Chapter 12

- Prediction. New value for the covariates is

$$\mathbf{x'}_0 = (1, x_{10}, x_{20}, ..., x_{k0}).$$

A $(1-\alpha)$ confidence interval for the mean response $E(Y_0|\mathbf{x}_0)$ is

$$\left[\hat{Y}_0 - t_{\alpha/2,n-k-1} S\sqrt{\mathbf{x'}_0 (\mathbf{X'X})^{-1}\mathbf{x}_0}, \hat{Y}_0 + t_{\alpha/2,n-k-1} S\sqrt{\mathbf{x'}_0 (\mathbf{X'X})^{-1}\mathbf{x}_0}\right],$$

where

$$\hat{Y}_0 = \hat{\beta}_0 + \sum_{j=1}^{k} \hat{\beta}_j x_{j0}.$$

A $(1-\alpha)$ prediction interval for the new observation Y_0 is

$$\left[\hat{Y}_0 - t_{\alpha/2,n-k-1} S\sqrt{1+\mathbf{x'}_0 (\mathbf{X'X})^{-1}\mathbf{x}_0}, \hat{Y}_0 + t_{\alpha/2,n-k-1} S\sqrt{1+\mathbf{x'}_0 (\mathbf{X'X})^{-1}\mathbf{x}_0}\right].$$

- Coefficient of determination

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}.$$

Adjusted coefficient of determination

$$R^2_{\text{adj}} = 1 - \frac{SSE/(n-k-1)}{SST/(n-1)}.$$

- C_p statistic

$$C_p = p + \frac{(s^2 - \hat{\sigma}^2)(n-p)}{\hat{\sigma}^2},$$

where p is the number of model parameters, s^2 is the mean square error for the candidate model, $\hat{\sigma}^2$ is the mean square error from the most complete model. One chooses the model with minimal C_p.