Stability Assessment for Trees and other Supervised Statistical Learning Results

Michel Philipp, Carolin Strobl, Thomas Rusch,
Kurt Hornik, Achim Zeileis

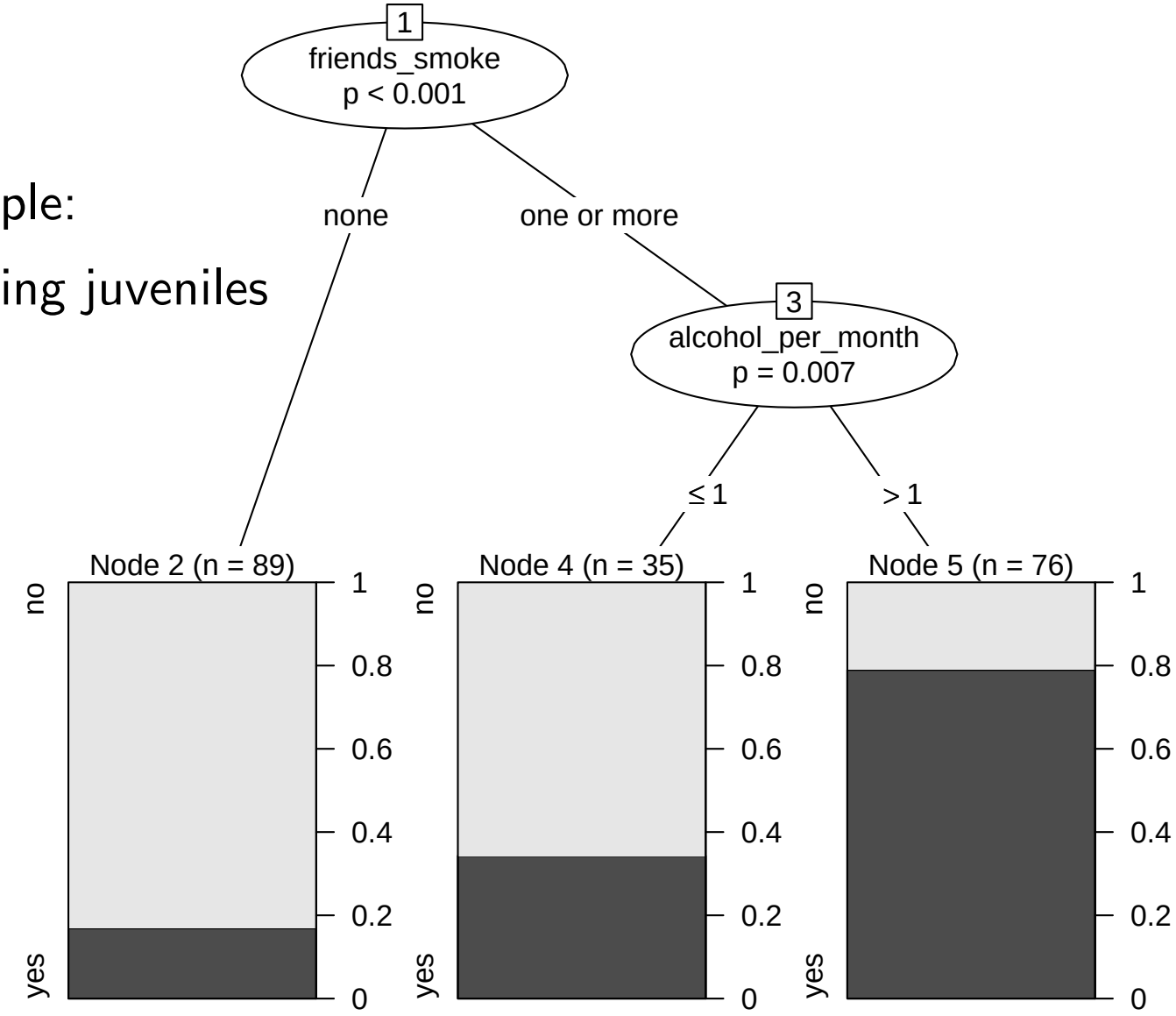
University of Zurich[UZH]

Stability Assessment
Trees and Co
CART
MOB
Random Forests
Stability of trees
Tree structure
stablelearner
Variable selection
Cutpoint selection
Stability for supervised learning in general
Summary
References

Classification and regression trees (CART)

Example:

Smoking juveniles



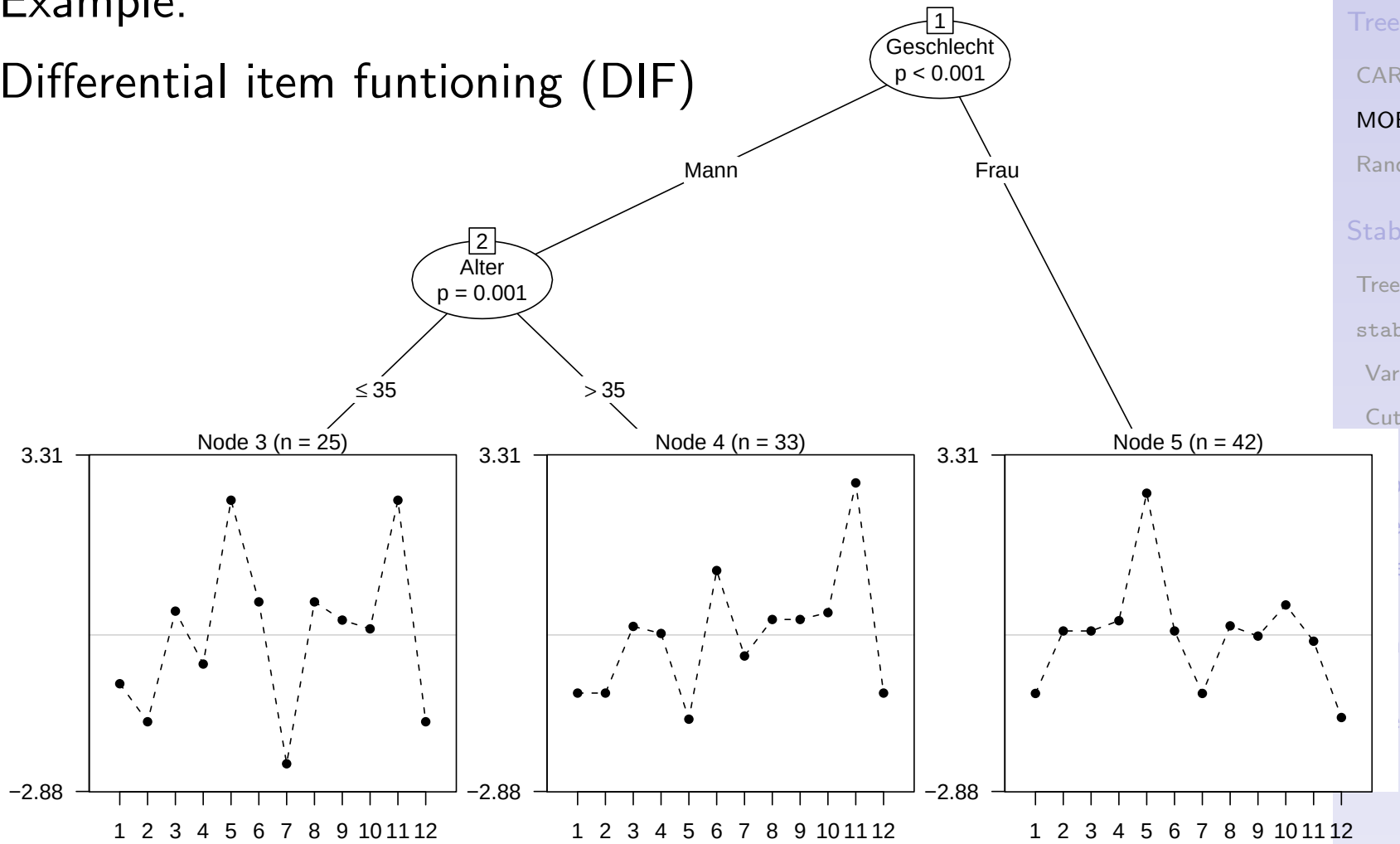
R-package `party` (Hothorn, Hornik, Strobl and Zeileis)

Model based recursive partitioning (MOB)

Trees and Co
CART
MOB
Random Forests
Stability of trees
Tree structure
stablelearner
Variable selection
Cutpoint selection
ility for
rvised learning
eneral
mary
rences

Example:

Differential item funtioning (DIF)



R-package `psychotree` (Zeileis, Strobl, Wickelmaier, Komboz, Kopf)

Trees are...

