ESTIMASI INTERVAL KEPERCAYAAN PARAMETER REGRESI DALAM MODEL *PENALIZED KUANTIL SPLINE* PADA PASIEN DEMAM BERDARAH *DENGUE*

ESTIMATION OF CONFIDENCE INTERVAL REGRESSION PARAMETERS IN THE PENALIZED QUANTILE SPLINE MODEL IN DENGUE HEMORRHAGIC FEVER PATIENT

HADIJAH

PROGRAM STUDI MAGISTER STATISTIKA

SEKOLAH PASCASARJANA

UNIVERSITAS HASANUDDIN

MAKASSAR

2022

ESTIMASI INTERVAL KEPERCAYAAN PARAMETER REGRESI DALAM MODEL *PENALIZED KUANTIL SPLINE* PADA PASIEN DEMAM BERDARAH *DENGUE*

Tesis

sebagai salah satu syarat untuk mencapai gelar magister

Program Studi Magister Statistika

Disusun dan diajukan oleh

HADIJAH

H062192005

kepada

PROGRAM STUDI MAGISTER STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

2022

TESIS

ESTIMASI INTERVAL KEPERCAYAAN PARAMETER REGRESI DALAM MODEL *PENALIZED KUANTIL SPLINE* PADA PASIEN DEMAM BERDARAH *DENGUE*

HADIJAH
H062192005

Telah dipertahankan di hadapan Panitia Ujian yang dibentuk dalam rangka Penyelesaian Program Studi Magister Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin pada tanggal 30 Maret 2022
dan dinyatakan telah memenuhi syarat kelulusan

Menyetujui,

Pembimbing Utama,

Dr. Anna Islamiyati, S.Si., M.Si
NIP. 19770808 200501 2 002

Pembimbing Pendamping

Dr. Nurtiti Sunusi, S.Si, M.Si
NIP. 1972017 199703 2 002

Ketua Program Studi
Magister Statistika,

Dr. Dr. Georgina M. Tinungki, M.Si
NIP. 19620926 198702 2 001

Dekan Fakultas Matematika dan
Ilmu Pengetahuan Alam
Universitas Hasanuddin

KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI UNIVERSITAS HASANUDDIN FAKULTAS MIPA

Dr. Eng. Amiruddin, M.Si
NIP. 19720515 199702 1 002

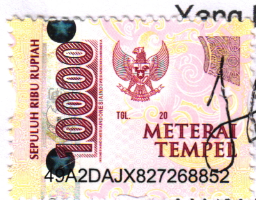
PERNYATAAN KEASLIAN TESIS
DAN PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa, tesis berjudul "Estimasi Interval Kepercayaan Parameter Regresi Dalam Model *Penalized Kuantil Spline* Pada Pasien Demam Berdarah *Dengue*" adalah benar karya saya dengan arahan dari komisi pembimbing (Dr. Anna Islamiyati, S.Si., M.Si sebagai Pembimbing Utama dan Dr. Nurtiti Sunusi, S.Si, M.Si sebagai Pembimbing Pendamping). Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka tesis ini. Sebagian dari isi tesis ini telah dipublikasikan di Jurnal (International Journal of Academic and Applied Research (IJAAR), ISSN: 2643-9603, Vol. 5 Issue 11, November – 2021, Pages 5-8) sebagai artikel dengan judul "*Estimation of Quantile Nonparametric Regression Model with Linear Penalized Spline*".

Dengan ini saya melimpahkan hak cipta dari karya tulis saya berupa tesis ini kepada Universitas Hasanuddin.

Makassar, 30 Maret 2022

Yang Menyatakan,



HADIJAH

NIM. H062192005

KATA PENGANTAR

Puji syukur atas kehadirat Tuhan Yang Maha Esa atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyusun dan menyelesaikan tesis ini. Penulis menyadari sepenuhnya bahwa apa yang dikemukakan dalam tesis ini masih jauh dari kesempurnaan yang merupakan sebagai akibat dari kererbatasan kemampuan serta berbagai kesulitan yang penulis hadapi dalam penyusunan tesis ini.

Keberhasilan penulis dalam Menyusun tesis ini tidak lepas dari kemudahan yang diberikan oleh Allah SWT dan bantuan dari berbagai pihak yang bersifat moril maupun materil. Oleh karena itu, penulis memanjatkan doa kepada Tuhan Yang Maha Esa agar memberikan rahmat-Nya kepada pihak yang banyak membantu dalam penyelesaian tesis ini. Penulis juga percaya bahwa tesis ini dapat selesai bukan hanya kekuatan pikiran penulis semata akan tetapi karena bantuan dari berbagai pihak juga, baik selama proses perkuliahan bahkan sampai proses pengerjaan tesis di Program Magister Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Hasanuddin. Namun demikian, penulis dengan senang hati menerima kritik dan saran yang bersifat membangun dari pembaca kerya tulis ini demi sempurnanya tesis ini.

Terima kasih yang tak terhingga kepada kedua orang tua tercinta, almarhum Ayah tercinta dan Ibu tercinta, suami dan anakku tercinta, mertua tercinta dan saudara-saudaraku atas doa yang tak pernah putus, dukungan serta segala kebaikan mereka yang sampai kapanpun takkan pernah bisa terbalaskan atas kasih sayang yang tiada henti dalam penyelesaian tesis ini. Selanjutnya saya ingin menyampaikan juga rasa hormat dan terima kasih kepada :

1. Ibu Prof. Dr. Dwia Aries Tina Palubuhu, MA. selaku Rektor Universitas Hasanuddin.
2. Bapak Dr. Eng. Amiruddin selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam beserta seluruh jajarannya.
3. Ibu Dr. Nurtiti Sunusi, S.Si., M.Si. selaku Ketua Departemen Statistika, dan Pembimbing Pertama penulis yang senantiasa memberikan ilmu, dukungan, dan motivasi serta kemudahan kepada penulis dalam berbagai hal selama menjalani Pendidikan di Departemen Statistika.

