Power Analysis

Meredith Fowlie
UC Berkeley

September 2021

Introduction

- An unbiased causal estimate is a necessary but not sufficient condition for a successful causal inference study.
- In order for our experimental results to be useful, we need to be able to quantify the uncertainty surrounding these estimates and estimate causal impacts with 'sufficient' precision.
- How we characterize uncertainty around our treatment effect estimates depends on how we think about the underlying data generating process.

Hypothesis testing and 'power analysis'

- In the design stages of a randomized program evaluation, the importance of statistical power calculations cannot be overstated.
- Statistical power refers to the probability that the researcher will reject the null hypothesis when the alternative hypothesis is true.
- It is hard to know what to do with a credible, unbiased estimate of an average treatment effect if it is surrounded by a vast confidence interval.
- An overpowered study may misallocate valuable resources that you could be investing in other research activities.

Overview

- When you are designing an experiment, you make critical decisions about intervention design, sampling frame, sample size.
- Power analysis informs the calibration of sample size and experimental design.
- The 'right' sample size will depend significantly on the nature of the questions we are asking.

Power analysis

Out discussion will focus on:

- Research designs that evaluate the causal impacts of a *binary* treatment on an outcome of interest.
- Testing the null hypothesis of zero average treatment effect in the population.
- Deriving closed form expressions for key research design parameters (such as sample size)
- Laying down foundations for performing simulation-based hypothesis testing.

The basics: What generates uncertainty in the first place?

Assume that the realizations $Y_i(1)$ and $Y_i(0)$ for a particular individual are fixed characteristics of that individual. The observed outcome at unit i can thus be written:

$$Y_i = Y_i(0) + (Y_i(1) - Y_i(0))D_i. \tag{1}$$

The average treatment effect $\hat{ATE} = \frac{1}{N}\sum Y_i(1) - Y_i(0)$ is defined over an underlying population:

$$\tau_{ATE} = E[Y_i(1) - Y_i(0)] \tag{2}$$

Assuming the sample has been randomly drawn from the population, we can make inferences about τ based on what we observe in our sample.

How to think about the process that generates variation in the empirical distribution of τ?

Two perspectives on causal analysis of experimental data:

1. The finite population/randomization inference view that the potential outcomes of the experimental units are fixed and the randomness comes solely from the physical randomization of the treatment assignment.
2. The super population view that the units are an independent sample, and thus the potential outcomes, are independent and identical draws from a hypothetical infinite population.

These two views differs conceptually and mathematically, resulting in different sampling variances of the usual difference-in-means estimator of the average causal effect.

Randomization inference perspective?

- The assignment mechanism is the only source of randomness.
- The sampling distribution of a test statistic under the null hypothesis can be constructed by simulating all possible random assignments.
- This generates a distribution of treatment effect estimates under the null. The standard deviation of this sampling distribution is the standard error.
- Expectations are taken over this *randomization distribution*.

Randomization inference perspective

- The only source of randomness is the draw of the assignment vector from the population of possible assignment vectors.
- Consider a canonical example: the 'sharp' null hypothesis of no effect of the treatment on any unit in a population (against the alternative hypothesis that the treatment impacts at least one unit).
- How might we use simulate the empirical distribution of a test statistic under this sharp null?

Randomization inference perspective

Infinite super-population perspective

- The researcher obtains a sample of size N from a finite population which is itself just one realization of a theoretical super-population.
- The assignment mechanism D determines which of the potential outcomes we observe for each individual in the sample.
- Under this perspective, potential outcomes $Y_i(0)$ and $Y_i(1)$ and any observed covariates used in the analysis are drawn from a joint distribution and are stochastic.
- The expectation is taken over the distribution that is generated by random sampling from this super-population.
- The sample joint probability distribution $\Pr(Y, D, X)$ converges to the true distribution in the super-population.

Does it matter what perspective you take?

- We will use both perspectives in our power analysis – different approaches lend themselves to different situations.
- The variance of the difference-in-means ATE estimator under an infinite population model is similar- but different- from the variance formulation implied by the randomization inference perspective.
- Strictly speaking, the infinite population variance estimate is a conservative (large) estimate of the finite population variance of the difference in average outcomes.
- The (small) difference in variance formulae is unidentifiable and typically assumed away.

