

Pemodelan Spline *Truncated* dalam Regresi Nonparametrik Birespon

Luh Putu Safitri Pratiwi

Program Studi Sistem Informasi

STMIK STIKOM Bali

Jl. Raya Puputan No. 68 Renon – Denpasar – Bali. Tlpn (0361) 244445

e-mail: safitri.pratiwi@yahoo.com

Abstrak

Derajat kesehatan menitikberatkan pada peningkatan kesejahteraan masyarakat. Terdapat dua indikator yang paling menggambarkan tingkat kesejahteraan masyarakat yaitu Angka Kematian Bayi (AKB) dan status gizi. Masih terdapatnya kabupaten yang memiliki AKB dan status gizi buruk balita tinggi tentunya tidak terlepas dari faktor-faktor yang mempengaruhi, sehingga secara tepat dapat dilakukan dengan pemodelan terhadap dua indikator tersebut yaitu dengan pemodelan Spline truncated dalam regresi nonparametrik birespon. Penelitian ini bertujuan untuk mendeskripsikan AKB dan status gizi buruk balita serta faktor – faktor yang diduga mempengaruhinya dan mendapatkan model Spline truncated dalam regresi nonparametrik birespon terbaik melalui hubungan antara variabel yang diduga berpengaruh dengan menggunakan metode GCV. Hasil yang didapat yaitu nilai GCV minimum yang terletak pada model Spline truncated linier satu knot yakni sebesar 78.7909 dengan MSE sebesar 77.4146.

Kata kunci: Spline, GCV, AKB, status gizi

1. Pendahuluan

Derajat kesehatan menitikberatkan pada peningkatan kesejahteraan masyarakat. Tujuan dari derajat kesehatan sebagai tolak ukur keberhasilan program pembangunan dan kesejahteraan nasional suatu masyarakat, bangsa, dan negara. Derajat kesehatan masyarakat digambarkan melalui Angka Kematian Bayi (AKB), Angka Kematian Balita (AKABA), Angka Kematian Ibu (AKI), angka morbiditas beberapa penyakit, dan status gizi balita [4]. Angka kematian bayi (AKB) dan status gizi balita merupakan indikator yang paling menggambarkan tingkat kesejahteraan masyarakat [5].

Masih terdapatnya Kabupaten/kota yang memiliki AKB dan status gizi buruk balita tinggi tentunya tidak terlepas dari faktor-faktor yang mempengaruhi, yang secara tepat dapat dilakukan dengan pemodelan terhadap dua indikator tersebut. Pada umumnya, pemodelan yang dapat dilakukan dengan menggunakan analisis regresi. Namun, tidak semua data yang diperoleh mengikuti pola tertentu sehingga jenis data ini menggunakan pendekatan regresi nonparametrik. Ada beberapa teknik yang dilakukan untuk estimasi dalam regresi nonparametrik yaitu Spline. Regresi Spline mempunyai interpretasi statistik dan interpretasi visual yang sangat khusus dan sangat baik [3], sehingga memiliki keistimewaan dibandingkan regresi lainnya. Spline bisa memodelkan data pada pola data yang berubah-ubah pada sub-sub interval tertentu karena Spline merupakan salah satu jenis potongan polinomial, yaitu polinomial yang memiliki sifat tersegmen.

Bentuk estimator Spline juga dipengaruhi oleh lokasi dan banyaknya titik-titik knot. Pemilihan parameter penghalus optimal dalam regresi Spline pada hakikatnya merupakan pemilihan lokasi titik knot [7]. [2] menyebutkan bentuk estimator Spline sangat dipengaruhi oleh nilai parameter penghalus, jika nilai parameter penghalus sangat kecil maka akan memberikan estimator kurva regresi yang sangat kasar. Salah satu metode untuk memilih parameter penghalus yaitu dengan metode *Generalized Cross Validation* (GCV).

Beberapa kasus dalam analisis regresi banyak dijumpai permasalahan yang tidak dapat diselesaikan dengan analisis regresi sederhana satu respon karena jika menggunakan dua variabel respon pada penelitian, maka harus dilihat nilai korelasi antar variabel. Korelasi antara AKB dan status gizi balita yang cukup tinggi sebesar 0.507 menyebabkan dua variabel tersebut dipergunakan sebagai respon pada penelitian ini. Apabila variabel respon dianalisis secara parsial atau satu-satu korelasi antara variabel respon yang akan diteliti tidak akan menghasilkan model yang optimal. Akibatnya, persoalan regresi harus diselesaikan dengan model regresi birespon.

