

2014년 한국아동·청소년패널조사 V 제1차 콜로키움 자료집

Mplus를 활용한 자료분석

- ▶ 일시 _ 2014년 7월 4일(금) 14:00~17:00
- ▶ 장소 _ 한국청소년정책연구원 10층 세미나실
- ▶ 주최 _ 한국청소년정책연구원

발 표



Mplus를 활용한 자료분석

조영일 교수

(성신여자대학교 심리학과)

M-plus를 활용한 자료분석

일시: 2014년 7월 4일
장소: 한국청소년정책연구원



성신여자대학교

성신여자대학교
심리학과
조영일, Ph.D.

1

Table of Contents

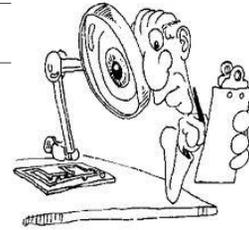
- I. M-plus의 소개
- II. 경로분석
- III. 확인적 요인분석(측정모형)
- IV. 구조방정식 모형
- V. 종단자료의 분석

M-plus의 소개

1. M-plus란?

2. M-plus의 기초 및 활용

3. M-plus의 예시



1. M-plus란?

기본정보

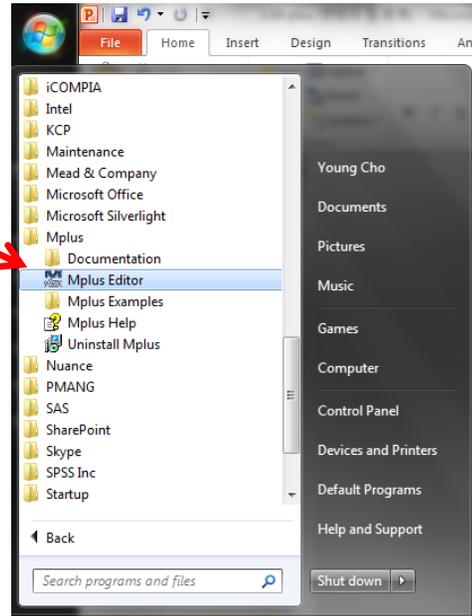
- "M-plus는 구조방정식 모형과 종단자료 분석(잠재성장모형)의 분석에 사용되기 위해서 고안된 프로그램임.
- 그러나 M-plus는 위에서 기술한 것 이외의 현재 심리학의 자료 분석에서 사용되는 거의 모든 통계 모형을 검증할 수 있다.
- 변수중심 분석모형 뿐만 아니라 사람중심 분석모형의 분석이 가능함.
- 명령어의 체계가 영문법에 기반을 두고 단순하여 배우기가 쉽다.
- input 파일은 .inp의 확장자를 output 파일은 .out의 확장자 명을 가진다.
- M-plus 홈페이지인 www.statmodel.com에서 Q&A를 통해서 질문에 답해준다.
- 특히, License를 구입한 경우(매년 일정 유지비용을 내는 경우)에는 문제가 발생시 이메일로 자료와 프로그램을 보내면 문제를 직접 해결해 주기도 한다.

1. M-plus란?

기본정보

- M-plus를 실행시키는 2가지 방법

1. Window에서 Mplus Editor를 선택한다.



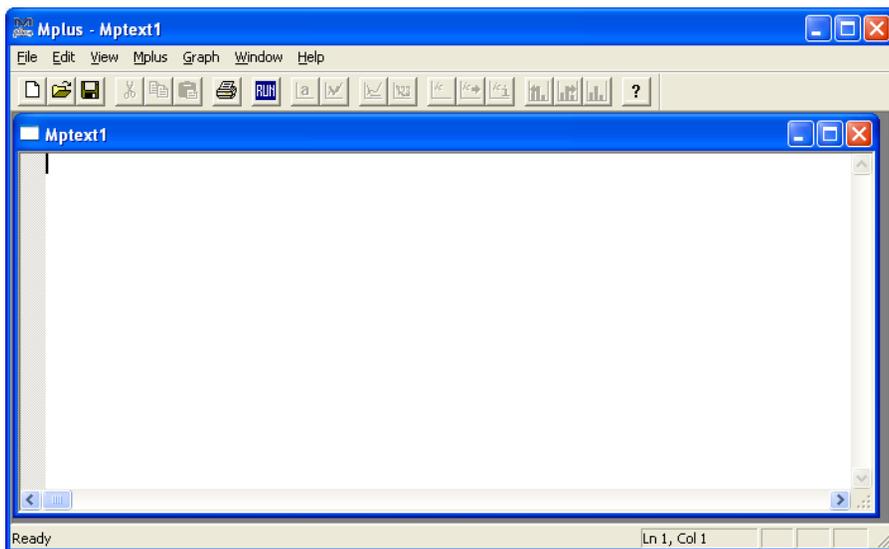
2. 바탕화면의 아래 아이콘을 클릭한다.



1. M-plus란?

M-plus 시작화면

- M-plus를 열면 다음과 같은 화면이 뜬다.

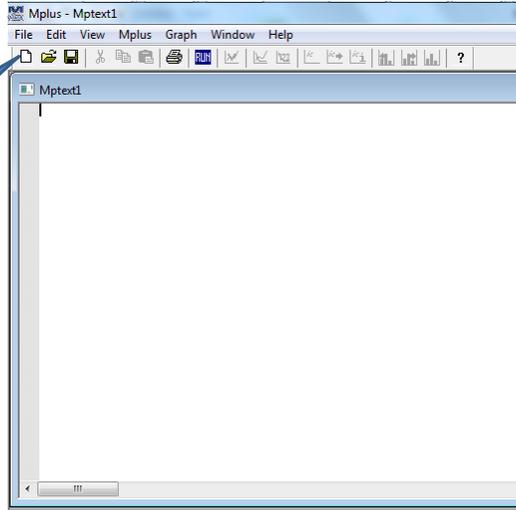


1. M-plus란?

M-plus 시작화면

- M-plus의 화면은 크기 2개로 구성됨
 - Input (프로그램의 편집)
 - Output(결과)

만약에 input화면이 나타나지 않으면, 이것을 클릭하면 새로운 Input 화면(Mptest1)이 뜬다.

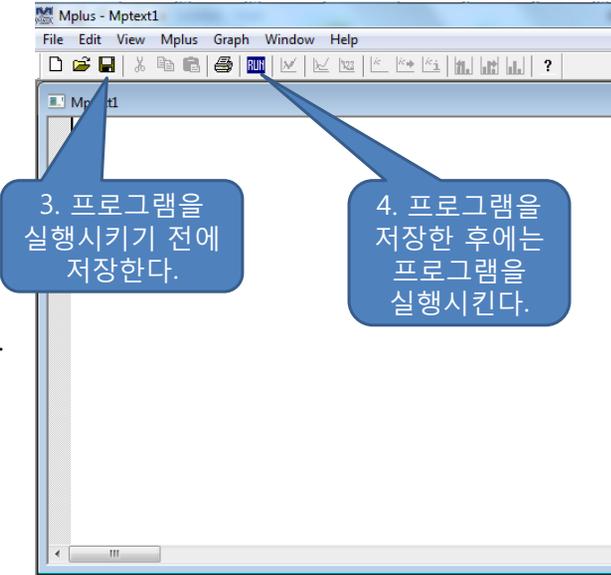


1. M-plus란?

M-plus의 실행

Input 창에서 다음과 같은 순서로 다양한 모형들을 검증할 수 있다.

1. Input 창을 연다.
2. Mplus의 형식에 맞게 검증하고자 하는 모형을 명령어를 사용하여 구현한다.
3. 프로그램을 저장한다.
4. 프로그램을 실행한다.
5. 결과창을 통해서 모형의 모수를 추정하고 모형을 검증한다.



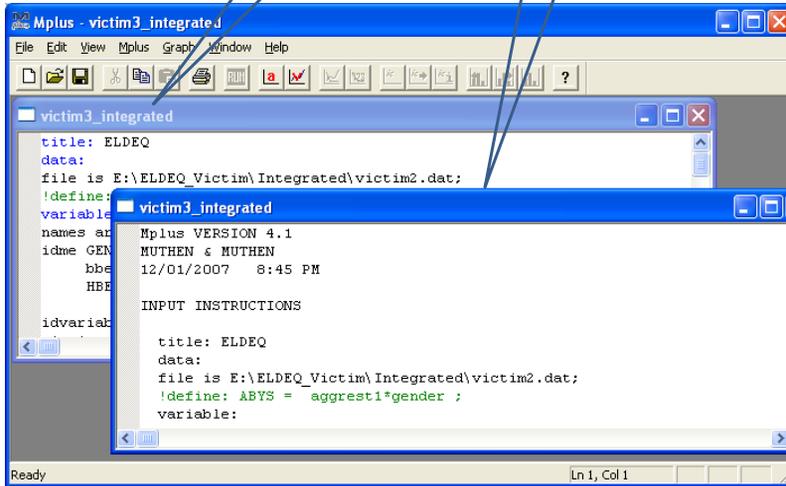
1. M-plus란?

M-plus의 실행

Input 창과 output 창.

Input 창

Output 창



1. M-plus란?

기본 명령어 체계

- M-plus는 AMOS처럼 path diagram을 직접 그리고 보다는 연구자가 명령어를 직접 input 창에 써야한다.
- M-plus는 모형을 명령어를 통해서 직접 구현해야 한다.
- 일반적으로, M-plus를 돌리기 전에 연구자가 해당하는 모형(구조방정식 모형 혹은 확인적 요인분석 모형)을 종이에 직접 그린 후에, 그 그림에 따라서 명령어를 작성한다.

1. M-plus란?

기본 명령어 체계

```

Mplus - KIM1
File Edit View Mplus Graph Window Help
KIM1
Title: Mediation Analysis
with Latent Growth Curve Model;
DATA:
FILE = kim1.dat;
VARIABLE:
NAMES = ID HA1 HA2 HA3 HA4
VARI68 VARI69 VARI70 VARI44 VARI45 DEP;
USEVARIABLES = HA1 HA2 HA3 HA4
VARI68 VARI69 VARI70 VARI44 VARI45 DEP;
MISSING = .;
DEFINE:
ANALYSIS:
BOOTSTRAP = 10000;
MODEL:
RISK BY HA1
HA2
HA3
HA4;
NON BY VARI68
VARI69
VARI70;
EMOTION BY VARI44
VARI45;
VARI44@0;
emotion on risk non;
DEP on emotion risk non;
model indirect:
dep ind risk;
dep ind non;
OUTPUT:
CINTERVAL(BCBOOTSTRAP);

```

- 1. 자료를 읽어들이고, 변수를 지정함
- 2. 모수를 추정할 방법을 지정해줌.
- 3. 검증하고자 하는 모형을 명령어를 사용하여 기술함.
- 4. 결과창에 대한 옵션을 지정해줌.

• 특히, M-plus는 input은 기본적으로 4개의 기본 section으로 나눌 수 있음.

1. 자료입력
2. 분석방법
3. 분석모형
4. 결과옵션

1. M-plus란?

기본 명령어 체계

• Different commands divided into a series of sections

- TITLE
- DATA (required)
- VARIABLE (required)
- DEFINE
- ANALYSIS
- MODEL
- OUTPUT
- SAVEDATA
- MONTECARLO

➤ ! 다음에 쓴 것은 프로그램 내용에 대한 주석 혹은 설명으로 들어감.

➤ 모든 명령어의 끝에는 ;(semi-colon)를 써준다.

1. M-plus란?

기본 명령어 체계

- **TITLE:**

- Everything after "*Title:*" is the title and the title ends when "*Data:*" appears

- 이 명령어를 사용해서 프로그램에 제목을 달거나, 해당 프로그램에서 분석하고자 하는 모형이 어떤 것인지에 대한 설명을 첨부할 수 있다.
- 한줄에 첫칸부터 80자를 넘지 말아야 한다.

1. M-plus란?

기본 명령어 체계

- **DATA:**

- Tells Mplus where to find the file containing the data.

- "E:\Growth_Curves\ClassData.dat"
 - Without a specific path, Mplus will look in the **same folder where the Mplus code is saved.**

DATA:

FILE = kim1.dat;

TYPE = INDIVIDUAL; ! DEFAULT

!TYPE = COVARIANCE;

- 일반적으로, Mplus에서는 스페이스(space)나 탭(tab), 쉼표(comma)로 분리된 원자료를 가지고 있는 text 파일을 불러들여 사용한다.
- 변수의 수는 500개로 제한된다.
- 공분산 행렬이나 상관행렬도 자료로 입력할 수 있다.
- 자료는 일반적으로 M-plus의 input 파일과 같은 폴더에 있는 것이 좋다.

1. M-plus란?

SPSS에서 M-plus 데이터 파일 만들기

(결측치를 특정한 수로 변환시키기)

	id	sexw2	scharew2	q33a01w2	q33a02w2	q33a03w2	q33a04w2	q33a05w2	q33a06w2	q35a4v
1	1	1	421	4	4	3	1	1	1	1
2	2	1	421	2	5	3	5	4	5	5
3	3	2	421	4	4	2	3	1	2	2
4	4	1	421	4	4	3	4	4	3	3
5	5	2	421	3	2	2	2	3	3	3
6	6	1	421	2	4	2	2	2	2	2
7	7
8	8	2	421	2	2	2	2	2	4	4
9	9	2	421	4	3	3	4	4	3	3
10	10	1	421	3	4	3	4	4	4	4
11	11	2	421	3	3	2	2	3	3	3
12	12
13	13
14	14	1	421	3	3	3	3	3	3	3
15	15	2	421	4	3	2	3	2	3	3
16	16	2	421	3	5	3	3	3	5	5
17	17	1	421	3	4	4	4	3	3	3

1. M-plus란?

SPSS에서 M-plus 데이터 파일 만들기

(결측치를 특정한 수로 변환시키기)

변환(T) 메뉴:

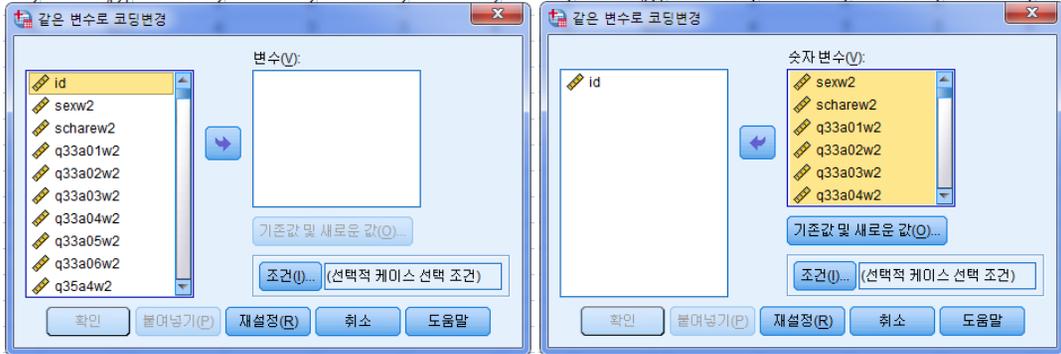
- 변수 계산(C)...
- 케이스 내의 값 반도(O)...
- 값 이동(E)...
- 같은 변수로 코딩변경(S)...**
- 다른 변수로 코딩변경(R)...
- 자동 코딩변경(A)...
- 비주열 빈 만들기(M)...
- 최적의 빈 만들기(O)...
- 모형화를 위한 데이터 준비(P) ▶
- 순위변수 생성(K)...
- 날짜 및 시간 마법사(D)...
- 시계열변수 생성(M)...
- 결측값 대체(V)...
- 난수 생성기(G)...
- 변환 중지(T) Ctrl+G

변환=>같은 변수로 코딩변경(S)를 선택.

