Chapter 3 HW

Fabiani Rafael

Conceptual Questions

Exercise 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.?

- The null hypothesis is that the variables TV, radio and newspaper have no effect on sales. From the table the p-values corresponding to TV and radio are less than 0.0001 indicating that the null hypothesis can be rejected. The p-value for newspaper is 0.8599 which is greater than 0.05, so the null hypothesis cannot be rejected. This would draw us to conclude that TV and radio have a significant effect on sales, while newspaper does not.

Exercise 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

$$\hat{\beta}_0 = 50, \ \hat{\beta}_1 = 20, \ \hat{\beta}_2 = 0.07, \ \hat{\beta}_3 = 35, \ \hat{\beta}_4 = 0.01, \ \hat{\beta}_5 = -10$$

.?

(a) Which answer is correct, and why?

   i. For a fixed value of IQ and GPA, high school graduates earn more, on average, than college graduates.

   ii. For a fixed value of IQ and GPA, college graduates earn more, on average, than high school graduates.

   iii. 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.

   iv. 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.

-From the problem :

$$X = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ \text{GPA}_1 \\ \text{IQ}_1 \\ \text{Level}_1 \\ (\text{GPA} \times \text{IQ})_1 \\ (\text{GPA} \times \text{Level})_1 \end{bmatrix}, \ \hat{\beta} = \begin{bmatrix} 50 \\ 20 \\ 0.07 \\ 35 \\ 0.01 \\ -10 \end{bmatrix}$$

Now Y is given by

$$Y = X^T \hat{\beta}$$

$$= \begin{bmatrix} 1 & \text{GPA} & \text{IQ} & \text{Level} & \text{GPA} \times \text{IQ} & \text{GPA} \times \text{Level} \end{bmatrix} \begin{bmatrix} 50 \\ 20 \\ 0.07 \\ 35 \\ 0.01 \\ -10 \end{bmatrix}$$

$$= 50 + 20(\text{GPA}) + 0.07(\text{IQ}) + 35(\text{Level}) + 0.01(\text{GPA} \times \text{IQ}) - 10(\text{GPA} \times \text{Level})$$

For fixed value of IQ, GPA, and letting Level $= 1$ for College, 0 for High School. the salary of a college graduate is given by

$$Y = 50 + 20(\text{GPA}_1) + 0.07(\text{IQ}_1) + 35(\text{Level}_1) + 0.01(\text{GPA} \times \text{IQ})_1 - 10(\text{GPA} \times \text{Level})_1$$
$$= 50 + 20(\text{GPA}_1) + 0.07(\text{IQ}_1) + 35(1) + 0.01(\text{GPA} \times \text{IQ})_1 - 10(\text{GPA} \times \text{Level})_1$$

and for high school graduates we have

$$Y = 50 + 20(\text{GPA}_1) + 0.07(\text{IQ}_1) + 35(0) + 0.01(\text{GPA} \times \text{IQ})_1 - 10(\text{GPA} \times \text{Level})_1$$
$$= 50 + 20(\text{GPA}_1) + 0.07(\text{IQ}_1) + 0.01(\text{GPA} \times \text{IQ})_1 - 10(\text{GPA} \times \text{Level})_1$$

We can then equate these equations to find

$$50 + 20(\text{GPA}_1) + 0.07(\text{IQ}_1) + 35 + 0.01(\text{GPA} \times \text{IQ})_1 - 10(\text{GPA} \times \text{Level})_1 = 50 + 20(\text{GPA}_1) + 0$$
$$50 - 50 + 20(\text{GPA}_1) - 20(\text{GPA}_1) + 0.07(\text{IQ}_1) - 0.07(\text{IQ}_1) + 35 + 0.01(\text{GPA} \times \text{IQ})_1 - 0.01(\text{GPA} \times \text{IQ})_1 = 10(\text{GPA} \times \text{Level})_1$$
$$35 = 10(\text{GPA} \times \text{Level})_1$$
$$3.5 = (\text{GPA} \times \text{Level})_1$$

(b) Predict the salary of a college graduate with IQ of 110 and a GPA of 4.0.

(c) 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.


Applied Questions

Exercise 10 : This question should be answered using the Carseats data set.

(a) Fit a multiple regression model to predict Sales using Price, Urban, and US.

(b) Provide an interpretation of each coefficient in the model. Be careful—some of the variables in the model are qualitative!

(c) Write out the model in equation form, being careful to handle the qualitative variables properly.

(d) For which of the predictors can you reject the null hypothesis H0 : $\beta j = 0$?

(e) On the basis of your response to the previous question, fit a smaller model that only uses the predictors for which there is evidence of association with the outcome.

(f) How well do the models in (a) and (e) fit the data?

(g) Using the model from (e), obtain 95 % confidence intervals for the coefficient(s).

