

# Application of Spatial Error Regression Method for Modeling The Percentage of Poor Population in Indonesia

Ameliatul 'Iffah<sup>1</sup>, Suliyanto<sup>2</sup>, Sediono<sup>3</sup>, Toha Saifudin<sup>4</sup>, Dita Amelia<sup>5</sup>, Elly Ana<sup>6</sup>  
<sup>1,2,3,4,5,6</sup> Statistics Study Program, Department of Mathematics, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia

<sup>1</sup>ameliatul.iffah-2019@fst.unair.ac.id

<sup>2</sup>Corresponding author: suliyanto@fst.unair.ac.id

<sup>3</sup>ahmad.s@fst.unair.ac.id

<sup>4</sup>toha.saifudin@fst.unair.ac.id

<sup>5</sup>dita.amelia@fst.unair.ac.id

<sup>6</sup>elly-a@fst.unair.ac.id

\*\*\*\*\*

## Abstract:

Poverty is one of the main problems still faced by various countries in the world. Poverty is a multidimensional problem so there are many factors that affect poverty including health, environment, education, and economy. This study aims to identify factors that affect poverty in Indonesia using classical regression and spatial error regression which will later select the best model. The best model will be based on the smallest Akaike Information Criterion (AIC) value and the biggest coefficient of determination. The variables used in the study are the percentage of poor people as the response variable and the predictor variables include life expectancy, percentage of households that have access to proper sanitation services, GRDP rate, and expected years of schooling. The results of the study show that the spatial error regression model is the best model. This shows that poverty in a region is influenced by poverty around the region. This study also found that the variables of sanitation quality, GRDP rate, and expected years of schooling have a significant influence on the percentage of poor people, while life expectancy does not have a significant influence on the percentage of poor people. Therefore, to reduce poverty, it is necessary to pay attention to sanitation, education, and economic factors.

**Keywords —Poverty, Classical Regression, Spatial Error Regression**

\*\*\*\*\*

## I. INTRODUCTION

Poverty is one of the main problems still faced by various countries in the world. The problem of poverty is related to various sectors of life such as education, health, environment, economy and other sectors so that to overcome poverty requires continuous improvement from related sectors [1]. Efforts to overcome poverty in the world have even become a major goal agreed upon in the Sustainable Development Goals (SDGs) [2]. Poverty is a

condition where people experience a lack of income and experience economic difficulties to meet minimum living standards [3]. As one of the countries that strives to reduce poverty, Indonesia implements several programs such as providing protection and affirmation through social assistance, social security, subsidizing basic goods, improving public services, creating jobs aimed at the poor, and creating economic growth that is able to expand employment opportunities in the formal sector [4]. However, based on data recorded as of September

2022, the poverty rate in Indonesia is 9.57% or 26.36 million people experiencing poverty, which means that the poverty rate has increased compared to March 2022 which amounted to 9.54%. Poverty in Indonesia experienced a slight increase in both urban and rural areas [5]. Therefore, a study on identifying factors that influence poverty is needed as an effort to reduce poverty in Indonesia.

Poverty is a multidimensional problem so there are many factors that affect poverty. Health is one of the factors that affect poverty. The low level of public health will lead to a decrease in productivity and human resources which causes the welfare of the community to decrease. Lack of knowledge about health is also often experienced by the poor so that these people are more vulnerable to disease. Diseases suffered by the community will certainly hamper their productivity, thus affecting the level of income [6]. The quality of health can be seen from life expectancy. Life expectancy is an estimate of the average age that a person is expected to live since birth. The calculation of life expectancy is based on age-specific mortality rates. With life expectancy, the government can evaluate its performance in improving the welfare of the population, especially in improving health status. A low life expectancy in an area indicates that the health status in the area is low, while an increasing life expectancy indicates that the population has more secure health and poverty has been overcome better. Increasing life expectancy in a region can be done by realizing health development programs, socialization programs on the importance of environmental health and nutritional adequacy, and followed by poverty eradication programs [7]. The higher life expectancy in a region indicates that the lower the number of poor people in the region, while the lower the life expectancy, the higher the number of poor people in the region.

