What is Econometrics?

Parameters

- The *parameters* we have studied so far are means or proportions (rates).

- We've also studied differences between means and proportions.

New Parameters

- But imagine a relationship between two random variables, labelled *y* and *x*, given by:

$$y = b_0 + b_1 x + e$$

where *e* is a random error with a mean of zero.

- And now imagine the population values of b_0 and b_1 are the parameters we are interested in.

Examples: Response/Explanatory

• *y*	*x*
• life expectancy	GDP per capita
• alcohol consumption	alcohol price index

(these vary in a cross-section)

• average temperature	CO_2 ppm
• exchange rate	previous day's value

(these vary over time)

Estimation

- We can estimate the two parameters with point estimates or form confidence intervals.

- The usual way to do this is called regression, or ordinary least squares regression (OLS)

Least Squares

- Visualize the scatter plot of *y* versus *x*.
- Imagine drawing a straight line through the points.
- OLS formulas, in software, minimize the sum of squared departures from the line so the line fits the data as closely as possible.

Applet

- Experiment with the applet

correlationregression

- You can see how the sum of squared residuals rises and falls as you adjust the line.
- Software performs this task for us with actual data.

- We also could fit curves, for example if we think *y* is related to the square of *x* or some other functional form.

- The results include estimates of b_0 and b_1, along with *standard errors*. Thus we can form confidence intervals or test hypotheses.

Hypothesis Tests

- *e.g.* Ho: The effect of alcohol prices on consumption is zero.

- *e.g.* Ho: $10000 of GDP per capita leads to 3 extra years of life expectancy.

- We can test these hypotheses with *t*-tests.

Fit and Prediction

- We also can report on the overall fit of the regression line, using a statistic called R-squared (R^2). It gives the percentage of the variation in *y* explained by *x*.

- And, if the line fits well, we can use it to predict or forecast values of *y*.

Applet

- We can use the applet

  twovar

  to see how well a regression fits.

- Select a data set, then look at the scatterplot (showing the least-squares line).

- Finally, look at the correlation and regression tab for the corresponding statistics.

Multiple Regression

- We can control for other influences on y, by using multiple regression:

$$y = b_0 + b_1x_1 + b_2x_2$$

e.g. Do wages (y) depend on education (x_1) or seniority (x_2) or both?

Warnings

- Remember to always plot your data (with a scatter plot or time series plot, or both). This helps you check for errors or outliers.

- Also remember that correlation (or a fitted regression) does not imply causation. We need theory or experiments to interpret the results.

Running a Regression

- In econometrics we usually use free software (like *gretl*) or widely-used professional software (like *stata*).

- But you can try your own examples in *excel*, with 2 columns of data and the *data* tab, then the *data analysis* pack, then the *regression* choice.