

驗證性因素分析(CFA)-觀念介紹

By
劉介宇

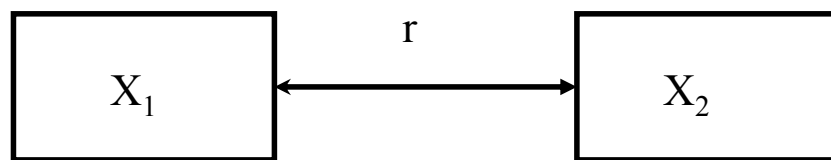
1

探索性與驗證性因素分析的差別

- 做探索性因素分析時，研究者對因素的結構未知或只有極少的理論根據。因此研究者收集資料，探求或尋找能解釋變間相關的因素結構或理論。此稱為探索性因素分析(**exploratory factor analysis**)。
- 驗證性因素分析，則假設因素結構已知或已有假設性的先驗理論。**驗證性因素分析(confirmatory factor analysis)**的目的是要證實或確認這樣的因素結構吻合手上蒐集的資料。

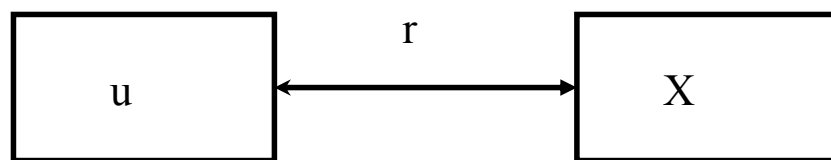
- 結構方程模式（Structural Equation Modeling, SEM）
- 潛在變數模式（Latent Variable Models, LVM）
- 線性結構關係模式（Linear Structural Relationship Model, LISREL）
- 共變數結構分析（Covariance Structure Analysis）
- 驗證性因素分析（Confirmatory Factor Analysis, CFA）

基本概念 從相關分析談起



相關分析（Correlational Analysis）

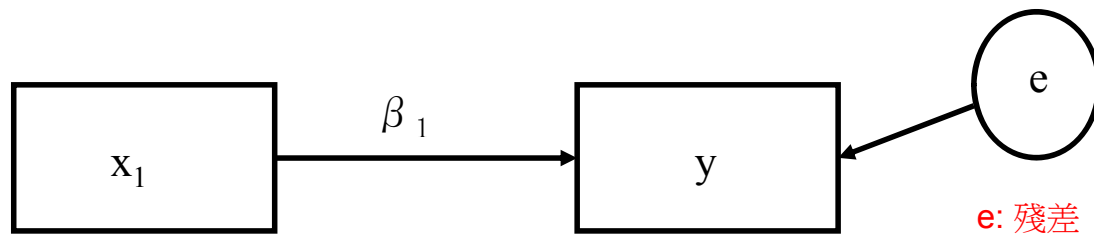
5



點二系列相關（ Point-biserial Correlation Coefficient ）

常用在成就測驗的項目分析，以進行各子題與總分之間的一致性分析

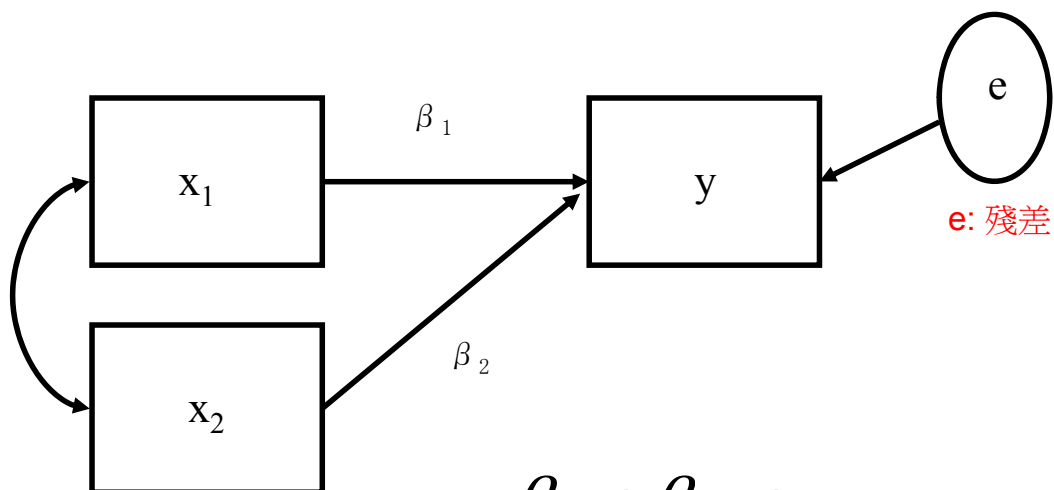
6



$$y = \beta_1 x + e$$

簡單線性迴歸 (Simple Linear Regression)

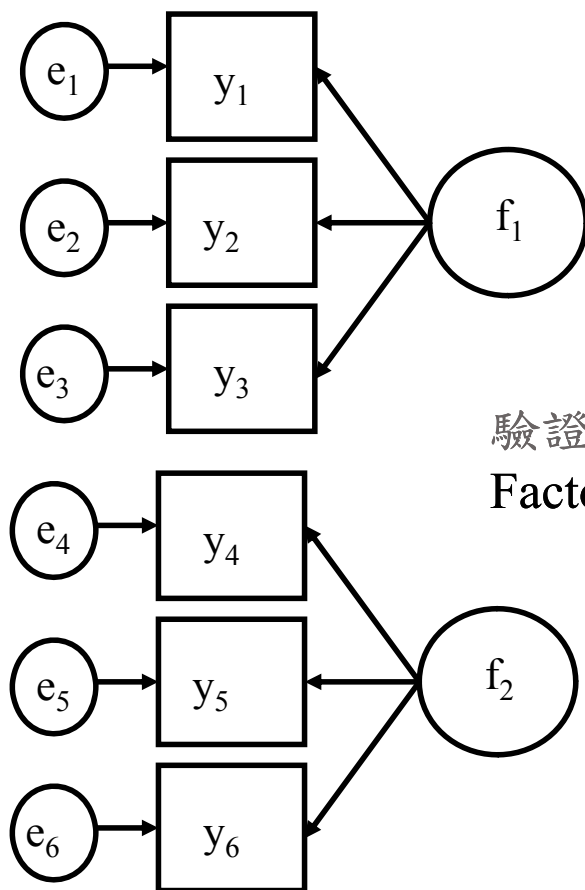
7



$$y = \beta_1 x_1 + \beta_2 x_2 + e$$

多元線性迴歸 (Multiple Linear Regression)

