

## Application of Log-Linear Model for Four Dimensional Contingency Table in Analysis of Students Anxiety in Mathematics

Obianuju Gloria Aziwe\* and Cecilia Nchedo Okoli

Department of Statistics, Chukwuemeka Odumegwu Ojukwu University, Uli, Anambra State, Nigeria.

Corresponding author email: [obianujuaziwe@gmail.com](mailto:obianujuaziwe@gmail.com)

### ABSTRACT

The study applied log-linear model for four contingency table in analysis of anxiety in mathematics among Nigerian students. The objectives are: to compare the log-linear models and select the best model, to determine whether mathematics anxiety level among university students is related to gender, course of study and mathematics performances using log-linear model and to examine the association between genders, department of the students, their first semester mathematics performance using the log-linear model. The population of the study is comprised of the undergraduate university students in Delta State, Nigeria. A total of one hundred and eighty (180) respondents constitute the sample size of the study and the data were collected using structured schedule. The independent model stated as;  $\log(\mu_{ijkl}) = \lambda + \lambda^M + \lambda^A + \lambda^K + \lambda^P$  was found to be the best model out of the models under consideration as it has the least AIC and BIC values. And the null hypothesis  $H_0$  which states that there is no relationship among the variables under study was not rejected as the  $p$ -value is greater than the critical value as revealed by the goodness-of-fit test statistic table for the best model. This indicates that mathematics anxiety among university students is independent of their gender, department and performance in mathematics.

**Keywords:** Log-linear models, Contingency table, Mathematics anxiety, Akaike Information criterion (AIC) Bayesian information criterion (BIC).

### Article information

Received 13 February 2025;

Accepted 8 March 2025;

Published 12 March 2025

<https://doi.org/10.26765/DRJMSS96951257>

Citation: Aziwe, O.G., and Okoli, C.N. (2025). Application of Log-Linear Model for Four Dimensional Contingency Table in Analysis of Students Anxiety in Mathematics. Direct Research Journal of Management and Strategic Studies. Vol. 6(1), Pp. 11-17.

This article is published under the terms of the Creative Commons Attribution License 4.0.

### INTRODUCTION

Log-linear models are generalized linear models commonly used to model the frequency in contingency table. It consists of number of parameters which are combined together to determine an estimated frequency for each cell in the contingency table. Frequency or count in each cell of the table is the response variable while the explanatory variables are usually categorical. Log-linear modelling is specifically used to identify the main effects and interactions required to express the joint distribution in multivariate cross-classifications. Applications of these models can be seen in wide variety of disciplines including marketing, social psychological

research (Brzezinska, 2012). Log-linear models can either be hierarchical or non-hierarchical in nature. Howell (2009), points that hierarchical log-linear models represent all the lower and main effects interactions in a model. Christensen (1997), also pointed that these models are very convenient and most widely used mathematical models to analyze categorical data represented in cross-classified nature in the contingency tables. The analysis of two dimensional tables and three dimensional contingency tables are less complex compared to four dimensional contingency tables as they involves less variables and parameters, that is to say that the complexity

of the model depends on the number of variables and parameters involved and as such explanation of results of analysis and interpretation will be more difficult for higher dimensional contingency tables like four dimensional tables.

Mathematics is an important subject in curriculum in every country (Mutawah, 2015). (Zakaira et al., 2012) assert that the study of mathematics was established to allow an individual to apply knowledge of mathematics effectively and responsibly in solving problems and making decisions in everyday life. Some student when heard of mathematics feel anxious as a result of the calculations involve in it.

Mathematics anxiety being ones negative reaction to situations dealing with numbers, equations and mathematical calculations is defined by (Richardson and Suinn, 1972) as “ a feeling of tension and anxiety that interfere with the manipulation of numbers and the solving of mathematical problem in a wide variety of ordinary life and academic situations.” Mathematics anxiety can go to severe from mild, from minor frustration to overwhelming emotional and physiological trouble (Ashcraft, 2012).

