

DEVELOPING AN INTERACTIVE LEARNING ENVIRONMENT FOR STATISTICS

Tanja Van Hecke, Luc Decombel
Department of Applied Engineering Sciences
Department of Biosciences and Landscape Architecture
University College Ghent

Abstract

Each year our master students struggled with the same problem: how to draw conclusions from the data related to their thesis in a scientifically correct way? This was the inspiration for our project where we wanted to construct an interactive learning environment for statistics to support the learning process of our students concerning data analysis. A classical course 'statistics' is part of the curriculum of the master in biosciences. However, it goes no further than the notion of hypothesis testing and is just an introduction to data analysis. Colleagues were consulted to inform us about the types of problems which arise in the field of their specialization (biosciences) and where statistics can play a supporting role.

Keywords: Statistics, ILO, Data analysis

An electronic learning environment starting from a library of cases

It was our aim to offer our students a tool which could teach them independently data analysis in a differentiated way. From the conviction that students can learn a lot from elaborated examples, we included a full data analysis of several cases from the field of biosciences and we hoped this would facilitate the application of the statistical knowledge to similar statistical problems. This forms the starting point of the process of learning data analysis.

Determination key

An important guidance tool of the learning environment is the determination key. By means of some basic questions, related to the student's dataset, a suitable statistical technique is offered. For every statistical technique, cases are elaborated in a detailed way. The theses of past years inspired us when choosing subjects for the different cases. The student determines the level of detail he wants to see concerning this technique.. Those who need and desire arithmetical and statistical details can reach them by several links. However, some students can be overwhelmed by this information. Therefore only the most relevant information is visible and more details can be reached by linked pages.

The cases serve as examples. We are convinced that understanding the cases is not enough to declare that the students have achieved the competence of being able to apply the statistical method appropriately. It is necessary that the student proves he can apply the statistical method to a similar dataset. That's why a set of exercises is available. There are two options: either they choose an exercise related to a well-named statistical method (for example linear regression, one-way anova, ...) or they choose randomly an exercise. The latter case supposes a student is yet able to identify the type of problem. This is of major importance in data analysis. We are aware of the need of the skill to discover the statistical method which is behind the data of an engineering problem. This is the reason we developed the determination key. Here the students

can see an overview of the different types of the most important statistical methods. It is a guide full of links but clear enough to discover the differences and resemblances of the statistical methods.

Determinatie sleutel

vergelijken van behandelingen groeperen van gegevens formuleren van verbanden kwantiteits controle

1 The main subjects I. Vergelijken van methoden, technieken of (be)ha

a. Vergelijken van twee gemiddelden 2. A clear and structured overview. The title hides a link to the related case. b. vergelijken van twee

=t-test =gepaarde t-test

behandeling	
A	B
observaties na behandeling A	observaties na behandeling B

3. Each test contains a link to a movie illustrating the use of a software program (SAS)

SASinput

c. Vergelijken van meerdere gemiddelden van één type behandeling d. Vergelijken van me

= One-way Anova =Two-way Anova (ge

Figure 1. Determination key

Incorporation of statistical software

Software is a very practical tool as a calculator, but drawing the right conclusions is still a human matter which should be trained. Therefore an extensive explanation of which conclusions can be drawn out of the data is added. In this part we also stress the importance of the conditions that have to be fulfilled to apply the test in a justified way.

A movie shows how to execute a statistical method applicable in the concerned case, in a statistical software program (SAS Learning Edition). The output of statistical software programs is rarely well documented: it only consists of figures and is sparing of comment. In addition we have chosen for a full explanation of the output in favour of a deeper understanding of the reason why a hypothesis is accepted or rejected.

