Ex. (Model Selection)

Suppose we are interested in predicting of surgery survival rate as a function of x_1 = blood clotting score, x_2 = prognostic index, x_3 = enzyme function test score, x_4 = liver function test score and y = log(surgery survival rate). The TSS for the full model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon \quad \text{is : TSS = 21.07733.}$$

We decided to screen the independent variables to determine the best set for predicting the surgery rates. The sums of squares for all possible regression models were found to be as follows:

Independent variables in the model	**SSR**	**SSE**	**d.f.** $_{SSE}$	**MSE**	R^2
X_1	2.52720	18.55013	52	0.3567332	0.1199013
X_2	7.39311	13.68422	52	0.263158	0.3507612
X_3	9.33966	11.73767	52	0.2257244	0.443114
X_4	11.10557	9.97176	52	0.1917646	0.5268964
X_1, X_2	12.76391	8.31342	51	0.1630082	0.6055752
X_1, X_3	13.62588	7.45145	51	0.1461068	0.6464708
X_1, X_4	11.11514	9.96219	51	0.195337	0.5273504
X_2, X_3	17.13462	3.94271	51	0.077308	0.8129407
X_2, X_4	13.67307	7.40426	51	0.1451815	0.6487097
X_3, X_4	14.47288	6.60445	51	0.129499	0.6866562
X_1, X_2, X_3	20.49376	0.58357	50	0.0116714	0.9723129
X_1, X_2, X_4	13.68169	7.39564	50	0.1479128	0.6491187
X_1, X_3, X_4	15.16439	5.91294	50	0.1182588	0.7194644
X_2, X_3, X_4	18.60417	2.47316	50	0.0494632	0.8826625
X_1, X_2, X_3, X_4	20.49413	0.5832	49	0.011902	0.9723304

1. Determine the subset of variables that is selected as best using **max R^2 criterion.** Show your steps.

$R^2 = \frac{SSR}{TSS}$ **, the set $\{X_1, X_2, X_3\}$ is selected as the best one. (Please note that the full model gives the highest R^2, however we prefer the second highest one other than the full model).**

2. Determine the subset of variables that is selected as best using **min MSE criterion.** Show your steps.

The best model is determined by the set $\{X_1, X_2, X_3\}$ (since the *min MSE* and *max* R^2 are equivalent).

3. Determine the subset of variables that is selected as best using **Mallows C_p criterion**. Show your steps.

We will select as the best model whose C_p is as close to *p* as possible.

$$C_p = \frac{SSE_p}{MSE(X_1, X_2, X_3, X_4)} - (n - 2p)$$

, n = 54 (since $d.f._{TSS} = d.f._{SSR} + d.f._{SSE}$, so then e.g. when $k = 1$, $d.f._{TSS} = 1 + 52 = 53 = (n - 1)$).

- **when $p = 2$ (i.e. one-variable models):**

for X_1: $$C_p = \frac{SSE(X_1)}{MSE(X_1, X_2, X_3, X_4)} - (54 - 2(2)) = \frac{18.55013}{0.011902} - 50 = \underline{\underline{\mathbf{1\,508.5725}}}$$

for X_2: $$C_p = \frac{SSE(X_2)}{MSE(X_1, X_2, X_3, X_4)} - (54 - 2(2)) = \frac{13.68422}{0.011902} - 50 = \underline{\underline{\mathbf{1\,099.7412}}}$$

for X_3: $$C_p = \frac{SSE(X_3)}{MSE(X_1, X_2, X_3, X_4)} - (54 - 2(2)) = \frac{11.73767}{0.011902} - 50 = \underline{\underline{\mathbf{936.19308}}}$$

for X_4: $$C_p = \frac{SSE(X_4)}{MSE(X_1, X_2, X_3, X_4)} - (54 - 2(2)) = \frac{9.97176}{0.011902} - 50 = \underline{\underline{\mathbf{787.82221}}}$$

