TESIS - SS14 2501

INFERENSI STATISTIK UNTUK KURVA REGRESI NONPARAMETRIK *SPLINE* KUADRATIK DAN APLIKASINYA PADA DATA ASFR (*AGE SPESIFIC FERTILITY RATE*) DI BALI

IDA AYU SEVITA INTANSARI
NRP. 1311 201 905

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

PROGRAM MAGISTER
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2016

TESIS - SS14 2501

STATISTICAL INFERENCE FOR QUADRATIC SPLINE NONPARAMETRIC REGRESSION CURVES AND THE APPLICATION ON ASFR (AGE SPESIFIC FERTILITY RATE) IN BALI

IDA AYU SEVITA INTANSARI
NRP. 1311 201 905

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

MAGISTER PROGRAM
DEPARTMENT OF STATISTIC
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
SEPULUH NOPEMBER INSTITUTE OF TECHNOLOGY
SURABAYA
2016

INFERENSI STATISTIK UNTUK KURVA REGRESI NONPARAMETRIK *SPLINE* KUADRATIK DAN APLIKASINYA PADA ASFR (*AGE SPESIFIC FERTILITY RATE*) DI BALI

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Sains (M.Si)
di
Institut Teknologi Sepuluh Nopember
Oleh :

IDA AYU SEVITA INTANSARI
NRP. 1311 201 905

Tanggal Ujian : 25 Januari 2016
Periode Wisuda : Maret 2016

Disetujui Oleh :

1. Prof. Dr. Drs. I Nyoman Budiantara, M.Si (Pembimbing)
NIP. 19650603 198903 1 003

2. Dr. Dra. Ismaini Zain, M.Si (Penguji)
NIP. 19600525 198803 2 001

3. Dr. Vita Ratnasari, S.Si, M.Si (Penguji)
NIP. 19700910 199702 2 001

Direktur Program Pascasarjana

Prof. Ir. Djauhar Manfaat, M.Sc, Ph.D
NIP 19601202 198701 1 001

INFERENSI STATISTIK UNTUK KURVA REGRESI NONPARAMETRIK SPLINE KUADRATIK DAN APLIKASINYA PADA ASFR (*AGE SPESIFIC FERTILITY RATE*) DI BALI

Nama	**:** Ida Ayu Sevita Intansari
NRP	**:** 1311 201 905
Jurusan	**:** Statistika FMIPA-ITS
Dosen Pembimbing	**:** Prof. Dr. Drs. I Nyoman Budiantara, M.Si

ABSTRAK

Permasalahan krusial yang dihadapi oleh negara-negara berkembang di dunia, khususnya negara-negara berpenduduk besar dan padat adalah laju pertumbuhan penduduk. Salah satu upaya yang dilakukan oleh Negara Republik Indonesia untuk menurunkan fertilitas wanita (jumlah anak yang dilahirkan hidup) adalah melalui program Keluarga Berencana (KB). Ukuran dasar fertilitas yang sering digunakan adalah ASFR (Age Spesific Fertility Rate), dimana ASFR sebagai angka yang menunjukkan banyaknya kelahiran selama setahun pada setiap 1000 wanita pada kelompok usia tertentu. Pemodelan data ASFR dengan regresi parametrik belum tentu sesuai diterapkan karena pola hubungan fertilitas memiliki pola yang berubah-ubah pada interval umur tertentu. Regresi Nonparametrik *Spline* adalah metode regresi yang dapat menangani data yang polanya berubah-ubah pada interval tertentu. *Spline* merupakan salah satu model dalam regresi nonparametik yang mempunyai interpretasi statistik secara visual yang sangat khusus dan sangat baik. Disamping itu, spline juga mampu menangani karakter data atau fungsi yang bersifat halus (*smooth*). Penelitian ini bertujuan menurunkan bentuk estimator dan interval konfidensi terpendek untuk model *spline* kuadratik dan memodelkan data ASFR di provinsi Bali.

Kata Kunci: ASFR, Fertilitas, Regresi Nonparametrik *Spline* Kuadratik, Inferensi Statistik.

STATISTICAL INFERENCE FOR QUADRATIC SPLINE NONPARAMETRIC REGRESSION CURVES AND THE APPLICATION ON ASFR (AGE SPESIFIC FERTILITY RATE) IN BALI

By	**:** Ida Ayu Sevita Intansari
Student Identity Number	**:** 1311 201 905
Supervisor	**:** Prof. Dr. Drs. I Nyoman Budiantara, M.Si

ABSTRACT

Crucial issues faced by developing countries in the world, especially countries with large and dense is the population growth rate. One of the efforts undertaken by the Republic of Indonesia to decrease women's fertility (number of children born alive) is through the Keluarga Berencana (KB) program. Basic measure of fertility is often used ASFR (Age Specific Fertility Rate), which ASFR as figures show the number of births during a year in every 1000 women in certain age groups. ASFR data modeling with the parametric regression is not necessarily appropriate to be applied because of the pattern of fertility relationship has a pattern changing at intervals of a certain age. Spline Nonparametric regression is a regression method that can handle data pattern changing at certain intervals. Is one of the spline regression models in statistical interpretation nonparametik that have visually very special and very good. In addition, the spline is also able to handle character data or functions that are smooth (smooth). This research aims to lower forms of estimators and confidence intervals for the shortest quadratic spline models and modeling the data ASFR in the province of Bali.

Keywords : *ASFR, Fertility, Quadratic Spline Nonparametric Regression, Statistical Inference.*

KATA PENGANTAR

Om Swastyastu,

Puji syukur kehadirat Ida Sang Hyang Widhi Wasa, Tuhan Yang Maha Esa atas berkat, rahmat, dan kelancaran yang diberikan kepada penulis, sehingga penulis dapat menyelesaikan proposal Tesis yang berjudul

INFERENSI STATISTIK UNTUK KURVA REGRESI NONPARAMETRIK SPLINE KUADRATIK DAN APLIKASINYA PADA ASFR (*AGE SPESIFIC FERTILITY RATE*) DI BALI

Keberhasilan dalam penyelesaian Tesis ini tidak lepas dari bantuan, arahan, bimbingan, serta dukungan dari berbagai pihak. Untuk itu, penulis tidak lupa menyampaikan ucapan terima kasih sebesar-besarnya kepada:

1. Bapak Dr. Suhartono, M.Sc selaku Ketua Jurusan dan Ketua Program Studi Pasca Sarjana Statistika-FMIPA ITS Surabaya.
2. Bapak Prof. Dr. Drs. I Nyoman Budiantara, M.Si, selaku dosen pembimbing Tesis atas ilmu, arahan, bimbingan, waktu serta semangat yang diberikan dalam diskusi yang telah dilalui dalam penyusunan proposal sehingga penulis dapat menyelesaikan proposal Tesis dengan baik.
3. Ibu Dr. Dra. Ismaini Zain, M.Si dan Ibu Dr. Vita Ratnasari, S.Si., M.Si selaku dosen penguji Tesis saya, terima kasih atas semua kritik dan masukan untuk Tesis saya sehingga dapat selesai dengan hasil yang baik.
4. Bapak Dr. Ir. Setiawan, M.Sc selaku dosen wali dari penulis.
5. Pemerintah, pimpinan Dikti, dan Direktur Pascasarjana ITS, Bapak Prof. Ir. Djauhar Manfaat, M.Sc, Ph.D melalui beasiswa *Fast Track* yang telah mendukung penulis untuk menyelesaikan studi Magister di Statistika ITS.
6. Segenap dosen dan karyawan di Jurusan Statistika ITS yang telah memberikan banyak sekali bantuan selama penulis berkuliah.
7. BKKBN Jatim dan Bali yang telah membantu dalam penyediaan data sehingga penelitian saya dapat selesai dan berguna bagi masyarakat, dan khususnya gek Anggie.

8. Bapak Ida Bagus Swela dan Ibu Bekti Sriyanti, orang tua tercinta atas kerja keras, kasih sayang sepanjang masa, doa yang tak ada hentinya serta dukungan dan nasehatnya untuk tetap berjalan pada kebaikan untuk menyelesaikan pendidikan Sarjana sampai Magister di ITS Surabaya.
9. Adikku, Ida Bagus Raditya Dewangga dan Ida Bagus Irenata Iswara serta segenap keluarga besar atas doa, dukungan dan semangatnya.Dan special buat Ibu(Eyang) dan Nenek di Bali terimakasih atas semua do'anya..
10. Calon suami tercinta, Ida Bagus Brahma Mahaputra yang selalu setia menemani, memberikan motivasi, dan mendoakan penulis.
11. Ida Bagus Oka Ari Adnyana, adek angkatan 2009 yang sudah membantu banyak dalam hal komputasi dan pembuatan program dalam Tesis ini.
12. Masnatul Laili dan Millatur Rodliyah, room mate saat seminar internasional di Yogyakarta yang sudah banyak membantu memberi informasi dan berjuang bersama untuk wisuda Maret ini.
13. Downline-downline di Vivacity "Happy" Team dan upline serta crossline di komunitas Glamorbiz terima kasih atas dukungan dan motivasinya.
14. Serta semua pihak yang turut berjasa dan tidak dapat disebutkan satu per satu.

Penulis menyadari sepenuhnya bahwa penyusunan Tesis ini masih jauh dari kesempurnaan, kritik maupun saran yang sifatnya membangun sangat diharapkan sebagai masukan dalam penelitian selanjutnya. Semoga penelitian ini bermanfaat bagi pembaca.

Om Shanti, Shanti, Shanti Om

Surabaya, 22 Januari 2016

Penulis

DAFTAR ISI

Halaman

DAFTAR GAMBAR

Halaman

DAFTAR TABEL

Halaman

`BAB 1
PENDAHULUAN

1.1 Latar Belakang

Laju pertumbuhan penduduk merupakan permasalahan krusial yang dihadapi oleh negara-negara berkembang di dunia. Salah satu upaya yang dilakukan oleh Indonesia untuk menurunkan fertilitas wanita (jumlah anak yang dilahirkan hidup) melalui program Keluarga Berencana (KB). Untuk itu Badan Koordinasi Keluarga Berencana Nasional (BKKBN) bekerjasama dengan Biro Pusat Statistik (BPS) melaksanakan Survey Sosial Ekonomi Nasional (SUSENAS) untuk mengetahui jumlah angka kelahiran di Indonesia. Dari data survey tersebut diharapkan diperoleh gambaran bagaimana model pola fertilitas wanita yang representatif untuk Indonesia. Dalam teori demografi, ASFR (*Age Spesific Fertility Rate*) termasuk dalam ukuran demografi yaitu "*rate*", dimana ASFR merupakan fungsi umur yang sering digunakan sebagai ukuran dasar fertilitas. ASFR adalah angka yang menunjukkan banyaknya kelahiran selama setahun pada setiap 1000 wanita pada kelompok usia tertentu.

Jika dibahas secara teori probabilistik, ASFR merupakan fungsi kuadratik, karena pola grafiknya yang berbentuk seperti gunung (*U-shape* terbalik) dan hampir seperti bentuk kurva distribusi normal tetapi tidak simetris. Di dalam statistika, fungsi kuadratik berbentuk *U-shape* terbalik, pada kasus data ASFR bentuk kurvanya cenderung kuadratik tetapi bentuknya agak condong ke kanan, tidak *smooth* dan tidak *truncated*,

Kurva yang belum tentu smooth bentuknya, perlu dibuktikan dan dianalisis menggunakan metode yang dapat secara peka mendeteksi adanya titik-titik yang merupakan potongan-potongan tersebut. Salah satu metode yang dapat digunakan untuk menyelesaikan masalah pemodelan adalah analisis regresi. Analisis regresi merupakan salah satu metode yang digunakan untuk mengetahui pola hubungan antara variabel respon dengan variabel prediktor. Pendekatan dalam metode analisis regresi ada tiga yaitu pendekatan parametrik, pendekatan nonparametrik, dan pendekatan semiparametrik. Jika suatu data bentuk kurva regresinya membentuk pola yang diketahui seperti bentuk linear, kuadratik, kubik maupun polynomial derajat *k* maka pendekatan yang tepat digunakan adalah pendekatan parametrik. Pendekatan regresi nonparametrik digunakan apabila bentuk kurva regresi dari suatu data tidak diketahui bentuknya. Sedangkan pendekatan semiparametrik digunakan apabila bentuk kurva regresi sebagian diketahui polanya dan sebagian lagi tidak diketahui polanya. Data yang terdapat di lapangan tidak semua diketahui pola hubungannya, sehingga pendekatan regresi nonparametrik dapat digunakan untuk penyelesaian permasalahan tersebut. Menurut Eubank (1988) mengatakan bahwa dalam pendekatan regresi nonparametrik data diharapkan mencari sendiri bentuk estimasinya tanpa dipengaruhi oleh subyektifitas peneliti. Dengan kata lain pendekatan regresi nonparametrik memiliki fleksibelitas yang tinggi. Pendekatan regresi nonparametrik spline lebih cocok diaplikasikan pada data ASFR dibandingkan dengan pendekatan parametrik

Beberapa pendekatan regresi nonparametrik yang sering digunakan adalah Deret *Fourier*, *Kernel*, *Spline*, dan *Wavelets*. Pendekatan *Spline* mempunyai kelebihan diantaranya model cenderung mencari sendiri estimasinya kemanapun data tersebut bergerak. Kelebihan ini diakibatkan karena dalam *Spline* terdapat titik knot yang merupakan titik perpaduan bersama yang menunjukkan terjadinya perubahan pola perilaku data (Budiantara, 2009). Pada umumnya *Spline* merupakan suatu estimator yang diperoleh dengan meminimumkan *Penalized Least Square* (PLS) yaitu kriteria optimasi penggabungan antara *goodness of fit* dan kemulusan kurva. Penyelesaian optimasi ini diperoleh dengan metode RKHS atau *Gateaux* yang secara matematika relatif sulit untuk dilakukan. Untuk mengatasinya digunakan metode estimator lain yang lebih sederhana yaitu dengan optimasi *Least Square* (LS), yang diharapkan memperoleh perhitungan matematik dan interpretasi matematik yang relatif lebih mudah dan sederhana (Budiantara, 2006). Inferensi statistik yang sangat penting dalam regresi *spline* adalah interval konfidensi dengan konsep interval konfidensi terpendek dengan *Pivotal Quantity*.

Penelitian tentang fertilitas telah dilakukan oleh beberapa peneliti diantaranya Sumarno, dkk (1997) melakukan penelitian mengenai penerapan model fertilitas perkawinan terhadap data Jawa-Bali menggunakan metode permodelan *Coale-Trussell*. Penelitian Mundiharno (1998) membahas tentang teknik-teknik estimasi fertilitas. Penelitian tentang metode Regresi *Spline* pernah dilakukan, diantaranya Kurnia (2012) pada kasus gizi buruk balita, Sari (2006) pada ASFR di Jawa Timur dan Basri (2008).

Provinsi Bali terdiri dari 8 kabupaten dan 1 kota. Penelitian ini dipusatkan pada kabupaten/kota yang memiliki nilai ASFR tertinggi dan terendah, serta data ASFR tingkat provinsi Bali. Kabupaten yang memiliki nilai ASFR tertinggi adalah kabupaten Karangasem yang terletak di wilayah timur pulau Bali, sedangkan yang terendah adalah kota Denpasar yang merupakan ibukota dari Provinsi Bali. Dari 3 wilayah tersebut akan dibuat model hubungan antara ASFR dan umur ibu sebagai aplikasi dari penelitian tentang inferensi statistik untuk kurva regresi nonparametrik *spline* kuadratik. Peneliti ingin melakukan penelitian yang mengupas secara matematis dan mendalam mengenai inferensi statistik regresi nonparametrik *spline* kuadratik khususnya mendapatkan interval konfidensinya yang diaplikasikan pada studi kasus data ASFR. Maka, pada Tesis ini akan digunakan metode regresi nonparametrik khususnya Regresi Nonparametrik *Spline* Kuadratik. Regresi *Spline* digunakan karena *Spline* memiliki kemampuan yang baik untuk digeneralisasikan pada pemodelan Statistika yang kompleks dan rumit (Budiantara, 2009).

1.2 Perumusan Masalah

Permasalahan yang akan dibahas dalam penelitian kali ini adalah sebagai berikut :

1. Bagaimana mendapatkan estimasi titik dan interval konfidensi untuk kurva regresi dalam regresi nonparametrik *spline* kuadratik?
2. Bagaimana memodelkan ASFR (*Age Spesific Fertility Rate*) di kabupaten Karangasem, kota Denpasar dan provinsi Bali menggunakan pendekatan regresi nonparametrik *spline* kuadratik?

1.3 Tujuan

Berdasarkan perumusan masalah yang telah dilakukan, maka tujuan yang ingin dicapai peneliti adalah

1. Mendapatkan estimasi titik dan interval konfidensi untuk kurva regresi dalam regresi nonparametrik *spline* kuadratik.
2. Memodelkan ASFR (*Age Spesific Fertility Rate*) untuk kabupaten Karangasem, kota Denpasar dan Provinsi Bali menggunakan pendekatan regresi nonparametrik *spline* kuadratik.

1.4 Manfaat

Manfaat yang diharapkan peneliti dari permasalahan mengenai ASFR (*Age Spesific Fertility Rate*) di Bali adalah

1. Penelitian ini diharapkan bisa memberikan masukan kepada seluruh pemerintah khususnya pemerintah daerah yang ada di Bali untuk dijadikan sebagai bahan pertimbangan dalam menentukan arah kebijakan atau melaksanakan program-program di bidang kesehatan khususnya KB (Keluarga Berencana) dan bermanfaat sebagai salah satu alat evaluasi terhadap keberhasilan Program Keluarga Berencana (KB) di Indonesia, khususnya Bali.
2. Mengetahui selang umur wanita di Indonesia, Khususnya di Bali yang mempunyai fertilitas tinggi dan rendah, sehingga bermanfaat untuk menduga umur masa subur dan masa *menopause*. Dengan demikian diharapkan para wanita bisa mengatur kelahiran anak pada usia yang aman secara medis Disamping itu juga dapat menyesuaikan diri menghadapi masa *menopause*, dimana pada masa ini wanita banyak mengalami perubahan perilaku seksual, psikologis dan fisik.
3. Memberikan informasi dan pengetahuan baru mengenai aplikasi penggunaan metode Regresi Nonparametrik *Spline* kuadratik untuk data yang tidak diketahui bentuk polanya.
4. Hasil penelitian diharapkan menjadi bahan masukan atau acuan dalam penelitian selanjutnya yang akan dilakukan.

1.5 Batasan Masalah

Penelitian ini dibatasi pada pendekatan *Spline* kuadratik dalam regresi nonparametrik dimana titik knot yang optimal diperoleh dari meminimumkan metode GCV. Disamping itu, untuk tujuan aplikasi digunakan data sekunder yang diperoleh dari Badan Koordinasi Keluarga Berencana Nasional (BKKBN) Propinsi Bali yaitu data tentang ASFR di Bali khususnya di kabupaten Karangasem, kota Denpasar dan provinsi Bali tahun 2013 dengan menggunakan satu variabel prediktor yang dikelompokan menurut usia dengan titik tengah sebagai wakil kelompok usia.

