

Analisis Pengaruh Angka Kematian Bayi Terhadap Angka Harapan Hidup di Provinsi Jawa Timur Berdasarkan Estimator *Least Square Spline*

Anies Yulinda W.¹, Trias Novia L.², Melati Tegarina³ & Nur Chamidah^{4,*}

^{1,2,3,4}Departemen Matematika, Fakultas Sains dan Teknologi, Kampus C
Universitas Airlangga, Jl. Mulyorejo, Surabaya

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

Abstract. Life expectancy can be used to evaluate the government's performance for improving the welfare of the population in the health sector. Life expectancy is closely related to infant mortality rate. Theoretically, decreasing of infant mortality rate will cause increasing of life expectancy. A statistical method that can be used to model life expectancy is nonparametric regression model based on least square spline estimator. This method provides high flexibility to accommodate pattern of data by using smoothing technique. The best estimated model is order one spline model with one knot based on minimum generalized cross validation (GCV) value of 0.607. Each increasing of one infant mortality rate unit will cause decreasing of life expectancy of 0.2314 for infant mortality rate less than 27, and of 0.0666 for infant mortality rate more than and equals to 27. In addition, based on mean square error (MSE) of 0.492 and R^2 value of 76.59% for nonparametric model approach compared with MSE of 0.634 and R^2 value of 71.8% for parametric model approach, we conclude that the use of nonparametric model approach based on least square spline estimator is better than that of parametric model approach.

Keywords: *infant mortality rate; least square spline estimator; life expectancy; nonparametric regression.*

1 Pendahuluan

Pemerintah Indonesia terus melakukan pembangunan di semua bidang kehidupan sebagai upaya dalam meningkatkan kesejahteraan masyarakat. Salah satu indikator yang digunakan untuk mengevaluasi pelaksanaan pembangunan tersebut adalah indikator kesehatan. Kesehatan merupakan hak dasar manusia dan merupakan salah satu aspek kualitas sumber daya manusia yang penting. Sumber daya manusia (SDM) yang sehat secara fisik diharapkan menjadi manusia berkualitas sehingga dapat ikut berperan aktif dalam pembangunan untuk mewujudkan kesejahteraan rakyat.

Secara umum, angka harapan hidup (AHH) merupakan salah satu alat yang dapat digunakan untuk mengevaluasi kinerja pemerintah dalam peningkatan kesejahteraan penduduk di bidang kesehatan. Angka harapan hidup adalah perkiraan lama hidup rata-rata penduduk dengan asumsi tidak ada perubahan pola mortalitas menurut umur. Angka harapan hidup (AHH) berhubungan erat dengan angka kematian bayi (AKB).

Secara teoritis, menurunnya angka kematian bayi akan menyebabkan meningkatnya angka harapan hidup. Angka kematian bayi merupakan banyaknya kematian bayi usia 0 tahun dari setiap 1000 kelahiran hidup. Angka kematian bayi mempunyai korelasi dengan angka harapan hidup karena dalam perhitungan angka harapan hidup digunakan indikator angka kematian bayi di dalamnya.

Berdasarkan data BPS tahun 2016, angka kematian bayi di Jawa Timur mencapai angka 23,6 bayi dari total 1000 kelahiran. Angka tersebut sudah di atas target pemerintah yaitu 24 bayi dari 1000 kelahiran. Menurut pengkategorian Pramono [1], angka tersebut tergolong sangat rendah, namun menurut pemerintah, angka yang mendekati target tersebut masih tergolong tinggi. Penelitian sebelumnya tentang angka kematian bayi dan angka harapan hidup di Jawa Timur telah dilakukan oleh Juliandari dan Budiantara [2]. Penelitian tersebut didasarkan pada data angka kematian bayi dan angka harapan hidup di Jawa Timur pada tahun 2011, sedangkan penelitian ini didasarkan pada data tahun 2016.

