

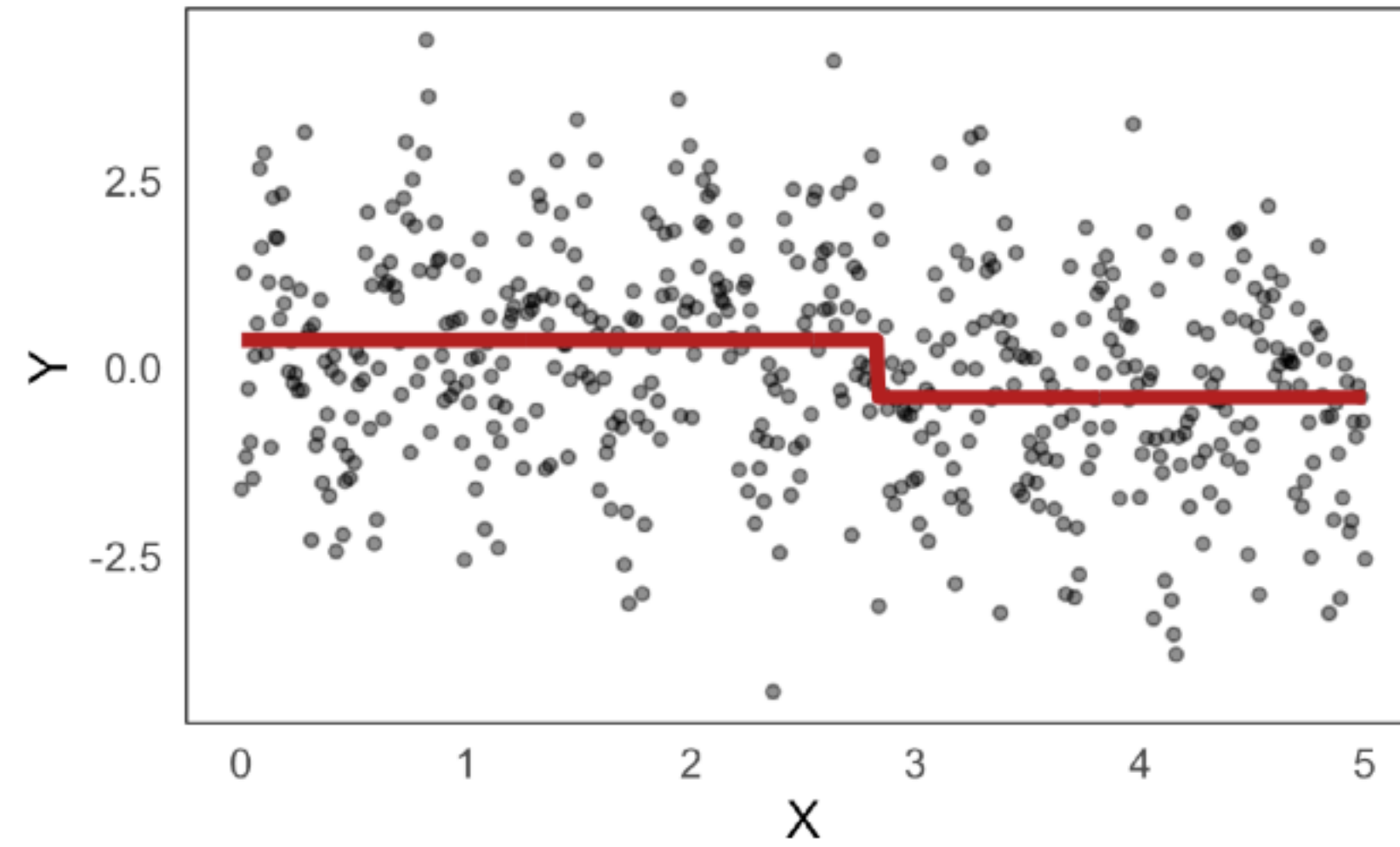
# Regression Trees

- For **classification**, purity of the regions is a good indicator the performance of the model
- For **regression**, we want to select a splitting criterion that promotes splits that improves the predictive accuracy of the model as measured by e.g. the MSE
  1. start with an empty decision tree
  2. choose a predictor on which to split and choose a threshold for splitting such that the weighted average MSE of the new region is as small as possible
  3. Recurse on each node until **stopping condition** is met
    - ▶ maximum depth
    - ▶ minimum number of points in region