- **when p = 3 (i.e. two-variable models):**

for X_1,X_2: $$C_p = \frac{SSE(X_1, X_2)}{MSE(X_1, X_2, X_3, X_4)} - (54 - 2(3)) = \frac{8.31342}{0.011902} - 48 = \underline{\underline{\mathbf{650.48933}}}$$

for X_1,X_3: $$C_p = \frac{SSE(X_1, X_3)}{MSE(X_1, X_2, X_3, X_4)} - (54 - 2(3)) = \frac{7.45145}{0.011902} - 48 = \underline{\underline{\mathbf{578.06705}}}$$

for X_1,X_4: $$C_p = \frac{SSE(X_1, X_4)}{MSE(X_1, X_2, X_3, X_4)} - (54 - 2(3)) = \frac{9.96219}{0.011902} - 48 = \underline{\underline{\mathbf{789.01815}}}$$

for X_2,X_3: $$C_p = \frac{SSE(X_2, X_3)}{MSE(X_1, X_2, X_3, X_4)} - (54 - 2(3)) = \frac{3.94271}{0.011902} - 48 = \underline{\underline{\mathbf{283.26449}}}$$

for X_2,X_4: $$C_p = \frac{SSE(X_2, X_4)}{MSE(X_1, X_2, X_3, X_4)} - (54 - 2(3)) = \frac{7.40426}{0.011902} - 48 = \underline{\underline{\mathbf{574.10217}}}$$

for X_3,X_4: $$C_p = \frac{SSE(X_3, X_4)}{MSE(X_1, X_2, X_3, X_4)} - (54 - 2(3)) = \frac{6.60445}{0.011902} - 48 = \underline{\underline{\mathbf{506.90254}}}$$

- **when p = 4 (i.e. three-variable models):**

for X_1,X_2,X_3: $$C_p = \frac{SSE(X_1, X_2, X_3)}{MSE(X_1, X_2, X_3, X_4)} - (54 - 2(4)) = \frac{0.58357}{0.011902} - 46 = \underline{\underline{\mathbf{3.0312553}}}$$

for X_1,X_2,X_4: $$C_p = \frac{SSE(X_1, X_2, X_4)}{MSE(X_1, X_2, X_3, X_4)} - (54 - 2(4)) = \frac{7.39564}{0.011902} - 46 = \underline{\underline{\mathbf{575.37792}}}$$

for X_1,X_3,X_4: $$C_p = \frac{SSE(X_1, X_3, X_4)}{MSE(X_1, X_2, X_3, X_4)} - (54 - 2(4)) = \frac{5.91294}{0.011902} - 46 = \underline{\underline{\mathbf{450.80222}}}$$

for X_2,X_3,X_4: $$C_p = \frac{SSE(X_2, X_3, X_4)}{MSE(X_1, X_2, X_3, X_4)} - (54 - 2(4)) = \frac{2.47316}{0.011902} - 46 = \underline{\underline{\mathbf{161.79365}}}$$

