Lecture 13
Model Diagnostics

STAT 8020 Statistical Methods II
September 18, 2019

Leverage
Studentized & Jackknife Residuals
DFFITS
Non-Constant Variance & Transformation

Whitney Huang
Clemson University

Agenda

1. **Leverage**

2. **Studentized & Jackknife Residuals**

3. **DFFITS**

4. **Non-Constant Variance & Transformation**

Leverage

Leverage
Studentized & Jackknife Residuals
DFFITS
Non-Constant Variance & Transformation

Recall in MLR that $\hat{Y} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T Y = \boldsymbol{H}\boldsymbol{Y}$ where $\boldsymbol{H}$ is the hat-matrix

Leverage

Leverage
Studentized & Jackknife Residuals
DFFITS
Non-Constant Variance & Transformation

Recall in MLR that $\hat{Y} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T Y = \boldsymbol{H}\boldsymbol{Y}$ where $\boldsymbol{H}$ is the hat-matrix

- The leverage score for the i_{th} observation is defined as:

$$h_i = \boldsymbol{H}_{ii}$$

Leverage

Studentized & Jackknife Residuals

DFFITS

Non-Constant Variance & Transformation

Leverage

Recall in MLR that $\hat{Y} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T Y = \boldsymbol{H}\boldsymbol{Y}$ where $\boldsymbol{H}$ is the hat-matrix

- The leverage score for the i_{th} observation is defined as:

$$h_i = \boldsymbol{H}_{ii}$$

Leverage

Recall in MLR that $\hat{Y} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T Y = \boldsymbol{H}\boldsymbol{Y}$ where $\boldsymbol{H}$ is the hat-matrix

- The leverage score for the i_{th} observation is defined as:

$$h_i = \boldsymbol{H}_{ii}$$

- Can show that $\text{Var}(e_i) = \sigma^2(1 - h_i)$, where $e_i = Y_i - \hat{Y}_i$ is the residual for the i_{th} observation

Leverage

Leverage
Studentized & Jackknife Residuals
DFFITS
Non-Constant Variance & Transformation

Recall in MLR that $\hat{Y} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^TY = \boldsymbol{H}\boldsymbol{Y}$ where $\boldsymbol{H}$ is the hat-matrix

- The leverage score for the i_{th} observation is defined as:

$$h_i = \boldsymbol{H}_{ii}$$

- Can show that Var$(e_i) = \sigma^2(1 - h_i)$, where $e_i = Y_i - \hat{Y}_i$ is the residual for the i_{th} observation

Leverage

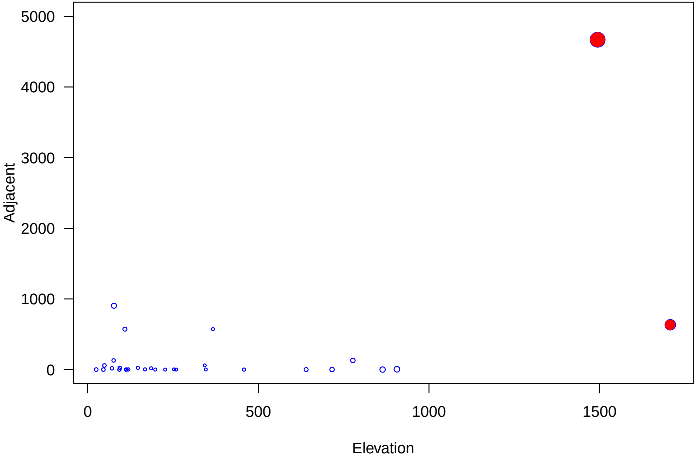
Recall in MLR that $\hat{Y} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^TY = \boldsymbol{H}\boldsymbol{Y}$ where $\boldsymbol{H}$ is the hat-matrix

- The leverage score for the i_{th} observation is defined as:

$$h_i = \boldsymbol{H}_{ii}$$

- Can show that Var$(e_i) = \sigma^2(1 - h_i)$, where $e_i = Y_i - \hat{Y}_i$ is the residual for the i_{th} observation
- $\frac{1}{n} \leq h_i \leq 1, \quad 1 \leq i \leq n$ and $\bar{h}_i = \frac{p}{n} \Rightarrow$ a “rule of thumb" is that leverages of more than $\frac{2p}{n}$ should be looked at more closely

