



TESIS-SS 142501

**PENGUJIAN HIPOTESIS PARSIAL
PADA REGRESI SEMIPARAMETRIK
SPLINE *TRUNCATED***
(Aplikasi: Data Angka Harapan Hidup di Indonesia)

DEWI FITRIANA
NRP.1315201713

DOSEN PEMBIMBING
Prof. Dr. Drs. I Nyoman Budiantara, M.Si
Dr. Vita Ratnasari, S.Si, M.Si

PROGRAM MAGISTER STATISTIK
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2017



TESIS-SS 142501

**PENGUJIAN HIPOTESIS PARSIAL
PADA REGRESI SEMIPARAMETRIK
SPLINE *TRUNCATED*
(Aplikasi: Data Angka Harapan Hidup di Indonesia)**

**DEWI FITRIANA
NRP.1315201713**

**DOSEN PEMBIMBING
Prof. Dr. Drs. I Nyoman Budiantara, M.Si
Dr. Vita Ratnasari, S.Si, M.Si**

**PROGRAM MAGISTER STATISTIK
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2017**



TESIS-SS 142501

**PARTIAL HYPOTHESIS TESTING IN TRUNCATED
SPLINE SEMIPARAMETRIC REGRESSION
(Application: Life Expectancy in Indonesia)**

**DEWI FITRIANA
NRP.1315201713**

**SUPERVISOR
Prof. Dr. Drs. I Nyoman Budiantara, M.Si
Dr. Vita Ratnasari, S.Si, M.Si**

**PROGRAM OF MAGISTER
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCE
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2017**

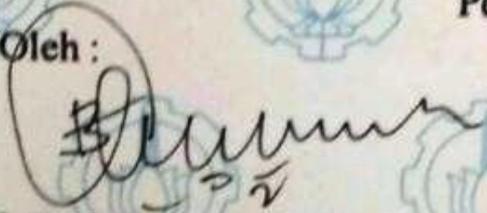
**PENGUJIAN HIPOTESIS PARSIAL PADA
REGRESI SEMIPARAMETRIK SPLINE TRUNCATED
(Aplikasi: Data Angka Harapan Hidup di Indonesia)**

Disusun untuk memenuhi syarat memperoleh gelar Magister Sains (M.Si)
di
Institut Teknologi Sepuluh Nopember

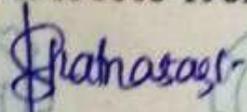
Oleh :
DEWI FITRIANA
NRP. 1315201713

Tanggal Ujian : 20 Juni 2017
Periode Wisuda : September 2017

Disetujui Oleh :


1. Prof. Dr. I Nyoman Budiantara, M.Si
NIP. 19650603 198903 1 003

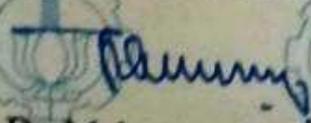
(Pembimbing I)


2. Dr. Vita Ratnasari, M.Si
NIP. 19700910 1997 2 001

(Pembimbing II)


3. Dr. Dra. Ismaini Zain, M.Si
NIP. 19600525 198803 2 001

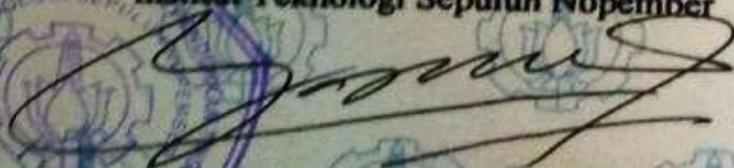
(Penguji I)


4. Dr. Muhammad Mashuri, MT
NIP. 19620408 198701 1 001

(Penguji II)

Dekan

Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember


Prof. Dr. Basuki Widodo, M.Sc
NIP. 19650605 198903 1 002

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini, mahasiswa Departemen Statistika FMIPA ITS:

Nama : Dewi Fitriana

NRP : 1315201713

Program Studi : Magister Statistika / Doktor Ilmu Statistik

menyatakan bahwa data yang digunakan dalam Tesis / Disertasi ini merupakan data sekunder yang diambil dari ~~penelitian / buku / Tugas Akhir / Thesis / Disertasi /~~ publikasi lainnya yaitu:

Sumber : Badan Pusat Statistik

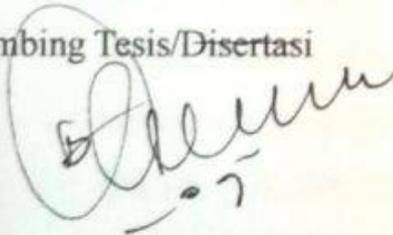
Keterangan : Publikasi Statistik Indonesia 2015
Publikasi Statistik Kesejahteraan Rakyat 2015
Publikasi Indeks Pembangunan Manusia 2015

Surat pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Surabaya, 10 - 07 - 2017

Mengetahui

Pembimbing Tesis/Disertasi



(Prof. Dr. I Nyoman Budiantara, M. Si)
NIP. 19650603 198903 1003

Mahasiswa



(Dewi Fitriana)
NRP. 1315201713

*(coret yang tidak perlu)

** Mohon dituliskan data primer apabila datanya dikumpulkan sendiri

**PENGUJIAN HIPOTESIS PARSIAL PADA REGRESI
SEMIPARAMETRIK SPLINE *TRUNCATED*
(Aplikasi : Data Angka Harapan Hidup di Indonesia)**

Mahasiswa Nama : Dewi Fitriana
Mahasiswa ID : 1315201713
Dosen Pembimbing I : Prof. Dr. Drs. I Nyoman Budiantara, M.Si
Dosen Pembimbing II : Dr. Vita Ratnasari, S.Si, M.Si

ABSTRAK

Salah satu metode dalam regresi semiparametrik untuk mengestimasi kurva regresi yang tidak diketahui bentuknya dapat dilakukan dengan pendekatan spline *truncated*. Pengujian hipotesis dalam model spline pada regresi semiparametrik bertujuan untuk mengidentifikasi apakah pola data dapat dimodelkan dengan pendekatan polinomial atau spline. Kajian teoritis tentang pengujian hipotesis dilakukan untuk mendapatkan statistik uji, distribusi statistik uji, dan daerah penolakan hipotesis nol. Kajian teoritis tentang pengujian hipotesis parsial pada regresi semiparametrik spline *truncated* belum pernah dilakukan. Oleh karena itu pada penelitian ini dilakukan pengembangan pengujian hipotesis secara parsial beserta terapannya. Metode yang digunakan adalah *Likelihood Ratio Test* (LRT). Dari kajian teori dihasilkan bahwa statistik uji pada pengujian hipotesis parsial adalah

$$W^* \sim F_{(1, n - (p+1 + (r+m)q))}$$

dimana n adalah banyaknya observasi, p adalah banyaknya variabel parametrik, r adalah banyaknya knot, m adalah ordo, dan q adalah banyaknya variabel nonparametrik.

Kajian terapan diaplikasikan pada data Angka Harapan Hidup di Indonesia pada Tahun 2015. AHH di Indonesia masih rendah jika dibandingkan dengan AHH negara berkembang lainnya. Metode pemilihan titik knot optimum menggunakan metode *Generalized Cross Validation* (GCV). Model terbaik yang terbentuk adalah model menggunakan tiga titik knot (3,3,3,3) dengan koefisien determinasi (R^2) sebesar 84,70 persen. Variabel yang signifikan adalah persentase penduduk miskin, persentase rumah tangga menggunakan air minum bersih, persentase penduduk yang mengalami keluhan kesehatan, persentase balita mendapat imunisasi lengkap, dan MYS. Dari hasil penelitian ini diharapkan dapat dijadikan masukan bagi Pemerintah untuk mengambil kebijakan dalam rangka meningkatkan AHH nasional secara menyeluruh.

Kata Kunci : Angka Harapan Hidup, Uji hipotesis parsial, Semiparametrik, spline *truncated*, *Likelihood Ratio Test* (LRT).

"Halaman ini sengaja dikosongkan"

**PARTIAL HYPOTHESIS TESTING
OF SPLINE TRUNCATED SEMIPARAMETRIC REGRESSION
(Application: Life Expectancy in Indonesia)**

By : Dewi Fitriana
Student Identity Number : 1315201713
Supervisor : Prof. Dr.Drs. I Nyoman Budiantara, M.Si
Co Supervisor : Dr. Vita Ratnasari, S. Si, M. Si

ABSTRACT

One of method in semiparametric regression to estimate unrecognized curve of regression can be conducted with spline truncated approach. Hypothesis testing for spline models in semiparametric regression is used to identify that data pattern can be modeled with polynomial or spline approach. Theoretical study of hypothesis testing used to obtain test statistics, distributions of test statistical, and rejection area of the hypothesis null. Actually theoretical study of partial hypothesis testing in truncated spline semiparametric regression has never been done. Therefore, in this research will develop the theoretical of partial hypothesis testing and application. In this study conducted a review again. The method used is Likelihood Ratio Test (LRT). From the theoretical study resulted that test statistical of partial hypothesis testing is

$$W^* \sim F_{((1), (n-(p+1+(r+m)q))}$$

where n is sum of observation, p is sum of parametric variable, r is sum of knot, m is ordo, and q is sum of nonparametric variable.

Applied study conducted on life expectancy data of Indonesia in 2015. AHH of Indonesia is still low if compared to AHH of other developing countries. The method of selecting optimum knots using Generalized Cross Validation (GCV) method. Best model which is obtained by using three knots (3,3,3,3) with coefficient of determination (R^2) is 84,70 percent. Significant variables were percentage of poor population, percentage of households using clean drinking water, percentage of population with health complaints, percentage of children under five were fully immunized and MYS. This result is expected to be used as input for Government to take policy in order to improve the national AHH overall in Indonesia .

Key words : Life Expectancy, Likelihood Ratio Test (LRT), partial hypothesis testing, semiparametric, spline *truncated*

"Halaman ini sengaja dikosongkan"

KATA PENGANTAR

Alhamdulillah Robbil Alamiin, segala puji bagi Allah SWT yang telah memberikan rahmat, hidayah, kekuatan, kesabaran, dan ilmu pengetahuan sehingga penulis diperkenankan menyelesaikan tesis yang berjudul **“Pengujian Hipotesis Parsial pada Regresi Semiparametrik Spline *Truncated* (Aplikasi : Angka Harapan Hidup di Indonesia)”**. Tesis ini merupakan salah satu syarat penyelesaian studi pada Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Program Pascasarjana Institut Teknologi Sepuluh Nopember (ITS) Surabaya.

Keberhasilan penulisan dan penyusunan tesis ini tidak terlepas dari bantuan, bimbingan, dan dukungan dari berbagai pihak. Untuk itu, teriring rasa syukur dan doa, penulis ingin menyampaikan penghargaan dan ucapan terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. Drs. I Nyoman Budiantara, M.Si. dan Ibu Dr. Vita Ratnasari, S.Si, M.Si selaku dosen pembimbing dengan segala kesibukannya dapat meluangkan waktu untuk memberikan bimbingan, saran, masukan serta motivasi selama penyusunan tesis ini.
2. Ibu Dr. Drs. Ismaini Zain, M.Si. dan Bapak Dr. Muhammad Mashuri, M.T selaku dosen penguji yang telah memberikan saran dan masukan untuk menjadikan tesis ini menjadi lebih baik.
3. Bapak Dr. Suhartono, M.Sc. selaku Ketua Jurusan Statistika sekaligus dosen wali, Bapak Dr. rer. pol. Heri Kuswanto, S.Si., M.Si. selaku Ketua Program Studi Pascasarjana Jurusan Statistika dan seluruh Bapak/Ibu dosen pengajar yang telah memberikan ilmu, saran, masukan dan pengalaman yang bermanfaat bagi penulis, serta segenap karyawan keluarga besar Jurusan Statistika FMIPA ITS Surabaya atas segala dukungan dan bantuannya selama ini.
4. Kepala BPS RI beserta jajarannya, Kepala Pusdiklat BPS beserta jajarannya, Kepala BPS Provinsi Provinsi Jawa Timur beserta jajarannya, Kepala BPS Kabupaten Mojokerto beserta jajarannya, yang telah memberikan kesempatan

dan kepercayaan kepada Penulis sehingga dapat melanjutkan studi program S2 di ITS.

5. Bapak dan Ibu tersayang yang tidak pernah letih berdoa, mendidik, mendampingi, mengasihi, dan terus memberi dorongan semangat dan kekuatan bagi Penulis.
6. Teristimewa, suami dan anak-anakku tercinta, Kanza Aulya Sabian, Kinantan Fairuz Chalissa, Bryan Rizqi Abdillah, dan Keanu Syarief Abdillah terima kasih atas segala doa, pengorbanan, pengertian, dukungan, senyum dan cinta yang tak pernah berhenti yang selalu menjadi semangat bagi penulis untuk menyelesaikan studi dengan baik.
7. Saudara-saudaraku Ika Fauziah, Z. Anisnawati, Abdul Azis, dan Zulia N. Azizah yang telah memberikan dukungan moril baik lahir maupun batin.
8. Saudara, sahabat, rekan-rekan seperjuangan ITS BPS Batch 9 yang telah banyak membantu dan bekerja sama dengan penulis dalam proses penyelesaian studi. Penulis sungguh bersyukur dapat menjalani masa studi dengan penuh sukacita bersama. Teman-teman batch 10, teman-teman reguler S2 2016, Fany, Sem, Mas Risky terima kasih atas kerjasamanya.
9. Ibu Ruliana di tengah kesibukannya mempersiapkan sidang disertasi masih sempat memberi masukan, bantuan, motivasi, dan ilmu dalam menyelesaikan tesis ini. Juga Mbak Mia terima kasih atas bantuannya selama ini.
10. Semua pihak yang tidak bisa disebutkan satu persatu. Penulis menyampaikan terima kasih atas semua dukungan dan bantuan yang diberikan selama menjalani studi.

Penulis menyadari bahwa tesis ini masih jauh dari sempurna. Oleh karena itu, kritik maupun saran yang bersifat membangun sangat penulis harapkan demi perbaikan tesis ini. Akhirnya, penulis berharap mudah-mudahan tesis ini bermanfaat untuk semua pihak yang memerlukan.

Surabaya, Agustus 2017

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN.....	i
KATA PENGANTAR	vii
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xv
DAFTAR NOTASI.....	xvii
DAFTAR LAMPIRAN.....	xix
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Permasalahan.....	7
1.3 Tujuan Penelitian	7
1.4 Manfaat Penelitian	7
1.5 Batasan Permasalahan	8
BAB 2 TINJAUAN PUSTAKA.....	9
2.1 Analisis Regresi	9
2.2 Regresi Parametrik.....	10
2.3 Regresi Nonparametrik	11
2.4 Regresi Semiparametrik	12
2.5 Spline <i>Truncated</i> dalam Regresi Semiparametrik	12
2.6 Pemilihan Titik Knot Optimal	16
2.7 <i>Likelihood Ratio Test</i> (LRT).....	16
2.8 Pengujian Parameter dalam Regresi Parametrik.....	17
2.8.1 Uji Simultan/Serentak	17
2.8.2 Uji Parsial/Individu	18
2.9 Pemeriksaan Asumsi Residual.....	19
2.9.1 Uji Asumsi Normalitas.....	19

2.9.2	Uji Asumsi Independen	20
2.9.3	Uji Asumsi Identik	20
2.10	Teorema Dasar Terkait dengan Aljabar Matriks	21
2.11	Tinjauan Angka Harapan Hidup	22
2.11.1	Angka Harapan Hidup (AHH)	22
2.11.2	Faktor-Faktor Yang Mempengaruhi Angka Harapan Hidup.....	25
BAB 3 METODOLOGI PENELITIAN.....		29
3.1	Sumber Data	29
3.2	Variabel Penelitian	29
3.3	Definisi Operasional Variabel Penelitian	29
3.4	Tahapan Penelitian	32
BAB 4 HASIL DAN PEMBAHASAN		35
4.1	Model Regresi Semiparametrik Spline <i>Truncated</i>	35
4.1.1	Estimasi Titik untuk Kurva Regresi Semiparametrik	35
4.1.2	Perumusan Uji Hipotesis	39
4.1.3	Menentukan Statistik Uji Hipotesis	42
4.1.4	Mendapatkan Distribusi Statistik Uji Hipotesis	52
4.1.5	Menentukan Daerah Penolakan Uji Hipotesis H_0	59
4.2	Aplikasi pada Data AHH Tahun 2015 di Indonesia	61
4.2.1	Analisis Deskriptif	61
4.2.2	Memodelkan AHH menggunakan Regresi Linier Berganda.....	64
4.2.2.1	Pengujian Signifikansi Parameter secara Simultan.....	65
4.2.2.2	Pengujian Signifikansi Parameter secara Parsial.....	65
4.2.2.3	Pengujian Asumsi Residual.....	66
4.2.2.3.1	Pengujian Asumsi Identik	66
4.2.2.3.2	Pemeriksaan Asumsi Independen	67

4.2.2.3.3	Pengujian Asumsi Distribusi Normal ...	67
4.2.3	Memodelkan AHH menggunakan Regresi Semiparametrik Spline <i>Truncated</i>	68
4.2.3.1	Menentukan Variabel Komponen Parametrik dan Nonparametrik.....	68
4.2.3.2	Memodelkan AHH Indonesia Menggunakan Satu Titik Knot.....	70
4.2.3.3	Memodelkan AHH Indonesia Menggunakan Dua Titik Knot.....	71
4.2.3.4	Memodelkan AHH Indonesia Menggunakan Tiga Titik Knot.....	72
4.2.3.5	Memodelkan AHH Indonesia Menggunakan Kombinasi Titik Knot.....	74
4.2.3.6	Memilih Titik Knot Optimum dengan Menggunakan GCV.....	75
4.2.3.7	Model Regresi Semiparametrik Spline <i>Truncated</i> dengan Titik Knot Optimum.....	75
4.2.3.8	Pengujian Signifikansi Parameter secara Simultan....	76
4.2.3.9	Pengujian Signifikansi Parameter secara Parsial....	77
4.2.3.10	Pengujian Asumsi Residual.....	78
4.2.3.10.1	Pengujian Asumsi Identik	78
4.2.3.10.2	Pengujian Asumsi Independen	79
4.2.3.10.3	Pengujian Asumsi Distribusi Normal ...	79
4.2.4	Perbandingan Model dan Interpretasi	80
BAB 5	KESIMPULAN DAN SARAN	87
5.1	Kesimpulan.....	87
5.2	Saran.....	89

DAFTAR PUSTAKA	91
LAMPIRAN	95
BIOGRAFI PENULIS	115

DAFTAR TABEL

Tabel 2.1 Analisis Varians Model Regresi.....	18
Tabel 4.1 Statistik Deskriptif Variabel Respon dan Variabel Prediktor	63
Tabel 4.2 ANOVA Hasil Regresi Linier Berganda	65
Tabel 4.3 Hasil Uji Hipotesis secara Parsial	66
Tabel 4.4 Hasil Uji <i>Glejser</i>	66
Tabel 4.5 Ringkasan Penentuan Komponen Parametrik dan Nonparametrik	70
Tabel 4.6 Nilai GCV dengan Satu Titik Knot	71
Tabel 4.7 Nilai GCV dengan Dua Titik Knot	72
Tabel 4.8 Nilai GCV dengan Tiga Titik Knot	73
Tabel 4.9 Nilai GCV dengan Kombinasi Titik Knot	74
Tabel 4.10 Nilai GCV Minimum pada Tiap Model	75
Tabel 4.11 ANOVA Hasil Regresi Semiparametrik Spline <i>Truncated</i>	76
Tabel 4.12 Hasil Uji Hipotesis secara Parsial	77
Tabel 4.13 Hasil Uji <i>Glejser</i>	78
Tabel 4.14 Nilai R^2 dan R^2 <i>adjusted</i>	80

”Halaman ini sengaja dikosongkan”

DAFTAR GAMBAR

Gambar 4. 1 Peta AHH menurut Provinsi di Indonesia.....	62
Gambar 4. 2 <i>Probability</i> Plot Residual	68
Gambar 4. 3 Scatter Plot antara AHH dengan Variabel-variabel Prediktor.....	69
Gambar 4. 4 <i>Probability</i> Plot Residual.....	80
Gambar 4. 5 Peta AHH terhadap Persentase Rumah Tangga yang Menggunakan Air Minum Bersih.....	82
Gambar 4. 6 Peta AHH terhadap Persentase Penduduk yang Mengalami Keluhan Kesehatan.....	84
Gambar 4. 7 Peta AHH terhadap Persentase Balita yang Mendapat Imunisasi Lengkap	85
Gambar 4. 8 Peta AHH terhadap MYS.....	86

"Halaman ini sengaja dikosongkan"

DAFTAR NOTASI

Simbol	Keterangan
$f(x)$	Kurva regresi
ε	Error random
$g(x)$	Fungsi regresi parametrik atau kurva regresi parametrik
$f(t)$	Kurva regresi nonparametrik
β	Parameter variabel parametrik
γ	Parameter variabel nonparametrik
K	Titik knot
y	Variabel respon
x	Variabel prediktor yang mengikuti pola parametrik
t	Variabel prediktor yang mengikuti pola nonparametrik
m	Banyaknya orde
n	Banyaknya observasi/data
p	Banyaknya variabel prediktor parametrik
q	Banyaknya variabel prediktor nonparametrik
r	Banyaknya knot
\mathbf{X}	Matriks yang memuat prediktor komponen parametrik
\mathbf{T}	Matriks yang memuat prediktor komponen nonparametrik
\mathbf{Z}	Matriks yang memuat prediktor komponen parametrik dan nonparametrik
$\tilde{\delta}$	Vektor parameter parametrik dan nonparametrik
λ	Likelihood ratio test
Ω	Ruang parameter di bawah populasi
ω	Ruang parameter di bawah Hipotesis nol

"Halaman ini sengaja dikosongkan"

DAFTAR LAMPIRAN

Lampiran 1 Data dan Struktur Data yang Digunakan.....	95
Lampiran 2 Program Regresi Semiparametrik Spline <i>Truncated</i> Satu Titik	
Knot	97
Lampiran 3 Program Regresi Semiparametrik Spline <i>Truncated</i> Dua Titik	
Knot.....	99
Lampiran 4 Program Regresi Semiparametrik Spline <i>Truncated</i> Tiga Titik	
Knot	102
Lampiran 5 Program Regresi Semiparametrik Spline <i>Truncated</i> Kombinasi	
Titik Knot.....	105
Lampiran 6 Uji Signifikansi Parameter	110
Lampiran 7 Uji <i>Glejser</i>	113

“Halaman ini sengaja dikosongkan”

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Analisis regresi merupakan metode dalam statistika yang dipergunakan untuk mengetahui pola hubungan antara variabel prediktor dengan variabel respon. Pendekatan dalam analisis regresi ada tiga yaitu pendekatan regresi parametrik, regresi nonparametrik, dan regresi semiparametrik. Apabila bentuk kurva regresi diketahui maka pendekatan model regresi tersebut dinamakan model regresi parametrik (Hardle, 1990; Islamiyati dan Budiantara, 2007). Jika pola data cenderung mengikuti model linear/kuadrat/kubik maka pendekatan regresi yang sesuai untuk data tersebut adalah regresi parametrik linear/kuadrat/kubik (Hardle, 1990). Apabila pola hubungan antara variabel prediktor dan variabel respon tidak diketahui bentuknya atau tidak terdapat informasi masa lalu yang lengkap tentang bentuk pola data maka regresi nonparametrik merupakan pendekatan regresi yang dianjurkan untuk digunakan (Budiantara, 2009). Model regresi nonparametrik yang sering mendapat perhatian dari para peneliti adalah *Kernel* (Hardle, 1990), *Spline* (Wahba, 1990), Deret *Fourier* (Antoniadis, 1994) dan *Wavelets* (Antoniadis, 2007).

Pendekatan regresi nonparametrik memiliki fleksibilitas yang tinggi, karena data diharapkan bisa mencari sendiri bentuk estimasi kurva regresinya tanpa dipengaruhi oleh faktor subyektifitas peneliti. Salah satu pendekatan yang dapat digunakan dalam mengestimasi model regresi nonparametrik adalah spline. Menurut Budiantara (2009), data yang memiliki pola yang berubah-ubah pada sub-sub interval tertentu sangat baik dimodelkan dengan menggunakan metode spline. Spline pada hakekatnya merupakan generalisasi dari fungsi polinomial biasa dengan cara-cara optimasi yang masih mengadopsi konsep dalam regresi parametrik. Spline merupakan salah satu model yang mempunyai interpretasi statistik dan interpretasi visual sangat khusus dan sangat baik (Eubank, 1988). Sementara itu, Wahba (1990) memberikan metode untuk memilih parameter penghalus optimal dalam estimator spline yaitu dengan *Generalized Cross*

Validation (GCV). Lee (2002) meneliti tentang regresi spline dengan *ordinary least square*.

Salah satu pendekatan yang dapat digunakan untuk estimasi parameter dalam model linear adalah pendekatan regresi spline *truncated*. Spline *truncated* dapat mengatasi pola data yang naik/turun tajam dengan bantuan titik-titik knot sehingga menghasilkan kurva yang relatif mulus atau *smooth* (Hardle, 1990). Budiantara (2005) mengembangkan estimator spline dalam regresi nonparametrik dengan menggunakan basis fungsi keluarga spline *truncated* tersebut. Pendekatan dengan menggunakan basis fungsi spline *truncated* ini memberikan perhitungan matematik yang lebih mudah dan sederhana serta optimasi yang digunakan tanpa melibatkan *penalty* yaitu optimasi *least square* (*LS*). Pada dasarnya, persoalan utama dalam estimasi parameter menggunakan regresi spline *truncated* adalah pemilihan titik-titik knot optimal.

Regresi spline banyak dikembangkan secara teoritis dan diterapkan dalam berbagai bidang baik dalam bidang ekonomi, industri, demografi, maupun kesehatan. Hal ini dapat dilihat dari berbagai penelitian tentang regresi spline diantaranya, Budiantara (2006) menggunakan spline terbobot untuk memantau pertumbuhan balita di Kota Surabaya. Purnomo (2016) melakukan pemodelan spline *truncated* campuran dengan kernel dalam regresi nonparametrik pada rata-rata lama sekolah di Provinsi Jawa Tengah. Bintariningrum (2014) juga melakukan pemodelan regresi nonparametrik spline *truncated* pada angka kelahiran kasar di Surabaya. Selain itu, Rory (2016) melakukan penelitian dengan menggunakan regresi campuran nonparametrik spline *truncated* dan kernel dalam data kemiskinan di Papua.

Dalam beberapa kasus, variabel respon dapat memiliki hubungan linear dengan salah satu variabel prediktor, tetapi dengan variabel prediktor yang lain tidak diketahui bentuk pola hubungannya. Dalam keadaan seperti ini, Wahba (1990) menyarankan penggunaan pendekatan regresi semiparametrik yang merupakan gabungan antara komponen parametrik dan komponen nonparametrik. Pemilihan model dalam regresi semiparametrik merupakan hal yang sangat penting dalam perkembangan selanjutnya. Untuk memilih model semiparametrik, fokus utama adalah pada pemilihan variabel dan parameter pemulus dari calon

model (*models candidate*) yang menggabungkan regresi parametrik dan nonparametrik. Beberapa penelitian tentang regresi semiparametrik adalah Engle (1986) memperkenalkan regresi semiparametrik untuk mengestimasi hubungan antara cuaca dan penjualan listrik dengan pendekatan spline linier. Heckman (1986) menyelesaikan model regresi semiparametrik dengan menggunakan spline. Mulianah (2006) meneliti tentang pendekatan kernel dalam regresi semiparametrik dan pemilihan bandwidth optimal dan Srinadi (2002) meneliti tentang estimator spline pada model semiparametrik, estimasinya dikaji dalam dua bagian yaitu estimator untuk model parametrik dan nonparametrik. Sriliana (2012) melakukan estimasi untuk regresi semiparametrik spline *truncated* dalam model linier parsial untuk data longitudinal. Solikhin (2014) mengkaji estimator spline *truncated* linier dalam regresi semiparametrik pada data persentase pengeluaran konsumsi padi-padian di Provinsi Jawa Tengah.

Inferensi statistik khususnya pengujian hipotesis terhadap parameter dalam model sangat penting dilakukan untuk mengetahui apakah variabel-variabel prediktor berpengaruh secara signifikan terhadap model. Wahba (1990) membuat suatu inferensi statistik khususnya interval konfidensi untuk model spline dalam regresi nonparameterik dengan pendekatan bayesian untuk smoothing spline, Zaki (2007) telah melakukan kajian tentang inferensi uji *Generalized Maximum Likelihood* (GML) untuk menguji hipotesis dalam model regresi nonparametrik spline dan menyelidiki perilaku uji GML yang dibandingkan dengan uji *Locally Most Powerful* (LMP) dengan menggunakan data simulasi. Tupen (2011) juga telah melakukan kajian tentang uji hipotesis dalam regresi nonparametrik spline *truncated*. Ruliana (2016) melakukan kajian tentang pengujian hipotesis simultan pada model spline pada *Structural Equation Modeling* (SEM) nonlinier. Sementara itu, pengembangan tentang pengujian hipotesis parsial khususnya pada regresi semiparametrik spline *truncated* belum pernah dilakukan. Kajian ini sangat penting karena dapat digunakan untuk mengetahui bentuk dari hipotesis, penurunan untuk mendapatkan statistik uji dan daerah penolakan dari pengujian hipotesis secara parsial khususnya pada regresi semiparametrik spline *truncated*.

Aplikasi pengujian hipotesis telah banyak dilakukan pada berbagai bidang keilmuan, salah satunya adalah pada bidang demografi. Penelitian bidang

demografi, diantaranya penelitian indikator derajat kesehatan yaitu Angka Harapan Hidup (AHH) merupakan bidang yang sesuai untuk penerapan regresi semiparametrik spline *truncated*. Hal ini disebabkan karena hubungan antara variabel respon dengan sebagian variabel prediktor cenderung tidak diketahui pola hubungannya, namun dengan sebagian variabel prediktor lainnya membentuk hubungan linier.

Jumlah penduduk merupakan salah satu kekayaan yang dimiliki oleh suatu bangsa. Indonesia adalah salah satu negara yang menempati posisi lima besar dengan jumlah penduduk terbanyak di dunia. Pada tahun 2015, Indonesia berada pada urutan ke empat setelah Cina, India, dan Amerika Serikat. Sebagai bangsa yang besar, Indonesia seharusnya memiliki kemampuan melakukan pembangunan dengan menjadikan penduduk sebagai subjek dan objek pembangunan. Keberhasilan pembangunan bangsa tidak terlepas dari pencapaian pembangunan manusia sebagai kekayaan yang sesungguhnya dimiliki oleh suatu bangsa. Pada prinsipnya, ada tiga pilihan yang paling mendasar pada semua level pembangunan yaitu masyarakat ingin berumur panjang dan hidup sehat, ingin memperoleh pendidikan, dan ingin memiliki akses terhadap sumber-sumber kebutuhan agar hidup secara layak. Konsep kehidupan tersebut menjadi cikal bakal terbentuknya Indeks Pembangunan Manusia (IPM) yang diluncurkan pertama kali oleh *United Nations Development Programme* (UNDP) pada tahun 1990. Oleh karena itu keberhasilan pembangunan suatu bangsa juga dapat dilihat dari pencapaian Indeks Pembangunan Manusia (IPM) tersebut.

Salah satu komponen IPM yang masih menjadi permasalahan dan perhatian dunia saat ini adalah Angka Harapan Hidup (AHH). UNDP menggunakan AHH untuk mengukur derajat kesehatan masyarakat sekaligus sebagai tolok ukur keberhasilan pembangunan. AHH yang rendah di suatu daerah harus diikuti program pembangunan kesehatan dan program sosial lainnya seperti kesehatan lingkungan, kecukupan gizi dan kalori termasuk program pemberantasan kemiskinan.

AHH di Indonesia terus mengalami peningkatan hampir sepanjang tahun. Artinya, harapan seorang bayi yang baru lahir untuk dapat hidup lebih lama menjadi semakin tinggi. Pada tahun 2010 AHH Indonesia adalah 69,81 tahun dan

terus mengalami kenaikan sampai pada tahun 2015 menjadi 70,78 tahun. Menurut Bappenas, pemerintah menargetkan AHH Indonesia akan mencapai 73,7 tahun pada tahun 2025. Namun jika dicermati, pertumbuhan AHH di Indonesia masih sangat rendah. Pada periode 1990 – 2015, secara rata-rata pertumbuhan AHH hanya sekitar 0,3 persen per tahun. Hal ini sejalan dengan pengamatan WHO pada AHH negara-negara berkembang yang bergerak lambat bahkan lebih lambat jika dibanding dengan negara tertinggal.

Seiring dengan peningkatan AHH secara nasional, provinsi-provinsi di Indonesia juga mengalami pertumbuhan AHH yang positif. Namun demikian masih terlihat ada disparitas AHH antar provinsi. Kenyataan ini tentu tidak sejalan dengan target pembangunan pemerintah yang tercantum dalam program NAWACITA yaitu pemerataan pembangunan antarwilayah di Indonesia. Pada tahun 2015, AHH provinsi di Indonesia berkisar antara 64 – 74 tahun. Dari 34 provinsi masih ada 73,53 persen provinsi yang memiliki AHH di bawah AHH nasional. Sebagian besar provinsi tersebut berada di wilayah Indonesia bagian timur, diantaranya adalah Provinsi Papua Barat 65,19 tahun diikuti oleh Provinsi Papua 65,09 tahun dan yang terendah adalah Provinsi Sulawesi Barat 64,22 tahun. Sementara itu 26,47 persen provinsi sudah memiliki AHH di atas AHH nasional. Sebagian besar provinsi tersebut berada di wilayah Indonesia bagian barat. Provinsi yang memiliki AHH tertinggi adalah Provinsi DI Yogyakarta yaitu 74,68 tahun, diikuti Provinsi Jawa Tengah 73,96 tahun dan Provinsi Kalimantan Timur 73,65 tahun. Jika dibandingkan dengan kondisi lima tahun yang lalu, pada tahun 2010 AHH terendah adalah Provinsi Sulawesi Barat yaitu 62,50 tahun, sedangkan AHH Provinsi DI Yogyakarta sudah mencapai 74,17 tahun pada tahun yang sama. Artinya bayi yang lahir pada tahun 2015 di Provinsi DI Yogyakarta secara rata-rata memiliki usia harapan hidup sekitar 12 tahun lebih lama dibanding bayi yang lahir di Provinsi Sulawesi Barat pada tahun yang sama.

