Current Science International
ISSN 2077-4435

Volume : 05 | Issue : 01 | Jan.-Mar. 2016
Pages: 75-82

Statistical Study for the Quality of the Higher Education

[1]Saleh, Saida and Helmy, Nahed

Faculty of Commerce, Al-Azhar University, Cairo Girls' Branch, Egypt

Received: 17 January 2016 / Accepted: 20 February 2016 / Publication date: 10 March 2016

ABSTRACT

This paper is concerned with evaluating the quality of the higher education. Multiple linear regression and binary logistic regression are used to analyze the relationship between single response variable with two or more controlled variables. In this paper, stepwise multiple regression will be used because this method is a combination of forward selection and backward elimination method. The main objective in this paper is to select the suitable controlled variables in the quality of the higher education. Data that has been used in this research were taken from the faculties of Al-Azhar University for girls. To satisfy this purpose a questionnaire is designed and given to a random sample of the staff and their assistants in the faculties of Al – Azhar University for girls. The data are analyzed by using SPSS 16.

Key words: Quality of the higher education, multiple linear regressions, stepwise multiple regressions, binary logistic regression

1-Introduction

Statistical analysis is widely used in all aspects such as in science, medicine, and also social science. There are many methods in statistics and one of them is regression. There are six types of linear regression analyses which are simple linear regression, multiple linear regression, logistic regression, ordinal regression, multinomial regression and discriminant analysis. Multiple linear regression was selected to build a model to evaluate the quality of the higher education, which is stepwise regression is used in this paper. The main objective in this paper is to select the suitable controlled variables to evaluate the quality of the higher education.

Heyneman, (2004) pointed out that education is linked with economic productivity and growth in personal income. He argued that the impact of education is derived primarily from its quality, but there are multiple indicators of educational quality which do not necessarily operate in uniform fashion. He reviewed what we know about educational quality from the evidence of the last two decades. He addressed some of the current debates surrounding investment in educational quality and introduced several issues which will drive these debates in the future.

Zhong, (2011) examined the relationship between returns to higher education and college quality in China which vary significantly depending on school quality. He found that this relationship between earnings and school quality is stronger for cohorts of workers that have entered the workforce more recently.

He also found that the earning gap between graduates from lower-quality colleges and those from vocational/technical schools decreases over time.

Alfahad, (2012) investigated the usefulness and efficiency of information technology in higher education in the Kingdom of Saudi Arabia. The study was conducted in the College of Education, King Saud University. The survey was distributed among 161 female college students selected randomly from a pool of 400 female students who were attending different courses in different areas. Results indicated that 61.5% of the participants used the electronic device in their course activities and 65.8% used information technology for blogging.

Zaim and Camgoz-Akdag, (2012) introduced a model of student with higher education experience based on the identification of the variable determinants of student perceived quality. They used 41 items instrument of service quality which is applicable in education industry taking the five dimensions for quality as basis. Exploratory and confirmatory factor analyses empirically verified and validated the underlying dimensions of perception of student satisfaction, organizational performance. Structural equation modelling was used to estimate the models and compare coefficients and latent means to test the model drawing on a sample of 1752 students from private universities.

This paper is divided into five Sections. The first Section contains the introduction. The second Section is devoted to the theoretical assumptions used in the study. The third Section contains the methodologies. An

Corresponding Author: Helmy, N., Faculty of Commerce, AL-Azhar University, Cairo Girls' Branch.
E-mail: nahedhelmy19@yahoo.com

application is presented in the fourth Section. The last Section is about the main results and discussion of the present paper.

2- Theoretical Assumptions Used in the Study

The present paper is concerned with statistical model of the quality of the higher education using both of multiple linear regression and binary logistic regression models. To satisfy this purpose a questionnaire is designed. This questionnaire is given to the staff and their assistants in the faculties of Al-Azhar University for girls. The staff's questionnaire is divided into five sections which contains 65 items which constitute variables under study. The first section is about the major management practices which contains 13 items. The second section is about the quality assurance system which contains 8 items. The third section is about the assurance of the quality of inputs which contains 17 items. The fourth section is about the assurance of the quality of operations which contains 18 items. The fifth section is about the assurance of quality of outputs which contains 9 items. A copy of the questionnaire is shown in the Appendix.

2-1. Nature and Sources of Data:

This study is concerned with the staff and their assistants in the faculties of Al-Azhar University for girls:the faculty of Commerce, the faculty of Medicine, the faculty of Dentistry, the faculty of Arabic and Islamic Studies, the faculty of Humanities, the faculty of Science and the faculty of Pharmacy. The number of valid questionnaires is 144.

