PO12Q: Introduction to Quantitative Political Analysis II

Week 4 - Worksheet Solutions

Dr Florian Linke
F.Linke@warwick.ac.uk

1. Plot the data in Table in a suitable scatter plot. Yes, on paper.

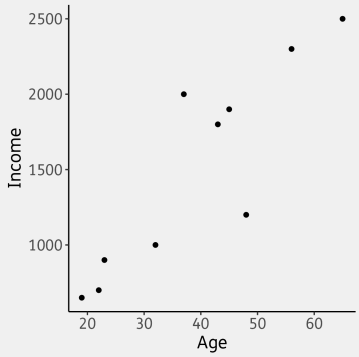

Figure 1: Scatterplot

2. Fit a line of best fit through the scatter plot (by eyeballing and a ruler).

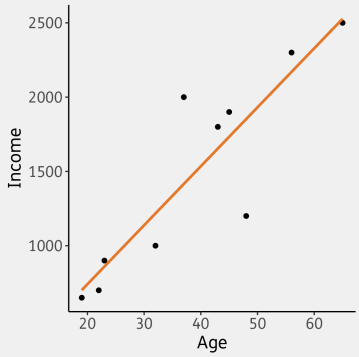

Figure 2: Regression Line

3. Assuming a regression model of the type $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$, calculate the estimators for β_0 and β_1.

For intermediate calculations see Excel Sheet.

$$\hat{\beta}_1 = \frac{83200}{2096} = 39.69465$$

$$\hat{\beta}_0 = 1495 - \hat{\beta}_1 \times 39 = -53.0916$$

4. Calculate the regression coefficients β_0 and β_1 using matrices.

Recall that

$$(\boldsymbol{X'X})^{-1} = \frac{1}{n\Sigma x_i^2 - \Sigma x_i \Sigma x_i} \begin{bmatrix} \Sigma x_i^2 & -\Sigma x_i \\ -\Sigma x_i & n \end{bmatrix}$$

and that

$$(\boldsymbol{X'Y}) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{bmatrix} \times \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \Sigma y_i \\ \Sigma x_i y_i \end{bmatrix}$$

Into this we plug the values from our Excel sheet:

$$(\boldsymbol{X'X})^{-1} = \frac{1}{20960} \begin{bmatrix} 17306 & -390 \\ -390 & 10 \end{bmatrix}$$

$$= \begin{bmatrix} 0.8256679 & -0.01860687 \\ -0.01860687 & 0.0004770992 \end{bmatrix}$$

$$(\boldsymbol{X'Y}) = \begin{bmatrix} 14950 \\ 666250 \end{bmatrix}$$

Which leads us to

$$\hat{\beta} = (\boldsymbol{X'X})^{-1}(\boldsymbol{X'Y}) = \begin{bmatrix} 0.8256679 & -0.01860687 \\ -0.01860687 & 0.0004770992 \end{bmatrix} \times \begin{bmatrix} 14950 \\ 666250 \end{bmatrix}$$

$$= \begin{bmatrix} -53.092 \\ 39.6946 \end{bmatrix}$$

5. Build the SRF and interpret the estimators of β_0 and β_1.

$$\widehat{\text{income}}_i = -53.0916 + 39.69465 \text{ age}_i$$

- **Intercept**: At age zero, a person would earn -53.09 units of income on average
- **Slope**: For every additional year of age, a person's income would increase by 39.69 units on average