

Pemodelan Faktor-Faktor yang Mempengaruhi Angka Prevalensi Kusta di Jawa Timur dengan Menggunakan Regresi Nonparametrik *Spline Truncated*

Ratri Galuh Pramesti, Madu Ratna, dan I Nyoman Budiantara
Departemen Statistika, Fakultas Matematika, Komputasi, dan Sains Data,
Institut Teknologi Sepuluh Nopember (ITS)
e-mail: i_nyoman_b@statistika.its.ac.id

Abstrak—Penyakit kusta atau lepra merupakan sebuah infeksi menular yang disebabkan oleh bakteri *Mycobacterium leprae*. Menurut laporan World Health Organization (WHO) pada tahun 2018, di tahun 2017 Indonesia menempati posisi ketiga setelah India dan Brazil sebagai negara dengan jumlah penderita kusta terbanyak di dunia. Pada tahun 2017, Jawa Timur menempati posisi pertama di Indonesia dengan jumlah penderita kusta terbanyak di Indonesia dengan jumlah 4.183 jiwa. Batas angka prevalensi kusta yang ingin dicapai Indonesia adalah 1. Sedangkan di Provinsi Jawa Timur, masih terdapat 10 Kabupaten dari 38 Kabupaten atau Kota yang memiliki angka prevalensi di atas 1. Metode analisis yang digunakan untuk mengetahui faktor-faktor yang mempengaruhi angka prevalensi kusta yaitu Regresi Nonparametrik *Spline Truncated* karena pola hubungan antara angka prevalensi kusta dengan masing-masing variabel prediktor yang didapatkan tidak membentuk suatu pola tertentu. Berdasarkan model yang diperoleh, hasilnya adalah semua variabel prediktor berpengaruh signifikan terhadap angka prevalensi kusta, yaitu persentase rumah tangga berPHBS, persentase sarana air bersih, persentase jamban sehat, persentase penduduk miskin, dan persentase puskesmas per 100.000 penduduk dengan nilai koefisien determinasi sebesar 95,34%.

Kata Kunci—Angka Prevalensi Kusta, Leprosy, Regresi Nonparametrik, *Spline Truncated*, Titik Knot.

I. PENDAHULUAN

KUSTA merupakan salah satu penyakit tertua di dunia yang hingga saat ini masih menjangkit pada jutaan orang di seluruh dunia. Penyakit kusta atau lepra adalah penyakit yang menyerang lapisan kulit, sistem saraf perifer, selaput lendir pada saluran pernapasan atas, serta mata (kecuali otak) yang disebabkan oleh bakteri *Mycobacterium leprae*. Bakteri tersebut keluar dari pernapasan, kulit yang terluka dan bisa hidup di luar tubuh manusia sehingga perlu daya tahan tubuh yang baik. Faktor sosial ekonomi memegang peranan penting, makin rendah sosial ekonomi makin subur penyakit kusta, sebaliknya sosial ekonomi tinggi membantu penyembuhan. Sehubungan dengan iklim, kusta tersebar di daerah tropis dan sub tropis yang panas dan lembab, terutama di Asia, Afrika dan Amerika Latin.

Menurut *World Health Organization* (WHO) pada tahun 2018 mengenai statistik kusta tahun 2017, Indonesia menempati posisi ketiga di dunia dengan jumlah penderita penyakit kusta yang tinggi sebanyak 15.910 kasus. India berada di posisi pertama dengan 126.124 kasus dan Brazil di posisi kedua dengan 26.875 kasus [1]. Sedangkan

menurut Dirjen Pencegahan dan Pengendalian Penyakit Kementerian Kesehatan RI, angka prevalensi penderita kusta di Indonesia pada tahun 2015 sebanyak 0,78 per 10.000 penduduk, sehingga jumlah penderita yang terdaftar sekitar 20.160 kasus [2]. Di Indonesia, wilayah yang memiliki jumlah penderita kusta terbanyak adalah Jawa Timur, yakni 4.183 jiwa. Dari 4.183 penderita kusta di Jawa Timur, sebanyak 1.464 jiwa menimpa masyarakat Madura. Jawa Timur memiliki lebih dari separuh kota dan kabupaten yang memiliki angka prevalensi di atas 1.

Pada penelitian ini bertujuan untuk mengetahui hubungan antara angka prevalensi kusta pada setiap kabupaten/kota di Jawa Timur pada tahun 2017 dengan faktor-faktor yang diduga mempengaruhinya diantaranya adalah persentase rumah tangga ber-PHBS, persentase sarana air bersih, persentase jamban sehat, persentase penduduk miskin, dan persentase puskesmas per 100.000 penduduk. Menggunakan data Jawa Timur karena Provinsi Jawa Timur adalah salah satu dari sembilan provinsi yang memiliki angka prevalensi terbesar di Indonesia berdasarkan Kemenkes RI, 2016 [2]. Salah satu metode yang berfungsi untuk mengetahui model hubungan antara variabel respon dan variabel prediktor adalah metode regresi. Metode yang paling fleksibel digunakan untuk mengetahui hubungan antara variabel respon dan prediktor jika tidak diketahui pola kurva regresinya adalah regresi nonparametrik. Metode regresi yang digunakan adalah regresi nonparametrik *Spline Truncated*, karena mampu mengestimasi data yang tidak memiliki pola tertentu dan memiliki kecenderungan dalam mencari estimasi data dari pola yang tidak mengikuti pola tertentu. Pada penelitian ini, pola yang didapatkan dari hubungan antara data angka prevalensi kusta dengan masing-masing variabel prediktornya tidak membentuk pola tertentu, sehingga metode yang akan digunakan sesuai. Hasil penelitian ini diharapkan dapat memodelkan faktor-faktor yang berpengaruh terhadap angka prevalensi kusta yang terjadi di Jawa Timur, dan mampu memberikan informasi sebagai pertimbangan dalam pengambilan kebijakan oleh pemerintah provinsi Jawa Timur.

II. TINJAUAN PUSTAKA

A. Statistika Deskriptif

Statistika deskriptif merupakan bagian statistika yang membahas tentang metode pengumpulan dan penyajian data yang mudah dipahami, menarik, dan memberikan informasi

yang berguna [3]. Statistika deskriptif memungkinkan seseorang untuk dapat menyampaikan data dengan visual sehingga dapat lebih mudah di pahami. Bentuk-bentuk visual tersebut dapat berupa diagram batang, diagram lingkaran, histogram, *scatter plot*, dan lain sebagainya.

Mean yaitu merupakan penjumlahan data dibagi dengan banyaknya data.

Rumus:
$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \tag{1}$$

Keterangan :

n : banyak observasi

x_i : banyaknya frekuensi

Varians dari sejumlah observasi adalah rata-rata kuadrat deviasi data dari rata-ratanya.

Rumus:
$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \tag{2}$$

B. Analisis Regresi

Analisis regresi merupakan suatu metode statistika yang menjelaskan tentang pola hubungan antara dua variabel atau lebih [4]. Dalam analisis regresi, variabel yang mempengaruhi adalah variabel prediktor (x) dan yang dipengaruhi adalah variabel respon (y).

C. Regresi Nonparametrik Spline Truncated

Pendekatan dengan menggunakan regresi nonparametrik digunakan jika bentuk kurva dari data adalah tidak memenuhi asumsi-asumsi klasik seperti pada regresi parametrik dimana bentuk kurva regresi diketahui atau informasi tentang pola data sebelumnya tidak lengkap [5].

