STAT 408/608 Homework 5 Solutions: Written Section

March 8, 2015

1. (a)

$$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_{n_1} \\ Y_{n_1+1} \\ \vdots \\ Y_{n_1+n_2} \end{bmatrix}_{n_1+n_2}, \; X = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{n_1+n_2}$$

Then,

$$(X'X)^{-1} = \frac{1}{n_1+n_2}, \; X'Y = \sum_{i=1}^{n_1+n_2} Y_i, \; \hat{\alpha} = (X'X)^{-1}X'Y = \frac{1}{n_1+n_2}\sum_{i=1}^{n_1+n_2} Y_i = \bar{Y}.$$

(b) $Var(\sqrt{w_i}e_i) = w_i\sigma^2 = \sigma^2$, so $w_i = 1$ for $i = 1, \cdots, n_1$

$Var(\sqrt{w_i}e_i) = w_i\frac{\sigma^2}{2} = \sigma^2$, so $w_i = 2$ for $i = n_1+1, \cdots, n_1+n_2$

$W = diag(1, \cdots, 1, 2, \cdots, 2)_{(n_1+n_2)\times(n_1+n_2)}$

$\hat{\alpha}_{WLS} = (X'WX)^{-1}X'WY = \frac{1}{n_1+2n_2}\sum_{i=1}^{n_1} Y_i + 2\sum_{i=n_1+1}^{n_1+n_2} Y_i$

(c) let $n_1 = n_2 = n$,

$E(\hat{\alpha}) = E(\frac{1}{2n}\sum_{i=1}^{2n}) = E(Y_i) = \alpha$

$Var(\hat{\alpha}) = Var(\frac{1}{2n}\sum_{i=1}^{2n} Y_i) = \frac{1}{4n^2}Var(\sum_{i=1}^{2n} Y_i) = \frac{n\sigma^2 + n\sigma^2/2}{4n^2} = \frac{3\sigma^2}{8n}$

For WLS estimator,

$E(\hat{\alpha}_{WLS}) = E(\frac{1}{3n}\sum_{i=1}^{n} Y_i + 2\sum_{i=n+1}^{2n} Y_i) = \frac{n\alpha + 2n\alpha}{3n} = \alpha$

$Var(\hat{\alpha}_{WLS}) = Var(\frac{1}{3n}\sum_{i=1}^{n} Y_i + 2\sum_{i=n+1}^{2n} Y_i) = \frac{1}{9n^2}Var(\frac{1}{3n}\sum_{i=1}^{n} Y_i + 2\sum_{i=n+1}^{2n} Y_i) = \frac{n\sigma^2 + 4n\sigma^2/2}{9n^2} = \frac{\sigma^2}{3n}$

Both estimators are unbiased. However, the weighted least squares estimator in part (b) has smaller variance. It is a better estimator.

2. The weighted least squares of $y_i = \beta x_i + e_i$ model is $WRSS = \sum w_i(y_i - \hat{\beta}x_i)^2$
Then taking partial derivatives with respect to β

$$\frac{dWRSS}{d\beta} = -2\sum w_i(y_i - \hat{\beta}x_i)x_i = 0$$
$$\Rightarrow \sum w_i x_i y_i = \sum w_i \hat{\beta} x_i^2$$
$$\Rightarrow \hat{\beta} = \frac{\sum w_i x_i y_i}{\sum w_i x_i^2}$$

Here, let $w_i = \frac{1}{x_i^2}$, so that $Var(w_i e_i) = \sigma^2$.

Hence, $\hat{\beta} = \frac{\sum w_i x_i y_i}{\sum w_i x_i^2} = \frac{\sum \frac{y_i}{x_i}}{n} = \frac{1}{n}\sum \frac{y_i}{x_i}$

3. (a) We use the weighted least squares model when the response is an average or median based on a given sample size at each point. This is because the variance of the responses are proportional to the sample size at each point, it makes sense to take the weights to be $w_i = n_i$. This will make the variance proportional to a constant $Var(\sqrt{w_i}Y_i) \propto 1$.

   (b) It does not show the linear relationship between the response and each of the explanatory variables, and it appears a relationship between the two explanatory variables which might cause multicolinearity problems. The residuals seem not be normally distributed and it shows non-constant variance. The standardized residuals plot reveals the outliers as well. Therefore, this model is not a valid model because of the lack of assumptions.

   (c) I would try to do the transformation for either y_i, x_1 or x_2 or all three. By tracking the diagnostic plots, we could find the best transformation result which is the best linear fit without breaking any assumptions.

