SELANG KEPERCAYAAN KURVA REGRESI NONPARAMETRIK BERDASARKAN ESTIMATOR *SPLINE TRUNCATED* PADA DATA LONGITUDINAL

SKRIPSI

MYANITA ASTUTI

PROGRAM STUDI S-1 STATISTIKA

DEPARTEMEN MATEMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS AIRLANGGA

2018

SELANG KEPERCAYAAN KURVA REGRESI NONPARAMETRIK BERDASARKAN ESTIMATOR *SPLINE TRUNCATED* PADA DATA LONGITUDINAL

SKRIPSI

Sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Statistika pada Fakultas Sains dan Teknologi
Universitas Airlangga

Myanita Astuti
NIM. 081411831034

Disetujui oleh:

Pembimbing I,

Dr. Nur Chamidah, M.Si
NIP. 197206021998022001

Pembimbing II,

Drs. Eko Tjahjono, M. Si
NIP. 196007061986011001

LEMBAR PENGESAHAN NASKAH SKRIPSI

Judul	: Selang Kepercayaan Kurva Regresi Nonparametrik berdasarkan Estimator *Spline Truncated* pada Data Longitudinal
Penyusun	: Myanita Astuti
NIM	: 081411831034
Pembimbing I	: Dr. Nur Chamidah, M.Si
Pembimbing II	: Drs. Eko Tjahjono, M.Si
Tanggal Ujian	: 27 April 2018

Disetujui oleh :

Pembimbing I	Pembimbing II
Dr. Nur Chamidah, M.Si	Drs. Eko Tjahjono, M.Si
NIP. 197206021998022001	NIP. 196007061986011001

Mengetahui :
Ketua Departemen Matematika
Fakultas Sains dan Teknologi
Universitas Airlangga

Drs. Eko Tjahjono, M.Si
NIP. 196007061986011001

PEDOMAN PENGGUNAAN SKRIPSI

Skripsi ini tidak dipublikasikan, namun tersedia di perpustakaan dalam lingkungan Universitas Airlangga, diperkenankan untuk dipakai sebagai referensi kepustakaan, tetapi pengutipan harus seizin penyusun dan harus menyebutkan sumbernya sesuai kebiasaan ilmiah. Dokumen skripsi ini merupakan hak milik Universitas Airlangga.

SURAT PERNYATAAN TENTANG ORISINALITAS

Yang bertanda tangan di bawah ini, saya:

Nama	: Myanita Astuti
NIM	: 081411831034
Program Studi	: Statistika
Fakultas	: Sains dan Teknologi
Jenjang	: Sarjana (S1)

Menyatakan bahwa saya tidak melakukan kegiatan plagiat dan penulisan skripsi saya yang berjudul:

Selang Kepercayaan Kurva Regresi Nonparametrik
berdasarkan Estimator *Spline Truncated*
pada Data Longitudinal

Apabila suatu saat nanti terbukti melakukan tindakan plagiat, maka saya akan menerima sanksi yang telah ditetapkan.

Demikian surat pernyataan ini saya buat dengan sebenar-sebenarnya.

Surabaya, 4 Mei 2018

METERAI TEMPEL
51F7CACF297298438
6000 DJP

Myanita Astuti
NIM. 081411831034

KATA PENGANTAR

بسم الله الرحمن الرحيم

Assalamu'alaikum warahmatullahi wabarakatuh.

Alhamdulillahirabbil 'alamin puji syukur kehadirat Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan proposal yang berjudul "Selang Kepercayaan Kurva Regresi Nonparametrik Berdasarkan Estimator *Spline Truncated* pada Data Longitudinal". Dalam menyusun proposal skripsi ini penulis mendapat bantuan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Kedua Orang tua (Ibu Kastiani dan Bapak Moch. Yasin), serta keluarga besar penulis yang telah mendoakan dan memberikan semangat, kasih sayang, kepercayaan, dan dan dukungan baik secara materil maupun moril.
2. Dr. Nur Chamidah, M.Si selaku dosen pembimbing I dan Drs. Eko Tjahjono, M.Si selaku dosen pembimbing II yang dengan tulus telah memberikan bimbingan dan arahan yang berguna kepada penulis.
3. Dr. Ardi Kurniawan selaku dosen wali selama menjadi mahasiswa yang telah memberikan nasehat, arahan dan saran agar menyelesaikan studi serta seluruh dosen Statistika yang telah membimbing dan memberikan pelajaran berharga dari semester awal hingga saat ini.
4. Semua pihak yang tidak dapat penulis sebutkan satu persatu, terima kasih atas segala bantuannya dalam penyelesaian skripsi ini.

Penulis menyadari bahwa dalam penulisan skripsi ini masih banyak kekurangan, sehingga penulis mengharapkan kritik dan saran yang membangun dari semua pihak. Semoga proposal ini menjadi sesuatu yang berguna dan senantiasa dirahmati dan diridhoi Allah SWT. Aamiin.

Surabaya, Mei 2018

Penulis,

Myanita Astuti

Myanita Astuti, 2018. **Selang Kepercayaan Kurva Regresi Nonparametrik berdasarkan Estimator *Spline Truncated* pada Data Longitudinal.** Skripsi dibawah bimbingan Dr. Nur Chamidah, M.Si dan Drs. Eko Tjahjono, M.Si. Program Studi S1-Statistika, Departemen Matematika, Fakultas Sains dan Teknologi, Universitas Airlangga, Surabaya.

ABSTRAK

Analisis regresi merupakan salah satu metode yang digunakan dalam statistika untuk mengetahui hubungan antara variabel respon dengan prediktor. Pendekatan regresi nonparametrik bentuk kurva regresi diasumsikan tidak diketahui dan hanya diasumsikan *smooth* (halus). Estimator yang digunakan adalah estimator *spline truncated*. Dalam kehidupan sehari-hari sering dijumpai bentuk data longitudinal. Data longitudinal lebih kompleks dibandingkan data *cross section* karena data juga diamati berdasarkan waktu. Persoalan inferensi yang sangat penting dalam regresi adalah selang kepercayaan, karena dapat meminimalkan kesalahan dalam mengestimasi dibandingkan estimasi titik. Namun penelitian mengenai selang kepercayaan untuk regresi nonparametrik belum banyak dikembangkan, sehingga secara teori menarik untuk mengembangkan estimasi selang kepercayaan kurva regresi nonparametrik. Tujuan penelitian ini adalah mengkaji estimasi selang kepercayaan untuk kurva regresi nonparametrik pada data longitudinal dengan menggunakan estimator *spline truncated*. Dalam menyelesaikan permasalahan tersebut digunakan metode *Weighted Least Square* untuk menentukan estimasi titik dan *pivotal quantity* untuk mengkonstruksi selang kepercayaan pada kasus variansi populasi $\left(\sigma^2\right)$ tidak diketahui. Hasil kajian teoritis diperoleh estimasi selang kepercayaan dengan distribusi-t. Penerapan estimasi selang kepercayaan dilakukan pada data pertumbuhan balita usia 0–24 bulan di Surabaya dengan berat badan balita sebagai variabel respon, dan usia balita sebagai variabel prediktor. Hasil estimasi data pertumbuhan berat badan balita menggunakan *software* OSS-R GCV minimum dan R^2 untuk balita laki-laki dan perempuan masing-masing yaitu 0,3151; 92,43%; 0,2648; dan 94,66%.

Kata Kunci : Nonparametrik, Data Longitudinal, *Spline Truncated*, Selang Kepercayaan, *Pivotal Quantity*, Pertumbuhan Balita.

Myanita Astuti, 2018. **Confidence Interval of Nonparametric Regression Model Based on Spline Truncated Estimator on Longitudinal Data**. This final project is under supervised by Dr. Nur Chamidah, M.Si and Drs. Eko Tjahjono, M.Si. S1 Statistics Study Program, Mathematics Department, Faculty of Science and Technology, Airlangga University, Surabaya.

ABSTRACT

Regression analysis is one of the methods in statistics that is used to know the functional relationship between the response with the predictor variables. Non-parametric regression approach, its curve assumed to be unknown and only assumed to be smooth. One Nonparametric regression estimator which often be used is spline truncated. In daily life common to see longitudinal data, its more complicated than cross-section data because the data is also observed based on time. Inferencing issue which is important in regression is confidence interval, because it can minimize errors in estimating than point estimates. However researches about non-parametric regression's confidence interval are not much being developed yet. So, theoretically, this study is fascinating to develop the confidence interval estimation on non-parametric regression curve. This research is meant to study the estimation of confidence interval for non-paramertric regression curve on longitudinal data by using spline truncated estimation. In order to solve the problem Weighted Least Square and pivotal quantity method are used in this unknown population variance (σ^2). The theoretical result obtained that confidence interval estimation is using t-distribution. The implementation of confidence interval estimation is done to data of 0-24 year-old child growth in Surabaya with child's body weight as response variable and child's age as predictor variable. The estimation result of child's body weight growth was used by OSS-R software and obtained minimum GCV and R^2 are 0,3151 and 92,43% for boy, also 0,2648 and 94,66% for girl.

Keywords : Nonparametric, Longitudinal Data, Spline Truncated, Confidence Interval, Pivotal Quantity, Infants Growth.

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR LAMPIRAN

BAB I
PENDAHULUAN

1.1 Latar Belakang

Analisis regresi merupakan salah satu metode yang digunakan dalam statistika untuk mengetahui hubungan antara variabel respon dengan variabel prediktor. Berdasarkan perkembangan ilmu pengetahuan, pendekatan yang digunakan dalam analisis regresi dapat dibedakan menjadi dua jenis, yaitu regresi parametrik dan regresi nonparametrik. Pendekatan regresi parametrik dilakukan apabila bentuk kurva regresinya telah diketahui, diasumsikan error berdistribusi normal, dan variansi homogen. Kemudian pendekatan regresi nonparametrik digunakan jika bentuk kurva regresinya tidak diketahui. Dalam pendekatan regresi nonparametrik bentuk kurva regresi diasumsikan tidak diketahui dan hanya diasumsikan smooth (halus) atau termuat dalam dalam suatu ruang fungsi tertentu. Regresi nonparametrik merupakan regresi yang sangat fleksibel dalam memodelkan pola data (Eubank, 1988).

Beberapa contoh estimator regresi nonparametrik yang telah banyak digunakan antara lain kernel, *spline*, polinomial lokal, dan deret *fourier*. Pada analisis regresi nonparametrik yang menggunakan estimator *spline* mampu mengatasi karakter data atau fungsi yang bersifat mulus dan perilakunya berubah-ubah pada sub interval tertentu, *spline* juga merupakan model polinomial yang tersegmen. Penggunaan *spline truncated* difokuskan kepada adanya pola data yang memiliki karakteristik berbeda antara interval satu dengan interval lain (Mubarak dan Budiantara, 2012). Salah satu kelebihan estimator *spline* adalah estimasi data pada model ini cenderung mengikuti pergerakan pola data tersebut. Kelebihan ini terjadi karena dalam *spline* terdapat titik-titik knot, yaitu titik perpaduan bersama yang menunjukkan terjadinya perubahan pola perilaku data (Budiantara, 2009). Kemudian *spline truncated* memiliki kelebihan dalam mengatasi pola data yang menunjukkan naik turun yang tajam dengan bantuan titik-titik knot, serta kurva yang dihasilkan relatif mulus (Hardle, 1990). Dengan titik knot ini, estimator *spline truncated* dapat memberikan fleksibilitas yang lebih baik dan memungkinkan untuk menyesuaikan diri terhadap karateristik lokal.

Beberapa penelitian yang telah dilakukan sebelumnya menggunakan estimator *spline* antara lain Malik (2014) tentang estimasi kurva regresi nonparametrik untuk data longitudinal dengan pendekatan s*pline*. Kemudian Bintariningrum dan Budiantara (2014) tentang pemodelan regresi nonparametrik *spline truncated* dan aplikasinya pada data *cross sectional*. Selain itu, Fadhilah *et al*. (2015) juga meneliti tentang pemodelan regresi *spline truncated* pada data longitudinal. Pendekatan *spline truncated* tidak hanya mampu mengestimasi kurva regresi nonparametrik untuk data *cross section* saja, tetapi juga dapat dikembangkan untuk data longitudinal. Data *cross sectional* merupakan jenis data yang pada umumnya digunakan oleh peneliti-peneliti dikarenakan hanya mengobservasi suatu kejadian pada titik tertentu sehingga lebih sederhana dan mudah. Namun, pada kehidupan sehari-hari sering dijumpai pula data longitudinal. Data longitudinal lebih kompleks dibandingkan data *cross section* karena data juga diamati berdasarkan waktu. Terdapat beberapa kelebihan data longitudinal dibandingkan data *cross section*, diantaranya adalah dapat mengetahui perubahan individu dan tidak membutuhkan subjek yang terlalu banyak karena pengamatan dilakukan secara berulang serta hasil estimasi yang diperoleh lebih efisien karena dilakukan disetiap pengamatan (Wu dan Zhang, 2006).

Berdasarkan beberapa penelitian tersebut, masih terbatas sampai estimasi titiknya dan belum membahas tentang estimasi selang kepercayaan, padahal salah satu persoalan pada inferensi statistik yang sangat penting yaitu selang kepercayaan. Selang kepercayaan merupakan rentang perkiraan nilai diantara dua angka, di mana nilai parameter sebuah populasi terletak di dalam interval tersebut. Estimasi selang kepercayaan ini merupakan pengembangan dari estimasi titik, bahwa nilai taksiran parameter tidak terfokus pada suatu titik tetapi didasarkan pada selang tertentu sehingga dapat meminimalkan kesalahan dalam mengestimasi dibandingkan estimasi titik. Selang kepercayaan untuk kurva dalam regresi nonparametrik dapat digunakan untuk menentukan rentang perkiraan nilai $\hat{f}(\mathrm{x})$.

Wang (1998) telah terlebih dahulu menggunakan pendekatan Bayesian dalam merancang interval konfidensi untuk kurva regresi. Pendekatan Bayesian ini memerlukan pemahaman tentang distribusi *prior improper*, fungsi risiko Bayes, distribusi *posterior*, proses stokastik *Winner*, mean *posterior* dan lainnya, yang umumnya kurang dimengerti oleh pengguna Statistika (Syaranamual dan Budiantara, 2011). Penelitian mengenai selang kepercayaan juga telah dilakukan oleh Nafi' (2010) yaitu interval konfidensi untuk kurva regresi nonparametrik *spline* menggunakan model *spline* kubik. Kemudian Syaranamual dan Budiantara (2011) yang mengkonstruksi interval konfidensi *spline* kuadrat dengan pendekatan *pivotal quantity* dan Intansari (2016) meneliti tentang estimasi titik dan interval konfidensi dalam kurva regresi nonparametrik menggunakan s*pline* kuadratik. Kemudian Setiawan (2017) juga telah melakukan penelitian tentang interval konfidensi untuk parameter model regresi nonparametrik *spline truncated* pada data indeks pembangunan gender. Namun dari penelitian-penelitian tersebut, pengaplikasian selang kepercayaan masih terbatas pada data *cross sectional* dan belum diterapkan pada data longitudinal.

Pada penerapannya ke dalam permasalahan baik itu data riil atau bangkitan, proses dalam perhitungannya tidak dapat dilakukan secara manual. Perhitungan harus memerlukan algoritma dan pemrograman dengan bantuan *software. Software* dengan bahasa pemrograman yang umumnya digunakan yaitu S-PLUS dan OSS-R. Keduanya memiliki cara penggunaan yang hampir sama. Namun OSS-R memiliki keunggalan yaitu merupakan *Open Source Software* atau software yang tidak berbayar/gratis. *Software* R dikembangkan oleh AT&T Bell Laboratories (sekarang Lucent Technologies) pada akhir tahun 1970 an.

Berdasarkan uraian yang telah dijelaskan, maka dalam skripsi ini akan mengkonstruksi selang kepercayaan kurva regresi nonparametrik menggunakan estimator *spline truncated* pada data longitudinal. Metode yang kerjakan pada skripsi ini akan diterapkan pada data riil atau bangkitan. Kemudian *Software* yang akan digunakan dalam skripsi ini yaitu *Open Source Software-R* (OSS-R), sehingga dapat lebih diketahui hasil *output* dari skripsi ini.

1.2 Rumusan Masalah

Berdasarkan uraian dari latar belakang, permasalahan yang perlu ditelaah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana mengkonstruksi selang kepercayaan kurva regresi nonparametrik berdasarkan estimator *spline truncated* pada data longitudinal?
2. Bagaimana membuat algoritma dan program untuk mengkonstruksi selang kepercayaan kurva regresi nonparametrik berdasarkan estimator *spline truncated* pada data longitudinal dan penerapannya pada data riil atau bangkitan?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan, maka tujuan yang ingin dicapai dalam penelitian adalah sebagai berikut:

1. Mengkonstruksi selang kepercayaan kurva regresi nonparametrik berdasarkan estimator *spline truncated* pada data longitudinal.
2. Membuat algoritma dan program untuk mengkonstruksi selang kepercayaan kurva regresi nonparametrik berdasarkan estimator *spline truncated* pada data longitudinal dan penerapannya pada data riil atau bangkitan.

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Menambah wawasan dan pengetahuan mahasiswa tentang selang kepercayaan kurva regresi nonparametrik berdasarkan estimator *spline truncated* pada data longitudinal.
2. Menambah wawasan bagi mahasiswa mengenai algoritma dan program yang diterapkan pada *software* OSS-R.

1.5 Batasan Masalah

Batasan yang digunakan pada penelitian ini yaitu menggunakan *pivotal quantity* untuk mendapatkan selang kepercayaan kurva regresi nonparametrik berdasarkan estimator *spline truncated.*

BAB II
TINJAUAN PUSTAKA

2.1 Teorema dan Definisi Dasar Aljabar Matriks

Beberapa teorema dan definisi dasar terkait aljabar matriks yang digunakan untuk menyelesaikan kajian selang kepercayaan kurva regresi nonparametrik berdasarkan estimator *spline truncated* pada data longitudinal berdasarkan Rencher dan Schaalje (2008) adalah sebagai berikut:

Definisi 2.1 Matriks

Matriks adalah susunan bilangan atau variabel dalam bentuk persegi panjang atau persegi. Bilangan-bilangan dalam susunan tersebut dinamakan entri matriks. Ukuran matriks dijelaskan dengan menyatakan banyaknya baris dan banyaknya kolom yang terdapat dalam matriks tersebut, sehingga suatu matriks dengan m baris dan n kolom dikatakan sebagai matriks dengan ukuran (*ordo*) $m \times n$. Bentuk umum matriks yang berukuran $m \times n$ adalah sebagai berikut.

$$\mathbf{A}_{mn} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Tiap-tiap bilangan a_{ij} yang berada di dalam matriks A disebut elemen. Indeks i dan j masing-masing menyatakan baris dan kolom tempat beradanya sebuah elemen dari matriks **A**. Vektor adalah matriks dengan baris tunggal atau kolom tunggal.

Definisi 2.2 Penjumlahan Dua Matriks atau Dua Vektor

Jika dua matriks atau dua vektor berukuran sama, maka dikatakan *conformal* untuk penjumlahan. Jika matriks **A** berukuran $n \times p$ dan matriks **B** berukuran $n \times p$, maka $\mathbf{C} = \mathbf{A} + \mathbf{B}$ juga berukuran $n \times p$ dan ditentukan $\mathbf{C} = \left(c_{ij}\right) = \left(a_{ij} + b_{ij}\right)$.

Definisi 2.3 Perkalian Dua Matriks atau Dua Vektor

Agar perkalian **AB** terdefinisi maka jumlah kolom dari **A** harus sama dengan jumlah baris dari **B**, dalam hal ini **A** dan **B** dikatakan *conformal* untuk perkalian. Kemudian elemen ke *ij* dari perkalian **C=AB** didefinisikan sebagai

$$c_{ij} = \sum_k a_{ik} b_{kj}$$

merupakan penjumlahan perkalian dari elemen-elemen dari baris ke i dari **A** dan elemen-elemen dari kolom j dari **B**. Jika **A** berukuran $n \times m$ dan **B** berukuran $m \times p$, maka **C** = **AB** berukuran $n \times p$.

Teorema 2.1

Apabila **A** suatu matriks berukuran $n \times p$ dan **B** berukuran $p \times n$, maka

$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$$

Definisi 2.4 Invers

Matriks persegi *full-rank* merupakan matriks *nonsingular*. Matriks *nonsingular* **A** mempunyai *invers* yang tunggal dan dinotasikan dengan $\mathbf{A}^{-1}$, berlaku

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

Dari definisi di atas, jelas bahwa **A** adalah *invers* dari $\mathbf{A}^{-1}$:

$$\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$$

Definisi 2.5 Matriks Idempoten

Sebuah matriks persegi **A** dikatakan idempoten jika $\mathbf{A}^2 = \mathbf{A}$.

Teorema 2.2

Jika **A** adalah matriks simetri dan idempoten dengan *rank r*, maka

$$rank(\mathbf{A}) = tr(\mathbf{A}) = r$$

Teorema 2.3

Jika **A** matriks berukuran $n \times p$ dan **B** matriks berukuran $p \times n$, maka

$$tr(\mathbf{AB}) = tr(\mathbf{BA})$$

Teorema 2.4

Jika **A** adalah mariks nonsingular, maka $\mathbf{A}^T$ nonsingular dan inversnya adalah

$$\left(\mathbf{A}^T\right)^{-1} = \left(\mathbf{A}^{-1}\right)^T$$

Teorema 2.5

Misalkan $u = \underset{\sim}{a}^T \underset{\sim}{x} = \underset{\sim}{x}^T \underset{\sim}{a}$, dengan $\underset{\sim}{a}^T = \left(a_1, a_2, \ldots, a_p\right)$ adalah vektor konstanta, dengan ini berlaku:

$$\frac{\partial u}{\partial \underset{\sim}{x}} = \frac{\partial\left(\underset{\sim}{a}^T \underset{\sim}{x}\right)}{\partial \underset{\sim}{x}} = \frac{\partial\left(\underset{\sim}{x}^T \underset{\sim}{a}\right)}{\partial \underset{\sim}{x}} = \underset{\sim}{a}$$

Teorema 2.6

Apabila $u = \underset{\sim}{x}^T \mathbf{A} \underset{\sim}{x}$, dengan **A** merupakan matriks simetri dari konstanta, maka

$$\frac{\partial u}{\partial \underset{\sim}{x}} = \frac{\partial\left(\underset{\sim}{x}^T \mathbf{A} \underset{\sim}{x}\right)}{\partial\left(\underset{\sim}{x}\right)} = 2\mathbf{A}\underset{\sim}{x}$$

Teorema 2.7

Diberikan **y** adalah vektor random, **X** adalah matriks random, **A** adalah matriks konstanta. Jika diasumsikan matriks dan vektor dalam setiap perkalian adalah *conformable*, maka berlaku

$$E\left(\mathbf{A}\underset{\sim}{y}\right) = \mathbf{A}E\left(\underset{\sim}{y}\right)$$

Teorema 2.8

Misalkan $\underset{\sim}{z} = \mathbf{A}\underset{\sim}{y}$, dengan **A** matriks konstanta berukuran $k \times p$ dan $\underset{\sim}{y}$ vektor random berdimensi $p \times 1$ dengan matriks covarians $\sum$, maka

$$\operatorname{cov}(\underset{\sim}{z}) = \operatorname{cov}\left(\mathbf{A}\underset{\sim}{y}\right) = \mathbf{A} \sum \mathbf{A}^T$$

Teorema 2.9

Diberikan vektor random $\underset{\sim}{y}$ berukuran $p \times 1$ yang berdistribusi $N_p\left(\mu, \mathbf{\Sigma}\right)$, misalkan diambil vektor konstanta $\underset{\sim}{a}$ berdimensi $p\backslash \times 1$ dan **A** matriks konstanta berukuran $k \times p$ dengan *rank* $k \leq p$, maka distribusi kombinasi linier dari $\underset{\sim}{a}^T \underset{\sim}{y}$ adalah $\underset{\sim}{a}^T \underset{\sim}{y} \sim N\left(\underset{\sim}{a}^T \underset{\sim}{\mu}, \underset{\sim}{a}^T \mathbf{\Sigma} \underset{\sim}{a}\right)$ serta distribusi kombinasi linier $\mathbf{A}\underset{\sim}{y} \sim N\left(\mathbf{A}\underset{\sim}{\mu}, \mathbf{A\Sigma A}^T\right)$.

