Hypothesis testing about a parameter of the population regression (end)

3. Pollution by paper mills: Are larger and smaller paper mills differentially dirty?

Data source: Data collected by Jay Shimshack, professor at Tulane University
Discharge of suspended solids in waterways by the 160 major pulp and paper plants located throughout the United States (in 23 states) in the month of January 1990

Emission: total suspended solids discharged, in pounds.
Permit: maximum allowance of suspended solids discharge under the law
Size: firm size - production capacity in kilotons/day
Pulp: a dummy variable equal to 1 for pulp manufacturer and 0 for paper manufacturer

$$\widehat{emission} = \underset{(1020)}{923} + \underset{(.81)}{4.74}\ size + \underset{(1347)}{3864}\ pulp \qquad R^2 = .25,\ n = 160$$

$$\widehat{\ln(emission)} = \underset{(.68)}{1.51} + \underset{(.110)}{.907}\ \ln(size) + \underset{(.29)}{.70}\ pulp \qquad R^2 = .37,\ n = 160$$

<u>The five steps of hypothesis testing:</u>

Testing whether emission increases proportionately to size, i.e., when plant size increases by 10%, emission also increases by 10%, i.e. the true value $\beta_{\ln size} = 1$, as opposed to larger plants being cleaner or dirtier

a. Set the hypotheses: Remember H0 is the hypothesis that you will attempt to reject in favor of H1

H0: $\beta_{\ln size} = 1$
H1: $\beta_{\ln size} \neq 1$

b. Construct the statistic

$$\text{t-stat: } t_{160-2-1} = \frac{\hat{\beta} - \beta_{(under\ H0)}}{se(\hat{\beta})} = \frac{.907 - 1}{.110} = -0.85$$

c. Select the significance level. Given the distribution (Student t in this case) and the degrees of freedom, find out the critical value.
At 5% significance level and 157 degrees of freedom, the critical value for a two-tailed test is 1.96

d. Decide whether to reject H0 or not.
Since $|t| < 1.96$, we cannot reject H0 that the true parameter $\beta_{\ln size} = 1$

e. Conclude with a (reader friendly) sentence:
We cannot reject the hypothesis that emission increases proportionately to plant size
There is no statistical evidence that emission does not increase proportionately to plant size. Or there is no statistical evidence that larger firm are differentially polluting the waterways.

<u>Stata output:</u>

```
. reg lwage educ exper female nonwhite

      Source |       SS       df       MS              Number of obs =    2000
-------------+------------------------------           F(  4,  1995) =  186.03
       Model |  182.711923     4  45.6779807           Prob > F      =  0.0000
    Residual |  489.864945  1995  .245546338           R-squared     =  0.2717
-------------+------------------------------           Adj R-squared =  0.2702
       Total |  672.576867  1999  .336456662           Root MSE      =  .49553

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .1166997   .0053153    21.96   0.000     .1062756    .1271237
       exper |   .0108872   .0008691    12.53   0.000     .0091827    .0125917
      female |  -.2533177   .0222198   -11.40   0.000    -.2968942   -.2097412
    nonwhite |  -.0374311   .0311452    -1.20   0.230    -.0985117    .0236495
       _cons |   1.061903   .0759003    13.99   0.000     .9130514    1.210756
------------------------------------------------------------------------------
```