\

(*Halaman ini sengaja dikosongkan*)

BAB 2
KAJIAN PUSTAKA

2.1 Analisis Regresi

Analisis regresi merupakan suatu metode statistika yang digunakan untuk mengetahui pola hubungan antara variabel respon dengan variabel prediktor. Dengan demikian analisis regresi merupakan suatu metode inferensi statistik untuk suatu fungsi regresi atau kurva regresi (Eubank, 1999). Tujuan utama dalam analisis regresi adalah bagaimana mencari bentuk estimasi untuk kurva regresi. Berkaitan dengan hal tersebut, terdapat tiga model pendekatan regresi, yaitu regresi parametrik, regresi semiparametrik, dan regresi nonparametrik. Apabila dalam analisis regresi bentuk kurva regresi diketahui maka pendekatan model regresi tersebut dinamakan model regresi parametrik (Budiantara, 2006). Sedangkan apabila bentuk kurva regresi tidak diketahui maka digunakan regresi nonparametrik. Regresi Semiparametrik digunakan jika dalam model regresi terdapat komponen Parametrik dan komponen Nonparametrik.:

2.2 Regresi Parametrik

Regresi parametrik merupakan metode statistik yang digunakan untuk mengetahui pola hubungan antara variabel respon dengan variabel prediktor dengan asumsi bentuk kurva diketahui. Model regresi linear sederhana yang hanya melibatkan satu variabel prediktor dapat dituliskan sebagai berikut.

$$y = \beta_0 + \beta_1 x + \varepsilon, \qquad (2.1)$$

dengan y adalah variabel respon, β_0 dan β_1 merupakan parameter, x adalah variabel prediktor, dan ε adalah error yang berdistribusi normal, independen dengan mean nol dan varians σ^2.

Selain model regresi derajat satu (2.1) terdapat pula model regresi derajat dua dengan satu variabel prediktor sebagai berikut.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x^2 + \varepsilon \ . \qquad (2.2)$$

Model regresi derajat tiga dengan satu variabel prediktor adalah :

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x^2 + \beta_3 x^3 + \varepsilon. \tag{2.3}$$

Regresi parametrik multivariabel merupakan perluasan dari analisis regresi parametrik univariabel yang bertujuan mengetahui hubungan linear antara satu variabel respon dengan beberapa variabel prediktor secara bersama-sama. Model regresi multivariabel dengan k variabel prediktor dapat ditulis sebagai berikut.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i, i = 1,2,\ldots,n, \tag{2.4}$$

dengan y_i adalah variabel respon, $\beta_0, \beta_1, \beta_2, \ldots, \beta_k$ adalah parameter, $x_{i1}, x_{i2}, \ldots, x_{ik}$ merupakan variabel prediktor, sedangkan ε_i adalah error yang berdistribusi normal, independen dengan mean nol dan varians σ^2. Penulisan persamaan linear (2.5) dapat ditulis ke dalam bentuk matriks seperti pada persamaan (2.6)

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \varepsilon \tag{2.5}$$

dengan y merupakan vektor respon berukuran $nx1$, X merupakan matriks berukuran $nx(k+1)$, β adalah vektor parameter yang akan diestimasi berukuran $(k+1)x1$, serta ε merupakan vektor error berukuran $nx1$. Secara lengkap matriks dan vektor-vektor tersebut diberikan oleh:

$$\boldsymbol{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} \; dan \; \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

2.3 Regresi Nonparametrik

Regresi nonparametrik merupakan pola hubungan antara variabel respon dan variabel prediktor dengan asumsi pola antara kedua variabel tersebut tidak diketahui bentuk fungsinya. Pendekatan model regresi nonparametrik digunakan jika bentuk kurva regresi dari pola data tidak diketahui bentuknya atau informasi tentang bentuk pola data dimasa lalu tidak lengkap (Eubank, 1988). Dalam pendekatan regresi nonparametrik data diharapkan mencari sendiri bentuk estimasinya sendiri tanpa harus dipengaruhi oleh faktor subyektifitas peneliti, sehingga pendekatan regresi nonparametrik memiliki fleksibilitas yang tinggi (Eubank, 1988). Kurva regresi hanya diasumsikan halus (*smooth*) dalam artian kurva termuat dalam suatu ruang fungsi tertentu. Regresi nonparametrik yang paling sering digunakan adalah regresi nonparametrik *Spline truncated*. Model regresi nonparametrik secara umum dituliskan sebagai berikut.

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1,2,\dots,n \tag{2.6}$$

dimana :

y_i : variabel respon

x_i : variabel prediktor

$f(x_i)$: fungsi regresi

ε_i : *error* yang berdistribusi normal, independent dengan mean nol dan varians σ^2

2.3.1 Fungsi *Spline* dalam Regresi Nonparametrik

Spline merupakan potongan polinomial (*piecewise polynomial)*, yaitu polinomial yang memiliki sifat tersegmen yang kontinu. Dalam regresi nonparametrik, *spline* mempunyai sifat flesibilitas yang tinggi dan mempunyai kemampuan mengestimasi perilaku data yang cenderung berbeda pada interval yang berlainan (Eubank, 1988 dan Budiantara, 2006). Kemampuan tersebut ditunjukkan dengan fungsi *truncated* (potongan-potongan) yang melekat pada estimator, dimana potongan-potongan itu disebut titik *knot*. Titik perpaduan bersama dari potongan-potongan tersebut yang menunjukan terjadinya perubahan pola perilaku fungsi *spline* pada selang yang berbeda disebut titik *knot* (Hardle, 1990).

Secara umum, fungsi $f(x_i)$ merupakan fungsi *Spline* dengan orde 2, dimana titik knot $k_1, k_2, \dots, k_r$ yang dapat dituliskan dengan persamaan berikut.

$$f(x_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \sum_{k=1}^{r} \beta_{2+k} (x_i - k_k)_+^2 \tag{2.7}$$

Dimana $(x_i - k_k)_+^2$ merupakan fungsi *truncated* (potongan) yang diuraikan sebagai berikut.

$$(x_i - k_k)_+^2 = \begin{cases} (x_i - k_k)^2, & x_i \geq k_k \\ 0, & x_i < k_k \end{cases} \tag{2.8}$$

Jika persamaan (2.8) disubstitusikan ke persamaan (2.7), maka akan dihasilkan model regresi nonparametrik *Spline* sebagai berikut.

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \sum_{k=1}^{r} \beta_{2+k} (x_i - k_k)_+^2 + \varepsilon_i \,, \qquad i = 0,1,2,\dots,n \tag{2.9}$$

2.3.2 Estimasi Parameter Regresi Nonparametrik Spline Kuadratik

Model regresi nonparametrik *spline* kuadratik secara umum dituliskan sebagai berikut.

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

dengan $f(x_i)$ merupakan fungsi *Spline* dengan orde 2, dengan titik *knot* $k_1, k_2, \dots, k_r$ dapat dituliskan dengan persamaan berikut.

$$f(x_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \sum_{k=1}^{r} \beta_{2+k} (x_i - k_k)_+^2$$

Model regresi *spline* kuadratik dapat dituliskan menjadi

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 (x_i - k_1)_+^2 + \cdots + \beta_{2+r} (x_i - k_r)_+^2 + \varepsilon_i \\ &= \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \sum_{k=1}^{r} \beta_{2+k} (x_i - k_k)_+^2 + \varepsilon_i \end{aligned}$$

Untuk i=1

$$Y_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \sum_{k=1}^{r} \beta_{2+k} (x_1 - k_k)_+^2 + \varepsilon_1$$

untuk i=2

$$Y_2 = \beta_0 + \beta_1 x_2 + \beta_2 x_2^2 + \sum_{k=1}^{r} \beta_{2+k} (x_2 - k_k)_+^2 + \varepsilon_2$$

dan seterusnya,
untuk i=n

$$Y_n = \beta_0 + \beta_1 x_n + \beta_2 x_n^2 + \sum_{k=1}^{r} \beta_{2+k} (x_n - k_k)_+^2 + \varepsilon_n$$

Jika ditulis dalam bentuk matriks menjadi:

$$\boldsymbol{Y} = \boldsymbol{X\beta}$$

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 & x_1^2 & (x_1 - k_1)_+^2 & \dots & (x_1 - k_r)_+^2 \\ 1 & x_2 & x_2^2 & (x_2 - k_1)_+^2 & \dots & (x_2 - k_r)_+^2 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & (x_3 - k_1)_+^2 & \dots & (x_n - k_r)_+^2 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{2+r} \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Jika diasumsikan error ε_i berdistribusi normal independen dengan mean nol dan variansi σ^2, yaitu :

$$\varepsilon_i \sim N(0, \sigma^2) \text{ atau}$$

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 I)$$

maka diperolah fungsi *Likelihood* :

$$L(\beta_0, \beta_1, \beta_2, \dots, \beta_{2+r}) = \prod_{i=1}^{n} f(\varepsilon_i)$$

$$= \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}\varepsilon_i^2}\right)$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} \varepsilon_i^2}$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - f(x_i))^2}$$

jika ditulis dalam bentuk matriks didapat:

$$L(\boldsymbol{\beta}) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} (Y-f)'(Y-f)}$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} (\boldsymbol{Y}-\boldsymbol{X\beta})'(\boldsymbol{Y}-\boldsymbol{X\beta})}$$

$$l(\boldsymbol{\beta}) = \log L(\boldsymbol{\beta}) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2} (\boldsymbol{Y}-\boldsymbol{X\beta})'(\boldsymbol{Y}-\boldsymbol{X\beta})$$

Jika persamaan di atas diderivatifkan parsial terhadap $\boldsymbol{\beta}$ kemudian hasilnya disamakan dengan nol, maka diperoleh :

$$\frac{\partial l(\beta_0)}{\partial \boldsymbol{\beta}} = 0$$

$$\frac{\partial((\boldsymbol{Y}-\boldsymbol{X\beta})'(\boldsymbol{Y}-\boldsymbol{X\beta}))}{\partial \boldsymbol{\beta}} = 0$$

$$\frac{\partial \boldsymbol{Y'Y}}{\partial \boldsymbol{\beta}} - \frac{\partial 2\boldsymbol{\beta' X' Y}}{\partial \boldsymbol{\beta}} + \frac{\partial \boldsymbol{\beta' X' X \beta}}{\partial \boldsymbol{\beta}} = 0$$

$$0 - 2\boldsymbol{X'Y} + 2\boldsymbol{X'X\beta} = 0$$

$$2\boldsymbol{X'X\beta} = 2\boldsymbol{X'Y}$$

$$\boldsymbol{X'X\beta} = \boldsymbol{X'Y}$$

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'Y}$$

Selanjutnya diperoleh estimasi fungsi *spline* kuadratik sebagai berikut:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2 + \hat{\beta}_3 (x_i - k_1)_+^2 + \cdots + \hat{\beta}_{2+r}(x_i - k_r)_+^2$$

Dalam bentuk matriks :

$\hat{Y} = \hat{f}(x)$, dimana $A(k) = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$ (disebut Matriks Koefisien)

$= \boldsymbol{X}\widehat{\boldsymbol{\beta}}$

$= \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}$

$= A(k)\boldsymbol{Y}$

Matriks ini memegang peranan penting dalam pemodelan regresi nonparametrik *spline*, matriks tersebut bersifat simetris dan idempoten.

2.3.3 Memilih Titik Knot Optimal

Spline sangat tergantung dengan titik knot. Model *Spline* yang terbaik adalah model dengan titik knot yang optimal, dimana model tersebut merupakan model yang paling sesuai dengan data. Menurut Wahba (1990) dan Wang (1998), salah satu metode yang baik dan banyak digunakan karena kelebihan yang dimiliki untuk pemilihan titik knot yang optimal adalah metode *Generalized Cross Validation* (GCV). Dibandingkan dengan metode lain misalnya *Cross Validation* (CV), metode GCV mempunyai sifat optimal asimtotik (Wahba, 1990). Metode GCV secara umum didefinisikan sebagai berikut.

$$GCV(k) = \frac{MSE(k)}{[n^{-1}trace(I - A(k))]^2} \tag{2.11}$$

dengan :

$MSE(k) = n^{-1}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$, dan

$A(k) = X(k)(X'(k)(X(k))^{-1}X'(k)$

Dalam analisis regresi, salah satu tujuan yang ingin dicapai adalah mendapatkan model terbaik yang mampu menjelaskan hubungan antara variabel prediktor dengan variabel respon berdasarkan kriteria tertentu. Kriteria yang sering digunakan untuk menentukannya adalah dengan GCV. Model dengan nilai GCV paling minimum dikatakan model yang terbaik.

Model regresi *spline* linear pada umumnya telah dituliskan pada persamaan sebelumnya. Adapun model regresi nonparametrik *spline* kuadratik adalah sebagai berikut.

$y_i = f(x_i) + \varepsilon_i$ dengan

$$f(x_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \sum_{k=1}^{r}\beta_{2+k}(x_i - K_k)_+^2, \tag{2.12}$$

oleh $\varepsilon_i \sim N(0, \sigma^2)$, $i = 1,2, \dots, n$

2.4 Identifikasi Asumsi Residual

Nilai residual adalah selisih dari nilai pengamatan sebenarnya (Y_i) dengan nilai dugaan $(\hat{Y}_i)$ (Drapper, 1996). Nilai residual dari hasil analisis regresi harus memenuhi asumsi *IIDN* yang berarti antar residual yang dihasilkan harus Identik (memiliki varians yang sama), Independen (tidak terdapat korelasi) dan berdistribusi Normal.

2.4.1 Uji Identik

Asumsi identik terpenuhi bila varians *error* sama dengan σ^2. Terpenuhi atau tidaknya asumsi identik dapat diketahui dengan melihat pola sebaran *scatter plot* antara residual *versus fits*. Asumsi terpenuhi jika sebaran plot tidak membentuk suatu pola tertentu. Jika plot membentuk pola sebaran tertentu maka dikatakan mengalami kasus heteroskedastisitas yaitu keadaan dimana varians residual yang diperoleh tidak homogen. Cara lain untuk identifikasi adanya kasus heteroskedastisitas adalah menggunakan uji *Glejser* dengan model yang terbentuk adalah sebagai berikut.

$|\varepsilon_i| = f(x_i) + u_i$

dengan hipotesis sebagai berikut:

$H_0: {\sigma_1}^2 = {\sigma_2}^2 = \cdots = {\sigma_3}^2 = \sigma^2$

$H_1: minimal\ ada\ satu\ {\sigma_i}^2 \neq \sigma^2, 1 = 1,2, \ldots, n$

Statistik Uji:

$$F_{hitung} = \frac{MSR}{MSE} = \frac{((\hat{\boldsymbol{Y}} - \bar{\boldsymbol{Y}})'(\hat{\boldsymbol{Y}} - \bar{\boldsymbol{Y}}))/p}{((\hat{\boldsymbol{Y}} - \boldsymbol{Y})\ (\hat{\boldsymbol{Y}} - \boldsymbol{Y})/(n-p)} \tag{2.13}$$

Keputusan tolak H_0 jika nilai F_{hitung} lebih besar dari nilai $F_{tabel}(F_{\alpha;(p,n-p)}$ atau $P_{value} < \alpha$.

2.4.2 Identifikasi Asumsi Independen

Asumsi kedua yang harus dipenuhi adalah asumsi independen pada residual yang ditunjukkan oleh nilai kovarian antara ε_i dan ε_j sama dengan nol. Independen adalah suatu keadaan dimana residual model pada periode tertentu tidak berkorelasi dengan residual pada periode lainnya. Apabila dalam persamaan regresi terdapat ketergantungan antara residual maka :

$$Cov(\varepsilon_i, \varepsilon_j) \neq 0$$

Untuk menguji asumsi ini dapat dilihat plot ACF dari residual, jika semua plot berada pada batas daerah signifikan yaitu $\pm Z_{\frac{\alpha}{2}}/\sqrt{n}$ maka dikatakan tidak terdapat kasus autokorelasi.

2.4.3 Uji Distribusi Normal

Asumsi ketiga yang harus dipenuhi adalah asumsi distribusi normal. Uji ini dilakukan untuk mengetahui bahwa asumsi distribusi normal sudah terpenuhi. Untuk melakukan pengujian ini dapat dilakukan dengan cara visual yaitu dilihat dari *normal probability plot residual*. Jika plot yang dihasilkan cenderung membentuk garis lurus dengan sudut 45° maka residual dikatakan berdistribusi normal. Cara lain adalah dengan melakukan pengujian hipotesis untuk uji Kolmogorov Smirnov, dengan hipotesis sebagai berikut.

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

$H_1: F_0(x) \neq F(x)$ (Residual tidak berdistribusi Normal)

Statistik Uji:

$$D = maks|F_0(x) - S_n(x)| \tag{2.14}$$

dengan $F_0(x)$ merupakan fungsi distribusi empirik sedangkan $S_n(x)$ adalah fungsi distribusi teoritis. Keputusan tolak H_0 jika $|D| > q_{(1-\alpha)}$ dengan nilai $q_{(1-\alpha)}$ didapatkan dari tabel Kolmogorov Smirnov (Daniel, 1989).

2.5 Koefisien Determinasi (R^2)

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

$$R^2 = \frac{SSR}{SST} = \frac{(\hat{Y}-\bar{Y})'(\hat{Y}-\bar{Y})}{(Y-\bar{Y})'(Y-\bar{Y})} \tag{2.15}$$

2.6 Interval Konfidensi

Estimasi interval adalah suatu interval tertentu yang memuat parameter dengan probabilitas tertentu. Misalkan $X_1, X_2, \dots, X_n$ adalah sampel random yang diambil dari populasi dengan parameter θ, maka interval konfidensi $1-\alpha$ untuk θ adalah $P(a \le \theta \le b) = 1-\alpha$.

Diberikan sampel random $X_1, X_2, \dots, X_n$ yang diambil dari populasi yang berdistribusi $N(\mu, \sigma^2)$ dengan σ^2 diketahui. Untuk mendapatkan interval konfidensi $1-\alpha$ untuk parameter μ dapat menggunakan *Pivotal Quantity* :

$$Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1) \tag{2.16}$$

Apabila ditentukan nilai $1-\alpha = 95\%$ maka interval konfidensi 95% untuk parameter μ diberikan oleh :

$$P(-1{,}96 \le Z \le 1{,}96) = 95\%$$

$$\Leftrightarrow P\left(-1{,}96 \le \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \le 1{,}96\right) = 95\%$$

$$\Leftrightarrow P\left(\bar{X} - 1{,}96.\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + 1{,}96.\frac{\sigma}{\sqrt{n}}\right) = 95\%$$

Jika σ tidak diketahui, diganti dengan *s* yaitu standar deviasi sampel. Dengan demikian, untuk dapat mendapatkan interval konfidensi $1-\alpha$ untuk μ digunakan *Pivotal Quantity*:

$$T = \frac{\bar{X}-\mu}{\frac{s}{\sqrt{n}}} \sim t_{(n-1)}$$

Dengan demikian, interval konfidensi $1-\alpha$ untuk parameter μ dapat diperoleh dari:

$$P\left(-t_{(n-1;\alpha/2)} \le T \le t_{(n-1;\alpha/2)}\right) = 1-\alpha$$

$$\Leftrightarrow P\left(-t_{(n-1;\alpha/2)} \le \frac{\bar{X}-\mu}{\frac{s}{\sqrt{n}}} \le t_{(n-1;\alpha/2)}\right) = 1-\alpha$$

$$\Leftrightarrow P\left(\bar{X} - t_{(n-1;\alpha/2)}.\frac{s}{\sqrt{n}} \le \mu \le \bar{X} + t_{(n-1;\alpha/2)}.\frac{s}{\sqrt{n}}\right) = 1-\alpha \tag{2.17}$$

2.7 Tinjauan Demografi

Demografi adalah suatu ilmu yang berkenaan dengan penduduk (manusia) (Sembiring, 1985). Demografi mempelajari secara statistik dan matematik tentang besar, komposisi, dan distribusi penduduk dan perubahan-perubahannya sepanjang masa melalui bekerjanya komponen-komponen demografi. Komponen demografi terdiri dari 5 hal, yaitu : Kelahiran (Fertilitas), Kematian (Mortalitas), Perkawinan, Migrasi dan Mobilitas Sosial. Dari kelima komponen demografi ini yang paling menjadi masalah bagi negara yang sedang berkembang adalah fertilitas, karena semakin tinggi angka kelahiran maka semakin meningkat pula laju pertumbuhan penduduknya.

