

Faktor-Faktor yang Mempengaruhi Angka Gizi Buruk Di Jawa Timur dengan Pendekatan Regresi Nonparametrik Spline

Riana Kurnia Dewi, I Nyoman Budiantara

Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember (ITS)

Jl. Arief Rahman Hakim, Surabaya 60111

E-mail: i_nyoman_b@statistika.its.ac.id

Abstrak—Salah satu permasalahan kesehatan di Indonesia adalah meningkatnya angka kematian balita. Salah satu penyebabnya adalah kebutuhan gizi yang tidak terpenuhi sehingga banyak balita mengidap gizi buruk. Pemodelan kejadian balita gizi buruk dengan regresi parametrik belum tentu cocok diterapkan karena pola hubungan antara angka gizi buruk dengan faktor-faktor yang mempengaruhinya memiliki bentuk pola tertentu. Regresi Nonparametrik Spline adalah metode regresi yang tidak memberikan asumsi terhadap bentuk kurva regresi. Penelitian ini bertujuan mengetahui faktor-faktor yang mempengaruhi kejadian balita gizi buruk di Provinsi Jawa Timur. Hasil penelitian menunjukkan bahwa dengan menggunakan Regresi Nonparametrik Spline, diperoleh nilai GCV minimum yaitu 3,943068 dan R^2 sebesar 88,77 persen. Kesimpulan lain diperoleh faktor-faktor yang berpengaruh terhadap kejadian balita gizi buruk di Jawa Timur tahun 2007 adalah persentase ibu yang memeriksakan kehamilan, persentase bayi mendapat vitamin A dan persentase rumah tangga miskin.

Kata Kunci—GCV, Gizi Buruk, R^2 , Regresi Nonparametrik Spline

I. PENDAHULUAN

SALAH satu permasalahan kesehatan di Indonesia adalah meningkatnya angka kematian balita. Salah satu penyebabnya adalah kebutuhan gizi yang tidak terpenuhi. Keadaan gizi balita akan mempengaruhi tingkat kesehatan dan harapan hidup yang merupakan salah satu unsur utama dalam penentuan keberhasilan pembangunan negara yang dikenal dengan istilah human development index (HDI). Status gizi buruk pada balita dapat menimbulkan pengaruh yang sangat menghambat pertumbuhan fisik, mental maupun kemampuan berpikir yang pada akhirnya akan menurunkan produktivitas kerja. Balita penderita gizi buruk dapat mengalami penurunan kecerdasan (IQ) hingga sepuluh persen. Selain itu, penyakit yang dapat diderita balita gizi buruk adalah diabetes (kencing manis) dan penyakit jantung koroner. Dampak paling buruk yang diterima adalah kematian pada umur yang sangat dini Samsul [1].

Menurut data Dinas Kesehatan RI tercatat sekitar 4 persen atau 900 ribu balita yang tersebar di seluruh Indonesia menyandang status gizi buruk. Hal ini mengakibatkan Indonesia menduduki peringkat lima besar pemilik gizi buruk balita Anonim [2]. Di antara provinsi-provinsi di Indonesia pada tahun 2007, posisi Jawa Timur jika dilihat dari tingkat status gizi buruk termasuk ke dalam kelompok menengah

dengan 4,8 persen balita digolongkan gizi buruk. Jawa Timur termasuk ke dalam empat provinsi yang selama 5 tahun berturut-turut (2005-2009) berada pada kategori 10 provinsi dengan kasus balita gizi buruk tertinggi Siswono [3].

Penelitian tentang gizi buruk telah dilakukan oleh beberapa peneliti diantaranya Paramita [4] melakukan klasifikasi terhadap status gizi balita di Kabupaten Nganjuk. Metode yang digunakan adalah bagging regresi logistik ordinal. Hayati [5] melakukan pengelompokan kabupaten/kota di Jawa Timur berdasarkan status gizi buruk balita. Riskiyanti [6] meneliti tentang faktor-faktor yang mempengaruhi derajat kesehatan di Provinsi Jawa Timur. Metode yang digunakan dalam penelitian tersebut adalah analisis regresi multivariat. Ayunin [7] meneliti tentang pemodelan balita gizi buruk di Kabupaten Ngawi menggunakan GWR (*Geographically Weighted Regression*). A'yunin [8] meneliti tentang pemodelan angka gizi buruk di Kota Surabaya menggunakan metode SAR (*Spatial Autoregressive*). Penelitian tentang metode Regresi Spline pernah dilakukan namun pada bidang lain, diantaranya Fridiati [9], Sutarsi [10], Basri [11], dan Federika [12]. Sejauh ini belum ada yang melakukan penelitian tentang angka gizi buruk di Jawa Timur dengan menggunakan Regresi Spline.

Tujuan penelitian ini adalah untuk mendeskripsikan kejadian balita gizi buruk, ibu yang memeriksakan kehamilannya, balita yang memenuhi kecukupan protein, balita yang mendapatkan vitamin A, Rumah Tangga yang mengakses air bersih, BBLR, Rumah Tangga miskin di Provinsi Jawa Timur serta memodelkan persentase gizi buruk di Jawa Timur dengan menggunakan pendekatan Spline. Manfaat penelitian ini bagi instansi pemerintah, dapat dijadikan sebagai tambahan informasi untuk menentukan kebijakan yang akan diambil perihal peningkatan kualitas perbaikan gizi khususnya di Provinsi Jawa Timur serta bagi pembaca dan mahasiswa, dapat dijadikan sebagai pengetahuan mengenai regresi nonparametrik dan aplikasinya dalam suatu permasalahan sosial.

