

Regresi Nonparametrik *Spline Truncated* untuk Pemodelan Persentase Penduduk Miskin di Jawa Barat Pada Tahun 2021

Rosse Millania Pichago Firpha^{*}, Anneke Iswani Achmad

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

^{*}rossemillaniapichagof@gmail.com, annekeiswani11@gmail.com

Abstract. Nonparametric regression is a statistical method used to model the relationship between response variables and predictor variables whose pattern shape is unknown. In nonparametric regression there are several approaches, one of which is the spline. Spline nonparametric regression, if there is one predictor variable then the regression is called univariable spline nonparametric regression. Conversely, if there is one response variable with more than one predictor variable then the regression is called a multivariable spline nonparametric regression. In nonparametric regression there is a truncated spline model. The truncated spline function is a polynomial function that is dismembered at a knot point. A knot point is a joint fusion point where the function is truncated, or a point that describes a change in data behavior at a certain sub-sub-interval. Therefore, truncated spline models have an excellent ability to handle data whose behavior is arbitrary at certain sub-sub intervals. This study will use truncated spline nonparametric regression to model the number of poor people in West Java in 2021. The data used is secondary data sourced from the publication of the Indonesian Central Statistics Agency (BPS).

Keywords: *Nonparametric regression, Spline Truncated, GCV, Poor Population.*

Abstrak. Regresi nonparametrik adalah suatu metode statistika yang digunakan untuk memodelkan hubungan antara variabel respon dengan variabel prediktor yang tidak diketahui bentuk polanya. Dalam regresi nonparametrik terdapat beberapa pendekatan salah satunya spline. Regresi nonparametrik spline, jika terdapat satu variabel prediktor maka regresi tersebut dinamakan regresi nonparametrik spline univariabel. Sebaliknya, apabila terdapat satu variabel respon dengan lebih dari satu variabel prediktor maka regresi tersebut disebut regresi nonparametrik spline multivariable. Dalam regresi nonparametrik terdapat model *Spline Truncated*. Fungsi *Spline Truncated* merupakan fungsi polinomial yang terpotong-potong pada suatu titik knot. Titik knot merupakan titik perpaduan bersama dimana fungsi tersebut terpotong, atau titik yang menggambarkan terjadinya perubahan perilaku data pada sub-sub interval tertentu. Oleh karena itu, model *Spline Truncated* memiliki kemampuan yang sangat baik untuk menangani data yang perilakunya berubah-ubah pada sub-sub interval tertentu. Pada penelitian ini akan menggunakan regresi nonparametrik *Spline Truncated* untuk memodelkan jumlah penduduk miskin di Jawa Barat pada tahun 2021. Data yang digunakan adalah data sekunder yang bersumber dari publikasi Badan Pusat Statistika (BPS) Indonesia.

Kata Kunci: *Regresi nonparametric, Spline Truncated, GCV, Penduduk Miskin.*

A. Pendahuluan

Statistika ialah ilmu yang mempelajari teknik pengumpulan, pengolahan, analisis, dan penarikan kesimpulan menurut data yang dikumpulkan dan dianalisis. Banyak masalah dan metode yang dapat diselesaikan dengan menggunakan analisis dari statistika, salah satunya analisis regresi. (Gujarati, 2006) mengatakan regresi didefinisikan sebagai pemeriksaan hubungan antara satu atau dua variabel yang menjelaskan dan yang dijelaskan. Variabel terikat adalah variabel pertama, sedangkan variabel bebas adalah variabel kedua.

Ada tiga model untuk analisis regresi yaitu parametrik, nonparametrik, dan semiparametrik. Regresi nonparametrik adalah teknik statistik untuk memodelkan hubungan antara variabel respons dan variabel prediktor yang tidak diketahui. Regresi nonparametrik adalah jenis regresi yang sangat fleksibel dalam hal pemodelan pola data. Secara umum model regresi nonparametrik yakni: $Y_i = f(x_i) + \varepsilon_i$, $i = 1, 2, 3, \dots, n$ (2.1) dimana Y_i variabel respons, x_i variabel prediktor, $f(x_i)$ variabel yang tidak diketahui fungsi regresi pola, dan $\varepsilon_i \sim \text{IIDN}(0, \sigma^2)$. Sebagian teknik regresi nonparametrik yaitu spline.