Peningkatan usia harapan hidup secara nasional di Indonesia masih menempatkan AHH Indonesia di bawah AHH negara-negara tetangga. Bersumber dari *United Nations* dalam publikasinya, BPS mencatat AHH beberapa negara di dunia selama periode 1990 - 2015. Pada periode tahun tersebut, AHH Indonesia masih berada di bawah Singapura (82,2), Malaysia (74,9), Thailand (74,3), dan

Kamboja (71,6). Pencapaian AHH Indonesia sedikit lebih baik dibandingkan Filipina (68,6) dan India (66,3). Menurut data *Central Intelligence Agency (CIA) World*, AHH di Indonesia masih masuk ke dalam kategori rendah jika dibanding dengan negara-negara di seluruh dunia. Berdasarkan kajian yang dihasilkan, AHH Indonesia menduduki peringkat 137 dari 223 negara di dunia. Hal ini menjadi salah satu cerminan akan masih rendahnya kualitas hidup penduduk di Indonesia. Kajian tersebut selaras dengan masalah dan tantangan yang dinyatakan dalam Rencana Pembangunan Jangka Panjang Nasional (RPJPN) 2005-2025, yaitu meningkatkan kualitas hidup di berbagai bidang pembangunan.

Studi *Global Burden Disease (GBD)* di bawah pengawasan *Institute for Health Metrics and Evaluation (IHME)* mengatakan bahwa negara-negara berkembang saat ini sedang menghadapi tantangan serius yaitu pola hidup dan penyakit mematikan yaitu jantung, stroke dan diabetes yang dapat menyerang semua kalangan umur. Menurut catatan *World Health Organization (WHO)*, salah satu negara yang sudah berhasil mengatasi penyakit mematikan tersebut adalah Korea Selatan, sehingga pada tahun 2015 Korea Selatan mampu mencapai AHH di atas 80 tahun dan pada tahun yang sama berhasil membuat sebuah komunitas penduduk yang berusia hampir mencapai 100 tahun.

Sebagai indikator kualitas kesehatan dan kesejahteraan, AHH tentu saja dipengaruhi oleh banyak faktor. Beberapa penelitian yang mengkaji faktor-faktor yang mempengaruhi AHH diantaranya adalah Andriani (2015) mengkaji faktor-faktor yang mempengaruhi angka harapan hidup di Provinsi Jawa Timur berdasarkan angka melek huruf, rata-rata lama sekolah, dan pengeluaran perkapita. Firdial (2010) menyebutkan faktor-faktor yang mempengaruhi AHH adalah Angka Kematian Bayi (AKB), persentase persalinan dengan bantuan non medis, rata-rata usia perkawinan pertama perempuan, rata-rata lama sekolah, rata-rata pengeluaran rumah tangga, persentase daerah berstatus desa, rata-rata lama pemberian Air Susu Ibu (ASI) eksklusif, dan persentase rumah tangga yang menggunakan air bersih.

1.2 Rumusan Permasalahan

Berdasarkan latar belakang yang telah dijelaskan di atas, pengujian hipotesis merupakan suatu inferensi yang sangat penting dalam suatu pemodelan. Selanjutnya pendekatan regresi semiparametrik spline *truncated* akan digunakan untuk memodelkan AHH karena diduga bentuk kurva regresi antara variabel respon dengan sebagian variabel prediktor tidak diketahui bentuk polanya dan sebagian lagi cenderung mempunyai pola hubungan linier.

Rumusan permasalahan dalam penelitian ini adalah bagaimana pengembangan statistik uji untuk pengujian hipotesis secara parsial pada regresi semiparametrik spline *truncated* dan bagaimana aplikasinya terhadap AHH di Indonesia

1.3 Tujuan Penelitian

Berdasarkan permasalahan di atas, tujuan dari penelitian ini adalah sebagai berikut.

1. Mengembangkan pengujian hipotesis secara parsial pada model regresi semiparametrik spline *truncated*.
2. Mengaplikasikan model regresi semiparametrik spline *truncated* pada data AHH di Indonesia.

1.4 Manfaat Penelitian

Manfaat yang ingin dicapai dari hasil penelitian ini adalah sebagai berikut.

1. Mengembangkan wawasan keilmuan terhadap regresi semiparametrik spline *truncated*.
2. Menghasilkan penurunan uji hipotesis regresi semiparametrik spline *truncated* sehingga dapat dijadikan tambahan referensi dalam melakukan analisis data.
3. Dapat memberikan informasi yang lebih rinci mengenai fenomena angka harapan hidup sehingga dapat membantu pemerintah dalam menentukan kebijakan di bidang kesehatan dan kesejahteraan masyarakat.

4. Memberikan alternatif bagi Badan Pusat Statistik (BPS) mengenai metode analisis data khususnya untuk data yang sesuai dengan regresi semiparametrik spline *truncated*.

1.5 Batasan Permasalahan

Bertolak dari perumusan masalah di atas, maka ruang lingkup dalam penelitian ini dibatasi pada beberapa hal, yaitu untuk mengembangkan penurunan statistik uji dalam penulisan ini menggunakan metode *Likelihood Ratio Test* (LRT). Metode yang digunakan untuk menentukan banyaknya titik knot dan lokasi titik knot optimum adalah dengan menggunakan metode *Generalized Cross Validation* (GCV). Pemilihan banyaknya titik knot dibatasi dengan menggunakan satu, dua, tiga dan kombinasi titik knot. Hasil penurunan uji hipotesis akan diaplikasikan pada data AHH di Indonesia dengan unit observasi adalah provinsi di Indonesia Tahun 2015. Dalam aplikasi model regresi semiparametrik spine *truncated*, fungsi spline yang digunakan adalah spline linier karena sesuai dengan pola data yang digunakan.

BAB 2

TINJAUAN PUSTAKA

2.1 Analisis Regresi

Analisis regresi merupakan suatu metode statistik yang menggambarkan pola hubungan antara variabel prediktor dan variabel respon. Salah satu tujuan lain dari analisis regresi adalah untuk memperkirakan atau meramalkan nilai dari variabel respon apabila nilai dari variabel prediktor diketahui (Draper dan Smith, 1992). Jika variabel respon adalah y_i , variabel prediktor adalah x_i , $i = 1, 2, \dots, n$, maka pasangan data (x_i, y_i) memiliki model hubungan fungsional.

$$y_i = f(x_i) + \varepsilon_i, i = 1, 2, \dots, n \quad (2.1)$$

dimana $f(x_i)$ adalah kurva regresi dan ε_i adalah error random yang diasumsikan identik, independen dan berdistribusi normal dengan mean nol dan varian σ^2 (Eubank, 1999).

Berdasarkan pola datanya, ada beberapa pendekatan analisis regresi yaitu pendekatan regresi parametrik, regresi semiparametrik, dan regresi nonparametrik. Jika pola hubungan variabel respon dan variabel prediktor diketahui, maka pendekatan analisis regresi tersebut dinamakan analisis regresi parametrik (Budiantara, 2009). Namun jika pola hubungan antara variabel prediktor dan respon yang tidak diketahui bentuk kurva regresinya atau tidak terdapat informasi masa lalu yang lengkap tentang bentuk pola data, salah satu pendekatan yang dapat digunakan adalah regresi nonparametrik (Eubank, 1988). Selain itu, apabila diasumsikan bahwa bentuk kurva regresi sebagian diketahui dan sebagian tidak diketahui, maka digunakan pendekatan regresi semiparametrik (Eubank, 1999; Budiantara, 2005). Beberapa konsep statistika yang akan digunakan dalam penelitian ini adalah regresi parametrik, regresi semiparametrik dan nonparametrik, spline dalam regresi nonparametrik, pemilihan titik knot optimal, dan pengujian parameter model.

2.2 Regresi Parametrik

Regresi parametrik merupakan suatu metode dalam analisis regresi. Dalam regresi parametrik menuntut terpenuhinya berbagai asumsi klasik yang sangat ketat, salah satunya adalah bentuk kurva regresi diketahui, misalnya linear, kuadratik, kubik, polinomial derajat-p, eksponen, dan lain-lain. Dengan mengetahui bentuk kurva regresi maka peneliti dapat memilih salah satu bentuk keluarga kurva atau fungsi regresi yang memungkinkan dari beberapa alternatif yang ada, kemudian menempatkan fungsi regresi tersebut dalam proses inferensi. Jika bentuk kurva atau fungsi regresi yang dipilih tepat, maka analisis regresi parametrik akan lebih menguntungkan, khususnya terkait dengan inferensinya dan interpretasi modelnya yang lebih sederhana. Oleh karena itu analisis regresi parametrik sering digunakan apabila terdapat informasi tentang bentuk kurva regresinya (Eubank, 1999). Pendekatan model regresi parametrik memiliki sifat yang sangat baik dari pandangan statistika inferensia, diantaranya adalah sederhana, mudah interpretasinya, parsimoni, tak bias, estimator linier, efisien, konsisten, dan *Best Linier Unbiased Estimator* (BLUE) yang sangat jarang dimiliki oleh pendekatan model regresi lainnya (Budiantara, 2009).

Secara matematis fungsi regresi parametrik bisa ditulis dengan persamaan sebagai berikut:

$$y_i = g(x_i) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.2)$$

fungsi $g(x_i)$ disebut fungsi regresi parametrik atau kurva regresi parametrik. *Error* random ε_i diasumsikan berdistribusi normal independen dengan rata-rata nol dan variansi σ^2 (Eubank, 1999). Fungsi $g(x_i)$ dapat juga dituliskan dalam bentuk matrik, sehingga persamaannya menjadi:

$$y_i = \mathbf{x}'_i \boldsymbol{\beta} + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (2.3)$$

dimana $\mathbf{x}'_i = [1 \quad x_{i1} \quad x_{i2} \quad \dots \quad x_{ip}]$, $i = 1, 2, \dots, n$; n adalah banyak data, p adalah

banyak variabel parametrik, dan $\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \\ \vdots \\ \beta_p \end{bmatrix}$.

Persamaan (2.3) dapat pula disajikan dalam bentuk :

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{i1} & x_{i2} & \dots & x_{ip} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i, i = 1, 2, \dots, n \quad (2.4)$$

2.3 Regresi Nonparametrik

Regresi nonparametrik adalah salah satu metode yang digunakan untuk mengestimasi pola hubungan antara variabel respon dan variabel prediktor yang tidak diketahui bentuk fungsinya, hanya diasumsikan *smooth* (mulus) dalam arti termuat dalam suatu ruang fungsi tertentu (Eubank, 1988). Pendekatan regresi nonparametrik memiliki fleksibilitas yang tinggi, karena data diharapkan mencari sendiri bentuk estimasi kurva regresinya tanpa dipengaruhi oleh faktor subyektifitas peneliti (Eubank, 1988).

Secara umum model regresi nonparametrik dapat ditulis sebagai berikut:

$$y_i = f(t_i) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.5)$$

dengan y_i adalah variabel respon ke- i , sedangkan fungsi $f(t_i)$ merupakan kurva regresi dan t_i sebagai variabel prediktor. ε_i adalah *error* random yang diasumsikan berdistribusi normal independen dengan mean nol dan variansi σ^2 (Wahba, 1990). Regresi nonparametrik ini disebut dengan regresi nonparametrik univariabel karena terdiri dari satu variabel respon dan satu variabel prediktor. Jika dalam regresi nonparametrik terdapat satu variabel respon dan lebih dari satu variabel prediktor, maka disebut regresi nonparametrik multivariabel (Budiantara, 2004). Jika diberikan data $(t_{i1}, t_{i2}, \dots, t_{iq}, y_i)$ dan hubungan antara $(t_{i1}, t_{i2}, \dots, t_{iq})$ dan y_i dapat dituliskan sebagai berikut:

$$y_i = f(t_{i1}, t_{i2}, \dots, t_{iq}) + \varepsilon_i \quad (2.6)$$

$$= f(t_{i1}) + f(t_{i2}) + \dots + f(t_{iq}) + \varepsilon_i \quad (2.7)$$

$$= \sum_{l=1}^q f(t_{il}) + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (2.8)$$

dengan y_i adalah variabel respon dan f adalah kurva regresi yang tidak diketahui bentuknya.

2.4 Regresi Semiparametrik

Regresi semiparametrik adalah gabungan antara regresi parametrik dan regresi nonparametrik. Menurut Eubank (1999) dan Budiantara (2005), apabila diasumsikan bentuk kurva regresi sebagian diketahui dan sebagian tidak diketahui, maka sebaiknya digunakan pendekatan regresi semiparametrik. Model regresi semiparametrik dapat ditulis sebagai berikut:

$$y_i = \mathbf{x}'_i \boldsymbol{\beta} + f(t_i) + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (2.9)$$

dimana y_i adalah variabel respon ke- i , $\mathbf{x}'_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{ip})'$ adalah variabel prediktor untuk komponen parametrik, t_i adalah variabel prediktor untuk komponen nonparametrik, $\mathbf{x}'_i \boldsymbol{\beta}$ merupakan komponen parametrik,

$\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_p)' \in R^p$ parameter yang tidak diketahui, $f(t_i)$ merupakan fungsi komponen nonparametrik yang tidak diketahui polanya dan ε_i adalah *error* random, dimana $\varepsilon_i \sim N(0, \sigma^2)$.

2.5 Spline *Truncated* dalam Regresi Semiparametrik

Spline *truncated* merupakan pendekatan regresi nonparametrik yang banyak digunakan. Spline *truncated* merupakan potongan-potongan polinomial yang memiliki sifat tersegmen dan kontinu. Salah satu kelebihan spline *truncated* adalah model ini cenderung mencari sendiri estimasi data kemanapun pola data tersebut bergerak. Kelebihan ini terjadi karena dalam spline *truncated* terdapat titik-titik knot, yaitu titik perpaduan bersama yang menunjukkan terjadinya perubahan pola perilaku data (Eubank, 1999; Budiantara, 2009).

Jika diberikan data berpasangan (y_i, x_i, t_i) , $i = 1, 2, \dots, n$, dimana y_i adalah variabel respon sedangkan x_i adalah variabel prediktor yang mengikuti pola parametrik dan t_i adalah variabel prediktor yang mengikuti pola nonparametrik. Pola hubungan y_i , x_i dan t_i dapat dinyatakan dalam model regresi seperti persamaan (2.9), yaitu:

$$y_i = \mathbf{x}'_i \boldsymbol{\beta}_i + f(t_i) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.10)$$

Selanjutnya jika kurva regresi $f(t_i)$ pada persamaan (2.10) dihipotesiskan dengan fungsi spline *truncated* dengan knot K_1, K_2, \dots, K_r maka:

$$f(t_i) = \gamma_1 t_i + \gamma_2 t_i^2 + \dots + \gamma_m t_i^m + \gamma_{1+m} (t_i - K_1)_+^m + \dots + \quad (2.11)$$

$$\begin{aligned} & \gamma_{r+m} (t_i - K_r)_+^m \\ & = \sum_{j=1}^m \gamma_j t_i^j + \sum_{k=1}^r \gamma_{k+m} (t_i - K_k)_+^m \end{aligned} \quad (2.12)$$

dengan

$$(t_i - K_k)_+^m = \begin{cases} 0, & t_i < K_k \\ (t_i - K_k)^m, & t_i \geq K_k \end{cases}$$

γ adalah parameter dan $(t_i - K_k)_+^m$ adalah fungsi *truncated*. Derajat m merupakan derajat pada persamaan polinomial. Biasanya digunakan $m = 1, 2$ atau 3 . Kurva regresi polinomial derajat 1 biasa disebut dengan kurva regresi linier, kurva regresi polinomial derajat 2 biasa disebut dengan kurva regresi kuadratik, sedangkan kurva regresi polinomial derajat 3 biasa disebut dengan kurva regresi kubik. Titik-titik knot K_1, K_2, \dots, K_r adalah titik-titik knot yang menunjukkan pola perilaku dari kurva pada sub-sub interval yang berbeda, dimana $K_1 < K_2 < \dots < K_r$.

Sebagai salah satu ilustrasi sederhana, misalkan diberikan spline linier *truncated* dengan $m = 1$, $r = 3$ atau tiga titik knots pada $t_{1i} = K_1$, $t_{2i} = K_2$, $t_{3i} = K_3$, maka dapat disajikan dalam bentuk:

$$f(t_i) = \gamma_1 t_i + \gamma_2 (t_i - K_1)_+^1 + \gamma_3 (t_i - K_2)_+^1 + \gamma_4 (t_i - K_3)_+^1 \quad (2.13)$$

Fungsi Spline $f(t_i)$ dapat pula disajikan dalam bentuk (Budiantara, 2005):

$$f(t_i) = \begin{cases} \gamma_1 t_i & , t_i < K_1 \\ \gamma_1 t_i + \gamma_2 (t_i - K_1) & , K_1 \leq t_i < K_2 \\ \gamma_1 t_i + \gamma_2 (t_i - K_1) + \gamma_3 (t_i - K_2) & , K_2 \leq t_i < K_3 \\ \gamma_1 t_i + \gamma_2 (t_i - K_1) + \gamma_3 (t_i - K_2) + \gamma_4 (t_i - K_3), & t_i \geq K_3 \end{cases} \quad (2.14)$$

Persamaan regresi semiparametrik spline *truncated* pada persamaan (2.10) dapat dinyatakan sebagai:

$$y_i = \mathbf{x}'_i \boldsymbol{\beta}_i + \sum_{j=1}^m \gamma_j t_i^j + \sum_{k=1}^r \gamma_{k+m} (t_i - K_k)_+^m + \varepsilon_i \quad (2.15)$$

Regresi semiparametrik spline *truncated* diatas terdiri dari satu variabel respon dengan satu atau lebih variabel prediktor parametrik dan hanya satu variabel prediktor nonparametrik. Jika dalam regresi semiparametrik terdiri dari

satu variabel respon dengan lebih dari satu variabel prediktor parametrik dan lebih dari satu variabel prediktor nonparametrik, dengan komposisi data seperti berikut

$$(x_{i1}, x_{i2}, \dots, x_{ip}, t_{i1}, t_{i2}, \dots, t_{iq}, y_i)$$

maka hubungan antara $(x_{i1}, x_{i2}, \dots, x_{ip}, t_{i1}, t_{i2}, \dots, t_{iq})$ dan y_i dapat dituliskan sebagai berikut:

$$y_i = \mathbf{x}'_i \boldsymbol{\beta} + f(t_{i1}, t_{i2}, \dots, t_{iq}) + \varepsilon_i \quad (2.16)$$

$$= \mathbf{x}'_i \boldsymbol{\beta} + f(t_{i1}) + f(t_{i2}) + \dots + f(t_{iq}) + \varepsilon_i \quad (2.17)$$

$$= \mathbf{x}'_i \boldsymbol{\beta} + \sum_{l=1}^q f(t_{il}) + \varepsilon_i, i = 1, 2, \dots, n \quad (2.18)$$

dengan y_i adalah variabel respon, $\mathbf{x}'_i \boldsymbol{\beta}$ adalah fungsi regresi parametrik dan $f(t_{il})$ adalah kurva regresi yang tidak diketahui bentuknya, dimana

$$\sum_{l=1}^q f(t_{il}) = \sum_{l=1}^q \left(\sum_{j=1}^m \gamma_{jl} t_{il}^j + \sum_{k=1}^r \gamma_{(k+m)l} (t_{il} - K_{kl})_+^m \right), \quad (2.19)$$

$$i = 1, 2, \dots, n$$

Persamaan (2.18) dapat ditulis dalam bentuk sebagai berikut:

$$\tilde{y} = \mathbf{Z}(\tilde{k}) \tilde{\delta} + \tilde{\varepsilon} \quad (2.20)$$

Respon y merupakan vektor berukuran $n \times 1$, matriks $\mathbf{Z}(\tilde{k}) = (\mathbf{X} : \mathbf{T})$, \mathbf{X} adalah matriks yang memuat prediktor komponen parametrik yang berukuran $n \times (p + 1)$ dan \mathbf{T} adalah matriks yang memuat komponen nonparametrik berukuran $n \times q(r + m)$ yang tergantung pada titik knot yang diberikan, dimana \tilde{k} adalah vektor titik knot berukuran $r \times 1$. Vektor parameter $\tilde{\delta} = (\beta_0, \dots, \beta_p, \gamma_{11}, \dots, \gamma_{(r+m)1}, \gamma_{12}, \dots, \gamma_{(r+m)2}, \dots, \gamma_{1q}, \dots, \gamma_{(r+m)q})'$ berukuran $((p + 1) + q(r + m)) \times 1$ dan $\tilde{\varepsilon}$ merupakan vektor *error*. Estimasi parameter model regresi spline *truncated* diperoleh dengan menggunakan metode *Ordinary Least Square* (OLS).

Berdasarkan model persamaan (2.20), diperoleh persamaan:

$$\tilde{\varepsilon} = \tilde{y} - \mathbf{Z}(\tilde{k}) \tilde{\delta} \quad (2.21)$$

Selanjutnya jumlah kuadrat *error* diberikan oleh:

$$\sum_{i=1}^n \varepsilon_i^2 = \tilde{\varepsilon}'\tilde{\varepsilon} \quad (2.22)$$

$$\sum_{i=1}^n \varepsilon_i^2 = (\tilde{y} - \mathbf{Z}(\tilde{k})\tilde{\delta})'(\tilde{y} - \mathbf{Z}(\tilde{k})\tilde{\delta}) \quad (2.23)$$

Dengan menggunakan sifat tranpose suatu matrik yaitu:

$$(\mathbf{Z}(\tilde{k})\tilde{\delta})' = \tilde{\delta}'\mathbf{Z}'(\tilde{k}) \quad (2.24)$$

Maka,

$$\sum_{i=1}^n \varepsilon_i^2 = (\tilde{y} - \mathbf{Z}(\tilde{k})\tilde{\delta})'(\tilde{y} - \mathbf{Z}(\tilde{k})\tilde{\delta}) \quad (2.25)$$

$$= \tilde{y}'\tilde{y} - \tilde{y}'\mathbf{Z}(\tilde{k})\tilde{\delta} - \tilde{\delta}'\mathbf{Z}'(\tilde{k})\tilde{y} + \tilde{\delta}'\mathbf{Z}'(\tilde{k})\mathbf{Z}(\tilde{k})\tilde{\delta} \quad (2.26)$$

$$= \tilde{y}'\tilde{y} - 2\tilde{\delta}'\mathbf{Z}'(\tilde{k})\tilde{y} + \tilde{\delta}'\mathbf{Z}'(\tilde{k})\mathbf{Z}(\tilde{k})\tilde{\delta} \quad (2.27)$$

$$= \mathbf{Q}(\tilde{\delta}) \quad (2.28)$$

Untuk mendapatkan estimator dari parameter $\tilde{\delta}$, dilakukan derivatif parsial $\mathbf{Q}(\tilde{\delta})$ terhadap $\tilde{\delta}$:

$$\begin{aligned} \frac{\partial \mathbf{Q}(\tilde{\delta})}{\partial \tilde{\delta}} &= \frac{\partial (\tilde{y}'\tilde{y} - 2\tilde{\delta}'\mathbf{Z}'(\tilde{k})\tilde{y} + \tilde{\delta}'\mathbf{Z}'(\tilde{k})\mathbf{Z}(\tilde{k})\tilde{\delta})}{\partial \tilde{\delta}} \\ &= -2\mathbf{Z}'(\tilde{k})\tilde{y} + 2\mathbf{Z}'(\tilde{k})\mathbf{Z}(\tilde{k})\tilde{\delta} \end{aligned}$$

Jika derivatif parsial di atas disamakan dengan nol, diperoleh persamaan:

$$-\mathbf{Z}'(\tilde{k})\tilde{y} + \mathbf{Z}'(\tilde{k})\mathbf{Z}(\tilde{k})\hat{\tilde{\delta}} = 0 \quad (2.29)$$

Persamaan (2.23) dapat pula ditulis dalam bentuk:

$$\mathbf{Z}'(\tilde{k})\mathbf{Z}(\tilde{k})\hat{\tilde{\delta}} = \mathbf{Z}'(\tilde{k})\tilde{y} \quad (2.30)$$

Sehingga diperoleh estimator $\hat{\tilde{\delta}}$ adalah sebagai berikut:

$$\hat{\tilde{\delta}} = (\mathbf{Z}'(\tilde{k})\mathbf{Z}(\tilde{k}))^{-1} \mathbf{Z}'(\tilde{k})\tilde{y} \quad (2.31)$$

Dan estimator kurva regresi semiparametrik adalah sebagai berikut:

$$\hat{f}(\mathbf{X}, \mathbf{T}) = \mathbf{Z}(\tilde{k}) (\mathbf{Z}'(\tilde{k})\mathbf{Z}(\tilde{k}))^{-1} \mathbf{Z}'(\tilde{k})\tilde{y} \quad (2.32)$$

$$= \mathbf{A}(\tilde{k})\tilde{y} \quad (2.33)$$

dimana $\mathbf{A}(\tilde{k}) = \mathbf{Z}(\tilde{k}) (\mathbf{Z}'(\tilde{k})\mathbf{Z}(\tilde{k}))^{-1} \mathbf{Z}'(\tilde{k})$

2.6 Pemilihan Titik Knot Optimal

Dalam regresi semiparametrik spline *truncated*, hal penting yang berperan dalam mendapatkan estimator spline *truncated* terbaik adalah pemilihan titik knot yang optimal. Salah satu metode yang sering digunakan dalam memilih titik knot optimal adalah *Generalized Cross Validation* (GCV). Jika dibandingkan dengan metode lain, misalnya *Cross Validation* (CV) dan metode *Unbiased Risk* (UBR) ataupun *Generalized Maximum Likelihood* (GML), GCV secara teoritis memiliki sifat optimal asyptotik. Metode GCV juga memiliki kelebihan tidak memerlukan pengetahuan terhadap variansi populasi σ^2 serta metode GCV invarians terhadap transformasi. Metode GCV merupakan pengembangan dari CV (Wahba, 1990). Fungsi GCV untuk pemilihan titik knot optimal pada regresi semiparametrik dapat ditunjukkan dalam persamaan berikut:

$$GCV(\tilde{k}) = \frac{n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2}{\left[n^{-1} \text{trace} (\mathbf{I} - \mathbf{A}(\tilde{k})) \right]^2} \quad (2.34)$$

dimana y_i adalah variabel respon, \hat{y}_i adalah nilai estimasi variabel respon, $i = 1, 2, \dots, n$ yang merupakan jumlah observasi, $\tilde{k} = (K_1, K_2, \dots, K_r)'$ merupakan titik-titik knot, matrik $\mathbf{A}(\tilde{k}) = \mathbf{Z}(\tilde{k}) \left(\mathbf{Z}'(\tilde{k}) \mathbf{Z}(\tilde{k}) \right)^{-1} \mathbf{Z}'(\tilde{k})$ diperoleh dari persamaan $\hat{\mathbf{y}} = \mathbf{A}(\tilde{k}) \mathbf{y}$, dan \mathbf{I} adalah matriks identitas.

Untuk mencari titik knot optimal $\mathbf{k}_{opt} = (k_{1(opt)}, \dots, k_{r(opt)})'$ diperoleh melalui optimasi:

$$\min_{k_1, k_2, \dots, k_r} \{GCV(\tilde{k})\} = \min_{k_1, k_2, \dots, k_r} \left\{ \frac{n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2}{\left[n^{-1} \text{trace} (\mathbf{I} - \mathbf{A}(\tilde{k})) \right]^2} \right\} \quad (2.35)$$

2.7 Likelihood Ratio Test (LRT)

Likelihood Ratio Test adalah uji yang mempunyai daerah penolakan dari bentuk $\{x: \lambda(x) \leq c\}$, dimana c adalah konstanta yang memenuhi $0 \leq c \leq 1$. Jika X_1, X_2, \dots, X_n adalah sampel random dari populasi dengan *probability density function* (pdf) atau *probability mass function* (pmf) $f(x|\theta)$, (θ mungkin suatu vektor), maka fungsi *likelihood* didefinisikan sebagai berikut (Casella, G dan Berger, R.L., 2001):

$$L(\theta|x_1, \dots, x_n) = L(\theta|x) = \prod_{i=1}^n f(x_i|\theta). \quad (2.36)$$

Selanjutnya *likelihood ratio test statistic* untuk pengujian $H_0: \theta \in \Theta_0$ melawan $H_1: \theta \in \Theta_0^c$ dengan Θ adalah seluruh ruang parameter adalah:

$$\lambda(x) = \frac{\sup_{\Theta_0} L(\theta|x)}{\sup_{\Theta} L(\theta|x)} \quad (2.37)$$

Jika dilakukan maksimalisasi baik itu pada seluruh ruang parameter (*unrestricted maximization*) maupun pada subset ruang parameter (*restricted maximization*), maka akan ada hubungan yang jelas antara LRT dan *Maximum Likelihood Estimators* (MLE). Anggap bahwa $\hat{\theta}$ merupakan estimator dari θ yang diperoleh dari metode *Maximum Likelihood Estimators* (MLE), dengan *unrestricted maximization* dari $L(\theta|x)$. Dan juga anggap bahwa $\hat{\theta}_0$ merupakan estimator dari θ yang diperoleh dari metode MLE, dengan *restricted maximization* pada ruang parameter Θ_0 dimana $\hat{\theta}_0 = \hat{\theta}_0(x)$ adalah nilai dari $\theta \in \Theta_0$ yang memaksimalkan $L(\theta|x)$. Sehingga formula LRT statistik adalah sebagai berikut:

$$\lambda(x) = \frac{L(\hat{\theta}_0|x)}{L(\hat{\theta}|x)} \quad (2.38)$$

2.8 Pengujian Parameter dalam Regresi Parametrik

Untuk mengetahui apakah suatu variabel memberikan pengaruh yang signifikan dalam model atau tidak maka dilakukanlah suatu uji parameter. Ada dua macam uji parameter, yaitu uji parameter simultan atau serentak dan uji parameter secara parsial atau individu.

2.8.1 Uji Simultan/Serentak

Uji serentak merupakan suatu uji parameter kurva regresi secara simultan dengan menggunakan uji F. Misalkan diberikan suatu model regresi seperti pada persamaan (2.3) maka hipotesis yang digunakan untuk menguji model secara simultan atau serentak adalah sebagai berikut.

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

$$H_1 : \text{Paling sedikit terdapat satu } \beta_h \neq 0, h = 1, 2, \dots, p$$

Statistik uji yang digunakan adalah uji F yaitu :

$$F_{hitung} = \frac{MSR}{MSE} \quad (2.39)$$

Banyaknya variabel prediktor adalah p , sedangkan observasi pengamatan sebanyak n , dengan daerah penolakannya adalah : tolak H_0 apabila $F_{hitung} > F_{\alpha; (p, n-p-1)}$ atau tolak H_0 apabila $p\text{-value} < \alpha$ yang mengindikasikan bahwa paling sedikit ada satu parameter yang tidak sama dengan nol atau paling sedikit ada satu prediktor yang berpengaruh signifikan terhadap respon.

Tabel 2.1 Analisis Varians Model Regresi

Sumber Variasi	Derajat Kebebasan	Jumlah Kuadrat	Means Square	F_{hitung}
Regresi	p	$\beta'X'y - n\bar{y}^2$	$\frac{\beta'X'y - n\bar{y}^2}{p}$	$\frac{MSR}{MSE}$
Error	$n-p-1$	$y'y - \beta'X'y$	$\frac{y'y - \beta'X'y}{n-p-1}$	
Total	$n-1$	$y'y - n\bar{y}^2$	-	

Sumber: Draper and Smith, 1992

2.8.2 Uji Parsial/Individu

Uji parsial atau individu merupakan suatu uji untuk parameter kurva regresi secara individu dengan menggunakan uji t . Hipotesis yang digunakan untuk menguji model secara parsial atau individu adalah sebagai berikut :

$$H_0 : \beta_1 = 0 ; H_1 : \beta_1 \neq 0$$

$$H_0 : \beta_2 = 0 ; H_1 : \beta_2 \neq 0$$

⋮

$$H_0 : \beta_p = 0 ; H_1 : \beta_p \neq 0$$

Statistik uji yang digunakan adalah dengan menggunakan uji t yaitu :

$$t_{hitung} = \frac{\hat{\beta}_h}{se(\hat{\beta}_h)}, h = 1, 2, \dots, p \quad (2.40)$$

Dimana $se(\hat{\beta}_h)$ adalah standart error dari $\hat{\beta}_h$, dan daerah penolakan pada uji t yaitu tolak H_0 apabila $|t_{hitung}| > t_{\frac{\alpha}{2}; (n-p-1)}$ atau tolak H_0 apabila $p\text{-value} < \alpha$.

Selanjutnya pengujian parameter secara parsial ini akan dikembangkan pada regresi semiparametrik spline *truncated* yang akan diuraikan pada Bab 4.

