Homework2

Israel Diego

8/29/2019

Part A (Maximum 2 pages).

The dataset uswages is drawn as a sample from the Current Population Survey in 1988.
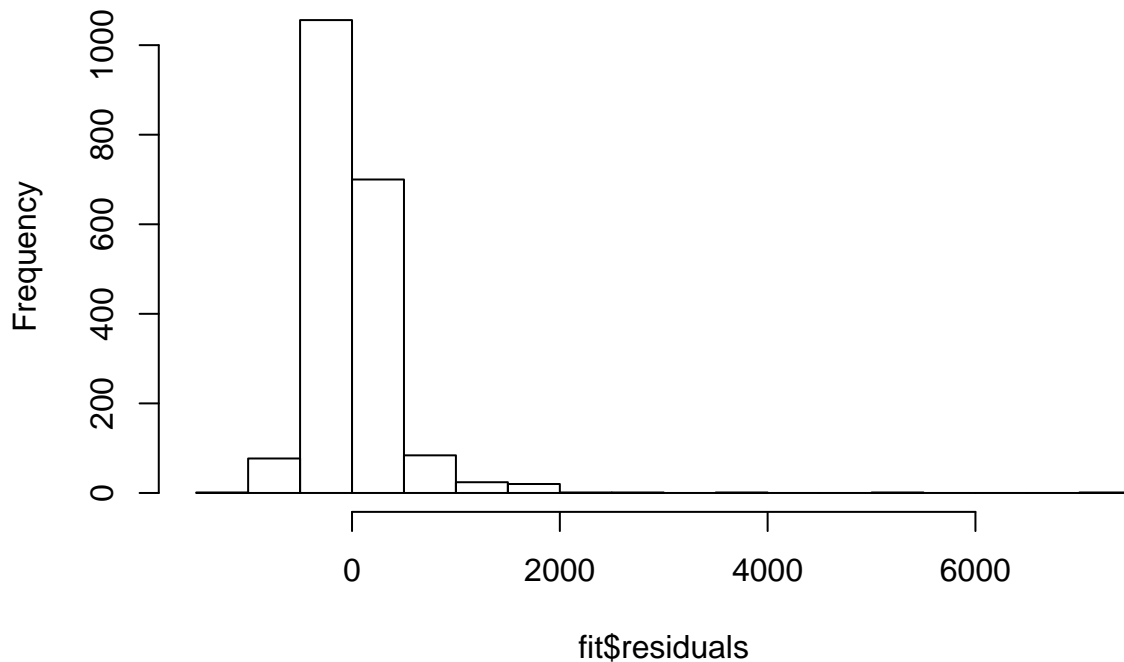
1. Fit a regression model with weekly wages as the response and years of education and experience as predictors. Present the output.

```r
# install.packages('faraway')
library(faraway)
data(uswages)

uswages = uswages[uswages$exper >=0,]

fit = lm(formula= wage ~ educ + exper, data= uswages)
hist(fit$residuals)
```

Histogram of fit$residuals


```r
summary(fit)
```

```
##
## Call:
## lm(formula = wage ~ educ + exper, data = uswages)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -1014.7  -235.2   -52.1    150.1  7249.2
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -239.1146    50.7111  -4.715 2.58e-06 ***
## educ          51.8654     3.3423  15.518  < 2e-16 ***
## exper          9.3287     0.7602  12.271  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 426.8 on 1964 degrees of freedom
## Multiple R-squared:  0.1348, Adjusted R-squared:  0.1339
## F-statistic:    153 on 2 and 1964 DF,  p-value: < 2.2e-16
```

2. What percentage of variation in the response is explained by these predictors? (Percentage variance explained is the same as coefficient of determination).

```
## [1] 0.133912
```

3. Which observation has the largest (positive) residual? Give the case number.

The index number occurs at 15387 and the residual value is 7249.174

4. Compute the mean and median of the residuals. Explain what the difference between the mean and the median indicates.

```
mean(fit$residuals)
```

```
## [1] -1.381535e-15
```

```
median(fit$residuals)
```

```
## [1] -52.14337
```

5. For two people with the same education and one year difference in experience, what would be the difference in predicted weekly wages?

The difference in weekly wages is the coefficient of `exper` which is 9.3287.

6. Compute the correlation of the residuals with the fitted values. Plot residuals against fitted values. Explain the value of this correlation using the geometric (projection) interpretation of least squares.

```
## [1] 6.35678e-17
```

Residuals Plotted against Fitted Values