## LITERATURE REVIEW

Several studies have been done on log-linear model, for instance, Onder and Adiguzel (2010) applied hierarchical log linear analysis method to occupational fatalities in the underground coal mines of Turkish Hardcoal Enterprises, the accidents records were evaluated and the main factors affecting the accidents were defined and the results found showed that the mostly affected job group by the fatality accidents was the production workers and these workers were mostly exposed to roof collapses and methane explosions. Olmus(2012) also presented a modeling effort in estimating the relationships between driver's fault and carelessness and the traffic variables, the result of the analysis showed that the best fit model regarding these variables was log-linear model, also the associations of the factors with the accident severity and the contributions of the various factors and interactions between these variables were assessed and the obtained results provided valuable information in regard to preventing undesired consequences of traffic accidents. Furthermore, Mustofa et al (2015) applied higher dimensional log-linear model to four categorical variables in education, it was found that the best model for data with three variable interactions are; Length of Study\*SEX\*GPA, Length of Study\*Science\*GPA, and Sex\*Science\*GPA. In addition, Okoli, Onyeagu and Osuji (2014), developed log-linear model for association among breed, age and chick loss of chicken; their findings revealed that all the two-factor interactions were significant except breed\*chick loss implying that the breed is independent of chick loss. It was also found in their study

that breed\*age, and age\*chick loss are the best model using likelihood ratio statistic ( $G^2$ ) and; Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) to checkmate the adequacy of the models. Also, Brzezinska (2012) utilizes likelihood ratio approach and after that the author discusses how to use AIC and BIC information criteria for selection of the best log-linear model. In construction and application of log-linear models to assess academic performance by Indika (2019) it was found that the number of hours the students sleep and the number of hours they watch TV significantly influence students' academic performance.

On the other hand, Harari, Vukovic, and Bailey (2013) did a study to explore the nature of mathematics anxiety in a sample of 106 first grade pupils. The result of their study showed that mathematics anxiety in first grade is a multidimensional construct encompassing: Negative reactions related specifically to foundational mathematical concepts. Numerical confidence related specifically to computation skill. And worry which was not related to any outcome. The levels of mathematics anxiety did not differ by sex or language background. Again, Karjanto and Yong (2013) studied the level of test anxiety in mathematics subjects among early undergraduate students at the University of Nottingham Malaysia Campus in 2013. The sample was 206 students completed questionnaires on test anxiety exactly before they write their midterm examinations. The level of test anxiety in mathematics was measured using seven Likert questionnaire statements adapted from the Test Anxiety Inventory describing one's emotional feeling before the start of an examination. The result shows that the students who had a lower score expectation were more anxious than those who had a higher score expectation, but they obtained a better score than the expected score. Ashcraft and Kirk (2001) studied the relationships among memory, math anxiety, and performance. They found that those individuals who have high math anxiety demonstrate smaller working memory spans, which mean when they are assigned a computation task their reduced working memory capacity will increase their errors. They suggested more empirical attention to be given to math anxiety. Andrews and Brown (2015), studied the effects of math anxiety on 80 freshmen students at a university in the southeastern United States. They conducted an observational study using pre-existing data from the Freshman Orientation Survey, which contained the 9-item Abbreviated Math Anxiety Scale, and institutional research data. The results suggest that standardized test scores and math anxiety had a moderate, negative relationship.

In another study of Taylor and Fraser (2013), who collected data from 745 high-school students from grades 9–12 in 34 classes from four schools in the Southern California area, used learning environment instrument and an updated Revised Mathematics Anxiety Rating scale.

They found significant associations between anxieties and learning environment, they defined the learning environment as 'the social, psychological, and pedagogical contexts in which learning occurs and which affect student achievement and attitudes'. Teachers also might face mathematics anxiety which may influence what and how they teach in their classrooms. Also, Geist (2015) examine training teachers' attitude toward mathematics and found that math anxiety affects how teachers assess their ability at mathematics. He also found that confident teachers are more likely to use developmentally appropriate methods of teaching mathematics. The study suggested that in-service trainings should take place to add a focus on teachers' confidence in mathematics to enhance the mathematics outcomes for children and their ability and enjoyment of mathematics.

(2)

**RESEARCH METHODOLOGY**

The data collected from the questionnaire were processed using Microsoft Excel and R Studio version 4.30. The obtained data were summarized in four dimensional contingency (Table 1). Below is the four dimensional table with the factors and categories under study.

