Journal of Education and Practice
ISSN 2222-1735 (Paper) ISSN 2222-288X (Online)
Vol.7, No.25, 2016

Binary Logistic Regression Analysis in Assessment and Identifying Factors That Influence Students' Academic Achievement: The Case of College of Natural and Computational Science, Wolaita Sodo University, Ethiopia

Bereket Tessema Zewude (MSc) 1* Kidus Meskele Ashine (Ass. Professor)2
1. Wolaita Sodo University, College of Natural and Computational Sciences, Wolaita Sodo, Ethiopia, P.O.Box 138
2. Wolaita Sodo University, School of Law, Wolaita Sodo, Ethiopia. P.O. Box 138

Abstract
An attempt has been made to assess and identify the major variables that influence student academic achievement at college of natural and computational science of Wolaita Sodo University in Ethiopia. Study time, peer influence, securing first choice of department, arranging study time outside class, amount of money received from family, good life later on and father's education level are major variables which influence the academic achievement of students at college of natural and computational science of Wolaita Sodo University, Ethiopia using binary logistic regression model.
Keywords: academic achievement, binary logistic regression, good life later on, peer influence, securing first choice of department, Wolaita Sodo University

Introduction

Student academic achievement measurement has received considerable attention in previous research, it is challenging aspects of academic literature, and science student achievement are affected due to social, psychological, economic, environmental and personal variables. These variables strongly influence on the student academic achievement but these variables vary from person to person and country to country. Indeed, student academic achievement can be influenced by some many variables these variables may be termed as student variables, family variables, school variables and peer variables(Crosnoe, Johnson & Elder, 2004).

It is assumed that the number of variables may significantly affect the student academic achievement in university. The variables might be the type and location of secondary school attended, type of admission, quality of teaching, life in university, study habit, economic and educational background of parents, references and textbook availability in a university, students placement by their first choice, peer influence, study time etc. For study purpose, we take Grade Point Average (GPA) of students to measure academicachievement. This idea supported by (Hijaz & Naqvi, 2006) stated that GPA in university is commonly used indicator of student academic achievement. Therefore, GPA can be influenced by above stated variables. The main objective of the study was to assess and identify the major variables which influence student academic achievement using binary logistic regression model.

Methodology

Description of study area and period

The study was carried out at college of natural and computational science in Wolaita Sodo University in the academic year of 2012. Wolaita Sodo University is one of the higher institutes of education in Ethiopia. It was established in 2007by the government of Ethiopia.It is found in temperate region of South Nation Nationalities and Peoples (SNNP) regional state in Wolaita zone capital town of Sodo. Sodo town is located (54^0N latitude and 380^0 S longitude) and 396km south of Addis Ababa and 130km from regional town Hawassa. Now the University is operating 3 campuses, 9 colleges and schools and more than 40 departments or programs.

Study Design

The research design was qualitative as well as quantitative research design can be employed.

Source of population

All college of Natural and Computational Science of Wolaita Sodo University students admitted in the academic year of 2012 were considered as population.

Sample Size Determination

Yamane (1967) provides a simplified formula to calculate sample sizes. This formula was used to calculate the sample size. Since it is simply to calculate the sample size. For our case, we uselevel of precision of 5%. Therefore, it is given by:

…….(1)

where:
N is total population which is 1,497
n is sample size to be determined
e is precision error with 5%
Based on the above formula, *n* can be calculated as follow:
Sampling procedure was done using simple random sampling technique to select the departments from nine departments we select five departments randomly.In order to select the students from the selected departments, stratification on the base of academic years was done on basis of proportional to size allocation method. It is given by:

$\frac{N_h}{N} = \frac{n_h}{n} = w_h\ (PPS) \frac{N_h}{N} = \frac{n_h}{n} = w_h\ (PPS)$ …… proportional to size allocation…….(2)

where: is population size in stratum h
is sample size in stratum h
Based on equation (2) the proportional to size allocation of selected department students to be sampled was shown in table 1.

