Journal of Education and Practice
ISSN 2222-1735 (Paper) ISSN 2222-288X (Online)
Vol.6, No.36, 2015

Modeling Student Performance in Mathematics Using Binary Logistic Regression at Selected Secondary Schools A Case Study of Mtwara Municipality and Ilemela District

[1,2*]Salyungu Mabula
1. Department of Economics and Mathematics, Stella Maris Mtwara University College (*A Constituent College of Saint Augustine University of Tanzania*), P.O.Box 674, Mtwara, Tanzania
2. Student of Masters of Science in Official Statistics, Eastern Africa Statistical Training Center, P.O.Box 35103, Dar es alaam - Tanzania

Abstract
This study investigated the performance of secondary school students in Mathematics at the Selected Secondary Schools in Mtwara Municipality and Ilemela District by Absenteeism, Conduct, Type of School and Gender as explanatory Factors. The data used in the study was collected from documented records of 250 form three students with 1:1 gender ratio – 50 students from each of the five selected secondary schools in the academic year 2011/2012. The sample was considered appropriate as they had covered more than half of the Mathematics syllabus in Ordinary Secondary Schools. Binary logistic regression was used to model a binary variable 'performance' (fail, pass) against a systematic component of linear combination predictors (absenteeism, conduct, type of school and gender). The model fitted for the log-odds in favour of poor performance is $\log_e\left(\frac{\pi(x)}{1-\pi(x)}\right) = -1.185 + 0.346\, x_1 + 1.137\, x_2$. The essence of this study is to provide student performance analysis method (Binary Logistic Regression) not commonly used in Tanzania. Findings show that two out of four explanatory factors used in the study (absenteeism and misconduct) significantly predict student performance in Mathematics based on binary logistic regression fitted. Absenteeism and misconduct predict the log-odds of poor performance by multiplicative effect of 1.414 and 3.137 respectively. Future work is recommended to focus on analysis using other Generalized Linear Models (GLM) as well considering other locations with more/other variables affecting performance of students in mathematics.
Keywords: Binary Logistic Modeling, Misconduct, Performance.

1. Introduction

Mathematics is recognized for playing a major role in development of science and technology making it an integral part of the world culture as evident in everyday life. For any nation to be relevant, it must not over-look the importance of Mathematics in its educational system.

The overall performance of students in Certificate of Secondary Education Examination (CSEE) has been steadily deteriorating in the last fifteen years. This deterioration specifically includes poor performance in Mathematics such that the subject has ranked the poorest performed followed by English subject. International Mathematical Union (IMU, 2009) observed that, poor performance in Mathematics has been a matter of serious concern to all well-meaning educators. Students' poor performance in Mathematics over the years has been attributed to the fact that the subject is considered difficult. Figure 1 presents percentage pass rates at CSEE from the year 2001 through 2012. The period had roughly two phases, one of slightly constant and higher pass rates (2002 – 2007) with a maximum pass rate of 91.5% in 2004 and the other with a more dramatic decrease in pass rates (2007 – 2012) reaching a minimum pass rate of 34.5% in 2012.

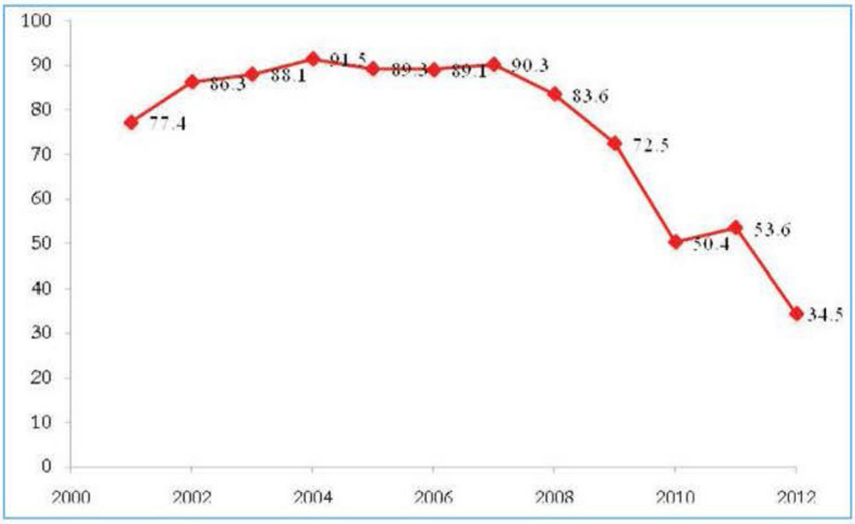

Figure 1: Percentage of candidates pass rates at CSEE 2000-2012 *(Source: Hakielimu, 2013)*

