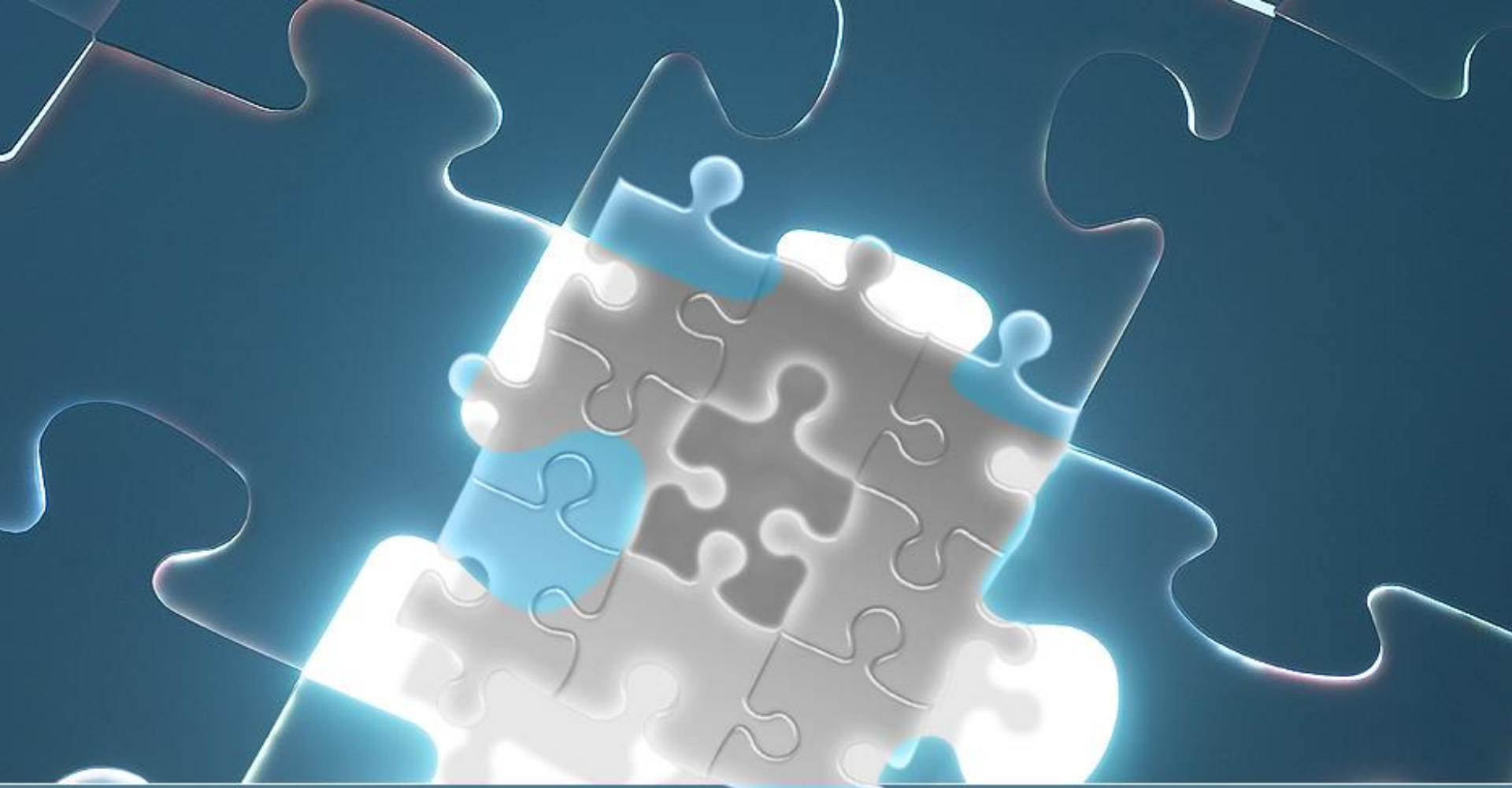
Study Design and Biostatistics for the Bedside Healthcare Professional.

Sherrine Eid MPH
Lehigh Valley Health Network, Sherrine.Eid@lvhn.org

Published In/Presented At

Eid, S. (2008, October). *Study Design and Biostatistics for the Bedside Healthcare Professional.* Presentation Presented at: Lehigh Valley Health Network, Allentown, PA.