4. Ibu Dr. Dr. Georgina Maria Tinungki, M.Si. selaku Ketua Program Studi Magister Statistika Universitas Hasanuddin yang juga menjadi salah satu tim penguji tesis.
5. Ibu Dr. Anna Islamiyati, S.Si., M.Si. selaku Pembimbing Utama yang telah bersabar dan bersedia meluangkan banyak waktunya untuk membimbing penulis dan memberikan masukan dalam penyelesaian tesis ini.
6. Ibu Dr. Erna Tri Herdiani, S.Si., M.Si. dan Bapak Dr. Nirwan, M.Si. selaku penguji penulis yang telah bersedia memberikan masukan-masukan dan arahan dalam penyusunan tesis ini.
7. Bapak dan Ibu Dosen di Departemen Statistika Fakultas MIPA Universitas Hasanuddin, yang dengan tulus ikhlas memberikan ilmu pengetahuan dan pengalaman yang dimilikinya selama perkuliahan berlangsung sehingga memberikan banyak manfaat bagi penulis untuk saat ini maupun di masa mendatang.
8. Teman-teman Mahasiswa Program Magsiter Statistika Angkatan 1, terima kasih atas nasehat dan dukungan luar biasa kepada penulis.
9. Sahabat=sahabatku semuanya, terima kasih banyak untuk semuanya, semoga silaturahmi kita tetap berlanjut.
10. Serta semua pihak yang telah membantu dalam penyelesaian skripsi ini yang tidak dapat penulis sebutkan satu per satu.

Akhir kata, penulis menyampaikan terima kasih kepada semua pihak yang telah banyak membantu semoga Tuhan Yang Maha Esa memberikan balasan yang berlipat ganda. Penulis juga berharap semoga tesis ini memberikan manfaat bagi pengembangan ilmu pengetahuan dan menambah referensi di bidang Stastistika serta berguna bagi masyarakat.

ABSTRAK

HADIJAH. Estimasi Interval Kepercayaan Parameter Regresi Dalam Model *Penalized Kuantil Spline* Pada Pasien Demam Berdarah *Dengue* (dibimbing oleh Anna Islamiyati and Nurtiti Sunusi**).**

Regresi nonparametrik merupakan salah satu pendekatan regresi yang digunakan ketika data yang bentuk fungsinya tidak mengikuti pola parametrik. Masalah lain yang sering ditemukan pada data riil adalah pencilan pada data. Salah satu metode regresi yang dikembangkan oleh beberapa peneliti untuk mengatasi data yang mengandung pencilan adalah regresi kuantil. Pada artikel ini, kami melakukan estimasi interval kepercayaan pada model regresi *penalized kuantil spline* dalam memperkirakan model regresi kuantil nonparametrik. Metode estimasi yang digunakan adalah least absolute deviation (LAD). Keunggulan *penalized kuantil spline* terletak pada kriteria estimasinya yaitu melibatkan fungsi *goodness of fit* dan fungsi penalti. Fungsi penalti yang memuat parameter penghalus bersama titik knot mengontrol kemulusan kurva regresi. Model diaplikasikan pada data pasien penderita Demam Berdarah *Dengue* (DBD) di Kota Makassar. Faktor yang dianalisis adalah jumlah trombosit, suhu badan dan sel darah putih pasien. Model optimal adalah model dengan kriteria nilai *generalized cross validation* (GCV) minimum. Berdasarkan nilai GCV minimum, secara simultan pada kuantil $\theta = 0{,}25$ diperoleh model optimal pada $\lambda = 1$. Pada kuantil $\theta = 0{,}50$ diperoleh model optimal pada $\lambda = 1{,}5$. Pada kuantil $\theta = 0{,}75$ diperoleh model optimal pada $\lambda = 1{,}5$. Estimasi interval kepercayaan yang terbentuk menunjukkan bahwa semua faktor dalam hal ini suhu badan dan sel darah putih signifikan dalam mengukur trombosit penderita DBD.

Kata Kunci: Regresi Nonparametrik, *Penalized Kuantil Spline,* Least Absolute Deviation, Interval Kepercayaan, Demam Berdarah *Dengue*

ABSTRACT

HADIJAH. Estimation Of Confidence Interval Regression Parameters In The Penalized Quantile Spline Model In Dengue Hemorrhagic Fever Patient (supervised by Anna Islamiyati and Nurtiti Sunusi**).**

Nonparametric regression is a regression approach that is used when the shape data does not follow a parametric pattern. Another problem that is often found in real data is outliers in the data. One of the regression methods developed by several researchers to overcome data containing outliers is quantile regression. In this article, we estimate the confidence interval in the penalized quantile spline regression model in estimating the nonparametric quantile regression model. The estimation method used is the least absolute deviation (LAD). The advantage of the penalized quantile spline lies in its estimation, which involves the goodness of fit function and the penalty function. A penalty function containing the smoothing parameter along with the smoothness control point of the regression curve. The model was applied to the data of patients with Dengue Hemorrhagic Fever (DHF) in Makassar City. Factors analyzed were the number of platelets, body temperature and white blood cells of the patient. The optimal model is a model with a minimum generalized cross validation (GCV) value criterion. Based on the minimum GCV value simultaneously at quantile = 0.25, the optimal model is obtained at $\lambda = 1$. At quantile $\theta = 0.50$, the optimal model is obtained at $\lambda = 1.5$. At quantile $\theta = 0.75$, the optimal model is obtained at $\lambda = 1.5$. The estimation of the confidence interval formed shows that all factors in terms of body temperature and white blood cells are significant in measuring the platelets of patients with DHF.

Keywords: Nonparametric Regression, Penalized Quantile Spline, Least Absolute Deviation, Confidence Interval, Dengue Hemorrhagic Fever

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR LAMPIRAN

BAB I
PENDAHULUAN

1.1 Latar Belakang

Analisis regresi merupakan alat dalam menilai secara spesifik pola hubungan dan pengaruh antar variabel. Metode analisis tersebut memiliki tujuan akhir untuk memperkirakan atau memprediksi nilai dari suatu variabel berdasarkan variabel lain (Jain dkk, 2016). Informasi mengenai hubungan fungsional antara variabel prediktor dengan variabel respon dapat diperkirakan dengan melihat bentuk pola hubungan pada diagram pencar (scatter plot). Apabila plot data mengikuti suatu bentuk pola parametrik misalnya linier, kuadratik atau kubik, maka pendekatan regresi yang digunakan adalah regresi parametrik. Budiantara dkk (2009) telah menjelaskan bahwa bentuk fungsi dapat diketahui bukan hanya dari pola data awal, tapi juga dapat diperoleh dari informasi penelitian sebelumnya. Selanjutnya, data yang bentuk fungsinya tidak mengikuti pola parametrik dianalisis dengan menggunakan pendekatan regresi nonparametrik. Beberapa estimator telah dikembangkan diantaranya *spline* (Budiantara, 2001; Huang, 2003; Lestari dkk, 2010, Islamiyati dkk, 2018), kernel (Aydin, 2007), polinomial lokal (Chamidah dkk, 2012), dan deret fourier (Mardianto, dkk, 2020).

