

PEMODELAN FAKTOR-FAKTOR YANG MEMENGARUHI KESEMBUHAN TUBERKULOSIS DI PROVINSI JAWA TIMUR DENGAN REGRESI NONPARAMETRIK *SPLINE TRUNCATED*

Niken Evitasari¹, Sri Sulistijowati Handajani², Hasih Pratiwi³

^{1,2,3}Progam Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Sebelas Maret

¹evitaniken1@student.uns.ac.id, ²rr_ssh@staff.uns.ac.id, ³hpratiwi@staff.uns.ac.id

ABSTRAK

Tuberkulosis (TBC) adalah penyakit yang disebabkan oleh kuman *Mycobacterium Tuberculosis*. Penularan TBC dapat melalui keadaan lingkungan (udara) dan perilaku masyarakat. Provinsi Jawa Timur memiliki 64.764 kasus tuberkulosis. Proses penyembuhan penyakit TBC merupakan hal yang penting guna mengurangi jumlah kasus atau memutuskan rantai penularan. Banyak faktor yang diduga memengaruhi angka kesembuhan tuberkulosis, sehingga dilakukan analisis yang bertujuan untuk memodelkan faktor-faktor yang berpengaruh secara signifikan. Penelitian ini terdiri dari enam variabel prediktor, diantaranya tempat pengelolaan makanan (TPM) memenuhi syarat kesehatan, rumah tangga memiliki sanitasi layak, pengobatan lengkap tuberkulosis, tempat umum (TU) memenuhi syarat kesehatan, jaminan kesehatan, dan penduduk merokok di atas 15 tahun. Metode yang digunakan dalam pemodelan angka kesembuhan tuberkulosis di Jawa Timur tahun 2020 adalah regresi nonparametrik *spline truncated* linier dengan kombinasi titik *knot* (2,3,1,3,2,3). Pemilihan titik *knot* optimal tersebut didapat dari nilai *Generalized Cross Validation* (GCV) terendah sebesar 30,7126. Hasil dari penelitian ini diperoleh nilai koefisien determinasi disesuaikan sebesar 82,46 persen dan nilai rata-rata kuadrat terkecil (RKS) sebesar 15,4028 yang menunjukkan bahwa model regresi *spline truncated* linier dengan kombinasi titik *knot* layak atau baik digunakan dalam pemodelan angka kesembuhan tuberkulosis di Jawa Timur tahun 2020.

Kata Kunci : *Spline truncated*, kesembuhan tuberkulosis, *knot*, *Generalized Cross Validation* (GCV)

ABSTRACT

Tuberculosis (TB) is a disease caused by *Mycobacterium tuberculosis*. TB transmission can be through environmental conditions (air) and community behavior. East Java Province has 64,764 cases of tuberculosis. The process of curing TB disease is essential to reduce the number of cases or break the chain of transmission. Many factors are thought to influence the tuberculosis cure rate, so an analysis was carried out that aims to model the factors that have a significant effect. This study consisted of six predictor variables, including food management places meeting health requirements, households having proper sanitation, complete tuberculosis treatment, public places meeting health requirements, health insurance, and smoking residents over 15 years. The method used in modeling the tuberculosis cure rate in East Java in 2020 is nonparametric spline truncated linear regression with a combination of knot points (2,3,1,3,2,3). The optimal knot point selection was obtained from the lowest Generalized Cross Validation (GCV) value of 30.7126. This results of study obtained the value of the adjusted coefficient of determination 82.46 percent and Mean Square Error (MSE) of 15.4028, which indicates that the linear truncated spline regression model with a combination of knot points is feasible or good to use in modeling the tuberculosis cure rate in East Java in 2020.

Keywords: *Spline truncated*, tuberculosis cure, knots, *Generalized Cross Validation* (GCV)

PENDAHULUAN

Tuberkulosis (TBC) adalah penyakit yang disebabkan oleh kuman *Mycobacterium Tuberculosis*, dapat menular melalui keadaan lingkungan (udara) dan perilaku masyarakat. Menurut Dinas Kesehatan Provinsi Jawa Timur (2020), tuberkulosis menjadi penyebab kematian ke-9 di dunia setelah HIV/AIDS. Penyebaran penyakit ini bermula dari orang yang sudah terkena TBC, bersin atau batuk, air liur terkontaminasi keluar dan dihirup oleh orang lain. Orang yang memiliki kekebalan tubuh rendah dapat dengan mudah tertular penyakit TBC. Organ yang biasanya diserang oleh TBC tidak hanya paru-paru, tetapi juga anggota tubuh lainnya, diantaranya sistem saraf pusat, kelenjar getah bening, jantung, dan lain-lain.

Menurut WHO dalam *Global TB Report* (2020), Indonesia memiliki jumlah kasus terinfeksi tuberkulosis urutan kedua terbesar di dunia setelah India. Menurut Dinas Kesehatan Provinsi Jawa Timur (2020), sebanyak 64.764 jumlah kasus tuberkulosis di Jawa Timur yang terdaftar dan diobati. Penderita tuberkulosis yang menerima pengobatan lengkap sebanyak 32.809 orang dan total penderita yang sembuh sebanyak 57.602 orang. Proses pengobatan penyakit tuberkulosis juga terdapat kasus kematian sebanyak 2.323 orang.

