NEW TECHNOLOGIES, SOFTWARE, AND E-LEARNING – ENRICHING COURSES IN INTRODUCTORY STATISTICS AND PROBABILITY[1]

Manfred BOROVCNIK[2]

Abstract

New Technologies changed the applications of statistics completely so that more statistics is applied nowadays than 25 years ago, which generates a need for a wider education in statistics even at an introductory level. At the same time, these technologies open the way to teach the subject matter of statistics completely differently than in the pre-computer era. We illustrate various approaches of technology-based teaching and identify key issues for their success. Large-scale e-learning projects, locally organised computer-based learning environments, and additional applets used to illustrate complex concepts, calculations outsourced to software are some of the many options to enrich the statistics course. The recommendations are for students of non-mathematical studies. The future class at university includes a diversity of information sources and a new role for academic teachers as organisers of the work in small groups rather than communicating the content in a big lecture hall. We illustrate the potential of systemic solutions for an introductory course in probability and statistics by a general discussion and by the extensive and long-term feedback from our own students.

Keywords: *Success criteria for e-learning, Feedback for students, Exemplary solutions to tasks, Statistical applets*

INTRODUCTION

We discuss various big-scale e-learning environments and draw some early conclusions about the setting of e-learning (Section 1). Blended learning courses, i.e., locally developed and administered courses that are enriched by e-learning elements, are more flexible (Section 2). We then summarise the debate on e-learning and New Technologies (NT) in university education to clarify the problems and the special challenges of the new capabilities for teaching (Section 3). The vital issues for students' progress are identified, as we see them and our decisions on the essential questions are presented and set to discussion for improving them or adapting them to other environments (Section 4). Our approach to blended learning is illustrated by some exemplary applets (Section 5) and an exemplary solution to a task that we hand out to the students after we have discussed the task in the class (Section 6).

In Sections 5 and 6, we illustrate our perception of teaching an introductory course in statistics for business students and give incentives to enrich teaching by interactive sequences, which may clarify crucial and complex concepts in statistics. Finally, some conclusions are presented (Section 7). A main consequence of our research on the course is that feedback between teachers and students, among students, and among staff is a vital issue for the success of blended courses.

---

[1] Extended version of an Invited Plenary, held at International Statistics Days Conference (ISDC/IGS 2016), Giresun, 7. 10. 2016.

[2] University of Klagenfurt, Austria, manfred.borovcnik@aau.at

The fact, that such feedback is lost in large systemic solutions may contribute to their relative failure.

E-learning was promoted strongly

There were great hopes to make use of the potential of New Technologies (NT) and e-learning. General learning platforms (like Moodle) for the exchange of learning material and organising discussion and peer work were promoted. Large inter-university projects in several disciplines were funded to develop courses that should be used all over a country. Furthermore, to boost the development, awards were organised; in Germany, the Medidaprix [11] for educational software had a yearly a prize of 100 000 €. More recently, massive open online courses (MOOCs) were developed and a great enthusiasm about their effect on university learning spread among the academia. Besides these endeavours, many university teachers tried to support their courses with elements of e-learning such as providing data, providing extra-material, or applets illustrating complex concepts, etc.

The rationale was partly driven by the need to reduce cost of teaching by using synergies between universities: why have a different introductory course on statistics at each of the universities? A further driving force may be seen in the need to adapt university education to a modern and flexible style of exchange of communication that students are familiar with. They use laptops and smartphones daily and are used to get more information on demand.

Aim

The aim of this paper is to discuss general criteria of e-learning (and also learn from larger projects from the past): What are the vital issues for success of e-learning endeavours? What are good criteria for NT learning materials? Furthermore, we want to give reasons for the decisions we made in the various options for our courses. Discussing about NT learning materials without showing some, might not be sufficient. Thus, we characterise indirectly the crucial criteria and our own approach by some applets we use to supplement teaching. Our decisions were influenced by the circumstance that we deal with business students who have a limited interest and pre-knowledge in mathematics. This is not really a restriction as courses in introductory mathematics or statistics for other studies (even for informatics students) experience a similar situation.

Large-Scale e-Learning Environments

The larger the scale of a learning environment, the more leading persons might be involved; the situation is similar with the financial resources. The attraction comes from the greatest centres and heads. However, larger systems involve other responsibilities and other, systemic solutions. They might be more effective with respect to a ready-made use. However, larger systems always lack flexibility. We inspect two extremes, massive open online courses and a country-wide project to unify a specific course.

Massive open online courses (MOOCs)

MOOCs combine traditional forms of teaching such as reading material, problems, discussion fora, video sequences of various types (video-taped lecture, animations, real environments of problems investigated, etc.). They are online courses with a great number of participants. A key advantage is that they are free of cost. Moreover, such MOOCs take advantage of the fact that a renowned person (which is often the case) rather than “someone from next door” explains the concepts, which – for many students – strongly increases their motivation. Moreover, videos are prepared, which would not be possible locally (not only for cost but also for facilities).

However, during the learning phases, feedback about steps taken and steps still required is vital, which is not so easy within such courses even if they provide discussion. Peers might not have an answer either, or their feedback may not be trustworthy; feedback from staff is usually not provided. On-going monitoring of learning is not only a control from outside but directs students’ efforts. Such monitoring cannot be granted within such MOOCs. Furthermore, we have to contrast the hype around MOOCs to their actual impact: "Registration [...] is around 50,000 [...] first lecture attendance is only 10%, finishers are only 1-2%," according to a comment in The Chronicle of Higher Education, a widely recognised forum for current problems in education [22].

Six key trends accelerating technology adoption in higher education are identified in the Horizon Report 2016 [16]:

- *Long-term* (advancing cultures of innovation; rethinking how institutions work).
- *Mid-term* (redesigning learning spaces; shift to deeper learning approaches that favour hands-on and student-centred experience).
- *Short-term* (measuring learning through data-driven practice and assessment; increasing use of blended learning designs).

MOOCs are no more listed as key trends there while redesigning learning spaces, deeper learning approaches and blended-learning designs are among the current trends. The focus of educational research in technology use includes "a visualized analytics system to assist students in their progression through core competencies, provide educators with dropout prediction, and evaluate ways to increase engagement in MOOCs" ([16], p. 53). Even the motivational aspect seems to be unsatisfactory. How difficult it is to develop a degree of personalised learning in such environments is also recognised in this report: "A major barrier, however, is that scientific, data-driven approaches to effectively facilitate personalization have only recently begun to emerge […]" ([16], p. 32).