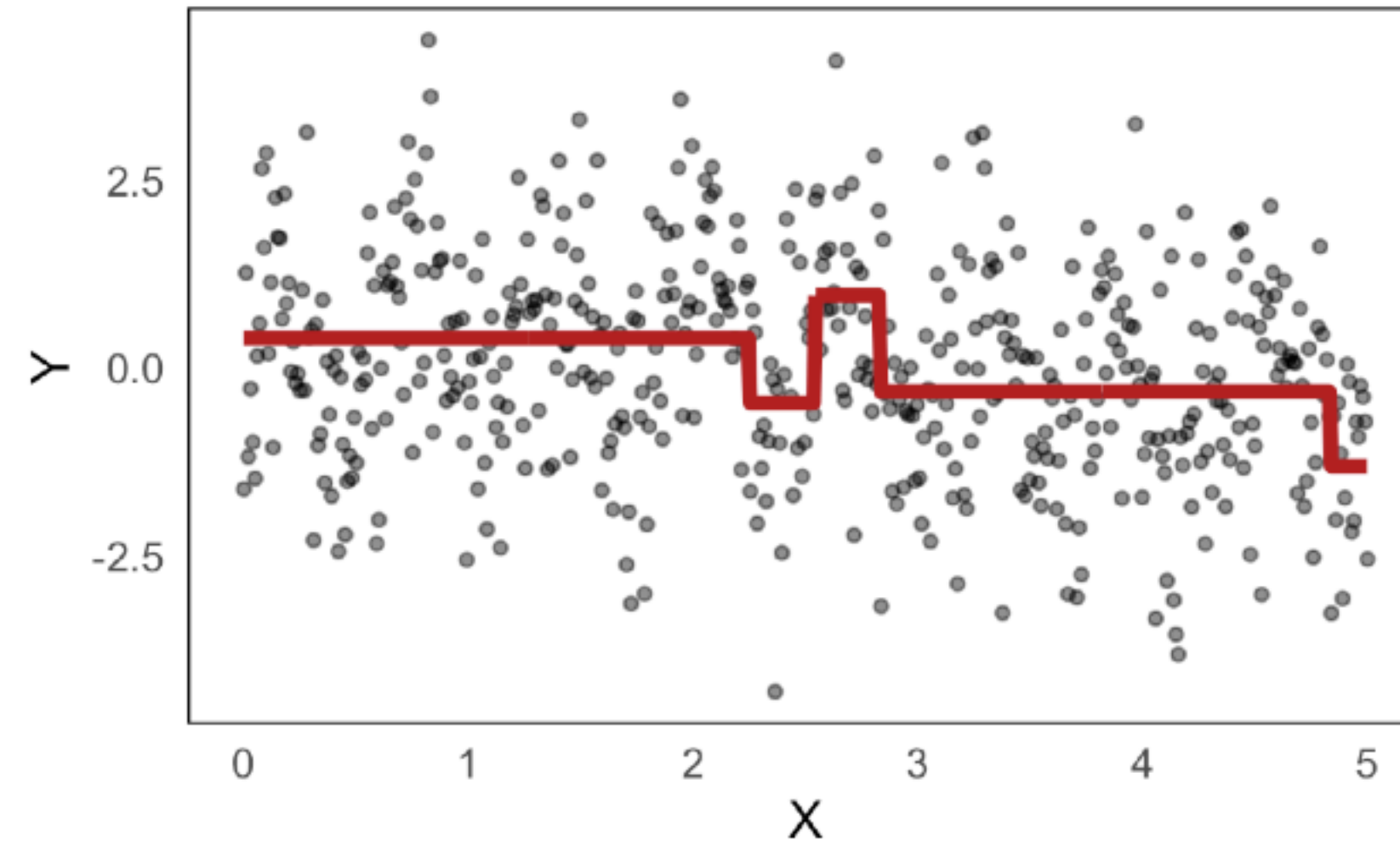
instead of purity gain, we instead compute accuracy gain

