Assignment - HW2

1. Exercises

# ISLR 2e

**Ex 6:** Using (3.4), argue that in the case of simple linear regression, the least squares line always passes through the point (¯x, y¯).

The expression (3.4) gives us the parameters for a simple linear regression line when minimizing RSS.

𝛽̂ =

𝑛

𝑖𝑖=1

∑

(𝑥𝑖𝑖 − 𝑥̅)(𝑦𝑖𝑖 − 𝑦�)

1 𝑛

∑

𝑖𝑖=1

(𝑥𝑖𝑖 − 𝑥̅)2

The least squares line equation is:

0

1

1

𝑖𝑖

𝛽̂

= 𝑦� − 𝛽̂ 𝑥̅

𝑦𝑖𝑖 = 𝛽̂

0

+ 𝛽̂ 𝑥

In order to proof this, we have to match 𝑥𝑖𝑖 = 𝑥̅ and then substitute 𝛽̂ in the previous equation.

0

1

𝑖𝑖

1

1

𝑦𝑖𝑖 = 𝛽̂

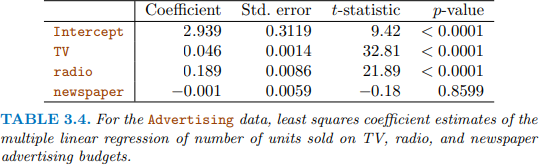
0

+ 𝛽̂ 𝑥

= �𝑦� − 𝛽̂ 𝑥̅� + 𝛽̂ 𝑥̅ = 𝑦�

# Section 3.7

**Ex 1:** Describe the null hypotheses to which the p-values given in Table 3.4 correspond. Explain what conclusions you can draw based on these p-values. Your explanation should be phrased in terms of sales, TV, radio, and newspaper, rather than in terms of the coefficients of the linear model.



1. Intercept

𝐻0: 𝛽0 = 0

𝐻𝑎: 𝛽0 ≠ 0

When x1=x2=x3=0 y takes the value of 𝛽0 . This is the sales we obtain when there are no adds on Tv, radio and in the newspapers. Because the p-value is so low we can reject the null hypothesis. This means that sales are not going to be 0 when there is not adds on Tv, radio and in the newspapers.

1. TV & Radio Hypothesis for TV:

Hypothesis for Radio:

𝐻0: 𝛽1 = 0

𝐻𝑎: 𝛽1 ≠ 0

𝐻0: 𝛽2 = 0

𝐻𝑎: 𝛽2 ≠ 0

The null hypothesis mean that the corresponding variable has no effect on the output when all the other predictors are fixed.

Because the p-value is so low we can reject the null hypothesis. This means that a change on the budget of adds in TV and radio have impact in the result of the sales.

1. Newspaper

Hypothesis for newspapers:

𝐻0: 𝛽3 = 0

𝐻𝑎: 𝛽3 ≠ 0

The null hypothesis mean that the corresponding variable has no effect on the output when all the other predictors are fixed.

Because p = 0.8599 we cannot reject H0. This means that changing the newspaper budget will have no impact on the sales.

**Ex 3:** Suppose we have a data set with five predictors, X1 = GPA, X2 = IQ, X3 = Level (1 for College and 0 for High School), X4 = Interaction between GPA and IQ, and X5 = Interaction between GPA and Level. The response is starting salary after graduation (in thousands of dollars). Suppose we use least squares to fit the model, and get βˆ0 = 50, βˆ1 = 20, βˆ2 = 0.07, βˆ3 = 35, βˆ4 = 0.01, βˆ5 = −10.

1. Which answer is correct, and why?
   1. For a fixed value of IQ and GPA, high school graduates earn more, on average, than college graduates.
   2. For a fixed value of IQ and GPA, college graduates earn more, on average, than high school graduates.
   3. For a fixed value of IQ and GPA, high school graduates earn more, on average, than college graduates provided that the GPA is high enough.
   4. For a fixed value of IQ and GPA, college graduates earn more, on average, than high school graduates provided that the GPA is high enough.

