

Faktor-faktor yang Mempengaruhi Persentase Kasus HIV di Indonesia Tahun 2017 Menggunakan Regresi Nonparametrik *Spline Truncated*

Azizah, dan I Nyoman Budiantara

Departemen Statistika, Fakultas Matematika, Komputasi, dan Sains Data,
Institut Teknologi Sepuluh Nopember (ITS)
e-mail: nyomanbudiantara65@gmail.com

Abstrak—*Human Immunodeficiency Virus (HIV)* secara berkelanjutan telah menjadi permasalahan utama bagi pembangunan dan kemajuan sosial. Di Indonesia, kasus HIV mengalami peningkatan tiap tahunnya. Pada tahun 2017 negara Indonesia menjadi negara ketiga dengan kasus HIV terbanyak di Asia-Pasifik. Penyebab meningkatnya kasus HIV dipengaruhi beberapa faktor, untuk mengetahui faktor yang mempengaruhinya dilakukan penelitian menggunakan Regresi Nonparametrik *Spline Truncated*. Karena pola hubungan yang ditunjukkan antar persentase kasus HIV dengan faktor-faktor yang diduga mempengaruhinya tidak mengikuti pola tertentu. Berdasarkan nilai GCV yang paling minimum, model terbaik adalah menggunakan kombinasi titik knot (3,3,2,2,1). Hasil pengujian signifikansi parameter menunjukkan bahwa seluruh variabel yang digunakan dalam penelitian berpengaruh signifikan terhadap persentase kasus HIV di Indonesia. Variabel yang digunakan adalah rasio layanan KT, persentase penggunaan kondom, persentase penduduk miskin, rasio layanan PDP, dan persentase penduduk laki-laki. Dan hasil pengujian asumsi residual menunjukkan semua asumsi terpenuhi dengan nilai koefisien determinasi dari model ini sama dengan 89,49%.

Kata Kunci— *GCV, HIV, Indonesia, Regresi Nonparametrik, Spline Truncated, Titik Knot.*

I. PENDAHULUAN

SUSTAINABLE *Development Goals* (SDGs) merupakan program pembangunan dunia untuk kepentingan manusia dan planet bumi, yang disepakati oleh 191 negara anggota Perserikatan Bangsa Bangsa (PBB) salah satunya Negara Indonesia. Dalam tujuan SDGs urutan yang ketiga, kesehatan memiliki tempat sentral yaitu mengenai memastikan hidup sehat dan meningkatkan kesejahteraan untuk semua usia. Hal tersebut dapat dicapai, salah satunya dengan memerangi kasus HIV dengan cara mengantisipasi peningkatan kasus HIV [1]. *Human Immunodeficiency Virus (HIV)* dapat meruntuhkan kemampuan tubuh untuk menangkis beberapa infeksi dan penyakit lainnya, atau mengakibatkan kekebalan tubuh manusia menurun dengan cara menyerang/menginfeksi sel darah putih [2].

Jumlah kasus HIV yang dinyatakan positif di Indonesia sejak tahun 2005 hingga tahun 2017 cenderung mengalami peningkatan yang tajam tiap tahunnya. Tahun 2017 Indonesia

dilaporkan berada di posisi ketiga setelah Negara India dan China dengan kasus HIV terbanyak di Asia-Pasifik [3]. Pemerintah Indonesia sebenarnya sudah melakukan berbagai strategi untuk mengurangi jumlah kasus HIV, yaitu dengan disediakan pengobatan terapi *Antiretroviral (ARV)* yang dapat menurunkan jumlah virus HIV agar mampu mencegah penularan terhadap orang lain. Kemudian dengan pemberian informasi mengenai HIV secara umum melalui kegiatan sekolah atau luar sekolah, pelayanan kesehatan dan strategi penggunaan kondom sebagai bentuk pencegahan [4]. Tetapi strategi yang dilakukan pemerintah Indonesia masih belum mampu menekan permasalahan kasus HIV di Indonesia tiap tahunnya.

Pengetahuan mengenai penyebab HIV sangat diperlukan bagi masyarakat Indonesia sebagai strategi untuk mencegah penularannya sekaligus menekan persentase kasus positif HIV di setiap provinsi di Indonesia, melihat bagaimana peningkatan jumlah kasus ini di Negara Indonesia tiap tahunnya, serta mewujudkan keberhasilan tujuan SDGs urutan ketiga. Penularan kasus HIV erat kaitannya dengan perilaku berisiko, sehingga perlu diketahui faktor-faktor yang mempengaruhi kasus HIV [5].

Berdasarkan referensi dan penelitian sebelumnya dengan kasus yang serupa dan metode yang berbeda, penelitian ini akan memodelkan faktor-faktor yang mempengaruhi persentase kasus HIV di Indonesia menggunakan regresi nonparametrik *Spline Truncated*. Hal ini dikarenakan kurva regresi yang dibentuk antara variabel prediktor dengan variabel respon pada data kasus HIV di Indonesia tahun 2017, tidak mengikuti pola tertentu. Dan pemilihan metode regresi nonparametrik *Spline Truncated* didasari oleh kelebihan dari metode ini dibandingkan metode regresi nonparametrik *Spline* lainnya, yaitu metode lebih sederhana, interpretasi mudah, dan perumusan matematis lebih sederhana. Diharapkan dapat bermanfaat dalam kemajuan ilmu pengetahuan dan mampu dijadikan bahan pertimbangan bagi pihak pemerintah Indonesia dalam pengambilan keputusan mengatasi peningkatan kasus HIV di Indonesia tiap tahunnya.