1. M-plus란?

SPSS에서 M-plus 데이터 파일 만들기

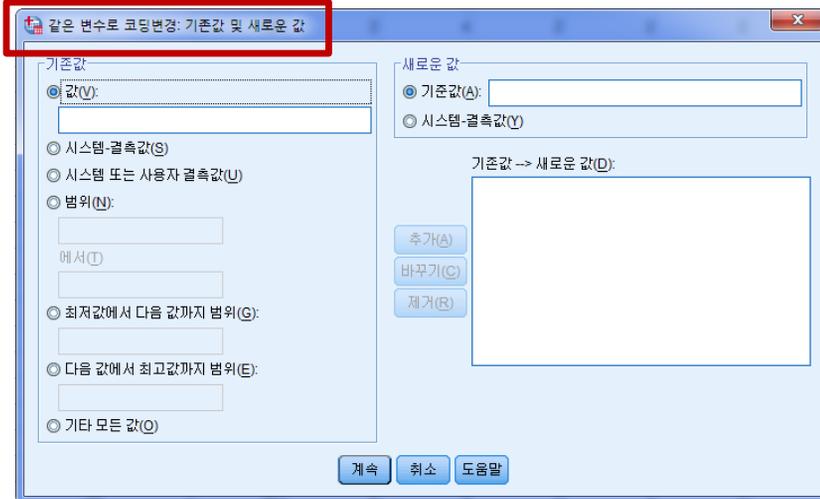
(결측치를 특정한 수로 변환시키기)



1. M-plus란?

SPSS에서 M-plus 데이터 파일 만들기

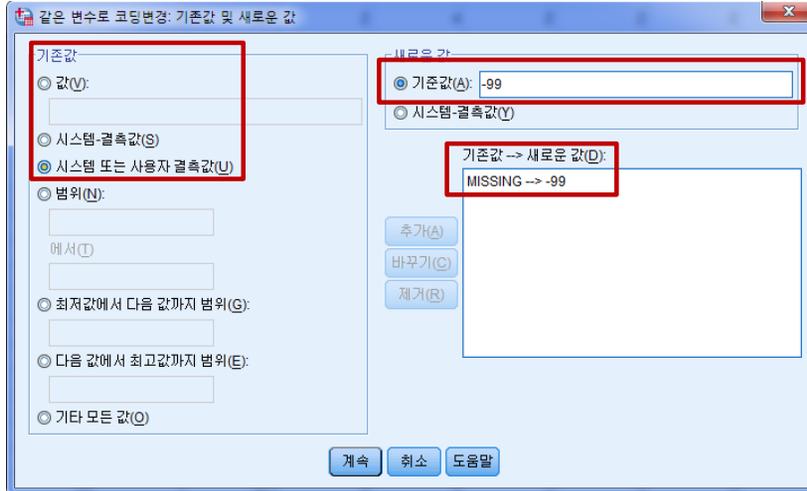
(결측치를 특정한 수로 변환시키기)



1. M-plus란?

SPSS에서 M-plus 데이터 파일 만들기

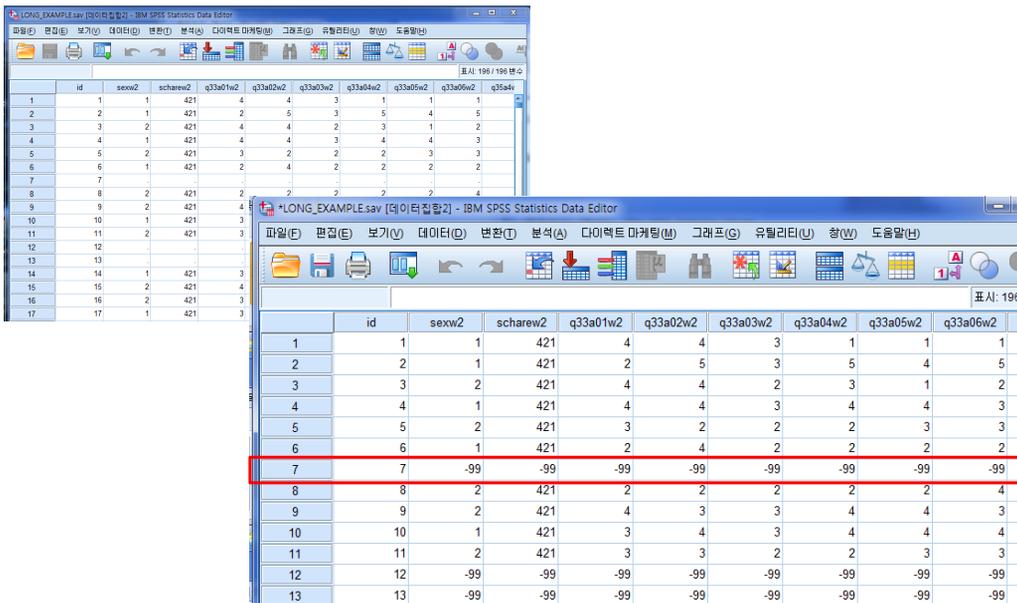
(결측치를 특정한 수로 변환시키기)



1. M-plus란?

SPSS에서 M-plus 데이터 파일 만들기

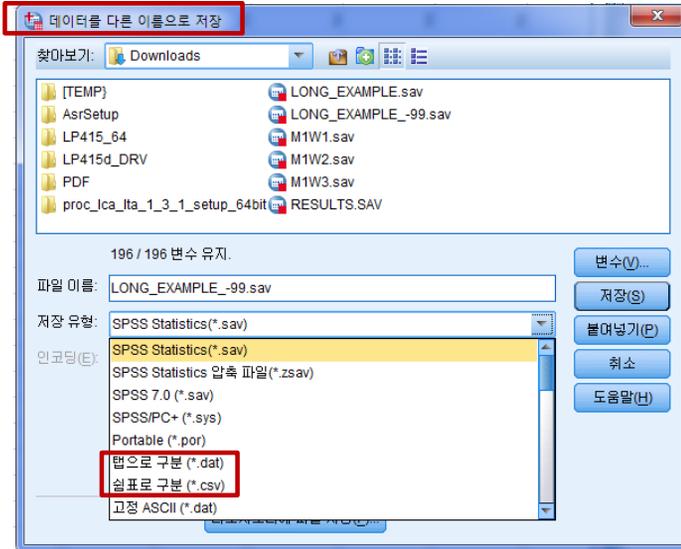
(결측치를 특정한 수로 변환시키기)



1. M-plus란?

SPSS에서 M-plus 데이터 파일 만들기

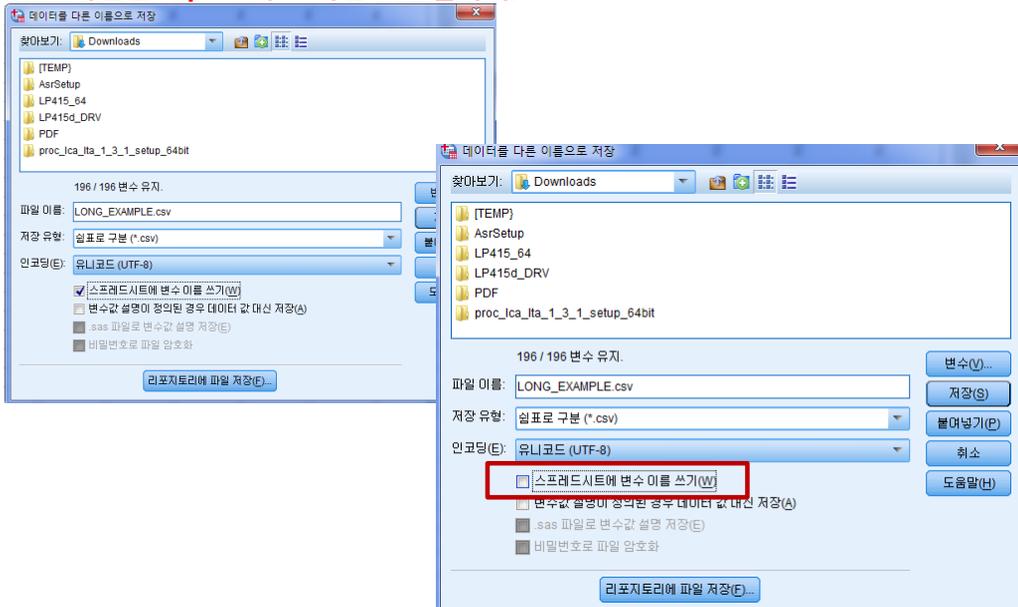
(SPSS파일을 Mplus 자료 파일로 변환하기)



1. M-plus란?

SPSS에서 M-plus 데이터 파일 만들기

(SPSS파일을 Mplus 자료 파일로 변환하기)



1. M-plus란?

SPSS에서 M-plus 데이터 파일 만들기

(SPSS파일을 Mplus 자료 파일로 변환하기)



1. M-plus란?

기본 명령어 체계

원자료를 포함한 text파일의 일반적인 형태

1) Space로 변수들이 구분됨

2) Tab으로 변수들이 구분됨

File	Edit	Format	View	Help						
2	6.0	5.0	4.0	3	20.00	20.00	18.00	3.1500	4.38	15
3	6.0	6.0	9.0	4	34.00	29.00	29.00	2.9200	3.25	22
4	7.0	7.0	14.0	2	22.00	21.00	23.00	2.8500	3.13	10
5	8.0	5.0	9.0	1	22.00	24.00	24.00	3.5400	4.38	10
6	6.0	7.0	10.0	5	21.00	20.00	20.00	2.9200	2.75	18
7	5.0	3.0	0.0	2	20.00	24.00	27.00	1.2300	1.38	16
8	9.0	6.0	10.0	4	18.00	31.00	30.00	3.6900	3.00	10
9	8.0	7.0	10.0	2	19.00	17.00	21.00	3.9200	2.88	43
10	6.0	4.0	13.0	1	38.00	45.00	36.00	2.7700	3.13	15
11	3.0	2.0	8.0	2	27.00	22.00	28.00	2.6200	3.75	5
12	4.0	5.0	5.0	0	27.00	31.00	27.00	3.3800	3.38	23
13	5.0	5.0	8.0	3	20.00	17.00	23.00	3.0000	3.50	17
14	8.0	3.0	3.0	2	24.00	21.00	25.00	2.3800	3.50	12
15	9.0	8.0	15.0	5	19.00	25.00	24.00	4.3100	3.88	32
16	3.0	3.0	1.0	2	22.00	23.00	23.00	2.6900	4.00	12
17	8.0	9.0	8.0	6	26.00	18.00	28.00	3.6900	3.25	27
18	7.0	4.0	10.0	2	19.00	25.00	24.00	3.6200	2.63	13
19	7.0	4.0	7.0	2	29.00	32.00	26.00	2.9200	4.00	5
20	6.0	7.0	10.0	4	27.00	28.00	31.00	2.5400	3.38	10
21	6.0	6.0	10.0	2	21.00	23.00	18.00	3.0000	3.25	12
22	9.0	8.0	7.0	5	19.00	26.00	17.00	3.8500	4.25	50

File	Edit	Format	View	Help						
2	6.0	5.0	4.0	3	20.00	20.00	18.00	3.1500	4.38	15
3	6.0	6.0	9.0	4	34.00	29.00	29.00	2.9200	3.25	22
4	7.0	7.0	14.0	2	22.00	21.00	23.00	2.8500	3.13	10
5	8.0	5.0	9.0	1	22.00	24.00	24.00	3.5400	4.38	10
6	6.0	7.0	10.0	5	21.00	20.00	20.00	2.9200	2.75	18
7	5.0	3.0	0.0	2	20.00	24.00	27.00	1.2300	1.38	16
8	9.0	6.0	10.0	4	18.00	31.00	30.00	3.6900	3.00	10
9	8.0	7.0	10.0	2	19.00	17.00	21.00	3.9200	2.88	43
10	6.0	4.0	13.0	1	38.00	45.00	36.00	2.7700	3.13	15
11	3.0	2.0	8.0	2	27.00	22.00	28.00	2.6200	3.75	5
12	4.0	5.0	5.0	0	27.00	31.00	27.00	3.3800	3.38	23
13	5.0	5.0	8.0	3	20.00	17.00	23.00	3.0000	3.50	17
14	8.0	3.0	3.0	2	24.00	21.00	25.00	2.3800	3.50	12
15	9.0	8.0	15.0	5	19.00	25.00	24.00	4.3100	3.88	32
16	3.0	3.0	1.0	2	22.00	23.00	23.00	2.6900	4.00	12
17	8.0	9.0	8.0	6	26.00	18.00	28.00	3.6900	3.25	27
18	7.0	4.0	10.0	2	19.00	25.00	24.00	3.6200	2.63	13
19	7.0	4.0	7.0	2	29.00	32.00	26.00	2.9200	4.00	5

1. M-plus란?

기본 명령어 체계

- **VARIABLE:**

- Series of subcommands that tell Mplus . . .

- **Names are** names of variables (8 characters max; case sensitive in certain versions)
- **Missing are all** (-99) ; tells Mplus user defined missing values
- **Use variables are** names variables to use in the analysis. Useful if have larger data file for multiple purposes/analysis. **IMPORTANT**

- **VARIABLE:**

```
NAMES = ID HA1 HA2 HA3 HA4  
VAR168 VAR169 VAR170 VAR144 VAR145 DEP;  
USEVARIABLES = HA1 HA2 HA3 HA4  
VAR168 VAR169 VAR170 VAR144 VAR145 DEP;  
MISSING = ;
```

- 읽어들이는 자료에 속한 변수들에 이름을 부여한다.
- 명명된 자료들 중에서 분석에 사용될 변수를 지정해준다.
- 특히, 어떠한 값 혹은 부호가 결측치를 나타내는지를 지정해준다.
- '='이나 'ARE'를 사용한다.

1. M-plus란?

기본 명령어 체계

- **VARIABLE:**

```
NAMES ARE gender age ordv1 ordv2 ordv3 ordv4 ordv5;  
USEVARIABLES ARE gender age ordv1 ordv2 ordv3 ordv4 ordv5;  
MISSING ARE ALL (.);  
CATEGORICAL ARE ordv1 ordv2 ordv3 ordv4 ordv5;  
NOMINAL ARE gender;
```

1. 범주 변인들은 ordv1 ordv2 ordv3 ordv4 ordv5.
2. 명명 변인은 gender.

- 이외에도 변수의 TYPE을 지정해 줄 수 있다.

→ CATEGORICAL, NOMINAL, & CENSORED

- 이외에 다양한 옵션이 있다.

→ IDVARIABLE, GROUPING, CLUSTERING, 등

1. M-plus란?

기본 명령어 체계

- **ANALYSIS:**

- Tells Mplus what **type of analysis and estimator** will be used
 - **Type = basic**; (default)
- 분석모형을 지정해주고,
- 모수의 추정방식을 정해준다.

ANALYSIS:

ESTIMATOR = ML;
BOOTSTRAP = 10000;

- **TYPE**를 이용하여 분석하고자 하는 통계 모형을 지정해줄수 있다. 지정해주지 않으면, 구조방정식 모형을 검증한다고 가정한다.
→ BASIC, EFA, MIXTURE, TWOLEVEL 등
- **ESTIMATOR**를 이용하여 모수 추정방식을 변경할 수 있음. 지정해주지 않으면, 최대우도법(Maximum Likelihood Estimation Method)를 사용함을 가정한다.
→ WLSMV(측정변수가 서열 혹은 이분변수일 경우), WLS 등

1. M-plus란?

기본 명령어 체계

- **MODEL:**

- This contains the basic model statements
 - Y **ON** X ; ! regression
 - F1 **BY** var1@1 var2 var3 var4 ; ! Latent factors
 - var1 **WITH** var2 ; !correlation
- 이 부분에서, **검증하고자 하는 모형을** 기술한다.
- 기본적으로 M-plus는 변수들간의 관계에 대해서 기술한다.
변수 [관계의 성격] 변수 (이름);
- 관계의 성격:
 - 잠재변수 와 잠재변수 간의 관계
 - 관찰변수 와 잠재변수 간의 관계
 - 관찰변수 와 관찰변수 간의 관계
- 이름은 선택사항이다.

1. M-plus란?

기본 명령어 체계

- 특히 **사용에 주의해야 할 2개의 명령어**

- @ : 모수 추정치의 값을 특정한 값으로 고정해 준다.

예: F1 by X1@1

X2

X3;

- * : 일정한 값으로 고정된 추정치를 자유롭게 추정하게 해준다. *뒤에 오는 숫자는 **시작값(starting value)**로 사용된다.

예: F1 by X1*.5

X2@1

X3;

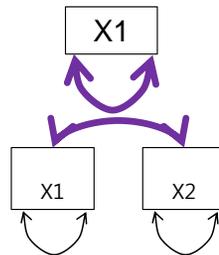
1. M-plus란?

기본 명령어 체계

- 분산(variance)

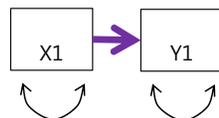
X1;

Path Diagram



- 공분산(covariance)

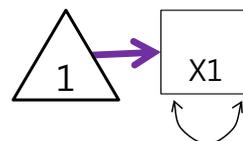
X1 with X2;



- 회귀계수(regression coefficient)
- 독립변수(예측변수)는 왼편에
- 종속변수(준거변수)는 오른편에

[X1];

- 평균(mean)
- M-plus에서는 측정변수의 평균을 자동으로 추정



1. M-plus란?

기본 명령어 체계

- **OUTPUT:**

- Lists specific statistical and graphical output wanted
 - Will get to this in the next section

OUTPUT:

CINTERVAL(BCBOOTSTRAP);

- 결과에서 보고되는 것들을 지정한다.
- 특히, 수정지수(Modification index)와 Bootstrapping의 신뢰구간 (Confidence interval)을 추가로 지정할 수 있다.
- 일반적으로, 아래와 같은 정보들이 출력된다.(default)
 - 자료에 대한 정보(변수의 개수, 사례 수, 평균, 공분산 등)
 - 공분산을 계산하는데 사용된 정보의 정도 (Covariance Coverage)
 - 모형 적합도 지수
 - 모수 추정치, 표준오차, 그리고 p-값 등.