TBC merupakan prioritas nasional program pengendalian penyakit, karena memberikan dampak terhadap ekonomi, kualitas hidup dan kematian. Proses penyembuhan penyakit tuberkulosis merupakan hal yang penting guna mengurangi jumlah kasus atau memutuskan rantai penularan penyakit ini. Proses penyembuhan penyakit tuberkulosis tidak hanya melalui faktor medis, melainkan kondisi lingkungan juga dapat memengaruhinya. Menurut Andriyati dan Rohaeti (2019), faktor-faktor yang berpengaruh terhadap ketahanan hidup pasien tuberkulosis yaitu, kebiasaan merokok, sanitasi lingkungan dan pencahayaan rumah. Menurut Anggreni dkk. (2016), dalam penelitiannya menggunakan regresi nonparametrik *spline truncated* menyatakan bahwa tempat umum sehat merupakan salah satu faktor yang berpengaruh terhadap kasus tuberkulosis di Provinsi Bali. Penelitian lainnya juga dilakukan oleh Murtantiningasih dan Wahyono (2010), mengenai faktor-faktor yang berpengaruh terhadap kesembuhan penderita tuberkulosis yaitu status gizi, pendapatan, dan keteraturan berobat. Berdasarkan penelitian sebelumnya, dapat disimpulkan banyak faktor yang memengaruhi kesembuhan penyakit tuberkulosis.

Analisis untuk melihat pengaruh variabel prediktor terhadap variabel respon adalah analisis regresi. Ada dua pendekatan dalam regresi, yaitu parametrik dan nonparametrik. *Spline* merupakan pendekatan regresi nonparametrik yang dilakukan karena bentuk kurva tidak diketahui polanya. *Truncated* diartikan sebagai fungsi potongan. *Spline truncated* adalah model polinomial yang terpotong-potong sehingga memiliki fleksibilitas yang lebih tinggi dibandingkan dengan model polinomial biasa, dikarenakan *spline* mampu menghasilkan fungsi regresi yang sesuai dengan data (Anggraeni dkk., 2016).

Penelitian kasus tuberkulosis sebelumnya dilakukan oleh Darma dkk. (2019), untuk memodelkan faktor-faktor yang berpengaruh signifikan terhadap kesembuhan tuberkulosis di Kota Surabaya menggunakan regresi nonparametrik *spline truncated* dengan menggunakan kombinasi titik *knot* dan nilai GCV minimum. Penelitian tentang tuberkulosis juga dilakukan oleh Khauliasari (2020), menggunakan metode *Mixed Geographically Weighted Poisson Regression* (MGWPR) untuk menentukan faktor-faktor dari persebaran jumlah penderita tuberkulosis di Surabaya. Keuntungan menggunakan metode MGWPR ini adalah dapat diketahui perbedaan di setiap lokasi secara lokal dan global. Hasil dari penelitian ini yaitu lebih baik model MGWPR jika dibandingkan dengan model GWPR dan regresi *poisson*, karena memiliki nilai AIC terkecil dan koefisien determinasi terbesar. Penelitian sebelumnya juga dilakukan oleh Fatmawati dkk. (2019), yang membahas tentang perbandingan menggunakan metode *spline truncated* dan *smoothing spline* berdasarkan nilai rata-rata kuadrat sesatan (RKS).

Berdasarkan penjelasan sebelumnya, dilakukan analisis yang bertujuan untuk memodelkan faktor-faktor yang berpengaruh signifikan terhadap kesembuhan penyakit tuberkulosis di Provinsi Jawa Timur tahun 2020 menggunakan regresi nonparametrik *spline truncated* linier. Terdapat enam variabel prediktor yang diduga memengaruhi kesembuhan tuberkulosis yaitu tempat pengelolaan makanan (TPM) memenuhi syarat kesehatan, rumah tangga memiliki sanitasi layak, pengobatan lengkap

tuberkulosis, tempat umum (TU) memenuhi syarat kesehatan, jaminan kesehatan, dan penduduk merokok di atas 15 tahun. Diduga pola data dari setiap variabel prediktor dengan variabel respon tidak mengikuti suatu pola tertentu sehingga dilakukan pemodelan menggunakan regresi nonparametrik, salah satunya adalah regresi *spline truncated*. Keuntungan menggunakan metode ini adalah dapat mengatasi pola data yang naik atau turun secara tajam dengan bantuan titik-titik *knot* sehingga sesatan yang dihasilkan pada model lebih rendah jika dibandingkan dengan metode lain (Hardle, 1990). Titik *knot* merupakan titik yang terdapat pola perubahan fungsi. Pemilihan model regresi terbaik berdasarkan titik *knot* yang optimum dapat menggunakan GCV yang minimum.

METODE PENELITIAN

Penelitian ini termasuk dalam penelitian kuantitatif pada data kesembuhan tuberkulosis di Provinsi Jawa Timur tahun 2020. Penelitian ini memiliki tujuan untuk melihat pengaruh variabel prediktor terhadap variabel respon menggunakan metode regresi nonparametrik *spline truncated*. Data merupakan data sekunder bersumber dari publikasi Dinas Kesehatan Provinsi Jawa Timur dan Badan Pusat Statistik Provinsi Jawa Timur melalui *website* yang terdiri dari satu variabel respon dan enam variabel prediktor yang dijelaskan pada Tabel 1.

Tabel 1. Variabel Penelitian

Variabel	Deskripsi
Y	Persentase kesembuhan tuberkulosis
X_1	Persentase tempat pengelolaan makanan (TPM) memenuhi syarat kesehatan
X_2	Persentase rumah tangga memiliki sanitasi layak
X_3	Persentase pengobatan lengkap tuberkulosis
X_4	Persentase tempat umum memenuhi syarat kesehatan
X_5	Persentase kepemilikan jaminan kesehatan
X_6	Persentase penduduk merokok di atas 15 tahun

Subjek penelitian berupa 38 kabupaten/kota yang ada di Provinsi Jawa Timur, dianalisis menggunakan *software* Minitab dan R dengan langkah-langkah sebagai berikut.

