Ecology Lab Introduction

- I thought this was ecology—not statistics

Modern Ecology = Statistics

- Overkill or essential?
- Regardless, you need a strong grounding in basic stats

Three points to remember…

- First, **look** at the data!
- Second, see if you can apply the **'O test'**
- Third, **results come first**, statistics second

Analysis = graphics & stats

- Magnusson's (1997) two axioms:
  - If you cannot present the results of your statistical test in simple graphical form, do not trust the statistics.
  - Do the graph and the statistics. If the statistics agree with the graph, then publish the statistics. If the statistics do not agree with the graph, then publish the graph and throw out the statistics.

Today's lab

- Good graphical techniques
- Basic statistics
- A brief write-up

Time to collect some data

On the board, record your:

- Height (in cm)
- Arm span (in cm) from left to right hand
- Academic class (e.g., junior, senior)

Look at the data—Graph it

- In groups, everyone should:
  - Copy the data from the board to an Excel spreadsheet
  - Calculate average height for each academic class =average(x1:x2)
  - Use a **column chart** to compare the mean height of one academic class vs. another (usually known as a bar chart)
  - Use **x-y scatter plot** graph to plot heights vs. arm spans for entire class (<u>all data, not means</u>)
  - Make graphs look as good as you can; you'll have a chance to edit them later

Graphical tips (1)

- Above all else, **show the data**
- Avoid distortion-- don't lie with numbers
- For *written* reports, use a figure caption; for *oral* reports, use a **descriptive** title

Graphical tips (2)

- Use space efficiently
- Draw axes to fit data
- Use 5 to 10 tick marks on the axes (fewer can sometimes be ok)
- Make sure the data points are big enough

Graphical tips (3)

- Label axes clearly and legibly (units!)
- Use labels on the graph itself, rather than putting a legend in the margin
- Use space efficiently

Graphical tips (4)

- Tufte's tips
    - Maximize the data-ink ratio
    - Erase non-data ink
    - Erase redundant data ink
    - **Revise and edit**

Two examples (1)

Two examples (2)

Now, some stats…

- Descriptive statistics
- Comparing groups
- Examining relationships

Descriptive statistics

- Mean or average: central tendency
- Standard deviation: dispersion about a mean

Comparing groups

- Comparing the means of <u>two</u> groups: **t-test**
- Comparing the means of <u>more than two</u> groups: analysis of variance (**ANOVA**)
- Recall, complete analysis: graphs AND stats…
  - Comparing groups: **bar charts** (= column charts in Excel)

Examining relationships

- 'Cause-effect' type of relationship between two variables: **regression**
  - Also: prediction, explaining variation
- Recall, complete analysis: graphs AND stats…
  - Relationships: **x-y scatter plots**

Significance

- In science, you <u>only</u> use the word '**significant**' if you have done statistical tests to see if your results are different than what would be expected from random
- So how do you determine significance from statistical output?
  - Usually through the use of 1 or 2 pieces of statistical information:
    - Calculated and critical test statistics
    - P-values

Interpreting significance: test statistics

- A result is significant if:
  - *The calculated test statistic is **greater** than the critical (or theoretical) test statistic*
  - Example:
    - T_{calc} = 5.34, T_{crit} = 1.99; Because $T_{calc} > T_{crit}$, there is a significant difference between the two means
- The values for the critical test statistics depend on:
  - the test being run,
  - the sample size,
  - and the probability level being used (usually $\alpha = 0.05$)

Interpreting significance: P-values

- A result is significant if:
  - *The computer-calculated P-value is **less** than a pre-defined α-value (usually 0.05)*
  - Example:
    - P-value = 0.01; Because the P-value < 0.05, then there is a significant relationship between the two sets of data
- P-values, or probability values, describe the probability that your results are due to chance
- The pre-defined α-value can vary depending on the needs of the researcher, but scientists often settle on 0.05 (i.e., reaching the wrong conclusion 5% of the time) as being an acceptable chance of error

Easier to observe significance if…

- You have a well-designed study
- You have many replicates
- You have *little variability within* the treatments or factors in your study
- You have large differences between the variables of interest (i.e., *more variability between* the treatments or factors)

Today's report

- Each group submits one column chart and one x-y scatter plot—edit your first attempts
- Briefly summarize the results and determine which test you would use to statistically examine the data shown in each graph