



## **Teaching Statistics with Raptors in IB Biology**

by Connie Scercy Wood, 2015 CTI Fellow  
East Mecklenburg High School

This curriculum unit is recommended for:  
AP and IB Biology classes

**Keywords:** Statistics, Sample Size, T-Test, Chi-Square, Correlation, Ecology, Carolina Raptor Center

**Teaching Standards:** See [Appendix 1](#) for teaching standards addressed in this unit.

**Synopsis:** Students will learn how to use statistical functions to analyze data as they mine a database from the Carolina Raptor Center on injured birds of prey. Background information is provided on basic statistics with examples of how to use them to analyze data mined from the database. Sample size selection, mean, median, mode, standard deviation, standard error, t-tests, chi-square tests, correlation and use of Excel spreadsheets are all topics covered in this unit. It is designed for teaching the statistics needed for an AP or IB Biology class.

*I plan to teach this unit during the coming year in to 70 students in AP and IB Biology courses*

*I give permission for Charlotte Teachers Institute to publish my curriculum unit in print and online. I understand that I will be credited as the author of my work.*

## Teaching Statistics with Raptors in IB Biology

Connie Scercy Wood

I teach an IB Biology course for seniors. A main component of the curriculum is to teach the students how to collect and analyze data and use it to form conclusions. This unit will be used to teach students the critical thinking skills they need to make and interpret graphs, design controlled experiments, and conduct statistical analysis of the data they collect. It has been difficult to help them with statistics because this is one of my own weak areas. I know content very well, but I do not have a good background in statistics. The research I have done for this unit has helped me to increase my understanding of statistics and how to apply them appropriately in different situations. There are many websites and resources on statistics; however, many of them are confusing or too technical. I have included a list of websites that I found very helpful for both teachers and students who just want a simple explanation of the statistic and what it can tell us about the data. If you have a good understanding of statistics, you might skip the background section of this paper. But as I pulled this information together to use for my own clarification, I looked at it from the perspective of how I would use the particular topics with my IB biology students. Students often fail to make the connections between data and the topics of biology. I have had students do a statistical test on their data, and then fail to even mention the results in their paper. They knew they were supposed to calculate standard deviation, or conduct a t-test on the data; checked that off their list, and moved on to the next part of their paper. In this unit there are activities to help students think critically about their data and assess its meaning and validity.

The unit includes a series of practice activities that will scaffold the acquisition of the statistical and data analysis skills the IB Biology students need as they design experiments throughout the year. The activities include how to select an appropriate sample size, perform simple statistical tests such as t-tests, chi-square tests and decide when to use each type. I will have them “dissect” an example of a published experiment related to topics they are learning about—finding the variables, how factors were controlled, how data were organized and how the data were used to support the conclusion. Students will learn about the topic but also get to see examples of how a good experiment should be carried out and how decisions were made about the analysis of the data.

Students will use a very rich database of information collected by the Carolina Raptor Center (CRC) on the injured birds of prey accumulated over 30 years. (See information under Teacher Resources for contacting CRC to obtain this database). Students will apply what they have learned to mine this database and analyze some of the data contained in it. As part of the IB Biology curriculum, the teacher chooses one topic, called an Option, to teach in greater depth. I will be teaching Ecology, which includes a section on human

impact. Raptors of all types are impacted by humans through illegal hunting, habitat destruction, poisoning, and injuries from cars and power lines. The Carolina Raptor Center has been treating injured birds of prey at its current location in Latta Plantation Nature Preserve since 1984.<sup>1</sup> The database of birds brought into CRC dates back to 1975 when the birds were being housed at UNCC under the direction of Dr. Richard Brown.<sup>2</sup> When the Center was established at Latta, Dr. David Scott was hired as a full time veterinarian. He developed the RaptorMed database to keep track of the hundreds of birds that come through the center each year. This database is a window into the impact that humans make in the lives of these birds.

## **Background Information**

### Sample Size

One of the issues I have had with student-designed experiments is the sample size they select. They have a difficult time understanding how important this is for setting the stage for having enough data from which they can make valid conclusions. There will be constraints on how big the sample size can be in a classroom situation, as there are in every experiment—time, costs, and available resources will limit the sample size. However, students need to understand that this means if their sample is too small, the confidence level for the data may not be very high. If nothing can be done about it, then it still needs to be considered as a factor when analyzing the data collected from the sample.

So, how to know what a good sample size is for a particular experiment? A lot depends on the type of experiment and the methods being used to collect data.<sup>3</sup> It may not be possible to have a large sample for some types of experiments because the test subjects are rare, difficult to obtain, or there may be ethical constraints, such as having to sacrifice organisms in order to collect data. The characteristics of the sample chosen are also important. Care must be taken to select samples in an unbiased manner and to select subjects that meet the criteria for the experiment. For example, when my students do kinesis experiments with mealworms, I tell them to avoid mealworms that do not appear healthy or active. After selection of the sample, it is important to keep track of any subjects that are dropped from the experiment and to make note of exactly why they were dropped. Basing conclusions only on those that completed the experiment may be misleading if there is a significant reason that some subjects were dropped from the experiment.<sup>4</sup> What if one of those mealworms suddenly stops moving during the course of the experiment and has to be omitted? The student has to address this when analyzing the data. There are statistical tests that will detect outliers, but rather than bog the students down with too much statistical work, I would have them justify the removal of a test subject, and possibly compare the results with and without the potential outlier.

A good experiment needs to have a large enough sample size so that the results can be accurately assumed to apply to the larger population.<sup>5</sup> A formula can be used to

determine the sample size needed for obtaining results within a particular confidence level for the population mean:

$$n = ((z_{\alpha/2} \times \sigma)/E)^2$$

where  $n$  = Sample size,  $z_{\alpha/2}$  = the two-tailed critical  $z$  value (2.58 for 99% confidence and 1.96 for 95% confidence),  $\sigma$  = the population standard deviation, and  $E$  = margin of error you are willing to accept the estimated mean being from the actual mean. The greater the margin of error you are willing to accept, the smaller the sample size can be.

For example, if you want to be 95% confident that the estimated mean is within 1 unit of the actual mean, and the standard deviation is 3, then...

$$n = (1.96 \times 3/1)^2 \quad n = 34.57 \text{ or at least 35 subjects in the sample}$$

If the population standard deviation is unknown, an estimate could be obtained by getting a sample's standard deviation,  $s$ . A rough alternative is to divide the range of the measurement in the population by 4 and use that as  $s$ .<sup>6</sup> Students would have to collect some data in order to have an idea about what the range was going to be. In the case of my student's mealworm kinesis experiment, if the range of square centimeters crossed by the mealworms is 8, and we want the estimated mean to be within one square centimeter of the actual mean, then the sample size would be calculated as follows:

$$n = (1.96 \times 2/1)^2 = 15: \text{ The student would need to test at least 15 mealworms.}$$

If they wanted the estimate to be within 0.5 square centimeters of the actual mean instead of 1, then they would need to have a sample size of 62 mealworms.

If the range of squares crossed was greater—say 12—then the sample size would need to be 35 if  $E=1$ . If students choose a smaller sample size, then they must be aware that this will lower the confidence level of their data. This method would give students a basis for selecting the size of their sample and a possible source of error if they are unable to select a sample as large as required to get the confidence level they want.

If students are forced to select a less than ideal sample size due to budgetary and time constraints, they can determine the effect on the confidence interval using the formulas found at <http://www.surveysystem.com/sscalc.htm>. They may end up with a large margin of error or a lower confidence level that they will have to address when they compare their results. This will help students realize why their sample size is so important.

### *Replication, Repeated measurements and Sample size*

If students measure something during an experiment several times, I have been guilty of saying they have multiple replications. However, a replication would actually be when

the experiment is completely done again from the beginning with new test subjects. Replication of the experiment is done to ensure that the method of measurement is accurate, that there is not some problem with the design of its collection and to be sure that confounding variables are under control. It might also be done to produce more data from which finer differences in the data could be detected.<sup>7</sup> Because there is usually a time interval between the initial experiment and the replication of the experiment, there may be some variation that appears because of equipment settings, environmental conditions, or from different people conducting the experiment and observing things differently.<sup>8</sup> In the classroom, there is often not enough time to replicate the experiment even once by the same student. If multiple groups are conducting the same experiment and the data from those groups is compared, this could be considered to be multiple replications. Consistent results through replication lend validity to the findings.<sup>9</sup> The ideal way to go would be to do a statistical analysis to see if the final results are significantly different. At the least, there could be a class discussion on what replications are in science and on how consistent the results were from different groups. Why might some groups have very different results? What might not be clear in the procedures, how could that lead to inconsistent results, and how could the procedures be improved? This might help students understand the importance of writing clear, detailed procedures. Discuss variation in the data, how much variation might be normal for that data, and how we can determine that using statistics.

Repeated measurements are measurements of the same factors done within the same experiment. So if 62 mealworms are tracked in a kinesis chamber and the number of squares they cross are counted, that is 62 repeated measurements—not 62 replications.

