IDENTIFYING FACTORS INFLUENCING THE NUMBER OF DIARRHEA CASES IN CHILDREN UNDER FIVE IN WEST JAVA USING NEGATIVE BINOMIAL REGRESSION

Akbar Rizki[1)]
Christin Halim[2)]
Utami Dyah Syafitri[3)]
[1,2,3)] Department of Statistics, Faculty of Mathematics and Natural Science, IPB University, Bogor, Indonesia
e-mail: akbar.ritzki@apps.ipb.ac.id

ABSTRACT

The WHO states that diarrhea is the leading killer of children under five worldwide, and Indonesia is no exception, where 10.3% of under-five deaths are caused by diarrhea. West Java Province, with the largest population in Indonesia, has the highest diarrhea cases under five. The potential for diarrhea to become an extraordinary event, which is often accompanied by death, is very likely to occur because diarrhea is an endemic disease in West Java. Therefore, analyzing the factors influencing the children under five diarrhea cases in West Java is essential. Negative binomial regression was used in this study because the response was to count data on the incidence of diarrhea in children under five in West Java. The analysis results show that an increase in the percentage of public premises (PPP) meeting health requirements and population density per km^2 will increase the number of diarrhea cases under five in West Java. However, an increase in the percentage of Community-Based Total Sanitation (CBTS), percentage of the population living in poverty, and percentage of households practicing Clean and Healthy Behavior (CHB) will decrease the number of diarrhea cases in West Java.

Keywords: diarrhea, negative binomial regression, overdispersion, under-five years old.

INTRODUCTION

Diarrhea is the leading cause of death for children under five in the world. According to calculations from data in 2019, more than 1,200 children die per day from diarrhea (UNICEF, 2024). Every year, there are about 1.7 billion cases of diarrhea among children under five years in the world. The disease causes 525,000 deaths of children under five in the world (WHO, 2017). Diarrhea is a disease characterized by the frequency of defecation of more than three times per day, and the concentration is watery (Dinas Kesehatan Jawa Barat, 2022). Diarrhea is the leading cause of death among children under five in Indonesia, at 10.3%. West Java Province has the highest number of cases of diarrhea under five (Kementerian Kesehatan Republik Indonesia, 2022). Diarrhea is an endemic disease in West Java and has the potential for extraordinary events, which are often accompanied by death (Dinas Kesehatan Jawa Barat, 2022). The disease is preventable and treatable, but it is the most significant cause of death in the world. Therefore, an analysis of the factors that influence children under-five diarrhea cases will be conducted to provide input to the local government of West Java Province to reduce the number of under-five diarrhea cases and under-five mortality rates in West Java.

Previous research on diarrhea was conducted by (Widyaningrum *et al*., 2021) by modeling children under-five diarrhea disease in East Java; it was found that the percentage of malnourished toddlers and the percentage of the population living in poverty had a significant effect on the number of under-five diarrhea cases. The Indonesian Ministry of Health (2022) provides information on how food management places and public premises that meeting health requirements can prevent the spread of diseases, such as diarrhea. Several studies have found the results of explanatory variables that have different significant effects, namely the percentage of community-based total sanitation in the research of Monica *et al*. (2020), the percentage of exclusive breastfeeding in research by Melvani *et al*. (2019), population density per km^2 in research by Sidqi *et al*. (2021), the percentage of households without toilet facilities in the study of Nurhayati *et al*. (2018), the percentage of households living in inadequate conditions in research by Cahyani *et al*. (2019), and the percentage of households practicing Clean and Healthy Behavior (CHB) in the study of Mas *et al*. (2017).

