

Multivariate methods for Investigating Academic Stumbling: A Case Study Faculty of Economics and Rural Development, University of Gezira State, Sudan (2017)

Adel Ali Ahmed Mohamed ⁽¹⁾

Mawda Salah Hassan Eltahir⁽²⁾

Abstract

The importance of factor analysis is to summarize many of the variables to a smaller number of variables known factors. The aim of the research to find out the academic stumble students of the Faculty of Economics and Rural Development, University of Gezira. A sample of 100 students was tested using A questionnaire form consisting of three axes applied to the area of the study. The analytical descriptive method was used and one of the methods of multiple variables (Factor Analysis) was used the Statistical Packages for Social Sciences (SPSS). The results indicate that the main weights affecting the first factor were weak motivation and desire to study and desire to work and work more than the study either affect the second factor is the disintegration of the family because of divorce or otherwise Accordingly, the third factor was affected by the weakness of the mental abilities and the difficulty in absorbing some subjects. The adoption of some faculty member's inappropriate methods in teaching led to a gap between the needs of the student and the content of the subject. Accordingly, the research came out with the recommendations of intensifying the process of guidance in the departments and taking care of academically stumbled students.

(1): Associate Professor, Faculty of Economics and Rural Development, Department of Applied Statistics and Demography.

(2): Lecturer, Faculty of Economics and Rural Development, Department of Applied Statistics and Demography.

1. Introduction:

University education in Sudan is great importance due to cultural factors, national factors, economic and political factors this is due to the large numbers of students who join universities to improve their social and living conditions, because education is the source of income.

There is no doubt that the academic stumbling in the request of the University of the Gezira is increasing in severity in the last years to the extent that the phenomenon is being talked about where the proportion of high and many researchers believe that this stumbling is due to personal reasons attributed to the student and social causes and the difficult political and economic conditions and others believe that the reason for this stumbling is due to purely administrative reasons practiced by the university right to demand in order to implement the regulations and laws .academic stumbling has become a concrete phenomenon Many students suffer from it and must find appropriate solutions. The researcher believes that information technology may contribute to solving a large part of the underlying causes of this problem especially with regard to administrative aspects, and this study was exhausted by this study, where the researcher developed a proposed experimental scenario for dealing with this phenomenon in the case of students who have failed studying at the university of Gezira.

2. Research Question:

- 1-Does the level of academic failure vary among students according to the type of students?
- 3-. Does the level of academic failure vary among students at different levels of study?

3. Research Objectives:

- 1-To estimate the level of academic failure vary among students according to the type of students?
- 2-To estimate the level of academic failure vary among students at different levels of study?

4. Importance of the Research:

It is a topic worthy of study and research, where academic stumbling has become suffers from large numbers of students, so it is considered standing on this phenomenon and its magnitude and knowledge of the factors affecting it is of great importance a concrete phenomenon It is clear in Sudanese universities.

Since academic stumbling is associated with the performance of university students, this study is gaining importance,

As it focuses on university youth as a target group

Society, and it has many responsibilities in the future of the country, so it is the treatment of the performance of students

Universities are always an urgent necessity.

This study may reveal the real reasons behind the students' academic stumbling, and therefore they are

Provide feedback to stakeholders, helping them to plan for appropriate solutions to address these causes to mitigate the phenomenon of stumbling.

5. Research Methodology:

The research mainly depends on descriptive analytical method and Factor Analysis.

5.1 Data Sources:-

Primary data is the main source of data collection. A questionnaire designed to solve the problem and include all the variables affecting (difficulties related to the academic side, Difficulties related to student personal circumstances, Difficulties related to economic and social factors).

6. Factor Analysis Model:

Multiple linear regression model:

$$x_1 = \lambda_{11}f_1 + \dots + \lambda_{1k}f_k + u_1 \quad (1)$$

$$x_2 = \lambda_{21}f_1 + \dots + \lambda_{2k}f_k + u_2 \quad (2)$$

$$x_p = \lambda_{p1}f_1 + \dots + \lambda_{pk}f_k + u_p \quad (3)$$

Where:

$x = (x_1, \dots, x_p)'$ are the observed variables (random)

$f = (f_1, \dots, f_k)'$ are the common factors (random)

$u = (u_1, \dots, u_p)'$ are called specific factors (random) λ_{ij} are called factor loadings (constants)

In short: $x = \Lambda f + u$,

where Λ is the $p \times k$ matrix containing the λ_{ij} 's.

Difference with multiple regression: common factors f_1, \dots, f_k are unobserved.

