SFB 649 Discussion Paper 2008-058

Statistics E-learning Platforms Evaluation: Case Study

Taleb Ahmad*
Wolfgang Härdle*

* Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

Statistics E-learning Platforms Evaluation: Case Study *

Taleb Ahmad, Wolfgang Härdle

CASE - Center for Applied Statistics and Economics

Humboldt-Universität zu Berlin

Spandauerstrasse 1, 10178 Berlin, Germany

August 31, 2008

Abstract

With the increase of e-learning by universities and educational institutes in the world through more electronic platforms, come the questions to researchers, educators and designers of electronic platforms about feasibility and using this method of learning. Are we achieving the desired goals and improving the quality of education? Are we improving their performance and ability to self-study without the need for a teacher? Is e-learning an effective and successful method from the students views? In this paper, we consider evaluate e-learning systems in statistics. We make an evaluation study, we analyze a students sample of the methods: Factor analysis, Logit model. The common aim of this evaluation is to provide data to justify the results or evidence to support that the e-learning platforms are helping the students to learn more effectively. The questionnaire covers information about e-learning evaluation criterias. Some of these criterias are: Navigability, applicability, instructional structure and interactivity.

Keywords: E-learning, Evaluation, Statistical software

JEL Classification I21, C19

*This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 'Economic Risk'. Corresponding author: Taleb Ahmad, CASE - Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Spandauer Str. 1, 10178 Berlin, Germany, email: tahmad@wiwi.hu-berlin.de

1 Introduction

"The evaluation of e-learning systems is important for all the actors involved in their development and use. Teachers and students need to evaluate the benefits of using e-learning in comparison with the classical methods of learning" (Posea, Matu and Cristea, 2007). "Evaluation is the interpretation of the data from the assessment in an institutional setting: The evaluators may be students, faculty and administrators. The results of an assessment process should provide information which can be used to determine whether or not intended outcomes are being achieved and how the e-learning material can be improved" (Falco and Soeiro, 2003).

With increasing growth in size e-learning applications, the demand on the technology is becoming more rigorous. This involves new ways to access, learn and prepare learning materials. There is need to find whether the technology infrastructure has the capacity to support the users and network load, or scalable enough to support growth. A good proof to improving any e-learning material is to have a clear description of the learning needs (Mungo, 2004). Some useful e-learning evaluation criteria include:

- Navigation: Navigation means: The ease with which students learn, and the ability to find their way within the learning package. The progress within course material, the necessary forward, backward option choices such as to skip ahead and go backwards to previously covered material.

- Instructional Structure: E-learning materials must relate to the reading level, depth and experience of the target learner. It also should include an introduction on the subject to be learned and its importance in the learning process (Mungo, 2004).

- Interactivity: Is seen as part of a system where learners are not passive recipients of information, but engage with material that is responsive to their actions. E-learning that merely allows the learner to navigate content or take an online test is often labelled as interactive (Thomas, 2001). There are identifiable types of interactivity like: Learner to content, learner to instructor, learner to computer (software/interface), learner to learner. This criteria focus on user participation in the learning process through interactive examples. Interaction is based on the knowledge and skill of the learner. The type of interaction may include simulations, free response, executing inherent software applications, drag and drop.

- Applicability: Applicability for e-learning evaluation involves how applicable the instructional content is to the specific need and situation the learner faces. It looks for a strong connection between the course content and how the learner benefits from the learning process (Mungo, 2004).

There are several electronic platforms at the institute of statistics-Humboldt University-Berlin, such:

- MM*Stat: http://www.quantlet.com/mdstat/products.html
- e-stat: http://www.e-stat.de
- Electronic books: http://mars.wiwi.hu-berlin.de/ebooks/html/
- Moodle: http://lms.hu-berlin.de/moodle/
- teachwiki: http://teachwiki.wiwi.hu-berlin.de

In this paper we try to show the effect of e-learning platforms to learn statistics through the student‘s views. The study was performed by students in the faculty of economics-Humboldt University- Berlin. This sample was drawn from statistics II course for the winter semester 2007- 2008. A total of 208 students took part in the study.

We have formulated the questionnaire to answer some of the following questions: Do e-learning platforms help the student to improve their study and their understanding the statistics course? Is e-learning a successful method for self-study? Can we always use e-learning platforms in statistics course?

2 The Model: Selecting the dependent and independent variables

The goal of the research here to examine the effect on the dependent variable: “Using e-learning platforms in statistics course”, through the independent variables: Study course, gender, previous experience with e-learning, understanding statistics, study level, structured format, self-study, the flexibility and freedom in dealing with the course, interactive environment and e-learning problems.

2.1 The dependent variable

The dependent variable: The decision to use e-learning platforms in a statistics course is a binary variable with two categories: Using or not using. The question was: Do you think that e-learning platforms should always be used in traditional courses? The answers were: a) yes, b) partial, c) no, d) I do not know. In order to satisfy the conditions of Logit Model, the answers are classified into two groups: Y = 1: Yes, i agree to using e-learning platforms in statistics course; b) Y = 0: No, I do not use e-learning platforms. The answers "yes" or "partial" are classified into the category: "I use" and the answer "I do not know" in whole sample 3.3% is defined as missing value. The valid percent in the sample for the decision: Using e-learning platforms in statistics course is $78,8\%$ and not using 21.2%.

2.2 The independent variables

Ten variables are chosen from the survey as independent variables. Basic descriptions to the characteristic of each variable are shown as followed:

- Gender: 59.6% of the valid responses, which were 124 observations, are female and 40.4% of the valid responses, which were 84 observations, are male.