Data on the number of cases of diarrhea in children under five years old is counting data whose occurrence is rare. One model approach that can be used is the Poisson distribution approach. Models with Poisson distribution can be used to deal with count data. This was shown in Faidah's research (2017), which analyzed the data of diarrhea cases children under five, which is discrete data. In addition, research by Yulianingsih *et al*. (2012) also used Poisson regression in modeling data on the incidence of not passing the national exam (UN) SMA / SMK in Bali, which is count data. The Poisson regression model has the assumption of equidispersion, which is a state with the same mean and variance. However, count data generally has a higher variance value than the mean. This situation is inversely proportional to the Poisson regression model assumptions, usually overdispersion. Violating the assumption of equidispersion in the Poisson regression model can cause bias in the parameter estimates obtained (S. P. Lestari & Wulandari, 2014). Overdispersion can be overcome with a negative binomial regression approach. The model can overcome overdispersion because it allows the parameters to describe the variation of the data (Hardin & Hilbe, 2014).

Previous research conducted by Fitrial & Fatikhurrizqi (2020) modeled the number of COVID-19 cases in Indonesia using the Poisson and negative binomial regression approaches. The results showed that the negative binomial regression model was better than the Poisson regression model. Another study with count data that experienced overdispersion conducted by Utami (2013) also found that the negative binomial regression model was the best. In addition, N. K. Lestari *et al*. (2019) research on environmental factors causing diarrhea cases in West Java also found that the negative binomial regression model best models count data with overdispersion conditions. Therefore, this study will use negative binomial regression to model the number of cases of diarrhea under five in West Java.

METHOD

This study used data from the West Java Central Bureau of Statistics and the West Java Health Profile in 2021 (Dinas Kesehatan Jawa Barat, 2022). The observation unit in this study is 27 districts/cities in West Java Province in 2021. The data consists of ten explanatory variables and one response variable. Details of the variables used in this study can be seen in Table 1.

This research employs negative binomial regression analysis utilizing Software R 4.1.1. The initial stages of data analysis in this study commenced with data exploration. Data exploration was conducted to observe data characteristics and patterns. The initial exploration involved computing descriptive statistics, including minimum, first quartile, median, mean, third quartile, maximum, and variance of the response variable, in which the number of diarrhea cases per 100,000 children under five in West Java. Subsequent exploration aimed to discern the distribution of the response variables using histograms.

Table 1. List of variables used in the study

Code	Variable	Reference[A1][A2]
Y	Number of under-five diarrhea cases per 100,000 children under five in West Java	
X_1	Percentage of food management places (FMP) meeting health requirements	Kementerian Kesehatan Republik Indonesia, (2022)
X_2	Percentage of public premises (PPP) meeting health requirements	Kementerian Kesehatan Republik Indonesia, (2022)
X_3	Percentage of Community-Based Total Sanitation (CBTS)	Monica *et al*. (2020)
X_4	Percentage of malnourished toddlers	Widyaningrum *et al*. (2021)
X_5	Percentage of exclusive breastfeeding	Melvani *et al*. (2019)
X_6	Percentage of the population living in poverty	Widyaningrum *et al*. (2021)
X_7	Population density per km^2	Sidqi *et al*. (2021)
X_8	Percentage of households without toilet facilities	Nurhayati *et al*. (2018)
X_9	Percentage of households living in inadequate conditions	Cahyani *et al*. (2019)
X_{10}	Percentage of households practicing Clean and Healthy Behavior (CHB)	Mas *et al*. (2017)

The last exploration process in this study is to check for multicollinearity using the variance inflation factor (VIF) of each explanatory variable. If the VIF value of the explanatory variable is more than 5, it indicates there is multicollinearity. Thus, it is necessary to select variables using stepwise selection. The VIF value can be calculated with the following formula:

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

where R_i^2 is the coefficient of determination between X_i with the other explanatory variables.

The next step is to check for overdispersion in the data. This step begins with performing Poisson regression model[A3][A4]. This is because the test for overdispersion identification uses the output generated from Poisson regression model[A5][A6]. The test is carried out by looking at the dispersion parameters, namely the deviance value and Pearson's Chi-Square value, each divided by the degree of freedom. Suppose D^2 is the deviance value, χ^2 is Pearson's Chi-Square value, y_i is the value of the i-th observation, $\hat{\mu}_i$ is the estimated value of the i-th response mean, df is the degree of freedom with the formula $df = n - k$, n is the number of observation, and k is the number of the regression coefficient. The following is the test statistic formula used in overdispersion testing.

