Printed by Kapoor, Rajan on Jun 28, 2017 10:48:02 AM

COVER PAGE

STAT 608 Homework 04, Summer 2017

Please TYPE your name and email address below, then convert to PDF and attach as the first page of your homework upload.

NAME: RAJAN KAPOOR

EMAIL: r.kapoor@tamu.edu

38/42

Solution 1

$$y_i = \alpha_1 x_{i1} + \alpha_2 x_{i2} + e_i$$

1.1 for first group, $x_{i1} = 1$, $x_{i2} = 0$

$$y_i = \alpha_1 + e_i \qquad i = 1 \rightarrow m$$

for second group,

$$y_i = \alpha_2 + e_i \qquad i = m+1 \rightarrow n$$

α_1 is the mean value of random variable Y (associated with observations y_i) for the first group of people while α_2 is that for the second group

1.2

$$\bar{Y} = \begin{bmatrix} 1_m \\ 0_{n-m} \end{bmatrix} \alpha_1 + \begin{bmatrix} 0_m \\ 1_{n-m} \end{bmatrix} \alpha_2 + \bar{e} \qquad — (1)$$

$$= \begin{bmatrix} 1_m \\ 1_{n-m} \end{bmatrix} \alpha_1 + \begin{bmatrix} 0_m \\ 1_{n-m} \end{bmatrix} (\alpha_2 - \alpha_1) + \bar{e}$$

$$= \begin{bmatrix} 1_m & 0_m \\ 1_{n-m} & 1_{n-m} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 - \alpha_1 \end{bmatrix} + \bar{e}$$

$$= \bar{X}\bar{\beta} + \bar{e} \qquad \bar{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 - \alpha_1 \end{bmatrix}$$

Define hypothesis $H_0 : \beta_1 = 0$

Least square estimate of $\vec{\beta}$ is given by

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

$$\bar{X}'\bar{X} = \begin{bmatrix} 1_m' & 1_{n-m}' \\ 0_m' & 1_{n-m}' \end{bmatrix} \begin{bmatrix} 1_m & 0_m \\ 1_{n-m} & 1_{n-m} \end{bmatrix}$$

$$= \begin{bmatrix} n & n-m \\ n-m & n-m \end{bmatrix}$$

$$|\bar{X}'\bar{X}| = n(n-m) - (n-m)^2$$

$$= (n-m)\ (n-n+m)$$

$$= (n-m)m$$

$$(\bar{X}'\bar{X})^{-1} = \frac{1}{|\bar{X}'\bar{X}|} \begin{bmatrix} n-m & -(n-m) \\ -(n-m) & n \end{bmatrix}$$

$$= \begin{bmatrix} 1/m & -1/m \\ -1/m & n/m(n-m) \end{bmatrix}$$

$$(\bar{X}'\bar{Y}) = \begin{bmatrix} 1_m' & 1_{n-m}' \\ 0_m' & 1_{n-m}' \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$= \begin{bmatrix} \sum_1^m y_i + \sum_{m+1}^n y_i \\ \sum_{m+1}^n y_i \end{bmatrix} = \begin{bmatrix} \sum_1^n y_i \\ \sum_{m+1}^n y_i \end{bmatrix}$$

$$\hat{\beta} = (\bar{X}'\bar{X})^{-1} \bar{X}' Y \quad -4$$

$$= \begin{bmatrix} 1/m & -1/m \\ -1/m & n/m(n-m) \end{bmatrix} \begin{bmatrix} \sum_1^n y_i \\ \sum_{m+1}^n y_i \end{bmatrix}$$

Solution 2

2.1

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix}$$

$$\bar{Y} = \bar{X}\bar{\beta} + \bar{e}$$

$$(\bar{X}'\bar{X}) = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$$

$$(\bar{X}'\bar{X})^{-1} = \frac{1}{(9-4)} \begin{bmatrix} 3 & -2 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 3/5 & -2/5 \\ -2/5 & 3/5 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} y_1 + y_3 + y_4 \\ y_2 + y_3 + y_4 \end{bmatrix}$$

