LATENT VARIABLE MODELING

Session 5: Multiple group analyses

Contents

- Measurement invariance: Differences in measurement models between groups
- Structural invariance: Differences in means, variances or regressions between groups
- Model fit comparisons

Multigroup SEM

Categorical variables can be included within a SEM as:

- Endogenous variables
  - Discussed in last session
- Exogenous variable
  - Model main effect only:
    - Include (several) 0-1 coded variable as a variable
  - Model possible interaction effects:
    - Use multi-group SEM: Fit same model in each group and compare parameter estimates
    - or: Create interaction(s) before analysis and include as variables in model (not possible for every research question)

Multigroup SEM

Multi-group SEM allows for assessing parameter differences between groups:

- Measurement parameters
  - A.k.a. measurement invariance, measurement equivalence, differential item functioning
- Structural parameters
  - Means and variances of LVs
  - Regression relationships between OVs and/or LVs, e.g.,
    - differential prediction (e.g., intelligence test predicts functioning in job differently among males/females, majority/minority, …)
    - genetically informative design
    - …

Measurement invariance (MI)

- Are measurement parameters equal across
  - two or more groups (between-group MI), or
  - two or more measurement occasions (longitudinal MI)
- If not, there is measurement bias (lack of measurement invariance)
  - Observed score differences do not only reflect true differences in the construct of interest, but also reflect group membership

(lack of) measurement invariance

- The observed score on observed variable *i* (i.e., item or subscale) of person *j* is given by: $Y_{ij} = \tau_i + \lambda_i \eta_j + \epsilon_{ij}$
- Therefore: $\mathrm{E}(Y_{ij}|\eta_j) = \tau_i + \lambda_i \eta_j$
- If intercepts or loadings differ between groups

$$\tau_{ig} \neq \tau_{ig'} \text{ or } \lambda_{ig} \neq \lambda_{ig'}$$

and

$$\mathrm{E}(Y_{ijg}|\eta_j) \neq \mathrm{E}(Y_{ijg'}|\eta_j)$$

- Thus, given the same latent trait value, we would expect a different item score for a person in group *g*, than a person in group *g'*
- *We say:* X_i is a biased indicator of η with respect to group
  - In other words: differences in item scores reflect group membership, in addition to true score differences

(lack of) measurement invariance

- The observed score on item or subscale i, of person j is given by $Y_{ij} = \tau_i + \lambda_i \eta_j + \epsilon_{ij}$
- By definition, ϵ_i follows a normal distribution with mean 0 and variance $\sigma^2_{\epsilon_i}$
- When variance of measurement error differs over groups
  - No systematic bias in observed scores, but
  - Construct is not measured with same precision across groups (i.e., different reliability across groups)

(lack of) measurement invariance

- With CFA, we can statistically test whether the parameters of the measurement model are equal across groups
- We subsequently test for equality across groups of:
  1. Pattern of zero and non-zero loadings
     - 'configural' invariance
  2. In additon to 1: loadings (λ_i's)
     - 'metric' or 'weak' invariance
  3. In addition to 2: intercepts (τ_i's)
     - 'scalar' or 'strong' invariance
  4. In addition to 3: errror variances ($\sigma^2_{\epsilon_i}$'s)
     - 'uniqueness' or 'strict' invariance

(lack of) structural invariance

- We can also test whether structural coefficients are equal across groups:
  - equality of $\boldsymbol{\beta}$ (structural or latent regressions)
  - equality of $\boldsymbol{\Psi}$ (structural or latent (co)variances)
  - equality of $\boldsymbol{\alpha}$ (structural or latent means)

Mean structure

- Graphically: mean structure is represented by one or more triangles, which
  - Denote a constant with a value of 1
  - Have outgoing, single-headed arrow(s), of which the corresponding coefficient is the value of the intercept
- Algebraically: mean structure is represented by two vectors in lavaan:
  - $\boldsymbol{\nu}$ (nu; contains intercepts of observed variables)
  - $\boldsymbol{\alpha}$ (contains intercepts of latent or variables)

Mean structure

- As discussed earlier, the model-implied covariance matrix is given by

$$\hat{\boldsymbol{\Sigma}} = \boldsymbol{\Lambda}(\mathbf{I}-\boldsymbol{\beta})^{-1}\boldsymbol{\Psi}\left[(\mathbf{I}-\boldsymbol{\beta})^{-1}\right]^{\mathrm{T}}\boldsymbol{\Lambda}^{\mathrm{T}} + \boldsymbol{\Theta}$$

- The model-implied mean vector is given by

$$\hat{\boldsymbol{\mu}} = \boldsymbol{\Lambda}(\boldsymbol{\alpha} + \boldsymbol{\beta}\boldsymbol{\alpha}) + \boldsymbol{\nu}$$

