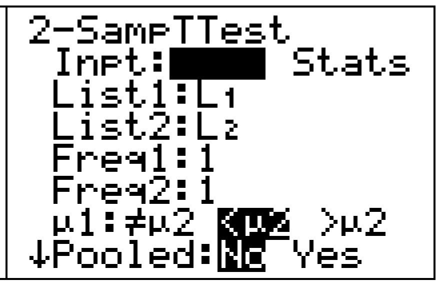
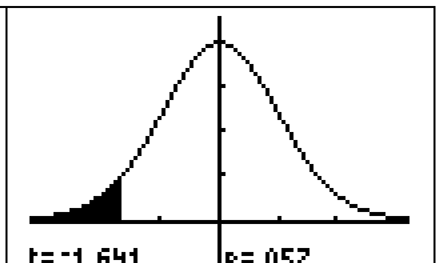
Two sample T test-the one you may well use for ISEF
Also a little on Anova

Yes, this is the statistics test most often and most often appropriately used in testing hypotheses for science fair and such experiments. The test is used to distinguish a control from the treated group.

Suppose you have 30 young carnivorous Phoenician Trifid plants to use to test your hypothesis that such plants will grow taller if one adult sewer fluke per week is added to their normal diet of one adult white mouse every two weeks. You divide your trifids into two groups of 15 to represent two much larger populations of trifids that would receive the same life experiences. Granted, larger samples would be better, but trifids, flukes, and mice are difficult to obtain and we really don't want to fool with more than 15 numbers in two lists anyway. So you establish your two groups in identical environments and proceed to feed on schedule as proposed. After 6 months (maturity), you measure the heights of both groups of plants. You find the control with the following heights in inches: 15.6, 15.7, 16.5, 16.8, 16.4, 13.4, 15.2, 15.9, 13.8, 11.9, 13.8, 17.2, 15.6, 14.7, and 14.6. The group with the fluke supplement measured: 17.3, 18.6, 16.1, 17.3, 12.8, 15.8, 17.1, 14.0, 17.2, 13.8, 14.8, 20.2, 13.9, 16.6, and 17.4. Place these numbers in L1 and L2, consecutively. Run 1 variable stats with **STAT**-*CALC* to see the difference in means. Surely the supplemented group is taller? But is the difference statistically significant? We will use the 2-SampTTest under **STAT**-*TESTS* to attempt to support your hypothesis.

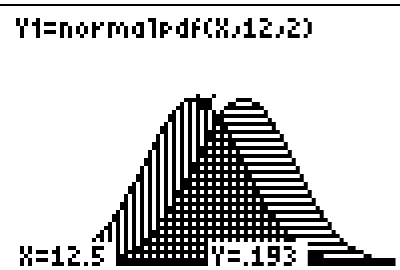
STAT-*TESTS*-4. Highlight *Data*, use L1 and L2 as shown for *List1* and *List2*, select your hypothesis that the mean of the first list is less than the mean of the second (<*μ2*), select *No* for *Pooled* because we don't know that the population variances are equal.	

Cursor down to the last line of the above window and select *Draw.*	

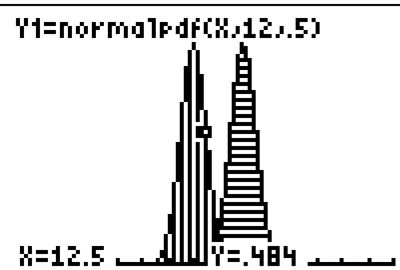
The p here, as in the one variable tests, is the probability associated with the null, the null hypothesis being that the mean of the control group is greater than the mean of the supplemented group, that μ1>μ2. Since this probability is not less than 0.05 or 5%, the null cannot be rejected and is therefore retained. Your hypothesis is rejected, even though you might argue that any fool could see that your supplemented trifids are taller. You results are simply not statistically significant within 95% certainty.

