**What is a hypothesis?**

Do we only use hypotheses in science, or are they actually a part of our everyday life? Consider this scenario:

You have studied and feel completely prepared for an exam that is scheduled for tomorrow. What time do you go to bed to feel rested for the next day? What are the constraints and variables that you can control that might play a role in a successful outcome of the exam? Is this a hypothesis? Why do you think this is or is not a hypothesis? Let’s look at another situation from daily life. Remember, a hypothesis is a **testable prediction based on previous observation**.

You would like to participate in the campus 5K run in a few months. Your friend runs regularly and you would like to run with her, but you are not in quite as good shape. What are some steps that you could take to improve your fitness to run with your friend?

If your fitness improves, is it enough to meet your goal? What could you do to continue this? Would you change your strategy if your fitness doesn’t improve? Can you think of a testable hypothesis that might allow you to make an educated decision concerning what steps you might take to prepare for the 5K?

A testable hypothesis must also be **evaluated**. How might you evaluate your hypothesis? In science, hypothesis testing is a bit more structured. The goal is to quantify the likelihood that the actual results are the same as the predicted ones. Using statistics is our way of quantifying the hypothesis testing. In statistical hypothesis testing, we try to estimate the P value, which is the probability of obtaining the observed results assuming the null hypothesis is true. If the observed results are unlikely under the null hypothesis, you would reject the null hypothesis.

**Null and alternative hypotheses**

The null hypothesis is a statement that you want to test. In general, the null hypothesis is that things are the same as each other, or the same as a theoretical expectation. For example, if you measure the size of the feet of male and female chickens, the null hypothesis could be that the average foot size in male chickens is the same as the average foot size in female chickens. If you count the number of male and female chickens born to a set of hens, the null hypothesis could be that the ratio of males to females is equal to a theoretical expectation of a 1:1 ratio.

**The alternative hypothesis is that things are different from each other, or different from a theoretical expectation.** For example, one alternative hypothesis would be that male chickens have a different average foot size than female chickens; another would be that the sex ratio is different from 1:1.

Usually, the null hypothesis is boring and the alternative hypothesis is interesting. For example, let’s say you feed chocolate to a bunch of chickens, then look at the sex ratio in their offspring. If you get more females than males, it would be a tremendously exciting discovery: it would be a fundamental discovery about the mechanism of sex determination, female chickens are more valuable than male chickens in egg-laying breeds, and you’d be able to publish your result in Science or Nature. Lots of people have spent a lot of time and money trying to change the sex ratio in chickens, and if you’re successful, you’ll be rich and famous. But if the chocolate doesn’t change the sex ratio, it would be an extremely boring result, and you’d have a hard time getting it published in the Eastern Delaware Journal of Chickenology. It’s therefore tempting to look for patterns in your data that support the exciting alternative hypothesis.

For example, you might look at 48 offspring of chocolate-fed chickens and see 31 females and only 17 males. This looks promising, but before you get all happy and start buying formal wear for the Nobel Prize ceremony, you need to ask

“What’s the probability of getting a deviation from the null expectation that large, just by chance, if the boring null hypothesis is really true?”

Only when that probability is low can you reject the null hypothesis.

The goal of hypothesis testing is to discover the likelihood that the result might be a result of random variation (in other words, just coincidence.) For example, let’s say you are testing your recent observation. You’ve fed chocolate to a bunch of female chickens (in birds, unlike mammals, the female parent determines the sex of the offspring), and get one of these three results:

• You get 25 female chicks and 23 male chicks. Anyone would look at those numbers and conclude that they could easily result from chance; there would be no reason to reject the null hypothesis of a 1:1 ratio of females to males. Probably can send back the formal clothes and avoid the Nobel party.

• You get 47 females and 1 male. Most people would look at those numbers and see that they would be extremely unlikely to happen due to luck, if the null hypothesis were true; you would reject the null hypothesis and conclude that chocolate really changed the sex ratio.

• You get 31 females and 17 males. Now what do you conclude? There are definitely more females than males, but is it really so unlikely to occur due to chance that you can reject the null hypothesis? To answer that, you need more than common sense; you need to calculate probability of getting a deviation that large by chance.

The significance level (also known as the “critical value” or “alpha”) you should use depends on the costs of different kinds of errors. With a significance level of 0.05, you have a 5% chance of rejecting the null hypothesis, even though it is true. That means that you would have a 1 in 20 chance of concluding falsely that your hypothesis was not supported. You must choose your significance level before you collect the data. Most scientists use alpha=0.05 by default, because it is generally accepted. If you choose to use a different significance level than the conventional 0.05, people will be skeptical; you must be able to justify your choice. It is very likely that you will use the conventional significance level of 0.05 for all of the experiments you will do in the core labs. However, you should understand why some scientists use other levels of significance so you can evaluate their conclusions effectively. For example, when reporting some medical experiments, the threshold of significance might be 0.01 to gain a higher level of plausibility of the stated conclusions. In other medical studies, 0.10 is acceptable, because scientists in that field do not want to reject the null hypothesis too soon; others use alpha =0.1 because the data they are using already has a lot of noise (variation) in it.

In the past, scientists reported the results of a statistical test as “P<0.05”, “P<0.01”, “P>0.10”, etc. Now, since almost all computer statistics programs give the exact P value resulting from a statistical test, such as P=0.029, that’s what you should report in your work.

You will conclude that the results are either significant or they’re not significant; they either reject the null hypothesis (if P is less than your pre-determined significance level) or don’t reject the null hypothesis (if P is greater than your significance level).

# Replicates and controls

**Replicates and Controls**  
When we perform experiments, we want to have some confidence that the results we observe are not due to random chance, and that the observations are due to the variable we wanted to test rather than some other factor. A well-designed experiment takes these into account.

REPLICATES are simply duplicates of the same treatment. Going back to our chocolate feeding experiment, you looked at 48 offspring of chocolate-fed chickens and saw 31 females and 17 males. This is one replicate of the experiment. You can increase your confidence in your results by repeating the experiment. Imagine you did the experiment a second time, and out of 50 offspring, 33 were females and 17 were males. You do it a third time, and for 45 offspring, 30 are females and 15 are males. The three replicates of the experiment show you that your first experimental results were not unusual, but typical.

A CONTROL GROUP is one that you use to determine what is normal for the study organism. In this example, a good control group would be chickens that are not fed chocolate. Usually the control group is what you will compare the results of your test group(s) to.

# Step by step guide to Experimental Design

You will use a systematic, step-by-step approach to decide how to design your experiment and analyze data. Working through the steps below should be part of your weekly lab routine. Using this approach will aid in your understanding of the pre-set labs and for labs in which you will create and analyze your own experimental studies.

1. Collect background knowledge and observations. What background information is vital to your study? What did you need to know before you could proceed with asking your question?
2. Specify the biological question you are asking. What is it that you want to know?
3. Put the question in the form of a biological null hypothesis and alternate hypothesis. The biological hypothesis refers to your expectation of what will happen in the physical world due to a biological process or mechanism. It is a generalized statement explaining how a biological mechanism of interest causes an independent variable to affect a dependent variable.
4. Put the question in the form of a statistical null hypothesis and alternate hypothesis. The statistical hypotheses refer to your expectation of what the data will show without stating what biological processes caused those data to be different. This may be expressed as a comparison of averages between or among treatment groups, or in comparison to a hypothesized value.
5. Determine which variables are relevant to the question. Specify exactly what variables play a part in the question you are asking.
6. For each variable, determine and explain what kind of variable each one is. (Many more details on variables are in the upcoming section.) Measurement, categorical, other? Independent or dependent?
7. Design an experiment that controls or randomizes the confounding variables. Sketch out a framework for your collected data. What will it look like in a spreadsheet? This should include everything you need as a data collection table. How will you describe your summarized data? Choose and describe the best statistical test to use.
8. Do the experiment. Carry out your plan, being careful to stay on track by reviewing your key objectives at regular intervals.
9. Apply the statistical test you have chosen and interpret the results. Don’t just look at the number and blindly assign a conclusion to your work. Think through the results of your experiment. How do they compare in the bigger picture?
10. Communicate your results effectively, usually with a graph or table. There are many different ways in which we will ask you to communicate your work, such as lab reports and presentations. The key is that this communication synthesizes all of the information gathered, from background to interpreting your results, and presents it clearly.

