# —9—

# Inferential Statistics and Data Interpretation

## Introduction

Once you have completed various descriptive statistics and organized your raw data into graphical representations, as described in Chapters 7 and 8, your teacher may want you to dive further into inferential statistical analysis of that data. This chapter is organized into two parts. In the first part, I briefly introduce inferential statistical tests that are commonly used in STEM research. The description of these tests is meant only to provide an overview and to help you determine which tests might apply to your research. You will need to do additional research and decide which tests to conduct, how to perform the test, and how to interpret the results. A good resource to use is Handbook of Biological Statistics (McDonald 2009), which you can download for free (*www.lulu.com/product/5507346*). Although written for biology students, the book is easy to understand and apply to any STEM field.

In the second part of this chapter, I provide questions you should be asking yourself about the results of your experiment to interpret your data. These questions will help you as you develop conclusions about your data and determine whether or not the data support your hypothesis.

---

### Learning Objectives

During the course of the chapter you should

1. familiarize yourself with the various inferential statistics,

2. determine which inferential statistics apply to your research data, and

3. determine whether or not differences in your data are statistically significant.

---

## Key Terms

**ANOVA (analysis of variance):** An inferential statistic used to determine whether the value of a single variable differs significantly among three or more levels of a factor.

**Chi-square:** An inferential statistic used with qualitative data to determine if differences between frequency distributions are statistically significant.

**Correlation:** An inferential statistical test used to determine whether there is a statistically significant connection, or a relationship, between two variables.

**Hypothesis testing:** The method for testing claims made about populations; also called *test of significance.*

**Inferential statistics:** Mathematical calculations performed to determine whether the differences between groups are due to chance or are a result of the treatment.

**Statistically significant:** When differences between groups are mathematically determined to be due to the change of the independent variable rather than to luck or chance.

***t*-test:** An inferential statistical test used to determine whether there is a real (statistically significant) difference between the means of two samples.

# Introduction to Inferential Statistical Tests

Inferential statistics are mathematical calculations performed to determine whether the differences between groups are due to chance or are a result of the treatment (Cothron, Giese, and Rezba 2006). For example, you may have calculated means for each of your experimental and control groups, and you have noticed differences, but now you need to know if those differences are significant. The term *statistical significance* is used when the mathematical differences between groups are more likely due to the change of the independent variable than to luck or chance. Raw data alone do not indicate whether the differences between your groups are due to the treatment that you introduced in your experiment (Statistics 2011). Inferential statistics are used to mathematically determine the significance of the raw data.

You will be deciding which statistics might be useful in your experiment and then researching that method in more detail. Use statistical software (see the list on p. 98 in Chapter 7) or online statistical calculators to help you perform the statistical calculations.

### Hypothesis Testing

Hypothesis testing is the "method for testing claims made about populations; also called test of significance" (Triola 2001, p. 809). Hypothesis testing consists of testing an alternative hypothesis against a null hypothesis. A null hypothesis proposes that there is no statistical significance in a set of observations. It makes the assumption that any difference in the data is due to chance, not to the treatment you introduced. Essentially, the null hypothesis is the statement you are trying to nullify, or reject, in your research. What you *really*

*think* is the cause of the phenomenon is stated in an alternative hypothesis. It is the statistical hypothesis statement that predicts a difference between two variables and is what the experiment was set up to establish. Therefore, the hypothesis that you wrote for your proposal is actually called the *alternative hypothesis.* The alternative hypothesis is only accepted if the null hypothesis is rejected. The types of statistical tests used in hypothesis testing may include those mentioned here or others, and they depend on many things, such as sample size and type of data.

In other words, you perform inferential statistics to test the significance of your data to either support or reject the null hypothesis. If the inferential statistics determine there are significant differences between the groups you are testing, then you have evidence that the null hypothesis is not correct. This is the statistical way to show that the groups are *unusually* different and that it is unlikely that the differences are merely a matter of chance (Index 2011). You therefore would accept your alternative hypothesis.

