European Journal of Statistics and Probability
Vol.4, No.2, pp.49-58, May 2016
Published by European Centre for Research Training and Development UK (www.eajournals.org)

STATISTICAL MODELING OF THE PERFORMANCE OF FIRST YEAR EVENING SCHOOL OF TAKORADI POLYTECHNIC: DOES PRIOR EDUCATION LEVEL, AGE AT THE TIME OF ADMISSION, AND COURSE LOAD MATTER?

Mohammed Frempong[1], Philomena Aboagye[2], Linda Duncan[3]

[1] Department of Mathematics and Statistics, Takoradi Polytechnic, Takoradi, Ghana
[2]School of Applied Science, Takoradi Polytechnic, P. O. Box 256, Takoradi, Ghana
[3]Department of Hospitality Management, Takoradi Polytechnic, Takoradi, Ghana

ABSTRACT: *Students' admission into tertiary institutions is an important issue that has generated a lot of public discussion in recent times. The pre-requisite criteria for admitting most students into regular school programmes at tertiary level in Ghana, is usually based on senior high school grades or scores. Having an in-depth knowledge about factors affecting students' performance more especially during the first year of the Evening School is very important. Knowledge acquired in this area will help school authorities design admission programmes that will render their regular and alternative school sections such as distance and evening school programmes more sustainable. This paper reports on analysis of factors affecting the academic performance of six hundred and ten (610) first year evening school students from 2011 to 2013 academic years of Takoradi Polytechnic. Among other things, the study revealed that* 25.4% of variance in evening school students' academic performance was explained by students' age at the time of admission, West African Senior School Certificate Examination (WASSCE) or other examination scores, and total credit hours taken during the first year. Students' total credit hours taken during the first year, was the best predictor of students' performance. Also, as students' total credit hours increases, students tend to perform poorly or students perform well as total credit hours decreases. Management of the Polytechnic are encouraged to explore other factors that significantly affect students' performance.

KEYWORDS: Evening School, Students' Cumulative Grade Point Average, Regression Analysis

INTRODUCTION

Many developing countries including Ghana have now cherished the need to commit themselves into science and technology education, as a way of creating wealth for poverty reduction. With Ghana Government intension of converting the ten polytechnics to technical universities, it has now formed a collaboration with the German Government in providing technical support to facilitate the training of middle level man power needs for socio – economic development (MOE, 2014). Products of Polytechnics in Ghana are normally graduates from Secondary, Technical, and Vocational Schools who have interest in acquiring knowledge and handy skills in disciplines such as; Applied Art, Business Study, Applied Science, and Engineering. These students are enrolled into regular programmes where they are trained to acquire practical skills in various disciplines. Polytechnic graduates are more practically trained and hence are generally, compared to offer quality industrial service than products of other tertiary institutions. Because of this quality, many people who have interest in providing practically-oriented middle level manpower have resorted to Polytechnic education. Large numbers of students keep on enrolling into polytechnic education. However,

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

both human and material resources are inadequate to absorb all of them at a go. There is therefore the need to provide alternative programmes that will meet other applicants' needs. This leads to creation of a second section of the programme called Evening School. This programme was designed for people who wanted to work and at the same time pursue tertiary education to improve upon their areas of expertise. In this respect, the evening school programme aims at providing skill training to students (typically workers) who do not find the regular session convenient but would want to further their education, making skill training more accessible to adult students. Although products from the polytechnics are ranked second to their counterparts from the universities by many Ghanaians, the services rendered by these graduates leaves much to be desired. Products from most polytechnics are found in industries and the service sectors of Ghana and they are usually identified by their excellent skills applied in various professions (Owusu- Agyeman, 2006). With such good performance in practical skills, one pertinent question that arises is; what is the academic output of the polytechnic students during the first year, more especially the evening school? Because, the academic achievement of evening school is deemed to link with a lot of factors. As many are mature students who have other family commitments and responsibilities, the issue that arises is what factors will significantly affect their academic achievement in the first year of the evening school. Having an in-depth knowledge about these factors will enable the polytechnic address the problem thereby making the programme more sustainable. This study therefore examines prior tertiary school factors such as; age of the student at the time of enrollment, WASSCE or other examinations score, and total credit hours taken during first year. In this respect, it seeks answers to; how these prior school factors predict students' cumulative grade point average, which variable is the best predictor of students' cumulative grade point average? And how the best predictor relates to students' academic achievement in the first year. The data was analysed using Statistical Products and Service Solutions (SPSS16). The regression analysis method used in this research is reviewed below.