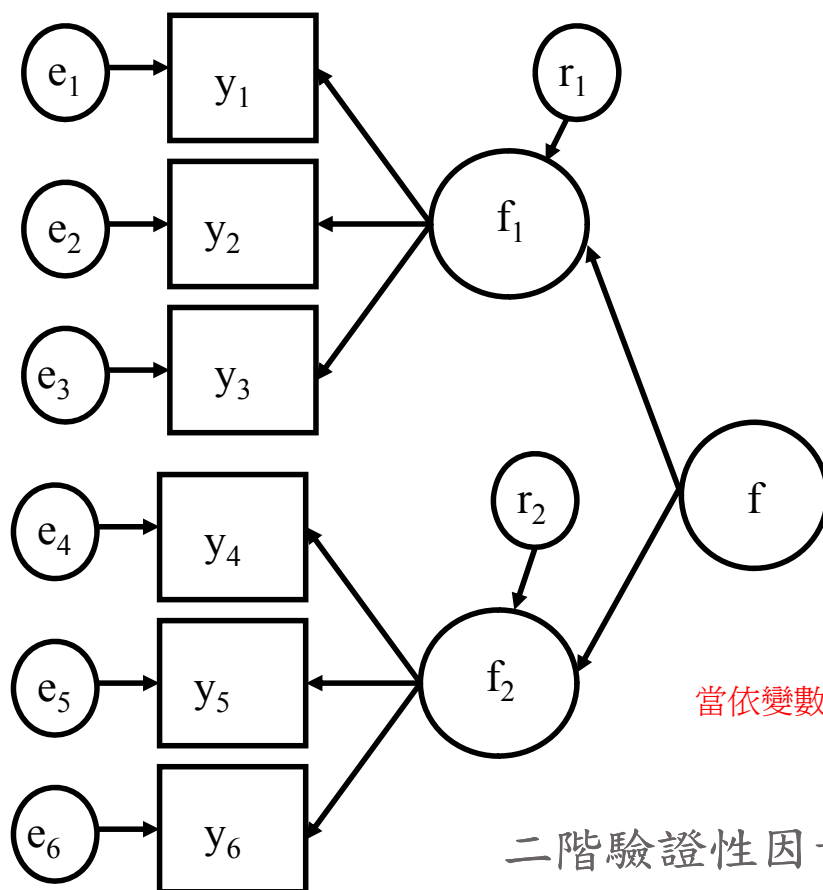
8



驗證性因素分析(Confirmatory Factor Analysis)

當依變數的才有殘差!!

9

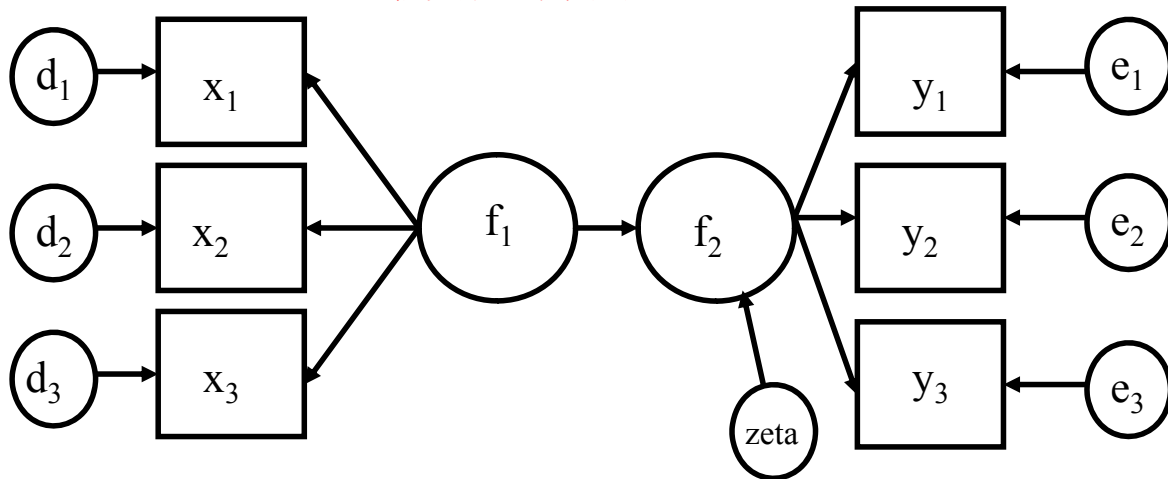


當依變數的才有殘差!!

二階驗證性因素分析 (2nd order Confirmatory Factor Analysis)

10

當依變數的才有殘差!!



結構方程模式(Structural Equation Model, SEM)的general case
潛在變項路徑分析(Path Analysis with Latent Variables, PA-LV)

基本概念

再談探索性因素分析

Exploratory Factor Analysis (EFA)

表 5.2 兩因素模式的共通性、型態與結構負荷及相關矩陣

共通性						
變數	共通性			總和	唯一性	
	Q	V				
M	0.640	0.040		0.680	0.320	
P	0.490	0.090		0.580	0.420	
C	0.360	0.090		0.450	0.550	
E	0.040	0.640		0.680	0.320	
H	0.023	0.672		0.695	0.305	
F	0.063	0.723		0.786	0.214	
總和	1.616	2.255		3.871	2.129	

型態及結構負荷與分享變異數						
變數	型態負荷		結構負荷		分享變異數	
	Q	V	Q	V	Q	V
M	0.800	0.200	0.800	0.200	0.640	0.040
P	0.700	0.300	0.700	0.300	0.490	0.090
C	0.600	0.300	0.600	0.300	0.360	0.090
E	0.200	0.800	0.200	0.800	0.040	0.640
H	0.150	0.820	0.150	0.820	0.023	0.672
F	0.250	0.850	0.250	0.850	0.063	0.723
總和					1.616	2.255

相關矩陣						
	M	P	C	E	H	F
M	1.000					
P	0.620	1.000				
C	0.540	0.510	1.000			
E	0.320	0.380	0.360	1.000		
H	0.284	0.351	0.336	0.686	1.000	
F	0.370	0.430	0.405	0.730	0.735	1.000

探索性因素分析

如何命名？

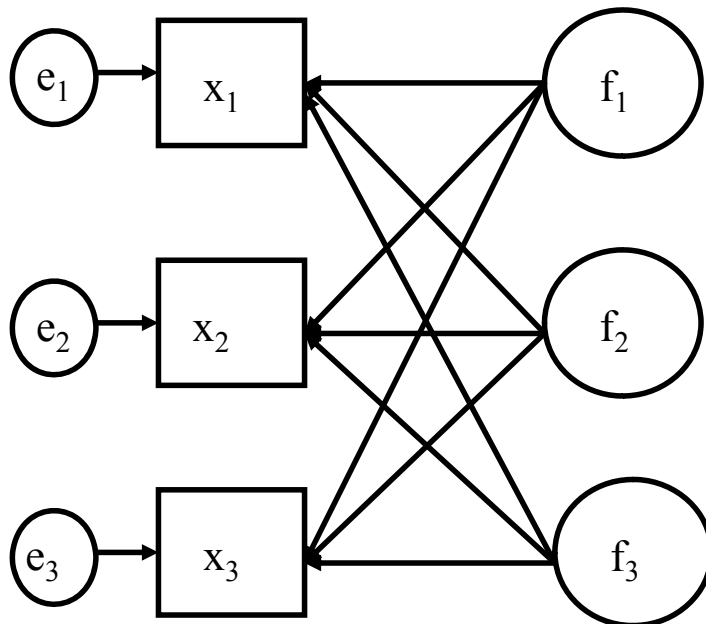
模式架構

- $x_1 = \mu_1 + l_{11}f_1 + l_{12}f_2 + \dots + l_{1q}f_q + \varepsilon_1$
- $x_2 = \mu_2 + l_{21}f_1 + l_{22}f_2 + \dots + l_{2q}f_q + \varepsilon_2$
- ...
- $x_p = \mu_p + l_{p1}f_1 + l_{p2}f_2 + \dots + l_{pq}f_q + \varepsilon_p$
- x 為觀察變項； f 為共同因素
- l 為因素負荷量； ε 為獨特因素(誤差項)

