

MODELING THE NUMBER OF MULTIBACILLARY LEPROSY USING NEGATIVE BINOMIAL REGRESSION TO OVERCOME OVERDISPERSION IN POISSON REGRESSION

Fitri Anugerahi Wibisono¹, Eva R. Kurniawati²

¹Department of Biostatistics and Population, Faculty of Public Health, Universitas Airlangga, 60115 Surabaya, East Java, Indonesia

²Department Cardiothoracic Surgery, Maastricht University, Medical Center+, Maastricht, The Netherlands

Corresponding Author: Fitri Anugerahi Wibisono

E-mail: fitrianugurkani@gmail.com

ABSTRACT

Poisson regression is used on discrete data (count) for the formation of the model. There is often a violation in Poisson regression analysis assumptions i.e., overdispersion, which means the average value of the data is smaller than the value of the variance. The number of multibacillary leprosy (MB) in 31 Surabaya districts orderly from 2015 to 2017 has increased as many as 127 cases, 140 cases, and 158 cases. This study aimed to model the number of MB leprosy in Surabaya in 2017 with a Negative Binomial regression in overdispersion. This was quantitative research with a descriptive method that uses secondary data. The data sourced from Surabaya City Health Profile in 2017. The independent variables studied include BCG immunization coverage, the percentage of healthy houses, the percentage of Households with Clean and Healthy Behavior (HCHB), the percentage of the male population, and the population density level. MB leprosy incidence modeling with Poisson regression proved to be overdispersed so that the Negative Binomial regression was used to overcome it. The variable that influenced the MB leprosy incidence with a Negative Binomial regression analysis was the percentage of healthy houses ($p = 0.019$). MB leprosy occurrence will decrease if the percentage of healthy houses increases. The percentage of healthy houses in Surabaya was 86.99%, which increased compared to the previous year with an increase of 1.78%. Public awareness about healthy houses is required to reduce the number of MB leprosy in Surabaya.

Keywords: poisson regression, overdispersion, negative binomial regression, multibacillary leprosy

ABSTRAK

Regresi Poisson digunakan pada data diskrit (count) untuk pembentukan modelnya. Pada kenyataannya sering kali terjadi pelanggaran asumsi pada analisis regresi Poisson yaitu terjadi overdispersi, yang artinya nilai rata-rata data lebih kecil dari nilai variannya. Jumlah angka kejadian kusta Multi Basiler (MB) di 31 kecamatan Kota Surabaya berturut-turut pada tahun 2015-2017 mengalami peningkatan, yaitu sebesar 127 kasus, 140 kasus, dan 158 kasus. Tujuan penelitian ini adalah untuk memodelkan jumlah kusta MB di Kota Surabaya tahun 2017 dengan regresi Binomial Negatif dalam mengatasi overdispersi. Penelitian ini merupakan penelitian kuantitatif dengan metode deskriptif yang menggunakan data sekunder. Data yang digunakan berasal dari Profil Kesehatan Kota Surabaya tahun 2017. Variabel independen yang diteliti meliputi cakupan imunisasi BCG, persentase rumah sehat, persentase rumah tangga ber-Perilaku Hidup Bersih dan Sehat (PHBS), persentase penduduk laki-laki, dan tingkat kepadatan penduduk. Pemodelan angka kejadian kusta MB dengan regresi Poisson terbukti mengalami overdispersi sehingga regresi Binomial Negatif digunakan dalam mengatasinya. Variabel yang berpengaruh terhadap angka kejadian kusta MB dengan analisis regresi Binomial Negatif adalah persentase rumah sehat ($p = 0,019$). Angka kejadian kusta MB akan semakin menurun apabila persentase rumah sehat semakin meningkat. Persentase rumah sehat di Kota Surabaya sebesar 86,99%, yang mana meningkat dibandingkan dengan tahun sebelumnya dengan peningkatan sebesar 1,78%. Kesadaran masyarakat akan rumah yang sehat diperlukan untuk menurunkan jumlah kusta MB di Kota Surabaya.

