

GENERALIZATIONS OF THE RECEIVER OPERATING CHARACTERISTIC CURVE



Universidad de Oviedo
Universidá d'Uviéu
University of Oviedo

Sonia Pérez-Fernández

University of Oviedo (Spain)

Department of Statistics and Operational Research and Mathematic Didactics





KEY MESSAGES & QUESTIONS

Key messages

- Generalization of ROC curve
- Restricted optimization
- Interpretability

Questions

- Optimize computational times:
R package construction
- Literature search and contact
authors



THE RECEIVER OPERATING CHARACTERISTIC (ROC) CURVE

Objective: Diagnostic accuracy of a marker

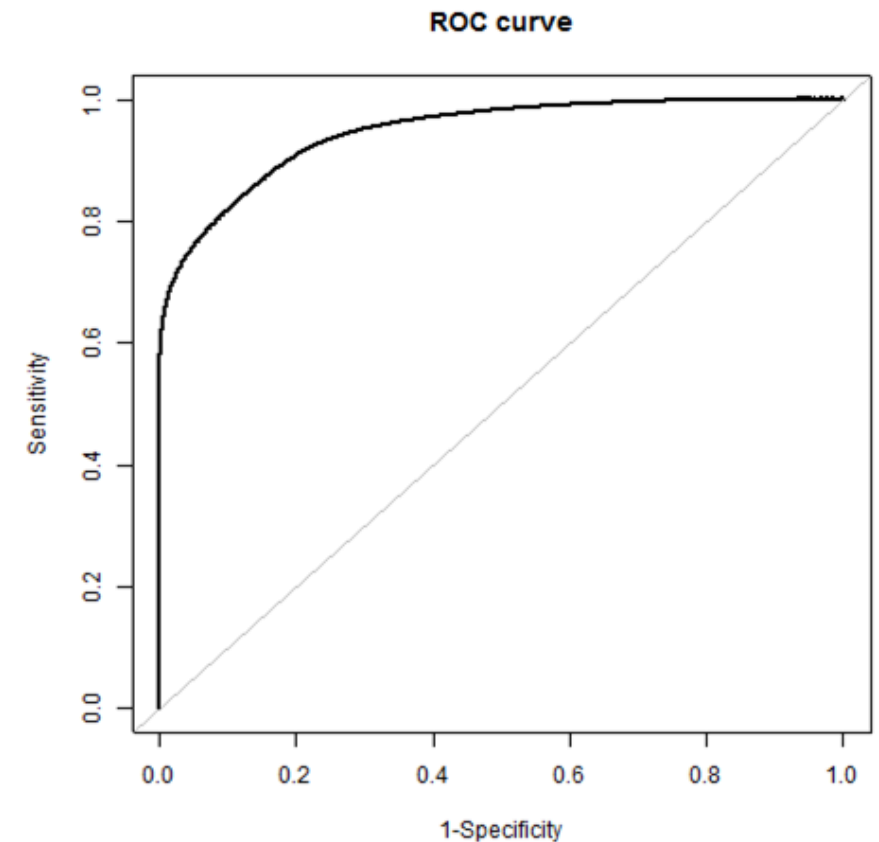
How good is a marker to distinguish between TWO groups in the population?

NOTATION:

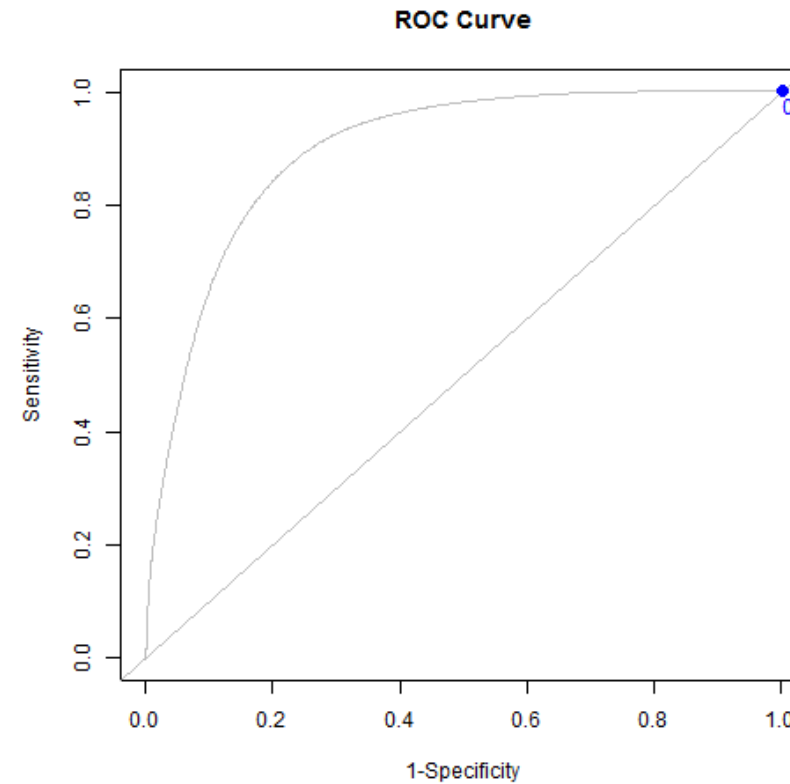
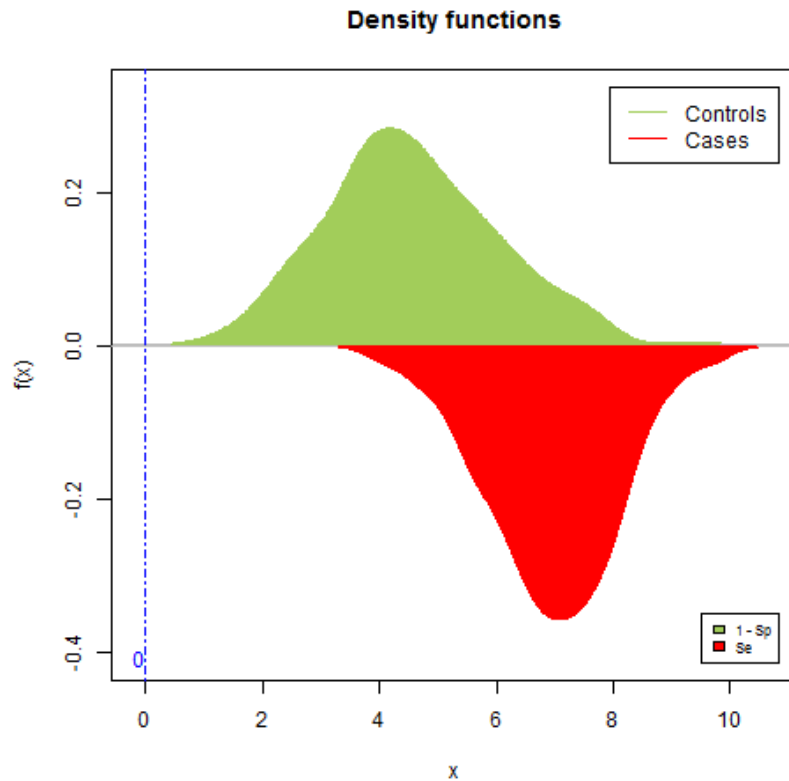
Groups: Controls vs. Cases

Cut-off point / Threshold (marker)

