Application of Three-Dimensional Log Linear Models in Analyzing Risk Factors for a History of Gastritis

Wita Resfi Ananta, Eujeniatul Jannah, Siti Nurhaliza, Yenni Kurniawati*, and Syafriandi Syafriandi
Departement of Statistics, Faculty Mathematich and Science, Padang State University, Padang, Indonesia
*Corresponding author: yennikurniawati@fmipa.unp.ac.id

ABSTRACT

Gastritis, or commonly known as an ulcer, is an inflammatory condition caused by excess stomach acid that irritates the stomach lining. This disease is one of the most common in Indonesia and often disrupts daily activities, especially among students who face academic pressure, stress, and irregular diet. Based on Indonesia's Health Profile Data, gastritis ranks sixth for inpatients with 330,580 cases, 60.86% of which occur in women, and seventh for outpatients with 201,083 cases, of which 77.74% occur in women. This study aims to examine the relationship between gastritis and demographic factors using a three-dimensional log-linear model. The method analyzes interactions between categorical variables to identify the best explanatory model. Results indicate that the most appropriate model involves the interaction between place of residence, gender, and history of stomach ulcers, showing that these factors collectively influence gastritis incidence. In conclusion, gastritis is not only related to physical health but also lifestyle and demographic factors. This study underlines the importance for students to manage stress, maintain healthy eating habits, and adopt preventive measures. The urgency of this research lies in raising awareness that untreated gastritis may reduce productivity and lead to more serious health problems.

Keywords: *Gastritis, Log Linear Model, Student, Three Dimensional*

I. INTRODUCTION

Gastritis, or what is better known as ulcers, is an inflammation of the stomach caused by excess stomach acid which irritates the stomach wall (Vanchapo, 2022). This condition can arise due to various factors, such as long-term use of painkillers, irregular eating patterns, alcohol consumption, smoking, and physical and psychological stress (WHO, 2020). Other contributing factors include autoimmune disorders, Crohn's disease, bile reflux, Helicobacter pylori bacterial infection, and other diseases such as HIV/AIDS, parasitic infections, and liver or kidney disorders (Jusuf et al., 2022).

According to the Indonesian Health Profile Data (Kemenkes RI, 2021), gastritis is among the ten most common diseases in hospitals. Among inpatients, gastritis ranked sixth with 330,580 cases, (60.86%) of which were experienced by women. Meanwhile, in outpatients, gastritis is ranked seventh with 201,083 cases, with the majority of sufferers (77.74%) also being women.

Alya et al., (2025) explain that gastritis is a disease that can interfere witth daily activities and has the potential to cause serious complications if not treated properly. Irregular eating patterns, consumption of foods that stimulate stomach acid production, and microbial infections are major risk factors (Zainurridha, 2019). Symptoms of gastritis include pain in the pit of the stomach, nausea, vomiting, weakness, decreased appetite, paleness, cold sweats, frequent belching, and vomiting blood in severe cases. According to Saraswati et al., (2022), these symptoms affect the quality of life of sufferers, including students.

Anika et al., (2024) mention that for students, a history of stomach ulcers can be a challenge in carrying out daily activities, especially in managing time, coping with stress, and maintaining a healthy diet. Students who live alone or in boarding houses generally have busy and irregular schedules, so they often skip meals or eat at inconsistent intervals. This type of diet can increase stomach acid production and cause irritation to the stomach lining (Sari et al., 2018). In addition, students often get caught up in unhealthy lifestyles such as lack of sleep, physical activity, smoking, or alcohol consumption. Academic pressure, assignments, exams, and the

implementation of the independent curriculum that requires students to be more independent through interdisciplinary projects can also increase stress levels (Suralaga et al., 2025). This condition was experienced by Statistics students from Padang State University, in the class of 2021. High stress, according to Sihotang et al., (2025), can trigger excess stomach acid production and disrupt stomach function.

Tsagris(2021) explains that research, many situations are found where the collected data can be categorized into in research, the collected data is often categorical and can be presented in the form of a contingency table. Contingency tables are used to summarize the combined frequencies between variable categories, while also measuring the existence of relationships or associations (Maryana, 2013). The advantages include simpler calculation preparation, easy to understand analysis results, and the ability to present complex situations more systematically. According to Yesiltas & Paek (2020), one of the statistical methods that can be used to analyse data in a contingency table is the log-linear model.

