



2nd International Conference on Higher Education Advances, HEAd'16, 21-23 June 2016,
València, Spain

Statistical training for a large number of students

Karl Moder^{a,*}

^a*Peter-Jordan-Strasse 82, A-1190 Vienna, Austria*

Abstract

At the University of Natural Resources and Life Sciences in Vienna the number of newly enrolled students increased from 696 to 3025 within the last 13 years, but the scientific staff of the statistical institute remained unchanged and consists of 4 people. About 60% of these students have to pass an introductory course in statistics at the Institute of Applied Statistics and Computing. So, depending on the branch of study, 200 to 500 students have to be supervised. As it is hardly possible to train such a high number of students, an online program was developed. It allows ongoing control of knowledge without excessively burdening the scientific staff. In addition, a program was developed, which creates test tasks for a final multiple choice test. These tasks are randomly selected from an appropriate database. Test results are automatically evaluated and graded. As this specific way of education is practiced for about 15 years, capabilities, advantages and disadvantages can be evaluated.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of the organizing committee of HEAd'16

Keywords: high number of students; online teaching; knowledge control.

1. Introduction

At the University of Natural Resources and Life Sciences in Vienna almost every student has to pass an introductory statistics course. About 60% of these students are supported by the Institute of Applied Statistics and Computing. So in 2015 about 1800 of new enrolled students and several others who did not pass the exam in previous years have to be supervised by the staff of the institutes. As the scientific staff of the institute consists of 4 permanently appointed persons only and additionally to basic statistical education several advanced statistical

* Corresponding author. Tel.: +43-1-47654-85112; fax: +43-1-47654-85109.

E-mail address: karl.moder@boku.ac.at

courses have to be given, introductory and some advanced courses are based on a combination of online and offline tools. The goal of this kind of instruction is on the one hand to keep the level of statistical knowledge high for students, on the other hand, to keep the burden of teaching low for the staff of the institute. The following description of the statistics course is based on lectures for students of "Landscape Architecture and Landscape Planning".

2. Audience of statistics course

Until 2013 students attend this course in their first semester of studies. About 350 (60% female, 40% male) were enrolled to the course. 25% of all participants abandoned the course within the first 5 weeks because they left university. The drop out rate for the remaining students was about 7% to 9% and rather stable over the last 10 years. Now this course takes place in the fifth semester. The number of supervised students by the author in 2015 was 230. Roughly 30% to 40% of students attend the lecture physically until the end of the semester. There is no obligation to attend the lecture because about 50% of students have a job beside their studies (The lecture takes two hours per week during the whole semester).

3. Short outline of statistics course

Ten years ago students had to pass 3 written tests and an oral exam. As this was very time consuming, an online program was developed for ongoing evaluation. Additionally an offline program creates multiple choice tests for the final examination and grading. Students have to complete 10 online tasks at home. Three of these tasks have to be completed by the use of a statistical package (SAS or R). Finally they have to pass a written test, consisting of 2 parts. Part 1 concerns statistical basics. In part 2 some data sets are provided and students have to determine which method to use to solve a certain statistical problem or they have to interpret results of certain statistical methods.

3.1. Data generation

Data generation for online tasks is done by Fortran programs. The basics of these data are real experiments which were conducted partly at our university. Based on the ACR-random number generator (Hörmann, 1990) data which are similar to these real experiments outcome is generated. So each student gets an individual data set with unique data values and data description. Besides data generation the program also creates corresponding solutions to these tasks. Data and solutions are stored in several tables of a MySQL-database.

3.2. Online-program

The online program is based on HTML and PHP – pages. To get access to data and tasks students have to register with their name and matriculation number. Then they have to complete 10 online tasks (so they do not need to be at the university physically). Each of these tasks is activated during a limited time slot (which corresponds to the actual content of the lecture). Within this period results can be entered into an input screen and are checked in respect to accuracy. The program displays whether results are right or wrong. Errors and mistakes can be corrected without any restriction within the set time window. As soon as students are satisfied, they have to store their results. For training purposes it is possible to enter a number between 1 and 100 as a matriculation number. Then students get access to free exercises without any time restriction. These exercises are very similar to those that must be solved obligatory. This feature offers a way to train skills for solving problems in connection with a specific online task. Approximately 10 % of students make use of this possibility at least once.

Figure 1 shows the start screen. If the time window is open for the program selected, it is possible to create a data set for a specific task or to enter results for this task. Without any time restriction one can control achieved scores or can practice his skills for the final test.

