Variance
Journal of Statistics and Its Applications

Volume 3 Nomor 1 | Juni 2021
p-ISSN: 2685-8738 e-ISSN: 2685-872X
: 10.30598/variancevol3iss1page1-5

ANALISIS ANGKA KEMATIAN BAYI DI KABUPATEN HALMAHERA UTARA DENGAN METODE REGRESI NONPARAMETRIK *SPLINE TRUNCATED*

Analysis of Infant Mortality in North Halmahera District Using Spline Truncated Nonparametric Regression Method

Meidy Kaseside[1], Samsul Bahri Loklomin[2*]
Program Studi Matematika Universitas Halmahera
Jl. Wari Ino, Kota Tobelo, 97762, Maluku Utara, Indonesia

e-mail: *kasesidemeidy@gmai.com[1],samsul.bahriloklomin@gmail.com[2*]*

Abstrak*: Angka Kematian Bayi merupakan cerminan kelangsungan hidup bayi yang akan menentukan kualitas hidup sumber daya manusia. Provinsi Maluku Utara merupakan Provinsi Kepulauan dengan jumlah Angka Kematian Bayi tergolong tinggi. Masih tingginya Angka Kematian Bayi di Halmahera Utara dipengaruhi oleh beberapa faktor, selain faktor rendahnya pemerataan dan keterjangkauan pelayanan kesehatan, faktor sosial ekonomi juga sangat memengaruhi Angka Kematian Bayi dan balita. Kabupaten Halmahera Utara merupakan salah satu daerah penyumbang Angka Kematian Bayi tertinggi di Provinsi Maluku Utara. Penelitian ini bertujuan untuk menganalisis Angka Kematian Bayi di Halmahera Utara secara matematika. Hasil penelitian menunjukkan bahwa semua variabel prediktor berpengaruh terhadap AKB. Hasil analisis didapati model terbaik dengan tiga titik knot dengan koefisen determinasi 97,92% dan nilai GCV minimum 1,52.*

Kata Kunci*: Angka Kematian Bayi, Regresi Nonparametrik, Spline Truncated*

Abstract*: Infant Mortality Rate is a reflection of infant survival which will determine the quality of life of human resources. North Maluku Province is an archipelago province with a relatively high infant mortality rate. The still high infant mortality rate in North Maluku is influenced by several factors, apart from the low equity and affordability of health services, socio-economic factors also greatly affect infant and under-five mortality rates. North Halmahera District is one of the highest IMR contributors in North Maluku Province. This study aims to mathematically analyze the infant mortality rate in North Halmahera. The results showed that all predictor variables had an effect on infant mortality. The analysis results found the best model with three point knots with a determination coefficient of 97.92% and a minimum GCV value of 1,52.*

Keywords*: Infant Mortality, Nonparametric Regression, Spline Truncated*

1. PENDAHULUAN

Kematian bayi merupakan salah satu indikator dalam menentukan derajat kesehatan. Apabila suatu daerah memiliki kematian bayi yang tinggi maka dapat dikatakan tingkat kesehatan anak pada daerah tersebut rendah. Kematian bayi adalah kematian yang terjadi saat setelah bayi lahir sampai bayi belum berusia tepat satu tahun [1]. Salah satu faktor tingginya Angka Kematian Bayi di Indonesia adalah masih rendahnya cakupan pertolongan oleh tenaga kesehatan. Rendahnya tingkat pendidikan di masyarakat, budaya dan ekonomi menjadikan sebagian masyarakat memilih bersalin pada tenaga non kesehatan salah satu contohnya adalah dukun. Pertolongan persalinan oleh tenaga kesehatan merupakan strategi untuk menangani masalah kesehatan ibu dan anak di Indonesia. Kemampuan dan keterampilan penolong persalinan sangat memengaruhi terjadinya kematian ibu maupun bayi.

Maluku Utara merupakan Provinsi Kepulauan dengan Angka Kematian Bayi masih tergolong tinggi jika dibandingkan daerah lain di Indonesia. Masih terdapatnya Kabupaten/Kota yang memiliki Angka Kematian Bayi dan status gizi buruk balita tinggi tentunya tidak terlepas dari faktor-faktor yang mempengaruhi.

