Common Mistakes

Contents

1 Common Mistakes **1**

Configuring R . . . 1

1.1 Parameters versus random variables . . . 2

1.2 R . . . 2

1.2.1 Don't copy-paste code . . . 2

1.3 Quarto . . . 2

1.3.1 Separate divs and slide breaks . . . 2

1.3.2 `library(printr)` currently breaks `df-print: paged` . . . 3

1.4 LaTeX . . . 3

1 Common Mistakes

Configuring R

Functions from these packages will be used throughout this document:

```
library(conflicted) # check for conflicting function definitions
# library(printr) # inserts help-file output into markdown output
library(rmarkdown) # Convert R Markdown documents into a variety of formats.
library(pander) # format tables for markdown
library(ggplot2) # graphics
library(ggfortify) # help with graphics
library(dplyr) # manipulate data
library(tibble) # `tibble`s extend `data.frame`s
library(magrittr) # `%>%` and other additional piping tools
library(haven) # import Stata files
library(knitr) # format R output for markdown
library(tidyr) # Tools to help to create tidy data
library(plotly) # interactive graphics
library(dobson) # datasets from Dobson and Barnett 2018
library(parameters) # format model output tables for markdown
library(haven) # import Stata files
library(latex2exp) # use LaTeX in R code (for figures and tables)
library(fs) # filesystem path manipulations
library(survival) # survival analysis
library(survminer) # survival analysis graphics
library(KMsurv) # datasets from Klein and Moeschberger
library(parameters) # format model output tables for
library(webshot2) # convert interactive content to static for pdf
library(forcats) # functions for categorical variables ("factors")
library(stringr) # functions for dealing with strings
library(lubridate) # functions for dealing with dates and times
```

Here are some R settings I use in this document:

```
rm(list = ls()) # delete any data that's already loaded into R

conflicts_prefer(dplyr::filter)
ggplot2::theme_set(
  ggplot2::theme_bw() +
        # ggplot2::labs(col = "") +
    ggplot2::theme(
      legend.position = "bottom",
      text = ggplot2::element_text(size = 12, family = "serif")))

knitr::opts_chunk$set(message = FALSE)
options('digits' = 6)

panderOptions("big.mark", ",")
pander::panderOptions("table.emphasize.rownames", FALSE)
pander::panderOptions("table.split.table", Inf)
conflicts_prefer(dplyr::filter) # use the `filter()` function from dplyr() by default
legend_text_size = 9
run_graphs = TRUE
```

1.1 Parameters versus random variables

The parameters of a probability distribution shouldn't involve the random variables being modeled:

> ⚠ This is wrong
>
> $$X \sim Pois(\lambda)$$
> $$\hat{\lambda}_{ML} \to_D N(\bar{X}, \lambda/n)$$

Solution.

$$\hat{\lambda}_{ML} \to_D N(\lambda, \lambda/n)$$

Expectations are means, not sums, despite the similarity of Σ and E. Really, we should use μ instead of E.

1.2 R

1.2.1 Don't copy-paste code

Successful programmers don't use copy-paste! Write functions instead.[1]

1.3 Quarto

1.3.1 Separate divs and slide breaks

Make sure not to put a div `:::` on the next line after a slide break `---`:

```
---
::: notes
:::
```

There needs to be an empty line between them:

```
---

::: notes
:::
```

[1] https://r4ds.hadley.nz/functions#introduction

1.3.2 `library(printr)` currently breaks `df-print: paged`

See https://github.com/yihui/printr/issues/41

1.4 LaTeX

- don't use `align*` or `align*` in quarto; only `aligned`

Double superscript issues: https://www.overleaf.com/learn/latex/Errors/Double_superscript