Analisis statistika yang dapat menggambarkan hubungan fungsional antara variabel respon dan variabel prediktor adalah analisis regresi. Ada dua pendekatan yang dapat dilakukan untuk mengestimasi hubungan fungsional tersebut, yaitu pendekatan model regresi parametrik dan nonparametrik. Pendekatan parametrik digunakan jika peneliti mengasumsikan bentuk tertentu dari fungsi regresinya seperti linier, kuadratik, kubik, eksponensial atau bisa disebut pendekatan global. Dalam pendekatan model regresi nonparametrik fungsi regresi hanya diasumsikan *smooth* sehingga memberikan fleksibilitas yang tinggi terhadap bentuk yang mungkin dari fungsi regresinya. Salah satu teknik *smoothing* yang dapat digunakan dalam pengestimasi model regresi nonparametrik adalah estimator spline yang terdiri dari potongan polinomial yang tersegmen oleh titik knot (fungsi *truncated*) sehingga dapat mengakomodasi sifat kelokalan (Chamidah, et al [3]). Estimator spline yang diperoleh dengan menggunakan metode *least square* disebut *least square Spline*. Peneliti yang menggunakan estimator spline dalam regresi nonparametrik antara lain Lestari et al [4], Islamiyati, et al [5] and Ayidin and Yilmaz [6].

Dalam hubungan fungsional antara AHH dengan angka kematian bayi menunjukkan pola yang tidak linier (tidak bersifat global) sehingga diperlukan model yang dapat mengakomodasi kelokalan yaitu estimator spline. Selain memodelkan AHH terhadap AKB berdasarkan estimator spline, dalam artikel ini juga ditunjukkan kelebihan model regresi nonparametrik berdasarkan estimator spline jika dibandingkan pendekatan parametrik global berdasarkan kriteria *goodness of fit* yaitu MSE dan R^2 . Untuk pengolahan dan analisis data digunakan program OSS-R.

2 Estimator Least Square Spline

Spline least square merupakan potongan-potongan polinomial dengan segmen-segmen polinomial berbeda yang digabungkan bersama pada knot dengan tetap menjamin sifat continuity (Chamidah[7]).

Titik knot merupakan titik perpaduan bersama dimana terjadi perubahan perilaku fungsi. Secara umum fungsi spline order p dengan titik knot $\tau_1, \tau_2, \dots, \tau_k$ pada sebarang fungsi $f(x)$ disajikan dalam persamaan sebagai berikut :

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

dengan

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

Jika diberikan untuk $\lambda = (\tau_1, \tau_2, \dots, \tau_k)$ maka estimator dari f dengan orde p dan banyaknya knot dapat dinyatakan dengan :

$$f_\lambda(x) = \sum_{j=0}^{p+k} \beta_{\lambda j} \varphi_j(x) \quad (2)$$

dengan

$$\varphi_j(x) = \begin{cases} x^j & \text{untuk } 0 \leq j \leq p \\ (x - \tau_{j-p})_+^p & \text{untuk } p + 1 \leq j \leq p + k \end{cases}$$

3 Pemilihan Titik Knot Optimum Estimator Least Square Spline

Agar diperoleh spline yang optimal perlu ditentukan titik knot yang optimal. Salah satu metode untuk mendapatkan titik knot optimal adalah metode *Generalized Cross Validation* (GCV) yang didefinisikan sebagai berikut :

$$GCV(\tau_1, \tau_2, \dots, \tau_k) = \frac{MSE(\tau_1, \tau_2, \dots, \tau_k)}{(n^{-1} \text{tr}[I - H(\tau_1, \tau_2, \dots, \tau_k)])^2} \quad (3)$$

dengan

$$MSE(\tau_1, \tau_2, \dots, \tau_k) = \frac{1}{n} \sum_{i=1}^n (y_i - f(x_i))^2 \quad (4)$$

dengan $\tau_1, \tau_2, \dots, \tau_k$ adalah titik knot pertama hingga knot ke- k .

Nilai titik knot yang optimal diperoleh berdasarkan nilai *GCV* yang minimum.

