การวิเคราะห์องค์ประกอบเชิงยืนยัน ตอนที่ 4 (Confirmatory Factor Analysis: Part 4)

สันทัด พรประเสริฐมานิต

โครงร่างการนำเสนอ

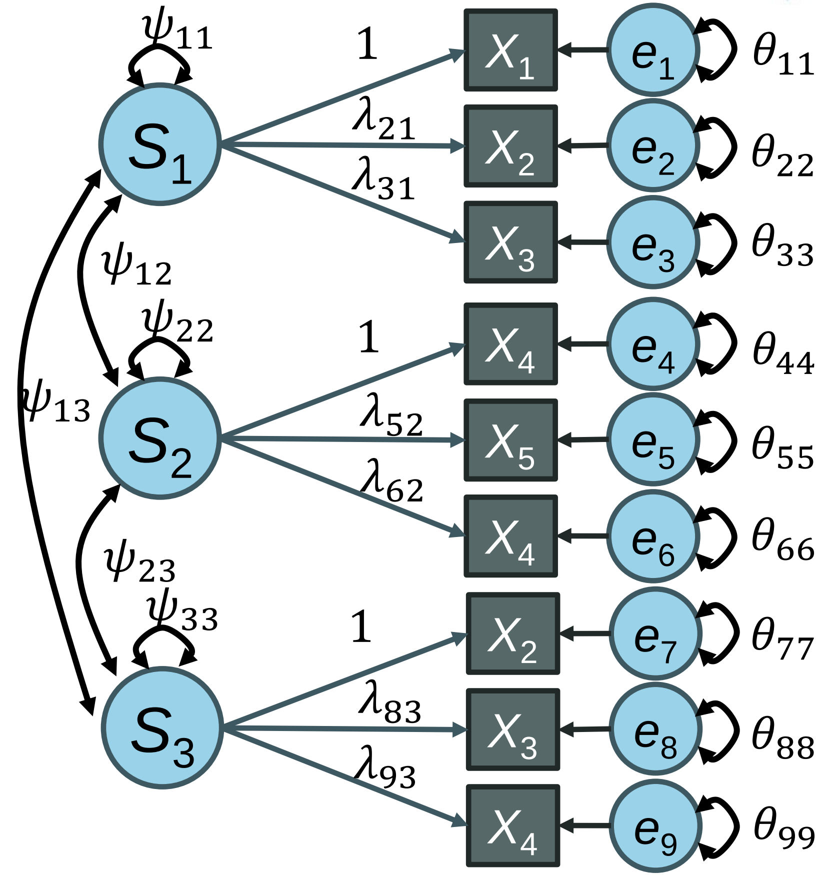
- โมเดลองค์ประกอบลำดับที่สอง (Second-order Factor Model)
- โมเดลองค์ประกอบสองด้าน (Bifactor Model)
- ตัวบ่งชี้แบบจัดอันดับ (Ordered Categorical Indicators)

โมเดลองค์ประกอบอันดับที่สอง

- CFA ในบทที่ผ่านมาจะเป็นการอธิบายว่าตัวบ่งชี้แต่ละตัว มีองค์ประกอบแฝงอะไรบ้าง

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \lambda_{21} & 0 & 0 \\ \lambda_{31} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda_{52} & 0 \\ 0 & \lambda_{62} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & \lambda_{83} \\ 0 & 0 & \lambda_{93} \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \\ e_8 \\ e_9 \end{bmatrix}$$

$$\mathbf{x} = \mathbf{\Lambda s} + \mathbf{e}$$

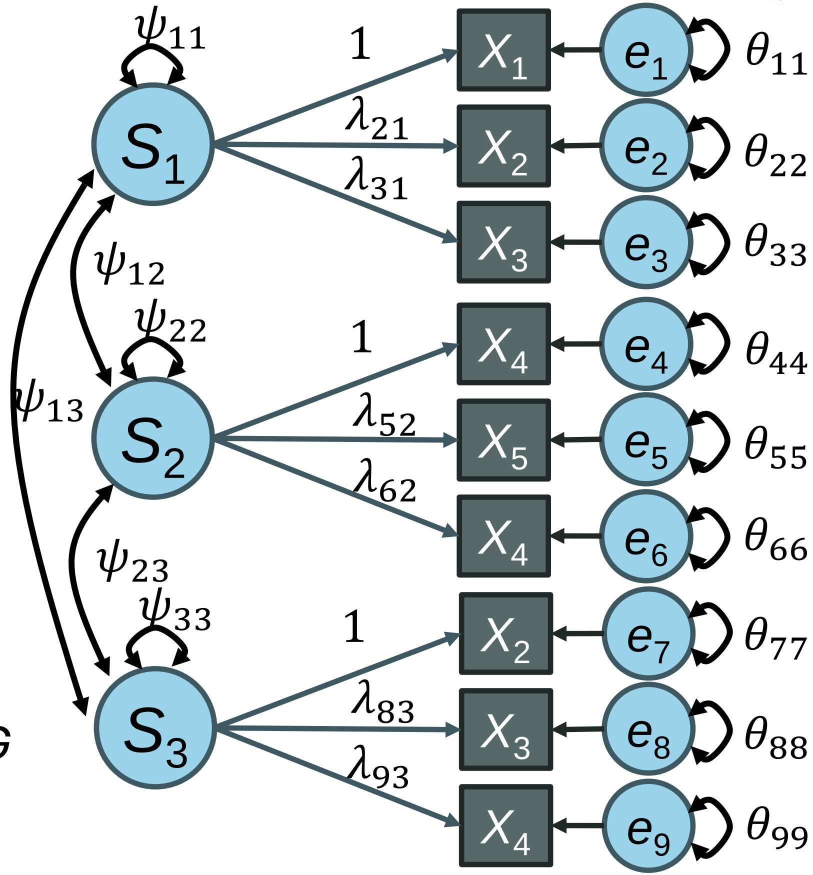


โมเดลองค์ประกอบอันดับที่สอง

- ในโมเดลนี้ **s** จะมีเมทริกซ์ความแปรปรวนร่วม คือ

$$\begin{bmatrix} \psi_{11} & & \\ \psi_{21} & \psi_{22} & \\ \psi_{31} & \psi_{32} & \psi_{33} \end{bmatrix}$$

- แนวคิดของโมเดลองค์ประกอบอันดับที่สอง (Second-order Factor) คือ องค์ประกอบ **s** เหล่านี้ที่เป็นองค์ประกอบอันดับที่หนึ่ง (First-order Factor) อาจมีองค์ประกอบที่อยู่เหนือขึ้นไปอธิบายการเปลี่ยนแปลงของ **s** ร่วมกัน เรียกว่า *G*



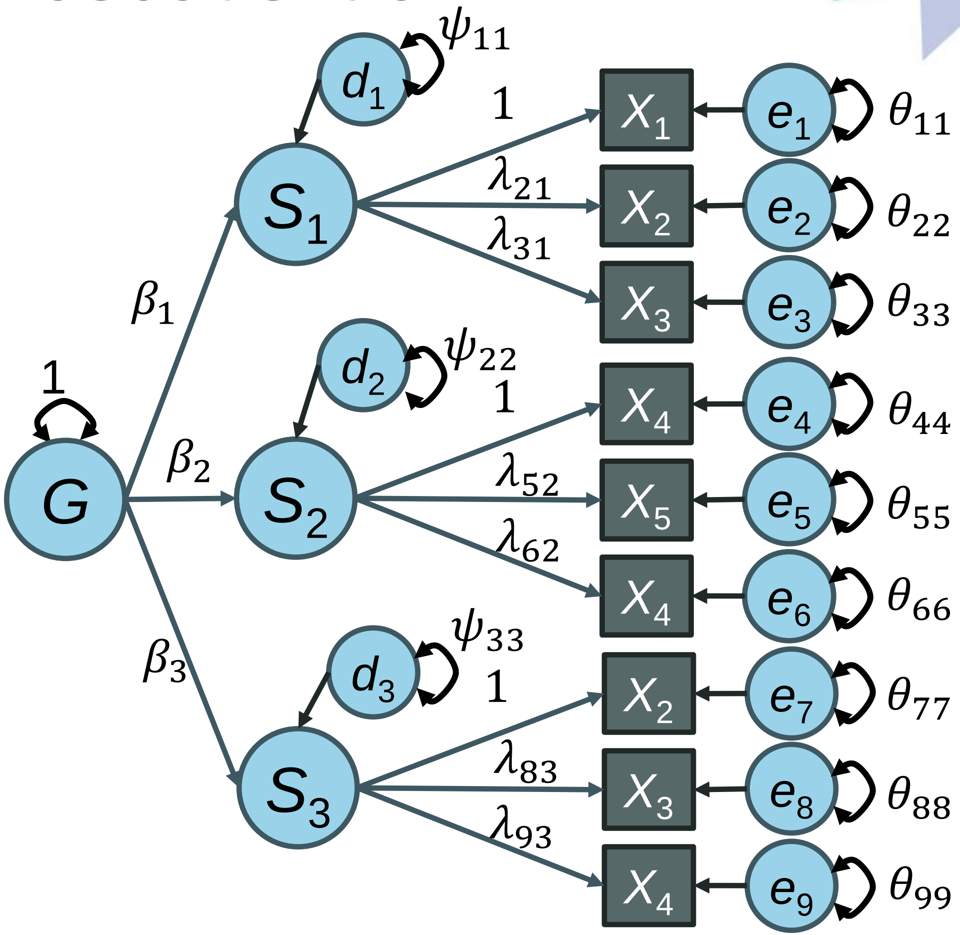
โมเดลองค์ประกอบอันดับที่สอง

- กล่าวคือ

$$S_1 = \beta_1 G + d_1$$
$$S_2 = \beta_2 G + d_2$$
$$S_3 = \beta_3 G + d_3$$

$$\begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} [G] + \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$$

$$\mathbf{s} = \mathbf{B}\mathbf{g} + \mathbf{d}$$



โมเดลองค์ประกอบอันดับที่สอง

- โมเดลที่มีองค์ประกอบอธิบายเป็นอันดับขั้น จะเรียกว่า โมเดลองค์ประกอบลำดับขั้น (Hierarchical Factor Model) และถ้ามีเพียงสองอันดับขั้น เรียกว่า โมเดลองค์ประกอบอันดับที่สอง (Second-order Factor Model)
- ตัวอย่างของโมเดลองค์ประกอบอันดับที่สอง เช่น บุคลิกภาพ ที่มีองค์ประกอบย่อย 30 องค์ประกอบ (อาจเรียกว่า Facets) และมีองค์ประกอบอันดับที่สอง 5 องค์ประกอบ ที่เรียกว่า Big Five
- องค์ประกอบอันดับที่สอง มีการระบุสเกลได้ทั้งแบบ Fixed-Factor Approach หรือ Marker "varia ble" (ควรเป็น first-order factor) approach
- โมเดลที่ระบุได้ ถ้ามีองค์ประกอบอันดับที่สองแค่องค์ประกอบเดียว ควรมีองค์ประกอบย่อยอย่างน้อย 3 องค์ประกอบเพื่อให้ระบุโมเดลได้ แต่หากมีองค์ประกอบอันดับที่สองมากกว่า 1 องค์ประกอบ ควรมีองค์ประกอบย่อย 2 องค์ประกอบต่อแต่ละองค์ประกอบอันดับที่สอง

โมเดลองค์ประกอบอันดับที่สอง

- การประเมินความเหมาะสมของโมเดล ควรทำเป็นสองขั้นตอน คือ
  - ประเมินโมเดลที่มีเฉพาะองค์ประกอบอันดับที่หนึ่งก่อน ซึ่งเป็นโมเดล CFA ปกติ ใช้วิธีการประเมินความเหมาะสมของโมเดลแบบเดิม
  - ประเมินโมเดลที่มีองค์ประกอบอันดับที่สอง เปรียบเทียบกับโมเดลที่แล้วด้วย Nested Model Comparison
- นักวิเคราะห์มุ่งที่จะหาทั้งคะแนนขององค์ประกอบทั้งอันดับแรกและอันดับที่สอง เช่น ต้องการรู้ทั้งคะแนนความรับผิดชอบ ที่เป็นองค์ประกอบอันดับที่สอง และการปฏิบัติตามหน้าที่ (Dutifulness) ที่เป็นองค์ประกอบอันดับที่หนึ่ง
  - ดังนั้น นักวิเคราะห์อาจต้องการสร้างคะแนนองค์ประกอบเพื่อประมาณค่าคะแนนองค์ประกอบเหล่านี้
  - วิธีการสร้างคะแนนองค์ประกอบที่ได้รับความนิยม คือ คะแนนรวมของข้อคำถาม ทำให้ต้องประเมินความเที่ยงของคะแนนรวมเหล่านั้น

คะแนนองค์ประกอบอันดับที่หนึ่ง

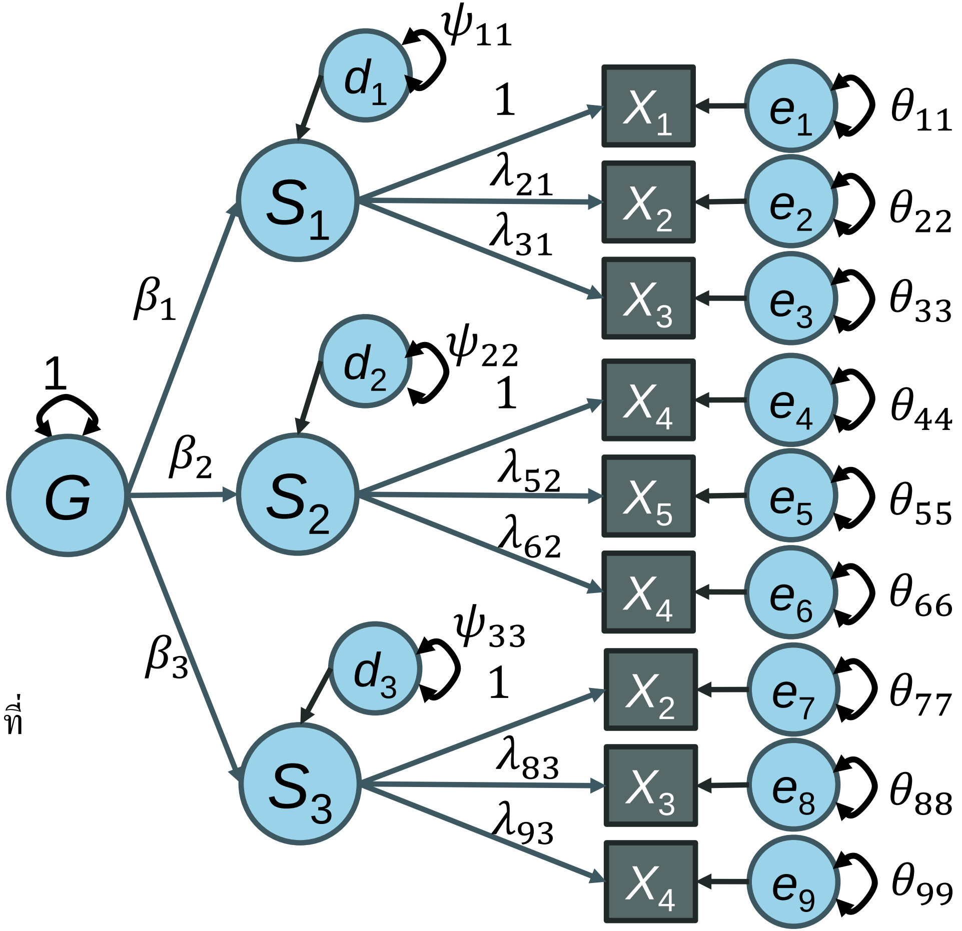
$$X_{ij} = \mu_j + \lambda_{jk} S_{ki} + e_{ij}$$

