36401 HW #9.
Sylvia (shuynan) Ding
shuynand

problem #2.

<u>Case 1.</u>

(p in this problem = 5)

$r_i = \hat{\varepsilon}_i / \sqrt{MSE} = \hat{\varepsilon}_i / 2 = 0.315.$

$D_i = \left(\frac{r_i^2}{p}\right) \cdot \left[\frac{h_{ii}}{(1-h_{ii})}\right] = \left(\frac{(0.315)^2}{5}\right) \cdot \left[\frac{0.9}{(0.1)}\right]$

$\approx \boxed{0.178}$

In this case Di is really small so this observation is not influential

(we compare Di with $F_{0.5}(p, n-p) \approx 1$, generally)

$SSE = MSE \cdot (n-p)$
$= 4 \cdot (49)$
$= 196$

$t_i = \hat{\varepsilon}_i \cdot \sqrt{\frac{(n-p-1)}{SSE(1-h_{ii}) - \hat{\varepsilon}_i^2}}$

$= 0.63 \cdot \sqrt{\frac{48}{196(0.1) - 0.63^2}} \approx \boxed{0.996}$

Test:

$t_{1-\alpha/2, n-p-1} = t_{0.995, 48} = 2.682$

Since t_i is less than 2.682, the first observation is not an outlier

<u>Case 2.</u>

$r_i = \hat{\varepsilon}_i / \sqrt{MSE} = \hat{\varepsilon}_i / 2 = 0.865$

$D_i = \left(\frac{r_i^2}{p}\right) \cdot \left[\frac{h_{ii}}{(1-h_{ii})}\right] = \left(\frac{(0.865)^2}{5}\right) \cdot \left[\frac{0.75}{(0.25)}\right]$

$= \boxed{0.449}$

In this case Di is again less than 1 so this observation is not influential

$t_i = \hat{\varepsilon}_i \cdot \sqrt{\frac{(n-p-1)}{SSE(1-h_{ii}) - \hat{\varepsilon}_i^2}}$

$= 1.73 \sqrt{\frac{48}{196(0.25) - (1.73)^2}} \approx \boxed{1.767}$

Test:

$t_{1-\alpha/2, n-p-1} = t_{0.995, 48} = 2.682$

Since t_i is less than 2.682, the second observation is not an outlier

<u>Case 3.</u>

$r_i = \hat{\varepsilon}_i / \sqrt{MSE} = \hat{\varepsilon}_i / 2 = 4.5$

$D_i = \left(\frac{r_i^2}{p}\right) \cdot \left[\frac{h_{ii}}{(1-h_{ii})}\right] = \left(\frac{4.5^2}{5}\right) \cdot \left[\frac{0.25}{(0.75)}\right]$

$= \boxed{1.35}$

In this case Di is greater than 1 so this observation is really influential

$$t_i = \hat{\varepsilon}_i \cdot \sqrt{\frac{(n-p-1)}{SSE(1-h_{ii}) - \hat{\varepsilon}_i^2}}$$

$$= 9 \cdot \sqrt{\frac{48}{196(0.75) - 81}} \approx \boxed{7.68}$$

Test:

$$t_{1-\alpha/2,\, n-p-1} = t_{0.995,48} = 2.682$$

Since t_i is greater than 2.682, the third observation is an outlier

Case 4

$$r_i = \hat{\varepsilon}_i / \sqrt{MSE} = \hat{\varepsilon}_i / 2 = 5.15$$

$$D_i = \left(\frac{r_i^2}{p}\right) \cdot \left[\frac{h_{ii}}{(1-h_{ii})}\right] = \left(\frac{5.15^2}{5}\right) \cdot \left[\frac{0.029}{(0.971)}\right]$$

$$\approx \boxed{0.158}$$

In this case Di is less than 1 and really small so this observation is not influential

$$t_i = \hat{\varepsilon}_i \cdot \sqrt{\frac{(n-p-1)}{SSE(1-h_{ii}) - \hat{\varepsilon}_i^2}}$$

$$= 10.3 \sqrt{\frac{48}{196(0.971) - 10.3^2}} \approx \boxed{7.78}$$

Test:

$$t_{1-\alpha/2,\, n-p-1} = t_{0.995,48} = 2.682.$$

Since t_i is greater than 2.682, this observation is an outlier

$$F^* = \frac{SSE(\text{Reduced}) - SSE(\text{Full})}{df_{\text{reduced}} - df_{\text{Full}}} \Big/ \frac{SSE(\text{full})}{df_{\text{Full}}}$$

$$= \frac{243006 - 193700}{48 - 46} \Big/ \frac{193700}{46} \approx 5.855$$

$F(0.05, 2, 46) \approx 3.2$, and therefore, $F^* > F(0.05, 2, 46)$.