BS2280 – Econometrics I
Homework 4: Hypothesis testing

1

A researcher hypothesizes that years of schooling, S, may be related to the number of siblings (brothers and sisters), SIBLINGS, according to the relationship

$$S = \beta_1 + \beta_2 SIBLINGS + u$$

She is prepared to test the null hypothesis $H0 : \beta_2 = 0$ against the alternative hypothesis $H1 : \beta_2 \neq 0$ at the 5 percent and 1 percent levels. She has a sample of 60 observations. The critical t values at the 5 percent and 1 percent significance level are 2.00 and 2.66 respectively. Undertake hypothesis tests for the following scenarios:

1. $\hat{\beta}_2 = -0.20$, s.e.$(\hat{\beta}_2) = 0.07$
2. $\hat{\beta}_2 = -0.12$, s.e.$(\hat{\beta}_2) = 0.07$
3. $\hat{\beta}_2 = 0.06$, s.e.$(\hat{\beta}_2) = 0.07$
4. $\hat{\beta}_2 = 0.20$, s.e.$(\hat{\beta}_2) = 0.07$

2

The number of cigarettes smoked per day is regressed on the price of cigarettes per pack in USD. The results are presented in the R output below.

1. Interpret the intercept and the coefficient of the independent variable.
2. Write down the test hypotheses for testing the significance of the intercept and coefficient.
3. Calculate t-statistics for the intercept and the coefficient of cigarette prices. The critical t-value at the 5% significance level is 1.96.

```
Call:
lm(formula = cigs ~ cigpric, data = smoke)

Residuals:
   Min     1Q Median     3Q    Max
-9.224 -8.678 -8.575 11.082 71.332

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 10.67457    6.17296
cigpric     -0.03297    0.10206
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 13.73 on 805 degrees of freedom
Multiple R-squared:  0.0001296, Adjusted R-squared:  -0.001112
F-statistic:        on 1 and 805 DF,  p-value:
```

3

Calculate the 95% confidence interval for the intercept as well as the coefficient of cigarette prices using the R output in Question 2.

4

Calculate the F statistic for the regression undertaken in Question 2 using ESS and RSS presented in the R anova output table above. Check that the F statistic derived from R^2 is the same. Perform the F test, whereby the critical F-value at the 5% significance level is approximately 3.8415.

The formula for the F-statistics are

$$F(1, 805) = \frac{ESS/(k-1)}{RSS/(n-k)}$$

or

$$F(1, 805) = \frac{R^2/(k-1)}{(1-R^2)/(n-k)}$$

```
Analysis of Variance Table

Response: cigs
            Df Sum Sq Mean Sq F value Pr(>F)
cigpric      1     20  19.672
Residuals  805 151734 188.489
```

5

The number of cigarettes per day is regressed on the age of participants. Use the R output tables below to answer Questions 2 – 4 again.

```
Call:
lm(formula = cigs ~ age, data = smoke)

Residuals:
   Min     1Q Median     3Q    Max
-9.498 -8.929 -7.991 10.669 71.372

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 10.06698    1.26597
age         -0.03348    0.02838
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 13.72 on 805 degrees of freedom
Multiple R-squared:  0.001726,  Adjusted R-squared:  0.0004856
F-statistic:      on 1 and 805 DF,  p-value:


Analysis of Variance Table

Response: cigs
           Df Sum Sq Mean Sq F value Pr(>F)
age         1    262  261.88
Residuals 805 151492  188.19
```