$$X_{ij} = \mu_j + \lambda_{jk}(\beta_k G + d_k) + e_{ij}$$

$$X_{ij} = \mu_j + \lambda_{jk}\beta_k G + \lambda_{jk} d_k + e_{ij}$$

คะแนนองค์ประกอบรวม

คะแนนองค์ประกอบย่อย
เมื่อควบคุมองค์ประกอบรวมให้คงที่



$$\hat{S}_{1i} = X_{i1} + X_{i2} + X_{i3} \qquad X_{ij} = \mu_j + \lambda_{jk}\beta_k G_i + \lambda_{jk} d_{ki} + e_{ij}$$

$$\hat{S}_{1i} = \mu_1 + \lambda_{11}\beta_1 G_i + \lambda_{11} d_{1i} + e_{i1} + \mu_2 + \lambda_{21}\beta_1 G_i + \lambda_{21} d_{1i} + e_{i2} + \mu_3 + \lambda_{31}\beta_1 G_i + \lambda_{31} d_{1i} + e_{i3}$$

$$\hat{S}_{1i} = \sum_{j=1}^{3} \mu_j + G_i \sum_{j=1}^{3} \lambda_{j1}\,\beta_1 + d_{1i} \sum_{j=1}^{3} \lambda_{j1} + \sum_{j=1}^{3} e_{ij}$$

$$Var(\hat{S}_{1i}) = Var\left(\sum_{j=1}^{3} \mu_j + G_i \sum_{j=1}^{3} \lambda_{j1}\,\beta_1 + d_{1i} \sum_{j=1}^{3} \lambda_{j1} + \sum_{j=1}^{3} e_{ij}\right) = Var\left(G_i \sum_{j=1}^{3} \lambda_{j1}\,\beta_1 + d_{1i} \sum_{j=1}^{3} \lambda_{j1} + \sum_{j=1}^{3} e_{ij}\right)$$

$$Var(\hat{S}_{1i}) = \left(\sum_{j=1}^{3} \lambda_{j1}\,\beta_1\right)^2 Var(G_i) + \left(\sum_{j=1}^{3} \lambda_{j1}\right)^2 Var(d_{1i}) + \sum_{j=1}^{3} Var(e_{ij})$$

$$Var(\hat{S}_{1i}) = \phi_{11}\left(\sum_{j=1}^{3} \lambda_{j1}\,\beta_1\right)^2 + \psi_{11}\left(\sum_{j=1}^{3} \lambda_{j1}\right)^2 + \sum_{j=1}^{3} \theta_{jj}$$

โมเดลองค์ประกอบอันดับที่สอง

- ค่าสัมประสิทธิ์โอเมก้า คือ สัดส่วนของความแปรปรวนองค์ประกอบทั้งหมดต่อความแปรปรวนของคะแนนรวม คือ

$$\omega_{S_1} = \frac{\phi_{11}\left(\sum_{j=1}^{3}\lambda_{j1}\beta_1\right)^2 + \psi_{11}\left(\sum_{j=1}^{3}\lambda_{j1}\right)^2}{Var(X_1 + X_2 + X_3)}$$

- ถ้านักวิเคราะห์ต้องการเอาคะแนนองค์ประกอบย่อย ไปอ้างอิงถึงคะแนนองค์ประกอบรวม ให้ใช้ค่าสัมประสิทธิ์โอเมก้าแบบลำดับขั้น (Hierarchical Omega) ได้ผลดังนี้ (Zinbarg et al., 2005)

$$\omega_{H(S_1)} = \frac{\phi_{11}\left(\sum_{j=1}^{3}\lambda_{j1}\beta_1\right)^2}{Var(X_1 + X_2 + X_3)}$$

คำนวณความเที่ยงของคะแนนรวมทั้งมตร

$$\hat{H}_i = \sum\nolimits_{j=1}^{9} X_{ij} \qquad X_{ij} = \mu_j + \lambda_{jk}\beta_k G_i + \lambda_{jk} d_{ki} + e_{ij}$$

$$\hat{H}_{1i} = \sum_{j=1}^{9} \mu_j + G_i \sum_{j=1}^{3} \lambda_{j1}\,\beta_1 + d_{1i} \sum_{j=1}^{3} \lambda_{j1} + G_i \sum_{j=4}^{6} \lambda_{j2}\,\beta_2 + d_{2i} \sum_{j=4}^{6} \lambda_{j2} + G_i \sum_{j=7}^{9} \lambda_{j3}\,\beta_3 + d_{3i} \sum_{j=7}^{9} \lambda_{j3} + \sum_{j=1}^{9} e_{ij}$$

$$\hat{H}_{1i} = \sum_{j=1}^{9} \mu_j + G_i \left(\sum_{j=1}^{3} \lambda_{j1}\,\beta_1 + \sum_{j=4}^{6} \lambda_{j2}\,\beta_2 + \sum_{j=7}^{9} \lambda_{j3}\,\beta_3 \right) + d_{1i} \sum_{j=1}^{3} \lambda_{j1} + d_{2i} \sum_{j=4}^{6} \lambda_{j2} + d_{3i} \sum_{j=7}^{9} \lambda_{j3} + \sum_{j=1}^{9} e_{ij}$$

$$Var\left(\hat{H}_{1i}\right) = \phi_{11} \left(\sum_{j=1}^{3} \lambda_{j1}\,\beta_1 + \sum_{j=4}^{6} \lambda_{j2}\,\beta_2 + \sum_{j=7}^{9} \lambda_{j3}\,\beta_3 \right)^2 + \psi_{11} \left(\sum_{j=1}^{3} \lambda_{j1} \right)^2 + \psi_{22} \left(\sum_{j=4}^{6} \lambda_{j2} \right)^2 + \psi_{33} \left(\sum_{j=7}^{9} \lambda_{j3} \right)^2 + \sum_{j=1}^{9} \theta_{jj}$$

โมเดลองค์ประกอบอันดับที่สอง

- ค่าสัมประสิทธ์โอเมก้าของคะแนนรวมทั้งมาตร คือ สัดส่วนของความแปรปรวนองค์ประกอบทั้งหมด (ทั้งอันดับที่หนึ่งและอันดับที่สอง) ต่อความแปรปรวนของคะแนนรวม คือ

$$\omega = \frac{\phi_{11}\left(\sum_{j=1}^{3}\lambda_{j1}\beta_1 + \sum_{j=4}^{6}\lambda_{j2}\beta_2 + \sum_{j=7}^{9}\lambda_{j3}\beta_3\right)^2 + \psi_{11}\left(\sum_{j=1}^{3}\lambda_{j1}\right)^2 + \psi_{22}\left(\sum_{j=4}^{6}\lambda_{j2}\right)^2 + \psi_{33}\left(\sum_{j=7}^{9}\lambda_{j3}\right)^2}{Var\left(\sum X_j\right)}$$

- อย่างไรก็ตาม นักวิเคราะห์มักต้องการนำคะแนนรวมทั้งมาตรไปอ้างถึงองค์ประกอบอันดับที่สองอยู่แล้ว ในที่นี้ ค่าสัมประสิทธิ์โอเมก้าแบบลำดับขั้น (Hierarchical Omega) จะมีประโยชน์กว่า

$$\omega_H = \frac{\phi_{11}\left(\sum_{j=1}^{3}\lambda_{j1}\beta_1 + \sum_{j=4}^{6}\lambda_{j2}\beta_2 + \sum_{j=7}^{9}\lambda_{j3}\beta_3\right)^2}{Var\left(\sum X_j\right)}$$

  - รากที่ 2 ของค่าสัมประสิทธิ์โอเมก้าแบบลำดับขั้น จะเท่ากับค่าสหสัมพันธ์ระหว่างคะแนนรวมและองค์ประกอบอันดับที่สอง

โมเดลองค์ประกอบอันดับที่สอง

- วิเคราะห์องค์ประกอบย่อยของความรับผิดชอบ ก่อนวิเคราะห์โมเดลองค์ประกอบอันดับที่สอง

```
> library(lavaan)
> datcon <- read.table("lecture11consci.csv", sep=",", header=TRUE, na.strings="999")
> mcon <- '
+ achi =~ c1 + c7 + c13 + c19 + c25 + c31 + c37 + c43 + c49 + c55
+ caut =~ c2 + c8 + c14 + c20 + c26 + c32 + c38 + c44 + c50 + c56
+ duti =~ c3 + c9 + c15 + c21 + c27 + c33 + c39 + c45 + c51 + c57
+ orde =~ c4 + c10 + c16 + c22 + c28 + c34 + c40 + c46 + c52 + c58
+ disc =~ c5 + c11 + c17 + c23 + c29 + c35 + c41 + c47 + c53 + c59
+ effi =~ c6 + c12 + c18 + c24 + c30 + c36 + c42 + c48 + c54 + c60'
> outcon <- cfa(mcon, data=datcon)
> summary(outcon, fit=TRUE, std=TRUE)
lavaan 0.6-12 ended normally after 193 iterations

  Estimator                                         ML
  Optimization method                           NLMINB
  Number of model parameters                       135

  Number of observations                           475

Model Test User Model:

  Test statistic                              5080.683
  Degrees of freedom                              1695
  P-value (Chi-square)                           0.000
```

```
User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    0.633
  Tucker-Lewis Index (TLI)                       0.616

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)             -32826.252
  Loglikelihood unrestricted model (H1)     -30285.910

  Akaike (AIC)                               65922.504
  Bayesian (BIC)                             66484.551
  Sample-size adjusted Bayesian (BIC)        66056.081

Root Mean Square Error of Approximation:

  RMSEA                                          0.065
  90 Percent confidence interval - lower         0.063
  90 Percent confidence interval - upper         0.067
  P-value RMSEA <= 0.05                          0.000

Standardized Root Mean Square Residual:

  SRMR                                           0.094
```

$$\chi^2(1695, N = 475) = 5080.68, p < .001, \mathrm{CFI} = .633, \mathrm{TLI} = .616, \mathrm{RMSEA} = .065, \mathrm{SRMR} = .094$$

โมเดลไม่ได้เหมาะสมกับข้อมูลดี แต่ใช้โมเดลนี้ต่อเพื่อเป็นตัวอย่างในการสอน

```
> inspect(outcon, "cov.lv")
     achi  caut  duti  orde  disc  effi
achi 0.204
caut 0.009 0.004
duti 0.148 0.015 0.245
orde 0.174 0.018 0.144 0.356
disc 0.138 0.013 0.114 0.171 0.145
effi 0.181 0.009 0.128 0.161 0.146 0.184
> inspect(outcon, "cor.lv")
     achi  caut  duti  orde  disc  effi
achi 1.000
caut 0.289 1.000
duti 0.663 0.450 1.000
orde 0.645 0.470 0.489 1.000
disc 0.802 0.506 0.603 0.753 1.000
effi 0.935 0.318 0.605 0.630 0.890 1.000
```

คำนวณความแปรปรวนร่วมระหว่างองค์ประกอบ

คำนวณสหสัมพันธ์ระหว่างองค์ประกอบ

องค์ประกอบย่อยทั้งหมดดูมีความสัมพันธ์ไปด้วยกัน
เป็นไปได้ว่ามีองค์ประกอบขั้นที่สอง

วิเคราะห์โมเดลองค์ประกอบลำดับที่สอง

```
> mconh <- '
+ achi =~ c1 + c7 + c13 + c19 + c25 + c31 + c37 + c43 + c49 + c55
+ caut =~ c2 + c8 + c14 + c20 + c26 + c32 + c38 + c44 + c50 + c56
+ duti =~ c3 + c9 + c15 + c21 + c27 + c33 + c39 + c45 + c51 + c57
+ orde =~ c4 + c10 + c16 + c22 + c28 + c34 + c40 + c46 + c52 + c58
+ disc =~ c5 + c11 + c17 + c23 + c29 + c35 + c41 + c47 + c53 + c59
+ effi =~ c6 + c12 + c18 + c24 + c30 + c36 + c42 + c48 + c54 + c60
+ con =~ achi + caut + duti + orde + disc + effi'
> outconh <- cfa(mconh, data=datcon)
> summary(outconh, fit=TRUE, std=TRUE)
lavaan 0.6-12 ended normally after 139

  Estimator
  Optimization method
  Number of model parameters

  Number of observations                              475

Model Test User Model:

  Test statistic                                 5175.991
  Degrees of freedom                                 1704
  P-value (Chi-square)                              0.000
```

นิยามองค์ประกอบลำดับที่สอง โดยเสมือนองค์ประกอบลำดับที่หนึ่งเป็นตัวบ่งชี้เลย

```
User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    0.623
  Tucker-Lewis Index (TLI)                       0.609

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)             -32873.906
  Loglikelihood unrestricted model (H1)     -30285.910

  Akaike (AIC)                               65999.812
  Bayesian (BIC)                             66524.390
  Sample-size adjusted Bayesian (BIC)        66124.484

Root Mean Square Error of Approximation:

  RMSEA                                          0.065
  90 Percent confidence interval - lower         0.063
  90 Percent confidence interval - upper         0.068
  P-value RMSEA <= 0.05                          0.000

Standardized Root Mean Square Residual:

  SRMR                                           0.096
```

$$\chi^2(1704, N = 475) = 5175.99, p < .001, \mathrm{CFI} = .623, \mathrm{TLI} = .609, \mathrm{RMSEA} = .065, \mathrm{SRMR} = .096$$

ค่าดัชนีความเหมาะสมตกลงเล็กน้อย เมื่อเทียบกับโมเดล CFA

```
                 Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
con =~
  achi              1.000                               0.935    0.935
  caut              0.072    0.043    1.686    0.092    0.441    0.441
  duti              0.811    0.088    9.217    0.000    0.670    0.670
  orde              1.060    0.105   10.115    0.000    0.723    0.723
  disc              0.846    0.110    7.687    0.000    0.935    0.935
  effi              0.961    0.096   10.028    0.000    0.949    0.949
```

ดูแล้วมี Cautiousness ตัวเดียวที่ Standardized Loadings ต่ำที่สุด องค์ประกอบย่อยอื่น มีขนาดดี

```
> library(nonnest2)
> vuongtest(outcon, outconh, nested=TRUE)

Model 1
 Class: lavaan
 Call: lavaan::lavaan(model = mcon, data = datcon, model.type = "cfa", ...

Model 2
 Class: lavaan
 Call: lavaan::lavaan(model = mconh, data = datcon, model.type = "cfa", ...

Variance test
  H0: Model 1 and Model 2 are indistinguishable
  H1: Model 1 and Model 2 are distinguishable
    w2 = 0.354,   p = <2e-16

Robust likelihood ratio test of distinguishable models
  H0: Model 2 fits as well as Model 1
  H1: Model 1 fits better than Model 2
    LR = 95.308,   p = 1.16e-08
```

