



Research Article

BINARY LOGISTIC REGRESSION MODEL WITH APPLICATION IN STUDENTS G.P.A PERFORMANCE

¹Sule Omeiza Bashiru, ²Aminu Ibrahim, ³Alhaji Modu Isa and ¹Ibrahim Ismaila Itopa

¹Department of Mathematical Sciences, Prince Abubakar Audu University, Anyigba, Kogi State

²Department of Statistics, Confluence University of Science and Technology, Osara, Kogi State.

³Department of Mathematics and Computer Science, Borno State University, Maiduguri, Nigeria

*Corresponding author: bash0140@gmail.com, doi.org/10.55639/607mlkji

ARTICLE INFO:

ABSTRACT

Keywords:

GPA

Academic Performance,

Binary Logistic Regression,

First Semester Courses,

The quality of results obtained by Statistics students of Prince Abubakar Audu University, Anyigba in their first semester 100 level examination has been called into question owing to their poor performances in some courses they offered. Considering the pivotal role of students' early academic achievement in determining the overall success of their university education, university educators must gain insight into how various courses influence students' main indicator of academic success which is the grade point average (GPA). This empirical study examined how the various courses offered by 100-level students of statistics in the aforementioned university influence students' academic performance at the end of the first year using the Binary logistic regression model. The study comprises nine eight (98) students. Negrerke R^2 test shows that 84.2% variation in the log odd ratio was explained by the fitted model. Also, 63.7% explained variation from Cox and Snell R test, the Hosmer and Lemeshow test and the classification results further confirmed the model is robust enough for such a study. Five courses namely MAT 101, PHY 101, STA 111, CHM 101 and BIO 101 were identified to have significantly influenced students' early academic achievement. One of the recommendations made is that an extra learning period for each of these courses should be added to the school lecture timetable to enable students to grasp the understanding of basic concepts of these courses

Corresponding author: Sule Omeiza Bashiru Email: bash0140@gmail.com

Department of Mathematical Sciences, Prince Abubakar Audu University, Anyigba, Kogi State

INTRODUCTION

Over the years, there has been an increasing interest to determine the factors that influence students' academic performance. Academic performance is considered the key feature of any educational system (Reno, 2013). Several factors are assumed to affect university students' academic achievement. Some of these could be peer influence, academic facilities at the university, nature of learning and teaching among others. School success also depends on some other non-cognitive factors which include parents' educational status, success motivation, inadequate educational conditions and the quality of the university (Ozguven, 1974).

Some research that has been conducted to unravel factors that influence the academic performance of students in the tertiary institution were: A study conducted by Jimoh *et al* (2021), to examine the factors that influence the academic performance of students in Kogi state university revealed that Jamb Score, number of friends students have on social media, how students relate theoretical knowledge of concepts taught in class to real-world experience, how often the students attend a lecture, discussing ideas from lectures, an additional job outside school were the significant variables that influence student's academic performance in Kogi State university and part of the recommendations made on how to improve student's academic performance based on the significant factors at the end of the study were: Intervention strategies that support the improvement of students' academic achievement should be initiated at the early stage of their university programme, students should be guided on proper usage of a social networking site and serious effort should be made to encourage and motivate students to work hard right from the beginning of their university programme. Oladebinu *et al* (2018) researched to examine factors that affect the academic performance of students in Colleges of Education in Southwest, Nigeria. The result of the analysis of the study

showed that the following variables which are factors relating to students and school, the background of parents and factors relating to teachers have a major influence on the academic performance of students. The study recommended that the provision of adequate facilities for the school and desired attention and appropriate funding should be given to colleges of education while government should also ensure the provision of alternative means of power supply to the school.

In the work of Sule and Saporu (2015), a logistic regression model was used to examine the factors that influence students' academic performance at the University of Maiduguri. The result of the study revealed that students G.P.A, Students attitude related to the course and the ability to apply the course to real-life experience contributed significantly to factors that affect students' academic achievement in the MTH 101 course. Suggestions were made at the end of the study on how to enhance students' performance in the course. Bereket and Kidus (2017), applied binary logistic regression to assess and unravel the important variables responsible for students' academic achievement at the college of natural and computation science of Wolaita Sodo University in Ethiopia. Among the variables investigated, time of the study, peer group influence, ability to secure the first choice of department, proper arrangement of study time outside class, amount of financial support received from family, good life later as well as father's educational level were discovered to be the important variables responsible for students' academic achievement in the school.

