GENDER ANALYSIS OF MATH EXAM RESULTS AT UNIVERSITY OF FINANCE AND ADMINISTRATION

Eva Ulrychová – Diana Bílková

Abstract

A student's success in an examination at a university may depend on various factors. It is debatable whether student's gender may be one of them. The gender gaps can depend on the subject of the examination. As far as mathematics is concerned, the tendency to accept gender stereotypes in favor of males may appear.

The aim of this paper is to analyze the results of the mathematics exam at the University of Finance and Administration in Prague, especially from the students' gender point of view. Two groups of students are included in the research – students who passed the mathematics exam in the academic year 2015/2016 and students who passed the exam in the academic year 2016/2017. For each of these years, the ratios of individual marks obtained by males and by females are compared and the conclusion is made: there are no substantial differences between the results of males and females in the tested groups.

Additionally, the students' results in the academic years 2015/2016 and 2016/2017 are compared regardless of students' gender. Students achieved better results in the year 2016/2017 than in 2015/2016.

Key words: mathematics, exam, mark, gender analysis

JEL Code: A22, I23, C12

Introduction

A student's performance during an examination may be affected by various factors. Apart from external conditions (different time of day, number of students in the classroom etc.) and individual personality traits, the exam result may also be influenced by other factors – by the field of study, see for example (Otavová & Sýkorová, 2015), the form of study – full-time or part-time (Joyce, Crockett, & Jaeger, 2015), student's gender etc., see for example (Ulrychová

& Bílková, 2016). Finally, the results may depend on the student's continuous preparation during the semester, see for example (Otavová & Sýkorová, 2014).

This paper analyzes the effect of student's gender on the result of mathematics exam at University of Finance and Administration in Prague. All students included in this research were students of the same field of study. No difference between full-time and part-time study is made because of the low number of part-time students.

The result of the exam could be affected by the teacher's personality and his/her teaching methods, see (Majovská, 2015), (Milková & Kořínek, 2015) or (Cho, Baek & Cho, 2015). For this reason, only the students who have studied and then passed the exam with the same teacher are included in the research.

However, according to the research (Falch & Naper, 2013), the gender grading gap is related to the characteristics of the teacher, suggesting that the teacher–student interaction during coursework favors girls in the teacher grading.

On the other hand, the literature related to gender issues in many cases indicates that teachers have different beliefs about male and female students and they tend to accept math-gender stereotypes in favor of males – see (Li, 1999).

According to research (Hyde, Fennema & Lamon, 1990), gender differences in mathematics performance are small in elementary or middle school, differences favoring males appear in high school and college. There are almost no differences among the general population.

There are different gender gaps in mathematics performance in various countries – see (Marks, 2008). It depends, among other things, on the organization of the school system and macro-societal factors. According to the 2000 knowledge survey of 15-year-olds in the principal industrialized countries, carried out by the Programme for International Student Assessment (PISA), females in all tested countries are on average better in reading literacy compared to males. Contrariwise, gender differences are smaller in mathematical and scientific literacy than they are in reading.

In this paper, the dependence of the exam result on the student's gender is tested in two groups – students who passed the mathematics exam in the academic year 2015/2016 and students who passed the exam in 2016/2017. In both the years 2015/2016 and 2016/2017, the ratios of marks A, B, C, D and E are calculated and the results of males and females are compared.

This paper has no ambition to make general conclusions in terms of gender impact on exam results. It is necessary to remember that a statistically valid comparison with international results cannot be completed due to the lack of knowledge of the demographic characteristics, field of study and the students' particular level of prior education, which undoubtedly affects the exam results (Exnarova, Dalihod &Mildeova, 2011).

In addition to the gender aspect, the difference between the years 2015/2016 and 2016/2017 is analyzed (independently of the students' gender) in this paper.

The approximate hypothesis test supposing the equation of the relative frequencies of the two alternative distributions was used for both analyses – for the gender analysis as well as for the comparison of results in the years 2015/2016 and 2016/2017.

1 Theory

Table 1 shows the sizes of individual analyzed sets. It is clear from this table that we can consider range sizes of all sets ($n \geq 30$). This means that we can use an approximate hypothesis test assuming the equation of the relative frequencies of the two alternative distributions.

Tab. 1: Sizes of individual analyzed sets

Year 2015/16		Year 2016/17		Years 2015/16 and 2016/17	
n_M	n_F	n_M	n_F	n_{2015}	n_{2016}
50	38	42	30	88	73

Source: Own research

We label n_M number of males, n_F number of females, m_M number of males with the mark A (or B, C, D, E) and m_F number of females with the mark A (or B, C, D, E). Similarly, n_{2015} and n_{2016} label the numbers of students in the years 2015/2016 and 2016/2017, respectively. Then the right-hand side alternative means that the proportion of males with the mark A (B, C, D, E) is statistically significantly higher than the proportion of females with the same mark.