Study Design and Biostatistics for the Bedside Healthcare Professional

Sherrine Eid, MPH
Research Day 2008
Biostatistician
Lehigh Valley Hospital and Health Network

Words of Wisdom

- *"[Biostatistics are] a whole lot of nothing!" (Gerard Dallal, 2000)*
- *"100% of all disasters are failures of design, not analysis." (Ron Marks, 1994)*
- *"Statistics is backwards!" (Gerard Dallal, 2000)*

Hypothesis Testing

- Begin with a Null and Alternative hypothesis
- Set the alpha level
- Perform the appropriate statistical test
- Calculate the p-value from the test statistic
- Decide if results are likely due to chance alone (not statistically significant) or unlikely due to chance alone (statistically significant)

Hypothesis Testing

- Uniform way to look at the results of the study.
- Null hypothesis H_O: There is no difference
- Alternative hypothesis H_A: There is a difference
- Since H_O and H_A are opposites we can just talk about H_O
- We talk about H_O by saying we 'accept' or 'reject' it
  - Accept means no difference or association
  - Reject means there is a difference or association

Setting it Up

- Sample Size
  - How many subjects are needed to truly answer the research question?
  - Is the study feasible?
  - How do we calculate it?
    - Hypothesis based
    - Confidence interval based

Sample Size Considerations

- Two fundamental concepts
  - Large differences between groups (if really there) are easier to find and take fewer subjects
  - Large variability in the variables you are studying make finding a difference harder and take more subjects

What I Need to Know to Compute Sample Size:

- A variance and mean assessment of the variable in question.
  - Can use old studies
  - Literature
  - Pilot study
- The Effect Size: What difference does the researcher consider significant?
- What p-value(s) the researcher wishes to use.

Effect Size

- Assume you are working in a diabetes program and the mean HgbA1c is 8.5, what does the effect size need to be to lower your patient population below risk?
- If the risk level is 7.5 then:
- 1-(7.5/8.5)=0.12 or 12 percent effect size is needed.

Power of a Statistical Test

- Power is the probability of rejecting the null hypothesis when the null hypothesis is false (1-ß).
- This is the correct decision. When the real world agrees with the statistics.
- Remember the table...

Power Analysis

- Each statistical procedure has its own formula for power determination.
- Effect Size - Difficult concept to teach because it is not determined by the statistician, but rather the clinician.
  - What do you consider meaningful?
- Must understand the difference between:
  - Clinical Significance
  - Statistical Significance

Significance

- Clinical - Change in a variable that has a physical advantage or disadvantage associated with it, i.e., an increase in risk of disease.
- Statistical - Mathematical change in a variable that achieves a probability threshold not due to chance alone.

Probability and Confidence

- "p values"-probability a particular outcome would have arisen by chance
- $p<0.05$ (1 in 20 chance)-statistically sig.
- $p<0.01$ (1 in 100 chance)- highly s.s.
- non-significant p value means no difference between groups or insufficient subject number-can't tell which

Step 1

- Descriptive Stats
  - Mean, Standard Deviation, Outliers
  - Summarize important features of numerical data
  - Identify errors
  - Characterized subjects
  - Assess assumptions for statistical tests

Step 2

- Analytical/Inferential Statistics
  - Looking at associations among two or more variables
  - Estimate pattern and strength of associations
  - Test hypotheses

How Do You Know
(some rules)

- The data that you have needs to be of the same quality as other studies
  - Collected at the same statistical level
  - Not viewed as secondary data
  - Needs reliability and validity indicators
  - Best if all the data reside in one system that is linkable

Statistical Tests

- Mathematical formulas that produce p-values that allow investigators to assess the likelihood that chance accounts for the results observed in the study.
- There are many different statistical tests.

