

**ECO 5720**  
**Problem Set 3**

1. Which of the following can cause the usual OLS  $t$  statistics to be invalid (that is, not to have  $t$  distributions under  $H_0$ )?

- Heteroskedasticity.
- A sample correlation coefficient of .95 between two independent variables that are in the model.
- Omitting an important explanatory variable.

Answer: Since the  $t$  statistic is a ratio of the form  $\hat{\beta}_j/se(\hat{\beta}_j)$ , both heteroskedasticity and omitted variables can render it invalid.

2. Are rent rates influenced by the student population in a college town? Let *rent* be the average monthly rent paid on rental units in a college town in the United States. Let *pop* denote the total city population, *avginc* the average city income, and *pctstu* the student population as a percentage of the total population. One model to test for a relationship is

$$\log(\text{rent}) = \beta_0 + \beta_1 \log(\text{pop}) + \beta_2 \log(\text{avginc}) + \beta_3 \text{pctstu} + u.$$

(i) State the null hypothesis that size of the student body relative to the population has no ceteris paribus effect on monthly rents. State the alternative that there is an effect.

Answer:  $H_0: \beta_3 = 0$  and  $H_1: \beta_3 \neq 0$ .

(ii) State the null hypothesis that the elasticity of *rent* with respect to *pop* (i.e., the percentage change in *rent* for a one percent change in *pop*) is 1. State the alternative that the elasticity is not 1.

Answer:  $H_0: \beta_1 = 1$  and  $H_1: \beta_1 \neq 1$ .

(iii) State the null hypothesis that the elasticity of *rent* with respect to *pop* is the same as the elasticity of *rent* with respect to *avginc*.

Answer:  $H_0: \beta_1 = \beta_2$ .

(iv) State the null hypothesis that both the elasticities of *rent* with respect to *pop* and *avginc* are jointly equal to zero. State the corresponding alternative.

Answer:  $H_0: \beta_1 = 0$  and  $\beta_2 = 0$ .  $H_1$ : At least  $\beta_1 \neq 0$  or  $\beta_2 \neq 0$ .

3. Consider the estimated equation below studying the effects of skipping class on college GPA:

$$\widehat{\text{colGPA}} = 1.39 + 0.412\text{hsGPA} + 0.015\text{ACT} - 0.083\text{skipped}$$

|                         |         |         |         |
|-------------------------|---------|---------|---------|
| (0.33)                  | (0.094) | (0.011) | (0.026) |
| $n = 141, R^2 = 0.234.$ |         |         |         |

Here, *colGPA* and *hsGPA* depict college and high school GPA, respectively; *skipped* depicts the average number of lectures missed per week; and *ACT* denotes the standardized college admission test's score. The standard errors are in parentheses.

(i) Using the standard normal approximation, find the 95% confidence interval for  $\beta_{\text{hsGPA}}$ .

Answer: Standard normal approximation for 95% CI implies the two-tailed critical value for 5% level of significance from Table G.2 with  $df = \infty$ . The interval is  $0.412 \pm 1.96(0.094)$ , or about 0.228 to 0.596.

(ii) Calculate the p-value for the hypothesis  $H_0: \beta_{hsGPA} = 0$  against the two-sided alternative.

Answer: The test statistic is given by  $0.412/0.094$ , i.e., 4.38. Now, since  $n-k-1$  is equal to 137, we can assume the test statistic to follow a standard normal distribution. The p-value is equal to 2 x (area to the right of 4.38 under a standard normal curve). From Table G.1, the p-value is less than 0.001.

(iii) Given the confidence interval in (i), can you reject the hypothesis  $H_0: \beta_{hsGPA} = 0.4$  against the two-sided alternative at the 5% level?

Answer: No, because the value 0.4 is inside the 95% CI.

Alternatively, the test statistic is  $(0.412 - 0.4)/0.094 = 0.128$ . This is inside the interval  $[-1.96, 1.96]$ .

Alternatively, while the test statistic is  $(0.412 - 0.4)/0.094 = 0.13$ , the p-value from Table G.1 is roughly  $2(1 - 0.5517) = 0.9$ . This is well above 0.05.

(iv) Dropping *hsGPA* and *ACT* from the equation gives

$$\widehat{colGPA} = 3.15 - 0.090skipped$$

$$(0.433) \quad (0.028)$$

$$n = 141, R^2 = 0.069.$$

Are *hsGPA* and *ACT* jointly significant in the original equation at the 5% level? Justify your answer by calculating the F statistic and comparing against the critical value.