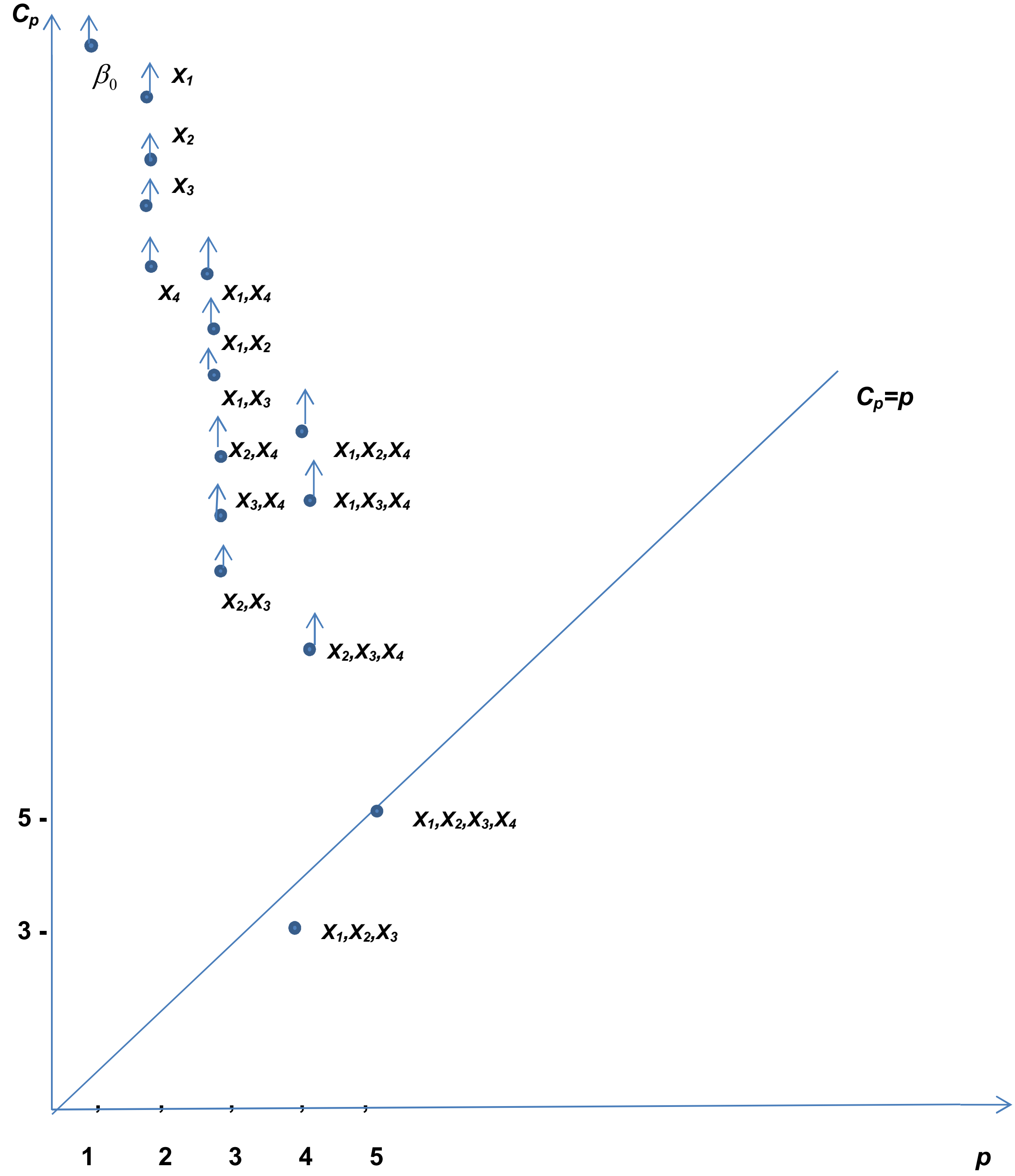
- **when p = 5 (i.e. four-variable model, i.e. the full model):**

for X_1,X_2,X_3,X_4: $$C_p = \frac{SSE(X_1, X_2, X_3, X_4)}{MSE(X_1, X_2, X_3, X_4)} - (54 - 2(5)) = \frac{0.5832}{0.011902} - 44 = \underline{\underline{\mathbf{5}}}$$

- **when *p* = 1 (i.e. no variables in the model, only β_0):**

$$C_p = \frac{TSS}{MSE(X_1, X_2, X_3, X_4)} - (54 - 2(1)) = \frac{21.07733}{0.011902} - 52 = \underline{\underline{\mathbf{1718.9066}}}$$

i.e.



$\therefore$ **the best set is given by $\{X_1, X_2, X_3\}$, since its C_p is closest to p (other than the full model). However, since in this case the full model's C_p is exactly equal to p, we may consider the full model as the best model, as well.**

4. Determine the subset of variables that is selected as best by the **Forward Selection Procedure** using $F_0^* = 4.2$ (to-add-variable). Show your steps.

(1) Fit all one-term models: $y = \beta_0 + \beta_1 x_j + \varepsilon$ **for** $j = 1, 2, 3, 4$
i.e.
SSR(X_1) = 2.52720
SSR(X_2) = 7.39311
SSR(X_3) = 9.33966
SSR(X_4) = 11.10557 ← max

$$\therefore F_4 = \frac{MSR(X_4)}{MSE(X_4)} = \frac{SSR(X_4)/1}{SSE(X_4)/52} = \frac{11.10557}{0.1917646} = \underline{\underline{57.9125}}$$

