

# **Spline *Truncated* Multivariabel pada Permodelan Nilai Ujian Nasional di Kabupaten Lombok Barat**

**Nurul Fitriyani**

Program Studi Matematika FMIPA Universitas Mataram, Mataram  
e-mail: [nurul.fitriyani@unram.ac.id](mailto:nurul.fitriyani@unram.ac.id)

**Lailia Awalushaumi**

Program Studi Matematika FMIPA Universitas Mataram, Mataram  
e-mail: [awalushaumi@unram.ac.id](mailto:awalushaumi@unram.ac.id)

**Agus Kurnia**

Program Studi Matematika FMIPA Universitas Mataram, Mataram  
e-mail: [aguskurnia@unram.ac.id](mailto:aguskurnia@unram.ac.id)

**Abstract.** *Regression model is used to analyze the relationship between dependent variable and independent variable. If the regression curve form is not known, then the regression curve estimation can be done by nonparametric regression approach. This study aimed to investigate the relationship between the value resulted by National Examination and the factors that influence it. The statistical analysis used was multivariable truncated spline, in order to analyze the relationship between variables. The research that has been done showed that the best model obtained by using three knot points. This model produced a minimum GCV value of 44.46 and the value of determination coefficient of 58.627%. The parameter test showed that all factors used were significantly influence the National Examination Score for Senior High School students in West Lombok Regency year 2017. The variables were as follows: National Examination Score of Junior High School; School or Madrasah Examination Score; the value of Student's Report Card; Student's House Distance to School; and Number of Student's Siblings.*

**Keywords:** *National Examination, Nonparametric Regression, Truncated Spline.*

## **1. Pendahuluan**

Pola hubungan antar variabel dapat dijelaskan melalui kurva regresi, yang estimasinya biasa didekati dengan pendekatan regresi parametrik dan nonparametrik. Apabila bentuk kurva regresi diketahui, maka dilakukan estimasi kurva regresi dengan pendekatan regresi parametrik dilakukan. Namun apabila bentuk kurva regresi tidak diketahui, maka estimasi kurva regresi dilakukan dengan pendekatan regresi nonparametrik [1].

Pada regresi parametrik, bentuk kurva diasumsikan diketahui, seperti kurva berbentuk linear, kuadrat, kubik, atau lainnya. Berbeda dengan regresi nonparametrik yang tidak mengasumsikan bentuk kurva regresi, sehingga regresi nonparametrik memiliki fleksibilitas tinggi dalam estimasi kurva regresi. Dalam pendekatan dengan regresi nonparametrik, data diharapkan mencari sendiri estimasi kurva regresi, tanpa adanya pengaruh subyektifitas peneliti [2]. Faktanya tidak semua model data dapat diestimasi dengan pendekatan regresi parametrik, dikarenakan kurangnya informasi mengenai bentuk kurva regresi. Pada kasus ini, dapat digunakan pendekatan regresi nonparametrik [3].

Terdapat beberapa model pendekatan dalam mengestimasi kurva regresi nonparametrik, seperti Kernel, Histogram, Deret Orthogonal, Deret Fourier, Wavelets, dan Spline [4]. Diantara pendekatan-pendekatan tersebut, pendekatan spline merupakan satu-satunya pendekatan dengan tata cara optimasi yang sangat baik. Optimasi pada spline diperoleh melalui perluasan dari optimasi-optimasi yang ada pada regresi parametrik. Tata cara optimasi seperti dalam spline ini tidak terdapat pada pendekatan regresi nonparametrik yang lain [5].

Spline dapat memodelkan data dengan pola berubah-ubah pada sub-sub interval tertentu, dikarenakan spline merupakan salah satu jenis potongan polinomial dengan sifat tersegmen. Sifat tersegmen memungkinkan adanya penyesuaian secara lebih efektif terhadap karakteristik lokal suatu data [6]. Seperti yang telah ditunjukkan oleh [7], spline memiliki sifat-sifat statistik yang berguna untuk menganalisis hubungan dalam regresi. Selanjutnya, regresi dengan satu variabel dependen dan lebih dari satu variabel independen disebut sebagai regresi spline multivariabel [4].

Pendidikan merupakan salah satu agenda utama dalam merancang pembangunan negara [8]. Pendidikan merupakan salah satu investasi jangka panjang yang mutunya harus selalu ditingkatkan. Mutu pendidikan yang rendah akan berdampak pada tidak tepatnya investasi pendidikan. Salah satu cara peningkatan mutu atau kualitas pendidikan adalah dengan memperhatikan penilaian hasil belajar, salah satunya adalah Ujian Nasional.

