



RESEARCH ARTICLE

THE IMPACT OF HEALTH-RELATED PROBLEM ON ACADEMIC ACHIEVEMENT OF STUDENTS: APPLICATION OF DUMMY VARIABLE REGRESSION MODEL

Sebsib Muanenda.

Department of Statistics, College of Natural and Computational Sciences, Wolkite University, Wolkite, Ethiopia.

Manuscript Info

Manuscript History

Received: 19 July 2018

Final Accepted: 25 August 2018

Published: September 2018

Keywords:-

Health-related problem, academic achievement, Regression, Samara University

Abstract

Health is a base for social, cultural and economic development of society. But sometimes attaining this health, unfavorable condition is impossible due to some health problems. The aim of this study is to identify the main impact health-related problems of student's academic achievement in Samara University.

The type of the data in this study is primary data collected from students. In this study, we used some statistical techniques such as stratified random sampling and some statistical models such as regression and chi-square which is used as inferential model to determine relationship between students' GPA and the impact of health related problems. Some statistical software packages like SPSS version 20 used in this study for analysis the data. There are 11 variables selected as health related problems that affect student academic achievement of student in Samara University those are sex, age, drug use, smoking cigarette, sleeping disorder, mental problem, TB, typhoid and gastric, but the study shows only there are five variables that affect student academic achievement of student in Samara University those are typhoid, TB, mental problem, sleeping disorder and gastric disease are the most variables that affect the students' academic achievement. Finally this study recommend that the University or the concern body will have to consider these problems and try to find the immediate solutions in order to the students to become successful in their academic achievement

Copy Right, IJAR, 2018,. All rights reserved.

Introduction:-

Background of the study:-

World Health Organization (WHO) is the base for social, cultural and economic development of the country. There are several health problems that affect University students in country special students who are do poorly in University will have more health problems which adversely affect their academic achievement. More health problem

Corresponding Author:-SebsibMuanenda.

Address:-Department of Statistics, College of Natural and Computational Sciences, Wolkite University, Wolkite, Ethiopia.

students have less likely they will succeed graduation time (Fossella Kitz, 2009). Each health problems that can be removed has the potential to positively influence academic achievement behaviors. They are many proven intervention that have a positive impact on the students health and academic achievements.

Most of the time, students are affected by different diseases like, TB, gastric, typhoid, sleeping disorders, mental problems and common cold. Not all students are able to succeed in the University and that certain group of students is consistently less likely to have success than others. Student's health behaviors and educational challenges may influence each other.

Based on seeming under recognized evidence over all finding. There is pre-season to believe health does have an impact on the academic achievement of students. Education and health are known to be highly correlated. The effect of health on academic achievement has been well researched in developing country. Students in university faced more series challenges (Miller, 2009).

For the past Decade University has understood and acted the need to dramatically increase students on academic achievement. In recent years, university has placed strong focus on academic achievement gap works across traditional jurisdiction to determine how, when and where to effectively and efficiently proved the educational (learns supports and services students need, and pay closer attention to how University can create more caring culturally prevalent and engaging environments for all students (Hillman *et al.*, 2009).

Students have the best chance to succeed when they are healthy. Health in this context includes nutrition diet, physical activity, emotional well beings satisfy security, the absence of chronic condition such as, TB, gastric, typhoid, sleeping disorder, mental problems and common cold etc. Students who are exposed to violence show pure academic achievement. Many of these students expressed past traumatic stress disorder feel hapless and depressed have reduced motivation and persistence to learn. Thus, this study has been intended to fill this gap with focus on the impact of health related problems on academic achievement of students in Samara University.

Statement of the problem:-

Health service is the basic need of the society. Health is a factor in the development of the country hence it determine the future of nation by affecting the well being of its population for these radiation of its human potential. Samara University clinic is one of the governmental health institution in which the campus students are treated. Some previous students shown factor affecting the academic achievement of students is:

TB, gastric, typhoid, sleeping disorders, mental problems and common cold are greatly affecting academic achievement of students. This study attempts to investigation the important and immediate health problems that affect the academic achievement of students in Samara University. The study attempts to address the following questions.

