ISSN 2347-1921
Volume 13 Number 3
Journal of Advance in Mathematics

A Statistical Approach to Determine Factors that Influence Students' Academic Performance

[1]F. Z. Okwonu and [2]N.O. Ogini
[1]Department of Mathematics
Faculty of Science, Delta State University, Abraka, Nigeria
[2]Department of Computer Science
Faculty of Science, Delta State University, Abraka, Nigeria
fzokwonu_delsu@yahoo.com

ABSTRACT

This paper provides answer to the following question. Does the performance of students' depend on the location of accommodation and facilities? The Chi square and the likelihood ratio Chi square statistics based on the contingency table are applied to answer the above question. The statistical analysis using the above techniques revealed that students' performance depends on the location of the accommodation and the facilities. Result based on the odd ratio indicated that there is no association between accommodation location and facilities.

Indexing terms/Keywords

Contingency Table; Chi squared; Likelihood ration Chi squared

Academic Discipline And Sub-Disciplines

Mathematics and Statistics

SUBJECT CLASSIFICATION

Mathematics Subject Classification 2010:03C45, 03C65

TYPE (METHOD/APPROACH)

Survey/Interview

INTRODUCTION

The Pearson Chi square test and the contingency table are inseparable. This is so because the later illustrate the diagrammatic features and the former is based on the decision rule, this implies using the decision rule based on the hypothesis testing. The process allows decision on whether the null hypothesis is rejected or accepted. The Chi square test can be applied to determine the fitness of a data set otherwise referred to as goodness of fit test. Contingency table is applied to organize categorical variables and testing hypothesis using the Chi-square test to determine independence. In general, the contingency table describes the relationship between two categorical variables. The contingency table is often applied to determine the effectiveness of a system under study or if the effectiveness of the system is based on certain profile variables[1]. The null hypothesis tested with a Chi-square test using the 2 x 2 contingency table is considered as test of independence [4,8-11].

The 2 x 2 contingency table combines association and the interactions of the attributes in pairs[6]. The degree of freedom of the Chi square suggests that there are different ways in which the probabilities differs from complete independence and is mathematically tractable to associate the profile variables[8-11]. The table below describes the contingency table. The illustration in [5] implies that the contingency table are not ordered but scaled based on the factors(resident and non-resident) and level.

Table1. 2 by 2 Contingency Table

Residential type\gender	Male	Female	Total
Resident	w	c	w+c
Non Resident	b	d	b+d
Row Total	w+b	c+d	w+b+c+d

Applying the probability concept, we have $\sum \varepsilon = 1, \varepsilon = w + b + c + d.$ Let $s(w, b, c, d)$ denotes the measure of association of resident and non resident subject to categorization as male or female then in similar formulation as in above table, we infer that since it is a 2 x 2 contingency table, then $s(w, b, c, d) = s(wp, bq, cp, dq).$ This system appears straight forward like the confusion matrix, such that

$$s(w, c, w\ p, cp) = s(b, d, bq, dq),$$
$$s(w, b, w\ p, bq) = s(c, d, cp, dq).$$

In dealing with problem associated with interaction and independence, the condition of symmetric is usually not fulfilled.

In this discussion, we consider other measure sure as the root mean square. This formulation can be expressed mathematically as follows;

$$R_w = \frac{wd - bc}{\sqrt{[(w+b)(w+c)(b+d)(c+d)]}}$$

However, the degree of association between profiles can be measured by the following

$$\delta = \frac{bc}{wd}.$$

This formula was famously considered by Fisher, 1935 [6] and described further that simple function such as the coefficient of association

$$\varpi = \frac{(1-\delta)}{(1+\delta)},$$

or the coefficient of colligation

$$\varpi_c = \frac{(1-\sqrt{\delta})}{(1+\sqrt{\delta})},$$

can be used to determine the factor of interest.

This paper investigates the effect of accommodation location and facilities on students'academic performance. The reminder of this paper is organized as follows. Section Two describes the methods which includes the Chi square and the likelihood ratio Chi square techniques. Results and discussion is contained in Section Three while conclusion is reported in Section Four.

METHODS

The Chi Square

The Chi square test is known for its applications. The formula is well described in different literature. However, for computational convienence we described the profile of interest as follows; let e denotes the expected value such that

$$e = \frac{t_n \times p_m}{K} = \sum \varepsilon_{nm},$$

where $t_n \text{ and}, p_m$ are the row and column and K is the sample size. Applying the expected value (e) and the observed value (β), the Chi square can be computed as follows;

$$\chi^2 = \frac{\sum_{n=1}^{R}\sum_{m=1}^{C}(e-\beta)^2}{e}.$$

The Chi square test statistics is approximately distributed as χ^2 on $(r-1)(c-1)$ degrees of freedom[3,12]. Oberve that the Chi square statistics is approximated by the Chi square distribution, thus this approximation is affected with small expected value. Howell[3] observed that when the expected value is small, the Chi square statistic is discrete. Thus, the accepted rule with respect to the minimum expected value should be at least five. However, the minimum of 'five' is a rule of thumb posited by [2] which he admitted can be "chosen arbitrarily"

The Likelihood Ratio Chi Square

Apart from the Chi square statistics, other test statistics exist. The likelihood ratio Chi square is designed based on the likelihood of the data set with respect to the null hypothesis[3]. This procedure is defined as

ISSN 2347-1921
Volume 13 Number 3
Journal of Advance in Mathematics

$$G^2 = 2\sum[\beta_{nm} \log\left(\frac{\beta_{nm}}{e_{nm}}\right)].$$

The uniqueness of this technique is that for large dimensional table, such contingency table can be decomposed into smaller tables. It is uniquely applied to log linear analysis[3]. It was observed that as the sample size increases the Chi square and the likelihood Chi square statistics converges or tends to infinity[7,10].

