

A Blended Learning Module in Statistics for Computer Science and Engineering Students Revisited

<https://doi.org/10.3991/ijep.v7i4.7441>

Christina Andersson, Doina Logofatu[✉]
Frankfurt University of Applied Sciences, Frankfurt/Main, Germany
logofatu@fb2.fra-uas.de

Abstract—Teaching a statistics course for undergraduate computer science students can be very challenging: As statistics teachers we are usually faced with problems ranging from a complete disinterest in the subject to lack of basic knowledge in mathematics and anxiety for failing the exam, since statistics has the reputation of having high failure rates. In our case, we additionally struggle with difficulties in the timing of the lectures as well as often occurring absence of the students due to spare-time jobs or a long traveling time to the university. This paper reveals how these issues can be addressed by the introduction of a blended learning module in statistics. In the following, we describe an e-learning development process used to implement time- and location-independent learning in statistics. The study focuses on a six-step-approach for developing the blended learning module. In addition, the teaching framework for the blended module is presented, including suggestions for increasing the interest in learning the course. Furthermore, the first experimental in-class usage, including evaluation of the students' expectations, has been completed and the outcome is discussed.

Keywords—blended learning; e-learning development process; statistics; lessons learned

1 Introduction

In the time era of data science, big amounts of data are collected and stored in immense databases, with the intention to be analyzed statistically. The demand of people with a profound knowledge in statistics is already now very large and rapidly increasing [1], [2], [3]. On the other hand, statistics as a university course for non-statistics students is usually not popular and often regarded just as something which “must be passed”. Statistics is a module in many quite different study programs at Frankfurt University of Applied Sciences (Fra-UAS): among others Industrial Engineering (B. Sc.), Social Work (B. Sc.), Business Information Systems (B. Sc.), Business Administration (B. Sc.), Computer Science (B. Sc.), High Integrity Systems (M. Sc.), etc.

Many empirical studies compare traditional face-to-face teaching with e-learning approaches, both concerning performance and students' attitudes towards the teaching

framework. However, the studies arrive at different conclusions. Evaluations of the performance show a range from worse achievements for students in online courses [4] through comparable results [5], [6], [7], [8], [9], [10], [11], [12] to significant better outcomes in the online teaching [13], [14]. The same holds for studies measuring the attitudes towards online teaching compared to traditional teaching. Also in this area we can find a range from negative attitudes towards online courses compared to traditional teaching [6], [8], [10] through similar satisfaction [12], [15] to more positive attitudes [16], [17], [18]. Despite this indeterminate picture, we decided to develop a blended learning module in statistics, since such a course addresses issues occurring in our teaching situation at FRA-UAS. The development of the blended learning module in statistics was realized as a part of the project MainCareer, supported by the initiative “Upward Mobility through Academic Training” at the German Federal Ministry of Education and Research [19].

2 Current Situation at FRA-UAS for Teaching Statistics

Teaching statistics for non-statistics students at a German University of Applied Sciences, we face different problematic issues, sometimes making the teaching a real challenge.

2.1 Reputation of Having High Failure Rate

Statistics courses are reputed with having high failure rates, i. e. the students are often scared of taking this kind of classes. For example, the statistics course for the program Computer Science is scheduled to be taken in the fourth semester. But the students have the possibility to take most exams in an arbitrary order. This means that many students, since they are afraid of the statistics course, choose to take this course much later than advised. The time schedule is very dense for the students and usually there is no possibility to make the lecture timing in statistics fit with the time schedule of the higher semesters. As a result, we noticed that when the students take the statistics course at a later time point, due to the scheduling, they cannot attend all lectures of the course.

2.2 Lack of Prior Knowledge in Mathematics

Without the knowledge of basic mathematics, it is difficult to learn statistics [20], [21], [22], [23]. The level of knowledge in mathematics, which the students bring from high school, is decreasing, i. e. we would need more time to repeat basic mathematics before we start the teaching of statistics itself. Unfortunately, there is no time to update a large amount of high school mathematics at the university.

2.3 Lack of Motivation

The students don't realize the use of the contents of statistics module in their later profession and therefore are not motivated to take it. The lecture time is limited and much time is wasted due to the lack of appropriate mathematics knowledge among the students. It is therefore difficult to find time in the lecture to arrive at real-world problems, which is expected to motivate the students.

