Introduction to Confirmatory Factor Analysis and Structural Equation Modeling

Lecture 12
August 7, 2011

Advanced Multivariate Statistical Methods

ICPSR Summer Session #2

Today's Class

- An Introduction to:
  - Confirmatory Factor Analysis (CFA)
  - Structural Equation Modeling (SEM)
- Placing both within the linear modeling framework
  - The return of the multivariate normal distribution
- A Description of how CFA and EFA differ statistically
- Showing how these methods have subsumed canonical correlation analysis

A Brief Review of Exploratory Factor Analysis

- **EFA**: “Determine nature and number of latent variables that account for observed variation and covariation among set of observed indicators (≈ items or variables)”
  - In other words, what causes these observed responses?
  - Summarize patterns of correlation among indicators
  - Solution is an end (i.e., is of interest) in and of itself

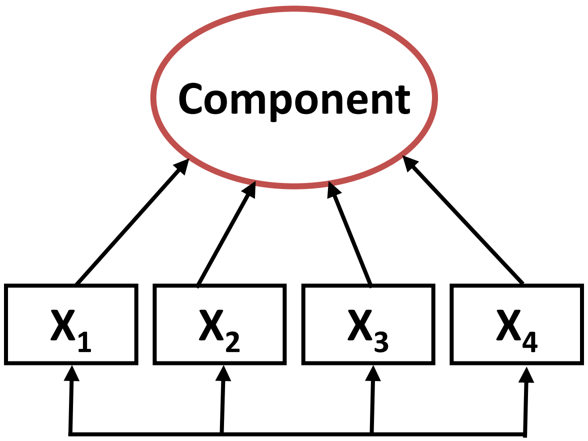
- **PCA**: “Reduce multiple observed variables into fewer components that summarize their variance”
  - In other words, how can I abbreviate this set of variables?
  - Solution is usually a means to an end

Big Conceptual Difference between PCA and EFA

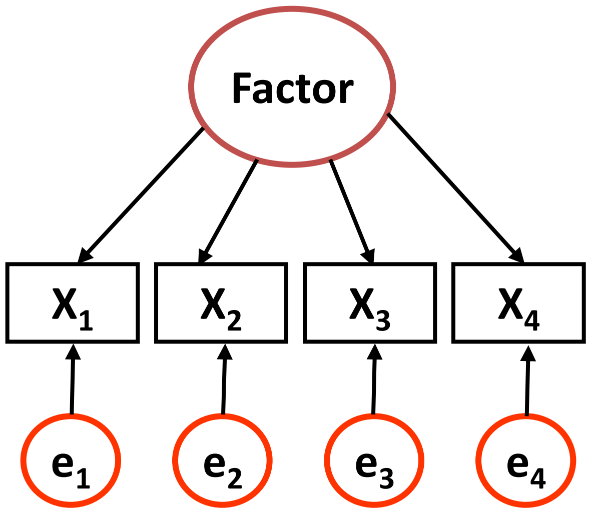
- In PCA, we get components that are **outcomes** built from linear combinations of the items:
  - $\mathbf{C_1} = L_{11}X_1 + L_{12}X_2 + L_{13}X_3 + L_{14}X_4 + L_{15}X_5$
  - $\mathbf{C_2} = L_{21}X_1 + L_{22}X_2 + L_{23}X_3 + L_{24}X_4 + L_{25}X_5$
  - … and so forth – note that C is the **OUTCOME**
    - **This is not a testable measurement model by itself**

- In EFA, we get factors that are thought to be the **cause** of the observed indicators (here, 5 indicators, 2 factors):
  - $X_1 = L_{11}\mathbf{F_1} + L_{12}\mathbf{F_2} + e_1$
  - $X_2 = L_{21}\mathbf{F_1} + L_{22}\mathbf{F_2} + e_1$
  - $X_3 = L_{31}\mathbf{F_1} + L_{32}\mathbf{F_2} + e_1$
  - … and so forth… but note that F is the **PREDICTOR → testable**

PCA vs. EFA/CFA



This is not a testable measurement model, because how do we know if we've combined items "correctly"?



This IS a testable measurement model, because we are trying to predict the observed covariances between the indicators by creating a factor – **the factor IS the reason for the covariance**

Big Conceptual Difference between PCA and EFA

- In **PCA**, the component is just the sum of the parts, and there is no inherent reason why the parts should be correlated (they just are)
  - But it's helpful if they are (otherwise, there's no point in trying to build components to summarize the variables)
    - "component" = "variable"
  - The type of construct measured by a component is often called an '**emergent**' construct – i.e., it emerges from the indicators ("**formative**")
  - Examples: "Lack of Free time", "SES", "Support/Resources"

- In **EFA**, the indicator responses are caused by the factors, and thus should be uncorrelated once controlling for the factor(s)
  - The type of construct that is measured by a factor is often called a '**reflective**' construct – i.e., the indicators are a reflection of your status on the latent variable
  - Examples: Any other hypothetical construct

Intermediate Summary…

- PCA and EFA are both exploratory techniques geared loosely towards examining the structure underneath a series of continuous indicators (items or subscales):
  - PCA: How do indicators linearly combine to produce a set of uncorrelated linear composite outcomes?
  - EFA: What is the structure of the latent factors that produced the covariances among the observed indicators (factor = predictor)?

- Involves sequence of sometimes ambiguous decisions:
  - Extraction method
  - Number of factors
  - And then: rotation, interpretation, and factor scores…

Factor Scores in EFA: Just Say No

- Factor Indeterminacy (e.g., Grice, 2001):
  - There is an infinite number of possible factor scores that all have the same mathematical characteristics
  - Different approaches can yield very different results
- A simple, yet effective solution is simply sum the items that load highly on a factor…"Unit-weighting"
  - Research has suggested that this 'simple' solution is more effective when applying the results of a factor analysis to different samples – factor loadings don't replicate all that well
  - Just make sure to standardize the indicators first if they are on different numerical scales
- Use CFA/SEM – you don't need the factor scores

CONFIRMATORY FACTOR ANALYSIS

Confirmatory Factor Analysis

- Rather than trying to determine the number of factors, and subsequently, what the factors mean (as in EFA), if you already know (or suspect) the structure of your data, you can use a confirmatory approach
- Confirmatory factor analysis (CFA) is a way to specify which variables load onto which factors
- The loadings of all variables not related to a given factor are then set to zero
- For a reasonable number of parameters, the factor correlation can be estimated directly from the analysis (rotations are not needed)

EFA vs. CFA, continued

- How we get an interpretable solution…
  - **EFA: Rotation**
    - All items load on all factors
    - Goal is to pick a rotation that gives closest approximation to simple structure (clear factors, fewest cross-loadings)
    - No way of separating ‘content’ from ‘method’ factors
  - **CFA: Your job in the first place!**
    - CFA must be theory-driven
    - You specify number of factors and their inter-correlations
    - You specify which items load on which factors (yes/no)
    - You specify any unique (error) relations for method variance

