

Basic Generalized Linear Modeling – Part 4: Exercises



This exercise is going to be the last exercise on Basic Generalized Linear Modeling (GLM). Please click [here](#) to find the other part of the Basic GLM Exercise that you've missed.

In this exercise, we will discuss Logistic Regression models as one of the GLM methods. The model is used where the response data is binary (ex. male or female, present or absence) or proportional (ex. percentage and ratio.)

```
M1 <- glm(response ~ Predictor1 + Predictor2, family =  
          binomial)
```

Data-sets are used based on Polis et al. (1998), which is recorded island characteristics in the Gulf of California. While the analysis is based on Quinn and Keough (2002), the data model presences/absences are of a spider predator against the perimeter to area ratio of the islands.

Answers to these exercises are available [here](#). If you obtained a different (correct) answer than those listed on the solutions page, please feel free to post your answer as a comment on that page. Load the data-set [here](#), call it 'spider' and load all the required packages before running the exercise.

Exercise 1

Visualize the data.

Exercise 2

Run the model.

Exercise 3

Check for over-dispersion.

Exercise 4

Use component+residual plots (crPlots) for further checking on dispersion.

Exercise 5

Check influential values.

Exercise 6

Check The Cooks Distance and the model summary.

Exercise 7

Check residuals.

Exercise 8

Plot and predict. Calculate the predicted values based on the fitted model.

Exercise 9

Produce a final plot, including the base plot, plot fitted model and 95% CI bands.

Exercise 10

Check the odds ratio to estimate the probability of presence, given the unit increases in perimeter or area ratio.

Exercise 11

Estimate the R2 value. What can be inferred?