The ratios

$$p_M = \frac{m_M}{n_M} \quad \text{and} \quad p_F = \frac{m_F}{n_F} \tag{1}$$

represent the ratios of males and females, respectively, with the mark A (B, C, D, E). The total ratio of students (males and females) with the mark A (B, C, D, E) is described by the formula

$$\overline{p} = \frac{m_M + m_F}{n_M + n_F}. \tag{2}$$

The statistic

$$U = \frac{p_M - p_F}{\sqrt{\overline{p} \cdot (1 - \overline{p})}} \cdot \sqrt{\frac{n_M \cdot n_F}{n_M + n_F}} \tag{3}$$

is the test criterion, which has asymptotically normal distribution supposing that the corresponding null hypothesis is valid. The null hypothesis supposes the equation H_0: $\pi_M = \pi_F$ (H_0: $\pi_{2015} = \pi_{2016}$), which means that the ratio of males with the mark A (B, C, D, E) is equal to the ratio of females with the corresponding mark A (B, C, D, E). The alternative hypothesis may be right-hand side or left-hand side according to the relation of corresponding ratio in Table 2. A right-hand side alternative means that the ratio of males with the mark A (B, C, D or E) is statistically significantly higher than the ratio of females with the corresponding mark at a given significance level, see Table 3.

Tab. 2: Relative frequencies

	Ratios								
	Year 2015/16			Year 2016/17			Years 2015/16 and 2016/17		
Mark	p_M		p_F	p_M		p_F	p_{2015}		p_{2016}
A	0.38	<	0.42	0.43	<	0.57	0.40	<	0.49
B	0.18	<	0.19	0.24	>	0.10	0.18	<	0.19
C	0.18	<	0.21	0.17	>	0.10	0.20	>	0.14
D	0.14	>	0.07	0.12	<	0.13	0.11	<	0.12
E	0.12	>	0.11	0.04	<	0.10	0.11	>	0.06
∑	**1.00**		**1.00**	**1.00**		**1.00**	**1.00**		**1.00**

Source: Own research

Tab. 3: Alternative hypotheses and the value of $\overline{p}$

	Year 2015/16		Year 2016/17		Years 2015/16 and 2016/17	
Mark	Alternative	$\overline{p}$	Alternative	$\overline{p}$	Alternative	$\overline{p}$
A	$H_1: \pi_M < \pi_F$	0.40	$H_1: \pi_M < \pi_F$	0.49	$H_1: \pi_{2015} < \pi_{2016}$	0.44
B	$H_1: \pi_M < \pi_F$	0.18	$H_1: \pi_M > \pi_F$	0.18	$H_1: \pi_{2015} < \pi_{2016}$	0.18
C	$H_1: \pi_M < \pi_F$	0.19	$H_1: \pi_M > \pi_F$	0.14	$H_1: \pi_{2015} > \pi_{2016}$	0.17
D	$H_1: \pi_M > \pi_F$	0.11	$H_1: \pi_M < \pi_F$	0.13	$H_1: \pi_{2015} < \pi_{2016}$	0.12
E	$H_1: \pi_M > \pi_F$	0.11	$H_1: \pi_M < \pi_F$	0.07	$H_1: \pi_{2015} > \pi_{2016}$	0.09

Source: Own research

Tab. 4: The value of test criterion and the conclusion of the test

	Year 2015/16				Year 2016/17				Years 2015/16 and 2016/17			
		Reject H_0 at α				Reject H_0 at α				Reject H_0 at α		
Mark	U	0.05	0.01	0.10	U	0.05	0.01	0.10	U	0.05	0.01	0.10
A	−0.62	no	no	no	−1.66	yes	no	yes	−1.83	yes	no	yes
B	−0.12	no	no	no	3.53	yes	yes	yes	−0.14	no	no	no
C	−0.82	no	no	no	2.16	yes	no	yes	2.30	yes	no	yes
D	2.65	yes	yes	yes	−0.51	no	no	no	−0.55	no	no	no
E	0.64	no	no	no	−3.27	yes	yes	yes	3.21	yes	yes	yes

Source: Own research

The critical value is $u_{0.95} = 1.645$ in the case of the right-hand side alternative and $u_{0.05} = -1.645$ in the case of the left-hand side alternative at $\alpha = 0.05$ significance level; $u_{0.99} = 2.326$ in the case of the right-hand side alternative and $u_{0.01} = -2.326$ in the case of the left-hand side at $\alpha = 0.01$ significance level and $u_{0.90} = 1.282$ in the case of the right-hand side alternative and $u_{0.10} = -1.282$ in the case of the left-hand side alternative at $\alpha = 0.10$ significance level.

2 Results

In the case of the marks A, B, C and E in 2015/16, D in 2016/17 and B and D for 2015/16 and 2016/17 together, we do not reject the null hypothesis which assumed the equality of ratios of males and females with the same mark even at 10 percent significance level. This means that the differences between the ratios of males and females with the marks mentioned above are not statistically significant even at 10 percent level of significance. Contrarily, in the case of the mark D in 2015/16, B and E in 2016/17 and E in 2015/16 and 2016/17 together, these

differences are statistically significant even at 1 percent significance level. In the case of the marks A and C in 2016/17 and in 2015/16 and 2016/17 together, these differences are not statistically significant only at 1 percent significance level, while at 5 and 10 percent level of significance they are statistically significant.

