

## Regresi Nonparametrik *Spline Truncated*

**Agustini Tripena**

Jurusan Matematika Universitas Jenderal Soedirman  
[agustini.brsurbakti@unsoed.ac.id](mailto:agustini.brsurbakti@unsoed.ac.id)

**Yosita Lianawati**

Sekolah Tinggi Ilmu Komputer Yos Sudarso  
Email: [yosita.lianawati@stikomys.ac.id](mailto:yosita.lianawati@stikomys.ac.id)

Correspondence author: [agustini.brsurbakti@unsoed.ac.id](mailto:agustini.brsurbakti@unsoed.ac.id)

**Abstract.** Nonparametric regression is used to determine the relationship between the response variable and the predictor variable whose shape of the regression curve is not known. *This study examines the factors that affect the number of tuberculosis cases in Central Java in 2021. The number of tuberculosis cases in Central Java reached 117 cases per 115,000 residents with a mean of 110,35 and a variance of 1360.74. Tuberculosis case rates and their affecting factors are modeled using a spline truncated nonparametric regression method. The relationship between tuberculosis case rates in Central Java and the factors that are thought to affect it does not form a specific plot. . Based on the research results, the best model of the minimum Mean Square Error (MSE) value is obtained at the knot point combination (1, 3, 2, 3, 3), with a MSE value of 0.57. The model, it is known that all factors have a significant effect on tuberculosis cases in Central Java, with a coefficient of determination of 93.41%. That is, the model can explain the diversity of tuberculosis cases in Central Java by 93.41%.*

**Keywords:** tuberculosis, nonparametric spline truncated, MSE, knot point

**Abstrak.** Regresi nonparametrik digunakan untuk mengetahui hubungan antara variabel respon dengan variabel prediktor yang tidak diketahui bentuk kurva regresinya. Pada penelitian ini dikaji faktor-faktor yang mempengaruhi angka kasus tuberkulosis di Jawa Tengah tahun 2021. Angka kasus tuberkulosis di Jawa Tengah mencapai 120 kasus per 115.000 penduduk dengan nilai rata-rata sebesar 110,35 dan nilai variansi 1360,74. Faktor yang mempengaruhinya dimodelkan menggunakan metode regresi nonparametrik *spline truncated*. Hubungan antara angka kasus tuberkulosis di Jawa Tengah dengan faktor yang diduga mempengaruhinya tidak membentuk pola tertentu. Angka kasus tuberkulosis beserta faktor yang mempengaruhinya dimodelkan menggunakan metode regresi nonparametrik *spline truncated* karena hubungan antara angka kasus tuberkulosis di Jawa Tengah dengan faktor yang diduga mempengaruhinya tidak membentuk pola. Berdasarkan hasil penelitian, model terbaik dari nilai *Mean Square Error* minimum diperoleh pada kombinasi titik *knot* (1, 3, 2, 3, 3) dengan nilai MSE sebesar 0,57. Model tersebut diketahui bahwa seluruh faktor berpengaruh signifikan terhadap kasus tuberkulosis di Jawa Tengah, dengan nilai koefisien determinasi sebesar 93,41%. Artinya, model mampu menjelaskan keragaman kasus tuberkulosis di Jawa Tengah sebesar 93,41%.

**Kata kunci:** tuberkulosis, nonparametrik *spline truncated*, MSE, titik *knot*

### LATAR BELAKANG

Analisis regresi digunakan untuk mengidentifikasi sejumlah variabel prediktor dalam memprediksi suatu variabel respon. Jika, data yang digunakan tidak mengikuti pola data tertentu, maka model yang digunakan adalah regresi nonparametrik. Regresi nonparametrik digunakan untuk memodelkan data apabila bentuk kurva dari fungsi regresinya tidak diketahui.

Received Juni 30, 2023; Revised Juli 15, 2023; Accepted Agustus 21, 2023

\* Agustini Tripena\*, [agustini.brsurbakti@unsoed.ac.id](mailto:agustini.brsurbakti@unsoed.ac.id)

Jadi, data diharapkan mencari sendiri bentuk estimasinya tanpa dipengaruhi oleh faktor subjektivitas (Rismal, 2016). Model regresi nonparametrik yang sering digunakan untuk melakukan estimasi terhadap kurva regresi salah satunya adalah regresi *spline truncated*. *Spline truncated* dapat mengatasi pola data yang menunjukkan signifikansi naik atau turun dengan bantuan titik-titik *knot* dan sangat baik digunakan untuk mengatasi pola data yang perilakunya berubah-ubah pada interval tertentu (Rahmadina, dkk., 2021).