$$D^2 = 2\sum_{i=1}^{n}\left\{y_i \ln\left(\frac{y_i}{\hat{\mu}_i}\right) - (y_i - \hat{\mu}_i)\right\} \quad (2)$$

$$\phi_1 = \frac{D^2}{df} \quad (3)$$

$$\chi^2 = \sum_{i=1}^{n}\frac{(y_i - \mu_i)^2}{var(y_i)} \quad (4)$$

$$\phi_2 = \frac{\chi^2}{df} \quad (5)$$

Data is said to be over disperses if ϕ_1 and $\phi_2$are greater than one.

Negative binomial regression model [A7][A8]is performed if there is overdispersion in the data. The first step in this modeling is to estimate the parameters. The method used in estimating the parameters of this model is the maximum likelihood estimation (MLE)method by maximizing the log-likelihood equation, with the equation:

$$\ln L(y;\mu,\alpha) = \sum_{i=1}^{n}\left\{y_i \ln\left(\frac{\alpha\mu_i}{1+\alpha\mu_i}\right) - \frac{1}{\alpha}\ln(1+\alpha\mu_i) + \ln\Gamma\left(y_i + \frac{1}{\alpha}\right) - \ln\Gamma(y_i+1) - \ln\Gamma\left(\frac{1}{\alpha}\right)\right\} \quad (6)$$

The regression coefficient (β) can be obtained from the equation by maximizing the log-likelihood equation using Newton-Raphson iteration. Furthermore, simultaneous parameter hypothesis testing is carried out using the G-test. This step is followed by partial parameter hypothesis testing using the Wald test. After conducting hypothesis testing, explanatory variables that are proven to have a significant effect at the 5% real level on the number of diarrhea cases among children under five in West Java will be interpreted.

RESULTS AND DISCUSSION

Data Exploration

Descriptive statistics of under-five diarrhea cases per 100,000 children under-fives in West Java are presented in Table 2. West Java Province has the highest total under-five diarrhea cases in Indonesia with a total of 666,244 that have occurred in 2021. Cases of under-five diarrhea in West Java districts/cities have a very large variation of 2,142,489 indicating the difference in the number of cases of under-five diarrhea between districts/cities is relatively large. The highest number of under-five diarrhea cases occurred in Depok City with the number of under-five diarrhea cases per 100,000 children under five amounting to 10,397 cases, while the lowest number of under-five diarrhea cases occurred in Sukabumi District with the number of under-five diarrhea cases per 100,000 children under five amounting to 5,153 cases. The mean value of cases is greater than the median, indicating that the distribution of under-five diarrhea cases in West Java is stretching to the right.

Table 2. Descriptive statistics of the number of under-five diarrhea cases per 100,000 under-fives in West Java

Descriptive statistic	Value
Minimum	5,153
First quartile	5,792
Median	6,298
Mean	6,714
Third quartile	6,924
Maximum	10,397
Variance	2,142,489

Figure 1 shows that the data distribution of the response variable skewed to the right according to the descriptive statistics obtained. The distribution looks like a Poisson regression distribution. Therefore, it is suspected that the data on under-five diarrhea per 100,000 children under-five has a Poisson distribution. This is in line with the research of Faidah & Wibowo (2017), Fuad *et al*. (2018), and Yulianingsih *et al*. (2012) who modeled count data using Poisson regression.

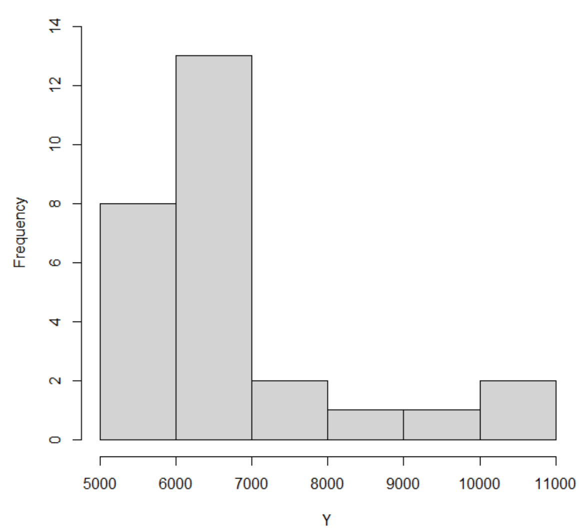

Figure 1. Histogram of the number of under-five diarrhea cases per 100,000 under five in West Java

