Problem 1.

a) Compare the RSS for the linear regression to the cubic regression

When we compare the training residual sum of squares (RSS) for the linear regression to that of the cubic regression, we expect the RSS for the cubic regression to be lower than the RSS for the linear regression. The reason is that since the polynomial regression can be more fitted to the training data, we are expecting that the training RSS for cubic regression would be lower than the training RSS for the linear regression.

b) Answer (a) using test rather than training RSS

When we use test data rather than training data, we expect the test RSS for cubic regression would be higher than the that for the linear regression. Since we are using test data, the polynomial regression has higher chance to be over-fitted, which will cause more error than the linear regression. Thus, when we use test data, the linear regression would have relatively lower test RSS.

c) Suppose that the true relationship between X and Y is not linear. Compare the training RSS for the linear regression and the cubic regression.

Since the true relationship between X and Y is not linear, the training RSS for the cubic regression will be more likely to have lower RSS than the linear regression. Compare to the linear regression model, the polynomial regression model tends to have more flexibility to cover the training data. Therefore, especially when the true relationship between X and Y is not linear, the training RSS for the cubic regression will be more closely fitted to the data and has lower training RSS.

d) Answer (c) using test rather than training RSS.

In (c), it says that we do not know how far it is from linear. Therefore, we do not have enough information to decide which test RSS would be lower or higher. If the true relationship between X and Y is close to the linear relationship, then the test RSS for the linear regression would be relatively lower. On the other hand, if the true relationship is close to the cubic relationship, then, obviously, the test RSS for the cubic regression would be relatively lower than the linear regression.

Problem 2.

(a)

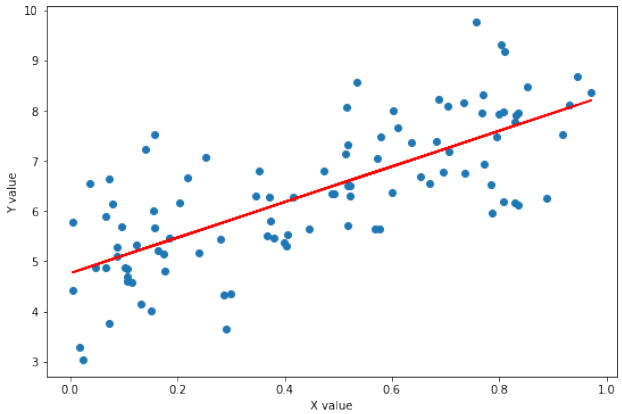

Figure 1: Linear Regression for part(a)

(b)

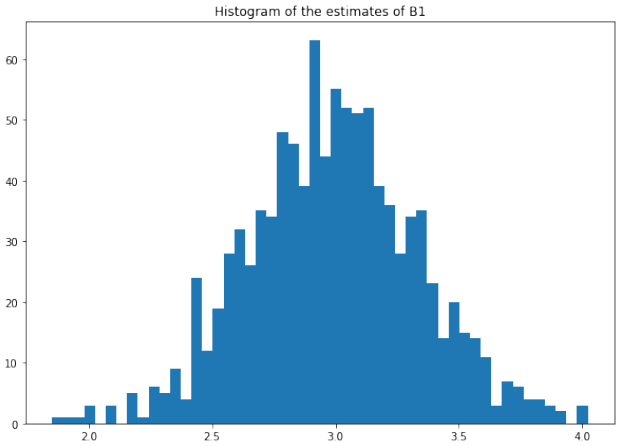

Figure 2: Histogram for part(b)

(c)

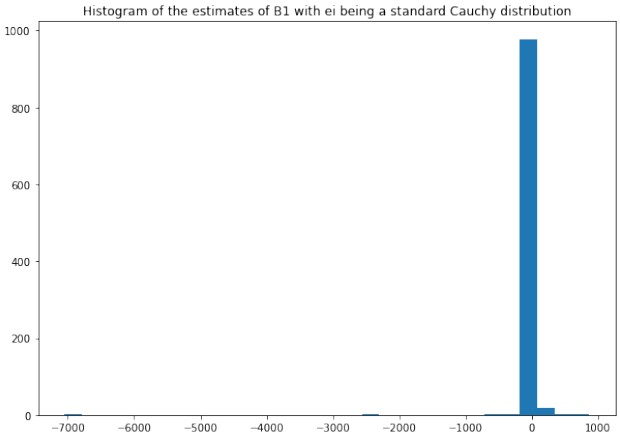

Figure 3: Histogram for part(c)

Here, the graph is not skewed to the left or right and the data is gathered in the center close to the mean.

