and β_2? What happens to the least squares estimates b_1 and b_2? What happens to the variance of the error term?

4.6 The fitted least squares line is $\hat{y}_i = b_1 + b_2 x_i$.
(a) Algebraically, show that the fitted line passes through the point of the means, $(\bar{x}, \bar{y})$.
(b) Algebraically show that the average value of $\hat{y}_i$ equals the sample average of y. That is, show that $\bar{\hat{y}} = \bar{y}$, where $\bar{\hat{y}} = \Sigma \hat{y}_i / N$.

4.7 In a simple linear regression model suppose we know that the intercept parameter is zero, so the model is $y_i = \beta_2 x_i + e_i$. The least squares estimator of β_2 is developed in Exercise 2.4.
(a) What is the least squares predictor of y in this case?
(b) When an intercept is not present in a model, R^2 is often defined to be $R_u^2 = 1 - SSE/\Sigma y_i^2$, where SSE is the usual sum of squared residuals. Compute R_u^2 for the data in Exercise 2.4.
(c) Compare the value of R_u^2 in part (b) to the generalized $R^2 = r_{y\hat{y}}^2$, where $\hat{y}$ is the predictor based on the restricted model in part (a).
(d) Compute $SST = \Sigma(y_i - \bar{y})^2$ and $SSR = \Sigma(\hat{y}_i - \bar{y})^2$, where $\hat{y}$ is the predictor based on the restricted model in part (a). Does the sum of squares decomposition $SST = SSR + SSE$ hold in this case?

4.5.2 Computer Exercises

4.8 The first three columns in the file *wa-wheat.dat* contain observations on wheat yield in the Western Australian shires Northampton, Chapman Valley, and Mullewa, respectively. There are 48 annual observations for the years 1950–1997. For the Chapman Valley shire, consider the three equations

$$y_t = \beta_0 + \beta_1 t + e_t$$
$$y_t = \alpha_0 + \alpha_1 \ln(t) + e_t$$
$$y_t = \gamma_0 + \gamma_1 t^2 + e_t$$

(a) Estimate each of the three equations.
(b) Taking into consideration (i) plots of the fitted equations, (ii) plots of the residuals, (iii) error normality tests, and (iii) values for R^2, which equation do you think is preferable? Explain.

4.9* For each of the three functions in Exercise 4.8
(a) Find the predicted value for yield when $t = 49$.
(b) Find estimates of the slopes dy_t/dt at the point $t = 49$.
(c) Find estimates of the elasticities $(dy_t/dt)(t/y_t)$ at the point $t = 49$.
(d) Comment on the estimates you obtained in parts (b) and (c). What is their importance?

4.10 The file *london.dat* is a cross section of 1519 households drawn from the 1980–1982 British Family Expenditure Surveys. Data have been selected to include only households with one or two children living in Greater London. Self-employed and retired households have been excluded. Variable definitions are in the file *london.def*. The budget share of a commodity, say food, is defined as

$$WFOOD = \frac{\text{expenditure on food}}{\text{total expenditure}}$$