+ easy to interpret

– instable

Trees are...

+ easy to interpret (really?)

– instable

Trees and Co
CART
MOB
Random Forests
Stability of trees
Tree structure
stablelearner
Variable selection
Cutpoint selection
Stability for supervised learning in general
Summary
References

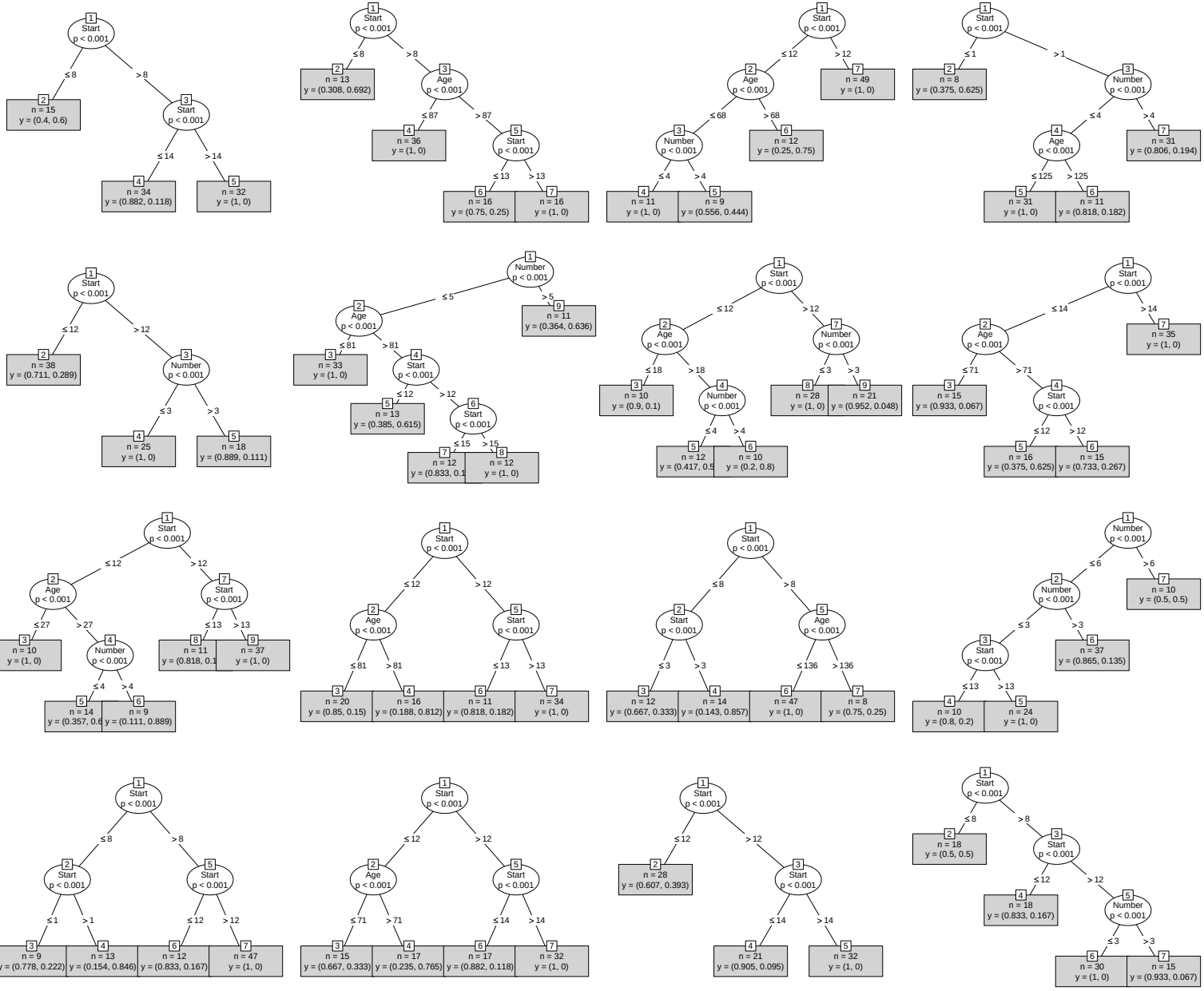
Trees are...

\+ easy to interpret (really?)

– instable (and no means of inference for complete tree)

Trees and Co
CART
MOB
Random Forests
Stability of trees
Tree structure
stablelearner
Variable selection
Cutpoint selection
Stability for supervised learning in general
Summary
References

Random Forests



Forests are...

\+ stable (for predictions)

– no longer interpretable

Forests are...

\+ stable (for predictions)

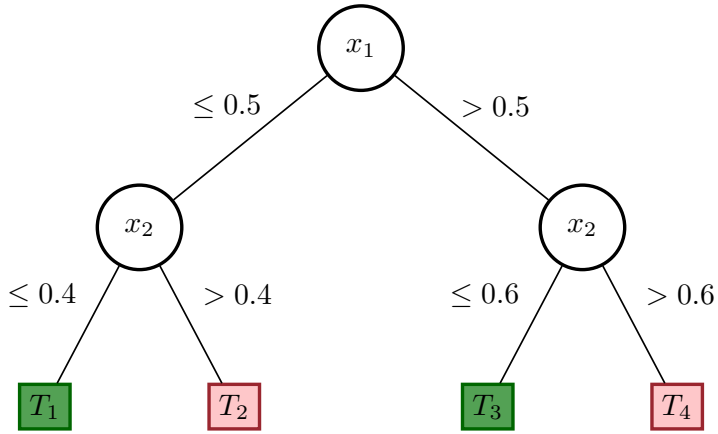
– no longer interpretable

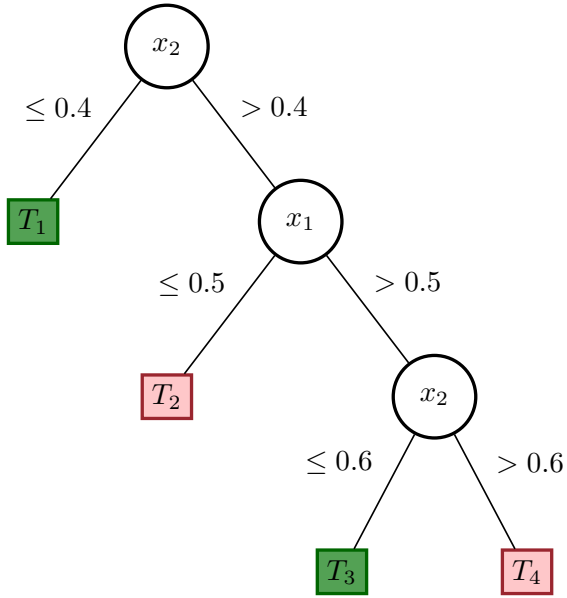
$\Rightarrow$ when is a single tree safe to interpret?