Since $F_4 = 57.9125 > F_0^* = 4.2$, we <u>keep X_4</u>

(2) Fit all two-term models: $y = \beta_0 + \beta_1 x_4 + \beta_2 x_j + \varepsilon$ **for** $j = 1, 2, 3$
Calculate SSR(X_j | X_4)

i.e.
SSR(X_1 | X_4) = SSR(X_1, X_4) – SSR(X_4) = 11.11514 – 11.10557 = 0.00957
SSR(X_2 | X_4) = SSR(X_2, X_4) – SSR(X_4) = 13.67307 – 11.10557 = 2.5675
SSR(X_3 | X_4) = SSR(X_3, X_4) – SSR(X_4) = 14.47288 – 11.10557 = 3.36731 ← max

$$\therefore F_3 = \frac{MSR(X_3 \mid X_4)}{MSE(X_3, X_4)} = \frac{[SSR(X_3, X_4) - SSR(X_4)]/[df_{SSR(X_3,X_4)} - df_{SSR(X_4)}]}{SSE(X_3, X_4)/df_{SSE(X_3,X_4)}} = \frac{3.36731/(2-1)}{6.60445/51} =$$

$$= \frac{3.36731}{0.129499} = \underline{\underline{26.00259}}$$

Since $F_3 = 26.00259 > F_0^* = 4.2$, we <u>keep X_3 & X_4</u>

(3) Fit all three-term models: $y = \beta_0 + \beta_1 x_4 + \beta_2 x_3 + \beta_3 x_j + \varepsilon$ **for** $j = 1, 2$
Calculate SSR(X_j | X_3, X_4)

i.e.
SSR(X_1 | X_3, X_4) = SSR(X_1, X_3, X_4) – SSR(X_3, X_4) = 15.16439 – 14.47288 = 0.69151

SSR(X_2 | X_3, X_4) = SSR(X_2, X_3, X_4) – SSR(X_3, X_4) = 18.60417 – 14.47288 = 4.13129

↑ **max**

$$\therefore F_2 = \frac{MSR(X_2 \mid X_3, X_4)}{MSE(X_2, X_3, X_4)} = \frac{[SSR(X_2, X_3, X_4) - SSR(X_3, X_4)]/[df_{SSR(X_2,X_3,X_4)} - df_{SSR(X_3,X_4)}]}{SSE(X_2, X_3, X_4)/df_{SSE(X_2,X_3,X_4)}} = \frac{4.13129/(3-2)}{2.47316/50} =$$

$$= \frac{4.13129}{0.0494632} = \underline{\underline{\mathbf{83.52249}}}$$

Since F_2 = 83.52249 > F_0^* = 4.2, we keep X_2, X_3 & X_4

(4) Fit the full model: $y = \beta_0 + \beta_1 x_4 + \beta_2 x_3 + \beta_3 x_2 + \beta_4 x_1 + \varepsilon$
Calculate SSR(X_1 | X_2, X_3, X_4)
i.e.
SSR(X_1 | X_2, X_3, X_4) = SSR(X_1, X_2, X_3, X_4) – SSR(X_2, X_3, X_4) = 20.49413 – 18.60417 =
= 1.88996

$$\therefore F_1 = \frac{MSR(X_1 \mid X_2, X_3, X_4)}{MSE(X_1, X_2, X_3, X_4)} = \frac{[SSR(X_1, X_2, X_3, X_4) - SSR(X_2, X_3, X_4)]/[df_{SSR(X_1,X_2,X_3,X_4)} - df_{SSR(X_2,X_3,X_4)}]}{SSE(X_1, X_2, X_3, X_4)/df_{SSE(X_1,X_2,X_3,X_4)}} = \frac{1.88996/(4-3)}{0.5832/49} =$$

$$= \frac{1.88996}{0.011902} = \underline{\underline{\mathbf{158.79348}}}$$

Since F_1 = 158.79348 > F_0^* = 4.2, we keep X_1, X_2, X_3 & X_4

∴ the best set is {X_1, X_2, X_3, X_4}, i.e. the full model.

5. Determine the subset of variables that is selected as best by the **Backward Elimination Procedure** using F_0^{**} = 4.1 (to-delete-variable). Show your steps.

