Mathematical Theory and Modeling
ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.3, No.11, 2013

Application of Ordinal Logistic Regression in the Study of Students' Performance

Adejumo, A. O.
Department of Statistics, University of Ilorin, Kwara State, Nigeria
aodejumo@unilorin.edu.ng
Adetunji, A. A.
Department of Mathematics and Statistics, Federal Polytechnic, Ado-Ekiti, Ekiti State, Nigeria
adecap4u@yahoo.com

Abstract
The problem of incessant decline in academic performance of Nigeria students in recent years cannot be over emphasized. Despite importance attached to academic performance, researchers have shown that students' performance is declining. Researches had also shown that there are a lot of factors responsible for this trend. Using data obtained from records of graduated students from the Faculty of Science, University of Ilorin for 2011/2012 academic session, mode of entry, age at entry, department, and sex of students are examined as factors that could contribute to students' performance. Ordinal Logistic Regression (Proportional Odds Model) is used to model the data and the results reveal that only sex of students is not a determinant factor of final grade that students may attained at graduation. This research also finds that there is equal chance for both male and female students to graduate from a university with First Class, hence, governments' policy on education should be focused on both gender instead of special attention usually given to female students. It has also been established that younger students perform better than the older ones; hence, age of students at entry into any of educational level should not be of major concerned but the ability of such student to cope with the demand of such level. It is also established that the highest odds of graduating with First Class is obtained by students who were admitted through Direct Entry (DE). Most of these students are educationally matured as they have spent at least two academic sessions in their previous school mostly Polytechnics, hence, in order to ensure that we have better graduates in our tertiary institutions, a policy that encouraged programmes such as Higher School Certificate (HSC) should be re-introduced into the system to ensure that only set of academically matured students get in different universities.
Keywords: Ordinal Logistic Regression, Odds Ratio, Link Function, Students' Performance

1.1 Introduction
Academic performance as a concept has become a source of concern to researchers, especially as the performance of the undergraduates is declining. The Nigerian society vests great emphasis on education because it is believed to be the only avenue for national development. This can only be achieved when undergraduates in various citadel of learning get actively involved in different academic activities which can enhance their academic performance. This is expected to lead to the technological advancement of the nation. In spite of these laudable values attached to academic performance, researchers (Ugoji 2008 and Egbule 2004) have shown that students' performance is declining. This could be because they are confronted with so many school and non-school related demands and responsibilities (Ukpong, 2007).
The measurement of student performance can be useful in universities, based on the notion that there are various factors that can contribute to the final grade attained by the students. A true and thorough understanding of the complex learning experience requires knowledge of various factors that determine the final grade of students. While examining the influence of age, financial status and gender on academic performance among undergraduates, Ebenuwa-Okoh (2010) used simple random sampling to select sample size of 175 respondents. The study revealed that gender, age and finance are not significant predictors of academic performance as there were no significant difference in academic performance based on age, gender and financial status. He then recommended that counselling centres should open to handle varying problems confronting student irrespective of age, financial status or gender. Also see Bedard and Dhuey (2006), Crawford et al. (2007), Pellizzari et al. (2011), Lao (1980), Kimball (1989), Wilberg and Lynn (1999), Stage and Kloosterman (1995), and Afuwape and Oludipe (2008) for more details.
This problem seems to be a major one that requires urgent and serious attention since students' academic performance affects the quality of human resources within the society, hence, the need for a research such as this. Using data obtained from records of graduated students from the Faculty of Science, University of Ilorin for 2011/2012 academic session, mode of entry, age at entry, and sex of students are examined as factors that could contribute to students' performance. Though, factors that could contribute to students' performance are inexhaustible.

When a continuous variable is treated as a categorical variable, it is called CATEGORIZATION of the continuous variable. In this research work, final grade of students is categorized, although continuous, grade is often treated as categorical in actual research for substantive and practical reasons.

