

MODEL REGRESI SPLINE KNOT OPTIMAL UNTUK MENGETAHUI FAKTOR – FAKTOR YANG MEMPENGARUHI JUMLAH KEMATIAN BAYI DI JAWA TIMUR

Elsha Puspitasari¹, Drs. Hery Tri Sutanto, M.Si²,

^{1,2} Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Surabaya
Jalan Ketintang, Surabaya

email : mark_elsha@yahoo.com¹, hery_trisutanto@gmail.com²

ABSTRAK

Model regresi spline merupakan model analisis dengan menggunakan pendekatan nonparametrik dimana menggunakan estimasi least square dengan titik knot optimal yang dipilih berdasarkan nilai GCV (*Generalized Cross Validation*) terkecil. Model regresi nonparametrik spline dengan titik knot optimal di aplikasikan pada jumlah kematian bayi di Jawa Timur dimana model spline dengan titik – titik knot terpilih merupakan titik knot optimal dengan nilai GCV paling kecil sebesar 1,00 dan nilai $R^2 = 85,00\%$. Dengan model regresi spline dan menguji masing – masing parameter diperoleh faktor – faktor yang mempengaruhi jumlah kematian bayi di Jawa Timur yaitu jumlah sarana medis (x_1), Persentase persalinan yang menggunakan tenaga non medis (x_3), Persentase ibu yang tidak melakukan kunjungan kehamilan (x_4), Persentase bayi yang tidak di beri ASI (x_5) . Dengan demikian model regresi nonparametrik spline dengan pemilihan titik knot optimal serta menggunakan estimasi least square mampu mengestimasi pola data nonparametrik.

Keywords: Regresi Nonparametrik, Spline, GCV, Knot, Parameter, Residual.

I.PENDAHULUAN

Regresi nonparametrik bentuk kurva regresi diasumsikan tidak diketahui sehingga dapat digunakan pada pendekatan regresi yang sesuai untuk pola data yang tidak diketahui bentuknya, atau tidak terdapat informasi masa lalu tentang pola data (Budiantara,2001b). Kurva regresi hanya diasumsikan mulus (smooth) dalam arti termuat di dalam suatu ruang fungsi tertentu. Model regresi nonparametrik yang sering mendapat perhatian dari para peneliti adalah Kernel, Spline, Deret Fourier dan Wavelets. Diantara model-model regresi nonparametrik di atas, Spline merupakan model regresi yang mempunyai interpretasi Statistik dan

Visual sangat khusus dan sangat baik (Budiantara, 2006). Model Spline diperoleh dari suatu optimasi *Penalized Least Square* (PLS) dan memiliki fleksibilitas yang tinggi, mampu menangani karakter data/fungsi yang mulus (*smooth*). Spline juga memiliki kemampuan yang sangat baik untuk menangani data yang perilakunya berubah – ubah pada sub-sub interval tertentu disebabkan spline merupakan model polinomial yang tersegmen/terbagi, dalam penulisan ini, akan ditampilkan pendekatan model spline yang optimal yang dapat memperlihatkan bentuk visual dari perubahan perilaku pola data pada interval yang berlainan yang tidak didasarkan pada parameter penghalus tetapi menggunakan titik knots dimana titik knot optimal dari model spline, dimana dalam penelitian ini yang akan dikaji adalah pendekatan spline pada pemodelan jumlah kematian bayi di Jawa Timur.

Tujuan Penelitian untuk mengetahui faktor – faktor yang mempengaruhi kematian bayi di Jawa Timur.

I.TINJAUAN PUSTAKA

A.Regresi Nonparametrik Spline

Regresi nonparametrik merupakan analisis regresi yang sesuai untuk pola data yang tidak diketahui bentuknya, atau tidak terdapat informasi masa lalu tentang pola data (Budiantara,2001b). Model umum persamaan regresi pada persamaan (1) yaitu

$$Y = f(x) + \varepsilon \quad (1)$$

Fungsi f yang belum diketahui bentuknya akan diduga dengan model regresi nonparametrik spline. Untuk tujuan tersebut kita akan mendefinisikan fungsi spline order r dengan titik knot yaitu $\xi_1, \xi_2, \dots, \xi_k$ pada sebarang fungsi $f(x)$ yang disajikan sebagai berikut :

$$f(x) = \beta_0 + \sum_{j=1}^p \left[\sum_{r=1}^{m-1} \beta_{j,r} X_j^r + \sum_{k=1}^{s_j} \delta_{j,(m-1),k} (X_j - \xi_{jk})_+^{m-1} \right] \quad (2)$$