Good health quality must be supported by adequate health facilities. One of the means of supporting health is access to proper sanitation. The poor quality of sanitation is one of the causes of poverty in Indonesia. If the quality of sanitation is poor, it will cause human welfare to decrease and adversely affect the health and productivity of the community. Some people are less concerned about

sanitation issues because their income is only enough to fulfill their basic needs [8]. The sanitation quality of Indonesian communities can be seen based on data on the percentage of households that have access to proper sanitation services. The percentage of households that have access to proper sanitation services is the number of people who have access to proper sanitation services compared to the total population expressed in percent. The requirements for sanitation facilities are considered proper if the facility is used by the household alone or with certain other households, the facility is equipped with a toilet, and is equipped with a final disposal site in the form of a septic tank. The percentage of households that have access to proper sanitation services can be used to measure households that already have access to proper sanitation services at both urban and rural levels. The percentage of households that have access to proper sanitation services can also indicate the welfare level of the population based on health aspects [9]. Improving the quality of sanitation in an area will affect the increase in community productivity so that it can move out of poverty.

Poverty is closely related to economic conditions. Good economic growth will certainly have an impact on the prosperity and welfare of the community. Economic growth will also increase various social developments that can overcome the problem of poverty. With good economic development, people will have adequate income so as to improve the standard of living of the community. Economic growth in Indonesia can be seen through the rate of Gross Regional Domestic Product (GRDP) in each region [10]. The GRDP rate is an indicator aimed at seeing the growth of production of goods and services at a certain interval of time in a region. The GRDP rate has several uses, namely it can be used to measure economic progress as a result of national development, can be used as a basis for government reference in estimating state revenues aimed at development planning on a national, sectoral and regional scale, and also the GRDP rate can be used to create sales equations that can be used as a basis for making business forecasts [11]. GRDP growth has a negative influence on poverty, which means

that the higher GRDP growth, the lower poverty will be.

In addition to health, environment and economy, education is also one of the factors that affect poverty. By pursuing education, people will have more insight, ability, skills and independence. Adequate education will improve the quality of human resources. The high quality of human resources will increase productivity. With a good education, people will also get a decent job so that they have adequate income [12]. The level of education in Indonesia can be seen from the expected years of schooling of its people. Expected years of schooling is the number of years expected to describe the length of schooling for the population aged seven years and over in the future. Expected years of schooling shows the opportunity for children aged seven years and over to pursue formal education at a certain time. Expected years of schooling can be used to see the state of development in the education system at various levels. Expected years of schooling depend on how many people are in school at any given time. The higher the dropout rate, the lower the expected years of schooling. An increase in expected years of schooling will lead to an increase in the Human Development Index (HDI) and a decrease in the poverty rate [13].

The statistical method that can be used to identify factors that affect poverty is regression analysis. Based on previous research, it shows that poverty in a region is influenced by poverty in the surrounding areas [14]. One type of regression analysis that is suitable for use in this study is spatial error regression analysis with an area approach. The area approach is used because with this approach, the dependence between poverty in a region and poverty in its surroundings can be known. The spatial error model is used in this study to determine the spatial influence on the error variable [15]. Therefore, this research was conducted by modeling the percentage of poor people in Indonesia using the spatial regression error analysis method of area approach as an effort to reduce poverty in Indonesia. The results of this study will produce several factors that affect poverty so that it can be a

recommendation for the government in making programs or policies aimed at reducing poverty.

## II. METHODOLOGY

The data used in this study is data related to the percentage of poor people in each province in Indonesia in 2022. Data on the percentage of poor people and data on factors that are thought to be influential are secondary data sourced from BPS Indonesia publications. In this study, the research units and data used are 34 provinces in Indonesia. The variables used in this study are divided into 2, namely the response variable ( $Y$ ) and the predictor variable ( $X$ ) with details of the variables used can be seen in Table 1.