พบว่าโมเดล CFA ปกติดีกว่าโมเดลองค์ประกอบสองชั้นแตกต่างอย่างมีนัยสำคัญ ตามนี้ ควรเลือกโมเดล CFA ปกติ หรือหาโมเดลองค์ประกอบอันดับที่สองอื่น

จากที่ผมลองตรวจสอบแล้ว พบว่าหากนำ Cautiousness ออกจากโมเดลแล้ว โมเดลองค์ประกอบแบบสองชั้นจะเหมาะสมดี

```
> detailscon <- data.frame(detailscon[,c("VariableName", "Direction")])
> detailscon <- detailscon[detailscon$VariableName %in% paste0("c", 1:60),]
>
> for(i in 1:nrow(detailscon)) {
+   if(detailscon[i, "Direction"] == "-"){
+     temp <- datcon[,detailscon[i, "VariableName"]]
+     temp <- 6 - temp
+     datcon[,detailscon[i, "VariableName"]] <- temp
+   }
+ }
> outcon <- cfa(mcon, data=datcon)
> outconh <- cfa(mconh, data=datcon)
>
> library(semTools)
> compRelSEM(outconh, higher="con")
 achi  caut  duti  orde  disc  effi   con
0.636 0.697 0.768 0.816 0.720 0.797 0.777
```

ก่อนหาความเที่ยง ต้องกลับคะแนนข้อคำถามทุกข้อให้เป็นทิศทางเดียวกันก่อน

ผมเขียน for loop เพื่อให้อ่านข้อมูลแล้วกลับคะแนนแต่ละข้ออัตโนมัติ ถ้าไม่ถนัด เขียนกลับคะแนน โดยนำคะแนนไปลบจาก 6 ทีละข้อได้เลย

ประมาณค่าโมเดลใหม่

หาความเที่ยงโดยกำหนดองค์ประกอบลำดับที่สองใน higher

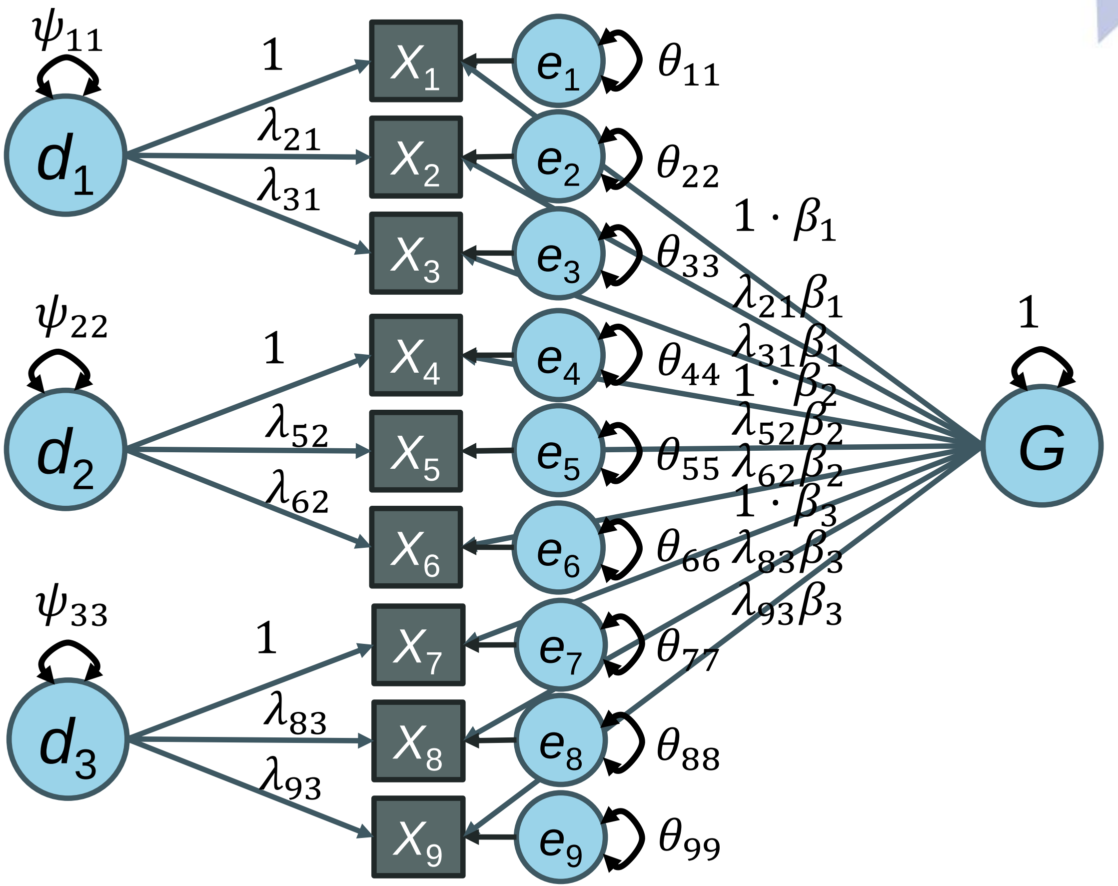
สูตรโอเมก้าปกติขององค์ประกอบย่อย (ผลรวมจะเฉพาะข้อที่อยู่ในแต่ละองค์ประกอบย่อย)

สูตรโอเมก้าลำดับขั้นของผลรวมคะแนนทุกข้อ

สังเกตที่สีของตัวอักษร ว่าตรงกับสูตรไหนในสไลด์ที่ผ่านมา

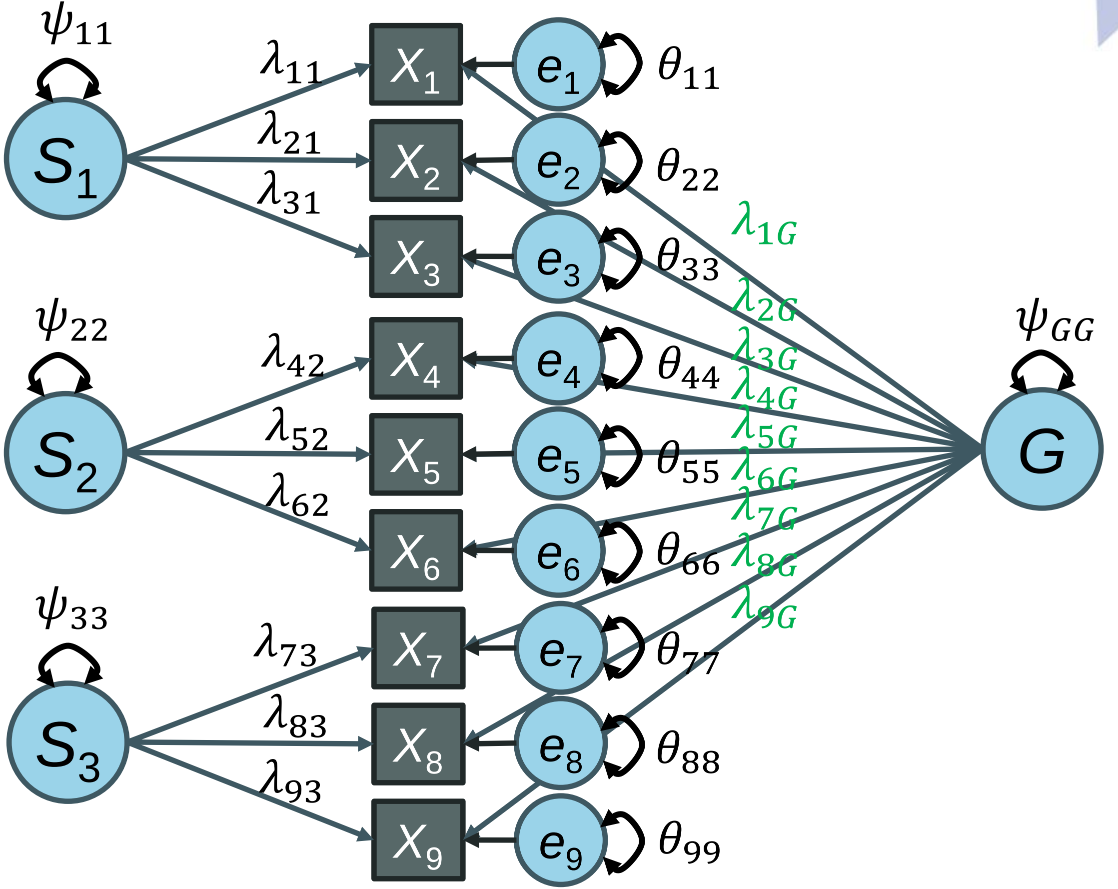
โมเดลองค์ประกอบสองด้าน

- โมเดลองค์ประกอบลำดับขั้นจะมองว่า องค์ประกอบอันดับที่หนึ่งถูกอธิบายด้วยองค์ประกอบอันดับที่สองซ้อนอีกทีหนึ่ง
- ถ้าเขียนโมเดลองค์ประกอบลำดับขั้นใหม่ จะสามารถเขียนได้ดังนี้
- เห็นว่า แต่ละข้อคำถามจะอธิบายด้วยองค์ประกอบรวม (*G*) และองค์ประกอบกลุ่มย่อย (*d*)



โมเดลองค์ประกอบสองด้าน

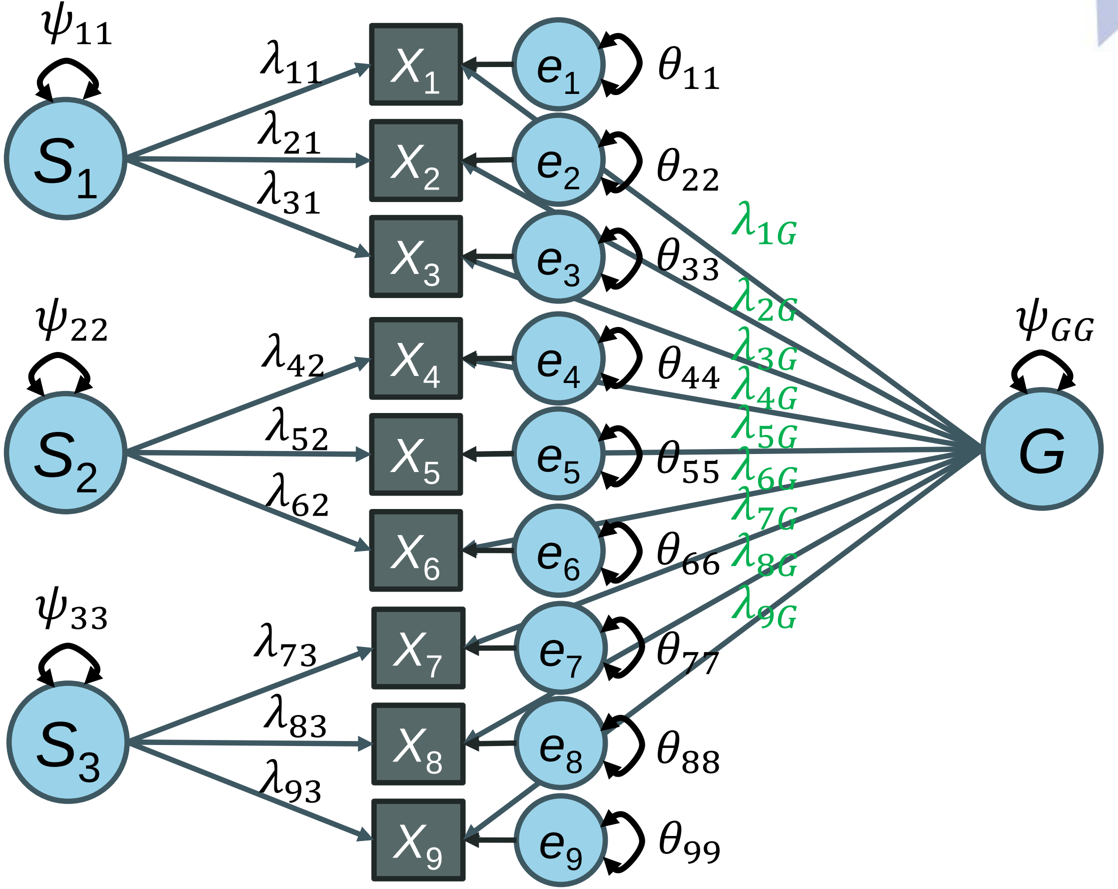
- อีกแนวคิดหนึ่ง มองว่าตัวบ่งชี้ทั้งหมด วัดเพียงองค์ประกอบองค์ประกอบเดียว แต่ตัวบ่งชี้บางข้อสามารถจับกลุ่มกัน ในสิ่งที่ไม่เกี่ยวข้องกับองค์ประกอบที่ต้องการวัด แนวคิดองค์ประกอบแบบนี้จะเรียกว่า โมเดลองค์ประกอบสองด้าน (Bifactor M odel) เกิดจาก
  - องค์ประกอบรวม (General Factor)
  - องค์ประกอบย่อย (Specific Factor)



Bifactor model

$$\mathbf{x} = \mathbf{\Lambda f} + \mathbf{e}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \end{bmatrix} = \begin{bmatrix} \lambda_{1G} & \lambda_{11} & 0 & 0 \\ \lambda_{2G} & \lambda_{21} & 0 & 0 \\ \lambda_{3G} & \lambda_{31} & 0 & 0 \\ \lambda_{4G} & 0 & \lambda_{42} & 0 \\ \lambda_{5G} & 0 & \lambda_{52} & 0 \\ \lambda_{6G} & 0 & \lambda_{62} & 0 \\ \lambda_{7G} & 0 & 0 & \lambda_{73} \\ \lambda_{8G} & 0 & 0 & \lambda_{83} \\ \lambda_{9G} & 0 & 0 & \lambda_{93} \end{bmatrix} \begin{bmatrix} G \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \\ e_8 \\ e_9 \end{bmatrix}$$



โมเดลองค์ประกอบสองด้าน

- การระบุโมเดลองค์ประกอบสองด้านและระบุสเกลองค์ประกอบใช้แนวทางเดียวกับ CFA เลย
- ส่วนการทดสอบความเหมาะสมของโมเดลองค์ประกอบสองด้าน ให้ใช้ดัชนีความเหมาะสมแบบปกติ แต่ DFI ยังไม่สนับสนุนการทดสอบโมเดลองค์ประกอบสองด้าน
- การเปรียบเทียบโมเดลองค์ประกอบสองด้านกับโมเดลอื่น
  - เปรียบเทียบกับกับ โมเดลองค์ประกอบแบบลำดับขั้น (Hierarchical Factor Model) ด้วย Nested Model Compa rison ได้ เพราะ Bifactor Model เป็นโมเดลตั้งต้น
  - หากต้องการเปรียบเทียบระหว่าง CFA ปกติ (ที่มี $S_1, S_2, \dots, S_k$ ที่สัมพันธ์กัน) และโมเดลองค์ประกอบสองด้านจะไม่ใช่โมเดลที่ซ้อนกันกับ CFA ต้องใช้ Nonnested Model Comparison

