Textbook Exercises: Chapters 16 and 17

Due on the 8th of May, 2025.

The owner of a furniture shop recorded the monthly advertising expenses (in thousand dollars) and sales revenues (in million dollars) of the last year. Let X and Y denote the advertising expense and sales revenue, respectively. The owner is now interested in predicting Y using X. Assume $Y = \alpha + \beta\sqrt{X} + \varepsilon$.

	Jan.	Feb.	Mar.	Apr.	May	Jan.
advertising expense	23.1	46.5	60.4	54.0	28.2	33.9
sales revenue	9.9	11.3	11.7	11.0	10.6	10.9
	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
advertising expense	24.7	31.5	36.2	87.8	90.1	99.3
sales revenue	10.2	10.8	10.4	12.4	12.0	12.5

(a) Estimate α and β using the method of ordinary least squares. (7 marks)

(b) Suppose the advertising expense for some future month is 70,000 dollars. Estimate with 90% confidence the expected sales revenue. (10 marks)

(c) Draw a scatter plot of the recorded data. Focus on $(X, Y) \in [0, 120] \times [8, 13]$. Include the fitted line. Also, visualise for any given future month the 90% confidence interval of the expected sales revenue. (7 marks)

(d) What your textbook calls standardised residuals is alternatively known as internally Studentised residuals. Calculate the internally Studentised residual for the month of January. (4 marks)

(e) Exclude the January data for now.

i. Re-answer Part (a): Estimate the coefficients. (4 marks)

ii. Calculate the standard error of the regression. (4 marks)

iii. Substitute your answer to Part ii. for its counterpart in your solution to Part (d), other things being equal. Re-calculate the Studentised residual for the month of January. (2 marks)

iv. What you have just calculated is called the externally Studentised residual for the month of January. Compare it with your answer to Part (d). Which is more appropriate, the internally or externally Studentised residual, in detecting outliers? Briefly explain. (4 marks)

(f) Let $r_{\text{in},\,i}$ and $r_{\text{ex},\,i}$ respectively denote the internally and externally Studentised residuals for the ith observation. Let n and p respectively denote the sample size and

the number of free parameters in a linear model, assuming the ordinary least-square method. Statisticians have already found,

$$r_{\text{ex},\,i} = r_{\text{in},\,i} \cdot \sqrt{\frac{n-p-1}{n-p-r_{\text{in},\,i}^2}}.$$

i. Calculate the internally Studentised residuals for all the other 11 months. (22 marks)

ii. Apply the given formula. Calculate the externally Studentised residuals for all the other 11 months. (11 marks)

iii. Draw a scatter plot of the externally Studentised residuals against the monthly advertising expenses. Tell from the plot whether or not the residuals appear to have a bell-shaped distribution. Briefly explain. (4 marks)

(g) Remember we are analysing times series data. Time series are often autocorrelated.

i. Define first-order autocorrelation. (4 marks)

ii. Determine with 90% confidence, using the Durbin−Watson test, whether or not the residuals are first-order autocorrelated. (6 marks)

iii. Suppose the residuals are indeed first-order autocorrelated. Do they appear to be positively or negatively so? How do you tell? (3 marks)

iv. Suppose you are testing whether or not $\beta > 0$. Following Part iii., what is the effect, if any, of the supposed autocorrelation, in the direction it appears, on the result of your test? Please do show your reasoning. (8 marks)