Adapun kriteria yang digunakan untuk menentukan kebaikan model regresi adalah koefisien determinasi (R^2). Berikut adalah rumus untuk menghitung R^2 .

$$R^2 = \frac{\hat{\beta}'X'y - n\bar{y}^2}{y'y - n\bar{y}^2} \times 100\% \quad (2.41)$$

2.9 Pemeriksaan Asumsi Residual

Pada model regresi semiparametrik spline *truncated* diasumsikan bahwa *error* random berdistribusi normal independent dengan mean nol dan variansi σ^2 (Wahba, 1990). Oleh karena itu sebelum melakukan analisis dan mengambil keputusan dari hasil pemodelan maka harus dilakukan uji asumsi residual terlebih dahulu. Uji dan deteksi asumsi residual yang dilakukan adalah uji normalitas, uji independen dan uji identik.

2.9.1 Uji Asumsi Normalitas

Pengujian hipotesis melibatkan distribusi tertentu, yaitu distribusi F dan distribusi *t-Student*. Distribusi tersebut digunakan untuk menguji signifikansi parameter baik secara simultan maupun parsial. Oleh karena itu diperlukan pengujian distribusi normal pada residualnya, jika residual tidak memenuhi asumsi normal maka pengujian parameter menjadi tidak akurat. Cara mendeteksi apakah residual berdistribusi normal dapat dilihat pada *normal probability plot residual*. Apabila plot yang dihasilkan membentuk garis lurus maka residual dari model regresi tersebut cenderung mengikuti distribusi normal.

Pengujian distribusi normal dapat dilakukan dengan metode Uji *Kolmogorov-Smirnov* yang juga dikenal dengan uji kesesuaian model (*Goodness of Fit test*). Hipotesis yang digunakan adalah sebagai berikut:

$$H_0 : F_0(x) = F(x) \text{ (Residual berdistribusi Normal)}$$

$$H_1 : F_0(x) \neq F(x) \text{ (Residual tidak berdistribusi Normal).}$$

Statistik Uji :

$$D = \max|F_0(x) - S_N(x)| \quad (2.42)$$

$F_0(x)$ adalah fungsi distribusi frekuensi kumulatif teoritis, sedangkan $S_N(x) = \frac{k}{n}$ merupakan fungsi peluang kumulatif yang diobservasi dari satu sampel random dengan N observasi. k adalah banyaknya observasi yang sama atau kurang dari x . Kesimpulan untuk menolak H_0 jika $D > q_{(1-\alpha)}$ dimana q adalah nilai berdasarkan tabel *Kolmogorov-Smirnov*.

2.9.2 Uji Asumsi Independen

Pemeriksaan asumsi independensi residual digunakan untuk mendeteksi korelasi antara residual. Salah satu metode yang dapat digunakan adalah dengan menggunakan statistik uji *Durbin-Watson*. Hipotesis yang digunakan adalah sebagai berikut:

$$H_0: \rho = 0 \text{ (tidak ada korelasi antara residual)}$$

$$H_1: \rho \neq 0 \text{ (ada korelasi antara residual)}$$

dengan statistik uji

$$d_{hitung} = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}$$

Adapun kesimpulannya adalah sebagai berikut :

- i. Tolak H_0 jika $d_{hitung} < d_{L, \frac{\alpha}{2}}$, dimana d_L adalah batas nilai bawah tabel.
- ii. Gagal tolak H_0 jika $d_{hitung} > d_{U, \frac{\alpha}{2}}$, dimana d_U adalah batas nilai atas tabel.
- iii. Jika $d_{L, \frac{\alpha}{2}} < d_{hitung} < d_{U, \frac{\alpha}{2}}$ maka tidak ada keputusan.

2.9.3 Uji Asumsi Identik

Uji identik dapat juga disebut uji homogenitas varians residual. Homogenitas varians residual didasarkan pada sifat $E(\varepsilon_i) = 0$ dimana $V(\varepsilon_i) = \sigma^2$. Pada kondisi ini, varians residual adalah konstan. Pelanggaran terhadap asumsi ini disebut heteroskedastis yaitu keadaan dimana variansi residual tidak homogen. Hal ini menyebabkan estimasi koefisien kurang akurat/tidak efisien (Gujarati, 1995). Cara mendeteksi heteroskedastisitas dapat dideteksi dengan membuat *scatter plot* antara residual dan estimasi respon (\hat{y}). Apabila plot menunjukkan sebaran data yang tidak random atau membentuk tren/pola tertentu,

maka terjadi kasus heteroskedastis residual, sehingga perlu diatasi antara lain dengan transformasi variabel menggunakan *Weighted Least Square* (WLS) (Gujarati, 1995).

Cara lain yang dapat digunakan dalam mendeteksi adanya heteroskedastisitas adalah dengan menggunakan Uji *Glejser*. Pengujian ini dilakukan dengan cara meregresikan harga mutlak residual dengan variabel prediktor (x).

$$|e_i| = G(x_i) + \varepsilon_i, i = 1, 2, \dots, n \quad (2.43)$$

Berdasarkan Persamaan (0.29) jika diketahui terdapat variabel prediktor yang signifikan dalam model maka hal ini mengindikasikan bahwa residual cenderung tidak homogen. Hipotesis yang digunakan dalam uji *Glejser* sebagai berikut:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1 : \text{minimal terdapat satu } \sigma_i^2 \neq \sigma^2, i = 1, 2, \dots, n$$

Statistik uji yang digunakan adalah sebagai berikut:

$$F_{hitung} = \frac{MSR}{MSE} = \left(\frac{\sum_{i=1}^n (|\hat{e}_i| - |\bar{e}|)^2}{p-1} \right) / \left(\frac{\sum_{i=1}^n (|e_i| - |\hat{e}_i|)^2}{n-p} \right) \quad (2.44)$$

p adalah banyaknya parameter model *glejser*. Daerah penolakan H_0 adalah jika nilai $F_{hitung} > F_{\alpha; ((p-1), (n-p))}$. Apabila pada kesimpulan dihasilkan penolakan H_0 , maka dapat dinyatakan bahwa artinya terdapat minimal satu $\sigma_i^2 \neq \sigma^2$ dan itu berarti terdapat heteroskedastisitas.

2.10 Teorema Dasar Terkait dengan Aljabar Matriks

Beberapa teorema dasar terkait dengan aljabar matriks yang digunakan untuk menyelesaikan estimasi parameter dan kajian pengujian hipotesis secara parsial berikut ini berdasarkan Rencher dan Scaalje (2007).

1. Idempoten (Rencher, 2007).

Matrik \mathbf{A} dikatakan idempoten jika $\mathbf{A}^2 = \mathbf{A}$.

2. Teorema 1 (Rencher, 2007)

Jika \mathbf{A} adalah $n \times p$ dan \mathbf{B} adalah $p \times n$, maka $tr(\mathbf{AB}) = tr(\mathbf{BA})$.

3. Teorema 2 (Rencher, 2007)

Jika matrik \mathbf{A} mempunyai rank r serta simetris dan *idempotent*, maka $rank(\mathbf{A}) = tr(\mathbf{A}) = r$.

4. Teorema 3 (Rencher, 2007)

Jika $u = \tilde{a}'\tilde{x} = \tilde{x}'\tilde{a}$, dimana $\tilde{a}' = (a_1, a_2, \dots, a_p)$ adalah vektor konstan.

Maka

$$\frac{\partial u}{\partial \tilde{x}} = \frac{\partial(\tilde{a}'\tilde{x})}{\partial \tilde{x}} = \frac{\partial(\tilde{x}'\tilde{a})}{\partial \tilde{x}} = \tilde{a}.$$

5. Teorema 4 (Rencher, 2007)

Jika $u = \tilde{x}'\mathbf{A}\tilde{x}$, dimana \mathbf{A} adalah matrik simetrik konstan. Maka

$$\frac{\partial u}{\partial \tilde{x}} = \frac{\partial(\tilde{x}'\mathbf{A}\tilde{x})}{\partial \tilde{x}} = 2\mathbf{A}\tilde{x}.$$

6. Teorema 5 (Rencher, 2007)

Jika \mathbf{A} dan \mathbf{B} adalah matrik dengan ukuran $n \times m$, maka $(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$.

7. Teorema 6 (Rencher, 2007)

Jika \mathbf{A} adalah matrik berukuran $n \times p$ dan \mathbf{B} adalah matrik berukuran $p \times n$ maka $(\mathbf{A}\mathbf{B})' = \mathbf{B}'\mathbf{A}'$.

8. Theorema 7 (Rencher, 2007)

Corollary 1 menyatakan bahwa jika $\tilde{y} \sim N(\tilde{\mu}, \sigma^2\mathbf{I})$ dan \mathbf{A} matrik simetris dengan rank r , maka $\tilde{y}'\mathbf{A}\tilde{y}/\sigma^2$ adalah $\chi^2(r, \tilde{\mu}'\mathbf{A}\tilde{\mu}/2\sigma^2)$, jika dan hanya jika \mathbf{A} adalah *idempotent*.

9. Teorema 8 (Rencher, 2007)

Corollary 2 menyatakan bahwa jika $\tilde{y} \sim N(\tilde{\mu}, \sigma^2\mathbf{I})$, maka $\tilde{y}'\mathbf{A}\tilde{y}$ dan $\tilde{y}'\mathbf{B}\tilde{y}$ adalah *independent* jika dan hanya jika $\mathbf{A}\mathbf{B} = 0$ (atau ekuivalen, $\mathbf{B}\mathbf{A} = 0$).

2.11 Tinjauan Angka Harapan Hidup

2.11.1 Angka Harapan Hidup (AHH)

Angka Harapan Hidup (*Life Expectancy at birth*) menurut *World Bank* adalah rata-rata jumlah tahun harapan hidup sekelompok orang yang lahir pada tahun yang sama, dengan asumsi kematian pada usia masing-masing tersebut tetap konstan di masa mendatang. Definisi itu juga digunakan Indonesia dalam menghitung AHH yang dilakukan oleh BPS. Selain itu, WHO menyatakan bahwa AHH sebenarnya merupakan alat ukur derajat kesehatan suatu negara yang

menjadi salah satu pilar utama penentu kualitas hidup manusia. Peningkatan derajat kesehatan dapat mengindikasikan keberhasilan program kesehatan dan pembangunan di bidang sosial ekonomi yang secara tidak langsung juga meningkatkan AHH. Menurut pakar ilmu kesehatan, Hendrick L. Blum, peningkatan derajat kesehatan masyarakat dapat diukur dari tingkat mortalitas/kematian dan morbiditas/kesakitan. Derajat kesehatan dipengaruhi oleh empat faktor yaitu faktor lingkungan, perilaku hidup sehat, pelayanan kesehatan, dan keturunan. Dari keempat faktor tersebut faktor lingkungan dan perilaku hidup sehat memberikan kontribusi pengaruh yang dominan terhadap derajat kesehatan. Sementara itu, Halicioglu (2011) melakukan penelitiannya di Turki dengan menggunakan metode *Autoregressive Distributed Lag (ARDL)* menyatakan terdapat tiga faktor utama yang mempengaruhi AHH yaitu faktor ekonomi, sosial, dan lingkungan. Faktor sosial yaitu banyaknya rokok yang dihisap per kapita dan tingkat buta huruf. Faktor lingkungan terdiri dari tingkat kejahatan dan urbanisasi. Sedangkan faktor ekonomi meliputi rata-rata pengeluaran real kesehatan per kapita. Singh (2006) dalam bukunya juga menyebutkan bahwa faktor-faktor yang diduga mempengaruhi AHH adalah pendidikan, pekerjaan, kemiskinan, pengangguran, pengeluaran, dan kualitas rumah.

Sebagai salah satu wujud keseriusan pemerintah dalam mengawal peningkatan kualitas hidup masyarakat dalam hal ini derajat kesehatan, pemerintah melalui Keputusan Menteri Kesehatan Nomor 1798/MENKES/SK/XII/2020 meluncurkan program Indeks Pembangunan Kesehatan Masyarakat (IPKM) pada tahun 2010. IPKM merupakan indeks komposit yang terdiri dari 24 indikator yang digunakan untuk mengetahui derajat kesehatan setiap provinsi. Adapun beberapa indikator yang digunakan diantaranya adalah imunisasi lengkap dan penggunaan air bersih.

Imunisasi merupakan salah satu upaya pencegahan yang dilakukan untuk mengurangi AKB. Menurunnya AKB dapat berpengaruh terhadap kenaikan AHH. Pada masa awal kehidupannya, bayi sangat rentan terkena penyakit berbahaya, seperti penyakit saluran pernapasan akut, polio, kerusakan hati, tetanus, dan campak. Anak yang terkena penyakit tersebut memiliki risiko kematian yang tinggi. Jika tidak sampai meninggal dunia, serangan virus dan penyakit tersebut

akan menyebabkan derita fisik dan mental berkepanjangan. Imunisasi merupakan perlindungan yang ampuh untuk mencegah beberapa penyakit berbahaya tersebut dengan cara merangsang kekebalan tubuh bayi.

Pola hidup sehat akan sangat mempengaruhi derajat kesehatan manusia dan berdampak pada usia harapan hidup. Sementara itu hasil studi IHME, tantangan serius yang sedang dihadapi oleh negara berkembang era sekarang adalah menjalankan pola hidup dan lingkungan yang sehat untuk mencapai usia harapan hidup yang lebih lama. Penyakit yang bermunculan belakangan ini tidak melihat faktor usia dan bukan dari penyakit keturunan, namun lebih banyak bersumber dari pola hidup yang tidak sehat, terutama faktor makanan dan minuman yang dikonsumsi.

Selain dipengaruhi oleh perilaku hidup sehat, AHH juga dipengaruhi oleh faktor sosial dan ekonomi. Salah satu indikator kesejahteraan masyarakat di suatu daerah adalah angka kemiskinan di daerah tersebut. Kemiskinan adalah keadaan dimana terjadi ketidakmampuan untuk memenuhi kebutuhan dasar seperti makanan, pakaian, tempat berlindung, pendidikan, dan kesehatan. Kemiskinan dapat disebabkan oleh kelangkaan alat pemenuh kebutuhan dasar, ataupun sulitnya akses terhadap pendidikan dan pekerjaan. Semakin tinggi pendapatan penduduk akan semakin mudah mengakses berbagai sumber kebutuhan termasuk akses fasilitas kesehatan. Penduduk yang mampu mengatasi kesakitannya akan menikmati kesehatannya lebih lama dan berumur panjang. Dengan demikian jika suatu wilayah memiliki pendapatan yang baik diharapkan dapat menikmati usia harapan hidup yang lebih lama. Belon (2011) dalam penelitiannya mengenai kesenjangan AHH di Brazil menyimpulkan bahwa kondisi sosial ekonomi yaitu pendapatan rumah tangga sangat mempengaruhi AHH di wilayah tersebut. Pengurangan kesenjangan sosial mengakibatkan kualitas hidup dan kesehatan meningkat dari waktu ke waktu. Perluasan cakupan dan kebijakan pemerintah memiliki efek positif pada pengurangan angka kematian dan konsekuensinya adalah peningkatan harapan hidup, terutama bagi penduduk miskin. Dalam tulisannya juga menyebutkan pendapat ahli kesehatan, Cockerham, dalam bukunya *Health Behavior* bahwa kepemilikan sumber daya, dipandang sebagai akumulasi sarana untuk membeli pilihan dan peluang dalam gaya hidup, sehingga

masyarakat dengan kondisi sosial ekonomi baik akan lebih mudah mengadopsi perilaku hidup sehat.

Penelitian mengenai AHH dengan berbagai metode sudah banyak dilakukan, antara lain: Ayuni (2015) melakukan pemodelan angka harapan hidup di Provinsi Jawa Timur dengan menggunakan metode Feed Forward Neural Network (FFNN). Sugiantari (2013) menganalisis faktor-faktor yang mempengaruhi angka harapan hidup di provinsi Jawa Timur dengan menggunakan regresi semiparametrik spline.

2.11.2 Faktor-Faktor Yang Mempengaruhi Angka Harapan Hidup

Berdasarkan kerangka konsep derajat kesehatan dan beberapa penelitian dalam melakukan kajian mengenai AHH maka penelitian ini akan menggunakan kerangka konseptual yang melibatkan faktor sosial, ekonomi dan perilaku hidup sehat. Faktor sosial ekonomi yang digunakan adalah persentase penduduk miskin dan pendidikan, sedangkan faktor perilaku hidup sehat adalah penggunaan air minum bersih, pemberian imunisasi lengkap, dan persentase penduduk yang mengalami keluhan kesehatan.

1. Persentase Penduduk Miskin

Faktor sosial yang digunakan dalam penelitian ini adalah persentase penduduk miskin. Pada tahun 2015 persentase penduduk miskin di Indonesia adalah 11,22 persen. Persentase penduduk miskin paling tinggi adalah Provinsi Papua yaitu 27,33 persen. Kondisi sebaliknya terjadi di Provinsi DKI Jakarta yang memiliki 3,93 persen penduduk miskin. Kesenjangan sosial ini diduga menjadi salah satu pemicu terjadinya kesenjangan AHH di Indonesia. Hipotesis antara persentase penduduk miskin dengan AHH di Indonesia memiliki hubungan negatif. Semakin sedikit jumlah penduduk miskin maka diharapkan AHH di Indonesia semakin meningkat. Penelitian sebelumnya yang menggunakan variabel angka kemiskinan adalah Singh (2006) dan Belon (2011).

2. Sumber Air Minum Bersih

Beberapa indikator yang menggambarkan kondisi lingkungan antara lain rumah sehat, air bersih, sarana sanitasi (pembuangan air limbah, tempat sampah,

dan kepemilikan jamban), tempat umum dan pengelolaan makanan, serta sarana pengolahan limbah di sarana pelayanan kesehatan. Air bersih merupakan kebutuhan yang sangat penting bagi rumah tangga dalam kehidupan sehari-hari. Ketersediaan dalam jumlah yang cukup terutama untuk keperluan minum dan masak merupakan tujuan dari program penyediaan air bersih yang terus menerus diupayakan oleh pemerintah.

Beberapa daerah di Indonesia masih ada yang mengalami kesulitan untuk mendapatkan air bersih. Selain faktor alam berupa kekeringan, kesulitan tersebut dikarenakan oleh semakin meningkatnya polusi yang mencemari air sehingga menjadi air yang tidak layak untuk dikonsumsi bahkan sangat berbahaya jika dikonsumsi oleh manusia. Pada tahun 2015, rumah tangga yang mengkonsumsi air bersih sebanyak 70,97 persen. Rumah tangga yang paling sedikit mengkonsumsi air bersih adalah Provinsi Bengkulu dan Sulawesi Barat yaitu 41,08 persen dan 53,89 persen. Sedangkan 93,40 persen rumah tangga di Provinsi DKI Jakarta sudah menggunakan air bersih. Untuk memperkecil terjadinya penyakit atau gangguan kesehatan sebagai akibat dari lingkungan yang kurang sehat, harus dilakukan berbagai upaya untuk meningkatkan kualitas lingkungan.

Penelitian sebelumnya yang menggunakan variabel air minum bersih adalah Firdial (2010) dan Riset Kesehatan Dasar (RIKESDAS) yang dilakukan oleh Kemenkes dalam menyusun IPKM Indonesia.

3. Keluhan Kesehatan

Keluhan kesehatan mengindikasikan tingkat kesakitan yang dialami oleh masyarakat. Pada umumnya, masyarakat yang memiliki derajat kesehatan tinggi akan memiliki kecenderungan hidup lebih lama. Sebaliknya, masyarakat dengan derajat kesehatan rendah dengan kata lain memiliki tingkat kesakitan yang tinggi memiliki kecenderungan berumur pendek. Menurut teori demografi, tingkat kesakitan/morbiditas kumulatif pada akhirnya akan menghasilkan sebuah peristiwa kematian/mortalitas. Dengan demikian, derajat kesehatan suatu bangsa dapat direpresentasikan oleh AHH bangsa tersebut. Penelitian sebelumnya yang menggunakan faktor kesehatan adalah *World Bank* (2015) melalui studi *Global Burden Disease (GBD)*, *Institute for Health Metrics and Evaluation (IHME)*.

4. Imunisasi Lengkap

Imunisasi boleh saja diberikan pada semua umur, namun beberapa imunisasi akan lebih efektif apabila diberikan pada usia tertentu. Misalnya ada yang efektif pada bayi, anak-anak, remaja, dewasa bahkan manula. Menurut Kementerian Kesehatan, ada lima vaksin wajib yang harus diberikan kepada bayi di bawah usia satu tahun. Anak yang sudah diberikan kelima vaksin tersebut dikatakan sudah mendapatkan vaksin lengkap. Adapun vaksin tersebut adalah :

- a. BCG (*Bacillus Calmette Guerin*) merupakan vaksinasi untuk mencegah penyakit TBC, diberikan pada bayi baru lahir atau anak, dengan suntikan pada kulit pangkal lengan atas. Bekas suntikan kemudian akan membentuk tonjolan kecil jaringan parut pada kulit lengan atas. Suntikan BCG diberikan kepada anak sebanyak 1 kali.
- b. DPT (*Difteri, Pertusis, Tetanus*) merupakan vaksin untuk mencegah penyakit Difteri, Pertusis, dan Tetanus yang diberikan pada bayi berumur 3 bulan ke atas, dengan suntikan pada paha, diulang 1 bulan dan 2 bulan kemudian, sehingga suntikan imunisasi DPT lengkap pada balita berjumlah 3 kali (kadang-kadang selang waktu antar suntikan bisa lebih dari 1 bulan).
- c. Polio merupakan vaksin untuk mencegah penyakit polio yang diberikan pada bayi berumur 3 bulan ke atas, dengan memberikan 3 tetes cairan vaksin berwarna merah muda atau putih ke dalam mulut anak diberikan biasanya bersama-sama DPT. Imunisasi polio lengkap pada balita berjumlah 3 kali.
- d. Campak/morbilli merupakan vaksin untuk mencegah penyakit campak/morbilli, biasanya diberikan pada bayi berumur 9 sampai 12 bulan, dengan suntikan di bawah kulit pada paha sebanyak 1 kali.
- e. Hepatitis B adalah suntikan secara intramuskular (suntikan ke dalam otot) biasanya di paha yang diberikan pada bayi untuk mencegah penyakit Hepatitis B. Suntikan ini diberikan sebanyak 3 kali. Suntikan pertama diberikan pada bayi berumur 0 bulan, kedua pada bayi 1 bulan, dan ke 3 pada bayi 6 bulan.

Penelitian sebelumnya yang menggunakan variabel imunisasi lengkap pada balita adalah Sugiantari (2013) dan Riset Kesehatan Dasar (RIKESDAS) yang dilakukan oleh Kemenkes dalam menyusun IPKM Indonesia.

5. Rata-rata Lama Sekolah (MYS)

Pendidikan akan memberikan pengaruh terhadap pola berpikir dan selanjutnya akan menentukan cara memberi respon terhadap sesuatu yang datang dari luar. Dengan semakin tingginya tingkat pendidikan, diharapkan seseorang akan semakin mudah dalam menyerap, memilih, beradaptasi atau mengembangkan segala bentuk informasi dan pengetahuan baru untuk kehidupannya. Salah satu indikator tingkat pendidikan adalah rata-rata lama sekolah (MYS). Sementara itu menurut catatan BPS, MYS Indonesia pada tahun 2015 adalah 7,84. MYS tertinggi di Provinsi DKI Jakarta yaitu 10,70 tahun. Dengan demikian penduduk DKI Jakarta yang berusia 15 tahun ke atas rata-rata berpendidikan sampai bangku Sekolah Menengah Umum. Sedangkan MYS terendah adalah 5,99 tahun di Provinsi Papua. Sehingga dapat dikatakan penduduk yang berusia 15 tahun ke atas di Provinsi Papua secara rata-rata berpendidikan sampai dengan Sekolah Dasar.

BAB 3

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder, yaitu data yang berasal dari Survei Sosial Ekonomi Nasional (Susenas) tahun 2015 yang telah dipublikasikan oleh Badan Pusat Statistik (BPS) dalam Statistik Kesejahteraan Rakyat 2015 dan data pendidikan yang telah dipublikasikan dalam Indeks Pembangunan Manusia 2015. Unit observasi yang digunakan dalam penelitian ini adalah seluruh provinsi yang ada di Indonesia, yaitu sebanyak 34 provinsi.

3.2 Variabel Penelitian

Dalam penelitian ini variabel respon dan variabel-variabel prediktor yang digunakan adalah sebagai berikut.

1. Variabel Respon:

y : Angka Harapan Hidup di Indonesia.

2. Variabel predictor:

x_1 : Persentase penduduk miskin.

x_2 : Persentase rumah tangga yang menggunakan air minum bersih.

x_3 : Persentase penduduk yang mengalami keluhan kesehatan.

x_4 : Persentase balita yang mendapat imunisasi lengkap.

x_5 : MYS.

3.3 Definisi Operasional Variabel Penelitian

Berikut adalah definisi operasional dari variabel-variabel yang digunakan dalam penelitian.

1. Angka Harapan Hidup

Angka Harapan Hidup pada suatu umur tertentu adalah rata-rata tahun hidup yang masih akan dijalani oleh seseorang yang telah berhasil mencapai umur tertentu pada suatu tahun tertentu dalam situasi mortalitas yang berlaku di

lingkungan masyarakatnya. Ukuran yang sering digunakan adalah AHH saat lahir yang mencerminkan kondisi kesehatan pada saat itu. Sehingga pada umumnya ketika membicarakan AHH yang dimaksud adalah rata-rata jumlah tahun yang akan dijalani oleh seseorang sejak orang tersebut lahir (BPS, 2010).

2. Persentase penduduk miskin

BPS menggunakan konsep kemampuan memenuhi kebutuhan dasar (*basic needs approach*) dalam mengukur kemiskinan. Dengan pendekatan ini, kemiskinan dipandang sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran (BPS, 2016). Penduduk yang dikategorikan sebagai penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran perkapita dibawah Garis Kemiskinan (GK). GK adalah penjumlahan dari Garis Kemiskinan Makanan (GKM) dan Garis Kemiskinan Non-Makanan (GKNM). Tahap pertama dalam penghitungan GK adalah dengan menentukan penduduk referensi yaitu 20 persen penduduk yang berada di atas GKS. GKS sendiri adalah GK periode sebelumnya yang di-inflate dengan inflasi umum (IHK). Dari penduduk referensi kemudian dihitung Garis Kemiskinan Makanan (GKM). GKM adalah jumlah nilai pengeluaran dari 52 komoditi dasar makanan yang riil dikonsumsi penduduk referensi yang kemudian disetarakan dengan 2100 kilokalori perkapita/hari. Penyetaraan nilai pengeluaran kebutuhan minimum makanan dilakukan dengan menghitung harga rata-rata kalori dari ke-52 komoditi tersebut. Selanjutnya GKM tersebut disetarakan dengan 2100 kilokalori dengan mengalikan 2100 terhadap harga implisit rata-rata kalori menurut daerah dari penduduk referensi. Sementara GKNM merupakan penjumlahan nilai kebutuhan minimum dari komoditi-komoditi non-makanan terpilih yang meliputi perumahan, sandang, pendidikan dan kesehatan. Nilai kebutuhan minimum per komoditi/sub-kelompok non-makanan dihitung dengan menggunakan suatu rasio pengeluaran komoditi/sub-kelompok tersebut terhadap total pengeluaran komoditi/sub-kelompok yang tercatat dalam data Susenas modul konsumsi. Rasio tersebut dihitung dari hasil Survei Paket Komoditi Kebutuhan Dasar 2004, yang

dilakukan untuk mengumpulkan data pengeluaran konsumsi rumah tangga per komoditi non-makanan yang lebih rinci dibandingkan data Susenas modul konsumsi. Data yang digunakan bersumber dari Susenas Konsumsi Pengeluaran Maret 2015 yang tercantum dalam Publikasi Statistik Indonesia Tahun 2015.

3. Persentase rumah tangga yang menggunakan air minum bersih
Air minum bersih adalah air minum yang berasal dari sumber air minum bersih yaitu air kemasan, air isi ulang, leding, sumur bor, sumur terlindung, dan mata air terlindung yang memiliki jarak ke tempat penampungan lebih dari 10 meter (BPS).
4. Persentase penduduk yang mengalami keluhan kesehatan
Keluhan kesehatan adalah keadaan seseorang yang mengalami gangguan kesehatan atau kejiwaan, baik karena penyakit akut, penyakit kronis (meskipun selama sebulan terakhir tidak mempunyai keluhan), kecelakaan, kriminal atau hal lain. Jenis keluhan kesehatan yang digunakan oleh BPS dalam membangun data Susenas adalah panas, batuk, pilek, asma, napas sesak/cepat, diare, sakit kepala berulang, sakit gigi, dan keluhan lainnya
5. Persentase balita yang mendapatkan imunisasi lengkap
Menurut konsep BPS dalam pengumpulan data Susenas, imunisasi adalah suatu upaya yang dilakukan untuk mencegah terjadinya penyakit dengan cara memberikan mikroorganisme bibit penyakit berbahaya yang telah dilemahkan (vaksin) ke dalam tubuh sehingga merangsang sistem kekebalan tubuh terhadap jenis antigen tersebut di masa yang akan datang. Anak dikatakan sudah mendapat imunisasi lengkap jika anak tersebut telah diimunisasi 1 kali BCG, minimal 3 kali hepatitis B, minimal 3 kali DPT, minimal 3 kali Polio, dan 1 kali campak.
6. MYS
Menurut konsep BPS, rata-rata lama sekolah (MYS) adalah jumlah tahun belajar penduduk usia 15 tahun ke atas yang telah diselesaikan dalam pendidikan formal (tidak termasuk tahun yang mengulang).
Rumusan yang digunakan untuk menghitung MYS adalah sebagai berikut :

$$MYS = \frac{1}{P_{15+}} \sum_{i=1}^{P_{15+}} (\text{Lama sekolah penduduk ke } - i)$$

P_{15+} adalah jumlah penduduk berusia 15 tahun ke atas.

Lama sekolah penduduk ke $- i$ adalah

- Tidak pernah sekolah = 0
- Masih sekolah di SD sampai dengan S1= konversi ijazah terakhir + kelas terakhir - 1
- Masih sekolah di S2/S3= konversi ijazah terakhir + 1
- Tidak sekolah lagi dan tamat di kelas terakhir = konversi ijazah terakhir
- Tidak sekolah lagi dan tidak tamat di kelas terakhir = konversi ijazah terakhir + kelas terakhir - 1.

3.4 Tahapan Penelitian

Tujuan pertama dalam penelitian ini adalah melakukan kajian tentang pengujian hipotesis parsial pada model regresi semiparametrik spline *truncated*. Tahapan penelitian yang dilakukan untuk menyelesaikan tujuan pertama ini adalah sebagai berikut.

1. Asumsikan data berpasangan $(x_1, x_2, \dots, x_p, t_1, t_2, \dots, t_q, y)$ mengikuti model regresi semiparametrik.

$$y = g(x_1, x_2, \dots, x_p) + f(t_1, t_2, \dots, t_q) + \varepsilon$$

2. Kurva regresi $g(x_1, x_2, \dots, x_p)$ merupakan komponen parametrik, dihampiri dengan fungsi linier.

$$g(x_1, \dots, x_p) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$$

3. Kurva regresi $f(t_1, t_2, \dots, t_q)$ merupakan komponen nonparametrik, dihampiri dengan fungsi spline *truncated* derajat m dan titik-titik knot K_1, K_2, \dots, K_r .

$$f(t_1, t_2, \dots, t_q) = \sum_{l=1}^q \left(\sum_{j=1}^m \gamma_{jl} t_l^j + \sum_{k=1}^r \gamma_{(m+k)l} (t_l - K_{kl})_+^m \right)$$

4. Diberikan model regresi semiparametrik spline *truncated*.

$$y = \sum_{h=0}^p \beta_h x_h + \sum_{l=1}^q \left(\sum_{j=1}^m \gamma_{jl} t_l^j + \sum_{k=1}^r \gamma_{(m+k)l} (t_l - K_{kl})_+^m \right) + \varepsilon$$

5. Merumuskan uji hipotesis parsial untuk parameter komponen parametrik dalam model regresi semiparametrik spline *truncated* sebagai berikut.

$$H_0 : \beta_h = 0 \text{ vs } H_1 : \beta_h \neq 0, h = 1, 2, \dots, p$$

6. Menentukan ruang parameter dibawah $H_0 (\omega)$.

$$\omega = \{(\beta_1, \beta_2, \dots, \beta_{h-1}, \beta_{h+1}, \dots, \beta_p, \gamma_{11}, \gamma_{21}, \dots, \gamma_{(r+m)q}, \sigma^2); \beta_h = 0\}$$

7. Mencari fungsi likelihood dibawah ruang $H_0 (\omega)$.

8. Memaksimumkan fungsi likelihood dibawah $\omega (L(\hat{\omega}))$.

9. Menentukan ruang parameter dibawah $H (\Omega)$.

$$\Omega = \{(\beta_1, \beta_2, \dots, \beta_p, \gamma_{11}, \gamma_{12}, \dots, \gamma_{(r+m)q}, \sigma^2)\}$$

10. Mencari fungsi likelihood dibawah ruang $H (\Omega)$.

11. Memaksimumkan fungsi likelihood dibawah $\Omega (L(\hat{\Omega}))$.

12. Membuat ratio likelihood.

$$\lambda(x_1, \dots, x_p, t_1, \dots, t_q) = \frac{L(\hat{\omega})}{L(\hat{\Omega})}, 0 \leq \lambda \leq 1$$

13. Mencari statistik uji berdasarkan langkah (12).

14. Mendapatkan distribusi statistik uji.

15. Menentukan daerah penolakan hipotesis H_0 , melalui

$$\lambda(x_1, \dots, x_p, t_1, \dots, t_q) < k, \text{ untuk suatu konstanta } k.$$

16. Merumuskan uji hipotesis parsial untuk parameter komponen nonparametrik dalam model regresi semiparametrik spline *truncated* sebagai berikut.

$$H_0: \gamma_{kl} = 0 \text{ vs } H_1: \gamma_{kl} \neq 0, l = 1, 2, \dots, q; k = 1, 2, \dots, r + m$$

dengan langkah serupa pada pengujian hipotesis parameter komponen parametrik.

