$$y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$$

a) $H_0 : \beta_2 = 0$

$$t = \frac{\hat{\beta}_2 - 0}{se(\hat{\beta}_2)} \overset{H_0}{\sim} t_{n-k}$$

k: # of est. coefficient

b) $H_0 : \beta_2 = 1$

$$t = \frac{\hat{\beta}_2 - 1}{se(\hat{\beta}_2)} \overset{H_0}{\sim} t_{n-3}$$

c) $H_0 : \beta_2 = \beta_3 = 0$

R: $y_i = \beta_1 + \varepsilon_i$

UR: $y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$

F-test for goodness of fit

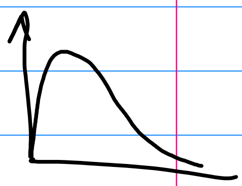
$$F = \frac{ESS / k-1}{RSS / n-k} \sim F(k-1, n-k)$$

$$= \frac{R^2 / 2}{(1-R^2) / n-3}$$

$H_0: \quad \beta_2 = \beta_3$

$$UR: \quad y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$$

$$R: \quad y_i = \beta_1 + \beta_2 \underbrace{(X_{2i} + X_{3i})}_{= z_i} + \varepsilon_i$$



$$F = \frac{(RSS_R - RSS_{UR})/q}{RSS_{UR}/n-k} \sim F(q, n-k)$$

$$F = \frac{(1 - R^2_R - 1 + R^2_{UR})/q}{(1 - R^2_{UR})/n-k} =$$

$$= \frac{(R^2_{UR} - R^2_R)/q}{(1 - R^2_{UR})/n-k} \sim F(q, n-k)$$

$H_0: \beta_2 + \beta_3 = 1$

UR: $y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$

R: $y_i = \beta_1 + \beta_2 X_{2i} + (1 - \beta_2) X_{3i} + \varepsilon_i$

$$\underbrace{y_i - X_{3i}}_{Z_i} = \beta_1 + \beta_2 (\underbrace{X_{2i} - X_{3i}}_{H_i}) + \varepsilon_i$$

$TSS_R \neq TSS_{UR} \Rightarrow$ F stat should be calculated using RSS (not R^2!)

c) $H_0: \beta_2 + \beta_3 = 1$ $\qquad \theta = \beta_2 + \beta_3 = 1$

$H_a: \beta_2 + \beta_3 < 1$

$$y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$$

$$y_i = \beta_1 + \beta_2 X_{2i} + (1 - \theta) X_{3i} + \varepsilon_i$$

$$y_i - X_{3i} = \beta_1 + \beta_2 X_{2i} - \theta X_{3i} + \varepsilon_i$$

$H_0: \theta = 1$

$H_a: \theta < 1$

Prediction error:

$$y_{n+1} = \beta_1 + \beta_2 x_{n+1} + \varepsilon_{n+1}$$

$$\hat{y}_{n+1} = \hat{\beta}_1 + \hat{\beta}_2 x_{n+1}$$

$$E(\hat{y}_{n+1}) = E(\hat{\beta}_1 + \hat{\beta}_2 x_{n+1}) = E(\hat{\beta}_1) + E(\hat{\beta}_2) \cdot x_{n+1}$$

$$= \beta_1 + \beta_2 \cdot x_{n+1}$$

$$E(y_{n+1}) = E(\beta_1 + \beta_2 x_{n+1} + \varepsilon_{n+1}) =$$

$$\beta_1 + \beta_2 x_{n+1} + \underset{0}{\underset{\shortparallel}{E(\varepsilon_{n+1})}}$$

$$E(y_{n+1}) = E(\hat{y}_{n+1})$$

$$Var(\hat{y}_{n+1} - y_{n+1}) = E(\hat{y}_{n+1} - y_{n+1})^2 =$$

$$= E(\hat{\beta}_1 + \hat{\beta}_2 x_{n+1} - \beta_1 - \beta_2 x_{n+1} - \varepsilon_{n+1})^2 =$$

$$= E((\hat{\beta}_1 - \beta_1) + (\hat{\beta}_2 - \beta_2) x_{n+1} - \varepsilon_{n+1})^2 =$$