EFA vs. CFA, continued

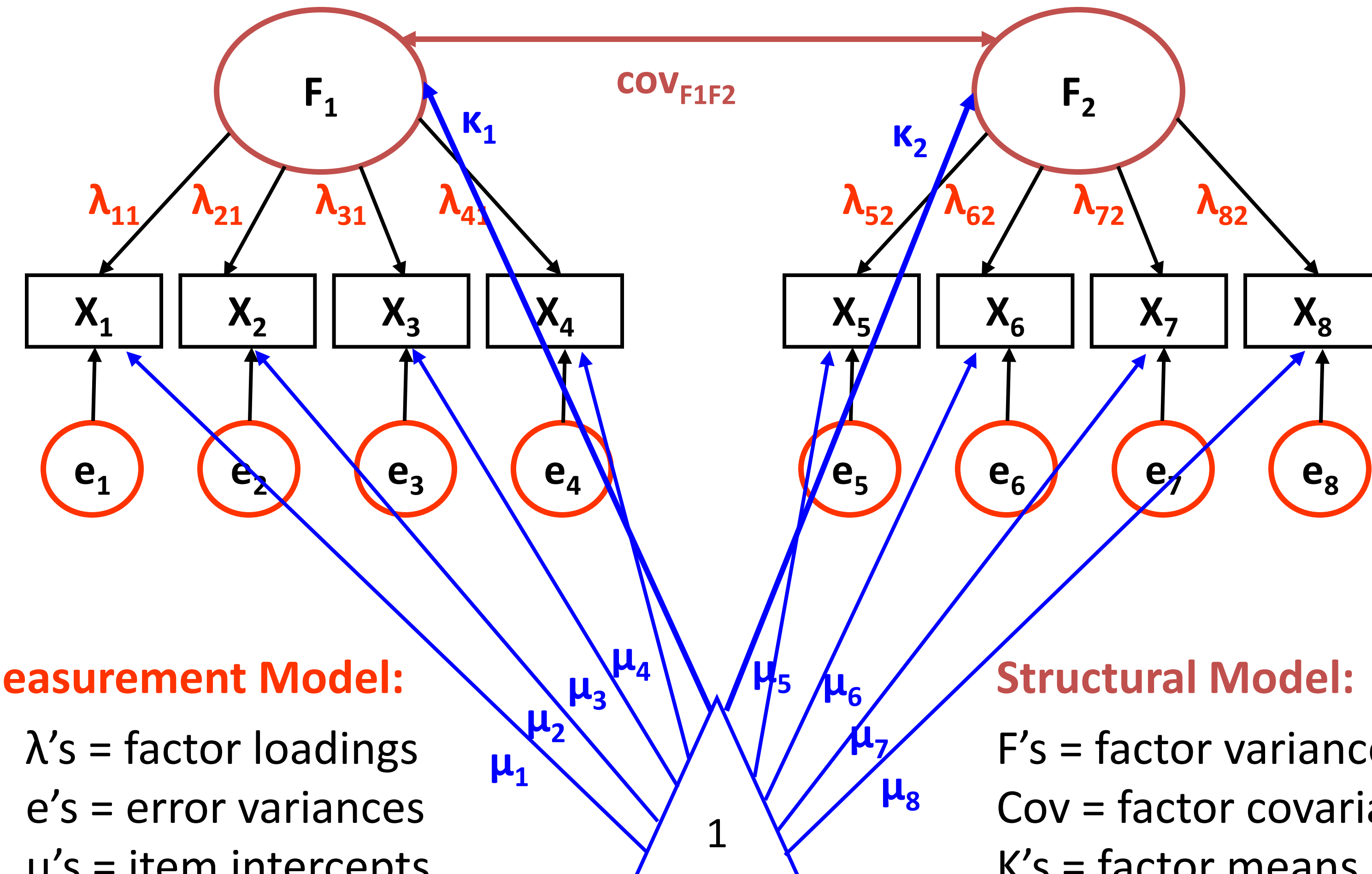
- How we judge model fit…
  - **EFA: Eye-balls and Opinion**
    - #Factors? Scree plots, interpretability…
    - Which rotation? Whichever makes most sense…
    - Which indicators load? Cut-off of .3-.4ish
  - **CFA: Inferential tests via of Maximum Likelihood**
    - Global model fit test
    - Significance of item loadings
    - Significance of error variances (and covariances)
    - Ability to test appropriateness of model constraints or model additions via tests for change in model fit

EFA vs. CFA, continued

- What we do with the latent factors…
  - **EFA: Don't compute factor scores…**
    - Factor indeterminacy issues
    - Inconsistency in how factor models are applied to data
      - Factor model based on common variance only
      - Summing items? That's using total variance (component)
  - **CFA: Let them be part of the model**
    - Don't need factor scores, but they are less indeterminate in CFA than in EFA (although still assumed perfect then)
    - Better: Test relations with latent factors directly through SEM
      - Factors can be predictors (exogenous) or outcomes (endogenous) or both at once as needed
      - Relationships will be disattenuated for measurement error

CFA Model WITH Factor Means and Item Intercepts

(But some of these values will have to be restricted for the model to be identified)



Measurement Model:

λ's = factor loadings

e's = error variances

μ's = item intercepts

Structural Model:

F's = factor variances

Cov = factor covariances

K's = factor means

2 Types of CFA Solutions

- CFA output comes in unstandardized and standardized versions:

- Unstandardized → predicts scale-sensitive original item response:
  - $\mathbf{X_{is} = \mu_i + \lambda_i F_s + e_{is}}$
  - *Useful when comparing solutions across groups or time*
  - Note the solution asymmetry: item parameters μ_i and λ_i will be given in the item metric, but $\mathbf{e_{is}}$ will be given as the error variance across persons for that item
  - $Var(X_i) = [\boldsymbol{\lambda_i}^2 * Var(F)] + Var(e_i)$

- Standardized → solution transformed to $Var(Y_i)$=1, $Var(F)$=1:
  - *Useful when comparing items within a solution (on same scale then)*
  - Standardized intercept = $\boldsymbol{\mu_i}$ **/ SD(Y)** → not typically reported
  - Standardized factor loading = **[$\boldsymbol{\lambda_i}$ * SD(F)] / SD(Y) = item correlation with factor**
  - Standardized error variance = 1 – standardized λ_i^2 = "variance due to *not* factor"
  - R^2 for item = standardized λ_i^2 = "variance due to the factor"

CFA Model Equations with Item Intercepts

- Measurement model per item (numbered) for subject *s*:
  - $X_{1s} = \mu_1 + \lambda_{11}F_{1s} + 0F_{2s} + e_{1s}$
  - $X_{2s} = \mu_2 + \lambda_{21}F_{1s} + 0F_{2s} + e_{2s}$
  - $X_{3s} = \mu_3 + \lambda_{31}F_{1s} + 0F_{2s} + e_{3s}$
  - $X_{4s} = \mu_4 + \lambda_{41}F_{1s} + 0F_{2s} + e_{4s}$
  - $X_{5s} = \mu_5 + 0F_{1s} + \lambda_{52}F_{2s} + e_{5s}$
  - $X_{6s} = \mu_6 + 0F_{1s} + \lambda_{62}F_{2s} + e_{6s}$
  - $X_{7s} = \mu_7 + 0F_{1s} + \lambda_{72}F_{2s} + e_{7s}$
  - $X_{8s} = \mu_8 + 0F_{1s} + \lambda_{82}F_{2s} + e_{8s}$

The equation predicting each item resembles a linear regression model:

$Y_{is} = B_{0i} + B_{1i}X_{1s} + B_{2i}X_{2s} + e_{is}$

You decide how many factors and whether each item loads (loading then estimated) or not.

Unstandardized loadings (λ) are the slopes of regressing the response (Y) on the factor (X).

Standardized loadings are the slopes in a correlation metric (and Std Loading^2 = reliability).

Intercepts (μ) are expected value of Y (item) when all factors (X's) are 0 (no misfit).

Expressing the CFA Model in Matrices: Factor Loadings

- If we put our loadings into a matrix $\boldsymbol{\Lambda}$ (size *p* items by *m* factors)

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{21} & 0 \\ \lambda_{31} & 0 \\ \lambda_{41} & 0 \\ 0 & \lambda_{52} \\ 0 & \lambda_{62} \\ 0 & \lambda_{72} \\ 0 & \lambda_{82} \end{bmatrix}$$

Expressing the CFA Model in Matrices: Unique Variances

- If we put our unique variances into a matrix $\mathbf{\Psi}$ (size *p* items by *p* items)

$$\mathbf{\Psi} = \begin{bmatrix} \psi_1^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \psi_2^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \psi_3^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \psi_4^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \psi_5^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \psi_6^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \psi_7^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \psi_8^2 \end{bmatrix}$$