Definisi 2.6 Distribusi *t* Sentral

Jika $z \sim N(0,1)$, $u \sim \chi^2_{(p)}$ dan *z* dan *u* independen maka

$$t = \frac{z}{\sqrt{u/p}} \sim t_{(p)}$$

berdistribusi *t* sentral dengan derajat bebas *p*.

Definisi 2.7 Distribusi *t* Non Sentral

Jika $y \sim N(\mu,1)$, $u \sim \chi^2_{(p)}$ dan *y* dan *u* independen maka

$$t = \frac{y}{\sqrt{u/p}} \sim t_{(p,\mu)}$$

berdistribusi *t* non sentral dengan derajat bebas *p* dan parameter nonsentral μ.

Definisi 2.8 Distribusi *t* Non Sentral

Jika $y \sim N(\mu,\sigma^2)$, $u \sim \chi^2_{(p)}$ dan *y* dan *u* independen maka

$$t = \frac{y/\sigma}{\sqrt{u/p}} \sim t_{\left(p,\mu/\sigma\right)}$$

y/σ berdistribusi $N\left(\mu/\sigma,1\right)$.

Teorema 2.10

Misalkan $\underset{\sim}{y}$ berdistribusi $N_p(\mu,\mathbf{\Sigma})$, $\mathbf{A}$ adalah matriks konstanta yang simetris dengan *rank r*, dan $\lambda=\frac{1}{2}\underset{\sim}{\mu}^T\mathbf{A}\underset{\sim}{\mu}$, maka bentuk kuadratik $\underset{\sim}{y}^T\mathbf{A}\underset{\sim}{y}$ berdistribusi $\chi^2_{(r,\lambda)}$ jika dan hanya jika $\mathbf{A\Sigma}$ idempoten.

Akibat 1.

Jika $\underset{\sim}{y}$ berdistribusi $N_p(\underset{\sim}{0},\mathbf{I})$ maka $\underset{\sim}{\mu}^T\mathbf{A}\underset{\sim}{\mu}$ berdistribusi $\chi^2_{(r)}$ jika dan hanya jika **A** idempoten dengan rank *r*.

Akibat 2.

Jika $\underset{\sim}{y}$ berdistribusi $N_p\left(\underset{\sim}{\mu},\sigma^2\mathbf{I}\right)$, maka $\frac{\underset{\sim}{\mu}^T\mathbf{A}\underset{\sim}{\mu}}{\sigma^2}$ berdistribusi $\chi^2_{\left(r,\frac{\underset{\sim}{\mu}^T\mathbf{A}\underset{\sim}{\mu}}{2\sigma^2}\right)}$ jika dan hanya jika **A** idempoten dengan rank *r*.

Teorema 2.11

Misalkan **B** matriks konstanta berukuran $k \times p$, **A** adalah matriks konstanta yang simetris, dan $\underset{\sim}{y}$ berdistribusi $N_p\left(\mu, \mathbf{\Sigma}\right)$. Maka **By** dan $\underset{\sim}{y}^T \mathbf{A} \underset{\sim}{y}$ independen jika dan hanya jika $\mathbf{B\Sigma A} = \mathbf{O}$.

Akibat 1.

Jika $\underset{\sim}{y}$ berdistribusi $N_p\left(\underset{\sim}{\mu}, \sigma^2 \mathbf{I}\right)$,maka **By** dan $\underset{\sim}{y}^T \mathbf{A} \underset{\sim}{y}$ independen jika dan hanya jika $\mathbf{BA} = \mathbf{O}$.

2.2 Analisis Regresi

Istilah regresi pertama kali digunakan oleh Francis Galton tahun 1877 yang saat itu sedang melakukan studi bahwa tinggi badan anak-anak yang dilahirkan dari orang tua yang tinggi cenderung bergerak (*regress*) kearah ketinggian rata-rata populasi secara keseluruhan. Kemudian Galton memperkenalkan kata regresi (*regression*) sebagai nama proses umum untuk memprediksi satu variabel anak dengan menggunakan variabel lain. Menurut Gujarati dan Porter (2009) analisis regresi merupakan kajian terhadap ketergantungan satu variabel yaitu variabel dependen dengan satu atau lebih variabel lain yaitu variabel independen dengan tujuan untuk membuat estimasi atau memprediksi. Berikut adalah model persamaan regresi :

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + \varepsilon_i; \quad i = 1, 2, \ldots, n \tag{2.1}$$

dengan y merupakan variabel dependen sedangkan x merupakan variabel independen. Persamaan (2.1) dapat dinyatakan dalam bentuk matriks sesuai persamaan (2.2).

$$\underset{\sim}{y} = \mathbf{X}\underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \tag{2.2}$$

2.3 Regresi Nonparametrik

Regresi nonparametrik merupakan metode statistika yang digunakan untuk mengetahui hubungan antara variabel respon dan variabel prediktor, namun tidak diketahui bentuk kurvanya. Menurut Hardle (1994), pendekatan regresi nonparametrik untuk mengestimasi kurva regresi memiliki beberapa tujuan utama. Diantaranya yaitu menyediakan metode yang dapat digunakan dalam berbagai kondisi untuk eksplorasi hubungan antara dua variabel. Memberikan prediksi pengamatan yang belum dapat dibuat tanpa adanya referensi untuk model parametrik tertentu. Dalam model regresi nonparametrik bentuk kurva regresi diasumsikan tidak diketahui. Kurva regresi hanya diasumsikan halus (*smooth*) dalam arti termuat di dalam suatu ruang fungsi tertentu. Misalkan x_i adalah variabel prediktior dan y_i adalah variabel respon dengan n buah subyek, maka bentuk umum dari regresi nonparametrik dapat ditulis sebagai berikut:

$$y_i = f(x_i) + \varepsilon_i, \qquad i = 1, 2, ..., n \tag{2.3}$$

y_i = variabel respon ke-i

$f(x_i)$ = fungsi regresi nonparametrik

ε_i = Error, iid dengan mean 0 dan variansi σ^2

2.4 Regresi Nonparametrik *Spline Truncated*

Beberapa model pendekatan regresi nonparametrik yang telah dikembangkan oleh para peneliti, salah satunya adalah spline. Spline merupakan salah satu teknik estimasi regresi nonparametrik yang pertama kali dikembangkan oleh Whittaker pada tahun 1923 (Hardle, 1990). Spline dalam regresi nonparametrik mempunyai kemampuan mengestimasi perilaku data yang cenderung berbeda pada interval yang berlainan (Eubank, 1988). Suatu basis untuk ruang *spline* berorde p dapat dinyatakan dalam bentuk (Budiantara, 2001):

$$\{1,\ x,\ x^2,\ \dots,\ x^p,\ \left(x - \tau_1\right)^p,\ \dots, \left(x - \tau_k\right)^p\}$$

dengan

$$(x-\tau_r)_+^p = \begin{cases} (x-\tau_r)_+^p & , x \geq \tau_r \\ 0 & , x < \tau_r \end{cases}$$

τ_1, τ_2, ..., τ_k merupakan titik-titik knot. Titik knot adalah titik perpaduan bersama yang memperlihatkan terjadinya perubahan pola perilaku dari fungsi *spline* pada interval-interval yang berbeda. Secara umum fungsi spline berorde p adalah sembarang fungsi yang dapat ditulis dalam bentuk (Eubank, 1988):

$$f(x_i) = \beta_0 + \sum_{d=1}^{p} \beta_d x_i^d + \sum_{r=1}^{k} \beta_{(p+r)} (x_i - \tau_r)_+^p \qquad (2.4)$$

Dengan β adalah parameter dari fungsi *spline*, $d = 1, ..., p, p+1, ..., p+k$. Model regresi spline dapat disajikan sebagaimana persamaan (2.4).

$$y_i = \beta_0 + \sum_{d=1}^{p} \beta_d x_i^d + \sum_{r=1}^{k} \beta_{(p+r)} (x_i - \tau_r)_+^p + \varepsilon_i \qquad (2.5)$$

Apabila ε_i diasumsikan berdistribusi normal independen dengan mean nol dan variansi σ^2, maka y_i juga berdistribusi normal dengan mean $f(x_i)$ dan variansi σ^2.

2.5 Data Longitudinal

Dalam analisis regresi, data yang sering digunakan adalah data *cross section*. Data *cross section* mengacu pada data yang dikumpulkan dengan mengamati sejumlah subjek seperti individu ataupun wilayah pada titik waktu yang sama, dengan kata lain tidak memperhatikan perbedaan waktu. Selain data *cross section* terdapat pula data longitudinal. Data longitudinal merupakan data *cross section* yang memperhatikan adanya perbedaan waktu. Data tersebut akan mengikuti perubahan suatu subjek dari waktu ke waktu. Penggunaan data longitudinal telah banyak dilakukan dalam beberapa waktu terakhir. Hal ini dikarenakan pada data longitudinal mempunyai beberapa aplikasi praktis dalam berbagai bidang seperti kedokteran, kesehatan, ekonomi, dan lain-lain.

Data longitudinal merupakan data yang diperoleh dari pengukuran atau pengamatan yang dilakukan sebanyak n subyek yang saling independen dengan setiap subyek diamati secara berulang dalam kurun waktu berbeda yang saling dependen (Wu dan Zhang, 2006). Sedangkan Diggle *et al.* (2002) mendefinisikan karakter dari suatu studi longitudinal yaitu bahwa suatu subjek individu diukur berulang kali dalam kurun waktu tertentu. Penggunaan data longitudinal memberikan beberapa keunggulan diantaranya mengurangi kolinieritas antar variabel sehingga dapat menghasilkan estimasi yang efisien.

2.6 Regresi Nonparametrik *Spline Truncated* untuk Data Longitudinal

Pada data longitudinal apabila terdapat i=1, 2, …, n subyek dan j=1, 2, …, m pengamatan dalam setiap subyek dengan satu variabel prediktor t maka fungsi *spline truncated* dapat didefinisikan sebagai fungsi *f(x)* dengan *p* orde dan *k* knot. Model regresi nonparametrik *spline truncated* pada data longitudinal yang dirumuskan adalah sebagai berikut (Sriliana, 2012).

$$\begin{aligned} y_{ij} &= \beta_0 + \sum_{d=1}^{p} \beta_d t_{ij}^d + \sum_{r=1}^{k} \beta_{(r+p)} (t_{ij} - \tau_r)_+^p + \varepsilon_{ij} \\ y_{ij} &= \beta_0 + \beta_1 t_{ij}^1 + \beta_2 t_{ij}^2 + \ldots + \beta_p t_{ij}^p + \beta_{(p+1)} (t_{ij} - \tau_1)_+^p + \beta_{(p+2)} (t_{ij} - \tau_2)_+^p \\ &\quad + \ldots + \beta_{(p+k)} (t_{ij} - \tau_k)_+^p + \varepsilon_{ij} \end{aligned} \tag{2.6}$$

dengan

$$(t_{ij} - \tau_r)_+^p = \begin{cases} (t_{ij} - \tau_r)_+^p, \mathrm{t}_{ij} \geq \tau_r \\ 0 \qquad\qquad , \mathrm{t}_{ij} < \tau_r \end{cases}$$

Persamaan (2.6) dapat dinyatakan dalam bentuk matriks yang dapat dikonstruksi dari vektor dan matriks berikut:

$$\underset{\sim}{y} = \mathbf{T}(\boldsymbol{\tau})\underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \tag{2.7}$$

$$\begin{bmatrix} \underset{\sim}{y}_1 \\ \underset{\sim}{y}_2 \\ \vdots \\ \underset{\sim}{y}_n \end{bmatrix} = \begin{bmatrix} \mathbf{T_1(\boldsymbol{\tau})} & 0 & \cdots & 0 \\ 0 & \mathbf{T_2(\boldsymbol{\tau})} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{T_n(\boldsymbol{\tau})} \end{bmatrix} \begin{bmatrix} \underset{\sim}{\beta}_1 \\ \underset{\sim}{\beta}_2 \\ \vdots \\ \underset{\sim}{\beta}_n \end{bmatrix} + \begin{bmatrix} \underset{\sim}{\varepsilon}_1 \\ \underset{\sim}{\varepsilon}_2 \\ \vdots \\ \underset{\sim}{\varepsilon}_n \end{bmatrix}$$

dengan vektor respon:

$$\begin{aligned} \underset{\sim}{y}_1 &= \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1m_1} \end{pmatrix}^T \\ \underset{\sim}{y}_2 &= \begin{pmatrix} y_{21} & y_{22} & \cdots & y_{2m_2} \end{pmatrix}^T \\ &\vdots \\ \underset{\sim}{y}_n &= \begin{pmatrix} y_{n1} & y_{n2} & \cdots & y_{nm_n} \end{pmatrix}^T \end{aligned}$$

matriks basis *spline truncated*:

$$\mathbf{T_1(\tau)} = \begin{bmatrix} 1 & t_{11} & t_{11}^2 & \cdots & t_{11}^p & (t_{11}-\tau_1)_+^p & (t_{11}-\tau_2)_+^p & \cdots & (t_{11}-\tau_k)_+^p \\ 1 & t_{12} & t_{12}^2 & \cdots & t_{12}^p & (t_{12}-\tau_1)_+^p & (t_{12}-\tau_2)_+^p & \cdots & (t_{12}-\tau_k)_+^p \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_{1m_1} & t_{1m_1}^2 & \cdots & t_{1m_1}^p & (t_{1m_1}-\tau_1)_+^p & (t_{1m_1}-\tau_2)_+^p & \cdots & (t_{1m_1}-\tau_k)_+^p \end{bmatrix}$$

$$\mathbf{T_2(\tau)} = \begin{bmatrix} 1 & t_{21} & t_{21}^2 & \cdots & t_{21}^p & (t_{21}-\tau_1)_+^p & (t_{21}-\tau_2)_+^p & \cdots & (t_{21}-\tau_k)_+^p \\ 1 & t_{22} & t_{22}^2 & \cdots & t_{22}^p & (t_{22}-\tau_1)_+^p & (t_{22}-\tau_2)_+^p & \cdots & (t_{22}-\tau_k)_+^p \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_{2m_2} & t_{2m_2}^2 & \cdots & t_{2m_2}^p & (t_{2m_2}-\tau_1)_+^p & (t_{2m_2}-\tau_2)_+^p & \cdots & (t_{2m_2}-\tau_k)_+^p \end{bmatrix}$$

$\vdots$

$$\mathbf{T_n(\tau)} = \begin{bmatrix} 1 & t_{n1} & t_{n1}^2 & \cdots & t_{n1}^p & (t_{n1}-\tau_1)_+^p & (t_{n1}-\tau_2)_+^p & \cdots & (t_{n1}-\tau_k)_+^p \\ 1 & t_{n2} & t_{n2}^2 & \cdots & t_{n2}^p & (t_{n2}-\tau_1)_+^p & (t_{n2}-\tau_2)_+^p & \cdots & (t_{n2}-\tau_k)_+^p \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_{nm_n} & t_{nm_n}^2 & \cdots & t_{nm_n}^p & (t_{nm_n}-\tau_1)_+^p & (t_{nm_n}-\tau_2)_+^p & \cdots & (t_{nm_n}-\tau_k)_+^p \end{bmatrix}$$

dengan vektor parameter,

$$\underset{\sim}{\beta}_1 = \begin{pmatrix} \beta_{01} & \beta_{11} & \beta_{21} & \cdots & \beta_{p1} & \beta_{(p+1)1} & \beta_{(p+2)1} & \cdots & \beta_{(p+k)1} \end{pmatrix}^T$$

$$\underset{\sim}{\beta}_2 = \begin{pmatrix} \beta_{02} & \beta_{12} & \beta_{22} & \cdots & \beta_{p2} & \beta_{(p+1)2} & \beta_{(p+2)2} & \cdots & \beta_{(p+k)2} \end{pmatrix}^T$$

$\vdots$

$$\underset{\sim}{\beta}_n = \begin{pmatrix} \beta_{0n} & \beta_{1n} & \beta_{2n} & \cdots & \beta_{pn} & \beta_{(p+1)n} & \beta_{(p+2)n} & \cdots & \beta_{(p+k)n} \end{pmatrix}^T$$

dan vektor error,

$$\underset{\sim}{\varepsilon}_1 = \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1m_1} \end{pmatrix}^T$$

$$\underset{\sim}{\varepsilon}_2 = \begin{pmatrix} \varepsilon_{21} & \varepsilon_{22} & \cdots & \varepsilon_{2m_2} \end{pmatrix}^T$$

$\vdots$

$$\underset{\sim}{\varepsilon}_n = \begin{pmatrix} \varepsilon_{n1} & \varepsilon_{n2} & \cdots & \varepsilon_{nm_n} \end{pmatrix}^T$$

Estimasi parameter pada persamaan regresi nonparametrik spline *truncated* pada data longitudinal dilakukan dengan menggunakan metode *least square*. Perbedaan metode *least square* antara data *cross section* dan data longitudinal adalah adanya bobot (*weight*). Pada data *cross section*, estimasi paramater dapat diperoleh dengan menggunakan *Ordinary Least Square* (OLS). Pada data longitudinal, estimasi parameter diperoleh menggunakan *Weighted Least Square* (WLS) untuk mengatasi korelasi dalam subyek pengamatan yang sama.

2.7 Weighted Least Square

Pada metode (*Ordinary Least Square)* OLS *error* diasumsikan identik atau homogenitas dalam variansi residual. Metode *Weighted Least Square* (WLS) dapat digunakan ketika asumsi OLS varians konstan dalam *error* tidak dipenuhi atau disebut dengan heteroskedastisitas (Greene, 2003). **W** adalah matriks pembobot (matriks varians kovarians) berukuran *nm* x *nm* dan berisi diagonal $(\mathbf{W}_1, \mathbf{W}_2, ..., \mathbf{W}_n)$. Wu dan Zhang (2006) menyatakan ada tiga metode dalam menentukan matriks pembobot, yaitu:

1. $\mathbf{W}_i = N^{-1}\mathbf{I}$ bobot imi memberikan perlakuan yang sama pada setiap pengamatan.
2. $\mathbf{W}_i = n^{-1}\mathbf{I}$ bobot ini emberikan perlakuan yang sama pada setiap pengamatan dalam subyek.
3. $\mathbf{W}_i = \text{cov}(\underset{\sim}{y}_i)$ bobot ini memperhitungkan korelasi dalam subyek pengamatan.

Jika didefinisikan invers dari matriks variansi kovariansi sebagai pembobot untuk estimasi parameter yaitu $\omega_i = \frac{1}{\sigma_i^2}$, maka matriks $\mathbf{W}^{-1}$ merupakan suatu matriks diagonal yang berisi pembobot untuk estimasi parameter sebagai berikut.

$$\mathbf{W}^{-1} = \begin{bmatrix} \omega_1 & 0 & \cdots & 0 \\ 0 & \omega_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \omega_n \end{bmatrix}$$

Pada metode WLS fungsi yang diminimumkan untuk mengestimasi parameter dirumuskan sebagai berikut.

$$Q = \left(\underset{\sim}{y} - \mathbf{X}\underset{\sim}{\beta}\right)^T \mathbf{W}^{-1}\left(\underset{\sim}{y} - \mathbf{X}\underset{\sim}{\beta}\right) \tag{2.8}$$

dengan mendiferensiasi persamaan (2.8) terhadap $\underset{\sim}{\beta}$ diperoleh estimator WLS sebagai berikut.

$$\underset{\sim}{\hat{\beta}} = \left(\mathbf{X}^T \mathbf{W}^{-1} \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{W}^{-1} \underset{\sim}{\mathrm{y}} \tag{2.9}$$

(Rencher dan Schaalje, 2008).

2.8 Pemilihan Titik Knot Optimal

Spline merupakan potongan polinomial yang memuat titik-titik knot. Titik-titik knot merupakan titik perpaduan bersama dimana terjadi perubahan pola perilaku fungsi. Oleh karena itu letak dan banyaknya titik knot merupakan hal penting dalam pemodelan regresi nonparametrik dengan pendekatan *spline truncated*. Salah satu metode yang digunakan untuk menentukan knot yang optimal adalah dengan metode GCV *(Generalized Cross Validation)* (Wahba, 1990). GCV merupakan pengembangan dari metode CV (*Cross Validation*), perbedaannya terletak pada faktor-faktor yang membagi residual. Pada metode GCV faktor merupakan nilai rata-rata dari faktor-faktor tersebut. Model *spline* terbaik diperoleh dari nilai GCV minimum. Fungsi GCV untuk model regresi nonparametrik *spline truncated* data longitudinal yaitu:

$$GCV(\tau) = \frac{MSE(\tau)}{\left(\frac{1}{\sum_{i=1}^{n} m_i} tr\left(\mathbf{I} - \mathbf{A}(\boldsymbol{\tau})\right)\right)^2} \tag{2.10}$$

dengan $MSE(\tau) = \frac{1}{\sum_{i=1}^{n} m_i} \sum_{i=1}^{n} \sum_{j=1}^{m} \left(y_{ij} - \hat{y}_{ij}\right)^2$, dan $\mathbf{A}(\boldsymbol{\tau})$ diperoleh dari hubungan $\underset{\sim}{\hat{y}} = \mathbf{A}(\boldsymbol{\tau})\underset{\sim}{y}$. Nilai knot optimal diberikan oleh nilai GCV terkecil.

2.9 Koefisien Determinasi

Koefisien determinasi yang disimbolkan dengan R^2 adalah alat untuk mengukur proporsi keragaman atau variansi total di sekitar nilai tengah y yang dapat dijelaskan oleh model regresi. Secara umum semakin besar nilai R^2, maka semakin baik pula model yang didapatkan karena mampu menjelaskan lebih banyak data (Draper dan Smith, 1996). Persamaan R^2 diberikan sebagai berikut.

$$R^2 = \frac{SSR}{SST} = \frac{\left(\hat{Y}-\overline{Y}\right)'\left(\hat{Y}-\overline{Y}\right)}{\left(Y-\overline{Y}\right)'\left(Y-\overline{Y}\right)} \tag{2.11}$$

Besaran nilai R^2 tidak akan negatif dan batasannya adalah $0 \leq R^2 \leq 1$ (Gujarati dan Porter, 2009).

