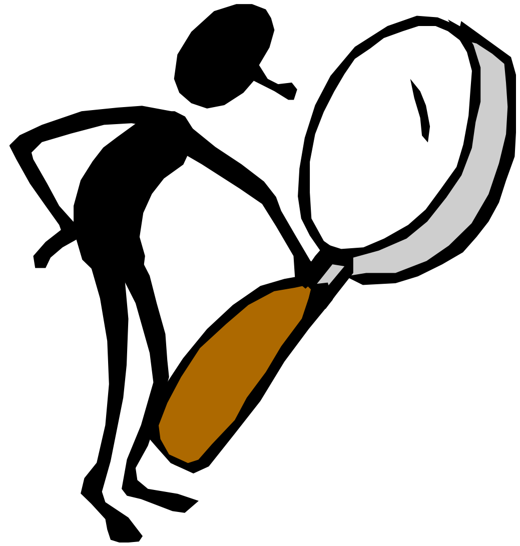
ANALYSIS OF VARIANCE



It’s been said that all statistics is about analyzing variance, i.e. why one measurement differs from another and what that means.

In this presentation we will discuss how to determine if means from several samples are all equal to one another or not. The technique is called a *one-way analysis of variance* or a *one-way ANOVA.*

Before we start, though, a good question is if we simply have three samples and if our null hypothesis is that the three sample means are all equal to one another, then why not simply do something like three separate *t-tests*?

Here is the answer. Suppose our null hypothesis is as below and that we want to do a test at the .05 level of significance.

$$H_0 : u_1 = \mu_2 = \mu_3$$

Recall, too, that a *Type I Error* is the probability of rejecting a true null hypothesis, and this is the same as our level of significance, *.05*.

$$H_0 : u_1 = \mu_2 = \mu_3$$

Hence, the probability of <u>not</u> making a *Type I Error* is *1 - .05 = .95.*

$$H_0 : u_1 = \mu_2 = \mu_3$$

$$P(\text{no Type I}) = .95$$

However, if we do three separate tests for equality of means, then the probability of not making a *Type I Error* changes.

$$H_0 : u_1 = \mu_2$$
$$\mu_2 = \mu_3$$
$$\mu_3 = \mu_1$$

$$P(\text{no Type I on test 1 \& no Type I on test 2 \& no Type I on test 2 })$$
$$= P(\text{no Type I on test 1 }) \cdot P(\text{no Type I on test 2 }) \cdot P(\text{no Type I on test 3 })$$
$$= (.95)(.95)(.95) = 0.857375$$

And the resulting *alpha* is *1 - .857375 = 0.142625.*

$$H_0 : u_1 = \mu_2$$
$$\mu_2 = \mu_3$$
$$\mu_3 = \mu_1$$

$$P(\text{no Type I on test 1 \& no Type I on test 2 \& no Type I on test 2 })$$
$$= P(\text{no Type I on test 1 }) \cdot P(\text{no Type I on test 2 }) \cdot P(\text{no Type I on test 3 })$$
$$= (.95)(.95)(.95) = 0.857375$$

Hence, our probability of making a *Type I Error* has gone from *5%* to *14.2625%*, and we call this *inflating the alpha*.

$$H_0 : u_1 = \mu_2$$
$$\mu_2 = \mu_3$$
$$\mu_3 = \mu_1$$

$$P(\text{no Type I on test 1 \& no Type I on test 2 \& no Type I on test 2 })$$
$$= P(\text{no Type I on test 1 }) \cdot P(\text{no Type I on test 2 }) \cdot P(\text{no Type I on test 3 })$$
$$= (.95)(.95)(.95) = 0.857375$$

Thus, if we can do just a single test, then we can minimize the probability of making a *Type I Error.*

$$H_0 : u_1 = \mu_2$$
$$\mu_2 = \mu_3$$
$$\mu_3 = \mu_1$$

$$P(\text{no Type I on test 1 \& no Type I on test 2 \& no Type I on test 3 })$$
$$= P(\text{no Type I on test 1 }) \cdot P(\text{no Type I on test 2 }) \cdot P(\text{no Type I on test 3 })$$
$$= (.95)(.95)(.95) = 0.857375$$

