EVALUATION OF DISTANCE EXAMS IN MATHEMATICS

Jindřich Klůfa

Abstract

The analysis of the results of examinations in mathematics at the Prague University of Economics and Business at the time of coronavirus will be performed in this paper. We shall compare the distance form and the full-time form (with personal participation) of mathematics exams. For the comparison we shall use different methods of mathematical statistics. The analysed data are the results of examinations in mathematics in basic course Mathematics for economists (ident 4MM101) in winter semester of the academic year 2019/2020 and winter semester of the academic year 2020/2021. Results of math exams in winter semester of the academic year 2019/2020 were obtained in the period before coronavirus (the full-time form of mathematics exams), results of math exams in winter semester of the academic year 2020/2021 were obtained in the period after coronavirus (the distance form of mathematics exams). The conclusions of this paper show that oral exams play an irreplaceable role during coronavirus.

Key words: Exams in mathematics, distance form of examinations, methods of mathematical statistics.

JEL Code: C12, I21

Introduction

Math examinations at the Prague University of Economics and Business in the coronavirus time will be analysed in present paper. Results of these exams will be compared with results of exams in mathematics in the period before coronavirus. Math exams at the university consist from three parts. The first part is a mid-term test, which is usually written in the ninth week of the semester. The share of this part of the exam in the final evaluation is 20 %. The second part of the exam in mathematics is the final test, which is written after the end of the semester in the term chosen by students. The share of this part of the exam in the final evaluation is 40 %. These tests are standard tests, no multiple choice question tests (see e.g. (Klůfa, 2015b), (Klůfa, 2015c), (Klůfa, 2016)) are used at the time of coronavirus. Within two days after the end of the final test, an oral exam is held, which also contributes 40 % to the final evaluation. The student

successfully completes the exam in mathematics if he obtains at least sixty points out of a hundred – see Table 1.

Tab. 1: Grade

Grade	Points
1 = Excellent	90 - 100
2 = Very good	75 - 89
3 = Good	60 - 74
4+ = Failed, eligible for retake	50 - 59
4 = Failed	0 - 49

Source: own construction

For comparison of the distance form and the full-time form (with personal participation) of mathematics exams we shall use the results of examinations in mathematics in basic course Mathematics for economists (ident 4MM101) in winter semester of the academic year 2019/2020 and in winter semester of the academic year 2020/2021. To eliminate influence different groups of examiners, we shell use results of exams of the same examiners. The analysed data are the results of exams in mathematics of 165 students in winter semester of the academic year 2019/2020 (without coronavirus) and 114 students in winter semester of the academic year 2020/2021 (with coronavirus).

The comparison of online and traditional exams we can find also in (Rovai, 2000). The using of an oral exam as a form of assessment in the online context at the Griffith University in Southport, Australia is studied in (Akimov and Malin, 2020). Student performance in and attitudes towards oral and written exams were compared in (Huxham, Campbell and Westwood, 2012). Oral examinations in German universities are studied in (Kehm, 2001). In (Okada, Mendonca and Scott, 2015) are mentioned practices on the use of a web videoconferencing application to quality control student assignments through online oral examination. Similar problems are studied in (Joughin, 2007), (Mesicek et al, 2017), (Klůfa, 2015a), (Fluck, 2019), (Kaspříková and Klůfa, 2011), (Salkova et al, 2020). Outcomes of this article can be used for improvement of the online education at University of Economics and Business.

1 Results

Now we shall compare the distance form and the full-time form (with personal participation) of mathematics exams. For the comparison we shall use basic statistical methods, especially Student *t* test for independent samples and the same variance. Statistic t (see e.g. (Anděl, 1978)) is

$$t = \frac{\bar{x} - \bar{y}}{s\sqrt{\frac{1}{m} + \frac{1}{n}}}, \quad \text{where } s = \sqrt{\frac{1}{m+n-2}\left[(m-1)s_x^2 + (n-1)s_y^2\right]} \qquad (1)$$

($\bar{x}$ is average number of points in the full-time form of the examination, $\bar{y}$ is average number of points in the distance form of the examination, s_x^2 is variance of points in the full-time form of the examination, s_y^2 is variance of points in the distance form of the examination, $m = 165$ (the sample size in winter semester of the academic year 2019/2020), $n = 114$ (the sample size in winter semester of the academic year 2020/2021)). When

$$|t| > t_\alpha(m+n-2), \qquad (2)$$

where $t_\alpha(m+n-2)$ is critical value of t distribution with $(m+n-2)$ degrees of freedom and significance level α, the hypothesis "mean number of points in both forms of the examination is the same" is rejected at significance level α.