Expressing the CFA Model in Matrices: Factor Covariances

- If we put our factor covariances into a matrix $\mathbf{\Phi}$ (size *m* factors by *m* factors):

$$\mathbf{\Phi} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{bmatrix}$$

The Result

- The CFA model then predicts the observed covariance matrix of the items by:

$$\mathbf{\Sigma} = \mathbf{\Lambda \Phi \Lambda}' + \mathbf{\Psi}$$

$$\mathbf{\Sigma} = \begin{bmatrix} \lambda_{11}^2 + \psi_1 & \lambda_{11}\lambda_{21} & \lambda_{11}\lambda_{31} & \lambda_{11}\phi_{12}\lambda_{42} & \lambda_{11}\phi_{12}\lambda_{52} & \lambda_{11}\phi_{12}\lambda_{62} \\ \lambda_{11}\lambda_{21} & \lambda_{21}^2 + \psi_2 & \lambda_{21}\lambda_{31} & \lambda_{21}\phi_{12}\lambda_{42} & \lambda_{21}\phi_{12}\lambda_{52} & \lambda_{21}\phi_{12}\lambda_{62} \\ \lambda_{11}\lambda_{31} & \lambda_{21}\lambda_{31} & \lambda_{31}^2 + \psi_3 & \lambda_{31}\phi_{12}\lambda_{42} & \lambda_{31}\phi_{12}\lambda_{52} & \lambda_{31}\phi_{12}\lambda_{62} \\ \lambda_{11}\phi_{12}\lambda_{42} & \lambda_{21}\phi_{12}\lambda_{42} & \lambda_{31}\phi_{12}\lambda_{42} & \lambda_{42}^2 + \psi_4 & \lambda_{42}\lambda_{52} & \lambda_{42}\lambda_{62} \\ \lambda_{11}\phi_{12}\lambda_{52} & \lambda_{21}\phi_{12}\lambda_{52} & \lambda_{31}\phi_{12}\lambda_{52} & \lambda_{42}\lambda_{52} & \lambda_{52}^2 + \psi_5 & \lambda_{52}\lambda_{62} \\ \lambda_{11}\phi_{12}\lambda_{62} & \lambda_{21}\phi_{12}\lambda_{62} & \lambda_{31}\phi_{12}\lambda_{62} & \lambda_{42}\lambda_{62} & \lambda_{52}\lambda_{62} & \lambda_{62}^2 + \psi_6 \end{bmatrix}$$

CFA Model Predictions

F_1 BY X_1-X_4, F_2 BY X_5-X_8

Two items from same factor (room for misfit):

- Unstandardized solution: Covariance$_{x1,x4}$ = λ_{11}*Var(F_1)*λ_{41}
- Standardized solution: Correlation$_{x1,x4}$ = λ_{11}*(1)*λ_{41} → std loadings
- ONLY reason for cor$_{x1,x4}$ is common factor (local independence, LI)

Two items from different factors (room for misfit):

- Unstandardized solution: Covariance$_{x1,x8}$ = λ_{11}*cov$_{F1,F2}$*λ_{82}
- Standardized solution: Correlation$_{x1,x8}$ = λ_{11}*cor$_{F1,F2}$*λ_{82} → std loadings
- ONLY reason for cor$_{x1,x8}$ is correlation between factors (again, LI)

Variances are additive (and will be reproduced correctly):

- **Var(X_1) = (λ_{11}^2)*Var(F_1) + Var(e_1) → note imbalance of λ^2 and e**

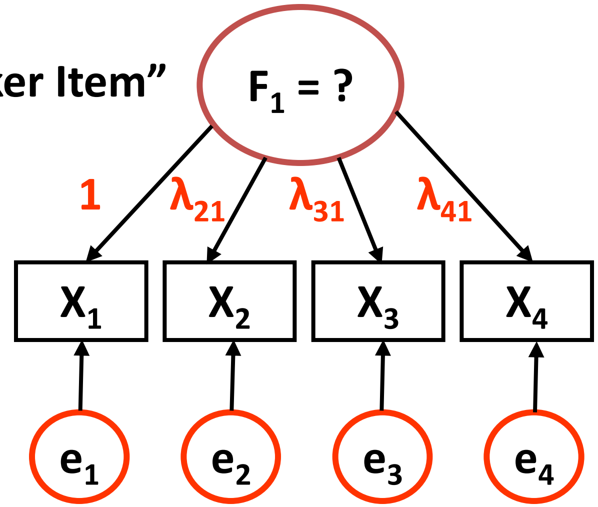
Assumptions of CFA Models

- **Dimensionality** is assumed known (from number of latent traits)
  - Local Independence → e's are independent after controlling for factor(s)
- **Linear model** → a one-unit change in latent trait/factor F has same increase in expected item response (Y) at all points of factor (X)
  - Won't work well for binary/ordinal data… thus, we need IRT
  - Often of questionable utility for Likert scale data (normality?)
- Goal is to **predict covariance** between items → basis of model fit
  - Variances will always be perfectly reproduced; covariances will not be
- CFA models are **usually presented without μ_i** (the item intercept)
  - μ_i doesn't really matter in CFA because it doesn't contribute to the covariance, but we will keep it for continuity with IRT
  - Item intercepts are also important when dealing with factor mean diffs

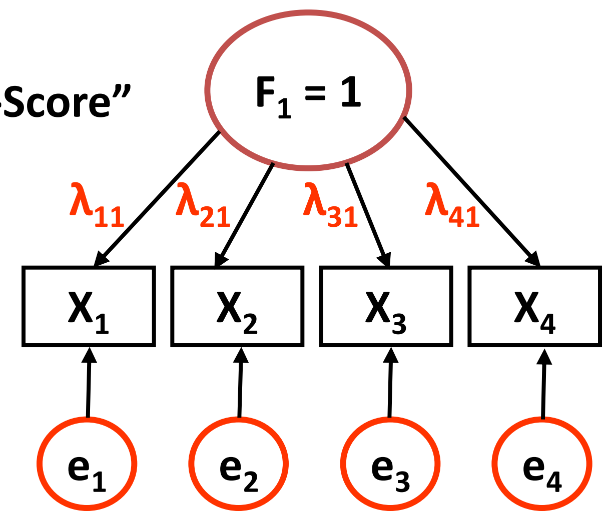
CFA Model Identification:

Create a Scale for the Latent Variable

"Marker Item"



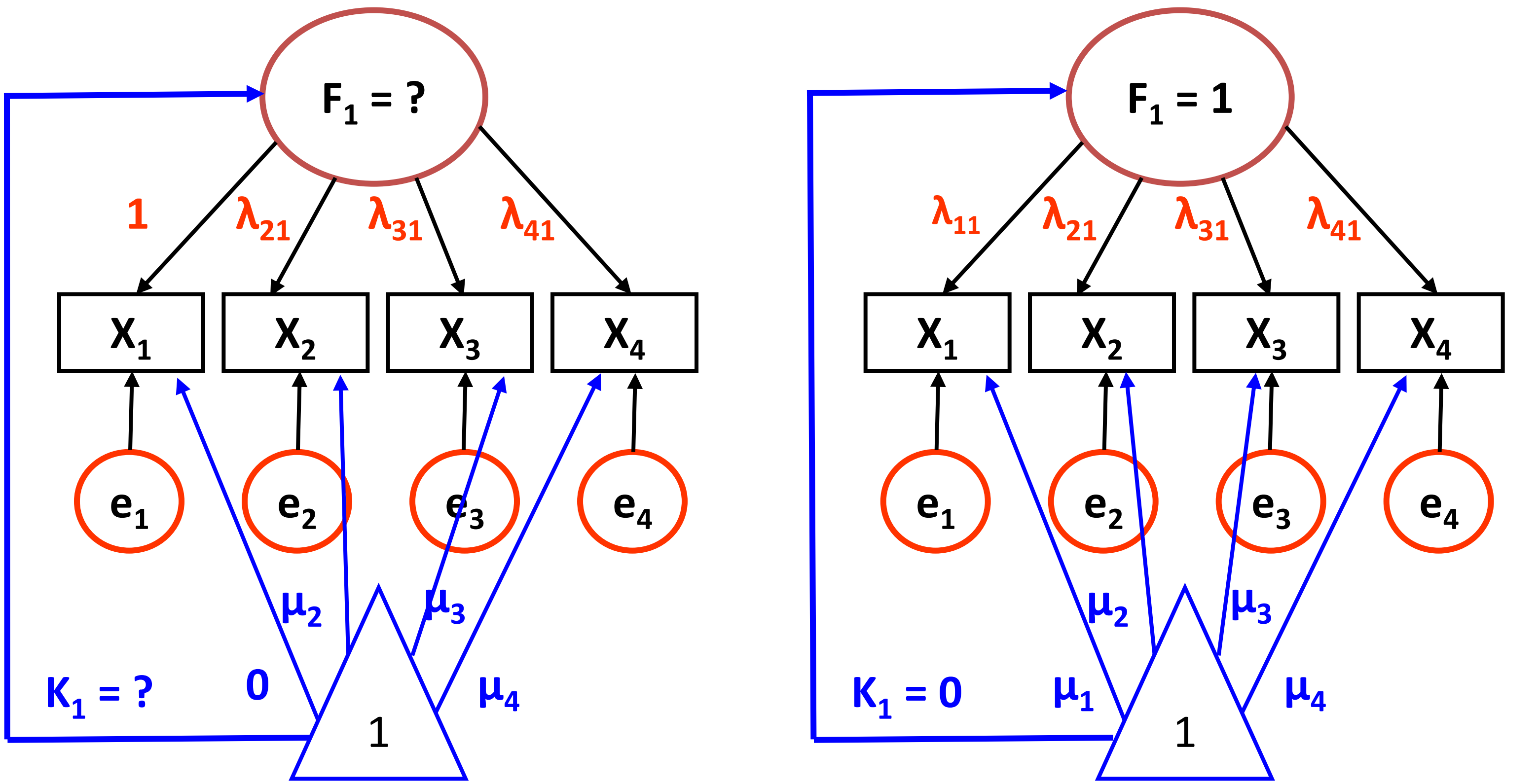
"Z-Score"



