**Using the Bootstrap to Test a Hypothesis**

Note that these differences are represented by a single column of data. *So, instead of viewing this as a problem involving a categorical predictor and a numerical response, you could view this as a problem involving a single numerical variable – the differences!* That is, the parameter of interest is the true population average of the differences which we will represent by μdifference.

* Our best estimate for this parameter is the sample mean of the observed differences. We’ll call this quantity.
* The sample standard deviation of the differences will be denoted by sdifference.

Before we discuss using a bootstrap sample to run a hypothesis test, let’s set up the null and alternative hypothesis for these data. (For convenience, the research questions is given below.)

**Research Question:** On average, does body weight of college students increase over the Thanksgiving holiday break?

**Step 1: Set up the null and alternative hypotheses**

Ho:

Ha:

Recall that we run a hypothesis test by performing the following two steps.

1. Simulate the null hypothesis
2. Determine if our statistic is unusual given this null distribution.

In this case, the null distribution is the *no average weight gain* distribution. To use the bootstrap to simulate this distribution, we will need to shift the mean of our data to simulate no weight gain. This is done by performing the following steps.

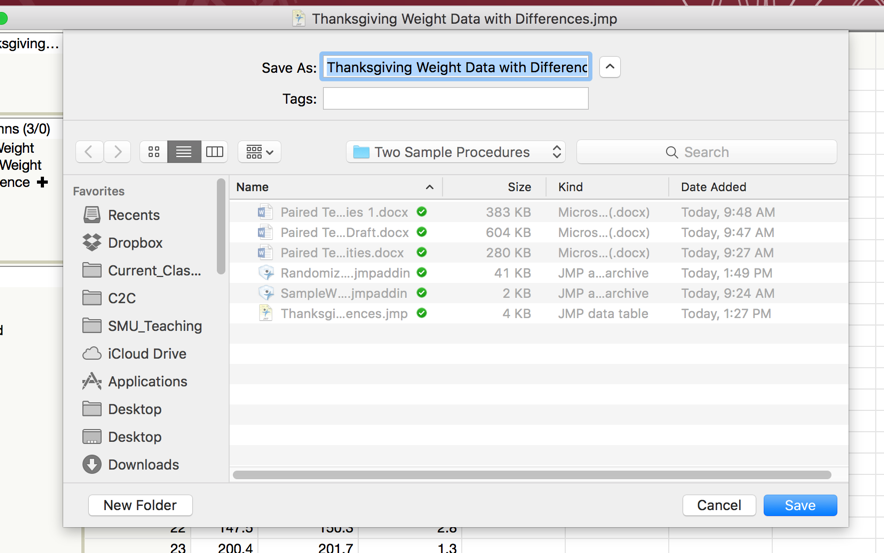
**Step 1:** Make a new column of differences by subtracting the original mean difference from the original differences. (You should verify that the resulting column has a mean of 0.)

**Step 2:** Take many bootstrap samples from this theoretical column to simulate the null distribution. For each sample compute the bootstrap mean.

**Step 3:** Use the resulting distribution to compute a p-value for the original statistic and make the appropriate conclusion to complete the testing process.

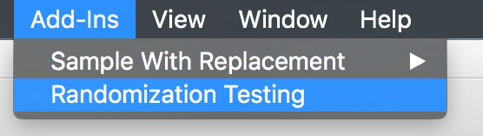
Luckily, there is a nice JMP add-in that automates this process. Please download and install the **Randomization Testing Beta 3.jmpaddin**

Let’s apply this process to the Thanksgiving weight gain data. We will use a nice add-in to JMP that automates these steps for use. To use this add-in, we need to save our data set ***with the differences***.