1. M-plus란?

기본 명령어 체계

- **OUTPUT:**

- Lists specific statistical and graphical output wanted
 - Will get to this in the next section

OUTPUT:

CINTERVAL(BCBOOTSTRAP);

- **활용 가능한 옵션:**
 - **SAMPSTAT** : 기술통계치를 출력하십시오.
 - **STANDARDIZED** : 표준화 계수를 출력하십시오.
 - **RESIDUAL** : 잔차 행렬을 출력하십시오.
 - **CINTERVAL(BCBOOTSTRAP)** : Bootstrapping에 의한 신뢰구간을 출력하십시오.
 - **TECH1** : 추정되는 모수들을 표시해준다.
 - **MODINDICES(숫자)**: 해당하는 숫자 이상의 수정지수를 가지는 값을 출력하십시오.

1. M-plus란?

기본 명령어 체계

- **SAVEDATA:**

- 분석 결과를 추후의 분석을 위해서 저장한다.

SAVEDATA:

```
SAMPLE IS SAMPLEDATA.DAT;  
FILE IS FILENAME.DAT;  
SAVE = FSCORES;  
      CPROBABILITIES;
```

1. M-plus란?

기본 명령어 체계

- **MONTECARLO:**

- Simulation 연구에서 활용될 수 있음.

MONTECARLO:

```
NAMES ARE y1-y3;  
NOBSERVATIONS = 100;  
NREPS = 100;  
SEED = 4533;  
repsave=all;  
save=con(3)*.dat;  
results=results.sav;
```

MODEL POPULATION:

```
i s | y1@0 y2@1 y3@2;  
[i*1 s*.5];      i*1; s*.2;  
i WITH s*.1;    y1-y3*.5;
```

2. M-plus의 기초 및 활용

모수추정 형태 지정 및 모수값 검증

사용되는 경우

1. M-plus에서 사용되는 default를 변경하고자 하는 경우
2. 전반적인 모형을 평가하는 것 이외에, 가정과 관련된 경로(예: 요인계수, 경로계수)의 값에 대한 통계적 추론을 하고자 하는 경우.

- Constraining a parameter to a given value
- Freeing a parameter
- Constraining parameters to equality
- Assigning names to parameters
- Testing hypotheses about parameter estimates using Wald tests

2. M-plus의 기초 및 활용

모수값 고정하기

사용법(in M-plus)

- @ 를 사용함.
- 변수@숫자
 - 고정하고자 하는 값과 관련된 변수(variable)을 선택함
 - 숫자(고정하고자 하는 값)을 @ 다음에 써준다.
 - 일반적으로 '0' 또는 '1'로 고정시킨다.

2. M-plus의 기초 및 활용

고정모수를 자유모수로 만들기 사용법(in M-plus)

- * 를 사용함.
- 변수*
 - 자유모수로 바꾸고자 하는 변수(variable)를 선택함
 - 변수 다음에 *를 붙여준다.
 - 일반적으로, M-plus의 default 방식을 변경시키기 위해서 사용된다.
 - *다음에 숫자가 주어지면, 숫자는 시작값(starting value)으로 사용된다.

2. M-plus의 기초 및 활용

모수에 동일화 제약가하기 사용법(in M-plus)

- **괄호()에 동일한 숫자나 영문이름을 넣어 고정시킴.**
- 변수들이 동일한 숫자 혹은 영문이름을 가지고 동일함을 가정함.
- 변수 (숫자 혹은 영문이름)
 - 숫자를 포함하는 괄호는 명령어 **해당 줄의 마지막에** 위치하여야 함.
 - 만약 2개의 변수가 하나의 괄호를 하나의 명령어 줄에서 고정하면 두 변수의 값들에는 동일화 제약이 가해짐.
 - 일반적으로 한번에 한 줄에 하나의 변수만을 고정해 주는 것이 좋음.

2. M-plus의 기초 및 활용

변수에 이름 부여하기

사용법(in M-plus)

- 괄호()에 동일한 영문이름을 넣어 고정시킴.
- 변수 (영문이름)
 - **Model Test:** 또는 **Model Constraint:** 에서 변수들을 사용하기 위해서 변수명을 지정함.

2. M-plus의 기초 및 활용

Wald 검증을 활용한 모수에 대한 가설 검증

사용법(in M-plus)

- **MODEL CONSTRAINT:** 를 사용하여 새로운 변수(가설에 기초한 제약을 가진 변수)를 생성하고 변수들 간의 관련성에 대한 검증을 할 수 있음.

- 예:

```
MODEL CONSTRAINT:  
NEW (A1B1 A2B2 A3B3);
```

```
A1B1 = A1*B1;
```

```
A2B2 = A2*B2;
```

```
A3B3 = A3*B3;
```

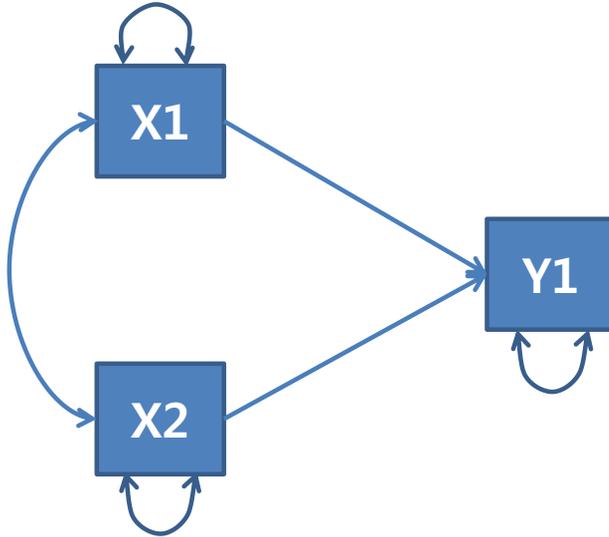
```
MODEL TEST:
```

```
A1B1 = A2B2;
```

- M-plus의 결과에서 합치도 지수를 보고하는 부분에서 위 가설검증에 대한 Wald test값을 보고함.

3. M-plus의 예시

통계 모형의 예 (경로분석)

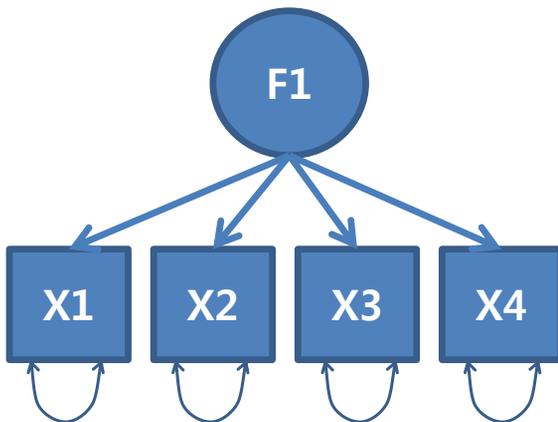


MODEL:

- ! Regression
Y1 ON X1 X2;
- ! Covariance
X1 WITH X2;
- ! Variance
Y1; X1; X2;
- ! Mean
[Y1 X1 X2];

3. M-plus의 예시

통계 모형의 예 (확인적 요인분석)

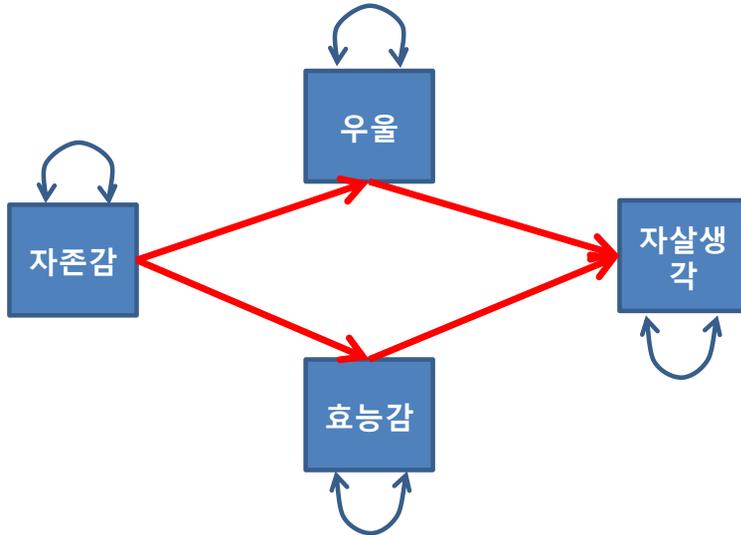


MODEL:

- F1 BY X1@1
- X2
- X3
- X4;

3. M-plus의 예시

통계 모형의 예 (중다매개모형)



3. M-plus의 예시

통계 모형의 예 (중다매개모형)

M-plus 파일

TITLE: MULTIPLE REGRESSION

In this file, you can find the variable names for the data file.

DATA: FILE IS PRACTICE.CSV;

VARIABLE: NAMES ARE

GENDER
 SE1 SE2 SE3
 EF1 EF2 EF3
 SS1 SS2 SS3
 DEP1 DEP2 DEP3
 SUICIDE
 SAT_LIFE STR_LIFE
 SE EF SS DEP;
 USEVARIABLES =
 SUICIDE
 SE DEP;
 MISSING = ALL(99);

M-plus 파일

ANALYSIS:

MODEL:

DEP ON SE;
 EF ON SE;
 SUICIDE ON DEP EF SE;

MODEL INDIRECT:

SUICIDE IND SE;

OUTPUT:

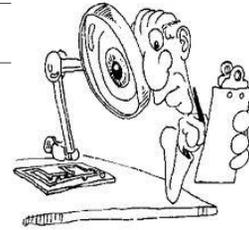
STANDARDIZED;

경로분석

1. 경로분석: 매개모형

2. 모형

3. M-plus의 예시



1. 경로분석: 매개모형

예제

• 연구문제:

직접효과

- 생활만족도 및 생활스트레스는 우울 수준에 영향을 주는가?
- 생활만족도 및 생활스트레스는 자기효능감 수준에 영향을 주는가?
- 우울과 자기효능감은 자살생각의 수준을 예측하는가?

간접효과

- 생활만족도 및 생활스트레스는 우울을 통해서 자살생각 수준에 영향을 주는가?
- 생활만족도 및 생활스트레스는 자기효능감을 통해서 자살생각 수준에 영향을 주는가?

• 자료프로그램 명:

PATH.DAT

모형 1
예제

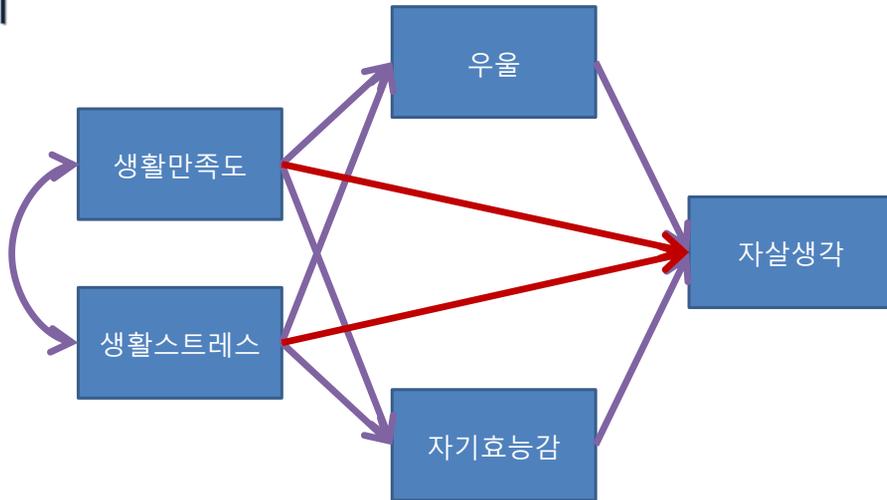


그림 1. 경로모형 5

모형 1
M-plus 프로그램

MODEL:

```
SUICIDE ON DEP SELF_EF;  
DEP ON SA_LIFE STR_LIFE;  
SELF_EF ON SA_LIFE STR_LIFE;  
SA_LIFE WITH STR_LIFE;
```

! THE BELOW LINES ARE ADDED

! DIRECT EFFECTS OF LIFE SATISFACTION AND LIFE STRESS ON SUICIDE IDEATION

SUICIDE ON SA_LIFE STR_LIFE;

OUTPUT:

```
MODINDICES;
```

모형 2

M-plus 프로그램(간접효과의 검증)

ANALYSIS:

BOOTSTRAP IS 5000;

MODEL:

SUICIDE ON DEP SELF_EF;

DEP ON SA_LIFE STR_LIFE;

SELF_EF ON SA_LIFE STR_LIFE;

SA_LIFE WITH STR_LIFE;

! THE BELOW LINES ARE ADDED

! DIRECT EFFECTS OF LIFE SATISFACTION AND LIFE STRESS ON SUICIDE IDEATION

SUICIDE ON SA_LIFE;

MODEL INDIRECT:

SUICIDE IND SA_LIFE;

SUICIDE IND STR_LIFE;

OUTPUT:

MODINDICES(10) **CINTERVAL(BCBOOTSTRAP);**

모형 2

M-plus 프로그램(간접효과의 검증)

CONFIDENCE INTERVALS OF TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT EFFECTS

	Lower .5%	Lower 2.5%	Lower 5%	Estimate	Upper 5%	Upper 2.5%	Upper .5%
Effects from SA_LIFE to SUICIDE							
Total	-0.557	-0.511	-0.491	-0.381	-0.275	-0.255	-0.211
Total indirect	-0.281	-0.252	-0.237	-0.165	-0.102	-0.091	-0.067
Specific indirect							
SUICIDE DEP SA_LIFE	-0.236	-0.213	-0.198	-0.126	-0.072	-0.062	-0.041
SUICIDE SELF_EF SA_LIFE	-0.107	-0.091	-0.082	-0.039	-0.012	-0.006	0.004
Direct							
SUICIDE SA_LIFE	-0.393	-0.347	-0.324	-0.216	-0.109	-0.089	-0.049

3. M-plus의 예시

간접효과의 추정 및 비교

• 연구문제:

직접효과

- 생활만족도는 우울 수준에 영향을 미치는가?
- 생활만족도는 자기효능감 수준에 영향을 미치는가?
- 우울과 자기효능감은 자살생각의 수준을 예측하는가?

간접효과

- 생활만족도는 우울을 통해서 자살생각 수준에 영향을 미치는가?
- 생활만족도는 자기효능감을 통해서 자살생각 수준에 영향을 미치는가?

• 자료프로그램 명:

PATH.DAT

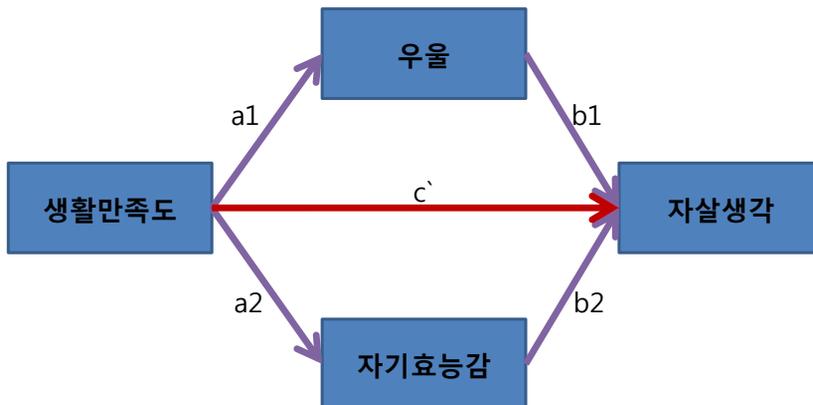
• 추가된 연구문제

- 우울의 매개효과와 자기효능감의 매개효과는 동일한가?

