Statistical Analysis

Why use statistical analysis?

- Scientists must assume that their variables are not connected until sufficient empirical evidence is gathered
  - "False until proven true"
- Provides numerical data that show probability that independent variable does have an effect on the dependent variable
  - Can be used as supporting empirical evidence
  - The more trials, the better
  - No hypothesis is ever "proven"

Types of Statistical Analysis

ARE YOUR RESULTS DUE TO CHANCE OR NOT? USE THESE TESTS TO FIND OUT!!!!

- Chi square test
  - Good for biological experiments in which expected result is easily determined
  - Can be done by hand or with a calculator
- T-test
  - Tells if the variation between two groups is significant
    - Control and experimental
- ANOVA
  - ANalysis Of VAriance between groups
  - Similar to t-test
  - Better for multiple groups of data

Chi-Square Tests

- Write null hypothesis
  - Assume there is no effect of independent variable on dependent variable
- Use null hypothesis to get expected value
  - You may need to research to find what the expected value should be
  - In the homework example, the expected value was determined using the expected assortment of phenotypes based on a dihybrid cross (parents that were heterozygous for the trait being observed)
  - Expected value is specific to what you are testing

Chi-Square Tests, continued

- Using the chi-square equation, plug in your numbers from the expected values (null hypothesis) and the observed values (experimental data)

$$\chi^2 = \sum \frac{(\text{observed value} - \text{expected value})^2}{(\text{expected value})}$$

Determine the number of degrees of freedom for your test by taking the total number of terms (columns from data table) and subtracting one

Chi-Square Tests, continued

- Then compare the value to the table below
  - The same table is used for all chi-square tests
  - ANOVA is usually more practical to use

d.f.	.90	.80	.70	.50	.30	.20	.10	.05	.01
1	.016	.064	.148	.455	1.07	1.64	2.71	3.84	6.64
2	.211	.446	.713	1.38	2.41	3.22	4.6	5.99	9.20
3	.548	1.00	1.42	2.37	3.66	4.64	6.25	7.82	11.3

Top row: probability (percent) that your results are due to chance

← Results due to chance. Accept null hypothesis. → Results due to variables. Reject null.

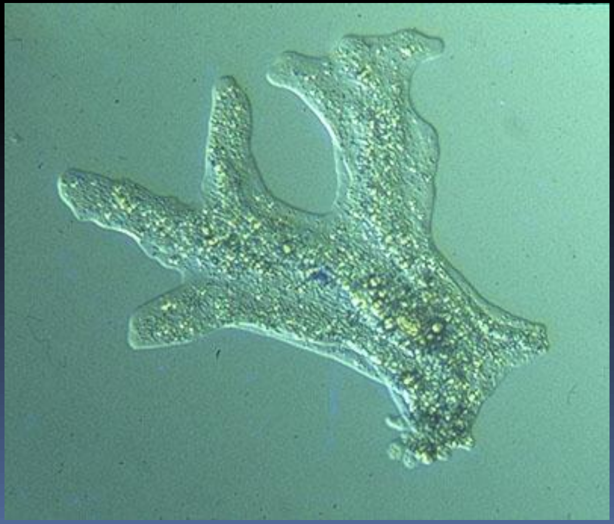
Sample T-Test

Table 1. Growth of *Amoeba proteus*

Time (days)	# individuals (luria broth)	# individuals (hayseed)
0 (initial)	100	100
1	110	115
2	121	132
3	133	152
4	146	175

Sample T-Test

- Your experiment will have four terms
  - Don't use initial entry
  - You will have three degrees of freedom
    - More individuals/groups/trials introduce a greater likelihood of change due to change, thus necessary chi-square values are higher to reject null hypothesis
- Use the luria broth growth rate as your "expected values"
  - This was the control group
  - Hayseed medium was experimental

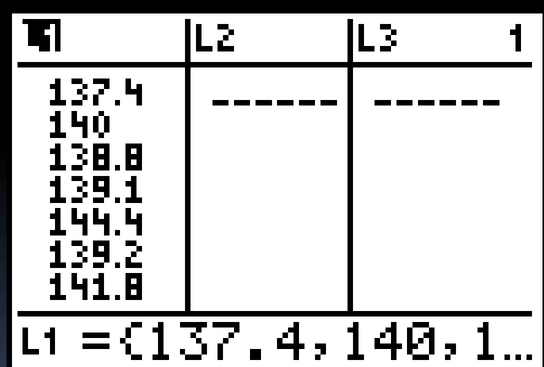


T-Test

- T-tests are used to compare the variation between *two groups*
  - For our purposes, these will be a “control” or “expected values” group
  - If you have more than two groups (or levels or independent variable), use ANOVA
- You can run a t-test on your calculator (TI-83 and above) or in Microsoft Excel
- For each statistical analysis test, you will use a table specifically designed for that test
  - Probabilities may be different

T-Test on your calculator

- Put data into two lists
  - Use L_1 and L_2 to represent control and experimental groups



To enter data into a list:

-Hit STAT button

-Select option 1, Edit

-Enter your values, hitting return after each

-L1 should be the control group

-Hit the right arrow key to move to List 2

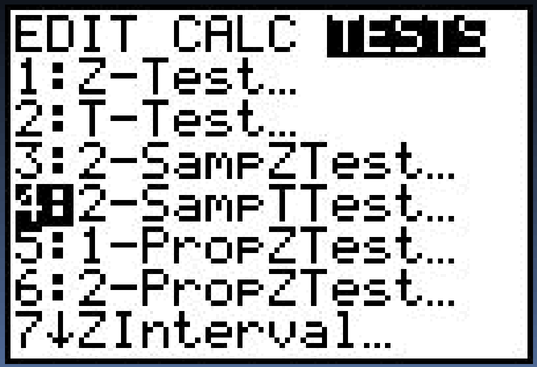
-Enter the experimental values there

-To clear out numbers already there, use arrows to move up to list name, then hit "clear" and enter

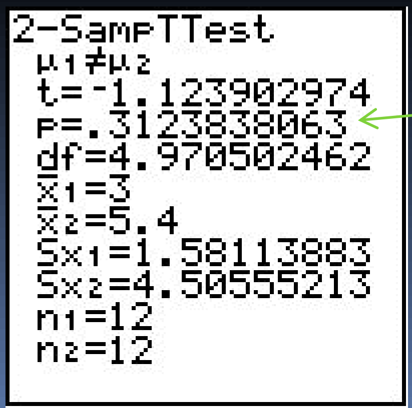
Do this with the data from the amoeba experiment!

T-Test on your calculator

- To run the t-test with the data you entered, hit STAT, then hit the right arrow twice to get to TESTS
- Select option 4, 2 sample t-test



- Under Inpt, select DATA
- L1 and L2 should be set as List 1 and List 2
- μ1 should be : ≠μ2
- Pooled: NO
- Hit CALCULATE



Probability-p-value

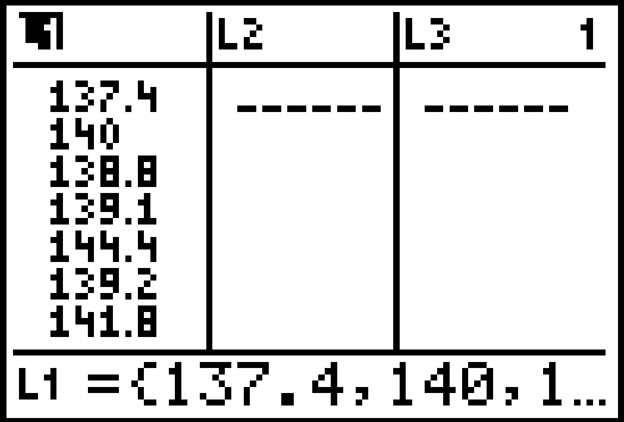
ANOVA- Analysis of Variance

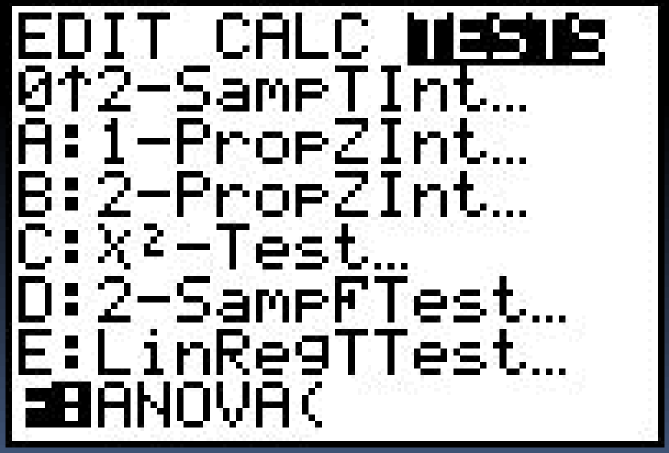
- Use ANOVA if you have multiple "levels" or groups of data
- Can be done in Excel or on calculator

Time (weeks)	Root Growth (cm)			
	Control Group (0 mL HCl)	5 mL HCl	10 mL HCl	15 mL HCl
1	10	5.75	2.2	1.3
2	12.5	6.6	2.4	1.4
3	15.6	7.6	2.7	1.5
4	19.5	8.7	2.9	1.7
5	24.4	10.0	3.2	1.9

ANOVA Test on your calculator

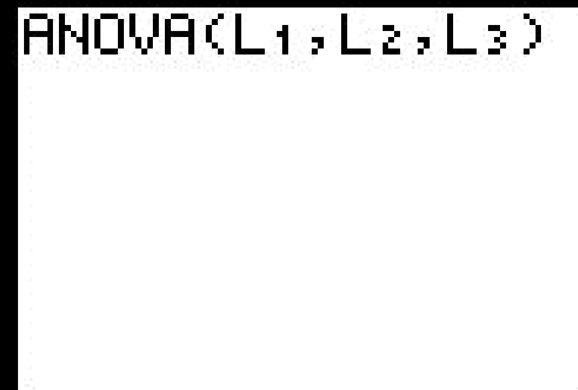
- Enter data into lists the same way as with the t-test
- Control group should be in L1
- Hit STAT button
- Use arrow keys to select F:ANOVA





ANOVA Test on your calculator

- You will see the word ANOVA with an open parentheses
- Hit 2nd STAT to get the LIST screen
- Hit 1 to enter L1
- Enter each list, separating each with a comma
- Close parentheses when you have finished
- Hit ENTER
- Record the p-value
  - The lower the p-value, the more certainty that you can REJECT the NULL hypothesis
- If $p \leq .05$, you can REJECT the NULL hypothesis and **accept the experimental hypothesis**



Bertram the Liger loves ANOVA