基本假定

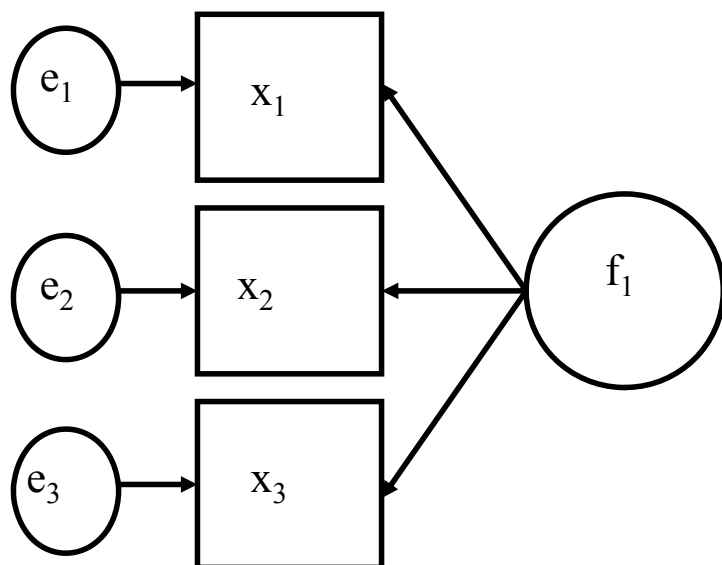
- 獨特因素互相獨立且為多變項常態分配。
- 共同因素間互相獨立且變異數為1。
- 共同因素與獨特因素互相獨立。
- **X**變項間有線性關係。

15



探索性因素分析
(Exploratory Factor
Analysis, EFA)

研究者可依某些萃取準
則，決定萃取因素數目



驗證性因素分析
(Confirmatory Factor
Analysis, CFA)

研究者依理論，指定因
素結構，獨特因素間亦
可指定相關

17

所需樣本大小

- 樣本大小不能少於**50**個，應該要超過**100**個。
- 至少要為**x**觀察變項數目的**5**倍量或**10**倍量。

Hair, J. F., Anderson, R. F., Tatham, R. L. & Black, W. C. (1998). *Multivariate data analysis 5th ed.* New Jersey: Prentice Hall Inc.

樣本大小亦取決於負荷量大小

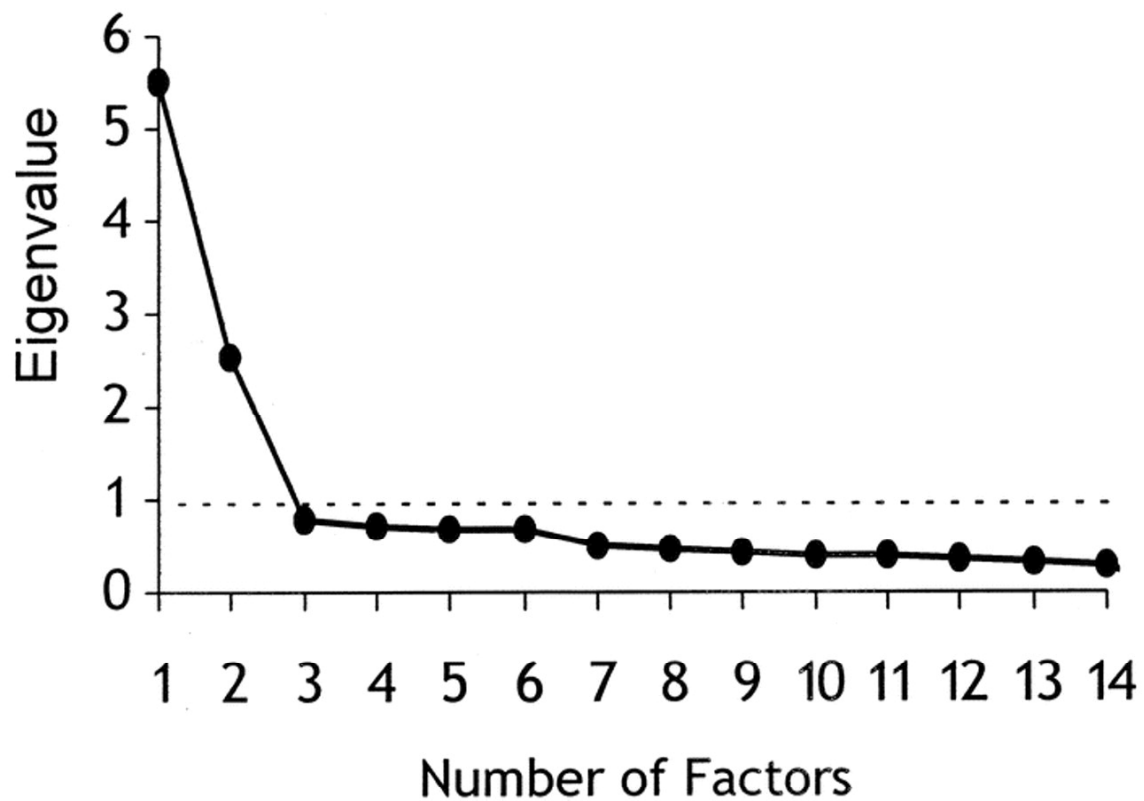
常見萃取方法（Extraction Method）

- 主要成份法（Principal components）
- 主要因素分析法（Principal factor analysis）
- 最大概似法（Maximum likelihood, ML）

x觀察變項要為連續變項，
若為二元資料（如：成就測驗）會有錯誤估計的可能

常見萃取標準（Extraction Criteria）

- 特徵值（eigenvalues）
- 先前文獻
- 累積變異量
- 陡坡圖

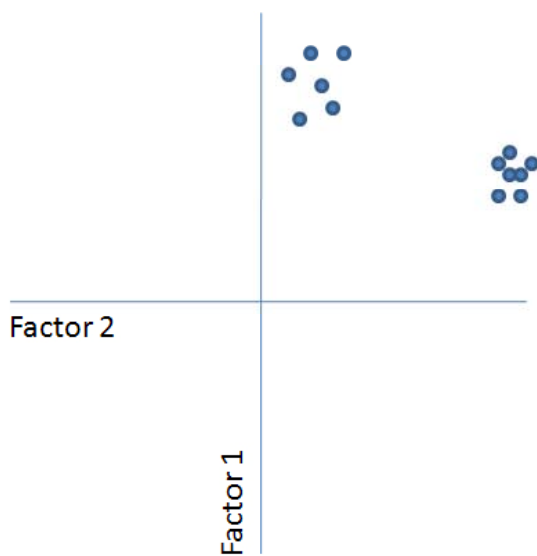


21

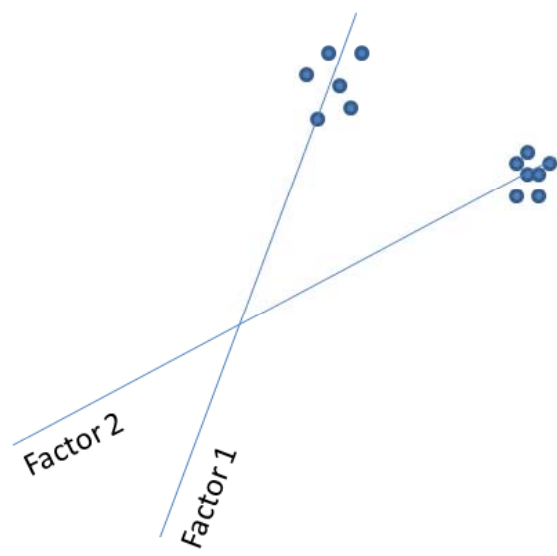
因素旋轉

- 正交旋轉(Orthogonal rotation methods)
 - Quartimax
 - Varimax
 - Equimax
- 斜交旋轉(Oblique rotation methods)
 - Oblimin
 - Promax
 - Orthoblique

