Statistical Uncertainty

Discovery of Sound in the Sea (www.dosits.org)

Part I

Introductory Activity

Student teams (three students each) choose a piece of string and a tape measure. The team creates a data table in which to record their measurements. Each student measures the string independently, and then passes it to another teammate who measures the string and passes it on until each has measured the string three times. The team records all nine measurements in the table.

Questions

- Were the numbers for all nine measurements exactly the same? If they are not identical, why?
- How can all the measurements be used to calculate the best answer?
- Describe the process used to obtain the answer.
- Summarize the discoveries made about making measurements.

Part II

Introduction

Scientists make measurements of the natural world — both to describe the natural world and to understand the fundamental laws that govern the operation of the natural world. These measurements allow scientists to quantify conditions and describe natural phenomena. However, there can be uncertainty in scientific findings about the natural world, because of limited accuracy and precision of measurements (Measurement Errors) and the variability that occurs in nature (Natural Variability). Scientists use statistical methods to help interpret their measurements and quantify the amount of uncertainty. Statistics involves analyzing numerical data from a specific set of observations in order to make broader generalizations about the natural world. The following activity demonstrates how scientific measurement data can be used to determine natural variability and its effect on statistical uncertainty.

Objectives

- To demonstrate methods for interpreting multiple measurements of a single quantity.
- To demonstrate natural variability and its effect on interpreting statistical uncertainty.

A. Measurement Errors

Background:

In order to understand the effects of measurement errors, consider the example of determining the speed of sound in seawater at a particular temperature, salinity, and pressure. This requires precise measurements of the distance between the sound source and receiver; the time that it takes for the sound to travel from the source to receiver; and the temperature, salinity, and pressure of the seawater through which the sound travels. Although there is only one true value for the speed of sound at a given temperature, salinity, and pressure, two measurements will almost always yield different values of sound speed because of inaccuracies that always occur in making measurements. If the measurements are repeated many times, a range of values for the speed of sound at a specific temperature, salinity, and pressure will be obtained. Scientists use statistics to help interpret and understand multiple measurements of a single condition or parameter.

Materials:

List of 20 measurements of the speed of sound at the surface of a fresh water lake (salinity 0, temperature of 8°C, and depth 0 m)

Graphing paper

Calculator

Procedure:

1) Work in groups. Using the list of measurements, discuss possible methods for determining a single value for the speed of sound at the surface of a fresh water lake (salinity 0, temperature of 8°C, and depth 0 m). You may choose to use the graph paper and/or a calculator in your procedures. Record the possible methods. As a group, agree on the procedure that you will use to perform your calculation.
2) Follow the procedure your group has selected in Step 1. Record your result.
3) Discuss the result within your group.
4) Write a one-paragraph summary of your calculations, why you selected this method, and the result.

B. Natural Variability

Background:

In addition to the statistical uncertainty associated with measurement errors, natural systems have variability. In such cases, there is no one "true" value, and scientists use statistical tools like those you developed in Part II Section A (Measurement Errors) to describe the natural variability that occurs.

Materials:

Measuring tape

Graphing paper

Calculator

Procedure:

1) Work in groups. Each group suggests three natural or biological attributes in the classroom to measure, such as student heights, arm lengths, etc. Write down the three natural or biological attributes.
2) Select one of your three natural or biological quantities and report it to the class.

3) Your teacher will make a list of the proposed natural or biological quantities. As a class, vote on the attribute that you will measure (majority rules).
4) As a class, develop a data sheet/table to record your measurements. Then in the group of three formed for Part I (Introductory Activity), follow the procedure for measuring the selected natural or biological attribute, and record the measurements on a data sheet.
5) Using the procedure selected in Part II Section A (Measurement Errors), calculate the value of the selected parameter.
6) Write a one-paragraph summary of the calculations and the result.
7) Each group will share their results with the class.
8) Complete the discussion questions.

Discussion Questions

1) How did the group select the natural or biological attributes to be measured?
2) How might the group's result have changed if you had conducted 200 measurements of the natural or biological attribute? 1,000 measurements?
3) How do you think your results would have differed if you had selected a different calculation method?