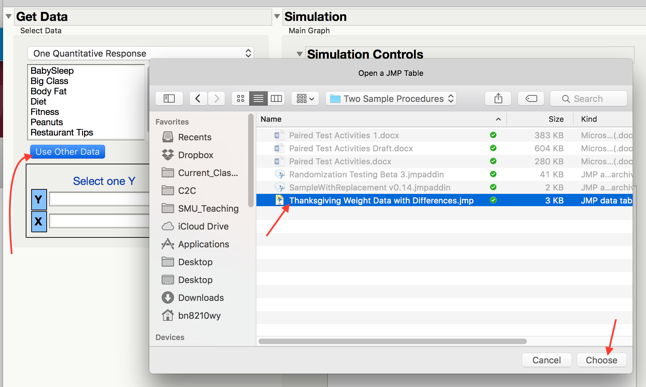


Now run the randomization addin by

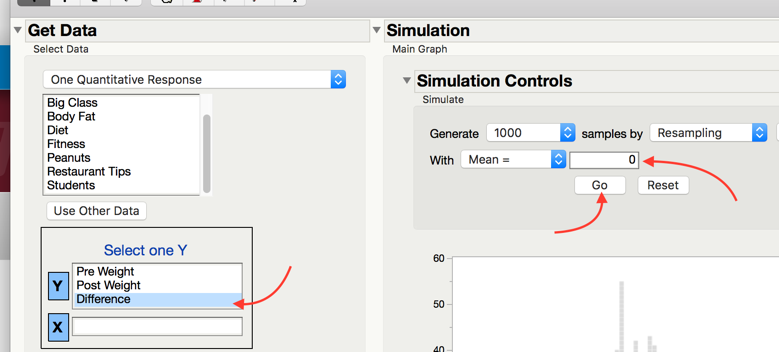
1. going to **Add-Ins > Randomization Testing**

****

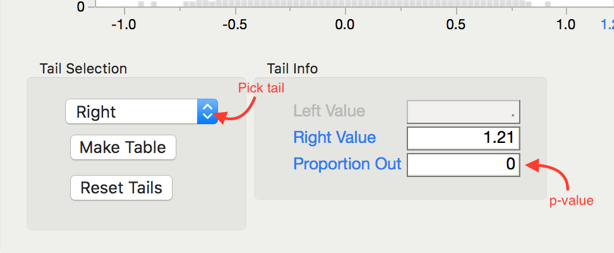
1. Click **Use Other Data** and open your JMP file ***with differences***.

****

1. Select the differences column, set the simulation controls to simulate **1000** samples by **Resampling** with **Mean =** 0 (The null mean).

****

1. Pick the correct tail for this test in **Test Selection** and note the p-value.

**** p-value =

**Step 3: Write a conclusion in the context of the problem**

**Paired T-Test in JMP**

Recall that we used a t-confidence interval in JMP as an alternative to the bootstrap interval. In the same way, we can use a t-test in JMP as an alternative to the bootstrap test. We first assess the normality of the data set. We can use the JMP test if (a) the original differences are normal or (b) the sample size is large enough (at least 30).

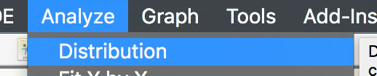
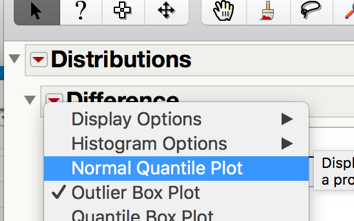
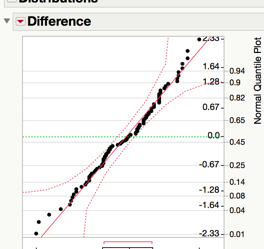
For the Thanksgiving weight gain data, run a JMP t-test by following these steps.

**Step 1: Set up the null and alternative hypotheses**

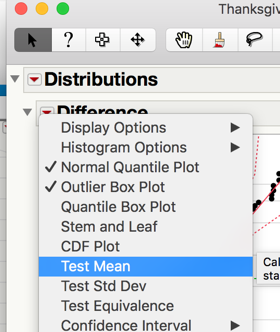
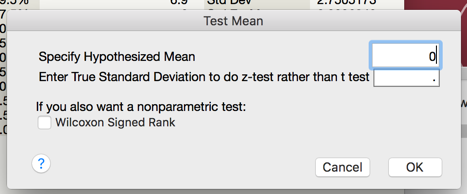
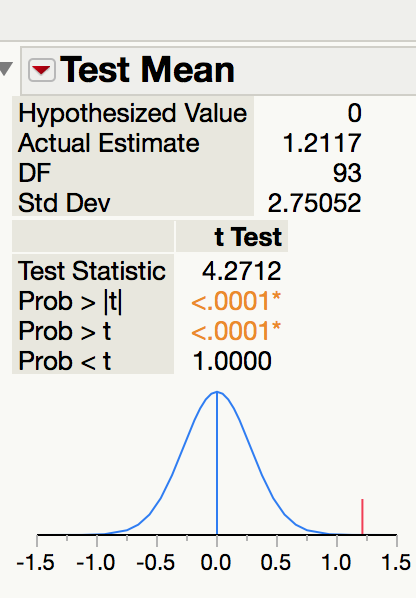
Ho:

Ha:

**Step 2: Make a difference column and assess the normality of the differences**

**  **

**Step 3: Find the t-statistic and the p-value**

To calculate this associated p-value in JMP, add the test probability to the difference distribution.  
   p-value =

**Step 3: Write a conclusion in the context of the problem**