

STAT 308 – Homework 8

For the problems in which calculations are needed, please include your R code with your answers, otherwise you will not be given full credit. Please upload your assignment by Thursday, November 10, 11:59 pm in a pdf file to Sakai.

This homework is meant to help you understand the interpretation behind the parameter estimates for categorical variables in linear regression. For this homework, you will use the `predict` function to obtain predictions from our linear model.

Consider the sleep study dataset uploaded to the course webpage, which contains information on a study where 18 students were deprived of sleep over 10 days, and their reaction times (in ms) to a particular stimulus were recorded. The dataset contains the following variables

- **Case:** The case number
 - **Reaction:** The reaction time (in milliseconds)
 - **Days:** Number of days which the subject was sleep deprived
 - **Subject:** Numeric ID of the subject being tested
1. Suppose we know that reaction time is related linearly with the number of days which the subject was sleep deprived. We wish to add the subject as a categorical variable into our linear model.

```
sleep <- read.csv("../data/sleepstudy.csv")
mod <- lm(Reaction ~ Days + as.character(Subject),sleep)
```

- a. What do we predict the reaction time to be for Subject 308 after 3 days of sleep deprivation? (Hint: use `newdat = data.frame(Days=3,Subject="308")` when using the `predict` function)

```
newdat = data.frame(Days=3,Subject="308")
predict(mod,newdat)
```

```
##          1
## 326.4329
```

326.43 ms

- b. What do we predict the reaction time to be for Subject 310 after 3 days of sleep deprivation?

```
newdat = data.frame(Days=3,Subject="310")
predict(mod,newdat)
```

```
##          1
## 215.3003
```

215.30 ms

- c. Compare the difference between your answers to (a) and (b) with the parameter associated with the indicator variable Subject310.

Subject 310's reaction time on average after 3 days of sleep deprivation is $215.3 - 326.43 = -111.13$ ms different than Subject 308.

- d. Does this relationship hold for any number of days of sleep deprivation?

Yes, changing the number of days does not change the difference in the predictions.

- e. Based on your answers to (c)-(d), how can we interpret the value of the parameter associated with the indicator variable Subject310?

For a given number of days of sleep deprivation, we expect Subject 310's reaction time to be 111.13 ms faster than Subject 308.

- f. What is the interpretation of the value of the parameter associated with the indicator variable Subject350?

For a given number of days of sleep deprivation, we expect Subject 350's reaction time to be 28.53 ms faster than Subject 308.

2. Now, suppose we wish to add an interaction term between subject and number of days of sleep deprivation to our linear model.

- a. What do we predict the reaction time to be for Subject 308 after 0 days of sleep deprivation?

```
mod <- lm(Reaction ~ Days*as.character(Subject),sleep)
newdat = data.frame(Days=0,Subject="308")
predict(mod,newdat)
```

```
##          1
## 244.1927
```

- b. What do we predict the reaction time to be for Subject 310 after 0 days of sleep deprivation?

```
newdat = data.frame(Days=0,Subject="310")
predict(mod,newdat)
```

```
##          1
## 203.4842
```

- c. Compare the difference between your answers to (a) and (b) with the parameter associated with the indicator variable Subject310.

Subject 310's reaction time is, on average $203.48 - 244.19 = -40.71$ ms different from Subject 308 at 0 days of sleep deprivation (i.e. Subject 310's reaction time is 40.71 ms faster on average than subject 308 at 0 days of sleep deprivation.)

- d. Based on your answer to (c), how can we interpret the value of the parameter associated with the indicator variable Subject310?

At 0 days of sleep deprivation, we expect Subject 310's reaction time to be 40.71 ms faster than Subject 308.

- e. What do we predict the increase in reaction time to be for Subject 308 going from 0 to 1 day of sleep deprivation?

```
newdat = data.frame(Days=0,Subject="308")
newdat2 = data.frame(Days=1,Subject="308")
predict(mod,newdat2) - predict(mod,newdat)
```

```
##          1
## 21.7647
```

21.76 ms

- f. What do we predict the increase in reaction time to be for Subject 310 going from 0 to 1 day of sleep deprivation?

```
newdat = data.frame(Days=0,Subject="310")
newdat2 = data.frame(Days=1,Subject="310")
predict(mod,newdat2) - predict(mod,newdat)
```

```
##          1
## 6.114899
```

6.11 ms

- g. Compare the difference between your answers to (e) and (f) with the parameter associated with the interaction between the indicator variable Subject310 and Days of sleep deprivation.

Subject 310's difference in reaction time going from 0 days of sleep deprivation to 1 day of sleep deprivation is $6.11 - 21.76 = -15.65$ ms

- h. Does this relationship hold for going from 3 to 4 days of sleep deprivation? Or 7 to 8 days of sleep deprivation?

Yes, we are fitting linear models to for each subject, so the change in expected reaction time will be the same for a one day increase in sleep deprivation.

- i. Based on your answers to (g) - (h), how can we interpret the value of the parameter associated with the interaction between the indicator variable Subject310 and Days of sleep deprivation?

As sleep deprivation increases by 1 day, we expect Subject 310's reaction time to decrease 15.65 ms slower when compared to Subject 308.