- Study course: The study course was divided into three kinds. Betriebswirtschaftslehre (BWL) being 110 students $52,9\%$. Volkswirtschaftslehre (VWL) students being 60 students $28,8\%$ and other courses which were 34 students $16,3\%$ and the missing values stood at 4 students $1,9\%$.

- Previous experience: The question was to students: Have you had any experience before with e-learning? This variable is re-classified into 2 categories "yes, i had" and "no, i did not have". The sub-category i had an experience with e-learning has 108 observations equal to 52.4% of total sample and the answer: I did not have any experience has 98 observations equal to 47.6% of the whole sample. 206 observations are valid for this variable.

- Flexibility: Is required for students to access the system at any point, to make their own way through the multimedia tool and to review the content at their own pace (Aydinli, Härdle and Rönz, 2003). Flexibility is a major benefit of e-learning. E-learning has to take place anytime anywhere. The question was to students: "Do you think that e-learning platforms give you the freedom and flexibility to work with the course? For example MM*Stat or moodle", this variable is re-classified into 4 categories: Yes, partial, no, i do not know. The sub-category "yes" has 104 observations which is 50% of the total sample, and the category "partial" has 50 observations equal to 24% of the total sample, the category "no" has 34 observations equal to 16.3% and the last category "I do not know" which has 20 observations equal to 9.6% of the sample.

- Understanding statistics: The question was to students: Do you think that using e-learning platforms helped you understand the statistics topics? This variable is re-classified into 4 categories: Yes, partial, no, i do not know. The category "yes" has 60 students which is 28.8% of the total sample, and the category "partial" 84 observations equal to 40.4% of the sample, the category "no" has 38 observations equal to 18.3% of the sample, and the category "I do not know" which has 26 students equal to 12.5% of the sample.

- Structured format: A clear structured format in presenting the statistical content must be maintained throughout the whole system. A well-designed e-learning platform can particularly make the learning process easier by allowing the students to develop their insights without getting bogged down in the mathematics (Aydinli, Härdle and Rönz, 2003). The question was to the students: What is your opinion about the structured format of e-learning platform which you have used? This variable is re-classified into 4 categories: Good, acceptable, bad, i do not know. The category "good" has 78 students which is 37.5% of the sample, and the category "acceptable" has 84 students equal to 40.4% of the sample, the category "bad" has 28 students equal to 13.5%. The category "I do not know" which has 16 students 9.6% of the sample, 206 observations are valid for this variable.

- Self-study: E-learning platforms offer the possibility to students to learn alone. Sometimes the students did not need the help and support from their teacher. Every thing is available, e-learning platforms must offer self-assessment components with automatic evaluation of the answers to give the students the opportunity of checking their acquired knowledge. The question was to students: Do you think that e-learning platforms improves and increases your self-study? This variable is re-classified into 4 categories: Yes, partial, no, i do not know. The category “i do not know” which has 8 students 3.8%. The category “yes” has 132 students which is 63.5% of the valid sample, and the category “partial” 58 students equal to 27.9% of the valid sample, the last category “no” has 4 students equal to 1.9%.

- Study level: The question was to the students: Do you think that e-learning platforms improves your study level and helps you to pass the examinations? This variable is re-classified into 4 categories: Yes, partial, no, i do not know. The category “yes” has 58 observations which is $28,2\%$ of the valid sample, and the category “partial” 66 observations equal to 32% of the valid sample, the category “no” has 42 observations equal to 20.4% of the valid sample. The category “I do not know” which has 40 observations 19.4%.

- Interactive environment: The question was to the students: What are your opinions about the interactive environment of e-learning platforms? This variable is re-classified into 4 categories: Good, acceptable, bad, i do not know. The category “good” has 144 observations which is $69,2\%$ of the valid sample, and the category “acceptable” 50 observations equal to 24% of the valid sample, the category “bad” has 6 observations equal to 2.9%. The category “I do not know” has 8 observations 3.8%.

- E-learning problems: This variable describes the problems during application e-learning platforms from the student‘s views. The question was to students: What are the important problems that are associated with e-learning platforms? This variable is re-classified into 4 categories: High costs, no presence of teachers, increased work time and other reasons. The sub-category “high costs” has 34 students which is 20.5% of the valid sample, and the category “no presence of teachers” 68 students equal to 41% of the valid sample, the category “increased work time” has 28 students equal to 16.9%. The category “other reasons” which has 36 students 21.7% of the valid sample, 166 observations are valid for this variable.

3 Applied statistical methods

3.1 Factor analysis

Factor analysis is a statistical data reduction technique. This method is used to find the factors in a group that consists of a large number of observed variables. The observed variables express as a functions of unobserved random variables called factors, or other words: The observed variables are modeled as linear combinations of the factors, plus "error" terms or not. The relation between the variables inside a single factor is stronger than the relation with the variables within other factors. Factor analysis helps us to understand the structure of correlation matrix through a few factors. We make exploratory factor analysis when we have no idea how the model looks like. The output equation to the model of explorative factor analysis is:

$$Z_{(n*m)} = F_{(n*Q)} A^{\top}_{(Q*m)} + U_{(n*m)} E_{(m*m)} \tag{3.1}$$

- Z Matrix of standardized variables z_j.
- F Matrix of factor values f_{iq} of each factor F_q for every case i with every Variables Z_j.
- A Matrix of factor loadings.
- U and E: Similar to the idea in linear regression model we include noise ϵ_j for every variable, we have the next form:

  $$\epsilon_j = e_j * U_j$$

  This noise indicates to the variance which are not explain by the common factors.