According to Anderson et al., (2021), the log-linear model is a special case of Generalized Linear Models (GLM) with a Poisson distribution that is used to analyze relationships between categorical variables. This analysis can evaluate the relationship patterns between two or more variables, and identify significant interactions. The log-linear method, as highlighted by Aliverti & Dunson (2022), is therefore suitable for use to see the relationship between gender and place of residence of students and history of stomach ulcers.

Based on this, this study aims to determine whether gender and residence of Statistics students in the 2021 intake are associated with a history of stomach ulcers, while also testing the accuracy of the classification so that it can provide relevant information regarding the factors causing gastritis in students.

II. METHODS

The object of this research is active students of the Statistics Department, class of 2021, Faculty of Mathematics and Natural Sciences, Padang State University. The population consists of 142 students, and because the number is relatively small, the total sampling technique was used so that the entire population became the research sample. Data was obtained from primary sources through an online population became the research sample. Data was obtained from primary sources through an online questionnaire distributed using Google Forms. The questionnaire contained questions regarding the student's residence and history of gastritis diagnosis. The research instrument used a nominal scale and was validated through content validity to ensure the suitability of the questions to the research objectives. The collected data were then analyzed using a three-dimensional log-linear model to assess the interaction between residence, gender, and the incidence of gastritis. The best model is determined based on significant interactions between observed variables.

(a) Data and Variable
This study used data from 142 students. The variables used consisted of:
1. Dependent variable: History of gastritis.
2. Independent variables: Gender and Place of residence.

These three variables are nominal scale with the following classification:
1. Gender (A): 1 (female), 2 (male).
2. Residence (B): 1 (living alone), 2 (living with parents).
3. History of gastritis (C): 1 (yes), 2 (no).
    These variables were selected because they are suspected to have a relationship in explaining the incidence of gastritis among stud

(b) The Steps of Analysis
The analysis was conducted using a three-dimensional log-linear model to assess interactions between categorical variables (Agresti, 2019). The analysis steps were:
1. Data Description
    At this stage, the data obtained from the questionnaire is described. The data comes from 142 active Statistics Department students, class of 2021. The variables used consist of three nominal variables: gender, residence, and history of gastritis.
2. Contingency Table
    Data for the three categorical variables is presented in a three-dimensional contingency table. This step aims to examine the frequency distribution of each combination of variable categories.
3. Model Specifications
    A prespecified log-linear model is used to identify possible interactions between variables. A 3-dimensional

log-linear model can be written in the general form:

$$log_{(\mu ijk)} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC} \qquad (1)$$

Description:
μijk : expected value of cell (i, j, k)
λ : average parameters
A : Gender (1 = Female, 2 = Male)
B : Residence (1 = Living alone/boarding house, 2 = Living with parents)
C : History of gastritis (1 = Yes, 2 = No)
$\lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$: two-way interactions
λ_{ijk}^{ABC} : three-way interaction among Gender, Residence, and Gastritis

4. Modeling with Log-Linear
At this stage, analysis was performed using a log-linear model to identify relationships between categorical variables, namely gender, place of residence, and history of gastritis. The modeling process begins with a saturated model, which includes all interactions between variables (main effect, two-way interaction, and three-way interaction). The initial model selection aims to ensure that all possible relationships between variables are taken into account. Next, the model will be simplified gradually by comparing candidate models using the Likelihood Ratio Test (LRT).

5. Goodness of Fit
Model suitability testing was conducted using (G^2) statistics (Likelihood Ratio Chi-Square) and Pearson Chi-Square. A p-value greater than 0.05 indicates that there is no significant difference between the observed frequency and the expected frequency, so the model is considered to fit the data. Conversely, if the p-value is less than 0.05, the model does not fit and adjustments need to be made while maintaining significant interactions.

$$G^2 = 2 \sum O_{ijk} \ln\left(\frac{O_{ijk}}{E_{ijk}}\right) \qquad (2)$$

Description:
O_{ijk} : frequency of observations
E_{ijk} : expected frequency of the model

6. Best Model Interpretation
After testing, the best model is determined based on the combination of significant interactions. If the three-way interaction is significant, then the saturated model will be the best model, which means that there is a reciprocal relationship between the three variables (gender, place of residence, and history of gastritis). However, if the three-way interaction is not significant, then the best model may be a model that only includes the main effect or relevant two-way interactions. Thus, the interpretation of the best model provides an overview of the pattern of interrelationships between the research variables.

