Regression Analysis - STAT 482 - Exam 2: Due Dec 14, 2021

Alex Towell (atowell@siue.edu)

Problem 1

A beer distributor is interested in the amount of time to service its retail outlets. Two factors are thought to influence the delivery time (y) in minutes: the number of cases delivered (x_1) and the distance traveled (x_2) in miles. A random sample of delivery time data has been collected. The data is available on Blackboard as a csv file.

(a)

Provide an interpretation of a regression coefficient in a multiple regression model.

(b)

Compute b_1, b_2, the estimated the regression coefficients for the delivery time data.

(c)

Compute t statistics for testing the effect of each input variable. Explain what type of effect is being tested here.

(d)

Compute $\mathrm{SS}_{\mathrm{R}}(X_1)$ and $\mathrm{SS}_{\mathrm{R}}(X_2|X_1)$. Explain what each sum of squares represents.

(e)

Test for a marginal effect of x_2 against a model which includes no other input variables. (Compute the test statistic and p-value.) Provide an interpretation of the result, stated in the context of the problem.

(f)

Test for a partial effect of x_2 against a model which includes x_1. (Compute the test statistic and p-value.) Provide an interpretation of the result, stated in the context of the problem.

(g)

Compute the correlation matrix. What feature of multidimensional modeling is illustrated in this problem?

Problem 2

A bakery is interested in the best formulation for a new product. A small-scale experiment is conducted to investigate the relationship between the product satisfaction (y), and the moisture content (input 1) and sweetness (input 2) of the product. The input variables have been coded (x_1, x_2) for ease of calculation. The data is available on Blackboard as a csv file.

(a)

Provide a definition for an orthogonal design. Discuss an advantage to using an orthogonal design.

(b)

Provide a definition for an interaction effect.

(c)

Fit an interaction model using the coded variables. Compute the regression coefficient estimates and their standard errors.

(d)

Write the estimated regression as a function of x_1 for $x_2 = 1, 0, -1$.

(e)

Create interaction plots for both the interaction model and the additive effects model.

(f)

Test for an interaction effect. (Compute the test statistic and p-value.)

Problem 3

An engineer is interested in comparing three chemical processes (categorical input with groups A,B,C) for manufacturing a compound. It is suspected that the impurity (continuous input x) of the raw material will affect the yield (response variable y) of the product. The data is available on Blackboard as a csv file.

(a)

Define indicator variables I_1 and I_2 using chemical process C as the baseline level.

(b)

Write an additive model for response y using continuous input variable x and indicator variables I_1, I_2.

(c)

Write a regression function for each of the chemical processes.

(d)

Provide an interpretation for each effect parameter, stated in the context of the problem.

(e)

Compute interval estimates for each of the effect parameters.

(f)

Create a scatterplot of the data with the estimated regression lines.