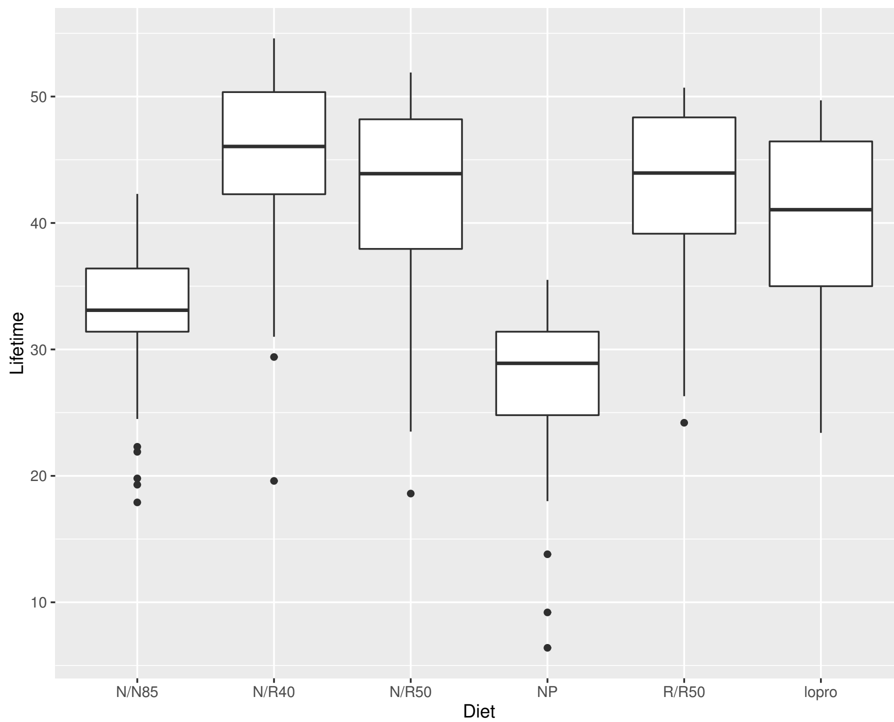
STAT 510 Homework 3

Yifan Zhu

January 31, 2017

1. (a)



(b)

$$SSE_{full} = 15297$$

(c)

$$\hat{\sigma^2} = \frac{SSE_{full}}{n-r} = \frac{15297}{343} = 44.60$$

(d)

$$SSE_{reduced} = 15511$$

(e)

$$F = \frac{(SSE_{reduced} - SSE_{full})/(DFE_{reduced} - DFE_{full})}{SSE_{full}/DFE_{full}} = \frac{(15511 - 15297)/1}{15297/343} = 4.8$$

(f) This test is to test the among the mice having these six diet plans, wheatehr the mice having diet plan N/R50 and N/R50 lopro have the same population life time.

(g) If the parameter vector $\boldsymbol{\beta} = \begin{bmatrix}\beta_1 & \beta_2 & \beta_2 & \beta_4 & \beta_5 & \beta_6\end{bmatrix}$ are in the order for N/N85, N/R40, N/R50, NP, R/R50 and N/R50 lopro, then

$$\boldsymbol{C} = \begin{bmatrix}0 & 0 & 1 & 0 & 0 & -1\end{bmatrix},\ d = 0$$

Then

$$F = \frac{(\boldsymbol{C}\hat{\boldsymbol{\beta}} - d)^T(\boldsymbol{C}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{C}^T)^{-1}(\boldsymbol{C}\hat{\boldsymbol{\beta}} - d)/rank(\boldsymbol{C})}{\hat{\sigma^2}} = 4.8$$

where $\boldsymbol{X}$ is the model matrix for full model.
The F statistic is the same as is computed in (e).

2. Let $\boldsymbol{X} = \begin{bmatrix}1 & 1\end{bmatrix}$, then $\boldsymbol{A} = \boldsymbol{X}^T\boldsymbol{X} = \begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$ is symmetric. Let $\boldsymbol{G} = \begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}$, then

$$\boldsymbol{AGA} = \begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} = \begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} = \begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} = \boldsymbol{A}$$

Thus $\boldsymbol{G}$ is a generalized inverse of the symmetric metrix $\boldsymbol{A}$, but it is not symmetric.

3. (a)

$$\begin{aligned}
&rank(\boldsymbol{X}^T\boldsymbol{X}) \le rank(\boldsymbol{X})\\
&rank(\boldsymbol{X}) = rank(\boldsymbol{P_X}) = rank(\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T\boldsymbol{X}) \le rank(\boldsymbol{X}^T\boldsymbol{X})\\
\Rightarrow &rank(\boldsymbol{X}) = rank(\boldsymbol{X}^T\boldsymbol{X})
\end{aligned}$$

(b)

$$\begin{aligned}
&rank(\boldsymbol{X}) = rank(\boldsymbol{P_X}\boldsymbol{X}) \le rank(\boldsymbol{P_X})\\
&rank(\boldsymbol{P_X}) = rank(\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}) \le rank(\boldsymbol{X})\\
\Rightarrow &rank(\boldsymbol{X}) = rank(\boldsymbol{P_X})
\end{aligned}$$

(c) First we can prove $rank(\boldsymbol{AP_X}) = rank(\boldsymbol{AP_X})$.