Multicollinearity can lead to incorrect interpretation and produce a significant variance in regression analysis. This causes the resulting model to be irrelevant. A common method to detect multicollinearity is using the variance inflation factor (VIF). The VIF value for each explanatory variable is no more than 5 (Table 3). This indicates that there is no multicollinearity in the model.

Table 3. VIF value of each explanatory variable

Explanatory variable	VIF
X_1	2.4645
X_2	3.3293
X_3	2.1698
X_4	2.0304
X_5	1.4181
X_6	3.5258
X_7	3.1895
X_8	1.9826
X_9	1.5687
X_{10}	2.3983

Overdispersion Identification

Poisson regression has an assumption that must be met: the assumption of equidispersion. Equidispersion is a condition where the expected value of the response variable is equal to the value of its variance. The test statistic used to check equidispersion is the deviance and Pearson's Chi-Square values, each divided by the degree of freedom (df) in Poisson regression. If the ratio of deviance and Pearson's Chi-Square value with the[A9][A10] degree of freedom are greater than one, the data is over dispersed so it can be said that the assumption of equidispersion is not met.

The results of calculating the deviance value/df and Pearson's Chi-Square value/df are 82.775 and 82.011, respectively. Thus, it can be concluded that there is a violation of assumptions in the Poisson regression model, which can result in underestimated standard errors. The small standard error causes the p-value to be smaller so that the partial test results will tend to reject H_0, which means that the explanatory variables have a real effect on the number of under-five diarrhea cases per 100,000 children under-five in West Java. In other words, the test results become invalid and irrelevant. Therefore, overdispersion will be handled in the Poisson regression model using the negative binomial distribution approach. Negative binomial regression modeling does not require the assumption of equidispersion and is suitable for count data.

Binomial Negative Regression Model

The negative binomial regression model is one of the models used to model count data with overdispersion conditions. Negative binomial regression can overcome overdispersion in Poisson regression because it has a dispersion parameter (α). The parameter is obtained from the relationship between mean and variance in the data. Negative binomial regression analysis will be conducted with ten explanatory variables. The model has a dispersion parameter (α) of 139.956. This model obtained simultaneous test statistics, namely the G-test of 45.795 greater than the value of $\chi^2_{(0{,}05,10)} = 18{,}31$, so the conclusion obtained is to reject H_0 which means that there is at least one explanatory variable that has a significant effect on the number of cases of under-five diarrhea per

100,000 children under five in West Java. The results of the parameter estimates generated from this model can be seen in Table 4.

Table 4. Parameter estimates of the negative binomial regression model with 10 explanatory variables

Parameter	Coefficient	Wald Statistic	P-value
$\boldsymbol{\beta_0}$	$\mathbf{9.62200}$	**859.08**	$\mathbf{0.000}$
β_1	-0.00206	1.87	0.172
$\boldsymbol{\beta_2}$	$\mathbf{0.00562}$	**9.13**	$\mathbf{0.003}$
$\boldsymbol{\beta_3}$	$\mathbf{-0.00513}$	**7.25**	$\mathbf{0.007}$
β_4	0.00318	0.04	0.850
β_5	-0.00001	0.08	0.782
$\boldsymbol{\beta_6}$	$\mathbf{-0.04345}$	**16.00**	$\mathbf{0.000}$
$\boldsymbol{\beta_7}$	$\mathbf{0.00001}$	**3.83**	$\mathbf{0.051}$
β_8	-0.00001	0.00	0.953
β_9	0.00001	0.06	0.804
$\boldsymbol{\beta_{10}}$	$\mathbf{-0.00450}$	**3.23**	$\mathbf{0.072}$

