**Math 152 Project 1** Due: September 19th in class

Great Smoky Mountains National Park is located about an hour south-east of Knoxville and is considered the ‘Salamander capital of the world’ as 31 species can be found throughout the park. The following Salamander data was collected by the Citizen Science program at the Great Smoky Mountains Institute at Tremont:

|  |  |  |
| --- | --- | --- |
| **Salamander Species** | **Location #1- Lower Dorsey Stream** | **Location #2- Pig Pen Stream** |
| Spotted Dusky Salamander | 7 | 18 |
| Imitator Salamander | 6 | 3 |
| Seal Salamander | 5 | 15 |
| Black-bellied Salamander | 7 | 11 |
| *Desmognathus spp.* Salamander | 4 | 17 |
| Blue Ridged Two-line Salamander | 1 | 31 |
| Spring Salamander | 2 | 1 |
| Northern Slimy Salamander | 0 | 1 |
| Santeetlah Salamander | 1 | 0 |
| Southern Red-backed Salamander | 2 | 0 |

Preliminary question:

1. Suppose a disaster struck the Lower Dorsey Stream and the only salamanders that survived were the imitator and the seal salamanders. Thus, we are considering 11 total individuals comprised of 6 imitator salamanders and 5 seal salamanders. Suppose you randomly remove an individual from that area and then randomly remove another individual, without replacement of the first individual.
	1. What is the probability that the two selected individuals are both seal salamanders? (reference your probability notes from 151 for help)
	2. What is the probability that the two selected individuals are both imitator salamanders? (reference your probability notes from 151 for help)

Some background information:

Biodiversity of an area is the degree of variation of life forms found within that area (usually an ecosystem in ecology). Biodiversity is a good measurement of the health and status of an area and has become a hot topic in various sciences.

Recall from chapter one that, in ecology, the two main factors taken into account when measuring diversity of an area are richness and evenness.

Richness- The number of different species present.

Evenness- A measurement of the relative abundance of the different species making up the richness of the area. That is, evenness describes the number of each species there are compared to the others. An area with the exact same number of every species would achieve maximum evenness. Similarly, an area with predominantly one species present while others are rare would have minimal evenness.

One way to measure these two parameters simultaneously is via the following function:

$$D=\sum\_{i=1}^{S}\frac{n\_{i}(n\_{i}-1)}{N(N-1)}$$

Where $n\_{i}$ is the number of individuals in species $i$, $N$ is the total number of individuals in the area and $S$ is the total number of species.

Notice that $D$ is always less than 1 and that the closer $D$ gets to zero, the higher number of equally represented species there are. That is, the closer $D$ is to 0, the more biodiversity there is in the area.

Thus, using $D$ (the combined measurement of richness and evenness), the Simpson’s Index of Diversity (SID) can be used to measure the biodiversity of an area using the following equation:

$SID=1-D$, where $D$ is defined above.

Considering the above information, the closer $SID$ is to 1, the more biodiversity there is in the area.

Objective: First see how biodiversity can be specifically calculated via Simpson’s Index of Diversity by doing a few concrete examples. Then, once you have a feel for how the $SID$ calculation works, examine how richness and evenness can individually impact $SID$. We will use MATLAB to assist in examining the second objective.

1. Calculate Simpson’s Index of Diversity ($SID$) for the data in the table for the Lower Dorsey Stream *and* for the Pig Pen Stream. Which stream area is more biodiverse with respect to salamanders just from this data?
2. Suppose a plot has 4 species with 100 individuals in each species. Calculate $SID$ for this plot.

Now that we have a handle on how $SID$ can be calculate, lets consider a scenario that will allow us to examine the individual impact of richness and evenness on $SID$. To make this a bit easier to handle, we will assume that there are an equal number of each species. (Note that this does not automatically give us a high evenness value since we need a significant number of each species in order for an area to have high evenness.)

Implication of this assumption:

Since we are interested in the individual impact of evenness and richness, what we really want to do is consider the following limits:

Before we take any limits we need use the previous implication to some simplification on the function $D$:

Now that we have a form of $D$ we can work with, you should be able to answer the following two questions by hand via calculating the limits in question using some of the techniques/properties we have previously discussed.

1. What happens to $SID$ as $n$ gets large, keeping $S$ fixed? (Here we see the influence of evenness)
2. What happens to $SID$ as $S$ gets large, keeping $n$ fixed? (Here we see the influence of richness)
3. Considering the previous two questions, which factor - evenness or richness – can bring $SID$ closer to 1? Does this make sense considering the definition of biodiversity? Explain/discuss…

Id like you to calculate 4 and 5 by hand as part of a homework assignment and then use MATLAB to verify your results two ways. The first method will be to simulate the two limits and the second way will be to graph the function for large values.

I will illustrate how to do this by completing the same problem but for a different function.

**NOTE: in each step you need to fix a number before taking the limit. That is, you need to first fix** $n$ **and then let** $S\rightarrow \infty $ **and in then you need to fix** $S$ **and let** $n\rightarrow \infty $**. To ensure everyone is doing their own work, when you fix a number let the number be the combination of your birthday.**

**For example: my birthday is June 2nd 1986. This can be written as 06/02/86. Thus, when fixing** $n$ **or** $S$**, I will let them equal** $060286=60,286$**.**









**Sequence of events:**

1. Answer questions 1-3
2. The limits implied by questions 4-5 will be assigned as homework. Make sure you obtain the correct answers so you know what exactly what MATLAB should be showing you.
3. Answer question 6 in light of your answers to questions 4 and 5.
4. Create a MATLAB code that will test the two limits and plot the results. Your code can look exactly like my example code except that you will need to change the function and the fixed values. Remember that your code name needs to have a “.m” at the end of it and that you need to set the folder correctly in order to run it.
5. Run the code. Print both the S and n limit vectors in the command window as described in the example code. Plot the S and n limit vectors as described in the example code. Note: first type “format long” into the command window before printing your values. This will show more decimal digits of your values.

**What to turn in:**

1. Questions 1-6 written up nicely and neatly. If you want to type it you can but it is not required.
2. A copy of your MATLAB code.
3. A copy of your S limit vector and n limit vector. That is, the values that MATLAB calculated as S goes to Infiniti with n being fixed and as n go to Infiniti with S being fixed. Note that you should already know what they each converge to based on your answers to questions 5 and 6. Everyone will have different limits since everyone will be fixing S and n different based on their own birthdays.
4. A copy of your two graphs. Be sure to give the graphs a title and to label the axis.
5. Make sure you name and birthday are both written on the final report.

Grading Rubric for Math 152 Project 1- Each of the following item is worth 5 points.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Item** | **Explanation** | **Points Earned** | **Item** | **Explanation** | **Points Earned** |
| **Q 1** | 2.5 points each. Partial credit for wrong formula |  | **Limit Vectors** | .5 points for each correct value |  |
| **Q 2 & 3** | Showed work- 2ptsCorrect calculation- 3 points |  | **MATLAB Code** | -2 points for each error  |  |
| **Q 4 & 5** | We did these in class so you should have the correct answers. 2.5 each |  | **Graphs** | Each axis title- 1 point eachCorrect graph title- .5 points |  |
| **Q 7** | 5 points. Based on your answers to 5 & 6 |  | **Neatness** | Clean paper, not messy, appropriate order etc |  |