We conclude that MOOCs are not sufficient to replace regular courses. They may best be used as an add-on for a traditional course; a source of information on content or examples that may be directed towards difficult concepts or to specific applications. If discussed in class (including a presentation from the students to summarise), they may enrich the local course.

Inter-university e-learning environments

It is tempting to design a course for a whole country. In Germany, e.g., the country is not too big and the cultural differences between the universities are not too large, to agree on contents and methods to cover and on the way how this is presented and organised (including classes with exercises and exam papers). One course at introductory level for statistics for a whole country could save a lot of effort that is invested by academic teachers year by year. That was one big incentive for research funds to provide money for a project intended to develop such a course.

The German example of EMILeA (funded by 3 million €) was one of the largest projects in the educational sector. For its aims, see [7] and [9]. The hopes accompanying that project were not realised. The learning environment is used only by the original project partners twelve years after its completion. This may also be seen from the critical attitude against e-learning in [14] and [15]. Some of the criteria that might account for failing are that educational considerations are still essential, the learning environment has to be flexible, and the decision to use a programming language has to consider the workload and flexibility of use for students. Even if programming in R could be avoided in a first approach, the decision for R decreased the flexibility of use of the course material and the flexibility to create one's own calculations for the students. And it meant an increase in workload that many students were not willing to accept.

Large projects such as EMILeA tend to focus on logistic problems. That is always the case with larger institutions. Too many technical issues of compatibility on all IT platforms and a widely usable exchange of materials used too many resources (human and money) so that by the end simple criteria of how the layout of content influences or restricts the way students can understand the concepts presented were neglected, or that students are discouraged if they do not get the required feedback on their (technical and learning) problems within a reasonable period. The following statements support a critical attitude towards e-learning ([14], p. 362): "We believe that e-learning is required in modern statistics education but we do not share a too optimistic view that it will also deepen the knowledge of students in statistics." Elsewhere Härdle, Klinke, & Ziegenhagen ([15], p. 12) conclude that

- "1. e-learning cannot replace the interaction of student, teacher and blackboard" and
- "2. e-learning tools can only be successful if they satisfy the need of all participants of the system."

They continue to name requirements for excellent electronic media in education: "Robust and reliable technology, high-quality contents and the willingness to adjust the [...] behaviour from both, the students and the lecturers." A decision for complex software can hardly be considered as appropriate for an introductory course; this may even be doubtful for mathematics students.

The Impact of New Technologies

We focus our discussion on the impact of New Technologies (NT) to teaching statistics. NT have changed statistical methods and the applications of statistics. Statistical methods spread over all branches of science, over all parts of administration, over all sectors of industries. On the Internet, rich sources of multi-variate data sets have become available and wait to be used for improving our views on problems and our decisions. Computer-intensive methods amend, improve, or replace traditional methods of statistics One branch is resampling and re-randomisation that can solve many problems of statistical inference satisfactorily. A different impact from NT may be seen in the new ways to communicate with each other and with problems – free of time and place, enriched with computing and graphing facility so the students may focus on the tasks that are set and have to be discussed and solved. It means that they may be much more involved in the systems analysis of the task and the joint effort to find suitable methods.

We discuss the issues of NT from two different sources, one is a session on NT at the *World Statistics Congress* in Lisboa [12] and the other is one volume of the *International Statistical Review* dedicated to e-learning in 2007 [14, 19, 20, 21, 24].

Computers and NT change applications and methods of statistics

We illustrate the changes in applications by two examples on selecting rocks and on buying behaviour of clients: Rocks are inspected in the quarry by infrared spectroscopy. The aim is to classify them for their quality (properties such as abrasiveness, brittleness, or compressive resistance). Light of different wavelengths sent through the rocks produces curves of absorption for a test rock. The massive data have to be used online while the rocks are passing by to give a decision about their quality and separate them accordingly. To make the problem tractable, the data have to be transformed into Fourier series and can be down-scaled to the Fourier coefficients. The classification method of traditional statistics is combined with complex mathematical methods to yield a new method. An implementation in real-time has to rely heavily on computer power.

The second topic is to investigate a stock of clients for their future buying behaviour, either real buying behaviour on existing goods, or potential for the development of new goods. To find segments of clients for a reliable prediction for their future demand for existing goods is already complicated enough. If such a modelling is successful, it may help planning. Consider online shops that have to design their reactions to the requests of customers. How to organise automatic collections of data, is the easiest but basic part of the modelling. How to introduce a hierarchy on goods and how to measure a conceptual distance between the goods to provide specific information for that client (and not for others) to influence his purchasing behaviour, is the more difficult part of this modelling.

Computational statistics revolutionises statistical methods and may replace sophisticated methods by easily understood methods (resampling can be done without deeper knowledge about probability distributions, e.g.). A discussion of the changes of applications of statistics by NT may be found in [2].

New tech open new ways to teach but require new forms of communication

Films like Moore's "Against all Odds" [17] spread easily; stored on DVD they do not require extra devices to look at them. Software helps to circumvent lengthy calculations and yet use the results so that educational efforts may be focused on understanding the method and the inherent concepts, or apply the method to various contexts so that the applications seem more realistic even in teaching. Software enlarges the pool of methods available but also prompts to use it blindly without careful thinking about the restrictive assumptions of the method or of the result.

Software brings graphing facilities into our reach and allows for multiple, exploratory analysis of data. That raises new methodological questions not so prevalent in the pre-computer era. Software allows for computer-intensive methods of computational statistics that are conceptually often more easy than traditional statistical methods.

Software allows studying the effects of the concept under investigation; it includes also the powerful tool of simulation. Any probability distribution – regardless of the interpretation of probability – may be reconstructed in material form by a simulation scenario that delivers relative frequencies that approximate the probabilities. In this way, e.g., the power of a test may be visualised in material form by simulating the situation under a specific alternative hypothesis. One may even change the alternative hypothesis (its expected value) gradually to show the effect. As simulation is done so fast, one can see the effect nearly in real time.

New technologies require new forms of communication and teaching. The usual linear sequencing of information to develop the concepts (as used in traditional textbooks) is quite ineffective for self-study, as explained below. To facilitate that learning becomes effective, e-learning requires a more sophisticated infrastructure than tutorial classes.

Several amendments have been introduced to break the sequencing of content such as glossaries, pop-up information about key concepts on demand, examples if required, extra explanation or an interactive applet that can be inserted if the students feels the need. These components are in stand-by mode and can be used to give feedback when necessary. Non-linear e-sequencing requires a more sophisticated infrastructure than a tutorial class to supply learning to become effective. In a personal discussion, the learner can ask for more and better explanation, the teacher may get feedback about something that has not been understood (e.g., [24]). In a systemic approach, such situations have to be anticipated and responses of the system have to be implemented that would flexibly react to the new request (e.g., [19], [20]).

