

Statistical Concepts for Clinical Research



- PJ Heagerty
- Department of Biostatistics
- University of Washington

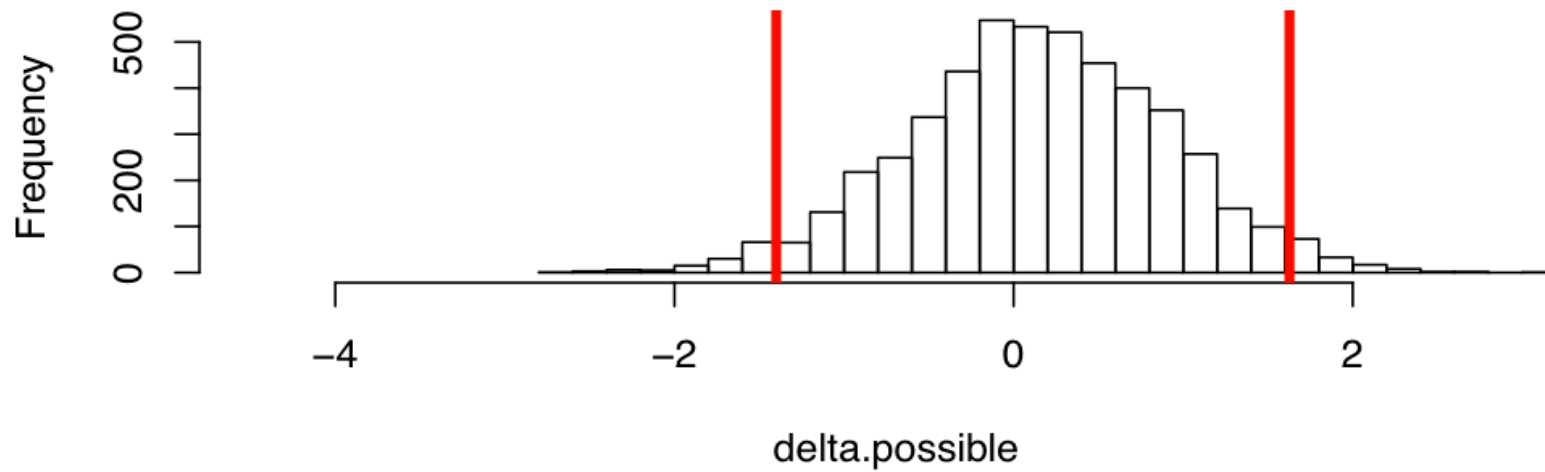
Session Three Outline

- Testing – the t-test
- Confidence intervals – the mathematical form
- Power
- Analysis of binary, event-time, and other outcomes
- What's next: Courses?
- What's next: Contacts?

Standard Hypothesis Testing

- Yesterday we say that in order to test a hypothesis we needed:
 - ▷ **EXPECTED** values for our comparison if the NULL WERE TRUE.
 - ▷ We then compared our **OBSERVED** difference to these “reference values” .
 - ▷ **p-value** shows placement of OBSERVED among EXPECTED values.

Expected under Null (10/group)



Standard Hypothesis Testing

- Same IDEA, just using a little more theory/math!

- **Common presentation of t-test:**

- ▷ Compute the OBSERVED difference and STANDARDIZE:

$$Z = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

- ▷ Compare this “distance” to **standard normal values** (e.g. expectations using a bell curve)
- Idea is exactly parallel to our permutation approach.

Standard Hypothesis Testing

- Needs measurement of **variation** captured in the data, s_1^2 (and s_2^2), or equivalently the square root of this sample variance – the sample **standard deviation**, s_1 (and s_2).

- | |
|---------|
| Define: |
|---------|

$$s_1^2 = \text{average} [(Y_i - \bar{Y})^2] \quad \text{in group 1}$$

- Summary of expected “distance.”
- **Our small surgical data:** $s_1 = 1.71$, $s_2 = 1.73$.