- The factor doesn't exist, so **it needs a scale** (a mean and variance):
- Two **equivalent** options to do so

- Create a scale for the VARIANCE:
  - **1) Scale using a marker item**
    - Fix one loading to 1; factor is scaled as reliable part of that marker item
    - Loading = .9, variance =16?
      $Var(F_1) = (.9^2)*16 = 12.96$
  - **2) Fix factor variance to 1**
    - Factor is interpreted as z-score
    - Can't be used in other models with higher-order factors

CFA Model Identification:
Two Options for Scaling the Factor Mean



"Marker Item" → Fix 1 item intercept to 0; estimate factor mean

Item intercept is expected outcome when factor = 0 (when item = 0)

"Z-Score" → Fix factor mean to 0, estimate all item intercepts

Item intercept is expected outcome when factor = 0 (when item = mean)

CFA Model Identification:
Two Options for Scaling the Factor

- <u>Summary: 2 options for giving the factor a scale:</u>
  - **Marker item**: Borrow a scale from one of the items
    - Fix that item's factor loading to 1 and its intercept to 0
    - Factor variance is interpreted using the "reliable" part of that item
  - **Z-score**: Put factor on scale of mean=0 and variance=1
    - Then all item factor loadings and all item intercepts are estimated
    - Can't be used in higher-order factor models
- Most common approach is a hybrid:
    - Fix factor mean to 0, estimate all item intercepts → "z-score"
    - Estimate factor variance, fix first item factor loading to 1 → "marker"
- In reality, all methods of scaling the factor will fit equivalently well, so long as the marker item loads at all

Factor Model Identification

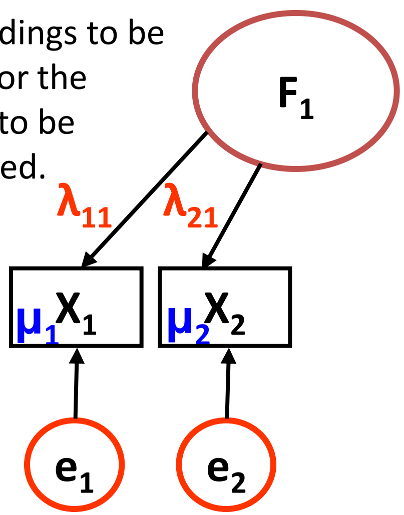
- Goal: Reproduce observed covariance matrix among items with as few estimated parameters as possible
  - Maximum likelihood usually used to estimate model parameters
    - **Measurement Model**: Factor loadings, item intercepts, error variances
    - **Structural Model**: Factor variances and covariances, factor means
  - Global model fit is evaluated as difference between model-predicted matrix and observed matrix (but only the covariances really contribute)

- How many possible parameters can you estimate (total DF)?
  - **Total DF depends on # ITEMS** → p (NOT on # people)
  - Total number of 'unique elements' in covariance matrix
    - Unique elements = each variance, each covariance, each mean
    - Total unique elements = $(p(p+1) / 2) + p$ → if 4 items, then $((4*5)/2) + 4 = 14$
- Model degrees of freedom (df)
  - Model df = # possible parameters − # estimated parameters

Under-Identified Factor: 2 Items

- Model is under-identified when there are more unknowns then pieces of information with which to estimate them
  - Cannot be solved because there are an infinite number of different parameter estimates that would result in perfect fit
  - Example: Solve x + y = 7 ??

You'd have to set the loadings to be equal for the model to be identified.



Total possible df = unique elements = **5**

0 factor variances		**1 factor variance**
0 factor means		**1 factor mean**
2 loadings	**OR**	**1 item loading**
2 item intercepts		**1 item intercept**
2 error variances		**2 error variances**

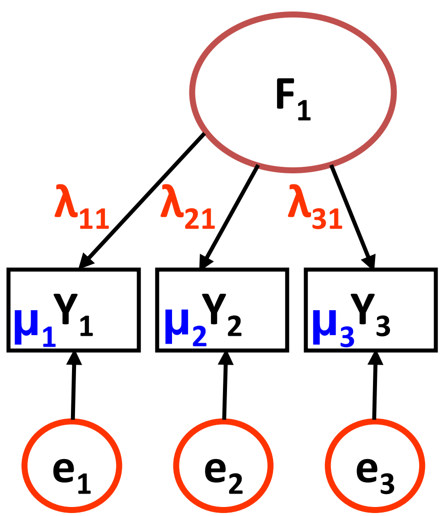
df = 5 – 6 = -1

If $r_{y1,y2} = .64$, then:

$\lambda_{11} = .800, \lambda_{21} = .800$??
$\lambda_{11} = .900, \lambda_{21} = .711$??
$\lambda_{11} = .750, \lambda_{21} = .853$??

Just-Identified Factor: 3 Items

- Model is just-identified when there are as many unknowns as pieces of information with which to estimate them
  - Parameter estimates have a unique solution that will perfectly reproduce the observed matrix
  - Example: Solve x + y = 7, 3x – y = 1



Total possible df = unique elements = **9**

0 factor variances
0 factor means
3 loadings
3 item intercepts
3 error variances

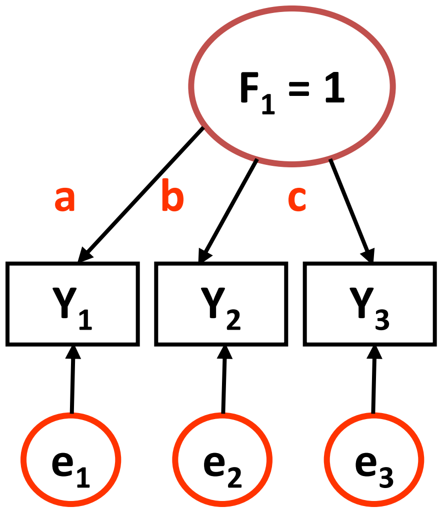
OR

1 factor variance
1 factor mean
2 item loadings
2 item intercepts
3 error variances

df = 9 – 9 = 0

Not really a model – more like a description

Solving a Just-Identified Model



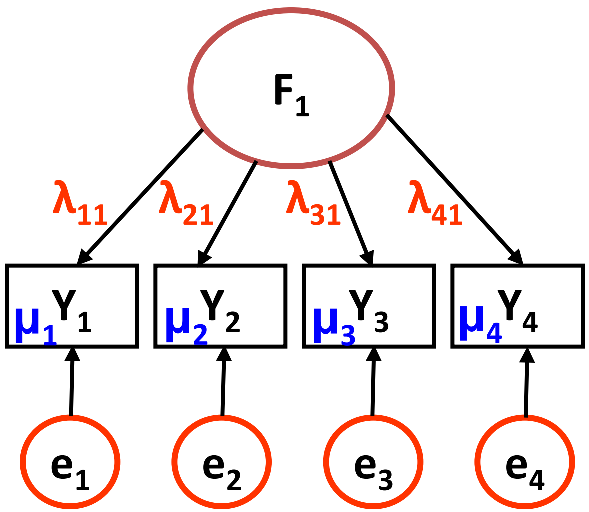
	Y_1	Y_2	Y_3
Y_1	1.00		
Y_2	.595	1.00	
Y_3	.448	.544	1.00

- Step 1: $ab = .595$
  $ac = .448$
  $bc = .544$
- Step 2: $b = .595/a$
  $c = .488/a$
  $(.595/a)(.448/a) = .544$
- Step 3: $.26656/a^2 = .544$
  $a = .70$
- Step 4: $.70b = .595 \quad b = .85$
  $.70c = .448 \quad c = .64$
- Step 5: $Var(e_1) = 1 - a^2 = .51$

Over-Identified Factor: 4+ Items

- Model is <u>over-identified</u> when there are fewer unknowns than pieces of information with which to estimate them
  - Parameter estimates have a unique solution that will NOT perfectly reproduce the observed matrix
  - **NOW we can test model fit**

Total possible df = unique elements = **14**



0 factor variances		**1 factor variance**
0 factor means		**1 factor mean**
4 loadings	**OR**	**3 item loadings**
4 item intercepts		**3 item intercepts**
4 error variances		**4 error variances**

df = 14 – 12 = 2

Did we do a 'good enough' job reproducing the matrix with 2 fewer parameters than was possible to use?