2.8 Fertilitas dan ASFR

Fertilitas merupakan taraf kelahiran penduduk yang sesungguhnya berdasarkan jumlah kelahiran yang terjadi (lahir hidup). Pengertian ini digunakan untuk menunjukkan pertambahan jumlah penduduk. Ukuran dasar fertilitas yang sering digunakan adalah angka kelahiran menurut usia (*Age Spesific Fertility Rate* atau ASFR). Sembiring (1985) mendefinisikan ASFR sebagai angka yang menunjukkan banyaknya kelahiran selama setahun, setiap 1000 wanita pada kelompok usia tertentu. Definisi ASFR harus dipahami dengan baik, karena ini merupakan penghitungan vital untuk mendapatkan angka nyata mengenai tingkat kelahiran di masyarakat. Mengetahui angka kelahiran di suatu negara menjadi hal yang penting untuk diketahui. Karena dengan itu, pemerintah bisa mengetahui bagaimana tingkat pertumbuhan dan pertambahan penduduknya yang nantinya tentu akan berpengaruh pada setiap kebijakan yang dibuatnya. Angka ini digunakan untuk membedakan dengan fertilitas, yang dirumuskan menjadi :

$$ASFR_i = \frac{B_i}{P_{fi}} . 1000 \qquad (2.18)$$

dimana :

B_i = banyaknya kelahiran di dalam kelompok usia selama 1 tahun.

P_{fi} = banyaknya perempuan pada kelompok usia tertentu selama 1 tahun.

i = kelompok umur, yaitu 15-19, 20-24,...,45-49.

Secara teori probabilistik, ASFR merupakan fungsi kuadratik, karena pola grafiknya yang berbentuk seperti gunung (*U-shape* terbalik) dan hampir seperti bentuk kurva distribusi normal tetapi tidak simetris. Hal itu diakibatkan karena kelahiran paling banyak terdapat di usia produktif wanita yaitu 20-29 tahun, karena usia tersebut normal tidak terlalu muda dan tidak terlalu tua dan merupakan rentang usia yang aman untuk hamil dan melahirkan.

a. Kelebihan dari ASFR

Kelebihan dari ASFR adalah.

1. Nilai ASFR ini lebih cermat dibandingkan GFR (General Fertility Rate) karena telah memperhitungkan kemampuan melahirkan seorang ibu dilihat dari umurnya atau sudah membagi penduduk yang "*exposed to risk*" dalam berbagai kelompok umur.
2. Bisa melakukan analisis perbedaan fertilitas menurut berbagai karakteristik wanita.
3. Bisa dilakukan studi fertilitas menurut kohor.
4. Sebagai dasar penghitungan ukuran fertilitas dan reproduksi selanjutnya.

b. Penghitungan ASFR

Pada penghitungan ASFR

1. Kelompok umur 5 tahunan yang paling sering digunakan.
2. Pola grafiknya seperti bentuk gunung, tidak simetris dan hampir seperti bentuk kurva distribusi normal.
3. Pola grafiknya untuk berbagai negara bentuknya hampir sama.
4. Dapat menggambarkan rata-rata usia kawin wanita yang ditunjukkan oleh letak puncak kurva.

\

(*Halaman ini sengaja dikosongkan*)

BAB 3
METODA PENELITIAN

3.1 Sumber Data Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Koordinasi Keluarga Berencana Nasional (BKKBN) Provinsi Bali tahun 2013.

3.2 Variabel Penelitian

Variabel-variabel yang digunakan pada penelitian kali ini terdiri dari variabel respon (Y) dan variabel prediktor (X). Variabel tersebut adalah :

1. Variabel respon (dependen)

   Y_i = Tingkat fertilitas menurut umur (ASFR)

   ASFR merupakan perbandingan antara jumlah kelahiran yang tercatat setahun pada ibu dalam kelompok usia (i) dengan jumlah ibu pada kelompok usia ke-i
2. Variabel prediktor (independen)

   X = Usia ibu (dalam tahun)

   Usia produktif ibu untuk melahirkan menurut BKKBN adalah usia 15 sampai dengan 49 tahun, dari rentang usia ini dibagi menjadi 7 pengamatan usia ibu yaitu dengan mengelompokkan usia ibu dalam interval 5 tahunan, sehingga dari pengelompokkan ini dapat diambil nilai tengah ibu yang akan menjadi variabel x dalam pengamatan.

3.3 Metode Penelitian

Adapun langkah-langkah yang harus dilakukan untuk analisis data untuk mencapai tujuan dalam penelitian ini adalah sebagai berikut.

A. Kajian Teoritis

1. Diberikan data berpasangan $(x_i, y_i), i = 1,2,3, \dots, n$ yang mengikuti model regresi nonparametrik :

   $$y_i = f(x_i) + \varepsilon_i$$

   dengan kurva regresi f didekati dengan fungsi *spline* kuadrat

$$f(x_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \sum_{k=1}^{r} \beta_{2+r}(x_i - K_k)_+^2\ , dan$$

$$\varepsilon_i \sim N(0, \sigma^2)$$

2. Mencari interval konfidensi 1-α pada regresi nonparametrik *spline* kuadratik dengan pendekatan *Pivotal Quantity*.

   Langkah-langkahnya adalah sebagai berikut :

   a. Menentukan basis fungsi *spline* kuadratik.

   b. Melakukan estimasi parameter dari model regresi nonparametrik *spline* kuadratik $Y_i = f(x_i) + \varepsilon_i$, dimana

   $$f(x_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \sum_{k=1}^{r} \beta_{2+r}(x_i - K_k)_+^2$$

   dengan meggunakan metode MLE.

   c. Menghitung panjang interval konfidensi $1 - \alpha$ dengan menggunakan konsep interval konfidensi terpendek.

   d. Menurunkan interval konfidensi terpendek untuk kurva regresi nonparametrik $f(x_i), i = 1,2, \ldots, n$.

B. Aplikasi pada data ASFR :

Pemodelan ASFR kabupaten Karangasem, kota Denpasar dan propinsi Bali dengan Regresi Nonparametrik *Spline* kuadratik :

1. Membuat *scatter plot* data ASFR terhadap usia ibu.
2. Memodelkan data menggunakan regresi *spline* kuadratik.
3. Menentukan penaksir parameter model regresi *spline* kuadratik.
4. Menghitung skor GCV (*Generalized Cross Validation*) untuk 1 *knots*, 2 *knots*, dan 3 *knots* pada model kuadratik.
5. Membuat plot untuk skor GCV.
6. Memilih skor GCV yang paling minimum.
7. Melakukan uji asumsi terhadap residual:

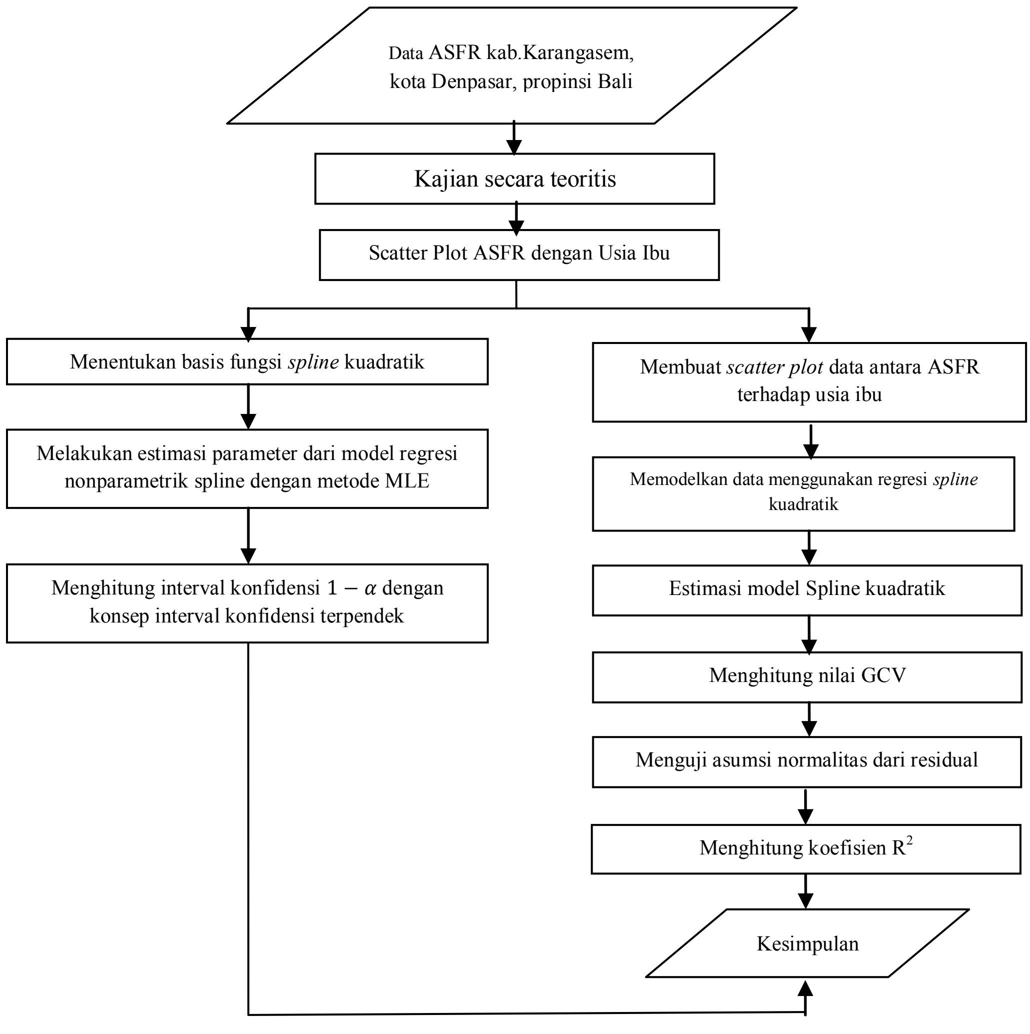
   a. Independensi.

   b. Identik/homogenitas.

   c. Kenormalan.

9. Menghitung nilai koefisien determinasi (R^2).
10. Menghitung interval konfidensi $(1-\alpha)$ untuk kurva regresi dengan menggunakan pendekatan *spline* kuadrat.

Untuk lebih jelasnya langkah-langkah penelitian akan disajikan dalam bentuk diagram alir penelitian seperti pada Gambar 3.1



Gambar 3.1 Diagram Alir Penelitian

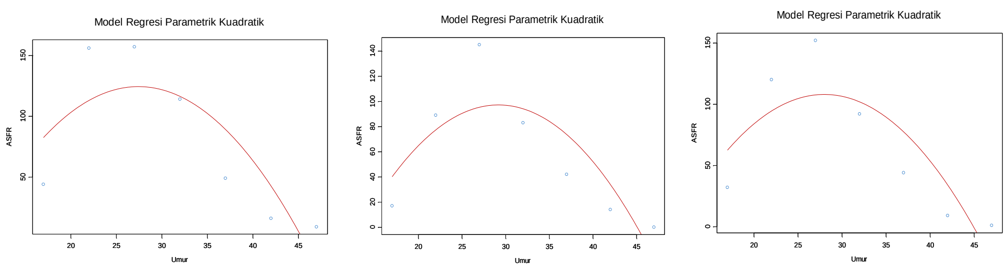
.

(*Halaman ini sengaja dikosongkan*)

BAB 4
HASIL DAN PEMBAHASAN

Salah satu faktor yang mempengaruhi angka kelahiran di Provinsi Bali adalah umur ibu, sesuai pada data BKKBN (Badan Koordinasi Keluarga Berencana Nasional) provinsi Bali dimana data ASFR (*Age Spesific Fertility Rate*) sebagai patokan angka kelahiran. Dalam penelitian ini akan dimodelkan pola hubungan antara umur ibu dengan ASFR untuk Kabupaten Karangasem, Kota Denpasar dan Provinsi Bali. Tetapi sebelumnya akan dikaji bentuk estimator dan interval konfidensi terpendek untuk kurva regresi nonparametrik *spline* kuadratik.

4.1 Model Regresi Parametrik Kuadratik ASFR Karangasem, Denpasar dan Bali



(a) (b) (c)

Gambar 4.1 Kurva Regresi Parametrik Kuadratik ASFR pada Karangasem, Denpasar, Bali

Dari gambar diatas menunjukkan bahwa regresi parametrik tidak cocok diaplikasikan pada data ASFR walaupun kurva ASFR berbentuk kuadratik, hal ini disebabkan karena regresi parametrik kuadratik tidak peka terhadap titik perubahan pola data. Jika dilihat dari nilai R^2 tidak terlalu tinggi dan nilai MSE nya pun teralu besar, berikut adalah rinciannya untuk ASFR Karangasem nilai R^2=88,37% dan nilai MSE=1937,77, untuk ASFR Denpasar nilai R^2=86,12% dan nilai MSE=1320,67, untuk ASFR Bali nilai R^2=87,26% dan nilai MSE=1560,63.

Maka, selanjutnya data ASFR tersebut akan dilakukan analisis dengan pendekatan regresi nonparametrik *spline* kuadratik karena lebih peka terhadap titik perubahan pola data dan kurva yang dihasilkan akan lebih *smooth*.

4.2 Interval Konfidensi untuk Kurva Regresi f

Dengan menggunakan konsep *Pivotal Quantity* yaitu suatu statistik yang distribusinya tidak memuat kurva regresi $f(x_i)$,

$$Z_i(x_1, x_2, \dots, x_n) = \frac{\hat{f}(x_i) - f(x_i)}{\sqrt{var(\hat{f}(x_i))}} \sim N(0,1) \ , i = 1,2, \dots, n$$

dengan fungsi probabilitas

$$f(Z_i) = \frac{1}{\sqrt{2\pi}} e^{\frac{1}{2}Z_i}, \ -\infty < z < \infty$$

tidak memuat kurva regresi $f(x_i)$, dimana i = 1,2,...,n.

Jadi, $Z_i(x_1, x_2, \dots, x_n)$ merupakan *Pivotal Quantity* untuk kurva regresi f. Interval konfidensi $1 - \alpha$ diperoleh dari menyelesaikan persamaan probabilitas :

$$P(-z\alpha/_2 < Z(x_1, \dots, x_n) < z\alpha/_2) = 1 - \propto$$

$$P(-z\alpha/_2 < \frac{\hat{f}(x_i) - f(x_i)}{\sqrt{var\left(\hat{f}(x_i)\right)}} < z\alpha/_2) = 1 - \propto$$

Persamaan diatas ekuivalen dengan

$-z\alpha/_2 < \frac{\hat{f}(x_i) - f(x_i)}{\sqrt{var(\hat{f}(x_i))}}$, atau

$$f(x_i) < \hat{f}(x_i) + z\alpha/_2\sqrt{var(\hat{f}(x_i))} \quad \text{(i)}$$

$\frac{\hat{f}(x_i) - f(x_i)}{\sqrt{var(\hat{f}(x_i))}} < z\alpha/_2$, atau

$$f(x_i) > \hat{f}(x_i) - z\alpha/_2\sqrt{var(\hat{f}(x_i))} \quad \text{(ii)}$$

Dari persamaan (i) dan (ii) didapat Interval Konfidensi untuk kurva regresi f

$$P\left(\left[\hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2 + \sum_{k=1}^{r} \hat{\beta}_{2+k}\,(x_i - k_k)_+^2\right] - z\alpha/_2\sqrt{var\hat{f}(x_i)} < f(x_i) < \right.$$
$$\left.\left[\hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2 + \sum_{k=1}^{r} \hat{\beta}_{2+k}\,(x_i - k_k)_+^2\right] + z\alpha/_2\sqrt{var\hat{f}(x_i)}\right) = 1 - \alpha$$

dimana $var\hat{f}(x_i)$ adalah elemen diagonal ke-i dari matriks:

$$var\hat{f}(x) = var\left(\boldsymbol{X}\widehat{\boldsymbol{\beta}}\right)$$

$$= var(\boldsymbol{X}(\boldsymbol{X}\,\boldsymbol{X})^{-1}\boldsymbol{X}\,\boldsymbol{Y})$$

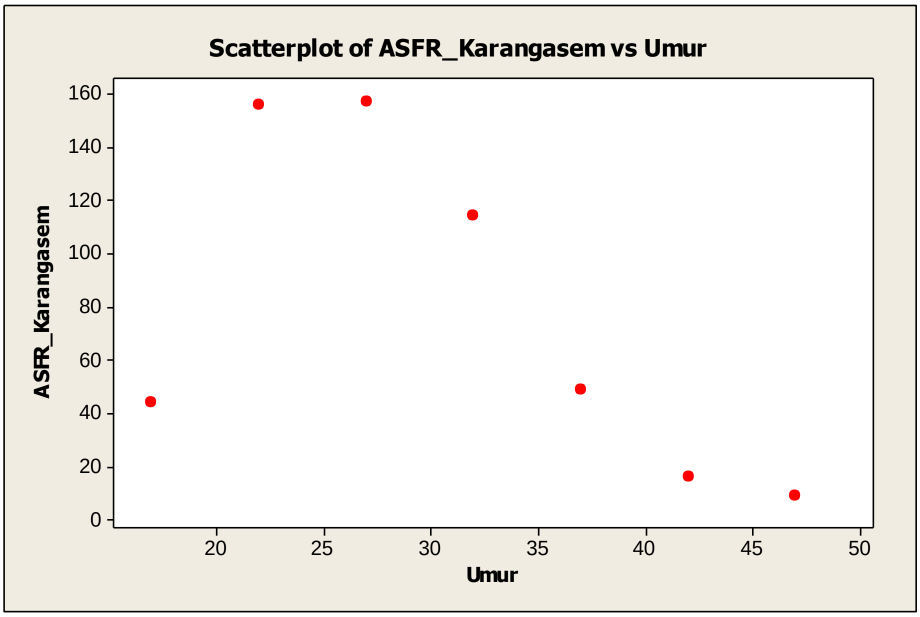
$$= \boldsymbol{X}(\boldsymbol{X}\,\boldsymbol{X})^{-1}\boldsymbol{X}\ \ var(\boldsymbol{Y})\ \boldsymbol{X}(\boldsymbol{X}\,\boldsymbol{X})^{-1}\boldsymbol{X}$$

$$= \ \sigma^2\boldsymbol{X}(\boldsymbol{X}\,\boldsymbol{X})^{-1}\boldsymbol{X}\,\boldsymbol{X}(\boldsymbol{X}\,\boldsymbol{X})^{-1}\boldsymbol{X}$$

$$= \ \sigma^2\boldsymbol{X}(\boldsymbol{X}\,\boldsymbol{X})^{-1}\boldsymbol{X}$$

4.3 Model *Spline* Kuadratik untuk ASFR Kabupaten Karangasem

Untuk melihat pola hubungan antara ASFR dengan umur ibu kabupaten Karangasem, diberikan plotnya sebagai berikut:



Gambar 4.2 Plot antara ASFR dengan umur ibu di Kab. Karangasem

Dari gambar di atas terdapat indikasi adanya perubahan pola perilaku data. Disamping itu, terlihat pola data yang cenderung kuadratik. Oleh karena itu, dalam penelitian ini penulis akan mencoba melakukan pendekatan menggunakan regresi nonparametrik *spline* kuadratik.

