LATENT VARIABLE MODELING

Session 5: Multiple group analyses

Multigroup SEM

Categorical variables can be included within a SEM as:

- Endogenous variables
  - Discussed in last session
- Exogenous variable
  - With main effect only:
    - Include (several) 0-1 coded variable as a variable
  - With possible interaction effect:
    - Or 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

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
  - Regression relationships between observed and/or latent variables, 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 no: There is measurement bias (lack of measurement invariance)
  - Observed score differences do not exclusively reflect true differences in the construct of interest, but also group membership

(lack of) measurement invariance

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

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

and

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

- 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 cannot only be attributed to true differences

(lack of) measurement invariance

- The observed score on item or subscale *i*, of person *j* is given by $X_{ij} = \tau_i + \lambda_i \eta_j + \epsilon_{ij}$
- By definition, ϵ follows a normal distribution with mean 0 and variance σ_{ϵ_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 (reliability differs 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's)
     - 'metric' or 'weak' invariance
  3. In addition to 2: intercepts (τ's)
     - 'scalar' or 'strong' invariance
  4. In addition to 3: errror variances (σ_ϵ'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:
  - $\boldsymbol{\nu}$ (contains intercepts of observed variables)
  - $\boldsymbol{\alpha}$ (contains means 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 value of the latent mean to 0 (in addition to setting variance of LV to 1)
- Marker variable: Set value of the intercept of an indicator variable to 0 (in addition to setting loading of indicator to 1)
- Effects coding: Set sum of intercept values to 0 (in addition to setting sum of loadings equal to number of items)

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 be equal across groups
3. Assess difference in model 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:
- Rules for assessing configural invariance (as usual):
  - non-significant χ^2-value; CFI > .95; RMSEA < .06; SRMS < .08
- For metric, scalar and uniqueness invariance, we have to assess difference in model fit using $\chi^2(df)$, CFI, AIC, BIC and/or SSABIC

Chi-square difference test

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

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

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

- Nested means that all free parameters in less complex model are also free in more complex model
  - More complex model can always approximate observed sample daat better, so χ^2 value of more complex model always lower
  - 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 most complext model
  - If not, retain least complex model

Testing invariance

- Delta chi-square 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 is robust statistic for testing tenablity of MI restrictions:
    - 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 (see github)

Example 4.6

Reporting your results

- When reporting on your SEM model, you should provide at least two tables:
  1. Table with indices of overal model fit
     In case of MI testing, also report differences in fit between models ($\Delta\chi^2$, Δdf, ΔCFI)
  2. Parameter estimates (from your final, best-fitting model)

Example table for presenting the results

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

- Useful examples: see Van de Schoot, Lugtig and Hox (2012); Vandenberg and Lance (2000)

Genetically informative design

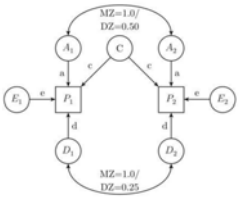
- Goal: estimate to what extent an observed characteristic (phenotype *P*; e.g., BMI) can be explained by
  1. additive effects of genes (*A*; narrow heritability)
  2. non-additive effects of genes (*D*) and shared environmental effects (C)
     - *D* and C cannot be distinguished if we only have twins that were raised together, in the same environment
     - *A* + *D* together are 'broad heritability'
  3. random environmental and measurement effects (*E*)

Genetically informative design

- P: phenotype
  - Variable / characteristic of interest
  - Only OV in the model
  - 2 variables: one for each sibling
- A: additive genetic effects
  - Assumed identical for MZ siblings (ρ = 1)
  - Assumed correlated between DZ siblings (ρ = .5)
- D: non-additive genetic effects
  - Assumed identical for MZ siblings (ρ = 1)
  - Assumed correlated between DZ siblings (ρ = .25)
- C: shared environment effects
  - Assumed identical for both MZ and DZ siblings (thus single variable)
- E: error
  - Non-shared environmental effects and measurement error
  - Assumed independent between siblings (thus two variables)

Genetically informative design

- Multigroup analysis: MZ vs DZ twins
- Note: C often cannot be distinguished, as often only twins raised in same environment
  - Then C not explicitly modelled
  - Estimates for A and/or D are then contaminated by C
- Note: Variances of A, C, D and E are unknown, but assumed equal for twins 1 and 2
  - Plausible, looking at Table 4.4?



P: phenotype
A: additive genetic effects
C: shared environment effects
D: non-additive genetic effects
E: non-shared environment effects + measurement error

Genetically informative design

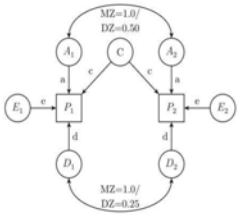
Table 4.4 Body Mass Index Covariances for Monozygotic (n = 534) and Dizygotic (n = 328) Twins Reared Together.

		Twin 1	Twin 2
MZ	Twin 1	0.725	0.589
	Twin 2	0.589	0.792
DZ	Twin 1	0.779	0.246
	Twin 2	0.246	0.837

Genetically informative design

Not all parameters can be estimated at the same time, so models:

- ADE model
- ACE model
- AE model
- CE model



P: phenotype
A: additive genetic effects
C: shared environment effects
D: non-additive genetic effects
E: non-shared environment effects + measurement error

Genetically informative design

- Conclusions ADE model:
  - A (additive gene effects) explain ($.636^2 \approx$) 40% of variance in BMI
  - D (non-additive gene effects) explain ($.615^2 \approx$) 38% of variance in BMI
    - But note that estimate of D is confounded by C (shared environmental effects)
  - E (random, unexplained influences) explain (.218 ≈) 22% of variance in BMI
  - Thus, 40% + 38% + 22% = 100% of variance in BMI is accounted for in the model

Exercises

See exercises on github:

- 4.2: Genetically informative design
- Additional: Testing measurement invariance with ordered categorical items + regression with latent variables

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.