

CURVE ESTIMATION AND ESTIMATOR PROPERTIES OF THE NONPARAMETRIC REGRESSION TRUNCATED SPLINE WITH A MATRIX APPROACH

Nurul Fitriyani^{1§}, I Nyoman Budiantara²

¹Department of Mathematics, Universitas Mataram, Mataram, Nusa Tenggara Barat, Indonesia
[Email: nurul.fitriyani@unram.ac.id]

²Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya, Jawa Timur, Indonesia
[Email: i_nyoman_b@statistika.its.ac.id]

[§]Corresponding Author

ABSTRACT

Regression analysis is one of the statistical analyses used to estimate the relationship between the predictor and the response variable. Data are given in pairs, and the relationship between the predictor and the response variable was assumed to follow a nonparametric regression model. This model is flexible in estimating the curve when a typical data pattern does not follow a specific pattern. The nonparametric regression curve was approached by using the truncated spline function with several knots. The truncated spline estimator in nonparametric regression is linear in the observation. It is highly dependent on the knot points. The regression model's random error is assumed to have an independent normal distribution with zero mean and equal variance. The truncated spline's curve estimate was obtained by minimizing the error model through the least squared optimization method. The nonparametric regression truncated spline's estimator properties are linear, unbiased, and if the error is normally distributed, the estimator is normally distributed.

Keywords: *curve estimation, estimator properties, matrix approach, nonparametric regression, truncated spline.*

1. INTRODUCTION

Nonparametric regression has received a lot of attention from researchers. This approach loosens assumptions about linearity, as well as information about functional form in regression analysis. Besides, this approach allows data to be explored more flexibly. It is a flexible regression in estimating the regression curve when, in some circumstances, a common data pattern does not follow a specific pattern (Mahmoud, 2019).

There are several approaches to nonparametric regression, one of which is the truncated spline approach. Among the nonparametric regression models, spline regression has several features, such as it is a model with particular and excellent statistical interpretation and visual interpretation. The spline can model data on changing data patterns at certain sub-intervals because the spline is a polynomial slice with segmented properties. This segmented property provides

more flexibility than ordinary polynomials, making it possible to adapt more effectively to the local characteristics of a function or data (Budiantara et al. (2015); Nurcahayani et al., (2019); Wening et al. (2020)).

Previous research has been carried out by applying truncated spline nonparametric regression to various case data, such as pattern data related to poverty, population, education, and health (Budiantara et al. (2012); Fitriyani et al. (2016); Chamidah et al. (2019); Murbarani et al. (2019); Nurcahayani et al. (2019)). This study is conducted related to the curve estimation and estimator properties of one of the nonparametric regression models with a matrix approach. It is expected to provide a scientific insight into the process of curve estimation, and the properties of the nonparametric regression model's estimators with the spline truncated approach.

2. CURVE ESTIMATION OF THE NONPARAMETRIC REGRESSION TRUNCATED SPLINE

The spline function is the sum of the polynomial functions with a truncated function. Consider the nonparametric regression model, where the g curve estimation is done using a spline. Given paired data (x_j, y_j) and the relationship between x_j and y_j assumed to follow a nonparametric regression model:

$$y_j = f(x_j) + \varepsilon_j, j = 1, 2, \dots, n \quad (1)$$

This study is carried out with the regression curve f approached with the spline function g with knots K . The spline function g approaches regression curve with r points of knot K , the spline regression model is obtained according to equation (1). If equation (1) is presented in a matrix form, the following form is obtained.

