

**Tutorial 9**

1. Let  $(x_i, y_i), 1 \leq i \leq n$  be a sequence of bivariate observations.

(a) Suppose we model the data by using

$$Y = \alpha + \beta x + \epsilon.$$

Show that the least squares estimators  $(\hat{\alpha}, \hat{\beta})$  of  $(\alpha, \beta)$  may be written as

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = (X^T X)^{-1} X^T Y,$$

where  $X^T = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{pmatrix}$  and  $Y^T = (y_1, y_2, \dots, y_n)$ .

(b) Suppose we model the data by using

$$Y = \alpha + \beta x^2 + \epsilon.$$

What are the least squares estimators of  $(\alpha, \beta)$ ?

2. The number of fatal road accidents in NSW over an eleven year period 1961-1971 were:

Year ( $x = \text{year}-1960$ )	1	2	3	4	5	6	7	8	9	10	11
Fatal accidents ( $y$ )	850	798	818	903	1026	1042	1022	1069	1070	1135	1096

(a) Fit a linear regression model to the data.

(b) What is your estimate for the expected change in the number of fatal accidents per year?

(c) Calculate two estimates for the variance,  $\sigma^2$ , of the number of fatal accidents per year: one with and the other without assuming the SLR model.

(d) Give a 95% confidence interval for the slope of the regression line.

(e) Use your regression line to predict the expected number of fatalities in 1972. (The actual number of fatal accidents was 981. This was the year that compulsory seat belt legislation was enforced.)

3. As an additional exercise do problem 2 on p. 592 of Rice's book.

## Computer Exercises week 9

Consider the data frame `fuel.frame`, which has information on makes of cars taken from the April 1990 issue of Consumer Reports.

- (b) Create two vectors  $x$  and  $y$  whose elements correspond to `Weight` (car weight in pounds) and `Fuel` (fuel economy in gallons per 100 miles) in the data `fuel.frame` respectively.
- (c) Plot the Fuel against Weight with Weight on the  $x$ -axis and comment on the plot.
- (d) Find the regression line for Fuel on Weight using `lm` and add this line on the first plot (Hint: use `abline`).
- (e) Find the  $p$ -value for testing whether the weight has an influence on the fuel economy.
- (f) Plot the residuals ( $y$ -axis) against the fitted values of fuel ( $x$ -axis), and add the horizontal line  $y = 0$ .
- (g) Obtain a boxplot of residuals and a normal qq-plot of the residuals.
- (h) Comment on the plots in (f) and (g), and explain whether the assumptions required in (e) are satisfied.