Vocabulary

Accuracy: the degree of closeness of measurements of a quantity to its actual value.

Precision: the degree to which repeated measurements under unchanged conditions show the same results; reproducibility; repeatability.

Mean: the average of a set of measurements defined to be the sum of all of the measurements divided by the number of measurements.

Standard deviation: an estimate of the variability of a set of measurements about the mean value; it is calculated by computing the square root of the variance.

Standard deviation of the mean: an estimate of the variability of the mean value computed from a specific set of measurements; it is calculated by dividing the standard deviation of the measurements by the square root of the number of measurements; also often called the standard error of the mean.

Sample: a subset of the population.

Population: the total collection of individuals or items from which samples are drawn.

Distributions: the frequency of occurrence of a specific value in a set of measurements.

Student's t-test: a statistic that compares the sample means with the standard deviations of the sample means to determine whether the two sample means are statistically different.

Statistically significant: findings of an experiment or study that have a low probability of being due to chance alone.

Histogram: a representation of a distribution by means of rectangles whose widths represent class intervals and whose areas are proportional to the corresponding frequencies of occurrence.

Variance: an estimate of the variability of a set of measurements about the mean value; it is calculated by subtracting the mean from each of the measurements, squaring the differences, adding all of the squared differences together, and dividing by one less than the total number of measurements; the square root of the variance is the standard deviation.

100 Measurements of the Speed of Sound (meters per second; m/s) in Fresh Water at 8°C, 0 m depth, and Atmospheric Pressure

1439.01	1440.95	1439.72
1443.55	1443.88	1440.35
1441.11	1439.68	1440.10
1436.71	1435.38	1438.27
1438.04	1436.16	1437.51
1439.45	1436.30	1438.18
1436.41	1440.53	1437.36
1436.29	1438.60	1438.39
1436.24	1431.32	1439.81
1439.54	1441.26	1440.00
1439.59	1441.77	1443.56
1438.68	1434.99	1437.17
1440.35	1441.04	1442.76
1439.57	1437.07	1443.52
1443.86	1442.06	1433.99
1440.98	1442.44	1439.99
1439.11	1441.99	1447.33
1439.74	1445.16	1437.31
1440.77	1439.88	1442.69
1442.42	1437.87	1442.03
1437.57	1437.09	1436.32
1442.74	1438.50	1440.40
1442.46	1437.52	1439.82
1440.02	1441.96	1435.29
1441.46	1435.70	1441.25
1440.89	1440.92	1445.45
1438.64	1435.36	1438.10
1440.23	1438.34	1437.45
1434.03	1438.40	1440.85
1439.92	1441.73	1439.35
1440.26	1438.88	1436.17
1437.88	1443.44	1439.44
1443.85	1441.77	
1442.99	1447.17	

TEACHER STRATEGIES

Background Information
(http://www.dosits.org/science/advancedtopics/statisticaluncertainty/)

Scientists make measurements of conditions or parameters of the natural world — first to describe it and then to understand the fundamental laws that govern its operation. There is uncertainty, however, due to the limited accuracy and precision of the measurements and to the variability that occurs in nature. Scientists use statistical methods to help interpret their measurements and quantify the amount of uncertainty. Statistics involves analyzing numerical data from a specific set of observations in order to make broader generalizations about the natural world.

Measurement Errors
(http://www.dosits.org/science/advancedtopics/statisticaluncertainty/measurementerrors/)

In order to understand the effects of measurement errors, consider the example of determining the speed of sound in seawater at a particular temperature, salinity, and pressure. This requires precise measurements of the distance between a sound source and receiver, the time that it takes the sound to travel from the source to receiver, and the temperature, salinity, and pressure of the seawater through which the sound travels. Although there is only one true value for the speed of sound at a given temperature, salinity, and pressure, two measurements will almost always yield different values of sound speed because of inaccuracies in the measurements. If the measurements are repeated many times, a range of values for the speed of sound at a specific temperature, salinity, and pressure will be obtained.