The questions to what is behind deterioration of performance and by what magnitude do the factors contribute is a major concern of stakeholders and decision makers in academia. Among others reasons, student absenteeism and misconduct contribute to poor performance (Miller 2009, Kaur 2005, Smith *et al* 2006 and Brock-Utne 2006). Hakielimu (2013) also show that the type of school (public, private) significantly contribute to student's performance. Other factors like intelligence, study habits, and attitudes of student towards school, different aspects of their personality, socio economic status also affect student performance, Adeyami *et al* (2013). The continuing decreasing trends in performance especially in Mathematics is a major threat to the education sector and development of the nation at large - calling upon stakeholders to utilize the expertise in collecting, analyzing and interpretation of results regarding poor performance henceforth recommending appropriate policies.

Authors use different models in simplifying the complexities of relationship among variables. Models range from simple mathematical relationships and conceptual frameworks with few explanatory variables predicting a response variable to more complicated and multi-variable models. In modeling student's performance, different mathematical models are used including Ordinary Least Square (OLS), Standard Multiple Regression (SMR) and Generalised Linear Models (GLM). The latter is more advanced than the others capable of handling more complicated situations and analyze the simultaneous effects of multiple variables, including mixtures of categorical and numerical variables (Agresti 2007, Agresti 2002, McCullagh & Nelder 1989). The term *generalized linear models* (GLM) usually refers to large class of conventional linear regression models for a continuous response variable given continuous and/or categorical predictors following McCullagh & Nellder (1989). In these models, the response variable is assumed to follow an exponential family distribution. Other authors would call GLM as "nonlinear" because the mean of the response variable is often a nonlinear function of the covariates, but *ibid* (1989) consider them to be linear, because the covariates affect the distribution of the response variable only through the linear combination explanatory variables in a transformed form.

The relationship between student performance in Mathematics as a continuous response variable coded as a categorical variable (fail, pass) and a combination of linear predictors like absenteeism, misconduct, school type and gender can be modeled using Generalized Linear Models following Adeyami *et al* (2013) and Adejumo & Adetunji (2013). Despite the merits of GLM over OLS and SMR in dealing with complex relationship of variables, the method is rarely used in Tanzania specifically when studying students' performance in secondary schools. It is, therefore, the purpose of this study to model students' performance in Mathematics at selected secondary schools in Mtwara Municipality and Ilemela District using performance as a continuous binary response variable (*fail, pass*) and a linear combination of predictors *viz:* absenteeism, misconduct, school type and gender as a mixture of continuous and categorical variables.

Adeyami *et al* (2013) conducted a survey study to determine the socio-economic factors influencing student's academic performance in Yaba College of Technology, Yaba, Lagos, Nigeria. The authors collected data from six hundred (600) students across the different departments and class study. Data collected was analyzed using

statistical package for social science by descriptive statistics, Pearson correlation test of existence of linear relationship between variable of study, chi square test for association of factors, and a four predictor binary logistic regression was fitted to the data. The students' academic performance was measured using variable GPA/CGPA categorized into two - poor (GPA/CGPA between 0 and 2.49) and - good (GPA/CGPA between 2.50 and 4.00). Four factors; mothers' education level (x_1) living togetherness of parents (x_2) student class (x_3) and weekly income/allowance (x_4) are found to influence students' academic performance after which the four factors were fitted into predictive binary logistic regression model for the log-odds in favour of poor performance as $log_e\left(\frac{\pi}{1-\pi}\right) = 0.122 - 0.092x_1 + 0.479x_2 - 0.383x_3 - 0.411x_4$. A number of recommendations like rendering of financial support to students in need of such, family planning orientation while in school, and teaching of effect demographic and socio-economic factors on student academic performance were emphasized**.**
The objective of the current study is to determine a binary logistic model explaining the relationship between students' performance in mathematics using predictors namely absenteeism, misconduct, school type and gender at selected secondary schools in Mtwara Municipality and Ilemela District.