The estimated parameter values obtained will be tested partially to determine the variables that significantly affect the model. The test results show that only three explanatory variables have a significant effect at the 5% real level on the variable number of under-five diarrhea cases per 100.000 in West Java. This can be seen from the value of the test statistics. Only three variables have test statistics greater than $\chi^2_{(0.05.1)} = 3.84$. The resulting p-values also show the same thing. However, when viewed from the p-values there are two explanatory variables that have a significant effect on the response variable at alpha 10%. Therefore, re-modeling will be carried out with variables that significantly affect the response variable both at alpha 5% and 10%.

The negative binomial regression model with five explanatory variables has a dispersion parameter (α) of 121.341. This model obtained simultaneous test statistics, namely the G-test of 42.005 greater than the value of $\chi^2_{(0.05.5)} = 11.07$, so the conclusion obtained is to reject H_0 which means that there is at least one explanatory variable that has a significant effect on the number of cases of diarrhea under five per 100.000 in West Java. The results of the parameter estimates generated from this model can be seen in Table 5.

Table 5. Parameter estimates of the negative binomial regression model with 5 explanatory variables

Parameter	Coefficient	Wald Statistic	P-value
β_0	9.56900	$1.220.59$	0.000
β_2	0.00435	7.81	0.005
β_3	-0.00478	6.11	0.013
β_6	-0.04154	13.45	0.000
β_7	0.00001	6.70	0.010
β_{10}	-0.00465	4.87	0.027

The partial test results show that all five variables significantly affect (α=5%) on the under-five diarrhea cases per 100,000 in West Java. This can be seen from the test statistic values for all variables greater than $\chi^2_{(0.05.1)} = 3.84$. The resulting p-values also show the same thing. The resulting deviance value in the negative binomial regression model with five variables is 27.027 with degree of freedom of 21. Thus, the ratio of the deviance value to the degree of freedom is obtained as 1.287. This value is close to one, much smaller than the value in the Poisson regression model. This shows that negative binomial regression can handle overdispersion in the data.

Model Interpretation

The negative binomial regression model shows that five factors affect the number of under-five diarrhea cases per 100,000 in West Java at α=5%. The factors are the percentage of public premises (PPP) meeting health requirements (X_2), percentage of Community-Based Total Sanitation (CBTS) (X_3), percentage of the population living in poverty (X_6), population density per km^2 (X_7), percentage of households practicing Clean and Healthy Behavior (CHB) (X_{10}). The model equation obtained is as follows.

$$\ln(\mu) = 9.56900 + 0.00435X_2 - 0.00478X_3 - 0.04154X_6 + 0.00001X_7 - 0.00465X_{10} \quad (7)$$

Interpretation of the negative binomial regression model obtained on the variable percentage of public premises (PPP) meeting health requirements that are significant at α = 5% on the number of cases of under-five diarrhea per 100,000 in West Java. that is, every increase in public premises (PPP) meeting health requirements by 1% will tend to increase the average number of cases of under-five diarrhea in West Java by exp(0.00435)=1.0044 times from before the rise with other variables assumed to be fixed. This is inversely proportional to the theory in general. The greater the percentage of public premises (PPP) meeting health requirements, the better in preventing health problems caused by bacteria. such as diarrhea (Kementerian Kesehatan Republik Indonesia, 2017). In other words, the more significant the percentage of public premises (PPP) meeting health requirements in West Java, the smaller the number of under-five diarrhea cases per 100,000 in West Java. However. the results of the model obtained are the opposite. This happens because an environment or a level of cleanliness that is too high will make someone easily sick. The less clean environment will make a person's body form antibodies or immunity to something that is less clean (Cicilia, 2019). The body of someone who is accustomed to a less clean environment will have good adaptability so that someone is not easily sick.

