International Journal of Soft Computing 12 (5-6): 318-321, 2017
ISSN: 1816-9503

Using Sweave and Random Number Generation to Create a Multitude of Statistical Examples for the Enhancement of Teaching Materials

Arthur L. Dryver
Graduate School of Business Administration, NIDA,
Seri Thai Road 118, Bangkapi, 10240 Bangkok, Thailand

Abstract: The function Sweave allows an R user to create LATEX code within R. Most studies discussing Sweave cover how to use Sweave for updating LATEX code within R for a single example. There is a R package called exams which allows teachers to create unique exams for students. There has been little research however, on how to leverage the function Sweave and other basic commands in R in order to easily offer students a virtually unlimited supply of statistical example problems to study from within a single PDF file. This study covers how to create a multitude of examples for each problem type covered in a statistical class.

Key words: Package, multitude, statistical, unlimited, PDF file, problems

INTRODUCTION

In theory, random number generation can be used to create a nearly infinite number of statistical examples. It is possible to do this using the typesetting system LATEX (Lamport, 1994; Mittelbach *et al*., 2004), combined with object-based scripting language Javascript (Flanagan, 2002). Unfortunately, the latter may be difficult to do especially for the more advanced statistical problems (Dryver, 2009). Using the R environment (Team, 1999) and a R package exams using Sweave can enable teachers to create unique problems for students (Grun and Zeileis, 2009; Leisch, 2002, 2003). The study Grun and Zeileis (2009) discusses this for exams and self study material. The R package exams will create multiple files for multiple problems of the same type. Unfortunately, it is not convenient using exams to create multiple examples of the exact same type within a single file. The following example code will yield 5 files each with 3 example problems in the file.

Exams(c("tstat.Rnw", "tstat.Rnw", "tstat.Rnw"), n = 5, dir = odir, template = c("plain")): The file 'tstat.Rnw' is an example file given by Grun and Zeileis (2009) and one can refer to Grun and Zeileis (2009) for further details. Unfortunately using the command exams as above produces 5 files with each of the 3 problems within a single file having all of the same numbers. The easiest way to fix this issue would be to create three files such as 'tstatA.Rnw', 'tstatB.Rnw' and 'tstatC.Rnw' all with the same content. For the latter reason this study introduces computer code to enable the create of a single file with multiple problem-solution sets all with different numbers. Not all students require the same amount of attention for each item to be covered in a class (Guskey, 2007). Ideally an infinite number of examples for students to learn from would be readily accessible. The examples, say for a t-test would offer multitudes of scenarios such as varying the alternative hypotheses, the sample sizes and the numbers. Unfortunately, it is very very tedious to create a multitude of truly different scenarios. Create a few scenarios are reasonable but tons of completely different examples would be very time consuming. Luckily, using random number generation and Sweave, it is possible to at least create numerous problems with different numbers. One can even use random number generation to randomly determine alternative hypothesis. This study covers how to leverage LATEX, Sweave, Story (1999) exerquiz in order to enhance instructional material by creating multiple examples within a single file. Using LATEX and Sweave alone can be used to create multiple examples within a single file. The exerquiz offers teachers a lot of flexibility in the creation of documents for class. The following is an excerpt from 1999 which highlights some of the possibilities of using the package exerquiz for teachers which combined with random number generation and Sweave is much more powerful. The exerquiz package provides environments for creating the following interactive elements in a PDF file:

- The exercise environment; macros for creating on-line exercises
- The shortquiz environment; Macros for creating interactive quizzes with immediate feedback
- Shortquiz with solutions; macros for creating quizzes with immediate feedback and a link to the solutions to the quizzes

- The quiz environment: macros for creating quizzes graded by JavaScript with an option to have the quizzes corrected using JavaScript

In each of the quiz environments, you can pose multiple choice, math fill-in or text fill-in (Story, 1999).

MATERIALS AND METHODS

In addition, the exerquiz package allows teachers the option of where to place the solutions in a PDF file, immediately after the problems or in a section at the end of a PDF file or even to not include solutions. These options on managing solutions can be done easily with a single command line. For example, for creating a student file without solutions, one can use the command\ usepackage[nosolutions]{exerquiz}. It is possible for placement of solutions in various locations using the commands\SolutionsAfter and\SolutionsAtEnd. Also, possible to "hide" some solutions in a file and to create a file with no solutions hidden. These options are for creating student's and teacher's versions of a textbook. For example, a teacher's version could contain all of the solutions while the student's version would contain only answers to selected problems.

A complete example: This study contains a complete example of how to create an almost infinite number of problem-solution sets within R leveraging LATEX, Sweave and exerquiz. In this example, the problem is to solve for the sample minimum, maximum, mean, median, variance and standard deviation of a randomly generated set of numbers.

