ISOQOL 2022
Prague, Czech Republic

Item-Focused Machine-Learning Models for Detection of Differential Item Functioning

Lisa Lix

Section Objectives

- To review concepts relevant to tree-based analyses
- To describe item-focused tree-based models
- To examine the results of analyses to test for uniform and non-uniform DIF using item-focused tree-based models

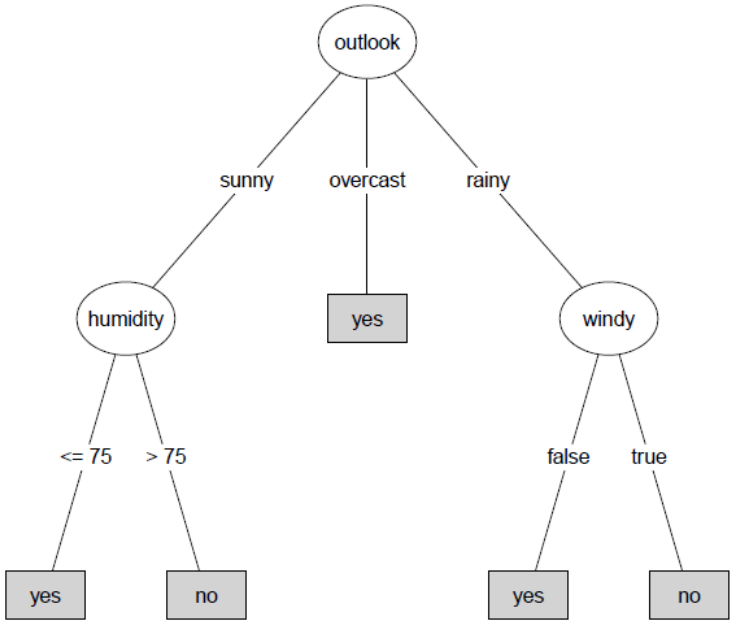
Trees are Easy to Interpret

- The ability to represent a DIF model as a tree contributes to its interpretability
  - Tree-based models mimic the approaches we use for decision making

- However, there are some limitations of tree-based models
  - Most notably, they are prone to over-fitting
  - Model validation approaches are essential to adopt