MATERIALS AND METHODS

This area briefly discusses multiple linear regression analysis in general, and specifically reviews simultaneous or standard regression, fundamental assumptions of linearity of the phenomenon under study, normality of the distribution of the residuals, independence of the residuals, and homoscedasticity of the residuals. It further discusses the basic equations that were used to analyse the data. Multiple regression analysis is a family of multivariate statistical techniques that can be used to examine relationship between single continuous dependent variable and a number of (generally continuous) independent variables. It is one of the most widely used statistical methods for solving real life problems such as the effect of one variable on another, describing the nature of relationship between variables, making predictions and so on. The technique involves developing a mathematical model that examines relationship between dependent variable and independent variables. In general, multiple regression is classified into three main types. These are; standard or simultaneous, stepwise and hierarchical regression. This paper employs concepts and theories of standard or simultaneous multiple regression.

Standard or Simultaneous Multiple Regression

In this technique, all the predictor (independent) variables are considered in the equation concurrently, with each predictor variable being assessed in respect of its predictive power over

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

the predictive power of all other predictor variables. It is employed in situations where a set of combined predictor variables accounts for variance in the dependent variable. Further, this technique also reports on the amount of unique variance of the dependent variable that each independent variable explains. Let $x_1, x_2, \ldots, x_k$ be k predictor variables considered to be related to a response (dependent) variable, Y . Then the standard linear regression model states that Y is composed of a mean, which depends in a continuous manner on the $x_i's$, and the random error, ε, which accounts for measurement error and the effects of other variables not explicitly expressed in the model (Johnson and Wichern, 1992).

The values of the predictor variables in this case are considered as fixed, while error is seen as random variable which is characterised by certain assumptions. The linear regression equation with single response in this situation is given by

$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$. This can ordinarily be seen as:

[Response]= mean (depending on $x_1, x_2, \ldots, x_k$)] + [error]. The term linear denotes the fact that the mean is a linear function of the unknown parameters; $\beta_0, \beta_1, \beta_2, \ldots, \beta_k$. If n independent observations made on Y are associated with values of x_i then the data setting can be modeled as follows:

$$
\begin{aligned}
Y_1 &= \beta_0 + \beta_1 x_{11} + \beta_1 x_{12} + \cdots + \beta_k x_{1k} + \varepsilon_1 \\
Y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \cdots \quad \beta_k x_{2k} + \varepsilon_2 \\
\vdots & \quad \vdots \quad \vdots \qquad \vdots \qquad \vdots \quad \vdots \\
Y_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_k x_{nk} + \varepsilon_n
\end{aligned}
$$

in this case one of the critical parts of the model, the error terms are assumed to: exhibit homoscedasticity in which case the variance of the errors is constant at all levels of the predictor variables, be normally distributed, and be independent of each other. Mathematically, we can express that as:

{Var $(\varepsilon_j) = \sigma^2$ (constant), $E(\varepsilon_j) = 0,$ and Cov $(\varepsilon_i, \varepsilon_k) = 0, j \neq k$} $\ldots\ldots\ldots\ldots(1)$.

