

Studentized Residuals

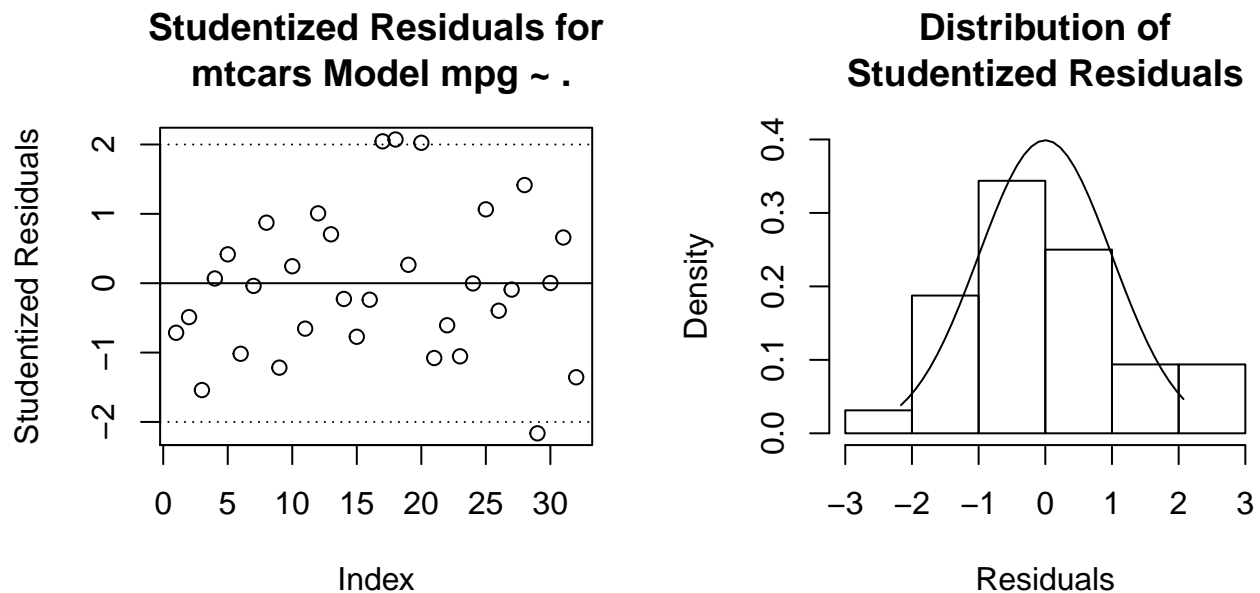
Definition

Studentized residuals are the standardized “jackknifed” residuals. “Jackknifing” is the process of *leave one out* such that the studentized residual is the residual where y_i is compared to the fitted model point \hat{y}_i where y_i is not included (equation 1). In equation 1, the denominator is no longer the Mean Square Error, but the Predicted Sum of Squares, or PRESS statistics. NOTE: h_{ii} is called “leverage” in that it measures the ability of a point to change the regression coefficient (slope) of the model. The greater the effect of moving y_i on the regression slope, the greater its “leverage.”

Equations

$$t_i = \frac{\hat{e}_i}{\hat{\sigma}\sqrt{1-h_{ii}}}$$
$$h_{ii} = \frac{d\hat{y}_i}{dy_i}$$

Example Plots



Interpretation and Use

Studentized residuals allow for evaluating how much of an outlier or leverage point that each y_i is in the data. It serves two purposes in diagnostics. First, the values of t_i indicate if any points are particularly influential in the model if their values lie outside ± 2 . Second, the

distribution of the studentized residuals can indicate if the data are non-normal when the distribution does not follow a normal distribution.

Further Avenues

QQ Plots - Studentized residuals provide a good method to identify if the residuals are non-normal. A QQ plot will better indicate if this is true for determining model goodness-of-fit.

R Code

```
library(MASS)
library(car)
data(mtcars)
model <- lm(mpg ~ ., data=mtcars)
studentResids <- studres(model)
#plot(studentResids, ylab="Studentized Residuals",
#      main="Studentized Residuals for\n mtcars Model mpg ~ .")
#abline(0,0)
#abline(h=2, lty = 3)
#abline(h=-2, lty = 3)
#hist(studentResids, freq=FALSE, main="Distribution of\nStudentized Residuals",
#      xlab="Residuals", ylim=c(0,0.4))
#xfit<-seq(min(studentResids),max(studentResids),length=40)
#yfit<-dnorm(xfit)
#lines(xfit, yfit)
```