> ### Free Online Statistical Calculators
>
> - GraphPad Software: *http://graphpad.com/quickcalcs*
>
> - Laerd Statistics: *http://statistics.laerd.com*
>
> - EasyCalculation.com: *http://easycalculation.com/statistics/statistics.php*
>
> - Stat Trek: *http://stattrek.com/Tables/StatTables.aspx*
>
> - Index of online stats calculators: *www.physics.csbsju.edu/stats/Index.html*
>
> - Interactive Statistic Calculators: *www.people.ku.edu/~preacher/calc.htm*

## *t*-tests

A *t*-test, also known as the student *t*-test, is a statistical test used to determine whether there is a statistically significant difference between the means of two samples. In your experiment, *t*-tests might be used to calculate whether or not differences seen between the control and each experimental group are a factor of your treatment or simply a result of chance (Gonzalez-Espada 2007). For example, if your experimental design included taking height measurements, you would enter the measurements for the control group and compare them to each experimental group. You would calculate the test several times as shown below.

| 1st *t*-test | Control | Experimental group 1 |
|---|---|---|
| 2nd *t*-test | Control | Experimental group 2 |

You would conduct a *t*-test of the control group versus each experimental group in your experiment. Each *t*-test would allow you to determine whether

there was any statistically significant difference between the control and that particular experimental group. If so, you could reject the null hypothesis and accept the alternative hypothesis. Using *t*-tests to compare control to experimental groups would allow you to make statements such as, "The mean heights from experimental group 1 were significantly higher than the control group." However, you would not be able to make any statistical comparisons between the experimental groups. For example, you could *not* state, "The mean heights from experimental group 1 are significantly different from the means of experimental group 2." To compare among experimental groups, you need additional statistical methods, such as adjusting for multiple comparisons or using an ANOVA if you have three or more treatments.

To use the *t*-test, each group should have at least 10 values. For example, values could be the total height that each of the 10 plant stems measured at the end of the experiment. Or the values might be the measurements that came from 10 individual trials conducted with the same conditions. Another type of *t*-test, called the paired *t*-test, is used when the two samples being compared are not randomly selected but are related, as when the same sample is measured before and after a treatment has been applied.

When you search on the internet for more information on the types of *t*-tests and how to calculate and interpret *t*-test results, use terms such as *student t-test, unpaired t-test, independent t-test*, or *paired t-test*. There are plenty of online statistic calculators that can be used for this test, such as *http://graphpad.com/quickcalcs/ttest1.cfm.*

## Analysis of Variance (ANOVA)

The analysis of variance, or the ANOVA, is a statistical test used to determine whether the value of a single dependent variable differs significantly among three or more levels of an independent variable (Creech 2011). There are several types of ANOVA tests, including the one-way ANOVA, two-way ANOVA, and nested ANOVA (McDonald 2009). In context of STEM research, there are several research designs that might use an ANOVA. The one-way ANOVA works well for a single factor with three or more levels and multiple observations at each level (Experiment Resources 2011). An experiment that includes a research design that changed the quantity or quality of the independent variable can be analyzed using an ANOVA. The one-way ANOVA is an extension of the independent two-sample *t*-test. For example, the ANOVA test could be used to compare mean distances that a four-kilogram projectile traveled for three levels of tension on a catapult. In this example, the various levels of the independent variable are the catapult tension settings and the single dependent variable is the measurement of how far the projectile

traveled. With an ANOVA test, if the results show a significant difference between the groups, this indicates only that at least two of the groups are different from each other. To determine which groups are different from which, post-hoc *t*-tests are performed.

ANOVA calculations can be done in Excel. There are numerous online tutorials that can help you perform the test in Excel, such as *www.stattutorials.com/ EXCEL/EXCEL_ANOVA.html.* There are also online calculators that can perform the ANOVA statistical test, such as *www.physics.csbsju.edu/stats/anova.html.*

## Chi-Square

Chi-square is a test for "qualitative data used to determine if differences between frequency distributions are statistically significant" (Cothron, Gieze, and Rezba 2006, p. 255). If you collected *categorical data* for your experiment, you might want to use this test. (Some examples of categorical data would be viable/nonviable; male/female: simple smooth-edge leaf/simple toothed leaf/ lobed simple leaf/compound leaf/evergreen leaf; or blond hair/brown hair/ red hair/black hair.) The chi-square test will determine whether the differences in your data are statistically different from what would normally be expected.