The above model can be transformed into matrix notation as follows:

$$
\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad \text{or algebraically}
$$

$$
\begin{array}{cccc} Y = & Z & \beta + & \varepsilon \\ (n \times 1) & (n \times (r+1) & ((r+1) \times 1) & (n \times 1) \end{array}
$$

(Johnson and Wichern, 1992).

The mathematical properties of the error variable in equation (1) can be reduced to:

$E(\varepsilon) = 0,$ and (b) $Cov(\varepsilon) = E(\varepsilon\varepsilon') = \sigma^2$I. It should be noted that one in the first column of design matrix, Z are the multiplier of the constant term β_0.

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

Inspection of the Assumptions

As stated earlier, one of the requirements of linear regression is that the variables under study display linearity. In this case, it is assumed that a linear relationship based on correlations exists among the predictor and response variables. Mathematically, if (x_1, y_1), (x_2, y_2)… (x_n, y_n) constitute n pairs of measurements on the two random variables X and Y, then the correlation coefficient, denoted by r is given by

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{[\sum_{i=1}^{n}(x_i-\bar{x})^2(y_i-\bar{y})^2]}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(1)}$$ (Gordor and Howard, 2000).

The value of r can assume any value from -1 to1, both inclusive $(-1 \leq r \leq 1)$. It is equal to-1 or 1, if and only if, all points of the n pairs of measurements lie exactly on a straight line. In this case we have perfect negative and perfect positive correlations respectively. If $r = 0$, we say that the two variables X and Y are uncorrelated or there is no linear relation between X and Y. In this respect we can use mathematical transformations such as; squaring, taking square root or logarithm of the variables depending on the nature of the data. Further, $r < 0$ implies a tendency for one value in the pair to be larger than its average when the other is smaller than its average, and $r > 0$ implies a tendency for one value of the pair to be large when the other value is large and also for both values to be small together (Johnson and Wichern, 1992). The sample correlation matrix,R, employed in the assessment of the linear relationship is given by

$$R = \begin{pmatrix} 1 & r_{12} & r_{13} & \cdots & r_{1p} \\ r_{21} & 1 & r_{23} & \cdots & r_{2p} \\ r_{31} & r_{32} & 1 & \cdots & r_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & r_{p3} & \cdots & 1 \end{pmatrix}$$

The array R consists of p $rows$ and p $columns$. The elements along the diagonal are one (1) each because they represent correlation between a variable and itself. The diagonal line serves a mirror line where elements above the upper diagonal are the same as elements below lower diagonal. Hence one half of the elements of the matrix can be used for interpretation without loss of information.

Alternatively, linearity can also be examined graphically by residual-predicted dependent values plots, in which case the residual points must be distributed evenly about the horizontal line zero, without displaying any trend or curvilinear shape. In situations where this assumption is violated, one can remedy it by; arithmetic transformation, direct inclusion of nonlinear relationship such as polynomial terms in the regression model, and the use of specialized nonlinear regression method designed to cater for curvilinear effect of the independent variables (Hair et al, 2006). Further, collinearity diagnostics to avert any problem of multicollinearity that is not noticeable in the correlation matrix is examined by two key statistics. These are tolerance level (TL) and variance inflation factor (VIF). The tolerance level measures the amount of variability of a specified predictor variable that is not accounted for by other predictor variables in the model. The tolerance level (TL) and the VIF values are calculated by $1 - R^2$ and $\frac{1}{TL}$ respectively (Pallant, 2005). Where R^2 statistic measures how much of the variance in the criterion variable is accounted for by the model. Pallant (2005) further clarifies that if; TL < 0.1 and VIF > 10 then multicollinearity exist between the

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

variables. Hence we need to remove one of the highly inter correlated predictor variables from our model. The next commonly considered assumption is the normality of the distribution of the residuals. This can be studied by the use of histogram of the residuals; in this case the distribution of the errors should approach normal distribution. Or a more enhanced method of normal probability plot is examined. In this respect, standardised residuals are compared with normal distribution. If the output of the standardised residuals traces a straight diagonal line from the bottom left to the top right corner then the residuals satisfy the normality assumption. Also, to fulfill the assumption that the residuals are independent, a graph of residuals versus predicted values of the dependent variable should show a null plot; in this situation the pattern should appear random rather than displaying a consistent pattern in the residuals. The final assumption is the ability of the residuals to show constancy at all levels of the independent variables. This is called homoscedasticity. This is checked by constructing a graph of residuals against predicted dependent values; in this case the variance of the error variable, ${\sigma_\varepsilon}^2$ is constant.