Spline unggul dalam menangani data yang perilakunya bervariasi pada sub-interval spline tertentu (Budiantara I. N., 2009). Jika hanya ada satu variabel prediktor dalam regresi spline nonparametrik, regresi tersebut disebut regresi nonparametrik spline univariabel. Sebaliknya, jika variabel respons dikaitkan dengan banyak variabel prediktor, regresi disebut sebagai regresi nonparametrik spline multivariabel (Budiantara I. N., 2005), menggunakan fondasi fungsi keluarga *Spline Truncated*, pengembangan penaksir spline untuk regresi nonparametrik. Fungsi spline terpotong adalah fungsi polinomial yang berakhir pada titik knot.

Titik knot adalah titik koneksi di mana fungsi terpotong, atau titik yang menunjukkan pergeseran perilaku data pada sub-interval tertentu (Budiantara I. N., 2009). Maka, model *Spline Truncated* unggul dalam menangani data yang perilakunya bervariasi pada sub-interval tertentu (Budiantara I. N., 2011). Model tersebut dapat diaplikasikan pada jumlah penduduk miskin di Jawa Barat.

Kepala BPS Provinsi Jawa Barat Dyah Anugrah Kuswardani berkata, pada Maret 2021, persentase penduduk miskin di Jabar terjadi kenaikan 6,82 ribu jiwa, sebelumnya 4,19 juta jiwa (8,43%) pada September 2020 menjadi 4,20 juta jiwa (8,40 persen) pada Maret 2021. Jika dibandingkan dengan kondisi pada bulan September 2020, ini mengalami kenaikan sebesar 6,82 ribu jiwa. Kenaikannya cukup signifikan jika dibandingkan dengan Maret 2020. Hal tersebut dapat mengakibatkan berhasilnya upaya pembangunan kualitas hidup manusia berkurang sehingga dikatakan persentase penduduk miskin di Jawa Barat meningkat.

Keberhasilan dalam upaya meningkatkan kualitas hidup manusia (BPS 2018) di (BPS 2015), IPM menguraikan bagaimana masyarakat dapat memperoleh manfaat dari pembangunan dalam hal kesehatan, pendidikan, dan pendapatan. Nilai IPM yang rendah mungkin menunjukkan produktivitas pekerja yang buruk. Produktivitas tenaga kerja yang rendah mengakibatkan peningkatan pendapatan yang rendah, yang mencerminkan tingginya tingkat kemiskinan di suatu wilayah. Pengangguran memiliki banyak konsekuensi terhadap berbagai masalah.

Dalam persentase penduduk miskin, pengangguran mengurangi kesejahteraan masyarakat karena pendapatan mereka berkurang. Kesejahteraan masyarakat yang semakin rendah akan menimbulkan persoalan baru yaitu kemiskinan (Sukirno, 2006).

Menurut Thomas Robert Malthus, upah adalah harga yang ditetapkan untuk pekerjaan, yang ditentukan oleh penawaran dan permintaan tenaga kerja. Penawaran tenaga kerja akan meningkat sejalan dengan pertumbuhan penduduk. Hal ini akan berpengaruh pada tingkat upah yang akan diterapkan (Mada, 2015). Sehingga jika upah yang diterima rendah maka akan mempengaruhi tingkat kemiskinan.

Berdasarkan latar belakang yang telah diuraikan, maka perumusan masalah dalam penelitian ini sebagai akan dilakukan pemodelan terhadap persentase penduduk miskin di Jawa Barat tahun 2021 sebagai variabel tak bebas dan untuk variabel bebas terdiri dari IPM tahun 2021, Tingkat Pengangguran Terbuka tahun 2021 dan Upah Minimum Kabupaten/Kota tahun 2021 menggunakan metode analisis regresi *Spline Truncated*. Dengan begitu, penelitian ini berupaya mendapatkan model regresi nonparametrik *Spline Truncated* pada pemodelan

persentase penduduk miskin di Jawa Barat tahun 2021. Tujuan dalam penelitian ini yaitu untuk memperoleh model model persentase penduduk miskin di Jawa Barat pada tahun 2021 berdasarkan Indeks Pembangunan Manusia, Tingkat Pengangguran Terbuka, Upah Minimum Kabupaten/Kota terhadap persentase penduduk miskin di Jawa Barat tahun 2021 menggunakan regresi nonparametrik *Spline Truncated*.