1. Melakukan analisis statistika deskriptif
2. Menentukan pola hubungan masing-masing variabel prediktor terhadap variabel respon
3. Melakukan pemodelan menggunakan regresi *spline truncated* dengan berbagai jumlah titik *knot*.
4. Menentukan titik *knot* optimal berdasarkan nilai GCV terkecil.
5. Melakukan pemodelan regresi *spline truncated* menggunakan titik *knot* optimal.
6. Menguji signifikansi parameter secara serentak dan parsial
7. Menguji asumsi sesatan
8. Menginterpretasi model regresi *spline truncated* terbaik yang diperoleh.

HASIL DAN PEMBAHASAN

Karakteristik Angka Kesembuhan Tuberkulosis di Jawa Timur

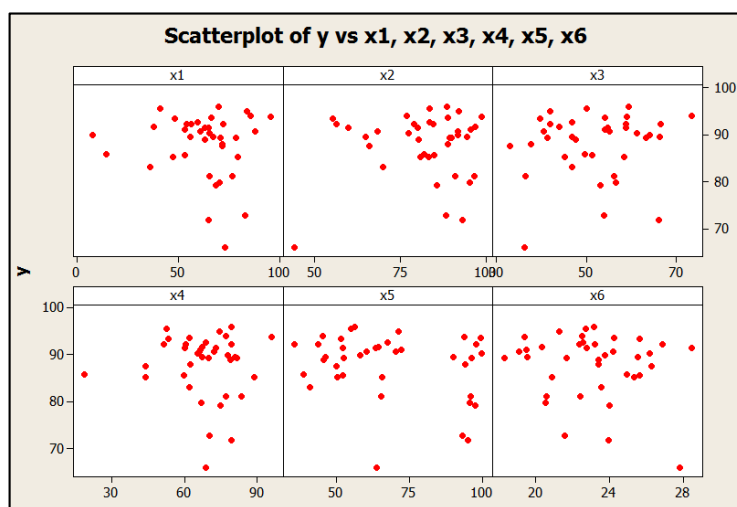
Faktor-faktor yang diduga mempengaruhi angka kesembuhan tuberkulosis di Jawa Timur beserta karakteristiknya dijelaskan oleh Tabel 2.

Tabel 2. Statistik Deskriptif

Variabel	Rata-rata	Variansi	Minimum	Maksimum
Persentase kesembuhan tuberkulosis	87,76	46,65	65,9	96
Persentase tempat pengelolaan makanan memenuhi syarat kesehatan	62,99	336,14	7,8	95,8
Persentase rumah tangga memiliki sanitasi layak	82,07	171,71	44,07	98,71
Persentase pengobatan lengkap tuberkulosis	51,76	98,31	32,8	73,4
Persentase tempat umum memenuhi syarat kesehatan	68,07	198,89	18,8	96
Persentase kepemilikan jaminan kesehatan	68,16	440,1	35,77	99,74
Persentase penduduk merokok di atas 15 tahun	23,084	6,242	18,29	28,45

Pola Hubungan

Scatter plot setiap variabel prediktor dengan variabel respon dijelaskan oleh Gambar 1.



Gambar 1. Pola Hubungan antar Variabel

Gambar 1 menunjukkan *scatter plot* untuk variabel respon dengan ke-enam variabel prediktor menyebar secara acak dan tidak diketahui bentuk polanya, sehingga digunakan pendekatan nonparametrik, salah satunya regresi nonparametrik *spline truncated*.

Pemilihan Titik *Knot* Optimal

Estimasi model terbaik didapatkan dari titik *knot* yang optimal. Titik *knot* adalah titik terjadinya perubahan pola kurva atau fungsi regresi. Titik *knot* optimal dihitung menggunakan GCV, karena sederhana (efisien dalam perhitungan), invarian terhadap transformasi dan tidak memerlukan

perhitungan σ^2 (Hardle, 1990). Titik *knot* optimum didapat dari nilai GCV minimum dimana sesatan harus berdistribusi normal (Nussbaum, 1985).

$$GCV_i(k) = \frac{RKS(k_1, k_2, \dots, k_r)}{(n^{-1} \text{trace}[I - A(k_1, k_2, \dots, k_r)])^2}$$

$$RKS(k_1, k_2, \dots, k_r) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$A(k) = X(X^T X)^{-1} X^T$$

Dengan k_1, k_2, \dots, k_r adalah titik-titik *knot* dan $A(k)$ adalah matriks gaussian.

Tabel 3. Titik *Knot* Optimal

Jumlah <i>Knot</i>	X_1	X_2	X_3	X_4	X_5	X_6	GCV
1 titik <i>knot</i>	89,5143	94,8071	70,5	90,4857	95,1707	27,7243	53,2482
2 titik <i>knot</i>	76,9423	87,0014	64,7	79,4571	86,0321	26,2727	50,3017
	89,5143	94,8071	70,5	90,4857	95,1707	27,7243	
3 titik <i>knot</i>	76,9423	87,0014	64,7	79,4571	86,0321	26,2727	40,7998
	83,2286	90,9043	67,6	84,9714	90,6014	26,9986	
	89,5143	94,8071	70,5	90,4857	95,1707	27,7243	
Kombinasi titik <i>knot</i> (2,3,1,3,2,3)	76,9423	87,0014	70,5	79,4571	86,0321	26,2727	30,7126
	89,5143	90,9043		84,9714	95,1707	26,9986	
		94,8071		90,4857		27,7243	

Tabel 3 menunjukkan model regresi nonparametrik *spline truncated* linier terbaik menggunakan jumlah kombinasi titik *knot* (2,3,1,3,2,3) karena memiliki nilai GCV terkecil sebesar 30,7126. Hasil ini selanjutnya digunakan dalam pemodelan angka kesembuhan tuberkulosis di Jawa Timur.