In early experiments to determine the speed of sound, such as that carried out by Colladon and Sturm in 1826, the measurements were not very accurate. The value that they reported for the speed of sound in Lake Geneva, Switzerland, at a temperature of 8°C was 1435 m/s, which is about 4 m/s slower than the modern value for fresh water at 8°C and atmospheric pressure (1439.07 m/s). If the measurement had been repeated 100 times, the measurement errors would have led to a range of values for the speed of sound, perhaps similar to those provided to the students for this part of the activity.

The most basic statistic that scientists compute to help interpret multiple measurements is the average value of the measurements. The average or mean (indicated by the symbol m) of a set of measurements is defined to be the sum of all of the measurements (indicated by the symbols x1, x2, … xn) divided by the number of measurements (indicated by the symbol n):

$$m = \frac{(X_1 + X_2 + \ldots + X_n)}{n}$$

The mean value of the 100 measurements provided to the students is 1439.69 m/s, which is very close to the “true” value of 1439.07 m/s.

The next statistic that scientists compute to help interpret multiple measurements is a measure of the variability in the measurements. The standard deviation, which is usually indicated by the Greek letter σ (pronounced "sigma") is a measure of the spread of the distribution about the mean. It is calculated by subtracting the mean from each of the measurements to obtain the differences between the measurements and the mean. One then computes the squares of the differences, adds all of the squared differences together, and divides by one less than the total number of measurements n. This calculation gives an estimate of the variance of the distribution when one has n measurements. The square root of the variance is the standard deviation:

$$\sigma = \sqrt{\frac{(x_1 - m)^2 + (x_2 - m)^2 + ... + (x_n - m)^2}{(n-1)}}$$

The standard deviation of the 100 measurements is 2.85m/s.

The mean value computed from a specific set of measurements is only an estimate of the "true" mean. Scientists want to know the statistical uncertainty of this estimate. How accurate is the estimate of the sound speed in fresh water at 8°C and atmospheric pressure computed from the 100 measurements in the simulated example, for example? The uncertainty in the estimate of the mean value is just the standard deviation of the measurements (σ) divided by the square root of the number of measurements (n). The standard deviation of the mean, which is written σ_m, is:

$$\sigma_m = \frac{\sigma}{\sqrt{n}}$$

The standard deviation of the mean is often also called the standard error of the mean. For the case shown above, the standard deviation of the mean is 0.285 m/s.

The standard deviation of the mean specifies the statistical uncertainty of the estimate of the mean, due to the limited accuracy and precision of the measurements. This set of simulated measurements gives the value of sound speed in fresh water at 8°C and atmospheric pressure as 1439.69 ± 0.29 m/s. This means that the true value of the mean is likely to fall within a range of values 0.29 m/s above or below the calculated mean of 1439.69 m/s.

Modern experiments in which sound speed is measured in the laboratory are much more accurate than the experiment that Colladon and Sturm conducted in Lake Geneva, but the measurements are still not perfect. The modern "true" value is also uncertain and is most properly written as 1439.07 ± 0.05 m/s.

Natural Variability
(http://www.dosits.org/science/advancedtopics/statisticaluncertainty/naturalvariability/)

In addition to the statistical uncertainty associated with measurement errors, biological and other natural systems have an uncertainty associated with the range of natural variability. In such cases, there is no one "true" value, and scientists use statistical tools such as the mean and standard deviation to describe the natural variability that occurs and answer questions about it.

When examining the effects of an underwater sound source on marine animals, for example, scientists might ask questions such as:

- Are there more or fewer animals near a sound source on average when it is transmitting than when it is not?
- Do the animals behave differently on average when the source is transmitting than when it is off?

The most accurate way to answer these questions would be to measure the number of animals around the source each time it transmits or to study the reaction of every single animal when the source is transmitting. However, this approach is not realistic. It would be impossible to follow a source every time it transmits and observe the number and behavior of all animals near it. Scientists must instead draw conclusions in such situations based on a limited number of observations.