Sweave file: In this subsection, the researcher will only cover the nonstandard LATEX and R commands. The code <<···>> = indicates that the following lines will be an R code. The @ symbol indicates that what will follow is LATEX documentation. These two codes inform Sweave when to switch between R and LATEX documentation. The command <<echo = FALSE>>= suppresses the R commands from being output into the LATEX file ultimately produced. The command <<echo = FALSE, results = tex>> = suppresses the R commands from being output but allows for the LATEX code to be output. In addition, without the option results = tex the output would be assumed to be R and output accordingly. The command\Sexpr{···} within LATEX tells Sweave to execute R within the curly brackets. In the example below, the command\Sexpr{whichy} is replaced by whatever whichy the string equals, either "cholesterol" or "weight" without the quotes. The LATEX commands\begin {exercise} and \end{exercise} signal the start and end of the exercise and the LATEX commands\begin{solution} and\end{solution} signal the start and end of the solution for the exercise. These commands are defined in the exerquiz (Story, 1999). Finally, command xtable (Dahl, 2009) can produce tables for LATEX from R variables in matrix form which the researcher uses in this example. The Sweave 'dstats.Rnw' file is given:

Algorithm 1 (The Sweave file):

```
<<echo = FALSE>>=
n = rpois(1, 5)+2
n = max(c(n, 6))
whaty = rbinom(1, 1, 0.5)
if (whaty == 0){y = rnorm(n, 180, 15)
whichy = "cholesterol"}else{
y = rnorm(n, 70, 5)
whichy = "weight"}
y = round(y, 1)
theans = rbind(min(y), max(y), mean(y), median(y)
var(y), sd(y))
rownames (theans) = c("min" =, "max " =, "mean" =
"median" =, "variance" =, "std deviation" =)
@
\begin{exercise}
Solve for the sample min, max, mean, median
variance and standard deviation of
men's\Sexpr{whichy} given
the data:
<<echo = FALSE, results = tex>>=
xt = xtable(t(as.matrix(y)))
print(xt, floating = FALSE, include.rownames = FALSE
include.colnames = FALSE)
@
\begin{solution}\\
<<echo = FALSE,results = tex>>=
xt2 = xtable (theans)
print(xt2, floating = FALSE, include.rownames = TRUE
include.colnames = FALSE)
@
\end{solution}
\end{exercise}
LaTeX file
The master LATEX file, called 'testR.tex' is give as
\documentclass{article}
\usepackage{exerquiz, longtable}
\begin{document}
\SolutionsAfter
\input{exercises}
\end{document}
```

RESULTS AND DISCUSSION

Core R code: The following code is an example of how to write an R code that allows one to create multiple examples from a single or multiple*.'Rnw' file(s) and output to a single '.tex' file. In this example, we use a single '*.Rnw' file; however, the code is written in a manner such that multiple files could be used. First the programmer should set a working directory, setwd. Loading the xtable package (Dahl, 2009) for use of the command xtable will enable easy production of tables for LATEX from R. The variable the files would contain a list of all of the Sweave files to use and from which examples can be made in this example, it contains a single file 'dstats.Rnw'. The examples to be produced will be output

to a file 'exercises.tex'. This file should be named something that does not already exist. First, the file 'exercises.tex' is removed and it is assumed, the user is updating this file as often happens when first produced, as mistakes are made along the way. Next, we create a file 'exercises.tex' using a nested for loop. The first loop is for the Sweave files and the second loop is for the number of repetitions desired for each example type; here reps = 3. Then, the command Sweave(thefiles[j], output = "temp.tex") is used to create a temporary LATEX file from the Sweave file. Next, we append 'temp.tex' to 'exercises.tex'. Finally, 'temp.tex' is removed and this process is repeated reps×length(thefiles) which is 3 times in this example:

Algorithm (Core R code):

```
wdir = "the desired directory"
setwd(wdir)
library(tools)
library(xtable)
thefiles = c("dstats.Rnw")
reps = 3
file.remove ("exercises.tex")
for (j in 1:length(thefiles)){
for (jj in 1:reps){
Sweave(thefiles[j],output = "temp.tex")
file.append("exercises.tex", "temp.tex")
file.remove("temp.tex")
} }
```

The library tools enables the use of texi2dvi. The texi2dvi calls a LATEX compiler from within R in order to:

Algorithm 3 (LATEX compiler):

```
produce a file
library(tools)
texi2dvi(file = "testR.tex", pdf = TRUE,
clean = TRUE, index = FALSE)
```

