Tutorial 3

Applied Stats/Quant Methods 1

Week 3, Fall 2021

1 Estimating Regression Coefficients "by hand"

Based on the values given in Table 1, find the estimates $\hat{\alpha}$, $\hat{\beta}_1$, and $\hat{\beta}_2$ of α, β_1, and β_2 in the following two population regression models:

$$y_i = \alpha + \beta_1 x_{1i}$$
$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i}$$

Do not use computer software to run a regression. Rather, estimate $\hat{\alpha}$, $\hat{\beta}_1$ and $\hat{\beta}_2$ "by hand", i.e., by plugging the relevant values into the estimator functions. Does the model provide a "reasonable" linear approximation to the data?

Table 1: Data for exercise 3

i	x_{1i}	x_{2i}	y_i
1	0.55	17.85	–21.84
2	35.18	23.25	79.13
3	31.75	19.92	77.11
4	29.20	9.36	92.36
5	–17.21	19.01	–94.33
6	7.06	10.07	18.48
7	15.26	8.68	52.34
8	20.47	18.10	41.44
9	22.17	8.07	79.28
10	11.89	15.40	10.11

2 OLS Regression Formula

Show that the formula for the LS estimator of the slope, which we saw in class as:

$$\hat{\beta} = \frac{n\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2}$$

Can be rewritten as:

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

3 Mechanics of OLS

A linear regression model of average income (y, log scale) conditional on prestige (x) for 102 professions yields the following quantities:

$$\bar{y} = 8.66 \qquad \Sigma_{i=1}^{n}(y_i - \bar{y})^2 = 35.33$$

$$\bar{x} = 46.83 \qquad \Sigma_{i=1}^{n}(x_i - \bar{x})^2 = 29895.43$$

$$\Sigma_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = 761.61$$

Compute the LS estimates and variances of β_0 and β_1 in the model $y = \beta_0 + \beta_1 x + \epsilon$.

4 Confidence Intervals of Coefficient Estimates

Load the `car` library. Construct 87% confidence intervals for all coefficient estimates of the following least-squares regression analyses (name of the dataset in parenthesis):

(a) `tfr` $\sim$ `contraceptors` (`Robey.dat`)

(b) `education` $\sim$ `income` $+$ `young` $+$ `urban` (`Anscombe.dat`)