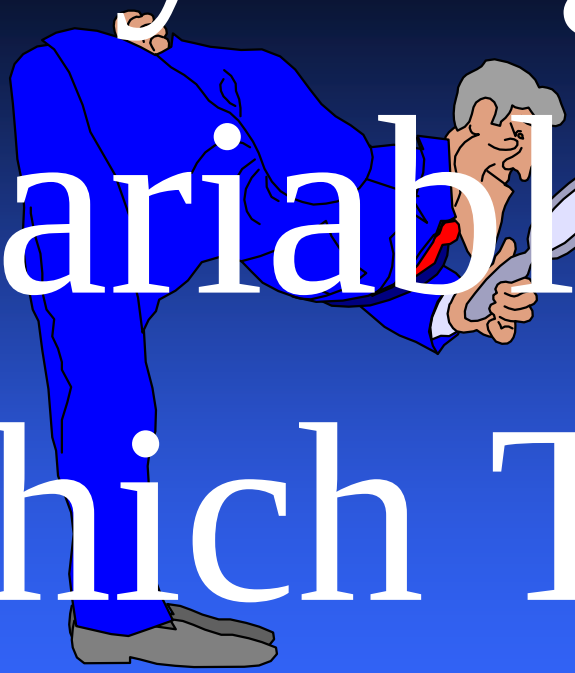
Analyzing the Variables: Which Test When?



Jargon

- We sheeted in the main and buried the lee rail. We made the mark without having to tack.
- He faded a wedge around the overhanging tree and cozied the ball right up to the hole.
- Oh, man! That was a great wreck!

A Language With a Purpose

- Describe data
- Compare groups of data

Significant Words and Phrases

- p-Value (alpha error)
- Power (1-beta error)
- Sample size
- Normal distribution
- Clinical significance
- Statistical significance

Quantitative Data

- Measured Numerically
  - Ratio data has a true zero
    - ex: # of cats owned by people in this room
  - Interval data does not have a true zero
    - ex: heart rates of people in this room

The Quantitative Dialect

- Mean
- Median*
- Standard Deviation
- Range
- Confidence Intervals*

Quantitative Tests

- Independent (or Student's) *t*-Test
  - means of two groups
- Dependent (or paired) *t*-Test
  - difference in means of two related groups
- Analysis of Variance (ANOVA)
  - means of more than two groups
- Confidence Intervals
  - accuracy of estimates of the mean

Qualitative Data

- Measured Descriptively
  - Nominal data can be specifically grouped
    - ex: people are mammals; frogs are amphibians
  - Ordinal data can be assigned to groups that have relative (but not absolute) value
    - ex: coffee is hot; McDonald's coffee is really really really hot

The Qualitative Dialect

- Frequency
- Rank
- Median*
- Agreement
- Confidence Intervals*

Qualitative Tests

- Chi-Square / Fisher's Exact Test
  - frequencies in two or more groups
- Wilcoxon Test / Kruskal-Wallis Test
  - median rankings of two (or more) groups
- Kappa
  - agreement
- Confidence Intervals
  - accuracy of estimates of proportions

Other Statistical Manipulations

- Repeated Measures
- Linear Regression
- Logistic Regression

Confidence Intervals

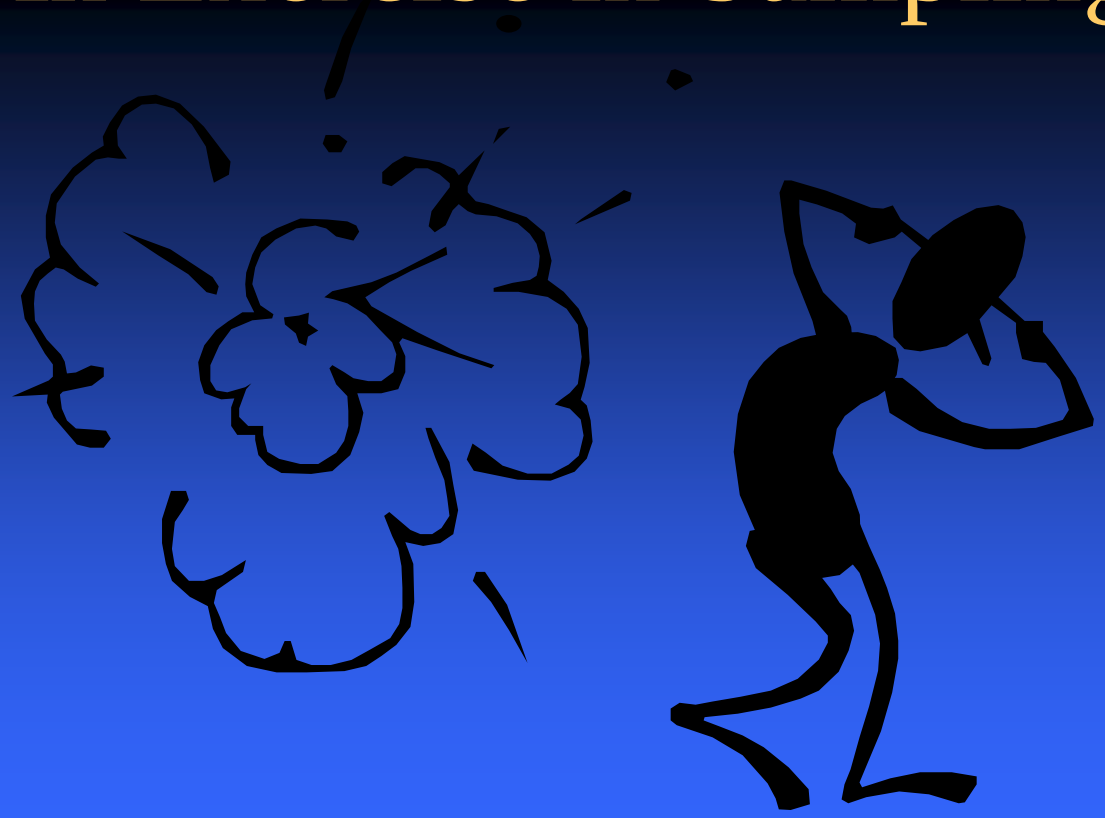
- 95% CI for a mean

  mean +/- 2.0 x ($sd / n^{1/2}$)

- 95% CI for a proportion

  proportion +/- 1.96 x ($sd / n^{1/2}$)

An Exercise in Sampling



Random Sample

RED		BLUE	
01	24	05	21
04	34	06	22
05	35	09	23
07	39	11	27
08	40	12	37
09	41	15	39
22	48	17	45
23	51	20	52

Reality

- Mean value of cards = 7.3
- SD of deck = 3.2
- Red vs. Black = 50%
- Median value = 5

Summary

- Statistics are about language and chance
  - Statistics describe data
  - Statistics compare data
- All sampling has error / bias
- EBM constructs can be helpful
- Don’t be afraid …. but ….
- Don’t be afraid to ask for help!