Copyright© 2021: Author
https://ojs3.unpatti.ac.id/index.php/variance/
Submitted: 2020-12-09 *Accepted::2021-12-15*

Kabupaten Halmahera Utara salah satu daerah penyumbang Angka Kematian Bayi tertinggi di Provinsi Maluku Utara. Pada tahun 2017 terdapat 3.351 kelahiran hidup di Kabupaten Halmahera Utara dengan jumlah Angka Kematian Bayi sebanyak 47 bayi, sedangkan jumlah balita 17.390 dengan kasus gizi buruk sebanyak 40 bayi [2]. Tujuan penelitian ini adalah untuk menganalisis Angka Kematian Bayi Kabupaten Halmahera Utara dengan metode dalam statistika. Salah satu metode yang digunakan adalah analisis regresi.

Analisis regresi merupakan analisis statistika yang digunakan untuk memodelkan pola hubungan antara variabel prediktor dan variabel respon. Pemodelan regresi dibagi menjadi tiga yaitu metode regresi parametrik, regresi nonparametrik, dan regresi semiparametrik. Pendekatan regresi parametrik digunakan jika bentuk kurva regresi diketahui. Jika pola hubungan data membentuk pola linear maka digunakan pendekatan regresi parametrik linear. Jika pola hubungan data membentuk pola kuadrat maka digunakan pendekatan regresi kuadratik, dan lain-lain. Bentuk pola hubungan dapat diidentifikasi berdasarkan pada informasi masa lalu atau *scatterplot* data. Pada regresi tidak semua variabel penjelas dapat didekati dengan pendekatan parametrik. Oleh karena tidak adanya informasi tentang bentuk hubungan variabel penjelas tersebut dengan variabel responnya, sehingga harus digunakan pendekatan nonparametrik. Dengan menggabungkan dua pendekatan tersebut dalam suatu pendekatan regresi akan didapatkan suatu model semiparametrik. Regresi semiparametrik digunakan jika salah satu kurva regresinya tidak diketahui sedangkan yang lainnya diketahui. Regresi nonparametrik merupakan pendekatan metode regresi dimana bentuk kurva dari fungsi regresinya tidak diketahui. Kurva fungsi diasumsikan termuat dalam ruang fungsi tertentu. Model regresi nonparametrik diberikan oleh persamaan berikut.

$$y_i = f(x_i) + \varepsilon \tag{1}$$

Spline dalam regresi nonparametrik mempunyai sifat fleksibilitas yang tinggi dan mempunyai kemampuan mengestimasi perilaku data yang cenderung berbeda pada interval yang berlainan. Kemampuan mengestimasi perilaku data ini ditunjukkan oleh fungsi *truncated* (potongan-potongan) yang melekat pada estimator dan potongan-potongan tersebut yang disebut titik knot. Titik knot merupakan titik perpaduan bersama yang menunjukkan perubahan pola perilaku fungsi pada selang yang berbeda [3]. Spline merupakan salah satu jenis *piecewise* polinomial, yaitu polinomial yang memiliki sifat tersegmen. Sifat tersegmen ini memberikan fleksibilitas lebih dari polinomial biasa, sehingga memungkinkan untuk menyesuaikan diri secara lebih efektif terhadap karakteristik lokal suatu fungsi atau data [4]. *Spline Truncated* merupakan model yang mempunyai interpretasi statistik dan interpretasi visual khusus. Kelebihan dari model *Spline Truncated* adalah jika data yang memiliki pola berubah-ubah pada sub-sub interval tertentu maka dapat dimodelkan dengan *Spline Truncated*. Ada beberapa metode yang dapat digunakan untuk memilih titik knot optimal dalam estimator Spline *Truncated*, antara lain metode *Cross Validation* (CV), *Unbiased Risk* (UBR), dan *Generalized Cross Validation* (GCV). Salah satu metode yang sering digunakan dalam pemilihan titik knot optimal adalah GCV. Metode GCV memberikan beberapa kelebihan dibandingkan metode lainnya, diantaranya memiliki sifat optimal asimptotik, invarian terhadap transformasi dan dalam perhitungannya varians populasi tidak perlu diketahui. Berdasarkan uraian tersebut, permasalahan dalam penelitian ini bagaimana memodelkan Angka Kematian Bayi dengan pendekatan estimasi interval parameter model regresi semiparametrik *Spline Truncated*. Manfaat yang dicapai dari penelitian ini adalah menganalisis angka kematian bayi dengan model regresi di Kabupaten Halmahera Utara dengan batasan masalah pada penelitian ini adalah pemilihan titik knot optimal menggunakan GCV satu knot, dua knot, tiga knot, dan kombinasi knot.