2:1 Logistic Regression Model: To fit a binary logistic regression model, a set of regression coefficients that predict the probability of the outcome of interest are estimated. The same logistic model can be written in different ways. The general linear logistic regression model is defined as:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = logit(\pi_i) = \beta_{i0} + \beta_{i1}x_{i1} + \cdots + \beta_{ip}x_{ip} \qquad (1)$$

Above model can also be rewritten in terms of the probability of a positive response:

$$\pi_i = \frac{\exp\left(\beta_{i0} + \beta_{i1}x_{i1} + \cdots + \beta_{ip}x_{ip}\right)}{1 + \exp\left(\beta_{i0} + \beta_{i1}x_{i1} + \cdots + \beta_{ip}x_{ip}\right)} \qquad (2)$$

The quantity to the left of the equal sign is called a LOGIT. It's the log of the odds that an event occurs. (The odds that an event occurs is the ratio of the number of people who experience the event to the number of people who do not. This is obtained when the probability that the event occurs is divided by the probability that the event does not occur, since both probabilities have the same denominator and it cancels, leaving the number of events divided by the number of non-events). When there are more than two events, binary logistics regression model can be extended to multinomial regression model. However, for ordinal categorical variables, the drawback of the multinomial regression model is that the ordering of the categories is ignored.

A general class of regression models for ordinal data that utilize the ordinal nature of the data by describing various modes of stochastic ordering and this estimates the need for assigning scores or otherwise assuming cardinality instead of ordinality was developed by McCullagh (1980).

2.2 Ordinal Logistic Regression (OLR): In OLR, the dependent variable is the ordered response category variable and the independent variable may be categorical, interval or a ratio scale variable. When the response categories have a natural ordering, model specification should take that into account so that the extra information is utilized in the model. This ordering is incorporated directly in the way the logits is specified. Sometimes the responses may be some continuous variable, which is difficult to measure, so that its range is divided into J ordinal categories with associated probabilities $\pi_1, \pi_2, \ldots, \pi_J$. There are two commonly used models for this situation. (Das and Rahman, 2011, Ombui and Gichuhi, 2011)

2.3 Cumulative Logit Model: The cumulative odds for the jth response category is given by:

$$\frac{\pi_1 + \pi_2 + \cdots + \pi_j}{\pi_{j+1} + \pi_{j+2} + \cdots + \pi_J}, \qquad (3)$$

where J is the total number of response categories.

The cumulative logit model is:

$$log\left(\frac{\pi_1 + \pi_2 + \cdots + \pi_j}{\pi_{j+1} + \pi_{j+2} + \cdots + \pi_J}\right) = \beta_{0j} + \beta_{1j}x_{1j} + \cdots + \beta_{pj}x_{pj} \qquad (4)$$

The odds ratio for each predictor is taken to be constant across all possible collapsing of the outcome variable. When a testable assumption is met, the odds ratios in a cumulative logit model are interpreted as the odds of being "lower" or "higher" on the outcome variable, across the entire range of the outcome. The wide applicability and intuitive interpretation of the cumulative logit model are two reasons for its being considered the most popular model for ordinal logistic regression.

2.4 Proportional Odds Model (POM): The POM for ordinal logistic regression provides a useful extension of the binary logistic model to situations where the response variable takes on values in a set of ordered categories. The model may be represented by a series of logistic regressions for dependent binary variables, with common regression parameters reflecting the proportional odds assumption. Key to the valid application of the model is the assessment of the proportionality assumption (Brant, 1990). In the cumulative logit model, all the parameters depend on the category j. The proportional odds model is based on the assumption that the effects of the covariates $x_1, x_2, \ldots, x_p$ are the same for all categories, on the logarithmic scale. Thus, in this model, only the intercept term β_{0j} depends on the category j so that the model is:

$$log\left(\frac{\pi_1 + \pi_2 + \cdots + \pi_j}{\pi_{j+1} + \pi_{j+2} + \cdots + \pi_J}\right) = \beta_{0j} + \beta_1 x_1 + \cdots + \beta_p x_p$$

(5)

The appropriateness of the model can be tested separately for each variable by comparing it to a cumulative odds model with the relevant parameter not depending on *j*.

3.0 Methodology

<u>Link Function</u>: The link function is a transformation of the cumulative probabilities that allows estimation of the model. In ordinal regression (OR) analysis, the major link functions, e.g. logit, complementary-log-log (continuation ratio or proportional hazard), negative log-log, probit and Cauchit are used to build specific models. There is currently no universal method to help researchers choose which link function best fits a given dataset - only basic heuristics.