Last but not least the presence of outliers which merit deletion is examined by the use of technique suggested by Tabachnick and Fidell (2001). They contend that outliers are observations with standardised residuals, $e_i, i = 1,2, \ldots, n$ of more than 3.3 $or\ less\ than\ -3.3$. In this connection, residuals from large samples may be ignored if there is a cause to justify their deletion. Further, they argue that outliers can also be assessed by the use of Mahanalobis distances produced from the data analysis. In this situation, critical chi-square values using the number of independent variables as degrees of freedom are employed. Analysis with three (3) independent variables at alpha level of 0.001, the cut- off critical value is 16.27. Also, pertinent information is the unusual cases obtained from the case wise diagnostics of the regression analysis. As mentioned earlier, Tabachnick and Fidell (2001), described outliers as observation with residuals, e_i, satisfying the set;$\{ e_i: e_i < -3.3 \cup e_i > 3.3\}$, where $e_i, i = 1,2, \ldots, n$ is standardised residual. It is expected that ideally, for normally distributed sample, one per cent (1%) of such unusual cases will fall outside this range. These unusual observations considered as having undue influence on the results for our model must be supported by Cook's Distance (D_i) cut-off value of unity.

This means that observations with $D_i > 1$ are considered to be influential. Mathematically, this can be expressed as $D_i = \frac{{r^2}_i}{p} \frac{h_{ii}}{(1-h_{ii})}$, $i = 1,2, \ldots, n$. Where $\sum_{i=1}^{n} h_{ii} = \text{rank}(H) = \text{rank}(X) = p,$ and r_i is the studentised residual. Note that, apart from the constant p, D_i is the product of the square of the i^{th} studentised residual and $\frac{h_{ii}}{(1-h_{ii})}$. This ratio can be shown to be the distance from a chosen vector, X_i to the centroid of the remaining data. Thus, D_i consists of a component that reflects how well the model fits the i^{th} observation y_i and a component that measures how far that point is from the rest of the data (Montgomery, 1996). Where h_{ii} is the ith diagonal element of an $n \times n$ matrix, H called "hat" where $H = (X'X)^{-1}X'$ which is very key in regression analysis. The results of these mathematical analyses applied to the data are discussed in the results of section 3 below.

RESULTS

Examining the assumptions

This section discusses and assesses some principal issues of; multicollinearity, outliers,

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

normality, linearity, homoscedasticity, and independence of residuals that govern regression analysis from the data output. The results are presented in Table 1 below.

Multicollinearity. The correlations between students' cumulative grade point average, CGPA (Y) , age of the student at the time of admission, (X_1), WASSCE or other examination score, (X_2), and the total credit hours taken during the first year, (X_3), in the model under discussion is shown in Table1. It is discovered that apart from total credit hours which correlates $(r = -0.469)$ substantially with CGPA, age at the time of admission, and WASSCE or other examination score, weakly relate to CGPA with correlations $(r = 0.076, and - 0.196)$ respectively**.** Interestingly $(\mathrm{r} = -0.469)$ means that as students' total credit hours increases, students tend to perform poorly or students perform well as total credit hours decreases.