Fit the full model: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon$ and check whether model is significant (at α = 5%)

i.e. $F = \dfrac{MSR_f}{MSE_f} = \dfrac{SSR_f/4}{SSE_f/49} = \dfrac{5.1235325}{0.011902} =$ **430.4766**

Since $F = 430.4766 > F_{0.05;(4,49)} = 2.57$**, we conclude that at 5% level of significance, the full model is significant (i.e. it can be used)**

(1) Calculate

$$F_j = (t_j)^2 = \frac{MSR(X_j \mid all \quad X's \quad except \quad X_j)}{MSE(X_1, X_2, X_3, X_4)} = \frac{[SSR_f - SSR(all \quad X's \quad except \quad X_j)]/d.f.}{MSE_f}$$

for *j* = 1, 2, 3, 4

i.e.

$$F_1 = \frac{MSR(X_1 \mid X_2, X_3, X_4)}{MSE(X_1, X_2, X_3, X_4)} = \frac{[SSR(X_1, X_2, X_3, X_4) - SSR(X_2, X_3, X_4)]/[df_{SSR(X_1,X_2,X_3,X_4)} - df_{SSR(X_2,X_3,X_4)}]}{SSE(X_1, X_2, X_3, X_4)/df_{SSE(X_1,X_2,X_3,X_4)}} =$$

$= \dfrac{20.49413 - 18.60417/(4-3)}{0.5832/49} = \dfrac{1.88996}{0.011902} =$ **158.79348**

$$F_2 = \frac{MSR(X_2 \mid X_1, X_3, X_4)}{MSE(X_1, X_2, X_3, X_4)} = \frac{[SSR(X_1, X_2, X_3, X_4) - SSR(X_1, X_3, X_4)]/[df_{SSR(X_1,X_2,X_3,X_4)} - df_{SSR(X_1,X_3,X_4)}]}{SSE(X_1, X_2, X_3, X_4)/df_{SSE(X_1,X_2,X_3,X_4)}} =$$

$= \dfrac{20.49413 - 15.16439/(4-3)}{0.5832/49} = \dfrac{5.32974}{0.011902} =$ **447.80205**

$$F_3 = \frac{MSR(X_3 \mid X_1, X_2, X_4)}{MSE(X_1, X_2, X_3, X_4)} = \frac{[SSR(X_1, X_2, X_3, X_4) - SSR(X_1, X_2, X_4)]/[df_{SSR(X_1,X_2,X_3,X_4)} - df_{SSR(X_1,X_2,X_4)}]}{SSE(X_1, X_2, X_3, X_4)/df_{SSE(X_1,X_2,X_3,X_4)}} =$$

$= \dfrac{20.49413 - 13.68169/(4-3)}{0.5832/49} = \dfrac{6.81244}{0.011902} =$ **572.37775**

$$F_4 = \frac{MSR(X_4 \mid X_1, X_2, X_3)}{MSE(X_1, X_2, X_3, X_4)} = \frac{[SSR(X_1, X_2, X_3, X_4) - SSR(X_1, X_2, X_3)]/[df_{SSR(X_1,X_2,X_3,X_4)} - df_{SSR(X_1,X_2,X_3)}]}{SSE(X_1, X_2, X_3, X_4)/df_{SSE(X_1,X_2,X_3,X_4)}} =$$

$= \dfrac{20.49413 - 20.49376/(4-3)}{0.5832/49} = \dfrac{0.00037}{0.011902} =$ **0.03108** $\leftarrow$ **min**

Since F_4 = 0.03108 < F_0^{**} = 4.1, we <u>delete X_4</u>

(2) Fit $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$ and calculate

$$F_j = \frac{MSR(X_j \mid all \quad X's \quad except \quad X_j)}{MSE(X_1, X_2, X_3)} \qquad \textbf{for } j = 1, 2, 3$$

i.e.

$$F_1 = \frac{MSR(X_1 \mid X_2, X_3)}{MSE(X_1, X_2, X_3)} = \frac{[SSR(X_1, X_2, X_3) - SSR(X_2, X_3)]/[df_{SSR(X_1,X_2,X_3)} - df_{SSR(X_2,X_3)}]}{SSE(X_1, X_2, X_3)/df_{SSE(X_1,X_2,X_3)}} =$$

$$= \frac{20.49376 - 17.13462/(3-2)}{0.58357/50} = \frac{3.35914}{0.0116714} = \mathbf{287.8095} \quad \leftarrow \textbf{min}$$

$$F_2 = \frac{MSR(X_2 \mid X_1, X_3)}{MSE(X_1, X_2, X_3)} = \frac{[SSR(X_1, X_2, X_3) - SSR(X_1, X_3)]/[df_{SSR(X_1,X_2,X_3)} - df_{SSR(X_1,X_3)}]}{SSE(X_1, X_2, X_3)/df_{SSE(X_1,X_2,X_3)}} =$$