Kata kunci: regresi poisson, overdispersi, regresi binomial negatif, kusta multi basiler

Received: September 23, 2019

Accepted: November 6, 2019

INTRODUCTION

The dependent variable can consist of count data in its regression analysis, which is

not always in the form of continuous data. The right regression to do data count analysis is Poisson regression. The conjecture that must be

fulfilled in Poisson regression is the similarity of values in the variance and the average data (dispersion). The possibility of fulfilling the liquidation assumption is minimal because overdispersion often occurs in the data count (Myers, 1990).

Poisson regression application, in fact, often violate the Poisson regression analysis assumptions. These violations include the higher value of the variance called overdispersion and the more excellent average value, also called underdispersion (Wang and Famoye, 2014). Overdispersion can be caused by a positive reciprocal relationship between the independent variables or the value of variance on a measureable independent variable. The existence of a first event that affects the next event can also cause overdispersion. In Poisson regression analysis, the estimated standard error in overdispersed data can be underestimated if it continues to use it. A variable parameter that can be known as significant but not significant is another impact of Poisson regression modeling that has a case of overdispersion (Hilbe, 2014).

A reasonably high error rate in overdispersion can result in a less useful Poisson regression modeling. The right alternative in overcoming the overdispersion that often occurs in Poisson regression modeling is the Negative Binomial distribution to replace the Poisson distribution. The value $\alpha = 0$ is a parameter value as a particular form of the Negative Binomial distribution. Hence, Negative Binomial distribution can be used as an alternative. The Negative Binomial regression model is formed from the distribution of the Negative Binomial distribution where the distribution does not emphasize the existence of an equidispersion assumption and the assumptions in the Poisson regression (Cameron and Trivedi, 2013).

World Health Organization noted that in 2017, the global prevalence of leprosy registered was 192,713 cases (0.25/10,000 population), an increase of 20,765 cases compared to 2016. Southeast Asia continued to be ranked highest in reporting new leprosy cases until 2017, which was 153,487 cases out of total reporting new cases in the world amounted to 210,671 cases (World Health Organization, 2014).

The number of new cases of leprosy recorded by the Republic of Indonesia Annual Health Report in 2017 was 15,910 cases, with

the percentage of multibacillary leprosy or wet leprosy of 86.12% of cases. East Java Provincial Health Office (2017) recorded the prevalence of leprosy cases of 1.04 per 10,000 people in 2016 and increased in 2017 to 1.06 per 10,000 people. The total number of multibacillary leprosy (MB) cases in 31 sub-districts in Surabaya that has been recorded in the Surabaya Health Profile in 2015, 2016 and 2017 has increased as many as 127 cases, 140 cases and 158 cases (Surabaya City Health Office, 2017).

Based on the previous description by analyzing Surabaya as a city in East Java with the number of MB leprosy that continues to increase every year. The research that will be conducted is modeling the number of MB leprosy in Surabaya in 2017 with Negative Binomial Regression because researchers suspect an overdispersion case in Poisson regression.

METHODS

This study used quantitative research data analysis with descriptive method that use secondary data. The secondary data was retrieved from Surabaya Health Profile in 2017. Population data recorded multibacillary leprosy patients in Surabaya in 2017 was the subject of research divided into each district. All sub-districts in Surabaya were used as research settings, and this research was conducted in July 2019.

The total population of those patients from multibacillary leprosy in Surabaya in 2017 was 158 population. All population was included as a sample in this study. The variables in this study were data on multibacillary leprosy incidents (Y), BCG immunization coverage (X_1), percentage of healthy houses (X_2), percentage of households having HCHB (X_3), percentage of the male population (X_4), and level of population density (X_5).

