1/24/08

→ Where we were last time:

I Properties of Expectations, (Co)Variances, & Conditioning

A Expectations

1) Laws of Iterated Expectations

x, y, z = R.V.'s

a) $E(E(x)) = E(x)$

a, b, c = constants

b) $E(E(y|x)) = E(y)$

2) Expectation of a sum is sum of the expectations:

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

3) Expectation of a constant is the constant

a) $E(a) = a$ b) $E(X|X) = X$

to "give" something renders it constant (it "conditions on that info")

4) Rob's Rule of Expections:

"E slides on through constants, snagging around random variables"

a) $E(a + bX + cY) = a + bE(x) + cE(y)$

b) $E\{(X'X)^{-1}X'[ee']X(X'X)^{-1}\} = (X'X)^{-1}X'E(ee')X(X'X)^{-1}$

constant matrix — random vector

B Conditioning:

1) Conditioning is distributive:

$$E(a + bX + cY|X) = E(a|X) + bE(X|X) + cE(Y|X)$$
$$= a + bX + cE(Y|X)$$

2) Conditioning More than Once is Redundant:

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

(on same info!)

3) Conditioning on X renders X constant (see A3b)

4) Conditioning & Expectation sort of undo each other (see A3b, A1b)

C (Co)Variances:

1) Rob's Rules of (Co)Variances:

a) Constants neither vary nor covary.

$$V(a) = 0 \qquad C(a, X) = 0$$

b) "Variance slides through sums, eliminating constants, snagging on random variables and squaring their coefficients, while spitting out 2 times all the covariances times their associated coefficients."

→ $V(a + bX + cY) = b^2 V(X) + c^2 V(Y) + 2bc \times Cov(X,Y)$

→ Notice that in Matrix notation, with $\underset{\sim}{x}$ a constant vector and $\underset{\sim}{b}$ a random vector, this becomes just:

$$V(\underset{\sim}{x}'\underset{\sim}{b}) = \underset{\sim}{x}' V(\underset{\sim}{b}) \underset{\sim}{x}$$

V-cov mat

"x squared" in matrix-land

→ Special Cases of the Rule:

- $V(X+Y) = V(X) + V(Y) + 2 \times Cov(X,Y)$
- $V(X-Y) = V(X) + V(Y) - 2 \times Cov(X,Y)$

$(-1)^2$ $\quad$ $2 \times (-1)$

c) The Analogous for Covariances: (becomes a kind of FOIL):

$$Cov(a + bX + cY,\ d + eW + fV) =$$

$$be\,Cov(X,W) + bf\,Cov(X,V) + ce\,Cov(Y,W) + cf\,Cov(V,Y)$$

D (Selected) Properties of Note Combining the Above:

1) Any R.V. y can be written: $y = E(y|x) + (y - E(y|x))$

$$\equiv E(y|x) + \varepsilon$$

n.b., not necessarily linear $\hat{y}$, & true $E(y|x)$ not literally $\hat{y}$

$$\equiv \hat{y} + \varepsilon$$

2) Properties that follow:

a) $Cov(X,Y) = Cov(X, E(Y|X))$

b) $V(Y) = E(V(Y|X)) + V(E(Y|X))$

$= \hat{\sigma}_\varepsilon^2 + V(\hat{y})$

$\Rightarrow TSS = RSS + ESS = ESS + RSS$

residual, explained, error, regression

$= SSR + SSE = SSE + SSR$

"variance decomposition"

Note: $\Rightarrow V(\hat{y}) \le V(y)$ (and also that $V(\hat{\varepsilon}) \le V(y)$ though that less notable)

I.D.2. (cont.) Properties that follow from $E(y) = E(y|x) + (y - E(y|x))$
$= \hat{y} + \varepsilon$

c) $E(XY) = E(X \cdot E(Y|X))$

d) $E(\varepsilon) = E(\varepsilon|X) = E(\varepsilon|Y) = 0$

i) $E(\varepsilon) = E(Y - E(Y|X)) = E(Y) - E(E(Y|X)) = E(Y) - E(Y) = 0_\square$

ii) $E(\varepsilon|X) = E(Y - E(Y|X)|X) = E(Y|X) - E(E(Y|X)|X) = E(Y|X) - E(Y|X) = 0_\square$

iii) $E(\varepsilon|Y) = E(Y - E(Y|X)|Y) = E(Y|Y) - E(E(Y|X)|Y) = Y - Y = 0_\square$

e) $Cov(X, \varepsilon) = 0$; in fact, $Cov(h(X), \varepsilon) = 0 \ \forall \ h(\cdot)$

$\hookrightarrow E(X\varepsilon) - E(X)E(\varepsilon) = E(XY - XE(Y|X)) - E(X) \times 0$
$= E(XY) - E(XE(Y|X))$
$= E(XY) - E(XY) = 0_\square$

f) $V(\varepsilon) = E(V(Y|X))$

$\hookrightarrow V(Y - E(Y|X)) = V(Y) + V(E(Y|X)) - 2Cov(Y, E(Y|X))$
$= V(Y) + V(E(Y|X)) - 2V(E(Y|X))$
$= V(Y) - V(E(Y|X))_\square$

using variance-decomposition result.

I.D.3. For Special Case where $E(Y|X)$ is linear, i.e., $E(Y|X) = a + bX$

a) $a = E(Y) - bE(X)$

b) $b = Cov(X,Y)/V(X)$ proof: $Cov(X,Y) = Cov(X, E(Y|X))$
$= Cov(X, a + bX)$
$= bV(X)$
$\Rightarrow b = Cov(X,Y)/V(X)_\square$

c) $V(Y) = E(V(Y|X)) + V(E(Y|X))$
$= E(V(Y - a - bX)) + V(a + bX)$
$= E(V(\varepsilon)) + b^2 V(X)$

recall that "giving" X, fixes it & constants don't vary or covary

d) $\rho_{X,Y} = \sqrt{b_{XY} b_{YX}} = \sqrt{\frac{Cov(X,Y)}{V(Y)} \times \frac{Cov(Y,X)}{V(X)}} = \sqrt{\frac{\{Cov(X,Y)\}^2}{V(Y)V(X)}}$

X reg Y Y reg X

$= \frac{Cov(X,Y)}{sd(X)sd(Y)}_\square$

e) $E(Y) = a + bE(X)$; i.e., $\bar{Y} = a + b\bar{X}$

f) "Coefficient of Determination", R^2, $\equiv \{\rho(y, \hat{y})\}^2$