Table 1. Shows Colleges, selected department and number of sampled students

No College	Department	Population Size	Sample Size
College of Natural & Comp.Sci	Biology	533	113
	Chemistry	348	73
	Statistics	138	29
	Environ'tal Sci.	200	42
	IT & Comp Sci.	278	59
		= 1,497	= 316

Source:*Wolaita Sodo University Registrar 2012.*

Variables Identification
The dependent variable of this study is "academic achievement" which has two binary outcomes if a student is not ok status () coded as 0 and ifa student ok status () coded as 1.

The predictor variables consider: age of student, parents' educational background, securing first choice of department, availability of textbooks and references, environmental factor, study habit, place of residence before joining university, peer influence, study time outside class, amount of money received from family, arranging study time and good life later on.

Data Collection Methods
Both primary as well as secondary source of data were used to collect data. Well prepared questionnaire and check list were designed to collect data by distributing to students.

Data Entry and Analysis
Data entry and cleaning were carried out using statistical software package for social science SPSS version 22.0 for the analysis. Descriptive statistics analysis was used to show the frequency distribution by using tables. Binary logistic regression model was used in order to assess and identify the influence of variables on student academic achievement.

Results and discussion
From table 2the age of students ranging from 18-23 years was about 270(90.3%). Regarding their sex, 177(56.1%) of them were males and only 126(39.8%) of them were females during the study period. Regarding place of high school were student attended account 238(79.1%) was urban and 61(20.3%) was rural, respectively.

On the same fashion, student mother's education level which assumed to influence student academic achievement account for illiterate 104(34.6%), for primary 116(38.5%), for secondary 39(13.0%) and followed certificate and above share 40(13.3%), respectively.

On the same manner, student father's education level for illiterate 63(21.0%), for primary 117(39.0%), for secondary 41(13.7%) and certificate and above share 76(25.3%), respectively. Peer influence of student in university stay on strongly agreed position account 64(21.5%, for agree 117(39.4%), for neutral 57(19.2%), for disagree 32(10.8%) and for strongly disagree account 19(6.4%), respectively.

Regarding student receive money from their family for the last four months which is assumed to influence student academic achievement account 189(64.5%)for less than 1500 birr position and 104(35.5%) for greater than 1500 birr, respectively. Study outside class for less than 48huors account 204(67.3%) and 99(32.7%), respectively.

Regarding good life later on account 64(55.1%) for strongly agree, 100(33.1%) for agree, 17(5.6%) for neutral, 10(3.3%) for disagree and 7(2.3%) for strongly disagree, respectively.

Table 2. Results of Descriptive statistics

Factors		Status			
		Not ok	Ok	Total	
Sex	Male	49(16.2%)	128(72.3%)	177(56.1%)	
	Female	34(11.5%)	92(30.5%)	126(39.8%)	
Place of High school	Urban	60(19.9%)	178(59.1%)	238(79.1%)	
	Rural	21(7.0%)	40(13.3%)	61(20.3%)	
Age	18-23 years	83(27.8%)	187(62.5%)	270(90.3%)	
24 years		0(0%)	29(9.7%)	29(9.7%)	
Father's edu. Level	Illiterate	20(6.7%)	43(14.3%)	63(21.0%)	
	Primary	31(10.3%)	86(28.7%)	117(39.0%)	
	Secondary	14(2.7%)	27 (10.3%)	41(13.7%)	
	Certificate & above	18(3.0%)	58(10.3%)	76(25.3%)	
Mother's edu. Level	Illiterate	25(8.3%)	79(26.2%)	104(34.6%)	
	Primary	41(13.6%)	75(24.9%)	116(38.5%)	
	Secondary	9(3%)		39(13.0%)	
	Certificate& above	30(10.0%)		40(13.3%)	
		9(3.0%)			
		31(10.3%)			
Good life	Strongly Agree	46(15.2%)	121(39.9%)	167(55.1%)	
	Agree	25(8.3%)	75(24.8%)	100(33.1%)	
	Neutral	4(1.3%)	13(4.3%)	17(5.6%)	
	Disagree	5(1.7%)	5(1.7%)	10(3.3%)	
	Strongly disagree	3(1.0%)	4(1.3%)	7(2.3%)	
Peer influence	Strongly Agree	19(6.4%)	45(15.2%)	64(21.5%)	
	Agree	29(6.4%)	88(29.6%)	117(39.4%)	
	Neutral	20(6.7%)	37(12.5%)	57(19.2%)	
	Disagree	6(2.0%)	26(8.8%)	32(10.8%)	
	Strongly disagree	7(2.4%)	12(4.0%)	19(6.4%)	
Study outside class	< 48hrs	81(26.7%)	123(40.6%)	204(67.3%)	
48hrs		2(0.7%)	97(32.0%)	99(32.7%)	
Money received	<1500birr	83(28.3%)	106(36.2%)	189(64.5%)	
1500birr		0(0%)	104(35.5%)	104(35.5%)	
Frustration	Yes	18(6.0%)	53(17.5%)	71(23.5%)	
	No	61(20.2%)	163(74.2%)	224(74.2%)	
Counseling & guidance	Yes	44(14.5%)	112(37.0%)	156(51.5%)	
	No	39(12.9%)	106(35.0%)	145(47.9%)	
1st choice dept	Yes	56(18.5%)	152(50.3%)	208(68.9%)	
	No	26(8.6%)	63(20.9%)	89(29.5%)	

