15 Sept

$$\hat{\underset{\sim}{\beta}} = (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} \underset{\sim}{X}^T \underset{\sim}{Y}$$

① $p = 1$ Model : $y_i = \beta_0 + \beta_1 x_{i1} + e_i \quad i = 1, \dots, n$

OR Model : $\underset{\sim}{Y} = \underset{\sim}{X} \underset{\sim}{\beta} + \underset{\sim}{e}$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad \begin{pmatrix} 1 & x_{11} \\ \vdots & \vdots \\ 1 & x_{n1} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \underset{\sim}{e}$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

$$\underset{\sim}{X}^T \underset{\sim}{X} = \begin{pmatrix} n & \sum_{i=1}^n x_{i1} \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_{i1}^2 \end{pmatrix} \qquad (\underset{\sim}{X}^T \underset{\sim}{X})^{-1} = \frac{1}{n \sum_{i=1}^n x_{i1}^2 - (\sum_{i=1}^n x_{i1})^2} \begin{pmatrix} \sum_{i=1}^n x_{i1}^2 & -\sum_{i=1}^n x_{i1} \\ -\sum_{i=1}^n x_{i1} & n \end{pmatrix}$$

$$\underset{\sim}{X}^T \underset{\sim}{Y} = \begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_{i1} y_i \end{pmatrix}$$

② ~~a~~ β_0 is known , $p = 1$

Model : $y_i = \boxed{\beta_0} + \beta_1 x_{i1} + e_i \quad i = 1, \dots, n$

known

$$\Rightarrow \underbrace{y_i - \beta_0}_{y_i'} = \beta_1 x_{i1} + e_i$$

$$\text{Model} = \begin{pmatrix} y_1 - \beta_0 \\ \vdots \\ y_n - \beta_0 \end{pmatrix} = \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix}_{n \times 1} \beta_1 + \underset{\sim}{e}$$

$$\Rightarrow \hat{\beta}_1 = \left(\sum_{i=1}^n x_{i1}^2 \right)^{-1} \sum_{i=1}^n x_{i1} (y_i - \beta_0)$$

$$= \frac{\sum_{i=1}^n x_{i1} y_i - \beta_0 \sum_{i=1}^n x_{i1}}{\sum_{i=1}^n x_{i1}^2}$$

(3) Example in p.7 $\qquad \beta_0 = 40$

Model : $y_i = \boxed{\beta_0} + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + e_i$

$$y_i - \beta_0 = \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + e_i$$

$$\underset{\sim}{Y} = \begin{pmatrix} y_1 - \beta_0 \\ \vdots \\ y_n - \beta_0 \end{pmatrix} \qquad \underset{n\times 3}{\underset{\sim}{X}} = \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ \vdots & \vdots & \vdots \\ X_{n1} & X_{n2} & X_{n3} \end{pmatrix} \qquad \underset{\sim}{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

Can I use the centered model ?

Original model $\qquad y_i' = \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + e_i$

Centered model $\qquad y_i' = \beta_1'(X_{i1} - \bar{X}_1) + \beta_2'(X_{i2} - \bar{X}_2) + \beta_3'(X_{i3} - \bar{X}_3) + e_i$

$$\Rightarrow y_i' = \underbrace{\boxed{-\beta_1'\bar{X}_1 - \beta_2'\bar{X}_2 - \beta_3'\bar{X}_3}}_{\beta_0'} + \beta_1' X_{i1} + \beta_2' X_{i2} + \beta_3' X_{i3} + e_i$$

$\beta_0 = 40$

Example 5: Example in Multiple Linear Regression

The percent survival of a certain type of animal semen after storage was measured at various combinations of concentrations of three materials used to increase chance of survival. The data are as follows:

y (% survival)	x_1 (weight %)	x_2 (weight %)	x_3 (weight %)
25.5	1.74	5.30	10.80
31.2	6.32	5.42	9.40
25.9	6.22	8.41	7.20
38.4	10.52	4.63	8.50
18.4	1.19	11.60	9.40
26.7	1.22	5.85	9.90
26.4	4.10	6.62	8.00
25.9	6.32	8.72	9.10
32.0	4.08	4.42	8.70
25.2	4.15	7.60	9.20
39.7	10.15	4.83	9.40
35.7	1.72	3.12	7.60
26.5	1.70	5.30	8.20

