

---

---

---

---

---



Q 1.

$$(2.7) \quad b_2 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$(2.8) \quad b_1 = \bar{y} - b_2 \bar{x}$$

let  $k=2$ 

$$Y_i = b_1 + b_2 X_i + \varepsilon_i \quad \text{and} \quad X = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \quad Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \quad \text{then} \quad X'X = \begin{bmatrix} 1 & 1 & \dots & 1 \\ X_1 & X_2 & \dots & X_n \end{bmatrix} \begin{bmatrix} 1 \\ X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 1 & 1 & \dots & 1 \\ X_1 & X_2 & \dots & X_n \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix}$$

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

$$\therefore b = (X'X)^{-1}(X'Y) = \frac{1}{n \sum X_i^2 - (\sum X_i)^2} \begin{bmatrix} \sum X_i^2 & -\sum X_i \\ -\sum X_i & n \end{bmatrix} \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix}$$

$$= \frac{1}{n \sum X_i^2 - (\sum X_i)^2} \begin{bmatrix} \sum X_i^2 \sum Y_i - \sum X_i \sum X_i Y_i \\ -\sum X_i \sum Y_i + n \sum X_i Y_i \end{bmatrix}$$

$$\text{where } b_1 = \frac{\sum X_i^2 \sum Y_i - \sum X_i \sum X_i Y_i}{n \sum X_i^2 - (\sum X_i)^2}$$

$$b_2 = \frac{-\sum X_i \sum Y_i + n \sum X_i Y_i}{n \sum X_i^2 - (\sum X_i)^2}$$

$$\text{from (2.7)} \quad b_2 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{\sum (x_i y_i - x_i \bar{y} - \bar{x} y_i + \bar{x} \bar{y})}{\sum (x_i^2 - 2x_i \bar{x} + \bar{x}^2)}$$

$$= \frac{\sum x_i y_i - \bar{y} \sum x_i - \bar{x} \sum y_i + n \bar{x} \bar{y}}{\sum x_i^2 - 2\bar{x} \sum x_i + n \bar{x}^2}$$

$$= \frac{\sum x_i y_i - 2n \bar{x} \bar{y} + n \bar{x} \bar{y}}{\sum x_i^2 - 2\bar{x} n \bar{x} + n \bar{x}^2}$$

$$= \frac{n \sum x_i y_i - \bar{x} \bar{y}}{\sum x_i^2 - n \bar{x}^2}$$

$$= \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - n^2 \bar{x}^2}$$

$$= \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} \quad \#$$

$$b_1 = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{n \sum x_i^2 - (\sum x_i)^2} \quad b_1 = \bar{y} - b_2 \bar{x} = \frac{\bar{y} [\sum (x_i - \bar{x})^2] - \bar{x} \sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$= \frac{\bar{y} [\sum (x_i^2 - 2x_i \bar{x} + \bar{x}^2)] - \bar{x} [\sum x_i y_i - \bar{y} \sum x_i - \bar{x} \sum y_i + \bar{x} \bar{y} n]}{\sum (x_i - \bar{x})^2}$$

$$= \frac{\bar{y} \sum x_i^2 - 2\bar{x} \bar{y} \sum x_i + \bar{y} \bar{x}^2 n - \bar{x} \sum x_i y_i + \bar{x} \bar{y} \sum x_i + \bar{x}^2 \bar{y} \sum y_i - \bar{x}^2 \bar{y} n}{\sum (x_i - \bar{x})^2}$$

$$= \frac{\bar{y} \sum x_i^2 - \bar{x} \bar{y} \sum x_i - \bar{x} \sum x_i y_i + \bar{x}^2 \bar{y} \sum y_i}{\sum (x_i - \bar{x})^2}$$

$$= \frac{\frac{\sum y_i}{n} \sum x_i^2 - \bar{x} \bar{y} \sum x_i - \frac{\sum x_i}{n} \sum x_i y_i + n \bar{y} \bar{x}^2}{\sum (x_i - \bar{x})^2}$$

$$= \frac{\sum x_i^2 \sum y_i - 2\bar{x} \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} \quad \#$$

$$\text{Q 2. (2.14)} \quad \text{var}(b_1 | x) = \sigma^2 \left[ \frac{\sum x_i^2}{n \sum (x_i - \bar{x})^2} \right]$$

$$(2.15) \quad \text{var}(b_2 | x) = \frac{\sigma^2}{\sum (x_i - \bar{x})^2}$$

$$(2.16) \quad \text{cov}(b_1, b_2 | x) = \sigma^2 \left[ \frac{-\bar{x}}{\sum (x_i - \bar{x})^2} \right]$$

