WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Analysis of Relation between Mental Health, Stress Level and Gender using Log-Linear Model

Rita Rahmawati[1], Yuciana Wilandari[2*], Triastuti Wuryandari[3]

[1,2,3]Departemen of Statistics, Faculty of Science and Mathematics, Diponegoro University, Semarang, 50275, Indonesia
**yuciana.wilandari@gmail.com*

A B S T R A C T

Mental health is a crucial dimension of individual well-being, influencing how a person thinks, feels, and acts. Stress has become a universal and unavoidable phenomenon, often a primary trigger for mental disorders. This study sought to explore in depth how stress dynamics interact with a person's gender to influence mental health status, specifically whether they experience a disorder, and whether the two are interrelated, using a log-linear model. The log-linear model is a model used to analyze the relationship between categorical response variables that form a contingency table and determine which variables tend to cause dependency. The results of this study, after conducting the K-way effect test and higher-order effects, the K-way effect test, and the partial association test, obtained the best log-linear model, namely the interaction of two variables, mental health and gender, as well as mental health and stress levels. The results also showed that of the variables mental health, stress levels, and gender, the dependent variable is mental health. Meanwhile, the goodness of fit is Likelihood Ratio of 7.549 and a Pearson Chi-Square of 7.157.

Keywords: Log-Linier Model; Mental Health, Stress Levels; Gender

1. Introduction

Mental health is an important dimension of individual well-being that affects the way a person thinks, feels, and acts. In modern society, stress has become a universal and inevitable phenomenon, which is often the main trigger for mental disorders. This study seeks to explore in depth how the dynamics of stress interact with an individual's gender in influencing a person's mental health status, specifically whether they have a disorder or not. In general, individuals with high levels of stress are more suspectible to a various of mental health conditions, such as depression and anxiety. However, empirical evidence suggests that there are differences in coping patterns, symptom manifestations, and even the prevalence of mental disorders between men and women. Biological, psychological, and sociocultural differences inherent in gender are thought to play a role in moderating or exacerbating the impact of stress on mental health. To determine whether or not there is a relationship between mental health, stress levels and gender, the log-linear model can be used. The log-linear model is a model used to analyze the relationship between the categorical response variables that form a contingency table and determine which variables tend to cause dependency.

One of the studies on mental health was conducted by Yulianti and Andika (2024), namely researching the factors that affect mental health are individual factors, family factors, environmental factors, and social factors using logistic regression. Another study on the relationship between gender and stress categories using the Chi-square test was conducted by Wilujeng et al. (2023). Meanwhile, research on the log-linear models has been conducted by Guntur (2012) and Sihotang (2020). Guntur (2012) researched a two-dimensional log-linear model based on a multinomial sampling design. Meanwhile, Sihotang (2023) researched a three-dimensional log-linear model using the forward method. Therefore, there has been no research on the relationship between mental health, stress levels and gender using the log-linear model. Therefore, this study aims to analyze how the relationship between mental health, stress levels and gender uses the log-linear model.

2. Literature Review

The log-linear models are used to analyze the relationships between categorical response variables that form the contingency table and determine which variables tend to cause dependency. For example, if a variable X consists of I categories, variable Y consists of J categories and variable Z consists of K categories, then the saturated log-linear model is:

$$\log m_{ijk} = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ} + \lambda_{ijk}^{XYZ} \quad (1)$$

with m_{ijk} : ecpected count in the cell-ijk

μ : mean

λ_i^X : main effect for variable X

λ_j^Y : main effect for variable Y

λ_k^Z : main effect for variable Z

λ_{ij}^{XY} : interaction effects for variables X and Y

λ_{ik}^{XZ} : interaction effects for variables X dan Z

λ_{jk}^{YZ} : interaction effects for variables Y dan Z

λ_{ijk}^{XYZ} : interaction effect for variables X, Y dan Z

According to Agresti (1990), there are several possible models in three-dimensional log-linear models. Some of these models are mentioned in Table 1.

