Homework Problem Set #1

'air.csv' includes a Civil Aeronautics Board report from SAS System for Regression, Third Edition. The data concern important predictors in predicting the cost of providing air service. The variables are

- CPM: cost per passenger mile (cents).

- UTL: average hours per day use of aircraft

- ASL: average length of nonstop legs of flights (1000 miles)

- SPA: average number of seats per aircraft (100 seats)

- ALF: average load factor (% of seats occupied by passengers).

CPM is a response and the other variables are predictors.

(1) Fit a linear regression model using SAS (PROC REG) and R (lm). State your linear regression model with assumptions. Are all predictors significant at a significance level $\alpha = 0.05$? Interpret the coefficient of UTL.

(2) Obtain $\hat{\boldsymbol{\beta}}$, its standard errors, and 95% CIs using matrix calculation in R.

(3) Fit H_0 and H_1 using SAS. Perform the t-test and the F-test (by hand) using the SAS outputs from fitting H_0 and H_1 to test

$$H_0 : CPM = \beta_0 + \beta_1 UTL + \beta_3 SPA + \beta_4 ALF + e, \quad \text{vs.}$$

$$H_1 : CPM = \beta_0 + \beta_1 UTL + \beta_2 ASL + \beta_3 SPA + \beta_4 ALF + e.$$

Compare the p-values from the t-test and the F-test. Confirm your p-values using PROC REG TEST statement. Are these two models significantly different?

(4) Fit H_0 and H_1 using SAS. Test the following hypothesis (by hand) using the SAS outputs from fitting H_0 and H_1:

$$H_0 : CPM = \beta_0 + \beta_1 UTL + e, \quad \text{vs.}$$

$$H_1 : CPM = \beta_0 + \beta_1 UTL + \beta_2 ASL + \beta_3 SPA + e.$$

Confirm your p-values using PROC REG TEST statement. Are they significantly different?

(5) Under the full model, i.e., $CPM = \beta_0 + \beta_1 UTL + \beta_2 ASL + \beta_3 SPA + \beta_4 ALF + e$, produce $Cov(\hat{\boldsymbol{\beta}})$. Using the t-test with $Cov(\hat{\boldsymbol{\beta}})$, test $H_0 : \beta_3 = \beta_4$ by hand. Compare the two models using the F-test by hand. Confirm your p-values using PROC REG TEST statement. What is your conclusion?

(6) Under the full model, i.e., $CPM = \beta_0 + \beta_1 UTL + \beta_2 ASL + \beta_3 SPA + \beta_4 ALF + e$, test $H_0 : \beta_3 + 2\beta_4 = -17$ using the t-test and the F-test (by hand). Confirm your p-values using PROC REG TEST statement. What is your conclusion?