International Journal of Applied Mathematics and Computing
Volume. 2, Nomor.1, 2025
e-ISSN: 3047-146X; Page 35-45
DOI ; https://doi.org/10.62951/ijamc.v2i1.92
Available online at: https://international.arimsi.or.id/index.php/IJAMC

Factors Influencing the Study Period of Students of the Mathematics Study Program at Udayana University

Ulfatun Farika Novitasari[1], Adinda Audy Sita Mayzandy[2], Miltiades Dewifortuna Pulo[3], Juliani Tandi Tumbiri[4], Nurul Ilma[5], Made Susilawati[6]
[1-6] Mathematics Study Program, Faculty of Mathematics and Natural Sciences, Udayana University, Indonesia
Email : ulfarika23@gmail.com[1], mayzandy05@gmail.com[2], dewifrtna123@gmail.com[3], julianitumbiri@gmail.com[4], ilmanurul497@gmail.com[5], mdsusilawati@unud.ac.id[6]

Corresponding Author: ulfarika23@gmail.com

Abstract The student study period is one of the important aspects to measure the quality of a higher education institution. The length of the student study period can be assumed to come from internal factors and external factors, so it is necessary to conduct research that aims to identify and model the factors that affect the student study period. The method used in this research is logistic regression and the data used is primary data obtained from distributing questionnaires. The target of this research is aimed at alumni students from the Mathematics Study Programme of Udayana University from 2011 to 2019. In this study, the best model produced has a classification accuracy of 98.17% and the independent variables that have a significant effect on the study period are gender, tuition fees and interest in majors.

Keywords: Logistic, Regression, Gender, Student

1. INTRODUCTION

The student's study period is the period of time required by a student to complete all academic requirements in a particular study program. The length of the study period can be influenced by various factors, both internal and external. These factors can have an impact on academic success and the quality of education.

One of the factors that can influence an education sector is Human Resources (HR). To obtain HR who have social opportunities and a better quality of life, the state should make education a top priority by prioritizing education will create prosperous HR (Wulandari et al., 2022).

Developing human resources through higher education is not an easy challenge. Various supporting factors in education are very important to create a conducive learning environment and provide the best opportunities for students to achieve academic success. Some of the main factors that support higher education include the quality of facilities and infrastructure, teacher competence, quality of education, and student intelligence and abilities.

Students are one aspect to measure the quality of a University. Assessment from the student aspect is the new student intake system for undergraduate programs, the average study period, and the Cumulative Achievement Index (GPA) obtained. Therefore, the punctuality of student graduation is one of the determining aspects of the quality of a University.

Received: September 15, 2024 Revised: September 28, 2024 Accepted: November 16, 2024 Online Available: November 19, 2024

The indicator of student success in obtaining a bachelor's degree is by graduating on time. Students can be said to have graduated on time if they complete their study period during $\leq$ 4 the year, and can be said not to have graduated on time if they complete their study period $>$ 4in the year.

The study period and GPA are interrelated. In general, a student who gets a low GPA usually takes longer to complete their study period, because the number of Semester Credit Systems (SKS) is taken based on the GPA obtained (Yelfera et al., 2022) . A low GPA means that students are only allowed to take low credits. As a result, students have to spend time improving unsatisfactory grades or taking credits that are still lacking. This can have an impact on the length of study that students must take.

Factors that influence students' study period can be classified into internal and external factors. Internal factors include students' self-motivation, study skills, and time management. While external factors include social support from family, interactions with lecturers, and resources available on campus. Previous research has shown that academic and social factors play a major role in determining timely graduation. For example, positive relationships with lecturers and effective academic support can help improve student performance and accelerate their study period (Gude et al., 2017). In addition, demographic factors such as age and employment status also have a significant influence, where students who work part-time are more likely to extend their study period due to limited time to study (Sverdlik et al., 2018). Robbins et al. (2022) added that psychosocial skill factors, such as emotional regulation and motivation, also contribute to accelerating graduation. By understanding these factors, universities can provide more specific support to help students complete their studies within the specified deadline.

One of the methods used to identify factors that affect the study period of students in the Mathematics Study Program at Udayana University is regression analysis. Regression analysis is an analysis technique that explains the form of the relationship between two or more variables that contain cause and effect. One method of regression analysis is logistic regression. Logistic regression, known as the logistic model or logit model, is used to predict the probability of an event based on data with the logit function of the logistic curve (Agresti, 2018).