Dengan

$$(X_j - \xi_{jk})_+^{m-1} = \begin{cases} (X_j - \xi_{jk})_+^{m-1} & ; X_j \geq \xi_{jk} \\ 0 & ; X_j < \xi_{jk} \end{cases} \quad (3)$$

Dimana :

$\beta_0, \beta_1, \beta_2, \dots, \beta_{m-1}$ dan $\delta_1, \delta_2, \delta_3, \dots, \delta_{s_j}$ adalah himpunan koefisien atau konstanta real, X_j peubah penjelas ke-j, ξ_{jk} adalah knot ke-k pada peubah X_j , s_j banyaknya knot dalam peubah penjelas ke-j.

Agar diperoleh spline yang optimal perlu dipilih titik knot yang optimal. Salah satu metode untuk mendapatkan titik knot optimal adalah dengan metode *Generalized Cross Validation* atau *GCV* (Budiantara, 2000) dimana kriteria *GCV* dapat didefinisikan sebagai berikut :

$$GCV(k_1, k_2, \dots, k_m) = \frac{MSE(k_1, k_2, \dots, k_m)}{\{n^{-1} \text{tr}(I - A(k_1, k_2, \dots, k_m))\}^2}$$

Dengan

$$MSE(k_1, k_2, \dots, k_m) = n^{-1} \sum_{j=1}^M (y_j - \hat{f}_{(k_1, k_2, \dots, k_m)}(t_j))^2, \quad k_1, k_2, \dots, k_m \text{ adalah titik knot dan matrik}$$

$$A(k_1, k_2, \dots, k_m) = T(k_1, k_2, \dots, k_m) (T'(k_1, k_2, \dots, k_m) T(k_1, k_2, \dots, k_m))^{-1} T'(k_1, k_2, \dots, k_m)$$

Nilai titik knot yang optimal diperoleh dari nilai *GCV* yang minimum.

B. Pengujian Signifikansi Parameter Model Terbaik dan Pengujian Asumsi Residual.

Untuk mengetahui signifikansi dari variabel prediktor maka dilakukan pengujian parameter secara serempak maupun secara individu.

a) Uji Serentak

Uji parameter dengan model regresi secara serentak dilakukan secara bersamaan terhadap model dengan hipotesis dari pengujian adalah :

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

$$H_1 : \text{paling sedikit ada satu } \beta_j \neq 0 \quad ;$$

$$j=1, 2, \dots, p$$

Statistik uji

$$F_{hitung} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2 / p}{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n - p - 1)} \quad (5)$$

Apabila $F_{hitung} > F_{\alpha(v_1, v_2)}$ maka H_0 ditolak artinya paling sedikit ada satu β_i yang memiliki pengaruh signifikan terhadap variabel respon.

b) Uji Individu

Uji parameter model regresi secara individu bertujuan untuk mengetahui seberapa jauh variabel prediktor secara individual dalam menerangkan variasi variabel respon.

Hipotesis dari pengujian secara individu adalah

$H_0 : \beta_j = 0$; artinya variabel prediktor tidak berpengaruh terhadap variabel respon

$H_1 : \beta_j \neq 0$; $j=1, 2, \dots, p$; artinya variabel prediktor berpengaruh terhadap variabel respon

Statistik uji

$$t_{hitung} = \frac{\hat{\beta}_j^2}{\hat{SE}(\beta_j)^2} \quad (6)$$

H_0 ditolak apabila $|t_{hitung}| > t_{(\frac{\alpha}{2}, n-p-1)}$ yang artinya ada pengaruh antara variabel prediktor terhadap variabel respon.

Pengujian asumsi residual erat kaitannya dengan kelayakan model regresi. Suatu model regresi dengan parameter signifikan dan memenuhi kriteria terbaik tetapi melanggar asumsi residual tidak disarankan untuk dipakai untuk menggambarkan pola hubungan antara variabel prediktor dan variabel respon. Untuk itu pengujian residual yang digunakan :

c) Uji Normalitas Residual

Uji Normalitas bertujuan untuk mengetahui apakah residual berdistribusi normal. Statistik uji yang digunakan adalah *Kolmogorov Smirnov*. Hipotesis yang digunakan adalah sebagai berikut.

