

Analysis of Factors Influencing The Number of Maternal Deaths in East Java Province 2023

M. Rizal Matofani¹, Muhammad Zulfadhli², Nurfadhlan Dhafin Nugroho³, Artika Putri Rahmawati⁴,
Fi Amanila Putri Kinanti⁵

¹Departement of Business Statistics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia

²Departement of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia

Article Info

Article history:

Received July 15, 2026

Revised August 7, 2026

Accepted September 3, 2026

Keywords:

Number of Maternal Deaths
Number of Midwives Health Workers
Number of Public Health Workers
Average School Length
Negative Binomial Regression
Poisson Regression

ABSTRACT

Maternal mortality is still a serious health problem in various regions, including in East Java Province. In 2019–2020, the number of maternal deaths nationally increased, which shows that there are still obstacles in access and quality of health services. Factors such as the number of midwives and public health workers, and the average length of school are suspected to have an effect on this condition. Midwives play a direct role in maternal and child services, while public health workers support promotive and preventive programs. Women's education, seen from the average length of schooling, also affects awareness of the importance of maintaining health during pregnancy. This study uses BPS 2023 data to analyze the influence of these three variables on maternal mortality rates in East Java districts/cities in 2023 using the Poisson regression method. Secondary data from BPS was analyzed and then a characteristic analysis, a multiliteracy examination, and an analysis of the Poisson regression method which included parameter estimation, simultaneous tests, partial tests, and model goodness. Furthermore, an examination of cases of overdispersion/underdispersion was carried out, and when there were cases, treatment would be carried out using negative binomial regression. The results of the analysis showed that maternal mortality was still relatively high. The number of midwives is quite large but the distribution is uneven. There are fewer public health workers. The average woman is educated up to the middle level, although the length of schooling varies. The multicollinearity test showed that there was no indication of multicollinearity between variables. The results of the Poisson regression showed that all variables, namely the number of midwives, the average length of female schooling, and the number of public health workers, had a significant effect on the number of maternal deaths per district/city in East Java in 2023. However, the overdispersion test showed an overdispersion in the data. Therefore, a negative binomial regression was used, the results of which showed only the number of midwives had a significant effect on maternal mortality, while the other two variables had no significant effect. The best model chosen between Poisson regression and negative binomial regression was the model on negative binomial regression.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Muhammad Zulfadhli
Departement of Statistics
Institut Teknologi Sepuluh Nopember
Surabaya 60111, Indonesia
Email: muhammadzulfadhli23@gmail.com

1. INTRODUCTION

East Java Province is one of the largest provinces in Indonesia, both in terms of area and population. Based on reports, in 2019-2020 there was an increase in maternal mortality from 4,221 to 4,627 maternal deaths in Indonesia [1]. One of the important indicators in assessing the quality of health services is maternal mortality. The high maternal mortality rate is often associated with limited access to health workers, service quality, and other socio-economic factors. The availability of health workers such as midwives and public health workers plays an important role in reducing maternal and child mortality rates. Midwives are often called the spearhead in maternal and child health services at the community level, while public health workers support broader promotive and preventive programs. In addition, education factors also play a role in women's health. The average length of women's schooling can be an indicator of women's knowledge and awareness of the importance of maintaining health, including in terms of safe pregnancy and childbirth. Based on this, it is important to know the factors that affect the number of maternal deaths in East Java Province [2]. This study uses secondary data from the Central Statistics Agency (BPS) in 2023 which includes several variables, namely the number of midwifery health workers, the average length of female schooling, and the number of public health workers.

Poisson regression is one of the models used for Modeling the response variable in the form of enumerated data, such as data with the response variable is abnormally distributed and discrete of type. Poisson regression based on the characteristics and uses of the poisson distribution. Poisson distribution is the distribution of a random variable expressed the number of results of the experiment that occurred in a certain time interval or in certain areas [3]. The Poisson regression model is a model used to model the number of occurrences of an event in a certain time interval. In Poisson regression, it is assumed that the dependent variable Y expresses the number (count) of events is distributed poisson. The Poisson regression model assumes equidispersion, which is a condition where the average value is equal to the variance value of the response variable of the same value [4,5].

This study will analyze the relationship between the number of midwives health workers, the number of public health workers, and the average length of female schooling to the number of maternal deaths by district/city in East Java Province in 2023 using the Poisson regression method with the aim of finding out how much these factors affect the number of maternal deaths by district/city in East Java in 2023. The data was obtained through the official BPS website. The data obtained will be processed using Minitab and R Studio software, then characteristic analysis, muktilinearly examination, and analysis of the Poisson regression method are carried out which include parameter estimation, simultaneous tests, partial tests, and model goodness. Furthermore, an examination of cases of overdispersion/underdispersion is carried out, and when there are cases, treatment will be carried out using negative binomial regression. The data that has been processed and analyzed will be interpreted so that a conclusion can be drawn.

2. METHOD

2.1. Data Source

Data on the number of maternal deaths, the number of midwives health workers, the average length of female schooling, and the number of public health workers by district/city in East Java Province in 2023 used are secondary data obtained from the official website of BPS (Central Statistics Agency) of East Java.

2.2. Research Variables

The research variables used are as shown in Table 1.

Table 1. Research Variables

Variable	Remarks	Data Scale	Units
Y	Number of maternal deaths	Ratio	Soul
X1	Number of midwives health workers	Ratio	People
X2	Average length of schooling for girls	Ratio	Year
X3	Number of public health workers	Ratio	People

2.3. Analysis Step

The analysis steps used are as follows.