Elements of a Tree

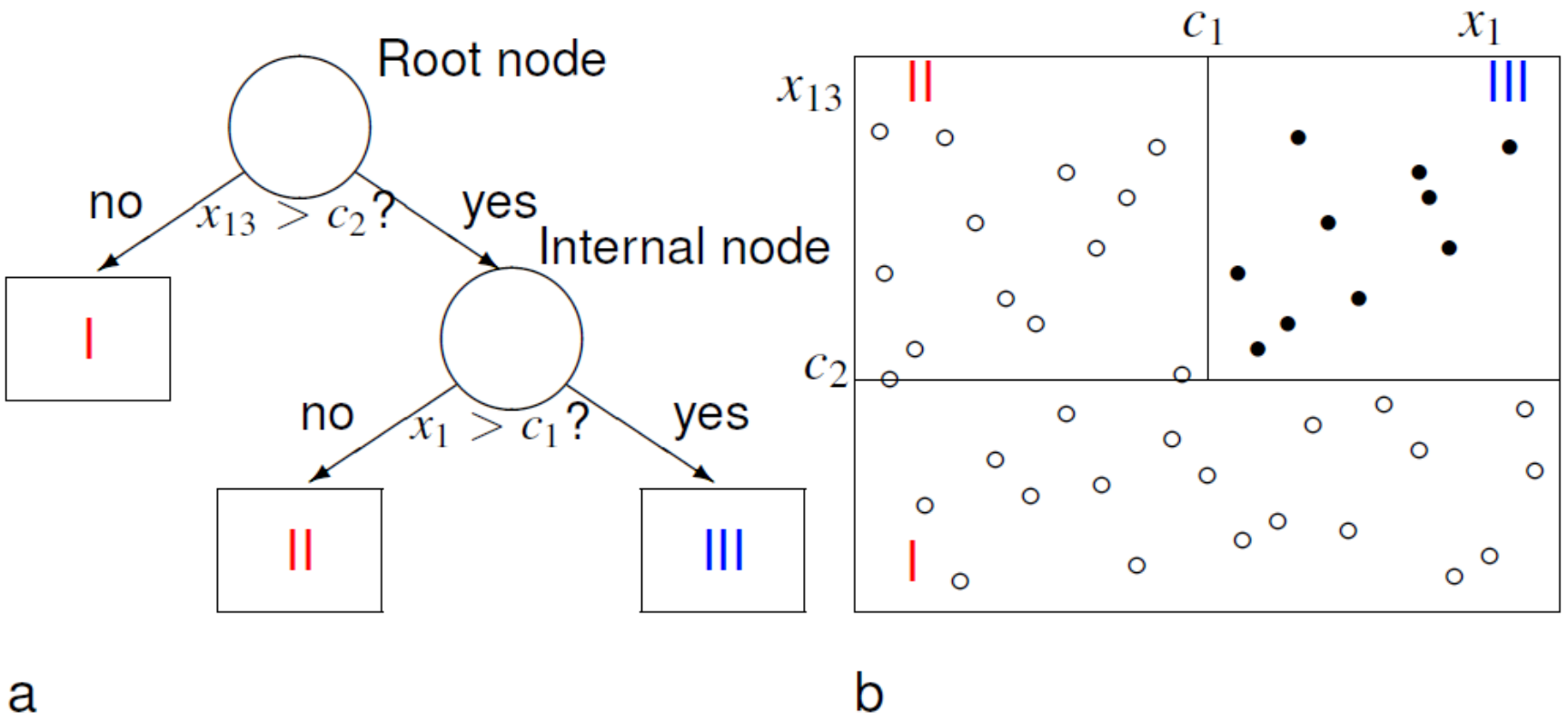
I know what the weather is like outside . . . should I play?



Can you identify the elements of this tree?

- Root node
- Terminal node
- Left and right daughter nodes
- Offspring nodes
- Split

Elements of a Tree



Interpreting a Tree

- Contents of the nodes
  - All nodes in the same layer constitute a partition of the root node
  - Every node in a tree is a subset of the dataset
- Splitting a parent node into daughter nodes
  - Splitting occurs when there is enough instability in the parameter estimate
- When does a node become a terminal node
  - Terminal nodes are homogeneous with respect to some criterion

Splitting a Node

- Consider the variable age, and assume that it has values that range from 17 to 48
  - There are 31 different values for age, and accordingly, 31-1 = 30 allowable splits on this variable
- For a binary variable, there is only one allowable split
- For a polytomous variable, the number of allowable splits is one less than the number of categories

Model-Based Recursive Partitioning

- The feature space (i.e. the space spanned by all predictor variables) is partitioned into a set of regions
- The partitions represent observations with similar responses

- After the data are partitioned, one model is fit to the data in each region

- **Assumptions:** A global model for the data does not fit the data well AND covariates are available to describe and partition the data

Logistic Regression Model for DIF Detection

Consider the conventional logistic regression model for the *i*th PRO item:

$$log\left(\frac{P(Y_{pi}=1|S_p,g_i)}{P(Y_{pi}=0|S_p,g_i)}\right) = \eta_{pi} = \beta_{0i} + \beta_{1i}S_p + \gamma_{ig_i}$$

$Y_{pi} \in \{0,1\}, p = 1,\ldots,P, i = 1,\ldots,I$ denote the response for person p on item i, g_i denotes group membership for the *i*th item, S_p is the overall score for person p, β_{0i} denotes the intercept for the *i*th item, β_{1i} denotes the slope of item i, and γ_{ig_i} denotes the group-specific parameter

The likelihood ratio test statistic is used for DIF detection. For example, to test for uniform DIF:

$H_0 : \gamma_{i1} = \gamma_{i2} = \ldots = \gamma_{iG} = 0$, where G is the total number of groups