Dengan diperolehnya gambaran mengenai faktor-faktor yang dapat mempengaruhi nilai Ujian Nasional di Kabupaten Lombok Barat khususnya, diharapkan selanjutnya dapat diambil langkah dan kebijakan oleh pihak-pihak terkait, dalam rangka meningkatkan pendidikan di Provinsi NTB umumnya, sehingga dapat sejajar dengan wilayah lain di Indonesia. Hal ini kemudian yang menjadi dasar dilakukannya pemodelan terhadap nilai Ujian Nasional siswa Sekolah Menengah Atas (SMA)/ MA sederajat, serta untuk mengetahui faktor-faktor yang secara signifikan mempengaruhi nilai Ujian Nasional.

Penelitian ini melakukan analisis mengenai hubungan antara nilai Ujian Nasional siswa dengan faktor-faktor yang diduga memberikan pengaruh. Selanjutnya, apabila

hubungan tersebut dapat diketahui, maka langkah-langkah selanjutnya yang dapat diambil oleh pihak-pihak terkait dapat menjadi lebih terarah. Oleh karena itu, dibutuhkan suatu analisis statistika, dalam hal ini dengan menggunakan pendekatan regresi spline *truncated* multivariabel, untuk menggambarkan pola hubungan antara nilai Ujian Nasional dengan faktor-faktor yang mempengaruhinya.

## 2. Tinjauan Pustaka

Salah satu analisis di dalam statistika yang dapat digunakan dalam mengestimasi pola hubungan antar variabel adalah analisis regresi. Hubungan antar variabel tersebut dapat dijelaskan melalui kurva regresi, yang estimasinya biasa didekati dengan pendekatan regresi parametrik dan nonparametrik. Apabila bentuk kurva regresi diketahui, maka dilakukan estimasi kurva regresi dengan pendekatan regresi parametrik dilakukan [2]. Namun apabila bentuk kurva regresi tidak diketahui, maka estimasi kurva regresi dilakukan dengan pendekatan regresi nonparametrik [5].

Jika  $x$  adalah variabel independen dan  $y$  adalah variabel dependen, maka terdapat hubungan fungsional antara  $x$  dan  $y$ , dengan penjabaran berikut.

$$y = f(x) + \varepsilon$$

dimana  $\varepsilon$  adalah *error* acak yang diasumsikan mengikuti distribusi normal [9].

Dalam analisis regresi nonparametrik, pola hubungan data tidak diketahui, sehingga informasi mengenai bentuk kurva regresi tidak diketahui. Data seolah-olah akan mencari bentuk estimasi dari kurva regresinya sendiri, sehingga pendekatan ini menjadi lebih fleksibel tanpa adanya subyektivitas peneliti. Berdasarkan kenyataan tersebut, maka secara visual tidak terdapat bentuk pola tertentu antar variabel. Terdapat banyak pendekatan model regresi nonparametrik yang telah dikembangkan seperti spline, Kernel, Wavelets, dan lain sebagainya, dengan model umum berikut.

$$y_i = g(x_i) + \varepsilon_i, \quad i = 1, 2, \dots, n$$

dengan  $g(x_i)$  fungsi regresi nonparametrik yang tidak diketahui bentuknya dan  $\varepsilon_i$  *error* acak yang diasumsikan berdistribusi normal independen dengan rata-rata nol dan variansi  $\sigma^2$  [2].

Salah satu pendekatan dalam regresi nonparametrik adalah pendekatan spline. Pendekatan ini memiliki sifat fleksibilitas yang tinggi dan kemampuan untuk melakukan estimasi perilaku data yang cenderung berbeda pada interval yang berlainan [2]. Kemampuan mengestimasi perilaku data ini ditunjukkan oleh fungsi *truncated* yang melekat pada estimator. Selanjutnya, titik knot merupakan titik perpaduan bersama yang menunjukkan perubahan pola perilaku fungsi pada selang yang berbeda. Spline merupakan salah satu jenis *piecewise* polinomial atau polinomial tersegmen. Sifat ini memberikan fleksibilitas lebih dari polinomial biasa, sehingga memungkinkan untuk

menyesuaikan diri secara lebih efektif terhadap karakteristik lokal suatu fungsi atau data [3].

Secara umum, fungsi spline berderajat  $m$  dapat disajikan dalam bentuk berikut.

$$g(x_i) = \sum_{j=0}^m \beta_j x_i^j + \sum_{j=1}^J \beta_{j+m} (x_i - k_j)_+^m$$

dengan  $\beta_j$  adalah konstanta riil, dan fungsi potongan:

$$(x_i - k_j)_+^m = \begin{cases} (x_i - k_j)^m & ; x \geq k_j \\ 0 & ; x < k_j \end{cases}$$

Apabila bilangan  $m = 1, 2,$  dan  $3,$  maka diperoleh berturut-turut spline linear, spline kuadrat, dan spline kubik serta  $k_j$  adalah titik knot.