The 2-SampTTest takes into consideration the size of the samples n, the dispersion of the samples (standard deviation), and the difference of the sample means to calculate the probability you used. Two samples are more easily distinguished the further their means are apart, the smaller their standard deviations, and the larger the samples.

Consider the overlap in these two distributions with means of 12 and 14 and both with standard deviations of 2



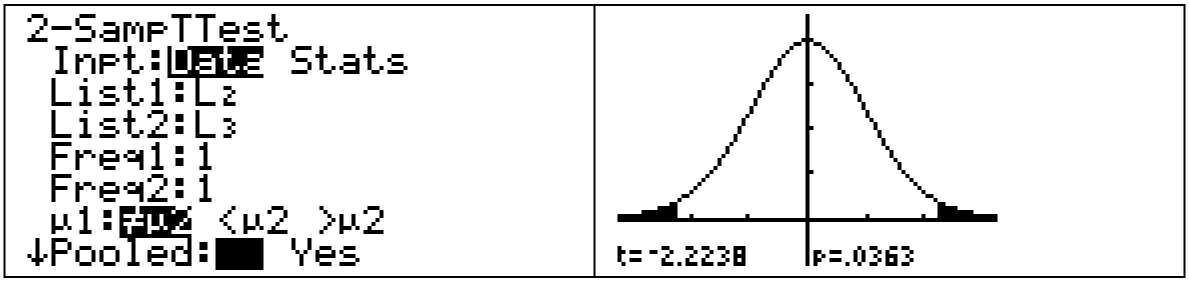
Now with a standard deviation of 0.5 instead of 2, changing only the Ymax window variable from .3 to .9, notice there is much less overlap and the two distributions are more easily distinguished even though the means are still 12 and 14.



Megaburger, Inc. is advertising that there is three times as much meat (not necessarily beef) in their standard $2 burger as there is in BurgerSwell's standard $3 burger. The little print at the bottom of the ad even gives additional information that this is before cooking. Somehow you obtain a random sample of 15 uncooked patties destined for the aforementioned burgers, 15 from each enterprise. Maybe you and/or your friends posed as government inspectors and crashed random locations with accurate gram scales. You have the following readings for BurgerSwell: 93, 89, 89, 93, 87, 91, 90, 89, 93, 91, 90, 84, 84, 93, and 90. MegaBurger: 251, 272, 252, 251, 272, 266, 267, 238, 283, 254, 234, 274, 277, 263, and 245. Place the first list in L1 and the larger numbers of the second list in L2.

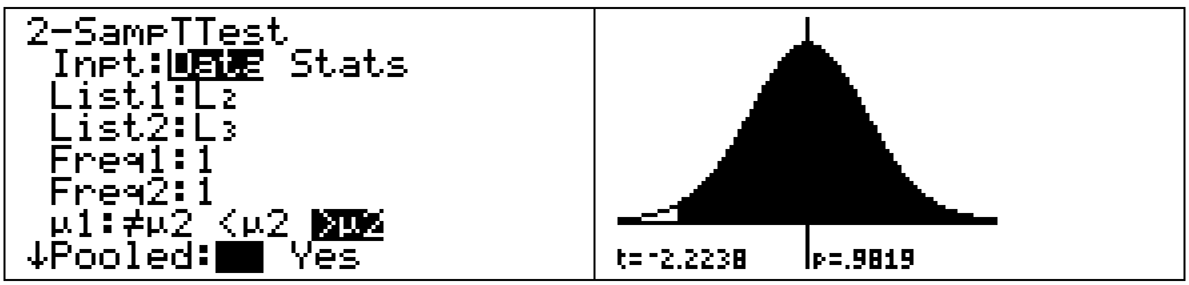
What we are about to do may not be approved by all (or any real) stat professors. We are going to multiply L1 by 3 and place the result in L3. Then we will

compare lists 2 and 3 with the 2SampTTest. Do 3*L1? L3. Remember that L1 is pasted with **2nd-List-L1** and the store arrow is **STO.**
Now let's test the hypothesis that L2 is not equal to L3:



