Review

1 The Summation Operator

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

For any constant :

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

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

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

These are NOT equal:

$$\sum_{i}^{n} \frac{x_i}{y_i} \neq \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} y_i}$$

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

We can use the summation operator to calculate the average:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

The sum of the deviations from the mean is always 0:

$$\sum_{i=1}^{n} (x_i - \bar{x}) = 0$$

Useful relations:

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

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

2 Proportion and Percentages

The proportionate change in x in moving from $x_1 \; to \; x_0$

$$(x_1 - x_0)/x_0 = \Delta x / x_0$$

The percentage change in x in going from x_0 to x_1 is simply 100 times the proportionate change ($\%\Delta x = 100$):

$$\%\Delta x = 100\,(\Delta x / x_0)$$

Logarithms
In particular, $\log(x)$ can be positive or negative. Some useful algebraic facts about the log function are

$$\log\left(x_1 \cdot x_2\right)=\log\left(x_1\right)+\log\left(x_2\right), x_1, x_2>0$$

$$\log\left(x_1 / x_2\right)=\log\left(x_1\right)-\log\left(x_2\right), x_1, x_2>0$$

$$\log\left(x^c\right)=c \log(x), x>0,\ c\ any\ number$$

$$\log(x)<0\ for\ 0<x<1$$

$$\log(1)=0$$

$$\log(x)>0\ for\ x>1$$

Occasionally, we will need to rely on these properties. The logarithm can be used for various approximations that arise in econometric applications. First, $\log(1+x) \approx x$ for $x \approx 0$.Then, it can be shown (using calculus) that:

$$\log\left(x_1\right)-\log\left(x_0\right) \approx\left(x_1-x_0\right) / x_0=\Delta x / x_0$$

for small changes in x.

$$\Delta \log(x)=\log\left(x_1\right) \log\left(x_0\right)$$

then

$$100 \cdot \Delta \log(x) \approx \% \Delta x$$

for small changes in x. The meaning of "small" depends on the context, and we will encounter several examples throughout this class.

Elasticity of y with respect to x as:

$$\frac{\Delta y}{\Delta x} \cdot \frac{x}{y}=\frac{\% \Delta y}{\% \Delta x}$$

3 Fundamentals of Probability

A **random variable** (X) is one that takes on numerical values and has an outcome that is determined by an experiment.

The **probability density function (pdf)** of X summarizes the information concerning the possible outcomes of X and the corresponding probabilities:

$$f\left(x_j\right)=p_j, j=1,2, \ldots, k$$

A variable X is a **continuous random variable** if it takes on any real value with zero probability. When computing probabilities for continuous random variables, it is easiest to work with the **cumulative distribution function (cdf)**. If X is any random variable, then its cdf is defined for any real number x by

$$F(x) \equiv \mathrm{P}(X \leq x)$$

Joint Distributions and Independence

Let X and Y be discrete random variables. Then, (X, Y) have a joint distribution:

$$f_{X, Y}(x, y)=\mathrm{P}(X=x, Y=y)$$

Random variables X and Y are said to be independent if, and only if,

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

This definition of independence is valid for discrete and continuous random variables.

4 Expected Value

Suppose that the variable X can take values $x_1, x_2 ... x_k$ each with probability $f(x_1), f(x_2) ... f(x_k)$

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

We can also express this using the expectation operator:

$$E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1} a_i E(X_i)$$

And in the special case where $a_i = 1$, then

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$$

For any constants a, b, and two random variables W and H

$$E(aW + b) = aE(W) + b$$

$$E(W + H) = E(W) + E(H)$$

$$E(W - E(W)) = 0$$

5 Variance

Consider the variance of a random variable, W :

$$V(W) = \sigma^2 = E\left[(W - E(W))^2\right]$$

We can show

$$V(W) = E\left(W^2\right) - E(W)^2$$

In a given sample of data, we can estimate the variance by the following calculation:

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

Properties of variance:

$$V(aX + b) = a^2 V(X)$$

And the variance of a constant is 0 (i.e., $V(c) = 0$ for any constant, c). The variance of the sum of two random variables is equal to:

$$V(X + Y) = V(X) + V(Y) + 2(E(XY) - E(X)E(Y))$$

If the two variables are independent, then $E(XY) = E(X)E(Y)$ and $V(X + Y)$ is equal to the sum of $V(X) + V(Y)$.

6 Standard Deviation

The standard deviation of a random variable, denoted $sd(X)$, is simply the positive square root of the variance: $sd(X) \equiv +\sqrt{Var(X)}$.

Property SD.1: For any constant $c, sd(c) = 0$

Property SD.2: For any constants a and b,

$$sd(aX + b) = |a| sd(X)$$

Standardizing a Random Variable

$$Z \equiv \frac{X-\mu}{\sigma}$$

which we can write as $Z = aX + b$, where $a \equiv (1/\sigma)$ and $b \equiv -(\mu/\sigma)$. Then, from Property E.2,

$$\mathrm{E}(Z) = a\mathrm{E}(X) + b = (\mu/\sigma) - (\mu/\sigma) = 0.$$

7 Covariance

The expression $C(X,Y) > 0$ indicates that two variables move in the same direction, whereas $C(X,Y) < 0$ indicates that they move in opposite directions.

$$V(X+Y) = V(X) + V(Y) + 2C(X,Y)$$

A zero covariance means that two random variables are unrelated, that is incorrect. They could have a nonlinear relationship.

The definition of covariance is:

$$C(X,Y) = E(XY) - E(X)E(Y)$$

As we said, if X and Y are independent, then $C(X,Y) = 0$ in the population.

$$Corr(X,Y) = Cov(W,Z) = \frac{C(X,Y)}{\sqrt{V(X)V(Y)}}$$

where

$$W = \frac{X - E(X)}{\sqrt{V(X)}} \; and \; Z = \frac{Y - E(Y)}{\sqrt{V(Y)}}$$

The correlation coefficient is bounded by -1 and 1. A positive (negative) correlation indicates that the variables move in the same (opposite) ways. The closer the coefficient is to 1 or -1, the stronger the linear relationship is.

8 Conditional Expectation

We denote this expected value by $\mathrm{E}(Y \mid X = x)$, or sometimes $\mathrm{E}(Y \mid x)$ for shorthand. Generally, as x changes, so does $\mathrm{E}(Y \mid x)$ When Y is a discrete random variable taking on values $\{y_1, \ldots, y_m\}$, then

$$\mathrm{E}(Y \mid x) = \sum_{j=1}^{m} y_j f_{Y \mid X}\left(y_j \mid x\right)$$

When Y is continuous, $\mathrm{E}(Y \mid x)$ is defined by integrating $v f_{y\mathrm{y}}(y \mid x)$ over all possible values of y.

Review of Mathematical Statistics

Population means any well-defined group of subjects, which could be individuals, firms, cities, or many other possibilities.

Random Sampling. If $Y_1, Y_2, \ldots, Y_n$ are independent random variables with a common probability density function $f(y;\theta)$, then $\{Y_1, \ldots, Y_n\}$ is said to be a random sample from $f(y;\theta)$