Economics 4811 - Handout #1

Math and Stats Review

Key Takeaways from Wooldridge Appendices A & B

Hop in your Way-Back Machine and refresh your mind about the following math and stats. Some keys:

- Properties of summation operators
- Expectations and conditional statements
- Correlation vs. Causation

1 Summation Operators

Given a sequence of numbers $\{x_i : i = 1, 2, ..., n\}$:

$$\sum_{i=1}^{n} x_i \equiv x_1 + x_2 + ... + x_n \tag{1}$$

if we have a constant, c:

$$\sum_{i=1}^{n} c = nc \tag{2}$$

$$\sum_{i=1}^{n} cx_i = c\sum_{i=1}^{n} x_i \tag{3}$$

What you *can* and *can't* do with summations - show for yourself with $n = 3$:

$$\sum_{i=1}^{n}(ax_i + by_i) = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} y_i \tag{4}$$

$$\sum_{i=1}^{n} \frac{x_i}{y_i} \neq \frac{(\sum_{i=1}^{n} x_i)}{(\sum_{i=1}^{n} y_i)} \tag{5}$$

$$\sum_{i=1}^{n} x_i^2 \neq \left(\sum_{i=1}^{n} x_i\right)^2 \tag{6}$$

Other important summation expressions to recall:

- Let $d_i \equiv x_i - \bar{x}$ where $\bar{x} = \frac{1}{n}\sum\limits_{i=1}^{n} x_i$, then $\sum\limits_{i=1}^{n} d_i = 0$

- We're going to care a lot about this expression:

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{7}$$

- Or more generally,

$$\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \tag{8}$$

- Can you show that equation (8) equals each of the following?

$$\sum_{i=1}^{n} x_i(y_i - \bar{y}) = \sum_{i=1}^{n} y_i(x_i - \bar{x}) = \sum_{i=1}^{n} x_i y_i - n(\bar{x}\bar{y})$$

2 Basic Derivatives

Remember the following and make your calculus teachers proud:

$$\frac{\partial}{\partial x}x^2 = \qquad\qquad \frac{\partial}{\partial x}\log(x) =$$

How is a $\log(x)$ function useful in calculating elasticities? (Price elasticity: $\frac{\partial Q^D}{\partial P}\frac{P}{Q^D}$)

$$\frac{\partial \log(Q^D)}{\partial \log(P)} =$$

3 Probability Concepts

3.1 PDFs and CDFs

You should have a good sense of the following terms:

- **Random Variables**
- **Discrete** vs. **continuous** random variables
- **Probability Density Function (pdf)** and **Cumultative Distribution Function (cdf)**
- **Joint Distribution**

3.2 Independence

Two variables X and Y are said to be independent iff:

$$f_{XY}(x,y) = f_X(x)f_Y(y)$$

3.3 Conditional Distribution

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)}$$

When X and Y are independent, knowing X doesn't tell us anything about Y:

$$f_{Y|X}(y|x) = f_Y(y)$$

Similar expressions exist for conditional probabilities (ie. $P(A|B) = \frac{P(A\cap B)}{P(B)}$)

3.4 Expectations and More

Recall that we'll be thinking **population** vs. **sample** all semester. So far in class we have been talking **sample**, but here are some **population** level properties.

Expected value - your "best guess" - or weighted average:

- For discrete random variables, with $f(x)$ denoting the pdf of X:

$$E(X) = x_1 f(x_1) + x_2 f(x_2) + ... + x_k f(x_k) \equiv \sum_{j=1}^{k} x_j f(x_j) = \mu$$

- Shares similarities with summation operator:

$$E(aX + b) = aE(X) + b \qquad E[g(X)] \neq g[E(X)]$$

Variance - dispersion - how spread out is the data?

$$Var(X) \equiv E[(X - \mu)^2] = \sigma^2$$

- The following is a useful expression - can you show it?:

$$Var(X) = E(X^2) - \mu^2$$

We'll care a lot about how random variables relate to each other.

Covariance - When X is increasing, is Y increasing? When X is decreasing, is Y decreasing?

$$Cov(X,Y) \equiv E[(X-\mu_X)(Y-\mu_Y)] = E(XY) - \mu_X \mu_Y$$

- Can you prove that last equality?
- What does it mean when $Cov(X,Y) < 0$? What about > 0 or $= 0$? Know how to draw examples of random variables that fall into these categories.
- Note that covariance informs you about *linear* relationships.

Correlation Coefficient (ρ) -

$$Corr(X,Y) \equiv \frac{Cov(X,Y)}{sd(X)sd(Y)} = \rho_{XY}$$

- Note that $-1 \leq Corr(X,Y) \leq 1$
- If $\rho = 1$, how can you describe the relationship between X and Y?
- What's the difference between **correlation** and **causation**? What do we care about? Why?

Conditional expectation - $E(Y|X=x)$ -What's your best guess for Y when you know the value of X? Useful for seeing how X can help explain Y:

- For example: what's the expected value of wages for those with college degrees?
- Or: what's the expected number of cookies left in Bob's if you show up 30 minutes late for lunch?
- Can you explain the **Law of Iterated Expectations** in words?

$$E[E(Y|X)] = E(Y)$$

3.5 Key Distributions

- **Normal** distribution - bell curve. A few rules of thumb (with a nice picture):
  - Approximately 68% of the observations lie within one standard deviation of the mean
  - Approximately 95% of the observations lie within approximately two (really 1.96) standard deviations of the mean
- ***t***, **F**, and **Chi-Square** (χ^2) distributions - useful for hypothesis testing