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

Statistik uji yang digunakan adalah :

$$D = \sup_x [|F_n(x) - F_0(x)|] \quad (7)$$

Dimana $F_n(x)$ adalah probabilitas kumulatif normal dan $F_0(x)$ adalah probabilitas kumulatif empiris. H_0 diterima

apabila nilai D lebih kecil dari D tabel yang artinya residual berdistribusi normal.

C. Kematian Bayi

Kematian bayi adalah kematian yang terjadi antara saat setelah bayi lahir sampai bayi belum berumur tepat satu tahun. Banyak faktor yang dikaitkan dengan kematian bayi. Secara garis besar, dari sisi penyebabnya, kematian bayi ada dua macam yaitu endogen dan eksogen.

Kematian bayi endogen atau yang umum disebut dengan kematian neonatal adalah kematian bayi yang terjadi pada bulan pertama setelah dilahirkan, dan umumnya disebabkan oleh faktor – faktor yang dibawa anak sejak lahir, yang diperoleh dari orang tuanya pada saat konsepsi atau didapat selama bulan kehamilan.

Kematian bayi eksogen atau kematian post neo – natal adalah kematian bayi yang terjadi setelah usia satu bulan sampai menjelang usia satu tahun yang disebabkan oleh faktor – faktor dari luar atau pengaruh lingkungan luar seperti keadaan sosial ekonomi, jumlah sarana medis, penolong pertama pada kelahiran, jumlah air bersih dan sebagainya.

III. METODOLOGI PENELITIAN

A. SUMBER DATA

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari hasil Susenas 2011 oleh Badan Pusat Statistika (BPS, 2012) dan Dinas Kesehatan Provinsi Jawa Timur 2011. Berupa data Kematian Bayi di 38 kota/ kabupaten di Jawa Timur.

B. VARIABEL PENELITIAN

Dalam penelitian ini terdapat satu variabel dependen dan lima variabel independen yaitu :

1. Jumlah Kematian Bayi (y)
2. Jumlah sarana medis (x_1),
3. Persentase bayi berat badan lahir rendah (x_2),
4. Persentase persalinan yang menggunakan tenaga non medis (x_3),
5. Persentase ibu yang tidak melakukan kunjungan kehamilan (x_4),
6. Persentase bayi yang tidak di beri ASI (x_5).

C. LANGKAH PENELITIAN

Untuk dapat memodelkan data Kematian Bayi dengan regresi nonparametrik spline maka akan dilakukan beberapa tahap sebagai berikut :

1. Membuat statistika deskriptif variabel dependent dalam hal ini jumlah kematian bayi dan variabel independen yaitu persentase jumlah sarana medis, persentase

wanita yang tidak pernah sekolah atau tidak tamat SD/MI, persentase persalinan yang menggunakan tenaga non medis, persentase wanita yang berumah tangga di bawah umur 17 tahun, persentase bayi yang tidak di beri ASI.

2. Membuat scatter plot antara jumlah kematian bayi dengan masing – masing variabel prediktor yang dijadikan deteksi awal pola hubungan antara variabel respon terhadap variabel prediktor.
3. Memodelkan jumlah kematian bayi dengan menggunakan spline linear dengan satu titik knot.
4. Memilih titik knot optimal berdasarkan nilai GCV minimal.
5. Memodelkan jumlah kematian bayi dengan variabel – variabel prediktornya dengan regresi spline dengan knot optimal.
6. Melakukan pengujian signifikansi parameter dan pengujian asumsi residual terbaik pada regresi spline.
7. Menginterpretasikan hasil analisis dan mengambil kesimpulan.

IV. ANALISIS DAN PEMBAHASAN

A. ESTIMASI PARAMETER MODEL REGRESI SPLINE

Hasil estimasi parameter model regresi spline linear dengan satu variabel prediktor menggunakan metode weighted least square adalah seperti berikut.