There are two types of chi-square tests: the chi-square test of goodness-of-fit and the chi-square test of independence (McDonald 2009; Preacher 2001). The goodness-of-fit test is used to compare frequencies of your dependent variable data to known outcomes (as in genetics or when theoretical outcomes are known). For example, if your data include the number of males and females in a normal population, you would expect the ratio to be 1:1; this test will determine if the differences in your data are statistically different from the expected 50% female and 50% male ratio.

In the chi-square test of independence, you are comparing frequencies of your dependent variable data for each of your independent variables or experimental groups. These frequencies are often put into contingency tables. For example, in the river otter example used in Chapter 2, the categories of active/nonactive behavior (dependent variable) are compared for the three temperature groups (independent variable). Therefore, the contingency table for this experiment would be a 3 × 2 table.

| | Active River Otter Behavior | Nonactive River Otter Behavior |
|---|---|---|
| Below-average temperature | | |
| Average (expected) temperature | | |
| Above-average temperature | | |

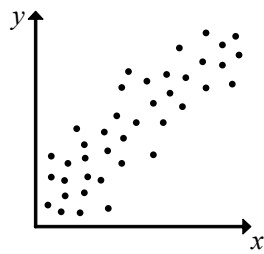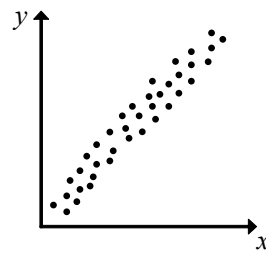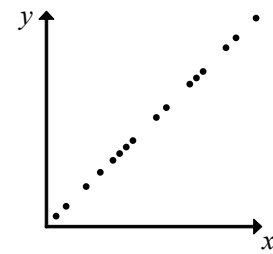You would organize the frequency data (total number of tallied behavior) into the 3 × 2 cells, which could then be entered into an online calculator, such as *http://people.ku.edu/~preacher/chisq/chisq.htm.*
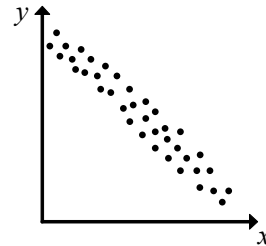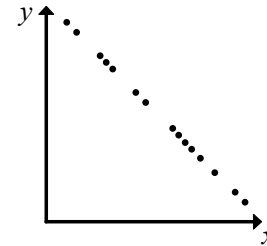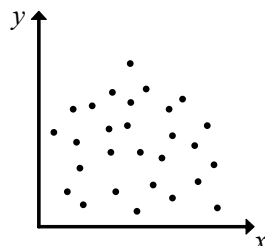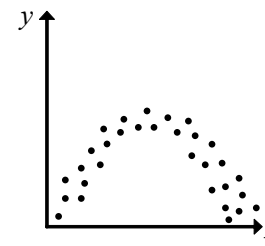
## Correlation

Correlation is a statistical technique used to determine whether there is a statistically significant relationship between two variables (Triola 2001). The results of this test will help you determine whether the relationship is significant, the strength of the relationship (weak or strong), and the direction of the relationship. Before calculating a correlation test, it is best to explore the data visually using a scatter plot.

Constructing a scatter plot will help you determine whether or not a correlation exists between two variables. The visual provided by a scatterplot allows you to see whether variables move in the same or opposite directions when they change. If the variables change in the same direction, the correlation is called a *positive correlation* and if they change in opposite directions, the correlation is called a *negative correlation*. If your scatter plot looks similar to the positive or negative correlation examples in Figure 9.1, then you might calculate a correlation statistic to determine how statistically strong the correlation is.