2. METODOLOGI

2.1. Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder bersumber dari Dinas Kesehatan Kabupaten Halmahera Utara dan Badan Pusat Statistik Kabupaten Halmahera Utara pada setiap kecamatan di Kabupaten Halmahera Utara [5].

2.2. Variabel Penelitian

Variabel yang digunakan dalam penelitian ini yaitu kejadian Angka Kematian Bayi di Kabupaten Halmahera Utara. Variabel prediktor dalam penelitian ini adalah multivariabel yaitu variabel-variabel yang mempengaruhi kejadian Angka Kematian Bayi. Variabel respon dan variabel prediktor dalam penelitian ini sebagai berikut:

1. Angka Kematian Bayi
2. Persentase bayi yang diberi ASI

3. Persentase Bayi Berat Badan Lahir Rendah

2.3. Tahapan Penelitian

Tahapan penelitian dalam analisis Angka Kematian Bayi di Kabupaten Halmahera Utara menggunakan metode regresi nonparametrik *Spline truncated* sebagai berikut:

1) Deskriptif data penelitian.

   Pada tahapan ini disekripsikan data penelitian yang menjadi fokus penelitian berupa intrepretasi sederhana.

2) Memodelkan data menggunakan regresi nonparametrik *Spline truncated*.

   Pada tahapan ini data penelitian dimodelkan dengan regresi nonparametrik *Spline truncated* dimana *Spline* yang digunakan adalah satu knot, dua knot, tiga knot, dan kombinasi knot.

3) Memilih titik knot optimal dengan metode GCV.

   Salah satu metode untuk mencari titik knot optimal yang sering digunakan adalah dengan metode GCV. Titik optimal diperoleh dari nilai GCV minimum. Metode GCV secara umum adalah sebagai berikut:

$$GCV(K) = \frac{MSE(K)}{\left[n^{-1}trace(I - A(K))\right]^2} \tag{2}$$

   dengan $MSE(K) = n^{-1}\sum_{i=1}^{n}\left(y_i - \hat{f}(x_i)\right)^2$, I adalah matriks identitas, n adalah jumlah pengamatan, $A(K)$ adalah matriks $\boldsymbol{X(X'X)^{-1}X'}$.

4) Menetapkan model terbaik dari nilai GCV terkecil.

   Pemilihan model regresi *Spline* yang terbaik dipilih berdasarkan lokasi dan banyaknya titik knot. Dalam pemilihan titik knot yang optimal sangatlah penting dalam menentukan model terbaik. Lokasi titik knot yang berbeda akan menghasilkan model yang berbeda. Titik knot yang optimal dilihat berdasarkan kriteria GCV yang paling minimum.

5) Memeriksa asumsi residual yang diisyaratkan yaitu asumsi identik, independen dan normalitas.

   Suatu model dengan kriteria terbaik adalah jika residual memenuhi asumsi klasik yaitu asumsi identik, independen dan berdistrbusi normal. Pengujian asumsi identik menggunakan uji varian dan pengujian asumsi independen menggunakan *Run Test* serta pengujian residual berditribusi normal menggunakan *Kolmogorov-Smirnov.*

6) Menghitung MSE dan R^2 sebagai bagian dari kriteria kebaikan model.

   Dalam pemodelan perlu dilakukan pengecekan apakah model tersebut baik atau tidak. Untuk melihat model terbaik pada umumnya lebih melihat nilai koefiien determinasi atau *R-square*. Koefisien determinasi mengintrepretasikan bahwa seberapa besar variabel prediktor memengaruhi Angka Kematian Bayi.

7) Mengambil kesimpulan.

3. HASIL DAN PEMBAHASAN

Kabupaten Halmahera Utara adalah salah satu daerah penyumbang kematian bayi tertinggi di Provinsi Maluku Utara. Pada tahun 2017, Dinas Kesehatan mencatat terdapat 3.351 kelahiran hidup di Kabupaten Halmahera Utara dengan jumlah kematian bayi sebanyak 47 bayi, sedangkan jumlah balita 17.390 dengan kasus gizi buruk sebanyak 40 bayi. Kabupaten Halmahera Utara dengan 17 kecamatan memiliki jumlah sarana kesehatan memadai sebanyak 21 yang terdiri dari 2 rumah sakit dan 19 puskesmas. Namun wilayah Halmahera Utara juga merupakan wilayah dengan akses ke sarana kesehatan yang termasuk sulit. Dari 17 kecamatan terdapat dua kecamatan yang tergolong sebagai kecamatan yang memiliki akses kesehatan yang sulit. Hal ini disebabkan karena dua kecamatan tersebut terletak jauh dari pusat pemerintahan Kabupaten Halmahera Utara dan salah satu dari kecamatan yakni Kecamatan Loloda Kepulauan letaknya terpisah dengan pulau Halmahera.