Different Link Functions and Forms in Ordinal Regression

Function	**Form**	**Typical Application**
Logit	$\log\left(\frac{F_k(x_i)}{1-F_k(x_i)}\right)$	Evenly distributed categories
Complementary log-log	$\log[(-\log(1-F_k(x_i))]$	Higher categories are more probable
Negative log-log	$-\log(F_k(x_i)))$	Lower categories are more probable
Probit	$\Phi^{-1}(F_k(x_i))$	Normally distributed latent variable
Cauchit	$\tan\left(\pi\left(F_k(x_i)-0.5\right)\right)$	Outcome with many extreme values

Hahn and Soyer (2010) found clear evidence that model fit can be improved by the selection of the appropriate link even in small data sets.

3.1 Choice of Baseline Category: Except for the dependent variable (final grade) where First Class (the desirable grade for any student) is used as the base or reference category, the baseline category for other factors are randomly chosen since ordinal regression provides a technique for interpreting for the factors irrespective of the choice of baseline category. Ordinal regression provides a set of coefficients from each comparison. The coefficients for the reference or baseline category are all zeros, similar to the coefficients for the baseline category for a dummy-coded variable.

3.2 Odds Ratios and Model Interpretation: Often, it is not easy to directly interpret the model parameters β_0, β_1, …, β_p. Odds ratios (OR) often provide a much easier interpretation. OR can be converted to its complement by dividing the OR by 1, e.g. OR (Category 1/ Category 2) = $\frac{1}{OR\ for\ category\ 1}$.

If the final grade of students can be ranked as Pass (1), Third Class (2), Second Class Lower (3), Second Class Upper (4) and First Class (5), in ordinal logistic regression, the event of interest is observing a particular grade. For the rating of students' grade, the following odds can be modelled:

$$\theta_1 = \frac{Prob\left(grade\ of\ 1\right)}{Prob\left(grade\ greater\ than\ 1\right)}$$

$$\theta_2 = \frac{Prob\left(grade\ of\ 1\ or\ 2\right)}{Prob\left(grade\ greater\ than\ 2\right)}$$

$$\theta_3 = \frac{Prob\left(grade\ of\ 1,\ 2\ or\ 3\right)}{Prob\left(grade\ greater\ than\ 3\right)}$$

$$\theta_4 = \frac{Prob\left(grade\ of\ 1,\ 2,\ 3\ or\ 4\right)}{Prob\left(grade\ greater\ than\ 4\right)}$$

The last category (First Class) does not have an odds associated with it since the probability of scoring up to and including the last score is 1. Hence, this category serves as the base or reference category. Also, since First Class is desirable grade for any reasonable students being admitted into any University, this research is limited to the fourth model; $\theta_4 = \frac{Prob(grade\ of\ 1,\ 2,\ 3\ or\ 4)}{Prob(grade\ greater\ than\ 4)}$

Often, it is not easy to directly interpret the model parameters $\beta_0, \beta_1, \ldots, \beta_p$. Odds ratios usually provide a much easier interpretation. With reference to this research, the dependent variable (Final Grade) has 5 categories, therefore,

$$[\ln(\theta]_4) = \ln[Prob(grade\ of\ 1,\ 2,\ 3\ or\ 4)] - \ln[Prob(grade\ greater\ than\ 4)] = \beta_{1j}$$

We estimate the odds ratio by $\hat{\theta}_j = \exp(b_{1j})$, where b_{1j} is the estimate of parameter β_{1j}. For instance, $\theta_j = 4$ implies that the odds for a student from a reference category graduating with first class is four times higher for response category.

3.3 Test of Parallelism: One of the assumptions of ordinal regression is that regression coefficients are the same for all categories. If the assumption of parallelism is rejected, it is better to consider using multinomial regression, which estimates separate coefficients for each category. If the null hypothesis is rejected at a specific level of significance, this implies that it is possible that the link function selected is incorrect for the data or that the relationships between the independent variables and logits are not the same for all logits. However, the test of the proportional odds assumption has been described as anti-conservative, that is, it nearly always results in rejection of the proportional odds assumption (O'Connell, 2006, p.29) particularly when the number of explanatory variables is large (Brant, 1990), the sample size is large (Allison, 1999; Clogg and Shihadeh, 1994)

3.4 Parameter Estimates of the Coefficients: The table of Parameter estimates tells us specifically about the relationship between our explanatory variables and the outcome. The threshold coefficients are not usually interpreted individually; they just represent the intercepts, specifically the point (in terms of a logit) where students might be predicted into the higher categories. While we do not usually have to interpret the threshold parameters directly, it can be used to explain how the model works (*in obtaining* R^2).