A 1% increase in Community-Based Total Sanitation (CBTS) in West Java will tend to reduce the average number of under-five diarrhea cases in West Java by exp(-0.00478)=0.9952 times from before the increase with other variables assumed to be constant. These results are in line with the results of Monica *et al*. (2020), namely the better the implementation of Community-Based Total Sanitation (CBTS) in an area, the more controlled the incidence of diarrhea will be. The implementation of Community-Based Total Sanitation (CBTS), which certainly has good provisions for health, will reduce the possibility of spreading bacteria through sanitation, especially bacteria that cause diarrhea. Therefore, the number of diarrhea cases will decrease if the Percentage of Community-Based Total Sanitation (CBTS) implementation in an area is good enough.

A 1% increase in the population living in poverty in West Java will tend to reduce the average number of under-five diarrhea cases in West Java by exp(-0.04154)=0.9593 times from before the increase with other variables assumed to be constant. These results are inversely proportional to the theory in general. According to research by Ariasih & Yuliarmi (2021), health level has a negative relationship with poverty. The higher the poverty level, the lower the health level. Supposedly, the higher the poverty rate, the higher the number of under-five diarrhea cases. However, the results show the opposite, as in Cirebon District, which has a high percentage of poverty, which is 12.3%, has a number of under-five diarrhea cases per 100,000 of 5,413, and Depok City, which has a relatively low percentage of poverty, which is 2.58%, has a number of under-five diarrhea cases per 100,000 of 10,397. This is because there is a relationship between poverty and the community's health level.

According to BPS (2023) poverty is seen as an economic inability to meet basic food and non-food needs measured in expenditure. Poverty has a negative relationship with the human development index (HDI) (Prasetyoningrum & Sukmawati, 2018). The lower the poverty in a region, the higher the HDI in that region. HDI describes the level of quality of life as well as the ability of Indonesian people to live decently (Ginting & Rasbin, 2010). Decent living can tell the level of cleanliness in the neighborhood. If many can live decently, it means that many live clean. A clean environment makes the human immune system unprepared to face immunity threats due to both bacteria and viruses (Richtel, 2019).

Any increase in population density per km^2 in a district/city in West Java will be followed by the rise in the average number of under-five diarrhea cases in West Java by exp(0.00001)=1.00001 times from before the increase with other variables assumed constant. This is by the research of Sulasih *et al*. (2021), namely if there is an increase in population density per km^2, there will be an increase in the number of under-five diarrhea cases. The more dense the population, the faster the spread of infectious diseases, such as diarrhea. This can be seen in the case of Bekasi City which has a population density of 14,795.66 per km^2 and the number of under-five diarrhea cases per 100,000 in the city is 10,165.

The raising percentage of households practicing Clean and Healthy Behavior (CHB) in West Java by 1% will tend to reduce the average number of under-five diarrhea cases in West Java by exp(-0.00465)=0.9954 times from before the increase with other variables assumed to be constant. These results are in line with the results of research by Mas *et al*. (2017) which states that the higher the percentage of households practicing Clean and Healthy Behavior (CHB), the lower the number of diarrhea cases among children under five. This happens because by living clean and healthy, will be able to reduce the breeding of bacteria or viruses that cause disease and will be able to reduce the spread of bacteria or viruses. The district/city that has a high percentage of households with clean and healthy behaviors is Banjar City, which is 80.77% with a fairly low number of under-five diarrhea cases per 100,000, which is 6,322.

CONCLUSION

This study result represent that there are five factors affect the number of children under five diarrhea cases in West Java. Factors that positively influence the number of under-five diarrhea cases per 100.000 in West Java are the percentage of public premises (PPP) meeting health requirements and population density per km^2. Meanwhile, the factors that have a negative effect are the percentage of Community-Based Total Sanitation (CBTS), the percentage of the population living in poverty, and the percentage of households practicing Clean and Healthy Behavior (CHB). This study also show that the negative binomial regression model has a ratio of deviance value to degree of freedom close to one. This indicates that negative binomial regression can handle overdispersion in count data.