The step of analysis carried out in this research was to describe the characteristics of multibacillary leprosy incidence rate and its factors. Testing multicollinearity cases must be conducted first based on the correlation criteria and VIF, to determine whether among the variables that affect the incidence of MB leprosy do not have a great reciprocal relationship. The test was continued with Poisson regression modeling if there was no case of multicollinearity. In Poisson regression,

Table 1. Descriptive Statistics

Variable	N	Average	Variance	Minimum	Maximum
Y	31	5.10	29.49	0	21
X ₁	31	100.26	445.47	78.82	198.57
X ₂	31	87.05	58.31	63.40	98.22
X ₃	31	72.63	194.33	30.47	96.96
X ₄	31	49.51	0.224	49.29	51.95
X ₅	31	11882.69	71795753.82	2293.88	37324.32

the occurrence of overdispersion was known by calculating the ratio value (Pearson Chi-Square value/degrees of freedom); if it was proven as overdispersion, then the analysis was continued being tested using Negative Binomial Regression (Wulandari, 2015). The final step taken was to determine the best model to determine which model was appropriate in modeling the occurrence of multibacillary leprosy in Surabaya by looking at the AIC value and the log-likelihood value (Pradawati and Sukarsa, 2013).

RESULT

Descriptive statistical calculations (Table 1) were used to see the characteristics of each variable. The occurrence of multibacillary leprosy in Surabaya in 2017 was recorded as 158 cases. Based on table 1, the highest distribution of multibacillary leprosy incidents in 2017 for each sub-district in Surabaya occurred in Kenjeran with a total of 21 cases. Bubutan, Tenggilis, Dukuh Pakis, and Wonocolo reported to be the lowest case, 0 case or there was no multibacillary leprosy case in 2017.

Multicollinearity Test

The multicollinearity test was performed as a first step towards the analysis phase with the Poisson regression method and the Negative Binomial regression. Multicollinearity test was conducted to prove that the independent variable did not have a great reciprocal relationship by taking into account the value of VIF (Variance Inflation Factor) and the Tolerance value. Tolerance value < 0.10 and VIF > 10 indicate high reciprocity in the independent variables.

Based on Table 2, the Tolerance value for all variables > 0.10 and VIF value < 10.0 means that there was no case of multicollinearity in the

regression model, or there was no excellent reciprocal relationship between independent variables. Therefore, Poisson regression modeling and Negative Binomials can be done.

Table 2. Multicollinearity Testing

Variable	Tolerance Value	VIF
X ₁	0.960	1.041
X ₂	0.938	1.066
X ₃	0.984	1.016
X ₄	0.972	1.029
X ₅	0.893	1.119

Poisson Regression

The incidence of multibacillary leprosy includes the count data. Therefore, Poisson regression used as a model of its formation. Poisson regression modeling was initiated with simultaneous testing of all model parameters. Concurrent test results were known to have a p-value of 0,000, which was no more than a significance level of 0.2, meaning that there was at least one variable that affects the number of MB leprosy.

Partial testing of Poisson regression model parameters was conducted to find out which were the influential variables. The conclusion was obtained by looking at the p-value and the significance level of 20%, which was 0.2. Based on Table 3, it can be seen that the variable with p-value > 0.2 was the percentage of the male population (X₄), so re-testing was required without including the insignificant variables.

Based on Table 4, it can be seen that the p-value < 0.2 . The results of partial testing form a decision with at least one variable affected the number of MB leprosy. Variable coverage of BCG immunization (X₁), percentage of healthy houses (X₂), percentage of households having HCHB (X₃), and population density (X₅) were stated to have a significant effect on MB leprosy

Table 3. Partial Testing of Poisson Regression Model Parameters

Parameter	B	Std.Error	80% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig
(Intercept)	12.806	9.4014	0.758	24.855	1.855	1	0.173
X ₁	0.005	0.0030	0.002	0.009	3.272	1	0.070
X ₂	-0.059	0.0099	-0.072	-0.046	35.867	1	0.000
X ₃	-0.012	0.0063	-0.020	-0.004	3.384	1	0.066
X ₄	-0.124	0.1893	-0.366	0.199	0.426	1	0.514
X ₅	2.068E-5	8.5847E-6	9.680E-6	3.168E-5	5.804	1	0.016

