Power analysis

Statistic Modeling & Causal Inference

Agenda

- Lecture review
  - Power analysis

- Power analysis in R

Some slides rely heavily on the following two resources, check them out if you want to dig deeper:

- https://www.povertyactionlab.org/resource/power-calculations
- https://nickch-k.github.io/EconometricsSlides/Week_08/Power_Simulations.html

What is statistical power?

- The probability of rejecting the null hypothesis (H_0) when it is in fact false (i.e., being correct by finding an effect that is there)

What is statistical power?

- The probability of rejecting the null hypothesis (H_0) when it is in fact false (i.e., **being correct by finding an effect that is there**)
- We can also think of it as the ability to distinguish a **signal** from **noise**

What is statistical power?

- The probability of rejecting the null hypothesis (H_0) when it is in fact false (i.e., **being correct by finding an effect that is there**)
- We can also think of it as the ability to distinguish a **signal** from **noise**

As prospective policy analysts, the signal we will be looking for is likely the **impact of a treatment** on some **outcome**.

What is statistical power?

As prospective policy analysts, the signal we will be looking for is likely the impact of a treatment on some outcome.

What is statistical power?

As prospective policy analysts, the **signal** we will be looking for is likely the **impact of a treatment** on some **outcome**.

Let's think about these **signals** and **noise**.

What is statistical power?

As prospective policy analysts, the signal we will be looking for is likely the impact of a treatment on some outcome.

Public health intervention and the incidence of a rare disease

Education intervention and income

What is statistical power?

As prospective policy analysts, the signal we will be looking for is likely the impact of a treatment on some outcome.

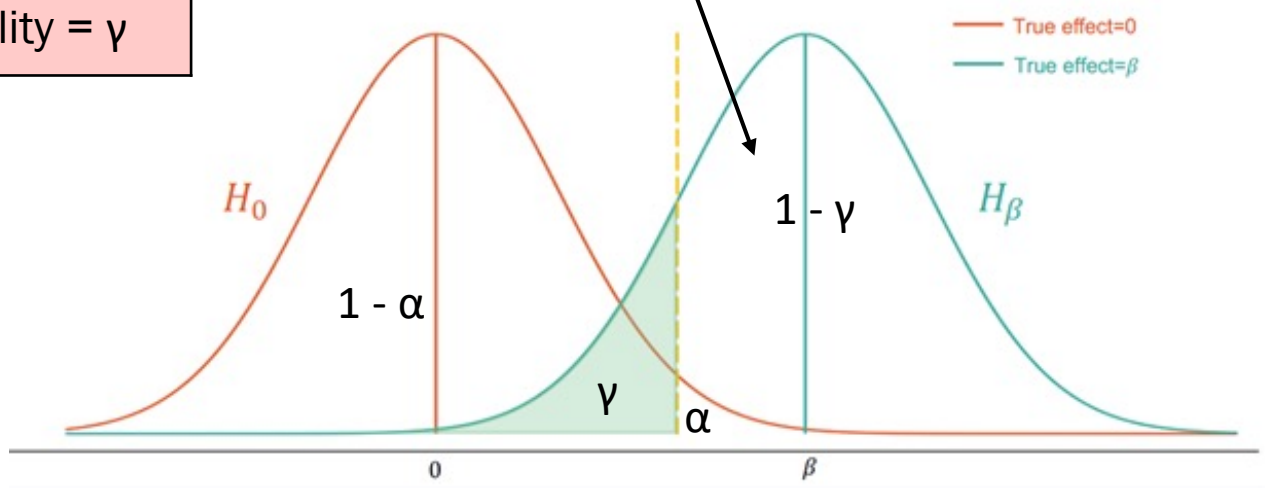
Public health intervention and the incidence of a rare disease

Education intervention and income

Statistical power

Null hypothesis is	TRUE	FALSE
Rejected	Type I error False positive Probability = α	Correct decision True positive Probability= 1 - γ
Not rejected	Correct decision True negative Probability = 1 - α	Type II error False negative Probability = γ

Statistical power:
The probability of rejecting H0 when H1 is indeed TRUE



What is power analysis about? I/III

Power analysis is about the probability of detecting a significant effect in your statistical analysis

In practice, power calculations are either used

- to find the minimum sample size needed to detect an effect of a certain size with a minimum statistical power (and given the other parameters of your analysis); or
- to determine the minimum detectable effect (MDE) size that you can detect given a sample size and a minimum level of statistical power that you want to achieve

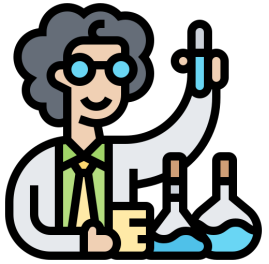
Power analysis is especially **relevant for experiments** in which you can control the sample size

What is power analysis about? II/III

There are a couple of things that determine how powered our experiments are:

Strength of the treatment

Amount of **noise**



Experimental **design features**

What is power analysis about? III/III

A power analysis for the study of the relationship between X (treatment) and Y (outcome) involves five things:

1. The size of the effect (e.g. coefficient in a regression)
2. The amount of variation in the treatment (the variance of X, say)
3. The amount of other variation in Y (the R2, or the variation from the residual after explaining Y with X)
4. Statistical precision (the standard error of the estimate, statistical power, i.e. the true-positive rate)
5. The sample size

A power analysis holds four of these constant and tells you what the fifth can be.