1.1 Mid-term test

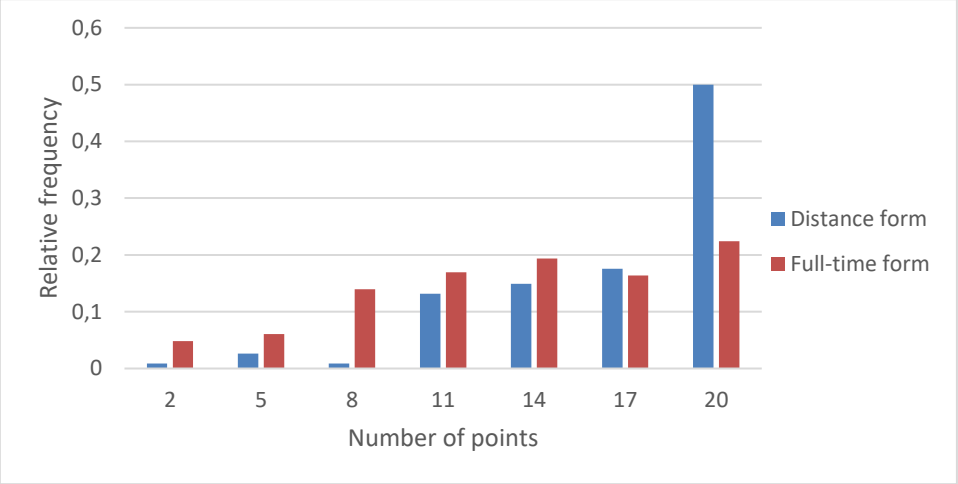
The number of points in the mid-term test can be in interval [0,20]. The mid-term test is usually written in the ninth week of the semester. The mid-term test in winter semester of the academic year 2020/2021 consisted of eight homework assignments for 1 point and a distance test for 12 points. Distributions of number of points in the mid-term test in math in winter semester of the 2019/2020 academic year and distributions of number of points in the mid-term test in math in winter semester of the 2020/2021 academic year in the course Mathematics for economists (ident 4MM101) are in Figure 1. Some fundamental descriptive statistics of these distributions are in Table 2.

Tab. 2: Descriptive statistics for number of points in mid-term test

	Full-time form (2019/2020)	Distance form (2020/2021)
Frequency	165	114
Average	12.430	16.070
Median	13	17.5
Mode	20	20
Std. Deviation	5.2860	4.424
Kurtosis	-0.831	1.157
Skewness	-0.321	-1.219

Source: own calculation

Fig. 1: Distribution of number of points in mid-term test in mathematics



Source: own construction

Average number of points in mid-term test for full-time form is $\bar{x} = 12.43$ and for distance form is $\bar{y} = 16.07$. The difference between these averages may be significant or may have occurred randomly. For objective decision we shall use t-test for independent samples. We shall test the null hypothesis $H_0 : \mu_1 = \mu_2$, i.e. mean number of points in both forms of mid-term test is the same, against the alternative hypothesis $H_1 : \mu_1 \neq \mu_2$. Since (see Tab. 3)

$$|t| = 6.035 > t_{0.05}(277) = 1.969, \tag{3}$$

the null hypothesis "mean number of points in both forms of mid-term test is the same" is rejected at significance level 0.05. The difference between averages $\bar{x} = 12.43$ and $\bar{y} = 16.07$ is statistical significant. The results of the full-time form and the distance form of the mid-term tests differ significantly (P value is 5.09×10^{-9}, see Tab. 3).

Remark. For the Student *t* test above, we assume that the variance of points in mid-term test in the full-time form and in the distance form is the same. We used the *F* test to verify this assumption. The result of the *F* test shows that assumption of the *t* test is fulfilled.