Tuberkulosis adalah suatu penyakit menular yang disebabkan oleh bakteri *Mycobacterium Tuberculosis*. Bakteri *Mycobacterium Tuberculosis* umumnya menyerang paru-paru, tetapi juga dapat menyerang organ dalam tubuh lainnya, seperti tulang, kelenjar getah bening, dan jantung. Penyakit ini dapat menular secara langsung melalui udara ketika penderita tuberkulosis batuk, bersin, atau meludah (Kementerian Kesehatan, 2018: 2). Kasus tuberkulosis di Indonesia menempati urutan kedua penyebab kematian tertinggi dari penyakit infeksi lainnya. Kasus tuberkulosis tertinggi terjadi di Provinsi Jawa Barat, diikuti dengan Jawa Timur, Jawa Tengah, dan Banten (Kementerian Kesehatan, 2021). Dengan tingginya kasus tuberkulosis di Jawa Tengah, diperlukan upaya preventif oleh pemerintah dengan cara mengendalikan faktor-faktor yang mempengaruhi tuberkulosis. Salah satu cara untuk mengetahui faktor yang berpengaruh dalam angka kasus tuberkulosis adalah dengan pemodelan menggunakan metode analisis regresi nonparametrik.

## KAJIAN TEORITIS

### Regresi Nonparametrik *Spline Truncated*

Regresi nonparametrik merupakan model regresi yang sangat fleksibel dalam memodelkan pola data. Regresi nonparametrik digunakan untuk mengetahui hubungan antara variabel respon dengan variabel prediktor yang tidak diketahui bentuk kurva regresinya. Secara umum, menurut Eubank (1999: 10) model regresi nonparametrik dapat ditulis sebagai berikut

$$y_i = f(x_i) + e_i, \quad i = 1, 2, 3, \dots, n$$

Diberikan data berpasangan  $(x_{1i}, x_{2i}, \dots, x_{si}, y_i)$  dan hubungan antar variabel  $(x_{1i}, x_{2i}, \dots, x_{si})$  dengan  $y_i$  diasumsikan mengikuti model regresi nonparametrik,

$$y_i = f(x_{1i}, x_{2i}, \dots, x_{si}) + e_i, \quad i = 1, 2, 3, \dots, n \quad (1)$$

Apabila kurva regresi  $f$  merupakan model aditif, maka dapat dijabarkan menjadi

$$\begin{aligned}
 f(x_{1i}, x_{2i}, \dots, x_{hi}) &= f_1(x_{1i}) + f_2(x_{2i}) + \dots + f_s(x_{si}) \\
 &= \sum_{u=1}^s f_u(x_{iu}), i = 1, 2, \dots, n
 \end{aligned}$$

Menurut Eubank (1999: 281), fungsi *spline* berorde  $m$  dengan titik *knot*  $K_{1r}, K_{2r}, \dots, K_{qr}$  dengan  $r = 1, 2, \dots, h$  dapat ditulis dalam persamaan

$$f_u(x_{iu}) = \sum_{j=0}^m \beta_{ju} x_{iu}^j + \sum_{q=1}^r \beta_{(m+q)u} (x_{iu} - K_{qu})_+^m \quad (2)$$

Jadi, berdasarkan persamaan (1) model regresi nonparametrik *spline truncated* linier multivariabel adalah

$$y_i = \sum_{u=1}^s \sum_{j=0}^m \beta_{ju} x_{iu}^j + \sum_{u=1}^s \sum_{q=1}^r \beta_{(m+q)u} (x_{iu} - K_{qu})_+^m + e_i, \quad i = 1, 2, \dots, n \quad (3)$$

Pada persamaan (3), fungsi  $(x_u - K_{qu})_+^m$  merupakan fungsi *truncated* yang disajikan dalam bentuk

$$(x_{iu} - K_{qu})_+^m = \begin{cases} (x_{iu} - K_{qu})^m, & x_i \geq K_{qu} \\ 0, & x_i < K_{qu} \end{cases} \quad (4)$$

## Tuberkulosis

Tuberkulosis merupakan penyakit infeksi kronik dan menular yang disebabkan oleh bakteri *Mycobacterium Tuberculosis*. *Mycobacterium Tuberculosis* termasuk bakteri tahan asam (BTA) yang dapat bertahan hidup pada suhu 4°C sampai -70°C (Kasaluhe, 2021: 1). Tuberkulosis dibedakan menjadi dua kategori, yaitu tuberkulosis paru dan ekstra paru. Tuberkulosis paru biasanya menyerang jaringan paru, sedangkan tuberkulosis ekstra paru menyerang organ lainnya seperti organ selaput otak, selaput jantung, kelenjar getah bening, ginjal, dan lain-lain (Kementerian Kesehatan, 2021).

