

Example 1 - 測定モデル (単一因子) 【評価版】

本 whitepaper は SEM を単一因子の測定モデルに適用したときの用例を紹介するものです。

1. データセット
 2. SEM モデル
 3. sem によるフィット
 4. Satorra-Bentler χ^2 検定
 5. gsem によるフィット
 6. 測定誤差モデル
- 補足 1

1. データセット

[SEM] Example 1 には Example データセット sem_1fmm.dta を用いた SEM の用例が紹介されています。

```
. use http://www.stata-press.com/data/r17/sem_1fmm.dta *1
(single-factor measurement model)
```

このデータセット中には潜在変数 X の測定値を表す変数 x_1, x_2, x_3, x_4 の値が 500 件記録されています。参考までに先頭 5 件のデータをリスト出力しておくと次のようになります。

```
. list in 1/5 *2
```

	x1	x2	x3	x4
1.	96	82	96	43
2.	106	81	93	40
3.	127	95	96	134
4.	123	98	94	109
5.	119	99	100	108

© Copyright Math 工房； 一部 © Copyright StataCorp LP (used with permission)

*1 メニュー操作：File▷Example Datasets▷Stata 17 manual datasets と操作、Structural Equation Modeling Reference Manual [SEM] の Example 1 の項よりダウンロードする。

*2 メニュー操作：Data▷Describe data▷List data

x_4 のみ平均値や標準偏差が他と異なっている点に注意してください。

```
. summarize *3
```

Variable	Obs	Mean	Std. dev.	Min	Max
x1	500	99.518	14.35402	60	137
x2	500	99.954	14.1939	52	140
x3	500	99.052	14.26395	59	150
x4	500	94.474	70.11603	-113	295

このデータセットには注釈 (notes) が付帯しています。その内容は次のようにして確認することができます。

```
. notes list *4
```


. notes list	
_dta:	
1.	fictional data
2.	Variables x1, x2, x3, and x4 each contain a test score designed to measure X. The test is scored to have mean 100.


要するに

- データは人為的に作成されたものであること。
- 変数 x_1, x_2, x_3, x_4 はいずれも潜在変数 X の値を計測すべく実施したテストのスコアであり、いずれのテストも平均値が 100 となるよう企図されている。

という趣旨のことが記されています。

2. SEM モデル

測定モデルを SEM Builder 上で設定するには  アイコンを使うのが便利です。

- Statistics ▷ SEM (structural equation modeling) ▷ Model building and estimation と操作
- SEM Builder 上の  アイコンを選択した後、描画領域内の任意の点をクリック、表示される Measurement component ダイアログ上で次の指定を行う
- Main タブ: Latent variable name: X
Measurement variables: x1 x2 x3 x4

*3 メニュー操作 : Statistics ▷ Summaries, tables and tests ▷ Summary and descriptive statistics ▷ Summary statistics