Pendekatan regresi nonparametrik dengan teknik estimator spline ini bertujuan untuk memperoleh kurva dengan titik knot yang optimal. Untuk tujuan memilih parameter optimal ini, telah dikembangkan beberapa metode dalam regresi nonparametrik, salah satunya adalah metode *Generalized Cross Validation* atau *GCV*. Metode ini merupakan metode populer dalam memilih titik knot yang optimal. Pemilihan titik knot  $k$  yang optimal dari *GCV* dapat diperoleh dengan meminimumkan fungsi *GCV* terhadap titik knot [10].

Sebagai syarat pada analisis regresi untuk kelayakan model dalam menggambarkan data yang sebenarnya, dibutuhkan pengecekan terhadap asumsi residual yang diperoleh. Asumsi residual yang harus dipenuhi adalah residual yang berdistribusi normal, identik, dan independen. Pengujian asumsi kenormalan residual dilakukan untuk mengetahui apakah residual telah memenuhi asumsi berdistribusi normal atau tidak. Pengujian residual berdistribusi normal dapat dilakukan dengan pengujian *Kolmogorov Smirnov*. Selanjutnya, residual dikatakan identik apabila residual memiliki variansi yang homogen. Untuk melihat apakah residual telah memiliki variansi yang homogen, dapat dilakukan dengan menggunakan uji *Glejser*. Apabila terjadi heteroskedastisitas, atau variansi residual tidak homogen, maka estimasi koefisien menjadi kurang akurat atau tidak efisien. Pengujian lain yang harus dilakukan adalah pengujian asumsi residual independen. Pengujian ini dilakukan dengan tujuan untuk mengetahui korelasi antar residual apakah sama dengan nol atau tidak. Pemeriksaan asumsi residual independen dilakukan untuk mengetahui apakah korelasi antar residual sama dengan nol [11]. Uji independen dapat dilakukan dengan cara membuat plot *Auto Correlation Function* dari residual [12].

Kriteria yang sering digunakan adalah pemilihan model terbaik adalah dengan menggunakan  $R^2$  dan *Mean Square Error* (MSE). Secara umum semakin kecil nilai MSE, maka semakin baik pula model yang didapatkan. Sebaliknya model yang didapatkan akan semakin baik untuk nilai  $R^2$  yang besar [9].

### 3. Metode Penelitian

Penelitian ini menggunakan data primer mengenai Rata-rata Nilai Ujian Nasional Siswa Sekolah Menengah Atas (SMA)/ MA sederajat yang berada di Kabupaten Lombok Barat, Nusa Tenggara Barat, Tahun Ajaran 2016/ 2017, beserta faktor-faktor yang mempengaruhinya. Variabel yang digunakan di dalam penelitian ini terdiri atas variabel dependen dan independennya. Variabel dependen ( $Y$ ) yang digunakan adalah Rata-rata Nilai Ujian Nasional Siswa, dengan variabel independen, yaitu: Rata-rata Nilai Ujian Nasional SMP Siswa ( $X_1$ ); Rata-rata Nilai Ujian Sekolah atau Ujian Madrasah Siswa ( $X_2$ ); Rata-rata Nilai Rapor Siswa ( $X_3$ ); Jarak yang Ditempuh Siswa ke Sekolah ( $X_4$ ); dan Jumlah Saudara Siswa ( $X_5$ ).

Selanjutnya, langkah-langkah analisis yang dilakukan dalam penelitian ini antara lain: (1) membuat *scatter plot* antar variabel; (2) memodelkan data dengan beberapa jenis titik knot; (3) menghitung nilai GCV untuk masing-masing model; (4) menentukan titik knot dan orde knot optimal berdasarkan nilai GCV minimum; (5) melakukan pengujian signifikansi parameter; (6) melakukan diagnostik residual; (7) membandingkan nilai  $R^2$  dan MSE estimasi model; dan (8) mengambil kesimpulan.

### 4. Hasil dan Pembahasan

Salah satu cara peningkatan mutu atau kualitas pendidikan adalah melalui penilaian hasil belajar. Salah satu bentuk penilaian hasil belajar adalah adanya Ujian Nasional. Hal ini kemudian yang menjadi dasar dilakukannya pemodelan terhadap nilai Ujian Nasional siswa Sekolah Menengah Atas (SMA)/ MA sederajat, serta untuk mengetahui faktor-faktor yang secara signifikan mempengaruhi nilai Ujian Nasional. Penelitian ini mengaplikasikan regresi nonparametrik spline untuk memodelkan nilai Ujian Nasional dengan variabel-variabel independennya.