II. TINJAUAN PUSTAKA

A. Regresi Nonparametrik *Spline Truncated*

Regresi nonparametrik *spline truncated* merupakan metode

yang paling banyak digunakan pada regresi nonparametrik. Bentuk kurva *spline* terpotong-potong sehingga mampu mengatasi perubahan pola data pada sub interval tertentu. Pada metode regresi nonparametrik *spline truncated* digunakan bantuan titik-titik knot. Titik knot merupakan titik dimana terjadi pola perubahan perilaku dari suatu fungsi pada selang yang berbeda [6]. Fungsi *spline truncated* multivariabel berorde q dengan titik knot K_1, K_2, \dots, K_r . Dapat dituliskan menjadi persamaan sebagai berikut :

$$\sum_{j=1}^p f_j(x_{ji}) = \sum_{j=1}^p \sum_{u=0}^q \beta_{ju} x_{ji}^u + \sum_{j=1}^p \sum_{k=1}^r \beta_{j(q+k)} (x_{ji} - K_{jk})_+^q \quad (1)$$

Model regresi nonparametrik *spline truncated* sebagai berikut,

$$y_i = \sum_{j=1}^p \sum_{u=0}^q \beta_{ju} x_{ji}^u + \sum_{j=1}^p \sum_{k=1}^r \beta_{j(q+k)} (x_{ji} - K_{jk})_+^q + \varepsilon_i \quad (2)$$

Fungsi *truncated* $(x_{ji} - K_{jk})_+^q$ akan menghasilkan persamaan (3) berikut,

$$(x_{ji} - K_{jk})_+^q = \begin{cases} (x_{ji} - K_{jk})^q, & x_{ji} > K_{jk} \\ 0, & x_{ji} \leq K_{jk} \end{cases} \quad (3)$$

Fungsi *Spline* yang digunakan adalah *Spline* linear dengan derajat $q = 1$.

B. Pemilihan Titik Knot Optimum

Model regresi *spline* terbaik merupakan model yang memiliki titik knot optimal. Titik knot merupakan titik yang terdapat pada perubahan pola perilaku fungsi. Salah satu metode yang biasa digunakan untuk memilih titik knot optimal adalah metode *Generalized Cross Validation* (GCV). Menggunakan metode GCV didasari oleh kelebihan yang dimiliki metode ini, yaitu mempunyai sifat optimal asimtotik, tidak memuat varians populasi (σ^2) yang tidak diketahui, *invariance* terhadap transformasi. Model regresi *spline* terbaik diperoleh dari titik knot optimal dengan melihat nilai GCV terkecil [7].

Metode GCV dapat dituliskan sebagai berikut.

$$GCV(K_1, K_2, \dots, K_r) = \frac{MSE(K_1, K_2, \dots, K_r)}{[n^{-1} \text{trace}(\mathbf{I} - \mathbf{A})]^2} \quad (4)$$

dimana \mathbf{I} merupakan matriks identitas, n adalah jumlah pengamatan, $K = (K_1, K_2, \dots, K_r)$ merupakan titik-titik knot,

$$MSE(K_1, K_2, \dots, K_r) = n^{-1} \sum_{i=1}^n (y_i - \hat{f}(x_i))^2 \quad (5)$$

serta $\mathbf{A} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$ [8].

C. Pengujian Parameter Model Regresi

Pengujian secara serentak dilakukan untuk mengetahui signifikansi parameter model regresi secara bersama-sama. Hipotesis untuk pengujian parameter secara serentak adalah sebagai berikut.

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

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

Dimana p adalah jumlah variabel prediktor dan r adalah jumlah titik knot. Statistik uji dalam pengujian serentak menggunakan uji F seperti pada persamaan 2.22 berikut ini.

$$F_{hitung} = \frac{MS_{regresi}}{MS_{residual}} \quad (6)$$

Pengujian parameter model secara serentak dapat disajikan menggunakan *Analysis of Varians* (ANOVA) yang disajikan dalam Tabel 1 sebagai berikut.

Tabel 1.
Analisis ragam (ANOVA) Uji Parameter

Sumber Variasi	Df	Sum of Square (SS)	Mean Square (MS)	F_{hitung}
Regresi	$p+r$	$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	$\frac{SS_{regresi}}{df_{regresi}}$	$\frac{MS_{regresi}}{MS_{error}}$
Error	$n-(p+r)-1$	$\sum_{i=1}^n (y_i - \hat{y}_i)^2$	$\frac{SS_{error}}{df_{error}}$	
Total	$n-1$	$\sum_{i=1}^n (y_i - \bar{y})^2$		

Nilai $p+r$ merupakan banyak parameter dalam model regresi nonparametrik *spline* kecuali β_0 . Tolak H_0 jika $F_{hitung} > F_{(\alpha; (p+r)n-(p+r)-1)}$ sehingga menghasilkan kesimpulan bahwa minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon [9].

Pengujian secara parsial bertujuan untuk mengetahui parameter mana yang berpengaruh dan tidak berpengaruh signifikan terhadap model regresi. Berikut merupakan hipotesis untuk pengujian secara parsial:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, j = 1, 2, \dots, p+r$$

Pengujian secara parsial dilakukan dengan menggunakan uji t [9]. Statistik uji yang digunakan adalah sebagai berikut:

$$t_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (7)$$

$SE(\hat{\beta}_j)$ adalah *standart error* $\hat{\beta}_j$ dimana $SE(\hat{\beta}_j) = \sqrt{\text{Var}(\hat{\beta}_j)}$ dengan $\text{Var}(\hat{\beta}_j)$ merupakan elemen diagonal utama ke- j dari matriks $\text{Var}(\hat{\beta})$, $j = 1, 2, \dots, p+r$.

Tolak H_0 jika $t > t_{\frac{\alpha}{2}; (n-(p+r)-1)}$ dan $t < -t_{\frac{\alpha}{2}; (n-(p+r)-1)}$ sehingga menghasilkan kesimpulan bahwa variabel prediktor ke- n berpengaruh signifikan terhadap variabel respon [9].

D. Koefisien Determinasi (R^2)

Koefisien determinasi adalah kuantitas yang dapat menjelaskan sumbangan variabel prediktor terhadap variabel respon. Semakin tinggi nilai R^2 yang dihasilkan suatu model, maka semakin baik pula variabel-variabel prediktor dalam model tersebut dalam menjelaskan variabilitas variabel respon. Berikut ini adalah rumus untuk mendapatkan nilai R^2 [10].

$$R^2 = \frac{SS_{regresi}}{SS_{total}} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (8)$$

E. Pengujian Asumsi Residual

Pengujian asumsi residual (*goodness of fit*) model regresi paling populer karena mudah digunakan. Residual yang dihasilkan harus memenuhi asumsi. Terdapat tiga asumsi yang

harus dipenuhi yaitu identik, independen, dan berdistribusi normal.