1. Collecting data on the number of maternal deaths, the number of midwives, the average length of women's schooling, and the number of public health workers by district/city in East Java Province in 2023.

2. Describe the data characteristics of the number of maternal deaths, the number of midwives, the average length of women's schooling, and the number of public health workers by district/city in East Java Province in 2023.
3. Conducting poisson distribution tests and examining multicollinearity assumptions on data on the number of midwives health workers, the average length of female schooling, and the number of public health workers, on the number of maternal deaths by district/city in East Java Province in 2023.
4. Conducting poisson regression analysis on data on the number of midwives health workers, the average length of female schooling, and the number of public health workers, on the number of maternal deaths by district/city in East Java Province in 2023.
5. Examining cases of *overdispersion* or *underdispersion* on data on the number of maternal deaths, the number of midwives, the average length of women's schooling, and the number of public health workers by district/city in East Java Province in 2023.
6. To counteract *overdispersion* using negative binomial regression on data on the number of maternal deaths, the number of midwives, the average length of female schooling, and the number of public health workers by district/city in East Java Province in 2023.
7. Choose the best model between *the poisson regression* and negative binomial regression methods.
8. Interpret the results of the analysis.
9. Draw conclusions and suggestions.

2.4. Descriptive Statistics

Descriptive statistics is statistics whose level of processing is to collect, organize, and process data to be able to be presented and provide a clear picture of a certain condition or event where data is taken or in other words, the task of descriptive statistics is to present data clearly so that certain meanings or meanings can be taken based on the description presented [6].

2.1.1 Average

How to calculate average, namely by adding up all data frequencies and dividing by the amount of data. *Red* This is very close to quantitative data, usually the value of this quantitative data with the symbol that is $x_1, x_2 \dots$ to x_n (if the data contains n values). For the symbol n , it is usually used to express a size in a sample, namely the amount of data, which will be studied, which is then used as a sample. N is also used to express the size of the population, which is the number of members as a whole in the data. Average or *red* is the ratio of the total value of observations to the number of observations [7].

The single data is as follows.

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n} \quad (1)$$

Description:

x_i = The value of the variable to $-i$
 i = $1, 2, 3, \dots, n$
 n = The amount of data of the value x

The group data is as follows [7].

$$\bar{x} = \sum_{i=1}^n \frac{x_i \cdot f_i}{n} \quad (2)$$

Description:

x_i = Midpoint of the i -value
 f_i = Frequency of class i
 n = The amount of data of the value x

2.2.2 Median

Median is one of the measures of data traceability, i.e. if a group of data is sorted from smallest to largest or largest to smallest, the value of the observation is right in the middle if the amount of data is odd, or the average of the two observations in the middle if the number of observations is even [8].

Formula *median* For a single data with n odd numbers is as follows.

$$Me = \frac{1}{2} x_{n+1} \quad (3)$$

Formula *median* For a single data with n even numbers is as follows.

$$Me = \frac{1}{2} x_{\frac{n}{2}} + x_{\frac{n}{2}} \quad (4)$$

Description:

Me = Median

x = Data value

n = Amount of data

Formula *median* for the group data is as follows.

$$Me = Tb \left[\frac{\frac{n-F}{2}}{f} \right] p \quad (5)$$

Description:

Tb = Bottom edge of the class *median*

F = Cumulative frequency before class *median*

f = Class frequency *median*

n = Lots of data

p = Class length

2.3.3 Variance

Variance is the sum of the squares of all the deviations of individual values from the mean of the group. *Variance* can describe how a quantitative data is scattered [9]. *Variance* (S_2) a single data is expressed as follows.

$$S_2 = \frac{\sum (x_i - \bar{x})^2}{n-1} \quad (6)$$

The formula has an explanation, namely i is the first observation and x_1 is the first observation, then x_2 is the second observation and so on. $\bar{x}$ is the average and n is the number of observations. As for the number of observations. $\bar{x}$ *variance* (S_2) a group data is stated as follows [9].

$$S_2 = \frac{\sum_{i=1}^n f_i (x_i - \bar{x})^2}{\sum_{i=1}^n f_i} \quad (7)$$

Description:

f_i = Frequency of class i

x_i = Middle grade of the i class

i = 1, 2, 3, ... n

$\bar{x}$ = Average

n = Amount of data

2.4.4 Standard Deviation

Standard deviation is the sum of the variation or spread in a range of values. *Standard deviation* low indicates that the value tends to be close to average. From this data, *Standard Deviation* indicates that the value of the data is getting farther or farther away from the average [10]. The formula used to calculate the value *Standard Deviation* shown as follows.

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (8)$$

Description:

s = *Standard Deviation*

$\bar{x}$ = average

x_i = value to $-x_i$

i = 1, 2, 3, ... n

n = a lot of data

2.5 Multicollinearity Examination

A multicollinearity equation is a condition in which there is a strong linear relationship between two or more independent variables in a regression model. This can cause problems in estimating the regression coefficient, as the variations that occur between the free variables become difficult to solve. There are several signs that can indicate the presence of multicollinearity in a regression model, such as a very high determination coefficient, a statistically insignificant regression coefficient, and a regression coefficient that has signs that do not match the initial hypothesis [11]. The multicollinearity equation will cause the variance of the regression coefficient to be very large and unstable. This means that errors in estimating the regression coefficient can be very large and can result in inaccurate predictions. Therefore, it is important to identify and address the problem of multicollinearity in regression models. There are several ways to detect the presence of multicollinearity, namely

- a. Checking the correlation strength between predictor variables
- b. Checking the value of VIF (*Variance Inflation Factor*)

If the VIF (*Variance Inflation Factor*) value is greater than 10, it indicates the existence of multicollinearity between the predictor variables.

$$VIF = \frac{1}{1 - R_j^2} \quad (9)$$

Description:

R_j^2 = The coefficient of determination between X_j and other predictor variables.