TABLE I  
 RESEARCH VARIABLES

Variables	Variable Description	Scale	Variable Type
$Y$	Percentage of Poor Population	Ratio	Continuous
$X_1$	Life Expectancy	Ratio	Continuous
$X_2$	Percentage of Households with Access to Adequate Sanitation Services	Ratio	Continuous
$X_3$	GRDP rate	Ratio	Continuous
$X_4$	Expected Years of Schooling	Ratio	Continuous

The methods used in this study are classical regression analysis and spatial error regression analysis. The classical regression model is a linear regression model consisting of one response variable and  $p$  predictor variables. The regression model is used to create a model that describes the relationship between the predictor variable and the response variable [16]. Meanwhile, spatial regression is one of the developments of classical

regression in which the spatial regression model has accommodated the occurrence of spatial autocorrelation in the observation data. Spatial error arises due to the dependence of the error value of a region with the error value of another adjacent region [17].

The stages of analysis in this study began with an analysis of the percentage of poor people in Indonesia using classical regression. Further analysis was carried out using the spatial error regression method. Spatial error regression analysis has several stages. First, the formation of a spatial weight matrix with an area approach is carried out. The weight matrix is used to determine the relationship between an area and its surrounding areas [18]. Furthermore, the spatial dependency test is carried out using Moran's index which is intended to determine the spatial relationship that occurs in an observation unit with other observation units. After that, the lagrange multiplier test is used to identify the spatial influence that occurs in the data in the spatial error regression model. If all tests have been met, then spatial error regression modeling is performed. A summary of the tests performed on error spatial regression is shown in Table 2.

TABLE 2

TEST ON SPATIAL REGRESSION ERROR

Test	Target	Description
Spatial Dependence Test	P-value < 0,05	There is spatial dependence between observation units
Lagrange Multiplier Test	P-value < 0,05	There is dependency between error units of observation

The selection of the best model between the classical regression model and the spatial error regression model is done by looking at the Akaike Information Criterion (AIC) value and the coefficient of determination of each model. The calculation of the AIC value is done with the following formula.

$$AIC = -2Lm + 2m$$

*Lm* : Maximum log-likelihood value  
*m* : number of model parameters

The best model is the one with the smallest AIC value [19]. The coefficient of determination is one of the measures that can be used to measure how well the model's ability to explain the diversity of the dependent variable. The coefficient of determination formula is as follows.

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

The range of the coefficient of determination is from 0 to 1. The smaller the coefficient of determination means that the ability of the resulting model to explain the variation in the response variable y is very limited, while for the coefficient of determination that is close to one means that the ability of the model obtained is very good [20].

### III. RESULT

This research aims to find out what factors affect poverty in Indonesia, based on data on the percentage of poor people in Indonesia. Descriptive statistical analysis of the percentage of poor people in Indonesia and its predicted factors was conducted as the initial stage of analysis. The results of the descriptive statistics are presented in Table 3.

TABLE 3

DESCRIPTIVE STATISTICS

Variable Description	Mean	Variance
Percentage of Poor Population	10.29	27.12
Life Expectancy	70.42	5.83
Percentage of Households with Access to Adequate Sanitation Services	80.99	92.93
GRDP rate	5.76	13.31
Expected Years of Schooling	13.24	0.53

Based on Table 3, it can be seen that the percentage of poor people in Indonesia in each province has an average of 10.29% and a variance of 27.12. In the life expectancy variable, it can be seen that the average life expectancy of the Indonesian people is 70.42 years and the variance is 5.83. The awareness of the importance of sanitation in Indonesia is still quite low because the average

percentage of households that have access to proper sanitation services is only 80.99% and the variance is 92.93. In economic factors, it can be seen that the economy in Indonesia is uneven because it has a fairly good average of 5.76, but has a large variance value of 13.31. The education factor shows that education in Indonesia is quite good as seen from the average length of schooling of 13.24 years with a variance of 0.52.

The identification of poverty in Indonesia can also be done by looking at the distribution through thematic maps. In this study, poverty in Indonesia is categorized into 4 categories, namely low, low medium, high medium, and high categories. The distribution of poverty through thematic maps is shown in Figure 1.