Tujuan kedua dari penelitian ini adalah memodelkan AHH dengan menggunakan regresi semiparametrik spline *truncated* dan selanjutnya melakukan pengujian hipotesis terhadap parameter-parameter model. Langkah-langkah yang dilakukan adalah:

1. Membuat plot variabel respon (y) dengan semua variabel-variabel prediktor $(x_1, t_1, t_2, t_3, t_4)$.
2. Menentukan variabel komponen parametrik dan nonparametrik.

3. Memodelkan variabel respon dan variabel prediktor menggunakan estimator semiparametrik spline *truncated*.
4. Memodelkan data dengan regresi semiparametrik spline *truncated* dengan satu, dua, tiga, dan kombinasi knot.
5. Memilih titik knot yang optimal berdasarkan metode GCV.
6. Melakukan pengujian signifikansi parameter secara simultan.
7. Melakukan pengujian signifikansi parameter secara parsial.
8. Melakukan pengujian terhadap asumsi independen, identik dan distribusi normal untuk residual.
9. Membuat interpretasi model regresi semiparametrik spline *truncated* dalam memodelkan Angka Harapan Hidup di Indonesia.

BAB 4

HASIL DAN PEMBAHASAN

Pada bab ini akan dilakukan kajian mengenai pengujian hipotesis secara parsial pada regresi semiparametrik spline *truncated*. Kemudian sesuai dengan tujuan pada penelitian ini, hasilnya akan diterapkan pada data AHH Tahun 2015 di Indonesia.

4.1 Model Regresi Semiparametrik Spline *Truncated*

Diberikan data berpasangan $(x_1, x_2, \dots, x_p, t_1, t_2, \dots, t_q, y)$ dan hubungan data berpasangan tersebut mengikuti model regresi semiparametrik spline *truncated* sebagai berikut:

$$y = g(x_1, x_2, \dots, x_p) + f(t_1, t_2, \dots, t_q) + \varepsilon \quad (4.1)$$

dimana $g(x_1, x_2, \dots, x_p)$ merupakan komponen parametrik, dan dihipotesis dengan fungsi linier sebagai berikut:

$$g(x_1, \dots, x_p) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (4.2)$$

sedangkan $f(t_1, t_2, \dots, t_q)$ merupakan komponen nonparametrik, dihipotesis dengan fungsi spline *truncated* derajat m dan titik-titik knot K_1, K_2, \dots, K_r sebagai berikut:

$$f(t_1, t_2, \dots, t_q) = \sum_{l=1}^q \left(\sum_{j=1}^m \gamma_{jl} t_l^j + \sum_{k=1}^r \gamma_{(m+k)l} (t_l - K_{kl})_+^m \right) \quad (4.3)$$

Berdasarkan persamaan (4.2) dan persamaan (4.3) maka persamaan (4.1) akan menjadi:

$$y = \sum_{h=0}^p \beta_h x_h + \sum_{l=1}^q \left(\sum_{j=1}^m \gamma_{jl} t_l^j + \sum_{k=1}^r \gamma_{(m+k)l} (t_l - K_{kl})_+^m \right) + \varepsilon \quad (4.4)$$

4.1.1 Estimasi Titik untuk Kurva Regresi Semiparametrik

Model persamaan (4.4) di atas dapat dijabarkan sebagai berikut.

$$\begin{aligned}
y_i = & \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \gamma_{11} t_{i1} + \gamma_{21} t_{i1}^2 + \dots + \gamma_{m1} t_{i1}^m + \\
& \gamma_{(m+1)1} (t_{i1} - K_{11})_+^m + \gamma_{(m+2)1} (t_{i1} - K_{21})_+^m + \dots + \\
& \gamma_{(m+r)1} (t_{i1} - K_{r1})_+^m + \dots + \gamma_{1q} t_{iq} + \gamma_{2q} t_{iq}^2 + \dots + \\
& \gamma_{mq} t_{iq}^m + \gamma_{(m+1)q} (t_{iq} - K_{1q})_+^m + \gamma_{(m+2)q} (t_{iq} - K_{2q})_+^m + \\
& \dots + \gamma_{(m+r)q} (t_{iq} - K_{rq})_+^m + \varepsilon_i
\end{aligned} \tag{4.5}$$

Persamaan (4.5) dapat pula disajikan dalam bentuk matirks berikut.

$$\tilde{y} = \mathbf{X}\tilde{\beta} + \mathbf{T}\tilde{\gamma} + \tilde{\varepsilon} \tag{4.6}$$

dengan

$$\tilde{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n \times 1} ; \mathbf{X} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{bmatrix}_{n \times (p+1)} ; \tilde{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}_{(p+1) \times 1} ; \tilde{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}_{n \times 1}$$

$$\mathbf{T} = \begin{bmatrix} t_{11} & \dots & t_{11}^m & (t_{11} - K_{11})_+^m & \dots & (t_{11} - K_{r1})_+^m & \dots & t_{1q} & \dots & t_{1q}^m & (t_{1q} - K_{1q})_+^m & \dots & (t_{1q} - K_{rq})_+^m \\ t_{21} & \dots & t_{21}^m & (t_{21} - K_{11})_+^m & \dots & (t_{21} - K_{r1})_+^m & \dots & t_{2q} & \dots & t_{2q}^m & (t_{2q} - K_{1q})_+^m & \dots & (t_{2q} - K_{rq})_+^m \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ t_{n1} & \dots & t_{n1}^m & (t_{n1} - K_{11})_+^m & \dots & (t_{n1} - K_{r1})_+^m & \dots & t_{nq} & \dots & t_{nq}^m & (t_{nq} - K_{1q})_+^m & \dots & (t_{nq} - K_{rq})_+^m \end{bmatrix}$$

Dengan ukuran $n \times (m+r)q$

$$\tilde{\gamma}' = [\gamma_{11} \dots \gamma_{m1} \gamma_{(m+1)1} \dots \gamma_{(m+r)1} \dots \gamma_{1q} \dots \gamma_{mq} \gamma_{(m+1)q} \dots \gamma_{(m+r)q}]_{1 \times (m+r)q}$$

Estimator kurva regresi semiparametrik mengikuti tahapan penyelesaian optimasi menggunakan metode *Ordinary Least Square* (OLS) sebagai berikut:

$$\tilde{\varepsilon} = \tilde{y} - \mathbf{X}\tilde{\beta} - \mathbf{T}\tilde{\gamma} \tag{4.7}$$

Selanjutnya jumlah kuadrat error diberikan oleh:

$$\begin{aligned}
\tilde{\varepsilon}'\tilde{\varepsilon} &= (\tilde{y} - \mathbf{X}\tilde{\beta} - \mathbf{T}\tilde{\gamma})'(\tilde{y} - \mathbf{X}\tilde{\beta} - \mathbf{T}\tilde{\gamma}) \\
&= (\tilde{y}' - \tilde{\beta}'\mathbf{X}' - \tilde{\gamma}'\mathbf{T}')(\tilde{y} - \mathbf{X}\tilde{\beta} - \mathbf{T}\tilde{\gamma}) \\
&= (\tilde{y}'\tilde{y} - \tilde{y}'\mathbf{X}\tilde{\beta} - \tilde{y}'\mathbf{T}\tilde{\gamma} - \tilde{\beta}'\mathbf{X}'\tilde{y} + \tilde{\beta}'\mathbf{X}'\mathbf{X}\tilde{\beta} + \tilde{\beta}'\mathbf{X}'\mathbf{T}\tilde{\gamma} - \tilde{\gamma}'\mathbf{T}'\tilde{y} + \\
&\quad \tilde{\gamma}'\mathbf{T}'\mathbf{X}\tilde{\beta} + \tilde{\gamma}'\mathbf{T}'\mathbf{T}\tilde{\gamma}) \\
&= (\tilde{y}'\tilde{y} - 2\tilde{\beta}'\mathbf{X}'\tilde{y} - 2\tilde{\gamma}'\mathbf{T}'\tilde{y} + 2\tilde{\beta}'\mathbf{X}'\mathbf{T}\tilde{\gamma} + \tilde{\beta}'\mathbf{X}'\mathbf{X}\tilde{\beta} + \\
&\quad \tilde{\gamma}'\mathbf{T}'\mathbf{T}\tilde{\gamma}) \\
&= Q(\tilde{\beta}, \tilde{\gamma})
\end{aligned} \tag{4.8}$$

Untuk mendapatkan estimasi dari parameter, dilakukan derivatif parsial. Derivatif parsial $Q(\tilde{\beta}, \tilde{\gamma})$ diberikan oleh:

$$\begin{aligned}\frac{\partial(Q(\tilde{\beta}, \tilde{\gamma}))}{\partial\tilde{\beta}} &= \frac{\partial(\tilde{\gamma}'\tilde{\gamma} - 2\tilde{\beta}'\mathbf{X}'\tilde{\gamma} - 2\tilde{\gamma}'\mathbf{T}'\tilde{\gamma} + 2\tilde{\beta}'\mathbf{X}'\mathbf{T}\tilde{\gamma} + \tilde{\beta}'\mathbf{X}'\mathbf{X}\tilde{\beta} + \tilde{\gamma}'\mathbf{T}'\mathbf{T}\tilde{\gamma})}{\partial\tilde{\beta}} \\ &= -2\mathbf{X}'\tilde{\gamma} + 2\mathbf{X}'\mathbf{T}\tilde{\gamma} + 2\mathbf{X}'\mathbf{X}\tilde{\beta}\end{aligned}\quad (4.9)$$

Jika derivatif parsial diatas disamakan dengan nol, diperoleh persamaan:

$$2\mathbf{X}'\mathbf{X}\tilde{\beta} = 2\mathbf{X}'\tilde{\gamma} - 2\mathbf{X}'\mathbf{T}\tilde{\gamma}$$

Sehingga estimator parameter $\tilde{\beta}$ adalah berikut:

$$\hat{\tilde{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\tilde{\gamma} - \mathbf{X}'\mathbf{T}\tilde{\gamma}) \quad (4.10)$$

Selanjutnya untuk mendapatkan estimator dari $\tilde{\gamma}$, kita lakukan derivatif parsial $Q(\tilde{\beta}, \tilde{\gamma})$ sebagai berikut.

$$\begin{aligned}\frac{\partial(Q(\tilde{\beta}, \tilde{\gamma}))}{\partial(\tilde{\gamma})} &= \frac{\partial(\tilde{\gamma}'\tilde{\gamma} - 2\tilde{\beta}'\mathbf{X}'\tilde{\gamma} - 2\tilde{\gamma}'\mathbf{T}'\tilde{\gamma} + 2\tilde{\beta}'\mathbf{X}'\mathbf{T}\tilde{\gamma} + \tilde{\beta}'\mathbf{X}'\mathbf{X}\tilde{\beta} + \tilde{\gamma}'\mathbf{T}'\mathbf{T}\tilde{\gamma})}{\partial(\tilde{\gamma})} \\ &= -2\mathbf{T}'\tilde{\gamma} + 2\mathbf{T}'\mathbf{X}\tilde{\beta} + 2\mathbf{T}'\mathbf{T}\tilde{\gamma}\end{aligned}\quad (4.11)$$

Jika derivatif parsial diatas disamakan dengan nol, diperoleh persamaan:

$$2\mathbf{T}'\mathbf{T}\tilde{\gamma} = 2\mathbf{T}'\tilde{\gamma} - 2\mathbf{T}'\mathbf{X}\tilde{\beta}$$

Sehingga estimator parameter $\tilde{\gamma}$ adalah sbagai berikut.

$$\hat{\tilde{\gamma}} = (\mathbf{T}'\mathbf{T})^{-1}(\mathbf{T}'\tilde{\gamma} - \mathbf{T}'\mathbf{X}\tilde{\beta}) \quad (4.12)$$

Karena estimator parameter $\tilde{\beta}$ masih mengandung $\tilde{\gamma}$, dan estimator $\tilde{\gamma}$ masih mengandung $\tilde{\beta}$ maka dilakukan substitusi antara persamaan (4.10) dan persamaan (4.12) sebagai berikut.

$$\begin{aligned}\hat{\tilde{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1} \left(\mathbf{X}'\tilde{\gamma} - \mathbf{X}'\mathbf{T} \left((\mathbf{T}'\mathbf{T})^{-1}(\mathbf{T}'\tilde{\gamma} - \mathbf{T}'\mathbf{X}\tilde{\beta}) \right) \right) \\ &= (\mathbf{X}'\mathbf{X})^{-1} \left(\mathbf{X}'\tilde{\gamma} - \mathbf{X}'\mathbf{T}((\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\tilde{\gamma} - (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}\tilde{\beta}) \right) \\ &= (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\tilde{\gamma} - \mathbf{X}'(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\tilde{\gamma} + \mathbf{X}'\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}\tilde{\beta}) \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\tilde{\gamma} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\tilde{\gamma} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}\tilde{\beta}\end{aligned}$$

Kemudian komponen yang mengandung $\hat{\tilde{\beta}}$ dijadikan satu sebagai berikut.

$$\hat{\tilde{\beta}} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}\tilde{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\tilde{\gamma} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\tilde{\gamma}$$

$$\hat{\tilde{\beta}}(\mathbf{I} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\tilde{\gamma} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\tilde{\gamma}$$

Sehingga estimator $\hat{\tilde{\beta}}$ diperoleh persamaan sebagai berikut.

$$\hat{\tilde{\beta}} = [\mathbf{I} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}]^{-1}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\tilde{\gamma} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\tilde{\gamma}]$$

$$\begin{aligned}
&= [\mathbf{I} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}]^{-1}(\mathbf{X}'\mathbf{X})^{-1}[\mathbf{X}'\tilde{\mathbf{y}} - \mathbf{X}'\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\tilde{\mathbf{y}}] \\
&= [\mathbf{I} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}]^{-1}(\mathbf{X}'\mathbf{X})^{-1}[\mathbf{X}' - \mathbf{X}'\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}']\tilde{\mathbf{y}} \\
&= \mathbf{A}(\tilde{k})\tilde{\mathbf{y}}
\end{aligned}$$

Dimana \mathbf{I} adalah matrik identitas, \tilde{k} adalah titik knot dari $1, 2, \dots, r$ dan matrik $\mathbf{A}(\tilde{k}) = [\mathbf{I} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}]^{-1}(\mathbf{X}'\mathbf{X})^{-1}[\mathbf{X}' - \mathbf{X}'\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}']$.

Begitu juga untuk estimator parameter $\tilde{\gamma}$, sehingga diperoleh :

$$\begin{aligned}
\hat{\tilde{\gamma}} &= (\mathbf{T}'\mathbf{T})^{-1}(\mathbf{T}'\tilde{\mathbf{y}} - \mathbf{T}'\mathbf{X}\tilde{\beta}) \\
&= (\mathbf{T}'\mathbf{T})^{-1}\left(\mathbf{T}'\tilde{\mathbf{y}} - \mathbf{T}'\mathbf{X}\left((\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\tilde{\mathbf{y}} - \mathbf{X}'\mathbf{T}\tilde{\gamma})\right)\right) \\
&= (\mathbf{T}'\mathbf{T})^{-1}\left(\mathbf{T}'\tilde{\mathbf{y}} - \mathbf{T}'\mathbf{X}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\tilde{\mathbf{y}} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}\tilde{\gamma})\right) \\
&= (\mathbf{T}'\mathbf{T})^{-1}(\mathbf{T}'\tilde{\mathbf{y}} - \mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\tilde{\mathbf{y}} + \mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}\tilde{\gamma}) \\
&= (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\tilde{\mathbf{y}} - (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\tilde{\mathbf{y}} + (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}\tilde{\gamma}
\end{aligned}$$

komponen yang mengandung $\hat{\tilde{\gamma}}$ dijadikan satu diruas kiri sebagai berikut:

$$\begin{aligned}
\hat{\tilde{\gamma}} - (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}\tilde{\gamma} &= (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\tilde{\mathbf{y}} - (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\tilde{\mathbf{y}} \\
\hat{\tilde{\gamma}}(\mathbf{I} - (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}) &= (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\tilde{\mathbf{y}} - (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\tilde{\mathbf{y}}
\end{aligned}$$

Sehingga estimator $\hat{\tilde{\gamma}}$ diperoleh persamaan sebagai berikut:

$$\begin{aligned}
\hat{\tilde{\gamma}} &= [\mathbf{I} - (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}]^{-1}[(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\tilde{\mathbf{y}} - (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\tilde{\mathbf{y}}] \\
&= [\mathbf{I} - (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}]^{-1}(\mathbf{T}'\mathbf{T})^{-1}[\mathbf{T}'\tilde{\mathbf{y}} - \mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\tilde{\mathbf{y}}] \\
&= [\mathbf{I} - (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}]^{-1}(\mathbf{T}'\mathbf{T})^{-1}[\mathbf{T}' - \mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\tilde{\mathbf{y}} \\
&= \mathbf{B}(\tilde{k})\tilde{\mathbf{y}} \tag{4.13}
\end{aligned}$$

Matrik \mathbf{I} adalah matrik identitas, \tilde{k} adalah titik knot yang bernilai $1, 2, \dots, r$ dan matrik $\mathbf{B}(\tilde{k}) = [\mathbf{I} - (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}]^{-1}(\mathbf{T}'\mathbf{T})^{-1}[\mathbf{T}' - \mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']$.

Setelah diperoleh estimator untuk komponen parametrik dan nonparametrik, selanjutnya menentukan estimator model regresi semiparametrik spline *truncated*.

$$\begin{aligned}
\hat{\mathbf{y}} &= \hat{f}(\mathbf{X}, \mathbf{T}) \\
&= \mathbf{X}\hat{\beta} + \mathbf{T}\hat{\tilde{\gamma}} \\
&= \mathbf{X}\mathbf{A}(\tilde{k})\tilde{\mathbf{y}} + \mathbf{T}\mathbf{B}(\tilde{k})\tilde{\mathbf{y}} \\
&= (\mathbf{X}\mathbf{A}(\tilde{k}) + \mathbf{T}\mathbf{B}(\tilde{k}))\tilde{\mathbf{y}}
\end{aligned}$$

$$= \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y} \quad (4.14)$$

dimana $\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) = \mathbf{XA}(\tilde{k}) + \mathbf{TB}(\tilde{k})$

4.1.2 Perumusan Uji Hipotesis

Persamaan (4.4) dalam bentuk matrik dapat pula dituliskan sebagai berikut:

$$\tilde{y} = \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\delta} + \tilde{\varepsilon} \quad (4.15)$$

dengan

$$\tilde{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n \times 1}; \quad \mathbf{Z} = [\mathbf{X} \quad \mathbf{T}]_{n \times (p+1+(m+r)q)}; \quad \tilde{\delta} = \begin{bmatrix} \tilde{\beta} \\ \dots \\ \tilde{\gamma} \end{bmatrix}_{((p+1)+(m+r)q) \times 1}$$

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{bmatrix}$$

$$\mathbf{T} = \begin{bmatrix} t_{11} & \dots & t_{11}^m & (t_{11} - K_{11})_+^m & \dots & (t_{11} - K_{r1})_+^m & \dots & t_{1q} & \dots & t_{1q}^m & (t_{1q} - K_{1q})_+^m & \dots & (t_{1q} - K_{rq})_+^m \\ t_{21} & \dots & t_{21}^m & (t_{21} - K_{11})_+^m & \dots & (t_{21} - K_{r1})_+^m & \dots & t_{2q} & \dots & t_{2q}^m & (t_{2q} - K_{1q})_+^m & \dots & (t_{2q} - K_{rq})_+^m \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ t_{n1} & \dots & t_{n1}^m & (t_{n1} - K_{11})_+^m & \dots & (t_{n1} - K_{r1})_+^m & \dots & t_{nq} & \dots & t_{nq}^m & (t_{nq} - K_{1q})_+^m & \dots & (t_{nq} - K_{rq})_+^m \end{bmatrix}$$

$$\tilde{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}; \quad \tilde{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}_{n \times 1}$$

$$\tilde{\gamma}' = [\gamma_{11} \dots \gamma_{m1} \gamma_{(m+1)1} \dots \gamma_{(m+r)1} \dots \gamma_{1q} \dots \gamma_{mq} \gamma_{(m+1)q} \dots \gamma_{(m+r)q}]_{1 \times (m+r)q}$$

Respon \tilde{y} merupakan vektor berukuran $n \times 1$, matriks $\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) = (\mathbf{X} : \mathbf{T})$, \mathbf{X} adalah matriks yang memuat prediktor komponen parametrik yang berukuran $n \times (p + 1)$ dan \mathbf{T} adalah matriks yang memuat komponen nonparametrik berukuran $n \times (m + r)q$ yang tergantung pada titik knot \tilde{k} yang diberikan yaitu $1, 2, \dots, r$.

$$\tilde{\delta}' = (\beta_0, \dots, \beta_p, \gamma_{11}, \dots, \gamma_{m1}, \gamma_{(m+1)1}, \dots, \gamma_{(m+r)1}, \dots, \gamma_{1q}, \dots, \gamma_{mq}, \gamma_{(m+1)q}, \dots, \gamma_{(m+r)q})$$

merupakan vektor parameter dengan ukuran $((p + 1) + (m + r)q) \times 1$ dan $\tilde{\varepsilon}$ adalah vektor *error*. Karena *error* $\tilde{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$ maka:

$$i. \quad \tilde{y} \sim N(\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\delta}, \sigma^2 \mathbf{I}) \quad (4.16)$$

$$\begin{aligned}
E(y) &= E(\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}) + E(\varepsilon) \\
&= E(\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}) + E(\varepsilon) \\
&= \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta} \\
Var(y) &= Var(\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}) + Var(\varepsilon) \\
&= Var(\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}) + Var(\varepsilon) \\
&= \sigma^2\mathbf{I}
\end{aligned}$$

Karena y adalah fungsi linier dalam ε , maka dapat dituliskan

$$\begin{aligned}
\tilde{y} &\sim N(\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}, \sigma^2\mathbf{I}) \\
\text{ii. } \hat{\tilde{\delta}} &\sim N\left(\tilde{\delta}, \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \sigma^2\mathbf{I}\right). \quad (4.17)
\end{aligned}$$

$$\tilde{y} = \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta} + \tilde{\varepsilon}$$

Diperoleh persamaan:

$$\tilde{\varepsilon} = y - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}$$

Selanjutnya jumlah kuadrat error diberikan oleh:

$$\begin{aligned}
\tilde{\varepsilon}'\tilde{\varepsilon} &= (\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta})'(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}) \\
&= (\tilde{y}' - \tilde{\delta}'\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}))(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}) \\
&= (\tilde{y}'\tilde{y} - \tilde{y}'\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta} - \tilde{\delta}'\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y} + \tilde{\delta}'\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}) \\
&= (\tilde{y}'\tilde{y} - 2\tilde{\delta}'\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y} + \tilde{\delta}'\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}) \quad (4.18)
\end{aligned}$$

$$= Q(\tilde{\delta}) \quad (4.19)$$

Untuk mendapatkan estimasi parameter dilakukan derivatif parsial. Derivatif parsial $Q(\tilde{\delta})$ diberikan oleh:

$$\begin{aligned}
\frac{\partial(Q(\tilde{\delta}))}{\partial\tilde{\delta}} &= \frac{\partial(\tilde{y}'\tilde{y} - 2\tilde{\delta}'\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y} + \tilde{\delta}'\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta})}{\partial\tilde{\delta}} \\
&= -2\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y} + 2\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\tilde{\delta}}
\end{aligned}$$

Jika derivatif parsial diatas disamakan dengan nol, diperoleh persamaan:

$$2\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\tilde{\delta}} = 2\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y}$$

sehingga akan diperoleh estimator parameter adalah berikut:

$$\hat{\tilde{\delta}} = \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y} \quad (4.20)$$

Sehingga diperoleh

$$\begin{aligned}
E(\hat{\delta}) &= E\left(\left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y}\right) \\
&= \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) E(\tilde{y}) \\
&= \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta} \\
&= \tilde{\delta} \\
Var(\hat{\delta}) &= Var\left(\left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y}\right) \\
&= \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) Var(\tilde{y}) \\
&\quad \left(\left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\right)' \\
&= \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \\
&\quad \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \sigma^2 \mathbf{I} \\
&= \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \sigma^2 \mathbf{I}
\end{aligned}$$

Karena $\hat{\delta}$ adalah fungsi linier dalam y , maka

$$\hat{\delta} \sim N\left(\tilde{\delta}, \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \sigma^2 \mathbf{I}\right)$$

iii. $\tilde{c}_j' \hat{\delta} \sim N\left(\tilde{c}_j' \tilde{\delta}, \tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \tilde{c}_j \sigma^2\right)$

dari persamaan

$$\begin{aligned}
\hat{\delta} &= \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y} \\
E(\tilde{c}_j' \hat{\delta}) &= E\left(\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y}\right) \\
&= E\left(\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\right) E(\tilde{y}) \\
&= \tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta} \\
&= \tilde{c}_j' \tilde{\delta}
\end{aligned}$$

$$\begin{aligned}
Var(\tilde{c}_j' \hat{\delta}) &= Var\left(\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y}\right) \\
&= \tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) Var(\tilde{y})
\end{aligned}$$

$$\begin{aligned}
& \left(\tilde{c}'_j \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)' \\
&= \tilde{c}'_j \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \\
& \quad \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \sigma^2 \mathbf{I} \\
&= \tilde{c}'_j \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \sigma^2
\end{aligned}$$

Karena $\tilde{c}'_j \hat{\delta}$ adalah fungsi linier dalam y , maka

$$\tilde{c}'_j \hat{\delta} \sim N \left(\tilde{c}'_j \tilde{\delta}, \tilde{c}'_j \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \sigma^2 \right)$$

Perumusan hipotesis yang digunakan untuk menguji signifikansi parameter model regresi semiparametrik spline *truncated* secara parsial adalah sebagai berikut:

$$H_0 : \tilde{c}'_j \tilde{\delta} = \tau \quad (4.21)$$

$$H_1 : \tilde{c}'_j \tilde{\delta} \neq \tau$$

Dimana \tilde{c}'_j adalah suatu vektor yang memiliki elemen bernilai nol kecuali pada elemen ke- h bernilai 1. $\tilde{c}'_j = (0 \ 0 \ \dots \ 1 \ \dots \ 0)$ berukuran $1 \times (p + 1 + (m + r)q)$ dan $\tilde{\delta}' = (\beta_0, \dots, \beta_p, \gamma_{11}, \dots, \gamma_{m1}, \gamma_{(m+1)1}, \dots, \gamma_{(m+r)1}, \dots, \gamma_{1q}, \dots, \gamma_{mq}, \gamma_{(m+1)q}, \dots, \gamma_{(m+r)q})$ sehingga $\tilde{\delta}$ berukuran $(p + 1 + (m + r)q) \times 1$, dan $\tau = 0$.

4.1.3 Menentukan Statistik Uji Hipotesis

Selanjutnya untuk mendapatkan statistik uji dari hipotesis persamaan (4.21) diselesaikan dengan menggunakan metode Likelihood Ratio Test (LRT). LRT digunakan untuk mencari rasio fungsi likelihood dari nilai maksimum dibawah ruang parameter ω terhadap nilai maksimum dibawah ruang parameter Ω . Berdasarkan ruang parameter pada persamaan (4.22) dan (4.23) maka dapat dicari fungsi likelihood dibawah Ω dan dibawah ω dengan langkah-langkah sebagai berikut.

a. Menentukan ruang parameter dibawah $H(\Omega)$ yaitu sebagai berikut.

$$\Omega = \left\{ \delta = \left(\beta_0, \dots, \beta_p, \gamma_{11}, \dots, \gamma_{(r+m)1}, \dots, \gamma_{1q}, \dots, \gamma_{(r+m)q} \right), \sigma_\Omega^2 \right\} \quad (4.22)$$

b. Menentukan ruang parameter dibawah $H_0(\omega)$ yaitu sebagai berikut.

$$\omega = \left\{ \delta = (\beta_0, \dots, \beta_{h-1}, \beta_{h+1}, \dots, \beta_p, \gamma_{11}, \dots, \gamma_{(r+m)1}, \gamma_{1q}, \dots, \gamma_{(r+m)q}), \sigma_\omega^2 | c_j' \tilde{\delta} = 0 \right\} \quad (4.23)$$

c. Menentukan fungsi likelihood di bawah $H(\Omega)$.

Karena regresi semiparametrik spline *truncated* seperti pada persamaan (4.4), $\tilde{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$ maka $\tilde{y} \sim N(\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\delta}_\Omega, \sigma_\Omega^2 \mathbf{I})$ sehingga fungsi likelihood dibawah ruang $H(\Omega)$ adalah sebagai berikut.

$$\begin{aligned} L(\tilde{\delta}_\Omega, \sigma_\Omega^2) &= \prod_{i=1}^n \left(\frac{1}{\sqrt{2\pi\sigma_\Omega^2}} \exp \left(-\frac{1}{2\sigma_\Omega^2} \left(\tilde{y} - \sum_{h=0}^p \beta_h x_{hi} + \sum_{l=1}^q \left(\sum_{j=1}^m \gamma_{jl} t_{li}^j + \sum_{k=1}^r \gamma_{(m+k)l} (t_{li} - K_{kl})_+^m \right) \right)^2 \right) \right) \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} \exp \left(-\frac{1}{2\sigma_\Omega^2} [(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\delta}_\Omega)' (\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\delta}_\Omega)] \right) \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} \exp \left(-\frac{1}{2\sigma_\Omega^2} [\tilde{y}' \tilde{y} - 2\tilde{\delta}_\Omega' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y} + \tilde{\delta}_\Omega' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\delta}_\Omega] \right) \end{aligned} \quad (4.24)$$

d. Menentukan fungsi likelihood di bawah ruang $H_0(\omega)$.

Karena regresi semiparametrik spline *truncated* seperti pada persamaan (4.4) $\tilde{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$ maka $\tilde{y} \sim N(\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\delta}_\omega, \sigma_\omega^2 \mathbf{I})$ sehingga fungsi likelihood dibawah ruang ω adalah sebagai berikut.

$$\begin{aligned} L(\tilde{\delta}_\omega, \sigma_\omega^2) &= \prod_{i=1}^n \left(\frac{1}{\sqrt{2\pi\sigma_\omega^2}} \exp \left(-\frac{1}{2\sigma_\omega^2} \left(\tilde{y} - \sum_{h=0}^p \beta_h x_{hi} + \sum_{l=1}^q \left(\sum_{j=1}^m \gamma_{jl} t_{li}^j + \sum_{k=1}^r \gamma_{(m+k)l} (t_{li} - K_{kl})_+^m \right) \right)^2 \right) \right) \end{aligned}$$

$$\begin{aligned}
&= (2\pi\sigma_\omega^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma_\omega^2} [(\tilde{\mathbf{y}} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\omega)'(\tilde{\mathbf{y}} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\omega)]\right) \\
&= (2\pi\sigma_\omega^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma_\omega^2} [\tilde{\mathbf{y}}'\tilde{\mathbf{y}} - 2\tilde{\delta}'_\omega \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\mathbf{y}} + \right. \\
&\quad \left. \tilde{\delta}'_\omega \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\omega]\right)
\end{aligned} \tag{4.25}$$

$$\begin{aligned}
\log L(\tilde{\delta}_\omega, \sigma_\omega^2) &= -\frac{n}{2} \log(2\pi\sigma_\omega^2) - \frac{1}{2\sigma_\omega^2} [\tilde{\mathbf{y}}'\tilde{\mathbf{y}} - 2\tilde{\delta}'_\omega \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\mathbf{y}} + \\
&\quad \tilde{\delta}'_\omega \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\omega]
\end{aligned} \tag{4.26}$$

e. Mencari penaksir parameter di bawah $H(\Omega)$.

