ESTIMASI MODEL REGRESI NONPARAMETRIK MENGGUNAKAN METODE SPLINE KUADRATIK

(Skripsi)

Oleh

LELVI FANESSA HALOHO

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2019

ABSTRAK

ESTIMASI MODEL REGRESI NONPARAMETRIK MENGGUNAKAN METODE SPLINE KUADRATIK

Oleh

Lelvi Fanessa Haloho

Spline merupakan potongan polinomial tersegmen yang memiliki sifat fleksibilitas yang tinggi, titik perpaduan bersama dari potongan –potongan tersebut atau titik yang menunjukan terjadinya perubahan prilaku kurva pada interval – interval yang berbeda disebut knot. Tujuan dari penelitian ini untuk mengestimasi model regresi nonparametrik dengan menggunakkan metode *spline* kuadratik pada data simulasi. Hasil penelitian menujukan spline kuadratik baik dalam mengestimasi fungsi kuadratik tetapi tidak untuk model linier dan kubik.

Kata kunci: Regresi Nonparametrik , *spline* , knot , GCV.

ABSTRACT

ESTIMATION OF NONPARAMETRIC REGRESSION MODEL USING QUADRATIC SPLINE METHOD

By

Lelvi Fanessa Haloho

Spline is a segmented polynomial piece that has high flexibility properties, the joint fusion point of the pieces or points that indicate changes in the behavior of the curve at different intervals called knots. The purpose of this research is to estimate a nonparametric regression model using the quadratic spline method on simulation data. The results show that quadratic spline is good in estimating the quadratic function model but not for linear and cubic model.

Keywords: *Spline*, knot, GCV, quadratic, linear, cubic.

ESTIMASI MODEL REGRESI NONPARAMETRIK MENGGUNAKAN METODE SPLINE KUADRATIK

Oleh

LELVI FANESSA HALOHO

Skripsi

Sebagai Salah Satu Syarat Untuk Memperoleh Gelar
SARJANA SAINS

Pada

Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2019

Judul Skripsi : **ESTIMASI MODEL REGRESI NONPARAMETRIK MENGGUNAKAN METODE SPLINE KUADRATIK**

Nama Mahasiswa : *Lelvi Fanessa Haloho*

No. Pokok Mahasiswa : 1517031097

Jurusan : Matematika

Fakultas : Matematika dan Ilmu Pengetahuan Alam

Menyetujui

1. Komisi Pembimbing

Ir. Netti Herawati, M.Sc., Ph.D.	**Dr. La Zakaria, S.Si., M.Sc.**
NIP. 196501251990032001	NIP. 196902131994021001

2. Ketua Jurusan

Prof. Dra. Wamiliana, M.A., Ph.D
NIP. 196311081989022001

MENGESAHKAN

1. Tim Penguji

   Ketua : **Ir. Netti Herawati, M.Sc., Ph.D.** ....

   Sekretaris : **Dr. La Zakaria, S.Si., M.Sc.** ....

   Penguji

   Bukan Pembimbing : **Drs. Nusyirwan, M.Si.** ....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Drs. Suratman, M.Sc.
NIP 19640604 199003 1 002

Tanggal Lulus Ujian Skripsi : **05 Agustus 2019**

PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini :

Nama Mahasiswa : **LELVI FANESSA HALOHO**

No. Pokok Mahasiswa : **1517031097**

Jurusan : **Matematika**

Judul Skripsi : **ESTIMASI MODEL REGRESI NONPARAMETRIK MENGGUNAKAN METODE SPLINE KUADRATIK**

Dengan ini menyatakan bahwa penelitian ini adalah hasil pekerjaan saya sendiri dan apabila kemudian hari terbukti bahwa skripsi ini merupakan hasil Salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 05 Agustus 2019
Penulis,

Lelvi Fanessa Haloho
NPM. 1517031097

RIWAYAT HIDUP

Penulis bernama lengkap Lelvi Fanessa Haloho, anak ketiga dari lima bersaudara yang dilahirkan di Bakauheni pada tanggal 06 Februari 1997 oleh pasangan Bapak Julianus Sihaloho dan Ibu Hotnauli Sijabat.

Penulis menempuh pendidikan di Taman Kanak-Kanak (TK) Bhakti Ibu pada tahuan 2002-2003, kemudian bersekolah di SD Bhakti Ibu pada tahun 2003-2009, kemudian melanjutkan sekolah di SMP Negeri 1 Bakauheni pada tahun 2009-2012, dan pada tahun 2012-2015 menempuh sekolah menengah atas di SMA Xaverius Pringsewu.

Pada tahun 2015, penulis terdaftar sebagai mahasiswa S1 jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam melalui jalur Mandiri. Pada bulan Januari-Februari tahun 2018 selama 30 hari melakukan Kerja Praktik (KP) di PT. Asuransi Jiwasraya (Persero) kantor cabang Bandar Lampung. Pada bulan Juli-Agustus tahun 2018, selama 40 hari penulis melakukan Kuliah Kerja Nyata (KKN) periode kedua di Desa Sukacari Kec. Batanghari Nuban.

MOTTO

“ Ingatlah selalu akan Dia, yang tekun menanggung bantahan yang sehebat itu terhadap diri-Nya dari pihak orang – orang berdosa, supaya jangan kamu menjadi lemah dan putus asa”

(Ibrani 12:3)

“Jangan selalu lihat keatas karna kau akan terjatuh, dan jangan selalu lihat kebawah karna kau akan kehilangan arah”

(BABE)

“Jangan terbuai dengan pujian, karna banyak nyamuk mati karena tepuk tangan”

(ANONIM)

PERSEMBAHAN

Kuucapkan syukurku dan terimakasihku untuk TUHAN YESUS KRISTUS

yang selalu menyertaiku untuk setiap proses jatuh bangun dalam perjuanganku

menyelesaikan skripsi.