Power analysis

Organizing framework for this discussion is the hypothesis test:

1. Define a specific conjecture about a population parameter and pre-specify your estimating equation.
2. State a null and alternative hypothesis clearly (one-tailed or two-tailed test)
3. Compute a test statistic using data.
4. Compare the test statistic against a critical value.
5. Draw a conclusion (reject or do not reject the null hypothesis) and interpret.

The basics: What are you estimating?

It is critical that you specify the equation(s) that you plan to estimate in the design stage of your experiment!

- The power calculations that inform key research design decisions are based on a specific estimating equation (or equations).
- Pre-analysis plans reduce problems associated with data mining and specification searching by setting out in advance exactly the specifications that will be run and with which variables.
- PAPs are the new industry standard!!

See the AEA registry: https://www.socialscienceregistry.org/

Motivating our estimating equation

Recall that the observed outcome at unit i can be written:

$$Y_i = Y_i(0) + (Y_i(1) - Y_i(0))D_i. \tag{3}$$

Assume (for now) that the estimand we are interested in is the average treatment effect $\hat{ATE} = \frac{1}{N}\sum Y_i(1) - Y_i(0)$.

This is defined over the underlying population:

$$\tau_{ATE} = E[Y_i(1) - Y_i(0)] \tag{4}$$

OLS meets RCT

Suppose that we estimate the following simple linear regression model:

$$Y_i = \alpha + \tau D_i + \varepsilon_i, \tag{5}$$

- This can be interpreted as a "structural" estimating equation with α representing $Y_i(0)$ and τ representing $(Y_i(1) - Y_i(0))$.

What is captured by ε?

OLS meets RCT

Suppose that we estimate the following simple linear regression model:

$$Y_i = \alpha + \tau D_i + \varepsilon_i, \tag{5}$$

- This can be interpreted as a "structural" estimating equation with α representing $Y_i(0)$ and τ representing $(Y_i(1) - Y_i(0))$.

What is captured by ε?

- The idiosyncratic variation in the control response plus the idiosyncratic variation in treatment effects.

τ an unbiased estimate of the ATE

OLS estimates are those that minimize the sum of squared residuals

$$\min : \sum_{i=1}^{N} (Y_i - \alpha - \tau D_i)^2 \tag{6}$$

Solving for $\hat{\tau}$:

$$\frac{\partial(\sum e^2)}{\partial \tau} = (Y_i - \alpha - \tau D_i) D_i = 0$$

After some rearranging:

$$\widehat{\tau} = \frac{(Y_i - \overline{Y})(D_i - \overline{D})}{(D_i - \overline{D})^2} = \frac{\sum_{i:D_i=1} Y_i}{PN} - \frac{\sum_{i:D_i=0} Y_i}{(1-P)N}.$$

This OLS regression (with no covariates) provides an unbiased estimate of τ.

Multiple regression meets RCT

- Regression methods provide a convenient way to incorporate "pre-treatment"covariates.
- In expectation, these covariates are identically distributed in our treatment and control groups. **Why include covariates??**

Multiple regression meets RCT

- Regression methods provide a convenient way to incorporate "pre-treatment" covariates.
- In expectation, these covariates are identically distributed in our treatment and control groups. **Why include covariates??**
  1. Increase precision by controlling for some of the variation in outcomes explained by random variation in observable attributes across groups.
  2. Test for heterogeneous treatment effects along observable dimensions.
  3. Stratified randomization was used.

Aside: Stratified randomizaton

- Stratified block randomization involves grouping patients into strata defined by baseline characteristics, and performing block randomization within each stratum.
- This ensures that the stratifying variables are balanced.
- If you use stratified randomization. adjusting for the baseline variables used in the stratification can yield valid standard error estimates.

RCT meets OLS?

Once you've implemented an RCT and you have data on Y_i, D_i, and X_i for a random sample of individuals, OLS is a standard go-to for estimating the average effect of the treatment on the outcome Y.