โมเดลองค์ประกอบสองด้าน

- โมเดลองค์ประกอบแบบลำดับขั้นจะบอกว่า อิทธิพลของ General Factor ที่ส่งไปสู่ตัวบ่งชี้ต่างๆ ถูก Fully Mediated โดย Specific Factor แต่ละตัว
  - ถ้าเลือกโมเดลองค์ประกอบแบบลำดับขั้น แสดงว่าองค์ประกอบรวม ส่งผลต่อองค์ประกอบย่อย แล้วค่อยส่งผลต่อตัวบ่งชี้เท่านั้น
  - แต่ถ้าปฏิเสธโมเดลองค์ประกอบแบบลำดับขั้น แสดงว่าองค์ประกอบรวมมีผลต่อตัวบ่งชี้โดยตรง ไม่จำเป็นต้องผ่านองค์ประกอบย่อย
  - แสดงว่าโมเดลแบบสองด้าน องค์ประกอบย่อยจะเกี่ยวข้อง เป็นภาวะสันนิษฐานเนื้อเดียวกับองค์ประกอบรวม (เช่น ด้านต่างๆ ของความพึงพอใจในชีวิต) หรือไม่เกี่ยวข้องใดๆ เลย (เช่น วิธีการวัด เครื่องมือ ฯลฯ) ก็ได้
- โมเดลองค์ประกอบสองด้านสามารถหาความเที่ยงของคะแนนรวมตัวบ่งชี้ภายในองค์ประกอบย่อยและคะแนนรวมทั้งมาตรได้เหมือนกัน เพื่อแทนค่าองค์ประกอบย่อยและองค์ประกอบรวม

คำนวณความเที่ยงของคะแนนองค์ประกอบย่อย

$$\hat{S}_{1i} = X_{i1} + X_{i2} + X_{i3} \qquad X_{ij} = \mu_j + \lambda_{jG} G_i + \lambda_{jk} S_{ki} + e_{ij}$$

$$\hat{S}_{1i} = \mu_1 + \lambda_{1G} G_i + \lambda_{11} S_{1i} + e_{i1} + \mu_2 + \lambda_{2G} G_i + \lambda_{21} S_{1i} + e_{i2} + \mu_3 + \lambda_{3G} G_i + \lambda_{31} d_{1i} + e_{i3}$$

$$\hat{S}_{1i} = \sum_{j=1}^{3} \mu_j + G_i \sum_{j=1}^{3} \lambda_{jG} + S_{1i} \sum_{j=1}^{3} \lambda_{j1} + \sum_{j=1}^{3} e_{ij}$$

$$Var(\hat{S}_{1i}) = Var\left(\sum_{j=1}^{3} \mu_j + G_i \sum_{j=1}^{3} \lambda_{jG} + S_{1i} \sum_{j=1}^{3} \lambda_{j1} + \sum_{j=1}^{3} e_{ij} \right) = Var\left(G_i \sum_{j=1}^{3} \lambda_{jG} + S_{1i} \sum_{j=1}^{3} \lambda_{j1} + \sum_{j=1}^{3} e_{ij} \right)$$

$$Var(\hat{S}_{1i}) = \left(\sum_{j=1}^{3} \lambda_{jG} \right)^2 Var(G_i) + \left(\sum_{j=1}^{3} \lambda_{j1} \right)^2 Var(S_{1i}) + \sum_{j=1}^{3} Var(e_{ij})$$

$$Var(\hat{S}_{1i}) = \psi_{GG} \left(\sum_{j=1}^{3} \lambda_{jG} \right)^2 + \psi_{11} \left(\sum_{j=1}^{3} \lambda_{j1} \right)^2 + \sum_{j=1}^{3} \theta_{jj}$$

โมเดลองค์ประกอบสองด้าน

- ค่าสัมประสิทธิ์โอเมก้า คือ สัดส่วนของความแปรปรวนองค์ประกอบทั้งหมดต่อความแปรปรวนของคะแนนรวม คือ

$$\omega_{S_1} = \frac{\psi_{GG}\left(\sum_{j=1}^{3}\lambda_{jG}\right)^2 + \psi_{11}\left(\sum_{j=1}^{3}\lambda_{j1}\right)^2}{Var(X_1 + X_2 + X_3)}$$

- ถ้านักวิเคราะห์ต้องการเอาคะแนนองค์ประกอบย่อย ไปอ้างอิงถึงคะแนนองค์ประกอบรวม ให้ใช้ค่าสัมประสิทธิ์โอเมก้าแบบลำดับขั้น (Hierarchical Omega) ได้ผลดังนี้ (Rodriguez et al., 2015)

$$\omega_{H(S_1)} = \frac{\psi_{GG}\left(\sum_{j=1}^{3}\lambda_{jG}\right)^2}{Var(X_1 + X_2 + X_3)}$$

โมเดลองค์ประกอบสองด้าน

- ถ้าเทียบระหว่างคนที่คะแนนองค์ประกอบรวมเท่ากัน (ควบคุมอิทธิพลของคะแนนรวม) ผลรวมคะแนนสามข้อนี้จะบ่งชี้ขนาดคะแนนองค์ประกอบย่อยได้มากน้อยเพียงใด ค่านี้เรียกว่าสัมประสิทธิ์โอเมก้าลำดับขั้นสำหรับองค์ประกอบย่อย (Omega Hierarchical Subscale) มีค่าดังนี้

$$\omega_{HS(S_1)} = \frac{\psi_{11}\left(\sum_{j=1}^{3}\lambda_{j1}\right)^2}{Var(X_1 + X_2 + X_3)}$$

คำนวณความเที่ยงของคะแนนรวมทั้งมตร

$$\widehat{H}_i = \sum_{j=1}^{9} X_{ij} \qquad X_{ij} = \mu_j + \lambda_{jG} G_i + \lambda_{jk} S_{ki} + e_{ij}$$

$$\widehat{H}_{1i} = \sum_{j=1}^{9} \mu_j + G_i \sum_{j=1}^{3} \lambda_{jG} + S_{1i} \sum_{j=1}^{3} \lambda_{j1} + G_i \sum_{j=4}^{6} \lambda_{jG} + S_{2i} \sum_{j=4}^{6} \lambda_{j2} + G_i \sum_{j=7}^{9} \lambda_{jG} + S_{3i} \sum_{j=7}^{9} \lambda_{j3} + \sum_{j=1}^{9} e_{ij}$$

$$\widehat{H}_{1i} = \sum_{j=1}^{9} \mu_j + G_i \left(\sum_{j=1}^{3} \lambda_{jG} + \sum_{j=4}^{6} \lambda_{jG} + \sum_{j=7}^{9} \lambda_{jG} \right) + S_{1i} \sum_{j=1}^{3} \lambda_{j1} + S_{2i} \sum_{j=4}^{6} \lambda_{j2} + S_{3i} \sum_{j=7}^{9} \lambda_{j3} + \sum_{j=1}^{9} e_{ij}$$

$$Var(\widehat{H}_{1i}) = \psi_{GG} \left(\sum_{j=1}^{9} \lambda_{jG} \right)^2 + \psi_{11} \left(\sum_{j=1}^{3} \lambda_{j1} \right)^2 + \psi_{22} \left(\sum_{j=4}^{6} \lambda_{j2} \right)^2 + \psi_{33} \left(\sum_{j=7}^{9} \lambda_{j3} \right)^2 + \sum_{j=1}^{9} \theta_{jj}$$

โมเดลองค์ประกอบสองด้าน

- ค่าสัมประสิทธิ์โอเมก้าของคะแนนรวมทั้งมาตร คือ สัดส่วนของความแปรปรวนองค์ประกอบทั้งหมด (ทั้งอันดับที่หนึ่งและอันดับที่สอง) ต่อความแปรปรวนของคะแนนรวม คือ

$$\omega = \frac{\psi_{GG}\left(\sum_{j=1}^{9}\lambda_{jG}\right)^2 + \psi_{11}\left(\sum_{j=1}^{3}\lambda_{j1}\right)^2 + \psi_{22}\left(\sum_{j=4}^{6}\lambda_{j2}\right)^2 + \psi_{33}\left(\sum_{j=7}^{9}\lambda_{j3}\right)^2}{Var\left(\sum X_j\right)}$$

- อย่างไรก็ตาม นักวิเคราะห์มักต้องการนำคะแนนรวมทั้งมาตรไปอ้างถึงองค์ประกอบอันดับที่สองอยู่แล้ว ในที่นี้ ค่าสัมประสิทธิ์โอเมก้าแบบลำดับขั้น (Hierarchical Omega) จะมีประโยชน์กว่า

$$\omega_H = \frac{\psi_{GG}\left(\sum_{j=1}^{9}\lambda_{jG}\right)^2}{Var\left(\sum X_j\right)}$$

  – รากที่ 2 ของค่าสัมประสิทธิ์โอเมก้าแบบลำดับขั้น จะเท่ากับค่าสหสัมพันธ์ระหว่างคะแนนรวมและองค์ประกอบอันดับที่สอง

```
> mconb <- '
+ achi =~ c1 + c7 + c13 + c19 + c25 + c31 + c37 + c43 + c49 + c55
+ caut =~ c2 + c8 + c14 + c20 + c26 + c32 + c38 + c44 + c50 + c56
+ duti =~ c3 + c9 + c15 + c21 + c27 + c33 + c39 + c45 + c51 + c57
+ orde =~ c4 + c10 + c16 + c22 + c28 + c34 + c40 + c46 + c52 + c58
+ disc =~ c5 + c11 + c17 + c23 + c29 + c35 + c41 + c47 + c53 + c59
+ effi =~ c6 + c12 + c18 + c24 + c30 + c36 + c42 + c48 + c54 + c60
+ con =~ c1 + c7 + c13 + c19 + c25 + c31 + c37 + c43 + c49 + c55
+ + c2 + c8 + c14 + c20 + c26 + c32 + c38 + c44 + c50 + c56
+ + c3 + c9 + c15 + c21 + c27 + c33 + c39 + c45 + c51 + c57
+ + c4 + c10 + c16 + c22 + c28 + c34 + c40 + c46 + c52 + c58
+ + c5 + c11 + c17 + c23 + c29 + c35 + c41 + c47 + c53 + c59
+ + c6 + c12 + c18 + c24 + c30 + c36 + c42 + c48 + c54 + c60
+ con ~~ 0*achi + 0*caut + 0*duti + 0*orde + 0*disc + 0*effi
+ achi ~~ 0*caut + 0*duti + 0*orde + 0*disc + 0*effi
+ caut ~~ 0*duti + 0*orde + 0*disc + 0*effi
+ duti ~~ 0*orde + 0*disc + 0*effi
+ orde ~~ 0*disc + 0*effi
+ disc ~~ 0*effi
+ '
> outconb <- cfa(mconb, data=datcon, std.lv=TRUE)
```

บังคับให้องค์ประกอบทั้งหมดไม่สัมพันธ์กัน

ใช้ Fixed Factor
เพราะ Marker ประมาณค่าไม่ได้

```
> summary(outconb, fit=TRUE, std=TRUE)
lavaan 0.6-12 ended normally after 34 iterations

  Estimator                                         ML
  Optimization method                           NLMINB
  Number of model parameters                       180

  Number of observations                           475

Model Test User Model:

  Test statistic                              4173.239
  Degrees of freedom                              1650
  P-value (Chi-square)                           0.000
```

```
User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    0.726
  Tucker-Lewis Index (TLI)                       0.706

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)             -32372.530
  Loglikelihood unrestricted model (H1)     -30285.910

  Akaike (AIC)                               65105.060
  Bayesian (BIC)                             65854.456
  Sample-size adjusted Bayesian (BIC)        65283.163

Root Mean Square Error of Approximation:

  RMSEA                                          0.057
  90 Percent confidence interval - lower         0.055
  90 Percent confidence interval - upper         0.059
  P-value RMSEA <= 0.05                          0.000

Standardized Root Mean Square Residual:

  SRMR                                           0.076
```

ค่าดัชนีความเหมาะสมดีกว่าโมเดล CFA แบบปกติด้วยซ้ำ
แต่ค่าดัชนีความเหมาะสมอาจยังไม่เพียงพอที่จะยอมรับโมเดล
องค์ประกอบสองด้าน แต่ใช้ผลนี้เป็นตัวอย่างไปก่อน

$$\chi^2(1650, N = 475) = 4173.24, p < .001, \mathrm{CFI} = .726, \mathrm{TLI} = .706, \mathrm{RMSEA} = .057, \mathrm{SRMR} = .076$$

con =~						
c1	0.432	0.031	13.916	0.000	0.432	0.599
c7	0.401	0.029	13.794	0.000	0.401	0.594
c13	0.342	0.028	12.048	0.000	0.342	0.531
c19	0.270	0.047	5.740	0.000	0.270	0.271
c25	-0.005	0.048	-0.114	0.910	-0.005	-0.005
c31	0.456	0.033	13.881	0.000	0.456	0.597
c37	0.248	0.036	6.867	0.000	0.248	0.319
c43	0.233	0.041	5.724	0.000	0.233	0.268
c49	0.454	0.042	10.838	0.000	0.454	0.485
c55	0.462	0.048	9.698	0.000	0.462	0.439
c2	0.332	0.035	9.536	0.000	0.332	0.432
c8	0.174	0.037	4.716	0.000	0.174	0.222
c14	0.253	0.037	6.762	0.000	0.253	0.314
c20	0.111	0.046	2.402	0.016	0.111	0.114
c26	0.393	0.043	9.208	0.000	0.393	0.419
c32	0.065	0.041	1.582	0.114	0.065	0.076
c38	0.163	0.042	3.855	0.000	0.163	0.183
c44	-0.114	0.042	-2.719	0.007	-0.114	-0.129
c50	0.446	0.046	9.764	0.000	0.446	0.442
c56	0.253	0.050	5.106	0.000	0.253	0.240
c3	0.298	0.035	8.584	0.000	0.298	0.393
c9	0.333	0.036	9.357	0.000	0.333	0.425
c15	0.244	0.038	6.359	0.000	0.244	0.297
c21	0.288	0.034	8.538	0.000	0.288	0.391
c27	0.269	0.034	8.029	0.000	0.269	0.370
c33	0.266	0.044	6.025	0.000	0.266	0.282
c39	0.354	0.041	8.614	0.000	0.354	0.394
c45	0.315	0.035	8.953	0.000	0.315	0.409

ข้อเหล่านี้ควรนำไปพิจารณา ว่าเกิดอะไรขึ้น ทำไป Standardized Loadings ต่ำใกล้ 0 หรือติดลบ

ส่วน Standardized Loadings ขององค์ประกอบจำเพาะ นักวิเคราะห์อาจจะไม่สนใจ หรือสนใจก็ได้ ขึ้นอยู่ว่าจะแปลความหมายองค์ประกอบจำเพาะหรือไม่

```
> vuongtest(outcon, outconb)

Model 1
 Class: lavaan
 Call: lavaan::lavaan(model = mcon, data = datcon, mod

Model 2
 Class: lavaan
 Call: lavaan::lavaan(model = mconb, data = datcon, st

Variance test
  H0: Model 1 and Model 2 are indistinguishable
  H1: Model 1 and Model 2 are distinguishable
    w2 = 4.921,   p = <2e-16

Non-nested likelihood ratio test
  H0: Model fits are equal for the focal population
  H1A: Model 1 fits better than Model 2
    z = -9.384,   p = 1
  H1B: Model 2 fits better than Model 1
    z = -9.384,   p = < 2.2e-16
```

