
1: Multiple regression models

(a) **R commands:** (with interpretation)

Brand preference. A small-scale experimental study investigates the relation between degree of brand liking (Y) and moisture content (X_1) and sweetness (X_2) of the product.

Data are available in the homework page (the link is near this homework).

- Obtain the scatter plot matrix and the correlation matrix. What information do these diagnostic aids provide here?
- Fit the regression model to the data, stating the estimated regression function. How are the coefficients interpreted here?
- Obtain the residuals and prepare a boxplot of the residuals. What information do this plot provide?
- Plot the residuals against \hat{Y} , \mathbf{X}_1 , \mathbf{X}_2 and $\mathbf{X}_1 \mathbf{X}_2$ on separate graphs. Also prepare a q-q plot. Interpret the plots and summarize your findings.
- test whether there is a regression relation, using $\alpha = .01$. State the alternatives, decision rule, and conclusion. What does your test imply about β_1 and β_2 ?
- Calculate the coefficient of multiple determination R^2 . How is it interpreted here?
- Calculate the coefficient of simple determination R^2 between Y_i and \hat{Y}_i . Does it equal the coefficient of multiple determination obtained in the previous point?
- Obtain an interval estimate of $E(Y_h)$ when $X_{h1} = 5$ and $X_{h2} = 4$. Use a 99% confidence coefficient. Interpret your interval estimate.
- Obtain a prediction interval for a new observation Y_h when $X_{h1} = 5$ and $X_{h2} = 4$. Use a 99% confidence coefficient. Interpret your interval estimate and compare it the interval estimate obtained at the previous point.