Asumsi identik atau biasa juga disebut homoskedastisitas yang berarti bahwa varians pada residual adalah identik. Kebalikanannya adalah kasus heteroskedastisitas, yaitu jika kondisi varians *residual* tidak identik [9].

$$\text{var}(y_i) = \text{var}(\varepsilon_i) = \sigma^2 ; i = 1, 2, \dots, n \quad (8)$$

Uji identik dapat menggunakan uji Glejser. Perumusan hipotesisnya adalah sebagai berikut:

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

$$H_1 : \text{Minimal ada satu } \sigma_i^2 \neq \sigma^2 ; i = 1, 2, \dots, n$$

Statistik uji yang digunakan adalah,

$$F_{hitung} = \frac{\sum_{i=1}^n (|\hat{\varepsilon}_i| - |\bar{\varepsilon}|)^2}{p+r} \div \frac{\sum_{i=1}^n (|\varepsilon_i| - |\hat{\varepsilon}_i|)^2}{n-(p+r)-1} \quad (9)$$

Tolak H_0 jika $F_{hitung} > F_{\alpha;(v,n-v-1)}$ atau $p\text{-value} < \alpha$ yang berarti bahwa tidak terindikasi terdapat kasus homoskedastisitas.

Asumsi independen merupakan asumsi dari model regresi yang mengharuskan tidak terdapat korelasi antar residual. Uji yang digunakan untuk mendeteksi kasus autokorelasi adalah *run test*. Dalam sampel yang lebih besar dari 20, maka uji yang digunakan adalah uji Z. Perumusan hipotesis yang digunakan adalah sebagai berikut.

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

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

Statistik Uji:

$$Z = \frac{\text{observed} - \text{expected}}{\sqrt{\text{variance}}} \quad (9)$$

dengan:

observed = banyaknya runtun dalam sampel

$$\text{expected} = 1 + \frac{2n_1n_2}{n_1 + n_2}$$

$$\text{variance} = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}$$

n_1 = jumlah residual bernilai positif

n_2 = jumlah residual bernilai negatif

Pengambilan keputusan dari *run test* akan dibandingkan dengan tabel distribusi normal standar, dengan keputusan Tolak H_0 jika nilai $Z > Z_{(1-\alpha/2)}$ dan $-Z < Z_{(\alpha/2)}$ atau $p\text{-value} < \alpha$ dengan taraf signifikansi sama dengan 5%, artinya terjadi kasus autokorelasi pada model [11].

Uji *Kolmogorov-Smirnov* dilakukan untuk mengetahui apakah suatu data telah mengikuti suatu distribusi normal [12].

Hipotesis :

$$H_0 : F_n(\varepsilon) = F_0(\varepsilon)$$

$$H_1 : F_n(\varepsilon) \neq F_0(\varepsilon)$$

Atau

H_0 : residual berdistribusi normal

H_1 : residual tidak berdistribusi normal

Statistik uji :

$$D = \text{Sup}_{\varepsilon} |F_n(\varepsilon) - F_0(\varepsilon)| \quad (10)$$

Tolak H_0 apabila $|D| > D_{(1-\alpha)}$. $D_{(1-\alpha)}$ adalah nilai kritis untuk uji *Kolmogorov Smirnov* satu sampel, diperoleh dari tabel *Kolmogorov Smirnov* satu sampel, $F_n(\varepsilon)$ merupakan nilai peluang kumulatif (fungsi distribusi kumulatif) berdasarkan data sampel, $F_0(\varepsilon)$ adalah nilai peluang kumulatif (fungsi distribusi kumulatif) dibawah H_0 .

F. Human Immunodeficiency Virus

Penularan HIV dapat melalui cairan semen, sekresi serviks/vagina, limfosit, sel-sel dalam plasma bebas, cairan serebrospinal, air mata, saliva, air seni, dan air susu. Semua cairan tersebut dapat menularkan infeksi dalam konsentrasi tertentu. Cairan yang memiliki resiko tertinggi dalam menularkan virus adalah darah, air mani, sekresi vagina, dan air susu ibu. Proses penularan dikaitkan dengan perilaku seks seperti homoseksual dan heteroseksual, penggunaan jarum suntik pada penyalahgunaan NAPZA, kesalahan medis dalam penggunaan jarum atau benda tajam yang telah tercemar, donor darah, transplantasi organ, tindakan medis *infasif* atau *in utero*, perinatal, dan ASI yang tersalur terhadap anak. Belum ada informasi yang dapat membuktikan mengenai penularan HIV melalui kontak sosial, alat makan, toilet, kolam renang, udara ruangan, maupun oleh nyamuk atau serangga [13].

III. METODOLOGI PENELITIAN

A. Sumber Data

Data yang digunakan pada penelitian ini merupakan data sekunder yang diperoleh dari Profil Kesehatan Indonesia tahun 2017. Data tersebut mengenai kasus HIV menurut provinsi di Indonesia tahun 2017. Unit observasi yang digunakan adalah 34 provinsi di Indonesia.

B. Variabel Penelitian

Variabel respon yang digunakan dalam penelitian dijelaskan pada Tabel 2 berikut.

Tabel 2.
Kebutuhan Data Penelitian

No.	Variabel	Keterangan
1	Y	Persentase kasus HIV
2	X ₁	Rasio layanan KT
3	X ₂	Persentase penggunaan kontrasepsi jenis kondom
4	X ₃	Persentase penduduk miskin
5	X ₄	Rasio layanan PDP
6	X ₅	Persentase penduduk laki-laki

C. Langkah Penelitian

Metode yang digunakan pada penelitian ini adalah Regresi Nonparametrik *Spline Truncated*. Berikut adalah langkah-langkah analisis yang digunakan dalam melakukan penelitian ini sesuai dengan tujuan penelitian pada Bab 1.