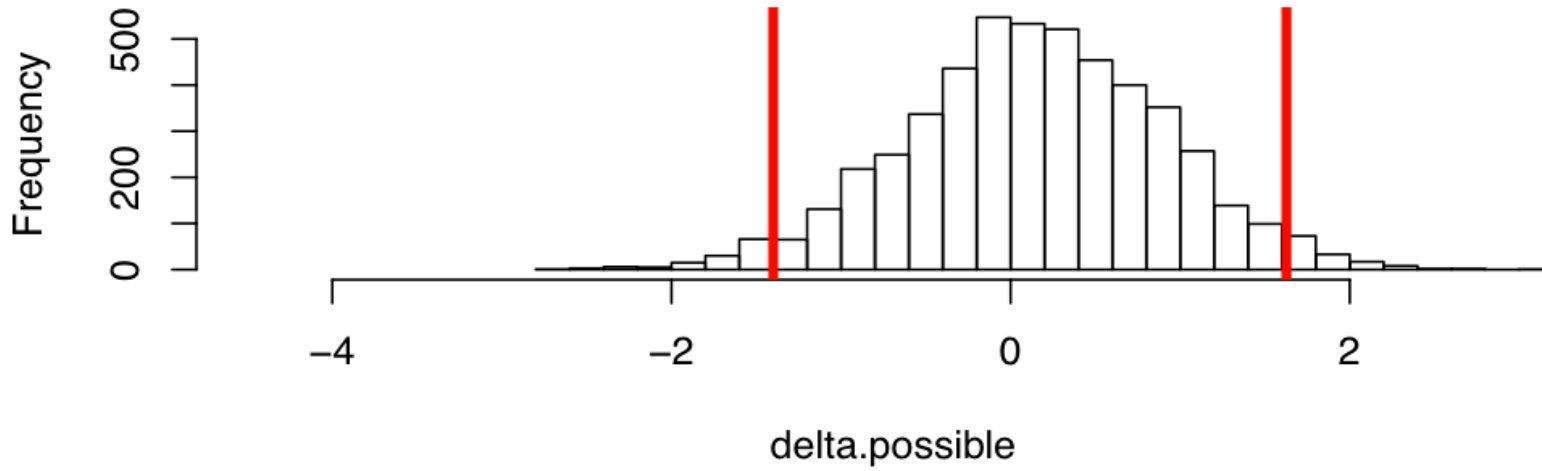
Principle: Plan your study to expect success!

- We have seen the EXPECTED VARIATION assuming the NULL determines the magnitude of a difference that is needed to call results significant.
- **Q:** If the NULL doesn't hold and we expect a difference then how likely are we to actually get a significant p-value?
- **A:** Depends on **sample size** and the **effect you think you might have** (e.g. your alternative hypothesis)

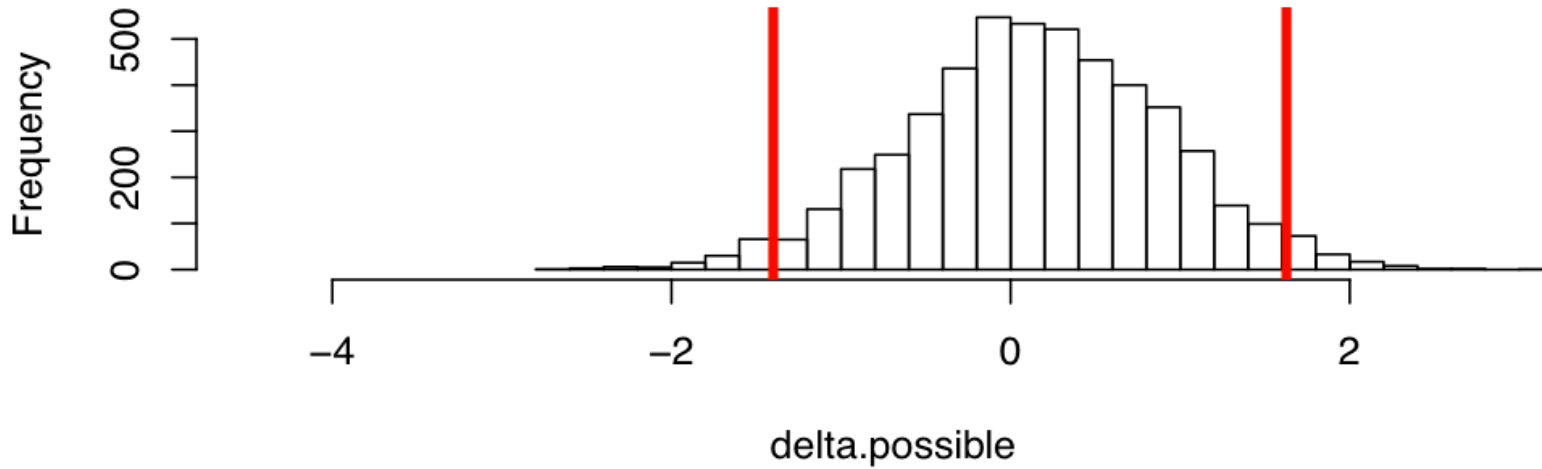
Power / Sample Size

- **A three step process:**
 - ▷ Calculate EXPECTED results if NULL IS TRUE (we did this already!)
 - ▷ Determine cut-off values that have only 5% outside these **critical values**.
 - ▷ Calculate ANTICIPATED results if THERE IS AN EFFECT (so you need to specify an effect size)
 - ▷ Count how often the ANTICIPATED results will reject the null – this is **power**!