Summary statistics:

$$\sum_{i=1}^{13} y_i = 377.5 \qquad \sum_{i=1}^{13} y_i^2 = 11,400.15 \qquad \sum_{i=1}^{13} x_{i1} = 59.43$$

$$\sum_{i=1}^{13} x_{i2} = 81.82 \qquad \sum_{i=1}^{13} x_{i3} = 115.40 \qquad \sum_{i=1}^{13} x_{i1}^2 = 394.7255$$

$$\sum_{i=1}^{13} x_{i2}^2 = 576.7264 \qquad \sum_{i=1}^{13} x_{i3}^2 = 1035.96 \qquad \sum_{i=1}^{13} x_{i1}y_i = 1877.567$$

$$\sum_{i=1}^{13} x_{i2}y_i = 2246.661 \qquad \sum_{i=1}^{13} x_{i3}y_i = 3337.78 \qquad \sum_{i=1}^{13} x_{i1}x_{i2} = 360.6621$$

$$\sum_{i=1}^{13} x_{i1}x_{i3} = 522.078 \qquad \sum_{i=1}^{13} x_{i2}x_{i3} = 728.31 \qquad n = 13$$

Model: $y_i - 40 = \beta_1' X_{i1} + \beta_2' X_{i2} + \beta_3' X_{i3} + e_i \quad i = 1, \dots, n$

$$\begin{pmatrix} 13 & 59.43 & 81.82 & 115.40 \\ 59.43 & 394.7255 & 360.6621 & 522.078 \\ 81.82 & 360.6621 & 576.7264 & 728.31 \\ 115.40 & 522.078 & 728.31 & 1035.96 \end{pmatrix}^{-1} = \begin{pmatrix} 8.06479 & -0.0825927 & -0.0941951 & -0.790527 \\ -0.0825927 & 0.00847982 & 0.00171669 & 0.00372002 \\ -0.0941951 & 0.00171669 & 0.0166294 & -0.00206331 \\ -0.790527 & 0.00372002 & -0.00206331 & 0.0886013 \end{pmatrix}$$

Or $\underset{\sim}{X}^T\underset{\sim}{X}$ $\quad (\underset{\sim}{X}^T\underset{\sim}{X})^{-1} = \begin{pmatrix} 0.0076339722 & 0.0075202117 & -0.04375877 \\ & 0.0155272451 & -0.011296496 \\ & & 0.0111123033 \end{pmatrix}$

$$(\underset{\sim}{X}_c^T \underset{\sim}{X}_c)^{-1} = \begin{pmatrix} 13 & 0 & 0 & 0 \\ 0 & 123.039 & -13.3812 & -5.4775 \\ 0 & -13.3812 & 61.7639 & 2.0002 \\ 0 & -5.4775 & 2.0002 & 11.5631 \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} 0.0769231 & 0 & 0 & 0 \\ 0 & 0.00847981 & 0.00171669 & 0.00371998 \\ 0 & 0.00171669 & 0.0166294 & -0.00206338 \\ 0 & 0.00371998 & -0.00206338 & 0.0886011 \end{pmatrix}$$

$\Rightarrow \hat{\beta}_0 = 39.1574$, $\hat{\beta}_1 = 1.0161$, $\hat{\beta}_2 = -1.8616$, $\hat{\beta}_3 = -0.3433$.