This equation is not solvable, because it contains many unknown parameters but step after step, the individual matrices based on the correlation matrix will estimate. This means we can calculate the correlation matrix R. The values are on the main diagonal:

$$r_{jj} = Cov(Z_j, Z_j) = Var(Z_j) = 1$$

The values besides the main diagonal are Bravais Pearson‘s correlation coefficients between the variables. The correlation matrix is the starting point for factor analysis. The common factors exist only for variables that are highly correlated with each other. The low correlations with other variables stay maybe unconsidered in the factor. The following conditions in the data are available to be able to apply factor analysis:

- Metric scaled variables.
- Independent observations.
- Approximate normal variables.
- Large samples size.

There are several methods to find the factors such as: Principal component analysis, maximum-likelihood method. In this analysis will use principal component analysis.

3.1.1 Kaiser-Meyer-Olkin measure and Bartlett's test of sphericity

Before using factor analysis method, we should first test if the sample is suitable for the analysis. We will apply Kaiser-Meyer-Olkin test. The Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) is a measure compares the observed correlation coefficients to the partial correlation coefficients, it is calculated by the next form:

$$KMO = \frac{\sum\sum r_{ij}^2}{\sum\sum r_{ij}^2 + \sum\sum a_{ij}^2} \tag{3.2}$$

r_{ij} is the observed correlation coefficients.

a_{ij} is the corresponding partial correlation coefficients.

KMO takes the values in the interval [0, 1]. If KMO value is near 1, the partial correlation coefficients are small. When the value of KMO is small, it means that factor analysis of the variables are not meaningful, in other words: If KMO value is bigger than 0.5, the implementation factor analysis is suitable with the given data. The Bartlett test checks the null hypothesis that the variables in the population correlation matrix are uncorrelated.

3.1.2 Anti-Image matrix

This matrix includes outside the diagonal (off-diagonal) the negative partial correlation coefficients between two variables. If we want to get a good factor model, the off-diagonal elements in Anti Image matrix should be small. The diagonal elements on anti-image matrix for each variable are a measure of sampling adequacy (MSA), in same time: MSA has the same meaning with KMO. We calculated MSA value for every variable using the next form:

$$MSA_i = \frac{\sum\sum r_{ij}^2}{\sum\sum r_{ij}^2 + \sum\sum a_{ij}^2} \tag{3.3}$$

As in KMO test, if MSA-value for each variable is bigger than 0,5 then these variables should stay in factor model.

3.2 Logit model

Logistic regression is a technique for analyzing problems in which there are one or more independent variables that determine an outcome. The outcome is measured with a dichotomous variable (in which there are only two possible outcomes). In logistic regression, the dependent variable is binary or dichotomous, it only contains data coded as 1 (True, success, pregnant, etc.) or 0 (False, failure, non-pregnant, etc.), http://www.medcalc.be/manual/logistic-regression.php.

We have a set of independent variables x_i that are included in the model. The dependent variable y can be described as a linear combination of the independent variables x_i and the parameters β plus the error term ε_i as in form 3.4

$$y_i = x_i^\top \beta + \varepsilon_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_j x_{ij} + \varepsilon_i \tag{3.4}$$

The linear regression consists of two parts: The mean value of the outcome variable that can be expressed as a linear function of the independent (predictor) variable and the error that attempts to describe how individual measurements vary around the mean value (Guan, 2006). This model can be expressed as:

Structure on the means:

$$E(Y_i|X_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_j x_{ij} \tag{3.5}$$

Error structure:

$$\varepsilon_i \sim N(0, \sigma^2) \tag{3.6}$$

.

There is a problem with the application of such models such as in eg; 3.5, for a dependent variable is a dichotom, however the variable Y can not take only the values from 0 to 1, but accept the values also between $-\infty$ and $+\infty$, a solution to this problem is the logistic regression.

The quantity $\pi_i = E(Y_i|X_i)$ is used for logistic distribution in order to simplify notation, and we have the form:

$$\pi_i = F(y_i) = \frac{e^{y_i}}{1 + e^{y_i}} \tag{3.7}$$

with:

$$y_i = x_i^\top \beta = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_j x_{ij}$$

The relationship occuring between the probability and the x-variable can then use a logistic distribution. The logistic distribution can only accept the values from 0 to 1. We get from 3.7:

$$e^{y_i} = \frac{\pi_i}{1 - \pi_i} \tag{3.8}$$

First the ratio between π_i and $1 - \pi_i$ is considered. This ratio denotes as “Odds”, this Odds accept the values between 0 and + ∞.

$$Odds(Y = 1) = \frac{\pi_i}{1 - \pi_i} \tag{3.9}$$

If we apply the logarithms of $\frac{\pi_i}{1-\pi_i}$ we get logit or log-odds.

$$Y_i = ln(\frac{\pi_i}{1 - \pi_i}) = X_i^\top \beta = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_j x_{ij} \tag{3.10}$$

With this transformation we find that the relationship between the variables X and the probability of the event in parameter will be linear. This expression may accept the values between $-\infty$ and $+\infty$.

“The Logit is also a link function, because it gives a connection (link) between the dichotomous variables and the linear regression expression on the right side of the equation” (Boyum, 2006).

As in the case of a linear regression the parameters will estimate, but in this case we use the maximum likelihood method, because the least squares method is not suitable and bring many statistical problems for logit model (Hosmer and Lemeshow, 1989). The estimated parameters indicated how the Logits change when the independent variables increase by one unit.

