



TESIS – SS 142501

**PERBANDINGAN ESTIMASI PARAMETER REGRESI QUANTIL
DENGAN REGRESI *ROBUST LEAST TRIMMED SQUARE* (LTS)**

**(Studi kasus : Faktor - faktor yang Mempengaruhi Penyebaran Penyakit
Malaria di Indonesia)**

Muhammad Yahya Matdoan

NRP. 1315 201 020

DOSEN PEMBIMBING

Dr. Bambang Widjanarko Otok, M.Si

R. Mohamad Atok, M.Si, Ph.D

PROGRAM PASCASARJANA

JURUSAN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

INSTITUT TEKNOLOGI SEPULUH NOPEMBER

SURABAYA

2017



TESIS – SS 142501

**COMPARISON OF PARAMETER ESTIMATION OF
QUANTILE REGRESSION WITH ROBUST REGRESSION
USING LEAST TRIMMED SQUARE (LTS)**

**(Case Study : Factors that Contribute to Spreading of Malaria
Desease in Indonesia)**

Muhammad Yahya Matdoan

NRP. 1315 201 020

DOSEN PEMBIMBING

Dr. Bambang Widjanarko Otok, M.Si

R. Mohamad Atok, M.Si, Ph.D

PROGRAM PASCASARJANA

JURUSAN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

INSTITUT TEKNOLOGI SEPULUH NOPEMBER

SURABAYA

2017


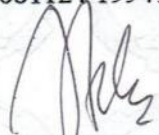

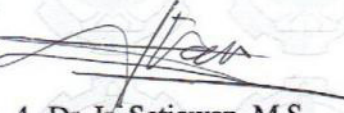
**PERBANDINGAN ESTIMASI PARAMETER REGRESI QUANTIL
DENGAN REGRESI *ROBUST LEAST TRIMMED SQUARE* (LTS)
(Studi Kasus : Faktor-faktor yang Mempengaruhi Penyebaran
Penyakit Malaria di Indonesia)**

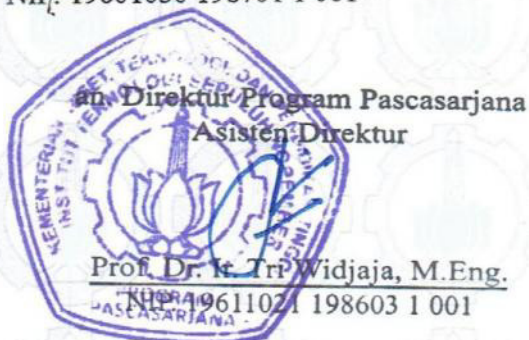
Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar Magister Sains (M.Si)
di
Institut Teknologi Sepuluh Nopember
Oleh :

**MUHAMMAD YAHYA MATDOAN
NRP. 1315 201 020**

Tanggal Ujian : 09 Januari 2017
Periode Wisuda : Maret 2017

Disetujui oleh :

- 
1. Dr. Bambang Widjanarko Otok, M.Si (Pembimbing I)
NIP. 19681124 199412 1 001
- 
2. R. Mohamad Atok, M.Si, Ph.D (Pembimbing II)
NIP. 19710915 199702 1 001
- 
3. Dr. Purnadi, M.Sc (Penguji)
NIP. 19620204 198701 1 001
- 
4. Dr. Ir. Setiawan, M.S (Penguji)
NIP. 19601030 198701 1 001



Direktur Program Pasca Sarjana,

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

Perbandingan Estimasi Parameter Regresi Quantil dengan Regresi Robust *Least Trimmed Square* (LTS)

(Studi Kasus : Faktor-faktor yang Mempengaruhi Penyebaran
Penyakit Malaria di Indonesia)

Nama Mahasiswa : Muhammad Yahya Matdoan
NRP : 1315201020
Pembimbing : Dr. Bambang Widjanarko Otok, M.Si
R. Mohamad Atok, M.Si, Ph.D

ABSTRAK

Estimasi metode OLS didasarkan pada distribusi normal, sehingga kurang tepat untuk menganalisis sejumlah data yang tidak simetris atau mengandung outlier. Oleh karena itu, dikembangkan Regresi Quantil yang tidak terpengaruh oleh adanya pencilan (outliers). Selain Regresi Quantil, Regresi *Robust* LTS juga dapat digunakan sebagai alternatif untuk menyelesaikan permasalahan data yang fluktuatif dan mengandung outlier. Penelitian ini membandingkan Regresi Quantil dengan Regresi *Robust* LTS pada kasus faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia. Diperoleh hasil bahwa dengan menggunakan metode OLS diperoleh R^2 sebesar 0.681, regresi *robust* LTS sebesar 0.732 dan regresi quantil dengan $\tau = 0.95$ sebesar 0.832, sehingga model terbaik yang diperoleh adalah model regresi quantil. Lebih lanjut diperoleh hasil bahwa faktor utama penyebab penyebaran penyakit malaria di Indonesia adalah faktor presentase rumah layak huni, faktor presentase penduduk miskin dan faktor rasio tenaga dokter.