2.4 No Hands-On with Real Data, Lack of Knowledge in Handling Statistical Software

Usually the students don't have any prior knowledge in using statistical analysis software when the course starts. The curriculum for engineering students is too tight to allow an introduction to statistical software to take place within the statistics course. On the other hand, hands-on experience in analyzing real-world data is an essential component when dealing with statistical problems [24], i. e. the absence of this component is actually a deficit of the statistics course.

2.5 General Problems

In addition to these problems, identified especially for the statistics lecture, we also face some common problems in teaching: time collisions, traveling time to the university, lectures in English without sufficient English knowledge.

3 Introducing Time- and Location-Independent Teaching in Statistics

The intention of the blended learning module in statistics is to provide the students with additional support in statistics, not to replace the conventional statistics course. Blended learning is an already well-known teaching method [25], [26]. In the Computer Science program, three groups are usually taught parallel in statistics. The students may choose the most suitable group, according to the time schedule, which depends on the other subjects they take. To introduce the blended learning course in statistics, one of the ordinary groups was replaced by the blended learning course. In this way, the students still have the possibility to choose if they want to join one of the two conventional statistics groups or if they join the blended learning course. In fact, the students also have the possibility to join both the conventional and the blended learning course. Already at the beginning, it was clearly announced that the exam would be identical for all students taking statistics, independent of teaching method.

4 E-Learning Development Process

The blended learning module in statistics consists of two main parts: A theoretical part and a practical part. The content of the theoretical part covers the theoretical background in basic statistics (descriptive statistics, basic probability calculations, statistical inference with confidence intervals and hypothesis testing for one sample) including problems to be solved only in a theoretical way (like derivations of formulas or solving tasks with a pocket calculator).

The second part contains online materials giving an introduction to the statistics software R as well as hands-on exercises based on real data. The course development process has taken place in two phases, partitioned into six steps. In the first phase, the theoretical part of the course has been developed. In the second phase, the practical part of the course has been developed.

After the development of a part, a test run is done during one semester with feedback from the students. Depending on the outcome of the feedback, adjustments of the course materials take place, in order to improve the achievements.

The statistics course is taught annually in the winter term for the Computer Science study program. The development of the theoretical part took place during the summer term 2015. The testing and evaluation took place during the following winter semester. Currently, we are in step 5 of the second phase.

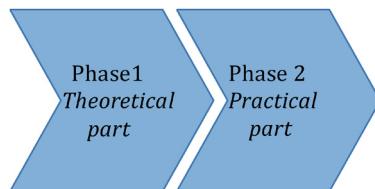


Fig. 1. The two development phases.

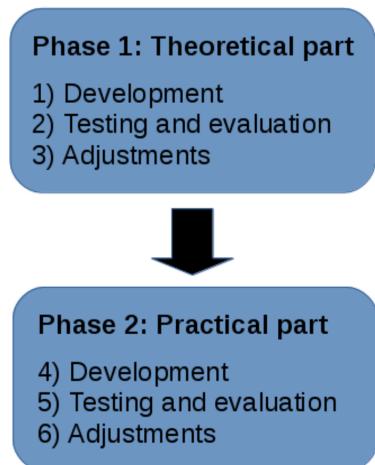


Fig. 2. The six steps within the two development phases.

5 Teaching Framework for the Statistics Module

The blended learning module Statistics for the study course Computer Science is taught during 15 weeks and is thus divided into 14 learning units with the last week reserved for recapitulation. Each learning unit starts with the definition of the objectives of the unit. The blended learning course has been implemented in the e-learning platform Moodle [27]. This platform has been chosen because this is the standard e-learning system at FRA-UAS, and also because it supports the implementation of mathematical formulas in LaTeX-notation.

5.1 Contents of the Module

The module consists of following components:

- **Online Lecture Notes.** The extensive lecture notes contain the complete theory covered in the course as well as examples with detailed solutions.
- **Time Table and Objectives.** A detailed timetable with learning objectives for each unit is provided.
- **Screencasts of the Lecture Slides.** To each learning unit, a set of lecture slides has been created. A separate screencast has been produced for each of the lecture slides. This means that the students can either play all screencasts as a long movie or select some especially difficult screencasts and recapitulate only these ones.
- **Exercises.** Totally, 158 exercises with different difficulty levels have been constructed. Three links belong to each exercise: the answer link, the hint link and the solution link. The answer link tells only the final result of the exercise without the solution, whereas the hint link gives a hint to where in the lecture notes help concerning the topic of the exercise can be found. This link should encourage the students to continue if they get stuck, without having to look at the complete solution at once. Finally, the solution link contains the complete solution to the exercise. In Fig. 3, an example of an exercise with the three different links is shown. For each of the exercises, screencasts telling the whole solution and the idea behind have been constructed.