<https://youtu.be/s-fVRJyEvS0>

# Kinds of variables (measurement, ratio, & categorical)

One of the first steps in deciding which statistical test to use is determining what kinds of variables you have. When you know what the relevant variables are, what kind of variables they are, and what the null and alternative hypotheses are, it’s usually pretty easy to figure out which test you should use. We classify variables into three types: measurement variables, which are expressed as numbers (such as 3.7 mm); categorical variables, which are expressed as names (such as “female”); and ranked variables, which are expressed as positions (such as “third”). Other names for these variable types and other ways of classifying variables are frequently used in statistics references, so be aware of this.

The statistical analysis of similar experiments, with similar null and alternative hypotheses, can be completely differently depending on which variable types are involved.

#### **Measurement variables**

Measurement variables are, as the name implies, things that are measurable. An individual observation of a measurement variable is always a number. Examples include length, weight, pH, and bone density. Other names for them include “numeric” or "quantitative” variables.

Some authors divide measurement variables into two types. One type is continuous variables, such as length of an isopod’s antenna, which in theory have an infinite number of possible values. The other is discrete (or meristic) variables, which only have whole number values; these are things you count, such as the number of spines on an isopod’s antenna. The mathematical theories underlying statistical tests involving measurement variables assume that the variables are continuous. Luckily, these statistical tests work well on discrete measurement variables, so you usually don’t need to worry about the difference between continuous and discrete measurement variables. The only exception would be if you have a very small number of possible values of a discrete variable, in which case you might want to treat it as a categorical variable instead.

Be careful when you count something, as it is sometimes should be considered a categorical variable and sometimes a measurement variable. For example, the number of bacteria colonies on a plate is a measurement variable; you count the number of colonies, and there are 87 colonies on one plate, 92 on another plate, etc. Each plate would have one data point, the number of colonies; that’s a number, so it’s a measurement variable. However, if the plate has red and white bacteria colonies and you count the number of each, it is a categorical variable. Now, each colony is a separate data point with one of two values of the variable, “red” or “white”; because that’s a word, not a number, it’s a categorical variable. In this case, you might summarize the categorical data with a number (the percentage of colonies that are red), but the underlying data are still categorical.

#### **Ratios**

Sometimes you can simplify your statistical analysis by taking the ratio of two measurement variables. For example, if you want to know whether male isopods have bigger heads, relative to body size, than female isopods, you could take the ratio of head width to body length for each isopod, and compare the mean ratios of males and females using a two-sample t–test. However, this assumes that the ratio is the same for different body sizes. We know that’s not true for humans—the head size/body size ratio in babies is freakishly large, compared to adults—so you should look at the regression of head width on body length and make sure the regression line goes pretty close to the origin, as a straight regression line through the origin means the ratios stay the same for different values of the X variable. If the regression line doesn’t go near the origin, it would be better to keep the two variables separate instead of calculating a ratio, and compare the regression line of head width on body length in males to that in females using an analysis of covariance.

#### **Categorical Variables**

Categorical variables classify observations into discrete groups. Examples of categorical variables include color (with possible values of while, pink, or red), genotype (values are AA, Aa, or aa), or ankle condition (values are normal, sprained, torn ligament, or broken). A good rule of thumb is that an individual observation of a categorical variable can be expressed as a word, not a number.

If you have just two values of what would normally be a measurement variable, it’s categorical instead: think of it as “present” vs. “absent” or “low” vs. “high.” Categorical variables are often used to divide individuals up into simple groups, so that other variables may be compared among the categories. In the comparison of head width in male vs. female isopods, the isopods are classified by sex, a categorical variable, and the measurement variable head width is compared between the sexes.

Categorical variables are also called nominal, discrete, qualitative, or attribute variables.

Categorical variables are often summarized as proportions or percentages. For example, if you count the number of male and female A. vulgare in a sample from Newark and a sample from Baltimore, you might say that 52.3% of the isopods in Newark and 62.1% of the isopods in Baltimore are female. These percentages may look like a measurement variable, but they really represent a categorical variable, sex. You determined the value of the categorical variable (male or female) on 65 isopods from Newark, of which 34 were female and 31 were male. You might plot 52.3% on a graph as a simple way of summarizing data, but you should use the 34 female and 31 male numbers in all statistical tests.

It may help to understand the difference between measurement and categorical variables if you imagine recording each observation in a lab notebook. If you are measuring head widths of isopods, an individual observation might be “3.41 mm.” That is clearly a measurement variable. An individual observation of sex might be “female,” which clearly is a categorical variable. Even if you don’t record the sex of each isopod individually, but just counted the number of males and females and wrote those two numbers down, the underlying variable is a series of observations of “male” and “female.”

#### **Word of warning about Categorizing**

It is possible to convert a measurement variable to a categorical variable, dividing individuals up into a two or more classes based on ranges of the variable. For example, if you are studying the relationship between levels of HDL (the “good cholesterol”) and blood pressure, you could measure the HDL level, then divide people into two groups, “low HDL” (less than 40 mg/dl) and “normal HDL” (40 or more mg/dl) and compare the mean blood pressures of the two groups, using a nice simple two-sample t–test. Converting measurement variables to categorical variables is common in epidemiology, psychology, and some other fields. However, there are several problems with categorizing measurement variables. One problem is that you’d be discarding a lot of information; in our blood pressure example, you’d be lumping together everyone with HDL from 0 to 39 mg/dl into one group. This reduces your statistical power, decreasing your chances of finding a relationship between the two variables if there really is one. Another problem is that it would be easy to consciously or subconsciously choose the dividing line (“cut point”) between low and normal HDL that gave an “interesting” result. For example, if you did the experiment thinking that low HDL caused high blood pressure, and a couple of people with HDL between 40 and 45 happened to have high blood pressure, you might put the dividing line between low and normal at 45 mg/dl. This would be cheating, because it would increase the chance of getting a “significant” difference if there really isn’t one.

Like measurement variables, if there are a very small number of possible values for a ranked variable, it would be better to treat it as a categorical variable. For example, if you let a honeybee sting people on one arm and a yellow jacket sting them on the other arm, then ask them “Was the honeybee sting the most painful or the second most painful?”, you are asking them for the rank of each sting. But you should treat the data as a nominal variable, one which has three values (“honeybee is worse” or “yellowjacket is worse” or “subject is so mad at your stupid, painful experiment that they refuse to answer”).

# Describing your data (Mean and standard deviation)

### Mean and standard deviation ([Excel template here](https://wakeforest.instructure.com/courses/27736/files/1155063/download?wrap=1) [download](https://wakeforest.instructure.com/courses/27736/files/1155063/download?download_frd=1) )

Descriptive statistics capture important quantitative features of frequency distributions. There are many ways to describe your summarized data (for example, mean, median, mode), but for the work that you are doing in this lab, you will be describing your data using an arithmetic mean.

The arithmetic **mean** (x) is the sum of all the observations divided by the number of observations. It is the most common statistic that describes data that is symmetrically distributed in a frequency graph. When someone says “the mean” or “the average,” this is what they are talking about.

It is sensitive to extreme values, which makes it not work well for data that are highly skewed. For example, imagine that you are measuring the heights of trees in two areas of equal size. Plot A is in a mature forest plot. Plot B experienced a disturbance by fire event that killed all but 2 very large trees a few years ago. In the time since, new seedlings have sprouted resulting in scores of small trees all about the same height.