B. Metodologi Penelitian

Peneliti menggunakan regresi nonparametrik *Spline Truncated* yang dilihat dari nilai GCV minimumnya. Populasi yang digunakan dalam penelitian ini adalah kabupaten/kota di Indonesia. Dengan sampel yang digunakan yaitu kabupaten/kota di Jawa Barat dengan jumlah sampel sebanyak 27 kabupaten/kota. Data yang digunakan merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS).

C. Hasil Penelitian dan Pembahasan

Deskripsi Data

Kondisi persentase penduduk miskin di Provinsi Jawa Barat beserta faktor yang mempengaruhi jumlah penduduk miskin pada penelitian ini jelaskan melalui statistik deskriptif seperti berikut:

Tabel 1. Statistik Deskriptif

	N	Minimum	Maksimum	Rata-Rata	Variansi
Persentase Penduduk Miskin (Y)	27	0,120	4,500	1,211	1,038
IPM (X ₁)	27	65,56	81,96	71,9407	22,859
TPT (X ₂)	27	3,25	13,07	9,4011	5,924
UMK (X ₃)	27	1831885	4798312	3038976	102003

Sumber: Data Penelitian yang Sudah Diolah, 2022

Dari tabel di atas, dapat diketahui bahwa pada diketahui bahwa rata-rata persentase penduduk miskin di Provinsi Jawa Barat adalah sebesar 1,211% dengan variansi 1,038%. Nilai rata-rata sebesar 1,211% menunjukkan bahwa persentase penduduk miskin yang tersebar di 27 kabupaten/kota di Jawa Barat adalah 1,211% dari total persentase penduduk miskin di masing-masing Kabupaten/Kota. Nilai tertinggi penduduk miskin di provinsi Jawa Barat adalah sebesar 4,500% dan terendah sebesar 0,120%. Karakteristik variabel X1 yaitu Indeks Pembangunan Manusia (IPM) yang memiliki rata-rata 71,9407 dengan variansi 22,859. Nilai rata-rata sebesar 71,9407% menunjukkan bahwa Indeks Pembangunan Manusia yang tersebar di 27 Kabupaten/Kota di Jawa Barat adalah 71,9407% dari total Indeks Pembangunan Manusia di masing-masing Kabupaten/Kota. Nilai tertinggi Indeks Pembangunan Manusia di provinsi Jawa Barat adalah sebesar 81,96% dan terendah sebesar 65,56%. Karakteristik variabel X2 yaitu Tingkat Pengangguran Terbuka (TPT) yang memiliki rata-rata 9,4011 dengan variansi 5,924. Nilai rata-rata sebesar 9,4011% menunjukkan bahwa Tingkat Pengangguran Terbuka yang tersebar di 27 Kabupaten/Kota di Jawa Barat adalah 9,4011% dari total Tingkat Pengangguran Terbuka di masing-masing Kabupaten/Kota. Nilai tertinggi Tingkat Pengangguran Terbuka di provinsi Jawa Barat adalah sebesar 9,4011% dan terendah sebesar 3,25%. Karakteristik variabel X3 yaitu Upah Minimum Kabupaten/Kota (UMK) yang memiliki rata-rata 3038976 dengan variansi 102003. Nilai rata-rata sebesar Rp. 3,038,976 menunjukkan bahwa Upah Minimum Kabupaten/Kota yang tersebar di 27 Kabupaten/Kota di Jawa Barat adalah Rp. 3,038,976 dari total Upah Minimum Kabupaten/Kota di masing-masing Kabupaten/Kota. Nilai tertinggi Upah Minimum Kabupaten/Kota di provinsi Jawa Barat adalah sebesar Rp. 4798,312 dan terendah sebesar Rp. 1,831,885.