2-2. Statistical and Analytical Procedure:

The third, the fourteenth, the forty three and the fifty seven variables are selected and replaced by its mean to constitute the dependent variable. This dependent variable is denoted by Ywhich takes value 0 or 1. The value 0 corresponds to (not satisfy quality of higher education) and the value 1 corresponds to (satisfy quality of higher education).The items in each section are replaced by its average. The averages of the five sections are constitute the independent variables which are continuous. The range of each variable is greater than 0 and less than ∞.Some questions in the questionnaire have no answers. So there are some missing values. It is suggested to calculate the mean of each variable and these missing values are replaced by this mean. The first variable represents the major management practices which is denoted by (X_1), the second variable represents the quality assurance system which is denoted by(X_2), the third variable represents the assurance of the quality of inputs which is denoted by(X_3), the fourth variable represents the assurance of the quality of the operations which is denoted by(X_4) and the fifth variable represents the assurance of quality of outputs which is denoted by (X_5).Descriptive statistics for these variables to calculate correlation coefficient, correlation matrix, multiple regression, estimating the regression through the stepwise method are obtained.

3. Methodologies

Variable selection in linear regression is a problem of great practical importance. There are various methods for selection and various selection criteria. The purpose of this paper is to apply an effective method to estimate or predict at a lower cost by reducing the number of variables on which data are to be collected, to predict more accurately by eliminating uninformative variable, to describe multivariate data sets parsimoniously and to estimate regression coefficients with smaller standard errors .

3.1. Multiple linear regression:

The multiple linear regression model represents a relationship between a dependent variable and number of independent variables. It can be defined as follows:

$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_j x_{ji} + \varepsilon_i,\ i = 1, 2 \dots n, j = 1,2, \dots, p$ ……..(1)

where

Y_i: represents the dependent variable for the $i-th$ observation,

x_{ji}: represent the independent variables for the $i-th$ observation,

p:is the number of independent variables,

n: represents the number of observations and

$\beta_0, \beta_1, \beta_2, \dots, \beta_p$: are regression coefficients indicating the relative effect of a particular explanatory variable on the outcome.

These parameters are unknown which to be estimated.

These parameters indicate the change in the mean of the probability distribution of the dependent variable Y per unit change in x_{ji}. ε_i is a vector of error with $E(\mathcal{E}) = 0$ and $v\ (\mathcal{E}) = I\sigma^2$(Draper and Smith 1981). This method can be classified as enumerative, sequential such as forward selection, backward elimination, and

stepwise method (Xu and Zhang (2001)).These methods are described as follows:

3.2. Forward selection:

The forward selection method adds variables to the model one at a time. The first variable included in the model is the one which has the highest correlation with the dependent variable Y . The variable that enters the model as the second variable is one which has the second highest correlation with Y after Y has been adjusted for the effect of the first variable. This process terminated when the last variable entering the model has insignificant regression coefficient or all the variables are included in the model.

3.3. Backward elimination:

Backward elimination begins with all the variables of the model and successively eliminates one at a time. The first variable deleted is the one with the smallest contribution to the reduction of predictive error sum of squares. Assuming that there are more variables that are insignificant, the process operates by eliminating the next most insignificant variable. The process is terminated when all the variables are significant or all but one variable has been deleted.

3.4. Stepwise method:

In stepwise procedure a variable that entered the model in the earlier stages of selection may be deleted at the later stages. The calculations made for inclusion and elimination of variables are the same as forward selection and backward procedures. That is, the stepwise method is essentially a forward selection procedure, but at each stage the possibility of deleting a variable, as in backward elimination is considered. The number of variables retained in the model is based on the levels of significance assumed for inclusion and exclusion of variables from the model.

The logistic regression model is a generalized linear model which consists of random component, systematic component and link function. In the random component the response variable is binary. $Y_i = 1\ or\ 0$ (an event occurs or doesn't). We are interesting in probability that $Y_i = 1, \pi(x_i)$. The distribution of Y_i is binomial. In the systematic component the linear predictor is

$\beta_0 + \beta_1 x_{1i} + \cdots + \beta_j x_{ji}\ , i = 1, \dots, n\ , j = 1, \dots, p$

The explanatory or predictor variables may be quantitative (continuous), qualitative (discrete), or both (mixed).