4 Metodologi

Data yang digunakan merupakan data sekunder tentang Angka Harapan Hidup (AHH) sebagai variabel respon dan Angka Kematian Bayi (AKB) sebagai variabel prediktor di 38 kabupaten/kota provinsi Jawa Timur tahun 2016. Dalam membangun model digunakan data *in sample* sebanyak 30 kabupaten/kota yang dipilih secara acak sedangkan 8 kabupaten/kota digunakan sebagai data *out sample* untuk prediksi. Data sekunder diperoleh melalui publikasi di website resmi Badan Pusat Statistik (BPS) Jawa

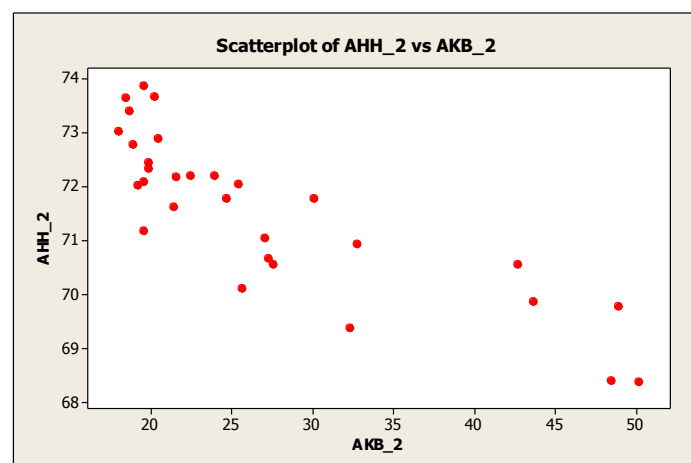
Timur yaitu <https://www.bps.go.id/>. Variabel dalam penelitian ini adalah angka kematian bayi (per 1000 kelahiran) yang menjadi variabel prediktor dan angka harapan hidup (tahun) yang menjadi variabel respon.

Langkah-langkah dalam memodelkan AHH berdasarkan AKB pendekatan regresi nonparametrik berdasarkan estimator *Least – Square Spline* adalah sebagai berikut:

- a. Melakukan scatter plot antara Angka Kematian Bayi (x) terhadap Angka Harapan Hidup (y).
- b. Melakukan pencocokan kurva dengan pendekatan parametrik linier dan kuadratik.
- c. Menentukan orde dan jumlah knot optimal menggunakan metode kuantil berdasarkan kriteria GCV pada persamaan (3).
- d. Mengestimasi model terbaik untuk AHH berdasarkan AKB berdasarkan orde dan jumlah knot optimal dari hasil (b)
- e. Menghitung nilai MSE dan R^2
- f. Membuat plot hasil estimasi dan data observasi
- g. Membandingkan hasil model regresi parametrik dan model regresi nonparametrik berdasarkan nilai MSE dan R^2

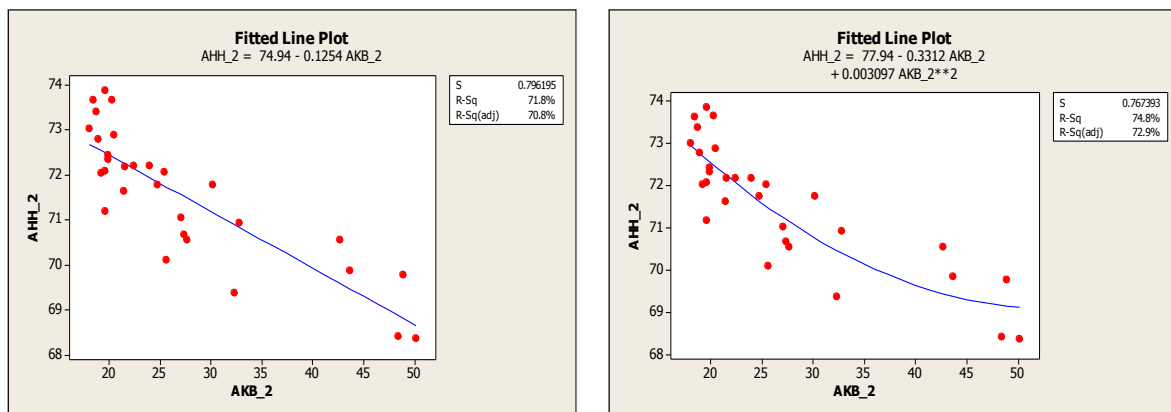
5 Pemodelan Angka Harapan Hidup berdasarkan Angka Kematian Bayi

Dalam pemodelan AHH berdasarkan AKB terlebih dahulu dibuat *scatter plot* AHH versus AKB untuk melihat pola hubungan kedua variabel tersebut yang ditunjukkan dalam Gambar 1 berikut ini:



Gambar 1. *Scatter plot* antara Angka Harapan Hidup (AHH) versus Angka Kematian Bayi (AKB)

Dalam Gambar 1 terlihat hubungan antara AHH dan AKB tidak membentuk pola tertentu seperti linier atau kuadratik. Berikut ini diberikan plot estimasi model dengan observasi dengan pendekatan parametrik linier dan kuadratik dalam Gambar 2.