Stability assessment for trees

- resample from the learning data (bootstrap/subsampling/sample splitting/jackknife)
- fit trees on all samples $\Rightarrow$ can look quite different

Tree structure





Stability Assessment

Trees and Co

CART

MOB

Random Forests

Stability of trees

Tree structure

stablelearner

Variable selection

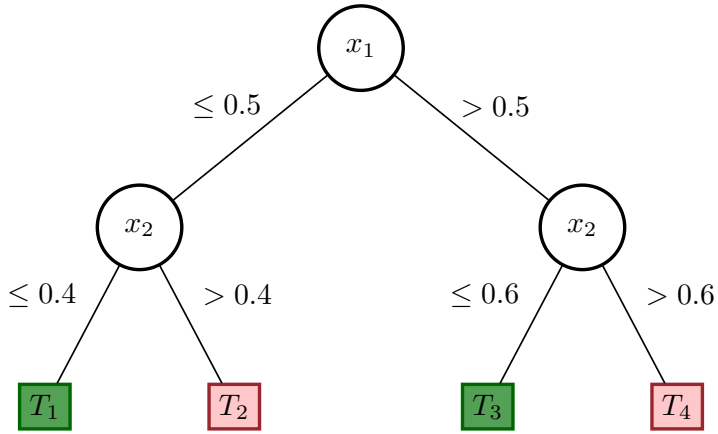
Cutpoint selection

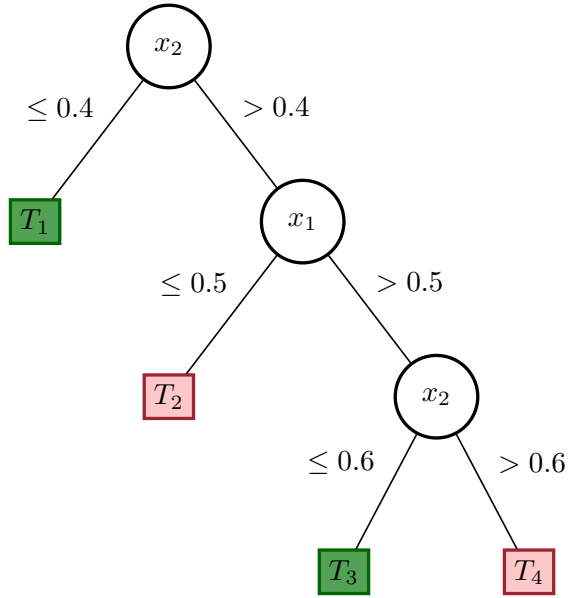
Stability for supervised learning in general

