Project Data Analysis

Congxing Zhu

11/12/2017

```r
library(readxl)
Data3 <- read_excel("/Users/congxingzhu/Desktop/Congxing Project/Data-Appendix 3.xls")
Data3 <- na.omit(Data3)
Data3 <- Data3[-1, ]
row.names(Data3) <- c()
R <- rep(1:5, length = 30)
TR <- rep(1:3, each = 5, length = 30)
D <- rep(c(7, 14), each = 15, length = 30)
N <- rep(20, length = 30)
W <- c(Data3$X__6)
Data4 <- data.frame(Day = D, Tre = TR, Rep = R, Num = N, WFT = W)
Data4$WFT <- as.numeric(as.character(Data4$WFT))
Data4$WFT[Data4$WFT > 20] <- 20
Data6 <- Data4
Data6$Num[10] <- 22
Data6$WFT[10] <- 22
Data6$Num[21] <- 23
Data6$WFT[21] <- 23
Data6
```

```
##     Day Tre Rep Num WFT
## 1     7   1   1  20  17
## 2     7   1   2  20  15
## 3     7   1   3  20  15
## 4     7   1   4  20  18
## 5     7   1   5  20  17
## 6     7   2   1  20  19
## 7     7   2   2  20  18
## 8     7   2   3  20  17
## 9     7   2   4  20  18
## 10    7   2   5  22  22
## 11    7   3   1  20  16
## 12    7   3   2  20  15
## 13    7   3   3  20  19
## 14    7   3   4  20  18
## 15    7   3   5  20  18
## 16   14   1   1  20  18
## 17   14   1   2  20  14
## 18   14   1   3  20  20
## 19   14   1   4  20  17
## 20   14   1   5  20  19
## 21   14   2   1  23  23
## 22   14   2   2  20  20
## 23   14   2   3  20  20
## 24   14   2   4  20  19
## 25   14   2   5  20  18
## 26   14   3   1  20  19
```

```
## 27  14   3   2  20  19
## 28  14   3   3  20  18
## 29  14   3   4  20  17
## 30  14   3   5  20  13
```

```r
m1 <- glm(cbind(WFT, Num - WFT) ~ as.factor(Day) * as.factor(Tre) - 1, family = binomial(link = "logit"),
    data = Data6)
summary(m1)
```

```
##
## Call:
## glm(formula = cbind(WFT, Num - WFT) ~ as.factor(Day) * as.factor(Tre) -
##     1, family = binomial(link = "logit"), data = Data6)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.3556  -0.6769   0.3574   1.0845   2.2613
##
## Coefficients:
##                                 Estimate Std. Error z value Pr(>|z|)
## as.factor(Day)7                   1.5163     0.2603   5.826 5.69e-09 ***
## as.factor(Day)14                  1.9924     0.3077   6.475 9.50e-11 ***
## as.factor(Tre)2                   0.9475     0.4510   2.101   0.0356 *
## as.factor(Tre)3                   0.2989     0.3883   0.770   0.4414
## as.factor(Day)14:as.factor(Tre)2  0.5666     0.8009   0.707   0.4793
## as.factor(Day)14:as.factor(Tre)3 -0.4761     0.5732  -0.831   0.4062
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 464.667  on 30  degrees of freedom
## Residual deviance:  38.821  on 24  degrees of freedom
## AIC: 116.06
##
## Number of Fisher Scoring iterations: 5
```

```r
-2 * logLik(m1)
```

```
## 'log Lik.' 104.0611 (df=6)
```

```r
BIC(m1)
```

```
## [1] 124.4683
```

```r
summary(aov(m1))
```

```
##  Response WFT :
##                            Df Sum Sq Mean Sq  F value  Pr(>F)
## as.factor(Day)              2 9581.3  4790.7 1238.9655 < 2e-16 ***
## as.factor(Tre)              2   35.5    17.7    4.5862 0.02057 *
## as.factor(Day):as.factor(Tre) 2    2.4     1.2    0.3103 0.73609
## Residuals                  24   92.8     3.9
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##  Response 2 :
```

```
##                                    Df  Sum Sq Mean Sq F value    Pr(>F)
## as.factor(Day)                      2 162.733  81.367 26.9724 7.263e-07 ***
## as.factor(Tre)                      2  21.800  10.900  3.6133   0.04249 *
## as.factor(Day):as.factor(Tre)       2   2.067   1.033  0.3425   0.71338
## Residuals                          24  72.400   3.017
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
t.test(WFT ~ Day, data = Data6)
```

```
##
##  Welch Two Sample t-test
##
## data:  WFT by Day
## t = -1.0142, df = 26.106, p-value = 0.3198
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -2.4211045  0.8211045
## sample estimates:
##   mean in group 7 mean in group 14
##          17.46667         18.26667
```

```
m2 <- lm(WFT ~ as.factor(Day) * as.factor(Tre), data = Data6)
summary(m2)
```

```
##
## Call:
## lm(formula = WFT ~ as.factor(Day) * as.factor(Tre), data = Data6)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -4.20  -1.15   0.10   1.25   3.20
##
## Coefficients:
##                                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)                      1.640e+01  8.794e-01  18.649  8.7e-16