1. Which factors have major effect on the academic achievements of students in Samar University?
2. Does health have a significant impact on educational performance?

Objective of the study:-

The general objective of this study is to investigate the impact of health-related problems on academic achievement of students in Samara University.

The specific objectives of this study are:

1. To assess the overall impact of health related problems on the academic achievement of students using descriptive statistics.
2. To identify the factors that has an impact on academic achievement of students using regression model.

Methods:-

Description of the study area:-

This study was conducted on the impact of health related problems on academic achievement of students in Samara University, which is found in Afar region. Samara University is one of governmental institution in Ethiopia. It located 588 Km North-East of the capital city of Ethiopia (Addis Ababa). The annual range of temperature ranges from 23 to 45 degree Celsius and the mean annual rain fall below 500milli meter.

Sampling techniques and sample size determination:-

Target population:-

Target populations of this study were first, second and third year students in Samara University in 2014 batch.

Sampling techniques:-

Sampling technique is a systematic or a system of taking small ratio of observation from large population with the aim of getting information of this large population from the small observation by using some statistical techniques (Cochran, 1977). A stratified random technique is the probability, which involves the stratification of population by portioning the sample frame in to non-over lapping and relatively homogenous groups called strata. In this study stratified random sampling was used to select the entire sample because of students are the homogeneity of the population with respect to the characteristics of interest within their batch.

Sample size determination:-

Even if there are a number of sample size determination formula approaches (like personal judgments and budgetary approach), the one in which the investigator interesting is based on precision with some confidence level. An appropriate sample is one of the means of gaining high precision and greater accuracy with minimum cost. That is in addition to the purpose of the study and population size, three criteria usually will need to be specified to determine the appropriate sample size: the level of precision, the level of confidence or risk, and the degree of variability in the attributes being measured (Cochran, 1977).

The Level of Precision:-

The level of precision, sometimes called sampling error, is the range in which the true value of the population is estimated to be.

Degree of Variability:-

The degree of variability in the attributes being measured refers to the distribution of attributes in the population. The more heterogeneous a population, the larger the sample size required to obtain a given level of precision. The less variable (more homogeneous) a population is, the smaller the sample size will be.

Confidence level:-

There is always a chance that the sample you obtain does not represent the true population value. This risk is reduced for 99% confidence levels and increased for 90% (or lower) confidence levels. The total number of Samara University students 1st, 2nd and 3rd year students are 4264. Therefore, the appropriate sample size can be determined by taking a pilot survey from 1st, 2nd and 3rd year student's cumulative grade point average (CGPA). The study uses stratified random sampling technique and accordingly the sample size determination formula adopted is given by [8].

$$n = \frac{\sum_{i=1}^h W_i * S_i^2}{\frac{e^2}{Z^2 \alpha/2} + \frac{\sum_{i=1}^h W_i * S_i^2}{N}} \text{-----[1]}$$

Where n is the sample size, N is total number of students in this study, h is number of strata, $Z\alpha/2$ is the value of standard normal distribution that gives an area of $\alpha/2$ to the right of it, e is the desired level of precision (in the same unit of measure as the variance), S_i^2 is the estimated variance of cumulative grade point of students for the i^{th} strata obtained through pilot survey, $W_i = \frac{N_h}{N}$, weight for the i^{th} strata N_h is total number of students in each strata. Thus using $\alpha = 5\%$ (95% confidence level), $e = 1\%$, $S_1^2 = 0.152$, $S_2^2 = 0.193$, $S_3^2 = 0.196$, $N_1 = 1097$, $N_2 = 1810$, $N_3 = 1357$, $N = 4264$, $W_1 = 0.257$, $W_2 = 0.424$ and $W_3 = 0.318$.