In the link function the log of the odds that an event occurs, otherwise known as the $logit$:

$$logit(\pi) = \log\left(\frac{\pi}{1-\pi}\right)$$

Putting this all together the logistic regression model is:

$$logit(\pi(x_i)) = \log\left(\frac{\pi(x_i)}{1-\pi(x_i)}\right) = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_j x_{ji}$$

The Wald test is obtained by comparing the maximum likelihood estimate of the slope parameter, to an estimate of its standard error. The Wald test for the logistic regression model is:

$$W = \frac{\hat{\beta}_1}{\widehat{SE}(\hat{\beta}_1)}$$

And the two tailed $p-$value is:

$W = (|Z| > W)$, where Z denotes a random variable following the standard normal distribution. Jennings (1986) has also looked at the adequacy of inferences in logistic regression based on Wald statistics.In logistic regression there are several possible measures of difference between the observed and fitted values. The fitted values in logistic regression are calculated for each covariate pattern and depend on the estimated probability for that covariate pattern.

4. Application of the Study

The questionnaire in the present study is given to a simple random sample $(n = 400)$ from the staff and their assistants in the faculties of Al-Azhar University for girls: the faculty of Commerce, the faculty of Medicine, the faculty of Dentistry, the faculty of Arabic and Islamic Studies, the faculty of Humanities, the faculty of Science and the faculty of Pharmacy. The number of valid questionnaires are 144.The results are shown in Table (4.1) as follows:

Table 4.1: Case Processing Summary

	N	%
Cases valid	144	100
Excluded[a]	0	0
Total	144	100

a. Listwise deletion based on all variables in the procedure.

Cronbach's α coefficient is used to test the reliability of each section in the questionnaire and the entire questionnaire. If the value of Cronbach's α coefficient is greater than 0.5 then the reliability is satisfied. The square root for the Cronbach's α coefficient is used to test the validity of each section in the questionnaire and the entire questionnaire. If the value of the validity is greater than 0.5, then the validity is satisfied. The results are shown in table (4.2).

Table 4.2: Reliability and Validity Statistics.

N	Cronbach's Alpha	Validity
65 (the entire questionnaire)	0.958	0.98
13 (the first section)	0.908	0.95
8 (the second section)	0.807	0.89
17 (the third section)	0.865	0.93
18 (the fourth section)	0.895	0.95
9 (the fifth section)	0.863	0.93

From Table (4.2) it is noticed that all data were analyzed using statistical package for the social sciences SPSS 16. It is noticed that the value of Cronbach's α (Cortina 1993) coefficient is 0.958 for all sections in the questionnaire. Also the value of Cronbach's α coefficient is greater than 0.5 for each section in the questionnaire. So that the reliability is satisfied for the data under study. The value of the measure of validity for all sections is 0.98 in the questionnaire. Also the value of the measure of the validity is greater than 0.5 for each section in the questionnaire. So that the validity is satisfied for the data under study.

The mean for the items of each section is calculated. So the spearman correlation coefficient (Hauke and Kossowski, 2011) is calculated between each section and its items. This measure is used to test the consistency of the questionnaire and the relationship between variables under study.

Likert scale measure is used to give the codes for the answers (responses) of the questionnaire as follows:

Code	5	4	3	2	1
Response	Agree very much	Agree	Indifferent	Not agree	Not agree absolutely

Multiple linear Regression

Table 4.3: The Model Summary

model	R	R square	Adjusted R square	Std. error of the estimate
1	0.803[a]	0.644	0.631	0.531

a. Predictors: (Constant), mean5, mean2, mean3, mean4, mean1

From Table (4.3) since $R^2 = 0.644$ this means that 64% of variability in the dependent variable refers to the effect of the independent variables, adjusted $R^2 = 0.631$, estimate standard error $= 0.531$ the model is significant.

Table 4.4: ANOVA[b.]

Model	Sum of Squares	$d.f.$	Mean Square	F	Sig.
1 Regression	70.293	5	14.059	49.951	0[a]
Residual	38.840	138	0.281		
Total	109.133	143			

a. Predictors: (Constant), mean5, mean2, mean3, mean4, mean1
b. Dependent Variable: mean

Table (4.4) presents the $ANOVA$ table since $p - value = 0$ which is less than the level of significance $\alpha = 0.05$ then the model is significant.

