

Social Analysis and Simulation in R

R Markdown Tutorial (O-Week Lab Session)

Yuki Atsusaka

August 5, 2019

1. R Markdown

1. R Markdown has been developed in order to enhance:
 - communication between analysts and audience
 - collaboration among researchers
 - research capabilities for data scientists
2. Using R Markdown, you can directly *combine* a document style explanation with R codes that you used to produce the results. In other words, you do not have to **copy and paste** text and graphs to make the final document.
3. You can output your document in various styles including **html**, **pdf** (you need a TeX environment), **docx** (you don't need Office), and others.
4. For reference, see *R Markdown: The Definitive Guide* (2019) by Yihui Xie, J. J. Allaire, Garrett Grolmund (<https://bookdown.org/yihui/rmarkdown/>).

2. Setting Up

In order to make well-reproducible documents in R Markdown, you need to install two programs/packages: R Markdown and TeX. (TeX is only required for pdf documentation.) First thing first, install the R Markdown **R** package via:

```
install.packages("rmarkdown")
```

If you have not installed any distribution of TeX, you can install tinytex that is a convenient TeX distribution recommended for R Markdown.

```
install.packages("tinytex")
tinytex::install_tinytex() # install TinyTeX
```

If you already have some TeX distribution installed, you can skip this part. Note that you can do the above process in **R**.

Now, if you finish installation, open up "SocialAnalysisSimulationR_RMarkdown.rmd" that is a source code for this pdf document. Then, I wanted to make sure that you can do the following.

1. Reproduce this pdf file by clicking Knit button on the tab or Ctrl + Shift + K (Windows OS) / Cmd + Shift + K (macOS).
2. Let me know (atsusaka@rice.edu) whether you successfully reproduced the document.

3. Basic Syntax

3.1. Writing and Comment Out

Inside .rmd files, you can write sentences as in text editor or microsoft word. If you want to indent, you can type `
`. If you want to comment out, you can use:

```
<!-- This is how we comment out -->
```

Other basic syntax include:

if you want to italicize words and also you can do like this

if you want to emphasize or and make words bold

superscript²

~~strikethrough because I want to emphasize that this is not a good idea~~

Finally, you can insert a web linke: link name

3.2. R Chunk

If you want to include R code, you can write like this:

```
x = 2
y = 3
x + y
```

```
## [1] 5
```

Observe that the result of the operation is automatically shown. If you want to only show the operation without showing results, you can type `eval=FALSE`:

```
x = 2
y = 3
x + y
```

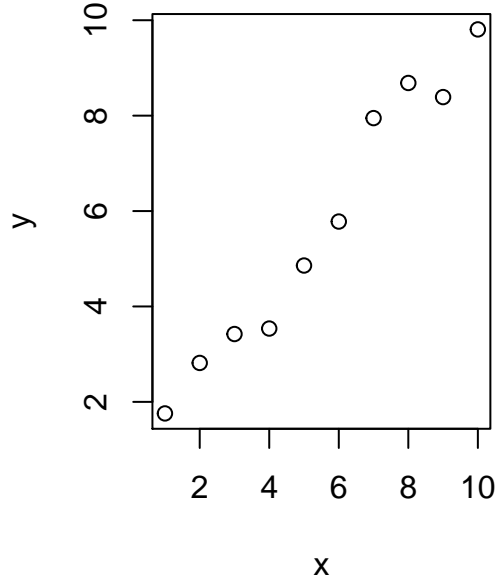
If you want to show the result only (by supressing the original code), you can type `echo=FALSE`

```
## [1] 5
```

If you feel like including R code inline, you can do like this: `x=2+y=3` (this doesn't make sense as R code, but it doesn't evoke any error).

You can also show a plot:

```
x = 1:10
y = x + rnorm(n=10)
plot(y ~ x)
```



For basic syntax for R Markdown, please see (<https://www.rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf>).