## Pemilihan Titik *Knot* Optimal

Titik *knot* merupakan titik perpaduan bersama ketika terdapat perubahan perilaku pola data pada suatu interval yang berlainan (Budiantara, 2006). Oleh karena itu, perlu dicari titik *knot* optimal untuk menentukan model regresi *spline* terbaik. Salah satu metode yang dapat

digunakan untuk menentukan titik *knot* optimal adalah *Mean Square Error* (MSE). Menurut Eubank (1999: 43), persamaan metode GCV dapat dituliskan sebagai

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (5)$$

### Kriteria Kebaikan Model

Salah satu ukuran yang sering digunakan untuk mengetahui kebaikan model yaitu koefisien determinasi, yang dilambangkan dengan  $R^2$ . Koefisien determinasi digunakan untuk mengetahui sampai sejauh mana kecocokan regresi yang terbentuk terhadap data pengamatan.

Rumus untuk menghitung  $R^2$  adalah

$$R^2 = \frac{\text{SSR}}{\text{SST}} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (6)$$

Semakin tinggi nilai  $R^2$  yang dihasilkan suatu model, maka semakin baik pula variabel-variabel prediktor dalam model tersebut dalam menjelaskan variabel respon (Gujarati, 2004: 217).

## METODE PENELITIAN

Penelitian ini dilakukan dengan metode studi pustaka dari buku teks dan jurnal. Data yang digunakan adalah data sekunder angka kasus tuberkulosis, persentase rumah tangga ber-PHBS, persentase akses layanan sanitasi layak, persentase tempat umum sehat, persentase Pendidikan terakhir SMA/ sederajat, dan persentase penduduk miskin di 33 kabupaten/kota di Provinsi Jawa Tengah tahun 2020. Data diperoleh dari Dinas Kesehatan Provinsi Jawa Tengah dan Badan Pusat Statistik (BPS) Provinsi Jawa Tengah. Langkah-langkah yang dilakukan dalam penelitian ini adalah:

1. Mendeskripsikan data.
2. Membuat *scatterplot* antara angka kasus tuberkulosis di Jawa Tengah dengan masing-masing variabel prediktor yang diduga mempengaruhinya untuk mengetahui pola hubungan yang terjadi.
3. Melakukan pemodelan angka kasus tuberkulosis menggunakan model regresi nonparametrik *spline truncated* dengan satu titik *knot*, dua titik *knot*, tiga titik *knot*, dan kombinasi titik *knot*.

Menurut Eubank (1999: 281), fungsi *spline* berorde  $m$  dengan titik *knot*  $K_{1r}, K_{2r}, \dots, K_{qr}$  dengan  $r = 1, 2, \dots, h$  dapat ditulis dalam persamaan

$$f_u(x_{iu}) = \sum_{j=0}^m \beta_{ju} x_{iu}^j + \sum_{q=1}^r \beta_{(m+q)u} (x_{iu} - K_{qu})_+^m$$

Jadi, berdasarkan persamaan diatas model regresi nonparametrik *spline truncated* linier multivariabel adalah

$$y_i = \sum_{u=1}^s \sum_{j=0}^m \beta_{ju} x_{iu}^j + \sum_{u=1}^s \sum_{q=1}^r \beta_{(m+q)u} (x_{iu} - K_{qu})_+^m + e_i, \quad i = 1, 2, \dots, n$$

4. Memilih titik *knot* optimal berdasarkan nilai GCV yang paling minimum. Model regresi Spline terbaik tergantung pada titik knot yang optimal. Titik knot optimal diperoleh dari nilai MSE minimum. Menurut Eubank (1999: 43), persamaan metode  $MSE = \frac{1}{n} \sum_{t=1}^n (y_i - \hat{y}_i)^2$
5. Melakukan pemodelan regresi *spline* terbaik dengan titik *knot* optimal.
6. Melakukan uji signifikansi parameter secara serentak dan parsial.