Untuk mendapatkan estimasi parameter δ di bawah populasi atau di bawah ruang $H(\Omega)$ digunakan metode MLE(*Maximum Likelihood Estimation*) yaitu diawali dengan mencari nilai logaritma dari persamaan (4.24).

$$\begin{aligned}
\log L(\tilde{\delta}_\Omega, \sigma_\Omega^2) &= -\frac{n}{2} \log(2\pi\sigma_\Omega^2) - \frac{1}{2\sigma_\Omega^2} [\tilde{\mathbf{y}}'\tilde{\mathbf{y}} - 2\tilde{\delta}'_\Omega \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\mathbf{y}} + \\
&\quad \tilde{\delta}'_\Omega \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\Omega]
\end{aligned} \tag{4.27}$$

Selanjutnya turunan parsial persamaan (4.27) terhadap parameter $\tilde{\delta}_\Omega$ didapatkan

$$\begin{aligned}
\frac{\partial(\log L(\tilde{\delta}_\Omega, \sigma_\Omega^2))}{\partial \tilde{\delta}_\Omega} &= \frac{\partial}{\partial \tilde{\delta}_\Omega} \left(-\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma_\Omega^2) - \right. \\
&\quad \left. \frac{1}{2\sigma_\Omega^2} (\tilde{\mathbf{y}}'\tilde{\mathbf{y}} - 2\tilde{\delta}'_\Omega \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\mathbf{y}} + \right. \\
&\quad \left. \tilde{\delta}'_\Omega \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\Omega) \right)
\end{aligned} \tag{4.28}$$

$$= 0 - \frac{1}{2\sigma_\Omega^2} (0 - 2\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\mathbf{y}} + 2\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\Omega) \tag{4.29}$$

Jika derivatif parsial diatas disamakan dengan nol, diperoleh persamaan:

$$\mathbf{0} = -2\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\mathbf{y}} + 2\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\Omega \tag{4.30}$$

$$2\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\Omega = 2\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\mathbf{y}} \tag{4.31}$$

sehingga akan diperoleh estimator parameter dibawah ruang Ω adalah berikut:

$$\hat{\tilde{\delta}}_\Omega = \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\mathbf{y}} \tag{4.32}$$

Selanjutnya untuk mendapatkan estimasi parameter σ_Ω^2 di bawah H dilakukan dengan mencari turunan parsial dari persamaan (4.27) terhadap σ_Ω^2 .

$$\frac{\partial(\log L(\tilde{\delta}_\Omega, \sigma_\Omega^2))}{\partial \sigma_\Omega^2} = -\frac{n}{2} \frac{2\pi}{2\pi\sigma_\Omega^2} + \frac{1}{2\sigma_\Omega^4} \left((\tilde{\mathbf{y}} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\Omega)' \right) \tag{4.33}$$

$$(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\delta_\Omega)$$

Jika persamaan (4.33) disamakan dengan nol, maka diperoleh persamaan:

$$\begin{aligned} \frac{n}{\sigma_\Omega^2} &= \frac{1}{\sigma_\Omega^4} \left[(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\delta_\Omega)' (\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\delta_\Omega) \right] \\ \hat{\sigma}_\Omega^2 &= \frac{(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\delta_\Omega)' (\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\delta_\Omega)}{n} \end{aligned} \quad (4.34)$$

Sehingga nilai maksimum dari fungsi likelihood pada persamaan (4.24) diberikan oleh:

$$\begin{aligned} \max_{\Omega} L(\tilde{\delta}_\Omega, \sigma_\Omega^2) &= L(\hat{\Omega}) = L(\hat{\tilde{\delta}}_\Omega, \hat{\sigma}_\Omega^2) \\ L(\hat{\Omega}) &= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma_\Omega^2} (\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\Omega)'\right. \\ &\quad \left. (\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\Omega)\right) \end{aligned} \quad (4.35)$$

Dengan mensubstitusikan persamaan (4.34) ke persamaan (4.35) maka akan menjadi:

$$\begin{aligned} L(\hat{\Omega}) &= (2\pi\hat{\sigma}_\Omega^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \frac{n (\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\tilde{\delta}}_\Omega)' (\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\tilde{\delta}}_\Omega)}{(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\tilde{\delta}}_\Omega)' (\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\tilde{\delta}}_\Omega)}\right) \\ &= (2\pi\hat{\sigma}_\Omega^2)^{-\frac{n}{2}} \exp\left(-\frac{n}{2}\right) \\ &= (2\pi\hat{\sigma}_\Omega^2)^{-\frac{n}{2}} e^{-\frac{n}{2}} \end{aligned} \quad (4.36)$$

f. Menentukan estimasi parameter di bawah $H_0(\omega)$.

Untuk mendapatkan estimasi parameter dibawah $H_0(\omega)$ yaitu $\hat{\tilde{\delta}}_\omega$ dengan menggunakan metode Fungsi *Multiplier Lagrange* (LM). Diberikan fungsi LM yaitu

$$F(\tilde{\delta}_\omega, \theta) = V(\tilde{\delta}_\omega) + 2\theta(\tilde{c}'_j \tilde{\delta}_\omega) \quad (4.37)$$

dimana

$$V(\tilde{\delta}_\omega) = (\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\omega)' (\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\omega) \quad (4.38)$$

dengan konstrain $\tilde{c}'_j \tilde{\delta}_\omega = 0$.

Persamaan (4.38) dapat dijabarkan sebagai berikut

$$\begin{aligned}
V(\tilde{\delta}_\omega) &= \tilde{y}'\tilde{y} - y'\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\omega - \tilde{\delta}_\omega'\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y} + \\
&\quad \tilde{\delta}_\omega'\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\omega \\
&= \tilde{y}'\tilde{y} - 2\tilde{\delta}_\omega'\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y} + \tilde{\delta}_\omega'\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\omega \quad (4.39)
\end{aligned}$$

Dengan mensubstitusi persamaan (4.37) ke persamaan (4.39) maka akan menjadi sebagai berikut:

$$\begin{aligned}
F(\tilde{\delta}_\omega, \theta) &= \tilde{y}'\tilde{y} - 2\tilde{\delta}_\omega'\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y} + \tilde{\delta}_\omega'\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \\
&\quad \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\omega + 2\theta(\tilde{c}'_j\tilde{\delta}_\omega) \\
&= \tilde{y}'\tilde{y} - 2\tilde{\delta}_\omega'\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y} + \tilde{\delta}_\omega'\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \\
&\quad \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\omega + 2\theta(\tilde{\delta}_\omega'\tilde{c}'_j) \quad (4.40)
\end{aligned}$$

Jika persamaan (4.40) diturunkan terhadap $\tilde{\delta}_\omega$ maka akan menghasilkan persamaan:

$$\frac{\partial (F(\tilde{\delta}_\omega, \theta))}{\partial (\tilde{\delta}_\omega)} = -2\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y} + 2\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\omega + 2\theta\tilde{c}_j \quad (4.41)$$

Jika persamaan (4.41) disamakan dengan nol, maka persamaan akan menjadi:

$$\begin{aligned}
-2\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y} + 2\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\omega + 2\theta\tilde{c}_j &= 0 \\
2\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{\delta}_\omega &= 2\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y} - 2\theta\tilde{c}_j
\end{aligned}$$

Sehingga akan didapatkan $\hat{\tilde{\delta}}_\omega$ sebagai berikut:

$$\begin{aligned}
\hat{\tilde{\delta}}_\omega &= (\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}))^{-1} (\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y} - \theta\tilde{c}_j) \\
&= (\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}))^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\tilde{y} - (\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}))^{-1} \theta\tilde{c}_j \\
&= \hat{\tilde{\delta}}_\Omega - (\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}))^{-1} \tilde{c}_j\theta \quad (4.42)
\end{aligned}$$

Jika persamaan (4.40) diturunkan terhadap θ maka akan menghasilkan persamaan:

$$\frac{\partial (F(\tilde{\delta}_\omega, \theta))}{\partial \theta} = 2\tilde{c}'_j\tilde{\delta}_\omega$$

Jika persamaan di atas disamakan dengan nol, maka persamaan akan menjadi:

$$\tilde{c}'_j\hat{\tilde{\delta}}_\omega = 0 \quad (4.43)$$

Dengan menstutstitusi persamaan (4.42) dan (4.43) akan mendapatkan θ , yaitu

$$\begin{aligned}
\tilde{c}'_j \left(\hat{\delta}_\Omega - \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \theta \tilde{c}_j \right) &= 0 \\
\tilde{c}'_j \hat{\delta}_\Omega - \tilde{c}'_j \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \theta &= 0 \\
\tilde{c}'_j \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \theta &= \tilde{c}'_j \hat{\delta}_\Omega \\
\theta &= \left(\tilde{c}'_j \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right)^{-1} \tilde{c}'_j \hat{\delta}_\Omega
\end{aligned} \tag{4.44}$$

Sehingga persamaan (4.42) akan menjadi

$$\begin{aligned}
\hat{\delta}_\omega &= \hat{\delta}_\Omega - \left(\mathbf{Z}'(\mathbf{X}, \mathbf{Z}, \tilde{k}) \mathbf{T}(\mathbf{X}, \mathbf{Z}, \tilde{k}) \right)^{-1} \\
&\quad \tilde{c}'_j \left(\tilde{c}'_j \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right)^{-1} \tilde{c}'_j \hat{\delta}_\Omega
\end{aligned} \tag{4.45}$$

$$\begin{aligned}
\hat{\delta}_\Omega - \hat{\delta}_\omega &= \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \\
&\quad \tilde{c}'_j \left(\tilde{c}'_j \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right)^{-1} \tilde{c}'_j \hat{\delta}_\Omega
\end{aligned} \tag{4.46}$$

Dengan menurunkan parsial pada persamaan (4.26) terhadap σ_ω^2 maka didapat persamaan berikut:

$$\begin{aligned}
\frac{\partial \left(\log L(\tilde{\delta}_\omega, \sigma_\omega^2) \right)}{\partial \sigma_\omega^2} &= -\frac{n}{2} \frac{2\pi}{2\pi\sigma_\omega^2} - \frac{1}{2\sigma_\omega^4} \left(\left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\delta}_\omega \right)' \right. \\
&\quad \left. \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\delta}_\omega \right) \right)
\end{aligned} \tag{4.47}$$

Jika persamaam (4.47) disamakan dengan nol, maka persamaan diberikan oleh:

$$\begin{aligned}
-\frac{n}{2} \frac{2\pi}{2\pi\sigma_\omega^2} - \frac{1}{2\sigma_\omega^4} \left[\left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\omega \right) \right] &= 0 \\
\frac{n}{\sigma_\omega^2} &= \frac{1}{\sigma_\omega^4} \left[\left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\omega \right) \right] \\
\hat{\sigma}_\omega^2 &= \frac{\left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\omega \right)}{n}
\end{aligned} \tag{4.48}$$

Sehingga nilai maksimum dari fungsi likelihood pada persamaan (4.25) diberikan oleh:

$$\begin{aligned}
\max_\omega L(\tilde{\delta}_\omega, \sigma_\omega^2) &= L(\hat{\omega}) = (2\pi\hat{\sigma}_\omega^2)^{-\frac{n}{2}} \\
&\quad \exp \left(-\frac{1}{2\hat{\sigma}_\omega^2} \left(\left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\omega \right) \right) \right)
\end{aligned} \tag{4.49}$$

Dengan mensubstitusikan persamaan (4.48) ke persamaan (4.49) maka akan diperoleh:

$$\begin{aligned}
L(\hat{\omega}) &= (2\pi\hat{\sigma}_{\omega}^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \frac{n(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\omega})'(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\omega})}{(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\omega})'(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\omega})}\right) \\
&= (2\pi\hat{\sigma}_{\omega}^2)^{-\frac{n}{2}} \exp\left(-\frac{n}{2}\right) \\
&= (2\pi\hat{\sigma}_{\omega}^2)^{-\frac{n}{2}} e^{-\frac{n}{2}} \tag{4.50}
\end{aligned}$$

Senjutnya menghitung ratio likelihood dengan cara membagi persamaan (4.36) dan persamaan (4.50) seperti berikut:

$$\begin{aligned}
L_{ratio} = \lambda(y) &= \frac{L(\hat{\omega})}{L(\hat{\Omega})} = \frac{(2\pi\hat{\sigma}_{\omega}^2)^{-\frac{n}{2}} e^{-\frac{n}{2}}}{(2\pi\hat{\sigma}_{\Omega}^2)^{-\frac{n}{2}} e^{-\frac{n}{2}}} \\
&= \left(\frac{\hat{\sigma}_{\omega}^2}{\hat{\sigma}_{\Omega}^2}\right)^{-\frac{n}{2}} \tag{4.51}
\end{aligned}$$

Dengan mensubstitusikan persamaan (4.34) dan persamaan (4.48) maka persamaan (4.51) akan menjadi:

$$\begin{aligned}
\lambda(y) &= \left(\frac{\left(\frac{(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\omega})'(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\omega})}{n}\right)^{-\frac{n}{2}}}{\left(\frac{(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\Omega})'(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\Omega})}{n}\right)^{-\frac{n}{2}}}\right)^{-\frac{n}{2}} \\
&= \left(\frac{\left(\frac{(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\omega})'(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\omega})}{(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\Omega})'(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\Omega})}\right)^{-\frac{n}{2}}}{\left(\frac{(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\Omega})'(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\Omega})}{(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\omega})'(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\omega})}\right)^{-\frac{n}{2}}}\right)^{-\frac{n}{2}} \tag{4.52}
\end{aligned}$$

$$\begin{aligned}
&= \left(\frac{\left(\frac{(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\Omega})'(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\Omega})}{(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\omega})'(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\omega})}\right)^{\frac{n}{2}}}{\left(\frac{(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\omega})'(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\omega})}{(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\Omega})'(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\Omega})}\right)^{\frac{n}{2}}}\right)^{\frac{n}{2}} \\
&= \left(\frac{Q_2}{D}\right)^{\frac{n}{2}} \tag{4.53}
\end{aligned}$$

Dengan $Q_2 = (\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\Omega})'(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\Omega})$ dan

$$D = (\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\omega})'(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_{\omega})$$

Selanjutnya persamaan D akan disederhanakan dengan cara mensubstitusikan $(-\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega + \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega)$ pada persamaan D sehingga akan didapat :

$$\begin{aligned}
D &= \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega + \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\omega \right)' \\
&\quad \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega + \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\omega \right) \\
&= \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega + \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\hat{\delta}_\Omega - \hat{\delta}_\omega \right) \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega + \right. \\
&\quad \left. \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\hat{\delta}_\Omega - \hat{\delta}_\omega \right) \right) \\
&= \left(\left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega \right)' + \left(\hat{\delta}_\Omega - \hat{\delta}_\omega \right)' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \right) \left(\left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega \right) + \right. \\
&\quad \left. \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\hat{\delta}_\Omega - \hat{\delta}_\omega \right) \right) \\
&= \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega \right) + \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega \right)' \\
&\quad \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\hat{\delta}_\Omega - \hat{\delta}_\omega \right) + \left(\hat{\delta}_\Omega - \hat{\delta}_\omega \right)' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega \right) + \\
&\quad \left(\hat{\delta}_\Omega - \hat{\delta}_\omega \right)' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\hat{\delta}_\Omega - \hat{\delta}_\omega \right) \\
&= \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega \right) + M \left(\hat{\delta}_\Omega - \hat{\delta}_\omega \right) + \\
&\quad \left(\hat{\delta}_\Omega - \hat{\delta}_\omega \right)' N + \left(\hat{\delta}_\Omega - \hat{\delta}_\omega \right)' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\hat{\delta}_\Omega - \hat{\delta}_\omega \right) \tag{4.54}
\end{aligned}$$

Akan dijabarkan ruas kedua dan ketiga yaitu M dan N pada persamaan (4.54).

$$M = \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega \right)' \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \quad \text{dan} \quad N = \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})\hat{\delta}_\Omega \right),$$

dengan cara mensubstitusikan nilai $\hat{\delta}_\Omega$ sebagai berikut:

$$\begin{aligned}
M &= \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y} \right)' \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \\
&= \tilde{y}' \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) - \tilde{y}' \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \\
&\quad \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \\
&= \tilde{y}' \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) - \tilde{y}' \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \\
&= 0
\end{aligned}$$

Dan untuk N sebagai berikut:

$$\begin{aligned}
N &= \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y} \right) \\
&= \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y} - \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})
\end{aligned}$$

$$\begin{aligned}
& \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y} \\
&= \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y} - \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y} \\
&= 0
\end{aligned}$$

Sehingga persamaan (4.54) akan menjadi:

$$\begin{aligned}
D &= \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right) + \\
&\quad \left(\hat{\delta}_\Omega - \hat{\delta}_\omega \right)' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\hat{\delta}_\Omega - \hat{\delta}_\omega \right)
\end{aligned}$$

jika

$$\begin{aligned}
B &= \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right) \\
C &= \left(\hat{\delta}_\Omega - \hat{\delta}_\omega \right)' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\hat{\delta}_\Omega - \hat{\delta}_\omega \right)
\end{aligned}$$

maka persamaan di atas dapat disederhanakan menjadi:

$$D = B + C$$

Selanjutnya akan dijabarkan persamaan C dengan mensubstitusikan persamaan (4.45) sebagai berikut:

$$\begin{aligned}
\left(\hat{\delta}_\Omega - \hat{\delta}_\omega \right) &= \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \left(\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right)^{-1} \\
&\quad \tilde{c}_j' \hat{\delta}_\Omega
\end{aligned}$$

maka akan diperoleh persamaan sebagai berikut:

$$\begin{aligned}
C &= \left(\left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \left(\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right)^{-1} \tilde{c}_j' \hat{\delta}_\Omega \right)' \\
&\quad \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \\
&\quad \tilde{c}_j \left(\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right)^{-1} \tilde{c}_j' \hat{\delta}_\Omega
\end{aligned}$$

Sehingga persamaan D menjadi

$$\begin{aligned}
D &= \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right) + \\
&\quad \left(\left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \left(\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right)^{-1} \tilde{c}_j' \hat{\delta}_\Omega \right)' \\
&\quad \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1}
\end{aligned}$$

$$\begin{aligned}
& \tilde{c}_j \left(\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right)^{-1} \tilde{c}_j' \hat{\delta}_\Omega \\
&= \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right) + \\
& \quad \left(\tilde{c}_j' \hat{\delta}_\Omega \right)' \left(\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right)^{-1} \tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{Z}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \\
& \quad \mathbf{Z}'(\mathbf{X}, \mathbf{Z}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{Z}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \\
& \quad \left(\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right)^{-1} \tilde{c}_j' \hat{\delta}_\Omega \\
&= \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right) + \left(\tilde{c}_j' \hat{\delta}_\Omega \right)' I \\
& \quad \left(\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right)^{-1} \tilde{c}_j' \hat{\delta}_\Omega \\
&= \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right) + \\
& \quad \left(\tilde{c}_j' \hat{\delta}_\Omega \right)' \left(\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right)^{-1} \tilde{c}_j' \hat{\delta}_\Omega \\
&= Q_2 + Q_1
\end{aligned}$$

dimana

$$\begin{aligned}
Q_2 &= \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right) \\
Q_1 &= \left(\tilde{c}_j' \hat{\delta}_\Omega \right)' \left(\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right)^{-1} \tilde{c}_j' \hat{\delta}_\Omega
\end{aligned}$$

Selanjutnya persamaan D akan disubstitusikan ke ratio likelihood yang sudah diperoleh sebelumnya yaitu pada persamaan (4.52) sebagai berikut.

$$\begin{aligned}
\lambda(y) &= \left(\frac{\left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right)}{\left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\omega \right)} \right)^{\frac{n}{2}} \\
&= \left(\frac{\left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right)}{\left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\omega \right) + \left(\tilde{c}_j' \hat{\delta}_\Omega \right)' \left(\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right)^{-1} \tilde{c}_j' \hat{\delta}_\Omega} \right)^{\frac{n}{2}}
\end{aligned}$$

$$\begin{aligned}
&= \left(\frac{1}{1 + \frac{(c_j' \hat{\delta}_\Omega)' (c_j' (\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}))^{-1} c_j)^{-1} c_j' \hat{\delta}_\Omega}{(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega)' (\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega)}}} \right)^{\frac{n}{2}} \\
&= \left(\frac{1}{1 + \frac{Q_1}{Q_2}} \right)^{\frac{n}{2}} \tag{4.55}
\end{aligned}$$

Berdasarkan persamaan (4.55) maka didapat statistik uji yang disimbolkan dengan W dari hipotesis $H_0: \tilde{c}_j' \tilde{\delta} = 0$ melawan $H_1: \tilde{c}_j' \tilde{\delta} \neq 0$ adalah:

$$W = \frac{Q_1}{Q_2} = \frac{(\tilde{c}_j' \hat{\delta}_\Omega)' (\tilde{c}_j' (\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}))^{-1} \tilde{c}_j)^{-1} \tilde{c}_j' \hat{\delta}_\Omega}{(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega)' (\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega)} \tag{4.56}$$

4.1.4 Mendapatkan Distribusi Statistik Uji Hipotesis

Untuk mendapatkan distribusi statistik uji maka akan dilakukan penjabaran pada persamaan Q_1 dan Q_2 . Langkah pertama adalah menjabarkan Q_1 sebagai berikut:

$$Q_1 = (\tilde{c}_j' \hat{\delta}_\Omega)' (\tilde{c}_j' (\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}))^{-1} \tilde{c}_j)^{-1} \tilde{c}_j' \hat{\delta}_\Omega \tag{4.57}$$

Dengan mensubstitusikan $\hat{\delta}_\Omega$ pada Q_1 maka persamaan (4.57) akan menjadi

$$\begin{aligned}
Q_1 &= (\tilde{c}_j' (\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}))^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y})' \\
&\quad \left[\tilde{c}_j' (\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}))^{-1} \tilde{c}_j \right]^{-1} \\
&\quad \tilde{c}_j' (\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}))^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y} \\
&= \tilde{y}' \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) (\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}))^{-1} \tilde{c}_j \left[\tilde{c}_j' (\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}))^{-1} \tilde{c}_j \right]^{-1} \\
&\quad \tilde{c}_j' (\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}))^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y} \tag{4.58}
\end{aligned}$$

Dari persamaan(4.58) dapat dinyatakan bahwa

$$\begin{aligned} & \tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \\ & = \mathbf{A} \end{aligned}$$

terbukti bahwa matrik \mathbf{A} adalah *idempotent*.

Karena \mathbf{A} adalah simetris dan *idempotent* maka statistik

$$\frac{\tilde{y}' \mathbf{A} \tilde{y}}{\sigma^2} \sim \chi^2(r_1, \tilde{\mu}' \mathbf{A} \tilde{\mu} / 2\sigma^2) \quad (4.59)$$

Selanjutnya akan didapatkan nilai r_1 dan $\tilde{\mu}' \mathbf{A} \tilde{\mu} / 2\sigma^2$.

Karena matrik \mathbf{A} adalah simetris dan *idempotent* maka $r_1 = \text{rank}(\mathbf{A}) = \text{tr}(\mathbf{A})$

$$\begin{aligned} \text{tr}(\mathbf{A}) &= \text{tr} \left(\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right. \\ & \left. \left[\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right]^{-1} \tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \right) \end{aligned}$$

Jika

$$\begin{aligned} \mathbf{B} &= \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \\ & \left[\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right]^{-1} \end{aligned}$$

dan

$$\mathbf{C} = \tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \text{ maka}$$

$$\text{tr}(\mathbf{A}) = \text{tr}(\mathbf{BC}) = \text{tr}(\mathbf{CB}) \text{ sehingga}$$

$$\text{tr}(\mathbf{A}) = \text{tr} \left(\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)$$

$$\left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \left[\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right]^{-1}$$

$$= \text{tr}(\mathbf{I}_{(1)})$$

$$= 1$$

Sehingga $r_1 = \text{rank}(\mathbf{A}) = \text{tr}(\mathbf{A}) = 1$

Selanjutnya akan didapatkan nilai dari $\frac{\tilde{\mu}' \mathbf{A} \tilde{\mu}}{2\sigma^2}$ sebagai berikut:

$$\frac{\tilde{\mu}' \mathbf{A} \tilde{\mu}}{2\sigma^2} = \frac{1}{2\sigma^2} \left(\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\delta} \right)' \left(\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \right)$$

$$\tilde{c}_j \left[\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right]^{-1} \tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1}$$

$$\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\delta}$$

Akan dilakukan penjabaran pada pembilang sebagai berikut:

$$\begin{aligned} \tilde{\mu}' \mathbf{A} \tilde{\mu} &= (\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\delta})' \left(\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \right. \\ &\quad \left. \tilde{c}_j \left[\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right]^{-1} \tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \right. \\ &\quad \left. \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \right) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\delta} \\ &= \tilde{\delta}' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \right. \\ &\quad \left. \tilde{c}_j \left[\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right]^{-1} \tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \right. \\ &\quad \left. \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \right) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\delta} \\ &= \tilde{\delta}' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \\ &\quad \tilde{c}_j \left[\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right]^{-1} \tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \\ &\quad \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\delta} \\ &= \left(\tilde{c}_j' \tilde{\delta} \right)' \left[\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right]^{-1} \tilde{c}_j \tilde{\delta} \\ &= 0 \end{aligned}$$

Sehingga dapat dinyatakan bahwa

$$\frac{Q_1}{\sigma^2} \sim \chi^2(1) \quad (4.60)$$

Langkah kedua yaitu menjabarkan Q_2 sebagai berikut:

$$\begin{aligned} Q_2 &= \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right)' \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right) \\ &= \left(\tilde{y}' - \hat{\delta}_\Omega' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \right) \left(\tilde{y} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \right) \\ &= \tilde{y}' \tilde{y} - \tilde{y}' \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega - \hat{\delta}_\Omega' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y} + \hat{\delta}_\Omega' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \\ &= \tilde{y}' \tilde{y} - 2 \hat{\delta}_\Omega' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y} + \hat{\delta}_\Omega' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \hat{\delta}_\Omega \end{aligned}$$

Dengan mensubstitusikan $\hat{\delta}_\Omega$ maka akan diperoleh:

$$Q_2 = \tilde{y}' \tilde{y} - 2 \left(\left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y} \right)' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y} +$$

$$\begin{aligned}
& \left(\left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\mathbf{y}} \right)' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \\
& \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\mathbf{y}} \\
&= \tilde{\mathbf{y}}' \mathbf{I} \tilde{\mathbf{y}} - 2 \tilde{\mathbf{y}}' \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\mathbf{y}} + \\
& \quad \tilde{\mathbf{y}}' \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \\
& \quad \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\mathbf{y}} \\
&= \tilde{\mathbf{y}}' \mathbf{I} \tilde{\mathbf{y}} - 2 \tilde{\mathbf{y}}' \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\mathbf{y}} + \\
& \quad \tilde{\mathbf{y}}' \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\mathbf{y}} \\
&= \tilde{\mathbf{y}}' \mathbf{I} \tilde{\mathbf{y}} - \tilde{\mathbf{y}}' \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\mathbf{y}} \\
&= \tilde{\mathbf{y}}' \left[\mathbf{I} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \right] \tilde{\mathbf{y}} \quad (4.61) \\
&= \tilde{\mathbf{y}}' \mathbf{B} \tilde{\mathbf{y}} \quad (4.62)
\end{aligned}$$

Dari persamaan (4.61) dapat dinyatakan bahwa

$$\mathbf{B} = \mathbf{I} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})$$

Selanjutnya akan dibuktikan bahwa \mathbf{B} adalah simetris dan *idempotent*. Matrik \mathbf{B} simetris jika $\mathbf{B}' = \mathbf{B}$, berikut penjabarannya

$$\begin{aligned}
\mathbf{B}' &= \left(\mathbf{I} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)' \\
&= \mathbf{I} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \\
&= \mathbf{B}
\end{aligned}$$

Dari penjabaran diatas terlihat bahwa $\mathbf{B}' = \mathbf{B}$ dengan ukuran $n \times n$ sehingga terbukti bahwa matrik \mathbf{B} adalah simetris. Selanjutnya akan dibuktikan bahwa matrik \mathbf{B} adalah *idempotent*. Berikut penjabarannya

$$\begin{aligned}
\mathbf{B}^2 &= \left[\mathbf{I} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \right]^2 \\
&= \mathbf{I} - 2 \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) + \\
& \quad \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k})
\end{aligned}$$

$$\begin{aligned}
& \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \\
&= \mathbf{I} - 2\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) + \\
& \quad \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \\
&= \mathbf{I} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \\
&= \mathbf{B}
\end{aligned}$$

terbukti bahwa matrik \mathbf{B} adalah matrik *idempotent*.

Karena \mathbf{B} adalah simetris dan *idempotent* maka dapat dinyatakan bahwa

$$\frac{\tilde{\mathbf{y}}' \mathbf{B} \tilde{\mathbf{y}}}{\sigma^2} \sim \chi^2(r_2, \tilde{\boldsymbol{\mu}}' \mathbf{B} \tilde{\boldsymbol{\mu}} / 2\sigma^2) \quad (4.63)$$

Selanjutnya akan didapatkan nilai r_2 dan $\tilde{\boldsymbol{\mu}}' \mathbf{B} \tilde{\boldsymbol{\mu}} / 2\sigma^2$.

Karena matrik \mathbf{B} adalah simetris dan *idempotent* maka $r_2 = \text{rank}(\mathbf{B}) = \text{tr}(\mathbf{B})$

$$\begin{aligned}
\text{tr}(\mathbf{B}) &= \text{tr} \left(\mathbf{I} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \right) \\
&= \text{tr}(\mathbf{I}) - \text{tr} \left(\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \right) \\
&= \text{tr}(\mathbf{I}_n) - [p + 1 + (m + r)q] \\
&= n - [p + 1 + (m + r)q]
\end{aligned}$$

Sehingga didapatkan

$$r_2 = \text{rank}(\mathbf{B}) = \text{tr}(\mathbf{B}) = n - [p + 1 + (m + r)q].$$

Selanjutnya akan didapatkan nilai $\frac{\tilde{\boldsymbol{\mu}}' \mathbf{B} \tilde{\boldsymbol{\mu}}}{2\sigma^2}$ sebagai berikut:

$$\begin{aligned}
\frac{\tilde{\boldsymbol{\mu}}' \mathbf{B} \tilde{\boldsymbol{\mu}}}{2\sigma^2} &= \frac{1}{2\sigma^2} (\mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\boldsymbol{\delta}})' \left(\mathbf{I} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \right. \\
& \quad \left. \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \right) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\boldsymbol{\delta}} \\
&= \frac{1}{2\sigma^2} (\tilde{\boldsymbol{\delta}}' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\boldsymbol{\delta}} - \tilde{\boldsymbol{\delta}}' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \\
& \quad \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\boldsymbol{\delta}}) \\
&= \frac{\tilde{\boldsymbol{\delta}}' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\boldsymbol{\delta}} - \tilde{\boldsymbol{\delta}}' \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{\boldsymbol{\delta}}}{2\sigma^2} \\
&= 0
\end{aligned}$$

Sehingga dapat dinyatakan bahwa

$$\frac{Q_2}{\sigma^2} \sim \chi^2(n - (p + 1 + (r + m)q)) \quad (4.64)$$

Langkah selanjutnya adalah membuktikan bahwa Q_1 dan Q_2 *independent*. Pembuktian ini menggunakan Teorema 8 (Rencher, 2007) Corollary 1 yang sudah dituliskan pada Bab 2. Sesuai dengan persamaan (4.58) dan persamaan (4.61) maka didapatkan

$$\mathbf{A} = \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \left[\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right]^{-1} \\ \tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \text{ dan}$$

$$\mathbf{B} = \mathbf{I} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})$$

Sehingga

$$\mathbf{AB} = \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \left[\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right]^{-1} \\ \tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \\ \left[\mathbf{I} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \right] \\ = \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \left[\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right]^{-1} \tilde{c}_j' \\ \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \\ \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \left[\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right]^{-1} \tilde{c}_j' \\ \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \\ \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \\ \mathbf{AB} = \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \left[\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right]^{-1} \\ \tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \\ \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \left[\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right]^{-1} \tilde{c}_j' \\ \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k})$$

= 0

Karena $\mathbf{AB} = 0$ maka dapat dikatakan bahwa Q_1 dan Q_2 *independent*.

Dari penjabaran di atas didapatkan hasil sebagai berikut: $\frac{Q_1}{\sigma^2} \sim \chi^2(1)$,
 $\frac{Q_2}{\sigma^2} \sim \chi^2(n - (p + 1 + (r + m)q))$, serta Q_1 dan Q_2 *independent*.

Berdasarkan teorema Rencher (2007) statistik

$$\begin{aligned} W^* &= \frac{Q_1/(1)}{Q_2/(n - (p + 1 + (r + m)q))} \sim F_{(1, (n - (p + 1 + (r + m)q)))} \\ &= \frac{(Q_1/\sigma^2)/1}{(Q_2/\sigma^2)/(n - (p + 1 + (r + m)q))} \sim F_{(1, (n - (p + 1 + (r + m)q)))} \end{aligned}$$

Sehingga distribusi statistik uji yang dihasilkan adalah

$$W^* \sim F_{(1, (n - (p + 1 + (r + m)q)))} \quad (4.65)$$

dimana

$$\begin{aligned} Q_1 &= \tilde{y}' \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \\ &\quad \left[\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right]^{-1} \\ &\quad \tilde{c}_j \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y} \\ Q_2 &= \tilde{y}' \left[\mathbf{I} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \right] \tilde{y} \end{aligned}$$

4.1.5 Menentukan Daerah Penolakan Uji Hipotesis H_0

Berdasarkan metode LRT, daerah penolakan $\lambda < k$, dimana $0 < \lambda < 1$, dan k adalah konstanta sehingga

$$\lambda = \frac{L(\hat{\omega})}{L(\hat{\Omega})} < k$$

Selanjutnya dengan mensubstitusikan ratio likelihood yang sudah dihasilkan sebelumnya yaitu persamaan (4.55) maka akan didapatkan

$$\left(\frac{1}{1 + \frac{Q_1}{Q_2}} \right)^{\frac{n}{2}} < k$$

$$\frac{1}{\left(1 + \frac{Q_1}{Q_2}\right)^{\frac{n}{2}}} < k$$

$$\frac{1}{k} < \left(1 + \frac{Q_1}{Q_2}\right)^{\frac{n}{2}}$$

Jika kedua ruas dipangkatkan dengan $\frac{2}{n}$ maka persamaan diatas akan menjadi

$$\left(\frac{1}{k}\right)^{\frac{2}{n}} < \left(1 + \frac{Q_1}{Q_2}\right)^1$$

$$1 + \frac{Q_1}{Q_2} > \left(\frac{1}{k}\right)^{\frac{2}{n}}$$

$$\frac{Q_1}{Q_2} > \left(\frac{1}{k}\right)^{\frac{2}{n}} - 1$$

Jika kedua ruas dikalikan dengan $(n - (p + 1 + (r + m)q))/1$ maka persamaan akan menjadi

$$\frac{Q_1/(1)}{Q_2/(n - (p + 1 + (r + m)q))} > \left(\left(\frac{1}{k}\right)^{\frac{2}{n}} - 1\right) \frac{(n - (p + 1 + (r + m)q))}{1}$$

Berdasarkan persamaan (4.65) maka dapat dituliskan

$$W^* > \left(\left(\frac{1}{k}\right)^{\frac{2}{n}} - 1\right) (n - (p + 1 + (r + m)q))$$

$$W^* > k^*$$

dimana $k^* = \left(\left(\frac{1}{k}\right)^{\frac{2}{n}} - 1\right) (n - (p + 1 + (r + m)q))$

Daerah kritik untuk uji $H_0 : \tilde{c}'_j \tilde{\delta} = 0, H_1 : \tilde{c}'_j \tilde{\delta} \neq 0$ adalah

$$C = \{(x_1, x_2, \dots, x_p, t_1, t_2, \dots, t_q, y); W^* > k^*\} \quad (4.66)$$

Jika diberikan taraf signifikansi α maka

$$\alpha = P(\text{menolak } H_0 | H_0 \text{ benar})$$

$$= P(W^* > k^* | \tilde{c}'_j \tilde{\delta} = 0) \quad (4.67)$$

dengan

$$W^* = \frac{Q_1/(1)}{Q_2/(n - (p + 1 + (r + m)q))} \sim F_{((1), (n - (p + 1 + (r + m)q)))}$$

Dimana

$$Q_1 = \tilde{y}' \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_i$$

$$\left[\tilde{c}_i' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_i \right]^{-1}$$

$$\tilde{c}_j' \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y}$$

$$Q_2 = \tilde{y}' \left[\mathbf{I} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \right] \tilde{y}$$

Nilai k^* (F tabel) didapat dari persamaan

$$\int_{k^*}^{\infty} f(w) dw = \alpha$$

Dimana $f(w)$ fungsi distribusi F dengan derajat bebas 1 dan $(n - (p + 1 + (r + m)q))$.