The sample size for this particular study is computed as

$$n = \frac{\sum_{i=1}^3 (w_1 s_1^2 + w_2 s_2^2 + w_3 s_3^2)}{\frac{d^2}{z^2 \alpha/2} + \sum_{i=1}^3 (w_1 s_1^2 + w_2 s_2^2 + w_3 s_3^2)}$$

$$n = \frac{0.257 * 0.152 + 0.424 * 0.193 + 0.318 * 0.196}{\frac{(0.1)^2}{(1.96)^2} + \frac{0.257*0.152+0.424*0.193+0.318*0.196}{4264}} = 70$$

Thus finally 70 students have been selected using random number table from the area to represent all of the students in the University.

Method of data collection:-

The study used primary data that were collected using structural self administer questionnaire would be used as instrumental tool of gathering the information from the respondents that is the select sample from impact of health related problems on academic achievement of students in samara University. The questionnaire involves both closed and open ended question and distributed to the respondents. The respondent filled the questions give to the researchers. The data from questionnaires were entered into SPSS and cleaned to eliminate errors and then analyzed.

Method of data analysis:-

After data collection is over, the collected data was carefully edited, coded, tabulated and organized depending on the type of questions and the nature of the data before analysis. The data gathered from the sample was analyzed by using both descriptive statistical methods and inferential statistical methods.

Descriptive statistics: -

it is statistical methods that helps us to summarize a given sets of data using tables, diagrams, graphs and summary measures such as mean, median and standard deviation (Bluman, 2012). For this particular study mainly frequency distribution table was used to summarize the most important features of the sample data.

Inferential statistics: -

it is statistical method that helps us to make inference about the population parameters depending on the results obtained from sample data. Making statistical estimation and conducting statistical hypothesis testing are methods through which inferences are made. This study has used the dummy variable regression model to identify the most significant predictors of the impact of health-related problems in the academic performance of students in Samara University.

Regression statistical method:-

is a conceptually simple method for investigating functional relationships among variables. The relationship is expressed in the form of an equation or a model connecting the response or dependent variable and one or more explanatory or predictor variables (Walter *et al.*, 2006).

Multiple linear regression models: -

is a model that involves more than one explanatory variable. The goal of multiple linear regression models is to model a linear relationship between the response variable and several independent variables. The multiple linear regressions assume that the response variable is a linear function of the model parameters.

In this study multiple linear regressions model was used which can be given by the following model

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k + \varepsilon \text{ ----- [2]}$$

Where Y_i = cumulative grade point of students,

Age of students, sex of students; batch of students; students suffered by typhoid; smoking cigarette; students suffered by TB; Mental problem; drinking of alcohol; sleeping disorder; drug use and gastric disease all are dummy independent variables. Dummy variable means variables that can be categorized in to two categories. ε is assumed to be a random error terms and it accounts for the failure of the model to fit the data exactly. Vector β are the coefficients of the explanatory variable $X_1, X_2 \dots X_k$; are unknown constants to be determined (estimated) from the sample data using the least square estimation technique. Least square estimation technique estimates unknown parameters by minimizing the sum square of distance of each value from the fitted value and α is multiplicative constant or intercept.

Assumption of the multiple linear regression models:-

The validity of a statistical method, such as regression analysis, depends on certain assumptions. Assumptions are usually made about the data and the model. The accuracy of the analysis and the conclusions derived from an analysis depends crucially on the validity of these assumptions. For regression model the following assumptions are to be addressed.

Assumption about the form of the model:-

Linearity: The model should specify a linear relationship between dependent variable and the predictor variables. The linearity assumption is not as narrow as it might first appear. In the regression context, linearity refers to the manner in which the parameters and the error term enter the equation, not necessarily to the relationship among the variables.

