

jurors

Chris Parrish
January 8, 2016

jurors

references:

- Cannon, et al., Stat2, chapter 03, examples 3.8

Import the data.

```
data <- read.csv("Jurors.csv", header=TRUE)
head(data)
```

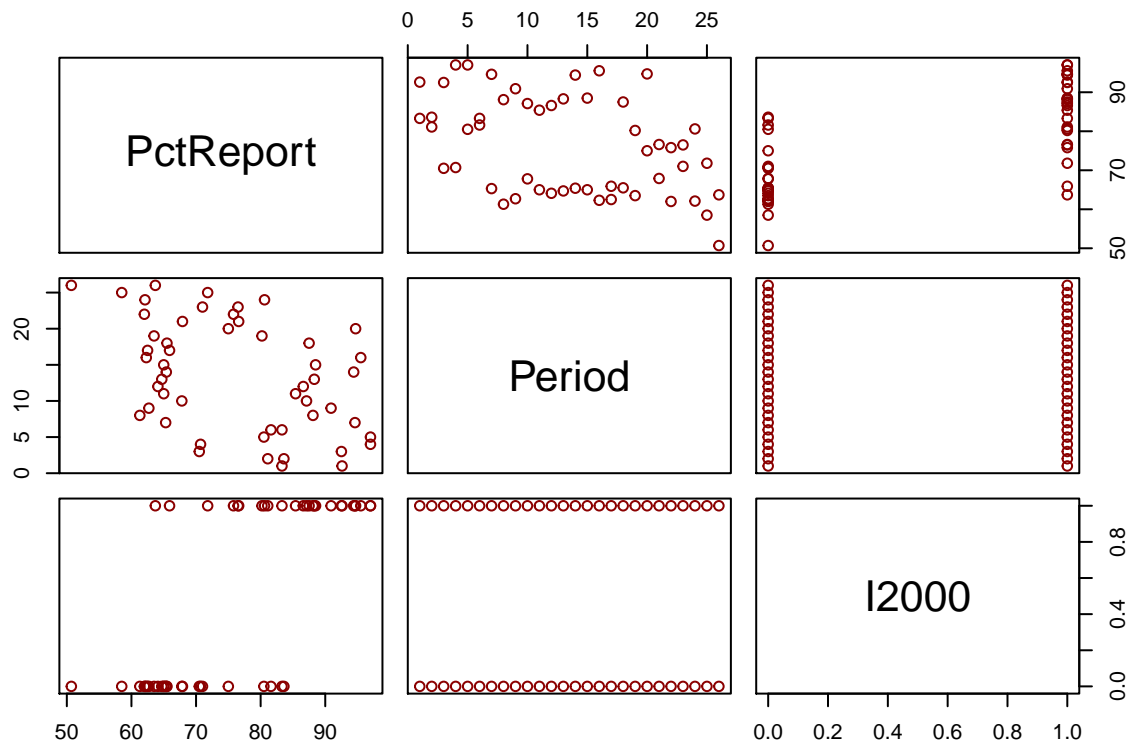
```
##   Period PctReport Year I2000
## 1      1      83.3 1998      0
## 2      2      83.6 1998      0
## 3      3      70.5 1998      0
## 4      4      70.7 1998      0
## 5      5      80.5 1998      0
## 6      6      81.6 1998      0
```

```
dim(data)
```

```
## [1] 52  4
```

Scatterplot matrix.

```
pairs(~ PctReport + Period + I2000, data=data, col="darkred")
```



Separate linear models for each year.

```
jury1998.lm <- lm(PctReport[I2000==0] ~ Period[I2000==0], data=data)
```

$$\widehat{PctReport} = 76.426 + -0.668 \text{ Period}$$

```
coef(jury1998.lm)
```

```
##      (Intercept) Period[I2000 == 0]  
##      76.4255385      -0.6682735
```

```
jury2000.lm <- lm(PctReport[I2000==1] ~ Period[I2000==1], data=data)
```