Indices of Global Model Fit

- Primary: obtained model $\boldsymbol{\chi^2 = F_{ML}(N\text{-}1)}$
  - X^2 is evaluated based on model df (# parameters left over)
  - Tests null hypothesis that $\Sigma = S$ (that model is perfect), so significance is undesirable (smaller χ^2, bigger p-value is better)
  - Just using χ^2 is insufficient, however:
    - Distribution doesn't behave like a true χ^2 if sample sizes are small or if items are non-normally distributed
    - Obtained χ^2 depends largely on sample size
    - Is unreasonable null hypothesis (perfect fit??)
- Because of these issues, alternative measures of fit are usually used in conjunction with the χ^2 test of model fit
  - Absolute Fit Indices (besides χ^2)
  - Parsimony-Corrected; Comparative (Incremental) Fit Indices

Indices of Global Model Fit

- Absolute Fit: $\boldsymbol{\chi^2}$
  - Don't use 'ratio rules' like $\chi^2/df > 2$ or $\chi^2/df > 3$

- Absolute Fit: **SRMR**
  - **Standardized Root Mean Square Residual**
  - Get difference of Σ and S → residual matrix
  - Sum the squared residuals in matrix, divide by number of residuals summed
  - Ranges from 0 to 1: smaller is better
  - ".08 or less" → good fit

- See also: **RMR** (**Root Mean Square Residual**)

Indices of Global Model Fit

Parsimony-Corrected: **RMSEA**

- **Root Mean Square Error of Approximation**
- Relies on a non-centrality parameter (NCP)
  - Indexes how far off your model is → χ^2 distribution shoved over
  - NCP → d = (χ^2 – df) / (N-1)   Then, RMSEA = SQRT(d/df)
  - RMSEA ranges from 0 to 1; smaller is better
  - < .05 or .06 = "good", .05 to .08 = "acceptable", .08 to .10 = "mediocre", and >.10 = "unacceptable"
  - In addition to point estimate, get 90% confidence interval
  - RMSEA penalizes for model complexity – it's discrepancy in fit per df left in model (but not sensitive to N, although CI can be)
  - Test of "close fit": null hypothesis that RMSEA ≤ .05

Indices of Global Model Fit

Comparative (Incremental) Fit Indices

- Fit evaluated relative to a 'null' model (of 0 covariances)
- Relative to that, your model should be great!
- **CFI: Comparative Fit Index**
  - Also based on idea of NCP ($\chi^2 - df$)
  - $CFI = 1 - \frac{\max[(\chi^2_T - df_T), 0]}{\max[(\chi^2_T - df_T), (\chi^2_N - df_N), 0]}$

    T = target model
    N = null model
  - From 0 to 1: bigger is better, > .90 = "acceptable", > .95 = "good"
- **TLI: Tucker-Lewis Index** (= **Non-Normed Fit Index**)
  - $TLI = \frac{(\chi^2_N/df_N) - (\chi^2_T/df_T)}{(\chi^2_N/df_N) - 1}$
  - From <0 to >1, bigger is better, >.95 = "good"

CFA THROUGH AN EXAMPLE

Software for CFA and SEM

- SAS has the CALIS procedure that will estimate the covariance-portion of a CFA/SEM model
  - Is somewhat dated now

- Instead, I recommend the use of Mplus for CFA and SEM
  - Has many options – fairly easy to use
  - Used in our examples

Our Data: Teacher Ratings

- To demonstrate CFA, we will return to the teacher ratings data we used last week
  - First question: does a one-factor model fit the data?

-----Alphabetic List of Variables and Attributes-----

Variable	Type	Len	Pos	Label
ITEM13	Num	8	0	INSTRUC WELL PREPARED
ITEM14	Num	8	8	INSTRUC SCHOLARLY GRASP
ITEM15	Num	8	16	INSTRUCTOR CONFIDENCE
ITEM16	Num	8	24	INSTRUCTOR FOCUS LECTURES
ITEM17	Num	8	32	INSTRUCTOR USES CLEAR RELEVANT EXAMPLES
ITEM18	Num	8	40	INSTRUCTOR SENSITIVE TO STUDENTS
ITEM19	Num	8	48	INSTRUCTOR ALLOWS ME TO ASK QUESTIONS
ITEM20	Num	8	56	INSTRUCTOR IS ACCESSIBLE TO STUDENTS OUTSIDE CLASS
ITEM21	Num	8	64	INSTRUCTOR AWARE OF STUDENTS UNDERSTANDING
ITEM22	Num	8	72	I AM SATISFIED WITH STUDENT PERFORMANCE EVALUATION
ITEM23	Num	8	80	COMPARED TO OTHER INSTRUCTORS, THIS INSTRUCTOR IS
ITEM24	Num	8	88	COMPARED TO OTHER COURSES THIS COURSE WAS

One Factor Results

Mplus - [teacher1.inp]

File Edit View Mplus Graph Window Hel

```
TITLE:
teachers
DATA:
    FILE = teachers.csv;
VARIABLE:
    NAMES = ITEM13-ITEM24;
    MISSING = ALL(99);


MODEL:
    F1 by ITEM13* ITEM14-ITEM24;

    F1@1;
OUTPUT:
    RESIDUAL MODINDICES STAND;
```

```
MODEL FIT INFORMATION

Number of Free Parameters                       36

Loglikelihood

          H0 Value                      -18337.949
          H1 Value                      -17707.861

Information Criteria

          Akaike (AIC)                   36747.899
          Bayesian (BIC)                 36937.379
          Sample-Size Adjusted BIC       36823.019
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                           1260.178
          Degrees of Freedom                    54
          P-Value                           0.0000

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                           0.125
          90 Percent C.I.                    0.119  0.131
          Probability RMSEA <= .05           0.000

CFI/TLI

          CFI                                0.867
          TLI                                0.837

Chi-Square Test of Model Fit for the Baseline Model

          Value                           9132.568
          Degrees of Freedom                    66
          P-Value                           0.0000

SRMR (Standardized Root Mean Square Residual)

          Value                              0.060
```

Results Interpretation

- Model parameters: 36
  - 12 item intercepts (means – just $\bar{X}$)
  - 12 factor loadings
  - 12 unique variances
  - 0 factor variances (factor variance set to one)
- Model fit:
  - RMSEA: 0.125 ("good" is < 0.05)
  - CFI: 0.867 ("good" > 0.95)
  - TLI: 0.837 ("good" > 0.95)
- Conclusion:
  - One factor model does not fit data very well

Two Factor Model

- What happens when we use some of the information from last week and build a two-factor model?
  - Competency factor: (13, 14, 15, 16, 17)
  - Friendliness factor: (18, 19, 20, 21)
- But what about items 22, 23, and 24?