Family name:	Miller
First name:	Emily
Registration number	1040999
Choose exercise	t-test
	<input checked="" type="radio"/> Create a dataset <input type="radio"/> Enter results <input type="radio"/> Control score <input type="radio"/> Test preparation
© Karl Moder	Start

Fig.1.

If one chooses to create a data set or to enter results, a window is displayed which shows the time limits for this specific task. For t-test a typical data set look like this:

Data for EMILY MILLER (0740999)

Within 2 European countries selenium content (ppm) in soil was investigated.
 Try to verify if selenium content in the second country is higher than in the first ($\alpha = 0.001$).

country 1: (n=8)	41.8	38.6	47.1	36.0	45.6	41.3	45.1	48.0	
country 2: (n=9)	46.6	44.1	46.2	36.5	44.4	53.4	39.5	45.6	55.8

Fig.2. Example data set for a t-test.

Students have to calculate key figures like means, sum of squares, test statistics, probabilities, and to enter these results into the input screen of the program. Figure 3 shows the corresponding website.

t-Test

		Sample 1	Sample 2
Mean	\bar{x}	42.94	45.79
Sum of Squares	SS	125.6	289.629
Mean difference	\bar{d}	-2.85	
Standard error of difference	$s_{\bar{d}}$	2.6	
Degrees of freedom	df	15	H_0
Test statistic	t	1.096	<input checked="" type="radio"/> accepted
Prob	p	0.1458	<input type="radio"/> rejected

prerequisites

- independent samples
- dependent samples
- arbitrary distributed population
- normal distributed population
- variances arbitrary
- variances homogeneous

Input:

1. Use a dot as decimal separator.
2. Enter at least 4 significant digits (e.g.: 0.001234, 123.4, 1.234).
3. If your calculated prob value is smaller than 0.0001, then enter 0.0001.

Control results
Save results
Cancel

Fig. 3. Website to enter results for a specific task like t-test

The way students achieve their results cannot be controlled, so theoretically each available software can be used. Based on comments at the online forum, obviously most students calculate these key figures for the first 6 exercises either by hand or by the use of spreadsheet programs whereas the remaining 4 exercises are normally solved by means of statistical packages like SAS and R. These packages are taught and have to be used to solve exercise 7 to 9. In the follow the last exercise is often solved with one of these packages, too. Most statistical methods are based on assumptions about distributions; dependencies...These prerequisites are checked by the online program by means of a small multiple choice test (Figure 3).

Once data entry is complete, one can either check results or save the score obtained. If one selects “Save results”, the achieved score is stored in the database, but can be overwritten as long as the time window is open. It is also

possible to check if calculated values are valid without storing them into the database (“Control results”). In this case a table containing input data as well as a color code is generated, showing correct/incorrect results or answers. Green fields mean, that a value is correct, red fields indicate a failure in a specific input field. A characteristic picture similar to that of figure 4 will be displayed.

Control of Input:

Identifier	Value	r/w
\bar{x}_1	42.94	
\bar{x}_2	45.79	
SS_1	125.6	
SS_2	289.629	
\bar{d}	-2.85	
sd	2.6	
df	15	
t value	1.096	
Prob	0.1458	
H_0	accepted	
Dependency	independent	
Population	normal	
Variance	homogeneous	

You achieved **10** out of **13** points.

Your results are not saved!

To save, go back one page (Back key of the browser), and press "Save results"!

Fig.4.

At any time a students can check their achieved scores by selecting “Control score” at the start screen. In the follow a table which displays scores for solved tasks and the maximum attainable scores is generated (Figure 5).

Score for Emily Miller (Registration number: 1040999)		
program	your score	max. score
Steam & Leaf	20	20
Boxplot	11	11
t-test	10	13
One-way analysis of variance		14
Two-way analysis of variance		25
Paired t-test		9
Block design		20
Regression analysis		23
Confidence interval		11
Contingency table		9
sum	41	155
End		

Fig.5.

If “Test preparation” is selected from the start screen, then the program allows to practice for the final test as shown in Figure 6a to 6b.

Ozone concentration was measured on 4 consecutive days at 6 monitoring stations each.

Check if ozone concentration is the same at all stations ($\alpha = 0.05$).