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1, 2, \dots, n \tag{3}$$

Kelebihan dari regresi nonparametrik jika dibandingkan dengan regresi parametrik adalah dalam hal fleksibilitas yang lebih baik [5]. Salah satu pendekatan yang dapat dilakukan untuk melakukan estimasi terhadap fungsi $f(x_i)$ dalam regresi nonparametrik adalah dengan menggunakan *spline* [6]. Secara umum $f(x_i)$ merupakan fungsi *spline* berorde r dengan titik knot k_1, k_2, \dots, k_m dan dapat ditampilkan dalam persamaan berikut.

$$f(x_i) = \sum_{p=0}^r \beta_p x_i^p + \sum_{q=1}^m \beta_{r+m}(x_i - k_m)_+^r \tag{4}$$

dengan,

$$(x_i - k_m)_+^r = \begin{cases} (x_i - k_m)^r, & x_i \geq k_m \\ 0, & x_i < k_m \end{cases} \tag{5}$$

β merupakan parameter-parameter model dengan r merupakan orde *Spline*.

D. Estimasi Parameter

Salah satu metode yang dapat digunakan untuk mengestimasi parameter model regresi nonparametrik *spline* adalah *Ordinary Least Square* (OLS). Metode OLS mengestimasi parameter model regresi dengan meminimumkan jumlah kuadrat residual. Berikut merupakan penyajian matriks dari model regresi nonparametrik *Spline* linier dengan p knot dan univariabel prediktor.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{6}$$

E. Pemilihan Titik Knot Optimum

Titik knot merupakan titik perpaduan bersama dimana terdapat perubahan pola perilaku kurva atau fungsi pada sub interval yang berbeda [7]. Salah satu metode yang dapat digunakan untuk menentukan titik knot yang optimal adalah dengan menggunakan metode GCV (*Generalized Cross Validation*). Kelebihan dari metode ini adalah GCV memiliki sifat optimal asimtotik, tidak memuat varians populasi (σ^2) yang tidak diketahui, dan *invariance*

terhadap transformasi dimana kelebihan tersebut tidak dimiliki oleh metode lainnya [6]. Titik knot optimal yang diperoleh dengan metode GCV dapat didefinisikan sebagai berikut.

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

dengan A merupakan matriks, $A(k_1, k_2, \dots, k_r) = X(X'X)^{-1}X'$ dan rumus MSE diperoleh dengan rumus (7) sebagai berikut.

$$MSE(k_1, k_2, \dots, k_r) = n^{-1} \sum_{i=1}^n (y_i - \hat{f}(x_i))^2 \tag{8}$$

Model *spline* dengan nilai GCV paling minimum merupakan model *spline* yang terbaik.

F. Ukuran Kesesuaian Model

Dalam analisis regresi, salah satu cara untuk mengetahui kesesuaian model adalah dengan menggunakan koefisien determinasi. Yaitu proporsi variansi total dari nilai variabel y yang dijelaskan oleh variabel x [4]. Nilai kesesuaian model ditandai dengan R^2 dan ditunjukkan dalam persamaan berikut.

$$R^2 = \frac{SSR}{SST} = \frac{\hat{\boldsymbol{\beta}}' \mathbf{X}' \mathbf{y} - n \bar{y}^2}{\mathbf{y}' \mathbf{y} - n \bar{y}^2} \times 100\% \tag{9}$$

Nilai R^2 menunjukkan seberapa besar model dapat menjelaskan variabel respon.

G. Uji Parameter

Ada dua tahapan uji yaitu secara serentak (simultan) dan individu (parsial).

1. Uji Serentak

Hipotesis yang digunakan adalah sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_{r+m} = 0$$

$$H_1: \text{minimal ada satu } \beta_p \neq 0, p = 1, 2, \dots, r, r + 1, \dots, r + m$$

dimana r merupakan jumlah orde polinomial dan m merupakan jumlah titik knot. Statistik uji yang digunakan adalah sebagai berikut.

$$F_{hitung} = \frac{MS_{regresi}}{MS_{error}} \tag{10}$$

Daerah penolakan yang digunakan adalah dengan membandingkan nilai F_{hitung} dengan $F_{\alpha, (r+m), n-(r+m)-1}$ keputusan tolak H_0 adalah jika $F_{hitung} > F_{tabel}$ atau jika $p\text{-value} < \alpha$. Artinya minimal terdapat satu parameter pada model regresi *spline* yang signifikan [4].

Tabel 1. ANOVA

Sumber Variasi	Derajat Bebas (df)	Jumlah Kuadrat (SS)	Rataan Kuadrat (MS)	F_{hitung}
Regresi	$r + m$	$\beta^1 X' y - n \bar{y}^2$	$\frac{SS_{regresi}}{df_{regresi}}$	$\frac{MS_{regresi}}{MS_{residual}}$
Residual	$n - (r + m) - 1$	$y' y - \beta^1 X' y$	$\frac{SS_{residual}}{df_{residual}}$	
Total	$n - 1$	$y' y - n \bar{y}^2$	-	

2. Uji Parsial

Hipotesis yang digunakan adalah sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_{r+m} = 0$$

$$H_1: \beta_p \neq 0, p = 1, 2, \dots, r, r + 1, r + 2, \dots, r + m$$

$$\text{statistik uji : } t_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}, j = 1, 2, \dots, r + m \tag{11}$$

Nilai $SE(\hat{\beta}_j)$ merupakan *standard error* dari $\sqrt{\text{var } \hat{\beta}_j}$. Keputusan bahwa H_0 akan ditolak jika nilai dari $t_{hitung} > t_{(\frac{\alpha}{2}, n-(r+m)-1)}$ atau $t_{hitung} < -t_{(\frac{\alpha}{2}, n-(r+m)-1)}$ atau jika $p -$

value < α. Artinya bahwa parameter berpengaruh secara signifikan terhadap model [4].

H. Uji Asumsi Residual IIDN

Uji asumsi residual yang harus dipenuhi adalah identik, independen, dan berdistribusi normal.

1. Uji Asumsi Identik

Asumsi identik harus terpenuhi, karena bertujuan untuk mengetahui homogenitas variansi residual. Jika asumsi ini tidak terpenuhi, berarti terdapat heteroskedastisitas yang dapat merugikan efisiensi estimator [8]. Biasanya, uji asumsi identik dilakukan dengan menggunakan visualisasi data dan uji *glejser*. Hipotesis yang digunakan adalah sebagai berikut.

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1: \text{minimal terdapat satu } \sigma_j \neq 0, j = 1, 2, \dots, n$$

Statistik uji yang digunakan adalah sebagai berikut.

$$F_{hitung} = \frac{\frac{\sum_{i=1}^n (|\hat{\epsilon}_i| - |\bar{\epsilon}|)^2}{(r+m)}}{\frac{\sum_{i=1}^n (|\epsilon_i| - |\bar{\epsilon}|)^2}{(n-(r+m)-1)}} \quad (12)$$

Dimana *n* adalah banyak observasi dan *(r+m)* adalah banyak parameter model *glejser*. Keputusan tolak *H₀* adalah jika *F_{hitung}* > *F_{α,((r+m),n-(r+m)-1)}* atau jika *p-value* < α. Kesimpulannya adalah tidak adanya indikasi heteroskedastisitas.