3.5 Measuring Strength of Association: Various R^2-like statistics can be used to measure the strength of the association between the dependent variable and the predictor variables. They are not as useful as the R^2 statistic in regression, since their interpretation is not straightforward. Three commonly used statistics are:

Cox and Snell R^2: $$R^2_{CS} = 1 - \left(\frac{L(B^{(0)})}{L(\hat{B})}\right)^{\frac{2}{n}}$$

Nagelkerke's R^2: $$R^2_N = \frac{R^2_{CS}}{1 - L(B^{(0)})^{\frac{2}{n}}}$$

McFadden's R^2: $$R^2_M = 1 - \left(\frac{L(\hat{B})}{L(B^{(0)})}\right)$$

where $L(\hat{B})$ is the log-likelihood function for the model with the estimated parameters and $L(B^{(0)})$ is the log-likelihood with just the thresholds, and n is the number of cases (sum of all weights). What constitutes a "good" R^2 value varies between different areas of application. While these statistics can be suggestive on their own, *they are most useful when comparing competing models for the same data. The model with the largest* R^2 *statistic is "best" according to this measure.*

4. Descriptive Analysis

Examining the data (appendix) before building the model, figures 1 to 4 show cumulative percentage plot of the four factors (Department, Mode of entry, Sex, and age at entry) against the final grade.

Figure 1 suggests that the odds of graduating with First Class is likely to be highest in the department of Microbiology while figure 2 shows that students who were admitted through DE are most likely to achieve higher final grade compared to those who are admitted through UME while those who are admitted through REM have the least odds of achieving higher final grade.

Though points for male students appear to be below those for female students at some points, these points do not completely dominate each other since there are some points where cumulative points for female appear below those for male. Hence, this gives an intuitive feeling that sex may not have effect on odds of achieving final grade. Figure 4 however suggests that students those whose ages at entry are above 20 years are above those of students whose ages at entry is below or equal to 20 years. This suggests that the odds of achieving First Class for students in lower age category.

Choosing Appropriate Link Function

The plot of the final grade (dependent variable) against the frequency of occurrence (appendix) gives the idea of the most appropriate link function to be used in modelling an ordinal categorical variable. The figure below suggests PROBIT link function is appropriate.

From Table 2 and 3, the model for Final Grade on Sex does not give a significant improvement over the baseline intercept-only model. This tells that the model does not give better predictions than if we just guessed based on the marginal probabilities for the outcome categories. All other models give a significant improvement over the baseline intercept-only model. This indicates that all the models except Final Grade on Sex give better predictions than if we just guessed based on the marginal probabilities for the outcome categories.

From Table 4, although, these R^2 appear very low, when comparing values with other link function, this value (under Probit Link Function) is the best obtainable. While sex of students does not explain any proportion of the variation between students in their attainment (final grade), mode of entry has the highest explanation for the variations in the final grade of students.

What constitutes a "good" R^2 value varies between different areas of application. While these statistics can be suggestive on their own, they are most useful when comparing competing models for the same data. The model with the largest R^2 statistic is "best" according to this measure. Pseudo R^2 is not as useful as the R^2 statistic in regression, since their interpretation is not straightforward.

Also from Table 5, the assumption of proportional odds is rejected for model of Final Grade on Sex. This means that the general model gives a significantly different fit to the data than model with a separate set of coefficients for each threshold. This assumption is however not rejected for other factors. This means that the ordinal regression model of Final Grade on Age at entry, department, and mode of entry gives a result that is not significantly different fit to the data than model with a separate set of coefficients for each threshold.

Table 6 gives the estimates of coefficient for male to be 0.031, taking its exponent to find the OR with female as the base: exp(0.031) = 1.0315. To find the complementary OR with male as the base, we either reverse the sign of the coefficient before taking the exponent, exp(-0.031) = 0.9695 or we find

$$\frac{1}{OR\ for\ female} = \frac{1}{1.0315} = 0.9695$$

.