2. Materials and Methods

2.1 The study area

This study was conducted in two locations, namely Mtwara Municipality located at $10^0 S, 40^0 E$ in Mtwara region - southern part of Tanzania and Ilemela district located at $2^0 S, 32^0 E$ in Mwanza region. In Mtwara Municipality three secondary schools, namely Umoja, Mtwara Technical and Shangani were selected as sampling sites. Two secondary schools namely, Loreto girls and Lukobe were selected as sampling sites from Ilemela. Among the five schools, only Loreto girls' is a private school and the rest are public.

2.2 Sampling techniques

The study generally targeted secondary school students in the districts of Mtwara Municipality and Ilemela District. Since the population was spread all over the country, thus we adopted an international standard of sampling with the rule of "exclusion and inclusion" following Gonzales *et al* (2008). According to this rule, population of the selected schools will exclude all other schools; population of the selected classes in each selected school will exclude all other classes. Using this rule it was possible to restrict the population of the present study to only form three students in the selected schools. This procedure of sampling is generally referred by Trends in International Mathematics and Science Study-TIMSS - (2007), to as a three-stage stratified cluster sampling technique. The sample size of the present study was 250 students - 50 students from each of the five selected schools. The sex ratio of the participants was 1:1 – that is, 125 boys and 125 girls of which 150 participants were from Mtwara Municipality and 100 from Ilemela District. Since Loreto is a single-sex school, all participants were girls. Furthermore, to maintain the sex ratio of 1:1, the researcher sampled boys only at Shangani secondary school. Finally, each participant in the study from each form three class of the selected schools was selected using simple random sampling by assigning serial numbers to all students in the class based on gender and then using table of random numbers to select participants based on the serial numbers. The purpose was to get a bias-free sample.
The main tool of data collection in the current study was a checklist (table 7) which was developed and filled using data documented by schools in attendance registers and students progress reports. Documentaries have an advantage of dealing with massive data of samples, Rootman *et al* (2010). The student's average of terminal and annual examinations grades in Mathematics was coded as performance with less than 30% being fail = 1 and greater or equal to 30% a pass = 0, as prescribed by Ministry of Education and Vocational Training, Mbelle & Katabaro (2003). The average grade of character assessment of a student coded misconduct = 1 for grade less than 30% and good conduct = 0 for grade greater than or equal to 30%. On the other hand, female was coded 1 and male 0 while with school type, public school was coded 1 and private school 0. Absenteeism was treated as a continuous explanatory variable presented as a percentage of number of days a student was absent on school open days in the academic year 2011/2012.