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

$$= \frac{1}{5} \begin{bmatrix} 3y_1 + 3y_3 + 3y_4 - 2y_2 - 2y_3 - 2y_4 \\ -2y_1 - 2y_3 - 2y_4 + 3y_2 + 3y_3 + 3y_4 \end{bmatrix}$$

$$= \frac{1}{5} \begin{bmatrix} 3y_1 - 2y_2 + y_3 + y_4 \\ -2y_1 + 3y_2 + y_3 + y_4 \end{bmatrix}$$

$$\begin{bmatrix} \hat\beta_0 \\ \hat\beta_1 \end{bmatrix} = \begin{bmatrix} \dfrac{\sum_1^m y_i}{m} \\ \dfrac{-(n-m)\sum_1^n y_i + n \sum_{M+1}^n y_i}{m(n-M)} \end{bmatrix}$$

$$\hat\beta_1 = \left[-(n-m)\sum_1^m y_i - (n-m)\sum_{m+1}^n y_i + n\sum_{m+1}^n y_i\right] \times \frac{1}{m(n-m)}$$

$$= -\frac{\sum_i^m y_i}{m} + \frac{m\sum_{m+1}^n y_i}{m(n-M)}$$

$$= -\frac{\sum_i^m y_i}{m} + \frac{\sum_{m+1}^n y_i}{(n-m)}$$

$$\boxed{\begin{aligned} \hat\alpha_1 &= \frac{\sum_1^m y_i}{m} \\ \hat\alpha_2 &= \hat\beta_1 + \hat\alpha_1 = \frac{\sum_{m+1}^n y_i}{n-m} \end{aligned}}$$

OR simply,

$$\hat\alpha = (\bar X'\bar X)^{-1}\bar X'\bar Y \quad \text{Where } \bar X = \begin{bmatrix} 1_M & 0_M \\ 0_{n-M} & 1_{n-M} \end{bmatrix}$$

from ①

$$(\bar X'\bar X) = \begin{bmatrix} m & 0 \\ 0 & n-M \end{bmatrix}$$

$$(\bar X'\bar X)^{-1} = \begin{bmatrix} \frac{1}{M} & 0 \\ 0 & \frac{1}{n-M} \end{bmatrix} \qquad X'Y = \begin{bmatrix} \sum_{i=1}^M y_i \\ \sum_{i=M+1}^n y_i \end{bmatrix}$$

$$(\bar X'\bar X)^{-1}\bar X'\bar Y = \begin{bmatrix} \sum_{i=1}^m y_i / M \\ \sum_{i=m+1}^n y_i / n-M \end{bmatrix}$$

2.2 For $\hat{\beta}_1$

$$y_3 = \beta_1 + \beta_2 + e_3 \quad y_4 = \beta_1 + \beta_2 + e_4 \quad y_1 = \beta_1 + e_1$$

Adding y_3 and y_4 diminishes the effect of e_3 and e_4

$$y_3 + y_4 = 2\beta_1 + 2\beta_2 + e_3 + e_4$$

β_2 needs to be removed

$$y_3 + y_4 - 2y_2 = 2\beta_1 + e_3 + e_4 - 2e_2$$

y_1 needs to be included with more weight

$$y_3 + y_4 - 2y_2 + 3y_1 = 5\beta_1 + e_3 + e_4 - 2e_2 + 3e_1$$

Normalizing

$$\frac{1}{5}(y_3 + y_4 - 2y_2 + 3y_1) = \beta_1 + \frac{1}{5}(e_3 + e_4 - 2e_2 + 3e_1)$$

So intuitively, the expressions make sense.

Similarly $\hat{\beta}_2$ can be intuitively explained.

Solution 3

Given,

$\bar{X}_1, \bar{X}_2$ are column matrices

$\bar{X}_1'\bar{X}_1 = 1 \quad \Rightarrow \sum x_{j1}^2 = 1$

$\bar{X}_1' J = 0 \quad \Rightarrow \sum x_{j1} = 0$

Similarly,

$\sum x_{j2}^2 = 1$

$\sum x_{j2} = 0$

$\rho = \bar{X}_1'\bar{X}_2 = \bar{X}_2'\bar{X}_1$

$X = [1_n \ \bar{X}_1 \ \bar{X}_2]$

(3.1) $\quad \bar{X}'\bar{X} = \begin{bmatrix} 1_n' \\ \bar{X}_1' \\ \bar{X}_2' \end{bmatrix} [1_n \ \bar{X}_1 \ \bar{X}_2]$

$$= \begin{bmatrix} n & \sum x_{j1} & \sum x_{j2} \\ \sum x_{j1} & \bar{X}_1'\bar{X}_1 & \bar{X}_1'\bar{X}_2 \\ \sum x_{j2} & \bar{X}_2'\bar{X}_1 & \bar{X}_2'\bar{X}_2 \end{bmatrix} = \begin{bmatrix} n & 0 & 0 \\ 0 & 1 & \rho \\ 0 & \rho & 1 \end{bmatrix}$$

