



TUGAS AKHIR - SS141501

FAKTOR-FAKTOR YANG MEMPENGARUHI *TOTAL FERTILITY RATE* (TFR) DI INDONESIA DENGAN MENGGUNAKAN REGRESI SEMIPARAMETRIK SPLINE TRUNCATED

KHUSNIYAH
NRP 1312 106 006

Dosen Pembimbing
Prof. Dr. Drs. I Nyoman Budiantara, M.Si.

JURUSAN STATISTIKA
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember
Surabaya 2015



FINAL PROJECT - SS141501

**FACTORS AFFECTING *TOTAL FERTILITY RATE*
(TFR) IN INDONESIA USING SEMIPARAMETRIC
REGRESSION SPLINE TRUNCATED**

KHUSNIYAH
NRP 1312 106 006

Supervisor
Prof. Dr. Drs. I Nyoman Budiantara, M.Si.

DEPARTMENT OF STATISTICS
Faculty of Mathematics and Natural Sciences
Sepuluh Nopember Institute of Technology
Surabaya 2015

KATA PENGANTAR

Assalamualaikum Wr. Wb.

Puji syukur kehadirat Allah SWT yang telah memberikan segala petunjuk, kemudahan, rahmat serta karunia-Nya, sehingga penulis dapat menyelesaikan laporan Tugas Akhir dengan judul

”FAKTOR-FAKTOR YANG MEMPENGARUHI *TOTAL FERTILTY RATE* (TFR) DI INDONESIA DENGAN MENGGUNAKAN REGRESI SEMIPARAMETRIK *SPLINE TRUNCATED*”

Sholawat dan salam tak lupa penulis sampaikan pada junjungan besar Nabi Muhammad SAW. Selesaiannya Tugas Akhir serta laporan ini tak lepas dari peranan berbagai pihak. Oleh karena itu penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada :

1. Orang Tua dan saudara-saudara tercinta yang sangat saya hormati, yang telah menjadi sumber kekuatan dan semangat tiada tara. Terimakasih atas segala doa dan dukungannya.
2. Bapak Prof. Dr. Drs. I Nyoman Budiantara, M.Si selaku dosen pembimbing yang telah menuntun saya, memberikan segala nasehat, serta pengetahuan baru demi terselesaikannya Tugas Akhir ini.
3. Ibu Dra. Madu Ratna, M.Si dan Dr. Vita Ratnasari, S.Si, M.Si selaku dosen penguji yang telah memberikan banyak saran, kritik dan masukan demi kesempurnaan Tugas Akhir ini.
4. Bapak Dr. Purhadi, M.Sc selaku dosen wali yang telah memberikan semangat dan memberikan bantuan kepada penulis untuk melanjutkan pendidikan di Jurusan Statistika ITS sampai dengan selesai.
5. Bapak Dr. Muhammad Mashuri, M.T selaku Ketua Jurusan Statistika ITS.

6. Ibu Dra. Lucia Aridinanti, M.T selaku kaprodi S1 Jurusan Statistika.
7. Seluruh Dosen Jurusan Statistika ITS yang telah memberikan ilmu dan membagi pengalamannya dengan penulis.
8. Seluruh staf Jurusan Statistika ITS yang telah membantu demi kelancaran perkuliahan.
9. Teman-teman seperjuangan dan sebingbingan, Diana, Karisma dan Fifi yang bersama-sama berjuang untuk menyelesaikan tugas akhir ini dan atas semangat berbagi kalian.
10. Teman-teman Kepegawaian atas dukungannya.
11. Teman-teman Lintas Jalur 2012 dan Lintas Jalur 2013 atas kebersamaan dan dukungannya.
12. Pihak-pihak lain yang telah mendukung dan membantu atas terselesaikannya Tugas Akhir ini yang tidak dapat disebutkan satu persatu.

Besar harapan penulis agar Tugas Akhir ini bermanfaat dan menambah wawasan keilmuan bagi berbagai pihak. Tugas Akhir ini masih belum sempurna, oleh karena itu saran dan kritik yang membangun sangat diharapkan demi sempurnanya Tugas Akhir ini.

WassalamualaikumWr. Wb.

Surabaya, Januari 2015

Penulis

LEMBAR PENGESAHAN

**FAKTOR – FAKTOR YANG MEMPENGARUHI
TOTAL FERTILITY RATE (TFR) DI INDONESIA
DENGAN MENGGUNAKAN REGRESI
SEMIPARAMETRIK SPLINE TRUNCATED**

TUGAS AKHIR

**Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains**

pada

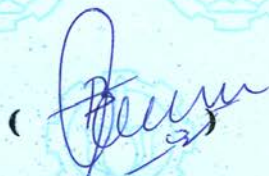
**Program Studi S-1 Jurusan Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember**

Oleh :

**KHUSNIYAH
NRP. 1312 106 006**

Disetujui oleh Pembimbing Tugas Akhir :

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



Mengetahui

Ketua Jurusan Statistika FMIPA-ITS



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

SURABAYA, JANUARI 2015

**FAKTOR-FAKTOR YANG MEMPENGARUHI TOTAL
FERTILITY RATE (TFR) DI INDONESIA DENGAN
MENGUNAKAN REGRESI SEMIPARAMETRIK
SPLINE TRUNCATED**

Nama : Khusniyah
NRP : 1312 106 006
Jurusan : Statistika
Dosen Pembimbing : Prof. Dr. Drs. I Nyoman Budiantara, M.Si

Abstrak

Indonesia merupakan negara dengan jumlah penduduk yang besar. Menurut BPS (2013) jumlah penduduk Indonesia cenderung meningkat dari tahun ke tahun, sehingga dapat menjadi penghambat keberhasilan pembangunan apabila peningkatan jumlah penduduk tidak terkontrol. Salah satu upaya pemerintah dalam rangka mengendalikan laju pertumbuhan penduduk adalah melalui program KB. Salah satu indikator yang dapat digunakan untuk mengukur keberhasilan program KB adalah Total Fertility Rate (TFR). Faktor-faktor yang mempengaruhi TFR, sebagian memiliki pola tertentu (parametrik) dan sebagian lagi tidak memiliki pola tertentu (nonparametrik) sehingga pemodelan dilakukan dengan metode regresi semiparametrik spline. Penelitian ini menggunakan pendekatan spline dengan titik knot optimum didapatkan dengan metode Generalized Cross Validation (GCV). Berdasarkan hasil analisis didapat faktor-faktor yang berpengaruh signifikan terhadap TFR antara lain persentase keinginan membatasi kelahiran setelah memiliki 2 anak (X_1), persentase Contraceptive Prevalence Rate (CPR) (X_3), dan persentase akses media pesan KB (X_4). Model regresi semiparametrik spline yang terbentuk memiliki nilai R^2 sebesar 84,31% dan MSE sebesar 0.0353.

Kata kunci: Total Fertility Rate (TFR), Regresi Semiparametrik Spline, GCV

(Halaman ini sengaja dikosongkan)

FACTORS AFFECTING *TOTAL FERTILITY RATE (TFR)* IN INDONESIA USING SEMIPARAMETRIC REGRESSION SPLINE TRUNCATED

Name of Student : Khusniyah
NRP : 1312 106 006
Department : Statistika
Supervisor : Prof. Dr. Drs. I Nyoman Budiantara, M.Si

Abstract

Indonesia is a country with a large population. According to BPS (2013) Indonesia's population is likely to increase from year to year, so it can become an obstacle to the success of development if the increase in population uncontrolled. One of the government efforts to control the population growth rate is through the family planning program . One of the indicators that can be used to measure the success of family planning programs is the Total Fertility Rate (TFR). Factors that influence the TFR, the majority have a certain pattern (parametric) and some do not have a specific pattern (nonparametric) so that the modeling was conducted using semiparametric regression spline. This study uses spline approach with optimum knots point obtained from Generalized Cross Validation (GCV) method. Based on the analysis results obtained factors that significantly influence the TFR among others, the percentage desire birth control after having 2 children (X_1), the percentage of Contraceptive Prevalence Rate (CPR) (X_3), and the percentage of media access message KB (X_4). Semiparametric regression model spline formed have R^2 value of 84.31 % and MSE of 0.0353.

Keywords: Total Fertility Rate (TFR), Spline Semiparametric Regression, GCV

(Halaman ini sengaja dikosongkan)

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR TABEL	xvii
DAFTAR GAMBAR	xix
DAFTAR LAMPIRAN	xxi
 BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah	4
 BAB II TINJAUAN PUSTAKA	
2.1 Analisis Regresi	5
2.1.1 Regresi Parametrik	5
2.1.2 Regresi Non Parametrik	6
2.1.3 Regresi Semiparametrik	7
2.2 Regresi Spline Univariabel	7
2.3 Regresi Spline Multivariabel	8
2.4 Pemilihan Titik Knot Optimal	9
2.5 Pengujian Parameter Model Regresi	9
2.6 Pemeriksaan Asumsi Residual dari Model regresi Spline	11
2.7 Kriteria Pemilihan Model Terbaik	13
2.8 Ferilitas (Kelahiran)	14
2.9 Age Specific Fertility Rate (ASFR)	14
2.10 Total Fertility Rate (TFR)	15

2.11	Penelitian Sebelumnya.....	16
------	----------------------------	----

BAB III METODOLOGI PENELITIAN

3.1	Sumber Data	19
3.2	Variabel Penelitian.....	19
3.3	Metode Analisis	22
3.4	Diagram Alir Penelitian	23

BAB IV ANALISIS DAN PEMBAHASAN

4.1	Karakteristik <i>Total Fertility Rate</i> (TFR) dengan Faktor-Faktor yang Diduga Mempengaruhi.....	25
4.2	Pola Hubungan <i>Total Fertility Rate</i> (TFR) dengan Faktor-Faktor yang Diduga Mempengaruhi ...	29
4.2.1	<i>Scatterplot</i> TFR dengan Persentase Keinginan Membatasi Kelahiran Setelah Mempunyai 2 Anak (X_1).....	29
4.2.2	<i>Scatterplot</i> TFR dengan Persentase Unmet Need (X_2).....	30
4.2.3	<i>Scatterplot</i> TFR dengan Persentase <i>Contraceptive Prevalence Rate</i> (X_3).....	31
4.2.4	<i>Scatterplot</i> TFR dengan Persentase Akses Media Pesan KB (X_4).....	32
4.3	Pemilihan Titik Knot Optimal.....	33
4.3.1	Pemilihan Titik Knot Optimal dengan Satu Titik Knot.....	34
4.3.2	Pemilihan Titik Knot Optimal dengan Dua Titik Knot	34
4.3.3	Pemilihan Titik Knot Optimal dengan Tiga Titik Knot.....	35
4.3.4	Pemilihan Titik Knot Optimal dengan Kombinasi Knot.....	36
4.4	Pemodelan Regresi Semiparametrik Spline dengan Titik Knot Optimal	37
4.5	Pengujian Signifikansi Parameter	39
4.5.1	Uji Serentak	39
4.5.2	Uji Parsial	39

4.6	Pemilihan Titik Knot Optimal Tanpa Menggunakan Variabel Persentase <i>Unmet Need</i> (X_2)	40
4.7	Pemodelan Regresi Semiparametrik Spline dengan Titik Knot Optimal Tanpa Menggunakan Variabel Persentase <i>Unmet Need</i> (X_2)	44
4.8	Pengujian Signifikansi Parameter Tanpa Menggunakan Variabel Persentase <i>Unmet Need</i> (X_2)	45
	4.8.1 Uji Serentak.....	45
	4.8.2 Uji Parsial.....	46
4.9	Uji Asumsi Residual	46
	4.9.1 Uji Asumsi Residual Identik	47
	4.9.2 Uji Asumsi Residual Independen	48
	4.9.3 Uji Asumsi Distribusi Normal.....	49
4.10	Interpretasi Model Regresi Semiparametrik Spline	50
BAB V KESIMPULAN DAN SARAN		
5.1	Kesimpulan.....	53
5.2	Saran.....	54
DAFTAR PUSTAKA.....		55
LAMPIRAN.....		59
BIODATA PENULIS.....		79

(Halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

	Halaman
Gambar 3.1 Diagram Alir Penelitian.....	23
Gambar 4.1 Nilai <i>Total Fertility Rate</i> (TFR) 33 Provinsi di Indonesia	26
Gambar 4.2 <i>Scatterplot</i> Variabel TFR dengan Persentase Keinginan Membatasi Kelahiran Setelah Mempunyai 2 Anak	30
Gambar 4.3 <i>Scatterplot</i> Variabel TFR dengan Persentase Unmet Need.....	31
Gambar 4.4 <i>Scatterplot</i> Variabel TFR dengan Variabel Persentase <i>Contraceptive Prevalence Rate</i> (CPR) .	32
Gambar 4.5 <i>Scatterplot</i> Variabel TFR dengan Variabel Persentase Akses Media Pesan KB	33
Gambar 4.6 <i>Scatterplot</i> antara Fits dan Residual.....	47
Gambar 4.7 Plot ACF dari Residual.....	48
Gambar 4.8 Uji Kolmogorov Smirnov.....	49

(Halaman ini sengaja dikosongkan)

DAFTAR TABEL

	Halaman
Tabel 2.1 ANOVA.....	10
Tabel 3.1 Struktur Data Penelitian	19
Tabel 3.1 Variabel Penelitian	20
Tabel 4.1 Statistik Deskriptif Variabel Prediktor	27
Tabel 4.2 Nilai GCV untuk Satu Titik Knot.....	34
Tabel 4.3 Nilai GCV untuk Dua Titik Knot	35
Tabel 4.4 Nilai GCV untuk Tiga Titik Knot	36
Tabel 4.5 Nilai GCV untuk Kombinasi Knot	37
Tabel 4.6 Estimasi Parameter Model Regresi Semiparametrik Spline	38
Tabel 4.7 ANOVA Model Regresi Semiparametrik Spline ...	39
Tabel 4.8 Uji Parsial Model Regresi Semiparametrik Spline.	40
Tabel 4.9 Nilai GCV untuk Satu Titik Knot.....	41
Tabel 4.10 Nilai GCV untuk Dua Titik Knot	41
Tabel 4.11 Nilai GCV untuk Tiga Titik Knot	42
Tabel 4.12 Nilai GCV untuk Kombinasi Knot	43
Tabel 4.13 Estimasi Parameter Model Regresi Semiparametrik Spline	44
Tabel 4.14 ANOVA Model Regresi Semiparametrik Spline ...	45
Tabel 4.15 Uji Parsial Model Regresi Semiparametrik Spline.	46
Tabel 4.16 ANOVA untuk Uji <i>Glejser</i>	48

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Data <i>Total Fertility Rate</i> (TFR) beserta Faktor-faktor yang diduga mempengaruhi	59
Lampiran 2 Program Regresi Semiparametrik Spline Dengan Software R	61
Lampiran 3 Uji Signifikansi Parameter	75
Lampiran 4 Uji Glejser	78
Lampiran 5 Output Uji Signifikansi Parameter Model Regresi Semiparametrik	81
Lampiran 6 Output Uji Glejser	83

(Halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Total Fertility Rate (TFR) atau angka fertilitas total erat kaitannya dengan penduduk. Penduduk merupakan objek dan subjek dalam pembangunan. Sebagai objek, penduduk merupakan sasaran pembangunan. Sedangkan sebagai subjek, penduduk merupakan pelaku pembangunan. Oleh karena itu, penduduk berperan penting dalam menentukan arah dan keberhasilan pembangunan. Akan tetapi, peran penting penduduk sebagai objek dan subjek dalam pembangunan tersebut dapat menjadi penghambat keberhasilan pembangunan apabila masih banyak terdapat permasalahan kependudukan. Permasalahan kependudukan di Indonesia salah satunya adalah jumlah penduduk yang besar. Indonesia merupakan negara dengan jumlah penduduk yang besar. Data tentang jumlah penduduk Indonesia dapat diketahui dari hasil Sensus Penduduk (SP). Berdasarkan sensus penduduk tahun 2010, jumlah penduduk di Indonesia mencapai 237 juta jiwa. Jumlah tersebut diproyeksikan akan meningkat mencapai 273 juta jiwa pada tahun 2025 (BAPPENAS, BPS, & UNFPA, 2005). Menurut BPS (2013), hasil sensus penduduk menunjukkan bahwa perkembangan jumlah penduduk Indonesia cenderung meningkat dari tahun ke tahun. Hal tersebut dapat menjadi penghambat keberhasilan pembangunan apabila peningkatan jumlah penduduk tidak terkontrol. Salah satu upaya pemerintah dalam rangka mengendalikan laju pertumbuhan penduduk adalah melalui program Keluarga Berencana (KB).