To help the students choose in which order they should solve the exercises, a traffic light system has been introduced: The red flagged exercises are basic exercises to start with, the yellow exercises are on an intermediate level and, finally, the green flagged exercises are exam-level compatible.

Online Quiz Questions with Automatic Correction. Integrated in the Moodle environment, 140 online questions help the students to confirm that the study objectives of each learning unit have been reached. To each of the learning units, 10 quiz questions have been constructed. In Fig. 4, an example of a typical online question is shown.

In Fig. 5, we see an example of the solution to an online task and observe that mathematical formulas are possible to use, since the LaTeX-notation is available in Moodle.

Exercise 1.4: Coins

On a table, there are some coins, having different currencies. For the coins, mention

- a. a qualitative variable,
- b. an ordinal variable,
- c. a quantitative, discrete variable,
- d. a quantitative, continuous variable.

Hints 1.4 Answer 1.4 Solution 1.4

Fig. 3. Example of an exercise question with the links to the hints, the answer and the complete solution.

Question 1:

10 observations from a normal distribution have the sample mean 10.1 and the sample variance 0.9.

Determine a 90% confidence interval for the mean.

Select one:

- a. [5.876, 9.888]
- b. [1.234, 8.987]
- c. [8.789, 10.987]
- d. [9.551, 10.649]
- e. [3.876, 9.876]

Check

Fig. 4. Example of an online quiz question.

$$\begin{aligned}\bar{x} \pm t \cdot \frac{s}{\sqrt{n}} &= 10.1 \pm 1.83 \cdot \sqrt{\frac{0.9}{10}} \\ &\Rightarrow 10.1 \pm 1.83 \cdot 0.3 \Rightarrow [9.551, 10.649]\end{aligned}$$

Fig. 5. Example of the solution to an online quiz task.

The 10 quiz questions, belonging to one learning unit, are organized in two sections with five questions in each section. The students are encouraged to test their immediately gained knowledge after having finished the learning unit by solving the first five quiz questions. The other five quiz questions are recommended to solve later as a recapitulation or even as exam preparation.

Test Exams. Furthermore, online mock exams, consisting of questions from earlier exams, are offered. These are the only materials, which are not available at course start, but appear online a few weeks before the exam will take place. This is done to reduce the risk that the students should try just to learn by looking at old exam problems.

Tutorial for the Statistics Software R. Even if statistics is a discipline where the application to real data plays a crucial role, the conventional statistics course does not contain any parts using statistical software. The reason for this is, unfortunately, lack of time for covering this important aspect in the conventional course. In the blended learning statistics course, the students are encouraged to start learning how to use the statistical software R to small, real problems. The R software [28] has been chosen, because it is free software and also possible to use with different operating systems (Windows, Linux, Mac). The R part of the blended learning course is optional, since this part doesn't exist in the conventional statistics course and according to the examination rules at Fra-UAS, the exam contents must be exactly the same for both courses.

5.2 Tuition of the Course

The tuition of the course is organized in the following way:

- **Self-Study of the Course Materials.** This is the main study method of the course. The students are encouraged to work independently with the provided course materials. As a help for an enriching communication among the students, which is important for the progress of the self-study, can a discussion forum in Moodle be used.
- **Weekly Web Conferences Using the Adobe Connect System.** The weekly web conferences are used to summarize important contents of the learning units, leaving time for questions and discussion. This means that the web conferences don't intend to be whole lectures of the total content of the module. This is already covered in the screencasts of the lecture slides.
- **Two Optional Assignments with Deadlines.** The students are twice offered to submit solutions to problems, which then are corrected by the responsible teacher. Even if the assignments are completely optional, they have deadlines in order to encourage the students to study continuously during the semester, not only in the end.
- **Three Half-Day Workshops.** To give the students the opportunity to discuss their open questions face-to-face with the responsible teacher, three optional half-day workshops are organized. The workshops take place during the weekend in order to avoid collisions with the regular time schedule of other courses.

6 Evaluation at Course Start

At course start, we asked the students to complete a survey about their reasons for taking the blended learning course in statistics. Especially, we wanted to compare the students' thoughts about their advantages of e-learning with the problems we expected to address by using the blended learning course.