-----Alphabetic List of Variables and Attributes-----

Variable	Type	Len	Pos	Label
ITEM13	Num	8	0	INSTRUC WELL PREPARED
ITEM14	Num	8	8	INSTRUC SCHOLARLY GRASP
ITEM15	Num	8	16	INSTRUCTOR CONFIDENCE
ITEM16	Num	8	24	INSTRUCTOR FOCUS LECTURES
ITEM17	Num	8	32	INSTRUCTOR USES CLEAR RELEVANT EXAMPLES
ITEM18	Num	8	40	INSTRUCTOR SENSITIVE TO STUDENTS
ITEM19	Num	8	48	INSTRUCTOR ALLOWS ME TO ASK QUESTIONS
ITEM20	Num	8	56	INSTRUCTOR IS ACCESSIBLE TO STUDENTS OUTSIDE CLASS
ITEM21	Num	8	64	INSTRUCTOR AWARE OF STUDENTS UNDERSTANDING
ITEM22	Num	8	72	I AM SATISFIED WITH STUDENT PERFORMANCE EVALUATION
ITEM23	Num	8	80	COMPARED TO OTHER INSTRUCTORS, THIS INSTRUCTOR IS
ITEM24	Num	8	88	COMPARED TO OTHER COURSES THIS COURSE WAS

Mixing Known Factors with Unknown Items

- Because we more-or-less know how 9 of our items work, we can be less specific about the rest of our items
  - Allow them to load onto both factors
  - See if any loadings are significantly different from zero
- There are other methods we could use to see how these items functioned
  - LaGrange multipliers (modification indices)

Two Factor Model #1

teacher2a.inp

```
TITLE:
teachers
DATA:
    FILE = teachers.csv;
VARIABLE:
    NAMES = ITEM13-ITEM24;
    MISSING = ALL(99);

MODEL:
    F1 by ITEM13* ITEM14-ITEM17 ITEM22-ITEM24;
    F2 by ITEM18* ITEM19-ITEM21 ITEM22-ITEM24;
    F1 with F2;
    F1@1; F2@1;
OUTPUT:
    RESIDUAL MODINDICES STAND;
```

```
MODEL FIT INFORMATION

Number of Free Parameters                        40

Loglikelihood

          H0 Value                       -18039.160
          H1 Value                       -17707.861

Information Criteria

          Akaike (AIC)                    36158.321
          Bayesian (BIC)                  36368.854
          Sample-Size Adjusted BIC        36241.788
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                             662.600
          Degrees of Freedom                     50
          P-Value                            0.0000

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                            0.093
          90 Percent C.I.                     0.086  0.099
          Probability RMSEA <= .05            0.000

CFI/TLI

          CFI                                 0.932
          TLI                                 0.911

Chi-Square Test of Model Fit for the Baseline Model

          Value                            9132.568
          Degrees of Freedom                     66
          P-Value                            0.0000

SRMR (Standardized Root Mean Square Residual)

          Value                               0.042
```

Factor Loadings:

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
F1 BY				
ITEM13	0.574	0.017	33.658	0.000
ITEM14	0.552	0.016	33.607	0.000
ITEM15	0.584	0.017	33.951	0.000
ITEM16	0.569	0.021	27.734	0.000
ITEM17	0.703	0.021	33.687	0.000
ITEM22	-0.065	0.051	-1.280	0.201
ITEM23	0.433	0.037	11.823	0.000
ITEM24	0.310	0.041	7.553	0.000
F2 BY				
ITEM18	0.820	0.024	33.681	0.000
ITEM19	0.672	0.024	27.742	0.000
ITEM20	0.552	0.024	23.423	0.000
ITEM21	0.735	0.023	31.306	0.000
ITEM22	0.814	0.052	15.609	0.000
ITEM23	0.394	0.037	10.745	0.000
ITEM24	0.373	0.042	8.996	0.000
F1 WITH				
F2	0.760	0.017	43.923	0.000

- Item 22:
  - No loading onto competency factor
  - Loading onto friendliness factor
- Items 23 & 24:
  - Loadings about the same magnitude on both factors
  - Inconclusive results
    - Perhaps we should omit the items?

Results Interpretation

- Model parameters: 40
  - 12 item intercepts (means – just $\bar{X}$)
  - 15 factor loadings
  - 12 unique variances
  - 1 factor covariance
  - 0 factor variances (factor variance set to one)
- Model fit:
  - RMSEA: 0.093 ("good" is < 0.05)
  - CFI: 0.932 ("good" > 0.95)
  - TLI: 0.911 ("good" > 0.95)
- Conclusion:
  - Two factor model does not fit data very well

Perhaps Another Factor?

- Items 23 and 24 seem to have another thing in common: the wording of their questions is very similar
  - Perhaps this indicates another factor

-----Alphabetic List of Variables and Attributes-----

Variable	Type	Len	Pos	Label
ITEM13	Num	8	0	INSTRUC WELL PREPARED
ITEM14	Num	8	8	INSTRUC SCHOLARLY GRASP
ITEM15	Num	8	16	INSTRUCTOR CONFIDENCE
ITEM16	Num	8	24	INSTRUCTOR FOCUS LECTURES
ITEM17	Num	8	32	INSTRUCTOR USES CLEAR RELEVANT EXAMPLES
ITEM18	Num	8	40	INSTRUCTOR SENSITIVE TO STUDENTS
ITEM19	Num	8	48	INSTRUCTOR ALLOWS ME TO ASK QUESTIONS
ITEM20	Num	8	56	INSTRUCTOR IS ACCESSIBLE TO STUDENTS OUTSIDE CLASS
ITEM21	Num	8	64	INSTRUCTOR AWARE OF STUDENTS UNDERSTANDING
ITEM22	Num	8	72	I AM SATISFIED WITH STUDENT PERFORMANCE EVALUATION
ITEM23	Num	8	80	COMPARED TO OTHER INSTRUCTORS, THIS INSTRUCTOR IS
ITEM24	Num	8	88	COMPARED TO OTHER COURSES THIS COURSE WAS

Three Factor Model

```
TITLE:
teachers
DATA:
    FILE = teachers.csv;
VARIABLE:
    NAMES = ITEM13-ITEM24;
    MISSING = ALL(99);


MODEL:
    F1 by ITEM13* ITEM14-ITEM16;
    F2 by ITEM18* ITEM17 ITEM19-ITEM21 ITEM22;
    F3 by ITEM23* ITEM24;
    F1 with F2; F1 with F3; F2 with F3;
    F1@1; F2@1; F3@1;
OUTPUT:
    RESIDUAL MODINDICES STAND;
```

```
MODEL FIT INFORMATION

Number of Free Parameters                        39

Loglikelihood

          H0 Value                       -17975.183
          H1 Value                       -17707.861

Information Criteria

          Akaike (AIC)                    36028.367
          Bayesian (BIC)                  36233.637
          Sample-Size Adjusted BIC        36109.747
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                             534.645
          Degrees of Freedom                     51
          P-Value                            0.0000

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                            0.082
          90 Percent C.I.                     0.075  0.088
          Probability RMSEA <= .05            0.000

CFI/TLI

          CFI                                 0.947
          TLI                                 0.931

Chi-Square Test of Model Fit for the Baseline Model

          Value                            9132.568
          Degrees of Freedom                     66
          P-Value                            0.0000

SRMR (Standardized Root Mean Square Residual)

          Value                               0.044
```

Results Interpretation

- Model parameters: 39
  - 12 item intercepts (means – just $\bar{X}$)
  - 12 factor loadings
  - 12 unique variances
  - 3 factor covariances
  - 0 factor variances (factor variance set to one)
- Model fit:
  - RMSEA: 0.082 ("good" is < 0.05) – borderline
  - CFI: 0.947 ("good" > 0.95) – acceptable
  - TLI: 0.931 ("good" > 0.95) – acceptable
- Conclusion:
  - Three factor model fits data adequately

Three Factor Model Estimates

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
F1 BY				
ITEM13	0.588	0.017	34.547	0.000
ITEM14	0.572	0.016	35.144	0.000
ITEM15	0.593	0.017	34.350	0.000
ITEM16	0.559	0.021	26.819	0.000
F2 BY				
ITEM18	0.803	0.024	33.205	0.000
ITEM17	0.685	0.021	32.263	0.000
ITEM19	0.651	0.024	26.969	0.000
ITEM20	0.544	0.023	23.196	0.000
ITEM21	0.745	0.023	32.272	0.000
ITEM22	0.738	0.028	26.506	0.000
F3 BY				
ITEM23	0.890	0.021	41.523	0.000
ITEM24	0.715	0.022	32.600	0.000
F1 WITH				
F2	0.774	0.016	48.402	0.000
F3	0.776	0.016	49.570	0.000
F2 WITH				
F3	0.822	0.014	56.998	0.000
Variances				
F1	1.000	0.000	999.000	999.000
F2	1.000	0.000	999.000	999.000
F3	1.000	0.000	999.000	999.000
Residual Variances				
ITEM13	0.197	0.010	20.156	0.000
ITEM14	0.176	0.009	19.913	0.000
ITEM15	0.207	0.010	20.459	0.000
ITEM16	0.393	0.017	23.733	0.000
ITEM17	0.338	0.016	21.482	0.000
ITEM18	0.423	0.020	21.038	0.000
ITEM19	0.515	0.022	23.464	0.000
ITEM20	0.540	0.022	24.582	0.000
ITEM21	0.415	0.019	22.259	0.000
ITEM22	0.717	0.030	24.046	0.000
ITEM23	0.141	0.017	8.409	0.000
ITEM24	0.355	0.017	21.098	0.000

