ECO220Y: Homework, Lecture 21

Readings: Sections 20.1-20.4

Exercises: 13 (skip part d), 14 (skip part b), 15

Problems:

(1) In doing a test of statistical significance, economists often use a simple "rule of thumb": Is slope coefficient divided by its standard error either > 2 or < -2. What is the sense of this rule of thumb?

(2) A researcher wishes to estimate the parameters of the following model.

$$\ln(q_{1i}) = \alpha_1 + \beta_1 \ln(p_{1i}) + \gamma_1 \ln(p_{2i}) + \varepsilon_{1i}$$
$$\ln(q_{2i}) = \alpha_2 + \beta_2 \ln(p_{2i}) + \gamma_2 \ln(p_{1i}) + \varepsilon_{2i}$$

Practice using the rule of thumb for the following estimation results:

	Parameter Estimate	Standard Error (s.e.)
$\hat{\alpha}_1$	10.0	6.3
$\hat{\alpha}_2$	11.0	4.8
$\hat{\beta}_1$	-2.5	0.8
$\hat{\beta}_2$	-2.0	0.6
$\hat{\gamma}_1$	1.0	0.4
$\hat{\gamma}_2$	1.5	0.5

(3) There is some limited (and not very convincing) evidence that sitting close to the front of the classroom improves a student's performance. Consider the research question: What is the effect of seat location on a student's performance in a course? The researcher obtains approval to conduct an experiment where students in ECO220Y are randomly assigned a seat in a classroom where they must sit for the entire course. Attendance is taken to ensure compliance in every lecture. The following variables are available in the data:

MARK_220	Student's percentage mark in ECO220Y
ROW	Row number of student (row 1 is first row at the front of the lecture hall)
MARK_100	Student's percentage mark in ECO100Y

```
. su MARK_220 MARK_100 ROW;

Variable | Obs    Mean       Std. Dev.            Min         Max
-------------+--------------------------------------------------------
MARK_220 | 250   66.42        12.35016         36.63087     97.20178
MARK_100 | 250   81.84        5.484407         67           98
ROW      | 250   13          7.225568         1            25
```

```
. regress MARK_220 ROW MARK_100;
Source   | SS              df            MS         Number of obs = 250
---------------------------------------------       F( 2, 247) = 166.23
Model    | 21790.3847     2      10895.1924         Prob > F = 0.0000
Residual | 16188.7148   247      65.541355         R-squared = 0.5737
---------------------------------------------       Adj R-squared = 0.5703
Total    | 37979.0996   249      152.526504         Root MSE = 8.0958
------------------------------------------------------------------------------
MARK_220 |    Coef.     Std. Err.      t     P>|t|        [95% Conf. Interval]
-------------+----------------------------------------------------------------
ROW      | -.4845315    .0710174     -6.82   0.000        -.6244085 -.3446546
MARK_100 | 1.56963      .0935637     16.78   0.000          1.385345 1.753914
_cons    | -55.73957    7.747089     -7.19   0.000        -70.99835 -40.48079
------------------------------------------------------------------------------
```

(a) Interpret the coefficient estimates (slopes and intercepts). Are they of the expected sign?
(b) Do we have sufficient evidence to infer that our research hypothesis is true? (Show your work and explain.)
(c) Considering the following simple regression with these same data, are you surprised by these results? If so, explain. If not, explain how these results are what you would expect.

```
. regress MARK_220 ROW;

Source   |     SS        df          MS                Number of obs = 250
---------------------------------------------          F( 1, 248) = 23.95
Model    | 3344.68018     1      3344.68018            Prob > F = 0.0000
Residual | 34634.4194   248      139.654917            R-squared = 0.0881
---------------------------------------------          Adj R-squared = 0.0844
Total    | 37979.0996   249      152.526504            Root MSE = 11.818

------------------------------------------------------------------------------
MARK_220 |    Coef.     Std. Err.     t      P>|t|       [95% Conf. Interval]
-------------------------------------------------------------------------------
     ROW | -.5072308    .1036469    -4.89    0.000       -.7113713 -.3030903
   _cons |    73.014    1.540822    47.39    0.000         69.97923 76.04877
------------------------------------------------------------------------------
```

(e) The reason that existing evidence is not very convincing is because it often relies on observational data. Describe the nature of observational data that would be available to answer the research question. Describe what would happen if a regression analysis were conducted using such data. Indicate the direction of bias on the coefficient of interest.