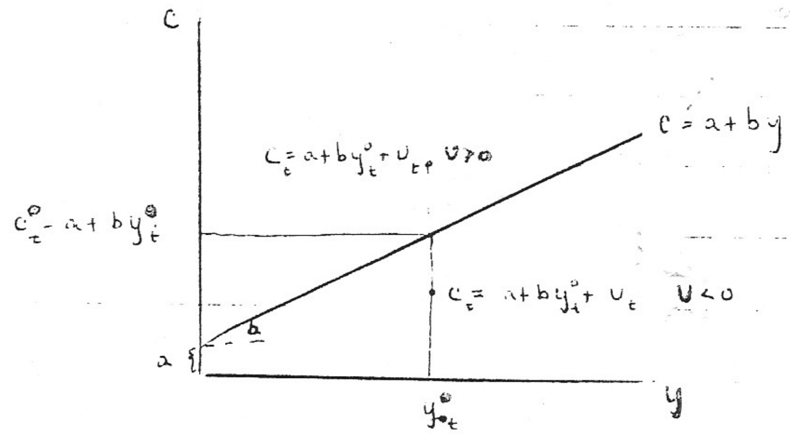
Lecture 2

Introduction to Behavioral Relations

We must postulate some "true" model, say

$$C_t = a + b y_t + U_t$$

The term U_t represents various factors which cause deviations from the average, and is called the disturbance or error term.



Reasons for disturbance term include left out variables, errors in measurement of regressand, "inherent" unpredictability of behavior. But if the model is well specified, the disturbance term should not contain any systematic behavior.

Bivariate Regression Model

Consider the model

$$y_t = a + b x_t + u_t \qquad t = 1, \ldots, T$$

where y is the regressand or dependent variable, and x is the regressor or independent variable. Obviously, y depends on x and u, so we must consider the specification of u.

Assumed Properties of Error

1. $E(u_t) = 0 \quad [= \mu_u]$

2. $E(u_t - \mu_u)^2 = E(u_t^2) = \sigma_u^2$ homoskedasticity
$E(u_t^2 + \underbrace{\mu_u^2}_{=0} - \underbrace{u_t \mu_u}_{=0})$

3. $E(u_t - \mu_u)(u_s - \mu_u) = E(u_t u_s) = \sigma_{u_t u_s} = 0.$

i.e. u_t is uncorrelated w/ u_s for $t \neq s$.

4. $\sigma(u_t x_t) = 0$

Note $\sigma(u_t x_t) = E(u_t - \mu_u)(x_t - \mu_x) = E(u_t x_t)$
$E(u_t x_t) - \mu_x \underbrace{E(u_t)}_{=0} - \underbrace{\mu_u}_{=0} E(x_t) + \underbrace{\mu_u}_{=0} \mu_x$

Comments on Assumptions

1. If $E(u) = \alpha$, including regression constant insures zero mean: E.g.

$$y = a + bx + u$$

define $v = u - \alpha$, $E(v) = 0$. Take

$$y = (a+\alpha) + bx + v$$

$E(u) = \alpha$
$y = bx + u$
$v = u - \alpha$
$y = -\alpha + bx + v$
$E(v) = E(u - \alpha) = 0$

2. Constant variance (homoskedastic errors) is related to all observations being "equally reliable". Note that if σ^2 varies then the info contained in observations regarding true relation also varies.

3. No serial correlation restricts systematic behavior of y to movements in x.

Direct Derivation Using Error Properties

$$\hat{u}_t = y_t - \hat{a} - \hat{b}x_t$$

$$\frac{\sum_t \hat{u}_t}{T} = 0 \Rightarrow \bar{y} = \hat{a} + \hat{b}\bar{x} \qquad (1)'$$

$$\frac{\sum_t \hat{u}_t x_t}{T} = 0 \Rightarrow \frac{\sum x_t y_t}{T} = \frac{\sum \hat{a} x_t}{T} + \frac{\sum \hat{b} x_t^2}{T} \qquad (2)'$$

Note that (1)' and (2)' are identical to (1) and (2). The **normal** equations **are** the **same**, hence so are the implied estimators. Two equations for two unknown parameters, $\hat{a}$ and $\hat{b}$.

