

Problem set 9, due Monday 11/16/2015

1. Simple linear regressions in Excel and Stata

Go to the course web page, <http://inside.bard.edu/~armytag/metricsf15/>

There, next to the link for this problem set, find two simulated data sets, each containing a column of n x_i values and n corresponding y_i values. For each of these two data sets, do the following:

a) In Excel, calculate the following values:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

b) In Excel, make a scatter plot of x and y , with a line representing the estimated linear relationship between x and y .

c) Import the x and y columns into Stata, and estimate the regression coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$.

d) In Stata, make a scatter plot of x and y , with a line representing the estimated linear relationship between x and y .

e) Verify that the answers you get in Excel and Stata are the same (or approximately the same). According to this analysis, what is the best linear estimate of the relationship between x and y ?

2. Reading and exercises from Studenmund

Read Chapters 1-2. Do exercises 1-3 at the end of Chapter 1