Automatic Generation of Simple (Statistical) Exams

Bettina Grün, Achim Zeileis

`http://statmath.wu-wien.ac.at/`

Overview

- Introduction
  - Challenges
  - Solution implemented in the R package **exams**
- Exercises
- Combining exercises: The master LaTeX file
- Application and customization: Function `exams()`
- Discussion

Introduction

Re-design of introductory statistics lecture at WU Wien:

- The course is attended each semester by 1,000–1,500 students (mostly first-year business students).
- Several lecturers from the Department of Statistics and Mathematics teach this course in parallel.
- All teaching materials are covered by the re-design: presentation slides, collections of exercises, exams, etc.
- The re-design was accomplished through a collaborative effort of all concerned faculty members working in small teams on different chapters.

Introduction / 2

Main challenges:

- *Scalable exams:* Automatic generation of a large number of different exams.
- *Associated self-study materials:* Collections of exercises and solutions from the same pool of examples.
- *Joint development:* Development and maintenance of a large pool of exercises in a multi-author and cross-platform setting.

Tools chosen:

- R (R Development Core Team 2008) and LaTeX(Knuth 1984; Lamport 1994) $\Rightarrow$ Sweave (Leisch 2002)
- Subversion (SVN, Pilato, Collins-Sussman, and Fitzpatrick 2004)

Introduction / 3

Design principles of package exams:

- *Maintenance:* Each exercise template is a single file (also just called “exercise”).
- *Variation:* Exercises are dynamic documents, containing a problem/solution along with a data-generating process (DGP) so that random samples can be drawn easily.
- *Correction:* Solutions for exercises are either multiple-choice answers (logical vectors) or numeric values (e.g., a test statistic or a confidence interval).

Exercises

Each exercise typically represents an exemplary application of a statistical procedure.

The exercise file consists of (at least):

- **Two environments:** a question and a solution description encapsulated in corresponding LaTeXenvironments.
- **Meta-information:** about type of questions (e.g. multiple-choice or numeric), the solution, a descriptive name and the allowed tolerance for numeric solutions.

An exercise file can be processed in R by:

```
R> set.seed(1090)

R> library("exams")
R> tstat_sol <- exams("tstat.Rnw")
R> tstat_sol

plain1
    1. t statistic: 15.958 (15.948--15.968)
```

A simple Sweave exercise: `tstat.Rnw`

```
<<echo=FALSE, results=hide>>=
## DATA GENERATION
n <- sample(120:250, 1)
mu <- sample(c(125, 200, 250, 500, 1000), 1)
y <- rnorm(n, mean = mu * runif(1, min = 0.9, max = 1.1),
           sd = mu * runif(1, min = 0.02, max = 0.06))
## QUESTION/ANSWER GENERATION
Mean <- round(mean(y), digits = 1)
Var <- round(var(y), digits = 2)
tstat <- round((Mean - mu)/sqrt(Var/n), digits = 3)
@
\begin{question}
  A machine fills milk into $\Sexpr{mu}$ml packages. It is suspected that the
  ...
\end{question}
\begin{solution}
  ...
\end{solution}

%% META-INFORMATION
%% \extype{num}
%% \exsolution{\Sexpr{format(abs(tstat), nsmall = 3)}}
%% \exname{t statistic}
%% \extol{0.01}
```

LaTeX output of `Sweave("tstat.Rnw")`

```
\begin{question}
  A machine fills milk into $500$ml packages. It is suspected that the
  machine is not working correctly and that the amount of milk filled differs
  from the setpoint $\mu_0 = 500$. A sample of $226$ packages filled by the
  machine are collected. The sample mean $\bar{y}$ is equal to $517.2$ and
  the sample variance $s^2_{n-1}$ is equal to $262.56$.

  Test the hypothesis that the amount filled corresponds on average to the
  setpoint. What is the absolute value of the $t$~test statistic?
\end{question}
\begin{solution}
  The $t$~test statistic is calculated by:
  \begin{eqnarray*}
    t & = & \frac{\bar y - \mu_0}{\sqrt{\frac{s^2_{n-1}}{n}}}
    = \frac{517.2 - 500}{\sqrt{\frac{262.56}{226}}} = 15.958.
  \end{eqnarray*}
  The absolute value of the $t$~test statistic is thus equal to $15.958$.
\end{solution}

%% META-INFORMATION
%% \extype{num}
%% \exsolution{15.958}
%% \exname{t statistic}
%% \extol{0.01}
```

Display of processed `tstat` exercise

- **Problem**
  A machine fills milk into 500ml packages. It is suspected that the machine is not working correctly and that the amount of milk filled differs from the setpoint $\mu_0 = 500$. A sample of 226 packages filled by the machine are collected. The sample mean $\bar{y}$ is equal to 517.2 and the sample variance s^2_{n-1} is equal to 262.56.