Model Predicted Covariance Matrix

- To show how CFA works…We can confirm in IML the model predicted covariances

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}$$

predcov 12 rows 12 cols (numeric)

	COL1	COL2	COL3	COL4	COL5	COL6
ROW1	0.542744	0.336336	0.348684	0.328692	0.3654549	0.3117517
ROW2	0.336336	0.503184	0.339196	0.319748	0.3555106	0.3032687
ROW3	0.348684	0.339196	0.558649	0.331487	0.3685625	0.3144027
ROW4	0.328692	0.319748	0.331487	0.705481	0.3474308	0.2963762
ROW5	0.3654549	0.3555106	0.3685625	0.3474308	0.982809	0.550055
ROW6	0.3117517	0.3032687	0.3144027	0.2963762	0.550055	0.892225
ROW7	0.2962779	0.2882159	0.2987973	0.2816656	0.522753	0.445935
ROW8	0.2475809	0.240844	0.2496862	0.2353703	0.436832	0.37264
ROW9	0.3390584	0.3298324	0.3419416	0.3223362	0.598235	0.510325
ROW10	0.3358727	0.3267333	0.3387287	0.3193075	0.592614	0.50553
ROW11	0.4060963	0.3950461	0.4095495	0.3860678	0.5874587	0.5011323

The SAS System

	COL1	COL2	COL3	COL4	COL5	COL6
ROW12	0.3262459	0.3173685	0.3290201	0.3101556	0.4719472	0.4025951

	COL7	COL8	COL9	COL10	COL11	COL12
ROW1	0.2962779	0.2475809	0.3390584	0.3358727	0.4060963	0.3262459
ROW2	0.2882159	0.240844	0.3298324	0.3267333	0.3950461	0.3173685
ROW3	0.2987973	0.2496862	0.3419416	0.3387287	0.4095495	0.3290201
ROW4	0.2816656	0.2353703	0.3223362	0.3193075	0.3860678	0.3101556
ROW5	0.522753	0.436832	0.598235	0.592614	0.5874587	0.4719472
ROW6	0.445935	0.37264	0.510325	0.50553	0.5011323	0.4025951
ROW7	0.938801	0.354144	0.484995	0.480438	0.4762586	0.3826122
ROW8	0.354144	0.835936	0.40528	0.401472	0.3979795	0.3197251
ROW9	0.484995	0.40528	0.970025	0.54981	0.5450271	0.4378589
ROW10	0.480438	0.401472	0.54981	1.261644	0.539906	0.4337447
ROW11	0.4762586	0.3979795	0.5450271	0.539906	0.9331	0.63635
ROW12	0.3826122	0.3197251	0.4378589	0.4337447	0.63635	0.866225

Model Estimated Covariances/Correlations/Residual Correlations

	ITEM13	ITEM14	ITEM15	ITEM16	ITEM17
ITEM13	0.543				
ITEM14	0.336	0.502			
ITEM15	0.349	0.339	0.559		
ITEM16	0.329	0.319	0.332	0.705	
ITEM17	0.312	0.303	0.315	0.297	0.807
ITEM18	0.366	0.356	0.369	0.348	0.551
ITEM19	0.296	0.288	0.299	0.282	0.446
ITEM20	0.248	0.241	0.250	0.235	0.373
ITEM21	0.340	0.330	0.342	0.323	0.511
ITEM22	0.336	0.327	0.339	0.319	0.505
ITEM23	0.407	0.395	0.410	0.386	0.502
ITEM24	0.327	0.317	0.329	0.310	0.403

Model Estimated Covariances/Correlations/Residual Correlations

	ITEM18	ITEM19	ITEM20	ITEM21	ITEM22
ITEM18	1.068				
ITEM19	0.523	0.939			
ITEM20	0.437	0.354	0.836		
ITEM21	0.599	0.485	0.405	0.971	
ITEM22	0.593	0.480	0.401	0.550	1.261
ITEM23	0.588	0.476	0.398	0.545	0.540
ITEM24	0.472	0.382	0.320	0.438	0.434

Model Estimated Covariances/Correlations/Residual Correlations

	ITEM23	ITEM24
ITEM23	0.934	
ITEM24	0.636	0.866

COMPARING CFA AND EFA

Comparing CFA and EFA

- Although CFA and EFA are very similar, their results can be very different for 2 or more factors
- Recall, EFA typically assumes uncorrelated factors
- If we fix our factor correlation to zero, a CFA model becomes very similar to an EFA model
  - But…with one exception…

EFA Model Constraints

- For more than one factor, the EFA model has too many parameters to estimate
  - Uses identification constraints:

$$\mathbf{\Lambda}'\mathbf{\Psi}\mathbf{\Lambda} = \mathbf{\Delta}$$

  where $\mathbf{\Delta}$ is diagonal

- This constraint puts m*(m-1)/2 constraints on the loadings and uniquenesses
  - Multivariate constraints

Model Likelihood Function

- Under maximum likelihood estimators, both EFA and CFA use **the same likelihood function**
  - Multivariate normal
  - Mplus: full information
  - SAS: sufficient statistics (i.e., Wishart distribution for the covariance matrix)

CFA Approaches to EFA

- Therefore, we can approach an EFA model using a CFA
  - We just need to set the right number of constraints for identification
  - We set the value of factor loadings for a few items on a few of the factors
    - Typically to zero
    - Sometimes to one
  - We keep the factor covariance matrix as an identity
- Benefits:
  - Our constraints remove rotational indeterminacy of factor loadings
  - Defines factors with *potentially* less ambiguity
    - Constraints are easy to see
  - For some software (SAS), we get more model fit information

EFA with CFA Constraints

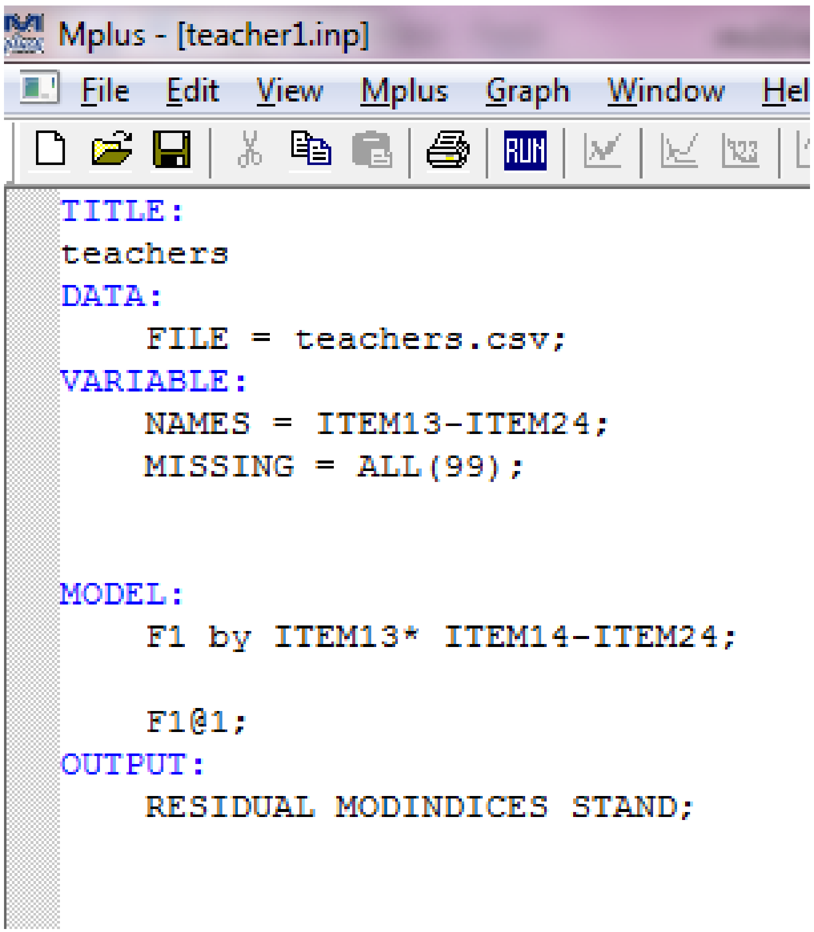
- The constraints in CFA are
  - Fixed factor loadings (set to either zero or one)
    - Use “row echelon” form :
    - One item has only one factor loading estimated
    - One item has only two factor loadings estimated
    - One item has only three factor loadings estimated
  - Fixed factor covariances
    - Set to zero

