Ordinal Logistic Regression Analysis of Factors Affecting the Length of Student Study

Baiq Rika Ayu Febrilia[1], Suning Rahayu[2], Baiq Dewi Korida[3]

Institut Keguruan dan Ilmu Pendidikan Mataram, *rikafebrilia@ikipmataram.ac.id[1]*
Institut Keguruan dan Ilmu Pendidikan Mataram, *suningrahayu1198@gmail.com[2]*
Institut Keguruan dan Ilmu Pendidikan Mataram, *dewikorida12@gmail.com[3]*

Abstrak: Lama waktu seorang mahasiswa untuk menyelesaikan masa studi menjadi ukuran prestasi mahasiswa tersebut dan kesuksesan program studinya. Oleh karena lama masa studi cukup berpengaruh terhadap kualitas suatu program studi dan proses pembelajaran di dalamnya, maka perlu dilakukan studi lebih mendalam mengenai faktor-faktor yang mempengaruhi lama studi mahasiswa. Penelitian ini bertujuan untuk memodelkan faktor-faktor yang mempengaruhi lama masa studi mahasiswa menggunakan regresi logistik ordinal. Faktor-faktor tersebut adalah IPK mahasiswa dan jenis kelaminnya. Data mengenai lama studi, IPK dan jenis kelamin diambil untuk mahasiswa Program Studi Pendidikan Matematika IKIP Mataram yang lulus pada tahun 2017 dan 2018. Hasil penelitian menunjukkan bahwa kedua faktor berpengaruh signifikan terhadap lama masa studi mahasiswa.

Kata kunci: regresi logistik ordinal, masa studi

Abstract: The length of time a student completes the study period is a measure of the student's achievement and the success of his study program. Because the duration of the study period is quite influential on the quality of a study program and the learning process in it, it is necessary to do a more in-depth study of the factors that influence the duration of student studies. This study aims to model the factors that influence the length of the study period of students using ordinal logistic regression. These factors are the student's GPA and gender. Data on the length of study, GPA and gender were taken for students of the Mathematics Education Department, IKIP Mataram, who graduated in 2017 and 2018. The results of the study showed that the two factors had a significant effect on the length of the study period.

Keywords: ordinal logistic regression, length of the study

1. Introduction

The duration of the student's study period is determined by how long a student can complete all his obligations in college until he is entitled to a bachelor's degree. The duration of the study period determines the success of learning at the college level, especially at the smallest unit level, especially on the study program [1]. The length of time a student completes the study period is also a measure of student achievement [2]. Under normal conditions, students must be able to complete their studies at the college level for eight semesters [3]. Thus, if there is a study program with an average length of the study period of more than eight semesters or four years, then the student achievement in the study program is not so good [4]. The duration of the study period is also one of the very important factors in determining the accreditation of study programs [5]. Therefore, managers of study programs need to pay special attention to the factors that affect the length of time student's study.

There are many factors that can affect the length of time student's study. Two factors include the Grade Point Average (GPA) and student gender. GPA is one of the determinants of one's study period [6] because the value of student achievement, in this case, the IP / GPA value, determines how many Semester Credit Systems (SKS) can be taken by the student. The higher the GP / GPA of a student, the student is allowed to take more credits. Thus, students with high GP / GPA will be able to shorten their study period or graduate for less than four years. In terms of student gender, female students are known to pay more attention to their study process than male students. Female students are more disciplined with time and they have sufficiently high motivation to be able to complete the study period on time. One of the results of the study also shows that female students' academic performance is better than male students [7]. Likewise, for learning motivation, female students have higher learning motivation than male students [8]. Therefore, gender is also a long-standing determinant of student studies. To determine whether GPA and gender are factors that significantly influence the length of a person's studies, further and more detailed statistical analysis is needed.