A group of trees

Description automatically generated with medium confidence

It is entirely possible that these two plots could have similar means, so that descriptive statistic does not provide enough information to compare the plots.

To better describe this data, you must describe the variability of those means as well. As with the means, there are many ways of describing this (think: standard deviation, standard error, variance. Collectively, these are called measures of dispersion.) For the work you are doing in the core labs, you will primarily be using standard deviation.

**Standard deviation** (SD) is a measure of the spread of a distribution in the same units as the distribution. It measures how far from the mean observations typically are. If the standard deviation is large, most observations are far from the mean. Conversely, if it is small, most measurements lie close to the mean.

From our example above, we can see that Plot A has a smaller SD and than plot B has. What does this tell us about the distribution of tree size in each plot?

#### **Normal distribution**

#### The data in Plot A has a distribution that resembles a bell curve, or normal distribution.

#### The standard deviation relates to the normal distribution. Sixty eight percent (68%) of a distribution lies within ±1 standard deviation of the mean. 95% of a distribution lies within ±2 standard deviations of the mean. And, 99.7% of a distribution lies within ±3 standard deviations of the mean. When graphing your summarized data, you should include mean and error bars representing one standard deviation.

# Analyzing your data with statistics

Place the summary table above the tabs?

[T-Tests: A Matched Pair Made in Heaven: Crash Course Statistics #27 (Links to an external site.)](https://www.youtube.com/watch?v=AGh66ZPpOSQ) [[](https://www.youtube.com/watch?v=AGh66ZPpOSQ)](https://www.youtube.com/watch?v=AGh66ZPpOSQ)

Two-sample t-tests are used to compare the means from two groups of data; specifically, the mean of a control group to the mean of an experimental group. Use the two-sample t–test when you have one categorical variable and one measurement variable, and you want to compare the mean values of the measurement variable. The categorical variable must have only two values, such as “male” and “female” or “treated” and “untreated.” To do this test, you use all the data for each group (not just the means alone).

This version of the t-test should only be used when you are comparing data collected from two independent groups. This mean that they were collected from completely different groups, i.e., one group of flowers gets normal water (control) while a completely separate group of plants get water and fertilizer (experimental).

Null hypothesis (H0)

The statistical null hypothesis is that the means of the measurement variable are equal for the two categories, i.e., there is no statistically significant difference between the control group’s mean and the experimental group’s mean.

H0: xC = xE

Alternative hypotheses (HA)

**One tailed**

The mean of one group is statistically significantly greater or less than that of the other group. (Here you must predict the directionality of the difference.)

HA: xC > xE or HA: xC < xE (depending on the direction you choose)

**Two tailed**

The means between the two groups are statistically significantly different. (Here you are not predicting the directionality of difference, just that the two means are statistically different.) HA: xC ≠ xE

Example

Is the average height of students in 2 sections of biology lecture significantly different?

Section A has 16 students with the following height in cm: 175, 177, 157, 160, 161, 170, 168,  
168, 169, 174, 178, 183, 187, 152, 182, 181.

Section B has 12 students with the following height in cm: 173, 170, 152, 160, 180, 170, 160,  
161, 187, 152, 182, 181.

#### Remember from our earlier discussion of statistical significance and P-values, you will be determining whether there is a significant difference between the two groups by seeing whether the calculated P-value is greater than 0.05. What is it in this case?

Reporting your results

When reporting the results of a two-sample t-test, you need to include the resulting t-statistic, the degrees of freedom (df), and the corresponding P-value. Your statement might look like this, “The mean of the control group was not statistically significantly different than that of the experimental group (t-stat = 0.57, df = 26, P = 0.574).”

Paired T-Test

[T-Tests: A Matched Pair Made in Heaven: Crash Course Statistics #27 (Links to an external site.)](https://www.youtube.com/watch?v=AGh66ZPpOSQ) [[](https://www.youtube.com/watch?v=AGh66ZPpOSQ)](https://www.youtube.com/watch?v=AGh66ZPpOSQ)

Paired t-test begins at time 6:54.

Use a paired t test when you have multiple pairs of observations of one group. It tests whether the mean difference in the pairs is different from 0. A paired t-test compares two means that come from the same group, but the group is measured twice. To do this test, you use all the data for each group (not just the means alone). The first time you measure the group is your control, then you do something to the group, and re-measure the same groups after the treatment (which is your experimental group), i.e., you measure a group of flowers the first time (control), then you give this same group of flowers water and fertilizer, and re-measure them (experimental).

Null hypothesis (H0): The mean of the group is not statistically significantly different between the first time you measure it, and the second time (after you added your treatment).

H0: x̄Time 1 = x̄Time 2

Alternative hypotheses (HA), One tailed: The mean of the group at one time point is statistically significantly greater or less than that of the mean at the other time point. (Here you are predicting the directionality of the difference.)

HA: x̄Time 1 > x̄Time 2 -or- HA: x̄Time 1 < x̄Time 2 (depending on the direction you choose)

Alternative hypotheses (HA), Two tailed: The means at the two time points are statistically significantly different. (Here you are not predicting the directionality of difference, just that the two means are statistically different.)

HA: x̄Time 1 ≠ x̄Time 2

Example

Is the average systolic blood pressure (the top BP number) of students different before and after an exam?

Here are the data (systolic BP in mm of mercury): Systolic Blood Pressure Student # Before Exam After Exam 1 120 114 2 121 125 3 125 120 4 110 111 5 124 120 6 150 111 7 130 121 8 131 110 9 148 121 10 129 111 Remember from our earlier discussion of statistical significance and P-values, you will be determining whether there is a significant difference between the two groups by seeing whether the calculated P-value is greater than 0.05.

Reporting Results When reporting the results of a paired t-test, you need to include the resulting t-statistic, the degrees of freedom, and the corresponding P-value. Your statement might look like this, “The mean BP of the group after the exam was significantly greater than its mean before the exam (t-stat = 2.62, df = 9, P = 0.028).”

Analysis of variance (ANOVA) is an approach that is used to simultaneously test whether the means of multiple (more than two) groups are equal. It works by assessing whether individuals chosen from different groups are, on average, more different than individuals chosen from the same group. It is used with one or more categorical variable.

One-way ANOVA

Used with one measurement variable and one categorical variable, a one-way ANOVA should be used when you want to compare the means of more than two independent groups. You make multiple observations of the measurement variable for each value of the categorical variable. One of the means will be your control group (xc ), and the others will be the means of your experimental groups (x1, x2, x3,…,xk). To do this test, you should use the raw data for each group (do not compare the actual means).

If your ANOVA tells you that at least one of the means is different from the other, you will need to perform additional post hoc tests to determine where a significant difference exists. You may choose to perform multiple comparisons using a two-sample t-test involving all possible pair-wise combinations of groups. However, this is generally not a recommended practice because it increases the possibility of detecting a statistical significance when one does not truly exist. An alternative is to use the Tukey-Kramer test. (The Tukey–Kramer test is also part of the template.)

In the Tukey–Kramer method, the minimum significant difference (MSD) is calculated for each pair of means. It depends on the sample size in each group, the average variation within the groups, and the total number of groups. For a balanced design, all of the MSDs will be the same; for an unbalanced design, pairs of groups with smaller sample sizes will have bigger MSDs. If the observed difference between a pair of means is greater than the MSD, the pair of means is significantly different.

#### Example

Is the average height of students in all 4 sections of biology lab significantly different?

Here are some example data that we could compare using ANOVA. From these we could determine whether the average height was significantly different between Sections A-D of the class. We would need to make pair-wise comparison after the initial ANOVA to determine which sections are different.