1. Mendeskripsikan karakteristik kasus HIV di Indonesia tahun 2017 dan faktor-faktor yang diduga mempengaruhinya
2. Mengidentifikasi pola hubungan antara variabel respon dengan masing-masing variabel prediktor
3. Menganalisis model kasus HIV di Indonesia tahun 2017 menggunakan Regresi Nonparametrik *Spline Truncated* dengan langkah sebagai berikut

- Memodelkan variabel respon menggunakan dengan satu, dua, tiga dan kombinasi knot
- Menentukan titik knot optimal berdasarkan nilai GCV yang paling minimum
- Memperoleh model terbaik dengan titik knot optimal
- Mendapatkan estimator parameter model
- Melakukan pengujian signifikansi parameter secara serentak dan parsial
- Melakukan uji asumsi residual identik, independen, dan berdistribusi normal (IIDN)
- Menghitung nilai R² dan MSE
- Membuat interpretasi model dan menarik kesimpulan

IV. ANALISIS DAN PEMBAHASAN

A. Karakteristik Persentase Kasus HIV di Indonesia dan Faktor-Faktor yang Mempengaruhinya.

Berikut karakteristik data berdasarkan variabel penelitian ditampilkan pada Tabel 3.

Tabel 3. Karakteristik Persentase Kasus dan Faktor-Faktor yang Diduga Berpengaruh

Variabel	Mean	Varians	Min	Maks
Y	1,943	0,969	0,600	5,070
X ₁	20,30	447,59	0,01	108,11
X ₂	1,344	1,080	0,110	5,210
X ₃	10,951	33,493	3,780	27,760
X ₄	2,637	4,717	0,463	8,500
X ₅	50,713	1,180	48,533	53,060

Persentase kasus HIV adalah angka untuk mengetahui seberapa banyak klien yang mengikuti test dinyatakan positif HIV. Berdasarkan Tabel 3 diketahui bahwa rata-rata persentase kasus HIV di Indonesia tahun 2017 sama dengan 1,943%, artinya setiap 100 klien yang berkunjung pada tahun 2017 untuk mengikuti Konseling dan Test HIV 2 orang diantaranya dinyatakan positif terjangkit HIV dengan nilai varians sama dengan 0,969. Hal ini menunjukkan bahwa pada tahun 2017, persentase kasus HIV di Indonesia cenderung merata disetiap provinsi di Indonesia dengan nilai antara 0,6% sampai 5,070%.

B. Pemilihan Titik Knot Optimum

Berdasarkan pemodelan yang telah dilakukan dengan berbagai titik knot (satu, dua, tiga, dan kombinasi), kemudian masing-masing nilai GCV pada tiap titik knot yang paling optimal tersebut dibandingkan. Berikut nilai GCV minimum pada tiap pemodelan menggunakan satu, dua, tiga dan kombinasi knot.

Tabel 4. Perbandingan Nilai GCV Minimum tiap Titik Knot

No.	Knot	GCV Minimum
1	Satu Titik Knot	0,698935504
2	Dua Titik Knot	0,604087385
3	Tiga Titik Knot	0,505032281
4	Kombinasi Knot (3, 3, 2, 2, 1)	0,395511182

Nilai GCV yang paling minimum pada tiap titik knot, menunjukkan bahwa pemodelan regresi nonparametrik *spline truncated* menggunakan kombinasi titik knot pada titik

(3,3,2,2,1) menghasilkan nilai yang paling optimal. Dengan tiga titik knot pada variabel rasio layanan KT, tiga titik knot pada variabel persentase penggunaan kondom, dua titik knot pada variabel persentase penduduk miskin, dua titik knot pada variabel rasio layanan PDP, dan satu titik knot pada variabel persentase penduduk laki-laki.

C. Penaksiran Parameter

Berdasarkan nilai GCV yang paling minimum titik knot paling optimal yang digunakan untuk memodelkan data persentase kasus HIV di Indonesia tahun 2017 adalah titik kombinasi, yaitu pada titik (3,3,2,2,1). Sehingga dapat diestimasi parameter untuk model regresi nonparametrik *spline truncated* dengan titik kombinasi. Estimasi parameter menggunakan metode *Ordinary Least Square* (OLS) adalah sebagai berikut.

$$\hat{y} = -56,33 + 2,27x_1 - 3,36(x_1 - 2,22)_+ + 1,50(x_1 - 6,63)_+ - 0,42(x_1 - 8,84)_+ - 11,92x_2 + 24,37(x_2 - 0,21)_+ - 29,12(x_2 - 0,42)_+ + 16,78(x_2 - 0,53)_+ + 7,62x_3 - 8,98(x_3 - 4,27)_+ + 1,36(x_3 - 5,74)_+ + 16,67x_4 - 20,55(x_4 - 0,60)_+ + 3,88(x_4 - 1,09)_+ + 0,39x_5 - 0,79(x_5 - 50,66)_+$$

E. Pengujian Signifikansi Parameter

Pengujian parameter model regresi nonparametrik *spline truncated* terdiri dari pengujian secara serentak dan pengujian secara parsial atau individu sebagai berikut. Hipotesis yang digunakan untuk pengujian parameter secara serentak adalah sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{16} = 0$$

$$H_1 : \text{Minimal terdapat satu } \beta_j \neq 0, j = 1, 2, \dots, 16$$

Dengan menggunakan taraf kepercayaan 95% diperoleh $F_{(0,05;16;17)} = 3,27$. Berikut hasil pengujian parameter secara serentak ditampilkan pada Tabel 5.

