Trees and Rasch-Trees

Carolin Strobl, Julia Kopf and Achim Zeileis

LMU

CEST
Center for Empirical Studies

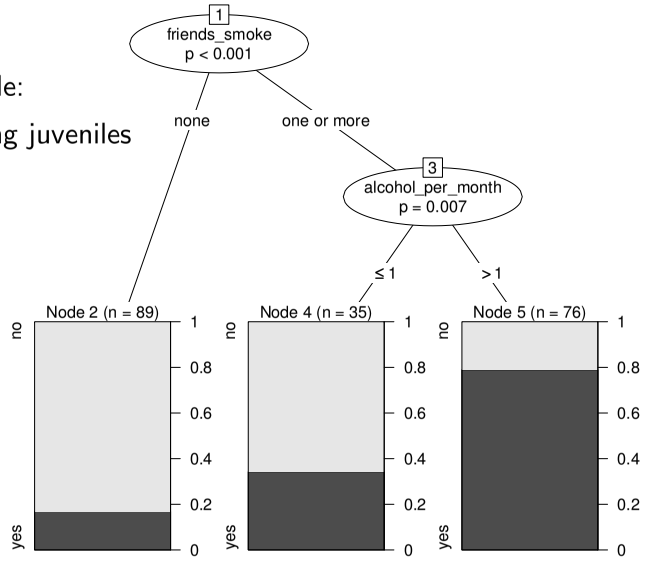
Outline

Trees and Rasch-Trees
Carolin Strobl, Julia Kopf, Achim Zeileis
CART
MOB
Testing for DIF
Graphical test
Latent classes
Model-based partitioning
Example
Conclusion
Literatur

Classification and regression trees (CART)

Example:

Smoking juveniles



Trees and Rasch-Trees
Carolin Strobl, Julia Kopf, Achim Zeileis
CART
MOB
Testing for DIF
Graphical test
Latent classes
Model-based partitioning
Example
Conclusion
Literatur

Optimal cutpoint selection

Example: $\chi^2 - test$

operationalization 1:

X binary, Y binary

		Y	
		1	2
X	1	h_{11}	h_{12}
	2	h_{21}	h_{22}

$$\sum_i \sum_j \frac{\left(h_{ij} - \tilde{h}_{ij}\right)^2}{\tilde{h}_{ij}} \sim \chi^2(1)$$

Trees and Rasch-Trees
Carolin Strobl, Julia Kopf, Achim Zeileis
CART
MOB
Testing for DIF
Graphical test
Latent classes
Model-based partitioning
Example
Conclusion
Literatur

Optimal cutpoint selection

Example: $\chi^2 - test$

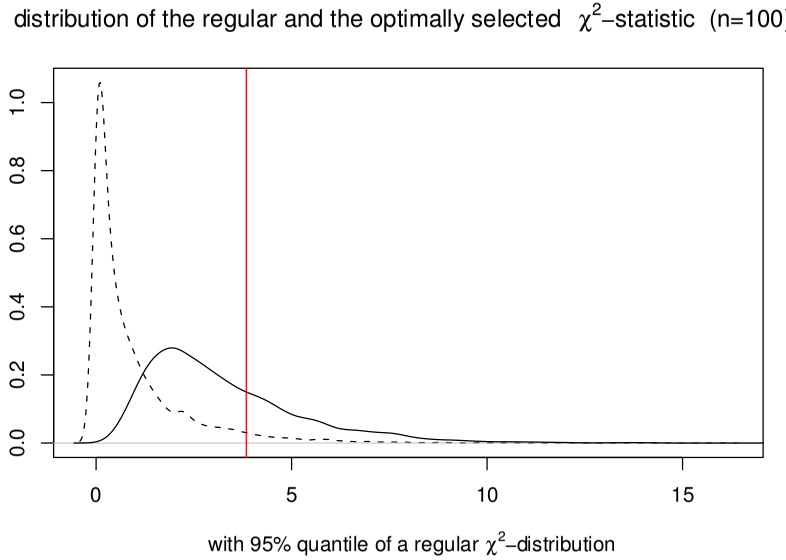
operationalization 2:

X continuous, Y binary

		Y	
		1	2
X	$x \leq c$	h_{11}	h_{12}
	$x > c$	h_{21}	h_{22}

$\sum_i \sum_j \frac{(h_{ij} - \tilde{h}_{ij})^2}{\tilde{h}_{ij}} \sim \chi^2(1)$???