In Fig. 6. we can see that most students estimate that they will have a personal advantage concerning time and location independence when taking the blended learning course in statistics. This is conforming with our expectations that the blended learning course would address some general problems occurring by the face-to-face courses, such as time collisions and that students not attend lectures due to a long traveling distance to the university. Furthermore, we see that also the possibility to apply an individual learning pace is considered to be an advantage by many students. This corresponds to our intention to use the blended learning course to facilitate the course for students with different backgrounds, especially for those with a low prior knowledge in mathematics.

In Fig. 7, we see that three factors have been selected most often by the students as the factors, which are most important for their personal learning process: Interest in topics, the professor and the media usage. Concerning the blended learning course in statistics, we have good opportunities to influence the outcome of the factor media usage, since a lot of different techniques are available in this course. To increase the interest in the topic statistics, coincides with our objective to use the course materials in such a way that the students can overcome their lack of motivation to take the course.

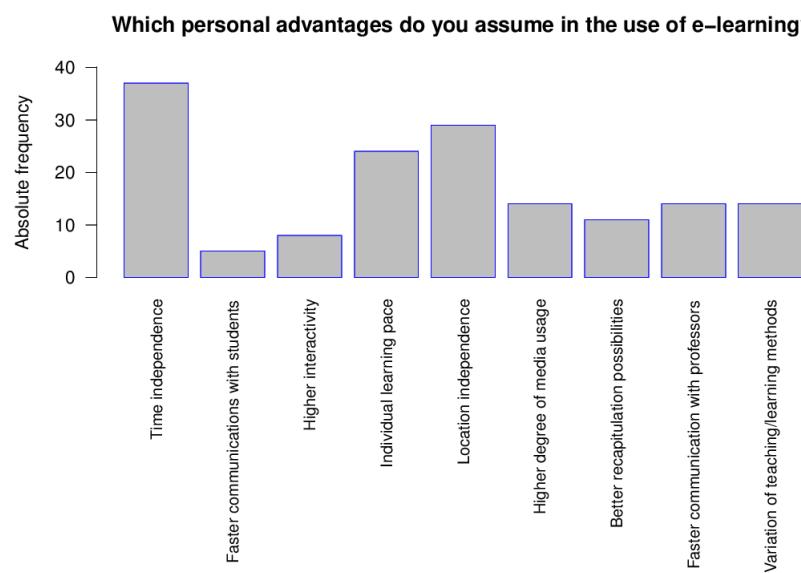


Fig. 6. Evaluation of the advantages the students estimate they will have by using the blended learning course in statistics.

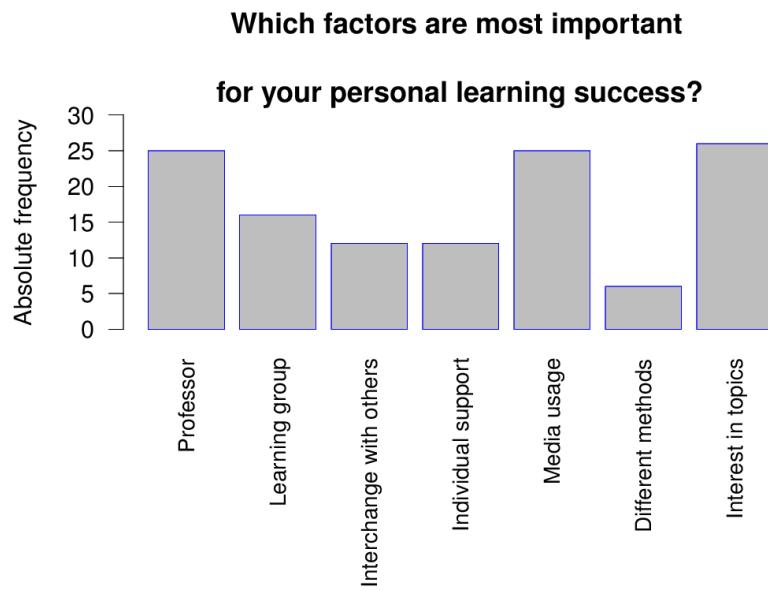


Fig. 7. Evaluation of the factors, which, according to the students, are most important for their personal learning process.

7 Experiences from Experimental In-Class Usage

From the first experimental in-class usage, we could see that the problems arising in the conventional teaching of statistics for engineering students at FRA-UAS could be addressed by the application of the e-learning module in the following way:

7.1 Addressing the Reputation of Having High Failure Rate

The problem with a high proportion of students taking the course at a later time point than planned in the curriculum, because they are afraid of failing, is decreasing, since the introduction of time- and location-independent teaching reduces the number of lecture collisions for later-taking students, i. e. they have now better possibilities to complete the whole course without large knowledge gaps.