4.2 Aplikasi pada Data AHH Tahun 2015 di Indonesia

Model dan pengujian hipotesis regresi semiparametrik spline *truncated* yang telah dibahas pada sub bab 4.1 akan diaplikasikan pada data AHH Tahun 2015 di Indonesia. Variabel respon yang digunakan adalah AHH (y), sedangkan variabel prediktor yang digunakan sebanyak lima variabel yaitu persentase penduduk miskin (x_1), persentase rumah tangga yang menggunakan air minum bersih (x_2), persentase penduduk yang mengalami keluhan kesehatan (x_3), persentase balita yang mendapat imunisasi lengkap (x_4) dan MYS (x_5). Unit observasi yang digunakan sebanyak 34 provinsi.

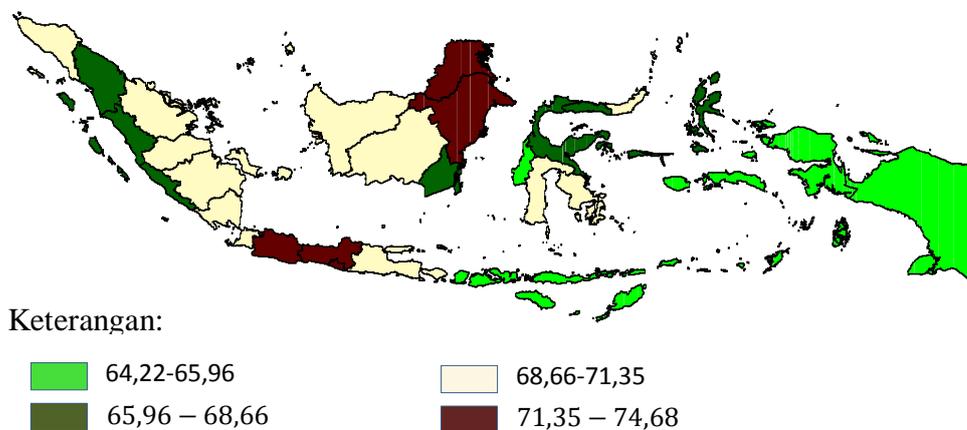
4.2.1 Analisis Deskriptif

Sebelum melakukan pemodelan dan pengujian hipotesis, akan dilakukan eksplorasi data melalui analisis deskriptif. Hasil dari statistik deskriptif dapat digunakan untuk inisiasi titik-titik knot pada tahap selanjutnya.

Pencapaian AHH sebagian besar provinsi di Indonesia, yaitu 73,53 persen dari 34 provinsi, masih di bawah pencapaian AHH nasional. Hampir seluruh provinsi tersebut berada wilayah Indonesia bagian timur. Adapun provinsi yang sudah mencapai AHH di atas AHH nasional adalah Provinsi Riau, Provinsi

Sulawesi Utara, Provinsi Bali, Provinsi Kalimantan Utara, Provinsi Jawa Barat, Provinsi Daerah Khusus Ibukota (DKI) Jakarta, Provinsi Kalimantan Timur, Provinsi Jawa Tengah, dan Provinsi Daerah Istimewa Yogyakarta (DIY).

Gambar 4.1 menunjukkan AHH Indonesia menurut provinsi yang dibagi ke dalam empat interval. Enam provinsi yang menempati posisi interval terbawah dengan nilai AHH 64,22 – 65,96 tahun adalah Provinsi Sulawesi Barat, Provinsi Papua, Provinsi Papua barat, Provinsi Maluku, Provinsi Nusa Tenggara Barat (NTB), dan Provinsi Nusa Tenggara Timur (NTT). Provinsi yang memiliki AHH pada interval 65,96 – 68,66 adalah Provinsi Gorontalo, Provinsi Sulawesi Tengah, Provinsi Maluku Utara, Provinsi Kalimantan Selatan, Provinsi Sumatera Utara, Provinsi Bengkulu, dan Provinsi Sumatera Barat. Provinsi yang berada pada interval AHH 68,66 – 71,35 tahun adalah Provinsi Sumatera Barat, Provinsi Sumatera Selatan, Provinsi Kepulauan Riau, Provinsi Aceh, Provinsi Kalimantan Tengah, Provinsi Sulawesi Selatan, Provinsi Kalimantan Barat, Provinsi Bangka Belitung, Provinsi Lampung, Provinsi Sulawesi Tenggara, Provinsi Jambi, Provinsi Jawa Timur, dan Provinsi Riau. Adapun provinsi yang berada pada interval tertinggi dengan nilai AHH 71,35 – 74,68 tahun adalah Provinsi Kalimantan Utara, Provinsi Jawa Barat, Provinsi DKI Jakarta, Provinsi Kalimantan Timur, dan Provinsi Jawa Tengah, dan Provinsi DIY.



Gambar 4. 1. Peta AHH menurut Provinsi di Indonesia

Karakteristik masing-masing variabel, baik variabel respon maupun variabel prediktor dapat dilihat pada Tabel 4.1. Berdasarkan Tabel 4.1, variabel respon (y)

yaitu AHH Indonesia pada tahun 2015 memiliki nilai rata-rata sebesar 69,32 tahun dengan standar deviasi sebesar 2,66. Adapun range nilai AHH di Indonesia meliputi 34 provinsi adalah 10,46. Nilai AHH tertinggi sebesar 74,68 tahun yaitu di Provinsi DIY. Sedangkan nilai AHH terendah dengan nilai 64,22 tahun terdapat di Provinsi Sulawesi Barat.

Tabel 4.1 Statistik Deskriptif Variabel Respon dan Variabel Prediktor

Variabel	Minimum	Maksimum	Range	Mean	Standar Deviasi
y	64,22	74,68	10,46	69,32	2,66
x_1	3,93	27,33	23,40	11,83	6,16
x_2	41,08	93,4	52,32	68,62	11,04
x_3	16,71	39,58	22,87	28,42	6,11
x_4	40,36	76,01	35,65	59,17	8,73
x_5	5,99	10,70	4,71	8,02	0,96

Selanjutnya gambaran karakteristik untuk masing-masing variabel prediktor, yaitu persentase penduduk miskin (x_1), persentase rumah tangga yang menggunakan air minum bersih (x_2), persentase penduduk yang mengalami keluhan kesehatan (x_3), persentase balita yang mendapat imunisasi lengkap (x_4), dan MYS (x_5) adalah sebagai berikut :

- a. Persentase penduduk miskin di Indonesia pada tahun 2015 yang tertinggi ada di Propinsi Papua dan yang paling rendah ada di Propinsi Daerah Khusus Ibukota (DKI) Jakarta, dengan range persentase penduduk miskin adalah sebesar 23,40 persen. Rata-rata persentase penduduk miskin di Indonesia pada tahun 2015 adalah 11,83 persen dengan standar deviasi sebesar 6,16.
- b. Persentase rumah tangga yang menggunakan air minum bersih terendah di Indonesia pada tahun 2015 adalah Propinsi Bengkulu yaitu 41,08 persen dan yang tertinggi adalah Propinsi DKI Jakarta yaitu 93,40. Sehingga range persentase rumah tangga yang menggunakan air minum bersih sebesar 52,32 persen. Rata-rata persentase rumah tangga yang menggunakan air minum bersih di Indonesia adalah 68,62 persen dengan standar deviasi sebesar 11,04.
- c. Persentase penduduk yang mengalami keluhan kesehatan di Indonesia secara rata-rata pada tahun 2015 adalah 28,42 persen dengan standar deviasi sebesar

- 6,11. Adapun range persentase penduduk yang mengalami keluhan kesehatan di Indonesia adalah 22,87 persen dengan persentase tertinggi adalah Propinsi DIY yaitu 39,58 persen dan persentase yang terendah adalah Propinsi Maluku Utara yaitu 16,71 persen.
- d. Persentase balita yang mendapatkan imunisasi lengkap terendah di Indonesia ada di Propinsi Aceh yaitu sebesar 40,36 persen, sedangkan persentase tertinggi adalah 76,01 persen di Provinsi DIY. Range persentase balita yang mendapatkan imunisasi lengkap di Indonesia sebesar 35,65. Rata-rata persentase balita yang mendapatkan imunisasi lengkap di Indonesia pada tahun 2015 sebanyak 59,17 persen.
 - e. Penduduk Indonesia yang berusia 15 tahun ke atas pada tahun 2015 secara rata-rata menempuh pendidikan selama 8 tahun. Hal ini ditunjukkan oleh angka MYS di Indonesia secara rata-rata adalah 8,02 tahun. Sehingga penduduk Indonesia yang berusia 15 tahun ke atas secara umum mengenyam pendidikan sampai dengan level Sekolah Menengah Pertama (SMP). Adapun range MYS di Indonesia adalah 4,71 tahun, yaitu berada di antara 5,99 tahun dan 10,70 tahun. MYS terendah ada di Provinsi Papua, sedangkan MYS tertinggi ada di Provinsi DKI Jakarta.

4.2.2 Memodelkan AHH menggunakan Regresi Linier Berganda

Model regresi parametrik seperti pada persamaan (2.4) dalam hal ini didekati dengan regresi linier berganda dengan y_i adalah variabel respon, $\beta_0, \beta_1, \dots, \beta_p$ adalah parameter yang tidak diketahui, $x_{i1}, x_{i2}, \dots, x_{ip}$ adalah variabel prediktor dan ε_i adalah error random yang identik, independen, berdistribusi normal dengan mean nol dan variansi σ^2 . Maka estimasi kurva regresi linier berganda untuk AHH Tahun 2015 di Indonesia diberikan oleh:

$$\hat{y}_i = 57,97 - 0,12x_{i1} + 0,05x_{i2} + 0,03x_{i3} + 0,06x_{i4} + 0,55x_{i5}$$

dimana \hat{y}_i adalah AHH, x_{i1} adalah persentase penduduk miskin, x_{i2} adalah persentase rumah tangga menggunakan air minum bersih, x_{i3} adalah persentase penduduk mengalami keluhan kesehatan, x_{i4} adalah persentase balita mendapat imunisasi lengkap dan x_{i5} adalah MYS.

4.2.2.1 Pengujian Signifikansi Parameter secara Simultan

Pengujian hipotesis untuk menguji signifikansi parameter secara simultan menggunakan hipotesis sebagai berikut:

$$H_0 : \beta_1 = \dots = \beta_5 = 0$$

$$H_1 : \text{paling sedikit terdapat satu } \beta_h \neq 0, \text{ dimana } h = 1, \dots, 5$$

Hasil ANOVA pada pengujian hipotesis secara simultan tersaji pada Tabel 4.2.

Tabel 4.2 ANOVA Hasil Regresi Linier Berganda

Sumber Variasi	Derajat Kebebasan	Jumlah Kuadrat	Rata-rata Kuadrat	F_{hitung}	p-value
Regresi	5	113,81	22,76	5,36	0,001
Error	28	118,89	4,25		
Total	33	232,70			

Berdasarkan Tabel 4.2 terlihat bahwa nilai *p-value* sebesar 0,00 lebih kecil dari nilai α yang ditentukan, sehingga diperoleh keputusan tolak H_0 . Oleh karena itu dapat disimpulkan bahwa paling sedikit terdapat satu variabel prediktor yang signifikan dalam model. Model regresi linier berganda ini memiliki koefisien determinasi (R^2) sebesar 48,90 persen. Hal ini berarti bahwa variasi dari variabel respon dapat dijelaskan oleh kelima variabel prediktor sebesar 48,90 persen, sedangkan sisanya 51,10 persen dijelaskan oleh faktor lain.

4.2.2.2 Pengujian Signifikansi Parameter secara Parsial

Untuk mengetahui variabel yang berpengaruh dalam model dilakukan pengujian signifikansi parameter secara parsial. Pengujian signifikansi parameter secara parsial menggunakan hipotesis sebagai berikut:

$$H_0 : \beta_h = 0$$

$$H_1 : \beta_h \neq 0 ; \text{ dimana } h = 0,1, \dots, 5$$

Hasil pengujian signifikansi parameter secara parsial tersaji pada Tabel 4.3.

Tabel 4.3 Hasil Uji Hipotesis secara Parsial

Variabel	Parameter	Estimator	t_{hitung}	P -value
Konstan	β_0	57,97	11,31	0,00
x_1	β_1	-0,12	-1,73	0,09
x_2	β_2	0,06	1,22	0,23
x_3	β_3	0,03	0,43	0,67
x_4	β_4	0,06	0,99	0,33
x_5	β_5	0,55	1,15	0,26

Berdasarkan Tabel 4.3 terlihat bahwa hanya variabel persentase penduduk miskin yang berpengaruh terhadap AHH.

4.2.2.3 Pengujian Asumsi Residual

Residual pada regresi linier berganda harus memenuhi identik, independen dan berdistribusi normal (IIDN). Pengujian asumsi residual identik menggunakan uji *glejser*. Pengujian asumsi independen menggunakan uji *Durbin-Watson*. Adapun pengujian asumsi normal menggunakan uji *Kolmogorov-Smirnov*.

4.2.2.3.1 Pengujian Asumsi Identik

Pengujian asumsi identik dimaksudkan untuk mendeteksi apakah varians dari residualnya memiliki varians yang homogen atau tidak. Pemeriksaan menggunakan uji *glejser*. Hipotesis yang digunakan adalah:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1 : \text{paling sedikit terdapat satu } \sigma_i^2 \neq \sigma^2 ; \text{ dimana } i = 1, 2, \dots, n$$

Hasil pengujian *glejser* tersaji pada Tabel 4.4.

Tabel 4.4 Hasil Uji *Glejser*

Sumber Variasi	Derajat Kebebasan	Jumlah Kuadrat	Rata-rata Kuadrat	F_{hitung}	p -value
Regresi	5	5,43	1,09	1,08	0,40
Error	28	28,16	1,01		
Total	33	33,59			

Dari Tabel 4.4 terlihat bahwa nilai p -value sebesar 0,40 lebih besar dari nilai α sehingga keputusan yang diambil adalah gagal tolak H_0 . Hal ini memberikan kesimpulan bahwa varians dari residual adalah homogen, sehingga asumsi identik pada model regresi linier berganda yang terbentuk telah terpenuhi.

4.2.2.3.2 Pemeriksaan Asumsi Independen

Pengujian asumsi independen dimaksudkan untuk mendeteksi korelasi antar residual. Pemeriksaan dilakukan dengan menggunakan uji *Durbin-Watson*. Hipotesis yang digunakan adalah sebagai berikut:

$$H_0: \rho = 0 \text{ (tidak ada korelasi antara residual)}$$

$$H_1: \rho \neq 0 \text{ (ada korelasi antara residual)}$$

dengan statistik uji sebagai berikut :

$$d_{hitung} = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}$$

Hasil penghitungan diperoleh $d_{hitung} = 1,13$, dengan $d_{L, \frac{0,05}{2}} = 1,14$ maka diperoleh keputusan tolak H_0 . Hal ini memberikan kesimpulan bahwa antar residual terjadi korelasi. Dengan demikian asumsi independen pada model regresi linier berganda tidak terpenuhi.

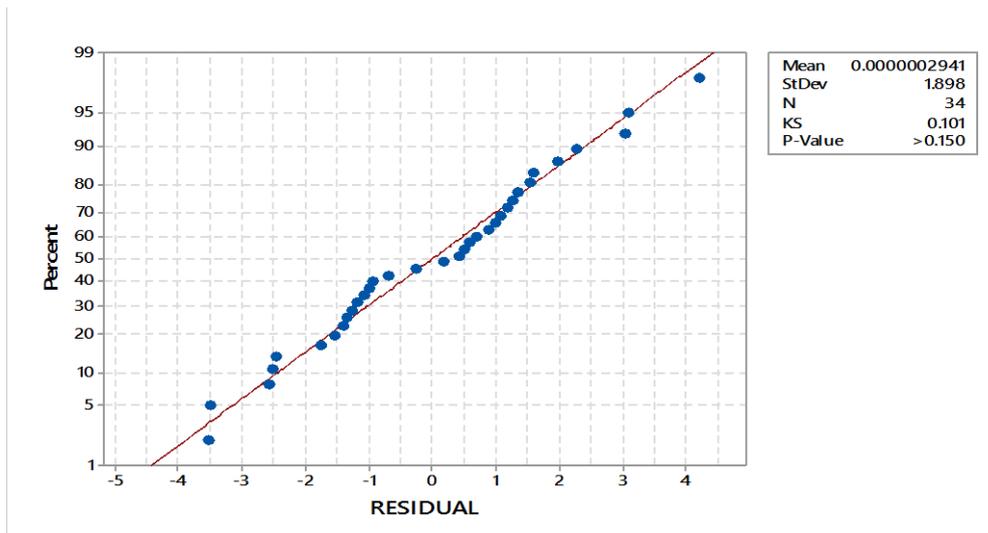
4.2.2.3.3 Pengujian Asumsi Distribusi Normal

Residual regresi linier berganda harus memenuhi asumsi distribusi normal. Oleh karena itu dilakukan pemeriksaan asumsi distribusi normal pada residual yang diperoleh. Pemeriksaan dilakukan dengan menggunakan uji *Kolmogorov-Smirnov*. Hipotesis yang digunakan adalah sebagai berikut:

$$H_0 : F_0(x) = F(x) \text{ (Residual berdistribusi Normal)}$$

$$H_1 : F_0(x) \neq F(x) \text{ (Residual tidak berdistribusi Normal).}$$

Hasil uji *Kolmogorov-Smirnov* disajikan dalam bentuk plot seperti terlihat pada Gambar 4. 2..



Gambar 4. 2 *Probability Plot Residual*

Plot residual seperti yang terlihat pada Gambar 4.2 menghasilkan *p-value* sebesar 0,150 yang lebih besar dari nilai α yang ditentukan, sehingga keputusan yang diambil adalah gagal tolak H_0 . Dari hasil uji ini dapat disimpulkan bahwa regresi linier berganda yang dihasilkan memenuhi asumsi berdistribusi normal.

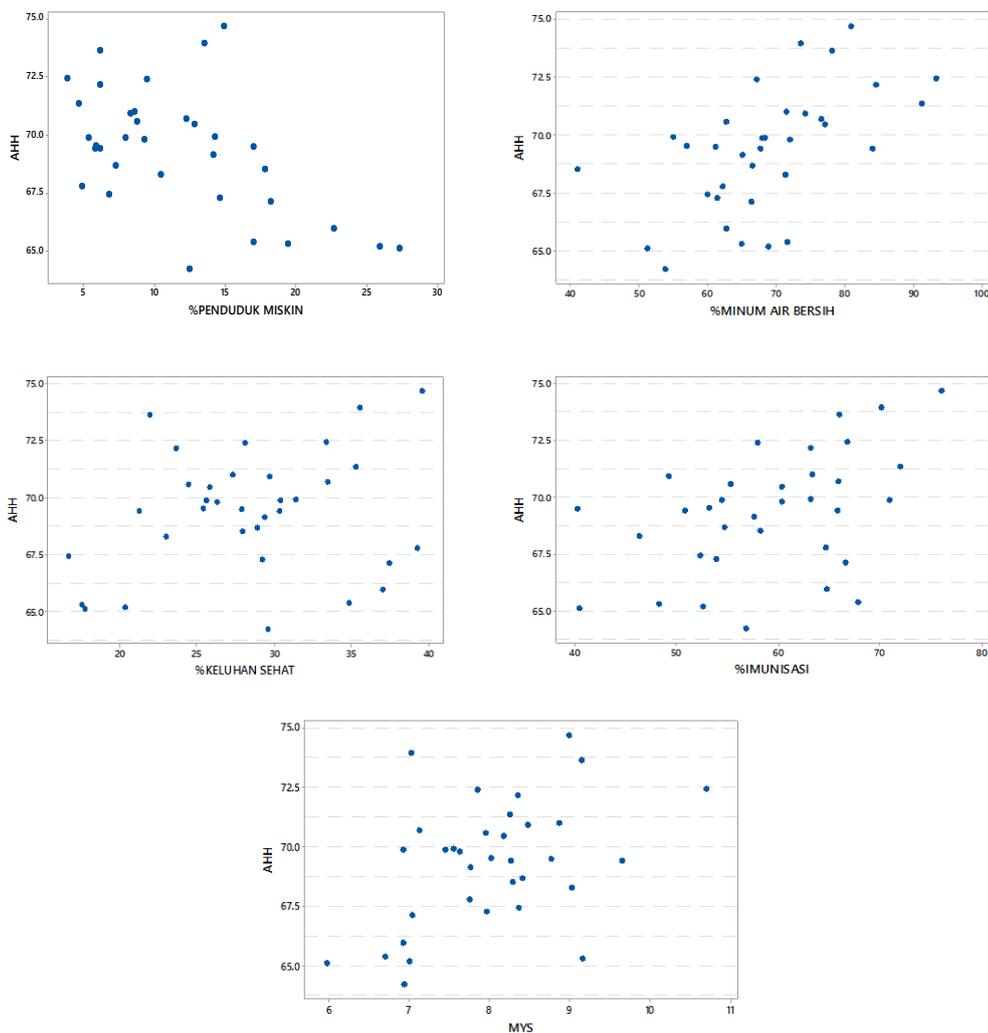
4.2.3 Memodelkan AHH menggunakan Regresi Semiparametrik Spline *Truncated*

Eksplorasi data selanjutnya adalah membuat *scatter plot*. Hal ini dilakukan untuk mengidentifikasi bentuk pola hubungan antara variabel respon dengan masing-masing variabel prediktor. Selanjutnya dapat digunakan untuk menentukan variabel sebagai komponen parametrik dan komponen nonparametrik.

4.2.3.1 Menentukan Variabel Komponen Parametrik dan Nonparametrik

Tahap awal yang dilakukan untuk melakukan proses pemodelan regresi semiparametrik spline *truncated* adalah dengan membuat *scatter plot* antara variabel respon dengan masing-masing variabel prediktor. Dari hasil *scatter plot* ini dapat dilihat bentuk pola hubungan antara variabel respon dengan masing-masing variabel prediktor. Dari informasi bentuk pola hubungan tersebut dapat

digunakan untuk menentukan variabel prediktor yang akan digunakan sebagai komponen parametrik dan variabel sebagai komponen nonparametrik.



Gambar 4. 3 Scatter Plot antara AHH dengan Variabel-variabel Prediktor

Pada Gambar 4.3 menunjukkan bentuk pola hubungan antara variabel AHH dengan variabel-variabel prediktor yaitu persentase penduduk miskin, persentase rumah tangga yang menggunakan air minum bersih, persentase penduduk yang mengalami keluhan kesehatan, persentase balita mendapat imunisasi lengkap, dan MYS.

Pola hubungan antara variabel AHH dengan variabel persentase penduduk miskin terlihat membentuk pola linier. Dari hasil *scatter plot* tersebut maka variabel persentase penduduk miskin didekati sebagai komponen parametrik. Sedangkan pola hubungan antara variabel AHH dengan persentase rumah tangga

yang menggunakan air minum bersih, AHH dengan persentase penduduk yang mengalami keluhan kesehatan, AHH dengan persentase balita yang mendapat imunisasi lengkap, dan AHH dengan MYS cenderung tidak membentuk pola tertentu. Pola hubungan yang terlihat cenderung mengalami perubahan perilaku pada sub interval tertentu. Dari hasil *scatter plot* pada gambar 4.3, maka variabel persentase rumah tangga yang menggunakan air minum bersih, persentase penduduk yang mengalami keluhan kesehatan, persentase balita yang mendapat imunisasi lengkap, dan MYS dapat didekati sebagai komponen nonparametrik.

Ringkasan hasil penentuan komponen parametrik dan komponen nonparametrik dituliskan pada Tabel 4. 5.

Tabel 4.5 Ringkasan Penentuan Komponen Parametrik dan Nonparametrik

Notasi	Nama Variabel	Komponen
x_1	Persentase Penduduk Miskin	Parametrik
t_1	Persentase Rumah Tangga yang Menggunakan Air Minum Bersih	Nonparametrik
t_2	Persentase Penduduk yang Mengalami Keluhan Kesehatan	Nonparametrik
t_3	Persentase Balita Mendapatkan Imunisasi Lengkap	Nonparametrik
t_4	MYS	Nonparametrik

4.2.3.2 Memodelkan AHH Indonesia Menggunakan Satu Titik Knot

Model regresi semiparametrik spline *truncated* dengan menggunakan satu titik knot dengan prediktor satu komponen parametrik dan empat komponen nonparametrik adalah sebagai berikut:

$$y = \beta_0 + \beta_1 x_1 + \gamma_{11} t_1 + \gamma_{21} (t_1 - K_{11})_+^1 + \gamma_{12} t_2 + \gamma_{22} (t_2 - K_{12})_+^1 + \gamma_{13} t_3 + \gamma_{23} (t_3 - K_{13})_+^1 + \gamma_{14} t_4 + \gamma_{24} (t_4 - K_{14})_+^1 + \varepsilon$$

Nilai GCV yang dihasilkan dengan menggunakan regresi semiparametrik spline *truncated* dengan satu knot disajikan dalam Tabel 4. 6. Adapun *Syntax* program R untuk mengolah dengan satu titik knot dapat dilihat pada Lampiran 2.

Tabel 4.6 Nilai GCV dengan Satu Titik Knot

Knot				GCV
t_1	t_2	t_3	t_4	
42.15	17.18	41.09	6.09	6.12
43.22	17.64	41.82	6.18	6.10
44.28	18.11	42.54	6.28	6.01
56.03	23.24	50.55	7.34	5.41
57.1	23.71	51.27	7.43	5.35
58.16	24.18	52	7.53	5.25
59.23	24.64	52.73	7.62	5.18
60.3	25.11	53.46	7.72	5.11
61.37	25.58	54.18	7.82	5.04
62.44	26.04	54.91	7.91	5.08
63.5	26.51	55.64	8.01	5.14
64.57	26.98	56.37	8.1	5.20

Berdasarkan Tabel 4.6 nilai GCV minimum yang dihasilkan adalah sebesar 5,04. Lokasi titik knot pada variabel air minum bersih (t_1) yaitu 61,37 (K_{11}), variabel Keluhan kesehatan (t_2) yaitu 25,58 (K_{12}), variabel imunisasi (t_3) yaitu 54,18 (K_{13}) dan variabel MYS (t_4) yaitu 7,82 (K_{14}).

4.2.3.3 Memodelkan AHH Indonesia Menggunakan Dua Titik Knot

Model regresi semiparametrik spline *truncated* dengan menggunakan dua titik knot dengan prediktor satu komponen parametrik dan empat komponen nonparametrik adalah sebagai berikut:

$$y = \beta_0 + \beta_1 x_1 + \gamma_{11} t_1 + \gamma_{21} (t_1 - K_{11})_+^1 + \gamma_{31} (t_1 - K_{21})_+^1 + \gamma_{12} t_2 + \gamma_{22} (t_2 - K_{12})_+^1 + \gamma_{32} (t_2 - K_{22})_+^1 + \gamma_{13} t_3 + \gamma_{23} (t_3 - K_{13})_+^1 + \gamma_{33} (t_3 - K_{23})_+^1 + \gamma_{14} t_4 + \gamma_{24} (t_4 - K_{14})_+^1 + \gamma_{34} (t_4 - K_{24})_+^1 + \varepsilon$$

Nilai GCV yang dihasilkan dengan menggunakan regresi semiparametrik spline *truncated* dengan dua knot disajikan dalam Tabel 4.7. Adapun *syntax* program R untuk mengolah dengan dua titik knot dapat dilihat pada Lampiran 3.

Tabel 4.7 Nilai GCV dengan Dua Titik Knot

Knot				GCV
t_1	t_2	t_3	t_4	
45.35	18.58	43.27	6.37	4.93
52.83	21.84	48.36	7.05	
45.35	18.58	43.27	6.37	5.05
53.89	22.31	49.09	7.14	
47.49	19.51	44.73	6.57	6.26
50.69	20.91	46.91	6.86	
47.49	19.51	44.73	6.57	5.85
51.76	21.38	47.64	6.95	
47.49	19.51	44.73	6.57	4.76
52.83	21.84	48.36	7.05	
47.49	19.51	44.73	6.57	5.00
53.89	22.31	49.09	7.14	
48.55	19.98	45.45	6.66	4.81
52.83	21.84	48.36	7.05	
52.83	21.84	48.36	7.05	7.41
70.98	29.78	60.73	8.68	

Berdasarkan Tabel 4.7 nilai GCV minimum yang dihasilkan adalah sebesar 4,76. Lokasi titik knot pada variabel air minum bersih (t_1) yaitu 47,49 (K_{11}) dan 52,83 (K_{21}), variabel keluhan kesehatan (t_2) yaitu 19,51 (K_{12}) dan 21,84 (K_{22}), variabel imunisasi (t_3) yaitu 44,73 (K_{13}) dan 48,36 (K_{23}) serta variabel MYS (t_4) yaitu 6,57 (K_{14}) dan 7,05 (K_{24}).

4.2.3.4 Memodelkan AHH Indonesia Menggunakan Tiga Titik Knot

Model regresi semiparametrik spline *truncated* dengan menggunakan tiga titik knot dengan prediktor satu komponen parametrik dan empat komponen nonparametrik adalah sebagai berikut:

$$\begin{aligned}
 y = & \beta_0 + \beta_1 x_1 + \gamma_{11} t_1 + \gamma_{21} (t_1 - K_{11})_+^1 + \gamma_{31} (t_1 - K_{21})_+^1 + \gamma_{41} (t_1 - K_{31})_+^1 + \\
 & \gamma_{12} t_2 + \gamma_{22} (t_2 - K_{12})_+^1 + \gamma_{32} (t_2 - K_{22})_+^1 + \gamma_{42} (t_2 - K_{32})_+^1 + \\
 & \gamma_{13} t_3 + \gamma_{23} (t_3 - K_{13})_+^1 + \gamma_{33} (t_3 - K_{23})_+^1 + \gamma_{43} (t_3 - K_{33})_+^1 + \gamma_{14} t_4 + \\
 & \gamma_{24} (t_4 - K_{14})_+^1 + \gamma_{34} (t_4 - K_{24})_+^1 + \gamma_{44} (t_4 - K_{34})_+^1 + \varepsilon.
 \end{aligned}$$

Tabel 4.8 Nilai GCV dengan Tiga Titik Knot

Knot				GCV
t_1	t_2	t_3	t_4	
60.30	25.11	53.46	7.72	6.58
81.65	34.45	68.01	9.64	
90.20	38.18	73.83	10.41	
60.30	25.11	53.46	7.72	6.56
81.65	34.45	68.01	9.64	
91.26	38.65	74.55	10.51	
60.30	25.11	53.46	7.72	6.57
81.65	34.45	68.01	9.64	
92.33	39.11	75.28	10.60	
60.30	25.11	53.46	7.72	4.19
82.72	34.91	68.73	9.74	
83.79	35.38	69.46	9.83	
60.30	25.11	53.46	7.72	5.28
82.72	34.91	68.73	9.74	
84.86	35.85	70.19	9.93	
60.30	25.11	53.46	7.72	5.62
82.72	34.91	68.73	9.74	
85.93	36.31	70.92	10.03	

Nilai GCV yang dihasilkan dengan menggunakan regresi semiparametrik spline *truncated* dengan tiga knot disajikan dalam Tabel 4.8. Berdasarkan Tabel 4.8 nilai GCV minimum yang dihasilkan adalah sebesar 4,19. Lokasi titik knot pada variabel air minum bersih (t_1) yaitu 60,30 (K_{11}), 82,72 (K_{21}) dan 83,79 (K_{31}), variabel keluhan kesehatan (t_2) yaitu 25,11 (K_{12}), 34,91 (K_{22}) dan 35,38 (K_{32}), variabel imunisasi (t_3) yaitu 53,46 (K_{13}), 68,73 (K_{23}) dan 69,46 (K_{33}) serta variabel MYS (t_4) yaitu 7,72 (K_{14}), 9,74 (K_{24}) dan 9,83 (K_{34}).

4.2.3.5 Memodelkan AHH Indonesia Menggunakan Kombinasi Titik Knot

Model regresi semiparametrik spline *truncated* dengan menggunakan kombinasi titik knot dengan prediktor satu komponen parametrik dan empat komponen nonparametrik tergantung dari kombinasi titik knot yang terbentuk. Nilai GCV yang dihasilkan dengan menggunakan regresi semiparametrik spline *truncated* dengan kombinasi titik knot disajikan dalam Tabel 4.9. *Syntax* program R untuk mengolah dengan dua titik knot dapat dilihat pada Lampiran 4.