## as.factor(Day)14                 1.200e+00  1.244e+00   0.965   0.3442
## as.factor(Tre)2                  2.400e+00  1.244e+00   1.930   0.0655
## as.factor(Tre)3                  8.000e-01  1.244e+00   0.643   0.5261
## as.factor(Day)14:as.factor(Tre)2 -8.600e-17  1.759e+00   0.000   1.0000
## as.factor(Day)14:as.factor(Tre)3 -1.200e+00  1.759e+00  -0.682   0.5016
##
## (Intercept)                      ***
## as.factor(Day)14
## as.factor(Tre)2                    .
## as.factor(Tre)3
## as.factor(Day)14:as.factor(Tre)2
## as.factor(Day)14:as.factor(Tre)3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.966 on 24 degrees of freedom
## Multiple R-squared:  0.315,  Adjusted R-squared:  0.1722
## F-statistic: 2.207 on 5 and 24 DF,  p-value: 0.0869
```

```r
AIC(m1)
```

```
## [1] 116.0611
```

```r
AIC(m2)
```

```
## [1] 133.0138
```

```r
shapiro.test(m1$residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  m1$residuals
## W = 0.92084, p-value = 0.0282
```

```r
shapiro.test(m2$residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  m2$residuals
## W = 0.98075, p-value = 0.8453
```
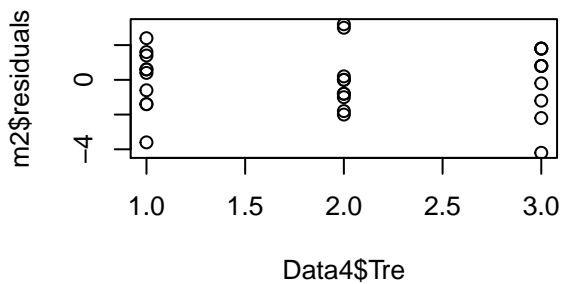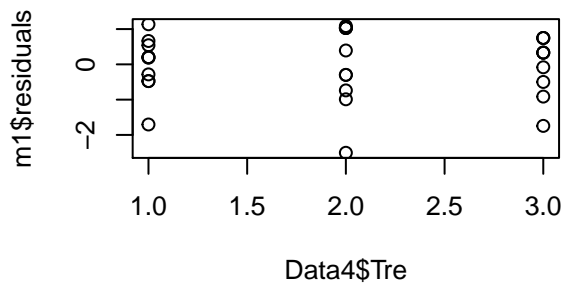
```r
par(mfrow = c(2, 2))
plot(Data4$Day, m1$residuals)
plot(Data4$Day, m2$residuals)
plot(Data4$Tre, m1$residuals)
plot(Data4$Tre, m2$residuals)
```
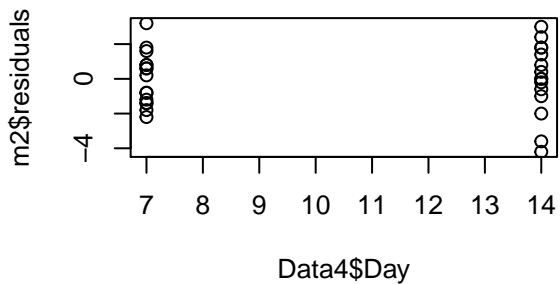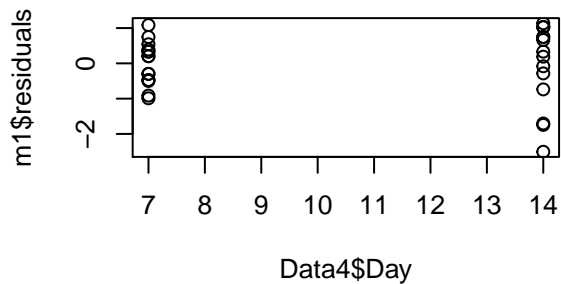
```r
cor(m1$residuals[-30], m1$residuals[-1])
```

```
## [1] 0.002774696
```

```r
cor(m2$residuals[-30], m2$residuals[-1])
```

```
## [1] -0.08376239
```