2. Uji Asumsi Independen

Pengujian asumsi independen digunakan untuk mengetahui apakah terdapat autokorelasi atau tidak. Metode yang digunakan untuk menguji asumsi ini adalah *Run test* [9]. Metode ini termasuk kedalam statistik nonparametrik yang dapat digunakan untuk mengetahui apakah terdapat residual yang berkorelasi tinggi atau tidak. Jika tidak terdapat korelasi, dapat diindikasikan bahwa residual bersifat acak. Berikut merupakan hipotesis yang digunakan.

$$H_0 : \rho = 0 \text{ (tidak terjadi autokorelasi)}$$

$$H_1 : \rho \neq 0 \text{ (terjadi autokorelasi)}$$

dengan statistik uji yang digunakan adalah sebagai berikut.

$$Z = \frac{R - \bar{R}}{s_R} \quad (13)$$

dan nilai dari *R̄* dan *s_R* adalah sebagai berikut.

$$\bar{R} = \frac{2n_1n_2}{n_1+n_2} + 1 \quad (14)$$

$$s_R^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)} \quad (15)$$

r : Banyaknya *run*

s_R : standar deviasi dari *run*

n₁ : Banyaknya anggota kelompok pada kategori 1

n₂ : Banyaknya anggota kelompok pada kategori 2

Keputusan tolak *H₀* apabila *Z* > *Z_{α/2}* atau *Z* < -*Z_{α/2}* atau jika *p-value* < α, yang berarti bahwa asumsi residual independen tidak terpenuhi.

3. Uji Asumsi Distribusi Normal

Uji asumsi berikutnya yang harus terpenuhi adalah uji asumsi residual berdistribusi normal. Tujuannya adalah untuk mengetahui apakah data tersebut telah berdistribusi normal atau tidak. Terdapat dua cara yang paling sering digunakan untuk melakukan uji normalitas. Yaitu uji secara visual dengan melihat normal probability plot residual dan lainnya adalah uji Kolmogorov-Smirnov. Hipotesis yang digunakan adalah sebagai berikut.

$$H_0 : F_n(\epsilon) = F_0(\epsilon) \text{ , residual berdistribusi normal}$$

$$H_1 : F_n(\epsilon) \neq F_0(\epsilon) \text{ , residual tidak berdistribusi normal}$$

Statistik uji yang digunakan adalah sebagai berikut.

$$D = \text{Sup}_{\epsilon} |F_n(\epsilon) - F_0(\epsilon)|$$

$$(16)$$

dengan *F_n(ε)* adalah fungsi distribusi kumulatif dan *F₀(ε)* adalah fungsi peluang kumulatif yang diamati dari suatu sampel acak. Pada uji *Kolmogorov-Smirnov*, daerah penolakannya adalah tolak *H₀* apabila *D* > *D_(1-α)* dimana didapatkan dari nilai tabel atau jika *p-value* < α. Artinya adalah residual tidak berdistribusi normal. Apabila asumsi berdistribusi normal tidak terpenuhi, maka dapat diatasi dengan menggunakan metode transformasi.

I. Kusta

Istilah kusta berasal dari bahasa sansekerta, yaitu *kushta* yang berarti kumpulan permasalahan kulit secara umum. Kusta juga dikenal dengan nama lepra atau penyakit Hansen. Penyakit ini menyerang kulit, sistem saraf perifer, selaput lendir pada saluran pernapasan atas, serta mata. Kusta disebabkan oleh bakteri *Mycrobacterium leprae*. Bakteri ini memerlukan waktu 6 bulan hingga 40 tahun untuk berkembang dalam tubuh. Tanda dan gejala kusta bisa saja muncul 1 hingga 20 tahun setelah bakteri tersebut tumbuh dan menginfeksi tubuh penderita. Kuman kusta dapat hidup di dalam tubuh manusia selama 24-48 jam dan akan berenang dalam 7-9 hari tergantung suhu dan cuaca [10].

Penyakit kusta dipengaruhi oleh beberapa aspek sosial ekonomi, aspek lingkungan, aspek demografi, dan aspek perilaku [11]. Selain itu, aspek fasilitas kesehatan juga mempengaruhi terjadinya kusta [12]. Pada aspek sosial ekonomi meliputi penduduk miskin dan rumah tangga dengan alas lantai tanah. Aspek lingkungan meliputi rumah sehat, dan jamban sehat. Aspek demografi meliputi kepadatan penduduk. Aspek perilaku meliputi rumah tangga ber-PHBS, dan sarana air bersih. Sedangkan aspek fasilitas kesehatan meliputi banyak *public health*, dan jumlah tenaga medis.

Angka prevalensi kusta merupakan jumlah kasus penyakit kusta per 10.000 penduduk. Dengan keadaan orang tersebut terinfeksi pada waktu tertentu, baik itu kusta kering maupun kusta basah. Rumus yang digunakan untuk menghitung angka prevalensi kusta adalah sebagai berikut.

$$Y = \frac{\text{Jumlah Kasus Kusta Basah} + \text{Jumlah Kasus Kusta Kering}}{\text{Jumlah penduduk}} \times 10.000 \text{ penduduk}$$

III. METODE PENELITIAN

A. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang didapatkan dari Profil Kesehatan Indonesia tahun 2017 oleh Dinas Kesehatan dan Provinsi Jawa Timur dalam Angka tahun 2017 oleh Badan Pusat Statistik. Unit penelitian yang diamati adalah kabupaten/kota di Jawa Timur yang terdiri dari 29 kabupaten dan 9 kota.

B. Variabel Penelitian

Variabel yang digunakan dalam penelitian ini ditampilkan pada Tabel 2.

Tabel 2.
Variabel Penelitian

Variabel	Nama Variabel
<i>y</i>	Angka Prevalensi Kusta
<i>x₁</i>	Persentase Rumah Tangga ber-PHBS
<i>x₂</i>	Persentase Sarana Air Bersih
<i>x₃</i>	Persentase Jamban Sehat
<i>x₄</i>	Persentase Penduduk Miskin
<i>x₅</i>	Persentase Puskesmas per 100.000 Penduduk

C. Langkah Penelitian

Adapun langkah-langkah analisis yang dilakukan dalam penelitian ini adalah sebagai berikut.

1. Membuat statistika deskriptif pada masing-masing variabel untuk mengetahui karakteristik masing-masing kabupaten/kota di Jawa Timur
2. Membuat *scatterplot* antara angka prevalensi kusta di Jawa Timur dengan masing-masing variabel prediktor
3. Memodelkan angka prevalensi kusta di Jawa Timur dengan menggunakan *spline* linier dengan satu, dua, tiga, dan kombinasi titik knot
4. Memilih titik knot yang optimal berdasarkan nilai GCV minimum
5. Memodelkan angka prevalensi kusta di Jawa Timur dengan masing-masing variabel prediktor menggunakan *spline* dengan titik knot yang optimal
6. Melakukan uji signifikansi parameter dan pengujian asumsi residual model *spline* terbaik
7. Menghitung koefisien determinasi
8. Membuat interpretasi dan kesimpulan.

IV. HASIL DAN PEMBAHASAN

A. Karakteristik Data Angka Prevalensi Kusta dan Faktor-Faktor yang Mempengaruhinya

Berikut merupakan karakteristik data angka prevalensi kusta dan faktor-faktor yang diduga mempengaruhinya

Tabel 3.
Karakteristik Angka Prevalensi Kusta dan Faktor-Faktor yang Diduga Berpengaruh

Variabel	Mean	Varians	Minimum	Maksimum
Y	0,866	1,003	0,000	3,968
X ₁	54,01	230,52	24,20	100,00
X ₂	78,70	367,85	17,00	100,00
X ₃	80,08	297,81	25,20	100,00
X ₄	11,625	22,284	4,172	23,562
X ₅	2,70	0,493	1,191	4,337

Berdasarkan Tabel 3, variabel Y merupakan variabel respon yaitu angka prevalensi kusta di Jawa Timur pada tahun 2017. Rata-rata dari angka prevalensi kusta di Jawa Timur pada tahun 2017 adalah sebesar 0,866 dengan nilai varians yang besar. Artinya bahwa setiap 10.000 penduduk di Jawa Timur, sebanyak 1 orang menderita penyakit kusta. Nilai minimum dan maksimum didapatkan angka 0 pada Kota Batu, dan 3,96 pada Kabupaten Sumenep.