II. TINJAUAN PUSTAKA

A. Regresi Nonparametrik Spline

Regresi nonparametrik merupakan suatu metode statistika yang digunakan untuk mengetahui hubungan antara variabel respon dengan variabel prediktor yang tidak diketahui bentuk fungsinya, hanya diasumsikan *smooth* (mulus) dalam arti termuat dalam suatu ruang fungsi tertentu. Regresi nonparametrik merupakan regresi yang sangat fleksibel dalam memodelkan pola data Eubank [13].

Model regresi nonparametrik secara umum sebagai berikut.

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

Diberikan data $(t_{1i}, t_{2i}, \dots, t_{pi}, y_i)$ dan hubungan antara $(t_{1i}, t_{2i}, \dots, t_{pi})$ dan y_i diasumsikan mengikuti model regresi nonparametrik, $y_i = g(t_{1i}, t_{2i}, \dots, t_{pi}) + \varepsilon_i$ untuk $i = 1, 2, \dots, n$ dengan y_i variabel respon, f kurva regresi yang tidak diketahui bentuknya. Apabila kurva regresi g merupakan model aditif dan dihipir dengan fungsi spline maka diperoleh model regresi

$$y_i = g(t_{1i}) + g(t_{2i}) + \dots + g(t_{pi}) + \varepsilon_i \\ = \sum_{j=1}^p f(t_{ji}) + \varepsilon_i, i = 1, 2, \dots, n, \quad (2)$$

dimana,

$$g_j(t_{ji}) = \sum_{h=1}^q \alpha_{hj} t_{ji}^h + \sum_{l=1}^m \beta_{lj} (t_{ji} - K_{lj})_+^q \\ = \alpha_{1j} t_{ji} + \dots + \alpha_{qj} t_{ji}^q + \beta_{1j} (t_{ji} - K_{1j})_+^q + \dots + \beta_{mj} (t_{ji} - K_{mj})_+^q,$$

$$\text{dengan } (t_{ji} - K_{lj})_+^q = \begin{cases} (t_{ji} - K_{lj})^q, & t_{ji} \geq K_{lj} \\ 0, & t_{ji} < K_{lj} \end{cases}$$

dan $K_{1j}, K_{2j}, \dots, K_{mj}$ adalah titik-titik knot yang memperlihatkan pola perubahan perilaku dari fungsi pada sub-sub interval yang berbeda. Nilai q pada persamaan di atas merupakan derajat dari polinomial. Kurva polinomial derajat satu disebut kurva linear, derajat dua disebut sebagai kurva kuadratik, serta derajat tiga disebut sebagai kurva kubik.

Agar diperoleh spline yang optimal perlu dipilih titik knot yang optimal. Salah satu metode untuk memilih titik knot optimal adalah dengan metode GCV (*Generalized Cross Validation*) Budiantara [14]. Model spline yang sesuai berkaitan dengan titik knot yang optimal didapat dari nilai GCV terkecil. Fungsi GCV didefinisikan sebagai

$$GCV(K_1, K_2, \dots, K_m) = \frac{MSE(K_1, K_2, \dots, K_m)}{(n^{-1} \text{tr}[I - A(K_1, K_2, \dots, K_m)])^2} \quad (3)$$

dimana,

$$MSE(K_1, K_2, \dots, K_m) = n^{-1} \sum_{j=1}^n (y_j - \hat{f}_{(K_1, K_2, \dots, K_m)}(t_j))^2,$$

K_1, K_2, \dots, K_m adalah titik knot dan matriks $A(K_1, K_2, \dots, K_m)$ diperoleh dari persamaan $\hat{y} = A(K_1, K_2, \dots, K_m)y$.

Setelah diperoleh model spline terbaik, selanjutnya akan dilakukan pengujian parameter yang terdiri dari uji serentak dan uji parsial serta pengujian distribusi normal pada residual.

B. Uji Simultan

Hipotesis yang digunakan pada uji simultan adalah sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1 : \text{Minimal ada satu } \beta_p \neq 0, p = 1, 2, \dots, k$$

Statistik uji yang digunakan adalah uji F :

$$F_{hitung} = \frac{MS_{Regresi}}{MS_{Residual}} \quad (4)$$

Tolak H_0 jika $F_{hitung} > F_{\alpha; (k, n-k-1)}$.

C. Uji Individu

Hipotesis yang digunakan pada uji individu adalah sebagai berikut.

$$H_0 : \beta_p = 0$$

$$H_1 : \beta_p \neq 0; p = 1, 2, \dots, k$$

Statistik uji yang digunakan adalah uji t :

$$t_{hitung} = \frac{\hat{\beta}_p}{se(\hat{\beta}_p)} \quad (5)$$

Tolak H_0 jika $|t_{hitung}| > t_{(\alpha/2, n-k)}$, dimana n adalah jumlah pengamatan dan k adalah jumlah parameter.

