Multiple Regression Part 2 Using RStudio for Ag Data

STAT 441/541 Statistical Methods II

Information in the Model Summary Table

```
Residual standard error: 7.642 on 43 degrees of freedom
Multiple R-squared:  0.7396,    Adjusted R-squared:  0.7275
F-statistic: 61.06 on 2 and 43 DF,  p-value: 2.737e-13
```

- Estimate of the model standard deviation

$s_\varepsilon = 7.642$

- Adjusted R-squared = 0.7275

Assess Multicollinearity

```
> VIF(model)
      x6       x9
1.753282 1.753282
```

- The Variance Inflation Factor for independent variable x6 is 1.75
- The Variance Inflation Factor for independent variable x9 is 1.75

Confidence Interval for the Mean, $E(y)$

```
> # confidence interval on E(y) for specified values of all
> #   independent variables
> forecastdata = data.frame(x6=90,
+                           x9=70)
> predict(model,newdata=forecastdata,interval="confidence")
       fit     lwr      upr
1 134.0359 99.9341 168.1377
```

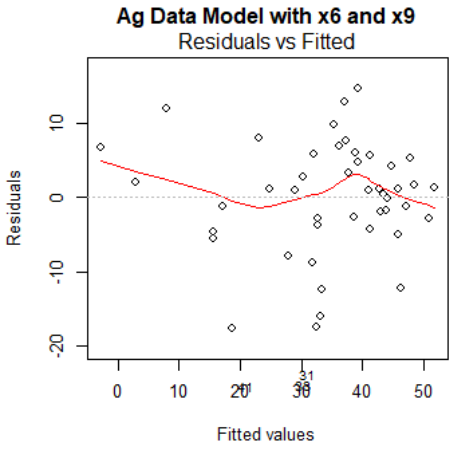
- The predicted value when x6=90 and x9=70 is 134.04
- The lower confidence limit is 99.93 and the upper confidence limit is 168.14 when x6=90 and x9=70

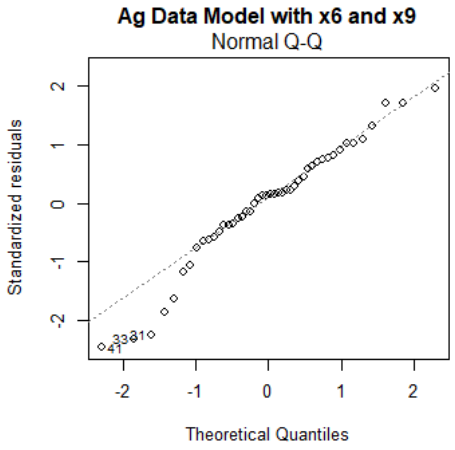
Prediction Interval for Individual Values of *y*

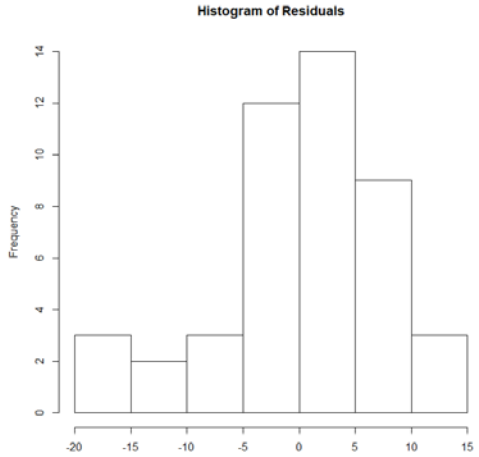
```
> # prediction interval on y for specified values of all
> #   independent variables
> predict(model,newdata=forecastdata,interval="prediction")
       fit      lwr      upr
1 134.0359 96.61352 171.4583
```

