P.1. $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$

a) $H_0: \beta_2 = 0$

$H_a: \beta_2 \neq 0$

$$t = \frac{\hat{\beta}_2 - 0}{se(\hat{\beta}_2)} \sim t_{n-3}$$

b) $H_0: \beta_2 = 1$

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

c) 1) $H_0: \beta_2 = \beta_3 = 0$

$$F = \frac{ESS/2}{RSS/n-3} \sim F(2, n-3)$$

d) $H_0: \beta_2 = \beta_3$

$$\left[\begin{array}{ll} UR: & Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i \\ R: & Y_i = \beta_1 + \beta_2 \underbrace{(X_{2i} + X_{3i})}_{t'} + \varepsilon_i \end{array}\right.$$

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

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

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

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

e) $H_0: \beta_2 + \beta_3 = 1$

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

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

$$\underbrace{y_i - X_{3i}}_{= Z} = \beta_1 + \beta_2 \underbrace{(X_{2i} - X_{3i})}_{= H} + \epsilon_i$$

$TSS_R \neq TSS_{UR} \Rightarrow$ F stat with RSS (not R^2!)

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

$$\beta_2 + \beta_3 = \theta = 1$$

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

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

$$H_0: \theta = 1$$

$$t = \frac{\hat{\theta} - 1}{se(\hat{\theta})} \sim t_{n-3}$$

Prediction error:

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

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

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

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

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

$$\beta_1 + \beta_2 x_{n+1} + \underbrace{E(\varepsilon_{n+1})}_{=0}$$

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

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

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

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

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

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

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

$$= var(\hat{\beta}_1) + x_{n+1}^2 \, Var(\hat{\beta}_2) + \sigma^2 +$$

$$+ 2x_{n+1} \cdot cov(\hat{\beta}_1, \hat{\beta}_2) + 0 + 0 =$$

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

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

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

$$\Updownarrow$$

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

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

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} \quad X'X = \begin{bmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{bmatrix}$$

Perfect multicollinearity:

one variable is a lin. comb of the other variables

$$\hat{\beta} = (X'X)^{-1} X' y$$

perf. MC $\Rightarrow$ rank$(X) < k$

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

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

Perf. MC examples:

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

$$1 = m_i + f_i$$

$$\beta_0 + \underset{=0}{\beta_m} \cdot m_i$$

$$\beta_m \cdot m_i + \beta_f \cdot f_i + 0$$

$$\beta_m = \beta_f$$

Multicollinearity:

- $se(\hat{\beta}_i)$ are inflated

  $\Rightarrow$ $\hat{\beta}$ - not efficient

  t - statistic are smaller

- $\hat{\beta}$ - are not stable

  (+ new obs.,
  + new regressor)

  $\Rightarrow$ $\hat{\beta}$ switches sign