The P-value of estimate of sex (0.728) indicates that sex of students is not a significant contributor to the final grade attained by the students. Though, the odds ratio (Female/Male) is slightly higher than that for (Male/Female), both are approximately equal i.e. the odds of male students achieving a higher final grade are approximately the odds for female students'. This indicates that neither of the odds dominates another (neither male nor female students achieve higher Final Grade than one another)

In Table 7, the Odds Ratio (> 20 years /≤ 20 years = 1.3499) indicates that the odds for a graduating student whose age is ≤ *20 years* to graduate with First Class is approximately 1.35 more likely compared to those whose ages above 20 years as at the time of entry. (DE/REM) = exp(-1.215) = 0.2967, (DE/UME) = exp(-0.417) = 0.6590 and

$$(REM/UME) = \exp\{-0.417-(-1.215)\} = \frac{\left(\frac{DE}{UME}\right)}{\left(\frac{DE}{REM}\right)} = 2.2211$$

Lastly in Table 8, the odds ratio (DE/REM = 0.2967) indicates the odds for a graduating student whose mode of entry is REM graduating with First Class is 0.2967 more likely compared to a student who got admitted through DE. This means that the odds for a DE student graduating with First Class is 3.37 (0.2967^{-1}) more likely compared to a student who is admitted through REM. Also, the odds for a DE student graduating with First Class is 1.52 (0.6590^{-1}) more likely compared to a student who is admitted through UME while it is 2.22 more likely for a student who is admitted through UME compared to a student who is admitted through REM.

5 Conclusion

From the analyses carried out, the following conclusions are drawn:

- Among the factors examined, only sex does not depend on the final grade.
- Models of all the factors except sex give better predictions than if we just guessed based on the marginal probabilities for the outcome categories.
- Sex gives the least explanation about the variation in the final grade while the mode of entry gives the highest.

- The assumption of proportional odds is only rejected for sex among all the factors considered. Hence, for sex, the general model gives a significant different fit to the data than model with a separate set of coefficients for each threshold.
- Only sex is also found not to be a significant to the final grade attained by the students.
- All these indicators reveal that sex of students is not a determinant factor of final grade that students may attained at graduation.
- Neither male nor female students achieve higher Final Grade than another.
- The odds that a student with age ≤ *20 years* graduating with First Class is approximately 1.35 times more likely compared to another who is above 20 years as at the time of entry.
- The odds for a DE student graduating with First Class is more than three times compared to a student who is admitted through REM.
- The odds for a DE student graduating with First Class is more than one and a half compared to a student who is admitted through UME.
- It is more than twice more likely for a student who is admitted through UME to graduate with First Class compared to a student who is admitted through REM.

Appendix
Table 1: Data Summary

	Variable	**Code**	**N**	**Marginal Percentage**
Final Grade	Pass	[CGPA = 1]	9	1.4%
	Third Class	[CGPA = 2]	129	20.7%
	Second Class Lower	[CGPA = 3]	270	43.3%
	Second Class Upper	[CGPA = 4]	203	32.6%
	First Class	[CGPA = 5]	12	1.9%
Department	Plant Biology	[dept=1]	63	10.1%
	Geology & Mineral Science	[dept=2]	63	10.1%
	Chemistry	[dept=3]	50	8.0%
	Microbiology	[dept=4]	91	14.6%
	Biochemistry	[dept=5]	69	11.1%
	Physics	[dept=6]	48	7.7%
	Industrial Chemistry	[dept=7]	50	8.0%
	Mathematics	[dept=8]	59	9.5%
	Zoology	[dept=9]	47	7.5%
	Statistics	[dept=10]	83	13.3%
Age at entry	≤ 20 years	[Age=1]	395	63.4%
	> 20 years	[Age=2]	228	36.6%
Sex	Male	[Sex=1]	370	59.4%
	Female	[Sex=2]	253	40.6%
Mode of Entry	REM (Remedial)	[Mode=1]	162	26.0%
	UME (JAMB Exam)	[Mode=2]	395	63.4%
	DE (Direct Entry)	[Mode=3]	66	10.6%

.

Table 2: Chi-Square Tests of Independence

Factor	Pearson χ^2 Value	df	P value	Decision
Sex * Final Grade	10.714	4	.030	Sex does not depend on the final grade obtained at 2.5% significance level.
Department * Final Grade	128.834	36	.000	Department is dependent on the final grade
Age at entry * Final Grade	23.524	4	.000	Age at entry is dependent on the final grade
Mode of entry * Final Grade	105.006	8	.000	Mode of entry is dependent on the final grade

