

PAPER • OPEN ACCESS

## Modelling of Poverty Percentage of Non-Food Per Capita Expenditures in Indonesia Using Least Square Spline Estimator

To cite this article: A. Massaid *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **546** 052044

View the [article online](#) for updates and enhancements.



**IOP | ebooks™**

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

# Modelling of Poverty Percentage of Non-Food Per Capita Expenditures in Indonesia Using Least Square Spline Estimator

A. Massaid<sup>1</sup>, M. Hanif<sup>1</sup>, D. Febrianti<sup>1</sup>, N. Chamidah<sup>2,\*</sup>

<sup>1</sup>Student of Study Program of Statistics, Department of Mathematics, Airlangga University, Surabaya, Indonesia

<sup>2</sup>Department of Mathematics, Faculty of Sciences and Technology, Airlangga University, Surabaya, Indonesia

\*Corresponding author: nur-c@fst.unair.ac.id

**Abstract.** The average per capita expenditure is the cost incurred for the consumption of all household members for a month divided by the number of household members. The higher expenditure per capita, the higher level of welfare of the population. In Indonesia, the number of non-food per capita expenditures increases every year and followed by a decrease in the number of poor people. Therefore, it is necessary to investigate the effect of changes in per capita non-food expenditure on the percentage of poor people in Indonesia. To investigate it, in this paper, we use nonparametric regression model based on least square spline estimator that could accommodate the data patterns changing at certain points. The results show that for every one-million-rupiah increase in per capita non-food expenditure, if the expenditure per capita is less than 0.47 million rupiah, the percentage of poverty will decrease 68.71%, and if per capita expenditure is more than 0.47, the percentage of poverty will decrease 5.96%.

## 1. Introduction

Development is the goal of a country, that the country will advance when there are an increase in its development. Rustam [1] studied one indicator success of development is increasing economic growth, the expectation by high economic growth can reduce unemployment or existing poverty. According to Christianto [2], the poverty condition of a country or region is a reflection of the level of welfare of the population living in the country / region. According to Sumodiningrat [3], absolute poverty is that when the level of income is below the poverty line and cannot fulfil needs to live and work. While per capita expenditure can be classified into food and non-food groups.

The changes in income of people affects to the shift in spending patterns. The higher income means the higher non-food expenditure. Thus, non-food expenditure patterns can be used as a tool to measure the level of welfare of the population. The size of the proportion of non-food expenditure is a reflection of the welfare of the population. This is compatible with Engel's theory, which states that the greater proportion of non-food expenditure is usually identified by increasing welfare of the community. While the greater proportion of non-food expenditure reflects the decreasing level of people's poverty. Therefore, we can model poverty based on non-food expenditure to find out how far it affects on decreasing poverty.

In statistics, one of methods that can be used to model is regression analysis. There are two model approaches, i.e., parametric and nonparametric regression models. In the nonparametric



regression model, there are many estimators to estimate the function, i.e., kernel, local linear, local polynomial and spline estimator. Researchers who have studied nonparametric regression model by using kernel (Chamidah & Saifudin [4]), local linear (Chamidah, et al [5]), local polynomial (Chamidah, et al [6]) and spline estimators (Chamidah & Lestari [7], Oktavetri et al [8], and Lestari et al [9]). The number of poor people tends to change based on per capita expenditure, that affected to the number of poor people based on per capita expenditure would be good if it is estimated locally. The spline estimator approach gives more flexible to adopt behaviour pattern change of data locally because in the spline estimator there are knots that can accommodate the data patterns changing at certain points. So, the use of least square method based spline estimator is appropriate to model of poverty percentage of non-food per capita expenditures in Indonesia. According to Islamiyati et al [10] least Estimator square spline is pieces with different polynomial segments that are combined by knots while ensuring continuity.

Data on the number of poor people on per capita expenditure has different functions in several domain functions, besides that it can be seen from plots of scattered and truncated data at certain points forming pieces with different polynomial segments. Based on the reason the Least Square Spline estimation can be used as a reference for forming local functions. Some previous studies related spline regression methods are: Islamiyati et al [10] studied Estimation of Covariance Matrix on Bi-Response longitudinal data analysis with Penalized Spline Regression, and Anggreni et al [11] design the number of tuberculosis cases in the province of Bali using nonparametric regression with spline truncated estimator, so the authors used a nonparametric regression approach to estimate the percentage of poverty based on per capita expenditure based on selected samples from 34 provinces in Indonesia in 2017 by using the least square spline estimator.