Tabel 5. Analysis of Variance

Sumber Variasi	df	SS	MS	F	P-Value
Regresi	16	28,62962	1,789351	9,05	2,0902x10 ⁻⁵
Error	17	3,361845	0,1977556		
Total	33	31,99146			

Berdasarkan Tabel 5 diperoleh nilai statistik uji F_{hitung} sama dengan 9,05. Karena nilai statistik uji F_{hitung} > F_(0,05;16;17), maka menghasilkan keputusan Tolak H₀. Hal ini menunjukkan bahwa minimal terdapat satu parameter yang berpengaruh signifikan pada model, atau secara bersama terdapat minimal satu variabel prediktor berpengaruh signifikan terhadap persentase kasus HIV di Indonesia. Berikut hipotesis untuk melakukan uji parameter secara parsial.

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, j = 1, 2, \dots, 16$$

Pengujian parameter secara parsial akan dibandingkan dengan $t_{0,025;17} = 2,109$ dan $-t_{0,025;17} = -2,109$ dengan menggunakan taraf kepercayaan 95%.

Tabel 6. Hasil Pengujian Signifikansi Parameter Secara Individu

Variabel	Parameter	Estimator	t	P-value
----------	-----------	-----------	---	---------

Konstan	β_0	-56,335	-4,246890	0,000543
x_1	β_1	2,271	5,220739	$6,922 \times 10^{-5}$
	β_2	-3,365	-5,121940	$8,497 \times 10^{-5}$
	β_3	1,502	3,258194	0,0046293
	β_4	-0,421	-1,943351	0,0687200

Tabel 6. Hasil Pengujian Signifikansi Parameter Secara Individu (Lanjutan)

Variabel	Parameter	Estimator	t	P-value
x_2	β_5	-11,916	-1,683312	0,1105871
	β_6	24,367	2,675150	0,0159821
	β_7	-29,125	-3,327255	0,0039883
	β_8	16,775	2,845642	0,0111766
x_3	β_9	7,618	3,253256	0,0046789
	β_{10}	-8,977	-3,453616	0,003034
	β_{11}	1,364	3,410308	0,003332
x_4	β_{12}	16,667	3,196157	0,005291
	β_{13}	-20,554	-3,166650	0,005637
	β_{14}	3,883	2,645446	0,017007

Berdasarkan Tabel 6 diperoleh dua nilai statistik uji $t > -t_{(0,025;17)}$ sehingga Gagal Tolak H_0 , yaitu parameter β_4 (pada variabel x_1) dan β_5 (pada variabel x_2). Hal ini menunjukkan parameter β_4 dan β_5 tidak berpengaruh signifikan terhadap model. Sedangkan lima belas parameter lainnya berpengaruh signifikan terhadap model, karena nilai statistik uji $t > t_{(0,025;17)}$ dan $t < -t_{(0,025;17)}$.

Walaupun terdapat parameter yang tidak signifikan pada variabel x_1 dan x_2 , variabel tersebut masih memiliki parameter yang signifikan dalam masing-masing variabel. Sehingga seluruh variabel prediktor berpengaruh signifikan terhadap persentase kasus HIV di Indonesia secara parsial.

D. Pengujian Asumsi Residual

Terdapat tiga asumsi residual yang harus dipenuhi sebuah model dalam analisis regresi nonparamterik *spline truncated*, yaitu asumsi residual identik, dan independen dan berdistribusi normal. Hasil pengujian asumsi residual adalah sebagai berikut. Berikut hipotesis pengujian asumsi identik.

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

$$H_1 : \text{Minimal ada satu } \sigma_i^2 \neq \sigma^2 ; i = 1, 2, \dots, 34$$

Dengan menggunakan taraf kepercayaan 95% diperoleh $F_{(0,05;16;17)} = 3,27$. Hasil pengujian asumsi identik ditampilkan pada Tabel 6.

Tabel 7. Hasil Pengujian Statistik Uji Glejser

Sumber	Df	SS	MS	Fhit	P-value
Regresi	16	0,579455	0,036216	0,889294	0,590714
Error	17	0,692314	0,040724		
Total	33	1,271769			

Berdasarkan Tabel 7 diperoleh nilai statistik uji F_{hitung} sama dengan 0,8892 dengan $p\text{-value} = 0,5907$. Karena nilai statistik uji $F_{hitung} < F_{(0,05;16;17)}$, maka menghasilkan keputusan Gagal Tolak H_0 . Hal ini menunjukkan tidak terjadi heterokedastisitas pada model, sehingga asumsi residual identik terpenuhi.

Hipotesis pengujian asumsi residual independen adalah sebagai berikut.

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

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

Dengan menggunakan taraf kepercayaan 95% diperoleh $Z_{(0,025)} = 1,96$. Hasil pengujian asumsi independen menggunakan *run test* adalah sebagai berikut.

$$expected = 1 + \frac{2n_1n_2}{n_1 + n_2} = 18$$

$$observed = 17$$

$$variance = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)} = 8.242424$$

$$Z = \frac{observed - expected}{\sqrt{variance}} = -0.34832$$

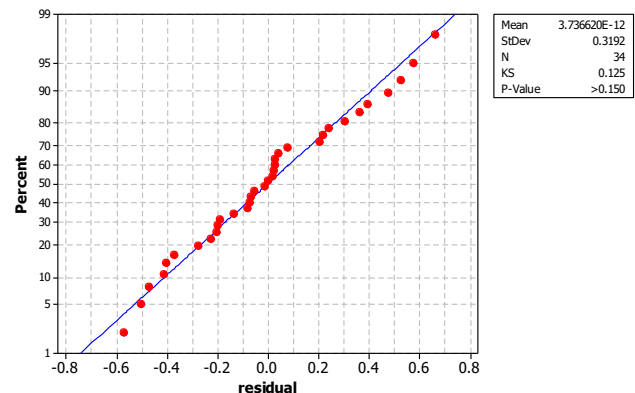
Berdasarkan perhitungan diperoleh nilai statistik uji Z sama dengan -0,34832. Karena nilai statistik uji $Z < Z_{(0,025)}$, maka menghasilkan keputusan Gagal Tolak H_0 . Hal ini menunjukkan tidak terjadi autokorelasi pada model, sehingga asumsi residual independen terpenuhi.