3. M-plus의 예시

간접효과의 추정 및 비교

경로모형(중다 부분 매개 모형)



3. M-plus의 예시

간접효과의 추정 및 비교

모형 1: M-plus 프로그램 (간접효과의 동등성 검증)

MODEL:

SUICIDE ON DEP (B1);
 SUICIDE ON SELF_EF (B2);
 DEP ON SA_LIFE (A1);
 SELF_EF ON SA_LIFE (A2);
 SUICIDE ON SA_LIFE;

MODEL CONSTRAINT:

NEW (A1B1 A2B2);
 A1B1 = A1*B1;
 A2B2 = A2*B2;

MODEL TEST:

A1B1 = A2B2;

OUTPUT:

MODINDICES;

3. M-plus의 예시

간접효과의 추정 및 비교

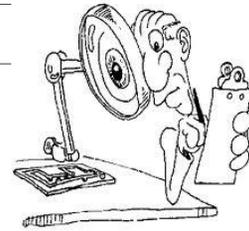
M-plus output

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
SUICIDE ON DEP	0.573	0.053	10.815	0.000
SELF_EF	-0.162	0.061	-2.648	0.008
SA_LIFE	-0.216	0.057	-3.810	0.000
DEP ON SA_LIFE	-0.410	0.048	-8.541	0.000
SELF_EF ON SA_LIFE	0.323	0.042	7.777	0.000
Intercepts				
DEP	4.144	0.177	23.428	0.000
SELF_EF	2.536	0.153	16.587	0.000
SUICIDE	1.727	0.344	5.017	0.000
Residual Variances				
DEP	0.539	0.038	14.125	0.000
SELF_EF	0.403	0.029	14.125	0.000
SUICIDE	0.588	0.041	14.125	0.000
New/Additional Parameters				
A1B1	-0.235	0.035	-6.702	0.000
A2B2	-0.052	0.021	-2.506	0.012

확인적 요인분석(측정모형)

1. 구조방정식 모형의 정의
2. 확인적 요인분석의 기초 및 개념
3. 확인적 요인분석의 예시(M-plus의 활용)
4. 다집단 분석의 개요
5. 다집단 분석의 순서
6. 모형 평가 기준
7. 다집단 분석에서의 유의사항
8. 다집단 분석의 예시(M-plus의 활용)



1. 구조방정식 모형이란?

구조방정식모형(Structural Equation Modeling, SEM)이란?

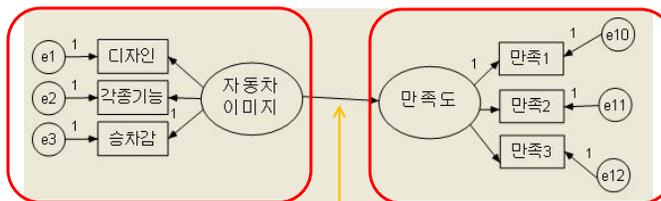
➢ 실험연구가 어렵거나 가능하지 않은 상황에서 **변수간의 인과관계에(추거적인 가정에 필요함)** 관한 추론을 가능하도록 해주는 통계기법

- 확인적 요인분석(Confirmatory Factor Analysis: CFA)과 경로분석(path analysis)이 결합된 형태이며,

모형 형태의 관점에서 보면 확인적 요인분석은 측정모형 (measurement model)

경로분석은 구조모형 (structural model, 이론모형이라고도 함)에 해당

- **AMOS(Arbuckle, 1997), Mplus(Muthen & Muthen, 1999) 등**



측정모형

구조모형

2. 확인적 요인분석의 기초 및 개념

기본개념

확인적 요인분석의 4단계 기본 과정

1. 요인모형을 이론에 맞게 설정(correctly specified-model).
 - ✓ 측정변수들을 그에 해당하는 잠재변수 (요인)들에 연결.
 - ✓ 가정된 요인모형과 일치하도록, 모수들을 추정하거나(freely estimated), 고정시키거나 (fixed), 혹은 다른 모수 값들과 동일시 시킴(equally constrained).
2. 요인모형을 자료로 적용(fitting the model to data).
3. 요인모형의 추정된 값과 적합도를 평가한다(evaluation of the model).
 - ✓ 요인모형을 자료에 적용할 때, 요인모형에 의해서 재생산된 분산행렬과 관찰된 분산행렬의 차이(discrepancy)를 최대한 감소시키는 모수가 추정됨.
 - ✓ 요인모형에 의해서 재생산된 분산행렬과 관찰된 분산행렬의 차이의 크기 (the amount of discrepancy)는 요인모형이 자료를 설명하는 정도를 나타냄.
 - ✓ 차이함수에 근거하여 합치도 지수 (Fit statistics)가 추정됨
 - ✓ 모수 추정치의 값이 사전에 가정된 크기의 정도와 방향이 같은지 확인.

2. 확인적 요인분석의 기초 및 개념

기본개념

확인적 요인분석의 4단계 기본 과정

4. 사전에 설정된 대안 모형들(다른 요인구조 가짐)을 비교.
 - ✓ 두 내재적(nested) 모형들을 비교할 때, 카이제곱 차이 검증 실시.
 - ✓ 두개의 비내재적(non-nested) 모형들을 비교할 때, RMSEA 또는 information criterion (e.g., the AIC and BIC)을 사용한다.
- 4-1. 한개의 요인모형이 검증이 되었고 적절한 수준의 합치도를 보이지 못하는 경우(즉, 모형이 자료를 충분히 설명하지 못하는 경우. 모형이 너무 간명한 경우),
 - ✓ 수정지수(modification index)를 활용하여 모형을 확장
 - ✓ 측정변수의 고유요인(unique factor) 간의 공분산을 고려(예, 동일한 형태의 질문, 비슷한 내용의 질문).

3. 확인적 요인분석의 예시(M-plus를 활용한 예시)

M-plus를 활용한 자료 분석

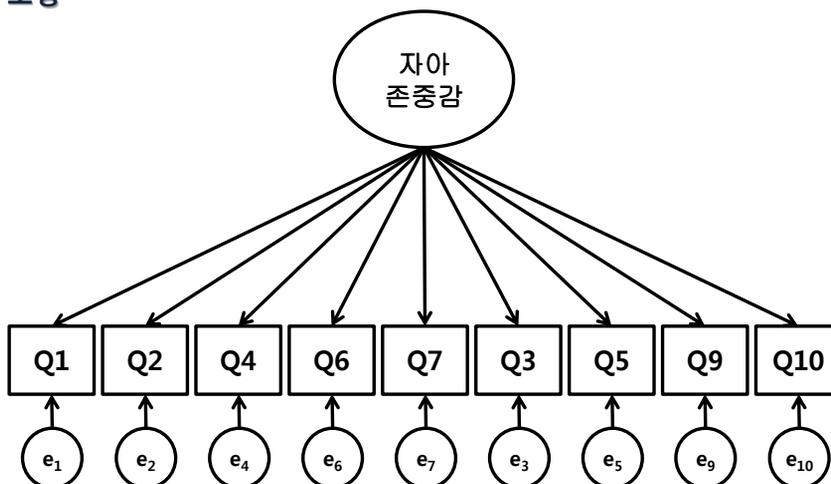
참고논문: 최수미, 조영일(2013), 부정문항이 포함된 척도의 요인구조 및 방법효과 검증과 남녀 간의 차이비교: Rosenberg 자기존중감 척도를 중심으로

- 표본크기: 2351명
- 변수: Rosenberg 자아존중감 척도 (10문항)
10개 문항이 사용되었지만, 1개 문항이 적절한 수준의 신뢰도를 보이지 않기 때문에 9개 문항만을 이번 예시에서 사용함.
- 자료이름: temp.dat
- 모형1: 1요인 모형(자존감)
- 모형2: 2요인 모형(정적 자존감 + 부정적 자존감)

3. 확인적 요인분석의 예시(M-plus를 활용한 예시)

M-plus를 활용한 자료 분석

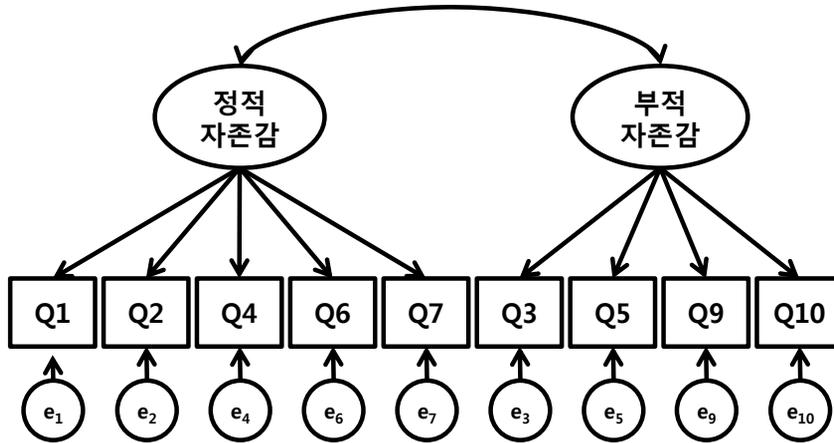
1요인 모형



3. 확인적 요인분석의 예시(M-plus를 활용한 예시)

M-plus를 활용한 자료 분석

2요인 모형



3. 확인적 요인분석의 예시(M-plus를 활용한 예시)

M-plus를 활용한 자료 분석

Input 파일(1요인 모형)

```

1Factor model
Title: 1요인 모형
DATA:
FILE = temp.dat;
NAMES = id female
q1-q10;
usevariable = q1-q7 q9 q10;
MISSING = q1 q2 q4 q6 q7 (-4)
q3 q5 q9 q10 (9);
DEFINE:
ANALYSIS:
MODEL:
SE BY q1
q2
q3
q4
q5
q6
q7
q9
q10;
OUTPUT:
SAMPSTAT STANDARDIZED TECH1 TECH4 MODINDICES (50);
  
```

1. 프로그램에 대한 설명을 기입한다. 이것은 1요인(자존감) 모형이다.
2. 원자료(raw data)가 들어있는 파일의 이름을 입력.
3. 화일에 들어 있는 변수들의 이름을 순서대로 기입.
4. 분석에서 사용될 변수들을 지정해 줌.
5. 각 변수들의 결측치에 해당하는 값을 지정해 줌.
6. 1요인 모형을 지정해 줌. 자존감을 9개의 문항으로 측정함.

3. 확인적 요인분석의 예시(M-plus를 활용한 예시)

M-plus를 활용한 자료 분석 Input 화할(모형 2: 2요인 모형)

```

Title:
2요인 모형(정적 자존감 vs. 부적자존감)

DATA:
FILE = temp.dat;

VARIABLE:
NAMES = id female
        q1-q10;

usevariable = q1-q7 q9 q10;

MISSING = q1 q2 q4 q6 q7 (-4)
          q3 q5 q9 q10 (9);

DEFINE:

ANALYSIS:

MODEL:
PE BY Q1 Q2 Q4 Q6 Q7;
NE BY Q3 Q5 Q9 Q10;

OUTPUT:
SAMPSTAT STANDARDIZED TECH1 TECH4 MODINDICES (50);
    
```

정적 자존감은 Q1, Q2, Q4, Q6, Q7
부적 자존감은 Q3, Q5, Q9, Q10
으로 측정된다.

MODEL:
PE BY Q1 Q2 Q4 Q6 Q7;
NE BY Q3 Q5 Q9 Q10;

3. 확인적 요인분석의 예시(M-plus를 활용한 예시)

M-plus를 활용한 자료 분석 모형 비교 결과 보고하기(모형 1, 2)

모형 합치도 지수

모형	$\chi^2(df)$	TLI	CFI	RMSEA [90% CI]
모형 1	1613.78 (27)	.761	.821	.158 [.152, .165]
모형 2	299.08 (26)	.957	.969	.067 [.060, .074]

두 개의 모형을 적합도 지수를 활용하여 비교할 때, 두 개의 모형들이 내재된 관계를 보이지 않기 때문에 카이 제곱 차이 검증을 사용할 수 없다.
대신에, 카이 제곱 및 다른 상대적 적합도 지수의 크기를 비교함으로써 자료를 보다 잘 설명하는 모형을 찾을 수 있다.
이 경우에는, 모형 2이 다른 1개의 모형들에 비해서 비교적 자료를 잘 설명함을 알 수 있다. 따라서 모형 2을 채택한다.

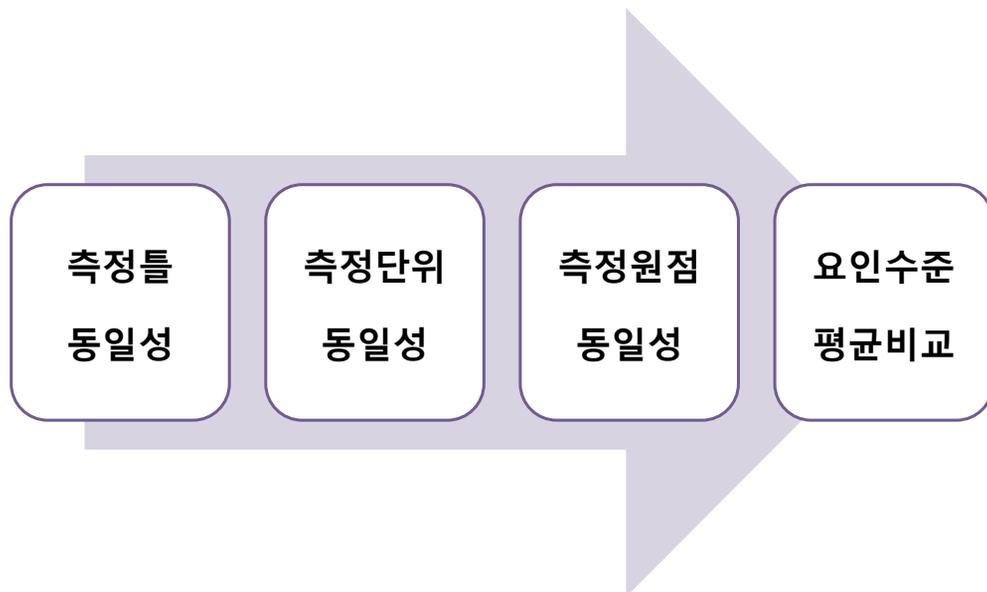
4. 다집단 분석의 개요

목적(Objectives) (이순목, 이찬순, 이현정, 여성철, 2012)

- 질적으로 구분되는 상이한 모집단을 대상으로 일반화를 하기 위한 분석 방법
- 동일한 측정치(일반적으로 심리검사나 심리척도)를 사용하여
 - 측정모형, 이론모형, 또는 측정모형과 이론모형을 포함하는 전체모형을 구성하고
 - 상이한 모집단들에 동시에 합치시켜 집단간 동등성을 분석하는 방법으로서
 - 상황(예: 집단, 시점, 장소 등)에 걸친 일반화 가능성 또는 집단차이를 연구하는 것.
- 추정하고자 하는 모형의 (상이한)집단 간의 차이가 표집오차(sampling error)에 기인한 것으로 (확률적 및 실제적) 판단되면 모형이 일반화된 것으로 결론짓는다(이순목, 김한조, 2011)

5. 다집단 분석의 순서

순서(Procedures) (이순목, 이찬순, 이현정, 여성철, 2012)



6. 다집단 분석의 모형평가

다집단 분석에서의 모형평가 (이순목, 이찬순, 이현정, 여성철, 2012)

- 다집단 분석에서의 모형의 동일성 검증
 - 두 개의 모형(이론적으로 중요한 가정을 검증, 내재된 모형)들 간의 합치도를 비교.
 - 또한, 두 모형에서의 합치도 값의 변화를 실용적인 관점(practical perspective)에서 평가함
- 합치도 평가의 기준(다집단 모형 평가, Cheung & Rensvold, 2002)
 - **동일화 제약을 추가한 모형의 합치도와 추가되기 전 모형(기저모형)의 합치도 값을 비교**
 - 카이제곱 차이검증(chi-square difference test): 카이제곱 분포를 따름
 - **CFI의 변화값 < -.01(보다 선호됨, Little, 2013)**
 - McDonald의 mc 변화값 < .02
 - Steiger의 Gamma hat 변화값 < -.001

7. 다집단 분석의 유의사항

참조지표 선정 (이순목, 이찬순, 이현정, 여성철, 2012)

- 요인분석에서 요인의 인식(identification)
 - 측정모형에서 사용된 관찰변수들 중에서 한 관찰변수의 요인계수 값을 1로 고정함(**참조지표** 사용)
 - 요인의 분산을 1로 고정함(표준화 시킴).
 - 많은 SEM 프로그램에서 참조지표 방식을 default로 사용함.
- 참조지표 방식을 사용하여 다집단 분석을 실시할 경우 참조지표가 집단간 동일하지 않음으로 인하여 문제가 발생함.
- 해결방안
 - 탐색적 요인분석(사각회전)에서 집단 간 요인계수의 값이 크고 유사한 값을 가지는 측정변수를 참조지표로 정함.
 - 유사함을 강조(Vandenberg, 2002)
 - 요인계수의 크기를 강조(Chan, 2000)
 - 요인계수의 값이 커서 그 변수의 구성개념타당도가 높아야 하고 집단 간에 유사한 값을 가짐으로써 측정단위가 동일할 가능성을 보여야 함.