A neural network of the relations between the concepts might support the individual learning paths as it is organised by conceptual distance as is discussed in [6]. In such a neural network, the usual linear sequencing of content is replaced by conceptual distance. That may assist the learners to organise their own learning paths through the scattered pieces of information. Such a new structure could function as organisational memory to enhance learning also in traditional environments. However, both teachers and learners have to get accustomed to use structures beyond mathematics. Further investigations are required before the idea of neural networks can effectively be used to support the comprehension of mathematical concepts; an application of neural networks seems to be more appropriate in a knowledge basis such as in medicine or law.

Computer-assisted learning environments

The key for designing effective hypermedia for teaching is undoubtedly *structure* according to Schuyten and Thas ([20], p. 368). It is difficult to anticipate individual learning paths and an interactive adaption to individual (singular) difficulties is missing. “Structure can be easily given by the lecturer, but more careful thinking is needed [... in] an e-learning environment. Learning paths may be part of the solution, but nowadays they are often still too linear ...” (p. 369). Dynamic environments as offered by Nolan & Lang [19] provide flexible documents, which can be enlarged by the users. However, the decisive step of embedding them into the structure of a statistics course requires much work, which has not yet been undertaken.

Making explicit the value of statistical methods for the students is an important step for any teaching in statistics as it is by no means obvious for many learners. Tamura [21] comes to the conclusion that missing this requirement is a common failure of teaching. In constructing e-learning environments, the focus is shifted to solve many other – partially technical – problems so that insufficient effort is invested to implement context and feedback to underpin how statistics may help the learners to develop promising strategies for their own future careers.

One specific difficulty of learning environments is to anticipate individual learning paths. Another one is to provide guidance for the students. Wild [24] names a further crucial point of any e-learning endeavour: With tutorial classes and the possibility for direct feedback, it is much easier to recognise when the (single) students *perceive* tasks or concepts differently from the *intention* and correct accordingly. He ([24], p. 329) states "Our experience with simple applets also gives transferable lessons for virtual environments. One lesson is that which seems to make something blindingly obvious to the initiated is often completely opaque to the novice." An e-learning environment has to *preview* such difficulties and design facilities to deal with such a break-down in communication. One element designed for that purpose is an organisational memory; others are to provide a clear structure, a distinct focus on the purpose of the content, or tasks for feedback to the student. Visualisations (and simulations) become much more easily available in a course with online elements than in a traditional course. However, that does not imply that they are used more effectively. With the "intention and perception" caveat of Wild in mind, a prevision of students' remote use of learning material enriched by visualisation should help to channel students' perceptions after study.

Computers might change teaching without e-learning

The computer may change the ways and the kind of communication completely, apart from the option of e-learning. One idea is to change the role of the teacher who traditionally transmits the content and determines the pace of the work of the learners. The students – in small groups – could sit in front of the computer and work on pre-set tasks at their own speed discussing the required steps within peers. The students would get their information about the concepts from a hypertext, or from an e-learning text, and work independently. Intermittently the teacher could assemble their approaches and introduce a classroom discussion by showing them where their different approaches would lead to. If the members of the group decide that they need extra information or help, they may address the teacher. The teacher would also go from group to group and oversee their progress and may occasionally give further hints, clarify misunderstandings about the task or the context, counsel about the required methods, assist with logistic problems in using the computer, etc. When the students then learn more about the concept, they learn it on demand so that the new knowledge becomes useful to solve the set task.

The teacher's role in such a setting is to organise students' individual or group work; to consult them how to find the required information or method; to discuss the difficulties of their approaches; or to use their comments and questions to follow-up students' work, e.g., to provide a list of frequently asked questions (FAQ) with appropriate explanations. In other words, the teacher resumes the role of an advisor [3].

Keys for designing effective hypermedia

It is important to note that low achievers benefit the least from such learning environments so that the gap between top and bottom achievers widens. Most learners need more structure in the learning materials, which becomes a key element for the success of e-learning elements [20]. Structure has to anticipate all learning difficulties – no flexible adaption to current needs can be given individually. As the materials are scattered with no overview given what else is available and no guidance where it is situated and how to get it, the students need a carefully designed structure of all materials.

Intention and actual perception can go apart. It is difficult to anticipate individual learning paths. However, it is vital to foresee the most frequent at least as there is no interactive adaption to individual (singular) difficulties once the materials are in use. Tutorial classes offer direct feedback to recognise when (single) students perceive tasks or concepts differently from the intention and correct accordingly.

Blended-learning courses

Early analyses of the impact of blended courses show [23] that there are no differences in performance of students, that there is a preference for the interactive way of the face-to-face course, finally, the problems of orientation in an electronic environment are frequent and hinder their success.

There is a wide range of measurements to implement e-learning that reaches from using a few technological add-ons, enriching hypertexts by applets to illustrate difficult notions to administering the course in mega classes (with video-taped lectures, weekly exercises assessed automatically, and tutoring in small groups by student tutors, see [10]).

We have implemented four levels of guidance in our introductory statistics for business and administration: the lecturer, the coursework in small groups tutored by academic staff, student tutors, and e-tutorials. We have monitored feedback in different ways: a general questionnaire at the end of the course is supplemented by a weekly questionnaire on the tasks when our solutions are handed out. Students judge the tasks on a 7-point scale with respect to task difficulty, motivation of task and its context. They also judge our specimen solution with respect to its comprehensibility and usefulness for learning. From an open category, we get their personal comments on the reasons why solutions are not convincing or understandable.

Vital Issues for Students' Progress

In this section, we deal with the following issues and identify them as keys for students' progress: the decision about software, the development of suitable tasks for the students, feedback for the students, orientation, developing suitable applets, co-ordinating the academic staff and the tutors. It is essential that the options are carefully compared and a sensible decision is made that addresses the special challenges and needs.

The decision about software

Software is useful for the calculations, for graphical displays, and to enhance complex notions by animations and simulation. It has a huge impact on the setting and on the workload of students. Extra-work on computing de-motivates many students. We chose Excel for various reasons:

- It is used in many other courses;
- It is wide-spread in workplaces;
- Students have rudimentary skills in Excel;
- It is easy for the basics; and
- It is very flexible to show some effect on demand if a question arises and needs a direct answer.

