Problem Set #3 Answer Key

Economics 435: Quantitative Methods

Fall 2011

1 Averages

a) The condition is $a_1 + a_2 + \cdots + a_n = 1$.

b) The answer is:

$$var(W_n) = (a_1^2 + a_2^2 + \cdots + a_n^2)\sigma^2$$

c) If W_n is an unbiased estimator then

$$\frac{(a_1 + a_2 + \cdots + a_n)^2}{n} = \frac{1^2}{n} = \frac{1}{n} \leq a_1^2 + a_2^2 + \cdots + a_n^2$$

Multiply both sides of the above by σ^2 and we get:

$$var(W_n) = (a_1^2 + a_2^2 + \cdots + a_n^2)\sigma^2 \geq \frac{\sigma^2}{n} = var(\bar{X}_n)$$

2 Consistency and unbiasedness

a) Yes it is:

$$\begin{aligned} E(\bar{x}) &= E\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} E(x_i) \\ &= E(x) = p \end{aligned}$$

b) Yes it is. By the Law of Large Numbers plim $\bar{x} = E(x) = p$.

c) No it is not. For a given n:

$$\Pr(\bar{x} = 0) = (1-p)^n > 0$$

Since $\ln 0$ does not exist, $E(\ln \bar{x})$ also does not exist. Therefore $E(\ln \bar{x})$ cannot equal p.

d) Yes, it is. We have already established that plim $\bar{x} = p$. Since $p > 0$, the function $\ln(.)$ is continuous at p. Therefore Slutsky's theorem implies:

$$\text{plim } \ln \bar{x} = \ln \text{plim } \bar{x} = \ln p$$

3 Estimating the covariance

a) The analog principle says that if you have an expression that defines your parameter of interest in terms of expected values, you can consistently estimating that parameter if you replace all unknown expected values with their corresponding sample average. This suggests two estimators for $cov(x, y)$:

$$\begin{aligned} \hat{cov}(x,y) &= \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n}x_i\right)\left(y_i - \frac{1}{n}\sum_{i=1}^{n}y_i\right) \\ &= \left(\frac{1}{n}\sum_{i=1}^{n}x_i y_i\right) - \left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)\left(\frac{1}{n}\sum_{i=1}^{n}y_i\right) \end{aligned}$$

A little algebra will reveal that these two estimators are actually identical.

b) I'll prove this for the bottom definition. By the law of large numbers:

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n}x_i y_i &\to^p E(xy) \\ \frac{1}{n}\sum_{i=1}^{n}x_i &\to^p E(x) \\ \frac{1}{n}\sum_{i=1}^{n}y_i &\to^p E(y) \end{aligned}$$

Since the function $f(a, b, c) = a - bc$ is continuous in all arguments, Slutsky's theorem implies:

$$\begin{aligned} \hat{cov}(x,y) &= \left(\frac{1}{n}\sum_{i=1}^{n}x_i y_i\right) - \left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)\left(\frac{1}{n}\sum_{i=1}^{n}y_i\right) \\ &\to^p E(xy) - E(x)E(y) \\ &= cov(x,y) \end{aligned}$$

c)

$$\begin{aligned} E\left[\hat{cov}(x,y)\right] &= E\left[\left(\frac{1}{n}\sum_{i=1}^{n}x_i y_i\right) - \left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)\left(\frac{1}{n}\sum_{i=1}^{n}y_i\right)\right] \\ &= \frac{1}{n}\sum_{i=1}^{n}E(x_i y_i) - \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}E(x_i y_j) \\ &= \frac{1}{n}\sum_{i=1}^{n}E(xy) - \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}E(x_i y_j) \\ &= \frac{1}{n}\sum_{i=1}^{n}E(xy) - \frac{1}{n^2}\sum_{i=1}^{n}E(xy) + (n-1)E(x)E(y) \\ &= \frac{1}{n}nE(xy) - \frac{1}{n^2}n\left(E(xy) + (n-1)E(x)E(y)\right) \\ &= E(xy) - \frac{1}{n}E(xy) - \frac{n-1}{n}E(x)E(y) \\ &= \frac{n-1}{n}\left(E(xy) - E(x)E(y)\right) \\ &= \frac{n-1}{n}cov(x,y) \end{aligned}$$

It is biased. Note that the bias goes away as n goes to infinity, so it is still consistent.

d) Yes, just multiply this estimator by $\frac{n}{n-1}$. This bias correction is standard practice when estimating variances and covariances.

4 Basic statistics in R

This code will work:

```
x <- rbern(1000,100,0.75)
m1000 <- apply(x,2,mean)
m25 <- apply(x[1:25,],2,mean)
m5 <- apply(x[1:5,],2,mean)
m1 <- x[1,]
par(mfrow=c(2,2))
hist(m1)
hist(m5)
hist(m25)
hist(m1000)
```