Biostats and Biases

Alex B. Haynes, MD, MPH, FACS
Division of Surgical Oncology, Massachusetts General Hospital
Associate Director for Safe Surgery, Ariadne Labs
Research Associate, Harvard TH Chan School of Public Health
October 21, 2017

Disclosures

- No relevant disclosures

What is “statistics”?

Merriam-Webster: a branch of mathematics dealing with the collection, analysis, interpretation, and presentation of masses of numerical data

Google: the practice or science of collecting and analyzing numerical data in large quantities, especially for the purpose of inferring proportions in a whole from those in a representative sample.

synonyms: data, facts and figures, numbers, information, details

Why Do Surgeons Need to Know About Statistics?

AAS Association for Academic Surgery

- Understand literature for clinical care
- Develop appropriate research questions and approaches
- Communicate with statisticians and other scientists

Goals of this talk

- Will help with some vocabulary
- Demonstrate some skills for communicating intelligently with biostatisticians, analysts, and others
- Encourage you to learn more

Disclaimers

- I am not a biostatistician
- Will not teach you to do statistics today
- There will be almost no numbers or complex equations presented
- No test at the end

From observation to understanding

- Observations create data
  - Experimental
  - Ecological
- Data can be truth but not necessarily so
  - Role of chance
  - Role of bias
  - Reliability of observations

Define your data

- Continuous
- Categorical
  - Binary
  - Ordinal
  - Nominal
- Dependent vs Independent Variables (e.g. Outcome vs Predictor)

Definitions

- Null hypothesis: there is no difference in outcome between the two exposures
- Type I error = false rejection of the null hypothesis (α)
- Type II error = false acceptance of the null hypothesis (β)
- Power = 1- β

When should you engage with a statistician?

- Early and often!
- When writing protocol or proposal
- Design analysis ahead of data collection
- Ensure that methods are able to answer the question that you have
  - Power
  - Types of data

How can my study be powerful?

- Power = 1- β
  - Set at 80 (pilot) – 95% (final, conclusive study)
  - 90% is accepted power for NIH funding
- Goal is to minimize both Type I & II error to maximize the chance that observation is true

Power/Sample Size Calculations

- You need 3 of the following 4:
  - Sample size (N)
  - Power (1 – β)
  - Pre-specified surety (α)
  - Effect size
- What is effect size?
  - Difference between 2 means, times, or proportions
  - Based on pilot data, best guess, "minimum clinical significance"

Power/Sample Size Calculations

- *A priori* analysis
  - Determine the sample size necessary to detect a clinically important effect with pre-specified surety
  - Determine the chance that with a given sample size, a clinically significant effect can be detected
  - Determine the magnitude of the effect size that can be detected with a pre-specified surety and sample size
- *Post hoc analysis*
  - Once the study is done, determine the probability that the conclusion is correct

*Orav, 2002

What is significant?

- Statistical significance
  - Data are extreme enough to reject null hypothesis at present level of significance (α), typically 0.05
- Clinical significance
  - Is the level of difference identified meaningful for clinical use
  - Varies by instance
  - Take into account imprecision and generalizability
  - Different for medicine and public health

0.05 or bust!

- A p-value is the conditional probability of obtaining the results in hand if the null hypothesis is true
- In plain English, how likely are you to get these results if there is truly no difference?
- Significance (α) is set *a priori* and represents the likelihood of falsely rejecting the null hypothesis

Some Basics for Analysis

- Is the outcome categorical or continuous?
- Is the predictor categorical or continuous?
- Is the data paired?
  - Sequential measurements made on the same subject
    - HR before and after medication X is given
  - Need to account for the fact that these are likely to be more highly correlated
  - If do not do paired analysis, you can introduce bias into significance of results

Some Basics for Analysis

- Is the data normal?
  - If yes, then
    - Parametric tests = Assumes that the data are drawn from a normal distribution
  - If no, then
    - Non-parametric = No such assumption
      - Based on the rank order
    - Transform data to make it normal and use a parametric test
      - Logarithm, square root, reciprocal