> *"There is a disconnect between the way the conventional assumptions in regression analyses are formulated and the implications of randomization. As a result it is easy for the researcher using regression methods to go beyond analyses that are justified by randomization, and end up with analyses that rely on a difficult-to-assess mix of randomization assumptions, modeling assumptions, and large sample approximations."*
>
> *Athey and Imbens (2016)*

What are they talking about??

Regression analysis of experimental data

$$Y_i = \alpha + \tau D_i + \beta X_i + \varepsilon_i, \tag{7}$$

Concerns?

Regression analysis of experimental data

$$Y_i = \alpha + \tau D_i + \beta X_i + \varepsilon_i, \tag{7}$$

Concerns?

- Once we introduce additional covariates, we introduce structure that does not immediately follow from the underlying model of the data generating process:
- Here we assume that the effect of the continuous covariate X_i on the outcome variable Y_i is linear and additive. ..but where does this come from?
- **Critique**: Including covariates in the regression equation can introduce bias in finite samples because randomization does not guarantee zero correlation between covariates and treatment assignment in finite samples.

Punchline

With reasonably sized samples, multivariate regression allows us to estimate treatment effects and standard errors using standard machinery while adjusting for covariates to improve precision.

- Include covariates and their interactions with treatment indicators in a fully interacted regression (treatment-by-covariate regression)
- Rather than adjusting for covariates afterwards, stratify your randomization on them beforehand (but make adjustments in the ex post analysis!).
- Report unadjusted estimates (even if your primary analysis conditions on covariates).
- Pre-specify your primary analysis in a pre-analysis plan at the design stage of your study.

Power analysis

Organizing framework for this discussion is the hypothesis test:

1. Define a specific conjecture about a population parameter and pre-specify your estimating equation.
2. State a null and alternative hypothesis clearly (one-tailed or two-tailed test)
3. Compute a test statistic using data.
4. Compare the test statistic against a critical value.
5. Draw a conclusion (reject or do not reject the null hypothesis) and interpret.

Step 2: Specify a null and alternative hypothesis

- For ease of exposition, we define $H_0 : \tau = 0$.
- Because the OLS estimate of τ is a random parameter, it would be an astonishing coincidence if $\widehat{\tau}$ were equal to the population average in any finite sample.
- We must take this sampling variation into account if we are to make inferences about τ.
- We construct/calibrate a test statistic that has a known distribution when the null hypothesis is true.
- If the value of the test statistic is an extreme one that would rarely be encountered by chance under the null, this is evidence against the null.

Step 3: Construct a test statistic

We have shown that the OLS estimate of τ is unbiased.
We assume $\varepsilon \sim iid\mathcal{N}(0, \sigma_\varepsilon^2)$. Using the standard z-score :

$$z = \frac{\widehat{\tau} - \tau}{\sqrt{var(\widehat{\tau})}},$$

where z measures the difference between the observed statistic and its hypothesized population parameter in units of the standard deviation. A test of our null hypothesis is based on the test statistic:

$$\frac{\widehat{\tau}}{\sqrt{var(\widehat{\tau})}}$$

How to calibrate this test statistic??

Sampling variance of the ATE estimate

Suppose the data are generated according to the following model:

$$Y_i = \alpha + \tau D_i + \varepsilon_i, \tag{8}$$

where ε_i is distributed $\sim iid\mathcal{N}(0, \sigma_\varepsilon^2)$.
The variance of this regression residual is estimated as:

$$\hat{\sigma}^2_{Y|D} = \frac{1}{N-2}\sum(Y_i^{obs} - \widehat{Y}_i)^2,$$

where the predicted value $\widehat{Y}_i$ is $\hat{\alpha}$ if $D_i = 0$ and $\hat{\alpha} + \hat{\tau}$ if $D_i = 1$. (This assumes homogeneous treatment effects).
The variance of the OLS estimate τ is given by :

$$var(\widehat{\tau}) = \frac{\hat{\sigma}^2_{Y|D}}{\sum(D_i - \overline{D})^2}$$

Sampling variance of the ATE estimate

Note that $(D_i - \overline{D}) = (1 - P)$ when $D_i = 1$ and $-P$ when $D_i = 0$. Thus:

$$\sum (D_i - \overline{D})^2 = PN(1 - P). \tag{9}$$

The sampling variance of $\hat{\tau}$ is estimated as:

$$var(\widehat{\tau}) = \frac{\sigma^2_{Y|D}}{PN(1 - P)}. \tag{10}$$

The variance is increasing with the residual variance (numerator) and decreasing with the variability in the treatment indicator (denominator).