(h) Is there evidence of outliers or high leverage observations in the model from (e)?

```r
summary(cars)
```

```
##      speed           dist
##  Min.   : 4.0   Min.   :  2.00
##  1st Qu.:12.0   1st Qu.: 26.00
##  Median :15.0   Median : 36.00
##  Mean   :15.4   Mean   : 42.98
##  3rd Qu.:19.0   3rd Qu.: 56.00
##  Max.   :25.0   Max.   :120.00
```

```r
plot(cars)
```
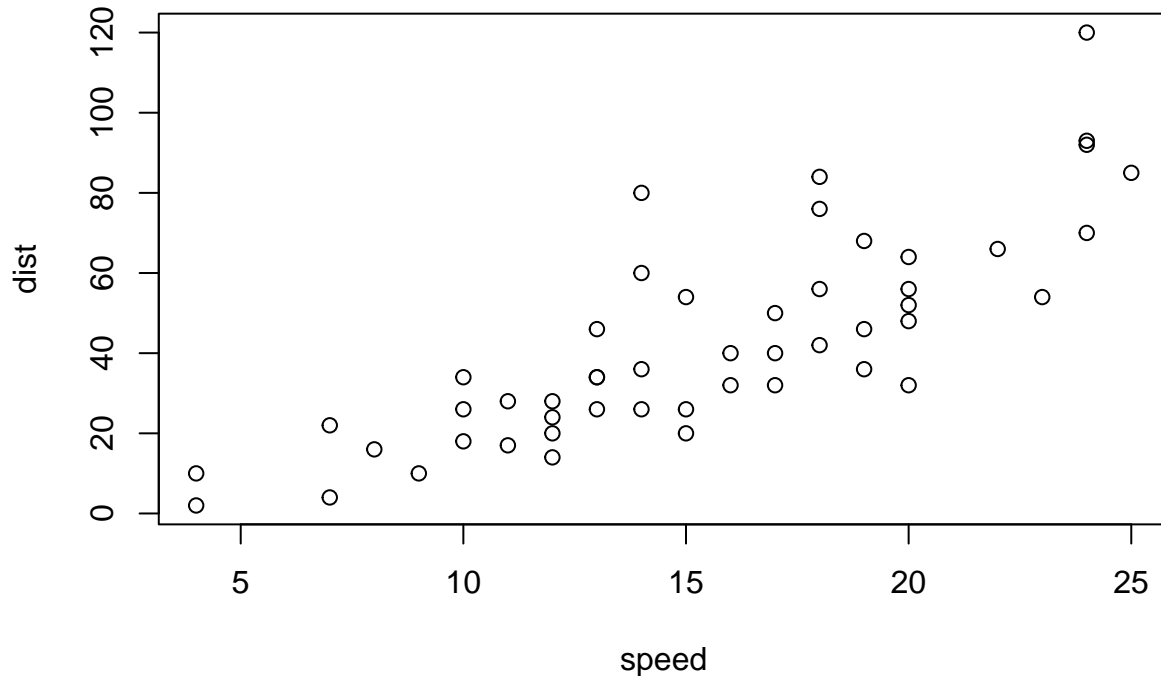

Exercise 13: In this exercise you will create some simulated data and will fit simple linear regression models to it. Make sure to use set.seed(1) prior to starting part (a) to ensure consistent results.

(a) Using the rnorm() function, create a vector, x, containing 100 observations drawn from a N (0, 1) distribution. This represents a feature, X.

(b) Using the rnorm() function, create a vector, eps, containing 100 observations drawn from a N (0, 0.25) distribution-a normal distribution with mean zero and variance 0.25.

(c) Using x and eps, generate a vector y according to the model

$$Y = -1 + 0.5X + \epsilon.$$

What is the length of the vector y? What are the values of

$$\beta_0$$

and

$$\beta_1$$

in this linear model?

(d) Create a scatterplot displaying the relationship between x and y. Comment on what you observe.

(e) Fit a least squares linear model to predict y using x. Comment on the model obtained. How do $\hat{\beta}_0$ and $\hat{\beta}_1$ compare to $\hat{\beta}_0$ and $\hat{\beta}_1$?

(f) Display the least squares line on the scatterplot obtained in (d). Draw the population regression line on the plot, in a different color. Use the legend() command to create an appropriate legend.

(g) Now fit a polynomial regression model that predicts y using x and x2 . Is there evidence that the quadratic term improves the model fit? Explain your answer.

(h) Repeat (a)-(f) after modifying the data generation process in such a way that there is less noise in the data. The model (3.39) should remain the same. You can do this by decreasing the variance of the normal distribution used to generate the error term ϵ in (b). Describe your results.

(i) Repeat (a)-(f) after modifying the data generation process in such a way that there is more noise in the data. The model (3.39) should remain the same. You can do this by increasing the variance of the normal distribution used to generate the error term ϵ in (b). Describe your results.

(j) What are the confidence intervals for $\hat{\beta}_0$ and $\hat{\beta}_1$ based on the original data set, the noisier data set, and the less noisy data set? Comment on your results.