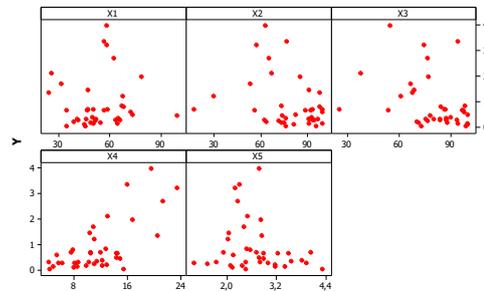
B. Pemodelan Angka Prevalensi Kusta di Jawa Timur dengan Menggunakan Pendekatan Regresi Nonparametrik Spline Truncated

Sebelum dilakukan analisis lebih lanjut, dilakukan analisis deskriptif dari variabel-variabel yang diduga mempengaruhi angka prevalensi kusta di Jawa Timur untuk memastikan bahwa komponen yang digunakan merupakan nonparametrik.

i. Pola Hubungan Angka Prevalensi Kusta di Jawa Timur dengan Faktor-faktor yang Diduga Berpengaruh

Analisis secara deskriptif mengenai angka prevalensi kusta dan masing-masing faktor yang diduga mempengaruhinya serta penentuan komponen parametrik dan nonparametrik untuk masing-masing variabel prediktor dengan menggunakan *scatter plot*.

Berdasarkan Gambar 1, diketahui bahwa pola yang terbentuk antara hubungan variabel respon dan prediktornya tidak membentuk pola tertentu. Sehingga variabel-variabel tersebut termasuk dalam komponen nonparametrik.



Gambar 1. Pola Hubungan Angka Prevalensi Kusta dengan Variabel Prediktor.

ii. Pemilihan Titik Knot Optimum

Langkah selanjutnya adalah melakukan estimasi model regresi nonparametrik *spline truncated* hingga 3 knot dan kombinasi knot. Banyaknya knot yang digunakan pada penelitian ini adalah 1 titik knot, 2 titik knot, 3 titik knot, dan kombinasinya.

1. Regresi Nonparametrik Spline dengan 1 Titik Knot

Estimasi model regresi nonparametrik *spline* dengan satu titik knot pada angka prevalensi kusta di Jawa Timur dimana semua komponennya adalah komponen nonparametrik secara umum adalah sebagai berikut.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 (x_{i1} - K_{11})_+^1 + \beta_3 x_{i2} + \beta_4 (x_{i2} - K_{21})_+^1 + \beta_5 x_{i3} + \beta_6 (x_{i3} - K_{31})_+^1 + \beta_7 x_{i4} + \beta_8 (x_{i4} - K_{41})_+^1 + \beta_9 x_{i5} + \beta_{10} (x_{i5} - K_{51})_+^1 + \epsilon_i$$

Nilai GCV paling optimum untuk model regresi nonparametrik *spline* dengan menggunakan satu titik knot adalah sebesar 0,596. Nilai ini didapatkan dari satu titik knot optimal pada setiap variabel. Besarnya titik knot untuk variabel persentase rumah tangga ber-PHBS (X₁) adalah 56,685; variabel persentase sarana air bersih (X₂) adalah 52,571; variabel persentase jamban sehat (X₃) adalah 57,342; variabel persentase penduduk miskin (X₄) adalah 12,48; serta variabel persentase puskesmas per 10.000 penduduk (X₅) adalah 2,539.

2. Regresi Nonparametrik Spline dengan 2 Titik Knot

Selanjutnya dilakukan pemodelan dengan dua titik knot bertujuan untuk mendapatkan perbandingan model yang lebih baik. Berikut adalah model regresi *spline* dengan dua titik knot.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 (x_{i1} - K_{11})_+^1 + \beta_3 (x_{i1} - K_{12})_+^1 + \beta_4 x_{i2} + \beta_5 (x_{i2} - K_{21})_+^1 + \beta_6 (x_{i2} - K_{22})_+^1 + \beta_7 x_{i3} + \beta_8 (x_{i3} - K_{31})_+^1 + \beta_9 (x_{i3} - K_{32})_+^1 + \beta_{10} x_{i4} + \beta_{11} (x_{i4} - K_{41})_+^1 + \beta_{12} (x_{i4} - K_{42})_+^1 + \beta_{13} x_{i5} + \beta_{14} (x_{i5} - K_{51})_+^1 + \beta_{15} (x_{i5} - K_{52})_+^1 + \epsilon_i$$

Nilai GCV minimum untuk model regresi nonparametrik *spline* dengan dua titik knot adalah sebesar 0,384. Titik knot optimal untuk variabel persentase rumah tangga ber-PHBS (X₁) adalah 69,061 dan 70,608; variabel persentase sarana air bersih (X₂) adalah 66,122 dan 67,816; variabel persentase jamban sehat (X₃) adalah 69,587 dan 71,118; variabel persentase penduduk miskin (X₄) adalah 15,645 dan 16,041; serta variabel persentase puskesmas per 10.000 penduduk (X₅) adalah 3,052 dan 3,116.

3. Regresi Nonparametrik Spline dengan 3 Titik Knot

Estimasi model regresi nonparametrik *spline* dengan menggunakan tiga titik knot untuk kasus angka prevalensi kusta di Jawa Timur adalah sebagai berikut.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 (x_{i1} - K_{11})_+^1 + \beta_3 (x_{i1} - K_{12})_+^1 + \beta_4 (x_{i1} - K_{13})_+^1 + \beta_5 x_{i2} + \beta_6 (x_{i2} - K_{21})_+^1 + \beta_7 (x_{i2} - K_{22})_+^1 + \beta_8 (x_{i2} - K_{23})_+^1 + \beta_9 x_{i3} + \beta_{10} (x_{i3} - K_{31})_+^1 + \beta_{11} (x_{i3} - K_{32})_+^1 + \beta_{12} (x_{i3} - K_{33})_+^1 + \beta_{13} x_{i4} + \beta_{14} (x_{i4} - K_{41})_+^1 + \beta_{15} (x_{i4} - K_{42})_+^1 + \beta_{16} (x_{i4} - K_{43})_+^1 + \beta_{17} x_{i5} + \beta_{18} (x_{i5} - K_{51})_+^1 + \beta_{19} (x_{i5} - K_{52})_+^1 + \beta_{20} (x_{i5} - K_{53})_+^1 + \varepsilon_i$$

Nilai GCV terkecil dengan menggunakan tiga titik knot yang diperoleh adalah sebesar 0,2237. Titik knot optimal untuk variabel persentase rumah tangga berperilaku hidup bersih dan sehat (X₁) adalah 39,669; 69,061; dan 70,608; kemudian untuk variabel persentase sarana air bersih (X₂) adalah sebesar 33,938; 66,122; dan 67,816; variabel persentase jamban sehat (X₃) adalah sebesar 40,506; 69,587; dan 71,118; variabel persentase penduduk miskin (X₄) adalah sebesar 8,127; 15,645; dan 16,041; serta variabel persentase puskesmas per 100.000 penduduk (X₅) adalah sebesar 1,832; 3,052; dan 3,116.