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} g(x_1) \\ g(x_2) \\ \vdots \\ g(x_n) \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad (2)$$

The shape of the above equation's regression curve $g(x_j)$ is assumed to be unknown, while the errors $\varepsilon_j, j = 1, 2, \dots, n$ are mutually independent with zero mean and σ^2 variance. The spline function in equation (2) can be described in the following form:

$$\begin{aligned} g(x_j) = & \alpha_0 + \alpha_1 x_j + \alpha_2 x_j^2 + \dots + \alpha_m x_j^m \\ & + \beta_1 (x_j - K_1)_+^m + \dots \\ & + \beta_r (x_j - K_r)_+^m \end{aligned} \quad (3)$$

where α_i and β_k are a real constant, and a truncated function as in equation (3). If the truncated spline regression model is presented in the form of a matrix, it is obtained:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^m & (x_1 - K_1)_+^m & \dots & (x_1 - K_r)_+^m \\ 1 & x_2 & x_2^2 & \dots & x_2^m & (x_2 - K_1)_+^m & \dots & (x_2 - K_r)_+^m \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^m & (x_n - K_1)_+^m & \dots & (x_n - K_r)_+^m \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \\ \beta_1 \\ \vdots \\ \beta_r \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

or it can be written as:

$$Y = X[K_1, K_2, \dots, K_r] \beta + \varepsilon$$

Furthermore, the parameter estimates of β are obtained using the least square method by completing the optimization.

$$\underset{\beta}{\text{Min}} \{ \varepsilon^T \varepsilon \} = \underset{\beta}{\text{Min}} \left\{ \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}^T \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \right\}$$

with,

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}^T = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} - \begin{pmatrix} 1 & x_1 & \dots & x_1^m & (x_1 - K_1)_+^m & \dots & (x_1 - K_r)_+^m \\ 1 & x_2 & \dots & x_2^m & (x_2 - K_1)_+^m & \dots & (x_2 - K_r)_+^m \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \dots & x_n^m & (x_n - K_1)_+^m & \dots & (x_n - K_r)_+^m \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \\ \beta_1 \\ \vdots \\ \beta_r \end{pmatrix}^T$$

and,

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} - \begin{pmatrix} 1 & x_1 & \cdots & x_1^m & (x_1 - K_1)_+^m & \cdots & (x_1 - K_r)_+^m \\ 1 & x_2 & \cdots & x_2^m & (x_2 - K_1)_+^m & \cdots & (x_2 - K_r)_+^m \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^m & (x_n - K_1)_+^m & \cdots & (x_n - K_r)_+^m \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \\ \beta_1 \\ \vdots \\ \beta_r \end{pmatrix}$$

In a more straightforward form of presentation, the estimator,

$$\hat{\beta} = (\hat{\alpha}_0 \quad \hat{\alpha}_1 \quad \cdots \quad \hat{\alpha}_m \quad \hat{\beta}_1 \quad \cdots \quad \hat{\beta}_r)^T$$

$$\text{Min}_{\beta} \{\varepsilon^T \varepsilon\} = \text{Min}_{\beta} \{(Y - X[K_1, K_2, \dots, K_r] \beta)^T (Y - X[K_1, K_2, \dots, K_r] \beta)\}$$

with $Y = (y_1 \quad y_2 \quad \cdots \quad y_n)^T$ is a vector of size $n \times 1$ and $X[K_1, K_2, \dots, K_r] = X[K]$ is a

matrix of size $n \times (m + 1 + r)$. The sum of squared error is given as follows.

$$\begin{aligned} \sum_{j=1}^n \varepsilon_j^2 &= \varepsilon^T \varepsilon \\ &= (Y - X[K] \beta)^T (Y - X[K] \beta) \\ &= Y^T Y - 2\beta^T X[K]^T Y + \beta^T X[K]^T X[K] \beta \end{aligned}$$

If the above equation is derived concerning vector β^T and the result is equalized to zero,

then the result of $\partial \varepsilon^T \varepsilon / \partial \beta^T = 0$ given as the following,

$$\begin{aligned} \frac{\partial \varepsilon^T \varepsilon}{\partial \beta^T} &= 0 \\ \frac{\partial (Y^T Y - 2\beta^T X[K]^T Y + \beta^T X[K]^T X[K] \beta)}{\partial \beta^T} &= 0 \\ -2X[K]^T Y + 2X[K]^T X[K] \beta &= 0 \\ X[K]^T X[K] \beta &= X[K]^T Y \\ \hat{\beta} &= (X[K]^T X[K])^{-1} X[K]^T Y \end{aligned}$$