Gambar 2 Plot antara estimasi (garis biru) dan observasi (titik merah) AHH versus AKB dengan pendekatan model regresi linier (kiri) dengan model regresi kuadratik (kanan)

Dengan pendekatan model regresi parametrik linier diperoleh R^2 sebesar 71,8% sedangkan dengan pendekatan model regresi parametrik kuadratik diperoleh R^2 sebesar 74,8%.

Selanjutnya untuk memodelkan dengan pendekatan regresi nonparametrik berdasarkan estimator *least square* spline ditentukan orde dan titik knot optimal berdasarkan kriteria GCV.

Berikut ini diberikan hasil perhitungan nilai GCV, MSE dan R^2 dalam Tabel 1 sebagai berikut:

Tabel 1. Hasil Perhitungan Nilai GCV, MSE dan R^2 berdasarkan jumlah knot dan Orde

Orde	Knot	Titik Knot	GCV	MSE	R^2
1	1	27	0,6072	0,4918	76,5933
2	1	39	0,6186	0,4646	77,8881
1	2	27 ; 43	0,6143	0,4614	78,0425
2	2	32 ; 33	0,6657	0,4623	78,0003
1	3	27 ; 33 ; 43	0,6597	0,4582	78,1963

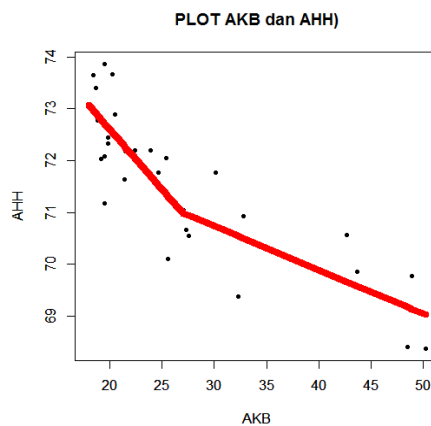
Berdasarkan Tabel 1, diperoleh model spline terbaik dengan orde optimal satu dan titik knot optimal 27 berdasarkan nilai GCV minimum sebesar 0,6072, MSE sebesar 0,4918 dan R^2 sebesar 76,59%. Estimasi model AHH berdasarkan AKB berdasarkan estimator least square spline ditunjukkan dalam persamaan (5) sebagai berikut:

$$\hat{y} = 77,2365 - 0,2314x + 0,1468(x - 27) \quad (5)$$

Persamaan (5) dapat diuraikan menjadi persamaan (6) berikut ini :

$$\hat{y} = \begin{cases} 77,2365 - 0,2314x & , \text{ untuk } x < 27 \\ 73,27729 - 0,0846x & , \text{ untuk } x \geq 27 \end{cases} \quad (6)$$

Untuk mengetahui seberapa besar perubahan Angka Harapan Hidup (AHH) akibat perubahan nilai Angka Kematian Bayi (AKB), perlu dilakukan interpretasi terhadap hasil dari model pada persamaan (6). Setiap kenaikan satu satuan angka kematian bayi (AKB) akan mengakibatkan angka harapan hidup (AHH) menurun sebesar 0,2314, jika AKB kurang dari 27 sedangkan jika AKB lebih dari sama dengan 27, maka setiap kenaikan satu satuan angka kematian bayi (AKB) akan mengakibatkan angka harapan hidup (AHH) menurun sebesar 0,0846. Berikut adalah plot observasi dan hasil estimasi untuk Angka Harapan Hidup (AHH) berdasarkan metode *least square spline*.