2.6 Poisson Regression Analysis

Poisson regression analysis is a statistical technique used to understand the influence of independent variables on the frequency of events calculated in the data [12]. Poisson regression assumes that the event follows the poisson distribution and its relationship with independent variables can be estimated through the regression model. Explanation of the analysis of poisson regression are as follows.

2.6.1 Parameter Estimation

Estimation of poisson regression parameters is the process of obtaining regression coefficient values that describe the relationship between independent variables and dependent variables in the form of data count. Commonly used estimation methods in poisson regression analysis is the Maximum Likelihood Estimation (MLE) method. Estimation of regression parameters *poisson* Equations are shown follows.

$$\hat{\mu}_i = \exp(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \dots + \hat{\beta}_k X_k) \quad (10)$$

Description:

Y = variable responds
 β = parameter regresi
 e = exponential
 X = variable predictor
 K = number of predictors

2.6.2 Concurrent Test

Simultaneous poisson regression test, also known as the Chi-Square test for poisson regression models, used to evaluate the suitability or validity of poisson regression models against the observed count data. Simultaneous test of poisson regression [13] focuses on testing the null hypothesis that the poisson regression model. Overall, it does not provide a significant explanation for the variation in dependent variables. The alternative hypothesis is that at least one independent variable has a significant influence on the dependent variable. Poisson regression concurrent test can also be complemented by model quality evaluation using methods such as deviance, Akaike Information Criterion (AIC), or Bayesian Information Criterion (BIC). These methods help in the selection of the best model by considering the suitability of the model and its complexity. This simultaneous test uses the Maximum Likelihood Ratio Test (MLRT). The MLRT formula is as follows.

Hipotesis:

$H_0 : \beta_1 = \beta_2 = \beta_3 = \dots \beta_i = 0$ (Together independent variables have no significant effect on dependent variables)

$H_1 : \beta_1 \neq \beta_2 \neq \dots \beta_i \neq 0, i = 1, 2, \dots p$ (Simultaneously independent variables have a significant effect on dependent variables)

The significant level used in the simultaneous test is α with the rejection area, minus H_0 if or p-value $G^2 > X^2_{(\alpha;p)} < \alpha$. Test statistics are shown in the following formula.

$$G^2 = -2 \ln \left(\frac{L(\omega)}{L(\Omega)} \right) \quad (11)$$

Description :

G^2 = test statistics
 $L(\omega)$ = maximum value *Likelihood* No predictor
 $L(\Omega)$ = Maximum value *likelihood* with predictors

2.6.3 Partial Test

Partial poisson regression test also known as the Wald test for poisson regression coefficients, used to evaluate the individual significance of regression coefficients associated with independent variables in poisson regression models [14]. Partial poisson regression test allows an assessment of the individual

significance of each independent variable, thus allowing the identification of variables that significantly contribute to the incidence rate in the poisson regression model. The partial test hypothesis is as follows.

Hipotesis:

H0 : $\beta_i = 0$ (independent variables have no significant effect on dependent variables)

H1: $\beta_j \neq 0$ (independent variables have a significant effect on dependent variables)

The significant level used in the partial test is α with the rejection area, minus H0 if or p-value $Z > Z_{\frac{\alpha}{2}} < \alpha$.

Test statistics are shown in the following formula.

$$Z = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \quad (12)$$

Description:

OR(β_i) = standar error parameter

2.7 Dispersion Testing

The Poverty Gap Index or poverty depth index is defined as a measure of the gap in the average expenditure of each population that is classified as poor against the poverty line. This index can function as a comparison between regions, where the higher the value of an area, the more severe the poverty depth index [15].

2.8 Poverty Severity Index

Dispersion testing in poisson regression aims to evaluate whether there is an excess of variability (overdispersi) or underdispersion (underdispersi) in the data count compared to the assumption of the poisson distribution assumed [16]. *Overdispersed* occurs when the diversity of the data exceeds its average value. If it occurs *overdispersed* on discrete data and fixed using poisson regression. As a method of analysis, the conclusions obtained will be invalid. If it is a disperse parameter, then: θ

$\theta > 1$ means *overdispersion*

$\theta = 1$ means *equidispersion*

$\theta < 1$ means *underdispersion*

2.9 Goodness Check Model

A model goodness check is carried out to determine the alignment of the model with the data or the alignment between the observation results and the model's prediction results [16]. By using only one independent variable, the model makes it possible to understand the linear relationship between dependent and independent variables easily. Determination coefficient is shown in the equation. Coefficient of determination (R2) has a value that ranges from 0 to 1. The formula for determination is as follows.

$$R^2 = \frac{SS \text{ Regresi}}{SS \text{ Total}} = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (13)$$

Description :

R2 = Coefficient of determination

y_i = Observation of the i-response

$\bar{y}$ = Variable average Y

$\hat{y}$ = Y is presumed to be i-i

2.10 Negatif Binomial Regresi Analysis

Negative binomial regression analysis is a statistical method used to analyze dependent variables in the form of data counts with negative binomial distributions [17]. This method models the number of successes in a number of experiments that follow a negative binomial distribution. Negative binomial regression takes into account overdispersion, i.e. greater than expected variation in data count.

