

ANALYZING THE IMPACT OF ONLINE EXAMS VS TRADITIONAL EXAMS ON THE GRADES OBTAINED

Beatriz Cobo¹, Ramón Ferri-García², Luis Castro-Martín^{3*}

¹Dr. Beatriz Cobo, University of Granada, SPAIN, beacr@ugr.es

²Dr. Ramón Ferri-García, University of Granada, SPAIN, rferri@ugr.es

³Mr. Luis Castro-Martín, University of Granada, SPAIN, luiscastro193@ugr.es

*Corresponding author

Abstract

The COVID-19 crisis forced a transition from conventional on-site teaching to online teaching. One of the main concerns was the evaluation process, since online exams greatly increase the risk of cheating. This risk is hard to avoid and therefore we may ask ourselves whether there are significant differences in the grades obtained by the students depending on the evaluation system (on-site or online).

Our objective is analyzing the differences between the grades obtained by the students via on-site exams and those obtained via online exams in order to determine if they are significant. We also analyze if, for a given gender of the students, there are significant differences between both evaluation systems. Finally, we analyze whether there are significant differences between male and female students for a given evaluation system.

The study participants are the students of the Quantitative Techniques I and Quantitative Techniques II courses at the Faculty of Economics and Business at the University of Granada from different academic years. A convenience sample of 299 grades obtained via traditional exams and 272 grades obtained via online exams is drawn.

The analysis of the data shows, via statistical hypothesis tests, significantly different means and variances between the grades obtained via traditional exams and those obtained via online exams. The mean grade via traditional exams is 4.28 and the mean grade via online exams is 5.09. When the same analysis is carried out for a fixed gender, the results show significant differences between the final grades obtained with each evaluation system by male students while there are no significant differences for the female students. When the evaluation system is fixed, we find significant differences by gender for traditional exams, being the mean grade obtained by females 4.56 and the mean grade obtained by males 3.97, while there are no significant differences for online exams.

In conclusion, the interesting results obtained invite us to further investigate the factors related to online teaching which are affecting the final grades of the students.

Keywords: online exams, COVID-19, gender.

1 INTRODUCTION

The COVID-19 crisis forced a transition from offline to online teaching and thus we had to reinvent ourselves although one of the biggest preoccupations for teachers is evaluations. Online exams imply a greater risk of students copying or even committing identity fraud during exams which is very difficult to deter. It is by means of the online teaching platform used by the university and with the help of other software for random exam generation that we have been able carry out an online evaluation in a secure and efficient way, even if

we still wonder if there are significant differences in the scores obtained by students depending on the evaluation system (offline or online).

Many studies have taken place in relation to the matter throughout the years. In 2009 Larson and Sung carried out a research in which they compared the final scores for online exams, blended exams and offline exams. In 2008, Harmon and Lambrinos conducted experiments in which they studied if online exams are an invitation to copying. We can even find projects that classify the different threats encountered during an online exam (Ullah et al. 2016). In 2013, Sarrayih and Ilyas proposed a system reinforcing security to improve online evaluations by using techniques based on biometric authentication or cryptography among others.

Since the beginning of the pandemic many authors have been concerned about it. In 2020, Eltayeb et al. carried out a survey in order to check the perception that students had about teaching and online exams. In the same year, Balderas and Caballero analyzed learning records to detect students copying in online exams. In 2021, Ebaid conducted a survey for students in order to see how easy it was to copy in online exams. Also in 2021, El Said compared the scores obtained both offline and online by evaluating globally and taking into account factors such as gender, course credits and student age.

2 OBJECTIVES

- Showing if there are significant differences between the average scores obtained by students in offline exams and the average scores obtained in online exams, considering subjects of a similar difficulty.
- Studying if there are significant differences among the examination modalities, depending on the student gender.
- Studying if there are significant differences taking into account student gender, depending on the examination modalities.

3 METHODOLOGY

3.1 Participants

The participants of this research are students for the subject of Quantitative Techniques I and II in the Faculty of Economic and Business Sciences of the University of Granada during different academic courses. In this research we will only include these subjects because our intention is to choose subjects of a similar difficulty. The purpose of this is to exclusively occasion variance because of the type of exam, online or offline, and not due to its difficulty.