Spline adalah salah satu estimator yang banyak dikembangkan oleh peneliti karena sifat fleksibilitas dari estimator tersebut. Penggunaan titik knot yang mampu menunjukkan pola perubahan pada data menjadi salah satu keunggulan dari estimator *spline* (Budiantara, 2006). Berbagai estimator *spline* yang telah dikembangkan oleh peneliti, salah satunya adalah *penalized spline*. Keunggulan *penalized spline* terletak pada kriteria estimasinya yaitu melibatkan fungsi goodness of fit dan fungsi penalti. Fungsi penalti yang memuat parameter penghalus bersama titik knot mengontrol kemulusan kurva regresi. Beberapa penelitian yang telah mengembangkan *penalized spline* adalah Claeskens dkk (2009) yang bekerja pada kasus uniprediktor, Montoya dkk (2014) yang membuat data simulasi untuk menunjukkan kemampuan estimator, dan Islamiyati dkk (2018) yang bekerja pada kasus birespon untuk data longitudinal. Akan tetapi, penelitian tersebut belum mempertimbangkan adanya pencilan yang dapat saja terjadi pada data.

Data pencilan adalah suatu nilai data ekstrem berbeda dengan yang lainnya. Pada kasus pencilan, beberapa pendekatan yang dapat digunakan, diantaranya M robust (Siswanto, 2017), Least Median Square (Ortiz, 1995), Least Trimmed Square (Karel, 2015), S robust (Zuni dan Endang, 2017), MM robust (Chen, 2002) dan regresi kuantil (Aprilia dkk, 2019). Koenker dan Bassett (1978) memperkenalkan regresi yang berdistribusi kuantil bersyarat dari variabel respon yang dinyatakan sebagai fungsi kovariat. Koenker dkk (1994) mengembangkan kuantil *spline* smoothing, Rossi dan Harvey (2009) mengembangkan kuantil *spline* kubik, dan Koenker dkk (2017) mengembangkan kuantil *penalized*. Akan tetapi, penelitian tersebut hanya mempertimbangkan estimasi titik untuk parameternya tanpa mempertimbangkan estimasi interval kepercayaan. Estimasi titik mendekati suatu nilai parameter dengan menggunakan satu nilai atau satu titik, sehingga menghasilkan suatu nilai tunggal sebagai estimasi parameter terbaik. Adapun estimasi interval mendekati suatu nilai parameter dengan menggunakan dua titik yang memungkinkan untuk mengukur derajat kepercayaan terhadap ketelitian dari estimasi (Ma'unnah, 2016). Penelitian tentang estimasi parameter telah dilakukan oleh Syaranamual (2011) yang mengestimasi interval konfidensi *spline* kuadratik. Intansari (2016) meneliti tentang estimasi titik dan interval konfidensi dalam kurva regresi *spline* kuadratik dan Sindi (2017) mengukur interval konfidensi pada data indeks pembangunan gender di Jawa Timur dengan estimator *spline*.

Penelitian ini menggunakan estimasi *Least Absolute Deviation* (LAD). LAD merupakan alternatif least squares untuk analisis statistik model regresi linier yang meminimumkan jumlah absolut error (Chen dkk, 2008). Penelitian tentang estimasi LAD telah dilakukan oleh Hananingrum dkk (2021) yang meneliti data jumlah kematian bayi di Jawa Timur, Setyono dkk (2018) meneliti tentang data tanaman cabai merah di Jawa Barat dan Amyad dkk (2018) meneliti tentang Indeks Pembangunan Manusia di Indonesia. Berdasarkan penelitian-penelitian tersebut, penulis mengkaji tentang estimasi interval kepercayaan pada model regresi *penalized kuantil spline* dengan metode estimasi Selanjutnya, model diaplikasikan pada data pasien penderita Demam Berdarah Dengue (DBD) di Kota Makassar (Puteri dkk, 2020). Penyakit DBD adalah salah satu penyakit menular yang dapat menimbulkan kematian dalam waktu singkat dan sering menimbulkan wabah pada suatu wilayah. Srichaikul dkk (1989) meneliti pasien penderita DBD dan menemukan bahwa trombositopenia muncul pada hari ketiga dan tetap bertahan selama perjalanan penyakit. Penurunan jumlah trombosit pada umumnya terjadi sebelum ada peningkatan hematokrit dan terjadi sebelum suhu tubuh turun. Namun selain jumlah trombosit dan suhu tubuh, diperlukan pemeriksaan medis lainnya, misalnya pemeriksaan kandungan sel darah putih (Lontoh dkk, 2016). Oleh sebab itu, dalam penelitian ini, faktor yang dianalisis dalam model *penalized kuantil spline* adalah jumlah trombosit, suhu badan dan sel darah putih pasien.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana estimasi interval kepercayaan parameter regresi dalam model *penalized kuantil spline*?
2. Bagaimana model *penalized kuantil spline* pada data pasien DBD dan estimasi interval kepercayaan dari parameter regresi yang diperoleh?

1.3 Batasan Masalah

Beberapa batasan masalah yang dibuat dalam penelitian ini adalah:

1. Kuantil yang digunakan terdiri dari $\boldsymbol{\theta}$ = 0,25; 0,50; 0,75.
2. Orde yang digunakan dalam memodelkan data dibatasi pada orde linear.
3. Metode pemilihan titik knot dan parameter penghalus optimal menggunakan kriteria *generalized cross validation* (GCV) minimum.
4. Data yang digunakan adalah data DBD di Kota Makassar dengan mempertimbangkan faktor suhu badan dan sel darah putih.

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Mendapatkan estimator interval kepercayaan parameter regresi dalam model *penalized kuantil spline*.
2. Mendapatkan model data pasien DBD dengan *penalized kuantil spline* dan mengestimasi interval kepercayaan dari parameter regresi yang diperoleh.

1.5 Manfaat Penelitian

Hasil penelitian ini diharapkan dapat memberi manfaat sebagai berikut:

1. Menambah wawasan keilmuan dan pengetahuan mengenai model pola perubahan trombosit pada pasien penderita DBD berdasarkan suhu badan dan jumlah sel darah putih melalui *penalized kuantil spline* dan estimasi interval kepercayaan dari parameter regresi yang diperoleh.

2. Memberikan informasi akurat dan lebih terperinci mengenai pola perubahan trombosit pada pasien penderita DBD berdasarkan suhu badan dan jumlah sel darah putih, sehingga dapat menjadi acuan bagi akademisi maupun praktisi dalam penanganan pasien penderita DBD.