Pemodelan Regresi *Spline Truncated* menggunakan Titik *Knot* Optimal

Analisis regresi *spline* digunakan ketika kurva regresi tidak membentuk sebuah pola tertentu. Regresi *spline* dapat mengatasi data yang berubah pada interval tertentu (naik turun secara tajam) sehingga dapat menyesuaikan dengan bantuan titik-titik *knot* (Eubank, 1999). Persamaan regresi *spline truncated* orde satu adalah (Sugiantari dan Budiantara, 2013)

$$y_i = \beta_0 + \sum_{j=1}^p \beta_{j1} x_{ji} + \sum_{j=1}^p \sum_{k=1}^r \beta_{j(1+k)} (x_{ji} - K_{jk})_+^1 + \varepsilon_i$$

Tabel 5 dan Tabel 7 menunjukkan model regresi *spline truncated* terbaik untuk angka kesembuhan tuberkulosis adalah

$$\hat{y} = \beta_0 + \beta_{11} x_{1i} + \beta_{12} (x_{1i} - k_{11})_+^1 + \beta_{13} (x_{1i} - k_{12})_+^1 + \beta_{21} x_{2i} + \beta_{22} (x_{2i} - k_{21})_+^1 + \beta_{23} (x_{2i} - k_{22})_+^1 + \beta_{24} (x_{2i} - k_{23})_+^1 + \beta_{31} x_{3i} + \beta_{32} (x_{3i} - k_{31})_+^1 + \beta_{41} x_{4i} +$$

$$\beta_{42}(x_{4i} - k_{41})_+^1 + \beta_{43}(x_{4i} - k_{42})_+^1 + \beta_{44}(x_{4i} - k_{43})_+^1 + \beta_{51}x_{5i} + \beta_{52}(x_{5i} - k_{51})_+^1 + \beta_{53}(x_{5i} - k_{52})_+^1 + \beta_{61}x_{6i} + \beta_{62}(x_{6i} - k_{61})_+^1 + \beta_{63}(x_{6i} - k_{62})_+^1 + \beta_{64}(x_{6i} - k_{63})_+^1$$

$$\hat{y} = 126,1944 + 0,1314x_{1i} - 1,727(x_{1i} - 76,9423)_+^1 + 2,4064(x_{1i} - 89,5143)_+^1 - 0,2461x_{2i} + 2,6868(x_{2i} - 87,0014)_+^1 - 6,0308(x_{2i} - 90,9043)_+^1 + 4,7451(x_{2i} - 94,8071)_+^1 + 0,171x_{3i} + 4,5322(x_{3i} - 70,5)_+^1 - 0,1589x_{4i} - 3,5911(x_{4i} - 79,5471)_+^1 + 6,7269(x_{4i} - 84,9714)_+^1 + 2,1111(x_{4i} - 90,4857)_+^1 + 0,2748x_{5i} - 2,9314(x_{5i} - 86,0321)_+^1 + 6,0157(x_{5i} - 95,1707)_+^1 - 1,5559x_{6i} + 12,6414(x_{6i} - 26,2727)_+^1 - 69,0342(x_{6i} - 26,9986)_+^1 + 97,8949(x_{6i} - 27,7243)_+^1$$

dengan

x_{ji} : variabel prediktor ke- j data ke- i

$\beta_{j(1+k)}$: estimasi parameter variabel prediktor ke- j ; (1+titik knot ke- k)

K_{jk} : titik knot variabel ke- j knot ke- k

Uji Signifikansi Parameter

Tujuan dari uji signifikansi adalah menentukan pengaruh variabel prediktor terhadap variabel respon.

Uji Serentak

Uji serentak bertujuan untuk mengetahui pengaruh sekelompok variabel prediktor terhadap variabel respon. Uji hipotesis yang digunakan adalah (Setyowati dkk., 2020).

H_0 : $\beta_{11} = \beta_{12} = \dots = \beta_{63} = 0$ (variabel prediktor secara serentak tidak memengaruhi variabel respon)

H_1 : minimal ada satu $\beta_{lj} \neq 0$; untuk $l = 1, 2, \dots, 1+r$ dan $j = 1, 2, \dots, 6$ (variabel prediktor secara serentak memengaruhi variabel respon)

Tabel 4. Analisis Ragam Regresi *Spline Truncated*

SV	DB	JK	RK	F_{hitung}
Regresi	$p(1+r)=20$	$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = 1464,103$	$\frac{JK_{reg}}{db_{re}} = 73,2052$	$\frac{RK_{reg}}{RK_{ses}} = 4,7524$
Sesatan	$n-(p(1+r))-1=17$	$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = 261,865$	$\frac{JK_{ses}}{db_{ses}} = 15,4038$	
Total	$n-1=37$	$\sum_{i=1}^n (y_i - \bar{y})^2 = 1725,968$		

Tabel 4 menunjukkan nilai $F_{hitung} = 4,7524$ dimana lebih besar dari $F_{tabel} = F_{0,05;20;17} = 2,2303$ dan nilai $p-value = 0,00102$ dimana nilai tersebut kurang dari $\alpha = 0,05$, maka H_0 ditolak yang berarti variabel prediktor secara serentak memengaruhi variabel respon.