22



正交旋轉



斜交旋轉

23

結構方程模式

驗證性因素分析

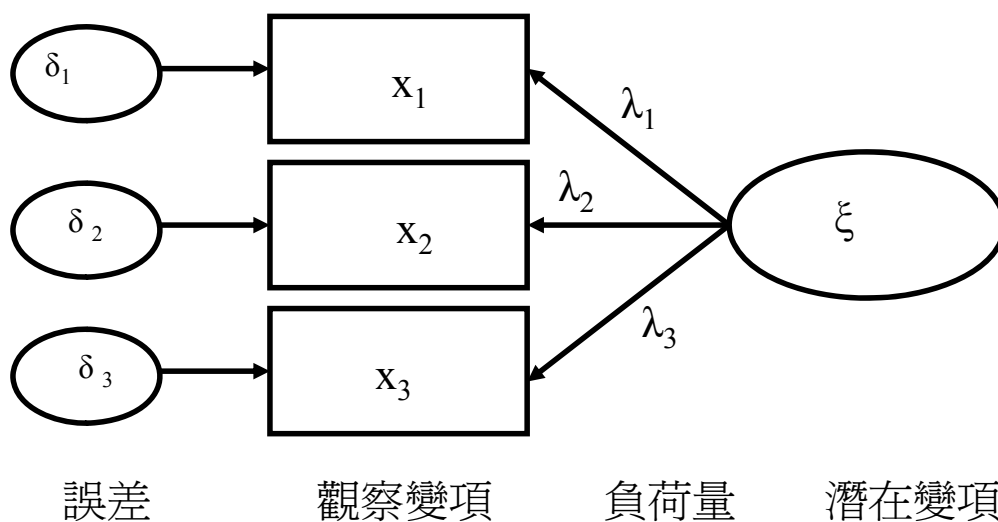
Confirmatory Factor Analysis (CFA)