The test hypothesis is designed as follows;

H_0 : students performance does not depends on accommodation location and facilities, that is,

$$H_0 : \chi^2_{comp} > \chi^2_{table(0.05)},.$$

H_A : students performance depends on accommodation location and facilities, that is,

$$H_A : \chi^2_{comp} < \chi^2_{table(0.05)}.$$

Therefore, the comparative acceptance of either of the above, lies strictly on computed Chi square value. In this discussion we use 5% level of significance to determine the acceptability or otherwise of the hypothesis.

RESULTS and DISCUSSION

This study was designed to investigate the influence of accommodation location and facilities on student academic performance. The residential locations are categorized into two groups. The first group comprises of students residing within the university premises (resident) and the other group of students that reside outside the university premises (non-resident). In each category, a total of 300 students participated in the study. The table below contains the detail of the study.

Table 2. Observed value

Residential location\gender	Male	Femal e	total
resident	118	127	245
Non resident	130	133	263
Row total	248	260	508

Table 3. Expected value

Residential location\gender	Male	Female	Total
Resident	119.61	125.39	245
Non resident	128.39	134.61	263
	248	260	508

Table 4: Chi square value

Observed value (β)	**Expected value** (e)	$\mathrm{d} = (\beta - \mathrm{e})$	d^2	$\chi^2 = \frac{d^2}{e}$
118	119.61	-1.61	2.5921	0.02167
127	125.39	1.61	2.5921	0.02067
130	128.39	1.61	15.76	0.02019
133	134.61	-1.61	1.21	0.01927
				$\chi^2 = 0.0818$

Since the computed Chi square value is less than the table value (3.84) at 5% level of significance, we accept the alternate hypothesis indicating that students' performance depends on the location of the accommodation and facilities. The null hypothesis tested with a Chi-square test based on a 2 x 2 contingency table is considered as test of independence [4]. The results based on the likelihood ratio Chi square support the acceptance of the alternate hypothesis, since the likelihood ratio Chi square is less than the Chi square value at 5% level of significance, that is 0.0818<3.84. Both techniques support the notion that students performance depends on the location of the accommodation and facilities. The odd ratio value for this data set equal to 1.05, which implies that the location of accommodation does not implies availability of facilities, as such no association between location and facilities.

CONCLUSIONS

This paper ascertains if the performance of student depends on the location of accommodation and facilities. The Chi square and the likelihood ratio Chi square statistics were applied to determine the factors that affect students' performance. These techniques revealed that students' performance strictly depends on the location of the accommodation and facilities. Therefore, the recommendation based on this study indicates that accommodation location and facilities enhance better performance. This implies that parents and the school authorities should provide conducive learning environment for their wards or students to enhance their academic performance.

ACKNOWLEDGMENTS

Our thanks to the experts who have contributed towards development of the template.

REFERENCES

[1] Albrecht, C., 2013. Contingency table and the chi square statistics: Interpreting computer printouts and constructing tables http://extension.usu.edu/evaluation/files/uploads/Start%20Your%20Engine/Study%20the%20Route/Analyze%20the%20Data/Interpreting_Chi_Square_Printouts.pdf .

[2] Cochran, W. G., 1952. The x^2 test of goodness of fit. Annals of Mathematical Statistics 25: 315-345.

[3] Howell, D. C., 2012. Chi-square test-analysis of contingency tables. http://www.uvm.edu/~dhowell/methods8/supplement .

[4] Ingersoll G.M., 2010. Analysis of 2x2 contingency tables in educational research and evaluation. International Journal for Research in Education,1-14.

[5] Lauritzen, S. L., 2002. Lectures on contingency tables, Aalborge University, 1-53.

[6] Simpson, E. H., 1951. The interpretation of interaction in contingency tables. Journal of the royal statistical society, Series B (Methodological) 13: 238-241.

[7] Wilks, S.S., 1935, The likelihood test of independence in contingency tables. Annal of Mathematical Statistics, 6:4, 190-196.

[8] Okwonu, F. Z., 2016. A Chi Square Approach to Determine the Effect of Students' Residence on Academic Performance; Journal of Basic and Applied Research International,15,4 :280-286.[9] [9]Okwonu, F. Z., 2015. Analyzing transportation preference using contingency table, Journal of statistics and application, 3:3, 25-28.

[10] Okwonu**,** F. Z., 2015. Analyzing the cause of vehicular breakdown using contingency table: A preliminary study on DBS/Core Area and Okpanam, Asaba, Delta State, Nigeria. Asian Journal of Mathematics and Computer Research (AJOMCOR), Vol. 4. No. 3, Pp. 164-168, 2015.

[11] Okwonu, F. Z.,2015. The ugly side of sentimental politics: The Nigeria situation using statistical model. Journal of Multidisciplinary Engineering Science and Technology(JMSET), Vol. 2, No. 8, 1983-1987, 2015.

[12] http://math.hws.edu/javamath/ryan/ChiSqu

Author' biography

F. Z. Okwonu, hold a Ph. D in robust computational statistics with special interest in robust classification and decision analysis. He is currently with the Delta State University, Abraka, Nigeria.