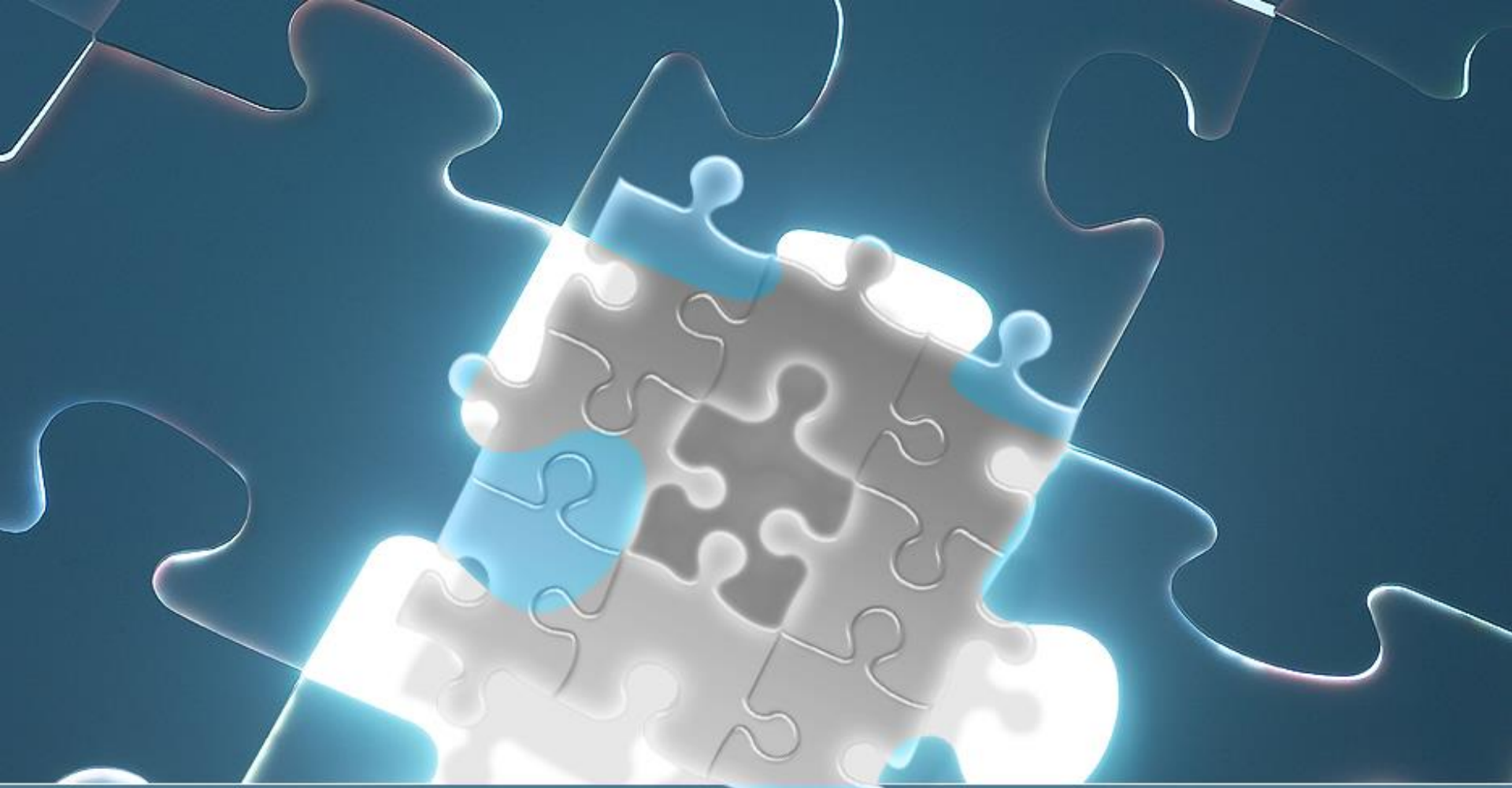
Which Statistical Test Do You Choose?

- This answer depends on:
  - The type of data
    - Continuous, Ordinal, or Nominal
  - The distribution of the data
    - Normally distributed, skewed
  - The type of study design
    - Means, proportions, number of groups, etc.

General Formula

- All statistical tests follow this general formula:

$$\text{Test Statistic} = \frac{\text{Observed val} - \text{Expected val}}{\text{Standard Error}}$$



