1. Explain what is:

i) the probability of type I error; ii) the 95% confidence interval; iii) the p value; iv) the probability of type II error; v) the power of a test.

Answer:

i) It is the probability of rejecting the null when the null is true.

iv) It is the probability of failing to reject the null when the null is false.

v) It is one minus the probability of type II error.

2. Solve exercise 3.11 in Johnston and Dinardo:

The following regression equation is estimated as a production function for Q:

$$
\begin{aligned}
\ln(Q) &= 1.37 + \underset{(0.257)}{0.632}\ln(K) + \underset{(0.219)}{0.452}\ln(L), \\
R^2 &= 0.98,\ \mathrm{Cov}(b_k, b_l) = 0.055,
\end{aligned}
$$

where the standard errors are given in parentheses. Test the following hypotheses:

i) The capital and labor elasticities of output are identical.

ii) There are constant returns to scale.

i) The capital elasticity, that is the elasticity of output, Q, with respect to capital, K, is given by

$$e_K^Q = \frac{d\,\mathsf{ln}(Q)}{d\,\mathsf{ln}(K)} = b_k,$$

The labour elasticity, that is the elasticity of output, Q, with respect to labour, L, is given by

$$e_L^Q = \frac{d\,\mathsf{ln}(Q)}{d\,\mathsf{ln}(L)} = b_l.$$

Hence, to test the hypothesis that $e_K^Q = e_L^Q$ is equivalent to test the hypothesis that $b_k = b_l$.

The t-test is given by

$$t = \frac{(\beta_k - \beta_l)}{SE(\beta_k - \beta_l)}.$$

The standard error in the dominator is given by:

$$\begin{aligned} SE(\beta_k - \beta_l) &= \sqrt{\text{Var}(\beta_k - \beta_l)} \\ &= \sqrt{\text{Var}(\beta_k) + \text{Var}(\beta_l) - 2\text{Cov}(b_k, b_l)} \\ &\sqrt{0.257^2 + 0.219^2 - 2 \times 0.055.} \end{aligned}$$

Thus

$$t = \frac{0.632 - 0.452}{\sqrt{0.257^2 + 0.219^2 - 2 \times 0.055}} = 2.846.$$

In order to get the critical values we need the sample size. For a large sample size the 5% critical value is 1.96. Thus we reject the null hypothesis.

ii) Similarly, to test the hypothesis that there are constant returns to scale is equivalent to test the hypothesis that $b_k + b_l = 1$. The t-test is given by

$$\begin{aligned} t &= \frac{(\beta_k + \beta_l) - 1}{SE(\beta_k + \beta_l)} = \\ &= \frac{(\beta_k + \beta_l) - 1}{\sqrt{\text{Var}(\beta_k) + \text{Var}(\beta_l) + 2\text{Cov}(b_k, b_l)}} = \\ &= \frac{(0.632 + 0.452) - 1}{\sqrt{0.257^2 + 0.219^2 + 2 \times 0.055}} = 0.178. \end{aligned}$$

Thus, we fail to reject the null hypothesis.

3. Solve exercise 3.19 in Johnston and Dinardo:

An economist is studying the variation in fatalities from road traffic in different states. She hypothesizes that the fatality rate depends on the average speed and the standard deviation of speed in each state. No data are available on the standard deviation, but in a normal distribution the standard deviation can be approximated by the difference between the 85th percentile and the 50th percentile. Thus, the specified model is

$$Y = b_1 + b_2X_2 + b_3(X_3 - X_2) + u,$$

where Y =fatality rate; X_2 =average speed; X_3 =85th percentile speed.

Instead of regressing Y on X_2 and $(X_3 - X_2)$ as in the specified model, the research assistant fits

$$Y = a_1 + a_2X_2 + a_3X_3 + u,$$

with the following results:

$$\widehat{Y} = \text{constant} - 0.24X_2 + 0.20X_3,$$

with $R^2 = 0.62$. The (absolute) t statistics for the two slope coefficients are 1.8 and 2.3, respectively, and the covariance of the regression slopes is 0.003. Use these regression results to calculate a point estimate of b_2 and test the hypothesis that average speed has no effect on fatalities.