Another problem can arise with repeated measurements when students are trying to collect enough data to be statistically valid. If students do the calculations for sample size of mealworms and discover that they need a sample of 62 mealworms, they may try to get that number by measuring the same 16 mealworms 4 times each. This is not a sample of 64; it is a sample of 16, measured 4 times. Those 16 mealworms are going to get tired and may not perform as well on the fourth trial as they did during the first trial. Or another result might be that they get the same 16 responses each time because they are collecting data from the same individuals. At best, the data will show the variation in the method of measurement. Replications should be done on new test subjects. A more accurate reflection of how mealworms behave would be to test 62 individual mealworms one time each.<sup>10</sup> Then of course, if there is time, replicate the entire experiment to verify the results from the first trial.

## Statistics

Statistics is the science of collecting, organizing, analyzing and drawing conclusions from data. Statistics can be descriptive or inferential. Descriptive statistics, as the name implies, tells you something about the data collected—for example the mean, mode, median, and standard deviation. Inferential statistics is used in forming conclusions about

the data. Using a t-test to compare data from different groups would be an example of inferential data.<sup>11</sup>

### *Mean, Mode, and Median*

Mean, median and mode are measures of central tendency which describe some central point in a data set.<sup>12</sup> The mean, or average, is the sum of the values for a variable divided by the number of values. Mode is the most frequently occurring value in a dataset. Median is the value in the center if all the data are lined up from lowest to highest. They will all be the same if there is a normal distribution to the data. I always knew what mean, median, and mode were — or at least how they were determined. However, I never really knew why median or mode would be used instead of the mean. Students most commonly use mean, but the mean can be misleading if there are outliers in the data that skew the results or if the data distribution is skewed. In that case, the median might be of more value in describing the central tendency of the data. For example, when describing income levels, if the mean income was used, it might appear that more people were making a good wage. But the data are skewed with many more people making low wages and a few making very high wages that skew the mean.<sup>13</sup> So instead you are more likely to hear the median income of a population discussed.

Mode is more useful when the data are nominal—in distinct categories—where it would be useful to know how frequently each category was showing up in the data.<sup>14</sup> For example in a taxis experiment with the mealworms, a student could look at how frequently the worms went toward particular stimuli, and to which one did they go most often. With the mode, there can be two or more categories with the same frequency. Students need to look at their data more closely and determine if the median or mode might not give them more information about their data.

### *Standard Deviation, Standard Error, and Range*

Standard deviation, standard error and range are used to describe the dispersion of data around the mean. Range is the simplest of the three to determine. It represents the difference between the highest and lowest values in the sample. I would recommend that students use range to describe the dispersion of their data rather than standard deviation if there is a small sample size. An example of using range is shown below:

Data—30, 28, 31, 30, 27, 29, 29, 33, 30, 28. The lowest value is 27 and the highest value is 33. For this data set, Range =  $33 - 27 = 6$ .

When graphing the mean of the data and adding error bars to represent the range, students would add error bars that reflect the high and low values of the range by subtracting the low value from the mean and adding the high value to the mean. The mean of the data in this example is 29.5. So the error bar would extend down 2.5 units ( $29.5 - 27$ ) to represent the lower end of the range and extend up 3.5 units ( $33 - 29.5$ ) from the mean to

represent the higher end of the range. Students have the option to do error bars this way when graphing with Excel.

Standard deviation is a measure of the variation in the data around the mean. There are two different formulas—one for the sample deviation and one for the population standard deviation. In most cases, students would use the sample standard deviation.

Sample standard deviation—

$$s = \sqrt{\frac{\sum (X - \bar{X})^2}{n - 1}}$$

where  $s$  is the standard deviation of the sample

$X$  is the individual values in the sample

$\bar{X}$  is the mean of the values in the sample

$n$  is the number of values in the samples

Population standard deviation—(used when the entire population is measured)

$$\sigma = \sqrt{\frac{\sum (X - \mu)^2}{N}}$$

--where  $\sigma$  is the population standard deviation

$X$  is the individual values in the sample

$\mu$  is the population mean

$N$  is the population size

Standard deviation can be calculated using the STATs functions on a graphing calculator or using Excel. However, students need to understand the formula and what it is telling them about the data.

When the data values are not all close to the mean, the standard deviation value is large. When the data values are close to the mean, the standard deviation will be small. Because the mean is part of the calculation for standard deviation, outliers can affect its value, making it larger. A large standard deviation is not always bad—it just says that the data are more spread out than data with a small standard deviation.<sup>15</sup> Standard deviation is helpful when comparing the amount of variation in multiple data sets. Graphing the data with error bars based on the standard deviation allows you to see how much that variation overlaps in different data sets.

If the data has a normal distribution; a bell-shaped curve, then 68% of the data would fall within one standard deviation of the mean; 95% of the data would fall within two standard deviations of the mean, and 99% would fall within three standard deviations of

the mean, as shown in Figure 1 below. This is true for a normally distributed population, and may hold true for a sample population if there is enough data collected to determine that the graph of the data would be bell-shaped.

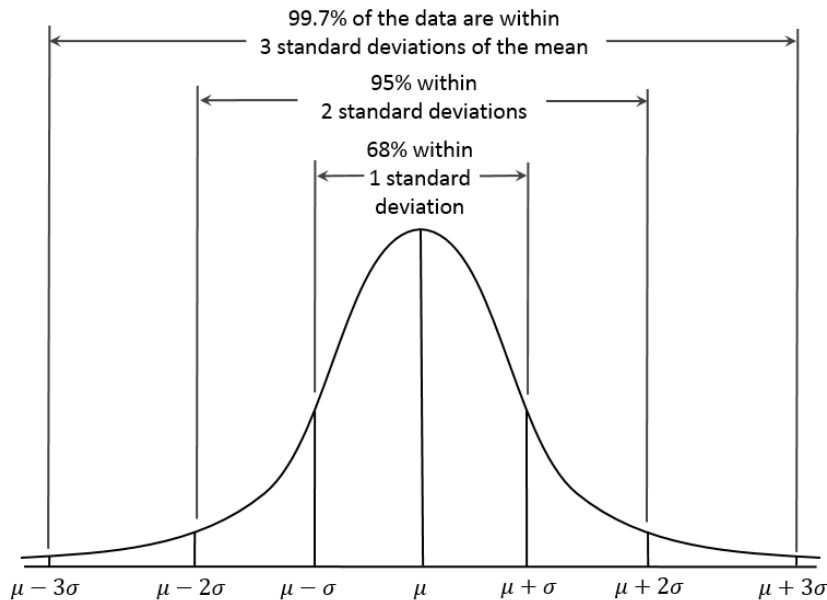


Figure 1. Normal curve. Source: [https://upload.wikimedia.org/wikipedia/commons/a/a9/Empirical\\_Rule.PNG](https://upload.wikimedia.org/wikipedia/commons/a/a9/Empirical_Rule.PNG)

Standard error of the mean (SEM) is a measure of the variation in sample means. If a number of samples are taken and sample means calculated, the sample means will probably not be the same. Each sample mean is an estimate of the population mean, but each sample will vary. The means are distributed around the true population mean just as the data are. But even if only one sample is taken, an estimate of the accuracy of the sample mean can be calculated using the standard deviation of the sample and the sample size).<sup>16</sup> The formula is as follows:

$$\text{SEM} = \frac{s}{\sqrt{n}} \quad \text{where } s = \text{sample standard deviation, } n = \text{the size of the sample}$$

The larger the sample size, the closer to the population mean the sample mean will be and the smaller the standard error. A small sample is more likely to produce a mean value different from the population mean. Error bars can be used to show standard error of the mean, but should not be interpreted to be significantly different if the bars from two different samples do not overlap. It is not the same as standard deviation and a statistical test such as a t-test should be used to determine the statistical significance of the differences.<sup>17</sup> Students could use this calculation to evaluate the mean from their sample.



*Hypothesis testing: t-tests, ANOVA, chi-square tests and correlations*

After collecting and processing their data, students need to determine if the data support their hypothesis. Are the differences they see in the data from different samples or test groups statistically significant? The type of test depends on how the experiment was set up. For an experiment where normally distributed data from two groups is being compared, a Student t-test should be used. For more than two groups, an ANOVA test should be used and if the data are in categories, then a chi-square test should be used.

There are several different types of t-tests. A one-sample t-test is used to compare a sample mean to a known population mean.<sup>18</sup> For example, you would use this test to compare the average test score in one class to the average test score in the entire district. A paired t-test, or dependent sample t-test, would be used when you are comparing data in a before and after situation, such as number of birds treated for gunshot wounds versus the number of those birds who were healthy enough to be released after treatment.<sup>19</sup>

The most commonly used t-test is the Students t-test. It is also sometimes called the independent sample t-test or the 2-sample t-test—good to know when you are searching for different calculators or functions. Whatever it’s called, it is done when the sample size is small (<30), the population standard deviation is unknown, there is continuous data with a normal distribution, and the means of the dependent variable of two groups are being compared. For example, this test would be used to compare the mean heights of pine trees on two different sites. The formulas for each type of t-test are a little different, so it is important to know which one will be required for the data you are analyzing or to select the right program if using a calculator or Excel spreadsheet. I found several different formulas that were all supposedly for the Student t-test--one of them is shown below.

