Clinical Statistical Communications

Perry K. Parendo
651-230-3861
Perry@PerrysSolutions.com

Objectives/ Agenda

- **Understand statistical analysis section**
- **Continuous versus categorical variables**
- **Overview of basic commonly used stats tools**
  - **How to read and understand basic hypothesis testing**
  - **What does "P" mean?**
  - **What it means when something is significant**
- **How do you get the "N"?**
- **Key questions when reviewing data**

Types of Data - Categorical

- **Data can come in different forms that can be treated differently. There are 4 different types of data**
- **Two are categorical. (easy to use, but not real powerful, need larger sample size for same significance)**
- **Nominal Data**
  - **This is yes/ no information. Did the patient live or die? This is discrete information, there is no continuum.**
  - **Keeps artery open**
  - **Complication free**
  - **No debris in blood**
- **Ordinal Data**
  - **Ranked information that has no mathematical meaning.**
  - **Examples**
    - **Grades A, B, C, D**
    - **Pain ranked from 1-10**
    - **Fewer clots**

Things that are counted.

Types of Data - Continuous

- **The other two types of data are Continuous data. It is more powerful and has more validity**
  - **Interval Data**
    - **Continuous and valid information**
    - **The size of the difference has calibrated meaning**
    - **May have an arbitrary zero point; Ratios have no meaning**
    - **Example: Temperature (F or C)**
  - **Ratio data**
    - **Continuous data where size and ratio are meaningful**
    - **Examples**
      - **Age**
      - **Linear dimension**
      - **Heart rate**
      - **Absolute temperature (R or K)**
      - **Ease of insertion**

Continuous Distributions

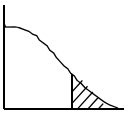
- **Descriptive Statistics - variable**
  - **Central Tendency**
    - **Sample mean (sum all/ number of values)**
    - **Median (middle value)**
    - **Mode (value occurring most often)**
  - **Spread**
    - **Range (maximum – minimum)**
    - **Variance (s^2, additive property)**
    - **Standard deviation or s (measure of spread)**

Continuous – Tools - Categoric Discrete

- **t-test**
  - **Z test**
- **F-test**
- **Proportions**

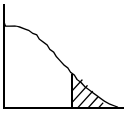
t-test

- **What is it**
  - **Comparison of means for sample data if normally distributed**
  - **Use it to compare two sample means within a confidence zone**
  - **Test of Hypothesis**
- **Why use it**
  - **Determine if means are similar relative to data spread**



t-test

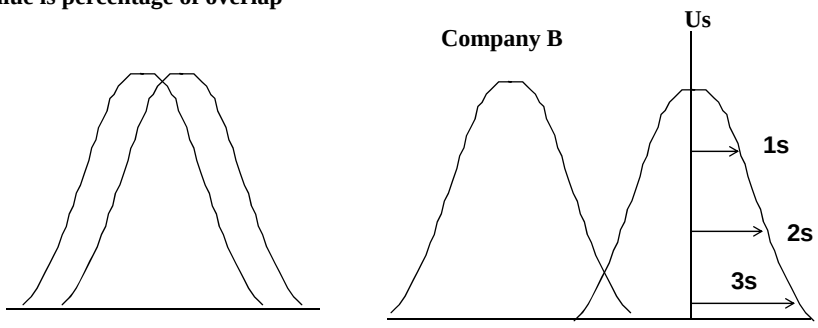
- **When not to use it**
  - **Non normal distribution**
- **Weaknesses**
  - **Practically significant**



t-test Pictorially

P-value is percentage of overlap

"We are better than company B."



same mean

different mean, "significant"

t-test: are means equal or not

How to Interpret p-values

- **Percent chance of misreading the null hypothesis**
  - **"Null" means "none" or no difference (i.e. same mean)**
- **Typical Rule of Thumb - less than 0.05 (5%) is significant, thus means would be different**
- **If not below threshold, implies they are the same**
  - **Actually only says "not enough evidence to say they are different"**
- **What are boundaries?**
  - **0.01 and 0.10 are sometimes used**

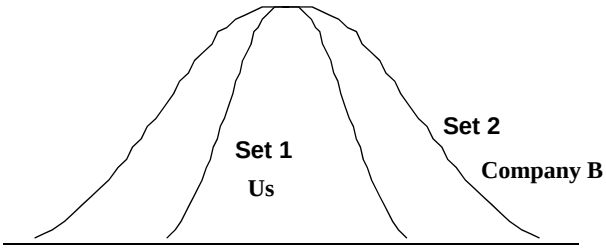
Don't just look at numbers, look at a graph too!