  Test the hypothesis that the amount filled corresponds on average to the setpoint. What is the absolute value of the t test statistic?

  Solution
  The t test statistic is calculated by:

$$t = \frac{\bar{y} - \mu_0}{\sqrt{\frac{s^2_{n-1}}{n}}} = \frac{517.2 - 500}{\sqrt{\frac{262.56}{226}}} = 15.958.$$

  The absolute value of the t test statistic is thus equal to 15.958.

Combining exercises: The master LaTeX file

`exams()` allows for

- construction of exams with stratified sampling of exercises,
- automatic generation of multiple copies (potentially of multiple layouts) with suitable names and storage,
- inclusion of a suitable cover page with answer fields, and
- collection of meta-information for problems and solutions in an R object.

Sequence of work steps for `exams()`

1. Collect all Sweave files for the exercises, the master LaTeX file(s) and potentially additionally specified input files.
2. Copy all files to a (temporary, by default) directory.
3. Run `Sweave()` for each exercise.
4. Produce a copy of the master LaTeX file(s) in which certain control structures are substituted by dynamically generated LaTeX commands (e.g., for including the exercises).
5. Run `texi2dvi()` for each master LaTeX file.
6. Store the resulting PDF file(s) in an output directory or display it on the screen (for a single file only, by default).

A simple master LaTeX file: `plain.tex`

```
\documentclass[a4paper]{article}

\usepackage{a4wide,Sweave}
\newenvironment{question}{\item \textbf{Problem}\newline}{}
\newenvironment{solution}{\textbf{Solution}\newline}{}

\begin{document}
\begin{enumerate}
%% \exinput{exercises}
\end{enumerate}
\end{document}
```

To hide the solution the corresponding environment needs to be defined as a comment:

```
\newenvironment{solution{\comment}{\endcomment}}
```

Possible dynamic modifications

- `\exinput{exercises}`: Inclusion of exercises.
  Replaced by: `\input{`*filename*`}` (one for each exercise).
  Example: `\input{tstat}`.
- `\exinput{questionnaire}`: Inclusion of questionnaires, e.g., for cover sheets.
  Replaced by: `\exnum{...}` or `\exmchoice{...}`, respectively (one for each exercise).
  Example: `\exnum{}{}{}{}{1}{5}{9}{5}{8}`.
- `\exinput{header}`: Further commands and definitions.
  Replaced by: `\`*command*`{`*value*`}` (one for each header command).
  Example: `\Date{2009-01-16}`.

Application and customization: Function `exams()`

Function `exams()` has the following arguments:

```
exams(file, n = 1, dir = NULL,
      template = "plain", inputs = NULL,
      header = list(Date = Sys.Date()),
      name = NULL, quiet = TRUE, edir = NULL,
      tdir = NULL, control = NULL)
```

Application and customization: Function `exams()` / 2

```
R> myexam <- list("boxplots",
+    c("confint", "ttest", "tstat"),
+    c("anova", "regression"),
+    "scatterplot",
+    "relfreq")
R> odir <- tempfile()
R> getID <- function(i)
+    paste("myexam", gsub(" ", "0", format(i, width = 2)),
+          sep = "")
R> getID(1)

[1] "myexam01"
```

Application and customization: Function `exams()` / 3

```
R> set.seed(1090)
R> sol <- exams(myexam, n = 3, dir = odir,
+    template = c("exam", "solution"),
+    header = list(ID = getID, Date = Sys.Date()))
R> list.files(odir)

[1] "exam1.pdf"     "exam2.pdf"     "exam3.pdf"
[4] "metainfo.rda"  "solution1.pdf" "solution2.pdf"
[7] "solution3.pdf"

R> print(sol, 1)

exam1
    1. Multiple choice: abde
    2. t statistic: 0.188 (0.178--0.198)
    3. Prediction: 236.678 (236.668--236.688)
    4. Multiple choice: abde
    5. Multiple choice: d
```