When you search on the internet for more information on how to calculate and interpret correlation statistics, use the terms *linear correlation, Pearson product moment correlation coefficient, linear correlation coefficient,* and *linear regression*. For an online calculator, I suggest using the Pearson Correlation Coefficient Calculator at *http://statistics.laerd.com.* Note that Excel both makes scatter plots and performs the correlation test. Search for online tutorials that provide step-by-step instructions on how to enter values and interpret the results, such as *www.internet4classrooms.com/excel_scatter.htm.*

As you organize both your descriptive and inferential statistics into tables and graphs, you should be thinking about the big picture of your research. Data interpretation is the next step you will take.

## Figure 9.1

### Explanation of Data in Scatter plots

#### I. Positive Correlations Between *x* and *y*



(a) weak positive correlation between *x* and *y*

(b) strong positive correlation between *x* and *y*

(c) perfect positive correlation between *x* and *y*

#### II. Negative Correlations Between *x* and *y*



(d) weak negative correlation between *x* and *y*

(e) strong negative correlation between *x* and *y*

(f) perfect negative correlation between *x* and *y*

#### III. No Correlation



(g) No correlation between *x* and *y*

#### IV. No Linear Correlation



(h) No linear correlation between *x* and *y*

# Data Interpretation

As you organize your raw data into tables and graphs and calculate any statistics in your laboratory notebook, you should also be evaluating the data and writing down your ideas next to the tables and graphs. You will use these evaluations to write the analysis and conclusions portion of your research paper or poster. Throughout the statistical analysis process, look at the data and keep asking yourself the following three questions.

1. What is true about my data? What additional questions come from the data?

2. How do the data describe the relationship between the two variables?

3. Do the data support the hypothesis (yes, no, or partially)?

## 1. What is true about my data? What additional questions come from the data?

With fresh eyes, look at the data and ask yourself questions. Below is a list of general questions to get you started. But your questions should be specific to *your* data and experiment. The answers to the questions will give you issues to address in both the results section and the analysis and conclusions section of your paper. Copy these questions and write your answers to them in your laboratory notebook.

- Why did certain groups perform better than others?

- Which group had the most drastic changes? Why?

- Why did one group do so much better than the others?

- How does the control group compare to the experimental groups?

- What *strange possible reasons* could explain the results?

- What trends or patterns are noticeable? Why did these occur?

- Are there any surprising outcomes of the experiment? Why did they occur?

- Are there some data that do not lie within the normal distribution of the majority of data? Why might this be? What could account for this?

- How might the procedure have influenced the results? (This question is huge!)

- How did any irregularities affect the results?

Use Student Handout #5, Interpreting Graphical and Statistical Data: A Peer Editing Exercise, pp. 142–144. Then, after you and your two peer editors have put together a list of questions, go find the answers. You will probably have to do more background research, and then refer to this research in the documentation of your paper/poster. There are two purposes of this peer editing exercise. The first is to get an outside perspective on your data. A fresh pair of eyes may help you see something you hadn't noticed before. The second is to provide you with practice in interpreting graphs and statistics from other research studies.

When writing the analysis and conclusions section, you will need to describe the relationship between an independent variable and the dependent variable and then explain how you know of this relationship. It all hangs on your hypothesis. What is the relationship between a and b? Describe the relationship and explain how you know the relationship exists.

## 2. How do the data describe the relationship between the two variables?

An important question to ask yourself is, *Did the change (independent variable) I made cause the effect that was measured (dependent variable)?* If the data support your hypothesis—but you believe it is not because of the independent variable—this distinction will need to be made.

You will answer these questions in the last few paragraphs of your analysis and conclusions section. You must either confirm or deny a relationship between the two variables and then describe the statistical data that support your final conclusion.