Consequently, the estimate for the spline regression curve with knots K is given by,

$$\begin{aligned} \hat{g}(x_j) &= X[K] \hat{\beta} \\ &= X[K] (X[K]^T X[K])^{-1} X[K]^T Y \end{aligned}$$

The equation can be written as the following form, with $X[K] (X[K]^T X[K])^{-1} X[K]^T = A[K]$ is a function of the knot points, and $K = (K_1, K_2, \dots, K_r)^T$ is the knot points.

$$\hat{g}(x_j) = A[K] Y$$

3. ESTIMATOR PROPERTIES OF THE NONPARAMETRIC REGRESSION TRUNCATED SPLINE

Further study is conducted regarding the properties of the truncated spline estimator in nonparametric regression.

a. Linear Estimator

In this section, a study of the linearity of the spline estimator in nonparametric regression is conducted. The spline

nonparametric regression model can be written in the following matrix form with $\mathbf{K} = (K_1, K_2, \dots, K_r)$.

$$\mathbf{Y} = \mathbf{X}[\mathbf{K}] \boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

If $\mathbf{X}[\mathbf{K}] \boldsymbol{\beta}$ is written by \mathbf{g} , with $\mathbf{X}[\mathbf{K}]$ is the function matrix of \mathbf{K} , we get:

$$\mathbf{Y} = \mathbf{g} + \boldsymbol{\varepsilon}$$

and,

$$\begin{aligned} \hat{\mathbf{g}} &= \mathbf{X}[\mathbf{K}] \hat{\boldsymbol{\beta}} \\ &= \mathbf{X}[\mathbf{K}] (\mathbf{X}[\mathbf{K}]^T \mathbf{X}[\mathbf{K}])^{-1} \mathbf{X}[\mathbf{K}]^T \mathbf{Y} \\ &= \mathbf{A}[\mathbf{K}] \mathbf{Y} \end{aligned}$$

Based on the above equation, it can be seen that the spline estimator $\hat{\mathbf{g}}$ is linear. This linearity can make it easier for researchers to form inference statistics for the spline approach.

b. Unbiased Estimator

In this section, the expected value of the estimator spline $\hat{\mathbf{g}}$ is determined to see whether the estimator is biased or not. Then obtained,

$$\begin{aligned} E(\hat{\mathbf{g}}) &= E(\mathbf{A}[\mathbf{K}] \mathbf{Y}) \\ &= \mathbf{A}[\mathbf{K}] E(\mathbf{Y}) \end{aligned}$$

where,

$$\begin{aligned} E(\mathbf{Y}) &= E(\mathbf{X}[\mathbf{K}] \boldsymbol{\beta} + \boldsymbol{\varepsilon}) \\ &= E(\mathbf{X}[\mathbf{K}] \boldsymbol{\beta}) + E(\boldsymbol{\varepsilon}) \end{aligned}$$

$\boldsymbol{\varepsilon}$ is a vector of random error $\varepsilon_j, j = 1, 2, \dots, n$ that is mutually independent with zero mean, $\mu_{\boldsymbol{\varepsilon}} = E(\boldsymbol{\varepsilon}) = \mathbf{0}$, and $\sigma^2_{\boldsymbol{\varepsilon}} = \sigma^2$ variance. The following form is obtained,

$$\begin{aligned} E(\mathbf{Y}) &= E(\mathbf{X}[\mathbf{K}] \boldsymbol{\beta}) \\ &= \mathbf{X}[\mathbf{K}] \boldsymbol{\beta} \end{aligned}$$

Therefore, the expectations of the estimator $\hat{\mathbf{g}}$ are obtained with the following equation,

$$\begin{aligned} E(\hat{\mathbf{g}}) &= \mathbf{A}[\mathbf{K}] \mathbf{X}[\mathbf{K}] \boldsymbol{\beta} \\ &= \mathbf{X}[\mathbf{K}] (\mathbf{X}[\mathbf{K}]^T \mathbf{X}[\mathbf{K}])^{-1} \mathbf{X}[\mathbf{K}]^T \mathbf{X}[\mathbf{K}] \boldsymbol{\beta} \\ &= \mathbf{X}[\mathbf{K}] \boldsymbol{\beta} \\ &= \mathbf{g} \end{aligned}$$