Table 4.5: The Descriptive Statistics

	Mean	Std. Deviation	N
Y (Mean)	2.981	0.874	144
X_1 (Mean 1)	2.729	0.817	144
X_2 (Mean 2)	2.483	0.779	144
X_3(Mean 3)	2.389	0.604	144
X_4 (Mean 4)	2.584	0.674	144
X_5 (Mean 5)	1.869	0.718	144

From Table (4.5) the descriptive statistics for the dependent variable and the independent variables are obtained. The first independent variable has the largest mean and the largest standard deviation.

Table 4.6: The Correlation Matrix.

	Constant	X_1	X_2	X_3	X_4	X_5
Step 1 constant	1.000	0.716	0.581	0.559	0.672	0.652
X_1	0.716	1.000	0.690	0.647	0.631	0.536
X_2	0.581	0.690	1.000	0.564	0.595	0.448
X_3	0.559	0.647	0.564	1.000	0.617	0.491
X_4	0.672	0.631	0.595	0.617	1.000	0.599
X_5	0.652	0.536	0.448	0.491	0.599	1.000

From Table (4.6) the correlation matrix for all independent variables, where represented positive relations between the constant and all variables. The largest correlation coefficient is 0.69 between the first and the second variables which indicates to strong relationship between them.

Table 4.7: The Coefficients[a].

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
1 (constant)	0.303	0.199		1.526	0.129	−0.090	0.697
X_1 (mean1)	0.407	0.087	0.381	4.682	0.000	0.235	0.579
X_2 (mean2)	0.061	0.083	0.054	0.733	0.465	−0.103	0.224
X_3(mean3)	0.001	0.104	0.001	0.011	0.991	0.205	0.207
X_4 (mean4)	0.295	0.099	0.227	0.227	0.003	0.099	0.490
X_5 (mean5)	0.349	0.080	0.287	0.287	0.000	0.190	0.508

a. Dependent Variable: mean

From Table (4.7) the linear regression model is given by:

$Y = 0.303 + 0.407\,X_1 + 0.061X_2 + 0.001X_3 + 0.295X_4 + 0.349X_5$

Note that the constant is positive and there are positive relationship between Y (dependent variable) and the independent variables in this method.

Logistic Regression:

Table 4.8: Case Processing Summary.

Unweighted Cases[a]	N	Percent
Selected Cases Included in Analysis	144	100
Missing Cases	0	0
Total	144	100.0
Unselected Cases	0	0.0
Total	144	100.0

a. *If weight is in effect, see classification table for the total number of cases.*

Dependent Variable Encoding:

Original Value	Internal Value
0	0
1	1

Block 0: beginning block:

Table 4.9: Classification Table[a,b].

Observed		Predicted		
		Y		Percentage Correct
		0	1	
Step 0 Y	0	113	0	100
	1	31	0	0
Overall Percentage				78.5

a. Constant is included in the model
b. The cut value is 0.5

From Table (4.9) the percent of cases for which the dependent variables was correctly predicted given the model is 78.5 = 113/144.

Table 4.10: Variables in the model

Step 0 constant	B	$S.E.$	Wald	$d.f.$	Sig.	$Exp(B)$
	−1.293	0.203	40.695	1	0.000	0.274

From Table (4.10) the Wald chi-square test tests the null hypothesis that the constant equals 0. This hypothesis is rejected because the p-value (listed in the column called "Sig.") is smaller than the level of significance α = 0.05 (or 0.01). Hence, we conclude that the constant is not 0. Exp (B) is the exponentiation of the B coefficient, which is an odds ratio. This value is given by default because odds ratios can be easier to interpret than the coefficient, which is in log-odds units. This is the odds: 31/113 = 0.274.

Table 4.11: Variables in the Model

	Score	$d.f.$	Sig.
Step 0 variables X_1 (mean1)	43.663	1	0.000
X_2 (Mean2)	31.283	1	0.000
X_3 (Mean3)	25.616	1	0.000
X_4 (Mean4)	28.460	1	0.000
X_5 (Mean5)	35.389	1	0.000
Overall Statistics	52.985	5	0.000

Block 1: method= forward stepwise (Wald):

From Table (4.12) the value given in the Sig. column is the probability of obtaining the chi-square statistic given that the null hypothesis is true if there no effect of the independent variables, taken together, on the dependent variable. Since the $p - value = 0$ is less than the level of significance $\alpha = 0.05\ or\ 0.01$ the model is statistically significant.

