LATENT VARIABLE MODELS

4: Ordered categorical indicator variables

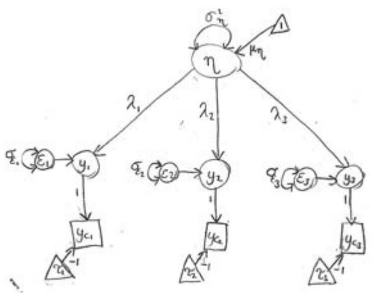
Ordered categorical indicator variables

- Up till now, endogenous variables have always been continuous
- Often variables in psychology will not be continuous, but (ordered) categorical
  - Exogenous ordered categorical variables:
    - Treat as if continuous
    - Code as (multiple) dummy (0-1) variables
  - Endogenous variables: need different model

Ordered categorical indicator variables

- The regression formula for the continuous item response of person j on item i is given by: $Y_{ij} = \tau_i + \lambda_i \eta_j + \epsilon_{ij}$
- A dichotomous (categorical) response $Y_{c,i}$ can only take values 0 or 1
  - Or 0, 1, 2, … for > 2 ordered categorical values
- Solution: we assume a continuous LV Y_i underlies categorical item response $Y_{c,i}$, which is linearly dependent on η
- The categorical $Y_{c,i}$ has a threshold τ_i
  - If $Y_i > \tau_i$ then $p(Y_{c,i}=1) > .5$
  - If $Y_i < \tau_i$ then $p(Y_{c,i}=1) < .5$
- We assume Y_i follows a normal distribution with $\mu = 0$ and $\sigma = 1$, to identify its scale and estimate the loading λ_i and threshold τ_i
- In words: instead of standard covariance matrix, we calculate a tetra- or polychoric correlation matrix and perform CFA as in the linear / continuous variable case (DWLS approach)

CFA with dichotomous indicator variables: graphical representation



Available sample statistics:
3+2+1 = 6 (co)variances
3 means

Model parameters:
1 mean of eta
1 variance of eta
3 loadings
3 error variances
3 thresholds

Identification restrictions:
- Standardize all Yi (σ=1, μ=0)

and

- Standardize η (σ=1, μ=0)

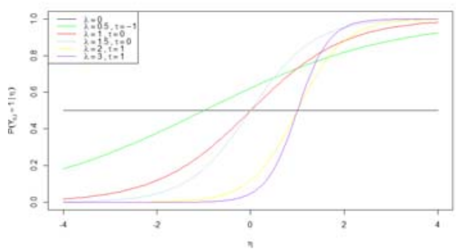
or

- Marker variable (λ=1, τ=0)

Item characteristic curves (ICCs)

To get from LV (η) to observed dichotomous response ($Y_{c,i}$):

$$p(Y_{c,i} = 1|\eta) = \frac{e^{\lambda(\eta-\tau)}}{1+e^{\lambda(\eta-\tau)}}$$



Identifying scale of underlying latent variable

- By definition:

$$\sigma_{y_i}^2 = \lambda_i^2 \sigma_{\eta_j}^2 + \sigma_{\varepsilon_i}^2 \qquad \Delta_i = \frac{1}{\sigma_{y_i}^2} \qquad \sigma_{y_i}^2 = 1$$

- 'Delta', or marginal, parameterization assumes $\Delta_i = 1$

  so $\sigma_{\varepsilon_i}^2 = 1 - \lambda_i^2 \sigma_{\eta_j}^2$

- 'Theta', or conditional, parameterization assumes $\sigma_{\varepsilon_i}^2 = 1$

  - Delta parameterization is more natural from FA viewpoint parameterization is more natural from IRT viewpoint

Common IRT and FA parameterizations

Parameterization	Latent Variable (θ) Identification: Marker Variable	Latent Variable (θ) Identification: Standardized
Delta or marginal:	$\alpha = \frac{\lambda\sqrt{\sigma_\theta^2}}{\sqrt{1-\lambda^2\sigma_\theta^2}}$	$\alpha = \frac{\lambda^2}{\sqrt{1-\lambda^2}}$
	$\beta = \frac{-(\tau-\lambda\mu_\theta)}{\sqrt{1-\lambda^2\sigma_\theta^2}}$	$\beta = \frac{-\tau}{\sqrt{1-\lambda^2}}$
Theta or conditional:	$\alpha = \lambda\sqrt{\sigma_\theta^2}$	$\alpha = \lambda$
	$\beta = -(\tau-\lambda\mu_\theta)$	$\beta = -\tau$

Some programs estimate $-\tau$ instead of τ. For those estimates, use τ instead of $-\tau$ in the conversion formulae. To obtain b, use the conversion: $b = -\beta/\alpha$

What to remember?