Tabel 4.9 Nilai GCV dengan Kombinasi Titik Knot

Kombinasi	Knot				GCV
	t_1	t_2	t_3	t_4	
(1,2,3,2)	61,37	19,51	53,46	6,57	4,88
		21,84	68,73	7,05	
			69,46		
(2,2,3,2)	47,49	19,51	53,46	6,57	4,83
	52,83	21,84	68,73	7,05	
			69,46		
(3,1,2,2)	60,3	25,58	44,73	6,57	4,47
	82,72		48,36	7,05	
	83,79				
(3,1,2,3)	60,3	25,1	44,73	7,72	4,48
	82,72		48,36	9,74	
	83,79			9,83	

Berdasarkan Tabel 4. 9 nilai GCV minimum yang dihasilkan adalah sebesar 4,47 yaitu pada kombinasi knot (3,1,2,2). Lokasi titik knot pada variabel persentase rumah tangga yang menggunakan air minum bersih (t_1) yaitu 60,30 (K_{11}), 82,72 (K_{21}) dan 83,79 (K_{31}), variabel persentase penduduk yang mengalami keluhan kesehatan (t_2) yaitu 25,58 (K_{12}), variabel persentase balita yang mendapat imunisasi lengkap (t_3) yaitu 44,73 (K_{13}) dan 48,36 (K_{23}) serta variabel MYS (t_4) yaitu 6,57 (K_{14}) dan 7,05 (K_{24}).

4.2.3.6 Memilih Titik Knot Optimum dengan Menggunakan GCV

Setelah didapatkan model regresi semiparametrik spline *truncated* dengan satu titik knot, dua titik knot, dan tiga titik knot, maka langkah selanjutnya adalah memilih model terbaik dengan membandingkan nilai GCV pada masing-masing model. Nilai GCV minimum pada masing-masing model disajikan pada tabel 4.10.

Tabel 4.10 Nilai GCV Minimum pada Tiap Model

Banyaknya Knot	GCV Minimum
Satu titik knot	5,04
Dua titik knot	4,76
Tiga titik knot	4,19
Kombinasi titik knot	4,47

Dari Tabel 4.10 terlihat bahwa GCV minimum terdapat pada tiga titik knot dengan nilai GCV sebesar 4,19. Model regresi semiparametrik spline *truncated* terbaik yaitu dengan tiga titik knot pada variabel air minum bersih, variabel keluhan kesehatan, variabel imunisasi dan variabel MYS. Sehingga dapat dituliskan bahwa model terbaik adalah dengan menggunakan titik knot optimum (3,3,3,3).

4.2.3.7 Model Regresi Semiparametrik Spline *Truncated* dengan Titik Knot Optimum

Titik knot optimum terbaik yang terpilih adalah tiga titik knot dengan titik knot optimum (3,3,3,3). Dengan demikian model regresi semiparametrik spline *truncated* yang terbentuk adalah sebagai berikut:

$$y = \beta_0 + \beta_1 x_1 + \gamma_{11} t_1 + \gamma_{21} (t_1 - K_{11})_+^1 + \gamma_{31} (t_1 - K_{21})_+^1 + \gamma_{41} (t_1 - K_{31})_+^1 + \gamma_{12} t_2 + \gamma_{22} (t_2 - K_{12})_+^1 + \gamma_{32} (t_2 - K_{22})_+^1 + \gamma_{42} (t_2 - K_{32})_+^1 + \gamma_{13} t_3 + \gamma_{23} (t_3 - K_{13})_+^1 + \gamma_{33} (t_3 - K_{23})_+^1 + \gamma_{43} (t_3 - K_{33})_+^1 + \gamma_{14} t_4 + \gamma_{24} (t_4 - K_{14})_+^1 + \gamma_{34} (t_4 - K_{24})_+^1 + \gamma_{44} (t_4 - K_{34})_+^1 + \varepsilon$$

Dari hasil estimasi parameter dan titik knot optimum yang diperoleh maka model yang terbentuk adalah sebagai berikut.

$$\hat{y} = 68,65 - 0,23x_1 - 0,12t_1 + 0,30(t_1 - 60,30)_+ - 1,81(t_1 - 82,72)_+ + 0,60(t_1 - 83,79)_+ + 0,06t_2 - 0,33(t_2 - 25,11)_+ + 17,09(t_2 - 34,91)_+ - 19,66(t_2 - 35,38)_+ - 0,36t_3 + 0,57(t_3 - 53,46)_+ - 6,94(t_3 - 68,73)_+ + 8,93(t_3 - 69,46)_+ + 3,63t_4 - 6,19(t_4 - 7,72)_+ + 9,42(t_4 - 9,74)_+ + 8,48(t_4 - 9,83)_+$$

4.2.3.8 Pengujian Signifikansi Parameter secara Simultan

Pengujian hipotesis untuk menguji signifikansi parameter secara simultan menggunakan hipotesis sebagai berikut:

$$H_0 : \beta_h = \gamma_{jl} = \gamma_{(m+k)l} = 0$$

$$H_1 : \text{paling sedikit terdapat satu } \beta_h \neq 0 \text{ atau } \gamma_{jl} \neq 0 \text{ atau } \gamma_{(m+k)l} \neq 0$$

Dimana $h = 1; j = 1; l = 1,2,3,4; m = 1; k = 1,2,3$

Tabel 4.11 ANOVA Hasil Regresi Semiparametrik Spline *Truncated*

Sumber Variasi	Derajat Kebebasan	Jumlah Kuadrat	Rata-rata Kuadrat	F _{hitung}	p-value
Regresi	17	197,10	11,59	5,21	0,000
Error	16	35,60	2,22		
Total	33	232,70			

Hasil ANOVA pada pengujian hipotesis secara simultan tersaji pada Tabel 4.11. *Syntax* program R untuk pengujian signifikansi parameter secara simultan dapat dilihat pada Lampiran 6. Berdasarkan Tabel 4.11 terlihat bahwa nilai F_{hitung} sebesar 5,21 lebih besar dari nilai F_(0,05,17,16) yaitu 2,32 dan nilai p-value sebesar 0,00 lebih kecil dari nilai α (0,05), sehingga diperoleh keputusan tolak H₀. Maka dapat disimpulkan bahwa paling sedikit terdapat satu variabel prediktor yang signifikan dalam model. Model regresi semiparametrik spline *truncated* dengan tiga titik knot ini memiliki koefisien determinasi (R²) sebesar 84,70 persen. Hal ini menunjukkan bahwa variasi dari variabel respon dapat dijelaskan oleh kelima variabel prediktor sebesar 84,70 persen, sedangkan sisanya 15,30 persen dijelaskan oleh faktor lain.

4.2.3.9 Pengujian Signifikansi Parameter secara Parsial

Untuk mengetahui parameter mana saja yang berpengaruh secara signifikan terhadap model maka dilakukan pengujian hipotesis secara parsial. Pengujian signifikansi parameter secara parsial menggunakan hipotesis sebagai berikut:

$$H_0 : \beta_h = 0 ; H_1 : \beta_h \neq 0 ; \text{dimana } h = 0,1$$

$$H_0 : \gamma_{jl} = 0 ; H_1 : \gamma_{jl} \neq 0 ; \text{dimana } j = 1 ; l = 1,2,3,4$$

$$H_0 : \gamma_{(m+k)l} = 0 ; H_1 : \gamma_{(m+k)l} \neq 0 ; \text{dimana } m = 1 ; k = 1,2,3 ; l = 1,2,3,4$$

Hasil pengujian signifikansi parameter secara parsial tersaji pada Tabel 4.12.

Tabel 4.12 Hasil Uji Hipotesis secara Parsial

Variabel	Parameter	Estimator	W^*	$P\text{-value}$
Konstan	β_0	68,65	49,74	0,00
x_1	β_1	-0,23	10,14	0,01
t_1	γ_{11}	-0,12	2,05	0,17
	γ_{21}	0,30	5,15	0,04
	γ_{31}	-1,81	1,56	0,23
	γ_{41}	0,60	0,13	0,72
t_2	γ_{12}	0,06	0,14	0,71
	γ_{22}	-0,33	1,58	0,23
	γ_{32}	17,09	12,63	0,00
	γ_{42}	-19,66	12,51	0,00
t_3	γ_{13}	-0,36	8,94	0,01
	γ_{23}	0,57	9,31	0,01
	γ_{33}	-6,94	3,88	0,06
	γ_{43}	8,93	4,64	0,05
t_4	γ_{14}	3,63	10,27	0,00
	γ_{24}	-6,19	10,26	0,00
	γ_{34}	9,42	13,75	0,00
	γ_{44}	8,48	13,75	0,00

Berdasarkan Tabel 4.12 terlihat bahwa dari 18 parameter terdapat 12 parameter yang signifikan dan 6 parameter yang tidak signifikan. Semua variabel prediktor berpengaruh dalam model yaitu persentase penduduk miskin, persentase rumah tangga yang menggunakan air minum bersih, persentase penduduk yang

mengalami keluhan kesehatan, persentase balita yang mendapat imunisasi lengkap, dan MYS.

4.2.3.10 Pengujian Asumsi Residual

Residual pada regresi semiparametrik spline *truncated* harus memenuhi asumsi identik, independen dan berdistribusi normal (IIDN). Pemeriksaan asumsi residual ini dilakukan untuk mengetahui apakah residual yang terbentuk telah memenuhi asumsi IIDN. Pengujian asumsi residual identik menggunakan uji *glejser*. Pengujian asumsi residual independen menggunakan uji *Durbin-Watson*. Sedangkan pengujian asumsi normal menggunakan uji *Kolmogorov-Smirnov*.

4.2.3.10.1 Pengujian Asumsi Identik

Pengujian asumsi identik dimaksudkan untuk mendeteksi varians dari residualnya memiliki varians yang homogen. Pemeriksaan menggunakan uji *glejser* dan hipotesis yang digunakan adalah sebagai berikut:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1 : \text{paling sedikit terdapat satu } \sigma_i^2 \neq \sigma^2 ; \text{ dimana } i = 1, 2, \dots, n$$

Hasil pengujian *glejser* tersaji pada Tabel 4.13. *Syntax* program R untuk Uji *Glejser* dapat dilihat pada Lampiran 7.

Tabel 4.13 Hasil Uji *Glejser*

Sumber Variasi	Derajat Kebebasan	Jumlah Kuadrat	Rata-rata Kuadrat	F _{hitung}	p-value
Regresi	17	7,50	0,44	1,13	0,40
Error	16	6,24	0,39		
Total	33	13,74			

Dari Tabel 4.13 terlihat bahwa nilai *p-value* sebesar 0,40 lebih besar dari nilai α yang ditetapkan sehingga keputusan yang diambil adalah gagal tolak H_0 . Hal ini memberikan kesimpulan bahwa varians dari residual adalah homogen, sehingga asumsi identik pada model regresi semiparametrik spline *truncated* yang terbentuk telah terpenuhi.

4.2.3.10.2 Pengujian Asumsi Independen

Pengujian asumsi independen dimaksudkan untuk mendeteksi korelasi antar residual. Pemeriksaan dilakukan dengan menggunakan uji *Durbin-Watson*.

Hipotesis yang digunakan adalah sebagai berikut:

$$H_0: \rho = 0 \text{ (tidak ada korelasi antara residual)}$$

$$H_1: \rho \neq 0 \text{ (ada korelasi antara residual)}$$

dengan statistik uji sebagai berikut :

$$d_{hitung} = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}$$

Keputusan tolak H_0 jika $d_{hitung} < d_{L, \frac{\alpha}{2}}$, dimana $d_{L, \frac{0,05}{2}} = 1,14$. Hasil penghitungan diperoleh $d_{hitung} = 2,33$, sehingga keputusannya adalah gagal tolak H_0 . Hal ini memberikan kesimpulan bahwa antar residual tidak terjadi korelasi. Dengan demikian asumsi independen pada model regresi semiparametrik spline *truncated* terpenuhi.

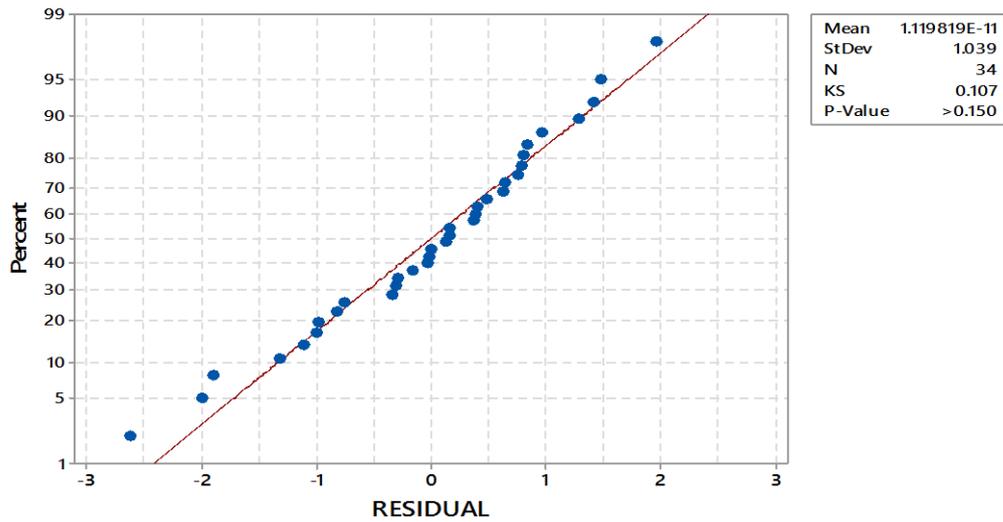
4.2.3.10.3 Pengujian Asumsi Distribusi Normal

Residual regresi semiparametrik spline *truncated* harus memenuhi asumsi distribusi normal. Oleh karenanya dilakukan pemeriksaan asumsi distribusi normal pada residual yang diperoleh. Pemeriksaan dilakukan dengan menggunakan uji *Kolmogorov-Smirnov*. Hipotesis yang digunakan adalah sebagai berikut:

$$H_0 : F_0(x) = F(x) \text{ (Residual berdistribusi Normal)}$$

$$H_1 : F_0(x) \neq F(x) \text{ (Residual tidak berdistribusi Normal).}$$

Hasil uji *Kolmogorov-Smirnov* disajikan dalam bentuk plot seperti terlihat pada gambar 4.4.



Gambar 4. 4 Probability Plot Residual

Plot residual seperti yang terlihat pada gambar 4.8 menghasilkan p -value sebesar 0,150 yang lebih besar dari nilai α yang ditentukan yaitu 0,05, sehingga keputusan yang diambil adalah gagal tolak H_0 . Oleh karenanya dapat disimpulkan bahwa residual berdistribusi normal. Dari hasil uji ini maka regresi semiparametrik spline *truncated* yang dihasilkan memenuhi asumsi berdistribusi normal.

4.2.4 Perbandingan Model dan Interpretasi

Perbandingan model antara metode regresi linier berganda dan regresi semiparametrik spline *truncated* berdasarkan koefisien determinasi (R^2) dan menggunakan koefisien determinasi yang disesuaikan (R^2 *adjusted*). Penggunaan R^2 *adjusted* karena antara regresi linier berganda dengan regresi semiparametrik spline *truncated* memiliki jumlah parameter yang berbeda. Berikut ini adalah nilai R^2 dan R^2 *adjusted* pada masing-masing model.

Tabel 4.14 Nilai R^2 dan R^2 *adjusted*

Model	R^2 (%)	R^2 <i>adjusted</i> (%)
Regresi Linier Berganda	48,10	39,80
Regresi Semiparametrik Spline <i>Truncated</i>	84,70	68,45

Pada Tabel 4.14 terlihat bahwa nilai R^2 maupun R^2 *adjusted* pada model regresi semiparametrik spline *truncated* menghasilkan nilai yang lebih besar yaitu berturut-turut adalah 84,70 persen dan 68,45 persen. Oleh karena itu metode yang dipilih adalah metode regresi semiparametrik spline *truncated*.

Model regresi semiparametrik spline *truncated* dengan tiga titik knot telah memenuhi asumsi residual IIDN, maka model yang dihasilkan dapat diinterpretasikan lebih lanjut. Model regresi semiparametrik spline *truncated* terbaik yang dihasilkan adalah sebagai berikut:

$$\begin{aligned} \hat{y} = & 68,65 - 0,23x_1 - 0,12t_1 + 0,30(t_1 - 60,30)_+ - 1,81(t_1 - 82,72)_+ + \\ & 0,60(t_1 - 83,79)_+ + 0,06t_2 - 0,33(t_2 - 25,11)_+ + 17,09(t_2 - 34,91)_+ - \\ & 19,66(t_2 - 35,38)_+ - 0,36t_3 + 0,57(t_3 - 53,46)_+ - 6,94(t_3 - 68,73)_+ + \\ & 8,93(t_3 - 69,46)_+ + 3,63t_4 - 6,19(t_4 - 7,72)_+ + 9,42(t_4 - 9,74)_+ + \\ & 8,48(t_4 - 9,83)_+ \end{aligned}$$

Interpretasi model terhadap variabel-variabel yang berpengaruh secara signifikan adalah sebagai berikut:

1. Interpretasi terhadap persentase penduduk miskin (x_1) dengan asumsi variabel lain konstan adalah sebagai berikut:

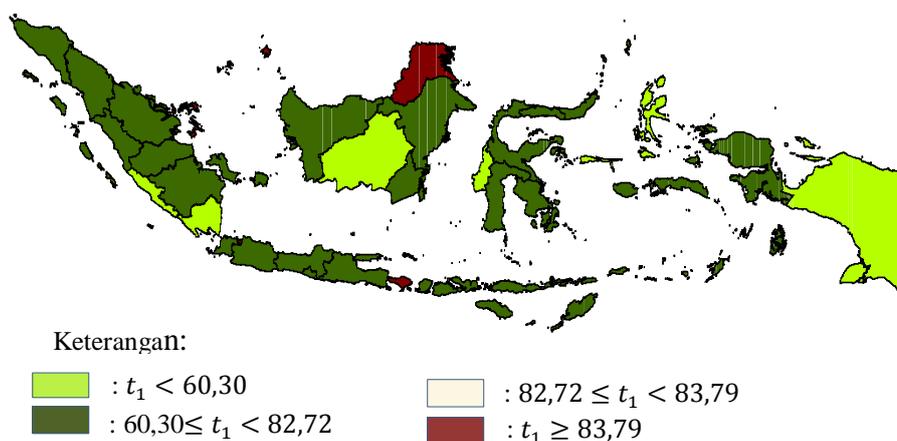
$$\hat{y} = 68,65 - 0,23x_1$$

Jika terjadi kenaikan persentase penduduk miskin sebanyak satu persen maka nilai AHH akan turun sebesar 0,23 persen. Hasil pemodelan tersebut sesuai dengan hipotesis awal, yaitu semakin bertambah persentase penduduk miskin maka AHH akan semakin menurun. Dengan indikasi arah hubungan antara persentase penduduk miskin dengan AHH adalah negatif. Penekanan terhadap persentase penduduk miskin secara umum dapat meningkatkan AHH di Indonesia.

2. Interpretasi terhadap persentase rumah tangga yang menggunakan air minum bersih (t_1) dengan asumsi variabel lain konstan adalah sebagai berikut.

$$\hat{y} = \begin{cases} -0,12t_1 & , t_1 < 60,30 \\ 0,18t_1 - 18,27 & , 60,30 \leq t_1 < 82,72 \\ -1,63t_1 + 131,61 & , 82,72 \leq t_1 < 83,79 \\ -1,03t_1 + 81,54 & , t_1 \geq 83,79 \end{cases}$$

Dengan regresi semiparametrik spline *truncated* maka variabel yang mempunyai pola spline *truncated* dapat diinterpretasikan sesuai dengan pola perubahan datanya atau pada setiap interval knotnya. Untuk nilai persentase rumah tangga yang menggunakan air minum bersih antara 60,30 sampai 82,7 persen, setiap kenaikan satu persen maka AHH akan mengalami kenaikan sebesar 0,18 tahun.



Gambar 4. 5 Peta AHH terhadap Persentase Rumah Tangga yang Menggunakan Air Minum Bersih

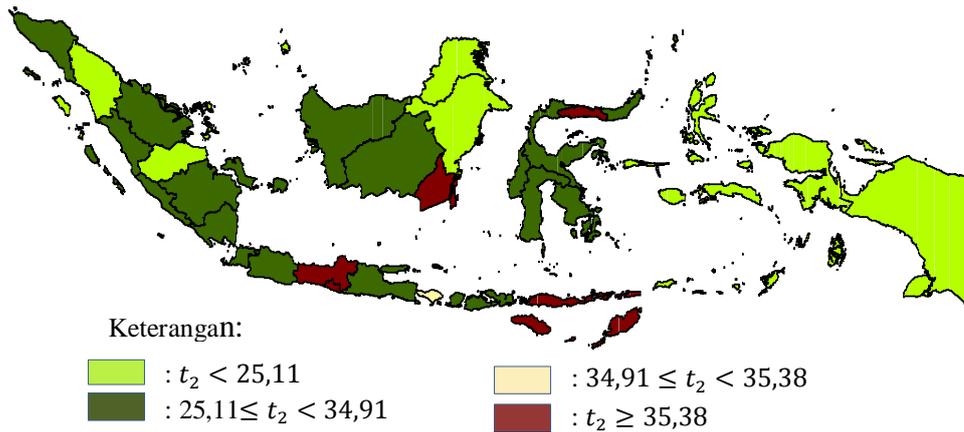
Hipotesis awal tentang persentase rumah tangga yang menggunakan air minum bersih terhadap AHH adalah positif, artinya jika persentase rumah tangga yang menggunakan air minum bersih meningkat maka AHH juga akan mengalami kenaikan. Dari hasil pemodelan menunjukkan bahwa 70,59 persen provinsi di Indonesia memiliki pola seperti ini. Provinsi tersebut adalah Provinsi Aceh, Sulawesi Tengah, Kalimantan Selatan, Nusa Tenggara Timur, Jambi, Maluku, Sumatera Selatan, Gorontalo, Sumatera Barat, Jawa Barat, Banten, Kepulauan Bangka Belitung, Kalimantan Barat, Papua Barat, Sumatera Utara, Nusa Tenggara Barat, Sulawesi Selatan, Jawa Tengah, Sulawesi Tenggara, Kalimantan Timur, Jawa Timur, DIY, dan Papua.

3. Interpretasi terhadap persentase penduduk yang mengalami keluhan kesehatan (t_2) dengan asumsi variabel lain konstan adalah sebagai berikut:

$$\hat{y} = \begin{cases} 0,06t_2 & , t_2 < 25,11 \\ -0,27t_2 + 8,41 & , 25,11 \leq t_2 < 34,91 \\ 16,82t_2 - 588,26 & , 34,91 \leq t_2 < 35,38 \\ -2,84t_2 + 107,31 & , t_2 \geq 35,38 \end{cases}$$

Pada variabel persentase penduduk yang mengalami keluhan kesehatan terdapat empat sub interval perubahan perilaku. Untuk persentase penduduk yang mengalami keluhan kesehatan antara 25,11 persen sampai 34,91 persen, setiap kenaikan satu persen penduduk yang mengalami keluhan kesehatan akan menyebabkan AHH mengalami penurunan sebesar 0,27 tahun. Sedangkan untuk nilai persentase penduduk yang mengalami keluhan kesehatan lebih dari 35,38 persen, maka setiap kenaikan satu persen nilai persentase penduduk yang mengalami keluhan kesehatan akan menyebabkan AHH mengalami penurunan sebesar 2,84 tahun. Daerah yang mempunyai pola perilaku seperti ini adalah Provinsi Jawa Tengah, Nusa Tenggara Timur, Gorontalo, Kalimantan Selatan, dan DIY.

Hipotesis awal tentang persentase penduduk yang mengalami keluhan kesehatan terhadap AHH adalah semakin banyak penduduk yang mengalami keluhan kesehatan dengan kata lain tingkat morbiditas tinggi akan memperbesar peluang terjadinya kematian dan berdampak pada AHH menurun. Dengan indikasi tersebut maka arah hubungan antara persentase penduduk yang mengalami keluhan kesehatan dengan AHH adalah negatif. Jika persentase penduduk yang mengalami keluhan kesehatan naik maka AHH cenderung turun. Berdasarkan model yang terbentuk, hampir seluruh daerah di Indonesia yaitu 70,58 persen provinsi di Indonesia mempunyai pola seperti ini.

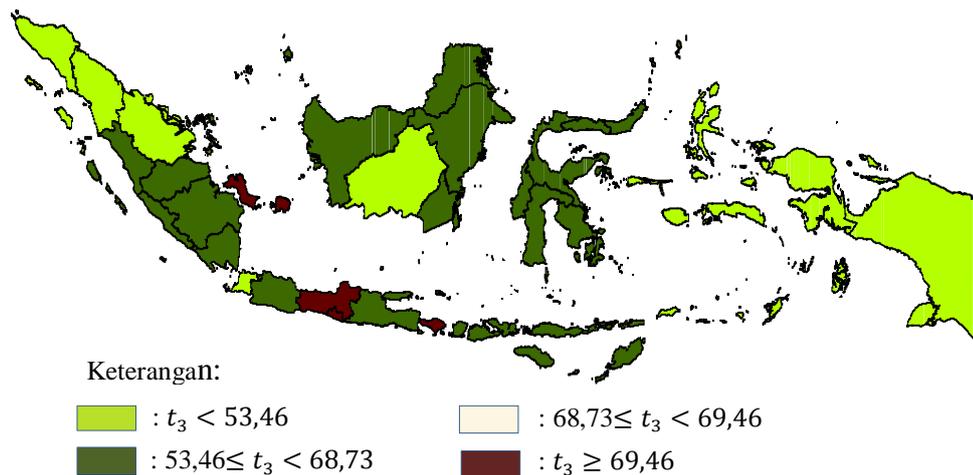


Gambar 4.6. Peta AHH terhadap Persentase Penduduk yang Mengalami Keluhan Kesehatan

4. Interpretasi terhadap persentase balita yang mendapat imunisasi lengkap dengan asumsi variabel yang lain konstan adalah sebagai berikut:

$$\hat{y} = \begin{cases} -0,36t_3 & , t_3 < 53,46 \\ 0,21t_3 - 14,22 & , 53,46 \leq t_3 < 68,73 \\ -6,73t_3 + 227,92 & , 68,73 \leq t_3 < 69,46 \\ 2,2t_3 - 88,02 & , t_3 \geq 69,46 \end{cases}$$

Pada variabel persentase balita yang mendapat imunisasi lengkap terdapat empat sub interval perubahan perilaku. Untuk nilai persentase balita yang mendapat imunisasi lengkap diantara 53,46 persen sampai 68,73 persen, maka jika terjadi kenaikan persentase balita yang mendapat imunisasi lengkap sebesar satu persen, nilai AHH akan naik 0,21 tahun. Terdapat 21 provinsi di Indonesia yang mempunyai perilaku seperti ini. Sedangkan untuk persentase balita yang mendapat imunisasi lengkap di atas 69,46 persen, jika terjadi kenaikan persentase balita yang mendapat imunisasi lengkap satu persen maka nilai AHH akan naik 2,2 tahun. Daerah yang mempunyai perilaku seperti ini adalah Provinsi Jawa Tengah, Bangka Belitung, Bali, dan DIY.



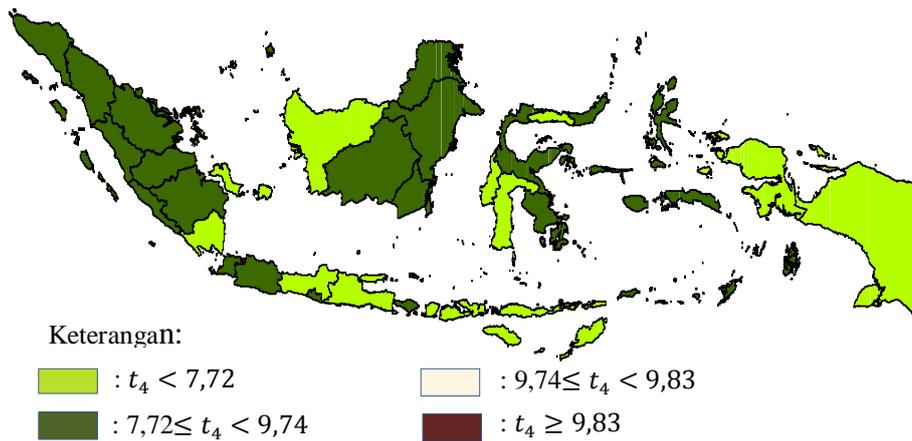
Gambar 4. 7. Peta AHH terhadap Persentase Balita yang Mendapat Imunisasi Lengkap

Hipotesis awal mengenai persentase balita yang mendapat imunisasi lengkap terhadap AHH adalah semakin besar persentase balita yang mendapat imunisasi lengkap maka AHH akan semakin meningkat. Dengan indikasi arah hubungan antara persentase balita yang mendapat imunisasi lengkap dan AHH adalah positif.

5. Interpretasi terhadap MYS dengan asumsi variabel yang lain konstan adalah sebagai berikut:

$$\hat{y} = \begin{cases} 3,63t_4 & , t_4 < 7,72 \\ -2,56t_4 + 47,79 & , 7,72 \leq t_4 < 9,74 \\ 6,86t_4 - 43,96 & , 9,74 \leq t_4 < 9,83 \\ 15,34t_4 - 127,32 & , t_4 \geq 9,83 \end{cases}$$

Variabel MYS mempunyai empat sub interval perilaku. Untuk nilai MYS dibawah 7,72 tahun maka perilakunya adalah jika nilai MYS naik satu tahun, AHH akan naik 3,63 tahun. Daerah yang mengalami pola perilaku seperti ini adalah Provinsi Papua, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Sulawesi Barat, Papua Barat, Jawa Tengah, Gorontalo, Jawa Timur, Bangka Belitung, Lampung, dan Sulawesi Selatan. Kebijakan untuk meningkatkan kualitas hidup seperti AHH melalui peningkatan pendidikan di provinsi-provinsi tersebut sangat tepat. Sementara itu ada satu provinsi yang memiliki pola setiap kenaikan 1 tahun MYS maka akan meningkatkan AHH sebesar 15,34 tahun, yaitu Provinsi DKI Jakarta.



Gambar 4. 8. Peta AHH terhadap MYS

Hipotesis awal tentang MYS terhadap AHH adalah semakin tinggi MYS maka AHH juga akan naik. Dengan indikasi arah hubungan antara MYS dengan AHH adalah positif.

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berikut ini adalah kesimpulan yang didapatkan berdasarkan dari hasil pembahasan pada Bab 4.

1. Pengujian hipotesis secara parsial pada regresi semiparametrik spline *truncated* dengan hipotesis adalah $H_0: \tilde{c}'_j \tilde{\delta} = 0$ dan $H_1: \tilde{c}'_j \tilde{\delta} \neq 0$ maka dapat disimpulkan:

a. Estimasi parameter model regresi semiparametrik spline *truncated* pada ruang parameter di bawah hipotesis Ω adalah

$\hat{\delta}_\Omega = \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y}$ dan estimasi parameter pada ruang parameter di bawah hipotesis nol ω adalah

$$\hat{\delta}_\omega = \hat{\delta}_\Omega - \left(\mathbf{Z}'(\mathbf{X}, \mathbf{Z}, \tilde{k}) \mathbf{T}(\mathbf{X}, \mathbf{Z}, \tilde{k}) \right)^{-1} \tilde{c}'_j \left(\tilde{c}'_j \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right)^{-1} \tilde{c}'_j \hat{\delta}_\Omega$$

b. Statistik uji untuk mengkaji hipotesis $H_0: \tilde{c}'_j \tilde{\delta} = 0$ melawan $H_1: \tilde{c}'_j \tilde{\delta} \neq 0$ diberikan oleh

$$W^* = \frac{Q_1/(1)}{Q_2/(n - (p + 1 + (r + m)q))}$$

dimana

$$Q_1 = \tilde{y}' \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j$$

$$\left[\tilde{c}'_j \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \tilde{c}_j \right]^{-1} \tilde{c}'_j \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \tilde{y}$$

$$Q_2 = \tilde{y}' \left[\mathbf{I} - \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \left(\mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \mathbf{Z}(\mathbf{X}, \mathbf{T}, \tilde{k}) \right)^{-1} \mathbf{Z}'(\mathbf{X}, \mathbf{T}, \tilde{k}) \right] \tilde{y}$$

Distribusi dari statistik uji yang diperoleh adalah

$$W^* = \frac{Q_1/(1)}{Q_2/(n - (p + 1 + (r + m)q))} \sim F_{((1), (n - (p + 1 + (r + m)q)))}$$

- c. Daerah penolakan H_0 adalah $C\{(x_1, x_2, \dots, x_p, t_1, t_2, \dots, t_q, y); W^* > k^*\}$, dimana k^* diperoleh dari persamaan $\alpha = P(W^* > k^* | \tilde{c}_j' \tilde{\delta} = 0)$ atau nilai k^* (F tabel) didapat dari persamaan $\int_{k^*}^{\infty} f(w) dw = \alpha$, dengan $f(w)$ adalah fungsi berdistribusi $F_{((1), (n-(p+1+(r+m)q))}$.
2. Aplikasi pada data AHH di Indonesia Tahun 2015 diperoleh hasil sebagai berikut.
- Model terbaik yang diperoleh adalah dengan menggunakan tiga titik knot.
 - Dari kelima variabel yang digunakan semua variabel memberikan pengaruh yang signifikan terhadap model.
 - Koefisien determinasi (R^2) yang diperoleh adalah sebesar 84,70 persen, sehingga model tersebut sudah layak untuk digunakan.
3. Berdasarkan model yang terbentuk, beberapa masukan untuk pemerintah adalah sebagai berikut.
- Pencapaian kestabilan dimensi kesehatan di Indonesia dapat dilakukan dengan program-program pengentasan kemiskinan di bidang kesehatan, seperti pemberian pelayanan kesehatan dengan biaya yang minim. Selain itu peningkatan pelayanan kesehatan dari segi kualitas dan kuantitas sangat diperlukan sehingga masyarakat dapat dengan mudah mendapatkan pelayanan kesehatan yang baik.
 - Pemerintah harus dapat menjamin ketersediaan air bersih untuk seluruh penduduk Indonesia terutama wilayah Indonesia bagian timur yang masih mengalami kelangkaan air bersih.
 - Upaya untuk mencapai indikator pendidikan di Indonesia dapat terus bertumbuh melalui peningkatan aksesibilitas dan kualitas pendidikan di semua kalangan masyarakat. Kemudahan masyarakat dalam mendapatkan pendidikan yang baik diharapkan dapat meningkatkan MYS dan berdampak pada kenaikan AHH di Indonesia.