Our own experience and feedback from students strongly support this approach. As Excel cannot directly be used for some tasks that are standard also in an introductory course of statistics for business students, it has to be backed up. Instead of buying ready-made add-ons, we developed templates for specific demands and made them accessible to the students (for box plot, histogram, analysis of tabular data, e.g.). Excel is of limited use for higher-level courses but only a small minority of the students attend such courses. Few students react positively to the use of a complicated programming demand such as R.

For postgraduate courses, a decision for the software is different; Mougeot [18], e.g., uses R and shows its effectiveness. At that stage, only those students who specialise in statistics and who have already understood the basic concepts use the software to write programmes that flexibly solve advanced tasks. As an alternative, one might opt for Minitab or other software that has been written with the idea of high user friendliness; Minitab is comparably easy to handle and the package opens the way to many statistical methods, which are used in practice. As the majority of the students will remain mere consumers of statistical results, the higher effort to learn new software might not reward in the end.

There are arguments pro and con for any choice of software depending on the level of the course and the horizon of the students. The balance between the effort to learn and the ease to use may differ from group to group. The ease to use may not be the sole criterion as this may also hide details, which are essential for understanding the related concepts. The applets in Section 6 highlight that Excel may also be used to illustrate crucial concepts interactively so that Excel becomes an educational tool rather than a statistical package.

The development of suitable tasks for the students

Tasks may make statistics teaching more vivid; they can be used to enhance its purpose; tasks may also serve to illustrate the meaning of the used or investigated statistical concepts. Furthermore, tasks and their context are vital for the motivation of the students. Last but not least, statistics is an applied field and it is essential to show to the students how the final results may be interpreted in the context. There seems to be a consensus that – especially introductory courses for other studies – are data-driven and draw upon accessible contexts.

Tasks may serve several purposes. In their simplest form, they ask for step 1, step 2, etc. in order to get accustomed to the method learned. Often, content does not play a substantial role for this type of tasks. Tasks may illustrate some theoretical properties of the concepts that remain obscure if they are not embedded into a suitable context. Tasks may also ask for modelling steps with an adequate interpretation of the final result. For the modelling type of tasks, one may find exemplars in [1] or [5].

What signifies challenging tasks? We get feedback from the students by the weekly evaluation of the tasks and then modify the questions accordingly. From feedback, we have learnt that the context of tasks can help to make subtle concepts more plausible but also may mislead the students. We see also that language often is a crucial issue. Moreover, the length of a text can have an adverse impact on understanding. We are still surprised how a task was (mis)perceived by the students.

Feedback for the students

Feedback is vital for the students to revise their preliminary and shaky knowledge about the concepts and methods. However, feedback about solutions of posed tasks has to be given with little delay so that students can continue their studies sensibly. Exemplary solutions are vital but how to decide what is good? Our first approach was to give all steps of the solution in full detail. Amend the solution by graphs to enable the students to see where the concrete numbers asked for do fit in. And extend the solution by remarks on alternative ways to solve the task, by interpreting the answer in the context, and to relate the task and the result to other contexts. Feedback from the students showed us that the extensive solution was perceived as confusing. We revised our design of the solutions accordingly:

- The solutions should be hierarchically split into levels.
- What is absolutely required to solve a task?
- How can the solution be graphically enhanced?
- How can we implement the solution in the used software?
- How can a variation of the task enhance the result and the method used?
- How can the solution be enhanced by comments on the context?
- Can the solution be transferred to other contexts?

Solutions show the students how they can progress and help them to correct their ideas. To establish feedback to students on their progress on a regular basis is crucial whether the course is traditional or e-learning based, as shown by feedback from our students. Rather than to give an online compulsory quiz [10] or weekly test, we administer 5 tasks weekly, which the students have to prepare (and present on the blackboard or on their PC). And we opted for one exam during the semester and one at the end.

We developed a style of specimen solution to exercises. Following students' feedback we clearly signify the minimum requirements of a solution; we separate side remarks, we mark alternative solutions; further comments on the context and links to related concepts are visibly kept apart from the main body of the solution.

Orientation – how to find information if you search for it

With a book, the scope of the content is clear. However, with a dynamic learning environment, the pieces of information are scattered. Search systems (supported by tagging with keywords), glossary and overviews with links might help. Hierarchical placement of information – enfold more only on demand – is often helpful but difficult and cost-intensive to implement.

Feedback – if negative – referred to the extent of solutions and to missing search functions (keywords, index, etc.). The content that is required for the exam has to be clear from the online study material. And it should be easy for the students to use the materials for their work on tasks. For clearer orientation, we developed a companion manuscript to the reference textbook, which we use in the course ([8], a concise, yet non-mathematical-style textbook).

Other materials remain scattered:

- Applets for illustrating crucial concepts,
- Templates for calculations in Excel, which provide facilities for basic tasks (such as calculating statistics from tabular data),
- Applets illustrating the use of special functions in Excel that go beyond basic Excel but are immensely useful to organise and reduce the implementation of calculations.

Only by using these sources regularly during the course, the students are able to orientate themselves during an exam.

Developing suitable applets

What are crucial criteria for good applets?

Any software used in the course has to be flexible enough to facilitate illustrating some (theoretical) property or effect on demand if a question arises and should be answered directly. Some examples of dynamic animations for enhancing properties of statistical concepts implemented in Excel are in [4] or [5]. "What is an effective visualisation?" is an essential question for developing applets. But the following example from [13] shows that proper criteria for good visualisation still have to be elaborated: in illustrating distributions, the scales are automatically adapted to changes of the parameters; this hinders – in the author's view – the *comparison* of the *shape after* some parameter has been changed to the *shape before*, which makes it difficult to judge the influence of the parameter and recognise its meaning.

This example shows that there is not yet a consensus on this issue; the authors of this applet [13] answered to a critique that the same shape before and after illustrates that the parameters have no influence as the shape remains the same (the example deals with the normal density and the change of the shape by shifting and scaling). How can the standard deviation be seen as scaling parameter if the scale of the diagram is always changed? How can the student see the effect of a shift of, e.g., 10 in the mean if he has to look first on the changed scale?

How to share ideas between the involved academic teachers? We back up Excel by special templates (box plot, histogram, and tabular data). We have developed dynamic animations enhancing properties of concepts (see Section 6, or [2] and [4]). Yet, what is an effective visualisation? For example, showing the influence of a change of a parameter, the diagram has to be scaled in the same way during all changes. This would allow tracing the changed place in the diagram interactively so that at least a kind of monotonicity may be recognised visually. That may enhance the meaning of the parameter under scrutiny.

