

Statistical Rethinking Chapter 1

Benjamin Wee

01/07/2018

Statistics Curriculum

- ▶ What are *your* thoughts on statistics at University?
 - ▶ Curriculum
 - ▶ Easy/hard/intuitive
 - ▶ Useful?
 - ▶ “Real world” application
 - ▶ p-values, confidence intervals, ‘tests’, linear models, GLMs, MLE, etc

“Instead of a single method for building, refining and critiquing statistical models, students are offered a zoo of pre-constructed golems known as “tests”. . . Teaching engineering backwards, start with bridge building then ending with basic physics.”

Chapter Overview

- ▶ Argues for the 'rethinking' of classical strategic procedures and the paradigms of empirically testing hypothesis
- ▶ Moreover, encourages us to think about *how* we quantify our results and what it means to use models in scientific endeavours
- ▶ Outlines the use of bayesian methods, multilevel modelling and information criteria as the foundation to analyse non-null models of natural phenomena
- ▶ An introduction to statistical *modelling* and *inference*

1.1 Statistical Golems

- ▶ Statistics as a branch of engineering, a diverse set of golems for various contexts.

“Eventually you will make a statistical model which will misbehave. But it will only misbehave according to your expectations. It is behaving exactly according to its design”

- ▶ Classical tools are not diverse enough to handle many common research questions. Need both statistical expertise and domain knowledge. Otherwise we may wreck Prague.
- ▶ A unified theory of model (golem) engineering. Principles for design, building, refining special purpose statistical procedures.

1.2 Statistical Rethinking

- ▶ Challenging the tacit belief that proper objective statistical inference is falsifying hypothesis

*Science is not desirived by the falsification standard. . . .
deductive falsification is impossible in nearly every
scientific context*

1. Hypothesis are not models.
2. Measurement matters

1.2.1 Hypothesis are not models

- ▶ Binary decisions based on null hypothesis consistent with several process models
- ▶ Hypothesis, process, statistical models.

Flasifying null models are insufficient for understanding how the world works.

1.2.2 Measurement Matters

- ▶ The logic of falsification – *modus tollens*, the method of destruction
 - ▶ Seeking disconfirming evidence is important, but prone to two kinds of error.
1. Observation Error (Type 1/2 error or false positives/negatives). Not always “black or white” outcome observed, what if we see a noisy measure or a “grey” picture? Measurement matters
 2. Continuous hypothesis - Frequency estimation vs distributions (e.g. 80% of swans are white). Hard to apply falsification in this case. Requires careful construction of null hypothesis

Falsification is consensual not logical.

1.3 Three tools for golem engineering

1. Bayesian Data Analysis (interpretation of probability vs frequentism, laplacian statistics)
2. Multilevel models
3. Model comparison using information criteria (cross validation metric)

Final thoughts

We construct models to do things which are difficult for us. We design them to be good at things we are bad at. But they are fantastically bad at the things people find easy. They are blind to our intentions.

Everything in Bayes is pseudo random, there is a deterministic process (probability distributions which produce random numbers). We taught computers to generate pseudo random numbers (we taught them a deterministic way to generate “randomness”)

Propositional logic. What is a valid deduction. Bayes extends this to continuous logic.

We want to understand “the golem’s belief”, not the scientist’s.