This study used the binary logistic regression method. Binary logistic regression is a statistical method used to model the relationship between one or more independent variables (predictors) with a binary response variable (two categories), such as the outcome "yes" or

"no". This method is based on the logit transformation of the probability of an event to predict the response variable more accurately in cases with non-linear relationships (Hosmer et al, 2013). Parameter estimation is carried out using the *Maximum Likelihood Estimation* (MLE) method, which seeks predictor coefficients that maximize the fit of the observed data to the model (Agresti, 2018). This model requires the fulfillment of certain assumptions, including independence of observations, the absence of perfect multicollinearity between predictors, and the existence of a linear relationship between continuous variables and *the log-odds* of the response variable (*BMJ Family Medicine and Community Health* , 2023). Recent research results show that binary logistic regression is very flexible in analyzing binary data in various fields, from health to education (Sverdlik et al., 2018)

This study aims to determine and model the factors that influence the study period of students at the Mathematics Study Program at Udayana University using binary logistic regression.

2. RESEARCH METHODS

In this study. the data used is primary data, obtained from the distribution of questionnaires containing factors that influence the study period of undergraduate (S1) Mathematics Study Program students. The respondents of this study were students of the Mathematics Study Program of Udayana University from 2011 to 2019 who had completed their study period in this case they had graduated. The following are the research variables used.

Table 1. Variables and Categories Used in the Research

No.	Variables	Variable Name	Category	Dummy
1	Y	Study Period	$> 4\ tahun$ $\leq 4\ tahun$	0 1
2	X_1	Gender	Man Woman	0 1
3	X_2	UKT	Full Scholarship UKT 1 UKT 2 UKT 3 UKT 4 UKT 5	0 1 2 3 4 5
4	X_3	GPA	< 2.00 $2.00 - 3.00$ > 3.00	0 1 2
5	X_4	Scholarship grantee	No Yes	0 1
6	X_5	Working While Studying	No Yes	0 1
7	X_6	Residence	Home (parents/siblings) Boarding house	0 1
8	X_7	Guidance System	Offline On line Hybrid	0 1 2
9	X_8	Organizational activity	Not active Active	0 1
10	X_9	Interest in Major	Do not like Just Like Really like	0 1 2
11	X_{10}	Entry Path	SNMPTN SBMPTN Independent Test Other	0 1 2 3
12	X_{11}	Choice of Majors	First Choice Second Option	0 1

Logistic Regression Method

The logistic regression method is a regression analysis used to see the relationship between categorical dependent variables based on one or more independent variables (Varamita, 2017).

The logistic regression model can be written in the following equation 1 (Homser aet al, 2013).

$$\pi(x) = \frac{exp(\beta_0 + \beta_1 x_1 + \beta_1 x_2 + \cdots + x_k x_k)}{1 + exp(\beta_0 + \beta_1 x_1 + \beta_1 x_2 + \cdots + \beta_k x_k)}$$

Where $\pi(x)$(odds) is the chance of a "successful" event occurring, namely Y = 1 with a probability value of, $0 \leq \pi(x) \leq 1$.

Logistic regression has the advantage of not having the assumption of normality of its error model. In binary logistic regression, the predictor variables owned can be of mixed scale, namely continuous variables and discrete variables.

The following are the stages in conducting data analysis in this study using SPSS 25 *software* .

1. Conduct descriptive analysis on data to find out the picture of the data obtained
2. Conducting multicollinearity test using VIF test
3. Carrying out the logistic regression method by conducting parameter tests, namely:
   a. Perform parameter testing simultaneously
   b. Perform partial parameter testing
4. Conduct a Wald test to determine the partial significance of the parameters.
5. Conducting a model feasibility test using the *Hosmer and Lameshow test*
6. Forming a logistic regression model
7. Calculating the results of classifying data on the length of study period of Mathematics Study Program students using the *APER test* .
8. Performing interpretation

3. RESULTS AND DISCUSSION

The data to be analyzed in this study were 59 data obtained through the distribution of questionnaires to students of the Mathematics Study Program who had completed their study period. Data processing in this study used SPSS. The results of the descriptive statistical analysis can be seen in Table 2 below.

Table 2. Student Study Period

	Frequency	Percentage
$\leq 4\ Tahun$	21	35%
$> 4\ Tahun$	39	65%
Total	60	100

Source: Processed data, 2023

Based on Table 2, the data obtained amounted to 59 data with 35% of students completing their study period in less than or equal to 4 years, and 60% of students completing their study in more than 4 years.

Next, conducting a multicollinearity test using the VIF test, the results obtained are in Table 3 below.