***Table: Height of Students in 4 Lab Sections (cm)***

|  |  |  |  |
| --- | --- | --- | --- |
| Section A | Section B | Section C | Section D |
| 175 | 173 | 157 | 190 |
| 177 | 170 | 166 | 185 |
| 157 | 152 | 176 | 197 |
| 160 | 160 | 180 | 190 |
| 161 | 180 | 160 | 190 |
| 170 | 170 | 165 | 179 |
| 168 | 160 | 175 | 185 |
| 168 | 161 | 155 | 192 |
| 169 | 187 | 159 | 180 |
| 174 | 152 |  |  |
| 178 | 182 |  |  |
| 183 | 181 |  |  |
| 187 |  |  |  |
| 152 |  |  |  |

Remember from our earlier discussion of statistical significance and P-values, you will be determining whether there is a significant difference between the groups by seeing whether the calculated P-value is greater than 0.05.

#### Reporting results

When reporting the results of a one-way ANOVA, you need to include the P-value. Your statement might look like this, “There was significant difference (P=0.000064) in the average height of the four lab sections. Further tests indicate that section D was significantly taller than all the other sections. No significant difference was indicated between other sections.”

The usual way to graph the results of a one-way ANOVA is with a bar graph. The heights of the bars indicate the means and there’s usually some kind of error bar. In this case, you should use SD.

Linear regression finds the line that best fits the data points. There are actually a number of different definitions of “best fit,” and therefore a number of different methods of linear regression that fit somewhat different lines. By far the most common is “ordinary least-squares regression”; when someone just says “least-squares regression” or “linear regression” or “regression,” they mean ordinary least-squares regression. This is a measure of association and can evaluate causal relationships because you selected the range of X variables for which you measured Y. Thus, you can say, with more confidence, that at least for the X variables you chose, the X variables caused the pattern in the Y variable.

Null Hypothesis H0: The slope of the best-fit line is equal to zero. (The variables are not associated. The strength of the association is close to 0. You cannot predict values of X and Y.)

Alternative hypotheses HA: There is a significant association between the X and Y variables you measured. (The slope of the best-fit line is not equal to zero.)

This is not a causal relationship, i.e., correlation does not imply causation!

#### Example

The ground cricket is known to change the rate of its call, or “trill” with ambient air temperature. The following data was collected in a lab:

|  |  |
| --- | --- |
| Chirps per second | Temperature in degree C |
| 15 | 21 |
| 20 | 24 |
| 12 | 20 |
| 21 | 25 |
| 18 | 24 |
| 16 | 23 |
| 13 | 22 |
| 14 | 22 |
| 20 | 25 |
| 23 | 26 |
| 12 | 21 |
| 15 | 22 |
| 18 | 24 |
| 20 | 24 |

Reporting Results

When reporting the results, you will need to include the resulting r2 (correlation), d.f., and P value. From the cricket example, “There was a significant association between the rate of cricket trills and air temperature measured in the lab (r2=0.903, d.f. = 12, P= 0.000000197).”

If you are interested in comparing the strength of the relationship (r2) to the strength of other relationships, you are doing a correlation and should design your experiment so that you measure X and Y on a random sample of individuals. This is a measure of association, and not an evaluation of causal relationships. Thus, you could say X and Y variables are associated, but you could not confidently say the X variable caused the pattern in the Y variable. Use correlation when you have 2 variables (independent and dependent) and you want to test the strength of the association between these variables, and you have measured these variables randomly from a population. The value of r2 will range from 0 to 1- the closer to 1 indicated a stronger relationship between the two variables.

# Data visualization

### Using Tables and Figures

Tables and figures are numbered separately, in the order they are referred to in your text. If you do not reference a table or figure, readers do not know where to look for the supporting data.

#### Tables

Tables should only contain summarized data, NEVER raw data. Each table should have a Table #, and short descriptive title, placed above the table itself. Tables should be numbered in the order they are referred to in your text. If a more extensive explanation of the table is needed, place the detailed explanation in a caption below the table.

Use table to visualize your data when the reader needs to know specific numbers. For example, suppose you did an experiment and determined the minimum and maximum amount of 6 different B–vitamins that lizards must consume each day. The relative amounts of B1 compared to B6 is not that important; it is the absolute quantity (i.e., 10 mg B1, versus 100 mg B6) that is important. These data are best presented in tables.

Clearly label the columns and rows of your table. Rows and columns must be neatly arranged.

Example

Table 5. Average height of med from the United Kingdom from 1810 - 1850.

|  |  |
| --- | --- |
| Year | Height (cm) |
| 1810 | 169.7 |
| 1820 | 169.1 |
| 1830 | 166.7 |
| 1840 | 166.5 |
| 1850 | 165.6 |

#### Graphs

Graphs are used to summarize and report numerical or statistical results. Graphs are the best way to show trends in numbers, or let a reader compare multiple numbers quickly. Tables are not as useful if we need to compare multiple numbers or look at trends across a range of conditions (time, temperature, etc.)

When creating figures for lab reports, research papers, or scientific articles, it is essential that you present numerical data properly. Failure to do so is misleading. The graph below was created in MS Excel. It summarizes the results of an experiment with descriptive statistics. The figure legend is the small box on the right side. The text below the image is the caption. It explains what the graph is showing. The caption includes a Figure #. All figures should be numbered in the order they are first referred to in the text.

When creating graphs:

* Label the x-axis (independent values) and y-axis (dependent values). DO NOT put a title above the figure.
* If possible, use very different textures or plot elements to indicate different sets of data (for example, the black vs. white bars in the figure shown below).
* If you plot means always include standard deviation error bars as well.
* Be sure your graph is legible. Do not skimp on space, or try to put too much data in a single figure. Clutter keeps readers from extracting information effectively.
* Number figures in the order you refer to them in your report. Place that number at the start of the caption.

Example

#### Photos and Diagrams

Photos and diagrams (drawings, maps, or other visual data) summarize data that are not described easily in words or cannot be presented in a graph. A map of the 2–dimensional distribution of organisms in a test site, or photos of the pattern of blue staining in control and cold–treated transgenic plants, are two examples. When using photos and diagrams, do not include every photo or diagram you have. Select just a few that will show the reader of the outcome of your experiment. Photos and diagrams should have a figure number and caption describing what is shown.

Formatting

The title should clearly communicate the topic to readers, indicating what organisms are being studied (scientific name), the biological property or system being studied, the particular stimulus, stress, or situation that is being applied to the system, and briefly what was found. A title more than 2 typed lines may be too long and probably should be shortened.

An abstract is a summary of your work that includes information from parts of each section of that work. Even though it is usually at the front of a report or available before a presentation, the abstract should be the last part of the report that you write. You do not know what you need to summarize until the rest of the report has been written. In about 200 words or less, it should summarize the study’s main objective(s), give the scientific name of the organism you studied, and state your hypothesis. It also will summarize the study’s background, the methods, major results, and conclusions. You should not include references in your abstract.

The introduction tells the reader (or listener) what topic you are addressing, presents the current state of knowledge of this topic (citing sources), and ties this information to your biological question.

* What do we already know about this particular organism, biochemical system, or experimental model?
* What questions come to mind when we think about this system? In other words, what is the biological question we are asking? (Getting more specific.)
* Why are we asking this question? What do we need to know to answer it? (Still more specific.)
* What do we expect to happen in this model system? What do we predict will happen (this is your hypothesis)?

The written version begins with a brief general introduction to the subject or problem, then moves on to more specific information that relates to your hypothesis. It will explain the underlying biological principles a reader needs to know to understand the purpose of the experiment. The last part of the introduction should clearly state the exact hypothesis that you tested. It says what the purpose of your experiment is, what specific function or phenomenon is being explored, and lays out a reasonable, testable hypothesis. Make it clear what the dependent and independent variables are. You can either say this as part of the hypothesis or explain it as part of the overall purpose of the experiment. Include the scientific name of the organism (in italics, with first letter of genus capitalized and the rest of the name in lower case), and a brief (1-2 sentences) explanation for why the model organism was good to use in your experiment. You will find it necessary to refer to the work of others as you describe your experiment. You must cite these sources as you use them in your work (using in text citations). Your textbook and lab manual are accepted sources, but you must include primary literature also. If you are unsure whether a source is an appropriate source, ask your GTA. Don’t try to find an experiment just like yours. You do not need a source to prove you “did the experiment right.” The goal of doing and reporting experiments is to widen our knowledge of the world around us. If you sent a paper out for review that is nearly identical to another scientist’s work, the editor of the journal will immediately return it with a rejection letter saying that you have not made any significant contribution to that particular field.