REFERENCE

Ariasih, N. L. M., & Yuliarmi, N. N. (2021). Pengaruh Tingkat Pendidikan, Tingkat Kesehatan dan Pengangguran Terbuka Terhadap Tingkat Kemiskinan di Provinsi Bali. *Cerdika: Jurnal Ilmiah Indonesia*, *1*(7), 821–839. https://doi.org/10.59141/cerdika.v1i7.131

BPS. (2023). *Kemiskinan dan Ketimpangan*. Badan Pusat Statistik. https://www.bps.go.id/subject/23/kemiskinan-dan-ketimpangan.html

Cahyani, S. D., Poerwoningsih, D., & Wahjutami, E. L. (2019). Konsep Hunian Adaptif Sebagai Upaya Penanganan Rumah Tinggal Tidak Layak Huni Terhadap Resistensi Penyakit Infeksi. *Mintakat: Jurnal Arsitektur*, *20*(2), 79–91. https://doi.org/10.26905/mj.v20i2.3800

Cicilia, M. (2019, June 18). Terlalu Higienis, Anak Malah Mudah Sakit. *Antara*. https://www.antaranews.com/berita/917744/terlalu-higienis-anak-malah-mudah-sakit

Dinas Kesehatan Jawa Barat. (2022). Profil Kesehatan Jawa Barat Tahun 2021. In *Dinas Kesehatan Provinsi Jawa Barat*. Dinas Kesehatan Jawa Barat.

Faidah, F., & Wibowo, W. (2017). *Analisis Faktor Risiko yang Mempengaruhi Pneumonia Pada Balita Di Kabupaten Bangkalan Menggunakan Regresi Poisson*. Institut Teknologi Sepuluh November.

Fitrial, N. H., & Fatikhurrizqi, A. (2020). Pemodelan Jumlah Kasus Covid-19 Di Indonesia Dengan Pendekatan Regresi Poisson Dan Regresi Binomial Negatif. *Seminar Nasional Official Statistics*, *2020*(1), 65–72. https://doi.org/10.34123/semnasoffstat.v2020i1.465

Fuad, M. S., Kudus, A., & Sunendiari, S. (2018). Pemodelan Data Kecelakaan Lalu Lintas Menggunakan Metode Regresi Poisson dan Regresi Binomial Negatif. *SPeSIA-UNISBA* , 26–33.

Ginting, A. M., & Rasbin. (2010). Pengaruh pertumbuhan ekonomi terhadap tingkat kemiskinan di Indonesia sebelum dan setelah krisis. *Jurnal Ekonomi & Kebijakan Publik*, *2*(1), 279–312.

Hardin, J. W., & Hilbe, J. M. (2014). Regression models for count data based on the negative binomial(p) distribution. *Stata Journal*, *14*(2), 280–291. https://doi.org/10.1177/1536867x1401400203

Kementerian Kesehatan Republik Indonesia. (2017). Profil Kesehatan Indonesia 2016. In R. Kurniawan, Yudiarto, B. Hardhana, & T. A. Soenardi (Eds.), *Profil Kesehatan Indonesia* (First Edit). Kementerian Kesehatan Republik Indonesia. http://www.depkes.go.id/resources/download/pusdatin/profil-kesehatan-indonesia/Profil-Kesehatan-Indonesia-2016.pdf

Kementerian Kesehatan Republik Indonesia. (2022). Profil Kesehatan Indonesia 2021. In F. Sibuea, B. Hardhana, & W. Widiantini (Eds.), *Profil Kesehatan Indonesia*. Kementerian Kesehatan Republik Indonesia.

Lestari, N. K., Hikmawan, D. W., Safitrah, I., & Widodo, E. (2019). Analisis Faktor Lingkungan Penyebab Kasus Diare Di Jawa Barat Tahun 2016 Menggunakan Regresi Binomial Negatif. *Prosiding-M19*.

Lestari, S. P., & Wulandari, S. P. (2014). Pemodelan Faktor-Faktor yang Mempengaruhi Jumlah Kasus Tetanus Neonatorum (TN) di Jawa Timur dengan Metode Regresi Zero- Inflated Generalized Poisson (ZIGP). *Sains Dan Seni Pomits*, *3*(2), 116–121.