Co-ordinating the academic staff and the tutors

To discuss the intention of the set tasks with tutors and staff reveals that our intentions remain all-too often too implicit. It is also vital to exchange experience with the students in class within the staff members. We have allocated academic staff for the exercises and additional student tutors to give a better chance to reach all the students. A tutor is approached quite differently by the students; there is no formal barrier, no fear that wrong questions will have a lasting effect, etc. As feedback from tutors is more authentic, it establishes a valuable additional source for revising our learning material.

Exemplary Applets

The applets presented here stem from a wider collection that we use in our courses to enhance complex concepts. They support our introductory statistics course for business and administration students and are available online (see also [4]). Key factors in the applets that illustrate crucial concepts are: to illustrate the effect of probability indices by relative frequencies, to show the relevance of parameters by changing their values to investigate the impact on other entities, to show theorems in action but to extrapolate limit behaviour by thought experiments. Rather than increasing sample size beyond limits, a fairly stable shape of some distribution is revealed by repeating the simulation several times with a fixed sample size.

Behaviour of control charts to prepare statistical tests

One key factor of many applets is that indices that designate some quality of a method that have an indirect (conditional) probability interpretation – such as type-II error – are illustrated by relative frequencies of a simulation scenario (Figure 1).

The control chart shows the individual values of samples of size 5 at repeated control points. The control lines are drawn for 1% and 5% excess of the process under regular conditions, that means that individual values of the items are modelled by a normal distribution with expected value 1000 (the target value) and a standard deviation of 25 (the process variability). Pressing the repeat-button, one may see that the percentage of inspections that lead to exceed the drawn control limits are quite stable around 1% and 5%, which gives the type-I error a meaning of unnecessary intervention in the production. In the middle part of Figure 1, one sees the behaviour of the control chart in case an actual deviation from the target value of 20 (20/25 of the standard deviation) has taken place: 37% of the inspections lead to exceeding the warning limits (and to further investigations for potential causes for the shift without interrupting the production, and in 18%, the control limits are exceeded so that the production is stopped. This reflects the type-II error for that deviation. This may be too large. However, a further investigation might lead to look at the inspection times how long no intervention takes place. While it may take more than 10 times to stop the process, rarely takes it more than 6 times that the warning limits are exceeded so that a search for potential causes starts. That means that such a deviation should not remain undetected too long. The space for further enquiries in the charts by repeating the scenario by the repeat-button seems infinite.

The link between the distributions of means and the population

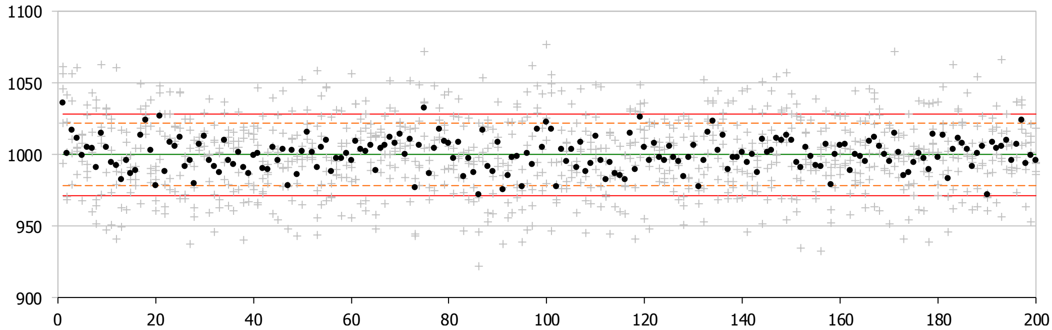
Another key element of some applets is that parameters of interest are dynamically changed and the influence of such changes may be traced in real time.

While the theorems for such relations are not proven, their impact is shown in the diagram of the density function. Shifting the expected value of the population (the single values) by the ruler shows that the expected value for the means is aligned with the first. Moreover, the distribution for the means of a random sample is much more restricted than in the population (roughly 92% instead of 68% within ±1 around the expected value). Increasing the sample size by the related ruler shows that the distribution of the mean constricts around the mean value (Law of Large Numbers), which shows that a larger sample provides more precise information about the central tendency of the population). This is unpacked from the formula and enacted in the dynamic visualisation. Such interrelations are invariant against a change of the standard deviation of the population, which again can easily be seen by the effect of changing this parameter.

Production under regular conditions - everything is under control, i. X ~ N(1000, σ^2=625)
Inspection involves at each control time 5 single items; 200 inspections are simulated
It is checked how the prescribed control (CL) and warn limits (WL) behave.

Chart for mean values of 5 items	CL	False alarms C	WL	False W
Number of samples	200		200	
Samples within the limits	198		184	
Proportion of samples within the limits	0.990	0.010	0.920	0.080

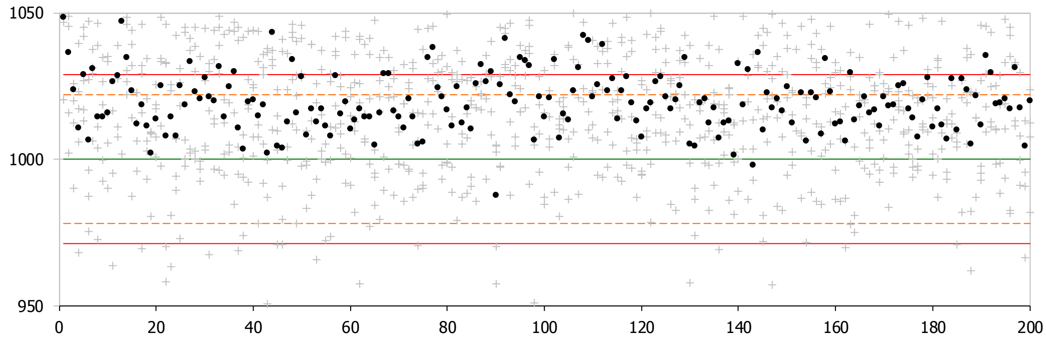
Control chart for mean values of samples of 5 items

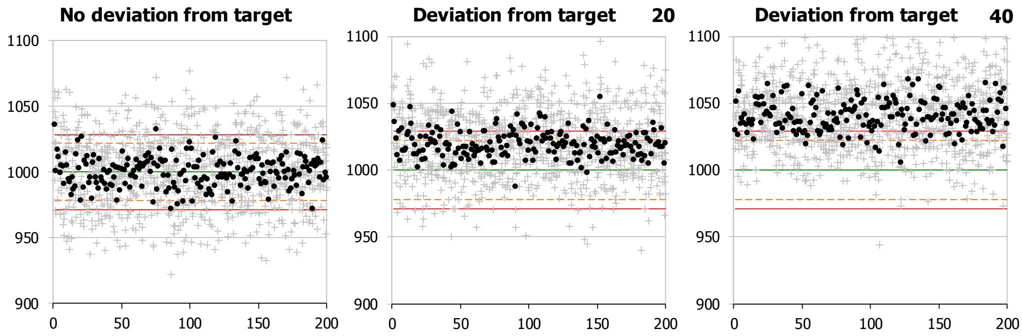


Production under deviation from regular conditions - deviation in mean of 20
Chart for mean values of 5 items

