Lecture 33: Assumptions and diagnostics

(randomized)

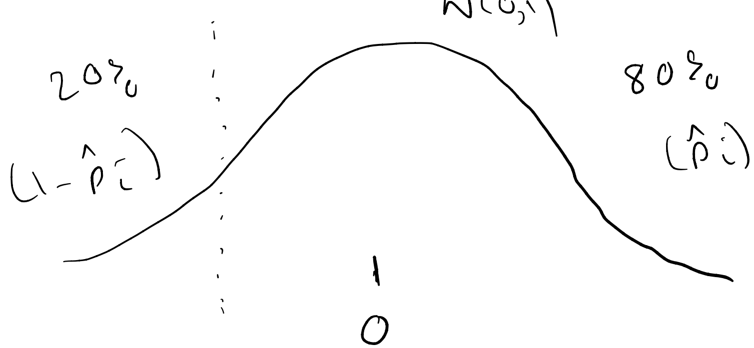
Quantile residuals

(checking the shape assumption)

Motivation: Suppose $\hat{p}_i = 0.8$. I want to create a residual r_i that behaves like residuals in linear regression. Want:

- If $\hat{p}_i \approx p_i$ (good estimate), $\mathbb{E}[r_i | X_i] \approx 0$
- If $\hat{p}_i > p_i$ (overestimate), $\mathbb{E}[r_i | X_i] < 0$
- If $\hat{p}_i < p_i$ (underestimate), $\mathbb{E}[r_i | X_i] > 0$
- want $r_i | X_i \approx$ Normal (if $\hat{p}_i \approx p_i$)

Idea: $\hat{p}_i = 0.8$. Divide $N(0,1)$ into 2 regions



Based on observed Y_i, sample a residual from one part:

- If $Y_i = 1$, sample r_i from the right side
- If $Y_i = 0$, sample r_i from the left side

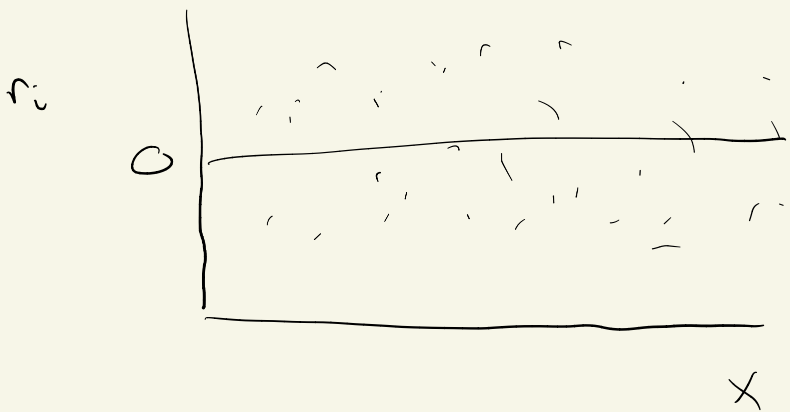
If $\hat{p}_i \approx p_i$, then on average I'm sampling from a $N(0,1)$

Quantile residual plot:

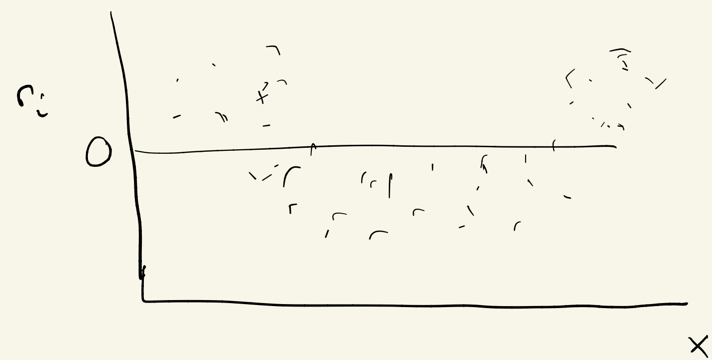
1) Calculates a quantile residual r_i for all observations

2) Plot against X_i

Shape assumption looks good!



Shape assumption is violated



Leverage and Cook's distance

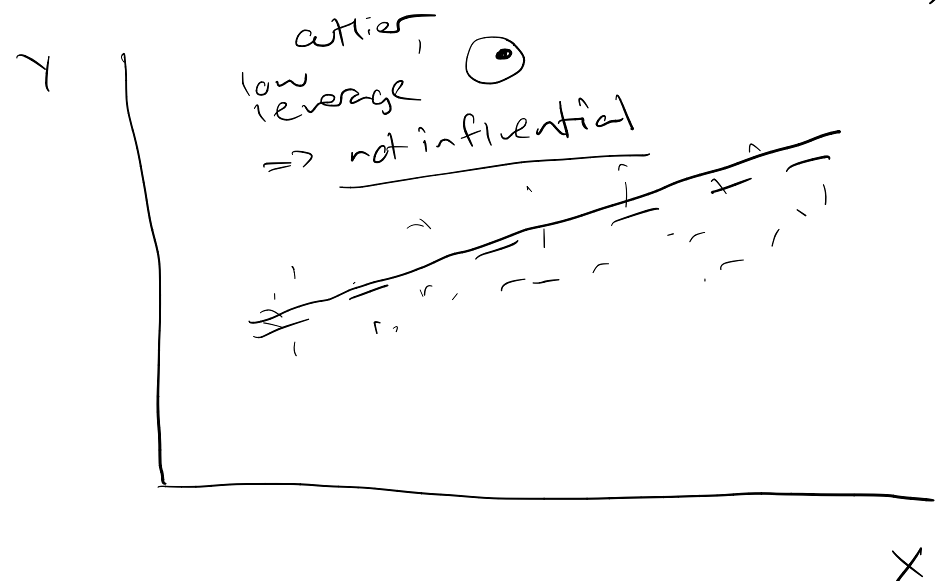
<u>Linear regression</u>:

$$\hat{Y} = X\hat{\beta}$$
$$= X(X^TX)^{-1}X^TY$$
$$= HY \quad \text{“hat matrix”}$$

$$\mathrm{Var}(Y-\hat{Y}) = \sigma^2(I-H)$$

$$\mathrm{Var}(Y_i-\hat{Y}_i) = \sigma^2(1-h_i)$$

$h_i = [H]_{ii}$ ← <u>leverage</u> of the ith observation (potential for a point to be influential)



Influence depends on both leverage & on $Y_i - \hat{Y}_i$

Cook's distance (linear regression)

$$D_i = \frac{(Y_i - \hat{Y}_i)^2}{\underbrace{(k+1)}_{\text{\# of }\beta\text{s in model}} \underbrace{\hat{\sigma}^2}_{\text{estimated variance of residuals}}} \cdot \underbrace{\frac{h_i}{(1-h_i)^2}}_{\text{high leverage?}}$$

outlier?

rule of thumb: concerned that a point might be influential when $D_i >$ threshold (usually 0.5 or 1)

Logistic regression:

Hat matrix $= W^{\frac{1}{2}} X (X^T W X)^{-1} X^T W^{\frac{1}{2}}$

H

$h_i =$ leverage

$W = \mathrm{diag}(p_i(1-p_i))$

$W^{\frac{1}{2}} = \mathrm{diag}(\sqrt{p_i(1-p_i)})$

$$D_i = \frac{(Y_i - \hat{p}_i)^2}{(k+1)\,\hat{p}_i(1-\hat{p}_i)} \cdot \frac{h_i}{(1-h_i)^2}$$

(concerned when $D_i > 0.5$ or 1)