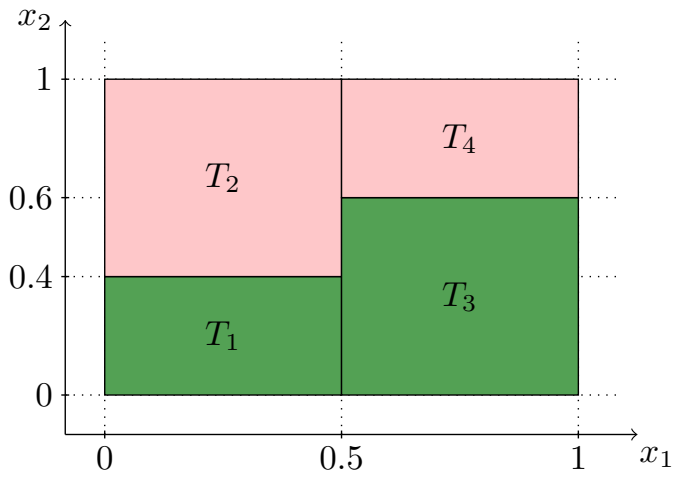
Summary

References

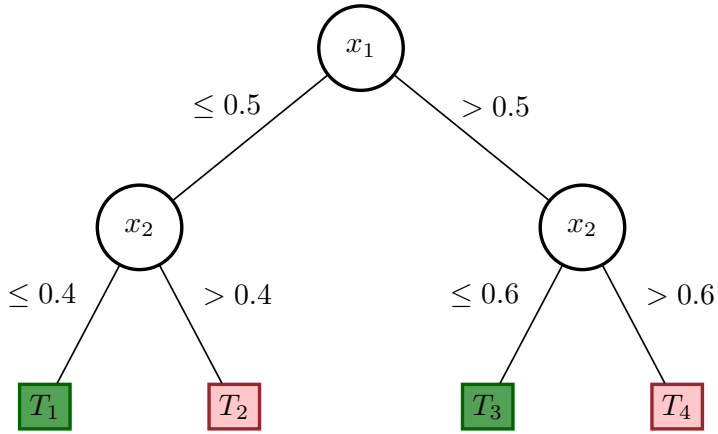
Tree structure

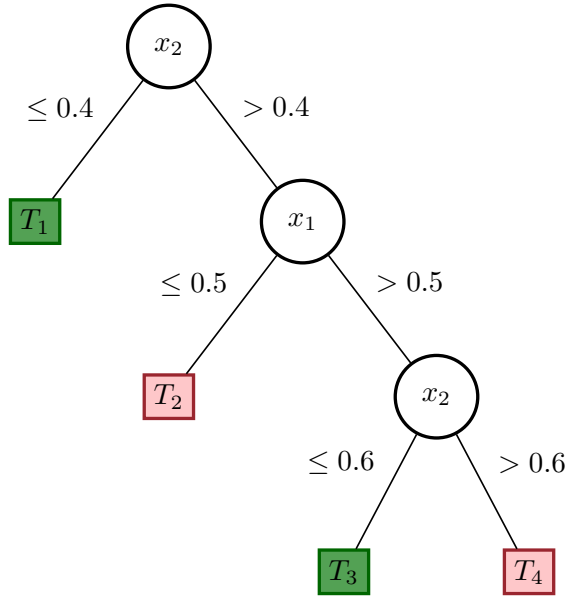


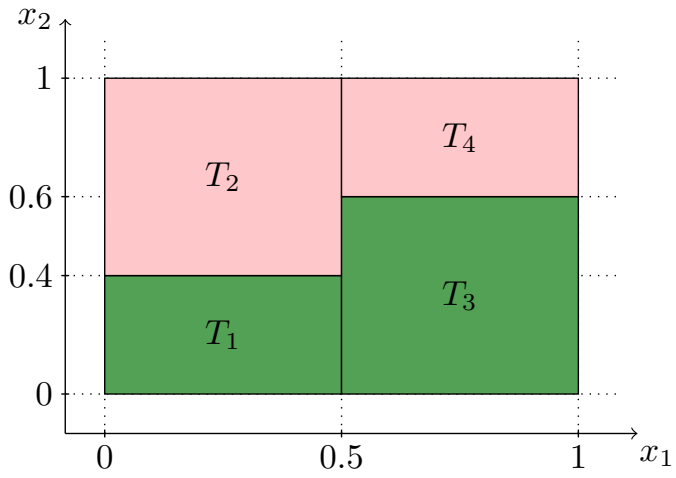


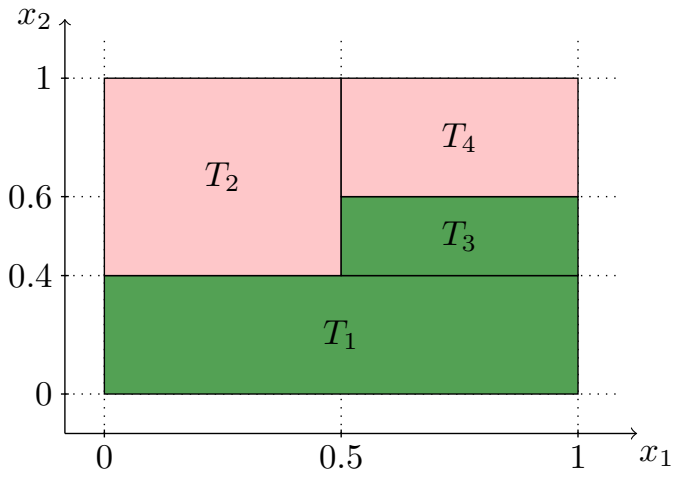


Tree structure

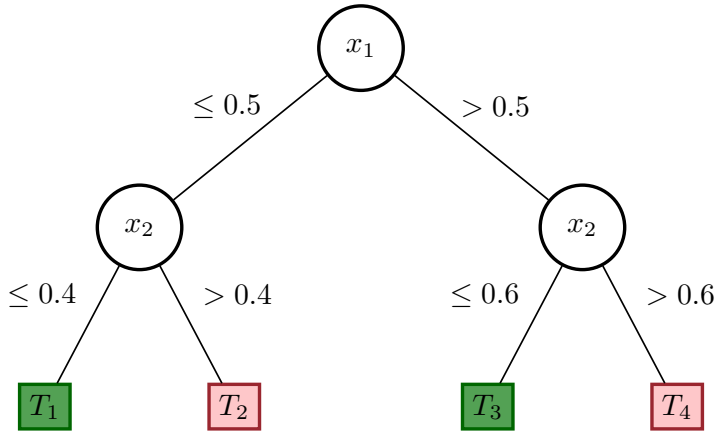


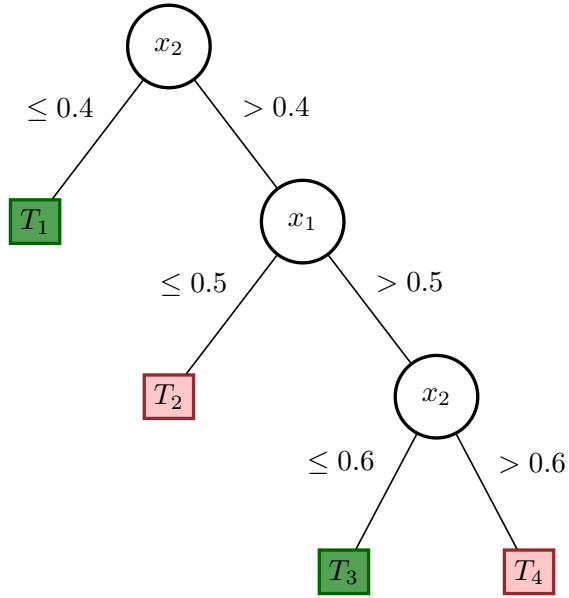


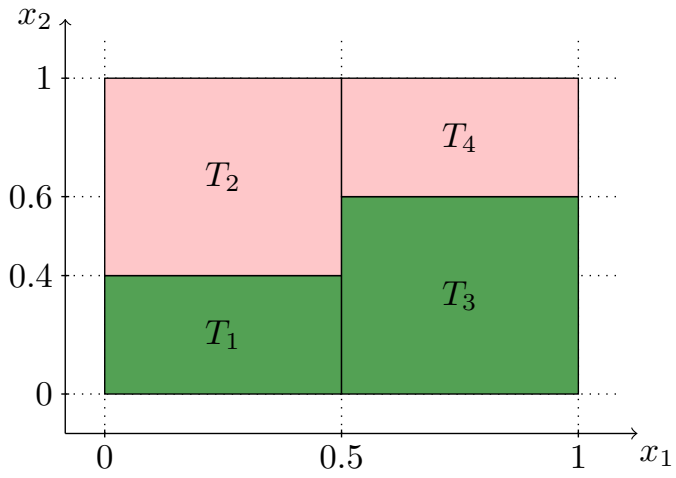


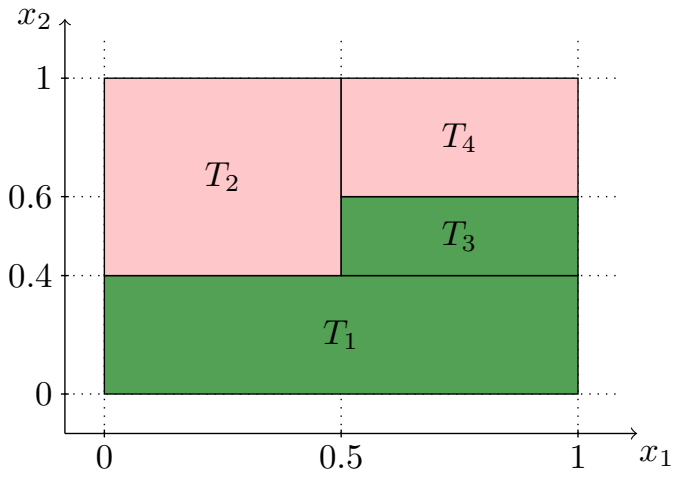


Tree structure – totally overrated!









Stability Assessment

Trees and Co

CART

MOB

Random Forests

Stability of trees

Tree structure

stablelearner

Variable selection

Cutpoint selection

Stability for supervised learning in general

Summary

References

Stability assessment for trees

- resample from the learning data (bootstrap/subsampling/sample splitting/jackknife)
- fit trees on all samples $\Rightarrow$ can look quite different

Trees and Co
CART
MOB
Random Forests
Stability of trees
Tree structure
stablelearner
Variable selection
Cutpoint selection
Stability for supervised learning in general
Summary
References

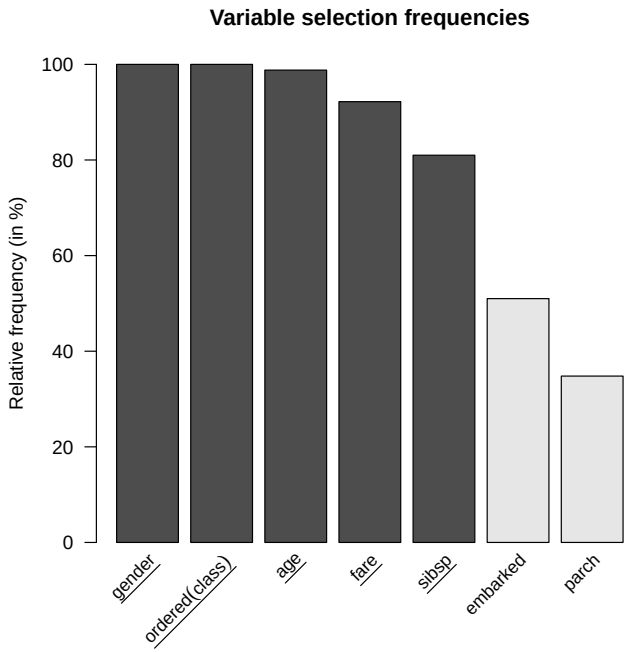
Stability assessment for trees

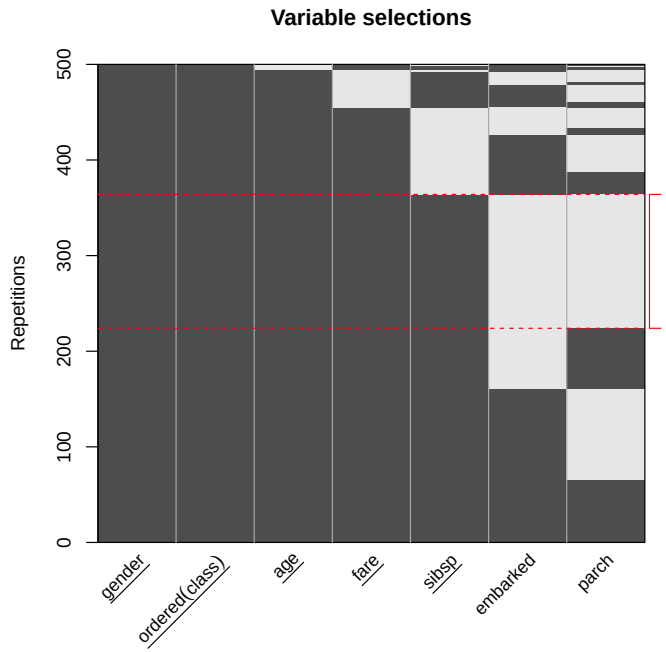
- resample from the learning data (bootstrap/subsampling/sample splitting/jackknife)
- fit trees on all samples $\Rightarrow$ can look quite different
- create summary statistics and plots for
  - variable selection
  - cutpoint selection

