Hypothesis Testing for Means:

Let's begin by reviewing hypothesis testing for the mean of a single population (Case 1).

Example 1: In a study with $n = 26$, the sample average amount of antitoxin needed to neutralize a snake poison was $\bar{x} = 1.89$ mg, with $s = .42$ mg. Previous research indicated that the true average amount needed was 1.75 mg. Does the new data suggest differently? Use $\alpha = 0.05$. (You may assume that amounts of antitoxins needed follow a normal distribution).

Step 1: Identify your parameter of interest.

Step 2: Write down your null and alternative hypotheses.

Step 3: What is your significance level?

Step 4: Determine the rejection region for your t-value. (Be careful! This is the first time we've seen a *two-tailed* test - how will this change your rejection region?)

Step 5: Calculate your test-statistic, and compare to the rejection region.

Step 6: Compute the p-value and compare to α. (Note: the fact that this is a two-tailed test will also have an effect on your p-value!)

Step 7: State your conclusion *in the context of the problem.*

Let's use what you know now to consider a problem where we want to draw conclusions about a mean of *paired differences*. Remember: in this case, we use the same procedure as in Case 1, but we use $\bar{d}$ in place of $\bar{x}$ and s_d in place of s, where $\bar{d}$ and s_d are the mean and standard deviation of the *differences* within the pairs.

Example 2: Suppose we are interested in the breaking load for fabrics that are in abraded versus unabraded conditions. We observe the following data:

	Fabric 1	Fabric 2	Fabric 3	Fabric 4	Fabric 5	Fabric 6	Fabric 7
Unabraded	36.4	51.5	38.7	43.2	48.8	25.6	49.8
Abraded	28.5	46.0	34.5	36.5	52.5	26.5	46.8

Is there sufficient evidence to suggest that the breaking load differs depending on whether a fabric is abraded or unabraded? Carry out a t-test at the $\alpha = 0.01$ level - you may assume that differences in breaking load are normally distributed. Make sure to show all of your work below.

Now you're ready to look at a difference in population means problem (Case 3). In this case, we will use a *two-sample t-test.* Let's walk through an example.

Example 3: Suppose we are interested in examining the effects of sleep on test scores. We gather 28 people and randomly assign half to Group 1 and half to Group 2. Group 1 is told to sleep 8 hours before they are given their exam. Group 2 sleeps only 4 hours before their exam. We observe the following results: $\bar{x}_1 = 83.5$ with $s_1^2 = 168.6$, and $\bar{x}_2 = 75.4$ with $s_2^2 = 73.32$. Is there sufficient evidence at the $\alpha = 0.05$ level to suggest that more sleep leads to better test scores?

Step 1: State your parameter of interest. In this case, your parameter of interest should look like "$\mu_1 - \mu_2$". Describe what this quantity represents below.

Step 2: Describe your null and alternative hypotheses. Here, these should take the form of "$H_0: \ \mu_1 = \mu_2$" and either "$H_a: \ \mu_1 \neq \mu_2$", "$H_a: \ \mu_1 > \mu_2$", or "$H_a: \ \mu_1 < \mu_2$". Choose an appropriate set of hypotheses and write them in below.

Step 3: Identify your significance level.

Step 4: We will still be using the t statistic in this case. Determine your rejection region for t as you would for any one-sample test (you can use your shortcut for unpooled variances to compute your degrees of freedom).

Step 5: Compute your test-statistic. In the two-sample case, this should look like

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}.$$

(You've seen that denominator before, when we did confidence intervals!). Remember, we are operating under the assumption that H_0 is true, so what can you say about $\mu_1 - \mu_2$? Compute t below.

Step 6: Compute the p-value as in any one-sample t test.

Step 7: State your conclusion in the context of the problem below.

Your turn! Consider the following examples, decide which case they fall under, and carry out the requested hypothesis test.

Example 4: Is there a significant difference in GPAs between sophomore and juniors? A random sample of 47 sophomores and 36 juniors was obtained at a large university. Researchers observed that the mean GPA for the sophomores was 2.840 with a standard deviation of 0.520. The mean GPA for the juniors was 2.981 with a standard deviation of 0.309. Is there sufficient evidence to conclude that there is a difference in GPA between the sophomore and junior classes? Carry out an appropriate hypothesis test at the $\alpha = 0.01$ level. Show all of your work.

Example 5: It is believed that the average time required to adapt to reduced light is 7 seconds. A sample was conducted with $n = 9$ people and it was observed that $\bar{x} = 6.32$ seconds, with $s = 1.65$ seconds. Assuming that adaptation time is normally distributed, does the data contradict the current belief? Using $\alpha = 0.1$, carry out an appropriate hypothesis test, showing all of your work.