Assumptions about the error term:-

1. **Normality:** The error term is a random variable and should have a normal distribution with mean vector equal to zero and variance-covariance matrix (δ^2). Symbolically, this is shown as $\mathcal{E} \sim N(0, \delta^2)$.
2. **Homoscedasticity:** The error variance (δ^2) is assumed to be constant. The variance should not increase as the values of the independent variables increases.
3. **Non autocorrelation:** The error term should also have no correlation to each other. The errors associated to different observations should be independent to each other. The deviation of y_i from the line should not affect the deviation of y_j from the line. That is correlation $(\mathcal{E}_i, \mathcal{E}_j) = 0$

Assumption about the independent variables :-

The X matrix should be full rank. That is the independent variables must be linearly independent to each other

Measuring the quality of fit (assessing model adequacy):-

After fitting a linear model relating Y to X 's, we are interested not only in knowing whether a linear relationship exists, but also in measuring the quality of the fit of the model to the data. The quality of the fit can be assessed by one of the following highly related mechanisms

Testing the regression coefficients associated with all the independent variables:-

Testing the null hypothesis that all the regression coefficients are zero against the alternative which state that at least one of the regression coefficients are different from zero. Rejection of the null hypothesis indicates the adequacy (quality) of the model to fit the data very well.

Determining the coefficient of determination:-

Coefficient of determination is the percentage of the total variation in the response variable which is explained by the presence of the independent variable in the model. Coefficient of determination has values in between 0 and 1 and can also be expressed in percent having values ranging from 0 to 100. Usually high value of R^2 indicates the quality of the model to fit the data very well and low values of R^2 indicates the poor model to fit the data.

Checking if there are any violation of regression assumptions :-

Adequacy of the model is also related with the regression assumptions set. We reemphasize that the regression assumptions should be checked before drawing statistical conclusions from the analysis (e.g., conducting tests of hypothesis or constructing confidence or prediction intervals) because the validity of these statistical procedures hinges on the validity of the assumptions.

Results:-

One of the most challenging or rewarding task in survey research comes after the data have been collected and coding, editing and preliminary processing have been completed. This section contains statistically analysis results and interpretation of the findings.

Summarization of most important characteristics:-

A data fom sample of 70 respondents in Samara University was collected to achieve the main objective of this study. The different characteristics of the sampled population, a table which contains frequency and percent was used to make summary of the variables.

Table 1:-Descriptive statistics for variables, Samara 2014

Variable	Category	Student's CGPA								Total
		2.00-2.50		2.50-3.00		3.00-3.50		Above 3.50		
		Count	Row %	Count	Row %	Count	Row %	Count	Row %	
Sex	Female	10	28.6%	13	37.1%	5	14.3%	7	20%	35
	Male	2	5.7%	14	40%	10	28.6%	9	25.7%	35
Age of student	17 to 19	6	37.5%	6	37.5%	1	6.3%	3	18.7%	16
	20 to 22	2	4.9%	17	41.4%	9	22%	13	31.7%	41
	Above22	4	30.8%	4	30.8%	5	38.4%	0	0%	13
Batch of student	First	4	23.5%	4	23.5%	5	29.5%	4	23.5%	17
	Second	6	27.3%	8	36.3%	2	9.1%	6	27.3%	22
	Third	2	6.4%	15	48.4%	8	25.8%	6	19.4%	31
Student suffered by typhoid	Yes	5	15.6%	16	50%	6	18.8%	5	15.6%	32
	No	7	18.4%	11	28.9%	9	23.8%	11	28.9%	38
Student Smoking cigarette	Yes	0	0%	2	40%	0	0%	3	60%	5
	No	12	18.5%	25	38.5%	15	23%	13	20%	65
Student suffered by TB	Yes	0	0%	4	100%	0	0%	0	0%	4
	No	12	18.3%	23	34.8%	15	22.7%	16	24.2%	66
Drinking alcohol	Yes	3	25%	3	25%	6	50%	0	0%	12
	No	9	15.5%	24	41.4%	9	15.5%	16	27.6%	58
Sleeping disorder	Yes	4	33.3%	2	16.7%	0	0%	6	50%	12
	No	6	10.3%	27	46.6%	15	25.9%	10	17.2%	58
Students suffered by drug use	Yes	8	20.5%	14	35.9%	10	25.7%	7	17.9%	39
	No	4	13.8%	13	44.8%	5	17.2%	7	24.2%	29
Students suffered by gastric disease	Yes	3	10.3%	13	44.8%	7	24.2%	6	20.7%	29
	No	9	22%	14	34.1%	8	19.5%	10	24.4%	41