Sensitivity & Specificity



CLASSICAL ROC CURVE – RIGHT SIDED



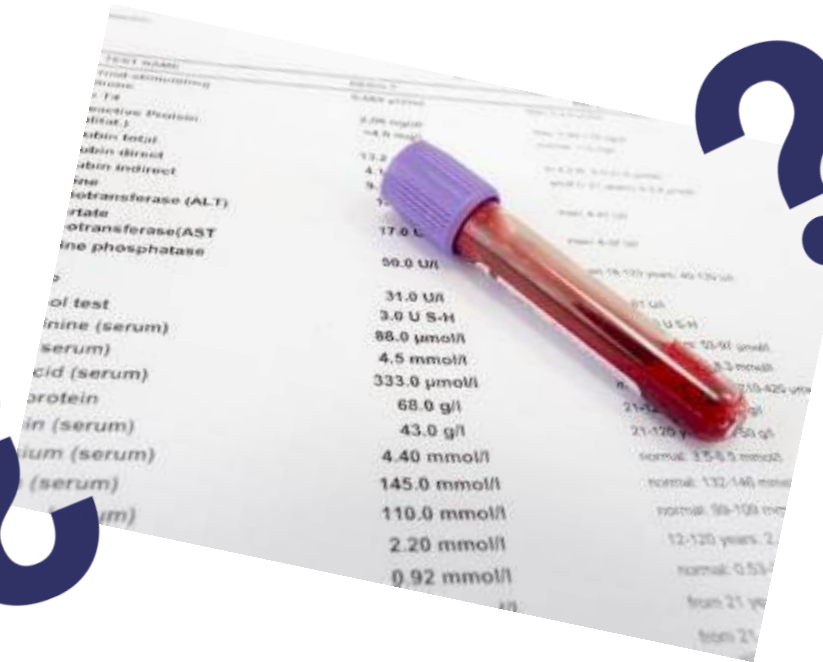
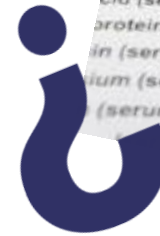
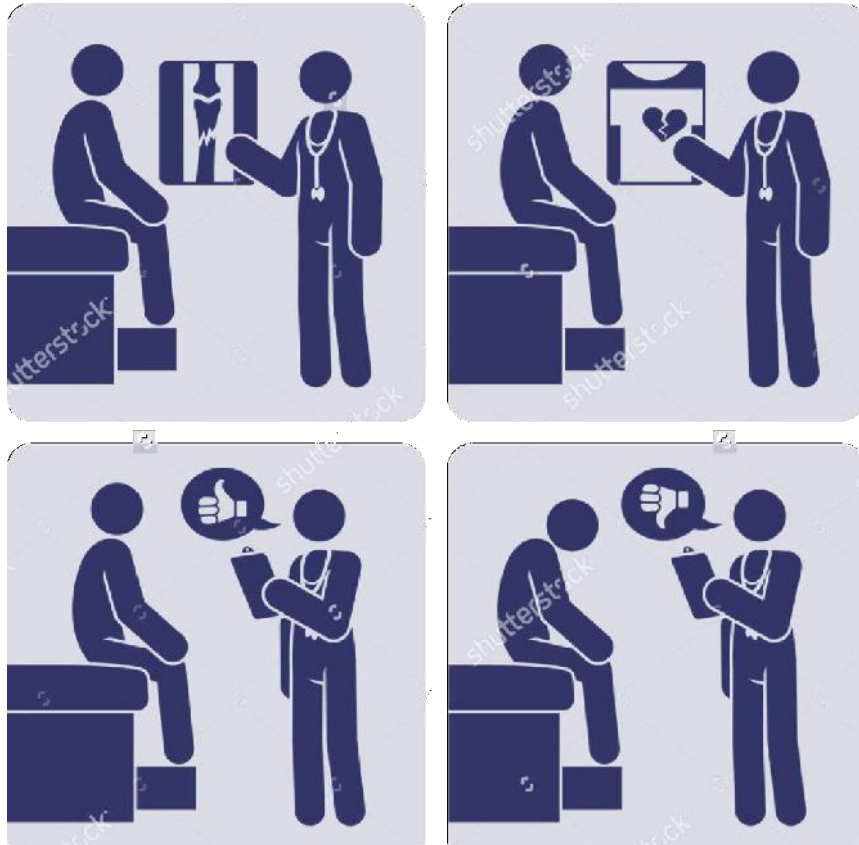
Marker values

