

Statistics with GraphPad Prism

Isn't it time to switch from Excel?

BaRC Hot Topics
Fran Lewitter
2/24/09

Features of Prism

- Prism is a powerful combination of biostatistics, curve fitting (nonlinear regression) and scientific graphing in one comprehensive program.
- Free Prism viewer (share results with non-Prism users)
- The Pros and Cons of Using Excel for Statistical Calculations (<http://www.graphpad.com/faq/viewfaq.cfm?faq=1406>)

Prism vs Excel: Excel deficiencies

- No automatic error bars on XY graphs--user can write functions to compute error bars and to graph them on XY graphs, but with difficulty
- No error bars on bar graphs
- No nonparametric statistics
- No normality testing
- No survival analysis
- No Fishers exact test

Prism vs Excel: Excel deficiencies

- Nonlinear regression is very difficult
- No standard errors or confidence intervals from nonlinear regression
- No statistical teaching on Help screens
- No analysis of survival curves
- No box-and-whiskers plots
- User can't export graphs in other formats (tiff, jpg, PDF, EPS, etc.)

Ways to get help

- <http://www.graphpad.com/prism/learnwatch.htm>
- http://graphpad.com/help/prism5/prism5help.html?using_tour_overview.htm
- Choosing a statistical test (<http://www.graphpad.com/www/book/Choose.htm>)
- Guided examples Statistical Analysis
- Online help and Tutorials
- Use sample data (a guide through analysis)
- Interpret results

Statistical Comparisons

- Paired or unpaired t tests
- Mann-Whitney or Wilcoxon tests
- Ordinary or repeated measures one-way ANOVA with Tukey, Newman-Keuls, Dunnett or Bonferroni post tests
- Kruskal-Wallis or Friedman nonparametric one-way ANOVA with Dunn's post test
- Fisher's exact test or the chi-square test. Calculate the relative risk and odds ratio with confidence intervals
- Two-way ANOVA, even with missing values with some post tests
- Repeated measures two-way ANOVA with some post tests
- Kaplan-Meier survival analysis. Compare curves with the log-rank test (including test for trend)

Column Statistics

- Calculate min, max, quartiles, mean, SD, SEM, CI, CV, Geometric mean with Confidence Intervals
- Specify desired level of confidence
- Frequency distributions (bin to histogram), including cumulative histograms.
- Kolgoromov-Smirnov normality test
- One sample t test or Wilcoxon test to compare the column mean (or median) with a theoretical value
- Skewness and Kurtosis

Linear regression and correlation

- Calculate slope and intercept with confidence intervals
- Force the regression line thru a specified point
- Compare slopes and intercepts of two or more regression lines
- Determine new points along the standard curve
- Pearson or Spearman (nonparametric) correlation

Contingency Tests

- 2 x 2 contingency table
- Comparing observed and expected counts
 - Use the chi-square test - χ^2
- Comparing two proportions
 - Use Fisher's Exact Test or the chi-square test - χ^2
 - Fisher's Exact Test is recommended

Contingency Tests

- **How the data are organized?** The columns represent presence or absence of HIV antigen among patients with symptoms suggestive of HIV infection. The rows represent the results of a simpler test. The values are the number of subjects in each group. Data from: Daar et. al., Annals of Internal Medicine, 134:25-29 (2001).
- **The goal** To quantify the sensitivity (what fraction of people with the disease are identified by the test) and specificity (what fraction of healthy people have a negative test result), with confidence intervals.

Contingency Tests

How to analyze the data?

1. Click Analyze
2. Choose "Chi-square (and Fisher's exact) test" from the list of analyses for contingency tables.
3. Click OK.
4. Choose Fisher's exact test and check the option to compute the sensitivity, specificity and predictive values.

One Sample t-Test

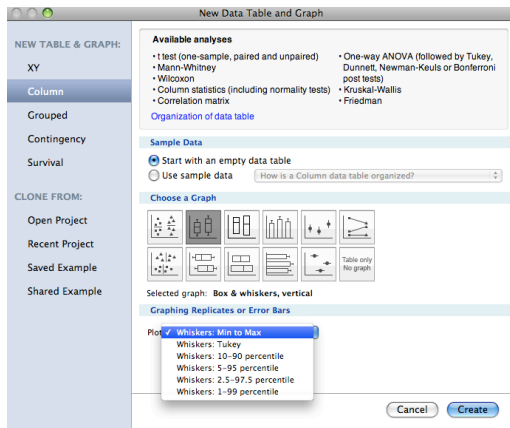
- **How the data are organized?** Each value is the result of one experiment, expressed as percent of control.
- **The goal.** If the experimental treatment was ineffective, the outcomes would all be about 100% of control. That is the null hypothesis. A one sample t test will determine if the mean of the data are further from that null hypothesis than you'd expect due to chance.

One Sample t-Test

- **How to analyze the data**
 1. Click Analyze, and choose Column Statistics from the list of Column analyses.
 2. On the Column Statistics dialog, check the option box to perform a one-sample t test. To do that test, you must enter a hypothesis. Enter 100.
 3. Review the rest of the dialog, and check any tests you want to perform, perhaps a normality test.

BoxPlot

- **Purpose:** Quantifying Variability with Percentiles
- **How column data are arranged?** Each column defines a group. In this example, there are five groups. Groups are defined as columns.
- Error bars come from a stack of values. If your graph includes an error bar, it will be computed from the scatter of values stacked in each column.



Box & Whisker Plots

There are five ways to plot box-and-whisker graphs. Traditional Tukey whiskers go 1.5 times the interquartile distance or to the highest or lowest point, whichever is shorter. Other choices show specified percentiles, which seems more logical. Any data beyond these whiskers are shown as points. Because whiskers can be done in various ways, it is essential to mention in a graph's figure legend which way they were made.

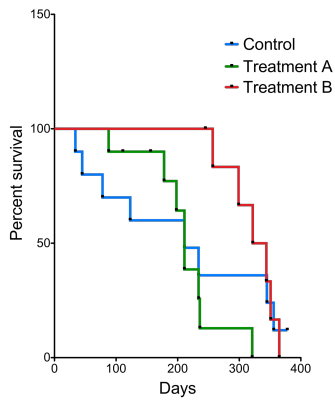
Survival Analysis

- Can't do with Excel
- **How the data are organized?**
 - Each row represents one subject. The X values are time. The Y values are entered into two (or more) columns that define the treatment groups. The Y value is "1" when the subject died at the specified time, and "0" when the subject's data was *censored* at that time (either because we don't know what happened after that time, or we do know but can't use the data because the protocol was not being followed).

Survival Analysis

- **The goals-** To create a Kaplan-Meier survival curve.- To determine whether the difference between survival curves is more than expected by chance.
- **Viewing the results.** Survival analysis is unique. You don't need to click the Analyze button because Prism automatically analyzes survival data. Simply view the linked results sheet and graph. Click below to learn more about survival analysis.

Survival of Three groups: Survival proportions



Other software @ WI for stats and graphing

- Igor Pro
- Kaleidograph
- Matlab
- R

Statistical expertise

- BaRC
- Statistical consultant @ HMS
- Books to loan

What's next?

- Download GraphPad Prism from the Whitehead Software Database
- Demo
 - Getting started, help info
 - Contingency Tests
 - One sample t-Test
 - BoxPlots
 - Kaplan-Meier survival analysis
 - Working with graphs
 - Cloning graphs