In order to understand the process for comparing two measurements in systems with natural variability given limited observations, one example of a natural or biological quantity that the students might choose to measure in the classroom is the height of boys versus the height of girls. The DOSITS website goes through the statistical calculations for a sample of 573 boys and 576 girls in the United States.

Approximate Time Required:

One class period.

Prior Preparation:

Gather all materials and review the activity and the DOSITS web pages for familiarity. For the Introductory Activity, make sure all strings are the same length and that students are measuring length to the finest level available with the tape measure.

Instructional Strategies:

Begin a class discussion with the inquiry questions provided (approximately 15 minutes).

Measurement Errors

Break students into groups of three. Once students are in their groups, move among the groups and guide students to:

- Encourage participation by all three students within the group as the three ways of calculating one number are developed.
- Encourage students to creatively brainstorm ways of calculating one number that may be different from statistical analyses such as mean, median, and mode.

Natural Variability

Break students into groups of three. Once students are in their groups, move among the groups and guide students to:

- Write down three natural or biological quantities and the procedure for measuring them. Sample ideas include height of students, length or width of students' feet, or length or width of plant leaves (from plants in the classroom or a plant you bring in).

Allow each group to present a compelling argument for why their natural or biological quantity should be selected. After all presentations are complete, conduct a class vote to determine which quantity the class will measure. Lead the class in developing a data sheet to measure the quantity and then record the measurements.

Separate students back into their groups and guide them to calculate a statistical parameter to summarize the measured data and answer the discussion questions.

Answers to Discussion Questions:

1. How did your group determine your calculation methods?

Answers will vary based on students' discussions. All reasonable answers are acceptable, as long as an explanation is provided.

2. Why did you select the calculation method that you did?

Answers will vary based on students' discussions. All reasonable answers are acceptable, as long as an explanation is provided.

3. How did your calculation method compare with others that were presented by your classmates?

Answers will vary based on students' discussions. All reasonable answers are acceptable, as long as an explanation is provided.

4. If you could select only one calculation to do, what would it be and why?

Answers will vary based on students' discussions. All reasonable answers are acceptable, as long as an explanation is provided.

5. How do you think your result would have changed if you had been given 200 measurements? 1,000 measurements?

With more samples, the one answer should be closer to the "true" answer with less variability (e.g., smaller standard deviation or standard error).

6. How did your group select the natural or biological quantities to be measured?

Answers will vary based on students' discussions. All reasonable answers are acceptable, as long as an explanation is provided.

7. How do you think your results would have differed if you had selected a different quantity?

Answers will vary based on students' discussions. All reasonable answers are acceptable, as long as an explanation is provided.

8. How do you think your results would have differed if you had selected a different calculation method?

Answers will vary based on students' discussions. All reasonable answers are acceptable, as long as an explanation is provided.

9. How do you think your results would have differed if you had been able to measure more of the natural or biological quantity that was selected?

With more samples, the result should have less variability (e.g., smaller standard deviation or standard error).

Extensions:

Expand to addresses outliers (use another set of numbers with a really obvious outlier)

Standards Addressed:

This activity is a mathematical activity in statistics with a solid connection to sciences where dealing with natural data sets. The standards listed are where this activity would intersect with those standards as part of the instruction in those standards.

Common Core Math Standards:

HSS.ID.A.2 - Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation) of two or more different data sets.

HSS.IC.A.1

Understand statistics as a process for making inferences about population parameters based on a random sample from that population.

HSS.IC.B.4

Use data from a sample survey to estimate a population mean or proportion; develop a margin of error through the use of simulation models for random sampling.

Next Generation Science Standards:

The data set here is a good representation of any natural data set and so this activity would be a good accompaniment to many high school science standards, particularly as expressed by the practices and the Cross Cutting Concepts in NGSS. The standards listed below are the ones with the most straightforward fits.

HS.PS4 - Waves and Electromagnetic Radiation

NGSS Practices:

• Analyzing and interpreting data

• Using mathematics and computational thinking

NGSS Cross Cutting Concepts:

• Cause and effect

• System and system models

• Scale, proportion and quantity -patterns