$$\begin{aligned}
&rank(\boldsymbol{AX}) = rank(\boldsymbol{AP_X X}) \le rank(\boldsymbol{AP_X})\\
&rank(\boldsymbol{AP_X}) = rank(\boldsymbol{AX}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T) \le rank(\boldsymbol{AX})\\
\Rightarrow &rank(\boldsymbol{AX}) = rank(\boldsymbol{AP_X})
\end{aligned}$$

Then we have

$$\boldsymbol{C}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{C}^T = \boldsymbol{AX}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T\boldsymbol{A}^T = \boldsymbol{AP_X A}^T = \boldsymbol{AP_X P_X A}^T$$

Thus $rank(\boldsymbol{C}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{C}^T) = rank(\boldsymbol{AP_X}(\boldsymbol{AP_X})^T) = rank(\boldsymbol{AP_X}) = rank(\boldsymbol{AX}) = rank(\boldsymbol{C}) = q$.

(d) From (c) we know $rank(\boldsymbol{AX}) = rank(\boldsymbol{AP_X}) = rank(\boldsymbol{A})$.

4. (a) Let the parameter vector in such an order

$$\begin{bmatrix}(S1, F1, 25\%)\\ (S2, F1, 25\%)\\ (S1, F2, 25\%)\\ (S2, F2, 25\%)\\ (S1, F1, 50\%)\\ (S2, F1, 50\%)\\ (S1, F2, 50\%)\\ (S2, F2, 50\%)\\ (S1, F1, 75\%)\\ (S2, F1, 75\%)\\ (S1, F2, 75\%)\\ (S2, F2, 75\%)\end{bmatrix}$$

Then

$$\boldsymbol{C}_{(a)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \; \hat{\mu}_{F2,S1,50\%} = \boldsymbol{C}_{(a)}\hat{\boldsymbol{\beta}} = 233.5$$

(b)

$$\boldsymbol{C}_{(b)} = \begin{bmatrix} 1/6 & 1/6 & 0 & 0 & 1/6 & 1/6 & 0 & 0 & 1/6 & 1/6 & 0 & 0 \\ 0 & 0 & 1/6 & 1/6 & 0 & 0 & 1/6 & 1/6 & 0 & 0 & 1/6 & 1/6 \end{bmatrix}, \; \begin{bmatrix} \hat{\bar{\mu}}_{F1,.,.} \\ \hat{\bar{\mu}}_{F2,.,.} \end{bmatrix} = C_{(b)}\hat{\boldsymbol{\beta}} = \begin{bmatrix} 214.75 \\ 181.08 \end{bmatrix}$$

(c)

$$\boldsymbol{C}_{(c)} = \begin{bmatrix} 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \; \hat{\bar{\mu}}_{.,S2,25\%} = \boldsymbol{C}_{(c)}\hat{\boldsymbol{\beta}} = 156.25$$

(d)

$$se_{(c)} = \sqrt{\hat{\sigma^2}\boldsymbol{C}_{(c)}(\boldsymbol{X}^T\boldsymbol{X})^{-}\boldsymbol{C}_{(c)}^T} = 8.20$$

(e)

$$H_0 : \bar{\mu}_{F1,.,.} = \bar{\mu}_{F2,.,.} (\boldsymbol{C\beta} - d = 0)$$
$$H_a : \bar{\mu}_{F1,.,.} \neq \bar{\mu}_{F2,.,.} (\boldsymbol{C\beta} - d \neq 0)$$

$$\boldsymbol{C} = \begin{bmatrix} 1/6 & 1/6 & -1/6 & -1/6 & 1/6 & 1/6 & -1/6 & -1/6 & 1/6 & 1/6 & -1/6 & -1/6 \end{bmatrix}, \; d = 0$$

$$F = \frac{(\boldsymbol{C}\hat{\boldsymbol{\beta}} - d)^T(\boldsymbol{C}(\boldsymbol{X}^T\boldsymbol{X})^{-}\boldsymbol{C}^T)^{-1}(\boldsymbol{C}\hat{\boldsymbol{\beta}} - d)/rank(\boldsymbol{C})}{\hat{\sigma^2}} = 25.3$$

p-value = 0.0003

Conclusion : Since p-value is small, we decide to reject the null hypothesis and conclude that there is filler type main effect.

(f)

$$H_0 : \text{there is no three way interaction.} \quad (\boldsymbol{C\beta} - \boldsymbol{d} = 0)$$
$$H_a : \text{there is three way interaction.} \quad (\boldsymbol{C\beta} - \boldsymbol{d} \neq 0)$$

$$\boldsymbol{C} = \begin{bmatrix} 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 & -1 & 1 & 1 & -1 \end{bmatrix}, \; \boldsymbol{d} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$F = \frac{(\boldsymbol{C}\hat{\boldsymbol{\beta}} - d)^T(\boldsymbol{C}(\boldsymbol{X}^T\boldsymbol{X})^{-}\boldsymbol{C}^T)^{-1}(\boldsymbol{C}\hat{\boldsymbol{\beta}} - d)/rank(\boldsymbol{C})}{\hat{\sigma^2}} = 0.89$$

p-value = 0.44

Conclusion : Since p-value is big, we fail to reject the null hypothesis and conclude that there is no three way interaction.