**Tabel 1. Analisis rvariansi (ANOVA) dengan uji F**

Sumber Variasi	Df	Sum of Square (SS)	Mean Square (MS)	F <sub>hitung</sub>
Regresi	$p + q$	$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	$\frac{SSR}{df_{regresi}}$	
Error	$n - (p + q) - 1$	$\sum_{i=1}^n (y_i - \hat{y}_i)^2$	$\frac{SSE}{df_{error}}$	$\frac{MSR}{MSE}$
Total	$n - 1$	$\sum_{i=1}^n (y_i - \bar{y})^2$		

$$t_{hitung} = \frac{\hat{\beta}_u}{se(\hat{\beta}_u)}$$

7. Melakukan uji asumsi residual identik, independen, dan distribusi normal (IIDN) dari model regresi *spline*.
8. Menghitung nilai koefisien determinasi  $R^2$ .

Koefisien determinasi digunakan untuk mengetahui sampai sejauh mana kecocokan regresi yang terbentuk terhadap data pengamatan. Rumus untuk menghitung  $R^2$  adalah

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Semakin tinggi nilai  $R^2$  yang dihasilkan suatu model, maka semakin baik pula variabel-variabel prediktor dalam model tersebut dalam menjelaskan variabel respon (Gujarati, 2004: 217).

9. Menginterpretasikan model.

## HASIL DAN PEMBAHASAN

### Karakteristik Data

Jawa Tengah merupakan salah satu provinsi di Indonesia yang memiliki luas wilayah sebesar 32.800,69 km<sup>2</sup> dengan jumlah penduduk sebanyak 38.686.024 jiwa. Angka kasus tuberkulosis di Jawa Tengah pada tahun 2021 mencapai 120 per 115.000 penduduk. Berikut ini disajikan diagram batang angka kasus tuberkulosis di Jawa Tengah menurut kabupaten/kota tahun 2021.

**Tabel 2. Karakteristik data kasus tuberkulosis dan faktor-faktor yang diduga mempengaruhinya**

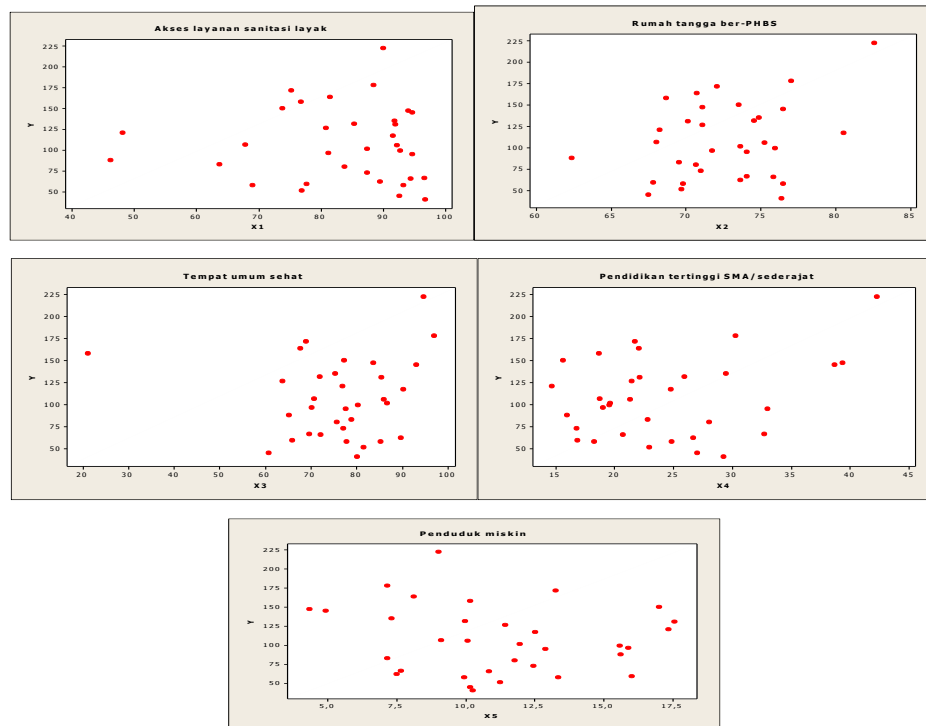
<i>Variable</i>	<i>Mean</i>	<i>Varians</i>	<i>Maximum</i>	<i>Minimum</i>
<i>Y</i>	110,35	1360,74	203,10	39,70
<i>X<sub>1</sub></i>	76,18	134,12	75,40	48,25
<i>X<sub>2</sub></i>	61,41	14,34	71,51	52,31
<i>X<sub>3</sub></i>	86,23	171,67	88,18	28,10
<i>X<sub>4</sub></i>	14,13	42,63	31,19	13,61
<i>X<sub>5</sub></i>	10,11	11,23	16,39	5,44