Table 1. Multiple Log-linear Models for Three Dimensions

Log-linear Model	Symbol
$\log m_{ijk} = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z$	(X, Y, Z)
$\log m_{ijk} = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY}$	(XY, Z)
$\log m_{ijk} = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{jk}^{YZ}$	(XY, YZ)
$\log m_{ijk} = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ}$	(XY, YZ, XZ)
$\log m_{ijk} = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ} + \lambda_{ijk}^{XYZ}$	(XYZ)

To determine the best model of the log-linear model, model selection performed by performing the following tests,

1. The K-way test for K-term or more interactions is zero

For k = 3 H_0: the effect of 3 interactions is equals to zero

H_1 : the effect of 3 interactions is not equal to zero

Test statistics : Pearson Chi-Square (χ^2) or LR Chi-Square (G^2)

The test criterion reject H_0 if $\chi^2 \geq \chi^2_{(\alpha,df)}$ or $G^2 \geq \chi^2_{(\alpha,df)}$ or Prob (sig) $\leq \alpha$

For k = 2 H_0 : the effect of 2 or more interactions is equals to zero

H_1 : at least one effect of 2 or more interactions is not equal to zero

Test statistics : Pearson Chi-Square (χ^2) or LR Chi-Square (G^2)

The test criterion reject H_0 if $\chi^2 \geq \chi^2_{(\alpha,df)}$ or $G^2 \geq \chi^2_{(\alpha,df)}$ or Prob (sig) $\leq \alpha$

For k = 1 H_0 : one or more are equals to zero

H_1 : at least one effect is not equal to zero

Test statistics : Pearson Chi-Square (χ^2) or LR Chi-Square (G^2)

The test criterion reject H_0 if $\chi^2 \geq \chi^2_{(\alpha,df)}$ or $G^2 \geq \chi^2_{(\alpha,df)}$ or Prob (sig) $\leq \alpha$

2. The K-way test for K-term interaction is zero

For k = 3 H_0 : the effect of 3 interactions are equals to zero

H_1 : the effect of 3 interactions is not equal to zero

Test statistics : Pearson Chi-Square (χ^2) or LR Chi-Square (G^2)

The test criterion reject H_0 if $\chi^2 \geq \chi^2_{(\alpha,df)}$ or $G^2 \geq \chi^2_{(\alpha,df)}$ or Prob (sig) $\leq \alpha$

For k = 2 H_0 : the effect of 2 interactions is zero

H_1 : at least one effect of 2 interactions is not zero

Test statistics : Pearson Chi-Square (χ^2) or LR Chi-Square (G^2)

Test criteria : H_0 is reject if $\chi^2 \geq \chi^2_{(\alpha,df)}$ or $G^2 \geq \chi^2_{(\alpha,df)}$ or Prob (sig) $\leq \alpha$

For k = 1 H_0 : the effect of 1 is zero

H_1 : at least one effect 1 is not zero

Test statistics : Pearson Chi-Square (χ^2) or LR Chi-Square (G^2)

The test criterion reject H_0 if $\chi^2 \geq \chi^2_{(\alpha,df)}$ or $G^2 \geq \chi^2_{(\alpha,df)}$ or Prob (sig) $\leq \alpha$

with Likelihood Ratio Test, $G^2 = 2\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} n_{ijk}\log(\frac{n_{ijk}}{\widehat{m}_{ijk}})$ (2)

Pearson Chi Square Test, $\chi^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\frac{(n_{ijk}-\widehat{m}_{ijk})^2}{\widehat{m}_{ijk}}$ (3)

3. Partial Association Test

The Partial Association Test aims to examine all possible parameters of a complete model both for a single independent variable and for the dependent relationship of several variables that are partial of a complete (*saturated*) model (Christensen, 1997).

Hypothesis H_0 : There is no interaction between the variables

H_1 : Interaction between the variables are contained in the model

The test statistics used are *the Partial Chi Square* value or significance value.