(3)

**Table 1:** Four dimensional table with the factors and categories

Factors	Categories					
Department of student (D)	AnimalSc	Biochem	CropSc	Microbiology	PoliticalSc	Sociology
Gender (G)	Male	Female				
Mathematics Performance (M)	Fail	Pass	Good	V. Good	Excellent	
Anxiety level in Mathematics (A)	Severe	High	Moderate	Low	Minimal	

(4)

The saturated model, 3 – way interaction model, 2 – way interaction model and independent model are possible for the above four dimensional contingency table. Saturated Model

$$\log(\mu_{ijkl}) = \lambda + \lambda_i^M + \lambda_j^A + \lambda_k^G + \lambda_l^P + \lambda_{ij}^{MA} + \lambda_{ik}^{MG} + \lambda_{il}^{MD} + \lambda_{jk}^{AG} + \lambda_{jl}^{AD} + \lambda_{kl}^{GP} + \lambda_{ijk}^{MAG} + \lambda_{ijl}^{MAD} + \lambda_{ikl}^{MGD} + \lambda_{jkl}^{AGD} + \lambda_{ijkl}^{MAGD},$$

with degree of freedom (df) = 0;

Three - way interaction model

$$\log(\mu_{ijkl}) = \lambda + \lambda_i^M + \lambda_j^A + \lambda_k^G + \lambda_l^P + \lambda_{ij}^{MA} + \lambda_{ik}^{MG} + \lambda_{il}^{MD} + \lambda_{jk}^{AG} + \lambda_{jl}^{AD} + \lambda_{kl}^{GP} + \lambda_{ijk}^{MAG} + \lambda_{ijl}^{MAD} + \lambda_{ikl}^{MGD} + \lambda_{jkl}^{AGD},$$

with degree of freedom (df) = [(I-1)(J-1)(K-1)(L-1)];

Two – way interaction model

$$\log(\mu_{ijkl}) = \lambda + \lambda_i^M + \lambda_j^A + \lambda_k^G + \lambda_l^P + \lambda_{ij}^{MA} + \lambda_{ik}^{MG} + \lambda_{il}^{MD} + \lambda_{jk}^{AG} + \lambda_{jl}^{AD} + \lambda_{kl}^{GP},$$

with degree of freedom (df) = [IJK-IJ-IK-IL-JK-JL-KL+2(I+J+K+L)];

Independence Model

$$\log(\mu_{ijkl}) = \lambda + \lambda_i^M + \lambda_j^A + \lambda_k^G + \lambda_l^P$$

with degree of freedom (df) = [IJKL-I-J-K-L+3]

∀ i = 1, 2, ..., 5.; j = 1, 2, ..., 5.; k = 1, 2.; and l = 1, 2, ... 6.

with the constraint that;

$$\sum_i \lambda_i = \sum_j \lambda_j = \sum_k \lambda_k = \sum_l \lambda_l = \sum_{ij} \lambda_{ij} = \sum_{ik} \lambda_{ik} = \sum_{il} \lambda_{il} = \sum_{jk} \lambda_{jk} = \sum_{jl} \lambda_{jl} = \sum_{kl} \lambda_{kl} = \sum_{ijk} \lambda_{ijk} = \sum_{ijl} \lambda_{ijl} = \sum_{ikl} \lambda_{ikl} = \sum_{jkl} \lambda_{jkl} = \sum_{ijkl} \lambda_{ijkl} = 0$$

where  $\mu_{ijkl} = E(n_{ijkl})$  is the expected cell count,  $\lambda$  is the overall effect,

$\lambda_i^M$  represents the main effect from Mathematics performance being in the i-th row,

$\lambda_j^A$  is the main effect from Anxiety level being in the j-th column,

$\lambda_k^G$  is the main effect from Gender being in the k-th layer,

$\lambda_l^D$  is the main effect from Department being in the l-th level,

$\lambda_{ij}^{MA}$  is the interaction between Mathematics performance in the i-th row and Anxiety level in the j-th column,

$\lambda_{ik}^{MG}$  is the interaction between Mathematics performance in the i-th row and Gender in the k-th layer,

$\lambda_{il}^{MD}$  is the interaction between Mathematics performance in the i-th row and Department in the l-th level,

$\lambda_{jk}^{AG}$  is the interaction between Anxiety level in the j-th column and Gender in the k-th layer,

$\lambda_{jl}^{AD}$  is the interaction between Anxiety level in the j-th column and Department in the l-th level,

$\lambda_{kl}^{GD}$  is the interaction between Gender in the k-th layer and Department in the l-th level,

$\lambda_{ijk}^{MAG}$  is the interaction among Mathematics performance in the i-th row, Anxiety level in the j-th column and Gender in the k-th layer,

$\lambda_{ijl}^{MAD}$  is the interaction among Mathematics performance in the i-th row, Anxiety level in the j-th column and Department in the l-th level,

$\lambda_{ikl}^{MGD}$  is the interaction among Mathematics performance in the i-th row, Gender in the k-th layer and Department in the l-th level,

$\lambda_{jkl}^{AGD}$  is the interaction among Anxiety level in the j-th column, Gender in the k-th layer and Department in the l-th level, and

$\lambda_{ijkl}^{MAGD}$  is the interaction among Mathematics performance in the i-th row, Anxiety level in the j-th column, Gender in the k-th layer and Department in the l-th level.