Berdasarkan Tabel 2, rata-rata dari angka kasus tuberkulosis di Jawa tengah terdapat 110,35 kasus tiap 115.000 penduduk, dengan variansi yang sangat besar yakni 1360,74, artinya di kota/kabupaten tertentu angka kasus tuberkulosisnya tidak merata, ada yang besar dan ada yang kecil. Angka kasus tuberkulosis tertinggi terdapat di Kota Surakarta dan yang terendah di Kab. Karanganyar.

Rata-rata dari persentase akses layanan sanitasi layak adalah sebesar 76,18, artinya sebagian besar kota/kabupaten di Jawa Tengah memiliki akses layanan sanitasi yang layak, akan tetapi persebarannya masih belum merata dikarenakan variansi yang tinggi. Begitupula dengan rumah tangga ber-PHBS memiliki rata-rata persentase yang tinggi sebesar 61,41 yang artinya rumah tangga di tiap kota/kabupaten di Jawa Tengah telah melakukan perilaku hidup bersih dan sehat. Diikuti oleh persentase tempat umum sehat di tiap kota/kecamatan di Jawa Tengah memiliki nilai rata-rata sebesar 86,23 dengan persebaran yang belum merata karena nilai variansi yang masih tinggi yaitu 171,67. Selanjutnya, persentase pendidikan tertinggi SMA/ sederajat, dengan nilai rata-rata 13,13% dan variansi 41,63%. Untuk persentase penduduk miskin, nilai rata-rata persentase penduduk miskin adalah 10,11% dengan variansi 11,23%.

### Scatterplot antara Angka Kasus Tuberkulosis dengan Variabel yang Diduga Mempengaruhinya

Identifikasi pola hubungan antar variabel prediktor terhadap variabel respon dilakukan untuk mengetahui bagaimana pola hubungan masing-masing variabel prediktor terhadap variabel respon. Dengan menggunakan *scatterplot* sebagai visualisasi data, berikut adalah hasil dari pola hubungan antar variabelnya ditunjukkan pada Gambar 1.



**Gambar 1.** Pola hubungan antar variabel prediktor terhadap variabel respon

### Pemilihan Titik *Knot* Optimum

Nilai MSE dari masing-masing pemodelan dengan satu titik *knot*, dua titik *knot*, tiga titik *knot*, dan kombinasi titik *knot*, dibandingkan untuk mendapatkan pemodelan yang optimal. Berikut tabel nilai MSE dari masing-masing pemodelan.

**Tabel 3.** Perbandingan nilai MSE minimum

Banyak Titik <i>Knot</i> Optimal	MSE
Satu titik <i>knot</i>	605,08
Dua titik <i>knot</i>	370,50
Tiga titik <i>knot</i>	153,31
<b>Kombinasi titik <i>knot</i></b>	<b>132,93</b>

Tabel 3 menunjukkan bahwa nilai MSE paling minimum diperoleh pada kombinasi titik *knot*, dengan nilai MSE sebesar 132,93. Oleh karena itu, diputuskan bahwa model terbaik yang akan dipilih adalah model regresi nonparametrik *spline* dengan menggunakan kombinasi titik *knot*.

**Pemodelan Menggunakan Titik *Knot* Optimum**

Untuk melakukan pemodelan angka kasus tuberkulosis di Provinsi Jawa Tengah, digunakan titik *knot* optimum berdasarkan nilai MSE paling minimum. Estimasi parameter untuk model *spline* terbaik diperoleh menggunakan metode *Ordinary Least Square* (OLS). Hasil estimasi parameter disajikan pada Tabel 4.

