

Study project: undamentals of Causal Inferences
With R

François Lefebvre

2022-02-01

Contents

Study Project

About

This is a study project of ?. Many thanks to Babette Brumback for a great book to tackle the question of causality. It is *extremely* in my professional life which is not about making predictions but rather come up with intervention plans (business intel).

The suggestions for errata are in the section *Errata*. Comments are also included in the section *Comments*.

Packages

The functions have been rewritten to simplify them and improve the learning experience (personal opinion) for newbies such as me.

The following packages are so useful that they could not be avoided. In particular, `dplyr` and `tidyr` make the coding experience so much more interesting that one could almost claim they have become the standard in R coding.

- `dplyr` for data wrangling. see ?
- `tidyr` for data wrangling. see ?
- `ggplot2` for plots, see ?
- `gt` for all tables, see ?
- `dagitty` for analysis of structural causal models, see ?
- `ggdag` for directed acyclic graphs, see ?
- `gee`: a generalized estimation equation solver. Introduced in chapter 4. See ?.
- `MonteCarlo`: To perform Monte Carlo simulations. Introduced in chapter 4, section 4.2. See ?.
- `simstudy`: To run all sorts of simulations. A great tool to learn right from the start since simulations are so important. See ?.

Errata

Preface

page xi, last word of first paragraph is **standardidzation**, s/b *standardization*

Chapter 1

Section 1.2.3.2, p. 11

The sentence of the 6th line on top of the page is “We simulated the data according to the **hyothetical**”, s/b *hypothetical*

Chapter 2

Figure 2.1, p. 30

This is really a small detail. The caption of the bottom plot is $\hat{E}_{np}(Y \mid A = 1, H = 1, T = 1)$, s/b \hat{E}_{np}

Chapter 3

Typography: section 3.2 p. 40, equation 3.1

The current latex expression of conditional independence used seems to be $(Y(0), Y(1)) \perp\!\!\!\perp \text{II} T$ with the output

$$(Y(0), Y(1)) \text{ II } T$$

a better typography would be $\perp\!\!\!\perp$ for the symbol \perp . When used for equation 3.1 as $(Y(0), Y(1)) \perp\!\!\!\perp T$ we obtain

$$(Y(0), Y(1)) \perp\!\!\!\perp T$$

Comments

Chapter 2

section 2.4 p. 31

The second sentence of the last paragraph on p. 33 says

We also need the `car` package in order for the `summary()` function to operate on boot objects the way we describe.

This sentence is **not required** if we use the `boot::boot.ci()` which simplifies `lmodboot.r()` and does not require the `car` package. See the code in this document for `lmodboot.r` in chapter 2.

Chapter 4

Section 4.1

See the plots in section 4.2. They could be helpful to visualize the changes in effect measures from one level of modifier to the other.

Section 4.2

Monte Carlo Simulation

A Monte Carlo is provided in section 4.2 and coded in a function called `betasim_effect_measures()`. It uses the *Beta* distribution. It is helpful in that it

- confirms the same results as in ?

- is less CPU intensive as it needs only 5000 iterations to confirm ?
- is easier to code than `java` and uses `R` which is the declared language of ?
- allows some extra flexibility with the shape parameters of *Beta* to investigate the conclusion with different curves. See the suggestion for applications below.

page 72, Figure 4.1

The probabilities shown in the Venn diagram do not add up to 100% because, for example, the event that RR changes in the same direction as RD but not in the same direction as the other two measures [...]. It would be awkward to arbitrarily set one of those 2 chances as zero.

? mentions that it is the result of *not mutually exclusive events*. That is true. Yet, these events, properly grouped are actually mutually exclusive. In section 4.2 they are called **Opposite pairwise events**. Using these definitions then yes, they are mutually exclusive but cannot be properly shown in the Venn diagram. This can be easily solved by splitting the probabilities. See section 4.2 for details.

The end result is a proper partitioning of the sample space Ω and is, in fact, a σ -field (See ?, section 1.2). Yet it does not change the conclusions reached in ?. Actually, it reinforces them as this point is **extremely important** when using probabilities and statistics.

Applications

See my sub-section 4.2 called *Applications* where 2 possible applications are mentioned.

- Data pre-processing (data cleaning)
- Bayesian prior for Beta-binomial model

Exercises

Exercise 1

Using the causal power, the conclusion is different than the official answer. It is not obvious why the official solution does not make use of the *causal power*.

Exercise 5

The official solution uses `gee` with the default family, that is `gaussian`.

Since the outcome *attend* is binary isn't it better to use the `binomial` family?

We quote p. 50 from chapter 3 in that respect

Because our outcome is binary, we choose to fit the logistic parametric model

Chapter 5

The `dagitty` and `ggdag` are used extensively.

Chapter 1

Hello bookdown

All chapters start with a first-level heading followed by your chapter title, like the line above. There should be only one first-level heading (#) per .Rmd file.

1.1 Setup

Make sure you tell GitHub that the web site is not to be build via Jekyll, since the **bookdown** HTML output is already a standalone website. See section 6.3 of bookdown for details.

```
# create a hidden file .nojekyll  
# to tell GitHub that the website is not to be build via Jekyll  
a_file <- file.path(getwd(), ".nojekyll")  
if (!file.exists(a_file)) file.create(a_file)
```

1.2 A section

All chapter sections start with a second-level (##) or higher heading followed by your section title, like the sections above and below here. You can have as many as you want within a chapter.

An unnumbered section

Chapters and sections are numbered by default. To un-number a heading, add a {.unnumbered} or the shorter {-} at the end of the heading, like in this section.

Chapter 2

Cross-references

Cross-references make it easier for your readers to find and link to elements in your book.