Analysis of Continuous Data: Z and t Tests

One Sample Z Test

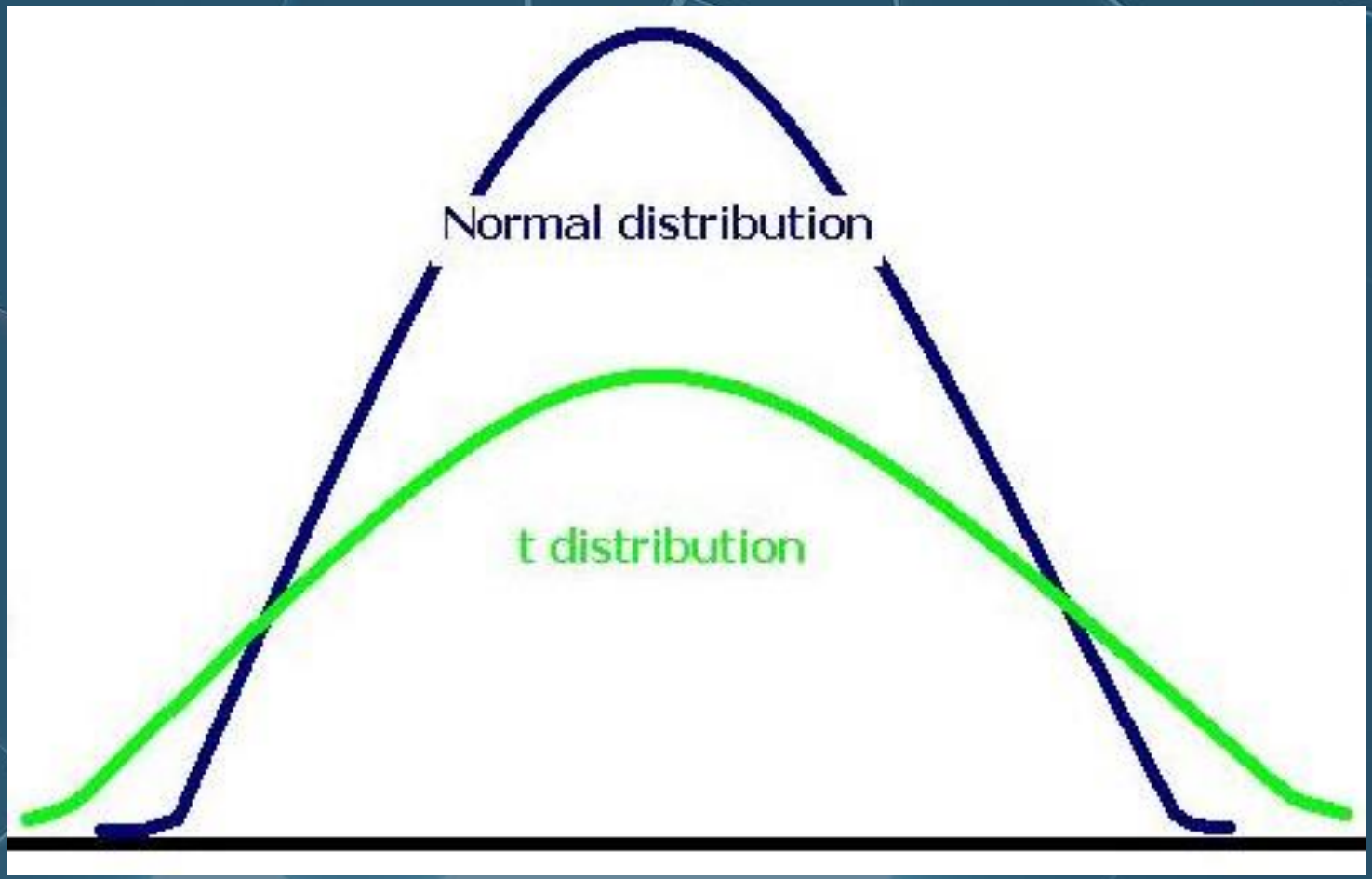
- Researchers are interested in whether the mean level of enzyme A in a certain population is different from 25. They measure levels of enzyme A in a sample of 10 individuals and find that the mean, $\bar{X} = 22$. Assume that the population has a known standard deviation, $\sigma = 6.7$
- H_0: $\mu = 25$ $\alpha = 0.05$

  H_A: $\mu \neq 25$
- Calculate test statistic: $Z = -1.41$
- Critical Value (from Z table): ± 1.96

t Tests

- Z test assumes population variance (or standard deviation) is known.
- When population standard deviation is unknown, but sample size is large (>30), then use sample standard deviation (s) to estimate population standard deviation and use normal distribution theory (Central Limit Theorem)
- If sample size is small and standard deviation is NOT known, use Student's t distribution theory.

Normal versus t Distribution



Properties of the t Distribution

- Family of distributions –different distribution for each sample value of n-1 (degrees of freedom)
- It has a mean of 0 and is symmetrical around the mean
- The t distribution approaches the normal distribution as n-1 approaches 30

One Sample t Test

- Example
  - Researchers are interested in whether the mean level enzyme B in a certain population is different from 120. They measure levels of enzyme B in a sample of 15 individuals and find that the mean =96 and the sample standard deviation is s=36.

Two Sample t Test

- Purpose is to compare the means of continuous variable in two independent samples.
- Example:
  - Researchers are interested in knowing whether people with diabetes have the same SBP as people without diabetes.

Paired t Test

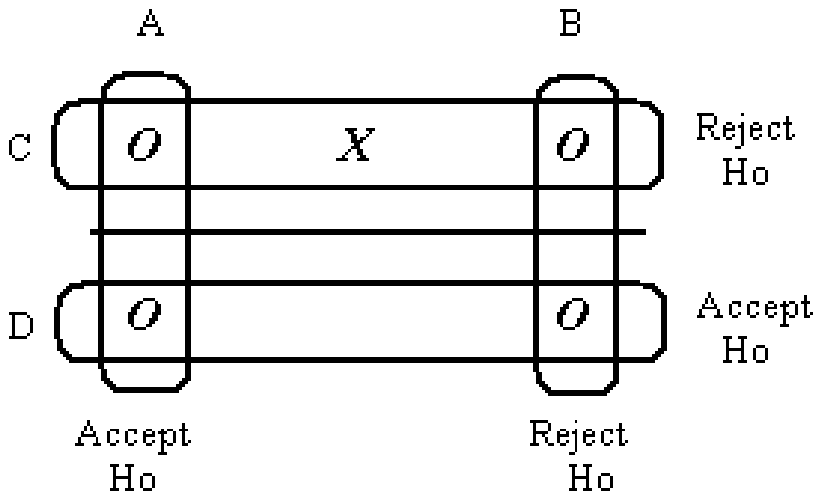
- Purpose is to compare means of two non-independent samples.
- Measurements on the same individuals before and after a treatment or intervention.
- Analysis performed on differences between individual pairs of observations.