Kata Kunci : Regresi Quantil, Regresi *Robust* LTS, *Outliers*, Malaria

**COMPARISON OF PARAMETER ESTIMATION OF
QUANTILE REGRESSION WITH ROBUST REGRESSION
USING LEAST TRIMMED SQUARE (LTS)
(Case Study : Factors that Contribute to Spreading of
Malaria Disease in Indonesia)**

Name : Muhammad Yahya Matdoan
NRP : 1315201020
Supervisor : Dr. Bambang Widjanarko Otok, M.Si
: R. Mohamad Atok, M.Si, Ph.D

ABSTRACT

OLS estimation method based on the normal distribution, therefore OLS is not suitable method to analyze unsymetric data or contain data outliers. Quantile regression and LTS Robust Regression are alternative methods to solve the problem. Aim of this study is to compare quantile regression and LTS robust regression in case of factors that contribute to spreading of malaria disease in Indonesia. The results obtain R^2 0.681, 0.732, and 0.832 using OLS, LTS, and Quantile Regression ($\tau = 0.95$) methods respectively. Conclusion of this study is quantile regression as the best model. Moreover the main factors that contribute to malaria disease spreading are the percentage of livable housing, the percentage of poor people, and doctors ratio.

Keywords : Quantile Regression, LTS Robust Regression , Outliers, Malaria

KATA PENGANTAR

Alhamdulillah Rabbil Alamin, puji syukur dan sujud kehadirat Allah SWT, karena hanya atas izin dan kuasa-Nya penulis dapat menyelesaikan tesis dengan judul “Perbandingan Estimasi Parameter Regresi Quantil dengan Regresi *Robust Least Trimmed Square* (Studi kasus : faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia)”.

Tesis ini tidak mungkin selesai tanpa bimbingan dan dukungan dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih yang ikhlas dan tulus kepada semua yang telah ikut memberikan dukungan, bantuan, baik moril maupun spiritual dalam menyelesaikan tesis ini. Ucapan terima kasih dan penghargaan ini ditujukan kepada :

1. Bapak. Dr. Bambang Widjanarko Otok M.Si, selaku dosen pembimbing I, atas segala ilmu dan pembelajaran hidupnya, sangat berarti dalam membentuk pribadi yang lebih baik dan tenang.
2. Bapak. R. Mohamad Atok, M.Si, Ph.D, selaku dosen pembimbing II yang telah meluangkan banyak waktu di tengah kesibukannya untuk memberikan bimbingan, dukungan dan perhatian kepada penulis untuk dapat menyelesaikan tesis ini.
3. Bapak. Dr. Purhadi, M.Sc dan Dr. Ir. Setiawan, M.S, selaku penguji I dan Penguji II yang telah banyak memberikan masukan yang sangat berharga bagi penyempurnaan tesis ini.
4. Bapak. Dr. Suhartono M.Sc dan Bapak. Dr. rer, pol. Heri Kuswanto, M.Si, selaku Ketua Jurusan dan Ketua Program Studi S2/S3 Statistika ITS yang telah memberikan arahan dan semangat kepada seluruh mahasiswa.

5. Bapak Ibu Dosen Pengajar serta seluruh jajaran Karyawan-Karyawati Jurusan Statistika F-MIPA ITS Surabaya, atas waktu dan tenaga yang diluangkan untuk penulis selama di Kampus.
6. Orang Tuaku Tercinta yang dengan segala kekuatan dan keterbatasannya mendukungku dari awal hingga akhir.
7. Saudara-saudaraku “Salma Matdoan beserta Adnan Rahanyamtel, Mutmainnah Matdoan beserta Rifai Matdoan, dan Nafisa Matdoan serta Fatima Matdoan yang telah memberikan dukungan dan motivasi sehingga penulis dapat menyelesaikan studi.
8. Keluarga Besar Jurusan Matematika Universitas Pattimurra yang telah memberikan dorongan dan arahan hingga penulis dapat menyelesaikan studi
9. Teman-teman Seperjuangan (Malik, Vongky, Willy, Sham, Agus, Ka Ardy, Ka Mario, Ka Kirek, Pa lucky dan Pa Rey, Tutus, Cinti, Mita beserta teman-teman mahasiswa regular dan BPS angkatan 2015 yang selalu menemani penulis baik suka maupun duka, dalam proses menempuh studi di Kota Surabaya.
10. Semua pihak yang tidak dapat disebutkan satu persatu.