Kepada Ibu yang ada di surga yang sangat kurindukan, dan Ayah yang selalu

kuandalkan dalam segala hal, terima kasih atas segala kasih sayang dan doa yang

kalian berikan, ***kupersembahkan sebuah karya sederhana ini yang menjadi bukti***

kudapatkan gelar, Lelvi Fanessa Haloho, S.Si .

semua usaha dan kerja keras kulakukan untuk membanggakan ibu dan ayah, semoga

dengan karya kecil ku ini bisa menambah kebahagiaan kalian

Kepada saudara – saudariku (Bang Lio, Ka cin, Rabekka , dan kesayangan kami

ZIO) yang menjadi penyemangat bagi penulis.

Serta,

Almamater tercinta, Universitas Lampung.

SANWACANA

Puji syukur kepada Tuhan Yang Maha Esa yang telah memberikan karunia serta kasih-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul "Estimasi Model Regresi Nonparametrik Menggunakan Metode Spline Kuadratik". Skripsi ini disusun sebagai salah satu syarat memperoleh gelar Sarjana Sains (S.Si.) pada Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

Dengan setulus hati penulis ingin mengucapkan terima kasih kepada :

1. Ibu Ir. Netti Herawati, M.Sc., Ph.D. selaku pembimbing utama atas kesediaan waktu, pemikiran dalam memberikan evaluasi, arahan dan saran yang membangun dalam proses penyusunan skripsi ini.
2. Bapak Dr. La Zakaria, S.Si., M.Sc. selaku pembimbing kedua atas kesediaan waktu dan saran serta kritik dalam proses penyusunan skripsi.
3. Bapak Drs. Nusyirwan, M.Si. selaku dosen penguji atas kesediannya menguji dan memberikan saran serta kritik yang membangun dalam proses penyusuna skripsi ini.
4. Bapak Drs. Eri Etiawan, M.Si. selaku pembimbing akademik yang telah memberikan arahan dan nasihat kepada penulis selama proses perkuliahan.
5. Ibu Dra. Wamiliana, M.A., Ph.D. selaku Ketua Jurusan Matematika atas izin dan bantuan selama masa pendidikan.

6. Bapak Drs. Suratman, M.Sc., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Keluarga Besar, terkhusus untuk Bapak, Mama, Bang Lio, Ka Cin, Rabekka, Zio. Untuk setiap dukungan begitu berdampak kepada semangatku menyelesaikan pendidikan ku di perkuliahan.
8. Sahabat-sahabatku : Abraham, Winda, Juni, Gina, Hani, Nda, Eva, Vina, Agung, Mute, Sela, Lipe, Wilma, Lita, Dino, Aldo, Rapli, Ikhsan, Debora, Rani, Hana, Kakak Yunitri, Kakak Enyka, KP (Agung, Dian, Birgita) , KKN (Poyenk, Andi, Candra, Deka, Anggita, Rohma). yang tetap setia dari masa seminar kuliah praktek hingga masa sidang skripsi memberikan doa, semangat dan bantuan dalam proses skripsweet ini. Terimakasih, aku mengasihi Kalian.
9. Angkatan 2015 Matematika UNILA serta kakak dan adik tingkat yang membuatku merasakan indahnya dan serunya dunia perkuliahan.
10. Serta semua pihak yang tak bisa disebutkan satu persatu yang telah menjadi bagian penting dari masa awal hingga akhir perkuliahanku.

Penulis berharap semoga Tuhan Yang Maha Esa membalas kebaikan kalian semua.

