간접적으로 미치는 효과는 약 .060이다. 따라서 교육수준(E)이 낙태에 대한 허용적 태도(A)에 미치는 효과계수는 .287이 되며, 이 두 변수간 공변량의 나머지 부분은 비인과효과로 더 이상 분석되지 않은 채로 남게 된다.

경로분석에서 변수들이 내생변수에 미치는 영향력을 서로 비교하기 위해서는 효과계수를 이용하면 된다(김두섭, 1994). 위의 수식에서 낙태에 대한 허용적 태도(A)에 대한 혼전 성관계(S), 교육수준(E), 종교적 신념(R)의 효과계수는 각각 .338, .287, -.183이므로, 이 경로모형에서 낙태에 대한 허용적 태도에 가장 큰 영향력을 행사하는 변수는 혼전 성관계에 대한 허용적 태도라 결론 지을 수 있다.

### 3) 경로분석의 한계

경로분석은 전제로 해야 할 기본 가정이 매우 까다롭고, 비현실적이다. 예를 들면, 경로분석에서는 인과구조의 흐름이 일방적인 것으로 가정하고 있다. 그러나 현실적으로 사회과학에서 관심의 대상이 되는 변수들은 인과적인 영향을 서로 주고 받거나 반복되는 경우가 있다. 또한 경로분석에서는 잔여원인변수들끼리 서로 상호 연관되어 있지 않을 것으로 가정하고 있는데, 이 또한 비현실적인 가정이라 할 수 있다.

이 밖에 경로분석은 측정변수들만을 다룬다는 점에서 문제가 있다. 측정변수만으로 모형을 세울 경우 어떤 이론적 개념의 정의와 측정이 동일시되는 문제가 있다. 예컨대, 우리는 “지능”이란 이론적 개념을 여러 개의 측정 가능한 변수들(수학능력, 국어 능력, 언어구사능력 등)을 측정해서 그 공통된 부분을 “지능”으로 조작적으로 정의한다. 여기서 공통부분을 뽑아 정의되는 이론적 개념인 “지능”은 이론변수이고, 이를 간접적으로 측정하는 데 사용된 측정 가능했던 변수들은 측정변수라고 한다(이순목, 1990).

이론변수는 잠재변수(latent variable), 요인(factor), 또는 가설적 구인(hypothetical construct) 등으로도 불리우며, 지능이나 태도와 같은 직접적으로 관찰이 불가능한 일차원적 개념이다. 직접적으로 관찰이 불가능하므로, 지능이나 태도와 같은 이론변수는 이를 나타내 주는 측정 가능한 다른 변수들을 통해 정의된다. 이렇게 이론변수를 나타내 주는 관찰 가능한 변수를 측정변수(measured variable), 또는 관찰변수(observed variable)라고 한다. 즉, 이론변수가 겉으로 드러나 관찰된 또는 측정된 상태를 말한다. 사회경제적 지위라는 이론변수가 소득이나 학력과 같은 변수에 의해서 측정되면, 소득이나 학력을 측정변수 또는 관찰변수라고 하는 것이다. 따라서 관찰 가능한 변수들의 공통부분을 뽑아 이론변수로 정의하는 경우, 각 측정변수들에서 이론변수를 정의하는 데 쓰이는 공통부분을 제외한 나머지 부분이 생기게 되는데, 이를 측정의 오차(measurement error)라고 한다. 이러한 측정의 오차는 측정변수를 관찰할 때 사용되는 측정도구의 미비함과 측정 당시의 통제할 수 없는 여러 요인들에서 비롯되는 것이므로, 타당도와 신뢰도가 높은 측정도구를 사용함으로써 오차의 크기를 줄일 수 있다(성태제, 1998).

그런데 경로분석은 이러한 이론변수와 측정변수가 동일시되므로, 보다 순수한 이론변수가 추출되지 않은 상태에서 모형이 정립되고, 가설이 검증된다.

경로분석의 이러한 한계점을 극복함으로써 사회과학에서 보다 정교한 인과구조를 밝히는데 공헌한 방법이 바로 구조방정식모형(structural equation models)이다. 구조방정식모형은 측정변수뿐만 아니라 이론변수까지 포함하므로 더 복잡하고 다양한 설명기능을 제공한다. 구조방정식모형에서는 경로분석에서 기본적으로 전제하였던 인과적 흐름의 순서를 고려할 필요가 없다. 즉, 구조방정식모형에서는 변수간 인과관계에 대한 일방성의 가정을 버리고, 이론변수간의 관계가 상호적으로 표시되는 것을 허용한다. 또한 구조방정식모형에서는 잔여원인변수의 무상관에 대한 가정 역시 필요없게 된다. 구조방정식모형은 비현실적인 경로분석의 가정을 버리고, 측정변수와 이론변수를 모두 모형에 포함함으로써 보다 현실적인 이론구조의 표시를 가능하게 해 준다(이순목, 1990).