Sebagai contoh diberikan data (x_a, y_a) dan diasumsikan hubungan antara x_a dan y_a mengikuti model regresi nonparametrik $y_a = f(x_a) + \varepsilon_a$, $a = 1, 2, 3, \dots, n$ kurva regresi $f(x_a)$ didekati dengan model spline linear dengan m knots k_1, k_2, \dots, k_m dengan $k_1 < k_2 < \dots < k_m$.

Kemudian diberikan suatu basis untuk ruang Spline berorde n (Budiantara, 2001) dengan bentuk:

$$\{1, x_{a1}, x_{a2}, \dots, x_a^n, (x_a - k_1)_+^n, \dots, (x_a - k_m)_+^n\}$$

Dengan fungsi truncated sebagai berikut :

$$(x_a - k)_+^n = \begin{cases} (x_a - k)^n, & x \geq k \\ 0, & x < k \end{cases}$$

Untuk setiap fungsi f dalam ruang Spline dapat dinyatakan menjadi :

$$f(x) = \sum_{j=0}^n \gamma_j x_a^j + \sum_{m=1}^m \gamma_{m+n} (x_a - k_m)_+^n \quad (8)$$

Dengan γ_j , $j = 0, 1, \dots, n+1, \dots, n+m$ merupakan konstanta yang bernilai real.

Model regresi spline dapat ditulis $y_a = f(x_a) + \varepsilon_a$ (persamaan 1) menjadi :

$$= \sum_{j=0}^n \gamma_j x_a^j + \sum_{m=1}^m \gamma_{m+n} (x_a - k_m)_+^n + \varepsilon_a$$

Dari regresi spline ini dapat ditulis :

$$\varepsilon_a = y_a - \sum_{j=0}^n \gamma_j x_a^j - \sum_{m=1}^m \gamma_{m+n} (x_a - k_m)_+^n$$

Untuk setiap $a = 1, 2, \dots, n$

Jika persamaan di atas dinyatakan dalam bentuk matriks, maka diperoleh :

$$\underline{\varepsilon} = \underline{y} - X(x, k)\underline{\gamma}$$

Selanjutnya dapat dibentuk suatu fungsi :

$$Q(\underline{\gamma}) = \underline{\varepsilon}^T V^{-1} \underline{\varepsilon}$$

Dengan $\underline{\gamma} = (\gamma_0, \gamma_1, \dots, \gamma_p, \gamma_{p+1}, \dots, \gamma_{p+m})^T$,

$\underline{y} = (y_1, y_2, \dots, y_a)^T$, $X(x, k)$ matriks berukuran $p \times (n+m+1)$, diberikan oleh :

$$X(x, k) = \begin{pmatrix} 1 & x_1 & \dots & x_1^n & (x_1 - k_1)_+^n & \dots & (x_1 - k_m)_+^n \\ 1 & x_2 & \dots & x_2^n & (x_2 - k_2)_+^n & \dots & (x_2 - k_m)_+^n \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 1 & x_a & \dots & x_a^n & (x_a - k_1)_+^n & \dots & (x_a - k_m)_+^n \end{pmatrix}$$

Dan

$$\underline{\gamma} = \begin{pmatrix} \gamma_0 \\ \vdots \\ \gamma_n \\ \dots \\ \gamma_{n+1} \\ \vdots \\ \gamma_{n+m} \\ \dots \\ \theta \end{pmatrix}$$

Apabila error random $\underline{\varepsilon} \sim N(0, V\sigma^2)$ maka estimasi $\underline{\gamma}$ dapat diselesaikan dengan optimasi sebagai berikut :

$$\begin{aligned} \underline{\gamma} &\in \underset{R^{n+m+2}}{\text{Min}} (\underline{\varepsilon}^T V^{-1} \underline{\varepsilon}) \\ &= \underset{R^{n+m+2}}{\text{Min}} \left\{ (\underline{y} - X\underline{\gamma})^T V^{-1} (\underline{y} - X\underline{\gamma}) \right\} \end{aligned}$$

Dengan menggunakan *derivatif parsial* dapat diselesaikan dengan cara :

$$\frac{\partial (\underline{\varepsilon}^T V^{-1} \underline{\varepsilon})}{\partial \underline{\gamma}} = 0 - 2X^T V^{-1} \underline{y} + 2X^T V^{-1} X \underline{\gamma} = 0$$