Table 1 summarizes the categorical characteristics by using percents. It can be observed that 28.6 % of the female respondents and 5.7% of male respondents were there CGPA between 2.00 up to 2.50. Surprisingly about 44.8% of the respondents whose CGPA is between 2.50 and 3.00 are suffered by gastric diseases. Most of the students whose CGPA is more than 3.00 are easily suffered by different health related diseases.

Multiple regressions; student's cumulative grade point average versus age of students, sex, student's batch:-

The multiple regression models were applied to identify the impacts of health related problems of student's cumulative grade point average (CGPA) with the selected dummy independent variables. Before the main parameter estimation table is presented the summary statistics of the model is given the table below.

Table 2:-model summary statistics, Samara 2014

Coefficient of determination	Adjusted Coefficient of determination	Standard Error of the Estimate
.87	.82	52.34

The value of Coefficient of determination is 0.87 which indicates that 87% of the total variation in student's academic achievement is attributed by the factors which are included in the model. This large value of coefficient of variation is an indication for the model to fit the data very well. On the contrary the standard error of the estimates is not small. This may indicate some problem in the model if it does; it will be find out during assumptions checking.

Table 3:-Regression results showing the effects of selected explanatory variables on academic performance of students, Samara 2015

Attributes	Estimates	Standard error of the estimates	t-value	P-value	VIF
(Constant)	2.086	.387	.221	.826	
Sex of students (male)	-.157	.210	-.746	.458	2.239

Age of students (≥ 20)	-.158	.153	-1.031	.307	1.442
Batch of students (above 2 nd)	-.052	.128	-.407	.685	1.483
Students suffered by typhoid (Yes)	.399	.195	2.050	.045	2.792
Students smoking cigarette (Yes)	.167	.331	.506	.615	1.136
Students suffered by TB (Yes)	.599	.295	2.029	.047	1.423
Students suffered by mental problems (Yes)	.703	.271	2.594	.012	1.217
Students drinking alcohol (Yes)	-.082	.387	-.213	.832	2.102
Students suffered by sleeping disorder (Yes)	.523	.217	2.413	.019	1.254
Students suffered by drug use (Yes)	-.266	.336	-.792	.431	1.118
Students suffered by gastric disease (Yes)	.436	.200	2.184	.033	2.143
p-value for analysis of variance table=0.000					

Statistical test of individual predictors:-

Compare the P-value with the given level of significance to test the significance of the individual predictors to the model. The hypotheses that are used to test individual predictors are stated as:

Ho: $\beta_i = 0$ (the i^{th} factor is not statistically significant impact on the academic performance of students)

Ha: $\beta_i \neq 0$ (the i^{th} factor is statistically significant impact on the academic performance of students)

Decision rule:-

Reject Ho if the P-value is less than the pre-selected level of significance (5%). Since the P-value associated with students suffered by typhoid (0.045) is less than the selected level of significance (0.05), there is not enough evidence not to reject the null hypotheses. Therefore, it can be concluded that students suffered by typhoid are statistically significant impact on the academic performance of students in Samara University.