Substitute (1) into (2):

$$\frac{\sum x_t y_t}{T} = (\bar{y} - \hat{b}\bar{x})\frac{\sum x_t}{T} + \hat{b}\frac{\sum x_t^2}{T}$$

$$= \bar{y}\frac{\sum x_t}{T} - \hat{b}\bar{x}\frac{\sum x_t}{T} + \hat{b}\frac{\sum x_t^2}{T}$$

$$= \bar{y}\bar{x} - \hat{b}\left(\bar{x}^2 - \frac{\sum x_t^2}{T}\right)$$

hence;

$$\frac{\sum x_t y_t}{T} - \bar{y}\bar{x} = \hat{b}\left(-\bar{x}^2 + \frac{\sum x_t^2}{T}\right)$$

solving for $\hat{b}$:

$$\hat{b} = \frac{\frac{\sum x_t y_t}{T} - \bar{y}\bar{x}}{\frac{\sum x^2}{T} - \bar{x}^2} = \frac{\frac{1}{T}\sum(x_t - \bar{x})(y_t - \bar{y})}{\frac{1}{T}\sum(x_t - \bar{x})^2}$$

Indeed note that:

$$\frac{1}{T}\sum\left(x_t^2+\bar{x}^2-2x_t\bar{x}\right)=$$

$$=\frac{\sum x_t^2}{T}+\underbrace{\frac{1}{T}\sum\bar{x}^2}_{\frac{1}{T}T\bar{x}^2}-2\underbrace{\frac{\sum x_t}{T}}_{\bar{x}}\bar{x}$$

$$=\frac{\sum x_t^2}{T}+\bar{x}^2-2\bar{x}^2$$

$$=\frac{\sum x^2}{T}-\bar{x}^2$$

Similarly:

$$\frac{1}{T}\sum\left(x_t y_t-x_t\bar{y}-\bar{x}y_t+\bar{x}\bar{y}\right)$$

$$=\frac{\sum x_t y_t}{T}-\frac{\bar{y}}{T}\sum x_t-\bar{x}\frac{\sum y_t}{T}+\frac{1}{T}\sum\bar{x}\bar{y}$$

$$=\frac{\sum x_t y_t}{T}-\bar{x}\bar{y}-\bar{x}y+\frac{1}{T}T\bar{x}\bar{y}$$

$$=\frac{\sum x_t y_t}{T}-\bar{x}\bar{y}$$

Least Squares Approach

Consider $\hat{a}, \hat{b}$

$$\hat{u}_t = y_t - \hat{a} - \hat{b} x_t$$

Minimize $\hat{a}, \hat{b}$ $\quad \sum \hat{u}_t^2 = \sum [y - \hat{a} - \hat{b} x]^2$

$$= \sum_t [y_t^2 + \hat{a}^2 + \hat{b}^2 x_t^2 - 2\hat{a} y_t - 2\hat{b} y_t x_t + \hat{a}\hat{b} x_t]$$

Necessary conditions:

$$\frac{\partial}{\partial \hat{a}} = 2 \sum (y_t - \hat{a} - \hat{b} x_t) = 0$$

$$\frac{\partial}{\partial \hat{b}} = 2 \sum (y_t - \hat{a} - \hat{b} x_t)(-x_t) = 0$$

Simplifying:

$$\bar{y} = \hat{a} + \hat{b} \bar{x} \qquad (1)$$

$$\sum x_t y_t = \hat{a} \sum x_t + \hat{b} \sum x_t^2 \qquad (2)$$

Solving first eq for $\hat{a}$

$$\boxed{\hat{a} = \bar{y} - \hat{b} \bar{x}}$$

and substituting into second, dividing by T

$$\frac{\sum x_t y_t}{T} = (\bar{y} - \hat{b}\,\bar{x})\frac{\sum x_t}{T} + \hat{b}\,\frac{\sum x_t^2}{T}$$

Solving for $\hat{b}$

$$\hat{b} = \frac{\dfrac{\sum x_t y_t}{T} - \bar{y}\,\bar{x}}{\sum x_t^2 - \bar{x}^2} = \frac{\frac{1}{T}\sum (x_t - \bar{x})(y_t - \bar{y})}{\frac{1}{T}\sum (x_t - \bar{x})^2}$$

or

$$\boxed{\hat{b} = \frac{\hat{\sigma}_{xy}}{\hat{\sigma}_x^2}}$$

This is the "least squares" estimator for $\hat{b}$.