4 Estimation the model and the results

4.1 Results of exploratory factor analysis

The result of the exploratory factor analysis for binary data using SPSS programm is explained in this section. We have done factor analysis using the method of principal components. We found on the diagonal anti-image correlation matrix in table 1: Measure of sampling adequacy, MSA-values for the variables: Gender, previous experience, study course and e-learning problems are smaller than 0,5 therefore we excluded these variables from factor analysis and we continued our analysis with the rest variables (six variables).

Variable	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
Study course	,31[a]	-,28	,06	,04	,05	,13	-,24	-,05	-,06	,03
Structured format	-,28	,61[a]	-,19	-,22	,31	-,15	,19	-,13	,04	,19
Interactive environment	,06	-,19	,68[a]	,08	,04	,02	,04	-,08	-,01	,02
Previous experience	,04	-,22	,08	,40[a]	-,14	,03	-,29	,07	-,15	,24
Understanding statistics	,05	,31	,04	-,14	,61[a]	-,03	,16	-,25	,09	,04
Flexibility	,13	-,15	,02	,03	-,03	,68[a]	,06	-,08	,24	,23
E-learning problems	-,05	-,13	-,08	,07	-,25	-,08	,35[a]	-,12	-,06	,03
Self-study	-,07	-,13	-,47	-,07	-,15	-,33	-,24	,73[a]	-,13	,11
Study level	-,02	-,43	,13	,17	-,36	,06	,11	,01	,68[a]	,05
Gender	-,06	,04	-,01	-,15	,09	,24	,21	-,06	,32	,23[a]

Table 1: Anti-Image correlation matrix.

Table 2 shows Bravais-Pearson correlation between the rest variables. The figure 1 presents scree plot that shows two eigenvalues are bigger than one.

Variable	V1	V2	V3	V4	V5	V6
Understanding statistics	1,0	,09	,04	,14	,34	,47
Structured format	,09	1,0	,38	,38	,56	,40
Interactive environment	,04	,38	1,0	,31	,35	,09
Flexibility	,14	,38	,31	1,0	,44	,22
Self-study	,34	,56	,35	,44	1,0	,39
Study level	,47	,40	,09	,22	,39	1,0

Table 2: Correlation matrix.

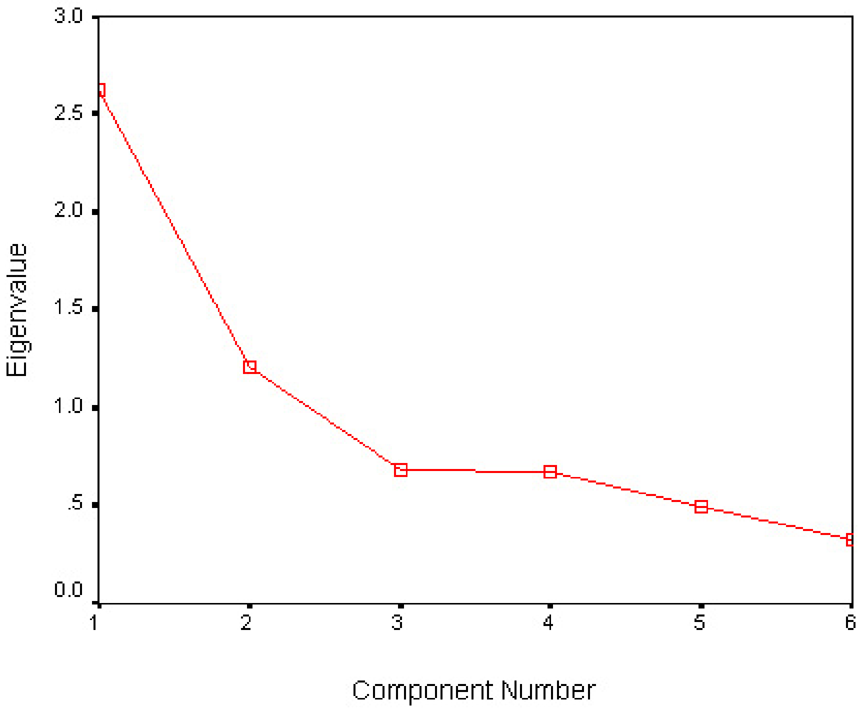

Figure 1: Scree plot.

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative%	Total	% of Variance	Cumulative%
1	2,62	43,68	43,68	2,62	43,68	43,68
2	1,20	20,13	63,81	1,20	20,13	63,81
3	,679	11,32	75,13			
4	,671	11,18	86,31			
5	,492	8,19	94,51			
6	,329	5,48	100,00			

Table 3: Total variance explained.

Two factors are extracted from six variables. The first factor explained 43.68% of the total variance and the second factor explained $20,13\%$ as in table 3 (Total variance explained). Two factors explain $63,81\%$ of total variance. The table 4 presents the communalities for every variable, the communalities mean proportion of variance in each variable explained by the two factors, for example: The communalities for variable “Understanding statistics” is 0,73, that means the two factors explain 73% from the variance of understanding statistics, it means square multiple correlation of variable with the factors, the value of communalities is between 0 and 1. In our analysis the value of communalities lies from 0,49 to 0,73.

Variable	Initial	Extraction
Understanding statistics	1,00	,73
Structured format	1,00	,64
Interactive environment	1,00	,57
Flexibility	1,00	,49
Self-study	1,00	,67
Study level	1,00	,70

Table 4: Communalities.

Variable	Component	
	1	2
Understanding statistics	,03	,70
Structured format	,76	-,25
Interactive environment	,53	-,33
Flexibility	,64	-,27
Self-study	,82	-,03
Study level	,05	,65

Table 5: Rotated component matrix.