Re-Examining Our EFA of the Teacher Ratings Data – with CFA

- We will fit a series of “just-identified” CFA models and examine our results

- NOTE: the one factor model CFA model will be identical to the one factor EFA model
  - The loadings and unique variances in the EFA model are the standardized versions from the CFA model

One Factor Results




```
TITLE:
teachers
DATA:
    FILE = teachers.csv;
VARIABLE:
    NAMES = ITEM13-ITEM24;
    MISSING = ALL(99);


MODEL:
    F1 by ITEM13* ITEM14-ITEM24;

    F1@1;
OUTPUT:
    RESIDUAL MODINDICES STAND;
```

```
MODEL FIT INFORMATION

Number of Free Parameters                       36

Loglikelihood

          H0 Value                      -18337.949
          H1 Value                      -17707.861

Information Criteria

          Akaike (AIC)                   36747.899
          Bayesian (BIC)                 36937.379
          Sample-Size Adjusted BIC       36823.019
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                           1260.178
          Degrees of Freedom                    54
          P-Value                           0.0000

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                           0.125
          90 Percent C.I.                    0.119  0.131
          Probability RMSEA <= .05           0.000

CFI/TLI

          CFI                                0.867
          TLI                                0.837

Chi-Square Test of Model Fit for the Baseline Model

          Value                           9132.568
          Degrees of Freedom                    66
          P-Value                           0.0000

SRMR (Standardized Root Mean Square Residual)

          Value                              0.060
```

Two Factor Results

```
TITLE:
teachers
DATA:
    FILE = teachers.csv;
VARIABLE:
    NAMES = ITEM13-ITEM24;
    MISSING = ALL(99);


MODEL:
    F1 by ITEM13* ITEM14-ITEM24;
    F2 by ITEM14* ITEM15-ITEM24;
    F1 with F2@0;
    F1@1; F2@1;
OUTPUT:
    RESIDUAL MODINDICES STAND;
```

```
MODEL FIT INFORMATION

Number of Free Parameters                       47

Loglikelihood

          H0 Value                      -17919.928
          H1 Value                      -17707.861

Information Criteria

          Akaike (AIC)                   35933.855
          Bayesian (BIC)                 36181.232
          Sample-Size Adjusted BIC       36031.929
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                            424.134
          Degrees of Freedom                    43
          P-Value                           0.0000

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                           0.079
          90 Percent C.I.                    0.072  0.086
          Probability RMSEA <= .05           0.000

CFI/TLI

          CFI                                0.958
          TLI                                0.935

Chi-Square Test of Model Fit for the Baseline Model

          Value                           9132.568
          Degrees of Freedom                    66
          P-Value                           0.0000

SRMR (Standardized Root Mean Square Residual)

          Value                              0.025
```

Three Factor Results

```
TITLE:
teachers
DATA:
    FILE = teachers.csv;
VARIABLE:
    NAMES = ITEM13-ITEM24;
    MISSING = ALL(99);


MODEL:
    F1 by ITEM13* ITEM14-ITEM24;
    F2 by ITEM14* ITEM15-ITEM24;
    F3 by ITEM15* ITEM16-ITEM24;
    F1 with F2@0; F1 with F3@0; F2 with F3@0;
    F1@1; F2@1; F3@1;
OUTPUT:
    RESIDUAL MODINDICES STAND;
```

```
MODEL FIT INFORMATION

Number of Free Parameters                       57

Loglikelihood

          H0 Value                      -17776.793
          H1 Value                      -17707.861

Information Criteria

          Akaike (AIC)                   35667.586
          Bayesian (BIC)                 35967.596
          Sample-Size Adjusted BIC       35786.527
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                            137.865
          Degrees of Freedom                    33
          P-Value                           0.0000

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                           0.047
          90 Percent C.I.                    0.039  0.055
          Probability RMSEA <= .05           0.701

CFI/TLI

          CFI                                0.988
          TLI                                0.977

Chi-Square Test of Model Fit for the Baseline Model

          Value                           9132.568
          Degrees of Freedom                    66
          P-Value                           0.0000

SRMR (Standardized Root Mean Square Residual)

          Value                              0.014
```

Four Factor Results

```
TITLE:
teachers
DATA:
    FILE = teachers.csv;
VARIABLE:
    NAMES = ITEM13-ITEM24;
    MISSING = ALL(99);


MODEL:
    F1 by ITEM13* ITEM14-ITEM24;
    F2 by ITEM14* ITEM15-ITEM24;
    F3 by ITEM15* ITEM16-ITEM24;
    F4 by ITEM16* ITEM17-ITEM24;

    F1 with F2@0; F1 with F3@0; F1 with F4@0;
    F2 with F3@0; F2 with F4@0;
    F3 with F4@0;
    F1@1; F2@1; F3@1; F4@1;
OUTPUT:
    RESIDUAL MODINDICES STAND;
```

```
MODEL FIT INFORMATION

Number of Free Parameters                       66

Loglikelihood

          H0 Value                      -17730.604
          H1 Value                      -17707.861

Information Criteria

          Akaike (AIC)                   35593.207
          Bayesian (BIC)                 35940.587
          Sample-Size Adjusted BIC       35730.928
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                             45.486
          Degrees of Freedom                    24
          P-Value                           0.0051

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                           0.025
          90 Percent C.I.                    0.013  0.036
          Probability RMSEA <= .05           1.000

CFI/TLI

          CFI                                0.998
          TLI                                0.993

Chi-Square Test of Model Fit for the Baseline Model

          Value                           9132.568
          Degrees of Freedom                    66
          P-Value                           0.0000

SRMR (Standardized Root Mean Square Residual)

          Value                              0.008
```

Model Comparison

Model	AIC	BIC	RMSEA	CFI
1 Factor	36,747.899	36,823.019	0.125	0.867
2 Factor	35,933.855	36,181.232	0.079	0.958
3 Factor	35,667.586	35,967.596	0.047	0.988
4 Factor	35,593.207	35,940.587	0.025	0.998

Three Factor Model Results

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
F1 BY				
ITEM13	0.602	0.017	34.803	0.000
ITEM14	0.569	0.017	33.780	0.000
ITEM15	0.576	0.018	31.687	0.000
ITEM16	0.555	0.022	25.392	0.000
ITEM17	0.645	0.023	28.318	0.000
ITEM18	0.511	0.031	16.672	0.000
ITEM19	0.346	0.030	11.556	0.000
ITEM20	0.350	0.027	13.069	0.000
ITEM21	0.568	0.027	21.005	0.000
ITEM22	0.463	0.033	14.075	0.000
ITEM23	0.676	0.027	25.443	0.000
ITEM24	0.525	0.027	19.731	0.000
F2 BY				
ITEM14	0.032	0.020	1.613	0.107
ITEM15	0.093	0.034	2.730	0.006
ITEM16	0.120	0.030	3.948	0.000
ITEM17	0.241	0.116	2.072	0.038
ITEM18	0.623	0.203	3.064	0.002
ITEM19	0.628	0.156	4.026	0.000
ITEM20	0.408	0.128	3.179	0.001
ITEM21	0.359	0.211	1.705	0.088
ITEM22	0.479	0.269	1.780	0.075
ITEM23	0.119	0.395	0.301	0.763
ITEM24	0.093	0.405	0.231	0.818

F3 BY		
ITEM15	-0.038	0.057
ITEM16	-0.009	0.105
ITEM17	-0.144	0.183
ITEM18	-0.266	0.465
ITEM19	-0.203	0.472
ITEM20	-0.162	0.309
ITEM21	-0.271	0.270
ITEM22	-0.343	0.370
ITEM23	-0.532	0.080
ITEM24	-0.507	0.117

Wrapping Up

- Today we covered an introduction CFA and SEM
- The main point of this lecture is to show how each of these
  - Fits into a mixed modeling framework
    - Multivariate normal
  - Subsumes the EFA and CCA techniques used in the past
- The link between CFA/SEM and mixed models is important to understand
  - Latent variables = random effects (broadly construed)

Up Next

- <u>Tuesday:</u> Multidimensional Scaling, Classical Clustering, Distance Methods
- <u>Wednesday:</u> Latent class/finite mixture models
- <u>Thursday:</u> Missing data methods