Table 4.12: Omnibus Tests of Model Coefficients

	Chi-square	df	Sig.
Step 1 step	55.106	1	0
Block	55.106	1	0
Model	55.106	1	0
Step 2 step	23.439	1	0
Block	78.545	2	0
Model	78.545	2	0

Table 4.13: Model Summary

Step	−2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	94.903[a]	0.318	0.491
2	71.464[b]	0.420	0.650

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.
b. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.

Table 4.14: Classification Table[a]**.**

Predicted			Observed
Percentage Correct	Y		
	1	0	
93.8	7	106	Step 1 Y 0
61.3	19	12	1
86.8			Overall Percentage
96.5	4	109	Step 2 Y 0
71	22	9	1
91			Overall Percentage

a. The cut value is 0.5

From Table (4.14) at step 1 the predicted values of the dependent variable is based on the full logistic regression model. This table shows how many cases are correctly predicted (106 cases are observed to be 0 and are correctly predicted to be 0; 19 cases are observed to be 1 and are correctly predicted to be 1), and how many cases are not correctly predicted (7 cases are observed to be 0 but are predicted to be 1; 12 cases are observed to be 1 but are predicted to be 0).At step 2 (109 cases are observed to be 0 and are correctly predicted to be 0; 22 cases are observed to be 1 and are correctly predicted to be 1), and how many cases are not correctly predicted (4 cases are observed to be 0 but are predicted to be 1; 9 cases are observed to be 1 but are predicted to be 0).The overall percent of cases that are correctly predicted by the model has increased from 78.5 for the null model to 86.8 at step1 and has increased from 78.5 to 91 at step 2.

Table 4.15: Variables in the Model

	B	$S.E.$	Wald	$d.f.$	Sig.	$Exp(B)$
Step 1[a] mean1	*−2.659*	0.486	29.902	1	0	0.07
Constant	5.003	1.081	21.433	1	0	148.817
Step 2[b] mean1	*−2.085*	0.529	15.668	1	0	0.124
Mean5	*−3.403*	0.9	14.296	1	0	0.033
Constant	8.720	1.725	25.563	1	0	6.123E3

a. Variable(s) entered on step 1: X_{1_1}(mean1).
b. Variable(s) entered on step 2: X_5(mean5).

From Table (4.15) B are the values for the logistic regression model for predicting the dependent variable from the independent variables. It is noticed that the two variables which affect the quality of the higher education are the major of management practices and the assurance of the quality of outputs. The logistic regression model is:
Y = 8.72-2.085 (the major management practices)-3.403 (assurance of the quality of outputs).

Table 4.16: Correlation Matrix.

	Constant	mean1	mean5
Step 1 constant	1	*−0.971*	
Mean1	*−0.971*	1	
Step 2 constant	1	*−0.695*	*−0.708*
Mean1	*−0.695*	1	0.014
Mean5	*−0.708*	0.014	1

Results and Discussion

Comparison between the multiple linear regression model and the logistic regression model is made. Both models are significant and the coefficient of determination in the linear regression model equals 64.4% and in the logistic model equals 65%. The classification predicted equals91% in the logistic model. The logistic regression model is highly significant than the linear regression model. When the logistic model is applied it is found that the most two independent variables affect the quality of the higher education in the university are the major management practices (mean1) and the assurance of quality of outputs (mean5).

Appendix:

The questionnaire which is given to the staff and their assistance in the faculties of Al Alzhar University for girls is:

The first section: the major management practices.

1.	The university administration evaluates the performance of each faculty member from students' perspective
2.	The university administration evaluates the performance of each faculty member from the perspective of direct boss
3.	Al-Azhar University focused on the quality of education
4.	The university adopts that the education meets the needs of students
5.	Al-Azhar University suggests educational programs that meet the requirements of the labor market
6.	University confirms the quality education should be shared by all employees
7.	University conducted a comprehensive strategic planning process for all colleges
8.	The university administration requests from each college to provide a comprehensive plan for the next period
9.	The university administration practice effective control of all faculties of the university
10.	University encourages creativity through programs that support the development of creators
11.	University encourages colleges to participate in specialized competitions
12.	The university supports scientific research
13.	The university adopts the use of the computer in the implementation of educational activities

The second section: the quality assurance system.