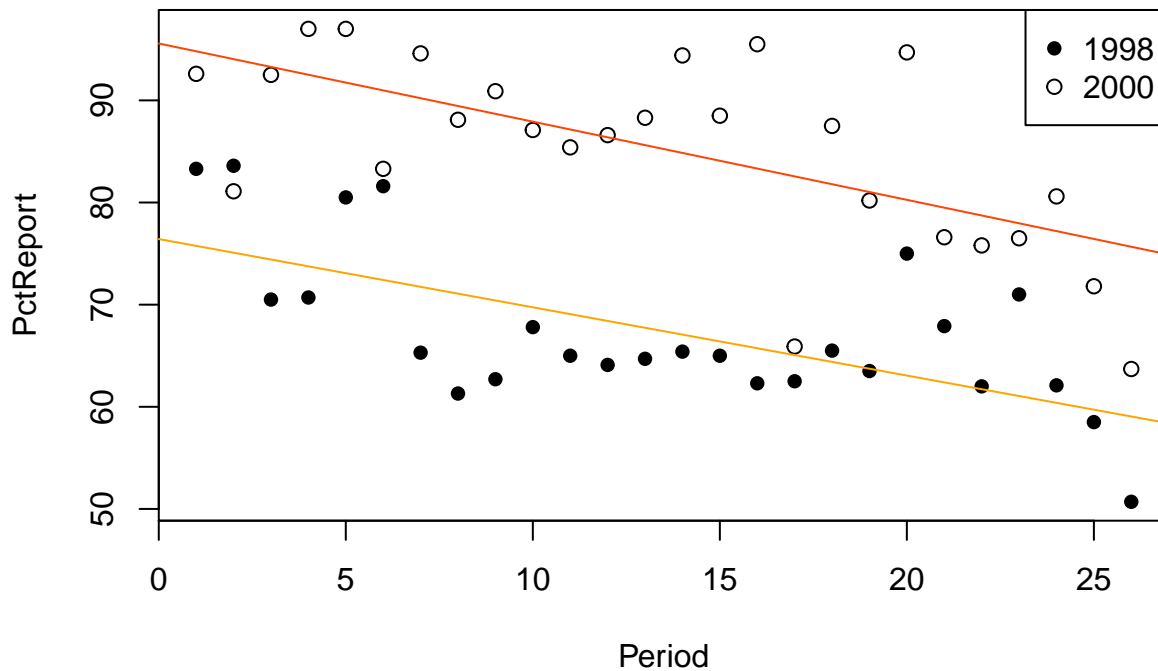
$$\widehat{PctReport} = 95.572 + -0.765 \text{ Period}$$

```
coef(jury2000.lm)
```

```
##      (Intercept) Period[I2000 == 1]  
##      95.5723077      -0.7654701
```

View the data (separate linear models for each year).

```
plot(PctReport ~ Period, data=data,  
     pch=16 - 15 * data$I2000)  
legend("topright", legend=c("1998", "2000"), pch=c(16, 1))  
abline(jury1998.lm, col="orange")  
abline(jury2000.lm, col="orangered")
```



Multiple regression with interaction (allowing different slopes and intercepts).

```
jury.lm1 <- lm(PctReport ~ Period * I2000, data=data)
options(show.signif.stars=FALSE)
summary(jury.lm1)
```

```
##
## Call:
## lm(formula = PctReport ~ Period * I2000, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16.659  -3.148  -1.024   4.430  14.437
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   76.4255     2.7331  27.963 < 2e-16
## Period        -0.6683     0.1770  -3.776 0.000439
## I2000          19.1468     3.8652   4.954 9.44e-06
## Period:I2000  -0.0972     0.2503  -0.388 0.699473
##
## Residual standard error: 6.768 on 48 degrees of freedom
## Multiple R-squared:  0.7197, Adjusted R-squared:  0.7022
## F-statistic: 41.08 on 3 and 48 DF,  p-value: 2.658e-13
```

Multiple regression without interaction (allowing different intercepts but enforcing identical slopes).