The regression equation is:

𝑠𝑎𝑙𝑎𝑟𝑦 = 50 + 20 ∙ 𝑔𝑝𝑎 + 0.07 ∙ 𝑖𝑖𝑞 + 35 ∙ 𝑙𝑒𝑣𝑒𝑙 + 0.01 ∙ 𝑔𝑝𝑎 ∙ 𝑖𝑖𝑞 − 10 ∙ 𝑔𝑝𝑎 ∙ 𝑙𝑒𝑣𝑒𝑙

Now we look our output when level = 1 (college) and when level is 0 (High School).

𝐸[𝑦 | 𝑙𝑒𝑣𝑒𝑙 = 1] = 50 + 20 ∙ 𝑔𝑝𝑎 + 0.07 ∙ 𝑖𝑖𝑞 + 35 + 0.01 ∙ 𝑔𝑝𝑎 ∙ 𝑖𝑖𝑞 − 10 ∙ 𝑔𝑝𝑎

= 85 + 10 ∙ 𝑔𝑝𝑎 + 0.07 ∙ 𝑖𝑖𝑞 + 0.01 ∙ 𝑔𝑝𝑎 ∙ 𝑖𝑖𝑞

𝐸[𝑦 | 𝑙𝑒𝑣𝑒𝑙 = 0] = 50 + 20 ∙ 𝑔𝑝𝑎 + 0.07 ∙ 𝑖𝑖𝑞 + 0.01 ∙ 𝑔𝑝𝑎 ∙ 𝑖𝑖𝑞

= 50 + 20 ∙ 𝑔𝑝𝑎 + 0.07 ∙ 𝑖𝑖𝑞 + 0.01 ∙ 𝑔𝑝𝑎 ∙ 𝑖𝑖𝑞

The starting salary is higher for college than for high-school when:

𝐸[𝑦 | 𝑙𝑒𝑣𝑒𝑙 = 1] > 𝐸[𝑦 | 𝑙𝑒𝑣𝑒𝑙 = 0]

85 + 10 ∙ 𝑔𝑝𝑎 + 0.07 ∙ 𝑖𝑖𝑞 + 0.01 ∙ 𝑔𝑝𝑎 ∙ 𝑖𝑖𝑞 > 50 + 20 ∙ 𝑔𝑝𝑎 + 0.07 ∙ 𝑖𝑖𝑞 + 0.01 ∙ 𝑔𝑝𝑎 ∙ 𝑖𝑖𝑞

𝑔𝑝𝑎 < 3.5

For fixed IQ and GPA , a salary with level of college is greater than with a level of high- school except when the gpa this last one is greater than 3.5.

This is the same as in the iii option.

1. Predict the salary of a college graduate with IQ of 110 and a GPA of 4.0.

If we put the values in the salary formula, we obtain:

Salary = 137 100 $

1. True or false: Since the coefficient for the GPA/IQ interaction term is very small, there is very little evidence of an interaction effect. Justify your answer

False: The coefficient value for an interaction term between x1 and x2 will depend on the scales of x1, x2 and y.

We may change these coefficients by changing the scale of particular variables and this will not be reflective of whether there is an interaction effect present. So the value of β4 is not informative for this.

1. I collect a set of data (n = 100 observations) containing a single predictor and a quantitative response. I then fit a linear regression model to the data, as well as a separate cubic regression, i.e. Y = β0 + β1X + β2X2 + β3X3 + ϵ.
   1. Suppose that the true relationship between X and Y is linear, i.e. Y = β0 + β1X + ϵ. Consider the training residual sum of squares (RSS) for the linear regression, and also the training RSS for the cubic regression. Would we expect one to be lower than the other, would we expect them to be the same, or is there not enough information to tell? Justify your answer.

In this case I expect similar results. But with the cubic regression with a lower training RSS. We are increasing the flexibility of the model going from linear to cubic and we are not working with enough data. So I would expect the model the cubic model to overfit the non-linearities and therefore have lower training RSS.