Diagnostics

$$Y = X\beta + \varepsilon$$

$$\mathbb{E}Y = X\beta ; \quad \hat{Y} = X\hat{\beta} = X(X^TX)^{-1}X^TY$$

$$e = Y - \hat{Y} = (I-P)Y = QY$$

- $\mathbb{E}e = (I-P)X\beta = QX\beta = 0$
- $\operatorname{Var} e = \sigma^2 Q$
- $\operatorname{Cov}(e, \hat{Y}) = \operatorname{Cov}(QY, PY) = 0$

Define ① $r_i = \dfrac{e_i}{\sqrt{\sigma^2(1-h_{ii})}}$ Internally Studentized Residual

② $t_i = \dfrac{e_i}{S_{(i)}\sqrt{1-h_{ii}}}$ where Externally Studentized residual

$S_{(i)}^2$ estimate σ^2 without the i^{th} case.

Q: Are they related & what are their distributions?

$$\hat{\beta} - \hat{\beta}_{(i)} = \frac{(X^TX)^{-1}X_i e_i}{1-h_{ii}}, \quad e_i = Y_i - X_i^T\hat{\beta}_{(i)}$$

Use this to show relationship between

$$S_{(i)}^2 \ \& \ S^2$$

Verify: $(n-p-1)S_{(i)}^2 = (n-p)S^2 - \dfrac{e_i^2}{1-h_{ii}}$

$$t_i^2 \overset{d}{=} \frac{B}{1-B}(n-p-1), \quad B \sim \text{Beta}\left(\tfrac{1}{2}, \tfrac{n-p-1}{2}\right)$$

Recall: $\left(\dfrac{\frac{a}{b}F}{1+\frac{a}{b}F} \sim B\left(\tfrac{a}{2}, \tfrac{b}{2}\right) \right)$

Exact Dist. of r_i^2 not available, but

$$\frac{r_i^2}{n-p} \sim B\left(\tfrac{1}{2}, \tfrac{n-p-1}{2}\right)$$

Test of outlier: Is the i^{th} case an outlier for the x-value.

$$Y = X\beta + \varepsilon$$

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} X_1^T\beta \\ \vdots \\ X_n^T\beta \end{pmatrix} + \delta \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i^{th}$$

Show the test statistic for $\delta = 0$ is

$$F = t_i^2 \ \&$$

t_i externally studentized residual.

We need:

$$P_{C(X)} = X(X^TX)^{-1}X^T$$

$$\underbrace{P_{C(X|(0,\dots,1,\dots,0)^T)}}\, y = P_{C(X|(0,\dots,1,\dots,0)^T)}\, y$$

$$P_{C(X)} + \frac{(I-P_X)e_ie_i^T(I-P_X)}{e_i^T(I-P_X)e_i}$$