Solutions to Exercises 4.12 and 4.14

1 Exercise 4.12: More Complex Multiple Regression Models

1.1 4.43 First-Order Model

A first-order linear model assumes a linear relationship between the dependent variable y and the independent variables.

1.1.1 (a) Two quantitative independent variables

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

1.1.2 (b) Four quantitative independent variables

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$$

1.2 4.44 Second-Order Model

A second-order model includes quadratic terms and interactions to capture non-linear relationships.

1.2.1 (a) Two quantitative independent variables

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_2^2 + \beta_5 x_1 x_2$$

1.2.2 (b) Three quantitative independent variables

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1^2 + \beta_5 x_2^2 + \beta_6 x_3^2 + \beta_7 x_1 x_2 + \beta_8 x_1 x_3 + \beta_9 x_2 x_3$$

1.3 4.45 Qualitative Predictors

Qualitative variables require the use of dummy variables.

1.3.1 (a) Two levels (A and B)

Define a dummy variable:

$$D = \begin{cases} 1, & \text{if A} \\ 0, & \text{if B} \end{cases}$$

The model is:

$$E(y) = \beta_0 + \beta_1 D$$

1.3.2 (b) Four levels (A, B, C, D)

Define three dummy variables:

$$E(y) = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3$$

where: - $D_1 = 1$ for A, 0 otherwise, - $D_2 = 1$ for B, 0 otherwise, - $D_3 = 1$ for C, 0 otherwise.

1.4 4.46 and 4.47 Graphing First-Order Models

These exercises involve graphing first-order regression models. The response surfaces are planes, and when plotted against one independent variable, they form straight lines.

1.5 4.48 and 4.49 Graphing Second-Order Models

These models contain squared terms and interaction terms, resulting in curved surfaces when plotted. The interaction term modifies the shape of the graph, making it non-parallel.

2 Exercise 4.14: Regression Model for Oil Removal

This problem involves fitting a first-order regression model with interaction terms for predicting voltage in an oil removal experiment.

2.1 Given Model

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_5 + \beta_4 x_1 x_2 + \beta_5 x_1 x_5$$

2.2 (a) Interaction Effects

The interaction terms $x_1 x_2$ and $x_1 x_5$ indicate that the effect of disperse phase volume on voltage is influenced by salinity and surfactant concentration.

2.3 (b) Model Fitting in Minitab

Regression analysis should be performed using **Minitab**. The dataset used is the WATEROIL dataset. The results should include: - Regression coefficients, - R^2 and R^2_{adj}, - Residual analysis.

2.4 (c) Interpretation of Coefficients

- **Main effects** ($\beta_1, \beta_2, \beta_3$) show the independent influence of each predictor.
- **Interaction terms** (β_4, β_5) indicate how predictors interact.

2.5 Minitab Analysis

1. Load the WATEROIL dataset into Minitab.
2. Run multiple regression:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_5 + \beta_4 x_1 x_2 + \beta_5 x_1 x_5$$

3. Extract:
   - Coefficient estimates,
   - R^2 for model fit,
   - Residuals to check assumptions.

2.6 Conclusion

If R^2 is high and p-values for interactions are low, the model is a strong predictor. Otherwise, model improvements are needed.