Table 4 Partial Testing of Poisson Regression Model Parameters without X₄

Parameter	B	Std.Error	80% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig
(Intercept)	6.720	1.0752	5.343	8.098	39.069	1	0.000
X ₁	0.005	0.0030	0.002	0.009	3.277	1	0.070
X ₂	-0.059	0.0099	-0.072	-0.047	36.294	1	0.000
X ₃	-0.012	0.0064	-0.020	-0.004	3.438	1	0.064
X ₅	2.034E-5	8.5970E-6	9.318E-6	3.135E-5	5.595	1	0.018

incidents. The following can be written as Poisson regression model:

$$(\hat{\mu}) = \exp[6.720 + 0.005(X_1) - 0.059(X_2) - 0.012(X_3) + 2.034E - 5(X_5)]$$

The interpretation of the Poisson regression model above was that every 1% increase in BCG (X₁) immunization coverage, the number of MB leprosy incidents will increase by $e^{(0.005)} = 1.005 \approx 1$ case if it was constantly under the assumptions of other variables. Increasing the percentage of healthy houses (X₂) every 1% assuming other variables were constant, the number of MB leprosy was reduced to $e^{(0.059)} = 1.006 \approx 1$ case. MB leprosy incidence rate decreased to $e^{(0.012)} = 1.012 \approx 1$ case, every 1% increase in the percentage of households having HCHB (X₃) if the assumption of other variables was constant. MB leprosy incidence rate will increase to $e^{(2.034E-5)} = 1$ case, each additional level of population density (X₅) assuming other variables were constant.

Overdispersion

Overdispersion in Poisson regression can be determined through the Pearson Chi-Square value degree of freedom. If the value > 1 , it can be concluded that there was an overdispersion in the data. Pearson Chi-Square Poisson regression model of MB leprosy incidents was

104.898 and the degree of freedom was 26, so we get a result of 4.035 which can be concluded that overdispersion occurred. The overdispersion results in the model being formed, which resulted in an estimated parameter bias. Modeling with Negative Binomial Regression was needed to overcome the overdispersion in the Poisson regression model of MB leprosy incidents.

Negative Binomial Regression

The distribution that was often used in overdispersion violations was Negative Binomial. Negative Binomial Regression was used in modeling by involving all significant variables in Poisson regression, namely BCG immunization coverage (X₁), percentage of healthy houses (X₂), percentage of households having HCHB (X₃), and population density (X₅).

The initial step in modeling Negative Binomial regression was the standardization of data. Data standardization was carried out if the data used in this study had different units, namely for the BCG immunization coverage variable (X₁), the percentage of healthy houses (X₂), and the percentage of households having HCHB (X₃). At the same time, the population density (X₅) was the unit of live/km².

Table 5. Partial Testing of Negative Binomial Regression Model Parameters $Y = f(X_1, X_2, X_3, X_5)$

Parameter	B	Std.Error	80% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig
(Intercept)	16.483	14.1511	-1.653	34.618	1.357	1	0.244
X ₁	0.580	1.1637	-0.911	2.072	0.249	1	0.618
X ₂	-4.577	2.2932	-7.516	-1.639	3.984	1	0.046
X ₃	-0.066	1.2033	-1.608	1.476	0.003	1	0.956
X ₅	0.333	0.3051	-0.058	0.724	1.194	1	0.275

Table 6. Partial Testing of Negative Binomial Regression Model Parameters $Y = f(X_2)$

Parameter	B	Std.Error	80% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig
(Intercept)	7.161	2.3375	4.165	10.157	9.385	1	0.002
X ₂	-0.065	0.0270	-0.099	-0.030	5.804	1	0.016

Negative Binomial regression modeling was started with simultaneous testing of significant variables in Poisson regression. The p-value of the simultaneous test results was 0.114 or less than the significance level of 0.2 so that there was at least one variable that affects the number of MB leprosy.

The partial test results were known to form a decision that at least one variable influences the MB leprosy incidents. Based on Table 6, the percentage of healthy houses (X₂) variable was stated to have a significant effect on the number of MB leprosy, so that re-testing was needed without including insignificant variables, namely X₁, X₃, and X₄ (Table 6). The following can be written for Negative Binomial Regression model:

$$(\hat{\mu}) = \exp[7.161 - 0.065(X_2)]$$

The interpretation of the Negative Binomial regression model that has been arranged was an increase of 1% per percentage of healthy houses (X₂), the number of MB leprosy was reduced by $e^{(-0.065)} = 1.067 \approx 1$ case, assuming other variables were constant.