Table 3: ORDINAL REGRESSION
Model Fitting Information

	Model	-2 Log Likelihood	Chi-Square	df	Sig.
Sex * Final Grade	Intercept Only	47.887			
	Final	47.766	.121	1	.728
Department * Final Grade	Intercept Only	231.703			
	Final	176.001	55.702	9	.000
Age at entry * Final Grade	Intercept Only	59.540			
	Final	48.454	11.086	1	.001
Mode of entry * Final Grade	Intercept Only	143.131			
	Final	65.198	77.933	2	.000

Table 4: Measure of Determination

Model	Nagelkerke R^2	Cox and Snell R^2	McFadden R^2
Sex * Final Grade	0.000	0.000	0.000
Department * Final Grade	0.094	0.086	0.038
Age at entry * Final Grade	0.019	0.018	0.007
Mode of entry * Final Grade	0.130	0.118	0.053

Table 5: Test of Parallel Lines

	Model	-2 Log Likelihood	Chi-Square	df	Sig.
Sex * Final Grade	Null Hypothesis	47.766			
	General	35.379	12.386	3	.006
Department * Final Grade	Null Hypothesis	176.002			
	General	169.096	6.905	27	.265
Age at entry * Final Grade	Null Hypothesis	48.454			
	General	46.066	2.388	3	.146
Mode of entry * Final Grade	Null Hypothesis	65.198			
	General	63.496	1.702	6	.576

Table 6: Parameter Estimates (Sex*Final Grade)

		Estimate	Std. Error	Wald	df	Sig.	95% C.I.	
							Lower	Upper
Threshold	[CGPA = 1]	-2.166	.140	239.404	1	.000	-2.440	-1.891
	[CGPA = 2]	-.749	.076	96.430	1	.000	-.899	-.600
	[CGPA = 3]	.417	.073	32.112	1	.000	.272	.561
	[CGPA = 4]	2.089	.129	262.550	1	.000	1.836	2.342
Location	[Sex=1]	.031	.088	.121	1	.728	-.141	.203
	[Sex=2]	0	.	.	0	.	.	.

Odds Ratio

Female/Male	1.0315
Male/Female	0.9695

Table 7: Parameter Estimates (Age at entry*Final Grade)

		Estimate	Std. Error	Wald	df	Sig.	95% C.I.	
							Lower	Upper
Threshold	[CGPA = 1]	-2.034	.142	205.069	1	.000	-2.312	-1.756
	[CGPA = 2]	-.582	.078	55.423	1	.000	-.735	-.429
	[CGPA = 3]	.595	.078	57.861	1	.000	.441	.748
	[CGPA = 4]	2.273	.134	287.977	1	.000	2.011	2.536
Location	[Age=1]	.300	.090	11.111	1	.001	.124	.476
	[Age=2]	0	.	.	0	.	.	.

Odds Ratio

> 20 years/≤ 20 years	1.3499
≤ 20 years/> 20 years	0.7408

Table 8: Parameter Estimates (Mode of entry * Final Grade)

		Estimate	Std. Error	Wald	df	Sig.	95% Con. Int	
							Lower	Upper
Threshold	[CGPA = 1]	-2.970	.199	221.890	1	.000	-3.361	-2.580
	[CGPA = 2]	-1.401	.145	92.854	1	.000	-1.686	-1.116
	[CGPA = 3]	-.141	.138	1.046	1	.307	-.412	.129
	[CGPA = 4]	1.611	.168	92.380	1	.000	1.283	1.940
Location	[Mode=1]	-1.215	.164	54.815	1	.000	-1.537	-.893
	[Mode=2]	-.417	.146	8.121	1	.004	-.704	-.130
	[Mode=3]	0	.	.	0	.	.	.