## 2. Method

### 2.1. Least Square Spline in Nonparametric Regression

One of the nonparametric regression models used is spline. Spline in nonparametric regression has the ability to estimate the behaviour of data that tends to differ at different intervals. The different intervals are joint fusion points where a pattern of behaviour changes from a function at different intervals. In general, the function  $f$  in spline space has the order  $p$  with points knots  $\tau_1, \tau_2, \dots, \tau_k$  is any function that can be expressed as the following equation:

$$f(x_i) = \sum_{j=0}^p \beta_j x^j + \sum_{j=1}^k \beta_{j+p} (x - \tau_j)_+^p \quad (1)$$

$$\text{where, } (x - \tau_{j-p})_+^p = \begin{cases} (x - \tau_{j-p})^p, & x \geq \tau_{j-p} \\ 0, & x < \tau_{j-p} \end{cases}$$

If given  $\lambda = (\tau_1, \tau_2, \dots, \tau_k)$  and  $\beta$  are the parameters of the model with  $p$  is the order of spline. The  $\beta$  parameter can be calculated by the following equation

$$\hat{\beta} = (X_\lambda^T X_\lambda)^{-1} X_\lambda^T y \quad (2)$$

Estimator of  $f(x_i)$  in (1) can be expressed as

$$\hat{f}_\lambda(x) = H(\lambda)y \quad (3)$$

$$\text{where } H(\lambda) = X_\lambda (X_\lambda^T X_\lambda)^{-1} X_\lambda^T.$$

### 2.2. Generalized Cross Validation

One of methods used to determine the optimum knots is the Generalized Cross Validation (GCV) method. The best least square spline estimator is obtained based on the minimum GCV value. The GCV function for nonparametric spline regression model is:

$$GCV(\tau_1, \tau_2, \dots, \tau_k) = \frac{MSE(\tau_1, \tau_2, \dots, \tau_k)}{\left( n^{-1} \text{tr}[I - H(\tau_1, \tau_2, \dots, \tau_k)] \right)^2} \quad (4)$$

$$\text{where } MSE(\tau_1, \tau_2, \dots, \tau_k) = n^{-1} \sum_{i=1}^n (y_i - \hat{f}(x_i))^2 \quad (5)$$

The optimum knots is obtained by minimizing GCV in (4).

### 3. Data and Steps of Analyze

The secondary data used in this study is obtained from the Central Bureau of Statistics. The data obtained is the percentage of poverty data and per capita non-food expenditure data from each Province in Indonesia in 2017. The analysis is carried out by creating the R code. The steps of the analysis used are as follows:

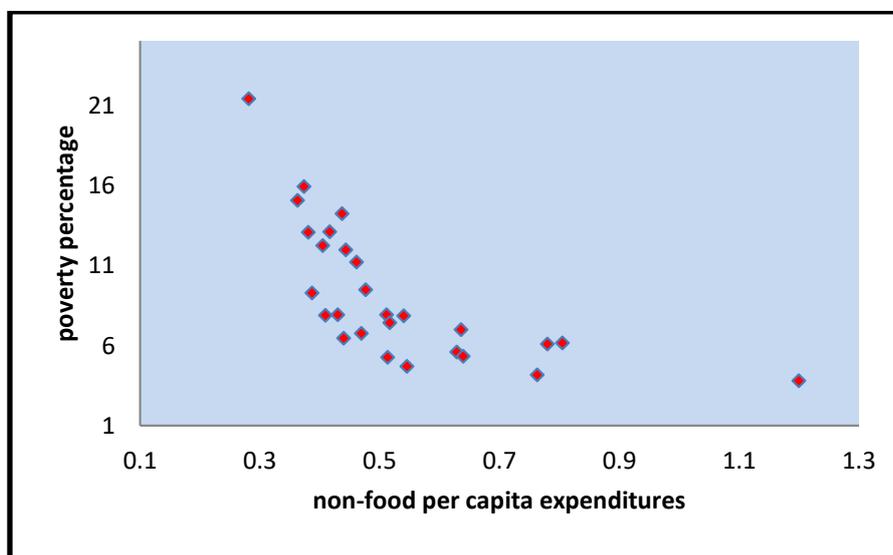
- Input data in pairs for predictor variable ( $x$ ), i.e., per capita non-food expenditure and response variable ( $y$ ), i.e., percentage of poverty in Indonesia.
- Determine optimum of the order, the number of knots and knots based on minimum GCV value in (4)
- Estimate parameters by using least square spline estimator in (3)
- Determine goodness of fit criterion for MSE (in 5) and R-square as follows:

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (6)$$

- Plot observed and the estimated of poverty percentage versus non-food per capita expenditures in Indonesia.
- Interpret the estimated model of poverty percentage of non-food per capita expenditures.

### 4. Result and Discussion

For estimating poverty percentage based on non-food per capita expenditures in Indonesia, we provide scatter plot of poverty percentage versus non-food per capita expenditures to find out pattern of behavior change of poverty percentage in the following Figure 1.



**Figure 1.** Scatter plot of poverty percentage versus non-food per capita expenditures

Figure 1 shows that there is behavior change of poverty percentage based on non-food per capita expenditures. Therefore, we use least square spline estimator that contains knots to estimate poverty percentage based on non-food per capita expenditures. After that, we determine the best knots by using quantile methods and seen the minimum GCV from each result.

**Table 1.** Number of knots, knots, order and GCV values

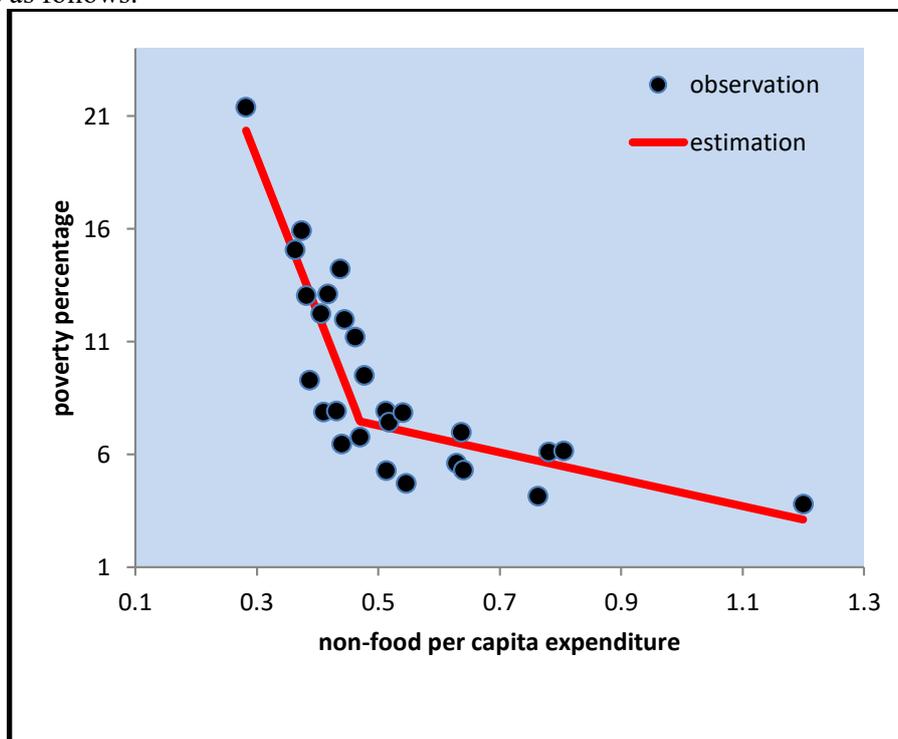
Number of Knots	Knots	Order	GCV Values
1	0.47	1	5.23
		2	5.57
2	0.44; 0.53	1	5.29
		2	5.68
3	0.41; 0.47; 0.59	1	5.51
		2	6.03

