Assignment - 2

① (a) $\hat{w}_1 \longrightarrow$ slope estimate   given $n = 250$

$\hat{w}_0 \longrightarrow$ intercept "

According to simple linear regression model, using least squares estimators, we know

$$\hat{w}_0 = \bar{y} - \hat{w}_1 \bar{x}$$

$$\hat{w}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n \bar{x} \bar{y}}{\sum_{i=1}^{n} x_i^2 - \frac{(n\bar{x})^2}{n}} = \frac{s_{xy}}{s_{xx}}$$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{11211}{250} = 44.844$$

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{44520.80}{250} = 178.0832$$

$$s_{xy} = 1996904.15 - 250 \times 44.844 \times 178.0832$$

$$= 413.3948$$

$$s_{xx} = 543503 - 250 \times (44.844)^2 = 40756.916$$

$$\hat{w}_1 = \frac{413.3948}{40756.916} = \boxed{0.010143}$$

$$\hat{w}_0 = 178.0832 - 0.010143 \times 44.844 = \boxed{177.62835}$$

(b) The regression line has an equation $\hat{y} = \hat{w}_0 + \hat{w}_1 x$

The predicted value of weight of a 25 year old man is $\hat{y} = 177.62835 + 0.010143 \times 25$ lbs

$$= \boxed{177.881925 \text{ lbs}}$$

(c) Residual for the observation is

$$y - \hat{y} = 170 - 177.881925 \text{ lbs}$$

$$= \boxed{-7.881925 \text{ lbs}}$$

(d) It is an $\boxed{\text{overestimate}}$, because the observed value is less than the predicted one.

(2) given $n = 14$

(a) $$\bar{x} = \frac{\Sigma x_i}{n} = \frac{43}{14} = 3.0714$$

$$\bar{y} = \frac{\Sigma y_i}{n} = \frac{572}{14} = 40.85714$$

$$s_{xy} = \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y} = 1697.8 - 14 \times 3.0714 \times 40.85714$$

$$= -59.040677$$

$$s_{xx} = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2 = 157.42 - 14 \times (3.0714)^2$$

$$= 25.351029$$

$$\hat{w}_1 = \frac{s_{xy}}{s_{xx}} = \frac{-59.040677}{25.351029} = \boxed{-2.3289}$$

$$\hat{w}_0 = \bar{y} - \hat{w}_1 \bar{x} = 40.85714 + 2.3289 \times 3.0714$$

$$= \boxed{48.01012}$$

(b) When $x = 4.3$, y predicted $= \hat{y} = \hat{w}_0 + \hat{w}_1 x$

$$= 48.01012 - 2.3289 \times 4.3$$

$$= \boxed{37.99585}$$

(c) Mean permeability at $x = 3.7$, $E(y \mid x{=}3.7) = \hat{w}_0 + \hat{w}_1 x$

$$= \boxed{39.39319}$$

(d) e_i = Residual = $y_i - \hat{y}_i$

$= 46.1 - 39.39319$

$= \boxed{6.70681}$

(3) $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$

Least sq. error func^n $g(\beta) = \dfrac{\sum_{i=1}^{n} (\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} - y_i)^2}{n}$

$= \frac{1}{n} \sum_{i=1}^{n} (\beta x_i^T - y_i)^2$

(a) Now least squares normal eq^{n}s :-

(i) $\left.\dfrac{\partial g(\beta)}{\partial \beta_0}\right|_{\hat{\beta}_1, \hat{\beta}_2} = 0 \Rightarrow \frac{2}{n} \sum_{i=1}^{n} (\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} - y_i) = 0$

$\Rightarrow n\hat{\beta}_0 + \hat{\beta}_1 \Sigma x_{1i} + \hat{\beta}_2 \Sigma x_{2i} - \Sigma y_i = 0$

$\Rightarrow \boxed{\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}_{1i} - \hat{\beta}_2 \bar{x}_{2i}}$

(ii) $\left.\dfrac{\partial g(\beta)}{\partial \beta_1}\right|_{\hat{\beta}_1, \hat{\beta}_2} = 0 \Rightarrow \frac{2}{n} \Sigma (\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} - y_i) x_{1i} = 0$