Paradigma baru program KB nasional telah mengalami perubahan visi dari mewujudkan Norma Keluarga Kecil Bahagia Sejahtera (NKKBS) menjadi visi 4 untuk mewujudkan “Keluarga Berkualitas tahun 2015” (Saifuddin, 2003). Program KB sebagai salah satu kebijakan pemerintah dalam bidang kependudukan

memiliki implikasi yang tinggi terhadap pembangunan kependudukan yang bersifat kuantitatif dan kualitatif (Putuamar, 2007). Oleh karena itu, program KB memiliki posisi strategis dalam upaya pengendalian laju pertumbuhan penduduk melalui pengendalian kelahiran dan pendewasaan usia perkawinan (secara kuantitatif), maupun pembinaan ketahanan dan peningkatan kesejahteraan keluarga (secara kualitatif). Karena vitalnya peranan program KB dalam mengatasi permasalahan kependudukan di Indonesia, maka perlu dilakukan penelitian yang dapat mengukur keberhasilan program KB. Salah satu indikator yang dapat digunakan untuk mengukur keberhasilan program KB adalah TFR.

TFR adalah rata-rata anak yang dilahirkan seorang wanita selama masa usia suburnya. TFR merupakan gambaran mengenai rata-rata jumlah anak yang dilahirkan seorang perempuan dari usia 15 sampai 49 tahun. Perbandingan angka TFR antar negara atau antar daerah dapat menunjukkan keberhasilan daerah dalam melaksanakan pembangunan sosial ekonominya. Angka TFR yang tinggi dapat merupakan cerminan rata-rata usia kawin yang rendah, tingkat pendidikan yang rendah terutama perempuannya, tingkat sosial ekonomi rendah atau tingkat kemiskinan yang tinggi. Selain itu tentu saja menunjukkan tingkat keberhasilan program KB yang dilaksanakan selama tiga dekade ini. Angka fertilitas total secara nasional di Indonesia berdasarkan hasil Survei Demografi dan Kesehatan Indonesia (SDKI) 2012, terjadi stagnasi tetap dalam angka fertilitas total (kelahiran) atau TFR yaitu 2,6. Nilai tersebut masih terlalu tinggi untuk dapat mencapai kondisi Penduduk Tumbuh Seimbang (PTS). Yang dimaksud PTS adalah kondisi dimana penduduk Indonesia akan mencapai tingkat penggantian manusia (*replacement level*) apabila angka fertilitas total turun mencapai 2,1 pada tahun 2015 (BPS, 2014). Berdasarkan latar belakang tersebut, maka perlu dilakukan penelitian mengenai faktor-faktor yang mempengaruhi TFR di Indonesia.

Penelitian sebelumnya tentang fertilitas pernah dilakukan oleh beberapa peneliti diantaranya yaitu, Sari (2006) melakukan pemodelan regresi spline untuk mengestimasi rata-rata angka kelahiran ASFR (*Age Spesific Fertility Rate*) di Jawa Timur. Radifan (2010) untuk menganalisis faktor-faktor yang mempengaruhi tingkat fertilitas atau angka kelahiran total di Indonesia tahun 2007 dengan menggunakan metode *Ordinary Least Squared* (OLS). Sulisyaningsih (2010) melakukan pemodelan TFR dan faktor-faktor yang mempengaruhi di Provinsi Jawa Timur. Selanjutnya Imam (2014) melakukan analisis regresi probit dengan efek interaksi untuk memodelkan angka fertilitas total di Indonesia.

Salah satu metode yang dapat digunakan untuk memodelkan TFR adalah dengan menggunakan regresi semiparametrik spline. Dalam regresi semiparametrik spline pola data antara variabel respon dan variabel prediktor yang terbentuk merupakan gabungan dari regresi parametrik dan regresi nonparametrik. Analisis dengan menggunakan metode regresi semiparametrik spline pernah dilakukan oleh Amelia (2013) yang melakukan analisis faktor-faktor yang mempengaruhi produksi kedelai di Provinsi Jawa Timur, Sugiantari (2013) yang melakukan penelitian untuk mengetahui faktor-faktor yang mempengaruhi angka harapan hidup di Jawa Timur. Lahthifah (2013) melakukan penelitian mengenai faktor-faktor yang mempengaruhi persentase penduduk miskin di Jawa Timur dan Consetta (2013) melakukan pemodelan faktor-faktor yang mempengaruhi Angka Buta Huruf (ABH) Kabupaten/Kota di Jawa Timur.

Dalam penelitian ini digunakan pendekatan regresi semiparametrik spline, karena setelah memperhatikan *scatter plot* dari *Total Fertility Rate* (TFR) terhadap variabel prediktor sebagian diketahui dan sebagian tidak diketahui bentuk polanya. Spline adalah suatu metode dalam analisis regresi yang merupakan potongan-potongan *polynomial* yang memiliki sifat tersegmen. Pendekatan Spline memiliki kelebihan antara lain

adalah model cenderung mencari sendiri estimasinya kemanapun data tersebut bergerak. Karena didalam Spline terdapat titik knot yang merupakan titik perpaduan bersama yang menunjukkan perubahan pola perilaku data (Budiantara, 2009).

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas maka pada penelitian ini dapat dirumuskan permasalahan sebagai berikut.

1. Bagaimana karakteristik TFR di Indonesia?
2. Bagaimana memodelkan dan menentukan faktor-faktor yang berpengaruh terhadap TFR di Indonesia dengan menggunakan regresi semiparametrik spline truncated?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah maka tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Mendeskripsikan karakteristik TFR di Indonesia.
2. Memodelkan dan menentukan faktor-faktor yang mempengaruhi TFR di Indonesia dengan menggunakan regresi semiparametrik spline truncated.

1.4 Manfaat Penelitian

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

1. Bagi pemerintah Indonesia dan instansi atau lembaga yang terkait, penelitian ini dapat dijadikan sebagai salah satu acuan dalam menentukan kebijakan untuk menurunkan TFR dalam upaya pengendalian laju pertumbuhan penduduk.
2. Memberikan wawasan keilmuan kepada peneliti yang berkaitan dengan regresi semiparametrik spline.

1.5 Batasan Masalah

Batasan masalah dalam penelitian ini adalah pemilihan titik knot optimal menggunakan metode *Generalized Cross Validation* (GCV) dan menggunakan metode Regresi Semiparametrik Spline linear dengan 1 knot, 2 knot, dan 3 knot, serta kombinasi knot.

BAB II

TINJAUAN PUSTAKA

2.1 Analisis Regresi

Analisis regresi merupakan salah satu metode statistik yang menggambarkan pola hubungan antara variabel respon dan variabel prediktor serta mampu meramalkan variabel respon dengan nilai variabel prediktor yang telah diketahui. Terdapat tiga model pendekatan regresi yaitu model regresi parametrik, regresi nonparametrik, dan regresi semiparametrik. Apabila dalam analisis regresi bentuk kurva regresi diketahui maka pendekatan model regresi disebut model regresi parametrik (Budiantara, 2006). Apabila bentuk kurva regresi tidak diketahui, maka pendekatan model regresi yang sesuai digunakan adalah model regresi nonparametrik.

2.1.1 Regresi Parametrik

Regresi parametrik adalah metode statistika yang digunakan untuk mengetahui pola hubungan antara variabel prediktor dan variabel respon dimana bentuk kurva regresi diketahui. Regresi parametrik memiliki asumsi yang sangat ketat, salah satunya adalah bentuk kurva regresi harus diketahui. Jika bentuk kurva regresi yang dipilih tepat, maka regresi parametrik akan menghasilkan inferensi dan interpretasi yang lebih baik dan sederhana. Oleh karena itu, analisis regresi parametrik sering digunakan jika terdapat informasi sebelumnya tentang bentuk kurva regresinya (Eubank, 1999). Model regresi parametrik linier dengan p variabel prediktor dapat ditulis sebagai berikut :

$$y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.1)$$

dengan *error* random ε_i dan diasumsikan independen, identik, dan mengikuti berdistribusi normal dengan mean nol dan varian σ^2 (Eubank, 1999).

Secara umum bentuk regresi parametrik dapat dirumuskan dalam bentuk matrik sebagai berikut :

$$y = X\beta + \varepsilon \quad (2.2)$$

persamaan model (2.2) dapat dituliskan secara lengkap menjadi :

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ 1 & x_{31} & x_{32} & \cdots & x_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad (2.3)$$

dimana y merupakan vektor variabel respon berukuran $n \times 1$, X merupakan matriks variabel prediktor berukuran $n \times (p+1)$, β merupakan vektor parameter berukuran $(p+1) \times 1$, dan ε merupakan vektor error berukuran $n \times 1$.

2.1.2 Regresi Nonparametrik

Regresi nonparametrik digunakan jika bentuk kurva regresi tidak diketahui. Bentuk kurva regresi nonparametrik hanya diasumsikan halus (*smooth*) yang termuat di dalam suatu ruang fungsi tertentu (ruang *Hilbert*, ruang *Sobolev*, ruang *Hilbert-Sobolev*, ruang *Banach*, ruang fungsi *kontinu*, ruang *Entropi*, dan lain-lain). Regresi nonparametrik memiliki fleksibilitas yang tinggi dalam mendekati pola data (Budiantara, 2009).

Model regresi nonparametrik secara umum dapat ditulis sebagai berikut :

$$y_i = f(t_i) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.4)$$

dimana y_i adalah variabel respon dan $f(t_i)$ merupakan fungsi regresi yang tidak diketahui bentuknya dengan t_i variabel prediktor dan *error* random ε_i diasumsikan independen, identik, dan mengikuti berdistribusi normal dengan mean nol dan varian σ^2 .

2.1.3 Regresi Semiparametrik

Regresi semiparametrik adalah metode statistika yang merupakan gabungan dari komponen parametrik dan komponen nonparametrik (Ruppert, 2003). Analisis regresi semiparametrik ini digunakan ketika bentuk pola hubungan antara variabel prediktor dan variabel respon sebagian diketahui bentuknya dan sebagian lagi tidak diketahui bentuknya. Berikut ini adalah model regresi semiparametrik:

$$y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} + f(t_i) + \varepsilon_i, \quad (2.5)$$

$i = 1, 2, \dots, n$ dan X_1, X_2, \dots, X_p merupakan variabel prediktor untuk komponen parametrik dan t_i sebagai variabel prediktor untuk komponen nonparametrik.

2.2 Regresi Spline Univariabel

Analisis regresi spline univariabel merupakan regresi *spline* yang memiliki satu variabel prediktor dan satu variabel respon. Fungsi *spline* berorde p dengan r knot adalah sebagai berikut:

$$f(t) = \sum_{k=0}^p \alpha_k X^k + \sum_{j=1}^r \beta_j (X - K_j)_+^p \quad (2.6)$$

$$\text{dimana } (X - K_j)_+^p = \begin{cases} (X - K_j)^p & ; X \geq K \\ 0 & ; X < K \end{cases}$$

α_k , $k = 0, 1, \dots, p$ dan β_j , $j = 1, 2, \dots, r$ merupakan parameter dari fungsi spline, X adalah variabel prediktor, dan K_j adalah titik knot. Titik knot adalah titik perpaduan bersama dimana terjadi pola perubahan perilaku dari suatu fungsi pada sub-sub interval tertentu. Misalkan pada model regresi (2.4) kurva regresi didekati dengan fungsi *spline* $f(X)$ maka diperoleh model regresi *spline* sebagai berikut :

$$y_i = \sum_{k=0}^p \alpha_k X_i^k + \sum_{j=1}^r \beta_j (X_i - K_j)_+^p + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.7)$$

Model (2.6) dapat dijabarkan sebagai berikut ini:

$$y_i = \alpha_0 + \alpha_1 X_i + \alpha_p X_i^p + \beta_1 (X_i - K_1)_+^p + \dots + \beta_r (X_i - K_r)_+^p + \varepsilon_i \quad (2.8)$$

Berikut ini adalah model regresi spline univariabel dalam bentuk matrik:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & X_1 & \dots & X_1^p & (X_1 - K_1)_+^p & \dots & (X_1 - K_r)_+^p \\ 1 & X_2 & \dots & X_2^p & (X_2 - K_1)_+^p & \dots & (X_2 - K_r)_+^p \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_n & \dots & X_n^p & (X_n - K_1)_+^p & \dots & (X_n - K_r)_+^p \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_p \\ \beta_1 \\ \vdots \\ \beta_r \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad (2.9)$$

2.3 Regresi Spline Multivariabel

Dalam regresi spline, jika terdapat satu variabel respon dan lebih dari satu variabel prediktor maka regresi ini disebut regresi spline multivariabel (Budiantara, 2004). Diberikan data $(X_{1i}, X_{2i}, \dots, X_{pi}, y_i)$ dan hubungan antara $(X_{1i}, X_{2i}, \dots, X_{pi})$ dan y_i diasumsikan mengikuti model regresi non-parametrik $y_i = f(X_{1i}, X_{2i}, \dots, X_{pi}) + \varepsilon_i$. Jika f didekati dengan fungsi spline multivariabel aditif maka diperoleh model regresi :

$$\begin{aligned} y_i &= f(X_{1i}) + f(X_{2i}) + \dots + f(X_{pi}) + \varepsilon_i \\ &= \sum_{l=1}^p f(X_{li}) + \varepsilon_i, \end{aligned} \quad (2.10)$$

dengan $f(X_{li}) = \sum_{k=0}^m \alpha_{kl} X_{li}^k + \sum_{j=1}^r \beta_{lj} (X_{li} - K_{lj})_+^p$, dan

$$(X_{li} - K_{lj})_+^p = \begin{cases} (X_{li} - K_{lj})^p & ; X_{li} \geq K_{lj} \\ 0 & ; X_{li} < K_{lj} \end{cases}$$

$i = 1, 2, \dots, n$ dan p adalah derajat spline. Spline derajat satu sering disebut spline linier, spline derajat dua disebut spline kuadratik, sedangkan spline derajat tiga disebut spline kubik.

2.4 Pemilihan Titik Knot Optimal

Dalam regresi nonparametrik sangat penting menentukan titik knot optimal. Jika titik knot yang optimal sudah diperoleh, maka akan memberikan spline yang terbaik. Salah satu metode pemilihan titik knot optimal adalah *Generalized Cross Validation* (GCV) (Budiantara, 2000). Model spline yang terbaik didapat dari nilai GCV terkecil. Fungsi GCV dapat dituliskan seperti di bawah ini:

$$\text{GCV}(K_1, K_2, \dots, K_r) = \frac{\text{MSE}(K_1, K_2, \dots, K_r)}{\left(n^{-1} \text{trac}[I - A(K_1, K_2, \dots, K_r)] \right)^2}, \quad (2.11)$$

dimana $\text{MSE}(K_1, K_2, \dots, K_r) = n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2$ dan (K_1, K_2, \dots, K_r) adalah titik knot dan matriks $A(K_1, K_2, \dots, K_r)$ diperoleh dari persamaan $\hat{y} = A(K_1, K_2, \dots, K_r)y$.

2.5 Pengujian Parameter Model Regresi

Uji parameter dilakukan terhadap model untuk menentukan variabel prediktor yang berpengaruh terhadap variabel respon. Pengujian hipotesis yang dilakukan antara lain:

1. Uji Serentak

Uji serentak digunakan untuk mengetahui apakah parameter-parameter model regresi sudah signifikan atau belum. Uji ini dilakukan secara serentak dengan parameter yang ada dalam model. Dalam pengujian parameter regresi secara serentak digunakan *analysis of variance* (ANOVA) dari model regresi sebagaimana ditunjukkan pada Tabel 2.1 berikut.