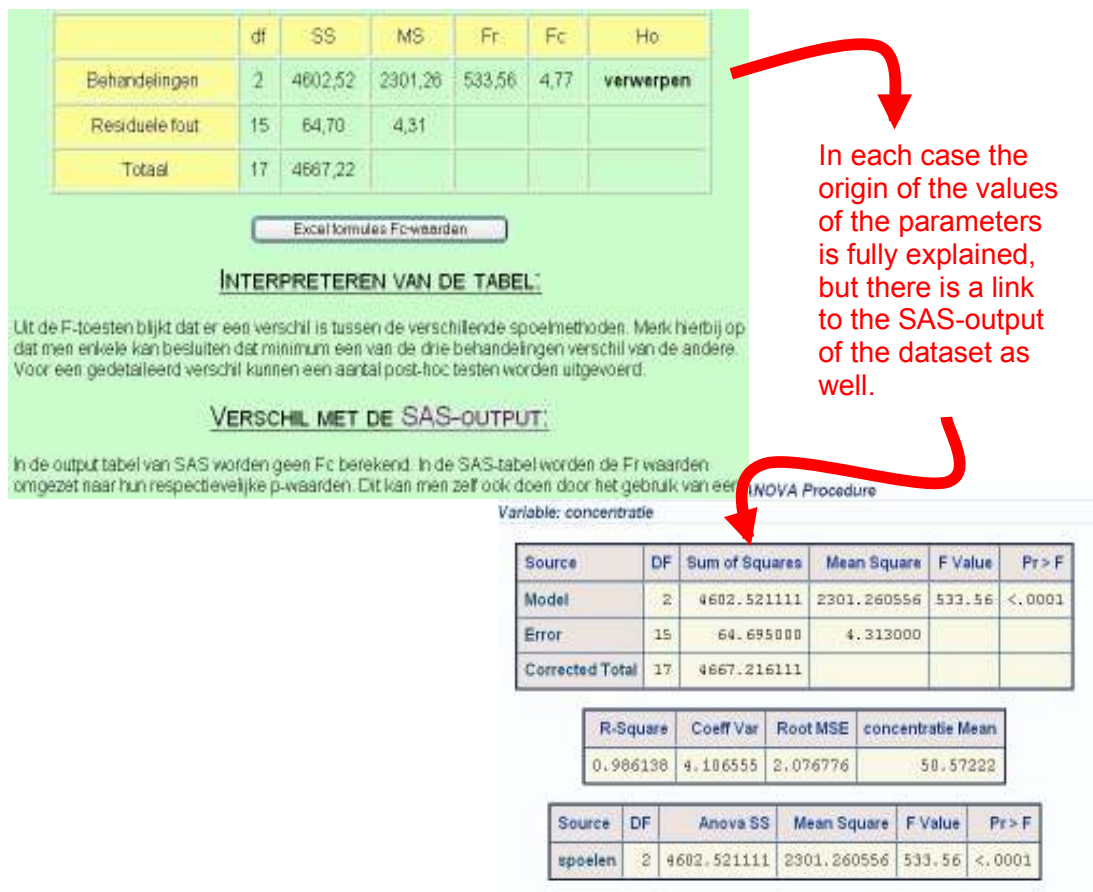


Figure 2. Link to statistical software

Role of the theory

The starting point of this concept of the learning process of data analysis is not some theoretical explanations as we wanted less emphasis on passive transmission of information. This does not mean we undervalue the mathematical theory. The students we train are engineering students and will obtain a master degree. So they should be able to get to the bottom of a theoretical exposition. Nevertheless, it could be discouraging to use the theory as a starting point. Therefore, practical cases were used to attract the interest of the students. The mathematical theory is offered when it is needed. It makes it easier to shape the theory by means of numbers related to the case. By links to excel-files the student can follow step by step the details of the calculations and formulas, which makes the theory livelier.

Dit is een tweezijdige t-toets. Een eenzijdige t-toets wordt enkel gebruikt wanneer weet dat het ene gemiddelde zeker niet kleiner (of groter) kan zijn dan het andere. Gegevens kan men de parameters, weergegeven in de onderstaande tabel, b

Parameter	Formule	Waarde
\bar{X}_1	$\frac{1}{n} \sum_{i=1}^n X_i$	2477,7
\bar{X}_2	$\frac{1}{n} \sum_{i=1}^n X_i$	2479,0
S^2_1	$\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$	26,48
S^2_2	$\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$	22,27
S^2_p	$\frac{(n_1-1)S^2_1 + (n_2-1)S^2_2}{n_1 + n_2 - 2}$	24,47
t_r	$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	35,54

In the examples on finds all formulas and the result of the application of these formulas.

The theory is linked at an excel-file where all calculations can be followed and extra comment can be found.

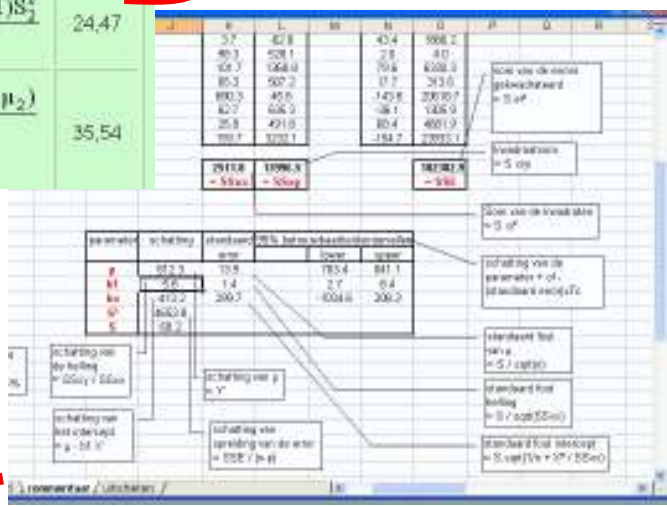


Figure 3. Position of the theory

Feedback

After answering the questions of the exercises, the student receives more than just a score. He also gets explanatory feedback on the answers. Moreover a visual presentation by means of an evaluation window is made: it displays the measure in which the student has already achieved the aimed competences concerning statistical topics. This should stimulate the student to focus on the topics where he obtained a lower score.