Dalam penggunaan regresi *Spline*, langkah awal yang harus dilakukan adalah menentukan lokasi dari titik *knot*. Cara pemilihan *knot* dengan cara mencoba-coba pada tempat yang terlihat adanya perubahan perilaku data pada selang tertentu. Dari berbagai *knot* dipilih nilai GCV minimum. Untuk mendapatkan hasil yang terbaik dalam regresi *spline* dicoba model *spline* kuadratik dengan satu *knot* sampai tiga *knot*. Selanjutnya menghitung nilai estimasi parameter dan membuat kurva estimasinya.

Untuk perhitungan GCV dilakukan dengan fasilitas Program S-PLUS. Akan dicari GCV untuk 1 *knot*, 2 *knot* dan 3 *knot* model *spline* kuadratik, sebagai berikut:

a. Model kuadratik 1 titik *knot*:

$$y = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 (x_i - k_1)_+^2 + \varepsilon$$

b. Model kuadratik 2 titik *knot*:

$$y = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 (x_i - k_1)_+^2 + \beta_4 (x_i - k_2)_+^2 + \varepsilon$$

c. Model kuadratik 3 titik *knot*:

$$y = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 (x_i - k_1)_+^2 + \beta_4 (x_i - k_2)_+^2 + \beta_5 (x_i - k_3)_+^2 + \varepsilon$$

Hasil perhitungan GCV bisa dilihat Tabel 4.1 di bawah ini.

Tabel 4.1 Skor GCV untuk 1,2 dan 3 *Knot* pada Model Kuadratik Karangasem

Banyak Titik *Knot*	**Lokasi Titik *Knot***			**Skor GCV**	**R^2**
	28			198,3554	
	29			169,3082	
1 *Knot*	**30**			**155,5447***	99,17%
	31			164,206	
	32			221,5404	
	22	30		227,0283	
	22	31		111,7987	
2 *Knot*	22	32		27,25547	
	22	33		4,927536*	99,98%
	22	34		27,47136	
	22	34	35	0,8189272	
	22	34	36	0,6708065*	99,99%
3 *Knot*	22	34	37	6,065825	
	22	34	38	12,03385	
	22	34	39	15,48113	

Berdasarkan Tabel 4.1 terlihat bahwa nilai GCV terkecil adalah 0,7608065 terdapat pada model *spline* kuadratik dengan tiga titik *knot* yaitu pada ibu yang berumur 22 tahun, 34 tahun dan 36 tahun. Tetapi untuk kesederhanaan model, dipilih model *spline* dengan satu *knot*. Hal ini dilakukan karena nilai R^2

untuk model *spline* kuadratik satu *knot* R^2=99,17% sudah relatif besar dan hampir sama dengan dua dan tiga *knot*.

Setelah diperoleh skor GCV minimum untuk model *spline* kuadratik, langkah selanjutnya yaitu menghitung estimasi untuk model *spline* kuadratik. Estimasi model *spline* kuadratik dengan satu titik *knot* (30) sebagai berikut:

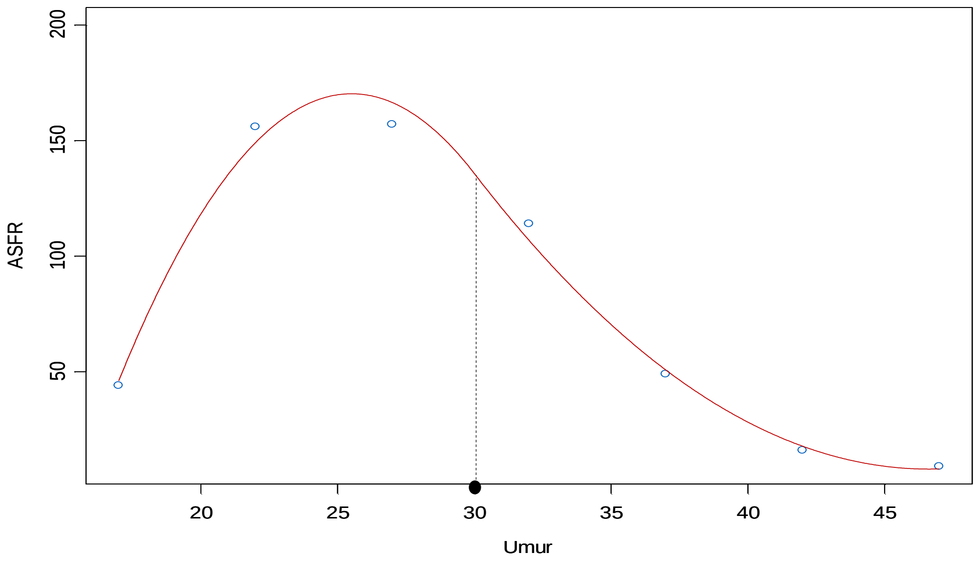
$$\hat{y} = -945{,}29 + 87{,}45x_i - 1{,}71x_i^2 + 2{,}17(x_i - 30)_+^2$$

Estimasi model tersebut dapat ditulis menjadi:

$$\hat{y} = \begin{cases} -945{,}29 + 87{,}45x_i - 1{,}71x_i^2, & x_i \leq 30 \\ 1007{,}71 - 42{,}75x_i + 0{,}46x_i^2, & x_i \geq 30 \end{cases}$$

Estimasi model spline di atas dapat diinterpretasikan bahwa di kabupaten Karangasem pada wanita/ibu yang berusia diatas 30 tahun, jumlah bayi yang dilahirkan meningkat 87 bayi dilahirkan oleh 1000 wanita dalam 1 tahun. Ditandai dengan bentuk kurva yang meningkat dimana puncaknya terletak pada titik di usia 25 tahun, yang berarti pada kabupaten Karangasem jumlah bayi paling banyak dilahirkan oleh wanita di usia 25 tahun. Sedangkan pada wanita/ibu yang usianya lebih dari 30 tahun masih produktif melahirkan, tetapi mulai mengalami penurunan sebesar 42 bayi yang dilahirkan oleh 1000 wanita dalam 1 tahun. Hal ini kemungkinan diakibatkan karena wilayah Karangasem sebagian besar desa dan agak terpelosok, sehingga kurang mendapat penyuluhan yang merata tentang KB dan banyak yang menikah di usia muda. Banyak wanita yang tidak berkesempatan mengenyam pendidikan tinggi karena tidak mempertimbangkan usia dalam merencanakan usia kelahiran maka di usia lebih dari 30 tahun pun masih aktif melahirkan.

Estimasi model spline kuadratik dengan satu titik *knot* (30) untuk ASFR Karangasem mempunyai nilai koefisien determinasi R^2 = 99,17% dan *Mean Square Error* = 28,56943. Jika ditampilkan dalam gambar, visualisasi model astimasi kurva *spline* kuadratik tersebut ditampilkan pada Gambar 4.3



Gambar 4.3 *Spline* Kuadratik ASFR di Kabupaten Karangasem dengan titik *knot* pada umur ibu 30 tahun.

a. Menguji Signifikansi Model

Pada tahap ini akan dilakukan pengujian terhadap koefisien-koefisien regresi pada model *spline* kuadratik, yang dilakukan dengan dua tahap yaitu uji serentak dan uji parsial.

(i) Uji Serentak (Uji F)

Akan diuji hipotesa :

$H_0 : \beta_1 = \beta_2 = \beta_3 = 0$

H_1 : minimal ada satu $\beta_j \neq 0$, j=1,2,3

Sratistik uji :

$$F = \frac{MSR}{MSE} = 120{,}123$$

Daerah penolakan : tolak H_0 apabila $F_{hitung} > F_{\alpha(p,n-p)}$

Keputusan :

Karena nilai F_{hitung} = 120,123 lebih besar dari F_{tabel} = 6,38823, maka tolak H_0 yang berarti terdapat parameter yang signifikan dalam model *spline* kuadratik pada taraf α=5%.

(ii) Uji t

Akan diuji hipotesa :

$H_0 : \beta_j = 0$

$H_1 : \beta_j \neq 0$, j=1,2,3

Sratistik uji : $t_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}$

Tabel 4.2 Uji Parsial model *spline* kuadratik Karangasem

Parameter	Estimasi	Nilai t_{hitung}
β_1	87,4542	15,2021
β_2	-1,7139	-15,2438
β_3	2,1765	12,2895

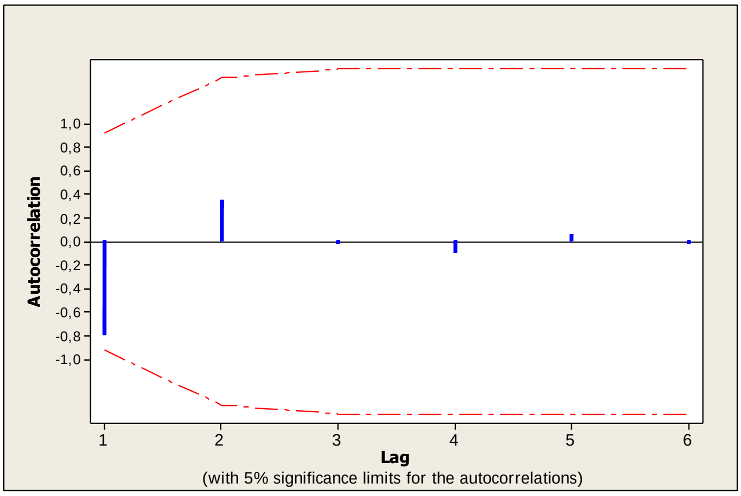
Daerah penolakan : H_0 ditolak jika $|t_{hitung}| > t_{\alpha/2,(n-p)}$

Keputusan :

Dengan menggunakan taraf α=5%., maka nilai t_{tabel} = 2,776445. Sehingga H_0 ditolak, karena nilai t_{hitung} semua parameter lebih besar dari t_{tabel}. Dengan demikian parameter $\beta_1, \beta_2,$ dan β_3 signifikan dalam model *spline* kuadratik dengan satu titik *knot*.

b. Analisa Residual

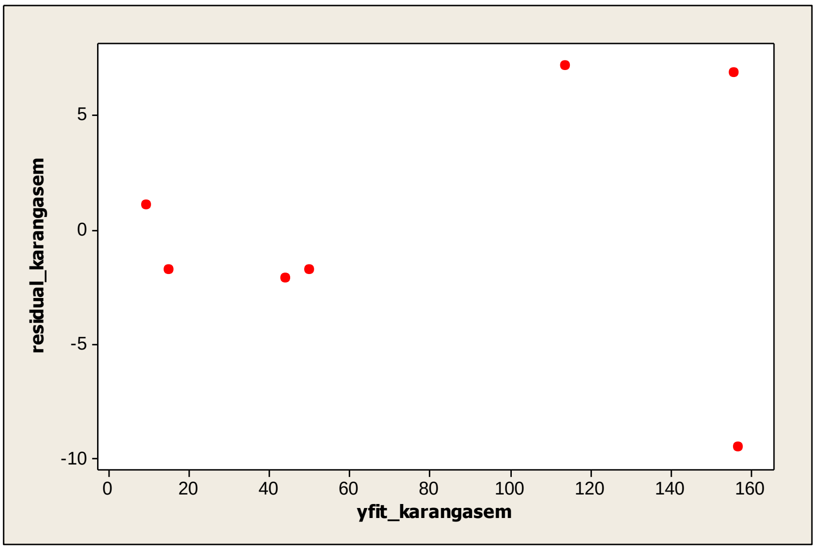
Model *spline* kuadratik dikatakan layak apabila residualnya telah memenuhi asumsi. Untuk memeriksa asumsi independen dari residual, dapat dilakukan dengan menganalisa plot ACF dari residual :



Gambar 4.4 Plot ACF residual model *spline* kuadratik satu *knot*

Dari Gambar 4.4 menunjukkan bahwa residual model *spline* kuadratik telah memenuhi asumsi independen, karena dari plot ACF tersebut tidak ada lag yang keluar dari batas.

Selanjutya dilakukan analisa terhadap residual apakah sudah memenuhi asumsi identik yaitu dengan melihat plot antara residual dengan y taksiran seperti yang terlihat pada Gambar 4.5. Dari Gambar 4.5 dapat dilihat bahwa data sudah menyebar secara acak di sekitar nilai nol dan tidak membentuk suatu pola sehingga dapat dikatakan bahwa residual telah memenuhi asumsi identik.



Gambar 4.5 Plot residual dengan y taksiran *spline* kuadratik

Asumsi terakhir yang harus dipenuhi oleh residual adalah kenormalan. Untuk mengetahui apakah residual telah berdistribusi normal dapat dilihat dari uji *Kolmogorv Smirnov* residual.

Hipotesanya :

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

Statistik Uji : $D = Sup\left|S(x) - F_{0(x)}\right| = 0{,}216$

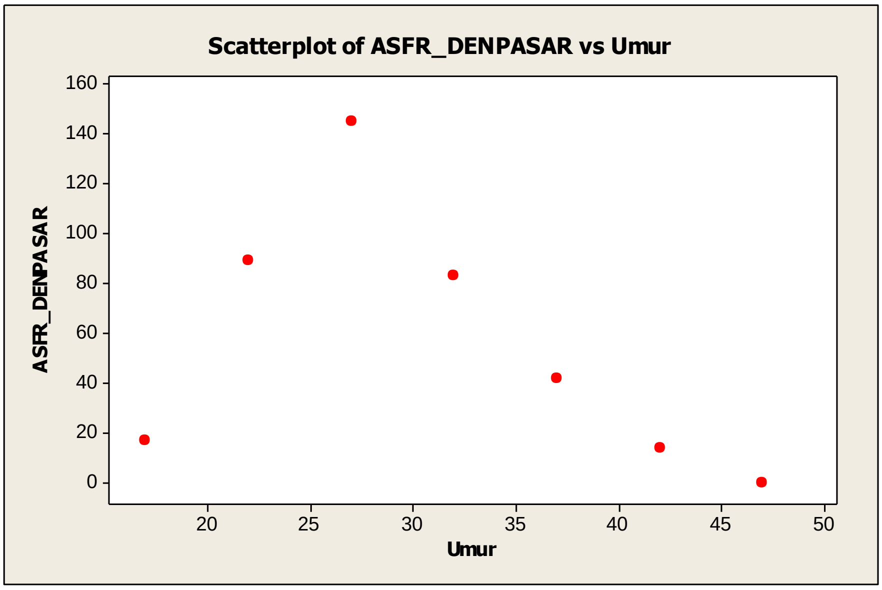
Daerah penolakan : tolak H_0 apabila $D_{hitung} > D_{1-\alpha(n)}$ tabel *Kolmogorv Smirnov*

Keputusan :

Dengan menggunakan taraf α=5%., maka gagal tolak H_0 karena nilai D_{hitung} = 0,216 kurang dari $D_{0,05(7)}$=.0,483 , sehingga dapat dinyatakan residual model *spline* kuadratik dengan satu *knot* berdistribusi normal.

4.4 Model *Spline* Kuadratik untuk ASFR Kota Denpasar

Untuk melihat pola hubungan antara ASFR dengan umur ibu kota Denpasar tahun 2013, diberikan plotnya sebagai berikut:



Gambar 4.6 Plot antara ASFR dengan umur ibu di Kota Denpasar

Dari gambar di atas terdapat indikasi adanya perubahan pola perilaku data. Disamping itu, terlihat pola data yang cenderrung kuadratik. Oleh karena itu, dalam penelitian ini penulis akan mencoba melakukan pendekatan menggunakan regresi nonparametrik *spline* kuadratik.

Dalam penggunaan regresi *Spline*, langkah awal yang harus dilakukan adalah menentukan lokasi dari titik *knot*. Cara pemilihan *knot* dengan cara mencoba-coba pada tempat yang terlihat adanya perubahan perilaku data pada selang tertentu. Dari berbagai *knot* dipilih nilai GCV minimum. Untuk mendapatkan hasil yang terbaik dalam regresi *spline* dicoba model *spline* kuadratik dengan satu *knot* sampai tiga *knot*. Selanjutnya menghitung nilai estimasi parameter dan membuat kurva estimasinya.

Untuk perhitungan GCV dilakukan dengan fasilitas Program S-PLUS. Akan dicari GCV untuk 1 *knot*, 2 *knot* dan 3 *knot* model *spline* kuadratik, sebagai berikut:

a. Model kuadratik 1 titik *knot*:

$$y = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 (x_i - k_1)_+^2 + \varepsilon$$

b. Model kuadratik 2 titik *knot*:

$$y = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 (x_i - k_1)_+^2 + \beta_4 (x_i - k_2)_+^2 + \varepsilon$$

c. Model kuadratik 3 titik *knot*:

$$y = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 (x_i - k_1)_+^2 + \beta_4 (x_i - k_2)_+^2 + \beta_5 (x_i - k_3)_+^2 + \varepsilon$$

Hasil perhitungan bisa dilihat Tabel 4.3 di bawah ini.

Tabel 4.3 Skor GCV untuk 1,2 dan 3 *Knot* pada Model Kuadratik Denpasar

Banyak Titik *Knot*	**Lokasi Titik *Knot***			**Skor GCV**	**Rsq**
1 *Knot*	29			572,2208	96,67%
	30			465,4573	
	31			423,4986*	
	32			459,0331	
	33			557,9929	
2 *Knot*	**24**	**27**		**10,6***	99,96%
	24	28		26,83843	
	24	29		87,01268	
	24	30		219,558	
	24	31		455,7134	
3 *Knot*	25	30	32	0,35*	99,99%
	25	30	33	2,386102	
	25	30	34	12,71374	
	25	30	35	40,63306	
	25	30	36	100,1435	

Berdasarkan Tabel 4.3 terlihat bahwa nilai GCV terkecil adalah 0,35 terdapat pada model *spline* kuadratik dengan tiga titik *knot* yaitu pada ibu yang berumur 25 tahun, 30 tahun dan 32 tahun. Untuk kesederhanaan model dipilih model *spline* dengan dua *knot*. Hal ini dilakukan karena nilai R^2 untuk model *spline* kuadratik dua *knot* R^2 = 99,96% sudah relatif besar dan hampir sama dengan tiga *knot*.

