13.1 The Will to Power

Zongyi Liu

2023-06-26

13.1 The Will to Power

Researchers collect data in order to achieve a goal, which can be used to explore the world.

Statistical power is the probability of achieving the goal of an empirical study, when a suspected underlying state of the world is true.

13.1.1 Goals and Obstacles

There are many possible goals for an experiment or observational study.

Example: we might want to show that the rate of recovery for patients who take a drug is higher than the rate of recovery for patients who take a placebo. We might want to show that a coin is biased, i.e., that its tendency to show heads is not equal to 0.5. Goals such as those involve showing that a suspected difference or value really is tenable, or, complementarily, showing that a "null" value is not tenable. Other goals do not necessarily have a suspected value in mind. Instead, the goal is merely to put a desired degree of precision on whatever tendencies happen to be observed.

- **Goal**: Demonstrate that believable parameter values exclude a "null" value.
- **Formal Expression**: Show that the 95% HDI excludes the "null" value, or its ROPE.
- **Goal**: Achieve precision in the estimate of believable values.
- **Formal Expression**: Show that the 95% HDI has some specified maximal width.

In some research, the goal might be to demonstrate that a "null" value is true, rather than false.

If we knew the benefits of achieving our goal, and the costs of pursuing it, and if we knew the penalties for making a mistake while interpreting the data, then we could express the results of the research in terms of the long-run expected payoff.

The crucial obstacle to the goals of research is that a random sample is only a probabilistic representation of the population from which it came.

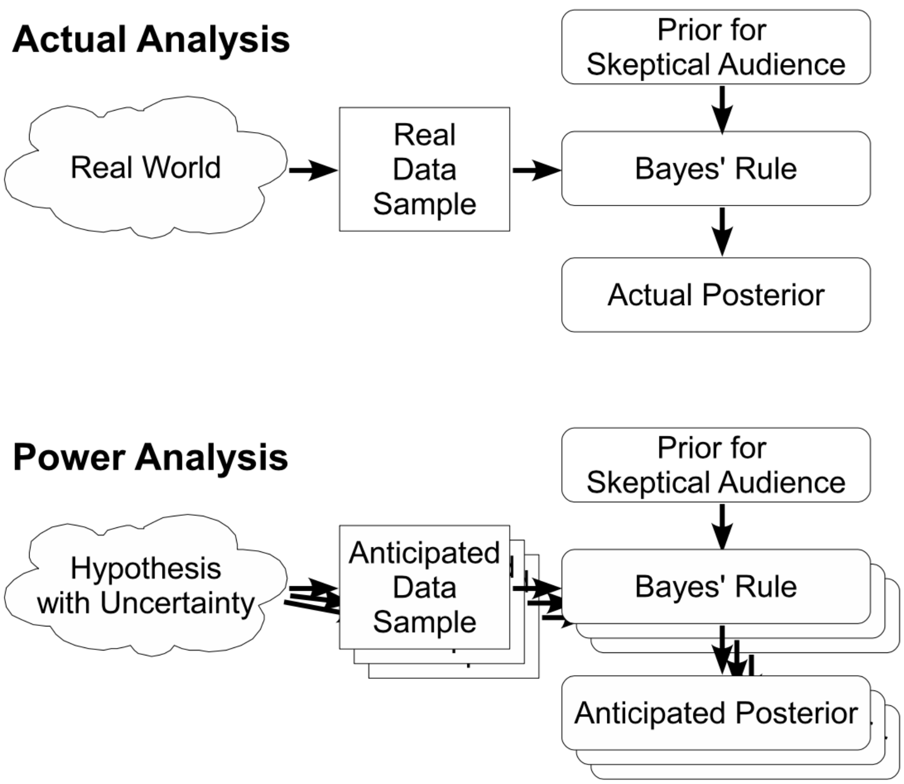
13.1.2 Power

Because of random noise, the goal of a study can be achieved only probabilistically. The probability of achieving the goal, given the (suspected) true state of the world, is called the **power** of the experiment.

In general, power can be approximated in those manner:

1. Generate a random sample of data points, using the data-generating hypothesis. The sample should be generated according to how actual data will be gathered in the eventual real experiment. For example, typically it is assumed that the number of data points is fixed at N. It might instead be assumed that data will be collected for a fixed interval of time T , during which data points appear randomly at a known mean rate n/T.
2. Compute the posterior estimate, using Bayesian analysis with skeptical-audience priors. The data analysis must be convincing to the audience, and therefore the analysis must use priors that are agreeable to that skeptical audience.
3. From the posterior estimate, tally whether or not the goal was attained. The goal could be any of those outlined previously, such as having the 95% HDI exclude a ROPE around the null value, or having the 95% HDI be narrower than a desired width.
4. Repeat above steps many times, to approximate the power. Power is, by definition, the proportion of times that the goal is attained.

The process of computing power can be illustrated as below:



13.1.3 Sample Size

Typically, power increases as sample size increases. Because gathering data is costly, we would like to know the minimal sample size, or sample duration, which is required to achieve a desired power.

Generally, the larger the sample size is, the more precise the estimation is.

13.1.4 Other Expressions of Goals

There are other ways to express mathematically the goal of precision in estimation. For example, another way of using HDIs was described as an "average length criterion", which requires that the average HDI width, across repeated simulated data, does not exceed some maximal value L.

A rather different mathematical expression of precision is entropy of a distribution. Entropy describes how spread out a distribution is, so smaller entropy connotes a more precise distribution. A distribution that consists of an infinitely dense spike, that has an infinitesimally narrow width, has zero entropy. At the opposite extreme, a uniform distribution has maximal entropy. The goal of high precision in the posterior distribution might be re-expressed as a goal of small entropy in the posterior distribution.

Also, goal of excluding a null value can be expressed as wanting a large Bayes factor in a model comparison between the spike-null prior and the automatic alternative prior.