Relationship between power and power components

Component	Relationship to power	Relationship to MDE
Sample size increases	Increases power	Decreases the MDE
Other variation in outcome increases	Decreases power	Increases the MDE
True effect size increases	Increases power	n/a
Confidence-level increases (e.g. from 95% to 99%)	Decreases power	Increases the MDE

Illustration 1: Changing variation/standard error of $\hat{\beta}$

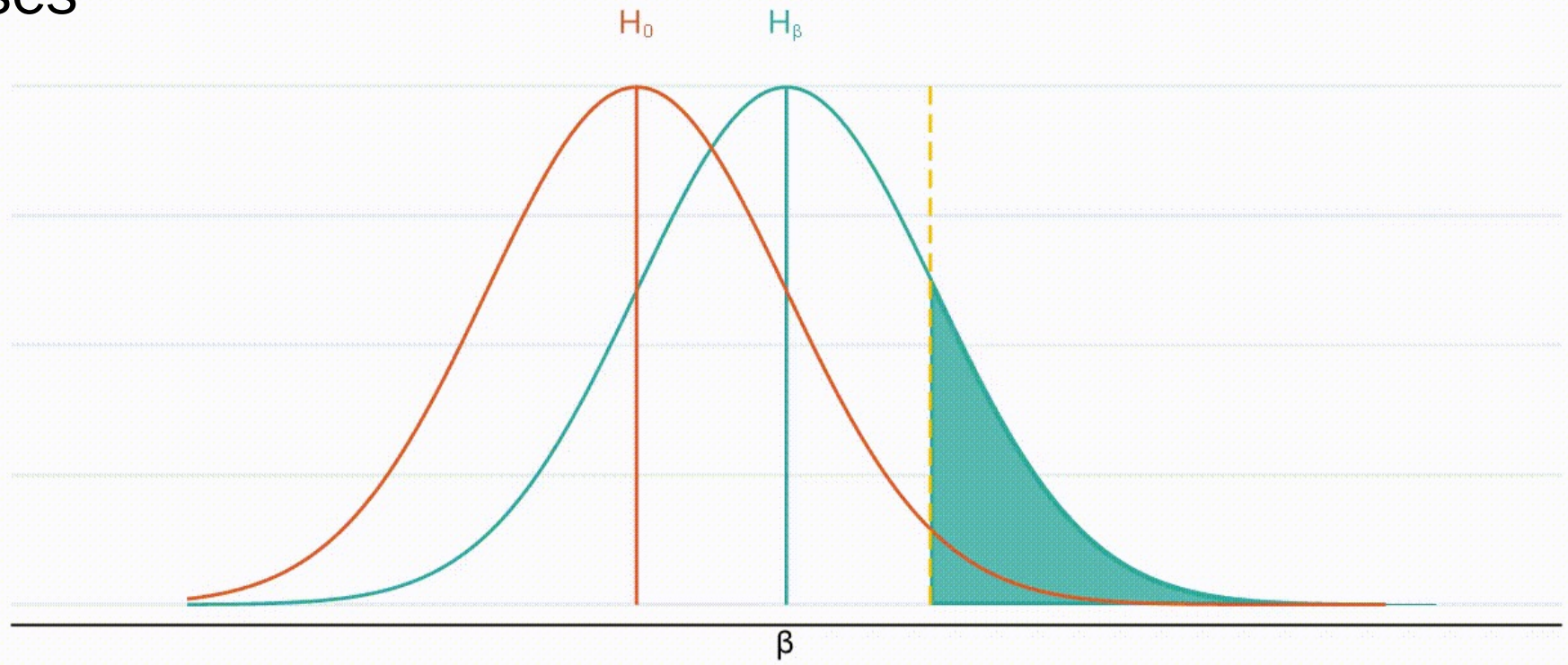
Variation/standard error of $\hat{\beta}$ decreases→ statistical power increases

Standard error of $\hat{\beta}$ decreases if:

- Sample size increases
- Other variation in Y decreases

Illustration 2: Changing value of true effect

Absolute value of the true effect increases → statistical power increases



Practical recommendations

- Power is **usually set at 80%**; for particularly important studies where you absolutely want to avoid a type II error, consider 90% or 95%.
- Do a power analysis **early in the research process**
- Run **sensitivity analyses** to learn which "ingredient" has the largest impact on the statistical power of your analysis
- Power analysis is **not an exact science**: If you don't have data on some components, make a guess based on the info you have. Power analysis is just supposed to provide orders of magnitudes of the sample size you need or the MDE that you can achieve.

Further Resources

For any coding issues – Stackoverflow

Hertie's Data Science Lab – Research Consulting