2.10.1 Parameter Estimation

Parameter estimation is the process of calculating numerical values for parameters that are not yet known in a statistical model, using the information available in the data samples that have been collected. The goal is to obtain the most accurate and reliable value of the parameter. The most commonly used parameter estimation methods are the Maximum Likelihood method and the Least Squared Method. Regression parameter estimation [18] *poisson* shown the following equation.

$$\hat{Y}_i = \exp[\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} \dots + \hat{\beta}_k X_{ki}] \quad (14)$$

Description:

Y = Respond variable

β = Parameter regresi

e = Exponential
 X = Variable predictor
 K = Number of predictors

2.10.2 Concurrent Test

The concomitant test of negative binomial regression is important for evaluating the overall contribution of independent variables in explaining variations in dependent variables. The results of these concurrent tests provide an understanding of the importance of each independent variable together in the regression model [19].

Hipotesis:

$H_0 : \beta_1 = \beta_2 = \beta_3 = \dots \beta_i = 0$ (Together independent variables have no significant effect on dependent variables)

H_1 : (simultaneously independent variables have a significant effect on dependent variables) $\beta_1 \neq \beta_2 \neq \dots \beta_i \neq 0, i = 1, 2, \dots p$

The significant level used in the simultaneous test is α with the rejection area, minus H_0 if or p-value $G^2 > X^2_{(\alpha; p)} < \alpha$.

Test statistics are shown in the following formula.

$$G^2 = -2 \ln \left(\frac{L(\omega)}{L(\Omega)} \right) \quad (15)$$

Description :

G^2 = test statistics

$L(\omega)$ = maximum value *Likelihood* No predictor

$L(\Omega)$ = Maximum value *likelihood* with predictors

2.10.3 Partial Test

A partial test commonly used in negative binomial regression is the Wald test. This test is performed to test the null hypothesis that the regression coefficient of each independent variable is zero, indicating that the variable has no significant influence on the dependent variable. The partial test of negative binomial regression allows the identification of independent variables that contribute significantly to the variation in the dependent variables individually. It is important to understand the specific influence of each independent variable in the regression model. The partial test hypothesis [20] is as follows.

Hipotesis:

$H_0 : \beta_i = 0$ (independent variables have no significant effect on dependent variables)

$H_1 : \beta_j \neq 0$ (independent variables have a significant effect on dependent variables)

The significant level used in the partial test is α with the rejection area, minus H_0 if or p-value $t_{hitung} > t_{(\frac{\alpha}{2}; n-k)} < \alpha$. Test statistics are shown in the following formula.

$$W_k = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \quad (16)$$

Description:

$\hat{\beta}_i$ = Parameter K-I

OR($\hat{\beta}_i$) = Paramater test statistics $\hat{\beta}_i$

2.11 Maternal death

Maternal death is the death of a woman during pregnancy or within 42 days of the end of pregnancy, regardless of the length and location of the pregnancy, for any cause related to or aggravated by the pregnancy or its treatment, but not by accidental or incidental cause [21].

2.12 Midwife Health Workers

Midwives are health professionals who have completed midwifery education and are recognized by the authorized health authorities to provide midwifery services, including care during pregnancy, childbirth, postpartum period, and other reproductive services [22].

2.13 Average Length of Schooling for Girls

The average length of schooling for girls is the total number of years spent by the female population aged 15 and over in formal education, divided by the number of female residents in that age group.

2.14 Public Health Workers

Public health workers are individuals who have an educational background in the field of public health and have the authority to carry out health promotion, disease prevention, and efforts to improve the quality of public health.

2.15 Previous Research

Research related to modeling socio-economic determinants, health status, and regional development in Indonesia has demonstrated the importance of multivariable and nonparametric statistical approaches in capturing the complexity of regional data, such as unmet need status [23],[24], nonparametric binary logistic regression inference frameworks for the status of disadvantaged areas [25], and some studies using classification models such as [26] as well as modeling of open unemployment status and public health [27],[28],[29].

3. RESULTS AND DISCUSSION

3.1. Data Characteristics of the Number of Maternal Deaths, the Number of Midwives Health Workers, the Average Length of Women's Schooling, and the Number of Public Health Workers by Regency/City in East Java Province in 2023

The characteristics of the number of maternal deaths, the number of midwives, the average length of female schooling, and the number of public health workers by district/city in East Java Province in 2023 are shown in Table 1 as follows.

Table 2. Data Characteristics

Variable	Varians	Mean	Maximum	Minimum
Y	170.397,2	840,7	1891	146
X_1	505.546	1.240	2895	168
X_2	3,047	7,949	11,390	4,640
X_3	5.521,1	80,3	477	20

Table 2 shows that the data characteristics on variable Y, which is the number of maternal deaths, have variance, *mean*, maximum value, and minimum values respectively of 170397.2 people, 840.7 people, 146 people, and 1891 people. The data characteristics in the variable, which is the number of midwives, have variance, X_1 *mean*, maximum value, and minimum values respectively of 505546 people, 1240 people, 168 people, and 2895 people. The data characteristics on the variable, which is the average length of women's schooling, have variance, X_2 *mean*, maximum value, and minimum values of 3,047 years, 7,949 years, 4,640 years, and 11.39 years, respectively. Finally, the data characteristics on the variable, which is the number of public health workers, have variance, X_3 *mean*, maximum value, and minimum values of 5521.1 people, 80.3 people, 20 people, and 477 people, respectively.