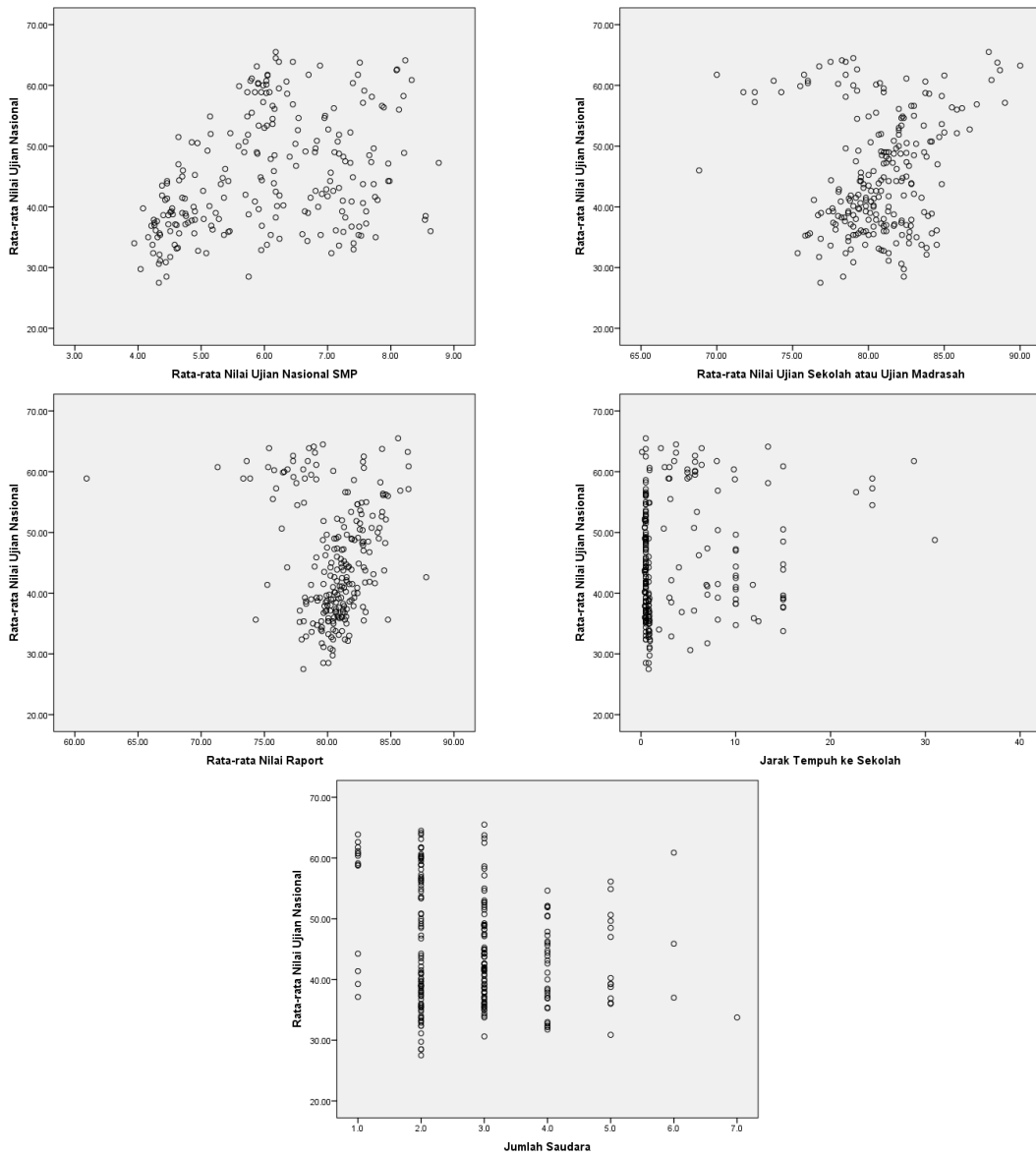
Tabel 4.1 Deskripsi Data Nilai Ujian Nasional Siswa dan Variabel Lainnya

Variabel	N	Range	Minimum	Maksimum	Rata-rata	Variansi
$Y$	230	38.00	27.50	65.50	45.16	89.97
$X_1$	230	4.82	3.94	8.76	6.04	1.48
$X_2$	230	21.17	68.83	90.00	80.69	9.43
$X_3$	230	26.87	60.93	87.80	80.64	7.57
$X_4$	230	31.00	0.10	31.00	3.70	31.40
$X_5$	230	6.00	1.00	7.00	2.81	1.16

Tabel 4.1 menunjukkan ringkasan data dari Rata-rata Nilai Ujian Nasional Siswa Sekolah Menengah Atas (SMA)/ MA sederajat di Kabupaten Lombok Barat, beserta variabel-variabel lainnya.