$$\underset{\sim}{X}^T\underset{\sim}{Y} = \begin{pmatrix} \sum_{i=1}^{n} X_{i1}(y_i - 40) \\ \sum_{i=1}^{n} X_{i2}(y_i - 40) \\ \sum_{i=1}^{n} X_{i3}(y_i - 40) \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} X_{i1} y_i - 40 \sum_{i=1}^{n} X_{i1} \\ \sum_{i=1}^{n} X_{i2} y_i - 40 \sum_{i=1}^{n} X_{i2} \\ \sum_{i=1}^{n} X_{i3} y_i - 40 \sum_{i=1}^{n} X_{i3} \end{pmatrix} = \begin{pmatrix} -499.633 \\ -1026.119 \\ -1278.22 \end{pmatrix}$$

$$\hat{\beta} = (\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{Y} = \begin{pmatrix} 1.00748 \\ -1.87149 \\ -0.425862 \end{pmatrix}$$

④ <u>Example in p.7</u> $\beta_1 = 1$ (one of reg. coeff. is known)

Model : $y_i = \beta_0 + \underset{\overset{\parallel}{1}}{\boxed{\beta_1}} X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + e_i$ $i = 1, —, n$

$$\underbrace{y_i - \underset{\overset{\parallel}{1}}{\boxed{\beta_1}} X_{i1}}_{y_i'} = \beta_0 + \beta_2 X_{i2} + \beta_3 X_{i3} + e_i$$

$$\Rightarrow \underset{\underset{\sim}{Y}}{\begin{pmatrix} y_1 - \beta_1 X_{11} \\ \vdots \\ y_n - \beta_1 X_{n1} \end{pmatrix}} = \underset{\underset{\sim}{X}}{\begin{pmatrix} 1 & X_{12} & X_{13} \\ 1 & X_{22} & X_{23} \\ \vdots & & \vdots \\ 1 & X_{n2} & X_{n3} \end{pmatrix}} \underset{\underset{\sim}{\beta}}{\begin{pmatrix} \beta_0 \\ \beta_2 \\ \beta_3 \end{pmatrix}} + \underset{\sim}{e}$$

$\beta_1 = 1$ Model: $y_i - x_{i1} = \beta_0 + \beta_2 x_{i2} + \beta_3 x_{i3} + e_i$

Centred model: $y_i - x_{i1} = \beta_0' + \beta_2'(x_{i2} - \bar{x}_2) + \beta_3'(x_{i3} - \bar{x}_3) + e_i$

$\beta_2' = \beta_2$, $\beta_3' = \beta_3$

Example 5: Example in Multiple Linear Regression

The percent survival of a certain type of animal semen after storage was measured at various combinations of concentrations of three materials used to increase chance of survival. The data are as follows:

$\hat{\beta}_0' = \bar{y}' \Rightarrow \hat{\beta}_0 = \bar{y}' - \beta_2\bar{x}_2 - \beta_3\bar{x}_3$

$\boxed{\bar{y}' = \bar{y} - \bar{x}_1}$

y (% survival)	x_1 (weight %)	x_2 (weight %)	x_3 (weight %)
25.5	1.74	5.30	10.80
31.2	6.32	5.42	9.40
25.9	6.22	8.41	7.20
38.4	10.52	4.63	8.50
18.4	1.19	11.60	9.40
26.7	1.22	5.85	9.90
26.4	4.10	6.62	8.00
25.9	6.32	8.72	9.10
32.0	4.08	4.42	8.70
25.2	4.15	7.60	9.20
39.7	10.15	4.83	9.40
35.7	1.72	3.12	7.60
26.5	1.70	5.30	8.20

Summary statistics:

$\sum_{i=1}^{13} y_i = 377.5$ $\quad \sum_{i=1}^{13} y_i^2 = 11,400.15$ $\quad \sum_{i=1}^{13} x_{i1} = 59.43$

$\sum_{i=1}^{13} x_{i2} = 81.82$ $\quad \sum_{i=1}^{13} x_{i3} = 115.40$ $\quad \sum_{i=1}^{13} x_{i1}^2 = 394.7255$

