Basic Statistics and how to calculate them in a TI-83 graphing calculator

The **sample mean** (average) is a common statistic used to describe a sample. It is computed as the sum of all observations divided by the number of observations:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

To calculate the mean in a TI-83, you press 2^{nd} **STAT**, then go to MATH (\rightarrow , \rightarrow), then "mean("(3). Then enter the number of the list in which your data is stored, for example L2.

The **sample variability** is another important statistic of interest, and it measures how much scatter there is around the mean. It is computed as the sum of the squared deviates, divided by the sample size minus one:

$$s^{2} = \frac{1}{N-1} \sum_{i=1}^{N} (y_{i} - \overline{y})^{2},$$

To calculate the variance you basically do the same as above, except instead of doing 3, for the mean you press 8, for variance.

The **standard deviation** is the most commonly used measure of statistical dispersion. Simply put, it measures how spread out the values in a data set are. The standard deviation is defined as the square root of the variance. This means it is the root mean square (RMS) deviation from the average. It is defined this way in order to give us a measure of dispersion that is (1) a non-negative number, and (2) has the same units as the data. For example, if the data are distance measurements in meters, the standard deviation will also be measured in meters. Standard deviations is often calculated using the following formula, specially for large populations:

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2}$$

To calculate the standard deviation you do the same as the two above, except you will now use 7, for stdDev.

Sometimes Standard deviation is also defined as:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \overline{x})^2}$$

Standard error of a measurement, value or quantity is the standard deviation of the process by which it was generated, after adjusting for sample size. In other words, it is a measure of how reliable the sample mean is as an approximation to the population mean.

$$\frac{\sigma}{\sqrt{n}}$$

Another important statistical method to know is the t-test. The **student's t-test** is done in order to compare the means of two samples and assess whether or not they differ enough to represent samples from different populations. In order to do a t-test you first need to find a t-value.

$$t = \frac{\overline{x}_{1} - \overline{x}_{2}}{\sqrt{\frac{\sigma_{1}^{2}}{n_{1}} + \frac{\sigma_{2}^{2}}{n_{2}}}}$$

To do that in the calculator you press **STAT**, and go to TESTS, there you will find all sorts of tests you can do. In this case we want #4 (2-SampTTest), where you can type which lists of data you want to compare and just calculate it. This should give you the mean for both lists and the t value of the comparison.

Now we need to decide the degrees of freedom of this sample, which is:

df = n1 + n2 - 2

You should also decide the probability level of this test, which indicates how much of the curve will be considered significant. It is usual to use a probability level of 5% (P=0.05) to establish significance. Therefore if the t value is even larger, note the smallest P value, the difference between the two samples are even more significant. Now we can look at a table and find out what the critical t-value is.

For 5% significant:

<u>DF</u>	<u>Value of t</u>	DF	<u>Value of t</u>	DF	<u>Value of t</u>
1	6.314	12	1.782	23	1.714
2	2.920	13	1.771	24	1.711
3	2.353	14	1.761	25	1.708
4	2.132	15	1.753	26	1.706
5	2.015	16	1.746	27	1.703
б	1.943	17	1.740	28	1.701
7	1.895	18	1.734	29	1.699
8	1.860	19	1.729	30	1.697
9	1.833	20	1.725	40	1.684
10	1.812	21	1.721	60	1.671
11	1.796	22	1.717	120	1.658
				Infinity	1.645

1	12.70	16	2.12
2	4.30	17	2.11
3	3.18	18	2.10
4	2.78	19	2.09
5	2.57	20	2.09
6	2.45	22	2.07
7	2.36	24	2.06
8	2.31	26	2.06
9	2.26	28	2.05
10	2.23	30	2.04
11	2.20	40	2.02
12	2.18	60	2.00
13	2.16	90	1.99
14	2.14	120	1.98
15	2.13	infinite	1.96

Note that the BIO150 table for t-values is slightly different, the values above correspond to 10% significance there. For 5% they use:

- Leonardo Silenieks