*4 メニュー操作 : Data ▷ Data utilities ▷ Notes utilities ▷ List or search notes

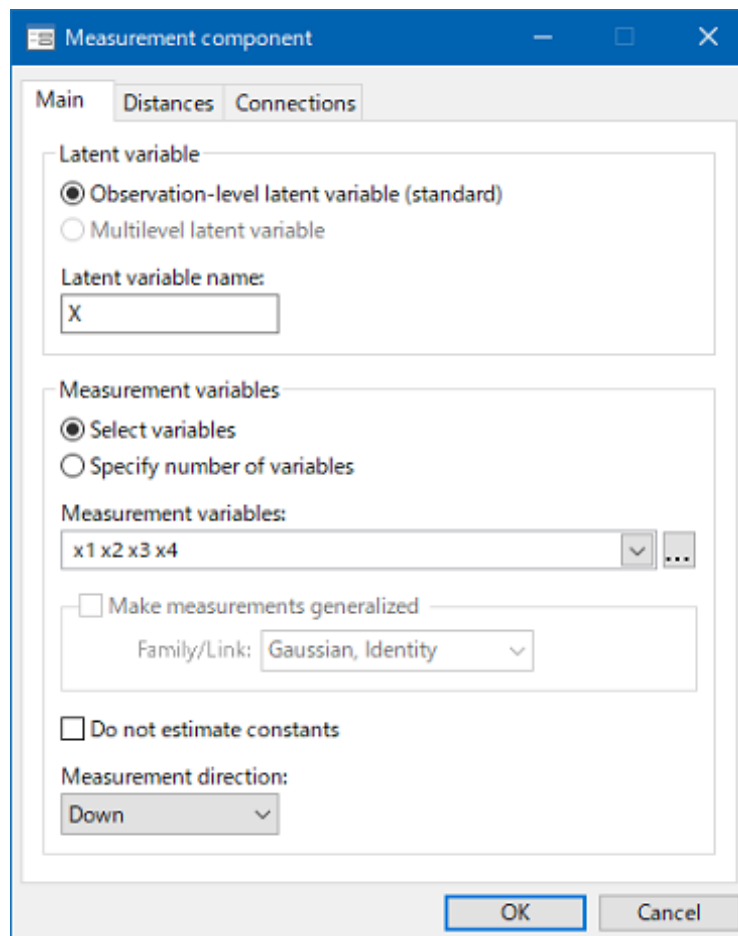


図1 Measurement component ダイアログ - Main タブ

SEM Builder 上の描画領域内には次のようなモデルが構成されます。

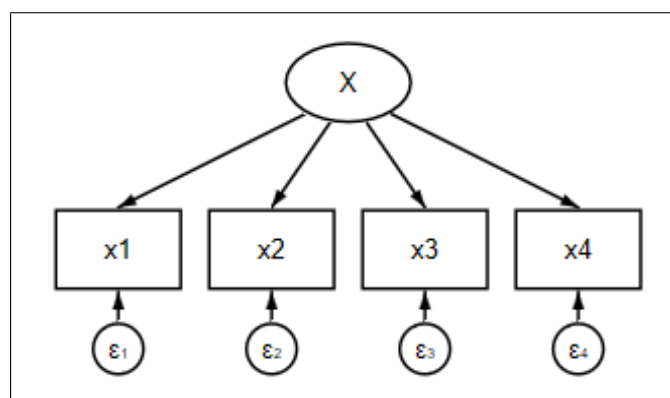





図2 測定モデル

 英文マニュアル上のモデルダイアグラムでは誤差変数からのパス係数が 1 と明示されています。しかし上記の操作のように内生変数に付帯する形で誤差変数 ε を自動生成した場合、そのパス係数は 1 と暗黙のうちに仮定されるため、明示の必要はありません。

  アイコンをクリックした上でモデルダイアグラム全体を選択すると、オブジェクトを描画領域内の任意の位置に移動させることができます。


コマンドインタフェースでモデルを規定する場合には次のいずれを用いても構いません。

(x1<-X) (x2<-X) (x3<-X) (x4<-X)

(x1 x2 x3 x4<-X)

(X->x1 x2 x3 x4)

3. sem によるフィット

モデルの推定を実行するには SEM Builder 上の  アイコンをクリックします。SEM estimation options ダイアログが表示されますが、デフォルト設定のまま OK をクリックします。

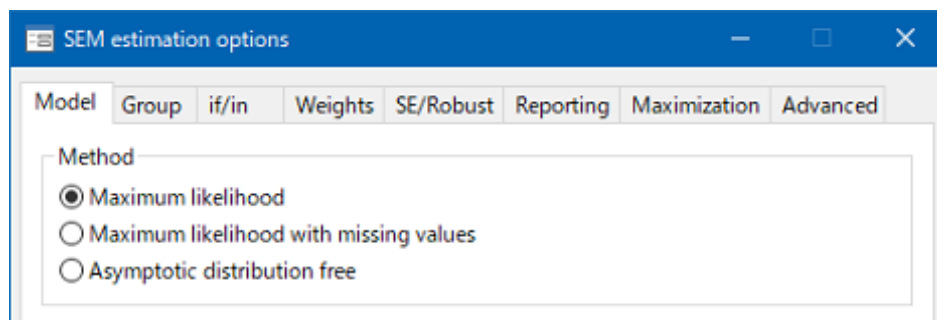


図 3 SEM estimation options ダイアログ - Model タブ

```

. sem (X -> x1, ) (X -> x2, ) (X -> x3, ) (X -> x4, ), latent(X ) nocapslatent

```

Endogenous variables
Measurement: x1 x2 x3 x4

Exogenous variables
Latent: X

Fitting target model:
Iteration 0: log likelihood = -8487.5905
Iteration 1: log likelihood = -8487.2358
Iteration 2: log likelihood = -8487.2337
Iteration 3: log likelihood = -8487.2337