The table 5 presents the rotated component matrix. The relationship between the variables and the common factors are described through factor loadings. The stronger variable correlated to the first factor is a “self study” variable, where loading this variable with the first factor is 0,82, then “structured format” variable 0,76 and the stronger variable correlated to the second factor is “Understanding statistics” 0,70, then the variable “Study level” 0,65.

Now we have two extracted factors:

Factor 1: Self study, structured format, flexibility and interactive environment.

Factor 2: Understanding statistics, study level.

In table 6: KMO value is 0,70, this value is suitable for doing factor analysis. On the other hand the significance level of the Bartlett’s test of sphericity is .000. We can reject the null hypothesis that the variables in the correlation matrix are uncorrelated.

The test	Value
KMO Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	,70
Bartlett’s Test of Sphericity Approx. Chi-Square	285,34
df	15,00
Sig.	,000

Table 6: KMO and Bartlett’s test.

Now come the question if there is a relationship between factor variables and the dependent variable "decision" (using or not using e-learning platforms). The independent-samples t-test for the comparison between "using" and "not using" is used, because the both factors are continuous and the dependent variable is dichotomous. From table 7: The difference between using and not using as shown in factor 1 is significant at the 5% level. Factor 2 is not significant at the 5% level. The "one way ANOVA" is used to test the relation between factor variables and the dependent variable. The table 8 shows the difference between using and not using e-learning platforms as shown in factor 1. This is significant at the 5% level, and to factor 2 is not significant at the 5% level.

factor score	Assumption	Levenes Test		T-test for Equality of Means		
		F	Sig.	t	df	Sig
factorscore 1	Equal variances assumed	6,00	,01	-3,95	198	,00
	Equal variances not assumed			-3,44	51,76	,00
factorscore 2	Equal variances assumed	3,86	,05	,92	198	,36
	Equal variances not assumed			,79	51,52	,43

Table 7: Independent samples test by decision.

Factor variable	Assumption	Sum of Squares	df	Mean Square	F	Sig.
	Between Groups	14,59	1	14,59	15,67	,000
factorscore1	Within Groups	184,4	198	,93		
	Total	199,00	199			
	Between Groups	,84	1	,84	,84	,36
factorscore2	Within Groups	198,16	198	1,00		
	Total	199,00	199			

Table 8: ANOVA by the decision.

Now we will check both factors as normal distribution, then we will do Kolmogorov-Smirnov test for both factors. The null hypothesis for this test ist: The distribution of both factors are normal. The p-value for the first factor is 0,16 as in table 9, then we can not reject that the distribution of the first factor is normal and the p-value for the second factor is 0,04, this factor is not normally distributed.

Value		Factor 1	Factor 2
N		162	162
Normal parameters	Mean	8.82	3.93
	Std. Deviation	1,00	1,00
Most Extreme	Absolute	,088	,139
Differences	Positive	,088	,139
	Negative	-,079	-,063
Kolmogorov-Smirnov Z		1,11	1,76
Asymp.Sig.(2 tailed)		,16	,04

Table 9: One-Sample Kolmogorov-Smirnov test for both factor.

4.2 Estimate and fit of the model using logit Model

We test now the statistical hypothesis to determine whether the independent variables included in the model are significantly associated with the dependent variable (Decision). There are two forms of stepwise logistic regression: Forward inclusion and backward elimination. We use backward elimination to estimate the parameters. Backward stepwise regression appears to be the preferred method of exploratory analyses, where the analysis begins with a full model and variables are eliminated from the model in an iterative process. The fit of the model is tested after the elimination of each variable to ensure that the model still adequately fits the data (http://userwww.sfsu.edu). We will prove that our model contains all of the independent variables. At each step, the significance of the independent variable being removed is tested using the Wald test (Hosmer and Lemeshow, 2000), (Duncan and Chapman, 2003). If the variables p-value are equal to or greater than the significant level, these variables should be eliminated from the model, otherwise, it remains in the model. A Wald test is used to test the statistical significance of each coefficient β in the model. A Wald test calculates a Z statistic, which is:

$Z = \frac{\beta}{SE}$

This Z value is then squared, yielding a Wald statistic with a chi-square distribution (http://userwww.sfsu.edu). Table 10 and 11 present the result for SPSS output of backward stepwise with Wald.

Variable	B	SE	Wald	Sig.	Exp(B)
Gender	-,058	,06	,93	,378	,944
Study course	,029	,26	,01	,737	1,02
Flexibility	,252	,04	39,69	,009	1,28
Understanding statistics	,192	,04	23,04	,033	1,21
Previous experience	-,072	,06	1,44	,284	,931
Self-study	,279	,02	194,6	,000	1,32
Structured format	,272	,03	82,2	,002	1,31
E-learning problems	-,004	1,3	,003	,988	,996
Study level	,18	,05	12,9	,041	1,19
Interactive environment	,176	,04	19,4	,028	1,19

Table 10: Backward stepwise with Wald, 1 step.

The model includes all the independent variables in the first step in table 10 and we see the p-value of the variables: Gender, study course, previous experience and e-learning problems are larger than the significant level 5%, this means these variables should be removed from the model, as we show in the second step in table 11 only the p-value of the variables are smaller than the significant level 5%.

Variable	B	SE	Wald	Sig.	Exp(B)
Flexibility	,294	,03	96,04	,00	1,34
Understanding statistics	,194	,04	23,5	,02	1,21
Self-study	,266	,02	176,9	,00	1,31
Structured format	,211	,03	49,5	,01	1,24
Study level	,167	,03	30,9	,02	1,18
Interactive environment	,145	,03	23,4	,03	1,16

Table 11: Backward stepwise with Wald, 2 step.

