

# Mulling Over McElreath's Statistical Rethinking

Dan Burrell

2020-07-28



# Contents

1	Before you start, you should do these things	5
2	The Golem of Prague	7



# Chapter 1

## Before you start, you should do these things

We're working through the second edition of McElreath's Statistical Rethinking text. We need to access the associated R package `rethinking` and also load the `tidyverse`. Use the following code:

```
if(!require(pacman)) install.packages("pacman")
library(pacman)

p_load(coda, mvtnorm, loo, dagitty)
p_load_gh("rmcelreath/rethinking")
```

Below is an example of the generic format for STANcode. The blocks need to occur in the order specified, however only the `model` block is necessary (the others are optional, depending on the needs of the modeler).

```
modelString = "

  data {
    ...declarations... // This is a comment
  }

  transformed data {
    ...declarations ... statements ...
  }

  parameters {
    ...declarations...
  }

  model {
    ...statements...
  }

}
```

```

transformed parameters {
  ... declarations ... statements ...
}

model {
  ...declarations ... statements ...
}

generated quantities {
  ...declarations ... statements ...
}
" # Close quote for modelString

```

Translate the model into C++ code and compile into an executable dynamic shared object (DSO) using `stan_model()` from the `rstan` package:

```

p_load(rstan)
stanDSO = stan_model( model_code=modelString )

```

Once the DSO is created, it can be used for generating a Monte Carlo sample from the posterior distribution. For example:

```

# Create some fictitious data:
N = 50; z = 10; y = c(rep(1,z), rep(0,N-z))
dataList = list( y=y, N=N)
stanFit = sampling( object=stanDSO, data=dataList,
                    chains=3, iter=1000, warmup=200, thin=1)

# Load rjags, coda, and DBDA2E functions
source("DBDA2E-utilities.R")

```

## Chapter 2

# The Golem of Prague

Statistical models are like powerful yet dangerous robots. They will do their task, but they have no discernment as to whether their task is appropriate. For their particular task, they are good. But set to work on a task they're not intended for, they yield untrustworthy results.

Classical statistical methods are fragile and inflexible. They are inflexible in that they exhibit very limited adaptability to unique research contexts. They are fragile in that they fail in unpredictable ways when applied to new contexts. This is important because at the cutting edge of research it's rarely ever clear which procedure is appropriate.

McElreath's point is that classical statistical tools are not diverse enough to handle many common research problems. Moreover, statistical tools on their own only understand association, and can tell us nothing about cause and effect. So, rather than having a tool kit of statistical "robots" and risk their misapplication, what is needed is a unified set of engineering principles for designing, constructing and refining purpose-built statistical procedures.

Some thoughts on this:

- You need to understand how a statistical procedure processes information in order to be able to reasonably interpret its output;
- detailed knowledge is a requirement;
- getting greasy under the hood is how you cultivate this sort of deep understanding — doing the computations the hard way (at least initially);
- we also need a statistical epistemology — an appreciation of how statistical models relate to hypotheses and the natural mechanisms of interest.
- we need to do away with the null hypothesis significance testing mindset that is born from unscrupulous application of Popper's falsificationism; good science is done by developing statistical procedures that can falsify hypotheses, but this ought to be done thoughtfully and with adequate

insight into the limitations of deductive falsification.

McElreath argues that deductive falsification is impossible in nearly every scientific context because:

1. Hypotheses are not models. There is not a one-to-one correspondence between hypotheses and models, so that strict falsification is impossible.
2. Measurement matters. Even when we think our data do falsify a model, the trustworthiness (or representativeness) of the data is always up for debate.

The upshot is that the scientific method cannot be reduced to a statistical procedure; and hence our scientific methods ought not pretend. That's the arrow shot at the heart of Null hypothesis significance testing (NHST), which is often identified with falsificationist, or Popperian, philosophy of science.

You can label chapter and section titles using `{#label}` after them, e.g., we can reference Chapter `??`. If you do not manually label them, there will be automatic labels anyway, e.g., Chapter `??`.

Figures and tables with captions will be placed in `figure` and `table` environments, respectively.

```
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

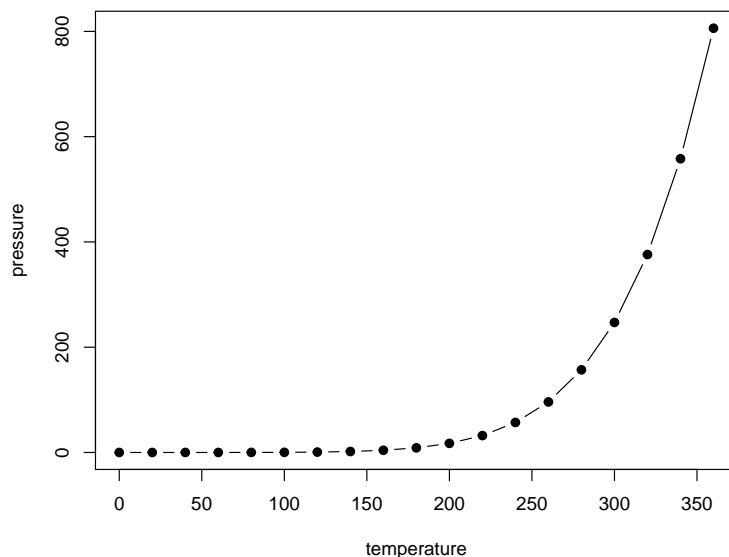


Figure 2.1: Here is a nice figure!

Reference a figure by its code chunk label with the `fig:` prefix, e.g., see Figure 2.1. Similarly, you can reference tables generated from `knitr::kable()`, e.g.,



Table 2.1: Here is a nice table!

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa
5.1	3.5	1.4	0.3	setosa
5.7	3.8	1.7	0.3	setosa
5.1	3.8	1.5	0.3	setosa

see Table 2.1.

```
knitr::kable(
  head(iris, 20), caption = 'Here is a nice table!',
  booktabs = TRUE
)
```

You can write citations, too. For example, we are using the **bookdown** package (?) in this sample book, which was built on top of R Markdown and **knitr** (Xie, 2015).



# Bibliography

Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.