D. Uji Normalitas Residual

Hipotesis yang digunakan pada uji normalitas residual adalah sebagai berikut :

$$H_0 : \text{Residual mengikuti distribusi normal}$$

$$H_1 : \text{Residual tidak mengikuti distribusi normal}$$

Statistik uji yang digunakan adalah :

$$Z_{hitung} = \text{Sup}_x |F_n(x) - F_0(x)| \quad (6)$$

Tolak H_0 jika $Z_{hitung} > Z_{\alpha}$.

III. METODOLOGI PENELITIAN

A. Data Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari data Riset Kesehatan Dasar (Riskesdas) tahun 2007 dan Survei Sosial Ekonomi Nasional (Susenas) tahun 2007. Riskesdas tahun 2007 mempunyai desain sampling yang sama dengan Susenas tahun 2007 dimana datanya dapat mengestimasi kesehatan masyarakat kabupaten/kota, provinsi, ataupun nasional. Variabel yang diambil dari Riskesdas adalah persentase balita gizi buruk, persentase ibu yang memeriksakan kehamilannya, persentase balita yang mencukupi kebutuhan protein, persentase balita yang mendapatkan vitamin A, persentase Rumah Tangga yang mengakses air bersih dan persentase bayi berat lahir rendah. Sedangkan variabel yang diambil dari Susenas adalah persentase rumah tangga miskin. Lokasi penelitian adalah di Provinsi Jawa Timur yang terdiri atas 38 kabupaten/kota.

B. Metode Analisis Data

Langkah-langkah analisis yang dilakukan pada penelitian ini adalah sebagai berikut.

1. Mendeskripsikan kejadian balita gizi buruk di Jawa Timur.
 - i. Mendeskripsikan kejadian balita gizi buruk, ibu yang memeriksakan kehamilannya, balita yang memenuhi kecukupan protein, balita yang mendapatkan vitamin A, rumah tangga yang mengakses air bersih, BBLR, rumah tangga miskin di Provinsi Jawa Timur.
 - ii. Menginterpretasikan hasil analisis dan mengambil kesimpulan.
2. Memodelkan kejadian balita gizi buruk di Provinsi Jawa Timur dengan pendekatan Spline.
 - i. Membuat *scatterplot* antara variabel respon dengan masing-masing variabel prediktor yang dijadikan

sebagai deteksi awal mengenai pola hubungan antara variabel respon dan variabel prediktor.

- ii. Memodelkan variabel respon dengan menggunakan Spline linear dan berbagai titik knot.
- iii. Menentukan titik-titik knot optimal yang didasarkan pada nilai GCV minimum.
- iv. Menetapkan model Spline terbaik.
- v. Menguji signifikansi parameter secara serentak dan parsial.
- vi. Melakukan uji Normalitas Residual.
- vii. Menginterpretasikan hasil analisis dan mengambil kesimpulan.

IV. ANALISA DAN PEMBAHASAN

A. Statistika Deskriptif

Provinsi Jawa Timur memiliki 29 Kabupaten dan 9 Kota atau secara administratif terdapat 38 Kabupaten/Kota. Setiap kabupaten/kota memiliki kondisi sosial dan ekonomi yang berbeda. Berikut hasil analisa deskriptif dari tiap variabel penelitian.

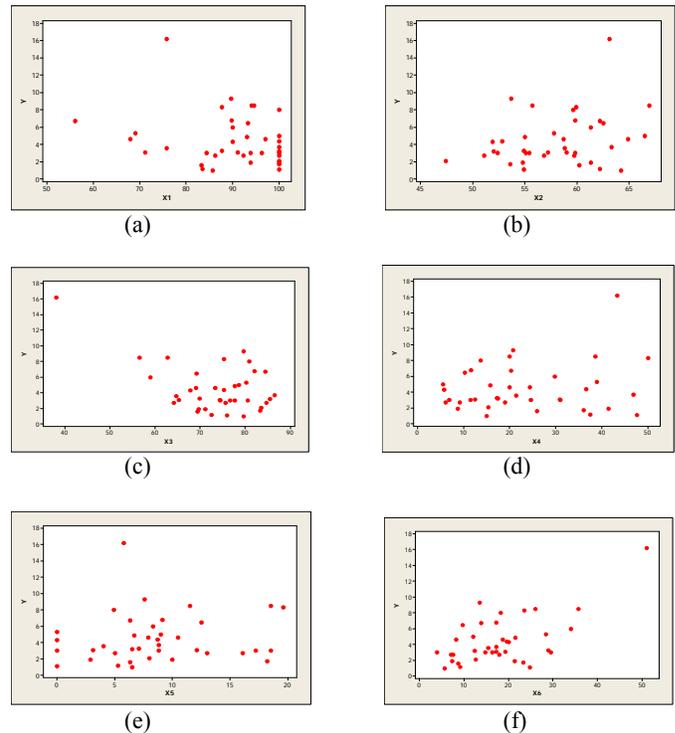
Tabel 1. Statistika Deskriptif

Variabel	Mean	Varians	Minimum	Maksimum
Y	4,49	9,02	1	16,2
X ₁	89,8	116	56,1	100
X ₂	58,1	21,1	47,4	66,9
X ₃	73,5	90,9	38,1	86,5
X ₄	23,3	172	5,5	50
X ₅	8,53	28,5	0	19,6
X ₆	18,19	92,01	3,95	51,02