โมเดลองค์ประกอบสองด้านดีกว่าโมเดล CFA ปกติ อย่างมีนัยสำคัญ

```
> vuongtest(outconb, outconh, nested=TRUE)

Model 1
 Class: lavaan
 Call: lavaan::lavaan(model = mconb, data = datcon, std.l

Model 2
 Class: lavaan
 Call: lavaan::lavaan(model = mconh, data = datcon, model

Variance test
  H0: Model 1 and Model 2 are indistinguishable
  H1: Model 1 and Model 2 are distinguishable
    w2 = 4.773,   p = 2.78e-08

Robust likelihood ratio test of distinguishable models
  H0: Model 2 fits as well as Model 1
  H1: Model 1 fits better than Model 2
    LR = 1002.752,   p = <2e-16
```

โมเดลองค์ประกอบสองด้านแตกต่างจากโมเดลองค์ประกอบสองขั้นอย่างมีนัยสำคัญ จึงเลือกโมเดลองค์ประกอบสองด้าน

```
> compRelSEM(outconb)
 achi  caut  duti  orde  disc  effi   con
0.183 0.491 0.314 0.385 0.200 0.073 0.816
```

สูตรโอเมก้าลำดับขั้นสำหรับองค์ประกอบย่อย
(ผลรวมจะเฉพาะข้อที่อยู่ในแต่ละองค์ประกอบย่อย)

สูตรโอเมก้าลำดับขั้นของผลรวมคะแนนทุกข้อ

สังเกตที่สีของตัวอักษร ว่าตรงกับสูตรไหนในสไลด์ที่ผ่านมา

ตัวบ่งชี้แบบจัดอันดับ

- ใน CFA หรือ SEM ที่ผ่านมาทั้งหมด จะกำหนดให้ตัวบ่งชี้เป็นตัวแปรต่อเนื่อง แต่มาตรวัดทางจิตวิทยาส่วนใหญ่ จะเป็นมาตรที่แบ่งเป็นกลุ่มๆ เช่น มาตรวัดแบบลิเคิร์ท (Likert) หรือมาตรวัดแบบใช่/ไม่ใช่ (Binary variables)
- แม้จะสามารถเอา CFA สำหรับตัวแปรต่อเนื่องมาวิเคราะห์ตัวบ่งชี้แบบจัดอันดับได้ แต่ค่าพารามิเตอร์จะมีค่าน้อยกว่าความเป็นจริง ส่งผลให้ค่าน้ำหนักองค์ประกอบมาตรฐานต่ำ ทำให้ดูเหมือนมาตรไม่มีคุณภาพ
- โมเดลที่เหมาะสมกับตัวบ่งชี้แบบจัดกลุ่มแบบเรียงลำดับ (Ordered Categorical Variables) นั้นจะต้องสร้างสมการที่เหมาะสมสำหรับความสัมพันธ์ระหว่างองค์ประกอบที่เป็นตัวแปรต่อเนื่อง และตัวบ่งชี้ที่เป็นตัวแปรแบบจัดกลุ่มเรียงลำดับ

ตัวบ่งชี้แบบจัดอันดับ

- ให้ x_j เป็นตัวแปรแบบจัดอันดับ ซึ่งมีค่าได้ตั้งแต่ $c = 0,1,\ldots,C-1$ โดย C คือจำนวนกลุ่ม
  - หากเป็นมาตรแบบทวินาม (Binary Variables) $C = 2$ และ $c = 0,1$
  - หากเป็นมาตรแบบลิเคิร์ต 5 ระดับ $C = 5$ และ $c = 0,1,2,3,4$
- มาตรจัดอันดับเหล่านี้ จะมีความสัมพันธ์กับองค์ประกอบผ่านจุดเปลี่ยน (Thresholds) เช่น ถ้าองค์ประกอบความรับผิดชอบ (*F*) อยู่ในระดับต่ำมากๆ จะตอบข้อคำถามทวินามที่ว่า “ฉันเป็นคนรับผิดชอบ” ว่า “ไม่ใช่” หากองค์ประกอบความรับผิดชอบ (*F*) มีค่ามากขึ้นเรื่อยๆ จนเกินระดับจุดเปลี่ยน จะตอบคำถามทวินามดังกล่าวว่า “ใช่”
  - เช่น ให้ตัวบ่งชี้มีน้ำหนักองค์ประกอบมาตรฐานเท่ากับ 0.8, $F \sim N(0,1), e \sim N(0,0.6)$ ถ้าตัวบ่งชี้เป็นตัวแปรต่อเนื่อง (ให้สัญลักษณ์แทนเป็น y^*) สมการที่อธิบายอิทธิพลขององค์ประกอบไปหา y^* เป็นดังนี้

$$y^* = 0.8F + e$$

$$y^* = 0.8F + e$$

คุณมีความรับผิดชอบ
(ไม่ใช่ / ใช่)

จุดเปลี่ยน = 0.25

```
> set.seed(123321)
> f <- rnorm(10, 0, 1)
> e <- rnorm(10, 0, 0.6)
> ystar <- 0.8*f + e
> ycut <- cut(ystar, c(-Inf, 0.25, Inf))
> y <- as.numeric(ycut) - 1
> data.frame(f, e, ystar, ycut, y)
            f           e        ystar        ycut y
1   0.7964754  1.24489371  1.882074039 (0.25, Inf] 1
2   0.9673793 -0.96534738 -0.191443928 (-Inf,0.25] 0
3   0.5404655 -0.42555126  0.006821144 (-Inf,0.25] 0
4   1.2876865 -0.76640969  0.263739523 (0.25, Inf] 1
5  -1.2772009 -0.04408187 -1.065842624 (-Inf,0.25] 0
6   0.3688038  0.62856449  0.923607537 (0.25, Inf] 1
7   0.1279525  0.15187945  0.254241457 (0.25, Inf] 1
8   0.3892638  0.90361548  1.215026503 (0.25, Inf] 1
9  -0.5727820  0.24722949 -0.210996113 (-Inf,0.25] 0
10 -0.4322062 -0.12711528 -0.472880265 (-Inf,0.25] 0
```

คุณมีความรับผิดชอบ
(ไม่เห็นด้วยอย่างยิ่ง / ไม่เห็นด้วย / เห็นด้วย / เห็นด้วยอย่างยิ่ง)

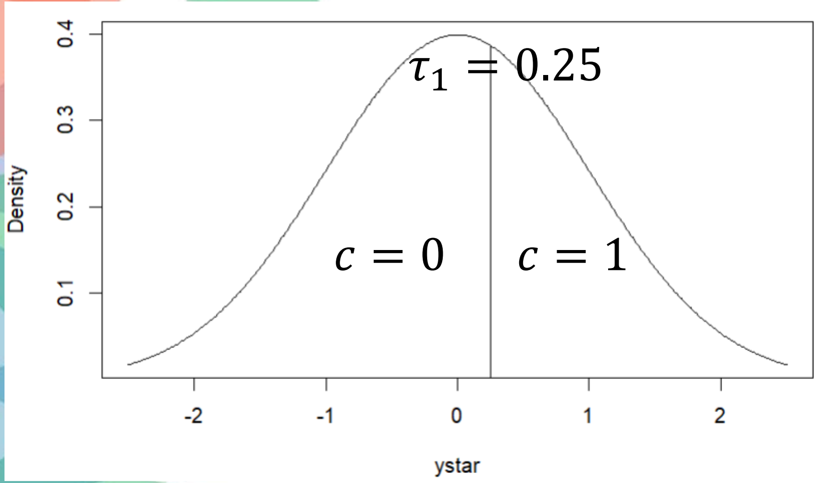
จุดเปลี่ยน = -1.25, -0.15, 0.90

```
> set.seed(113113)
> f <- rnorm(10, 0, 1)
> e <- rnorm(10, 0, 0.6)
> ystar <- 0.8*f + e
> ycut <- cut(ystar, c(-Inf, -1.25, -0.15, 0.90, Inf))
> y <- as.numeric(ycut) - 1
> data.frame(f, e, ystar, ycut, y)
            f           e       ystar          ycut y
1   0.6971884 -0.24077974  0.31697100   (-0.15,0.9] 2
2  -0.2086236  1.04486490  0.87796602   (-0.15,0.9] 2
3   0.7894843  0.40425715  1.03584460    (0.9, Inf] 3
4  -0.9645252 -1.50746738 -2.27908750  (-Inf,-1.25] 0
5  -1.7787699 -0.76340560 -2.18642148  (-Inf,-1.25] 0
6  -0.1376532 -0.33488174 -0.44500431 (-1.25,-0.15] 1
7   1.2409768 -0.45733803  0.53544342   (-0.15,0.9] 2
8  -0.3898191  0.31775020  0.00589495   (-0.15,0.9] 2
9   0.9007824 -0.07124724  0.64937870   (-0.15,0.9] 2
10  0.1700058  0.64739970  0.78340437   (-0.15,0.9] 2
```

$$y^* = 0.8F + e$$

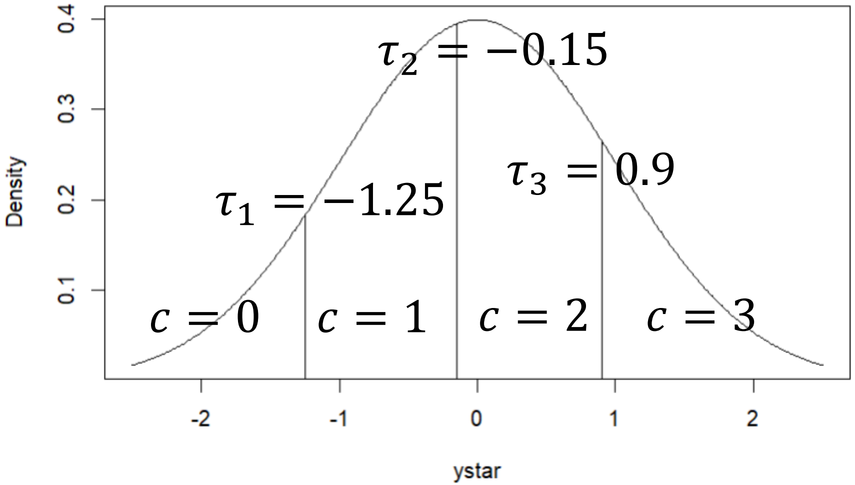
คุณมีความรับผิดชอบ
(ไม่ใช่ / ใช่)

จุดเปลี่ยน = 0.25



คุณมีความรับผิดชอบ
(ไม่เห็นด้วยอย่างยิ่ง / ไม่เห็นด้วย / เห็นด้วย / เห็นด้วยอย่างยิ่ง)

จุดเปลี่ยน = -1.25, -0.15, 0.90

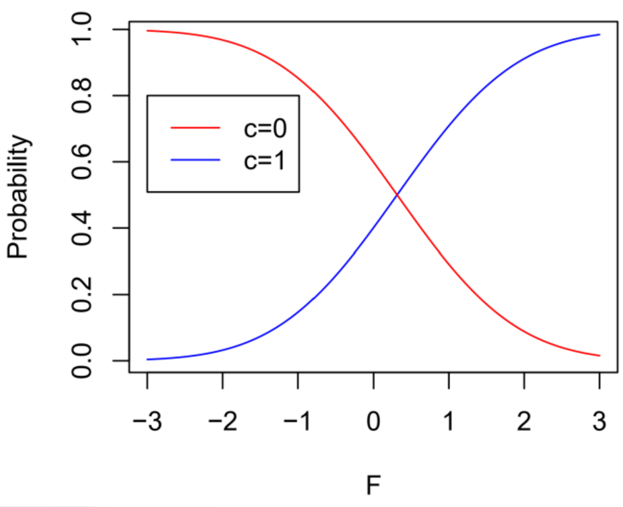


$$y^* = 0.8F + e$$

e จะเป็นส่วนค่าคงเหลือที่ทำให้ค่า *F* หนึ่งบางครั้งอาจเกินจุดเปลี่ยนหรือบางครั้งอาจต่ำกว่าจุดเปลี่ยน
หาโอกาสที่จะออกค่า *c* แต่ละค่า ในแต่ละระดับของ *F*

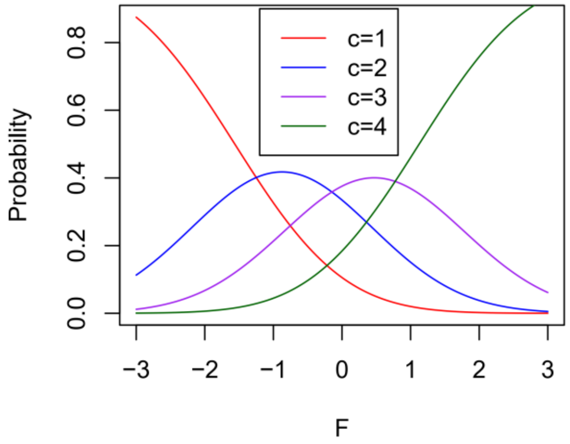
คุณมีความรับผิดชอบ
(ไม่ใช่ / ใช่)

จุดเปลี่ยน = 0.25



คุณมีความรับผิดชอบ
(ไม่เห็นด้วยอย่างยิ่ง / ไม่เห็นด้วย / เห็นด้วย / เห็นด้วยอย่างยิ่ง)

จุดเปลี่ยน = -1.25, -0.15, 0.90



ตัวบ่งชี้แบบจัดอันดับ

- ให้ y^* เป็นตัวแปรแฝงที่อยู่ใต้ตัวบ่งชี้แบบจัดกลุ่ม y ที่มีจำนวนกลุ่ม *C* กลุ่ม มีค่า *c* = 0, 1, … , *C* – 1 แล้ว โมเดล CFA สำหรับตัวบ่งชี้แบบจัดกลุ่มแบบจัดอันดับจะเป็นดังนี้

$$\boldsymbol{y}^* = \boldsymbol{\Lambda}\mathbf{f} + \mathbf{e} \qquad \mathbf{f} \sim \mathrm{MVN}(0, \boldsymbol{\Psi}) \qquad \mathbf{e} \sim \mathrm{MVN}(0, \boldsymbol{\Theta})$$

$$y_j = c \quad \text{if} \quad \tau_{j,c-1} < y_j^* < \tau_{j,c} \qquad Var\left(y_j^*\right) = 1 \qquad \tau_0 = -\infty \qquad \tau_C = \infty$$

- สเกลขององค์ประกอบ สามารถกำหนดได้ทั้ง Fixed Factor Approach หรือ Marker Variable Approach
- สเกลของ y^* กำหนดโดยให้ $Var\left(y_j^*\right) = 1$ ทำให้ไม่ได้ประมาณค่า $Var\left(e_j\right)$ หรือ θ_{jj}
- ค่าเฉลี่ยขององค์ประกอบ กำหนดให้เป็น 0 และจุดตัด (Intercept; ν_j) ของ y_j^* กำหนดให้เป็น 0