4. Regresi Nonparametrik *Spline* dengan Kombinasi knot

Nilai GCV terkecil dengan menggunakan tiga titik knot yang diperoleh adalah sebesar 0,201. Besarnya titik knot untuk variabel persentase rumah tangga ber-PHBS (X₁) adalah k₁ = 39,669; k₂ = 69,061; dan k₃ = 70,608. Nilai titik knot untuk variabel persentase sarana air bersih (X₂) adalah k₁ = 33,938; k₂ = 66,122; dan k₃ = 67,816. Nilai titik knot untuk persentase jamban sehat (X₃) adalah k₁ = 40,506; k₂ = 69,587; dan k₃ = 71,118. Nilai titik knot untuk variabel persentase penduduk miskin (X₄) adalah k₁ = 15,647 dan k₂ = 16,043. Sedangkan Nilai titik knot untuk variabel persentase puskesmas per 100.000 penduduk (X₅) adalah k₁ = 1,832; k₂ = 3,052; dan k₃ = 3,116.

iii. Pemilihan Model Terbaik

Pemilihan model terbaik dilakukan dengan cara membandingkan nilai GCV yang paling optimum untuk masing-masing variabel prediktor dari masing-masing titik knot.

Tabel 4.
Nilai GCV Terkecil dari Pemodelan Tiap Knot

Jumlah Knot	GCV
1 Titik Knot	0,5968
2 Titik Knot	0,3840
3 Titik Knot	0,2237
Kombinasi Knot (3,3,3,2,3)	0,2015

Berdasarkan Tabel 4 diketahui bahwa pemodelan yang menghasilkan nilai GCV paing minimum adalah pemodelan regresi nonparametrik *spline* dengan menggunakan kombinasi knot (3,3,3,2,3).

iv. Penaksiran Parameter Model Regresi Nonparametrik *Spline Truncated*

Berdasarkan kriteria pemilihan model terbaik, diketahui bahwa nilai GCV yang paling minimum dihasilkan oleh model regresi nonparametrik *spline* dengan kombinasi knot. Kombinasi titik knot tersebut yaitu untuk variabel x₁, x₂, x₃, dan x₅ adalah tiga titik knot, sedangkan x₄ adalah dua titik knot.

$$y_i = -3,001 + 0,10323x_{i1} - 0,08449(x_{i1} - 39,669)_+^1 - 0,5379(x_{i1} - 69,061)_+^1 + 0,5437(x_{i1} - 70,608)_+^1 + 0,9352x_{i2} - 1,0328(x_{i2} - 33,938)_+^1 - 0,2063(x_{i2} - 66,122)_+^1 + 0,3234(x_{i2} - 67,816)_+^1 - 0,7588x_{i3} + 0,7546(x_{i3} - 40,506)_+^1 - 0,6441(x_{i3} - 69,587)_+^1 + 0,6676(x_{i3} - 71,118)_+^1 + 0,0539x_{i4} + 5,9317(x_{i4} - 15,647)_+^1 - 6,1265(x_{i4} - 16,043)_+^1 + 1,1299x_{i5} - 1,7047(x_{i5} - 1,832)_+^1 - 5,4126(x_{i5} - 3,052)_+^1 + 6,5467(x_{i5} - 3,116)_+^1$$

Nilai koefisien determinasi yang didapatkan dari model regresi nonparametrik *spline truncated* adalah sebesar 95,34%.

v. Pengujian Signifikansi Parameter

Setelah mendapatkan model terbaik, akan dilanjutkan dengan melakukan pengujian terhadap parameter model baik secara serentak maupun parsial.

1. Pengujian Serentak

Hipotesis yang digunakan adalah sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_{r+m} = 0$$

$$H_1: \text{minimal ada satu } \beta_p \neq 0, p = 1, 2, \dots, r, r + 1, \dots, r + m$$

Tabel 5.
Hasil Uji Serentak

Sumber Variasi	df	SS	MS	F	P-Value
Regresi	19	35,393	1,8628	19,41	2,46x10 ⁻¹¹
Error	18	1,7274	0,0959		
Total	37	37,1208			

Berdasarkan Tabel 5 diketahui bahwa *p-value* yang didapatkan adalah sebesar 2,463x10⁻¹¹, *P-value* < α=0,05. Nilai statistik uji *F* yang didapatkan adalah sebesar 19,411. Nilai tabel *F* untuk derajat bebas v₁=19 dan v₂=18 dengan α=0,05 adalah 2,203. Maka nilai statistik *F* > *F*_{0,05;(19,18)}. Sehingga dapat disimpulkan Tolak *H*₀ yang artinya minimal terdapat satu variabel prediktor yang memberikan pengaruh signifikan terhadap model.

2. Pengujian Parsial

Hipotesis yang digunakan adalah sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_{r+m} = 0$$

$$H_1: \beta_p \neq 0, p = 1, 2, \dots, r, r + 1, r + 2, \dots, r + m$$

Tabel 6.
Hasil Uji Parsial

Variabel	Parameter	Estimator	t	Keputusan
Konstan	β ₀	-3,001	-1,477	0,15695
	β ₁	0,10323	3,374	0,003374
	β ₂	-0,08449	-2,401	0,027368
	β ₃	-0,5379	-2,955	0,008456
	β ₄	0,5437	2,874	0,010080
x ₁	β ₅	0,9352	4,640	0,000203
	β ₆	-1,0328	-4,430	0,000323
	β ₇	-0,2063	-0,770	0,451030
	β ₈	0,3234	1,275	0,218414
	β ₉	-0,7588	-4,055	0,000742
x ₂	β ₁₀	0,7546	3,791	0,001336
	β ₁₁	-0,6441	-3,01	0,007510
	β ₁₂	0,6676	3,227	0,004671
	β ₁₃	0,0539	1,666	0,112933
x ₃	β ₁₄	5,9317	7,855	3,17x10 ⁻¹¹
	β ₁₅	-6,1265	-7,839	3,26x10 ⁻¹¹
	β ₁₆	1,1299	1,926	0,069966
	β ₁₇	-1,7047	-2,168	0,043794
x ₄	β ₁₈	-5,4126	-1,399	0,178610
	β ₁₉	6,5467	1,694	0,107402

Berdasarkan hasil pengujian parameter secara parsial dengan membandingkan statistik uji *t* dengan nilai *t*_{tabel}, yaitu *t*_{0,025;18} sebesar 2,101. Apabila nilai statistik uji |*t*| > *t*_{0,025;18}, maka keputusannya adalah Tolak *H*₀. Terdapat 7 parameter yang tidak signifikan dan 13 parameter lainnya telah signifikan.

vi. Pengujian Asumsi Residual

1. Uji Asumsi Identik

Hasil statistik uji *Glejser* adalah sebagai berikut.

Tabel 7.

Hasil Pengujian Statistik Uji Glejser					
Sumber	Df	SS	MS	Fhit	P-value
Regresi	19	0,38915	0,0204	1,2836	0,300095
Error	18	0,2872	0,0159		
Total	37	0,67635			

Pada Tabel 7 menunjukkan hasil pengujian identik dengan menggunakan uji Glejser didapatkan nilai F_{hitung} sebesar 1,2836. Residual dikatakan identik apabila nilai $F_{hitung} > F_{tab}$ atau $F_{\alpha,((r+m),n-(r+m)-1)}$. Nilai $F_{0,05;(19,18)}$ sebesar 2,203 dan $p-value > \alpha$. Karena nilai $F_{hitung} > F_{tab}$, dan $p-value > \alpha$, maka keputusannya adalah Gagal Tolak H_0 . Artinya adalah tidak terjadi heteroskedastisitas dalam model, sehingga asumsi residual identik telah terpenuhi.

2. Uji Asumsi Independen

Hipotesis yang digunakan dalam pengujian asumsi residual independen adalah sebagai berikut.

