Class Note and Slides Using Sweave in R

Mohammed Moinuddin, PhD
m.moinuddin@ucl.ac.uk

PPP Dept, UCL Institute of Child Health

6 May 2020, ICH CodeClub

Contents

1 Introduction

Sweave is a powerful and flexible system for creating dynamic reports and reproducible research using LaTeX. Sweave enables the embedding of R code within LaTeXdocuments to generate a PDF file that includes:

- narrative, analysis and equations,
- graphics,
- code,
- and the results of computations.

Requirements: To use Sweave for creating PDF reports LaTeXneed to have installed on your system. It is free. Any package of LaTeXcan be used. The installation guide can be found in https://www.latex-project.org/get/.

2 Getting started, text and equations

Assuming you have LaTeXand RStudio installed and are fully functional, open a new R Sweave script (go to **File | New** and select "R Sweave") in your RStudio. Set your document class, paper size, font, language and other packages as the priority. You can define your choice of colour and new commands. For details please follows the LaTeXmanual. For class note type simple documents I prefer `scrartcl` which is the coma-script article class more flexible than `article` class. You can choose `article` or `report` class as well with no problem. DO NOT change other options that come default unless you are too sure. For presentation slides the class is `beamer`.

Now fix your title with other relevants entities such as authors, institution, date etc. The option `maketitle` will make the title page. The `maketitle` must be placed after `begin{documents}`. All the LaTeXcommands should be start with a `forward slash`.

Once everything is set, you can compile the Sweave document into a PDF using the **Compile PDF** button on the toolbar.

Numerical expression inside text (for example, $(\alpha + \beta)$) can be inserted between two `dollor`($\cdot$) sign. To write a seperate equation or a set of equation there are several options. You can write between double `dollor`($\$\cdot\$$) sign or use the command `begin{equation}`. An example equation is,

To test this hypotheis of equality of two normal percentiles one of the suitable tests is Behrens-Fisher problem. The test statistic for the above-mentioned hypotheis is,

$$U = \frac{(\bar{X}_1 + z_{1-p}S_1) - (\bar{X}_2 + z_{1-p}S_2)}{\sqrt{\left(1 + \frac{1}{2}z_{1-p}^2\right)\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)}} \tag{1}$$

Under the null hypothesis the test statistic U follows a Student's t distribution with degress of freedom,

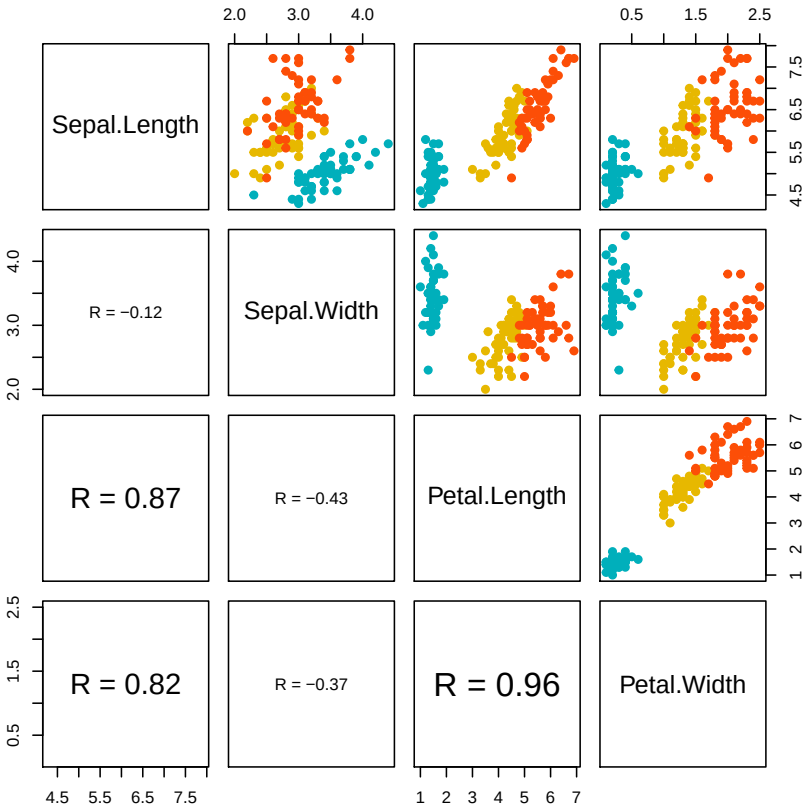
$$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\left(\frac{S_1^4}{n_1^2(n_1-1)} + \frac{S_2^4}{n_2^2(n_2-1)}\right)} \tag{2}$$

The decision about the relevant hypothesis follows the standard procedure of hypothesis.

3 R Code, graphs and output

The R-code in Sweave is called **chunk**. To insert an R chunk, use the `Chunks` menu at the top right of the source editor(**green** colored). Put the cursor at your desired point and clink the **chunk** button. To manage what you would like to see in the output file set the options *echo, fig, message, warning* etc.

```
> my_cols <- c("#00AFBB", "#E7B800", "#FC4E07")
> # Correlation panel
> panel.cor <- function(x, y){
+     usr <- par("usr"); on.exit(par(usr))
+     par(usr = c(0, 1, 0, 1))
+     r <- round(cor(x, y), digits=2)
+     txt <- paste0("R = ", r)
+     cex.cor <- 0.8/strwidth(txt)
+     text(0.5, 0.5, txt, cex = cex.cor * r)
+ }
> # Customize upper panel
> upper.panel<-function(x, y){
+   points(x,y, pch = 19, col = my_cols[iris$Species])
+ }
> # Create the plots
> pairs(iris[,1:4],
+       lower.panel = panel.cor,
+       upper.panel = upper.panel)
```



The figures can also be inserted from external source. An example,

Thanks for not doing this

The correlation and regression analysis of these data set are as below:

```
> #Correlation
> cor(iris[,1:4])
```

```
             Sepal.Length Sepal.Width Petal.Length Petal.Width
Sepal.Length    1.0000000  -0.1175698    0.8717538   0.8179411
Sepal.Width    -0.1175698   1.0000000   -0.4284401  -0.3661259
```

```
Petal.Length    0.8717538  -0.4284401    1.0000000   0.9628654
Petal.Width     0.8179411  -0.3661259    0.9628654   1.0000000
```

```
> #Regression
> m1 <- lm(Sepal.Length~Species, data = iris, family='gaussian')
> summary(m1)

Call:
lm(formula = Sepal.Length ~ Species, data = iris, family = "gaussian")

Residuals:
    Min      1Q  Median      3Q     Max
-1.6880 -0.3285 -0.0060  0.3120  1.3120

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)         5.0060     0.0728  68.762  < 2e-16 ***
Speciesversicolor   0.9300     0.1030   9.033 8.77e-16 ***
Speciesvirginica    1.5820     0.1030  15.366  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.5148 on 147 degrees of freedom
Multiple R-squared:  0.6187,        Adjusted R-squared:  0.6135
F-statistic: 119.3 on 2 and 147 DF,  p-value: < 2.2e-16
```

This is a very brief introduction of Sweave. To be proficient please visit the details LaTeXuser manual.