Paired t Test

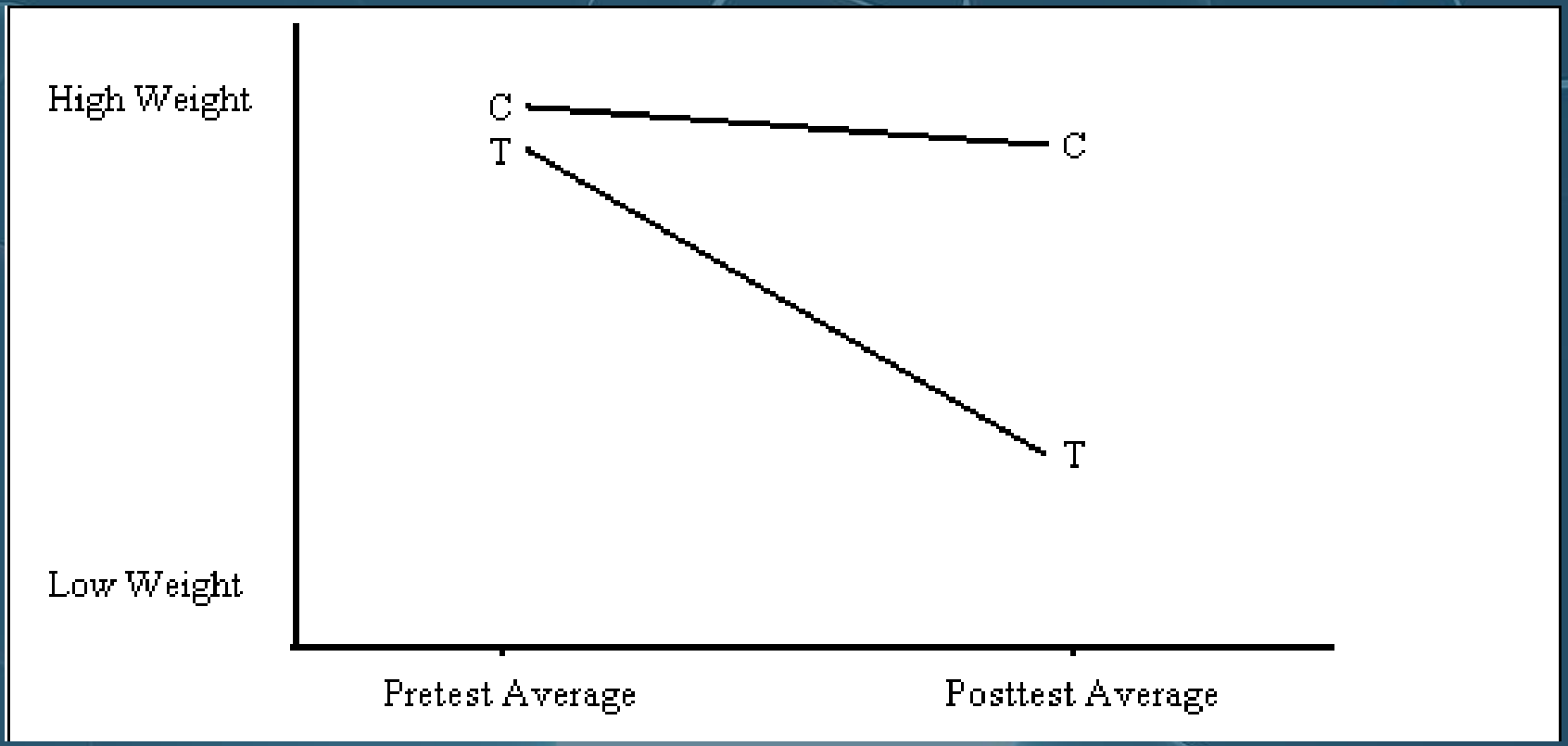
- Example
  - 12 subjects participated in a study on the effectiveness of a certain diet on serum cholesterol levels.