7. Conclusion
This stage aims to summarize the modeling results by emphasizing whether there is a significant relationship between gender, place of residence, and history of gastritis. Conclusions are drawn based on the best model obtained in the previous stage, so that it can answer the research objective, which is to explore the relationship between categorical variables using a log-linear model approach.

III. RESULT AND DISCUSSION
Descriptive statistics of this data are presented in Table 1:
Table 1. Descriptive Statistics of Data

No	Characteristics	Category	Frequency
1	Gender	Male	29
		Famale	113
2	Living alone	Yes	92
		No	50
3	History of gastritis	Yes	64
		No	78

Table 1 provides information that female respondents had the highest proportion of 113 people compared to

male respondents of 29 people. The respondents with the largest proportion living alone (boarding) were in the "yes" category, namely 92 people compared to the "no" category with 50 people. A total of 78 people answered yes to the characteristics of stomach ulcers and 64 people answered no.

Table 2. Three-way contingency

Gender	Living alone	History of gastritis	
		Yes	No
Female	Yes	30	44
	No	26	13
Male	Yes	6	12
	No	2	9

Based on Table 2, the highest proportion of respondents were female students living in boarding houses without gastritis, totaling 44 people (44.31%). Meanwhile, the lowest proportion was male students not living in boarding houses with gastritis, totaling only 2 people (2.02%). This indicates that the distribution of gastritis incidence varies across gender and residential status, with female boarding students showing the highest frequency.

Table 3. Convergence Information

Generating Class	Gender*boarding house*gastritis
Number of Iterations	1
Max. Difference between Observed and Fitted Marginals	.000
Convergence Criterion	1.936

Based on the Generating Class output, it was found that the interactions formed were between the variables of Gender, Place of Residence (boarding house), and History of Stomach Ulcers. This indicates that the appropriate model is a saturated model, because in a three-dimensional log-linear model, the existence of three-way interactions indicates that all main effects, two-way interactions, and three-way interactions are included in the model. To prove this, a goodness-of-fit test was conducted using chi-square statistics. The test results show a significance value of p-value = 0.000 (< 0.05), so the simple model is rejected and it is decided that the saturated model is the best model to explain the relationship between variables. Thus, it can be concluded that the relationship between gender, place of residence, and history of stomach ulcers is interdependent, so that these three variables cannot be analyzed separately but must be considered together in explaining the variation in the incidence of gastritis among students.

Hypothesis:
H_0 : There was no association between gender, place of residence (boarding house), and history of gastric ulcer.
H_1 : There is a relationship between gender, place of residence (boarding house), and history of stomach ulcers.
Test Statistics:
To test whether a simple model (independent or two-way interaction) is sufficient or whether a saturated model is required, use equation (2.2).
Significance level (∝): 5% = 0.05
Real level $\alpha = 0{,}05$
Reject H_0 if p-value < 0,05
Accept H_1 if p-value ≥ 0,05

Table 4. Goodness-of-fit-test

	Value	Df	Sig.
Likelihood Ratio	.000	0	.000
Pearson Chi-Square	.000	0	.000

Based on Table 4, the likelihood ratio p-value is 0.000, which is smaller than the α = 0.05. The P-value obtained was < 0.05, so the information obtained was to reject H_0, so it can be said that there is a relationship between gender, place of residence (boarding house), and history of gastritis. If the decision is to reject H₀, then

we can proceed to create the best model using the 3-dimensional log-linear method.

Table 5. Choosing the best model

Model	Saturated	Complete Independence	Joint Independence	Conditional Independence	Homogeneous Association
Value AIC	51.04989	55.72358	54.00163	55.88443	52.82785

Table 5 shows the AIC values for each model tested. A lower AIC value is considered better, so from the table above, the lowest value is the saturated model with an AIC value of 51.05.