Tabel 2.1 ANOVA

Sumber Variasi	df	Sum of Sqaure	Mean Square	F_{hitung}
Regresi	$p(m+r)$	$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	$\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{p(m+r)}$	$\frac{MS_{regresi}}{MS_{error}}$
Error	$n - p(m+r) - 1$	$\sum_{i=1}^n (y_i - \hat{y}_i)^2$	$\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - p(m+r) - 1}$	
Total	$(n-1)$	$\sum_{i=1}^n (y_i - \bar{y})^2$		

Hipotesis yang digunakan untuk uji model secara serentak sebagai berikut :

$$H_0 : \alpha_{11} = \alpha_{12} = \dots = \alpha_{pm} = \beta_{11} = \beta_{12} = \dots = \beta_{pr} = 0$$

$$H_1 : \text{Minimal ada satu } \alpha_{kj} \neq 0, \text{ atau } \beta_{kj} \neq 0$$

$$k = 1, 2, \dots, m, \quad j = 1, 2, \dots, p$$

Statistik uji yang digunakan adalah :

$$F_{hitung} = \frac{MS_{regresi}}{MS_{error}} \quad (2.12)$$

Daerah penolakan : tolak H_0 jika F -hitung lebih besar daripada F -tabel ($F_{\alpha; (p(m+r), n-p(m+r)-1)}$) atau nilai p -value $< \alpha$.

2. Uji Parsial

Uji parsial digunakan untuk mengetahui parameter yang berpengaruh signifikan secara parsial terhadap model. Hipotesis yang digunakan untuk uji secara parsial sebagai berikut:

$$H_0 : \alpha_{kj} = 0$$

$$H_1 : \alpha_{kj} \neq 0, \quad k = 1, 2, \dots, m, \quad j = 1, 2, \dots, p$$

dan

$$H_0 : \beta_{kj} = 0$$

$$H_1 : \beta_{kj} \neq 0, \quad k = 1, 2, \dots, m, j = 1, 2, \dots, p$$

Statistik uji yang digunakan adalah:

$$t_{hitung} = \frac{\hat{\beta}_{kj}}{SE(\hat{\beta}_{kj})} \quad (2.13)$$

Daerah penolakan : tolak H_0 jika $|t\text{-hitung}|$ lebih besar daripada $t\text{-tabel}$ $\left(|t_{hitung}| > t_{\left(\frac{\alpha}{2}; n-p(m+r)-1\right)} \right)$, dimana n adalah jumlah pengamatan dan p adalah banyak parameter.

2.6 Pemeriksaan Asumsi Residual dari Model Regresi Spline

Setelah didapatkan model terbaik dari regresi spline maka diperlukan beberapa pemeriksaan asumsi yang harus dipenuhi dari model tersebut. Pemeriksaan asumsi yang dilakukan dalam regresi spline ini serupa dengan pemeriksaan regresi parametrik pada umumnya. Berikut ini asumsi yang harus dipenuhi.

a. Asumsi Residual Identik

Uji asumsi identik digunakan untuk melihat homogenitas dari variansi residual. Jika terjadi pelanggaran dalam asumsi residual identik disebut heterokedastisitas yang mengakibatkan kerugian bagi efisiensi estimator (Eubank, 1988). Oleh karena itu, untuk mendeteksi adanya heterokedastisitas terdapat beberapa cara yaitu:

1. Secara Visual

Cara ini mendeteksi adanya kasus heterokedastisitas residual dengan membuat *scatter plot* antara residual dengan estimasi variabel respon (\hat{y}). Jika hasil plot menunjukkan suatu pola yang membentuk corong, atau garis melengkung maka ada indikasi adanya heterokedastisitas.

2. Dilakukan uji *glejser*

Uji *glejser* dilakukan dengan meregresikan absolut dari residual dengan variabel prediktornya (Gujarati, 2006). Jika variabel prediktornya signifikan terhadap model berarti residual sebanding dengan nilai x dimana terjadi kasus heterokedastisitas. Hipotesis yang digunakan untuk uji gletser adalah sebagai berikut :

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

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

Statistik uji yang digunakan adalah :

$$F_{hitung} = \frac{\left[\sum_{i=1}^n (|\hat{e}_i| - |\bar{e}|)^2 \right] / (k-1)}{\left[\sum_{i=1}^n (|e_i| - |\hat{e}_i|)^2 \right] / (n-k)} \quad (2.14)$$

Daerah penolakan untuk pengujian *glejser* adalah tolak H_0 jika F -hitung lebih besar daripada F -tabel ($F_{\alpha; (k-1, n-k)}$) atau nilai p -value $< \alpha$. Nilai k adalah banyaknya parameter model gletser.

b. Asumsi Residual Independen

Asumsi kedua yang harus dipenuhi adalah tidak adanya autokorelasi pada residual atau residual bersifat saling independen yang ditunjukkan oleh nilai kovarian antara ε_i dan ε_j adalah sama dengan nol. Adanya autokorelasi pada residual dapat dideteksi dengan plot ACF. Berikut ini adalah persamaan untuk ACF.

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{n-k} (e_t - \bar{e})(e_{t+k} - \bar{e})}{\sum_{t=1}^n (e_t - \bar{e})^2}, \quad k = 1, 2, 3, \dots \quad (2.15)$$

dimana

$\hat{\rho}_k$ = korelasi antara e_t dan e_{t+k}

$\hat{\gamma}_k$ = kovarian antara e_t dan e_{t+k}

$\hat{\gamma}_0 = Var(e_t) = Var(e_{t+k})$

k = lag ke- k

Batas signifikansi atas dan bawah untuk interval konfidensi adalah sebagai berikut.

$$-\frac{z_{\alpha/2}}{\sqrt{n}} < \rho_k < \frac{z_{\alpha/2}}{\sqrt{n}} \quad (2.16)$$

Adanya autokorelasi dapat diketahui jika terdapat lag yang keluar dari batas signifikansi, sehingga asumsi independen tidak terpenuhi. Sebaliknya, asumsi independen terpenuhi jika tidak terdapat lag yang keluar dari batas signifikansi (Wei, 2006).

c. Asumsi Normalitas Residual

Pengujian normalitas residual dilakukan untuk melihat apakah residual mengikuti distribusi normal atau tidak dengan menggunakan uji *Kolmogorov-Smirnov*. Hipotesis yang digunakan untuk pengujian normalitas residual sebagai berikut :

H_0 : Residual mengikuti distribusi normal

H_1 : Residual tidak mengikuti distribusi normal

Statistik uji yang digunakan adalah :

$$D = \sup_x |F_n(x) - F_0(x)| \quad (2.17)$$

Daerah penolakan untuk pengujian asumsi berdistribusi normal adalah tolak H_0 jika $D > D_\alpha$ atau nilai *p-value* $< \alpha$.

2.7 Kriteria Pemilihan Model Terbaik

Tujuan dari analisis regresi adalah memperoleh pola hubungan antara variabel respon terhadap variabel prediktor. Beberapa kriteria untuk mengetahui model regresi terbaik diantaranya adalah *Mean Square Error* (MSE), Koefisien Determinasi (R^2), *Akaike's Information Criterion* (AIC), *Schwartz's Bayesian Criterion* (SBC), dan lain sebagainya. Dalam penelitian ini dibatasi menggunakan kriteria pemilihan model terbaik dengan menggunakan nilai MSE dan R^2 . Nilai MSE adalah nilai taksiran dari varians residual, sehingga model regresi terbaik adalah model dengan MSE minimum. Koefisien determinasi adalah nilai dari proporsi keragaman total disekitar

nilai tengah \bar{y} yang dijelaskan dari model regresi (Draper dan Smith, 1992).

2.8 Fertilitas (Kelahiran)

Istilah *fertilitas* sama dengan kelahiran hidup (*live birth*), yaitu terlepasnya bayi dari rahim perempuan dengan ada tanda-tanda kehidupan, misalnya berteriak, bernafas, jantung berdenyut, dan sebagainya (Mantra, 1985). *Fertilitas* atau kelahiran merupakan salah satu faktor penambah jumlah penduduk selain migrasi masuk. Tingkat kelahiran di masa lalu mempengaruhi tingginya tingkat *fertilitas* masa kini. Jumlah kelahiran yang besar dimasa lalu akan menyebabkan bayi-bayi tersebut tetap hidup dalam jumlah yang lebih banyak dibandingkan tahun sebelumnya dan lima belas tahun kemudian akan membentuk kelompok perempuan usia subur.

Menurut Mantra (1985), terdapat sejumlah faktor yang dapat mempengaruhi *fertilitas* yang dibedakan atas faktor demografi dan faktor non demografi. Faktor demografi antara lain struktur atau komposisi umur, status perkawinan, umur kawin pertama, keperidinan atau fekunditas, dan proporsi penduduk yang kawin. Sedangkan faktor non demografi diantaranya adalah keadaan ekonomi penduduk, tingkat pendidikan, perbaikan status wanita, urbanisasi dan industrialisasi.

2.9 Age Specific Fertility Rate (ASFR)

Angka kelahiran menurut umur atau (ASFR) merupakan indikator kelahiran yang memperhitungkan pembedaan *fertilitas* dari perempuan yang terpapar untuk melahirkan, yaitu perempuan usia subur dengan memperhatikan kelompok umurnya. Secara ilmiah potensi perempuan untuk melahirkan berbeda menurut umur, dan stril setelah usia 49 tahun. Berikut ini merupakan rumus perhitungan dari ASFR :

$$ASFR_i = \frac{B_i}{P_{fi}} \times 1000$$

Dengan $ASFR_i$ merupakan *Age Specific Fertility Rate* pada kelompok umur i , B_i merupakan jumlah kelahiran pada kelompok umur i , dan P_{fi} merupakan jumlah penduduk perempuan pada kelompok umur i (BPS, 2013).

2.10 *Total Fertility Rate (TFR)*

Indonesia merupakan negara dengan yang memiliki jumlah penduduk terbesar keempat di dunia dengan sejumlah permasalahan kependudukan diantaranya adalah jumlah penduduk yang besar dan laju pertumbuhan penduduk yang cenderung meningkat dari tahun ke tahun. Oleh karena itu, pengendalian laju pertumbuhan penduduk merupakan faktor utama penunjang keberhasilan pembangunan.

Menurut Putuamar (2007), program KB memiliki posisi strategis dalam upaya pengendalian laju pertumbuhan penduduk melalui pengendalian kelahiran dan pendewasaan usia perkawinan (secara kuantitatif), maupun pembinaan ketahanan dan peningkatan kesejahteraan keluarga (secara kualitatif). Salah satu indikator yang dapat digunakan untuk mengukur keberhasilan program KB adalah Angka Fertilitas Total.

Angka *fertilitas* total atau (TFR) merupakan gambaran mengenai rata-rata jumlah anak yang dilahirkan seorang perempuan dari usia 15 sampai 49 tahun. Perbandingan angka TFR antar daerah dapat menunjukkan keberhasilan daerah dalam melaksanakan pembangunan sosial ekonominya. Angka TFR yang tinggi dapat mencerminkan rata-rata usia kawin yang rendah, tingkat pendidikan yang rendah, tingkat sosial ekonomi yang rendah atau tingkat kemiskinan yang tinggi. Berikut merupakan rumus perhitungan TFR :

$$TFR = 5 \sum_{i=15-19}^{45-49} ASFR_i$$

dimana TFR merupakan *Total Fertility Rate*, $ASFR_i$ merupakan *Age Specific Fertility Rate* pada kelompok umur i , dan i

merupakan kelompok umur, yaitu 15-19, 20-24,...., 45-49 (BPS, 2013).

2.11 Penelitian Sebelumnya

Penelitian sebelumnya tentang fertilitas pernah dilakukan oleh beberapa peneliti diantaranya yaitu, Sari (2006) melakukan pemodelan regresi spline untuk mengestimasi rata-rata angka kelahiran ASFR (*Age Spesific Fertility Rate*) di Jawa Timur, dalam penelitiannya menyebutkan bahwa faktor yang mempengaruhi rata-rata angka kelahiran ASFR (*Age Spesific Fertility Rate*) adalah umur ibu. Radifan (2010) untuk menganalisis faktor-faktor yang mempengaruhi tingkat fertilitas atau angka kelahiran total di Indonesia tahun 2007 dengan menggunakan metode *Ordinary Least Squared* (OLS), hasil penelitian menyebutkan bahwa variabel yang berpengaruh signifikan terhadap TFR di Indonesia adalah indeks pendidikan dan persentase wanita 15-49 tahun berstatus kawin yang memakai alat kontrasepsi. Sulisyarningsih (2010) melakukan pemodelan TFR dan faktor-faktor yang mempengaruhi di Provinsi Jawa Timur, hasil penelitian menyebutkan bahwa variabel yang berpengaruh signifikan terhadap TFR di Jawa Timur adalah pertumbuhan ekonomi, angka harapan hidup wanita, rata-rata lama sekolah, persentase wanita usia ≥ 15 tahun yang bekerja, persentase wanita usia 15-49 tahun berstatus kawin yang memakai alat kontrasepsi, dan persentase wanita kawin pertama usia 19-24 tahun. Imam (2014) melakukan analisis regresi probit dengan efek interaksi untuk memodelkan angka fertilitas total di Indonesia, hasil penelitian menyebutkan bahwa variabel yang berpengaruh signifikan terhadap angka fertilitas total di Indonesia adalah persentase *Contraceptive Prevalence Rate* (CPR), persentase akses terhadap media yang memuat pesan Keluarga Berencana, persentase keinginan membatasi kelahiran setelah mempunyai 2 anak, rata-rata jumlah anak ideal, dan persentase *unmet need*.

Analisis dengan menggunakan metode regresi semiparametrik spline pernah dilakukan oleh Amelia (2013) yang

melakukan analisis faktor-faktor yang mempengaruhi produksi kedelai di Provinsi Jawa Timur, Sugiantari (2013) yang melakukan penelitian untuk mengetahui faktor-faktor yang mempengaruhi angka harapan hidup di Jawa Timur. Lahthifah (2013) melakukan penelitian mengenai faktor-faktor yang mempengaruhi persentase penduduk miskin di Jawa Timur dan Consetta (2013) melakukan pemodelan faktor-faktor yang mempengaruhi Angka Buta Huruf (ABH) Kabupaten/Kota di Jawa Timur.

(Halaman ini sengaja dikosongkan)

BAB III METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari hasil Survey Demografi dan Kesehatan Indonesia (SDKI) tahun 2012. Data yang digunakan meliputi data Angka Fertilitas Total (TFR), persentase keinginan membatasi kelahiran setelah mempunyai 2 anak, persentase *unmet need*, persentase *Contraceptive Prevalence Rate* (CPR), dan persentase akses terhadap media yang memuat pesan Keluarga Berencana (KB). Adapun unit observasi yang digunakan dalam penelitian ini adalah 33 provinsi di Indonesia. Berikut adalah struktur data pada penelitian ini.

Tabel 3.1 Struktur Data Penelitian

Observasi	Y	X_1	X_2	X_3	X_4
1	Y_1	$X_{1,1}$	$X_{2,1}$	$X_{3,1}$	$X_{4,1}$
2	Y_2	$X_{1,2}$	$X_{2,2}$	$X_{3,2}$	$X_{4,2}$
3	Y_3	$X_{1,3}$	$X_{2,3}$	$X_{3,3}$	$X_{4,3}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
33	Y_{33}	$X_{1,33}$	$X_{2,33}$	$X_{3,33}$	$X_{4,33}$

3.2 Variabel Penelitian

Variabel yang digunakan untuk penelitian ini terdiri atas variabel respon dan variabel prediktor. Variabel respon (Y) adalah Angka Fertilitas Total, sedangkan variabel-variabel prediktor untuk mengetahui faktor-faktor yang berpengaruh terhadap angka fertilitas total. Pemilihan variabel-variabel tersebut sebagai variabel prediktor didasarkan pada beberapa sumber referensi atau penelitian terdahulu dimana variabel prediktor terpilih merupakan variabel prediktor yang secara statistik berpengaruh signifikan terhadap variabel respon. Secara ringkas variabel pada penelitian ini disajikan dalam Tabel 3.2 berikut.

Tabel 3.2 Variabel Penelitian

Variabel	Keterangan
Y	Angka Fertilitas Total (TFR)
X_1	Persentase keinginan membatasi kelahiran setelah memiliki 2 anak
X_2	Persentase <i>Unmet need</i>
X_3	Persentase <i>Contraceptive Prevalence Rate</i> (CPR)
X_4	Persentase akses media pesan KB

Berikut adalah penjelasan dari variabel-variabel yang digunakan dalam penelitian ini.