Hipotesis pengujian asumsi berdistribusi normal adalah sebagai berikut.

$$H_0 : F_n(\epsilon) = F_0(\epsilon)$$

$$H_1 : F_n(\epsilon) \neq F_0(\epsilon)$$

Menggunakan taraf kepercayaan 95% diperoleh $D_{(1-0,05)} = 0,227$. Dan secara visual, pengujian asumsi distribusi normal dapat dideteksi melalui penyebaran titik residual yang berada sekitar sumbu diagonal dari grafik. Hasil pengujian asumsi distribusi normal menggunakan uji *Kolmogorov Smirnov* yang ditunjukkan pada Gambar 1.



Gambar 1. Hasil Uji Kolmogorov Smirnov dan Plot Normalitas Residual.

Titik plot residual yang ditunjukkan pada Gambar 1 menunjukkan penyebaran titik plot residual berada pada sumbu diagonal dari grafik, sehingga hal ini mengindikasikan residual berdistribusi normal. Hasil pengujian *Kolmogorov Smirnov* diperoleh $D = 0,125$. Karena nilai statistik uji $D < D_{(0,05)}$ menghasilkan keputusan Gagal Tolak H_0 . Sehingga dapat ditarik kesimpulan bahwa residual mengikuti pola distribusi normal, maka model telah memenuhi asumsi residual distribusi normal.

E. Koefisien Determinasi (R^2)

Nilai koefisien determinasi menunjukkan seberapa baiknya model yang digunakan. Perhitungan nilai koefisien determinasi R^2 diperoleh sebagai berikut.

$$R^2 = \frac{SS_{Regrasi}}{SS_{Total}} \times 100\% = 89,49\%$$

Berdasarkan hasil perhitungan koefisien determinasi (R^2) diperoleh sama dengan 89,49%. Artinya, model regresi nonparametrik nonparametrik *spline truncate* dengan titik kombinasi (3,3,2,2,1) mampu menjelaskan variabilitas persentase kasus HIV di Indonesia sama dengan 89,49%. Atau, persentase kasus HIV di Indonesia dapat dijelaskan kelima variabel prediktor yang digunakan dalam penelitian ini, sama dengan 89,49%. Sedangkan sisanya dapat dijelaskan oleh variabel-variabel lain yang tidak terdapat dalam model.

F. Interpretasi Model

Seperti yang sudah dibahas model telah memenuhi ketiga asumsi residual. Sehingga model regresi nonparametrik *spline truncated* menggunakan titik kombinasi 3, 3, 2, 2, 1 sudah layak digunakan untuk memodelkan persentase kasus HIV di Indonesia. Dengan semua variabel merupakan komponen nonparametrik. Interpretasi dari model tersebut adalah sebagai berikut.

1. Mengasumsikan semua variabel selain variabel X_1 konstan, maka persamaan regresi dari rasio layanan KT (X_1) terhadap persentase kasus HIV di Indonesia (Y) adalah sebagai berikut.

$$\hat{y} = \begin{cases} -56,33 + 2,27057x_1 & ; x_1 < 2,22 \\ -48,87 - 1,09395x_1 & ; 2,22 \leq x_1 < 6,63 \\ -58,83 + 0,40863x_1 & ; 6,63 \leq x_1 < 8,84 \\ -55,12 - 0,01205x_1 & ; x_1 \geq 8,84 \end{cases}$$

Pada interval $x_1 < 2,22$ menunjukkan bahwa apabila rasio layanan KT mengalami kenaikan rasio layanan KT sebesar satu satuan akan meningkatkan kasus HIV sebanyak 23 kasus per 1000 klien yang mengikuti tes. Provinsi yang termasuk ke dalam interval $x_1 < 2,22$ adalah Provinsi DKI Jakarta dan Jawa Timur.

Rasio layanan KT di suatu provinsi berada pada interval $2,22 \leq x_1 < 6,63$ maka setiap kenaikan rasio layanan KT sebesar satu satuan mengakibatkan kasus HIV turun sebanyak 11 kasus per 1000 klien yang mengikuti tes. Provinsi yang tergolong dalam interval ini adalah Kepulauan Riau, Bali, Nusa Tenggara Timur, Kalimantan Selatan, Sulawesi Utara, dan Papua.

Dan provinsi lainnya yang berada pada interval $x_1 \geq 8,84$ menunjukkan bahwa apabila rasio layanan KT mengalami kenaikan sebesar satu satuan mengakibatkan kasus HIV turun sebanyak 1 kasus per 1000 klien yang mengikuti tes.



Gambar 2. Peta Persebaran Provinsi Berdasarkan Rasio Layanan KT

Keterangan:

- : $x_1 < 2,22$
- : $2,22 \leq x_1 < 6,63$
- : $6,63 \leq x_1 < 8,84$
- : $x_1 \geq 8,84$

2. Semua variabel selain variabel X_2 diasumsikan konstan. Berikut adalah persamaan persamaan regresi dari persentase penggunaan kondom (X_2) terhadap persentase kasus HIV di Indonesia (Y).

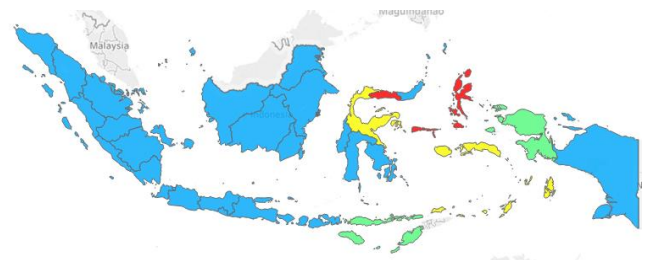
$$\hat{y} = \begin{cases} -56,33 - 11,9156x_2 & ; x_2 < 0,21 \\ -61,55 + 12,4518x_2 & ; 0,21 \leq x_2 < 0,42 \\ -49,25 - 16,673x_2 & ; 0,42 \leq x_2 < 0,53 \\ -58,08 + 0,1018x_2 & ; x_2 \geq 0,53 \end{cases}$$

Pada interval $x_2 < 0,21$ menunjukkan apabila penggunaan kondom disuatu provinsi mengalami kenaikan sebesar 1% akan mengakibatkan kasus HIV di Indonesia turun sebanyak 120 kasus per 1000 klien yang mengikuti tes. Provinsi yang tergolong dalam interval ini adalah Gorontalo, dan Maluku Utara.