Probability of being a case

1 cut-off point



DAILY BINARY CLASSIFICATIONS WHICH LEAD TO AN ROC CURVE GENERALIZATION

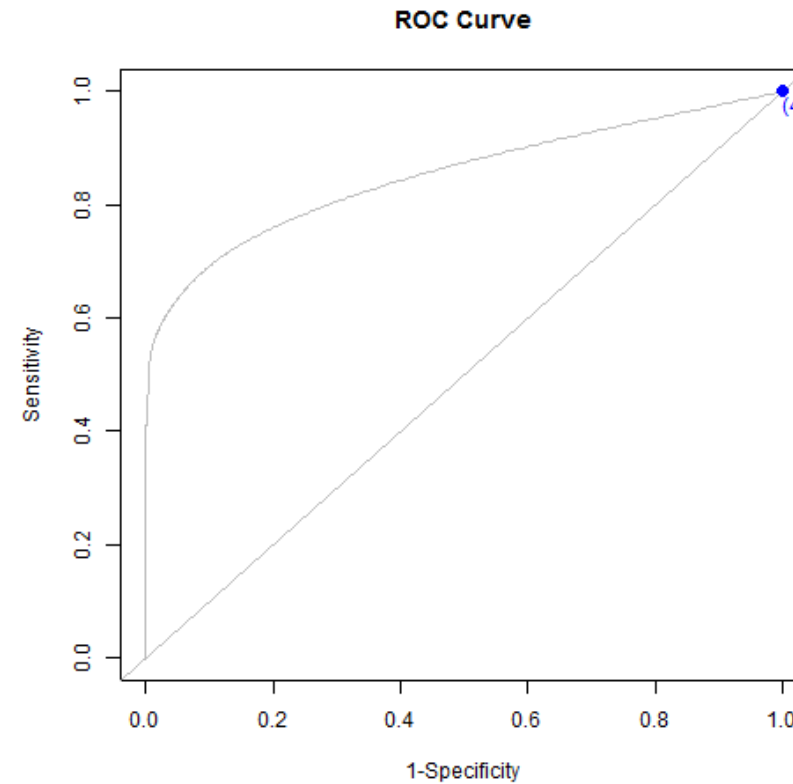
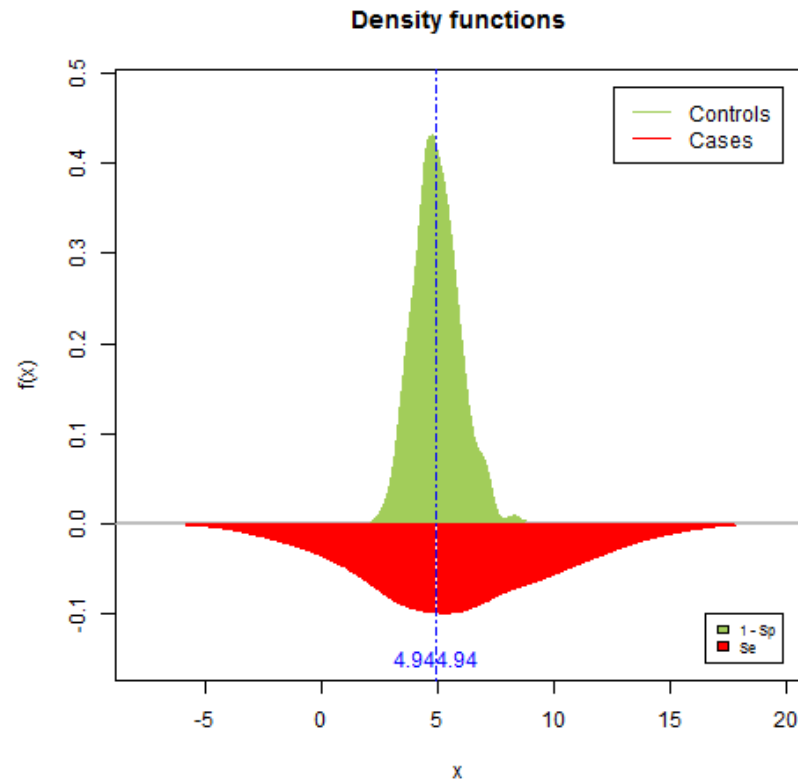