Setelah diperoleh skor GCV minimum untuk model *spline* kuadratik, langkah selanjutnya yaitu menghitung estimasi untuk model *spline* kuadratik. Estimasi model *spline* kuadratik dengan dua titik *knot* (24;27) sebagai berikut:

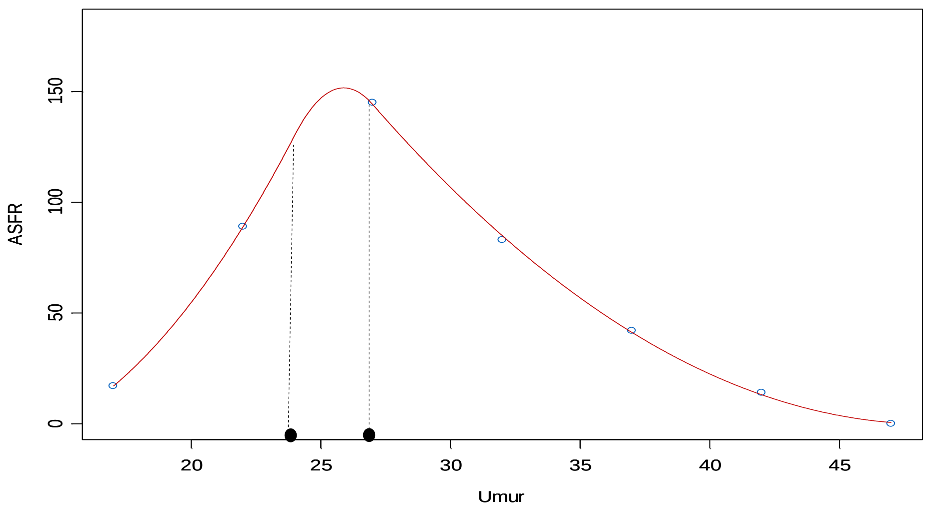
$$\hat{y} = 110{,}63 - 20{,}89x_i + 0{,}90x_i^2 - 6{,}89(x_i - 24)_+^2 + 6{,}30(x_i - 27)_+^2$$

Estimasi model tersebut dapat ditulis menjadi:

$$\hat{y} = \begin{cases} 110{,}63 - 20{,}89x_i + 0{,}90x_i^2, & x_i \leq 24 \\ -3858{,}01 + 309{,}83x_i - 5{,}99x_i^2, & 24 \leq x_i \leq 27 \\ 734{,}69 - 30{,}37x_i + 0{,}31x_i^2, & x_i \geq 27 \end{cases}$$

Estimasi model *spline* di atas dapat diinterpretasikan bahwa di kota Denpasar pada wanita/ibu yang berusia kurang dari 24 tahun banyaknya anak yang dilahirkan sebesar 20 bayi oleh 1000 wanita selama 1 tahun, ditandai dengan bentuk kurva yang naik perlahan. Untuk wanita yang berusia antara 24 sampai 27 tahun banyaknya anak yang dilahirkan meningkat yaiu 309 bayi oleh 1000 wanita selama 1 tahun, dapat dilihat pula pada kurva puncaknya terletak di titik usia 26 tahun, hal ini disebabkan wanita di kota Denpasar banyak yang telah mengenyam pendidikan yang tinggi dan rata-rata wanita di Denpasar adalah wanita pekerja, sehingga banyak wanita yang merencanakan kelahiran di usia antara 24 sampai 27 tahun. Maka jumlah anak yang dilahirkan paling banyak terdapat di selang usia tersebut. Sedangkan untuk wanita/ibu yang berusia di atas 27 tahun jumlah anak yang dilahirkan sudah mulai mengalami penurunan, yaitu 30 bayi oleh 1000 wanita selama 1 tahun dan terus menurun seiring dengan bertambahnya usia, karena wanita di kota Denpasar sudah memiliki pengetahuan bahwa maksimal usia melahirkan yang baik adalah di usia 30 tahun.

Estimasi model *spline* kuadratik dengan dua titik *knot* (24;27) untuj ASFR Denpasar mempunyai nilai koefisien determinasi R^2 = 99,96% dan *Mean Square Error* = 0,8653061. Jika ditampilkan dalam gambar, visualisasi model astimasi kurva *spline* kuadratik tersebut ditampilkan pada Gambar 4.7.



Gambar 4.7 Kurva *Spline* Kuadratik Kota Denpasar dengan titik *knot* pada umur ibu 24 tahun dan 27 tahun.

a. Menguji Signifikansi Model

Pada tahap ini akan dilakukan pengujian terhadap koefisien-koefisien regresi pada model *spline* kuadratik, yang dilakukan dengan dua tahap yaitu uji serentak dan uji parsial.

(i) Uji Serentak (Uji F)

Akan diuji hipotesa :

$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$

H_1 : minimal ada satu $\beta_j \neq 0$, j=1,2,3,4

Sratistik uji :

$$F = \frac{MSR}{MSE} = 1619{,}513$$

Daerah penolakan : tolak H_0 apabila $F_{hitung} > F_{\alpha(p,n-p)}$

Keputusan :

Karena nilai F_{hitung} = 1619,513 lebih besar dari F_{tabel} = 9,013455, maka tolak H_0 yang berarti terdapat parameter yang signifikan dalam model *spline* kuadratik pada taraf α=5%.

(ii) Uji t

Akan diuji hipotesa :

$H_0 : \beta_j = 0$

$H_1 : \beta_j \neq 0$, j=1,2,3,4

Sratistik uji : $t_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}$

Tabel 4.4 Uji Parsial model *spline* kuadratik Denpasar

Parameter	Estimasi	Nilai t_{hitung}
β_1	-20,8904	-4,3976
β_2	0,9048	7,9599
β_3	-6,8969	-25,5832
β_4	6,3035	35,8506

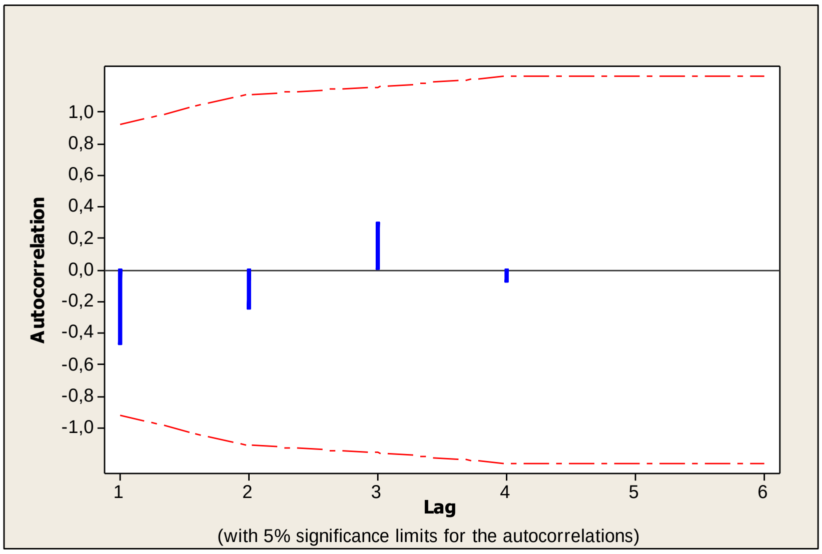
Daerah penolakan : H_0 ditolak jika $|t_{hitung}| > t_{\alpha/2,(n-p)}$

Keputusan :

Dengan menggunakan taraf α=5%, maka nilai t_{tabel} = 3,182446. Sehingga H_0 ditolak, karena nilai t_{hitung} semua parameter lebih besar dari t_{tabel}. Dengan demikian parameter β_1, β_2, β_3 dan β_4 signifikan dalam model *spline* kuadratik dengan dua titik *knot*.

b. Analisa Residual

Model *spline* kuadratik dikatakan layak apabila residualnya telah memenuhi asumsi. Untuk memeriksa asumsi independen dari residual, dapat dilakukan dengan menganalisa plot ACF dari residual :

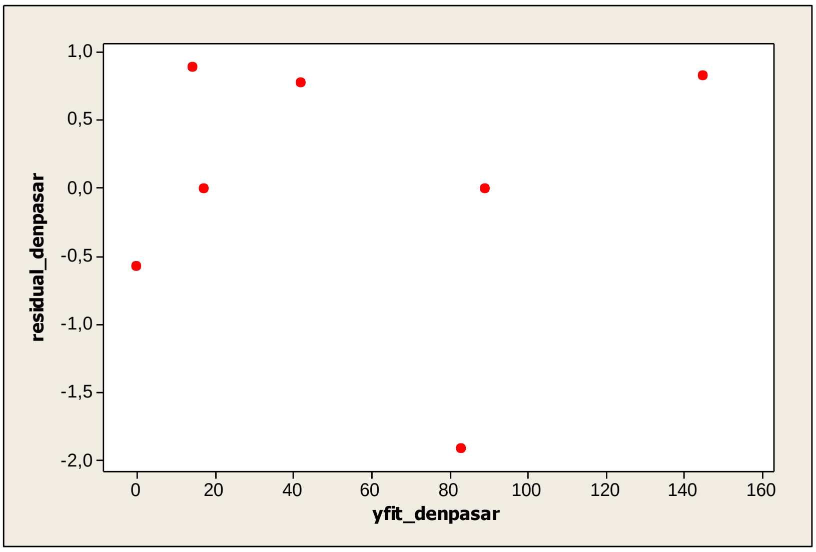


Gambar 4.8 Plot ACF residual model *spline* kuadratik dua *knot*

Dari Gambar 4.8 menunjukkan bahwa residual model *spline* kuadratik telah memenuhi asumsi independen, karena dari plot ACF tersebut tidak ada lag yang keluar dari batas.

Selanjutya dilakukan analisa terhadap residual apakah sudah memenuhi asumsi identik yaitu dengan melihat plot antara residual dengan y taksiran seperti yang terlihat pada Gambar 4.9.

Dari Gambar 4.9 dapat dilihat bahwa data sudah menyebar secara acak di sekitar nilai nol dan tidak membentuk suatu pola sehingga dapat dikatakan bahwa residual telah memenuhi asumsi identik.



Gambar 4.9 Plot residual dengan y taksiran *spline* kuadratik

Asumsi terakhir yang harus dipenuhi oleh residual adalah kenormalan. Untuk mengetahui apakah residual telah berdistribusi normal dapat dilihat dari uji *Kolmogorv Smirnov* residual.

Hipotesanya :

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

Statistik Uji : $D = Sup\ |S(x) - F_{0(x)}| = 0{,}214$

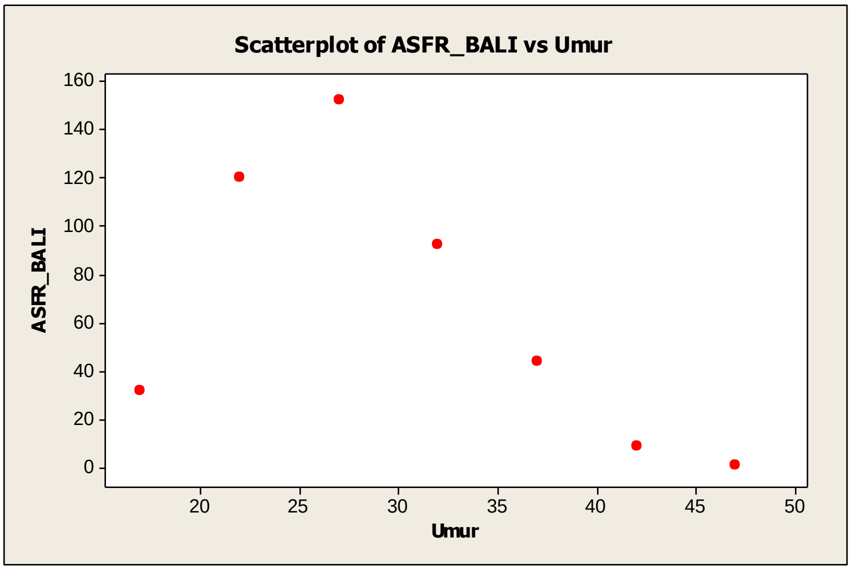
Daerah penolakan : tolak H_o apabila $D_{hitung} > D_{1-\alpha(n)}$ tabel *Kolmogorv Smirnov*

Keputusan :

Dengan menggunakan taraf α=5%., maka gagal tolak H_0 karena nilai D_{hitung} = 0,214 kurang dari $D_{0,05(7)}$=.0,483 , sehingga dapat dinyatakan residual model *spline* kuadratik dengan dua *knot* berdistribusi normal.

4.5 Model *Spline* Kuadratik untuk ASFR Prov Bali

Untuk melihat pola hubungan antara ASFR dengan umur ibu Provinsi Bali tahun 2013, diberikan plotnya sebagai berikut:



Gambar 4.10 Plot antara ASFR dengan umur ibu di Provinsi Bali

Dari gambar di atas terdapat indikasi adanya perubahan pola perilaku data. Disamping itu, terlihat pola data yang cenderrung kuadratik. Oleh karena itu, dalam penelitian ini penulis akan mencoba melakukan pendekatan menggunakan regresi nonparametrik *spline* kuadratik.

Dalam penggunaan regresi *Spline*, langkah awal yang harus dilakukan adalah menentukan lokasi dari titik *knot*. Cara pemilihan *knot* dengan cara mencoba-coba pada tempat yang terlihat adanya perubahan perilaku data pada selang tertentu. Dari berbagai *knot* dipilih nilai GCV minimum. Untuk mendapatkan hasil yang terbaik dalam regresi *spline* dicoba model *spline* kuadratik dengan satu *knot* sampai tiga *knot*. Selanjutnya menghitung nilai estimasi parameter dan membuat kurva estimasinya.

Untuk perhitungan GCV dilakukan dengan fasilitas Program S-PLUS. Akan dicari GCV untuk 1 *knot*, 2 *knot* dan 3 *knot* model *spline* kuadratik, sebagai berikut:

a. Model kuadratik 1 titik *knot*:

$$y = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 (x_i - k_1)_+^2 + \varepsilon$$

b. Model kuadratik 2 titik *knot*:

$$y = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 (x_i - k_1)_+^2 + \beta_4 (x_i - k_2)_+^2 + \varepsilon$$

c. Model kuadratik 3 titik *knot*:

$$y = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 (x_i - k_1)_+^2 + \beta_4 (x_i - k_2)_+^2 + \beta_5 (x_i - k_3)_+^2 + \varepsilon$$

Hasil perhitungan bisa dilihat Tabel 4.5 di bawah ini.

Tabel 4.5 Skor GCV untuk 1,2 dan 3 *Knot* pada Model Kuadratik Bali

Banyak Titik *Knot*	**Lokasi Titik *Knot***			**Skor GCV**	**Rsq**
1 *Knot*	29			152,1885	99,54%
	30			83,06019	
	31			**71,15876***	
	32			134,2305	
	33			265,3245	
2 *Knot*	24	27		43,6	99,95%
	24	28		29,42144	
	24	29		17,15*	
	24	30		39,27533	
	24	31		124,,6784	
3 *Knot*	25	27	38	5,974895	99,99%
	25	27	39	1,422083	
	25	27	40	0,0786247*	
	25	27	41	0,110822	
	25	27	42	0,35	

Berdasarkan Tabel 4.5 terlihat bahwa nilai GCV terkecil adalah 0,07862477 terdapat pada model *spline* kuadratik dengan tiga titik *knot* yaitu pada ibu yang berumur 25 tahun, 27 tahun dan 40 tahun. Untuk kesederhanaan model, dipilih model *spline* dengan satu titik *knot* pada ibu berumur 31 tahun. Hal ini dilakukan karena nilai R^2 untuk model *spline* kuadratik satu *knot* R^2 = 99,54% sudah relaitf besar dan hampir sama dengan dua dan tiga *knot*.

Setelah diperoleh skor GCV minimum untuk model *spline* kuadratik, langkah selanjutnya yaitu menghitung estimasi untuk model *spline* kuadratik. Estimasi model *spline* kuadratik dengan satu titik *knot* (31) sebagai berikut:

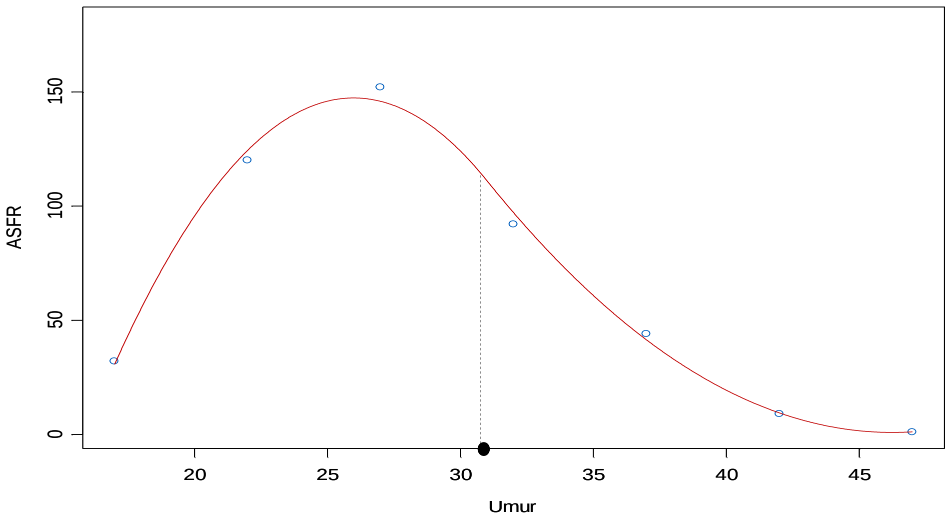
$$\hat{y} = -827{,}55 + 75{,}02x_i - 1{,}44x_i^2 - 1{,}92(x_i - 31)_+^2$$

Estimasi model tersebut dapat ditulis menjadi:

$$\hat{y} = \begin{cases} -827{,}55 + 75{,}02x_i - 1{,}44x_i^2, & x_i \le 31 \\ 1017{,}57 - 44{,}02x_i + 0{,}48x_i^2, & x_i \ge 31 \end{cases}$$

Estimasi model *spline* di atas dapat diinterpretasikan bahwa di provinsi Bali pada wanita/ibu yang berusia diatas 31 tahun, jumlah bayi yang dilahirkan meningkat 75 bayi dilahirkan oleh 1000 wanita dalam 1 tahun. Ditandai dengan bentuk kurva yang meningkat dimana puncaknya terletak pada titik di usia 25 tahun, yang berarti pada kabupaten Karangasem jumlah bayi paling banyak dilahirkan oleh wanita di usia 26 tahun. Sedangkan pada wanita/ibu yang usianya lebih dari 31 tahun masih produktif melahirkan, tetapi mulai mengalami penurunan sebesar 44 bayi yang dilahirkan oleh 1000 wanita dalam 1 tahun. Hal ini kemungkinan diakibatkan karena masih ada wilayah-wilayah di Bali yang masi terpelosok, sehingga kurang mendapat penyuluhan yang merata tentang KB dan banyak yang menikah di usia muda. Masih banyak wanita yang tidak berkesempatan mengenyam pendidikan tinggi karena tidak mempertimbangkan usia dalam merencanakan usia kelahiran maka di usia lebih dari 31 tahun pun masih aktif melahirkan.

Estimasi model *spline* kuadratik dengan satu titik *knot* (31) untuk ASFR Bali mempunyai nilai koefisien determinasi R^2 = 99,54% dan *Mean Square Error* = 13,06998. Jika ditampilkan dalam gambar, visualisasi model astimasi kurva *spline* kuadratik tersebut ditampilkan pada Gambar 4.11.



Gambar 4.11 Kurva *Spline* Kuadratik Provinsi Bali dengan titik *knot* pada umur ibu 31 tahun

a. Menguji Signifikansi Model

Pada tahap ini akan dilakukan pengujian terhadap koefisien-koefisien regresi pada model *spline* kuadratik, yang dilakukan dengan dua tahap yaitu uji serentak dan uji parsial.

(i) Uji Serentak (Uji F)

Akan diuji hipotesa :

$H_0 : \beta_1 = \beta_2 = \beta_3 = 0$

H_1 : minimal ada satu $\beta_j \neq 0$, j=1,2,3

Statistik uji :

$$F = \frac{MSR}{MSE} = 218{,}4935$$

Daerah penolakan : tolak H_0 apabila $F_{hitung} > F_{\alpha(p,n-p)}$

Keputusan :

Karena nilai F_{hitung} = 218,4935 lebih besar dari F_{tabel} = 6,38823, maka tolak H_0 yang berarti terdapat parameter yang signifikan dalam model *spline* kuadratik pada taraf α=5%.

