Lecture 2 - Economics 113

1. Variance, Standard Deviation
2. Covariance
3. Correlation

I'm a great believer in luck, and I find the harder I work, the more I have of it.
-Thomas Jefferson

Course information

- The syllabus is the main source of information for the course. Please check the syllabus before asking questions.
- Course website is available at http://people.ucsc.edu/~aspearot/Teaching.html
- Book: *Introductory Econometrics* by Jeffrey Wooldridge
- Office hours: 3PM-5PM on Mondays, 453 Engineering 2
- Email: aspearot@ucsc.edu
- Sections start this week!

Course information

Grades

- Yes, I do curve.
- No, it is not consistent from quarter to quarter.

- The curve will be worse if you disrespect your classmates, TAs, or me.
- Don't cheat. It will not be tolerated. If you're caught, you will receive a failing grade and be reported for academic misconduct.
- Regrades require a written explanation of your answer, and why you think it is correct.
- Regrades require unanimous agreement among TAs that more points are deserved.

Course wisdom

On Soapbox

- Exams won't (and shouldn't) be the homework with different numbers.
- Now is the time to work hard.
- You will solve problems in this course. Don't give up if things are difficult to begin with.

If you think you can, or you think you can't, you're probably right.
-Henry Ford

Summary measures

Measures of variation

- **Range**

  A measure of data dispersion, though not used for many applications

  1. Identify the largest observation
  2. Identify the smallest observation
  3. Take the difference

Summary measures

Measures of variation (cont.)

- **Sample Variance**
  - The most commonly used measure of dispersion.
  - Summarizes the how far a typical observation is from the mean

$$\widehat{\sigma}_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \widehat{\mu}_x)^2$$

- Why do we divide by $n-1$ instead of n?

$\Rightarrow$ *We used one "piece" of information to calculate* $\widehat{\mu}_x$

Summary measures

Measures of variation (cont.)

- **Sample Standard deviation**

$$\widehat{\sigma}_x = \sqrt{\widehat{\sigma}_x^2}$$

- This is more desirable than the sample variance. Why?

$\Rightarrow$ On the same scale as $\widehat{\mu}_x$.

Covariance

Relationships

- Covariance describes the relationship between two random variables
- $\widehat{\sigma}_{xy} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \widehat{\mu}_x)(y_i - \widehat{\mu}_y)$
- When will covariance be positive/negative?
- $\widehat{\sigma}_{xy} > 0 \Rightarrow$ tends to have $x_i > \widehat{\mu}_x$ when $y_i > \widehat{\mu}_y$ (and vice versa)
- $\widehat{\sigma}_{xy} < 0 \Rightarrow$ tends to have $x_i > \widehat{\mu}_x$ when $y_i < \widehat{\mu}_y$ (and vice versa)
- Covariance describes a "linear" relationship
- Any non-linear relationships with zero covariance?

Covariance

Example

- US Imports and Exports, 2007, percentage change
- Do you think they positively or negatively covary?
- $IM = \{4, -3, 4-1\}$, $EX = \{1, 8, 19, 6\}$
  (rounded percentage changes)
- $\widehat{\sigma}_{EX,IM} = \frac{1}{n-1}\sum_{i=1}^{n}\left(EX_i - \widehat{\mu}_{EX}\right)\left(IM_i - \widehat{\mu}_{IM}\right)$
- Compute mean of $\widehat{\mu}_{IM}$:

$$\widehat{\mu}_{IM} = \frac{1}{4}(4-3+4-1) = \frac{1}{4}(4) = 1$$

- Compute mean of $\widehat{\mu}_{EX}$:

$$\widehat{\mu}_{EX} = \frac{1}{4}(1+8+19+6) = \frac{1}{4}(34) = 8.5$$

Covariance

Example

$$
\begin{aligned}
\widehat{\sigma}_{EX,IM} &= \frac{1}{4-1}\sum_{i=1}^{4}\left(EX_i-\widehat{\mu}_{EX}\right)\left(IM_i-\widehat{\mu}_{IM}\right) \\
&= \frac{1}{4-1}\left(\begin{array}{c} (4-1)(1-8.5)+(-3-1)(8-8.5)+ \\ (4-1)(19-8.5)+(-1-1)(6-8.5) \end{array}\right) \\
&= \frac{1}{4-1}\left(-3\cdot 7.5+4\cdot 0.5+3\cdot 10.5+2\cdot 2.5\right) \\
&= \frac{1}{4-1}\left(-22.5+2+31.5+5\right)=5.33
\end{aligned}
$$

Correlation

Basic

- Correlation describes a linear relationship
- $\widehat{\rho}_{xy} = \frac{\widehat{\sigma}_{xy}}{\widehat{\sigma}_x \widehat{\sigma}_y}$
- $\widehat{\rho}_{xy} \in [-1, 1]$

Correlation

Import-Export Example

- $IM = \{4, -3, 4, -1\}, EX = \{1, 8, 19, 6\}$
- $\widehat{\mu}_{IM} = 1, \widehat{\mu}_{EX} = 8.5, \widehat{\sigma}_{EX,IM} = 5.33$
- Compute mean of $\widehat{\sigma}_{IM}$:

$$\begin{aligned}\widehat{\sigma}_{IM} &= \sqrt{\frac{1}{3}((4-1)^2 + (-3-1)^2 + (4-1)^2 + (-1-1)^2)} \\ &= \sqrt{\frac{1}{3}(9+16+9+4)} = \sqrt{\frac{1}{3}(38)} = 3.559026\end{aligned}$$

- Compute mean of $\widehat{\sigma}_{EX}$:

$$\begin{aligned}\widehat{\sigma}_{EX} &= \sqrt{\frac{1}{3}((1-8.5)^2 + (8-8.5)^2 + (19-8.5)^2 + (6-8.5)^2)} \\ &= \sqrt{\frac{1}{3}((-7.5)^2 + (-0.5)^2 + (10.5)^2 + (-2.5)^2)} \\ &= 7.593857\end{aligned}$$

- $\widehat{\rho}_{zy} = \frac{\widehat{\sigma}_{zy}}{\widehat{\sigma}_z \widehat{\sigma}_y} = \frac{5.33}{3.559 * 7.59} = 0.1973124$