Let’s continue with the assumption that we have three samples with three means that are equal to one another.

$$H_0 : u_1 = \mu_2 = \mu_3$$

The idea behind our procedure is that if we combine the three samples together, then we can compute the overall variance in two ways.

$$H_0 : u_1 = \mu_2 = \mu_3$$

The first way would be to simply compute the variance of the sample means. This is called the *variance between samples*. It is also known as the *mean square treatment* or *mean square factor*.

$$H_0 : u_1 = \mu_2 = \mu_3$$

$$MS(\text{treatment}) = MS(\text{factor})$$

Treatment or *factor* refers to what is being studied such as the effects of different medications or other medical treatments.

$$H_0 : u_1 = \mu_2 = \mu_3$$

$$MS(\text{treatment}) = MS(\text{factor})$$

Square refers to the fact that the variance is the square of the standard deviation.

$$H_0 : u_1 = \mu_2 = \mu_3$$

$$MS(\text{treatment}) = MS(\text{factor})$$

And finally, *mean* refers to the fact that we are trying to find the mean or average squared deviation between the sample means.

$$H_0 : u_1 = \mu_2 = \mu_3$$

$$MS(\text{treatment}) = MS(\text{factor})$$

The second way to estimate the overall variance is to find the variance within each sample and then take an average. We call this the *variance within samples* or the *mean square error*.

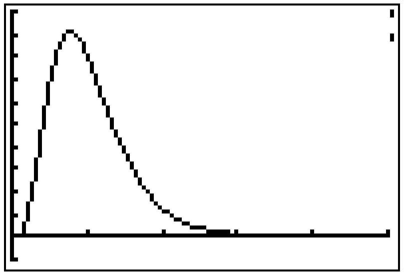
$$H_0 : u_1 = \mu_2 = \mu_3$$

$$MS(\text{error})$$

If we now take the ratio of the two estimates, then we get a value that pertains to an *F distribution*, named after Sir Roland Fisher who pioneered this technique.

$$H_0 : u_1 = \mu_2 = \mu_3$$

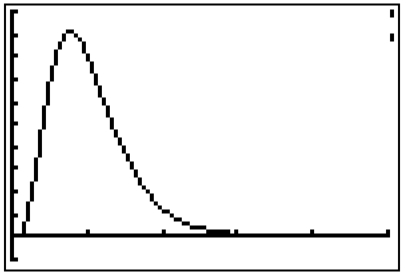
$$F = \frac{\text{variance between samples}}{\text{variance within samples}} = \frac{MS(\text{factor})}{MS(\text{error})}$$



We will only reject our null hypothesis if the numerator is too large, and this will happen only if there are "large" differences between the sample means. Thus, we do a one-tailed test.

$$H_0 : u_1 = \mu_2 = \mu_3$$

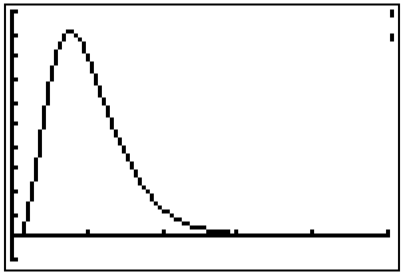
$$F = \frac{\text{variance between samples}}{\text{variance within samples}} = \frac{MS(\text{factor})}{MS(\text{error})}$$



Also, we have both *numerator degrees of freedom* and *denominator degrees of freedom*.

$$H_0 : u_1 = \mu_2 = \mu_3$$

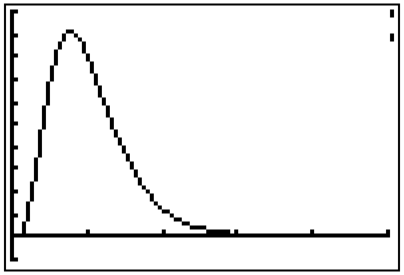
$$F = \frac{\text{variance between samples}}{\text{variance within samples}} = \frac{MS(\text{factor})}{MS(\text{error})}$$



The *numerator degrees of freedom* is equal to one less than the number of levels or categories in our factor (treatment). Or in other words, one less than the number of samples.