(ii) Uji t

Akan diuji hipotesa :

$H_0 : \beta_j = 0$

$H_1 : \beta_j \neq 0$, j=1,2,3

Sratistik uji : $t_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}$

Tabel 4.6 Uji Parsial model *spline* kuadratik Bali

Parameter	Estimasi	Nilai t_{hitung}
β_1	75,0209	21,1059
β_2	-1,4432	-21,1288
β_3	1,9177	16,3990

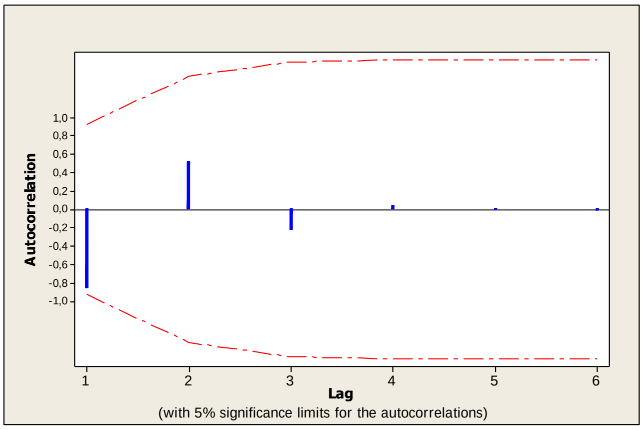
Daerah penolakan : H_0 ditolak jika $|t_{hitung}| > t_{\alpha/2,(n-p)}$

Keputusan :

Dengan menggunakan taraf α=5%, maka nilai t_{tabel} = 2,776445. Sehingga H_0 ditolak, karena nilai t_{hitung} semua parameter lebih besar dari t_{tabel}. Dengan demikian parameter β_1, β_2 dan β_3 signifikan dalam model *spline* kuadratik dengan satu titik *knot*.

b. Analisa Residual

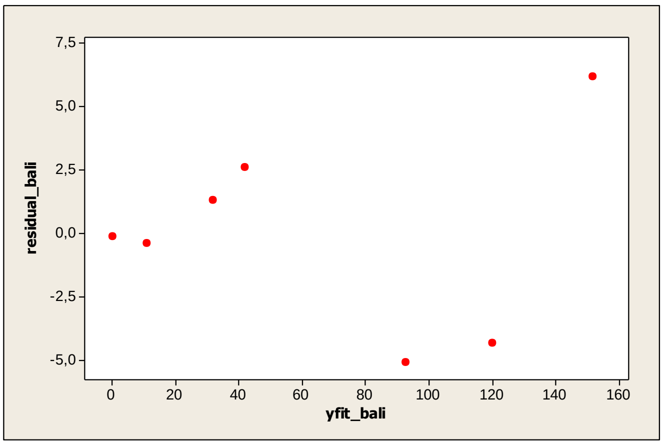
Model *spline* kuadratik dikatakan layak apabila residualnya telah memenuhi asumsi. Untuk memeriksa asumsi independen dari residual, dapat dilakukan dengan menganalisa plot ACF dari residual :



Gambar 4.12 Plot ACF residual model *spline* kuadratik satu *knot*

Dari Gambar 4.12 menunjukkan bahwa residual model *spline* kuadratik telah memenuhi asumsi independen, karena dari plot ACF tersebut tidak ada lag yang keluar dari batas.

Selanjutya dilakukan analisa terhadap residual apakah sudah memenuhi asumsi identik yaitu dengan melihat plot antara residual dengan y taksiran seperti yang terlihat pada Gambar 4.13. Dari Gambar 4.13 dapat dilihat bahwa data sudah menyebar secara acak di sekitar nilai nol dan tidak membentuk suatu pola sehingga dapat dikatakan bahwa residual telah memenuhi asumsi identik.



Gambar 4.13 Plot residual dengan y taksiran *spline* kuadratik

Asumsi terakhir yang harus dipenuhi oleh residual adalah kenormalan. Untuk mengetahui apakah residual telah berdistribusi normal dapat dilihat dari uji *Kolmogorv Smirnov* residual.

Hipotesanya :

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

Statistik Uji : $D = Sup\ |S(x) - F_{0(x)}| = 0{,}172$

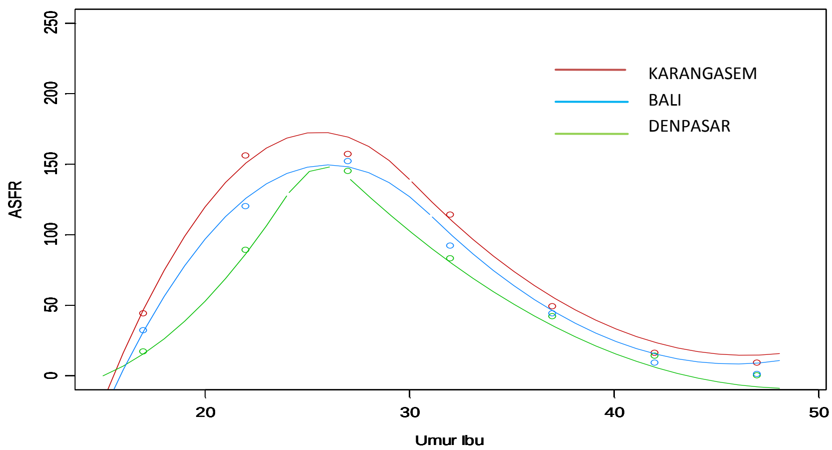
Daerah penolakan : tolak H_o apabila $D_{hitung} > D_{1-\alpha(n)}$ tabel *Kolmogorv Smirnov*

Keputusan :

Dengan menggunakan taraf α=5%., maka gagal tolak H_0 karena nilai D_{hitung} = 0,172 kurang dari $D_{0,05(7)}$=.0,483, sehingga dapat dinyatakan residual model *spline* kuadratik dengan satu *knot* berdistribusi normal.

4.6 Kurva Gabungan *Spline* Kuadratik ASFR di Karangasem, Denpasar, dan Bali

Dari analisa data menggunakan regresi nonparamterik *spline* kuadratik terhadap ASFR di Karangasem, Denpasar dan Bali telihat bahwa terdapat perbedaan pola dan lokasi titik knot yang merupakan letak banyaknya jumlah bayi yang dilahirkan di usia tertentu. Menunjukkan bahwa masi ada wilayah di Bali yang masi belum mendapatkan penyuluhan yang merata dari BKKBN Bali tentang program KB. Visualisasinya dapat dilihat pada gambar berikut:



Gambar 4.14 Kurva *Spline* Kuadratik ASFR di Karangasem, Denpasar dan Bali

Pada Gambar 4.14 terdapat 3 kurva, yaitu kurva ASFR Karangasem yang berwarna merah. kurva ASFR Denpasar yang berwarna hjau dan kurva ASFR Bali yang berwarna biru. Dari gambar tersebut terlihat bahwa pada kurva Karangasem dan Bali mengalami banyak kelahiran di usia ibu/wanita kurang dari 30 dan 31 tahun dilihat dari bentuk kurvanya yang melengkung naik dengan cepat, sedangkan pada kurva Denpasar dapat dilihat pada usia ibu/wanita kurang dari 24 tahun mengalami peningkatan yang sedikit sehingga bentuk kurvanya yang naik dengan lambat. Untuk ketiga kurva tersebut puncaknya sama-sama terdapat di titik usia 26 tahun. Dan untuk usia ibu/wanita lebih dari 30 dan 31 tahun di Karangasem dan Bali mengalami penurunan yang melambat berarti di usia tersebut masi banyak wanita yang melahirkan, sedangkan untuk Denpasar untuk usia ibu/wanita diatas 27 tahun mulai mengalami penurunan yang cepat berarti wanita di Denpasar di usia tersebut sudah tidak aktif melahirkan lagi.

Hal seperti ini menunjukkan bahwa pergerakan ASFR di provinsi Bali belum baik dan program KB di Bali belum berhasil, karena aktifitas melahirkannya masih relatif tinggi, khususnya di daerah-daerah yang agak terpelosok. Sehingga, ke depannya pemerintahan daerah bekerjasama dengan BKKBN, lebih fokus mengedukasi dengan memberikan banyak penyuluhan merata di seluruh daerah-daerah yang ada di provinsi Bali.

4.7 Interval Konfidensi untuk ASFR di Karangasem, Denpasar dan Bali

Setelah mendapatkan model *spline* terbaik untuk ASFR di kabupaten Karngasem, kota Denpasar dan provinsi Bali, selanjutnya akan dibangun konfidensi intervalnya yang berfungsi sebagai prediksi batas maksimal dan minimal jumlah anak yang sesuai dengan rata-rata nya pada data di 3 wilayah tersebut.

4.7.1 Interval Konfidensi untuk ASFR di Kabupaten Karangasem

Dengan model *spline* kuadratik dengan satu titik *knot* yaitu di umur ibu 30 tahun, akan dibangun interval konfidensi 95% dengan persamaan:

$$P\left(\hat{f}(x_i) - z_{\alpha/2}\sqrt{var\hat{f}(x_i)} < f(x_i) < \hat{f}(x_i) + z_{\alpha/2}\sqrt{var\hat{f}(x_i)}\right) = 0{,}95$$

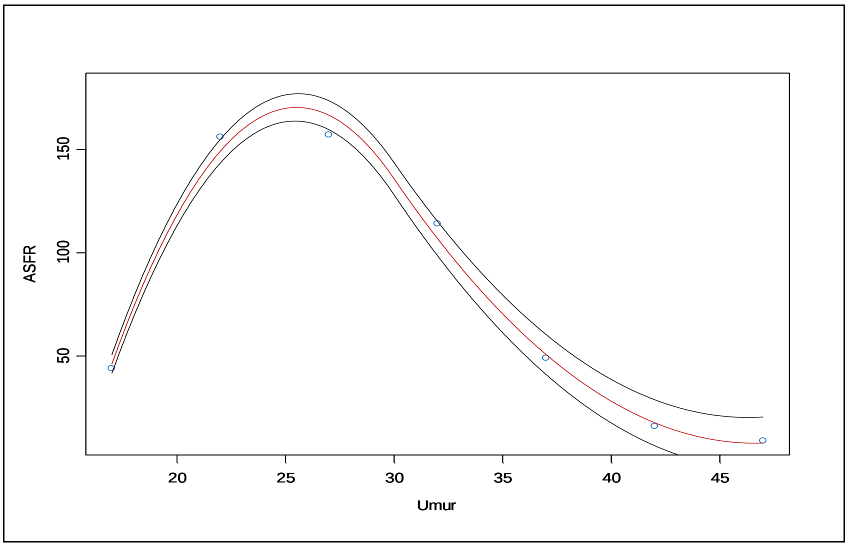
$$P\left(\left[\hat{\beta}_0+\hat{\beta}_1x_i+\hat{\beta}_2x_i^2+\hat{\beta}_3(x_i-k_1)_+^2\right]-z_{\alpha/2}\sqrt{var\hat{f}(x_i)}<f(x_i)<\right.$$

$$\left.\left[\hat{\beta}_0+\hat{\beta}_1x_i+\hat{\beta}_2x_i^2+\hat{\beta}_3(x_i-k_1)_+^2\right]+z_{\alpha/2}\sqrt{var\hat{f}(x_i)}\right)=0{,}95$$

$$P\left(\left[\left(-945{,}29+87{,}45x_i-1{,}71x_i^2+2{,}17(x_i-30)_+^2\right)\right]-1{,}96\sqrt{var\hat{f}(x_i)}<f(x_i)<\right.$$

$$\left.\left[\left(-945{,}29+87{,}45x_i-1{,}71x_i^2+2{,}17(x_i-30)_+^2\right)\right]+1{,}96\sqrt{var\hat{f}(x_i)}\right)=0{,}95$$

Interval konfidensi 95% untuk kurva regresi $\hat{f}(x_i)$ ditunjukkan oleh Gambar 4.13



Gambar 4.15 Interval Konfidensi *spline* kuadratik untuk ASFR Karangasem dengan 1 titik *knot*

Berdasarkan Gambar 4.15 dan Tabel 4.7 dari model *spline* kuadratik yang diperoleh $\hat{y}=-945{,}29+87{,}45x_i-1{,}71x_i^2+2{,}17(x_i-30)_+^2$, diketahui bahwa pada tingkat signifikansi 95%, ASFR di kabupaten Karangasem masih berada di dalam kisaran *mean* nya. Dari gambar di atas dan tabel di bawah ini dapat dilihat batas atas dan bawahnya, yang dapat membantu pembaca khususnya pemerintah daerah dalam memprediksi jumlah kelahiran bayi yang baik adalah yang tidak lebih dan tidak kurang dari batas atas dan batas bawah konfidensi interval yang didapat.

Tabel 4.7 Batas atas dan batas bawah Konfidensi Interval ASFR Karangasem

Umur (x_i)	x_1	x_2	x_3	x_4	x_5	x_6	x_7
Batas atas	59,51	158,52	176,49	114,99	60,35	26,17	20,99
$\hat{f}(x_i)$	46,09	149,14	166,49	106,84	50,74	17,77	7,92
Batas bawah	32,67	139,76	156,49	98,70	41,13	9,36	-5,15

4.7.2 Interval Konfidensi untuk ASFR di Kota Denpasar

Dengan model *spline* kuadratik dengan dua titik *knot* yaitu di umur ibu 24 tahun dan 27 tahun, akan dibangun interval konfidensi 95% dengan persamaan:

$$P(\hat{f}(x_i) - z_{\alpha/2}\sqrt{var\hat{f}(x_i)} < f(x_i) < \hat{f}(x_i) + z_{\alpha/2}\sqrt{var\hat{f}(x_i)}) = 0{,}95$$

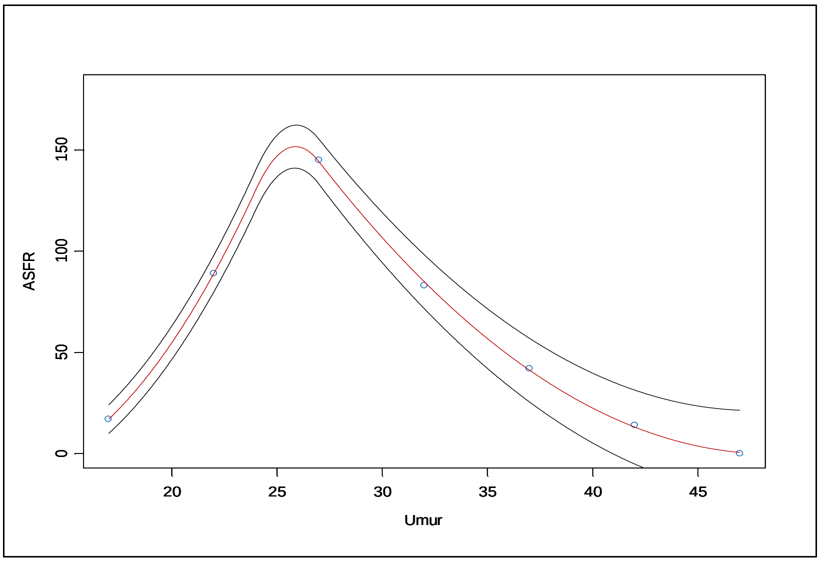
$$P\Big([\hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2 + \hat{\beta}_3 (x_i - k_1)_+^2 + \hat{\beta}_4 (x_i - k_2)_+^2] - z_{\alpha/2}\sqrt{var\hat{f}(x_i)} < f(x_i) <$$

$$[\hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2 + \hat{\beta}_3 (x_i - k_1)_+^2 + \hat{\beta}_4 (x_i - k_2)_+^2] + z_{\alpha/2}\sqrt{var\hat{f}(x_i)}\Big) = 0{,}95$$

$$P\Big([(110{,}63 - 20{,}89x_i + 0{,}90x_i^2 - 6{,}89(x_i - 24)_+^2 + 6{,}30(x_i - 27)_+^2)] - 1{,}96\sqrt{var\hat{f}(x_i)} < f(x_i) <$$

$$[(110{,}63 - 20{,}89x_i + 0{,}90x_i^2 - 6{,}89(x_i - 24)_+^2 + 6{,}30(x_i - 27)_+^2)] + 1{,}96\sqrt{var\hat{f}(x_i)}\Big) = 0{,}95$$

Interval konfidensi 95% untuk kurva regresi $\hat{f}(x_i)$ ditunjukkan oleh Gambar 4.14



Gambar 4.16 Interval Konfidensi *spline* kuadratik untuk ASFR Denpasar dengan 2 titik *knot*

Berdasarkan Gambar 4.16 dan Tabel 4.8 dari model *spline* kuadratik yang diperoleh yaitu $\hat{y} = 110{,}63 - 20{,}89x_i + 0{,}90x_i^2 - 6{,}89(x_i - 24)_+^2 + 6{,}30(x_i - 27)_+^2$, diketahui bahwa pada tingkat signifikansi 95%, ASFR di kota Denpasar masih berada di dalam kisaran *mean* nya. Dari gambar di atas dan tabel di bawah ini dapat dilihat batas atas dan bawahnya, yang dapat membantu pembaca khususnya pemerintah daerah dalam memprediksi jumlah kelahiran bayi yang baik adalah yang tidak lebih dan tidak kurang dari batas atas dan batas bawah konfidensi interval yang didapat.

Tabel 4.8 Batas atas dan batas bawah Konfidensi Interval ASFR Denpasar

Umur (x_i)	x_1	x_2	x_3	x_4	x_5	x_6	x_7
Batas atas	19,78	91,78	146,79	86,61	43,17	14,81	3,19
$\hat{f}(x_i)$	17	88,99	144,17	84,91	41,23	13,11	0,57
Batas bawah	14,21	86,21	141,55	83,22	39,29	11,42	-2,05

4.7.3 Interval Konfidensi untuk ASFR di Provinsi Bali

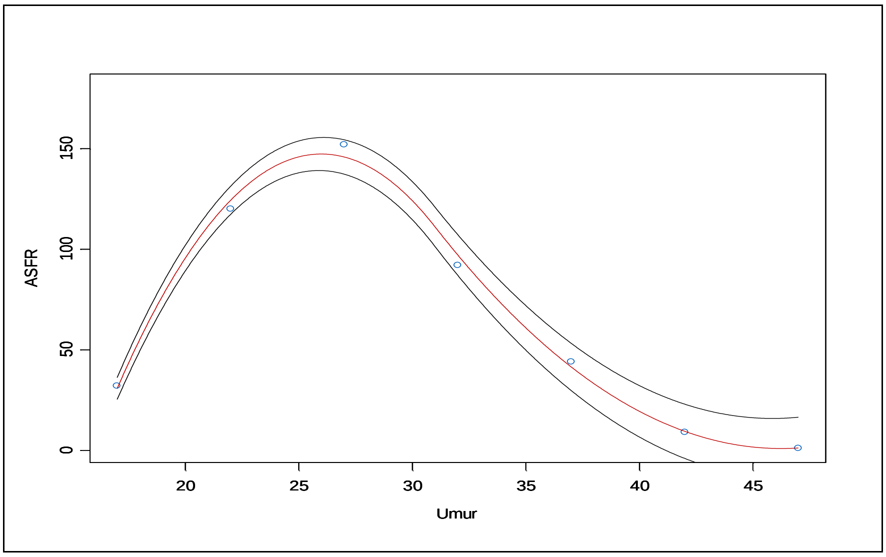
Dengan model *spline* kuadratik dengan satu titik *knot* yaitu di umur ibu 31 tahun, akan dibangun interval konfidensi 95% dengan persamaan:

$$P(\hat{f}(x_i) - z_{\alpha/2}\sqrt{var\hat{f}(x_i)} < f(x_i) < \hat{f}(x_i) + z_{\alpha/2}\sqrt{var\hat{f}(x_i)}) = 0{,}95$$

$$P\Big([\hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2 + \hat{\beta}_3 (x_i - k_1)_+^2] - z_{\alpha/2}\sqrt{var\hat{f}(x_i)} < f(x_i) <$$
$$[\hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2 + \hat{\beta}_3 (x_i - k_1)_+^2] + z_{\alpha/2}\sqrt{var\hat{f}(x_i)}\Big) = 0{,}95$$

$$P\Big([(-827{,}547 + 75{,}021x_i - 1{,}443x_i^2 - 1{,}917(x_i - 31)_+^2)] - 1{,}96\sqrt{var\hat{f}(x_i)} < f(x_i) <$$
$$[(-827{,}547 + 75{,}021x_i - 1{,}443x_i^2 - 1{,}917(x_i - 31)_+^2)] + 1{,}96\sqrt{var\hat{f}(x_i)}\Big) = 0{,}95$$

Interval konfidensi 95% untuk kurva regresi $\hat{f}(x_i)$ ditunjukkan oleh Gambar 4.15



. **Gambar 4.17** Interval Konfidensi *spline* kuadratik untuk ASFR Bali dengan 1 titik *knot*.

