# INTRODUCTION TO ECONOMETRICS II

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + u$

$\bar{y} = \beta_0 + \beta_1 \bar{x}_1 + \beta_2 \bar{x}_2 + \dots \beta_k \bar{x}_k + \bar{u}$ (Mean)

$y - \bar{y} = \beta_1 (x_1 - \bar{x}_1) + \beta_2 (x_2 - \bar{x}_2) + \dots$

$y = \beta_1 x_1 + \beta_2 x_2 + \cancel{\beta} \dots \beta_k x_k + u$ (small x)

$y = X\beta + u$

$$y = X\beta + u$$

$$u = y - X\beta \qquad u' = (y - X\beta)'$$

* Square errors and then minimize (get derivative the equate to 0)

$$u'u = (y - X\beta)'(y - X\beta)$$

$$u'u = [y' - \beta' X'][y - X\beta]$$

* 2nd & 3rd are transpose of each other
* Are scalars (1x1 matrix)

∴ $y'X\beta$ & $\beta'X'y$ are the same

Rules of transpose $u'u = y'y - \underline{y'X\beta} - \underline{\beta'X'y} + \beta'X'X\beta$ $\longrightarrow$ NEXT PAGE

1 The transpose of the sum is the sum of the transpose

$$(A+B)' = A' + B'$$

$$\underset{A}{\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}} + \underset{B}{\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}} = \begin{bmatrix} 6 & 8 \\ 10 & 9 \end{bmatrix}$$

$$A' = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \quad B' = \begin{bmatrix} 5 & 7 \\ 6 & 5 \end{bmatrix} = \begin{bmatrix} 6 & 10 \\ 8 & 9 \end{bmatrix}$$

2 The transpose of the product is the product of the transpose with the order reversed

$$[AB]' = B'A'$$

$$y' - [X\beta]'$$

$$y' - \beta' X'$$

$$\sum u^2 = \sum (y-\beta_1 x_1-\beta_2 x_2)^2$$

$1(\beta')\beta^{1-1}\ 2(\ )\beta^{2-1}$

$$u'u = y'y - 2\beta'x'y + \beta'x'x\beta$$

$$\frac{\partial u'u}{\partial \beta} = 0 - 2x'y + 2xx'\beta = 0$$

$$= -2x'y + 2xx'\beta = 0$$

$$[x'x]\beta = [x'y] \quad \text{(Multiply by inverse on both sides.}$$

$$\hat{\beta} = [x'x]^{-1}x'y$$

$$\boxed{\hat{\beta} = [x'x]^{-1}x'y}$$

$$\hat{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix} \quad \text{Matrix}$$

$$[xx']^{-1} = \begin{bmatrix} \sum x_1^2 & \sum x_1x_2 \\ \sum x_1x_2 & \sum x_2^2 \end{bmatrix}$$

$$x'y = \begin{bmatrix} \sum yx_1 \\ \sum yx_2 \\ \vdots \\ \sum yx_K \end{bmatrix}$$

Example

| Y | $X_1$ | $X_2$ | $y$ | $x_1$ | $x_2$ |
|---|---|---|---|---|---|
| 10 | 1 | 3 | -10 | -2 | 0 |
| 15 | 2 | 4 | -5 | -1 | 1 |
| 20 | 3 | 6 | 0 | 0 | 3 |
| 25 | 4 | 2 | 5 | 1 | -1 |
| 30 | 5 | 1 | 10 | 2 | -2 |

$\sum Y = 100$   $\sum X_1 = 15$   $\sum X_2 = 16$

$\bar{Y} = 20$   $\bar{X}_1 = 3$   $\bar{X}_2 = 3.2$

$$\hat{\beta} = [x'x]^{-1}x'y$$

$$\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \sum x_1^2 & \sum x_1x_2 \\ \sum x_1x_2 & \sum x_2^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum yx_1 \\ \sum yx_2 \end{bmatrix}$$

$Y = \beta_0 + \beta_1 X_1 + u$

$$\begin{bmatrix} \cancel{N} & \cancel{\Sigma X_1} \\ \Sigma X_1 & \Sigma X_1^2 \end{bmatrix} \quad \begin{bmatrix} \cancel{\Sigma y} \\ \Sigma y X_1 \end{bmatrix}$$

$$[\Sigma X_1^2] = [\Sigma y X_1]$$

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$

$$\begin{bmatrix} \cancel{N} & \cancel{\Sigma X_1} & \cancel{\Sigma X_2} \\ \Sigma X_1 & \Sigma X_1^2 & \Sigma X_1 X_2 \\ \Sigma X_2 & \Sigma X_2 X_1 & \Sigma X_2^2 \end{bmatrix} \quad \begin{bmatrix} \cancel{\Sigma y} \\ \Sigma y X_1 \\ \Sigma y X_2 \end{bmatrix}$$

| $\cancel{Y}$ | $X_1$ | $X_2$ | $y$ | $x_1$ ($X_1-\bar{X}_1$) | $x_2$ ($X_2-\bar{X}_2$) | $yx_1$ | $yx_2$ | $x_1^2$ | $x_1x_2$ | $x_2^2$ | $y^2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 1 | 3 | -10 | -2 | 0 | 20 | 0 | 4 | 0 | 0 | 100 |
| 15 | 2 | 4 | -5 | -1 | 1 | 5 | -5 | 1 | -1 | 1 | 25 |
| 20 | 3 | 6 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 9 | 0 |
| 25 | 4 | 2 | 5 | 1 | -1 | 5 | -5 | 1 | -1 | 1 | 25 |
| 30 | 5 | 1 | 10 | 2 | -2 | 20 | -20 | 4 | -4 | 4 | 100 |
| $\Sigma = 100$ | | | 0 | 0 | 1 | 50 | -30 | 10 | -6 | 15 | 250 |

$\bar{Y} = 20$

$$\hat{\beta} = [X'X]^{-1} X'y$$

$$\hat{\beta} = \begin{bmatrix} \Sigma x_1^2 & \Sigma x_1 x_2 \\ \Sigma x_1 x_2 & \Sigma x_2^2 \end{bmatrix}^{-1} \begin{bmatrix} \Sigma y x_1 \\ \Sigma y x_2 \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

$$= \begin{bmatrix} 10 & -6 \\ -6 & 15 \end{bmatrix}^{-1} \begin{bmatrix} 50 \\ -30 \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

$$\begin{bmatrix} 10 & -6 \\ -6 & 15 \end{bmatrix}^{-1} = (10 \times 15) - 36 = 114 \text{ (Determinant)}$$

$$\text{Cofactor} = \begin{bmatrix} 15 & 6 \\ 6 & 10 \end{bmatrix}$$

$$\frac{1}{114} \begin{bmatrix} 15 & 6 \\ 6 & 10 \end{bmatrix}$$

$$\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \frac{1}{114} \begin{bmatrix} 15 & 6 \\ 6 & 10 \end{bmatrix} \begin{bmatrix} 50 \\ -30 \end{bmatrix}$$

$$\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \frac{15}{114} & \frac{6}{114} \\ \frac{6}{114} & \frac{10}{114} \end{bmatrix} \begin{bmatrix} 50 \\ -30 \end{bmatrix}$$

$$\beta_1 = \left(\frac{15}{114} \times 50\right) + \left(\frac{6}{114} \times -30\right) = 5$$

$$\beta_2 = \left(\frac{6}{114} \times 50\right) + \left(\frac{10}{114} \times -30\right) = 0$$

$\beta_0 = \bar{y} - \beta_1\bar{x}_1 - \beta_2\bar{x}_2$
$= 20 - 5(3) - 0$
$\beta_0 = 5$

Assignment

| $y$ | $x_1$ | $x_2$ | $x_3$ |
|---|---|---|---|
| 10 | 3 | 14 | 1 |
| 30 | 5 | 18 | 2 |
| 12 | 6 | 16 | 3 |
| 18 | 2 | 20 | 4 |
| 4 | 10 | 21 | 5 |

$\hat{\beta} = [x'x]^{-1} x'y$

$$\begin{bmatrix} \Sigma x_1^2 & \Sigma x_1x_2 & \Sigma x_1x_3 \\ \Sigma x_1x_2 & \Sigma x_2^2 & \Sigma x_2x_3 \\ \Sigma x_1x_3 & \Sigma x_2x_3 & \Sigma x_3^2 \end{bmatrix}^{-1} \begin{bmatrix} \Sigma yx_1 \\ \Sigma yx_2 \\ \Sigma yx_3 \end{bmatrix}$$

Example 3

| $Y$ | $X_1$ | $X_2$ | $y_1$ ($Y-\bar{Y}$) | $x_1$ ($X_1-\bar{X}_1$) | $x_2$ ($X_2-\bar{X}_2$) | $yx_1$ | $yx_2$ | $x_1x_2$ | $x_1^2$ | $x_2^2$ | $y^2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 30 | 4 | 10 | 0 | -1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 20 | 3 | 8 | -10 | -2 | -2 | 20 | 20 | 4 | 4 | 4 | 100 |
| 36 | 6 | 11 | 6 | 1 | 1 | 6 | 6 | 1 | 1 | 1 | 36 |
| 24 | 4 | 9 | -6 | -1 | -1 | 6 | 6 | 1 | 1 | 1 | 36 |
| 40 | 8 | 12 | 10 | 3 | 2 | 30 | 20 | 6 | 9 | 4 | 100 |
| $\Sigma$ = 150 | 25 | 50 | | | | 62 | 52 | 12 | 16 | 10 | 272 |
| $\bar{Y}$ = 30 | $\bar{X}_1 = 5$ | $\bar{X}_2 = 10$ | | | | | | | | | |

~~Y = 6 + 0.2X1 − 0.3X2 + 10X3~~

$Y = \beta_0 + \beta_1 X_1 + \ldots \beta_n X_n + u$

* To solve MLR we use the compact form.

$Y = X\beta + u$

$\beta = [x'x]^{-1} x'y$

$$\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \Sigma x_1^2 & \Sigma x_1 x_2 \\ \Sigma x_1 x_2 & \Sigma x_2^2 \end{bmatrix} \begin{bmatrix} \Sigma y x_1 \\ \Sigma y x_2 \end{bmatrix}$$

$$= \begin{bmatrix} 16 & 12 \\ 12 & 10 \end{bmatrix}^{-1} \begin{bmatrix} 62 \\ 52 \end{bmatrix}$$

$\beta_0 = \bar{Y} - \beta_1 \bar{X}_1 - \beta_2 \bar{X}_2$

$= 30 + 0.25(5) - 5.5(10)$

$\beta_0 = -23.75$

$Y = -23.75 - 0.25X_1 + 5.5X_2$

Det $= 160 - 144$

$= 16$

Cofactor $= \begin{bmatrix} 10 & -12 \\ -12 & 16 \end{bmatrix}$

$$\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \frac{1}{16} \begin{bmatrix} 10 & -12 \\ -12 & 16 \end{bmatrix} \begin{bmatrix} 62 \\ 52 \end{bmatrix}$$

$\beta_1 = \frac{10}{16}[62] - \frac{12}{16}[52]$

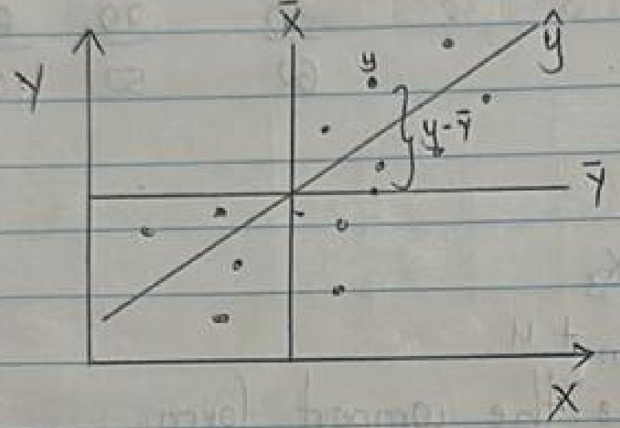
$\beta_1 = -1/4 \quad = -0.25$

$\beta_2 = \frac{-12}{16}[62] + 1[52]$

$\beta_2 = 5\,1/2 \quad = 5.5$

## $R^2$ - Coeffiaent of Determination

- It is a measure of goodness of fit.
- It explains the proportion of the variation of the dependent variable explained jointly by the independent variables.
- If $R^2 = 0.9$, it means, out of 100, 90 is estimated correctly
- A bigger $R^2$ means a bigger goodness of fit.




$\Sigma(y-\bar{y})^2 = TSS$
Total Sum of Squares
$\Sigma(\hat{y}-\bar{y})^2 = ESS$
Explained sum of squares
$\Sigma(y-\hat{y})^2 = RSS$
Residual Sum of squares

- $R^2$ = the proportion of the total variation of the dependent variable explained by the regression value.

$$Y - \bar{Y} = y$$
$$\hat{y} - \bar{y} = \hat{y}$$

$$R^2 = \frac{ESS}{TSS} = \frac{\Sigma(\hat{y})^2}{\Sigma y^2}$$

$$y = \beta_1 X + u$$
$$\hat{y} = \hat{\beta}_1 x_1$$

$$\boxed{R^2 = \frac{\Sigma \hat{y}^2}{\Sigma y^2} = \frac{\Sigma(\beta_1 x_1)^2}{\Sigma y^2} = \frac{\beta_1^2 \Sigma x_1^2}{\Sigma y^2} = \frac{\beta_1 \frac{\Sigma yx}{\Sigma x^2} \cdot \Sigma x^2}{\Sigma y^2} = \frac{\beta_1 \Sigma yx}{\Sigma y^2}}$$

$$TSS = ESS + RSS$$
$$\Sigma y^2 = \Sigma \hat{y}^2 + \Sigma u^2$$
$$ESS = TSS - RSS$$

$$R^2 = \frac{TSS - RSS}{TSS} = 1 - \frac{RSS}{TSS}$$

$$\boxed{R^2 = \frac{\beta_1 \Sigma yx}{\Sigma y^2} \qquad R^2 = 1 - \frac{\Sigma u^2}{\Sigma y^2}}$$

For MLR;

$$R^2 = \frac{\beta_1 \Sigma yx_1 + \beta_2 \Sigma yx_2 + \beta_3 \Sigma yx_3 + \dots + \beta_k \Sigma yx_k}{\Sigma y^2} = \frac{\beta' x' y}{y' y}$$

$$\begin{bmatrix} \beta_1 & \beta_2 & \dots & \beta_k \end{bmatrix} \begin{bmatrix} \Sigma yx_1 \\ \Sigma yx_2 \\ \vdots \\ \Sigma yx_k \end{bmatrix} = \beta' x' y \qquad \begin{bmatrix} y_1 & y_2 & \dots & y_k \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} y' y$$

* $R^2$ is non-decreasing in k. It changes with the independent variables irregardless of whether the independent variable is important.
- Since $R^2$ is biased, it is normally adjusted to reflect its true position.

$\bar{R}^2$ - Adjusted $R^2$

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k}$$ , where k is the no. of independent variables

$$R^2 = \frac{\beta_1 \Sigma yx_1 + \beta_2 \Sigma yx_2}{\Sigma y^2} = \frac{-0.25[\Sigma yx_1] + 5.5[\Sigma yx_2]}{\Sigma y^2}$$

* Highest value of $R^2$ when $ESS = TSS$ $(Y = \hat{y})$

$$R^2 = \frac{TSS}{TSS} \qquad 0 \le R^2 \le 1$$

$$R^2 = 1$$

* Lowest value of $R^2$ when $RSS = TSS$ $(\hat{y} = \bar{Y})$

$$R^2 = 0$$

$$\boxed{R^2 = \frac{\beta_1 \Sigma yx_1 + \beta_2 \Sigma yx_2}{\Sigma y_2}} \qquad \boxed{R^2 = 1 - \frac{\Sigma u^2}{\Sigma y^2}}$$

## Test of Significance

- It shows us whether each independent variable is a good determinant of the dependent variable.
- It is a specific test of hypothesis where; $H_0 : \beta_1 = 0$, $H_1 : \beta_1 \neq 0$, which tests whether the coefficients are individually equal to zero.
- It is a test of mean (t-test) or (z-test) $H_0 : \beta_2 = 0$, $H_1 : \beta_2 \neq 0$

$$t = \frac{\beta_i}{S.E(\beta_i)}$$

$$Var(x) = \frac{\Sigma(x_1 - \bar{x})^2}{N} = E(x_1 - \bar{x})^2 = E[x_1 - E(x)]^2$$

$$\frac{\Sigma x}{N} = \bar{x} = E(x)$$

$$\boxed{Var(\beta) = E[\beta_1 - E(\beta)]^2}$$

Matrix

$$Var(\beta) = E[\beta - E(\beta)][\beta' - E(\beta)']'$$

$$\hat{\beta} = [X'X]^{-1}X'Y$$

$$Y = X\beta + u$$

$$\hat{\beta} = [X'X]^{-1}X'[X\beta + u]$$

$$= \underbrace{[X'X]^{-1}X'X\beta}_{} + [X'X]^{-1}X'u$$

$$= \beta + [X'X]^{-1}X'u$$

$$\hat{\beta} = \beta + [X'X]^{-1}X'u$$

$$E(\hat{\beta}) = \beta + [X'X]^{-1}E(X'u)$$

$$E(X'u) = 0$$

$$E(\hat{\beta}) = \beta + 0$$

$$\hat{\beta} - E(\hat{\beta}) = [X'X]^{-1}X'u$$

$$Var(\hat{\beta}) = \mathcal{E}\left[[X'X]^{-1}X'u\right]\left[[X'X]^{-1}X'u\right]$$

$$Var(\hat{\beta}) = \sigma^2[X'X]^{-1}$$

$$\sigma^2 = \frac{\Sigma u^2}{n-k} = \frac{u'u}{n-k}$$

Var

$$Var(\hat{\beta}) = \mathcal{E}[\hat{\beta} - \mathcal{E}(\beta)]'[\hat{\beta} - \mathcal{E}(\hat{\beta})]^{-1}$$

$$\hat{\beta} = [X'X]^{-1}X'y$$

$$\beta = [X'X]^{-1}X[X\beta + u]$$

$$\underbrace{[X'X]^{-1}X'X}_{I}\beta + [X'X]^{-1}X'u$$

$$\beta + [X'X]^{-1}X'u$$

$$\hat{\beta} = \beta + [X'X]^{-1}X'u$$

$$\mathcal{E}(\hat{\beta}) = \beta + [X'X]^{-1}\mathcal{E}(X'u)$$

$$\mathcal{E}(\hat{\beta}) = \beta$$

$\Sigma x_1^2 = 16$
$\Sigma x_2^2 = 10$
$\Sigma x_1 x_2 = 12$
$\Sigma y x_1 = 62$
$\Sigma y x_2 = 52$
$\Sigma y^2 = 272$

$Var(\hat{\beta}) = \mathcal{E}(\hat{\beta} - \mathcal{E}(\beta)][\hat{\beta} - \mathcal{E}(\beta)]$

$$\hat{\beta} = \begin{bmatrix} 16 & 12 \\ 12 & 10 \end{bmatrix}^{-1} \begin{bmatrix} 62 \\ 52 \end{bmatrix}$$

$Det = [160 - 144]$

* CONTNUATION (Example 3)

$Var(\beta) = \sigma^2 [X'X]^{-1}$

$$\sigma^2 \begin{bmatrix} 10/16 & -12/16 \\ -12/16 & 16/16 \end{bmatrix}$$

$$\sigma^2 \begin{bmatrix} 0.625 & -0.75 \\ -0.75 & 1 \end{bmatrix}$$

$$\sigma^2 = \frac{\Sigma u^2}{n-2}$$

$$R^2 = 1 - \frac{\Sigma u^2}{\Sigma y^2}$$

$$R^2 = \frac{\beta_1 \Sigma y x_1 + \beta_2 \Sigma y x_2}{\Sigma y^2}$$

$$= \frac{-0.25(62) + 5.5(52)}{272}$$

$R^2 = 0.99$

$$\sigma^2 = \frac{\Sigma u^2}{5-2}$$

$\Sigma u^2 = (1-R^2)\Sigma y^2 = 0.01(272)$

$\Sigma u^2 = 2.72$

$$\sigma^2 = \frac{2.72}{3} = 0.9066$$

$$Var(\beta) = \sigma^2 \begin{bmatrix} 0.625 & -0.75 \\ -0.75 & 1 \end{bmatrix}$$

$$Var(\beta) = 0.906 \begin{bmatrix} 0.625 & -0.75 \\ -0.75 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0.56 & -0.68 \\ -0.68 & 0.906 \end{bmatrix}$$

$$Var(\beta) = \begin{bmatrix} var\,\beta_1 & cov(\beta_1, \beta_2) \\ Cov(\beta_1, \beta_2) & var\,\beta_2 \end{bmatrix}$$

$Var\,\beta_1 = 0.56$
$Var\,\beta_2 = 0.906$

$R^2 = 0.99$
$y = -0.25x_1 + 5.5x_2$

$S.E\,\beta_1 = \sqrt{var\beta_1} = \sqrt{0.56}$
$S.E\,\beta_1 = 0.75$

$S.E\,\beta_2 = \sqrt{var\beta_2} = \sqrt{0.906}$
$S.E\,\beta_2 = 0.95$

t-value

$$t_1 = \frac{\beta_1}{S.E\beta_1} = \frac{-0.25}{0.75}$$

$t_1 = -0.33$

$$t_2 = \frac{\beta_2}{S.E\beta_2} = \frac{5.5}{0.95} \quad t_2 = 5.8$$

Criteria.
1. Significant.
2. Goodness of fit
3. Sign of coefficient.

$R^2 = 0.99$ means that 99% of the variations in y is jointly explained by $x_1$ and $x_2$

* If $t \geq 2$, reject null hypothesis. or if $p \leq \alpha$ (significance level), reject null hypothesis
5% or 0.05
* $t_{computed} > t_{\alpha/2}$, reject null hypothesis.

| Variable | Coeff | S.E | t-value | Range |
|---|---|---|---|---|
| $X_1$ | −0.25 | 0.75 | −0.33 | |
| $X_2$ | 5.5 | 0.95 | 5.8 | |

* $X_1$ is not an important determinant of y.

* ~~In~~

$H_0: \beta_0 = 0$
$H_1: \beta_1 \neq 0$
$t < 2$ Do not reject null hypothesis.
$X_1$ is not an important determinant of y

$H_0: \beta_0 = 0$
$H_2: \beta_1 \neq 0$
$t > 2$ reject null hypothesis.
$X_2$ is an important determinant of y.

$R^2$ (sign of coefficient)
- If the sign of the coefficient is negative, it implies an inverse relationship.
- If the coefficient is positive, then the dependent and independent variables are directly related.
e.g $X_2$ has a direct r/ship with y (if $X_2$ goes up, y goes up).

Properties of OLS estimates ($\hat{\beta}$)

1. Linearity $\hat{\beta} = [X'X]^{-1}X'u$
2. Unbiasedness $E(\hat{\beta}) = \beta$
3. They have minimum variance.
   - OLS gives the smallest variance among others.

Q. Prove minimum variance.
Q. Prove unbiasedness.

Example 4

| $y$ | $X_1$ | $X_2$ |
|---|---|---|
| 10 | 1 | 8 |
| 12 | 2 | 7 |
| 15 | 3 | 4 |
| 17 | 4 | 2 |
| 3 | 5 | 10 |
| 6 | 6 | 7 |

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

1. Find $\beta_0, \beta_1, \beta_2$.

2. Test whether $\beta_1$ and $\beta_2$ are equal to zero

Proof of minimum variance:

$y = \beta x + u$

$\hat{\beta} = \frac{\Sigma yx}{\Sigma x^2}$

$var\,\hat{\beta} = \frac{\sigma^2}{\Sigma x^2}$

Let $\beta^*$ be linear and unbiased ($\mathcal{E}(\beta^*) = \beta$)

$$\beta^* = \frac{\Sigma (x+d)y}{\Sigma x^2}$$

$$\beta^* = \frac{\Sigma (x+d)[\beta x + u]}{\Sigma x^2}$$

$$= \frac{\beta \Sigma x^2 + \Sigma xu + \Sigma d\beta x + du}{\Sigma x^2}$$

$$\mathcal{E}(\beta^*) = \beta + \frac{\Sigma (xu) + \Sigma d\beta x + d\mathcal{E}u}{\Sigma x^2}$$

$$\mathcal{E}(\beta^*) = \beta + \frac{d\beta \Sigma x}{\Sigma x^2}$$

$\mathcal{E}(\beta^*) = \beta$ if & only if $dx = 0$

$$Var(\beta^*) = \mathcal{E}[\beta^* - \mathcal{E}(\beta^*)]^2$$

$$Var(\beta^*) = \mathcal{E}\left[\frac{\Sigma xu + du}{\Sigma x^2}\right]^2$$

$$= \left[\frac{\Sigma xu}{\Sigma x^2} + \frac{du}{\Sigma x^2}\right]^2$$

$$= \left[\frac{\Sigma xu}{\Sigma x^2}\right]^2 + \left[\frac{du}{\Sigma x^2}\right] + 2\frac{\Sigma xu}{\Sigma x^2}\frac{du}{\Sigma x^2} = 0$$

$$Var(\beta^*) = \mathcal{E}\frac{[u^2 \Sigma x^2]}{\Sigma}$$

$$Var(\beta^*) = \frac{\sigma^2}{\Sigma x^2} + \text{positive}$$

$$Var(\beta^*) = Var(\beta)_{OLS} + \text{positive}$$

$$Var(\beta)_{OLS} < Var(\beta^*)$$

Using Matrix

$$\beta^* = [x'x]^{-1}(x'+D)Y$$

$$\beta^* = [x'x]^{-1}(x'+D](x\beta+u)$$

$$= [x'x]^{-1}x'x\beta + [x'x]^{-1}x'u + Dx\beta + Du$$

$$\beta^* = \beta + [x'x]^{-1}x'u + Dx\beta + Du \quad \ldots \text{(i)}$$

$$\mathcal{E}(\beta^*) = \beta + [x'x]^{-1}\underbrace{\mathcal{E}(x|u)}_{0} + Dx\beta + D\underbrace{\mathcal{E}(u)}_{0}$$

$$\mathcal{E}(\beta^*) = \beta + Dx\beta \quad \ldots \text{(ii)}$$

$\mathcal{E}(\beta^*) = \beta$ if and only if $Dx = x'D' = 0$

$$Var(\beta^*) = \mathcal{E}[\underbrace{\beta^* - (\mathcal{E}\beta^*)}_{eq.(i) - eq(ii)}) (\beta^* - \mathcal{E}(\beta^*))^{-1}]$$

$$\cancel{Var}$$

$$Var(\beta^*) = \mathcal{E}[[X'X]^{-1}X'u + bu][u'X[X'X]^{-1} + u'b']$$

*Transpose, you write in reverse order.

$*bu + u'b = 0$

$$= \mathcal{E}[[X'X]^{-1}X'\underbrace{uu'}_{\sigma^2}X[X'X]^{-1} + b\underbrace{uu'}_{\sigma^2}b'$$

$$Var(\beta^*) = \sigma^2[X'X]^{-1}\underbrace{X'X[X'X]^{-1}}_{\text{Identity matrix}} + \sigma^2 bb'$$

$$Var(\beta^*) = \sigma^2[X'X]^{-1} + \sigma^2 bb'$$

$Var(\beta^*) = Var\,\hat{\beta}_{OLS}$ + positive

$$Var(\hat{\beta}^*_{OLS}) < Var\,\beta^*$$

TOPIC 2 / SIMULTANE Concept of unbiasedness

$y = \beta_1 x + u$

$\hat{\beta} = \frac{\sum yx}{\sum x^2}$

$\hat{\beta} = \frac{\sum x[\beta x + u]}{\sum x^2} = \frac{\beta \sum x^2 + \sum xu}{\sum x^2}$

$\hat{\beta} = \beta + \frac{\sum xu}{\sum x^2}$

$E(\hat{\beta}) = \beta + \frac{\sum \overbrace{E(xu)}^{0}}{\sum x^2}$

$E(\hat{\beta}) = \beta$

$y = X\beta + u$

$\beta = [X'X]^{-1} X'y$

$[\sum x^2]^{-1} \sum yx$

$\hat{\beta} = [X'X]^{-1} X'[X\beta + u]$

$\hat{\beta} = \underbrace{[X'X]^{-1} X'X}_{I} \beta + [X'X]^{-1} X'u$

$\hat{\beta} = \beta + [X'X]^{-1} X'u$

$E(\hat{\beta}) = \beta + [X'X]^{-1} E\underbrace{(X'u)}_{0}$

$E(\hat{\beta}) = \beta$

* $u \sim IID(0, \sigma^2)$

$E(u) = 0$

$E(xu) = 0$

- Simultaneous equations - a group of equations that give complete relationship btwn econ. variables

TOPIC 2: SIMULTANEOUS EQUATION MODELS

- A single equation model $y = f(x)$
- When we apply regression to a single equation, we assume:
  i) One-way causation i.e it is x influencing y and not vice-versa.
- In economic theory, there are so many cases of 2 way causalities.
  ii) The complete relationship between economic variables can be summarized by one equation.
- In macro economics, economic variables are summarized by ~~more~~ several equations. $C = Y$ $I = Y$ $Y = C + I + G$
- When we use regression, we make very stringent assumptions that are violated.

What are the consequences?

The consequences of running a regression on a single equation when there is more than one way causality is simultaneity bias.

- How then do we estimate simultaneous equations?
  i) Identify the equations
  ii) Estimation methods

* Consider the following definitions:

1. Endogenous variables - its a variable whose values is determined within the model (explained variable)
   - Occupy the left hand side of the equation. ($C$, $I$, $Y$)
2. Exogenous variables - is a variable whose values is determined outside that particular equation (explanatory variable)
   - They cannot occupy the right hand-side of the equation.
   - A lagged value of an endogenous variable is considered as exogenous: $G_t$, $Y_{t-1}$
3. Structural model - it expresses endogenous variables as a function of other endogenous variables, exogenous variables (pre-determined) and the disturbances (errors). $y = f(Y, X, u)$

$C_t = a_0 + a_1 Y_1 + u_i$
$I_t = b_0 + b_1 Y_1 + b_2 Y_{t-1} + u_2$
$Y_t = C_t + I_t + G_t$

- A structural model is informed by theory
- The ~~parameters~~ in coefficient of the structural equation, the parameters in general are considered as propensity, elasticity or any other and are generally direct effects $(a_0, a_1, b_1 b_2)$
- They are called direct effects because they express direct effects of each explanatory variables.

4 Reduced form model - Endogenous variables are expressed as a function of pre-determined (exogenous) variables only.

- Endogenous variables is expressed as a function of all exogenous variables in the model. $y_1 = f(x, x_2, u)$

$$C_t = \Pi_{01} + \Pi_{11} Y_{t-1} + \Pi_{21} G_t + V_1$$

Q. Express the consumption and investment function in their reduced form

$$I_t = \Pi_{02} + \Pi_{12} Y_{t-1} + \Pi_{22} G_t + V_2$$

$$Y_t = \Pi_{03} + \Pi_{13} Y_{t-1} + \Pi_{23} G_t + V_3$$

$$Y_t = C_t + I_t + G_t$$

$$Y_t = a_0 + a_1 Y_t + u_1 + b_0 + b_1 Y_t + b_2 Y_{t-1} + u_2 + G_t$$

$$Y_t - a_1 Y_t - b_1 Y_t = (a_0 + b_0) + b_2 Y_{t-1} + G_t + \underbrace{(u_1 + u_2)}_{V}$$

$$Y_t = \underbrace{\left(\frac{a_0 + b_0}{1 - a_1 - b_1}\right)}_{\Pi_{03}} + \underbrace{\left(\frac{b_2}{1 - a_1 - b_1}\right)}_{\Pi_{13}} Y_{t-1} + \underbrace{\left[\frac{1}{1 - a_1 - b_1}\right]}_{\Pi_{23}} G_t + V$$

- The parameters of the reduced form are called total effec[ts] because they are composed of direct and indirect effects
- This is because they capture direct and indirect change [of] pre-determined (exogenous) variables on endogenous variables

5 Recursive model - In a recursive model, the first endogenous variable is expressed as a function of exogenous variables only. The second endogenous variables is expressed as a function of the first ~~exogenous~~ endogenous variable and exogenous variables only.

$$y_1 = f(x_1, x_2 \ldots x_{n-1}, u)$$

$$y_2 = f(y_1, x_1, x_2 \ldots x_{n-1}, u)$$

$$y_3 = f(y_1, y_2, x_1 \ldots x_{n-1}, u)$$

$$y_n = f(y_1, y_2, \ldots y_{n-1}, x_1, x_2 \ldots, u)$$

- Recursive model A.k.A triangular systems.

1) Identify the equation.
- An equation is identified if it is in a unique structural form enabling unique estimates from its parameters to be subsequently made from a sample data.
- An equation is not identified if estimates of parameters of relationships between variables measured in samples may relate to the model in the equation to another model or a mixture of models.
- For any regression $Y = \xi X\beta + u$, independent variables should not be zero.
- $\mathcal{E}(x, u) = 0$ expected value of coefficients indicates biased

$$C_t = \alpha_0 + \alpha_1 Y_t + u_1$$
$$I_t = \beta_0 + \beta_1 Y_t + \beta_2 Y_{t-1} + u_2$$
$$Y_t = I_t + C_t + G_t$$

- $C_t$, $I_t$ and $Y_t$ are endogenous variables. A lagged value is any variable on the right hand side.
- An equation can be identified or over identified.

Identification.

- This is knowing which equation is from a set of similar equation. Demand and supply are both functions of price.
- Which equation is which when given a set of equations.

1. A paradox of identification is used. We expose the demand curve to supply curve. If the demand curve shifts slightly then this is a demand curve. If it continuously changes upon applying supply variables then it is a mongrel.

2. If you suspect it is a supply curve, apply demand variables. If there is no change then it is not sensitive to demand variables therefore it is a supply curve.

* Estimatability is the second level of identification. We already have this definition from the previous topic. It can be exactly identified, over-identified or unidentified. An equation is over-identified if we can still get estimates.
- If we find unique estimates then it is exactly identified

- Two conditions for identification:
  i) ORDER condition.
  ii) RANK condition.

- Order condition is a neccessary condition. Rank is a sufficient condition
- A sufficient condition is a proof. The rank condition is a proof so we consider it more than the order.

i) Order Condition
- An equation is identified if and only if the total no. of variables excluded from it must be equal to or greater than the no. of endogenous variables in the model less 1.

Let k - no. of variables in the model.
Let $k_i$ - no. of variables in equation i
M - no. of variable of endogenous variables in the model

$$K - k_i \geq M - 1$$

$$k - k_i = M - 1 \quad \text{(Exatly (just)}$$

$$K - k_i > M - 1 \quad \text{(Over identified)}$$

$$K - k_i < M - 1 \quad \text{(Not identified)}$$

Example.

$$C_t = \alpha_0 + \alpha_1 Y_t + u_1 \quad \ldots \text{(i) (Over identifified)}$$
$$I_t = \beta_0 + \beta_1 Y_t + \beta_2 Y_{t-1} + u_2 \quad \ldots \text{(ii) (Exactly identified)}$$
$$Y_t = I_t + C_t + G_t \quad \ldots \text{(iii) (Not identified)}$$

$C_t, Y_t, I_t$ = Endogenous

$Y_{t-1}, G_t$ = Exogenous

$k = 5$
$M = 3$

(i)
$k - k_i \qquad M - 1$
$5 - 2 \qquad 3 - 1$
$3 \quad > \quad 2$

(ii)
$k - k_i \qquad M - 1$
$5 - 3 \qquad 3 - 1$
$2 \quad = \quad 2$

(iii)
$k - k_i \qquad M - 1$
$5 - 4 \qquad 3 - 1$
$1 \quad < \quad 2$

Example.

$k = 5$

$$C_t = a_0 + a_1 Y_t - a_2 I_t + u_1$$
$$I_t = b_0 + b_1 Y_t + u_2$$
$$T_t = c_0 + c_1 Y_t + u_3$$
$$Y_t = C_t + I_t + G_t$$

$C_t, I_t, T_t, Y_t$ = Endogenous

$G_t$ = Exogenous

i) $k - k_i \qquad M - 1$
$5 - 3 \qquad 4 - 1$
$2 \quad < \quad 3$ (~~Exactly identified~~) Not identified.

ii) $k - k_i \qquad M - 1$
$5 - 2 \qquad 4 - 1$
$3 \quad = \quad 3$ (~~Over~~ Exactly identified)

iii) $k - k_i$ $\quad$ $M-1$

$5-2$ $\quad$ $4-1$

$3 = 3$ $\quad$ Exactly identified.

iv) $k - k_i$ $\quad$ $M-1$

$5-4$ $\quad$ $4-1$

$1 < 3$ $\quad$ Not identified

* Equation (ii) and (iii) can be estimated using econometric methods.

ii) Rank Condition

- In a system of m equations any particular equation is identified if and only if it is possible to construct atleast one non-zero determinant of order M-1 from coefficients of variables excluded from that particular equation but contained in other equations of the model

Example 1

$$C_t = \alpha_0 + \alpha_1 Y_t + u_1 \quad \dots \quad (i)$$

$$I_t = \beta_0 + \beta_1 Y_t + \beta_2 Y_{t-1} + u_2 \quad \dots \quad (ii)$$

$$Y_t = C_t + I_t + G_t \quad \dots \quad (iii)$$

Step I: Form a matrix of parameters of all equations.

| $C_t$ | $Y_t$ | $I_t$ | $Y_{t-1}$ | $G_t$ |
|---|---|---|---|---|
| ~~-1~~ | ~~$\alpha_1$~~ | ~~0~~ | ~~0~~ | ~~0~~ |
| 0 | $\beta_1$ | -1 | $\beta_2$ | 0 |
| 1 | -1 | 1 | 0 | 1 |

Step II: ~~Straight~~ Strike out the row representing the equation to be identified.

Step III: Strike out the columns in which non-zero coefficients appear on the basis of the equation to be identified.

Equation (i)

$$
\begin{array}{ccccc}
C_t & Y_t & I_t & Y_{t-1} & G_t \\
1 & \alpha_1 & 0 & 0 & 0 \\
0 & \beta_1 & -1 & \beta_2 & 0 \\
1 & -1 & 1 & 0 & 1
\end{array}
$$

Step IV: From ii and iii, form as many 2x2 matrices

$\begin{vmatrix} -1 & \beta_2 \\ 1 & 0 \end{vmatrix} \neq 0$ $\begin{vmatrix} -1 & 0 \\ 1 & 1 \end{vmatrix} \neq 0$ $\begin{vmatrix} \beta_2 & 0 \\ 0 & 1 \end{vmatrix} \neq 0$ get the determinant

* Equation (i) has more than one non-zero on the basis of determinants, therefore it is identified.

Equation (ii)

$$
\begin{array}{ccccc}
C_t & Y_t & I_t & Y_{t-1} & G_t \\
1 & \alpha_1 & 0 & 0 & 0 \\
0 & \beta_1 & -1 & \beta_2 & 0 \\
1 & -1 & 1 & 0 & 1
\end{array}
$$

$\begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} \neq 0$ Exactly identified; one non-zero determinant ~~that is~~

Equation (iii)

$$
\begin{array}{ccccc}
C_t & Y_t & I_t & Y_{t-1} & G_t \\
1 & \alpha_1 & 0 & 0 & 0 \\
0 & \beta_1 & -1 & \beta_2 & 0 \\
1 & -1 & 1 & 0 & 1
\end{array}
$$

$\begin{bmatrix} 0 \\ \beta_2 \end{bmatrix}$ * It is not possible to form a 2 X 2 matrix
* Therefore, it is not identified.

Example 2.

$C_t = a_0 + a_1 Y_t - a_2 T_t + u_1$

$I_t = b_0 + b_1 Y_t + u_2$

$T_t = c_0 + c_1 Y_t + u_3$

$Y_t = C_t + I_t + G_t$

| $C_t$ | $Y_t$ | $T_t$ | $I_t$ | $G_t$ |
|---|---|---|---|---|
| -1 | $a_1$ | $-a_2$ | 0 | 0 |
| 0 | $b_1$ | 0 | -1 | 0 |
| 0 | $c_1$ | -1 | 0 | 0 |
| 1 | -1 | 0 | 1 | 1 |

* Form as many 3x3 matrices as possible.

Eq. (i)

| $C_t$ | $Y_t$ | $T_t$ | $I_t$ | $G_t$ |
|---|---|---|---|---|
| -1 | $a_1$ | $-a_2$ | 0 | 0 |
| 0 | $b_1$ | 0 | -1 | 0 |
| 0 | $c_1$ | -1 | 0 | 0 |
| 1 | -1 | 0 | 1 | 1 |

$$\begin{bmatrix} -1 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} \quad 3 \times 2$$

* Not possible to form 3X3 matrices
* Hence, not identified.

Eq. (ii)

| $C_t$ | $Y_t$ | $T_t$ | $I_t$ | $G_t$ |
|---|---|---|---|---|
| -1 | $a_1$ | $-a_2$ | 0 | 0 |
| 0 | $b_1$ | 0 | -1 | 0 |
| 0 | $c_1$ | -1 | 0 | 0 |
| 1 | -1 | 0 | 1 | 1 |

+1

$$\begin{bmatrix} -1 & -a_2 & 0 \\ 0 & -1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

* Only one 3X3 matrix

Get the determinant.

$$-1\begin{vmatrix}-1 & 0\\ 0 & 1\end{vmatrix} + a_0\begin{vmatrix}0 & 0\\ 1 & 1\end{vmatrix} + 0\begin{vmatrix}0 & -1\\ 1 & 0\end{vmatrix}$$

$$\qquad 1 \qquad\qquad 0 \qquad\qquad 0 \qquad \neq 0$$

$\neq 0$, so it can be identified.

Estimation Methods

* 1. Indirect Least Squares
* 2. Instrumental Variable
* 3. Two ~~squared~~ stage least square
4. Limited information Maximum Likelihood

] Single equation method.

5. Full information Maximum Likelihood
6. Three stage Least square

] System methods

7. Mixed Method

1. Indirect Least Squares

- It is applied only to equations that are exactly identified
- The procedure is;
  - i) Express that particular equation in its reduced form
  - ii) Apply OLS in its estimation.
  - iii) Compute the direct effect.

Properties of ILS coefficients.

1. ~~Unbiased~~ Coefficients of total effects are unbiased.

$$\mathcal{E}(\hat{\Pi}_i) = \Pi_i$$

2. In small samples, $\mathcal{E}(\hat{\beta}) \neq \beta$ $\therefore$ the direct effects are biased in small numbers.

Criticisms

1. Can only be used when the equation is exactly identified.
2. And it is biased in small samples.

2. Instrumental Variable Method.

- It can be used when an equation is ~~e~~ either identified or over identified.
- The method involves replacing the endogenous variables on the right hand side with a variable we call an instrument then run a regression

$C_t = a_0 + a_1 Y_t + u$ Over identified

Replace;

$C_t = a_0 + a_1 Z_t + u$

$Z_t$ = an instrument.

Qualities of a good instrument

1. Should not be related to the error term
$\mathcal{E}(z:u) = 0$
2. All the good qualities must be retained
$\mathcal{E}(y:u) \neq 0$
3. Should not be strongly correlated to the other variables; no multi-collinearity.

* ILS method on its own provides us with the instrument. hints to us how to get an instrument, but gives biased results.
* IV method guarantees us good results if we get an ~~go~~ instrument, but does not provide us with the instrument.
* Therefore, we combine both ILS and IV method to get the Two Stage Least Square Method

$C_t = a_0 + a_1 Y_t + u$

Stage I: Express endogenous variable on the RHS in its reduced form and run a regression (use OLS)

$Y_t = \Pi_0 + \Pi_1 Y_{t-1} + \Pi_2 G_t + u$

$Y = 10 + 0.2 Y_{t-1} + 0.4 G_t$

~~Stage II~~: Substitute log of $\hat{Y}$

Stage II: Use the new variable in your equation then run a regression.

# QUALITATIVE RESPONSE REGRESSION MODELS

- Dependent variable is not qualitative.

Presentations of data.

Quantitative { 1. Ratio. $y_1/y_2$
2. Interval $y_1 - y_2$

Qualitative { 3. Ordinal $y_1 > y_2$
4. Nominal - lowest level, distinguishes terms

* Arithmetic mean can be used in ratio and interval.
- Regression method assumes dependent variable is ratio or scale, can be quantified.
- Dependent variable is continuous (assumption).

- Qualitative response models refers to a situation where the regressand (dependent, explained, endogenous) is qualitative in nature. (or the regressand is the dummy variable)
- Suppose we know that the dependent variable is binary, what is the problem in the regression method? The problem is in the characteristics
- When the regressand is qualitative, there is need for special regression methods which include:
  - i) Linear Probability Model.
  - ii) Logit.
  - iii) Probit.
  - iv) Tobit.

## 1. Binary Models.

- This is where the regressand is qualitative but has only 2 choices: the event occurring or not occurring.

| Y | X |
|---|---|
| 0 | 8 |
| 1 | 16 |
| 1 | 18 |
| 0 | 11 |
| 0 | 12 |
| 1 | 19 |
| 1 | 20 |
| 0 | 9 |

$\Sigma Y = 4$

$\mathcal{E}(Y) = \frac{\Sigma Y}{n}$

$= 4/8 = 0.5$

* We replace Y with the probability (P) to make it continuous so you can run a line.

- What is the probability that somebody chosen randomly has a house? 0.5

Linear Probability Model

$Y = \beta_0 + \beta_1 X + u$

$\mathcal{E}[Y|X] = \beta_0 + \beta_1 X$ ; assume $E(u) = 0$

* If $P_i$ = probability that $Y_i = 1$ (event occurs) and $(1-P_i)$ = probability that $Y_i = 0$ (event does not occur), the variable $Y_i$ has the following probability distribution:

| Y | Probability |
|---|---|
| 0 | $1-P_1$ |
| 1 | $P_1$ |

$E(Y_i) = 0(1-P_i) + 1(P_i) = P_i$

$E(Y/X) = \beta_1 + \beta_2 X_i = P_i$

$\mathcal{E}(Y) = 0(1-P_1) + 1(P_1)$

$\mathcal{E}(Y) = P_1$

* The mean of the binomial distribution is $np$
* Variance $= np(1-p)$

* Convert Y into the expected values and then run a regressio

-

Characteristic of LPM coefficient

1. Non-normality of the error term - the error is not normally distributed; it follows a Bernoulli distribution.
2. The error term is heteroskedastic. (the variance is not constan can be solved
3. The slope is constant. (the change in probability is going to b the same at all points of x)

4. Non-fulfillment of the order $0 \leq E(y|x) \leq 1$
- solved by truncation.

* If we have binary choice models, we can use LPM.

5. $R^2$ may not be a good measure of goodness of fit.

- ~~Suppose~~;

ii) Logit. Model

$$P_i = \frac{1}{1+e^{-z}} \quad ; z = \beta_0 + \beta_1 X$$

$$\ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X + u$$

X - income bracket

N - no. of people in the income bracket

n - no. of people who own houses

| $\ln\left(\frac{P}{1-P}\right)$ | $\left(\frac{P}{1-P}\right)$ | 1 - P | X | N | n | P |
|---|---|---|---|---|---|---|
| | | 1 - 8/40 | 6 | 40 | 8 | 8/40 |
| | | | 8 | 50 | 12 | 12/50 |
| | | | 10 | 60 | 18 | 18/60 |
| | | | 13 | 80 | 28 | 28/80 |
| | | | 15 | 100 | 45 | 45/100 |
| | | | 20 | 70 | 36 | 36/70 |
| | | | 25 | 65 | 39 | 39/65 |

$\frac{P}{1-P}$ is an odds ratio. Ratio of occurrence divided by ratio of non-occurrence

- So we take the logs of both sides

$$\ln\left(\frac{P}{1-P}\right) = z \ln e = Z = \beta_0 + \beta_1 X$$

$$\ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X + u$$

$L = \beta_0 + \beta_1 X_1 + u$ where $L$ is the log of the odds ratio.

Features of Logit Model

1. As P goes from 0 to 1, the logit goes from $-\infty$ to $\infty$. So P is actually bounded.

2. Although L is linear in X, the probabilities are not linear. There is no constant slope. For every change in X, there will be a different probability

If $L = 0.4 + 0.002X$, you have to translate it to $\frac{\partial P}{\partial X}$

3. If L is positive, the value of a regressor increases the odds that that regressant equals 1 also increases

- So we interpret a logit model in terms of:
  i) sign.
  ii) goodness of fit but not $R^2$ but pseudo $R^2$.
  iii) convert the slope coefficient into a change of probability.
- We take the log of the odds ratio
- We can substitute and get P for every value.

$$L = 0.4 + 0.02(100)$$
$$= 2.4$$

* 2.4 is an increase in odds

- So $\ln\left(\frac{P}{1-P}\right) = 2.4$
- So we take the anti log of 2.4
- The anti-log of 2.4 = 251

$$\left(\frac{P}{1-P}\right) = 251 \times (1-P)$$

$$P = 251 - 251P$$

$$\frac{252P}{252} = \frac{251}{252}$$

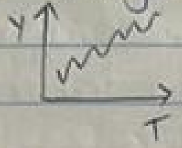
$$P = 0.99$$

Estimation of the Logit Model

1. The method used is maximum likelihood and not OLS. Because P can be 1 or 0. This can lead to us finding log of 0 or log 100 which we do not know.

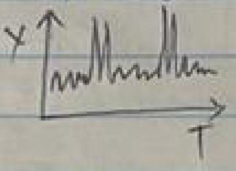
Facts about Time Series Data (Components)

1. Trend - average changes in variable per unit time.

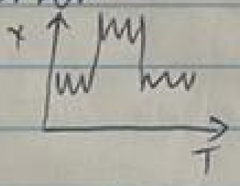



2. Seasonal variations - periodic variations that occur with some degree of regularity.

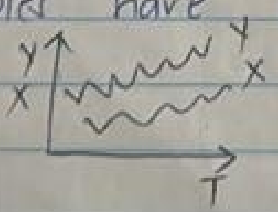

3. Cyclical variations - recurring up and down movements which are extended over a period of time (volatility clust




4. Irregular variations - random fluctuations that happen due to error.




5. Some variables have co movements



- From the components of the time series data, some of the assumptions of OLS may be violated. As a result, we may have spurious / nonsensical results.
(spurious - looking good but not mean much)

Stationarity and non-stationarity.
- The time-series variable is said to be stationary if all the moments are constant, meaning, it has constant mean, constant variance and constant covariance.
- A variable is non-stationary if any of these moments are not constant.
Weak - covariance does not depend on time but depends on time differen
Strong - covariance does not depend on time and does not depend on time difference.

How do we model time series data?
- Models of time-series data:
i) AR - Autoregressive.
ii) MA - Moving Average.
iii) ARMA - Autoregressive Moving Average.
iv) ARIMA - Autoregressive Integrated Moving Average.

i) AR.
- A variable is modeled as a function of its first behaviours and the error term.

$$Y_t = \rho_0 + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \ldots \rho_q y_{t-q} + u_t \quad \ldots\ldots AR(q)$$

$$Y_t = \rho_0 + \rho_1 y_{t-1} + u_t \quad \ldots\ldots AR(1)$$

$$Y_t = \rho_0 + \rho_1 y_{t-1} + \rho_2 y_{t-2} + u_t \quad \ldots\ldots AR(2)$$

$$Y_t = \sum_{i=1}^{q} \rho_i Y_{t-i}$$

ii) MA.

- A variable is modeled as a function of the current and previous error terms.

$$Y_t = \theta_0 \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_p \varepsilon_{t-p}$$

$$Y_t = \sum_{i=0}^{p} \theta \varepsilon_{t-i}$$

$$Y_t = \theta \varepsilon_t + \theta_1 \varepsilon_{t-1} \quad \dots \text{MA}(1)$$

iii). ARMA

- A variable is modeled as a function of its first behaviour, its past and current error terms.

$$Y_t = \rho_1 y_{t-1} + \dots + \rho_q y_{t-q} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_p \varepsilon_{t-p}$$

$$Y_t = \sum_{i=1}^{q} \rho Y_{t-i} + \sum_{i=1}^{p} \theta_i \varepsilon_{t-i} + \varepsilon_t$$

Stationarity - $E(y)$ - constant, $Var(y)$ - constant, $cov(Y_t, Y_{t-1})$ - constant.

$$Y_t = \rho y_{t-1} + \varepsilon_t$$

$$0 \leq \rho \leq 1$$

$$Y_t = y_{t-1} + \varepsilon_t$$

$$Y_1 = y_0 + \varepsilon_1$$

$$Y_2 = y_1 + \varepsilon_2 = y_0 + \varepsilon_1 + \varepsilon_2$$

$$Y_3 = y_2 + \varepsilon_3 = y_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3$$

$$Y_4 = y_3 + \varepsilon_4 = y_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \varepsilon_4$$

$$Y_n = y_0 + \sum_{i=1}^{n} \varepsilon_i$$

$$E(Y) = E[y_0 + \varepsilon_1 + \varepsilon_2 + \dots \varepsilon_n]$$

$$= y_0 + \underbrace{E(\varepsilon_1)}_{0} + \underbrace{E(\varepsilon_2)}_{0} + \dots + \underbrace{E(\varepsilon_n)}_{0}$$

* Expected value of an error = 0

$$Var(Y) = E[y - E(y)]^2$$

$$= E[y_0 + \sum_{i=1}^{n} \varepsilon_i - y_0]^2$$

$$Var(Y) = E[\sum_{i=1}^{n} \varepsilon_i]^2$$

$$Var(Y) = E[\varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \dots \varepsilon_{t-n}]^2$$

$$= \sigma^2 + \sigma^2 + \sigma^2 + \sigma^2 \quad \text{(t times)}$$

$$Var(Y) = t\sigma^2$$

$$Cor(Y_t, Y_{t-1})$$

$$Cor(x, y) = E[(x - \bar{x})(y - \bar{y})]$$

$$Cor(Y_t, Y_{t-1}) = E[(y_t - E(y_t))(y_{t-1} - E(y_{t-1}))]$$

$$= E[\varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \dots][\varepsilon_{t-1} + \varepsilon_{t-2} + \dots]$$

* Any error $(\varepsilon)^2 = \sigma^2$
* Any multiplication of $(\varepsilon)$ that is not squared $= 0$

$$= 0 + 0 + 0 \dots$$

$$= \sigma^2 + 0 + 0 \dots$$

$$= 0 + \sigma^2 + 0 \dots$$

$$Cor(y_t, y_{t-1}) = (t-1)\sigma^2$$

$y_t = y_{t-1} + \varepsilon$

$y_t - y_{t-1} = y_{t-1} - y_{t-1} + \varepsilon_t$

$\Delta Y_t = \varepsilon_t$

$E(\Delta Y) = E(\varepsilon_t) = 0$

$Var(\Delta Y) = E[\Delta Y - E(\Delta Y)]^2$

$Var(\Delta Y) = E(\varepsilon_t)^2 = \sigma^2$

$$Cor = E(\underbrace{\Delta Y_t - E(\Delta Y_t)}_{\varepsilon_t})(\underbrace{\Delta Y_{t-1} - E(\Delta Y_{t-1})}_{\varepsilon_{t-1}})$$

$Cor(\Delta Y, \Delta Y_{t-1}) = 0$

---

$y_t = y_{t-1} + \varepsilon$

$y_t = \rho y_{t-1} + \varepsilon_t$

$0 \leq \rho \leq 1$

$y_t - y_{t-1} = \rho y_{t-1} - y_{t-1} + \varepsilon_t$

$y_t - y_{t-1} = y_{t-1}(\rho - 1)$

$= (\rho - 1) y_{t-1}$

$y_t = \varepsilon_t$

$y_t = \rho y_{t-1} + \varepsilon_t$

- If $\rho = 1$, y is non-stationary. If $\rho = 0$, y is non-stationary.

# TEST FOR STATIONARITY

$y_t = \rho y_{t-1} + \varepsilon$

$H_0 : \rho = 1$ - Non-stationary
$H_a : \rho < 1$ - Stationary

$y_t - y_{t-1} = \rho y_{t-1} - y_{t-1} + e_t$

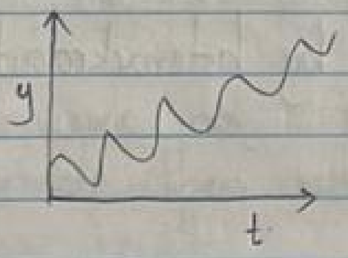
$\Delta y_t = (\rho - 1) y_{t-1} + e_t$

$\Delta y_t = \beta y_{t-1} + e_t$

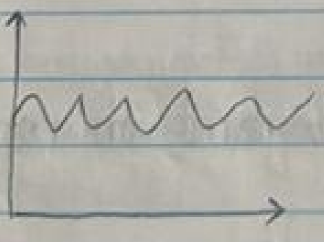
$H_0 : \beta = 0$
$H_a : \beta < 0$

1. Dickey - Fuller Test.

a) $\Delta y_t = \beta y_{t-1} + e_t$ - Random walk.




b) $\Delta y_t = \alpha_0 + \beta y_{t-1} + e_t$ - Random walk with a drift.



c) $\Delta y_t = \alpha_0 + \alpha_1 t + \beta y_{t-1} + e_t$ - Random walk with a drift and a trend

* The $H_0$ : for a dickey - fuller test is that the variable is non-stationary and the $H_a$ is stationary.

* If the first test is non-stationary, we move to the next and test. It is the results of all the 3 tests that would conclude that it is non-stationary.
- However, if you find it is stationary from the first test, there is no need to do the other tests.

Weakness

1. It assumes ~~that~~ an $AR_1$ process. Works only if it has one-lag. ~~Modif~~ Dickey-Fuller modified it to accommodate a larger lag.

2. Augmented Dickey Fuller.

a) $\Delta y_t = \beta_1 y_{t-1} + \beta_2 \Delta y_{t-1} + \beta_3 \Delta y_{t-2} + e_t$ Random walk.

b) $\Delta y_t = \alpha + \beta_1 y_{t-1} + \beta_2 \Delta y_{t-1} + \ldots\ldots$ Random walk with a drift

c) $\Delta y_t = \alpha + \alpha_1 t + \beta_1 y_{t-1} + \ldots\ldots$ Random walk with a drift and trend

- Augmented DF can be affected by heteroskedasticity and autocorrelation.

3. Phillips - Perron Test.
- It is a complimentary test to ADF to test for heteroskedasticity and autocorrelation.

* The DF test is more-likely to insist on non-stationarity.

Multivariate Analysis

- In Timeseries, other variables can help us predict/explain a variable

$y = f(y_{t-}, x_1, x_2)$

- In a situation where we have more than one variable, we must test for stationarity of each variable.

$y = f(x_1$

$\Delta y_t = \alpha_0 + \alpha_1 \Delta x_t + \varepsilon$ - (short term, $\alpha_0, \alpha_1$)

$y_t = \beta_0 + \beta_1 x_t + \varepsilon$ - (long term, $\beta_0, \beta_1$)

1. $I_0$ $I_0$ - Stationary; there is both short-and long-term r/ship

2. $I_1$ $I_1$ - Non-stationary; but integrated of the same order - become stationary after differencing
   a) Have a short term r/ship ($\Delta y_t = \alpha_0 + \alpha_1 \Delta x_t + e$

3. $I_0$ $I_1$ - Non-stationary - mixed; integrated of different order
   - Don't have both a long-term and short-term r/ship

b) If they are co-integrated, the long-term relationship still exists

$y_t = \beta_0 + \beta_1 X_t + e$ - will get good results

c) If they are not co-integrated then there is no long-term relationship then you cannot run a long-term r/ship equation below without spurious results

$y_t = \beta_0 + \beta_1 x_t + e$

* If all the variables are stationary, we have both short and long-term relationship but we can have a VAR (Vector Auto Regression)

- A VAR borrows from the simr relationship where, all the variables are treated the same way.
- VAR models a variable as a function of its past values (lags) and also a function of its present and lagged values of other variables and the error term

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots \alpha_p Y_{t-p} + \beta_0 Y_t + \beta_1 Y_{t-1} + \beta_r Y_{t-r} + u_t$$

$$Y_t = \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=0}^{k} \beta_j Y_{t-j} + u_t$$

$\underline{Y \times Z}$

$$Y_t = \sum_{i=1}^{p} \beta_i Y_{t-i} + \sum_{j=0}^{q} \alpha_j X_{t-j} + \sum_{i=0}^{r} \gamma_i Z_{t-i} + u_t$$

$$X_t = \sum_{i=1}^{u} \beta_i \cancel{Y} X_{t-i} + \sum_{j=0}^{v} \alpha_j Y_{t-j} + \sum_{i=0}^{w} \gamma_i Z_{t-i} + u_t$$

$$Z_t = \sum_{i=1}^{x} \beta_i Z_{t-i} + \sum_{j=0}^{y} \alpha_j Y_{t-j} + \sum_{i=0}^{z} \gamma_i X_{t-i} + u_t$$

* Apart from just getting the coefficients, VAR has 3 applicat
  i) GRANGER CAUSALITY
  ii) IMPULSE RESPONSE
  iii) VARIANCE DECOMPOSITION.

i) GRANGER CAUSALITY
- Variable X granger causes variable Y if the past values of X helps in predicting ~~the~~ Y.
- Granger causality requires that lagged values of variable X are related to subsequent values in variable Y.

$$Y_t = \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum \beta_i X_{t-i} + u \quad \dots (i)$$

| | |
|---|---|
| $H_0 : \beta_i = 0$ | $H_0$ is rejected ; X granger causes Y |
| $H_a : \beta_i \neq 0$ | $H_0$ is not rejected ; X does not granger cause Y<br>- X should not be on the right side. |

* If X does not granger cause Y ;

$$Y_t = \sum \alpha_i Y_{t-i} + u$$

- If Y granger causes X;

$$X_t = \Sigma \beta_i X_{t-i} + \Sigma \alpha_i Y_{t-i} + u \quad \ldots (ii)$$

* Granger causality of VAR dictates which variables should be on the right hand side.
* If X granger causes Y and Y granger causes X; then we can run equations (i) and (ii)

ii) IMPULSE RESPONSE
- It traces the expected response of current and future values of each of the variables to a shock in one of the VAR equations.
- The impulse response traces the impact of a shock on a variable on other variables.

Y and X impulse response tells us how;

i) shocks in X affect Y
ii) shocks in X affect X
iii) shocks in Y affect X
iv) shocks in Y affect Y

- Some shocks affect instantaneously, others don't affect at all while other shocks affect after a period of time.

iii) VARIANCE DECOMPOSITION
- The aim is to decompose the variances of each element of Y into components due to each of the elements of the error term.

$$Y_t = \Sigma \beta_i Y_{t-1} + \Sigma \alpha_i X_{t-1} + \Sigma \gamma_i Z_{t-1} + e$$

NON-STATIONARY

$y = f(x)$

$I_1 \quad I_1$

- Non-stationary but integrated of the same order.

*Co-integration means the existence of a long-run equilibrium relationship among non-stationary time-series variables.

- Co-integration is a property of two or more variables moving together through time and despite following their own individual trends will not drift too far apart since they are linked together in some sense.

Tests for Co-integration

1. Engel-Granger Test
2. Johansen Test
3. Dubin-Watson Test (Co-integration Relation)

*Test for co-integration; if they are co-integrated they have a long-run relation and you can run a regression.

- If the variables are co-integrated, there exists a long-term relationship and a short-term relationship.
- If they are not co-integrated, get the difference and run a short-term relationship $\Delta Y = \alpha_0 + \alpha_1 \Delta X$

$\alpha_0, \alpha_1$ (short-run)

- If they are co-integrated, run an error correction model (ECM)

i) Long term

$$Y_{t-1} = \beta_0 + \beta_1 Y_{t-1} + e_{t-1}$$

$\beta_0, \beta_1$ - Long term

ii) Short term

$$\Delta Y_t = \alpha_0 + \alpha_1 \Delta X$$

$\alpha_0, \alpha_1$ - Short term

iii) Combine

$$\Delta Y_t = \alpha_0 + \alpha_1 \Delta X_t + \Delta e_{t-1} + e_t$$ (Speed of adjustment)