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

--where  $\bar{x}$  = the sample mean,  $\mu$  = the population mean,  $s$  = the sample standard deviation and  $n$  = sample size

Or there is this one, which I find a lot more user friendly since it does not require the population mean which you probably won't have:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2 + s_2^2}{n_1 + n_2}}}$$

--where  $\bar{x}_1$  = mean of the first sample,  $\bar{x}_2$  = mean of the second sample,  $s_1$  and  $s_2$  are the standard deviations of sample one and two, and  $n_1$  and  $n_2$  are the sample sizes. The t-value is a measure of the difference in the means of the samples you are comparing relative to the variation in those samples. The greater the t-value, the more likely you are to reject the null hypothesis that there is no significant difference in the means. The closer to zero the t-value is, the more likely you are to accept the null hypothesis that there is no significant difference in the means.<sup>20</sup>

In Figure 2 below, several t-distributions are shown, each based on the sample size. The larger the sample the more the t-distribution approaches the normal distribution. The p-values are the probabilities of obtaining those t-values. As t-values approach zero, it means that the means being compared show very little difference from each other. But with a smaller sample size, as shown in the green distribution line, the higher the probability that values fall in the two tail areas, and there is a decreased chance that the t-values fall near zero difference.

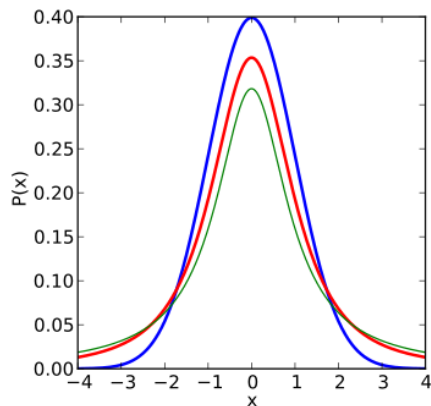


Figure 2. T-distribution graph-

[https://commons.wikimedia.org/wiki/File:T\\_distribution\\_2df\\_enhanced.svg](https://commons.wikimedia.org/wiki/File:T_distribution_2df_enhanced.svg)

Once the  $t$  value has been calculated, the degrees of freedom must be determined so that the correct t-distribution can be used. Degrees of freedom represent the number of values that can vary for a given statistic, such as the mean. So if you had 10 values, you could vary any 9 of them, but to obtain the particular mean, the last value would have to be a certain value in order to result in that mean. When using a one-sample t-test, the degrees of freedom would be the number of values in that one sample minus one. For a Student t-test, comparing the means of two samples, the degrees of freedom is  $(n_1 + n_2) - 2$ , as long as the samples are of similar sizes and have similar variances.<sup>21</sup> The larger the sample size, the more degrees of freedom, and the closer the t-distribution will be to the

normal distribution. Notice in Figure 2 that as the t-values becomes larger—either in a negative or positive direction—the probability of that t-value occurring decreases. The chance of obtaining those higher t-values is very low if the means are statistically the same as proposed by the null hypothesis. What is more likely if the t-value is high is that the means are different because the samples are different.<sup>22</sup>

With the *t* value and the degrees of freedom for the data, rather than use a distribution graph, a t-table is used to test the hypothesis. T-tables can be found online and in any good statistics book. A portion of a t-table is shown in Figure 3 below. One Tail and Two Tail refer to the choice you make to determine if the differences in your means are greater than or lesser than, or different (either way) from each other. The “tails” refer to the end section of the normal curve where only extreme values are located.

<b>t- Table</b>											
<b>One-tail α</b>	<b>0.50</b>	<b>0.25</b>	<b>0.20</b>	<b>0.15</b>	<b>0.10</b>	<b>0.05</b>	<b>0.025</b>	<b>0.01</b>	<b>0.005</b>	<b>0.001</b>	<b>0.0005</b>
<b>Two-tail α</b>	<b>1.0</b>	<b>0.50</b>	<b>0.40</b>	<b>0.30</b>	<b>0.20</b>	<b>0.10</b>	<b>0.05</b>	<b>0.02</b>	<b>0.01</b>	<b>0.002</b>	<b>0.001</b>
<b>Conf.Level:</b>	<b>0%</b>	<b>50%</b>	<b>60%</b>	<b>70%</b>	<b>80%</b>	<b>90%</b>	<b>95%</b>	<b>98%</b>	<b>99%</b>	<b>99.8%</b>	<b>99.9%</b>
<b>df</b>											
<b>1</b>	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
<b>2</b>	0.000	0.816	1.061	1.368	1.886	2.920	4.303	6.965	9.925	22.327	31.599
<b>3</b>	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
<b>4</b>	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
<b>5</b>	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
<b>6</b>	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
<b>7</b>	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
<b>8</b>	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
<b>9</b>	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
<b>10</b>	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587

Figure 3. T-Table. Source of values: <http://www.sjsu.edu/faculty/gerstman/StatPrimer/t-table.pdf>

First, establish the null and alternative hypotheses. The null hypothesis ( $H_0$ ) is that there is no significant difference in the means being compared. If the means of two samples are being compared then  $H_0: \bar{x}_A = \bar{x}_B$ . The alternative hypothesis ( $H_a$ ) could be one of the following:  $H_a: \bar{x}_A \neq \bar{x}_B$ , or  $H_a: \bar{x}_A > \bar{x}_B$ , or  $H_a: \bar{x}_A < \bar{x}_B$ , depending on what you are interested in concluding.<sup>23</sup>

The second step is to select the confidence level that you want. In most cases, this will be 95% confidence, or a 0.05 significance level. The significance level or  $\alpha$  or *p value*, is the chance that that a true null hypothesis would be falsely rejected.

After calculating the t-value, you compare that value with the table values for the degrees of freedom for the data. For example, in the t-table above, if our degrees of freedom were

7, we wanted a 95% confidence level, and the calculated t-value was 3.115, then we would reject the null hypothesis because 3.115 is greater than the table value of 2.365. This means that there is a significant difference between the means of the two samples. If however, the t-value was 0.015—less than the table value of 2.365—then we cannot reject the null hypothesis, and we can be 95% certain that there is no significant difference between the two means.<sup>24</sup>

There are also online calculators such as the one at <http://www.graphpad.com/quickcalcs/ttest1.cfm> which do the calculations and provide the results for the p-value. These allow students to enter their data and the heavy lifting of the math is done for them. Students still need to understand what the feedback from the calculations tells them about their data.

The one-way ANOVA (Analysis of Variance) test is another test that can be used in place of the t-test if there are more than two groups to be compared. The samples must have similar variances and there must be a normal distribution of the dependent variable in order to use this test. The null hypothesis is that none of the samples are significantly different. The alternative hypothesis is that at least one of the samples is significantly different from the others. Instead of a t-value, you will get an F-value which is compared to critical values in an F-Table.<sup>25</sup> A p-value is produced—if it is large, the null hypothesis is accepted. If the p-value is small, then the null hypothesis is rejected.<sup>26</sup> The ANOVA test will tell you if there are any differences between the means of three or more samples, but not which ones were different.<sup>27</sup> The math is a bit more than I would want to go into in a biology class, but if students understand what it can do and what it can tell them about their data, they can use Excel or an online ANOVA calculator to test their data. One calculator is at <http://danielsoper.com/statcalc3/calc.aspx?id=43> . To use it, students need to have the means and standard deviations of their data.

The Chi-Square test is used to compare the frequencies of observed categories or ranked data to predicted frequencies. The formula for chi-square is:

$$X^2 = \sum \frac{(O - E)^2}{E}$$

where  $O$  is the observed frequency for each category and  $E$  is the expected frequency for each category. The expected values may be some known distribution—such as a ratio of phenotypes from a genetic cross. In a case like that, the expected frequencies would vary. Or, if the hypothesis is that all of the categories should be the same, then the expected values would be the average of the observed values. In both cases, the observed and the expected values must add up to the same total. Degrees of freedom are the number of categories minus 1. Using the degrees of freedom, students compare the calculated chi-square value to the critical chi-square value for an  $\alpha$  of 0.05 in the chi-square table. If the calculated chi-square value is greater than the table value, the null hypothesis is rejected—meaning that the observed frequencies are significantly different from what

was expected. If the calculated chi-square value is less than the critical value, the null hypothesis is accepted that the observed frequencies are not significantly different than the expected frequencies.

This calculation is not difficult to do by hand and I often teach it using the frequencies of colors of M&M candies in a bag. The Biology Corner website has a good version of this: [http://www.biologycorner.com/worksheets/chi\\_square\\_candy.html](http://www.biologycorner.com/worksheets/chi_square_candy.html) . Chi-square test may be done on graphing calculators, Excel, or using online chi-square calculators. In the Strategies and Lessons section you will find “**Chi-Square Activity—Frequency of Owl Species Hit by Cars**”—an activity using the RaptorMed database from the Carolina Raptor Center. Students will filter the database to gather results on the number of owls of different species who have been hit by cars. The partial results for one species look like this:

	A	B	C	I	J	K
1	patient_id	admission_date	species	injury_cause	status	final_date
334	18680-2015-017	3/23/2015	GHOW	HBC	EOA	3/23/2015
365	18648-2015-014	3/6/2015	GHOW	HBC	D	3/10/2015
486	18527-2015-002	1/13/2015	GHOW	HBC	R	7/11/2015
570	18443-2014-083	12/7/2014	GHOW	HBC	R	12/19/2014
584	18428-2014-081	12/3/2014	GHOW	HBC	R	3/8/2015
626	18382-2014-077	11/11/2014	GHOW	HBC	R	12/14/2014
911	18101-2014-049	7/1/2014	GHOW	HBC	DOA	7/1/2014

Figure 4. Source—RaptorMed, historical database, obtained from Dr. David Scott, Carolina Raptor Center.