3.2. Multicollinearity Examination of the Number of Maternal Deaths, the Number of Midwives Health Workers, the Average Length of Women's Schooling, and the Number of Public Health Workers by Regency/City in East Java Province in 2023

The multicollinearity examination of the number of maternal deaths, the number of midwives health workers, the average length of female schooling, and the number of public health workers by district/city in East Java Province in 2023 uses the following *Variance Inflation Factor (VIF)* values.

Table 3. Multicollinearity Detection

Variable	Remarks	VIF
X_1	Number of midwives health workers	1,74
X_2	Average length of schooling for girls	1,42
X_3	Number of public health workers	1,95

Table 3 shows that the VIF value on the number of midwives, the average length of female schooling, and the number of public health workers are 1.74, 1.42, and 1.95 respectively, where the three have a value of less than 10. Therefore, it can be concluded that there is no multicollinearity in the three variables.

3.3. Poisson Regression Analysis of Factors Affecting the Number of Maternal Deaths by Regency/City in East Java Province in 2023

Poisson regression analysis was carried out to determine the influence of several predictor variables on response variables. This analysis included the process of estimating parameters, simultaneous tests, partial

tests, and checking the goodness of the model carried out on the number of maternal deaths, the number of midwives, the average length of female schooling, and the number of public health workers as follows.

3.3.1 Estimation of Poisson Regression Parameters of Total Maternal Mortality

Parameter estimation from the model of the number of maternal deaths, the number of midwives, the average length of women's schooling, and the number of public health workers is shown through the poisson regression model as follows.

$$\hat{y} = \exp(6,1897 + 0,000508 x_1 - 0,02816 x_2 + 0,000659 x_3)$$

The poisson regression model above shows that the number of midwives health workers, the average length of female schooling, and the number of public health workers are 0, then the number of maternal deaths increases by $\exp(6.1897)$ or by 487.7 people. If the number of midwives health workers increases by 1 person, then the number of maternal deaths increases by $\exp(0.000508)$ or by 1.000508 people provided that other variables are constant. If the average length of female schooling increases by 1 year, then the number of maternal deaths decreases by $\exp(0.02816)$ or by 1.02856 people. If the number of public health workers increases by 1 person, then the number of maternal deaths increases by $\exp(0.000659)$ or by 1.00066 people.

3.3.2 Simultaneous Test of the Poisson Regression Model of Total Maternal Mortality

Simultaneous tests from the modeling of the number of maternal deaths, the number of midwives, the average length of schooling for women, and the number of public health workers were shown through the following poisson regression model.

Hipotesis:

$H_0: \beta_1 = \beta_2 = \beta_3 = 0$ (VPredictor variabel has no significant effect on total maternal mortality)

$H_1: \text{at least } 1; j=1,2,3 \beta_j \neq 0$ (There is at least one predictor variable that has a significant effect on total maternal mortality)

If a significant level of α of 5% or 0.05 is used, then a rejection criterion is obtained, namely H_0 rejected if or $G^2 > X^2_{(0,05;3)} P_{value} < \alpha$.

Table 4. Results of Simultaneous Test

G^2	$X^2_{(0,05;3)}$	P_{value}
6.016,02	7,81	0,00

Table 4 shows that statistical calculations produce a value G^2 6016.02 is greater than 7.81, and strengthened by a value of 0.00 which is smaller than the α value of 0.05. So that a rejection decision can be made, which means that there is at least one predictor variable that has a significant effect on the number of maternal deaths.

3.3.3 Partial Test of the Poisson Regression Model

Partial tests from the modeling of the number of maternal deaths, the number of midwives, the average length of female schooling, and the number of public health workers were shown through the following poisson regression model.

Hipotesis:

$H_0: \beta_1 = 0$ (The number of midwives health workers does not have a significant effect on total maternal mortality)

$H_1: \beta_1 \neq 0$ (The number of midwives has a significant effect on total maternal mortality)

$H_0: \beta_2 = 0$ (The average length of girls' schooling does not has a significant effect on total maternal mortality)

$H_1: \beta_2 \neq 0$ (Average length of schooling for girls has a significant effect on total maternal mortality)

$H_0: \beta_3 = 0$ (Number of public health workers has no significant effect on total maternal mortality)

$H_1: \beta_3 \neq 0$ (Number of public health workers has a significant effect on total maternal mortality)

If a significant level of α of 5% or 0.05 is used, the rejection criterion is obtained, namely H_0 rejected if or $|Z_{hitung}| > Z_{0,975} P_{value} < \alpha$.

Table 5. Partial Test Results of Regression Models

Parameter	$ Z_{hitung} $	$Z_{0,975}$	P_{value}
β_1	52,51		0,00
β_2	6,96	1,96	0,00
β_3	8,73		0,00

Table 5 shows that the statistical calculation of partial tests for the number of midwives resulted in a value of 52.51 which is greater than 1.96, and strengthened by 0.00 which is smaller than the significant level of α of 0.05. So that a rejection decision can be made, which means that the number of midwife health workers has a significant effect on the number of maternal deaths.

Furthermore, for the average length of women's schooling, a score of 6.96 is greater than 1.96, and strengthened by 0.00 which is smaller than the significant level of α of 0.05. So that a rejection decision can be made, which means that the average length of women's schooling has a significant effect on the number of maternal deaths.

