**Exercise 9: Inference for Categorical Responses**

**9.1 Single Proportion**

1. Load the data Catfish.csv from <https://tinyurl.com/4pxcj74m> and assign it a variable name.

2. Examine the data structure and determine the proportion of successful pigeon hunting of the total recorded stranding behavior of the wild catfish and its 95% Confidence Interval.

3. Remember to set.seed() before you start working on your code for the bootstrap sample.

4. Create a matrix to store the bootstrapped samples. The matrix should be of the same number of rows as the number of observations in the original data and number of columns should be equal to number of bootstrap events. Assign the matrix an object name.

5. Resample from the original samples with replacement 1000 times to create a bootstrap sample. Use the basic functions in R to write your own bootstrapping function.

6. Create vectors for the bootstrap proportions.

7. Visualize the bootstrap sample distribution of the proportions in a histogram.

8. Determine the standard error and the 95% confidence interval of the proportion of the successful hunting from the beaching behavior of the wild catfish based on the bootstrap sample distribution. What inference can you make from this value of the confidence interval? How can you use the bootstrap confidence interval to test for a hypothesis that hunting success in wild catfish is greater than the recorded 33% of veteran human hunters?

9. Test using the one proportion z-test if the bootstrapped proportion of hunting success of the wild catfish on pigeon is equal to the hunting success of the veteran hunters on pigeons of 33%.

**9.2 Difference of Two Proportions**

1. Load the data on Green Turtle.csv from <https://tinyurl.com/e3weddfn> and assign it a variable name.

2. Examine the data structure and determine the proportion of females hatched in each nest depth, the standard error, and the 95% Confidence Interval of each group.

3. State the hypothesis that you will test.

4. Apply what you have learned on bootstrapping techniques in the previous section of this exercise and resample with replacement 1000 times the data on sex of sea turtle hatchlings to create a bootstrap samples. Note that you now have a data set with two groups in the explanatory variable.

5. Create a vector of the bootstrap proportions of females grouped according to nest depth. Assign the vector an object name.

6. Visualize the bootstrap distribution of the proportion of females from each nest depth.

7. Determine the bootstrap standard error and the bootstrap 95% confidence interval of the proportion of females in each nest depth. Interpret your results. How do the bootstrap estimates differ from those of the original data set? Do your results support the hypothesis that shallow nests are female biased?

**9.3 Independence in Two-Way Table**

1. Using the same data set on hatching of sea turtles in 9.2, state the hypothesis that you will test about the effect of nest temperature on sex of hatchlings.

2. Perform the chi-square test on the original data on proportions of females and males hatched from nests to different temperatures.

3. Use the bootstrap sample matrix from section 9.2. and determine the proportion of females in each group of nest temperature per bootstrap event .

4. Compute the chi-squared test statistic for each bootstrap event and store the results in a vector with an assigned object name.

5. Visualize the bootstrap distribution of the chi-square values and determine the 95% confidence interval.

6. How does the bootstrap 95% confidence interval of chi-square values compare with that of the original data? Is there a statistical evidence to support your hypothesis?