14.	An organizational unit in the university is working to ensure the quality of education at the university
15.	University establishes a mechanism for the conduct of all activities at the university
16.	University defines the duties , responsibilities and roles clearly
17.	University held specialized courses to improve the performance of the university teaching staff
18.	University provides an information network linking with other universities
19.	University provides an information network linking with foreign universities
20.	Any candidate for employment at the university is subject to a personal interview
21.	Conduct between applicants for employment on objective grounds based on the efficiency in the field of specialization

The third section: the insurance ofthe quality of inputs

22.	University uses special forms to evaluate the performance of faculty members
23.	University provides inputs (materials / tools / devices) necessary for the educational process
24.	University provides inputs (materials / tools / devices) necessary to carry out research
25.	University provides adequate financial support for the educational process
26.	University provides adequate financial support to conduct research
27.	The number of computers in university laboratories is enough for students
28.	The number of computers in university laboratories is enough for professors
29.	The university computer labs meet the needs effectively
30.	Network online is sufficient to serve the university faculty members
31.	Network online is sufficient to serve the students
32.	There is a sufficient number of university libraries
33.	The number of sections in the university libraries meet the needs of beneficiaries
34.	Specialized periodicals are available in university libraries
35.	New periodicals are available in university libraries
36.	Faculty members know how to use a computer

37.	The number of teaching staff fits the number of students
38.	The number of employees is higher than the needs of the university

The fourth section: the insurance of the quality of operations.

39.	Evaluating the performance of the teaching staff periodically
40.	Computer is used in implementing the educational activities
41.	The participation of faculty staff in specialized scientific conferences
42.	Cooperation between faculty members to conduct joint research
43.	Development of courses and curricula
44.	Development of study plans
45.	Development of a list of approved books as a source for each section
46.	Each section has a file includes complete data about its plan of study and models for exams
47.	Encouraging the teaching staff for research
48.	The conferences are held in the field of specialization
49.	There is a full and detailed documentation of the procedures required for all aspects of the educational process
50.	The annual increase on the salary is consistent with the adequate level of cost of living
51.	The prices of health insurance are suitable for workers
52.	The health insurance conditions suited workers
53.	The value of the management bonus match with the administrative work
54.	University gives housing expenses for workers living outside their places of residence
55.	University gives transport expenses for workers living away from the university.
56.	The level of wages and the standard cadre level is fair.

The fifth section: the insurance of quality of outputs.

57.	Taking the views of students in teaching staff periodically
58.	Taking the views of students in the physical facilities necessary for the educational process
59.	Taking the views of students in the supplementary services at the university (cafeteria, parking lots, sports activities ...)
60.	Analyze the feasibility of currently subjects proposed according to the quality of graduate level
61.	Development of materials and programs of study according to the quality of graduates periodically
62.	Benchmarking conducted educational programs with educational programs for educational institutions leading in similar areas
63.	The training programs are designed for the students to serve them in the labor market
64.	Specialists from institutions where graduates are working be hosted to evaluate the quality of university graduates
65.	University seeks to insure the relationship of the students with the institutions before going out into the labor market

References

Alfahad, F.N., 2012. Effectiveness of using information technology in higher education in Saudi Arabia. Procedia - Social and Behavioral Sciences, 46: 1268 – 1278.

Cortina, J.M., 1993. What is coefficient Alpha? An Examination of Theory and Applications. Journal of Applied Psychology, 78(1): 98 – 104.

Draper, N.R. and H. Smith, 1981. Applied Regression Analysis, 2nd Edition, John and wiley, New York.

Hauke, J. and T. Kossowski, 2011. 'Comparison of Values of Pearson's and Spearman's correlation Coefficients on the Sets of Data' Quaestiones Geographicae, 30(2): 1 – 7.

Heyneman, S.P., 2004. International Education Quality. Economics of Education Review, 23: 441–452.

Jennings, D.E., 1986. Trudging Inference Adequacy in Logistic Regression.' *American Statistical Association*, 81: 471-476.

Kobayashi, K., K.S. Pillai, Y. Sakuratani, M. Suzuki, W. Jie, 2008. Do we need to examine the quantitative data obtained from toxicity studies for both normality and homogeneity of variance?. Journal of Environmental Biology, 29(1): 47 – 52.

Xu, L. and H. Zhang, 2001. Comparison of different methods for Variable Selection.' Analytica Chimica Acta, 446: 477-483.

Yang, L. and B. McCall, 2013. World education finance policies and higher education access: A statistical analysis of World Development Indicators for 86 countries. International Journal of Educational Development, 1503: 1 – 12.

Zaim, S. and H. Camgoz-Akdag, 2012. Education: a Comparative Structural Equation Modelling Study. Procedia - Social and Behavioral Sciences, 47: 874 – 880.

Zhong, H., 2011. Returns to Higher Education in China: What is the role of college quality?. China Economic Review, 22: 260–275.