$\sum_{i=1}^{13} x_{i2}^2 = 576.7264$ $\quad \sum_{i=1}^{13} x_{i3}^2 = 1035.96$ $\quad \sum_{i=1}^{13} x_{i1}y_i = 1877.567$

$\sum_{i=1}^{13} x_{i2}y_i = 2246.661$ $\quad \sum_{i=1}^{13} x_{i3}y_i = 3337.78$ $\quad \sum_{i=1}^{13} x_{i1}x_{i2} = 360.6621$

$\sum_{i=1}^{13} x_{i1}x_{i3} = 522.078$ $\quad \sum_{i=1}^{13} x_{i2}x_{i3} = 728.31$ $\quad n = 13$

Original model

$$\begin{pmatrix} 13 & 59.43 & 81.82 & 115.40 \\ \cancel{59.43} & \cancel{394.7255} & \cancel{360.6621} & \cancel{522.078} \\ 81.82 & 360.6621 & 576.7264 & 728.31 \\ 115.40 & 522.078 & 728.31 & 1035.96 \end{pmatrix}^{-1} = \begin{pmatrix} 8.06479 & -0.0825927 & -0.0941951 & -0.790527 \\ -0.0825927 & 0.00847982 & 0.00171669 & 0.00372002 \\ -0.0941951 & 0.00171669 & 0.0166294 & -0.00206331 \\ -0.790527 & 0.00372002 & -0.00206331 & 0.0886013 \end{pmatrix}$$

Or $\underset{\sim}{X}^T\underset{\sim}{X} = \begin{pmatrix} 13 & 81.82 & 115.40 \\ 81.82 & 576.7264 & 728.31 \\ 115.40 & 728.31 & 1035.96 \end{pmatrix}$ $\quad (\underset{\sim}{X}^T\underset{\sim}{X})^{-1}$ $\quad \underset{\sim}{X}^T\underset{\sim}{Y} = \begin{pmatrix} \sum_{i=1}^{13}(y_i - x_{i1}) \\ \sum_{i=1}^{13} x_{i2}(y_i - x_{i1}) \\ \sum_{i=1}^{13} x_{i3}(y_i - x_{i1}) \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{13} y_i - \sum_{i=1}^{13} x_{i1} \\ \sum_{i=1}^{13} x_{i2}y_i - \sum_{i=1}^{13} x_{i1}x_{i2} \\ \sum_{i=1}^{13} x_{i3}y_i - \sum_{i=1}^{13} x_{i1}x_{i3} \end{pmatrix}$

Centred model

$$(\underset{\sim}{X}_c^T\underset{\sim}{X}_c)^{-1} = \begin{pmatrix} 13 & 0 & 0 & 0 \\ \cancel{0} & \cancel{123.039} & \cancel{-13.3812} & \cancel{-5.4775} \\ 0 & -13.3812 & 61.7639 & 2.0002 \\ 0 & -5.4775 & 2.0002 & 11.5631 \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} 0.0769231 & 0 & 0 & 0 \\ 0 & 0.00847981 & 0.00171669 & 0.00371998 \\ 0 & 0.00171669 & 0.0166294 & -0.00206338 \\ 0 & 0.00371998 & -0.00206338 & 0.0886011 \end{pmatrix}$$

$\Rightarrow \hat{\beta}_0 = 39.1574$, $\hat{\beta}_1 = 1.0161$, $\hat{\beta}_2 = -1.8616$, $\hat{\beta}_3 = -0.3433$.

