**Community Ecology Pre-Lab**

**Introduction to Community Ecology**

Biological communities examine the relationship of different species within the same geographical area. The abundance of one species may impact another, or the **richness** (number of species) of an area can give scientists an idea on the biodiversity of a region. Typically, the higher the biodiversity, the more resilient the community is to change. Humans also benefit from increased biodiversity as we rely on it for medicines, ecotourism, soil and water management, and a plethora of more services. Richness does not provide the whole picture though, **evenness** examines the abundance of each species in relation to each other. Communities that are more even have populations of relatively equal size. While both of these metrics in isolation can provide insight into a community, scientists use the **Shannon-Wiener Diversity Index (*H’*)** as a way to combine these metrics.

This tool allows for a quick comparison between communities to determine which is more biodiverse in order to make conservation decisions. To calculate the Shannon-Wiener Diversity Index (denoted as *H’*), use the following formula:

= The proportion of the entire community made up of species *i*

*ln* = the natural log

**Calculating the Shannon-Wiener Diversity Index (*H’*)**

A wildlife ecologist wanted to compare the diversity of a wetland that was downstream of a new housing development to that of a non-disturbed wetland found upstream to see if there were any potential impacts of the construction. She went out and surveyed all the plants found within a 5m x 5m plot. The information of the species she found is in the Prelab\_Data Excel Workbook in the Diversity Index worksheet. Use the information in this worksheet to answer the following questions.

1. What is the species richness of the disturbed wetland (downstream of construction)? The wetland with no disturbance (upstream)?
2. Which community is more even?
3. Using the formula above and the notes in the Excel sheet, what is the *H’* of the disturbed wetland? Of the undisturbed wetland? *Round to 2 decimal places.*
4. Which community has a greater diversity?

**Using Statistics to Answer Problems**

While a scientist could do one transect or one plot to gather data on species and their abundance, we cannot be sure that it is a truly representative sample of the community as a whole. To be more thorough, scientists will do many samplings to take into account outliers and natural variation. Scientists can use the data from multiple samplings to develop conclusions comparing different communities to each other using **statistics**.

In wildlife ecology, researchers would likely take multiple samples across the landscape to get an average diversity since some areas within the region may have more diversity than others. Researchers can then use this dataset and compare them to a different region in order to focus their efforts on a region that needs more resources to improve its biodiversity. We can use a statistical test, in this case a t-test, to compare the two datasets to see if they are statistically different from each other. The output of a t-test is a p-value, which will tell us the likelihood that the differences in values we see is explained by the null hypothesis. In science, we usually use a threshold (the alpha value, α) of 0.05.

Thus, we can accept the following:

* **If p<0.05, we can reject the null hypothesis & accept the alternative hypothesis; the data is significantly different between groups**
* **If p>0.05, we fail to reject the null hypothesis; the data is not significantly difference between groups**

Let’s say the scientist above wants to expand her study to provide statistical rigor to her results. She goes back out to both habitats and samples nine additional 5m x 5m plots and calculates the Shannon-Wiener Diversity Index (*H’*). Her results are posted in the Excel Workbook in the Graphing worksheet.

1. What is the null hypothesis of this study?
2. What is the alternative hypothesis of this study?
3. What is the average diversity (H’) in the disturbed wetland? What is the average diversity (H’) in the undisturbed wetland?
4. In the first box in the “Graphing” worksheet, create a bar graph comparing the averages of the two groups.

We have learned about standard deviation, which is a measure of the spread in a dataset. Standard error is similar, but takes into account the impact of the sample size. **Larger sample sizes make standard error smaller, meaning we are more confident that our true value is closer to the mean.** To calculate standard error, divide the standard deviation by the square root of the number of samples. In Excel, this would look like:

= STDEV(*range)*/SQRT(N)

where *range* is the cells that contain the data and N is the number of samples

1. What is the standard error for the disturbed wetland? What is the standard error for the undisturbed wetland?
2. What is the range for the H’ values in the disturbed wetland? What is the range for the H’ values in the undisturbed wetland?
3. Using the values from above and the instructions in the “Graphing” worksheet, add error bars to your current bar graph.
4. Do the error bars represent the range you calculated in #10 above? If not, what values do the error bars represent?
5. What do the error bars on the graph tell you about the data?
6. Would having a larger sample size (i.e. more plots) increase or decrease the size of the error bars? Mathematically, how do you know this?

We will now run a t-test on the data, generating a p-value that will evaluate the **null hypothesis** that there is no significant difference in plant diversity between the disturbed wetland and the undisturbed wetland. Begin by opening the “T-test” worksheet.

* Select cell C16. This should be a blank cell next to the word “P-value”.
* Type in the following command in cell C16. After typing in the command, hit enter, and a p value should be displayed\*

**=TTEST((C2:C11),(D2:D11),2,3)**

1. What p value did you get for “Comparing Shannon Wiener Diversity Indices”? What conclusions can you make based on this? Be sure to reference the null and alternative hypotheses in your answer.

\*What does the “2” and “3” represent at the end of these excel formulas? The “2” represents a two tailed t-test. T-tests can be one-tailed or two-tailed. A two tailed test is used if there is a possibility that the difference in the data can be positive or negative (meaning, on either end of “tail” of the data’s distribution). A one-tail test is only used if the difference is in one specific direction (either positive OR negative). The number “3” in this equation is used for a two-tailed test with unequal variances.