H_0 : Residual bersifat independen

H_1 : Residual bersifat dependen

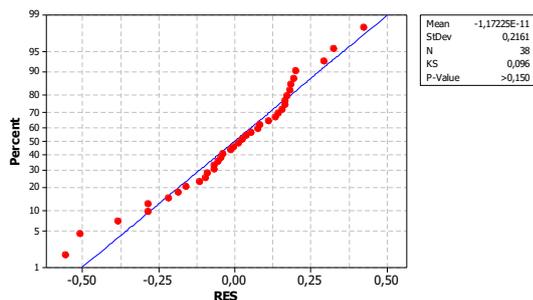
Tabel 8.

Hasil Pengujian Run Test	
Test Value	P-Value
0,677	0,622

Karena Z hitung $< Z_{\frac{\alpha}{2}}$, yaitu $0,677 < 1,96$ dan $p-value$ yang dihasilkan sebesar $0,622$ dan jika dibandingkan dengan taraf signifikansi α ($\alpha=0,05$), $p-value > \alpha$, sehingga keputusannya adalah Gagal Tolak H_0 . Kesimpulannya adalah telah terbukti bahwa residual independen atau tidak terdapat autokorelasi.

3. Uji Asumsi Distribusi Normal

Nilai α yang digunakan dalam pengujian asumsi residual independen adalah $0,05$. Berikut adalah hasil statistik uji Kolmogorov-Smirnov yang ditunjukkan pada Gambar 2.



Gambar 2. Plot Normalitas Residual.

Pada Gambar 2 dapat dilihat bahwa titik-titik residual berada di sekitar garis linear, sehingga dapat disimpulkan bahwa residual telah berdistribusi normal. Berdasarkan nilai D diketahui bahwa $0,096 < D_{(1-\alpha)}$ yaitu $0,194$. Apabila dilihat dari nilai $p-value$, didapatkan bahwa $p-value > 0,150$, dimana nilai tersebut lebih besar dari nilai α dengan taraf sigifikansi $\alpha(\alpha=0,05)$, sehingga keputusannya adalah Gagal Tolak H_0 , yang artinya adalah residual hasil pemodelan telah berdistribusi normal.

vii. Interpretasi Model

Model dan interpretasi dari model regresi nonparametrik *spline truncated* terbaik pada angka prevalensi kusta di Jawa Timur dengan kelima variabel prediktor yang signifikan adalah sebagai berikut.

1. Hubungan antara rumah tangga berperilaku hidup bersih dan sehat (PHBS) (x_1) dengan angka prevalensi kusta (y)

dan diasumsikan variabel prediktor selain x_1 dianggap konstan, model regresi dapat ditulis sebagai berikut.

$$y = -3,001 + 0,10323x_{i1} - 0,08449(x_{i1} - 39,669)_+^1 - 0,5379(x_{i1} - 69,061)_+^1 + 0,5437(x_{i1} - 70,608)_+^1$$

$$\hat{y} = \begin{cases} -3,001 + 0,10323x_1 & ; x_1 < 39,669 \\ 0,3506 + 0,01874x_1 & ; 39,669 \leq x_1 < 69,061 \\ 37,498 - 0,51916x_1 & ; 69,061 \leq x_1 < 70,608 \\ -0,891 + 0,02545x_1 & ; x_1 \geq 70,608 \end{cases}$$

Pada kondisi (x_1) yang kurang dari 39,669 jika terjadi penambahan sebesar 1 satuan maka angka prevalensi kusta akan mengalami peningkatan sebesar 0,1032 per 10.000 penduduk. Kabupaten/kota yang termasuk dalam kategori (x_1) yang kurang dari 39,669 adalah Kabupaten Trenggalek, Kota Batu, Kabupaten Bondowoso, Kabupaten Lumajang, Kabupaten Situbondo, Kabupaten Probolinggo. Apabila (x_1) di antara 39,669 dan 69,061 terjadi penambahan sebesar 1 satuan maka angka prevalensi kusta akan mengalami kenaikan sebesar 0,01874 per 10.000 penduduk. Kabupaten/kota yang termasuk dalam kategori (x_1) di antara 39,669 dan 69,061 adalah Kota Malang, Kabupaten Tulungagung, Kota Blitar, Kabupaten Pacitan, Kabupaten Banyuwangi, Kabupaten Magetan, Kabupaten Nganjuk, Kabupaten Pasuruan, Kabupaten Blitar, Kabupaten Madiun, Kota Pasuruan, Kota Kediri, Kabupaten Malang, Kabupaten Mojokerto, Kabupaten Kediri, Kabupaten Bojonegoro, Kabupaten Pamekasan, Kabupaten Sampang, Kabupaten Sumenep, Kabupaten Jombang, Kabupaten Bangkalan, Kota Mojokerto, Kota Madiun, Kabupaten Ponorogo, Kabupaten Sidoarjo, Kabupaten Gresik, Kabupaten Jember, dan Kota Probolinggo. Apabila (x_1) yang lebih besar dari 70,608 terjadi penambahan sebesar 1 satuan maka angka prevalensi kusta akan mengalami kenaikan sebesar 0,0254 per 10.000 penduduk. Kabupaten/kota yang termasuk dalam kategori persentase rumah tangga ber-PHBS yang lebih besar dari 70,608 adalah Kota Surabaya, Kabupaten Lamongan, Kabupaten Tuban, dan Kabupaten Ngawi. Pertambahan angka prevalensi kusta pada setiap kenaikan variabel rumah tangga yang menerapkan perilaku hidup bersih disebabkan oleh faktor-faktor lain yang mempengaruhi angka prevalensi kusta selain rumah tangga yang menerapkan hidup bersih dan sehat. Jika dilihat dan dibandingkan dari statistika deskriptif kedua variabel, diketahui bahwa angka prevalensi kusta akan tetap sama meskipun rumah tangga yang menerapkan hidup bersih dan sehat ditingkatkan.

2. Hubungan antara persentase sarana air bersih (x_2) dengan angka prevalensi kusta (y) dan diasumsikan variabel prediktor selain x_2 dianggap konstan, model regresi dapat ditulis sebagai berikut.

$$y = -3,001 + 0,9352x_{i2} - 1,0328(x_{i2} - 33,938)_+^1 - 0,2063(x_{i2} - 66,122)_+^1 + 0,3234(x_{i2} - 67,816)_+^1$$

$$\hat{y} = \begin{cases} -3,001 + 0,9352x_2 & ; x_2 < 33,938 \\ 31,969 - 0,0976x_2 & ; 33,938 \leq x_2 < 66,122 \\ 45,609 - 0,3039x_2 & ; 66,122 \leq x_2 < 67,816 \\ -23,678 - 0,0195x_2 & ; x_2 \geq 67,816 \end{cases}$$

Pada (x_2) yang kurang dari 33,938 jika terjadi penambahan sebesar 1 satuan maka angka prevalensi kusta akan mengalami kenaikan sebesar 0,9352 per 10.000 penduduk. Kabupaten atau kota yang termasuk dalam kategori (x_2) yang kurang dari 33,938 adalah Kabupaten Jember dan Kabupaten Nganjuk. Apabila persentase sarana air bersih berada di antara 33,938 dan 66,122 terjadi

penambahan sebesar 1 satuan maka angka prevalensi kusta akan mengalami penurunan sebesar 0,0976 per 10.000 penduduk. Kabupaten atau kota yang termasuk dalam kategori ini adalah Kabupaten Lumajang, Kabupaten Bondowoso, Kabupaten Sampang, Kabupaten Tulungagung, Kabupaten Sumenep, Kabupaten Trenggalek, dan Kabupaten Bangkalan.