$$= \frac{20.49376 - 13.62588/(3-2)}{0.58357/50} = \frac{6.86788}{0.0116714} = \mathbf{588.43669}$$

$$F_3 = \frac{MSR(X_3 \mid X_1, X_2)}{MSE(X_1, X_2, X_3)} = \frac{[SSR(X_1, X_2, X_3) - SSR(X_1, X_2)]/[df_{SSR(X_1,X_2,X_3)} - df_{SSR(X_1,X_2)}]}{SSE(X_1, X_2, X_3)/df_{SSE(X_1,X_2,X_3)}} =$$

$$= \frac{20.49376 - 12.76391/(3-2)}{0.58357/50} = \frac{7.72985}{0.0116714} = \mathbf{662.2898}$$

Since F_1 = 287.8095 $\nless$ F_0^{**} = 4.1, we can not delete X_1 (i.e. we keep X_1) and we <u>stop.</u>

∴ the best set is <u>{X_1, X_2, X_3 }</u>

6. Determine the subset of variables that is selected as best by the **Stepwise Regression Procedure** using $F_0^{*} = 4.2$ (to-add) and $F_0^{**} = 4.1$ (to-delete). Show your steps.

(1) Fit all one-term models: $y = \beta_0 + \beta_1 x_j + \varepsilon$ **for** $j = 1, 2, 3, 4$

- **as in Forward Selection in part (a), we know that we keep** X_4

(2) Fit all two-term models: $y = \beta_0 + \beta_1 x_4 + \beta_2 x_j + \varepsilon$ **for** $j = 1, 2, 3$

- **as in Forward Selection in part (a), we know that we keep** X_3 ***and*** X_4

➢ **Is X_4 redundant when X_3 is in the model?**

i.e. SSR(X_4 | X_3) = SSR(X_3, X_4) – SSR(X_3) = 14.47288 – 9.33966 = 5.13322

$$\therefore F_4 = \frac{MSR(X_4 \mid X_3)}{MSE(X_3, X_4)} = \frac{[SSR(X_3, X_4) - SSR(X_3)]/[df_{SSR(X_3,X_4)} - df_{SSR(X_3)}]}{SSE(X_3, X_4)/df_{SSE(X_3,X_4)}} = \frac{5.13322/(2-1)}{6.60445/51} =$$

$$= \frac{5.13322}{0.129499} = \mathbf{39.639}$$

Since F_4 = 39.639 $\nless$ F_0^{**} = 4.1, we keep X_3 & X_4

(3) Fit all three-term models: $y = \beta_0 + \beta_1 x_4 + \beta_2 x_3 + \beta_3 x_j + \varepsilon$ **for** $j = 1, 2$

- **as in Forward Selection in part (a), we know that we keep** X_2, X_3 ***and*** X_4

➢ **Is X_3 redundant when X_2 & X_4 are in the model?**

i.e. SSR(X_3 | X_2, X_4) = SSR(X_2, X_3, X_4) – SSR(X_2, X_4) = 18.60417 – 13.67307 = 4.9311

$$\therefore F_3 = \frac{MSR(X_3 \mid X_2, X_4)}{MSE(X_2, X_3, X_4)} = \frac{[SSR(X_2, X_3, X_4) - SSR(X_2, X_4)]/[df_{SSR(X_2,X_3,X_4)} - df_{SSR(X_2,X_4)}]}{SSE(X_2, X_3, X_4)/df_{SSE(X_2,X_3,X_4)}} =$$

$$= \frac{4.9311/(3-2)}{2.47316/50} = \frac{4.9311}{0.0494632} = \mathbf{99.6922}$$

Since F_3 = 99.6922 ≤ F_0^{**} = 4.1, we <u>keep X_3 | X_2, X_4</u>

- **<u>Is X_4 redundant when X_2 & X_3 are in the model?</u>**

i.e. SSR(X_4| X_2, X_3) = SSR(X_2, X_3, X_4) – SSR(X_2, X_3) = 18.60417 – 17.13462 = 1.46955

$$\therefore F_4 = \frac{MSR(X_4 \mid X_2, X_3)}{MSE(X_2, X_3, X_4)} = \frac{[SSR(X_2, X_3, X_4) - SSR(X_2, X_3)]/[df_{SSR(X_2,X_3,X_4)} - df_{SSR(X_2,X_3)}]}{SSE(X_2, X_3, X_4)/df_{SSE(X_2,X_3,X_4)}} =$$

$$= \frac{1.46955/(3-2)}{2.47316/50} = \frac{1.46955}{0.0494632} = \mathbf{29.70996}$$