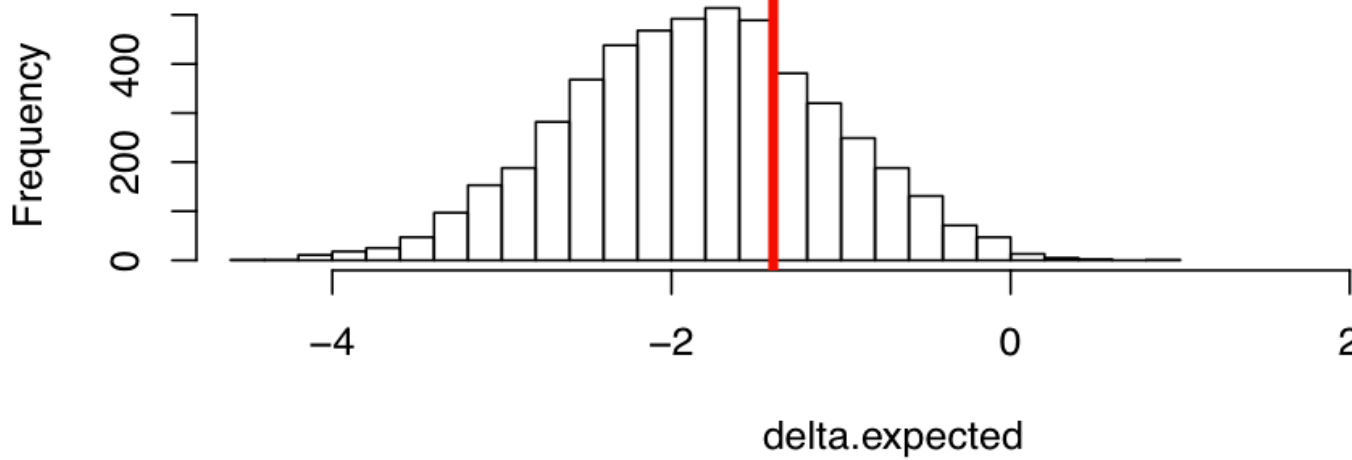
Expected under Null (10/group)



Expected under Null (10/group)



Expected if Benefit = 2pts (10/group)



Power

- **NULL EXPECTED**: Here we have assumed that we had a sample of $n=20$ untreated subjects available (e.g. pilot data) – made 1,000 copies of each person – and then we used these to simulate samples of $n=10$ subjects assigned to treatment and $n=10$ assigned to control – but all created equal!
- Calculate threshold to call significance: **critical value**.

Power

- **AN EFFECT EXPECTED:** Here we decided that we expected a surgical effect of **-2.0** points – our “alternative hypothesis” – and use one population like above to represent control subjects, and created another population where we subtracted **2.0** points to represent treated subjects.
- Simulate many studies with $n=10$ surgical and $n=10$ non-surgical subjects.
- Calculate how often the null is rejected: **power**.

Power

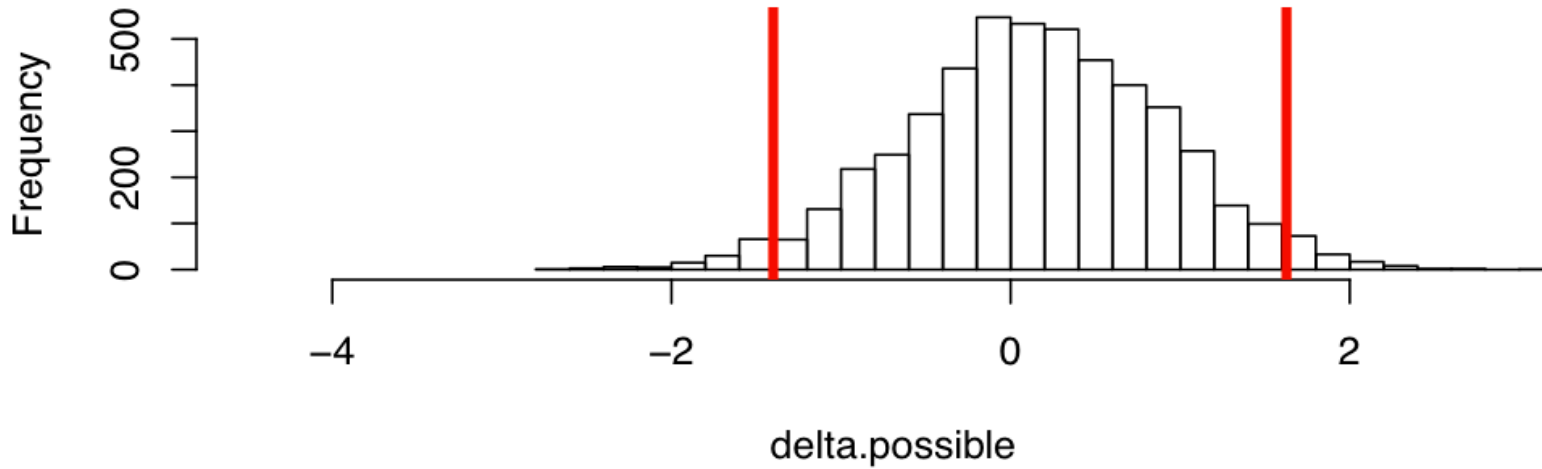
	Don't Reject	Reject Null
Null True	good	false positive
Alternative True	false negative	good power

- We usually set **false positive** at **5%**.
- We choose a sample size to get power \geq **80%**.