Bandar Lampung, 05 Agustus 2019
Penulis,

Lelvi Fanessa Haloho

DAFTAR ISI

DAFTAR TABEL i

DAFTAR GAMBAR.. iii

I. PENDAHULUAN 1

1.1 Latar Belakang dan Masalah 1

1.2 Tujuan Penelitian 2

1.3 Manfaat Penelitian... 3

II. TINJAUAN PUSTAKA 4

2.1 Analisis Regresi 4

2.2 Kegunaan Analisis Regresi 5

2.3 Regresi Parametrik 5

2.4 Regresi Nonparametrik 6

2.5 Spline.. 7

2.6 Regresi Spline.. 8

2.7 Estimasi Parameter Regresi Nonparametrik Spline.. 9

2.8 Estimasi Kurva Regresi Nonprametrik Spline.. 11

2.9 Pemilihan Titik Knot Optimal 12

2.10 Koefisien Determinasi 13

III. METODOLOGI PENELITIAN 14

3.1 Waktu Dan Tempat Penelitian 14

3.2 Data Penelitian 14

3.3 Metode Penelitian.. 15

IV. HASIL DAN PEMBAHASAN 16
4.1 Plot Data Pengamatan.. 16
4.2 Pemilihan Titik Knot 23
4.2.1 Menentukan Nilai GCV Pada Fungsi Linear $y = 2x + 1 + \varepsilon_1$ 24
4.2.2 Menentukan Nilai GCV Pada Fungsi Linear $y = 2x + 1 + \varepsilon_2$ 26
4.2.3 Menentukan Nilai GCV Pada Fungsi Linear $y = 2x + 1 + \varepsilon_3$ 28
4.2.4 Menentukan Nilai GCV Pada Fungsi Linear $y = 2x + 1 + \varepsilon_4$ 30
4.2.5 Menentukan Nilai GCV Pada Fungsi Linear $y = 2x + 1 + \varepsilon_5$ 32
4.2.6 Menentukan Nilai GCV Pada Fungsi Kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_1$ 34
4.2.7 Menentukan Nilai GCV Pada Fungsi Kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_2$ 36
4.2.8 Menentukan Nilai GCV Pada Fungsi Kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_3$ 38
4.2.9 Menentukan Nilai GCV Pada Fungsi Kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_4$ 40
4.2.10 Menentukan Nilai GCV Pada Fungsi Kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_5$ ... 42
4.2.11 Menentukan Nilai GCV Pada Fungsi kubik $y = sin(4x) + \varepsilon_1$ 44
4.2.12 Menentukan Nilai GCV Pada Fungsi kubik $y = sin(4x) + \varepsilon_2$ 46
4.2.13 Menentukan Nilai GCV Pada Fungsi kubik $y = sin(4x) + \varepsilon_3$ 48
4.2.14 Menentukan Nilai GCV Pada Fungsi kubik $y = sin(4x) + \varepsilon_4$ 50
4.2.15 Menentukan Nilai GCV Pada Fungsi kubik $y = sin(4x) + \varepsilon_5$ 52
4.2.16 Menentukan Nilai GCV Pada Fungsi kubik $y = cos(4x) + \varepsilon_1$ 54
4.2.17 Menentukan Nilai GCV Pada Fungsi kubik $y = cos(4x) + \varepsilon_2$ 56
4.2.18 Menentukan Nilai GCV Pada Fungsi kubik $y = cos(4x) + \varepsilon_3$ 58
4.2.19 Menentukan Nilai GCV Pada Fungsi kubik $y = cos(4x) + \varepsilon_4$ 60
4.2.20 Menentukan Nilai GCV Pada Fungsi kubik $y = cos(4x) + \varepsilon_5$ 62

V. KESIMPULAN.. 63

DAFTAR PUSTAKA 64

LAMPIRAN

DAFTAR TABEL

Tabel Halaman

4.1 Nilai GCV pada fungsi linear $y = 2x + 1 + \varepsilon_1$ 24

4.2 Estimasi model fungsi linear $y = 2x + 1 + \varepsilon_1$ dengan tiga titik *knot*........ 24

4.3 Nilai GCV pada fungsi linear $y = 2x + 1 + \varepsilon_2$ 26

4.4 Estimasi model fungsi linear $y = 2x + 1 + \varepsilon_2$ dengan tiga titik *knot*........ 26

4.5 Nilai GCV pada fungsi linear $y = 2x + 1 + \varepsilon_3$ 28

4.6 Estimasi model fungsi linear $y = 2x + 1 + \varepsilon_3$ dengan tiga titik *knot*........ 28

4.7 Nilai GCV pada fungsi linear $y = 2x + 1 + \varepsilon_4$ 30

4.8 Estimasi model fungsi linear $y = 2x + 1 + \varepsilon_4$ dengan tiga titik *knot*........ 30

4.9 Nilai GCV pada fungsi linear $y = 2x + 1 + \varepsilon_5$ 32

4.10 Estimasi model fungsi linear $y = 2x + 1 + \varepsilon_5$ dengan tiga titik *knot*........ 32

4.11 Nilai GCV pada fungsi kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_1$ 34

4.12 Estimasi model fungsi kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_1$ dengan tiga titik *knot*.. 34

4.13 Nilai GCV pada fungsi kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_2$ 36

4.14 Estimasi model fungsi kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_2$ dengan tiga titik *knot*.. 36

4.15 Nilai GCV pada fungsi kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_3$ 38

4.16 Estimasi model fungsi kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_3$ dengan tiga titik *knot*.. 38

4.17 ilai GCV pada fungsi kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_4$ 40