Y = Angka Fertilitas Total atau (TFR), yaitu rata-rata jumlah anak yang dilahirkan seorang wanita usia 15- 49 tahun.

$$TFR = 5 \sum_{i=15-19}^{45-49} ASFR_i$$

Keterangan:

$i=15-19, 20-24, \dots, 45-49$ (kelompok umur)

$$ASFR_i = \frac{\text{Jumlah kelahiran pada kelompok umur ke-}i}{\text{Jumlah perempuan pada kelompok umur ke-}i}$$

X_1 = Persentase keinginan membatasi kelahiran setelah memiliki 2 anak, yaitu persentase wanita menikah usia 15-49 tahun yang tidak ingin melahirkan anak lagi setelah memiliki 2 anak.

$$X_1 = \frac{a}{b} \times 100\%$$

Keterangan:

X_1 = Persentase keinginan membatasi kelahiran setelah memiliki 2 anak.

a = Jumlah wanita usia 15-49 tahun yang telah menikah yang tidak ingin melahirkan anak lagi setelah memiliki 2 anak.

b = Jumlah wanita usia 15-49 tahun yang telah menikah.

X_2 = Persentase *Unmet need*, yaitu persentase suatu kondisi dimana keinginan menghindari atau menunda kelahiran tidak diiringi dengan upaya penggunaan alat kontrasepsi (Ardiana, 2012). Dalam hal ini persentase *unmet need* dinyatakan dalam persentase wanita usia 15-49 tahun yang telah menikah yang ingin menghindari atau menunda kelahiran akan tetapi tidak menggunakan alat kontrasepsi.

$$\text{Persentase } unmet \text{ need} = \frac{c+d}{b} \times 100\%$$

Keterangan:

b = Jumlah wanita usia 15-49 tahun yang telah menikah.

c = Jumlah wanita usia 15-49 tahun yang telah menikah yang ingin menunda kelahiran tetapi tidak menggunakan alat kontrasepsi.

d = Jumlah wanita usia 15-49 tahun yang telah menikah yang ingin membatasi kelahiran tetapi tidak menggunakan alat kontrasepsi.

X_3 = Persentase *Contraceptive Prevalence Rate (CPR)* atau angka prevalensi pemakaian kontrasepsi, yaitu angka yang menunjukkan persentase banyaknya PUS (Pasangan Usia Subur) yang sedang menggunakan alat kontrasepsi.

$$CPR = \frac{\text{Jumlah PUS yang menggunakan alat kontrasepsi}}{\text{Jumlah PUS}}$$

X_4 = Persentase akses media pesan KB, yaitu persentase wanita usia 15-49 tahun yang telah menikah yang memperoleh informasi tentang program KB baik melalui media cetak maupun media elektronik. Media memiliki peran penting dalam mengakrabkan masyarakat dengan informasi terkini tentang KB. Media yang berperan dalam KIE (Komunikasi, Informasi, dan Edukasi) program KB meliputi radio, televisi, surat kabar atau majalah, poster, dan pamflet.

$$\text{Persentase akses media pesan KB} = \frac{e}{b} \times 100\%$$

Keterangan:

b = Jumlah wanita usia 15-49 tahun yang telah menikah.

e = Jumlah wanita usia 15-49 tahun yang telah menikah yang memperoleh informasi tentang program KB dari media.

3.3 Metode Analisis

Metode analisis yang digunakan dalam penelitian ini adalah metode statistika deskriptif, dan analisis regresi semiparametrik spline. Langkah-langkah yang dilakukan dalam menganalisis data adalah sebagai berikut.

1. Statistika Deskriptif

Melakukan analisis deskriptif untuk mengetahui karakteristik TFR di Indonesia berdasarkan variabel X_1 , X_2 , X_3 dan X_4 .

2. Analisis Regresi Semiparametrik Spline.

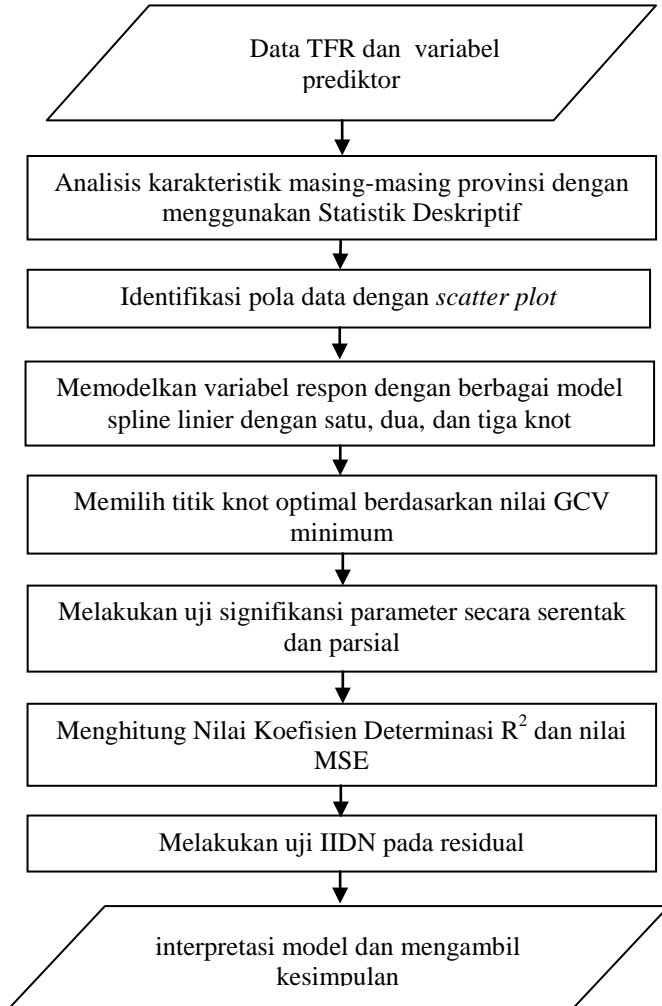
Memodelkan faktor TFR dengan pendekatan regresi semiparametrik spline. Langkah-langkah yang dilakukan dalam analisis regresi semiparametrik spline adalah sebagai berikut.

Membuat *scatterplot* antara variabel respon dengan semua variabel prediktor (X_1 , X_2 , X_3 dan X_4).

- a. Memodelkan variabel respon dengan berbagai model spline linier dengan satu knot, dua knot, tiga knot dan kombinasi knot.
- b. Memilih titik knot optimal berdasarkan nilai GCV minimum.
- d. Melakukan uji signifikansi parameter secara serentak dan parsial .
- e. Menghitung Nilai Koefisien Determinasi R^2 dan nilai MSE.
- f. Melakukan uji IIDN pada residual.
- g. Melakukan interpretasi model dan mengambil kesimpulan.

3.4 Diagram Alir Penelitian

Diagram alir tahapan penelitian dapat dijelaskan pada Gambar 3.1.



Gambar 3.1 Diagram Alir Penelitian

(Halaman ini sengaja dikosongkan)

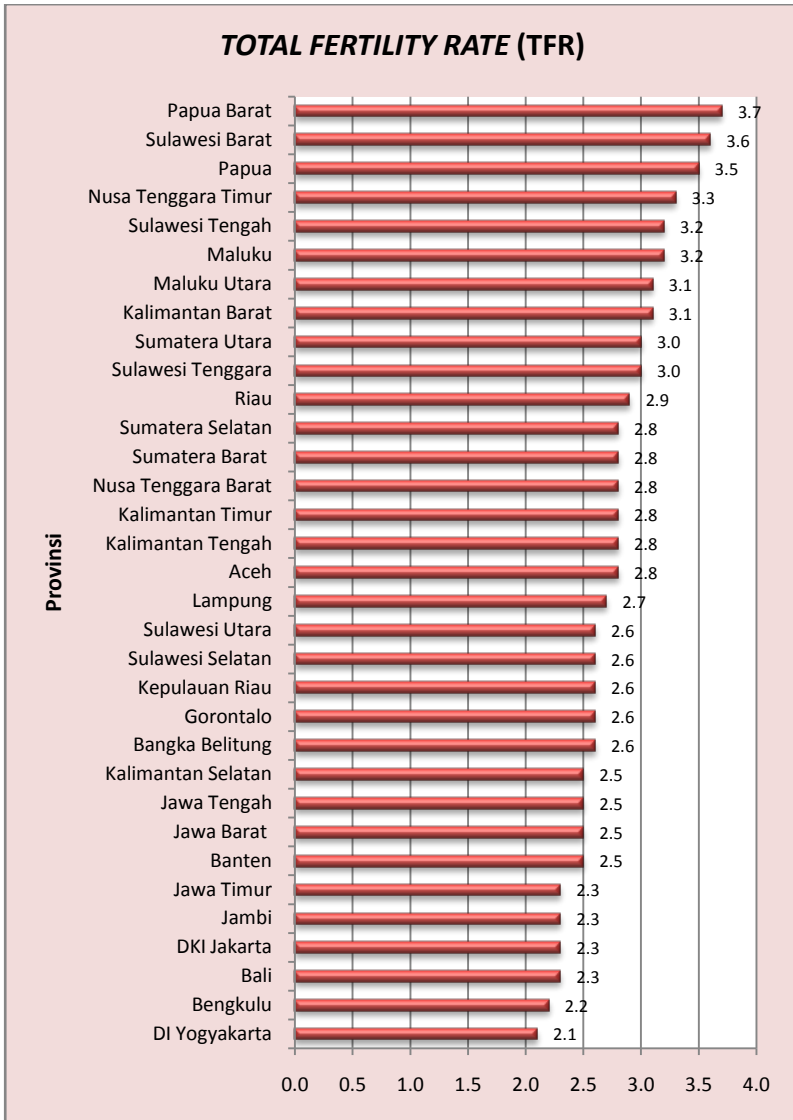
BAB IV

ANALISIS DAN PEMBAHASAN

Pada bagian ini akan dibahas karakteristik dari variabel-variabel penelitian yaitu Angka Fertilitas Total atau TFR (Y), beserta variabel-variabel yang diduga mempengaruhi. Variabel-variabel yang diduga mempengaruhi tersebut adalah persentase keinginan membatasi kelahiran setelah mempunyai 2 anak (X_1), persentase *unmet need* (X_2), persentase *Contraceptive Prevalence Rate* (X_3), dan persentase akses media KB (X_4) dengan menggunakan statistika deskriptif. Selanjutnya menyusun model regresi spline yang menyatakan pola hubungan antara angka fertilitas total dengan faktor-faktor yang mempengaruhinya. Pada penelitian ini, penggunaan nilai GCV minimum dihasilkan dari titik knot optimum dari model regresi spline linier dengan membatasi 3 titik knot.

4.1 Karakteristik *Total Fertility Rate* (TFR) dengan Faktor-Faktor yang Diduga Mempengaruhi

Angka fertilitas total secara nasional di Indonesia berdasarkan hasil Survey Demografi dan Kesehatan Indonesia (SDKI) tahun 2012 sebesar 2,6. Nilai tersebut menunjukkan rata-rata jumlah anak yang dilahirkan seorang wanita usia 15-49 tahun di Indonesia. Berdasarkan nilai tersebut dapat dikatakan bahwa penduduk di Indonesia cenderung untuk memiliki 3 anak. Jumlah tersebut tentu harus dapat diminimalkan agar laju pertumbuhan penduduk di Indonesia dapat dikendalikan. Hal tersebut dikarenakan penduduk Indonesia akan mencapai kondisi Penduduk Tumbuh Seimbang (PTS) apabila nilai angka fertilitas total dapat mencapai nilai 2,1. Dengan kata lain, rata-rata jumlah anak yang dilahirkan seorang wanita usia 15-49 tahun di Indonesia adalah 2 anak. Nilai Total Fertility Rate (TFR) pada setiap propinsi dapat dilihat pada Gambar 4.1 berikut.



Gambar 4.1 Nilai *Total Fertility Rate* (TFR) 33 Provinsi di Indonesia

Berdasarkan Gambar 4.1 diatas dapat diketahui bahwa Provinsi Papua Barat memiliki nilai TFR tertinggi, yaitu sebesar 3,7. Hal tersebut menunjukkan bahwa rata-rata jumlah anak yang dilahirkan seorang wanita usia 15-49 tahun di Provinsi Papua Barat adalah 4 anak. Sedangkan Provinsi DI Yogyakarta memiliki nilai TFR terkecil, yaitu sebesar 2,1. Hal tersebut menunjukkan rata-rata jumlah anak yang dilahirkan seorang wanita usia 15-49 tahun di Provinsi DI Yogyakarta adalah 2 anak. Dengan demikian dapat dikatakan bahwa Provinsi DI Yogyakarta telah mencapai kondisi penduduk tumbuh seimbang. Selanjutnya akan dilakukan analisis deskriptif dari variabel-variabel yang diduga berpengaruh terhadap TFR di Indonesia. Variabel-variabel prediktor tersebut adalah persentase keinginan membatasi kelahiran setelah mempunyai 2 anak (X_1), persentase *unmet need* (X_2), *Contraceptive Prevalence Rate* (X_3), dan persentase akses terhadap media yang memuat pesan Keluarga Berencana (X_4). Berikut adalah statistika deskriptif untuk keempat variabel prediktor tersebut.

Tabel 4.1 Statistik Deskriptif Variabel Prediktor

Variabel	Mean	Minimum	Maximum	Provinsi Terendah	Provinsi Tertinggi
X_1	49,81	21,4	87,9	Aceh	DI Yogyakarta
X_2	12,81	7,6	23,8	Kalimantan Tengah	Papua
X_3	58,73	21,8	70,3	Papua	Lampung
X_4	50,74	23,2	81,6	Kalimantan Barat	DKI Jakarta

Berdasarkan Tabel 4.1 diatas dapat diketahui bahwa variabel persentase keinginan membatasi kelahiran setelah mempunyai 2 anak (X_1) memiliki rata-rata sebesar 49,81. Nilai tersebut menunjukkan bahwa persentase wanita usia 15-49 tahun yang telah menikah dan tidak ingin melahirkan anak lagi setelah memiliki 2 anak secara rata-rata di 33 provinsi di Indonesia,

sebanyak 49,81%. Provinsi yang memiliki persentase keinginan membatasi kelahiran setelah mempunyai 2 anak terendah adalah Provinsi Aceh, yaitu sebesar 21,4%. Sedangkan persentase keinginan membatasi kelahiran setelah mempunyai 2 anak tertinggi adalah Provinsi DI Yogyakarta, yaitu sebesar 87,9%.

Persentase *unmet need* merupakan persentase suatu kondisi dimana keinginan untuk membatasi atau menunda kelahiran tidak diiringi dengan upaya penggunaan alat kontrasepsi. Dari Tabel 4.1 dapat diketahui bahwa rata-rata nilai persentase *unmet need* di 33 provinsi di Indonesia adalah sebesar 12,81. Nilai tersebut menunjukkan bahwa secara rata-rata di 33 provinsi di Indonesia, sebanyak 12,81% wanita usia 15-49 yang telah menikah yang ingin membatasi atau menunda kelahiran tidak menggunakan alat kontrasepsi. Provinsi dengan persentase *unmet need* tertinggi adalah Provinsi Papua, yaitu sebesar 23,8%. Sedangkan provinsi dengan persentase *unmet need* terendah adalah Provinsi Kalimantan Tengah, yaitu sebesar 7,6%.

Variabel persentase *Contraceptive Prevalence Rate* (X_3) memiliki rata-rata sebesar 58,73. Nilai tersebut menunjukkan bahwa secara rata-rata, persentase banyaknya Pasangan Usia Subur (PUS) di 33 provinsi di Indonesia yang sedang menggunakan alat kontrasepsi sebesar 58,73%. Provinsi yang memiliki persentase CPR tertinggi adalah Provinsi Lampung dengan nilai persentase CPR sebesar 70,3%. Sedangkan provinsi yang memiliki persentase CPR terendah adalah Provinsi Papua dengan nilai persentase CPR sebesar 21,8%.