Table 1: Correlations between the studying variables

	CGPA	Age	WASSCE/Other Score	Total Credit HRS
CGPA	1.000			
Age	0.076	1.000		
WASSCE/Other Score	-0.196	-0.035	1.000	
Total Credit HRS	-0.469	0.013	0.062	1.000

Further, as part of the regression process, collinearity diagnostics on the variables were examined to ensure any problem of multicollinearity that was not noticeable in the correlation matrix is not missed. The output is displayed in Table 2 below. Two key statistics; tolerance level (TL) and variance inflation factor (VIF) are considered in this respect. The VIF for the predictor variables are 0.999, 0.995, and 0.996, and the corresponding TL values are 1.001, 1.005, and 1.004. These values satisfy the cut-off criteria values of not less than 0.1 for TL, and not more than 10 for VIF (Pallant, 2005), as clarified earlier in section 2.2.

Table 2: Coefficients[a]

Model	Unstandardised Coefficients		Std. Coffs			95% C. I for B		Correlations			Collinearity Statistics	
	B	Std. Error	Beta	T	Sig.	Lower bound	Upper bound	Zero Order	Partial	Part	Tolerance	VIF
1Constant	6.194	0.301		20.58	0.000	5.603	6.785					
(X_1)	0.010	0.005	0.076	2.16	0.031	0.001	0.019	0.076	0.087	0.076	0.999	1.001
(X_2)	-0.019	0.004	-0.165	-0.47	0.000	-0.027	-0.011	-0.196	-0.187	-1E-1	0.995	1.005
(X_3)	-0.074	0.006	-0.460	-13.08	0.000	-0.085	-0.063	-0.469	-0.469	-4E-1	0.996	1.004

Outliers, Normality, Linearity, Homoscedasticity, and Independence of Residuals

As one of the assumption requirements, normal probability plot of the regression standardised residuals, and residuals scatter plots were examined. The results of Figure 1 indicate that the normal probability plot displays points threading a reasonably straight line moving from bottom left to top right. This indicates that no major violation of normality assumption is breached. Further, the distribution of the standardised residuals as shown in Figure 2 does not reveal any curvilinear pattern or points distributed highly to one side than the other. Rather, the residual scores tend to concentrate at the centre which concurs with theoretical consideration. This suggests that residuals of the variables are independent. Also, case wise diagnostics that assess the merit of standardised residual as being influential on the results of our model or not, revealed that apart from the observation 3.330 which appears to be an outlier, the remaining three values; $-3.042, -3.135, -3.252,$ all appear not to be influential. The observation 3.330 constitute 1 out of the 610 cases $(0.1639\%, i.e. not\ more\ than\ 1\%)$ was identified as an outlier. These cases had maximum Cooke's Distance, $D_i = 0.042 < 1$. Judging from the cut-off values discussed in section 2 earlier, it implies that the outlier could cause no major problems