4.18 Estimasi model fungsi kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_4$ dengan tiga titik *knot* 40

4.19 Nilai GCV pada fungsi kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_5$ 42

4.20 Estimasi model fungsi kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_5$ dengan tiga titik *knot* 42

4.21 Nilai GCV pada fungsi kubik $y = sin(4x) + \varepsilon_1$ 44

4.22 Estimasi model fungsi kubik $y = sin(4x) + \varepsilon_1$ dengan tiga titik *knot* 44

4.23 Nilai GCV pada fungsi kubik $y = sin(4x) + \varepsilon_2$ 46

4.24 Estimasi model fungsi kubik $y = sin(4x) + \varepsilon_2$ dengan tiga titik *knot* 46

4.25 Nilai GCV pada fungsi kubik $y = sin(4x) + \varepsilon_3$ 48

4.26 Estimasi model fungsi kubik $y = sin(4x) + \varepsilon_3$ dengan tiga titik *knot* 48

4.27 Nilai GCV pada fungsi kubik $y = sin(4x) + \varepsilon_4$ 50

4.28 Estimasi model fungsi kubik $y = sin(4x) + \varepsilon_4$ dengan tiga titik *knot* 50

4.29 Nilai GCV pada fungsi kubik $y = sin(4x) + \varepsilon_5$ 52

4.30 Estimasi model fungsi kubik $y = sin(4x) + \varepsilon_5$ dengan tiga titik *knot* 52

4.31 Nilai GCV pada fungsi kubik $y = \ cos\ (4x) + \varepsilon_1$ 54

4.32 Estimasi model fungsi kubik $y = \ cos\ (4x) + \varepsilon_1$ dengan tiga titik *knot* 54

4.33 Nilai GCV pada fungsi kubik $y = \ cos\ (4x) + \varepsilon_2$ 56

4.34 Estimasi model fungsi kubik $y = \ cos\ (4x) + \varepsilon_2$ dengan tiga titik *knot* 56

4.35 Nilai GCV pada fungsi kubik $y = \ cos\ (4x) + \varepsilon_3$ 58

4.36 Estimasi model fungsi kubik $y = \ cos\ (4x) + \varepsilon_3$ dengan tiga titik *knot* 58

4.37 Nilai GCV pada fungsi kubik $y = \ cos\ (4x) + \varepsilon_4$ 60

4.38 Estimasi model fungsi kubik $y = \ cos\ (4x) + \varepsilon_4$dengan tiga titik *knot* 60

4.39 Nilai GCV pada fungsi kubik $y = \ cos\ (4x) + \varepsilon_5$ 62

4.40 Estimasi model fungsi kubik $y = \ cos\ (4x) + \varepsilon_5$ dengan tiga titik *knot* 62

DAFTAR GAMBAR

Gambar Halaman

4.1 Plot data $X \sim U(0,3)$ dengan fungsi linear $y = 2x + 1 + \varepsilon_1$ 16

4.2 Plot data $X \sim U(0,3)$ dengan fungsi linear $y = 2x + 1 + \varepsilon_2$ 17

4.3 Plot data $X \sim U(0,3)$ dengan fungsi linear $y = 2x + 1 + \varepsilon_3$ 17

4.4 Plot data $X \sim U(0,3)$ dengan fungsi linear $y = 2x + 1 + \varepsilon_4$ 17

4.5 Plot data $X \sim U(0,3)$ dengan fungsi linear $y = 2x + 1 + \varepsilon_5$ 18

4.6 Plot data $X \sim U(0,3)$ dengan fungsi kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_1$ ... 18

4.7 Plot data $X \sim U(0,3)$ dengan fungsi kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_2$ ... 18

4.8 Plot data $X \sim U(0,3)$ dengan fungsi kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_3$ ... 19

4.9 Plot data $X \sim U(0,3)$ dengan fungsi kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_4$ ... 19

4.10 Plot data $X \sim U(0,3)$ dengan fungsi kuadratik $y = 2x^2 - 6x + 16 + \varepsilon_5$ ... 19

4.11 Plot data $X \sim U(0,3)$ dengan fungsi kubik $y = \sin(4x) + \varepsilon_1$ 20

4.12 Plot data $X \sim U(0,3)$ dengan fungsi kubik $y = \sin(4x) + \varepsilon_2$ 20

4.13 Plot data $X \sim U(0,3)$ dengan fungsi kubik $y = \sin(4x) + \varepsilon_3$ 20

4.14 Plot data $X \sim U(0,3)$ dengan fungsi kubik $y = \sin(4x) + \varepsilon_4$ 21

4.15 Plot data $X \sim U(0,3)$ dengan fungsi kubik $y = \sin\,(4x) + \varepsilon_5$ 21

4.16 Plot data $X \sim U(0,3)$ dengan fungsi kubik $y = \cos(4x) + \varepsilon_1$ 21

4.17 Plot data $X \sim U(0,3)$ dengan fungsi kubik $y = \cos(4x) + \varepsilon_2$ 22

4.18 Plot data $X \sim U(0,3)$ dengan fungsi kubik $y = \cos(4x) + \varepsilon_3$ 22

4.19 Plot data $X \sim U(0,3)$ dengan fungsi kubik $y = \cos(4x) + \varepsilon_4$ 22

4.20 Plot data $X \sim U(0,3)$ dengan fungsi kubik $y = \cos(4x) + \varepsilon_5$ 23

4.21 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *25*

4.22 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *27*

4.23 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *29*

4.24 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *31*

4.25 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *33*

4.26 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *35*

4.27 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *37*

4.28 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *39*

4.29 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *41*

4.30 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *43*

4.31 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *45*

4.32 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *47*

4.33 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *49*

4.34 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *51*

4.35 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *53*

4.36 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *55*

4.37 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *57*

4.38 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *59*

4.39 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *61*

4.40 (a) 1 *knot*, (b) 2 *knot*, (c) 3 *knot* *63*

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Analisis regresi merupakan salah satu teknik analisis data dalam statistika yang paling banyak digunakan untuk mengetahui hubungan antara variabel prediktor (X) dengan variabel responnya (Y). Meskipun analisis regresi sangat bervariasi ragamnya, namun keseluruhan analisis regresi dalam statistik bertujuan untuk prediksi atau peramalan data. Dalam analisis regresi terdapat dua pendekatan yaitu regresi parametrik dan regresi nonparametrik, pada model parametrik diasumsikan bahwa pola fungsi diketahui seperti linear, kuadratik, kubik, polinomial derajat-p, eksponensial, dan lain sebagainya. Asumsi pada pendekatan parametrik tersebut juga tersedia sumber-sumber lain yang memberikan suatu informasi yang rinci. Sebaliknya dengan regresi nonparametrik, pada pendekatan model regresinya tidak diketahui analisis regresi bentuk atau pola.