Determination of the Best Model

In order to determine the best model to use, AIC (*Akaike's Information Criterion*) and log-likelihood values were applied. AIC was a relative measure of the goodness of fit of the statistical model to get the best regression

model. AIC was a measure of the quality estimates of each available statistical model because it was interconnected with each other for a particular data set, making it an ideal method for model selection. The model with the smallest AIC value was said to be better in modeling the data, while the log-likelihood value was to look at the log-likelihood of each model where the model with the most considerable log-likelihood value was said to be better in modeling the data.

Based on the analysis that has been done, it was known that from the two models, the value of AIC is 207.797. Poisson regression was higher than the AIC value of Negative Binomial regression (166.621). The log-likelihood value of the Negative Binomial regression model was stated to be better in modeling the number of multibacillary leprosy incidents for each sub-district in Surabaya in 2017, with a value of -81.311 which was known to be smaller than the Poisson regression model.

DISCUSSION

The MB leprosy incidents variance value of 29.49 is considered significant because it is caused by the gap, which is also quite large, resulting from the difference between the lowest and highest MB leprosy incidence rates. The significant variations between sub-districts in the MB leprosy incidents in Surabaya are indicated by the value of a significant dependent variable variance. Data overdispersion is

determined if the average value of data is smaller than the data variant (Hilbe, 2014).

The independent variables included in the modeling of MB leprosy incidents did not produce any Tolerance value < 0.10 and VIF > 10 , it is known that there is no reciprocal relationship between independent variables. Estimated parameters in the regression will produce an error so that positive (+) or negative (-) signs contained in the coefficient of the regression model conflict with the theory if the reciprocal relationship between the independent variables occurs in regression modeling (Widarjono, 2007). According to Asnawi and Wijaya (2005), the higher the covariance and variance, the greater the standard error so that the higher the confidence interval is a consequence of multicollinearity.

Negative Binomial Regression in Overcoming Overdispersion

The assumption that must be fulfilled in Poisson regression is the mean of the mean value and variance on the independent variable, also called equidispersion. Testing is enough to administer with Poisson regression if the equidispersion assumptions have been met. The fulfillment of equidispersion assumptions can be done by looking at the value of deviance or df or Pearson chi-square value or degrees of freedom to detect it. Pearson chi-square value or degrees of freedom MB leprosy incidence has a value of more than one, so it can be said that there has been a case of overdispersion (Maziyah, 2018). According to McCullagh and Nelder (1989), if a population under study occurs grouping, and there are many zero values, it can lead to overdispersion.

According to Hilbe (2014), if the overdispersion case is still used in Poisson regression, it can cause an insignificant or invalid variable. The method using Negative Binomial regression was done as an effort to overcome this. The Poisson distribution and Gamma distribution are the origins of the idea in developing the Negative Binomial distribution. In the Poisson regression assumption, there is equal value in the value of the variance and the average value if the value of the dispersion parameter is zero in the Negative Binomial regression (Cameron and Trivedi, 2013).

Modeling the number of MB leprosy incidence involving five factors, four of which

affect the Poisson distribution. The number of factors that are considered influential on the incidence of MB leprosy is one of the effects to the overdispersion occurrence.

Relationship between BCG Immunization Coverage with MB Leprosy Incidence Rate

One strain of weakened *Mycobacterium bovis* can form a vaccine called *Bacille Calmette Guerin* (BCG). The BCG vaccine as a form of disease prevention caused by *Mycobacterium tuberculosis* (TB). BCG said to be a protective force against leprosy is a conjecture that emerged in the late 1930s.