6.1. Assumptions of Factor Analysis Model

Factor analysis is designed for interval data, although it can also be used for ordinal data (e.g. scores assigned to Likert scales). The variables used in factor analysis should be linearly related to each other. This can be checked by looking at scatter plots of pairs of variables. Obviously the variables must also be at least moderately correlated to each other, otherwise the number of factors will be almost the same as the number of original variables, which means that carrying out a factor analysis would be pointless (Cornish, 2007).

$E(x) = 0$ (if this is not the case, simply subtract the mean vector)

$$E(f) = 0, \text{Cov}(f) = I \quad (4)$$

$$E(u) = 0, \text{Cov}(u_i, u_j) = 0 \text{ for } i \neq j \quad (5)$$

$$\text{Cov}(f, u) = 0 \quad (6)$$

Variance of x_i

Notation:

$$\text{Cov}(u) = \Psi = \text{diag}(\Psi_{11} \dots \Psi_{kk}) \quad (7)$$

$\text{Cov}(x) = \Sigma$

$$\sigma_{ii} = \text{var} \sum_{j=1}^k \lambda_{ij}^2 + \psi_{ii} \quad (8)$$

$\text{Var}(x_i)$ consists of two parts:

$$h_i^2 = \sum_{j=1}^k \lambda_{ij}^2 \quad (9)$$

called communality of x_i , represents variance of x_i that is shared with the other variables via the common factors. Ψ_{ii} , called the specific or unique variance, represents the variance of x_i that is not shared with the other variables.

Covariance matrix of x

$$\sigma_{ii} = \text{cov}(x_i, x_j) = \sum_{j=1}^k \lambda_{ie} \lambda_{je} \quad (10)$$

Hence, the factor models leads to

$$\Sigma = \Lambda \Lambda' + \Psi$$

6.2 Analysis of Co-variances and Correlations

Since factor analysis usually works with the variances and co-variances of the observed x variables, it is sometimes referred to as “the analysis of covariance structures”. Some hint of this is apparent in equation, where the absence of an intercept term suggests that the means of the observed variables are either zero or of no direct interest. Indeed, this is typically the case in factor analysis, where the task is to learn about inter-relationships among variables rather than model the levels of each variable. Moreover, it is generally not possible to estimate both the factor loadings and intercept terms (Joreskog and Sorbom, 1979, cited in (Cornish, 2007)). also Bollen, (1989),cited in (Cornish, 2007).. Consequently, all the x variables and the unobserved n are presumed to have zero means, constraining any intercept term in equation to zero. In addition, for the ordinal variables frequently encountered in surveys, the latent variable approach to generating a correlation matrix posits the variances of the latent variables to be 1, making the all co-variances between the latent variables interpretable as correlations,(Cornish, 2007).

6.3 Steps in Factor Analysis:

The factor analysis model can be written algebraically as follows. If you have p variables x_1, x_2, \dots, x_p measured on a sample of n subjects, then variable i can be written as a linear combination of m factors f_1, f_2, \dots, f_m where, as explained above $m < p$. Thus,

$$x_i = \lambda_{i1}f_1 + \lambda_{i2}f_2 + \dots + \lambda_{im}f_m + e_i \quad (11)$$

where the λ_{is} are the factor loadings (or scores) for variable i and e_i is the part of variable x_i that cannot be 'explained' by the factors.

There are three main steps in a factor analysis:

Step 1 Calculate Initial Factor Loadings

The factor loadings λ are parameters to be estimated that tap how the unobserved factors account for the observed variables: the larger the values of λ , the more a particular variable is said to “load” on the corresponding factor. Note that the factor loadings λ vary across survey items, but not across individuals. Put differently, items vary in the way they are explained by the underlying factors, but the relationships between underlying factors and observed responses is constant across individuals (hence the absence of an i subscript indexing λ). Note also that there

are fewer underlying factors than there are variables ($p < k$), consistent with the notion that like any statistical procedure, factor analysis is a device for “data reduction” taking a possibly rich though unwieldy set of survey responses and summarizing them with a simpler underlying structure. (Cornish, 2007).

Calculate initial factor loadings can be done in a number of different ways:

Principal Component Method: As the name suggests, this method uses the method used to carry out a principal components analysis. However, the factors obtained will not actually be the principal components (although the loadings for the k^{th} factor will be proportional to the coefficients of the k^{th} principal component) (Cornish, 2007).

(Child D, 1973) in Principal components method: found the linear combination that explains the maximum variance from the X's. This is the first factor. then find the next combination that explains the maximum proportion of the remaining variance and is orthogonal to the next factor, etc. (proceed until all variance is explained).