Odds Ratio

(DE/REM)	0.2967
(DE/UME)	0.6590
(REM/UME)	2.2211

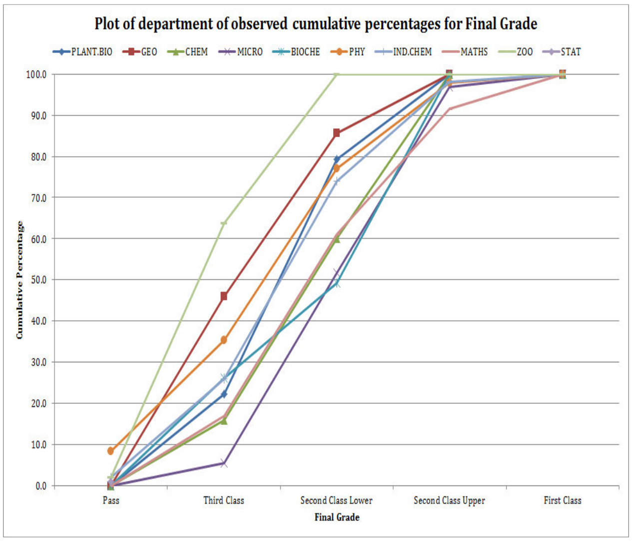

Figure 1: Plot of observed cumulative percentages for department

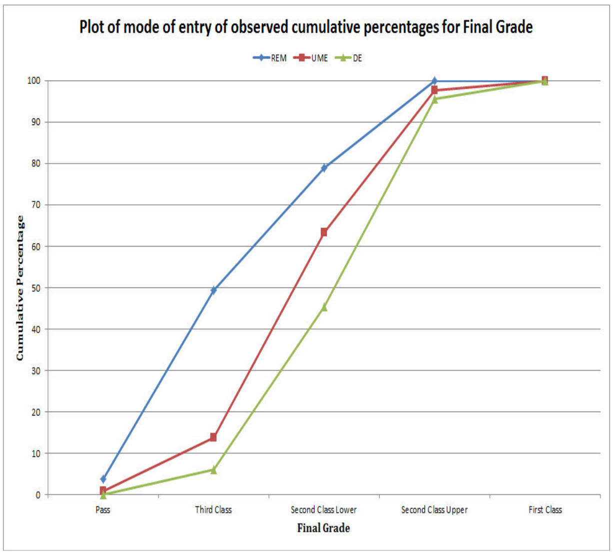

Figure 2: Plot of observed cumulative percentages for mode of entry

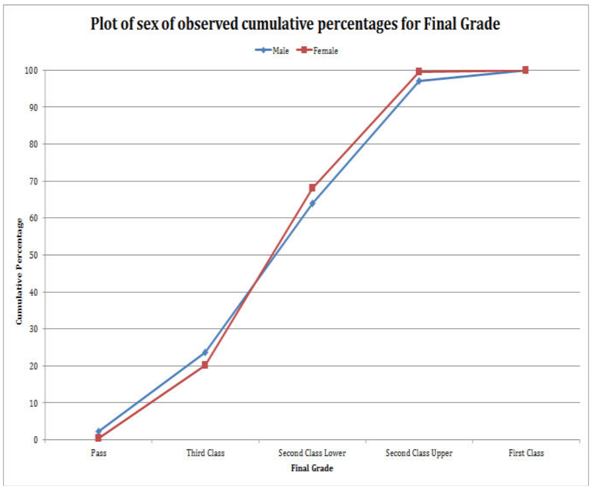

Figure 3: Plot of observed cumulative percentages for gender

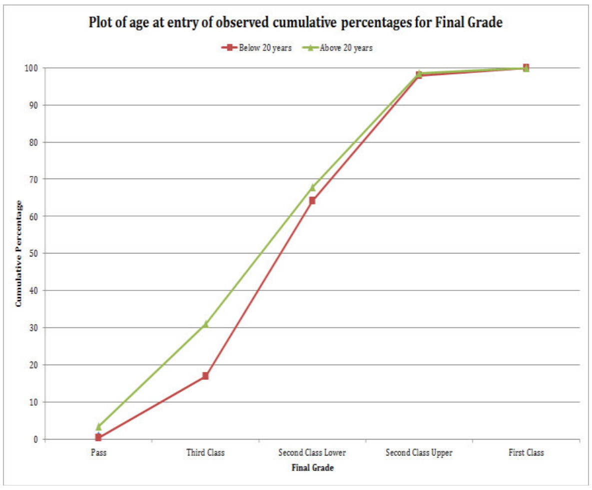

Figure 4: Plot of observed cumulative percentages for age at entry

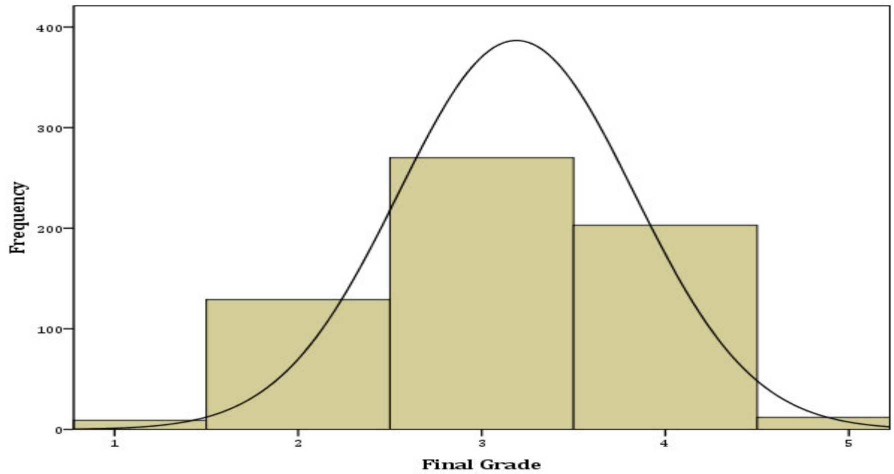

Figure5: Plot of Final grade against the frequency

References

Afuwape, M. and Oludipe, D. (2008): *Gender differences in Integrated Science achievement among pre-service teachers in Nigeria.* Educational Research and Review, 3 (7), 242-245.
Allison, P. D. (1999): *Logistic regression using the SAS system: Theory and application.* Cary, NC.: SAS Institute.
Bedard K. and Dhuey E. (2006): *The persistence of early childhood maturity: International evidence of long-run age effects.* Quarterly Journal of Economics, 121(4):1437–1472.
Brant, R. (1990): *Assessing proportionality in the proportional odds model for ordinal logistic regression*, Biometrics 46(4), 1171–1178.
Clogg, C., and Shihadeh, E. S. (1994): *Statistical models for ordinal variables.* Thousand Oaks, California: Sage publications.
Cox, D. R., and Snell, E. J. (1989): *The Analysis of Binary Data*, 2nd ed. London: Chapman and Hall.
Crawford C., Dearden L., and Meghir C. (2007): *When you are born matters: The impact of the date of birth on child cognitive outcomes in England.* Institute for Fiscal Studies.
Das S. and Rahman M. R. (2011): *Application of Ordinal Logistic Regression Analysis in determining risk factors of child malnutrition in Bangladesh* Nutrition Journal 2011, (10): 124