Answer: Using the  $R^2$  version of the formula for the F statistic, we have

$$F = \frac{(0.234 - 0.069)/2}{(1 - 0.234)/(141 - 3 - 1)} = 14.76.$$

From Table G.3b, the critical value for 5% level of significance with numerator  $df$  as 2 and denominator  $df$  as  $\infty$  is 3. Accordingly, we reject  $H_0$ , i.e., find *hsGPA* and *ACT* to be jointly significant.

4. Use the data in MLB1 and estimate the following model:

$$\log(\text{salary}) = \beta_0 + \beta_1 \text{years} + \beta_2 \text{gamesyr} + \beta_3 \text{bavg} + \beta_4 \text{hrunsyr} + \beta_5 \text{rbisyr} + u.$$

(i) Test the joint significance of *hrunsyr* and *rbisyr*. Please paste your R script.

Answer:

#####

```
library(wooldridge)
library(car)
library(stargazer)
data(mlb1)
View(mlb1)
```

```
modell1 <- lm(lsalary ~ years + gamesyr + bavg + hrunsyr + rbisyr, data=mlb1)
```

```
# Summary of regression results
summary(modell1)
stargazer(modell1, type="text")
```

```
=====
                        Dependent variable:
                        -----
                        lsalary
-----
years                    0.069***
                        (0.012)

gamesyr                  0.013***
                        (0.003)

bavg                     0.001
                        (0.001)

hrunsyr                  0.014
                        (0.016)

rbisyr                   0.011
                        (0.007)

Constant                 11.192***
                        (0.289)

-----
Observations              353
R2                        0.628
Adjusted R2               0.622
Residual Std. Error      0.727 (df = 347)
F Statistic              117.060*** (df = 5; 347)
=====
Note:                    *p<0.1; **p<0.05; ***p<0.01
```

```
# Joint test that the coefficients for hrunsyr and rbisyr are both zero #
linearHypothesis(modell1, c("hrunsyr = 0", "rbisyr = 0"))
```

```
Linear hypothesis test:
hrunsyr = 0
rbisyr = 0
```

```
Model 1: restricted model
Model 2: lsalary ~ years + gamesyr + bavg + hrunsyr + rbisyr
```

```
   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
1     349 197.43
2     347 183.19  2     14.244 13.491 2.279e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The variables *hrunsyr* and *rbisyr* are jointly significant.

(ii) State the null hypothesis that *hrunsyr* and *rbisyr* have the same effect on  $\log(\text{wage})$ . Test this null hypothesis against a two-sided alternative at the 5% significance level. Please paste your R script.

Answer:  $H_0: \beta_4 = \beta_5$ .

```
# Test that the coefficients for hrunsyr and rbisyr are equal *  
linearHypothesis(modell, "hrunsyr - rbisyr = 0")
```

```
Linear hypothesis test:  
hrunsyr - rbisyr = 0
```

```
Model 1: restricted model
```

```
Model 2: lsalary ~ years + gamesyr + bavg + hrunsyr + rbisyr
```

|   | Res.Df | RSS    | Df | Sum of Sq | F      | Pr(>F) |
|---|--------|--------|----|-----------|--------|--------|
| 1 | 348    | 183.20 |    |           |        |        |
| 2 | 347    | 183.19 | 1  | 0.013758  | 0.0261 | 0.8718 |

We fail to reject  $H_0$ .

5. In the article “Statistical Significance, p-Values, and the Reporting of Uncertainty” what does the term p-hacking on page 158 imply?

Answer: p-hacking refers to the practice whereby researchers choose to report specifications on the basis of p-values less than 0.05.

6. Please visit the link: [Understanding Statistical Power and Significance Testing — an Interactive Visualization](#). Under settings, for two-tailed tests, solve for power when  $\alpha = 0.05$ ,  $n = 100$ , and  $d = 1$ . How does power change when  $\alpha = 0.05$ ,  $n = 100$ , and  $d = 0.1$ ? What happens to power when  $\alpha = 0.05$ ,  $n = 10$ , and  $d = 1$ ?

Answer: When  $\alpha = 0.05$ ,  $n = 100$ , and  $d = 1$ , power = 100%. Next, when  $\alpha = 0.05$ ,  $n = 100$ , and  $d = 0.1$ , power = 17%. Finally, when  $\alpha = 0.05$ ,  $n = 10$ , and  $d = 1$ , power = 89%.