What We Hope to See

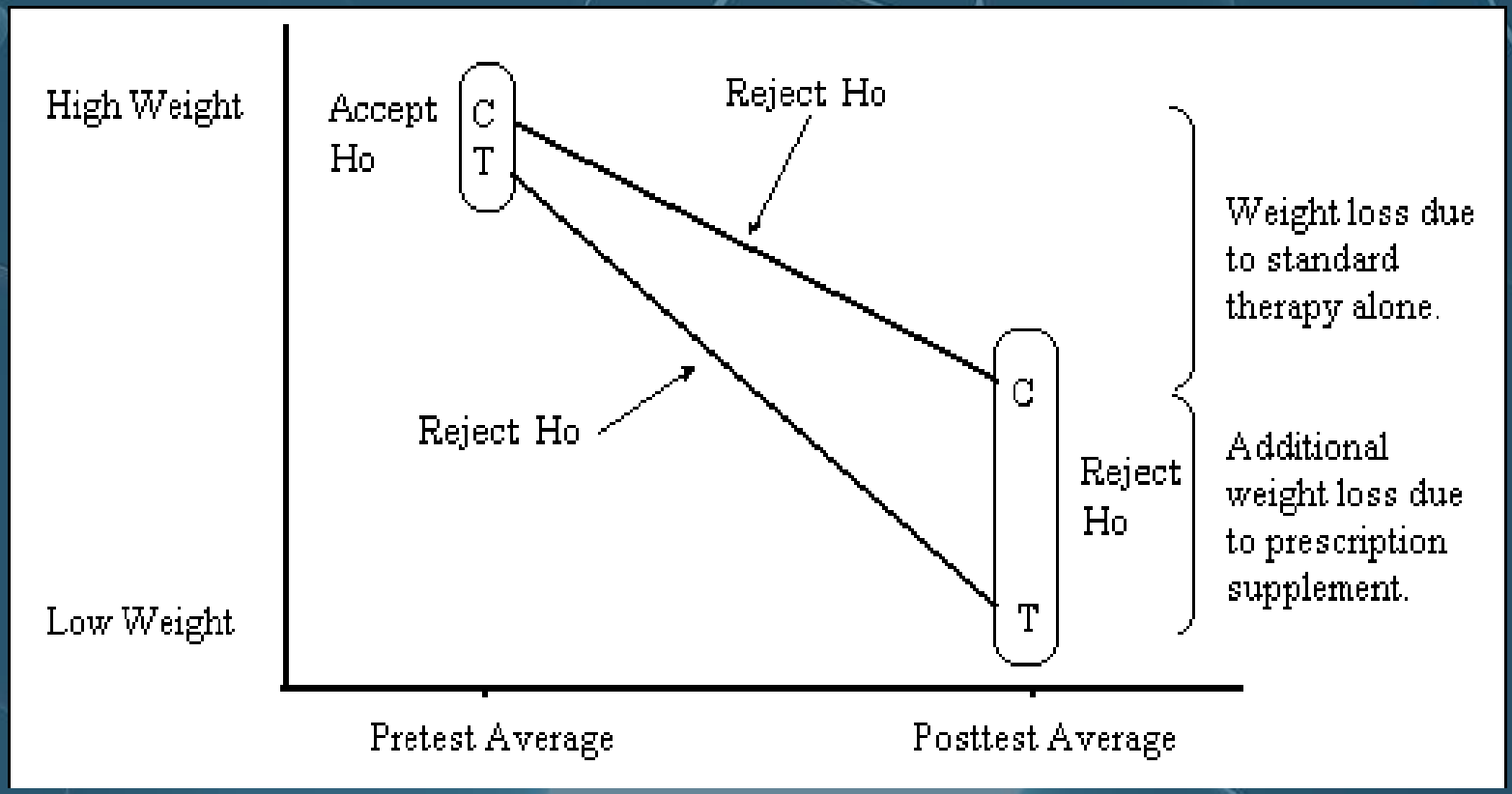
Theoretical testing for a NECG design.