Trees and Co
CART
MOB
Random Forests
Stability of trees
Tree structure
stablelearner
Variable selection
Cutpoint selection
Stability for supervised learning in general
Summary
References

Stability assessment for trees

```
> mytree <- ctree(y ~ x1 + x2 + ..., data = ...)
> stab_mytree <- stabletree(mytree, B = 500)
> barplot(stab_mytree)    > image(stab_mytree)
```

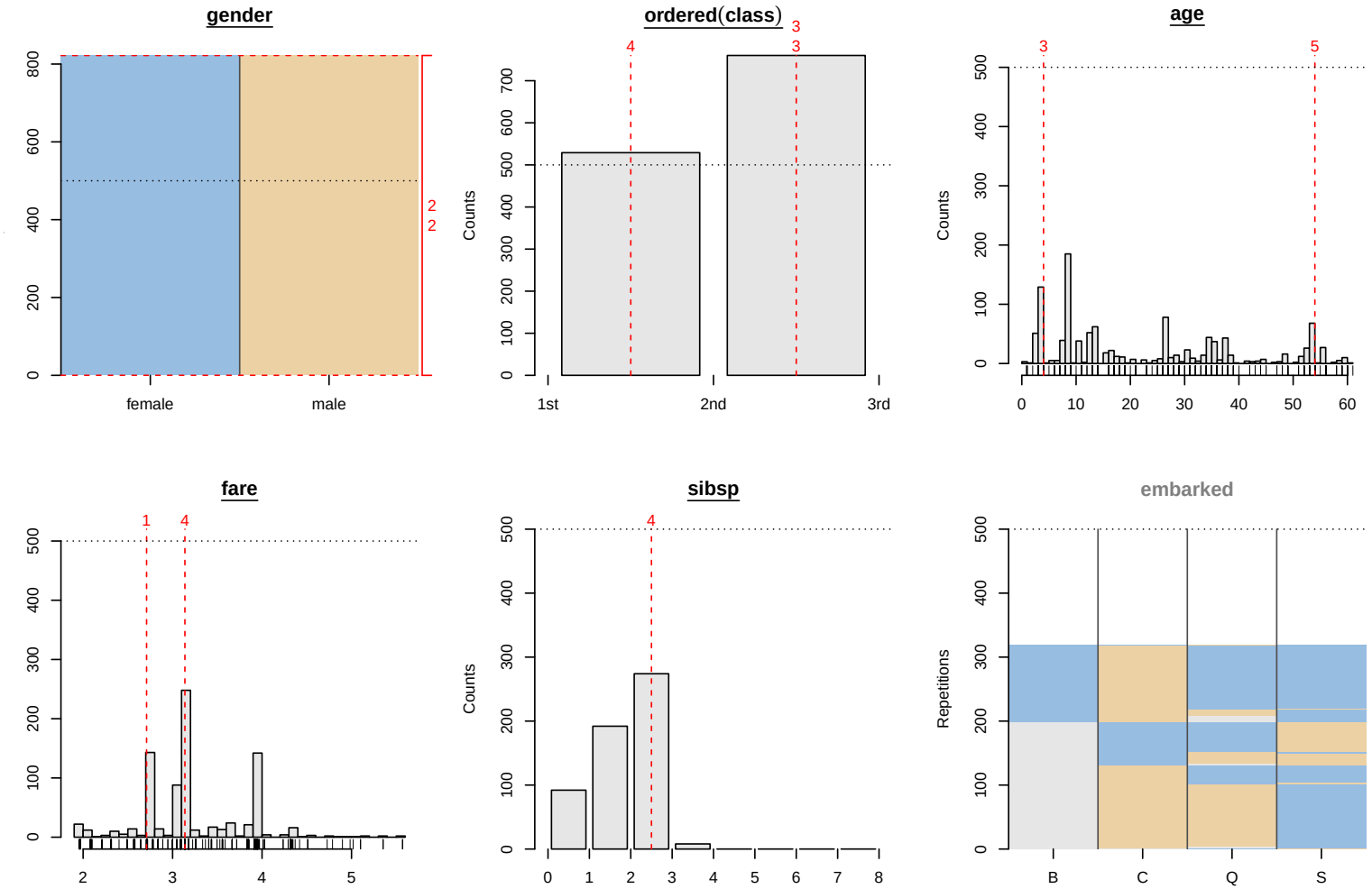




R-package `stablelearner` (Philipp, Strobl, Zeileis, Rusch, Hornik)

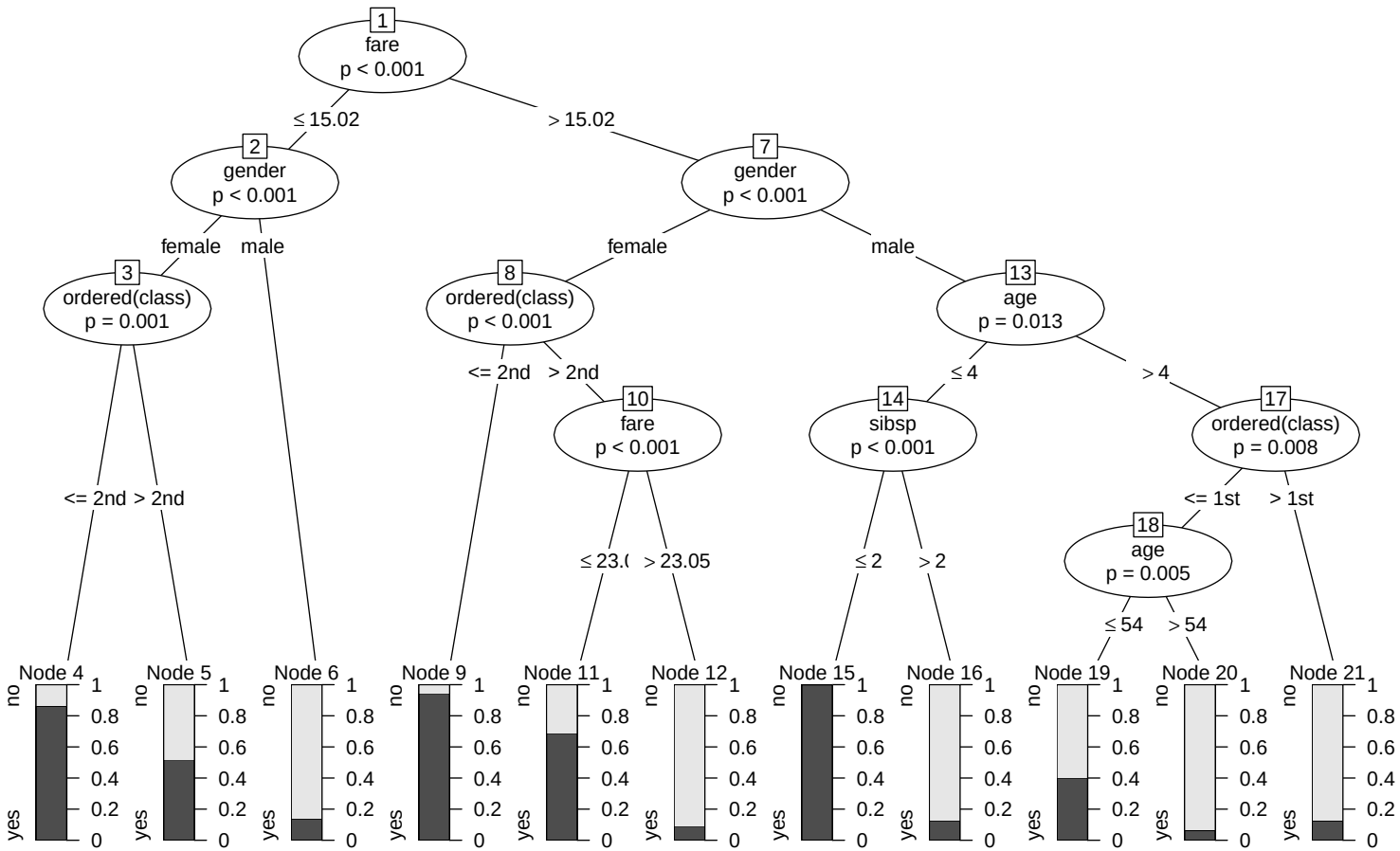
Stability assessment for trees

```
> plot(stab_mytree)
```