A review of the factors that influence the length of the study period can be analyzed using regression analysis. Regression analysis is a tool in statistics to show a causal relationship between two or more variables, where one variable acts as a response variable and another variable as a predictor variable so that one variable can be predicted through another variable [2]. Regression analysis can be either linear regression or nonlinear regression. Because the data used is categorical or ordinal data, then the regression analysis that can be used is ordinal logistic regression [9]. Previous research shows that ordinal logistic regression analysis can provide information that the factors that influence the predicate of student graduation are faculty, gender, admission path, father's job, mother's work, and income [2]. Using the same analysis, [3] indicates that the student graduation time is influenced by the department taken and its gender and provides a good model in explaining the relationship between the three variables. In contrast to previous research, this study will analyze the extent of the influence of GPA and student gender on the duration of student studies

This study aims to describe the old characteristics of the student's study period, and the model is based on the factors that influence it. The results of this study are expected to contribute knowledge to managers of study programs regarding the study period of students as well as a reflection in improving the management system that is at the level of study programs.

2. Ordinal Logistic Regression

In the regression model, if the response variable is binary or dichotomous, then the analysis will use a logistic regression model, which in its parameter estimation must use a

connecting function [10]. Ordinal logistic regression has the characteristics of response variables and predictors that have an ordinal scale. Suppose that the Y response variable has k categories with $x_i = (x_{i1}, x_{i2}, \ldots, x_{ip})$ is a vector of predictor variables, p, at the value of the first observation with $i = 1,2,\ldots,n$. Opportunities of the k-response variable with predictor X are expressed as $P[Y \leq k|x]$ [11]. Cumulative opportunities are as follows [12].

$$P[Y \leq k|x] = \pi(x) = \frac{\exp\left(\beta_{0k} + \sum_{r=1}^{p} \beta_r x_{ir}\right)}{1 + \exp\left(\beta_{0k} + \sum_{r=1}^{p} \beta_r x_{ir}\right)}$$

The cumulative logit model is defined using the following formula.

$$
\begin{aligned}
logit\ (P[Y \leq k|x]) &= \log\left(\frac{P[Y \leq k|x]}{P[Y > k|x]}\right) \\
&= \log\left(\frac{P[Y \leq k|x]}{1 - P[Y \leq k|x]}\right) \\
&= \beta_{0k} + \sum_{r=1}^{p} \beta_r x_{ir}
\end{aligned}
$$

which are $k = 1,2,\ldots,K-1$ and β are vector regression coefficients.

3. Research Methods

This study uses secondary data obtained from the academic section of the Mathematics Education Study Program, IKIP Mataram. This type of research is a quantitative type with a case study of factors that influence the duration of study of undergraduate students in the Mathematics Education Study Program year passed 2017 and 2018 through an ordinal logistic regression approach. The variables used in this study were divided into two types, namely the duration of the study period as the response variable (Y) and GPA and gender as predictor variables (X_1 and X_2). This response variable consists of three categories, namely:

$$1 = Y < 4 \text{ years}$$
$$2 = 4 \geq Y \leq 4 \text{ years 6 months}$$
$$3 = Y > 4 \text{ years 6 months.}$$

The 4 years 6 months is adjusted to the second period of graduation at the IKIP Mataram. This period is still assumed to be a normal category.

Each predictor variable that will be used in this study consists of several categories, such as the following.

$x_{1,i}$ = Student GPA,

which is i:　　$1 = GPA > 3{,}50$

　　　　　　$2 = 3{,}00 < GPA \leq 3{,}50$

　　　　　　$3 = GPA \leq 3{,}00$

$x_{2,i}$ = gender

which is i:　　$1 = $ male

　　　　　　$2 = $ female.

This study consists of several steps. The first step is to make cross-tabulations to analyze student characteristics. The next step is followed by estimating the ordinal logistic regression parameter. Then test the significance of the parameters simultaneously and partially. After that, the test of model suitability was followed by concluding the results of the study.

4. Results and Discussion

Before estimating parameters, the collected data were analyzed for characteristics based on the length of the study period on the GPA and length of study period by gender.

Table 1. Characteristics based on the length of the study period and student GPA

Student GPA range	Length of study			Total
	$Y < 4$ year	$4 \geq Y \leq 4$ years 6 months	$Y > 4$ years 6 months	
$GPA > 3,50$	2 (28.6%)	3 (42.8%)	2 (28.6%)	7 (100%)
$3,00 < GPA \leq 3,50$	1 (0.95%)	42 (40%)	62 (59.05%)	105 (100%)
$GPA \leq 3,00$	0 (0%)	0 (0%)	22 (100%)	22 (100%)

Based on Table 1, information was obtained that the characteristics of students of Mathematics Education Study Program in the years of 2017 and 2018 had the most study period of more than 4 years 6 months and were dominated by graduates with GPA in the range of 3.00 to 3.50. This shows that the study period of Mathematics Education Study Program students in graduating 2017-2018 is quite long. This is possible due to a number of things such as, students having difficulty in receiving the given lecture material or the final project being worked on, students have external problems, such as family problems or lack of tuition fees which require them to work while studying.