Dari Tabel 1 diperoleh informasi bahwa pada variabel Y (persentase angka gizi buruk) memiliki nilai tertinggi sebesar 16,2 persen yaitu Kabupaten Sampang. Pada variabel X₁ (persentase ibu yang memeriksakan kehamilannya) memiliki nilai terendah sebesar 56,1 persen yaitu Kabupaten Mojokerto. Pada variabel X₂ (persentase balita yang memenuhi kecukupan protein) memiliki nilai keragaman data sebesar 21,1. Untuk variabel X₃ (persentase balita yang mendapatkan vitamin A) memiliki nilai rata-rata sebesar 73,5 persen. Pada variabel X₄ (persentase rumah tangga yang mengakses air bersih) memiliki nilai terendah sebesar 5,5 persen yaitu Kota Probolinggo. Untuk variabel X₅ (persentase bayi berat lahir rendah) memiliki nilai tertinggi sebesar 19,6 persen yaitu Kabupaten Probolinggo. Sedangkan untuk variabel X₆ (persentase rumah tangga miskin) memiliki nilai tertinggi sebesar 51,02 persen yaitu Kabupaten Sampang.

B. Scatterplot untuk Variabel Respon dengan Variabel Prediktor

Analisis regresi merupakan salah satu metode statistika yang digunakan untuk menyelidiki pola hubungan antara variabel prediktor dengan variabel respon. Bentuk pola hubungan fungsional antara variabel prediktor dengan variabel respon dapat diperkirakan dengan membuat diagram pencar (*scatter plot*) yang memuat informasi tentang kedua hubungan tersebut. Bentuk pola hubungan antara variabel respon dengan variabel prediktor disajikan pada Gambar 1.



Gambar. 1. Scatterplot antara Variabel Respon dengan Variabel Prediktor.

Gambar 1 menunjukkan bahwa pola hubungan antara variabel respon dengan variabel prediktor memiliki pola yang menyebar, sehingga tidak memiliki kecenderungan membentuk suatu pola tertentu.

Pola hubungan antara setiap variabel prediktor terhadap variabel respon memiliki pola yang tidak mengikuti pola tertentu. Oleh karena itu sulit digunakan pemodelan dengan pendekatan regresi parametrik. Selanjutnya pola data akan didekati dengan menggunakan regresi nonparametrik spline.

C. Regresi Nonparametrik Spline Linier 1 Titik Knot

Model regresi spline linier dengan satu titik knot (K) sebagai berikut.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - K_1)_+^1 + \dots + \hat{\beta}_{11} x_6 + \hat{\beta}_{12} (x_6 - K_6)_+^1,$$

Nilai GCV untuk model 1 titik knot disajikan pada Tabel 2.

Tabel 2. Nilai GCV Model Spline Linier Satu Titik Knot

No.	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	GCV
1	83	66	47	27	17	27	7,92
2	90	51	72	46	3	24	7,56
3	58	52	70	44	7	5	7,22
4	96	57	82	20	14	5	9,34
5	79	54	54	48	15	20	6,25
6	84	56	66	48	12	11	5,28
7	93	62	43	13	6	24	7,38
8	79	53	78	41	5	33	7,47
9	61	60	55	35	6	36	8,30
10	73	55	48	38	16	48	7,25

Dari Tabel 2 didapatkan nilai GCV minimum sebesar 5,28 yang bersesuaian dengan knot K₁ = 84, K₂ = 56, K₃ = 66, K₄ = 48, K₅ = 12, K₆ = 11.

D. Regresi Nonparametrik Spline Linier 2 Titik Knot

Model regresi spline linier dua knot adalah sebagai berikut.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - K_1)_+ + \hat{\beta}_3 (x_1 - K_2)_+ + \dots + \hat{\beta}_{16} x_6 + \hat{\beta}_{17} (x_6 - K_{11})_+ + \hat{\beta}_{18} (x_6 - K_{12})_+$$

Nilai GCV untuk model dua titik knot disajikan pada Tabel 3.

Tabel 3.

Nilai GCV Model Spline Linier Dua Titik Knot

No	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	GCV
1	81;93	49;52	59;71	8;20	6;8	15;33	10,17
2	93;95	58;66	53;75	28;33	2;13	18;38	10,60
3	66;70	52;53	44;70	17;18	1;11	6;43	11,68
4	65;72	49;54	69;84	15;20	7;18	28;30	9,33
5	62;94	59;60	50;63	46;49	2;13	27;28	7,86
6	80;95	58;59	58;69	17;33	12;15	15;18	8,83
7	86;98	52;62	61;80	20;30	3;5	15;21	7,64
8	81;92	48;54	79;83	13;36	2;13	26;30	9,99
9	72;87	53;57	51;75	29;32	15;19	32;41	9,61
10	84;85	51;59	52;58	39;43	8;17	7;39	9,16

Tabel 3 menunjukkan bahwa nilai GCV minimum adalah sebesar 7,64 yang bersesuaian dengan knot K₁ = 86, K₂ = 98, K₃ = 52, K₄ = 62, K₅ = 61, K₆ = 80, K₇ = 20, K₈ = 30, K₉ = 3, K₁₀ = 5, K₁₁ = 15, K₁₂ = 21.