Recursive Partitioning with Logistic Regression for DIF Detection

- A single global logistic regression model does not fit the data well in all regions

- The data are partitioned into regions, and in each region a logistic regression model is fit to the data
- To assess whether splitting of a node is necessary, a fluctuation test for parameter instability is performed. If there is significant instability with respect to any of the partitioning variables, split the node into B locally optimal segments and repeat the procedure
- Given the correct partition of the data, the parameter estimates that minimize the global objective function can be achieved by computing the locally optimal parameter estimates in each segment

Item-Focused Trees versus Person-Focused Trees

- **Item-Focused Trees:** use recursive partitioning at the item level; flags DIF items and the variables that contribute to DIF on individual items
- **Person-Focused Trees:** the covariate space is recursively partitioned to identify regions of the covariate space in which item parameters differ; in the investigated regions a parametric model that includes covariates is fit to the data

Choosing between Item-Focused Trees and Person-Focused Trees

- The choice between an item-focused method and a person-focused method depends on the question being asked and the characteristics of the data

- If you expect that a single item may be contributing to DIF, use the item-focused method
  - You may end up removing or rewording an item based on the outcome of the analysis
  - This method may be beneficial to the development of item banks

- If you generally believe that DIF is present in an instrument, then use a person-focused method
- Practical issues: The person-focused method is computationally more efficient to implement than the item-focused method; the algorithm for the former method will be easier to run than the item-focused algorithm if sample size is large

Item-Focused Tree Model

- The algorithm recursively splits data across the items and stops once there are no more statistically significant splits
- First split: Examine all items, variables and possible splits in a variable, and fit logistic regression models with predictor

$$\eta_{pi} = \beta_i \theta_p + \left[\gamma_{i1} I\left(x_{pj} \leq c_{ijk}\right) + \gamma_{i2} I\left(x_{pj} > c_{ijk}\right)\right],$$

$$j = 1, \ldots, m;\ k = 1, \ldots, K_j$$

  - $I(\,)$: an indicator function
  - c_{ijk}: kth cutpoint in the *j*th variable for the *i*th item

