Family Studies Center Methods Workshop
Statistical Power Analysis

Adam Davey
Temple University

November 14, 2014
Overview

- Understand the role of statistical power analysis in family studies research
- Introduce concept of statistical power
- Develop intuitions about factors affecting statistical power
- Learn applications of power analysis when sample size is fixed
Motivating Problem

- Research is *difficult, time-consuming, and expensive* to conduct
  
- *Before* we conduct a study, we want to be assured that we have a reasonable change of finding an effect if, in fact, one exists

- We must recruit *sufficient numbers* of subjects into our study

- We must *also* consider efforts (and potential for risk) of study participants

- Nearly all studies entail at least some *risk* for participants (even after data are collected!)

- We must not recruit *too many* research subjects into our study
Motivating Problem

- When number of potential subjects is \textit{limited}, need to \textit{identify design} that gives us the best chance of answering our question

- When number of subjects is \textit{fixed} in advance, need to know \textit{how big an effect} we can detect in our data with desired probability
### Ways to be Wrong in Hypothesis Testing

<table>
<thead>
<tr>
<th>Decision</th>
<th>True State of Affairs</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Accept $H_0$</td>
<td>$H_0$ True</td>
<td>$H_0$ False</td>
</tr>
<tr>
<td></td>
<td>Correct $(1 - \alpha)$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>Reject $H_0$</td>
<td>$\alpha$</td>
<td>Correct $(1 - \beta)$</td>
</tr>
</tbody>
</table>
Ways to be Wrong in Hypothesis Testing

<table>
<thead>
<tr>
<th>Decision</th>
<th>True State of Affairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Effect</td>
<td>No Effect</td>
</tr>
<tr>
<td>Effect</td>
<td>Effect</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Central and Non-Central Distributions

- Noncentral Distributions
- Statistical Power Analysis

Adam Davey (Temple University)  Statistical Power Analysis  November 14, 2014
Central and Non-Central Distributions

- **Central** distributions apply when the null hypothesis ($H_0$) is true
- They are *standardized*
- **Non-Central** distributions apply when $H_0$ is false
- They are *not standardized*
- Non-Centrality Parameter ($NCP, \lambda$) reflects *degree* to which ($H_0$) is false
- Non-centrality parameter can affect both *location* and *shape* of distribution.
Central and Non-Central Distributions

- Central $\chi^2$ distribution with $df$ degrees of freedom can be generated by squaring and summing $df$ different random normal variates with means of 0 and variances of 1.

- Non-central $\chi^2$ distribution with $df$ degrees of freedom and $NCP = \lambda$ can be generated by squaring and summing $df$ different random normal variates with means of

$$\mu = \sqrt{\frac{\lambda}{df}}$$
• Four variables are important for power analysis
  • $\alpha$
  • Power, $(1 - \beta)$
  • $N$
  • Effect Size, $(ES, \lambda)$

Knowing any 3, solve for fourth

Two other factors include choice of $H_0$ and $Pr(H_0$ is false)
Conventions
- $\alpha = .05$
- $\text{Power} \geq .80$
- $N$ (Some applications, may define minimum acceptable standards or heuristics for overall sample size, distinct from power conventions)

Effect Size
- Power analyses are invaluable $a$ priori, not so useful $a$ posteriori

(http://www.stat.uiowa.edu/files/stat/techrep/tr378.pdf)
Visualizing Statistical Power

alpha = 0.05
Visualizing Statistical Power
Visualizing Statistical Power

\[ alpha = 0.05 \]
Visualizing Statistical Power

\[
\text{Area under } H_0 = 0.05 \\
\text{Area under } H_A = 0.08
\]
Effects of Increasing Alpha

\[ \alpha = 0.001 \]

Area under H0 = 0.001

Area under HA = 0
Effects of Increasing Alpha

alpha = 0.01

Area under H0 = 0.01
Area under HA = 0.02

Density

0.4
0.3
0.2
0.1
0.0

-4 -2 0 2 4

ES
N
Variance
Summary
Resources
Effects of Increasing Alpha

\[\text{Area under } H_0 = 0.025\]

\[\text{Area under } H_A = 0.04\]
Effects of Increasing Alpha
Effects of Increasing Alpha

![Graph showing the effects of increasing alpha](image)

- Area under $H_0 = 0.1$
- Area under $H_A = 0.14$
Effects of Increasing Alpha

alpha = 0.2

Area under H0 = 0.2

Area under HA = 0.25
Effects of Increasing Effect Size

ES = 0.2

Area under H0 = 0.05

Area under HA = 0.05
Effects of Increasing Effect Size

ES = 0.4

Area under H0 = 0.05

Area under HA = 0.07
Effects of Increasing Effect Size

ES = 0.6

Area under H0 = 0.05
Area under HA = 0.09
Effects of Increasing Effect Size

![Graph showing the effects of increasing effect size](image)

