

11.3: Factors influencing statistical power

Introduction

Components of an experimental design that may influence statistical power include:

1. α (probability of committing a Type I error).
As the probability of a Type I error increases, the probability of a Type II error decreases. Therefore, as α gets bigger, power gets decreases.
2. σ , the variance in the population
Statistical power decreases as variability increases.
3. Effect size
The effect size is a measure of the differences between two or more groups considered biologically important; its difficult to have lots of power to detect very small differences.
4. Sample size
As n increases, power increases.

Size of alpha (Type I error)

When we introduced the idea of Type I error, alpha, it followed a story about the *Challenger*, the space shuttle that disintegrated after the failure of O-rings allowed hot gasses to cross joints in the rockets. If we adopt a Type I error rate of 5% in our rocket designs, then failure is expected at one in twenty launches. That rate clearly is unacceptable, so the logical extension of this thinking would be to decrease the Type I error rate. Unfortunately, this comes at the expense of increasing our risk of Type II error. Thus, decreasing Type I error decreases statistical power.

Variance

If the range of values for individuals in the samples are great, then it should not be surprising that the power to distinguish between sample means from the groups will be low. Conversely, if the variance for a sample is small, then the precision of the estimated sample mean will increase, and, therefore, the power will increase.

Effect size

Effect size deserves some more comments. If we are thinking cause/effect, then we are asking whether or not our independent variable explains a lot of variation in the response (dependent) variable. If there is a strong link between the independent and dependent variable, then the effect size will be large and small numbers of observations will be needed.

There are various ways to estimate effect size, but the simplest is a variation of the t-test (Cohen's d),

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s}$$

The formula would be different for ANOVA (involves the mean squares), but you get the idea. By convention, an effect size of about 0.2 would be "small," 0.5 would be "medium," and an effect size greater than 0.8 would be "large."

Sample size

This is the area of course where the experimenter has control. We can choose how many individuals are assigned to treatment groups. The Central Limit Theorem basically states that you can use the normal distribution to predict how likely an individual observation is in relation to a sample mean even if the sample distribution is not normally distributed. The larger the number of individuals in the sample from a population, the more confidence we have about making this assumption. Translation: sample size directly impacts standard error of the calculated statistic. Recall our equation for the standard error of the mean.

Alternatives to Cohen's d

Glass's g

Hedges' g_u

Keselman and colleagues' d_j

estimators based on **trimmed mean** and **Winsorized variances**.

Questions

1. The Central limit Theorem can be invoked when we have large samples of observations. In this subchapter we state that increase sample size increases statistical power. Discuss and contrast these two characteristics of large sample size on statistical inference.
2. How many samples are enough? Some statistical textbooks will cite a rule of 30. With respect to factors that affect statistical power, discuss the limitations of adopting such a rule to design an experiment.

This page titled [11.3: Factors influencing statistical power](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Michael R Dohm](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.