E. Regresi Nonparametrik Spline Linier 3 Titik Knot

Model regresi spline linier dua knot adalah sebagai berikut.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - K_1)_+ + \hat{\beta}_3 (x_1 - K_2)_+ + \hat{\beta}_4 (x_1 - K_3)_+ + \dots + \hat{\beta}_{21} x_6 + \hat{\beta}_{22} (x_6 - K_{16})_+ + \hat{\beta}_{23} (x_6 - K_{17})_+ + \hat{\beta}_{24} (x_6 - K_{18})_+$$

Nilai GCV untuk model tiga titik knot disajikan pada Tabel 4.

Tabel 4.

Nilai GCV Model Spline Linier Tiga Titik Knot

No.	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	GCV
1	80;84;99	54;58;64	55;79;86	35;39;47	6;9;12	20;25;49	21,63
2	69;85;91	50;52;61	42;62;86	31;40;46	13;14;15	17;31;33	12,03
3	72;74;92	49;56;60	53;54;73	37;38;49	2;9;12	11;17;29	13,11
4	57;58;77	51;56;59	42;49;78	16;21;24	2;8;13	21;30;35	1302,88
5	56;57;91	53;56;64	52;69;70	13;17;22	11;15;18	4;10;11	11,71
6	59;70;98	48;51;62	63;67;78	37;38;47	15;17;19	17;47;50	13,08
7	63;73;92	54;62;64	47;85;86	12;24;47	3;4;5	8;30;32	17,10
8	57;71;76	53;56;66	45;57;70	10;19;42	4;12;19	25;29;43	102,32
9	61;86;96	50;57;66	41;56;61	16;17;20	1;18;19	21;35;49	12,32
10	60;74;86	50;56;57	52;70;81	34;35;45	7;8;17	42;45;47	1511,59

Dari Tabel 4 didapatkan informasi bahwa nilai GCV minimum yang didapatkan sebesar 11,71 yang bersesuaian dengan titik-titik knot K₁ = 56, K₂ = 57, K₃ = 91, K₄ = 53, K₅ = 56, K₆ = 64, K₇ = 52, K₈ = 69, K₉ = 70, K₁₀ = 13, K₁₁ = 17, K₁₂ = 22, K₁₃ = 11, K₁₄ = 15, K₁₅ = 18, K₁₆ = 4, K₁₇ = 10 dan K₁₈ = 11.

F. Regresi Nonparametrik Spline Linier dengan Kombinasi Knot

Nilai GCV dari model spline dengan berbagai kombinasi knot disajikan pada Tabel 5.

Tabel 5 menunjukkan bahwa nilai GCV minimum dari model spline linier menggunakan kombinasi knot sebesar 3,94 dengan banyak knot optimum pada X₁ = 3 knot, X₂ = 1 knot, X₃ = 3 knot, X₄ = 1 knot, X₅ = 1 knot serta X₆ = 3 knot. Nilai GCV minimum tersebut bersesuaian dengan titik-titik knot K₁

= 67, K₂ = 85, K₃ = 87, K₄ = 49, K₅ = 65, K₆ = 74, K₇ = 78, K₈ = 8, K₉ = 5, K₁₀ = 6, K₁₁ = 10, dan K₁₂ = 29.

Tabel 5.
Nilai GCV Model Spline Linier dengan Kombinasi Knot

No	Variabel Prediktor	Jumlah Titik Knot Optimum	Titik-Titik Knot	GCV
1	X ₁	1	80	6,97
	X ₂	1	48	
	X ₃	1	72	
	X ₄	3	18,46,48	
	X ₅	1	9	
	X ₆	1	22	
2	X ₁	1	73	7,21
	X ₂	1	49	
	X ₃	1	39	
	X ₄	1	6	
	X ₅	3	4,15,19	
	X ₆	1	18	
3	X ₁	2	86,94	7,39
	X ₂	1	57	
	X ₃	1	41	
	X ₄	1	10	
	X ₅	1	5	
	X ₆	2	7,16	
4	X ₁	3	67,85,87	3,94
	X ₂	1	49	
	X ₃	3	65,74,78	
	X ₄	1	8	
	X ₅	1	5	
	X ₆	3	6,10,29	
5	X ₁	1	61	6,81
	X ₂	1	51	
	X ₃	1	53	
	X ₄	1	13	
	X ₅	3	12,16,17	
	X ₆	1	4	

G. Model Spline Terbaik

Dapat disimpulkan dari model spline linier satu titik knot, dua titik knot, tiga titik knot, serta kombinasi titik knot bahwa nilai GCV yang paling minimum adalah nilai GCV pada kombinasi titik knot.

Estimasi parameter untuk model spline terbaik akan disajikan pada Tabel 6.

Tabel 6.