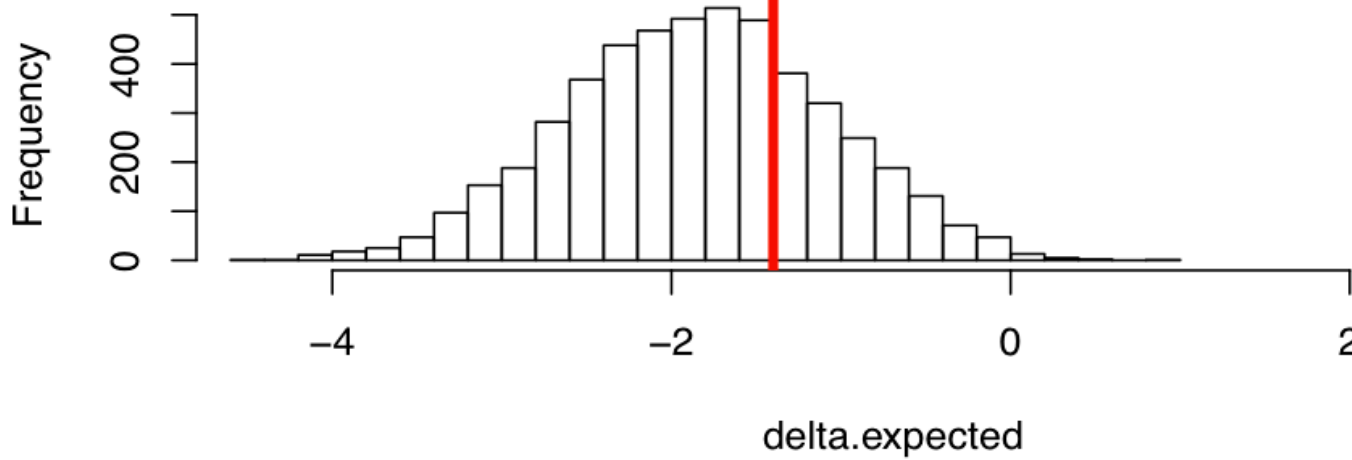
Our Planned Surgical Trial

- Using these ideas we find that $n=10$ / group would yield **power = 74%** if we anticipate a -2 point treatment effect.
- Let's consider a different sample size, $n=20$ / group, and repeat the exercise. Here's what will happen:
 - ▷ If NULL IS TRUE we will have less variability around 0.0
 - ▷ When ALTERNATIVE IS TRUE we will have less variability around -2.0

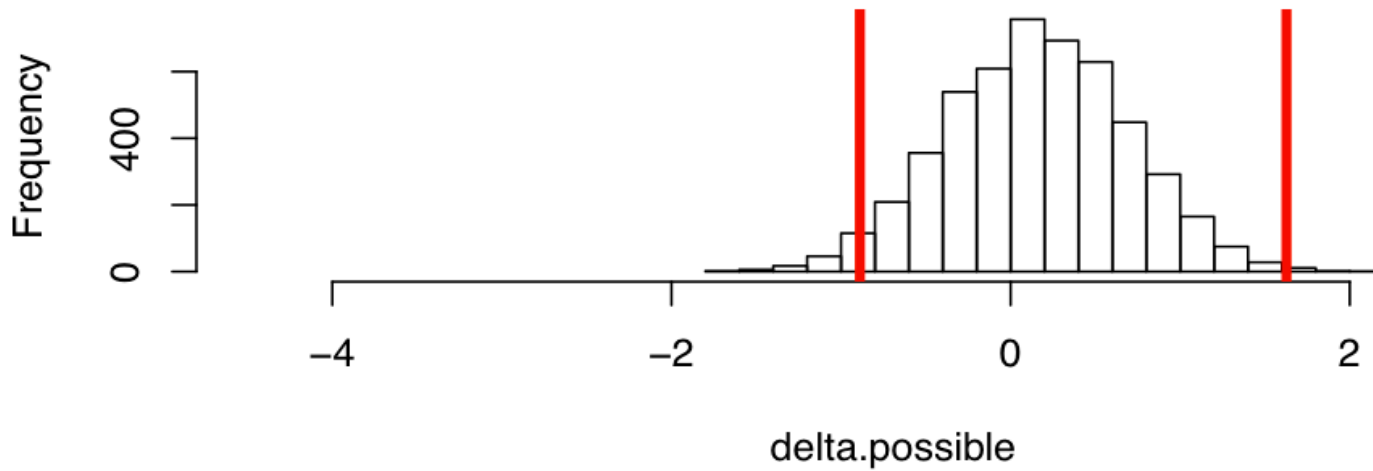
Expected under Null (10/group)