ตัวบ่งชี้แบบจัดอันดับ

- ให้ y^* เป็นตัวแปรแฝงที่อยู่ใต้ตัวบ่งชี้แบบจัดกลุ่ม y ที่มีจำนวนกลุ่ม *C* กลุ่ม มีค่า *c* = 0, 1, … , *C* – 1 แล้ว โมเดล CFA สำหรับตัวบ่งชี้แบบจัดกลุ่มแบบจัดอันดับจะเป็นดังนี้

$$\boldsymbol{y}^* = \mathbf{\Lambda f} + \mathbf{e} \qquad \mathbf{f} \sim \mathrm{MVN}(0, \mathbf{\Psi}) \qquad \mathbf{e} \sim \mathrm{MVN}(0, \mathbf{\Theta})$$

$$y_j = c \quad \text{if} \quad \tau_{j,c-1} < y_j^* < \tau_{j,c} \qquad Var\left(y_j^*\right) = 1 \qquad \tau_0 = -\infty \qquad \tau_C = \infty$$

- ส่วนโมเดลค่าเฉลี่ยจะถูกผลักให้ไปอยู่ในจุดเปลี่ยนทั้งหมด เนื่องจากค่าเฉลี่ยขององค์ประกอบกำหนดให้เป็น 0 และจุดตัด (Intercept; ν_j) ของ y_j^* กำหนดให้เป็น 0 เช่นกัน
- การกำหนดค่าคงที่ในโมเดลแบบนี้ เรียกว่า การกำหนดพารามิเตอร์แบบเดลต้า (Delta Parameterization)

ตัวบ่งชี้แบบจัดอันดับ

- กำหนดพารามิเตอร์อีกแบบหนึ่ง จะให้ $Var(e_j) = \theta_{jj} = 1$ โดยไม่กำหนด $Var(y_j^*)$ การกำหนดพารามิเตอร์แบบนี้เรียกว่า การกำหนดพารามิเตอร์แบบเทต้า (Theta Parameterization)

$$\boldsymbol{y}^* = \boldsymbol{\Lambda}\mathbf{f} + \mathbf{e} \qquad \mathbf{f} \sim \mathrm{MVN}(0, \boldsymbol{\Psi}) \qquad \mathbf{e} \sim \mathrm{MVN}(0, \boldsymbol{\Theta})$$

$$y_j = c \quad \text{if} \quad \tau_{j,c-1} < y_j^* < \tau_{j,c} \qquad \theta_{jj} = 1 \qquad \tau_0 = -\infty \qquad \tau_C = \infty$$

- Theta Parameterization จะใช้ในตอนเปรียบเทียบความคงทนของคุณสมบัติการวัด (Measurement Invariance) ระหว่างกลุ่ม เวลา ฯลฯ เนื่องจากใน Metric Invariance ต้องกำหนดให้น้ำหนักองค์ประกอบเท่ากัน แต่ต้องยอมรับว่าขนาดความแปรปรวนของค่าคงเหลือของทั้งสองกลุ่มแตกต่างกัน จึงต้องกำหนดให้ $\theta_{jj} = 1$ ในกลุ่มหนึ่ง และประมาณค่าอย่างอิสระในอีกกลุ่มหนึ่ง
- ค่าปกติจะเป็น delta parameterization เพราะเข้าใจง่ายกว่าเยอะ

ตัวบ่งชี้แบบจัดอันดับ

- สมมติว่าโครงสร้างองค์ประกอบเป็นแบบง่ายมาก (Very Simple Structure) ใช้ Fixed Factor Approach และ Delta Parameterization จำนวนพารามิเตอร์จะเป็นดังนี้
  - ความสัมพันธ์ระหว่างองค์ประกอบ (กรณีมากกว่า 1 องค์ประกอบ) : $k(k-1)/2$
  - น้ำหนักองค์ประกอบ เท่ากับจำนวนตัวบ่งชี้ : J
  - จำนวนจุดเปลี่ยน : $\sum_{j=1}^{J} C_j - 1$
- จำนวนข้อมูลเท่ากับจำนวนสหสัมพันธ์ที่เป็นไปได้ระหว่างตัวบ่งชี้ : $J(J-1)/2$ รวมกับจำนวนจุดเปลี่ยนรวมจากแต่ละข้อ $\sum_{j=1}^{J} C_j - 1$
- *df* = จำนวนข้อมูล - จำนวนพารามิเตอร์

ตัวบ่งชี้แบบจัดอันดับ

- การประมาณค่าจะไม่สามารถประมาณค่าผ่าน Maximum Likelihood ได้แล้ว เนื่องจากการกระจายของตัวบ่งชี้ไม่ได้เป็นตัวแปรต่อเนื่องที่มีการกระจายเป็น Multivariate Normal Distribution
- วิธีการประมาณค่าพารามิเตอร์ที่ไม่สนใจการกระจายของตัวบ่งชี้ที่เหมาะสมในกรณีนี้ คือ การทำให้ผลต่างกำลังสองต่ำสุดแบบถ่วงน้ำหนัก (Weighted Least Square; WLS)
- ให้ $\boldsymbol{v}$ และ $\boldsymbol{w}$ เป็นเวคเตอร์ที่มีสมาชิกเป็นจุดเปลี่ยนและสหสัมพันธ์ระหว่าง $\boldsymbol{y}^*$ ของกลุ่มตัวอย่าง (Saturated Model) และสร้างจากโมเดล (Model-implied) แล้ว ค่า Discrepancy Function สามารถนิยามได้ดังนี้

$$F_{WLS}(\boldsymbol{\theta}) = (\boldsymbol{v} - \boldsymbol{w})'\mathbf{W}^{-1}(\boldsymbol{v} - \boldsymbol{w})$$

- โดย **W** เป็นเมทริกซ์ความแปรปรวนร่วมของ Sampling Distribution ของ $\boldsymbol{v}$ โดยรากที่สองของสมาชิกแนวทแยงมุมของ **W** คือ ค่าความผิดพลาดมาตรฐาน (Standard Errors) ของจุดเปลี่ยนและสหสัมพันธ์ระหว่างตัวบ่งชี้

ตัวบ่งชี้แบบจัดอันดับ

- หากมีกลุ่มตัวอย่างเยอะมากๆ WLS จะสามารถประมาณค่าพารามิเตอร์ได้ถูกต้อง
- อย่างไรก็ตาม WLS ในกลุ่มตัวอย่างขนาดน้อยจะไม่เหมาะสม สมมติว่ามีตัวบ่งชี้ทวินาม 10 ตัว จะมีจุดเปลี่ยน 10 ตัว ความแปรปรวนร่วม 45 ตัว **W** จะมีขนาด 55 x 55 ซึ่งการหา Inverse ของเมทริกซ์ขนาด 55 x 55 เป็นเรื่องที่เปลืองทรัพยากรในการคำนวณมาก ค่าที่ได้มีโอกาสที่ไม่เสถียรถ้ากลุ่มตัวอย่างน้อย
- Muthén (1984) จึงเสนอให้ใช้เฉพาะสมาชิกแนวทแยงมุมของ **W** มาใช้ในการประมาณค่าพารามิเตอร์ เรียกว่า Diagonally Weighted Least Square (DWLS) ซึ่งจะลดจำนวนสมาชิกให้น้อยลงมาก การหา Matrix Inverse ก็ง่ายมากโดยการกลับเศษเป็นส่วนเท่านั้น
- ค่า Chi-square จะสามารถคำนวณได้ดังนี้

$$\chi^2 = N \cdot F_{WLS}(\boldsymbol{\theta})$$

ตัวบ่งชี้แบบจัดอันดับ

- เมื่อการกระจายของตัวบ่งชี้ไม่ได้เป็น MVN และกลุ่มตัวอย่างขนาดน้อย ค่า Chi-square เมื่อ Null Hypothesis เป็นจริงจะไม่ได้มีการกระจายเป็น Chi-square นักวิจัยจึงเสนอค่า Scaled Chi-squa re เพื่อปรับค่าของ Chi-square ให้มีการกระจายเป็น Chi-square Distribution มากขึ้น เมื่อตัวบ่งชี้ไม่ได้เป็น MVN (Savalei & Rosseel, 2022)
  - Satorra & Bentler (1995) เสนอสูตร Scaled Chi-square ปรับให้มีค่าเฉลี่ยเท่ากับ *df*
  - Asparouhov & Muthén (2010) เสนอสูตร Scaled Chi-square ปรับให้ค่าเฉลี่ยเท่ากับ *df* และความแปรปรวนเท่ากับ 2*df* ตาม Chi-square distribution ด้วย
- ค่าปกติของ lavaan จะใช้ DWLS และ Scaled Chi-square ตามสูตรของ Asparouhov & Muthén (2010) ซึ่งตรงกับการกำหนด ESTIMATOR = WLSMV ใน Mplus ซึ่งเป็นวิธีที่ได้รับความนิยมในปัจจุบัน

ตัวบ่งชี้แบบจัดอันดับ

- อย่างไรก็ตาม DWLS จะไม่สามารถจัดการค่าสูญหายได้ ไม่เหมือนกับ ML ที่สามารถปรับสูตรให้สนใจเฉพาะค่าที่ไม่สูญหายในแต่ละ Case ได้ ดังนั้นการแทนค่าสูญหายสำหรับ Categorical CFA จึงควรใช้แบบ Listwise หรือ Multiple Imputation (ซึ่งดีกว่า) เท่านั้น
- บางครั้ง WLS จะเรียกว่า Asymptotic Distribution Free (ADF) เพราะไม่สนใจการกระจายของตัวบ่งชี้เลย แต่ต้องใช้กลุ่มตัวอย่างจำนวนมาก ($N > 5000$) ถึงจะได้การประมาณค่าพารามิเตอร์ที่ถูกต้อง และ SE ที่ถูกต้อง (จึงเป็นสาเหตุที่ Muthén เสนอ DWLS)
- ULS เป็นอีกหนึ่งทางเลือกในการประมาณค่าพารามิเตอร์ ที่ **W** = **I** แต่ไม่ค่อยได้รับความนิยม
- ดูรายละเอียดเพิ่มเติมได้ที่ rosseel_sem_cat.pdf (personality-project.org)

```
> head(datdicho)
  y1 y2 y3 y4 y5 y6 y7 y8
1  1  2  2  2  2  1  1  1
2  2  1  2  2  2  1  1  1
3  2  1  1  2  2  1  1  2
4  2  2  2  2  2  1  1  1
5  1  1  2  2  2  2  2  2
6  1  2  2  1  1  1  1  1
```

ตัวแปรเป็นแบบทวินาท มี 2 ค่า
(1 = ไม่ใช่, 2 = ใช่)

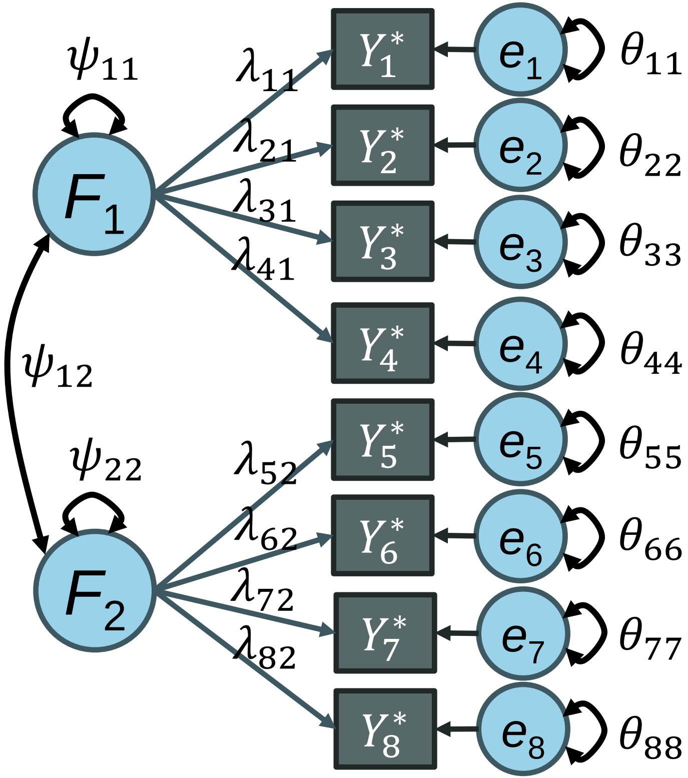
กำหนดว่าเป็นตัวบ่งชี้แบบจัดอันดับ

```
> mdicho <- '
+ f1 =~ y1 + y2 + y3 + y4
+ f2 =~ y5 + y6 + y7 + y8
+ '
> outdicho <- cfa(mdicho, data=datdicho, ordered=TRUE)
> summary(outdicho, std=TRUE, fit=TRUE)
lavaan 0.6-12 ended normally after 29 iterations

  Estimator                                         DWLS
  Optimization method                             NLMINB
  Number of model parameters                          17

  Number of observations                             200

Model Test User Model:
                                              Standard      Robust
  Test Statistic                                24.102      27.909
  Degrees of freedom                                19          19
  P-value (Chi-square)                           0.192       0.085
  Scaling correction factor                                  0.961
  Shift parameter                                            2.835
    simple second-order correction
```

$$\chi^2(19, N = 200) = 27.91, p = .085$$

พารามิเตอร์:
ความแปรปรวนร่วม 1 ตัว,
น้ำหนักองค์ประกอบ 8 ตัว,
จุดเปลี่ยน 8 ตัว
รวม 17 ตัว

ดู Scaled Chi-square



ข้อมูล: สหสัมพันธ์ (8)(7)/2 = 28, จุดเปลี่ยน 8 ตัว รวม 36 ตัว

df = 36 – 17 = 19

```
User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    0.985       0.967
  Tucker-Lewis Index (TLI)                       0.977       0.951

  Robust Comparative Fit Index (CFI)                            NA
  Robust Tucker-Lewis Index (TLI)                               NA

Root Mean Square Error of Approximation:

  RMSEA                                          0.037       0.049
  90 Percent confidence interval - lower         0.000       0.000
  90 Percent confidence interval - upper         0.076       0.085
  P-value RMSEA <= 0.05                          0.665       0.487

  Robust RMSEA                                                  NA
  90 Percent confidence interval - lower                     0.000
  90 Percent confidence interval - upper                        NA

Standardized Root Mean Square Residual:

  SRMR                                           0.109       0.109
```

แม้ว่าดัชนีความเหมาะสมจะเหมือนกับกรณีตัวแปรต่อเนื่อง แต่จุดตัดที่หามาจากตัวแปรต่อเนื่องไม่เหมาะสมกับตัวแปรจัดกลุ่มแบบจัดอันดับ (Nye & Drasgow, 2010)

อย่างไรก็ตาม คนก็ใช้จุดตัดจากตัวแปรต่อเนื่องอยู่ดี

ดังนั้น DFI น่าจะเหมาะสมในกรณีนี้มากที่สุด แต่ฟังก์ชั่น cfaHB ในปัจจุบัน (ต.ค. 2566) ยังทำได้เฉพาะการประมาณค่าแบบ ML

RMSEA = .049, SRMR = .109, CFA = .967, TLI = .951

```
Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  f1 =~
    y1                1.000                               0.672    0.672
    y2                0.890    0.192    4.639    0.000    0.598    0.598
    y3                0.809    0.193    4.198    0.000    0.544    0.544
    y4                1.416    0.287    4.937    0.000    0.951    0.951
  f2 =~
    y5                1.000                               0.549    0.549
    y6                1.297    0.319    4.069    0.000    0.713    0.713
    y7                1.603    0.375    4.280    0.000    0.881    0.881
    y8                1.542    0.387    3.985    0.000    0.847    0.847
```

Standardized Loadings

แปลความหมายเหมือนปกติ ทุกตัวบ่งชี้มีค่าสูง

```
Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  f1 ~~
    f2                0.135    0.052    2.614    0.009    0.365    0.365
```