Uji Parsial

Hasil uji serentak menunjukkan bahwa variabel prediktor secara serentak memengaruhi variabel respon, sehingga dilakukan uji parsial untuk mengetahui pengaruh setiap variabel prediktor terhadap variabel respon. Uji Hipotesis yang digunakan adalah (Setyowati dkk., 2020)

H_0 : $\beta_{jl} = 0$ (variabel prediktor ke- j memengaruhi variabel respon)

H_1 : $\beta_{jl} \neq 0$, $j = 1, 2, \dots, p$, $l = 1, 2, \dots, 1+r$ (variabel prediktor ke- j memengaruhi variabel respon)

Statistik uji parsial adalah

$$t_{hitung} = \frac{\widehat{\beta}_{jl}}{\sqrt{\text{var}(\widehat{\beta}_{jl})}}$$

$$\text{var}(\widehat{\beta}_j) = \text{diag}[(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{RKS})^2]$$

dengan $\text{var}(\widehat{\beta}_{jl})$ adalah elemen diagonal utama ke- j dari matriks $\text{var}(\widehat{\beta})$.

Hasil dari statistik uji parsial dijelaskan oleh Tabel 5.

Tabel 5. Hasil Uji Parsial Regresi *Spline Truncated*

Variabel	Parameter	Estimasi	t-hitung	p-value	Keterangan
	β_0	126,1944	6,9673	$3,13 \times 10^{-8}$	H_0 ditolak
X_1	β_{11}	0,1314	1,8839	0,0674	H_0 tidak ditolak
	β_{12}	-1,7269	-3,4455	0,0014	H_0 ditolak
	β_{13}	2,4064	1,4953	0,1433	H_0 tidak ditolak
X_2	β_{21}	-0,2461	-1,8077	0,0788	H_0 tidak ditolak
	β_{22}	2,6868	2,3149	0,0263	H_0 ditolak
	β_{23}	-6,0308	-2,5195	0,0162	H_0 ditolak
	β_{24}	4,7451	1,2244	0,2285	H_0 tidak ditolak
X_3	β_{31}	0,171	1,2766	0,2097	H_0 tidak ditolak
	β_{32}	4,5322	2,2712	0,029	H_0 ditolak
X_4	β_{41}	-0,1589	-1,8012	0,0798	H_0 tidak ditolak
	β_{42}	-3,5911	-3,3305	0,002	H_0 ditolak
	β_{43}	6,7269	2,3373	0,02494	H_0 ditolak
	β_{44}	2,1111	1,4953	0,1433	H_0 tidak ditolak
X_5	β_{51}	0,2748	2,1208	0,0407	H_0 ditolak
	β_{52}	-2,9314	-5,0892	$1,07 \times 10^{-5}$	H_0 ditolak
	β_{53}	6,0157	5,4473	$3,5 \times 10^{-6}$	H_0 ditolak
X_6	β_{61}	-1,5559	-2,9211	0,0059	H_0 ditolak
	β_{62}	12,6414	1,6083	0,1163	H_0 tidak ditolak
	β_{63}	-69,0343	-3,7447	0,0006	H_0 ditolak
	β_{64}	97,8949	5,1758	$8,19 \times 10^{-6}$	H_0 ditolak

Cara pengambilan keputusannya adalah H_0 ditolak apabila $|t_{hitung}| > t_{\frac{\alpha}{2}, n-(p+(1+r))-1} = t_{0,025;17} = 2,1098$ atau $pvalue < \alpha = 0,05$. Tabel 5 menunjukkan setiap variabel prediktor minimal ada satu yang H_0 ditolak maka ke-enam variabel prediktor signifikan berpengaruh terhadap kesembuhan tuberkulosis.

Pengujian Asumsi Sesatan

Tujuan pengujian ini adalah untuk mendapatkan model yang memiliki kriteria BLUE (Best Linier Unbiased Estimator). Model regresi yang memenuhi kriteria BLUE ketika digunakan sebagai estimator tidak bias, konsisten berdistribusi normal dan juga efisien.

Normalitas

Pengujian asumsi normalitas pada penelitian ini menggunakan *Kolmogorov-Smirnov*.

H_0 : sesatan mengikuti distribusi normal

H_1 : sesatan tidak mengikuti distribusi normal

Pengujian *Kolmogorov-Smirnov* menghasilkan $D_{hitung} = maks|F_s(x_i) - F_t(x_i)| = 0,1164$ dimana lebih kecil dari $D_{(n,\alpha)} = D_{(38;0,05)} = 0,215$ dan nilai $p-value = 0,6386$ dimana lebih dari $\alpha = 0,05$, maka H_0 tidak ditolak, artinya sesatan mengikuti distribusi normal.

Heteroskedastisitas

Asumsi identik terpenuhi apabila variansi antarsesatan homogen atau dapat disebut tidak terjadi heteroskedastisitas (Widajarno, 2007). Pengujian asumsi heteroskedastisitas pada penelitian ini menggunakan uji *Glejser* dengan cara meregresikan nilai mutlak sesatan dengan variabel prediktor (Nisa' dan Budiantara, 2016).

H_0 : $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_{38}^2 = \sigma^2$ (tidak terdapat heteroskedastisitas)

H_1 : minimal terdapat satu $\sigma_i^2 \neq \sigma^2$, untuk $i = 1, 2, \dots, n$ (terdapat heteroskedastisitas)

$$F_{hitung} = \frac{\frac{[\sum_{i=1}^n (|\varepsilon_i| - |\bar{\varepsilon}|)^2]}{v}}{[\sum_{i=1}^n (|\varepsilon_i| - |\hat{\varepsilon}_i|)^2]} \frac{v}{n - v - 1}$$

Nilai v adalah jumlah parameter, $v = p + (1 + r)$.