Variabel persentase akses media pesan KB (X_4) memiliki rata-rata sebesar 50,74. Nilai tersebut menunjukkan bahwa persentase wanita usia 15-49 tahun yang telah menikah yang memperoleh informasi tentang program Keluarga Berencana (KB) baik melalui media cetak maupun media elektronik secara rata-rata di 33 provinsi di Indonesia, sebanyak 50,74%. Persentase akses terhadap media yang memuat pesan tentang program KB paling banyak ada di Provinsi DKI Jakarta, yaitu sebesar 81,6%. Sedangkan persentase akses terhadap media yang

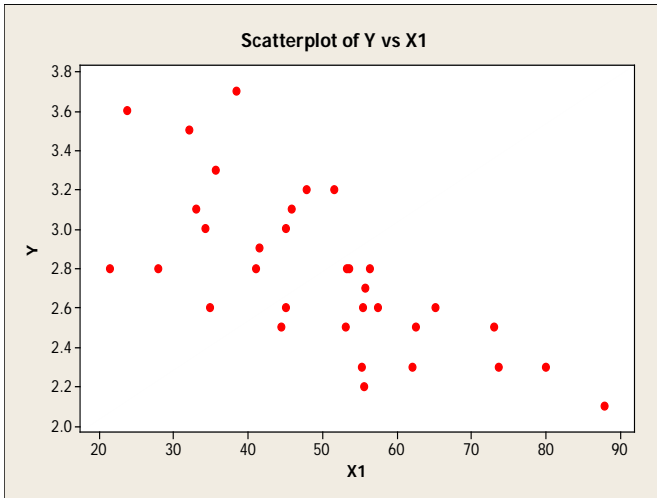
memuat pesan tentang program KB paling sedikit ada di Provinsi Kalimantan Barat, yaitu sebesar 23,2%.

4.2 Pola Hubungan *Total Fertility Rate* (TFR) dengan Faktor-Faktor yang Diduga Mempengaruhi

Dalam melakukan pemodelan TFR harus diketahui pola data dari variabel tersebut. Pola dari suatu data dapat berbentuk linear, kuadratik, kubik, polinomial derajat k , dan lain sebagainya. Jika data memiliki pola hubungan antara variabel respon dan variabel prediktor diketahui bentuknya maka dapat dilakukan pemodelan dengan menggunakan model regresi parametrik. Sebaliknya, apabila pola hubungan antara variabel respon dan variabel prediktor tidak diketahui bentuknya maka pendekatan yang dilakukan adalah dengan pendekatan regresi nonparametrik. Namun, jika hubungan data dari satu atau sebagian diantara variabel respon dan variabel prediktor diketahui dan sebagian lagi tidak, maka dilakukan pendekatan regresi semiparametrik. Berikut akan dilakukan analisa terhadap pola data antara variabel respon TFR dengan masing–masing variabel–variabel prediktor yakni persentase keinginan membatasi kelahiran setelah mempunyai 2 anak (X_1), persentase *unmet need* (X_2), persentase *Contraceptive Prevalence Rate* (X_3) dan persentase akses media pesan KB (X_4) dengan menggunakan *scatterplot*.

4.2.1 *Scatterplot* TFR dengan Persentase Keinginan Membatasi Kelahiran Setelah Mempunyai 2 Anak (X_1)

Identifikasi menggunakan *scatterplot* akan dapat diketahui pola hubungan antara variabel respon dan variabel prediktor. Variabel yang diduga mempengaruhi variabel respon TFR adalah persentase keinginan membatasi kelahiran setelah mempunyai 2 anak. Berikut ini adalah gambar yang menunjukkan pola hubungan dari variabel respon TFR dengan variabel persentase keinginan membatasi kelahiran setelah mempunyai 2 anak.

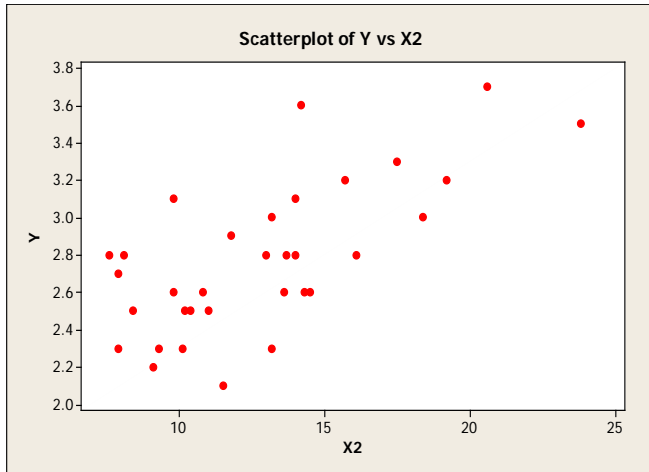


Gambar 4.2 *Scatterplot* Variabel TFR dengan Persentase Keinginan Membatasi Kelahiran Setelah Mempunyai 2 Anak

Berdasarkan Gambar 4.2 terlihat bahwa variabel respon TFR dengan variabel prediktor persentase keinginan membatasi kelahiran setelah mempunyai 2 anak berbentuk garis lurus atau menunjukkan pola hubungan yang cenderung linier. Hal ini menunjukkan semakin besar persentase keinginan membatasi kelahiran setelah mempunyai 2 anak pada suatu provinsi, maka akan mengindikasikan TFR yang relatif kecil pada provinsi tersebut. Sehingga estimasi model dilakukan dengan model parametrik linier.

4.2.2 *Scatterplot* TFR dengan Persentase *Unmet Need* (X_2)

Variabel kedua yang diduga mempengaruhi variabel respon TFR adalah persentase *unmet need*, yaitu persentase suatu kondisi dimana keinginan menghindari atau menunda kelahiran tidak diiringi dengan upaya penggunaan alat kontrasepsi. Berikut ini adalah gambar yang menunjukkan pola hubungan dari variabel respon TFR dengan variabel persentase *unmet need*.

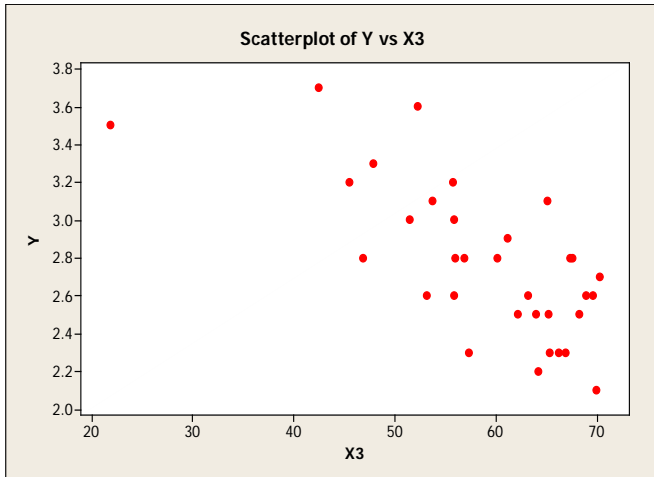


Gambar 4.3 Scatterplot Variabel TFR dengan Persentase Unmet Need

Berdasarkan Gambar 4.3 terlihat bahwa variabel respon TFR dengan variabel prediktor persentase unmet need menunjukkan pola hubungan yang membentuk suatu pola linier. Hal ini menunjukkan bahwa semakin tinggi persentase unmet need atau persentase wanita usia 15-49 tahun yang telah menikah yang ingin menghindari atau menunda kelahiran akan tetapi tidak menggunakan alat kontrasepsi pada suatu provinsi, maka semakin tinggi pula TFR di provinsi tersebut. Bentuk kurva antara variabel TFR dengan persentase unmet need diketahui, sehingga estimasi model dilakukan dengan pendekatan regresi parametrik linier.

4.2.3 Scatterplot TFR dengan Persentase Contraceptive Prevalence Rate (X_3)

Variabel ketiga yang diduga mempengaruhi variabel respon TFR adalah persentase *Contraceptive Prevalence Rate* (CPR). Berikut ini adalah gambar yang menunjukkan pola hubungan antara variabel respon TFR dengan persentase *Contraceptive Prevalence Rate* (CPR).



Gambar 4.4 Scatterplot Variabel TFR dengan Variabel Persentase Contraceptive Prevalence Rate (CPR)

Dari Gambar 4.4 terlihat bahwa dari *scatterplot* variabel respon TFR dengan variabel prediktor persentase *Contraceptive Prevalence Rate* (CPR) atau angka prevalensi pemakaian kontrasepsi, yaitu angka yang menunjukkan persentase banyaknya PUS (Pasangan Usia Subur) yang sedang menggunakan alat kontrasepsi menunjukkan pola hubungan yang tidak mengikuti suatu pola tertentu. Karena tidak diketahui bentuk pola datanya, sehingga estimasi model dilakukan dengan pendekatan regresi nonparametrik.

4.2.4 Scatterplot TFR dengan Persentase Akses Media Pesan KB (X_4)

Variabel keempat yang diduga mempengaruhi variabel respon TFR adalah persentase akses media yang Keluarga Berencana (KB). Berikut ini adalah gambar yang menunjukkan pola hubungan antara variabel respon TFR dengan persentase akses terhadap media yang memuat pesan Keluarga Berencana (KB).

minimum yaitu dengan menggunakan satu titik knot, dua titik knot, tiga titik knot, dan kombinasi antara ketiganya. Titik-titik knot tersebut terdapat pada komponen nonparametrik yaitu pada variabel X_3 dan X_4 .

4.3.1 Pemilihan Titik Knot Optimal dengan Satu Titik Knot

Kedua komponen nonparametrik dicobakan dengan menggunakan regresi spline linier satu knot terlebih dahulu. Nilai GCV dengan menggunakan satu titik knot ditunjukkan pada Tabel 4.2 berikut.

Tabel 4.2 Nilai GCV untuk Satu Titik Knot

No	X_3	X_4	GCV
1	24.7694	26.7755	0.0598
2	31.6980	35.1184	0.0633
3	22.7898	24.3918	0.0585
4	36.6469	41.0776	0.0630
5	41.5959	47.0367	0.0634
6	46.5449	52.9959	0.0635
7	51.4939	58.9551	0.0635
8	56.4429	64.9143	0.0633
9	61.3918	70.8735	0.0636
10	66.3408	76.8327	0.0634

Berdasarkan Tabel 4.2 menunjukkan bahwa pemilihan titik knot optimal dengan satu titik knot diperoleh nilai GCV minimum sebesar 0,0585. Nilai GCV tersebut bersesuaian dengan titik knot optimum, dimana letak titik knot untuk masing-masing prediktor sebagai berikut.

1. Titik knot pada variabel X_3 , $K_1 = 22.7898$
2. Titik knot pada variabel X_4 , $K_2 = 24.3918$

4.3.2 Pemilihan Titik Knot Optimal dengan Dua Titik Knot

Setelah mendapatkan knot optimum dari satu titik knot, maka selanjutnya dilakukan pemilihan titik knot optimal dengan dua titik knot. Nilai GCV dengan menggunakan dua titik knot ditunjukkan pada Tabel 4.3 berikut.

Tabel 4.3 Nilai GCV untuk Dua Titik Knot

No	X_3		X_4		GCV
	K_1	K_2	K_3	K_4	
1	25.7592	53.4735	27.9673	61.3388	0.0717
2	27.7388	68.3204	30.3510	79.2163	0.0734
3	30.7082	48.5245	33.9265	55.3796	0.0689
4	33.6776	37.6367	37.5020	42.2694	0.0681
5	35.6571	68.3204	39.8857	79.2163	0.0734
6	39.6163	42.5857	44.6531	48.2286	0.0534
7	42.5857	58.4224	48.2286	67.2980	0.0736
8	46.5449	60.4020	52.9959	69.6816	0.0746
9	51.4939	60.4020	58.9551	69.6816	0.0744
10	58.4224	62.3816	67.2980	72.0653	0.0739

Berdasarkan Tabel 4.3 diperoleh titik knot optimal dengan dua titik knot dengan nilai GCV minimum sebesar 0,0534. Nilai GCV tersebut menghasilkan titik knot optimum, dimana letak titik knot untuk masing-masing prediktor sebagai berikut.

1. Titik knot pada variabel X_3 , $K_1 = 39,6163$; $K_2 = 42,5857$
2. Titik knot pada variabel X_4 , $K_3 = 44,6531$; $K_4 = 48,2286$

Nilai GCV minimum dengan menggunakan dua knot ini lebih kecil apabila dibandingkan dengan GCV minimum yang dihasilkan oleh satu knot. Sehingga dapat dikatakan bila model yang dihasilkan oleh regresi semiparametrik spline linier dengan dua knot lebih baik daripada dengan hanya menggunakan satu knot.

4.3.3 Pemilihan Titik Knot Optimal dengan Tiga Titik Knot

Setelah mendapatkan knot optimum dari satu titik knot dan dua titik knot, maka selanjutnya dilakukan pemilihan titik knot optimal dengan tiga titik knot. Nilai GCV dengan menggunakan tiga titik knot ditunjukkan pada Tabel 4.4 berikut.

Tabel 4.4 Nilai GCV untuk Tiga Titik Knot

No	X_3			X_4			GCV
	K_1	K_2	K_3	K_4	K_5	K_6	
1	24.7694	49.5143	58.4224	26.7755	56.5714	67.2980	0.0772
2	26.7490	59.4122	64.3612	29.1592	68.4898	74.4490	0.0784
3	28.7286	31.6980	41.5959	31.5429	35.1184	47.0367	0.0732
4	30.7082	47.5347	48.5245	33.9265	54.1878	55.3796	0.0747
5	32.6878	34.6673	43.5755	36.3102	38.6939	49.4204	0.0436
6	35.6571	40.6061	68.3204	39.8857	45.8449	79.2163	0.0767
7	37.6367	40.6061	44.5653	42.2694	45.8449	50.6122	0.0634
8	39.6163	43.5755	63.3714	44.6531	49.4204	73.2571	0.0637
9	48.5245	58.4224	63.3714	55.3796	67.2980	73.2571	0.0873
10	55.4531	62.3816	63.3714	63.7224	72.0653	73.2571	0.0745

Berdasarkan Tabel 4.4 menunjukkan bahwa pemilihan titik knot optimal dengan tiga titik knot diperoleh nilai GCV minimum sebesar 0,0436. Nilai GCV tersebut dihasilkan pada titik knot optimum, dimana letak titik knot untuk masing-masing prediktor sebagai berikut.

1. Titik knot pada variabel X_3 ,
 $K_1 = 32,6878$; $K_2 = 34,6673$; $K_3 = 43,5755$
2. Titik knot pada variabel X_4 ,
 $K_4 = 36,3102$; $K_5 = 38,6939$; $K_6 = 49,4204$

Nilai GCV minimum yang dihasilkan dengan menggunakan tiga titik knot lebih kecil dibandingkan dengan menggunakan satu knot dan dua titik knot.

4.3.4 Pemilihan Titik Knot Optimal dengan Kombinasi Knot

Titik-titik knot optimum yang telah ditemukan dengan menggunakan spline satu knot, dua knot, dan tiga knot kemudian akan dikombinasikan. Kombinasi knot ini maksudnya adalah setiap variabel tidak selalu memiliki jumlah titik knot yang sama. Sehingga dengan adanya kombinasi knot ini diharapkan didapat

model terbaik. Nilai GCV untuk knot kombinasi terdapat pada Tabel 4.5 berikut.

Tabel 4.5 Nilai GCV untuk Kombinasi Knot

No	X_3	X_4	GCV
1	22.7898	36.3102	0.0554
		38.6939	
		49.4204	
2	39.6163	36.3102	0.0569
	42.5857	38.6939	
		49.4204	
3	32.6878	44.6531	0.0534
	34.6673	48.2286	
	43.5755		
4	32.6878	24.3918	0.0576
	34.6673		
	43.5755		
5	39.6163	24.3918	0.0588
	42.5857		
6	22.7898	44.6531	0.0599
		48.2286	

Nilai GCV minimum dengan kombinasi titik knot yang ditunjukkan pada Tabel 4.5 adalah sebesar 0,0534. Berdasarkan hasil pemilihan titik knot optimum, maka model regresi semiparametrik spline dengan tiga titik knot adalah model yang terbaik karena memiliki nilai GCV yang paling minimum apabila dibandingkan dengan nilai GCV minimum jika menggunakan satu, dua dan kombinasi titik knot.

