Bulletin UASVM Horticulture, 66(2)/2009
Print ISSN 1843-5254; Electronic ISSN 1843-5394

Knowledge Assessment: Distribution of Answers to an Online Quizzed System

Lorentz JÄNTSCHI, Carmen E. STOENOIU, Sorana D. BOLBOACĂ, Mugur C. BĂLAN, Liviu C. BOLUNDUȚ, Violeta POPESCU, Horea I. NAŞCU, Ioan ABRUDAN

Technical University of Cluj-Napoca, 400641 Cluj, Romania
correspondence: lori@academicdirect.org

Abstract. Number of correct answers and time spent by the students enrolled at Faculty of Materials Science and Engineering from Technical University of Cluj-Napoca and attending at Materials Chemistry discipline answering to a multiple choice multiple answer online evaluation system were analyzed. The analysis shown that the evaluation characteristics it follow a lifetime distribution.

Keywords: knowledge assessment; online evaluation; distribution fitting; goodness of fit; lognormal distribution; fatigue life distribution

INTRODUCTION

An auto-calibrated online system for students evaluation was previously designed (Jäntschi and Bolboacă, 2006) as alternative to a classical multiple choice examination system (Naşcu and Jäntschi, 2004). The system were implemented for a series of disciplines (Bolboacă and Jäntschi, 2007; Jäntschi and others, 2007). Descriptive statistics of students performances were the subject of another paper (Jäntschi and others, 2008).

The present study are based on the evaluation results obtained by the students attending at Materials chemistry discipline in first year of study at Faculty of Materials Science and Engineering from Technical University of Cluj-Napoca during the second semester of 2008-2009 year of study. The aim of the study was obtaining of the distribution law for the number of correct answers and for the time necessary to answer to the quizzes.

MATERIALS AND METHODS

A number of 59 students participated to the supervised online evaluation from June 5 to June 21, 2009. The system allows multiple evaluations during a period of evaluation. The students used this feature of the system, giving a total number of 83 evaluations. When a student participate more than once at evaluation, the lowest score obtained does not enter in the computation of the average, giving thus to the students the opportunity to first accommodate with the system. A number of 63 evaluations meet these criteria from all 83 evaluations. The obtained scores are online available by querying the database:

http://l.academicdirect.org/Education/Evaluation/Chemistry/Materials/statistics.php?punctaje=09

The search of the statistical distribution of the points (from 0 to 30) and of the times spend (from near 0 to near infinity) were conducted using Maximum likelihood estimation (Fisher, 1912) to obtain population parameters and using Pearson-Fisher (Pearson, 1900; Fisher, 1922; Fisher, 1924), Kolmogorov-Smirnov (Kolmogorov, 1941; Smirnov, 1948) and Anderson-Darling (Anderson and Darling, 1952) statistics to measure the goodness of fit. A number of 61 alternatives of probability density function were available (by using EasyFit software) and were included in the analysis.

RESULTS AND DISCUSSION

Data were subject of analysis according to the design of the experiment from Table 1. The histograms of the observables are given in Figure 1.

Table 1. Quizzes answers analysis: design of the experiment

Number of observables	Correct answers	Times spent
All evaluations	O1 (N=83)	O2 (N=83)
Evaluations contributing to the final scores	O3 (N=63)	O4 (N=63)