Mycobacterium leprosy incidents and BCG immunization from negative Binomial modeling results are known to have no significant relationship. This result is different from the results of research in Malawi in 1996. Giving one dose of BCG vaccination can protect against leprosy by 50% and an increase of 0.6 times each increase of one dose of BCG vaccination. According to Hariyadi (2010), with BCG immunization, 20-80% of a person will be protected from exposure to leprosy clinical symptoms in various places.

Relationship Percentage of Healthy Houses to MB Leprosy Incidence Rate

The house as a place to live and take shelter is a basic need for everyone. A healthy house is a condition where both the area and the environment around the house allow residents and the surrounding community to obtain optimal health degree. Having a house that is healthy, clean and comfortable will increase productivity for each individual.

The percentage of whole houses in the Negative Binomial model that increases will reduce the number of MB leprosy as much as 0.004 times if it occurs continuously on the assumption of other variables. In line with the existing theory that if the house meets the requirements of a healthy house, indirectly, the environmental conditions and sanitation are better.

According to Sya'diana (2018), the chances of people who stay in the house with the condition of the house far from healthy, have a risk of 7,875 times higher than the ratio of people who live in the house with the condition of the house that meets the recommended health requirements. In line with Sya'diana, according to Norlatifah, Sutomo and Solikhah (2010), the

chance of someone who settled with the physical condition of the house far from healthy can be attacked by leprosy 3.169 times higher than the ratio of someone who settled in a healthy house.

Relationship Percentage of HCHB Households to MB Leprosy Incidence Rate

The risk of an illness can be prevented, one of them is by empowering family members. Clean and Healthy Behavior (CHB) is an effort that is expected to encourage family members to understand, aim, and able to carry out the community health movement. Realizing an increase in health status requires each individual to practice many of the behaviors covered by HCHB. Based on the Negative Binomial Regression test, the percentage of HCHB households did not significantly relate to the number of MB leprosy in Surabaya.

This result does not follow Hendrik L. Blum's theory, which explains that the environment, behavior, health services, and heredity are four factors that affect public health. One of the influences in the environment and behavior. Mainly behavior in preventing the spread of diseases caused by a less healthy house environment. According to Rismawati (2013), poor house sanitation can cause leprosy, especially multibacillary types.

Relationship Percentage of Male Population to MB Leprosy Incidence Rate

The highest number of the male population was in Sawahan (sub-district). Based on the number of MB leprosy incidents in Surabaya, the number of men patients was higher than women even though MB leprosy generally can infect both women and men.

In the Negative Binomial model, the percentage of the male population stated no significant relationship to the number of MB leprosy. This result was different from research in most parts of the world: women are less exposed to leprosy than men, except for a few countries in Africa. According to Juniardi (2015) in his research, for every 1% increase in the male population, leprosy will increase by 2 cases. Environmental and sociocultural factors are likely to be known as factors causing the low incidence rate of leprosy in women. That is because women's access to health services is very limited to specific cultures.

Relationship of Population Density Level to MB Leprosy Incidence Rate

Population density can have impacts, including social, economic, and health problems. The slightly more dense population of houses indirectly results in slums and poor sanitation, which can cause disease, especially infectious diseases, including MB leprosy.

Population density level in the Negative Binomial model showed to have no significant relationship to the number of MB leprosy in Surabaya City. In contrast to existing research, according to Ulfah (2012), population density which requirements do not meet health standard. Is proved to be one of the causes of the transmission of leprosy to healthy people who live at houses with leprosy.

The highest population density in Surabaya was in Simokerto (sub-district). The existence of the highest number can be influenced by several factors such as the high birth rate when compared to the mortality rate, the number of people moving to Simokerto, population mobility, and the characteristics of the Simokerto itself. Increasing population density will cause an imbalance between the population and the environment and the facilities available to solve the problem of expanding settlements to health services.

CONCLUSIONS AND SUGGESTIONS

Conclusion

MB leprosy incidence rate data can be proven that true overdispersion was known where the Pearson Chi-Square or degrees of freedom $(4,035) > 1$. The Negative Binomial Model was an alternative in dealing with overdispersion by producing one significant variable, namely the percentage of healthy houses (X_2) with the Negative Binomial regression model as followed:

$$(\hat{\mu}) = \exp[7,161 - 0,065(X_2)]$$

Suggestion

Based on the conclusions above, it is necessary to increase public awareness about the importance of maintaining a healthy house. In addition, Surabaya City Health Office can increase efforts to reduce the number of MB leprosy incidents by conducting counseling and other programs related to the variable of healthy houses.