Leverage Scores of `Species` $\sim$ `Elev` $+$ `Adj`



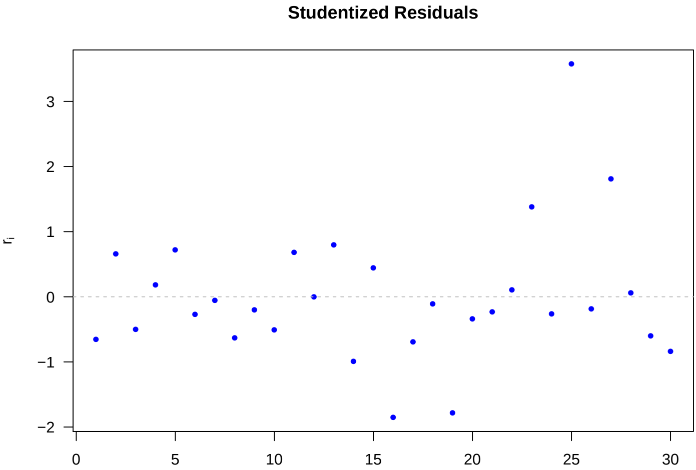
Studentized Residuals

Leverage
Studentized & Jackknife Residuals
DFFITS
Non-Constant Variance & Transformation

As we have seen Var$(e_i) = \sigma^2(1 - h_i)$, this suggests the use of $r_i = \frac{e_i}{\hat{\sigma}\sqrt{(1-h_i)}}$

- r_i's are called **studentized residuals**. r_i's are sometimes preferred in residual plots as they have been standardized to have equal variance.
- If the model assumptions are correct then Var$(r_i) = 1$ and Corr(e_i, e_j) tends to be small

Studentized Residuals of `Species` $\sim$ `Elev` + `Adj`



Leverage
Studentized & Jackknife Residuals
DFFITS
Non-Constant Variance & Transformation

Jackknife Residuals

- For a given model, exclude the observation i and recompute $\hat{\boldsymbol{\beta}}_{(i)}$, $\hat{\sigma}_{(i)}$ to obtain $\hat{Y}_{i(i)}$

Jackknife Residuals

- For a given model, exclude the observation i and recompute $\hat{\boldsymbol{\beta}}_{(i)}$, $\hat{\sigma}_{(i)}$ to obtain $\hat{Y}_{i(i)}$

Jackknife Residuals

- For a given model, exclude the observation i and recompute $\hat{\boldsymbol{\beta}}_{(i)}$, $\hat{\sigma}_{(i)}$ to obtain $\hat{Y}_{i(i)}$
- The observation i is an outlier if $\hat{Y}_{i(i)} - Y_i$ is “large”

Jackknife Residuals

Leverage
Studentized & Jackknife Residuals
DFFITS
Non-Constant Variance & Transformation

- For a given model, exclude the observation i and recompute $\hat{\boldsymbol{\beta}}_{(i)}$, $\hat{\sigma}_{(i)}$ to obtain $\hat{Y}_{i(i)}$
- The observation i is an outlier if $\hat{Y}_{i(i)} - Y_i$ is "large"

Jackknife Residuals

- For a given model, exclude the observation i and recompute $\hat{\boldsymbol{\beta}}_{(i)}$, $\hat{\sigma}_{(i)}$ to obtain $\hat{Y}_{i(i)}$
- The observation i is an outlier if $\hat{Y}_{i(i)} - Y_i$ is "large"
- Can show Var$(\hat{Y}_{i(i)} - Y_i) = \sigma^2 \left(1 + \boldsymbol{x}_i^T (\boldsymbol{X}_{(i)}^T \boldsymbol{X}_i)^{-1} \boldsymbol{x}_i\right)$

Jackknife Residuals

- For a given model, exclude the observation i and recompute $\hat{\boldsymbol{\beta}}_{(i)}$, $\hat{\sigma}_{(i)}$ to obtain $\hat{Y}_{i(i)}$
- The observation i is an outlier if $\hat{Y}_{i(i)} - Y_i$ is "large"
- Can show Var$(\hat{Y}_{i(i)} - Y_i) = \sigma^2 \left(1 + \boldsymbol{x}_i^T (\boldsymbol{X}_{(i)}^T \boldsymbol{X}_i)^{-1} \boldsymbol{x}_i\right)$