Tabel 4.1 menunjukkan deskripsi masing-masing variabel, meliputi jumlah data sampel ( $N$ ), *range*, nilai minimum dan maksimum data, serta nilai rata-rata dan variansi data yang digunakan.

Selanjutnya, berikut diberikan penyebaran data yang disajikan dalam bentuk *scatter plot*. Pola hubungan tersebut dapat memudahkan peneliti dalam menentukan model regresi yang akan digunakan. Gambar 4.1 berikut menunjukkan pola hubungan antara Rata-rata Nilai Ujian Nasional ( $Y$ ). siswa SMA/ MA sederajat di Kabupaten Lombok Barat dengan masing-masing variabel independennya.



Gambar 4.1. *Scatter Plot* antara Variabel Respon dan Variabel-variabel Prediktornya

*Scatter plot* yang disajikan pada Gambar 4.1 mengidentifikasi bahwa hubungan antara Rata-rata Nilai Ujian Nasional ( $Y$ ) Siswa SMA/ MA sederajat di Kabupaten Lombok Barat dengan masing-masing variabel independennya tidak mengikuti pola tertentu. Oleh karena itu, perlu dilakukan permodelan pola data dengan menggunakan regresi nonparametrik spline.

Sebagai bahan pertimbangan lain, peneliti terlebih dahulu memodelkan data dengan menggunakan pendekatan parametrik dengan menggunakan model regresi linear berganda untuk Rata-rata Nilai Ujian Nasional ( $Y$ ). Permodelan dengan regresi linear berganda ini digunakan dengan asumsi bahwa pola hubungan antara variabel dependen ( $Y$ ) dengan masing-masing variabel-variabel independennya ( $X$ ) telah diketahui. Permodelan ini dilakukan dengan tujuan untuk lebih meyakinkan bahwa regresi nonparametrik memberikan hasil yang lebih baik untuk memodelkan data tersebut.

Pemodelan menggunakan regresi linear berganda tersebut menghasilkan nilai koefisien determinasi  $R^2$  sebesar 24.80 % dengan nilai *Mean Square Error* (MSE) sebesar 69.18. Nilai  $R^2$  yang relatif rendah ini menunjukkan hasil yang kurang baik, sehingga metode parametrik kurang baik dalam memodelkan Rata-rata Nilai Ujian Nasional ( $Y$ ) siswa SMA/ MA sederajat di Kabupaten Lombok Barat dengan variabel-variabel independen yang diteliti. Oleh karena itu, pendekatan nonparametrik merupakan salah satu solusi yang lebih baik untuk digunakan. Dalam hal ini, penelitian menggunakan fungsi spline *truncated* linear.

Pemilihan model regresi spline *truncated* linear terbaik dilakukan dengan berdasarkan pada titik knot optimal, sesuai dengan lokasi titik knot serta jumlah titik knot yang menghasilkan nilai *Generalized Cross Validation* (GCV) minimum. Titik knot optimal ditentukan dengan memperhatikan nilai GCV yang paling minimum. Ringkasan dari keseluruhan pemilihan titik knot optimal yang dilakukan dengan memperhatikan nilai GCV minimum, baik dengan menggunakan satu (1) titik knot, dua (2) titik knot, tiga (3) titik knot, dan empat (4) titik knot, tertera pada Tabel 4.2 berikut.

Tabel 4.2 Nilai *Generalized Cross Validation* (GCV) dan Koefisien Determinasi  $R^2$

Jumlah Titik Knot	Nilai GCV	Nilai $R^2$
1	52.75	46.61 %
2	46.43	55.13 %
<b>3</b>	<b>44.46</b>	<b>58.62 %</b>
4	44.63	60.42 %

Tabel 4.2 menunjukkan bahwa nilai GCV minimum sebesar 44.46 diperoleh pada model regresi nonparametrik spline *truncated* linear dengan menggunakan tiga (3) buah titik knot. Tiga (3) titik knot bagi masing-masing variabel independen (X) tersebut merupakan titik-titik knot optimal yang selanjutnya digunakan dalam pemodelan Rata-rata Nilai Ujian Nasional (Y) siswa SMA/ MA sederajat di Kabupaten Lombok Barat.