Optimal cutpoint selection



Trees and Rasch-Trees
Carolin Strobl, Julia Kopf, Achim Zeileis
CART
MOB
Testing for DIF
Graphical test
Latent classes
Model-based partitioning
Example
Conclusion
Literatur

Variable selection bias

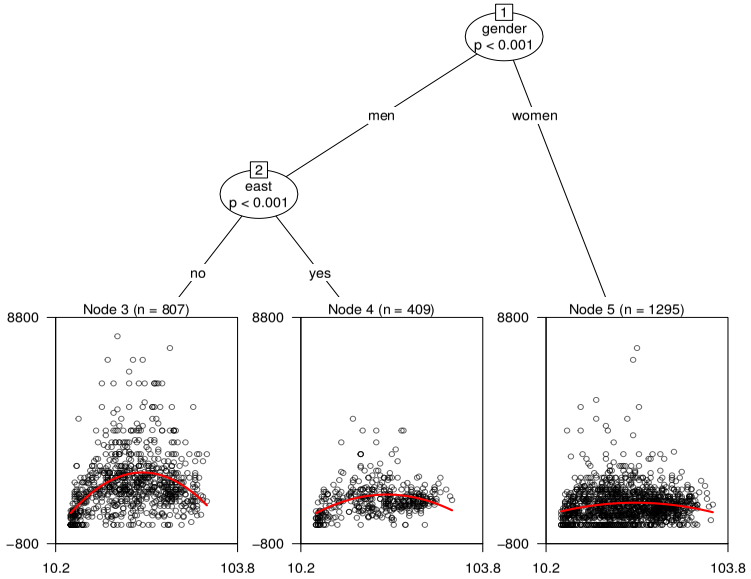
combination of optimal cutpoint and variable selection leads to bias in the standard algorithms:

variables with many categories and continuous variables are systematically preferred (cf., e.g., White & Liu, 1994; Loh & Shih, 1997; Jensen & Cohen, 2000; Kim & Loh, 2001; Dobra & Gehrke, 2001; Strobl, Boulesteix & Augustin, 2007)

this bias affects the R-packages `tree` (Ripley, 2007) and `rpart` (Therneau & Atkinson, 2006), but not function `ctree` in package `party` (Hothorn, Hornik, Strobl & Zeileis, 2011)

Trees and Rasch-Trees
Carolin Strobl, Julia Kopf, Achim Zeileis
CART
MOB
Testing for DIF
Graphical test
Latent classes
Model-based partitioning
Example
Conclusion
Literatur

Model-based recursive partitioning (MOB)



Trees and Rasch-Trees
Carolin Strobl, Julia Kopf, Achim Zeileis
CART
MOB
Testing for DIF
Graphical test
Latent classes
Model-based partitioning
Example
Conclusion
Literatur

Model-based recursive partitioning (MOB)

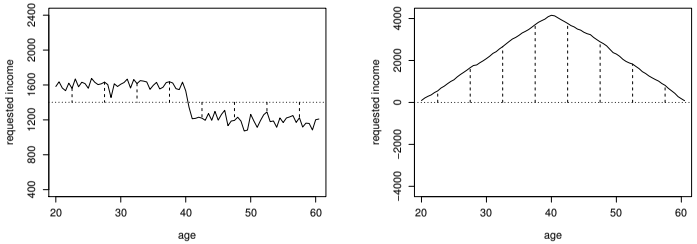
identify groups of people with different parameters by means of tests for parameter instability:

- individual contributions to the score-funktion

$$\psi(y_i, \theta) = \frac{\partial \Psi(y_i, \theta)}{\partial \theta}$$

- cumulated over all values of covariate ℓ

$$W_\ell(t) = \hat{V}^{-1/2} n^{-1/2} \sum_{i=1}^{\lfloor n \cdot t \rfloor} \psi(y_{(i|\ell)}, \hat{\theta})$$



- under H_0 the path fluctuates randomly around zero ($\rightarrow$ Brownian bridge; Zeileis & Hornik, 2007)

Model-based recursive partitioning (MOB)

optimally selected Lagrange-Multiplier test-statistics

- for continuous covariates

$$S_\ell = \max_{i=\underline{i},\ldots,\overline{i}} \left(\frac{i}{n} \cdot \frac{n-i}{n}\right)^{-1} \left\| W_\ell\left(\frac{i}{n}\right) \right\|_2^2$$