2.10 *Pivotal Quantity*

Statistik $Q(x_1, x_2, ..., x_n)$ disebut *pivotal quantity* apabila distribusi Q independen dari parameter.

Contoh:

Z= N(0,1)

Fungsi distribusi probabilitas $f(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$

Maka *pivotal quantity* yang dimaksud adalah jika distribusi Z independen dari parameter-parameter yang terdapat pada f (Z).

2.11 Selang Kepercayaan

Suatu selang kepercayaan merupakan suatu alat yang penting untuk mengamati variabilitas sampling dari suatu estimator. Dalam konteks regresi nonparametrik membentuk suatu selang kepercayaan merupakan hal yang sulit karena adanya bias dari estimator (Eubank dan Speckman, 1993). Estimasi selang adalah suatu selang tertentu yang memuat parameter dengan probabilitas tertentu.

Selang kepercayaan dari *f(x)* memungkinkan untuk menilai seberapa akurat dari estimator $\hat{f}(x)$ berada pada lokasi yang berbeda pada kisaran *range* tertentu. Selang kepercayaan kurva dapat dikonstruksi dengan mengestimasi variansi dari error (Wu dan Zhang, 2006).

Langkah-langkah untuk mendapatkan estimasi selang untuk suatu parameter dalam regresi dijelaskan oleh Montgomery *et al.* (2012) dengan menggunakan metode *pivotal quantity*. Misalkan X_1, X_2, …, X_n adalah sampel random yang diambil dari populasi dengan parameter θ, maka selang kepercayaan 1-α untuk θ adalah $P(a \le \theta \le b) = 1 - \alpha$. Diberikan sampel random X_1, X_2, …, X_n yang diambil dari populasi yang berdistribusi $N(\mu, \sigma^2)$ dengan σ^2 diketahui, maka untuk mendapatkan selang kepercayaan 1-α untuk parameter μ dapat menggunakan *pivotal quantity*.

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

Maka selang kepercayaan 1-α adalah sebagai berikut.

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

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

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

Namun, apabila σ^2 tidak diketahui, maka diganti dengan s^2 yaitu variansi sampel. Dengan demikian, untuk mendapatkan selang kepercayaan 1-α untuk μ digunakan *pivotal quantity*, sehingga selang kepercayaan 1-α adalah sebagai berikut.

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

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

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

2.12 *Open Source Software* R

R merupakan salah satu *software* yang sering digunakan dalam statistika dan termasuk dalam kategori *Open Source Software* (OSS) untuk memanipulasi data, simulasi, kalkulasi, dan peragaan *graphic*. Bahasa R berbasis bahasa S yang dibangun di Bell Laboratories di tahun 80 – an sehingga *syntax* R memiliki perbedaan yang tidak terlalu banyak atau hampir identik jika dibandingkan dengan *syntax* pada *software* S-Plus (Sawitzki, 2009). R mempunyai beberapa kelebihan dan fitur-fitur yang canggih dan berguna, diantaranya:

a. Efektif dalam pengolahan data dan fasilitas penyimpanan. Ukuran file yang disimpan jauh lebih kecil dibanding *software* lainnya.
b. Lengkap dalam perhitungan *array*.
c. Lengkap dan terdiri dari koleksi *tools* statistik yang terintegrasi untuk analisis data, diantaranya mulai statistik deskriptif, fungsi probabilitas, berbagai macam uji statistik hingga *time series*.
d. Tampilan grafik yang menarik dan fleksibel ataupun *costumized*.
e. Dapat dikembangkan sesuai keperluan dan kebutuhan karena sifatnya yang terbuka, setiap orang dapat menambahkan fitur-fitur tambahan dalan bentuk paket ke dalam *software* R.

Beberapa perintah internal yang digunakan dalam OSS-R adalah sebagai berikut.

1) function( ), merupakan perintah untuk menunjukkan kumpulan dari beberapa fungsi yang digunakan dalam program. Fungsi dipanggil dengan format *nama fungsi*(*daftar argumen*).
2) length( ), merupakan perintah yang digunakan untuk menghitung banyaknya data. Misalkan terdapat perintah *length*(*vector*), maka akan diperoleh hasil yaitu panjang dari *vector* tersebut.
3) plot( ), digunakan untuk membuat plot data. Beberapa penggunaan perintah ini diantaranya adalah sebagai berikut.
   a. plot(X,Y), artinya bahwa akan dibuat plot data berupa titik dengan sumbu datar X dan sumbu tegak Y.
   b. plot(X,Y,type=”l”) memberikan hasil plot bertipe garis.

c. plot(X,Y,type=”b”) memberikan hasil plot bertipe garis dan titik.

4) rep(a,b), merupakan perintah yang digunakan untuk membentuk suatu vektor dengan anggota a sebanyak b.
5) matrix(a, b, c), merupakan perintah yang digunakan untuk membentuk suatu matriks berukuran $b \times c$ dengan elemen a.
6) print( ), digunakan untuk menampilkan hasil atau *output* dari program.
7) cat(“...”), merupakan perintah untuk menuliskan kemudian menampilkan argumen dalam bentuk karakter.
8) for( ), merupakan perintah yang digunakan untuk mengulang satu blok pernyataan berulang kali hingga memenuhi kondisi yang telah ditentukan. Format penulisan perintah ini adalah for(*kondisi*) {*pernyataan*}.
9) repeat( ), hampir mirip dengan for( ), apabila kondisi sudah terpenuhi maka proses pengulangan akan dihentikan. Struktur penulisan *statment* repeat dalam R yaitu repeat{command if(*kondisi*) break}.
10) if-else, merupakan perintah yang digunakan untuk seleksi kondisi. Jika suatu kondisi bernilai benar, maka pernyataan pertama akan dijalankan, namun jika kondisi bernilai salah maka pernyataan kedua yang akan dijalankan. Struktur penulisan perintah ini adalah sebagai berikut.
    if(*kondisi*) {*pernyataan pertama*}
    else {*pernyataan kedua*}
11) solve(A), digunakan untuk menghitung *invers* dari suatu matriks A.
12) sum( ), digunakan untuk menghitung jumlah dari keseluruhan data.
13) rbind( ), digunakan untuk menggabungkan suatu matriks atau vektor berdasarkan baris.
14) cbind( ), digunakan untuk menggabungkan suatu matriks atau vektor berdasarkan kolom.

15) diag(*a*), merupakan perintah yang digunakan untuk membentuk suatu vektor *a* menjadi suatu matriks diagonal dengan elemen diagonal utamanya adalah elemen dari *a* dan elemen lain bernilai nol.
16) sort( ), merupakan perintah yang digunakan untuk mengurutkan sekumpulan data.
17) unique( ), digunakan untuk menentukan nilai tunggal dari suatu data.
18) quantile(..., ...), merupakan perintah untuk menentukan sampel kuantil.
19) order( ), merupakan perintah untuk menunjukkan vektor posisi data apabila data tersebut diurutkan.
20) var( ), merupakan perintah untuk menghitung nilai *varians* dari suatu vektor atau matriks variansi-kovariansi dari suatu matriks.
21) shapiro.test( ), merupakan perintah untuk menguji asumsi normalitas dengan *Shapiro wilks*.

BAB III
METODOLOGI PENELITIAN

3.1 Mengkonstruksi Selang Kepercayaan Kurva Regresi Nonparametrik berdasarkan Estimator *Spline Truncated* pada Data Longitudinal

Langkah- langkah yang dilakukan untuk menjawab tujuan dalam penelitian ini adalah sebagai berikut:

1. Diberikan data berpasangan $(y_{ij}, \mathrm{t}_{ij})$ dengan i = 1, 2, ..., n menyatakan indeks jumlah subjek yang diamati dan j = 1, 2,..., m_i menyatakan indeks jumlah pengamatan pada tiap subjek.
2. Hubungan antara t_{ij} dan y_{ij} diasumsikan mengikuti model regresi nonparametrik berdasarkan estimator *spline truncated* pada data longitudinal sebagai berikut :

   $y_{ij} = f(\mathrm{t}_{ij}) + \varepsilon_{ij}, \quad i = 1,2,...,n, \quad j = 1,2,...,m$ dengan $\varepsilon_{ij} \sim N(0, \sigma_i^2)$
3. Menghampiri kurva regresi $f(\mathrm{t}_{ij})$ dengan menggunakan *spline truncated* dengan p orde dan k knot:

   $f(\mathrm{t}_{ij}) = \beta_0 + \beta_1 t_{ij} + \beta_2 t_{ij}^2 + ... + \beta_p t_{ij}^p + \sum_{r=1}^{k} \beta_{(\mathrm{r}+p)} (\mathrm{t}_{ij} - \tau_r)_+^p$
4. Model regresi pada langkah (2) dapat disajikan dalam persamaan berikut.

   $y_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_2 t_{ij}^2 + ... + \beta_p t_{ij}^p + \sum_{r=1}^{k} \beta_{(\mathrm{r}+p)} (\mathrm{t}_{ij} - \tau_k)_+^p + \varepsilon_{ij}$
5. Menyajikan model regresi pada langkah (4) dalam bentuk matriks $\underset{\sim}{y} = \mathbf{T}\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$
6. Mengestimasi untuk parameter $\underset{\sim}{\beta}$ menggunakan metode *Weighted Least Square* dengan langkah- langkah sebagai berikut:
   i. Membentuk persamaan $Q = (\underset{\sim}{\mathrm{y}} - \mathbf{T}\underset{\sim}{\beta})^T \mathbf{W}^{-1} (\underset{\sim}{y} - \mathbf{T}\underset{\sim}{\beta})$, dengan $\mathbf{W}$ merupakan suatu matriks diagonal yang berisi pembobot untuk estimasi parameter.

$$\mathbf{W} = \text{cov}(\underset{\sim}{y}) = \begin{bmatrix} \mathbf{W}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{W}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{W}_n \end{bmatrix}$$

$$\mathbf{W}_i = \begin{bmatrix} \sigma_i^2 & 0 & \cdots & 0 \\ 0 & \sigma_i^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_i^2 \end{bmatrix}, \quad i = 1, 2, ..., n$$

ii. Meminimumkan persamaan Q dengan menyelesaikan persamaan berikut.

$$\frac{\partial Q}{\partial \underset{\sim}{\beta}^T} = 0$$

iii. Membuktikan turunan kedua dari Q terhadap $\underset{\sim}{\beta}$ bernilai positif.

$$\frac{\partial Q}{\partial \underset{\sim}{\beta}^T \underset{\sim}{\beta}} > 0$$

iv. Mendapatkan estimasi dari $\underset{\sim}{\beta}$ yaitu $\hat{\underset{\sim}{\beta}}$

7. Mendapatkan distribusi dari $\hat{\underset{\sim}{\beta}}$.
8. Kemudian memperoleh mean dan variansi dari $\hat{\underset{\sim}{f}}(t)$ yaitu $E(\hat{f}(\text{t}))$ dan $Var(\hat{f}(\text{t}))$, serta distribusi dari $\hat{\underset{\sim}{f}}(t)$.
9. Mendapatkan *pivotal quantity* untuk kurva regresi $\hat{\underset{\sim}{f}}(\text{t})$.

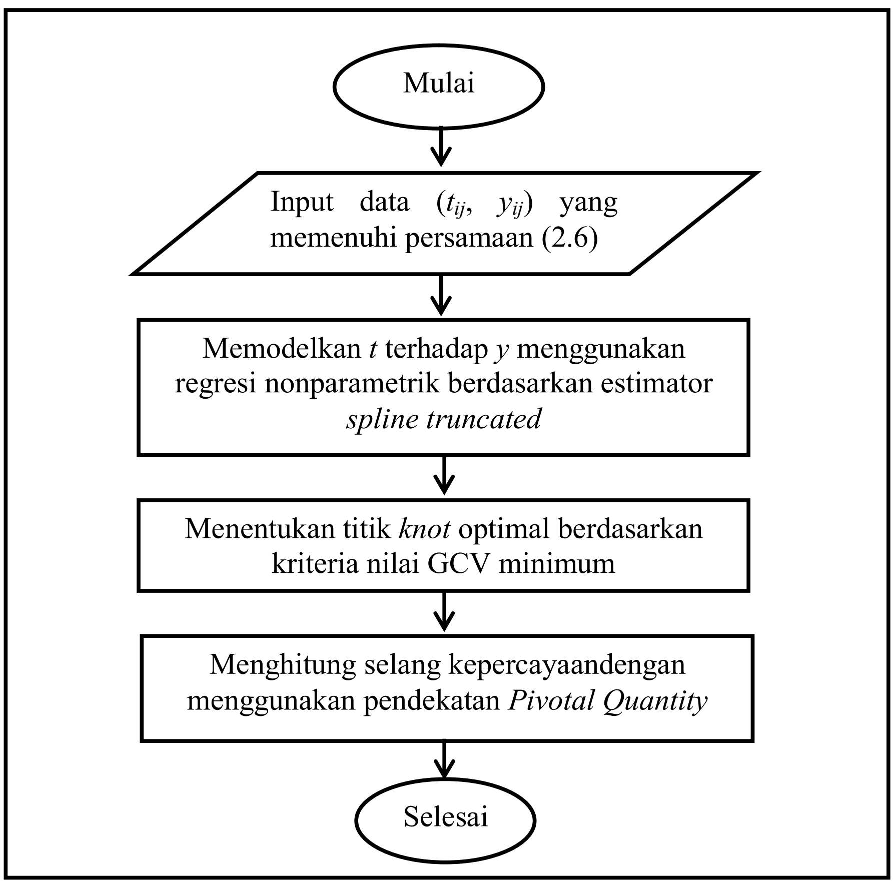
   Misalkan $U_{ij}(t)$ adalah *pivotal quantity* untuk $\hat{f}(\text{t}_{ij})$ dengan i = 1, 2, … , n dan j = 1, 2, … , m_i.
10. Menyelesaikan persamaan dalam probabilitas

    $P(a_{ij} \le U_{ij}(\text{t}) \le b_{ij}) = 1 - \alpha$

3.2 Membuat Algoritma dan Program untuk Selang Kepercayaan Kurva Regresi Nonparametrik Berdasarkan Estimator *Spline Truncated* pada Data Longitudinal dan Menerapkan pada Data

3.2.1 Algoritma dan Program untuk Mengkonstruksi Selang Kepercayaan Kurva Regresi Nonparametrik berdasarkan Estimator *Spline Truncated* pada Data Longitudinal

Penerapan model regresi nonparametrik berdasarkan estimator *Spline Truncated* pada data longitudinal tidak dapat dilakukan secara manual sehingga diperlukan bantuan *software* statistika untuk memperoleh penyelesaiannya. Langkah - langkah dalam merancang algoritma program untuk mengestimasi selang kepercayaan kurva regresi nonparametrik berdasarkan estimator *Spline Truncated* pada data longitudinal dapat digambarkan dalam **Gambar 3.1** berikut.



Gambar 3.1 Diagram Alir Algoritma dan Program

3.2.2 Penerapan Selang Kepercayaan Kurva Regresi Nonparametrik berdasarkan Estimator *Spline Truncated* pada Data Longitudinal pada Data Rill

a. Data dan Sumber Data

Data yang digunakan dalam skripsi ini adalah data sekunder pertumbuhan berat badan balita di Surabaya tahun 2015 dari penelitian yang dilakukan Azizah (2016). Data berjumlah 46 subjek dengan 23 balita laki-laki dan 23 balita perempuan, namun dalam skripsi ini hanya digunakan 16 subjek yaitu 8 balita laki-laki dan 8 balita perempuan. Unit observasi pada skripsi ini adalah balita di Surabaya yang telah memenuhi *screening* berdasarkan standar WHO.

b. Variabel Penelitian

Variabel-variabel penelitian yang digunakan pada penelitian ini disajikan pada **Tabel 3.1** berikut.

Tabel 3.1 Variabel-variabel Penelitian

No	Variabel	Keterangan	Satuan
1	t_{ij}	Usia balita pada subjek ke $-\ i$ dan pengamatan ke $-\ j$	Bulan
2	y_{ij}	Berat badan balita yang diukur pada saat t_{ij}	Kilogram (kg)

Struktur data yang digunakan dalam skripsi ini disajikan dalam **Tabel 3.2** berikut.

Tabel 3.2 Struktur Data Longitudinal

Subjek ke-i (i=1, 2, …, n)	Pengamatan ke-j (j=1, 2, …, m_i)	Variabel Respon (y)	Variabel Prediktor (t)
1	1	y_{11}	t_{11}
	2	y_{12}	t_{12}
	⋮	⋮	⋮
	m_1	y_{1m_1}	t_{1m_1}

Subjek ke-i (i=1, 2, …, n)	Pengamatan ke-j (j=1, 2, …, m_i)	Variabel Respon (y)	Variabel Prediktor (t)
2	1	y_{21}	t_{21}
	2	y_{22}	t_{22}
	⋮	⋮	⋮
	m_2	y_{2m_2}	t_{2m_2}
⋮	⋮	⋮	⋮
n	1	y_{n1}	t_{n1}
	2	y_{n2}	t_{n2}
	⋮	⋮	⋮
	m_n	y_{nm_n}	t_{nm_n}

BAB IV
HASIL DAN PEMBAHASAN

5.1 Konstruksi Selang Kepercayaan Kurva Regresi Nonparametrik berdasarkan Estimator *Spline Truncated* pada Data Longitudinal

Pada sub bab ini dibahas estimasi selang kepercayaan untuk kurva regresi nonparametrik dengan estimator *Spline Truncated* pada data longitudinal. Diberikan data berpasangan (t_{ij}, y_{ij}) , $i = 1, 2, ..., n; j = 1, 2, ..., m_i$ dengan y_{ij} adalah respon untuk subjek ke- i dan pengamatan ke- j dengan m_i pengamatan, yang mengikuti model regresi nonparametrik pada data longitudinal sebagai berikut.

$$y_{ij} = f(t_{ij}) + \varepsilon_{ij} \tag{4.1}$$

dengan $f\left(t_{ij}\right)$ adalah fungsi regresi yang diasumsikan *smooth* atau belum diketahui bentuk kurvanya sehingga didekati dengan pendekatan nonparametrik, dan ε_{ij} (error) berdistribusi normal dengan mean 0 dan variansi σ_i^2. Model regresi nonparametrik pada persamaan (4.1) untuk $i = 1, 2, ..., n; j = 1, 2, ..., m_i$ dapat dijabarkan sebagai berikut.

$$\left.\begin{aligned}
y_{11} &= f(t_{11}) + \varepsilon_{11} \\
y_{12} &= f(t_{12}) + \varepsilon_{12} \\
&\vdots \\
y_{1m_1} &= f(t_{1m_1}) + \varepsilon_{1m_1} \\
y_{21} &= f(t_{21}) + \varepsilon_{21} \\
y_{22} &= f(t_{22}) + \varepsilon_{22} \\
&\vdots \\
y_{2m_2} &= f(t_{2m_2}) + \varepsilon_{2m_2} \\
&\vdots \\
y_{n1} &= f(t_{n1}) + \varepsilon_{n1} \\
y_{n2} &= f(t_{n2}) + \varepsilon_{n2} \\
&\vdots \\
y_{nm_n} &= f(t_{nm_n}) + \varepsilon_{nm_n}
\end{aligned}\right\} \tag{4.2}$$

Berdasarkan persamaan (4.2) dapat ditulis dalam bentuk persamaan umum sebagai berikut.

$$\underset{\sim}{y} = f(\underset{\sim}{t}) + \underset{\sim}{\varepsilon} \tag{4.3}$$

Pada persamaan (4.2), kurva regresi $f(t_{ij})$ didekati dengan fungsi *spline truncated* berorde p dan knot $\tau_1, \tau_2, ..., \tau_k$ untuk $i = 1,2,...,n; j = 1,2,...,m_i$ diuraikan sebagai berikut.

$$\left.\begin{aligned}
f(t_{11}) &= \beta_0 + \beta_1 t_{11} + \beta_2 t_{11}^2 + \ldots + \beta_p t_{11}^p + \beta_{(p+1)}(t_{11} - \tau_1)_+^p + \ldots + \beta_{(p+k)}(t_{11} - \tau_k)_+^p \\
f(t_{12}) &= \beta_0 + \beta_1 t_{12} + \beta_2 t_{12}^2 + \ldots + \beta_p t_{12}^p + \beta_{(p+1)}(t_{12} - \tau_1)_+^p + \ldots + \beta_{(p+k)}(t_{12} - \tau_k)_+^p \\
&\vdots \\
f(t_{1m_1}) &= \beta_0 + \beta_1 t_{1m_1} + \beta_2 t_{1m_1}^2 + \ldots + \beta_p t_{1m_1}^p + \beta_{(p+1)}(t_{1m_1} - \tau_1)_+^p + \ldots + \beta_{(p+k)}(t_{1m_1} - \tau_k)_+^p \\
f(t_{21}) &= \beta_0 + \beta_1 t_{21} + \beta_2 t_{21}^2 + \ldots + \beta_p t_{21}^p + \beta_{(p+1)}(t_{21} - \tau_1)_+^p + \ldots + \beta_{(p+k)}(t_{21} - \tau_k)_+^p \\
f(t_{22}) &= \beta_0 + \beta_1 t_{22} + \beta_2 t_{22}^2 + \ldots + \beta_p t_{22}^p + \beta_{(p+1)}(t_{22} - \tau_1)_+^p + \ldots + \beta_{(p+k)}(t_{22} - \tau_k)_+^p \\
&\vdots \\
f(t_{2m_2}) &= \beta_0 + \beta_1 t_{2m_2} + \beta_2 t_{2m_2}^2 + \ldots + \beta_p t_{2m_2}^p + \beta_{(p+1)}(t_{2m_2} - \tau_1)_+^p + \ldots + \beta_{(p+k)}(t_{2m_2} - \tau_k)_+^p \\
&\vdots \\
f(t_{n1}) &= \beta_0 + \beta_1 t_{n1} + \beta_2 t_{n1}^2 + \ldots + \beta_p t_{n1}^p + \beta_{(p+1)}(t_{n1} - \tau_1)_+^p + \ldots + \beta_{(p+k)}(t_{n1} - \tau_k)_+^p \\
f(t_{n2}) &= \beta_0 + \beta_1 t_{n2} + \beta_2 t_{n2}^2 + \ldots + \beta_p t_{n2}^p + \beta_{(p+1)}(t_{n2} - \tau_1)_+^p + \ldots + \beta_{(p+k)}(t_{n2} - \tau_k)_+^p \\
&\vdots \\
f(t_{nm_n}) &= \beta_0 + \beta_1 t_{nm_n} + \beta_2 t_{nm_n}^2 + \ldots + \beta_p t_{nm_n}^p + \beta_{(p+1)}(t_{nm_n} - \tau_1)_+^p + \ldots + \beta_{(p+k)}(t_{nm_n} - \tau_k)_+^p
\end{aligned}\right\} \tag{4.4}$$

Persamaan (4.4) dapat dinyatakan dalam bentuk notasi sigma yaitu:

$$f(t_{ij}) = \sum_{d=0}^{p} \beta_d t_{ij}^d + \sum_{r=1}^{k} \beta_{(p+r)} (t_{ij} - \tau_r)_+^p \tag{4.5}$$

dengan $(t_{ij} - \tau_r)_+^p$ memenuhi persamaan berikut.

$$(t_{ij} - \tau_r)_+^p = \begin{cases} (t_{ij} - \tau_r)^p & , t_{ij} \geq \tau_r \\ 0 & , t_{ij} < \tau_r \end{cases} \tag{4.6}$$

Berdasarkan persamaan (4.5) apabila disubtitusikan ke persamaan (4.3) maka model regresi dapat ditulis sebagai berikut :

$$y_{ij} = \sum_{d=0}^{p} \beta_d t_{ij}^d + \sum_{r=1}^{k} \beta_{(p+r)} \left(t_{ij} - \tau_r \right)_+^p + \varepsilon_{ij} \tag{4.7}$$

Berdasarkan persamaan (4.7) untuk setiap subjek ke i = 1, 2, ..., n dan j = 1, 2, ..., m_i model regresi nonparametrik *Spline Truncated* untuk data longitudinal dapat disajikan dalam bentuk matriks sebagai berikut.

$$\begin{bmatrix} \underset{\sim}{y}_1 \\ \underset{\sim}{y}_2 \\ \vdots \\ \underset{\sim}{y}_n \end{bmatrix} = \begin{bmatrix} \mathbf{T}_1 \\ \mathbf{T}_2 \\ \vdots \\ \mathbf{T}_n \end{bmatrix} \underset{\sim}{\beta} + \begin{bmatrix} \underset{\sim}{\varepsilon}_1 \\ \underset{\sim}{\varepsilon}_2 \\ \vdots \\ \underset{\sim}{\varepsilon}_n \end{bmatrix} \tag{4.8}$$

dengan

$$\underset{\sim}{y}_i = \begin{bmatrix} y_{i1} & y_{i2} & \cdots & y_{im_i} \end{bmatrix}^T, i = 1, 2, ..., n$$

$$\mathbf{T}_i = \begin{bmatrix} 1 & t_{i1} & t_{i1}^2 & \cdots & t_{i1}^p & \left(t_{i1} - \tau_1\right)_+^p & \cdots & \left(t_{i1} - \tau_k\right)_+^p \\ 1 & t_{i2} & t_{i2}^2 & \cdots & t_{i2}^p & \left(t_{i2} - \tau_1\right)_+^p & \cdots & \left(t_{i2} - \tau_k\right)_+^p \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_{im_i} & t_{im_i}^2 & \cdots & t_{im_i}^p & \left(t_{im_i} - \tau_1\right)_+^p & \cdots & \left(t_{im_i} - \tau_k\right)_+^p \end{bmatrix},$$

$$\underset{\sim}{\beta} = \begin{bmatrix} \beta_0 & \beta_1 & \beta_2 & \cdots & \beta_p & \beta_{(p+1)} & \cdots & \beta_{(p+k)} \end{bmatrix}^T,$$

dan $\underset{\sim}{\varepsilon}_i = \begin{bmatrix} \varepsilon_{i1} & \varepsilon_{i2} & \cdots & \varepsilon_{im_i} \end{bmatrix}^T$.