Regresi nonparametrik memiliki fleksibilitas yang tinggi dan tidak memerlukan adanya asumsi-asumsi mengenai sebaran data. Dalam regresi nonparametrik ada beberapa teknik estimasi seperti pendekatan histogram, estimator kernel, estimator deret *orthogonal*, analis *wavelet*, estimator MARS, estimator *Fourier*, estimator *spline* dan lain-lain.

spline merupakan salah satu model yang mempunyai interpretasi statistik dan interpretasi visual yang baik dan memiliki kemampuan yang sangat kecil untuk menangani data yang perilakunya berubah pada sub interval tertentu. Masalah utama pada saat menduga fungsi regresi menggunakan *smoothing spline* adalah memilih dan menentukan parameter pemulus. Metode *Spline* juga merupakan model polinomial yang tersegmen atau terbagi dimana sifat segmen inilah yang memberikan eksibelitas yang lebih baik dibanding model polinomial biasa. Sifat ini memungkinkan model regresi *spline* menyesuaikan diri secara efektif terhadap karakteristik lokal dari data. Dalam penelitian- penelitian sebelumya sudah banyak yang menggunakan regresi *spline*, dengan berbagai jenis *spline*. Penulis tertarik untuk menggunakan *spline* kuadratik untuk mencari model regresi dari data simulasi.

1.2 Tujuan Penelitian

Tujuan dari penelitian ini untuk mengestimasi model regresi nonparametrik dengan menggunakkan metode *spline* kuadratik pada data simulasi.

1.3 Manfaat Penelitian

Manfaat dilakukannya penelitian ini yaitu:

1. Memberikan sumbangan pemikiran bagi penelitian lain mengenai model regresi nonparametrik.
2. Menambah pengetahuan tentang penggunakkan metode *spline* kuadratik pada data simulasi

II. TINJAUAN PUSTAKA

2.1 Analisis Regresi

Analisis regresi merupakan analisis hubungan antara variabel variabel respon (*Y*) dengan prediktor (*X*). Pendekatan regresi dibedakan menjadi dua yaitu pendekatan secara parametrik dan pendekatan nonparametrik. Pendekatan parametrik merupakan pemodelan regresi yang terikat dengan asumsi-asumsi dalam regresi. Asumsi-asumsi tersebut antara lain multikolinearitas, residual normalitas, homokedastisitas residual, dan nonautokorelasi. Sedangkan pendekatan regresi nonparametrik tidak ada asumsi-asumsi yang harus dipenuhi dalam pemodelan. Regresi parametrik dilakukan apabila bentuk kurva regresinya diketahui. Sedangkan regresi nonparametrik dilakukan jika bentuk kurva regresinya tidak diketahui. Regresi semiparametrik digunakan jika sebagian bentuk kurva regresinya tidak diketahui sedangkan yang sebagian lainnya diketahui (Wati, 2014).

2.2 Kegunaan Analisis Regresi

Menurut Iriawan (2006), analisis Regresi sangat berguna dalam penelitian antara lain model regresi dapat digunakan untuk mengukur kekuatan hubungan antara variabel respons dan variabel prediktor, model regresi dapat digunakan untuk mengetahui pengaruh suatu atau beberapa variabel prediktor terhadap variabel respon, model regresi berguna untuk memprediksi pengaruh suatu variabel atau beberapa variabel predictor terhadap variabel respons. Kedua variabel tersebut dihubungkan dalam bentuk persamaan matematika, secara umum, bentuk persamaan regresi dinyatakan sebagai berikut:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \qquad (2.1)$$

dengan :

β_0 = intersep (konstanta)

$\beta_1 \dots \beta_k$ = parameter model regresi untuk $x_1 \dots x_k$

ε = galat

2.3 Regresi Parametrik

Analisis regresi parametrik Analisis regresi merupakan suatu studi yang digunakan untuk melihat ketergantungan atau hubungan antara suatu variabel respons (variabel terikat) pada satu atau lebih variabel prediktor (variabel bebas). Analisis regresi terdiri dari dua jenis variabel yaitu variabel tak bebas atau variabel respons yang disebut juga sebagai variabel dependen. Hubungan antara variabel respons (Y_i) dengan variabel prediktor (X_i) dapat dinyatakan dalam persamaan berikut (Kutner, *et al*., 2004):

$$Y_i = f(x_i) + \varepsilon_i \;\; ; i = 1,2,3, \dots , n \qquad (2.2)$$

dalam hal ini Y_i adalah nilai variabel respons dalam amatan ke-i, X adalah peubah bebas atau variabel prediktor dalam amatan ke-i, $f(x_i)$ adalah regresi yang telah diketahui bentuknya dengan asumsi data berdistribusi normal, varians homogen, jenis data yang digunakan rasio atau interval. ε_i adalah suku galat atau sisaan yang diasumsikan independen dan bersifat acak dengan nilai tengah nol dan variansi σ^2dalam amatan ke-i, dan *n* adalah banyaknya amatan Model antara dua atau lebih variabel independen atau variabel prediktor ($x_1, x_2, x_3, \dots , x_n$) dengan variabel dependen (Y_i).