Note: ① $\hat{b}$ defined only if $\sum (x_t - \bar{x})^2 \neq 0$.

② Since $\sum (x_t - \bar{x})(y_t - \bar{y}) = \sum (x_t - \bar{x})\,y_t$,

$$\hat{b} = \sum w_t y_t, \quad w_t = \frac{x_t - \bar{x}}{\sum (x_t - \bar{x})^2}$$

i.e. $\hat{b}$ is a **linear** function of y, where **weights** depend on x.

as $\bar{y}\,\underbrace{\sum (x_t - \bar{x})}_{\text{deviations from mean sum to zero}} = 0$

Properties of OLS estimators

Unbiasedness

$$\hat{b} = \frac{\sum(x_t-\bar{x})(y_t-\bar{y})}{\sum(x_t-\bar{x})^2} = \frac{\sum(x_t-\bar{x})(\overbrace{a+bx_t+u_t}^{y_t}-\overbrace{a-b\bar{x}}^{\bar{y}})}{\sum(x_t-\bar{x})^2}$$

$$= \frac{b\sum(x_t-\bar{x})^2+\sum(x_t-\bar{x})u_t}{\sum(x_t-\bar{x})^2} = b + \sum\underbrace{\frac{(x_t-\bar{x})}{\sum(x_t-\bar{x})^2}}_{=\omega_t}u_t$$

$$\boxed{E(\hat{b})} = b + \frac{\sum(x_t-\bar{x})E(u)}{\sum(x_t-\bar{x})^2} = b \quad (3)$$

$$\hat{a} = \bar{y}-\hat{b}\bar{x} = \underbrace{a+b\bar{x}}_{\bar{y}}-\hat{b}\bar{x}$$

$$\boxed{E(\hat{a})} = a + b\bar{x} - E(\hat{b})\bar{x} = a \quad (4)$$

Variances of Estimators

1. Variance of $\hat{b}$

Note from (2) we can write

$$\hat{b} = b + \sum w_t u_t \qquad w_t = \frac{x_t-\bar{x}}{\sum(x_t-\bar{x})^2}$$

$$\hat{b}-E(\hat{b}) = \sum w_t u_t$$

Compute variance

Hence, since U_t are independent r.v.'s

$$\mathrm{Var}(\hat{b}) = \overbrace{E(\hat{b}-b)^2}^{E(\hat{b}-E(\hat{b}))^2} = \overbrace{V}^{E\left[\sum w_t u_t\right]^2} = \sum w_t^2 \, E(u_t - \underbrace{E(u_t)}_{=0})^2$$

$$= \sum w_t^2 \sigma_u^2 = \sigma_u^2 \sum w_t^2$$

But

$$\sum w_t^2 = \sum \left[\frac{(x_t-\bar{x})^2}{\underbrace{\left[\sum (x_t-\bar{x})^2\right]^2}_{\text{it is a number}}} \right] = \frac{\sum (x_t-\bar{x})^2}{\left(\sum (x_t-\bar{x})^2\right)^2}$$

$$= \frac{1}{\sum (x_t-\bar{x})^2}$$

Hence

$$\boxed{\mathrm{var}(\hat{b}) \equiv \sigma_{\hat{b}}^2 = \frac{\sigma_u^2}{\sum (x_t-\bar{x})^2}}$$

Variance of $\hat{b}$ depends on Var(u) and on X's

2. Variance of $\hat{a}$

Note: $y = a + bx + u \Rightarrow \bar{y} = a + b\bar{x} + \underbrace{\bar{u}}_{=0} = a + b\bar{x}$

$$\hat{a} = \bar{y} - \hat{b}\bar{x}$$

Substituting $\hat{b} = b + \sum w_t u_t$ $\qquad w_t = \dfrac{x_t - \bar{x}}{\sum (x_t - \bar{x})^2}$

$$\hat{a} = \underbrace{a + b\bar{x} + \bar{u}}_{\bar{y}} - \underbrace{b\bar{x} - \bar{x}\sum w_t u_t}_{\hat{b}\bar{x}}$$