$$H_0 : u_1 = \mu_2 = \mu_3$$

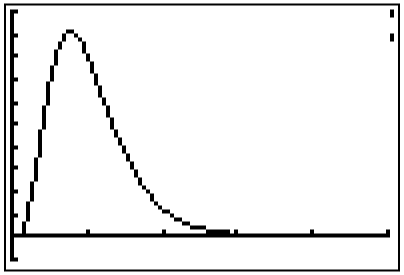
$$F = \frac{\text{variance between samples}}{\text{variance within samples}} = \frac{MS(\text{factor})}{MS(\text{error})}$$



The *denominator degrees of freedom* is equal to the total number of elements of data minus the number of categories.

$$H_0 : u_1 = \mu_2 = \mu_3$$

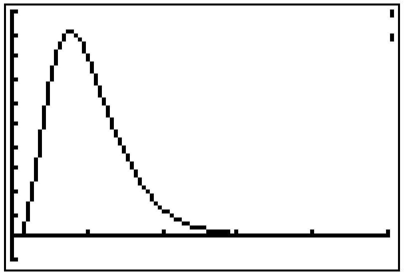
$$F = \frac{\text{variance between samples}}{\text{variance within samples}} = \frac{MS(\text{factor})}{MS(\text{error})}$$



The is the same thing you would get if you subtracted one from the size of each sample and then added up the results.

$$H_0 : u_1 = \mu_2 = \mu_3$$

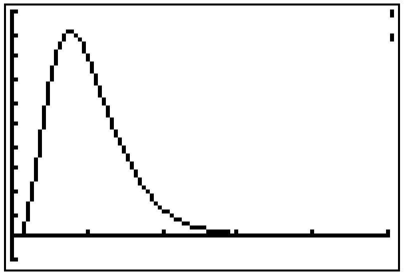
$$F = \frac{\text{variance between samples}}{\text{variance within samples}} = \frac{MS(\text{factor})}{MS(\text{error})}$$



Fortunately, our calculator will do all of this for us automatically.

$$H_0 : u_1 = \mu_2 = \mu_3$$

$$F = \frac{\text{variance between samples}}{\text{variance within samples}} = \frac{MS(\text{factor})}{MS(\text{error})}$$

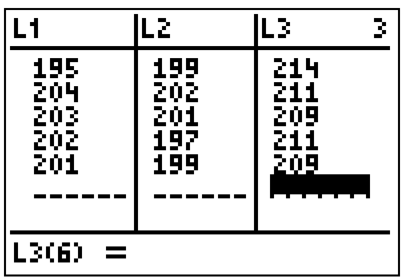


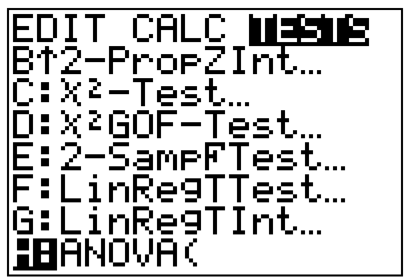
<u>Example</u>: A cyclist rides in a three mile race, and each mile is composed of 5 laps. The cyclist records in seconds the amount of time that each lap takes. Is there a difference in the amount of time it takes him to do each mile? Use *alpha = .05.*

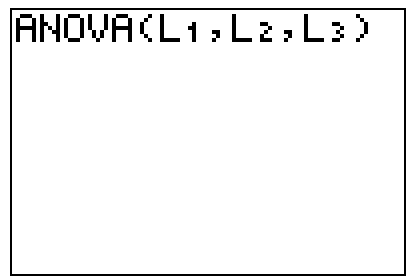
Mile 1	195	204	203	202	201
Mile 2	199	202	201	197	199
Mile 3	214	211	209	211	209