Let $(\lambda_i ; e_i)$ be the eigenvalue-eigenvector pairs of Σ , with

$$\lambda_1 \geq \lambda_2 \dots \lambda_p > 0$$

From the spectral theorem:

$$\Sigma = \lambda_1 e_1 e_1' + \lambda_2 e_2 e_2' + \dots + \lambda_p e_p e_p' \quad (12)$$

$$\text{Let } L = \sqrt{\lambda_1} e_1, \sqrt{\lambda_2} e_2, \dots, \sqrt{\lambda_p} e_p$$

Then:

$$\Sigma = \lambda_1 e_1 e_1' + \dots + \lambda_p e_p e_p' = LL' = LL' + 0$$

Thus L is given by $\sqrt{\lambda_i}$ times the coefficient of principles components, $\psi = 0$

Now, if $\lambda_{m+1}, \lambda_{m+2}, \dots, \lambda_p$ are small, then the first m principle components explain most Σ .

Thus , with $L_m = \sqrt{\lambda_1}e_1, \dots, \sqrt{\lambda_m}e_m$

$$\Sigma \approx L_M L_M'$$

With specific factors this becomes

$$\Sigma \approx L_M L_M' + \Psi \quad (3.43)$$

Where:

$$\Psi = \sigma_{ii} - \sum_{j=1}^m \lambda_{ij}^2 \quad (13)$$

As estimator for the factor loadings and specific variances, we take:

$$\tilde{L} = \tilde{L}_m = (\sqrt{\hat{\lambda}_1} \hat{e}_1, \dots, \sqrt{\hat{\lambda}_m} \hat{e}_m) \quad (3.45)$$

where $(\hat{\lambda}_i, \hat{e}_i)$ are the eigenvalue-eigenvector pairs of the sample covariance matrix S , and

Where:

$$\Psi = S_{ii} - \sum_{i=1}^m \lambda_{ij}^2$$

Ψ : diagonal matrix

In many cases, the correlation matrix R (which is also the covariance matrix of the standardized data) is used instead of S , to avoid problems related to measurements being in different scales (Thulin, 2011).

7. Data Analysis and Discussion

7.1 Introduction:

The study revealed the causes of academic stumbling for students of university of Gazira (Faculty of Economics and Rural Development).

The data on which this research was based is preliminary data and the analysis of the data was applied factor analysis of these data This was done using the statistical packages program for social sciences (SPSS) In order to reach the results of the targeted analysis.

Table:(1) Kaiser And Bartlett's Test

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.666
Bartlett's Test of Sphericity	Approx. Chi-Square	884.806
	Df	253
	Sig.	.000

source: prepared by the researcher of the outputs of the program (Spss).

Table (1) **shows** that the KOM is more than 0.5 which indicates an adequate sample size the Bartlett test is test values are equal 884.80 Significant level (0.000) This indicates That the test (significance) statistically Significant (0.01) Thus we conclude the original matrix relational is not the type of matrix unit therefore, the requirement to use the main components method has been met factor analysis.

7.2 Check the Correlation Matrix: -

In order to be valid for factor analysis they must produce at least some of the matrix relationships of 0.3 or higher. If we do not find in the matrix some of the relations of 0.3 or higher, we judge the matrix unfit for analysis factor analysis Thus, we ignore the use of the main components method, the researcher find that the matrix of correlative relations contained some relations reached 0.3 It is also possible to use the main components method in analysis.

Table:(2) Total Variance Explained

Component	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.355	24.194	24.194	2.570	14.277	14.277
2	1.961	10.893	35.087	2.299	12.773	27.050
3	1.857	10.315	45.402	1.994	11.080	38.129
4	1.471	8.170	53.572	1.864	10.358	48.487
5	1.307	7.263	60.835	1.758	9.766	58.252
6	.985	5.471	78.235			
7	.803	4.464	82.699			
8	.656	3.644	86.343			
9	.514	2.854	89.197			
10	.394	2.192	91.388			
11	.377	2.093	93.481			
12	.318	1.765	95.246			
13	.270	1.499	96.745			
14	.224	1.244	97.989			
15	.190	1.055	99.044			
16	.172	.956	100.000			

source: prepared by the researcher of the outputs of the program (Spss).

Table (2) presents the total variance of the variable in initial solution. The results of initial solution extracted 16 factors the same number as number of variables factored. The factors extracted accounted for more than 1.0 unites of variance. The initial solution extracted 5 factors from accounting for more than 1.0 unit of variance.

Factor 1: The factor has an Eigen value equals (2.570) since this is greater than 1.0 it explains more variance than a single variable in fact (2.570) times as much. The percent a variance explained $(2.570/16 \text{ unit of variance}) * 100 = (14.277\%)$.