File produce

The output created for within R

Exercise 1: Solve for the sample min, mean, median, variance and stan-dard deviation of mean's cholesterol given the data as 181.30, 201.10, 191.40, 173.60, 185.40, 171.60:

Solution:

- Min. = 171.60
- Max. = 201.10
- Mean = 184.07
- Median = 183.35
- Variance = 123.66
- SD = 11.12

Exercise 2: Solve for the sample min, mean, median, variance and stan-dard deviation of mean's weight given the data as 65.90, 73.80, 74.30, 77.70, 63.80, 66.40, 67.50, 72.30, 75.40:

Solution:

- Min. = 63.80
- Max. = 77.70
- Mean = 70.79
- Median = 72.30
- Variance = 24.44
- SD = 4.94

Exercise 3: Solve for the sample min, mean, median, variance and standard deviation of mean's weight given the data as 68.00, 69.20, 75.70, 57.50, 70.40, 73.60:

Solution:

- Min. = 57.50
- Max. = 75.70
- Mean = 69.07
- Median = 69.80
- Variance = 40.25
- SD = 6.34

CONCLUSION

Possibilities and summary: Combining LATEX, Sweave and exerquiz can create an almost infinite number of problem-solutions sets. A teacher can easily create one file with solutions for him or herself and another without solutions for students. With minimal additional effort using exerquiz, a teacher can produce a file with problem solution sets as examples and questions without solutions for homework exercises. The R code presented in this study is just a few lines long and thus it is very easy for one to use in order to accomplish the latter for anyone already using Sweave to produce examples for class. Using random-number generation can go beyond just creating multiple examples of the same problem which is not the best method for teaching statistics, rote learning (Hulsizer and Woolf, 2008). Random number generation can be used to easily modify problems in various manners. For example:

- An odd and even number of data points for calculating the median
- A greater than to a less than alternative hypothesis
- From a fail to reject null hypothesis to a reject null hypothesis

All the way to story modification such as:

- Men's cholesterol to their weight
- A company's revenue numbers to net profit
- Exam scores to credit scores, etc.
- Others
- etc.

Although, for minor story modification random-number generation is fine for major story modification the programmer may prefer to simply write another 'Sweave' file. In conclusion, those unfamiliar and uninterested in Sweave, thinking that it is mainly for updating examples (Leisch, 2002), should reconsider learning Sweave given that it is much more than just that. In addition, as electronic books (e-Books) have become more popular (Nelson, 2008), having a plethora of examples does not have to end with heavy textbooks but merely a bigger computer file. Finally, given the extensive use of e-Books not only is it possible to create large numbers examples for students and teachers and it is practical to distribute and consume the material.

REFERENCES

Dahl, D.B., 2009. Xtable: Export tables to LaTeX or HTML. R. Package, 1: 1-5.

Dryver, A., 2009. The enhancement of teaching materials for applied statistics courses by combining random number generation and portable document format files via LaTEX. J. Stat. Software, 31: 1-9.

Flanagan, D., 2002. JavaScript: The Definitive Guide. 4th Edn., O'Reilly, Boston, Massachusetts, ISBN:0-596-00048-0, Pages: 871.

Grun, B. and A. Zeileis, 2009. Automatic generation of exams in R. J. Stat. Software, 29: 1-14.

Guskey, T.R., 2007. Closing achievement gaps: Revisiting Benjamin S, bloom's learning for mastery. J. Adv. Acad., 19: 8-31.

Hulsizer, M.R. and L.M. Woolf, 2008. A Guide to Teaching Statistics: Innovations and Best Practices. Wiley, Hoboken, New Jersey, USA., ISBN:9781405155748, Pages: 280.

Lamport, L., 1994. LaTeX: A Document Preparation System. 2nd Edn., Addison-Wesley Professional, Reading, Massachusetts, ISBN-13: 978-0201529838.

Leisch, F., 2002. Sweave, part I: Mixing R and latex. R. News, 2: 28-31.

Leisch, F., 2003. Sweave, part II: Package vignettes. R. News, 3: 21-24.

Mittelbach, F., M. Goossens, J. Braams, D. Carlisle and C. Rowley, 2004. The LATEX Companion. 2nd Edn., Addison-Wesley, Boston, Massachusetts, ISBN:0-201-36299-6, Pages: 1089.

Nelson, M.R., 2008. E-Books in higher education: Nearing the end of the era of hype?. Educause Rev., 43: 40-56.

Story, D.P., 1999. The AcroTEX eDucation bundle. Master Thesis, AcroT E X Software Development Team, AcroTEX Ltd., South East England.

Team, R.C., 1999. Writing R Extensions. R Core Team, New Zealand, Pages: 185.