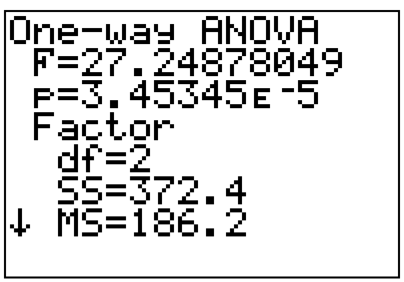
Example: A cyclist rides in a three mile race, and each mile is composed of 5 laps. The cyclist records in seconds the amount of time that each lap takes. Is there a difference in the amount of time it takes him to do each mile? Use *alpha = .05.*

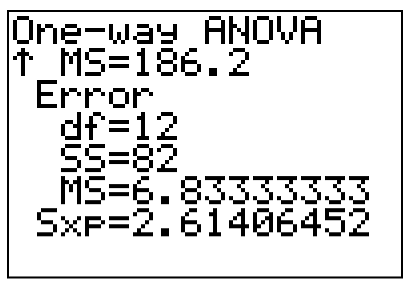
Mile 1	195	204	203	202	201
Mile 2	199	202	201	197	199
Mile 3	214	211	209	211	209





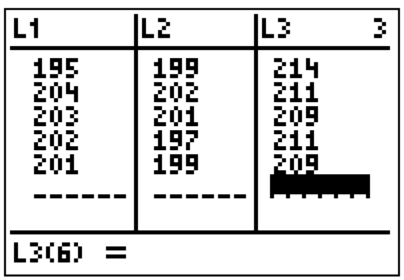


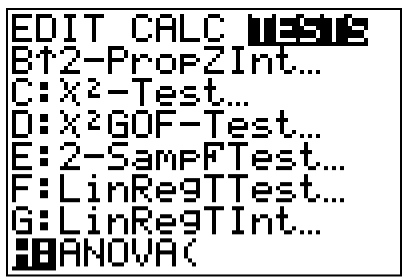


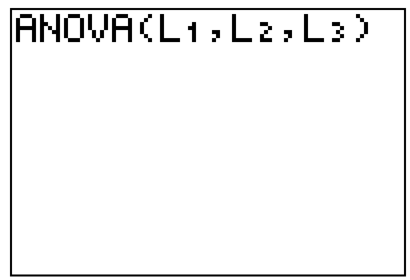


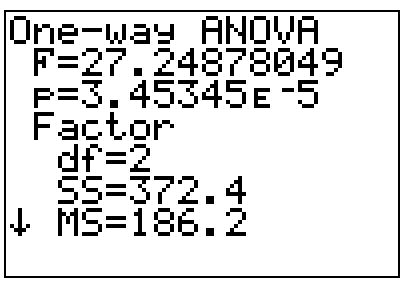
Example: A cyclist rides in a three mile race, and each mile is composed of 5 laps. The cyclist records in seconds the amount of time that each lap takes. Is there a difference in the amount of time it takes him to do each mile? Use *alpha = .05.*

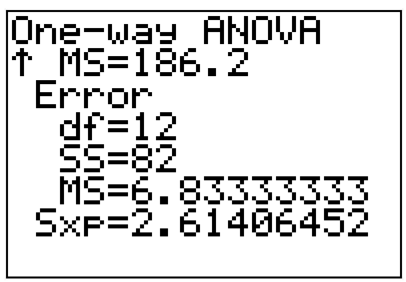
Mile 1	195	204	203	202	201
Mile 2	199	202	201	197	199
Mile 3	214	211	209	211	209











The *P* is low, the null must go!

Example: A cyclist rides in a three mile race, and each mile is composed of 5 laps. The cyclist records in seconds the amount of time that each lap takes. Is there a difference in the amount of time it takes him to do each mile? Use *alpha = .05.*

$H_0 : \mu_1 = \mu_2 = \mu_3$

H_1 : At least one of the means is different

Test Statistic Formula: $F = \dfrac{\text{variance between samples}}{\text{variance within samples}} = \dfrac{MS(\text{factor})}{MS(\text{error})}$

Numerator Degrees of Freedom: $ndf = 2$

Denominator Degrees of Freedom: $ddf = 12$

Level of Significance: $\alpha = .05$

P-Value: $P = 0.00003$

Decision: Reject H_0

The question now is which mean is different from the others?

Mile 1	195	204	203	202	201
Mile 2	199	202	201	197	199
Mile 3	214	211	209	211	209

If we have sophisticated software for ANOVA, we can follow up with a *post-hoc test*.