24

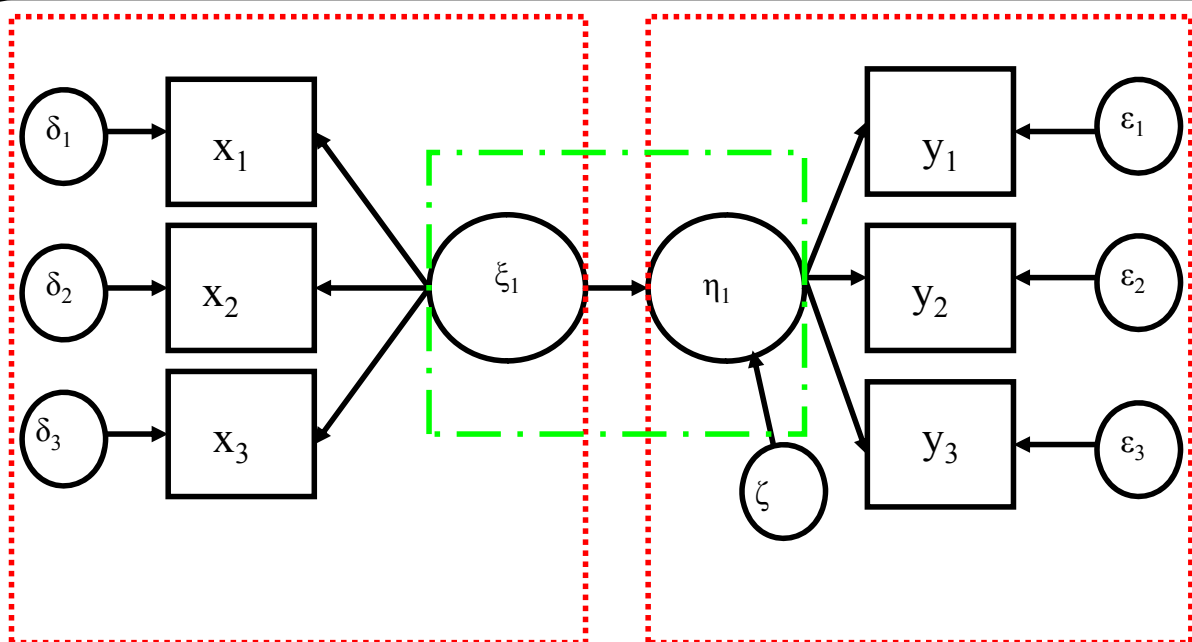
基本概念

25

符號介紹



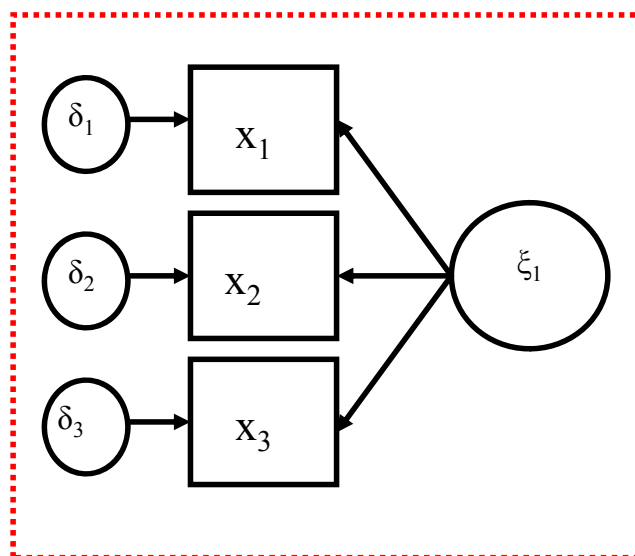
26



測量模式

結構模式

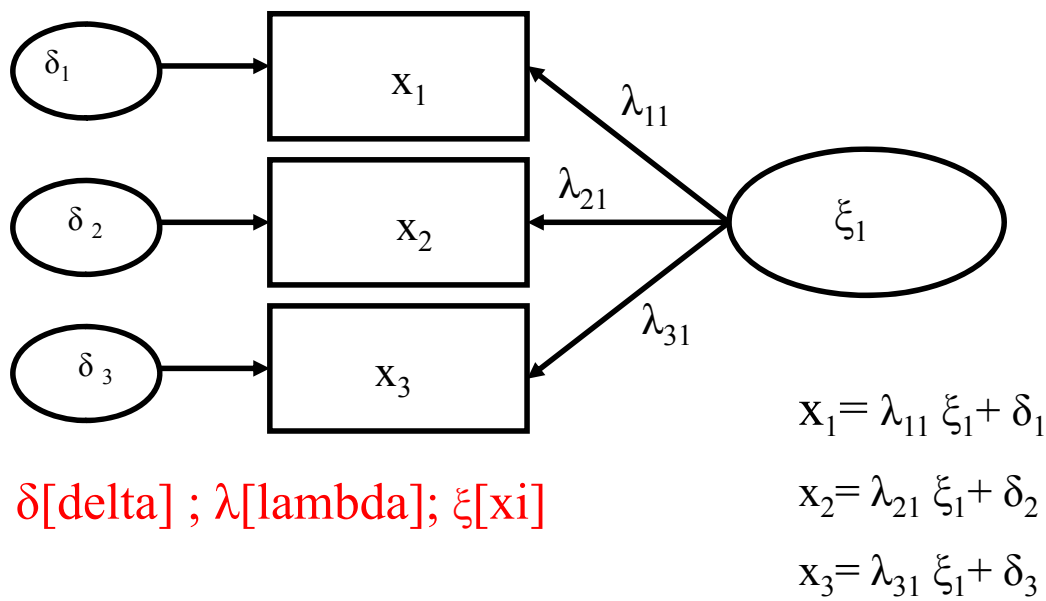
27



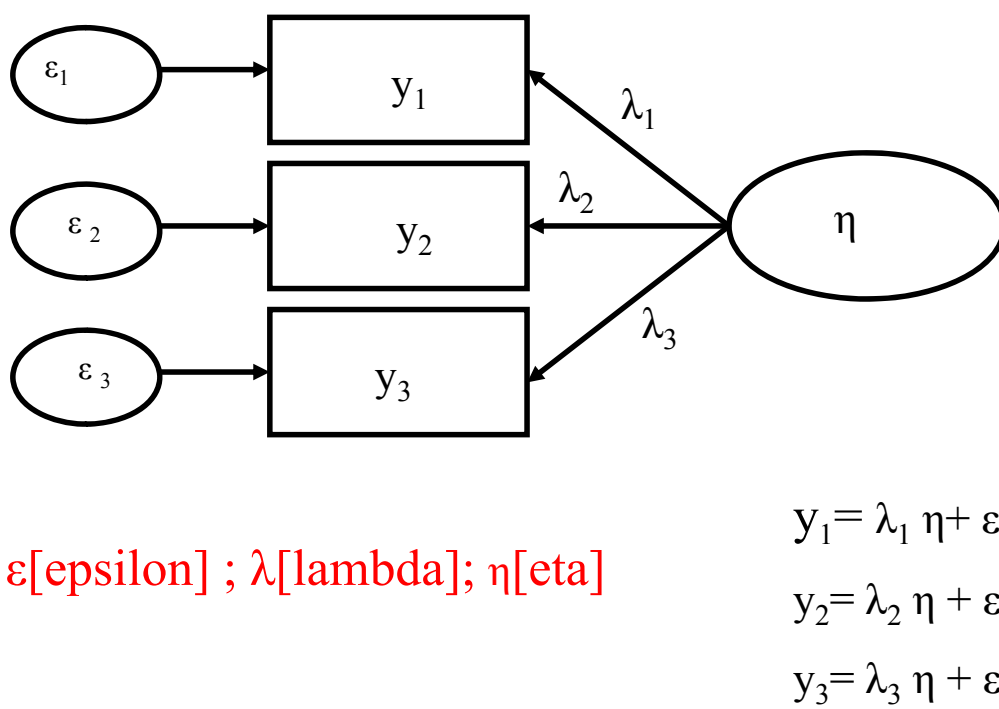
測量模式

僅有測量模式就是CFA

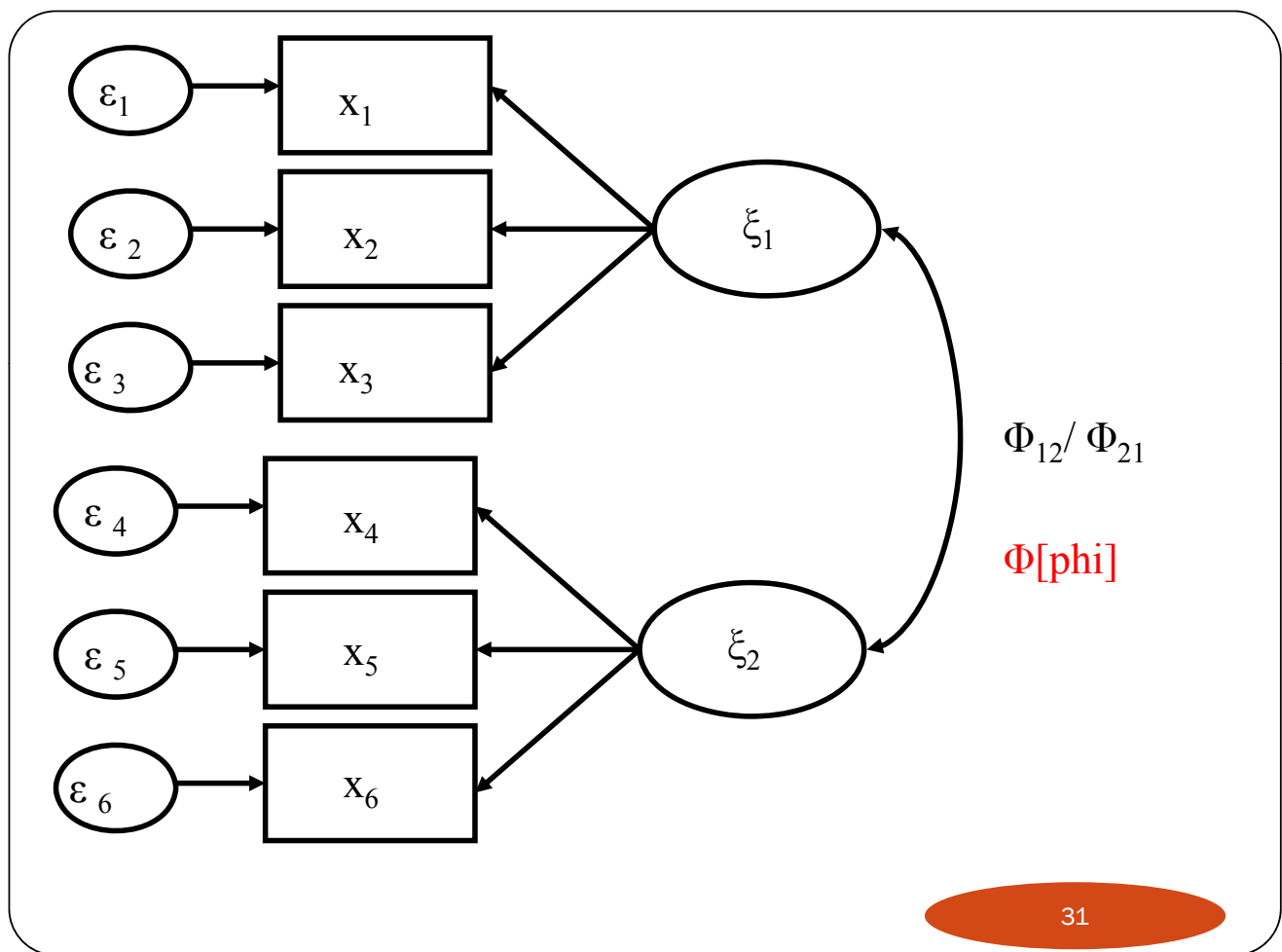
28



29



30



基本假定

- ε 與 η 、 ξ 以及 δ 相互獨立。
- δ 與 ξ 、 η 以及 ε 相互獨立。
- 樣本資料要服膺多變量常態分配。
- 觀察變項間有線性關係。

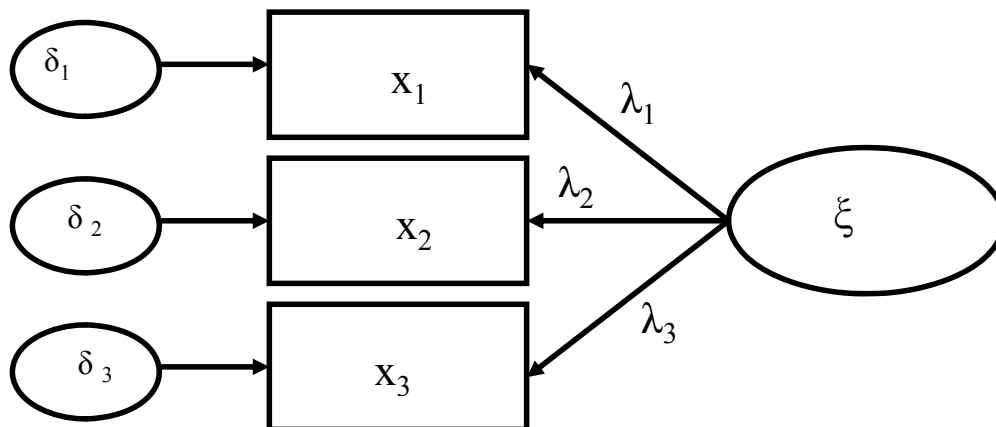
指標變項的討論

- 以觀察變項作為潛在變項的指標變項時，要幾個觀察變項才夠？
- 多元指標原則：一個潛在變項必須有兩個以上的觀察變項來估計
- 愈多愈好嗎？一個可不可以？

應回歸到工具設計與施測實務以及樣本大小、負荷量大小等問題

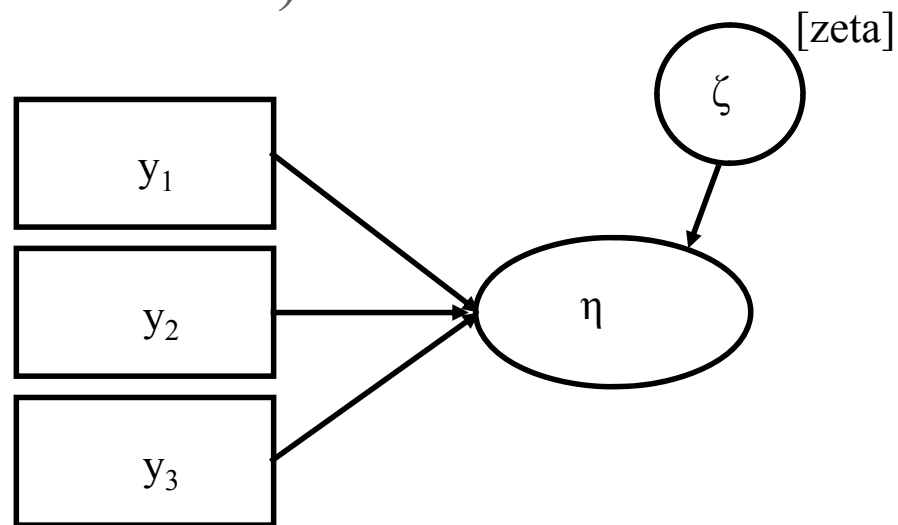
33

指標變項的討論：反應性指標 (reflective indicators)



34

指標變項的討論：形成性指標 (formative indicators)



主成分分析：多合一!! (多個變數合成一個指標)

35

樣本大小的討論

- 樣本大小至少超過150個。

Rigdon, E. (2005). *SEM FAQ*. from
<http://www.gsu.edu/~mkteer/html>

- 至少要為x觀察變項數目的10倍量或15倍量。