BAB II
TINJAUAN PUSTAKA

2.1 Regresi Nonparametrik

Regresi nonparametrik merupakan analisis regresi yang digunakan untuk mengetahui hubungan antara variabel respon dengan variabel prediktor yang tidak diketahui fungsinya, hanya diasumsikan smooth (mulus). Kurva regresi hanya diasumsikan smooth (mulus) dalam arti termuat dalam suatu ruang fungsi tertentu. Regresi nonparametrik merupakan regresi yang sangat fleksibel dalam memodelkan pola data (Eubank, 1989).

Jika diberikan pasangan data (x_i, y_i) dengan $i = 1,2,3, \dots, n,$ dan pola hubungan antar variabel respon dan variabel prediktor tidak diketahui bentuknya, maka dapat digunakan pendekatan nonparametrik. Model regresi nonparamterik secara umum adalah sebagai berikut:

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

dengan y_i merupakan variabel respon pada pengamatan ke- i, $f(x_i)$ merupakan persamaan kurva regresi dengan x_i sebagai variabel prediktor pada pengamatan ke- i, sedangkan ε_i merupakan error yang berdistribusi normal, independent dengan mean 0 dan variansi σ^2 (Eubank, 1999).

2.2 Estimasi Parameter Regresi Kuantil

Estimasi parameter dalam regresi dengan metode *Ordinary Least Square* (OLS), hanya dapat digunakan untuk memberi solusi pada masalah *mean,* sehingga Koenker dan Basset (1978) mengembangkan metode alternatif, yaitu regresi kuantil. Regresi dengan metode OLS diestimasi dengan meminimumkan jumlah kuantil *error,* sedangkan regresi kuantil akan meminimumkan jumlah absolut *error* yang lebih dikenal dengan *least absolute deviation (*LAD).

Pada regresi kuantil, *error* diberi bobot yang berbeda. Bobot yang digunakan, yaitu θ untuk nilai *error* yang lebih besar atau sama dengan nol, dan 1- θ untuk *error* nyang kurang dari nol. Perkalian antara *error* dengan bobot yang diberikan disebut sebagai fungsi *loss* (ρ_θ), yaitu:

$$\rho_\theta = \sum_{i=1,\varepsilon_i \geq 0}^{n} \theta|\varepsilon_i| + \sum_{i=1,\varepsilon_i < 0}^{n} (1-\theta)|\varepsilon_i| \tag{2.2}$$

dengan demikian, dalam regresi kuantil terdapat fungsi kuantil ke-θ dari variabel y dengan syarat x yang mempertimbangkan penduga $\hat{\beta}(\theta)$, sehingga diperoleh solusi untuk permasalahan tersebut yang dinyatakan sebagai berikut:

$$\min_{\beta \epsilon R^{\rho}} \sum_{i=1}^{n} \rho_{\theta}(\varepsilon) = \min_{\beta \epsilon R^{\rho}} \sum_{i=1}^{n} \rho_{\theta}\left(y_i - Q_{y|x}(\theta)\right) \tag{2.3}$$

dengan:

$\rho_{\theta}(\varepsilon)$: fungsi *loss*

θ : indeks kuantil dengan $\theta \epsilon (0,1)$

$Q_{y|x}(\theta)$: fungsi kuantil ke- θ dari variabel y dengan syarat x

Fungsi untuk kuantil bersyarat $Q_{y|x}(\theta)$ didefinisikan sebagai berikut:

$$Q_{y|x}(\theta) = \boldsymbol{X^T \beta(\theta)} \tag{2.4}$$

Dalam regresi kuantil, pada kuantil ke-θ dari F_y meminimumkan fungsi *loss* dari persamaan (2.3) adalah

$$\hat{\beta}(\theta) = \min_{\beta \epsilon R^{\rho}} \sum_{i=1}^{n} \rho_{\theta}(\varepsilon) = \min_{\beta \epsilon R^{\rho}} \sum_{i=1}^{n} \rho_{\theta}\left(\boldsymbol{y}_i - \boldsymbol{X^T \beta}(\theta)\right) \tag{2.5}$$

dimana $\rho_{\theta}(\varepsilon)$ pada persamaan (2.5) didefinisikan

$$\rho_{\theta}(\varepsilon) = \begin{cases} \theta_{\varepsilon}, \varepsilon \geq 0 \\ (1-\theta)_{\varepsilon}, \varepsilon < 0 \end{cases}$$

Selanjutnya, akan dibuktikan bahwa fungsi *loss* berbentuk asimetris dengan penjelasan sebagai berikut:

$$\begin{aligned} \rho_{\theta} &= \left[\theta I(\varepsilon \geq 0) + (1-\theta) I\left((\varepsilon < 0)\right)\right]|\varepsilon| \\ &= [\theta - I(\varepsilon < 0)]\, \varepsilon \end{aligned} \tag{2.6}$$

dengan

$$I(\varepsilon \geq 0) = \begin{cases} 1, \varepsilon \geq 0 \\ 0, \varepsilon < 0 \end{cases}$$

ε : residual dari estimasi parameter

$I(\varepsilon)$: fungsi indikator yang terdefinisikan

Sehingga dapat dibuktikan:

$$\rho_{\theta}(\varepsilon) = \begin{cases} \theta_{\varepsilon}, \varepsilon \geq 0 \\ (1-\theta)_{\varepsilon}, \varepsilon < 0 \end{cases}$$

a. Untuk $\varepsilon \geq 0$

$$\begin{aligned} \rho_{\theta} &= [\theta I(\varepsilon \geq 0) + (1-\theta) I(\varepsilon < 0)]|\varepsilon| \\ &= \left[\theta I(\varepsilon \geq 0) + (1-\theta) I\left((\varepsilon < 0)\right)\right] \varepsilon \\ &= [\theta I + (1-\theta) I(\varepsilon < 0)]\, \varepsilon \end{aligned}$$

$= [\theta + I(\varepsilon < 0) - \theta I(\varepsilon < 0)]\ \varepsilon$

$= \left[\theta + \left(1 - I(\varepsilon \geq 0)\right) - \theta\left(1 - I(\varepsilon \geq 0)\right)\right]\ \varepsilon$

$= [\theta + (1 - 1) - \theta(1 - 1)]\ \varepsilon$

$= \theta\varepsilon$

atau

$\rho_\theta = [\theta - I(\varepsilon < 0)]|\varepsilon|$

$= \left[\theta - \left(1 - I(\varepsilon \geq 0)\right)\right]\ \varepsilon$

$= [\theta - (1 - 1)]\ \varepsilon$

$= \theta\varepsilon$

b. Untuk $\varepsilon < 0$

$\rho_\theta = [\theta I(\varepsilon \geq 0) + (1 - \theta)I(\varepsilon < 0)]|\varepsilon|$