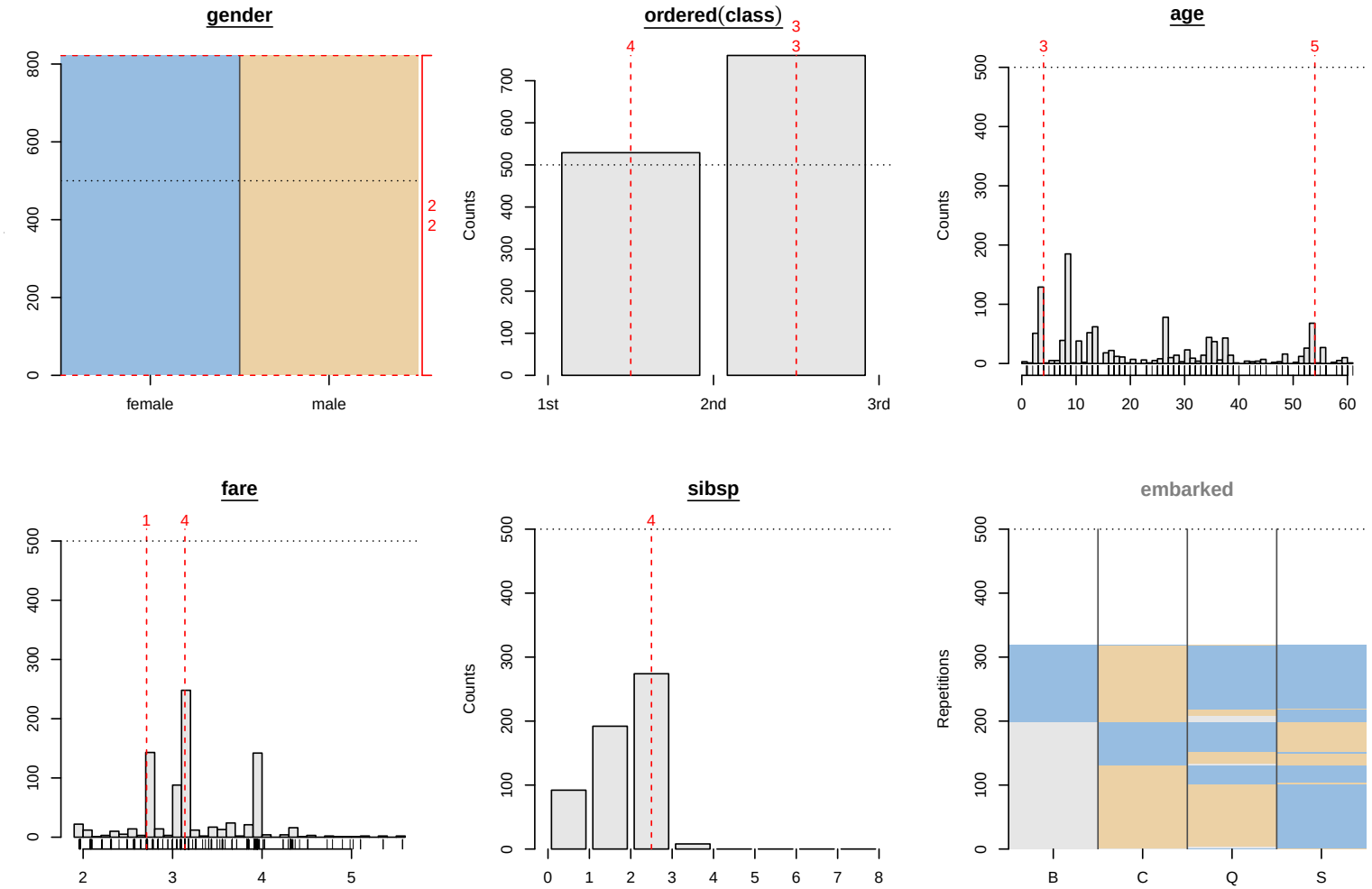
Stability assessment for trees

```
> plot(mytree)
```