$E(\hat{\mathbf{g}}) = \mathbf{g}$ indicates that the estimator $\hat{\mathbf{g}}$ is unbiased.

c. Normally Distributed Estimator

In this section, a study is conducted regarding the distribution of the spline estimator in nonparametric regression. In inference statistics, random error $\boldsymbol{\varepsilon}$ is assumed to follow a normal distribution with zero mean, $E(\boldsymbol{\varepsilon}) = \mu_{\boldsymbol{\varepsilon}} = \mathbf{0}$, and $\sigma^2 \mathbf{I}$ variance. Since the spline nonparametric regression model's error is normally distributed, written as $N(\mathbf{0}, \sigma^2 \mathbf{I})$, the Moment Generating Function (MGF) of the \mathbf{Y} vector is given as follows.

$$\begin{aligned} M_{\mathbf{Y}}(\mathbf{t}) &= M_{\mathbf{g}+\boldsymbol{\varepsilon}}(\mathbf{t}) \\ &= M_{\mathbf{g}}(\mathbf{t}) \cdot M_{\boldsymbol{\varepsilon}}(\mathbf{t}) \\ &= E(\exp(\mathbf{t}^T \mathbf{g})) \cdot M_{\boldsymbol{\varepsilon}}(\mathbf{t}) \\ &= \exp(\mathbf{t}^T \mathbf{g}) \cdot M_{\boldsymbol{\varepsilon}}(\mathbf{t}) \\ &= \exp(\mathbf{t}^T \mathbf{X}[\mathbf{K}] \boldsymbol{\beta}) \cdot M_{\boldsymbol{\varepsilon}}(\mathbf{t}) \end{aligned}$$

Since the spline nonparametric regression model's random error is normally distributed, the Moment Generating Function (or written as MGF) of the vector $\boldsymbol{\varepsilon}$ is given as follows.

$$\begin{aligned} M_{\boldsymbol{\varepsilon}}(\mathbf{t}) &= E(\exp(\mathbf{t}^T \boldsymbol{\varepsilon})) \\ &= \exp\left(\mathbf{t}^T \mu_{\boldsymbol{\varepsilon}} + \frac{1}{2} \mathbf{t}^T \sigma^2_{\boldsymbol{\varepsilon}} \mathbf{I} \mathbf{t}\right) \\ &= \exp\left(\mathbf{t}^T (\mathbf{0}) + \frac{1}{2} \mathbf{t}^T \sigma^2 \mathbf{I} \mathbf{t}\right) \\ &= \exp\left(\frac{1}{2} \mathbf{t}^T \sigma^2 \mathbf{I} \mathbf{t}\right) \end{aligned}$$

So we get MGF from the \mathbf{Y} vector as the following.

$$\begin{aligned} M_{\mathbf{Y}}(\mathbf{t}) &= \exp(\mathbf{t}^T \mathbf{X}[\mathbf{K}] \boldsymbol{\beta}) \cdot \exp\left(\frac{1}{2} \mathbf{t}^T \sigma^2 \mathbf{I} \mathbf{t}\right) \\ &= \exp\left(\mathbf{t}^T \mathbf{X}[\mathbf{K}] \boldsymbol{\beta} + \frac{1}{2} \mathbf{t}^T \sigma^2 \mathbf{I} \mathbf{t}\right) \end{aligned}$$

The above results show the MGF of the normal distribution with mean $\mathbf{X}[\mathbf{K}] \boldsymbol{\beta}$ and variance $\sigma^2 \mathbf{I}$. This means that the vector \mathbf{Y} of the spline nonparametric regression model follows a normal distribution with $\mathbf{X}[\mathbf{K}] \boldsymbol{\beta}$ mean and $\sigma^2 \mathbf{I}$ variance.