Berdasarkan persamaan (4.8) diperoleh model regresi sebagai berikut:

$$\underset{\sim}{y} = \mathbf{T}\underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \tag{4.9}$$

dengan $\underset{\sim}{\varepsilon}$ berdistribusi $N(\underset{\sim}{0}, \sigma^2\mathbf{W})$, dan $\underset{\sim}{y}$ adalah vektor berdimensi $M \times 1$, **T** adalah matriks berordo $M \times (1 + p + k)$, $\underset{\sim}{\beta}$ adalah vektor berdimensi $(1 + p + k) \times 1$, dan $\underset{\sim}{\varepsilon}$ adalah vektor berdimensi $M \times 1$ dengan $M = \sum_{i=1}^{n} m_i$.

Pada penelitian menggunakan data longitudinal, estimasi parameter diperoleh dengan menggunakan *Weighted Least Square* (WLS) untuk mengatasi kasus heteroskesdastisitas. Estimasi $\underset{\sim}{\beta}$ pada persamaan (4.9) dapat diperoleh dengan menyelesaikan optimasi WLS berikut.

$$\min_{\beta}\left\{\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{\beta}\right)^{T}\mathbf{W}^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{\beta}\right)\right\} \tag{4.10}$$

dengan matriks **W** merupakan matriks pembobot (matriks variansi kovariansi $\underset{\sim}{y}$) berukuran $M \times M$.

$$\mathbf{W}=\operatorname{cov}(\underset{\sim}{y})=\begin{bmatrix}\mathbf{W}_1 & 0 & \cdots & 0\\ 0 & \mathbf{W}_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \mathbf{W}_n\end{bmatrix} \tag{4.11}$$

dengan $\mathbf{W}_i = diag\left(\sigma_i^2,\sigma_i^2,...,\sigma_i^2\right)$, untuk $i=1,2,...,n$ dan $\operatorname{cov}\left(y_{ik},y_{ir}\right)=0, \mathrm{k}\neq r$.

Penyelesaian optimasi (4.10) dilakukan dengan meminimumkan nilai **Q** sebagai berikut.

$$\begin{aligned}
\mathbf{Q} &= \left\{\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{\beta}\right)^{T}\mathbf{W}^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{\beta}\right)\right\}\\
&= \underset{\sim}{y}^{T}\mathbf{W}^{-1}\underset{\sim}{y}-2\underset{\sim}{y}^{T}\mathbf{W}^{-1}\mathbf{T}\underset{\sim}{\beta}+\underset{\sim}{\beta}^{T}\mathbf{T}^{T}\mathbf{W}^{-1}\mathbf{T}\underset{\sim}{\beta}\\
\frac{\partial\mathbf{Q}}{\partial\underset{\sim}{\beta}^{T}} &= -2\mathbf{T}^{T}\mathbf{W}^{-1}\underset{\sim}{y}+2\mathbf{T}^{T}\mathbf{W}^{-1}\mathbf{T}\underset{\sim}{\beta}\\
0 &= -2\mathbf{T}^{T}\mathbf{W}^{-1}\underset{\sim}{y}+2\mathbf{T}^{T}\mathbf{W}^{-1}\mathbf{T}\underset{\sim}{\beta}\\
2\mathbf{T}^{T}\mathbf{W}^{-1}\mathbf{T}\underset{\sim}{\beta} &= 2\mathbf{T}^{T}\mathbf{W}^{-1}\underset{\sim}{y}\\
\mathbf{T}^{T}\mathbf{W}^{-1}\mathbf{T}\underset{\sim}{\beta} &= \mathbf{T}^{T}\mathbf{W}^{-1}\underset{\sim}{y}
\end{aligned}$$

sehingga diperoleh:

$$\hat{\underset{\sim}{\beta}}=\left(\mathbf{T}^{T}\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{T}\mathbf{W}^{-1}\underset{\sim}{y} \tag{4.12}$$

Kemudian membuktikan bahwa turunan kedua dari Q terhadap $\hat{\underset{\sim}{\beta}}$ bernilai positif.

$$\frac{\partial\mathbf{Q}}{\partial\underset{\sim}{\beta}^{T}} = -2\mathbf{T}^{T}\mathbf{W}^{-1}\underset{\sim}{y}+2\mathbf{T}^{T}\mathbf{W}^{-1}\mathbf{T}\underset{\sim}{\beta}$$

$$\frac{\partial\mathbf{Q}}{\partial\underset{\sim}{\beta}^{T}\underset{\sim}{\beta}} = 2\mathbf{T}^{T}\mathbf{W}^{-1}\mathbf{T}$$

$$2\mathbf{T}^{T}\mathbf{W}^{-1}\mathbf{T}\geq 0$$

Beberapa sifat estimator untuk parameter $\hat{\underset{\sim}{\beta}}$ dapat diturunkan sebagai berikut:

1. $\hat{\underset{\sim}{\beta}}$ bersifat linier terhadap $\underset{\sim}{y}$

   Sifat linieritas $\hat{\underset{\sim}{\beta}}$ ditunjukkan oleh persamaan (4.12) yang memperlihatkan bahwa $\hat{\underset{\sim}{\beta}}$ merupakan hasil kali suatu matriks yang bersifat tetap dengan suatu vektor $\underset{\sim}{y}$ yang merupakan suatu variabel random.

2. $\hat{\underset{\sim}{\beta}}$ mempunyai sifat tak bias

   Pada persamaan (4.9) diketahui $\underset{\sim}{\varepsilon} \sim N(\underset{\sim}{0}, \sigma^2\mathbf{W})$ dan $\underset{\sim}{y}$ merupakan kombinasi linier dari $\underset{\sim}{\varepsilon}$, maka $\underset{\sim}{y}$ berdistribusi $N\left(\mathbf{T}\underset{\sim}{\beta}, \sigma^2\mathbf{W}\right)$ yang dapat dibuktikan menggunakan MGF sebagai berikut.

   $$M_{\underset{\sim}{\varepsilon}}(t) = e^{t'\underset{\sim}{0} + t'\sigma^2\mathbf{W}t/2} = e^{\sigma^2 t'\mathbf{W}t/2} \tag{4.13}$$

   Sehingga MGF $\underset{\sim}{y}$ diperoleh sebagai berikut.

   $$\begin{aligned} M_{\underset{\sim}{y}}(t) &= M_{\mathbf{T}\underset{\sim}{\beta}+\underset{\sim}{\varepsilon}}(t) \\ &= M_{\mathbf{T}\underset{\sim}{\beta}}(t) M_{\underset{\sim}{\varepsilon}}(t) \\ &= e^{t'\mathbf{T}\underset{\sim}{\beta}} e^{\sigma^2 t'\mathbf{W}t/2} \\ &= e^{t'\mathbf{T}\underset{\sim}{\beta} + \sigma^2 t'\mathbf{W}t/2} \end{aligned} \tag{4.14}$$

   Berdasarkan persamaan (4.14) maka dapat dituliskan bahwa:

   $$\underset{\sim}{y} \sim N\left(\mathbf{T}\underset{\sim}{\beta}, \sigma^2\mathbf{W}\right) \tag{4.15}$$

   Berdasarkan persamaan (4.15) maka dapat diperoleh ekspektasi dari $\hat{\underset{\sim}{\beta}}$ yaitu sebagai berikut.

   $$\begin{aligned} E\left(\hat{\underset{\sim}{\beta}}\right) &= E\left[\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\underset{\sim}{y}\right] \\ &= \left[\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}E\left(\underset{\sim}{y}\right)\right] \text{ (berdasarkan Teorema (2.7))} \\ &= \left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\underset{\sim}{\beta} \\ &= \underset{\sim}{\beta} \end{aligned} \tag{4.16}$$

Berdasarkan persamaan (4.16) diperoleh bahwa sifat estimator $\hat{\underset{\sim}{\beta}}$ adalah tak bias karena $E\left(\hat{\underset{\sim}{\beta}}\right)=\underset{\sim}{\beta}$.

3. $\hat{\underset{\sim}{\beta}}$ berdistribusi normal

Distribusi dari estimator parameter $\hat{\underset{\sim}{\beta}}$ dapat diturunkan dari sifat linieritas $\hat{\underset{\sim}{\beta}}$ terhadap $\underset{\sim}{y}$, karena $\underset{\sim}{\varepsilon}$ berdistribusi normal maka $\underset{\sim}{y}$ berdistribusi normal. Oleh karena, $\hat{\underset{\sim}{\beta}}$ mempunyai sifat linieritas terhadap $\underset{\sim}{y}$ maka dapat disimpulkan bahwa $\hat{\underset{\sim}{\beta}}$ berdistribusi normal atau dapat dituliskan $\hat{\underset{\sim}{\beta}}$ ~ $N\left(E\left(\hat{\underset{\sim}{\beta}}\right),Var\left(\hat{\underset{\sim}{\beta}}\right)\right)$. Nilai ekspektasi dari $\hat{\underset{\sim}{\beta}}$ disajikan dalam persamaan (4.16) dan variansi dari $\hat{\underset{\sim}{\beta}}$ dapat diperoleh sebagai berikut:

$$Var\left(\hat{\underset{\sim}{\beta}}\right)=Var\left[\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\underset{\sim}{y}\right] \tag{4.17}$$

berdasarkan Teorema (2.8) persamaan (4.17) menjadi

$$\begin{aligned}Var\left(\hat{\underset{\sim}{\beta}}\right)&=\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}Var\left(\underset{\sim}{y}\right)\left(\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\right)^T\\&=\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\sigma^2\mathbf{W}\mathbf{W}^{-1}\mathbf{T}\left(\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\right)^T\\&=\sigma^2\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\left(\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\right)^T\\&=\sigma^2\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\end{aligned} \tag{4.18}$$

Berdasarkan persamaan (4.16) dan (4.18), maka dapat disimpulkan distribusi dari estimator parameter $\hat{\underset{\sim}{\beta}}$ adalah :

$$\hat{\underset{\sim}{\beta}} \sim N\left(\underset{\sim}{\beta},\sigma^2\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\right) \tag{4.19}$$

Estimasi titik kurva regresi nonparametrik *Spline Truncated* pada data longitudinal dapat diperoleh dengan mensubstitusikan persamaan (4.12) ke dalam persamaan (4.9), sehingga diperoleh estimasi titik kurva kurva regresi nonparametrik *Spline Truncated* pada data longitudinal $\hat{\underset{\sim}{f}}(t)$ sebagai berikut.

$$
\begin{aligned}
\hat{\underset{\sim}{y}} = \hat{\underset{\sim}{f}}(t) &= \mathbf{T}\hat{\underset{\sim}{\beta}} \\
&= \mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\underset{\sim}{y} \\
&= \mathbf{A}\underset{\sim}{y} \qquad (4.20)
\end{aligned}
$$

dengan

$$\mathbf{A} = \mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}$$

Beberapa sifat estimator titik kurva regresi nonparametrik *Spline Truncated* $\hat{\underset{\sim}{f}}(t)$ pada data longitudinal dapat diturunkan sebagai berikut:

1. $\hat{\underset{\sim}{f}}(t)$ bersifat linier terhadap $\underset{\sim}{y}$

   Sifat linieritas $\hat{\underset{\sim}{f}}(t)$ ditunjukkan oleh persamaan (4.20) yang memperlihatkan bahwa $\hat{\underset{\sim}{f}}(t)$ merupakan hasil kali suatu matriks yang bersifat tetap dengan suatu vektor $\underset{\sim}{y}$ yang merupakan suatu variabel random.

2. $\hat{\underset{\sim}{f}}(t)$ mempunyai sifat tak bias

   Estimator $\hat{\underset{\sim}{f}}(t)$ bersifat tak bias apabila $E\left(\hat{\underset{\sim}{f}}(t)\right) = \underset{\sim}{f}(t)$ yang dapat ditunjukkan sebagai berikut:

$$
\begin{aligned}
E\left(\hat{\underset{\sim}{f}}(t)\right) &= E\left(\mathbf{T}\hat{\underset{\sim}{\beta}}\right) \\
&= \mathbf{T}E\left(\hat{\underset{\sim}{\beta}}\right) \quad \text{(berdasarkan Teorema (2.7))} \\
&= \mathbf{T}\underset{\sim}{\beta} \\
&= \underset{\sim}{f}(t) \qquad (4.21)
\end{aligned}
$$

3. $\hat{\underset{\sim}{f}}(t)$ berdistribusi normal

   Distribusi dari estimator kurva $\hat{\underset{\sim}{f}}(t)$ dapat diturunkan dari sifat linieritas $\hat{\underset{\sim}{f}}(t)$ terhadap $\underset{\sim}{y}$ yaitu berdistribusi normal atau dapat dituliskan dengan $\hat{\underset{\sim}{f}}(t) \sim N\left(E\left(\hat{\underset{\sim}{f}}(t)\right), Var\left(\hat{\underset{\sim}{f}}(t)\right)\right)$. Nilai ekspektasi dari $\hat{\underset{\sim}{f}}(t)$ tersebut dalam persamaan (4.21) dan variansi dari $\hat{\underset{\sim}{f}}(t)$ dapat diperoleh sebagai berikut:

$$\begin{aligned} Var\left(\hat{\underset{\sim}{f}}(t)\right) &= Var\left(\mathbf{T}\hat{\underset{\sim}{\beta}}\right) \\ &= \mathbf{T}Var\left(\hat{\underset{\sim}{\beta}}\right)\mathbf{T}^T \quad \text{(berdasarkan Teorema (2.8))} \\ &= \mathbf{T}\sigma^2\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T \\ &= \sigma^2\mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T \end{aligned} \tag{4.22}$$

Berdasarkan persamaan (4.21) dan (4.22) diperoleh distribusi $\hat{\underset{\sim}{f}}(t)$ yaitu

$$\hat{\underset{\sim}{f}}(t) \sim N\left(\underset{\sim}{f}(t), \sigma^2\mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\right) \tag{4.23}$$

karena $\hat{\underset{\sim}{f}}(t) = \left[\hat{f}(t_{11}) \quad \hat{f}(t_{12}) \quad \cdots \quad \hat{f}(t_{1m_1})\cdots \quad \hat{f}(t_{n1}) \quad \hat{f}(t_{n2})^T \quad \cdots\hat{f}(t_{nm_n})\right]^T$

sehingga var($\hat{f}(t_{ij})$) diperoleh dari elemen diagonal utama ke-*ij* matriks $\sigma^2\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)$ maka $\hat{f}(t_{ij}) \sim N(f(t_{ij}), \sigma^2\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij})$ dengan $i = 1, 2, ..., n$ dan $j = 1, 2, ..., m_i$.

Selanjutnya dikonstruksi selang kepercayaan $(1-\alpha)100\%$ untuk $f(t_{ij})$. Estimasi selang kepercayaan kurva regresi nonparametrik *spline truncated* pada data longitudinal digunakan *pivotal quantity*, sehingga didapatkan statistik sebagai berikut.

$$U_{ij}(t) = \frac{\hat{f}(t_{ij}) - E(\hat{f}(t_{ij}))}{\sqrt{Var(\hat{f}(t_{ij}))}} \tag{4.24}$$

dengan mensubtitusikan persamaan (4.23) ke dalam persamaan (4.24) maka diperoleh

$$U_{ij}(t) = \frac{\hat{f}(t_{ij}) - f(t_{ij})}{\sqrt{\sigma^2\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}}} \tag{4.25}$$

Dari persamaan (4.25) karena variansi tidak diketahui maka diestimasi dengan variansi sampel $(\hat{\sigma}^2)$, sehingga didapatkan *pivotal quantity* sebagai berikut.

$$U_{ij}(t) = \frac{\hat{f}(t_{ij}) - f(t_{ij})}{\sqrt{\hat{\sigma}^2 \left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}}} \tag{4.26}$$

$U_{ij}(t)$ merupakan *pivotal quantity* untuk kurva regresi $f(t_{ij})$. Kemudian $\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}$ menyatakan elemen diagonal utama ke-*ij* dari matriks $\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)$ dengan $i = 1,2,...,n$ dan $j = 1,2,...,m_i$ dan karena variansi populasi σ^2 tidak diketahui sehingga σ^2 diestimasi dengan $\hat{\sigma}^2$.

$$\begin{aligned}
\hat{\sigma}^2 &= \frac{\underset{\sim}{\varepsilon}^T \underset{\sim}{\varepsilon}}{M-(1+p+k)} \\
&= \frac{\left(\underset{\sim}{y}-\underset{\sim}{\hat{y}}\right)^T \left(\underset{\sim}{y}-\underset{\sim}{\hat{y}}\right)}{M-(1+p+k)} \\
&= \frac{\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{\hat{\beta}}\right)^T \left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{\hat{\beta}}\right)}{M-(1+p+k)} \\
&= \frac{\left(\underset{\sim}{y}-\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1}\underset{\sim}{y}\right)^T \left(\underset{\sim}{y}-\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1}\underset{\sim}{y}\right)}{M-(1+p+k)} \\
&= \frac{\left(\underset{\sim}{y}-\mathbf{A}\underset{\sim}{y}\right)^T \left(\underset{\sim}{y}-\mathbf{A}\underset{\sim}{y}\right)}{M-(1+p+k)} \\
&= \frac{\underset{\sim}{y}^T \underset{\sim}{y} - \underset{\sim}{y}^T \mathbf{A}\underset{\sim}{y} - \underset{\sim}{y}^T \mathbf{A}^T \underset{\sim}{y} + \left(\mathbf{A}\underset{\sim}{y}\right)^T \mathbf{A}\underset{\sim}{y}}{M-(1+p+k)}
\end{aligned}$$

Karena matriks $\mathbf{A}$ simetris dan idempoten maka $\left(\mathbf{A}\underset{\sim}{y}\right)^T \mathbf{A}\underset{\sim}{y} = \underset{\sim}{y}^T \mathbf{A}\underset{\sim}{y}$, sehingga

$$\begin{aligned}
\hat{\sigma}^2 &= \frac{\underset{\sim}{y}^T \underset{\sim}{y} - \underset{\sim}{y}^T \mathbf{A}\underset{\sim}{y} - \underset{\sim}{y}^T \mathbf{A}^T \underset{\sim}{y} + \underset{\sim}{y}^T \mathbf{A}\underset{\sim}{y}}{M-(1+p+k)} \\
&= \frac{\underset{\sim}{y}^T \underset{\sim}{y} - \underset{\sim}{y}^T \mathbf{A}^T \underset{\sim}{y}}{M-(1+p+k)}, \quad \text{karena matriks } \mathbf{A} \text{ simetris maka } \mathbf{A}^T = \mathbf{A} \\
&= \frac{\underset{\sim}{y}^T (\mathbf{I}-\mathbf{A})\underset{\sim}{y}}{M-(1+p+k)}
\end{aligned}$$

Maka diperoleh nilai $\hat{\sigma}^2$ sebagai berikut:

$$\hat{\sigma}^2 = \frac{\underset{\sim}{y}^T(\mathbf{I}-\mathbf{A})\underset{\sim}{y}}{M-(1+p+k)} \tag{4.27}$$

Berikut ini diberikan sebuah Lemma 4.1 tentang distribusi dari $U_{ij}(t)$.

Lemma 4.1.

Jika

$$U_{ij}(t) = \frac{\hat{f}(t_{ij}) - f(t_{ij})}{\sqrt{\hat{\sigma}^2\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}}}, \text{ maka } U_{ij}(t) \sim t_{(M-(1+p+k))}$$

Bukti

$$U_{ij}(t) = \frac{\hat{f}(t_{ij}) - f(t_{ij})}{\sqrt{\hat{\sigma}^2\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}}}$$

$$= \frac{\hat{f}(t_{ij}) - f(t_{ij})}{\sqrt{\frac{\underset{\sim}{y}^T(\mathbf{I}-\mathbf{A})\underset{\sim}{y}}{M-(1+p+k)}\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}}}$$

$$= \frac{\dfrac{\hat{f}(t_{ij}) - f(t_{ij})}{\sqrt{\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}}}}{\sqrt{\dfrac{\underset{\sim}{y}^T(\mathbf{I}-\mathbf{A})\underset{\sim}{y}}{M-(1+p+k)}}} \tag{4.28}$$

Kemudian pada persamaan (4.28) masing-masing pembilang dan penyebut dibagi dengan akar dari variansi populasi σ^2.

$$U_{ij}(t) = \frac{\dfrac{\hat{f}(t_{ij}) - f(t_{ij})}{\sqrt{\sigma^2\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}}}}{\sqrt{\dfrac{\underset{\sim}{y}^T(\mathbf{I}-\mathbf{A})\underset{\sim}{y}}{\sigma^2 M-(1+p+k)}}}$$

$$= \frac{Q_{ij}}{\sqrt{\dfrac{S}{M-(1+p+k)}}} \sim t_{(M-(1+p+k))} \tag{4.29}$$

dengan

$$Q_{ij} = \frac{\hat{f}(t_{ij}) - f(t_{ij})}{\sqrt{\sigma^2 \left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}}}$$

$$S = \frac{\underset{\sim}{y}^T(\mathbf{I} - \mathbf{A})\underset{\sim}{y}}{\sigma^2}$$

Untuk membuktikan kebenaran pada persamaan (4.29), maka harus dipenuhi Definisi (2.6). Langkah-langkah pembuktian sebagai berikut:

i. $Q_{ij} = \dfrac{\hat{f}(t_{ij}) - f(t_{ij})}{\sqrt{\sigma^2 \left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}}} \sim \mathrm{N}(0,\ 1)$

Dari persamaan (4.17) diketahui bahwa $\underset{\sim}{\hat{f}}(t)$ berdistribusi normal dan oleh karena Q merupakan kombinasi linier dari $\underset{\sim}{\hat{f}}(t)$, maka Q juga berdistribusi normal.

