Person-Centered Polytomous IRT for Detection of Differential Item Functioning

Tolu Sajobi

Rasch Model

- The partial credit model describes the probability that the ith subject with person parameter θ_i endorse the kth $(k = 1, \dots, K)$ response category on the jth $(j = 1, \dots, J)$ item

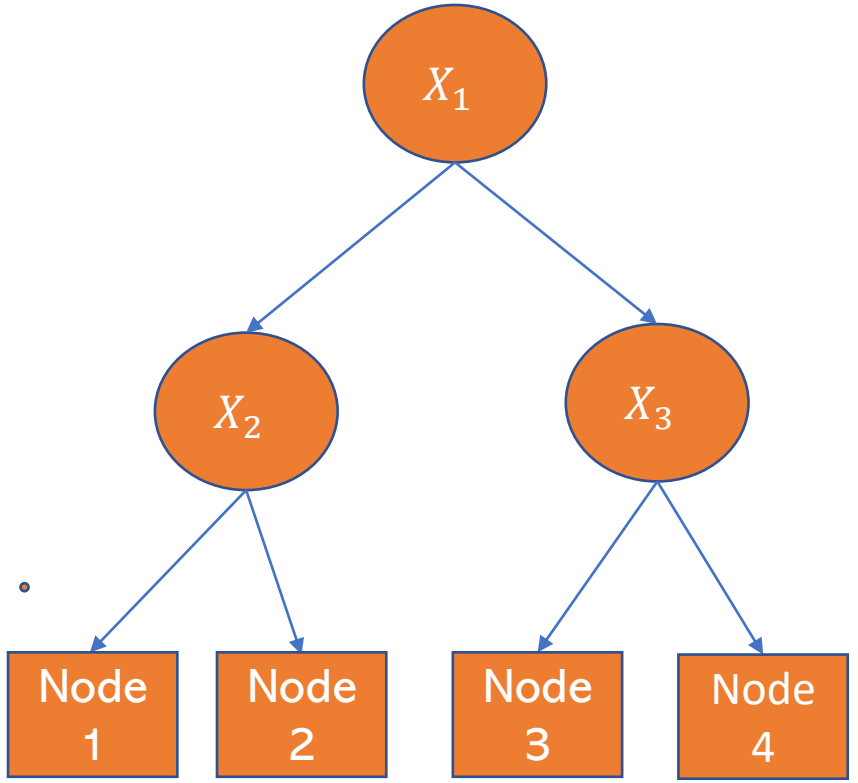
$$P\left(Y_{ij} = k \middle| \theta_i, \tau_j\right) = \frac{e^{\sum_{k=1}^{K}(\theta_i - \tau_{jk})}}{\sum_{l=0}^{K} e^{\sum_{k=0}^{l}(\theta_i - \tau_{jk})}},$$

where τ_{jk} is the kth threshold for item j.

