/* Influence Statistics */

[Leverage]

$$h_{ii} = \frac{1}{n} + \frac{(x_i - \overline{X}_i)^2}{\sum_{i=1}^{n}(X_i - \overline{X}_i)^2}$$

$h_{ii} > \frac{2(k+1)}{n}$: high leverage(Belsley, Kuh, Welsch)

cf) Hat Diag H

[RSTUDENT]

→ discrepancy

$$t_i = \frac{e_{i(i)}}{\sqrt{MSE_{(i)}(1-h_{ii})}}$$

$MSE_{(i)}$: the mean square error with obs i deleted

$|t_i| > 2.0$: high discrepancy

[Influence]

(Cook'sD)

→ Cook's D : information about how the regression eq. changes

$$D_i = \frac{\sum_{i=1}^{n}(\hat{Y} - \hat{Y}_{i(i)})^2}{k\,MSE}$$

$D_i > 1.0$: potential infulential obs. (Cook and Weisberg, 1982)

or $D_i > \frac{4}{n}$ (Bollen and Jackman, 1990)

(DFFITS)

$$DFFITS_i = \frac{\hat{Y}_i - \hat{Y}_{i(i)}}{\sqrt{MSE_{(i)}h_{ii}}} = t_i\sqrt{\frac{h_{ii}}{(1-h_{ii})}}$$

$$|DFFITS| > 2\sqrt{\frac{k+1}{n}}$$

k : number of predictors