Tab. 3: Student *t* test for independent samples (mid-term tests)

Significance level 0.05	2020/2021	2019/2020
Average number of points	16.07018	12.4303
Frequency	114	165
Degrees of freedom	277	
t Stat	6.034749	
P value	5.09E-09	
t krit	1.968565	

Source: own calculation

1.2 Final test

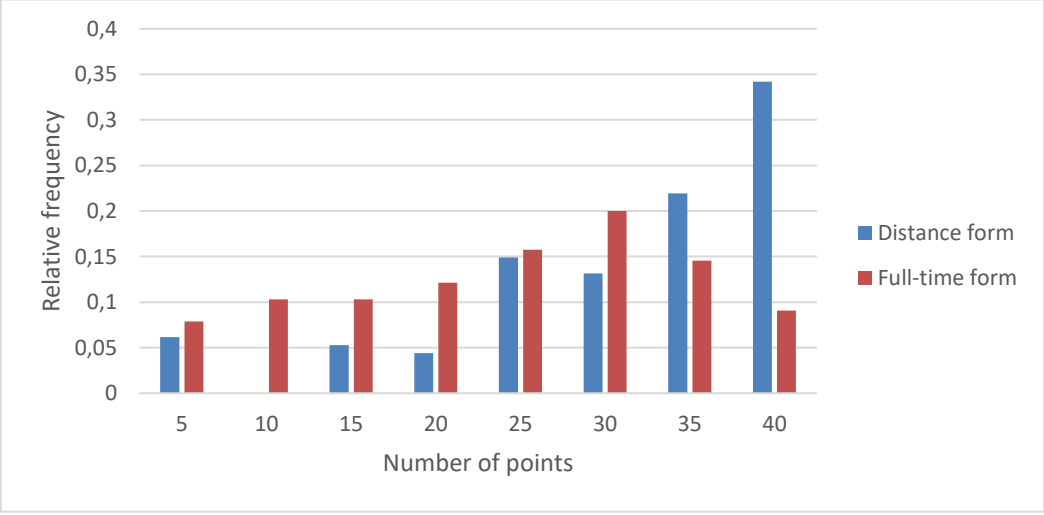
The number of points in the final test can be in interval [0,40]. The final test is written after the end of the semester in the term chosen by students. The final tests in winter semester of the academic year 2020/2021 were performed in distance form with a strictly limited processing time. Distributions of number of points in the final test in math in winter semester of the 2019/2020 academic year and distributions of number of points in the final test in math in winter semester of the 2020/2021 academic are in Figure 2. Some fundamental descriptive statistics of these distributions are in Table 4.

Tab. 4: Descriptive statistics for number of points in final test

	Full-time form (2019/2020)	Distance form (2020/2021)
Frequency	165	114
Average	22.127	29.053
Median	24	31.5
Mode	29	40
Std. Deviation	10.622	10.488
Kurtosis	-0.867	1.373
Skewness	-0.298	-1.352

Source: own calculation

Fig. 2: Distribution of number of points in final test in mathematics



Source: own construction

Average number of points in final test for full-time form is $\bar{x} = 22.13$ and for distance form is $\bar{y} = 29.05$. The difference between these averages may be significant or may have occurred randomly. For objective decision we shall use t-test for independent samples. We shall test

$$H_0 : \mu_1 = \mu_2, \tag{4}$$

i.e. mean number of points in both forms of final test is the same, against the alternative hypothesis $H_1 : \mu_1 \neq \mu_2$. Because (see Tab. 5)

$$|t| = 5.381 > t_{0.05}(277) = 1.969, \tag{5}$$

the null hypothesis "mean number of points in both forms of final test is the same" is rejected at significance level 0.05. The difference between averages $\bar{x} = 22.13$ and $\bar{y} = 29.05$ is statistical significant. The results of the full-time form and the distance form of the final tests differ significantly (P value is 1.58×10^{-7}, see Tab. 5).

Remark. For the Student *t* test above, we assume that the variance of points in final test in the full-time form and in the distance form is the same. We used the *F* test to verify this assumption. Statistics $F = s_x^2 / s_y^2 = 1.025654$. Since *F* is not greater than critical value of *F* distribution with *m*-1=164 and *n*-1=113 degrees of freedom and significance level 0.025 (see e.g. (Anděl, 1978)), we can say that assumption of the *t* test is fulfilled.