Tabel 6. Analisis Variansi Uji *Glesjer*

SV	DB	JK	RK	F_{hitung}	$P-value$
Regresi	20	77,223	3,8611	1,4774	0,2102
Sesatan	17	44,428	2,6134		
Total	37	121,65			

Tabel 6 menunjukkan nilai $F_{hitung} = 1,4774$ dimana lebih kecil dari $F_{\alpha;(v,n-v-1)} = F_{0,05;20,17} = 2,2303$ dan $p-value = 0,2102$ dimana lebih dari $\alpha = 0,05$ maka H_0 tidak ditolak, yang artinya sesatan tidak terdapat heteroskedastisitas.

Autokorelasi

Pengujian ada tidaknya autokorelasi salah satunya dapat menggunakan Uji *Durbin Watson*.

H_0 : $\rho = 0$ (tidak terdapat autokorelasi)

H_1 : $\rho \neq 0$ (terdapat autokorelasi)

Statistik uji yang digunakan adalah

$$d_{hit} = \frac{\sum_{i=2}^n (\varepsilon_i - \varepsilon_{i-1})^2}{\sum_{i=2}^n (\varepsilon_i^2)}$$

Tabel *Durbin Watson* menunjukkan nilai $d_L=1,1463$ dan $d_U=1,8641$. Hasil dari statistik uji didapat nilai d_{hit} sebesar 1,8283 dimana $d_L \leq d_{hit} \leq d_u$ dimana tidak dapat ditarik kesimpulan ada tidaknya autokorelasi pada sesatan. Cara pengambilan keputusan selain menggunakan nilai d_{hitung} dapat menggunakan p -value. Nilai p -value = 0,716 dimana lebih dari $\alpha = 0,05$ maka H_0 tidak ditolak, yang artinya tidak terdapat autokorelasi pada sesatan.

Kebaikan Model Regresi

Langkah ini bertujuan untuk melihat kebaikan model regresi nonparametrik *spline truncated* berdasarkan nilai rata-rata kuadrat terkecil (RKS) dan koefisien determinasi yang disesuaikan R_{Adj}^2 .

$$R_{Adj}^2 = 1 - \frac{n-1}{n-p} (1 - R^2)$$

Tabel 7. Kebaikan Model Regresi *Spline Truncated*

R_{Adj}^2	RKS
82,46 persen	15,4038

Tabel 9 menunjukkan bahwa angka kesembuhan tuberkulosis di Jawa Timur sesuai apabila dianalisis menggunakan regresi *spline truncated* karena memiliki nilai koefisien determinasi yang tinggi. Ke-enam variabel prediktor yang digunakan dalam penelitian ini secara serentak berpengaruh terhadap angka kesembuhan tuberkulosis sebesar 82,46 persen, sedangkan 17,54 persen dipengaruhi oleh variabel yang lain diluar persamaa atau variabel yang tidak digunakan dalam penelitian.

Interpretasi Model Regresi Nonparametrik *Spline Truncated*.

Fungsi *truncated* yang dibentuk (Sugiantara dan Budiantara, 2013)

$$(x_{ji} - K_{jk})_+ = \begin{cases} (x_{ji} - K_{jk})^q, & x_{ji} - K_k \geq 0 \\ 0, & x_{ji} - K_k < 0 \end{cases}$$

1. Apabila variabel selain X_1 konstan, pengaruh TPM memenuhi syarat kesehatan (X_1) terhadap kesembuhan tuberkulosis adalah

$$\hat{y} = 126,1944 + 0,1314x_{1i} - 1,727(x_{1i} - 76,9423)_+^1 + 2,4064(x_{1i} - 89,5143)_+^1$$

$$\hat{y} = \begin{cases} 126,1944 + 0,1314x_{1i}, & x_{1i} < 76,9423 \\ 132,8794 - 1,5956x_{1i}, & 76,9423 \leq x_{1i} < 89,5143 \\ 43,6665 + 0,8108x_{1i}, & x_{1i} \geq 89,5143 \end{cases}$$

Apabila TPM memenuhi syarat kesehatan kurang dari 76,9423 persen dan jika naik sebesar 1 persen, maka kesembuhan tuberkulosis meningkat 0,1314 persen. Apabila TPM memenuhi syarat kesehatan 76,9423 sampai 89,5143 persen dan jika naik 1 persen, maka kesembuhan tuberkulosis turun sebesar 1,5956 persen. Apabila TPM memenuhi syarat kesehatan 89,5143 persen atau lebih dan jika naik 1 persen, maka kesembuhan tuberkulosis meningkat 0,8108 persen.

2. Apabila variabel selain X_2 konstan, pengaruh rumah tangga yang memiliki sanitasi layak (X_2) terhadap kesembuhan tuberkulosis adalah

$$\hat{y} = 126,1944 - 0,2461x_{2i} + 2,6868(x_{2i} - 87,0014)_+^1 - 6,0308(x_{2i} - 90,9043)_+^1 + 4,7451(x_{2i} - 94,8071)_+^1$$

$$\hat{y} = \begin{cases} 126,1944 - 0,2461x_{2i}, & x_{2i} < 87,0014 \\ -107,561 + 2,4407x_{2i}, & 87,0014 \leq x_{2i} < 90,9043 \\ 440,6647 - 3,5901x_{2i}, & 90,9043 \leq x_{2i} < 94,8071 \\ -9,2048 + 1,155x_{2i}, & x_{2i} \geq 94,8071 \end{cases}$$

Apabila sanitasi layak kurang dari 87,0014 persen dan jika naik 1 persen, maka kesembuhan tuberkulosis turun 0,2461 persen. Apabila sanitasi layak 87,0014 sampai 90,9043 persen dan jika naik 1 persen, maka kesembuhan tuberkulosis meningkat 2,4407 persen. Apabila sanitasi layak 90,9043 sampai 94,8071 persen dan jika naik 1 persen, maka kesembuhan tuberkulosis turun 3,5901 persen. Apabila sanitasi layak 94,8071 persen atau lebih dan jika naik 1 persen, maka kesembuhan tuberkulosis meningkat 1,155 persen.

