Statistical Analysis: Using IRP Data

Using your data: The first step is deciding which type of data you have, and what test would best make sense of the results. You need to use one of the statistical tests in order to show that there is a *statistically significant connection* between your data and your hypothesis; in other words, you will get a number value as to whether or not your data support your hypothesis.

You should already have your hypotheses written and your data organized into tables. Refer to the PowerPoint to help you decide which test to use. In general, if you have different "levels" of variables (different tanks with salinities or pH values) you could use a chi-square test. If you are comparing one control to one experimental group, you would use a t-test, and for multiple groups you would use an ANOVA.

The following sample data will help you use the tests:

Table 1: Growth of *Amoeba proteus*

Time (days)	# individuals (control)	# individuals (luria broth)	# individuals (hayseed)
0 (initial)	100	100	100
1	101	110	115
2	104	121	132
3	107	133	152
4	112	146	175

Part I: T-tests and ANOVA Tests

a. T-tests are used to compare the variation between *two groups*

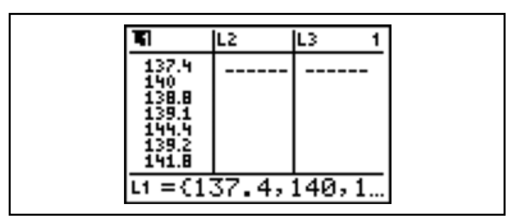
Possible groups:

Control vs. experimental

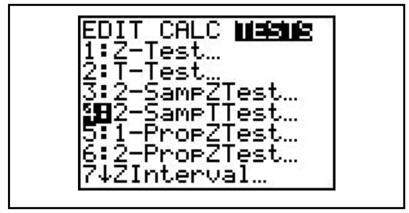
Experimental vs. experimental

- ➢ You can run a t-test on your calculator (TI-83 and above) or in Microsoft Excel. Follow the steps below to do the test on a TI-83 or similar.
  - ✓ Put data into two lists
  - ✓ *Careful not to put in time if it's not a variable!*

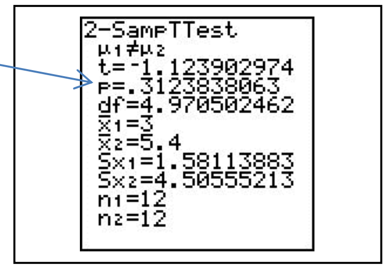
  To enter data into a list:
  - ✓ -Hit STAT button
  - ✓ -Select option 1, Edit
  - ✓ -Enter your values, hitting return after each
  - ✓ -L1 should be the control group
  - ✓ -Hit the right arrow key to move to List 2 and enter the experimental values there



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

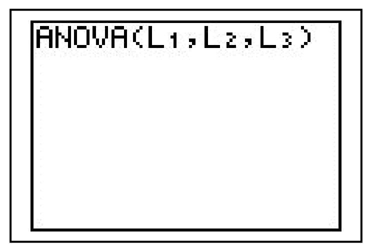


  - ✓ Select option 4, 2 sample t-test
  - ✓ Under Inpt, select DATA
  - ✓ L1 and L2 should be set as List 1 and List 2
  - ✓ µ1 should be : ≠µ2
  - ✓ Pooled: NO
  - ✓ Hit CALCULATE
- On the results screen, the number you are looking for is the "p-value." *This is the probability that your results are NOT significant and they are due to chance. You "want" the p-value to be as small as possible.*
  - ✓ 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 alternate (experimental) hypothesis**
    - o A p-value of less than or equal to .05 means there was a 95% or greater chance that your independent variable DID effect your dependent variable.



<u>b. Analysis of Variance (ANOVA):</u> Use ANOVA if you have multiple "levels" or groups of data. To perform an ANOVA on a TI-83:

- ✓ Enter data into lists the same way as with the t-test
- ✓ Control group should be in L1 and experimental groups in L2, L3, etc.
- ✓ Hit STAT button
- ✓ Use arrow keys to select F:ANOVA
- ✓ 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 and analyze the same way as with t-tests.



