Print

The Leverage Statistic and the DFBETAS Statistic

Leverage Statistic

Consider a regression model with two independent variables, $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$. The vector of observed values $\boldsymbol{Y}$ is not in the plane defined by the vectors $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$. The perpendicular projection from $\boldsymbol{Y}$ onto the plane of $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ defines the vector $\hat{\boldsymbol{Y}}$ of predicted values. The projection matrix, also called the *hat matrix*, is shown below:

$$\boldsymbol{P} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$$

The leverage values are the diagonal elements of the projection matrix.

DFBETAS Statistic

DFBETAs measure the change in each parameter estimate when an observation is deleted from the model.

$$DFBETA_{j(i)} = \frac{b_j - b_{j(i)}}{\hat{\sigma}(b_j)}$$

where

- b_j is the parameter estimate for the j^{th} independent variable.
- $b_{j(i)}$ is the parameter estimate for the j^{th} independent variable with the i^{th} observation deleted from the analysis.
- $\hat{\sigma}(b_j)$ is the standard error of the j^{th} parameter estimate when all observations are included in the analysis.

Close