# Regression Trees

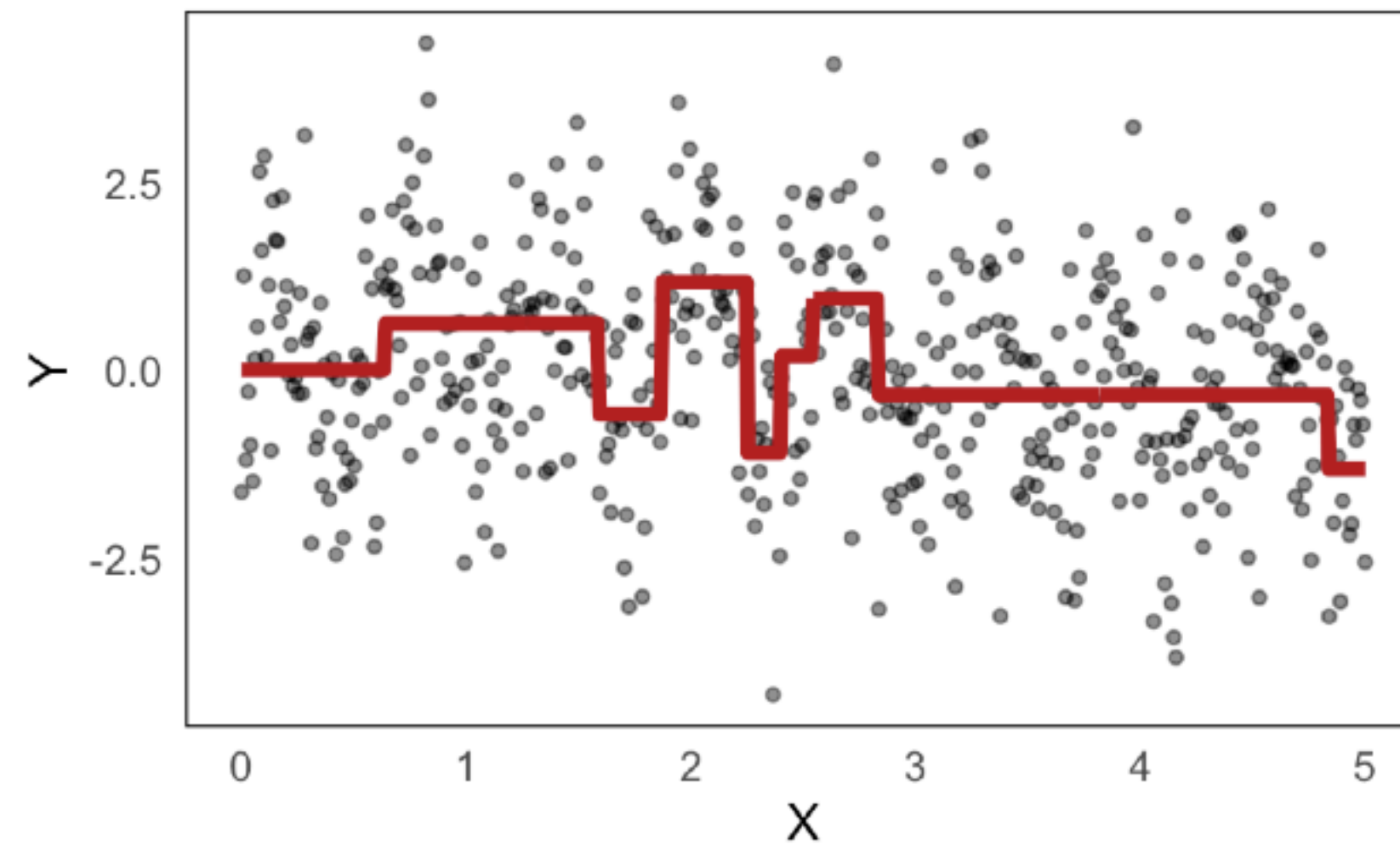
Regression Tree (Depth 1)



Regression Tree (Depth 3)



Regression Tree (Depth 6)



Regression Tree (Depth 10)