Leverage

Studentized & Jackknife Residuals

DFFITS

Non-Constant Variance & Transformation

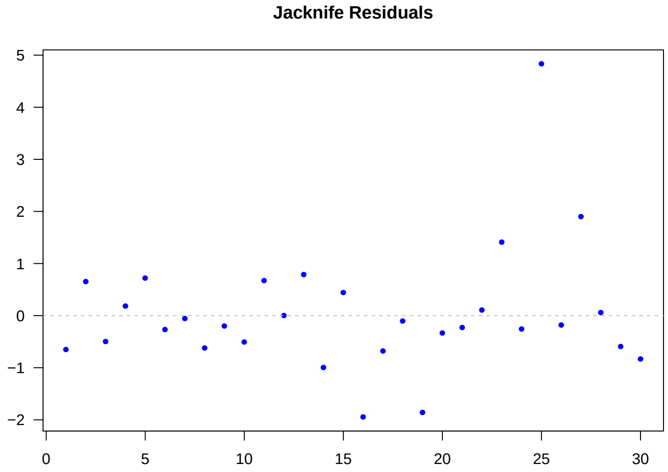
Jackknife Residuals

- For a given model, exclude the observation i and recompute $\hat{\boldsymbol{\beta}}_{(i)}$, $\hat{\sigma}_{(i)}$ to obtain $\hat{Y}_{i(i)}$
- The observation i is an outlier if $\hat{Y}_{i(i)} - Y_i$ is "large"
- Can show Var$(\hat{Y}_{i(i)} - Y_i) = \sigma^2 \left(1 + \boldsymbol{x}_i^T (\boldsymbol{X}_{(i)}^T \boldsymbol{X}_i)^{-1} \boldsymbol{x}_i\right)$
- Define the **jackknife residuals** as

$$t_i = \frac{\hat{Y}_{i(i)} - Y_i}{\sqrt{\hat{\sigma}^2 \left(1 + \boldsymbol{x}_i^T (\boldsymbol{X}_{(i)}^T \boldsymbol{X}_i)^{-1} \boldsymbol{x}_i\right)}}$$

  which are distributed as a t_{n-p} if the model is correct and $\varepsilon \sim \mathrm{N}(\mathbf{0}, \sigma^2 \boldsymbol{I})$

Jackknife Residuals of `Species` $\sim$ `Elev` + `Adj`

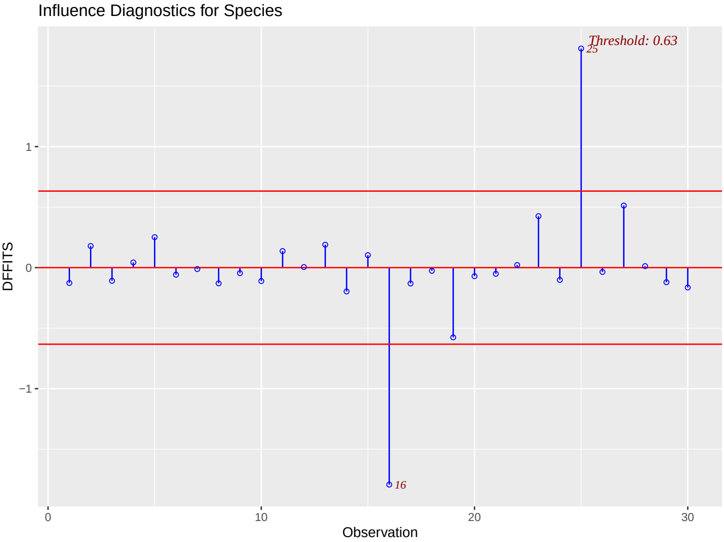


Influential Observations

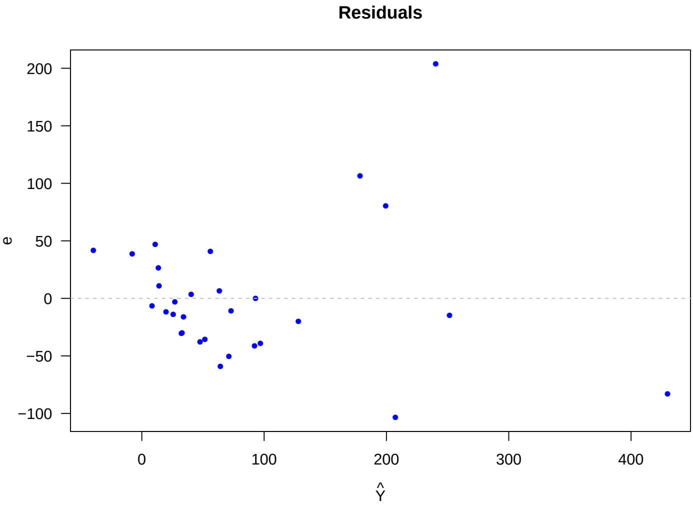
DFFITS

- Difference between the fitted values $\hat{Y}_i$ and the predicted values $\hat{Y}_{i(i)}$
- DFFITS$_i = \frac{\hat{Y}_i - \hat{Y}_{i(i)}}{\sqrt{\text{MSE}_{(i)} h_i}}$
- Concern if absolute value greater than 1 for small data sets, or greater than $2\sqrt{p/n}$ for large data sets

DFFITS of `Species` $\sim$ `Elev` + `Adj`



Residual Plot of $\texttt{Species} \sim \texttt{Elev} + \texttt{Adj}$



Residual Plot After Square Root Transformation