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

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

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

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

$$= \begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 \end{bmatrix}$$

$$\text{where } \sigma_{11}^2 = \text{Var}(b_1 | x)$$

$$\sigma_{22}^2 = \text{Var}(b_2 | x)$$

$$\sigma_{12}^2 = \sigma_{21}^2 = \text{Cov}(b_1, b_2 | x) \quad \#$$

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  $\beta_4$ . 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.

C.  $CI = [-0.1503 - 1.96(0.0235), -0.1503 + 1.96(0.0235)] = [-0.1964, -0.1042]$ 

increase in 1 year of age the share of the alcohol expenditure is estimated decrease by an amount between 0.1042 & 0.1964 units

d. Except the intercept, all coefficient estimates are significantly different from 0 at 5% level. ( $\because$  p-values < 0.05)

e.  $H_0: \beta_3 = -2$        $t = \frac{-1.4515 + 2}{0.3695} = 1.495 < 1.96$ . We fail to reject  $H_0$  since there's no evidence to say that the additional of an extra children leads to a decline in the alcohol expenditure is different from 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  $\beta_2$ ,  $\beta_3$ , and  $\beta_4$ ?
- 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?

As quantity increase 1 unit, the price will decrease by 0.05997, and the others condition is held.

As quality increase 1 unit, the price will increase by 0.11621, and the others condition is held.

As time increase 1 year, the price will decrease by -0.35458, and the others condition is held.

} All the signs are same as expected.

a.

i.  $t$  of  $b_1 = \frac{1.4515}{2.2019} = 0.6592$

ii.  $se$  of  $b_2 = \frac{2.7648}{5.7103} = 0.4842$

iii.  $b_3 = 0.3695 \times (-3.9376) = -1.4549$

iv.  $R^2 = 1 - \frac{SS_E}{SS_T} = 1 - \frac{46221.62}{49041.5418} = 0.0575$ 

$SS_T = 6.39547^2 \times (1200 - 1) = 49041.5418$

v.  $\hat{\sigma} = \sqrt{\frac{SS_E}{N - k}} = \sqrt{\frac{46221.62}{1200 - 4}} = 6.217$

b.

$b_2 = 2.965$  :

1 unit increase in total expenditure will increase the share of alcohol expenditure by approximately 2.965 units, and the others condition is held.

$b_3 = -1.45494$ 

increase in # of children will decrease the share of alcohol expenditure by 1.45494 units, and the others condition is held.

$b_4 = -0.1503$ 

increase in 1 year of age will decrease the share of alcohol expenditure by 0.1503 units, and the others condition is held.

a. We expect  $\beta_2$  is negative because as # of grams increase, the price should increase.

$\beta_3$  is positive because the purer cocaine leads the higher price.

$\beta_4$  would depend on the demand & supply over the time. Fixed demand & increasing supply lead fall in price, fixed supply & increasing demand lead rise in price.

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				

$\hat{PRICE} = 90.84669 - 0.05997 QUANT + 0.11621 QUAL - 2.35458 TREND$

c.  $R^2 = 0.5097$

d.  $H_0: \beta_2 \geq 0$      $t = -5.892 < t_{(0.05, 54)} = -1.6736$ , we fail to reject  $H_0$ , means that sellers are willing to accept a lower price if they can make sales in larger quantities.  
 $H_1: \beta_2 < 0$

e.  $H_0: \beta_3 \leq 0$      $t = 0.572 < t_{(0.95, 54)} = 1.6736$ , we fail to reject  $H_0$ , means that we can't conclude that a premium is paid for better quality cocaine.  
 $\beta_3 > 0$

f. Ave. annual change in cocaine price is  $-2.35458 = b_4$ . the possible reason for decreasing price is the development of improved technology for producing.