The above model (1), (2), (3) and (4) can be written respectively as

$$[MAGD],$$

$$[MAG][MGD][AGD][MAD],$$

$$[MA][MG][MD][AG][AD][GD],$$

(5)

And

$$[M][A][G][D].$$

### Model Selection and Goodness of Fit Measure of the Models

In order to find the best model from a set of possible models, some measurements should be considered. Two of the popular information criterion called Akaike Information Criteria (AIC) and Bayesian Information Criterion (BIC) are necessary in finding the best model. They will be used in comparing the model under study, so as to select the best model out of the models under investigation.

The information contained in the following equations below are used to estimate the AIC and BIC (Stone, 1979; Akaike, 1974) of a model.

$$AIC = -2 \ln(L) + 2K \tag{6}$$

$$BIC = -2 \ln(L) + 2 \ln(N)K \tag{7}$$

Where L is the value of the likelihood, N is the number of recorded measurements and k is the number of model parameters. The model with minimal AIC or BIC value is chosen as the best model.

**Null Deviance:** This tells us how well the response variable can be predicted by a model with only an intercept term.

**Residual Deviance:** This tells us how well the response variable can be predicted by a model with p predictor variables. The lower the value of residual deviance the better the model is able to predict the response variables. High residual deviance shows that the model cannot be accepted.

Recall also that the aim of log-linear analysis is to find the smallest model that fits the data under study. The overall goodness-of-fit of a model is assessed by comparing the expected frequencies to the observed cell frequencies for each model. The goodness of fit of the best model will be tested with the Pearson chi-squared test statistic  $\chi^2$ . For four dimensional table the Pearson chi-squared test statistic  $\chi^2$  is given by;

$$\chi^2 = 2 \sum n_{ijkl} \log \left( \frac{n_{ijkl}}{\hat{\mu}_{ijkl}} \right), \tag{8}$$

where  $n_{ijkl}$  is the observed cell counts in the cell (i,j,k,l) and  $\hat{\mu}_{ijkl}$  is the estimated expected counts in the cell (i,j,k,l) for the four dimensional table.

Chi-squared test statistic  $\chi^2$  can also be calculated as;

$$\chi^2 = \text{Null deviance value} - \text{Residual deviance value}$$

with p predictor variables as its degree of freedom.

The chi-square  $\chi^2$  test of independence seek to reject the null hypothesis of no association.

### Research Questions

The study is guided by the following research questions;

- i. What is the best model among the models under investigation?
- ii. Is there any relationship among the four dimensional categorical variables?

### Statement of Hypothesis

**H<sub>0</sub>** : there is no relationship among the four dimensional categorical variables.

vs

**H<sub>a</sub>** : there is relationship among the four dimensional categorical variables.

The research questions were answered using Akaike information criteria (AIC), Bayesian information criteria (BIC) of model selection and Chi-square test statistic, while the decision on the acceptance or rejection of stated hypothesis were made at  $\alpha = 0.05$ .

## RESULTS AND DISCUSSION

### Model comparison and model selection

**Table 2: Log-linear model analysis**

Model	Null df	Null deviance	Residual df	Residual deviance	AIC	BIC
[MAGD]	0	-	0	-	-	-
[MAG][MGD][AGD][MAD]	126	29.92	126	29.92	344.26	384.26
[MA][MG][MD][AG][AD][GD]	121	29.31	121	29.31	336.70	365.45
[M][A][G][D]	117	28.87	117	28.87	327.30	341.68

From (Table 2), the results of analysis of the model (1), (2), (3) and (4), model (4) is the best model among the models under investigation as it has the lowest residual deviance value, minimum AIC and BIC value.