2.4 Regresi Nonparametrik

Regresi nonparametrik merupakan salah satu metode statistik yang digunakan untuk mengestimasi bentuk hubungan antara variabel respon dengan variabel prediktor tidak diketahui bentuk kurva regresinya. Dalam regresi nonparametrik kurva regresi hanya diasumsikan mulus (*smooth*) dalam arti termuat dalam suatu ruang fungsi tertentu sehingga mempunyai fleksibilitas yang tinggi (Eubank, 1999). Menurut Wahba (1990). Secara umum model regresi nonparametrik dapat dituliskan sebagai berikut:

$$y_i = f(x_i) + \varepsilon_i \qquad ; i = 1,2, \dots n \qquad (2.3)$$

dengan:

y_i : variabel respon dari data ke-(i)

$f(x_i)$: fungsi regresi yang tidak diketahui bentuknya (fungsi mulus)

t_i : variabel prediktor dari data ke-*i*

ε_i : galat ke-i yang diasumsikan menyebar $N\ (0,\sigma^2)$

2.5 *Spline*

Spline merupakan potongan polinomial (*pieceweise polynomial*) tersegmen yang memiliki sifat fleksibilitas. Titik perpaduan bersama dari potongan –potongan tersebut atau titik yang menunjukan terjadinya perubahan – perubahan prilaku kurva pada interval – interval yang berbeda disebut knot (Fan & Yao, 2005). Secara umum fungsi *spline* berorde *m* adalah sebuah fungsi yang didefinisikan dengan (Wu & Zang ,2006):

$$f(x) = \sum_{j=0}^{m-1} \beta_j x_i^j + \sum_{k=1}^{r} \beta_{m-1+k} (x_i - K_k)_+^{m-1} \quad , \quad i = 1,2,\dots,n \tag{2.4}$$

dengan:

$$(x - K_k)_+^{m-1} = \begin{cases} (x_i - K_k)^{m-1} & , x \geq K \\ 0 & , x \geq K \end{cases}$$

$\beta_j$dan β_{m-1+k} adalah parameter, $\lambda = \{K_1, K_2, \dots K_n\}$ adalah titik knot $f(x)$ adalah fungsi regresi, dan x adalah variabel prediktor, serta $(x_i - K_k)_+^{m-1}$ merupakan fungsi potongan polinomial tersegmen kontinu.

2.6 Regresi *spline*

Dalam regresi nonparametrik, bentuk fungsi hubungan antara variabel bebas dan variabel tak bebas tidak diketahui. Fungsi hubungan tersebut adalah sembarang fungsi yang biasa dari kelas tertentu. Regresi *spline* adalah regresi dimana kurva regresinya berupa fungsi *spline*. Secara umum model regresi nonparametrik *spline* dapat ditulis sebagai berikut:

$$y_i = \sum_{j=0}^{m-1} \beta_j x_i^j + \sum_{k=1}^{r} \beta_{(m-1)+k}(x_i - k_k)^{m-1} + \varepsilon_i\,, \quad i = 1,2,\dots,n \qquad (2.5)$$

Regresi *spline* memungkinkan untuk berbagai macam orde sehingga dapat dibentuk regresi *spline* linear, kuadrat, kubik maupun orde ke-*m*. Secara umum model spline kuadratik sebagai berikut :

$$Y_i = \beta_0 + \beta_1 x_i + \beta_2 {x_i}^2 + \sum_{k=1}^{r} \beta_{(2)+k}\,(x_i - K_k)^2 \qquad (2.6)$$

Dengan data amatan sebanyak n, bentuk persamaan (2.5) dituliskan dengan mtriks:

$$\tilde{y} = X_1\tilde{\delta}_1 + X_2\tilde{\delta}_2 + \tilde{\varepsilon}$$

dimana:

$$\tilde{y} = \begin{bmatrix} y_1 \\ y_1 \\ \vdots \\ y_n \end{bmatrix}; X_1 = \begin{bmatrix} 1 & x_1 & \dots & x_1^{m-1} \\ 1 & x_2 & \dots & x_2^{m-1} \\ 1 & \vdots & \ddots & \vdots \\ 1 & x_n & \dots & x_n^{m-1} \end{bmatrix};\ \tilde{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$X_2 = \begin{bmatrix} (x_1 - K_1)_+^{m-1} & (x_1 - K_2)_+^{m-1} & \dots & (x_1 - K_r)_+^{m-1} \\ (x_2 - K_1)_+^{m-1} & (x_2 - K_2)_+^{m-1} & \dots & (x_2 - K_r)_+^{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ (x_n - K_1)_+^{m-1} & (x_n - K_2)_+^{m-1} & \dots & (x_n - K_r)_+^{m-1} \end{bmatrix};\ \tilde{\delta}_1 = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix};$$

$$\tilde{\delta}_1 = \begin{bmatrix} \beta_{(m-1)+1} \\ \beta_{(m-1)+2} \\ \vdots \\ \beta_{(m-1)+r} \end{bmatrix}$$

Bentuk matriks dari persamaan (2.5) dapat disederhanakan menjadi

$$\tilde{y} = X\tilde{\beta} + \tilde{\varepsilon}$$

Dengan X = [X_1 ⋮ X_2] dan $\tilde{\beta} = \begin{bmatrix} \tilde{\delta}_1 \\ \cdots \\ \tilde{\delta}_2 \end{bmatrix}$