Table 7: Data Summary

STUDENT S/NO	LORETO SS				LUKOBE SS				UMOJA SS				MTWARA TECH. SS				SHANGANI SS			
	AB	SC	CD	GN	AB	SC	CD	GN	AB	SC	CD	GN	AB	SC	CD	GN	AB	SC	CD	GN
1	0	41	80	F	10	7	40	M	8	9	60	M	1	79	80	M	9	15	60	M
2	5	38	80	F	13	15	40	M	12	6	60	M	7	28	40	M	6	8	80	M
3	0	36	80	F	7	18	40	M	11	5	60	M	16	22	40	M	6	10	40	M
4	6	14	60	F	5	19	40	M	16	0	60	M	4	89	80	M	4	23	40	M
5	1	20	60	F	1	61	80	M	6	8	60	M	9	26	40	M	5	4	40	M
6	0	47	80	F	2	49	40	M	20	18	40	M	5	25	40	M	14	6	40	M
7	1	24	80	F	9	13	60	M	11	3	40	M	4	79	100	M	5	28	60	M
8	8	12	60	F	8	23	60	M	7	4	60	M	10	41	60	M	4	21	60	M
9	0	80	80	F	9	16	60	M	21	6	60	M	6	61	80	M	5	13	80	M
10	0	45	80	F	5	32	60	M	11	12	60	M	1	64	80	M	5	4	60	M
11	1	25	80	F	5	36	40	M	33	3	60	M	1	42	80	M	7	29	60	M
12	0	24	60	F	7	25	60	M	29	0	40	M	6	23	60	M	7	26	40	M
13	3	25	60	F	6	35	40	M	35	0	40	M	5	11	40	M	4	8	60	M
14	0	10	60	F	18	8	40	M	11	7	40	M	2	55	80	M	2	55	60	M
15	0	35	80	F	11	15	60	M	27	6	60	M	6	36	80	M	6	19	80	M
16	3	37	80	F	4	38	60	M	19	0	60	M	5	32	40	M	6	25	80	M
17	6	61	80	F	5	30	60	M	17	3	40	M	3	17	40	M	13	2	60	M
18	0	38	80	F	4	35	60	M	10	9	40	M	5	10	40	M	11	18	60	M
19	0	40	80	F	8	19	60	M	37	0	60	M	23	32	40	M	7	4	40	M
20	0	19	60	F	9	22	80	M	9	5	40	M	6	70	80	M	4	8	40	M
21	0	21	60	F	2	62	60	M	34	20	60	M	2	4	40	M	4	6	60	M
22	0	57	80	F	4	52	60	M	20	3	60	M	1	68	80	M	8	1	60	M
23	0	64	80	F	11	17	60	M	12	3	40	M	2	51	60	M	7	2	60	M
24	0	40	80	F	10	16	40	M	1	7	60	M	2	37	80	M	6	9	60	M
25	0	48	60	F	10	14	60	M	22	0	60	M	12	10	60	M	5	37	60	M
26	0	56	80	F	6	32	60	F	7	19	60	F	7	34	80	F	5	35	60	F
27	0	57	60	F	7	20	40	F	6	10	60	F	6	64	80	F	2	59	60	F
28	9	22	60	F	3	61	80	F	5	14	60	F	8	12	80	F	4	20	60	F
29	1	21	60	F	6	19	40	F	8	11	60	F	3	35	40	F	6	14	60	F
30	0	52	40	F	11	16	40	F	5	10	80	F	3	93	100	F	7	10	60	F
31	0	7	40	F	0	41	60	F	10	12	60	F	5	22	40	F	7	0	60	F
32	0	30	60	F	10	15	40	F	4	14	60	F	10	28	40	F	6	19	60	F
33	15	7	60	F	6	29	60	F	6	7	60	F	6	12	80	F	11	5	60	F
34	0	62	80	F	2	58	60	F	6	17	60	F	3	62	60	F	2	54	40	F
35	0	74	80	F	5	47	80	F	22	0	40	F	2	73	80	F	4	22	40	F
36	0	73	60	F	11	19	40	F	22	0	40	F	4	15	80	F	3	21	80	F
37	1	35	80	F	6	31	60	F	13	8	40	F	7	16	60	F	8	8	80	F
38	1	48	80	F	5	31	60	F	26	3	60	F	5	2	60	F	5	8	40	F
39	6	10	60	F	5	29	40	F	5	3	40	F	6	8	60	F	7	48	60	F
40	9	5	60	F	6	27	60	F	11	3	40	F	4	56	80	F	4	41	40	F
41	5	31	80	F	7	22	40	F	9	3	60	F	4	4	40	F	4	50	60	F
42	0	14	60	F	12	21	40	F	5	5	60	F	6	5	60	F	5	22	60	F
43	0	25	60	F	18	16	40	F	5	4	60	F	3	62	80	F	6	23	40	F
44	2	38	80	F	19	15	40	F	5	3	60	F	9	38	60	F	10	27	40	F
45	0	94	80	F	6	30	60	F	7	6	60	F	11	50	60	F	4	8	60	F
46	0	30	60	F	7	19	60	F	2	5	60	F	6	15	80	F	3	20	60	F
47	5	7	60	F	6	32	60	F	13	3	60	F	2	33	80	F	18	0	40	F
48	2	11	60	F	7	22	40	F	18	0	60	F	4	48	80	F	8	8	60	F
49	1	46	60	F	18	10	40	F	20	3	60	F	4	24	40	F	4	22	60	F
50	1	24	40	F	5	30	60	F	3	4	60	F	8	49	80	F	9	6	80	F

Key: AB=Absenteeism (%), CD=Student Conduct (%) , SC = Grade (%), GN = Gender, F=Female, M=Male