$\underset{\sim}{X}^T\underset{\sim}{Y} = \begin{pmatrix} \sum_{i=1}^{13}(y_i - x_{i1}) \\ \boxed{\sum_{i=1}^{13}(x_{i2} - \bar{x}_2)(y_i - x_{i1})} \\ \sum_{i=1}^{13}(x_{i3} - \bar{x}_3)(y_i - x_{i1}) \end{pmatrix}$

$\sum_{i=1}^{13}(x_{i2} - \bar{x}_2)(y_i - x_{i1})$

$= \boxed{\sum_{i=1}^{13}(x_{i2} - \bar{x}_2)y_i} - \sum_{i=1}^{13}(x_{i2} - \bar{x}_2)x_{i1}$

$= \sum_{i=1}^{13}(x_{i2} - \bar{x}_2) y_i - \boxed{\bar{y}\sum_{i=1}^{13}(x_{i2} - \bar{x}_2)} = 0$

$= S_{x_2y} - S_{x_1x_2}$

$\hat{\beta}_0 = 39.3147$, $\hat{\beta}_2 = -1.86491$, $\hat{\beta}_3 = -0.35032$

Fitted line: $\hat{y}' = y \boxed{- x_1} = 39.3147 - 1.86491 x_2 - 0.35032 x_3$

Method of maximum likelihood (3423)

$\underline{p=1}$ Model : $y_i = \beta_0 + \beta_1 x_{i1} + e_i$ Assume $e_i \sim N(0, \sigma^2)$

$\mathrm{cov}(e_i, e_j) = 0$

$\Rightarrow y_i \sim N(\beta_0 + \beta_1 x_{i1}, \sigma^2)$ $i = 1, \text{---}, n$

likelihood $(\beta_0, \beta_1, \sigma^2)$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{(y_i - (\beta_0 + \beta_1 x_{i1}))^2}{2\sigma^2}\right\}$$

increasing function

$$= \left(\frac{1}{2\pi}\right)^{n/2} (\sigma^2)^{-n/2} \exp\left\{-\frac{\sum_{i=1}^{n}(y_i - (\beta_0 + \beta_1 x_{i1}))^2}{2\sigma^2}\right\}$$

$\boxed{\log}$ - likelihood

$$= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 \cancel{\exp\left\{-\sum_{i=1}^{n}(y_i\right.}$$

$$- \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - (\beta_0 + \beta_1 x_{i1}))^2$$

$$\frac{\partial \log L}{\partial \beta_0} = -\frac{1}{2\sigma^2} * 2 \boxed{\sum_{i=1}^{n}(y_i - (\beta_0 + \beta_1 x_{i1}))(-1) = 0}$$

$$\frac{\partial \log L}{\partial \beta_1} = -\frac{1}{2\sigma^2} * 2 \boxed{\sum_{i=1}^{n}(y_i - (\beta_0 + \beta_1 x_{i1}))(-x_{i1}) = 0}$$

$\Leftarrow$ from method of least squares $\Downarrow$ $\hat{\beta}_0, \hat{\beta}_1$

$$\frac{\partial \log L}{\partial \sigma^2} = -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(y_i - (\beta_0 + \beta_1 x_{i1}))^2 = 0$$

$$\tilde{\sigma}^2 = \frac{\sum_{i=1}^{n}(y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1}))^2}{n} \Leftarrow \frac{\sum_{i=1}^{n}\hat{e}_i^2}{n} = \frac{\text{Res S.S.}}{n}$$

Is $\tilde{\sigma}^2$ unbiased ? $\boxed{\text{NO}}$

$$\text{Res S.S.} = \sum_{i=1}^{n}\hat{e}_i^2 \qquad \hat{e}_i = y_i - \hat{y}_i$$

$$= \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

$$= (\underset{\sim}{Y} - \hat{\underset{\sim}{Y}})^T(\underset{\sim}{Y} - \hat{\underset{\sim}{Y}}) \qquad \hat{\underset{\sim}{Y}} = \underset{\sim}{X}\hat{\underset{\sim}{\beta}}$$

$$= ((\underset{\sim}{I} - \underset{\sim}{H})\underset{\sim}{Y})^T((\underset{\sim}{I} - \underset{\sim}{H})\underset{\sim}{Y}) \qquad = \underbrace{\boxed{\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T}}_{= \underset{\sim}{H}}\underset{\sim}{Y}$$

$$= \underset{\sim}{Y}^T\boxed{(\underset{\sim}{I} - \underset{\sim}{H})^T(\underset{\sim}{I} - \underset{\sim}{H})}\underset{\sim}{Y}$$

$$\hat{\underset{\sim}{Y}} = \underset{\sim}{H}\,\underset{\sim}{Y} \qquad \underset{\sim}{H} \overset{\text{def}}{=} \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T$$