Despite the overwhelming literature on factors that influences students' academic performance, students tend to still perform low at the end of their study. Thus, the need to look at other factors, especially the cognitive ones that could enhance students' academic performance becomes very necessary. This empirical study aims to investigate the courses that contribute

significantly to the academic performance of students at their first semester 100 level using the Binary logistic regression model. The researchers considered the first-semester 100-level course because these are the first grades that undergraduate students will be obtaining at the beginning of their academic life. Grades obtained at the beginning of students' academic journey have a great influence on how they will succeed at the end of their academic studies. Good grades can be a source of motivation for the student to do better while bad grades can demoralize the student and thus lead to a weak motivation to do better. The study will be considered grades obtained in courses by 100-level students in the department of Statistics, Prince Abubakar Audu University Anyigba. Statistics is the science of collecting, analyzing and interpreting data. Statistics is indispensable in monitoring and evaluating economic reforms and programmes of government. The various critical issues Nigeria is currently facing such as unemployment, insecurity, HIV/AIDS pandemic, dearth of skilled manpower in technology, Science and medicine and globalization requires sound and well-trained statisticians that will provide insight into the underlying realities (Shangodoyin *et al.*, 2011). The need to receive a quality education in the area of statistics at various tertiary institutions in Nigeria becomes more necessary at this point to guarantee any meaningful social and economic development. A nice transition from secondary school to university goes a long way in increasing the chances of student's realization of their educational goals in life (Lowe and Cook 2003; Renties *et al.*, 2012). University educators must

The model equation is as given below in logit form.

$$\text{Log} \left[\frac{p(x)}{1-p(x)} \right] = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (1)$$

where $X_i, i=1, \dots, k$ are the factors which are being investigated, the vector, $X = (x_1, x_2, \dots, x_k)$ and

strive hard to understand ways of improving this transition to enable students to start their university education on a good note. The performances of students at the end of their first year in the university are good measures of how well they have adjusted to their new environment and also play a major role in determining students' overall performance in their various degree programmes. Undergraduate statistics students are faced with the realities of offering many courses in their first year. Over the years, Statistics students find it difficult to obtain good grades in some of these courses. Much attention has not been paid to how the level of their performance in each of these courses affects their G.P.A at the end of the first semester, 100 level. This study investigated courses that significantly determine students' G.P.A in their first year in the university. The first semester results of ninety-eight (98) year one students of Statistics in Prince Abubakar Audu University, Anyigba were used for the current study.

LOGISTIC REGRESSION MODEL

Binary Logistic regression analysis is an extension of multiple regression techniques to research situations whereby the dependent variable has a binary outcome.

Students with a GPA between 0 and 2.39 were coded as **0**; while students with a GPA between 2.40 and 5.00 were coded as **1**.

The factors being investigated are considered as the predictor variables to establish possible influence on the response variable.

$$\begin{aligned}
 P(x) &= P(Y=1/X=x) \\
 &= \frac{1}{1 + e^{\sum_{i=1}^k \beta_i X_i}}
 \end{aligned} \tag{2}$$

Hence the odds ratio is

$$\Pi(x) = \frac{p(x)}{1 - p(x)} \tag{3}$$

MATERIALS AND METHODS

The data set used in this study is secondary data obtained from the examination office of the Department of Mathematical Sciences, Prince Abubakar Audu University, Anyigba. Kogi State. The data set contains the grades of 100 level ninety-eight (98) students in Statistics in their first semester, 100 Level. The GPA of the student which is the binary dependent variable is coded as mentioned earlier. The grades of the students in their first semester courses covered largely the ten independent variables that were investigated. The explanatory variables are the ten courses offered in the first semester which are: MAT 101 (Elementary Mathematics I), PHY 101 (General Physics I), PHY 107 (General Physics Laboratory I), STA 111 (Probability I), STA 113 (Basic Statistical Methods), CSC 101 (Introduction to Computer Science), GST 101 (Use of English I), GST 107 (Introduction to Library), CHM 101 (General Chemistry) and BIO 101 (General Biology I).

The data were analysed using the statistical package for social sciences (SPSS) version 26. The estimates of the parameters were obtained by the method of maximum likelihood. The fitted logistic regression model was evaluated using the Neglerke R^2 , Hosmer - Lemeshow test, Chi-Square Test and the Classification Percentage Accuracy. The odds ratio was used to examine the significance of the coefficient of the respective explanatory variables (the ten courses).