$= \left[\theta I(\varepsilon \geq 0) + (1 - \theta)I\left((\varepsilon < 0)\right)\right]\ (-\varepsilon)$

$= [\theta 0 + (1 - \theta)I(\varepsilon < 0)]\ (-\varepsilon)$

$= [(\theta - 1)I(\varepsilon < 0)]\ \varepsilon$

$= \left[(\theta - 1)\left(1 - I(\varepsilon < 0)\right)\right]\ \varepsilon$

$= [(\theta - 1)(1 - 0)]\ \varepsilon$

$= (\theta - 1)\ \varepsilon$

atau

$\rho_\theta = [\theta - I(\varepsilon < 0)]\varepsilon$

$= \left[\theta - \left(1 - I(\varepsilon \geq 0)\right)\right]\varepsilon$

$= [\theta - (1 - 0)]\varepsilon$

$= (\theta - 1)\ \varepsilon$

Sehingga menjadi,

$$\rho_\theta = [\theta I(\varepsilon \geq 0) + (1 - \theta)I(\varepsilon < 0)]|\varepsilon| = [\theta - I(\varepsilon < 0)]\ \varepsilon, \forall_\varepsilon$$

Jika y merupakan fungsi x yang diketahui dan memiliki fungsi probabilitas $F_{y|x}(y)$ maka kuantil ke-θ dari fungsi tersebut dapat dituliskan seperti pada persamaan berikut:

$$\min_\beta \theta \int_{i=1:\varepsilon_i \geq 0}^{n} |\varepsilon_i| dF_y(y) + (1 - \theta) \int_{i=1:\varepsilon_i < 0}^{n} |\varepsilon_i| dF_y(y) \qquad (2.7)$$

dengan mempertimbangkan $\hat{\beta}(\theta)$, sehingga diperoleh solusi untuk permasalahan tersebut yang dinyatakan

$$\hat{\beta}(\theta) = \min_{\beta \epsilon R^{\rho}} \left\{ \theta \sum_{i=1:\varepsilon_i \geq 0}^{n} |\boldsymbol{y}_i - \boldsymbol{X}^T \boldsymbol{\beta}(\theta)| + (1-\theta) \sum_{i=1:\varepsilon_i < 0}^{n} |\boldsymbol{y}_i - \boldsymbol{X}^T \boldsymbol{\beta}(\theta)| \right\} \quad (2.8)$$

(Koenker, 2005).

Solusi dari persamaan (2.8) tidak dapat diperoleh secara analitik, tetapi secara numerik. Salah satu metode numerik yang digunakan adalah algoritma simpleks.

2.2.1 Optimasi dengan Algoritma Simpleks pada Model Regresi Kuantil

Algoritma simpleks adalah salah satu metode pendugaan parameter secara numerik untuk regresi kuantil yang telah dikembangkan oleh Barrodale dan Robert pada tahun 1974. Algoritma simpleks memberikan solusi permasalahan program linier yang melibatkan beberapa variabel keputusan dengan bantuan komputasi (Davino dkk, 2014). Berikut ini akan diberikan proses algoritma simpleks, yaitu:

1. Menyajikan data ke dalam bentuk tabel awal, yaitu dengan mendefinisikan data variable x ke dalam bentuk *spline* $(\boldsymbol{X}[K])$, yaitu dengan penambahan titik knot pada data variabel x.

2. Menentukan nilai θ yang digunakan, yaitu $\theta =$ 0,25; 0,50; 0,75 dan 0,95.

3. Mengubah terlebih dahulu masalah optimal linear ke bentuk standar, fungsi tujuan dan kendala-kendala diubah ke dalam bentuk persamaan.

   Minimalkan : $\theta \sum_{i=1}^{n} |\varepsilon_{i1}(\theta)| + (1-\theta) \sum_{i=1}^{n} |\varepsilon_{i2}(\theta)|$

   Dengan kendala : $\boldsymbol{X}\hat{\boldsymbol{\beta}}(\theta) + \varepsilon_1(\theta) - \varepsilon_2(\theta) = \boldsymbol{y}, \varepsilon_1, \varepsilon_2 \geq 0$

   Bentuk tersebut diubah ke dalam bentuk standar sesuai dengan data yang disajikan.

4. Persamaan fungsi tujuan dan kendala dimasukkan ke dalam bentuk tabel awal simpleks.

5. Menentukan kolom kunci (variabel masuk), yaitu untuk masalah maksimum memilih $c_j - z_j$ yang terbesar, sedangkan untuk masalah minimum memilih $c_j - z_j$ yang terkecil.

6. Menentukan baris kunci (variabel keluar), yaitu dari nilai rasio antara nilai ruas kiri b_i dengan koefisien kolom kunci (a_{ij}), pilih yang terkecil (untuk masalah minimum atau maksimum). Rasio $= \frac{b_i}{a_{ij}}$, dimana rasio > 0.

7. Menentukan pivot dari pepotongan antara kolom kunci dan baris kunci yang dinamakan elemen kunci atau elemen penentu iterasi algoritma simpleks dan akan diubah nilainya menjadi 1.
8. Selanjutnya, melakukan operasi dasar (OBD) berdasarkan pivot untuk baris lainnya, termasuk baris $c_j - z_j$ dengan nilai elemen-elemen yang termasuk di dalam kunci dijadikan nol (selain elemen yang dijadikan pivot).
9. Proses iterasi untuk masalah maksimum berhenti jika semua nilai pada baris $c_j - z_j \leq 0$, berarti solusi sudah optimal. Apabila masih ada $c_j - z_j > 0$ (positif), maka iterasi algoritma simpleks masih berlanjut. Untuk masalah minimum berhenti jika semua nilai pada baris $c_j - z_j \geq 0$. Apabila masih ada $c_j - z_j < 0$ (negatif), maka iterasi algoritma simpleks masih berlanjut.

