Problem Set 2

Due: September 25, 2017

Groundrules *You are allowed to discuss the problem set with your classmates but every student must submit his\her writeup. The problem set can be handed to the TA (Chiara Felli) during her TA session on Monday September 25 or scanned and uploaded on the LEARN platform.*
The problem sets will not be graded. You should check the detailed answer sheet that will be posted on the course website after the deadline for the problem set submission has passed. If you have troubles understanding the problem set solutions, you are encouraged to attend office hours. In some cases, the TA may go over some of the questions appearing in a problem set during the TA session.

1 Question 1

The tables below are summary statistics from the California test scores dataset often referenced in the textbook.
The variables we are interested in are: the number of students enrolled on average at schools in the district (*Enrolment*), the average number of teachers working at schools in the district (*Teachers*), the average number of computers available at schools in the district (*Computer*) and the average standardized test score for schools in the district (*Test score*).
Below you find descriptive statistics for the full sample, which includes both K-8 and K-6 schools:

Variable	*N*	*Mean*	*Std. dev*
Enrolment	420	2628.8	3913.1
Teachers	420	129.1	187.9
Computer	420	303.4	441.3
Test score	420	654.2	19.1

Here are summary statistics for the subsample of K-6 schools:

Variable	*N*	*Mean*	*Std. dev*
Enrolment	61	3181.3	4380.9
Teachers	61	157.7	210.3
Computer	61	338.4	476.7
Test score	61	661.5	19.7

Here are summary statistics for the subsample of K-8 schools:

Variable	*N*	*Mean*	*Std. dev*
Enrolment	359	2534.9	3826.8
Teachers	359	124.2	183.7
Computer	359	297.4	435.5
Test score	359	652.9	18.7

1. Can you reject $H_0 : \mu^{Enrolment} = 2250$ vs $H_1 : \mu^{Enrolment} \neq 2250$ at 5% in the overall sample? What if you performed the same test for the sample of K-6 schools? How would your answers change if you were to perform a 1% test?

2. Find the p-value for the test of $H_0 : \mu^{Teachers} = 130$ vs $H_0 : \mu^{Teachers} > 130$ for the sample of K-8 schools. Based on this p-value, would you reject the null hypothesis at 1%? *[You can do this using the Standard Normal tables or taking advantage of the Stata command normal(x) which returns the value of the Standard Normal cdf at x.]*

3. Compute the 90%, 95% and 99% confidence interval for $\mu^{Computers}$ in the sample of K-6 school.

4. Can you reject the null $H_0 : \mu_{K-6}^{TestScore} = \mu_{K-8}^{TestScore}$ vs the alternative $H_1 : \mu_{K-6}^{TestScore} \neq \mu_{K-8}^{TestScore}$ at 1%?

5. Can you reject the null $H_0 : \mu_{K-6}^{Computer} = \mu_{K-8}^{Computer}$ vs the alternative $H_1 : \mu_{K-6}^{Computer} \neq \mu_{K-8}^{Computer}$ at 1%?

2 Question 2

Below you can see the Stata output of a regression estimating a linear relationship between the average math test score (*MathScore*) and the average students-teachers ratio (*STR*) in a district as in the model below:

$$MathScore_i = \alpha + \beta STR_i + u_i$$

```
      Source |       SS           df       MS              Number of obs =     420
-------------+------------------------------           F(  1,   418) =   16.62
       Model |  5635.62443     1  5635.62443           Prob > F      =  0.0001
    Residual |  141735.097   418   339.07918
-------------+------------------------------
       Total |  147370.722   419   351.720099

------------------------------------------------------------------------------
math_scr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         str |  -1.938591   .4755165    -4.08   0.000    -2.873292   -1.003889
       _cons |   691.4174   9.382469    73.69   0.000     672.9747    709.8601
------------------------------------------------------------------------------
```

1. What would be the change in math test score associated with a reduction of the student-teachers ratio of 3.7?

2. How would you argue whether the effect calculated above is large or small?

3. Compute the R^2 of this regression

4. Compute the Standard Error of the Regression