Introduction to Power Analysis and Sample Size Calculation

Liang Shan, PhD
Research Biostatistician
Assistant Professor, Biomedical Affairs
Email: lshan@vcom.vt.edu
Consulting and Collaboration meetings

Visit our website to request a personalized statistical assistance with

- study design
- grant proposals (e.g. sample size calculation)
- data visualization and analysis

http://www.vcom.edu/statistical-consulting-and-collaboration
Outline

• Define power analysis and distinguish it from sample size calculation

• Identify information needed for power analysis in analytic studies and experimental studies

• Get to know some resources for simple power analysis
Population and Sampling

• Population: a collection of all individuals about whom we wish to make a statement.

• Research question determines the composition of population.

• But how many samples does one need to answer a given question?
# Example 1: Prevalence of Type II diabetes

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Research Question</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>To evaluate the prevalence of Type II diabetes among all adults living in the <strong>United States</strong></td>
<td>All adults living in the <strong>United States</strong></td>
</tr>
<tr>
<td>2</td>
<td>To evaluate the prevalence of Type II diabetes among all adults living in <strong>Virginia</strong></td>
<td>All adults living in <strong>Virginia</strong></td>
</tr>
<tr>
<td>3</td>
<td>To evaluate the prevalence of Type II diabetes among all adults living in <strong>Blacksburg</strong></td>
<td>All adults living in <strong>Blacksburg</strong></td>
</tr>
</tbody>
</table>
Population size is the number of subjects in the population, usually denoted by \( N \).

Sample Size is the number of subjects in the sample, usually denoted by \( n \).

How many samples, \( n \), are needed to answer a question that correctly represents the whole population, \( N \)?
Key Question

Many research projects focus on changes relative to a control.

How do you know how big your sample sizes need to be to adequately represent the test group and control group – and provide clear information on whether the two groups are the same or different?
Sample Size Calculation vs. Power Analysis

- The so called “power analysis” (“prior power analysis”), i.e., sample size calculation using power analysis.
- Prior power analysis is used when we have research hypothesis to test.
- Power analysis tells us how many samples are needed.
What is Power Analysis?

“Power analysis is an important aspect of experimental design. It allows us to determine the sample size required to detect an effect of a given size with a given degree of confidence. Conversely, it allows us to determine the probability of detecting an effect of a given size with a given level of confidence, under sample size constraints.

From: the website “Quick R,” https://www.statmethods.net/stats/power.html
Why is Sample Size Important?

Too large?
• difficult to get data on every member
• more costly than necessary
• may put more subjects “at risk”

Too Small?
• fail to answer research question, or
• fail to detect important effects even for a well designed study
What Type of Project?  
When Do You Ask These Questions?

Descriptive/Exploratory Observational Studies

- No need for sample size calculation or estimate sample size by another approach (e.g., CI approach)
- Data mining process: explore the lay of the land and lead to generation of research hypotheses
- Example: projects originates from QA/QI projects (e.g. educational & primary care) _ information from the whole population are available

Analytical Observational Studies or Experimental Studies (Clinical Trial)

- Power analysis is mostly needed due to predetermined research hypotheses
- 1, Analytical Observational Study: hypotheses are about association
- 2, Experimental Study (Clinical Trial): hypotheses are about cause-and effect relationship
How do these questions affect project design?

Step 1: Research Aims/Questions
- Population of interest is determined

Step 2: Study Design
- Possible samples to consider

Step 3: Statistical Consideration
- Sample size Calculation
- Data Analytical Approach
Narrow Definition: Power Analysis

Given power, significance level, and standardized effect size, we can determine sample size.
Broad Definition of Power Analysis

Given any three, we can determine the fourth.

- **Prior power analysis (ideal):**
  determine sample size
- **Post-hoc power analyses:**
  determine power
- **Criterion power analysis:**
  determine $\alpha$
- **Sensitivity power analysis:**
  determine standardized effect size
Null and Alternative Hypothesis

• Set up study and proposed hypothesis:
  – Clarify research question
  – Identify the parameter of interest (e.g. population mean “μ”, population proportion “p”)

• Null and research hypotheses
  – Null hypothesis (H₀) reflects no difference, no association, or no effect
  – Alternative hypothesis (H₁ or Hₐ) reflects the investigator’s belief

• Example:
  – H₀: There is no difference in mean blood-pressure “μ” between smokers and non-smokers, i.e., \( \mu_{\text{smoker}} = \mu_{\text{non-smoker}} \)
  – H₁: There is a difference in mean blood-pressure between smokers and non-smokers, i.e., \( \mu_{\text{smoker}} \neq \mu_{\text{non-smoker}} \)
## Type I and Type II error