Berdasarkan Gambar 4.17 dan Tabel 4.9 dari model *spline* kuadratik yang diperoleh $\hat{y} = -827{,}55 + 75{,}02x_i - 1{,}44x_i^2 - 1{,}92(x_i - 31)_+^2$, diketahui bahwa pada tingkat signifikansi 95%, ASFR di kota Denpasar masih berada di dalam kisaran *mean*-nya. Dari gambar di atas dan tabel di bawah ini dapat dilihat batas atas dan bawahnya, yang dapat membantu pembaca khususnya pemerintah daerah dalam memprediksi jumlah kelahiran bayi yang baik adalah yang tidak lebih dan tidak kurang dari batas atas dan batas bawah konfidensi interval yang didapat.

Tabel 4.9 Batas atas dan batas bawah Konfidensi Interval ASFR Bali

Umur (x_i)	x_1	x_2	x_3	x_4	x_5	x_6	x_7
Batas atas	39,72	130,51	152,64	102,55	48,01	15,23	10,05
$\hat{f}(x_i)$	30,69	124,36	145,86	97,11	41,40	9,41	1,15
Batas bawah	21,67	118,21	139,08	91,67	34,79	3,60	-7,75

\

(*Halaman ini sengaja dikosongkan*)

BAB 5
KESIMPULAN

5.2 Kesimpulan

Berdasarkan analisis dan pembahasan yang telah dilakukan pada bab 4 dpat diambil kesimpulan sebagai berikut:

1. Konfidensi interval dalam regresi nonparametrik *spline* kuadratik dibangun dengan menggunakan konsep Pivotal *Quantity*. *Pivotal Quantity* untuk kurva regresi f adalah :

   $$Z(x_1, x_2, \dots, x_n) = \frac{\hat{f}(x_i) - f(x_i)}{\sqrt{var(\hat{f}(x_i))}} \sim N(0,1)\,,\; i = 1,2,\dots,n,$$

   Interval konfidensi $1 - \alpha$ diperoleh dari menyelesaikan persamaan probabilitas :

   $$P(-z\alpha/_2 < \frac{\hat{f}(x_i) - f(x_i)}{\sqrt{var(\hat{f}(x_i))}} < z\alpha/_2 = 1 - \propto.$$

   Interval Konfidensi untuk kurva regresi f diberikan oleh:

   $$P\left(\left[\hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2 + \sum_{k=1}^{r} \hat{\beta}_{2+k}\,(x_i - k_k)_+^2\right] - z\alpha/_2\sqrt{var\hat{f}(x_i)} < f(x_i) < \left[\hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2 + \sum_{k=1}^{r} \hat{\beta}_{2+k}\,(x_i - k_k)_+^2\right] + z\alpha/_2\sqrt{var\hat{f}(x_i)}\right) = 1 - \alpha$$

2. Model *spline* kuadratik terbaik untuk mengestimasi hubungan antara ASFR dan umur ibu adalah model *spline* kuadratik dengan dua titik *knot* untuk kota Denpasar dan model *spline* kuadratik dengan satu titik *knot* untuk provinsi Bali dan kabupaten Karangasem.

   a. Model estimasi ASFR di kabupaten Karangasem

   $$\hat{y} = -945{,}29 + 87{,}45x_i - 1{,}71x_i^2 + 2{,}17(x_i - 30)_+^2$$
   $$= \begin{cases} -945{,}29 + 87{,}45x_i - 1{,}71x_i^2, & x_i \le 30 \\ 1007{,}71 - 42{,}75x_i + s0{,}46x_i^2, & x_i \ge 30 \end{cases}$$

   Dengan R^2 sebesar 99,17%

   b. Model estimasi ASFR di kota Denpasar

   $$\hat{y} = 110{,}63 - 20{,}89x_i + 0{,}90x_i^2 - 6{,}89(x_i - 24)_+^2 + 6{,}30(x_i - 27)_+^2$$
   $$\hat{y} = \begin{cases} 110{,}63 - 20{,}89x_i + 0{,}90x_i^2, & x_i \le 24 \\ -3858{,}01 + 309{,}83x_i - 5{,}99x_i^2, & 24 \le x_i \le 27 \\ 734{,}69 - 30{,}37x_i + 0{,}31x_i^2, & x_i \ge 27 \end{cases}$$

   Dengan R^2 sebesar 99,96%

c. Model estimasi ASFR di provinsi Bali

$\hat{y} = -827{,}55 + 75{,}02x_i - 1{,}44x_i^2 - 1{,}92(x_i - 31)_+^2$

$= \begin{cases} -827{,}55 + 75{,}02x_i - 1{,}44x_i^2, & x_i \leq 31 \\ 1017{,}57 - 44{,}02x_i + 0{,}48x_i^2, & x_i \geq 31 \end{cases}$

Dengan R^2 sebesar 99,54%

1. Setelah mendapatkan model *spline* terbaik untuk ASFR di kabupaten Karangasem, kota Denpasar dan provinsi Bali, selanjutnya akan dibangun konfidensi intervalnya.

a. Dengan model *spline* kuadratik untuk ASFR di Karangasem dengan satu titik *knot* yaitu di umur ibu 30 tahun, akan dibangun interval konfidensi 95% dengan persamaan:

$$P(\hat{f}(x_i) - z\alpha_{/2}\sqrt{var\hat{f}(x_i)} < f(x_i) < \hat{f}(x_i) + z\alpha_{/2}\sqrt{var\hat{f}(x_i)}) = 0{,}95$$

$$P\Big(\left[\hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2 + \hat{\beta}_3 (x_i - k_1)_+^2\right] - z\alpha_{/2}\sqrt{var\hat{f}(x_i)} < f(x_i) <$$
$$\left[\hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2 + \hat{\beta}_3 (x_i - k_1)_+^2\right] + z\alpha_{/2}\sqrt{var\hat{f}(x_i)}\Big) = 0{,}95$$

$$P\Big(\left[\left(-945{,}29 + 87{,}45x_i - 1{,}71x_i^2 + 2{,}17(x_i - 30)_+^2\right)\right] - 1{,}96\sqrt{var\hat{f}(x_i)} < f(x_i) <$$
$$\left[\left(-945{,}29 + 87{,}45x_i - 1{,}71x_i^2 + 2{,}17(x_i - 30)_+^2\right)\right] + 1{,}96\sqrt{var\hat{f}(x_i)}\Big) = 0{,}95$$

b. Dengan model *spline* kuadratik untuk ASFR di Denpasar dengan model *spline* kuadratik dengan dua titik *knot* yaitu di umur ibu 24 tahun dan 27 tahun, akan dibangun interval konfidensi 95% dengan persamaan:

$$P(\hat{f}(x_i) - z\alpha_{/2}\sqrt{var\hat{f}(x_i)} < f(x_i) < \hat{f}(x_i) + z\alpha_{/2}\sqrt{var\hat{f}(x_i)}) = 0{,}95$$

$$P\Big(\left[\hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2 + \hat{\beta}_3 (x_i - k_1)_+^2 + \hat{\beta}_4 (x_i - k_2)_+^2\right] - z\alpha_{/2}\sqrt{var\hat{f}(x_i)} < f(x_i) <$$
$$\left[\hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2 + \hat{\beta}_3 (x_i - k_1)_+^2 + \hat{\beta}_4 (x_i - k_2)_+^2\right] + z\alpha_{/2}\sqrt{var\hat{f}(x_i)}\Big) = 0{,}95$$

$$P\Big(\left[\left(110{,}63 - 20{,}89x_i + 0{,}90x_i^2 - 6{,}89(x_i - 24)_+^2 + 6{,}30(x_i - 27)_+^2\right)\right] - 1{,}96\sqrt{var\hat{f}(x_i)} < f(x_i) <$$
$$\left[\left(110{,}63 - 20{,}89x_i + 0{,}90x_i^2 - 6{,}89(x_i - 24)_+^2 + 6{,}30(x_i - 27)_+^2\right)\right] + 1{,}96\sqrt{var\hat{f}(x_i)}\Big) = 0{,}95$$

c. Dengan model *spline* kuadratik untuk ASFR di Bali dengan model *spline* kuadratik dengan satu titik *knot* yaitu di umur ibu 24 tahun dan 29 tahun, akan dibangun interval konfidensi 95% dengan persamaan:

$$P(\hat{f}(x_i) - z\alpha_{/2}\sqrt{var\hat{f}(x_i)} < f(x_i) < \hat{f}(x_i) + z\alpha_{/2}\sqrt{var\hat{f}(x_i)}) = 0{,}95$$

$$P\left(\left[\hat{\beta}_0+\hat{\beta}_1 x_i+\hat{\beta}_2 x_i^2+\hat{\beta}_3 (x_i-k_1)_+^2\right]-z\alpha_{/2}\sqrt{var\hat{f}(x_i)}<f(x_i)<\right.$$

$$\left.\left[\hat{\beta}_0+\hat{\beta}_1 x_i+\hat{\beta}_2 x_i^2+\hat{\beta}_3 (x_i-k_1)_+^2\right]+z\alpha_{/2}\sqrt{var\hat{f}(x_i)}\right)=0{,}95$$

$$P\left(\left[\left(-827{,}547+75{,}021x_i-1{,}443x_i^2-1{,}917(x_i-31)_+^2\right)\right]-1{,}96\sqrt{var\hat{f}(x_i)}<f(x_i)<\right.$$

$$\left.\left[\left(-827{,}547+75{,}021x_i-1{,}443x_i^2-1{,}917(x_i-31)_+^2\right)\right]+1{,}96\sqrt{var\hat{f}(x_i)}\right)=0{,}95$$

5.2 Saran

Berdasarkan hasil penelitian yang diperoleh, saran yang dapat diberikan penulis adalah sebagai berikut:

1. Untuk penelitian selanjutnya mengenai pendekatan regresi nonparametik *spline* menggunakan *spline* umum berorde m > 2 untuk mengestimasi data ASFR baik di provinsi Bali maupun provinsi-provinsi lainnya.
2. Dapat menggunakan inferensi statistik lainnya selain interval konfidensi pada penelitian berikutnya.

\

(*Halaman ini sengaja dikosongkan*)

DAFTAR PUSTAKA

Basri,H. (2008). *Estimasi Kurva Regresi Nonparametrik dengan Pendekatan Spline (Studi Kasus pada Data Murid Madrasah Ibtidaiyah dan Keluarga Prasejahtera Setiap Kecamatan di Kabupaten Bone).* Didaktika Jurnal Kependidikan, 3, 2.

Budiantara, I.N. (2000). *Metode UBR, GML, CV, dan GCV dalam Regresi Nonparametrik Spline. Majalah Ilmiah Himpunan Matematika Indonesia (MIHMI)*, 6, 285-290.

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 Knots Optimal*. Jurnal Ilmu Dasar. FMIPA Universitas Jember, Vol.7, Hal. 77-85.

Budiantara, I. N. (2009). *Spline dalam Regresi Nonparametrik dan Semiparametrik : Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatan".* Pidato Pengukuhan untuk Jabatan Guru Besar dalam Bidang Ilmu Matematika Statistika dan Probabilitas, pada Jurusan Statistika, Fakultas MIPA. Surabaya: ITS Press.

Daniel, W., Wayne, (1989) *Statistika Nonparametrik Terapan*, PT Gramedia Pustaka Utama Jakarta. Diterjemahkan oleh Alex Tri Kantjono W.

Dewi, R. K. (2012). *Faktor-Faktor yang Mempengaruhi Angka Gizi Buruk di Jawa Timur dengan Pendekatan Regresi Nonparametrik Spline*. Tugas Akhir, Jurusan Statistika Institut Teknologi Sepuluh Nopember, Surabaya.

Draper, N.R. and Smith H. (1996). *Applied Regression Analysis,* Second Edition. John Wiley & Sons, Inc, New York.

Eubank, R.L. (1988). *Spline Smoothing and Nonparametric Regression*. New York : Marcel Deker.

Gujarati, D.N., 2004, *Basic Econometrics 4^{th} edition*, The Mc Graw Hill Companies, New York. Wahba, G. 1990. *Spline Models For Observational Data.* Pennysylvania: SIAM.

Hardle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press, New York.

Mundiharno. (1998). *Beberapa Teknik Estimasi Fertilitas*. Jakarta.

Sari, D. W. (2006). Model Regresi *Spline* untuk Mengestimasi Rata-rata Angka Kelahiran ASFR (*Age Spesific Fertility Rate*) di Jawa Timur. Skripsi, Jurusan Statistika Institut Teknologi Sepuluh Nopember, Surabaya.

Sumarno, H., Jemain, A. Z., Mahir, A., dan Norsiah, W. (1997). *Penerapan Model Fertilitas Perkawinan terhadap Data Jawa-Bali*. Forum Statistika dan Komputasi, Maret 1997, ISSN 0853-8115.

Sembiring, R.K. (1985). *DEMOGRAFI*. Fakultas Pasca Sarjana IKIP Jakarta dan BKKBN Jakarta, Jakarta.

Wahba, G. (1990). *Spline Models For Observational Data.* Pennysylvania: SIAM.

Wang, Y. (1998). *Spline Smoothing Models With Correlated Error*. *Journal of The American Statistical Association.* 93, 341-348.

LAMPIRAN

Lampiran 1.

Data ASFR Tahun 2013

Umur Ibu	ASFR Karangasem	ASFR Denpasar	ASFR Bali
17	44	17	32
22	156	89	120
27	157	145	152
32	114	83	92
37	49	42	44
42	16	14	9
47	9	0	1

Lampiran 2.

Program *Spline* Kuadratik 1 Knots

```
trun <- function (x, a, polinomial)
{
 x[x<a] <- a
 (x-a)^polinomial
}
spline.1.kuadratik<-function(umur,ASFR,k1)
{
        y<-ASFR
        n<-length(y)
        trun<-function(data,knots,kuadratik)
        {((data-knots)^kuadratik)*(data>=knots)}
        derajat<-2
        jumk<-1
        p<-derajat+jumk
        x<-matrix(0,ncol=4,nrow=n)
        x[,1]<-1
        x[,2]<-umur
        x[,3]<-umur^2
        x[,4]<-trun(umur,k1,2)
        beta<-solve(t(x)%*%x)%*%t(x)%*%y
        ytopi<-x%*%beta
        A<-x%*%solve(t(x)%*%x)%*%t(x)
        I<-matrix(0, ncol=n, nrow=n)
        for(i in 1:n)
                I[i, i]<-1
        ia<-diag(I)-A
        k<-((sum(diag(ia)))/n)^2
        msek<-(t(y-ytopi)%*%(y-ytopi))/n
```

```
        gcv<-msek/k
        residual<-y-ytopi
        ybar<-sum(y)/n
        sse<-t(y-ytopi)%*%(y-ytopi)
        sst<-t(y-ybar)%*%(y-ybar)
        ssr<-t(ytopi-ybar)%*%(ytopi-ybar)
        rsq<-ssr/sst
        mse<-as.vector(sse)/(n-p-1)
        msr<-as.vector(ssr)/p
        fhitung<-msr/mse
        q <- seq(min(umur), max(umur), length=1000)
        Q <- matrix(0, ncol=4, nrow=1000)
        Q[,1] <-1
        Q[,2] <- q
        Q[,3] <- q^2
        Q[,4] <- trun (q,k1,2)
        f <- Q %*% beta

graphsheet()
  plot(umur,ASFR,type="p",col=2,xlim=c(min(umur),max(umur)),
ylim=c(min(ASFR),max(180)), xlab="Umur", ylab="ASFR")
  par(new=T)
  plot(q,f,type="l",col=8,xlim=c(min(q),max(q)),    ylim=c(min(ASFR),max(180)),
xlab="Umur", ylab="ASFR")

return(beta,ytopi,msek,gcv,rsq)
}


umur<-c(17,22,27,32,37,42,47)
ASFRK<-c(44,156,157,114,49,16,9)
ASFRB<-c(32,120,152,92,44,9,1)
ASFRD<-c(17,89,145,83,42,14,0)

spline.1.kuadratik(umur,ASFR,k1)
```

Lampiran 3.

Program *Spline* Kuadratik 2 Knots

```
trun <- function (x, a, polinomial)
{
 x[x<a] <- a
 (x-a)^polinomial
}
spline.2.kuadratik<-function(umur,ASFR,k1,k2)
{
	y<-ASFR
	n<-length(y)
	trun<-function(data,knots,kuadratik)
	{((data-knots)^kuadratik)*(data>=knots)}
	derajat<-2
	jumk<-2
	p<-derajat+jumk
	x<-matrix(0,ncol=5,nrow=n)
	x[,1]<-1
	x[,2]<-umur
	x[,3]<-umur^2
	x[,4]<-trun(umur,k1,2)
	x[,5]<-trun(umur,k2,2)
	beta<-solve(t(x)%*%x)%*%t(x)%*%y
	ytopi<-x%*%beta
	A<-x%*%solve(t(x)%*%x)%*%t(x)
	I<-matrix(0, ncol=n, nrow=n)
	for(i in 1:n)
		I[i, i]<-1
	ia<-diag(I)-A
	k<-((sum(diag(ia)))/n)^2
	msek<-(t(y-ytopi)%*%(y-ytopi))/n
	gcv<-msek/k
	residual<-y-ytopi
	ybar<-sum(y)/n
	sse<-t(y-ytopi)%*%(y-ytopi)
	sst<-t(y-ybar)%*%(y-ybar)
	ssr<-t(ytopi-ybar)%*%(ytopi-ybar)
	rsq<-ssr/sst
	mse<-as.vector(sse)/(n-p-1)
	msr<-as.vector(ssr)/p
	fhitung<-msr/mse
	q <- seq(min(umur), max(umur), length=1000)
	Q <- matrix(0, ncol=5, nrow=1000)
	Q[,1] <-1
	Q[,2] <- q
	Q[,3] <- q^2
```

```
        Q[,4] <- trun (q,k1,2)
        Q[,5] <- trun (q,k2,2)
        f <- Q %*% beta

 graphsheet()
  plot(umur,ASFR,type="p",col=2,xlim=c(min(umur),max(umur)),
ylim=c(min(ASFR),max(180)), xlab = "Umur", ylab = "ASFR")
  par(new=T)
  plot(q,f,type="l",col=8,xlim=c(min(q),max(q)),    ylim=c(min(ASFR),max(180)),
xlab = "Umur", ylab = "ASFR")

return(beta,ytopi,msek,gcv,rsq)
}


umur<-c(17,22,27,32,37,42,47)
ASFRK<-c(44,156,157,114,49,16,9)
ASFRB<-c(32,120,152,92,44,9,1)
ASFRD<-c(17,89,145,83,42,14,0)

spline.2.kuadratik(umur,ASFR,k1,k2)
```

Lampiran 4.