Pemilihan Model Terbaik

Pada penelitian ini dilakukan pemilihan model terbaik berdasarkan nilai GCV minimum dari masing-masing titik knot untuk menentukan model terbaik yang disajikan pada Tabel 2 berikut:

Tabel 2. Nilai GCV masing-masing knot

Jumlah Knot	GCV Minimum
Satu knot	1,341801
Dua knot	1,238639
Tiga knot	0,990854*
Kombinasi knot (3,2,3)	0.990854

Pada hasil nilai GCV minimum dari masing-masing knot. Pemodelan dilihat dari nilai GCV minimum. Tabel 4.2 menunjukkan bahwa nilai GCV paling minimum yaitu pada tiga knot, sehingga satu knot digunakan sebagai model terbaik regresi nonparametrik *Spline Truncated*.

Adapun model regresi nonparametrik Spline yang terbentuk dari tiga titik knot pada faktor-faktor yang diduga mempengaruhi persentase penduduk miskin di Jawa Barat sebagai berikut.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - K_1)_+^1 + \hat{\beta}_3 (x_1 - K_2)_+^1 + \hat{\beta}_4 (x_1 - K_3)_+^1 + \hat{\beta}_5 x_2 + \hat{\beta}_6 (x_2 - K_4)_+^1 + \hat{\beta}_7 (x_2 - K_5)_+^1 + \hat{\beta}_8 (x_2 - K_6)_+^1 + \hat{\beta}_9 x_3 + \hat{\beta}_{10} (x_3 - K_7)_+^1 + \hat{\beta}_{11} (x_3 - K_8)_+^1 + \hat{\beta}_{12} (x_3 - K_9)_+^1$$

Interpretasikan model regresi nonparametrik spline terbaik

Berikut merupakan model terbaik dengan satu knot yang telah didapatkan.

1. Jika variabel selain X_1 , yaitu X_2 , dan X_3 dianggap konstan, maka pengaruh Indeks Pembangunan Manusia (X_1) persentase penduduk miskin di Jawa Barat adalah sebagai berikut.

$$\hat{y} = -7,231456 + 1,674705x_1 + 7,534624(x_1 - 76,712)_+^1 + 6,644366(x_1 - 77,368)_+^1 + 5,754111(x_1 - 78,024)_+^1$$

$$= \left. \begin{cases} -7,231456 + 1,674705x_1 & ; x_1 < 76,712 \\ -585,227532 + 9,208674x_1 & ; 76,712 \leq x_1 < 77,368 \\ -1099,2884 + 15,85304x_1 & ; 77,368 \leq x_1 < 78,024 \\ -1663,51594 + 21,607151x_1 & ; x_1 \geq 78,024 \end{cases} \right\}$$

2. Dengan mengasumsikan variabel prediktor selain X_2 konstan, maka persamaan regresi dari variabel Tingkat Pengangguran Terbuka dan variabel tingkat pengangguran terbuka termasuk ke dalam komponen nonparametrik, sehingga modelnya dapat ditulis sebagai berikut.

$$\hat{y} = -7,231456 + 4,89143x_2 + 1,064705(x_2 - 99,276)_+^1 - 3,691056(x_2 - 10,3204)_+^1 + 3,58888(x_2 - 10,7132)_+^1$$

$$= \left. \begin{cases} -7,231456 + 4,89143x_2 & ; x_2 < 99,276 \\ -112,93111 + 5,956118x_2 & ; 99,276 \leq x_2 < 10,3204 \\ -74,837932 + 2,265062x_2 & ; 10,3204 \leq x_2 < 10,7132 \\ -38,448389 + 3,58888x_2 & ; x_2 \geq 10,7132 \end{cases} \right\}$$

3. Dengan mengasumsikan variabel prediktor selain X_3 konstan, maka persamaan regresi dari variabel Upah Minimum Kabupaten/Kota dan variabel upah minimum kabupaten/kota termasuk ke dalam komponen nonparametrik, sehingga modelnya dapat ditulis sebagai berikut.

