

CHE384 Data to Decisions
Chris Mack, University of Texas at Austin

Homework #9 – Logistic Regression

Turn in your solution with the answers to the questions below. Also, email to me the supporting spreadsheet and/or R script that you used to perform the analysis. (Please name the file using this format: HW9_yourname.R).

1. To test the toxicity of a particular substance, six groups of 250 insects each were given different doses of the toxic substance. The dose (x) is given in arbitrary units on a logarithmic scale. One day after exposure, the number of insects that were dead were counted (y).

Use R to perform a logistic regression:

- Plot the sample proportion that died (y/n) as a function of x . Does the plot suggest that the logistic regression function is appropriate?
- Find the maximum likelihood estimates for β_0 and β_1 . State the fitted response function.
- Plot the sample data together with the fitted response function. Does the fit appear to be a good one?
- What is the estimated probability that an insect dies when the dose is $x = 3.5$?
- What is the estimated median lethal dose (the dose where there is a 50% chance of the insect dying)?

2. A random sample 33 families were surveyed to determine their annual family income (x_1 , in thousands of dollars) as well as the age of their oldest car (x_2 , in years). One year later, a follow-up survey determined if they bought a car ($y = 1$) or did not purchase a car ($y = 0$) in the past year. Assume that a multiple logistic regression model first order in the two predictor variables is appropriate.

Use R to perform a logistic regression:

- Find the maximum likelihood estimates for β_0 , β_1 and β_2 . State the fitted response function.
- Obtain $\exp(b_1)$ and $\exp(b_2)$ and explain these numbers.
- What is the estimated probability that a family with an income of \$50,000 and an oldest car that is 3 years old will buy a car in the next year?