Persentase penggunaan kontrasepsi kondom suatu provinsi berada pada interval $0,21 \leq x_2 < 0,42$ maka setiap kenaikan penggunaan kondom sebesar 1% akan meningkatkan kasus HIV sebanyak 125 kasus per 1000 klien yang mengikuti tes. Provinsi yang tergolong dalam interval ini adalah Sulawesi Tengah dan Maluku Utara.

Pada interval $0,42 \leq x_2 < 0,53$ menunjukkan bahwa apabila persentase penggunaan kondom suatu provinsi berada interval tersebut, maka setiap kenaikan penggunaan kondom sebesar 1% akan mengakibatkan kasus HIV turun sebanyak 167 kasus per 1000 klien yang mengikuti tes. Provinsi yang tergolong dalam interval tersebut adalah Nusa Tenggara Barat dan Papua Barat.

Dan provinsi lainnya yang berada pada interval $x_2 \geq 0,53$ apabila persentase penggunaan kondom berada pada interval ini, maka setiap kenaikan persentase penggunaan kondom naik 1% akan mengakibatkan kasus HIV meningkat sebanyak 1 kasus per 1000 klien yang mengikuti tes.



Gambar 3. Peta Persebaran Provinsi Berdasarkan Persentase Penggunaan Kondom

Keterangan:

- : $x_2 < 0,21$
- : $0,21 \leq x_2 < 0,42$
- : $0,42 \leq x_2 < 0,53$
- : $x_2 \geq 0,53$

3. Variabel selain variabel X_3 diasumsikan konstan, maka berikut adalah persamaan regresi dari persentase penduduk miskin terhadap persentase kasus HIV (Y).

$$\hat{y} = \begin{cases} -56,33 + 7,62618x_3 & ; x_3 < 4,27 \\ -18,01 - 1,35117x_3 & ; 4,27 \leq x_3 < 5,74 \\ -25,83 + 0,01247x_3 & ; x_3 \geq 5,74 \end{cases}$$

Pada interval $x_3 < 4,27$ menunjukkan apabila persentase penduduk miskin suatu provinsi $< 4,27\%$, maka setiap kenaikan persentase penduduk miskin sebesar 1% akan meningkatkan kasus HIV sebanyak 77 kasus per 1000 klien yang mengikuti tes. Provinsi yang tergolong dalam interval tersebut adalah DKI Jakarta dan Bali.

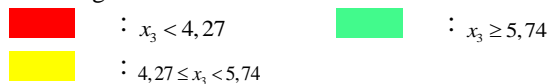
Persentase penduduk miskin suatu provinsi yang berada pada interval $4,27 \leq x_3 < 5,74$ menunjukkan setiap kenaikan persentase penduduk miskin sebesar 1% akan mengakibatkan kasus HIV turun sebanyak 14 kasus per 1000 klien yang mengikuti tes. Provinsi yang tergolong dalam daerah ini adalah Kepulauan Bangka Belitung, Banten, Kalimantan Tengah, dan Kalimantan Selatan.

Dan provinsi lainnya yang berada pada interval $x_3 \geq 5,74$, maka setiap kenaikan persentase penduduk miskin sebesar 1% pada interval tersebut akan mengakibatkan kasus HIV meningkat sebanyak 1 kasus per 1000 klien yang mengikuti tes.



Gambar 4. Peta Persebaran Provinsi Berdasarkan Persentase Penduduk Miskin

Keterangan :



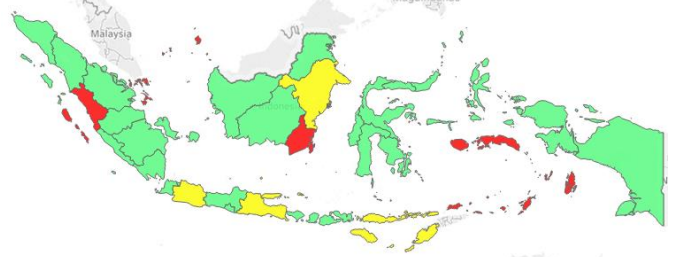
4. Dengan mengasumsi variabel selain variabel X_4 konstan, maka berikut adalah persamaan regresi dari rasio layanan PDP (X_4) terhadap persentase kasus HIV (Y).

$$\hat{y} = \begin{cases} -56,33 + 16,6666x_4 & ; x_4 < 0,6 \\ -43,99 - 3,88733x_4 & ; 0,6 \leq x_4 < 1,09 \\ -48,24 - 0,00422x_4 & ; x_4 \geq 1,09 \end{cases}$$

Pada interval $x_4 < 0,6$ menunjukkan bahwa apabila rasio layanan mengalami sebesar satu satuan akan meningkatkan kasus HIV sebanyak 167 kasus per 1000 klien yang berkunjung. Provinsi yang termasuk dalam interval ini adalah Sumatera Barat, Kepulauan Riau, Kalimantan Selatan, dan Maluku.

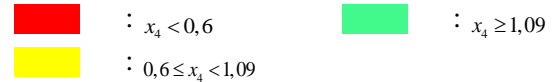
Pada interval $0,6 \leq x_4 < 1,09$ menunjukkan bahwa jika rasio layanan PDP berada pada interval tersebut maka setiap kenaikan rasio layanan PDP sebesar satu satuan akan mengakibatkan kasus HIV turun sebesar 39 kasus per 1000 klien yang mengikuti tes. Provinsi yang berada pada interval tersebut adalah provinsi Jawa Barat, Jawa Timur, Nusa Tenggara Timur, dan Kalimantan Timur. Dan

provinsi lainnya yang berada pada interval $x_4 \geq 1,09$, maka setiap kenaikan rasio layanan PDP sebesar satu satuan akan mengakibatkan kasus HIV turun sebanyak 1 kasus per 1000 klien yang mengikuti tes.