$\Rightarrow \hat{\beta}_0 \Sigma x_{1i} + \hat{\beta}_1 \Sigma (x_{1i})^2 + \hat{\beta}_2 \Sigma x_{1i} x_{2i} - \Sigma y_i x_i = 0$

$\Rightarrow \bar{y} \Sigma x_{1i} - \hat{\beta}_2 \bar{x}_{2i} \Sigma x_{1i} - \hat{\beta}_1 \bar{x}_{1i} \Sigma x_{1i} + \hat{\beta}_1 \Sigma (x_{1i})^2 + \hat{\beta}_2 \Sigma x_{1i} x_{2i} - \Sigma y_i x_i = 0$

(iii) $\left.\dfrac{\partial g(\beta)}{\partial \beta_2}\right|_{\hat{\beta}_1, \hat{\beta}_0} = 0 \Rightarrow \frac{2}{n} \Sigma (\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} - y_i) x_{2i} = 0$

$$\Rightarrow \hat\beta_0 \sum x_{2i} + \hat\beta_1 \sum x_{1i} x_{2i} + \hat\beta_2 (\sum x_{2i})^2 - \sum y_i x_{2i} = 0$$

$$\Rightarrow \bar y \sum x_{2i} - \hat\beta_2 \bar x_{2i} \sum x_{2i} - \hat\beta_1 \bar x_{1i} \sum x_{2i} + \hat\beta_1 \sum x_{1i} x_{2i} + \hat\beta_2 \sum (x_{2i})^2 - \sum y_i x_{2i} = 0$$

(b) From (ii),

$$\hat\beta_1 \underbrace{\left(\sum x_{1i}^2 - \bar x_{1i} \sum x_{1i}\right)}_{a} + \hat\beta_2 \underbrace{\left(\sum x_{1i} x_{2i} - \bar x_{2i} \sum x_{1i}\right)}_{b} = \underbrace{\sum y_i x_{1i} - \bar y \sum x_{1i}}_{e}$$

From (iii),

$$\hat\beta_1 \underbrace{\left(\sum x_{1i} x_{2i} - \sum x_{2i} \bar x_{1i}\right)}_{b} + \hat\beta_2 \underbrace{\left(\sum x_{2i}^2 - \bar x_{2i} \sum x_{2i}\right)}_{c} = \underbrace{\sum y_i x_{2i} - \bar y \sum x_{2i}}_{f}$$

$$(a\hat\beta_1 + b\hat\beta_2 = e) \times c$$

$$(b\hat\beta_1 + c\hat\beta_2 = f) \times (-b)$$

$$(ac - b^2)\hat\beta_1 = ce - bf$$

$$\boxed{\hat\beta_1 = \frac{ce - bf}{ac - b^2}}$$

$$a\left(\frac{ce - bf}{ac - b^2}\right) + b\hat\beta_2 = e$$

$$\boxed{\hat\beta_2 = \frac{af - be}{ac - b^2}}$$

$$\boxed{\hat\beta_0 = \bar y - \left(\frac{af - be}{ac - b^2}\right)\bar x_{2i} - \left(\frac{ce - bf}{ac - b^2}\right)\bar x_{1i}}$$

(c) $a = \sum x_{1i}^2 - \bar x_{1i} \sum x_{1i}$

$$= \sum x_{1i}^2 - \frac{(\sum x_{1i})^2}{n} \qquad [n = 10]$$

$$\boxed{a = 228}$$

$$b = \sum x_{1i} \cdot x_{2i} - \frac{\sum x_{2i} \sum x_{1i}}{n}$$

$$\boxed{b = 20.1}$$

$c = \sum x_{2i}^2 - \bar{x}_{2i} \sum x_{2i} = \sum x_{2i}^2 - \frac{(\sum x_{2i})^2}{n}$

$\boxed{c = 1148.1}$

$\hat{\beta}_1 = \frac{ce - bf}{ac - b^2} = 3.7133$

$e = \sum y_i x_i - \bar{y} \sum x_{1i}$

$\boxed{e = 824}$

$\hat{\beta}_2 = \frac{e - a\hat{\beta}_1}{b} = -1.1259$

$f = \sum y_i x_{2i} - \bar{y} \sum x_{2i}$

$\boxed{f = -1218}$

$\hat{\beta}_0 = \bar{y} - \hat{\beta}_2 \bar{x}_{2i} - \hat{\beta}_1 \bar{x}_{1i}$