Interval marker

2 cut-off points



ROC CURVE GENERALIZATION (Martínez-Camblor et al., 2017)



&

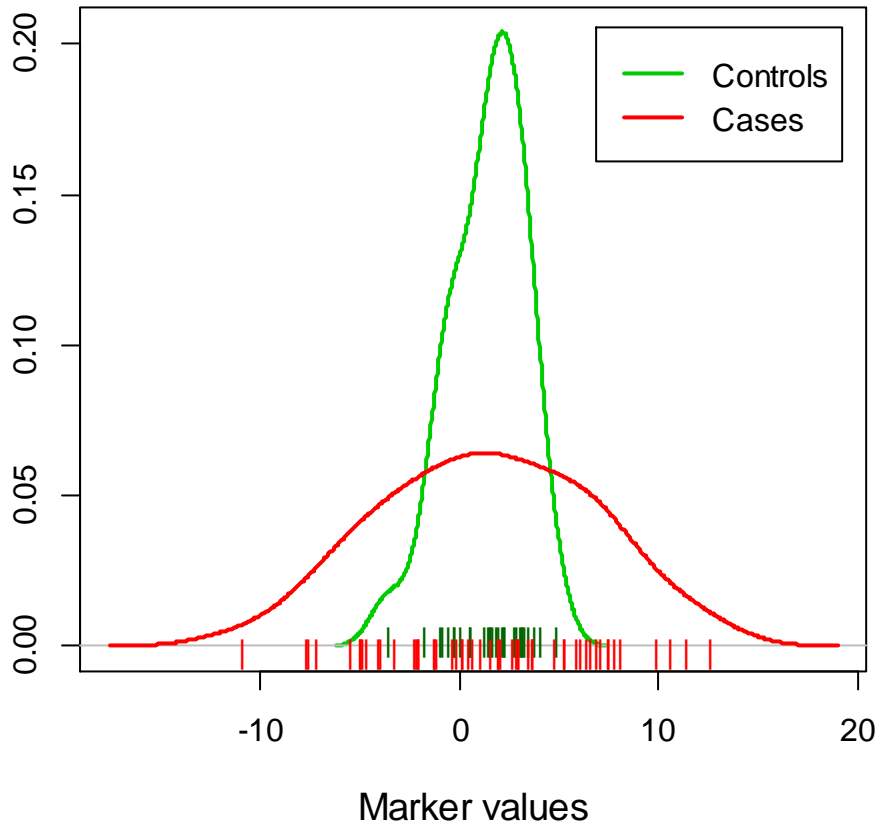
Marker values

Probability of being a case

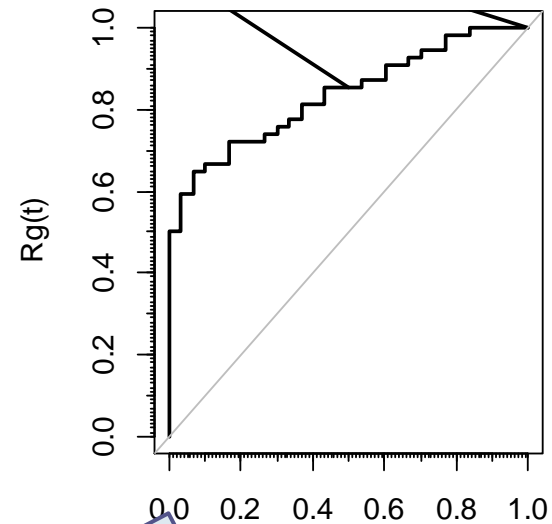
2 cut-off points

ROC CURVE GENERALIZATION – What happens with the intervals?