Weight Loss Example

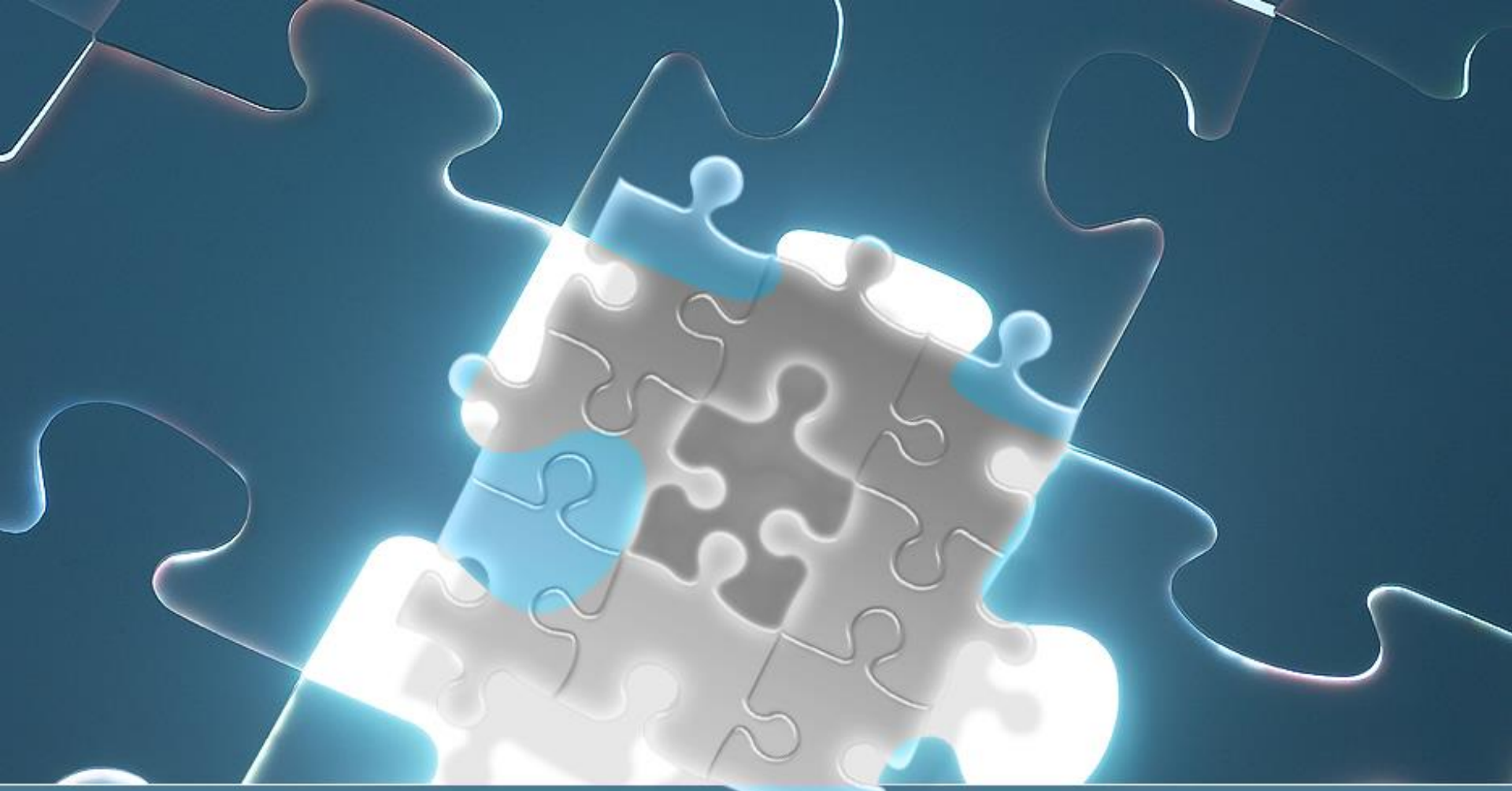


Effective Therapies



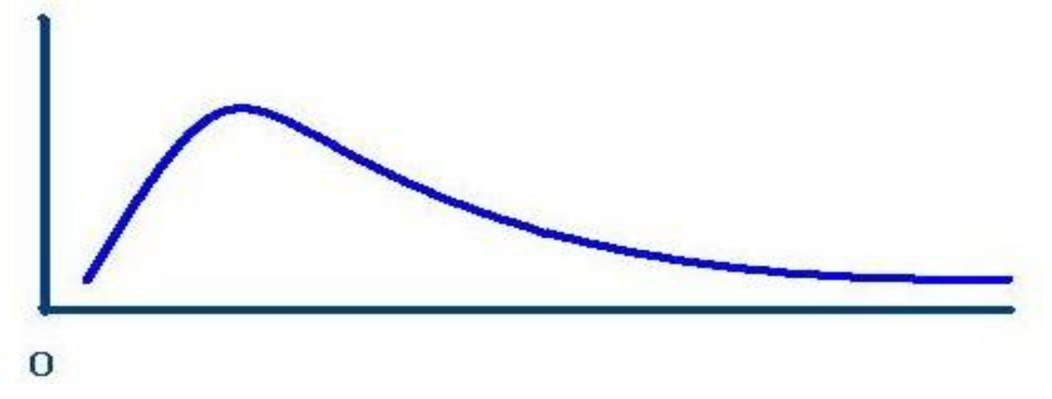
What if Your Data Do Not Meet Assumptions?

- Mathematically transform the data into a normal distribution (take the log or square root of values).
- Use a different class of tests called "non-parametric" tests. These tests are based on the ranking or ordering of data rather than their numerical values.