$= 171.0557$

y predicted $= \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$

$= \boxed{189.4814}$

(4) $y = \beta_0 + \beta_1 \underset{\text{height}}{x_1} + \beta_2 \underset{\text{waist}}{x_2} + \epsilon$ given $n = 250$

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix}$$

$X'X\hat{\beta} = X'Y$

$\hat{\beta} = (X'X)^{-1}(X'Y)$

$$\hat{\beta} = \begin{bmatrix} 2.9705 & -4.0042 \times 10^{-2} & -4.1679 \times 10^{-2} \\ -0.4004 & 6.0774 \times 10^{-4} & -7.3875 \times 10^{-5} \\ -0.00417 & -7.3875 \times 10^{-5} & 2.5766 \times 10^{-4} \end{bmatrix} \begin{bmatrix} 4757.9 \\ 334335.8 \\ 179706.7 \end{bmatrix}$$

$$\hat{\beta} = \begin{bmatrix} -6744.1277 \\ -1715.14975 \\ 1.7637281 \end{bmatrix}$$

Model eq^n, $Y = -6744.1277 - 1715.14975\, x_1 + 1.7637281\, x_2$

(5) Cost func^n, $G(W) = \sum \left(y_i - \left(W_0 + W_1 x_{1i} + W_2 x_{2i} + W_3 x_{1i} x_{2i} + W_4 x_{1i}^2 + W_5 x_{2i}^2\right)\right)^2$

$$\frac{\partial G}{\partial W_0} = 0 \Rightarrow \sum y_i = W_0 \sum 1 + W_1 \sum x_{1i} + W_2 \sum x_{2i} + W_3 \sum x_{1i} x_{2i} + W_4 \sum x_{1i}^2 + W_5 \sum x_{2i}^2$$

$$\frac{\partial G}{\partial W_1} = 0 \Rightarrow \sum y_i x_{1i} = W_0 \sum x_{1i} + W_1 \sum x_{1i}^2 + W_2 \sum x_{2i} + W_3 \sum x_{1i} x_{2i} + W_4 \sum x_{1i}^2 + W_5 \sum x_{2i}^2$$

$$\frac{\partial G}{\partial W_2} = 0 \Rightarrow \sum y_i x_{2i} = W_0 \sum x_{2i} + W_1 \sum x_{1i} x_{2i} + W_2 \sum x_{2i}^2 + W_3 \sum x_{2i}^2 x_{1i} + W_4 \sum x_{1i}^2 x_{2i} + W_5 \sum x_{2i}^3$$

$$\frac{\partial G}{\partial W_3} = 0 \Rightarrow \sum y_i x_{1i} x_{2i} = W_0 \sum x_{1i} x_{2i} + W_1 \sum x_{1i}^2 x_{2i} + W_2 \sum x_{1i} x_{2i}^2 + W_3 \sum x_{1i}^2 x_{2i}^2 + W_4 \sum x_{1i}^3 x_{2i} + W_5 \sum x_{1i} x_{2i}^3$$

$$\frac{\partial G}{\partial W_4} = 0 \Rightarrow \sum y_i x_{1i}^2 = W_0 \sum x_{1i}^2 + W_1 \sum x_{1i}^3 + W_2 \sum x_{2i} x_{1i}^2 + W_3 \sum x_{1i}^3 x_{2i} + W_4 \sum x_{1i}^4 + W_5 \sum x_{1i}^2 x_{2i}^2$$

$$\frac{\partial G}{\partial W_5} = 0 \Rightarrow \sum y_i x_{2i}^2 = W_0 \sum x_{2i}^2 + W_1 \sum x_{1i} x_{2i}^2 + W_2 \sum x_{2i}^3 + W_3 \sum x_{1i} x_{2i}^3 + W_4 \sum x_{1i}^2 x_{2i}^2 + W_5 \sum x_{2i}^4$$