Since F_4 = 29.70996 ≤ F_0^{**} = 4.1, we <u>keep X_4 | X_2, X_3</u>

(4) Fit the full model: $y = \beta_0 + \beta_1 x_4 + \beta_2 x_3 + \beta_3 x_2 + \beta_4 x_1 + \varepsilon$
- **as in Forward Selection in part (a), we know that we <u>keep X_1, X_2, X_3 and X_4</u>**

now we need to check for redundancy of previously entered variables when X_1 is in the model:

- **<u>Is X_2 redundant when X_1, X_3 & X_4 are in the model?</u>**

i.e. SSR(X_2| X_1, X_3, X_4) = SSR(X_1, X_2, X_3, X_4) – SSR(X_1, X_3, X_4) = 20.49413 – 15.16439 =

= 5.32974

$$\therefore F_2 = \frac{MSR(X_2 \mid X_1, X_3, X_4)}{MSE(X_1, X_2, X_3, X_4)} = \frac{[SSR(X_1, X_2, X_3, X_4) - SSR(X_1, X_3, X_4)]/[df_{SSR(X_1,X_2,X_3,X_4)} - df_{SSR(X_1,X_3,X_4)}]}{SSE(X_1, X_2, X_3, X_4)/df_{SSE(X_1,X_2,X_3,X_4)}} =$$

$$= \frac{5.32974/(4-3)}{0.5832/49} = \frac{5.32974}{0.011902} = \mathbf{447.80205}$$

Since F_2 = 447.80205 ≤ F_0^{**} = 4.1, we <u>keep X_2 | X_1, X_3, X_4</u>

- **Is X_3 redundant when X_1, X_2 & X_4 are in the model?**

i.e. SSR(X_3| X_1, X_2, X_4) = SSR(X_1, X_2, X_3, X_4) – SSR(X_1, X_2, X_4) = 20.49413 – 13.68169 =

= 6.81244

$$\therefore F_3 = \frac{MSR(X_3 \mid X_1, X_2, X_4)}{MSE(X_1, X_2, X_3, X_4)} = \frac{[SSR(X_1, X_2, X_3, X_4) - SSR(X_1, X_2, X_4)]/[df_{SSR(X_1,X_2,X_3,X_4)} - df_{SSR(X_1,X_2,X_4)}]}{SSE(X_1, X_2, X_3, X_4)/df_{SSE(X_1,X_2,X_3,X_4)}} =$$

$$= \frac{6.81244/(4-3)}{0.5832/49} = \frac{6.81244}{0.011902} = \mathbf{572.37775}$$

Since F_3 = 572.37775 ≤ F_0^{**} = 4.1, we keep X_3 | X_1, X_2, X_4

- **Is X_4 redundant when X_1, X_2 & X_3 are in the model?**

i.e. SSR(X_4| X_1, X_2, X_3) = SSR(X_1, X_2, X_3, X_4) – SSR(X_1, X_2, X_3) = 20.49413 – 20.49376 =

= 0.00037

$$\therefore F_4 = \frac{MSR(X_4 \mid X_1, X_2, X_3)}{MSE(X_1, X_2, X_3, X_4)} = \frac{[SSR(X_1, X_2, X_3, X_4) - SSR(X_1, X_2, X_3)]/[df_{SSR(X_1,X_2,X_3,X_4)} - df_{SSR(X_1,X_2,X_3)}]}{SSE(X_1, X_2, X_3, X_4)/df_{SSE(X_1,X_2,X_3,X_4)}} =$$

$$= \frac{0.00037/(4-3)}{0.5832/49} = \frac{0.00037}{0.011902} = \mathbf{0.03108}$$

Since F_4 = 0.03108 < F_0^{**} = 4.1, we delete X_4 when X_1, X_2 & X_3 are in the model.

∴ the best set is {X_1, X_2, X_3 }