<table>
<thead>
<tr>
<th>Decision</th>
<th>$H_0$</th>
<th>$H_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail to reject $H_0$</td>
<td><strong>CORRECT DECISION</strong></td>
<td>Type II error</td>
</tr>
<tr>
<td>Reject $H_0$</td>
<td>Type I error</td>
<td><strong>CORRECT DECISION</strong></td>
</tr>
</tbody>
</table>
# Medical Test

<table>
<thead>
<tr>
<th>Test Result</th>
<th>The truth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Disease</td>
</tr>
<tr>
<td></td>
<td>Disease</td>
</tr>
<tr>
<td>Negative</td>
<td>False negative</td>
</tr>
<tr>
<td>Positive</td>
<td>False positive</td>
</tr>
</tbody>
</table>

*CORRECT DECISION*
## Power & Significance Level

### The truth

<table>
<thead>
<tr>
<th>Decision</th>
<th>$H_0$ (no difference/effect)</th>
<th>$H_1$ (some difference/effect)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fail to reject $H_0$</td>
<td><img src="https://example.com/correct" alt="Correct Decision" /></td>
<td>Type II error</td>
</tr>
<tr>
<td>Reject $H_0$</td>
<td>Type I error</td>
<td><img src="https://example.com/correct" alt="Correct Decision" /></td>
</tr>
</tbody>
</table>

- **Significance level ($\alpha$)** = Probability (Type I error) = Probability (Reject $H_0$ | $H_0$ is true) = Probability (False positive)
- **$\alpha=0.05$** is chosen most commonly

- **$\beta$** = Probability (Type II error) = Probability (Fail to reject $H_0$ | $H_1$ is true) = Probability (False negative)
- **Power** = $1-\beta$ = Probability (Reject $H_0$ | $H_1$ is true) = Probability (True positive)
- **Power=0.8 to 0.9** is most commonly used
Effect size (ES) is the difference in the parameter of interest that represents a *clinically/practically meaningful difference.*

Related to the difference in two measurement mean values and their corresponding standard deviations

Often referred to as “small,” “medium,” or “large” effects

- **Small Effect:** the difference between two means is 0-0.2 standard deviations
  - The difference is real, but it can only be discerned or measured through careful and precise study.

- **Large Effect:** the difference between two means is greater than 0.5 standard deviations
  - The difference is easily observable with even the most rudimentary study methods

- **Medium Effect:** the difference between two means is between 0.2 to 0.5 standard deviations
  - Small but easily discernable, easily reproducible but with some care, little doubt about outcome
Effect Size and Variability

• Clearly variability in a measurement will affect the Effect Size.

• **Standardized Effect Size = Effect Size / Variability**

• The *effect size* and *variability* can be estimated from the following:
  - Previous studies in literature
  - Consultation with domain experts
  - Small pilot study

• Sometimes, an investigator cannot obtain any meaningful information about the standard deviation of a variable. In that situation, it’s worthwhile to use the unit less **standardized effect size**.
  - The larger the standardized effect size, the smaller the required sample size.
  - For most studies in clinical research, the standardized effect size will be >0.1.
Power Analysis Determinant

• **Effect size** \(\uparrow\), **sample size** \(\downarrow\)  Bigger effects are easier to measure with small numbers of samples.

• \(\alpha\) \(\uparrow\), **sample size** \(\downarrow\)  More samples are needed to have a smaller confidence interval.

• **Power** \(\uparrow\), **sample size** \(\uparrow\)  To ensure the alternative hypothesis is correct, more samples are needed.

• **Variability** \(\uparrow\), **sample size** \(\uparrow\)  Variability increases standard deviation, so more samples are needed.
Steps for Prior Power Analysis

1. State the **null hypothesis** and **alternative hypothesis** (either one-sided or two-sided).

2. Select the **appropriate statistical test** based on the type of predictor variable and outcome variable in the hypotheses.

3. Choose a **reasonable effect size (and variability)** or **standardized effect size**.

4. Set **significance level** (α) and **power** (1-β).

5. Use G*Power, an online calculator, or a statistical package to estimate the sample size.
Practical Consideration

Account for dropouts and subjects with missing data.

- For instance, if it is estimated that 20% of the sample will be lost to follow up, then the sample size should be increased by a factor of $1/(1-0.2)=1.25$

From Dr. Reilly’s R15 Grant (Mock Review)

**Statistical analysis and power of the study:** We will consider a >50% change in any parameter to be significant. To achieve a power of 0.8 with an α value of 0.05 we will need to have 15 animals in each treatment group. Three additional mice will be included in each group to account for deaths due to disease.
Resources for Prior Power Analysis

• G*Power, a free and comprehensive software for power analysis.
  ❖ Manual
    [link]
  ❖ Short Tutorial
    [link]

• [link], good for sample size calculations in simple situations.

• Other statistical packages, for instance, JMP, R, SAS, and STATA.
References


Take Home Message

• Understand the connection and difference between power analysis and sample size calculation
• Know what information is needed for prior power analysis and where to find them
• Get to know some resources for simple power analysis