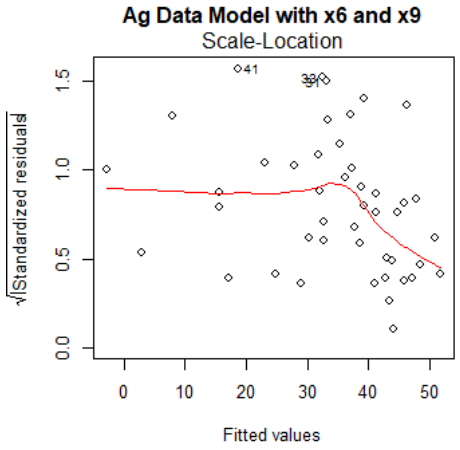
- The predicted value when x6=90 and x9=70 is 134.04
- The lower prediction limit is 96.61 and the upper prediction limit is 171.46 when x6=90 and x9=70

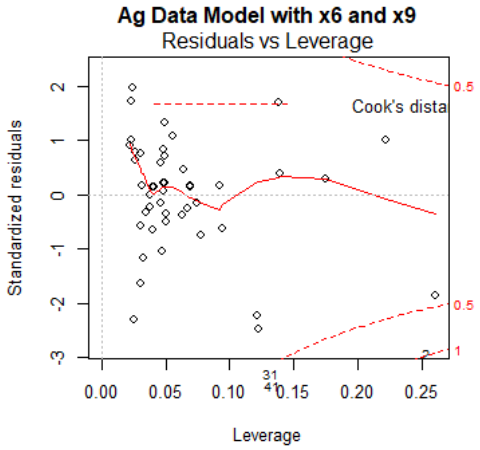
Various Diagnostic Plots for Checking Assumptions

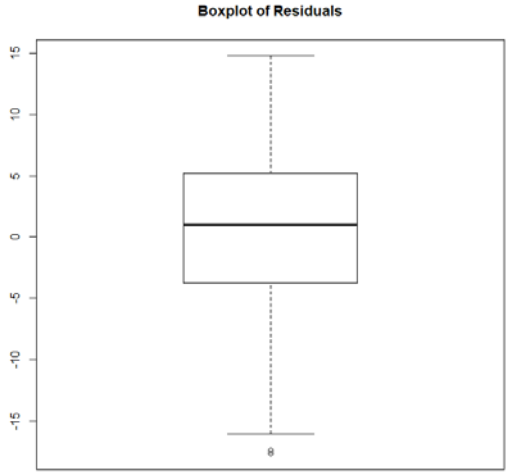












Shapiro-Wilk Test

```
> # Shapiro-Wilk test for normality of errors and use alpha=0.01
> shapiro.test(resid(model))

        Shapiro-Wilk normality test

data:  resid(model)
W = 0.96418, p-value = 0.1665
```

Test Statistic

Breusch-Pagan Test

```
> # Breusch-Pagan Test for a common error variance and use alpha=0.01
> bptest(model)

        studentized Breusch-Pagan test

data:  model
BP = 4.6148, df = 2, p-value = 0.09952
```

Test Statistic

Potential Outliers

dfbetas

Observations flagged as influential

```
> influence.measures(model)
Influence measures of
         lm(formula = y ~ x6 + x9, data = dataobj) :

     dfb.1_   dfb.x6    dfb.x9   dffit cov.r   cook.d    hat inf
1   0.027301 -0.02895 -0.018427  0.03491 1.152 4.16e-04 0.0697
2  -1.073012  0.91723  1.042055 -1.13557 1.130 4.05e-01 0.2608   *
3   0.139330 -0.14036 -0.106024  0.15406 1.235 8.07e-03 0.1398   *
4   0.041139 -0.04640 -0.024258  0.05387 1.180 9.90e-04 0.0923
5   0.097695 -0.08655 -0.088674  0.11841 1.130 4.76e-03 0.0643
6   0.029586 -0.20565  0.220084  0.54005 1.284 9.72e-02 0.2221   *
7   0.026174 -0.06583  0.034510  0.13050 1.292 5.80e-03 0.1745   *
8  -0.274608  0.05972  0.519511  0.69611 1.010 1.54e-01 0.1378
9  -0.143161  0.08664  0.198147  0.26409 1.046 2.32e-02 0.0557
10 -0.046490  0.03927  0.052977  0.12740 1.056 5.46e-03 0.0264
```