Program *Spline* Kuadratik 3 Knots

```
trun <- function (x, a, polinomial)
{
 x[x<a] <- a
 (x-a)^polinomial
}
spline.3.kuadratik<-function(umur,ASFR,k1,k2,k3)
{
	y<-ASFR
	n<-length(y)
	trun<-function(data,knots,kuadratik)
	{((data-knots)^kuadratik)*(data>=knots)}
	derajat<-2
	jumk<-3
	p<-derajat+jumk
	x<-matrix(0,ncol=6,nrow=n)
	x[,1]<-1
	x[,2]<-umur
	x[,3]<-umur^2
	x[,4]<-trun(umur,k1,2)
	x[,5]<-trun(umur,k2,2)
	x[,6]<-trun(umur,k3,2)
	B<-solve(t(x)%*%x)
	beta<-B%*%t(x)%*%y
	ytopi<-x%*%beta
	A<-x%*%(solve(t(x)%*%x))%*%t(x)
	I<-matrix(0, ncol=n, nrow=n)
	for(i in 1:n)
		I[i, i]<-1
	ia<-diag(I)-A
	k<-((sum(diag(ia)))/n)^2
	msek<-(t(y-ytopi)%*%(y-ytopi))/n
	gcv<-msek/k
	residual<-y-ytopi
	ybar<-sum(y)/n
	sse<-t(y-ytopi)%*%(y-ytopi)
	sst<-t(y-ybar)%*%(y-ybar)
	ssr<-t(ytopi-ybar)%*%(ytopi-ybar)
	rsq<-ssr/sst
	mse<-as.vector(sse)/(n-p-1)
	msr<-as.vector(ssr)/p
	fhitung<-msr/mse
	q <- seq(min(umur), max(umur), length=1000)
	Q <- matrix(0, ncol=6, nrow=1000)
	Q[,1] <-1
```

```
        Q[,2] <- q
        Q[,3] <- q^2
        Q[,4] <- trun (q,k1,2)
        Q[,5] <- trun (q,k2,2)
        Q[,6] <- trun (q,k3,2)
        f <- Q %*% beta

  graphsheet()
  plot(umur,ASFR,type="p",col=2,xlim=c(min(umur),max(umur)),
ylim=c(min(ASFR),max(180)), xlab = "Umur", ylab = "ASFR")
  par(new=T)
  plot(q,f,type="l",col=8,xlim=c(min(q),max(q)),    ylim=c(min(ASFR),max(180)),
xlab = "Umur", ylab = "ASFR")

return(beta,ytopi,msek,gcv,rsq)
}

umur<-c(17,22,27,32,37,42,47)
ASFRK<-c(44,156,157,114,49,16,9)
ASFRB<-c(32,120,152,92,44,9,1)
ASFRD<-c(17,89,145,83,42,14,0)

spline.3.kuadratik(umur,ASFR,k1,k2,k3)
```

Lampiran 5.

Program Interval Konfidensi *Spline* 1 Knots

```
trun <- function (x, a, polinomial)
{
 x[x<a] <- a
 (x-a)^polinomial
}
CI.1.knot<-function(umur,ASFR,k1)
{
        y<-ASFR
        n<-length(y)
        trun<-function(data,knots,kuadratik)
        {((data-knots)^kuadratik)*(data>=knots)}
        derajat<-2
        jumk<-1
        p<-derajat+jumk
        x<-matrix(0,ncol=4,nrow=n)
        x[,1]<-1
        x[,2]<-umur
        x[,3]<-umur^2
        x[,4]<-trun(umur,k1,2)
        beta<-solve(t(x)%*%x)%*%t(x)%*%y
        ytopi<-x%*%beta
        A<-x%*%solve(t(x)%*%x)%*%t(x)
        I<-matrix(0, ncol=n, nrow=n)
        for(i in 1:n)
                I[i, i]<-1
        ia<-diag(I)-A
        k<-((sum(diag(ia)))/n)^2
        msek<-(t(y-ytopi)%*%(y-ytopi))/n
        gcv<-msek/k
        residual<-y-ytopi
        ybar<-sum(y)/n
        sse<-t(y-ytopi)%*%(y-ytopi)
        sst<-t(y-ybar)%*%(y-ybar)
        ssr<-t(ytopi-ybar)%*%(ytopi-ybar)
        rsq<-ssr/sst
        dbreg <- 4
        MSreg <- ssr/dbreg
        dbres <- n-3
        MSres <- sse/dbres
        dbtot <- dbreg+dbres
        varian <- (diag(A) %*% MSres)
        q <- seq(min(umur), max(umur), length=1000)
        f <- beta[1]+beta[2]*q+beta[3]*(q^2)+beta[4]*trun(q,k1,2)
        var <- varian[1]+varian[2]*q+varian[3]*(q^2)+varian[4]*trun(q,k1,2)
```

```
        akarvar <- sqrt(var)
        uppr <- f+0.2*akarvar
        lowr <- f-0.2*akarvar
        i <- seq(min(umur),max(umur),length=n)
        fest <- beta[1]+beta[2]*i+beta[3]*i^2+beta[4]*trun(i,k1,2)
        upper <- fest+1.96*sqrt(varian)
        lower <- fest-1.96*sqrt(varian)

         graphsheet()
  plot(umur,ASFR,type="p",col=2,xlim=c(min(umur),max(umur)),
ylim=c(min(ASFR),max(180)), xlab "Umur", ylab="ASFR")
  par(new=T)
  plot(q,f,type="l",col=8,xlim=c(min(q),max(q)), ylim=c(min(ASFR),max(180)),
xlab="Umur",ylab="ASFR")
par(new=T)
plot(q,uppr,type="l",xlim=c(min(q),max(q)), ylim=c(min(ASFR),max(180)),
xlab="Umur", ylab="ASFR")
par(new=T)
plot(q,lowr,type="l",xlim=c(min(q),max(q)), ylim=c(min(ASFR),max(180)),
xlab="Umur",ylab="ASFR")

return(upper,lower)

}

umur<-c(17,22,27,32,37,42,47)
ASFRK<-c(44,156,157,114,49,16,9)
ASFRB<-c(32,120,152,92,44,9,1)
ASFRD<-c(17,89,145,83,42,14,0)

CI.1.knot(umur,ASFR,k1)
```

Lampiran 6

Program Interval Konfidensi *Spline* 2 Knots

```
trun <- function (x, a, polinomial)
{
 x[x<a] <- a
 (x-a)^polinomial
}
CI.2.knot<-function(umur,ASFR,k1,k2)
{
	y<-ASFR
	n<-length(y)
	trun<-function(data,knots,kuadratik)
	{((data-knots)^kuadratik)*(data>=knots)}
	derajat<-2
	jumk<-2
	p<-derajat+jumk
	x<-matrix(0,ncol=5,nrow=n)
	x[,1]<-1
	x[,2]<-umur
	x[,3]<-umur^2
	x[,4]<-trun(umur,k1,2)
	x[,5]<-trun(umur,k2,2)
	beta<-solve(t(x)%*%x)%*%t(x)%*%y
	ytopi<-x%*%beta
	A<-x%*%solve(t(x)%*%x)%*%t(x)
	I<-matrix(0, ncol=n, nrow=n)
	for(i in 1:n)
		I[i, i]<-1
	ia<-diag(I)-A
	k<-((sum(diag(ia)))/n)^2
	msek<-(t(y-ytopi)%*%(y-ytopi))/n
	gcv<-msek/k
	residual<-y-ytopi
	ybar<-sum(y)/n
	sse<-t(y-ytopi)%*%(y-ytopi)
	sst<-t(y-ybar)%*%(y-ybar)
	ssr<-t(ytopi-ybar)%*%(ytopi-ybar)
	rsq<-ssr/sst
	dbreg <- 5
	MSreg <- ssr/dbreg
	dbres <- n-4
	MSres <- sse/dbres
	dbtot <- dbreg+dbres
	varian <- (diag(A) %*% MSres)
	q <- seq(min(umur), max(umur), length=1000)
```

```
f <- beta[1]+beta[2]*q+beta[3]*(q^2)+beta[4]*trun(q,k1,2)+
beta[5]*trun(q,k2,2)
var <- varian[1]+varian[2]*q+varian[3]*(q^2)+varian[4]*trun(q,k1,2)+
varian[5]*trun(q,k2,2)
akarvar <- sqrt(var)
uppr <- f+0.2*akarvar
lowr <- f-0.2*akarvar
i <- seq(min(umur),max(umur),length=n)
fest <- beta[1]+beta[2]*i+beta[3]*i^2+beta[4]*trun(i,k1,2)+
beta[5]*trun(i,k2,2)
upper <- fest+1.96*sqrt(varian)
lower <- fest-1.96*sqrt(varian)

graphsheet()
plot(umur,ASFR,type="p",col=2,xlim = c(min(umur),max(umur)), ylim =
c(min(ASFR),max(180)), xlab = "Umur", ylab = "ASFR")
par(new=T)
plot(q,f,type="l",col=8,xlim = c(min(q),max(q)), ylim =
c(min(ASFR),max(180)), xlab = "Umur",ylab = "ASFR")
par(new=T)
plot(q,uppr,type="l",xlim=c(min(q),max(q)), ylim=c(min(ASFR),max(180)), xlab
= "Umur", ylab="ASFR")
par(new=T)
plot(q,lowr,type="l",xlim=c(min(q),max(q)), ylim=c(min(ASFR),max(180)), xlab
= "Umur",ylab="ASFR")

return(upper,lower)

}

umur<-c(17,22,27,32,37,42,47)
ASFRK<-c(44,156,157,114,49,16,9)
ASFRB<-c(32,120,152,92,44,9,1)
ASFRD<-c(17,89,145,83,42,14,0)

CI.2.knot(umur,ASFR,k1,k2)
```

Lampiran 7

Program Gabungan Kurva

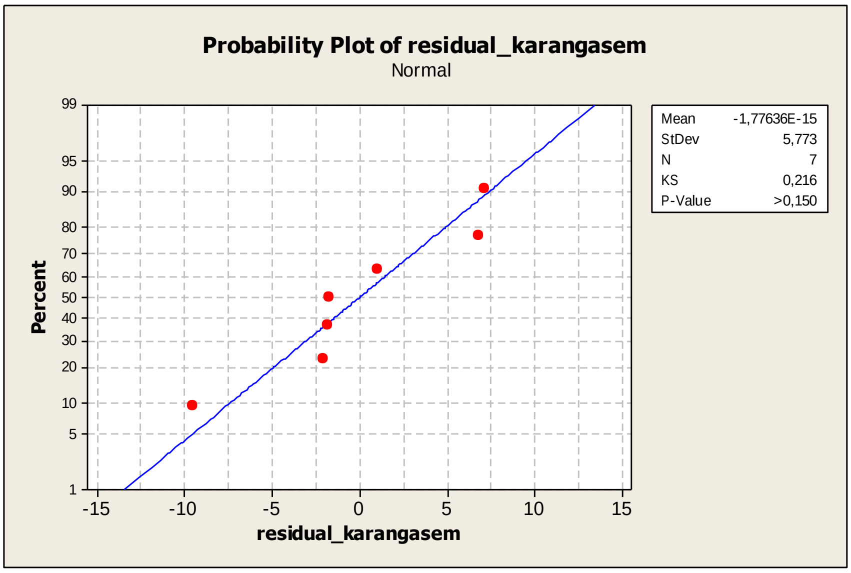
```
# KURVA ESTIMASI REGRESI SPLINE KUADRATIK GABUNGAN ASFR
KAB KARANGASEM, KOTA DENPASAR, PROV. BALI TAHUN 2013 #
# A = UMUR IBU #
# B = ASFR KARANGASEM #
# C = ASFR DENPASAR #
# D = ASFR BALI #
A <- c(17,22,27,32,37,42,47)
B <- c(44,156,157,114,49,16,9)
C <- c(17,89,145,83,42,14,0)
D <- c(32,120,152,92,44,9,1)
  k <- c(15:30)
  l <- c(30.1:49)
  m <- c(15:24)
  n <- c(24.1:27)
  o <- c(27.1:49)
  p <- c(15:31)
  q <- c(31.1:49)
K <- -945.29 + (87.45*k) - (1.71*k^2)
L <- 1007.71 - (42.75*l) + (0.46*l^2)
M <- 110.63 - (20.89*m) + (0.90*m^2)
N <- -3858.01 + (309.83*n) - (5.99*n^2)
O <- 734.69 - (30.37*o) + (0.31*o^2)
P <- -827.55 + (75.02*p) - (1.44*p^2)
Q <- 1017.57 - (44.02*q) + (0.48*q^2)

plot(A,B,type="p",col=8,xlim=c(15,49),ylim=c(0,250),xlab="Umur
Ibu",ylab="ASFR")
par (new=T)
plot(A,C,type="p",col=4,xlim=c(15,49),ylim=c(0,250),xlab="Umur
Ibu",ylab="ASFR")
par (new=T)
plot(A,D,type="p",col=6,xlim=c(15,49),ylim=c(0,250),xlab="Umur
Ibu",ylab="ASFR")
par (new=T)
plot(k,K,type="l",col=8,xlim=c(15,49),ylim=c(0,250),xlab="Umur
Ibu",ylab="ASFR")
par (new=T)
```

```
plot(l,L,type="l",col=8,xlim=c(15,49),ylim=c(0,250),xlab="Umur
Ibu",ylab="ASFR")
par (new=T)
plot(m,M,
type="l",col=4,xlim=c(15,49),xlim=c(15,49),ylim=c(0,250),xlab="Umur
Ibu",ylab="ASFR")
par (new=T)
plot(n,N,
type="l",col=4,xlim=c(15,49),xlim=c(15,49),ylim=c(0,250),xlab="Umur
Ibu",ylab="ASFR")
par (new=T)
plot(o,O,
type="l",col=4,xlim=c(15,49),xlim=c(15,49),ylim=c(0,250),xlab="Umur
Ibu",ylab="ASFR")
par (new=T)
plot(p,P, type="l",col=6,xlim=c(15,49),xlim=c(15,49),ylim=c(0,250),xlab="Umur
Ibu",ylab="ASFR")
par (new=T)
plot(q,Q,
type="l",col=6,xlim=c(15,49),xlim=c(15,49),ylim=c(0,250),xlab="Umur
Ibu",ylab="ASFR")
```

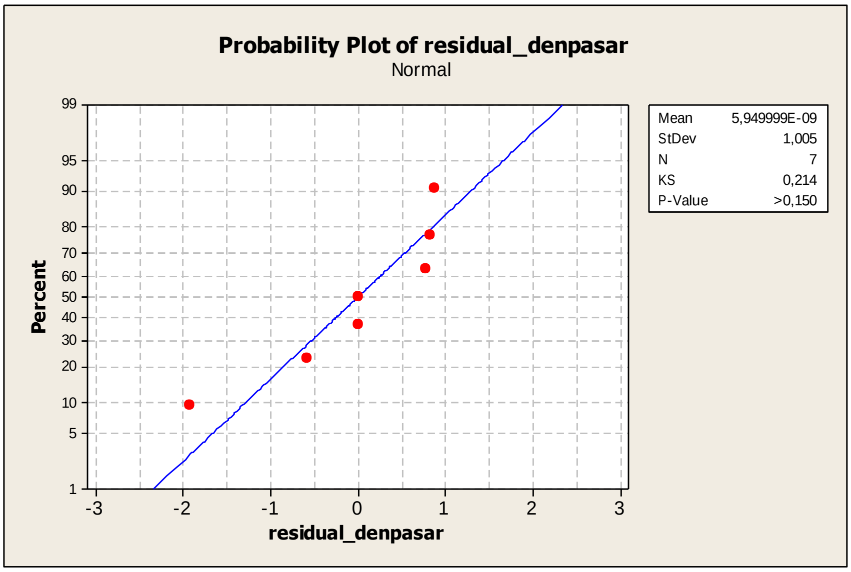
Lampiran 8

Plot Uji Kenormalan Residual data ASFR kab. Karangasem, kota Denpasar dan Propinsi Bali

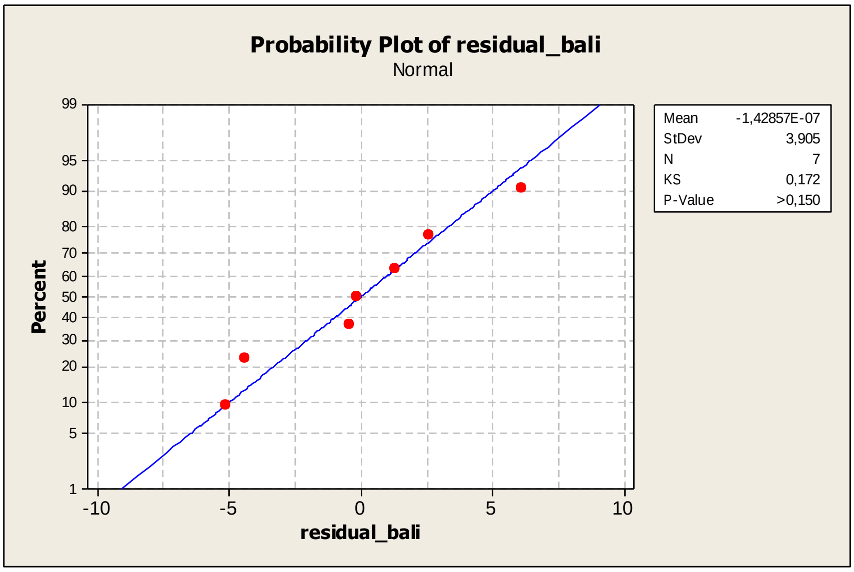
1. Kabupaten Karangasem



2. Kota Denpasar



3. Propinsi Bali



BIODATA PENULIS

Ida Ayu Sevita Intansari atau biasa dipanggil Vita atau Gek Vita lahir di Surabaya, 9 Juli 1990, merupakan anak pertama dari tiga bersaudara dari pasangan Ida Bagus Swela dan Bekti Sriyanti. Penulis menyelesaikan pendidik-an dasar di SDN Dukuh Kupang V Surabaya, sedangkan pendidikan menengah pertama dan pendidikan menengah atas di SMPN 6 Surabaya dan SMAN 6 Surabaya. Tahun 2008, Penulis melanjutkan pendidikan tinggi di Institut Teknologi Sepuluh Nopember Surabaya Jurusan Statistika yang diterima melalui jalur SNMPTN dan lulus Sarjana pada tahun 2012. Kemudian melanjutkan pendidikan Magister di Statistika ITS dan lulus pada tahun 2016.

Selama kuliah S2, penulis juga sempat bekerja sebagai *Research and Development Executive* di PT. Radio Baturiti Menaraswara (Hard Rock FM Bali). Dan pada pengerjaan Tugas Akhir ini Penulis mengambil bidang Sosial Pemerintahan dan diaplikasikan pada data ASFR di Bali. Untuk semua informasi, masukan maupun diskusi mengenai penelitian ini dapat menghubungi penulis melalui email idaayusevita@yahoo.com atau FB : IdaAyu Sevita Intansari (gek_dayoe_tha2@yahoo.com), LINE : idaayusevita.