After elimination the variables has a weak effect on the dependent variable (Decision). We use the likelihood ratio test to check the fit of the model to data. This test checks the null hypothesis that all coefficients of the explanatory variables are zero. The likelihood-ratio test uses the ratio of the maximized value of the likelihood function for the full model (L1) over the maximized value of the likelihood function for the simpler model (L0). The likelihood-ratio test statistic equals:

$$D = -2log\frac{L0}{L1} = -2[log(L0) - log(L1)] = -2(L0 - L1)$$

If the difference is small, the independent variables contribute with little effect to explain the dependent variable and if the difference is large, we have a good model. Now we compare the two models with and without eliminated variables to see the goodness-of-fit of the reduced model L0. From table 12 and 13 we have the form:

D= 2[likelihood of full model - likelihood of reduced model]

= 1427,29 - 1415,77 = 11,52

Step	-2 Log likelihood	Cox Snell R Square	Nagelkerke R Square
1	1427,29[a]	,034	,047

Table 12: Full model.

Step	-2 Log likelihood	Cox Snell R Square	Nagelkerke R Square
1	1415,77[a]	,034	,047

Table 13: Reduced model.

When we remove four variables from the model the degree of freedom is 4. The statistic value D is larger than the critical value at 10$\%$ significant level equal to 7.78. The null hypothesis can then be rejected and then we will have a good model fit the data.

Through analysis a students sample is shown using the statistical analysis of the methods: Factor analysis, Logit model. We have noted the best variables that effect on the dependent variable and that have contributed to the students decision when using e-learning platforms in traditional courses. These variables were found to be: Flexibility, understanding statistics, self-study, structured format, study level and an interactive environment.

A user friendly e-learning platform with attractive design and clear screen readability, effective platform browsability and organization and interactivity is important for the user.

5 Conclusion

We have analyzed in this paper a study of a students sample in the faculty of economics -Humboldt university in Berlin. This sample was drawn from statistics II course. In summarizing i discuss the best variables that effect the students decision to use e-learning platforms in traditional courses: Flexibility, understanding statistics, self-study, structured format, study level and an interactive environment.

Flexibility is an important element in every e-learning system. This allows us to enter the system at anytime and anywhere and the e-learning platform should have a clear structured format to present the statistical content to the student without difficulties in understanding. The interactive environment should help the students and users to develope their skills and experiences. These reasons help the students to understand the course and increase self-learning while decreasing the need for teachers.

The evaluation result positively affirms that, the concise structural format, interactivity between platform content, hypertext functionality as well as the combination of statistical computing language like XploRe and R are the practical and innovative requirements for which users of the statistics e-learning platform can successfully learn and interact with statistical applications in the learning process. This helps the students to utilize e-learning platforms. It is a successful and effective method in education, but must only be an assistant factor to increase understanding along with the traditional courses and it is not an alternative method.

Through out the evaluation study we found: “That using e-learning platforms offer an effective way of learning statistics”. On the other hand one of the limitations of this study has been how to handle the number of variables that affect the evaluation of an e-learning platforms. Therefor it is particularly difficult to estimate whether an e-learning system fulfils the need of every learner.

References

Aydynly, G., Härdle, W. and Rönz, B. (2003). E-learning/e-teaching of statistics: A new challenge, Humboldt-Universität zu Berlin, Germany.

Boyum, S. (2006). Furcht vor dem Verlust der Arbeitsstelle: Eine Analyse mittels Logit Modellen, Humboldt-Universität zu Berlin, Germany.

Falco, R. and Soeiro, A. (2003). Measuring Impact of e-Learning on Quality and Learning Outcomes. A Proposal for a PhD Project, Universidade do Porto, Portugal.

Guan, S. (2006). A Binary Logistic Analysis of Car Consumer Behavior in China, Humboldt-Universität zu Berlin, Germany.

Hosmer, D. and Lemeshow, S. (1989). Applied Logistic Regression. New York: John Wiley Sons.

Krauss, F. and Ally, M. (2005). A Study of the Design and Evaluation of a Learning Object and Implications for Content Development, University of Toronto, Canada.

Lieu, V. (2007). Asking PISA-Questions to Bachelor students, Humboldt-Universität zu Berlin, Germany.

Mungo, J. (2004). E-learning/e-teaching: An Implementation and Evaluation of a Finance Introductory Course, Humboldt-Universität zu Berlin. Germany.

Posea, V., Matu, S. and Cristea, V. (2007). Online Evaluation of Collaborative Learning Platforms, Politehnica University of Bucharest, Computer Science Department, Bucharest, Romania.

Thomas, R. (2001). Interactivity and Simulations in e-Learning, http://www.multiverse.co.uk/whitepaper.pdf.

Appendix

Fragebogen für e-learning

1- Geschlecht: männlich ☐ weiblich ☐ 2-Studiengang: BWL ☐ VWL ☐ MTH ☐ Andere ☐

3- Alter: ☐ Welches Semester ☐

4- Haben Sie eine klare Vorstellung bezüglich des e-learning oder interessieren Sie sich für e- learning?

Ja ☐ Nein ☐

5- Haben Sie bereits Erfahrungen mit e-learning im Allgemeinen gehabt ?

Ja ☐ Nein ☐

6-Haben Sie bereits im Rahmen des regulären Uni-Studiums eine e-learning Methode (Z.B: CD, Audio, Video, Online Kurse) benutzt?

Ja ☐ Teilweise ☐ Nein ☐
Wenn Nein , warum nicht?