4.4 Pemodelan Regresi Semiparametrik Spline dengan Titik Knot Optimal

Titik knot optimal yang diperoleh dari tiga titik knot digunakan dalam pemodelan regresi semiparametrik spline. Model regresi semiparametrik spline yang terbentuk adalah.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 + \hat{\beta}_{31}(X_3 - K_1)_+ + \hat{\beta}_{32}(X_3 - K_2)_+ + \hat{\beta}_{33}(X_3 - K_3)_+ + \hat{\beta}_4 X_4 + \hat{\beta}_{41}(X_4 - K_4)_+ + \hat{\beta}_{42}(X_4 - K_5)_+ + \hat{\beta}_{43}(X_4 - K_6)_+$$

Estimasi parameter regresi dapat dilihat pada Tabel 4.6 berikut.

Tabel 4.6 Estimasi Parameter Model Regresi Semiparametrik Spline

Variabel	Parameter	Estimasi Parameter
Constant	β_0	-1.1839
X_1	β_1	-0.0136
X_2	β_2	0.0472
X_3	β_3	0.1863
	β_{31}	1.7863
	β_{32}	-2.6154
X_4	β_{33}	0.6521
	β_4	-0.0023
	β_{41}	-0.3540
	β_{42}	0.4123
	β_{43}	-0.0701

Estimasi parameter pada Tabel 4.6 membentuk model regresi semiparametrik spline sebagai berikut.

$$\hat{y} = -1,1839 - 0,0136X_1 + 0,0472X_2 + 0,1863X_3 + 1,7863(X_3 - 32,6878)_+ - 2,6154(X_3 - 34,6673)_+ + 0,6521(X_3 - 43,5755)_+ - 0,0023X_4 + -0,3540(X_4 - 36,3102)_+ + 0,4123(X_4 - 38,6939)_+ - 0,0701(X_4 - 49,4204)_+$$

Model regresi semiparametrik spline dengan tiga titik knot ini memiliki nilai R^2 sebesar 86,5142. Dengan demikian variabilitas variabel respon atau angka kelahiran total (TFR) dapat dijelaskan model adalah sebesar 86,51%.

4.5 Pengujian Signifikansi Parameter

Pengujian signifikansi parameter ini dilakukan untuk mengetahui apakah parameter yang didapatkan dari hasil pemodelan dengan regresi semiparametrik spline signifikan terhadap TFR. Pengujian signifikansi parameter ini dimulai dengan pengujian parameter secara serentak. Apabila parameter signifikan secara serentak, maka selanjutnya dilakukan pengujian signifikansi parameter secara parsial.

4.5.1 Uji Serentak

Uji secara serentak dilakukan untuk melihat signifikansi parameter terhadap variabel respon secara keseluruhan dengan melibatkan seluruh variabel prediktor. Hasil dari uji serentak untuk model regresi semiparametrik spline dapat dilihat pada Tabel 4.7 berikut.

Tabel 4.7 ANOVA Model Regresi Semiparametrik Spline

Sumber Variasi	df	SS	MS	F _{hitung}	P-value
Regresi	10	4.4819	0.4482	14.1135	0,000
Error	22	0.6986	0.0318		
Total	32	5.1806			

Berdasarkan Tabel 4.7 diatas dapat diketahui bahwa nilai F-hitung sebesar 14,1135 dan nilai p-value sebesar 0,000. Jika dibandingkan dengan $F_{(10,22,0.05)}$ sebesar 2,30, maka diperoleh keputusan tolak H_0 karena nilai F-hitung lebih besar dari $F_{(10,22,0.05)}$. Sehingga dapat disimpulkan bahwa minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon. Untuk mengetahui variabel mana yang berpengaruh maka dilakukan pengujian secara parsial.

4.5.2 Uji Parsial

Uji parsial digunakan untuk mengetahui parameter-parameter mana yang signifikan terhadap model. Hasil uji secara parsial dapat dilihat pada Tabel 4.8 berikut.

Tabel 4.8 Uji Parsial Model Regresi Semiparametrik Spline

No	Variabel	Parameter	Estimator	P-value	Keputusan	Kesimpulan
1	Constant	β_0	-1.1839	0.0193	Tolak H_0	Signifikan
2	X_1	β_1	-0.0136	0.0003	Tolak H_0	Signifikan
3	X_2	β_2	0.0472	0.0713	Gagal Tolak H_0	Tidak Signifikan
4	X_3	β_3	0.1863	0.0001	Tolak H_0	Signifikan
		β_{31}	1.7863	0.0285	Tolak H_0	
		β_{32}	-2.6154	0.0127	Tolak H_0	
		β_{33}	0.6521	0.0034	Tolak H_0	
5	X_4	β_4	-0.0023	0.8895	Gagal Tolak H_0	Signifikan
		β_{41}	-0.3540	0.0030	Tolak H_0	
		β_{42}	0.4123	0.0021	Tolak H_0	
		β_{43}	-0.0701	0.0078	Tolak H_0	

Berdasarkan Tabel 4.8 dapat diketahui nilai p-value untuk variabel X_1 , X_3 , dan X_4 kurang dari α (0,05), maka diperoleh keputusan tolak H_0 . Sehingga dapat disimpulkan bahwa variabel persentase keinginan membatasi kelahiran setelah memiliki 2 anak (X_1), persentase *Contraceptive Prevalence Rate* (CPR) (X_3), dan persentase akses media pesan KB (X_4) berpengaruh secara signifikan terhadap variabel respon. Sedangkan untuk variabel persentase *unmet need* (X_2) tidak berpengaruh secara signifikan terhadap variabel respon karena memiliki nilai p-value lebih besar dari α (0,05). Setelah diketahui bahwa variabel persentase *unmet need* (X_2) tidak berpengaruh signifikan terhadap variabel respon, maka selanjutnya dilakukan pemodelan regresi semiparametrik spline tanpa menggunakan variabel persentase *unmet need* (X_2).

4.6 Pemilihan Titik Knot Optimal Tanpa Menggunakan Variabel Persentase *Unmet Need* (X_2)

Dengan cara yang sama pemilihan titik knot optimal dilakukan kembali tanpa menggunakan variabel persentase *unmet*

need (X_2) karena tidak berpengaruh secara signifikan terhadap variabel respon. Nilai GCV dengan menggunakan satu titik knot ditunjukkan pada Tabel 4.9 berikut.

Tabel 4.9 Nilai GCV untuk Satu Titik Knot

No	X_3	X_4	GCV
1	27.7388	30.3510	0.0683
2	32.6878	36.3102	0.0686
3	22.7898	24.3918	0.0636
4	37.6367	42.2694	0.0690
5	42.5857	48.2286	0.0687
6	47.5347	54.1878	0.0711
7	52.4837	60.1469	0.0709
8	57.4327	66.1061	0.0718
9	62.3816	72.0653	0.0724
10	66.3408	76.8327	0.0714

Berdasarkan Tabel 4.9 menunjukkan bahwa pemilihan titik knot optimal dengan satu titik knot diperoleh nilai GCV minimum sebesar 0,0636. Dimana letak titik knot untuk masing-masing prediktor sebagai berikut.

1. Titik knot pada variabel X_3 , $K_1 = 22.7898$
2. Titik knot pada variabel X_4 , $K_2 = 24.3918$

Setelah mendapatkan knot optimum dari satu titik knot, maka selanjutnya kembali dilakukan pemilihan titik knot optimal dengan dua titik knot. Nilai GCV dengan menggunakan dua titik knot ditunjukkan pada Tabel 4.10 berikut.

Tabel 4.10 Nilai GCV untuk Dua Titik Knot

No	X_3		X_4		GCV
	K_1	K_2	K_3	K_4	
1	22.7898	23.7796	24.3918	25.5837	0.0683
2	24.7694	31.6980	26.7755	35.1184	0.0733
3	26.7490	44.5653	29.1592	50.6122	0.0665
4	28.7286	61.3918	31.5429	70.8735	0.0775
5	31.6980	44.5653	35.1184	50.6122	0.0660
6	34.6673	36.6469	38.6939	41.0776	0.0703

Tabel 4.10 Nilai GCV untuk Dua Titik Knot (lanjutan)

No	X_3		X_4		GCV
	K_1	K_2	K_3	K_4	
7	39.6163	42.5857	44.6531	48.2286	0.0579
8	40.6061	47.5347	45.8449	54.1878	0.0711
9	43.5755	66.3408	49.4204	76.8327	0.0775
10	48.5245	51.4939	55.3796	58.9551	0.0814

Berdasarkan Tabel 4.10 diperoleh titik knot optimal dengan dua titik knot dengan nilai GCV minimum sebesar 0,0579. Nilai GCV tersebut menghasilkan titik knot optimum, dimana letak titik knot untuk masing-masing prediktor sebagai berikut.

1. Titik knot pada variabel X_3 , $K_1 = 39,6163$; $K_2 = 42,5857$
2. Titik knot pada variabel X_4 , $K_3 = 44,6531$; $K_4 = 48,2286$

Setelah mendapatkan knot optimum dari satu titik knot dan dua titik knot, maka selanjutnya dilakukan pemilihan titik knot optimal dengan tiga titik knot. Nilai GCV dengan menggunakan tiga titik knot ditunjukkan pada Tabel 4.11 berikut.

Tabel 4.11 Nilai GCV untuk Tiga Titik Knot

No	X_3			X_4			GCV
	K_1	K_2	K_3	K_4	K_5	K_6	
1	23.7796	48.5245	68.3204	25.5837	55.3796	79.2163	0.0800
2	24.7694	49.5143	58.4224	26.7755	56.5714	67.2980	0.0815
3	25.7592	52.4837	58.4224	27.9673	60.1469	67.2980	0.0831
4	26.7490	59.4122	68.3204	29.1592	68.4898	79.2163	0.0819
5	28.7286	31.6980	41.5959	31.5429	35.1184	47.0367	0.0778
6	29.7184	37.6367	54.4633	32.7347	42.2694	62.5306	0.0838
7	30.7082	47.5347	48.5245	33.9265	54.1878	55.3796	0.0741
8	32.6878	34.6673	43.5755	36.3102	38.6939	49.4204	0.0466
9	33.6776	47.5347	49.5143	37.5020	54.1878	56.5714	0.0762
10	35.6571	40.6061	68.3204	39.8857	45.8449	79.2163	0.0752

Berdasarkan Tabel 4.11 menunjukkan bahwa pemilihan titik knot optimal dengan tiga titik knot diperoleh nilai GCV minimum sebesar 0,0466. Nilai GCV tersebut dihasilkan pada

titik knot optimum, dimana letak titik knot untuk masing-masing prediktor sebagai berikut.

1. Titik knot pada variabel X_3 ,
 $K_1 = 32,6878$; $K_2 = 34,6673$; $K_3 = 43,5755$
2. Titik knot pada variabel X_4 ,
 $K_4 = 36.3102$; $K_5 = 38.6939$; $K_6 = 49.4204$

Nilai GCV minimum yang dihasilkan dengan menggunakan tiga titik knot lebih kecil dibandingkan dengan menggunakan satu knot dan dua titik knot. Titik-titik knot optimum yang telah ditemukan dengan menggunakan spline satu knot, dua knot, dan tiga knot selanjutnya akan dikombinasikan. Nilai GCV untuk knot kombinasi terdapat pada Tabel 4.12 berikut.

Tabel 4.12 Nilai GCV untuk Kombinasi Knot

No	X_3	X_4	GCV
1	39.6163	36.3102	0.0586
	42.5857	38.6939	
		49.4204	
2	22.7898	36.3102	0.0606
		38.6939	
		49.4204	
3	32.6878	44.6531	0.0579
	34.6673	48.2286	
	43.5755		
4	32.6878	22.7898	0.0616
	34.6673		
	43.5755		
5	39.6163	22.7898	0.0625
	42.5857		
6	22.7898	44.6531	0.0662
		48.2286	

Nilai GCV minimum dengan kombinasi titik knot yang ditunjukkan pada Tabel 4.12 adalah sebesar 0,0579. Berdasarkan

hasil pemilihan titik knot optimum, maka model regresi semiparametrik spline dengan tiga titik knot adalah model yang terbaik karena memiliki nilai GCV yang paling minimum apabila dibandingkan dengan nilai GCV minimum jika menggunakan satu, dua dan kombinasi titik knot.

4.7 Pemodelan Regresi Semiparametrik Spline dengan Titik Knot Optimal Tanpa Menggunakan Variabel Persentase *Unmet Need* (X_2)

Titik knot optimal yang diperoleh dari tiga titik knot tanpa menggunakan variabel persentase *unmet need* (X_2) digunakan dalam pemodelan regresi semiparametrik spline. Model regresi semiparametrik spline yang terbentuk adalah.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_3 X_3 + \hat{\beta}_{31}(X_3 - K_1)_+ + \hat{\beta}_{32}(X_3 - K_2)_+ + \hat{\beta}_{33}(X_3 - K_3)_+ + \hat{\beta}_4 X_4 + \hat{\beta}_{41}(X_4 - K_4)_+ + \hat{\beta}_{42}(X_4 - K_5)_+ + \hat{\beta}_{43}(X_4 - K_6)_+$$

Estimasi parameter regresi dapat dilihat pada Tabel 4.13 berikut.

Tabel 4.13 Estimasi Parameter Model Regresi Semiparametrik Spline

Variabel	Parameter	Estimasi Parameter
Constant	β_0	-1.0837
X_1	β_1	-0.0126
X_3	β_3	0.2385
	β_{31}	1.5614
	β_{32}	-2.4935
	β_{33}	0.6814
X_4	β_4	-0.0073
	β_{41}	-0.3573
	β_{42}	0.4303
	β_{43}	-0.0807

Estimasi parameter pada Tabel 4.13 membentuk model regresi semiparametrik spline sebagai berikut.

$$\begin{aligned} \hat{y} = & -1,0837 - 0,0126X_1 + 0,2385X_3 + 1,5614(X_3 - 32,6878)_+ + \\ & - 2,4935(X_3 - 34,6673)_+ + 0,6814(X_3 - 43,5755)_+ - 0,0073X_4 + \\ & - 0,3573(X_4 - 36,3102)_+ + 0,4303(X_4 - 38,6939)_+ + \\ & - 0,0807(X_4 - 49,4204)_+ \end{aligned}$$

Model regresi semiparametrik spline dengan tiga titik knot tanpa menggunakan variabel persentase *unmet need* (X_2) ini memiliki nilai R^2 sebesar 84.3129. Dengan demikian variabilitas variabel respon atau angka kelahiran total (TFR) dapat dijelaskan model adalah sebesar 84,31%.

4.8 Pengujian Signifikansi Parameter Tanpa Menggunakan Variabel Persentase *Unmet Need* (X_2)

Selanjutnya dilakukan pengujian signifikansi parameter secara serentak dan parsial kembali setelah mengeluarkan variabel persentase *unmet need* (X_2) yang tidak berpengaruh signifikan terhadap variabel respon.

4.8.1 Uji Serentak

Uji secara serentak dilakukan untuk melihat signifikansi parameter terhadap variabel respon secara keseluruhan dengan melibatkan seluruh variabel prediktor yang berpengaruh signifikan. Hasil dari uji serentak untuk model regresi semiparametrik spline dapat dilihat pada Tabel 4.14 berikut.

Tabel 4.14 ANOVA Model Regresi Semiparametrik Spline

Sumber Variasi	df	SS	MS	F _{hitung}	P-value
Regresi	9	4.3679	0.4853	13.7352	0,000
Error	23	0.8127	0.0353		
Total	32	5.1806			

Berdasarkan Tabel 4.14 diatas dapat diketahui bahwa nilai F-hitung yang diperoleh sebesar 13,7352 dan nilai p-value sebesar 0,000. Jika dibandingkan dengan $F_{(9,23,0.05)}$ sebesar 2,34, maka

diperoleh keputusan tolak H_0 karena nilai F-hitung lebih besar dari $F_{(10,22,0.05)}$. Sehingga dapat disimpulkan bahwa minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon. Untuk mengetahui variabel mana yang berpengaruh maka dilakukan pengujian secara parsial.

4.8.2 Uji Parsial

Uji parsial digunakan untuk mengetahui parameter-parameter mana yang signifikan terhadap model. Hasil uji secara parsial dapat dilihat pada Tabel 4.15 berikut.