Table 2. Characteristics based on the length of the study period and student gender

Gender	Length of study			Total
	$Y < 4$ year	$Y < 4$ year	$Y < 4$ year	
Male	0 (0%)	8 (15,4%)	44 (84,6%)	52 (100%)
Female	3 (3,66%)	37 (45,12%)	42 (51,22%)	82 (100%)

Table 2 explains that the characteristics of the students of Mathematics Education Study Program in the year of 2017 and 2018 in terms of length of the study period and gender are more than those who studied for more than 4.6 years and were dominated by male students at that time. Female students dominated more during the study period under 4 years or in the range of 4 to 4 years 6 months. This phenomenon is possible because usually female students pay more attention to the study process than male students.

Then the parameter estimation of the ordinal logistic regression model is performed using the help of SPSS software. The estimation results can be seen in more detail in Table 3 below.

Table 3. Parameter estimation of ordinal logistic regression model

Variable	Category	Coefficient (β)	Standard Error Coefficient	Exp (β)	P-value
Length of study	Constant 1	-22,158	0,633		0,000
	Constant 2	-18,540	0,245		0,000

Variable	Category	Coefficient (β)	Standard Error Coefficient	Exp (β)	P-value
Range of students GPA	$GPA > 3,50$	-20,349	0,907	1,454	0,000
	$3,00 < GPA \leq 3,50$	-18,632	0,000	8,095	.
	$GPA \leq 3,00$	0	.	1	.
Gender	Male	1,455	0,467	4,284	0,002
	Female	0	.	1	.

Table 3 shows there are two constant values obtained in the model because there are three response variables that have coefficient values. This will produce two logit models. Table 3 also shows the odds ratio ($Exp(\beta)$) of students who graduate with a GPA from 3.00 to 3.50 is at 8,095. This shows that students with a GPA of 3.00 to 3.50 have an opportunity of 8.095 times greater than other GPA categories. Based on gender, students with male gender have an odds ratio of 4.284 or in other words, male students have a chance of 4.284 times greater than women. Next will be the suitability test of the model and the significance of the model both simultaneously and partially.

Model suitability test is needed to find out whether the model obtained is appropriate or not. The results of this test are presented in Table 4. This test uses the null hypothesis (H0) and the alternative hypothesis (H1) as follows.

H_0 : Model is fit

H_1 : Model is not fit

Table 4. Test goodness of fit model using *Deviance*

	Chi-Square	df	P-value	Decision
Deviance	3,676	5	0,597	Accepted H_0

Table 4 shows that the P-value = 0.597 > 0.05, then the decision is H_0 accepted or it can be concluded that the ordinal logistic regression model obtained is appropriate.

Table 5. Model test using Likelihood Ratio

Model	G^2	*Chi-Square*	df	P-value	Decision
Intercept only	57,000				
Final	18,414	38,586	3	0,000	Rejected H_0

To test the meaning of β coefficients together, simultaneous testing was carried out on the ordinal logistic regression model. This test is carried out using the G^2 Test or also known as the Likelihood Ratio Test. Based on Table 4, it is known that the value of $G^2 > \chi_{(0.05;3)} = 7.815$, which means that H0 is rejected or in other words, the β coefficient is significant for the ordinal logistic regression model. This result is also in accordance with the P-value = 0.000 < 0.05.

Partial testing using the Wald test is intended to determine whether the variables in the ordinal logistic regression model are significant or not. This test uses the following statistical hypothesis.

$H_0 : \beta_k = 0$

$H_1 : \beta_k \neq 0, k = 1, 2$

Partial test results can be seen in Table 3 of the sixth column through the P-value value. Through Table 3, it is known that all variables have variable values less than 0.05. These

results indicate that these variables are significant.