Estimasi Parameter untuk Model Spline Terbaik

Variabel	Parameter	Estimasi
Intersep	β ₀	140,7182
	β ₁	0,0575
X ₁	β ₂	-0,3389
	β ₃	3,6364
	β ₄	-3,6669
	β ₅	1,0007
X ₂	β ₆	-0,8790
	β ₇	-2,8957
X ₃	β ₈	2,9946
	β ₉	0,4925
	β ₁₀	-0,8602

Tabel 6.
Estimasi Parameter untuk Model Spline Terbaik (Lanjutan)

Variabel	Parameter	Estimasi
X ₄	β ₁₁	0,3657
	β ₁₂	-0,3415
X ₅	β ₁₃	-0,0013
	β ₁₄	-0,0063
X ₆	β ₁₅	-15,502
	β ₁₆	29,880
	β ₁₇	-16,634
	β ₁₈	-25,989

Dari Tabel 6 diperoleh nilai-nilai estimasi parameter dari model spline terbaik yang dapat ditulis ke dalam bentuk persamaan sebagai berikut.

$$\hat{y} = 140,7182 + 0,0575x_1 - 0,3389(x_1 - 67)_+^1 + 3,6364(x_1 - 85)_+^1 +$$

$$- 3,6669(x_1 - 87)_+^1 + 1,0007x_2 - 0,8790(x_2 - 49)_+^1 +$$

$$- 2,8957x_3 + 2,9946(x_3 - 65)_+^1 + 0,4925(x_3 - 74)_+^1 +$$

$$- 0,8602(x_3 - 78)_+^1 + 0,3657x_4 - 0,3415(x_4 - 8)_+^1 +$$

$$- 0,0013x_5 - 0,0063(x_5 - 5)_+^1 + -1,5502x_6 + 2,9880(x_6 - 6)_+^1 +$$

$$- 1,6634(x_6 - 10)_+^1 - 2,5089(x_6 - 29)_+^1$$

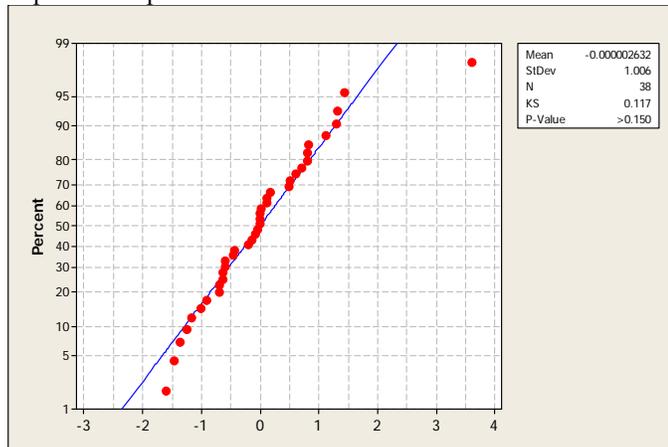
Dari model tersebut didapatkan nilai R² sebesar 88,77 persen yang berarti keenam variabel prediktor mampu menjelaskan sebesar 88,77 persen terhadap kejadian angka gizi buruk di Jawa Timur tahun 2007. Selanjutnya akan diuji apakah residual dari model tersebut berdistribusi normal dan apakah parameter-parameter model signifikan.

H. Uji Normalitas Residual

Untuk menguji asumsi ini digunakan uji Kolmogorov-Smirnov dengan hipotesis :

- H₀ : Residual berdistribusi normal
- H₁ : Residual tidak berdistribusi normal

Dengan menggunakan α = 0.05, maka hasil pengujian residual dapat dilihat pada Gambar 2.



Gambar. 2. Uji Normalitas Residual.

Dari Gambar 2 diperoleh informasi bahwa p-value > 0.15 yang nilainya lebih besar dari α = 0.05, maka gagal tolak H₀ yang artinya residual telah berdistribusi normal.

I. Uji Serentak

Untuk mengetahui pengaruh parameter secara serentak terhadap model maka dilakukan uji simultan dengan hipotesis:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{18} = 0$$

H₁ : Minimal ada satu β_p ≠ 0 ; p = 1,2,...,18
 Dengan menggunakan α = 0,05, maka diperoleh Tabel ANOVA yang dapat dilihat pada Tabel 7.

Tabel 7.
Tabel ANOVA Model Spline

Source of Variation	df	Sum of Square	Mean Square	F _{hitung}	F _{tabel}
Regression	18	280,5192	15,5844	7,4885	2,2172
Residual	19	39,5409	2,0811		
Total	37				

Tabel 7 menunjukkan bahwa nilai F hitung sebesar 7,4885 yang nilainya lebih besar dari F tabel sebesar 2,2172. Hal ini mengindikasikan bahwa H₀ ditolak, yang artinya minimal ada satu β_p ≠ 0.

J. Uji Parsial

Untuk mengetahui parameter mana saja yang berpengaruh terhadap model spline disajikan pada Tabel 8.

Tabel 8.
Keputusan pada Uji Parsial

Variabel	Parameter	t _{hitung}	t _{tabel}	Keputusan
X ₁	β ₁	0.3288	2.093	Tidak Signifikan
	β ₂	1.4745	2.093	Tidak Signifikan
	β ₃	4.7115	2.093	Signifikan
	β ₄	4.9027	2.093	Signifikan
X ₂	β ₅	0.93625	2.093	Tidak Signifikan
	β ₆	0.8047	2.093	Tidak Signifikan
X ₃	β ₇	3.7402	2.093	Signifikan
	β ₈	3.5157	2.093	Signifikan
	β ₉	1.2902	2.093	Tidak Signifikan
X ₄	β ₁₀	2.1309	2.093	Signifikan
	β ₁₁	0.7575	2.093	Tidak Signifikan
X ₅	β ₁₂	0.69051	2.093	Tidak Signifikan
	β ₁₃	0.0053293	2.093	Tidak Signifikan
X ₆	β ₁₄	0.021415	2.093	Tidak Signifikan
	β ₁₅	1.5363	2.093	Tidak Signifikan
	β ₁₆	2.3941	2.093	Signifikan
	β ₁₇	4.0353	2.093	Signifikan
	β ₁₈	2.8464	2.093	Signifikan