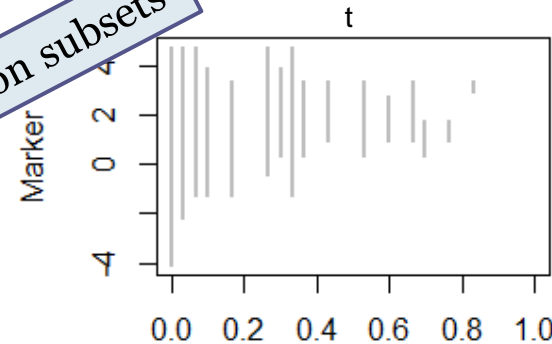
Density estimation



ROC Curve
(without restrictions)



Classification subsets



Some individuals classified as cases for a considered specificity are classified as controls for a lower specificity

ROC CURVE GENERALIZATION – What happens with the intervals?

To avoid losing interpretability Restriction: Self-contained marker intervals

$$R_g(t) = \sup_{(x_l, x_u) \in \mathcal{F}_t} \{F_\xi(x_l) + 1 - F_\chi(x_u)\}$$

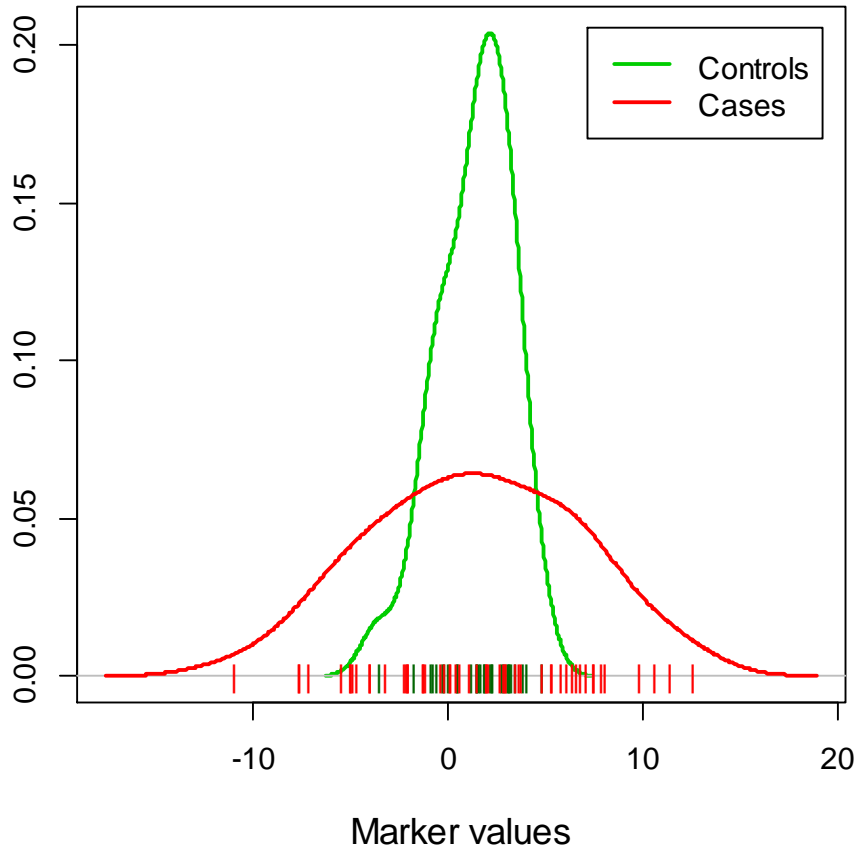
where F_ξ and F_χ denote the CDF of cases and controls, respectively, and

$$\mathcal{F}_t := \{(x_l, x_u) \in \mathbb{R}^2 \text{ such that } F_\chi(x_u) - F_\chi(x_l) = 1 - t\}.$$