Add covariates

Adding covariates to this model can reduce the residual variance:

$$Y_i = \alpha + \tau D_i + \beta X_i + \gamma X_i \cdot D_i + \varepsilon_i, \tag{11}$$

If we maintain our $\sim iid\mathcal{N}(0, \sigma_\varepsilon^2)$ assumption, the variance of our OLS estimate:

$$var(\widehat{\tau}) = \frac{\sigma^2_{Y|D,X}}{PN(1-P)}. \tag{12}$$

Including covariates can reduce the residual variance (but eats up degrees of freedom).

Step 4: Compare the test statistic against a critical value

Our hypothesis test combines this test statistic with a rejection rule. Reject the null if:

$$(\mathrm{t}_{1-\kappa} + t_{\alpha}) < \frac{\widehat{\tau}}{\sqrt{var(\widehat{\tau})}}$$

$$\widehat{\tau} > (t_{1-\kappa} + t_{\alpha})\sqrt{var(\widehat{\tau})}$$

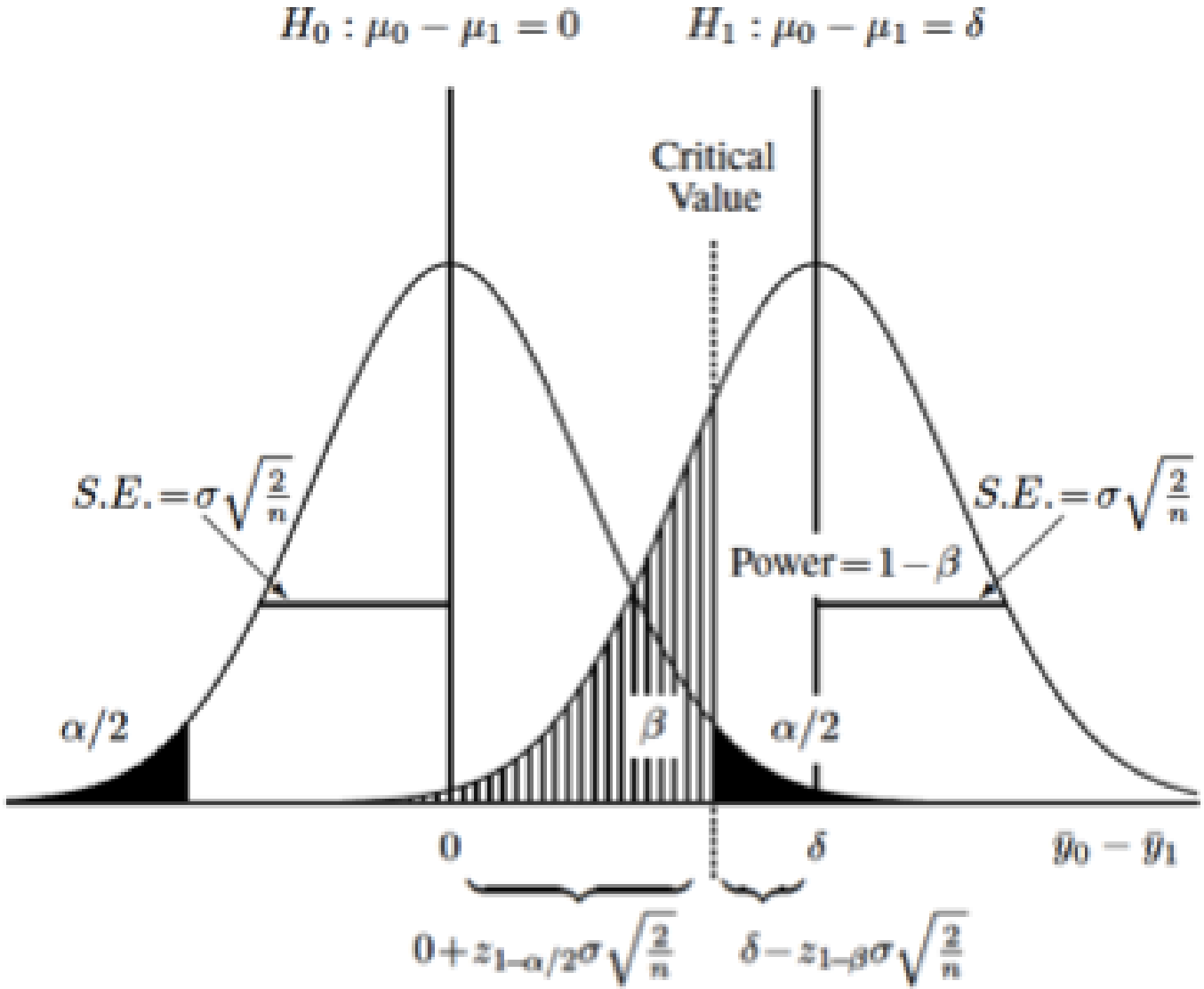
$$\widehat{\tau} > (t_{1-\kappa} + t_{\alpha})\sqrt{\frac{\sigma^2_{Y|D,X}}{PN(1-P)}}$$