3.2 Variables and instruments

In this investigation the gathered variables were gender and final exam scores without considering the scores of the continuous assessment since these values could affect our results.

3.3 Procedures

A convenience sample of size 570 was obtained, of which 299 correspond to the scores obtained during offline exams and 271 for online exams.

3.4 Data analysis

The data analysis that we carry out is done with the statistical software R. Firstly, we run a normality test to see if the data has a normal distribution and to observe if we can conduct a parametric study, accompanied by a descriptive research afterwards. Considering the scores obtained in final exams as the target study variable, and whether the exam is offline or online as the grouping variable, we contrast the hypothesis with the variable quotient among the different examination mechanisms and see if we can consider them as equal or not. Lastly, a hypothesis contrast is done for the comparison of the average scores between both mechanisms. These contrasts are done globally as well as separated by gender and taking into account the examination modality.

4 RESULTS

We have to take into account that the online exams have been carried out with the software R, specifically with the Exams package (Zeilis et al., 2014) which helps generate exams with random data, and therefore every single student has a different exam. The idea of using the software came to us because of the need to avoid students copying in online exams, or at least, increasing the difficulty of it. They might be able to help

themselves with the procedure, but all of them will have to carry out different calculations. This, together with the functionality that the platform Prado provides us with the creation of time-limited questionnaires, enables us to have an online evaluation.

In this research we can find interesting results which will be developed hereunder and that can be seen in Table 1.

Table 1: Sample size, average and standard deviation of the final score. Value of the contrast statistical and p-value for the contrast of the quotient of variances and for the contrast of the difference in averages both in their totality and according to gender and the examination modality

		s	A	SD	F	p	t	p
	Offline	299	4.28	2.34	4.97	0.03	-3.88	<0.001
	Online	271	5.09	2.60				
Men	Offline	155	3.97	2.41	0.42	0.52	-4.02	<0.001
	Online	121	5.18	2.58				
Women	Offline	144	4.62	2.23	7.60	0.01	-1.40	0.16
	Online	150	5.02	2.62				
Offline	Men	155	3.97	2.41	0.42	0.03	-2.43	0.02
	Women	144	4.62	2.23				
Online	Men	121	5.18	2.58	7.60	0.91	0.52	0.61
	Women	150	5.02	2.62				

Note: s = Sample size; A= Final score average; SD = Standard deviation.

To start with, we consider the normality test of Kolmogorov-Smirnov in which we contrast if the data has a normal distribution. After carrying it out we obtain a p-value of 0.105. Thus, if we consider a significance level of 5%, we do not have enough sample evidence to reject the normality of the data and we can conduct a parametric analysis.

As previously stated, we will firstly work on the complete sample and we carry out a hypothesis contrast concerning the variable quotient

$$H_0: \sigma_{offline}^2 = \sigma_{online}^2$$

$$H_1: \sigma_{offline}^2 \neq \sigma_{online}^2$$

obtaining a p-value of 0.026, which means that we reject the null hypothesis and we consider that the variance of offline exam scores differs from the obtained in online exams.

Taking into account this result we carry out a contrast of the difference in averages with unknown and different variances

$$H_0: \mu_{offline} = \mu_{online}$$

$$H_1: \mu_{offline} \neq \mu_{online}$$

obtaining a p-value < 0.001, hence rejecting the null hypothesis and considering that the average scores obtained for both offline and online exams are significantly different, being 4.28 the average for offline and 5.09 for online.

If we take gender into account and we carry out the same analysis spoken about, for males we consider:

$$H_0: \mu_{men-offline} = \mu_{men-online}$$

$$H_1: \mu_{men-offline} \neq \mu_{men-online}$$

In the case of the quotient of variances the p-value obtained is 0.52, therefore we do not have enough sample evidence to reject the null hypothesis and thus considering that the variances of the scores can be equal. We then carry out the contrast of the difference in the average with unknown and equal variances obtaining a p-value < 0.001 so the average scores obtained for both examination modalities differ significantly.

