# **Of Frogs and Frequency Analysis**

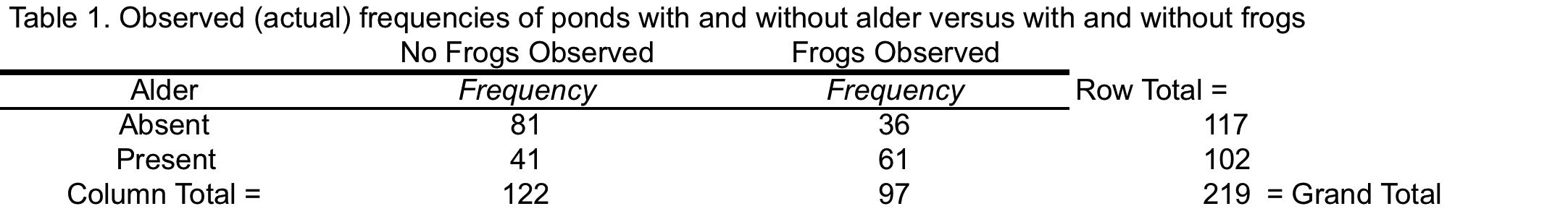
**Background:**

Long thought to be restricted to the realm of science fiction, cryobiologists (biologists who study of how living organisms respond to cold temperatures) have discovered many organisms that can survive the effects of freezing. One famous example is the wood frog (*Rana sylvatica*) that can survive being frozen for months out of the year, only to thaw and resume normal life functions during the summer. Carroll biologists studied the wood frog in Denali National Park trying to determine the habitat requirements for this unusual amphibian (see Hokit and Brown 2006). The central method involved counting pond sites with versus without wood frogs with respect to categorized habitat conditions.

This exercise will introduce you to a common type of inferential statistic known as frequency analysis (also known as Chi-square Test for Independence). Frequency analysis allows you to test for differences between two or more data sets that are nothing more than counts of observed events. Anything you can count can be tested with frequency analysis.

**Question 1:**

The data below include the number of times biologists observed wood frogs in ponds with alder versus without alder vegetation. Alder is a dominant shrub in the Alaskan tundra and because the wood frog prefers areas with a forest canopy, biologists want to test the hypothesis that wood frogs would be found more frequently in ponds that had alder vegetation around the perimeter versus ponds that did not have alder.

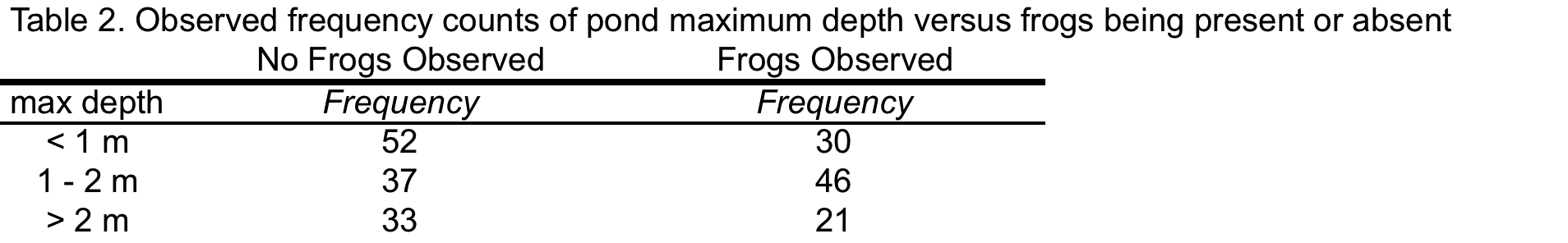


Create a Chi-square template in EXCEL to compare the observed frequencies with the expected frequencies assuming that alder and frog presence are independent of each other. Note that there should be 3 tables. Table 1 includes the observed data: the counts of sites with and without frogs (in columns) crossed with sites that did and did-not have alder (in rows). Table 2 includes the expected frequencies with the same column and row headings. Note that expected frequencies for each cell = (the row total) x (the column total) / (the grand total) based on the observed data. Table 3 is the difference between the observed and the expected squared and then divided by the expected. Add a cell below table 3 that automatically calculates the Chi-Square statistic as the sum of all the values in table 3. Insert an EXCEL function called the CHITEST that provides a p-value, the probability that the observed values are consistent with a null distribution based on the expected frequencies.

Note that if the p-value is less than 0.05, by convention the biologists should reject the null hypothesis and conclude that the presence of alder and the observation of frogs are not independent of each other. In other words, they should conclude that the observation of frogs depends on the presence or absence of alder. Based on the observed frequency counts, which sites are more likely to have frogs: those with or without alder?

**Question 2:**

The data below count the number of ponds in 3 depth (shallow, medium, deep) categories versus counts of sites with and without frogs. Calculate a Chi-square statistic and associated p-value for the depth data to test the hypothesis that depth and the presence of frogs are independent of each other. Note: because this data set has 3 levels of pond depth, the tables will be 2 x 3. Are the two variables independent (p > 0.05) or not (p < 0.05)? In other words, does the depth of the pond influence whether or not we observed frogs at a pond? Ponds of which depth category are more likely to have frogs?



**References:**

Hokit, D.G. and A. Brown. 2006. Distribution patterns of wood frogs (*Rana sylvatica*) in Denali National Park. Northwest Naturalist 87:128-137.