$$\hat{y} = -7,231456 + 1,225778x_3 + 2,46681(x_3 - 33962859776)_+^1 + 1,244978(x_3 - 357153923041)_+^1 + 1,384853(x_3 - 37467924832)_+^1$$

$$\left(\begin{array}{ll} -7,231456 + 1,225778x_3 & ; x_3 < 33962859776 \\ 1,146536 - 1,241032x_3 & ; 33962859776 \leq x_3 < 357153923041 \\ -3,2999951768 + 3,946x_3 & ; 357153923041 \leq x_3 < 3746792483 \\ -4,042220811 + 1,388799x_3 & ; x_3 \geq 37467924832 \end{array} \right)$$

D. Kesimpulan

Berdasarkan hasil dan pembahasan pada penelitian ini dapat disimpulkan bahwa, Model regresi nonparametrik *Spline Truncated* untuk persentase penduduk miskin di Jawa Barat pada tahun 2021 merupakan model tiga knot dengan nilai GCV minimum 0,990854. Sebanyak 3 variabel prediktor yang digunakan berpengaruh signifikan. Sehingga diperoleh model regresi nonparametrik *Spline Truncated* berikut.

$$\hat{y} = -7,231456 + 1,674705x_1 + 7,534624(x_1 - 76,712)_+^1 + 6,644366(x_1 - 77,368)_+^1 + 5,754111(x_1 - 78,024)_+^1 + 4,894143x_2 + 1,064705(x_2 - 99,276)_+^1 - 3,691056 + (x_2 - 10,3204)_+^1 + 3,58888(x_2 - 10,7132)_+^1 + 1,225378x_3 - 2,46681(x_3 - 33962859776)_+^1 + 1,244978(x_3 - 35715392304)_+^1 + 1,384853(x_3 - 37467924832)_+^1$$

Dari model di atas dapat dilihat bahwa nilai konstanta memiliki nilai negatif sebesar -7,231456. Yang artinya menunjukkan jika semua variabel independen yang meliputi Indeks Pembangunan Manusia (X1), Tingkat Pengangguran Terbuka (X2), dan Upah Minimum Kabupten/Kota (X3) bernilai 0 atau tidak mengalami perubahan, maka persentase penduduk miskin adalah -7,231456. Nilai koefisien Indeks Pembangunan Manusia (X1) sebesar 1,674705. Yang artinya jika Indeks Pembangunan Manusia naik sebesar 1% maka persentase penduduk miskin akan naik sebesar 1,674705. Nilai koefisien Tingkat Pengangguran Terbuka (X2) sebesar 4,894143. Yang artinya jika Tingkat Pengangguran Terbuka naik sebesar 1% maka persentase penduduk miskin akan naik sebesar 4,894143. Nilai koefisien Upah Minimum Kabupten/Kota (X3) sebesar 1,225378. Yang artinya jika Tingkat Pengangguran Terbuka naik sebesar 1% maka persentase penduduk miskin akan naik sebesar 1,225378.

Daftar Pustaka

- [1] Anindita, S. D. (2018). Pemodelan Balita Stunting di Indonesia Menggunakan Regresi Non Parametrik *Spline Truncated*.
- [2] BPS. (2020). <https://bps.go.id/subject/23/kemiskinan-dan-ketimpangan.html>.
- [3] Budiantara, I. (2011). Infants' Weight Growth Model in Surabaya International Journal of Basic & Applied Sciences (pp. 151-156). Surabaya: Using Weighted Spline Regression.
- [4] Budiantara, I. N. (2005). Berkala MIPA. Yogyakarta: Universitas Gadjah Mada.
- [5] Budiantara, I. N. (2009). Spline dalam Regresi Nonparametrik dan Semiparametrik. Surabaya: ITS Press.
- [6] Budiantara, I. N. (2012). Modeling the Percentage of Poor People in Indonesia Using Spline. International Journal of Basic & Applied Sciences 12, 119–124.
- [7] Eubank, R. (1999). Nonparametric Regression and Spline Smoothing. New York: Marcel Dekker Inc.
- [8] Fitriani, Anisa Nur. (2021). *Selang Kepercayaan Koefisien Korelasi Berdasarkan Empirical Likelihood dan Penerapannya pada Data Rata-Rata Lama Sekolah dan Penduduk Miskin Kota/Kabupaten di Indonesia*, Jurnal Riset Statistika, 1(1), 51-56.