2.7 Estimasi parameter regresi nonparametrik *Spline*

Hidayat (2017) melakukan suatu kajian dengan kurva dengan kurva regresi *f* dihampiri dengan fungsi *spline f* dengan titik knot K untuk mendapatkan estimator *spline*. Dalam bentuk matriks disajikan sebagai berkut:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Apabila model regresi *spline* disajikan dalam matriks, diperoleh :

$$\begin{bmatrix} y_1 \\ y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & \cdots & x_1^{m-1} & (x_1 - K_1)_+^{m-1} & (x_1 - K_2)_+^{m-1} & \cdots & (x_1 - K_r)_+^{m-1} \\ 1 & x_2 & \cdots & x_2^{m-1} & (x_2 - K_1)_+^{m-1} & (x_2 - K_2)_+^{m-1} & \cdots & (x_2 - K_r)_+^{m-1} \\ 1 & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^{m-1} & (x_n - K_1)_+^{m-1} & (x_3 - K_2)_+^{m-1} & \cdots & (x_n - K_r)_+^{m-1} \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{m-1} \\ \beta_{(m-1)+1} \\ \beta_{(m-1)+2} \\ \beta_{(m-1)+r} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Atau dapat ditulis dengan:

$$\tilde{y} = X[K_1, K_2, \dots, K_r]\tilde{\beta} + \tilde{\varepsilon} \tag{2.7}$$

Selanjutnya, estimasi parameter $\tilde{\beta} = [\ \beta_0\beta_1\ \dots \beta_m\beta_{m+1}\beta_{m+2}\ \dots\ \beta_{m+r}]^T$ diperoleh melalui least square, yaitu dengan meminimumkan jumlah kuadrat error sebagai berikut:

$$\min_{\tilde{\beta}\epsilon R^{m+1+r}} \{\tilde{\varepsilon}^T\tilde{\varepsilon}\} = \min_{\tilde{\beta}\epsilon R^{\tilde{\beta}\epsilon R^{m+1+r}}} \left(\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}^T \begin{bmatrix} \varepsilon_1 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} \right)$$

$$\min_{\tilde{\beta}\epsilon R^{m+1+r}} \{\tilde{\varepsilon}^T\tilde{\varepsilon}\} = \min_{\tilde{\beta}\epsilon R^{\tilde{\beta}\epsilon R^{m+1+r}}} \{[[\tilde{y} - X[K_1, K_2, \dots, K_r]\tilde{\beta}]^T \left[[\tilde{y} - X[K_1, K_2, \dots, K_r]\tilde{\beta}]\right]\}$$

Dengan penyajian matriks diberikan sebagai berikut :

$$\sum_{i=1}^{n} \varepsilon_i^2 = \tilde{\varepsilon}^T\tilde{\varepsilon}$$

$$= ([\tilde{y} - X[K_1, K_2, \dots, K_r]\tilde{\beta}]^T \left[[\tilde{y} - X[K_1, K_2, \dots, K_r]\tilde{\beta}]\right])$$

$$= ([\tilde{y} - X[\lambda]\tilde{\beta}]^T[[\tilde{y} - X[\lambda]\tilde{\beta}])$$

$$= \tilde{y}^T\tilde{y} - 2\tilde{\beta}^T X^T[\lambda\ ]\tilde{y} + \tilde{\beta}^T X^T[\lambda]X[\lambda]\tilde{\beta} \tag{2.8}$$

Kemudian, persamaan diturunkan terhadap vector $\tilde{\beta}^T$ dan disamakan dengan nol, maka diperoleh:

$$\frac{\partial(\tilde{\varepsilon}^T\widetilde{\varepsilon})}{\partial\tilde{\beta}^T} = \frac{\partial(\tilde{y}^T\tilde{y} - 2\tilde{\beta}^T X^T[\lambda]\tilde{y} + \tilde{\beta}^T X^T[\lambda]X[\lambda]\tilde{\beta}}{\partial\tilde{\beta}^T}$$

$$\hat{\tilde{\beta}} = (X^T[\lambda]X[\lambda])^{-1}X^T[\lambda]\tilde{y} \tag{2.9}$$

dengan:

$$X[\lambda] = X[K_1, K_2, \dots, K_r]$$

$$= \begin{bmatrix} 1 & x_1 & \dots & x_1^{m-1} & (x_1 - K_1)_+^{m-1} & (x_1 - K_2)_+^{m-1} & \dots & (x_1 - K_r)_+^{m-1} \\ 1 & x_2 & \dots & x_1^{m-1} & (x_2 - K_1)_+^{m-1} & (x_2 - K_2)_+^{m-1} & \dots & (x_2 - K_r)_+^{m-1} \\ 1 & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \dots & x_1^{m-1} & (x_n - K_1)_+^{m-1} & (x_3 - K_2)_+^{m-1} & \dots & (x_n - K_r)_+^{m-1} \end{bmatrix}$$

2.8 Estimasi kurva regresi nonparametrik spline

setelah didapatkan estimasi parameter seperti pada persamaan, maka estimasi tititk kurva regresi nonparametrik spline dapat diperoleh dengan mensubstitusikan Persamaan a persamaan b. akibatnya, estimasi untuk kurva regresi spline dengan knot λ diberikan oleh:

$$\hat{y}[\lambda](x_i) = \hat{\tilde{f}}[\lambda](x_i) = X[\lambda]\hat{\tilde{\beta}}$$

$$\hat{\tilde{f}}[\lambda](x_i) = X[\lambda](X^T[\lambda]X[\lambda)^{-1}X^T[\lambda]\tilde{y}$$

Sehingga diperoleh:

$$\hat{\hat{f}}[\lambda](x_i) = A[\lambda]\tilde{y}$$

Dengan, $A[\lambda] = X[\lambda](X^T[\lambda]X[\lambda])^{-1}X^T[\lambda]$ dan $X[\lambda]$ adalah matriks dari model yang bergantung pada titik knot dan λ= $[K_1, K_1, \dots, K_1]^T$