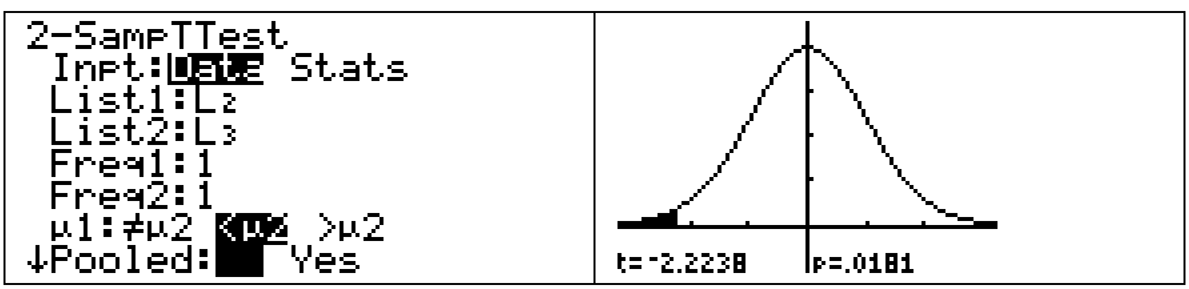
Cursor down to *Draw* and **ENTER** to get the final screen. The probability of the null (again, an improper thing to say in some circles) is less than 0.05 and you have "proven" that the means are not equal. So? Which company is happy?

Now test the hypothesis than L2 is greater than L3.



So you can definitely not prove your hypothesis that Megaburgers have 3 times the meat or more.

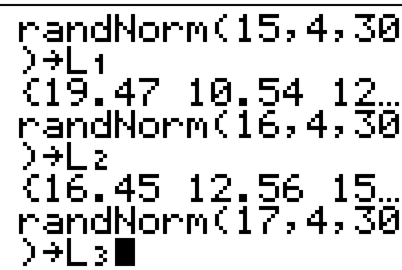
Now test the hypothesis that Megaburgers have less than 3 times the meat:



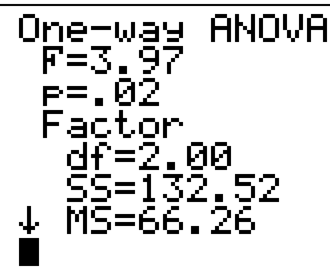
So you have defeated the null and "proven" that Megaburgers have less than 3 times the meat of the BurgerSwells because the p of 0.0181 is less than 0.05 or 5% and this is associated with the null that Megaburgers have more meat.

So what is this “ANOVA” you’ve heard so little about? Is it a glowing region in space? A Russian princess? Actually, it is “analysis of variance”. Remember that variance is the square of standard deviation, both being a measure of dispersion of a population or sample. Anyway, the test is used similarly to the 2SampTTest except you can use this one for more than 2 samples or lists. The hypothesis tested is that not all the means of the populations represented by the samples are equal. The null is that the means μ are all equal. Your TI83plus is good for up to 20 lists. Let’s do 3. Fill L1 through L3 with normally distributed random numbers and then run the ANOVA test as partially guided below:

MATH-*PRB*-6. Finish with means of 15, 16, and 17, standard deviations of 4 in all three samples, and n of 30 for each sample. You could use different standard deviations and different n’s for each list.



STAT-*TESTS*-*F* (or select ANOVA) and finish with “L1,L2,L3)” and **ENTER** to get screen as shown. Don’t worry about “Error” on down the screen. That’s Error data. For now just check out the *p* of 0.02 in this run. You may have a different result.



This p of 0.02 means that you have a difference in means since the probability associated with the null is less than 0.05. Experiment with lower n’s to see that this probability will increase. Also larger standard deviations will increase this probability, decreasing your chances of showing a difference in means and your chances of demonstrating that not all the samples represent the same population.