	Tag			
	1	2	3	4
Laaer Berg	60.2	60.0	48.5	65.2
Währinger Gürtel	43.9	39.7	35.5	47.3
Hietzinger Kai	52.1	53.6	60.0	49.7
Lobau	57.3	58.4	53.9	56.4
Taborstraße	54.6	50.0	40.8	56.7
Stephansdom	59.2	58.3	59.5	57.4

Multiple answers are possible	Test situation
<input type="checkbox"/> t-test <input type="checkbox"/> One-way analysis of variance <input type="checkbox"/> Two-way analysis of variance <input type="checkbox"/> Paired t-test <input checked="" type="checkbox"/> Block design <input type="checkbox"/> Regression analysis <input type="checkbox"/> Contingency table	<input type="radio"/> one sided test <input checked="" type="radio"/> two sided test
Next question	End

Fig.6a

Question number	Score	Maximum score	Answer (last question)	Hypothesis (last question)
5	8	10	correct	correct
		Correct answer:	Block design	two sided

Fig.6b. Practising for the final test - correct answer for the previous question, achieved and maximal score.

As soon as a student answers at least one of the presented questions a table like that in figure 6b is displayed. It shows the number of correctly answered questions as well as the last answer, gained score (e.g. 8) and maximum attainable (e.g. 10) score for all questions up to the last one (e.g. question 5).

3.3. Final test

To be allowed to take the test, one has to complete each of these 10 tasks with at least 50% of the maximum attainable score. Additionally 3 of these tasks have to be solved by using a statistical package like SAS or R.

The final test is designed as a multiple choice test. It is subdivided into 2 parts. In the first part students have to answer several theoretical questions (like prerequisites of tests, characteristics of distributions, ...). The second part checks if students are able to find appropriate methods which can be applied to a given dataset (as can be trained with the online program) or they are asked for a correct interpretation of a given analysis. A question concerning interpretation of a Tukey-test looks like this:

Question 1:

Different varieties of wine were tested in regard to acid content (gram per litre). Results based on a Bonferroni-Holm-Test are shown beneath:

Traminer	Grüner Veltliner	Blauburgunder Rose	Welschriesling	Riesling
4.3	5.8	6.4	7.0	7.7

What are correct answers?

- a) There is nor difference between Grüner Veltliner and Blauburgunder Rose in regard to acid content.
- b) Grüner Veltliner and Welschriesling show different acid contents.
- c) Blauburgunder Rose and Welschriesling show different acid contents.
- d) Riesling is different to all other wines in regard to acid content.
- e) Traminer is different to all other wines in regard to acid content.

Questionnaire boxes according to these questions are located on a different sheet and identified by the question number and the according letters of answers.

4. Grading

Grading is based on scores of the online program. At least 50% of the maximum score for each task has to be achieved. Almost all students reach at least 90% of the total score. Three of these tasks have to be accomplished by means of SAS or R. Again 50% of the maximum score has to be achieved. If these requirements are not fulfilled students are not allowed to take part in the final test. For each part of the final test 50% of the maximum score is necessary to get a positive mark. These offline surveys are scanned by a scanner and evaluated by an appropriate program. Based on all results a mark is calculated in the follow.

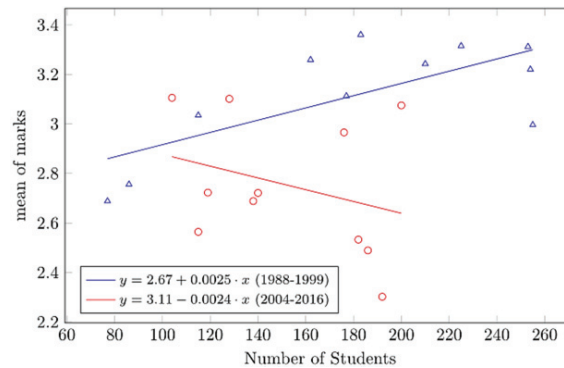


Fig. 7. Marks in dependence of group size

Figure 7 shows mean marks in dependence of the number of students who finished the course (This is less than the number of enrolled students, as not all take part in the final exam immediately at the end of the lecture).

Blue triangles and blue regression line refer to the situation before using the online program, (years 1988 to 1999) red dots and red regression line are based on results since the implementation of this program (years 2004 to 2016). Only those results were used where more than 70 students attended the lecture. Results became significantly worse for increasing group sizes before the online program was used. Since it was used mean grading seems to decrease with a higher number of students, but regression is not significant for this situation.

Figure 8 shows the distributions of marks before (blue bars) and after (red bars) implementation of the online program.