Table 3. VIF test results

Variables	Tolerance	VIF
X_1	0,841	1,189
X_2	0,579	1,726
X_3	0,840	1,190
X_4	0,534	1,874
X_5	0,754	1,326
X_6	0,869	1,151
X_7	0,850	0,177
X_8	0,880	1,136
X_9	0.867	1,153
X_{10}	0.779	1,252
X_{11}	0.668	1,497

Based on Table 3, it can be seen that there is no multicollinearity, which can be seen from the VIF value for the independent variable being less $< 10$and *the Tolerance value* $> 0{,}100$.

Then, a simultaneous parameter significance test is performed. The hypothesis for the simultaneous parameter significance test is as follows:

$H_0 = \beta_i = 0$, with $i = 1{,}2, \cdots, 11$(there is no significant influence of the independent variables $i$simultaneously on the length of the study period)

H_1 =there is at least one $\beta_i \neq 0$, with $i = 1{,}2, \cdots, 11$(there is a significant influence of the independent variable $i$simultaneously on the length of the study period).

The decision rule used is if it is a value $-2\ Log\ likelihood < \chi^2_{0{,}05,11}$then reject H_0it and vice versa.

The results of this test are shown in Table 4 below.

Table 4. Results of Simultaneous Parameter Significance Test

Step	$-2\ Log\ likelihood$
1	44,215

Source: Processed data, 2023

Based on Table 4, it can be seen that the value of $-2\ Log\ likelihood = 44{,}215$. The chi-square value obtained is $\chi^2_{0{,}05,11} = 19{,}68$, so the value is $-2\ Log\ likelihood = 44{,}215 < \chi^2_{0{,}05,11} = 19{,}68$ then the decision is rejected H_0, which means that there is at least one independent variable that has a significant effect on the dependent variable.

Then a partial parameter significance test is performed. The hypothesis in the partial parameter significance test is as follows:

$H_0 = \beta_i = 0$, with $i = 1{,}2, \cdots, 11$(there is no significant influence of the independent variable $i$on the length of the study period)

$H_0 = \beta_i \neq 0$, with $i = 1{,}2, \cdots, 11$(there is a significant influence of the independent variable $i$on the length of the study period).

The decision rule used is if the Sign result $< \alpha = 0{,}05$or Wald test result $> \chi^2_{0{,}05,1}$ means that the relevant variable does not have a significant effect on the model.

The results of this test are shown in the following table:

Table 5. Results of Partial Parameter Significance Test

Variables	Wald	Sign.
X_1	8,021	0,005
X_2	6,196	0,013
X_3	0,007	0,933
X_4	1,751	0,186
X_5	0,144	0,705
X_6	0,199	0,655
X_7	2,279	0,131
X_8	0,000	0,999
X_9	7,936	0.005
X_{10}	1,604	0.205
X_{11}	0.636	0.425
Constants	0,00	0,999

Source: Processed data, 2023

Based on Table 5 obtained using the Wald test, it was found that X_1, X_2, X_9greater than the value $\chi^2_{0{,}05,1}$ $(3{,}841)$so that the decision is to reject H_0. Thus, it can be concluded that there is a significant influence on the variables X_1, X_2, X_9on the study period.

Next, a model feasibility test was conducted using the *Hosmer and Lameshow test* . The hypotheses used are as follows:

H_0: The hypothesized model fits the data

H_1: The hypothesized model does not fit the data

Decision rule if the Sig value. in the table$\geq 0{,}05$ then H_0it is accepted and the hypothesized

model fits the data or in other words the model obtained is feasible.

The results of the *Hosmer and Lameshow test* can be seen in Table 6 below.

Table 6. Results of the Model Suitability Test (*Goodness of Fit*)

Step	*Chi-square*	df	Sig.
1	3,219	8	0,920

Source: Processed data, 2023

Based on Table 6, the sig. value is obtained $0{,}920 \geq 0{,}05$and the *chi-square value* is obtained $3{,}2198 \leq \chi^2_{0{,}05{,}4} = 9{,}49$. Thus, it can be said that there is no difference between the model and the data so that the model obtained is feasible.

The following is the formation of a logistic regression model which is shown in Table 7 below.