5.2 Saran

Dari hasil analisis dan pembahasan yang dilakukan pada Bab 4 selain kesimpulan terdapat beberapa saran yang dapat digunakan untuk penelitian selanjutnya yaitu:

1. Melakukan analisis dengan analisis inferensi untuk kasus multi respon AHH laki-laki dan perempuan. Mengingat AHH perempuan selalu lebih tinggi dibanding AHH laki-laki maka perlu dilakukan kajian tentang faktor yang mempengaruhi AHH menurut jenis kelamin.
2. Mengembangkan regresi semiparametrik spline *truncated* dengan spatial, untuk mengetahui pengaruh perbedaan antar wilayah. Sehingga input terhadap pemerintah mengenai kebijakan suatu wilayah akan lebih tepat.
3. Mengembangkan regresi semiparametrik spline *truncated* non linier, untuk mengantisipasi bentuk pola hubungan antara variabel respon dan variabel prediktor yang mengikuti pola non linier.

“Halaman ini sengaja dikosongkan”

DAFTAR PUSTAKA

- Andriani, H. (2015), *Estimator Spline Terpinalti pada Regresi Nonparametrik (Studi Kasus: Angka Harapan Hidup di Jawa Timur)*, Tesis, Institut Teknologi Sepuluh Nopember, Surabaya.
- Antoniadis, A., Gregore, G., and Mckeague, W., Wavelet. (1994), "Methods for Curve Estimation", *JASA*, Vol.89, No.428, 1994, 1340-1353.
- Antoniadis, A. (2007), "Wavelet methods in statistics: Some recent developments and their applications", *Statistics Survey*, Vol. 1, 16–55.
- Ayuni, N.W.D. (2015), "Pemodelan Angka Harapan Hidup di Provinsi Jawa Timur dengan Menggunakan Metode Feed Forward Neural Network (FFNN)", *Jurnal Sosial dan Humaniora*, Vol. 5, No. 2.
- Badan Pusat Statistik. (2013), *Proyeksi Penduduk Indonesia 2010-2035*, BPS RI, Jakarta.
- _____. (2016a), *Statistik Kesejahteraan Rakyat 2015*, BPS RI, Jakarta.
- _____. (2016b), *Statistik Indonesia 2015*, BPS RI, Jakarta.
- _____. (2016c), *Indeks Pembangunan Manusia 2015*, BPS RI, Jakarta.
- Belon, A.P. (2011). "Reduction of social inequalities in life expectancy in a city of Southeastern Brazil", *International Journal for Equity in Health*, State of Campinas, Sao Paulo, Brazil.
- Bintariningrum, M.F. (2014), "Pemodelan Regresi Nonparametrik Spline *Truncated* dan Aplikasinya pada Angka Kelahiran Kasar di Surabaya", *Jurnal Sains dan Seni Promits*, Vol. 3, hal. D7-D12.
- Budiantara, I.N. (2001), "Estimasi Parametrik dan Nonparametrik Untuk Pendekatan Kurva Regresi", *Makalah Pembicara Utama Pada Seminar Nasional Statistika V*, Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember, Surabaya.
- _____. (2004), "Spline: Historis, Motivasi dan Perannya Dalam Regresi Nonparametrik", *Makalah Pembicara Utama pada Konferensi Nasional Matematika XII*, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana, Denpasar.

- _____. (2005), "Regresi Spline Linier", *Makalah Pembicara Utama pada Seminar Nasional Matematika*, FMIPA Universitas Diponegoro, Semarang.
- _____. (2006), "Model Spline Dengan Knots Optimal", *Jurnal Ilmu Dasar*, FMIPA Universitas Jember, Vol. 7, 77-85.
- _____. (2007), "Model Keluarga Spline Polinomial *Truncated* Dalam Regresi Semiparametrik", *Jurnal Matematika, Pengetahuan Alam dan Pengajarannya*, Vol. 36, hal. 1-16.
- _____. (2009), "Spline Dalam Regresi Nonparametrik dan Semiparametrik : Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang", *Pidato Pengukuhan untuk Jabatan Guru Besar*, Institut Teknologi Sepuluh Nopember, ITS Press, Surabaya.
- Casella, G dan Berger R.L. (2001), *Statistical Inference*, 2nd edition, Duxbury Press, California.
- Draper, N.R., Smith, H. (1992), *Applied Regression Analysis*, 2nd Edition, John Wiley & Sons, Chapman and Hall, New York.
- Engle, R.F., Granger, C.W.J., Rice, J, dan Weiss, A. (1986), "Semiparametric Estimates of the Relation Between Weather and Electricity", *Journal of the American Statistical Association*, Vol. 81, hal. 310-320.
- Eubank, R.L. (1988), *Spline Smoothing and Nonparametric Regression*, Marcel Dekker, Inc., New York.
- Eubank, R. (1999), *Nonparametric Regression and Spline Smoothing*, Marcel Dekker, New York.
- Gujarati, D.N. (2003), *Basic Econometrics*, 4th edititon, McGraw-Hill, New York.
- Halicioglu, F. (2011), "Modeling Life Expectancy in Turkey", *Journal Economic Modelling*, Vol. 28, 2075-2082.
- Hardle, W. (1990), "Aplied Nonparametrik Regression", *Cambridge University Press*, New York.
- Heckman, N. (1986). "Spline smoothing in partly linear models", *J. Roy. Statist. Soc. Ser. B*, 48, 244-248.
- Islamiyati, A., dan Budiantara, I.N. (2007), "Model Spline dengan Titik-Titik Knots dalam Regresi Non Parametrik", *Jurnal Inferensi*, 3, 11-21.

- Kementerian Kesehatan. (2010). *Indeks Pembangunan Kesehatan Masyarakat*.
Kemenkes RI, Jakarta.
- Firdial, L. (2010), *Pemodelan Angka Harapan Hidup di Provinsi Jawa Timur dan Jawa Tengah dengan Model Geographically Weighted Regression (GWR)*, Tesis, Institut Teknologi Sepuluh Nopember, Surabaya.
- Lee, T.C.M. (2002),” On Algorithms for Ordinary Least Squares Regression Spline Fitting : A Comparative Study”, *Journal Statist. Comput. Simul.*, Vol. 72 (8), pp. 647–663
- Mulianah. (2006). *Pendekatan Kernel dalam Regresi Semiparametrik dan Pemilihan Bandwith Optimal*, Tesis, Institut Teknologi Sepuluh Nopember, Surabaya.
- Purnomo, A.A.S.I. (2016), *Estimator Campuran Kernel dan Spline Truncated Linier Multivariabel dalam Regresi Nonparametrik*, Tesis, Institut Teknologi Sepuluh Nopember, Surabaya.
- Rencher, A.C. dan Scaalje, G.B., (2007), *Linear Models in Statistics*. 2nd Edition, John Wiley & Sons, New Jersey.
- Rory. (2016), *Regresi Campuran Nonparametrikj Spline Truncated dan Kernel dalam Data Kemiskinan di Papua*, Tesis, Institut Teknologi Sepuluh Nopember, Surabaya.
- Ruliana, Budiantara, I.N., Otok, B.W. dan Wibowo, W. (2016), "Simultaneous Hypothesis Testing of Spline Truncated Model in Nonlinear Structural Equation Modeling (SEM)", *Journal of Theoretical and Applied Information Technology*, Vol. 89, Hal. 371-380.
- Singh, K.G., Siahpush, M. (2006), “Widening Socioeconomic inequalities in US Life Expectancy, 1980-2000”, *International Journal of Epidemiology*, 35, 969-979.
- Solikhin.(2014),*Estimator Spline Truncated Linier dalam Regresi Semiparametrik Pada Data Persentase Pengeluaran Konsumsi Padi-Padian di Provinsi Jawa Tengah*, Tesis, Institut Teknologi Sepuluh Nopember, Surabaya.
- Sriliiana, I. (2012), *Regresi Spline Truncated Dalam Model Linear Parsial untuk Data Longitudinal*, Tesis, Institut Teknologi Sepuluh Nopember, Surabaya.

- Srinadi, I.A.M. (2002), *Estimator Spline pada Model Semiparametrik*, Tesis, Institut Teknologi Sepuluh Nopember, Surabaya.
- Sugiantari, A.P. (2013), “Analisis Faktor-faktor Yang Mempengaruhi Angka Harapan Hidup di Jawa Timur Menggunakan Regresi Semiparametrik Spline”, *Jurnal sains dan seni pomits Vol.2, No.1*, Institut Teknologi Sepuluh Nopember, Surabaya
- Tupen, S.N. (2011), *Uji Hipotesis Dalam Regresi Nonparametrik Spline*, Tesis, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember, Surabaya.
- Wahba, G. (1990), *Spline Models for Observasional Data*, University of Winconsin at Madison, Phensylvania.
- Zaki, A. (2007), *Pengujian Hipotesis dalam Model Spline pada Regresi Nonparametrik*, Tesis, Institut Teknologi Sepuluh Nopember, Surabaya.

LAMPIRAN

Lampiran 1 Data dan Struktur Data yang Digunakan

y	x	t_1	t_2	t_3	t_4
65.09	27.33	51.27	17.79	40.48	5.99
65.38	17.06	71.70	34.87	67.84	6.71
65.96	22.69	62.72	37.03	64.78	6.93
69.87	8.02	68.39	25.62	54.49	6.93
64.22	12.54	53.89	29.6	56.83	6.94
65.19	25.94	68.85	20.39	52.63	7.01
73.96	13.56	73.63	35.52	70.18	7.03
67.12	18.28	66.47	37.43	66.64	7.05
70.68	12.33	76.64	33.45	65.94	7.14
69.88	5.4	68.03	30.42	70.99	7.46
69.9	14.34	55.06	31.42	63.22	7.56
69.8	9.37	72.07	26.3	60.39	7.64
67.8	4.98	62.23	39.27	64.68	7.76
69.14	14.24	65.16	29.41	57.63	7.77
72.41	9.5	67.20	28.11	58.04	7.86
70.56	8.85	62.75	24.45	55.38	7.96
67.26	14.68	61.49	29.21	53.93	7.97
69.54	5.93	57.01	25.39	53.29	8.03
70.44	12.9	77.19	25.86	60.41	8.18
71.35	4.74	91.27	35.29	72.03	8.26
69.43	5.89	67.68	30.34	50.91	8.27
68.5	17.84	41.08	27.96	58.29	8.29
72.16	6.21	84.59	23.65	63.17	8.36
67.44	6.89	60.07	16.71	52.39	8.37
68.66	7.31	66.58	28.92	54.76	8.42
70.93	8.39	74.24	29.71	49.30	8.49
69.5	17.05	61.23	27.92	40.36	8.77
70.99	8.65	71.53	27.31	63.41	8.88
74.68	14.97	80.99	39.58	76.01	9.00
68.29	10.51	71.41	23.04	46.44	9.03
73.65	6.22	78.13	21.98	66.03	9.15
65.31	19.5	64.96	17.59	48.30	9.16
69.41	6.22	84.12	21.28	65.81	9.65
72.43	3.93	93.40	33.39	66.82	10.70

Keterangan:

y : AHH (tahun)

x_1 : Persentase penduduk miskin (persen)

t_1 : Rumah tangga yang menggunakan air minum bersih (persen)

t_2 : Penduduk yang mengalami keluhan kesehatan (persen)

t_3 : Balita yang mendapatkan imunisasi lengkap (persen)

t_4 : MYS (tahun)

Lampiran 2 Program Regresi Semiparametrik Spline *Truncated* Satu Titik Knot

```

GCV1=function(para)
{
data=read.csv("D:// data.csv",header=T)
data=as.matrix(data)
p=length(data[,1])
q=length(data[1,])
m=ncol(data)-para-1
dataA=data[(para+2):q]
F=matrix(0,nrow=p,ncol=p)
diag(F)=1
nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
knot1=matrix(ncol=m,nrow=nk)
for (i in (1:m))
{
for (j in (1:nk))
{
a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
knot1[j,i]=a[j]
}
}
a1=length(knot1[,1])
knot1=knot1[2:(a1-1),]
aa=rep(1,p)
data1=matrix(ncol=m,nrow=p)
data2=data[,2:q]
a2=nrow(knot1)
GCV=rep(NA,a2)
Rsq=rep(NA,a2)
for (i in 1:a2)
{
for (j in 1:m)
{
for (k in 1:p)
{if(data[k,(j+para+1)]<knot1[i,j])data1[k,j]=0
else
data1[k,j]=data[k,(j+para+1)]-knot1[i,j]
}
}
}
mx=cbind(aa,data2,data1)
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0

```

Lampiran 2 Program Regresi Semiparametrik Spline *Truncated* Satu Titik Knot
(Lanjutan)

```

for (r in (1:p))
{
sum=(data[r,1]-yhat[r,])^2
sum1=(yhat[r,]-mean(data[,1]))^2
SSE=SSE+sum
SSR=SSR+sum1
}
Rsqr[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsqr=as.matrix(Rsqr)
cat("=====", "\n")
cat("Nilai Knot dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
print (knot1)
cat("=====", "\n")
cat("Rsqr dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
print (Rsqr)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
print(max(Rsqr))
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
cat(" GCV =", s1, "\n")
write.csv(GCV, file="D://output GCV1.csv")
write.csv(Rsqr, file="D:// output Rsqr1.csv")
write.csv(knot1, file="D:// output knot1.csv")
knota1=cbind(knot1,GCV) #gabungGCVdanknot
knotorder1=knota1[order(knota1[,ncol(knota1)]),] #mengurutkan
write.csv(knota1, file="D:// output knota1.csv")
write.csv(knotorder1, file="D:// output knota1 order.csv")
}

```

Lampiran 3 Program Regresi Semiparametrik Spline *Truncated* Dua Titik Knot

```

GCV2=function(para)
{
data=read.csv("D:// data.csv")
data=as.matrix(data)
p=length(data[,1])
q=length(data[1,])
m=ncol(data)-para-1
dataA=data[, (para+2):q]
F=matrix(0,nrow=p,ncol=p)
diag(F)=1
nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
knot=matrix(ncol=m,nrow=nk)
for (i in (1:m))
{
for (j in (1:nk))
{
a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
knot[j,i]=a[j]
}
}
a1=nrow(knot)
knot=knot[2:(a1-1),]
a2=nrow(knot)
z=(a2*(a2-1)/2)
knot2=cbind(rep(NA,(z+1)))
for (i in (1:m))
{
knot1=rbind(rep(NA,2))
for ( j in 1:(a2-1))
{
for (k in (j+1):a2)
{
xx=cbind(knot[j,i],knot[k,i])
knot1=rbind(knot1,xx)
}
}
}
knot2=cbind(knot2,knot1)
}
knot2=knot2[2:(z+1),2:(2*m+1)]
aa=rep(1,p)
data2=matrix(ncol=(2*m),nrow=p)
data1=data[, (para+2):q]
data3=data[,2:q]
a3=length(knot2[,1])
GCV=rep(NA,a3)
Rsqr=rep(NA,a3)

```

Lampiran 3 Program Regresi Semiparametrik Spline *Truncated* Dua Titik Knot
(Lanjutan)

```

for (i in 1:a3)
{
for (j in 1:(2*m))
{ if (mod(j,2)==1) b=floor(j/2)+1 else b=j/2
for (k in 1:p)
{ if (data1[k,b]<knot2[i,j]) data2[k,j]=0 else data2[k,j]=data1[k,b]-
knot2[i,j]
} }
mx=cbind(aa,data3,data2)
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0
for (r in (1:p))
{
sum=(data[r,1]-yhat[r,])^2
sum1=(yhat[r,]-mean(data[,1]))^2
SSE=SSE+sum
SSR=SSR+sum1
}
Rsq[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsq=as.matrix(Rsq)
cat("=====", "\n")
cat("Nilai Knot dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
print(knot1)
print (knot2)
cat("=====", "\n")
cat("Rsq dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
print (Rsq)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
print (GCV)

```

Lampiran 3 Program Regresi Semiparametrik Spline *Truncated* Dua Titik Knot
(Lanjutan)

```
s1=min(GCV)
print(max(Rsq))
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
cat(" GCV =", s1, "\n")
write.csv(GCV, file="D://output GCV2.csv")
write.csv(Rsq, file="D:// output Rsq2.csv")
write.csv(knot2, file="D:// output knot2.csv")
knota2=cbind(knot2, GCV)
knotorder2=knota2[order(knota2[,ncol(knota2)]),]
write.csv(knota2, file="D:// output knota2.csv")
write.csv(knotorder2, file="D:// output knota2 order.csv")
}
```

Lampiran 4 Program Regresi Semiparametrik Spline *Truncated* Tiga Titik Knot

```

GCV3=function(para)
{
data=read.csv("D:// data.csv",header=T)
data=as.matrix(data)
p=length(data[,1])
q=length(data[1,])
m=ncol(data)-para-1
F=matrix(0,nrow=p,ncol=p)
dataA=data[(para+2):q]
diag(F)=1
nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
knot=matrix(ncol=m,nrow=nk)
for (i in (1:m))
{
for (j in (1:nk))
{
a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
knot[j,i]=a[j]
}
}
knot=knot[2:(nk-1),]
a2=nrow(knot)
z=(a2*(a2-1)*(a2-2)/6)
knot1=cbind(rep(NA,(z+1)))
for (i in (1:m))
{
knot2=rbind(rep(NA,3))
for (j in 1:(a2-2))
{
for (k in (j+1):(a2-1))
{
for (g in (k+1):a2)
{
xx=cbind(knot[j,i],knot[k,i],knot[g,i])
knot2=rbind(knot2,xx)
}
}
}
knot1=cbind(knot1,knot2)
}
knot1=knot1[2:(z+1),2:(3*m+1)]
aa=rep(1,p)
data1=matrix(ncol=(3*m),nrow=p)
data2=data[(para+2):q]

```

Lampiran 4 Program Regresi Semiparametrik Spline *Truncated* Tiga Titik Knot
(Lanjutan)

```

a1=length(knot1[,1])
GCV=rep(NA,a1)
Rsqr=rep(NA,a1)
for (i in 1:a1)
{
for (j in 1:ncol(knot1))
{
b=ceiling(j/3)
for (k in 1:p)
{
if (data2[k,b]<knot1[i,j]) data1[k,j]=0 else data1[k,j]=data2[k,b]-knot1[i,j]
}
}
}
mx=cbind(aa,data[,2:q],data1)
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0
for (r in (1:p))
{
sum=(data[r,1]-yhat[r,])^2
sum1=(yhat[r,]-mean(data[,1]))^2
SSE=SSE+sum
SSR=SSR+sum1
}
Rsqr[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsqr=as.matrix(Rsqr)
cat("=====\n")
cat("Nilai Knot dengan Spline linear 3 knot","\n")
cat("=====\n")
print (knot1)
cat("=====\n")
cat("Rsqr dengan Spline linear 3 knot","\n")
cat("=====\n")
print (Rsqr)

```

Lampiran 4 Program Regresi Semiparametrik Spline *Truncated* Tiga Titik Knot
(Lanjutan)

```
r=max(Rsq)
print (r)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
print (r)
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
cat(" GCV =", s1, "\n")
write.csv(GCV, file="D:// output GCV3.csv")
write.csv(Rsq, file="D:// output Rsq3.csv")
write.csv(knot1, file="D:// output knot3.csv")
knota3=cbind(knot1, GCV)
knotorder3=knota3[order(knota3[, ncol(knota3)]), ]
write.csv(knota3, file="D:// output knota3.csv")
write.csv(knotorder3, file="D:// output knota3 order.csv")
}
```

Lampiran 5 Program Regresi Semiparametrik Spline *Truncated* Kombinasi Titik Knot

```

GCVKK=function(para)
{
data=read.csv("D:// data.csv",header=T)
data=as.matrix(data)
p1=length(data[,1])
q1=length(data[1,])
v=para+2
F=matrix(0,nrow=p1,ncol=p1)
diag(F)=1
x1=read.table("D:// x1.txt")
x2=read.table("D:// x2.txt")
x3=read.table("D:// x3.txt")
x4=read.table("D:// x4.txt")
n2=nrow(x1)
a=matrix(nrow=4,ncol=3^4)
m=0
for (i in 1:3)
for (j in 1:3)
for (k in 1:3)
for (l in 1:3)
{
m=m+1
a[,m]=c(i,j,k,l)
}
a=t(a)
GCV=matrix(nrow=nrow(x1),ncol=3^4)
for (i in 1:3^4)
{
for (h in 1:nrow(x1))
{
if (a[i,1]==1)
{
gab=as.matrix(x1[,1])
gen=as.matrix(data[,v])
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
}
else
if (a[i,1]==2)
{

```

Lampiran 5 Program Regresi Semiparametrik Spline *Truncated* Kombinasi Titik Knot (Lanjutan)

```

gab=as.matrix(x1[,2:3])
gen=as.matrix(cbind(data[,v],data[,v]))
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
else
gab=as.matrix(x1[,4:6])
gen=as.matrix(cbind(data[,v],data[,v],data[,v]))
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,2]==1)
{
gab=as.matrix(x2[,1] )
gen=as.matrix(data[,v+1])
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,2]==2)
{
gab=as.matrix(x2[,2:3] )
gen=as.matrix(cbind(data[,v+1],data[,v+1]))
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
}
else
{

```

Lampiran 5 Program Regresi Semiparametrik Spline *Truncated* Kombinasi Titik Knot (Lanjutan)

```

gab=as.matrix(x2[,4:6])
gen=as.matrix(cbind(data[,v+1],data[,v+1],data[,v+1]))
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,3]==1)
{
gab=as.matrix(x3[,1] )
gen=as.matrix(data[,v+2])
cc=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,3]==2)
{
gab=as.matrix(x3[,2:3] )
gen=as.matrix(cbind(data[,v+2],data[,v+2]))
cc=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x3[,4:6])
gen=as.matrix(cbind(data[,v+2],data[,v+2],data[,v+2]))
cc=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
}

```

Lampiran 5 Program Regresi Semiparametrik Spline *Truncated* Kombinasi Titik Knot (Lanjutan)

```

if (a[i,4]==1)
{
gab=as.matrix(x4[,1] )
gen=as.matrix(data[, (v+3)])
dd=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,4]==2)
{
gab=as.matrix(x4[,2:3] )
gen=as.matrix(cbind(data[, (v+3)],data[, (v+3)]))
dd=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x4[,4:6])
gen=as.matrix(cbind(data[, (v+3)],data[, (v+3)],data[, (v+3)]))
dd=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
ma=as.matrix(cbind(aa,bb,cc,dd))
mx=cbind(rep(1,nrow(data)),data[,2:q1],na.omit(ma))
mx=as.matrix(mx)
C=pinv(t(mx)%*% mx)
B=C%*%(t(mx)%*% data[,1])
yhat=mx%*% B
SSE=0
SSR=0
for (r in 1:nrow(data))
{

```

Lampiran 5 Program Regresi Semiparametrik Spline *Truncated* Kombinasi Titik Knot (Lanjutan)

```

sum=(data[r,1]-yhat[r,])^2
sum1=(yhat[r,]-mean(data[,1]))^2
SSE=SSE+sum
SSR=SSR+sum1
}
Rsq=(SSR/(SSE+SSR))*100
MSE=SSE/p1
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p1)^2
GCV[h,i]=MSE/A2
}

if (a[i,1]==1) sp=x1[,1] else
if (a[i,1]==2) sp=x1[,2:3] else
sp=x1[,4:6]
if (a[i,2]==1) spl=x2[,1] else
if (a[i,2]==2) spl=x2[,2:3] else
spl=x2[,4:6]
if (a[i,3]==1) splin=x3[,1] else
if (a[i,3]==2) splin=x3[,2:3] else
splin=x3[,4:6]
if (a[i,4]==1) spline=x4[,1] else
if (a[i,4]==2) spline=x4[,2:3] else
spline=x4[,4:6]
kkk=cbind(sp,spl,splin,spline)
cat("=====", "\n")
print(i)
print(kkk)
print(Rsq)
}
GCVt=t(GCV)
GCVmin=min(GCVt)
print (GCVmin)
GCVorder=order(GCVt)
write.csv(GCV,file="D:// output GCV kombinasi.csv")
write.csv(GCVt,file="D:// output GCV t kombinasi.csv")
write.csv(GCVorder,file="D:// output GCV t order kombinasi.csv")
write.csv(Rsq,file="D:// output Rsq kombinasi.csv")
}

```

Lampiran 6 Uji Signifikansi Parameter

```
uji=function(alpha,para)
{
data=read.csv("D:// coba.csv")
knot=read.table("D://knot.txt")
data=as.matrix(data)
knot=as.matrix(knot)
ybar=mean(data[,1])
m=para+2
p=nrow(data)
q=ncol(data)
dataA=cbind(data[,m],data[,m],data[,m],data[,m+1],data[,m+1],data[,m+1],data[,
m+2],data[,m+2],data[,m+2],data[,m+3],data[,m+3],data[,m+3])
dataA=as.matrix(dataA)
satu=rep(1,p)
n1=ncol(knot)
data.knot=matrix(ncol=n1,nrow=p)
for (i in 1:n1)
{
for(j in 1:p)
{
if(dataA[j,i]<knot[1,i])data.knot[j,i]=0 else
data.knot[j,i]=dataA[j,i]-knot[1,i]
}
}
mx=cbind(satu,data[,2],data[,3],data.knot[,1:3],data[,4],data.knot[,4:6],data[,5],da
ta.knot[,7:9],data[,6],data.knot[,10:12])
mx=as.matrix(mx)
B=(pinv(t(mx)%*%mx))%*%t(mx)%*%data[,1]
cat("=====", "\n")
cat("Estimasi Parameter", "\n")
cat("=====", "\n")
print (B)
n1=nrow(B)
yhat=mx%*%B
res=data[,1]-yhat
SSE=sum((data[,1]-yhat)^2)
SSR=sum((yhat-ybar)^2)
SST=SSR+SSE
MSE=SSE/(p-n1)
MSR=SSR/(n1-1)
Rsqr=(SSR/(SSR+SSE))*100
```

Lampiran 6 Uji Signifikansi Parameter (Lanjutan)

```

#uji F (uji serentak)
Fhit=MSR/MSE
pvalue=pf(Fhit,(n1-1),(p-n1),lower.tail=FALSE)
if (pvalue<=alpha)
{
cat("-----", "\n")
cat("Kesimpulan hasil uji serentak", "\n")
cat("-----", "\n")
cat("Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan", "\n")
cat("", "\n")
}
else
{
cat("-----", "\n")
cat("Kesimpulan hasil uji serentak", "\n")
cat("-----", "\n")
cat("Gagal Tolak Ho yakni semua prediktor tidak berpengaruh signifikan", "\n")
cat("", "\n")
}

#uji F (uji individu)
F.parsial=rep(NA,n1)
q1=rep(NA,n1)
q2=rep(NA,n1)
y=data[,1]
c=diag(1,nrow=n1,ncol=n1)
pval=rep(NA,n1)
cat("-----", "\n")
cat("Kesimpulan hasil uji individu", "\n")
cat("-----", "\n")
for (i in 1:n1)
{
q1[i]=t(y)%*%mx%*%pinv(t(mx)%*%mx)%*%c[,i]%*%pinv(t(c[,i])%*%pinv(t
(mx)%*%mx)%*%c[,i])%*%t(c[,i])%*%pinv(t(mx)%*%mx)%*%t(mx)%*%y
q2[i]=t(y)%*%(diag(1,nrow=nrow(mx),ncol=nrow(mx))-
(mx)%*%pinv(t(mx)%*%mx)%*%t(mx)))%*%y
F.parsial[i]=q1[i]/(q2[i]/(p-n1))
pval[i]=pf(F.parsial[i],1,n1,lower.tail=FALSE)
if (pval[i]<=alpha) cat("Tolak Ho yakni prediktor signifikan dengan
pvalue",pval[i], "\n") else cat("Gagal tolak Ho yakni prediktor tidak signifikan
dengan pvalue",pval[i], "\n")
}
F.parsial=as.matrix(F.parsial)
cat("=====", "\n")
cat("nilai F hitung uji individu", "\n")

```

Lampiran 6 Uji Signifikansi Parameter (Lanjutan)

```
cat("=====", "\n")
print (F.parsial)
cat("Analysis of Variance", "\n")
cat("=====", "\n")
  cat("Sumber    df    SS    MS    Fhit", "\n")
  cat("Regresi    ,(n1-1), " ",SSR," ",MSR," ",Fhit," "\n")
  cat("Error      ",p-n1," ",SSE," ",MSE," "\n")
  cat("Total      ",p-1," ",SST," "\n")
cat("=====", "\n")
  cat("s=",sqrt(MSE)," Rsq=",Rsq," "\n")
  cat("pvalue(F)=",pvalue," "\n")
Rsqadj=(1-((SSE/SST)*((p-1)/(p-n1))))*100
  cat("Rsqadj=",Rsqadj," "\n")
write.csv(res,file="D://output uji residual knot.csv")
write.csv(pval,file="D://output uji pvalue knot.csv")
write.csv(mx,file="D://output uji mx knot.csv")
write.csv(yhat,file="D://output uji yhat knot.csv")
}
```

Lampiran 7 Uji *Glejser*

```

glejser=function(data,knot,res,alpha,para)
{
data=read.csv("D:// data.csv")
knot=read.table("D://knot.txt")
res=read.table("D://residual.txt")
data=as.matrix(data)
knot=as.matrix(knot)
res=abs(res)
res=as.matrix(res)
rbar=mean(res)
m=para+2
p=nrow(data)
q=ncol(data)
dataA=cbind(data[,m],data[,m],data[,m],data[,m+1],data[,m+1],data[,m+1],data[,
m+2],data[,m+2],data[,m+2],data[,m+3],data[,m+3],data[,m+3])
dataA=as.matrix(dataA)
satu=rep(1,p)
n1=ncol(knot)
data.knot=matrix(ncol=n1,nrow=p)
for (i in 1:n1)
{
  for(j in 1:p)
  {
    if (dataA[j,i]<knot[1,i]) data.knot[j,i]=0 else data.knot[j,i]=dataA[j,i]-
knot[1,i]
  }
}
mx=cbind(satu,data[,2],data[,3],data.knot[,1:3],data[,4],data.knot[,4:6],data[,5],da
ta.knot[,7:9],data[,6],data.knot[,10:12])
mx=as.matrix(mx)
B=(ginv(t(mx)%*%mx))%*%t(mx)%*%res
n1=nrow(B)
yhat=mx%*%B
residual=res-yhat
SSE=sum((res-yhat)^2)
SSR=sum((yhat-rbar)^2)
SST=SSR+SSE
MSE=SSE/(p-n1)
MSR=SSR/(n1-1)
Rsq=(SSR/SST)*100
#uji F (uji serentak)
Fhit=MSR/MSE
pvalue=pf(Fhit,(n1-1),(p-n1),lower.tail=FALSE)
if (pvalue<=alpha)
}

```

Lampiran 7 Uji Glejser (Lanjutan)

```
{
cat("-----", "\n")
cat("Kesimpulan hasil uji serentak", "\n")
cat("-----", "\n")
cat("Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan atau terjadi
heteroskedastisitas", "\n")
cat("", "\n")
}
else
{
cat("-----", "\n")
cat("Kesimpulan hasil uji serentak", "\n")
cat("-----", "\n")
cat("Gagal Tolak Ho yakni semua prediktor tidak berpengaruh signifikan atau
tidak terjadi heteroskedastisitas", "\n")
cat("", "\n")
}
cat("Analysis of Variance", "\n")
cat("=====", "\n")
    cat("Sumber    df    SS    MS    Fhit", "\n")
    cat("Regresi    ", (n1-1), " ", "SSR, " ", "MSR, """, "Fhit, "\n")
    cat("Error      ", p-n1, " ", "SSE, """, "MSE, "\n")
    cat("Total      ", p-1, " ", "SST, "\n")
cat("=====", "\n")
    cat("s=", sqrt(MSE), " Rsq=", Rsq, "\n")
    cat("pvalue(F)=", pvalue, "\n")
```

BIOGRAFI PENULIS



Dewi Fitriana lahir di Boyolali, Jawa Tengah pada tanggal 27 Juli 1982, merupakan putri ketiga dari lima bersaudara, pasangan Bapak H. Slamet dan Ibu Hj. Maemunah. Pendidikan formal yang pernah ditempuh oleh Penulis adalah MI Negeri Boyolali (1988-1994), SLTP Negeri 2 Boyolali (1994- 1997), SMA Negeri 1 Boyolali (1997-2000), Sekolah Tinggi Ilmu Statistik Jakarta (2000-2004). Setelah menyelesaikan program studi D-IV STIS, penulis bekerja di Badan Pusat Statistik (BPS) Kabupaten Ende, Provinsi Nusa Tenggara Timur. Pada tahun 2007, penulis ditugaskan di BPS Kabupaten Mojokerto, Provinsi Jawa Timur. Pada tahun 2015, penulis mendapatkan kesempatan dan kepercayaan untuk melanjutkan pendidikan Magister (S2) di Program Pascasarjana Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam (MIPA), Institut Teknologi Sepuluh Nopember (ITS) Surabaya. Pembaca yang ingin memberikan kritik, saran dan berdiskusi mengenai penelitian ini dapat menghubungi penulis melalui email dewi.fitriana@bps.go.id.