Ebenuwa-Okoh E. E. (2010): *Influence of Age, Financial Status, and Gender on Academic Performance among Undergraduates,* Journal of Psychology, 1 (2), pp. 99 – 103
Egbule J. F. (2004): *Practical Guide to a Successful Project or Thesis in Writing and Defence. Owerri:* Whyte and Whyte Publishers.
Hahn D. Eugene and Soyer Refik (2010): *Probit and Logit Models: Differences in the Multivariate Realm* Salisbury University, Salisbury, MD, USA. , The George Washington University, Washington, DC, USA.
Kimball, M. M. (1989): *A New Perspective on Women's Math Achievement,* Psychological Bulletin, 105, pp. 198-214.
Lao, R. C. (1980): *Differential Factors Affecting Male and Female Academic Performance in High School,* The Journal of Psychology, 104, pp.119-127.
O'Connell, A. (2006): *Logistic regression models for ordinal response variables*. Thousand Oaks: Sage publications.
Ombui, G. M. and Gichuhi, A. W. (2011): *Using Ordinal Regression Modelling to evaluate the satisfaction of Jomo Kenyatta University of Agriculture and Technology, Faculty of Science Students,* JAGST Vol 13(1)
Pellizzari M. and Billari F. C. (2012): *The younger, the better? Age related differences in academic performance at University,* Journal of Population Economics, vol 25 (2), pp 697 - 739
Stage, F. K. and Kloosterman, P. (1995): *Gender, Beliefs, and Achievement in Remedial College-Level Mathematics,* Journal of Higher Education, 66, pp. 294-311.
Ugoji F. N. (2008): *The impact of Counselling on the Academic Performance of Secondary School Students.* Africa Journal for Inter-Disciplinary Studies, 8(2): 67 – 73
Ukpong D. E. (2007): *Teachers' Social acceptance and Junior Secondary Student Academic Performance in Social Studies in Uyo Metropolis.* Educational Insight. 20(12): 84-90.
Wilberg, S. and Lynn, R. (1999): *Sex Differences in Historical Knowledge and School Grades: A 26 Nation Study,* Personality and Individual Differences, 27, pp. 1221 – 1229.