Table 6. Parameter Estimates

Parameter	Estimate	Std. Error	Z	Sig.	95%CI Lower Bound	95%CI Upper Bound
Constant	2.251	.324	6.936	.000	1.615	2.887
[Gender = 1]	.351	.423	.830	.407	-.479	1.181
[Gender = 2]	0ª	.	.	.	.	.
[Boarding house = 1]	.274	.430	.638	.524	-.569	1.118
[Boarding house = 2]	0ª	.	.	.	.	.
[gastric pains = 1]	-1.335	.711	-1.878	.060	-2.728	.058
[gastric pains = 2]	0ª	.	.	.	.	.
[Gender = 1] * [Boarding house = 1] * [gastric pains = 1]	1.876	1.004	1.868	.062	-.092	3.843
[Gender = 1] * [Boarding house = 1] * [gastric pains = 2]	.918	.531	1.730	.084	-.122	1.959
[Gender = 1] * [Boarding house = 2] * [gastric pains = 1]	2.009	.786	2.558	.011	.470	3.549
[Gender = 1] * [Boarding house = 2] * [gastric pains = 2]	0ª	.	.	.	.	.
[Gender = 2] * [Boarding house = 1] * [gastric pains =1]	.681	.860	.792	.428	-1.004	2.366
[Gender = 2] * [Boarding house = 1] * [gastric pains =2]	0ª	.	.	.	.	.
[Gender = 2] * [Boarding house = 2] * [gastric pains =1]	0ª	.	.	.	.	.
[Gender = 2] * [Boarding house = 2] * [gastric pains =2]	0ª	.	.	.	.	.

Based on the analysis results in Table 6, it can be seen that significant interactions are indicated by parameters with p-values < 0.05. Although not all interactions meet the significance criteria at the 5% level, there are several variable combinations that show strong trends.

parameter interpretation:

The parameter values λ in the log-linear model cannot be directly summed, but rather interpreted through exponential form (e^λ) so that the odds ratio (OR) value is obtained. Here is the interpretation of some important parameters:

1. Interaction of Gender, Boarding House, and Gastric Ulcers ($\lambda_{111}^{ABC} = 1,876$) .The oods ratio is a $e^{1,876} = 6,53$. This means that women who live in boarding houses and have a history of stomach ulcers have a 6.53 times greater chance of developing stomach ulcers than other combinations. This shows that the boarding house factor can worsen the risk of ulcers, especially in female respondents.

2. Interaction Between Gender, Boarding House, and Not Having Stomach Ulcers ($\lambda_{112}^{ABC} = 0,918$). The oods ratio is $e^{0,918} = 2,50$. This means that women who live in boarding houses but do not have a history of stomach ulcers have a tendency 2.50 times higher than other combinations. This shows that even though boarding houses

are risky, there are still some female boarding house respondents who do not experience ulcers.

3. Interaction Between Gender, Non-Boarding, and Stomach Ulcers ($\lambda_{121}^{ABC} = 2,009$). The oods ratio is $e^{2,009} = 7,46$. This means that women who live with their parents/family but have a history of ulcers have a 7.46 times higher chance than other combinations. This indicates that the ulcer factor is not solely influenced by the boarding environment, but can also occur in female respondents at home.

4. Interaction between Men, Boarding Houses, and Stomach Ulcers ($\lambda_{211}^{ABC} = 0,681$). The oods ratio is $e^{0,681} = 1,98$. This means that men who live in boarding houses and suffer from stomach ulcers have a risk almost twice that of other combinations. However, this effect is relatively lower compared to female respondents.

These results show that gender and place of residence (boarding house vs. non-boarding house) play an important role in explaining the occurrence of stomach ulcers. Women, especially those who live in boarding houses, have a higher tendency to experience stomach ulcers.

In addition, although single variables (gender or boarding house) were not statistically significant, three-dimensional interactions between variables were more informative in explaining variations in the incidence of gastric ulcers. This finding confirms that the use of the log-linear model is appropriate for multivariate categorical data, because it is able to capture the influence of interactions between factors.

IV. CONCLUSION

Based on the results of log-linear analysis, it was found that the model that best fits the data is the saturated model. This indicates a relationship or dependence between the variables of gender, place of residence, and history of stomach ulcers in influencing the incidence of gastritis in students. The interaction between biological factors (gender and history of stomach ulcers) and environmental factors (living in a boarding house or not) plays an important role in increasing the risk of gastritis.

These findings confirm that students, especially women who live in dormitories with a history of gastric ulcers, have a higher tendency to experience recurrence than other groups. This condition shows that environmental and social factors, such as where students live, can affect their health, especially in relation to the risk of gastritis.

This study shows that students need to pay more attention to a healthy lifestyle, including regular eating patterns, stress management, and routine health checkups, especially for those who live far from home. For further research, it is recommended to develop this model by adding other variables, such as eating patterns, stress levels, or physical activity, in order to obtain more comprehensive and useful results in efforts to prevent gastritis among students.