$\uparrow$ hat matrix

Symmetric ? $\boxed{\text{Yes}}$

$$\underset{\sim}{H}^T = (\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)^T$$

$$= (\underset{\sim}{X}^T)^T((\underset{\sim}{X}^T\underset{\sim}{X})^{-1})^T\underset{\sim}{X}^T$$

+

$$= \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T$$

$$= \underset{\sim}{H}$$

Is $\underset{\sim}{H}\,\underset{\sim}{H}^T = \underset{\sim}{H}$? $\boxed{\text{Yes}}$

$$\underset{\sim}{H}\,\underset{\sim}{H}^T = (\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\,(\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)^T$$

$$= \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\boxed{\underset{\sim}{X}^T\underset{\sim}{X}}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T$$

$$= \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T$$

$$= \underset{\sim}{H} \quad \Rightarrow \boxed{\text{Idempotent}}$$

Is $(\underset{\sim}{I} - \underset{\sim}{H})$ idempotent ?

$$\text{Res S.S.} = \underset{\sim}{Y}^T(\underset{\sim}{I} - \underset{\sim}{H})^T(\underset{\sim}{I} - \underset{\sim}{H})\underset{\sim}{Y}$$

$$= \underset{\sim}{Y}^T(\underset{\sim}{I} - \underset{\sim}{H})\underset{\sim}{Y}$$

$$= \underset{\sim}{Y}^T(\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T)\underset{\sim}{Y}$$

$$= \underset{\sim}{Y}^T\underset{\sim}{Y} - \underbrace{\boxed{\underset{\sim}{Y}^T\underset{\sim}{X}(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}}}_{=\,\hat{\underset{\sim}{\beta}}^T}\underset{\sim}{X}^T\underset{\sim}{Y} \qquad \hat{\underset{\sim}{\beta}} = (\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{X}^T\underset{\sim}{Y}$$

$$= \underset{\sim}{Y}^T\underset{\sim}{Y} - \hat{\underset{\sim}{\beta}}^T\underset{\sim}{X}^T\underset{\sim}{Y}$$

① Assuming β_0 is unknown

$$\text{Res S.S.} = \sum_{i=1}^n y_i^2 - (\hat\beta_0, \hat\beta_1, \cdots, \hat\beta_p)\begin{pmatrix}\sum_{i=1}^n y_i \\ \sum_{i=1}^n x_{i1} y_i \\ \vdots \\ \sum_{i=1}^n x_{ip} y_i\end{pmatrix}$$

$$= \sum_{i=1}^n y_i^2 - \boxed{\hat\beta_0} \sum_{i=1}^n y_i - \hat\beta_1 \sum_{i=1}^n x_{i1} y_i - \cdots - \hat\beta_p \sum_{i=1}^n x_{ip} y_i$$

Note that $\hat\beta_0 = \bar y - \hat\beta_1 \bar x_1 - \cdots - \hat\beta_p \bar x_p$

$$\text{Res S.S.} = \sum_{i=1}^n y_i^2 - (\bar y - \hat\beta_1 \bar x_1 - \cdots - \hat\beta_p \bar x_p) \sum_{i=1}^n y_i \quad \leftarrow n\bar y$$

$$- \hat\beta_1 \sum_{i=1}^n x_{i1} y_i - \cdots - \hat\beta_p \sum_{i=1}^n x_{ip} y_i$$

$$= \boxed{\left(\sum_{i=1}^n y_i^2 - n\bar y^2\right)} - \hat\beta_1 \boxed{\left(\sum_{i=1}^n x_{i1} y_i - n \bar x_1 \bar y\right)}$$

