review_hypothesis_testing

October 29, 2020

Review Basics of Hypothesis Testing

1 1. Resources

- Discovering Statistics using R, Field, A. *et al.*, 2012
- Statistical Rethinking, McElreath, R., 2015
- All of Statistics, Wasserman, L., 2004

1.1 1.1 Review of Core Concepts

- Bayesians, Frequentists, and Likelihoodists
- There are a few approaches to statistical inference:
  - Bayesian
  - Likelihoodist
  - Frequentist

We will be concerned primarily with the frequentist approach.

1.1.1 1.1.1 What is hypothesis testing?

Hypothesis testing is the process of using data to make decisions under uncertainty.

1.1.2 1.1.2 What is hypothesis testing? (cont.)

The frequentist approach is typically choosing between 2 competing hypotheses.

- Null hypothesis (usually written H_0)
- Alternative hypothesis (usually written H_1 or sometimes H_A)

1.1.3 1.1.3 What is hypothesis testing? (cont.)

For example, we might be interested in whether some new medication, M, reduces cholesterol. Here the competing hypotheses are:

H_0: $\mu_1 = \mu_2$ M does not reduce cholesterol (null hypothesis)

H_1: $\mu_1 < \mu_2$ M reduces cholesterol (alternative hypothesis)

where μ_1 is mean cholesterol for those receiving M in the population, and μ_2 is mean cholesterol for those *not* receiving M in the population.

1.1.4 1.1.4 Notes on hypothesis testing

Some important things to note:

1. Previous example is one-sided test; two-sided tests generally look like:

- H_0: $\mu_1 = \mu_2$
- H_1: $\mu_1 \neq \mu_2$

2. Two-sided tests tend to be more common
3. You should clearly articulate hypotheses *priori to conducting statistical tests*

1.1.5 1.1.5 Notes on Hypothesis Testing (cont.)

General process of hypothesis testing:

1. Specify the null and alternative hypotheses, H_0 and H_1
2. Determine the test to be used, which gives us:

- Our test statistic
- Corresponding probability distribution

3. Set a level of significance (e.g., $\alpha = 0.05$)
4. Use our data to compute our test statistic (and perhaps its standard error)
5. Use test statistic and its accompanying distribution to obtain p-value

2 2. Review of p-values

What is a p-value?

2.1 2.1. Understanding p-values

A p-value is a probability. In particular, it is the probability of finding data as extreme or more extreme than what he have observed, given that the null hypothesis is true.

2.1.1 2.1.1 Understanding p-values (cont.)

In other words, a p-value can be used to answer this question:

If the null hypothesis is true, are my data unusual?

When a p-value is small, our answer is "yes". And when the answer is "yes", we are generally inclined to take this as evidence against the null hypothesis.

2.1.2 2.1.2 Understanding p-values (cont.)

A p-value is **NOT**:

- The probability the null hypothesis is true
- The probability that the data were produced by chance alone
- A measure of effect size
  - Be wary of papers discussing "highly" or "extremely" significant results based p-values
  - Also beware of studies using p-values as inputs to subsequent computations or tests

image

2.1.3 2.1.3 Understanding *p*-values (cont.)

Other notes on *p*-values:

1. Their use is controversial in some circles
2. Can be easily abused to show significant results
3. Despite limitations, they are ubiquitous in science

- We have used them for so long, it's hard to change course (but Bayesians are trying!)
- For many applied researchers and practitioners, they are a convenient way to turn observed data in to a "yes"/"no" decision

3 3. The Decision Problem

Ultimately, we want to be able to draw conclusions and make decisions based on data

3.1 3.1 Deciding between H_0 and H_1

So, how do we choose between our hypotheses?

1. Our default is to believe H_0

2. We use our data to determine if we have sufficient reason to reject H_0

3. This is where we rely on work from probability theory

3.1.1 3.1.1 Deciding between H_0 and H_1 (cont.)

Because we are relying on probabilistic reasoning about whether or not to reject H_0, we can be wrong.

3.1.2 3.1.2 Deciding between H_0 and H_1 (cont.)

Question:
How do we know when we have committed a Type I error or a Type II error?

3.1.3 3.1.3 Deciding between H_0 and H_1 (cont.)

Answer:

In general, we cannot know unequivocally when we have committed a Type I error or a Type II error.

This has important implications:

1. Replication is *absolutely crucial* in science
2. Must be *hyper vigilant* about inflated Type I error from repeated testing (more on this later)
3. Should be generally skeptical, and especially so for low power studies with "sexy" results