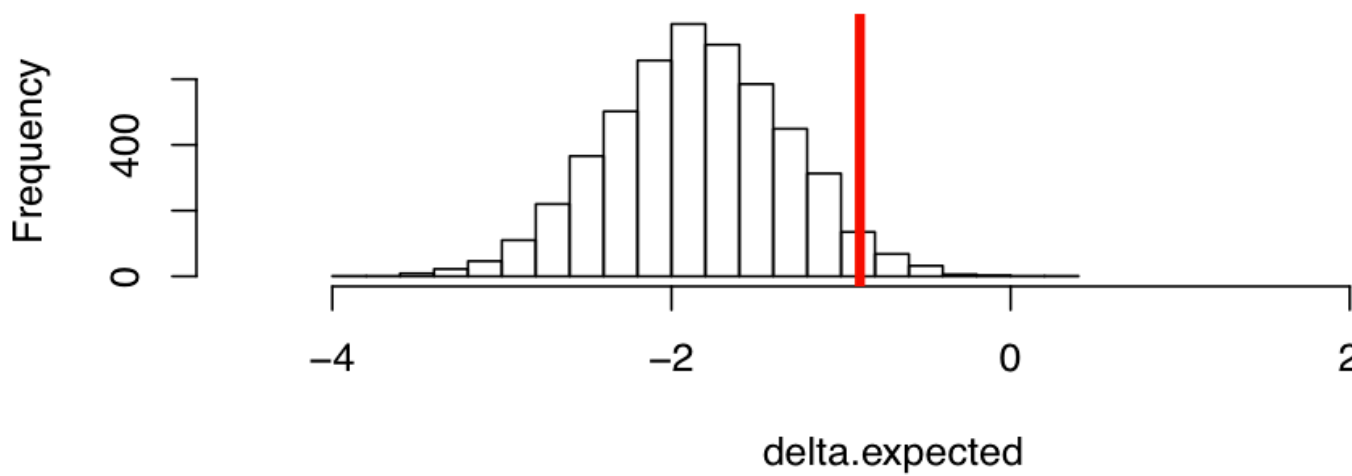
Expected if Benefit = 2pts (10/group)



Expected under Null (20/group)



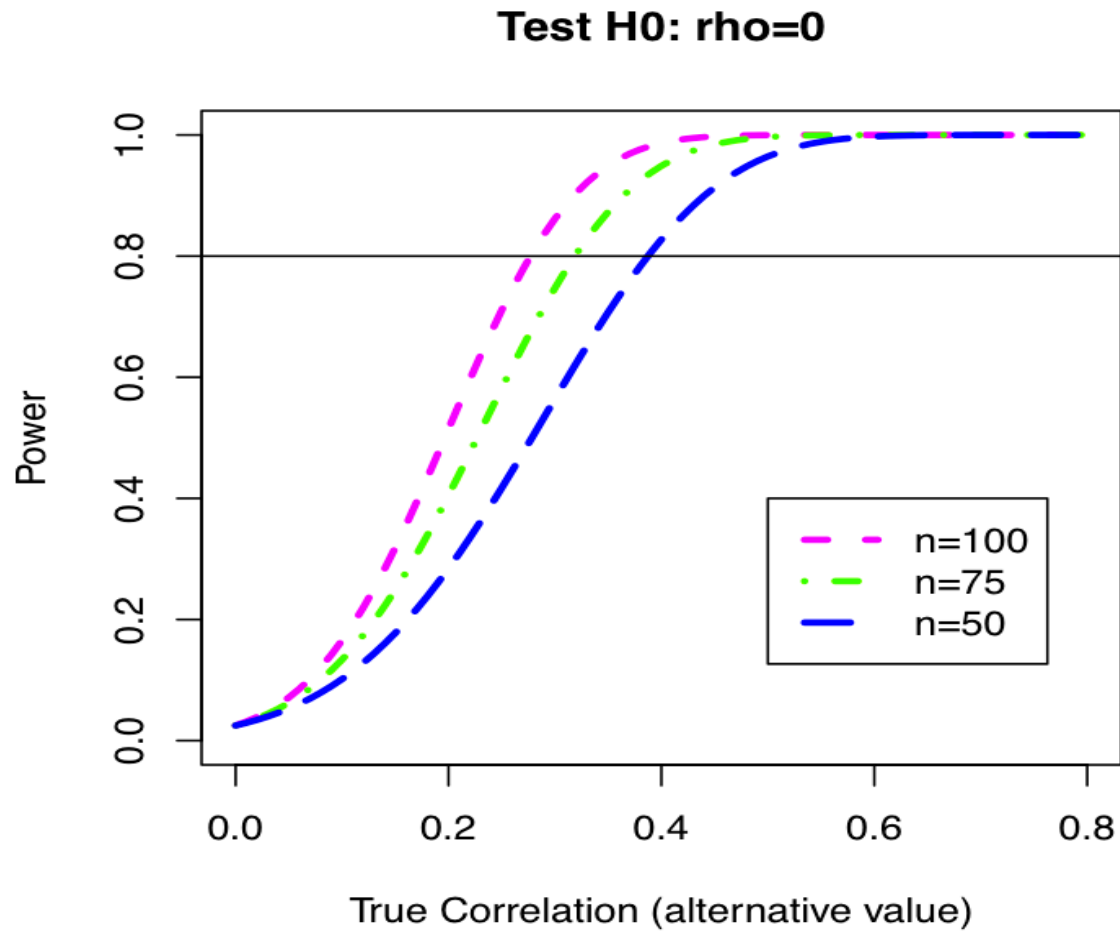
Expected if Benefit = 2pts (20/group)