We rewrite the initial regression as

$$Y = b_1 + (b_2 - b_3)X_2 + b_3X_3 + u.$$

This is the regression the research assistant estimated.

$$Y = a_1 + a_2X_2 + a_3X_3 + u.$$

Comparing the coefficients, we have

$$\beta_2 = \alpha_2 + \beta_3 = \alpha_2 + \alpha_3 = -0.04.$$

To carry out the test we need an estimate of the variance of β_2:

$$\begin{aligned}
\text{Var}(\beta_2) &= \text{Var}(\alpha_2 + \alpha_3) \\
&= \text{Var}(\alpha_2) + \text{Var}(\alpha_3) + 2\text{Cov}(\alpha_2, \alpha_3).
\end{aligned}$$

The variance of α_2 is given by

$$\begin{aligned} t_2 &= \frac{\alpha_2}{SE(\alpha_2)} \Rightarrow \\ SE(\alpha_2) &= \frac{\alpha_2}{t_2} \Rightarrow \\ \text{Var}(\alpha_2) &= (\frac{\alpha_2}{t_2})^2 = (\frac{-0.24}{1.8})^2. \end{aligned}$$

Similarly,

$$\text{Var}(\alpha_3) = (\frac{\alpha_3}{t_3})^2 = (\frac{0.20}{2.3})^2.$$

Thus the variance of β_2 is

$$\text{Var}(\beta_2) = (\frac{-0.24}{1.8})^2 + (\frac{0.20}{2.3})^2 + 2 \times 0.003 = 0.0313.$$

The test statistic for the significance of b_2 is then

$$t = \frac{\beta_2}{SE(\beta_2)} = \frac{-0.04}{\sqrt{0.0313}} = -0.226.$$

This is not significant at conventional significance levels and for any sample size.

We can not reject the hypothesis that average speed has no effect on fatalities.

ONLY FOR EC5501

4. Solve exercise 3.13 in Johnston and Dinardo:

One aspect of the rational expectation hypothesis involves the claim that expectations are unbiased, that is, that the average prediction is equal to the observed realization of the variable under investigation. This claim can be tested by reference to announced predictions and to actual values of the rate of interest on three-month U.S. Treasury Bills published in the *Goldsmith-Nagan Bond and Money Market Letter*. The results of least-squares estimation (based on 30 quarterly observations) of the regression of the actual on the predicted interest rates were as follows:

$$r_t = \underset{(0.86)}{0.24} + \underset{(0.14)}{0.94}\, r_t^* + e_t, \; RSS = 28.56,$$

where r_t is the observed interest rate, and r_t^* is the average expectation of r_t held at the end of the preceding quarter. Figures in parentheses are estimated standard errors. The sample data on r^* give

$$\sum r_t^*/30 = 10, \ \sum(r_t^* - \overline{r}_t^*)^2 = 52.$$

Carry out the test assuming that all basic assumptions of the classical regression model are satisfied.

Under the rational hypothesis, the regression

$$r_t = b_1 + b_2 r_t^* + e_t,$$

should yield $b_1 = 0$ and $b_2 = 1$.

We can write these two hypotheses in a matrix form as

$$Rb = r,$$

where

$$R = I, \ b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \ r = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Thus we can carry out this join test by the $F-$test

$$(R\beta - r)'[s^2 q R(X'X)^{-1}R']^{-1}(R\beta - r) \sim F(2, n-k).$$

Since $R = I$, that is R is an identity matrix, we have

$$\begin{aligned}(\beta - r)'[s^2 q(X'X)^{-1}]^{-1}(\beta - r) &\sim F(2, n-k) \text{ or} \\ \frac{(\beta - r)'(X'X)(\beta - r)}{s^2 q} &\sim F(2, n-k).\end{aligned}$$

Next, $\beta - r$ is given by

$$\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.24 \\ 0.94 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.24 \\ -0.06 \end{bmatrix}.$$

Moreover, $X'X$ is given by

$$\begin{bmatrix} N & \sum r_t^* \\ \sum r_t^* & \sum (r_t^*)^2 \end{bmatrix}.$$

We know that:

$$\sum r_t^*/30 = 10, \ \ \sum (r_t^* - \overline{r}_t^*)^2 = 52.$$

It follows that

$$\begin{aligned} \sum r_t^* &= 300, \ N = 30, \\ \sum (r_t^* - \overline{r}_t^*)^2 &= \sum (r_t^*)^2 - N(\overline{r}_t^*)^2 \Longrightarrow \\ \sum (r_t^*)^2 &= \sum (r_t^* - \overline{r}_t^*)^2 + N(\overline{r}_t^*)^2 = \\ 52 + 30 \times 100 &= 3052. \end{aligned}$$

In addition,

$$s^2 = \frac{RSS}{N-k} = \frac{28.56}{28}.$$

Thus

$$
\begin{aligned}
\frac{(\beta - r)'(X'X)(\beta - r)}{s^2 q} &= \\
\frac{\begin{bmatrix} 0.24 \\ -0.06 \end{bmatrix}' \begin{bmatrix} 30 & 300 \\ 300 & 3052 \end{bmatrix} \begin{bmatrix} 0.24 \\ -0.06 \end{bmatrix}}{28.56/28 \times 2} &= \\
&= 1.998.
\end{aligned}
$$

The 5% critical value of $F(2, 28)$ is 3.34.

So we fail to reject the null hypothesis.