Test criteria : H_0 is rejected if the value *Partial Chi Square* $\geq \chi^2_{(\alpha,df)}$ or Prob (sig) $\leq \alpha$

After the model selection is carried out and the best model has been obtained, an estimate of the model parameters can be carried out according to the model that has been obtained. For example, n_{ijk} and m_{ijk} are the observation data and expected values of each cell in contingency tables and log-linear models generally defined:

$\log m_{ijk} = \boldsymbol{X}^T\boldsymbol{\beta}$

with $\boldsymbol{\beta} = \left(\mu, \lambda_i^X, \lambda_j^Y, \lambda_k^Z, \lambda_{ij}^{XY}, \lambda_{ik}^{XZ}, \lambda_{jk}^{YZ}, \lambda_{ijk}^{XYZ}\right)$

Assuming a Poisson distribution, then the likelihood function is:

$$l(m) = \prod_i\prod_j\prod_k \frac{e^{-m_{ijk}}{m_{ijk}}^{n_{ijk}}}{n_{ijk}!}$$

so the log likelihood is:

$$L(m) = \sum_i\sum_j\sum_k n_{ijk}\log(m_{ijk}) - \sum_i\sum_j\sum_k m_{ijk}$$

The first partial derivatives in the form of a matrix are: $\boldsymbol{X}^T(\boldsymbol{n}-\widehat{\boldsymbol{m}})$

The second partial derivative are : $\boldsymbol{I}(\boldsymbol{\beta}) = \boldsymbol{X}^T\boldsymbol{Diag}(\widehat{\boldsymbol{m}})\boldsymbol{X}$

Using the Newton-Raphson iteration method

1. An initial estimate for $\widehat{\boldsymbol{\beta}}^{(0)}$ is chosen, suppose $\widehat{\boldsymbol{\beta}}^{(0)} = \boldsymbol{0}$
2. On each $t+1$, a new estimate is calculated:

$$\widehat{\boldsymbol{\beta}}^{(t+1)} = \widehat{\boldsymbol{\beta}}^{(t)} + \boldsymbol{I}(\widehat{\boldsymbol{\beta}})^{-1}(\boldsymbol{X}^T(\boldsymbol{n}-\widehat{\boldsymbol{m}}))$$

3. The iteration ends when $\widehat{\boldsymbol{\beta}}^{(t+1)} \approx \widehat{\boldsymbol{\beta}}^{(t)}$

After obtaining the model parameter estimate, the next step is test the model's significance. The Chi-square statistics are used to test the hypothesis that the expected population frequencies conform to a particular model, using the Likelihood Ratio Test (G^2) or the Pearson Chi Square (χ^2) test (Agresti, 1990).

Hypothesis: H_0 : The model is fit

H_1 : The model is not fit

Test Statistics: Likelihood Ratio Test, $G^2 = 2\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} n_{ijk}\log(\frac{n_{ijk}}{\widehat{m}_{ijk}})$ (4)

Pearson Chi Square Test, $\chi^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\frac{(n_{ijk}-\widehat{m}_{ijk})^2}{\widehat{m}_{ijk}}$ (5)

Test criteria: H_0 is reject if $G^2 \geq \chi^2_{(\alpha,df)}$ or sig value $< \alpha$, with a value df given as in Table 2 (Agresti, 1990)

Table 2. df value for 3-Dimensional Log-linear Model

Type	*Degrees of Freedom*
(X,Y,Z)	IJK-1-J-K+2
(XY,Z)	(K-1)(IJ-1)
(XZ,Y)	(J-1)(IK-1)
(YZ. X)	(I-1)(JK-1)
(XY,YZ)	J(I-1)(K-1)
(XZ,YZ)	K(I-1)(J-1)
(XY,XZ)	I(J-1)(K-1)
(XY,XZ,YZ)	(I-1)(J-1)(K-1)
(XYZ)	0

3. Research Methodology

The data used in this study is data on the relationship between mental health, stress level and gender. Mental health is categorized into three categories: low, medium and high. The stress level are categorized into two categories: undisturbed and disturbed, and gender is categorized into two categories: male and female. The following steps taken to obtain the best log-linear model,

1. Input the data and create a contingency tables
2. Create a saturated model using Equation (1)
3. Conduct K-way and Higher Order Effects Tests, as well as the K-way Effects Tests
4. Conduct the Partial Association Test
5. After obtaining the best model, determine the best model estimate.
6. Conduct the goodness of fit test using Equation (4) and (5)

4. Results And Discussion

The data used in this study consisted of three variables: Mental Health (M) which consist of three categories and their codes: low (1), medium (2) and high (3). The second variable is the stress level (S), which consists of two categories and their codes: undisturbed (1) and disturbed (2). While the third variable is gender (G), which is divided into wto categories and their codes: male (1) and female (2).

