

## INTISARI

### MODEL REGRESI SEMIPARAMETRIK CAMPURAN SPLINE TRUNCATED DAN DERET FOURIER

Oleh

FARIDA RIZA UMAMI

19/448799/PPA/05882

Regresi semiparametrik merupakan salah satu pendekatan untuk mengestimasi kurva regresi. Diberikan data berpasangan  $(x_{1i}, x_{2i}, \dots, x_{pi}, t_{1i}, t_{2i}, \dots, t_{qi}, z_{1i}, z_{2i}, \dots, z_{ri})$ . Hubungan antar variabel prediktor dengan variabel respon mengikuti model regresi semiparametrik campuran. Model regresi semiparametrik yang memuat variabel tersebut dinyatakan sebagai  $y_i = \mu(x_{1i}, x_{2i}, \dots, x_{pi}, t_{1i}, t_{2i}, \dots, t_{qi}, z_{1i}, z_{2i}, \dots, z_{ri}) + \epsilon_i$

Kurva regresi  $\mu$  diasumsikan bersifat additif dapat ditulis menjadi :  $y_i = \sum_{j=1}^p f_j(x_{ji}) +$

$\sum_{s=1}^q g_s(t_{si}) + \sum_{m=1}^r h_m(z_{mi}) + \epsilon_i$ . Secara teoritis, kurva regresi  $f_j(x_{ji})$  dapat didekati dengan fungsi linier, kurva regresi  $g_s(t_{si})$  didekati dengan fungsi Spline Truncated linier. Sementara itu, kurva regresi  $h_m(z_{mi})$  dapat dihipotesis dengan Deret Fourier. Tujuan dari penelitian ini adalah memperoleh bentuk estimator dalam regresi semiparametrik dengan menggunakan estimator campuran *spline truncated* dan deret Fourier menggunakan metode *Ordinary Least Square*. Berdasarkan hasil kajian diperoleh estimator kurva regresi parametrik linear adalah  $\tilde{f}(x) = C(k(\tilde{y}), \tilde{a})$ , estimator kurva regresi *spline truncated* adalah  $\tilde{g}(t) = K(k(\tilde{y}), \tilde{a})$  dan estimator kurva regresi deret Fourier adalah  $\tilde{h}(z) = D\tilde{a}$ . Selanjutnya diperoleh estimator kurva regresi semiparametrik campuran *spline truncated* dan deret Fourier  $\mu(x, t, z) = M(k(\tilde{y}), \tilde{a})$  dimana  $M(k(\tilde{y}), \tilde{a}) = C(k(\tilde{y}), \tilde{a}) + K(k(\tilde{y}), \tilde{a}) + D\tilde{a}$ . Estimator campuran ini tergantung pada titik knot dan parameter *bandwith*. Estimator terbaik diperoleh dengan cara meminimumkan fungsi *Generalized Cross Validation*. Model regresi semiparametrik campuran yang diperoleh digunakan untuk memodelkan data kasus Angka Harapan Hidup (AHH) di Provinsi Jawa Tengah. Model estimator campuran tersebut menghasilkan  $R^2$  sebesar 82.22%.

**Kata Kunci:** Regresi Semiparametrik, *Spline Truncated*, Deret Fourier, OLS.

## ABSTRACT

### MIXTURE SEMIPARAMETRIC REGRESSION MODEL USING TRUNCATED SPLINE AND FOURIER SERIES

By

FARIDA RIZA UMAMI

19/448799/PPA/05882

Semiparametric regression approach is one of approach to estimate the regression curve. Given paired data

$(x_{1i}, x_{2i}, \dots, x_{pi}, t_{1i}, t_{2i}, \dots, t_{qi}, z_{1i}, z_{2i}, \dots, z_{ri})$ . The relationship between predictor variables and response variables follows a mixed semiparametric regression model. The semiparametric regression model that contains these variables is expressed as  $y_i = \mu(x_{1i}, x_{2i}, \dots, x_{pi}, t_{1i}, t_{2i}, \dots, t_{qi}, z_{1i}, z_{2i}, \dots, z_{ri}) + \epsilon_i$ . The regression curve  $\mu$  assumed to be additive can be written as:  $y_i = \sum_{j=1}^p f_j(x_{ji}) + \sum_{s=1}^q g_s(t_{si}) + \sum_{m=1}^r h_m(z_{mi}) + \epsilon_i$ .

Theoretically, the  $f_j(x_{ji})$  regression curve can be approximated by a linear function, the  $g_s(t_{si})$  regression curve is approximated by a linear truncated spline function. Meanwhile, the  $h_m(z_{mi})$  regression curve can be approximated by a Fourier Series. The aim of this research is to get an estimator in semiparametric regression using the estimator mixture *spline truncated* and Fourier series using *Ordinary Least Square* method. Based on the results of the study obtained linear parametric regression curve estimator is  $\tilde{f}(x) = C(k(\tilde{y}), \tilde{a})$ , the regression curve estimator *spline truncated* is  $\tilde{g}(t) = K(k(\tilde{y}), \tilde{a})$ , and the Fourier series regression curve estimator is  $\tilde{h}(z) = D\tilde{a}$ . Furthermore, we get the mixture semiparametric regression curve estimator *spline truncated* and Fourier series  $\mu(x, t, z) = M(k(\tilde{y}), \tilde{a})$  where dimana  $M(k(\tilde{y}), \tilde{a}) = C(k(\tilde{y}), \tilde{a}) + K(k(\tilde{y}), \tilde{a}) + D\tilde{a}$ . This mixture estimator depends on the knot points and *bandwith* parameter. The best estimator is obtained by minimizing the *Generalized Cross Validation* function. The obtained mixture semiparametric regression model will be used to modeling case data on Life Expectancy Rates (AHH) in Central Java Province. The result of mixture estimator model is  $R^2$  sebesar 82.22%.

**Keywords:** Semiparametric Regression, *Spline Truncated*, Fourier Series, OLS.