Mile 1	195	204	203	202	201
Mile 2	199	202	201	197	199
Mile 3	214	211	209	211	209

ANOVA Table for TIMES

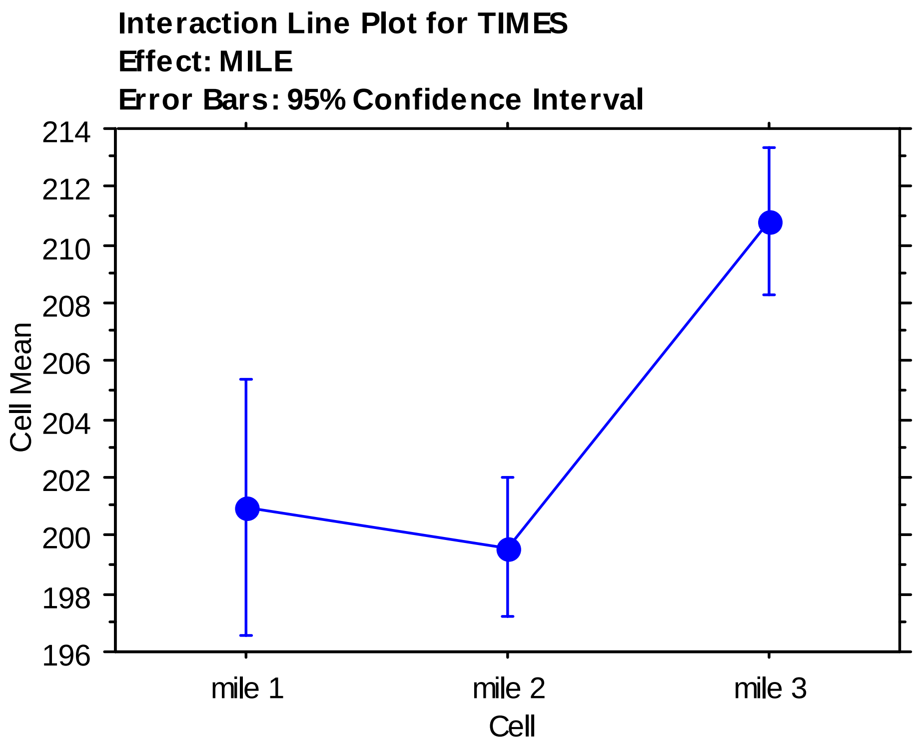
	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
MILE	2	372.400	186.200	27.249	<.0001	54.498	1.000
Residual	12	82.000	6.833				

Games/Howell for TIMES
Effect: MILE
Significance Level: 5 %

	Mean Diff.	Crit. Diff	
mile 1, mile 2	1.400	5.487	
mile 1, mile 3	-9.800	5.510	S
mile 2, mile 3	-11.200	3.616	S

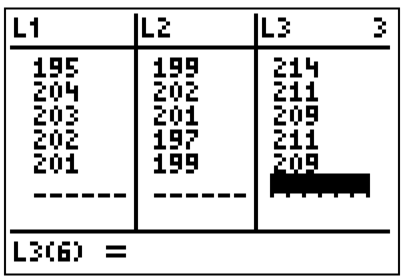
We can also create plots with *95% confidence interval* bars in order to see where the differences are.

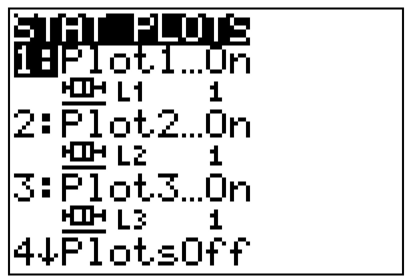
Mile 1	195	204	203	202	201
Mile 2	199	202	201	197	199
Mile 3	214	211	209	211	209

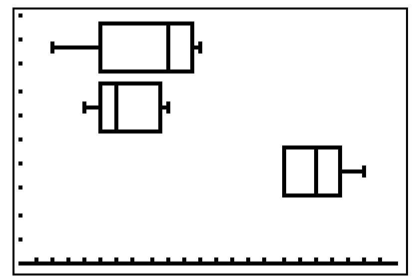


With a TI-83/84 calculator, though, we can use boxplots to estimate what's going on.