$$= a + \frac{\sum u_t}{T} - \bar{x}\sum w_t u_t$$

$$= a + \sum \gamma_t u_t \qquad \gamma_t = \frac{1}{T} - \bar{x} w_t$$

whence:

$$\therefore \quad \boxed{\sigma^2_{\hat{a}} = \sigma^2_u \sum \gamma_t^2}$$

Consider:

$$\sum \gamma_t^2 = \sum \left[\frac{1}{T^2} + \bar{x}^2 w_t^2 - 2\bar{x} w_t / T \right]$$

$$= \sum \left[\frac{1}{T^2} + \frac{\bar{x}^2 (x_t - \bar{x})^2}{\left[\sum (x_t - \bar{x})^2\right]^2} - 2\frac{\bar{x}}{T} \frac{(x_t - \bar{x})}{\sum (x_t - \bar{x})^2} \right].$$

(*)

$$= \frac{1}{T} + \frac{\bar{x}^2}{\sum (x - \bar{x})^2}$$

since $\sum (x_t - \bar{x}) = 0$.

Simplifying we get:

(**)

$$\circledast \quad \sum \frac{1}{T^2} = T\frac{1}{T^2} = \frac{1}{T}$$

$$\sum \frac{\bar{x}^2(x_t-\bar{x})^2}{\left[\sum(x_t-\bar{x})^2\right]^2} = \frac{\bar{x}^2\sum(x_t-\bar{x})^2}{\left[\sum(x_t-\bar{x})^2\right]^2} = \frac{\bar{x}^2}{\sum(x_t-\bar{x})^2}$$

$$\sum \frac{2\bar{x}}{T}\frac{(x_t-\bar{x})}{\sum(x_t-\bar{x})^2} = \frac{2\bar{x}}{T}\frac{\sum(x_t-\bar{x})}{\sum(x_t-\bar{x})^2} = 0 \text{ as } \sum(x_t-\bar{x}) = 0$$

$$\circledast\circledast \quad \sum \gamma_t^2 = \frac{1}{T} + \frac{\bar{x}^2}{\sum(x_t-\bar{x})^2}$$

$$= \frac{\sum(x_t-\bar{x})^2 + T\bar{x}^2}{T\sum(x_t-\bar{x})^2}$$

$$= \frac{\sum\left(x_t^2+\bar{x}^2-2x_t\bar{x}\right)+T\bar{x}^2}{T\sum(x_t-\bar{x})^2}$$

$$= \frac{\sum x_t^2 + T\bar{x}^2 - 2\bar{x}\sum x_t + T\bar{x}^2}{T\sum(x_t-\bar{x})^2}$$

noting that $\sum x_t = T\bar{x}$

$$= \frac{\sum x_t^2 + T\bar{x}^2 - 2T\bar{x}^2 + T\bar{x}^2}{T\sum(x_t-\bar{x})^2}$$

$$= \frac{\sum x_t^2}{T\sum(x_t-\bar{x})^2}$$

Hence

$$\sigma^2_{\hat{a}} = \sigma^2_u \frac{\sum x_t^2}{T \sum (x_t - \bar{x})^2}$$

Problem: not operational formulas as σ_u^2 not observed: need to estimate it

Variance of disturbance σ_u^2

$\sigma_u^2 = E(u_t^2)$ since $E(u_t) = 0$. Consider:

$$\hat{\sigma}_u^2 = \frac{\sum_t \hat{u}_t^2}{T-2} = \frac{\sum (y_t - \hat{a} - \hat{b} x_t)^2}{T-2}$$

where d.f. = T-2 because two degrees are lost in estimating $\hat{a}$ and $\hat{b}$. It can be shown that

$$E(\hat{\sigma}_u^2) = \sigma_u^2$$

Then we use (replace σ_u^2 with $\hat{\sigma}_u^2$ in $\hat{\sigma}_a^2$ and $\hat{\sigma}_b^2$)

$$\hat{\sigma}^2_{\hat{a}} = \frac{\hat{\sigma}_u^2 \sum x^2}{T \sum (x_t - \bar{x})^2} \; ; \; \hat{\sigma}^2_{\hat{b}} = \frac{\hat{\sigma}_u^2}{\sum (x_t - \bar{x})^2}$$