Pada kondisi (x_2) berada di antara 66,122 dan 67,816 jika terjadi penambahan sebesar 1 satuan maka angka prevalensi kusta akan mengalami penurunan sebesar 0,3039 per 10.000 penduduk. Kabupaten kota yang termasuk dalam kategori ini adalah Kabupaten Situnondo. Kemudian apabila kondisi (x_2) yang lebih besar dari 67,816 terjadi penambahan sebesar 1 satuan maka angka prevalensi kusta akan mengalami penurunan sebesar 0,0195 per 10.000 penduduk. Kabupaten atau kota yang termasuk dalam kategori tersebut adalah Kabupaten Malang, Kabupaten Gresik, Kabupaten Kediri, Kabupaten Ngawi, Kabupaten Blitar, Kabupaten Ponorogo, Kabupaten Pacitan, Kabupaten Pamekasan, Kota Blitar, Kabupaten Lamongan, Kabupaten Tuban, Kabupaten Bojonegoro, Kabupaten Pasuruan, Kabupaten Banyuwangi, Kabupaten Mojokerto, Kota Kediri, Kabupaten Magetan, Kabupaten Madiun, Kota Mojokerto, Kabupaten Jombang, Kabupaten Sidoarjo, Kota Batu, Kota Malang, Kabupaten Probolinggo, Kota Probolinggo, Kota Surabaya, Kota Pasuruan, dan Kota Madiun.

3. Hubungan antara persentase jamban sehat (x_3) dengan angka prevalensi kusta (y) dan diasumsikan variabel prediktor selain x_3 dianggap konstan, model regresi dapat ditulis sebagai berikut.

$$y_i = -3,001 - 0,7588x_{i3} + 0,7546(x_{i3} - 40,506)_+^1 - 0,6441(x_{i3} - 69,587)_+^1 + 0,6676(x_{i3} - 71,118)_+^1$$

$$\hat{y} = \begin{cases} -3,001 - 0,7588x_3 & ; x_3 < 40,506 \\ 33,566 - 0,0042x_3 & ; 40,506 \leq x_3 < 69,587 \\ 78,387 - 0,6483x_3 & ; 69,587 \leq x_3 < 71,118 \\ 30,909 + 0,0193x_3 & ; x_3 \geq 71,118 \end{cases}$$

Pada kondisi (x_3) yang kurang dari 40,506 jika terjadi penambahan sebesar 1 satuan maka angka prevalensi kusta akan mengalami penurunan sebesar 0,7588 per 10.000 penduduk. Kabupaten atau kota yang termasuk dalam kategori ini adalah Kabupaten Nganjuk dan Kabupaten Situbondo. Jika pada kondisi (x_3) berada di antara 40,506 dan 69,587 terjadi penambahan sebesar 1 satuan maka angka prevalensi kusta akan mengalami penurunan sebesar 0,0042 per 10.000 penduduk. Kabupaten atau kota yang termasuk dalam kategori tersebut adalah Kabupaten Bondowoso, Kabupaten Sumenep, Kabupaten Jember, Kabupaten Lumajang, Kabupaten Probolinggo, dan Kabupaten Pasuruan.

Apabila (x_3) berada di antara 69,587 dan 71,118 terjadi penambahan sebesar 1 satuan maka angka prevalensi kusta akan mengalami penurunan sebesar 0,6483 per 10.000 penduduk. Kabupaten atau kota yang termasuk dalam kategori ini adalah Kabupaten Trenggalek. Pada kondisi (x_3) yang lebih besar dari 71,118 jika terjadi penambahan sebesar 1 satuan maka angka prevalensi kusta akan mengalami penambahan sebesar 0,0193 per 10.000 penduduk. Kabupaten atau kota yang termasuk dalam kategori tersebut adalah Kabupaten Kediri, Kabupaten Pacitan, Kabupaten Sampang, Kota Malang, Kabupaten Tuban, Kabupaten Ngawi, Kabupaten Mojokerto, Kabupaten Tulungagung, Kabupaten Banyuwangi, Kota

Probolinggo, Kota Pasuruan, Kabupaten Jombang, Kabupaten Ponorogo, Kabupaten Sidoarjo, Kabupaten Blitar, Kabupaten Madiun, Kota Mojokerto, Kota Kediri, Kabupaten Pamekasan, Kabupaten Bojonegoro, Kabupaten Magetan, Kota Surabaya, Kabupaten Gresik, Kota Batu, Kabupaten Lamongan, Kota Madiun, dan Kota Blitar. Pertambahan angka prevalensi kusta pada setiap kenaikan variabel persentase jamban sehat pada kabupaten atau kota yang berada di interval lebih dari 71,118 disebabkan oleh faktor-faktor lain yang mempengaruhi angka prevalensi kusta selain persentase jamban sehat. Jika dilihat dan dibandingkan dari statistika deskriptif kedua variabel, diketahui bahwa angka prevalensi kusta akan tetap sama meskipun rumah tangga yang menggunakan jamban sehat ditingkatkan.

4. Hubungan antara persentase penduduk miskin (x_4) dengan angka prevalensi kusta (y) dan diasumsikan variabel prediktor selain x_4 dianggap konstan, model regresi dapat ditulis sebagai berikut.

$$y_i = -3,001 + 0,0539x_{i4} + 5,9317(x_{i4} - 15,647)_+^1 - 6,1265(x_{i4} - 16,043)_+^1$$

$$\hat{y} = \begin{cases} -3,001 + 0,0539x_4 & ; x_4 < 15,647 \\ -95,803 + 5,9856x_4 & ; 15,647 \leq x_4 < 16,043 \\ 2,4841 - 0,1409x_4 & ; x_4 \geq 16,043 \end{cases}$$

Pada kondisi (x_4) yang kurang dari 15,647 jika terjadi penambahan sebesar 1 satuan maka angka prevalensi kusta akan mengalami kenaikan sebesar 0,0539 per 10.000 penduduk. Kabupaten atau kota yang termasuk dalam kategori tersebut adalah Kota Malang, Kota Batu, Kota Madiun, Kota Surabaya, Kota Mojokerto, Kabupaten Sidoarjo, Kota Pasuruan, Kota Probolinggo, Kota Blitar, Kabupaten Tulungagung, Kota Kediri, Kabupaten Banyuwangi, Kabupaten Blitar, Kabupaten Mojokerto, Kabupaten Pasuruan, Kabupaten Jombang, Kabupaten Magetan, Kabupaten Lumajang, Kabupaten Jember, Kabupaten Malang, Kabupaten Ponorogo, Kabupaten Nganjuk, Kabupaten Kediri, Kabupaten Madiun, Kabupaten Gresik, Kabupaten Trenggalek, Kabupaten Situbondo, Kabupaten Bojonegoro, Kabupaten Lamongan, Kabupaten Bondowoso, Kabupaten Ngawi, dan Kabupaten Pacitan.

Sedangkan jika pada kondisi (x_4) di antara 15,647 dan 16,043 terjadi penambahan sebesar 1 satuan maka angka prevalensi kusta akan mengalami kenaikan sebesar 5,9856 per 10.000 penduduk. Kabupaten/kota yang termasuk dalam kategori tersebut adalah Kabupaten Pamekasan. Jika pada kondisi (x_4) lebih dari 16,043 jika terjadi penambahan sebesar 1 satuan maka angka prevalensi kusta akan mengalami penurunan sebesar 0,1409 per 10.000 penduduk. Kabupaten atau kota yang termasuk dalam kategori tersebut adalah Kabupaten Tuban, Kabupaten Sumenep, Kabupaten Probolinggo, Kabupaten Bangkalan, dan Kabupaten Sampang.