Based on **Table 1**, we obtain the optimal model of poverty percentage by using least square spline estimator, i.e., order one, knot of 0.47 and the minimum GCV value of 5.23. The estimated model gives mean square error (MSE) value and  $R^2$ , i.e., 4.13 and 76.3%, respectively. The estimated model can be expressed as follows:

$$y = 39.75 - 68.71x + 62.75(x - 0.47)_+$$

$$y = \begin{cases} 39.75 - 68.71x, & \text{for } x < 0.47 \\ 10.26 - 5.96x, & \text{for } x \geq 0.47 \end{cases} \quad (7)$$

Next, we show the plot of observed and estimated of poverty percentage versus non-food per capita expenditures as follows:



**Figure 2.** Plot estimation poverty percentage of non-food per capita expenditures in Indonesia.

Based on the estimated model in (7), if the non-food per capita expenditures less than 0.47 million rupiahs, then each increasing one million rupiahs non-food per capita expenditure can reduce the poverty percentage in Indonesia of 68.71%. Also, if the non-food expenditures equal or more than 0.47 million rupiahs, then each increasing one million rupiahs non-food per capita expenditure can reduce the poverty percentage in Indonesia of 5.96%.

## 5. Conclusion

Based on the results of the per capita non-food expenditure plot with the percentage of poverty in Indonesia there is a change in behavior of the data pattern. So, least square spline estimator is suitable and can estimate the pattern of poverty percentage well. The biggest rate of decline of the poverty percentage occurs if non-food per capita expenditure is less than 0.47 million rupiahs.

## 6. Acknowledgment

The authors would like to thank to Faculty of Sciences and Technology, Airlangga University for supporting financial in 9<sup>th</sup> Annual Basic Science International Conference 2019.

## References

- [1] Rustam 2010 Perencanaan Pertumbuhan Ekonomi Jawa Timur Dalam Rangka Mengurangi Angka Pengangguran dan Kemiskinan *Media Soerjo*6(1) p 1
- [2] Christianto T 2013 Determinan dan Karakteristik Kemiskinan di Provinsi Maluku *Cita Ekonomika*7(2) p 1
- [3] Sumodiningrat 1999 Kemiskinan : teori, fakta dan kebijakan (Indonesia: IMPAC) p 81
- [4] Chamidah N, Saifudin T 2013 Estimation of children growth based on kernel smoothing in multi response nonparametric regression *Applied Mathematical Sciences*7(37)pp 1839-1847
- [5] Chamidah N, Tjahjono E, Fadilah A.R., Lestari B 2018 Standard Growth Charts for Weight of Children in East Java Using Local Linear Estimator. *J.Phys: Conf.Ser*1097 012092.
- [6] Chamidah N, Gusti K.H, Tjahjono E, Lestari B 2019 Improving of Classification Accuracy of Cyst and Tumor Using Local Polynomial Estimator *Telkomnika* in press
- [7] Chamidah N, Lestari B 2016 Spline estimator in homoscedastic multi-response nonparametric regression model in case of unbalanced number of observations *Far East Journal of Mathematical Sciences (FJMS)*100(9) pp: 1433-1453
- [8] Oktavetri N.I, Kuncoro E.P, Purnobasuki H, Chamidah N 2019 Prediction of suspended and attached process behavior in anaerobic batch reactor using nonparametric regression model approach based on spline estimator *Eco. Env. & Cons.*25pp: 103-107
- [9] Lestari B, Fatmawati, Budiantara I.N, Chamidah, N 2018 Estimation of regression function in multi-response nonparametric regression model using smoothing spline and kernel estimators *J. Phys: Conf.Ser*1097 012091
- [10] Islamiyati A, Fatmawati, Chamidah N 2018 Estimation of Covariance Matrix on Bi-Response Longitudinal Data Analysis with Penalized Spline Regression *J. Phys.: Conf.Ser*979 012093
- [11] Anggreni, N.P.R., Suciptawati, N.L.P., Srinadi, I.G.A.M. 2018 Model Regresi Nonparametrik Spline Truncated Pada Jumlah Kasus Tuberkulosis di Provinsi Bali tahun 2016 E-jurnal *Matematika* 7(3) pp: 211-218