2.3 Fungsi *Spline*

Spline merupakan salah satu model regresi nonparametrik yang terdiri dari potongan-potongan polynomial yang memiliki sifat tersegmen pada titik knot. Titik knot merupakan titik perpaduan bersama yang memperlihatkan terjadinya perubahan perilaku dari fungsi *spline* pada interval-interval yang berbeda (Eubank, 1999). Secara umum, fungsi *spline* berorde q dengan titik knot $K_1, K_2, \dots, K_r$ dapat ditulis sebagai berikut:

$$f(x_i) = \sum_{l=0}^{q} \beta_l x_i^l + \sum_{h=1}^{r} \beta_{q+h}(x_i - K_h)_+^q \tag{2.9}$$

Dengan fungsi *truncated*

$$(x_i - K_h)_+^q = \begin{cases} (x_i - K_h)^q, jika\ x \geq K_h \\ 0, jika\ x < K_h \end{cases} \tag{2.10}$$

Dengan:

$f(x_i)$	: fungsi regresi
x_i	: variabel prediktor
$\beta_0, \beta_1, \dots, \beta_{q+r}$	: parameter regresi
K_h	: titik knot ke- h, $(h = 1,2, \dots, r)$
$(x_i - K_h)_+^q$	: fungsi polynomial *truncated*

Bentuk matriks $f(x_i)$ pada persamaan (2.9) untuk n sampel berpasangan dapat dinyatakan sebagai berikut:

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

Dengan uraian sebagai berikut:

$$\boldsymbol{f} = \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{q+h} \end{bmatrix}$$

$$\boldsymbol{X} = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^q & (x_1 - K_1)_+^q & \cdots & (x_1 - K_r)_+^q \\ 1 & x_2 & x_2^2 & \cdots & x_2^q & (x_2 - K_1)_+^q & \cdots & (x_2 - K_r)_+^q \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^q & (x_n - K_1)_+^q & \cdots & (x_n - K_r)_+^q \end{bmatrix}$$

2.4 *Penalized Kuantil Spline*

Regresi kuantil pertama kali diperkenalkan oleh Koenker dan Bassett pada tahun 1978. Regresi kuantil adalah suatu pendekatan dalam analisis regresi yang digunakan untuk menduga hubungan antara variabel respon dengan variabel prediktor pada fungsi kuantil bersyarat tertentu. Dengan pendekatan ini, fungsi kuantil dapat diduga dari sebaran bersyarat variabel respon pada setiap nilai kuantil sesuai dengan kuantil yang diinginkan. Regresi kuantil sangat berguna jika data tidak homogen (heterogenous) dan tidak berbentuk standar, seperti tidak simetris atau tidak berdistribusi normal, terdapat ekor pada sebaran atau *penalized* distribution. Keuntungan dari regresi kuantil, yaitu dapat meminimumkan pengaruh dari pencilan (Koenker dan Bassett, 1978). Regresi kuantil meminimumkan jumlah error mutlak terboboti dan menduga model dengan menggunakan fungsi kuantil bersyarat pada suatu sebaran data (Buhai, 2014). Persamaan umum regresi kuantil linier untuk kuantil bersyarat $Q_{y|x}(\theta)$ dari variable respon y_i, yaitu:

$$y_i = \beta_0(\theta) + \beta_1(\theta)x_{i1} + \cdots + \beta_k(\theta)x_{ik} + \varepsilon_i(\theta) \tag{2.11}$$

Apabila hubungan antara variabel respon dan prediktor dinyatakan dalam fungsi f yang tidak diketahui bentuknya dan dapat dihampiri dengan model regresi kuantil nonparametrik, maka hubungan dapat dinyatakan dalam bentuk persamaan regresi sebagai berikut:

$$y_i(\theta) = f(x_{i1}, x_{i2}, \dots, x_{ik}) + \varepsilon_i(\theta) \tag{2.12}$$

Jika $f(x_{i1}, x_{i2}, \dots, x_{ik})$ diasumsikan bersifat aditif dan dihampiri dengan fungsi regresi kuantil dengan *Penalized Spline* yang merupakan salah satu penduga yang digunakan pada regresi nonparametrik dalam penaksiran fungsi regresi nonparametrik. Penaksir *Penalized Spline* melibatkan titik simpul dan pemulusan parameter secara bersamaan dalam mengendalikan kelancaran kurva. Penaksir *Penalized Spline* diperoleh dengan meminimukan fungsi *Penalized Least Square* (PLS) yang merupakan kriteria estimasi dengan menggabungkan fungsi *goodness* *of fit* dengan *penalty.* Fungsi *goodness of fit* yang dimaksud adalah *Least Square* $\sum_{i=1}^{n}\left(y_i - f(x_i)\right)$ dan fungsi *penalty* adalah fungsi yang mengontrol kemulusan kurva yaitu $\sum_{l=1}^{k}\beta_{p+l}^{2}$ maka diperoleh model regresi kuantil dengan *Penalized Spline* sebagai berikut:

$$\begin{aligned} y_i(\theta) &= f(x_{i1}) + f(x_{i2}) + \cdots + f(x_{ik}) + \varepsilon_i(\theta) \\ &= \textstyle\sum_{j=1}^{k} f\left(x_{ij}\right) + \varepsilon_i(\theta) \end{aligned} \tag{2.13}$$

Berdasarkan persamaan (2.13) fungsi f dihampiri dengan fungsi regresi kuantil dengan *penalized spline* dengan orde q dan jumlah knot sebanyak r sebagai berikut:

$$f\left(x_{ij}\right) = \sum_{l=0}^{q} \beta_{ij}(\theta)x_{ij}^{l} + \sum_{h=1}^{r} \beta_{(q+h)j}(\theta)\left(x_{ij} - K_{hj}\right)_{+}^{q} \tag{2.14}$$

dengan

$$\left(x_{ij} - k_{hj}\right)_{+}^{q} = \begin{cases} \left(x_{ij} - K_{hj}\right)^{q}, jika\ x_{ij} \geq K_{hj} \\ \quad 0 \ \ , jika\ x_{ij} < K_{hj} \end{cases}$$

$y_i(\theta)$	: respon kuantil ke- θ pada pengamatan ke-i
$f\left(x_{ij}\right)$	: fungsi regresi kuantil dengan *spline* dari pengamatan ke- i pada prediktor ke-j
x_{ij}	: pengamatan ke- i pada prediktor ke-j
$\beta_{ij}(\theta)$	: parameter polinomial kuantil ke- θ pada orde ke l dan prediktor ke-j
$\beta_{(q+h)j}(\theta)$	: parameter *penalized* kuantil ke- θ pada titik knot ke-$(q+h)$ dan prediktor ke-j
K_{hj}	: nilai titik knot ke-h pada predikor ke-j
r	: banyaknya titik knot
q	: orde polynomial *spline penalized*
k	: banyaknya prediktor
ε_i	: *error* kuantil ke- θ pada persamaan ke- i

Uraian fungsi pada persamaan (2.14) dinyatakan sebagai berikut:

$$\sum_{j=1}^{k} f\left(x_{ij}\right) = f(x_{i1}) + f(x_{i2}) + \cdots + f(x_{ik}) \tag{2.15}$$