4. (a) $Var(y) \propto$ number of coins:
$$W = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

   (b) $\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = (X'WX)^{-1}(X'WY) = \frac{1}{3}\begin{bmatrix} 2y_1 - y_2 + 0.5y_3 + 0.5y_4 \\ 2y_2 - y_1 + 0.5y_3 + 0.5y_4 \end{bmatrix}$

(c)

$$E\begin{bmatrix}\hat{\beta}_1\\ \hat{\beta}_2\end{bmatrix} = \frac{1}{3}\begin{bmatrix}2E(y_1) - E(y_2) + 0.5E(y_3) + 0.5E(y_4)\\ 2E(y_2) - E(y_1) + 0.5E(y_3) + 0.5E(y_4)\end{bmatrix}$$
$$= \frac{1}{3}\begin{bmatrix}2\beta_1 - \beta_2 + 0.5(\beta_1+\beta_2) + 0.5(\beta_1+\beta_2)\\ 2\beta_2 - \beta_1 + 0.5(\beta_1+\beta_2) + 0.5(\beta_1+\beta_2)\end{bmatrix}$$
$$= \frac{1}{3}\begin{bmatrix}3\beta_1\\ 3\beta_2\end{bmatrix}$$
$$= \begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}$$

Therefore the WLS estimates are unbiased.
The individual observations are weighted more heavily than the combined observations to stabilize the variance. In other words, if the error variance for measuring each coin is constant, then the error variance for measuring two coins together will be larger. Therefore, we should give more weight to individual measurements.

5. (a) Not BLUE. Order statistics are not linear. The estimator is not a linear combination of y.

   (b) Not BLUE. $E(3y_{11} - y_{12} - y_{13} - y_{14} - y_{15}) = 3\alpha_1 - \alpha_1 - \alpha_1 - \alpha_1 - \alpha_1 = -\alpha_1$. Biased.

   (c) Not BLUE.

   i. Unbiased. $E(\tilde{\beta_1}) = \frac{1}{5}(\beta_0 + 2\beta_0 - 2\beta_0 - \beta_0 + 4\beta_1 + 2\times 3\beta_1 - 2\times 2\beta_1 - \beta_1) = \beta_1$.

   ii. $Var(\tilde{\beta_1}) = \frac{1}{25}Var(e_4 + 2e_3 - 2e_2 - e_1) = \frac{1}{25}(1+4+4+1)\sigma^2 = \frac{2}{5}\sigma^2$.

   iii. $(X'X)^{-1} = \frac{1}{20}\begin{bmatrix}30 & -10\\ -10 & 4\end{bmatrix}$

   $\hat{\beta} = (X'X)^{-1}X'Y$, $Var(\hat{\beta}) = (X'X)^{-1}\sigma^2$. Thus $Var(\hat{\beta_1}) = \frac{4}{20}\sigma^2 = 0.2\sigma^2$.

   (d) $Var(\tilde{\beta_1}) > Var(\hat{\beta_1})$. $\hat{\beta}$ is BLUE.

6. (a) No. The question asked that after taking into account foot length, whether there are gender differences in foot width. An interaction term means the relationship between foot width and foot length is different across gender.

   (b) $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + e$, where,
   Y: foot width (inches)
   x_1: foot length (inches)
   x_2: gender $\begin{cases}1, boys\\ 0, girls\end{cases}$
   β_0: Mean foot width for girls with foot length of 0
   β_1: By given other variables, 1 unit increase in foot length will increase 1 unit in

foot width
β_2: By given other variables, the foot width difference between boys and girls

(c) Here is the made up data for the 5 observations:
Observation 1: width = 3, length = 8, gender = F
Observation 2: width = 3.5, length = 9, gender = M
Observation 3: width = 4, length = 10.5, gender = M
Observation 4: width = 4.5, length = 11.5, gender = M
Observation 5: width = 3.25, length = 9.5, gender = F

$$X = \begin{bmatrix} 1 & 8 & 0 \\ 1 & 9 & 1 \\ 1 & 10.5 & 1 \\ 1 & 11.5 & 1 \\ 1 & 9.5 & 0 \end{bmatrix}$$

(d) First, we test the model. The hypotheses:
$H_0 : \beta_0 = \beta_1 = \beta_2 = 0$

H_1: at least one $beta_i \neq 0$
The test Statistics $F = \dfrac{SSReg/p}{SSE/n-p-1}$. If this test is significant, then can do the following test.
The researcher wondered whether 4th grade girls actually had narrower feet than boys. We should test hypothesis $H_0 : \beta_2 = 0 v.s. H_1 : \beta_2 > 0$.
The test Statistics $T = \dfrac{\hat{\beta}_i - \beta}{se(\hat{\beta}_i)}$. If this test is significant, then we can say that girls will have a smaller foot width in average compared to boys when their foot length are the same.