- and similarly for categorical covariates

with known distributions

$\Rightarrow$ p-values = split selection criteria

function `mob` in package `party` (Hothorn et al., 2011)

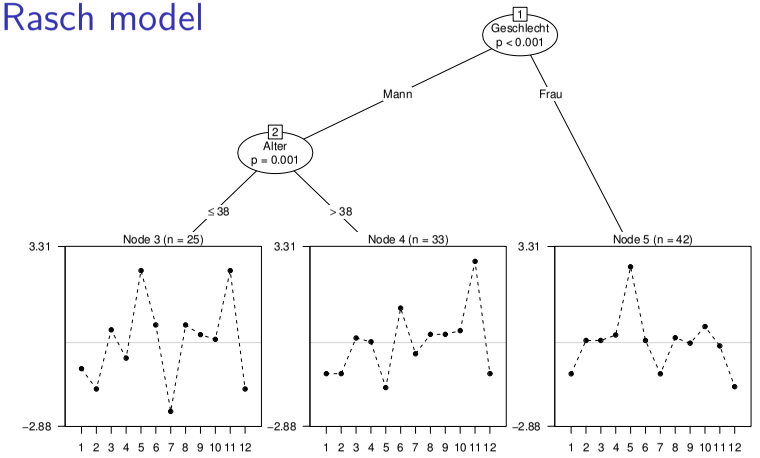
MOB versus CART

Berk (Berk, 2006, p. 236) about algorithmic methods such as classification and regression trees:

> *"With algorithmic methods, there is no statistical model in the usual sense; no effort has been made to represent how the data were generated. And no apologies are offered for the absence of a model. There is a practical data analysis problem to solve that is attacked directly with procedures designed specifically for that purpose."*

Model-based recursive partitioning combines theory-based and data-driven approach

New: Model-based recursive partitioning for the Rasch model



Strobl, Wickelmaier & Zeileis (2010); Strobl, Kopf & Zeileis (2010)

R-package `psychotree` (Zeileis, Strobl & Wickelmaier)

Differential Item Functioning (DIF)

is present when one or more items of a test

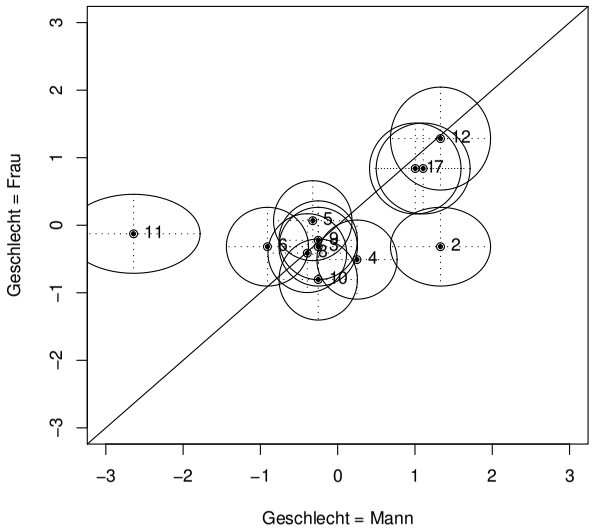
- are easier or harder to solve for certain subjects
- even though they have the same latent trait

Standard model tests

- tests for k given groups
  graphical test, Andersen's Likelihood-Ratio Test, Wald Tests
  - \+ straightforward interpretation
  - – only detect DIF in specified groups
- latent-class approach
  Rost's "Mixed" Rasch model
  - \+ identifies previously unknown groups with DIF
  - – groups are not directly interpretable
  - ⇒ 2nd step: describe groups with covariates

Trees and Rasch-Trees
Carolin Strobl, Julia Kopf, Achim Zeileis
CART
MOB
Testing for DIF
Graphical test
Latent classes
Model-based partitioning
Example
Conclusion
Literatur

Standard model tests



R-package eRm (Mair, Hatzinger & Maier, 2010)

Trees and Rasch-Trees
Carolin Strobl, Julia Kopf, Achim Zeileis
CART
MOB
Testing for DIF
Graphical test
Latent classes
Model-based partitioning
Example
Conclusion
Literatur