Selanjutnya, estimasi parameter-parameter model spline *truncated* linear dengan menggunakan tiga (3) titik knot dilakukan dengan menggunakan metode kuadrat terkecil. Diperoleh model spline *truncated* linear dengan menggunakan tiga (3) titik knot berikut.

$$\begin{aligned} \hat{y}_i = & \hat{\beta}_0 + \hat{\beta}_{11}x_{i1} + \hat{\beta}_{12}x_{i2} + \hat{\beta}_{13}x_{i3} + \hat{\beta}_{14}x_{i4} + \hat{\beta}_{15}x_{i5} + \hat{\beta}_{21}(x_{i1} - 4.928718)_+^1 \\ & + \hat{\beta}_{22}(x_{i2} - 73.17256)_+^1 + \hat{\beta}_{23}(x_{i3} - 66.44179)_+^1 + \hat{\beta}_{24}(x_{i4} - 6.438462)_+^1 \\ & + \hat{\beta}_{25}(x_{i5} - 2.230769)_+^1 + \hat{\beta}_{31}(x_{i1} - 5.793846)_+^1 + \hat{\beta}_{32}(x_{i2} - 76.97231)_+^1 \\ & + \hat{\beta}_{33}(x_{i3} - 71.26462)_+^1 + \hat{\beta}_{34}(x_{i4} - 11.98462)_+^1 + \hat{\beta}_{35}(x_{i5} - 3.307692)_+^1 \\ & + \hat{\beta}_{41}(x_{i1} - 7.276923)_+^1 + \hat{\beta}_{42}(x_{i2} - 83.48615)_+^1 + \hat{\beta}_{43}(x_{i3} - 79.53231)_+^1 \\ & + \hat{\beta}_{44}(x_{i4} - 21.49231)_+^1 + \hat{\beta}_{45}(x_{i5} - 5.153846)_+^1 \end{aligned}$$

dengan fungsi *truncated* (potongan) linear,

$$(x_{ip} - K_{kp})_+^1 = \begin{cases} (x_{ip} - K_{kp}), & x \geq K_{kp} \\ 0 & , x < K_{kp} \end{cases}$$

Tabel 4.3 berikut menunjukkan masing-masing nilai estimasi seluruh parameter model spline *truncated* linear dengan tiga (3) titik knot.

Tabel 4.3 Hasil Estimasi Parameter Model Spline *Truncated* Linear

Parameter	Nilai Estimasi	Parameter	Nilai Estimasi
$\beta_0$	0.11234	$\beta_{31}$	1.81689
$\beta_{11}$	6.55477	$\beta_{32}$	-6.03445
$\beta_{12}$	1.42995	$\beta_{33}$	4.31693
$\beta_{13}$	0.78039	$\beta_{34}$	2.20534
$\beta_{14}$	2.88455	$\beta_{35}$	0.05706
$\beta_{15}$	-1.17069	$\beta_{41}$	-0.56602
$\beta_{21}$	3.69946	$\beta_{42}$	-1.63416
$\beta_{22}$	-13.38321	$\beta_{43}$	-7.19088
$\beta_{23}$	6.27789	$\beta_{44}$	6.41302
$\beta_{24}$	-0.00000	$\beta_{45}$	-2.06875
$\beta_{25}$	0.84011		



Berdasarkan nilai estimasi parameter pada Tabel 4.3, serta masing-masing nilai knot yang diperoleh dengan bantuan *software R*, diperoleh model spline *truncated* linear sebagai berikut.

$$\begin{aligned} \hat{y}_i = & 0.11234 + 6.55477x_{i1} + 1.42995x_{i2} + 0.78039x_{i3} + 2.88455x_{i4} \\ & - 1.17069x_{i5} + 3.69946(x_{i1} - 4.928718)_+^1 - 13.38321(x_{i2} - 73.17256)_+^1 \\ & + 6.27789(x_{i3} - 66.44179)_+^1 - 0.00000(x_{i4} - 6.438462)_+^1 \\ & + 0.84011(x_{i5} - 2.230769)_+^1 + 1.81689(x_{i1} - 5.793846)_+^1 \\ & - 6.03445(x_{i2} - 76.97231)_+^1 + 4.31693(x_{i3} - 71.26462)_+^1 \\ & + 2.20534(x_{i4} - 11.98462)_+^1 + 0.05706(x_{i5} - 3.307692)_+^1 \\ & - 0.56602(x_{i1} - 7.276923)_+^1 - 1.62416(x_{i2} - 83.48615)_+^1 \\ & - 7.19088(x_{i3} - 79.53231)_+^1 + 6.41302(x_{i4} - 21.49231)_+^1 \\ & - 2.06875(x_{i5} - 5.153846)_+^1 \end{aligned}$$

Langkah selanjutnya yang dilakukan setelah diperoleh model regresi nonparametrik spline *truncated* linear dengan titik knot optimal adalah melakukan pengujian signifikansi parameter model. Pengujian signifikansi parameter ini dilakukan secara simultan (serentak) dan secara parsial (individu). Pengujian parameter secara simultan menunjukkan bahwa minimal terdapat satu parameter model yang tidak nol, sehingga dapat dikatakan bahwa model yang diperoleh tepat.