Table 3. Contingency Table of Mental Health, Stress Level and Gender

Stress levels (S)	Gender (G)	Mental Health (M)	
		Undisturbed	Disturbed
Low	Male	42	2
	Female	29	9
Medium	Male	30	11
	Female	13	12
High	Male	9	22
	Female	6	15

The first step is to determine the saturated model,

$$\log m_{ijk} = \mu + \lambda_i^M + \lambda_j^S + \lambda_k^G + \lambda_{ij}^{MS} + \lambda_{ik}^{MG} + \lambda_{jk}^{SG} + \lambda_{ijk}^{MSG} \quad (6)$$

Next, the K-Way and Higher-Order Effects Test was conducted, with the following hypothesis,

H_0 : the effect of K or more interaction is equal to zero, for K = 1, 2, 3

H_1 : the effect of K or more interaction is not equal to zero, for K = 1, 2, 3

Test Statistics: Likelihood *Ratio Test*, $G^2 = 2\sum_{i=1}^{3}\sum_{j=1}^{2}\sum_{k=1}^{2} n_{ijk}\log(\frac{n_{ijk}}{\widehat{m}_{ijk}})$

Pearson Chi Square *Test*,$\chi^2 = \sum_{i=1}^{3}\sum_{j=1}^{2}\sum_{k=1}^{2}\frac{\left(n_{ijk-}\widehat{m}_{ijk}\right)^2}{\widehat{m}_{ijk}}$

From the results of processing using SPSS, the results were obtained as seen in Table 4.

Table 4. K-way and Higher Order Effects Test and K-way Effects Test

K-way and Higher Order Effects						
K	Likelihood Ratio		Pearson		df	Decision
	Chi-Square	Sig	Chi-Square	Sig		
1	87.676	0.000	91.000	0.000	11	H_0 rejected
2	58.686	0.000	52.610	0.000	7	H_0 rejected
3	3.555	0.169	3.492	0.174	2	H_0 Failed to Reject
K-way Effects						
K	Likelihood Ratio		Pearson		df	Decision
	Chi-Square	Sig	Chi-Square	Sig		
1	28.990	0.000	38.390	0.000	4	H_0 rejected
2	55.132	0.000	49.118	0.000	5	H_0 rejected
3	3.555	0.169	3.492	0.174	2	H_0 Failed to Reject

Using α = 0,05, Table 4 shows that in the K-way and higher order effects test, there are no three interactions effects equal to zero, but there are two or more interactions effects that are not equal to zero. Meanwhile, the K-way effect test shows no three interactions effect, but two interactions effect are not equal to zero. To determine which variables interacted with each other, the partial association test can be used.

Hypothesis: H_0 : There is no interaction between the variables

H_1 : The interaction between the variables contained in the model

The test statistics used are *Partial Chi Square* (χ^2) or Prob (sig) values

Test criteria : H_0 is rejected if the value *Partial Chi Square* $\geq \chi^2_{(\alpha,df)}$ or Prob (sig) $\leq \alpha$

From the results of processing using SPSS, the results are as shown in Table 5.

Table 5. Partial Association Test

Effect	Partial Chi-Square	df	$\chi^2_{(0{,}05,df)}$	Sig.	Decision
level_stress*mental	50.573	2	5.991	0.000	H_0 rejected
level_stress*gender	3.995	2	5.991	0.136	H_0 Failed to Reject
mental*gender	6.253	1	3.841	0.012	H_0 rejected
level_stress	6.784	2	5.991	0.034	H_0 rejected
Mental	17.064	1	3.841	0.000	H_0 rejected
Gender	5.142	1	3.841	0.023	H_0 rejected

The partial association test shown that the only variables that do not interact are stress and gender. Therefore there are only two interactions of two variables: stress*mental and mental*gender. Therefore, the best model is a model with two interactions of two variables: stress*mental and mental*gender. From the output of the model selection, the best models are:

$$\log m_{ijk} = \mu + \lambda_i^M + \lambda_j^S + \lambda_k^G + \lambda_{ij}^{MS} + \lambda_{ik}^{MG} \quad (7)$$

Next, the best model parameter estimates were found. The following is the output in Table 6.