- α: the probability, under the null, that τ falls into the rejection region.
- κ: the probability that a test will correctly reject the null (the power of the test).

Key parameter value choices

- **Type I error rate** (α). This is the probability of rejecting the null hypothesis when it is true. Popular values of α include .05 and .01.
- **Type II error rate** (β). Type II error: fail to reject a false null. Note that estimating the Type II error rate for a given test requires estimating the distribution of the alternative hypothesis (which is usually unknown).
- **The "power" of a test** (κ): The ability to reject the null hypothesis (i.e. detect an effect) when the alternative hypothesis is true. Thus $\kappa = 1 - \beta$.

Popular value of $\kappa = .80$ implies that 80% of the experiments with sample size N will correctly reject the null hypothesis when the null is false.


$H_0 : \mu_0 - \mu_1 = 0$
$H_1 : \mu_0 - \mu_1 = \delta$
Critical Value
$S.E. = \sigma\sqrt{\frac{2}{n}}$
$S.E. = \sigma\sqrt{\frac{2}{n}}$
Power $= 1 - \beta$
$\alpha/2$
β
$\alpha/2$
0
δ
$\bar{y}_0 - \bar{y}_1$
$0 + z_{1-\alpha/2}\sigma\sqrt{\frac{2}{n}}$
$\delta - z_{1-\beta}\sigma\sqrt{\frac{2}{n}}$

Minimum detectable effect (MDE)

- To achieve a desired level of power κ , given our chosen α and the sample size N, it must be that the "true" population average effect is greater than or equal to some minimum parameter value τ.
- Mechanically, the MDE measures the smallest treatment effect τ that a research design can detect given α, κ, N.
- How do we choose this MDE?

Putting the pieces together

Given the desired level of power κ given our chosen significance level α, we will reject the null if the true effect is:

$$\tau > (t_{1-\kappa} + t_{\alpha})\sqrt{var(\widehat{\tau})}, \tag{13}$$

where $t_{1-\kappa}$ and t_{α} are critical values found in a t-table.

Note that once we have selected the desired power κ, size α, P, and N, and variance estimate, we can derive the minimum detectable effect as a function of:

- The number of units in the study (N).
- Type I error rate (α).
- The desired level of statistical power (κ).
- The proportion of the sample receiving the treatment (P). Note that power is maximized by setting $P = 0.5$.
- The residual variance σ^2).

Putting the pieces together

$$MDE = (t_{1-\kappa} + t_{\alpha}) \sqrt{\frac{1}{P(1-P)}} \cdot \sqrt{\frac{\sigma_{\varepsilon}^2}{N}}$$

Rearranging, we can solve for the required sample size given α, κ, MDE, P, and σ^2.

$$N = \frac{(t_{1-\kappa} + t_{\alpha})^2}{P(1-P)} \frac{\sigma_{\varepsilon}^2}{MDE^2}.$$

This tells us how many observations we need to test a pre-specified null hypothesis.

Calibration?