Application and customization: Function `exams()` / 4

R University
Statistics Exam 2009-01-15 **Exam ID myexam01**

Name: _____

Student ID: _____

Signature: _____

1. (a) ☐ (b) ☐ (c) ☐ (d) ☐ (e) ☐
2. ☐☐☐☐☐☐.☐☐☐
3. ☐☐☐☐☐☐.☐☐☐
4. (a) ☐ (b) ☐ (c) ☐ (d) ☐ (e) ☐
5. (a) ☐ (b) ☐ (c) ☐ (d) ☐ (e) ☐

1. In Figure 1 the distributions of a variable given by two samples (A und B) are represented by parallel boxplots. Which of the following statements are correct? *(Comment: The statements are either about correct or clearly wrong.)*

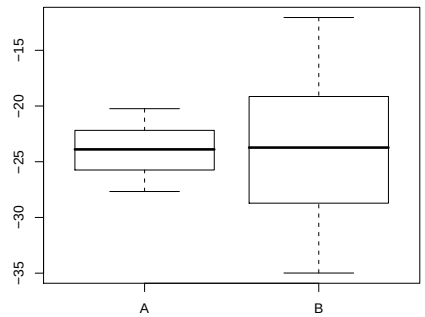

Figure 1: Parallel boxplots.

   (a) The location of both distributions is about the same.
   (b) Both distributions contain no outliers.
   (c) The spread in sample A is clearly bigger than in B.
   (d) The skewness of both samples is similar.
   (e) Distribution A is about symmetric.

2. A machine fills milk into 500ml packages. It is suspected that the machine is not working correctly and that the amount of milk filled differs from the setpoint $\mu_0 = 500$. A sample of 226 packages filled by the machine are collected. The sample mean $\bar{y}$ is equal to 499.7 and the sample variance s^2_{n-1} is equal to 576.1.

   Test the hypothesis that the amount filled corresponds on average to the setpoint. What is the absolute value of the t test statistic?

3. For 49 firms the number of employees X and the amount of expenses for continuing education Y (in EUR) were recorded. The statistical summary of the data set is given by:

	Variable X	Variable Y
Mean	58	232
Variance	124	1606

   The correlation between X and Y is equal to 0.65.

   Estimate the expected amount of money spent for continuing education by a firm with 60 employees using least squares regression.

4. Figure 2 shows a scatterplot. Which of the following statements are correct?

Application and customization: Function `exams()` / 5

Several arguments allow for a fine control, e.g. to modify the print output:

```
R> mycontrol <- list(mchoice.print =
+    list(True = LETTERS[1:5], False = "_"))
R> (exams(myexam, n = 1, template = "exam",
+    control = mycontrol))

exam1
    1. Multiple choice: _B_D_
    2. Multiple choice: AB_D_
    3. Prediction: 208.13 (208.12--208.14)
    4. Multiple choice: __C__
    5. Multiple choice: _____
```

Discussion

- Package **exams** provides a framework for automatic generation of simple (statistical) exams and associated self-study materials.
- It is based on independent exercises in Sweave format which can be compiled in exams (or other collections of exercises) by providing one (or more) master LaTeX template(s).
- Contributing to the pool of exercises only requires knowledge of Sweave and minimal markup for meta-information.
- Since Spring 2008, **exams** is used at the WU Wien for generating collections of exercises, trial exams, exams and solutions.

Package **exams** is available from the Comprehensive R Archive Network at

`http://CRAN.R-project.org/package=exams`

References

Grün B, Zeileis A (2008). "Automatic Generation of Simple (Statistical) Exams." *Report 77*, Department of Statistics and Mathematics, WU Wien, Research Report Series. `http://epub.wu-wien.ac.at/dyn/openURL?id=oai:epub.wu-wien.ac.at:epub-wu-01_e1d`.

Knuth DE (1984). *The TEXbook*, volume A of *Computers and Typesetting*. Addison-Wesley, Reading, Massachusetts.

Lamport L (1994). *LATEX: A Document Preparation System*. Addison-Wesley, Reading, Massachusetts, 2nd edition.

Leisch F (2002). "Dynamic Generation of Statistical Reports Using Literate Data Analysis." In W Härdle, B Rönz (eds.), "COMPSTAT 2002 – Proceedings in Computational Statistics," pp. 575–580. Physica Verlag, Heidelberg.

Pilato CM, Collins-Sussman B, Fitzpatrick BW (2004). *Version Control with **Subversion***. O'Reilly. Full book available online at `http://svnbook.red-bean.com/`.

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, `http://www.R-project.org/`.