Dengan menguraikan fungsi f dan memisahkan antara parameter dan variabel, model regresi kuantil *penalized spline* dapat dinyatakan ke dalam bentuk matriks sebagai berikut:

$$\boldsymbol{y}(\theta) = \boldsymbol{X}[K]\boldsymbol{\beta}(\theta) + \boldsymbol{\varepsilon}(\theta) \tag{2.16}$$

dengan:

$\boldsymbol{y}(\theta) = [y_1(\theta), y_2(\theta), \ldots, y_n(\theta)]^T$ merupakan vektor kolom berukuran $n \times 1$ dari variabel respon y pada kuantil ke- θ;

$\boldsymbol{X}[K] = [\mathbf{1} \quad \boldsymbol{X_1} \quad \cdots \quad \boldsymbol{X_k}]$ merupakan matriks $\boldsymbol{X}$ dalam bentuk *spline* dengan orde q dan r knot berukuran $n \times (k+1)$ dengan n observasi pada k variabel x; $\boldsymbol{\beta}(\theta)$ merupakan vector kolom berukuran $(k+1) \times 1$ dari parameter β pada kuantil ke- θ; dan $\boldsymbol{\varepsilon}(\theta)$ merupakan vektor kolom berukuran $n \times 1$ dari *error* (ε) pada kuantil ke- θ.

Estimator *penalized spline* melalui kriteria *penalized* least square (PLS) yang terbentuk dari fungsi *penalized* adalah sebagai berikut:

$$\text{PLS} = \left(\boldsymbol{y}(\theta) - \boldsymbol{X}[K]\boldsymbol{\beta}(\theta)\right)^T\left(\boldsymbol{y}(\theta) - \boldsymbol{X}[K]\boldsymbol{\beta}(\theta)\right) + \lambda\boldsymbol{\beta}(\theta)^T\boldsymbol{D}\boldsymbol{\beta}(\theta) \tag{2.17}$$

2.5 *Least Absolute Deviation* di dalam Regresi Kuantil

Regresi dengan metode OLS diestimasi dengan meminimumkan jumlah kuadrat error, sedangkan regresi kuantil akan meminimumkan jumlah absolut error yang lebih dikenal dengan least absolute deviation (LAD). Pada regresi kuantil, error diberi bobot yang berbeda. Bobot yang digunakan, yaitu θ untuk nilai error yang lebih besar atau sama dengan nol, dan $1 - \theta$ untuk error yang kurang dari nol. Perkalian antara error dengan bobot yang diberikan disebut sebagai fungsi loss (ρ_θ) yaitu:

$$\rho_\theta(\varepsilon) = \sum_{i=1,\varepsilon_{i\geq 0}}^{n} \theta|\varepsilon_i| + \sum_{i=1,\varepsilon_i<0}^{n} (1-\theta)|\varepsilon_i| \tag{2.18}$$

dengan demikian, dalam regresi kuantil terdapat fungsi kuantil ke− θ dari variabel y dengan syarat x yang mempertimbangkan penduga $\beta(\theta)$, sehingga diperoleh solusi untuk permasalahan tersebut yang dinyatakan sebagai berikut:

$$\min_{\beta \epsilon R^p} \sum_{i=1}^{n} \rho_\theta(\varepsilon) = \min_{\beta \epsilon R^p} \sum_{i=1}^{n} \rho_\theta\left(y_i - Q_{y|x}(\theta)\right) \tag{2.19}$$

dengan:

$\rho_\theta(\varepsilon)$: fungsi loss

θ : indeks kuantil dengan $\theta \in (0,1)$

$Q_{y|x}(\theta)$: fungsi kuantil dari variable dengan syarat x

Dimana $\rho_\theta(\varepsilon)$ pada persamaan (2.19) didefinisikan

$$\rho_\theta(\varepsilon) = \begin{cases} \theta\varepsilon, & \varepsilon \geq 0 \\ (1-\theta)\varepsilon, & \varepsilon < 0 \end{cases}$$

Fungsi untuk kuantil bersyarat $Q_{y|x}(\theta)$ didefinisikan sebagai berikut:

$$Q_{y|x}(\theta) = \boldsymbol{X^T\beta}(\theta) \tag{2.20}$$

Jika y merupakan fungsi x yang diketahui dan memiliki fungsi probabilitas $F_{y|x}(y)$, maka kuantil ke− θ dari fungsi tersebut dapat dituliskan seperti pada persamaan berikut:

$$\min_\beta \theta = \int_{i=1;\varepsilon_i \geq 0}^{n} |\varepsilon_i| dF_y(y) + (1-\theta) \int_{i=1;\varepsilon_i < 0}^{n} |\varepsilon_i| dF_y(y) \tag{2.21}$$

dengan mempertimbangkan $\hat{\beta}(\theta)$, sehingga diperoleh solusi untuk permasalahan tersebut yang dinyatakan:

$$\hat{\beta}(\theta) = \min_{\beta \epsilon R^p} \left\{ \theta \sum_{i=1;\varepsilon_i \geq 0}^{n} \left| y_i - \boldsymbol{X^T\beta}(\theta) \right| + (1-\theta) \sum_{i=1;\varepsilon_i < 0}^{n} \left| y_i - \boldsymbol{X^T\beta}(\theta) \right| \right\} \tag{2.22}$$

2.6 Pemilihan Titik Knot Optimal

Salah satu langkah penting dalam pendekatan *spline* adalah memilih titik knot yang optimal. Titik knot merupakan perpaduan dua kurva yang menunjukkan pola perubahan perilaku kurva pada selang yang berbeda. Titik knot optimal dibutuhkan untuk mendapatkan model regresi *spline* terbaik yang paling sesuai dengan data. Menurut Wahba dan Wang pada 1998 *dalam* Marina (2013) bahwa salah satu metode yang banyak dipakai dan disukai dalam memilih titik knot optimal adalah *Generalized Cross Validation (GCV)* minimum. Model *spline* yang terbaik dengan titik knot optimal didapat dari nilai GCV yang terkecil. Metode GCV dapat dituliskan seperti berikut:

$$GCV\ (K) = \frac{MSE\ (K)}{(n^{-1} trace[\boldsymbol{I} - \boldsymbol{A}(\boldsymbol{K})])^2} \tag{2.23}$$

dengan:

$K = \left[K_{11}, K_{11}, \ldots, K_{1r}, \ldots, K_{p1}, \ldots, K_{pr}\right]$ adalah titik knot.