Mas, E. M., Yudiernawati, A., & Maemunah, N. (2017). Hubungan Perilaku Hidup Bersih Dan Sehat (PHBS) Ibu Dengan Kejadian Diare Pada Anak Balita (1-5 Tahun) Di Posyandu Mawar Kelurahan

Merjosari Wilayah Puskesmas Dinoyo Kota Malang. *Journal Nursing News*, *2*(1), 31–37.

Melvani, R. P., Zulkifli, H., & Faizal, M. (2019). Analisis Faktor Yang Berhubungan Dengan Kejadian Diare Balita Di Kelurahan Karyajaya Kota Palembang. *JUMANTIK (Jurnal Ilmiah Penelitian Kesehatan)*, *4*(1), 57–68. https://doi.org/10.30829/jumantik.v4i1.4052

Monica, D. Z., Ahyanti, M., & Prianto, N. (2020). Hubungan Penerapan 5 Pilar Sanitasi Total Berbasis Masyarakat (STBM) Dan Kejadian Diare Di Desa Taman Baru Kecamatan Penengahan Kabupaten Lampung Selatan. *Ruwa Jurai: Jurnal Kesehatan Lingkungan*, *14*(2), 71–77. https://doi.org/10.26630/rj.v14i2.2183

Nurhayati, I. C., Rusgiyono, A., & Yasin, H. (2018). Robust Geographically Weighted Regression Dengan Metode Mutlak Simpangan Terkecil Pada Pemodelan Kejadian Diare Di Kota Semarang. *Jurnal Gaussian*, *7*(2), 143–152. https://doi.org/10.14710/j.gauss.v7i2.26646

Prasetyoningrum, A. K., & Sukmawati, U. S. (2018). Analisis Pengaruh Indeks Pembangunan Manusia (IPM), Pertumbuhan Ekonomi dan Pengagguran Terhadap Kemiskinan di Indonesia. *EQUILIBRIUM: Jurnal Ekonomi Syariah*, *6*(2), 217–240. https://www.academia.edu/download/76624968/pdf.pdf

Richtel, M. (2019, March 12). Your Environment Is Cleaner. Your Immune System Has Never Been So Unprepared. *The New York Times*. https://www.nytimes.com/2019/03/12/health/immune-system-allergies.html

Sidqi, D. N. S., Anasta, N., & Mufidah, P. K. (2021). Analisis Spasial Kasus Diare pada Balita di Kabupaten Banyumas Tahun 2019. *Jurnal Biostatistik, Kependudukan, Dan Informatika Kesehatan*, *1*(3), 135. https://doi.org/10.51181/bikfokes.v1i3.4920

Sulasih, I. G. A. D., Susilawati, M., & Suciptawati, N. L. P. (2021). Pemodelan Kasus Diare Di Provinsi Bali Dengan Metode Analisis Regresi Spasial. *E-Jurnal Matematika*, *10*(2), 95–102. https://doi.org/10.24843/mtk.2021.v10.i02.p327

UNICEF. (2024). *Diarrhoea*. Unicef Data. https://data.unicef.org/topic/child-health/diarrhoeal-disease/

Utami, T. W. (2013). Analisis Regresi Binomial Negatif Untuk Mengatasi Overdispersion Regresi Poisson Pada Kasus Demam Berdarah Dengue. *Jurnal Statistika*, *1*(2), 59–65. https://jurnal.unimus.ac.id/index.php/statistik/article/view/961

WHO. (2017). *Diarrhoeal disease*. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease

Widyaningrum, A. R., Susanti, Y., & Slamet, I. (2021). Pemodelan Penyakit Diare Balita Di Jawa Timur Menggunakan Regresi Robust. *Prosiding Seminar Nasional Sains*, *2*(1), 522–528.

Yulianingsih, K. A., Sukarsa, K. G., & Suciptawati, L. P. (2012). Penerapan Regresi Poisson Untuk Mengetahui Faktor-Faktor Yang Memengaruhi Jumlah Siswa SMA/SMK yang Tidak Lulus UN di Bali. *E-Jurnal Matematika*, *1*(1), 59–63.