Dengan sulitnya akses ke sarana kesehatan hal ini mengakibatkan jumlah kematian ibu dan bayi semakin banyak di Kabupaten Halmahera Utara.

3.1. *Spline Truncated* Dalam Regresi Nonparametrik

Secara umum, fungsi *Spline* linier dengan titik-titik knot $K_1, K_2, .., K_r$ adalah suatu fungsi yang dapat ditulis dalam bentuk sebagai berikut:

$$y_i = f(x_i) + \varepsilon_i,\ i = 1,2,..,n$$
$$= \beta_0 + \sum_{j=1}^{p}\left(\beta_{1j}x_{ij} + \sum_{k=1}^{r}\beta_{(1+j)k}\left(x_{ij} - K_{kr}\right)_+\right) + \varepsilon_i \quad (3)$$

dan fungsi *truncated* diberikan oleh:

$$\left(x_{ij} - K_j\right)_+ = \begin{cases}\left(x_{ij} - K_{kr}\right), x_{ij} \geq K_{kr}\\ 0 \qquad\quad , x_{ij} < K_{kr}\end{cases} \quad (4)$$

Kurva regresi $f(x_i)$ merupakan kurva regresi nonparametrik *Spline truncated* derajat satu dengan banyaknya titik knot *r*.

3.2. Pemilihan Titik Knot Optimum Dalam *Spline Truncated*

Titik knot merupakan titik perubahan perilaku data pada sub-sub interval tertentu. Model regresi nonparametrik *Spline truncated* terbaik didapatkan dari titik knot yang optimal. Untuk mendapatkan titik knot yang optimal digunakan metode *Generalize Cross Validation* (GCV). Metode GCV digunakan untuk menentukan titik knot dengan beberapa kombinasi titik yaitu pengujian dengan satu titik knot, dua titik knot, tiga titik knot dan kombinasi dari ketiga titik tersebut. Dalam penentuan titik knot optimal didasari pada nilai GCV yang minimum dari pengujian dengan satu titik knot, dua titik knot, tiga titik knot dan kombinasi dari ketiga titik tersebut. Berikut ini merupakan tabel perbandingan nilai GCV untuk masing-masing model.

Tabel 1. Perbandingan Nilai GCV untuk Pemilihan Model

Knot	GCV	*R-Square*
1 titik knot	10,548	95,16%
2 titik knot	**1,52**	**97,92%**
3 titik knot	5,309	92,99%
Kombinasi Knot	10,075	97,97%

Pada Tabel 1 diatas diperoleh knot dengan GCV minimum 1,52 berada pada dua titik knot. Hal ini menyatakan bahwa knot dengan dua titik adalah model yang terbaik dengan koefisien determinasi yang diperoleh 97,92%. Koefisien determinasi 97,92% dapat disimpulkan bahwa 97,92% variabel prediktor mampu menjelaskan varians dari variabel respon. Terdapat 2,08% varians yang tidak dapat dijelaskan dalam penelitian ini atau 2,08% varians yang dijelaskan oleh faktor lain. Dari titik knot optimal terpilih yaitu dua titik knot maka model persamaan regresi nonparametrik *Spline truncated* linier yang terbentuk adalah:

$$\hat{y} = 11{,}23 - 19{,}05X_1 + 3{,}52\,(X_1 - 54{,}32) + 5{,}27\,(X_1 - 91{,}24) + 27{,}54\,X_2 + 5{,}27\,(X_2 - 86{,}33) + 5{,}27\,(X_2 - 67{,}32) \quad (5)$$

Pada model persamaan yang terbentuk diatas dapat disimpulkan bahwa jika terjadi kenaikan persentase bayi yang diberi ASI maka Angka Kematian Bayi akan menurun dan jika persentase Bayi Berat Badan Lahir Rendah semakin turun maka akan terjadi penuruan Angka Kematian Bayi.