Conclusion

With respect to the gender aspect, the results of males and females are very similar in several cases (see Table 2), especially the mark B and E in 2015/2016 and D in 2016/2017. Not only in these cases, the effect of gender on the results is not statistically proved. The most distinct difference between males and females is proved in the case of the mark D in 2015/2016 (in favor of males) and the mark E (in favor of females). On the other hand, the result in the corresponding cases – D in 2016/2017 and E in 2015/2016 – is the opposite (D in favor of females and E in favor of males), though the difference between males and females is negligible and not statistically significant. In the case of the marks A and C in 2016/2017, the differences are statistically significant at 5 percent level of significance – A in favor of females and C in favor of males (not so in 2015/2016).

It is interesting (though not statistically significant) that only in the case of the mark A, the inequality between ratios of males and females is the same in 2015/2016 and 2016/2017 (both in favor of females). In other cases, the individual inequalities in 2015/2016 and 2016/2017 are opposite to each other (see Table 2).

A conclusion can be drawn: overall there are not very substantial differences between the results of males and females in the tested groups.

With respect to comparison of both the years 2015/2016 and 2016/2017 (regardless of gender), the difference between these years is most statistically significant in the case of the mark E (in favor of 2015/2016), less in A (in favor of 2016/2017) and C (in favor of 2015/2017). In the cases B and D, the differences are very small and are not statistically significant. Comparing the results of A (the best rating) and E (the worst rating) in 2015/2016 and 2016/2017, the conclusion can be made: students achieved better results in 2016/2017 than in 2015/2016.

Acknowledgment

The paper is a result of institutional research project no. 7429/2018/08 "Analysis of ICT startups" supported by University of Finance and Administration, Prague, Czech Republic.

This paper was subsidized by the funds of institutional support of a long-term conceptual advancement of science and research number IP400040 at the Faculty of Informatics and Statistics, University of Economics, Prague, Czech Republic.

References

Cho, D., Baek, W. & Cho, J. (2015). Why do good performing students highly rate their instructors? Evidence from a natural experiment. In *Economics of Education Review, 49 (December 2015)*, 172-179.

Exnarova, A., Dalihod, M. & Mildeova, S. (2011). Measuring systems thinking ability. In *Proceedings of the 8th International Conference on Efficiency and Responsibility in Education (ERIE 2011)*, Prague, 66-74.

Falch, T. & Naper, L. R. (2013). Educational evaluation schemes and gender gaps in student achievement. In *Economics of Education Review, 36 (October 2013)*, 12-25.

Hyde, J. S., Fennema, E. & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. In *Psychological Bulletin, 107*(2), 139-155.

Joyce, T., Crockett, S. & Jaeger, D. A. (2015). Onur Altindag, Stephen D. O'Connell (2015). Does classroom time matter? In *Economics of Education Review*, *47 (June 2015)*, 64-77.

Li, Q. (1999). Teachers' beliefs and gender differences in mathematics: a review. In *Educational research, 41 (1999)*, 63-76.

Majovská, R. (2015). Open Source Programmes as the Support in Mathematical Education. In *Proceedings of the 12th International Conference on Efficiency and Responsibility in Education (ERIE 2015)*, Prague, 341-347.

Marks, G. N. (2008). Accounting for the gender gaps in student performance in reading and mathematics: evidence from 31 countries. In *Oxford Review of Education , 34 (2008)*, 89-109.

Milková, E. & Kořínek, O. (2014). Future ICT Teachers - Programming Aptitude. In *Proceedings of the 11th International Conference Efficiency and Responsibility in Education (ERIE 2014)*, Prague, 456-462.

Otavová, M. & Sýkorová, I. (2014). Analysis of Scores from Mid-Term and Final Test by a Contingency Table. In *Proceedings of the 11th International Conference on Efficiency and Responsibility in Education (ERIE 2014)*, Prague, 527-533.

Otavová, M. & Sýkorová, I. (2015). Analysis of Differences in the Scores of Students from Different Faculties. In *Proceedings of the 12th International Conference on Efficiency and Responsibility in Education (ERIE 2015)*, Prague, 422-429.

Ulrychová, E. & Bílková, D. (2016). Analysis of Dependence between the Results of Marking of Selected Compulsory Courses at University of Finance and Administration. In *10th International Days of Statistics and Economics*, Prague, 1926-1937.

Contact

RNDr. Eva Ulrychová, Ph.D.
University of Finance and Administration
Faculty of Economic Studies
Department of Informatics and Mathematics
Street Estonian 500/3
101 00 Prague 10
Czech Republic
Mail: ulrychova@mail.vsfs.cz

doc. Ing. Diana Bílková, Dr.
University of Economics, Prague
Faculty of Informatics and Statistics
Department of Statistics and Probability
Sq. W. Churchill 1938/4
130 67 Prague 3
Czech Republic
Mail: bilkova@vse.cz

doc. Ing. Diana Bílková, Dr.
University of Finance and Administration
Faculty of Economic Studies
Department of Informatics and Mathematics
Street Estonian 500/3
101 00 Prague 10
Czech Republic
Mail: diana.bilkova@vsfs.cz