- ES = 0.8

Area under H0 = 0.05

Area under HA = 0.13
Effects of Increasing Effect Size

ES = 1

Area under H0 = 0.05
Area under HA = 0.17

Density

0.0 0.1 0.2 0.3 0.4

-4  -2   0   2   4

ES
N
Variance
Summary
Resources
Discerning Patterns: Large $N$

- Clockwise: (None, Small, Large, Moderate, $N = 1000$)
Discerning Patterns: Small $N$

- Clockwise: (None, Small, Large, Moderate, $N = 10$)
Effects of Increasing $N$ or Decreasing Variance/SE

Area under $H_0 = 0.05$
Area under $H_A = 0.17$
Effects of Increasing $N$ or Decreasing Variance/SE

$SD = 0.71$

Area under $H_0 = 0.05$

Area under $HA = 0.29$
Effects of Increasing $N$ or Decreasing Variance/SE

$SD = 0.58$

Area under $H0 = 0.05$

Area under $HA = 0.41$
Effects of Increasing $N$ or Decreasing Variance/SE

SD = 0.5

Area under $H_0 = 0.05$
Area under $HA = 0.52$
Effects of Increasing $N$ or Decreasing Variance/SE

$SD = 0.45$

Area under $H0 = 0.05$

Area under $HA = 0.61$
Effects of Increasing $N$ or Decreasing Variance/SE

$$SD = 0.41$$
Effects of Increasing $N$ or Decreasing Variance/SE

![Graph showing the effects of increasing N or decreasing variance/SE. The graph illustrates the area under the curve for H0 = 0.05 and HA = 0.75 with SD = 0.38.](image-url)
Effects of Increasing $N$ or Decreasing Variance/SE
Effects of Increasing $N$ or Decreasing Variance/SE

![Graph showing the effect of increasing sample size or decreasing standard deviation on statistical power. The graph illustrates the decrease in variance and increase in sample size leading to a higher peak and wider distribution of the test statistic. The areas under the curves for the null hypothesis ($H_0$) and alternative hypothesis ($H_A$) are also shown.]
Effects of Increasing $N$ or Decreasing Variance/SE
Summing Up

- Power of hypothesis test with significance level $\alpha$ is probability we reject null hypothesis when the alternative is true
- Power is probability that data gathered will be sufficient to reject null hypothesis when it is false
- Power is of *critical importance*
Summing Up

- Uses of power
  - *A priori*: When designing study, select a sample size large enough to detect and effect of magnitude you believe is meaningful
  - *A posteriori*: When test finds no significant difference/association, was there enough power to detect effect of meaningful magnitude?

- (Too little, too late. Can still be used to properly power next study.)
- See http://www.ats.ucla.edu/stat/seminars/Intro_power/ for more.
Statistical Power of a Test

- Significance testing is a balancing act
  - Chance $\alpha$ of making Type I error
  - Chance $\beta$ of making Type II error
  - Reducing $\alpha$ increases $\beta$, and thus reduces the power of a test. It might be tempting to emphasize greater power (the more the better)

- With “too much power” statistical significance may be clinically inconsequential

- A Type II error is not definitive since a failure to reject the null hypothesis does not imply that the null hypothesis is correct

- Since $H_0$ is either always true or false, we are only in danger of making one kind of error or the other (but we have no idea which one)
Factors Affecting Power

- Size of effect an important factor in determining power. Higher probability of detecting larger effects.
- More conservative significance levels (lower $\alpha$) yield lower power. Less power with $\alpha = .01$ than with $\alpha = .05$.
- Increasing the sample size decreases the spread of the sampling distribution and increases power, but there is a trade-off between gain in power and the time/expense of testing a larger sample.
- Larger variance ($\sigma^2$) implies a larger spread of the sampling distribution, ($\sigma/\sqrt{N}$). The larger the variance, the lower the power.
- Variance is partly a property of the population, but can be reduced through careful study design.
Power with Fixed Sample Size

- Many times, $N$ is fixed, either by resource constraints or with secondary data analysis
- In this context, power analysis serves a different function
- Minimum detectable effect (MDE)
- What is the smallest effect size I can detect with power $= (1 - \beta)$, sample size $= N$, and alpha $= \alpha$?
- (Stata Users: db power)
Power with Fixed Sample Size

- Accuracy in parameter estimation (AIPE: http://www.ats.ucla.edu/stat/stata/dae/aipe.htm)
- Bracketing effect sizes (half-width, $w$). For sample size $N$, find range that give $p\%$ chance that the estimated interval will be $\leq 2 \times w$
- The AIPE paradigm is a framework for managing width of confidence interval, independent of effect size
- (Stata Users: findit aipe)
Additional Resources for Power (Books)

- (Too) Simple

- Just Sufficient
Additional Resources for Power (Books)

- **More Contemporary**

- **Extensions**
Additional Resources for Power (Software)

- G*Power
  http://www.gpower.hhu.de/en.html

- R Package pwr
  http://www.statmethods.net/stats/power.html
  http://cran.r-project.org/web/packages/pwr/pwr.pdf

- R Package powerMediation
  http://cran.r-project.org/web/packages/powerMediation/powerMediation.pdf