REFERENCES

- Asnawi, K. and Wijaya, C., 2005. *Riset Keuangan: Pengujian-Pengujian Empiris*. Jakarta: Gramedia Pustaka Indonesia.
- Cameron, A.C. and Trivedi, P.K., 2013. *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.
- Hariyadi, P., 2010. Hubungan Status Vaksinasi BCG dengan Kejadian Kusta di Kabupaten Brebes Provinsi Jawa Tengah. *Thesis*. Universitas Gajah Mada.
- Hilbe, J.M., 2014. *Modeling Count Data*. Cambridge: Cambridge University Press.
- Juniardi, L.C., 2015. Analisis Faktor-Faktor yang Mempengaruhi Jumlah Kasus Kusta di Jawa Timur pada Tahun 2013 Menggunakan Geographically Weighted Negative Binomial Regression (GWNBR). *Jurnal Sains dan Seni ITS*, 4(1), pp.55–60.
- Maziyah, A.I. 2018. Pemodelan Generalized Poisson Regression (GPR) dalam Mengatasi Pelanggaran Equidispersi Pada Regresi Poisson Kasus Kusta di Kabupaten Bangkalan 2016. *Thesis*. Universitas Airlangga.
- McCullagh, P. and Nelder, J.A., 1989. *Generalized Linier Models*. Second ed. London: Chapman & Hall.
- Myers, R.H., 1990. *Classical and Modern Regression with Applicaton*. Boston: PWS-KENT.
- Norlatifah, Sutomo, A.H. and Solikhah, 2010. Hubungan Kondisi Fisik Rumah, Sarana Air Bersih dan Karakteristik Masyarakat dengan Kejadian Kusta di Kabupaten Tapin Kalimantan Selatan. *Jurnal KesMas FKM UAD*, 4(3), pp.182–195.
- Pradawati, P.S. and Sukarsa, K.G., 2013. Penerapan Regresi Binomial Negatif untuk Mengatasi Overdispersi pada Regresi Poisson. *E-Jurnal Matematika*, 2(2), pp. 6–10.
- Rismawati, D., 2013. Hubungan antara Sanitasi Rumah dan Personal Hygiene dengan Kejadian Kusta Multibasiler. *Unnes Journal of Public Health*, 2(1), pp.1–10.
- Surabaya City Health Office, 2017. *Profil Kesehatan Kota Surabaya Tahun 2017*. Surabaya.
- Sya'diana, U., 2018. Hubungan Kondisi Fisik Rumah dan Personal Hygiene dengan Kejadian Kusta di Wilayah Ponorogo. *Thesis*. Stikes Bhakti Husada Mulia Madiun.
- Ulfah, F., 2012. Faktor Kondisi Fisik Rumah dan Kepadatan Penghuni Hubungannya dengan Kejadian Kusta (Studi di Wilayah Kecamatan Talango Kabupaten Sumenep). *Thesis*. Universitas Airlangga.
- Wang, W. and Famoye, F., 1997. Modeling Household Fertility Decisions with Generalized Poisson Regression. *Journal of Population Economics*, 10(3), pp.273–283.
- World Health Organization, 2014. Relevé Épidémiologique Hebdomadaire. *Weekly Epidemiological Record*. 89(36), pp. 389–400.
- Widarjono, A., 2007. *Ekonometrika: Teori dan Aplikasi untuk Ekonomi dan Bisnis*. 2nd ed. Yogyakarta: Ekonisia Fakultas Ekonomi Universitas Islam Indonesia.
- Wulandari, I.R., 2015. Pemodelan dengan Pendekatan Regresi Binomial Negatif untuk Mengatasi Overdispersi Pada Regresi Poisson. *Thesis*. Universitas Airlangga.