As for the number of public health workers, a value of 8.73 was greater than 1.96, and strengthened by 0.00 which was smaller than the significant level of α of 0.05. So that a rejection decision can be made, which means that the number of public health workers has a significant effect on the number of maternal deaths.

3.3.4 Examination of the Benefits of the Poisson Regression Model

The goodness of the regression model from the modeling of the number of maternal deaths, the number of midwives, the average length of female schooling, and the number of public health workers produced a determination coefficient of R^2 72.32%, which means that the proportion of variability of the three predictor variables was able to explain the model of 72.32% and the remaining 27.68% was explained by other variables outside the model.

3.4. Overdispersion or Underdispersion Check

The overdispersion or underdispersion examination of data on the number of midwives health workers, the average length of women's schooling, and the number of public health workers on the number of maternal deaths by district/city in East Java Province in 2023 is shown in Table 6 as follows.

Table 6. Overdispersion or Underdispersion Check

<i>Deviance</i>	<i>df</i>	ϕ
2191,9	34	64,47

Table 6 shows that a *deviance* value of 2191.9 with a degree of freedom of 34 was obtained, so that the ratio between *deviance* and degree of freedom was obtained of 64.67 which is greater than 1. This means that data on the number of midwives health workers, average school age, and the number of public health workers to the number of maternal deaths by district/city in East Java Province in 2023 are overdispersed, so it is necessary to handle using negative binomial regression.

3.5. Handling Overdispersion Using Negative Binomial Regression

The results of overdispersion handling using negative binomial regression on data on the number of midwives, the average length of schooling, the number of public health workers to the number of maternal deaths by district/city in East Java Province in 2023 are explained as follows.

3.5.1 Parameter Estimation

The estimation of parameters from the modeling of the number of maternal deaths, the number of midwives, the average length of female schooling, and the number of public health workers is shown through the negative binomial regression model as follows.

$$\hat{y} = \exp(6,087 + 0,000642 x_1 - 0,03889 x_2 + 0,000674 x_3)$$

The negative binomial regression model above shows that the number of midwives health workers, the average length of female schooling, and the number of public health workers are 0, then the number of maternal deaths increases by $\exp(6.087)$ or by 440.1 people. If the number of midwife health workers increases by 1 person, then the number of maternal deaths increases by $\exp(0.000642)$ or by 1.00064 people provided that other variables are constant. If the average length of female schooling increases by 1 year, then the number of maternal deaths decreases by $\exp(0.03889)$ or by 1.03966 people. If the number of public health workers increases by 1 person, then the number of maternal deaths increases by $\exp(0.000674)$ or by 1.0006741 people.

3.5.2 Negative Binomial Regression Simultaneous Test

Simultaneous tests from the modeling of the number of maternal deaths, the number of midwives health workers, the average length of female schooling, and the number of public health workers were shown through the negative binomial regression model as follows.

Hipotesis:

$H_0: \beta_1 = \beta_2 = \beta_3 = 0$ (Predictor variable has no significant effect on total maternal mortality)

$H_1: \text{at least } 1; j=1,2,3 \beta_j \neq 0$ (There is at least one predictor variable that has a significant effect on total maternal mortality)

If a significant level of α of 5% or 0.05 is used, then a rejection criterion is obtained, namely H_0 rejected if or $G^2 > X^2_{(0,05;3)} P_{value} < \alpha$.

Table 7. Results of Simultaneous Test of Negative Binomial Regression Model

G^2	$X^2_{(0,05;3)}$	P_{value}
81,18	7,81	0,00

Table 7 shows that statistical calculations produce a value G^2 of 81.18 is greater than 7.81, and strengthened by a value of 0.000 which is smaller than the α value of 0.05. So that a rejection decision can be made, which means that there is at least one predictor variable that has a significant effect on the number of maternal deaths by district/city in East Java Province in 2023

3.5.3 Negative Binomial Regression Simultaneous Test

Partial tests from the modeling of the number of maternal deaths, the number of midwives, the average length of female schooling, and the number of public health workers were shown through the negative binomial regression model as follows.

Hipotesis:

$H_0: \beta_1 = 0$ (The number of midwives health workers does not have a significant effect on total maternal mortality)

$H_1: \beta_1 \neq 0$ (The number of midwives has a significant effect on total maternal mortality)

$H_0: \beta_2 = 0$ (The average length of girls' schooling does not has a significant effect on total maternal mortality)

$H_1: \beta_2 \neq 0$ (Average length of schooling for girls has a significant effect on total maternal mortality)

$H_0: \beta_3 = 0$ (Number of public health workers has no significant effect on total maternal mortality)

$H_1: \beta_3 \neq 0$ (Number of public health workers has no significant effect on total maternal mortality)

If a significant level of α of 5% or 0.05 is used, the rejection criterion is obtained, namely H_0 rejected if or $|T_{hitung}| > T_{(\frac{\alpha}{2}, n-k)} P_{value} < \alpha$.