ACKNOWLEDGMENT

REFERENCES

Aliverti, E., & Dunson, D. B. (2022). Composite Mixture of Log-Linear Models With Application To Psychiatric Studies. *Annals of Applied Statistics*, *16*(2), 765–790. https://doi.org/10.1214/21-AOAS1515

Alya, M., Haryanik, P., & Istiaji, E. (2025). Occurrence of Gastritis Is Related to Stress among Health Science Students. *Althea Medical Journal*, *12*(1), 14–21. https://doi.org/10.15850/amj.v12n1.3421

Anderson, C. J., Kateri, M., & Moustaki, I. (2021). *Log-Linear and Log-Multiplicative Association Models for Categorical Data*. *January*, 1–55.

Anika, M., Neri, E. L., & Ligita, T. (2024). Hubungan Tingkat Pengetahuan Gastritis Dengan Kejadian Gastritis Pada Mahasiswa Mahasiswi Program Studi Keperawatan Fk Untan. *Tanjungpura Journal of Nursing Practice and Education*, *6*(2), 28–34.

Jusuf, H., Adityaningrum, A., Yunus, R., Olahraga, F., Kesehatan, D., & Gorontalo, U. N. (2022). Determinan Kejadian Gastritis Pada Mahasiswa the Determinants of Gastritis Among Students. *Jambura Health and Sport Journal*, *4*(2), 108–118.

Kemenkes RI. (2021). Profil Kesehatan Indonesia 2021. In *Pusdatin.Kemenkes.Go.Id*.

Maryana. (2013). Model Log Linier yang Terbaik untuk Analisis Data Kualitatif pada Tabel Kontingensi Tiga Arah. *Malikussaleh Indutrial Engineering*, *2*(2), 32–37. https://www.semanticscholar.org/paper/Model-Log-Linier-yang-Terbaik-untuk-Analisis-Data-Maryana/9aabe47cf7b21a3a58caaab33c972271e85779d8

Saraswati, P. A., Suyasa, I. gede P. darma, & Wulandari, I. A. (2022). Hubungan Tingkat Stres Dengan Gejala GastritisMahasiswa Tingkat IV. *Jurnal Gema Keperawatan*, *15*, 207–216. https://ejournal.poltekkes-denpasar.ac.id/index.php/JGK/article/view/2021

Sari, R. D. T., Anggeny, Y., & Lita. (2018). Hubungan Stres Akademik dan Pola Makan dengan Kejadian Gastritis. *Jurnal Keperawatan STIkes Hang Tuah Tanjungpinang*, *14*(2), 10–16.

Sihotang, W., Hura, F. W., Karo – Karo, S., Nasution, M., Budi, A., Nadapdap, M. J., & Simanjuntak, R. S. (2025). *Factors Associated with Gastritis Symptoms Among College Students*. *Icolifemed 2024*, 35–43. https://doi.org/10.2991/978-94-6463-664-2_5

Suralaga, C., Merrin, & Usman, A. M. (2025). Unhealthy Mind, Unhealthy Gut: The Impact of Stress and Eating Patterns on Gastritis in Early Adolescents in Depok, Indonesia. *Malaysian Journal of Medical Research*, *09*(03), 37–44. https://doi.org/10.31674/mjmr.2025.v09i03.005

Tsagris, M. (2021). Conditional Independence Test for Categorical Data Using Poisson Log-Linear Model. *Journal of Data Science*, *15*(3), 345–354. https://doi.org/10.6339/jds.201707_15(3).0002

Vanchapo, A. R. (2022). *Pendidikan Kesehatan dalam Meningkatkan Pengetahuan Tentang Gastritis* (Issue October).

WHO. (2020). World Health Statistics. In *Block Caving – A Viable Alternative?* (Vol. 21, Issue 1). https://www.golder.com/insights/block-caving-a-viable-alternative/

Yesiltas, G., & Paek, I. (2020). A Log-Linear Modeling Approach for Differential Item Functioning Detection in Polytomously Scored Items. *Educational and Psychological Measurement*, *80*(1), 145–162. https://doi.org/10.1177/0013164419853000

Zainurridha, A. Y. (2019). Stres Dan Pola Makan Terhadap Kejadian Gastritis. *Jurnal Keperawatan Dan Kebidanan*, *4*(1), 15–22. http://jurnal.stikesalqodiri.ac.id/index.php/Jurnal_STIKESAlQodiri