

# Chapter 6 of MD, Homework 4: Due Thursday April 8

Your Name Here

Mar 28, 2021

Note: You are responsible for using the relevant libraries in the global options.

## Question 1

For this problem we use the `birthwt` data frame from the `MASS` library.

- a) Load the data and Use `?birthwt` in the console to look at the description of the data frame.

# Codes here...

- b) We would like to predict the birth weight using the mother's age and mother's race.

Note: Notice that the `race` is coded using 1, 2 and 3. For convenience I have added a new variable called `newRace` that has the races as `white`, `black` and `other`. Also, I created a new dataframe named `newBirthWt` which has the `newRace` variable in it.

Copy and paste the following commands to create the new variable and the data frame. *Use this data frame throughout the rest of the problem.*

```
newBirthWt <- birthwt %>%
  mutate(newRace = ifelse(race==1, "white",
    ifelse(race==2, "black", "other")))

head(newBirthWt)

##      low age lwt race smoke ptl ht ui ftv  bwt newRace
## 85    0  19 182    2     0  0  0  1  0 2523   black
## 86    0  33 155    3     0  0  0  0  3 2551   other
## 87    0  20 105    1     1  0  0  0  1 2557   white
## 88    0  21 108    1     1  0  0  1  2 2594   white
## 89    0  18 107    1     1  0  0  1  0 2600   white
## 91    0  21 124    3     0  0  0  0  0 2622   other
```

- c) Conduct an EDA on the variables that we are interested in.

# Codes here... Hint: EDA consists of 3 parts

- d) As a part of your EDA in part c) you should have created a scatterplot of `age` vs. `bwt` with colors and lines according to your `newRace` variable. Using this graph, decide whether to use a `parallel line model` or an `interaction model`. Explain.

# Codes here...

- e) Build the model you that you decided in part d). Name your model `wtModel`

# Codes here...

- f) Write the model equation (the whole model).

Answer here...

i. Write the model for the **black** mothers:

Answer here...

ii. Write the model for the **white** mothers:

Answer here...

iii. Write the model for the **other** mothers:

Answer here...

g) Create a new scatterplot of **age** vs. **bwt** with colors and lines according to your **newRace** variable using the models you came up with in part f).

*# Codes here...*

h) Create a residual plot and comment about the model.

*# Codes here...*

i) Example: Assume that a white female is going to have a baby at the age of 25. Predict the baby's birth weight in grams. Remove the **#** in the following code.

```
#newdata = data.frame(age=25, newRace = "white") # Define the new data point  
#BabyWt <- predict(wtModel, newdata) # Use the predict function with the model  
#cat("The baby will weigh about ", BabyWt, "grams") # Print the prediction
```

Use the above example to predict the the baby's birth weight in grams for

i) White female who is going to have a baby at the age of 36.

Answer here...

ii) Black female who is going to have a baby at the age of 55.

Answer here...

iii) A female from a different race than white or black who is going to have a baby at the age of 1

Answer here...