Thompson, B. (2000). Ten commandments of structural equation modeling. In L. G. Grimm & P. R. Yarnold (eds.), *Reading and understanding more multivariate statistics* (pp. 261-283). Washington, DC: APA.

樣本大小亦取決於潛在變項的數目

36

分析程序

- 模式概念化
- 路徑圖建構
- 模式確認
- 模式識別
- 參數估計
- 模式適配度評估
- 模式修正
- 模式複核效化 (cross-validation)

37

常見電腦軟體

- LISREL
- SIMPLIS
- AMOS
- EQS
- Mplus
- Mx
- Statistica
- SAS PROC CALIS
- COSAN
- LVPLS
- ...

38

模式確認與識別

39

路徑圖與參數設定

- 潛在變項的變異數為1，平均數為0（AMOS預設值）。
- 誤差項的路徑係數為1（AMOS預設值），或其變異數設定為1（研究者可自行更改）。
- 設定一個測量指標變項的路徑係數為1（AMOS預設值）。
- 內因變項要有誤差項（研究者自行設定）。

40

t法則(過度識別)

$$t < \frac{1}{2}(p+q)(p+q+1)$$

樣本共變矩陣
的數目

$$df = \frac{1}{2}(p+q)(p+q+1) - t, df > 0$$

- 過度識別，方可進行模式適配考驗。
- t 為待估計的自由參數。
- $p+q$ 為所有觀察變項數目。

t法則(低度識別)

$$t > \frac{1}{2}(p+q)(p+q+1)$$

$$df = \frac{1}{2}(p+q)(p+q+1) - t, df < 0$$

- 低度識別，參數估計有無限多個解。
- 減少自由參數數目（增加固定參數數目），以取得過度識別。

t法則(正好識別)