3. Apabila variabel selain X_3 konstan, pengaruh pengobatan lengkap semua kasus tuberkulosis (X_3) terhadap kesembuhan tuberkulosis adalah

$$\hat{y} = 126,1944 + 0,171x_{3i} + 4,5322(x_{3i} - 70,5)_+^1$$

$$\hat{y} = \begin{cases} 126,1944 + 0,171x_{3i}, & x_{3i} < 70,5 \\ -193,3257 + 4,7032x_{3i}, & x_{3i} \geq 70,5 \end{cases}$$

Apabila pengobatan lengkap kurang dari 70,5 persen dan jika naik sebesar 1 persen, maka kesembuhan tuberkulosis meningkat 0,171 persen. Apabila pengobatan lengkap 70,5 persen atau lebih dan jika naik sebesar 1 persen, maka kesembuhan tuberkulosis meningkat 4,7032 persen.

4. Apabila variabel selain X_4 konstan, pengaruh tempat umum memenuhi syarat kesehatan (X_4) terhadap kesembuhan tuberkulosis adalah

$$\hat{y} = 126,1944 - 0,1589x_{4i} - 3,5911(x_{4i} - 79,5471)_+^1 + 6,7269(x_{4i} - 84,9714)_+^1 + 2,1111(x_{4i} - 90,4857)_+^1$$

$$\hat{y} = \begin{cases} 126,1944 - 0,1589x_{4i}, & x_{4i} < 79,5471 \\ 411,856 - 3,75x_{4i}, & 79,5471 \leq x_{4i} < 84,9714 \\ -159,7381 + 2,9769x_{4i}, & 84,9714 \leq x_{4i} < 90,4857 \\ -350,7625 + 5,0088x_{4i}, & x_{4i} \geq 90,4857 \end{cases}$$

Apabila tempat umum sehat kurang dari 79,5471 persen dan jika naik 1 persen, maka kesembuhan tuberkulosis turun 0,1589 persen. Apabila tempat umum sehat 79,5471 sampai 84,9714 persen dan jika naik 1 persen, maka kesembuhan tuberkulosis turun sebesar 3,75 persen. Apabila tempat umum sehat 84,9714 sampai 90,4857 persen dan jika naik 1 persen, maka kesembuhan tuberkulosis meningkat 2,9769 persen. Apabila tempat umum sehat 90,4857 persen atau lebih dan jika naik 1 persen, maka kesembuhan tuberkulosis meningkat 5,0088 persen.

5. Apabila variabel selain X_5 konstan, pengaruh kepemilikan jaminan kesehatan (X_5) terhadap kesembuhan tuberkulosis adalah

$$\hat{y} = 126,1944 + 0,2748x_{5i} - 2,9314(x_{5i} - 86,0321)_+^1 + 6,0157(x_{5i} - 95,1707)_+^1$$

$$\hat{y} = \begin{cases} 126,1944 + 0,2748x_{5i}, & x_{5i} < 86,0321 \\ 378,3889 - 2,6566x_{5i}, & 86,0321 \leq x_{5i} < 95,1707 \\ -194,1295 + 3,3591x_{5i}, & x_{5i} \geq 95,1707 \end{cases}$$

Apabila kepemilikan jaminan kesehatan kurang dari 86,0321 persen dan jika naik 1 persen, maka kesembuhan tuberkulosis meningkat 0,2748 persen. Apabila kepemilikan jaminan kesehatan

86,0321 sampai 95,1707 persen dan jika naik 1 persen, maka kesembuhan tuberkulosis turun 2,6555 persen. Apabila kepemilikan jaminan kesehatan 95,1707 persen atau lebih dan jika naik 1 persen, maka kesembuhan tuberkulosis meningkat 3,3591 persen.

6. Apabila variabel selain X_6 konstan, pengaruh penduduk merokok di atas 15 tahun (X_6) terhadap kesembuhan tuberkulosis adalah

$$\hat{y} = 126,1944 - 1,5559x_{6i} + 12,6414(x_{6i} - 26,2727)_+^1 - 69,0342(x_{6i} - 26,9986)_+^1 + 97,8949(x_{6i} - 27,7243)_+^1$$

$$\hat{y} = \begin{cases} 126,1944 - 1,5559x_{6i}, & x_{6i} < 26,2727 \\ -205,9293 + 11,0855x_{6i}, & 26,2727 \leq x_{6i} < 26,9986 \\ 1657,8974 - 579487x_{6i}, & 26,9986 \leq x_{6i} < 27,7243 \\ -1056,1706 + 39,9462x_{6i}, & x_{6i} \geq 27,7243 \end{cases}$$

Apabila penduduk merokok kurang dari 26,2727 persen dan naik 1 persen, maka kesembuhan tuberkulosis turun 1,5559 persen. Apabila penduduk merokok 26,2727 sampai 26,9986 persen dan naik 1 persen, maka kesembuhan tuberkulosis meningkat 11,0855 persen. Apabila penduduk merokok 27,7243 persen atau lebih dan naik 1 persen, maka kesembuhan tuberkulosis meningkat 39,9462 persen.