Evaluation

Students welcomed the developed tool and appreciated the access is flexible in time and place. It appeared the students discovered the tool independently. We experimented with the use of the electronic learning environment as an evaluation instrument. By means of logins and passwords the students could identify themselves when they brought in the answers to the questions related to the exercises. Unfortunately we could not always could on the honesty of our users who sometimes forgot to work individually.

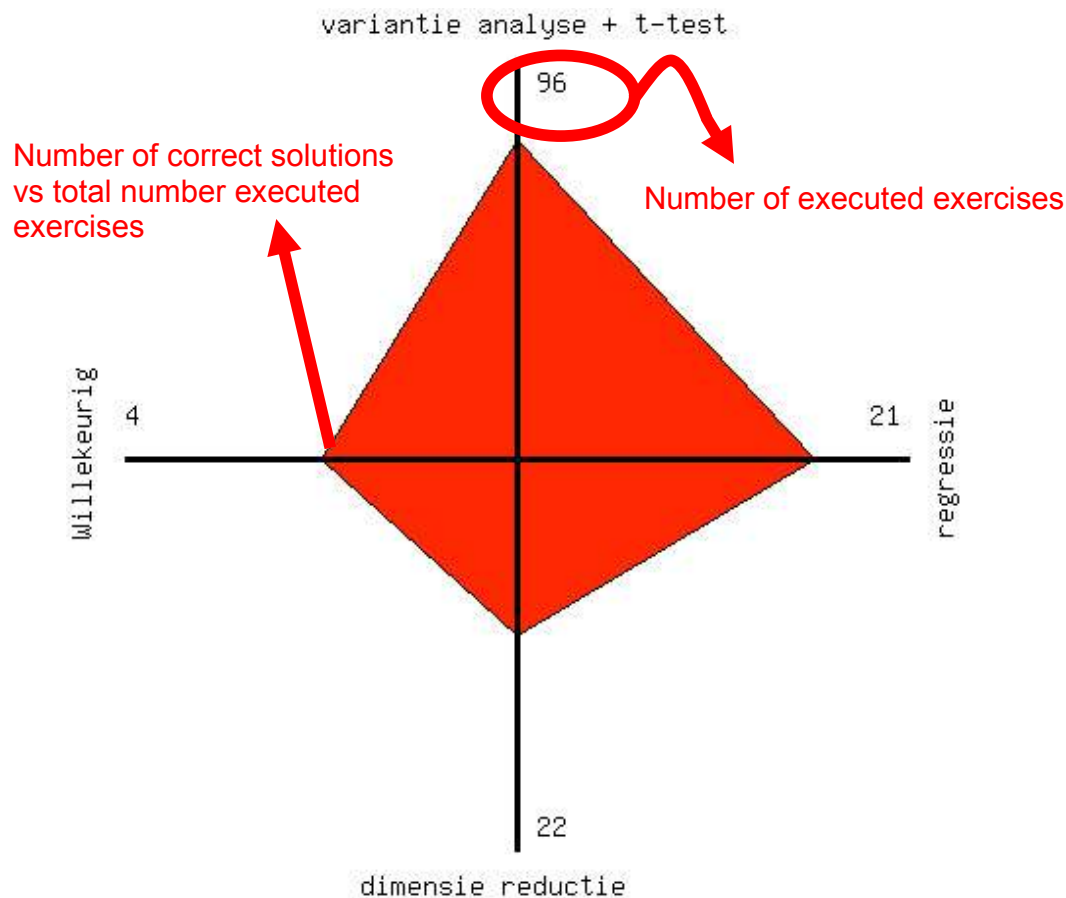


Figure 4. Evaluation window

Evolution

There is a never ending evolution in science. So it would be appropriated that our examples can follow this evolution. This is the reason why we decided in favour of a dynamical set of exercises. Teachers can add new exercises of present interest.

Biographical Information

Tanja Van Hecke (1973) obtained the master degree in applied mathematics in 1995 at the university of Ghent. In 1998 she finished here doctoral theses in numerical analysis. Since 2000 she works at the faculty of applied engineering sciences at the university college Ghent where she gives several courses in mathematics and statistics.

Luc Decombel (1977) obtained the master in bioscience engineering in 2000 at the university of Ghent. In 2004 he finished his doctoral thesis in crop protection. Since 2004, he works at the faculty of biosciences and landscape architecture at the university college Ghent in the domain of applied biological sciences. ...

Corresponding authors

Dr. Tanja Van Hecke
University College Ghent
Department of Applied Engineering Sciences
Schoonmeersstraat 52
9000 Gent
Belgium
Tel: 0032 9 2424228
Tanja.VanHecke@hogent.be

Dr.ir Luc Decombel
University College Ghent
Department of Biosciences and Landscape Architecture
Schoonmeersstraat 52
9000 Gent
Belgium
Tel: 0032 9 2104506
Luc.Decombel@hogent.be