ค่าสหสัมพันธ์ระหว่างองค์ประกอบทางบวกอย่างมีนัยสำคัญ

```
Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
   .y1                0.000                               0.000    0.000
   .y2                0.000                               0.000    0.000
   .y3                0.000                               0.000    0.000
   .y4                0.000                               0.000    0.000
   .y5                0.000                               0.000    0.000
   .y6                0.000                               0.000    0.000
   .y7                0.000                               0.000    0.000
   .y8                0.000                               0.000    0.000
    f1                0.000                               0.000    0.000
    f2                0.000                               0.000    0.000
```

ค่าเฉลี่ยองค์ประกอบ, จุดตัดตัวบ่งชี้มีค่าคงที่เท่ากับ 0

Thresholds:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
y1|t1	0.100	0.089	1.128	0.259	0.100	0.100
y2|t1	-0.385	0.091	-4.221	0.000	-0.385	-0.385
y3|t1	-0.628	0.096	-6.575	0.000	-0.628	-0.628
y4|t1	-0.739	0.098	-7.520	0.000	-0.739	-0.739
y5|t1	-0.994	0.107	-9.311	0.000	-0.994	-0.994
y6|t1	0.842	0.101	8.310	0.000	0.842	0.842
y7|t1	1.282	0.121	10.576	0.000	1.282	1.282
y8|t1	0.332	0.091	3.661	0.000	0.332	0.332

Variances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
.y1	0.549				0.549	0.549
.y2	0.643				0.643	0.643
.y3	0.705				0.705	0.705
.y4	0.096				0.096	0.096
.y5	0.698				0.698	0.698
.y6	0.492				0.492	0.492
.y7	0.224				0.224	0.224
.y8	0.283				0.283	0.283
f1	0.451	0.128	3.520	0.000	1.000	1.000
f2	0.302	0.126	2.394	0.017	1.000	1.000

จุดเปลี่ยนมีเพียงค่าเดียวในแต่ละตัวบ่งชี้ (เพราะเป็นตัวแปรทวินาม)

ค่าความแปรปรวนของค่าคงเหลือ แม้จะมีค่ามาให้ ได้จะถูกจำกัด เพราะกำหนด Var(y*) = 1

ความแปรปรวนขององค์ประกอบ มีการประมาณค่า เพราะใช้ Marker Variable Approach ในการกำหนดสเกลองค์ประกอบ

```
Scales y*:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
    y1                1.000                               1.000    1.000
    y2                1.000                               1.000    1.000
    y3                1.000                               1.000    1.000
    y4                1.000                               1.000    1.000
    y5                1.000                               1.000    1.000
    y6                1.000                               1.000    1.000
    y7                1.000                               1.000    1.000
    y8                1.000                               1.000    1.000
```

กำหนดสเกลของตัวแปรต่อเนื่องที่อยู่ภายใต้ของตัวบ่งชี้เท่ากับ 1

$Var(y^*) = 1$

```
> outdicho <- cfa(mdicho, data=datdicho, ordered=TRUE)
> outdicho2 <- cfa(mdicho, data=datdicho, ordered=TRUE, estimator="WLSMV")
> outdicho3 <- cfa(mdicho, data=datdicho, ordered=TRUE, estimator="DWLS",
+                  se="standard", test="scaled.shifted")
> outdicho3
lavaan 0.6-12 ended normally after 29 iterations

  Estimator                                         DWLS
  Optimization method                             NLMINB
  Number of model parameters                          17

  Number of observations                             200

Model Test User Model:
                                              Standard      Robust
  Test Statistic                                24.102      27.909
  Degrees of freedom                                19          19
  P-value (Chi-square)                           0.192       0.085
  Scaling correction factor                                  0.961
  Shift parameter                                            2.835
    simple second-order correction
```

การกำหนด arguments ทั้งสามรูปแบบได้ผลลัพธ์แบบเดียวกัน

สังเกตว่าค่าปกติของ Categorical CFA คือการกำหนด estimator="WLSMV" หรือการกำหนดตัวเลือกอย่างละเอียด คือ

- estimator = "DWLS"
- se = "standard"
- test = "scaled.shifted" คือ ปรับแบบ Asparouhov & Muthen (2010)

```
> outdichotheta <- cfa(mdicho, data=datdicho, ordered=TRUE,
+                      parameterization="theta")
> summary(outdichotheta, fit=TRUE, std=TRUE)
lavaan 0.6-12 ended normally after 75 iterations

  Estimator                                         DWLS
  Optimization method                             NLMINB
  Number of model parameters                          17

  Number of observations                             200

Model Test User Model:
                                              Standard      Robust
  Test Statistic                                24.102      27.909
  Degrees of freedom                                19          19
  P-value (Chi-square)                           0.192       0.085
  Scaling correction factor                                  0.961
  Shift parameter                                            2.835
    simple second-order correction
```

กำหนดให้ใช้ theta parameterization

ค่าดัชนีความเหมาะสมเหมือนเดิมท

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
f1 =~						
y1	1.000				0.907	0.672
y2	0.822	0.297	2.766	0.006	0.746	0.598
y3	0.714	0.268	2.661	0.008	0.648	0.544
y4	3.390	3.972	0.854	0.393	3.074	0.951
f2 =~						
y5	1.000				0.657	0.549
y6	1.545	0.620	2.492	0.013	1.016	0.713
y7	2.828	1.605	1.762	0.078	1.859	0.881
y8	2.423	1.293	1.874	0.061	1.593	0.847

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
f1 ~~						
f2	0.218	0.109	1.995	0.046	0.365	0.365

Intercepts:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
.y1	0.000				0.000	0.000
.y2	0.000				0.000	0.000
.y3	0.000				0.000	0.000
.y4	0.000				0.000	0.000
.y5	0.000				0.000	0.000
.y6	0.000				0.000	0.000
.y7	0.000				0.000	0.000
.y8	0.000				0.000	0.000
f1	0.000				0.000	0.000
f2	0.000				0.000	0.000

Thresholds:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
y1|t1	0.136	0.121	1.120	0.263	0.136	0.100
y2|t1	-0.481	0.119	-4.053	0.000	-0.481	-0.385
y3|t1	-0.748	0.123	-6.065	0.000	-0.748	-0.628
y4|t1	-2.388	2.259	-1.057	0.290	-2.388	-0.739
y5|t1	-1.190	0.156	-7.620	0.000	-1.190	-0.994
y6|t1	1.200	0.217	5.522	0.000	1.200	0.842
y7|t1	2.705	1.029	2.629	0.009	2.705	1.282
y8|t1	0.624	0.248	2.514	0.012	0.624	0.332

Variances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
.y1	1.000				1.000	0.549
.y2	1.000				1.000	0.643
.y3	1.000				1.000	0.705
.y4	1.000				1.000	0.096
.y5	1.000				1.000	0.698
.y6	1.000				1.000	0.492
.y7	1.000				1.000	0.224
.y8	1.000				1.000	0.283
f1	0.822	0.426	1.932	0.053	1.000	1.000
f2	0.432	0.258	1.672	0.095	1.000	1.000

$$\theta_{jj} = 1$$

Scales y*:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
y1	0.741				0.741	1.000
y2	0.802				0.802	1.000
y3	0.839				0.839	1.000
y4	0.309				0.309	1.000
y5	0.836				0.836	1.000
y6	0.702				0.702	1.000
y7	0.474				0.474	1.000
y8	0.532				0.532	1.000

$$Var(y^*) \neq 1$$

ค่าพารามิเตอร์แตกต่างจากเดิมนิดหน่อย เพราะ y* มีความแปรปรวนไม่เท่ากับ 1 แต่ standardized estimates เหมือนเดิม

วัดองค์ประกอบเดียวกัน ใน 4 ช่วงเวลาในระยะห่างของเวลาใกล้เคียงกัน

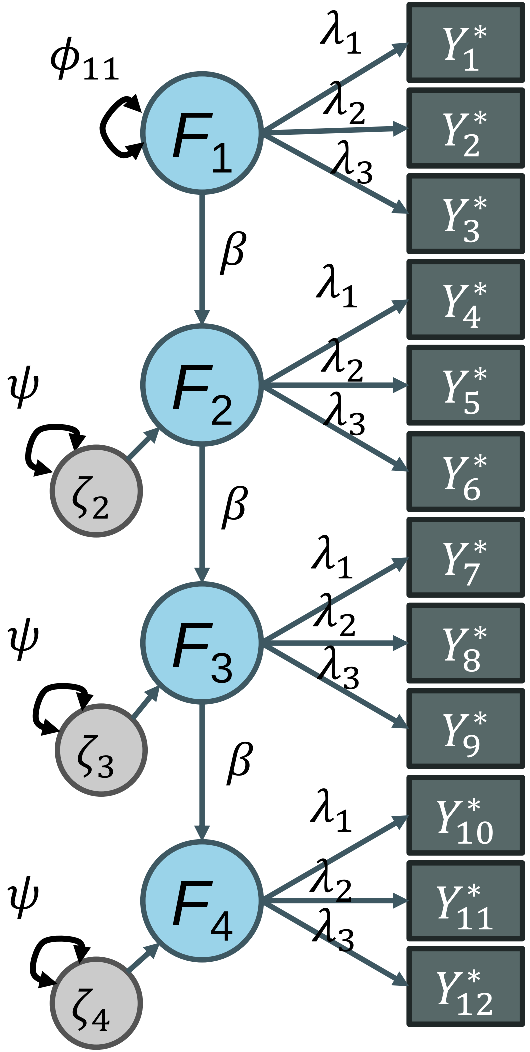
โมเดลที่ตัวแปรในแต่ละเวลามีอิทธิพลต่อตัวแปรในเวลาถัดไป เรียกว่า โมเดลวิเคราะห์ถดถอยต่อตนเอง (Autoregressive Models)

กรณีที่โมเดลมีอิทธิพลต่อกันเพียงแค่ 1 ช่วงเวลา จะเรียกว่า First-order autoregressive models หรือบางทีจะเรียกว่า simplex model แต่ในที่นี้จะเป็นในระดับองค์ประกอบ

เขียนสมการได้คือ $F_t = \beta F_{t-1} + \zeta_t$ โดย $Var(\zeta_t) = \psi$
(โมเดลเต็มจะมีจุดตัด เพื่อจัดการการเปลี่ยนแปลงของค่าเฉลี่ย)

```
> library(lavaan)
> datlikert <- read.table("lecture12ex2.csv", sep=",", header=TRUE)
> head(datlikert)
  y1 y2 y3 y4 y5 y6 y7 y8 y9 y10 y11 y12
1  1  3  2  3  4  3  4  3  4   3   3   3
2  3  4  3  2  2  3  2  2  3   4   4   2
3  2  2  2  1  3  1  3  3  3   3   2   3
4  3  3  2  2  1  2  2  2  3   3   2   2
5  4  4  3  4  4  2  3  3  4   3   4   2
6  2  2  2  3  4  3  2  2  3   1   1   1
```

ตัวแปรแต่ละตัวมีค่า 1,2,3,4



สร้างโมเดล CFA

```
> mlikert <- '
+ f1 =~ y1 + y2 + y3
+ f2 =~ y4 + y5 + y6
+ f3 =~ y7 + y8 + y9
+ f4 =~ y10 + y11 + y12
+ '
> outlikert <- cfa(mlikert, data=datlikert, ordered=TRUE)
> summary(outlikert, fit=TRUE, std=TRUE)
lavaan 0.6-12 ended normally after 32 iterations

  Estimator                                         DWLS
  Optimization method                             NLMINB
  Number of model parameters                          54

  Number of observations                             700

Model Test User Model:
                                              Standard      Robust
  Test Statistic                                27.668      46.653
  Degrees of freedom                                48          48
  P-value (Chi-square)                           0.992       0.528
  Scaling correction factor                                  0.716
  Shift parameter                                            8.033
    simple second-order correction

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    1.000       1.000
  Tucker-Lewis Index (TLI)                       1.005       1.001

  Robust Comparative Fit Index (CFI)                            NA
  Robust Tucker-Lewis Index (TLI)                               NA

Root Mean Square Error of Approximation:

  RMSEA                                          0.000       0.000
  90 Percent confidence interval - lower         0.000       0.000
  90 Percent confidence interval - upper         0.000       0.024
  P-value RMSEA <= 0.05                          1.000       1.000

  Robust RMSEA                                                  NA
  90 Percent confidence interval - lower                     0.000
  90 Percent confidence interval - upper                        NA

Standardized Root Mean Square Residual:

  SRMR                                           0.025       0.025
```

ใส่ ordered=TRUE เพื่อระบุว่าตัวบ่งชี้เป็นแบบจัดอันดับ

$$\chi^2(48, N = 700) = 46.65, p = .53$$

โมเดลเหมาะสมกับข้อมูลดี

RMSEA = 0, SRMR = .025, CFI = 1.00, TLI = 1.00

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
f1 =~						
y1	1.000				0.715	0.715
y2	1.117	0.054	20.649	0.000	0.799	0.799
y3	1.237	0.066	18.826	0.000	0.885	0.885
f2 =~						
y4	1.000				0.920	0.920
y5	0.609	0.051	11.885	0.000	0.560	0.560
y6	0.640	0.053	11.970	0.000	0.589	0.589
f3 =~						
y7	1.000				0.812	0.812
y8	0.760	0.054	14.164	0.000	0.617	0.617
y9	0.890	0.056	16.017	0.000	0.722	0.722
f4 =~						
y10	1.000				0.855	0.855
y11	0.829	0.050	16.578	0.000	0.709	0.709
y12	0.791	0.049	16.145	0.000	0.677	0.677

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
f1 ~~						
f2	0.294	0.031	9.477	0.000	0.448	0.448
f3	0.151	0.029	5.218	0.000	0.261	0.261
f4	0.115	0.030	3.890	0.000	0.189	0.189
f2 ~~						
f3	0.403	0.034	11.737	0.000	0.541	0.541
f4	0.259	0.037	7.069	0.000	0.330	0.330
f3 ~~						
f4	0.394	0.033	11.911	0.000	0.567	0.567

Intercepts:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
.y1	0.000				0.000	0.000
.y2	0.000				0.000	0.000
.y3	0.000				0.000	0.000
.y4	0.000				0.000	0.000
.y5	0.000				0.000	0.000
.y6	0.000				0.000	0.000
.y7	0.000				0.000	0.000
.y8	0.000				0.000	0.000
.y9	0.000				0.000	0.000
.y10	0.000				0.000	0.000
.y11	0.000				0.000	0.000
.y12	0.000				0.000	0.000
f1	0.000				0.000	0.000
f2	0.000				0.000	0.000
f3	0.000				0.000	0.000
f4	0.000				0.000	0.000