Putting the pieces together

$$MDE = (t_{1-\kappa} + t_{\alpha}) \sqrt{\frac{1}{P(1-P)}} \cdot \sqrt{\frac{\sigma_{\varepsilon}^2}{N}}$$

Rearranging, we can solve for the required sample size given α, κ, MDE, P, and σ^2.

$$N = \frac{(t_{1-\kappa} + t_{\alpha})^2}{P(1-P)} \frac{\sigma_{\varepsilon}^2}{MDE^2}.$$

This tells us how many observations we need to test a pre-specified null hypothesis.

Calibration?
Note that N, α, κ, P, MDE are all research design parameters chosen by the researcher. The variance is typically estimated using any potentially informative data you can find.

RED power calculations look a little different..

- We only use the variation in the treatment assignment variable D_i that is generated by the encouragement to identify the causal effect of the treatment.
- Let c denote the share of households that receive treatment in the encouraged group.
- Let s denote the share treated in the control group.

The OLS coefficient in a regression of outcomes on the encouragement indicator is used to construct the LATE estimate:

$$Y_i = \alpha + \pi Z_i + \varepsilon_i, \tag{14}$$

Recall the point estimate of π is divided by $(c - s)$ to obtain an unbiased estimate of the LATE.

Power calculations and REDs

The variance of this LATE estimator is:

$$\begin{aligned} var(LATE) &= var(\frac{\pi}{c-s}) \\ &= \frac{1}{(c-s)^2}\frac{1}{P(1-P)}\frac{\sigma^2}{N} \end{aligned}$$

The minimum detectable effect can be computed as:

$$MDE = (t_{1-\kappa} + t_{\alpha})\sqrt{var(\frac{\widehat{\pi}}{c-s})} \quad (15)$$

$$= (t_{1-\kappa} + t_{\alpha})\sqrt{\frac{1}{P(1-P)}\frac{\sigma^2}{N}\left(\frac{1}{(c-s)}\right)} \quad (16)$$

RED power analysis

$$MDE = (t_{1-\kappa} + t_{\alpha})\sqrt{var(\frac{\widehat{\pi}}{c-s})} \tag{17}$$

$$= (t_{1-\kappa} + t_{\alpha})\sqrt{\frac{1}{P(1-P)}\frac{\sigma^2}{N}\left(\frac{1}{(c-s)}\right)} \tag{18}$$

Rearranging, we can solve for the required sample size given α, κ, MDE, P, c, s, and σ^2.

$$N = \frac{(t_{1-\kappa} + t_{\alpha})^2}{P(1-P)}\frac{\sigma^2}{MDE^2}\cdot\frac{1}{(c-s)^2}. \tag{19}$$

Take note! Imperfect compliance sucks (power)

$$N = \frac{(t_{1-\kappa} + t_{\alpha})^2}{P(1-P)} \frac{\sigma^2}{MDE^2} \cdot \frac{1}{(c-s)^2}. \tag{20}$$

As compared to an RCT in which all households comply with their treatment assignment, the number of households required to obtain a given level of statistical power increases by a factor of $\frac{1}{(c-s)^2}$.

Cluster randomization

- Many research designs involve randomizing groups (or clusters) of units to treatments.
- Suppose proportion P of clusters are randomly assigned to treatment.
- Let T be the number of time periods in which data are collected.
- We observe outcome Y_{it} for cluster i in time t .
- Observations within a cluster may be exposed to a common shock that induces correlation among the observed outcomes at these units (but is not caused by the intervention).

Cluster randomization

The estimating equation can now be represented as a two level hierarchical model with observations nested in clusters indexed by i:

$$\begin{aligned} Y_{it} &= \alpha + \tau D_{it} + \varepsilon_{it}, \\ \varepsilon_{it} &= v_i + \epsilon_{it} \end{aligned}$$

Simplifying assumptions (to get us started...)

- The v_i (cluster-specific error component) are distributed $\sim iid\mathcal{N}(0,\sigma_v^2)$
- The ϵ_{it} idiosyncratic error component is distributed $\epsilon_i \sim iid\mathcal{N}(0, \sigma_\epsilon^2)$.
- Homogeneous treatment effects.