Fig 1 Normal

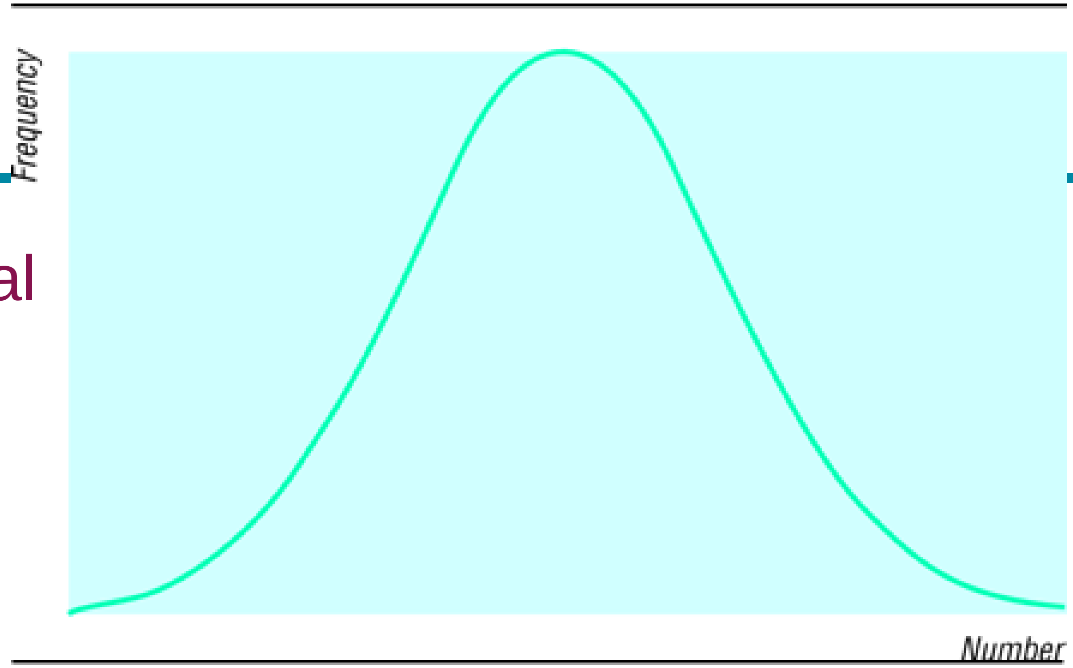
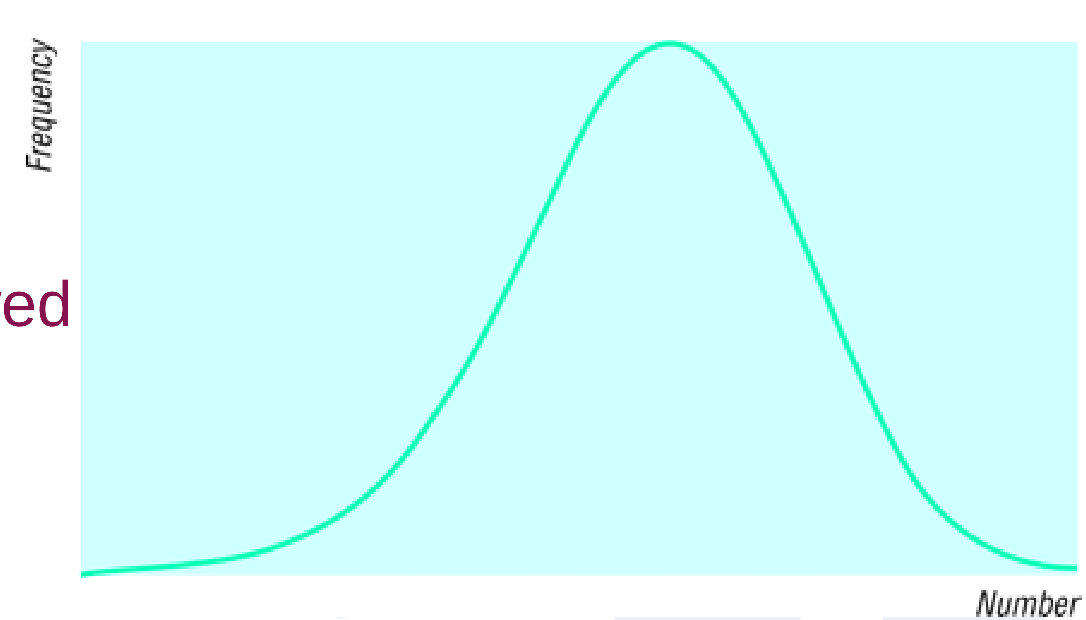


Fig 2 Skewed



Greenhalgh, BMJ 1997;315:364-366

Univariate Analysis

- Single outcome with single predictor
- Many different tests, depending upon:
  - Number of groups
  - Relationship of groups
  - Nature of outcome and predictor (discrete, normal, non-normal)
  - Examples: Chi-Square, Fisher exact, t-test, Wilcoxan Rank Sum, ANOVA, Spearman, Pearson, Kruskal-Wallis, McNemar

	1 Sample (Paired)	2 Samples	3 or more Samples	Continuous Predictor
Discrete Outcome	McNemar’s Test	Fisher/Chi -square	Fisher/Chi-square	T-test/ ANOVA
Normal Outcome	Paired t-Test	t-test	ANOVA	Pearson Spearman
Non-Normal Outcome	Wilcoxon Sign Rank	Wilcoxon Rank Sum	Kruskal-Wallis	Spearman

Multivariable modeling

- Build a model that relates the predictors to the outcome through a mathematical equation
  - Adjust for confounders
  - Build a prediction model
- Types of multivariable modeling
  - Linear regression
  - Logistic regression
  - Cox regression

Validity

- Internal validity – do the results represent a real relationship?
- External validity – can the study be applied to "real life"?

Biases

- Selection bias
  - Selection affects external validity
  - Selection bias affects internal validity
- Information bias
  - Misclassification bias
    - Differential misclassification changes findings
    - Random misclassification reduces power
  - Recall bias
  - Reporting bias
  - Surveillance bias

Confounding

- An exposure is a confounder if:
  1. It is a risk factor for disease and
  2. It is associated with A but not a result of A
- Example
  - Coffee drinkers have a higher rate of pancreatic cancer
  - However, this is because cigarette smoking, a known risk factor for pancreatic cancer, is more prevalent amongst heavy coffee drinkers

Approaches to confounding

- Matching
- Stratification
- Adjustment
- Capture of additional data

Statistics are a lens

- Analysis isn't something that happens after data collection
- Early input from statistical colleagues important
- Understanding biases and confounding essential to interpreting results
- Many opportunities to expand understanding

Thank You

- abhaynes@mgh.harvard.edu

Use this slide if you need to transition between topics.

Color Palette

Association for Academic Surgery

Pantone 7435 R-134 G-18 B-78	
Pantone 7455 R-18 G-104 B-179	Pantone 7527 R-226 G-221 B-203
Pantone 640 R-0 G-136 B-164	Pantone 5523 R-196 G-223 B-226
Pantone 2607 R-81 G-43 B-139	Pantone 650 R-199 G-217 B-234
Pantone 2746 R-26 G-59 B-142	Pantone 428 R-197 G-203 B-209