จุดตัดเป็น 0

ค่าเฉลี่ยองค์ประกอบ
เป็น 0

Thresholds:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
y1|t1	-0.937	0.056	-16.796	0.000	-0.937	-0.937
y1|t2	0.000	0.047	0.000	1.000	0.000	0.000
y1|t3	1.043	0.058	17.955	0.000	1.043	1.043
y2|t1	-1.211	0.062	-19.387	0.000	-1.211	-1.211
y2|t2	-0.298	0.048	-6.185	0.000	-0.298	-0.298
y2|t3	0.674	0.052	13.087	0.000	0.674	0.674
y3|t1	-1.425	0.070	-20.415	0.000	-1.425	-1.425
y3|t2	0.021	0.047	0.453	0.650	0.021	0.021
y3|t3	1.405	0.069	20.357	0.000	1.405	1.405
y4|t1	-0.706	0.052	-13.593	0.000	-0.706	-0.706
y4|t2	0.239	0.048	4.981	0.000	0.239	0.239
y4|t3	1.250	0.064	19.635	0.000	1.250	1.250
y5|t1	-1.250	0.064	-19.635	0.000	-1.250	-1.250
y5|t2	-0.309	0.048	-6.411	0.000	-0.309	-0.309
y5|t3	0.693	0.052	13.376	0.000	0.693	0.693
y6|t1	-1.087	0.059	-18.381	0.000	-1.087	-1.087
y6|t2	0.180	0.048	3.775	0.000	0.180	0.180
y6|t3	1.197	0.062	19.283	0.000	1.197	1.197
y7|t1	-1.368	0.068	-20.225	0.000	-1.368	-1.368
y7|t2	0.061	0.047	1.284	0.199	0.061	0.061
y7|t3	0.821	0.054	15.296	0.000	0.821	0.821
y8|t1	-0.412	0.049	-8.435	0.000	-0.412	-0.412
y8|t2	0.496	0.050	10.000	0.000	0.496	0.496
y8|t3	1.405	0.069	20.357	0.000	1.405	1.405
y9|t1	-0.943	0.056	-16.863	0.000	-0.943	-0.943
y9|t2	0.250	0.048	5.207	0.000	0.250	0.250
y9|t3	1.227	0.063	19.488	0.000	1.227	1.227
y10|t1	-1.386	0.068	-20.294	0.000	-1.386	-1.386
y10|t2	-0.476	0.049	-9.628	0.000	-0.476	-0.476
y10|t3	0.496	0.050	10.000	0.000	0.496	0.496
y11|t1	-1.377	0.068	-20.260	0.000	-1.377	-1.377
y11|t2	-0.202	0.048	-4.227	0.000	-0.202	-0.202
y11|t3	0.837	0.054	15.504	0.000	0.837	0.837
y12|t1	-0.883	0.055	-16.123	0.000	-0.883	-0.883
y12|t2	0.294	0.048	6.110	0.000	0.294	0.294
y12|t3	1.161	0.061	19.014	0.000	1.161	1.161

ตัวบ่งชี้แต่ละตัว
มี 3 จุดเปลี่ยน

Variances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
.y1	0.489				0.489	0.489
.y2	0.362				0.362	0.362
.y3	0.217				0.217	0.217
.y4	0.154				0.154	0.154
.y5	0.686				0.686	0.686
.y6	0.654				0.654	0.654
.y7	0.341				0.341	0.341
.y8	0.620				0.620	0.620
.y9	0.478				0.478	0.478
.y10	0.269				0.269	0.269
.y11	0.497				0.497	0.497
.y12	0.542				0.542	0.542
f1	0.511	0.041	12.477	0.000	1.000	1.000
f2	0.846	0.063	13.405	0.000	1.000	1.000
f3	0.659	0.049	13.468	0.000	1.000	1.000
f4	0.731	0.047	15.499	0.000	1.000	1.000

ค่าคงเหลือ ถูกจำกัดเพื่อทำให้ $Var(y^*) = 1$

ความแปรปรวนองค์ประกอบได้ประมาณค่า เพราะใช้ Marker Variable Approach

Scales y*:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
y1	1.000				1.000	1.000
y2	1.000				1.000	1.000
y3	1.000				1.000	1.000
y4	1.000				1.000	1.000
y5	1.000				1.000	1.000
y6	1.000				1.000	1.000
y7	1.000				1.000	1.000
y8	1.000				1.000	1.000
y9	1.000				1.000	1.000
y10	1.000				1.000	1.000
y11	1.000				1.000	1.000
y12	1.000				1.000	1.000

$$Var(y^*) = 1$$

ทำโมเดลที่น้ำหนักองค์ประกอบเท่ากันทั้ง 4 ช่วงเวลา

(Metric Invariance)

```
> mlikert2 <- '
+ f1 =~ y1 + lambda2*y2 + lambda3*y3
+ f2 =~ y4 + lambda2*y5 + lambda3*y6
+ f3 =~ y7 + lambda2*y8 + lambda3*y9
+ f4 =~ y10 + lambda2*y11 + lambda3*y12
+ '
> outlikert2 <- cfa(mlikert2, data=datlikert, ordered=TRUE)
> summary(outlikert2, fit=TRUE, std=TRUE)
lavaan 0.6-12 ended normally after 28 iterations

  Estimator                                         DWLS
  Optimization method                             NLMINB
  Number of model parameters                          54
  Number of equality constraints                       6

  Number of observations                             700

Model Test User Model:
                                              Standard      Robust
  Test Statistic                                40.939      58.469
  Degrees of freedom                                54          54
  P-value (Chi-square)                           0.905       0.315
  Scaling correction factor                                  0.852
  Shift parameter                                           10.431
    simple second-order correction
```

$$\chi^2(54, N = 700) = 58.47, p = .32$$

โมเดลเหมาะสมกับข้อมูลดี

```
Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  f1 =~
    y1                1.000                               0.713    0.713
    y2      (lmb2)    1.157    0.040   28.780    0.000    0.825    0.825
    y3      (lmb3)    0.855    0.033   26.223    0.000    0.609    0.609
  f2 =~
    y4                1.000                               0.709    0.709
    y5      (lmb2)    1.157    0.040   28.780    0.000    0.820    0.820
    y6      (lmb3)    0.855    0.033   26.223    0.000    0.606    0.606
  f3 =~
    y7                1.000                               0.729    0.729
    y8      (lmb2)    1.157    0.040   28.780    0.000    0.843    0.843
    y9      (lmb3)    0.855    0.033   26.223    0.000    0.623    0.623
  f4 =~
    y10               1.000                               0.688    0.688
    y11     (lmb2)    1.157    0.040   28.780    0.000    0.796    0.796
    y12     (lmb3)    0.855    0.033   26.223    0.000    0.588    0.588
```

Unstandardized Loadings เท่ากันทุกช่วงเวลา

```
> anova(outlikert, outlikert2)
Scaled Chi-Squared Difference Test (method = “satorra.2000”)

lavaan NOTE:
    The “Chisq” column contains standard test statistics, not the
    robust test that should be reported per model. A robust difference
    test is a function of two standard (not robust) statistics.

           Df AIC BIC  Chisq Chisq diff Df diff Pr(>Chisq)
outlikert  48         33.425
outlikert2 54         40.939     5.4759       6     0.4844
```

ทำ Nested Model Comparison พบผลไม่แตกต่างกันอย่างมีนัยสำคัญ เลือกโมเดล Metric Invariance

```
> mlikert3 <- '
+ f1 =~ y1 + lambda2*y2 + lambda3*y3
+ f2 =~ y4 + lambda2*y5 + lambda3*y6
+ f3 =~ y7 + lambda2*y8 + lambda3*y9
+ f4 =~ y10 + lambda2*y11 + lambda3*y12
+ f2 ~ f1
+ f3 ~ f2
+ f4 ~ f3
+ '
> outlikert3 <- cfa(mlikert3, data=datlikert, ordered=TRUE)
> summary(outlikert3, fit=TRUE, std=TRUE)
lavaan 0.6-12 ended normally after 20 iterations

  Estimator                                         DWLS
  Optimization method                             NLMINB
  Number of model parameters                          51
  Number of equality constraints                       6

  Number of observations                             700

Model Test User Model:
                                              Standard      Robust
  Test Statistic                                55.516      70.414
  Degrees of freedom                                57          57
  P-value (Chi-square)                           0.531       0.109
  Scaling correction factor                                  0.951
  Shift parameter                                           12.034
    simple second-order correction
```

ปรับให้เป็น Autoregressive Model

$$\chi^2(57, N = 700) = 70.41, p = .11$$

โมเดลเหมาะสมกับข้อมูลดี

```
> anova(outlikert2, outlikert3)
Scaled Chi-Squared Difference Test (method = "satorra.2000")

lavaan NOTE:
    The "Chisq" column contains standard test statistics, not the
    robust test that should be reported per model. A robust difference
    test is a function of two standard (not robust) statistics.

           Df AIC BIC  Chisq Chisq diff Df diff Pr(>Chisq)
outlikert2 54          40.939
outlikert3 57          55.516     7.6174       3    0.05462 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  f2 ~
    f1                0.593    0.044   13.513    0.000    0.608    0.608
  f3 ~
    f2                0.586    0.046   12.779    0.000    0.564    0.564
  f4 ~
    f3                0.453    0.047    9.561    0.000    0.475    0.475
```

ผลการวิเคราะห์ถดถอยในแต่ละช่วงเวลา
ที่ติดกัน

ผลการทำ Nested Model Comparison
ไม่แตกต่างอย่างมีนัยสำคัญ เลือกโมเดล AR(1)

```
> mlikert4 <- '
+ f1 =~ y1 + lambda2*y2 + lambda3*y3
+ f2 =~ y4 + lambda2*y5 + lambda3*y6
+ f3 =~ y7 + lambda2*y8 + lambda3*y9
+ f4 =~ y10 + lambda2*y11 + lambda3*y12
+ f2 ~ beta*f1
+ f3 ~ beta*f2
+ f4 ~ beta*f3
+ f2 ~~ psi*f2
+ f3 ~~ psi*f3
+ f4 ~~ psi*f4
+ '
> outlikert4 <- cfa(mlikert4, data=datlikert, ordered=TRUE)
> summary(outlikert4, fit=TRUE, std=TRUE)
lavaan 0.6-12 ended normally after 16 iterations

  Estimator                                         DWLS
  Optimization method                             NLMINB
  Number of model parameters                          51
  Number of equality constraints                      10

  Number of observations                             700

Model Test User Model:
                                              Standard      Robust
  Test Statistic                                72.202      79.205
  Degrees of freedom                                61          61
  P-value (Chi-square)                           0.154       0.059
  Scaling correction factor                                  1.134
  Shift parameter                                           15.520
    simple second-order correction
```

ตรวจสอบ Stationary Assumption

$$\chi^2(61, N = 700) = 79.21, p = .059$$

โมเดลเหมาะสมกับข้อมูลดี

```
> anova(outlikert3, outlikert4)
Scaled Chi-Squared Difference Test (method = "satorra.2000")

lavaan NOTE:
    The "Chisq" column contains standard test statistics, not the
    robust test that should be reported per model. A robust difference
    test is a function of two standard (not robust) statistics.

           Df AIC BIC  Chisq Chisq diff Df diff Pr(>Chisq)
outlikert3 57         55.516
outlikert4 61         72.202     7.3776       4     0.1172
```

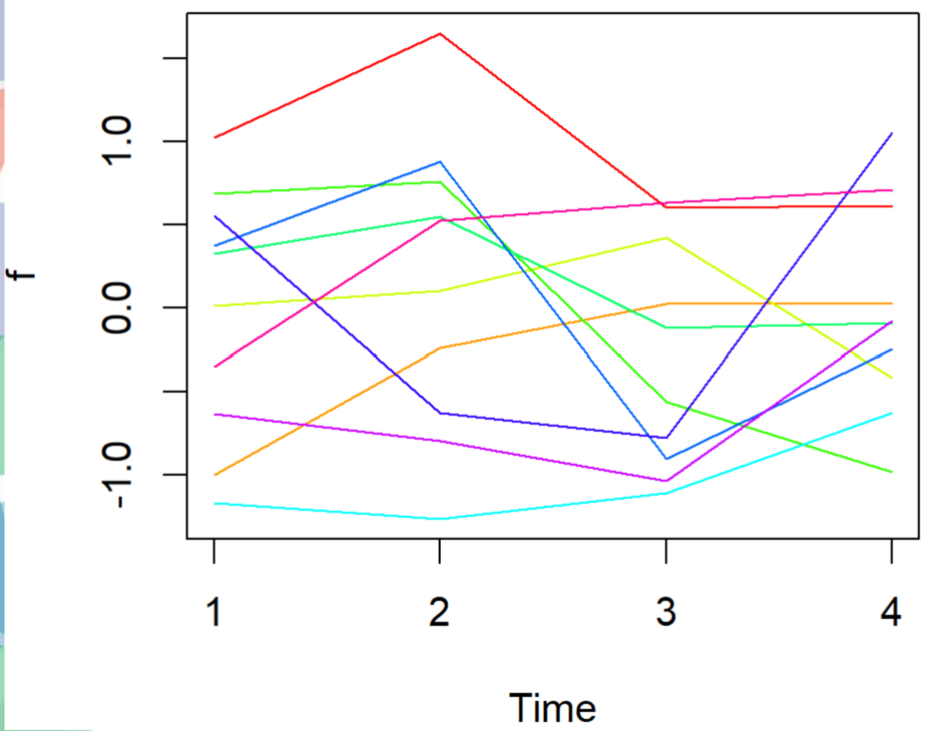
```
Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  f2 ~
    f1      (beta)    0.545    0.026   20.725    0.000    0.565    0.565
  f3 ~
    f2      (beta)    0.545    0.026   20.725    0.000    0.551    0.551
  f4 ~
    f3      (beta)    0.545    0.026   20.725    0.000    0.547    0.547
```

```
Variances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
   .f2       (psi)    0.338    0.021   15.840    0.000    0.681    0.681
   .f3       (psi)    0.338    0.021   15.840    0.000    0.696    0.696
   .f4       (psi)    0.338    0.021   15.840    0.000    0.701    0.701
```

Nested Model Comparison ไม่แตกต่างกันอย่างมีนัยสำคัญ ยอมรับ Stationary Assumption

ขนาดอิทธิพล = .545, ความแปรปรวน = .338

$$f_t = 0.545 f_{t-1} + e_t; e_t \sim N(0,0.581)$$



ตัวอย่างของข้อมูล 10 ชุด ที่มีการเปลี่ยนแปลงจากเวลาที่ 1 ไปยังเวลาที่ 4

ตัวบ่งชี้แบบจัดอันดับ

- การทำให้ตัวบ่งชี้เป็นแบบจัดอันดับจะทำให้
  - การวิเคราะห์ผลซับซ้อนมากขึ้น (เช่น การทดสอบความคงทนของการวัด [Measurement Invariance])
  - ไม่สามารถใช้ FIML ได้
  - ไม่สามารถใช้สถิติที่ออกแบบสำหรับตัวแปรต่อเนื่องบางอย่างได้ เช่น Vuong Test สำหรับการเปรียบเทียบโมเดล
- หากมีกลุ่มแบบจัดอันดับจำนวนมาก อาจเหมารวมให้ตัวบ่งชี้เป็นตัวแปรต่อเนื่องได้ โดย Rhemtulla และคณะ (2012) ได้ทำ Simulation Study พบว่าถ้ามีกลุ่ม 5 กลุ่มขึ้นไป (ที่มีข้อมูลคนตอบ) การใช้ ML สำหรับตัวแปรต่อเนื่อง จะประมาณค่าพารามิเตอร์ได้ถูกต้องเทียบเท่ากับกับใช้ CFA สำหรับตัวบ่งชี้แบบจัดอันดับ