In this section it is discussed as the model summary, in order to identify variables associated with student academic achievement binary logistic regression model was used. Moreover, the joint impact of all predictor variables on the dependent variables also determine by using the concept of Nagelkerke R2which is explained in the model summary (table 3).

Table 3. Model summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	111.00[a]	0.51	0.45

The most common assessment of overall model fit in logistic regression is the likelihood ratio test, which is simply the chi-square difference between the null model (i.e., with the constant only) and the model containing the predictors. Under Model Summary we see that the -2 Log Likelihood statistics is 111.00. This statistic measures how poorly the model predicts the student academic achievement in ok status, the smaller the statistic the better the model. The Cox and Snell or Nagelkerke R^2 is an analogous statistic in logistic regression to the coefficient of determination R^2 in linear regression, but not close analogy. The model summary provides some approximation of R^2 statistic in logistic regression. Cox and Snell's R^2 attempts to imitate multiple R^2 based on likelihood. The result of Cox and Snell R^2 indicates that 51% of the variation in the dependent variable is explained by the predictorvariable which is assumed to be good enough(table 3). It is well known that,

however, the big problem with Cox-Snell R^2 is that it has an upper bound that is less than 1.0. Specifically, the upper bound depends only on , the marginal proportion of cases of events given by: ……..(3)

To calculate the upper bound our p value is equal to 0.64. By replacing this value in equation (3) we get the upper bound as:

Table 4. Goodness of fit (Model Diagnostic)

HosmerLemeshow Teat			
Step	Chi-square	df	Sig.
1	45.000	7	0.980

As it is observed from the table above since P-value is 0.980 is greater than the level of significance at 5%. We can conclude that the data fits the model well. Since the p-value is 0.980 which is insignificant therefore our fitted logistic regression model is good fit (Table 4).

Table 5. Results of binary logistic regression model

Table 5. Binary Logistic regression model									
								95.0% C.I.forEXP(B)	
		B	S.E.	Wald	Df	Sig.	Exp(B)	Lower	Upper
Step 1[a]	Study time (> 48 hours= ref.)	3.968	1.215	10.671	1	.001*	***5.853***	4.890	571.318
	Amount of money received from family (>1500 birr ref.)	31.247	4.126E3	26.79	1	.009*	***3.720***	.678	3.880
	Peer influence (Yes=ref.)	-2.356	1.157	4.146	1	.042*	***.095***	.010	.916
	No	.765	1.347	.323	1	.570	2.149	.153	30.089
	Securing 1st choice of department (Yes=ref.)	2.44	1.888	1.290	1	.034*	***11.473***	.010	.916
	No	.965	1.447	.667	1	.570	2.149	.153	30.089
	Father Education			5.734	4	.022*	***2.512***		
	Illiterate (ref.)	1.838	4.019E4	.000	1	1.000	6.281	.000	.
	Primary	-.626	1.092	.328	1	.0467*	2.535	.063	4.547
	Secondary	.790	.893	.783	1	.376	2.204	.383	12.695
	Certificate and above	-.485	1.053	.212	1	.645	.616	.078	4.853
	Good life later on			6.938	4	.001*			
	Strongly agree	3.226	1.785	3.266	1	.071	25.182	.761	833.018
	Agree	4.063	1.839	4.882	1	.027*	58.135	1.582	2.136E3
	Neutral	5.358	2.332	5.278	1	.022*	212.330	2.196	2.053E4
	Disagree	2.976	2.111	1.986	1	.159	19.599	.313	1.229E3
	Strongly disagree			6.209	5	.286			
	Arranging study time outside class (strongly agree=ref.)	0..997	3.343	4.381	1	.036*	***2.001***	.000	.641
	Agree	-5.246	3.179	2.724	1	.099	.005	.000	2.675
	Neutral	-6.835	3.553	3.701	1	.044*	.001	.000	1.137
	Disagree	-3.361	4.119	.666	1	.414	.035	.000	111.239
	Constant	15.708	5.699E4	.000	1	1.000	6.639E6		