Just from this partial look at the data, some interesting results pop out that raise new questions about why some owls might be more prone to getting hit by cars.

A common type of experiment involves comparing two measurable variables to determine if one is related to the other. For example, is there a relationship between percent body fat and blood pressure? The data are graphed as a scatter plot of ordered pairs of the two variables for each individual sampled. If the relationship is linear, then correlation and linear regression can be performed on the data. Correlation is a test to determine if there is a relationship between the two variables, and if so, whether it is a positive or negative relationship. If the value of one variable increases proportionally to the increase in the second variable, then there is a positive correlation, as shown in Figure 5 below. If the value of one variable decreases as the other increases, there is a negative correlation. The strength of the correlation is shown by the level of scattering of the data. The tighter the data points, the stronger the correlation; the more spread in the data points, the weaker the correlation.

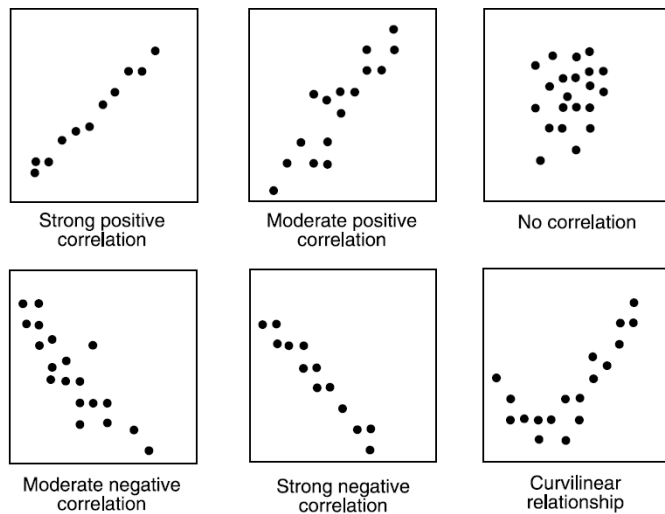


Figure 5-Correlation graphs. Source: <http://www.learnerator.com/blog/sat-review/analyzing-and-describing-scatterplots/>

This test is used when neither of the two variables has been manipulated and you are trying to determine if there is any relationship between the two. A correlation coefficient,  $r$ , can be calculated that measures the strength of the relationship for a sample population. Some sources referred to this as Pearson's correlation.<sup>28</sup>

$$r_{xy} = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{(n-1) s_x s_y}$$

where  $\bar{x}$  and  $\bar{y}$  are the sample means of X and Y and  $s_x$  and  $s_y$  are the standard deviations of X and Y, respectively.<sup>29</sup>

The coefficient  $r$  will have a value from -1 to +1, with -1 being a perfectly negative correlation or inverse relationship, and +1 being a perfectly positive correlation. A correlation coefficient of 0 indicates no correlation between the variables.

The  $r$ -value can then be compared to an  $r$ -table of critical values to determine if the correlation is significant. Degrees of freedom is equal to the number of subjects minus 2. Look up the critical value for  $r$  for an  $\alpha$  of 0.05. If the  $r$ -value is greater than the critical value, you would reject the null hypothesis that the variables are not correlated.<sup>30</sup>

It is important to note, and to stress to students, that just because you may find a correlation, that does not mean that one of the variables caused the effect in the other. There may be other factors that affect both of the variables being tested which resulted in the correlation, or the correlation may be entirely coincidental. Once a correlation is determined, it is important to consider all other possible variables before assuming causation.<sup>31</sup>

Linear regression is used when one of the variables is manipulated and the values are selected as the independent variable. An equation can be produced that describes the linear relationship of the two variables—one that describes the best-fit line through the data. This equation can be used to make predictions about the value of one of the variables if another is known. Most of my students have done this type of graphing and understand regression probably better than I do. But I do get some students who may not have grasped the concept in their math class, so need to see how it is actually used in science.<sup>32</sup>

### **Strategies and Activities**

Ideally, I will teach this unit early in the year, so that students will feel comfortable using what they learn about statistics throughout the year with other lab reports. But let's face it, statistics is not the most exciting topic in Biology. To engage my students in wanting to explore this database and dig out interesting information and relationships, I will take them to the Carolina Raptor Center in Huntersville, N.C. before we even begin talking about statistics. There they will get to see all the different kinds of birds represented in the database and learn more about them. CRC has a wonderful education program, as well as many resident birds, too damaged or imprinted to be released. The students will be able to visit the Rehabilitation facility and talk to the people and volunteers who work there and see how they work with the birds 365 days a year. I will also use this day as a service-learning opportunity. IB students are required to do community service and we will spend part of our day there doing volunteer work to help out--clearing trails, cleaning cages, and other tasks.

Having met the birds, and having seen some of the problems these birds have encountered, I hope that the students will now look at the database not as a lot of letters and numbers, but as animals that they care about.

Before beginning a lesson on sample size, I will give students the example of a kinesis experiment with mealworms and ask them to write down the number of mealworms they would use to do the experiment. We would then go through the calculations for sample size using the formula  $n = ((z_{\alpha/2} \times \sigma)/E)$  and compare their answers. Then to show how a sample size that is too small can affect the conclusions, students will use Excel and the RaptorMed database to compare results from a large sample and a small sample. See Activity 1—Selecting a Sample Size. The results of using the different sample sizes in this activity will change the confidence level from nearly 100% to 70%. Some of my students may not have used Excel before, so I will walk them through some of the basics with this activity—setting up a database, entering formulas for calculations, hiding columns, etc.

In Activity 2, students will take the larger sample from Activity 1 and use it to perform a t-test to determine if there is a significant difference in the number of window collision injuries between sharp-shinned hawks and Coopers hawks. Then in Activity 3, students will perform a chi-square test on owl data. Finally, they will look at correlation and lines of regression for the number of Red-tailed hawks hit by cars in different months in Activity 4. While doing the statistical work, I will require students to do research on the species of birds whose data we are analyzing—the physical characteristics, breeding seasons, and ecology that might be related to the impact that humans make on their lives by disrupting habitats, introducing non-native predators, illegal hunting, and obstacles we leave in their way such as fishing line and barbed wire.

Students will read a scientific article, “Modelling human impacts on the Tasmanian wedge-tailed eagle (*Aquila audax fleayi*)”<sup>33</sup> and use guided notes to dissect the article to determine how the scientists planned the experiment and analyzed the data. They will look for examples of the kinds of statistical analysis we have been learning about and work in pairs to compare their answers and discuss the data. Clear comparisons can be drawn with the dangers our own eagles face to the problems for eagles in Tasmania. A copy of the guided notes—Activity 5—is in the Handouts section.

As a final assessment, each student will write their own report based on data from RaptorMed. We will brainstorm all of the ways that the data could be looked at—geographically, by species, by injury cause, length of stay and survival success. Each student will be required to come up with a research question, sort through the data, do some type of statistical analysis appropriate to the data they mine, and present their findings to the class. I will assess their individual ability to select and perform the correct statistical calculations for their data.

Handouts

### **Activity 1-Selecting a Sample Size: Coopers Hawks and Sharp-shinned Hawks vs. Windows**

A good experiment needs to have a large enough sample size so that the results can be accurately assumed to apply to the larger population.<sup>34</sup> You will often be limited by time and materials when conducting experiments in class. However, you should be aware of how a sample that is too small can affect your data. The following is a simplified, down and dirty, equation for estimating the sample size you will need:

$$n = (1.96 \times s) / E$$

--where  $n$  is the sample size,  $s$  is the standard deviation, and  $E$  is the margin of error you are willing to accept. (If you do not know the standard deviation, you can use the range divided by 4 for a rough estimate.) The value 1.96 represents a critical value for a 95%



confidence level. We would like to be 95% confident that our sample results are representative of the whole population.

Question 1.—In what way would standard deviation affect the sample size?

---

---

Question 2—What will happen to the sample size if you decrease the margin of error that you will accept? \_\_\_\_\_

You are going to be using the RaptorMed database as the source of your samples to test some question that you have about birds that they treat. At some point you will have to decide how much of the over 30 years of data you will use—5 years, ten years. 25 years? What is enough? The database makes it easy to use it all if you need to, but right now we will use it as an example of what can happen to the significance level of your data if you choose a sample that's too small.

Problem 1: What size sample should I have in order to have a confidence level of 95% and a margin of error of +/- 1?

Problem 2: How will changing the sample size affect descriptive statistics such as mean, median, mode and standard deviation?