Selanjutnya, pengujian parameter model secara parsial (individu) dilakukan dengan tujuan untuk menguji pengaruh masing-masing parameter regresi yang signifikan terhadap model regresi spline *truncated* linear yang diperoleh. Hasil yang diperoleh menunjukkan bahwa terdapat 10 parameter model yang signifikan. Signifikansi parameter menunjukkan bahwa variabel independen ( $X$ ) yang digunakan berpengaruh terhadap variabel dependen ( $Y$ ) yang diteliti. Dalam hal ini, seluruh variabel-variabel independen ( $X_1, X_2, X_3, X_4$ , dan  $X_5$ ) yang digunakan, masing-masing yaitu Rata-rata Nilai Ujian Nasional SMP siswa, Rata-rata Nilai Ujian Sekolah atau Ujian Madrasah siswa, Rata-rata Nilai Rapor siswa, Jarak yang Ditempuh siswa ke sekolah, dan Jumlah Saudara siswa, masing-masing memiliki minimal sebuah parameter yang secara signifikan berpengaruh terhadap Rata-rata Nilai Ujian Nasional ( $Y$ ) siswa SMA/ MA sederajat di Kabupaten Lombok Barat.

Hasil pengujian parameter secara parsial juga menunjukkan bahwa terdapat 11 parameter model yang tidak berpengaruh secara signifikan. Parameter-parameter tersebut tidak perlu dikeluarkan secara langsung dari model, dikarenakan variabel independen yang berpengaruh dapat ditentukan berdasarkan parameter yang signifikan. Dalam [13] dikatakan bahwa, apabila terdapat paling tidak satu parameter yang signifikan untuk setiap variabel independen, maka variabel independen tersebut

dikatakan berpengaruh terhadap variabel dependennya. Oleh karena itu, semua parameter digunakan di dalam model. Sebagai contoh, variabel independen  $X_7$ , memiliki minimal satu parameter yang signifikan diantara empat (4) parameter bagi variabel independen  $X_I$ , yaitu:  $\beta_{11}$ ,  $\beta_{21}$ ,  $\beta_{31}$ , dan  $\beta_{41}$ . Hal tersebut berlaku pula bagi variabel-variabel independen lainnya. Jadi, dapat dikatakan bahwa semua variabel independen ( $X$ ) memiliki berpengaruh yang signifikan terhadap variabel dependen ( $Y$ ).

Selanjutnya, dilakukan pengecekan terhadap asumsi residual yang diperoleh. Asumsi residual yang harus dipenuhi adalah residual yang berdistribusi normal dengan rata-rata nol dan variansi  $\sigma^2$ , identik, dan independen. Hasil residual yang diperoleh menunjukkan bahwa residual data telah memenuhi asumsi distribusi normal, identik, namun tidak independen. Model regresi nonparametrik spline *truncated* linear yang diperoleh dengan menggunakan tiga (3) titik knot memiliki nilai koefisien determinasi  $R^2$  sebesar 58.62%. Hal ini menunjukkan bahwa model yang diperoleh dapat menjelaskan keragaman Rata-rata Nilai Ujian Nasional ( $Y$ ) siswa SMA/ MA sederajat di Kabupaten Lombok Barat sebesar 58.62%, dengan nilai *Mean Square Error* (MSE) sebesar 62.54.

Dengan memperhatikan hasil yang telah diperoleh, rendahnya nilai koefisien determinasi  $R^2$ , serta tidak terpenuhinya salah satu asumsi ini menunjukkan bahwa pendekatan dengan menggunakan regresi nonparametrik spline *truncated* linear tidak menunjukkan hasil terbaik. Namun, permodelan data dengan menggunakan pendekatan regresi nonparametrik spline *truncated* linear ini memberikan hasil yang lebih baik apabila dibandingkan dengan pendekatan parametrik dengan regresi linear berganda yang juga telah dilakukan sebelumnya. Hal ini terlihat dari meningkatnya nilai koefisien determinasi  $R^2$  dan menurunnya nilai *Mean Square Error* (MSE) yang dihasilkan oleh permodelan dengan pendekatan regresi nonparametrik spline *truncated* linear.