F-test

- **What is it**
  - **Two areas of use**
  - **Comparison of variances for sample data**
  - **Comparison of means for sample data**
  - **Sometimes also called "ANalysis Of VAriance" – or ANOVA**
- **Why use it**
  - **Know if a data set is different than random noise**
  - **Comparison of unequal sample sizes**

F-Test on Variance Pictorially

"We are more consistent compared to company B."



different variance (spread), "significant"

F-test: are variances equal (two sided)

Proportion test

- **What is it**
  - **Test for binomial data, categoric data**
- **Why use it**
  - **See if 2 percentages are the same**

"No more then 5% Adverse Events"

Sample Size

- **What is it**
  - **Representative of population**
- **Why use it**
  - **Size and cost the clinical study**
- **Weaknesses**
  - **Sampling error**
  - **Valid, reliable, unbiased**
  - **When do you know you are right?**

Sample Size and Sampling Plans

- **Many calculators exist to generate sample size**
- **Depends on what you desire to have confidence and power in**
- **Depends on your history with that product, if any**
- **Currently, interest is on the rise**

Drop-outs and non-compliance

- **There will be patients who start a study that will not finish it for a variety of reasons:**

  (Remember, anyone can drop out at any time; that is their right as a patient)
  - **Move away**
  - **Die of unrelated causes before completing study**
  - **Tired of examinations**
  - **Change in status were patient meets exclusion criteria**
  - **Does not comply with the intervention protocol for any reason**

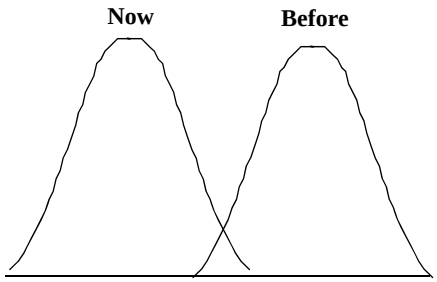
Sample Size

- **The thought pattern is:**
  - **Determine objective**
  - **Determine hypothesis with outcome (delta)**
  - **Know historical studies as much as reasonable (sample size and standard deviation estimates)**
  - **Understand protocol effect on potential standard deviation for your study**
    - **Can exclusion criteria make your effect easier to see?**
  - **Select confidence and power (alpha and beta)**
  - **Calculate sample size (using tables)**
  - **Execute protocol with adjusted sample size (for drop outs)**
  - **Use statistical tools**

Hypothesis Testing

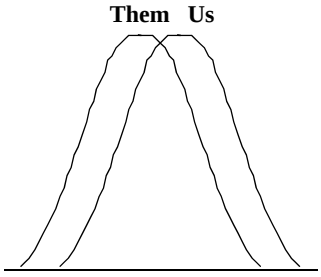
- **What questions are you trying to answer?**
- **Question 1: We have a shorter recovery time than previous products**
  - **Example: p=0.03 would be good**
- **Question 2: We are as safe as competitor**
  - **Example: p=0.55 would be good**
- **Statistically significant does not always mean "good"**

Question 1 – Recovery Time



different mean

Question 2 - Safety



same mean

Key Questions

- How do you know it is normally distributed?
- Did you perform residual analysis?
- Do you have a plot of the data?
  - Does data and predictive lines match?
- So the p-value is good, but is it practically significant?
- How did you account for potential bias?
- Could the measurements be continuous?
- Did any outliers occur? How did you handle them?

Conclusion

- Asking a few key questions can quickly assess the information you are being presented
- Perspective on sample size and hypothesis testing can help level the playing field for discussions
- If you want to see the reference materials – visit our website
  - www.PerrysSolutions.com
  - If interested, email us to be on our quarterly newsletter where we share recent trends and learning points
    - Newsletters are all archived on our publications page