2.9 Pemilihan Titik Knot Optimal

Pemilihan λ optimal dalam regresi nonparametrik dengan metode *spline* pada hakekatnya merupakan pemilihan lokasi titik knot. Knot diartikan sebagai titik fokus dalam fungsi *spline*, sehingga kurva yang dibentuk tersegmen pada titik tersebut. Titik knot merupakan perpaduan bersama yang menunjukkan pola perilaku fungsi *spline* pada selang yang berbeda (Hardle, 1990). Untuk nilai λ yang sangat besar akan menghasilkan bentuk kurva regresi yang sangat halus. Sebaliknya untuk nilai λ yang kecil akan memberikan bentuk kurva regresi yang sangat kasar (Wahba, 1990). Pemilihan *knot* pada dapat dilakukan secara *trial error* atau uji coba dan *Generalized Cross Validation* (GCV). Pada model spline kuadratik kriteria GCV didefinisikan sebagai berikut:

$$\text{GCV}(\lambda) = \frac{n^{-1}\sum_{i=0}^{n}(y_i - f_\lambda(x_i))^2}{[n^1\, tr(I - A(\lambda))]^2} \tag{2.10}$$

dengan:

K : titik knot $(K_1, K_2, \dots, K_n)$

n : banyaknya data

I : matriks identitas

Kriteria GCV(λ) diharapkan memiliki nilai yang minimum, sehingga model regresi nonparametrik spline dapat dikatakan memiliki λ yang optimal.

2.10 Koefisien Determinasi

Koefisien determinasi (R^2) merupakan ukuran ketelitian atau ketepatan model regresi, atau besarnya kontribusi *x* terhadap perubahan *y*. Semakin tinggi nilai R^2 maka model akan semakin baik. Rumus koefisien determinasi diberikan oleh:

$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} = \frac{JKR}{JKT} \qquad (2.11)$$

Menurut Gujarati (2003), besaran nilai R^2 tidak pernah negative dan batasannya adalah $0 \leq R^2 \leq 1$.

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester Ganjil Tahun Ajaran 2018/2019 di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung.

3.2 Data Penelitian

Data yang digunakan dalam penelitian ini yaitu data bangkitan menggunakan *software R*, dengan $x \sim U$ $(0,3)$ dan $n = 100$. y didapatkan dari fungsi:

1. Fungsi linear : $y = (2x + 1) + \varepsilon_i$
2. Fungsi kuadratik : $y = 2x^2 - 6x + 16 + \varepsilon_i$
3. Fungsi kubik : $Sin: \ y = Sin(4x) + \varepsilon_i$

   $Cos: \ y = cos(4x) + \varepsilon_i$

dengan ε_i:

$\varepsilon_1 \sim N(0{,}0.3)$, $\varepsilon_2 \sim N(0{,}1), \varepsilon_3 \sim N(0{,}1.3)$, $\varepsilon_4 \sim 0.95N(0{,}1) + 0.05N(5{,}1)$, dan $\varepsilon_5 \sim 0.9N(0{,}1) + 0.1N(5{,}1)$.

3.3 Metode Penelitian

Dalam penelitian ini dilakukan estimasi model regresi nonparametrik menggunakan metode *spline* kuadratik dengan menggunakan *software* R, adapun langkah-langkah analisis yang dilakukan dalam penelitian ini yaitu:

1. Membuat diagram pencar dari setiap fungsi yang di dapat.
2. Memilih titik *knot* optimal dengan menggunakan *Generalized cross validation (GCV)* untuk setiap fungsi yang diteliti.
3. Menghitung nilai koefisien determinasi *R* untuk setiap hasil dengan fungsi *spline* yang didapat.
4. Mengestimasi model dan menggambarkan model terbaik.
5. Menginterpretasikan model yang diperoleh dan menarik kesimpulan.

V. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan dapat disimpulkan bahwa:

1. Metode *spline* kuadratik dapat mengestimasi fungsi kuadratik sangat baik pada galat normal.
2. Metode *spline* kuadratik kurang baik digunakan untuk mengestimasi fungsi *linear* dan kubik pada galat normal dan galat yang mengandung pencilan.

DAFTAR PUSTAKA

Fan, J. & Yao, Q. 2005. *Nonlinear Time Series Nonparametric And Parametric Method*. Springer Science, Canada.

Gujarati, D N. 2003. *Basic Econometrics*. The McGraw-Hill Companies, New York.

Hardle,W. 1990. *Applied Nonparametric Regression*. Cambrigde University Press, New York.

Hidayat, R., *et al*. 2017. Model Regresi Nonparametrik dengan Pendekatan Spline Truncated. *Prosiding Seminar Nasional Institut Teknologo Sepuluh November*, Surabaya: Hal. 203-352.

Iriawan, N. 2006. *Mengolah Data Statistik Dengan Mudah Menggunakan Minitab 14*. Andi, Yogyakarta.

Kutner, *et al*. 2004. *Applied Linear Regression Models*. 4th Editions. The McGraw-Hill Company, New York.

Wahba, G. 1990. Spline model for observational data. *Society For Industrial and Applied Mathematics*, Philadelphia.

Wati, L. 2014. Penentuan Parameter Penghalus Smoothing Spline dalam Regresi Semiparametrik dengan GCV. Skripsi. Jurusan Matematika FMIPA UIN Maulana Malik Ibrahim, Malang.

Wu, H. & Zang, J.T. 2006. *Nonparametric Regression Methods for Longitudinal Data Analisys*. John Wiley & Sons, New Jersey