Lecture 5

Chapters 1.5-1.6

1.5 Results from OLS

Review from yesterday:

1. Simple linear regression model –

2. Fitted regression model –

3. Normal equations for $\widehat{\beta_1}$ and $\widehat{\beta_2}$ –

Exercise 1: Show that the mean value of the residuals is zero.

Exercise 2: Show that the sample correlation coefficient between X and e is zero.

Note that we can also show that $\rho_{e,\hat{Y}} = 0$.

1.6 Goodness of Fit: R^2

Often, we want to know how good a job the OLS estimates do at fitting Y.

Total sum of squares – sum of the squared deviations about the sample mean of Y. (TSS)

Formula – TSS

Exercise 3: Deconstruct the TSS formula into its explained and unexplained parts.

Explained sum of squares – sum of squared deviations of fitted Y about its sample mean. (ESS)

Coefficient of determination – the proportion of the total sum of squares that is explained by the regression line, R^2.

Formula – R^2

Properties of R^2

1. It always lies between 0 and 1.
2. When it is 1, RSS is 0.
3. When it is 0, ESS is 0, and TSS=RSS.

Exercise 4: Rewrite the R^2 formula in terms of the RSS.

Exercise 5: Find the R^2 given the following information.

Observation	X	Y	$\hat{Y}$
1	1	4	2.9
2	2	3	4.3
3	3	5	5.7
4	4	8	7.1
Mean	2.5	5	5

Exercise 6: If the fitted regression line does a good job, then the fitted values of Y should be highly correlated with the true values of Y. Show that $r_{Y,\hat{Y}} = \sqrt{R^2}$