(3.2) $\quad (\bar{X}_1'\bar{X})^{-1}$

$$|\bar{X}_1'\bar{X}| = n \begin{vmatrix} 1 & \rho \\ \rho & 1 \end{vmatrix} = n(1-\rho^2)$$

$$(\bar{X}'\bar{X})^{-1} = \frac{1}{|\bar{X}'\bar{X}|}\begin{bmatrix} 1-\rho^2 & 0 & 0 \\ 0 & n & n\rho \\ 0 & n\rho & n \end{bmatrix}$$

(Using adjugate/adjoint)

$$= \begin{bmatrix} 1/n & 0 & 0 \\ 0 & 1/1-\rho^2 & \rho/1-\rho^2 \\ 0 & \rho/1-\rho^2 & 1/1-\rho^2 \end{bmatrix}$$

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

$$Var(\hat{\beta}_1) = Var(\hat{\beta}_2) = \frac{\sigma^2}{1-\rho^2} > 5\sigma^2$$

$$\frac{1}{1-\rho^2} > 5$$

$$1-\rho^2 < 1/5$$

$$\frac{4}{5} < \rho^2$$

$$\rho^2 > \frac{4}{5}$$

$$\rho \in \left[-1, \frac{-2}{\sqrt{5}}\right) \cup \left(\frac{2}{\sqrt{5}}, 1\right]$$

Solution 4

4.1

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix} = \begin{array}{c} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 3 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 3 & 1 \end{bmatrix} \\ \begin{array}{ccc} & \uparrow & \uparrow \\ & \bar{X}_1 & \bar{X}_2 \end{array} \end{array} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \bar{e}$$

To compute variance inflation factor of covariate $\bar{X}_1$, regress $\bar{X}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \\ 2 \\ 3 \end{bmatrix}$ on $\bar{X}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$. The

coefficient of determination R_1^2, calculated using R is $R_1^2 = 2.958e-31 \approx 0$

$\Rightarrow$ Variance Inflation factor for $\bar{X}_1 = \dfrac{1}{1-R_1^2} \approx 1$

4.2

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix} = \begin{array}{c} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 3 & 0 \\ 1 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 4 & 1 \end{bmatrix} \\ \begin{array}{ccc} & \uparrow & \uparrow \\ & \bar{X}_1 & \bar{X}_2 \end{array} \end{array} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \bar{e}$$

$R_1^2 = 0.2727$

$$V.I.f. \text{ for } X_1 = \frac{1}{1-R_1^2} = \frac{1}{1-0.2727} = 1.374948$$

It is larger than 4.1 because there is positive correlation between $\bar{X}_1$ and $\bar{X}_2$ in this case. The correlation among the predictors increases the variance of the estimated regression coefficients.

Solution 5

5.1 From Table 7.4, AIC and BIC have minimum value for subset size 2, while R^2_{adj} is max for size 2 as well as size 3.

So, subset size 2 with predictors X1 and X2 is the optimal model. $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + e$

5.2 From AIC based forward selection X3 is the only predictor because addition of any other predictor after that increases AIC value from -0.3087 to a positive value.

$$y = \beta_0 + \beta_3 X_3 + e$$

Similarly from BIC based forward selection, information criterion (BIC) increases if any other predictor is selected after X_3. from -1.089 to a positive value

$$y = \beta_0 + \beta_3 X_3 + e$$

5.3 Forward selection starts with one variable at a time. The variable which lowers AIC/BIC the most at that step is selected. X_3 was selected first because it explained the most of the variability in data among all predictors. This is also in agreement with R^2_{adj} data from table 7.4 where X_3 is best predictor for subset size 1. However, once X_3 is selected, the additional variability cannot be explained by (X_3, X_2) or (X_3, X_1) so $y \sim X_3$ is selected.

Since forward selection does not search over all possible subsets, (X_1, X_2) combination was never a choice.

On the other hand, when all subset combinations are tested (X_2, X_3) and $(X_1 X_2 X_3)$ are found to have higher R^2_{adj} and lower AIC and BIC than (X_3). So different results are obtained.

5.4 I would recommend $y = \beta_0 + \beta_3 X_3 + e$. Firstly, it can be seen directly from data (Table 7.3). Secondly, X_1 and X_2 have correlation ≈ -1, which inflates the variance of estimated coefficients thus inflating SSreg. Since $R^2 = \frac{SSreg}{SSreg + RSS} \approx \frac{SSreg}{SSreg} \approx 1$. So any comparison using increased R^2 is misleading. This is also true for $(X_1 X_2 X_3)$. Also, results from the fit $y \sim X1 + X2$ indicate -1000 of intercept and unusually low p-values, indicating possible overfitting. The overfitting leads to $RSS \to 0$, thus very low AIC, BIC values.