Mengingat X merupakan matriks dengan rank penuh, maka diperoleh estimasi $\underline{\gamma}$ adalah :

$$\underline{\gamma} = (X^T V^{-1} X)^{-1} X^T V^{-1} \underline{y}$$

Estimator kurva regresi $f(x)$ diberikan oleh :

$$\begin{aligned} f(x, y) &= X(X^T V^{-1} X)^{-1} X^T V^{-1} \underline{y} \\ &= \beta(x, y) y \end{aligned}$$

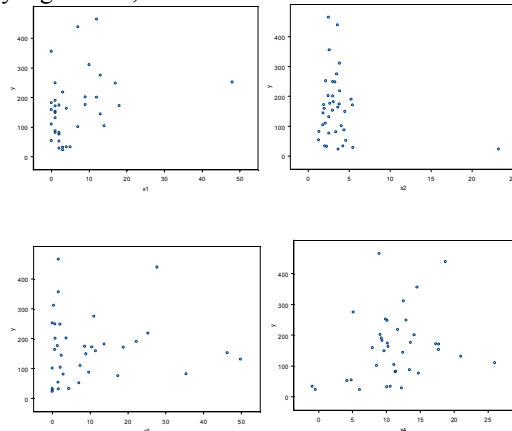
B. STATISTIKA DESKRIPTIF

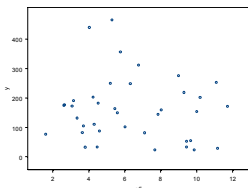
Provinsi Jawa Timur memiliki 29 Kabupaten dan 9 Kota atau secara administratif terdapat 38 Kabupaten/Kota. Setiap kabupaten/kota memiliki kondisi sosial dan ekonomi yang berbeda. Berikut hasil analisa deskriptif dari tiap variabel penelitian.

Tabel 1 .Deskripsi Faktor - Faktor yang Berpengaruh Terhadap Jumlah Kematian Bayi

Variabel	Mean	Min	Maks	Varians
y	160,50	23,00	465,00	11890,42
x1	6,11	0,00	48,00	75,77
x2	3,66	1,27	23,24	11,85
x3	9,41	0,00	49,78	162,66
x4	11,23	-0,93	26,00	27,68
x5	6,37	1,61	11,72	8,04

Dari Tabel 1 diketahui bahwa jumlah kematian bayi terbesar berada di daerah Surabaya sebanyak 465 kasus sedangkan jumlah kematian bayi terkecil berada di daerah Mojokerto dan Madiun sebanyak 23 kasus. Jumlah sarana medis dalam hal ini jumlah rumah bersalin (x_1) memiliki varians atau keragaman yang cukup besar antara kabupaten/kota yaitu 75,77 sehingga menunjukkan kesejangan antar kabupaten/kota. Persentase bayi berat badan lahir rendah (x_2) memiliki mean 3,66 dengan keragaman data sebesar 11,85, sedangkan persentase persalinan yang menggunakan tenaga non medis (x_3) dalam hal ini penolong pertamanya adalah dukun paling besar berada di kabupaten Sampang sebesar 49,78 persen artinya dari 100 kelahiran hampir 50 orang menggunakan jasa dukun dalam persalinan. Persentase ibu yang tidak melakukan kunjungan kehamilan (x_4) memiliki mean 11,23 dan persentase bayi yang tidak diberi ASI eksklusif memiliki keragaman atau varians yang kecil 8,04.





Gambar 1. Scatter Plot Antara Jumlah Kematian Bayi dengan masing – masing 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 atau 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 LINEAR

Berdasarkan pada persamaan (2) maka dapat dibentuk model regresi spline linear dengan satu titik knot dengan lima variabel prediktor sebagai berikut :

$$y = \beta_0 + \beta_1 x_1 + \beta_2(x_1 - k_1) + \beta_3 x_2 + \beta_4(x_2 - k_2) + \beta_5 x_3 + \beta_6(x_3 - k_3) + \beta_7 x_4 + \beta_8(x_4 - k_4) + \beta_9 x_5 + \beta_{10}(x_5 - k_5)$$

Dengan beberapa kali percobaan dalam membentuk model *spline* linier dengan 1 knot diperoleh titik knot yang optimum berdasarkan GCV minimum. Beberapa titik-titik knot dan GCV minimum yang terbentuk disajikan dalam Tabel 4.2