Person-Focused IRT Tree

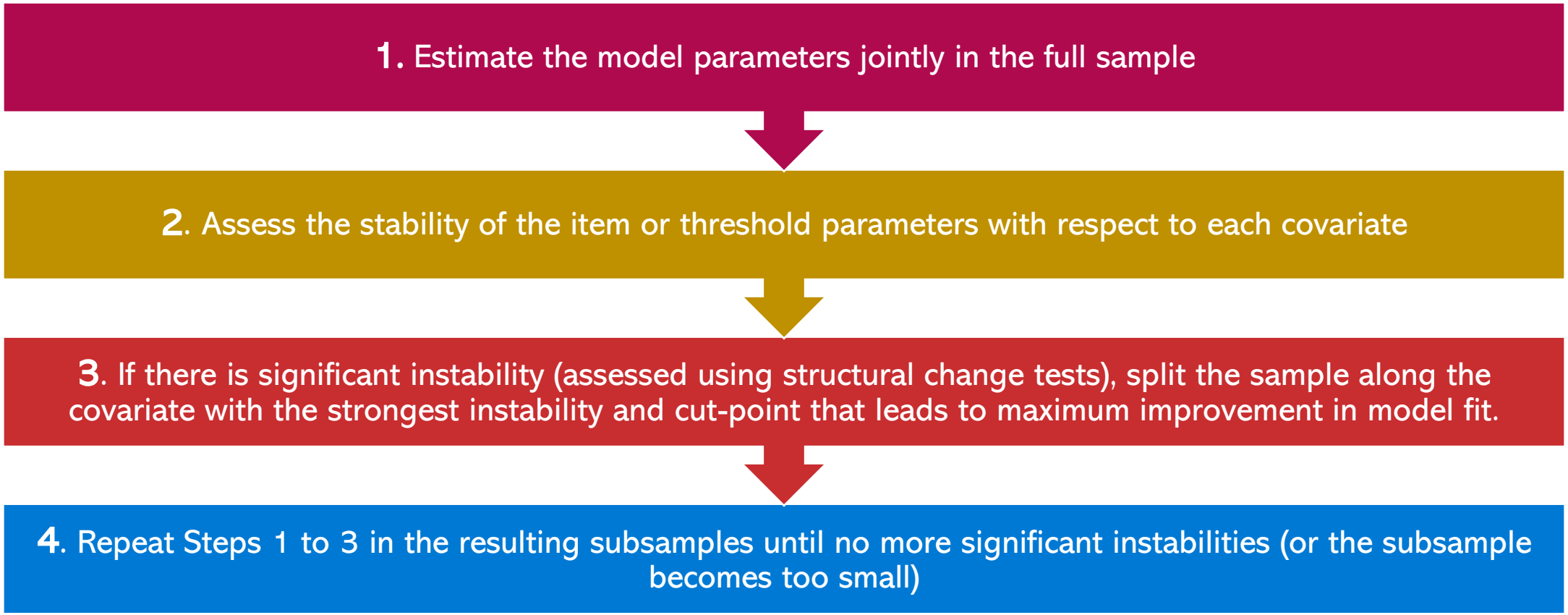
- A semi-parametric methodology that combines structural change tests and model-based on recursive partitioning framework to identify homogenous subgroups based on a set of covariates

- In variable-focused IRTree model, the sample space is partitioned in such a away that the IRT model parameters are non-invariant across the subgroups



Zeileis A, Hothorn T, Hornik K. Model-based recursive partitioning. Journal of Computational and Graphical Statistics 2008; 17: 492-514.

Polytomous Rasch Tree - Implementation



Strobl C, Kopf J, Zeileis A. Rasch trees: a new method for detecting differential item functioning in the Rasch model. Psychometrika 2015; 80:289-316.

Stopping Criteria

- Parameter Instability
  - Stop splitting if there is no significant parameter instability on any of the covariates
  - Statistical significance usually evaluated at 5%.
  - Bonferroni adjustment for the p-values for each test is recommended to control the over familywise Type I error

- Minimum sample size per node
  - Minimal node side chosen to provide sufficient sample size within each node for parameter estimation

  - Rule of Thumb: no fewer than 10 subjects per model parameter (Komboz, Strobl, & Zeileis, 2018)

Structural Change Tests

- Originally developed in Econometrics for detecting signals (e.g., drop in stock prices)

- Used to test whether the item parameters vary between groups of subjects defined by covariates

1. Estimate the item parameters jointly from the the original entire sample.

2. The individual deviations from this joint model are ordered with respect to a covariate

3. If there is systematic DIF with respect to the covariate, the ordering will exhibit a systematic change in the individual deviations.

4. If, on the other hand, no DIF is present, the values will merely fluctuate randomly.

Selecting Optimal Cutpoints for Covariates

- After a variable has been selected for splitting, the optimal cutpoint is determined by maximizing the partitioned log-likelihood over all candidate cutpoints within the range of this splitting variable.
  - The partitioned log-likelihood is the sum of the log-likelihoods for two separate models
  - Also equivalent to using the maximum likelihood ratio statistic that compares the joint model (for the entire sample) to the partitioned model.

Important Considerations

- Simulation studies have shown that the person-focused IRTree method exhibits comparable Type I error and statistical power to the multigroup IRT method (Komboz et al, 2018).

