1.

$$\beta = (X'X)^{-1}(X'Y) \qquad Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & X_1 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix}$$

$$(X'X) = \begin{bmatrix} 1 & X_1 \\ \vdots & \vdots \\ 1 & X_N \end{bmatrix}^{-1} \begin{bmatrix} 1 & X_1 \\ \vdots & \vdots \\ 1 & X_N \end{bmatrix} = \begin{bmatrix} N & \sum_{i=1}^{N} X_i \\ \sum_{i=1}^{N} X_i & \sum_{i=1}^{N} X_i^2 \end{bmatrix}$$

$$(X'X)^{-1} = \frac{1}{N\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2} \begin{bmatrix} \sum_{i=1}^{n} X_i^2 & -\sum_{i=1}^{n} X_i \\ -\sum_{i=1}^{n} X_i & N \end{bmatrix}$$

$$(X'Y) = \begin{bmatrix} 1 & X_1 \\ \vdots & \vdots \\ 1 & X_N \end{bmatrix}^{-1} \begin{bmatrix} Y_1 \\ \vdots \\ Y_N \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} X_i Y_i \end{bmatrix}$$

$$\hat{\beta} = (X'X)^{-1}(X'Y) = \frac{1}{N\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2} \begin{bmatrix} \sum_{i=1}^{n} X_i^2 & -\sum_{i=1}^{n} X_i \\ -\sum_{i=1}^{n} X_i & N \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} X_i Y_i \end{bmatrix}$$

$$= \frac{1}{N\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2} \begin{bmatrix} \sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} X_i Y_i \\ N\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i \end{bmatrix}$$

$$\hat{\beta}_1 = \frac{N\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{N\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2} = \frac{\sum_{i=1}^{n} X_i Y_i - n\bar{X}\bar{Y}}{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2} = \underline{\frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}} \quad 2\text{-}7$$

$$\hat{\beta}_0 = \frac{\sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} X_i Y_i}{N\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2} = \frac{\sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} Y_i \left(\sum_{i=1}^{n} X_i\right)^2/n + \sum_{i=1}^{n} Y_i \left(\sum_{i=1}^{n} X_i\right)^2/n - \sum_{i=1}^{n} X_i Y_i \sum_{i=1}^{n} X_i}{N\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2}$$

$$= \frac{\overbrace{\sum_{i=1}^{n} Y_i}^{n\bar{Y}} \left[\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2/n\right]}{N\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2} - \frac{\overbrace{\sum_{i=1}^{n} X_i}^{n\bar{X}} \left[\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i/n\right)}{N\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2} = \underline{\bar{Y} - \bar{X}\hat{\beta}_1} \quad 2\text{-}8$$

2.

$$Var(\hat{\beta}) = \sigma^2 (X'X)^{-1}$$

$$\sigma^2 (X'X)^{-1} = \frac{\sigma^2}{N\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \begin{bmatrix} \sum_{i=1}^{n} x_i^2 & -\sum_{i=1}^{N} x_i \\ -\sum_{i=1}^{n} x_i & N \end{bmatrix}$$

$$N\left(\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right)$$

$$= \begin{bmatrix} \dfrac{\sigma^2 \sum_{i=1}^{n} x_i^2}{N\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} & \dfrac{-\sigma^2 \sum_{i=1}^{N} x_i}{N\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \\ \dfrac{-\sigma^2 \sum_{i=1}^{n} x_i}{N\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} & \dfrac{N\sigma^2}{N\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \end{bmatrix}$$

$$Var(\hat{\beta}_1 | X) = \sigma^2 \left[\frac{\sum_{i=1}^{n} x_i^2}{N \sum_{i=1}^{n} (x_i - \bar{x})^2} \right]$$ 2-14

$$Var(\hat{\beta}_0 | X) = \sigma^2 \left[\frac{1}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \right]$$ 2-15

$$Cov(\hat{\beta}_0, \hat{\beta}_1 | X) = \sigma^2 \left[\frac{-\bar{x}}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \right]$$ 2-16

5.3.

The linear regression is:

$$WALC = \beta_1 + \beta_2 \ln(TOTEXP) + \beta_3 NK + \beta_4 AGE + e$$

a

i the t-statistic for β_1: $\dfrac{1.4515}{2.2019} = 0.6592034$

ii the standard error for β_2: $\dfrac{2.7648}{5.7103} = 0.4841777$

iii The estimate β_3: $0.3695 \times -3.9376 = -1.454943$

iv

$$R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{46221.62}{49041.5418} = 0.05750068.$$

$$SST = 6.39547^2 \times 1199 = 49{,}041.5418$$

v. $\hat{\sigma}^2 = \dfrac{SSE}{1200-4} = \dfrac{46221.62}{1196} = 38.64684$ $\quad \hat{\sigma} = 6.216658$

b.

Ch 4. p44

β_2: Holding other variable unchange, when total expenditure change 1%. the budget spent on alcohol went up to 0.027648 percentage point, on average,

β_3: Holding other variable unchange, when the number of children increase 1, the budget spent on alcohol went down to 1.454943 percentage point.

β_4: Holding other variable unchange, when the age go up by 1 age, the bedget spent on alcohol went down to 0.1503 percentage point.

c.

$\hat{\beta}_4$ 95% C.I. $[-0.1503 \pm 1.96 \times 0.0235] = [-0.10424, -0.19636]$

This interval don't contain 0, so we could reject Ho, that β_4 is not 0.

d.

The p-value are all small than 0.05, it means we can reject Ho, the coefficient are all different from 0, they are siginificant relative than y.

e.

1. Ho: $\beta_3 = 2$ against $H_1: \beta_3 \neq 2$.

2. t-statistic $= \dfrac{-1.454943 + 2}{0.3695} = 1.47512$

3. reject region: $\{ t > 1.96$ or $t < -1.96 \}$

We fail to reject Ho, there is no prove that the addition of an extra children decrease 2 percentage point to the budget of alcohol.