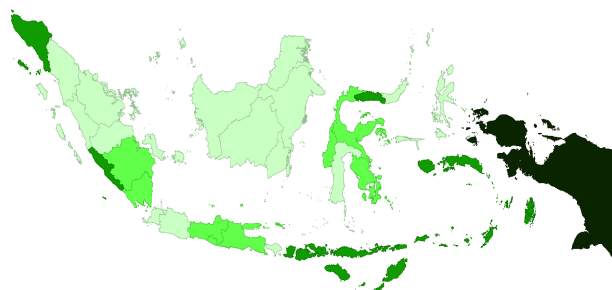


FIGURE 1

DISTRIBUTION OF POVERTY IN INDONESIA

The low category is the provinces that have a percentage of the poor population below 10.49%. In the low category there are 18 provinces which include Sumatera Utara, Riau, Sumatera Barat, Jambi, Bangka Belitung, Kepulauan Riau, DKI Jakarta, Jawa Barat, Banten, Bali, Kalimantan Barat, Kalimantan Timur, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Utara, Sulawesi Selatan, Sulawesi Utara, and Maluku Utara. The medium-low category are provinces that have a percentage of poor people between 10.49% and 13.82%. In the medium-low category, there are 8 provinces, namely Sumatera Selatan, Lampung, Jawa Tengah, DI Yogyakarta, Jawa Timur, Sulawesi Barat, Sulawesi Tengah, and Sulawesi Tenggara. The medium-high category are provinces that have a percentage of poor people from 13.82% to 20.32%. In the medium-high category, there are 6 provinces, namely Aceh, Bengkulu, Nusa Tenggara Barat, Nusa Tenggara Timur, Gorontalo, and Maluku. The

high category is a province that has a percentage of the poor population of more than 20.23%. In the high category there are 2 provinces, namely Papua and Papua Barat.

Identification of the factors that affect poverty in Indonesia is done using regression analysis. First, the analysis was conducted using classical regression. First, simultaneous testing using classical regression was conducted. The simultaneous test results of the classical regression analysis are presented in Table 4.

TABLE 4  
SIMULTANEOUS TESTING ON CLASSICAL REGRESSION

F-statistic	9.7159
Probability (F-statistic)	0.0000
Description	All variables have an effect simultaneously

Based on Table 4, it can be seen that simultaneously the variables of Life Expectancy, Percentage of Households with Access to Proper Sanitation Services, GRDP Rate, and Expected Years of Schooling have a significant effect on the Percentage of Poor Population in Indonesia. Furthermore, testing on each variable is presented in Table 5.

TABLE 5  
INDIVIDUAL TESTS ON CLASSICAL REGRESSION

Variable	Coefficient	Standar Error	t-Statistic	Probability
Constant	60.73	22.87	2.656	0.013
Life Expectancy	-0.666	0.323	-2.059	0.049
Percentage of Households with Access to Adequate Sanitation Services	-0.338	0.091	-3.709	0.001
GRDP rate	-0.319	0.177	-1.797	0.083

Expected Years of Schooling	1.936	0.992	1.952	0.061
-----------------------------	-------	-------	-------	-------

Tests on each variable resulted in the life expectancy variable having a significant effect on the percentage of poor people in Indonesia. Life expectancy and the percentage of poor people have a negative relationship, which means that when life expectancy increases, the percentage of poor people will decrease. The variable percentage of households that have access to proper sanitation also has a significant effect on the percentage of poor people. The percentage of households that have access to proper sanitation has a negative relationship with the percentage of poor people in Indonesia, which means that if the percentage of households that have access to proper sanitation increases, the percentage of poor people in Indonesia will decrease. Meanwhile, the GRDP rate and expected years of schooling have no significant effect on poverty in Indonesia. Based on Table 5, the classical regression model for the percentage of poor people in Indonesia is as follows.

$$\hat{Y} = 60,73 - 0,666X_1 - 0,338 X_2 - 0,319X_3 + 1,936X_4$$

Furthermore, the analysis is carried out using the spatial error regression method. The first stage before performing spatial error regression is to conduct a spatial dependency test. The results of the spatial dependency test are presented in Table 6.