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

$\boldsymbol{A}[K] = \boldsymbol{X}[K](\boldsymbol{X}^T[K]\boldsymbol{X}[K])^{-1}\boldsymbol{X}^T[K]$

(Budiantara, 2006)

2.7 Interval Kepercayaan

Interval kepercayaan adalah suatu interval tertentu yang memuat parameter dengan probabilitas tertentu. Beberapa model dalam regresi kuantil yang digunakan untuk menghitung selang kepercayaan β salah satunya yaitu fungsi *sparsity* (Chen dkk, 2005). Kelebihan metode ini adalah dapat menjelaskan penyebaran data dengan menggunakan ukuran kuantil sesuai sebaran data. Selain itu, moetode ini dapat menghasilkan hasil paling efisien untuk taksiran interval yang lebih sempit secara komputasi (Zhou dkk, 1996). Pada setiap θ yang sudah ditentukan, taksiran interval dengan formula sebagai berikut (Davino dkk, 2014):

$$P\left(\hat{\beta}(\theta) - t_{(\alpha/2,df)}se\left(\hat{\beta}(\theta)\right) \leq \beta(\theta) \leq \hat{\beta}(\theta) - t_{(\alpha/2,df)}se\left(\hat{\beta}(\theta)\right)\right) \quad (2.24)$$

$$= 1 - \alpha$$

Nilai $se\left(\hat{\beta}(\theta)\right)$ adalah nilai yang diperoleh dari perhitungan matrik $\widehat{w}^2\boldsymbol{D}^{-1}$, dimana $\boldsymbol{D}^{-1} = (X^TX)^{-1}$ dan $\widehat{w}^2(\theta) = \theta(1-\theta)/f\left(F^{-1}(\theta)\right)^2$. Untuk F adalah fungsi distribusi kumulatif $f = F^{-1}$ adalah fungsi densitas, maka model sparsity dinotasikan sebagai berikut:

$$S(\theta) = f\left(F^{-1}(\theta)\right)^{-1} \quad (2.25)$$

2.8 Pencilan

Secara umum, pencilan merupakan data yang tidak mengikuti pola umum pada model atau data yang keluar dari model dan tidak berada dalam daerah selang kepercayaan (Sembiring, 1995:62). Menurut Draper dan smith (1992), adakalanya pencilan memberikan informasi yang tidak dapat diberikan oleh titik data lainnya yang merupakan suatu keganjilan dan menandakan suatu titik data yang sama sekali tidak tipikal dibanding data lainnya.

Berdasarkan Montgomery and Peck (1992), sebagai kaidah umum, pencilan baru ditolak setelah ditelusuri ternyata merupakan akibat dari kesalahan-kesalahan seperti memasukkan ukuran atau analisis yang salah, ketidaktepatan pencatatan data, dan terjadi kerusakan alat pengukuran. Bila ternyata bukan akibat dari kesalahan-kesalahan semacam itu, penyelidikan yang seksama harus dilakukan. Menghapus data tersebut untuk memperbaiki persamaan yang cocok dapat berbahaya. Tindakan tersebut dapat menimbulkan kesalahan ketelitian dalam mengestimasi atau memprediksi.

2.9 Demam Berdarah *Dengue*

Demam Berdarah *Dengue* (DBD) atau *Dengue hemorrhagic fever* (DHF) adalah penyakit infeksi yang disebabkan oleh virus *Dengue* dengan manifestasi klinis demam, nyeri otot dan/atau nyeri sendi yang disertai *leukopenia*, ruam, *limfadenopati*, *trombositopenia* dan *diatesis hemoragik*. Pada DBD terjadi perembesan plasma yang ditandai oleh *hemokonsentrasi* (peningkatan hematokrit) atau penumpukan cairan di rongga tubuh. Virus *Dengue* masuk ke dalam tubuh melalui gigitan nyamuk *Ae. Aegypti* atau *Ae albopictus*. Nyamuk mendapatkan virus *Dengue* setelah menggigit orang yang terinfeksi virus *Dengue*. Organ sasaran dari virus ini adalah organ *hepar, nodus limfaticus*, sumsum tulang serta paru-paru. Virus yang berada di kelenjar liur berkembang biak dalam waktu 8-10 hari (*extrinsic incubation period*) sebelum dapat ditularkan Kembali kepada manusia pada saat gigitan berikutnya. Infeksi pertama kali mungkin memberi gejala *Dengue fever*. DBD baru dapat terjadi bila seseorang setelah terinfeksi *Dengue* pertama kali, kemudian mendapat infeksi berulang akibat virus *Dengue* lainnya, yang disebut dengan *the secondary heterologous infection* atau *the sequential infection hypothesis* (Suhendro dkk.,2009).

DBD dapat berkembang menjadi *Sindrom Syok Dengue* (SSD) yang berakibat fatal bila tidak ditangani secara baik dan tepat. Diagnosis penyakit DBD dan pemantauan perjalanan penyakit harus dilakukan secara tepat dan akurat. Pada awal penyakit DBD didapatkan *leukopenia*, namun kemudian dapat normal dengan dominasi dari sel *neutrophil*. Mendekati fase akhir penyakit akan terjadi penurunan jumlah total *leukosit* bersamaan dengan penurunan sel *polimorfonuklear*. Ini akan tampak tanda limfositosis pada akhir fase demam dan pada awal terjadinya syok terdapat monositosis, sehingga didapatkan *diff count* leukosit terjadi pergeseran ke kanan (*shift to the right*). Leukosit adalah sel darah yang mengandung inti, disebut juga sel darah putih. Leukosit merupakan komponen darah yang berperan dalam memerangi infeksi yang disebabkan oleh virus, bakteri, ataupun proses maetabolik toksin. Terdapat lima jenis leukosit, yang masing-masing memiliki fungsi khusus dalam melawan pathogen. Sel-sel itu adalah *neutrophil, eosinophil, basophil, monosit*, dan *kimfosit*. Hasil hitung jenis leukosit membantu menegakkan diagnosis, memberikan informasi yang lebih spesifik mengenai infeksi dan proses penyakit (Purwanto, 2002).

Kriteria laboratorium WHO tentang jumlah trombosit yang rendah (*trombositopenia*) dan kebocoran plasma yang ditandai dengan hemokonsentrasi merupakan indicator keparahan penyakit demam berdarag *Dengue*. Trombositopenia dikarenakan oleh suspense sumsung tulang, sedangkan mekanisme yang menginduksi destruksi perifer atau peningkatan penggunaan trombosit lebih penting dan memainkan peran utama dalam induksi trombositopenia pada DBD. Penurunan jumlah trombosit pada umumnya terjadi sebelum ada peningkatan hematokrit dan terjadi sebelum suhu turun. Dikatakan trombositopenia jika jumlah trombosit di bawah 10^3 sel/ul, biasanya dapat dijumpai antara hari sakit ketiga sampai ketujuh (Suhendro dkk., 2009).