SIMPULAN

Model yang baik untuk angka kesembuhan tuberkulosis di Provinsi Jawa Timur yaitu menggunakan model regresi nonparametrik spline truncated linier dengan kombinasi titik *knot* ((2,3,1,3,2,3) yang didapat dari nilai GCV minimum sebesar 30,7126. Variabel yang berpengaruh terhadap angka kesembuhan tuberkulosis adalah TPM memenuhi syarat kesehatan, rumah tangga memiliki sanitasi layak, pengobatan lengkap tuberkulosis, tempat umum memenuhi syarat kesehatan, kepemilikan jaminan kesehatan, dan penduduk merokok di atas 15 tahun. Model ini menghasilkan R_{Adj}^2 sebesar 82,46 persen dan RKS sebesar 15,4028.

REFERENSI

- Afa, I. B., Suparti, dan Rahmawati, R. (2018). Perbandingan Metode regresi Linier Multivariabel dan Regresi Spline Multivariabel dalam Pemodelan Indeks Harga Saham Gabungan. *Media Statistika*, 11(2), 147-158. Diakses dari https://ejournal.undip.ac.id/index.php/media_statistika/article/view/19991
- Andriyanti, A., dan Rohaeti, E. (2019). Analisis Faktor-Faktor yang Memengaruhi Ketahanan Hidup Pasien Tuberkulosis dengan Model Regresi Cox (Studi Kasus: Rumah Sakit Paru Bogor). *Kubik Jurnal Publikasi Ilmiah Matematika*, 4(1), 140-148. Diakses dari <https://journal.uinsgd.ac.id/index.php/kubik/article/view/5674>
- Anggreni, N. P. R., Suciptowati, N. L. P., dan 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), 211-218. Diakses dari <https://ojs.unud.ac.id/index.php/mtk/article/download/41894/25478>
- Darma, I. G. P. S., Ratna, M., dan Budiantara, I. N. (2019). Pemodelan Angka Kasus Tuberkulosis di Surabaya menggunakan Pendekatan Regresi Nonparametrik Spline Truncated. *Jurnal Sains*

- dan Seni ITS, 8(2), 2337-3530. Diakses dari http://ejurnal.its.ac.id/index.php/sains_seni/article/view/45404
- Dinas Kesehatan Provinsi Jawa Tengah. (2018). *Profil Kesehatan Provinsi Jawa Tengah Tahun 2018*. Semarang: Kementerian Kesehatan.
- Eubank, R. L. (1999). *Nonparametric Regression and Spline Smoothing Second Edition*. New York: Marcel Dekker, Inc.
- Fatmawati, Budiantara, I. N., and Lestari, B. (2019). Comparison of Smoothing and Truncated Spline Estimators in Estimating Blood Pressure Models. *International Journal of Innovation, Creativity and Change*, 5(3). Diakses dari https://ijicc.net/images/Vol_5_Iss_3/Part_2_2020/5320_Fatmawati_2019_E_R.pdf
- Hardle, W. 1990. *Applied Nonparametric Regression*. New York: Cambridge University Press.
- Khaulasari, H. (2020). Modelling Mixed Geographically Weighted Poisson Regression for Tuberculosis Disease in Surabaya. *Journal of Physics: Conference Series*, 1490(1), 012004. Diakses dari <https://iopscience.iop.org/article/10.1088/1742-6596/1490/1/012004/meta>
- Murtantiningsih, dan Wahyono, B. (2010). Faktor-Faktor yang Berhubungan dengan Kesembuhan Penderita TB Paru. *Jurnal Kesehatan Masyarakat*, 6(1), 44-50. Diakses dari <https://media.neliti.com/media/publications/25300-ID-faktor-faktor-yang-berhubungan-dengan-kesembuhan-penderita-tuberkulosis-paru.pdf>
- Nussbaum, M. (1985). Spline Smoothing in Regression Models and Asymptotic Efficiency L_2 . *Annals of Statistics*, 13(2), 984 – 997. Diakses dari <https://www.jstor.org/stable/2241120>
- Setyowati, D.W., Rumiati, A.T., and Budiantara, I.N. (2020). Pemodelan Contraceptive Prevalence Rate (CPR) di Provinsi Sulawesi Selatan Menggunakan Regresi Nonparametrik Spline Truncated. *Jurnal Sains dan Seni ITS*, 9(1), D72-D78. Diakses dari http://ejurnal.its.ac.id/index.php/sains_seni/article/view/51045
- Sugiantari, A. P., dan Budiantara, I. N. (2013). Analisis Faktor-faktor yang Mempengaruhi Angka Harapan Hidup di Jawa Timur Menggunakan Regresi Semiparametrik Spline. *Jurnal Sains dan Seni ITS*, 2(1), D37-D41. Diakses dari http://ejurnal.its.ac.id/index.php/sains_seni/article/view/3132
- Suyono. 2018. *Analisis Regresi untuk Penelitian untuk Penelitian*. Sleman: Deepublish.
- Widarjono, A. (2007). *Teori dan Aplikasi untuk Ekonomi dan Bisnis*. Yogyakarta; Ekonisia Fakultas Ekonomi Universitas Islam Indonesia
- Yayasan KNCV Indonesia. (2020). Global TB Report 2020: *Bagaimana Capaian Penanganan Tuberculosis di Tengah Pandemi Covid-19*. <https://yki4tbc.org/news-default/411-global-tb-report-2020-bagaimana-capaian-penanganan-tuberculosis-di-tengah-pandemi-covid-19.html>. Diakses tanggal 11 Juli 2021.