Penulis menyadari bahwa tesis ini masih terdapat kelemahan dan kekurangan, oleh karena itu saran dan kritik membangun akan diterima dengan tangan terbuka. Akhir kata, penulis berharap semoga tesis ini dapat menjadi sebuah sumbangan yang berarti bagi ilmu pengetahuan dan bermanfaat bagi banyak pihak.

Surabaya, Januari 2017

Penulis

Muhammad Yahya Matdoan

DAFTAR ISI

| | |
|---|------|
| LEMBAR PENGESAHAN | i |
| ABSTRAK | iii |
| ABSTRACT | v |
| KATA PENGANTAR | vii |
| DAFTAR ISI | ix |
| DAFTAR TABEL | xi |
| DAFTAR GAMBAR | xiii |
| DAFTAR LAMPIRAN | xv |
| | |
| BAB 1 PENDAHULUAN | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Rumusan Masalah | 5 |
| 1.3 Tujuan Penelitian | 5 |
| 1.4 Manfaat Penelitian | 5 |
| 1.5 Batasan Masalah | 5 |
| | |
| BAB 2 TINJAUAN PUSTAKA | 7 |
| 2.1 Analisis Regresi | 7 |
| 2.2 Pengujian Outlier | 7 |
| 2.3 Metode <i>Ordinary Least Square</i> (OLS) | 8 |
| 2.4 Metode <i>Least Trimmed Square</i> (LTS) | 10 |
| 2.5 Metode Regresi Quantil | 13 |
| 2.6 <i>Koefisien Korelasi</i> | 21 |
| 2.7 Multikolinieritas | 22 |
| 2.8 <i>Koefisien Determinasi</i> | 23 |
| 2.9 Penyakit malaria | 23 |

| | |
|---|----|
| BAB 3 METODOLOGI PENELITIAN..... | 29 |
| 3.1 Sumber Data | 29 |
| 3.2 Variabel Penelitian | 29 |
| 3.3 Struktur Data..... | 32 |
| 3.4 Langkah-langkah Penelitian..... | 32 |
| | |
| BAB 4 HASIL DAN PEMBAHASAN..... | 37 |
| 4.1 Pemodelan Penyebaran Penyakit Malaria di Indonesia..... | 37 |
| 4.1.1 Statistika Deskriptif | 37 |
| 4.1.2 Identifikasi Pola Hubungan dan Deteksi Outlier..... | 44 |
| 4.1.3 Estimasi Parameter dengan Menggunakan OLS | 48 |
| 4.1.4 Estimasi Parameter dengan Menggunakan LTS | 59 |
| 4.1.5 Estimasi Parameter Menggunakan Regresi Quantil.... | 50 |
| 4.2 Pemilihan Model Terbaik | 56 |
| 4.3 Faktor-faktor Penyebab Penyakit Malaria..... | 57 |
| | |
| BAB 5 KESIMPULAN DAN SARAN..... | 61 |
| 5.1 Kesimpulan..... | 61 |
| 5.2 Saran | 62 |
| | |
| DAFTAR PUSTAKA..... | 63 |
| LAMPIRAN | 67 |

DAFTAR TABEL

| Judul Tabel | Halaman |
|--|---------|
| Tabel 3.1 Variabel Penelitian..... | 29 |
| Tabel 3.2 Struktur Data..... | 32 |
| Tabel 4.1 Statistika Deskriptif | 37 |
| Tabel 4.2 Nilai <i>Koefisien Korelasi</i> | 46 |
| Tabel 4.3 Nilai <i>Variance Inflation Factor (VIF)</i> | 47 |
| Tabel 4.4 Pengujian Outlier | 47 |
| Tabel 4.5 Estimasi Parameter dengan OLS | 48 |
| Tabel 4.6 Estimasi Parameter dengan <i>Robust LTS</i> | 49 |
| Tabel 4.7 Estimasi Regresi Kuantil pada kuantil $\tau = 0.05$ | 51 |
| Tabel 4.8 Estimasi Regresi Kuantil pada kuantil $\tau = 0.25$ | 52 |
| Tabel 4.9 Estimasi Regresi Kuantil pada kuantil $\tau = 0.50$ | 53 |
| Tabel 4.10 Estimasi Regresi Kuantil pada kuantil $\tau = 0.75$ | 54 |
| Tabel 4.11 Estimasi Regresi Kuantil pada kuantil $\tau = 0.95$ | 55 |
| Tabel 4.12 Hasil <i>Koefisien Determinasi</i> | 56 |