Dengan menggunakan tingkat signifikansi sebesar 5 persen didapatkan parameter-parameter yang signifikan yaitu β₃, β₄, β₇, β₈, β₁₀, β₁₆, β₁₇ dan β₁₈. Dari Tabel 8 didapatkan kesimpulan bahwa faktor-faktor yang mempengaruhi angka gizi buruk di Jawa Timur tahun 2007 adalah persentase ibu yang memeriksakan kehamilannya (X₁), persentase balita yang mendapatkan vitamin A (X₃), dan persentase rumah tangga miskin (X₆). Sehingga model spline dari faktor-faktor yang mempengaruhi angka gizi buruk di Jawa Timur pada tahun 2007 adalah sebagai berikut.

$$\hat{y} = 3,6364(x_1 - 85)_+^1 - 3,6669(x_1 - 87)_+^1 - 2,8957x_3 +$$

$$2,9946(x_3 - 65)_+^1 - 0,8602(x_3 - 78)_+^1 + 2,9980(x_6 - 6)_+^1 +$$

$$- 1,6634(x_6 - 10)_+^1 - 2,5089(x_6 - 29)_+^1$$

Interpretasi dari model spline terbaik tersebut adalah.

Apabila variabel X₃ dan X₆ konstan, maka hubungan antara persentase ibu yang memeriksakan kehamilan (X₁) terhadap persentase angka gizi buruk (Y),

$$\hat{y} = 3,6364(x_1 - 85)_+^1 - 3,6669(x_1 - 87)_+^1$$

$$= \begin{cases} 3,6364x_1 - 309,094, & 85 \leq x_1 < 87 \\ -0,0305x_1 + 9,9263, & x_1 \geq 87 \end{cases}$$

Dari model di atas dapat diinterpretasikan, Pada saat persentase ibu yang memeriksakan kehamilannya (X_1) antara 85 persen sampai 87 persen, apabila persentase ibu yang memeriksakan kehamilannya naik sebesar satu persen, maka persentase angka gizi buruk akan naik sebesar 3,6364 persen. Hal ini dikarenakan pada tahun 2007, persentase balita gizi buruk juga tinggi, sehingga faktor ibu yang memeriksakan kehamilannya tidak terlalu berpengaruh signifikan. Persentase ibu yang memeriksakan kehamilannya lebih dari 87 persen, apabila persentase ibu yang memeriksakan kehamilannya naik sebesar satu persen, maka persentase angka gizi buruk akan turun sebesar 0,0305 persen.

2. Apabila variabel X_1 dan X_6 konstan, maka hubungan antara persentase balita yang mendapatkan vitamin A (X_3) terhadap persentase angka gizi buruk (Y),

$$\hat{y} = -2,8957x_3 + 2,9946(x_3 - 65)_+^1 - 0,8602(x_3 - 78)_+^1$$

$$= \begin{cases} -2,8957x_3, & x_3 < 65 \\ 0,0989x_3 - 194,649, & 65 \leq x_3 < 78 \\ -0,7613x_3 - 127,5534, & x_3 \geq 78 \end{cases}$$

Dari model di atas dapat diinterpretasikan, Pada saat persentase balita yang mendapatkan vitamin A (X_3) kurang dari 65 persen, apabila persentase balita yang mendapatkan vitamin A naik sebesar satu persen, maka persentase angka gizi buruk akan turun sebesar 2,8957 persen. Persentase balita yang mendapatkan vitamin A antara 65 persen sampai 78 persen, apabila persentase balita yang mendapatkan vitamin A naik sebesar satu persen, maka persentase angka gizi buruk akan naik sebesar 0,0989 persen. Persentase balita yang mendapatkan vitamin A lebih dari 78 persen, apabila persentase balita yang mendapatkan vitamin A naik sebesar satu persen, maka persentase angka gizi buruk akan turun sebesar 0,7613 persen.

3. Apabila variabel X_1 dan X_3 konstan, maka hubungan antara persentase rumah tangga miskin (X_6) terhadap persentase angka gizi buruk (Y),

$$\hat{y} = 2,9980(x_6 - 6)_+^1 - 1,6634(x_6 - 10)_+^1 - 2,5089(x_6 - 29)_+^1$$

$$= \begin{cases} 2,9980x_6 - 17,988, & 6 \leq x_6 < 10 \\ 1,3346x_6 - 1,354, & 10 \leq x_6 < 29 \\ -1,1743x_6 + 71,4041, & x_6 \geq 29 \end{cases}$$

Dari model di atas dapat diinterpretasikan, Pada saat persentase rumah tangga miskin (X_6) antara 6 persen sampai 10 persen, apabila persentase rumah tangga miskin naik sebesar satu persen, maka persentase angka gizi buruk akan naik sebesar 2,9980 persen. Persentase rumah tangga miskin antara 10 persen sampai 29 persen, apabila persentase rumah tangga miskin naik sebesar satu persen, maka persentase angka gizi buruk akan naik sebesar 1,3346 persen. Persentase rumah tangga miskin lebih dari 29 persen, apabila persentase rumah tangga miskin naik sebesar satu persen, maka persentase angka gizi buruk akan turun sebesar 1,1743 persen. Hal ini dikarenakan pada tahun 2007, persentase balita gizi buruk juga tinggi, sehingga faktor rumah tangga miskin tidak terlalu berpengaruh signifikan.