3.3. Pengujian Asumsi Residual

Suatu model dengan kriteria terbaik adalah jika residual memenuhi klasik yaitu asumsi identik, independen dan berdistrbusi normal. Langkah pertama dalam pengujian asumsi identik dengan tujuan tidak terdapat heteroskedastisitas . Pengujian asumsi identik dilakukan dengan hipotesis:

H_0: $\sigma^2{}_1 = \sigma^2{}_2 = \cdots = \sigma^2{}_n = \sigma^2$

H_0: minimal ada satu $\sigma^2{}_i \neq \sigma^2$

Pada uji identik, nilai p-*value* 0, 21 lebih besar dari $\alpha = 0{,}05$, maka dapat disimpulkan bahwa tidak terdapat kasus heteroskedastisitas, jadi varians residual memenuhi asumsi identik. Selanjutnya adalah pengujian asumsi independen. Pengujian asumsi independen menggunakan *Run Test* dengan hipotesis:

H_0: Residual random

H_1: Residual tidak random

dengan *Run Test* terhadap residual menghasilkan nilai 1,00 lebih besar dari $\alpha = 0{,}05$ maka gagal tolak H_0 sehingga disimpulkan residual bersifat random atau tidak ada korelasi antar residual, maka residual memenuhi asumsi independen. Langkah selanjutnya atau langkah terakhir dilakukan pengujian asumsi residual. Pengujian ini mengharuskan residual berdistribusi normal dengan hipotesis:

$H_0: F_0(x) = F(x)$ (Residual berdistribusi normal)

$H_1: F_1(x) \neq F(x)$ (Residual tidak berdistribusi normal)

Pada pengujian residual berditribusi normal menggunakan *Kolmogorov-Smirnov* atau K-S. Hasil uji K-S diperoleh p-*value* $> 0{,}167$ lebih besar dari 0,05 maka residual berdistribusi normal, maka asumsi residual berdistribusi normal terpenuhi. Pengujian residual telah memenuhi ketentuan tiga syarat klasik tersebut sehingga model tersebut adalah model terbaik karena residual memenuhi asumsi identik, independen dan berdistrbusi normal.

4. KESIMPULAN

Angka Kematian Bayi masih dipengaruhi oleh akses sarana kesehatan dan faktor pengetahuan atau tingkat pendidikan masyarakat di Kabupaten Halmahera Utara. Pada penelitian Angka Kematian Bayi di Kabupaten Halmahera Utara diperoleh model terbaik dengan dua titik knot. Model terbaik dengan dua titik knot dengan nilai GCV minimum 1,52 dan koefisien determinasi 97,92%. Pada model regresi nonparametrik *Spline truncated* terbentuk tersebut bermakna bahwa semua variabel prediktor berpengaruh terhadap Angka Kematian Bayi yang mampu dijelaskan sebesar 97,92% dan terdapat 2,08% faktor lain di luar model.

DAFTAR PUSTAKA

[1] Aulele, S. N. (2012). *Pemodelan Jumlah Kematian Bayi Di Provinsi Maluku Tahun 2010 dengan Menggunakan Regresi Poisson*. Jurnal Barekeng, 23 – 27.

[2] Dinkes Kabupaten Halmahera Utara. *Laporan Akuntabilitas Kinerja Instansi Pemerintah.* Tobelo: 2018

[3] Budiantara, I.N., *Spline dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang*. Surabaya: ITS Press. 2009.

[4] Budiantara, I.N., *Model Spline dengan Knots Optimal*. Jurnal FMIPA Universitas Jember, Jember: 2006

[5] Dinkes Kabupaten Halmahera Utara, *Laporan Kesehatan Kabupaten Halmahera Utara*, BPS,Tobelo: 2019.

[6] Eubank, R. L. & Speckman, P L., *Confidence bands in Nonparametric Regression*. Journal Am. Statist. Ass., 88, 1287-1301, New York, 1993.

[7] Intansari, I.A., *Inferensi Statistik Untuk Kurva Regresi Nonparametrik Spline Kuadratik dan Aplikasinya Pada Data ASFR Di Provinsi Bali*, Tesis FMIPA ITS, Surabaya, 2016

[8] Loklomin, S.B., Budiantara, I.N., & Zain,I.,*Estimation of Interval Parameters Semiparametric Regression Models Spline Truncated*. Thesis Departement Of Statistics. ITS Surabaya, 2017.

[9] Okumura, H & Naito, K., *Non-Parametrik Kernel Regression for Multinomial Data,* Journal of Multivariate Analysis, Vol. 97, No. 9, hal. 2009-2022, 2006