Tutorial 7

1. A student stated: "Adding predictor variables to a regression model can never reduce R^2, so we should include all available predictor variables in the model." Comment.

2. For a model with X_1, X_2, X_3, X_4 predictors, we have $n = 30$ and

$$SSE(X_1) = 161.081, SSE(X_2) = 195.846, SSE(X_3) = 56.432, SSE(X_4) = 225.584$$

$$SSE(X_1, X_2) = 146.635, SSE(X_1, X_3) = 16.579, SSE(X_1, X_4) = 161.044,$$

$$SSE(X_2, X_3) = 45.660, SSE(X_2, X_4) = 195.403, SSE(X_3, X_4) = 56.431$$

$$SSE(X_1, X_2, X_3) = 12.436, SSE(X_1, X_2, X_4) = 146.604, SSE(X_1, X_3, X_4) = 16.383,$$

$$SSE(X_2, X_3, X_4) = 45.656, SSE(X_1, X_2, X_3, X_4) = 12.246, SST = 226.189$$

   (a) find $SSR(X_1, X_2|X_3, X_4)$

   (b) In model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$, test $H_0 : \beta_1 = \beta_2 = 0$ with $\alpha = 0.05$.

   (c) Find the largest model in which every predictor variable is not significant at $\alpha = 0.05$.

3. State the number of degrees of freedom that are associated with each of the following extra sums of squares: (1) $SSR(X_1 \mid X_2)$; (2) $SSR(X_2 \mid X_1, X_3)$ (3) $SSR(X_1, X_2 \mid X_3, X_4)$; (4) $SSR(X_1, X_2, X_3 \mid X_4, X_5)$

4. Refer to **Patient satisfaction** date (see the **Example** of Chapter 2)

   a. Obtain the analysis of variance table that decomposes the regression sum of squares into extra sums of squares associated with X_2; with X_1, given X_2; and with X_3, given X_2 and X_1.

   b. Test whether X_3 can be dropped from the regression model given that X_1 and X_2 are retained. Use the F^* test statistic and level of significance 0.025. State the alternatives, decision rule, and conclusion. What is the P-value of the test?

5. Refer to **Patient satisfaction** data above, test whether $\beta_1 = -1.0$ and $\beta_2 = 0$; use $\alpha = 0.025$. State the alternatives, full and reduced models, decision rule and conclusion.

6. Show that:

   a. $SSR(X_1, X_2, X_3, X_4) = SSR(X_1) + SSR(X_2, X_3 \mid X_1) + SSR(X_4 \mid X_1, X_2, X_3)$.

   b. $SSR(X_1, X_2, X_3, X_4) = SSR(X_2, X_3) + SSR(X_1 \mid X_2, X_3) + SSR(X_4 \mid X_1, X_2, X_3)$.

7. The following regression model is being considered in a water resources study: $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1} X_{i2} + \beta_4 \sqrt{X_{i3}} + \varepsilon_i$. State the reduced model for testing whether or not: (1) $\beta_3 = \beta_4 = 0$, (2) $\beta_3 = 0$, (3) $\beta_1 = \beta_2 = 5$, (4) $\beta_4 = 7$.