2.1 Chapters and sub-chapters

There are two steps to cross-reference any heading:

1. Label the heading: `# Hello world {#nice-label}`.
 - Leave the label off if you like the automated heading generated based on your heading title: for example, `# Hello world = # Hello world {#hello-world}`.
 - To label an un-numbered heading, use: `# Hello world {-#nice-label}` or `{# Hello world .unnumbered}`.
2. Next, reference the labeled heading anywhere in the text using `\@ref(nice-label)`; for example, please see Chapter `??`.
 - If you prefer text as the link instead of a numbered reference use: any text you want can go here.

2.2 Captioned figures and tables

Figures and tables *with captions* can also be cross-referenced from elsewhere in your book using `\@ref(fig:chunk-label)` and `\@ref(tab:chunk-label)`, respectively.

See Figure `??`.

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

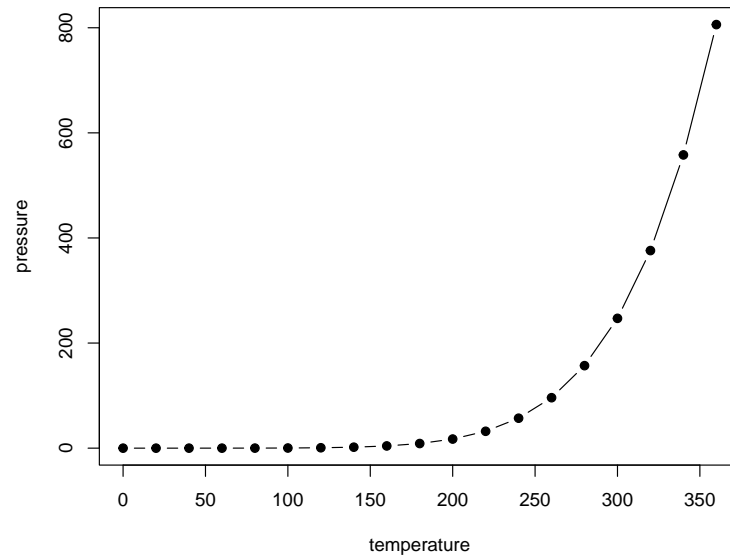


Figure 2.1: Here is a nice figure!

Don't miss Table ??.

```
knitr::kable(  
  head(pressure, 10), caption = 'Here is a nice table!',  
  booktabs = TRUE  
)
```


Table 2.1: Here is a nice table!

temperature	pressure
0	0.0002
20	0.0012
40	0.0060
60	0.0300
80	0.0900
100	0.2700
120	0.7500
140	1.8500
160	4.2000
180	8.8000

Chapter 3

Parts

You can add parts to organize one or more book chapters together. Parts can be inserted at the top of an .Rmd file, before the first-level chapter heading in that same file.

Add a numbered part: `# (PART) Act one {-}` (followed by `# A chapter`)

Add an unnumbered part: `# (PART*) Act one {-}` (followed by `# A chapter`)

Add an appendix as a special kind of un-numbered part: `# (APPENDIX) Other stuff {-}` (followed by `# A chapter`). Chapters in an appendix are prepended with letters instead of numbers.

Chapter 4

Footnotes and citations

4.1 Footnotes

Footnotes are put inside the square brackets after a caret `^[]`. Like this one ¹.

4.2 Citations

Reference items in your bibliography file(s) using `@key`.

For example, we are using the **bookdown** package (?) (check out the last code chunk in `index.Rmd` to see how this citation key was added) in this sample book, which was built on top of R Markdown and **knitr** (?) (this citation was added manually in an external file `book.bib`). Note that the `.bib` files need to be listed in the `index.Rmd` with the YAML `bibliography` key.

The RStudio Visual Markdown Editor can also make it easier to insert citations: <https://rstudio.github.io/visual-markdown-editing/#/citations>

¹This is a footnote.

Chapter 5

Blocks

5.1 Equations

Here is an equation.

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad (5.1)$$

You may refer to using `\@ref{eq:binom}`, like see Equation (??).

5.2 Theorems and proofs

Labeled theorems can be referenced in text using `\@ref{thm:tri}`, for example, check out this smart theorem ??.

Theorem 5.1. *For a right triangle, if c denotes the length of the hypotenuse and a and b denote the lengths of the **other** two sides, we have*

$$a^2 + b^2 = c^2$$

Read more here <https://bookdown.org/yihui/bookdown/markdown-extensions-by-bookdown.html>.

5.3 Callout blocks

The R Markdown Cookbook provides more help on how to use custom blocks to design your own callouts: <https://bookdown.org/yihui/rmarkdown-cookbook/custom-blocks.html>

Chapter 6

Sharing your book

6.1 Publishing

HTML books can be published online, see: <https://bookdown.org/yihui/bookdown/publishing.html>

6.2 404 pages

By default, users will be directed to a 404 page if they try to access a webpage that cannot be found. If you'd like to customize your 404 page instead of using the default, you may add either a `_404.Rmd` or `_404.md` file to your project root and use code and/or Markdown syntax.

6.3 Metadata for sharing

Bookdown HTML books will provide HTML metadata for social sharing on platforms like Twitter, Facebook, and LinkedIn, using information you provide in the `index.Rmd` YAML. To setup, set the `url` for your book and the path to your `cover-image` file. Your book's `title` and `description` are also used.

This `gitbook` uses the same social sharing data across all chapters in your book—all links shared will look the same.

Specify your book's source repository on GitHub using the `edit` key under the configuration options in the `_output.yml` file, which allows users to suggest an edit by linking to a chapter's source file.

Read more about the features of this output format here:

<https://pkgs.rstudio.com/bookdown/reference/gitbook.html>

Or use:

```
?bookdown::gitbook
```