The best model which is the independent can be stated below as;

$$\log(\mu_{ijkl}) = \lambda + \lambda_I^M + \lambda_J^A + \lambda_K^G + \lambda_L^D \quad (6)$$

### Goodness - of - fit test statistic for the best model

**Table 3:** Goodness – of – fittest statistic for the best model

Test statistic	Value	Df	P-value
Pearson Chi-square	2.059	4	0.7249

From (Table 3) above, the Pearson Chi-squared test statistic value is 2.059 with  $df = 4$  and  $p\text{-value} = 0.7249$ . It also necessary to note that the null hypothesis  $H_0$ , is not rejected since the  $p\text{-value} = 0.7249$  is greater than the critical value ( $\alpha = 0.05$ ) looking at table 4 above. Hence, the results of the sample data are not significant. There is not sufficient evidence to conclude that the alternative hypothesis  $H_a$  maybe correct. This implies that the there is no relationship among the variables in the sample data under study, that is to say, that mathematics anxiety level, mathematics performance, gender and department of students are independent of each other.

### Conclusion

The findings from the analysis showed that best model out of the models under investigation is the independent model as it has the minimal AIC and BIC value of 327.30 and 341.68 respectively.

And the model is stated below as;

$$\log(\mu_{ijkl}) = \lambda + \lambda_I^M + \lambda_J^A + \lambda_K^G + \lambda_L^D$$

The model has the least residual deviance value as to compare with other models. The test of hypothesis for study asserts that there is no relationship among the four categorical variables as the null hypothesis was not rejected, meaning that the variables are independent of each other. This finally indicate that mathematics anxiety among university students is independent of their gender, department and performances in mathematics.

### REFERENCES

- Andrews, A., & Brown, J. (2015). The effects of math anxiety. *Education*, 135(3), 362-370.
- Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology*, 130(2), 224-237. <http://dx.doi.org/10.1037/0096-3445.130.2.224>.
- Ashcraft, M. H., Krause, J. A., and Hopko, D. R. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin & Review*, 14(2), 246-247.
- Brzezinska, J. (2012). Hierarchical log-linear models for contingency tables. *FoL Oecon*. 269: 123– 129.
- Christensen, R. (1997). *Log-linear models and logistic regression* (2<sup>nd</sup> ed.). New York, NY: Springer. <https://doi.org/10.1007/b97647>.

- Geist, E. (2015). Math anxiety and the "math gap": How attitudes toward mathematics disadvantages students as early as preschool. *Education*, 135(3), 328-336.
- Harari, R. R., Vukovic, R. K., & Bailey, S. P. (2013). Mathematics anxiety in young children: an exploratory study. *The journal of experimental education*, 81(4), 538-555. <http://dx.doi.org/10.1080/00220973.2012.727888>.
- Howell, D. C. (2009). *Statistical methods for psychology* (7<sup>th</sup> ed.). Belmont, CA: Cengage Learning. 630–655.
- Indika, R.W. (2019). Construction and Application of Log-linear Models to Access Academic Performances. *An International Journal of Applied Mathematics*.
- Karjanto, N., & Yong, S. T. (2013). Test anxiety in mathematics among early undergraduate students in a British university in Malaysia. *European Journal of Engineering Education*, 38(1), 11-37. <http://dx.doi.org/10.1080/03043797.2012.742867>.
- Mustofa, U., Widiarti, Rudi, R., Faiz, E., and Jamal, I.D. (2015). Higher dimensional log-linear model and its application. *Sci. Int (Lahore)*, 27(5), 3989-3993.
- Mutawah, M. (2015). The influence of mathematics anxiety in middle & high school students' math achievement. *International Education Studies*, 8(11), 240. doi:10.5539/ies.v8n11p239.
- Okoli, C.N., Onyeagu, S.I., and Osuji, G.A. (2014). Information Criteria and Log-linear Models Selection for Contingency Tables. *Journal of Natural Science Research*. [www.iiste.org](http://www.iiste.org).
- Olmus H.(2012). Analysis of traffic accidents caused by drivers using log-linear models 504 *Promet – Traffic & Transportation*, vol 24, No 6, 495-504.
- Onder, M. and Adiguzel, E.(2010) Evaluation of occupational facilities among underground Coal mine workers through hierarchical log-linear models. *Industrial health* 48, 872- 878.
- Richardson, F. C., & Suinn, R. M. (1972). The mathematics anxiety rating scale: psychometric data. *Journal of Counseling Psychology*, 19(6), 551.
- Taylor, B. A., & Fraser, B. J. (2013). Relationships between learning environment and mathematics anxiety, *Learning Environ Res*, 16, 297-313. <http://dx.doi.org/10.1007/s10984-013-9134-x>.
- Zakaira, E., Zain, N. M., Ahmad, N. A., & Erlina, A. (2012). Mathematics anxiety and achievement among secondary school students. *American Journal of Applied Sciences*, 9(11), 1828.