**Within the introduction, cite all pertinent literature using the [Last name of first author: year] citation format. All information that is not common knowledge needs to be cited.**

A methods section is a description of the procedures of the study, including an explanation of how the hypothesis was tested. Provide enough detailed information so that someone not directly involved in the project (but works in a lab setting) could repeat the experiments. Do not just recopy the lab manual.

Summarize your methods, use diagrams, and find other ways to make your reader understand how you did your studies. Try to answer as many of these questions as possible:

* Briefly, how did you do your experiment? What are the essential details that someone else would need to know to repeat it?
* If appropriate, where, how were the organisms collected and maintained?
* What volume and concentration of key reagents (drugs added, volumes injected, etc.) were used in the experiment?
* If you used an unusual method to create the independent variable, describe how you did it. If you used commonly available methods, state them without explanation.
* What statistical tests were used? What groups were compared? Be careful not to add useless details. Examples of TOO much detail are “…I used a paint brush to place a red dot on each grasshopper…” or “...we graphed our data...”

The purpose of the results section is to state all results clearly, so that the reader can create an independent picture of the work. State your observations and describe the tables or figures that summarize your results and statistical test. Describe the general trends and features of the data, then refer to a specific figure or table that contains summarized numerical results to support that statement. Raw data should only be found in your lab notebook. When writing a lab report or preparing a presentation, you should summarize your data and prepare tables or other visual displays that describe the summarized data. **Do not list raw data numbers in lab reports or presentations.** Do not repeat details you already provided in the Materials and Methods. Similarly, do not try to interpret or discuss your data just yet. Try to answer as many of these questions as possible:

• What general trends did you see? Did specific trials or runs within the larger experiment come out differently? If so, how were they different?

• How can you summarize your numbers so they are easier to understand?

• If you tested more than one group (or had test and control groups) are they statistically different from one another?

**Be careful to refrain from any kind of interpretation of the data in this section. Only state what your results were, not what they might mean.**

First, you should state whether your results supported your hypothesis, providing a biological explanation for why the manipulation you made produced the results you observed. Relate your findings back to the introductory information at the beginning of the paper; this helps readers see how your experiment increases our understanding of the world around us. Compare your explanation to the results and interpretations of previously published studies. You should refer to published primary literature; it provides the supporting context for how you interpreted results. You will find it necessary to refer to the work of others as you interpret the results of your experiment. You must cite these sources as you use them in your work (using in-text citations). If you are unsure whether a source is an appropriate source, ask your GTA.

If you have evidence of an error in your experiment, say so, but do not make up possible sources of error if you do not think they could have occurred. If appropriate, describe the alternative explanations of your data. If there is another possible explanation for your data, what future experiments are needed to determine the best explanation? What is the next logical step for this study? Modify the procedure and repeat it? Try again, to increase the number of replicates with the existing methods? Test an entirely new hypothesis? Are there different methods or procedures you would use if you repeated the study again? This section also is where you relate your results back to the larger body of scientific knowledge.

By the end of the Discussion section, you should be talking in general, broad terms once again. Be careful not to get too broad though; your experiment is not the explanation for everything. Within the discussion, cite all pertinent literature using the [Last name of first author: year] citation format. **All information that is not common knowledge needs to be cited**.

The last section lists the sources that you used in your research and cited in your paper using the [Last name of first author: year] in-text citation format. Number your citations (1., 2., etc.) in the order they were used in the report and list them using the Harvard citation format. You probably used MLA, APA, or some other citation style in the past. The CSE (Council of Science Editors) format that we use does not dictate one single standard style. Instead CSE recommends what information citations should contain, then leaves the details of styling up to each journal or publisher. That means if you randomly select primary articles from 10 different life science journals, you will see 10 slightly to very different citation formats. For Biology Core Labs we use a simplified “Harvard Name-Year” format for the in-text citations, and for the Literature Cited. We require (in order) the author names, year a source was released, name of the source, and enough information to locate the resource again. Some online databases and citation generators can format citations for you and put them in a downloadable format. If you are using this option, look for the **Harvard citation format**.

Citing Electronic Materials

Be VERY careful about using electronic sources; they are not equally reliable. In general, you can safely use electronic materials obtained from official publications of government agencies (site URL usually ends with “.gov”). Scholarly research projects associated with a research institution or university are allowed but should never be the sole source of information. Information from general access web pages is NOT acceptable because the content is not peer reviewed for accuracy by subject matter experts. Wikipedia should not be used as a sources for that reason. If you decide to use electronic materials as part of your sources, the in–text citation still is the last name of the first author of the source, followed by the year of publication.

**How to cite this CANVAS site**

WFU Biology department. (2020). Biology 150 Laboratory Manual [Online]. Available at: <https://wakeforest.instructure.com/courses/27681>.

**Common Mistakes Students Make With Electronic Sources**

* Do not use just the URL or web address as a citation. The URL or web address for an article is NEVER an acceptable citation on its own. A valid citation for an electronic source has the names of the authors, name of the resource, and when and where it was accessed.
* Do not use a URL of a database page as your citation. If you use JSTOR, Medline, Web of Science, Google Scholar, or another database to find a printed source, cite that source using the appropriate format (book, book chapter, or journal article.)
* Do not use just a digital object identifier (DOI) to identify or cite a source. A DOI has a format similar to this: 10.1007/s11162-008-9088-5.

**Other Common Mistakes Students Make With Sources**

* DO NOT use extended quotes from sources. Any phrase more than 4–5 words long must be paraphrased and cited. Failing to paraphrase is a form of plagiarism. You are using another person’s words rather than your own.
* NEVER list references in the Literature Cited that are not used in the text of a report (or vice versa). Every reference in the text should be in the Literature Cited, and vice versa. Failing to list all cited sources properly is academic fraud.
* Use the correct citation format. We can overlook minor errors like a misplaced comma, but omitting important information is more serious. OMITTING ESSENTIAL INFORMATION IS TREATED THE SAME AS NOT CITING A SOURCE AT ALL.

<https://youtu.be/9lL6xDKpHxw>

Scientific Writing Made Easy: A Step- by- Step Guide to Undergraduate Writing in the Biological SciencesSheela P. Turbek,1 Taylor M. Chock,1 Kyle Donahue,1 Caroline A. Havrilla,1 Angela M. Oliverio,1,2Stephanie K. P

**Style of science communication**

As a summary to the previous item in this module, "Scientific Writing Made Easy: A Step-by-Step Guide to Undergraduate

Writing in the Biological Sciences," please note these important items.

All writing in this lab should:

* Be concise and direct,
* Have supporting literature with proper citations (No quotes) for all information that is not common knowledge,
* Use past tense when referring to your work and the work of others that you may cite,
* Use active verbs,
* Begins with a broader ideas (to put your work into context), narrows to describe your hypothesis of your study,
* Describes how your work was done in a way in which other scientists of your level could replicate it,
* Continues with a clear description of your results (tables and graphs - no raw data),
* Interprets your results with supporting literature.

An experimental design plan is a written explanation of your study before you do it. The following questions need to be addressed: What is your hypothesis? What is your rationale for your hypothesis (with supporting literature and citations)? How do you plan on testing your hypothesis? What statistical test will you use?

Each lab group will be expected to produce one of these for each experiment that it designs. Your GTA will evaluate your plan and may ask questions or help you refine your work before moving forward with the study. Your GTA will need to approve your final plan. You should also find that this experimental design plan will help to frame up some of your thoughts on the introduction, methods, and other portions of your work. It can be a less formal document, but still requires complete documentation of the study plan before it is executed.