Intra-cluster correlation

- The implications of using cluster-level (versus unit-level) randomization depends on how much of the variation in the outcome Y_{it} is explained by group-level variation versus the variation in ϵ_{it}.
- Accounting for similarities among subjects (i.e. correlated residuals) within a cluster almost always results in a net loss of power, requiring increased total subject recruitment (to achieve a given level of statistical precision).
- The intracluster correlation coefficient relates the within-group variance with the between-group variance:

$$\rho = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_\epsilon^2}. \tag{21}$$

Intra-cluster correlation

The intra-cluster correlation coefficient,

$$\rho = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_\epsilon^2}. \tag{22}$$

- In the extreme case of multiple observations per unit, and no within unit variation in potential outcomes $(Y_i(0), Y_i(1))$, $\rho = 1$ and statistical power is not increasing in T.
- Additional observations on a household buy us nothing. So we should conduct the power calculation as if we have only J observations (versus JxT).

Power calculations that accommodate cluster design:

$$MDE = (t_{1-\kappa} + t_{\alpha})\sqrt{var(\widehat{\tau})}. \quad (23)$$

$$= (t_{1-\kappa} + t_{\alpha})\sqrt{\frac{1}{P(1-P)}\frac{\sigma_e^2}{JT}(1 + T\rho - \rho)} \quad (24)$$

$$= \frac{(t_{1-\kappa} + t_{\alpha})}{\sqrt{P(1-P)J}}\sqrt{\left(\rho + \frac{1-\rho}{T}\right)\sigma_e} \quad (25)$$

- If there is no within group correlation, ρ=0 and we are back to the RCT power calculation we started with.
- If $\rho = 1$, additional observations within a cluster do not buy you anything!

Power calculations that accommodate this cluster randomized design:

Calculate the number of clusters required to detect *MDE* with the desired level of precision:

$$J = \frac{(t_{1-\kappa} + t_{\alpha})^2}{P(1-P)} \frac{\sigma_e^2}{MDE^2} \left(\rho + \frac{1-\rho}{T} \right) \tag{26}$$

When ρ is small, there is more to be gained from collecting repeated measures on the outcome of interest.

Panel data?

- McKenzie(2012) extends this analysis to accommodate pre-treatment data and cluster-level fixed effects.
- But his approach assumes constant serial correlation (i.e. iid errors after removing unit fixed effects).
- And with panel data, we might expect serial correlation within individuals (even after sweeping out individual fixed effects).
- Burlig, Preonas, and Woerman (2019) derive the variance of panel estimators under arbitrary error structures.
- They give us formulas (and Stata code!) for difference-in-difference estimators that are robust to arbitrary serial correlation.

Data generating process:

They assume that the data are generated according to the following:

$$Y_{it} = \beta + \tau D_{it} + \upsilon_i + \delta_t + \omega_{it}$$

Assumptions:

- Strict exogeneity: $E[\omega_{it}|X] = 0$. This follows from random assignment.
- Balanced panel
- Independence across units: $E[\omega_{it}\omega_{js}|X] = 0 \forall i \neq j$
- Symmetric covariance structures.
- Let m denote pre-treatment periods and r denote post-treatment periods.

$$var(\hat{\tau}) = \frac{1}{P(1-P)J}\left[\underbrace{\left(\frac{m+r}{mr}\right)\sigma^2_\omega}_{\text{idiosyncratic}}\right.$$

...and extend it for arbitrary serial correlation

$$var(\hat{\tau}) = \frac{1}{P(1-P)J}\left[\underbrace{\left(\frac{m+r}{mr}\right)\sigma_\omega^2}_{\text{idiosyncratic}} + \underbrace{\left(\frac{m-1}{m}\right)\psi^B}_{\text{pre-period}} + \underbrace{\left(\frac{r-1}{r}\right)\psi^A}_{\text{post-period}} - \underbrace{2\psi^X}_{\text{across periods}}\right]$$