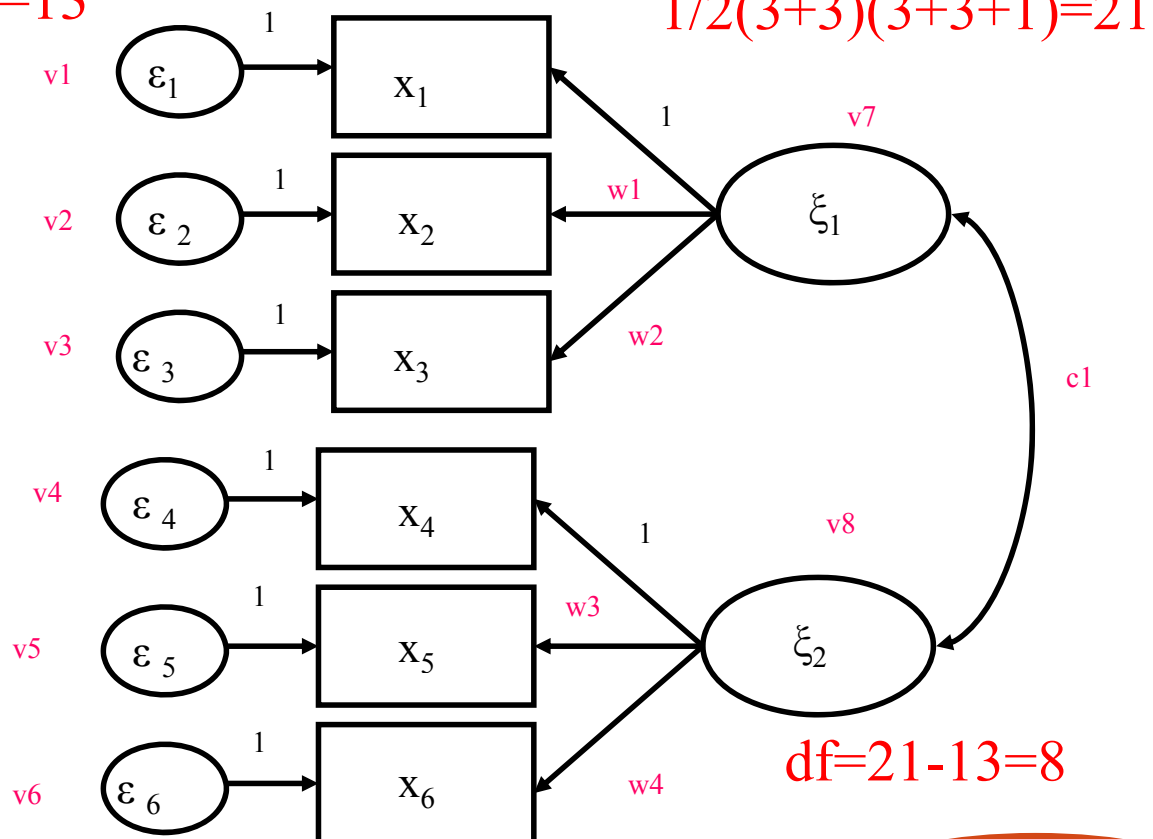
$$t = \frac{1}{2}(p+q)(p+q+1)$$

$$df = \frac{1}{2}(p+q)(p+q+1) - t, df = 0$$

- 正好識別，參數估計有唯一解，即為飽和模式。完美適配在實務不具參考價值。
- 減少自由參數數目（增加固定參數數目），以取得過度識別。

43

$t=13$



44

參數估計

45

常見參數估算方法

- 最大概似法（Maximum likelihood, ML）：樣本大於500且觀察資料為多變量常態分配。
- 一般化最小平方法（Generalized least Squares, GLS）：樣本小於500亦可；如果樣本夠大 (say >500)，觀察資料沒有服膺多變量常態分配亦可。
- 未加權最小平方法（Unweighted Least Squares, ULS）：不需符合某種分配假定。
- 貝氏估計法（Bayesian Estimation）：適用小樣本 (say <100)。

46

模式適配度檢核指標

47

模式基本適配指標

- 誤差變異數不能為負值。
- 所有誤差變異數須達顯著水準。
- 參數統計量間的相關絕對值不能太接近1。
- 因素負荷量約在0.5-0.95間。
- 標準誤不能太大。

Bagozzi, R. P. & Yi, Y. (1988). On the evaluation of structural equation models. *Academic of Marketing Science*, 16, 76-94.

48

整體模式適配度指標：外在品質評估

絕對適配度指標	標準
χ^2	$> \alpha$
χ^2 / df (NC)	1~3
GFI	> 0.9
AGFI	> 0.9
RMR	< 0.05
SRMR(AMOS須另外計算)	< 0.05
RMSEA	< 0.05
NCP	愈小愈好，信賴區間含0
ECVI	用於不同模式的相對比較，相對小者較好 理論模式比飽和模式與獨立模式的值還小

49

整體模式適配度指標：外在品質評估

增值適配度指數	標準
NFI	> 0.9
RFI	> 0.9
IFI	> 0.9
TLI/NNFI	> 0.9
CFI	> 0.9

Diamantopoulos, A. & Siguaw, J. A. (2000). Introducing LISREL: A guide for the uninitiated. Thousand Oaks, CA: Sage.

50

整體模式適配度指標：外在品質評估

簡約適配度指數	標準
PGFI	> 0.5
PNFI	> 0.5
CN	> 200
AIC	用於不同模式的相對比較，相對小者較好 理論模式比飽和模式與獨立模式的值還小
CAIC	用於不同模式的相對比較，相對小者較好 理論模式比飽和模式與獨立模式的值還小

51

模式內在結構適配度評估（內在品質評估）

- 除整體模式適配之外，深入探討每一個參數，對理論的驗證更能獲得保障。
- 測量模式的評鑑：觀察變項與潛在變項的關聯，即潛在變項的信度與效度考驗。
- 結構模式的評鑑：確認外衍與內衍變項的解釋或預測關係是否成立。

Bagozzi, R. P. & Yi, Y. (1988). On the evaluation of structural equation models. *Academic of Marketing Science*, 16, 76-94.

測量模式的評鑑

- 個別觀察變項的項目信度（individual item reliability）在0.5以上，即因素負荷量的平方值。
- 潛在變項的組合信度（composite reliability）在0.6以上。

$$\rho_c = \frac{(\sum \lambda)^2}{[(\sum \lambda)^2 + \sum \theta]}$$

因素負荷量

觀察變項的誤差

測量模式的評鑑

- 潛在變項的平均變異數萃取量（average variance extracted）在0.5以上。

$$\rho_v = \frac{(\sum \lambda^2)}{[(\sum \lambda^2) + \sum \theta]}$$

測量模式的評鑑

- 參數統計量的估計值達顯著水準（**t or CR**大於**1.96**）。
- 標準化殘差的絕對值小於**2.58/3**（可於**AMOS**的**Residual Covariance**矩陣獲得）。
- 修正指標（**modification indices**）小於**3.84**（**AMOS**的預設值為**4**）。

結構模式的評鑑

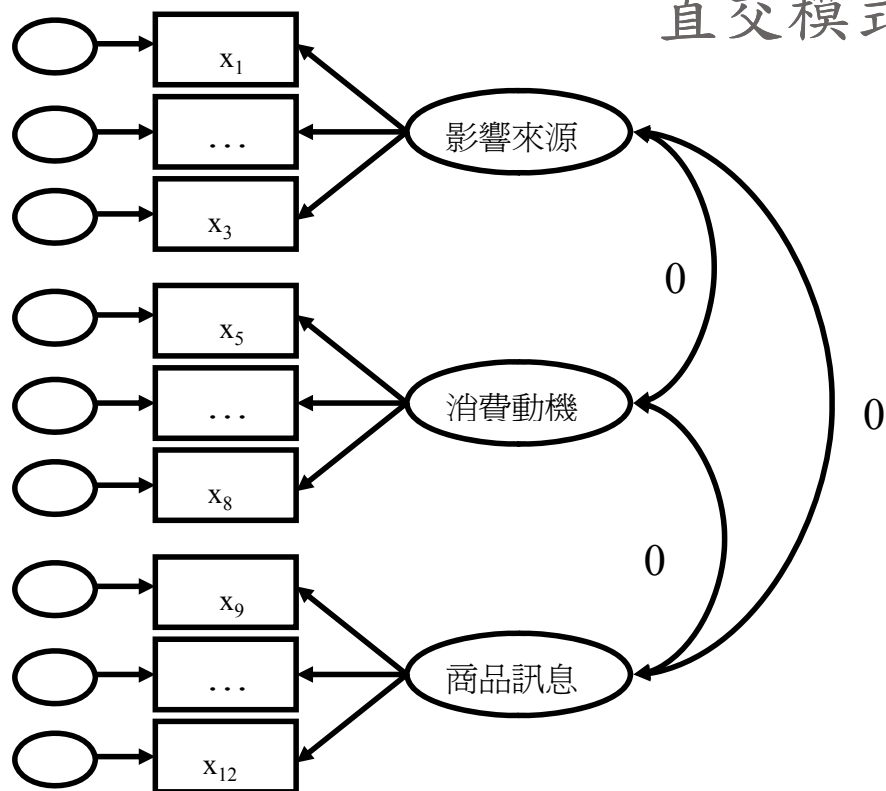
- 外衍與內衍變項的路徑係數是否顯著，正負向關係是否與理論相符。
- 參數統計量的估計值達顯著水準（**t or CR**大於**1.96**）。
- **R²**愈高，則解釋力/預測力愈高。