Comparing different types of communication

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Introduction | Methods | Results | Discussion | Citations |
| Experimental design plan | Preliminary, with hypothesis | Complete | None | None | Yes |
| Lab notebook entry | Preliminary, with hypothesis, well understood  Written in bullet form | Written completely, in bullet form | Written clearly in table, hand sketched graphs  has raw data | Preliminary, in bulleted form  no statistics | Yes |
| Lab report | Written completely in paragraph form | Written completely in paragraph form | Written completely in paragraph form  NO raw data | Written completely in paragraph form | Yes |

Lab notebooks are how scientists document routine work and their progress on specific projects. You will be producing individual lab notebook entries that describe your experiments from beginning to end. Each experiment will be documented in a notebook entry with the following sections: introduction, methods, results, and discussion, along with sources and citations.

Before you begin a lab, you will write the introduction and methods of the experiment. Not only will this help to prepare you for the work to be done that week, you will also be starting the process of preparing for a more formal communication of your work at a later time, such as in a lab report. You will need to use outside sources (other than this site) to complete a good introduction. Be sure to cite these sources using the same format as for lab reports. You should outline tables for your upcoming data collection as well. During lab, you will be writing your observations during your experiment and data collected (in outlined tables). These will be your notebook results section. After completing the procedure, you will need to write a quick, preliminary description of your data (with a hand drawn graph) and a short discussion of your findings. These will not be statistical tests, but observations that you can make with what you are able to collect during lab.

A lab report is a formal written report of your study. It is our department’s equivalent to a journal article. The report should include all sections of the parts of a study. Although performing the experiment is group work, lab reports are writing individually.

<https://guides.zsr.wfu.edu/zotero>

# Revising Reports Strategy

BIO214

Build a training module on how to read and revise lab reports

Annotate a bad lab report

Discussion with peers

Our review

Here is a strategy for self-review

Here is how I can practice the strategy

Module on how to interpret data tables, graphs, quantitative figures.

# What is a Hypothesis?

## We Make Hypotheses in Everyday Life

A hypothesis simply is a **testable prediction based on previous observation**. Consider this scenario: You have studied and feel completely prepared for an exam that is scheduled for tomorrow.

* What time do you go to bed to feel rested for the next day?
* What are the constraints and variables that you can control that might play a role in a successful outcome of the exam? Is this a hypothesis?
* Why do you think this is or is not a hypothesis?

Let’s look at another situation from daily life. Remember, a hypothesis is a **testable prediction based on previous observation**. You would like to participate in the campus 5K run in a few months. Your friend runs regularly and you would like to run with her, but you are not in quite as good shape.

* What are some steps that you could take to improve your fitness to run with your friend?
* If your fitness improves, is it enough to meet your goal? What could you do to continue this?
* Would you change your strategy if your fitness doesn’t improve?
* Can you think of a testable hypothesis that might allow you to make an educated decision concerning what steps you might take to prepare for the 5K?
* A testable hypothesis must also be **evaluated**. How might you evaluate your hypothesis?

In science, **hypothesis testing** is a bit more structured. The goal is to quantify the likelihood that the actual results are the same as the predicted ones. Using statistics is our way of quantifying hypothesis testing. In statistical hypothesis testing, we try to estimate the **P value**, which is the **probability of obtaining the observed results assuming the null hypothesis is true**. If the observed results are unlikely under the null hypothesis, you would reject the null hypothesis.

## Hypotheses Are Described As Null and Alternative Hypotheses

The **null hypothesis** is a statement that you want to test. In general, **the null hypothesis is that things are the same as each other, or the same as a theoretical expectation**. For example, if you measure the size of the feet of male and female chickens, the null hypothesis could be that the average foot size in male chickens is the same as the average foot size in female chickens. If you count the number of male and female chickens born to a set of hens, the null hypothesis could be that the ratio of males to females is equal to a theoretical expectation of a 1:1 ratio.

**The alternative hypothesis is that things are different from each other, or different from a theoretical expectation.** For example, one alternative hypothesis would be that male chickens have a different average foot size than female chickens; another would be that the sex ratio is different from 1:1.

Usually, the null hypothesis is boring and the alternative hypothesis is interesting. For example, let’s say you feed chocolate to a bunch of chickens, then look at the sex ratio in their offspring. If you get more females than males, it would be a tremendously exciting discovery: it would be a fundamental discovery about the mechanism of sex determination, because, female chickens are more valuable than male chickens in egg-laying breeds, and you’d be able to publish your result in Science or Nature. Lots of people have spent a lot of time and money trying to change the sex ratio in chickens, and if you’re successful, you’ll be rich and famous. But if the chocolate doesn’t change the sex ratio, it would be an extremely boring result, and you’d have a hard time getting it published in the Eastern Delaware Journal of Chickenology. It’s therefore tempting to look for patterns in your data that support the exciting alternative hypothesis.

For example, you might look at 48 offspring of chocolate-fed chickens and see 31 females and only 17 males. This looks promising, but before you get all happy and start buying formal wear for the Nobel Prize ceremony, you need to ask: ***“What’s the probability of getting a deviation from the null expectation that large, just by chance, if the boring null hypothesis is really true?”*** Only when that probability is low can you reject the null hypothesis.

The goal of hypothesis testing is to discover the likelihood that the result might be a result of random variation (in other words, just coincidence.) For example, let’s say you are testing your recent observation. You’ve fed chocolate to a bunch of female chickens (in birds, unlike mammals, the female parent determines the sex of the offspring), and get one of these three results:

* You get 25 female chicks and 23 male chicks. Anyone would look at those numbers and conclude that they could easily result from chance; there would be no reason to reject the null hypothesis of a 1:1 ratio of females to males. Probably can send back the formal clothes and avoid the Nobel party.
* You get 47 females and 1 male. Most people would look at those numbers and see that they would be extremely unlikely to happen due to luck, if the null hypothesis were true; you would reject the null hypothesis and conclude that chocolate really changed the sex ratio.
* You get 31 females and 17 males. Now what do you conclude? There are definitely more females than males, but is it really so unlikely to occur due to chance that you can reject the null hypothesis? To answer that, you need more than common sense; you need to calculate probability of getting a deviation that large by chance.

The **significance level** (also known as the “critical value” or “**alpha**”) you should use depends on the costs of different kinds of errors. With a significance level of 0.05, you have a 5% chance of rejecting the null hypothesis, even though it is true. That means that you would have a 1 in 20 chance of concluding falsely that your hypothesis was not supported. You must choose your significance level before you collect the data.

**Most scientists use alpha=0.05 by default**, because it is generally accepted. If you choose to use a different significance level than the conventional 0.05, people will be skeptical; you must be able to justify your choice. It is very likely that you will use the conventional significance level of 0.05 for all of the experiments you will do in the core labs.

However, you should understand why some scientists use other levels of significance so you can evaluate their conclusions effectively. For example, when reporting some medical experiments, the threshold of significance might be 0.01 to gain a higher level of plausibility of the stated conclusions. In other medical studies, 0.10 is acceptable, because scientists in that field do not want to reject the null hypothesis too soon; others use alpha =0.1 because the data they are using already has a lot of noise (variation) in it.

In the past, scientists reported the results of a statistical test as “P<0.05”, “P<0.01”, “P>0.10”, etc. Now, since almost all computer statistics programs give the exact P value resulting from a statistical test, such as P=0.029, that’s what you should report in your work.

You will conclude that the results are either significant or they’re not significant; they either reject the null hypothesis (if P is less than your pre-determined significance level) or don’t reject the null hypothesis (if P is greater than your significance level).

**Stepwise Guide to Experimental Design**

You will use a systematic, step-by-step approach to decide how to design your experiment and analyze data. Working through the steps below should be part of your weekly lab routine. Using this approach will aid in your understanding of the pre-set labs and for labs in which you will create and analyze your own experimental studies.