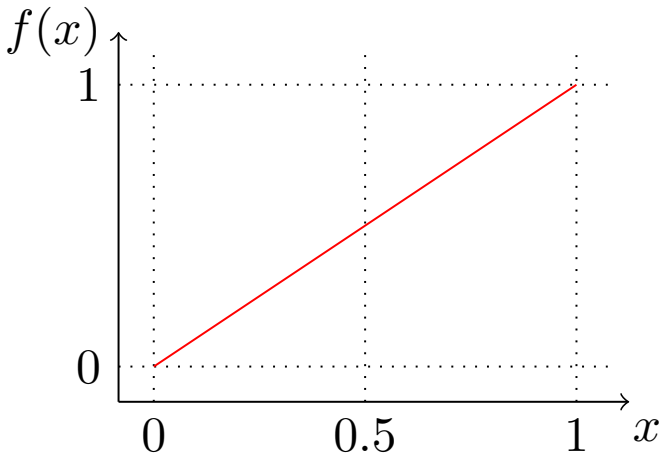


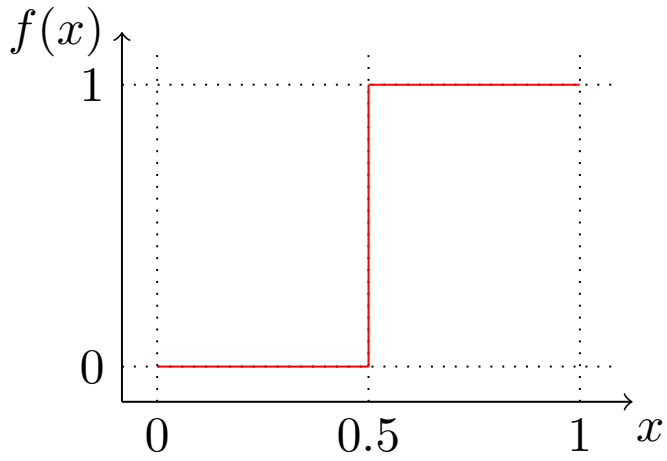
Stability assessment for trees

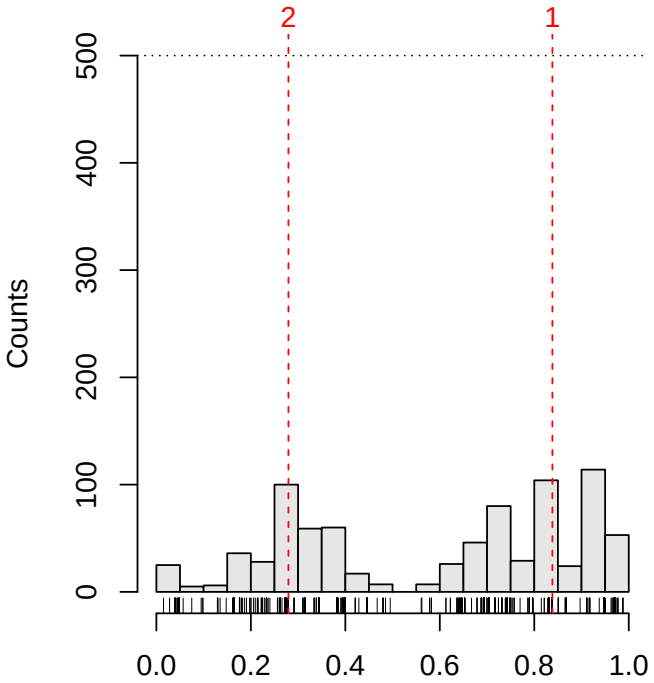
```
> plot(stab_mytree)
```

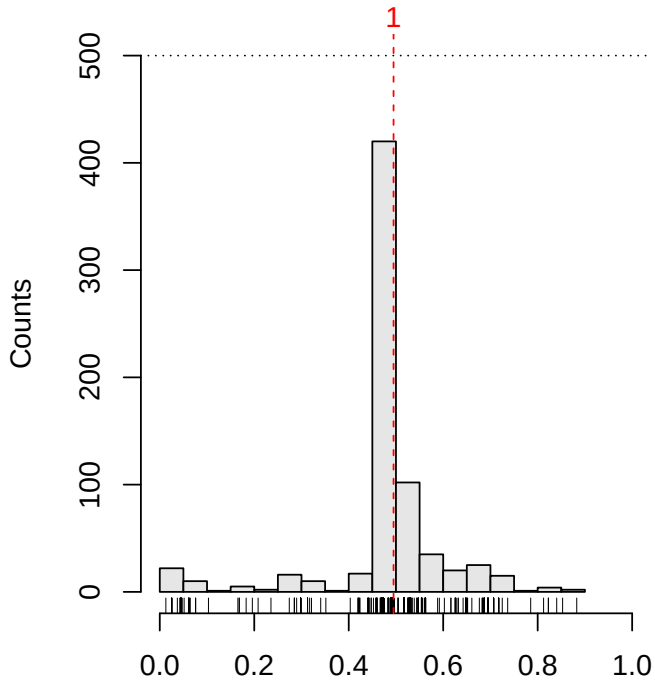


Stability assessment for trees







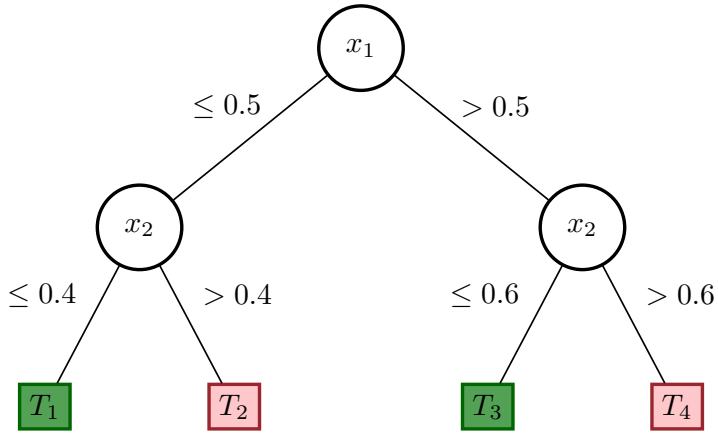


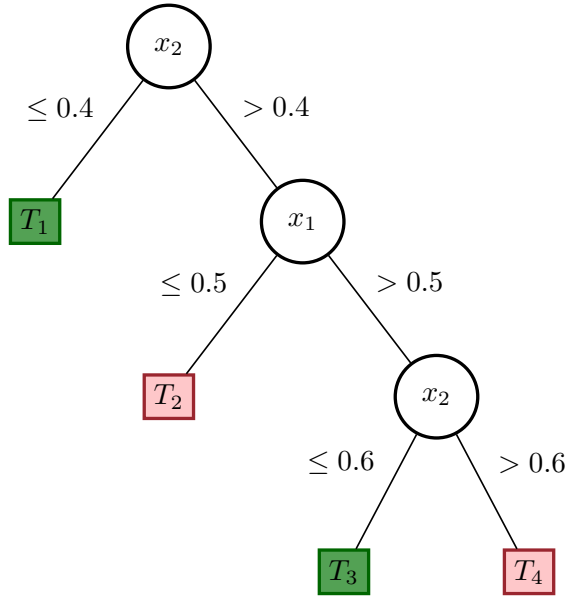
Stability assessment in general

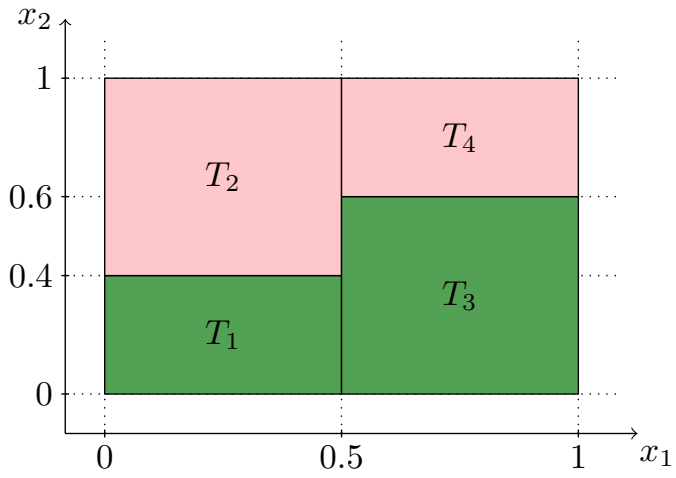
- structural vs. semantic stability

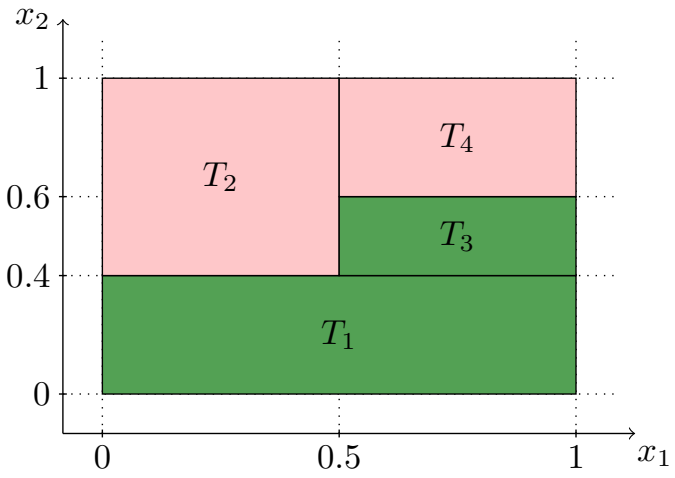
Trees and Co
CART
MOB
Random Forests
Stability of trees
Tree structure
stablelearner
Variable selection
Cutpoint selection
Stability for supervised learning in general
Summary
References

Tree structure – totally overrated!









Stability Assessment

Trees and Co

CART

MOB

Random Forests

Stability of trees

Tree structure

stablelearner

Variable selection

Cutpoint selection

Stability for supervised learning in general

Summary

References

Stability assessment in general

- structural vs. semantic stability

Trees and Co
CART
MOB
Random Forests
Stability of trees
Tree structure
stablelearner
Variable selection
Cutpoint selection
Stability for supervised learning in general
Summary
References

Stability assessment in general

- structural vs. semantic stability $\Rightarrow$ compare predictions

Trees and Co
CART
MOB
Random Forests
Stability of trees
Tree structure
stablelearner
Variable selection
Cutpoint selection
Stability for supervised learning in general
Summary
References