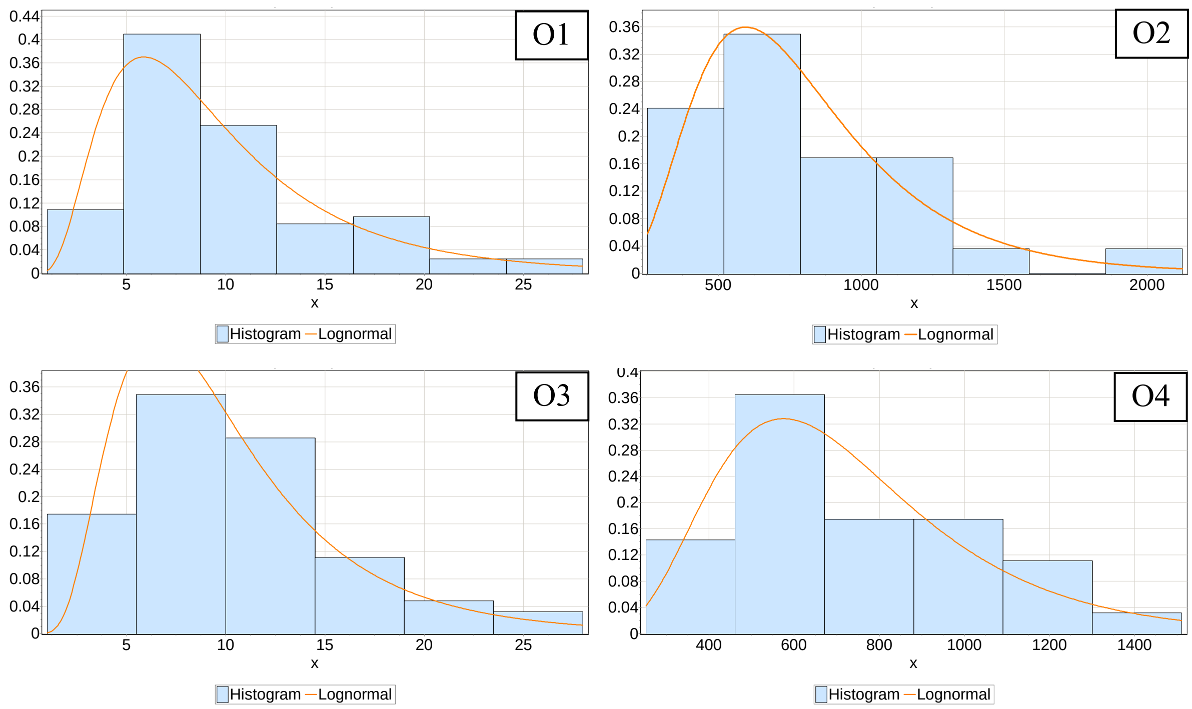

Figure 1. Histograms of the correct answers (O1 and O3) and time spent (O2 and O4)

Table 2. Probabilities (in percents) from goodness of fit statistics for five alternatives of distribution

Obs.	O1					O2					O3					O4				
S\D	BS	IG	LN	NG	WB	BS	IG	LN	NG	WB	BS	IG	LN	NG	WB	BS	IG	LN	NG	WB
PF	49	2.80	66	0.10	13	14	14	16	0.10	1	90	99	87	2	79	60	8	60	2	2
KS	27	66	56	4.60	35	56	63	55	9	13	43	84	74	8	42	61	51	59	11	24
AD	36	44	52	6.10	31	57	60	55	7	12	39	54	55	11	31	49	39	47	23	25
Min.	***27***	3	**52**	0.10	13	***14***	14	**16**	0.10	1	***39***	54	**55**	2	31	**49**	8	***47***	2	2

Obs.: Observables; S\D: Statistic vs. Distribution; Min.: Minimum of the probability
Distributions: BS - Fatigue life; IG - Inverse Gaussian; LN - Lognormal; NG - Normal; WB - Weibull
Statistics: PF - Pearson-Fiher; KS - Kolmogorov-Smirnov; AD - Anderson-Darling

Table 2 contains the results of the goodness of fit for five alternatives of two parameters probability density functions: Fatigue life or Birnbaum-Saunders (Birnbaum and Saunders, 1969), Inverse Gaussian (Chhikara and Folks, 1977), Lognormal (Aitchison and Brown, 1957), Normal (Gauss, 1809), and Weibull (Weibull, 1951).

As can be seen from Table 2, two two-parametric distribution functions have good agreements with the observed data: Fatigue life distribution (27% of their samples are worst than the O1 observable; 14% of their samples are worst than the O2 observable; 39% of their samples are worst than the O3 observable; 49% of their samples are worst than the O4 observable; a geometric mean of 29% for all O1-O4 observables) and Lognormal distribution (52% of their samples are worst than the O1 observable; 16% of their samples are worst than the O2 observable; 55% of their samples are worst than the O3 observable; 47% of their samples are worst than the O4 observable; a geometric mean of 38% for all O1-O4 observables). Since is no reason to reject any of Fatigue life and Lognormal distribution as population distribution hypothesis, Table 3 contains the distribution parameters as was obtained from maximum likelihood estimation procedure for these two distributions and for all four observables.

Table 3. O1-O4: population parameters and statistics for fatigue life and lognormal distributions

Distribution	Parameter	O1	O2	O3	O4
Fatigue life	Shape	0.631	0.460	0.581	0.414
	Scale	8.159	731	8.619	678
	Mean	9.784	808	10.08	737
	St.Dev.	6.301	378	5.974	310
	Mode	5.324	586	6.013	568
	C.Var.	64.4%	46.8%	59.3%	42.1%
Lognormal	Scale	0.593	0.450	0.542	0.408
	Location	2.123	6.592	2.178	6.522
	Mean	9.966	809	10.23	739
	St.Dev.	6.473	382	5.969	314
	Mode	5.877	595	6.586	575
	C.Var.	65.0%	47.4%	58.4%	42.5%

The population statistics from Table 3 can serve for further analysis. Thus, is expected that the population true distribution to have highest accuracy around the mode. A ±5% interval can be constructed around the mode for both theoretical distributions and in it should be about 10% of the sample. Chi Square statistic can be used to check this hypothesis for both distributions; then the probability from Chi Square distribution measures the likelihood of the observation. Table 4 contain this procedure.