Problem 3.

a) Estimate the probability that a student who studies for 40h and has an undergrad GPA of 3.5 gets and A in the class

The formula for the logistic regression with two explanatory variables and two categories is

$$p = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2}}{1+e^{\beta_0+\beta_1 x_1+\beta_2 x_2}}$$

Therefore, if we substitute $\hat{\beta}_0 = -6$, $\hat{\beta}_1 = 0.05$, $\hat{\beta}_2 = 1$, x_1 as 40, and x_2 as 3.5, we get the equation

$$P(Y) = \frac{e^{-6+(0.05(40))+3.5}}{1+e^{-6+(0.05(40))+3.5}} = 0.3775$$

Hence, the probability that a student who studies for 40 hours and get 3.5 GPA gets an A in the class is approximately 37.75%.

b) How many hours would the student in part(a) need to study to have a 50% chance of getting an A in the class?

If we substitute the values and solve for x_1 to get 0.5,

$$P(Y) = \frac{e^{-6+(0.05(x_1))+3.5}}{1+e^{-6+(0.05(x_1))+3.5}} = \frac{1}{2}$$
$$=> 2(e^{-6+(0.05(x_1))+3.5}) = 1 + e^{-6+(0.05(x_1))+3.5}$$
$$=> e^{-6+(0.05(x_1))+3.5} = 1$$

By taking ln, we get

$$-6 + 0.05(x_1) + 3.5 = 0$$
$$x_1 = \frac{2.5}{0.05} = 50$$

Therefore, the student in part(a) need to study 50 hours to get a 50% Chance of getting an A in the class.

Problem 4. Bayes Classifier

In the textbook, the Bayes classifier (also called as Bayes optimal classification rule) is referred as if $P(Y = 1|X = x) > \frac{1}{2}$, it corresponds to class one, and otherwise, it corresponds to class two.
Also, since $Y \in 0, 1$, and $P(Y = 1) = \frac{1}{2}$, $P(Y = 0)$ is also $\frac{1}{2}$.
Therefore, when we calculate the probability for each $P(X = x)$ for x = 1, 2, 3, by the Bayesian rule, we get

$$P(X = 1) = P(X = 1|Y = 0) * P(Y = 0) = \frac{1}{3} * \frac{1}{2} = \frac{1}{6}$$
$$P(X = 2) = P(X = 2|Y = 0) * P(Y = 0) = \frac{2}{3} * \frac{1}{2} = \frac{1}{3}$$
$$P(X = 3) = P(X = 3|Y = 1) * P(Y = 1) = \frac{2}{3} * \frac{1}{2} = \frac{1}{3}$$

Now, if we convert this to $Pr(Y = 1|X = x)$,

$$P(Y = 1|X = 1) = \frac{P(X = 1|Y = 1) * P(Y = 1)}{P(X = 1)} = \frac{\frac{2}{3} * \frac{1}{2}}{\frac{1}{6}} = 2$$
$$P(Y = 1|X = 2) = \frac{P(X = 2|Y = 1) * P(Y = 1)}{P(X = 2)} = \frac{\frac{1}{3} * \frac{1}{2}}{\frac{1}{3}} = \frac{1}{2}$$
$$P(Y = 1|X = 3) = \frac{P(X = 3|Y = 1) * P(Y = 1)}{P(X = 3)} = \frac{\frac{2}{3} * \frac{1}{2}}{\frac{1}{3}} = 1$$

Thus, we can say that the Bayes Classifier is class 1.

Pledge:
Please sign below (print full name) after checking (✓) the following. If you can not honestly check each of these responses, please email me at kbala@ucdavis.edu to explain your situation.

- We pledge that we are honest students with academic integrity and we have not cheated on this homework. ✓
- These answers are our own work. ✓
- We did not give any other students assistance on this homework. ✓
- We understand that to submit work that is not our own and pretend that it is our is a violation of the UC Davis code of conduct and will be reported to Student Judicial Affairs. ✓
- We understand that suspected misconduct on this homework will be reported to the Office of Student Support and Judicial Affairs and, if established, will result in disciplinary sanctions up through Dismissal from the University and a grade penalty up to a grade of "F" for the course. ✓

Team Member 1 Sungwon Lee Team Member 2