Gambar 5. Peta Persebaran Provinsi Berdasarkan Rasio Layanan PDP

Keterangan :

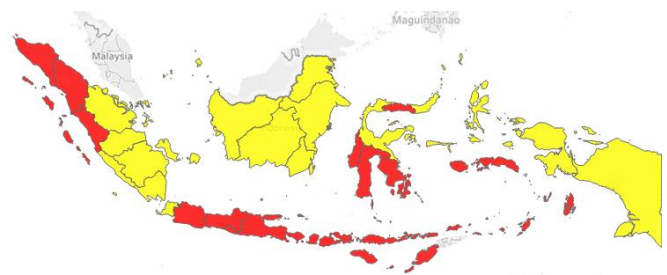


5. Mengasumsikan semua variabel (X_1, X_2, X_3 , dan X_4) selain variabel X_5 konstan, maka persamaan regresi dari persentase penduduk laki-laki (X_5) terhadap persentase kasus HIV di Indonesia (Y) adalah sebagai berikut.

$$\hat{y} = \begin{cases} -56,33 + 0,39855x_5 & ; x_5 < 50,66 \\ -16,55 - 0,38686x_5 & ; x_5 \geq 50,66 \end{cases}$$

Pada interval $x_5 < 50,66$ menunjukkan bahwa apabila persentase penduduk laki-laki mengalami kenaikan penduduk laki-laki sebesar 1% akan meningkatkan kasus HIV sebanyak 4 kasus per 1000 klien yang mengikuti tes. Provinsi yang termasuk ke dalam interval $x_5 < 50,66$ ada 16 provinsi yaitu Aceh, Sumatera Utara, Sumatera Barat, DKI Jakarta, Jawa Barat, Jawa Tengah, Yogyakarta, Jawa Timur, Bali, Nusa Tenggara Timur, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, dan Maluku.

Delapan belas provinsi lainnya yang berada pada interval $x_5 \geq 50,66$ menunjukkan bahwa apabila persentase penduduk laki-laki di suatu provinsi mengalami kenaikan sebesar 1% akan mengakibatkan kasus HIV turun sebanyak 4 kasus per 1000 klien yang berkunjung.



Gambar 6. Peta Persebaran Provinsi Berdasarkan Persentase Penduduk Laki-laki

Keterangan :



V. KESIMPULAN DAN SARAN

Berdasarkan analisis yang telah dilakukan, diperoleh kesimpulan. Tahun 2017 menunjukkan setiap 100 klien yang berkunjung untuk mengikuti tes rata-rata terdapat 2 orang diantaranya yang dinyatakan positif HIV. Persentase kasus HIV tertinggi berada pada Provinsi Kalimantan Selatan, kemudian Provinsi Papua berada di posisi ketiga tertinggi. Apabila dilakukan penelusuran lebih lanjut terkait kasus HIV ini dan terkait penyebabnya, keadaan ini menunjukkan bahwa perlunya ketersediaan layanan (KT dan PDP) yang memfasilitasi pengobatan dan dukungan bagi seorang yang dinyatakan positif HIV (ODHA), penggunaan kondom, penduduk miskin dan penduduk laki-laki. Model terbaik yang digunakan model Regresi Nonparametrik *Spline Truncated* dengan titik kombinasi 3,3,2,2,1, dengan semua variabel berpengaruh secara signifikan. Nilai koefisien determinasi dari model tersebut sama dengan 89,49%.

Saran yang dapat direkomendasikan untuk penelitian selanjutnya adalah pertimbangan dalam memilih variabel prediktor yang diduga berpengaruh terhadap persentase kasus HIV di Indonesia. Dan mempertimbangkan pola hubungan antara variabel prediktor dengan variabel respon, untuk dapat menentukan metode yang tepat digunakan untuk memodelkan persentase kasus HIV di Indonesia tahun 2017. Sedangkan bagi pemerintah, untuk meningkatkan ketersediaan layanan (layanan KT dan layanan PDP) yang memfasilitasi kebutuhan para ODHA (orang dengan HIV dan AIDS) baik secara

konseling dan pengobatan di setiap provinsi di Indonesia. Kemudian meningkatkan sosialisasi atau kampanye mengenai pentingnya penggunaan kondom untuk menekan resiko penyebaran HIV, dan mengenai bahayanya HIV pada daerah yang memiliki persentase penduduk miskin dan persentase penduduk laki-laki yang besar.

DAFTAR PUSTAKA

- [1] World Health Organization, "Sustainable Development Goals (SDGs)," 2018. [Online]. Available: <https://www.who.int/sdg/en/>.
- [2] Kementerian Kesehatan Republik Indonesia, "Profil Kesehatan Indonesia Tahun 2016," Jakarta, 2017.
- [3] UNAIDS, "UNAIDS Data 2018," Geneva, 2018.
- [4] Komisi Penanggulangan HIV dan AIDS Nasional, "Strategi dan Rencana Aksi Nasional 2015-2019 Penanggulangan HIV dan AIDS di Indonesia," Indonesia, 2015.
- [5] H. Oktarian and Budisuari, "Hubungan Antara Karakteristik Responden, Keadaan Wilayah dengan Pengetahuan, Sikap Terhadap HIV/AIDS pada Masyarakat Indonesia," Surabaya, 2009.
- [6] W. Hardle, *Applied Nonparametric Regression*. New York: Cambridge University Press, 1994.
- [7] G. Wahba, *Spline Models for Observation Data. Dalam CBMS-NSF Regional Conference Series in Applied Mathematic*. Philadelphia: SIAM, 1990.
- [8] R. Eubank, *Nonparametric Regression and Spline Smoothing*, 2nd ed. Texas: Department of Statistics Southern Methodist Dallas University, 1999.
- [9] D. Gujarati, *Basic Econometrics*, 4th ed. New York: Mc Graw Hill Companies, 2004.
- [10] N. Draper and H. Smith, *Applied Regression Analysis*. New York: John Wiley & Sons, Inc, 1998.