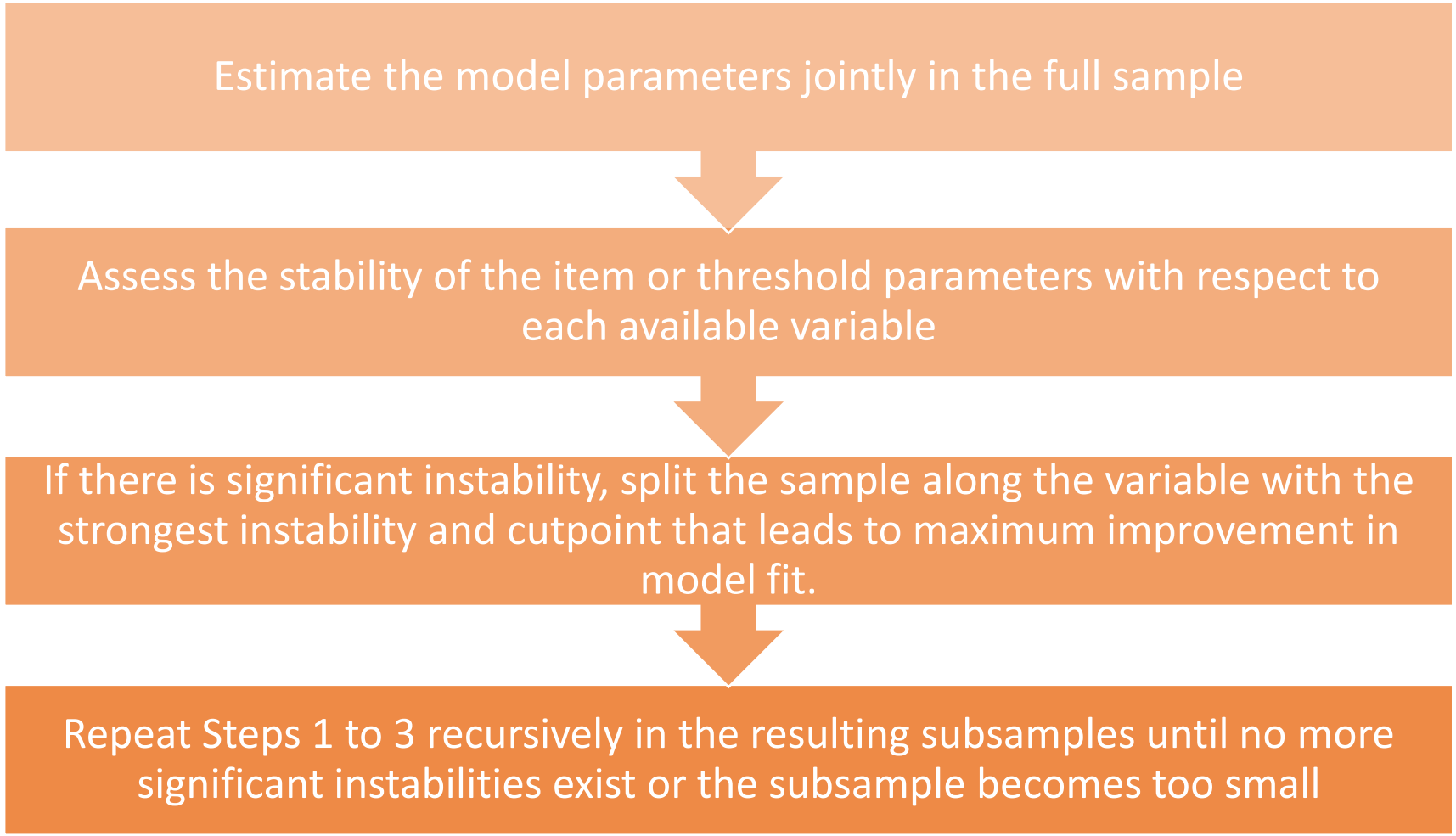
Item-Focused Tree Model

- The model does not require *a priori* theoretical or clinical evidence about heterogeneity in the population of interest, unlike conventional DIF models such as the logistic regression model

- The model combines a tree-based recursive partitioning model and an IRT model to detect DIF

- This model produces trees to visualize the structure of DIF in an item and potential interactions of the variables that generate DIF

Item-Focused Tree Model Implementation



Multiple Testing

- There are many significance tests being performed! The familywise error rate, that is, the overall probability of a Type I error, can therefore be large
- It may be beneficial to adopt a method to control the familywise error rates
  - For example, the Benjamini-Hochberg procedure is one option to evaluate the significance of variables for DIF on each item
  - This procedure is used for simultaneous testing of multiple independent hypotheses; it can control the false discovery rate or the expected ratio of false rejections to the number of total rejections

Numeric Example

- Model-based recursive partitioning using logistic regression was applied to physical health and mental health component items of the SF-12
- Refer to R script and output for:
  - Recoding of the data – to create a binary response variable
  - Covariates:
    - Age: continuous
    - Sex: 0 = female; 1 = males
  - Both uniform and non-uniform DIF models were considered

References

Berger M, Tutz G. Detection of uniform and nonuniform differential item functioning by item-focused trees. J Educ Behav Stat. 2016;41(6):559–92.

Komboz B, Strobl C, Zeileis A. Tree-based global model tests for polytomous Rasch models. Educ Psychol Meas. 2018;78(1):128–66.

Magis D, Tuerlinckx F, De Boeck P. Detection of differential item functioning using the lasso approach. J Educ Behav Stat. 2015;40(2):111–35.

Tutz G, Berger M. Item-focussed trees for the identification of items in differential item functioning. Psychometrika. 2016;81(3):727–50.

Yadegari I, Bohm E, Ayilara OF, Zhang L, Sawatzky R, Sajobi TT, et al. Differential item functioning of the SF-12 in a population-based regional joint replacement registry. Health Qual Life Outcomes. 2019;17(1):1–11.