ST466 Assignment 2

Nishant Kumar

27 March, 2020

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(GGally)
```

```
## Loading required package: ggplot2
```

```
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2
```

```
##
## Attaching package: 'GGally'
```

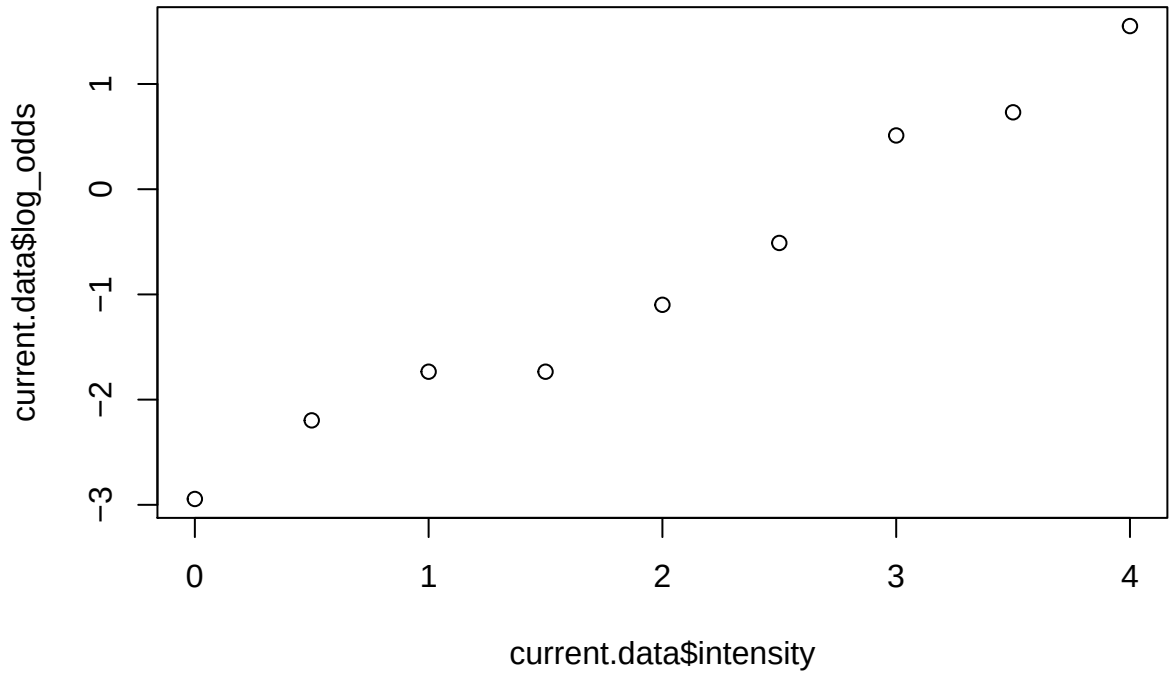
```
## The following object is masked from 'package:dplyr':
##
##     nasa
```

```
intensity <- c(0.0,0.5,1.0,1.5,2.0,2.5,3.0,3.5,4.0)
m <- c(40,40,40,40,40,40,40,40,40)
Y <- c(2,4,6,6,10,15,25,27,33)
current.data <- data.frame(intensity,m,Y)
current.data$p_yes <- current.data$Y/current.data$m
current.data$odds <- current.data$p_yes/(1-current.data$p_yes)
current.data$log_odds <- log(current.data$odds)
current.data$diff <- current.data$m - current.data$Y
current.data
```

```
##   intensity  m  Y p_yes       odds   log_odds diff
## 1       0.0 40  2 0.050 0.05263158 -2.9444390   38
## 2       0.5 40  4 0.100 0.11111111 -2.1972246   36
## 3       1.0 40  6 0.150 0.17647059 -1.7346011   34
## 4       1.5 40  6 0.150 0.17647059 -1.7346011   34
## 5       2.0 40 10 0.250 0.33333333 -1.0986123   30
## 6       2.5 40 15 0.375 0.60000000 -0.5108256   25
## 7       3.0 40 25 0.625 1.66666667  0.5108256   15
## 8       3.5 40 27 0.675 2.07692308  0.7308875   13
```

```
## 9       4.0 40 33 0.825 4.71428571  1.5505974    7
```

```
plot(current.data$log_odds ~ current.data$intensity)
```




```
glm_data <- current.data %>% mutate(failure = m - Y) %>% select(Y,failure)%>% as.matrix()
glm_data
```

```
##        Y failure
##  [1,]  2      38
##  [2,]  4      36
##  [3,]  6      34
##  [4,]  6      34
##  [5,] 10      30
##  [6,] 15      25
##  [7,] 25      15
##  [8,] 27      13
##  [9,] 33       7
```

```
fit_logreg <- glm(glm_data ~ intensity ,data = current.data,family = binomial())
summary(fit_logreg)
```

```
##
## Call:
## glm(formula = glm_data ~ intensity, family = binomial(), data = current.data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.7099  -0.6231   0.2283   0.5643   0.7993
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -3.114      0.333  -9.353   <2e-16 ***
## intensity      1.122      0.125   8.979   <2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 120.6862  on 8  degrees of freedom
## Residual deviance:   2.9689  on 7  degrees of freedom
## AIC: 39.374
##
## Number of Fisher Scoring iterations: 4
```

```
wald_test <- 1.122/0.125
wald_test
```

```
## [1] 8.976
```

```
fit0 <- glm(Y ~ 1, data = current.data)
fit0
```

```
##
## Call:  glm(formula = Y ~ 1, data = current.data)
##
## Coefficients:
## (Intercept)
##       14.22
##
## Degrees of Freedom: 8 Total (i.e. Null);  8 Residual
## Null Deviance:       1040
## Residual Deviance: 1040  AIC: 72.28
```

```
fit1 <- glm(Y ~ intensity, data = current.data)
fit1
```

```
##
## Call:  glm(formula = Y ~ intensity, data = current.data)
##
## Coefficients:
## (Intercept)    intensity
##      -1.778        8.000
##
## Degrees of Freedom: 8 Total (i.e. Null);  7 Residual
## Null Deviance:       1040
## Residual Deviance: 79.56     AIC: 51.15
```

```
current_res <- current.data %>%
mutate(pred_fit0 = predict(fit0, type = "response"),
pred_fit1 = predict(fit1, type = "response"))
current_res
```

```
##   intensity  m  Y p_yes       odds   log_odds diff pred_fit0 pred_fit1
## 1       0.0 40  2 0.050 0.05263158 -2.9444390   38  14.22222 -1.777778
## 2       0.5 40  4 0.100 0.11111111 -2.1972246   36  14.22222  2.222222
## 3       1.0 40  6 0.150 0.17647059 -1.7346011   34  14.22222  6.222222
## 4       1.5 40  6 0.150 0.17647059 -1.7346011   34  14.22222 10.222222
## 5       2.0 40 10 0.250 0.33333333 -1.0986123   30  14.22222 14.222222
## 6       2.5 40 15 0.375 0.60000000 -0.5108256   25  14.22222 18.222222
## 7       3.0 40 25 0.625 1.66666667  0.5108256   15  14.22222 22.222222
## 8       3.5 40 27 0.675 2.07692308  0.7308875   13  14.22222 26.222222
```

```
## 9       4.0 40 33 0.825 4.71428571  1.5505974    7  14.22222 30.222222
```

```
ggplot(current_res, aes(x = intensity, y = Y)) +
geom_point()+
geom_line(aes(x = intensity, pred_fit0, colour = "fit0")) +
geom_line(aes(x = intensity, pred_fit1, colour = "fit1")) +
labs(colour = "")
```