8. M-plus를 활용한 자료분석 예시

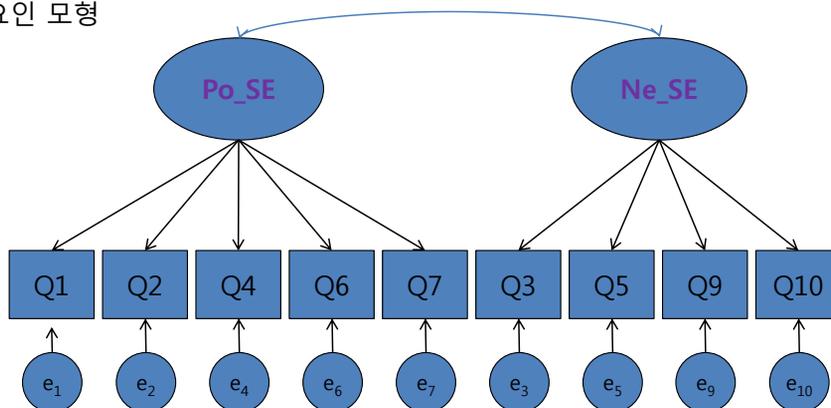
참고논문: 최수미, 조영일(2013), 부정문항이 포함된 척도의 요인구조 및 방법효과 검증과 남녀 간의 차이비교: Rosenberg 자기존중감 척도를 중심으로, 한국심리학회지: 일반.

- **측정모형의 동일성 검증**
- 연구문제 1: 남녀 중학생 간의 측정모형은 동일할 것이다.
- **요인평균의 비교**
- 연구문제 2: 남자 중학생의 자기 존중감은 여자 중학생의 자기 존중감에 비하여 높을 것이다.

8. M-plus를 활용한 자료분석 예시

측정틀 동일성 검증

- 2요인 모형



8. M-plus를 활용한 자료분석 예시

측정틀 동일성 검증

- 남녀 중학생 집단을 **동시에** 추정

```
Title:
DATA:
  FILE = rosenberg.dat;
VARIABLE:
  NAMES = id female
         q1-q10;
  usevariable = q1-q7 q9 q10;
  GROUPING IS FEMALE (0=MALE 1=FEMALE);
  MISSING = ALL(9);DEFINE:
ANALYSIS:
MODEL:
  F1 BY Q1 Q2 Q4 Q6 Q7;
  F2 BY Q3 Q5 Q9 Q10;

  [F1@0 F2@0];

MODEL FEMALE:
F1 BY Q2 Q4 Q6 Q7;
F2 BY Q5 Q9 Q10;
[Q1-Q7 Q9 Q10];
OUTPUT:
  SAMPSTAT STANDARDIZED TECH1 TECH4 MODINDICES (50);
```

조영일, Ph.D

71/75

8. M-plus를 활용한 자료분석 예시

측정단위 동일성 검증

MODEL:

```
F1 BY Q1 Q2 Q4 Q6 Q7;
F2 BY Q3 Q5 Q9 Q10;
```

```
[F1@0 F2@0];
F1 F2;
F1 WITH F2;
```

MODEL FEMALE:

```
!F1 BY Q1@1 Q2 Q4 Q6 Q7;
!F2 BY Q3@1 Q5 Q9 Q10;
```

```
F1 F2;
[Q1-Q7 Q9 Q10];
```

OUTPUT:

```
SAMPSTAT STANDARDIZED TECH1 TECH4 MODINDICES;
```

조영일, Ph.D

72/75

8. M-plus를 활용한 자료분석 예시 측정원점 동일성 검증

MODEL:

F1 BY Q1 Q2 Q4 Q6 Q7;
F2 BY Q3 Q5 Q9 Q10;

[F1@0 F2@0];
F1 F2;
F1 WITH F2;

MODEL FEMALE:

!F1 BY Q1@1 Q2 Q4 Q6 Q7;
!F2 BY Q3@1 Q5 Q9 Q10;

F1 F2;
[F1 F2];
![Q1-Q7 Q9 Q10];

OUTPUT:

SAMPSTAT STANDARDIZED TECH1 TECH4 MODINDICES;

조영일, Ph.D

73/75

8. M-plus를 활용한 자료분석 예시 측정오차 동일성 검증

MODEL:

F1 BY Q1 Q2 Q4 Q6 Q7;
F2 BY Q3 Q5 Q9 Q10;

[F1@0 F2@0];
F1 F2;
F1 WITH F2;

Q1 (11); Q2 (12); Q3 (13); Q4 (14);
Q5 (15); Q6 (16); Q7 (17); Q9 (18);
Q10 (19);

MODEL FEMALE:

!F1 BY Q1@1 Q2 Q4 Q6 Q7;
!F2 BY Q3@1 Q5 Q9 Q10;

F1 F2;
[F1 F2];
![Q1-Q7 Q9 Q10];

OUTPUT:

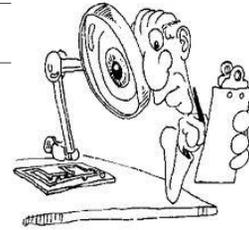
SAMPSTAT STANDARDIZED TECH1 TECH4 MODINDICES;

조영일, Ph.D

74/75

구조방정식 모형

1. 구조방정식모형의 기초
2. 구조방정식모형의 활용
3. M-plus를 활용한 예시



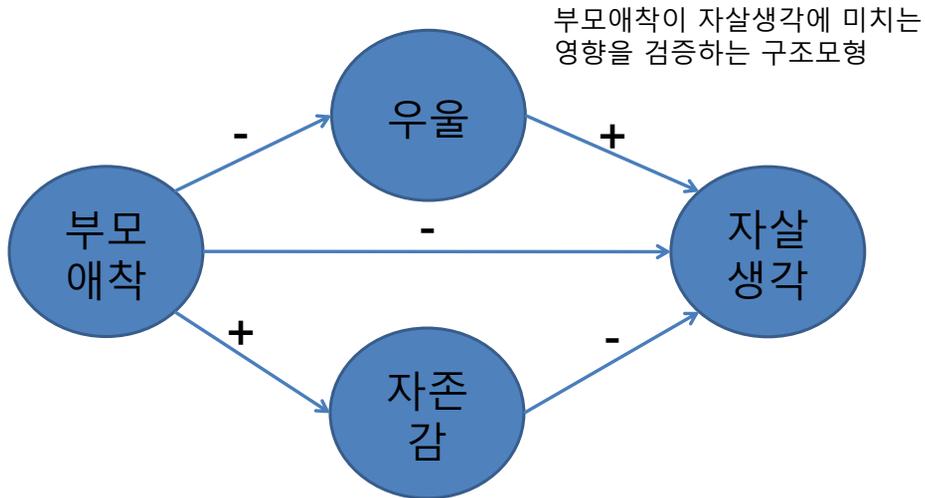
1. 구조방정식모형의 기초

구조방정식모형(Structural Equation Modeling, SEM)이란?

- 구조방정식 모형은 측정오차를 제거한 잠재변수에 바탕을 둔 경로분석이다.
 - 측정모형: 잠재변수와 측정변수의 관계
 - ✓ 확인적 요인분석
 - 구조모형: 잠재변수 간의 관계
 - ✓ 경로분석

1. 구조방정식모형의 기초

구조모형(structural model)



2. 구조방정식모형의 활용

문항꾸러미(parcel)의 사용

- 문항꾸러미를 사용하는 이유(Bandalos & Finney, 2001)
 - increasing the stability of the parameter estimates (29%)
 - improving the variable to sample size ratio (22.6%)
 - remedying small sample sizes (21%).
- 문항꾸러미 작성의 논리
 - ✓ To reduce the effects of nonnormality and likelihood of forming difficulty factors in factor analyses with binary items.
 - ✓ Using item parceling with nonnormal data did result in more normally distributed item parcels and improved model fit.
 - ✓ Their method of item parceling was to create parcels of items with opposite skew in an iterative procedure that resulted in parcels with less skew than the original items.

2. 구조방정식모형의 활용

Factor-based unidimensional parcels (FBP)

- Form parcels with items that relate to specific subdomains of scale
- E.G., $FBP1=A1+A2+A3$

A1: I felt depressed.

A2: I felt that I could not shake off the blues even with help from my family or friends.

A3: I felt sad.

Domain Representative parcels (DRP)

- Form parcels with items that span the subdomains of item set
- E.G., $DRP1=A1+B1+C1$

A1: I felt depressed.

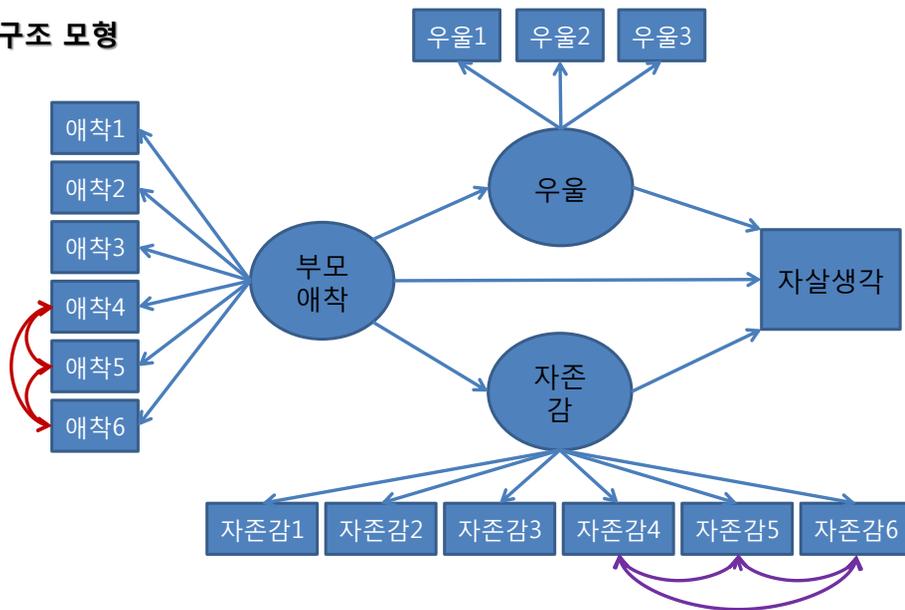
B1: I could not get going.

C1: I felt that people dislike me.

- Parcels are designed to be close to parallel measures

3. Mplus를 활용한 예시

구조 모형



3. Mplus를 활용한 예시

M-plus input 화일

Title: Structural Equation Modeling;

DATA:

FILE = SEM.dat;

VARIABLE:

NAMES = id female

ATT1-ATT6

EST1-EST6

DEP1-DEP3

SUICIDE;

USEVARIABLES = ATT1-ATT6

EST1-EST6

DEP1-DEP3

SUICIDE;

MISSING = ALL(999);

ANALYSIS:

MODEL:

ATT BY ATT1 ATT2 ATT3

ATT4 ATT5 ATT6;

EST BY EST1 EST2 EST3

EST4 EST5 EST6;

DEP BY DEP1 DEP2 DEP3;

EST4 WITH EST5 EST6;

EST5 WITH EST6;

ATT4 WITH ATT5 ATT6;

ATT5 WITH ATT6;

SUICIDE ON ATT EST DEP;

EST ON ATT;

DEP ON ATT;

OUTPUT:

SAMPSTAT STANDARDIZED TECH1

TECH4 MODINDICES (10);

3. Mplus를 활용한 예시

최종 구조모형(structural model)의 보고

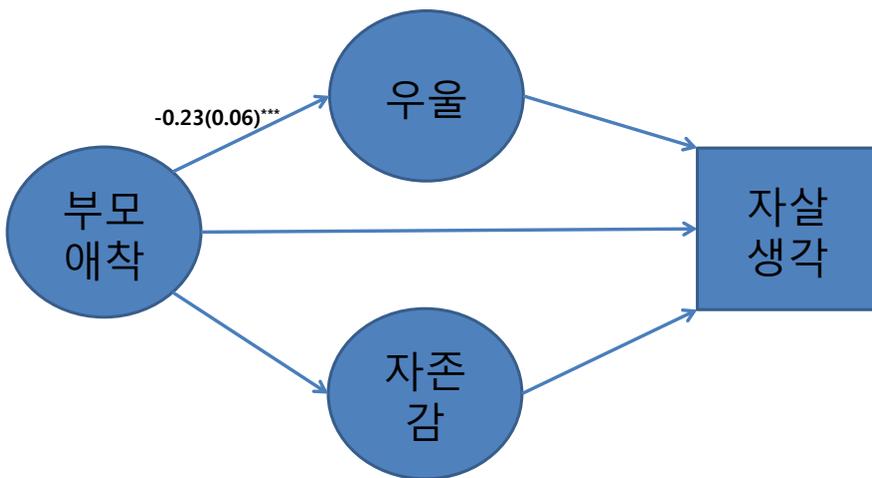
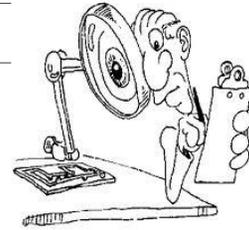


그림 1. 부모애착이 자살생각에 미치는 영향을 검증하는 구조모형

참조. 표준화된 경로계수이고 괄호안의 숫자는 표준오차임.

종단자료의 분석

1. 종단자료(Longitudinal data)란?
2. 자료분석의 2가지 접근법
3. 종단자료에서의 측정모형
4. 종단자료에서의 측정모형의 동일성 검증
5. M-plus를 활용한 자료 분석 예시-측정모형의 동일성 검증
6. M-plus를 활용한 자료 분석 예시-척도의 연결(Linking)



1. 종단자료란?

● 횡단자료(cross-sectional data)

- 다양한 집단을 하나의 시점에서 동시에 측정 혹은 평가하여 얻어진 자료.
- 각각의 연구 참여자들을 한번만 측정함.
- 집단간 평균 비교에 용이함.

● 종단자료(longitudinal data)

- 다양한(혹은 하나의) 집단을 여러 시점에 걸쳐서 여러 번 측정 혹은 평가하여 얻어진 자료
- 각각의 연구 참여자들을 시간의 흐름에 따라서 여러 번에 걸쳐서 측정함.
- 집단간 평균 비교 뿐만 아니라 집단의 변화를 추정할 수 있음.

1. 종단자료란?

- 종단자료의 특징

- Some of the same entities (at least some of them) are observed at repeated occasions.
- The measurement and scaling of observations are known.
- The ordering or time underlying the observations is known.

1. 종단자료란?

- 종단자료를 사용한 연구의 목적(Baltes & Nesselrode, 1979)

1. Identification of **intra**-individual change
2. Direct identification of **inter**-individual differences in intra-individual change
3. Analysis of **inter-relationships** in change
4. Analysis of **causes (determinants)** of intra-individual change
5. Analysis of **causes (determinants)** of inter-individual differences in intra-change

1. 종단자료란?

● 종단자료를 사용한 연구의 목적(Baltes & Nesselroade, 1979)

1. Identification of **intra**-individual change ⇒ 변화 함수의 식별 (선형 또는 비선형)
2. Direct identification of **inter**-individual differences in intra-individual change ⇒ 변화 함수내 모수의 분산 추정 (univariate unconditional LGM)
3. Analysis of **inter-relationships** in change ⇒ 두 변수의 변화 형태의 관련성 연구 (multivariate unconditional LGM)
4. Analysis of **causes (determinants)** of intra-individual change
5. Analysis of **causes (determinants)** of inter-individual differences in intra-change ⇒ 변화 형태의 개인 차의 원인에 대한 연구 (conditional LGM)

1. 종단자료란?

Cohort (Birth Year)	Time of Measurement														
	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Gen-Xers 1975	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
1976	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
1977	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
1978	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1979	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
1980	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1981	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1982	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1983	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1984	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1985	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1986	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1987	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1988	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1989	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1990	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

FIGURE 2.1. Traditional sequential designs. Ages are entered in gray in the body of the table. Any given row would be an example of a longitudinal design, and any given column would be an example of a cross-sectional study. A cohort-sequential design would consist of starting a new cohort at a certain age and then following longitudinally. A cross-sequential design starts with a traditional cross-sectional study and then follows all participants longitudinally. A time-sequential design is a repeated cross-sectional design, with some participants followed longitudinally.

