Common Statistical Tests Used to Analyze Data for Science Fair Experiments

Standard deviation is a measurement of the variability of a population. It shows how much variability, or dispersion, there is from the mean (average). A low standard deviation means that all the data points are very close to the mean, or that there is little variability. A high standard deviation means that the data points are spread out over a large area, or that there is greater variability in the population. In addition to expressing the variability of a population, standard deviation is commonly used to measure confidence in statistical conclusions. Science researchers commonly report the standard deviation of experimental data, and only effects that fall far outside the range of standard deviation are considered statistically significant.

Standard deviation is calculated in one of two ways, depending on if the data to be analyzed is a random sample of a large population (such as 100 sophomores from LRCHS) or the entire population (such as all 30 members of the wrestling team).

Population Standard Deviation (when you have data from the entire population) is calculated in the following way: Consider a population consisting of the following eight values:

$$2,\ 4,\ 4,\ 4,\ 5,\ 5,\ 7,\ 9$$

These eight data points have the mean (average) of 5:

$$\frac{2+4+4+4+5+5+7+9}{8}=5$$

To calculate the population standard deviation, first compute the difference of each data point from the mean, and square the result of each:

$$\begin{array}{ll}(2-5)^2=(-3)^2=9 & (5-5)^2=0^2=0\\ (4-5)^2=(-1)^2=1 & (5-5)^2=0^2=0\\ (4-5)^2=(-1)^2=1 & (7-5)^2=2^2=4\\ (4-5)^2=(-1)^2=1 & (9-5)^2=4^2=16\end{array}$$

Next compute the average of these values, and take the square root:

$$\sqrt{\frac{(9+1+1+1+0+0+4+16)}{8}}=2$$

The formula is valid *only* if the eight values we began with form the *complete* population. If they instead were a random sample, drawn from some larger, "parent" population, then we should have used 7 (which is $n - 1$) instead of 8 (which is n) in the denominator of the last formula, and then the quantity thus obtained would have been called the **sample standard deviation**

Standard Deviation, as with all statistical tests, is more accurate if your sample size or population size is large. So, keep this in mind when designing your experiment. Your results are more valid if you sample 100 sophomores at LRCHS, for example, rather than only 20. The more the better!

Common Statistical Tests Used to Analyze Data for Science Fair Experiments

How to do T-tests and ANOVAs to determine if there is a statistically significant difference between groups.

Which test do I use? If you are comparing data from two groups, say a control group of 5 plants and an experimental group of 5 plants, you can do a T-test to determine if there is a mathematically significant difference between the groups. If you are comparing 3 or more groups, and you just want to know if there is a statistically significant difference between any of the groups, you can us an Analysis of Variance Test to do this.

How to do a T-test
Hit the **STAT** button
Select **EDIT**
Enter the data from Group 1 in the column labeled L_1
Enter the data from Group 2 in the column labeled L_2

Hit the **STAT** button
Scroll over to **TESTS** then down to **2-Sample T-Test** and hit **ENTER**
Make sure the **DATA** option is selected (click on **DATA** if needed) then
Scroll down to **CALCULATE** and hit **ENTER**

Examine the results for the 2-Sample T-Test. Specifically, look at your *p* value.

If your *p* value is less than 0.05, the difference between the two groups *is* statistically significant. In other words, the chance that the difference between the groups is due to chance or error is < 5%. In other words, we are 95% certain that the difference between the groups is due to the independent variable tested.

How to do an ANOVA
Hit the **STAT** button
Select **EDIT**
Enter the data from Group 1 in the column labeled L_1
Enter the data from Group 2 in the column labeled L_2
Enter the data from Group 3 in the column labeled L_3

Hit the **STAT** button
Scroll over to **TESTS** then down to **ANOVA** then hit **ENTER.** On your screen you should see *ANOVA(*
Hit **2nd** button then hit the **1** button. On your screen you should now see *ANOVA(L1*
Hit the **,** button (comma) then hit **2nd** and **2.** You should now see *ANOVA (L1, L2*
Put in another comma then hit **2nd** and **3** so that on your screen you see *ANOVA (L1, L2, L3*
Now hit the **)** button (close the parentheses) so that you see *ANOVA (L1, L2, L3)* then hit **ENTER**

You should now see the results of your One-way ANOVA.
Look at the F and the p values. If F is larger than p, there is a statistically significant difference among your data. If F is not larger than p, then you do not have a statistically significant difference.

If it is not obvious which group is statistically significant from the others, you can do a T-Test to compare them two by two till you figure out which group is significantly different from the others.

CAUTION: watch out for scientific notation when reading your results!