V. KESIMPULAN DAN SARAN

Berdasarkan deskripsi kejadian balita gizi buruk tiap kabupaten/kota di Jawa Timur, didapatkan persentase angka gizi buruk tertinggi adalah Kabupaten Sampang sebesar 16,2 persen. Dari hasil analisa yang telah dilakukan, didapatkan kesimpulan bahwa faktor-faktor yang mempengaruhi angka gizi buruk di Jawa Timur pada Tahun 2007 adalah persentase ibu yang memeriksakan kehamilannya (X_1), persentase balita yang mendapatkan vitamin A (X_3) dan persentase rumah tangga miskin (X_6), dengan model spline terbaiknya adalah sebagai berikut.

$$\hat{y} = 3,6364(x_1 - 85)_+^1 - 3,6669(x_1 - 87)_+^1 - 2,8957x_3 + 2,9946(x_3 - 65)_+^1 - 0,8602(x_3 - 78)_+^1 + 2,9980(x_6 - 6)_+^1 - 1,6634(x_6 - 10)_+^1 - 2,5089(x_6 - 29)_+^1$$

dari model di atas didapatkan nilai R^2 sebesar 88,77 persen sehingga dapat dikatakan baik dalam pemodelan.

DAFTAR PUSTAKA

- [1] Samsul. (2011). *Dampak Gizi Buruk Bagi Anak-Anak Penerus Bangsa*. Accessed Januari 2, 2012, from <http://samsuljoker.blogspot.com/2011/01/dampak-gizi-buruk-bagi-anak-anak.html>
- [2] Anonim. (2012). *900 Ribu Balita Indonesia Bergizi Buruk*. Accessed January 19, 2012, from <http://www.tribunnews.com/2012/01/18/900-ribu-balita-indonesia-bergizi-buruk>
- [3] Siswono. (2010). *Kasus Gizi Buruk : Empat Provinsi Tak Pernah Absen*. Accessed Januari 2, 2012, from <http://gizi.net/2010/07/kasus-gizi-buruk-empat-provinsi-tak-pernah-absen.html>
- [4] Paramita, L. (2008). *Bagging Regresi Logistik Ordinal pada Klasifikasi Status Gizi Balita (Studi Kasus Kabupaten Nganjuk)*. Tugas Akhir, Jurusan Statistika Institut Teknologi Sepuluh Nopember, Surabaya.
- [5] Hayati, M. (2009). *Analisis Diskriminan Pada Faktor-Faktor yang Mempengaruhi Gizi Buruk Balita di Jawa Timur*. Tugas Akhir, Jurusan Statistika Institut Teknologi Sepuluh Nopember, Surabaya.
- [6] Riskiyanti, R. (2010). *Analisis Regresi Multivariat Berdasarkan Faktor-faktor yang Mempengaruhi Derajat Kesehatan di Provinsi Jawa Timur*. Tugas Akhir, Jurusan Statistika Institut Teknologi Sepuluh Nopember, Surabaya.
- [7] Ayunin, L. (2011). *Pemodelan Balita Gizi Buruk Di Kabupaten Ngawi Dengan Geographically Weighted Regression*. Tugas Akhir, Jurusan Statistika Institut Teknologi Sepuluh Nopember, Surabaya.
- [8] A'yumin, Q. (2011). *Pemodelan Angka Gizi Buruk pada Balita di Kota Surabaya dengan Spatial Autoregressive Model (SAR)*. Tugas Akhir, Jurusan Statistika Institut Teknologi Sepuluh Nopember, Surabaya.
- [9] Fridiati, I.D. (2009). *Pendekatan Mars Untuk Pemodelan Gas Buang Kendaraan Dengan Bahan Bakar Solar*. Tugas Akhir, Jurusan Statistika Institut Teknologi Sepuluh Nopember, Surabaya.
- [10] Sutarsi, S. (2008). *Pendekatan Regresi Spline untuk Memodelkan Nilai UNAS Siswa SMK Negeri 3 Buduran Sidoarjo*. Tugas Akhir, Jurusan Statistika Institut Teknologi Sepuluh Nopember, Surabaya.
- [11] Basri, H. (2008). *Estimasi Kurva Regresi Nonparametrik dengan Pendekatan Spline (Studi Kasus pada Data Murid Madrasah Ibtidaiyah dan Keluarga Prasejahtera Setiap Kecamatan di Kabupaten Bone)*. *Didaktika Jurnal Kependidikan*, Vol. 3 No.2.
- [12] Federika, Y. (2011). *Regresi Nonparametrik Spline untuk Data Berat Badan Balita Menurut Umur di Kabupaten Bojonegoro Tahun 2010*. Tugas Akhir, Jurusan Statistika Institut Teknologi Sepuluh Nopember, Surabaya.
- [13] Eubank, R.L. (1988). *Spline Smoothing and Nonparametric Regression*. New York : Marcel Dekker.
- [14] Budiantara, I.N. (2000). *Metode UBR, GML, CV, dan GCV dalam Regresi Nonparametrik Spline*. *Majalah Ilmiah Himpunan Matematika Indonesia (MIHMI)*. 6,285-290.