$$= E(\hat{\beta}_1 - \beta_1)^2 + x_{n+1}^2 E(\hat{\beta}_2 - \beta_2)^2 + E(\varepsilon_{n+1}^2) +$$

$$+ 2 x_{n+1} E(\hat{\beta}_1 - \beta_1)(\hat{\beta}_2 - \beta_2) - 2E((\hat{\beta}_1 - \beta_1)\varepsilon_{n+1}) +$$

$$- 2 x_{n+1} E((\hat{\beta}_2 - \beta_2)\varepsilon_{n+1}) = Var(\hat{\beta}_1) + x_{n+1}^2 \cdot Var(\hat{\beta}_2) +$$

$$\sigma^2 + 2 x_{n+1} \cdot Cov(\hat{\beta}_1, \hat{\beta}_2) - 0 - 0 =$$

$$= \frac{\sigma^2 \sum x_i^2}{\sum (x_i-\bar{x})^2} + x_{n+1}^2 \frac{\sigma^2}{\sum (x_i-\bar{x})^2} + \sigma^2 +$$

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

$$\delta$$

$$y_{n+1} : \left\{ \hat{y}_{n+1} \pm t_{\alpha/2, n-2} \cdot \delta \right\} \quad \hat{\delta} \Leftarrow s^2$$

$$\sigma^2(\hat{\beta}) = \sigma^2 (X'X)^{-1} = \begin{bmatrix} \frac{\sigma^2 \sum x_i^2}{n \sum (x_i - \bar{x})^2} & \frac{-\bar{x}\sigma^2}{\sum (x_i - \bar{x})^2} \\ \frac{-\bar{x}\sigma^2}{\sum (x_i - \bar{x})^2} & \frac{\sigma^2}{\sum (x_i - \bar{x})^2} \end{bmatrix}$$

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \qquad X'X = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix}$$

Multicollinearity

Perfect multicollinearity :

one of regressors is a lin. comb

from other regressors

$\operatorname{rank}(X) < k$ $\qquad \hat{\beta} = (X'X)^{-1}X'y$

$(X'X)^{-1}$ - is not invertible

$\Rightarrow$ can't obtain $\hat{\beta}$

Perfect MC example: dummy-variable trap

$$\begin{pmatrix} 1 & 1 & 0 & x_1 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & 1 & 0 & \vdots \\ \vdots & 0 & 1 & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 1 & x_n \end{pmatrix}$$

$1 = m_i + f_i$

$\beta_0 + \beta_f f_i$

$\beta_m \cdot m_i + \beta_f f_i$

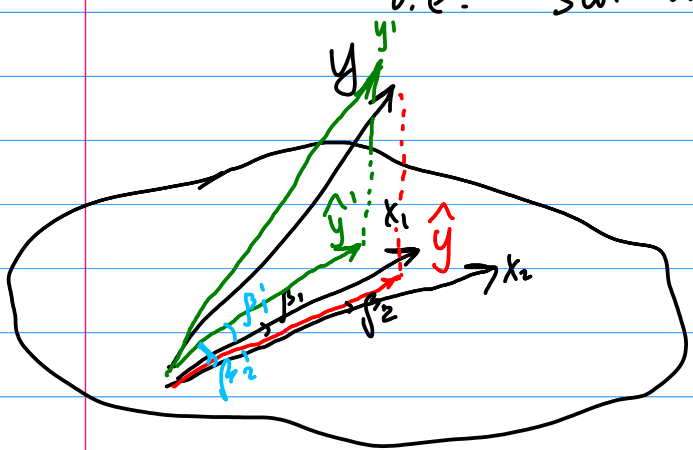
Multicollinearity

- $\hat{\beta}$ - unbiased, not effective
- $se(\hat{\beta})$ are inflated

  t-statistics are smaller

- instability of estimates

  i.e. switching sign



(add new variable /

add new regressor)

Testing for multicollinearity

$$y_i = \beta_0 + \beta_1 \cdot X_{1i} + \ldots + \beta_k X_{ki} + \varepsilon_i$$

VIF (variance inflation factor)

$X_i \mid X_{-i}$ $\quad (X_i \mid X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_k)$

$\Rightarrow R_i^2$

$$VIF(X_i) = \frac{1}{1 - R_i^2} > 10 \Rightarrow \text{multicollinearity}$$