Table 7. Estimation of Logistic Regression Model Parameters

Variables	β	$Exp\ (\beta))$
X_1	2,847	17,230
X_2	1,012	2,751
X_3	−0,131	0,877
X_4	1,590	4,906
X_5	0,427	1,533
X_6	0,394	1,483
X_7	−0,919	0,399
X_8	22,458	56,685
X_9	2,707	14,992
X_{10}	−0,808	0,446
X_{11}	0,808	2,244
Constants	−31,507	0,000

Source: Processed data, 2023

Based on Table 7, the logistic regression model obtained is as follows:

$$g(x) = ln\ ln\ \left(\frac{\pi(x)}{1-\pi(x)}\right) = -31{,}507 + 2{,}847X_1 + 1{,}012X_2 + 2{,}707X_9$$

The variables that have a significant influence are as follows:

X_1= gender

X_2= UKT

X_9= interest in the major

The value $Exp\ (\beta)$of the gender variable (X_1) is 17.230, which means that female students are 17.230 times more likely to graduate on time than male students in terms of length of study. The value $Exp\ (\beta)$of the variable, X_2namely UKT, is 2.751, which means that students in the UKT 1 to UKT 5 categories are 2.751 times more likely to complete their studies than students in the full scholarship category. The value $Exp\ (\beta)$of the variable, X_9namely interest in the major, is 14.992, which means that if the student in question likes the major they are taking, it will increase the study period of more than 4 years by 14.992.

The following is a table of the results of classifying data on the length of study period of students in the Mathematics Study Program using the *APER test*.

Table 8. Results of Classification Accuracy of Study Period Length

Observation	Prediction		Amount
	$\leq 4\ Tahun$	$> 4\ Tahun$	
$\leq 4\ Tahun$	15	6	21
$> 4\ Tahun$	5	34	39
Amount	20	40	60

Source: Processed data, 2023

Based on Table 8, the level of classification error can be calculated using the *APER test* , namely:

$$APER = \frac{5 + 6}{15 + 6 + 5 + 34} = 0{,}183$$

The classification accuracy of the data is obtained as follows:

$$(1 - 0.183) = 0.9817 = 98{,}17\%$$

So the overall classification accuracy percentage is 98.17% with a classification error of 1.83%. It can be concluded that the classification accuracy of students who graduated more than 4 years is very high.

4. CONCLUSION AND SUGGESTIONS

Research Conclusion

Based on the results of the analysis, the following conclusions can be drawn:

1. The best model resulting from modeling using logistic regression is:

   $g(x) = -31{,}507 + 2{,}847X_1 + \qquad 1{,}012X_2 + 2{,}707X_9$

   with a classification accuracy of 98.17%.
2. The variables that significantly influence the study period of students in the Mathematics Study Program at Udayana University from 2011 to 2019 are gender, UKT and interest in the major.

Suggestion

In further research, researchers can use variables other than the variables in this study. In addition, researchers can also use alumni data with different batches or even expanded samples.

BIBLIOGRAPHY

Agresti, A. (2018). An introduction to categorical data analysis (3rd ed.). Hoboken, NJ: Wiley.

BMJ Family Medicine and Community Health. (2023). Primer on binary logistic regression. Retrieved from BMJ Family Medicine and Community Health.

Gude, V., Anderson, C., & Roetzel, P. (2017). Factors influencing time to degree completion in doctoral programs: A systematic review. Journal of Higher Education Research, 88(3), 234–256. https://doi.org/10.1080/00221546.2017.1288092

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied logistic regression (3rd ed.). Hoboken, NJ: Wiley.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skills factors predict college outcomes? A meta-analysis. Educational Psychologist, 57(1), 1–18. https://doi.org/10.1080/00461520.2021.2014931

Sverdlik, A., Hall, N. C., McAlpine, L., & Hubbard, K. (2018). The PhD experience: A review of factors influencing doctoral students' completion, achievement, and well-being. International Journal of Doctoral Studies, 13, 361–388. https://doi.org/10.28945/4113

Sverdlik, A., Oreg, S., Berson, Y., & Dvir, T. (2018). Transformational leadership and followers' personal values: The mediating role of leader and follower regulatory focus. Journal of Organizational Behavior, 39(5), 639–653. https://doi.org/10.1002/job.2228

Varamita, A. (2017). Analisis regresi logistik dan aplikasinya pada penyakit anemia untuk ibu hamil di RSKD Ibu dan Anak Siti Fatimah Makassar (Unpublished undergraduate thesis). Universitas Negeri Makassar.

Wulandari, D. Y., Ismaraidha, & Amelia, O. (2022). Analisis model regresi logistik biner

dalam menentukan faktor-faktor yang mempengaruhi masa studi mahasiswa. Journal of Mathematics Education and Science, 8(1), 128–134.

Yelfera, S. R., Yozza, H., & Rahmi, I. (2022). Faktor-faktor yang mempengaruhi masa studi dan indeks prestasi kumulatif mahasiswa dengan model regresi logistik biner bivariat. Jurnal Matematika UNAND, 11(1), 82–94.