Table 8. Partial Test Results of Negative Binomial Regression Model

Parameter	$ T_{hitung} $	$T_{(0,025;35)}$	P_{value}
β_1	6,458		0,000
β_2	1,095	2,030	0,281
β_3	0,742		0,463

Table 8 shows that the statistical calculation of partial tests for the number of midwife health workers produced a value of 6,458 which is greater than 2,030, and strengthened by 0.000 which is smaller than the significant level of α 0.05. So that a rejection decision can be made, which means that the number of midwife health workers has a significant effect on the number of maternal deaths. $|T_{hitung}| T_{(0,025;35)} P_{value} H_0$ Furthermore, for the average length of school, women produced a score of 1.095 which was smaller than 2.030, and strengthened by 0.281 which was greater than the significant level of α of 0.05. So that a decision can be made to fail to reject which means that the average length of school for women does not have a significant effect on the number of maternal deaths.

As for the number of public health workers, a value of $|T_{hitung}| 0.742$ was produced which was greater than $T_{(0,025;35)} 2.030$, and strengthened by 0.463 P_{value} which was smaller than the significant level of α of 0.05. So that a decision can be made to fail to refuse, which means that the number of public health workers does not have a significant effect on the number of maternal deaths

3.6. Best Model Selection

The selection of the best model will be carried out by comparing the results of the poisson regression analysis and the negative binomial regression shown in Table 9.

Table 9. Selection of the Best Regression Model

Regression Method	Significantly Influential Parameters	AIC
Poisson	$\beta_1, \beta_2, \beta_3$	2519,9
Binomial Negative	β_1	525,06

Table 9 shows the AIC value on the factors that affect the number of maternal deaths in East Java Province in 2023 on the negative binomial regression method shows a value of 525.06 which means it is smaller than the AIC in the poisson regression analysis of 2519.9. Therefore, it can be decided that the best model of the factors affecting maternal mortality in East Java Province in 2023 is the negative binomial regression method.

4. CONCLUSION

Based on the results of the analysis and discussion, the following conclusions were obtained.

1. The data characteristics of the number of maternal deaths have an average of 840.7 people with a variance of 170397.2 people. The number of midwives has an average of 1240 people with a variance of 505546 people, while the number of public health workers has an average of 80.3 people with a variance of 5521.2 people. The average length of schooling for women is 7,949 years with a variance of 3,047 years.
2. The results of the multicollinearity examination did not indicate multicollinearity.
3. The results of *the poisson regression* show that the number of midwives health workers, the average length of women's schooling, and the number of public health workers have a significant effect on the number of maternal deaths by district/city in East Java Province in 2023
4. The results of the overdispersion examination show that overdispersion occurs.
5. The results of the negative binomial regression analysis showed that the number of midwives had a significant effect on the number of maternal deaths by district/city in East Java Province in 2023, while the average length of women's schooling and the number of public health had no significant effect on the number of maternal deaths.

The best model chosen between poisson regression and negative binomial regression is the model on negative binomial regression.

ACKNOWLEDGMENTS

Thank you to all parties involved in the compilation of this article.

FUNDING INFORMATION

There is no funding in this article.

AUTHOR CONTRIBUTIONS STATEMENT

Specify the contributions of each author and the role of using limited assistance from Artificial Intelligence (AI) as follows:

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
M. Rizal Matofani	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Muhammad Zulfadhli		✓				✓		✓	✓	✓	✓	✓		
Nurfadhlan Dhafin	✓		✓	✓			✓			✓	✓		✓	✓
Nugroho														
Artika Putri		✓				✓		✓	✓	✓	✓			
Rahmawati														
Fi Amanila Putri	✓		✓	✓			✓			✓	✓		✓	✓
Kinanti														

C : Conceptualization

M : Methodology

So : Software

Va : Validation

Fo : Formal analysis

I : Investigation

R : Resources

D : Data Curation

O : Writing - Original Draft

E : Writing - Review & Editing

Vi : Visualization

Su : Supervision

P : Project administration

Fu : Funding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

- [1] Afryanthy, et al., "Analisis Diskriminan pada Indikator IPM Kabupaten/Kota di Provinsi Sulawesi Selatan," VARIANSI: Journal