** (Private school) * (Public School)

The relationship between performance and the predicting factors was modeled using binary logistic model of the form;

$$P = \text{logit}(\pi(x)) = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k \text{, for k >2} \quad (1)$$

From which we get

$$\pi(x) = \frac{\exp(\alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots \beta_k x_k)}{1 + \exp(\alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k)} \quad (2)$$

Where P = scores
$\pi(x) =$ Pi = the probability of the student having poor performance in mathematics, $\alpha =$ constant, $x_1, x_2 ... x_k = explanatory\ variables$ and $\beta_1, \beta_2, ..., \beta_k = coefficients$

3. Results and Discussion

3.1 Classification Table

The benchmark that was used to characterize a logistic regression model as useful is a 25% improvement over the rate of accuracy achievable by chance alone following (*SW388R7 Data Analysis & Computers II*). Even if the independent variables had no relationship to the groups defined by the dependent variable, we would still expect to be correct in our predictions. This is referred to as by chance accuracy. Table 1 shows the number of cases in each group at Step 0 (before any independent variables are included). The proportion of cases in the largest group is equal to the overall percentage (66%). The proportional by chance accuracy rate was computed by calculating the proportion of cases for each group based on the number of cases in each group in the classification table at Step 0, and then squaring and summing the proportion of cases in each group ($0.34^2 + 0.66^2 = 0.55 = 55\%$). The proportional by chance accuracy criteria is 68.75% (i.e. 1.25 x 55% = 68.75%).

Table 1: Classification Table[a,b]

Observed			Predicted		
			performance		Percentage Correct
			pass	fail	
Step 0	performance	pass	0	85	.0
		fail	0	165	100.0
	Overall Percentage				66.0

a. Constant is included in the model. b. The cut value is .500

On the other hand, the comparison between predicted scores and the actual scores is presented in table 2. It indicates the number of cases in each group at Step 1 (after independent variables are included).

Table 2: Classification Table[a]

Observed		Predicted		
		performance		Percentage Correct
		pass	fail	
performance	pass	48	37	56.5
	fail	19	146	88.5
Overall Percentage				77.6

a. The cut value is .500

We can see from table 2 that 146 students who were predicted to fail by our model (with the four predictors) did indeed fail, while 37 were predicted to fail and in fact passed. Also, 48 students were predicted to pass and indeed passed while 19 were predicted to pass but failed. In total, 77.6% of our predictions were accurate, which though not perfect, it is a clear improvement over the baseline model, where 66% of predictions were accurate.

3.2 Omnibus Tests of Model Coefficients

The presence of a relationship between the response variable and combination of explanatory variables is based on the statistical significance of the model chi-square at step 1 after the predictors have been added to the analysis. Table 3 shows an indication of whether or not the model with our explanatory variables fits the data better (i.e. gives us a better prediction of individual performance) than the baseline model.

Table 3: Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	78.297	4	.000
	Block	78.297	4	.000
	Model	78.297	4	.000

We can find significance in the final column on table 3. In this analysis, the probability of the model chi-square (78.297) was < 0.001, less than or equal to the level of significance of 0.05. The existence of a relationship between absenteeism, misconduct, school type and gender as predictors on one hand and performance as a response on the other hand was supported, which means that our model with the four predictors (at step 1) fits better than a model with no predictors.

3.3 Model summary

The Pseudo R square statistics in table 4 show two measures, Cox & Snell and Nagelkerke. The measures use somewhat different formula, but both are equally valid. In this case Cox & Snell is 0.269, and Nagelkerke is 0.372.

Table 4: Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	242.221[a]	.269	.372

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.

These numbers indicate modest and moderate improvement in fit respectively over the baseline model, (based on 0-0.1 = poor improvement, 0.1-.3 modest improvement, 0.3-0.5 moderate improvement and more than 0.5 strong improvement).

3.4 Hosmer and Lemeshow Test

The Hosmer and Lemeshow Test in table 5 is the preferred test of goodness-of-fit. As with most chi-square based tests however, it is subject to inflation as sample size increases.

Table 5: Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	2.524	7	.925

Here, we see model fit is acceptable χ^2 (8) = 2.524, p = 0.925, which indicates our model predicts values not significantly different from what we observed. The p-value should be greater than the established cutoff value (generally 0.05) in table 5 to indicate goodness-of-fit.

