APU2141- Regression Analysis I

Open Book Test 2016/2017

Solutions

1. a) False; Pearson correlation coefficient measures the strength of linear association between two variables. There can be relationships between two variables of the form $y = \beta_0 + \beta_1 x^2 + \varepsilon$, where a simple linear regression model may fit the data well but with the Pearson correlation coefficient near zero.

   b) True; Pearson correlation coefficient *r* and the estimate for the slope parameter satisfy the relationship $\hat{\beta} = \frac{rs_y}{s_x}$ where s_y and s_x denote the standard deviations of data on response and the predictor. Hence, $\hat{\beta}$ and r have the same sign.

   c) False; Tightly clustered about a line does not guarantee that the Pearson correlation coefficient is close to ± 1. For instance, if the points are tightly clustered about a line parallel to the axis representing the predictor variable, Pearson correlation coefficient will be near zero.

   d) False; The parameter β_1 represents the expected change in the response associated with one unit of change in the explanatory variable. Change in the response associated with two points corresponding to one unit of change in the explanatory variable is equal to the difference in the corresponding error terms.

   e) False; Random errors are deviations of observable responses about the unknown population regression line. An observation is above the fitted line does not give any information about whether the point is above or below the population regression line.

2. a) The objective can be achieved by fitting a simple linear regression model.

I will choose $(t-20)^2$, where t is the temperature as the explanatory variable.

b) The objective can be achieved by fitting a simple linear regression model.

I will choose $\sqrt{age}$ as the explanatory variable.

c) The objective cannot be achieved by fitting a simple linear regression model.

Note that the partial derivative of μ with respect to α_1 involves parameters. Hence, the regression function is nonlinear.

d) The objective cannot be achieved by fitting a simple linear regression model.

Here, the model involves more than one predictor variable. Hence, this is not a simple linear regression model.

e) The objective cannot be achieved by fitting a simple linear regression model.

Even though the regression function can be written in the linear form, the resulting equation is not linear in the given parameters. Hence, the objective cannot be achieved.