

# Randomization Test for 2 Population Means

## Randomization for Two Means

We begin with two samples of quantitative data.

We use StatCrunch to pool all of the sample data and then randomly redistribute all of them to one of the two samples.

We then take note of the difference between the two sample means, and repeat the process over again many times.

We are looking to determine what proportion of the trials results in a difference that is at least as extreme as the difference between the original two sample means.

## StatCrunch Applet Steps

1. Load the data into a StatCrunch spreadsheet. Type the data from the first sample in the first column, and the data from the second sample in the second column.
2. Press Applets; highlight Resampling, then Randomization Test for Two Means.  
Select the column containing the first sample, and enter a label for that sample.  
Select the column containing the second sample, and enter a label for that sample.  
Click Compute!.
3. Make note of the difference between the two original samples.
4. Click 1000 times to randomize the two samples 1000 times.  
Repeat this process to obtain 10,000 runs.
5. Make note of the proportion of the 10,000 runs that produced a difference that was at least as extreme as the difference between the two original samples. (Use the "Total row.")
6. If the proportion is less than 0.05 then the difference between the two samples is unusual and that is evidence that there is a difference between the two population means.  
Otherwise, it is plausible that the two population means could be equal to each other.

## Investigation 1

A math instructor gave two versions of an exam to two groups of students. Version A was given to 36 students and they had a mean score of 79.3. Version B was given to 41 students and they had a mean score of 84.1. That is a difference of 4.8 points between the two versions of the exam.

*(Data can be found as "AMATYC 2020" in the "Woodbury Math 21" StatCrunch group.)*

Is there enough evidence to conclude that the two versions are not at the same level of difficulty? We will use a randomization test to answer that question.

**(a)** Use StatCrunch's randomization applet to generate 10,000 sample differences. Fill in the table.

Event	Count	Proportion of 10,000 runs
-4.844173 or below		
4.844173 or above		
Total		

**(b)** Circle the appropriate conclusion from your randomization test.

There is sufficient evidence to conclude that the two population means are different and that the two exams have different levels of difficulty.

OR

It is plausible that the two population means are equal and the two exams have the same level of difficulty.