Tab. 5: Student *t* test for independent samples (final tests)

Significance level 0.05	2020/2021	2019/2020
Average number of points	29.05263	22.12727
Frequency	114	165
Degrees of freedom	277	
t Stat	5.381092	
P value	1.58E-07	
t krit	1.968565	

Source: own calculation

1.3 Oral examination

The number of points in the oral exam can be in interval [0,40]. The oral examinations follow the exam tests within two days after the exam tests. The distance form of the oral exams in winter semester of the academic year 2020/2021 was performed using MS Teams. Distributions of number of points in the oral exam in mathematics in winter semester of the 2019/2020 academic year and distributions of number of points in the oral exam in winter semester of the 2020/2021 academic year are in Figure 3. Some descriptive statistics of these distributions are

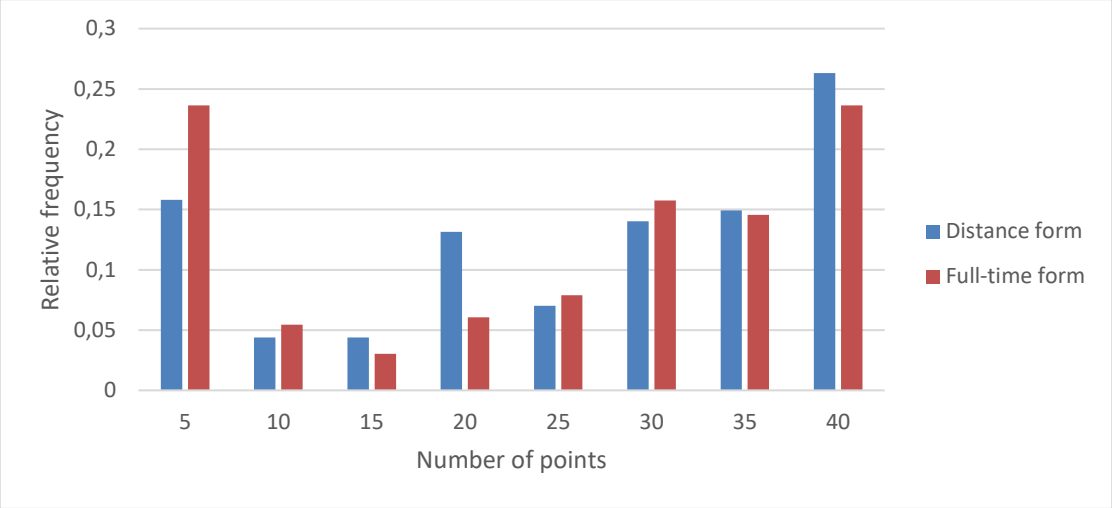
in Table 6. The mode 0 in winter semester of the 2019/2020 academic year is due to the fact that a large number of students did not come to the oral exam after unsuccessful tests.

Tab. 6: Descriptive statistics for number of points in oral exam

	Full-time form (2019/2020)	Distance form (2020/2021)
Frequency	165	114
Average	22.873	24.886
Median	28	30
Mode	0	40
Std. Deviation	14.930	13.822
Kurtosis	-1.294	-0.881
Skewness	-0.492	-0.645

Source: own calculation

Fig. 3: Distribution of number of points in oral examination in mathematics



Source: own construction

Average number of points in oral exam for full-time form is $\bar{x} = 22.87$ and for distance form is $\bar{y} = 24.89$. The difference between these averages may be significant or may have occurred randomly. For objective decision we shall use t-test for independent samples. We shall test the null hypothesis $H_0 : \mu_1 = \mu_2$, i.e. mean number of points in both forms of oral exam is the same, against the alternative hypothesis $H_1 : \mu_1 \neq \mu_2$. Because (see Tab. 7)

$$|t| = 1.141 < t_{0.05}(277) = 1.969, \qquad (6)$$

the null hypothesis "mean number of points in both forms of oral exam is the same" is not rejected at significance level 0.05 (P value is 0.255). The difference between averages $\bar{x} = 22.87$ and $\bar{y} = 24.89$ is not statistical significant.

Remark. For the Student *t* test above, we assume that the variance of points in final test in the full-time form and in the distance form is the same. We used the *F* test to verify this assumption. Statistics $F = s_x^2 / s_y^2 = 1.166687$. Since *F* is not greater than critical value of *F* distribution with *m*-1=164 and *n*-1=113 degrees of freedom and significance level 0.025 (see e.g. (Anděl, 1978)), we can say that assumption of the *t* test is fulfilled.