Procedure:

1. We want to compare the number of Coopers Hawks and Sharp-shinned Hawks per year that come into the Carolina Raptor Center with injuries from crashing into windows. Filter the database by species and injury cause by selecting Data in the menu bar and clicking on Filter. Use the drop-down menus at the top of each column to select for SSHA only and for window collisions. Also filter for Admission date—deselecting 2015, which was not a whole year's data and anything before 1985. The Center opened in its present location in 1984 and there may not have been a whole year's data that year. Hide any columns that are not needed for this activity.

2. At the bottom of the sheet, it will tell you how many records appeared for this search—that is the number of Sharp-shinned Hawks brought in after window collisions in CRC history. Count the number of records for each year. Record this data in an Excel spreadsheet.

3. Repeat for Coopers hawks, COHA.

Number of birds per year injured by window collisions

Year	COHA	SSHA
2014	8	2
2013	11	2

2012	11	1
2011	10	1
...	...	...
1985	1	1

4. Calculate the average number of hawks per year, the standard deviation, median and mode by entering the following formulas into cells at the end of your data for both species. These can be found under the “Formulas” tab in Excel, then “More Formulas”, and “Statistical”:

For Average per year: =AVERAGE( )

Standard Deviation: =STDEV.S( )

Median: =MEDIAN( )

Mode: =MODE.SGL( )

In the parenthesis enter the range of cells to do the function on. For example, =AVERAGE(B3:B32). Press ENTER and the calculation should be done.

5. Let’s say that you did not have time or resources to collect a large amount of data. How would you decide what individuals to test? Your selection should be random. Bias, even unintentional bias, can introduce error into your results. Statistical analysis is not reliable if the sampling method is not random. A random table of numbers can be used to objectively select your sample. An online version can be found at <http://stattrek.com/statistics/random-number-generator.aspx> . Use the years as your numbers and select 10 different years at random to collect counts on.

6. Create a new data table that uses the counts from the ten randomly selected years and do the same calculations on the data. Use the Confidence Interval calculator to see how the smaller sample size affects the confidence interval.

Conclusion—

Compare the results of all statistics and the confidence intervals for the two samples. How might the conclusions be different if drawn from the different sample sizes? Turn in copies of your data tables and statistics, neatly formatted and clearly labeled, with the conclusion.

## Activity 2—T-test Activity: Significant Differences in Hawk Data

Once data has been collected, averaged, and graphed, it is time to draw some conclusions. So the sample populations look different—but how different do they have to be to really be different? What is the least difference you can have and still call them different? It's hard to say unless you have an objective way to do it. That is inferential statistics—where you ask yourself, are the differences significant? Do I have the data to support my hypothesis?

One test that can be done is the t-test, which compares the means, and the variation in the data for those means of two samples, to determine if they are different. The null hypothesis is that there is no difference in the two samples. We are going to calculate the t value two ways—once by hand, and then using an online calculator. We'll be using the Coopers Hawk and Sharp-shinned Hawk data, so you already have some of the work done!

Procedure 1—The Hard Way

The formula for obtaining the t-value is  $t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2 + s_2^2}{n_1 + n_2}}}$ . Now it's not as bad as it looks.

Let's set up a data table. This data is from the previous activity:

	CoopersHawk	SharpshinHawk
Means ( $\bar{x}$ )		
Standard Dev.(s)		
( $s^2$ )		
Sample size ( $n$ )		

Solve for the parts of the equation and then put it all together:

$$\bar{x}_1 - \bar{x}_2 = \underline{\hspace{2cm}}$$

$$s_1^2 + s_2^2 = \underline{\hspace{2cm}} \quad \sqrt{\frac{s_1^2 + s_2^2}{n_1 + n_2}} = \sqrt{\hspace{1cm}} = \underline{\hspace{1cm}} \quad \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2 + s_2^2}{n_1 + n_2}}} = \underline{\hspace{1cm}} = t$$

$$n_1 + n_2 = \underline{\hspace{2cm}}$$

Once you have the t-value, all you need are the degrees of freedom and a t-table.

$$\text{Degrees of freedom} = n_1 + n_2 - 2 = \underline{\hspace{2cm}}$$

With this database, you will need a big t-table. Go to <http://www.sjsu.edu/faculty/gerstman/StatPrimer/t-table.pdf> Find your degrees of freedom and the column for a significance value of .05. Look up the critical t-value on the t-table and compare your calculated value.

Critical value for \_\_\_\_degrees of freedom and 0.05 significance level is \_\_\_\_\_

If your t-value is higher than the critical value, then you can reject the null hypothesis. If the t-value is less than the critical value, we cannot reject the null hypothesis.

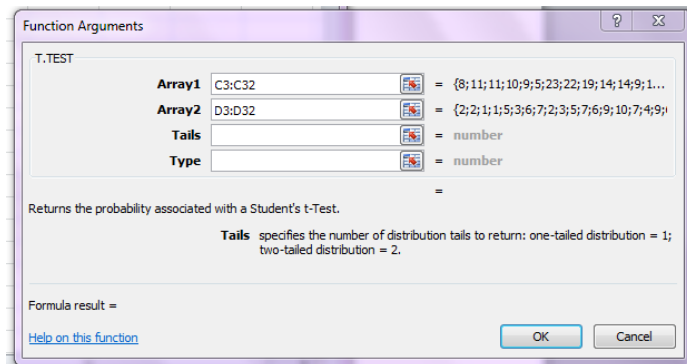
Do you accept or reject the null hypothesis?\_\_\_\_\_

What does this mean about the difference in the mean number of each species of hawks that ran into windows?\_\_\_\_\_

Procedure 2--Now the easy way...

1. You can also do this in Excel.(Note: Different versions of Excel may appear slightly different) Use the data tables for number of birds injured per year that you used in the Sample Size activity. In the menu bar, under Formulas, select More Functions...Statistical...T.TEST

2. You will see this screen--



Click in the Array 1 block and then highlight the cells with the COHA values. Click on Array 2 and then highlight the cells that contain the SSHA values. Click on Tails and enter 2. Click on Type and enter 2. Click OK in the box. The result will be the  $p$ -value, NOT the t-value. Excel

has already compared your t-value to the chart and has given you the probability that the two means are different. If it is less than .05, then we reject the null hypothesis; if greater than .05, we accept the null hypothesis.

3. The  $p$  value is \_\_\_\_\_

4. What can you conclude about the differences in the means of the Coopers Hawk and Sharp-shinned Hawk data?

### **Activity 3--Chi-Square Activity: Frequency of Owl Species Hit by Cars**

Problem: Do different species of owls get hit by cars at higher rates?

We will look at the species of owls that came into the Carolina Raptor Center from 2010 to 2014 that had been hit by cars and compare the frequencies of these incidents by species. Because there may be greater numbers of some species than others, we will compare the number hit by cars to the total number of each species that were brought to the CRC.

Procedure:

1. Filter the RaptorMed database by admission date: 2010, 2011, 2012, 2013, 2014
2. Filter by Injury Cause: HBC for hit by car.
3. Now Filter for each of the owl species that show up. (BDOW-barred owl, BNOW-barn owl, GHOW-great horned owl, and EASO-eastern screech owl)
4. Record the observed number of owls hit by cars for each species.
5. Remove the filter for Injury Cause—Select all types. Re-filter by species and record the total numbers of each owl species that came into CRC. Calculate the frequency of each species admitted by dividing the number admitted for each individual species by the total number of owls.
6. Calculate the Expected number of owls hit by cars by multiplying the total number of owls hit by cars by the frequency of each species.
7. Subtract the Expected number of owls from the Observed number.
8. Square the value obtained from step 7 for  $(\text{Observed} - \text{Expected})^2$ .
9. Divide the  $(\text{Observed} - \text{Expected})^2$  values by the Expected values, and add up the column. This total is your chi-square value.
10. Degrees of freedom is equal to the number of categories minus 1.
11. Compare the chi-square value,  $X^2$ , to the chi-square table of critical values, using the correct line for degrees of freedom and the column for a significance level of 0.05. If our chi-square value is less than the critical value, we accept the null hypothesis that there is no difference in the observed and expected values for number of owls hit by cars. If our chi-square value is greater than the table value, then we reject the null hypothesis and conclude that there is a significant difference between the observed and expected values.

Data--

Number of Owls Hit by Cars (HBC) from 2010 to 2014							
	Total Admitted	Frequency Admitted	Observed HBC	Expected HBC	Observed -Expected	(Observed -Expected) <sup>2</sup>	(O-E) <sup>2</sup> /E
BDOW	(839)	(0.42)	(71)	(47.46)	(23.54)	(554.13)	(11.68)
BNOW	(62)	(0.03)	(2)	(3.39)	(-1.39)	(1.93)	(0.57)
GHOW	(367)	(0.18)	(12)	(20.34)	(-8.34)	(69.56)	(3.42)
EASO	(730)	(0.37)	(28)	(41.81)	(-13.81)	(190.72)	(4.56)
Totals:	(1998)	(1.00)	(113)	(113)		$\chi^2 =$	(20.23)

Degrees of freedom = n-1= (3)

$$\chi^2 = \sum \frac{(O - E)^2}{E} \quad \chi^2 = 20.23$$

Critical value for a significance level of 0.05 and \_\_\_\_\_ degrees of freedom is \_\_\_\_\_

Degrees of Freedom	good fit					poor fit		
	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01
1	—	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345
4	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277
5	0.554	0.831	1.145	1.610	9.236	11.071	12.833	15.086

Source: [http://www.biologycorner.com/worksheets/chi\\_square\\_candy.html](http://www.biologycorner.com/worksheets/chi_square_candy.html)

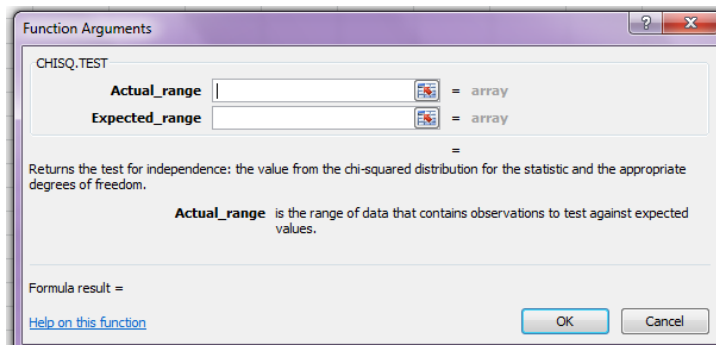
The null hypothesis is rejected / accepted ?