DAFTAR GAMBAR

| Judul Gambar | Halaman |
|--|---------|
| Gambar 3.1 Diagram Alir Pemodelan..... | 35 |
| Gambar 4.1 Diagram <i>Annual Parasite Incidence</i> Malaria..... | 38 |
| Gambar 4.2 Diagram Presentase Akses Sanitasi Layak | 39 |
| Gambar 4.3 Diagram Presentase Rumah Layak Huni..... | 40 |
| Gambar 4.4 Diagram Presentase Penduduk Miskin..... | 41 |
| Gambar 4.5 Diagram Rasio Tenaga Dokter..... | 42 |
| Gambar 4.6 Grafik Presentase Akses Air Bersih..... | 43 |
| Gambar 4.7 Diagram Rasio Keberadaan Puskesmas..... | 43 |
| Gambar 4.8 <i>Scatter Plot</i> Antara Variabel Respon dengan Variabel Prediktor..... | 45 |

DAFTAR LAMPIRAN

| Judul Lampiran | Halaman |
|---|---------|
| Lampiran 1. Data Penelitian | 67 |
| Lampiran 2. Hasil Pengujian Data Pencilan (<i>Outlier</i>)..... | 79 |
| Lampiran 3. Ouput <i>Koefisien Korelasi</i> dan Nilai <i>Variance Inflation Factor</i> (VIF)... | 71 |
| Lampiran 4. Syntax Metode <i>Ordinary Least Square</i> (OLS) | 72 |
| Lampiran 5. Syntax Metode Regresi <i>Robust Least Trimed Square</i> (LTS)..... | 73 |
| Lampiran 6. Ouput Metode <i>Robust Least Trimed Square</i> (LTS) | 75 |
| Lampiran 7. Syntax Metode Regresi Kuantil | 76 |
| Lampiran 8. Ouput Metode Regresi Kuantil | 81 |
| Lampiran 9. Ouput Plot Regresi Kuantil | 84 |

(Halaman Ini Sengaja Dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis regresi merupakan salah satu metode statistika yang digunakan untuk menggambarkan hubungan antara dua atau lebih variabel, sehingga suatu variabel dapat diprediksikan dari variabel yang lain. Variabel yang ditaksir nilainya disebut variabel respon (y), sedangkan variabel penaksir disebut sebagai variabel prediktor (x). Hal tersebut dapat dilakukan melalui tiga pendekatan yaitu pendekatan parametrik, nonparametrik dan semiparametrik. Pendekatan yang paling umum dan seringkali digunakan yaitu pendekatan parametrik. Pendekatan parametrik mengasumsikan bentuk model sudah ditentukan atau diketahui. Dengan kata lain, bentuk hubungan antara variabel respon dengan variabel prediktor diketahui atau diperkirakan dari bentuk kurvanya. Misalnya, diasumsikan berbentuk pola linear, kuadrat, eksponensial, dan polinomial dengan derajat tertentu. Apabila tidak ada informasi apapun tentang bentuk dari kurva regresi, maka pendekatan yang digunakan yaitu pendekatan nonparametrik, Sedangkan apabila bentuk kurva regresi terdiri dari komponen parametrik dan nonparametrik maka digunakan pendekatan regresi semiparametrik (Subanar dan Budiantara, 1998).

Dalam analisis regresi parametrik (khususnya regresi klasik), Gujarati (2004) menyatakan bahwa ada beberapa asumsi yang harus dipenuhi dalam melakukan estimasi, yaitu tidak terjadi multikolinieritas antar variabel prediktor, tidak terjadi otokorelasi dan heteroskedastisitas, serta residual berdistribusi normal. Jika semua asumsi tersebut terpenuhi, maka hasil estimasi dengan menggunakan metode OLS bersifat *Best Linear Unbiased Estimator* (BLUE). Untuk mengatasi penyimpangan asumsi dalam model regresi linear, dapat dilakukan transformasi terhadap data sehingga diperoleh model regresi yang sesuai dengan data yang telah ditransformasi. Pendekatan klasik untuk mendapatkan nilai dugaan parameter dari model regresi linier adalah metode