Tabel 2 Nilai GCV minimum model spline linear satu titik knot

No	x1	x2	x3	x4	x5	GCV
1	2,00	1,31	8,87	10,27	4,53	1,02
2	2,00	2,99	8,87	13,57	7,86	1,07
3	2,00	2,99	22,22	10,27	3,08	1,09
4	2,00	4,48	22,22	10,27	7,86	1,03
5	12,00	1,31	0,61	13,57	7,86	1,13
6	12,00	1,31	8,87	10,27	3,08	1,14
7	12,00	2,99	7,32	13,57	7,86	1,01
8	12,00	2,99	8,87	10,27	3,08	1,15
9	12,00	4,48	22,22	10,27	4,53	1,00
10	17,00	2,99	8,87	13,57	3,08	1,04

Dari tabel 2 terlihat bahwa nilai GCV minimum bersesuaian dengan titik – titik knot yaitu $k_1 = 12,00$; $k_2 = 4,48$; $k_3 = 22,22$; $k_4 = 10,27$; dan $k_5 = 4,53$ dimana nilai GCV nya adalah 1,00.

Estimasi parameter yang terbentuk dari pemilihan titik – titik knot optimal berdasarkan GCV minimum disajikan pada tabel 4.3 dimana banyak estimasi parameter yang terbentuk sebanyak 11 termasuk didalamnya adalah *intercept*.

Tabel 3 Estimasi parameter model spline satu titik knot

Xi	Estimasi Parameter
Intercept	$\beta_0 = 110.8963$
1	$\beta_1 = 9.1741$ $\beta_2 = 16.3098$
2	$\beta_3 = 6.4942$ $\beta_4 = 13.5420$
3	$\beta_5 = 87.0579$ $\beta_6 = -0.0764$
4	$\beta_7 = -0.6556$ $\beta_8 = -0.1164$
5	$\beta_9 = -0.3107$ $\beta_{10} = -6.6709$

Dari tabel 3 diperoleh nilai – nilai estimasi model regresi spline terbaik dengan satu titik knot dimana telah diperoleh titik knot optimal dengan GCV minimum pada tabel 2 sehingga dapat dibentuk model persamaan regresi spline terbaik dengan satu titik knot sebagai berikut :

$$y = 110,8963 + 9,1741x_1 + 16,3098(x_1 - 12,00) + 6,4942x_2 + 13,5420(x_2 - 4,48) + 87,0579x_3 - 0,0764(x_3 - 22,22) - 0,6556x_4 - 0,1164(x_4 - 10,27) - 0,3107x_5 - 6,6709(x_5 - 4,53)$$

Dari model tersebut didapatkan nilai R^2 sebesar 85,00 persen yang berarti keenam variabel prediktor mampu menjelaskan sebesar 85,00 persen terhadap jumlah kematian bayi di Jawa Timur tahun 2011. Selanjutnya akan diuji apakah residual dari model tersebut berdistribusi normal dan apakah parameter-parameter model signifikan.

D. ANALISIS REGRESI NONPARAMETRIK SPLINE LINEAR

1. Pengujian Signifikansi Parameter Model Terbaik

Terdapat dua tahap pengujian parameter regresi, yaitu pengujian secara simultan (uji serentak) dan secara parsial. Uji simultan merupakan pengujian parameter model regresi secara bersamaan. Sedangkan uji parsial adalah pengujian parameter model regresi secara satu persatu.

a) Uji Serentak

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

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{10} = 0$$

$$H_1 : \text{paling sedikit ada satu } \beta_j \neq 0 ;$$

$$j=1,2,\dots,10$$

Tabel 4. Analisis varians model spline

Sumber Variansi	Derajat Bebas	Jumlah Kuadrat	Rataan Kuadrat	F Hitung	F Tabel
Regresi	10	17,8468	4,4982	4,1001	2,143
Residual	27	26,1485	7,263		
Total	37				