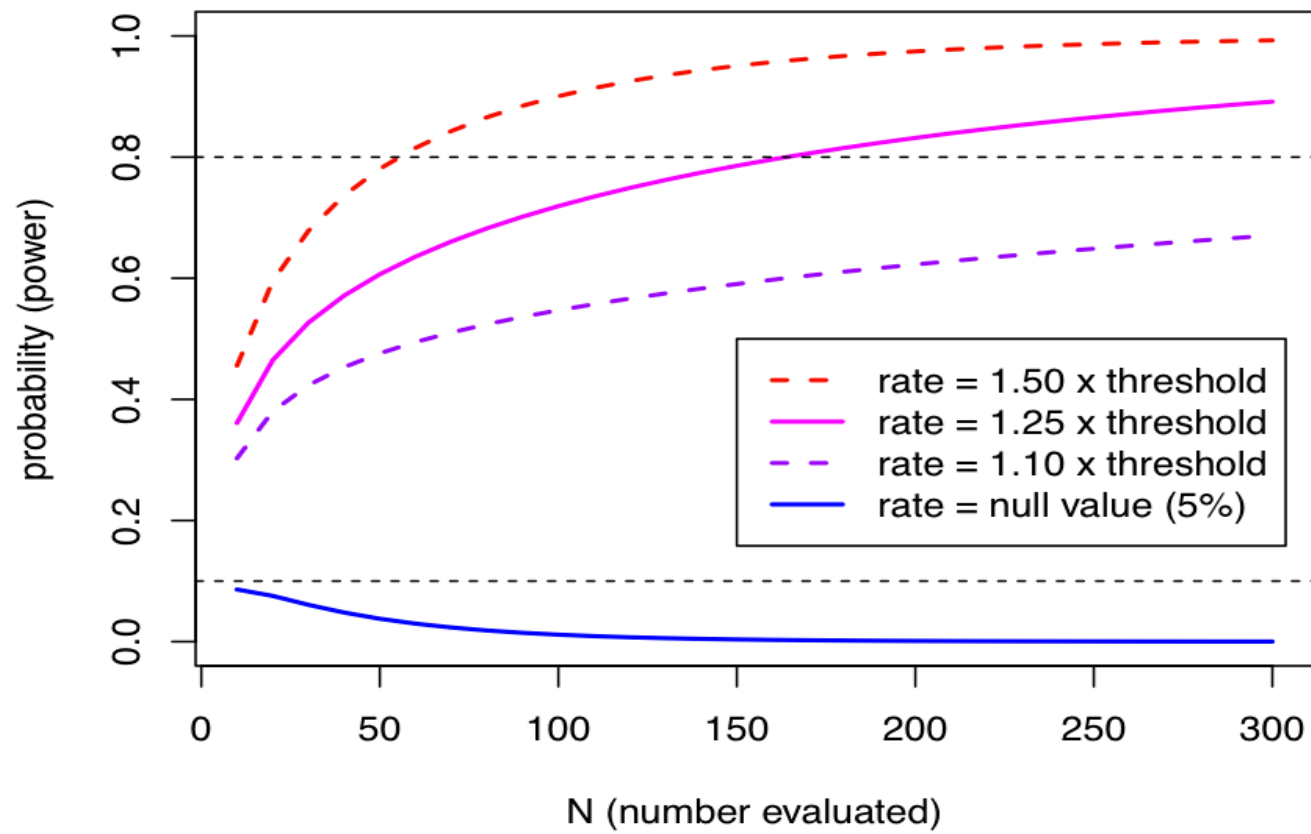
Our Planned Surgical Trial

- We now find that with $n=20$ /group we obtain **power = 96%**.
- Notice that the **critical value** moved closer to 0.0 since the sample size increased.
- We can make a couple of graphs to help inform sample size:
 - ▷ **vary sample size** but hold alternative fixed and display power.
 - ▷ **vary effect size** but hold sample size fixed and display power.