7.2 Addressing Lack of Sufficient Prior Knowledge in Mathematics

The problem with the low level of knowledge in mathematics could be approached with extra online teaching materials, covering basic concepts. The use of these materials is, of course, optional, since the mathematics knowledge among the students is heterogeneous.

7.3 Addressing Lack of Motivation

To increase the interest in taking the course, real-world problems were integrated in the course materials. The objective was here to provide the students with such problems that they could experience in daily life, both currently and later as engineering professionals. Such kind of problems need time to be described and processed and are therefore not possible to include in the tight schedule of the conventional statistics course.

7.4 Addressing No Hands-On with Real Data, Lack of Knowledge of Handling of Statistical Software

Due to the relaxed time constraints in the online course, teaching materials thoroughly covering the statistical software R are provided. This gives an immediate possibility to let the students explore real-data sets, which makes the understanding for statistical methods deeper.

7.5 Addressing General Problems

By the introduction of time- and location-independent teaching, the scheduling problems with colliding lecture times as well as the absence of the students due to a long traveling time to the university could be avoided to the greatest possible extent.

8 Conclusions and Future Work

In this paper, we have presented the development of a blended learning module in statistics for undergraduate engineering students. The module has been developed in two phases, the theoretical and the practical. Each phase consisted of three steps: Development, testing including evaluation and adjustments according to the results of the testing and evaluation.

We conclude that a lot of the problems frequently occurring in teaching statistics in a traditional way to non-statistical students at FRA-UAS could be addressed by the blended learning module in statistics, such as time collisions, a long traveling time to the university, lack of sufficient prior knowledge in mathematics, lack of interest due to less time to treat real-world examples in the traditional statistics course, lack of hands-on experience with statistical software, etc.

Future work would be to assess the learning success among the students taking the blended learning course and compare these results to those of the students taught in a traditional way. This has not been done so far, since the traditional statistics course and the blended learning course have until now been taught by different lecturers, which would make direct comparisons difficult. Another factor, which should be taken into account in the comparison, is the allowed mixing of the two course structures, i. e. that a lot of students partially take the blended learning course as well as partially the traditional statistics course.

9 Acknowledgment

This work has been supported by the project MainCareer at FRA-UAS.

10 References

- [1] Manyika, J., Cxhui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Byers, A.H., “Big data: The next frontier for Innovation, Competition, and Productivity”, [http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_i nnovation](http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation), last accessed November 7, 2016.
- [2] Minton, P. (1983), “The Visibility of Statistics as a Discipline”, *The American Statistician*, vol 37.4a, pp. 284-289. <https://doi.org/10.1080/00031305.1983.10483121>
- [3] Wasserstein, R. (2015). “Communicating the Power and Impact of Our Profession: A Heads up for the Next Executive Directors of the ASA,” *The American Statistician*, vol. 69, pp. 96-99. <https://doi.org/10.1080/00031305.2015.1031283>
- [4] Carnevale, D. (2002). “Online Students Don’t Fare as Well as Classroom Counterparts, Study Finds,” *The Chronicle of Higher Education*, vol. 48.27, 38.
- [5] Russell, T. (1999). “The no significant difference phenomenon”. Chapel Hill, NC: Office of Instructional Telecommunications, University of North Carolina.
- [6] Summers, J., Waigandt, A., Whittaker, T. (2005). “A Comparison of Student Achievement and Satisfaction in an Online versus a Traditional Face-to-Face Statistics Class”, *Innovative Higher Education*, vol. 29.3, pp. 233-250. <https://doi.org/10.1007/s10755-005-1938-x>
- [7] Stephenson, W. R. (2001). “Statistics at a distance “. *Journal of Statistics Education*, vol. 9.3, 24-5.
- [8] Yablon, Y. B., Katz, Y. J. (2001). “Statistics through the medium of internet: What students think and achieve “. *Academic Exchange Quarterly*, vol. 5.4, pp. 17-23.
- [9] Ward, B. (2004). “The Best of Both Worlds: A Hybrid Statistics Course “, *Journal of Statistics Education*, vol. 12.3, pp. 74-79.
- [10] Utts, J., Sommer, B., Acredolo, C., Maher, M. W., Matthews, H. R. (2003). “A study comparing traditional and hybrid internet-based instruction in introductory statistics classes “, *Journal of Statistics Education*, vol. 11.3, pp. 171-173.
- [11] Everson, M., Garfield, J. (2008). “An innovative approach to teaching online statistics courses “. *Technology Innovations in Statistics Education*, vol. 2.1.
- [12] Symanzik, J., Vukasinovic, N. (2006). “Teaching an introductory statistics course with CyberStats, an electronic textbook “. *Journal of Statistics Education*, vol. 14.1, 1.
- [13] Dutton, J., Dutton, M., Perry, J. (2001). “Do online students perform as well as lecture students?”, *Journal of Engineering Education*, vol. 90.1, p. 131. <https://doi.org/10.1002/j.2168-9830.2001.tb00580.x>
- [14] Brown, D. G. (2001). “Hybrid Courses Are Best,” *Syllabus Magazine*, vol. 15.3, p. 22.
- [15] MacGregor, C. J. (2001). “A comparison of student perceptions in traditional and online classes”. *Academic Exchange Quarterly*, vol. 5.4, pp. 143-149.
- [16] Navarro P. and Shoemaker, J. (2000), “Performance and Perceptions of Distance Learners in Cyberspace,” *American Journal of Distance Education*, vol. 14.2, pp. 15-35. <https://doi.org/10.1080/08923640009527052>
- [17] Tudor, G. (2006). “Teaching introductory statistics online—satisfying the students “. *Journal of Statistics Education*, vol. 14.3, n3.