12. Now let's try it on Excel and see if we get the same results. Open an Excel spreadsheet and create a data table with your observed and expected values.

13. In the menu bar, under Formulas, select More Functions...Statistical...CHISQ.TEST

You will see this screen--

Click in the Actual range block and then highlight the cells with the Observed column values. Click on Expected range and then highlight the cells that contain the Expected values. Click OK in the box.



The result will be the  $p$ -value. If it is less than 0.05, then we reject the null hypothesis; if greater than 0.05, we accept the null hypothesis.

What was the  $p$ -value from Excel? \_\_\_\_\_

The null hypothesis is rejected / accepted ?

Conclusion:

What can you conclude about our research question? What are the data and statistics that support your conclusion? Do some research on these species of owls to see if you can find something about their physical characteristics, behavior, ecology, or other traits that might explain the results. Look back at the database to see what other information might be searched to help provide a possible explanation. What are some limitations of using this database? How might that affect or restrict your conclusion?

#### **Activity 4—Correlation Activity: Number of Red-tailed Hawks Hit by Cars By Month of the Year**

Problem: Is there a correlation between the time of the year and the chances for Red-tailed Hawks to be hit by cars?

We will look at the Red-tailed Hawks that came into the Carolina Raptor Center for five years, from 2010 to 2014 that had been hit by cars and compare the frequencies of these incidents by month of the year to determine if there might be seasonal variations in these types of injuries.

Procedure:

1. Under Data in the menu bar, select Filter and filter the RaptorMed database by admission date: 2010, 2011, 2012, 2013, 2014
2. Filter by Injury Cause: HBC for hit by car.
3. Filter the species by RTHA—Red-tailed Hawk. Hide other columns that you are not using.
4. To make it easier to count number of hawks by month, insert a column beside the Admission date column. In the first cell next to the first date, enter =MONTH( ) with the cell number of the first date. This formula will put the date in as just the month. Select the entire column and Fill Down the equation. You should now have a column with just months

5. Create a data table and record the observed number of hawks hit by cars by month of the year.
6. Use these data to create a scatter graph using the graphing instructions for Excel. Add appropriate titles to the graph and axes.
7. Click on the graph and under Chart Tools, select Layout and add a Trendline.
8. Right click on just the trendline created and select Format Trendline. In the Format box, at the bottom, select “Display Equation on chart”. This will give you the equation for predicting the Y value (number of birds hit by car) for X value (month). A value,  $r^2$ , is also shown on the trendline. This is the coefficient of determination. The closer it is to 1, the tighter the correlation. The closer to 0, the less likely that there is a correlation between the 2 variables.<sup>35</sup>
9. Another way to get the correlation coefficient is to use the correlation formula. Click on Formulas in the menu bar...then More Functions...Statistical...CORREL. In the dialog box, click on Array 1 and then select the month data. Click on Array 2 and select the number of birds data. The value that is produced is r—the correlation coefficient.
10. Compare the calculated r-value to the critical value in the table below. Degrees of freedom for correlation is the number of subjects minus 2. So  $12 + 12 - 2 = 22$  degrees of freedom.

Is the r-value less than or greater than the critical value?

Degrees of Freedom: $n - 2$	Critical Values: (+ and -)
1	0.997
2	0.950
3	0.878
4	0.811
5	0.754
6	0.707
7	0.666
8	0.632
9	0.602
10	0.576
11	0.555
12	0.532
13	0.514
14	0.497
15	0.482
16	0.468
17	0.456
18	0.444
19	0.433
20	0.423
21	0.413
22	0.404
23	0.396
24	0.388
25	0.381

---

If it is less than the critical value, you cannot reject the null hypothesis that the two variables are not correlated. If the r-value is higher than the critical value, then you can conclude that the variables are correlated.

Do you accept or reject the null hypothesis?

---

95% critical Values of the Sample Correlation Coefficient



Source: <https://www.boundless.com/statistics/textbooks/boundless-statistics-textbook/estimation-and-hypothesis-testing-12/hypothesis-testing-correlations-56/hypothesis-tests-with-the-pearson-correlation-277-11392/>

11. Copy and paste your data tables, graphs and  $r$ -values into a Word document and add the following:

Conclusion—

Discuss the results of the statistical analysis and how it relates to our original question.

**Activity 5--Dissection of an Article: “Modelling human impacts on the Tasmanian wedge-tailed eagle (*Aquila audax fleayi*)”<sup>36</sup>**

1. Just looking at the title, what might be the independent variable and the dependent variable?
2. Why is it important to have some idea of those variables in the title?
3. Read through the Abstract before beginning the article it will give you an overall sense of what the article is about. Write a Tweet-sized version of the abstract that tells someone what its about.

Introduction

4. What does PVA stand for? Now, what does that mean?
5. What is PVA modeling being used for in this study

Methods

6. Why was it important to describe the area of the study?
7. The cards seem stacked against this species. How many issues can you find that have led to its placement on the Endangered Species list?
8. The eagles are not alone in their troubles. What other species of predators have bitten the dust or are in danger of it?

9. What are...

Meso-predators?

Meta-populations?

10. The computer model they used to predict what might happen to the eagles in different scenarios had two parts, what did they do?:

RAMAS GIS4--

LANDIS 3.7—

11. What was the 5-step approach to solving the problem that they used? Describe the purpose of each of the steps in predicting what might happen to these birds.


12. What are the habitat requirements for these birds to build nests and reproduce?

13. Vocab Break! Did you look these up when you came across them?

Edaphic--

Coastal heath—

Sclerophyll--

Emergent trees—

Fecundity—

14. How many eggs are laid each year by a breeding pair? What is the survival rate?

15. Stochasticity???

16. “Fecundity and survival were assumed to be perfectly positively correlated.”—what might be a reason for this correlation?

17. How is eagle population growth linked to forest growth and change?

18. What were raster maps used for?

19. Describe the scenario of forest management that would have been the best for the eagles—

The worst scenario—

20. What does Figure 3. Show happening to the forested areas in the future?

21. What is EMP a measure of?

22. This is a computer model, but list as many variables as you can that they simulated changing:

23. List as many variables as you can that they controlled for in the model:

Results—

24. In Figure 4, in what year of model 3 does the eagle population reach a new carrying capacity? What is that new carrying capacity?

25. What might cause that little bump in the eagle population around year 120?

26. What is the lowest the eagle population could get to if none of the trees are harvested?

If there is harvesting (Scenario 3)?

27. Using Figure 6. What is the probability that the eagle population will reach 80 birds in each scenario?

28. In Figure 7, to what 3 factors do the eagle populations show the most sensitivity to?

Discussion—

29. What factors did the model show had the most impact on the eagle population?

39. List some of the recommendations made by the scientists:

40. What might be the saving grace for this population if there is emigration?

41. What did they have to say in the evaluation of their modeling procedure?

## **Appendix 1—Implementing Teaching Standards**

### Mathematical Requirements for IB Biology:

Carry out calculations involving means, decimals, fractions, percentages and ratios.—this is covered in all of the activities.

Represent and interpret frequency data in the form of bar charts, graphs, and histograms—students will deal with frequencies when using chi-square.

Plot graphs (with suitable scales and axes) involving two variables that show linear or non-linear relationships—students are plotting graphs for all of these activities.

Plot and interpret scatter graphs to identify a correlation between two variables, and appreciate that the existence of a correlation does not establish a causal relationship—students will do this when doing the correlation and linear regression activity.

Determine the mode and median of a set of data, calculate and analyze standard deviation—Students are calculating standard deviation for all of the activities.

Select statistical tests appropriate for the analysis of particular data and interpret the results—Students have to select which of the statistical tests are appropriate to use for the final assessment research report.<sup>37</sup>

## Appendix 2.--Teacher and Student Resources

Bekessy, Sarah A., Brendan A. Wintle, Ascelin Gordon, Julian C. Fox, Ryan Chisholm, Bill Brown, Tracey Regan, Nick Mooney, Steve M. Read, and Mark A. Burgman. "Modelling Human Impacts on the Tasmanian Wedge-tailed Eagle (*Aquila Audax Fleayi*)." *Biological Conservation* 142, no. 11 (2009): 2438-448. Accessed November 29, 2015. [www.elsevier.com/locate/biocon](http://www.elsevier.com/locate/biocon). --This is the article I used for the "article dissection" activity. It is about endangered eagles in Tasmania.