## 3. Do the data support the hypothesis (yes, no, or partially)?

Review the data you organized, graphed, calculated, and analyzed to determine whether or not your hypothesis is supported. Your answer will most likely be a simple yes or no (although it's possible that the hypothesis is partially supported). Don't write that the hypothesis was *proved* or *disproved;* instead, say that the hypothesis was *supported* or *not supported.* This distinction is important. The word *prove* is much too strong for a single study on a topic. STEM researchers prefer that a collection of research studies be carried out on a topic before saying that anything has been "proved."

It might be that all your data except one piece support the hypothesis. You may choose to make a general statement of support of your hypothesis, but in the results section in the final paper as well as the analysis and conclusions section, you will need to address exactly which data support the hypothesis and which data do not. (See Chapter 11 for details of what goes into the results section and the analysis and conclusions section.)

## What if the results do NOT support your hypothesis?

Search for explanations for the results. It is likely that you will have a general sense of why you got the results you did. Either something happened during the experiment to influence the results, or maybe you now know something new about your topic and realize that the experimental design was faulty. Either way, any concrete statement about the relationship between the two variables becomes impractical.

Whatever the case, do not change your hypothesis to match the data or your data to match your hypothesis. Instead, look for possible explanations of why the results may have occurred the way they did. An important question to ask is, *Does the lack of support for my hypothesis mean that I can say that there is **not** a relationship between the independent variable and the dependent variable?* If this is a possibility, it would be valuable to mention this conclusion in the analysis and conclusions section of your paper. You could suggest another research project, with a different hypothesis, to better determine the lack of relationship between the two variables.

On the other hand, it may be that your hypothesis was not supported because other variables that were not expected influenced the results. Review your laboratory notebook. Maybe an assumption you made became an extraneous variable that you did not account for. Did any procedures you performed during the experiment adversely affect the outcome? Conduct more background research. Did you miss something that would have helped you design a better experiment or collect data more efficiently? Your conclusion and analysis section will include suggestions for future studies. You may want to suggest alternative methods or analysis for a similar research project that would retest the same hypothesis.

*Remember:* Having data that do not support a hypothesis is just as valuable as a supported hypothesis. An unsupported hypothesis may highlight either a possible nonrelationship between variables or an issue with the research design. *What defines a project's success is not necessarily that the hypothesis is supported but rather what it is that the researcher learns from the research process, something that he or she shares with others in the final paper or poster.*

# Chapter Applications

It is important that you research various statistical tests and determine which ones to use on your data. Calculate statistics and write the results in your laboratory notebook. For each statistical test, answer the three questions on page 140 to help analyze your data.

At this point in the handbook, we detour briefly from "science talk" to "documentation"—that is, the rules you need to follow to make sure you do not plagiarize. Chapter 10 explains how to properly give credit to the resources you used in your background research and that you will be referencing in your STEM research paper, poster, and/or presentation.

# References

Cothron, J. H., R. N. Giese, and R. J. Rezba. 2006. *Science experiments and projects for students: Student version of students and research*. Dubuque, IA: Kendall/Hunt.

Creech, S. 2011. Statistically significant consulting. Retrieved May 28, 2011, from *www.statisticallysignificantconsulting.com*.

Experiment Resources. 2011. Statistics tutorial. Retrieved May 21, 2011, from *www.experiment-resources.com/statistics-tutorial.html*.

Gonzalez-Espada, W. 2007. Using simple statistics to ensure science-fair success. *Science Scope* 8 (30): 48–50.

Index of online stats calculators. 2011. Retrieved May 30, 2011, from the College of Saint Benedict and Saint John's University's physics website: *ww.physics.csbsju.edu/stats/Index.html*.

McDonald, J. H. 2009. *Handbook of biological statistics*. Baltimore: Sparky House Publishing.

Preacher, K. J. 2001. Calculation for the chi-square test: An interactive calculation tool for chi-square tests of goodness of fit and independence [Computer software]. Available from *http://quantpsy.org*.

Statistics in science fair projects. 2011. Retrieved May 20, 2011, from *www.scitechfestival.com/pdf/statisticsdocument.pdf*.

Triola, M. F. 2001. *Elementary statistics*. New York: Addison-Wesley.