Table 6. Parameter Estimates for The Best Log-Linear Model

Variable	Symbol	Estimation
Constant	μ	2.932
[mental = 1]	λ_1^M	-1.212
[level_stress = 1]	λ_1^S	-1.213
[level_stress = 2]	λ_2^S	-0.475
[gender = 1]	λ_1^G	-0.028
[mental = 1]* [level_stress = 1]	λ_{11}^{MS}	2.768
[mental = 1]*[level_stress = 2]	λ_{12}^{MS}	1.529
[mental = 1]*[gender = 1]	λ_{11}^{MG}	0.551

The final step is to conduct a goodness of fit test as follows:

Hypothesis: H_0 : Model is fit

H_1 : Models is not fit

Test Statistics: Likelihood Ratio Test, $G^2 = 2\sum_{i=1}^{3}\sum_{j=1}^{2}\sum_{k=1}^{2} n_{ijk}\log(\frac{n_{ijk}}{\hat{m}_{ijk}})$

Pearson Chi Square Test, $\chi^2 = \sum_{i=1}^{3}\sum_{j=1}^{2}\sum_{k=1}^{2}\frac{\left(n_{ijk-}\hat{m}_{ijk}\right)^2}{\hat{m}_{ijk}}$

H_0 is reject if $G^2 \geq \chi^2_{(\alpha,df)}$ or sig value $< \alpha$, with a value df given as in Table 2 (Agresti, 1990).

The SPPS analysis yielded a Likelihood Ratio value of 7.549 with a sig. 0.110, and a Pearson Chi-Square value of 7.157 with a sig. 0.128. Therefore, Ho was rejected. Therefore, it can be concluded that the model is fit.

Next, calculate the expected count for the best model based on the obtained model, using the equation (6), shown in Table 7.

Table 7. Expected Count For The Best Model

Stress levels	Gender	Mental Health	Observed Count	Expected Count
Low	Male	Undisturbed	42	44.581
		Disturbed	2	5.423
	Female	Undisturbed	29	26.419
		Disturbed	9	5.577
Medium	Male	Undisturbed	30	27.000
		Disturbed	11	11.338
	Female	Undisturbed	13	16.000
		Disturbed	12	11.662
High	Male	Undisturbed	9	9.419
		Disturbed	22	18.239
	Female	Undisturbed	6	5.581
		Disturbed	15	18.761

5. Conclusion

Based on the analysis of the relationship between mental health, stress levels, and gender using a log-linear model, and after conducting the K-way and higher order effects test, K-way effect test and partial association test, the best log-linear model was obtained,

$$\log m_{ijk} = \mu + \lambda_i^M + \lambda_j^S + \lambda_k^G + \lambda_{ij}^{MS} + \lambda_{ik}^{MG}$$

Based on this best model, the dependent variable is Mental Health. The goodness of fit test of the best log-linear model showed a goodness of fit, with a Likelihood Ratio of 7.549 and a Pearson Chi-Square of 7.157.

References

Agresti, A, 2007, *An Introduction to Categorical Data Analysis.* John Willey & Sons Inc., New Jersey.

Christensen, R.1997. *Log-Linear Models and Logistic Regression*

Guntur, R.D. 2020. The Log Linear Models for Two Dimensional Contingency Tables Under The Multinomial Sampling Design**.** *Indonesian Journal of Statistics,* Vol. 17 No.2, Oktober 2012, p: 9-14

Hosmer, D.W., and Lemeshow, S, 2000, *Applied Logistic Regression.* John Willey & Sons Inc., New York.

Sihotang, S.F. 2020. Analisis Model Log Linier Tiga Dimensi Untuk Data Kualitatif dengan Metode Forward. *Journal of Mathematics Education and Science*, Vol. 6 No. 1

Wilujeng, C.S., Habibie, Y., and Ventyaningsih, A.D.I. 2023. Hubungan antara Jenis Kelamin dengan Kategori Stres pada Remaja di SMP Brawijaya Smart School. *Smart Society Empowermwnt Journal* 3(1):6

Yulianti, T.S. and Andika, I.P.J. 2024. Analisis Kesehatan Mental Mahasiswa Sekolah Tinggi Ilmu Kesehatan Panti Kosala. *Jurnal Ilmu Kesehatan*, Vo. 12 No. 2