Biology For Life <http://www.biologyforlife.com/statistics.html> --Teacher Gretel von Barga has set up this amazing website for her IB Biology classes. It contains great stuff on statistics and links to calculators in addition to everything else you need to teach IB Bio.

Biology Statistics Made Simple Using Excel.  
[http://www.utdallas.edu/~serfling/3332/Biology\\_statistics\\_made\\_simple\\_using\\_Excel.pdf](http://www.utdallas.edu/~serfling/3332/Biology_statistics_made_simple_using_Excel.pdf) --a good handout on using Excel that includes many kinds of statistical tests and when to use them.

"GraphPad Statistics Guide." GraphPad Statistics Guide. 2015.  
[http://www.graphpad.com/guides/prism/6/statistics/index.htm?f\\_ratio\\_and\\_anova\\_table\\_\(one-way\\_anova\).htm](http://www.graphpad.com/guides/prism/6/statistics/index.htm?f_ratio_and_anova_table_(one-way_anova).htm). --a great online resource for understanding statistics.

McDonald, John H. "Handbook of Biological Statistics."  
<http://www.biostathandbook.com/> --very good resource on statistics with good examples and a tongue-in-cheek attitude.

Scott, David. "RaptorMed." RaptorMed. 2015. <http://raptormed.carolinaraptorcenter.org/>.  
--You can access this site to see information on current patients at the raptor center. For access to the historical database used in these activities, contact Dr. Scott at [dscott@carolinaraptorcenter.org](mailto:dscott@carolinaraptorcenter.org)

StatisticsLectures.com <http://statisticslectures.com/topics/statistics/> --a good website with tutorials and also statistics calculators.

Laerd Statistics <https://statistics.laerd.com/statistical-guides/measure-central-tendency-mean-mode-median.php>. --if you are googling for info on statistics and this site pops up, it is good. I never could see how to get to the info from their main website, it's a company that sells its services, but you can get to their topics by googling them. They have to make it clear for their clients, so I found they made things clear for me.

Statistics Solutions <http://www.statisticssolutions.com/> --another company with good clear info on statistics.

<http://www.ncsu.edu/labwrite/pl/pl-homepage.htm> NC State University Lab website to help with writing lab reports including graphs. If every student would just follow this...

[http://www.surveysystem.com/sscalc.htm?sm\\_au=iVVVqtP0SJrDR5rM](http://www.surveysystem.com/sscalc.htm?sm_au=iVVVqtP0SJrDR5rM) Calculator for determining sample size or confidence level from a sample size. Awesome!

<http://stattrek.com/> online resource for AP Statistics with tutorials, stat tables, calculators, etc.

[http://abacus.bates.edu/~ganderso/biology/resources/stats\\_flow\\_chart\\_v2014.gif](http://abacus.bates.edu/~ganderso/biology/resources/stats_flow_chart_v2014.gif) flow chart for type of statistics to use.

<https://www.wilderness.net/toolboxes/documents/vum/Basic%20statistics%20vocabulary.pdf> statistics vocab list.

Campbell AM, Heyer LJ, Paradise CJ (2014) Integrating Concepts in Biology. Trunity.net. publisher. Portsmouth, NH.

<http://www.trunity.net/IntegratingConceptsinBiology/> --new online textbook for AP Biology. Emphasis on data analysis to understand concepts in biology. Check out the sample chapters available online.

### Notes—

<sup>1</sup>"CRC History." CRC History. Accessed November 27, 2015.  
<http://carolinaraptorcenter.org/about/crc-history>.

<sup>2</sup>Ibid.

<sup>3</sup>"3.2 - Sample Size Determination." Penn State Eberly College of Science STAT 503 Design of Experiments. 2015. Accessed September 19, 2015.  
<https://onlinecourses.science.psu.edu/stat503/node/14>.

<sup>4</sup> Rumsey, Deborah J. *Statistics for Dummies*. 2nd ed. Hoboken, NJ: Wiley, 2011, p 268.

<sup>5</sup> Ibid, p267.

<sup>6</sup> Bluman, Allan G. *Elementary Statistics: A Step by Step Approach*. 5th ed. Boston: McGraw-Hill, 2004.

<sup>7</sup>"Replicates and Repeats in Designed Experiments." Minitab 17 Support. 2015. Accessed

November 1, 2015. <http://support.minitab.com/en-us/minitab/17/topic-library/modeling-statistics/doe/basics/replicates-and-repeats-in-designed-experiments/>.

<sup>8</sup>Ibid.

<sup>9</sup>"Design of Experiments." Wikipedia. October 15, 2015. Accessed November 1, 2015. [https://en.wikipedia.org/wiki/Design\\_of\\_experiments](https://en.wikipedia.org/wiki/Design_of_experiments).

<sup>10</sup> Reinhart, Alex. "Pseudoreplication: Choose Your Data Wisely¶." Pseudoreplication: Choose Your Data Wisely — Statistics Done Wrong. March 1, 2015. Accessed September 28, 2015. <http://www.statisticsonwrong.com/pseudoreplication.html>.

<sup>11</sup> "Descriptive and Inferential Statistics." Descriptive and Inferential Statistics. 2010. Accessed November 2, 2015. <http://www.statisticslectures.com/topics/descriptiveinferential/>.

<sup>12</sup> Donnelly, Robert A. *The Complete Idiot's Guide to Statistics*. 2nd ed. New York, NY: Alpha, 2007.

<sup>13</sup>"Measures of Spread." Measures of Spread. 2013. Accessed November 22, 2015. <https://statistics.laerd.com/statistical-guides/measures-of-spread-range-quartiles.php>.

<sup>14</sup>Ibid.

<sup>15</sup> Rumsey, Deborah J. *Statistics for Dummies*. 2nd ed. Hoboken, NJ: Wiley, 2011, p78.

<sup>16</sup> McDonald, John H. "Handbook of Biological Statistics." Standard Error of the Mean -. June 20, 2015. Accessed November 23, 2015. <http://www.biostathandbook.com/standarderror.html>.

<sup>17</sup> Ibid.

<sup>18</sup>"One Sample T-Test - Statistics Solutions." Statistics Solutions. 2013. Accessed November 25, 2015. <http://www.statisticssolutions.com/manova-analysis-one-sample-t-test/>.

<sup>19</sup>"Paired Sample T-Test - Statistics Solutions." Statistics Solutions. 2013. Accessed November 25, 2015. <http://www.statisticssolutions.com/manova-analysis-paired-sample-t-test/>.

- <sup>20</sup> Runkel, Patrick. "What Are T Values and P Values in Statistics?" What Are T Values and P Values in Statistics? January 27, 2015. Accessed November 25, 2015. <http://blog.minitab.com/blog/statistics-and-quality-data-analysis/what-are-t-values-and-p-values-in-statistics>.
- <sup>21</sup> Zhang, Joy Ying. "Confidence Interval and the Student's T-test ." All About Student's T-test. December 1, 2006. Accessed November 25, 2015. <http://projectile.sv.cmu.edu/research/public/talks/t-test.htm>.
- <sup>22</sup>Runkel.
- <sup>23</sup>Rumsey, p23.
- <sup>24</sup>"Paired Sample T-Test - Statistics Solutions." Statistics Solutions. 2013. Accessed November 25, 2015. <http://www.statisticssolutions.com/manova-analysis-paired-sample-t-test/>.
- <sup>25</sup>"One-way ANOVA." Laerd Statistics. 2013. Accessed November 26, 2015. <https://statistics.laerd.com/statistical-guides/one-way-anova-statistical-guide.php>.
- <sup>26</sup>"GraphPad Statistics Guide." GraphPad Statistics Guide. 2015. Accessed November 26, 2015. [http://www.graphpad.com/guides/prism/6/statistics/index.htm?f\\_ratio\\_and\\_anova\\_table\\_\(one-way\\_anova\).htm](http://www.graphpad.com/guides/prism/6/statistics/index.htm?f_ratio_and_anova_table_(one-way_anova).htm).
- <sup>27</sup>"One-way ANOVA." Laerd Statistics. 2013. Accessed November 26, 2015. <https://statistics.laerd.com/statistical-guides/one-way-anova-statistical-guide.php>.
- <sup>28</sup>"What Is the Difference between Correlation and Linear Regression?" GraphPad. 2009. Accessed November 28, 2015. <http://www.graphpad.com/support/faqid/1141/>.
- <sup>29</sup>"Correlation and Dependence." Wikipedia. November 10, 2015. Accessed November 28, 2015. [https://en.wikipedia.org/wiki/Correlation\\_and\\_dependence](https://en.wikipedia.org/wiki/Correlation_and_dependence).
- <sup>30</sup>"Hypothesis Testing with Pearson's R." Hypothesis Testing with Pearson's R. 2012. Accessed November 28, 2015. <http://statisticslectures.com/topics/hypothesispearsonr/>.
- <sup>31</sup>Bluman, p506.
- <sup>32</sup> McDonald, John H. "Handbook of Biological Statistics." Correlation and Linear Regression -. July 20, 2015. Accessed November 28, 2015. <http://www.biostathandbook.com/linearregression.html>.