TABLE 6  
 SPATIAL DEPENDENCE TEST

Value	3,143
Probability	0,0017
Description	There is spatial dependence between observation units

Based on Table 6, it can be seen that there is spatial dependence between one observation unit and another observation unit. Because there is spatial dependence between observation units, then the lagrange multiplier test for spatial error regression is conducted. The results of the lagrange multiplier test for spatial error regression are presented in Table 7.

TABLE 7  
 LAGRANGE MULTIPLIER TEST

Value	6.492
Probability	0.0084
Description	There is dependency between error units of observation

Based on Table 7, it can be seen that there is a spatial dependence on the error of the spatial error regression model so that the next step can be modeled using the spatial error regression method. The results of spatial error regression analysis are presented in Table 8 as follows.

TABLE 8  
 RESULTS OF SPATIAL REGRESSION ERROR ANALYSIS

Variable	Coefficient	Standar d Error	t-Statistic	Proba bility
Constant	47.57	19.94	2.38	0.017
Life Expectancy	-0.49	0.29	-1.69	0.091
Percentage of Households with Access to Adequate Sanitation Services	-0.29	0.07	-4.34	0.000
GRDP rate	-0.35	0.12	-2.90	0.004
Expected Years of Schooling	1.77	0.73	2.43	0.015

Testing with the classical regression method resulted in the life expectancy variable not having a significant influence on the percentage of poor people in Indonesia. The variable percentage of households that have access to proper sanitation, the rate of GRDP, and expected years of schooling have a significant effect on the percentage of poor people. The percentage of households that have access to proper sanitation and the GRDP rate have a negative relationship with the percentage of poor people in Indonesia, which means that if the percentage of

households that have access to proper sanitation increases, it will reduce the percentage of poor people in Indonesia, as well as if the GRDP rate increases, the percentage of poor people in Indonesia will decrease. Meanwhile, school expectancy has a positive relationship with the percentage of poor people in Indonesia, which means that if school expectancy increases, the percentage of poor people in Indonesia will also increase. Based on Table 8, the spatial error regression model for the percentage of poor people in Indonesia is obtained as follows:

$$\hat{Y} = 86,321 + 2,275X_1 - 0,609 X_2 - 0,182X_3 - 0,903X_4 + u$$

$$u = 0,529 \sum_{i=1, j=1}^n W_{ij}u_j$$

The selection of the best model for the percentage of poor people in Indonesia between the classical regression model and the spatial error regression model is done by reviewing the coefficient of determination. The model with the highest coefficient of determination will be chosen as the best model to model the percentage of poor people in Indonesia. The coefficient of determination is presented in Table 9.

TABLE 9  
 AIC VALUE AND DETERMINATION COEFFICIENT VALUE

Model	AIC	Coefficient of Determination
Classical Regression	189.785	0.573
Spatial Regression Error	180.989	0.708

Based on Table 9, it can be seen that the AIC value of the spatial error regression model is smaller than the classical regression and the coefficient of determination of the spatial error regression model is bigger than the classical regression model, therefore it can be decided that the best model for the percentage of poor people in Indonesia is using the spatial error regression model.

#### IV. CONCLUSIONS

Modeling the percentage of poor people in Indonesia using spatial error regression is better than using classical regression. This is because there is a relationship between poverty in a region and poverty in the surrounding areas. From the results of the study, several factors were found to influence the percentage of poor people in Indonesia. Factors that have a significant effect on the percentage of poor people in Indonesia are the percentage of households that have access to proper sanitation, GRDP rate, and expected years of schooling, while the life expectancy factor does not have a significant effect on the percentage of poor people in Indonesia. In addition, poverty in a region is also influenced by poverty in surrounding areas. Based on the results of this study, efforts to reduce poverty in Indonesia can pay attention to influential factors, namely the quality of sanitation, economic growth, and the quality of education.