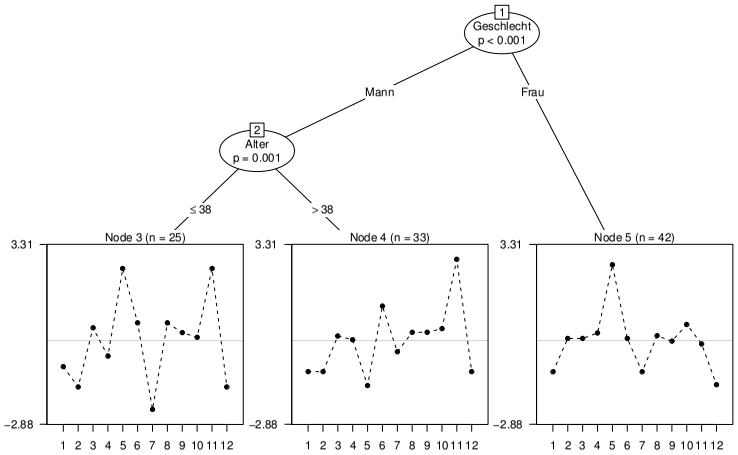
Standard model tests

- tests for k given groups
  graphical test, Andersen's Likelihood-Ratio Test, Wald Tests
  - \+ straightforward interpretation
  - – only detect DIF in specified groups
- latent-class approach
  Rost's "Mixed" Rasch model
  - \+ identifies previously unknown groups with DIF
  - – groups are not directly interpretable
  - ⇒ 2nd step: describe groups with covariates

Trees and Rasch-Trees
Carolin Strobl, Julia Kopf, Achim Zeileis
CART
MOB
Testing for DIF
Graphical test
Latent classes
Model-based partitioning
Example
Conclusion
Literatur

New: Model-based recursive partitioning

+ identifies previously unknown groups with DIF
+ straightforward interpretation



Example: "Students-PISA"

data from the SPIEGEL "Students-PISA" survey

- open-access online survey on general education
- each partcipant was randomly assigned one of 24 questionnaires, consisting of 45 items from 6 topics: politics, history, economics, culture and natural sciences
- questions were either multiple-choice or open
- recorded response: correct/wrong

results presented here are for one exemplary questionnaire, N = 30 188

Example: "Students-PISA"

exemplary questions:

- history: "In which century did the Thirty Years' War take place?" 17th
- economics: "Which internet-company took over the media-group Time Warner?" AOL
- culture: "Which city is the setting for the novel 'Buddenbrooks'?" Lübeck

Example: "Students-PISA"

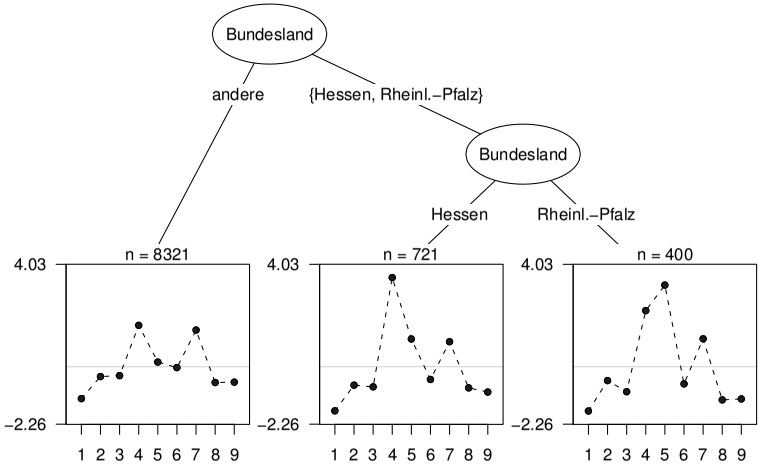
curious finding:

those participants who received their Abitur in Rheinland-Pfalz perform significantly better in the test

possible explanations:

- they are just smarter
- they have an unfair advantage ⇒ DIF

Example: "Students-PISA"



Nr. 4: Where is Hessen? (indicate location on a map)

Nr. 5: What is the capital of Rheinland-Pfalz? (Mainz)

psychological impact of DIF

- test is no longer specifically objective
- fair comparisons between the groups are impossible

⇒ eliminate DIF-items from the test

in our example:
eliminating items 4 and 5 eliminates group differences, i.e.,

the supposed group difference was only an artefact of the test construction!