Structural equation model Number of obs = 500
Estimation method: ml

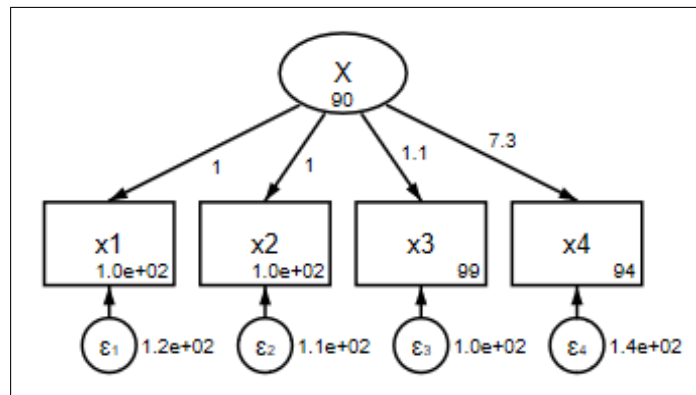
Log likelihood = -8487.2337

(1) [x1]X = 1

	OIM					[95% conf. interval]	
	Coefficient	std. err.	z	P> z			
Measurement							
x1							
X	1	(constrained)					
_cons	99.518	.6412888	155.18	0.000	98.2611	100.7749	
x2							
X	1.033249	.0723898	14.27	0.000	.8913676	1.17513	
_cons	99.954	.6341354	157.62	0.000	98.71112	101.1969	
x3							
X	1.063876	.0729725	14.58	0.000	.9208526	1.2069	
_cons	99.052	.6372649	155.43	0.000	97.80298	100.301	
x4							
X	7.276754	.4277638	17.01	0.000	6.438353	8.115156	
_cons	94.474	3.132547	30.16	0.000	88.33432	100.6137	
var(e.x1)	115.6865	7.790423			101.3823	132.0089	
var(e.x2)	105.0445	7.38755			91.51873	120.5692	
var(e.x3)	101.2572	7.17635			88.12499	116.3463	
var(e.x4)	144.0406	145.2887			19.94838	1040.069	
var(X)	89.93921	11.07933			70.64676	114.5001	

LR test of model vs. saturated: chi2(2) = 1.46 Prob > chi2 = 0.4827

結果の一部はモデルダイアグラム上にも表示されます。



なお、このモデルのフィットをコマンドインタフェースで実行する場合には

```
. sem (x1 x2 x3 x4 <- X)
```

の入力します。

図 2 に示された測定モデルは数式で次のように表現できます。

$$\begin{aligned}
 x_1 &= \alpha_1 + X\beta_1 + \varepsilon_1 \\
 x_2 &= \alpha_2 + X\beta_2 + \varepsilon_2 \\
 x_3 &= \alpha_3 + X\beta_3 + \varepsilon_3 \\
 x_4 &= \alpha_4 + X\beta_4 + \varepsilon_4
 \end{aligned}
 \tag{1}$$

実行結果に関しては次の点に注意してください。

1. 変数 X は潜在外生変数であるため、正規化 (normalization) という形の制約が自動的に課されます。変数 X は第 1 の観測可能変数 x_1 にアンカーされている (anchored) わけですが、それに対するパス係数 β_1 は 1 に設定されています。詳しくは [SEM] Intro 4 (*mwp-122*) をご参照ください。
2. $X \rightarrow x_1$, $X \rightarrow x_2$, $X \rightarrow x_3$ へのパス係数値、すなわち $\beta_1, \beta_2, \beta_3$ はそれぞれ 1, 1.03, 1.06 と推定されています。一方、 $X \rightarrow x_4$ へのパス係数値 β_4 は 7.28 と推定されています。このデータセットは元々擬似乱数を発生させる形で用意されたものであり、本来の係数値としてはそれぞれ 1, 1, 1, 7 が想定されていたことを考えると、推定結果は順当なものと言えます。
3. 出力末尾に $\chi^2(2)$ 統計量が 1.46 で有意水準が 0.4827 という検定結果が表示されています。 p 値が $\gg 0.05$ であるため、帰無仮説は棄却できないわけです。この検定は badness-of-fit を単位とした適合度検定 (goodness-of-fit test) の一種ですが、結果が有意と判定された場合にはモデルのフィットが十分とは言えないことになります。
より数学的な表現を使用するなら、この検定の帰無仮説は観測可能変数に対するフィットされた共分散行列と平均値ベクトルが、母集団中で観測されるであろう行列とベクトルに等しいとするものです。

4. Satorra-Bentler χ^2 検定

評価版では割愛しています。

5. gsem によるフィット

評価版では割愛しています。

6. 測定誤差モデル

評価版では割愛しています。

補足 1 - シミュレーション

評価版では割愛しています。