The next step to determine the properties of the spline estimator in nonparametric regression is to find the distribution of the estimator $\hat{\boldsymbol{\beta}}$. If $\hat{\boldsymbol{\beta}} = (\mathbf{X}[\mathbf{K}]^T \mathbf{X}[\mathbf{K}])^{-1} \mathbf{X}[\mathbf{K}]^T \mathbf{Y}$ is known, then the MGF of the estimator $\hat{\boldsymbol{\beta}}$ is given as follows.

$$\begin{aligned}
 M_{\hat{\beta}}(\mathbf{t}) &= M_{(X[K]^T X[K])^{-1} X[K]^T Y}(\mathbf{t}) \\
 &= M_Y((X[K]^T X[K])^{-1} X[K]^T \mathbf{t}) \\
 &= \exp \left\{ ((X[K]^T X[K])^{-1} X[K]^T \mathbf{t})^T X[K] \boldsymbol{\beta} + \right. \\
 &\quad \left. \frac{1}{2} ((X[K]^T X[K])^{-1} X[K]^T \mathbf{t})^T \sigma^2 I (X[K]^T X[K])^{-1} X[K]^T \mathbf{t} \right\} \\
 &= \exp \left\{ (X[K]^T \mathbf{t})^T ((X[K]^T X[K])^{-1})^T X[K] \boldsymbol{\beta} + \right. \\
 &\quad \left. \frac{1}{2} (X[K]^T \mathbf{t})^T ((X[K]^T X[K])^{-1})^T \sigma^2 I (X[K]^T X[K])^{-1} X[K]^T \mathbf{t} \right\} \\
 &= \exp \left\{ \mathbf{t}^T X[K] (X[K]^T X[K])^{-1} X[K] \boldsymbol{\beta} + \right. \\
 &\quad \left. \frac{1}{2} \mathbf{t}^T X[K] (X[K]^T X[K])^{-1} \sigma^2 I (X[K]^T X[K])^{-1} X[K]^T \mathbf{t} \right\} \\
 &= \exp \left\{ \mathbf{t}^T [X[K] (X[K]^T X[K])^{-1} X[K] \boldsymbol{\beta}] + \frac{1}{2} \mathbf{t}^T [X[K] (X[K]^T X[K])^{-2} X[K]^T \sigma^2] \mathbf{t} \right\}
 \end{aligned}$$

It is found that $M_{\hat{\beta}}(\mathbf{t})$ is the MGF of the normal distribution with the mean and variance are given by,

$$\text{Mean} = X[K] (X[K]^T X[K])^{-1} X[K] \boldsymbol{\beta}$$

$$\text{Variance} = X[K] (X[K]^T X[K])^{-2} X[K]^T \sigma^2$$

Furthermore, taking into account the equation $\hat{\mathbf{g}} = A[K] Y$, we will look for the distribution

of the estimator $\hat{\mathbf{g}}$. By using MGF, the following results were obtained.

$$\begin{aligned}
 M_{\hat{\mathbf{g}}}(\mathbf{t}) &= M_{A[K]Y}(\mathbf{t}) \\
 &= M_Y(A[K]\mathbf{t})
 \end{aligned}$$

In the previous section, we got $M_Y(\mathbf{t}) = \exp \left\{ \mathbf{t}^T X[K] \boldsymbol{\beta} + \frac{1}{2} \mathbf{t}^T \sigma^2 I \mathbf{t} \right\}$, so,

$$\begin{aligned}
 M_{\hat{\mathbf{g}}}(\mathbf{t}) &= \exp \left\{ (A[K]\mathbf{t})^T X[K] \boldsymbol{\beta} + \frac{1}{2} (A[K]\mathbf{t})^T \sigma^2 I A[K] \mathbf{t} \right\} \\
 &= \exp \left\{ \mathbf{t}^T (A[K])^T X[K] \boldsymbol{\beta} + \frac{1}{2} \mathbf{t}^T (A[K])^T \sigma^2 I A[K] \mathbf{t} \right\} \\
 &= \exp \left\{ \mathbf{t}^T [(A[K])^T X[K] \boldsymbol{\beta}] + \frac{1}{2} \mathbf{t}^T [(A[K])^T \sigma^2 A[K]] \mathbf{t} \right\}
 \end{aligned}$$