3.5 Variables in the Equation

Numerical problems such as multicollinearity among the independent variables (constant excluded) are checked by examining the value of standard errors for the B-coefficients in table 6. A standard error larger than 2 indicates numerical problems. None of the independent variables in this analysis had a standard error larger than 2 indicating that there is no numerical problem of the model in the current study.

Table 6: Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	absenteeism	.346	.066	27.519	1	.000	1.414	1.242	1.609
	misconduct	1.137	.431	6.956	1	.008	3.118	1.339	7.260
	School type	-0.388	.474	.671	1	.413	.678	.268	1.717
	gender	0.202	.364	.308	1	.579	1.224	.600	2.499
	Constant	-1.185	.497	5.680	1	.017	.306		

a. Variable(s) entered on step 1: absenteeism, misconduct, school type, gender.

The probabilities of the Wald statistic at 95% confidence interval for the variables absenteeism and misconduct are < 0.001 and 0.008 respectively which is less than or equal to the level of significance of 0.05. This suggests the relationship that "students who are frequently absent and misbehaving are more likely to perform poorly in mathematics compared to their counterparts" and concur with finding of Erick *et al* (2013) and Smith *et al* (2006) who contend in their separate studies that absenteeism and misconduct, respectively, are among factors that affect performance of a student. The insignificant relationship between performance and the independent variables school type and gender in table 6 may be attributed by the fact that both public and private schools in Tanzania use the same curriculum. In the current study, performance is a categorical variable that is coded so that 1 is associated with students who had failed mathematics and 0 associated with students who had passed. From table 6, the model relating student performance and absenteeism and misconduct is given by;

$$\text{logit}(\pi(x)) = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = -1.185 + 0.346x_1 + 1.137x_2 \tag{3}$$

From which we get $$\pi(x) = \frac{\exp(-1.185 + 0.346x_1 + 1.137x_2)}{1 + \exp(-1.185 + 0.346x_1 + 1.137x_2)} \tag{4}$$

In equations (3) and (4), $x_1 = absenteeism$ and $x_2 = misconduct$. School type (x_3) and gender (x_4) are excluded in the model because they are found to be statistically insignificant in table 6. The values of β_1 and β_2 in the model are positive which means that a unit increase in absenteeism and misconduct will increase the probability of a fail on the performance variable (fail in this case).

3.6 Odds Ratios

The value of exp(B) in table 6 for misconduct was 3.118 which implies that a unit increase in misconduct rate would increased the odds that a student would fail in mathematics by 3.118 more times. Likewise, the incident rate of absenteeism is 1.414 which implies that a unit increase in absenteeism rate would increase the odds of a student to fail in mathematics by 1.414 more times, holding other variables constant.

4. Conclusion and Recommendations

The current study was conducted to model the factors that influence student performance in mathematics at selected secondary schools in Mtwara Municipality and Ilemela District, Tanzania. Findings show that two out of four explanatory factors used in the study (absenteeism and misconduct) significantly predict student performance in mathematics based on binary logistic regression fitted in (3) and (4). This result is consistent with Erick *et al* (2013) and Smith *et al* (2006) who contend in their separate studies that absenteeism and misconduct, respectively, are among factors that affect performance of a student.

Based on the findings and implications of the present study, the following recommendations are worth:-

- The findings of this study show the crucial roles that parents, educators, government and community at large must play – through cooperation - to help learners achieve academically by inculcating high moral values to children. Parents must ensure availability of positive family environment, teachers should create conducive classroom environments and the government should, not only make good educational policies available, but also implement them accordingly. All the efforts should be towards eliminating student absenteeism and improving students' behaviour. The process can take two approaches, (1) by providing guidance and counseling where students are advised to attend classes and to behave well and (2) by enforcing regulations and rules thereafter implementing them effectively in all education institutions.
- Stakeholders like policy-makers and educators should attach great importance to mathematics education. Addressing the increasing poor performance in mathematics should go hand in hand with demand for increased mathematical skills throughout education systems, and it is therefore essential to monitor how well schools provide students with fundamental skills in this area.
- Impact of student absenteeism and misconduct, economically, is at both family and national level. Parents invest much in educating their children by paying school fees, transport and material support as well as government building infrastructures like classrooms and laboratories. Therefore, all stakeholders should see that students are present in classrooms for lessons and they are not misbehaving to avoid wastage of family and national resources.