σ_ω^2 : Idiosyncratic variance

ψ^B : Average covariance between pre-treatment observations

ψ^A : Average covariance between post-treatment observations

ψ^X : Average covariance across pre- and post-treatment observations

BPW Contributions

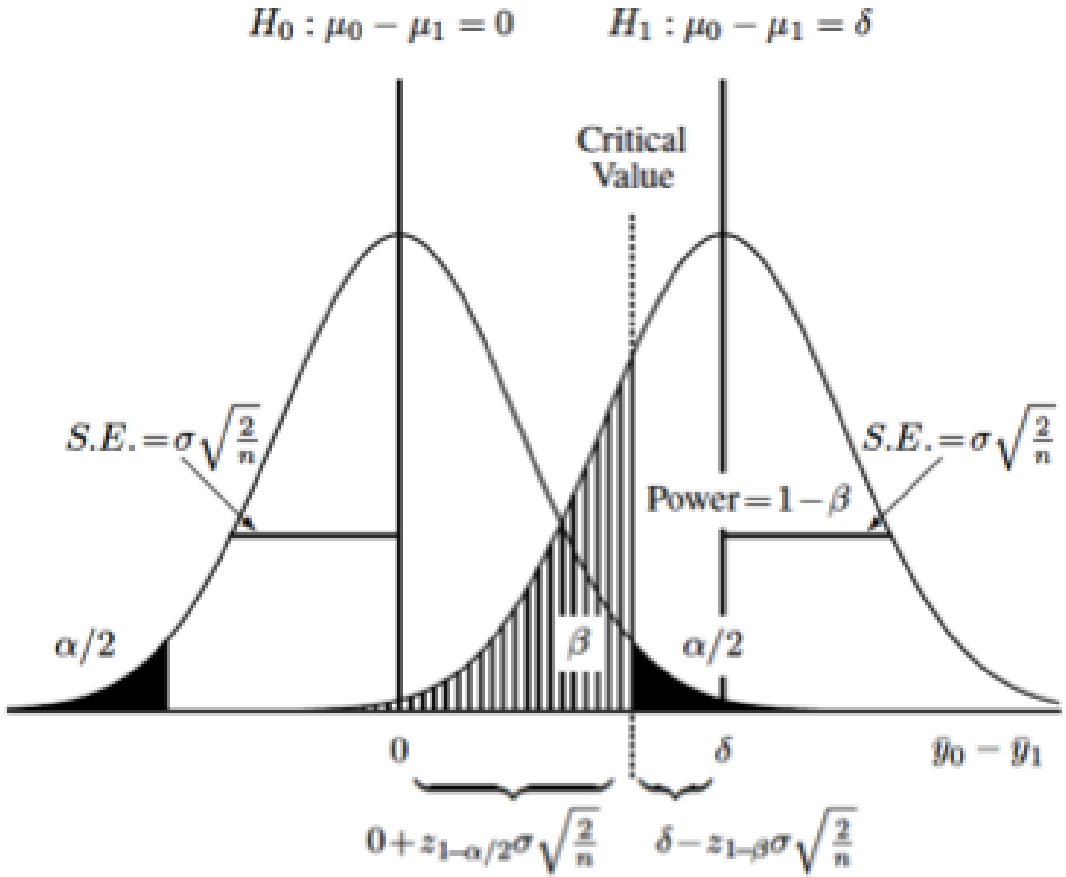
Analytics provide a way to do power calculations with non-i.i.d. errors...

But calibration not so straightforward! If you have relevant data, you should also pursue a simulation-based approach.

Benefits:

- Easily accommodates more complicated designs
- Allows for testing multiple models
- Avoids (large) headaches involved in calibrating general analytical formulas

In a data-rich environment, opt-for simulation-based power analysis!



The optimal treatment fraction?

- If the objective is simply to maximize power, we will generally want to set the treatment fraction P =0.5 so as to minimize $\frac{1}{P(1-P)}$.
- There are exceptions to this rule of thumb (e.g. when ρ varies across treatment and control groups).
- In practice, we typically find ourselves maximizing power subject to a binding budget constraint.
- The constrained optimum treatment fraction may well be less than 0.5....why?

The best way to understand power analysis...

Is to do it!

- Assignment 1 posted on bcourses.
- Read lecture notes and BPW(2019).
- Two different sets of energy consumption data. Use these to inform experimental research designs.
- Asks you to play around with simulation-based analysis (along with analytical).