Factor 2: The factor has an Eigen value equals (2.299) it also greater than 1.0 and there for explains more variance than a single variable.

The percent a variance explained $(2.299/16 \text{ units of variance}) * 100 = (12.773\%)$.

Factor 3: The factor has an Eigen value equals (1.994) it also greater than 1.0 and there for explains more variance than a single variable.

The percent a variance explained $(1.994/16 \text{ units of variance}) * 100 = (11.080\%)$.

Factor 4: The factor has an Eigen value equals (1.864) it also greater than 1.0 and there for explains more variance than a single variable.

The percent a variance explained $(1.864 /16 \text{ units of variance}) * 100 = (10.358 \%)$.

Factor 5: The factor has an Eigen value equals (1.758) it also greater than 1.0 and there for explains more variance than a single variable.

The percent a variance explained $(1.758 /16 \text{ units of variance}) * 100 = (9.766 \%)$.

The sum of the Eigen value associated with each factor (component) sum to 16.

$(4.355+1.961+1.857+1.471+1.307+0.985+$
 $0.0803+0.514+0.394+0.377+0.318+0.270+0.224+0.190+0.172)=16$

The cumulative percentage of variance explained by the first factors was (58.252%).

Communalities: The quantity of communalities of variables represents the percentage of variance explained by the factors derived from these variables.

The high this indicates that they explain the high proportion of the variance table no(3-3) Note that the factors explain the high proportion of variance where the lowest ratio is(.66)

University Connections and Confusion and shame in front of others.

Table (3) Communalities

	Initial	Extraction
Weak direction and guidance in the process section	1.000	.835
Different school system in secondary education	1.000	.688
Some faculty members have inappropriate teaching methods	1.000	.861
gap between the needs of students and the content of the course material	1.000	.819
-Lack of appropriate teaching aids in the educational process	1.000	.679
The classrooms are not suitable for lighting and ventilation	1.000	.744
Weak motivation and desire to study	1.000	.768
Desire to work and work more than study	1.000	.763
Weak self-confidence	1.000	.794
Confusion and shame in front of others	1.000	.663
Weakness in mental abilities	1.000	.781
Difficulty absorbing some subject	1.000	.773
University Connections	1.000	.660
There are differences between family members	1.000	.748
Disintegration of the family as a result of students or otherwise	1.000	.813
Weakness in the cultural level of the family	1.000	.758

-Marriage in school years takes responsibility	1.000	.835
--	-------	------

source: prepared by the researcher of the outputs of the program (Spss).

Explanation of Factors: We find that the method of analysis of the main components aims to summarize a set of variables in a smaller number of factors, but there is no guarantee that we always get factors that can be easily explained by their associations with variables, and a solution to forms uses the method of rotation axes, under a process called axial rotation. This method aims at creating new factors of the factors that have already been obtained, in which their associations with the original variables must be distributed in such a way as to facilitate their interpretation. This means that the method of rotating the axes seeks to collect variables similar in nature in a way that facilitates interpretation There are many methods to rotate the axes most important and most widely used varimax variation method, which aims to rotate the axes in such a way that makes the variance of the degrees of saturation of each factor as large as possible, that is to make the larger large and smaller specimens smaller than the design before the rotation, The variables are linked to the factors. by rotating the axes, the rotated component matrix is obtained, which shows the weights of the variable Loading. By examining the matrix of factors circulating in Table (3-4) we observe the following:-

A- The main factors on the first factor(Student) were weak motivation and desire to study and Desire to work and work more than study.

B- The main factors on the second factor(Family), there are difference between family members, disintegration of the family because of students or otherwise, weakness in the cultural level of the family, Marriage in school years takes responsibility

C. The main factors on the third factor (Behavior of student) Weak self-confidence, Confusion and shame in front of others.

D. The main factors on the four factor (Academic System) Different school system in secondary education, Difficulty absorbing some subjects.

E . The main factors on the fifth factor (Department) , Weak direction and guidance in the process section , Lack of appropriate teaching aids in the educational process .

