errors e_i and to have the same mean 0 and variance σ^2. Under these assumptions, the best linear unbiased predictor of y_0 is given by

$$\hat{y}_0 = b_1 + x_{02}b_2 + x_{03}b_3$$

where b_k's are the least squares estimators. This predictor is unbiased in the sense that the average value of the forecast error is zero. That is, if $f = (y_0 - \hat{y}_0)$ is the forecast error, then $E(f) = 0$. The predictor is best in the sense that the variance of the forecast error for all other linear and unbiased predictors of y_0 is not less than $\text{var}(y_0 - \hat{y}_0)$.

The variance of forecast error $\text{var}(y_0 - \hat{y}_0)$ contains two components. One component occurs because b_1, b_2, and b_3 are estimates of the true parameters, and the other component is a consequence of the unknown random error e_0. The expression for $\text{var}(y_0 - \hat{y}_0)$ is given by

$$\begin{aligned}\text{var}(f) &= \text{var}[(\beta_1 + \beta_2 x_{02} + \beta_3 x_{03} + e_0) - (b_1 + b_2 x_{02} + b_3 x_{03})] \\ &= \text{var}(e_0 - b_1 - b_2 x_{02} - b_3 x_{03}) \\ &= \text{var}(e_0) + \text{var}(b_1) + x_{02}^2 \text{var}(b_2) + x_{03}^2 \text{var}(b_3) \\ &\quad + 2x_{02}\,\text{cov}(b_1, b_2) + 2x_{03}\,\text{cov}(b_1, b_3) + 2x_{02}x_{03}\,\text{cov}(b_2, b_3)\end{aligned}$$

To obtain $\text{var}(f)$ we recognized that the unknown parameters and the values of the explanatory variables are constants, and that e_0 is uncorrelated with the sample data, and thus is uncorrelated with the least squares estimators (b_1, b_2, b_3). The remaining variances and covariances of the least squares estimators are obtained using the rule for calculating the variance of a weighted sum in Appendix B.4.3.

Each of the terms in the expression for $\text{var}(f)$ involves σ^2. To obtain the estimated variance of the forecast error $\widehat{\text{var}(f)}$, we replace σ^2 with its estimator $\hat{\sigma}^2$. The standard error of the forecast is given by $\text{se}(f) = \sqrt{\widehat{\text{var}(f)}}$. If the random errors e_i and e_0 are normally distributed, or if the sample is large, then

$$\frac{f}{\text{se}(f)} = \frac{y_0 - \hat{y}_0}{\sqrt{\widehat{\text{var}(y_0 - \hat{y}_0)}}} \sim t_{(N-K)}$$

Following the steps we have used many times, a $100(1-\alpha)\%$ interval predictor for y_0 is $\hat{y}_0 \pm t_c \text{se}(f)$, where t_c is a critical value from the $t_{(N-K)}$-distribution.

Thus, the methods for prediction in the model with $K = 3$ are straightforward extensions of the results from the simple linear regression model. For $K > 3$, the methods extend in a similar way.

6.9 Exercises

Answers to exercises marked * appear in Appendix D at the end of the book.

6.9.1 Problems

6.1 When using $N = 40$ observations to estimate the model

$$y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + e_i$$

you obtain $SSE = 979.830$ and $\hat{\sigma}_y = 13.45222$. Find

(a) R^2.

(b) The value of the F-statistic for testing $H_0: \beta_2 = \beta_3 = 0$. Do you reject or fail to reject H_0?