<u>Part II: Chi-Squared Tests</u>

THE BASICS:

1) Write hypotheses
2) Find expected data values
3) Find chi-squared (x^2)value
4) Use x^2 value to find p-value
5) Use p-value to evaluate hypothesis

Chi-squared tests are used to evaluate experimental hypotheses and determine if data are significant. Just like ANOVA and t-tests, chi-squared tests give a probability value that the independent variable DID NOT have an effect on the dependent variable. For all statistical analysis tests, the rule is that the p-value must be ***less than or equal to .05*** in order for the experimental hypothesis to be accepted. In other words, there must be a 5% or less chance that the IV had no effect if the experimental hypothesis is to be accepted.

1) Doing a chi-squared test must begin with writing the two hypotheses. You will need to be able to write an experimental and null hypothesis for the given situation. The data below shows the numbers of frogs of assorted tastiness eaten by some sort of frog-eating bird. The purpose of this test would be to determine if the delicious chocolate frog is more prone to being eaten that the other species.

   NULL HYPOTHESIS (H_0): There will be no significant difference between the numbers of each frog species eaten.
   OR
   The taste of a frog species has no effect on the amount of frogs eaten by the predator.

   EXPERIMENTAL HYPOTHESIS (H_A): It was hypothesized that the delicious chocolate frog will be eaten most frequently by the predator species.

	Observed	**Expected**
Gross poison frog	5	
Regular green frog	10	
Pretty blue frog	15	
Delicious chocolate frog	20	
Total Tasty Frogs	50	50

2) In order to perform the test, you must determine the expected values. Observed values are the experimental data. Expected values are not actually seen, but are instead the data that "would be" seen if there was really no difference between the species. In this case, if the tastiness of a species really had no effect than all of the species should have the same number eaten. The expected values would therefore be 50/4, or 12.5 each.

	Observed	**Expected**
Gross poison frog	5	12.5
Regular green frog	10	12.5
Pretty blue frog	15	12.5
Delicious chocolate frog	20	12.5
Total Tasty Frogs	**50**	**50**

3) You can now use the chi-squared equation to find the chi-squared value.

$$\chi^2 = \Sigma \frac{(observed - expected)^2}{expected}$$

$$\chi^2 = \frac{(5-12.5)^2}{12.5} + \frac{(10-12.5)^2}{12.5} + \frac{(15-12.5)^2}{12.5} + \frac{(20-12.5)^2}{12.5} = 10$$

4) You will now take the chi-squared (x^2) value and use it to find the p-value on the table provided. You need to know what row of the table to look on- the row will be the degrees of freedom, which is equal to the number of species (or data points) minus one. Since we have four species, there will be three degrees of freedom and you will look in that row to find the p-value. Find the numbers closest to your value (in this case 10) on that row.

Degrees of Freedom	Probability										
	0.95	0.90	0.80	0.70	0.50	0.30	0.20	0.10	0.05	0.01	0.001
1	0.004	0.02	0.06	0.15	0.46	1.07	1.64	2.71	3.84	6.64	10.83
2	0.10	0.21	0.45	0.71	1.39	2.41	3.22	4.60	5.99	9.21	13.82
3	0.35	0.58	1.01	1.42	2.37	3.66	4.64	6.25	7.82	11.34	16.27
4	0.71	1.06	1.65	2.20	3.36	4.88	5.99	7.78	9.49	13.28	18.47
5	1.14	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.07	15.09	20.52
6	1.63	2.20	3.07	3.83	5.35	7.23	8.56	10.64	12.59	16.81	22.46
7	2.17	2.83	3.82	4.67	6.35	8.38	9.80	12.02	14.07	18.48	24.32
8	2.73	3.49	4.59	5.53	7.34	9.52	11.03	13.36	15.51	20.09	26.12
9	3.32	4.17	5.38	6.39	8.34	10.66	12.24	14.68	16.92	21.67	27.88
10	3.94	4.86	6.18	7.27	9.34	11.78	13.44	15.99	18.31	23.21	29.59
	Nonsignificant								Significant		

p-values

5) **You can now evaluate your hypothesis.** ***According to the results of the test, a chi-squared value of 10 indicates that the p-value is less than .05/ This means that the data is significant and the <u>alternate hypothesis can be accepted.</u>*** You may now also eat the tasty frog of your choice as a reward for completing this calculation. ☺