2. 자료 분석의 두 가지 접근법

변수중심 방법(variable-centered approach)

- 집단의 구성원들이 동질적이라고 가정함
- The description of **associations between variables**.
- They are well suited for addressing questions that concern **the relative contributions** that predictor variables make to an outcome.
- 예: 요인분석(factor analysis)

사람중심 방법(person-centered approach)

- 집단의 구성원들이 이질적이라고 가정함.
- The identification of **groups of individuals** who share particular attributes or relations among attributes.
- They are well suited for addressing questions that concern **group differences in patterns of development**.
- 예: 군집분석(cluster analysis)

2. 자료 분석의 두 가지 접근법

변수중심 방법(variable-centered approach)-요인분석

수리능력과 언어능력 요인

지능검사의 상관행렬

어휘력 언어추리 수계산 수추리

어휘력	1.00			
언어추리	.80	1.00		
수계산	.25	.33	1.00	
수추리	.45	.38	.73	1.00

어휘력 언어추리 수계산 수추리

언어능력

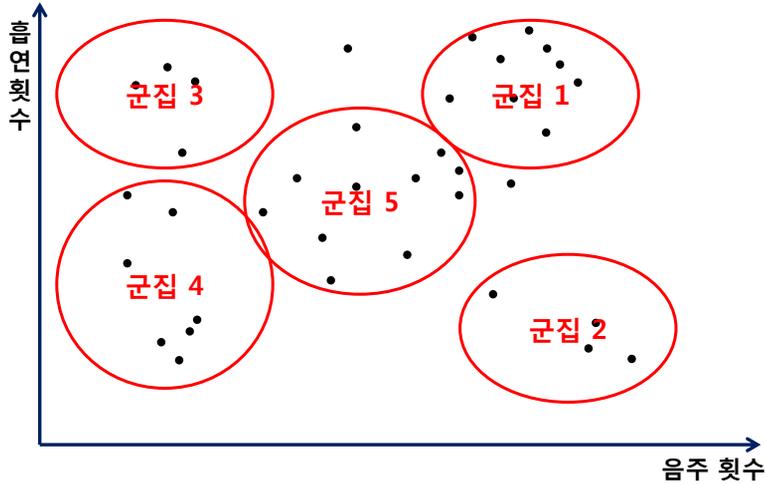
수리능력

▶ 요인분석을 활용해서 변수들 간의 관련성을 설명하고 요약할 수 있는 요인(factor)를 추정하고, 측정변수에서의 개인차를 요인으로 설명함.

2. 자료 분석의 두 가지 접근법

사람중심 방법(person-centered approach)-군집분석

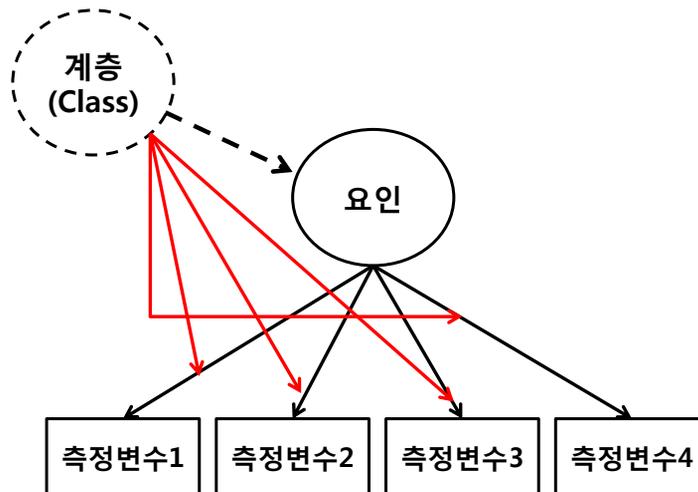
음주 및 흡연 정도에 따른 군집의 인식 및 구분



▶ 측정변수에서 유사한 반응형태를 보이는 개인들을 하나의 소집단으로 묶음. 소집단(군집 혹은 계층)으로 묶음으로써 측정변수에서의 개인차를 설명함.

2. 자료 분석의 두 가지 접근법

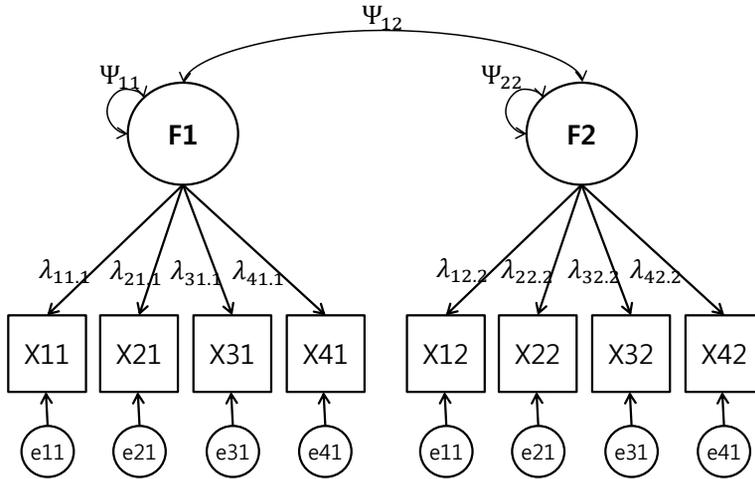
잠재변수모형(latent variable modeling)



▶ 계층 혹은 군집을 확인한 후에는 소집단의 특성을 추정하기 위해서 측정변수의 점수를 바탕으로 profile을 작성하는 것이 필요.

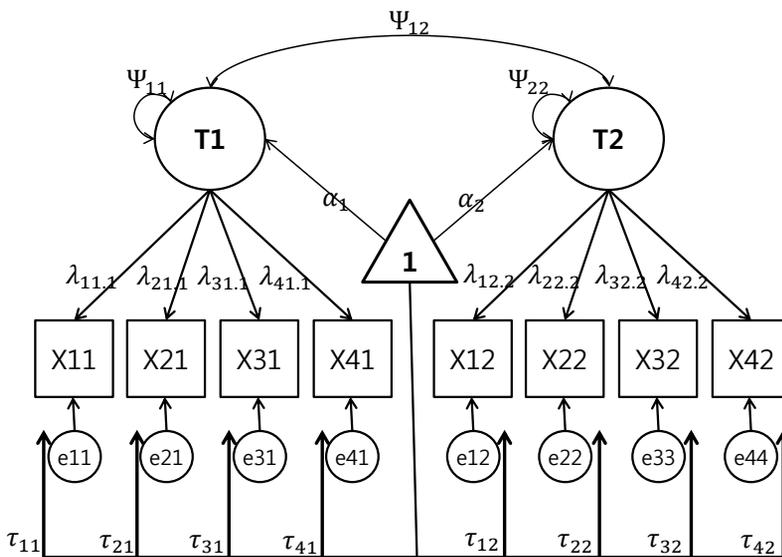
3. 종단자료에서의 측정모형

- 횡단자료에서의 측정모형(확인적 요인분석)



3. 종단자료에서의 측정모형

- 종단자료에서의 측정모형(확인적 요인분석)



3. 종단자료에서의 측정모형

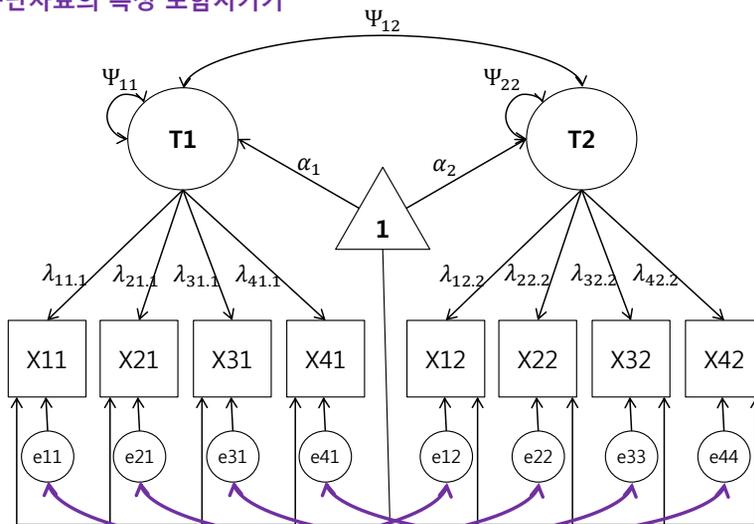
● 종단자료에서의 측정모형(확인적 요인분석)

- 종단자료의 특성 포함시키기
 - 측정모형(확인적 요인분석)에서
 - 측정변수 분산=공통요인분산 + 오차분산
 - 오차분산=고유요인분산(unique factor variance) + 무선오차(random error)
 - 고유요인 분산(item-specific variance)
 - 측정변수의 분산 중에서 공통요인에 의해서 설명되지 않지만,
 - 측정변수의 특수성을 반영하여 일관되게 관찰되는 분산
- 종단자료에서는 **동일한 측정변수가 여러 시점에 걸쳐서 사용됨.**
 - 동일한 측정변수의 고유요인 분산이 시점 간의 **공분산**을 가짐
 - 시점 간의 측정오차의 공분산이 통계적으로 유의하지 않아도 유지하는 것이 필요(Little, 2013)

3. 종단자료에서의 측정모형

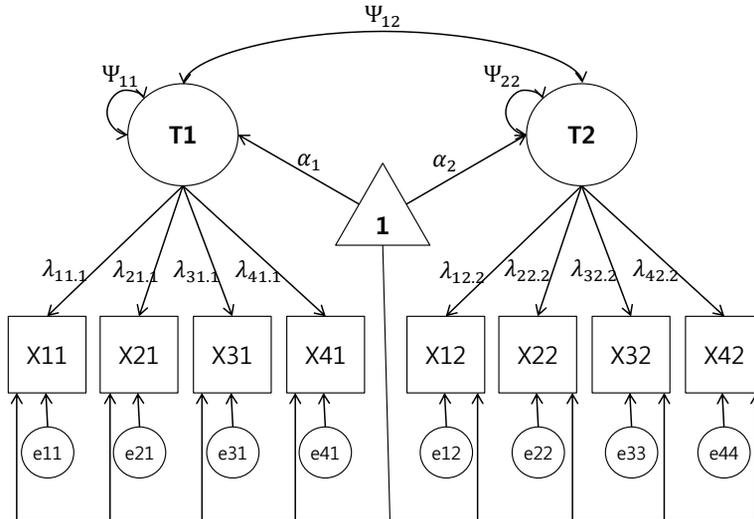
● 종단자료에서의 측정모형(확인적 요인분석)

- 종단자료의 특성 포함시키기



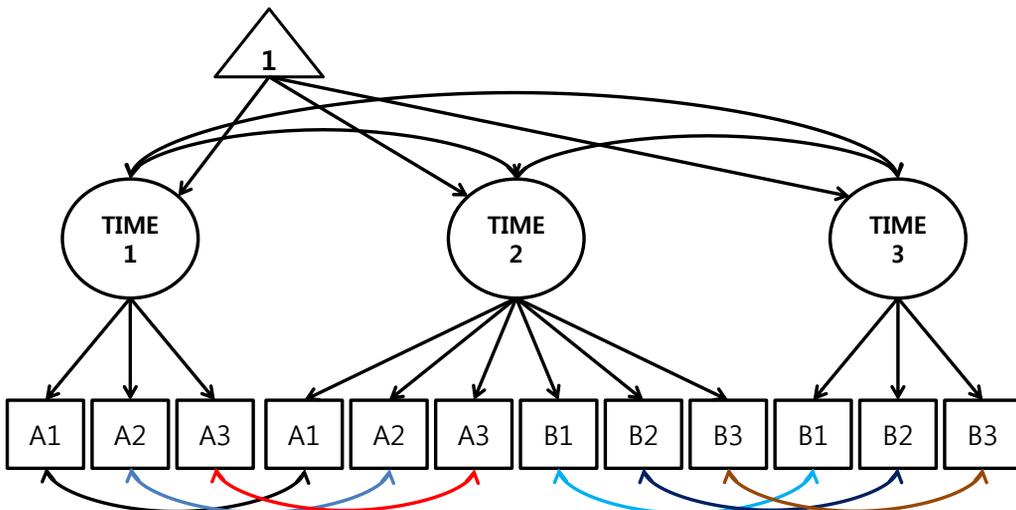
3. 종단자료에서의 측정모형

- 측정모형의 동일성 검증(확인적 요인분석)



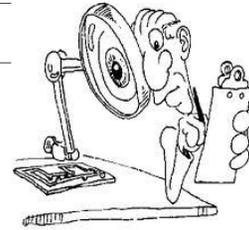
3. 측정모형의 동일성 검증

- Linking: 두 개의 다른 척도 연결시키기



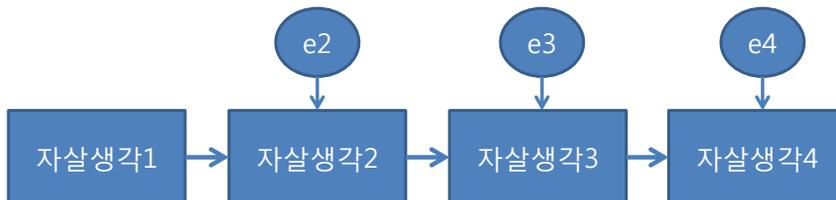
자기회귀모형

1. 자기회귀모형
2. 1차 자기회귀모형
3. 2차 자기회귀모형
4. 교차지연자기회귀모형



4. 자기회귀모형 자기회귀모형

$$y_{it} = \alpha_{yt} + \rho_{yt, yt-1} y_{i, t-1} + \varepsilon_{yit}$$



4. 자기회귀모형

1st-order 자기회귀모형

M-plus input 화일

Title: 1st-order autoregressive model;

DATA:

```
FILE = nbully_suicide.dat;
```

VARIABLE:

```
NAMES = id female nv1-nv5 VICTIM1-VICTIM5 SUICIDE1-SUICIDE5 DEP1-DEP5 ESTEEM1-ESTEEM5;
```

```
USEVARIABLES = SUICIDE1-SUICIDE4;
```

```
MISSING = ALL(999);
```

ANALYSIS:

MODEL:

```
SUICIDE2 ON SUICIDE1 (3); SUICIDE3 ON SUICIDE2 (3);
SUICIDE4 ON SUICIDE3 (3);
```

OUTPUT:

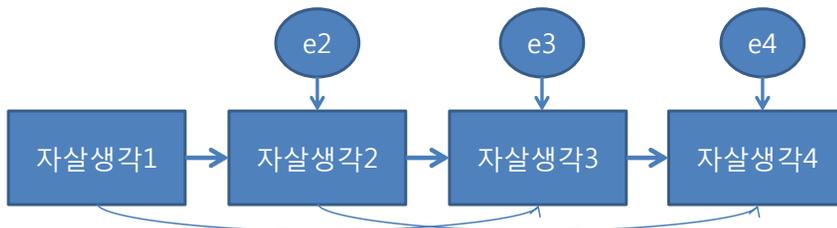
```
SAMPSTAT STANDARDIZED TECH1 TECH4 MODINDICES (10);
```

101/127

4. 자기회귀모형

2nd-order 자기회귀모형

$$y_{it} = \alpha_{yt} + \rho_{yt, yt-1} y_{i,t-1} + \rho_{yt, yt-2} y_{i,t-2} + \varepsilon_{yit}$$



한 시점의 값(t)이 이전 시점의 값(t-1)과 그 이전 시점의 값(t-2)에 의해 설명된다.

102/127

4. 자기회귀모형

2nd-order 자기회귀모형

M-plus input 화일

MODEL :

```

SUICIDE2 ON SUICIDE1 (7) ;
SUICIDE3 ON SUICIDE2 (7) ; SUICIDE3 ON SUICIDE1 (8) ;
SUICIDE4 ON SUICIDE3 (7) ; SUICIDE4 ON SUICIDE2 (8) ;
    
```

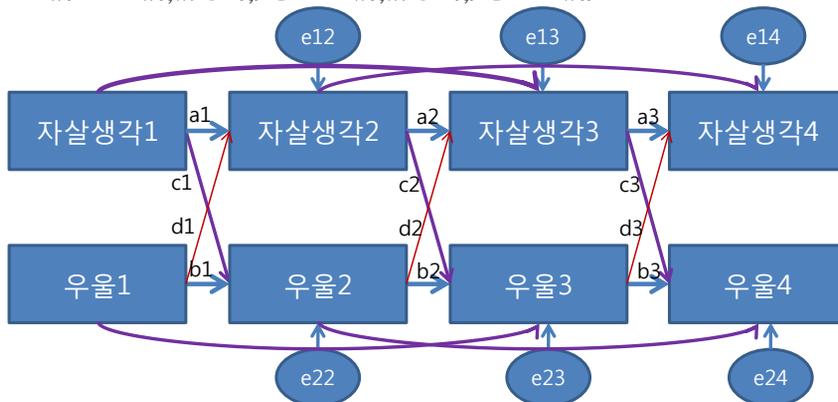
2nd-order 자기회귀계수를 추가하였다.