Example: Power for Grant Application



Example: Power for Grant Application



Power / Sample Size Tools

- You need some input information:
 - ▷ **standard deviation** for each group (we had data to suggest 1.7)
 - ▷ **anticipated group means** – the alternative hypothesis
- STATA has a simple tool `sampsi`.
- There are many online tools:
 - ▷ **n=12** stated using our parameters and:

<http://www.stat.ubc.ca/~rollin/stats/ssize/>

Power Calculation with STATA

The screenshot displays the Stata 12.0 software interface. The top menu bar includes icons for Open, Save, Print, Log, Viewer, Graph, Do-file Editor, Data Editor, Data Browser, More, and Break. The main window is divided into three panes: Review, Results, and Variables.

Review Pane: Shows the command entered: `. sampsi 5.5 3.5, sd1(1.71) sd2(1.71) n1(10) n2(10)`

Results Pane: Displays the output of the command:

```
. sampsi 5.5 3.5, sd1(1.71) sd2(1.71) n1(10) n2(10)

Estimated power for two-sample comparison of means

Test Ho: m1 = m2, where m1 is the mean in population 1
              and m2 is the mean in population 2

Assumptions:

              alpha = 0.0500 (two-sided)
              m1 = 5.5
              m2 = 3.5
              sd1 = 1.71
              sd2 = 1.71
sample size n1 = 10
              n2 = 10
              n2/n1 = 1.00

Estimated power:

              power = 0.7439
```

Power Calculation with STATA

The screenshot displays the Stata 12.0 software interface. The top menu bar includes icons for Open, Save, Print, Log, Viewer, Graph, Do-file Editor, Data Editor, Data Browser, More, and Break. The main window is divided into three panes: Review, Results, and Variables.

Review Pane: Shows the command entered: `1 sampsi 5.5 3.5, sd1(...)`

Results Pane: Displays the output of the `sampsi` command. The command used is `. sampsi 5.5 3.5, sd1(1.71) sd2(1.71) n1(20) n2(20)`. The output indicates an estimated power for a two-sample comparison of means. The test hypothesis is $H_0: m_1 = m_2$, where m_1 is the mean in population 1 and m_2 is the mean in population 2. The assumptions are listed as follows:

alpha =	0.0500	(two-sided)
m1 =	5.5	
m2 =	3.5	
sd1 =	1.71	
sd2 =	1.71	
sample size n1 =	20	
n2 =	20	
n2/n1 =	1.00	

The estimated power is **0.9589**.

Variables Pane: Shows a table with columns for Name and Label, which is currently empty.

Sample Size Calculation with STATA

Stata 12.0

Open Save Print Log Viewer Graph Do-file Editor Data Editor Data Browser More Break

Review Results

Command _rc

```
. sampsi 5.5 3.5, sd1(1.71) sd2(1.71) p(0.80)
```

Estimated sample size for two-sample comparison of means

Test Ho: $m_1 = m_2$, where m_1 is the mean in population 1
and m_2 is the mean in population 2

Assumptions:

alpha = 0.0500 (two-sided)
power = 0.8000
m1 = 5.5
m2 = 3.5
sd1 = 1.71
sd2 = 1.71
n2/n1 = 1.00

Estimated required sample sizes:

n1 = 12
n2 = 12

Variables

Name	Label

Summary of Power / Sample Size

- Planning to be confident that your study will yield a definitive result if the effect of treatment is what you anticipate.
- Requires input information / pilot data.
- Issues around selection of alternative hypothesis – what is smallest clinically meaningful effect?
- Tools exist for quantitative outcomes, binary outcomes, and event-time end-points.

(*) Confidence Intervals

- **common presentation:**

$$(\bar{Y}_1 - \bar{Y}_2) - 1.96 \times SE, \quad (\bar{Y}_1 - \bar{Y}_2) + 1.96 \times SE$$

- Where **SE** is the “standard error” defined as

$$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

- The interval above is a **95% confidence interval** for the true difference of means.