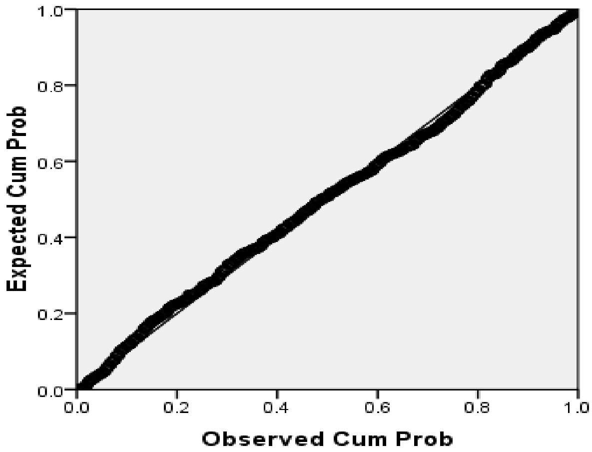

Figure 1: Normal P-P Plot of Regression Standardised Residuals

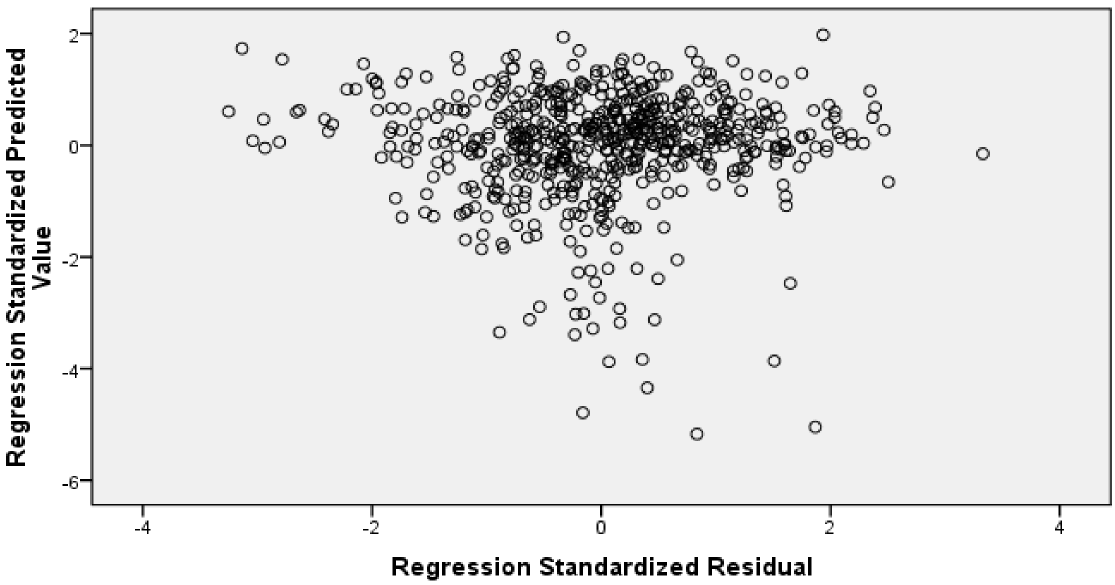

Figure 2: Distribution of the Standardised Residuals (Scatter Plot)

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

Evaluating the Model

Having examined the model with respect to the underlying assumptions, the next step is to evaluate the model with regard to the extent to which it explains the variability in the original data. This is achieved by considering the R^2 statistic. The value R^2 statistic obtained in this research is 0.254 which translates to 25.4%. This indicates that our model (which includes students' age at the time of admission, WASSCE or other examination score, and total credit hours taken during the first year) accounts for 25.4% of variance in cumulative grade point average. This is quite poor. This means that much of the variation in students' CGPA is unexplained by students' age at the time of admission, WASSCE or other examination scores, and credit hours taken during the first year.

Table 3: Model Summary[b]

Model	R	R- Square	Adjusted R Square	Std. Error of the Estimate
1	0.504[a]	0.254	0.250	0.63330

Predictors: (constant), total credit hours during the first year, students' grade point average, and students' age at the time of enrollment, WASSCE or other score.

Dependent variable: students' cumulative grade point average (CGPA).

Further, to measure the statistical significance of our results, it is imperative to consider the results of the analysis of variance, ANOVA. This examines the proposition that multiple R in the population equals 0. The output of this test contained in Table 4 indicates that the test is significant at 5%, (that is at the 0.05 level of significance).

Table 5: ANOVA[b]

Model	Sum of squares	Df	Mean square	F	Sig.
Regression	82.657	3	27.552	6.696	0.000a
Residual	243.053	606	0.401		
Total	325.708	609			

Predictors: (constant), total credit hours during the first year, students' grade point average, and students' age at the time of enrollment, WASSCE or other score.