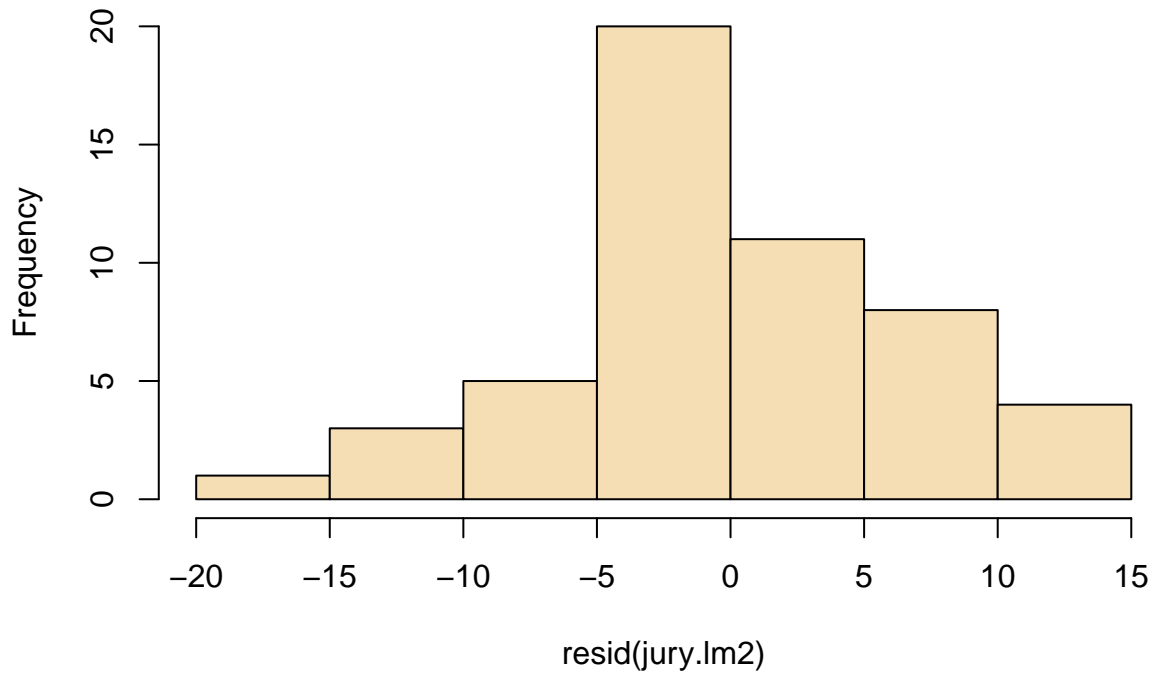
```
jury.lm2 <- lm(PctReport ~ Period + I2000, data=data)
summary(jury.lm2)
```

```
##
## Call:
## lm(formula = PctReport ~ Period + I2000, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16.8294  -3.3873  -0.8705   4.7642  14.1212
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   77.0816     2.1297  36.193 < 2e-16
## Period        -0.7169     0.1241  -5.779 5.12e-07
## I2000          17.8346     1.8608   9.585 8.08e-13
##
## Residual standard error: 6.709 on 49 degrees of freedom
## Multiple R-squared:  0.7188, Adjusted R-squared:  0.7073
## F-statistic: 62.63 on 2 and 49 DF,  p-value: 3.166e-14
```

Residuals.