Beberapa penelitian tentang regresi nonparametrik birespon pernah dilakukan oleh [6] dengan pendekatan deret Fourier. [8] menggunakan regresi nonparametrik birespon Spline pada data persentase penduduk miskin dan pengeluaran perkapita makanan di Jawa Timur dan [1] menggunakan regresi nonparametrik Spline dengan dua variabel prediktor pada kasus tingkat kesejahteraan Indonesia tahun 2009.

Penelitian ini bertujuan untuk melihat karakteristik AKB dan status gizi balita beserta faktor-faktor yang diduga berpengaruh dan memodelkan kedua indikator tersebut dengan menggunakan pendekatan regresi nonparametrik birespon Spline *truncated* dengan metode GCV.

2. Metode Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Dinas Kesehatan Provinsi Bali tahun 2015 Provinsi Bali tahun 2015. Adapun variabel bebas dan variabel respon

- Variabel respon yaitu Angka Kematian Bayi (AKB) (y_1) dan angka gizi buruk balita (y_2)
- Variabel bebasnya yaitu terdiri dari: persentase bayi yang diberi ASI (x_1), dan persentase bayi berat badan lahir rendah (BBLR) (x_2).

Adapun langkah-langkah untuk mencapai tujuan penelitian tersebut sebagai berikut:

- Mendeskripsikan AKB dan angka gizi buruk balita di Bali serta faktor – faktor yang diduga mempengaruhinya dengan membuat plot antara variabel prediktor dengan variabel respon
- Memodelkan data dengan model regresi nonparametrik birespon Spline *truncated* dengan satu titik knot dan dua titik knot
- Menghitung nilai GCV untuk masing-masing model regresi Spline *truncated*.
- Menentukan titik knot dan orde knot optimal berdasarkan nilai GCV minimum.
- Melakukan pengujian signifikansi parameter yang dihasilkan dari estimasi model regresi nonparametrik spline dengan metode GCV minimum.
- Membandingkan nilai MSE estimasi model regresi nonparametrik Spline dengan titik knot optimal menggunakan metode GCV.
- Menginterpretasikan hasil analisis dan mengambil kesimpulan.

3. Hasil dan Pembahasan

3.1 Karakteristik AKB dan Angka Gizi Buruk Balita serta Faktor yang Diduga Mempengaruhi

Karakteristik AKB dan angka gizi buruk balita beserta faktor-faktor yang diduga mempengaruhi di Provinsi Bali meliputi nilai rata-rata, variansi, nilai minimum, dan nilai maksimum, yang ditunjukkan pada Tabel 1.

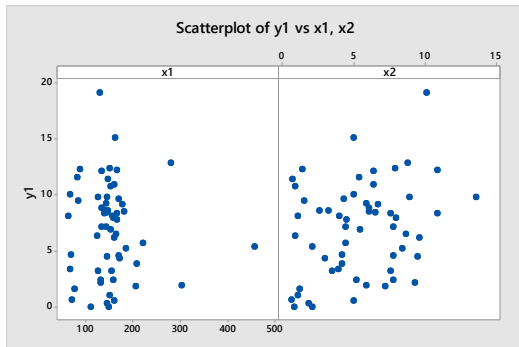
Tabel 1. Karakteristik AKB dan Angka Gizi Buruk Balita Faktor yang diduga Mempengaruhi

Variabel	Rata-rata	Variansi	Minimum	Maksimum
y_1	6.9	17.7	0.0	19.1
y_2	1.8	6.6	0.0	14.0
x_1	152.1	3774.1	63.9	458.4
x_2	5.4	9.9	0.6	13.6

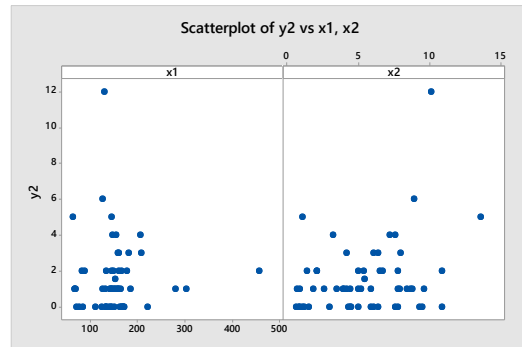
Tabel 1. menunjukkan karakteristik AKB dan angka gizi buruk balita yakni nilai rata – rata, variansi, minimum, dan maksimum. Berdasarkan Tabel 1 terlihat bahwa rata – rata AKB di provinsi Bali yang tersebar pada 57 Kecamatan tiap 1000 penduduk adalah sebanyak 6 atau 7 dengan keragaman sebesar 17.7. Terlihat dari nilai minimum dan maksimum pada Tabel 1. mengindikasikan bahwa AKB di Provinsi Bali tertinggi sebesar 19.1% dan terendah sebesar 0%. Sementara itu, angka gizi buruk balita di Provinsi Bali adalah sebesar 1.8% dengan keragaman sebesar 6.6. Angka gizi buruk balita di Provinsi Bali tertinggi sebesar 14.0% dan terendah sebesar 0%.