As for the female case:

$$H_0: \mu_{women-offline} = \mu_{women-online}$$

$$H_1: \mu_{women-offline} \neq \mu_{women-online}$$

In the contrast of the variance quotient, with a p-value of 0.01 we can conclude that significant differences exist concerning the variances of the scores thus making us carry out a contrast of the difference of the averages with unknown and different variances by means of p-value. In this case it is equal to 0.16 leading to the conclusion of there not being significant differences among the score averages with both mechanisms.

After obtaining these results, we ponder if the differences found could be due to gender. Focusing on the offline case we consider different variances, thus making us carry out a contrast of the differences in averages for both unknown and different variances

$$H_0: \mu_{offline-men} = \mu_{offline-women}$$

$$H_1: \mu_{offline-men} \neq \mu_{offline-women}$$

obtaining the result of score averages between men and women being statistically different.

If we focus on the online case considering that variances are equal and we therefore carry out a contrast of the differences in averages with unknown but equal variances

$$H_0: \mu_{online-men} = \mu_{online-women}$$

$$H_1: \mu_{online-men} \neq \mu_{online-women}$$

obtaining that the average scores between men and women can be considered equal.

5 DISCUSSION

The transition to online teaching that the COVID-19 pandemic brought implies a virtualisation of evaluations. According to our results, the scores obtained by students by means of a virtual evaluation system differ from those obtained in similar conditions in an offline context. It would hence be interesting to investigate if there are any related factors that exist with virtuality that condition the results of the students.

Depending on gender and carrying out the same research, the result obtained is that there are significant differences in final exam scores for both modalities in the case of men but not in the case of women. Focusing on the type of exam, we find significant differences concerning gender for the offline modality but not in the online modality.

After carrying out the research, we find the results very interesting and we thus will keep on investigating the subject considering students from other faculties.

REFERENCE LIST

Balderas, A., & Caballero-Hernández, J. A. (2020). Analysis of learning records to detect student cheating on online exams: Case study during COVID-19 pandemic. In Eighth international conference on technological ecosystems for enhancing multiculturalism (pp. 752-757).
<https://doi.org/10.1145/3434780.3436662>

Eltayeb, L. B., Alharthi, N. S., Elmosaad, Y. M., & Waggiallah, H. A. (2020). Students' perception on E-Learning and Remote Exams during COVID 19 Outbreak 2020. International Journal of

Pharmaceutical and Phytopharmacological Research (eIJPPR), 10(5), 142-148.
<https://eijppr.com/hMottTN>

- El Said, G. R. (2021). How did the COVID-19 pandemic affect higher education learning experience? An empirical investigation of learners' academic performance at a university in a developing country. *Advances in Human-Computer Interaction*, 2021. <https://doi.org/10.1155/2021/6649524>
- Harmon, O. R., & Lambrinos, J. (2008). Are online exams an invitation to cheat?. *The Journal of Economic Education*, 39(2), 116-125. <https://doi.org/10.3200/JECE.39.2.116-125>
- Larson, D. K., & Sung, C. H. (2009). Comparing student performance: Online versus blended versus face-to-face. *Journal of Asynchronous Learning Networks*, 13(1), 31-42.
<https://dx.doi.org/10.24059/olj.v13i1.1675>
- Sarrayrih, M. A., & Ilyas, M. (2013). Challenges of online exam, performances and problems for online university exam. *International Journal of Computer Science Issues (IJCSI)*, 10(1), 439.
- Ullah, A., Xiao, H., & Barker, T. (2016). A classification of threats to remote online examinations. In 2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON) (pp. 1-7). IEEE. <https://doi.org/10.1109/IEMCON.2016.7746085>
- Zeileis, A., Umlauf, N., & Leisch, F. (2014). Flexible generation of e-learning exams in R: Moodle quizzes, OLAT assessments, and beyond. *Journal of Statistical Software*, 58, 1-36.
<https://doi.org/10.18637/jss.v058.i01>