Table 4. Likelihood of the theoretical distribution

Distribution	Parameter	O1	O2	O3	O4
Fatigue life	Mode-5%	4.80	547.9	5.49	535.2
	Mode+5%	5.85	623.6	6.54	599.8
	Observations falling in	8	14	7	10
	Expected falling in	6.3 (10%)	8.3 (10%)	6.3 (10%)	8.3 (10%)
$\chi^2(6.17,4)=19\%$	$X^2=(O-E)^2/E$	0.01	3.91	0.08	2.17
$\chi^2(6.18,4)=2\%$	$X^2=(O-E)^2/E$	0.88	0.35	0.84	9.41
Lognormal	Expected falling in	6.3 (10%)	8.3 (10%)	6.3 (10%)	8.3 (10%)
	Observations falling in	11	10	4	14
	Mode+5%	6.40	632.8	7.11	607.5
	Mode-5%	5.35	558.2	6.07	543.4

Results from table 4 suggest that is more likely fatigue life than lognormal the distribution of the observables (19% vs. 2%).

CONCLUSIONS

The correct answers and the spent time of students solving the proposed online evaluation with multiple choice multiple answer are distributed by a lifetime law. Most likely this law is the fatigue life distribution. For the students which use the feature to be tested more than once the mode of correct answers increases significantly with 13% (from 5.324 to 6.013), and the mode of spent time decreases insignificantly with 3% (from 586s to 568s).

REFERENCES

Aitchison, J., J.A.C. Brown (1957). The lognormal distribution. Cambridge University Press, Cambridge.

Anderson, T. W., D. A. Darling (1952). Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. Annals of Mathematical Statistics 23(2):193-212.

Birnbaum, Z. W., S. C. Saunders (1969). A new family of life distributions. Journal of Applied Probability 6(2):319-327.

Bolboacă, S. D., L. Jäntschi (2007). Computer-based testing on physical chemistry topic: A case study, International Journal of Education and Development using Information and Communication Technology, 3(1):#242.

Chhikara, R. S., J. L. Folks (1977). The inverse Gaussian distribution as lifetime model. Technometrics 19:461-468.

Fisher R. A. (1924). The Conditions Under Which $\chi 2$ Measures the Discrepancy Between Observation and Hypothesis. Journal of the Royal Statistical Society 87:442-450.

Fisher, R. A. (1912). On an Absolute Criterion for Fitting Frequency Curves. Messenger of Mathematics 41:155-160.

Fisher, R. A. (1922). On the Interpretation of $\chi 2$ from Contingency Tables, and the Calculation of P. Journal of the Royal Statistical Society 85:87-94.

Gauss, C. F. (1809). Theoria Motus Corporum Coelestium. Perthes et Besser, Hamburg. Translated, 1857, as Theory of Motion of the Heavenly Bodies Moving about the Sun in Conic Sections, trans. C. H. Davis. Little, Brown; Boston. Reprinted, 1963, Dover, New York.

Jäntschi, L., C. E. Stoenoiu, S. D. Bolboacă (2007). Linking Assessment to e-Learning in Microbiology and Toxicology for Undergraduate Students, The International Conference on "Computer as a Tool", IEEE 07EX1617C, 2447-2452.

Jäntschi, L., S. D. Bolboacă (2006). Auto-calibrated Online Evaluation: Database Design and Implementation, Leonardo Electronic Journal of Practices and Technologies, 5(9):179-192.

Jäntschi, L., S. D. Bolboacă, M. M. Marta, A. Laszlo (2008). E-Learning and E-Evaluation: A Case Study. Conference on Human System Interaction, IEEE CN:08EX1995C, #4-KF-003247.

Kolmogorov, A. (1941). Confidence Limits for an Unknown Distribution Function. The Annals of Mathematical Statistics 12(4):461-463.

Naşcu, H. I., L. Jäntschi (2004). Multiple Choice Examination System 2. Online Quizzes for General Chemistry, Leonardo Electronic Journal of Practices and Technologies, 3(5):26-36.

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. Philosophical Magazine 5th Ser 50:157-175.

Smirnov, N. V. (1948). Table for estimating the goodness of fit of empirical distributions. The Annals of Mathematical Statistics 19(2):279-281.

Weibull, W. (1951). A statistical distribution function of wide applicability. Journal of Applied Mechanics -Transactions ASME 18(3):293-297.