- [18] Johnson, H. D., Dasgupta, N. (2005). "Traditional versus non-traditional teaching: Perspectives of students in introductory statistics classes". *Journal of Statistics Education*, vol. 13.2, pp. 1-12.
- [19] Wissenschaftliche Begleitung (2015). "Aufstieg durch Bildung: offene Hochschulen" (Upward mobility through academic training), <http://www.offene-hochschule.de>, last accessed November 8, 2016.
- [20] Johnson, M., Kuennen, E. (2006). "Basic math skills and performance in an introductory statistics course ". *Journal of Statistics Education*, vol. 14.2, 14.
- [21] Lunsford, M. L., Poplin, P. (2011). "From research to practice: Basic mathematics skills and success in introductory statistics ". *Journal of Statistics Education*, vol. 19.1, pp. 1-22.
- [22] Ely, D. P., Hittle, L. (1990). "The impact of math background on performance in managerial economics and basic finance courses. *Journal of Financial Education*", vol. 19, pp. 59-61.
- [23] Garfield, J., Ahlgren, A. (1988). "Difficulties in learning basic concepts in probability and statistics: Implications for research ". *Journal for research in Mathematics Education*, pp. 44-63. <https://doi.org/10.2307/749110>
- [24] Yilmaz, M. R. (1996). "The challenge of teaching statistics to non-specialists ". *Journal of Statistics Education*, vol. 4.1, pp. 1-9.
- [25] Fulton, K. (2012). "Upside down and inside out: Flip your classroom to improve student learning ". *Learning & Leading with Technology*, vol. 39.8, pp. 12-17.
- [26] Graham, Charles R. (2013). "Emerging Practice and Research in Blended Learning" in *Handbook of Distance Education*, Moore, M.G. (Ed.), Taylor & Francis, 3rd edition, pp. 333-350 <https://doi.org/10.4324/9780203803738.ch21>
- [27] Moodle (2015). Moodle – Learning Management System, <http://moodle.de>, last accessed November 8, 2016.
- [28] R Development Core Team (2014). R: A Language and Environment for Statistical Computing. The R Foundation for Statistical Computing, Vienna, Austria.

11 Authors

Christina Andersson holds a PhD degree in Mathematics from University of Kaiserslautern, Germany. She has been working in industry as a trainer in statistics and data mining for the SAS Institute GmbH. Currently Christina Andersson is a professor for mathematics and computer science at Frankfurt University of Applied Sciences, Dept. of Computer Science and Engineering, Frankfurt/Main, Germany.

Doina Logofătu holds a PhD degree in Computer Science from Babes-Bolyai University in Cluj-Napoca, Romania. She worked in the educational field as well in the industry side. She wrote some successful programming and mathematics books in Germany and Romania. At present she is professor for mathematics and computer science at Frankfurt am Main University of Applied Sciences Dept. of Computer Science and Engineering, Frankfurt/Main, Germany.

This article is a revised version of a paper presented at the EDUCON2017 conference held in Athens, Greece, 25-28 April 2017. Article submitted 17 July 2017. Published as resubmitted by the authors 18 August 2017.