Power

Matthew Shane Loop, PhD

2021-08-26

You will often be consulting with a statistician to justify your sample size, particularly through a power calculation. However, it can be helpful for early study planning for you to understand the components of power and how to calculate it for common scenarios. Therefore, we will go over statistical power and its uses, its limitations, and the `simglm` package for calculating it.

I will teach you to calculate power through a completely new framework. You won't have to memorize any equations.

Learning objectives

1. List the components of power and explain how changing each component, keeping all others equal, will affect the power.
2. Describe the limitations of using only power to design a study.
3. Using `R`, calculate power for a t-test, ANOVA, ANCOVA, and logistic regression model.

List the components of power and explain how changing each component, keeping all others equal, will affect the power.

Power is the probability[1] of rejecting[2] a specific null hypothesis, given a specific alternative hypothesis is true. $P[p < 0.05|\delta = 5]$ is different from $P[p < 0.05|\delta = 10]$. The problem is, we don't know δ. Estimating δ is the point of the study. For a power analysis δ is chosen to be: (a) not obviously wrong; and (b) large enough to lead to a small enough sample size / large enough power to be attractive to a funding agency or feasible. Therefore, estimates of power are often *optimistic*. This optimism is one reason why many trials fail to reject the null hypothesis.

[1] Why is it a probability and not a certainty?

[2] How do you reject a null hypothesis?

Power calculations rely on many assumptions, so any specific estimate of power is likely to be wrong. But an advantage of a power calculation is it helps you think ahead about your study design. The components needed to calculate power are:

- uncertainty in your estimate[3];
- sample size;
- the false positive rate you can tolerate; and
- magnitude of the treatment effect[4].

[3] This quantity is often calculated by software by asking for the variance of the outcome.

[4] What are some examples of how we measure the magnitude of a treatment effect?

How might you change each of these to increase power?

Describe the limitations of using only power to design a study.

When you focus on "rejecting" or "failing to reject" a null hypothesis, you focus on only a subset of the ways you can be wrong. Other ways to be wrong are errors in estimating the magnitude (type M) and sign (type S) of the treatment effect.[5] These errors are probably more important to the clinician and patient and type 1 and 2 errors.

We tend to think a trial with high power will produce the "correct" answer, but remember that power relies on *many* assumptions.

Equations for power must be designed "bespoke", so you often use a formula that is simplistic compared to your actual study.

Other quantities might be more interesting when designing a study besides the probability of rejecting a specific null hypothesis given a specific alternative hypothesis is true.[6]

[5] Gelman A, Carlin J. Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors. *Perspect Psychol Sci*. 2014;9: 641–651.

[6] Try to think of as many ways to finish the question, "What's the probability that. . . ?". Simulation studies must be used to answer these other, possibly more meaningful, questions.

Using `R`, calculate power for a t-test, ANOVA, ANCOVA, and logistic regression model.

Almost all of your power calculations will revolve around a treatment, a set of covariates to adjust for, and an outcome that is either continuous or binary. You usually will consult with a statistician to obtain a sample size justification, but it can help you to perform back-of-the-envelope power calculations on your own during the very early stages of designing a study.

A benefit of the simulation method we will demonstrate is that assumptions are *highly explicit*.

We will use the `simglm` package in `R` to simulate data and calculate power.[7]

[7] If you haven't used `R` before, watch this video to get started.