Tabel 4.15 Uji Parsial Model Regresi Semiparametrik Spline

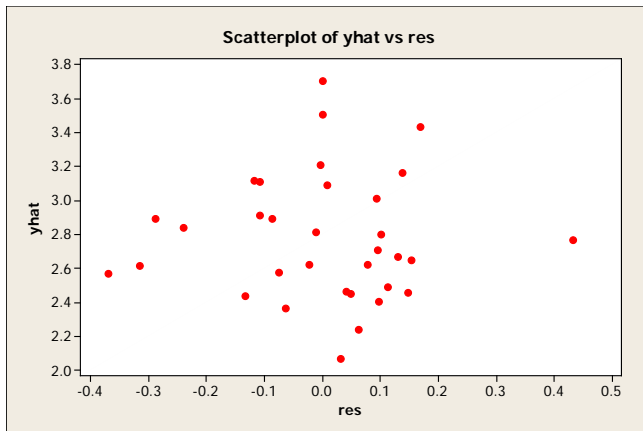
No	Variabel	Parameter Estimator	P-value	Keputusan	Kesimpulan	
1	Constant	β_0	-1.0837	0.0378	Tolak H_0	Signifikan
2	X_1	β_1	-0.0126	0.0009	Tolak H_0	
3	X_3	β_3	0.2385	0.0000	Tolak H_0	Signifikan
		β_{31}	1.5614	0.0614	Gagal Tolak H_0	
		β_{32}	-2.4935	0.0219	Tolak H_0	
		β_{33}	0.6814	0.0035	Tolak H_0	
		β_4	-0.0073	0.6707	Gagal Tolak H_0	
4	X_4	β_{41}	-0.3573	0.0039	Tolak H_0	Signifikan
		β_{42}	0.4303	0.0021	Tolak H_0	
		β_{43}	-0.0807	0.0032	Tolak H_0	

4.9 Uji Asumsi Residual

Pada analisis regresi semiparametrik spline, residual yang terbentuk harus memenuhi asumsi IIDN. Pengujian asumsi dilakukan untuk mengidentifikasi apakah residual yang terbentuk identik, independen, dan normal. Berikut adalah hasil pengujian asumsi residual.

4.9.1 Uji Asumsi Residual Identik

Asumsi yang harus terpenuhi dari analisis regresi adalah homogenitas variansi dari *error*. Untuk mengetahui apakah residual identik atau tidak, dapat dilakukan dengan uji secara visual dan uji *Glejser*. Disamping uji *Glejser*, dapat pula dilakukan secara visual, dilakukan dengan cara membuat plot antara nilai dugaan respon \hat{y} dengan residual. Apabila tidak terdapat pola pada plot maka tidak terjadi heterokedastisitas. Berikut adalah hasil plot antara nilai dugaan respond \hat{y} dan residual.



Gambar 4.6 Scatterplot antara Fits dan Residual

Berdasarkan Gambar 4.6 dapat dilihat bahwa plot menyebar tidak membentuk adanya suatu pola, sehingga secara visual asumsi identik telah terpenuhi. Selain itu untuk mengetahui terjadi heterokedastisitas dapat dilakukan dengan uji *Glejser*. Uji *Glejser* dilakukan dengan cara meregresikan nilai mutlak dari residual dengan variabel prediktor yang signifikan terhadap model. Berikut adalah hasil dari uji *Glejser*.

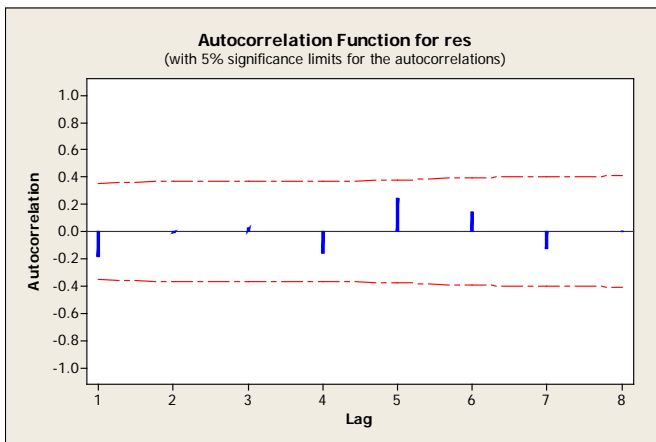
Tabel 4.16 ANOVA untuk Uji *Glejser*

Sumber Variasi	df	SS	MS	F _{hitung}	P-value
Regresi	9	0.1046	0.0116	1.0655	0.4229
Error	23	0.2509	0.0109		
Total	32	0.3556			

Berdasarkan Tabel 4.16 dapat diketahui bahwa nilai F_{hitung} uji *Glejser* sebesar 1.0655 dengan nilai p-value sebesar 0.4229 yang lebih besar dari nilai $\alpha(0.05)$. Sehingga dapat diputuskan bahwa gagal ditolak H_0 , yang berarti bahwa tidak terjadi heteroskedastisitas atau residual telah memenuhi asumsi identik.

4.9.2 Uji Asumsi Residual Independen

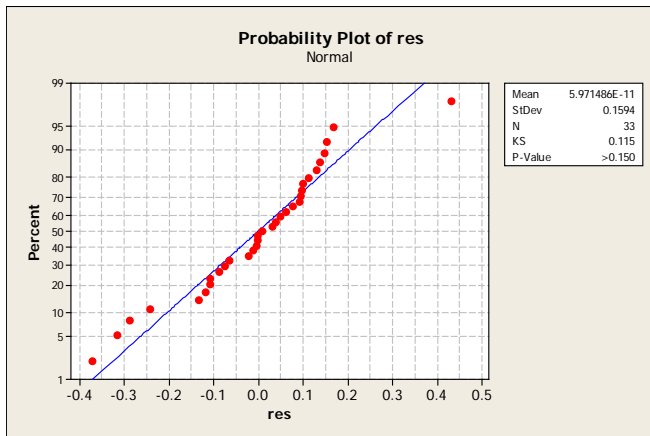
Asumsi residual selanjutnya yang harus dipenuhi dalam pemodelan regresi adalah asumsi independen pada residual. Asumsi residual independen yaitu asumsi bahwa tidak ada korelasi antar residual. Salah satu cara untuk mengetahui ada tidaknya korelasi dari suatu residual yaitu dengan melihat plot ACF atau *Autocorrelation Function*. Plot ACF dari residual dapat dilihat pada Gambar 4.7 berikut.

**Gambar 4.7** Plot ACF dari Residual

Dari Gambar 4.7 dapat dilihat bahwa tidak terdapat satupun autokorelasi yang keluar dari batas atas dan batas bawah, atau dapat dikatakan autokorelasi pada semua lag berada didalam batas. Sehingga dapat disimpulkan bahwa residual sudah memenuhi asumsi independen.

4.9.3 Uji Asumsi Distribusi Normal

Selain memenuhi asumsi identik dan independen, residual juga harus memenuhi asumsi distribusi normal. Pengujian asumsi residual berdistribusi normal dilakukan dengan menggunakan uji Kolmogorov Smirnov sebagai berikut.



Gambar 4.8 Uji Kolmogorov Smirnov

Berdasarkan Gambar 4.8 dapat diketahui bahwa nilai p-value pada uji Kolmogorov Smirnov adalah sebesar $> 0,150$. Karena nilai p-value pada uji Kolmogorov Smirnov lebih besar dari nilai α (0,05), maka dapat disimpulkan bahwa residual data telah memenuhi distribusi normal. Dengan terpenuhinya asumsi residual berdistribusi normal, maka residual dari model regresi semiparametrik spline telah memenuhi semua asumsi residual yang disyaratkan.

4.10 Interpretasi Model Regresi Semiparametrik Spline

Model regresi semiparametrik spline terbaik yang dihasilkan adalah dengan tiga titik knot setelah mengeluarkan variabel persentase *unmet need* (X_2). Berikut adalah model regresi semiparametrik spline yang terbentuk.

$$\begin{aligned}\hat{y} = & -1,0837 - 0,0126X_1 + 0,2385X_3 + 1,5614(X_3 - 32,6878)_+ + \\ & - 2,4935(X_3 - 34,6673)_+ + 0,6814(X_3 - 43,5755)_+ - 0,0073X_4 + \\ & - 0,3573(X_4 - 36,3102)_+ + 0,4303(X_4 - 38,6939)_+ + \\ & - 0,0807(X_4 - 49,4204)_+\end{aligned}$$

Dari model di atas dapat diketahui variabel yang signifikan adalah variabel persentase keinginan membatasi kelahiran setelah mempunyai 2 anak (X_1), persentase *Contraceptive Prevalence Rate* (X_3) dan persentase akses media pesan KB (X_4). Interpretasi model terhadap masing-masing variabel yang berpengaruh signifikan adalah sebagai berikut.

1. Apabila X_3 , dan X_4 diasumsikan konstan, maka model persentase keinginan membatasi kelahiran setelah mempunyai 2 anak (X_1) terhadap TFR sebagai berikut.

$$\hat{y} = -0,0126X_1$$

Dari model di atas dapat diketahui bahwa apabila persentase wanita menikah usia 15-49 tahun yang tidak ingin melahirkan anak lagi setelah memiliki 2 anak naik sebesar satu persen, maka *Total Fertility Rate* (TFR) di Indonesia akan berkurang sebesar 0,0126 persen.

2. Apabila X_1 dan X_4 diasumsikan konstan, maka model persentase *Contraceptive Prevalence Rate* (X_3) terhadap TFR sebagai berikut.

$$\begin{aligned}\hat{y} = & 0,2385X_3 + 1,5614(X_3 - 32,6878)_+ - 2,4935(X_3 - 34,6673)_+ + \\ & 0,6814(X_3 - 43,5755)_+\end{aligned}$$

$$\hat{y} = \begin{cases} 0,2385X_3 & ; & X_3 < 32,6878 \\ 1,7999X_3 - 51,0387 & ; & 32,6878 \leq X_3 < 34,6673 \\ -0,6936X_3 + 35,4042 & ; & 34,6673 \leq X_3 < 43,5755 \\ -0,0122X_3 + 5,7118 & ; & X_3 \geq 43,5755 \end{cases}$$

Dari model tersebut dapat diinterpretasikan apabila persentase *Contraceptive Prevalence Rate* (CPR) atau angka prevalensi pemakaian kontrasepsi, yaitu angka yang menunjukkan persentase banyaknya PUS (Pasangan Usia Subur) yang sedang menggunakan alat kontrasepsi kurang dari 32,6878 persen, maka apabila persentase CPR naik sebesar satu persen, maka TFR akan naik sebesar 0.2385 persen. Apabila persentase CPR antara 34,6673 persen hingga 43,5755 persen, maka apabila persentase CPR naik sebesar satu persen, maka TFR akan turun sebesar 0,6936 persen. Sedangkan apabila persentase CPR lebih dari 43,5755 persen, maka apabila persentase CPR naik sebesar satu persen, maka TFR akan turun sebesar 0,0122 persen. Sebagian besar wilayah provinsi di Indonesia memiliki persentase CPR yang lebih dari 43,5755 persen kecuali provinsi Papua Barat dan Papua.

3. Apabila X_1 dan X_3 diasumsikan konstan, maka model persentase akses media pesan KB (X_4) terhadap TFR sebagai berikut.

$$\hat{y} = -0,0073X_4 - 0,3573(X_4 - 36,3102)_+ + 0,4303(X_4 - 38,6939)_+ - 0,0807(X_4 - 49,4204)_+$$

$$\hat{y} = \begin{cases} -0,0073X_4 & ; & X_4 < 36,3102 \\ -0,3646X_4 - 12,9736 & ; & 36,3102 \leq X_4 < 38,6939 \\ 0,0657X_4 + 3,6763 & ; & 38,6939 \leq X_4 < 49,4204 \\ -0,0150X_4 - 0,3119 & ; & X_4 \geq 49,4204 \end{cases}$$

Dari model tersebut dapat diinterpretasikan apabila persentase akses media pesan KB antara 36,3102 persen hingga 38,6939

persen, maka apabila persentase akses media pesan KB naik sebesar satu persen, maka TFR akan turun sebesar 0,3646 persen. Wilayah yang termasuk dalam interval ini adalah provinsi Aceh dan Maluku Utara. Apabila persentase akses media pesan KB antara 38,6939 persen hingga 49,4204 persen, maka apabila persentase akses media pesan KB naik sebesar satu persen, maka TFR akan naik sebesar 0,0657 persen. Wilayah yang termasuk dalam interval ini adalah provinsi Sumatera Selatan, Lampung, Jawa Tengah, Nusa Tenggara Barat, Kalimantan Selatan dan Papua Barat. Sedangkan apabila persentase akses media pesan KB lebih dari 49,4204 persen, maka apabila persentase akses media pesan KB naik sebesar satu persen, maka TFR akan turun sebesar 0,0150 persen.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Kesimpulan yang dapat diambil dari penelitian ini adalah sebagai berikut.

1. Berdasarkan analisis karakteristik dapat diketahui bahwa *Total Fertility Rate* (TFR) tertinggi di Indonesia adalah provinsi Papua Barat yaitu sebesar 3,7. Sedangkan provinsi yang memiliki nilai TFR terkecil adalah DI Yogyakarta yaitu sebesar 2,1. Provinsi yang memiliki persentase keinginan membatasi kelahiran setelah mempunyai 2 anak terendah adalah Provinsi Aceh, yaitu sebesar 21,4%. Sedangkan persentase keinginan membatasi kelahiran setelah mempunyai 2 anak tertinggi adalah Provinsi DI Yogyakarta, yaitu sebesar 87,9%. Provinsi Papua Barat merupakan provinsi dengan persentase *unmet need* tertinggi yaitu sebesar 23,8%, sedangkan persentase *unmeet need* terendah terdapat pada provinsi Kalimantan Tengah, yaitu sebesar 7,6%. Provinsi Lampung merupakan provinsi yang memiliki persentase CPR tertinggi yaitu sebesar 70,3%. Sedangkan provinsi yang memiliki persentase CPR terendah adalah provinsi Papua yaitu sebesar 21,8%. Persentase akses terhadap media yang memuat pesan tentang program KB paling banyak ada di Provinsi DKI Jakarta, yaitu sebesar 81,6% dan persentase akses terhadap media yang memuat pesan tentang program KB paling sedikit ada di Provinsi Kalimantan Barat, yaitu sebesar 23,2%.
2. Model regresi yang diperoleh merupakan model regresi semiparametrik spline terbaik dengan menggunakan tiga titik knot. Model ini memiliki nilai GCV paling minimum yaitu 0,0466 dengan nilai R^2 sebesar 84,31%. Model regresi semiparametrik spline yang terbentuk adalah.

$$\begin{aligned} \hat{y} = & -1,0837 - 0,0126X_1 + 0,2385X_3 + 1,5614(X_3 - 32,6878)_+ + \\ & - 2,4935(X_3 - 34,6673)_+ + 0,6814(X_3 - 43,5755)_+ - 0,0073X_4 + \\ & - 0,3573(X_4 - 36,3102)_+ + 0,4303(X_4 - 38,6939)_+ + \\ & - 0,0807(X_4 - 49,4204)_+ \end{aligned}$$

3. Berdasarkan analisis pemodelan dengan metode regresi semiparametrik spline dapat diketahui bahwa variabel yang berpengaruh signifikan adalah variabel persentase keinginan membatasi kelahiran setelah memiliki 2 anak (X_1), persentase *Contraceptive Prevalence Rate* (CPR) (X_3), dan persentase akses media pesan KB (X_4).

5.2 Saran

Berikut adalah saran yang dapat disampaikan untuk pemerintah dan penelitian selanjutnya.

1. Berdasarkan hasil analisis dan pembahasan diharapkan pemerintah dapat menurunkan TFR dengan cara meningkatkan persentase keinginan membatasi kelahiran setelah memiliki 2 anak, persentase *Contraceptive Prevalence Rate* (CPR) dan persentase akses media pesan KB.
2. Pada penelitian selanjutnya disarankan dapat menambah variabel yang belum terdapat dalam penelitian ini yang diduga berpengaruh terhadap TFR dimana untuk saat ini masih terkendala dalam hal penyediaan data.