3.3. LaTeX Code

If you want to use mathematical expressions and other TeX types of expressions (e.g., greek letters), you can directly use LaTeX codes in .rmd files. For reference, please see existing tutorials such as (<http://www.malinc.se/math/latex/basiccodeen.php>) and (<http://web.ift.uib.no/Teori/KURS/WRK/TeX/symALL.html>).

4. Example: Quantitatively Predictive Logical Models (Day 3)

Based on these assumptions, we now consider the rational model of minority candidate entry. The literature of political ambition often considers the following model of candidate entry:

$$u_{it} = \hat{P}_{it}B_{it} - C_{it}, \quad (1)$$

where u_{it} is the utility that candidate i obtain from running for office at time t , \hat{P}_{it} is the candidate's estimate of the probability of winning if they enter the race, B_{it} is the benefits of holding the office, and C_{it} is the cost of running which is greater than zero, all at time t .

[...]

To formalize the idea that we just described, we introduce the following:

$$P_{it} = \mathbf{F}\left(\underbrace{\alpha_{it} \left(\frac{V_{it-1}^B - V_{it-1}^W}{2} \right)}_{\text{Last elections}} + (1 - \alpha_{it}) \underbrace{\beta_{it}}_{\text{Best guess}}\right), \quad (2)$$

where $\alpha_{it} = 4\lambda_{it}(1 - \lambda_{it}) \in [0, 1]$ is a weight parameter on the past information with $\lambda_{it} \in [0, 1]$ as the proportion of black voters and $\beta_{it} = 50(2\lambda_{it} - 1) \in [-50, 50]$ denotes an speculated marginal vote share under the specified assumptions. V_{it-1}^B and V_{it-1}^W express the vote share of the top black candidate and white candidate in the last election (i.e., at time $t - 1$) and they are known quantities.

```
rm(list=ls())
par(mfrow=c(4,3), mar = c(3, 3, 2, 2)) # Default (5,4,4,2) c(bottom, left, top, right)

for(i in 0:10){
    # We explore (0,1) in lambda

    lambda_val = i * 0.1
    SD = 17
    SD_small = 8
    SD_large = 25
    # Parameters to be varied
    # Parameters to be varied
    # Parameters to be varied
    # Parameters to be varied

    delta = seq(from=-50, to=50, by=0.01)
    alpha = 4*lambda_val*(1-lambda_val)
    beta = 50*(2*lambda_val - 1)
    q = alpha * delta + (1-alpha)*beta
    # Fixed quantities
    # Function of lambda_val
    # Function of lambda_val
    # Function of lambda_val

    probb <- pnorm(q=q, mean=0, sd=SD)
    probb2 <- pnorm(q=q, mean=0, sd=SD_small)
    probb3 <- pnorm(q=q, mean=0, sd=SD_large)
    probb4 <- pnorm(q=q, mean=-20, sd=SD_small)
    # Probability of winning
    # Probability of winning
    # Probability of winning
    # Probability of winning

    plot(probb ~ delta, type="l", lwd=2, ylim=c(0,1), ylab="",
         xlab=expression(Delta), cex.main=1.5,
         main=substitute(paste(lambda, sep=" = ", v), list(v=lambda_val)))
    lines(probb2 ~ delta, type="l", lty=2, col="firebrick", lwd=2)
    lines(probb3 ~ delta, type="l", lty=3, col="navy", lwd=2)
    lines(probb4 ~ delta, type="l", lty=4, col="forestgreen", lwd=2)
}

# We want to put a legend on this combined graph
plot(1, type = "n", axes=FALSE, xlab="", ylab="") # No plotting
legend(x = "topleft",
       legend = c("Normal(0,17)", "Normal(0,8)", "Normal(0,25)",
                  "Normal(-20,8)"),
       col=c("black", "firebrick", "navy", "forestgreen"),
       lty=c(1,2,3,4), ncol=2,
       lwd=2, cex=1.1, horiz = FALSE, text.width=0.2, box.col = "white")
```