## 5. Kesimpulan

Berdasarkan penelitian yang telah dilakukan, dapat ditarik kesimpulan bahwa model regresi nonparametrik spline *truncated* linear untuk memodelkan Rata-rata Nilai Ujian Nasional ( $Y$ ) siswa SMA/ MA sederajat di Kabupaten Lombok Barat terbaik diperoleh dengan menggunakan tiga (3) titik knot. Model ini menghasilkan nilai GCV minimum sebesar 44.46 dan nilai koefisien determinasi  $R^2$  sebesar 58.62%. Pengujian parameter menunjukkan bahwa semua variabel independen ( $X$ ) yang digunakan berpengaruh secara signifikan terhadap Rata-rata Nilai Ujian Nasional ( $Y$ ) siswa SMA/ MA sederajat di Kabupaten Lombok Barat. Variabel-variabel tersebut antara lain Rata-rata Nilai Ujian Nasional SMP siswa ( $X_I$ ), Rata-rata Nilai Ujian Sekolah atau Ujian

Madrasah siswa ( $X_2$ ), Rata-rata Nilai Rapor siswa ( $X_3$ ), Jarak yang Ditempuh siswa ke sekolah ( $X_4$ ), dan Jumlah Saudara siswa ( $X_5$ ).

### Ucapan Terima Kasih

Tim Peneliti menyampaikan ucapan terima kasih kepada Universitas Mataram yang telah mendanai penelitian, yaitu melalui sumber dana PNBK Tahun Anggaran 2017. Tim Peneliti menyampaikan ucapan terima kasih pula kepada pihak-pihak terkait, Dinas Pendidikan Pemuda dan Olahraga Provinsi NTB serta seluruh sekolah sampel penelitian, atas kerjasama yang baik hingga terselesaikannya penelitian ini.

### Daftar Pustaka

- [1] Budiantara, I.N. 2005. "Model Keluarga Spline Polinomial Truncated dalam Regresi Semiparametrik". *Berkala MIPA Institut Teknologi Sepuluh Nopember Surabaya*.
- [2] Eubank, R.L. 1988. *Spline Smoothing and Nonparametric Regression Second Edition*. New York: Marcel Dekker
- [3] Hardle, W.. 1990. *Applied Nonparametric Regression*. New York: Cambridge University Press.
- [4] Budiantara, I.N. 2004. "Konsistensi Estimator Spline Terbobot Berdasarkan Kriteria Integrated Mean Square Error". *Prosiding Konferensi Nasional Matematika, XII Himpunan Matematika Indonesia Jurusan Matematika FMIPA UNUD Bali, ISBN : 979-99592-0-9*.
- [5] Budiantara, I.N. 2006. "Model Spline dengan Knots Optimal". *Jurnal Ilmu Dasar FMIPA Universitas Jember, 7, pp.77-85*.
- [6] Budiantara, I.N. 2007. "Pendugaan Model Fertilitas Wanita di Indonesia dengan Menggunakan Regresi Spline". *Laporan Akhir Pelaksanaan Penelitian Studi Kajian Wanita Tahun Anggaran 2007, LPPM Teknologi Sepuluh Nopember Surabaya*.
- [7] Wahba, G. 1990 "Spline Models For Observational Data". *SIAM, CBMS-NSF Regional Conference Series and Applied Mathematics*.
- [8] Hasmori, A.A., dkk. 2011. "Pendidikan, Kurikulum dan Masyarakat: Satu Integrasi". *Journal of Edupres, Vol. 1 September 2011, pp. 350-356*.
- [9] Draper, N.R. and Smith, H. 1992. *Applied Regression Analysis, Second Edition*, New York: John Wiley and Sons, Inc.
- [10] Green, P.J. and Silverman, B.W. 1994. *Nonparametric Regression and Generalized Linear Model*. London: Chapman & Hall.

- [11] Gujarati, D.. 2004. *Basic Econometric*. New York: The McGraw-Hill Companies.
- [12] Wei, W.W.S. 2006. *Time Series Analysis, Univariate and Multivariate Methods, 2nd Edition*. USA : Pearson Education, Inc.
- [13] Sari, R.S. dan Budiantara, I.N. 2012. “Pemodelan Pengangguran Terbuka di Jawa Timur dengan Menggunakan Pendekatan Regresi Spline Multivariabel”. *Jurnal Sains dan Seni ITS Surabaya, Vol.1 No.1*.