Power Analysis for PIs

Center for Aging in Diverse Communities Scientist Meeting

January 9, 2019

Steve Gregorich

Motivation

Power analysis is a key aspect of quantitative research design

Appropriately, it serves a critical function in research proposals

If reviewers find fault—rightly or wrongly—with your power analysis, then your proposal will not get a fundable score

Power analyses also are described when reporting results of an RCT

Statisticians conduct power analyses, but need input from you (the PI)

This talk summarizes basic types of input needed for power analyses

I consider continuous (approx. symm. distributed) and binary outcomes

The basic ideas apply to other types of outcomes, e.g., counts.

Overview

Conceptual approaches to power analysis

Independent vs. clustered observations

Independent observations: Continuous *Y*: Needed inputs

Independent observations: Binary *Y*: Needed inputs

Clustered observations: Additional inputs

Clustered observations: Basic concepts

Adjusting for covariates in multivariate models: Additional inputs

Adjusting for non-response: Additional inputs

Reporting effect sizes

Advanced topics

Typical conceptual approaches to power analysis

Specify level of power (e.g., 80%) and alpha (e.g., 0.05)

Option 1

PI specifies

- sampling methodology
- distribution of the *X* and *Y* variables
- effect size

PI seeks estimated sample size needed to detect the specified effect

Option 2

PI specifies

- sampling methodology
- distribution of the *X* and *Y* variables
- anticipated sample size

PI seeks estimated minimum detectable effect size

Option 2 is more common in my experience

Independent vs. clustered observations

If, by design, observations can be considered independent of each other,
then power calculations will be simpler

Example designs where observations are assumed independent

. Sample study participants in an individual/independent fashion

E.g., simple random or convenience sampling design

. *And*, model cross-sectional outcomes

Example designs where observations are clustered

. Sampling clusters, e.g., married couples; doctors & many patients/doctor

. And/or longitudinal assessments made on sampled individuals

If you sample, e.g., 100 couples, then your data set will be of size *N*=200

However, couple members are not independent

Likely, the effective sample size for power analysis will not equal 200

Here, I assume a 2-level nested structure for clustered observations

Independent observations: Continuous *Y*: Needed inputs

1. Continuous *Y*: approx. symmetrically distributed
   Estimated standard deviation of *Y* can be useful
   If not available, you could assume std dev of *Y* equals 1.0

2a. For binary *X*
. Estimated distribution of X, e.g., 50/50, 20/80, 60/40, etc.

2b. For categorical *X*
. Estimated distribution of *X*, e.g., 20/20/60, 40/20/40
. Specify which comparisons are of primary importance
An *X* variable with 3 categories has 3-1=2 degrees of freedom (*df*)
But there are 3 possible pairwise comparisons
For proposals, good to limit the number of comparisons to equal *df*

2c. For continuous *X*: approx. symmetrically distributed
Estimated standard deviation of *X* can be useful
If not available, you could assume std dev of *X* equals 1.0

3. Specify statistical test. e.g., *t*-test, ANOVA, correlation, linear regression

Independent observations: Binary Y: Needed inputs

Same as 2a, 2b, and 2c (above) for continuous *Y* plus…

For binary *Y*

. Provide an estimate of the distribution of Y, e.g., 50/50, 20/80, 60/40 etc.

Specify statistical test. e.g., chi-square, logistic regression

Clustered observations: Additional Inputs

Same needed inputs as for independent observations,
plus the following additional inputs

1. Estimated intra-cluster correlation (ICC) of Y: ρ_Y

2. Two of the following three
   . Estimated number of clusters
   . Estimated average cluster size: B
   . Estimated total sample size

Clustered observations: Basic concepts

Effect of ρ_Y on statistical power depends on the comparison proposed.

E.g., if sampling intact couples, there are two basic comparison types

Between-couple comparisons, e.g., rural vs. urban couples

Here, positive ρ_Y will tend to reduce statistical power

E.g., sampled 100 couples (200 participants)

Members within a couple tend to be more alike

So, for between-couple comparisons, $\rho_Y>0$ will effectively yield $N<200$

Within-couple comparisons, e.g., younger partner vs. *their* older partner

Here, positive ρ_Y will tend to increase statistical power

Positive ρ_Y tends to increase precision of intra-couple differences

So, for within-couple comparisons, $\rho_Y>0$ will effectively yield $N>200$

Adjusting for covariates in multivariate models: Additional Inputs

. Choose a focal *X* variable for your analysis.

. Provide an estimate of the *R*-squared that would be obtained if the focal *X* was regressed onto all other *X* variables planned to be included in the model: R^2_X

In many cases, you will not have an estimate of R^2_X

Your statistician can assume, e.g., $R^2_X = 0.10$

Adjusting for non-response: Additional inputs

At minimum, provide an estimate of the proportion of sample members who will have provided all data needed for any particular analysis. (casewise deletion of missing data): Pr_{comp}

If you have more precise estimates of the proportionate breakdown of sample members with particular patterns of non-response, then that can be useful, as well.

Reporting Effect Sizes

An effect size describes the expected change in *Y* for a given change in *X*

<u>Effect size types</u>
(Minimum) Clinically Important Difference: (M)CID
AKA Clinically Meaningful Difference (CMD)
Usually, CIDs are established by consensus

Effect size expressed in natural units of *X* and *Y*, e.g.,
. +10 years of age (*X*) is expected to decrease QOL (*Y*) by 5 points
. Intervention expected to decrease CAD prevalence from 50% to 40%

Standardized effect sizes (standardized *X*, *Y*, or both)
. Intervention expected to decrease QOL by 0.5 standard deviations
. Intervention expected to halve odds of CAD: OR=0.5
. +1 std dev of age expected to decrease QOL by 0.5 std dev

Good to use (M)CID/CMD, if accessible

Some reviewers don't like standardized effects for continuous *X* and/or *Y*

Advanced. Specific adjustments used in power analysis

Adjust your observed sample size (*N*) for the effects of

- . clustered sampling,
- . covariates, and/or
- . non-response

The adjustment(s) result in the estimated effective sample size: *N*eff

*N*eff is plugged into standard power analysis software, instead of *N*

Adjustment for non-response: $Neff = N \times Pr_{comp}$

Adjustment for covariates: $Neff = N \times (1-R^2_X)$

Adjustment: between-cluster comparison: $Neff = N \div [1+(B-1)\times\rho_Y]$

Adjustment: within-cluster comparison: $Neff = N \div (1-\rho_Y)$

where Pr_{comp}, R^2_X, B, and ρ_Y are defined as above

Multiple adjustments are typical, e.g.,

$$Neff = N \times Pr_{comp} \times (1-R^2_X) \div [1+(B-1)\times\rho_Y],$$

(*B* is re-estimated after first adjusting *N* for non-response & covariates)

Advanced. Clustered sampling designs: ICC of *X* (ρ_X)

Previously, I mentioned ρ_Y

An estimate of ρ_X also may be important

Q: Is your focal *X* variable a subject variable or a design variable?

Subject variables are participant-determined, not investigator-determined

E.g., in a convenience sample, respondent age is a subject variable

Design variables are investigator-determined, e.g.,

. Assignment to intervention versus control group in an RCT

. Scheduled assessment time in a longitudinal study

If your focal *X* is a subject variable, then the statistician also will need an estimate of the intra-cluster correlation of *X* (ρ_X)

…however, many statisticians do not know about this…

END