適配指標的討論

- 沒有單一指標可以作為唯一明確的標準。
- 沒有單獨一種指標即可涵蓋或完全取代其他指標。
- 有些指標的計算是建立在另一些指標的基礎上，亦即有些指標間是相依並有線性關係。
- 多數決的結論並不能保證所得模式是最好的適配模式。
- 因此，參採某特定文獻的標準是較可行的方式。

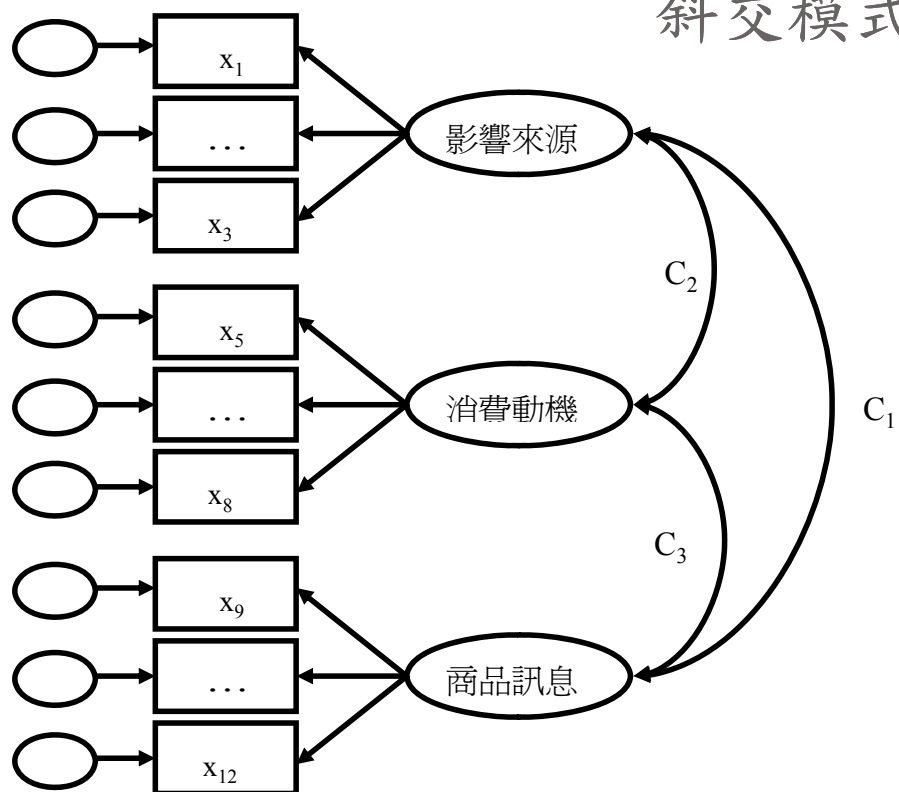
直交模式與斜交模式

直交模式



59

斜交模式



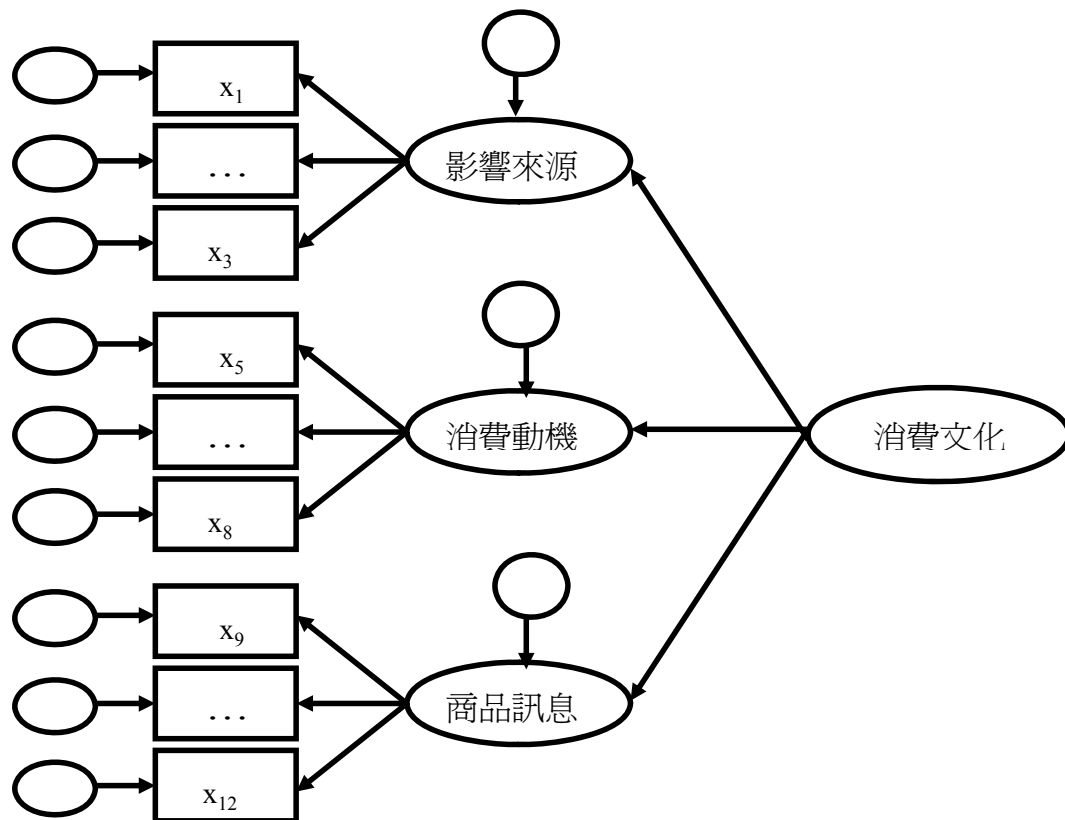
60

二階驗證性因分析

61

- 二階驗證性因素分析模式（**second-order CFA model**）是一階驗證性因素分析模式（**first-order CFA model**）的特例，又稱高階因素分析。
- 通常用於理論主張一階因素構念間有高度關聯，因此可進一步假定這幾個一階因素構念在測量更高一階的因素構念。也就是說，某一較高階結構的潛在變項可以解釋所有一階的因素構念。

62



63

收斂效度與區別效度：

- 收斂效度：一份量表的收斂效度是以標準化的測量之路徑係數(或稱負荷)為判斷依據，當所有 **負荷大於0.7**時，表示此份量表有收斂效度
- 區別效度：一份量表區別效度是指所有兩兩配對因素間都沒有完全相關，即各個因素間是有所區別的
 - t 檢定 for 相關係數

64

Thank you for your attention!!

