Handout for Lecture 17

Inference in Regression Models

ECON 340: Economic Research Methods   Instructor: Div Bhagia

We are interested in the following relationship:

$$CrimeRate = \beta_0 + \beta_1 UnemploymentRate + u$$

We run a regression of *Crime Rate* on *Unemployment Rate* across US counties. The output from this regression is presented below. [1]

	Crime Rate
Intercept	0.003***
	(0.0002)
Unemployment Rate	0.05***
	(0.004)
Observations	2,957
R^2	0.05

Note: *p<0.1; **p<0.05; ***p<0.01

1. Interpret the slope and intercept coefficient.

   The intercept coefficient is 0.003, which means that when the unemployment rate is zero, the predicted crime rate is 0.003. The slope coefficient for the unemployment rate is 0.05, indicating that for each percentage point increase in the unemployment rate, the predicted crime rate increases by 0.05.

---

[1]Crime rate is defined as the ratio of crimes in a specific area to the population of that area, expressed as incidents per 1,000 people per year. The unemployment rate represents the number of unemployed people as a percentage of the labor force. The average crime rate in the data is 0.0057, while the average unemployment rate is 0.05.

2. Interpret the R^2.

   The R^2 value of 0.05 implies that 5% of the variability in crime rates across counties can be explained by unemployment rate.

3. According to this model, what is the predicted Crime Rate for a county with an unemployment rate of 0.06 (6 percentage points)?

   We can find the predicted crime rate for a county with an unemployment rate of 0.06, as follows

   $$0.006 + 0.05 \times 0.06 = 0.009$$

4. What is the t-statistic associated with the hypothesis test that examines whether the coefficient on the unemployment rate is zero ($H_0 : \beta_1 = 0$)?

   $$t = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)} = \frac{0.05}{0.004} = 12.5$$

   Here, $\hat{\beta}_1$ is the standard error of $\hat{\beta}_1$, which is presented in parenthesis below the coefficient.

5. What is the smallest level of significance at which you would reject the hypothesis in question 4?

   The coefficient on the unemployment rate is marked with ***, which here indicates that the p-value is less than 0.001. Therefore, we can reject the null at any significance level greater than 0.001. So here we can say that the impact of unemployment on crime is statistically significant at the 1% level of significance.

   Technically, the smallest level of significance corresponds to the exact p-value, which is very close to zero in this case, as $Pr(|Z| > 12.5)$ is also near zero.

6. Construct a 95% confidence interval for β_1. (Note: $Pr(|Z| > 1.96) = 0.95$.) What is the interpretation of this interval?

   The 95% confidence interval for β_1 can be calculated as follows:

   $$\hat{\beta}_1 \pm 1.96 \times SE(\hat{\beta}_1) = 0.05 \pm 1.96 \times 0.004 = [0.042, 0.058]$$

   We are 95% confident that the true slope coefficient β_1 lies in this interval.

7. Can we attach a causal interpretation to the estimates here? Why or why not?

   No, because the exogeneity assumption may be violated here, which is key to causal inference. Exogeneity assumes that omitted factors captured in u are independent of X. However, it's possible that factors like educational attainment could be correlated with unemployment (X) and directly impact the crime rate, thereby violating the exogeneity assumption.