4. 자기회귀모형

교차연자기회귀모형

$$y_{it} = \alpha_{yt} + \rho_{yt,yt-1} y_{i,t-1} + \rho_{yt,wt-1} w_{i,t-1} + \varepsilon_{yit}$$

$$w_{it} = \alpha_{wt} + \rho_{wt,wt-1} w_{i,t-1} + \rho_{wt,yt-1} y_{i,t-1} + \varepsilon_{wit}$$



일련의 카이제곱차이 검증을 이용하여 회귀계수의 동일성을 검증하고 보다 간명한 모형을 선택할 수 있다. 예를 들어서, $c1 = c2 = c3$, or $d1 = d2 = d3 = 0$.

4. 자기회귀모형

M-plus input 화일

MODEL :

```
DEP2 ON DEP1 (A3);  
DEP3 ON DEP2 (A3); DEP3 ON DEP1 (A4);  
DEP4 ON DEP3 (A3); DEP4 ON DEP2 (A4);
```

우울의 자기회귀모형

```
SUICIDE2 ON SUICIDE1 (A7);  
SUICIDE3 ON SUICIDE2 (A7); SUICIDE3 ON SUICIDE1 (A8);  
SUICIDE4 ON SUICIDE3 (A7); SUICIDE4 ON SUICIDE2 (A8);
```

자살생각 자기회귀모형

```
SUICIDE2 ON DEP1 (A6);  
SUICIDE3 ON DEP2 (A6);  
SUICIDE4 ON DEP3 (A6);
```

자살생각에 대한 우울의 효과

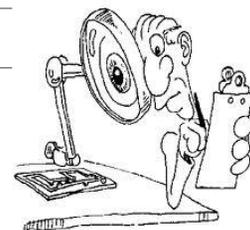
```
DEP2 ON SUICIDE1 (C1);  
DEP3 ON SUICIDE2 (C1);  
DEP4 ON SUICIDE3 (C1);
```

우울에 대한 자살생각의 효과

```
DEP4 WITH SUICIDE4 (B3);  
DEP2 WITH SUICIDE2 (B3);  
DEP3 WITH SUICIDE3 (B3);
```

잠재성장모형

1. LGM 추정의 4단계
2. 무조건화 잠재성장모형
3. 조건화 잠재성장모형
4. 2nd-order 잠재성장모형
5. 다변량 잠재성장모형



5. 잠재성장모형

LGM 추정의 4단계

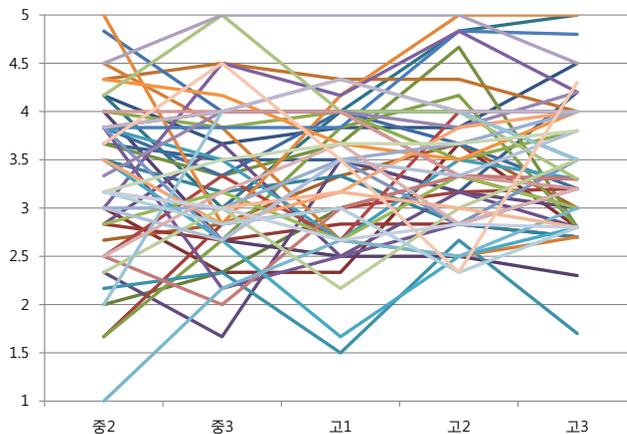
1. Step 1: A model of change is specified
2. Step 2: Expectations about means and covariances are generated based on the specified model
3. Step 3: Parameters are estimated
4. Step 4: Model fit is evaluated (discrepancy between model expectations and observed data)

107/127

5. 잠재성장모형

Step 1.

- 시간을 X축으로 하고 반복측정된 점수를 Y축으로 각 개인들마다 그린다.
- 이를 통해서, 개인들의 변화함수의 형태를 추측한다.



108/127

5. 잠재성장모형

Step 2.

- 개인 내 모형

$$y_{it} = \alpha_i + \beta_i x_{it} + \varepsilon_{it}$$

- 개인 간 모형

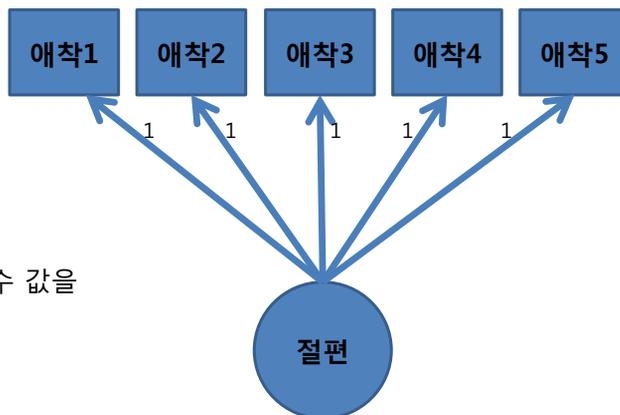
$$\alpha_i = \mu_\alpha + \zeta_{\alpha i}$$

$$\beta_i = \mu_\beta + \zeta_{\beta i}$$

5. 잠재성장모형

Step 2.

무변화모형(No growth)

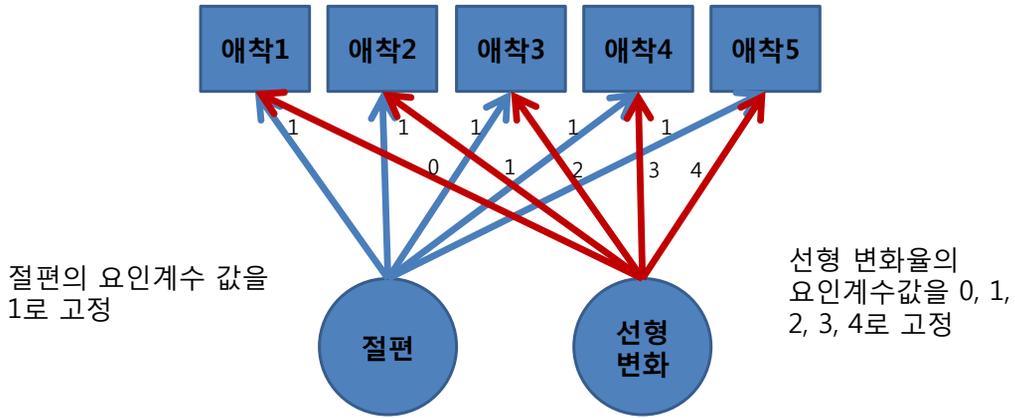


절편의 요인계수 값을
1로 고정

5. 잠재성장모형

Step 2.

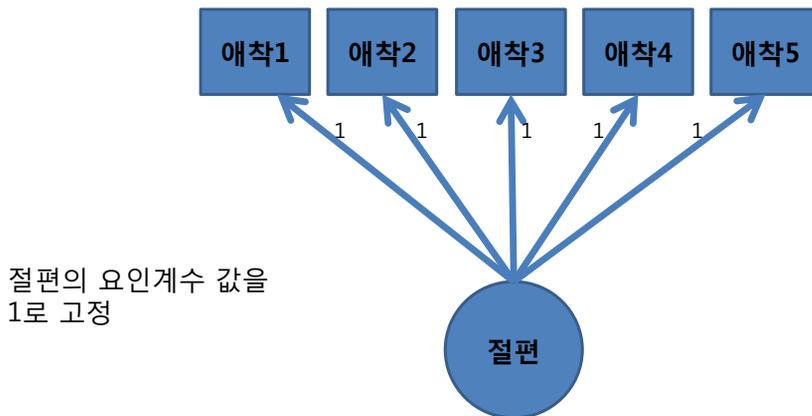
선형변화모형



5. 잠재성장모형

Step 3. 모형 비교

무변화모형(No growth)



5. 잠재성장모형

Step 3. 모형 비교(무성장 모형)

M-Plus 프로그램의 예

TITLE: Unconditional LGM

DATA: FILE IS ATTACH.DAT;

VARIABLE: NAMES ARE ID GENDER

ATT1 ATT2 ATT3 ATT4 ATT5;

USEVARIABLE ARE ATT1 ATT2 ATT3 ATT4 ATT5;

MISSING = ALL(99);

ANALYSIS:

MODEL:

I | ATT1@0 ATT2@1 ATT3@2 ATT4@3 ATT5@4;

OUTPUT:

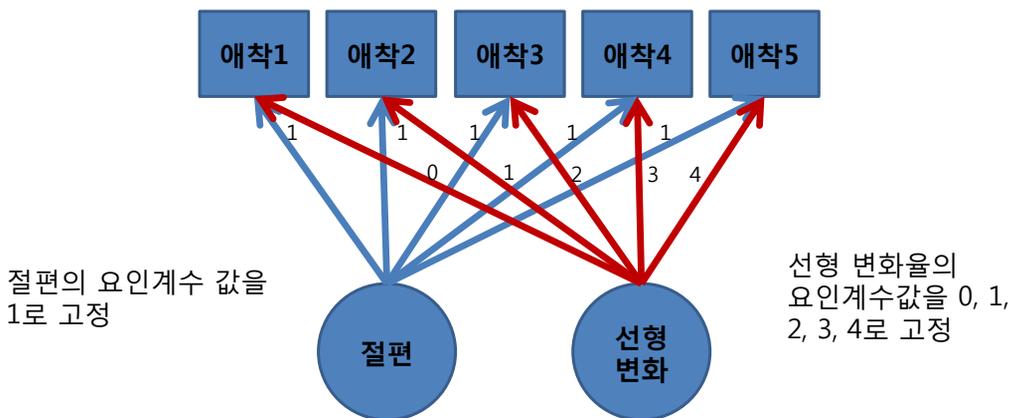
MODINDICES STANDARDIZED;

무성장 잠재성장모형을 specification하기 위해서는 'I'라는 명령어가 필요함.

5. 잠재성장모형

Step 3. 모형 비교

선형변화모형



절편의 요인계수 값을 1로 고정

선형 변화율의 요인계수값을 0, 1, 2, 3, 4로 고정

5. 잠재성장모형

Step 3. 모형 비교(선형 성장)

M-Plus 프로그램의 예

TITLE: Unconditional LGM

DATA: FILE IS ATTACH.DAT;

VARIABLE: NAMES ARE ID GENDER

ATT1 ATT2 ATT3 ATT4 ATT5;

USEVARIABLE ARE ATT1 ATT2 ATT3 ATT4 ATT5;

MISSING = ALL(99);

ANALYSIS:

MODEL:

I S | ATT1@0 ATT2@1 ATT3@2 ATT4@3 ATT5@4;

OUTPUT:

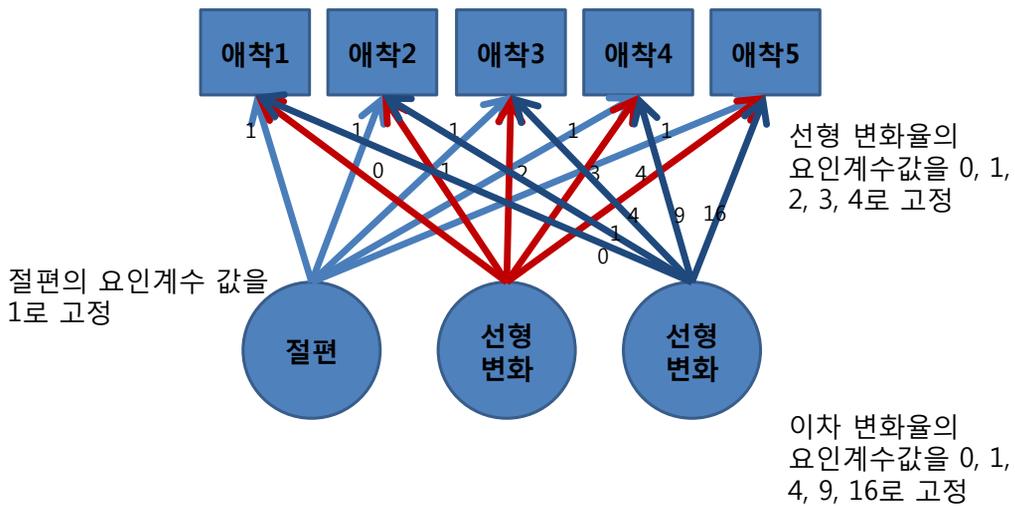
MODINDICES STANDARDIZED;

선형 잠재성장모형을 specification하기 위해서는 'I'라는 명령어가 필요함.

5. 잠재성장모형

Step 3. 모형 비교(2차 선형 성장모형)

2차선형변화모형



5. 잠재성장모형

Step 3. 모형 비교(2차 선형 성장모형)

M-Plus 프로그램의 예

TITLE: Unconditional LGM

DATA: FILE IS ATTACH.DAT;

VARIABLE: NAMES ARE ID GENDER

ATT1 ATT2 ATT3 ATT4 ATT5;

USEVARIABLE ARE ATT1 ATT2 ATT3 ATT4 ATT5;

MISSING = ALL(99);

ANALYSIS:

MODEL:

I S Q | ATT1@0 ATT2@1 ATT3@2 ATT4@3 ATT5@4;

OUTPUT:

MODINDICES STANDARDIZED;

2차선형 잠재성장모형을 specification하기 위해서는 'I'라는 명령어가 필요함.

5. 잠재성장모형

Step 3. 모형 비교(모형의 합치도 지수 비교)

결과 보고

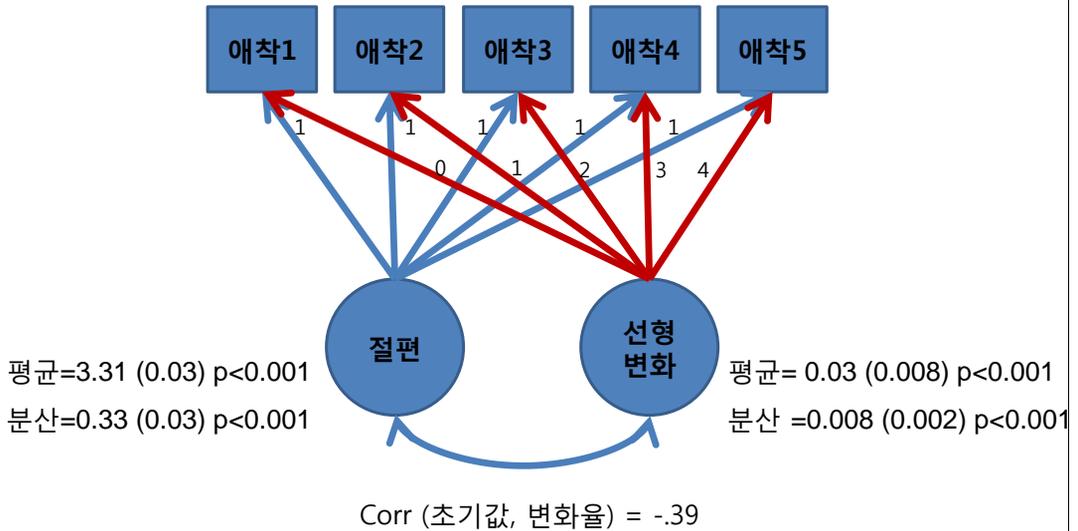
표2. 모형의 합치도 지수

모형	χ^2	df	Δdf	$\Delta \chi^2$	CFI	TLI	RMSEA
무성장	64.40	13			.943	.956	.087
1차 선형	34.42	10	3	29.98***	.975	.975	.067
2차 선형	18.72	6	4	15.70***	.987	.978	.062

3개의 성장모형의 합치도 지수를 비교한 결과, 2차 선형 변화 모형이 종단자료를 보다 잘 설명하고 있음을 알 수 있다.

