STAT 408/608 Homework 3 Solutions: Written Section

February 15, 2015

1. Geometrically $\hat{e}$ is orthogonal to the vector space spanned by the design matrix X. $v = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}'$ is in the space spanned by X, so v is orthogonal to the vector $\hat{e}$. That is, $\hat{e}'v = \sum_{i=1}^{5} e_i = 0$.

2. In this problem, the model $Y = X\beta + e$ can be written as:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 1 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} \beta_A \\ \beta_B \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix}$$

3. (a) In this problem, α_1 is the mean of the first group, and α_2 is the mean of the second group.

   (b) The design matrix:

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

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = (X'X)^{-1}X'Y = \begin{bmatrix} \frac{1}{m} & 0 \\ 0 & \frac{1}{n-m} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{m} y_i \\ \sum_{i=m+1}^{n} y_i \end{bmatrix} = \begin{bmatrix} \bar{y_1} \\ \bar{y_2} \end{bmatrix}$$

4. (a) In this problem, the model $Y = X\beta + e$ can be written as:

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

$$\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = (X'X)^{-1}X'Y = \begin{bmatrix} 3/5 & -2/5 \\ -2/5 & 3/5 \end{bmatrix} \begin{bmatrix} y_1 + y_3 + y_4 \\ y_2 + y_3 + y_4 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 3y_1 + (y_3 - y_2) + (y_4 - y_2) \\ 3y_2 + (y_3 - y_1) + (y_4 - y_1) \end{bmatrix}$$

(b) The result makes sense because the best estimates for each of the coins would be a combination of the four weights. It measured with more weight when the individual coins were measured on their own; less weight when the coins were measured together; and negative weight when the opposite coin was measured on its own.

5. D.
We know $SST = SSReg + RSS$. It is clear that $SSreg$ is close to zero if for each i, $\hat{y}_i$ is close to $\bar{y}$, while $SSreg$ is large if $\hat{y}_i$ differs from $\bar{y}$ for most values of x. Therefore, $SSreg$ for model 1 is greater than $SSreg$ for model 2, and RSS for model 1 is less than RSS for model 2.

6. (a) $y_i - \hat{y}_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i) = y_i - (\bar{y} - \hat{\beta}_1 \bar{x}) - \hat{\beta}_1 x_i = (y_i - \bar{y}) - \hat{\beta}_1 (x_i - \bar{x})$

   (b) $\hat{y}_i - \bar{y} = (\hat{\beta}_0 + \hat{\beta}_1 x_i) - \bar{y} = \bar{y} - \hat{\beta}_1 \bar{x} + \hat{\beta}_1 x_i - \bar{y} = \hat{\beta}_1 (x_i - \bar{x})$

   (c)
$$\begin{aligned}
\sum (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) &= \sum ((y_i - \bar{y}) - \hat{\beta}_1 (x_i - \bar{x}))(\hat{\beta}_1 (x_i - \bar{x})) \\
&= \sum (\hat{\beta}_1 (x_i - \bar{x})(y_i - \bar{y}) - \hat{\beta}_1^2 (x_i - \bar{x})^2) \\
&= \hat{\beta}_1 SXY - \hat{\beta}_1^2 SXX \\
&= \frac{SXY}{SXX} SXY - \frac{SXY}{SXX}^2 SXX \\
&= 0
\end{aligned}$$

7. (a) The assumptions is: the random errors e_i are i.i.d normal distributed, so Y is normal distributed. $\hat{\beta}$ is a linear combination of Y so $\hat{\beta}$ is normal.

   (b) If the sample size is larger, t distribution is quite robust. Even the model violates the i.i.d normality, we will see the distribution still approximate to t distribution. If the sample size is large enough, according to Central Limit Theory, even though the i.i.d normal assumption for the error term may not be met, we will have approximate normal distribution.

8.
$$\begin{aligned}
\Sigma &= E[(X - \mu)(X - \mu)'] \\
&= E[(X - \mu)(X' - \mu')] \\
&= E[XX' - \mu X' - X\mu' + \mu\mu'] \\
&= E[XX'] - \mu E[X'] - E[X]\mu' + \mu\mu' \\
&= E[XX'] - \mu\mu' - \mu\mu' + \mu\mu' \\
&= E[XX'] - \mu\mu'
\end{aligned}$$