The p-values associated with gender, age of student, batch of students, smoking cigarette, drinking alcohol and drug usage are all not less than the pre-selected level of significance. Thus gender, age of student, batch of students, smoking cigarette, drinking alcohol and drug usage all are found to be statistically insignificant predictors of academic performance of students at 5% level of significance in Samara University.

Testing the overall model adequacy:-

The analysis of variance table gives us the sources of variation in a regression model with their respective degree of freedom, sum of squares of variation, mean square of the variation, the corresponding F-statistic and the p value. This table is used to test the overall model adequacy at 5 % level of significance. The hypothesis

Ho; all the model parameters are zero (the model is not adequate)

Ha; at least one of the model parameter is different from zero (the model is adequate)

The p-value for the analysis of variance is (0.00) which is less than the pre-selected level of significance (5%), the null hypothesis is rejected and it can be concluded that the model is adequate to fit the data.

Checking the assumptions of multiple linear regression :-

Test for Multicollinearity:-

One of the strong assumptions that lie under multiple regression models is no Multicollinearity. It means that the independent variables must not be correlated with each other. One method of detecting Multicollinearity is variance inflation factor (VIF). If the VIF of each independent variable is less than 5, there is no correlation among the explanatory variables in the model. From the table 3 of the last column, one can see that the variance inflation factor (VIF) of each explanatory variable is less than 5; this implies that there is no Multicollinearity problem in the model.

Residual plots:-

The residual plots such as histogram of the residuals and scatter plot of fitted values against the residuals have been used in this study to check if there is any violation of the model assumptions. Both plots indicate that there are no series model defects in the above multiple regression model.

Conclusion and Discussion:-

In previous sections, the results obtained from both descriptive statistics and the multiple linear regressions were presented. This section tries to conclude and discuss the finding of this paper by relating with other similar works. Among the predictor variables that considered in this study, the student academic performance were affected for student's that have distinction grade point average. The cross-sectional study confounded by the fact that students suffered by mental disorders tend to significant effect on the academic performance students. This finding is to the finding in previous study (Halmi, 2000). The link between education and health has different potentials that learned in University provide an incentive to individuals to study health and rap in order to achieve their academic achievement. In this regard high light difference in preferences and individuals valuation of their future that may be affected by levels of education as relevant factors explained health .As key input in the health production function, students full health helps individuals maintain and to achieve the academic achievement .One factors which is affected the academic achievement of students are un health ,which mean that affected by different diseases like typhoid ,TB (Gang, 2010). This finding is consistent with this study. Based on the finding of this study the variable students suffered by typhoid, TB, mental problem, sleeping disorder and gastric disease are found to be statistically significant impact on academic performance of students in Samara University. But the variables gender, age of student, batch of students, smoking cigarette, drinking alcohol and drug usage have no any impact on academic performance of students in the University.

Recommendation:-

The University should plan to build big clinic in the University for student because most students are affected by different diseases and leads to significantly impact on student academic achievement.

References:-

1. Bluman: Elementary statistics, A step by step approach ,8th edition;2012
2. Cochran, W. G.: Sampling Techniques, 3rd Ed., New York: John Wiley and Sons, Inc.;1977
3. Fossella K. and Kitz man: Disparities in academic achievement and health the instruction of student education and health policy. Pediatrics 123; 1073-80; 2009
4. Gang l.: Estimating interdependence between health and education in a dynamic model .Cambridge MA; NBER; 2010
5. Halmi K...: Perfectionism in Anorexia Nervosa: Variation by Clinical Subtype, Obsessionality, and Pathological Eating Behavior. The American Journal of Psychiatry, 157(11); 2000
6. Hillman C. Potific B. and A.F Kramer: The effect of acute treadmill walhing on cognitive control and academic achievement of students .159.1044.54; 2009
7. Miller J.: Being drinking and associated health risk behaviors among high school student. pediatrics 119(1); 2009
8. Walter A. Shewhart and Samuel S.: Regression analysis by example,4th edition, john Wiley & sons, inc., publication;2006.