- The person-focused IRTree is sensitive to the type and number of covariates.
  - Changing the number and type of covariates could result in different conclusions about the number of subgroups identified
  - Prior knowledge of potentially relevant covariates might be useful

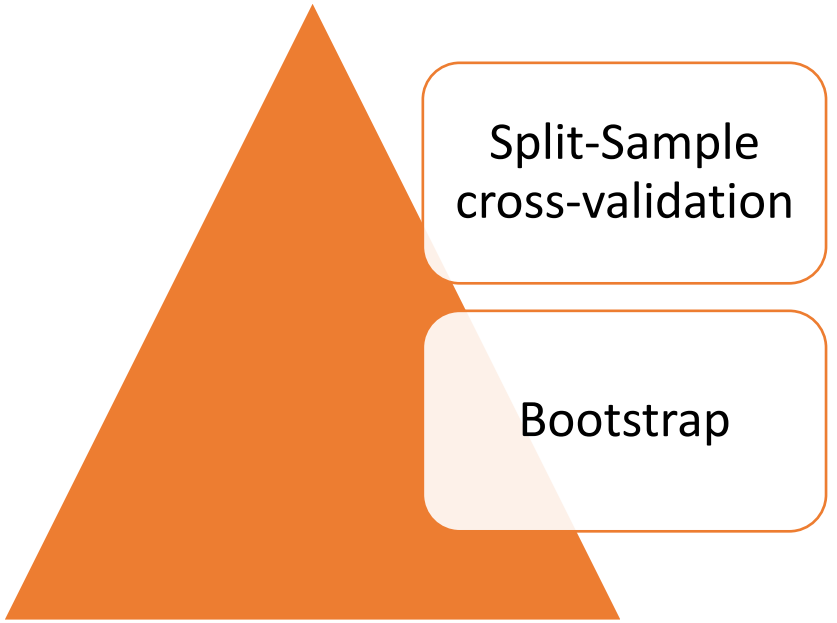
- The default minimum terminal node size might result in low statistical power to detect non-invariant subgroups.
  - Rule of thumb: At least 10 individuals per parameter
  - Depending on the sample size, larger terminal node size is recommended.
  - Further simulations are still needed to determine the optimal minimum node size

- Person-focused IRTrees are prone to overfitting. Model validation approaches are recommended for handling overfitting issues.

Potential Internal Cross-validation Approaches for IRTrees

- Split sample cross-validation (e.g., training and test data)
  - Randomly split the original sample into train and test data (60:40, 70:30)
  - Train the person-focused IRTree on a portion of the data. Use the identified splitting variables from the training data as input variables when applying the model to the test data
  - Spilt-sample cross-validation may be less powerful in small-sampled studies

- Resampling-based cross-validation (bootstrap validation, repeated cross-validation)
  - Generate bootstrap samples from the original data
  - Apply the IRTree model to each bootstrap sample and document splitting variables in each bootstrap sample.
  - Rank the input variables according to the proportion of times they are ranked the most important splitting variable across the bootstrap samples



External Validation

- External validation of unsupervised machine learning models remains a challenge
  - Unlike supervised learning methods, there is no known “gold standard” for the identified subgroups in unsupervised learning methods
- The key goal of person-focused IRTree models is the identification of variables that contribute to heterogeneity in the data
- It is unclear how this class of models could be external validated in other samples with different degrees and magnitude of sample heterogeneity

Strengths and Limitations of Person-focused IRTrees

- Strengths
  - The person-focused IRTree method is useful for exploratory analyses for hypothesis generation
  - The method not require a priori specification of group-variables associated with DIF
  - The method is appropriate for identifying subgroups of individuals with differential responses to PROMs
    - Can be extended to detect DIF items too!

- Limitations
  - Sensitive to the type and number of covariates
  - Does not identify the PROMs items that exhibited DIF
  - The IRTree method relies on the assumption of unidimensionality of the item responses. Not applicable for testing DIF in multidimensional patient-reported outcomes measures items
  - Prone to overfitting. Internal validation is recommended.

References

- Strobl C, Kopf J, Zeileis A. Rasch trees: a new method for detecting differential item functioning in the Rasch model. Psychometrika 2015; 80:289-316.

- Komboz B, Strobl C, Zeileis A. Tree-based global model tests for polytomous Rasch models. Educational and Psychological Measurement 2018; 78: 128 -166.

- Merkle EC, Zeileis A. Tests of measurement invariance without subgroups: A generalization of classical methods. Psychometrika 2013;78:59-82.

- Zeileis A, Hothorn T, Hornik K. Model-based recursive partitioning. Journal of Computational and Graphical Statistics 2008; 17: 492-514.