Mean values within	CL	WL
Number of samples	200	200
Samples within limits	165	127
Sample rate within limits	**0.825**	**0.635**

Deviation from target value 20





Deviation from target	0	20	40
SD of process	25	25	25
Samples	200	200	200
Within limits W	184	127	10
Within limits C	198	165	34
W alarms rate	0.08	0.365	0.95
C alarm rate	0.010	0.175	0.830

Figure 1. Behaviour of control charts for the mean of 5 data in 200 samples

Furthermore, one may use another applet that illustrates that the basic facts remain robust if the distribution of the population is changed to either highly skewed or scattered distributions (due to the Central Limit Theorem). Here, the lesson may teach that the mean of a sample does not vary much around the expected value of the population and its distribution tends to this value with increasing sample size. That is why we use larger samples to estimate the unknown mean of a population. We can extend our explorations (Figure 3) to statistical tests (judge the plausibility of various values for the population parameter) or confidence intervals (contain all plausible values of the parameter).

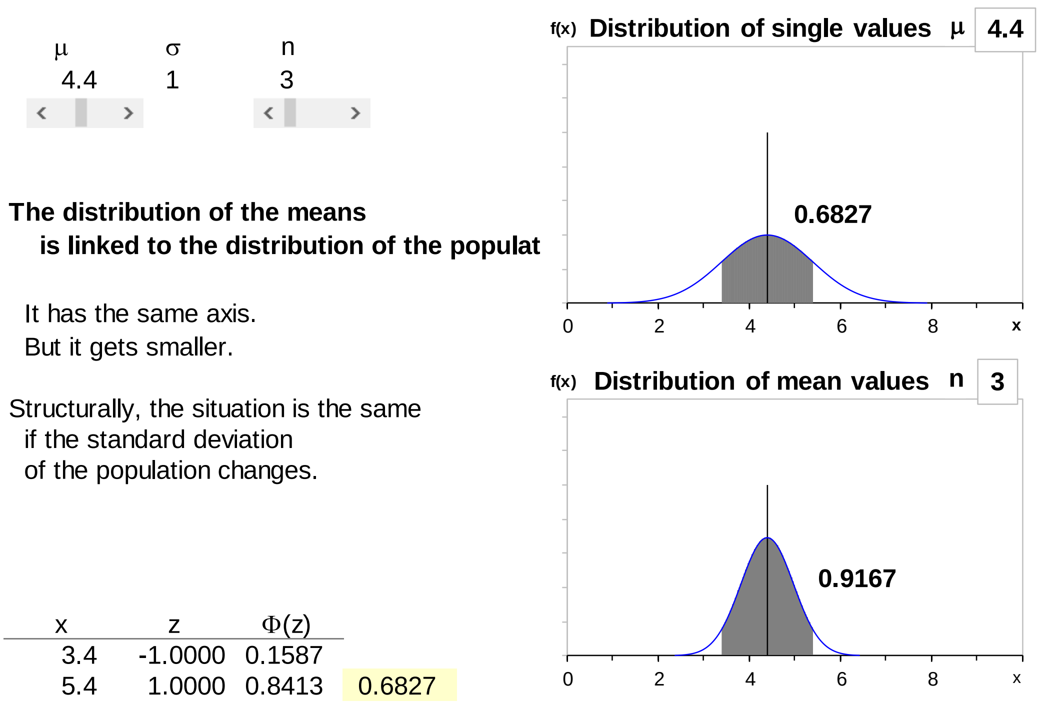


Figure 2. Link between the distribution of the mean and the population

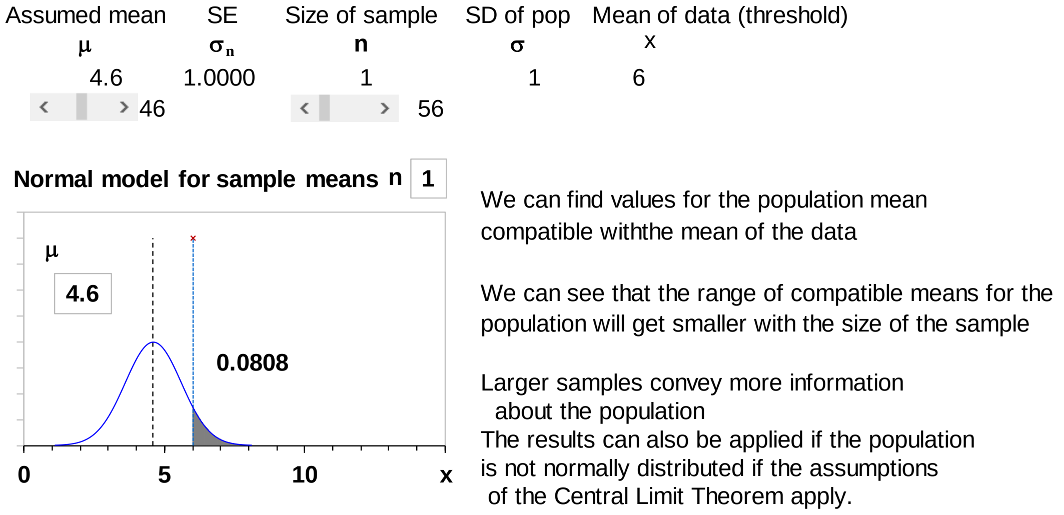


Figure 3. The mean of data compared to a changing population distribution

A finite version of the Central Limit Theorem

In the Central Limit Theorem – rather than increasing the number of random variables involved – the progressing process of the distribution to get a symmetric shape is illustrated for moderate sample sizes by the method of simulation (Figure 4). To repeat the whole simulation scenario shows that the shapes remain nearly the same from simulation to simulation. The rest of the heuristic argument is then an extrapolation of the observed phenomenon: if a highly skewed population leads to a fairly symmetric distribution for the mean of a random sample already for moderate sample sizes, this process of symmetrisation should continue and become "perfect" if we increase the sample size above all limits.