7- Das Statistik-Institut bietet verschiedene e-learning Plattformen an , z.B: MM*Stat, e-stat, Electronic Books, Moodle, TeachWiki.

Haben Sie bereits einige dieser Plattform benutzt ? Ja ☐ Nein ☐

Wenn Ja, welche : MM*Stat ☐ e-stat ☐ Electronic Books ☐ Moodle ☐ TeachWiki ☐

8- Hat die Nutzung der e-learning Plattform beim Verständnis des statistischen Themas geholfen ?

Ja ☐ Teilweise ☐ Nein ☐ Weiß nicht ☐

9 - Glauben Sie, dass die Nutzung von e-learning Plattformen Ihnen im Studium und beim Bestehen von Prüfung geholfen hat?

Ja ☐ Teilweise ☐ Nein ☐ Weiß nicht ☐

10 - Glauben Sie, dass e-learning Technologien Ihnen größere Freiheit bei der Bearbeitung des Stoffes gibt? Z.B : Moodle oder MM*Stat:

Ja ☐ Teilweise ☐ Nein ☐ Weiß nicht ☐

11- Konnten Sie den Vorlesungsstoff durch die Benutzung der e-learning-Methode ohne die Hilfe von Lehrkräften besser verstehen?

Ja ☐ Teilweise ☐ Nein ☐ Weiß nicht ☐

12 – Was ist Ihre meinung über die Strukturierte Form der e-learning plattform , die Sie verwendet haben ? Gut☐ Teilweise ☐

Schlecht ☐ Weiß nicht ☐

13- E-learning hilft beim Selbststudium: Ja ☐ Teilweise ☐ Nein ☐ Weiß nicht ☐

14- Was ist Ihre Meinung über die interaktive Umgebung von e-learning plattformen? Gut ☐ Teilweise ☐ Schlecht ☐ Weiß nicht ☐

15- Welche Probleme sind im Zusammenhang mit e-learning besonders wichtig?

Hohe Kosten ☐ Keine Anwesenheit von Lehrkräften ☐ Erhöhter Zeitaufwand ☐ Andere Gründe ☐

16- Glauben Sie, dass die e-learning Technologie in den traditionellen Lehrveranstaltungen immer Anwendung finden können?

Ja ☐ Teilweise ☐ Nein ☐ Weiß nicht ☐
17- Was sind Ihre Vorschläge, um das e-learning weiter zu verbessern ?

1- ………………………………………………………………………………………………..

2- ………………………………………………………………………………………………

3- ………………………………………………………………………………………………

Vielen Dank für Ihre Antwort

SFB 649 Discussion Paper Series 2008

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Testing Monotonicity of Pricing Kernels" by Yuri Golubev, Wolfgang Härdle and Roman Timonfeev, January 2008.
002 "Adaptive pointwise estimation in time-inhomogeneous time-series models" by Pavel Cizek, Wolfgang Härdle and Vladimir Spokoiny, January 2008.
003 "The Bayesian Additive Classification Tree Applied to Credit Risk Modelling" by Junni L. Zhang and Wolfgang Härdle, January 2008.
004 "Independent Component Analysis Via Copula Techniques" by Ray-Bing Chen, Meihui Guo, Wolfgang Härdle and Shih-Feng Huang, January 2008.
005 "The Default Risk of Firms Examined with Smooth Support Vector Machines" by Wolfgang Härdle, Yuh-Jye Lee, Dorothea Schäfer and Yi-Ren Yeh, January 2008.
006 "Value-at-Risk and Expected Shortfall when there is long range dependence" by Wolfgang Härdle and Julius Mungo, Januray 2008.
007 "A Consistent Nonparametric Test for Causality in Quantile" by Kiho Jeong and Wolfgang Härdle, January 2008.
008 "Do Legal Standards Affect Ethical Concerns of Consumers?" by Dirk Engelmann and Dorothea Kübler, January 2008.
009 "Recursive Portfolio Selection with Decision Trees" by Anton Andriyashin, Wolfgang Härdle and Roman Timofeev, January 2008.
010 "Do Public Banks have a Competitive Advantage?" by Astrid Matthey, January 2008.
011 "Don't aim too high: the potential costs of high aspirations" by Astrid Matthey and Nadja Dwenger, January 2008.
012 "Visualizing exploratory factor analysis models" by Sigbert Klinke and Cornelia Wagner, January 2008.
013 "House Prices and Replacement Cost: A Micro-Level Analysis" by Rainer Schulz and Axel Werwatz, January 2008.
014 "Support Vector Regression Based GARCH Model with Application to Forecasting Volatility of Financial Returns" by Shiyi Chen, Kiho Jeong and Wolfgang Härdle, January 2008.
015 "Structural Constant Conditional Correlation" by Enzo Weber, January 2008.
016 "Estimating Investment Equations in Imperfect Capital Markets" by Silke Hüttel, Oliver Mußhoff, Martin Odening and Nataliya Zinych, January 2008.
017 "Adaptive Forecasting of the EURIBOR Swap Term Structure" by Oliver Blaskowitz and Helmut Herwatz, January 2008.
018 "Solving, Estimating and Selecting Nonlinear Dynamic Models without the Curse of Dimensionality" by Viktor Winschel and Markus Krätzig, February 2008.
019 "The Accuracy of Long-term Real Estate Valuations" by Rainer Schulz, Markus Staiber, Martin Wersing and Axel Werwatz, February 2008.
020 "The Impact of International Outsourcing on Labour Market Dynamics in Germany" by Ronald Bachmann and Sebastian Braun, February 2008.
021 "Preferences for Collective versus Individualised Wage Setting" by Tito Boeri and Michael C. Burda, February 2008.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