3.2 Scatterplot AKB dan Angka Gizi Buruk Balita dengan Faktor yang diduga Mempengaruhi

Pola hubungan yang terbentuk antara AKB dan angka gizi buruk balita serta dua variabel prediktor divisualisasikan pada Gambar 1. dan Gambar 2.



Gambar 1. Scatterplot antara y_1 dengan x_1 dan x_2 ,



Gambar 2. Scatterplot antara y_2 dengan x_1 dan x_2

Gambar 1. dan Gambar 2. menunjukkan pola hubungan yang terbentuk antara y_1 dan y_2 dengan keempat variabel tersebut yang tidak membentuk pola tertentu. Sehingga mengindikasikan bahwa terdapat komponen nonparametrik dimana fungsi dari kurva regresi tidak diketahui, maka pemodelan yang tepat adalah memodelkan dengan regresi nonparametrik birespon dengan estimator yang digunakan adalah Spline *truncated*. Dari variabel tersebut akan dibuat model dari model Spline linier dengan jumlah titik knot satu dan dua knot. Hasilnya kemudian dibandingkan nilai GCV dan dipilih nilai GCV terkecil diantara model yang terbentuk.

3.3 Model Regresi Nonparametrik Birespon Spline *Truncated*

Model regresi nonparametrik birespon Spline *truncated* dengan dua variabel prediktor dapat dituliskan dalam persamaan berikut ini.

$$\hat{y}_1 = \alpha_0^{(1)} + \alpha_{11}^{(1)} x_1 + \hat{\beta}_{11}^{(1)} (x_1 - K_{11}^{(1)})_+ + \dots + \hat{\beta}_{k1}^{(1)} (x_1 - K_{k1}^{(1)})_+ + \alpha_{12}^{(1)} x_1 + \hat{\beta}_{12}^{(1)} (x_2 - K_{12}^{(1)})_+ + \dots + \hat{\beta}_{k2}^{(1)} (x_2 - K_{k2}^{(1)})_+$$

$$\hat{y}_2 = \alpha_0^{(2)} + \alpha_{11}^{(2)} x_1 + \hat{\beta}_{11}^{(2)} (x_1 - K_{11}^{(2)})_+ + \dots + \hat{\beta}_{k1}^{(2)} (x_1 - K_{k1}^{(2)})_+ + \alpha_{12}^{(2)} x_1 + \hat{\beta}_{12}^{(2)} (x_2 - K_{12}^{(2)})_+ + \dots + \hat{\beta}_{k2}^{(2)} (x_2 - K_{k2}^{(2)})_+$$

Setelah membuat model umum dari regresi nonparametrik birespon Spline maka kemudian akan dilakukan pembuatan model dengan pendekatan regresi nonparametrik birespon Spline dengan melakukan pemilihan titik knot optimal. Dalam penelitian ini dilakukan dengan menggunakan satu titik knot dan dua titik knot.

3.4 Pemilihan Titik Knot Optimal

Pemilihan titik knot optimal dilakukan dengan mencari nilai GCV terendah yang dihasilkan. GCV yang dihasilkan dengan menggunakan satu titik knot dan dua titik knot, ditunjukkan pada Tabel 2.

Tabel 2. Nilai GCV Minimum dan MSE Masing-masing Titik Knot

Variabel Prediktor	GCV Minimum	MSE
1 Titik knot	78.7909	77.4146
2 Titik knot	82.3998	80.9605

Tabel 2 menunjukkan bahwa nilai GCV minimum dihasilkan pada saat menggunakan knot satu yakni sebesar 78.7909 dengan MSE sebesar 77.4146.

3.5 Pemodelan AKB dan Angka Gizi Buruk Balita dengan Titik Knot Optimal

Nilai GCV minimum dihasilkan dengan menggunakan knot satu yakni sebesar 0,004. Pemodelan AKB dan angka status gizi buruk balita menggunakan titik knot optimal ditunjukkan pada persamaan berikut.