DISCUSSION OF RESULTS

A binary logistic model was fitted to data on some variables provided by the grades of statistics students in their first Semester 100 level. The Chi-Square test in table 1 shows that the model is significant. The Hosmer and Lemeshow test in table 2 also shows that the model is a good fit. From table 3, the Cox and Snell R^2 , 63. % variation in the log odd ratio was adequately explained by the various independent variables, and the Neglereke R Square, 84.2% variation in the log odd ratio was explained by fitting the independent variables. Also, the model correctly classifies 89.6% of the entire cases which is shown in table 4; this also further shows that the model is a good fit.

The results as presented in table 5 shows those five (5) variables were identified to have significant regression coefficient β . They are MAT 101, PHY 101, STA 111, CHM 101 and BIO 101. Hence, there are five (5) variables that significantly contributed to the model's predictive ability. The values of the respective estimated odd ratios (e^β) are given in Table 6 under column seven. The estimated odd ratio of 3.186 indicates that students that perform well or have high grades in MAT101 courses are 3.186 times more likely to have a higher GPA than students with a lower grade in MAT 101; controlling for other variables in the model. Also, an odd ratio of 2.098 means that students with higher grades in PHY 101 are 2.098 times more likely to have a higher

GPA than students with lower grades in PHY 101; controlling for other variables in the model. Similarly, an estimated odd ratio of 2.186 indicates that a student who obtains better grades in STA 111 is 2.186 times more likely to have a higher GPA than students with lower grades in STA 111 controlling for other variables in the model. Furthermore, an estimated odd ratio of 3.29 indicates that a student who performs well in

CHM 101 course is 3.29 times more likely to perform better or have a higher GPA than students who perform poorly in CHM 101 grade. Finally, an estimated odd ratio of 2.208 indicates that students with higher grades in BIO 101 are 2.208 times more likely to perform better than students with lower grades in BIO 101 controlling for other variables in the model.

The fitted model in logit form is as given below;

$$\log \frac{p(x)}{1-p(x)} = -27.972 + 0.171X_1 + 0.094X_2 + 0.17X_3 + 0.254X_4 + 0.189X_5 \tag{4}$$

where

$$p(x) = \frac{e^{(-27.972 + 0.171X_1 + 0.094X_2 + 0.17X_3 + 0.254X_4 + 0.189X_5)}}{1 + e^{(-27.972 + 0.171X_1 + 0.094X_2 + 0.17X_3 + 0.254X_4 + 0.189X_5)}} \tag{5}$$

and

- X₁ = MAT 101
- X₂ = PHY 101
- X₃ = STA 111
- X₄ = CHM 101
- X₅ = BIO 101

These five variables significantly affect the GPA of the students with low academic performance. This reveals that any intervention strategy aimed at enhancing the performance of the students with a low GPA in the Statistics department, 100 level of Prince Abubakar Audu University could be focused on MAT 101, PHY 101, STA 111, CHM 101 and BIO 101

Table 1: Omnibus Tests of Model Coefficients

| | | Chi-square | Df | Sig. |
|--------|-------|------------|----|------|
| Step 1 | Step | 95.491 | 10 | .000 |
| | Block | 95.491 | 10 | .000 |
| | Model | 95.491 | 10 | .000 |

Table 2: Hosmer and Lemeshow Test

| Step | Chi-square | Df | Sig. |
|------|------------|----|------|
| 1 | 1.507 | 8 | .993 |

Table 3: Model Summary

| Step | -2 Log likelihood | Cox & Snell R Square | Nagelkerke R Square |
|------|---------------------|----------------------|---------------------|
| 1 | 36.925 ^a | .630 | .842 |

Table 4: Classification Table

| | | | 0 - 2.49 | 2.50 – 5.0 | |
|--------------------|-----|------------|----------|------------|------|
| Step 1 | GPA | 0 - 2.39 | 39 | 5 | 88.6 |
| | | 2.40 – 5.0 | 5 | 47 | 90.4 |
| Overall Percentage | | | | | 89.6 |