Pada Tabel 4 dapat dilihat bahwa nilai F_{hitung} sebesar 4,1001. Pada pengujian ini, taraf signifikansi adalah 5% dan $F_{0,05(10,27)} = 2,143$. Daerah penolakan H_0 dilakukan apabila nilai $F_{hitung} > F_{0,05(10,27)}$. Berdasarkan nilai F_{hitung} dan $F_{0,05(10,27)}$ maka keputusan yang diambil H_0 ditolak karena $F_{hitung} > F_{0,05(10,27)}$. Sehingga, dapat ditarik kesimpulan bahwa di dalam parameter pada model ini terdapat minimal satu parameter yang signifikan.

b) Uji Individu

Untuk mengetahui parameter mana saja yang berpengaruh terhadap model spline dilakukan uji individu dengan hipotesis sebagai berikut :

$H_0 : \beta_j = 0$; artinya variabel prediktor tidak berpengaruh terhadap variabel respon

$H_1 : \beta_j \neq 0 ; j=1,2,\dots,10$; artinya variabel prediktor berpengaruh terhadap variabel respon

H_0 ditolak apabila $|t_{hit}| > t_{((\alpha/2),(n-m))}$. Dengan taraf signifikansi $\alpha = 5\%$, maka diperoleh nilai $t_{tabel} = 2.05$, untuk mengetahui parameter yang signifikan terlihat pada tabel 4.5

Tabel 5. Uji individu estimasi parameter spline linear satu titik knot

Xi	Estimasi Parameter	t hitung	Keputusan
Intercept	$\beta_0 = 110.8963$	6,0597	Signifikan
1	$\beta_1 = 9.1741$	2,0985	Signifikan
	$\beta_2 = 16.3098$	1,7849	Tidak Signifikan
2	$\beta_3 = 6.4942$	1,6353	Tidak Signifikan
	$\beta_4 = 13.5420$	1,6771	Tidak Signifikan
3	$\beta_5 = 87.0579$	2,1432	Signifikan
	$\beta_6 = -0.0764$	2,5116	Signifikan
4	$\beta_7 = -0.6556$	1,7602	Tidak Signifikan
	$\beta_8 = -0.1164$	2,9210	Signifikan
5	$\beta_9 = -0.3107$	1,5025	Tidak Signifikan
	$\beta_{10} = -6.6709$	2,2755	Signifikan

Dengan menggunakan tingkat signifikansi sebesar 5 persen didapatkan parameter-parameter yang signifikan yaitu $\beta_0, \beta_1, \beta_5, \beta_6, \beta_8,$ dan β_{10} . Dari Tabel 5 didapatkan kesimpulan bahwa faktor-faktor yang mempengaruhi jumlah kematian bayi di Jawa Timur tahun 2011 adalah jumlah sarana medis (x_1), Persentase persalinan yang menggunakan tenaga non medis (x_3), Persentase ibu yang tidak melakukan kunjungan kehamilan (x_4), Persentase bayi yang tidak di beri ASI (x_5).

Sehingga model spline dari faktor-faktor yang mempengaruhi jumlah kematian bayi di Jawa Timur pada tahun 2011 adalah sebagai berikut.

$$y = 110,8963 + 9,1741x_1 + 87,0579x_3 - 0,0764(x_3 - 22,22) - 0,1164(x_4 - 10,27) - 6,6709(x_5 - 4,53)$$

2. Pengujian Asumsi Residual

Pengujian asumsi residual erat kaitannya dengan kelayakan model regresi. Suatu model regresi dengan parameter signifikan dan memenuhi kriteria terbaik tetapi melanggar asumsi residual tidak disarankan untuk dipakai untuk menggambarkan pola hubungan antara variabel

prediktor dan variabel respon. Uji Asumsi Residual yang digunakan adalah

a) Uji Normalitas Residual

Untuk menguji asumsi ini digunakan uji Kolmogorov-

Smirnov dengan hipotesis :

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

Dengan menggunakan $\alpha = 0.05$ dan daerah penolakan H_0 jika $P_{\text{value}} < \alpha$, dengan menggunakan macro R menghasilkan P_{value} sebesar 0,08913. Karena $P_{\text{value}} > \alpha$ maka terima H_0 artinya model spline telah memenuhi asumsi normal.