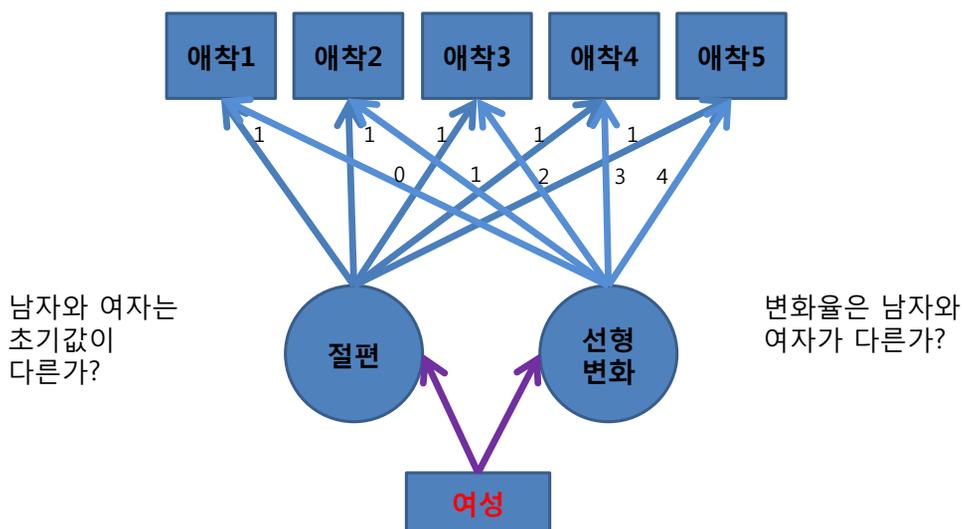
5. 잠재성장모형

Step 3. 모형 결과 보고
선형변화모형의 예



5. 잠재성장모형

Conditional LGM



5. 잠재성장모형

Conditional LGM

M-plus 프로그램

```
TITLE: Unconditional LGM
DATA: FILE IS ATTACH.DAT;
VARIABLE: NAMES ARE ID FEMALE ATT1 ATT2 ATT3 ATT4 ATT5;
USEVARIABLE ARE ATT1 ATT2 ATT3 ATT4 ATT5 FEMALE;
MISSING = ALL(99);

ANALYSIS:
MODEL:
I S | ATT1@0 ATT2@1 ATT3@2 ATT4@3 ATT5@4;
I ON FEMALE; S ON FEMALE;
OUTPUT:
MODINDICES STANDARDIZED;
```

성별이 초기값과 선형변화율에 미치는 영향력을 검증한다.

5. 잠재성장모형

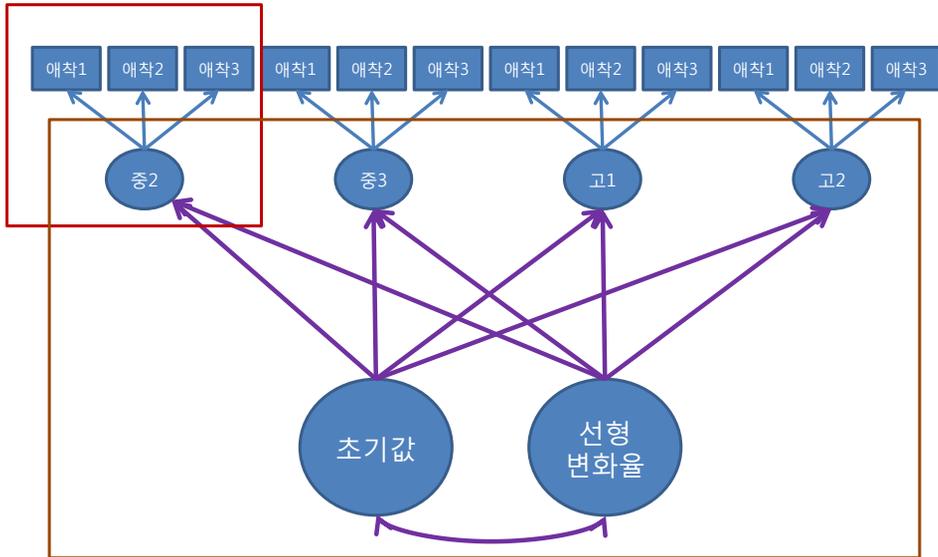
2nd-order LGM

여러 개의 측정변수를 활용한 잠재성장모형

- 한개의 측정변수가 사용되는 경우는, 관찰변수가 오차가 없이 측정되었다고 가정 (비현실적)
- 여러 개의 측정변수(multiple indicators)를 사용함으로써 고유분산을 제거하고 공통요인만을 가지고 성장모형을 설정할 수 있음.
- Type II errors의 가능성을 줄인다.
- 종단 자료를 사용하여 시점 간의 측정모형의 동일성을 검증할 수 있다.
- 고유분산간의 복잡한 관계를 가정할 수 있다.

5. 잠재성장모형

2nd-order LGM

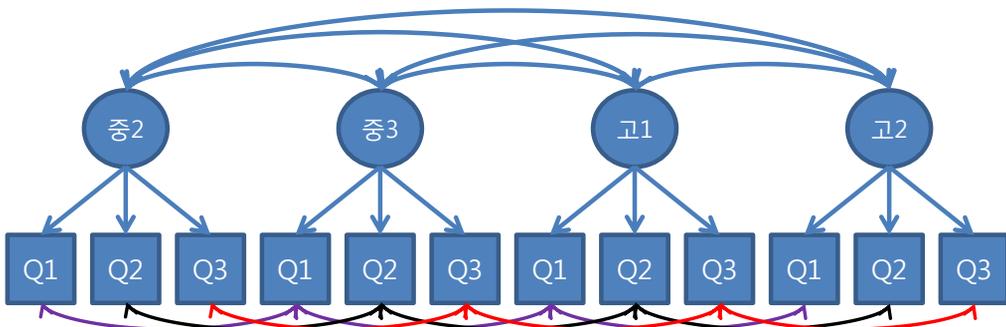


5. 잠재성장모형

M-plus 화일(예제, 무변화 2nd-order LGM)

측정단위 동일성

- 요인계수(Factor loadings)의 값이 각 시점 간의 동일하다고 가정함.



5. 잠재성장모형

M-plus 화일(예제, 무변화 2nd-order LGM)

측정단위 동일성

TITLE:

DATA:

FILE IS ATTACH_3ITEMS.DAT;

VARIABLE: NAMES ARE

ID FEMALE

A11 A21 A31

A12 A22 A32

A13 A23 A33

A14 A24 A34

A15 A25 A35;

USEVARIABLE ARE

A11 A21 A31

A12 A22 A32

A13 A23 A33

A14 A24 A34;

MISSING = ALL(99);

ANALYSIS:

MODEL:

ATT1 BY A11 (1)

A21 (2)

A31 (3);

ATT2 BY A12 (1)

A22 (2)

A32 (3);

ATT3 BY A13 (1)

A23 (2)

A33 (3);

ATT4 BY A14 (1)

A24 (2)

A34 (3);

A11 WITH A12; A12 WITH A13; A13 WITH A14;

A21 WITH A22; A22 WITH A23; A23 WITH A24;

A31 WITH A32; A32 WITH A33; A33 WITH A34;

OUTPUT:

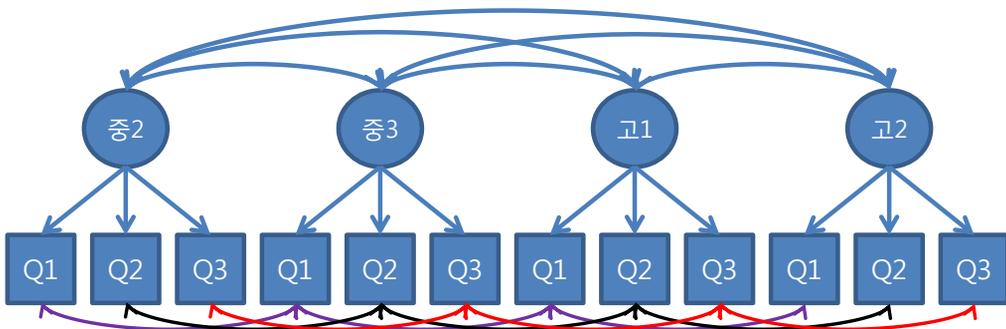
MODINDICES STANDARDIZED TECH1 ;

5. 잠재성장모형

M-plus 화일(예제, 무변화 2nd-order LGM)

측정원점 동일성(기저모형)

- 측정변수의 절편이 측정 시점 간의 동일하다고 가정한다.



5. 잠재성장모형

M-plus 화일(예제, 무변화 2nd-order LGM)

측정원점 동일성

- 측정변수의 절편이 측정 시점 간의 동일하다고 가정한다.

TITLE:

DATA:

FILE IS ATTACH_3ITEMS.DAT;

VARIABLE: NAMES ARE

ID FEMALE

A11 A21 A31

A12 A22 A32

A13 A23 A33

A14 A24 A34

A15 A25 A35;

USEVARIABLE ARE

A11 A21 A31

A12 A22 A32

A13 A23 A33

A14 A24 A34;

MISSING = ALL(99);

ANALYSIS:

MODEL:

ATT1 BY A11 (1) A21 (2) A31 (3);

ATT2 BY A12 (1) A22 (2) A32 (3);

ATT3 BY A13 (1) A23 (2) A33 (3);

ATT4 BY A14 (1) A24 (2) A34 (3);

A11 WITH A12; A12 WITH A13; A13 WITH A14;

A21 WITH A22; A22 WITH A23; A23 WITH A24;

A31 WITH A32; A32 WITH A33; A33 WITH A34;

[A11] (11); [A12] (11); [A13] (11); [A14] (11);

[A21] (12); [A22] (12); [A23] (12); [A24] (12);

[A31] (13); [A32] (13); [A33] (13); [A34] (13);

OUTPUT:

MODINDICES STANDARDIZED TECH1 ;

127/127

5. 잠재성장모형

M-plus 화일(예제, 무변화 2nd-order LGM)

모형 비교

표1. 모형의 합치도 지수

모형	χ^2	df	Δdf	$\Delta\chi^2$	CFI	TLI	RMSEA
모형 1	93.69	39			.980	.965	.051
모형 2	106.09	45	6	12.40	.977	.967	.050
모형 3	120.35	54	9	14.26	.975	.970	.048

주. 모형 1은 측정틀 동일성을 가정함. 모형 2는 측정단위 동일성을 가정함. 모형 3은 측정원점 동일성을 가정함.

모형들 간의 합치도 지수를 비교한 결과 측정원점 동일성 가정이 지지됨.

128/127

5. 잠재성장모형

M-plus 화일(예제, 무변화 2nd-order LGM)

무변화 모형

TITLE:

DATA:

FILE IS ATTACH_3ITEMS.DAT;

VARIABLE: NAMES ARE

ID FEMALE

A11 A21 A31

A12 A22 A32

A13 A23 A33

A14 A24 A34

A15 A25 A35;

USEVARIABLE ARE

A11 A21 A31

A12 A22 A32

A13 A23 A33

A14 A24 A34;

MISSING = ALL(99);

ANALYSIS:

MODEL:

ATT1 BY A11 (1) A21 (2) A31 (3);

ATT2 BY A12 (1) A22 (2) A32 (3);

ATT3 BY A13 (1) A23 (2) A33 (3);

ATT4 BY A14 (1) A24 (2) A34 (3);

A11 WITH A12; A12 WITH A13; A13 WITH A14;

A21 WITH A22; A22 WITH A23; A23 WITH A24;

A31 WITH A32; A32 WITH A33; A33 WITH A34;

[A11] (11); [A12] (11); [A13] (11); [A14] (11);

[A21] (12); [A22] (12); [A23] (12); [A24] (12);

[A31] (13); [A32] (13); [A33] (13); [A34] (13);

I | ATT1@0 ATT2@1 ATT3@2 ATT4@3;

OUTPUT:

MODINDICES STANDARDIZED TECH1 ;

5. 잠재성장모형

M-plus 화일(예제, 2nd-order LGM)

모형비교

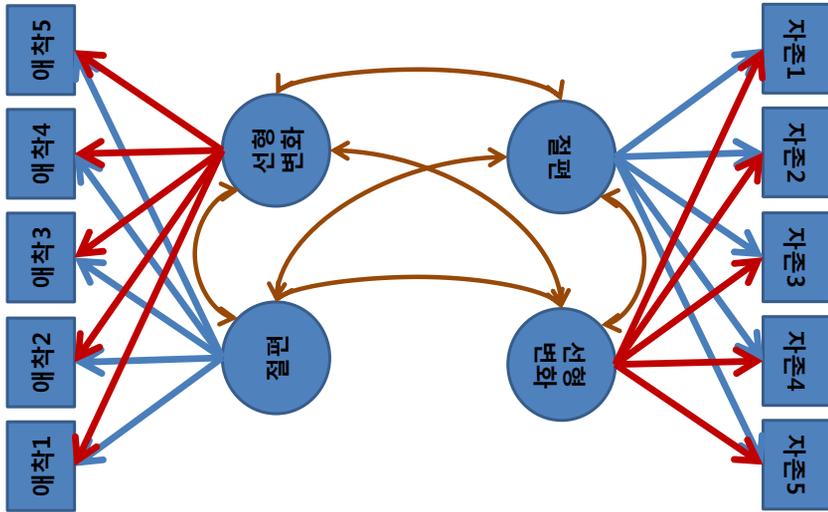
표1. 모형의 합치도 지수

모형	χ^2	df	Δdf	$\Delta\chi^2$	CFI	TLI	RMSEA
무변화	141.95	59			.969	.965	.051
선형 변화	125.74	56	3	16.21	.974	.969	.048
2차 변화	116.52	52	4	9.22	.976	.969	.048

5. 잠재성장모형

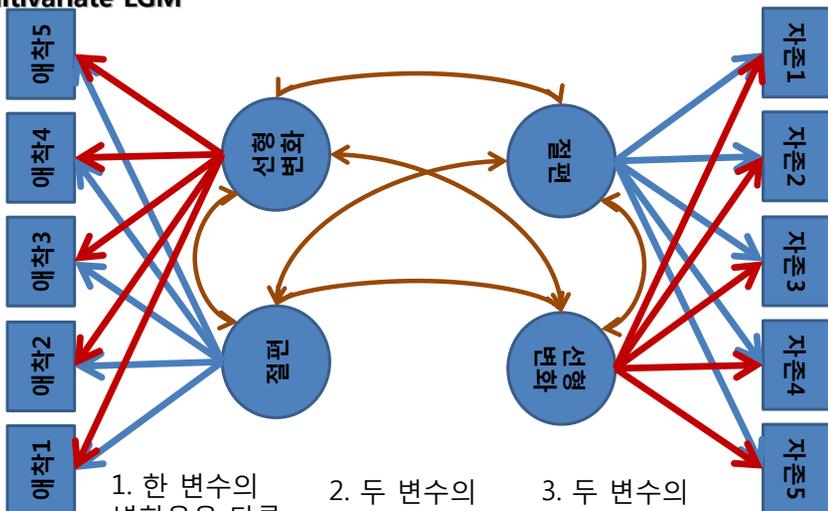
Multivariate LGM

두 변수의 성장모형 간의 관련성의 연구가 주 목적임.



5. 잠재성장모형

Multivariate LGM



1. 한 변수의 변화율을 다른 변수의 변화율과 관련이 있는가?
2. 두 변수의 초기값들은 서로 연관되어 있는가?
3. 두 변수의 초기값들은 변화율과 관련이 있는가?

5. 잠재성장모형

Multivariate LGM
Mplus input 화일

자료파일: M_lgm.dat

```
TITLE: Unconditional LGM
DATA: FILE IS M_LGM.DAT;
VARIABLE: NAMES ARE
        ATT1 ATT2 ATT3
        ATT4 ATT5
        EST1 EST2 EST3
        EST4 EST5;
USEVARIABLE ARE
        ATT1 ATT2 ATT3
        ATT4 ATT5
        EST1 EST2 EST3
        EST4 EST5;
MISSING = ALL(99);

ANALYSIS:
MODEL:
I_A S_A | ATT1@0 ATT2@1 ATT3@2 ATT4@3 ATT5@4;
I_E S_E | EST1@0 EST2@1 EST3@2 EST4@3 EST5@4;

OUTPUT:
MODINDICES STANDARDIZED TECH1 ;
```

5. 잠재성장모형

Multivariate LGM
Mplus input 화일

자료파일: M_lgm.dat

MODEL:

I_A S_A | ATT1@0 ATT2@1 ATT3@2 ATT4@3 ATT5@4;

I_E S_E | EST1@0 EST2@1 EST3@2 EST4@3 EST5@4;

OUTPUT:
MODINDICES STANDARDIZED TECH1 ;

5. 잠재성장모형

Multivariate LGM

Mplus 결과 화일

S_A	WITH				
I_A		-0.394	0.089	-4.423	0.000
I_E	WITH				
I_A		0.002	0.081	0.023	0.982
S_A		-0.073	0.139	-0.522	0.602
S_E	WITH				
I_A		0.066	0.153	0.432	0.666
S_A		-0.111	0.263	-0.422	0.673
I_E		-0.359	0.175	-2.055	0.040

135/127

감사합니다

조영일, Ph.D

136

■ 2014년 한국아동·청소년패널조사 V 제1차 콜로키움 자료집
Mplus를 활용한 자료분석

인 쇄 2014년 6월 30일

발 행 2014년 7월 1일

발행처 한국청소년정책연구원
서울특별시 서초구 태봉로 114

발행인 이 재 연

등 록 1993. 10. 23. 제 21-500호

인쇄처 문영사 전화 02)2263-5087 김희자

사전 승인없이 보고서 내용의 무단전재·복제를 금함.
구독문의 : (02) 2188-8844(정보자료·전산보안팀)