Table 5: Binary Logistic Regression Model

| | B | S.E. | Wald | Df | Sig. | Exp(B) | 95% C.I. for EXP(B) | |
|---------------------|---------|--------|--------|----|------|--------|---------------------|-------|
| | | | | | | | Lower | Upper |
| Step 1 ^a | | | | | | | | |
| MAT 101 | .171 | .064 | 7.190 | 1 | .007 | 3.186 | 1.047 | 1.344 |
| PHY 101 | .094 | .043 | 4.658 | 1 | .031 | 2.098 | 1.009 | 1.196 |
| PHY 107 | -.046 | .079 | .347 | 1 | .556 | .955 | .818 | 1.114 |
| STA 111 | .170 | .072 | 5.533 | 1 | .019 | 2.186 | 1.029 | 1.366 |
| STA 113 | .056 | .044 | 1.602 | 1 | .206 | 1.058 | .970 | 1.154 |
| CSC 101 | .045 | .075 | .356 | 1 | .551 | 1.046 | .902 | 1.212 |
| GST 101 | .102 | .057 | 3.225 | 1 | .073 | 1.107 | .991 | 1.237 |
| GST 107 | .141 | .095 | 2.195 | 1 | .138 | 1.152 | .955 | 1.389 |
| CHM 101 | .254 | .090 | 8.062 | 1 | .005 | 3.290 | 1.082 | 1.537 |
| BIO101 | .189 | .074 | 6.606 | 1 | .010 | 2.208 | 1.046 | 1.396 |
| Constant | -27.972 | 13.879 | 17.446 | 1 | .000 | .000 | | |

CONCLUSION

A robust logistic regression model was fitted to the raw scores obtained by 100-level statistics students of Prince Audu Abubakar University, Ayingba. Negerke R2 test shows that 84.2% variation in the log odd ratio was explained by the fitted model. Also, the percentage of unexplained variation, the Hosmer Lemeshow test and the classification results further confirmed the model is a good fit. From the results of the model, the odds of doing well in courses like MAT 101, PHY 101, STA 111, CHM 101 and BIO 101 are a significant predictor of a student's GPA. Students with low GPAs performed poorly in these courses.

RECOMMENDATIONS

Based on the results obtained from this empirical study, the following recommendations are put forward for possible consideration:

1. Extra learning periods for each of these courses should be added to the school lecture timetable to enable students to grasp the understanding of basic concepts of these courses.
2. Experienced educators should be assigned to handle these courses to enable them to simplify difficult concepts for students, especially the academically weak ones.

3. The university authority should ensure that the teaching and learning of courses like PHY 101, BIO 101, STA 111, MAT 101 and CHM 101 are adequately complemented with practical classes to concretize the learning experience.
4. The lecturers handling these courses should be enlightened on the significance of these courses in determining students' early academic achievement. This will go a long way in making them more focused on their tasks while also improving their pedagogy.
5. This empirical study should be carried out in other departments of the university to generally address the issue of the early abysmal academic performance of undergraduate students.

REFERENCES

- Bereket T.Z and Kidus M.A (2017). Binary Logistic Regression in Assessment and Identifying Factors that Influence Students' Academic Achievement: The Case of College of Natural and Computational Science, Wolaita Sodo University, Ethiopia. *Mathematical Theory and Modeling*. 7(1) 1 - 6
- Jimoh A., Sule O.B., Salawu A.S., Owonipa O.R and Ibrahim I.I (2021). A Study on Factors that influence students' Academic Performance in Kogi State University Anyigba: A Binary Logistic Regression Approach. *International Journal of Mathematics and Statistics Studies*. 9(4) 14 – 26
- Lowe, H., Cook, A. (2003). Mind the gap: are students prepared for higher education? *Journal of Further and Higher Education*, 27(1), 53–76.
- Oladebinu T.O., Amos A.A and Oyediran (2018). Factors Affecting Students' Academic Performance in Colleges of Education in SouthWest, Nigeria. *British Journal of Education*. 6(10) 43 – 56
- Ozguven, I. E. (1974). Non-cognitive factors affecting university students' Academic Achievements. *Hacettepe University Journal of Education*, 41, 313–324.
- Reno R. (2013). Factors Affecting Pupil's Performance in Public Primary Schools in Kenya Certificate of Primary Education Examinatin (KCPE) in Emgwen Division, Nandi District Kenya. *A Masters Project in Education Foundation of the University of Nairobi*
- Shangodoyin, D. K and Lasisi, T. A (2011). The Role of Statistics in National Development with Reference to Botswana and Nigeria Statistical Systems. *Journal of Sustainable Development*. 4(3), 131.
- Sule B.O, and Saporu F.W.O (2015). A Logistic Regression Model of Students' Academic Performance in University of Maiduguri, Maiduguri, Nigeria. *Mathematical Theory and Modeling*. 5(10), 12-1

