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) $t = \dfrac{b_1}{se(b_1)} = \dfrac{1.4515}{2.2019} = 0.6592$

(ii) $t = \dfrac{2.7648}{se(b_2)} = 5.7103$

$\Rightarrow se(b_2) = \dfrac{2.7648}{5.7103} = 0.4842$

(iii) $t = \dfrac{b_3}{0.3695} = -3.9376$

$\Rightarrow b_3 = 0.3695 \times -3.9376 = -1.4549$

(iv) $SSE = 46221.62$

$SST = (N-1)\, s_y^2$

$= (1200-1)\cdot 6.39547^2$

$= 49041.54$

$R = 1 - \dfrac{SSE}{SST} = 1 - \dfrac{46221.62}{49041.54}$

$= 0.0575$

(v). $\hat{\sigma}^2 = \dfrac{\sum_{i=1}^{N} \hat{e}_i^2}{N-K} = \dfrac{46221.62}{2000-4}$

$= 38.6468$

$\hat{\sigma} = \sqrt{38.6468} = 6.2167$

b. β_2：其他變數維持不變，TOTEXP 增加 1%，WALC 增加 0.027648%（$2.7648 \div 100$）
β_3：其他變數維持不變，每增加 1 名 Child，WALC 減少 1.4549 %
β_4：其他變數維持不變，每增加一單位 AGE，WALC 減少 0.1503 %

c. $[-0.1503 \pm 1.96 \times 0.0235] = [-0.19636, -0.10424]$

d. β_1 的 p-value 大於 0.05 ⇒ 不拒絕 H_0 ⇒ b_1 沒有顯著異於 0
$\beta_2, \beta_3, \beta_4$ 的 p-value 小於 0.05 ⇒ 拒絕 H_0 ⇒ b_2, b_3, b_4 皆顯著異於 0

e. $\begin{cases} H_0: \beta_3 = -2 \\ H_1: \beta_3 \neq 2 \end{cases}$

$$t = \frac{-1.4549 - (-2)}{0.3695} = 1.4752$$

$RR: \{t \mid t > 1.96 \text{ or } t < 1.96\}$

⇒ $t < 1.96$ ⇒ 不拒絕 H_0 ⇒ 每增加 1 名 Child，WALC 增加異於 2%.

5.23 The file *cocaine* contains 56 observations on variables related to sales of cocaine powder in northeastern California over the period 1984–1991. The data are a subset of those used in the study Caulkins, J. P. and R. Padman (1993), "Quantity Discounts and Quality Premia for Illicit Drugs," *Journal of the American Statistical Association*, 88, 748–757. The variables are

$PRICE$ = price per gram in dollars for a cocaine sale
$QUANT$ = number of grams of cocaine in a given sale
$QUAL$ = quality of the cocaine expressed as percentage purity
$TREND$ = a time variable with 1984 = 1 up to 1991 = 8
Consider the regression model

$$PRICE = \beta_1 + \beta_2 QUANT + \beta_3 QUAL + \beta_4 TREND + e$$

a. What signs would you expect on the coefficients β_2, β_3, and β_4?

b. Use your computer software to estimate the equation. Report the results and interpret the coefficient estimates. Have the signs turned out as you expected?

c. What proportion of variation in cocaine price is explained jointly by variation in quantity, quality, and time?

d. It is claimed that the greater the number of sales, the higher the risk of getting caught. Thus, sellers are willing to accept a lower price if they can make sales in larger quantities. Set up H_0 and H_1 that would be appropriate to test this hypothesis. Carry out the hypothesis test.

e. Test the hypothesis that the quality of cocaine has no influence on expected price against the alternative that a premium is paid for better-quality cocaine.

f. What is the average annual change in the cocaine price? Can you suggest why price might be changing in this direction?

a. β_2 預期為負，大量購買 cocaine 會有折扣使單價較低
β_3 預期為正，cocaine 的品質越好，單價越高
β_4 取決於市場趨勢，若 cocaine 價格隨時間上升，β_3 為正，若 cocaine 價格隨時間下降，β_4 為負。

b.

```
Call:
lm(formula = price ~ quant + qual + trend, data = cocaine)

Residuals:
    Min      1Q  Median      3Q     Max
-43.479 -12.014  -3.743  13.969  43.753

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 90.84669    8.58025  10.588 1.39e-14 ***
quant       -0.05997    0.01018  -5.892 2.85e-07 ***
qual         0.11621    0.20326   0.572   0.5700
trend       -2.35458    1.38612  -1.699   0.0954 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 20.06 on 52 degrees of freedom
Multiple R-squared:  0.5097,    Adjusted R-squared:  0.4814
F-statistic: 18.02 on 3 and 52 DF,  p-value: 3.806e-08
```

$\beta_2 = -0.05997$，符合預期，表示每購買多一單位的 cocine，單價下降約 \$0.06.
$\beta_3 = 0.11621$，符合預期，表示 cacine 每增加 1 單位的 quality，單價上升 \$0.11621.
$\beta_4 = -2.35458$，當時間增加 1 年，單價減少 \$2.3546，代表 cocaine 的價格隨時間下降，表示 cocaine 的供給速度大於需求。

c. $R = 0.5097$ ⇒ 模型可以解釋 50.97% 的 cocaine 價格變異。

d. $\begin{cases} H_0: \beta_2 \geq 0 \\ H_1: \beta_2 < 0 \end{cases}$ RR: $\{t : t < -1.674689\}$

t-value $= -5.892 < -1.675$ ⇒ 拒絕 H_0.
⇒ 若賣家可以一次賣出大量 cocaine，賣家更願意接受低價

e. $\begin{cases} H_0: \beta_3 \le 0 \\ H_1: \beta_3 > 0 \end{cases}$ RR: $\{t : t > 1.675\}$

t-value = 0.572 < 1.675 ⇒ 不拒絕 H_0.

⇒ 高品質的 cocaine 並沒有溢價

f. $b_4 = -2.35458$ ⇒ cocaine 價格隨時間下降.

⇒ 供給隨時間的增加速度大於需求

⇒ 可能原因是生產技術的改善