ISLR Chapter 3 Exercises

Anthony Chau

Last compiled: May 30 2020

Contents

Conceptual Exercises **2**

Question 1 . . . 2
Question 2 . . . 3
Question 3 . . . 4
Question 4 . . . 5
Question 5 . . . 6
Question 6 . . . 7
Question 7 . . . 8
Question 8 . . . 9
Question 9 . . . 10
Question 10 . . . 11

Conceptual Exercises

Question 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.

Solution:

The null hypotheses to which these p-values correspond to is if the true regression coefficients are zero. Note that there is nothing said about a specified significance levels. We will not state that some coefficients are "significant" vs "non-significant". However, the regression results suggest that for for every $1000 increase in television, radio, and newspaper advertising, there will be an average change in sales by 46, 189, and -1 units, respectively.

Question 2

Carefully explain the differences between the KNN classifier and KNN regression methods.

Solution:

KNN regression is a regression method used to estimate a functional form for a given dataset. It does this by identifying K training observations that are closest to x_0 (represented by N_0) and then estimates $f(x_0)$ by the following formula: $f(x_0) = \frac{1}{K} * \sum_{x_i \in N_0} y_i$. That is, KNN regressions averages the response values from all the points in N_0.

KNN classification is a classification algorithm used to classify observations into different classes within the dataset. The algorithm assigns the observation to the class corresponding to the most common class of its neighboring points.

Question 3

Suppose we have a data set with five predictors, X1 = GPA, X2 = IQ, X3 = Gender (1 for Female and 0 for Male), X4 = Interaction between GPA and IQ, and X5 = Interaction between GPA and Gender. The response is starting salary after graduation (in thousands of dollars). Suppose we use least squares to fit the model, and get

$$\begin{aligned}\beta_0 &= 50\\ \beta_1 &= 20\\ \beta_2 &= 0.07\\ \beta_3 &= 35\\ \beta_4 &= 0.01\\ \beta_5 &= -10\end{aligned}$$

Solution:

We write out the equation:

$$Salary = 50 + 20 \cdot GPA + 0.07 \cdot IQ + 35 \cdot Gender + 0.01 \cdot (GPA \cdot IQ) - 10 \cdot (GPA \cdot Gender)$$

a)

To answer this question, we assume that IQ and GPA is fixed. Then, we write the equation for males and for females

Equation for males given fixed IQ and GPA:

$$Salary = 50 + 20 \cdot GPA + 0.07 \cdot IQ + 0.01 \cdot (GPA \cdot IQ)$$

Equation for females given fixed IQ and GPA:

$$\begin{aligned}Salary &= 50 + 20 \cdot GPA + 0.07 \cdot IQ + 35 + 0.01 \cdot (GPA \cdot IQ) - 10 \cdot GPA\\ &= 85 + 20 \cdot GPA + 0.07 \cdot IQ + 0.01 \cdot (GPA \cdot IQ) - 10 \cdot GPA\end{aligned}$$

Question 4

Question 5

Question 6

Question 7

Question 8

Question 9

Question 10