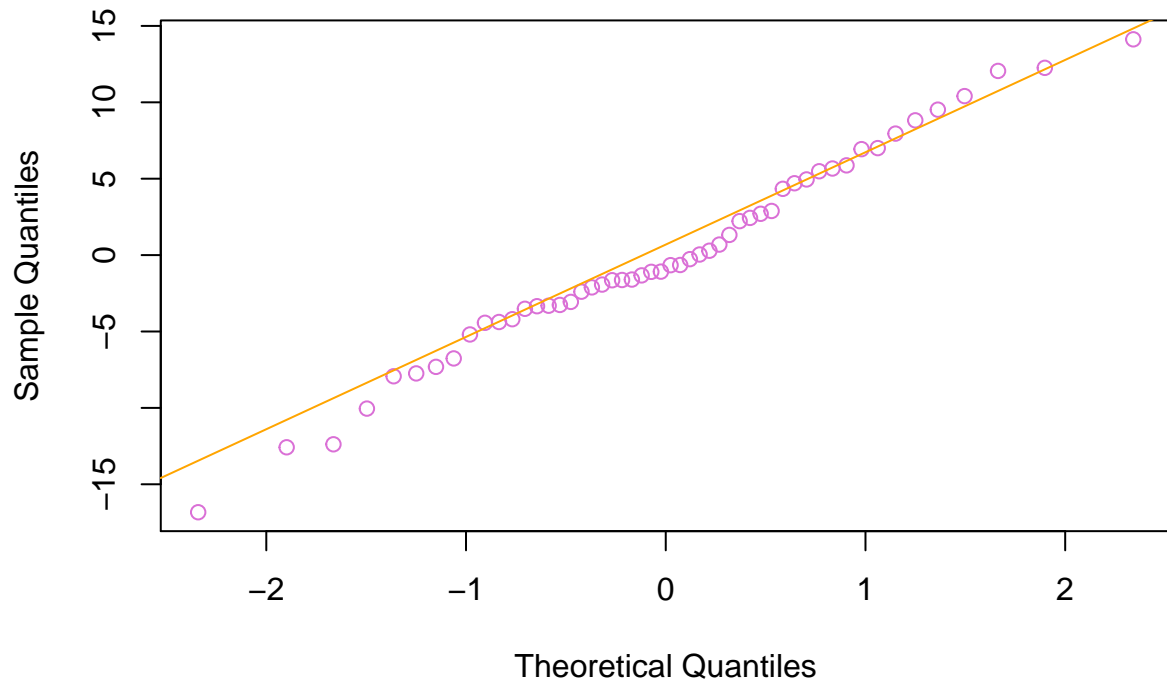
```
hist(resid(jury.lm2),
     col="wheat")
```

Histogram of resid(jury.lm2)

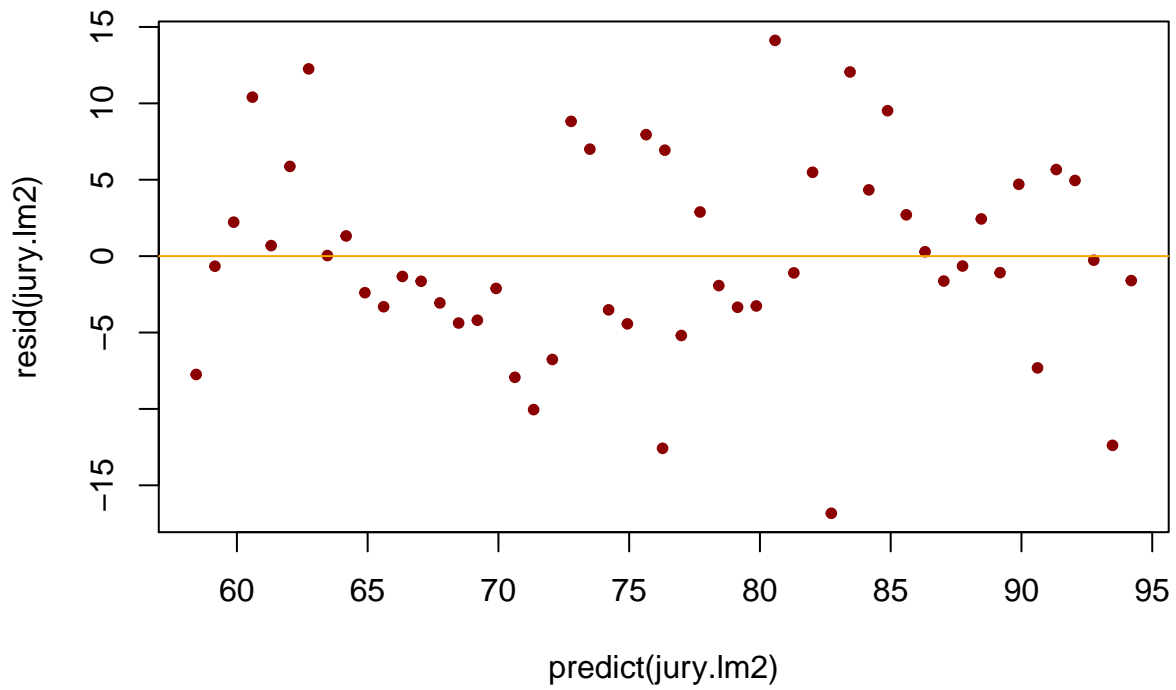


```
qqnorm(resid(jury.lm2),  
       col="orchid")  
qqline(resid(jury.lm2),  
       col="orange")
```