Setelah model spline linear dengan satu titik knot telah memenuhi uji signifikansi parameter dan uji asumsi residual maka hasil estimasi parameter model spline linear dengan satu titik knot terbaik diperoleh persamaan sebagai berikut :

$$y = 110,8963 + 9,1741x_1 + 87,0579x_3 - 0,0764(x_3 - 22,22) - 0,1164(x_4 - 10,27) - 6,6709(x_5 - 4,53)$$

Dari model regresi diatas dapat diartikan bahwa jumlah sarana medis (x_1), Persentase persalinan yang menggunakan tenaga non medis (x_3), Persentase ibu yang tidak melakukan kunjungan kehamilan (x_4), Persentase bayi yang tidak di beri ASI (x_5) memiliki pengaruh terhadap jumlah kematian bayi di Jawa Timur pada tahun 2011.

E.KESIMPULAN

Pemilihan lokasi atau titik knot dalam model spline mampu menginterpretasikan pola data (kurva regresi) sehingga diperoleh model regresi terbaik, dimana model spline dengan pemilihan titik knot optimal mudah menggambarkan pola data sehingga bisa dipahami sifat – sifat datanya dan memiliki perhitungan matematika maupun statistika yang mudah dan sederhana.

Penerapan pada jumlah kematian bayi di Jawa Timur menggambarkan hal itu dimana diperoleh model regresi nonparametrik spline linear dengan satu titik knot optimal sebagai berikut :

$$y = 110,8963 + 9,1741x_1 + 87,0579x_3 - 0,0764(x_3 - 22,22) - 0,1164(x_4 - 10,27) - 6,6709(x_5 - 4,53)$$

Model regresi nonparametrik spline diatas merupakan model yang terbaik dimana nilai $R^2 =$

85,00 % dan nilai GCV minimal 1,00 dengan faktor – faktor yang berpengaruh pada jumlah kematian bayi di Jawa Timur adalah jumlah sarana medis (x_1), Persentase persalinan yang menggunakan tenaga non medis (x_3), Persentase ibu yang tidak melakukan kunjungan kehamilan (x_4), Persentase bayi yang tidak di beri ASI (x_5).

DAFTAR PUSTAKA

[1] Agung, Ngurah IG.1988.*Garis-Patah Paritas*.Yogyakarta: Pusat Penelitian Kependudukan UGM.

[2] Budiantara, I.N.*Model Spline dengan Knot Optimal*.(online:<http://www.resources.id/M0787687623.pdf>).diakses tanggal 27 Juli 2011.

[3] Budiantara, I.N, Dewi, Riana Kurnia. *Faktor-Faktor yang Mempengaruhi Angka Gizi Buruk Di Jawa Timur dengan Pendekatan Regresi NonparametrikSpline*.(online:http://ejurnal.its.ac.id/index.php/sains_seni/article/download/1968/323.pdf). Diakses tanggal 7 Desember 2012.

[4] Budiantara, I.N, Mubarak, Reza. *Analisis Regresi Spline Multivariabel untuk Pemodelan Kematian Penderita Demam Berdarah Dengue (DBD) di Jawa Timur*.(online:http://ejurnal.its.ac.id/index.php/sains_seni/article/download/2025/335.pdf).Diakses tanggal 29 Desember 2012.

[5] Elyna, Arrie Kunilasari.Srinadi, I Gusti Ayu Made.Susilawati, Made.*Pemodelan Angka Kematian Bayi dengan Pendekatan Geographically Weighted Poisson Regression diPropinsiBali*.(online:<http://ojs.unud.ac.id/index.php/mtk/article/download/1790/1101.pdf>). Diakses tanggal 7 Desember 2012.

[6] Eubank, Randall.1988.*Spline Smoothing and Nonparametric Regression*. Marcel Dekker. New York.

[7] Green, P. J. and Silverman, B.W. 1994.*NonParametric Regression and Generalized Linier Models–A Roughness Penalty Approach*. Chapman & Hall. London.

[8] Hardle,Wolfgang.1990.*SmoothingTechniques with Implementation in S*.Springer Verlag. New York.

[9] Sasmitoadi, Didip. *Kajian Penggunaan Knot dan Orde pada Regresi Spline*.(online:http://www.resources.id/makalah/Kajian_Knot_dan_Orde.pdf) diakses tanggal 26 November 2010.

[10] Wahba,Grace.1990.*Spline Models for ObservationalData*.Philadelphia.Pennsylvania.

[11] Wegman, J Edward dkk.*Spline in Statistics*.Journal of the American Statistical Assosiation : America.