- of Statistics and Its Application on Teaching and Research, 2020.
- [2] S. Eskani, "Pengaruh Karbon Aktif dari Batok Kelapa untuk Menurunkan Zat Besi dan Mangan dalam Air Sumur," *Jurnal Ilmiah Teknik Lingkungan*, 2017.
 - [3] Á. García del Castillo-López and M. P. Domínguez, "Employees' Emotional Intelligence and Job Satisfaction: The Mediating Role of Work Climate and Job Stress," *Administrative Sciences*, 2024.
 - [4] M. S. Ginting, Hubungan Pilihan Mahasiswa Unja Dalam Pembayaran Uang Kuliah Tunggal Berdasarkan Jenis Bank Dengan Tabel Kontingensi Uji Statistik Chi-Square, Jambi: Universitas Jambi, 2023.
 - [5] R. Hoppock, "Job satisfaction of psychologists," *J. Appl. Psychol.*, vol. 21, pp. 300-303, 1937.
 - [6] T. Hidayati, *Statistika Dasar*, Purwokerto: CV. Pena Persada, 2019.
 - [7] S. M. Nisak Ruwah Ibnatur Husnul, *Statistika Deskriptif*, Pamulang - Tangerang Selatan: Unpam Press, 2020.
 - [8] R. Walpole and M. Myers, *Probability & Statistics for Engineers & Scientists*, London: Pearson Education, 2016.
 - [9] M. Maswar, "Analisis Statistik Deskriptif Nilai UAS Ekonometrika Mahasiswa dengan Program SPSS 23 & Eviews 8.1," *Jurnal Pendidikan Islam Indonesia*, pp. 273–292, 2017.
 - [10] Indwiarti, R. P. H. G. et al., *Pengantar Statistika Eksplorasi Data dengan MS.Excel dan R*, Yogyakarta: KBM Indonesia, 2021.
 - [11] N. R. Asnidar, "Penerapan Analisis Diskriminan Kuadratik Robust Pada Klasifikasi Desa," *ESTIMASI: Journal of Statistics and Its Application*, vol. 5, no. 2, pp. 196–240, 2024.
 - [12] D. N. Putri, *Rancangan Penelitian Bidang Teknologi Pangan*, Malang: UMMPress, 2020.
 - [13] N. E. C. Qonitatin Nafisah, "Analisis Cluster Average Linkage Berdasarkan Faktor-Faktor Kemiskinan di Provinsi Jawa Timur," *Zeta – Math Journal*, pp. 31–36, 2017.
 - [14] Usmadi, "PENGUJIAN PERSYARATAN ANALISIS (Uji Homogenitas dan Uji Normalitas)," *Inovasi Pendidikan*, vol. 7, no. 1, 2020.
 - [15] Barnabas Anthony Philbert Marbun, "Penentuan Kategori Kelulusan Mahasiswa Menggunakan Metode Analisis Diskriminan," *Zeta-Math Journal*, vol. 9, no. 1, pp. 31–40, 2024.
 - [16] S. Nia Shantian Lumbantobing, "Confirmatory Factor Analysis to Measure the Satisfaction Level of Air Soda Tourists in North Tapanuli," *Formosa Journal of Science and Technology (FJST)*, vol. 2, no. 2, pp. 475–492, 2023.
 - [17] Rizka Widya Berliana et al., "Analisis Diskriminan Pada Indikator yang Memengaruhi Indeks Ketahanan Pangan Menurut Kabupaten/Kota di Indonesia Tahun 2022," *Seminar Nasional Official Statistics*, pp. 11–20, 2023.
 - [18] M. T. Ashshiddiqi et al., "Strategi Program Indeks Pembangunan Manusia (IPM) di Desa Pakuaon Kecamatan Sukaresmi Kabupaten Cianjur," *Jurnal Inovasi Penelitian (JIP)*, vol. 2, no. 4, pp. 1153–1162, 2021.
 - [19] S. Sitompul, "Peran Tenaga Kerja Formal dan Upah Minimum Dalam Mengatasi Kemiskinan Pulau Jawa Tahun 2015-2022," *e-journal UAJY*, 2023.
 - [20] Grace Wulandari, "Pemodelan Persentase Penduduk Miskin Di Indonesia Menggunakan Regresi Probit dan Regresi Logistik," *Prosiding Seminar Nasional Matematika, Statistika, dan Aplikasinya*, pp. 354–368, 2022.
 - [21] A. U. Yogi Ramadhan, "Pendugaan Persentase Penduduk Miskin Ekstrem Menggunakan Small Area Estimation dengan Partitioning Around Medoids Clustering," *Politeknik Statistika STIS*, pp. 557–568, 2022.
 - [22] T. A. N. A. Ayu Noviatamara, "Analisis Pengaruh Pertumbuhan Ekonomi dan Tingkat Pengangguran Terbuka di Daerah Istimewa Yogyakarta," *Jurnal REP (Riset Ekonomi Pembangunan)*, pp. 53–60, 2019.
 - [23] M. Zulfadhli, I. N. Budiantara, V. Ratnasari, and A. S. Suriaslan, "Fourier series nonparametric regression estimator for modeling status of unmet need in East Java Province in 2023," *Commun. Math. Biol. Neurosci.*, vol. 2025, Art. ID 2025.
 - [24] A. S. Suriaslan, I. N. Budiantara, V. Ratnasari, M. Zulfadhli, and R. C. Dhewy, "An Empirical Study on Unmet Need: A Statistical Inference Framework for Truncated Spline Nonparametric Binary Logistic Regression (TSNLBR)," *J. Cultural Anal. Social Change*, pp. 868–877, 2025.
 - [25] M. Zulfadhli, I. N. Budiantara, V. Ratnasari, A. S. Suriaslan, and R. C. Dhewy, "A statistical inference framework for FSNBLR: Modeling underdeveloped regional status in Eastern Indonesia," *MethodsX*, p. 103746, 2025.
 - [26] M. Zulfadhli, I. N. Budiantara, V. Ratnasari, and A. S. Suriaslan, "Semiparametric Regression Estimator of Fourier Series For Categorical Data," *IAENG Int. J. Appl. Math.*, vol. 55, no. 7, pp. 2157–2164, 2025.
 - [27] B. W. Otok, M. Zulfadhli, R. D. Pangesti, M. I. Kurniawan, A. E. P. Haryanto, and I. Kurniawan, "Comparison of Binary Probit Regression and Fourier Series Nonparametric Logistic Regression in Modeling Diabetes Status at Hajj General Hospital Surabaya," *Barekeng*, vol. 20, no. 1, pp. 255–270, 2026.
 - [28] A. M. Achmad, M. Zulfadhli, and N. Nurfaudzan, "Binary Logistic Regression on the Status of Unemployment Rate in West Java Province in 2023," *Likelihood: J. Stat. Its Appl.*, vol. 1, no. 1, pp. 30–34, 2025.
 - [29] M. Zulfadhli, I. N. Budiantara, and V. Ratnasari, "Nonparametric regression estimator of multivariable Fourier series for categorical data," *MethodsX*, vol. 13, p. 102983, 2024.