Tab. 7: Student *t* test for independent samples (oral examination)

Significance level 0.05	2020/2021	2019/2020
Average number of points	24.88596	22.87273
Frequency	114	165
Degrees of freedom	277	
t Stat	1.140953	
P value	0.254875	
t krit	1.968565	

Source: own calculation

Conclusion

The evaluation of tests written by distance form is significantly better than the evaluation of tests written in full-time form (see Table 2, Table 4, Figure 1 and Figure 2). For example, 50% of students obtained the maximum number of points from the mid-term test, which was realized in the distance form, while in the full-time form it was less than 25% of students (see Figure 1). However, there are no significant differences between the evaluation of the oral exams (see Table 7). The results of this paper show that oral examinations are becoming more and more important in present time, when the objectivity of exam tests written in a distance form is not always ensured.

Acknowledgment

The paper was processed with contribution of long term support of scientific work on Faculty of Informatics and Statistics, University of Economics and Business, Prague (IP 400040).

References

Akimov, A., Malin, M. (2020) When old becomes new: a case study of oral examination as an online assessment tool, *Assessment & Evaluation in Higher Education*, vol. 45, no 8, pp. 1205-1221. https://doi.org/10.1080/02602938.2020.1730301

Anděl, J. (1978) *Matematická statistika*. Praha: SNTL/ALFA.

Fluck, A. E. (2019) An International Review of eExam Technologies and Impact, *Computers & Education*, vol. 132, pp. 1–15. doi:10.1016/j.compedu.2018.12.008

Huxham, M., Campbell, F., Westwood, J. (2012) Oral versus Written Assessments: A Test of Student Performance and Attitudes, *Assessment & Evaluation in Higher Education*, vol. 37 (1), pp. 125–136. doi:10.1080/02602938.2010.515012

Joughin, G. (2007) Student Conceptions of Oral Presentations, *Studies in Higher Education*, vol. 32 (3), pp. 323–336. doi:10.1080/03075070701346873

Kaspříková, N., Klůfa, J. (2011) *Calculation of LTPD single sampling plans for inspection by variables and its software implementation, The 5th International Days of Statistics and Economics*, Prague, pp. 266-276. ISBN 978-80-86175-86-7.

Kehm, B. (2001) Oral Examinations in German Universities, *Assessment in Education,* vol. 8 (1), pp. 25–31. doi:10.1080/09695940120033234.

Klůfa, J. (2015a) Economic aspects of the LTPD single sampling inspection plans. *Agricultural Economics - Zemědělská ekonomika,* vol. 61, no 7, pp. 326-331. https://doi.org/10.17221/186/2014-AGRICECON

Klůfa, J. (2015b) Comparison of the ways of acceptance students at university. *Journal on Efficiency and Responsibility in Education and Science*, vol. 8, no 3, pp. 72-76. https://doi.org/10.7160/eriesj.2015.080304

Klůfa, J. (2015c) Dependence of the Results of Entrance Examinations on Test Variants. *Procedia - Social and Behavioural Sciences*, vol. 174, pp. 3565-3571. https://doi.org/10.1016/j.sbspro.2015.01.1073

Klůfa, J. (2016) *Analysis of the differences between results of test variants, Efficiency and Responsibility in Education, Proceedings of the 13th International Conference*, Prague, pp. 279–285.

Mesicek, L., Petrus, P. and Kovarova, K. (2017) *What can we learn from students' tests results in subject management? Efficiency and Responsibility in Education, Proceedings of the 14th International Conference*, Prague, pp. 247–254.

Okada, A., Mendonca, M. and Scott, P. (2015) Effective Web Videoconferencing for Proctoring Online Oral Exams: A Case Study at Scale in Brazil, *Open Praxis*, vol. 7 (3), pp. 227–242. doi:10.5944/openpraxis.7.3.215.

Rovai, A. P. (2000) Online and Traditional Assessments: What is the Difference?, *The Internet and Higher Education*, vol. 3 (3), pp. 141–151. doi:10.1016/S1096-7516(01)00028-8

Salkova, D., Sanova, P. and Regnerova, O. (2020) *University students' experience with professional social networks, Efficiency and Responsibility in Education, Proceedings of the 17th International Conference*, Prague, pp. 255–262.

Contact

Prof. RNDr. Jindřich Klůfa, CSc.
Prague University of Economics and Business,
Department of Mathematics,
W. Churchill Sq.4, 13067 Prague 3
klufa@vse.cz