Courses at UW

BIOST 511 Medical Biometry I (4)

Presentation of the principles and methods of data description and elementary parametric and nonparametric statistical analysis. Examples are drawn from the biomedical literature, and real data sets are analyzed by the students after a brief introduction to the use of standard statistical computer packages. Statistical techniques covered include description of samples, comparison of two sample means and proportions, simple linear regression and correlation. Offered: A.

BIOST 512 Medical Biometry II (4)

Multiple regression, analysis of covariance, and an introduction to one-way and two-way analyses of variance: including assumptions, transformations, outlier detection, dummy variables, and variable selection procedures. Examples drawn from the biomedical literature with computer assignments using standard statistical computer packages. Prerequisite: either BIOST 511 or BIOST 517, or equivalent. Offered: W.

BIOST 513 Medical Biometry III (4)

Analysis of categorical data including two sample methods, sets of 2 x 2 tables, R x C tables, and logistic regression. Classification and discrimination techniques. Survival analysis including product limit estimates and the Cox proportional hazards model. Prerequisite: BIOST 512 or permission of instructor. Offered: Sp.

Instructor Course Description: [Norbert David Yanez Iii](#)

Courses at UW

BIOST 536 Categorical Data Analysis in Epidemiology (4)

Summary of univariate categorical data analysis; introduction to multivariate analysis of categorical epidemiologic and health sciences data using multiplicative models. Experience at interpretation; familiarity with available software programs gained by analysis of bona fide data and critiques of published analyses appearing in literature. Prerequisite: BIOST 515; EPI 513 and either BIOST 513 or BIOST 518; or permission of instructor. Offered: jointly with EPI 536; A.

Instructor Course Description: [*Norman Breslow*](#)

BIOST 537 Survival Data Analysis in Epidemiology (4)

Introduction to multivariate analysis of survival data using multiplicative models. Application to epidemiologic and health sciences studies. Familiarity with interpretation and available software computer programs gained by analysis of bona fide sets of data and critiques of published analyses appearing in the literature. Prerequisite: BIOST 536 or permission of instructor. Offered: jointly with EPI 537; W.

Instructor Course Description: [*Norman Breslow*](#)

BIOST 540 Correlated Data Regression (3)

Introduction to regression modeling of longitudinal and clustered data from epidemiology and health sciences. Interpretation and familiarity with software gained by analysis of data and critiques of published analyses. Prerequisite: Either BIOST 513, BIOST 515, BIOST 518, BIOST 536, or permission of instructor. Offered: Sp.

Instructor Course Description: [*Brian Leroux Norman Breslow Norbert David Yanez Iii*](#)

Courses at UW

Summer Institute in Biostatistics: Calendar and Daily Schedule

Calendar			Click Here to go to Daily Schedule		
Monday 8:30 am- 5:00 pm	Tuesday 8:30 am- 6:00 pm	Wednesday 8:30 am- noon	Wednesday 1:30 pm- 5:00 pm	Thursday 8:30 am- 6:00 pm	Friday 8:30 am- 5:00 pm
August 6	August 7	August 8	August 8	August 9	August 10
Mod 1: Introduction to Clinical Trials Mod 2: Longitudinal Data Mod 3: Introduction to Survival Analysis			Mod 4: Advanced Topics in Design of Clinical Trials Mod 5: Sequential Monitoring of Clinical Trials Mod 6: Design of Clinical Trials in Oncology		

Ask for help

ITHS Services

At ITHS, we take advantage of our multi-institutional staff to offer researchers a wide range of services. Below are general areas of ITHS expertise along with examples of the services they offer.

Clinical Research Studies and Support Services

- [Clinical Research facilities \(Pediatric and Adult CRCs\)](#)
- [Bionutrition Research](#) - Research diet design and implementation, diet intake analysis, diet assessment and instruction
- [Body Composition/Bioenergetics and Exercise Research](#) - Body composition, energy expenditure and activity measurement and analysis
- [Dental clinical research facility](#)
- [Research Coordinator Core](#) - Trained research support staff
- [Data safety and monitoring plans](#)
- [Research bioethics consultation](#)
- [Regulatory auditing services](#)

Study and Data Management

- [Statistical methods](#): Study design, sample size determination, randomization, and blinding
- [Electronic medical record \(EMR\) data extraction](#)
- [Biospecimen acquisition and management tools](#)
- [Study data management and electronic data capture including REDCap](#)

Core Facility and Technology Access

- [Directory of Regional Research Facilities and Services](#)
- [Gene and Cell Therapy Lab](#) - Manufacturing of clinical cell products
- [Center for Clinical Genomics](#)
- [Metabolomics User Group](#)

Outline for ITHS Bootcamp

- Why you might want to talk to a statistician.
- How a statistician approaches research aims
- Key considerations in study design
- What you need to know to “bridge the gap”
 - ▷ What data should arise if there is **no association**?
 - ▷ What data should arise if there is **an association**?