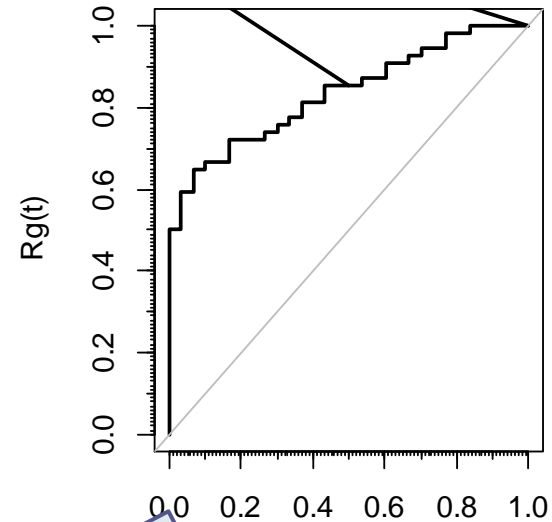
New problem: Restricted optimization

ROC CURVE GENERALIZATION – What happens with the intervals?

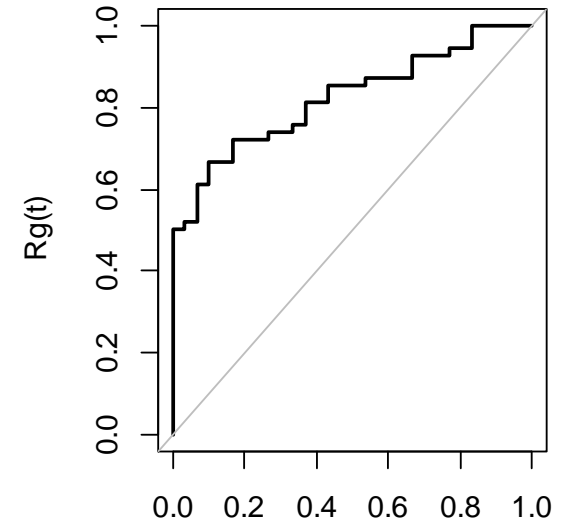
Density estimation



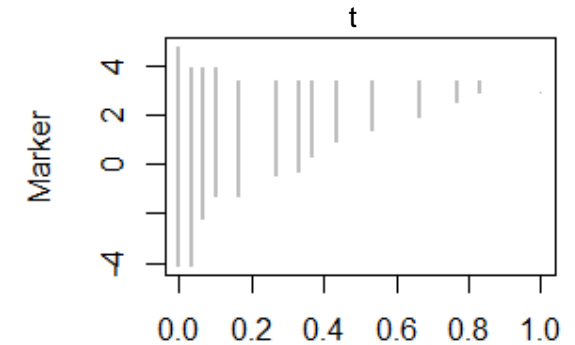
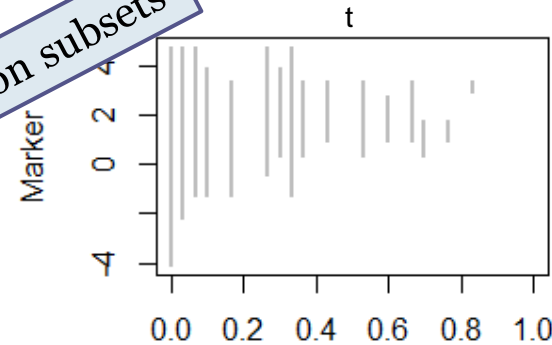
ROC Curve (without restrictions)



ROC Curve (with restrictions)



Classification subsets





ROC CURVE GENERALIZATION

New problems to deal with = Open Questions

- I. The generalization requires many pair-of-points evaluations
Computational times.
Example: replace *for* loops with *sapply* or *matrix operations* if they are non-iterative. For iterative methods (general ROC curve with self-contained intervals restriction) any other solution?
- II. Any experience with *animation* R package or any other library for creating videos/GIFs?
- III. When you are involved in a new problem "away from your main field", how to deal with the literature searching?
- IV. Literature searching: possible overlapping?
- V. Contacting authors of scientific papers: experiences.



THANK YOU!



Contact:
perezsonia@uniovi.es