The role of other predictors of performance such as home environment, curricular intentions and school and classroom environments have not been included although their impact on performance has been noted, IEA (2004). In order to improve on this, it is suggested that further studies be carried out to ascertain the potential effect of predictors – other than absenteeism and misconduct - on performance and as well using different locations.

5. Acknowledgements

The author is grateful to The Germany Academic Exchange Services for financial support in pursuing a Masters in Official Statistics (MOS) at Eastern Africa Statistical Training Center (EASTC) where he is acquiring, among others, skills in Categorical Data Analysis. The community of Stella Maris Mtwara University College (STEMMUCO) is acknowledged for encouragement and special thanks to Mr. Buteta whom trained the author on computer skills and continually encouraged him in vesting time for academic writings. Last but not least, fellow students at EASTC for observations and suggestions in the course of preparing the manuscript.

References

Adejumo, A.O & Adetunji, A.A (2013). Application of Ordinal Logistic Regression in the Study of Students' Performance. *Journal of Mathematical Theory and Modeling, Volume 3, No. 11.*

Adeyami, D. et al (2013). Binary Logistic Regression of Students Academic Performance in Tertiary Institutions in Nigeria by Socio-Demographic and Economic Factors. *International Journal of Engineering Science and Innovative Technology, Issue 4.*

Agresti, A. (2007). An Introduction to Caategorical Data Analysis. Second Edition. Florida: Wiley-interscience, A John Wiley & Sons,INC, Publication

Agresti, A. (2002). Caategorical Data Analysis. Second Edition. Florida: Wiley-interscience, A John Wiley & Sons,INC, Publication

Brock-Utne, B. (2006). Language of Instruction and Poverty Alleviation: Papers in Education and Development. *Journal of the Faculty of Education*, University of Dar es Salaam.

Erick, M.E *et al*. (2013). Factors Contributing to Students Poor Performance in Mathematics in Public Secondary Schools in Tharaka South District, Kenya. *Journal of Education and Practice*, Vol. 4, No. 7

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D. & Brenwald, S. (2008). Highlights from TIMSS-2007: Mathematics and Science. *Achievement of US Fourth and Eighth-Grade Students in an International Context*. Washington DC: National Center for Educational Statistics.

Hakielimu. (2013). Does School Environment Affect Student Achievement? *An Investigation into the Relationship Between Secondary School Characteristics and Academic Performance in Tanzania*. Dar es Salaam: Hakielimu

IEA. (2004). Results from Ghanaian Junior Secondary 2 Students' Participation in TIMSS-2003 in Mathematics and Science. Accra: Ministry of Education Youth and Sports.

International Mathematical Union. (2009). Mathematics in Africa: *Challenges and Opportunities*. *A Report to the John Templeton Foundation from the Developing Countries Strategies Group.* Trieste.

Kaur, B. (2005). Schools in Singapore with High Performance in Mathematics at the Eighth Grade Level. Nanyang Technological University, Singapore

Mbelle, A. & Katabaro, J. (2003). Environment, Performance and Access to Education in Tanzania. Dar es Salaam: Mkuki na Nyota Publishers Ltd.

McCullagh P. & Nelder J.A. (1989). Generalized Linear Models. Second Edition. London:CRC Press Publishers

Miller, R.E. (2009). Student Performance. Conduct and Behavior Concerns. *International Journal of Teaching and Learning in Higher Education*. Volume 21, Number 2, 248-251

Rootman, C., Tait, M. & Sharp, G. (2010). Customer Retention in British and Canadian Retail Banks: *Lessons for South Africa*. *A Research Paper*. Nelson Mandela Metropolitan University (NMMU), South Africa.

Smith, B.S., Posamentier, A.S. and Stepelman, J. (2006). *Teaching Secondary Mathematics. Teaching and Enriching Units*. (Seventh Ed.). New Jersey: Merrill Prentice Hall

SW388R7. Data Analysis & Computers II. Logistic regression – Basic Relationships, Retrived from http://www.utexas.edu/../LRBR on 29th November 2015 at 08:20am.