

1.

We first have,

$$F = \frac{SS_{\text{exp}}/p}{SS_{\text{res}}/(n-p-1)}$$

hence

$$\frac{F}{F+c} = \frac{1}{1+c/F}$$

and

$$1 - \frac{SS_{\text{res}}}{SS_{\text{total}}} = \frac{SS_{\text{exp}}}{SS_{\text{total}}}$$

Therefore,

$$\begin{aligned} 1 + \frac{c}{F} &= \frac{SS_{\text{total}}}{SS_{\text{exp}}} \\ \implies 1 + \frac{c(n-p-1)}{p} \frac{SS_{\text{res}}}{SS_{\text{exp}}} &= \frac{SS_{\text{total}}}{SS_{\text{exp}}} \\ \implies SS_{\text{exp}} + \frac{c(n-p-1)}{p} SS_{\text{res}} &= SS_{\text{total}} \\ \implies c &= \frac{p}{n-p-1} \end{aligned}$$

We then have

$$5 = \frac{8}{n-8-1} \implies n = 10.6$$

Hence, the population need to be ≥ 11 .

2.

We first have that

$$X^T X = \begin{pmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix} = \begin{pmatrix} 45 & 0 \\ 0 & \sum x_i^2 \end{pmatrix}$$

and thus the inverse is

$$\begin{pmatrix} \frac{1}{45} & 0 \\ 0 & \frac{1}{\sum x_i^2} \end{pmatrix}$$

then we can calculate

$$\text{se}(c) = \sqrt{\frac{1}{45} SS_{\text{res}}/43}$$

where

$$SS_{\text{res}} = \sum_{i=1}^{45} (Y_i - \hat{Y}_i)^2$$

Thus, we can find the confidence interval for c

$$\hat{c} - t_{\alpha/2, 43} \sqrt{\frac{SS_{\text{res}}}{1935}} < c < \hat{c} + t_{\alpha/2, 43} \sqrt{\frac{SS_{\text{res}}}{1935}}$$

Similarly,

$$\text{se}(c) = \sqrt{\frac{1}{\sum x_i^2} SS_{\text{res}}/43}$$

and the confidence for d is

$$\hat{d} - t_{\alpha/2, 43} \sqrt{\frac{SS_{\text{res}}}{43 \sum_{i=1}^{45} x_i^2}} < d < \hat{d} + t_{\alpha/2, 43} \sqrt{\frac{SS_{\text{res}}}{43 \sum_{i=1}^{45} x_i^2}}$$

Thus the $1 - \alpha$ confidence interval for $c + 5d$ is

$$\hat{c} + 5\hat{d} - t_{\alpha/2, 43} \left(\sqrt{\frac{SS_{\text{res}}}{1935}} + 5 \sqrt{\frac{SS_{\text{res}}}{43 \sum x_i^2}} \right) < c + 5d < \hat{c} + 5\hat{d} + t_{\alpha/2, 43} \left(\sqrt{\frac{SS_{\text{res}}}{1935}} + 5 \sqrt{\frac{SS_{\text{res}}}{43 \sum x_i^2}} \right)$$

where \hat{c} and \hat{d} can be estimated by MLE

$$\begin{aligned} \mathcal{L}(c, d, \sigma^2 | X) &= \prod_{i=1}^{45} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(y_i - c - dx_i)^2}{2\sigma^2}\right) \\ &= \frac{1}{\sqrt{(2\pi\sigma^2)^{45}}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{45} (y_i - c - dx_i)^2\right) \end{aligned}$$

then

$$l(c, d, \sigma^2 | X) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{45} (y_i - c - dx_i)^2$$

Thus

$$\frac{dl}{dc} = \frac{1}{\sigma^2} \sum_{i=1}^{45} (y_i - c - dx_i) \implies \hat{c} = \bar{y} - \hat{d}\bar{x}$$

and

$$\begin{aligned} \frac{dl}{dd} &= \frac{1}{\sigma^2} \sum_{i=1}^{45} (x_i y_i - cx_i - dx_i^2) \\ \implies 0 &= \frac{1}{\hat{\sigma}^2} \sum_{i=1}^{45} (x_i y_i - \hat{c}x_i - \hat{d}x_i^2) \\ \implies 0 &= \sum_{i=1}^{45} (x_i y_i - (\bar{y} - \hat{d}\bar{x})x_i - \hat{d}x_i^2) \\ \implies 0 &= \sum_{i=1}^{45} x_i y_i - 45\bar{x}\bar{y} + \hat{d}45\bar{x}^2 - \hat{d} \sum_{i=1}^{45} x_i^2 \\ \implies \hat{d} &= \frac{\sum_{i=1}^{45} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{45} (x_i - \bar{x})^2} \end{aligned}$$

3.

First, let's calculate

$$x^T(X^TX)^{-1}x = (1 \quad 7) \begin{pmatrix} \frac{1}{n} & 0 \\ 0 & \frac{1}{\sum x_i^2} \end{pmatrix} \begin{pmatrix} 1 \\ 7 \end{pmatrix} = \frac{1}{n} + \frac{1}{49 \sum x_i^2}$$

We can just apply the formula and get the prediction interval for when $x = 7$,

$$\hat{c} + 7\hat{d} - t_{\alpha/2, 43} \sqrt{\frac{SS_{\text{res}}}{43} \left(\frac{1}{n} + \frac{1}{49 \sum x_i^2} \right)} \leq E(\hat{Y}_0 | X_0 = x) \leq \hat{c} + 7\hat{d} + t_{\alpha/2, 43} \sqrt{\frac{SS_{\text{res}}}{43} \left(\frac{1}{n} + \frac{1}{49 \sum x_i^2} \right)}$$