**Tabel 4. Hasil estimasi parameter**

Parameter	Estimasi	Parameter	Estimasi
$\hat{\beta}_0$	50,25	$\hat{\beta}_{32}$	-1,86
$\hat{\beta}_{11}$	-0,88	$\hat{\beta}_{33}$	4,08
$\hat{\beta}_{12}$	-17,76	$\hat{\beta}_{41}$	5,40
$\hat{\beta}_{21}$	0,21	$\hat{\beta}_{42}$	-59,63
$\hat{\beta}_{22}$	6,22	$\hat{\beta}_{43}$	63,52
$\hat{\beta}_{23}$	2,11	$\hat{\beta}_{44}$	-27,69
$\hat{\beta}_{24}$	-19,40	$\hat{\beta}_{51}$	-1,83
$\hat{\beta}_{31}$	-0,91	$\hat{\beta}_{52}$	63,49

Dari Tabel 4 diperoleh nilai estimasi parameter dari masing-masing variabel prediktor. Jadi, model regresi nonparametrik *spline* dengan kombinasi titik *knot* yaitu

$$\hat{Y} = 50,25 - 0,88X_1 - 17,76(X_1 - 91,90)_+ + 0,21X_2 + 6,22(X_2 - 69,15)_+ + 2,11(X_2 - 70,27)_+ - 19,40(X_2 - 80,40)_+ - 0,91X_3 - 1,86(X_3 - 42,80)_+ + 4,08(X_3 - 85,52)_+ + 5,40x_4 - 59,63(X_4 - 26,43)_+ + 63,52(X_4 - 27,83)_+ - 27,69(X_4 - 38,55)_+ - 1,83X_5 + 63,49(X_5 - 15,39)_+$$

**Pengujian Signifikansi Parameter Model Regresi Nonparametrik *Spline Truncated***

**Uji Serentak**

Uji serentak menggunakan tabel *Anova* dengan nilai  $\alpha$  yang digunakan sebesar 0,05.

Hipotesis yang digunakan adalah

$$H_0 : \beta_1 = \dots = \beta_{1+s} = 0$$

$$H_1 : \text{terdapat } \beta_u \neq 0 \text{ untuk suatu } u = 1, 2, \dots, s$$

Hasil analisis ragam model regresi nonparametrik *spline* disajikan dalam Tabel 5.

**Tabel 5. Hasil uji serentak menggunakan uji *F***

Sumber variasi	Df	Sum of Square (SS)	Mean Square (MS)	$F_{hitung}$	$p\text{-value}$
----------------	----	--------------------	------------------	--------------	------------------



Regresi	15	51,15	3,41	14,21	0,00000000
Error	17	4,13	0,24		
Total	32	55,28			

Dari Tabel 5 diketahui bahwa nilai  $p$ -value sebesar 0,00000000, yang lebih kecil daripada  $\alpha = 0,05$ . Jadi, dapat disimpulkan bahwa  $H_0$  ditolak, yang berarti minimal terdapat satu parameter yang berpengaruh signifikan terhadap angka kasus tuberkulosis.

### Uji Parsial

Dengan menggunakan hipotesis berikut:

$$H_0 : \beta_u = 0 \text{ untuk suatu } u = 1, 2, \dots, s$$

$$H_1 : \beta_u \neq 0 \text{ untuk suatu } u = 1, 2, \dots, s$$

Hasil uji  $t$  dengan bantuan *software R* dapat dilihat pada Tabel 6.

**Tabel 6.** Hasil uji parsial menggunakan uji  $t$

Variabel	Parameter	$t_{hit}$	$p$ -value	Keputusan
Konstanta	$\beta_0$	-4,72	0,0006	Signifikan
$X_1$	$\beta_{11}$	12,53	0,0003	Signifikan
	$\beta_{12}$	9,31	0,0006	Signifikan
$X_2$	$\beta_{21}$	-1,66	0,12	Tidak Signifikan
	$\beta_{22}$	-7,32	0,0007	Signifikan
	$\beta_{23}$	-9,33	0,0005	Signifikan
	$\beta_{24}$	-6,88	0,0006	Signifikan
$X_3$	$\beta_{31}$	-1,78	0,008	Signifikan
	$\beta_{32}$	-1,59	0,08	Tidak Signifikan
	$\beta_{33}$	-1,59	0,09	Tidak Signifikan
$X_4$	$\beta_{41}$	5,87	0,007	Signifikan
	$\beta_{42}$	-8,95	0,001	Signifikan
	$\beta_{43}$	3,49	0,004	Signifikan
	$\beta_{44}$	-1,79	0,09	Tidak Signifikan
$X_5$	$\beta_{51}$	-1,79	0,09	Tidak Signifikan
	$\beta_{52}$	-6,03	0,0005	Signifikan

Pada Tabel 6 diketahui bahwa kelima variabel prediktor mempunyai parameter yang signifikan terhadap model karena memiliki  $p$ -value lebih dari 0,05. Jadi, seluruh variabel yang digunakan berpengaruh secara signifikan terhadap angka kasus tuberkulosis di Jawa Tengah.