Gambar 3. Plot estimasi AHH terhadap AKB berdasarkan estimator *least square spline*

Berdasarkan hasil estimasi model terbaik pada persamaan (6), dilakukan pengestimasi AHH untuk data *outsample* yaitu data yang tidak digunakan untuk membangun model (*in sample*) dan kemudian dibandingkan hasil estimasi AHH dengan pendekatan model parametrik linier sehingga didapat hasil sebagai berikut:

Tabel 2. Perbandingan hasil estimasi AHH berdasarkan estimator least square spline dengan parametrik linier pada data *out sample*

AKB (x)	AHH (y)	(\hat{y})	
		<i>Least Square Spline</i>	<i>Parametrik</i>
15,60	73,09	73,62	72,98376
17,04	72,68	73,29	72,803184
35,00	70,67	70,94	70,551
17,16	69,79	73,27	72,788136
42,34	67,62	70,45	69,630564
41,90	66,95	70,48	69,68574
45,40	65,89	70,25	69,24684
56,13	66,31	69,53	67,901298

Sesuai kriteria *goodness of fit* yaitu MSE dan R^2 dapat ditunjukkan pada Tabel 3 berikut bahwa pemodelan AHH berdasarkan AKB dengan pendekatan model regresi nonparametrik menggunakan estimator spline lebih baik daripada pendekatan parametrik linier baik untuk *in sample* maupun *out sample*.

Tabel 3. Perbandingan *Goodness of Fit* pada Model Parametrik dan Nonparametrik untuk Data *In Sample* dan *Out Sample*

		MSE	R^2
<i>IN SAMPLE</i>	<i>Least Square Spline</i>	0,492	76,593
	Parametrik	0,634	71,8
<i>OUT SAMPLE</i>	<i>Least Square Spline</i>	1,512	78,622
	Parametrik	2,021	78,6

6 Kesimpulan

Hasil pemodelan AHH berdasarkan AKB di provinsi Jawa Timur menggunakan estimator *least square* spline diperoleh hasil bahwa apabila nilai AKB bertambah sebesar 1 satuan, maka nilai AHH akan berkurang sebesar 0,2314 untuk nilai AKB kurang dari 27 dan berkurang sebesar 0,0846 ketika nilai AKB lebih dari sama dengan 27. Berdasarkan kriteria *goodness of fit* yaitu MSE dan R^2 model dengan pendekatan nonparametrik berdasarkan estimator spline lebih baik dibandingkan model regresi parametrik linier dalam pemodelan AHH berdasarkan AKB di propinsi Jawa Timur.

7 Daftar Pustaka

- [1] Pramono, M.S., Suci, Wulansari., dan Sutikno., 2012, Pemetaan Determinan Angka Kematian Bayi di Jawa Timur Berdasarkan Indikator Indeks Pembangunan

- Kesehatan Masyarakat., *Bulletin of Health System Research*, **15**: 38-46. Juliandari, N.Y.N., dan Budiantara, I.N., 2014, *Pemodelan Angka Harapan Hidup dan Angka Kematian Bayi di Jawa Timur dengan Pendekatan Regresi Nonparametrik Spline Birespon*, ITS Paper, Surabaya.
- [2] Chamidah, N., Kurniawan, A., Zaman, B., Muniroh, L., 2018, Least square-spline estimator in Multiresponse semiparametric regression model for estimating median growth charts of children in East Java Indonesia, *Far East Journal of Mathematical Sciences (FJMS)*, 107 (2), 295-307.
- [3] Lestari, B., Budiantara, I.N., Sunaryo, S., and Mashuri, M., 2012, Spline smoothing for multi-response nonparametric regression model in case of heteroscedasticity of variance. *Journal of Mathematics and Statistics*, 8 (3), 337-384.
- [4] Islamiyati A., Fatmawati and Chamidah N., 2018, Estimation of Covariance Matrix on Bi-Response Longitudinal Data Analysis with Penalized Spline Regression. *Journal of Physics: Conf. Series*, 979 012093, IOP Publishing. doi :10.1088/1742-6596/979/1/012093.
- [5] Aydin, D. and Yilmaz, E., 2018, Modified spline regression based on randomly right-censored data: A comparative study, *Communications in Statistics-Simulation and Computation*, Volume 47, Issue 9.
- [6] Chamidah, N., and 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), 1433-1453.