5. Hubungan antara persentase puskesmas per 100.000 penduduk (x_5) dengan angka prevalensi kusta (y) dan diasumsikan variabel prediktor selain x_5 dianggap konstan, model regresi dapat ditulis sebagai berikut.

$$y_i = -3,001 + 1,1299x_{i5} - 1,7047(x_{i5} - 1,832)_+^1 - 5,4126(x_{i5} - 3,052)_+^1 + 6,5467(x_{i5} - 3,116)_+^1$$

$$\hat{y} = \begin{cases} -3,001 + 1,1299x_5 & ; x_5 < 1,832 \\ 0,12201 - 0,57475x_5 & ; 1,832 \leq x_5 < 3,052 \\ 16,6394 - 5,98735x_5 & ; 3,052 \leq x_5 < 3,116 \\ -3,760 + 0,5593x_5 & ; x_5 \geq 3,116 \end{cases}$$

Pada kondisi (x_5) yang kurang dari 1,832 jika terjadi penambahan sebesar 1 satuan maka angka prevalensi kusta akan mengalami kenaikan sebesar 1,129 per 10.000 penduduk. Kabupaten atau kota yang termasuk dalam kategori tersebut adalah Kabupaten Sidoarjo, Kabupaten Malang, dan Kota Malang. Selain itu, jika pada kondisi (x_5) di antara 1,832 dan 3,052 terjadi penambahan sebesar 1 satuan maka angka prevalensi kusta akan mengalami penurunan sebesar 0,574 per 10.000 penduduk. Kabupaten atau kota yang termasuk dalam kategori ini adalah Kabupaten Nganjuk, Kabupaten Jember, Kabupaten Pasuruan, Kabupaten Blitar, Kota Blitar, Kota Surabaya, Kabupaten Sampang, Kabupaten Bangkalan, Kabupaten Pamekasan, Kabupaten Kediri, Kabupaten Lumajang, Kota Batu, Kabupaten Mojokerto, Kabupaten Gresik, Kabupaten Situbondo, Kota Probolinggo, Kabupaten Jombang, Kabupaten Sumenep, Kabupaten Lamongan, Kabupaten Banyuwangi, Kabupaten Tuban, Kabupaten Probolinggo, Kabupaten Ngawi, dan Kabupaten Tulungagung.

Kemudian pada kondisi (x_5) yang lebih besar dari 3,116 terjadi penambahan sebesar 1 satuan maka angka prevalensi kusta akan mengalami peningkatan sebesar 0,5593 per 10.000 penduduk. Kabupaten atau kota yang termasuk dalam kategori tersebut adalah Kota Kediri, Kabupaten Trenggalek, Kabupaten Bondowoso, Kota Madiun, Kabupaten Magetan, Kabupaten Ponorogo, Kabupaten Madiun, Kota Mojokerto, Kota Pasuruan, dan Kabupaten Pacitan. Pertambahan angka prevalensi kusta pada setiap kenaikan variabel persentase jamban sehat pada kabupaten atau kota yang berada di interval lebih dari 71,118 disebabkan oleh faktor-faktor lain yang mempengaruhi angka prevalensi kusta selain persentase puskesmas per 100.000 penduduk. Jika dilihat dan dibandingkan dari statistika deskriptif kedua variabel, diketahui bahwa angka prevalensi kusta akan tetap sama meskipun persentase puskesmas per 100.000 penduduk ditingkatkan.

V. KESIMPULAN DAN SARAN

A. Kesimpulan

Terdapat 10 Kabupaten yang masih berada di atas batas capaian angka prevalensi kusta sebesar 1. Kabupaten tersebut adalah Kabupaten Sumenep, Kabupaten Pamekasan, Kabupaten Sampang, Kabupaten Bangkalan, Kabupaten Situbondo, Kabupaten Tuban, Kabupaten Lumajang, Kabupaten Pasuruan, Kabupaten Probolinggo, dan Kabupaten Jember. Sedangkan 28 Kabupaten atau Kota

lainnya di Jawa Timur telah melebihi capaian batas angka prevalensi kusta.

Model regresi nonparametrik *spline truncated* terbaik yang dihasilkan adalah dengan menggunakan kombinasi knot (3,3,3,2,3) dengan nilai GCV paling minimum adalah 0,201. Koefisien determinasi atau R^2 yang menunjukkan kebaikan model adalah sebesar 95,34%. Berikut adalah model terbaik yang diperoleh.

$$y_i = -3,001 + 0,10323x_{i1} - 0,08449(x_{i1} - 39,669)_+^1 - 0,5379(x_{i1} - 69,061)_+^1 + \\ 0,5437(x_{i1} - 70,608)_+^1 + 0,9352x_{i2} - 1,0328(x_{i2} - 33,938)_+^1 - 0,2063(x_{i2} - 66,122)_+^1 + \\ 0,3234(x_{i2} - 67,816)_+^1 - 0,7588x_{i3} + 0,7546(x_{i3} - 40,506)_+^1 - 0,6441(x_{i3} - 69,587)_+^1 + \\ 0,6676(x_{i3} - 71,118)_+^1 + 0,0539x_{i4} + 5,9317(x_{i4} - 15,647)_+^1 - 6,1265(x_{i4} - 16,043)_+^1 + \\ 1,1299x_{i5} - 1,7047(x_{i5} - 1,832)_+^1 - 5,4126(x_{i5} - 3,052)_+^1 + 6,5467(x_{i5} - 3,116)_+^1$$

B. Saran

Bagi penelitian selanjutnya, sebaiknya menambah jumlah variabel yang diduga berpengaruh terhadap angka prevalensi kusta di Jawa Timur, sehingga diharapkan akan mendapatkan model yang lebih baik. Sebaiknya pemerintah lebih memperhatikan faktor-faktor apa saja yang berpengaruh untuk mengurangi angka prevalensi kusta di Jawa Timur.

DAFTAR PUSTAKA

- [1] WHO, "Leprosy," 2017. [Online]. Available: <https://www.who.int/lep/en/>.
- [2] Kemenkes, "Profil Kesehatan Indonesia 2016," Jakarta, 2016.
- [3] R. Walpole, *Pengantar Metode Statistika*, 3rd ed. Jakarta: Bumi Aksara, 1995.
- [4] D. N.R and S. H., *Analisis Regresi Terapan*. Jakarta: Gramedia Pustaka Utama, 1992.
- [5] R. Eubank, *Nonparametric Regression and Spline Smoothing*, 2nd ed. Texas: Department of Statistics Southern Methodist Dallas University, 1999.
- [6] Wahba, *Spline Models for Observational Data*. Pennsylvania, 1990.
- [7] I. Budiantara, "Regresi Nonparametrik Dalam Statistika," in *Seminar Nasional Matematika*, 2006.
- [8] R. L. Eubank, *Nonparametric Regression and Spline Smoothing, Second Edition*. New York: Taylor & Francis, 1999.
- [9] W. W. Daniel, *Statistik Non Parametrik Terapan. Diterjemahkan oleh Alex Tri Kantjono W*. Jakarta: PT.Gramedia, 1989.
- [10] Kemenkes RI, "Profil Kesehatan Indonesia Tahun 2017," Jakarta, 2018.
- [11] D. Jariwala, "Socio-Demographic and Environmental Correlates of Leprosy: A Hospital Based Cases Control Study," *Natl. J. Community Med.*, vol. 4, no. 3, 2013.
- [12] L. Kerr-Pontes and et al, "Socioeconomic, Environmental, and Behavioural Risk Factors for Leprosy in North-east Brazil: an Ecological Study," *Int. J. Epidemiol.*, vol. 33, no. 2, pp. 262-9, 2004.