Q_{ij} berdistribusi $N\left(E(Q_{ij}), Var(Q_{ij})\right),\ i = 1, 2, ..., \mathrm{n},\ j = 1, 2, ..., \mathrm{m}_i$

Dengan

$$\begin{aligned}
E(Q_{ij}) &= E\left(\frac{\hat{f}(t_{ij}) - f(t_{ij})}{\sqrt{\sigma^2 \left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}}}\right) \\
&= \frac{1}{\sqrt{\sigma^2 \left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}}} E\left(\hat{f}(t_{ij}) - f(t_{ij})\right) \\
&= \frac{1}{\sqrt{\sigma^2 \left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}}} E\left(\hat{f}(t_{ij})\right) - f(t_{ij}) \\
&= \frac{1}{\sqrt{\sigma^2 \left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}}} (f(t_{ij}) - f(t_{ij})) \\
&= \frac{1}{\sqrt{\sigma^2 \left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}}} .0 \\
&= 0,\ i = 1, 2, ..., n \quad \mathrm{j} = 1, 2, ..., \mathrm{m}_i
\end{aligned} \tag{4.30}$$

$$Var(Q_{ij}) = Var\left(\frac{\hat{f}(t_{ij}) - f(t_{ij})}{\sqrt{\sigma^2\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}}}\right)$$

$$= \frac{1}{\sigma^2\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}} Var\left(\hat{f}(t_{ij}) - f(t_{ij})\right)$$

$$= \frac{1}{\sigma^2\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}} Var(\hat{f}(t_{ij})) - f(t_{ij})$$

$$= \frac{1}{\sigma^2\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}} Var(\hat{f}(t_{ij})) - 0$$

$$= \frac{1}{\sigma^2\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}} \sigma^2\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}$$

$$= 1, \quad i = 1, 2, ..., n \quad j = 1, 2, ..., m_i \tag{4.31}$$

Dari hasil perhitungan pada persamaan (4.30) dan (4.31) dapat disimpulkan bahwa

$$Q_{ij} = \frac{\hat{f}(t_{ij}) - f(t_{ij})}{\sqrt{\sigma^2\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}}} \sim N(0,1),\ i = 1, 2, ..., \mathrm{n}, \quad j = 1, 2, ..., m_i \tag{4.32}$$

ii. $S = \dfrac{\underset{\sim}{y}^T(\mathbf{I} - \mathbf{A})\underset{\sim}{y}}{\sigma^2} \sim \chi_{(M-(1+p+k))}$

Untuk menunjukkan bahwa S $\sim \chi^2_{(M-(1+p+k))}$ maka misalkan diambil

$$S = \frac{\underset{\sim}{y}^T(\mathbf{I} - \mathbf{A})\underset{\sim}{y}}{\sigma^2} = \frac{\underset{\sim}{y}^T\mathbf{D}\underset{\sim}{y}}{\sigma^2} \text{ dan } \lambda = \frac{1}{2}\left(\mathbf{T}\underset{\sim}{\beta}\right)^{\mathrm{T}}\mathbf{D}\left(\mathbf{T}\underset{\sim}{\beta}\right)$$

$$\mathbf{D} = \mathbf{I} - \mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1}$$

adalah matriks simetris dan idempoten. Kemudian akan dibuktikan berdasarkan Teorema (2.10) yaitu untuk

$$\underset{\sim}{y} \sim \mathrm{N}(\mathbf{T}\underset{\sim}{\beta}, \sigma^2\mathbf{W}) \Rightarrow \left[T = \underset{\sim}{y}^{\mathrm{T}}\mathbf{D}\underset{\sim}{y} \sim \chi^2_{(c,\lambda)} \Leftrightarrow \begin{matrix}\mathbf{D} \text{ adalah idempoten} \\ \text{c adalah } rank \text{ dari } \mathbf{D}\end{matrix}\right]$$

Matriks **D** dikatakan simetris jika $\mathbf{D}^{\mathbf{T}} = \mathbf{D}$**.** Berikut ini adalah pembuktian bahwa **D** adalah simetris.

$$\mathbf{D}^T = \left(\mathbf{I} - \mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1}\right)^T$$

Misal

$$\mathbf{A} = \mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1}$$

Merupakan matriks simetris dan idempoten maka

$$\mathbf{A}^T = \mathbf{A}$$

Sehingga

$$\begin{aligned}\mathbf{D}^T &= (\mathbf{I} - \mathbf{A})^T \\ &= (\mathbf{I} - \mathbf{A})\end{aligned}$$

$$\mathbf{D}^T = \mathbf{I} - \mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1} = \mathbf{D} \tag{4.33}$$

Berdasarkan persamaan (4.33) terlihat bahwa matriks **D** simetris dengan ukuran M x M. Selanjutnya dibuktikan juga bahwa matriks **D** idempoten. Matriks **D** dikatakan idempoten jika $\mathbf{D}^2 = \mathbf{D}$. Berikut ini adalah pembuktian bahwa **D** adalah idempoten.

$$\begin{aligned}\mathbf{D}^2 &= \left(\mathbf{I} - \mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1}\right)\left(\mathbf{I} - \mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1}\right) \\ &= \mathbf{I} - \mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1} - \mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1} \\ &\quad + \left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1}\right).\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1}\right) \\ &= \mathbf{I} - 2\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1} + \left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1}\right).\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1}\right) \\ &= \mathbf{I} - 2\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1} + \mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1} \\ &= \mathbf{I} - \mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{W}^{-1} \\ &= \mathbf{D}\end{aligned} \tag{4.34}$$

Terbukti bahwa **D** adalah matriks idempoten, karena **D** adalah simetris dan idempoten maka dapat dinyatakan bahwa

$$\mathrm{S} \sim \chi^2_{(rank(\mathrm{D}), \underset{\sim}{y}^T \mathrm{D}\underset{\sim}{y}/2\sigma^2)}$$

Selanjutnya akan didapatkan rank(**D)** dan $\underset{\sim}{y}^T \mathbf{D}\underset{\sim}{y}/2$.

Langkah pertama yaitu mendapatkan rank (**D)** berdasarkan Teorema (2.2), yaitu sebagai berikut.

$$
\begin{aligned}
c = rank(\mathbf{D}) &= rank\left(\mathbf{I}-\mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\right) \\
&= trace\left(\mathbf{I}-\mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\right) \\
&= trace\left(\mathbf{I}_M\right) - trace\left(\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right) \\
&= trace\left(\mathbf{I}_M\right) - trace\left(\mathbf{I}_{(1+p+k)}\right) \\
&= M-(1+p+k) \qquad (4.35)
\end{aligned}
$$

Selanjutnya akan dihitung nilai dari λ sebagai berikut:

$$
\begin{aligned}
\lambda &= \frac{\underset{\sim}{y}^T\mathbf{D}\underset{\sim}{y}}{2} \\
&= \frac{1}{2}\left(\mathbf{T}\underset{\sim}{\beta}\right)^T\mathbf{D}\left(\mathbf{T}\underset{\sim}{\beta}\right) \\
&= \frac{1}{2}\left(\left(\mathbf{T}\underset{\sim}{\beta}\right)^T\left(\mathbf{I}-\mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\right)\left(\mathbf{T}\underset{\sim}{\beta}\right)\right) \\
&= \frac{1}{2}\left(\left(\mathbf{T}\underset{\sim}{\beta}\right)^T\left(\mathbf{T}\underset{\sim}{\beta}\right)-\left(\mathbf{T}\underset{\sim}{\beta}\right)^T\left(\mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\right)\left(\mathbf{T}\underset{\sim}{\beta}\right)\right) \\
&= \frac{1}{2}\left(\underset{\sim}{\beta}^T\mathbf{T}^T\mathbf{T}\underset{\sim}{\beta}-\underset{\sim}{\beta}^T\mathbf{T}^T\mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)\underset{\sim}{\beta}\right) \\
&= \frac{1}{2}\left(\underset{\sim}{\beta}^T\mathbf{T}^T\mathbf{T}\underset{\sim}{\beta}-\underset{\sim}{\beta}^T\mathbf{T}^T\mathbf{T}\underset{\sim}{\beta}\right) \\
&= \frac{1}{2}.0 \\
&= 0 \qquad (4.36)
\end{aligned}
$$

Oleh karena hasil perhitungan dari persamaan (4.34), (4.35), dan (4.36), maka terbukti bahwa

$$
S = \frac{\underset{\sim}{y}^T\left(\mathbf{I}-\mathbf{A}\right)\underset{\sim}{y}}{\sigma^2} \sim \chi^2_{\left(M-(1+p+k)\right)} \qquad (4.37)
$$

iii. Langkah selanjutnya membuktikan $\mathbf{A}\underset{\sim}{y}$ dan $\underset{\sim}{y}^T\mathbf{D}\underset{\sim}{y}$ independen, dengan

$$
\mathbf{A}\underset{\sim}{y} = \underset{\sim}{\hat{f}}(t) = \mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\underset{\sim}{y}
$$

$$
\begin{aligned}
\underset{\sim}{y}^T\mathbf{D}\underset{\sim}{y} &= \underset{\sim}{y}^T\left(\mathbf{I}-\mathbf{A}\right)\underset{\sim}{y} \\
&= \underset{\sim}{y}^T\left(\mathbf{I}-\mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\right)\underset{\sim}{y}
\end{aligned}
$$

sehingga diperiksa independensi dari $\mathbf{A}\underset{\sim}{y}$ dan $\underset{\sim}{y}^T\mathbf{D}\underset{\sim}{y}$ yang memenuhi syarat pada Teorema (2.11). Apabila **A** matriks berukuran *M x M*, **D** adalah matriks konstanta yang simetris, dan $\underset{\sim}{y}$ berdistribusi $N(\mathbf{T}\hat{\underset{\sim}{\beta}},\sigma^2\mathbf{W})$, maka $\mathbf{A}\underset{\sim}{y}$ dan $\underset{\sim}{y}^T\mathbf{D}\underset{\sim}{y}$ independen jika dan hanya jika $\mathbf{A}\sigma^2\mathbf{W}\mathbf{D}=\mathbf{O}$.

$$\begin{aligned}
\mathbf{A}\sigma^2\mathbf{W}\mathbf{D} &= \left(\mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\right)\sigma^2\mathbf{W}\left(\mathbf{I}-\mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\right)\\
&=\sigma^2\left(\mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\mathbf{W}-\right.\\
&\quad\left.\mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\mathbf{W}^{-1}\mathbf{W}\right)\\
&=\sigma^2\left(\mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T-\mathbf{T}\left(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^T\right)\\
&=\sigma^2\mathbf{O}
\end{aligned}
\qquad =\mathbf{O}\ (4.38)$$

Berdasarkan persamaan (4.32), (4.37), dan (4.38) maka dapat disimpulkan bahwa persamaan (4.29) adalah benar. Selanjutnya persamaan (4.29) dapat dituliskan menjadi:

$$U_{ij}(t)=\frac{\dfrac{\hat{f}(t_{ij})-f(t_{ij})}{\sqrt{\sigma^2\left(\mathbf{T}(\mathbf{T}^T\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\right)_{ij}}}}{\sqrt{\dfrac{\dfrac{\underset{\sim}{y}^{\mathrm{T}}\left(\mathbf{I}-\mathbf{A}\right)\underset{\sim}{y}}{\sigma^2}}{M\text{ - }(1+\mathrm{p}+\mathrm{k})}}} \tag{4.39}$$

Proses untuk memperoleh *pivotal quantity* seperti pada persamaan (4.39) dapat dituliskan sebagai Teorema 4.1.

Teorema 4.1.

Misalkan **T** merupakan matriks berukuran M x (1+p+k) dan vektor $\underset{\sim}{\beta}$ berukuran $(1+p+k)$ x 1, model regresi nonparametrik *spline truncated* pada data longitudinal $\underset{\sim}{y}=\mathbf{T}\underset{\sim}{\beta}+\underset{\sim}{\varepsilon}$, $\underset{\sim}{\varepsilon}\sim$ $\mathrm{N}(\underset{\sim}{0},\sigma^2\mathbf{W})$ dengan σ^2 tidak diketahui dan estimasi kurva regresinya $\hat{\underset{\sim}{f}}(t)=\mathbf{A}\underset{\sim}{y}$ dengan $\mathbf{A}=\mathbf{T}(\mathbf{T}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{W}^{-1}$ adalah matriks yang simetris dan idempoten maka *pivotal quantity* untuk kurva regresi nonparametrik *spline truncated* pada data longitudinal pada subjek ke-i, pengamatan ke-j berdistribusi *t-student* dengan derajat bebas M - $(1+p+k)$ adalah sebagai berikut.

$$U_{ij}(t)=\frac{\hat{f}(t_{ij})-f(t_{ij})}{\sqrt{\frac{\underset{\sim}{y}^{\mathrm{T}}\left(\mathbf{I}-\mathbf{A}\right)\underset{\sim}{y}}{M\text{-}(1+p+k)}\left(\mathbf{T}(\mathbf{T}^{T}\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^{T}\right)_{ij}}}\quad i=1,2,...,n;\ j=1,2,...,m_i \tag{4.40}$$

Langkah yang dilakukan setelah memperoleh *pivotal quantity* adalah mengkonstruksi selang kepercayaan kurva regresi nonparametrik berdasarkan estimator *spline truncated* pada data longitudinal. Misalkan diambil suatu kontanta tertentu $a_{ij}(t)=a_{ij}, b_{ij}(t)=b_{ij}$, suatu *pivotal quantity* $U_{ij}(t)$, dan α merupakan taraf kesalahan maka dapat disusun persamaan probabilitas sebagai berikut.

$$P(a_{ij}\le U_{ij}(t)\le b_{ij})=1-\alpha,\ 0<\alpha<1,\ i=1,2,...,n; j=1,2,...,m_i \tag{4.41}$$

Berdasarkan persamaan (4.40) maka persamaan (4.41) dapat dituliskan menjadi.

$$P(a_{ij}\le\frac{\hat{f}(t_{ij})-f(t_{ij})}{\sqrt{\frac{\underset{\sim}{y}^{\mathrm{T}}\left(\mathbf{I}-\mathbf{A}\right)\underset{\sim}{y}}{M\text{-}(1+p+k)}\left(\mathbf{T}(\mathbf{T}^{T}\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^{T}\right)_{ij}}}\le b_{ij})=1-\alpha \tag{4.42}$$

Persamaan (4.42) dapat dituliskan kembali menjadi:

$$f(t_{ij})\le\hat{f}(t_{ij})-a_{ij}\sqrt{s_{ij}}\ \text{ dan }\ f(t_{ij})\ge\hat{f}(t_{ij})-b_{ij}\sqrt{s_{ij}} \tag{4.43}$$

dengan $s_{ij}=\dfrac{\underset{\sim}{y}^{T}\left(\mathbf{I}-\mathbf{A}\right)\underset{\sim}{y}}{M-(1+p+k)}\left(\mathbf{T}\left(\mathbf{T}^{T}\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{T}\right)_{ij}$

Persamaan (4.43) dapat dituliskan kembali menjadi:

$$P\left(\hat{f}(t_{ij})-b_{ij}\sqrt{s_{ij}}\le f(t_{ij})\le\hat{f}(t_{ij})-a_{ij}\sqrt{s_{ij}}\right)=1-\alpha \tag{4.44}$$

Kemudian $a_{ij}=-b_{ij}$ dan b_{ij} diperoleh dari distribusi *student-t* dengan $\frac{\alpha}{2}$ dan derajat bebas $M-(1+p+k)$ sehingga selang kepercayaan kurva regresi nonparametrik berdasarkan estimator *spline truncated* pada data longitudinal menjadi seperti berikut.

$$P\left(\hat{f}(t_{ij})-t_{\left(\frac{\alpha}{2},M-(1+p+k)\right)}\sqrt{s_{ij}}\leq f(t_{ij})\leq\hat{f}(t_{ij})+t_{\left(\frac{\alpha}{2},M-(1+p+k)\right)}\sqrt{s_{ij}}\right)=1-\alpha \tag{4.45}$$

5.2 Algoritma dan Program untuk Konstruksi Selang Kepercayaan Kurva Regresi Nonparametrik berdasarkan Estimator *Spline Truncated* pada Data Longitudinal menggunakan *Software* OSS-R yang diterapkan pada Data Rill

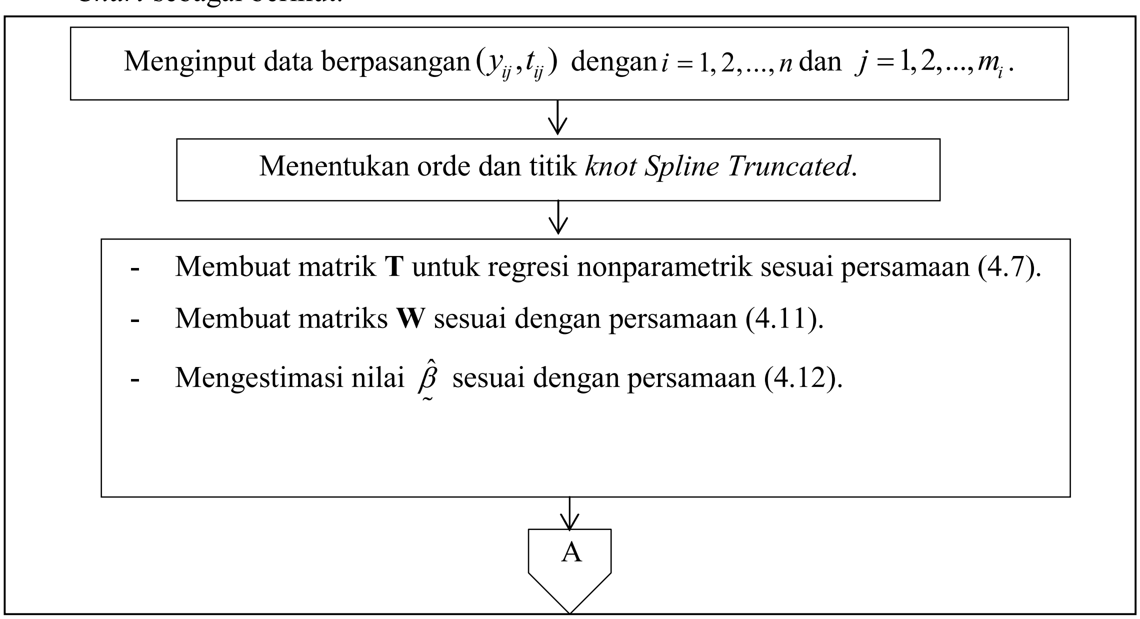
4.2.1 Algoritma untuk Konstruksi Selang Kepercayaan Kurva Regresi Nonparametrik berdasarkan Estimator *Spline Truncated* pada Data Longitudinal

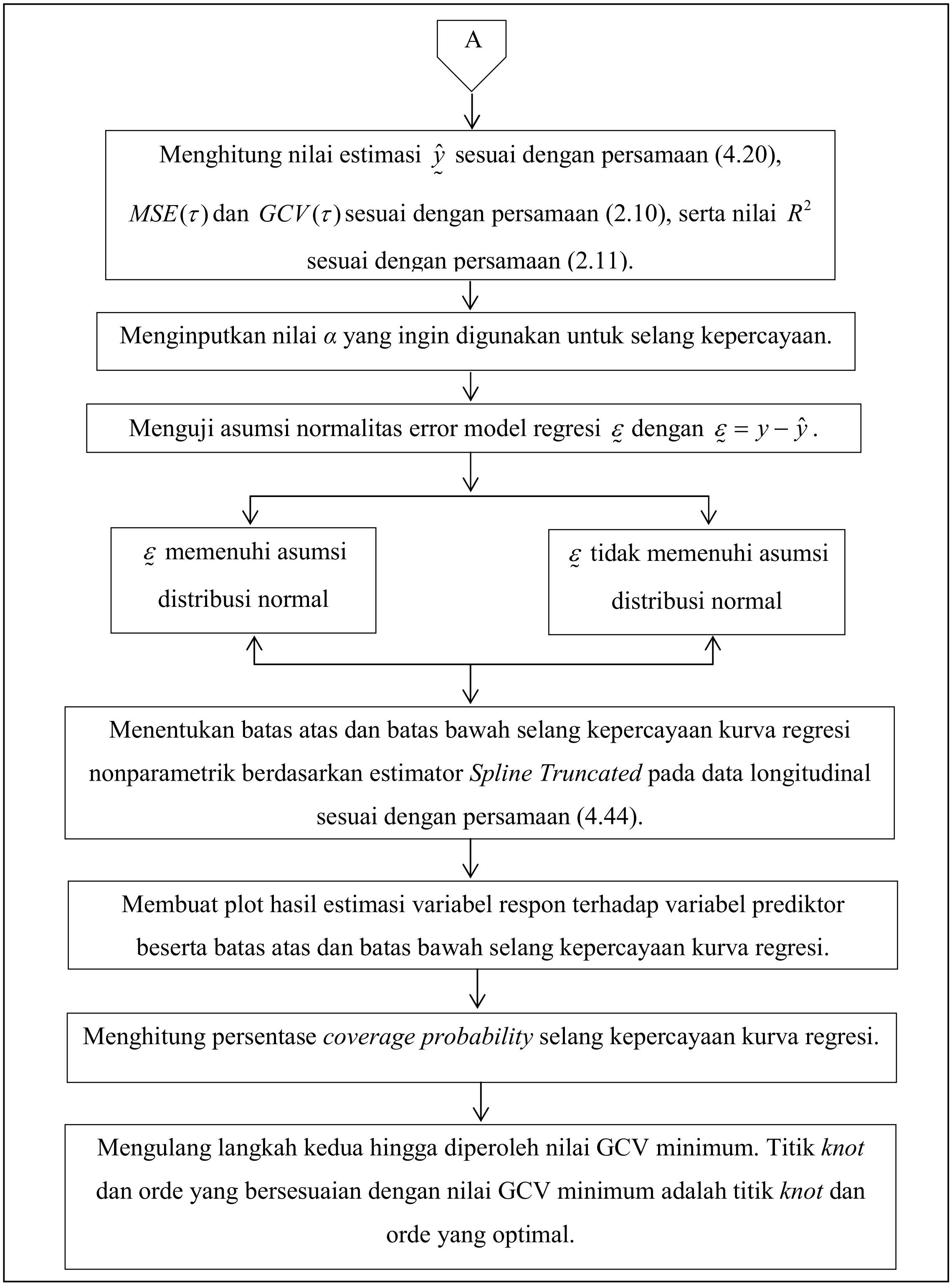
Algoritma untuk mengkonstruksi selang kepercayaan kurva regresi nonparametrik berdasarkan estimator *spline truncated* pada data longitudinal dengan menggunaan OSS-R dengan langkah-langkah sebagai berikut:

1) Menginput data berpasangan (y_{ij}, t_{ij}) dengan $i = 1, 2, ..., n$ dan $j = 1, 2, ..., m_i$.
2) Menentukan orde dan titik knot *spline truncated*.
3) Membuat matrik **T** untuk regresi nonparametrik sesuai persamaan (4.7)
4) Membuat matrik pembobot **W** dari kovariansi y sesuai dengan persamaan (4.11).
5) Mengestimasi nilai $\hat{\underset{\sim}{\beta}}$ berdasarkan persamaan (4.12)
6) Mengestimasi nilai $\hat{\underset{\sim}{y}}$ berdasarkan persamaan (4.20)
7) Menghitung nilai $MSE(\tau)$ sesuai dengan persamaan (2.10).
8) Menghitung nilai $GCV(\tau)$ sesuai dengan persamaan (2.10).
9) Menghitung nilai R^2 sesuai dengan persamaan (2.11).
10) Menginputkan nilai α yang ingin digunakan untuk selang kepercayaan.