Several ways to scale and identify the model, but all are alike:

- Values of slope, discrimination parameters a and α increase together
  - Higher values: better discrimination
- Values of threshold, difficulty parameters b and β increase together
  - Higher values: more difficult item

Examples and exercises

- Examples chapter 6 (see github)
- Exercise 6.1 + additional questions a & b (see github)

Ordinal responses

- We may have items with four ordered response options: a < b < c < d
- We model the probability of endorsing each response option:

$$P(Y_i \geq a \mid \eta) = 1 \qquad P(Y_i \geq c \mid \eta) = \frac{e^{\lambda_i(\eta-\tau_{i_c})}}{1+e^{\lambda(\eta-\tau_{i_c})}}$$

$$P(Y_i \geq b \mid \eta) = \frac{e^{\lambda_i(\eta-\tau_{i_b})}}{1+e^{\lambda(\eta-\tau_{i_b})}} \qquad P(Y_i \geq d \mid \eta) = \frac{e^{\lambda_i(\eta-\tau_{i_d})}}{1+e^{\lambda(\eta-\tau_{i_d})}}$$

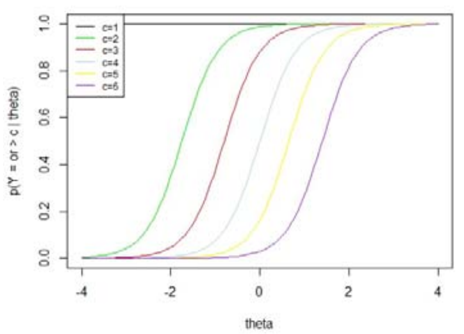
- This gives us the probabilities for each response option:

$$P(Y_i = a) = P(Y_i \geq a) - P(Y_i \geq b) \qquad P(Y_i = c) = P(Y_i \geq c) - P(Y_i \geq d)$$

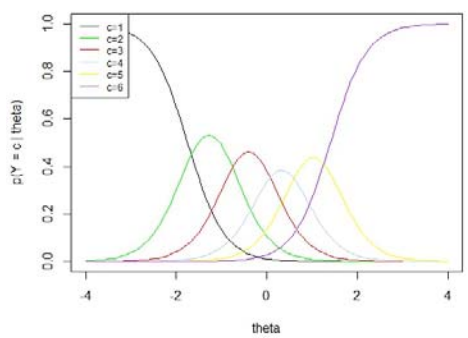
$$P(Y_i = b) = P(Y_i \geq b) - P(Y_i \geq c) \qquad P(Y_i = d) = P(Y_i \geq d)$$

- So, for every item with *k* ordered categories, we need to estimate one loading, and *k-1* thresholds

Ordinal responses



Ordinal responses



Ordinal responses

- For every item with *k* ordered categories, we need to estimate one loading, and *k-1* thresholds
- In lavaan, we use the same approach as with dichotomous data: use 'ordered = ….' argument
  - For every item declared ordered, lavaan checks number of categories, and estimates *k*-1 thresholds

Categorical FA vs. IRT

Historical differences:

- Estimation:
  - Item response theory: maximum likelihood (ML)
    - Estimates model parameters in one step
    - Not available for ordered categorical indicators in lavaan
  - Factor analysis: diagonally weighted least squares (DWLS)
    - Estimates tetra- or polychoric correlation matrix, performs continuous variable CFA on that matrix
    - Only option for ordered categorical indicators in lavaan
- Parameterization:
  - In IRT, latent trait is scaled by assuming mean 0 and variance 1
  - In CFA, latent trait is often scaled by setting loading of first item to 1
  - What we call loadings and thresholds in CFA, we call discrimination and difficulty parameters in IRT

IRT models

- Binary items:
  - 1PL, or Rasch model (loadings equal, thresholds free)
  - 2PL (loadings free, thresholds free)
  - …
- Polytomous items:
  - Partial credit mo[…]sholds free)
  - Graded response[…]thresholds free)
  - …

Georg Rasch (1901-1980)

No Rasch, no good?

- Often in psychology, we want to use the test score: the (unweighted) sum of item scores
  - Easy to calculate, you need no IRT or SEM software to estimate it
- In the Rasch model, all item loadings are equal, so all item scores contribute equally to estimation of the latent trait
  - Test score is 'sufficient statistic' for eta (latent trait)
    - "no other statistic that can be calculated from the same sample provides any additional information as to the value of the parameter"
  - Well-fitting Rasch model: test score contains all information about latent trait

IRT: from binary to ordered categorical responses

- Partial credit model is the Rasch model generalized to polytomous items
  - Same loadings for all items
  - Freely estimates thresholds for all categories and items
- Graded response model is the 2pl model generalized to polytomous items
  - Freely estimates loadings for all items
  - Freely estimates thresholds for all categories and items
- Note: Unlike in Rasch model, in PCM test score is not a sufficient statistic (does not contain all information about) for the latent trait (eta)

Examples and exercises

- Exercise 6.1 additional c & d (see github)
- Exercise 6.2 (see github)
- Additional exercise HADS (see github)