Mile 1	195	204	203	202	201
Mile 2	199	202	201	197	199
Mile 3	214	211	209	211	209







There are some basic assumptions that we make, though, whenever we do an ANOVA. Here are the main ones:

There are some basic assumptions that we make, though, whenever we do an ANOVA. Here are the main ones:

- The samples are independent of one another.

There are some basic assumptions that we make, though, whenever we do an ANOVA. Here are the main ones:

- The samples are independent of one another.
- The samples come from populations that are normally distributed.

There are some basic assumptions that we make, though, whenever we do an ANOVA. Here are the main ones:

- The samples are independent of one another.
- The samples come from populations that are normally distributed.
- The samples come from populations that have the same variance.

ANOVA is fairly robust and tends to work well even when normality and homogeneity of variance (*homoscedasticity*) are violated.

- The samples are independent of one another.

- The samples come from populations that are normally distributed.

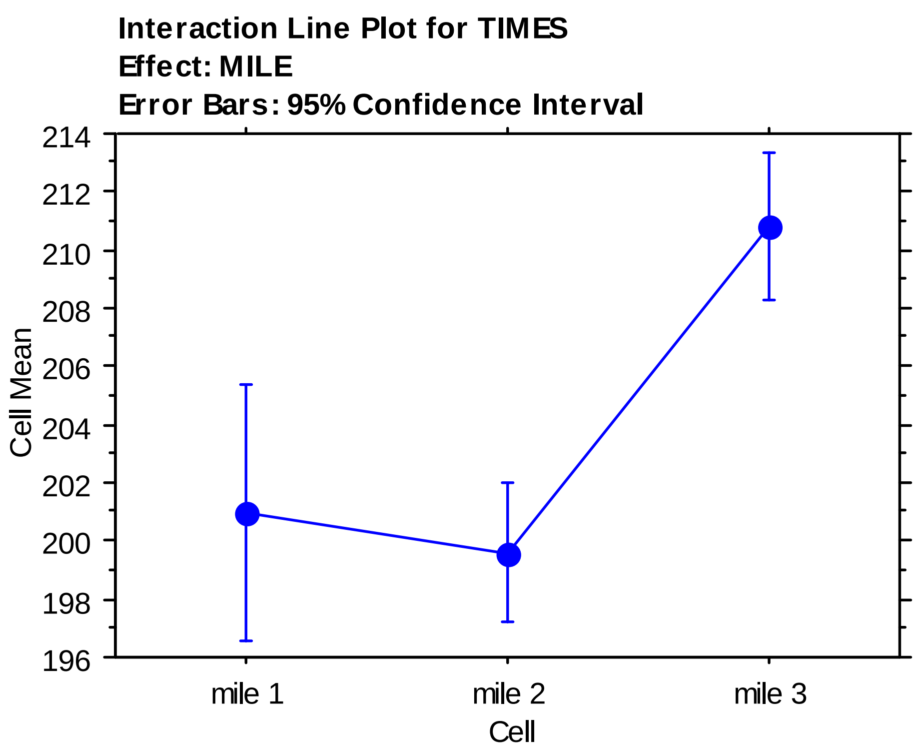
- The samples come from populations that have the same variance.

The one exception is when both sample variances and sample sizes are unequal.

- The samples are independent of one another.
- The samples come from populations that are normally distributed.
- The samples come from populations that have the same variance.

In this situation, it may be better to simply do an analysis using a graph of sample means with appropriate error bars.

Mile 1	195	204	203	202	201
Mile 2	199	202	201	197	199
Mile 3	214	211	209	211	209



mile 1

TInterval
(196.61,205.39)
$\bar{x}$=201
Sx=3.535533906
n=5

mile 2

TInterval
(197.18,202.02)
$\bar{x}$=199.6
Sx=1.949358869
n=5

mile 3

TInterval
(208.26,213.34)
$\bar{x}$=210.8
Sx=2.049390153
n=5