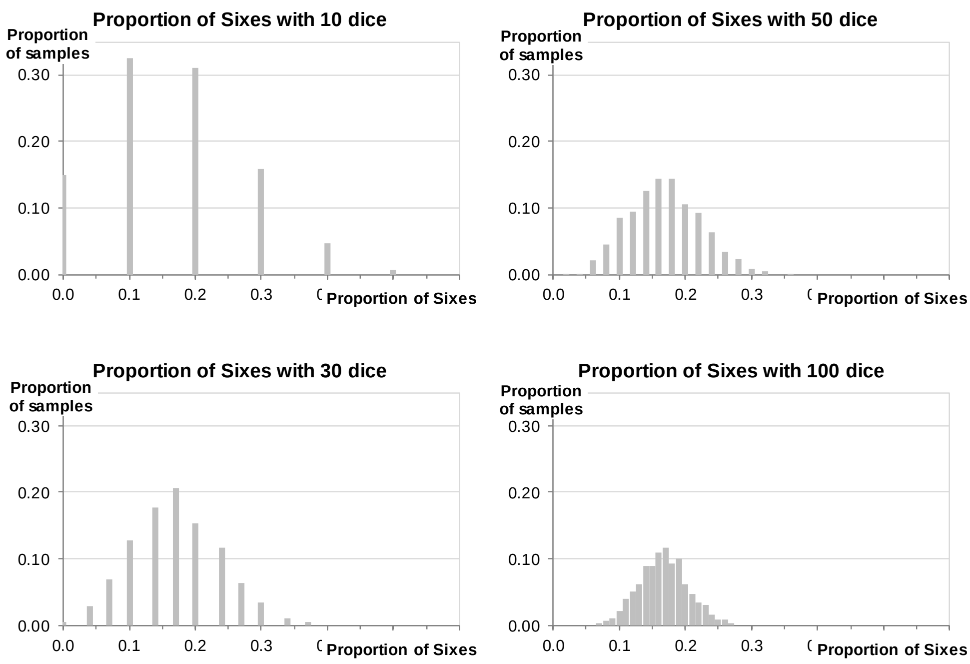


Figure 4. Normalising of the non-standardised proportion – CLT in action

Exemplary Task to illustrate our design of solutions for the students

We show one of the course tasks and the exemplary solution that we hand out to the students to promote their learning.

The task

In an artisan factory one wants to investigate whether production time is related to the achieved quality of the produced piece. In a test, 8 workers make a painting according to a template. They are not informed about the target of the investigation. The time to finish the painting is measured in minutes and the achieved quality is rated on a scale of 1 (very bad) to 20 (very good quality). The following data have been collected:

Worker no.	1	2	3	4	5	6	7	8
Time	55	51	64	63	64	78	66	69
Quality	15	14	12	14	17	20	12	15

a) Use an adequate statistical measure to answer the question whether there is a relation between time and achieved quality.

b) The question is also whether all data have to be included into consideration. Justify your approach. What are the consequences of your decision?

Exemplary solution for students' progress – basic requirements of a solution

a) We measure the strength of the co-relation by Spearman's rank correlation coefficient.

First we rank both data sets (separately within the related variable) and check whether the rankings in time needed and quality achieved are condordant. If we have ties in the data, we attribute mean ranks. The Spearman correlation yields $r_S(\text{ranks } r, \text{ranks } s) = 0.3110$;

the correlation is just above the threshold between low and medium positive correlation.

Worker	Original data		Ranks	
No.	Time	Quality	Time r_i	Quality s_i
2	51	14	8	5.5
1	55	15	7	3.5
4	63	14	6	5.5
3	64	12	4.5	7.5
5	64	17	4.5	2
7	66	12	3	7.5
8	69	15	2	3.5
6	78	20	1	1

b) A scatter plot of the data shows: there is no excess of points in quadrants I and III – the higher the rank in time (shorter time), the higher the rank in quality (the worse the quality). Removing the shaded point, we obtain a very low correlation $r_S = -0.0463$.

Looking back to the original data, one might see that only one worker (no. 6) has achieved a 20 in quality with an excessively long time to produce it. For all other workers, there seems hardly an influence of production time on the achieved quality.

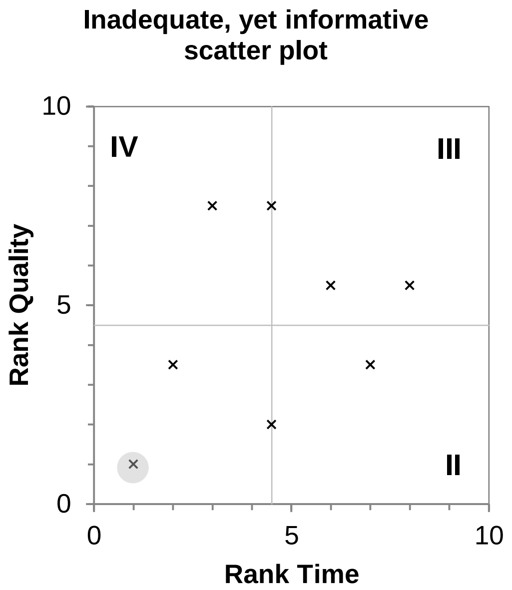


Figure 5. Inadequate, yet informative scatter plot to detect patterns

Extended remarks on the solution

- Time for production is measured on a metric scale but quality is measured only on an ordinal scale. Thus, we have to adapt our calculations to this lower scale. That implies we cannot use the Pearson correlation coefficient to measure the strength of the co-relation between the variables. That means also that we cannot use a regression model to describe the type of dependence between achieved quality and needed time.
- Rank 1 in time means longest production time (that might be confusing); rank 1 in quality means highest quality achieved by all workers (here 20).
- As there are ties, we cannot apply the Spearman formula with the squared differences in ranks. The more complicated formula for such cases yields the same value as the usual (Pearson) correlation (applied to the ranks rather than to the original data).
- The Excel function `Correl` is applied to ranked rather than original data.
- With quality, e.g., the 3rd and 4th ranks are tied; we determine a mean rank of 3.5. In sport competitions, rather than attributing mean ranks, we would have two third places (ranks).
- As quality is measured only on an ordinal scale, a scatter plot is not feasible. However, it provided essentially new insights into the situation. Often one has to be very flexible in data analysis.

CONCLUSIONS

There has been a marked endeavour of universities to build up e-learning environments. We chose one project (EMILeA) as an exemplar to illustrate the problematic of large-scale projects. Of course, this choice was influenced by the fact that EMILeA was greatly promoted by the publication in the *International Statistical Review*. The discussion about drawbacks of large-scale courses reveals that the reasons for its failure are generic and are embedded in its size.