LAMPIRAN

Lampiran 1. Data *Total Fertility Rate* (TFR) beserta Faktor - faktor yang diduga mempengaruhi

No	Provinsi	Y	X ₁	X ₂	X ₃	X ₄
1	Aceh	2,8	21,4	14,0	46,8	38,0
2	Sumatera Utara	3,0	45,1	13,2	55,9	36,1
3	Sumatera Barat	2,8	41,1	13,7	56,9	68,4
4	Riau	2,9	41,5	11,8	61,1	58,4
5	Jambi	2,3	55,3	7,9	66,9	54,3
6	Sumatera Selatan	2,8	53,5	8,1	67,6	48,9
7	Bengkulu	2,2	55,6	9,1	64,2	59,3
8	Lampung	2,7	55,7	7,9	70,3	49,1
9	Bangka Belitung	2,6	57,4	9,8	69,6	58,9
10	Kepulauan Riau	2,6	45,1	14,5	53,1	55,9
11	DKI Jakarta	2,3	62,0	13,2	57,3	81,6
12	Jawa Barat	2,5	62,5	11,0	62,2	63,1
13	Jawa Tengah	2,5	73,0	10,4	65,2	49,0
14	DI Yogyakarta	2,1	87,9	11,5	69,9	61,0
15	Jawa Timur	2,3	73,7	10,1	65,3	52,4
16	Banten	2,5	44,4	10,2	64,0	68,5
17	Bali	2,3	80,0	9,3	66,2	51,0
18	Nusa Tenggara Barat	2,8	27,9	16,1	56,0	44,0
19	Nusa Tenggara Timur	3,3	35,7	17,5	47,9	49,7
20	Kalimantan Barat	3,1	45,8	9,8	65,1	23,2
21	Kalimantan Tengah	2,8	53,2	7,6	67,3	34,7
22	Kalimantan Selatan	2,5	53,1	8,4	68,3	44,9
23	Kalimantan Timur	2,8	56,4	13,0	60,1	56,9

No	Provinsi	Y	X ₁	X ₂	X ₃	X ₄
24	Sulawesi Utara	2,6	65,1	10,8	68,9	55,4
25	Sulawesi Tengah	3,2	47,9	15,7	55,7	59,5
26	Sulawesi Selatan	2,6	34,8	14,3	55,8	65,5
27	Sulawesi Tenggara	3,0	34,2	18,4	51,5	51,7
28	Gorontalo	2,6	55,4	13,6	63,2	56,8
29	Sulawesi Barat	3,6	23,8	14,2	52,2	35,9
30	Maluku	3,2	51,6	19,2	45,5	30,3
31	Maluku Utara	3,1	33,0	14,0	53,7	37,1
32	Papua Barat	3,7	38,4	20,6	42,5	45,9
33	Papua	3,5	32,1	23,8	21,8	29,1

Keterangan:

Variabel	Keterangan
Y	Angka fertilitas total (TFR)
X ₁	Persentase keinginan membatasi kelahiran setelah memiliki 2 anak
X ₂	Persentase <i>Unmet need</i>
X ₃	Persentase <i>Contraceptive Prevalence Rate (CPR)</i>
X ₄	Persentase akses media pesan KB

Lampiran 2. Program Regresi Semiparametrik Spline Dengan Software R

a. GCV untuk 1 knot

```
GCV1=function(data,para)
{
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[,])
  m=para+2
  dataA=data[,m:q]
  mm=ncol(dataA)
  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=mm,nrow=nk)
  for (i in (1:mm))
  {
    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=mm,nrow=p)
  data2=data[,2:q]
  a2=nrow(knot1)
  GCV=rep(NA,a2)
  Rsq=rep(NA,a2)
  for (i in 1:a2)
  {
```

```

    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=ginv(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 1 knot","\n")
cat("=====\n")
print (knot1)
cat("=====")

```

```
cat("Rsq dengan Spline linear 1 knot","\n")
cat("=====", "\n")
print (Rsq)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 1 knot","\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
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 fix.csv")
write.csv(Rsq,file="d:/output Rsq1 fix.csv")
write.csv(knot1,file="d:/output knot1 fix.csv")
}
```

b. GCV untuk 2 knot

```

GCV2=function(data,para)
{
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=para+2
  dataA=data[,m:q]
  mm=ncol(dataA)
  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=mm,nrow=nk)
  for (i in (1:mm))
  {
    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:mm))
  {
    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*mm+1)]
aa=rep(1,p)
data2=matrix(ncol=(2*mm),nrow=p)
data1=data[,m:q]
data3=data[,2:q]
a3=length(knot2[,1])
GCV=rep(NA,a3)
Rsqr=rep(NA,a3)
for (i in 1:a3)
{
  for (j in 1:(2*mm))
    {
      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 (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)
s1=min(GCV)
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")
}

```

c. GCV untuk 3 knot

```

GCV3=function(data,para)
{
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=para+2
  F=matrix(0,nrow=p,ncol=p)
  dataA=data[,m:q]
  mm=ncol(dataA)
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot=matrix(ncol=mm,nrow=nk)
  for (i in (1:mm))
  {
    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:mm))
  {
    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*mm+1)]
aa=rep(1,p)
data1=matrix(ncol=(3*mm),nrow=p)
data2=data[,m:q]
a1=length(knot1[,1])
GCV=rep(NA,a1)
Rsq=rep(NA,a1)
for (i in 1:a1)
{
  for (j in 1:(3*mm))
  {
    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=ginv(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 3 knot", "\n")
cat("=====", "\n")
print (knot1)
cat("=====", "\n")
cat("Rsq dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (Rsq)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
cat(" GCV =",s1, "\n")
write.csv(GCV,file="d:/output GCV3fx.csv")
write.csv(Rsq,file="d:/output Rsq3fx.csv")
write.csv(knot1,file="d:/output knot3fx.csv")
}

```

d. GCV untuk Knot Kombinasi

```

GCVkom=function(data,para)
{
  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")
  n2=nrow(x1)
  a=matrix(nrow=2,ncol=3^2)
  m=0
  for (i in 1:3)
  for (j in 1:3)
  {
    m=m+1
    a[,m]=c(i,j)
  }
  a=t(a)
  GCV=matrix(nrow=nrow(x1),ncol=3^2)
  for (i in 1:3^2)
  {
    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)
{
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
{
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]
}
}
ma=as.matrix(cbind(aa,bb))
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))
  {
    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)
spl=x2[,4:6]
kkk=cbind(sp,spl)
cat("=====", "\n")
print(i)
print(kkk)
print(Rsq)
}
write.csv(GCV,file="d:/GCVkom.csv")
write.csv(Rsq,file="d:/Rsqkom.csv")
}
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
A1=(F-A)
A2=(sum(diag(A1))/p1)^2
GCV[h,i]=MSE/A2
}

```



```
}  
for (i in 1:3^2)  
{  
  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]  
  kkk=cbind(sp,spl)  
  cat("=====", "\n")  
  print(i)  
  print(kkk)  
  print(Rsq)  
}  
write.csv(GCV,file="d:/GCVkom.csv")  
write.csv(Rsq,file="d:/Rsqkom.csv")  
}
```

Lampiran 3. Uji Signifikansi Parameter

```

uji=function(data,knot,alpha,para)
{
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])
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:m],data.knot[,1:3],data[,m+1],data.knot[,4:6])
mx=as.matrix(mx)
B=(ginv(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/SST)*100

#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 t (uji individu)
thit=rep(NA,n1)
pval=rep(NA,n1)
SE=sqrt(diag(MSE*(ginv(t(mx)%*%mx))))

```

```

cat("-----", "\n")
cat("Kesimpulan hasil uji individu", "\n")
cat("-----", "\n")
thit=rep(NA,n1)
pval=rep(NA,n1)
for (i in 1:n1)
{
thit[i]=B[i,1]/SE[i]
pval[i]=2*(pt(abs(thit[i]),(p-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")
}
thit=as.matrix(thit)
cat("=====","\n")
cat("nilai t hitung", "\n")
cat("=====","\n")
print (thit)
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")
write.csv(res,file="d:/output residual.csv")
write.csv(pval,file="d:/output pvalue.csv")
write.csv(mx,file="d:/output mx.csv")
write.csv(yhat,file="d:/output yhat.csv")
}

```

Lampiran 4. Uji Glejser

```

glejser=function(data,knot,res,alpha,para)
{
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])
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:m],data.knot[,1:3],data[,m+1],data.knot[,4:6])
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)
Rsqr=(SSR/SST)*100

#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
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")
}
```

Lampiran 5. Output Uji Signifikansi Parameter Model Regresi Semiparametrik

=====
 Estimasi Parameter
 =====

[,1]

[1,] -1.083713961

[2,] -0.012579509

[3,] 0.238526076

[4,] 1.561407928

[5,] -2.493516170

[6,] 0.681431752

[7,] -0.007297329

[8,] -0.357287628

[9,] 0.430273256

[10,] -0.080711404

 Kesimpulan hasil uji serentak

Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan

 Kesimpulan hasil uji individu

Tolak Ho yakni prediktor signifikan dengan pvalue 0.03776034

Tolak Ho yakni prediktor signifikan dengan pvalue 0.0008878336

Tolak Ho yakni prediktor signifikan dengan pvalue 5.337208e-08

Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue
 0.06139196

Tolak Ho yakni prediktor signifikan dengan pvalue 0.02193758

Tolak Ho yakni prediktor signifikan dengan pvalue 0.003463885
 Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.6706829

Tolak Ho yakni prediktor signifikan dengan pvalue 0.003998944

Tolak Ho yakni prediktor signifikan dengan pvalue 0.002086418

Tolak Ho yakni prediktor signifikan dengan pvalue 0.003234438

=====
 nilai t hitung
 =====

[,1]

[1,] -2.2045389

[2,] -3.8158148

[3,] 7.8983896

[4,] 1.9667439

[5,] -2.4578048

[6,] 3.2578597

[7,] -0.4307182

[8,] -3.1978731

[9,] 3.4675864

[10,] -3.2863870

Analysis of Variance

=====

Sumber	df	SS	MS	Fhit
Regresi	9	4.367918	0.4853242	13.73523
Error	23	0.8126881	0.03533427	
Total	32	5.180606		

 =====

s= 0.1879741 Rsq= 84.31288

pvalue(F)= 2.671892e-07

Lampiran 6. Output Uji Glejser

 Kesimpulan hasil uji serentak

Gagal Tolak Ho yakni semua prediktor tidak berpengaruh signifikan atau tidak terjadi heteroskedastisitas

Analysis of Variance

```
=====
Sumber      df      SS      MS      Fhit
-----
Regresi     9  0.1046091  0.01162323  1.065519
Error       23  0.250896  0.01090852
Total       32  0.3555051
=====
```

s= 0.1044439 Rsq= 29.42549

pvalue(F)= 0.4228646

(Halaman ini sengaja dikosongkan)

DAFTAR PUSTAKA

- Al Fattah, I.A. 2014. *Analisis Regresi Probit Dengan Efek Interaksi Untuk Memodelkan Angka Fertilitas Total di Indonesia*. Surabaya : Tesis Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Teknologi Sepuluh Nopember.
- Amalia, D. 2013. *Analisis Faktor-Faktor Yang Mempengaruhi Produksi Kedelai di Provinsi Jawa Timur Dengan Metode Regresi Semiparametrik Spline*. Surabaya: Tugas Akhir Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Teknologi Sepuluh Nopember.
- BAPPENAS, BPS, & UNFPA. 2005. *Proyeksi Penduduk Indonesia 2000-2025*. Jakarta: Badan Perencanaan Pembangunan Nasional, Badan Pusat Statistik, United Nations Population Fund.
- BPS, BKKBN, & KEMENKES. 2013. *Survey Demografi dan Kesehatan Indonesia 2012*. Jakarta: Badan Pusat Statistik, Badan Kependudukan dan Keluarga Berencana Nasional, Kementerian Kesehatan.
- BPS. 2013. Statistik Indonesia 2013. <http://www.bps.go.id>, 24 August, 2014, 09:00
- BPS. 2014. Angka Fertilitas Total. <http://www.datastatistik-indonesia.com/portal>, 24 August, 2014, 08:04
- Budiantara, I N. 2000. Metode U, GLM, CV, dan GCV dalam Regresi Nonparametrik *Spline*. *Majalah Ilmiah Himpunan Matematika Indonesia (MIHMI)* , 6,41-45.
- Budiantara, I N. 2004. Model *Spline* Multivariabel dalam Regresi Nonparametrik. *Makalah Seminar Nasional Matematika, Jurusan Matematika, ITS, Surabaya*.

- Budiantara, I N. 2006. Model *Spline* dengan Knot Optimal. *Jurnal Ilmu Dasar* , Vol 7,77-85.
- Budiantara, I N. 2009. "*Spline dalam Regresi Nonparametrik dan Semiparametrik : Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang*", Pidato Pengukuhan untuk Jabatan Guru Besar dalam Bidang Ilmu Matematika Statistika dan Probabilitas, pada Jurusan Statistika, Fakultas MIPA. Surabaya: ITS Press.
- Consetta, E.G. 2013. *Pemodelan Faktor-Faktor yang Mempengaruhi Angka Buta Huruf (ABH) Kabupaten/Kota di Jawa Timur dengan Regresi Spline Semiparametrik*. Surabaya: Tugas Akhir Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Teknologi Sepuluh Nopember.
- Draper, N.R., and Smith, H. 1992, *Analisis Regresi Terapan*, PT Gramedia Pustaka Utama Jakarta.
- Eubank, R. 1988. *Spline Smoothing and Nonparametric Regression*. New York: Marcel Deker.
- Eubank, R. 1999. *Nonparametric Regression and Spline Smoothing 2nd Edition*. New York: Marcel Deker.
- Gujarati, D. N. 2006. *Dasar-Dasar Ekonometrika Edisi Pertama*, Jakarta: Erlangga.
- Lalhthifah, S. 2013. *Faktor-Faktor yang Mempengaruhi Persentase Penduduk Miskin dengan Pendekatan Regresi Semiparametrik Spline di Jawa Timur*. Surabaya: Tugas Akhir Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Teknologi Sepuluh Nopember.
- Mantra, I. B. 1985. *Demografi Umum*. Yogyakarta: Pustaka Pelajar.
- Putuamar, H.F. (2007). *37 Tahun Program KB*. Cirebon: Pikiran Rakyat.

- Radifan, M. 2010. *Analisis faktor-faktor yang mempengaruhi fertilitas di Indonesia*. Medan : Skripsi Universitas Sumatra Utara.
- Ruppert, M., & Carroll, R. 2003. *Semiparametric Regression*. New York: Cambridge University Press.
- Saifuddin, A.B., & Affandi, B. (2003). *Buku Panduan Praktis Pelayanan Kontrasepsi*. Jakarta: Yayasan Bina Pustaka Sarwono Prawirohardjo.
- Sari, D. N. 2006. *Model Regresi Spline Untuk Mengestimasi Rata-rata Angka Kelahiran ASFR (Age Specific Fertility Rate) Di Jawa Timur*. Surabaya: Tugas Akhir Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Teknologi Sepuluh Nopember.
- Sugiantari. 2013. *Analisis Faktor-Faktor yang Mempengaruhi Angka Harapan Hidup di Jawa Timur Menggunakan Regresi Semiparametrik Spline*. Surabaya: Tugas Akhir Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Teknologi Sepuluh Nopember.
- Sulistyaningsih, S. 2006. *Pemodelan Total Fertility Rate Dan Faktor-faktor yang Mempengaruhi Di Provinsi Jawa Timur*. Surabaya: Tugas Akhir Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Teknologi Sepuluh Nopember.

(Halaman ini sengaja dikosongkan)

BIODATA PENULIS



Penulis memiliki nama lengkap Khusniyah lahir di Surabaya, pada tanggal 06 Juli 1982. Penulis merupakan anak ketiga dari empat bersaudara. Penulis telah menempuh pendidikan formal yaitu di SDI Yapita Surabaya, SMP Negeri 30 Surabaya, dan SMA Negeri 10 Surabaya. Setelah lulus dari SMA penulis diterima di Program Studi Diploma III Jurusan Statistika ITS pada Tahun 2000 dengan NRP. 1300030013. Selanjutnya pada Semester Genap Tahun 2012 penulis diterima di Program Studi S1 Lintas Jalur Jurusan Statistika FMIPA ITS dengan NRP. 1312106006. Segala saran dan kritik yang membangun selalu penulis harapkan untuk kebaikan kedepannya. Penulis dapat dihubungi melalui email: khusniyah@statistika.its.ac.id atau gmkhusniyah@gmail.com.

(Halaman ini sengaja dikosongkan)