$M_{\hat{\mathbf{g}}}(\mathbf{t})$ is the MGF of the normal distribution, with the mean (expected value) and variance

are given in the following formulas, respectively.

$$\begin{aligned}
 \text{Mean} &= (A[K])^T X[K] \boldsymbol{\beta} \\
 &= (X[K] (X[K]^T X[K])^{-1} X[K]^T)^T X[K] \boldsymbol{\beta} \\
 &= X[K] (X[K]^T X[K])^{-1} X[K]^T X[K] \boldsymbol{\beta}
 \end{aligned}$$

and,

$$\begin{aligned}
 \text{Variance} &= (X[K] (X[K]^T X[K])^{-1} X[K]^T)^T \sigma^2 X[K] (X[K]^T X[K])^{-1} X[K]^T \\
 &= X[K] (X[K]^T X[K])^{-1} X[K]^T X[K] (X[K]^T X[K])^{-1} X[K]^T \sigma^2 \\
 &= X[K] (X[K]^T X[K])^{-1} X[K]^T \sigma^2
 \end{aligned}$$

4. CONCLUSIONS

From the studies that have been done, it can be seen that the spline estimator in nonparametric regression is linear in the observation $\mathbf{Y} = (y_1 \ y_2 \ \dots \ y_n)^T$ and is highly dependent on the knot points K_1, K_2, \dots, K_r . The curve estimate for the truncated spline was obtained, namely $\hat{g}(x) = \mathbf{A}[\mathbf{K}] \mathbf{Y}$, with $\mathbf{A}[\mathbf{K}]$ indicates the function of knot points $\mathbf{K} = (K_1, K_2, \dots, K_r)^T$. The estimator's properties are linear, unbiased, and if the error is normally distributed, the estimator is normally distributed.

REFERENCES

- Budiantara, I. N., Ratna, M., Zain, I., & Wibowo, W. 2012. Modeling the Percentage of Poor People in Indonesia Using Spline Nonparametric Regression Approach. *International Journal of Basic & Applied Sciences IJBAS-IJENS*, 12(06), 119–124.
- Budiantara, I. N., Ratnasari, V., Ratna, M., & Zain, I. 2015. The Combination of Spline and Kernel Estimator for Nonparametric Regression and its Properties. *Applied Mathematical Sciences*, 9(122), 6083–6094.
- Chamidah, N., Lestari, B., & Saifudin, T. 2019. Modeling of Blood Pressures Based on Stress Score using Least Square Spline Estimator in Bi-response Non-parametric Regression. *International Journal of Innovation, Creativity and Change*, 5(3), 1200–1216.
- Fitriyani, N., Budiantara, I. N., Zain, I., & Ratnasari, V. 2016. Nonparametric Regression Spline in The Estimation of The Average Number of Children Born Alive Per Woman. *The 1st International Conference on Science and Technology (ICST)*, 169–172.
- Mahmoud, H. F. F. 2019. Parametric versus Semi and Nonparametric Regression Models. *Preprint*, 1–24. <https://arxiv.org/abs/1906.10221>
- Murbarani, N., Swastika, Y., Dwi, A., Aris, B., & Chamidah, N. 2019. Modeling of the Percentage of AIDS Sufferers in East Java Province using Nonparametric Regression Approach based on Truncated Spline Estimator. *Indonesian Journal of Statistics and Its Applications. Vol 3 No 2 (2019)*, 139 - 147, 3(2), 139–147.
- Nurchayani, H., Budiantara, I. N., & Zain, I. 2019. Nonparametric Truncated Spline Regression on Modelling Mean Years Schooling of Regencies in Java. *The 2nd International Conference on Science, Mathematics, Environment, and Education, AIP Conf. Proc. 2194*, 1–8.
- Wening, A. W., Budiantara, I. N., & Zain, I. 2020. Semiparametric Regression Curve Estimation for Longitudinal Data using Mixed Spline Truncated and Fourier Series Estimator. *ICCGANT 2019, Journal of Physics: Conference Series*, 1538, 1–10.