1. Collect **background knowledge and observations**. What background information is vital to your study? What did you need to know before you could proceed with asking your question?
2. Specify the **biological question** you are asking. What is it that you want to know?
3. Put the question in the form of a **biological null hypothesis and alternate hypothesis**. The biological hypothesis refers to your expectation of what will happen in the physical world due to a biological process or mechanism. It is a generalized statement explaining how a biological mechanism of interest causes an independent variable to affect a dependent variable.
4. Put the question in the form of a **statistical null hypothesis and alternate hypothesis**. The statistical hypotheses refer to your expectation of what the data will show without stating what biological processes caused those data to be different. This may be expressed as a comparison of averages between or among treatment groups, or in comparison to a hypothesized value.
5. **Determine which variables are relevant** to the question. Specify exactly what variables play a part in the question you are asking.
6. For each variable, determine and explain **what kind of variable each one is**. Measurement, categorical, other? Independent or dependent?
7. Design an experiment that **controls or randomizes the confounding variables**. Sketch out a framework for your collected data. What will it look like in a spreadsheet? This should include everything you need as a data collection table. How will you describe your summarized data? Choose and describe the best statistical test to use.
8. Do the experiment. **Carry out your plan,** being careful to stay on track by reviewing your key objectives at regular intervals.
9. **Apply the statistical test you have chosen and interpret the results.** Don’t just look at the number and blindly assign a conclusion to your work. Think through the results of your experiment. How do they compare in the bigger picture?
10. **Communicate your results** effectively, usually with a graph or table. There are many different ways in which we will ask you to communicate your work, such as lab reports and presentations. The key is that this communication synthesizes all of the information gathered, from background to interpreting your results, and presents it clearly.

# Learning to Use Excel

Microsoft Excel is the go-to spreadsheet program for business and science both. Some people use Google Sheets, but it has a VERY limited toolset that will create problems almost immediately for you. For instance, Sheets can make column bar graphs, but does not let you adjust the size of the error bars. **That means graphs made using Google Sheets CANNOT comply with our format requirements.**

Excel is used almost universally; if you do not already know how to use it already, it's time to learn. We have chosen several good starting tutorials. Every minute you invest in learning to use Excel efficiently pays you back in hours of timed saved.

## Tutorials from LinkedIn Learning

LinkedIn Learning (formerly Lynda.com) is THE online source for software tutorials. Their tutorials are so good that Adobe, Microsoft, and dozens of other big companies give their customers free 6-month accounts rather than produce their own software tutorials.

WFU has a campus-wide subscription for students and faculty to the entire online library. To access LinkedIn Learning:

1. Go to the Google App Launcher in your email. Or, visit [**lil.wfu.edu** (Links to an external site.)](http://lil.wfu.edu).
2. Sign on with your WFU credentials.
3. Complete a short enrollment process the first time you access the new platform.
4. If you have an existing LinkedIn account, you may choose to connect it with your new LinkedInLearning account, to highlight your training and certifications on your LinkedIn profile. You DO NOT have to link your accounts and can choose to skip this step.
5. If you had a Lynda.com account, check that your completed training and course certifications are showing in LinkedIn Learning.

We've picked some Excel tutorials that focus on specific skills you will need for lab. Don't just try these though; look around for others that answer your specific needs or interests.

### Completely New to Excel?

* + Learning Excel with David Rivers (1 hr, 15 min). [**LINK** (Links to an external site.)](https://www.linkedin.com/learning/learning-excel-desktop-office-365/create-a-new-workbook?u=67856482)
  + Excel Essential Training with Dennis Taylor (2 hr, 10 min). [**LINK** (Links to an external site.)](https://www.linkedin.com/learning/excel-2019-essential-training?u=67856482)

### Need Help Writing Functions for Calculations?

* + Intro to Formulas and Functions with Curt Frye (2 hr 58 min). [**LINK** (Links to an external site.)](https://www.linkedin.com/learning/excel-2016-introduction-to-formulas-and-functions?u=67856482)

### Want to Visualize Data in Graphs?

* + Excel: Introduction to Charts and Graphs with Dennis Taylor (56 min). [**LINK** (Links to an external site.)](https://www.linkedin.com/learning/excel-introduction-to-charts-and-graphs?u=67856482)

### Keep Having Errors?

* + Excel 2016: Avoiding Common Mistakes with Dennis Taylor (1 hr, 44 min). [**LINK** (Links to an external site.)](https://www.linkedin.com/learning/excel-2016-avoiding-common-mistakes?u=67856482)

### Want a Deeper Dive Into ALL That Excel Can Do?

* + Excel for Mac Essential Training with Curt Frye (4 hr, 42 min). [**LINK** (Links to an external site.)](https://www.linkedin.com/learning/excel-for-mac-essential-training-office-365?u=67856482)
  + Excel 2019 Essential Training (PC) with Dennis Taylor (2 hr, 8 min). [**LINK**](https://www.linkedin.com/learning/excel-2019-essential-training?u=67856482)

## Chi square goodness of fit

## Additional Resources

This part of the Guide is a starting point to help you learn basic principles of biostatistics. There are many good books and articles on biostatistics written for different sub-fields that go into greater depth. Here are some we like to recommend.

1. Nuzzo R. 2014. Statistical errors: P values, the ‘gold standard’ of statistical validity, are not as reliable as many scientists assume. *Nature*, 506:150-152.
2. Motulsky H. 2013. *Intuitive Biostatistics: A Non-Mathematical Guide to Statistical Thinking*, 3rd edition. Oxford University Press, 576 pp.
3. *MacDonald’s Biostatistics Handbook*. This is the original source for much of the material in this handbook, which we have re-used with Dr. MacDonald’s kind permission. [**http://www.biostathandbook.com/**](http://www.biostathandbook.com/)

## 

The description of **paired t-test** begins at time 6:54.

Use a paired t test when you have multiple pairs of observations of one group. It tests whether the mean difference in the pairs is different from 0. A paired t-test compares two means that come from the same group, but the group is measured twice. To do this test, you use all the data for each group (not just the means alone). The first time you measure the group is your control, then you do something to the group, and re-measure the same groups after the treatment (which is your experimental group), i.e., you measure a group of flowers the first time (control), then you give this same group of flowers water and fertilizer, and re-measure them (experimental).

### **Null hypothesis (H0)**

The mean of the group is not statistically significantly different between the first time you measure it, and the second time (after you added your treatment).

H0: x̄Time 1 = x̄Time 2

### **Alternative hypotheses (HA)**

**One tailed**

The mean of the group at one time point is statistically significantly greater or less than that of the mean at the other time point. (Here you are predicting the directionality of the difference.)

HA: x̄Time 1 > x̄Time 2

-or-

HA: x̄Time 1 < x̄Time 2 (depending on the direction you choose)

**Two tailed**

The means at the two time points are statistically significantly different. (Here you are not predicting the directionality of difference, just that the two means are statistically different.)

HA: x̄Time 1 ≠ x̄Time 2

### Example

Is the average systolic blood pressure (the top BP number) of students different before and after an exam? Here are the data (systolic BP in mm of mercury):

|  |  |  |
| --- | --- | --- |
|  | Systolic Blood Pressure |  |
| Student # | Before Exam | After Exam |
| 1 | 120 | 114 |
| 2 | 121 | 125 |
| 3 | 125 | 120 |
| 4 | 110 | 111 |
| 5 | 124 | 120 |
| 6 | 150 | 111 |
| 7 | 130 | 121 |
| 8 | 131 | 110 |
| 9 | 148 | 121 |
| 10 | 129 | 111 |

Remember from our earlier discussion of statistical significance and P-values, you will be determining whether there is a significant difference between the two groups by seeing whether the calculated P-value is greater than 0.05.

### Reporting your results

When reporting the results of a paired t-test, you need to include the resulting t-statistic, the degrees of freedom, and the corresponding P-value. Your statement might look like this, “The mean BP of the group after the exam was significantly greater than its mean before the exam (t-stat = 2.62, df = 9, P = 0.028).”