11) Menguji asumsi normalitas error model regresi $\underset{\sim}{\varepsilon}$ dengan $\underset{\sim}{\varepsilon} = \underset{\sim}{y} - \underset{\sim}{\hat{y}}$. Kriteria pengujiannya yaitu jika *p-Value* < α maka $\underset{\sim}{\varepsilon}$ tidak memenuhi asumsi, sedangkan apabila *p-Value* > α maka $\underset{\sim}{\varepsilon}$ memenuhi asumsi pada persamaan (4.9).
12) Menentukan batas atas dan batas bawah selang kepercayaan kurva regresi nonparametrik berdasarkan estimator *Spline Truncated* pada data longitudinal sesuai dengan persamaan (4.45).
13) Membuat plot hasil estimasi variabel respon terhadap variabel prediktor beserta batas atas dan batas bawah selang kepercayaan kurva regresi nonparametrik berdasarkan estimator *Spline Truncated* pada data longitudinal.
14) Menghitung persentase *coverage probability* selang kepercayaan kurva regresi nonparametrik berdasarkan estimator *Spline Truncated* pada data longitudinal.
15) Mengulang langkah kedua hingga diperoleh nilai GCV minimum. Titik *knot* dan orde yang bersesuaian dengan nilai GCV minimum adalah titik *knot* dan orde yang optimal.

Berdasarkan algoritma yang telah dibuat, lebih jelasnya dapat dilihat pada *Flow Chart* sebagai berikut:

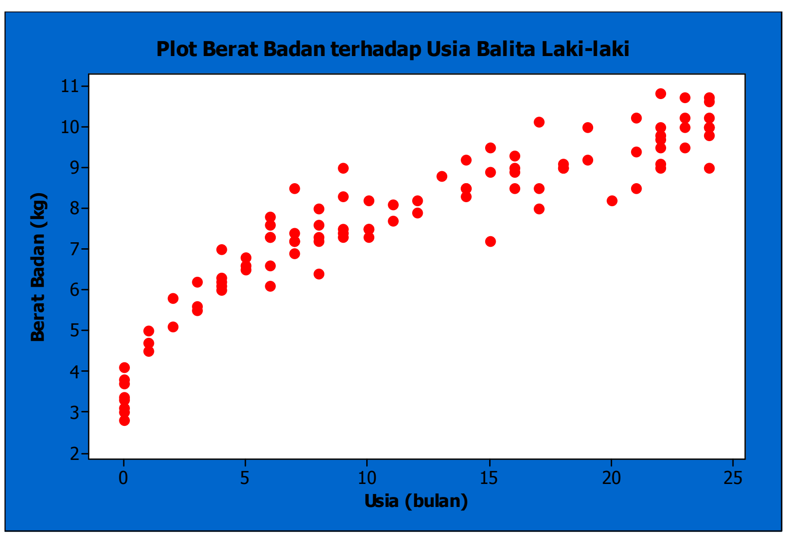




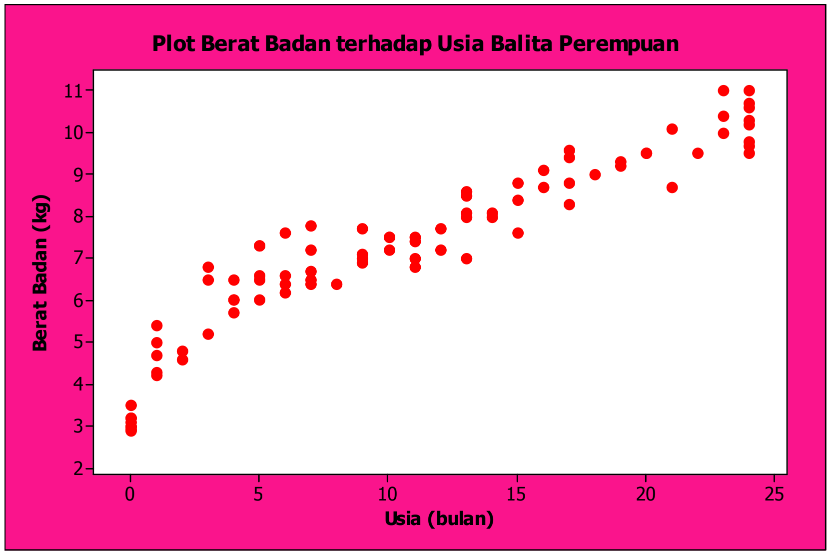
Gambar 4.1 *Flowchart* Algoritma dan Program

4.2.2 Penerapan Selang Kepercayaan Kurva Regresi Nonparametrik pada Data Longitudinal Berdasarkan Estimator Spline Truncated pada Data Pertumbuhan Berat Badan Balita Usia 0-24 Bulan

Data yang digunakan untuk penerapan model regresi nonparametrik pada data longitudinal berdasarkan estimator *spline truncated* adalah data pertumbuhan balita (0-24 bulan) di Surabaya tahun 2015 yang didapat dari Azizah (2016) dengan jumlah balita sebanyak 8 balita laki-laki dan 8 balita perempuan (Lampiran 1). Variabel yang digunakan dalam pengestimasian pertumbuhan balita (0-24 bulan) tersebut diantaranya variabel prediktor yang diasumsikan nonparametrik (t_{ij}) yaitu usia balita pada subjek ke-*i*, pengamatan ke-*j*, dan variabel respon (y_{ij}) yaitu berat badan balita (kg) yang diukur saat t_{ij}. Model regresi nonparametrik *spline truncated* data longitudinal pada berat badan balita tersebut kemudian dianalisis lebih lanjut dalam bentuk estimasi selang model berat badan balita dengan probabilitas sebesar 95%. Estimasi selang batas bawah dan batas atas dari model regresi nonparametrik *spline truncated* pada data longitudinal berat badan balita usia 0-24 bulan. Plot antara berat badan dan usia balita laki-laki serta perempuan disajikan dalam **Gambar 4.2** dan **4.3** berikut.



Gambar 4.2 Plot Berat Badan terhadap Usia (0-24 bulan) Balita Laki-laki di Surabaya Tahun 2015



Gambar 4.3 Plot Berat Badan terhadap Usia (0-24 bulan) Balita Perempuan di Surabaya Tahun 2015

Berdasarkan **Gambar 4.2** dan **Gambar 4.3** terlihat bahwa terjadi perubahan pola pertumbuhan balita laki-laki maupun perempuan pada interval usia tertentu. Oleh karena itu, untuk memodelkan pola hubungan antara berat badan dan usia ini, digunakan model *Spline Truncated***.** Pemilihan titik knot optimum dalam model regresi nonparametrik *spline truncated* pada data longitudinal yang terbaik untuk studi kasus data pertumbuhan balita laki-laki dan perempuan digunakan metode GCV minimum. Kemudian pemodelan dibagi menjadi dua, yaitu pertumbuhan balita laki-laki dan pertumbuhan balita perempuan.

a. Pertumbuhan Balita Laki-laki

Berikut merupakan hasil pemilihan titik knot optimum dalam model regresi nonparametrik *spline truncated* pada data longitudinal yang terbaik untuk studi kasus data pertumbuhan berat badan balita laki-laki.

Tabel 4.1 Pemilihan Orde Optimum berdasarkan Jumlah Knot Optimum Berat Badan Balita Laki-laki

Orde	Jumlah Knot Optimum	Titik Knot Optimum	GCV	R-Square
1	1	6	0,3295	0,9191
	1	12	0,4845	0,8811
	1	18	0,6285	0,8457
	2	6 ; 12	0,3365	0,9192
	2	6 ; 18	0,3365	0,9192
	2	12 ; 18	0,4696	0,8872
	3	6 ; 12 ; 18	0,3440	0,9192
2	1	6	0,3151	0,9243
	1	12	0,3406	0,9182
	1	18	0,3748	0,9100
	2	6 ; 12	0,3211	0,9246
	2	6 ; 18	0,3191	0,9251
	2	12 ; 18	0,3478	0,9183
	3	6 ; 12 ; 18	0,3259	0,9251

Berdasarkan kriteria nilai GCV yang minimum, dapat dipilih model regresi nonparametrik pada data berat badan balita laki-laki berdasarkan estimator *spline truncated* terbaik dengan orde respon yaitu orde 2 dengan jumlah titik knot sebanyak 1. Titik knot optimum terletak pada titik knot 6, maka diperoleh estimasi model rata-rata pertumbuhan berat badan balita laki-laki sebagai berikut:

$$\hat{y}_{lk} = 3,47650 + 1,03131t - 0,07151t^2 + 0,07061(t-6)_+^2 \qquad (4.46)$$

Persamaan (4.46) dapat diinterpretasikan dalam bentuk fungsi sebagai berikut:

$$\hat{y}_{lk} = \begin{cases} 3,47650 + 1,03131t - 0,07151t^2 & ; 0 \le t < 6 \\ 6,01846 + 0,18399t - 0,00090t^2 & ; \quad t \ge 6 \end{cases} \qquad (4.47)$$

Berdasarkan model pada persamaan (4.46), diperoleh nilai diperoleh nilai *R-square* sebesar 92,43%. Berdasarkan potongan polinomial persamaan (4.47), untuk menduga berat badan balita laki-laki gambarannya adalah jika ingin mengetahui berat badan balita laki-laki pada usia 0 bulan maka perhitungaannya dilakukan dengan cara melihat interval $0 \leq t < 6$ sehingga nilai estimasi berat badan balita laki-laki tersebut adalah 3,48 kg.

b. Pertumbuhan Balita Perempuan

Berikut merupakan hasil pemilihan titik knot optimum dalam model regresi nonparametrik *spline truncated* pada data longitudinal yang terbaik untuk studi kasus data pertumbuhan berat badan balita perempuan.

Tabel 4.2 Pemilihan Orde Optimum berdasarkan Jumlah Knot Optimum Berat Badan Balita Perempuan

Orde	Jumlah Knot Optimum	Titik Knot Optimum	GCV	R-Square
1	1	6	0,3499	0,9258
	1	12	0,5098	0,8919
	1	18	0,5471	0,8840
	2	6 ; 12	0,3284	0,9321
	2	6 ; 18	0,3519	0,9272
	2	12 ; 18	0,5196	0,8925
	3	6 ; 12 ; 18	0,3294	0,9336
2	1	6	0,2737	0,9434
	1	12	0,3692	0,9236
	1	18	0,4280	0,9115
	2	6 ; 12	0,2648	0,9466
	2	6 ; 18	0,2722	0,9451
	2	12 ; 18	0,3597	0,9274
	3	6 ; 12 ; 18	0,2713	0,9466

Berdasarkan kriteria nilai GCV yang minimum, dapat dipilih model regresi nonparametrik pada data berat badan balita perempuan berdasarkan estimator *spline truncated* terbaik dengan orde respon yaitu orde 2 dengan jumlah titik knot sebanyak 2. Titik knot optimum terletak pada titik knot 6 dan 12, maka diperoleh estimasi model rata-rata pertumbuhan berat badan balita perempuan sebagai berikut:

$$\begin{aligned}\hat{y}_{pr} &= 3,24354 + 1,09231t - 0,08730t^2 + 0,10519(t-6)_+^2 \\ &\quad - 0,02185(t-12)_+^2\end{aligned} \tag{4.48}$$

Persamaan (4.48) dapat diinterpretasikan dalam bentuk fungsi sebagai berikut:

$$\hat{y}_{pr} = \begin{cases} 3,24354 + 1,09231t - 0,08730t^2 & ;\ 0 \le t < 6 \\ 7,03038 - 0,16997t + 0,01789t^2 & ;\ 6 \le t < 12 \\ 3,88398 + 0,35443t - 0,00396t^2 & ;\quad t \ge 12 \end{cases} \tag{4.49}$$

Berdasarkan model pada persamaan (4.48), diperoleh nilai diperoleh nilai *R-square* sebesar 94,66%. Berdasarkan potongan polinomial persamaan (4.49), untuk menduga berat badan balita perempuan gambarannya adalah jika ingin mengetahui berat badan balita perempuan pada usia 0 bulan maka perhitungaannya dilakukan dengan cara melihat interval $0 \le t < 6$ sehingga nilai estimasi berat badan balita perempuan tersebut adalah 3,24 kg.

Persamaan (4.46) dan (4.48) merupakan model *spline* kuadratik terbaik dengan 1 titik knot yaitu 6 bulan untuk balita laki-laki dan 2 titik knot yaitu 6 dan 12 bulan untuk balita perempuan. Selanjutnya dibentuk selang kepercayaan 95% untuk kurva pertumbuhan berat badan balita laki-laki dan perempuan yang mengacu pada persamaan (4.46) dan (4.48). Namun sebelum membentuk selang kepercayaan, error model harus memenuhi asumsi distribusi normal. Uji asumsi distribusi normal dengan hipotesis sebagai berikut:

H_0 : Error berdistribusi normal

H_1 : Error tidak berdistribusi normal

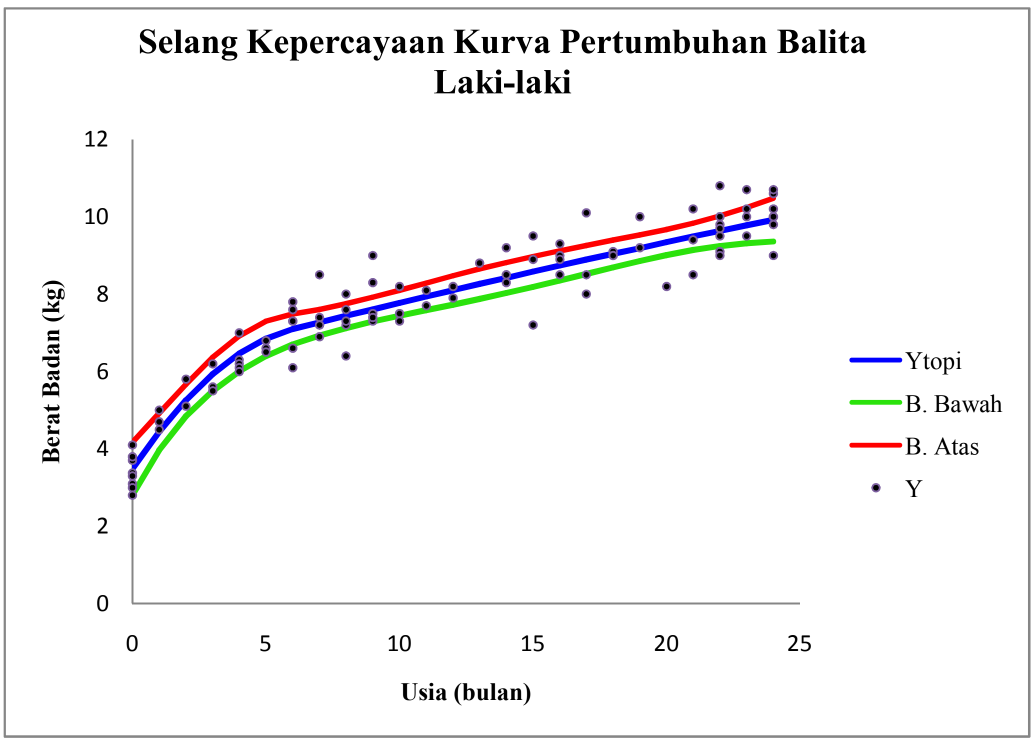
Berdasarkan uji *shapiro wilk* pada Lampiran 3 dan Lampiran 4 diperoleh nilai *p-value* sebesar 0,5825 untuk error berat badan balita laki-laki dan 0,101 untuk error berat badan balita perempuan. Kedua nilai *p-value* lebih besar dari nilai α sebesar 0,05 sehingga keputusannya H_0 diterima artinya error memenuhi asumsi normal. Batas bawah dan batas atas selang kepercayaan kurva pertumbuhan berat badan balita laki-laki dan perempuan data longitudinal disajikan sebagai berikut.

Tabel 4.3 Batas Bawah dan Batas Atas Selang Kepercayaan Kurva Pertumbuhan Berat Badan Balita Data Longitudinal

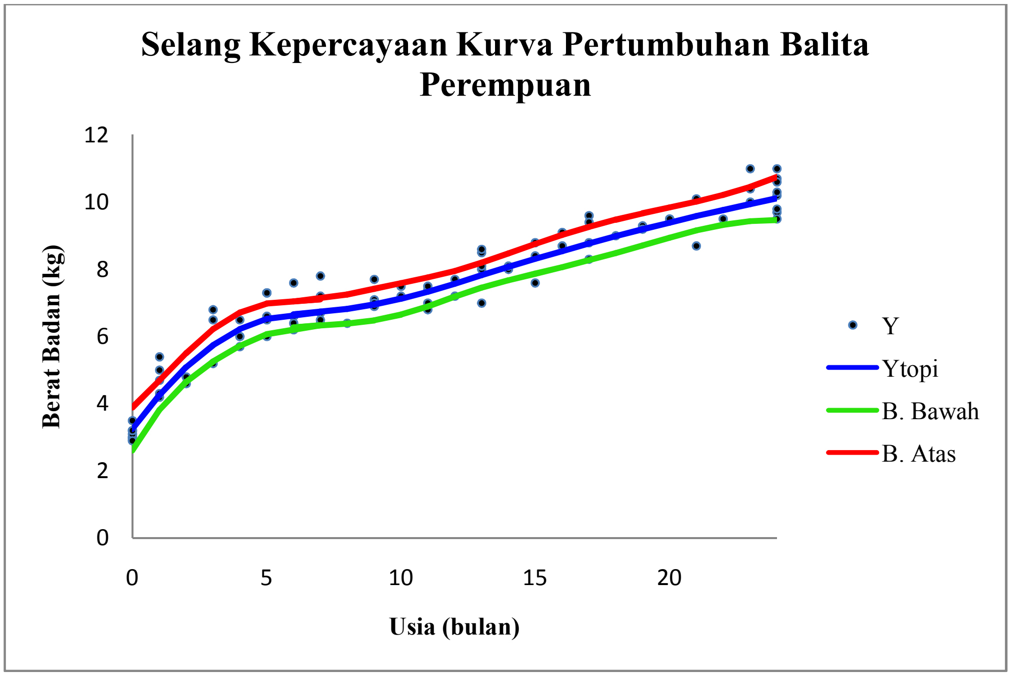
Balita Laki-laki			Balita Perempuan		
Usia	Batas Bawah	Batas Atas	Usia	Batas Bawah	Batas Atas
0	2,7922	4,1608	0	2,6001	3,8869
1	3,9581	4,9145	1	3,8080	4,6891
2	4,8415	5,6646	2	4,6511	5,5068
3	5,4946	6,3590	3	5,2548	6,2146
4	5,9989	6,9162	4	5,7184	6,7134
5	6,3958	7,2947	5	6,0659	6,9791
6	6,6943	7,4856	6	6,3309	7,1031
7	6,9236	7,6008	7	6,2683	7,0407
8	7,1197	7,7456	8	6,3805	7,2502
9	7,2878	7,9148	9	6,4793	7,4197
10	7,4384	8,0980	10	6,6494	7,5894
11	7,5815	8,2849	11	6,8916	7,7586
12	7,7243	8,4686	12	7,1850	7,9482
13	7,8707	8,6450	13	7,4510	8,1930
14	8,0227	8,8122	14	7,6662	8,4728
15	8,1811	8,9694	15	7,8632	8,7549
16	8,3454	9,1171	16	8,0610	9,0204
17	8,5142	9,2566	17	8,2675	9,2613
18	8,6843	9,3913	18	8,4848	9,4755

Balita Laki-laki			Balita Perempuan		
Usia	Batas Bawah	Batas Atas	Usia	Batas Bawah	Batas Atas
19	8,8503	9,5264	19	8,7106	9,6655
20	9,0041	9,6702	20	8,9375	9,8386
21	9,1360	9,8322	21	9,1499	10,0102
22	9,2384	10,0200	22	9,3231	10,2052
23	9,3097	10,2354	23	9,4322	10,4485
24	9,3521	10,4761	24	9,4693	10,7479

Dari model estimasi, batas bawah serta batas atas selang kepercayaan kurva, selanjutnya dapat diplotkan kurva hasil pengamatan (observasi) serta rata-rata estimasi titik dan estimasi selang pertumbuhan berat badan balita laki-laki dan perempuan yang dapat dilihat pada **Gambar 4.4** dan **Gambar 4.5** berikut.



Gambar 4.4 Kurva Pertumbuhan Berat Badan Balita Laki-laki di Surabaya Tahun 2015



Gambar 4.5 Kurva Pertumbuhan Berat Badan Balita Perempuan di Surabaya Tahun 2015

Berdasarkan **Gambar 4.4** dan **Gambar 4.5** dapat dilihat bahwa beberapa observasi masuk dalam selang kepercayaan. Pada kurva pertumbuhan berat badan balita laki-laki dari 95 observasi, yang berada di dalam selang sebanyak 61 observasi dan 34 observasi berada diluar selang. Sehingga dapat diketahui bahwa dengan selang kepercayaan sebesar 95%, peluang observasi yang masuk ke dalam selang yaitu sebesar 64,21 %. Demikian pula untuk kurva pertumbuhan berat badan balita perempuan dari 85 observasi, observasi yang berada di dalam selang sebanyak 56 dan 29 observasi berada diluar selang. Sehingga dapat diketahui bahwa dengan selang kepercayaan sebesar 95%, peluang observasi yang masuk ke dalam selang yaitu sebesar 65,88 %. Selain itu, berdasarkan **Gambar 4.4** dan **Gambar 4.5** serta menghitung rata-rata kenaikan di setiap usia dari model estimasi persamaan (4.46) dan (4.48) menunjukkan bahwa kenaikan berat badan balita laki-laki dan perempuan tertinggi terletak pada interval usia 0 – 6 bulan, yaitu setiap kenaikan 1 bulan, maka secara rata-rata berat badan balita mengalami kenaikan sebesar 0,60224 kg bagi balita laki-laki dan 0,5789 kg untuk balita perembuan. Hal tersebut menunjukkan bahwa pada batas usia ini pertumbuhan balita memasuki tahap awal pertumbuhan, sehingga balita mengalami pertumbuhan yang sangat pesat.

BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan analisis dan pembahasan pada bab sebelumnya, dapat ditarik kesimpulan sebagai berikut.

1. Selang kepercayaan kurva regresi nonparametrik *spline truncated* untuk data longitudinal pada penelitian ini dibangun menggunakan metode *pivotal quantity*. *Pivotal quantity* untuk kurva $\underset{\sim}{f}(t)$ saat variansi tidak diketahui adalah

$$U_{ij}(t)=\frac{\hat{f}(t_{ij})-f(t_{ij})}{\sqrt{\hat{\sigma}^{2}\left(\mathbf{T}(\mathbf{T}^{T}\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^{T}\right)_{ij}}}$$

$$=\frac{\hat{f}(t_{ij})-f(t_{ij})}{\sqrt{\frac{\underset{\sim}{y}^{T}(\mathbf{I}-\mathbf{A})\underset{\sim}{y}}{M-(1+p+k)}\left(\mathbf{T}(\mathbf{T}^{T}\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^{T}\right)_{ij}}}$$

Selang kepercayaan kurva regresi nonparametrik *spline truncated* untuk data longitudinal saat variansi tidak diketahui adalah sebagai berikut

$$P\left(\hat{f}(t_{ij})-t_{\left(\frac{\alpha}{2},M-(1+p+k)\right)}\sqrt{s_{ij}}\leq f(t_{ij})\leq\hat{f}(t_{ij})+t_{\left(\frac{\alpha}{2},M-(1+p+k)\right)}\sqrt{s_{ij}}\right)=1-\alpha$$

dengan

$$s_{ij}=\frac{\underset{\sim}{y}^{T}\left(\mathbf{I}-\mathbf{A}\right)\underset{\sim}{y}}{M-(1+p+k)}\left(\mathbf{T}\left(\mathbf{T}^{T}\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{T}\right)_{ij},\ \mathbf{A}=\mathbf{T}\left(\mathbf{T}^{T}\mathbf{W}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{T}\mathbf{W}^{-1},$$

$i=1,2,...,\mathrm{n}$ dan $j=1,2,...,m_i$ serta ij merupakan indeks elemen diagonal utama dari $\left(\mathbf{T}(\mathbf{T}^{T}\mathbf{W}^{-1}\mathbf{T})^{-1}\mathbf{T}^{T}\right)$.

2. Aplikasi model untuk regresi nonparametrik *spline truncated* untuk data longitudinal pada data pertumbuhan berat badan balita usia (0-24 bulan) di Surabaya tahun 2015. Model terbaik yang dihasilkan berdasarkan kriteria GCV terkecil yaitu orde 2 dengan titik knot 6 untuk jenis kelamin laki-laki dan titik knot 6 dan 12 untuk jenis kelamin perempuan. Secara umum, model terbaik yang diperoleh dapat dituliskan sebagai berikut:

   a. Balita laki-laki

   Estimasi model rata-rata pertumbuhan berat badan balita laki-laki sebagai berikut:

   $$\hat{y}_{lk} = 3,47650 + 1,03131t - 0,07151t^2 + 0,07061(t-6)_+^2$$

   Model tersebut dapat diinterpretasikan dalam bentuk fungsi sebagai berikut:

   $$\hat{y}_{lk} = \begin{cases} 3,47650 + 1,03131t - 0,07151t^2 & ;\ 0 \le t < 6 \\ 6,01846 + 0,18399t - 0,00090t^2 & ;\quad t \ge 6 \end{cases}$$

   Model yang terpilih memiliki nilai *R-square* sebesar 92,43%, dan dari 95 observasi, yang berada di dalam selang sebanyak 61 observasi dan 34 observasi berada diluar selang. Sehingga dapat diketahui bahwa dengan selang kepercayaan sebesar 95%, peluang observasi yang masuk ke dalam selang yaitu sebesar 64,21 %. Berdasarkan potongan polinomial tersebut, untuk menduga berat badan balita laki-laki gambarannya adalah jika ingin memprediksi berat badan balita laki-laki pada usia 0 bulan maka perhitungannya dilakukan dengan cara melihat interval $0 \le t < 6$ sehingga nilai estimasi berat badan balita laki-laki tersebut adalah 3,48 kg.

   b. Balita perempuan

   Estimasi model rata-rata pertumbuhan berat badan balita perempuan sebagai berikut:

   $$\hat{y}_{pr} = 3,24354 + 1,09231t - 0,08730t^2 + 0,10519(t-6)_+^2 - 0,02185(t-12)_+^2$$

Model tersebut dapat diinterpretasikan dalam bentuk fungsi sebagai berikut:

$$\hat{y}_{pr} = \begin{cases} 3,24354 + 1,09231t - 0,08730t^2 & ; 0 \leq t < 6 \\ 7,03038 - 0,16997t + 0,01789t^2 & ; 6 \leq t < 12 \\ 3,88398 + 0,35443t - 0,00396t^2 & ; \quad t \geq 12 \end{cases}$$

Model yang terpilih memiliki nilai *R-square* sebesar 94,66%, dan dari 85 observasi, yang berada di dalam selang sebanyak 61 observasi dan 34 observasi berada diluar selang. Sehingga dapat diketahui bahwa dengan selang kepercayaan sebesar 95%, peluang observasi yang masuk ke dalam selang yaitu sebesar 65,88 %. Berdasarkan potongan polinomial tersebut, untuk menduga berat badan balita perempuan gambarannya adalah jika ingin memprediksi berat badan balita perempuan pada usia 0 bulan maka perhitungannya dilakukan dengan cara melihat interval $0 \leq t < 6$ sehingga nilai estimasi berat badan balita perempuan tersebut adalah 3,24 kg.

5.2 Saran

Berdasarkan hasil yang telah diperoleh dari skripsi ini, saran yang dapat diberikan untuk penelitian selanjutnya adalah sebagai berikut.

1. Pada penelitian ini, selang kepercayaan kurva regresi nonparametrik *spline truncated* untuk data longitudinal dibangun dengan menggunakan metode *pivotal quantity*. Pada penelitian selanjutnya dapat dikembangkan selang kepercayaan yang dibangun dengan pendekatan *Bayesian*.
2. Variabel-variabel prediktor yang digunakan dalam penelitian ini terbatas hanya pada kasus pertumbuhan berat badan balita dan hanya menggunakan satu variabel prediktor. Pada penelitian selanjutnya dapat digunakan kasus data lainnya, serta menggunakan variabel prediktor yang lebih banyak lagi.

DAFTAR PUSTAKA

Azizah, Z., 2016, Estimasi Model Regresi Semiparametrik Birespon pada Data Longitudinal berdasarkan Estimator Lokal Linier, *Skripsi*, Fakultas Sains dan Teknologi Universitas Airlangga, Surabaya.

Bintariningrum, M. F. dan Budiantara, I. N., 2014, Pemodelan Regresi Nonparametrik Spline Truuncated dan Aplikasinya pada Angka Kelahiran Kasar di Surabaya, *Jurnal Sains dan Seni POMITS Vol. 3*, No.1, 7-12.

Budiantara, I. N., 2001, Estimasi Parametrik dan Nonparametrik untuk Pendekatan Kurva Regresi, *Makalah Pembicara Utama pada Seminar Nasional Statistika V,* Jurusan Statistika ITS Surabaya.

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

Diggle, P. J., Heagerty, P., Liang, K.Y. dan Zelger, S.L., 2002, *Analysis of Longitudinal Data*, Oxford University Press, Oxford.

Draper, N. dan Smith, H., 1992, *Analisis Regresi Terapan*, Edisi ke 2, Diterjemahkan oleh Bambang Sumantri, Gramedia, Jakarta.

Eubank, R. L., 1988, Spline *Smoothing and Nonparametric Regression*, Marcel Dekker, Inc., New York.

Eubank, R. L. dan Speckman, P. L., 1991, A Bias Reduction Theorem with Applications in Nonparametric regression, *Scand. J. Statist. 18*, 211-222.

Fadhilah, K.N., Suparti, Tarno, 2016, Pemodelan Regresi Spline Truncated untuk Data Longitudinal (Studi Kasus : Harga Saham pada Kelompok Saham Perbankan Periode Januari 2009 – Desember 2015), *Jurnal Gaussian,* Volume 5*,* No 3, Tahun 2016, Hal. 447 – 454, ISSN : 2339 – 2541.

Frees, E. W., 2003, *Longitudinal and Panel Data: Analysis and Applications for the Social Sciences,* Cambridge: Cambridge University Press.

Greene, W., 2003, *Econometric Analysis Fifth Edition*, Prentice Hall, New York.

Gujarati, D.N., dan Porter, D.C., 2009, *Basic Econometric 5th Edition,* McGraw-Hill Companies, Inc., New York.

Hardle W., 1990, *Applied Non Parametric Regression*, Cambridge University Press, New York.

Hardle, W., 1994, *Applied Nonparametric Regression*, Humboldt-University of Berlin, Berlin.

Hogg, R. V., McKean, J.W., dan Craig , A.T., 2005, *Introduction to Mathematical Statistics, 6thedition*, Pearson Prentice Hall, USA.

Intansari, I. A. S., 2016, Inferensi Statistik untuk Kurva Regresi Nonparametrik Spline Kuadratik dan Aplikasinya pada Data ASFR (Age Spesific Fertility Rate) di Bali, *Tesis*, Jurusan Statistika ITS Surabaya.

Malik, S., 2014, Estimasi Kurva Regresi Nonparametrik Multivariabel untuk Data Longitudinal dengan Pendekatan Spline, *Tesis*, Jurusan Statistika ITS Surabaya.

Montgomery, D. C., Peck., E. A., dan Vining, G. G., 2012, *Introduction to Linier Regression Analysis Fifth Edition*, New Jersey: John Wiley & Sons.

Mubarak, R. dan Budiantara, I. N., 2012, Analisis Regresi Spline Multivariabel untuk Pemodelan Kematian Penderita Demam Berdarah Dengue (DBD) di Jawa Timur, *Jurnal Sains dan Seni POMITS*, Surabaya.

Nafi', M., 2010, Estimasi Interval Spline dalam Regresi Nonparametrik, *Tesis*, Jurusan Statistika ITS Surabaya.

Rencher, A.C. and Schaalje, G.B., 2008, *Linier Models in Statistic*, Second Edition, John Wiley and Sons Inc., USA.

Sawitzki, G., 2009, *Computational Statistics: An Introduction to R*, Chapman & Hall/CRC Press, USA.

Setiawan, R. N. S., 2017, Interval Konfidensi untuk Parameter Model Regresi Nonparametrik Spline Truncated Multivariabel, *Tesis*, Jurusan Statistika ITS Surabaya.

Sriliana, I., 2012, Regresi Spline Tuncated dalam Model Linear Parsial untuk Data Longitudinal, *Tesis*, Jurusan Statistika ITS Surabaya.

Syaranamual, R.D. dan Budiantara, I. N., 2011, Interval Konvidensi Spline Kuadrat dengan Pendekatan Pivotal Quantity, *Prosiding Seminar Nasional Statistika Universitas Diponegoro 2011*, ISBN: 978-979-097-142-4.

Wahba, G., 1983, Bayesian "Confidence Interval" for the Cross-validated Smoothing Spline, *J.R. Statist. Sc. B.,* 45(1):133-150.

Wang, Y., 1998, Spline Smoothing Models with Correlated Errors. *Journal of the American Statistical Association*, 93, 341-348.

Wu, H. & Zhang, J., 2006, *Nonparametric Regression Methods for Longitudinal Data Analysis*, New Jersey : John Willey & Sons, Inc., Publication.

Lampiran 1. Data Pertumbuhan Balita di Surabaya Tahun 2015

Laki-Laki		
Subyek	Usia	BB(kg)
1	0	3,38
	2	5,8
	5	6,8
	6	7,3
	7	7,4
	9	7,3
	12	7,9
	18	9,1
	22	9,1
	24	10
2	0	4,1
	3	6,2
	4	6,3
	6	7,8
	8	8
	9	8,3
	11	7,7
	16	9,3
	22	10
	23	10,2
3	0	3,7
	10	7,5
	16	9
	18	9
	19	10
	22	9,8
	24	10,6
4	0	3,8
	6	7,6
	7	8,5
	9	9
	21	10,2
	23	10
	24	10,2

Laki-Laki		
Subyek	Usia	BB(kg)
5	0	3,3
	4	6,2
	5	6,6
	7	7,2
	8	7,2
	9	7,5
	10	7,5
	14	8,5
	16	8,5
	17	8,5
	22	9,5
	23	9,5
	24	9,8
6	0	3,1
	1	4,7
	3	5,6
	4	6,1
	6	6,1
	8	6,4
	14	8,3
	15	7,2
	17	8
	20	8,2
	21	8,5
	22	9
	24	9
7	0	3
	1	5
	4	7
	6	7,3
	7	7,2
	8	7,6
	10	8,2
	11	8,1

Laki-Laki		
Subyek	Usia	BB(kg)
7	12	8,2
	14	8,5
	15	8,9
	16	8,9
	18	9
	19	9,2
	21	9,4
	22	9,7
	24	10
8	0	2,8
	1	4,5
	2	5,1
	3	5,5
	4	6
	5	6,5
	6	6,6
	7	6,9
	8	7,3
	9	7,4
	10	7,3
	13	8,8
	14	9,2
	15	9,5
	17	10,1
	22	10,8
	23	10,7
	24	10,7

Perempuan		
Subyek	Usia	BB(kg)
1	0	3,5
	1	5
	3	6,5
	5	7,3
	6	7,6
	13	8,5
	18	9
	24	10,2
2	0	3,1
	7	7,2
	13	8
	16	9,1
	19	9,3
	24	10,7
3	0	3
	1	4,2
	2	4,8
	3	5,2
	4	5,7
	5	6
	6	6,2
	7	6,4
	9	6,9
	11	7
	12	7,2
	14	8
	24	9,7
4	0	2,93
	9	7,7
	11	7,4
	15	8,8
	17	9,4
	21	10,1
	24	10,3

Perempuan		
Subyek	Usia	BB(kg)
5	0	3,2
	1	5,4
	4	6,5
	5	6,5
	9	7,1
	10	7,5
	20	9,5
	23	10,4
	24	10,6
6	0	2,9
	1	4,3
	2	4,6
	4	6
	5	6,6
	6	6,6
	7	6,7
	8	6,4
	9	7
	10	7,2
	11	7,5
	12	7,7
	13	8,1
	14	8,1
	15	8,4
	16	8,7
	17	8,8
	19	9,2
	20	9,5
	22	9,5
	23	10
	24	9,8

Perempuan		
Subyek	Usia	BB(kg)
7	0	3
	3	6,8
	5	7,3
	7	7,8
	13	8,6
	17	9,6
	23	11
	24	11
8	0	3,2
	1	4,7
	4	6
	6	6,4
	7	6,5
	10	7,5
	11	6,8
	13	7
	15	7,6
	17	8,3
	21	8,7
	24	9,5

Lampiran 2. Program R Selang Kepercayaan Kurva Regresi Nonparametrik berdasarkan Estimator Spline Truncated pada Data Longitudinal

```
mp<-function(x,eps=1e-006)
{
	x<-as.matrix(x)
	xsvd<-svd(x)
	diago<-xsvd$d[xsvd$d>eps]
	if(length(diago)==1)
	{
	xplus<-
as.matrix(xsvd$v[,1])%*%t(as.matrix(xsvd$u[,1])/diago)
	}
	else
	{
xplus<-
xsvd$v[,1:length(diago)]%*%diag(1/diago)%*%t(xsvd$u[,1:length(diag
o)])
	}
	return(xplus)
}

trun<-function(prediktor,knot,orde)
{
	prediktor[prediktor<knot]<-knot
	b<-(prediktor-knot)^orde
	return(b)
}


#estimasi dengan pembobot
estimasi_pembobot<-function(data)
{
	t<-data[,1]
	y<-data[,2]
	M<-length(y)
	n<-as.numeric(readline("Inputkan banyak subyek : "))
	p<-as.numeric(readline("Input orde : "))
	k<-as.numeric(readline("Input jumlah knot optimum : "))
	w<-rep(0,k)
	for(i in 1:k)
	{
		cat("Input titik knot optimum  ke-",i)
		w[i]<-as.numeric(readline(" = "))
	}
	alfa<-as.numeric(readline("alfa : "))
	gm<-rep(0,(n+1))
	gm[1]<-0
	jpmn <-rep(0,n)
	for(i in 2:n)
	{
		s<-gm[i-1]+1
		repeat
		{
			s<-s+1
			if(t[s]<t[s-1])break
```

```
            g<-s
        }
        gm[i]<-g
        jpmn [i-1]<-gm[i]-gm[i-1]
}
jpmn [n]<-M-sum(jpmn [1:(n-1)])
cat("balita ke-\tjumlah pengamatan\n")
for( i in 1:n)
cat(i,"\t\t", jpmn [i],"\n")
v1<-matrix(0,M,p)
for(i in 1:p)
        {
            v1[,i]<-t^(i)
        }
v2<-matrix(0,M,k)
for(j in 1:k)
        {
            v2[,j]<-trun(t,w[j],p)
        }
T=cbind(v1,v2)
t1=rep(1,M)
C=cbind(t1,T)
wadah<-matrix(0,M,1)
c<-rep(0,(M+1))
c[1]<-0
for(i in 1:n)
{
        c[i+1]<-sum(jpmn[1:i])
        wadah[(c[i]+1):c[i+1],1]<-rep(i,jpmn[i])

}
vr=rep(0,n)
z=1
for(i in 1:n)
{
        zt=z+jpmn[i]-1
        vr[i]=var(data[(z:zt),2])
        z=zt+1
}
V=rep(0,M)
for(i in 1:M)
{
        V[i]=vr[wadah[i,1]]
}
S=diag(V)
W=solve(S)
beta<-mp(t(C)%*%W%*%C)%*%t(C)%*%W%*%y
ytopi=C%*%beta
print(beta)
Anon=C%*%mp(t(C)%*%W%*%C)%*%t(C)%*%W
MSE<-(t(y-ytopi)%*%(y-ytopi))/(M)
cat("MSE=",MSE,"\n")
GCV<-MSE/(((1/M)%*%(sum(diag(1-Anon))))^2)
cat("GCV=",GCV,"\n")
JKT<-t(y-(mean(y)))%*%(y-(mean(y)))
cat("JKT=",JKT,"\n")
```

```
      JKG<-t(y-ytopi)%*%(y-ytopi)
      cat("JKG=",JKG,"\n")
      RK<-1-(JKG/JKT)
      cat("R-square=",RK,"\n")
      error<-y-ytopi
      ER<-matrix(0,M,2)
      ER[,1]<-error[1:M]
      c<-rep(0,(M+1))
      c[1]<-0
      for(i in 1:n)
      {
            c[i+1]<-sum(jpmn[1:i])
            ER[(c[i]+1):c[i+1],2]<-rep(i,jpmn[i])

      }
      cat("error \t\t subyek\n")
      print(ER)
      H=shapiro.test(error)
      j=as.numeric(H[2])
      if(j<alfa)
      {
            cat("Data tidak berdistribusi normal karena ",j," <
",alfa,"\n")
      }
      else
      {
            cat("Data berdistribusi normal karena ",j," >
",alfa,"\n")
      }

#selang kepercayaan
      sigmatopi=(t(y-ytopi)%*%(y-ytopi))/(M-(1+p+k))
      ss=C%*%mp(t(C)%*%W%*%C)%*%t(C)
      Sij=sqrt(sigmatopi*diag(ss))
      yatas=ytopi+(qt(1-alfa/2,M-(1+p+k))*Sij)
      ybawah=ytopi-(qt(1-alfa/2,M-(1+p+k))*Sij)
      SKP=matrix(0,M,5)
      SKP[,1]<-y[1:M]
      SKP[,2]<-ytopi[1:M]
      SKP[,3]<-ybawah[1:M]
      SKP[,4]<-yatas[1:M]
      b=rep(0,M)
      b[1]=0
      for(i in 1:n)
      {
            b[i+1]<-sum(jpmn[1:i])
            SKP[(b[i]+1):b[i+1],5]<-rep(i,jpmn[i])

      }
      cat("====================================================\n
")
      cat("\tY    YTOPI  BTS BAWAH  BTS ATAS  subyek\n")
      cat("====================================================\n
")
      print(SKP)
      t=data[,1]
```

```
      st=sort(t)
      y=data[,2]
      sy=y[order(t)]
      sytopi=ytopi[order(t)]
      sybawah=ybawah[order(t)]
      syatas=yatas[order(t)]
      win.graph()
      plot(st,sy,xlab="Usia",ylab="berat badan")
      lines(st,sytopi,type="l",col="red",lwd=3)
      lines(st,sybawah,type="l",col="black",lwd=2)
      lines(st,syatas,type="l",col="black",lwd=2)
      title(main="PLOT USIA TERHADAP BERAT BADAN")
      pdy=length(data[,2])
      sm=0
      sk=0
      for(i in 1:pdy)
      {
            if(ybawah[i]<=data[i,2]&&data[i,2]<=yatas[i])sm=sm+1
            else sk=sk+1
      }
      print(sm)
      print(sk)
      cp=(sm/pdy)*100
      cat("GCV=",GCV,"\n")
      cat("R-square=",RK,"\n")
      cat("Coverage probability : ",cp,"%\n")

}
```

Lampiran 3. Output Estimasi Selang Kepercayaan Kurva Pertumbuhan Berat Badan Balita Laki-laki pada Orde 2 dan Titik Knot 6

```
Inputkan banyak subyek : 8
Input orde : 2
Input jumlah knot optimum : 1
Input titik knot optimum  ke- 1 = 6
alfa : 0.05
balita ke-      jumlah pengamatan
1                10
2                10
3                7
4                7
5                13
6                13
7                17
8                18
            [,1]
[1,]  3.47650015
[2,]  1.03130670
[3,] -0.07151120
[4,]  0.07060711
MSE= 0.2890983
GCV= 0.3150721
JKT= 363.0016
JKG= 27.46433
R-square= 0.924341
error            subyek
             [,1] [,2]
 [1,] -0.09650015    1
 [2,]  0.54693126    1
 [3,] -0.04525363    1
 [4,]  0.21006288    1
 [5,]  0.13779467    1
 [6,] -0.30131717    1
 [7,] -0.19642355    1
 [8,]  0.06218473    1
 [9,] -0.52924602    1
[10,]  0.08588772    1
[11,]  0.62349985    2
[12,]  0.27318056    2
[13,] -0.15754773    2
[14,]  0.71006288    2
[15,]  0.56733466    2
[16,]  0.69868283    2
[17,] -0.23319628    2
[18,]  0.56874923    2
[19,]  0.37075398    2
[20,]  0.42741676    2
[21,]  0.22349985    3
[22,] -0.26816082    3
[23,]  0.26874923    3
[24,] -0.03781527    3
[25,]  0.81161477    3
[26,]  0.17075398    3
[27,]  0.68588772    3
```

