Teaching Statistics: An example of "How to" improve the students' statistical skills using individualized assignments

Joan Vila[1,2,3*], Montse Rue[4], Nuria Codern[2], Albert Sorribas,[4] Cristina Rodríguez[2], Anna Foraster[2]

1. Institut Municipal d'Investigació Mèdica (IMIM), Barcelona, Spain
2. Escola Universitària d'Infermeria Creu Roja, Terrassa, Spain
3. CIBER Salud Pública, Barcelona, Spain
4. Universitat de Lleida, Lleida, Spain

*Contact author: jvila@imim.es

Keywords: Teaching Materials, Software Design, LaTeX , Sweave.

Sweave is a well known tool that makes it possible to embed R code in LaTeX documents[1]. When the code is run and the resulting output, figures and tables are automatically inserted into the final document.

We present the results of using Sweave for designing individualized assignments for undergraduate and graduate students of:

- Human Nutrition (HN). University of LLeida, Spain.
- Medicine (ME). University of LLeida, Spain.
- Nursing (NU). Autonomous University of Barcelona, Spain.
- Occupational Therapy (OT). Autonomous University of Barcelona, Spain.

in two different universities:

For each course we chose a clinical trial (CT). The main results of the CT were simulated to create a dataset for each student: 50 HN, 120 ME, 176 NU and 79 OT. Data were *similar* but different for each student in the course. Each student received a spreadsheet with his/her particular data and a PDF with his/her name and id. The assignment consisted of 50 exercises that covered most of the concepts taught during the course.

The exercises and the methods needed to solve them were exactly the same in each discipline, but answers could be the same, similar or absolutely different depending on each dataset.

The exercices consisted of open questions that required calculating a figure (i.e. a Student-t statistic) as well as multiple choice questions for which the answers were randomly ordered. So, not only could answers be different for different students, but even if the answers were the same, the choice could be a different item.

The students provide their answers in a free and open-source e-learning software platform (Moodle or Sakai, depending on the university).

By the assignment deadline, each student received a detailed and personalized answer, showing the correct and detailed explanation of how to solve the problem. In the multiple choice questions we explained why the remaining answers were not correct.

At the end of the semester the students had an exam. We compared the exam grades after applying this new system with historic exams performed two years before with the same teachers and similar students. Results showed dramatic improvement. For example, in the NU program, 52.7% of 186 students in the academic year 2007-08 and 70.6% of the 177 in 2009-10 passed the final exam (p-value <0.001). Furthemore, although the percentage of students who took the exam was similar, 146 (78.5%) in 2007-08 versus 136 (76.8%) in 2009-10, the students that took the exam received higher grades: $\bar{X}$ (SD): 5.99 (2.12) out of 10 versus 7.43 (1.54) (p-value < 0.001) and the percentage of students that passed the exam was higher: 67.1% versus 91.9% (p-value < 0.001).

Students were surveyed about about strengths and weaknesses of the teaching method. The results of this analysis, obtained using open questions of the type "positive / negative" with qualitative research methodology, provided descriptions and interpretations of the experience of the student in learning statistics on an innovative context. It also identified some room for improvement in future courses.

In conclusion, we present instruction on **how to** perform:

- Individualized assignments
- Simulated data from a Clinical Trial
- Open or multiple random choice questions
- Individualized, detailed, explanatory answers
- Embedding  code in LaTeX documents with `Sweave`

as well as the main quantitative and qualitative results of applying the new method.

References

1. Leisch F (2003). Sweave and Beyond:: Computations on text documents. In Proceedings of the 3rd International Workshop on Distributed Statistical Computing, Vienna, Austria, 2003. `http://www.ci.tuwien.ac.at/Conferences/DSC-2003/Proceedings/Leisch.pdf/`.