Analysis of Nominal Data: Chi-square Test

Chi square Distribution



- Not a symmetrical distribution
- Ranges from 0 to infinity –no negative values
- Skewed to the right
- Total area under the curve =1
- Family of distributions –different chi square distribution for each value of the degrees of freedom
- Degrees of freedom = (#rows-1) x (#columns -1)

Chi square Test

- Used when both exposure and outcome are nominal
- Can be 2 x 2 or larger tables
- Null hypothesis: No difference in proportions between groups

Chi square test

- Example:
  - In an air pollution study, a random sample 200 households were selected from each of 2 communities. A respondent in each household was asked whether or not anyone in the household was bothered by air pollution.
  - Data:

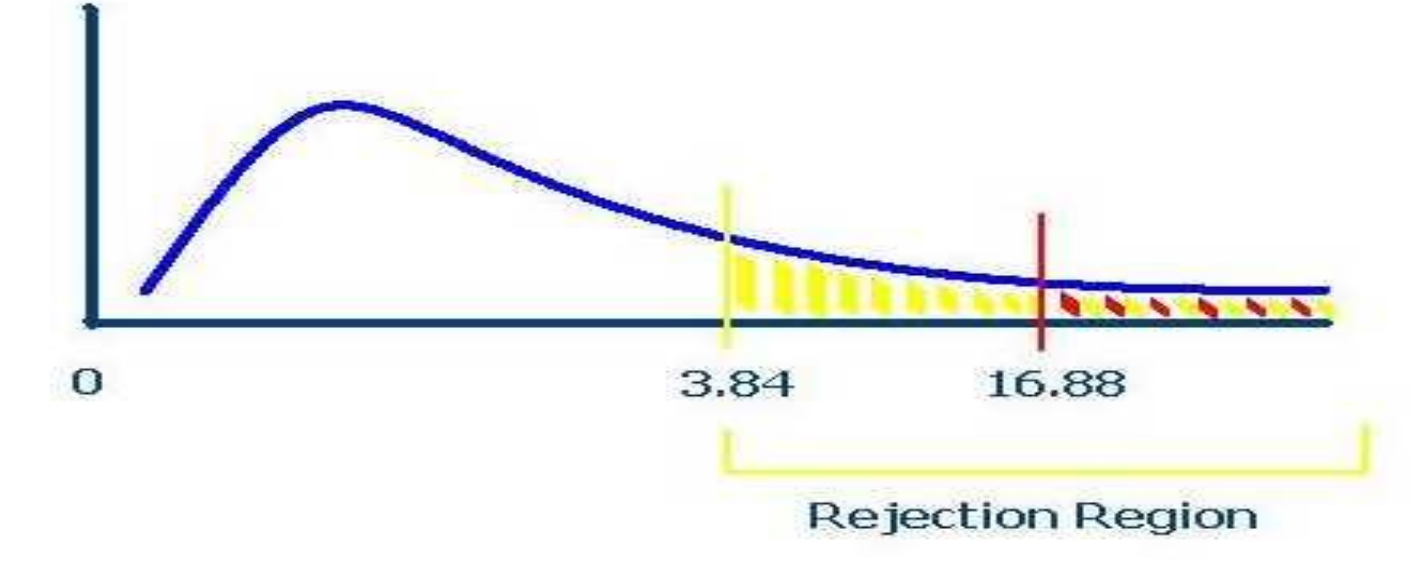
Bothered by air pollution	Community		Total
	A	B	
Yes	43	81	124
No	157	119	276
Total	200	200	400

Chi square Hypothesis

Bothered by air pollution	Community		Total
	A	B	
Yes	62 43	62 81	124
No	138 157	138 119	276
Total	200	200	400

- Expected = (row total x column total)/total population
  (124 x 200)/400 = 62 OR (124/400) x 200 = 62
- H_0: prop. bothered in Comm A = prop. bothered in Comm B.
- H_A: prop. bothered in Comm A ≠ prop. bothered in Comm B.
- α (alpha) =0.05

Chi square Results



- Critical value for Chi Square with 1 df =3.84
- 16.88 lies within the rejection region – reject the null hypothesis.
- $p<0.005$

Definitions

- Same statistical level (Nominal, Ordinal, Interval, Ratio)
- Secondary data (data collected for some other reason than what you need)
- Reliability and Validity (are you sure that coders/nurses/physicians see the same thing you do)
- Data that reside on one data system or, better yet, one data file are preferred.

Statistical Methods: What to Report

- Methods
  - Statistical procedure used
  - Significance level (Predetermined α level that will be considered statistically significant).
- Results
  - Sample size (n)
  - Means $\pm$ Standard Deviation
  - Observed significance level for each statistical test done
  - Confidence Interval for each mean

Have fun with stats!!!