**Two-sample t-tests** are used to compare the means from two groups of data; specifically, the mean of a control group to the mean of an experimental group.

Use the two-sample t–test when you have one categorical variable and one measurement variable, and you want to compare the mean values of the measurement variable. The categorical variable must have only two values, such as “male” and “female” or “treated” and “untreated.” To do this test, you use all the data for each group (not just the means alone).

This version of the t-test should only be used when you are comparing data collected from two independent groups. This mean that they were collected from completely different groups, i.e., one group of flowers gets normal water (control) while a completely separate group of plants get water and fertilizer (experimental).

### **Null hypothesis (H0)**

The statistical null hypothesis is that the means of the measurement variable are equal for the two categories, i.e., there is no statistically significant difference between the control group’s mean and the experimental group’s mean.

H0: xC = xE

### **Alternative hypotheses (HA)**

**One tailed**  
The mean of one group is statistically significantly greater or less than that of the other group. (Here you must predict the directionality of the difference.)

HA: xC > xE or HA: xC < xE (depending on the direction you choose)

**Two tailed**  
The means between the two groups are statistically significantly different. (Here you are not predicting  
the directionality of difference, just that the two means are statistically different.)

HA: xC ≠ xE

### Example

Is the average height of students in 2 sections of biology lecture significantly different?

* Section A has 16 students with the following height in cm: 175, 177, 157, 160, 161, 170, 168,  
  168, 169, 174, 178, 183, 187, 152, 182, 181.
* Section B has 12 students with the following height in cm: 173, 170, 152, 160, 180, 170, 160,  
  161, 187, 152, 182, 181.

#### Remember from our earlier discussion of statistical significance and P-values, you will be determining whether there is a significant difference between the two groups by seeing whether the calculated P-value is greater than 0.05. What is it in this case?

### Reporting your results

When reporting the results of a two-sample t-test, you need to include the resulting t-statistic, the degrees of freedom (df), and the corresponding P-value. Your statement might look like this, “The mean of the control group was not statistically significantly different than that of the experimental group (t-stat = 0.57, df = 26, P = 0.574).”

## Chi square independence

**Analysis of variance (ANOVA)** is an approach that is used to simultaneously test whether the means of multiple (more than two) groups are equal. It works by assessing whether individuals chosen from different groups are, on average, more different than individuals chosen from the same group. It is used with one or more categorical variable.

### One Way ANOVA

Used with one measurement variable and one categorical variable, a one-way ANOVA should be used when you want to compare the means of more than two independent groups. You make multiple observations of the measurement variable for each value of the categorical variable. One of the means will be your control group (xc ), and the others will be the means of your experimental  
groups (x1, x2, x3,…,xk).

To do this test, you should use the raw data for each group (do not compare the actual means).

If your ANOVA tells you that at least one of the means is different from the other, you will need to perform additional **post hoc tests** to determine where a significant difference exists. You may choose to perform multiple comparisons using a two-sample t-test involving all possible pair-wise combinations of groups. However, this is generally not a recommended practice because it increases the possibility of detecting a statistical significance when one does not truly exist. An alternative is to use the **Tukey-Kramer test.** (The Tukey–Kramer test is also part of the template.)

In the Tukey–Kramer method, the minimum significant difference (MSD) is calculated for each pair of means. It depends on the sample size in each group, the average variation within the groups, and the total number of groups. For a balanced design, all of the MSDs will be the same; for an unbalanced design, pairs of groups with smaller sample sizes will have bigger MSDs. If the observed difference between a pair of means is greater than the MSD, the pair of means is significantly different.

### **Null hypothesis (H0)**

The means of the control and 2 or more experimental groups are not statistically significantly different.

H0: = x̄Time 1 = x̄Time 2 = x̄Time 3

### **Alternative hypotheses (HA)**

The mean of **at least one** of the groups is statistically significantly greater or less than the means for the other groups.

HA: x̄Time 1 ≠ x̄Time 2

-and/or-

HA: x̄Time 2 ≠ x̄Time 3

-and/or-

HA: x̄Time 1 ≠ x̄Time

Like t-tests, ANOVA can test directionality (one-tailed) or simply difference (two-tailed.)

### Example

Is the average height of students in 4 sections of biology lab significantly different?

Here are some example data that we could compare using ANOVA. From these we could determine whether the average height was significantly different between Sections A-D of the class. We would need to make pair-wise comparison after the initial ANOVA to determine which sections are different.

***Table: Height of Students in 4 Lab Sections (cm)***

|  |  |  |  |
| --- | --- | --- | --- |
| Section A | Section B | Section C | Section D |
| 175 | 173 | 157 | 190 |
| 177 | 170 | 166 | 185 |
| 157 | 152 | 176 | 197 |
| 160 | 160 | 180 | 190 |
| 161 | 180 | 160 | 190 |
| 170 | 170 | 165 | 179 |
| 168 | 160 | 175 | 185 |
| 168 | 161 | 155 | 192 |
| 169 | 187 | 159 | 180 |
| 174 | 152 |  |  |
| 178 | 182 |  |  |
| 183 | 181 |  |  |
| 187 |  |  |  |
| 152 |  |  |  |

Remember from our earlier discussion of statistical significance and P-values, you will be determining whether there is a significant difference between the groups by seeing whether the calculated P-value is greater than 0.05.

Reporting Results

When reporting the results of a one-way ANOVA, you need to include the P-value. Your statement might look like this:

“There was significant difference (P=0.000064) in the average height of the four lab sections. Further tests indicate that section D was significantly taller than all the other sections. No significant difference was indicated between other sections.”

The usual way to graph the results of a one-way ANOVA is with a bar graph. The heights of the bars indicate the means and there’s usually some kind of error bar. In this case, you should use SD (standard deviation). Also, you should use the SD for each group, not one value for all bars.

**Linear regression** finds the line that best fits the data points. There are actually a number of different definitions of “best fit,” and therefore a number of different methods of linear regression that fit somewhat different lines. By far the most common is “ordinary least-squares regression”; when someone just says “least-squares regression” or “linear regression” or “regression,” they mean ordinary least-squares regression. This is a measure of association and can evaluate causal relationships because you selected the range of X variables for which you measured Y. Thus, you can say, with more confidence, that at least for the X variables you chose, the X variables caused the pattern in the Y variable.

### **Null hypothesis (H0)**

The slope of the best-fit line is equal to zero. (The variables are not associated. The strength of the association is close to 0. You cannot predict values of Y using X .)

### **Alternative hypotheses (HA)**

There is a significant association between the X and Y variables you measured. (The slope of the best-fit line is not equal to zero.)

This is not a causal relationship, i.e., correlation does not imply causation!

### Example

The ground cricket is known to change the rate of its call, or “trill” with ambient air temperature. The following data was collected in a lab:

|  |  |
| --- | --- |
| Chirps per second | Temperature in degree C |
| 15 | 21 |
| 20 | 24 |
| 12 | 20 |
| 21 | 25 |
| 18 | 24 |
| 16 | 23 |
| 13 | 22 |
| 14 | 22 |
| 20 | 25 |
| 23 | 26 |
| 12 | 21 |
| 15 | 22 |
| 18 | 24 |
| 20 | 24 |

Reporting Results

When reporting the results, you will need to include the resulting r2 (correlation), d.f., and P value. From the cricket example,

“There was a significant association between the rate of cricket trills and air temperature measured in the lab (r2=0.903, d.f. = 12, P= 0.000000197).”

Correlation

https://youtu.be/GtV-VYdNt\_g

Regression

https://youtu.be/WWqE7YHR4Jc

ANOVA

https://youtu.be/oOuu8IBd-yo

Chi-square

https://youtu.be/7\_cs1YlZoug

T-tests

https://youtu.be/AGh66ZPpOSQ