- <sup>33</sup>Bekessy, Sarah A., Brendan A. Wintle, Ascelin Gordon, Julian C. Fox, Ryan Chisholm, Bill Brown, Tracey Regan, Nick Mooney, Steve M. Read, and Mark A. Burgman. "Modelling Human Impacts on the Tasmanian Wedge-tailed Eagle (Aquila Audax Fleayi)." *Biological Conservation* 142, no. 11 (2009): 2438-448. Accessed November 29, 2015. [www.elsevier.com/locate/biocon](http://www.elsevier.com/locate/biocon).
- <sup>34</sup>Rumsey, p267.
- <sup>35</sup> McDonald, John H. "Handbook of Biological Statistics." Correlation and Linear Regression -. July 20, 2015. Accessed November 28, 2015. <http://www.biostathandbook.com/linearregression.html>.
- <sup>36</sup>Bekessey, et al.
- <sup>37</sup> *Biology Guide: First Examinations 2016*. Cardiff, Wales: International Baccalaureate Organisation, 2014. Print.

## **Bibliography**

- Bekessy, Sarah A., Brendan A. Wintle, Ascelin Gordon, Julian C. Fox, Ryan Chisholm, Bill Brown, Tracey Regan, Nick Mooney, Steve M. Read, and Mark A. Burgman. "Modelling Human Impacts on the Tasmanian Wedge-tailed Eagle (Aquila Audax Fleayi)." *Biological Conservation* 142, no. 11 (2009): 2438-448. Accessed November 29, 2015. [www.elsevier.com/locate/biocon](http://www.elsevier.com/locate/biocon).
- Biology Guide: First Examinations 2016*. Cardiff, Wales: International Baccalaureate Organisation, 2014. Print.
- Bluman, Allan G. *Elementary Statistics: A Step by Step Approach*. 5th ed. Boston: McGraw-Hill, 2004.
- "CRC History." CRC History. Accessed November 27, 2015. <http://carolinaraptorcenter.org/about/crc-history>.
- Cai, Eric. "Applied Statistics Lesson of the Day: Sample Size and Replication in Experimental Design." *The Chemical Statistician*. January 6, 2014. Accessed September 28, 2015. <https://chemicalstatistician.wordpress.com/2014/01/06/applied-statistics-lesson-of-the-day-sample-size-and-replication-in-experimental-design/>.
- "Correlation and Dependence." Wikipedia. November 10, 2015. Accessed November 28,

2015. [https://en.wikipedia.org/wiki/Correlation\\_and\\_dependence](https://en.wikipedia.org/wiki/Correlation_and_dependence).
- "Descriptive and Inferential Statistics." Descriptive and Inferential Statistics. 2010. Accessed November 2, 2015. <http://www.statisticslectures.com/topics/descriptiveinferential/>.
- "Design of Experiments." Wikipedia. October 15, 2015. Accessed November 1, 2015. [https://en.wikipedia.org/wiki/Design\\_of\\_experiments](https://en.wikipedia.org/wiki/Design_of_experiments).
- Donnelly, Robert A. *The Complete Idiot's Guide to Statistics*. 2nd ed. New York, NY: Alpha, 2007.
- Gerstman, B. Bert. "T-Table." StatPrimer. July 14, 2007. Accessed November 25, 2015. <http://www.sjsu.edu/faculty/gerstman/StatPrimer/t-table.pdf>.
- "GraphPad Statistics Guide." GraphPad Statistics Guide. 2015. Accessed November 26, 2015. [http://www.graphpad.com/guides/prism/6/statistics/index.htm?f\\_ratio\\_and\\_anova\\_table\\_\(one-way\\_anova\).htm](http://www.graphpad.com/guides/prism/6/statistics/index.htm?f_ratio_and_anova_table_(one-way_anova).htm).
- "Hypothesis Testing with Pearson's R." Hypothesis Testing with Pearson's R. 2012. Accessed November 28, 2015. <http://statisticslectures.com/topics/hypothesispearsonr/>.
- "Hypothesis Tests with the Pearson Correlation - Boundless Open Textbook." Boundless Statistics. July 21, 2015. Accessed November 28, 2015. <https://www.boundless.com/statistics/textbooks/boundless-statistics-textbook/estimation-and-hypothesis-testing-12/hypothesis-testing-correlations-56/hypothesis-tests-with-the-pearson-correlation-277-11392/>.
- "Inferential Statistics Video." Films On Demand Digital Educational Videos. 2005. Accessed October 17, 2015. <http://digital.films.com/PortalViewVideo.aspx?xtid=36201>.
- Liu, Luke. "Analyzing and Describing Scatterplots | Learnerator Blog." Learnerator Blog. 2015. Accessed November 28, 2015. <http://www.learnerator.com/blog/sat-review/analyzing-and-describing-scatterplots/>.
- McDonald, John H. "Handbook of Biological Statistics." Standard Error of the Mean -. June 20, 2015. Accessed November 23, 2015. <http://www.biostathandbook.com/standarderror.html>.

- McDonald, John H. "Handbook of Biological Statistics." Correlation and Linear Regression -. July 20, 2015. Accessed November 28, 2015.  
<http://www.biostathandbook.com/linearregression.html>.
- "Measures of Central Tendency." Laerd Statistics. 2013. Accessed November 2, 2015.  
<https://statistics.laerd.com/statistical-guides/measures-central-tendency-mean-mode-median.php>.
- "Measures of Spread." Measures of Spread. 2013. Accessed November 22, 2015.  
<https://statistics.laerd.com/statistical-guides/measures-of-spread-range-quartiles.php>.
- Muskopf, Shannon. "Chi Square (X2) Modeling Using Candy." Chi Square Modeling Using Candy. Accessed November 28, 2015.  
[http://www.biologycorner.com/worksheets/chi\\_square\\_candy.html](http://www.biologycorner.com/worksheets/chi_square_candy.html).
- National Research Council (US) Committee on Guidelines for the Use of Animals in Neuroscience and Behavioral Research. Sample Size Determination. Accessed September 19, 2015.
- "Normal Distribution." Wikipedia. November 20, 2015. Accessed November 23, 2015.  
[https://upload.wikimedia.org/wikipedia/commons/a/a9/Empirical\\_Rule.PNG](https://upload.wikimedia.org/wikipedia/commons/a/a9/Empirical_Rule.PNG).
- "One Sample T-Test - Statistics Solutions." Statistics Solutions. 2013. Accessed November 25, 2015. <http://www.statisticssolutions.com/manova-analysis-one-sample-t-test/>.
- "One-way ANOVA." Laerd Statistics. 2013. Accessed November 26, 2015.  
<https://statistics.laerd.com/statistical-guides/one-way-anova-statistical-guide.php>.
- "Paired Sample T-Test - Statistics Solutions." Statistics Solutions. Accessed November 26, 2015. <http://www.statisticssolutions.com/manova-analysis-paired-sample-t-test/>.
- "Paired Sample T-Test - Statistics Solutions." Statistics Solutions. 2013. Accessed November 25, 2015. <http://www.statisticssolutions.com/manova-analysis-paired-sample-t-test/>.
- Reinhart, Alex. "Pseudoreplication: Choose Your Data Wisely¶." Pseudoreplication: Choose Your Data Wisely — Statistics Done Wrong. March 1, 2015. Accessed September 28, 2015.  
<http://www.statisticsdonewrong.com/pseudoreplication.html>.

"Replicates and Repeats in Designed Experiments." Minitab 17 Support. 2015. Accessed November 1, 2015. <http://support.minitab.com/en-us/minitab/17/topic-library/modeling-statistics/doe/basics/replicates-and-repeats-in-designed-experiments/>.

Rumsey, Deborah J. *Statistics for Dummies*. 2nd ed. Hoboken, NJ: Wiley, 2011.

Runkel, Patrick. "What Are T Values and P Values in Statistics?" What Are T Values and P Values in Statistics? January 27, 2015. Accessed November 25, 2015. <http://blog.minitab.com/blog/statistics-and-quality-data-analysis/what-are-t-values-and-p-values-in-statistics>.

"3.2 - Sample Size Determination." Penn State Eberly College of Science STAT 503 Design of Experiments. 2015. Accessed September 19, 2015. <https://onlinecourses.science.psu.edu/stat503/node/14>.

Scott, David. "RaptorMed." RaptorMed. 2015. Accessed November 28, 2015. <http://raptormed.carolinaraptorcenter.org/>.

"The Effects of Outliers." The Effects of Outliers. July 1, 2010. Accessed October 31, 2015. <http://www.statisticslectures.com/topics/outliereffects/#video>.

"What Is the Difference between Correlation and Linear Regression?" GraphPad. 2009. Accessed November 28, 2015. <http://www.graphpad.com/support/faqid/1141/>.

Zhang, Joy Ying. "Confidence Interval and the Student's T-test ." All About Student's T-test. December 1, 2006. Accessed November 25, 2015. <http://projectile.sv.cmu.edu/research/public/talks/t-test.htm>.