$$- \hat\beta_2 \left(\sum_{i=1}^n x_{i2} y_i - n \bar x_2 \bar y\right) - \cdots - \hat\beta_p \left(\sum_{i=1}^n x_{ip} y_i - n \bar x_p \bar y\right)$$

$$= S_{yy} - \left(\hat\beta_1 S_{x_1 y} + \cdots + \hat\beta_p S_{x_p y}\right)$$

$$\boxed{S_{uv} = \sum_{i=1}^n (u_i - \bar u)(v_i - \bar v) = \sum_{i=1}^n u_i v_i - n \bar u \bar v = \sum_{i=1}^n (u_i - \bar u) v_i}$$

$$\text{Res S.S.} = \boxed{S_{yy}} - \boxed{\left(\hat\beta_1 S_{x_1 y} + \cdots + \hat\beta_p S_{x_p y}\right)}$$

S_{yy} = corrected total S.S. $= \sum_{i=1}^n (y_i - \bar y)^2$

$\left(\hat\beta_1 S_{x_1 y} + \cdots + \hat\beta_p S_{x_p y}\right)$ = Reg S.S

$\beta_1, \cdots, \beta_p$ reg. coeff of $x_1, \cdots, x_p$, respectively

$\underline{p=1}$ $\quad \text{Res S.S.} = S_{yy} - \hat{\beta}_1 S_{x_1 y}$ $\qquad \hat{\beta}_1 = \dfrac{S_{x_1 y}}{S_{x_1 x_1}}$

$$\Rightarrow S_{x_1 y} = \boxed{\hat{\beta}_1 S_{x_1 x_1}}$$

$$= S_{yy} - \hat{\beta}_1 (\hat{\beta}_1 S_{x_1 x_1})$$

$$= S_{yy} - \hat{\beta}_1^2 S_{x_1 x_1}$$

$$\boxed{\text{Res S.S.} = S_{yy} - (\hat{\beta}_1^2 S_{x_1 x_1} + \cdots + \hat{\beta}_p^2 S_{x_p x_p})} \quad \times$$

$$E(\text{Res S.S.}) = E(S_{yy}) - S_{x_1 x_1} E(\hat{\beta}_1^2)$$

$$E(\hat{\beta}_1^2) = Var(\hat{\beta}_1) + E(\hat{\beta}_1)^2 *$$

$$E(S_{yy}) = E\left(\sum_{i=1}^n (y_i - \bar{y})^2\right) \neq (n-1)\sigma^2$$

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - n\bar{y}^2$$

diff. means

Then, $E(S_{yy}) = ?$

(2) $\underline{\beta_0 \text{ is known}}$

$$\underset{\sim}{Y} = \begin{pmatrix} y_1' \\ \vdots \\ y_n' \end{pmatrix} = \begin{pmatrix} y_1 - \beta_0 \\ \vdots \\ y_n - \beta_0 \end{pmatrix}$$

$$\text{Res S.S.} = \underset{\sim}{Y}^T \underset{\sim}{Y} - \hat{\underset{\sim}{\beta}}^T X^T \underset{\sim}{Y}$$

$$\underset{\sim}{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}$$

$$= \sum_{i=1}^n y_i'^2 - \hat{\beta}_1 \sum_{i=1}^n x_{i1} y_i' - \cdots - \hat{\beta}_p \sum_{i=1}^n x_{ip} y_i'$$

$$\underset{p\times n}{X^T} \underset{n\times 1}{\underset{\sim}{Y}} = \begin{pmatrix} \sum_{i=1}^n x_{i1} y_i' \\ \vdots \\ \sum_{i=1}^n x_{ip} y_i' \end{pmatrix}$$

$$= \sum_{i=1}^n (y_i - \beta_0)^2 - \hat{\beta}_1 \sum_{i=1}^n x_{i1}(y_i - \beta_0) * - \cdots - \hat{\beta}_p \sum_{i=1}^n x_{ip}(y_i - \beta_0)$$

$$= \sum_{i=1}^n y_i^2 - 2\boxed{\beta_0}\sum_{i=1}^n y_i + \boxed{\beta_0^2} - \hat{\beta}_1 \left(\sum_{i=1}^n x_{i1} y_i - \boxed{\beta_0}\sum_{i=1}^n x_{i1}\right) - \cdots - \hat{\beta}_p \left(\sum_{i=1}^n x_{ip} y_i - \boxed{\beta_0}\sum_{i=1}^n x_{ip}\right)$$

