Boosting

- add new trees depending on the previous tree

residuals of the previous tree

- (Y, X)
- $\hat{f}^{(b)}(x_i) = \hat{f}^{(b-1)}(x_i) + \lambda T_b(x_i)$

want to predict e_i^{b-1} at x_i

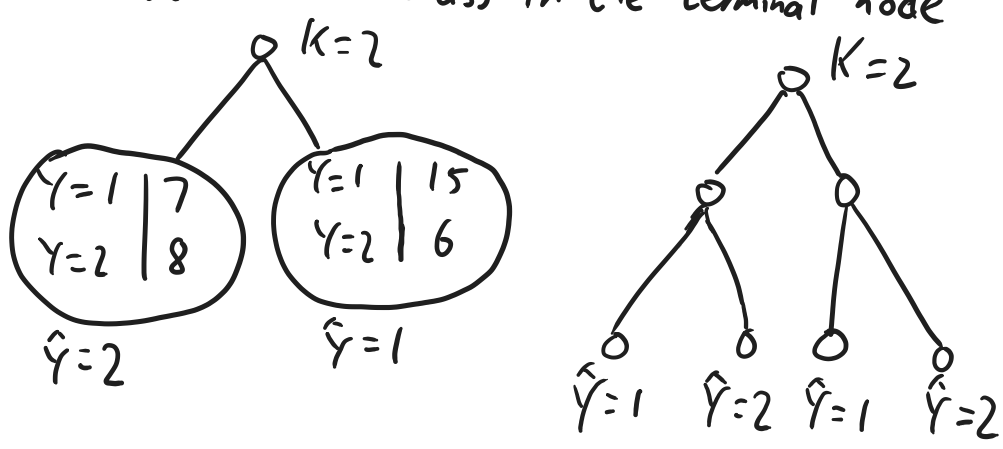
start boosting $\bar{Y}$

$e_i = \bar{Y} - Y_i$ → want a tree to predict these residuals,

classification trees

Y-category/class
↳ K possible values

- how to get "best" split
- $\hat{Y}$ for terminal nodes
- how we prune (only prune if only using one tree)
- $\hat{Y}$ → most common class in the terminal node



how to choose "best" split

- misclassification rate not the right answer
- want to split to get "pure" nodes

n_t = # obs in node t

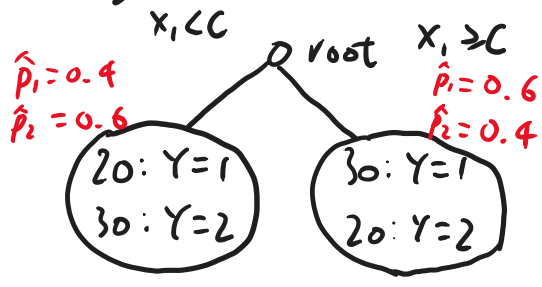
$\hat{p}_{tk}$ = proportion of obs in node t from class k

$= \frac{n_{tk}}{n_t}$

Gini Index want to minimize

$$\sum_{k=1}^{K} \hat{p}_{tk}(1 - \hat{p}_{tk})$$ across the proposed split

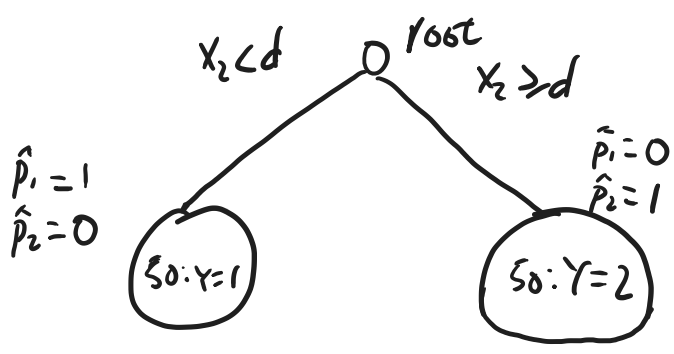
e.g: $k=2$, $n=100$



→ Gini = 0.4(0.6) + 0.6(0.4) + 0.6(0.4) + 0.4(0.6)

left right

= 0.96



→ Gini = 1(0) + 0(1) + 0(1) + 1(0)

= 0