The high expectations in the potential of e-learning have to be contrasted to the actual results and some critique from the community. If it can be realised with regards to financial and staff resources, blended-learning environments at smaller scale, locally organised, offer more possibilities for direct feedback from students and more incentives for interaction of teaching staff with the students.

We have elaborated a list of essential questions for a blended course, which are important for electronic learning environments but partially also apply for traditional face-to-face courses. One key element is to learn about the needs of students and staff and provide opportunities for such feedback also outside the tutorial classes. Our weekly questionnaires are very helpful for that purpose. We learn from them:

- Tops of our setting: Many examples, specimen solutions, manuscript coordinated with the textbook, permanent help with Excel, eTutorial.
- Still open issues: More orientation and links between the sources of information.

Orientation can be more easily adapted in smaller courses. Communication within staff helps to improve the processes. However, it is a real challenge to engage staff as our approach is work-intensive. Three final remarks: Education has to be adapted to the new challenges and technology. There are wider criteria than to reduce cost but cost considerations are ubiquitous and have a severe negative impact. One should not forget: Whatever system is introduced, its success or failure should be measured to the extent to which it supports that "Learning and teaching are interactive processes".

REFERENCES

[1] C. Batanero, M. Borovcnik, "Statistics and Probability in High School", Sense Publishers, Rotterdam (2016).

[2] M. Borovcnik, "The Influence of Software Support on Stochastics Courses", Edited M. I. Gomes, J. A. Pinto Martins and J. A. Silva, Bulletin 56. Session of the International Statistical Institute (ISI), ISI, Voorburg, **(2007)**, pp. 4602-4605.

[3] M. Borovcnik, "New Technologies Revolutionize the Applications of Statistics and its Teaching". Edited M. I. Gomes, J. A. Pinto Martins and J. A. Silva, Bulletin 56. Session of the International Statistical Institute (ISI), ISI, Voorburg, **(2007)**, pp. 823-826.

[4] M. Borovcnik, "Key Properties and Central Theorems in Probability and Statistics – Corroborated by Simulations and Animations", Selçuk Journal of Applied Mathematics. Special Issue of Statistics, **(2011)**, pp. 3-19.

[5] M. Borovcnik, R. Kapadia, "Modelling in Probability and Statistics – Key Ideas and Innovative Examples", Edited J. Maaß, J. O'Donoghue, Real-World Problems for Secondary School Students – Case Studies, Sense Publishers, Rotterdam (2011).

[6] B. Chaput, "An Organizational Memory as Support of Learning in Applied Mathematics", Edited M. I. Gomes, J. A. Pinto Martins and J. A. Silva, Bulletin 56. Session of the International Statistical Institute (ISI), ISI, Voorburg, **(2007)**, pp. 808-812.

[7] K. Cramer, U. Kamps, C. Zuckschwerdt, St Apps and EMILeA-Stat: "Interactive Visualizations in Descriptive statistics", Edited J. Antoch, Proceedings in Computational Statistics, Physica, Heidelberg, **(2004)**, pp. 101-112.

[8] C. Duller, „Einführung in die Statistik mit Excel und SPSS", Physica, Heidelberg, **(2007)**.

[9] EMILeA, „Eine multimediale, internetbasierte, interaktive Lehr- und Lernumgebung in der angewandten Statistik", Online: www.emilea.de/, **(n.d.)**.

[10] M. Esfandiari, M., H. Nguyen, Y. Yaglovskaya, R. Gould., "Enhancing Statistical Literacy through Short Open-Ended Questions that Involve Context, Data, and upper Level Thinking", Edited C. Reading, Data and Context in Statistical Education: Towards an Evidence-based Society, ISI, Voorburg, **(2010)**.

[11] GMW. Gesellschaft für Medien in der Wissenschaft, „Medidaprix", **(n.d.)**.

[12] M. I. Gomes, J. A. Pinto Martins and J. A. Silva, "Bulletin 56. Session of the International Statistical Institute (ISI)", ISI, Voorburg, **(2007)**.

[13] R. Gould, "Statistics Online Computational Resource (SOCR). Visualisation of the Normal Distribution", Online: socr.stat.ucla.edu/htmls/SOCR_Distributions.html, **(n.d.)**.

[14] W. Härdle, S. Klinke, and U. Ziegenhagen, "On the Utility of e-Learning in Statistics", International Statistical Review, vol 75, no. 3, **(2007)**, pp. 355–364.

[15] W. Härdle, S. Klinke, & U. Ziegenhagen, "e-Learning Statistics – A Selective Review", SFB 649 Discussion Paper 2006-024, Humboldt University, Berlin (**2006**).

[16] L. Johnson, S. Adams Becker, M. Cummins, V. Estrada, A. Freeman, A., C. Hall, "NMC Horizon Report: 2016 Higher Education Edition". Austin, Texas: The New Media Consortium. (**2016**), Online: cdn.nmc.org/media/2016-nmc-horizon-report-he-EN.pdf.

[17] D. Moore, S. Braun, J. Blatt, P. Girshman, J. Singer, J. Rubin, S. Stewart, D. Huntley, R.Buderi, "Against all Odds: Inside Statistics", (1995),
Online: www.learner.org/courses/againstallodds/unitpages/index.html.

[18] M. Mougeot, "Impact of New Technologies for Teaching Statistics". Edited M. I. Gomes, J. A. Pinto Martins and J. A. Silva, Bulletin 56. Session of the International Statistical Institute (ISI), ISI, Voorburg, **(2007)**, pp. 813-816.

[19] D. Nolan; D. Temple Lang, "Dynamic, Interactive Documents for Teaching Statistical Practice", International Statistical Review, vol. 75, no. 3, (**2007**), pp. 295–321.

[20] G. Schuyten and O. Thas, "Statistical Thinking in computer-Based Learning Environments", International Statistical Review, vol. 75, no. 3, (2007), pp. 365–371.

[21] H. Tamura, "Foundational Value of Statistics Education for Management Curriculum", International Statistical Review, vol. 75, no. 3, (**2007**), pp. 397–405.

[22] The Chronicle of Higher Education, "The Minds behind the MOOCs", Online: chronicle.com/article/The-Professors-Behind-the-MOOC/137905/?cid=at&utm_source=at&utm_medium=en#id=overview, **(n.d.)**.

[23] B.Ward, "The Best of both Worlds: A Hybrid Statistics Course", Journal of Statistics Education, vol 12, no. 3, **(2004)**, pp. 74–79.

[24] C. Wild, "Virtual Environments and the Acceleration of Experiential Learning", International Statistical Review, vol. 75, no. 3, **(2007)**, pp. 322–335.