If $\beta_0 = 0$, $\text{Res S.S.} = \sum_{i=1}^n y_i^2 - \hat{\beta}_1 \sum_{i=1}^n x_{i1} y_i - \cdots - \hat{\beta}_p \sum_{i=1}^n x_{ip} y_i$

Example in p.7

Assume $\beta_0, \beta_1, \beta_2, \beta_3$ are unknown

$$\text{Res S.S.} = S_{yy} - \underset{\substack{\uparrow \\ 1.0161}}{\hat{\beta}_1} S_{x_1y} - \underset{\substack{\uparrow \\ -1.8616}}{\hat{\beta}_2} S_{x_2y} - \underset{\substack{\uparrow \\ -0.3433}}{\hat{\beta}_3} S_{x_3y}$$

$$= 38.68289$$

Assume $\beta_0 = 40$ $\qquad \hat{\beta}_1 = 1.00748, \quad \hat{\beta}_2 = -1.871149, \quad \hat{\beta}_3 = -0.425863$

$$\text{Res S.S.} = \sum_{i=1}^{n} (y_i - \beta_0)^2 - \hat{\beta}_1 \sum_{i=1}^{n} x_{i1}(y_i - \beta_0) - \hat{\beta}_2 \sum_{i=1}^{n} x_{i3}(y_i - \beta_0)$$

$$- \hat{\beta}_3 \sum_{i=1}^{n} x_{i3}(y_i - \beta_0)$$

$$= \sum_{i=1}^{n} y_i^2 - 2\beta_0 \sum_{i=1}^{n} y_i + \beta_0^2 - \hat{\beta}_1 \left(\sum_{i=1}^{n} x_{i1} y_i - \beta_0 \sum_{i=1}^{n} x_{i1} \right)$$

$$- \hat{\beta}_2 \left(\sum_{i=1}^{n} x_{i2} y_i - \beta_0 \sum_{i=1}^{n} x_{i2} \right) - \hat{\beta}_3 \left(\sum_{i=1}^{n} x_{i3} y_i - \beta_0 \sum_{i=1}^{n} x_{i3} \right)$$

$$= 38.76445$$

Assume $\beta_1 = 1$ $\qquad \hat{\beta}_0 = 39.3147, \quad \hat{\beta}_2 = -1.86491, \quad \hat{\beta}_3 = -0.35032$

$\uparrow$ model with intercept

$$\text{Res S.S.} = S_{y'y'} - \hat{\beta}_2 \boxed{S_{x_2y'}} - \hat{\beta}_3 S_{x_3y'}$$

$$S_{x_2y'} = \sum_{i=1}^{n} (x_{i2} - \bar{x}_2)(y'_i - \bar{y}') = \sum_{i=1}^{n} (x_{i2} - \bar{x}_2)\left[(y_i - \bar{y}) - (x_{i1} - \bar{x}_1)\right] = S_{x_2y} - S_{x_1x_2}$$

$$S_{y'y'} = \sum_{i=1}^{n} (y'_i - \bar{y}')^2$$

$$= \sum_{i=1}^{n} \left[(y_i - x_{i1}) - (\bar{y} - \bar{x}_1)\right]^2$$

$$= \sum_{i=1}^{n} \left[(y_i - \bar{y}) - (x_{i1} - \bar{x}_1)\right]^2$$

$$= S_{yy} - 2S_{x_1y} + S_{x_1x_1}$$

$$\text{Res S.S.} = 38.70697$$