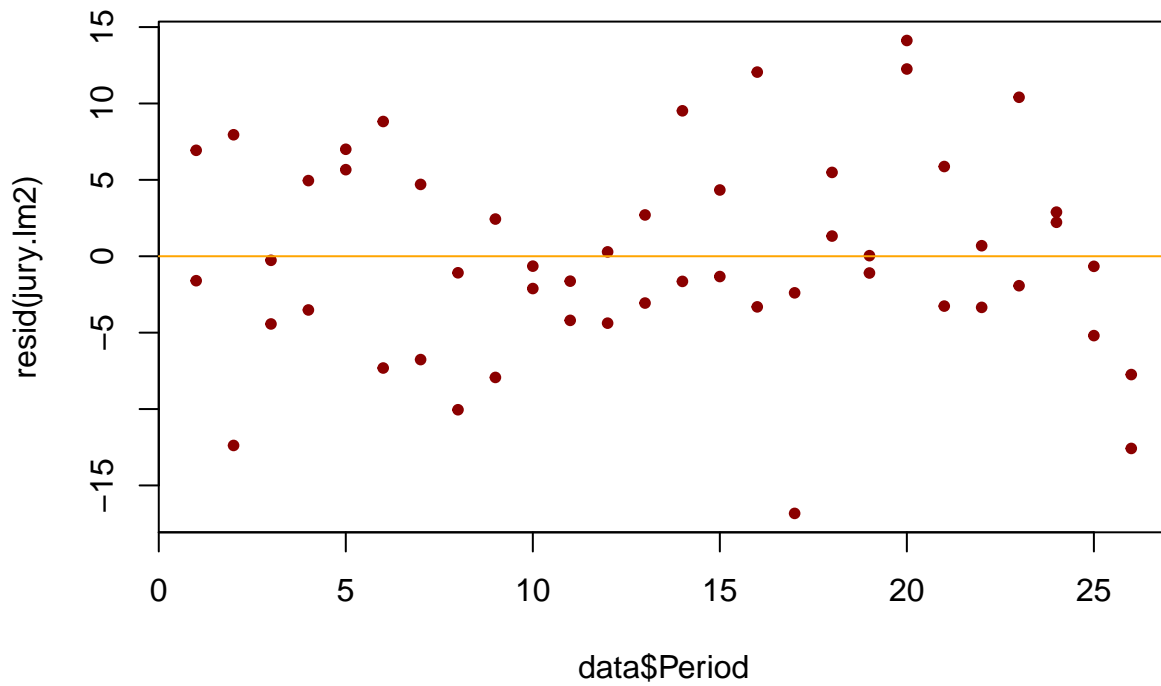
Normal Q-Q Plot



```
plot(predict(jury.lm2), resid(jury.lm2),  
      pch=20, col="darkred")  
abline(h=0, col="orange")
```



```
plot(data$Period, resid(jury.lm2),  
      pch=20, col="darkred")  
abline(h=0, col="orange")
```



CI.

```
confint(jury.lm2)
```

```
##              2.5 %   97.5 %  
## (Intercept) 72.8017366 81.361494  
## Period      -0.9661616 -0.467582  
## I2000       14.0952680 21.573963
```