Untuk variabel respon pertama yaitu AKB sebagai berikut.

$$\hat{y}_1 = 0.0004 + 0.0482x_1 - 0.0107(x_1 - 176.61)_+ + 0.0029x_2 + 0.0026(x_2 - 0.60)_+$$

Untuk variabel respon kedua angka gizi buruk balita sebagai berikut.

$$\hat{y}_2 = 0.0007 + 0.0802x_1 - 0.0209(x_1 - 176.61)_+ + 0.0054x_2 + 0.0049(x_2 - 0.60)_+$$

Dari model tersebut maka model derajat kesehatan di Indonesia dapat diinterpretasikan kedalam regresi nonparametrik birespon yaitu:

1. Model yang terbaik yang menjelaskan derajat kesehatan yaitu dengan variabel respon AKB dan angka gizi buruk balita di Indonesia adalah model Spline linier dengan satu titik knot.
2. Pada respon AKB dan respon angka gizi buruk balita, perubahan pola perilaku data pada variabel persentase bayi diberikan ASI terjadi pada titik 176.61, dimana jika nilai variabel tersebut dibawah 176.61 persen maka persentase AKB dan angka gizi buruk balita memiliki pola yang berbeda dengan setelah persentase bayi diberikan ASI bernilai 176.61 persen dan lebih.
3. Pada respon AKB dan respon angka gizi buruk balita, perubahan pola perilaku data pada variabel BBLR terjadi pada titik 0.60, yaitu pola persentase AKB dan angka gizi buruk balita berubah setelah titik 0.60.

4. Simpulan

Berdasarkan hasil dan pembahasan dapat diambil kesimpulan sebagai berikut:

Model yang terbaik yang dapat menggambarkan tingkat derajat kesehatan di Indonesia yaitu AKB dan angka gizi buruk balita adalah dengan model Spline *truncated* linier satu titik knot. Model yang terbentuk yaitu:

$$\hat{y}_1 = 0.0004 + 0.0482x_1 - 0.0107(x_1 - 176.61)_+ + 0.0029x_2 + 0.0026(x_2 - 0.60)_+$$

$$\hat{y}_2 = 0.0007 + 0.0802x_1 - 0.0209(x_1 - 176.61)_+ + 0.0054x_2 + 0.0049(x_2 - 0.60)_+$$

dengan Nilai GCV : 78.7909 dan nilai MSE sebesar: 77.4146.

Saran yang dapat diberikan dalam penelitian ini adalah perlu dilakukan penelitian lebih lanjut untuk menentukan uji statistik dari model yang terbentuk misalnya uji hipotesis dan interval konfidensi serta dalam memodelkan derajat kesehatan di Indonesia, penulis hanya menggunakan dua variabel respon dan dua variabel prediktor, dengan jumlah knot terbanyak hanya dua titik knot sedangkan masih banyak respon dan prediktor yang mungkin mempengaruhi derajat kesehatan di Indonesia serta jumlah knot yang mungkin akan meminimumkan GCV

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada STIKOM Bali yang telah memberi dukungan financial terhadap penelitian dengan judul *Pemodelan Spline Truncated* dalam Regresi Nonparametrik Birespon

Daftar Pustaka

- [1] I Gde Adnyana. Estimator Spline dalam Regresi Nonparametrik Multirespon (Studi Kasus Tingkat Kesejahteraan di Indonesia Tahun 2009). Tesis. Surabaya, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember (ITS); 2011.
- [2] I Nyoman Budiantara. *Metode U, GML, CV dan GCV dalam Regresi Nonparametrik Spline*. Majalah Ilmiah Himpunan Matematika Indonesia (MIHMI). 2000; Vol 6: hal. 41-45.
- [3] I Nyoman Budiantara. *Spline dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang*. Surabaya: ITS Press. 2009.
- [4] Dinkes Bali (Dinas Kesehatan Provinsi Bali). *Profil Kesehatan Provinsi Bali Tahun 2015*. Bali: Pemerintah Provinsi Bali. 2015.
- [5] Depkes RI (Departemen Kesehatan RI). *Indikator Indonesia Sehat 2010 dan Pedoman Penetapan Indikator Provinsi Sehat dan Kabupaten/Kota Sehat: Keputusan Menteri Kesehatan Nomor 1202/Menkes/SK/VIII/2003*. Jakarta: Departemen Kesehatan RI. 2003.
- [6] Rini Semiati. *Regresi Nonparametrik Deret Fourier Birespon*. Tesis. Surabaya, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember (ITS), Surabaya; 2010.
- [7] Agustin Tripena. *Penentuan Model Regresi Spline Terbaik. Prosiding Seminar Nasional Statistika Universitas Diponegoro 2011*. 2011.
- [8] I Dewa Ayu Made Istri Wulandari. *Analisis Faktor – Faktor Yang Mempengaruhi Persentase Penduduk Miskin Dan Pengeluaran Perkapita Makanan Di Jawa Timur Dengan Metode Regresi*

Nonparametrik Birespon Spline. Tugas Akhir. Surabaya, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember (ITS); 20