5.3 Consider the following model that relates the percentage of a household's budget spent on alcohol *WALC* to total expenditure *TOTEXP*, age of the household head *AGE*, and the number of children in the household *NK*.

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

This model was estimated using 1200 observations from London. An incomplete version of this output is provided in Table 5.6.

TABLE 5.6 **Output for Exercise 5.3**

Dependent Variable: *WALC*
Included observations: 1200

Variable	Coefficient	Std. Error	*t*-Statistic	Prob.
C	1.4515	2.2019		0.5099
ln(*TOTEXP*)	2.7648		5.7103	0.0000
NK		0.3695	−3.9376	0.0001
AGE	−0.1503	0.0235	−6.4019	0.0000

R-squared		Mean dependent var	6.19434
S.E. of regression		S.D. dependent var	6.39547
Sum squared resid	46221.62		

a. Fill in the following blank spaces that appear in this table.
  i. The *t*-statistic for b_1.
  ii. The standard error for b_2.
  iii. The estimate b_3.
  iv. R^2.
  v. $\hat{\sigma}$.

b. Interpret each of the estimates b_2, b_3, and b_4.
c. Compute a 95% interval estimate for β_4. What does this interval tell you?
d. Are each of the coefficient estimates significant at a 5% level? Why?
e. Test the hypothesis that the addition of an extra child decreases the mean budget share of alcohol by 2 percentage points against the alternative that the decrease is not equal to 2 percentage points. Use a 5% significance level.

(a). (i) $t = \frac{1.4515}{2.2019} = 0.6592.$

(ii) Se: $\frac{2.7648}{5.7103} = 0.4842$

(iii) $b_3 = 0.3695 \cdot -3.9376 = -1.4549$

(iv.) $R^2 = 1 - \frac{\sum \hat{e}_i^2}{\sum (y_i - \bar{y})^2}$

$= 1 - \frac{\sum \hat{e}_i^2}{(n-1)\,var(y_i)} = 0.0575.$

(v.) $\hat{\sigma} = \frac{\sum \hat{e}_i^2}{1200-4} = 6.2167.$

(b) Holding other constants:
Increase +e in TOTEXP, increase in 2.7648 in WALC
Increase +1 in NK, decrease in 1.4549 in WALC
Increase +1 in AGE, decrease in 0.1503 in WALC.

(c) 95% interval: $b_4 \pm 1.993 \times 0.0235$: $[-0.1971, -1.035]$
⇒ 95% of this interval will cover the true parameter.

(d) No, we see that the p-value for β_1 is $0.5099 > 0.05$.

(e) $H_0: \beta_3 = -0.1239$. $(6.19434 \times 2\%)$, $H_a: \beta_3 \neq -0.1239$.

$\Rightarrow t = \frac{b_3 - (-0.1239)}{0.3695} = -3.6022 < -t_{0.975,\infty} = -1.993$

⇒ We reject H_0 ⇒ $\beta_3 \neq -0.1239$.

- $K=2$, $y = X\beta + \varepsilon$, $y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, $X = \begin{bmatrix} 1 & X_{12} \\ 1 & X_{22} \\ \vdots & \vdots \\ 1 & X_{n2} \end{bmatrix}$.

$\varepsilon = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$, $(X'X)^{-1} = \begin{bmatrix} n & n\bar{x} \\ n\bar{x} & \Sigma x^2 \end{bmatrix}^{-1}$

$$= \frac{1}{n\Sigma x^2 - (n\bar{x})^2} \begin{bmatrix} \Sigma x^2 & -n\bar{x} \\ -n\bar{x} & n \end{bmatrix}$$

$$b = (X'X)^{-1}X'y = \frac{1}{n\Sigma x^2 - (n\bar{x})^2} \begin{bmatrix} \Sigma x^2 & -n\bar{x} \\ -n\bar{x} & n \end{bmatrix} \begin{bmatrix} n\bar{y} \\ \Sigma xy \end{bmatrix}$$

$$= \frac{1}{n\Sigma x^2 - (n\bar{x})^2} \begin{bmatrix} n\bar{y}\Sigma x^2 - n\bar{x}\Sigma xy \\ -n^2\bar{x}\bar{y} + n\Sigma xy \end{bmatrix}. \quad \ldots (*)$$

$$\Sigma(x_i - \bar{x})^2 = \Sigma(x_i^2 - 2\bar{x}x_i + \bar{x}^2) = \Sigma(x_i^2 - \bar{x})$$

$$= \Sigma x^2 - n\bar{x}, \quad \Sigma(x_i - \bar{x})(y_i - \bar{y}) = \Sigma(xy - \bar{y}x - \bar{x}y + \bar{x}\bar{y})$$

$$= \Sigma(xy - \bar{x}\bar{y}) = \Sigma xy - n\bar{x}\bar{y},$$

$$b_2 = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})^2} = \frac{n\Sigma xy - n^2\bar{x}\bar{y}}{n\Sigma x^2 - (n\bar{x})^2} = \frac{\Sigma xy - n\bar{x}\bar{y}}{\Sigma x^2 - n\bar{x}^2} \quad \#$$

$$b_1 = \frac{\cancel{n}\bar{y}\Sigma x^2 - \cancel{n}\bar{x}\Sigma xy}{\cancel{n}\Sigma x^2 - n^{\cancel{2}}\bar{x}^2} = \frac{\bar{y}(\Sigma x^2 - n\bar{x}^2) + \bar{y}n\bar{x}^2 - \bar{x}\Sigma xy}{\Sigma x^2 - n\bar{x}^2}$$

$$= \bar{y} - \bar{x}\left(\frac{\Sigma xy - n\bar{x}\bar{y}}{\Sigma x^2 - n\bar{x}^2}\right) = \bar{y} - b_2\bar{x} \quad \#$$

$\cdot\ K=2,\ var(b) = \sigma^2 (X'X)^{-1}$

$$= \frac{\sigma^2}{n\sum x^2 - n^2\bar{x}^2} \begin{bmatrix} \sum x^2 & -n\bar{x} \\ -n\bar{x} & n \end{bmatrix}$$

$$= \sigma^2 \begin{bmatrix} \frac{\sum x^2}{n\sum x^2 - n^2\bar{x}^2} & -\frac{n\bar{x}}{n\sum x^2 - n^2\bar{x}^2} \\ -\frac{n\bar{x}}{n\sum x^2 - n^2\bar{x}^2} & \frac{n}{n\sum x^2 - n^2\bar{x}^2} \end{bmatrix}$$

$$= \sigma^2 \begin{bmatrix} \frac{\sum x^2}{n\sum (X_i - \bar{X})^2} & -\frac{\bar{X}}{\sum (X_i - \bar{X})^2} \\ -\frac{\bar{x}}{\sum (X_i - \bar{X})^2} & \frac{1}{\sum (X_i - \bar{X})^2} \end{bmatrix}$$

$$\Rightarrow var(b_1|X) = \sigma^2 \cdot \frac{\sum X^2}{n\sum (X_i - \bar{X})^2},\ var(b_2|X) = \sigma^2 \frac{1}{\sum (X_i - \bar{X})^2}$$

$$\Rightarrow cov(b_1, b_2|X) = \sigma^2 \cdot \frac{-\bar{X}}{\sum (X_i - \bar{X})^2}$$ #