STAT 201: Homework 9 (R)

Your name here

2025-11-17

We will examine data that were collected at Baystate Medical Center in Springfield, MA during 1986. We have the birth weights of 189 babies at the time of sampling, along with descriptive information about the mother. Researchers wanted to learn about risk factors for underweight babies. The variables are:

- `low`: indicator of `bwt` less than 2.5 kg
- `age`: mother's age in years
- `lwt`: mother's weight in pounds at last menstrual period
- `race`: mother's race (`1` = white, `2` = black, `3` = other)
- `smoke`: smoking status during pregnancy (no or yes)
- `ptl`: number of previous premature labours
- `ht`: history of hypertension (0 = no, 1 = yes)
- `ui`: presence of uterine irritability (0 = no, 1 = yes)
- `ftv`: number of physician visits during the first trimester
- `bwt`: birth weight in grams

0. Load in the `tidyverse` and `broom` packages below. Then run the code chunk and take a look at the data.

```
library(readr)
birthwt <- read.csv("https://raw.githubusercontent.com/midd-stat201-spring2025/midd-stat201-s
# load other necessary packages here
```

We intend to fit the following linear regression model:

$$bwt = \beta_0 + \beta_1 * lwt + \epsilon$$

1. Assess if the linearity condition is met.

Answer:

2. Provide an example/scenario in context where the independence condition would be violated.

Answer:

3. Using your new data frame, fit a linear regression model for the birth weight of the baby using the mother's weight at last menstrual period as the explanatory variable. Store this as an object called `birth_lm`, then display summary information of the coefficients of `birth_lm` as as a beautiful and tidy table (remember the `kable()` function from awhile go…).

4. Check if the remaining conditions for SLR are satisfied. Briefly explain in words why or why not, using visuals as support as necessary. Be sure to have informative axis labels and title.

Answer:

We will continue assuming all conditions are met.

5. Interpret the estimated slope coefficient.

Answer:

6. Based on the model output, is there evidence of a linear relationship between the mother's weight at last period and the birth weight of the baby? Why or why not?

Answer:

7. Using code, obtain the R^2 of the model and store it as a variable called `r2`. Using in-line code, report the the R^2 of the model and interpret what it means in context.

Answer:

8. Does a significant linear relationship between x and y guarantee that a model is actually useful? Use your answers above to answer this question.

Answer: