$$\hat{\beta}_1 = \frac{Cov(X,Y)}{Var(X)} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} =$$

$$= \frac{\sum (x_i - \bar{x})(\cancel{\beta_0} + \beta_1 x_i + \varepsilon_i - \cancel{\beta_0} - \beta_1 \bar{x} - \bar{\varepsilon})}{\sum (x_i - \bar{x})^2} =$$

$$= \frac{\sum (x_i - \bar{x}) \beta_1 (x_i - \bar{x})}{\sum (x_i - \bar{x})^2} + \frac{\sum_{i=1}^{n} (x_i - \bar{x})(\varepsilon_i - \bar{\varepsilon})}{\sum (x_i - \bar{x})^2} =$$

$$= \beta_1 + \sum \frac{(x_i - \bar{x})\varepsilon_i}{\sum (x_i - \bar{x})^2} - \frac{\bar{\varepsilon} \overbrace{\sum (x_i - \bar{x})}^{=0}}{\sum (x_i - \bar{x})^2}$$

$\sum (x_i - \bar{x}) = 0$

$\sum x_i = n\bar{x}$

$$= \beta_1 + \sum a_i \cdot \varepsilon_i , \quad a_i = \frac{x_i - \bar{x}}{\sum (x_i - \bar{x})^2}$$

1) Equation: $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ – error (u_i)

2) Assumptions: $E(\varepsilon_i) = 0$ $\quad E(\varepsilon_i \varepsilon_j) = 0$ $\quad E(\varepsilon_i^2) = \sigma_\varepsilon^2$

3) Method: OLS $\quad y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\varepsilon}_i$

residual (e_i)

4) Properties.

Assumptions

of linear regression
for Stochastic Regressors

1) $E(\varepsilon_i | X) = 0$ $\quad E(\varepsilon_i) = 0$

2) (X_i, Y_i) i.i.d.

3) $E(\varepsilon_i^4) < \infty$
$E(X_i^4) < \infty$

4) $rank(X) = k$ no perfect multicollinearity

5) $Var(\varepsilon_i | X) = \sigma_\varepsilon^2$ $\quad Var(\varepsilon_i) = \sigma_\varepsilon^2$

6) $\varepsilon_i | X_i \sim N(0, \sigma_\varepsilon^2)$ $\quad \varepsilon_i \sim N(0, \sigma_\varepsilon^2)$

Assumption GMT

1) $E(\varepsilon | X) = 0$

2) $E(\varepsilon \varepsilon' | X) = \sigma_\varepsilon^2 \cdot \mathbb{I}$

3) $rank(X) = k$

$\hat{\beta}_{OLS}$ - BLUE

$$a_i = \frac{x_i - \bar{x}}{\Sigma (x_i - \bar{x})^2}$$

1) $\Sigma a_i = 0 \qquad \bar{x} \, \Sigma (x_i - \bar{x}) = 0$

2) $$\Sigma a_i^2 = \Sigma \left(\frac{x_i - \bar{x}}{\Sigma (x_i - \bar{x})^2} \right)^2 = \frac{\cancel{\Sigma (x_i - \bar{x})^2}}{\left(\Sigma (x_i - \bar{x})^2 \right)^{\cancel{2}}} = \frac{1}{\Sigma (x_i - \bar{x})^2}$$

3) $$\Sigma a_i x_i = \frac{\Sigma (x_i - \bar{x}) x_i - \Sigma (x_i - \bar{x}) \bar{x}}{\Sigma (x_i - \bar{x})^2} = 1$$

$$\hat{\beta}_1 = \beta_1 + \sum a_i \varepsilon_i$$

$$E(\hat{\beta}_1) = \beta_1 + \sum a_i \underbrace{E(\varepsilon_i)}_{=0}$$

$$\hat{\beta}_1 = \frac{Cov(X,Y)}{Var(X)} = \frac{Cov(X, \beta_0 + \beta_1 X + \varepsilon_i)}{Var(X)} =$$

$$= 0 + \beta_1 + \frac{Cov(X, \varepsilon)}{Var(X)}$$

$$E(\hat{\beta}_1) = \beta_1 + \frac{Cov(X, E\varepsilon)}{Var(X)} \overset{=0}{} = \beta_1$$

$$\sigma^2_{\hat{\beta}_1} = \frac{\sigma^2_\varepsilon}{\sum(X_i - \bar{X})^2} \qquad \hat{\beta}_1 = \beta_1 + \sum a_i \varepsilon_i$$

$$\sigma^2_{\hat{\beta}_1} = E\left(\left(\hat{\beta}_1 - E(\hat{\beta}_1)\right)^2\right) = E\left((\hat{\beta}_1 - \beta_1)^2\right) = E\left(\left(\sum a_i \varepsilon_i\right)^2\right) =$$

$$= E\left(\sum_{i=1}^{n} a_i^2 \varepsilon_i^2 + \sum_{i=1}^{n}\sum_{j \neq i} a_i a_j \underbrace{\varepsilon_i \varepsilon_j}_{E(\varepsilon_i \varepsilon_j) = 0}\right) = \sigma^2_\varepsilon \sum a_i^2 = \frac{\sigma^2_\varepsilon}{\sum(X_i - \bar{X})^2}$$

$$\sigma^2_{\hat{\beta}_0} = \sigma^2_\varepsilon \left(\frac{1}{n} + \frac{\bar{X}^2}{\sum(X_i - \bar{X})^2}\right)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = \beta_0 + \beta_1 \bar{x} + \bar{\varepsilon} - \hat{\beta}_1 \bar{x} =$$

$$\beta_0 + \bar{x}(\beta_1 - \hat{\beta}) + \bar{\varepsilon} =$$

$$\beta_0 \underbrace{- \bar{x} \sum a_i \cdot \varepsilon_i + \frac{\sum \varepsilon_i}{n}} = \left\{ c_i = \frac{1}{n} - a_i \bar{x} \right\}$$

$$= \beta_0 + \sum c_i \varepsilon_i$$

$$\sigma^2_{\hat{\beta}_0} = E\left((\hat{\beta}_0 - \beta_0)^2 \right) = E\left(\left(\sum c_i \varepsilon_i \right)^2 \right) =$$

$$\sigma^2_\varepsilon \sum c_i^2 = \sigma^2_\varepsilon \sum \left(\frac{1}{n} - a_i \bar{x} \right)^2 =$$

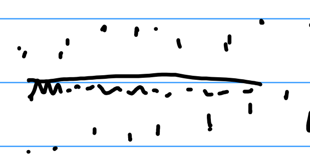
$$= \sigma^2_\varepsilon \left(n \cdot \frac{1}{n^2} - 2 \frac{\bar{x}}{n} \underbrace{\sum a_i}_{\overset{\shortparallel}{0}} + \bar{x}^2 \cdot \underbrace{\sum a_i^2}_{\overset{\shortparallel}{\frac{1}{\sum (x_i - \bar{x})^2}}} \right) =$$

$$= \sigma^2_\varepsilon \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2} \right)$$

$$Var(\hat{\beta}_j) = \frac{\sigma^2_\varepsilon}{TSS_j (1 - R^2_j)}$$

$Var(\hat{\beta}_j) \uparrow$

1) level of noise $(\sigma^2_\varepsilon) \uparrow$

2) 

regressor close to const $(TSS \approx 0)$

3) multicollinearity $R^2_j \approx 1$

$$TSS_j = \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2$$

R^2_j – coefficient of determination

$$X_j \mid X_{-j}$$

$$X_j \mid X_1, \dots, X_{j-1}, X_{j+1}, \dots X_n$$

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

$$(X'X) = \begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix}$$

$$\sigma^2(\hat{\beta}) = \begin{bmatrix} \operatorname{Var}(\hat{\beta}_0) & \operatorname{Cov}(\hat{\beta}_0, \hat{\beta}_1) \\ \operatorname{Cov}(\hat{\beta}_0, \hat{\beta}_1) & \operatorname{Var}(\hat{\beta}_1) \end{bmatrix} =$$

$$= \begin{bmatrix} \dfrac{\sigma^2 \sum x_i^2}{n \sum (x_i - \bar{x})^2} & \dfrac{-\bar{x} \cdot \sigma^2}{\sum (x_i - \bar{x})^2} \\ \dfrac{-\bar{x}\, \sigma^2}{\sum (x_i - \bar{x})^2} & \dfrac{\sigma^2}{\sum (x_i - \bar{x})^2} \end{bmatrix}$$