Trees and Rasch-Trees
Carolin Strobl, Julia Kopf, Achim Zeileis
CART
MOB
Testing for DIF
Graphical test
Latent classes
Model-based partitioning
Example
Conclusion
Literatur

Conclusion

- DIF can lead to severe misinterpretations
- statistical tools for DIF-diagnosis
  - approaches for given groups
    - graphical test
    - Andersen's Likelihood-Ratio Test
  - Rost's latent-class approach
  - model-based recursive partitioning („Rasch-trees")
- interpretability of diagnostic tools can help understand the psychological sources of DIF
  but: observed covariates may be proxies for true causes

Trees and Rasch-Trees
Carolin Strobl, Julia Kopf, Achim Zeileis
CART
MOB
Testing for DIF
Graphical test
Latent classes
Model-based partitioning
Example
Conclusion
Literatur

References I

Berk, R. A. (2006). An introduction to ensemble methods for data analysis. *Sociological Methods & Research 34*(3), 263–295.

Dobra, A. und Gehrke, J. (2001). Bias correction in classification tree construction. In C. E. Brodley & A. P. Danyluk (Hrsg.), *Proceedings of the Seventeenth International Conference on Machine Learning (ICML 2001), Williams College, Williamstown, MA, USA*, S. 90–97. Morgan Kaufmann.

Hothorn, T., Hornik, K., Strobl, C., und Zeileis, A. (2011). `party`: A laboratory for recursive part(y)itioning. R package version 0.9-995.

Trees and Rasch-Trees
Carolin Strobl, Julia Kopf, Achim Zeileis
CART
MOB
Testing for DIF
Graphical test
Latent classes
Model-based partitioning
Example
Conclusion
Literatur

References II

Jensen, D. D. und Cohen, P. R. (2000). Multiple comparisons in induction algorithms. *Machine Learning 38*(3), 309–338.

Kim, H. und Loh, W. (2001). Classification trees with unbiased multiway splits. *Journal of the American Statistical Association* **96**, 589–604.

Loh, W. und Shih, Y. (1997). Split selection methods for classification trees. *Statistica Sinica* **7**, 815–840.

Mair, P. und Hatzinger, R. (2007). Extended Rasch modeling: The eRm package for the application of IRT models in R. *Journal of Statistical Software 20*(9).

References III

Mair, P., Hatzinger, R., und Maier, M. (2010). eRm*: Extended Rasch Modeling*. R package version 0.13-0.

Ripley, B. (2007). tree*: Classification and Regression Trees*. R package version 1.0-26.

Strobl, C. (2010). *Das Rasch-Modell – Eine verständliche Einführung für Studium und Praxis*. München, Mering: Rainer Hampp Verlag.

Strobl, C., Boulesteix, A.-L., und Augustin, T. (2007). Unbiased split selection for classification trees based on the Gini Index. *Computational Statistics & Data Analysis 52*(1), 483–501.

Trees and Rasch-Trees

Carolin Strobl, Julia Kopf, Achim Zeileis

CART

MOB

Testing for DIF

Graphical test

Latent classes

Model-based partitioning

Example

Conclusion

Literatur

References IV

Strobl, C., Kopf, J., und Zeileis, A. (2010a). A new method for detecting differential item functioning in the Rasch model. Technical Report 92, Department of Statistics, Ludwig-Maximilians-Universität München, Germany.

Strobl, C., Kopf, J., und Zeileis, A. (2010b). Wissen Frauen weniger oder nur das Falsche? – Ein statistisches Modell für unterschiedliche Aufgaben-Schwierigkeiten in Teilstichproben. In S. Trepte & M. Verbeet (Hrsg.), *Allgemeinbildung in Deutschland – Erkenntnisse aus dem SPIEGEL Studentenpisa-Test*, Wiesbaden, S. 255–272. VS Verlag.

Trees and Rasch-Trees

Carolin Strobl, Julia Kopf, Achim Zeileis

CART

MOB

Testing for DIF

Graphical test

Latent classes

Model-based partitioning

Example

Conclusion

Literatur

References V

Strobl, C., Malley, J., und Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application and characteristics of classification and regression trees, bagging and random forests. *Psychological Methods 14*(4), 323–348.

Therneau, T. und Atkinson, B. (2006). rpart: Recursive partitioning. R package version 3.1-30.

White, A. und Liu, W. (1994). Bias in information based measures in decision tree induction. *Machine Learning 15*(3), 321–329.

Zeileis, A. und Hornik, K. (2007). Generalized M-fluctuation tests for parameter instability. *Statistica Neerlandica 61*(4), 488–508.

Trees and Rasch-Trees

Carolin Strobl, Julia Kopf, Achim Zeileis

CART

MOB

Testing for DIF

Graphical test

Latent classes

Model-based partitioning

Example

Conclusion

Literatur