Dependent variable: students' cumulative grade point average

Evaluating Each of the Independent Variables

This section examines how each of the predictor variables; age of the student at the time of enrollment, WASSCE or other examination score, and total credit hours taken during the first year, considered in our regression model, contributes to the prediction of the dependent variable (cumulative grade point average). Table 2 shows an important column (standardised coefficients) which ensures that all the values of the different variables being compared have the same scale. Hence we observe that the largest beta coefficient (neglecting any negative sign) is -0.460 which is for total credit hours taken during the first year. This indicates that total credit hours taken during the first year, comparatively contributes substantially to explaining students' cumulative grade point average, when the variance accounted for by all other

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

predictors in the model is controlled. Similarly, the beta coefficients for WASSCE or other score, age of the student at the time of enrollment in descending order of contributions are -0.165, and 0.076 respectively. The corresponding statistical significances made by these predictor variables in respect of unique contribution to the equation are 0.031, 0.000, and 0.000. This is shown in the column headed Sig. in Table 2. Further, it indicates that age of the student at the time of enrollment, WASSCE or other examination score, and total credit hours taken during the first year make a significant contribution to the prediction of students' cumulative grade point average.

CONCLUSION

Summarising the above analyses, we conclude that for Takoradi Polytechnic to succeed in identifying pre-tertiary school factors that have a bearing on the performance of evening school students, it must go beyond the traditional admission pre-requisites such as; age of the student at the time of enrollment, WASSCE or other examinations score, and total credit hours taken during first year. Also, total credit hours (course load) for the evening students during the first year must not be overloaded; as this tends to affect students' performance adversely. One of the findings of this research about total credit hours, disagrees with the results of Szatran (2001), but concurs with the result of Sansgiry et al (2006). Whiles

Szatran revealed that students who register for more credits hours tend to earn higher GPAs, Sansgiry et al concluded in their research about pharmacy students that, time management was not important variable that was related to students' academic achievement. In this research, data were collected from only first year Evening School. In given directions for future research, we suggest that the entire results of the 3-year HND programme are used. This will enable us obtain information about the pre-tertiary factors in terms of the overall Evening School. In addition, this study was limited to Takoradi Polytechnic. We therefore recommend that in future, this research is replicated with more variables in other polytechnics that run Evening School so that the external validity of the study results can be firmly established.

REFERENCES

Gordor, B.K. & Howard, N. K. (2000). *Elements of Statistical Analysis.* The City Printings, Accra, Ghana.

Hair, J. F. et al. (2006). *Multivariate Data Analysis*. 6th ed. Upper Saddle River, New Jersey, USA: Pearson Prentice Hall.

Johnson, R. A. & Wichern, D. W. (1992). *Applied Multivariate Statistical Analysis*. 3rd ed. New Jersey, USA: Prentice-Hall, Inc.

Montgomery, D.C.(1996). *Design and Analysis of Experiments*. 4th ed. John Wiley & Sons, Inc. USA.

Ministry of Education (2014). Report of the technical committee on conversion of the Polytechnics in Ghana to Technical Universities. hptt://www.moe.gov.gh/assets/media/docs/conversionofPolytechnicsToTUs-FinalReport.pdf. date: 05-05-16; time: 11:20 pm

Owusu-Agyeman, Y. (2006). Government Access- Policies on Polytechnic Education in Ghana: The relevance for Cape- Coast Polytechnic

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

http://essay.utwene.nl/590/1/scriptie-owusu.pdf date: 05-05-16; time: 10:20 pm

Pallant, J. (2005). SPSS survival manual: A step by step guide to data analysis using SPSS version 12. 2nd ed. Maidenhead, Berkshire: Open University Press.

Sansgiry,S.S, Bhole, M. & Sail, K. (2006). Factors Affecting the Academic Performance Among Pharmacy Students. American Journal of Pharmaceutical Education; 70(5):104

Szatran, F.R. (2001). The Effect of Academic Load on Success for New College Student: Is Lighter Better? Research in higher education; 42(1):27-50

Tabachnick, B.G. & Fidell, L.S. (2001). Using Multivariate Statistics. 4th ed. New York: Allyn & Bacon.

ISSN 2055-0154(Print), ISSN 2055-0162(Online)