Econ 140 - Spring 2016
Section 5

GSI: Fenella Carpena

February 18, 2016

1 Hypothesis Testing for Regression Coefficients

Exercise 1.1. (Adapted from Stock & Watson, Exercise 5.1) Suppose that a researcher, using data on class size (CS) and average test scores from 100 third-grade classes, estimates the OLS regression

$$\widehat{TestScore} = \underset{(20.4)}{520.4} - \underset{(2.21)}{5.82} * CS, \quad R^2 = 0.08, \quad SER = 11.5$$

(a) Test $H_0 : \beta_1 = -1.5$ vs. $H_a : \beta_1 < -1.5$ at 5% significance level using a t-statistic.

(b) Calculate the p-value for the two-sided test of the null hypothesis $H_0 : \beta_1 = 0$. Do you reject the null hypothesis at the 5% level? At the 1% level?

(c) Calculate the p-value for the two-sided test of the null hypothesis $H_0 : \beta_1 = -5.6$. Without doing any additional calculations, determine whether -5.6 is contained in the 95% confidence interval for β_1.

(d) Construct a 95% confidence interval for β_1.

(e) Construct a 99% confidence interval for β_0.

2 Dummy Variables

Exercise 2.1. Suppose that the Berkeley Undergraduate Program Office wishes to examine the potential educational benefits of providing students with an iPad. Using data from 141 students, the table below summarizes the mean and SD of the variable *collGPA* (college GPA) for students with and without an iPad.

	n	Mean	SD
Without iPad	85	2.989	0.321
With iPad	56	3.159	0.421

(a) Let $\overline{X}_1$ be the sample mean collage GPA of students with an iPad, and $\overline{X}_0$ be the sample mean collage GPA of students without an iPad. Calculate:

(i) $\overline{X}_1 - \overline{X}_0$

(ii) $SE(\overline{X}_1 - \overline{X}_0)$

(b) The program office does not plan to provide iPads to students unless there is evidence that having an iPad increases students' GPAs. What is the one-sided null and alternative hypothesis that the program office would like to test?

(c) Carry out the test from part (b) at the 1% level. Do you reject or fail to reject the null hypothesis?

(d) Can the difference in the GPA between students with and without an iPad be attributed to iPad ownership alone? Explain why or why not.

(e) Suppose that an analyst at the program office instead estimated a regression of college GPA on $HasIPAD$, a dummy variable equal to 1 if the student has an iPad, and 0 otherwise, obtaining the following regression results. Interpret each of the coefficients in the regression.

```
Linear regression                                      Number of obs =     141
                                                       F(  1,   139) =    6.57
                                                       Prob > F      =  0.0114
                                                       R-squared     =  0.0500
                                                       Root MSE      =  .36419

------------------------------------------------------------------------------
             |               Robust
      colGPA |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     HasIPAD |   .1695168   .0661386     2.56   0.011      .038749    .3002846
       _cons |   2.989412   .0349104    85.63   0.000     2.920388    3.058436
------------------------------------------------------------------------------
```

(f) Consider again the null and alternative hypothesis from part (b).

(i) How can you use the regression coefficients from part (e) to test the same null and alternative hypothesis as in part (b)?

(ii) Carry out the hypothesis test using the regression coefficients at the 1% level.

(iii) What does question (f)(i) and (f)(ii) tell us about how we can use regressions to carry out a test for the differences of means across two groups?

3 Heteroskedasticity and Homoskedasticity

Exercise 3.1. (Stock & Watson, Review the Concepts 5.3) Define *homoskedasticity* and *heteroskedasticity*. Provide a hypothetical empirical example in which you think the errors would be heteroskedastic and explain your reasoning.

Exercise 3.2. (Stock & Watson, Exercise 5.5, 5.6) In the 1980s, Tennessee conducted an experiment in which kindergarten students were randomly assigned to "regular" and "small" classes and given standardized tests at the end of the year. (Regular classes contained approximately 24 students, and small classes contained approximately 15 students.) Suppose that, in the population, the standardized tests have a mean score of 925 points and a standard deviation of 75 points. Let $SmallClass$ denote a binary variable equal to 1 if the student is assigned to a small class, and 0 otherwise. A regression of $TestScore$ on $SmallClass$ yields

$$\widehat{TestScore} = \underset{(1.6)}{918.0} + \underset{(2.5)}{13.9} * SmallClass, \quad R^2 = 0.01, \quad SER = 74.6$$

(a) Do small classes improve test scores? By how much? Is this effect large? Explain.

(b) Is the estimated effect of a class size on test scores statistically significant? Carry out the test at the 5% level.

(c) Construct a 99% confidence interval for the effect of $SmallClass$ on $TestScore$.

(d) Do you think that the regression errors are plausibly homoskedastic? Explain.

(e) $SE(\hat{\beta}_1)$ was computed using heteroskedastic standard errors. Suppose that the regression errors were homoskedastic: Would this affect the validity of the confidence interval constructed in part (c)? Explain.