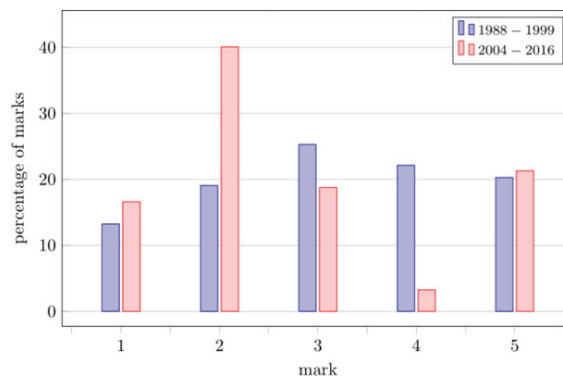


Fig. 8. Distribution of marks before and after implementation of the online program

Obviously distributions differ significantly. The percentage of students which positively/negatively complete the course is very similar, but the distribution of positive marks is quite different for both situations. Whereas marks before the online program (1988-2000) are nearly symmetric distributed, this is not the case after implementing of the program. Data are skewed. This is because results of the online program are taken into account at grading if

students pass the final test. So most of them are motivated to complete this online tasks with a very high score and use the opportunity of correcting erroneous results within the time period. The number of those who do not pass the course is almost identical as one has to reach at least 50 % of the maximum score at the final multiple choice test to get a positive mark.

5. Discussion of results

Implementation of the online program and the use of multiple choice tests for grading was motivated by the high number of students in basic statistical teaching and the enormous amount of time to supervise them. The goal was to reduce this time exposure. This goal was achieved to a 100%. The time needed for correcting tests is very low, and does not depend on the number of students, as this is done either by the online program or by automatically evaluating multiple choice tests. The proportion of students who pass the exam to those who do not, did not change by using this new kind of teaching, but distribution of positive marks changed significantly and are shifted to better grading. Online surveys of the course are really positive. So almost every year the online program is regarded to be a helpful tool for understanding statistical methods, especially as it allows to correct errors. 77% of students believe that the program encourages to participate actively in the class (11% are of a different opinion). 70% think, that they are well taken care of and well supported by the course (10% think that they are not). Based on entries in the accompanying online-forum the program seems to motivate students to cooperate, as there are several questions which indicate this cooperation (E.g. I had a problem with the F- value and some of my colleges had just the same. Is this a failure in the program?). As mentioned above students are allowed to enter results without any restriction within a given time period. The average number of storing results lies between 1.2 to 2.0 depending on the task, but there are up to 20 individual attempts for certain students. Failure is instructive (Dewey) and so it seems that students are motivated to deal with a problem intensively.

6. Conclusions

Before this program started 3 written tests had to be positive completed. In 2000 the online program was implemented and started in one branch of study (Landscape Architecture and Landscape Planning). Gradually at intervals of 2 years, all branches of studies, which are managed by the Institute of Statistics, have been converted to this system (which can be used from inside the elearn-platform (ACM, 2014)). This permits a comparison between these different systems of teaching. The main results are summarized in the follow.

- Distribution of scores for students who pass the course positively is skewed for the online system in compare to the old one (but the number of students which did not pass remains constant).
- Students are motivated to cooperate.
- Evaluation showed that students like this online system as it is not necessary to come to university (this is especially important, as more than 50% of all students have a job in addition to their study).
- Within the given time window students are free to manage solution of these online tasks.
- Time required for statistics training is low for the staff of the institute and does not depend on the number of students.
- An accompanying seminar (which is conducted by 3 tutors) where all of these online tasks tasks are explained by means of several examples is regarded to be very helpful by students in this respect. Additionally there is an online forum to support students with their questions regarding tasks and statistical problems.

Overall, the experiences with this program and kind of teaching are very positive. In the meantime other institutes of our university use this program in a somehow modified form. The original motivation for implementing this program was to keep the burden of supporting a high number of students low, but based on experiences with the program it is planed to use it with other courses, too (eg. multivariate statistics, design of experiments).

References

- Hörmann, W. & Derflinger G. (1989). The ACR Method for Generating Random Variables. *OR Spectrum* 12, 181-185.
- ACM (2014). <https://www.academic-moodle-cooperation.org>.
- Dewey, John. BrainyQuote.com, Xplore Inc, 2016. <http://www.brainyquote.com/quotes/quotes/j/johndewey121337.html>.