- In CFA, there are no structural regression parameters, and the equations simplify to

$$\hat{\boldsymbol{\Sigma}} = \boldsymbol{\Lambda}\boldsymbol{\Psi}\boldsymbol{\Lambda}^{\mathrm{T}} + \boldsymbol{\Theta}$$

$$\hat{\boldsymbol{\mu}} = \boldsymbol{\Lambda}\boldsymbol{\alpha} + \boldsymbol{\nu}$$

Mean structure: Identification

For identification of the mean structure of a latent variable, we take a similar approach as for identification of the covariance structure:

- Standardized latent variable: Set intercept of LV to 0 (in addition to setting variance of LV to 1)
- Marker variable: Set intercept of an indicator variable to 0 (in addition to setting loading of indicator to 1)

Testing invariance

To test whether a set of parameters (loadings, intercepts, residual variances, latent (co)variances, or latent means) are equal across groups, we fit two models:

1. Model with parameters of interest **estimated freely** in both groups
2. Model with parameters of interest **restricted to equality** across groups
3. Assess difference in fit between models 1 and 2
   - $\chi^2(df)$, CFI, AIC, BIC and/or SSABIC

More restricted model will (almost always) have worse fit, but is it significantly or substantially worse?

Testing invariance

- Like model fit, tenability of MI is not an all-or-nothing question, researcher should make informed decision
- Rules-of-thumb offer a good starting point:
  - For evaluating configural invariance (as usual):
    Non-significant χ^2-value; CFI > .95; RMSEA < .06; SRMS < .08 (for good fit; can also use more lenient 'acceptable-fit' criteria)
  - For evaluating metric, scalar and uniqueness invariance:
    Non-significant $\Delta\chi^2(df)$, ΔCFI
    Model with lowest AIC or BIC value fits best

Chi-square difference test

- Statistical significance of difference in fit between two nested models can be assessed using $\Delta\chi^2(\Delta df)$ test

$$\Delta\chi^2 = \chi^2_{\text{model2}} - \chi^2_{\text{model1}}$$

$$\Delta df = df_{\text{model2}} - df_{\text{model1}}$$

- Nested model: all free parameters in less complex model are also free in more complex model
  - More complex model will have χ^2 equal to or lower than that of less complex model, by definition
  - More complex model also has lower *df*
- $\Delta\chi^2$ tests whether more complex model fits significantly better than less complex model
  - If so, retain more complex model
  - If not, retain less complex model

Testing invariance

- $\Delta\chi^2$ test often significant with larger sample sizes. Alternatives:
- Use AIC, BIC or RMSEA (lower value is better model)
- Use difference in CFI values:
  - Cheung and Rensvold (2000): ΔCFI > .01 indicates that null hypothesis of invariance should be rejected
  - Meade et al. (2008): ΔCFI > .002 indicates that null hypothesis of invariance should be rejected

Examples and exercises

- Example 4.4
- Exercise 4.1
  - Assessing measurement and structural invariance of the WISC (continuous indicators)

Reporting your results

- When reporting on your SEM model, you should provide at least two tables:
  - Table with indices of overal model fit indices
    - When doing model comparisons, also report differences in fit between models ($\Delta\chi^2$, *Δdf*, *ΔCFI*)

Example table for presenting the results

	χ^2	*df*	*p*	*CFI*	*TLI*	*RMSEA*	*BIC*	*AIC*
Model 1								
Model 2								
Model 3								
Model 4								

  - Table with parameter estimates (from your final, best-fitting model)
- Useful examples: see Van de Schoot, Lugtig and Hox (2012); Vandenberg and Lance (2000)

invariance testing with ordered-categorical indicators

- We subsequently test for equality across groups of:
  1. Pattern of zero and non-zero loadings
     - 'configural' invariance
  2. In additon to 1: loadings (λ_i's)
     - 'metric' or 'weak' invariance
  3. In addition to 2: threshold (τ_i's)
     - threshold invariance
  - Equality of residual variances is not tested
    - Delta parameterization: $\sigma_{\epsilon_i} = 1 - \lambda_i^2$. Thus, test of equal loadings also tests equality of residual variances.
    - Theta parameterization: $\sigma_{\epsilon_i} = 1$. All equal by default.
- Structural invariance tests as in continuous-indicator case:
  - equality of β (structural or latent regressions)
  - equality of Ψ (structural or latent (co)variances)
  - equality of α (structural or latent means)

Examples and exercises

- Additional Exercise: Invariance of the HADS
  - ordered categorical indicators
  - adding regressions and interactions to an LVM

References

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural equation modeling*, 9(2), 233-255.

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology*, 93(3), 568.

Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, 9(4), 486-492.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational research methods*, 3(1), 4-70.