From the table 5, it is observed that the estimated odds ratio 5.85 indicates those students who study more than 48 hours are 5.85 times more likely to perform better in academic achievement compared to those student study less than 48 hours controlling for other variables in the model. On the same fashion, the estimated odds ratio 11.47 indicates that student securing first choice of department are 11.47 more likely to perform better in academic achievement compared to those student who have not secured first choice of department controlling for other variables in the model. Similarly, the estimated odds ratio 0.095 indicates that student who are not influenced by peer are 9.5% more likely to perform better as compared to those student influenced by peer controlling for other variables in the model. Based on the above table, the estimated odds ratio 3.72 indicates that those students who receive more than 1500 birr money from their family are 3.72 times more likely to perform better compared to those who receive less than 1500 birr money from their family controlling for other variables in the model. Arranging study outside class also one of the determine factor for academic achievement from the

above result revealed that the estimated odds ratio 2.00 indicates that those student who arrange study outside class are 2.00 times more likely to perform better compared to their counterparts controlling for other variables in the model. Regarding father's education level the estimated odds ratio 2.51 indicates that those student whose father's level of education certificate and above level are 2.51 more likely to perform better compared to those counterparts controlling for other variables in the model.

Conclusions

From the logistic regression analysis it was also concluded that the odds of securing first choice of department, peer influence, father's education level, study time, arranging study outside class and amount of money received from family are significant predictor variables seems to indicate better academic achievement of students in ok status as compared to their counterparts in college of natural and computational science of Wolaita Sodo University situation.

Recommendations

- A lot should be done towards developing the academic achievement of students by counseling and guiding about peer influence at university level.
- It can be also recommended that the university should set programs to strength self-concept or motivation to make them confident on their potential.
- The stalk holders should secure student first choice of department.
- Further studywith additional predictor variables have to be made so as to address the issues raised in this study.

References

AdemKedir (2005). Factors Affecting Students' Academic Performance in Higher Institutions. Journal of Ethiopian Statistical Association 14, 73-82.

Agus and Makhbul, Z.K. (2002): An empirical study on academic achievement business students in pursuing higher education: An emphasis on the influence of family backgrounds, paper presented at International conference on the challenges of learning and teaching in a brave new world: Issues and opportunities in bordless education, Hotyai, Thailand.

Battle, J., & Lewis, M. (2002). The increasing significance of class: The relative effects of race and socioeconomic status on academic achievement. Journal of Poverty, 6(2), 21-35.

Checchi, D. (2008): University education in Italy. International journal of manpower 21(4), 177-205.

Crosnoe, R., Johnson, M. K., & Elder, G. H. (2004). School size and the interpersonal side of education: An examination of race/ethnicity and organizational context. Social Science Quarterly, 85(5), 1259-1274.

Hanushek, E.A (2006): assessing the effect of school resources on students' performance. Educational evaluation and policy analysis 19(2), 141-164.

Rothstein, R. (2007): Finance Fungibility: Investing Relative Impacts of investments in schools and non-schools educational institution to improve student achievement. Center of educational policy publications, Washington DC.

Syed Tahir Hijaz and S.M.M. Raza Naqvi (2006): Factors affecting students' performance: A case of private colleges in Bangladesh. Journal of sociology 3(1), 44-45.