#### REFERENCES

- [1] Munandar, A. I., Darjono, A. H. & Aprilasani, Z., *Pembangunan Berkelanjutan: Studi Kasus Di Indonesia*, Jakarta: Bypass, 2019.
- [2] United Nations Development Programme, *Sustainable Development Goals*, 2022.
- [3] Akhmad, K. A., Peran Pendidikan Kewirausahaan untuk Mengatasi Kemiskinan, *Intelektiva: Jurnal Ekonomi, Sosial & Humaniora*, 2(6), pp. 173-181, 2021.
- [4] Kementerian Pendayagunaan Aparatur Negara dan Reformasi Birokrasi, *Beragam Upaya Mengurai Permasalahan Pengentasan Kemiskinan Ekstrem di Indonesia*, 2022.
- [5] Kementerian Keuangan Republik Indonesia, *Peranan APBN Berhasil Menahan Kenaikan Angka Kemiskinan*, 2023.
- [6] Adhitya, B., Prabawa, A. & Kencana, H., Analisis Pengaruh Pendidikan, Kesehatan, Sanitasi dan Rata-Rata Jumlah Anggota Keluarga Per Rumah Tangga terhadap Kemiskinan di Indonesia, *Ekonomis: Journal of Economics and Business*, 6(1), pp. 288-295, 2022.
- [7] Badan Pusat Statistik, *Angka Harapan Hidup*, 2012.
- [8] Azizah, S. P. N., Pratiwi, L. S., Amaliah, I. & Fitriyana, F., Sanitasi Dan Kepadatan Penduduk Sebagai Dinamika Kemiskinan Kota: Studi Kasus Provinsi Jawa Barat. *Jurnal Nuansa Akademik*, 7(1), pp. 55-70, 2022.
- [9] Badan Pusat Statistik, *Persentase Rumah Tangga yang Memiliki Akses terhadap Layanan Sanitasi Layak*, 2020.
- [10] Proseptian, L. & Primandhana, W. P., Analisis faktor-faktor yang mempengaruhi kemiskinan. *Jurnal Ekonomi Manajemen dan Akuntansi*, 24(1), pp. 45-53, 2022.
- [11] Badan Pusat Statistik, *Laju Pertumbuhan PDB / PDRB*, 2022.
- [12] Iksan, M. & Arka, S., Pengaruh Upah, Pendidikan, dan Pertumbuhan Ekonomi terhadap Kesempatan Kerja serta Kemiskinan Provinsi Jabar Bagian Selatan. *E-Jurnal EP Unud*, 11(1), pp. 147-175, 2022.

- [13] Sabrina, R., Manurung, A. I. & Sirait, B. A., Peningkatan Rata-Rata Lama Sekolah (RLS) dari Harapan Lama Sekolah (HLS) di Sumatera Utara. *Jurnal Pendidikan Tambusai*, 6(1), pp. 4784-4792, 2022.
- [14] Nirmala, K. L., Pramesti, W. & Fitriani, F., Spatial Autoregressive Moving Average Pada Pemodelan Persentase Penduduk Miskin. *J Statistika*, 15(1), pp. 158-166, 2022.
- [15] Anselin, L., *Spatial Econometric: Methods and Models*. Dordrecht: Academic Publishers, 1988.
- [16] Rencher, A. C. & Schaalje, G. B., *Linear Model in Statistics*. Inc, USA: John Wiley and Sons, 2008.
- [17] LeSage, J. P. & Pace, R. K., *Introduction to Spatial Econometrics*. Boca Ration: R Press, 2009.
- [18] Lee, J. & Wong, S. W. D., *Statistical Analysis with Arcview GIS*. United Stated of America: John Willey & Sons, 2001.
- [19] Grasa, A. A., *Econometric Model Selection: A New Approach*. Kluwer: Springer Science and Business Media, 1989.
- [20] Walpole, R. E., Myers, R. H., Myers, S. L. & Ye, K., *Probability & Statistics for Engineers & Scientists*. United States of America: Prentice Hall, 2011.