## 1.1 Pengujian Asumsi Residual

### 1.1.1 Asumsi Identik

Pengujian asumsi *error* identik dilakukan dengan menggunakan uji Glejser. Hasil uji Glejser menggunakan *software R* dapat dilihat sebagai berikut.

**Tabel 7. Hasil uji asumsi identik menggunakan uji Glejser**

Sumber variasi	Df	Sum of Square (SS)	Mean Square (MS)	$F_{hitung}$
Regresi	20	331,60	16,58	0,56
Error	12	344,64	28,72	
Total	32	676,04		

Dari Tabel 7 diketahui bahwa nilai  $F_{hitung} = 0,56 < 2,54 = F_{(0,05;20;12)}$  sehingga  $H_0$  ditolak. Jadi, dapat disimpulkan bahwa tidak terjadi heteroskedastisitas dalam model. Hal ini berarti bahwa asumsi *error* yang identik terpenuhi.

### Asumsi Independen

Asumsi *error* berikutnya yang harus terpenuhi adalah asumsi independen. Pengujian asumsi *error* yang independen dilakukan dengan menggunakan uji Durbin-Watson.

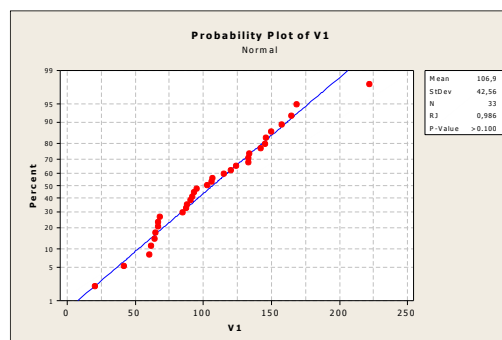
**Tabel 8. Hasil uji Durbin-Watson**

Nilai $d$	Nilai $d_L$	Nilai $d_U$	Nilai $4 - d_U$
1,71	1,11	1,91	2,18

Dari Tabel 8 diketahui bahwa nilai  $d_U < d < 4 - d_U$  sehingga  $H_0$  ditolak. Jadi, dapat dikatakan bahwa tidak terdapat autokorelasi antar residual. Hal ini membuktikan bahwa asumsi independen pada model terpenuhi.

### Asumsi Distribusi Normal

Hasil uji distribusi normal dengan uji Shapiro-Wilk menggunakan grafik PP-Plot diberikan pada Gambar 2.



**Gambar 2. Grafik plot residual**

Gambar 2 menunjukkan bahwa *error* model telah mendekati garis linier. Ini berarti bahwa *error* model regresi nonparametrik *spline truncated* telah memenuhi asumsi distribusi normal.

### Koefisien Determinasi

Nilai koefisien determinasi ( $R^2$ ) adalah sebesar 93,45%. Ini berarti bahwa variabel prediktor yang digunakan pada model regresi nonparametrik *spline truncated* mampu menjelaskan model sebesar 93,35%, dan sisanya sebesar 6,65% dijelaskan oleh variabel lain yang tidak masuk ke dalam model.

### Interpretasi Model Regresi Nonparametrik *Spline Truncated*

Interpretasi model regresi nonparametrik *spline* terbaik menggunakan kombinasi titik *knot*, yaitu

$$\begin{aligned} \hat{Y} = & 50,25 - 0,88X_1 - 17,76(X_1 - 91,90)_+ + 0,21X_2 + 6,22(X_2 - 69,15)_+ \\ & + 2,11(X_2 - 70,27)_+ - 19,40(X_2 - 80,40)_+ - 0,91X_3 \\ & - 1,86(X_3 - 42,80)_+ + 4,08(X_3 - 85,52)_+ + 6,46x_4 \\ & - 59,63(X_4 - 26,43)_+ + 63,52(X_4 - 27,83)_+ - 27,69(X_4 - 38,55)_+ \\ & - 1,83X_5 + 63,49(X_5 - 15,39)_+ \end{aligned}$$

Model tersebut memiliki 5 variabel prediktor, yaitu persentase akses layanan sanitasi layak, persentase rumah tangga ber-PHBS, persentase tempat umum sehat, persentase pendidikan tertinggi SMA/ sederajat, dan persentase penduduk miskin yaitu:

1. Pengaruh persentase akses layanan sanitasi layak ( $X_1$ ) terhadap angka kasus tuberkulosis tiap kabupaten/kota di Jawa Tengah dengan mengasumsikan variabel lainnya konstan.
2. Pengaruh persentase rumah tangga ber-PHBS ( $X_2$ ) terhadap angka kasus tuberkulosis tiap kabupaten/kota di Jawa Tengah dengan mengasumsikan variabel lainnya konstan
3. Pengaruh persentase tempat umum sehat ( $X_3$ ) terhadap angka kasus tuberkulosis tiap kabupaten/kota di Jawa Tengah dengan mengasumsikan variabel lainnya konstan
4. Pengaruh persentase pendidikan tertinggi SMA/ sederajat ( $X_4$ ) terhadap angka kasus tuberkulosis tiap kabupaten/kota di Jawa Tengah dengan mengasumsikan variabel lainnya konstan,
5. Pengaruh persentase penduduk miskin ( $X_5$ ) terhadap angka kasus tuberkulosis tiap kabupaten/kota di Jawa Tengah dengan mengasumsikan variabel lainnya konstan

## KESIMPULAN DAN SARAN

Kesimpulan yang dapat diperoleh dari penelitian tentang angka kasus tuberkulosis di Jawa Tengah tahun 2021 adalah angka kasus tuberkulosis di Jawa Tengah pada tahun 2021 mencapai 115 per 115.000 penduduk. Kabupaten/kota yang memiliki kasus tuberkulosis tertinggi adalah Kota Surakarta, sedangkan kasus tuberkulosis terendah terdapat di Kabupaten Karanganyar. Rata-rata kasus tuberkulosis di Jawa Tengah yaitu 110,35 per 115.000 penduduk. Keragaman data ditunjukkan oleh nilai variansi sebesar 1360,74. Model regresi nonparametrik *spline truncated* terbaik yang terpilih dalam angka kasus tuberkulosis di Jawa Tengah tahun 2020 adalah model dengan kombinasi titik *knot* (1, 3, 2, 3, 3).

Berdasarkan hasil pengujian signifikansi parameter model, diketahui bahwa seluruh variabel berpengaruh signifikan terhadap kasus tuberkulosis di Jawa Tengah. Nilai koefisien determinasi yang diperoleh sebesar 93,45%. Hal ini menunjukkan bahwa model mampu menjelaskan keragaman kasus tuberkulosis di Jawa Tengah sebesar 93,45%, sedangkan sisanya sebesar 6,35% dijelaskan oleh variabel lain yang tidak masuk dalam model. Untuk penelitian selanjutnya, disarankan menggunakan regresi nonparametrik *spline truncated* kuadratik agar lebih optimal terhadap model, serta menambah jumlah variabel yang diduga berpengaruh terhadap angka kasus tuberkulosis di Provinsi Jawa Tengah.

## DAFTAR REFERENSI

### Buku Teks

- Eubank, R. L. (1999). *Nonparametric Regression and Spline Smoothing Second Edition*. New York: Marcel Dekker, Inc.
- Gujarati, D. N. (2004). *Basic Econometrics Fourth Edition*. New York: The McGraw–Hill Companies.
- Hardle, W. (1994). *Applied Nonparametric Regression*. New York: Cambridge Univesity Press.
- Peck, R., Olsen, C., dan Devore, J. (2010). *Introduction to Statistics & Data Analysis Fourth Edition*. US: Cengage Learning.

### Artikel Jurnal

- Budiantara, I. N. (2006). Model *Spline* dengan *Knots* Optimal. *Jurnal Ilmu Dasar*. 7(6), 77-85.
- Rahmadina, R. L., Ratna, M., dan Budiantara, I. N. (2021). Pemodelan Faktor yang Mempengaruhi Angka Partisipasi Kasar SMA/ sederajat di Papua Menggunakan Regresi Nonparametrik *Spline Truncated*. *Jurnal Sains dan Seni ITS*, 10(1), 2337-3520.

### **Laporan Instansi/Lembaga/Organisasi/Perusahaan**

Kementerian Kesehatan. (2018). *Pusat Data dan Informasi Tuberkulosis*. Jakarta Selatan: Kementerian Kesehatan RI.

Kementerian Kesehatan. (2021). *TB Indonesia*. Diambil dari Kementerian Kesehatan RI: [www.tbindonesia.or.id](http://www.tbindonesia.or.id). Diakses pada tanggal: 20 Oktober 2021.

### **Artikel Prosiding**

Tripena, A. (2011). Penentuan Model Regresi Spline Terbaik. Seminar Nasional Statistika. *Prosiding*. Semarang: Universitas Diponegoro.