Table (4) Factor one: -Student

Variable	Variable name	Loading
X ₁	Weak motivation and desire to study	.735
X ₂	Desire to work and work more than study	.812

source: prepared by the researcher of the outputs of the program (Spss).

$$P_{C1} = .735 X_1 + .812 X_2$$

Table (5) factor two: -Family

Variable	Variable name	Loading
X ₁	there are difference between family members	.708
X ₂	disintegration of the family as a result of students or otherwise	.774
X ₃	weakness in the cultural el of the family	.638
X ₄	Marriage in school years takes responsibility	.706

source: prepared by the researcher of the outputs of the program (Spss).

$$P_{C2} = .708 X_1 + .774 X_2 + .638 X_3 + .706 X_4$$

Table (6) factor three: -Behavior of student

Variable	Variable name	Loading
X ₁	Weak self-confidence	.812
X ₂	Confusion and shame in front of others	.743

source: prepared by the researcher of the outputs of the program (Spss).

$$P_{C3} = .812X_1 + .743X_2$$

Table (7) factor four: -Academic system

Variable	Variable name	Loading
X ₁	Different school system in secondary education	.730
X ₂	Difficulty absorbing some subjects	.801

source: prepared by the researcher of the outputs of the program (Spss).

$$P_{C4} = .730X_1 + .801 X_2$$

Table (8) factor five:-Department

Variable	Variable name	Loading
X ₁	Weak direction and guidance in the process section	.801
X ₂	Lack of appropriate teaching aids in the educational process	.669

source: prepared by the researcher of the outputs of the program (Spss).

8. Conclusion:

The aim of the research is to use statistical methods to reach results that contribute to the knowledge of the most important factors that affect In the academic stumble **The** three study variables were subjected to analysis using the main components method using the SPSS program The research found follows to measure the adequacy of the sample size used in the study, use the KMO scale of .666 , indicating the adequacy of the sample size.

- The Bartlett test was applied, which revealed that the correlation between the variables is significant, indicating the possibility of using the method of the main components in the analysis.

-Standard Kaiser has revealed the existence of seven factors that increase its roots distinct from the correct one the seven factors accounted for about 72.764 of the total variance So that the first factor of about 14.277, the second factor of about 12.773

third factor about 11.080, fourth factor about 10.358, fifth factor about 9.766, sixth factor about 7.551 and seven factor about 6.960.

-To explain these factors used style Varimax the biggest contrast method we find that the first factor saturation variables (Weak motivation and desire to study, Desire to work and work more than study).

- The most significant variables contributing to itself in the formation of the second factor Disintegration of the family because of students or otherwise.

- The most significant variables contributing to itself in the formation of the third factor Weak self-confidence.

- The most significant variables contributing to itself in the formation of the four-factor difficulty absorbing some subjects.

- The most significant variables contributing to itself in the formation of the five-factor weak direction and guidance in the process section.

9. Recommendations:

1-Increase the number of hours of academic supervision.

2- There necessary follow-up from the family within the university to see the status of the student.

3-Considering the appropriate teaching methods by the professors.

References:

1.Basiony study (2006) The role of Mansoura University in solving the academic stumbling problem of university students Seventh Conference twenty of the Arab organization responsible for admissions and registration in universities in Arab countries, University of Sharja folder two number 77.

- 2- Barth, p., Haycock, H.Jackson, K.Mora, p.Ruiz, s. Robinson & a.willins, Eds.(1999) *Dispelling the myth : High Poverty Schools Exceeding exectations*. Washington, D.C.: Report of the Education Trust in cooperation with the council of chief state school officers and partially funded by the U.S. development of education.
- 3- Bradley, a (2000). "The Gatekeeping Challenge" *Education Week*, Vol XIX, Number 18, January.
- 4- Hassan's study(2001) the reasons of the failure of students at the Faculty of Education Zagazig University. *College Journal* 285. Education in Zagazig, January, No. 37, p. 197.
- 5- Hassayneh and Kassem (2006) The academic reasons for the failure of the university student Twenty - seventh Conference Arab Organization for Admissions and Registration in Universities in Arab Countries, University of Sharjah. (2006).

4.3References

1. Hassayneh and Kassem (2006) The academic reasons for the failure of the university student Twenty - seventh Conference Arab Organization for Admissions and Registration in Universities in Arab Countries, University of Sharjah. (2006).
 - 2.
 - 3.
 4. Basiony study (2006) The role of Mansoura University in solving the academic stumbling problem of university students Seventh Conference Twenty of the Arab organization responsible for admissions and registration in universities in Arab countries, University of Sharjah .
 5. Hassan's study(2001) the reasons of the failure of students at the Faculty of Education Zagazig University. *College Journal* 285. Education in Zagazig, January, No. 37, p. 197.
 6. Bradley, a (2000). "The Gatekeeping Challenge" *Education Week*, Vol XIX, Number 18, January.
 7. Barth, p., Haycock, H.Jackson, K.Mora, p.Ruiz, s. Robinson & a.willins, Eds.(1999) *Dispelling the myth : High Poverty Schools Exceeding exectations*. Washington, D.C.: Report of the Education Trust in cooperation with the council of chief state school officers and partially funded by the U.S. development of education.
- 6-www.uofg.edu.sd .