Stability assessment in general

- structural vs. semantic stability $\Rightarrow$ compare predictions

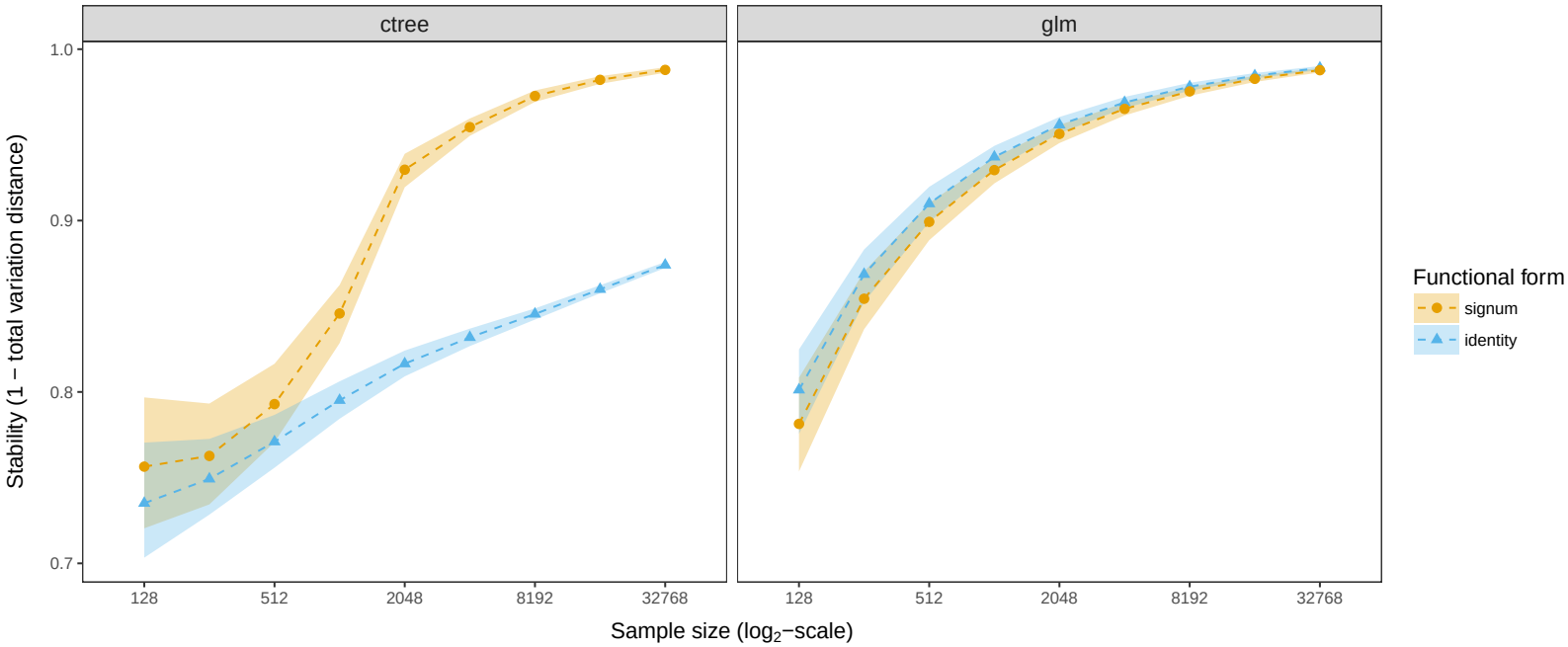
Trees and Co
CART
MOB
Random Forests
Stability of trees
Tree structure
stablelearner
Variable selection
Cutpoint selection
Stability for supervised learning in general
Summary
References

Stability assessment in general

- structural vs. semantic stability $\Rightarrow$ compare predictions
- stability depends on
  - algorithm, model
  - data generating process

Trees and Co
CART
MOB
Random Forests
Stability of trees
Tree structure
stablelearner
Variable selection
Cutpoint selection
Stability for supervised learning in general
Summary
References

Match between algorithm and dgp



Stability assessment in general

- structural vs. semantic stability $\Rightarrow$ compare predictions
- stability depends on
  - algorithm, model
  - data generating process

Trees and Co
CART
MOB
Random Forests
Stability of trees
Tree structure
stablelearner
Variable selection
Cutpoint selection
Stability for supervised learning in general
Summary
References

Stability assessment in general

- structural vs. semantic stability $\Rightarrow$ compare predictions
- stability depends on
  - algorithm, model
  - data generating process
  - sample size
  - resampling strategy, distance measure
- works for real and simulation settings

Example: Benchmark study

stability and prediction accuracy of `ctree`

Data set	n	p	K	Stability		Accuracy		CPU time [sec]
				Median	IQR	Median	IQR	
Iris	150	4	3	0.956	0.048	0.919	0.146	4.7
Breast Cancer	699	9	2	0.933	0.029	0.864	0.080	13.4
Titanic	1317	7	2	0.925	0.027	0.561	0.086	13.6
Ionosphere	351	34	2	0.900	0.058	0.788	0.132	10.8
Pima	768	8	2	0.835	0.036	0.403	0.130	13.8
Satellite	6435	36	6	0.819	0.013	0.363	0.025	631.3
Sonar	208	60	2	0.728	0.098	0.412	0.231	12.6
Vehicle	846	18	4	0.723	0.050	0.552	0.084	50.3
Glass	214	9	6	0.694	0.107	0.219	0.185	12.9

Summary

framework for assessing the stability of trees

- descriptive statistics and plots for variable and cutpoint selection
- works for all kinds of trees and MOBs (incl. `ctree`, `(g)lmtree`, `raschtree` etc. via `partykit`)

framework for assessing the stability of supervised learning

- one number/distribution for measuring stability
- works for classification and regression problems
- future work: extend to MOB

available in R-package `stablelearner` (Philipp, Strobl, Zeileis, Rusch, Hornik)

Trees and Co
CART
MOB
Random Forests
Stability of trees
Tree structure
stablelearner
Variable selection
Cutpoint selection
Stability for supervised learning in general
Summary
References

References

Philipp M, Zeileis A, Strobl C (2016). A toolkit for stability assessment of tree-based learners. In Colubi et al. (Eds.) *Proceedings of COMPSTAT 2016 – 22nd International Conference on Computational Statistics*, 315–325.

Philipp M, Rusch T, Hornik K, Strobl C (2018). Measuring the stability of results from statistical learning. *Journal of Computational and Graphical Statistics* (forthcoming).

Strobl C, Malley J, 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.

Stability assessment for trees

```
> summary(stab_mytree)

                    freq *  mean *
gender             1.000 1 1.644 2
ordered(class)     1.000 1 2.578 3
age                0.988 1 2.316 2
fare               0.922 1 1.676 2
sibsp              0.810 1 1.132 1
embarked           0.510 0 0.638 0
parch              0.348 0 0.444 0
(* = original tree)
```

freq: variable selection frequency
mean: average number of appearances per tree

Trees and Co
CART
MOB
Random Forests
Stability of trees
Tree structure
stablelearner
Variable selection
Cutpoint selection
Stability for supervised learning in general
Summary
References

Stability assessment in general

for $b = 1, \ldots B$

(1) Generate two learning samples $\mathfrak{L}_b'$ and $\mathfrak{L}_b''$ plus an evaluation sample $\mathfrak{E}_b$ by sampling from $\mathcal{F}$ that is a proxy for P_Z. To do this in practice, the plug-in principle can be used. Thus, for the

$$\begin{aligned} \text{real data problem:} \quad \mathcal{F} &\mathrel{\widehat{=}} \widehat{F}_n, \\ \text{simulation study problem:} \quad \mathcal{F} &\mathrel{\widehat{=}} \mathsf{P}_Z, \end{aligned}$$

where $\widehat{F}_n$ is an approximation of P_Z that represents the DGP well when n is large.

(2) Generate the results $r_{\mathcal{A},\mathcal{M}}(x; \mathfrak{L}_b')$ and $r_{\mathcal{A},\mathcal{M}}(x; \mathfrak{L}_b'')$ by training the algorithm on both learning samples.

(3) Compute $s(\hat{y}_b', \hat{y}_b'')$ using the evaluation sample $\mathfrak{E}_b$.

Stability Assessment
Trees and Co
CART
MOB
Random Forests
Stability of trees
Tree structure
stablelearner
Variable selection
Cutpoint selection
Stability for supervised learning in general
Summary
References