022 "Lumpy Labor Adjustment as a Propagation Mechanism of Business Cycles" by Fang Yao, February 2008.
023 "Family Management, Family Ownership and Downsizing: Evidence from S&P 500 Firms" by Jörn Hendrich Block, February 2008.
024 "Skill Specific Unemployment with Imperfect Substitution of Skills" by Runli Xie, March 2008.
025 "Price Adjustment to News with Uncertain Precision" by Nikolaus Hautsch, Dieter Hess and Christoph Müller, March 2008.
026 "Information and Beliefs in a Repeated Normal-form Game" by Dietmar Fehr, Dorothea Kübler and David Danz, March 2008.
027 "The Stochastic Fluctuation of the Quantile Regression Curve" by Wolfgang Härdle and Song Song, March 2008.
028 "Are stewardship and valuation usefulness compatible or alternative objectives of financial accounting?" by Joachim Gassen, March 2008.
029 "Genetic Codes of Mergers, Post Merger Technology Evolution and Why Mergers Fail" by Alexander Cuntz, April 2008.
030 "Using R, LaTeX and Wiki for an Arabic e-learning platform" by Taleb Ahmad, Wolfgang Härdle, Sigbert Klinke and Shafeeqah Al Awadhi, April 2008.
031 "Beyond the business cycle – factors driving aggregate mortality rates" by Katja Hanewald, April 2008.
032 "Against All Odds? National Sentiment and Wagering on European Football" by Sebastian Braun and Michael Kvasnicka, April 2008.
033 "Are CEOs in Family Firms Paid Like Bureaucrats? Evidence from Bayesian and Frequentist Analyses" by Jörn Hendrich Block, April 2008.
034 "JBendge: An Object-Oriented System for Solving, Estimating and Selecting Nonlinear Dynamic Models" by Viktor Winschel and Markus Krätzig, April 2008.
035 "Stock Picking via Nonsymmetrically Pruned Binary Decision Trees" by Anton Andriyashin, May 2008.
036 "Expected Inflation, Expected Stock Returns, and Money Illusion: What can we learn from Survey Expectations?" by Maik Schmeling and Andreas Schrimpf, May 2008.
037 "The Impact of Individual Investment Behavior for Retirement Welfare: Evidence from the United States and Germany" by Thomas Post, Helmut Gründl, Joan T. Schmit and Anja Zimmer, May 2008.
038 "Dynamic Semiparametric Factor Models in Risk Neutral Density Estimation" by Enzo Giacomini, Wolfgang Härdle and Volker Krätschmer, May 2008.
039 "Can Education Save Europe From High Unemployment?" by Nicole Walter and Runli Xie, June 2008.
040 "Solow Residuals without Capital Stocks" by Michael C. Burda and Battista Severgnini, August 2008.
041 "Unionization, Stochastic Dominance, and Compression of the Wage Distribution: Evidence from Germany" by Michael C. Burda, Bernd Fitzenberger, Alexander Lembcke and Thorsten Vogel, March 2008
042 "Gruppenvergleiche bei hypothetischen Konstrukten – Die Prüfung der Übereinstimmung von Messmodellen mit der Strukturgleichungs-methodik" by Dirk Temme and Lutz Hildebrandt, June 2008.
043 "Modeling Dependencies in Finance using Copulae" by Wolfgang Härdle, Ostap Okhrin and Yarema Okhrin, June 2008.
044 "Numerics of Implied Binomial Trees" by Wolfgang Härdle and Alena Mysickova, June 2008.

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

045 "Measuring and Modeling Risk Using High-Frequency Data" by Wolfgang Härdle, Nikolaus Hautsch and Uta Pigorsch, June 2008.
046 "Links between sustainability-related innovation and sustainability management" by Marcus Wagner, June 2008.
047 "Modelling High-Frequency Volatility and Liquidity Using Multiplicative Error Models" by Nikolaus Hautsch and Vahidin Jeleskovic, July 2008.
048 "Macro Wine in Financial Skins: The Oil-FX Interdependence" by Enzo Weber, July 2008.
049 "Simultaneous Stochastic Volatility Transmission Across American Equity Markets" by Enzo Weber, July 2008.
050 "A semiparametric factor model for electricity forward curve dynamics" by Szymon Borak and Rafał Weron, July 2008.
051 "Recurrent Support Vector Regreson for a Nonlinear ARMA Model with Applications to Forecasting Financial Returns" by Shiyi Chen, Kiho Jeong and Wolfgang K. Härdle, July 2008.
052 "Bayesian Demographic Modeling and Forecasting: An Application to U.S. Mortality" by Wolfgang Reichmuth and Samad Sarferaz, July 2008.
053 "Yield Curve Factors, Term Structure Volatility, and Bond Risk Premia" by Nikolaus Hautsch and Yangguoyi Ou, July 2008.
054 "The Natural Rate Hypothesis and Real Determinacy" by Alexander Meyer-Gohde, July 2008.
055 "Technology sourcing by large incumbents through acquisition of small firms" by Marcus Wagner, July 2008.
056 "Lumpy Labor Adjustment as a Propagation Mechanism of Business Cycle" by Fang Yao, August 2008.
057 "Measuring changes in preferences and perception due to the entry of a new brand with choice data" by Lutz Hildebrandt and Lea Kalweit, August 2008.
058 "Statistics E-learning Platforms: Evaluation Case Studies" by Taleb Ahmad and Wolfgang Härdle, August 2008.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".