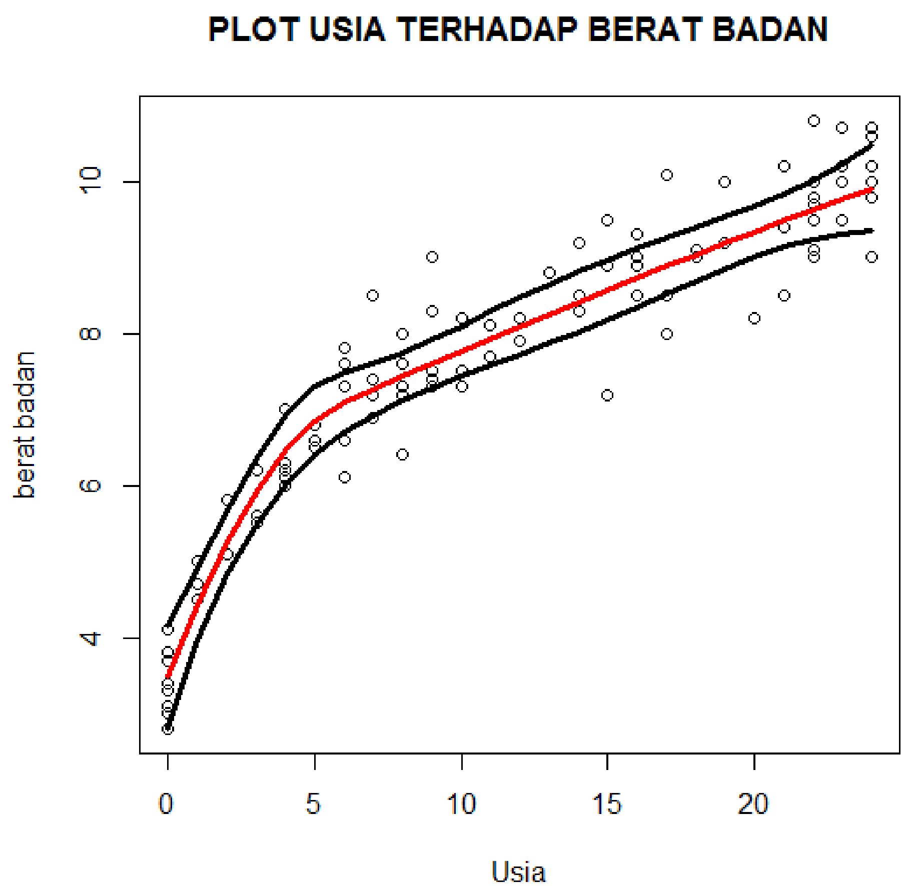
```
[28,]  0.32349985    4
[29,]  0.51006288    4
[30,]  1.23779467    4
[31,]  1.39868283    4
[32,]  0.71589939    4
[33,]  0.22741676    4
[34,]  0.28588772    4
[35,] -0.17650015    5
[36,] -0.25754773    5
[37,] -0.24525363    5
[38,] -0.06220533    5
[39,] -0.23266534    5
[40,] -0.10131717    5
[41,] -0.26816082    5
[42,]  0.08254647    5
[43,] -0.23125077    5
[44,] -0.38543711    5
[45,] -0.12924602    5
[46,] -0.27258324    5
[47,] -0.11411228    5
[48,] -0.37650015    6
[49,]  0.26370436    6
[50,] -0.32681944    6
[51,] -0.35754773    6
[52,] -0.98993712    6
[53,] -1.03266534    6
[54,] -0.11745353    6
[55,] -1.37525625    6
[56,] -0.88543711    6
[57,] -1.13714702    6
[58,] -0.98410061    6
[59,] -0.62924602    6
[60,] -0.91411228    6
[61,] -0.47650015    7
[62,]  0.56370436    7
[63,]  0.54245227    7
[64,]  0.21006288    7
[65,] -0.06220533    7
[66,]  0.16733466    7
[67,]  0.43183918    7
[68,]  0.16680372    7
[69,]  0.10357645    7
[70,]  0.08254647    7
[71,]  0.32474375    7
[72,]  0.16874923    7
[73,] -0.03781527    7
[74,]  0.01161477    7
[75,] -0.08410061    7
[76,]  0.07075398    7
[77,]  0.08588772    7
[78,] -0.67650015    8
[79,]  0.06370436    8
[80,] -0.15306874    8
[81,] -0.42681944    8
[82,] -0.45754773    8
[83,] -0.34525363    8
```

```
[84,] -0.48993712    8
[85,] -0.36220533    8
[86,] -0.13266534    8
[87,] -0.20131717    8
[88,] -0.46816082    8
[89,]  0.54215737    8
[90,]  0.78254647    8
[91,]  0.92474375    8
[92,]  1.21456289    8
[93,]  1.17075398    8
[94,]  0.92741676    8
[95,]  0.78588772    8
Data berdistribusi normal karena  0.5823524  >  0.05
Warning in sigmatopi * diag(ss) :
  Recycling array of length 1 in array-vector arithmetic is
deprecated.
  Use c() or as.vector() instead.

=====================================================
        Y    YTOPI  BTS BAWAH  BTS ATAS  subyek
=====================================================
       [,1]     [,2]     [,3]      [,4] [,5]
 [1,]  3.38 3.476500 2.792231  4.160769    1
 [2,]  5.80 5.253069 4.841523  5.664615    1
 [3,]  6.80 6.845254 6.395841  7.294666    1
 [4,]  7.30 7.089937 6.694294  7.485580    1
 [5,]  7.40 7.262205 6.923639  7.600772    1
 [6,]  7.30 7.601317 7.287826  7.914808    1
 [7,]  7.90 8.096424 7.724280  8.468567    1
 [8,]  9.10 9.037815 8.684314  9.391316    1
 [9,]  9.10 9.629246 9.238455 10.020037    1
[10,] 10.00 9.914112 9.352136 10.476089    1
[11,]  4.10 3.476500 2.792231  4.160769    2
[12,]  6.20 5.926819 5.494595  6.359044    2
[13,]  6.30 6.457548 5.998925  6.916170    2
[14,]  7.80 7.089937 6.694294  7.485580    2
[15,]  8.00 7.432665 7.119743  7.745587    2
[16,]  8.30 7.601317 7.287826  7.914808    2
[17,]  7.70 7.933196 7.581518  8.284874    2
[18,]  9.30 8.731251 8.345439  9.117062    2
[19,] 10.00 9.629246 9.238455 10.020037    2
[20,] 10.20 9.772583 9.309695 10.235471    2
[21,]  3.70 3.476500 2.792231  4.160769    3
[22,]  7.50 7.768161 7.438360  8.097962    3
[23,]  9.00 8.731251 8.345439  9.117062    3
[24,]  9.00 9.037815 8.684314  9.391316    3
[25,] 10.00 9.188385 8.850337  9.526434    3
[26,]  9.80 9.629246 9.238455 10.020037    3
[27,] 10.60 9.914112 9.352136 10.476089    3
[28,]  3.80 3.476500 2.792231  4.160769    4
[29,]  7.60 7.089937 6.694294  7.485580    4
[30,]  8.50 7.262205 6.923639  7.600772    4
[31,]  9.00 7.601317 7.287826  7.914808    4
[32,] 10.20 9.484101 9.135980  9.832221    4
[33,] 10.00 9.772583 9.309695 10.235471    4
[34,] 10.20 9.914112 9.352136 10.476089    4
```

[35,]	3.30	3.476500	2.792231	4.160769	5
[36,]	6.20	6.457548	5.998925	6.916170	5
[37,]	6.60	6.845254	6.395841	7.294666	5
[38,]	7.20	7.262205	6.923639	7.600772	5
[39,]	7.20	7.432665	7.119743	7.745587	5
[40,]	7.50	7.601317	7.287826	7.914808	5
[41,]	7.50	7.768161	7.438360	8.097962	5
[42,]	8.50	8.417454	8.022704	8.812203	5
[43,]	8.50	8.731251	8.345439	9.117062	5
[44,]	8.50	8.885437	8.514219	9.256655	5
[45,]	9.50	9.629246	9.238455	10.020037	5
[46,]	9.50	9.772583	9.309695	10.235471	5
[47,]	9.80	9.914112	9.352136	10.476089	5
[48,]	3.10	3.476500	2.792231	4.160769	6
[49,]	4.70	4.436296	3.958082	4.914509	6
[50,]	5.60	5.926819	5.494595	6.359044	6
[51,]	6.10	6.457548	5.998925	6.916170	6
[52,]	6.10	7.089937	6.694294	7.485580	6
[53,]	6.40	7.432665	7.119743	7.745587	6
[54,]	8.30	8.417454	8.022704	8.812203	6
[55,]	7.20	8.575256	8.181088	8.969424	6
[56,]	8.00	8.885437	8.514219	9.256655	6
[57,]	8.20	9.337147	9.004134	9.670160	6
[58,]	8.50	9.484101	9.135980	9.832221	6
[59,]	9.00	9.629246	9.238455	10.020037	6
[60,]	9.00	9.914112	9.352136	10.476089	6
[61,]	3.00	3.476500	2.792231	4.160769	7
[62,]	5.00	4.436296	3.958082	4.914509	7
[63,]	7.00	6.457548	5.998925	6.916170	7
[64,]	7.30	7.089937	6.694294	7.485580	7
[65,]	7.20	7.262205	6.923639	7.600772	7
[66,]	7.60	7.432665	7.119743	7.745587	7
[67,]	8.20	7.768161	7.438360	8.097962	7
[68,]	8.10	7.933196	7.581518	8.284874	7
[69,]	8.20	8.096424	7.724280	8.468567	7
[70,]	8.50	8.417454	8.022704	8.812203	7
[71,]	8.90	8.575256	8.181088	8.969424	7
[72,]	8.90	8.731251	8.345439	9.117062	7
[73,]	9.00	9.037815	8.684314	9.391316	7
[74,]	9.20	9.188385	8.850337	9.526434	7
[75,]	9.40	9.484101	9.135980	9.832221	7
[76,]	9.70	9.629246	9.238455	10.020037	7
[77,]	10.00	9.914112	9.352136	10.476089	7
[78,]	2.80	3.476500	2.792231	4.160769	8
[79,]	4.50	4.436296	3.958082	4.914509	8
[80,]	5.10	5.253069	4.841523	5.664615	8
[81,]	5.50	5.926819	5.494595	6.359044	8
[82,]	6.00	6.457548	5.998925	6.916170	8
[83,]	6.50	6.845254	6.395841	7.294666	8
[84,]	6.60	7.089937	6.694294	7.485580	8
[85,]	6.90	7.262205	6.923639	7.600772	8
[86,]	7.30	7.432665	7.119743	7.745587	8
[87,]	7.40	7.601317	7.287826	7.914808	8
[88,]	7.30	7.768161	7.438360	8.097962	8
[89,]	8.80	8.257843	7.870666	8.645019	8
[90,]	9.20	8.417454	8.022704	8.812203	8

```
[91,]  9.50 8.575256 8.181088  8.969424    8
[92,] 10.10 8.885437 8.514219  9.256655    8
[93,] 10.80 9.629246 9.238455 10.020037    8
[94,] 10.70 9.772583 9.309695 10.235471    8
[95,] 10.70 9.914112 9.352136 10.476089    8
[1] 61
[1] 34
GCV= 0.3150721
R-square= 0.924341
Coverage probability :  64.21053 %
```



Lampiran 4. Output Estimasi Selang Kepercayaan Kurva Pertumbuhan Berat Badan Balita Perempuan pada Orde 2 dan Titik Knot 6, 12

```
Inputkan banyak subyek : 8
Input orde : 2
Input jumlah knot optimum : 2
Input titik knot optimum  ke- 1 = 6
Input titik knot optimum  ke- 2 = 12
alfa : 0.05
balita ke-      jumlah pengamatan
1                8
2                8
3                13
4                7
5                6
6                9
7                22
8                12
            [,1]
[1,]  3.24354121
[2,]  1.09230673
[3,] -0.08730318
[4,]  0.10519494
[5,] -0.02185297
MSE= 0.2345537
GCV= 0.2647892
JKT= 373.2058
JKG= 19.93707
R-square= 0.9465789
error            subyek
             [,1] [,2]
 [1,]  0.25645879    1
 [2,]  0.75145524    1
 [3,]  0.76526719    1
 [4,]  0.77750455    1
 [5,]  0.94553277    1
 [6,]  0.67800891    1
 [7,]  0.01980187    1
 [8,]  0.09139263    1
 [9,] -0.24354121    2
[10,]  1.06526719    2
[11,]  0.77750455    2
[12,]  1.08297239    2
[13,]  0.77800891    2
[14,]  0.83559848    2
[15,]  1.05965484    2
[16,]  0.89139263    2
[17,] -0.24354121    3
[18,] -0.04854476    3
[19,] -0.27894197    3
[20,] -0.53473281    3
[21,] -0.51591731    3
[22,] -0.52249545    3
[23,] -0.45446723    3
[24,] -0.31702761    3
[25,] -0.04949896    3
```

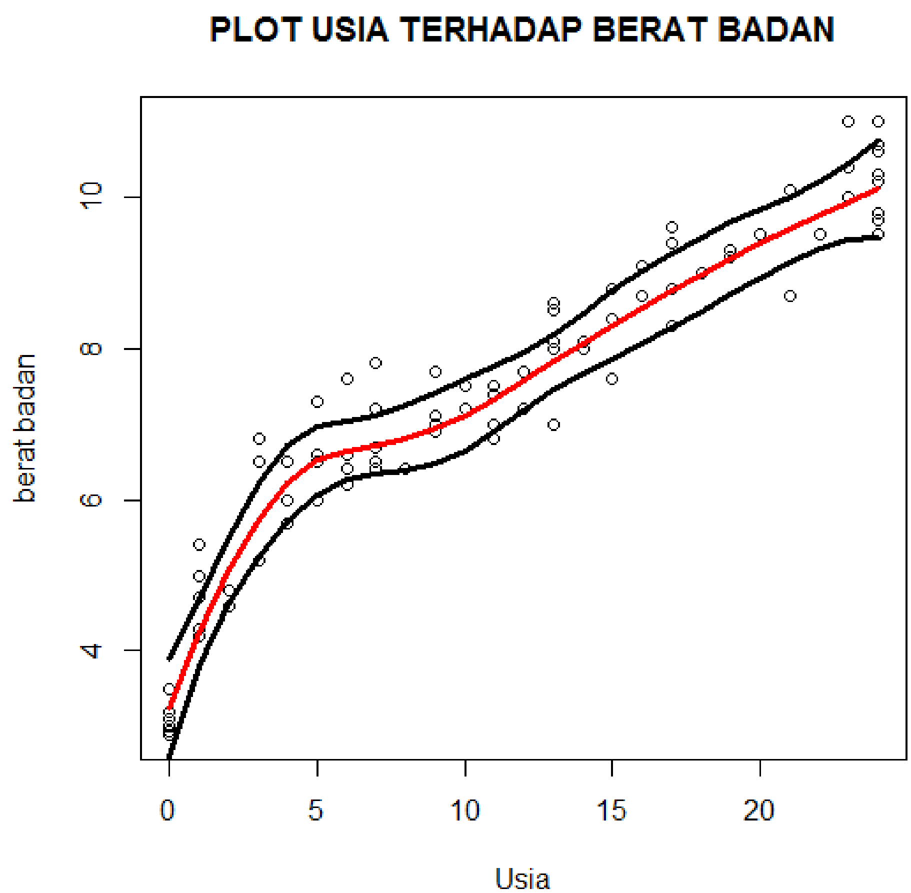
```
[26,] -0.32510444    3
[27,] -0.36658248    3
[28,] -0.06947730    3
[29,] -0.40860737    3
[30,] -0.31354121    4
[31,]  0.75050104    4
[32,]  0.07489556    4
[33,]  0.49095889    4
[34,]  0.63559848    4
[35,]  0.51994645    4
[36,]  0.19139263    4
[37,] -0.14354121    5
[38,]  0.48297239    5
[39,]  0.17800891    5
[40,]  0.55931748    5
[41,]  0.11192766    5
[42,]  0.59139263    5
[43,] -0.04354121    6
[44,]  1.15145524    6
[45,]  0.28408269    6
[46,] -0.02249545    6
[47,]  0.15050104    6
[48,]  0.38059007    6
[49,]  0.11197586    6
[50,]  0.45965484    6
[51,]  0.49139263    6
[52,] -0.34354121    7
[53,]  0.05145524    7
[54,] -0.47894197    7
[55,] -0.21591731    7
[56,]  0.07750455    7
[57,] -0.05446723    7
[58,] -0.01702761    7
[59,] -0.41537151    7
[60,]  0.05050104    7
[61,]  0.08059007    7
[62,]  0.17489556    7
[63,]  0.13341752    7
[64,]  0.27800891    7
[65,]  0.03052270    7
[66,]  0.09095889    7
[67,]  0.15931748    7
[68,]  0.03559848    7
[69,]  0.01192766    7
[70,]  0.11197586    7
[71,] -0.26416056    7
[72,]  0.05965484    7
[73,] -0.30860737    7
[74,] -0.04354121    8
[75,]  0.45145524    8
[76,] -0.21591731    8
[77,] -0.25446723    8
[78,] -0.21702761    8
[79,]  0.38059007    8
[80,] -0.52510444    8
[81,] -0.82199109    8
```

```
[82,] -0.70904111    8
[83,] -0.46440152    8
[84,] -0.88005355    8
[85,] -0.60860737    8
Data berdistribusi normal karena  0.1011474  >  0.05
Warning in sigmatopi * diag(ss) :
  Recycling array of length 1 in array-vector arithmetic is
deprecated.
  Use c() or as.vector() instead.

=====================================================
        Y    YTOPI  BTS BAWAH  BTS ATAS  subyek
=====================================================
       [,1]      [,2]     [,3]      [,4] [,5]
 [1,]  3.50  3.243541 2.600137  3.886946    1
 [2,]  5.00  4.248545 3.807964  4.689125    1
 [3,]  6.50  5.734733 5.254834  6.214632    1
 [4,]  7.30  6.522495 6.065855  6.979136    1
 [5,]  7.60  6.654467 6.268276  7.040659    1
 [6,]  8.50  7.821991 7.450979  8.193004    1
 [7,]  9.00  8.980198 8.484849  9.475547    1
 [8,] 10.20 10.108607 9.469348 10.747867    1
 [9,]  3.00  3.243541 2.600137  3.886946    2
[10,]  6.80  5.734733 5.254834  6.214632    2
[11,]  7.30  6.522495 6.065855  6.979136    2
[12,]  7.80  6.717028 6.330916  7.103139    2
[13,]  8.60  7.821991 7.450979  8.193004    2
[14,]  9.60  8.764402 8.267539  9.261264    2
[15,] 11.00  9.940345 9.432216 10.448475    2
[16,] 11.00 10.108607 9.469348 10.747867    2
[17,]  3.00  3.243541 2.600137  3.886946    3
[18,]  4.20  4.248545 3.807964  4.689125    3
[19,]  4.80  5.078942 4.651105  5.506779    3
[20,]  5.20  5.734733 5.254834  6.214632    3
[21,]  5.70  6.215917 5.718386  6.713449    3
[22,]  6.00  6.522495 6.065855  6.979136    3
[23,]  6.20  6.654467 6.268276  7.040659    3
[24,]  6.40  6.717028 6.330916  7.103139    3
[25,]  6.90  6.949499 6.479267  7.419731    3
[26,]  7.00  7.325104 6.891556  7.758653    3
[27,]  7.20  7.566582 7.184962  7.948203    3
[28,]  8.00  8.069477 7.666156  8.472799    3
[29,]  9.70 10.108607 9.469348 10.747867    3
[30,]  2.93  3.243541 2.600137  3.886946    4
[31,]  7.70  6.949499 6.479267  7.419731    4
[32,]  7.40  7.325104 6.891556  7.758653    4
[33,]  8.80  8.309041 7.863207  8.754876    4
[34,]  9.40  8.764402 8.267539  9.261264    4
[35,] 10.10  9.580054 9.149923 10.010184    4
[36,] 10.30 10.108607 9.469348 10.747867    4
[37,]  3.10  3.243541 2.600137  3.886946    5
[38,]  7.20  6.717028 6.330916  7.103139    5
[39,]  8.00  7.821991 7.450979  8.193004    5
[40,]  9.10  8.540683 8.060997  9.020368    5
[41,]  9.30  9.188072 8.710642  9.665503    5
[42,] 10.70 10.108607 9.469348 10.747867    5
```

```
[43,]  3.20  3.243541 2.600137  3.886946    6
[44,]  5.40  4.248545 3.807964  4.689125    6
[45,]  6.50  6.215917 5.718386  6.713449    6
[46,]  6.50  6.522495 6.065855  6.979136    6
[47,]  7.10  6.949499 6.479267  7.419731    6
[48,]  7.50  7.119410 6.649411  7.589409    6
[49,]  9.50  9.388024 8.937476  9.838573    6
[50,] 10.40  9.940345 9.432216 10.448475    6
[51,] 10.60 10.108607 9.469348 10.747867    6
[52,]  2.90  3.243541 2.600137  3.886946    7
[53,]  4.30  4.248545 3.807964  4.689125    7
[54,]  4.60  5.078942 4.651105  5.506779    7
[55,]  6.00  6.215917 5.718386  6.713449    7
[56,]  6.60  6.522495 6.065855  6.979136    7
[57,]  6.60  6.654467 6.268276  7.040659    7
[58,]  6.70  6.717028 6.330916  7.103139    7
[59,]  6.40  6.815372 6.380525  7.250218    7
[60,]  7.00  6.949499 6.479267  7.419731    7
[61,]  7.20  7.119410 6.649411  7.589409    7
[62,]  7.50  7.325104 6.891556  7.758653    7
[63,]  7.70  7.566582 7.184962  7.948203    7
[64,]  8.10  7.821991 7.450979  8.193004    7
[65,]  8.10  8.069477 7.666156  8.472799    7
[66,]  8.40  8.309041 7.863207  8.754876    7
[67,]  8.70  8.540683 8.060997  9.020368    7
[68,]  8.80  8.764402 8.267539  9.261264    7
[69,]  9.20  9.188072 8.710642  9.665503    7
[70,]  9.50  9.388024 8.937476  9.838573    7
[71,]  9.50  9.764161 9.323091 10.205230    7
[72,] 10.00  9.940345 9.432216 10.448475    7
[73,]  9.80 10.108607 9.469348 10.747867    7
[74,]  3.20  3.243541 2.600137  3.886946    8
[75,]  4.70  4.248545 3.807964  4.689125    8
[76,]  6.00  6.215917 5.718386  6.713449    8
[77,]  6.40  6.654467 6.268276  7.040659    8
[78,]  6.50  6.717028 6.330916  7.103139    8
[79,]  7.50  7.119410 6.649411  7.589409    8
[80,]  6.80  7.325104 6.891556  7.758653    8
[81,]  7.00  7.821991 7.450979  8.193004    8
[82,]  7.60  8.309041 7.863207  8.754876    8
[83,]  8.30  8.764402 8.267539  9.261264    8
[84,]  8.70  9.580054 9.149923 10.010184    8
[85,]  9.50 10.108607 9.469348 10.747867    8

[1] 56
[1] 29
GCV= 0.2647892
R-square= 0.9465789
Coverage probability :  65.88235 %
```


PLOT USIA TERHADAP BERAT BADAN
berat badan
4
6
8
10
0
5
10
15
20
Usia