Based on the results of data analysis, the logit model of the length of the study period in terms of student GPA and student gender as many as 2 (two) models, as follows.

$$Logit\ (y_1) = log\left(\frac{y_1}{1 - y_1}\right) = -22{,}158 + (-20{,}349X_{1,1}) + (1{,}455X_{2,1})$$

$$Logit\ (y_2) = log\left(\frac{y_2}{1 - y_2}\right) = -18{,}540 + (-20{,}349X_{1,1}) + (1{,}455X_{2,1})$$

5. Conclusion

Based on the results of the research and discussion, it was found that the characteristics of the students of the Mathematics Education Study Program were dominated by graduates with a GPA from 3.00 to 3.50, but the length of the study period of students was more than 4 years 6 months. The length of the study period for more than 4 years and 6 months is dominated by male students. The test results simultaneously and partially indicate that the factors that affect the length of the study period are GPA and gender.

References

[1]   Z. Zakariyah and I. Zain, "Analisis Regresi Logistik Ordinal pada Prestasi Belajar Lulusan Mahasiswa di ITS Berbasis SKEM," *J. Sains dan Seni ITS*, vol. 4, no. 1, pp. 121–126, 2015.

[2]   S. Imaslihkah, M. Ratna, and V. Ratnasari, "Analisis regresi logistik ordinal terhadap faktor-faktor yang mempengaruhi predikat kelulusan mahasiswa S1 di ITS Surabaya," *J. Sains dan Seni Pomits ISSN 2337-3520 (2301-928X Print)*, vol. 2, no. 2, pp. 177–182, 2013.

[3]   M. W. Talakua, A. Ratuanak, and V. Y. I. Ilwaru, "Analisis regresi logistik ordinal terhadap faktor-faktor yang mempengaruhi waktu kelulusan mahasiswa S1 di FMIPA Unpatti Ambon tahun 2016 dan 2017," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 13, no. 1, pp. 033–038, 2019.

[4]   D. Samuel, "Pengaruh budaya belajar, strategi pengajaran dosen, dan motivasi belajar terhadap hasil belajar mahasiswa Pendidikan Ekonomi FKIP UKSW," *J. Educ. Soc. Stud.*, vol. 3, no. 2, pp. 6–12, 2014.

[5]   U. Zulfa and N. Kurniasih, "Efektivitas strategi percepatan penyelesaian Program S1 Fakultas Tarbiyah IAIIG Cilacap," *J. Tawadhu*, vol. 1, no. 2, pp. 167–218, 2017.

[6]   D. Samekto, H. A. Syafrudie, and S. Sutrisno, "Kecenderungan lama studi dan prestasi belajar mahasiswa jalur reguler dan non-reguler Program Studi Pendidikan Teknik Bangunan," *Teknol. dan Kejuru.*, vol. 37, no. 2, pp. 153–166, 2014.

[7]   S. Nuryoto, "Perbedaan prestasi akademik antara laki-laki dan perempuan studi di wilayah Yogyakarta," *J. Psikol.*, no. 2, pp. 16–24, 1998.

[8]   I. W. Anita, "Pengaruh motivasi belajar ditinjau dari jenis kelamin terhadap kemampuan berpikir kritis matematis mahasiswa," *J. Ilm. UPT P2M STKIP Siliwangi*, vol. 2, no. 2, pp. 246–251, 2015.

[9]   L. Fatonah, S. Sanapiah, and B. R. A. Febrilia, "Regresi logistik ordinal (studi kasus faktor-faktor yang mempengaruhi tingkat stres mahasiswa dalam menyelesaikan skripsi di IKIP Mataram)," *Media Pendidik. Mat.*, vol. 5, no. 2, pp. 206–216, 2017.

[10]   E. Yulian and G. Pawitan, "Pemodelan status usaha (pengusaha dan

pekerja/karyawan) menggunakan regresi logistik multilevel," *J. Mat. "MANTIK,"* vol. 3, no. 1, pp. 30–38, 2017.

[11] D. W. Hosmer and S. Lemeshow, *Applied logistic regression second edition*. Canada: John Wiley & Sons, Inc., 2000.

[12] A. Agresti, *Categorical data analysis 2nd ed*, vol. 45, no. 1. 2002.