Ordinary Least Square (OLS). Penduga parameter dari metode OLS diperoleh dengan meminimumkan jumlah kuadrat sisaan (residual). Namun, metode OLS dianggap kurang tepat untuk menganalisis sejumlah data yang tidak simetris atau mengandung outliers. Masalah ini disebabkan karena outliers dapat menyebabkan hasil estimasi parameter menjadi tidak stabil. Selain itu, analisis regresi dengan menggunakan metode OLS didasarkan pada data yang distribusi di mean, dimana nilai mean menunjukkan ukuran pemusatan dari suatu distribusi data. Oleh sebab itu, pendekatan dengan menggunakan metode OLS hanya mampu menduga model dari fungsi bersyarat mean dan tidak merepresentasikan keseluruhan data. Untuk mengatasi masalah tersebut, maka dikembangkanlah metode analisis regresi kuantil.

Metode regresi kuantil pertama kali diperkenalkan oleh Koenker dan Basset (1978). Metode ini merupakan perluasan dari model regresi pada kuantil bersyarat. Distribusi kuantil bersyarat dari variabel respon dinyatakan sebagai fungsi dari kovariat yang diamati. Koenker dan Hallock (2005) menyatakan bahwa regresi kuantil berguna untuk data yang memiliki pola sebaran bersyarat yang tidak simetris, data yang memiliki sebaran padat di bagian ekor (atas atau bawah), dan sebaran yang terpotong. Selain itu, regresi kuantil juga bermanfaat digunakan ketika ingin mengetahui tingkat perubahan pada kuantil bersyarat yang dinyatakan oleh koefisien regresi pada kuantil tertentu. Untuk mendapatkan estimasi parameter model regresi kuantil yaitu dengan metode pemrograman linier diantaranya yaitu metode simpleks, *interior-point*, dan *smoothing*.

Metode regresi *robust* pertama kali diperkenalkan oleh Rousseeuw pada tahun 1984. Metode ini sebagai alternatif untuk mengatasi kelemahan metode *Ordinary Least Squares* (OLS). Regresi *robust* merupakan metode yang digunakan untuk mengatasi permasalahan pencilan dan dikembangkan dengan tujuan untuk mampu mengenali dan mengatasi adanya pencilan. Ada beberapa metode estimasi yang bisa digunakan dalam regresi *robust* yaitu penaksir *Least Median of Squares* (LMS), *Least Trimmed Squares* (LTS), penaksir M, penaksir S dan penaksir MM. Menurut Rousseeuw (1984), Metode LTS memiliki kemampuan yang lebih baik dibandingkan dengan metode-metode lainnya karena mampu mengatasi pencilan yang disebabkan baik oleh variabel bebas maupun

variabel terikat. Selain itu, kelebihan dari metode LTS adalah algoritmanya yang lebih mudah dibandingkan metode lainnya. Hal ini disebabkan karena dalam proses estimasi LTS hanya akan memangkas sebaran data berdasarkan jumlah pencilan yang teramati sehingga akan menghasilkan fungsi objektif yang mengecil dan konvergen.

Metode regresi kuantil dan regresi *robust* LTS dapat diaplikasikan di berbagai bidang kehidupan. Semakin berkembangnya kedua metode tersebut, sehingga menarik perhatian para ilmuwan untuk melakukan riset. Sebagai contoh, banyak permasalahan dalam pembangunan Nasional, mulai dari bidang pendidikan, kesehatan, ekonomi, maupun bidang sosial kependudukan dan lain sebagainya. Masalah utama yang dihadapi oleh semua Negara di Dunia yaitu masalah kesehatan, dalam hal ini masalah penyebaran penyakit malaria.

Malaria adalah penyakit infeksi yang disebabkan oleh parasit *Plasmodium* yang hidup dan berkembang biak dalam sel darah merah manusia. Penyakit malaria disebabkan oleh nyamuk (*Anopheles*) betina dan dapat menyerang semua orang, baik laki-laki maupun perempuan pada semua golongan umur. Penyakit ini secara langsung dapat menyebabkan seseorang sering sakit, daya tahan tubuh rendah dan menurunkan produktivitas kerja serta menyebabkan kematian (Kemenkes RI, 2015).

Dalam rangka pengendalian penyakit malaria, banyak hal yang sudah maupun yang sedang dilakukan oleh semua pihak. Malaria merupakan salah satu indikator dari target Pembangunan Milenium (MDGs), dimana ditargetkan untuk menghentikan penyebaran dan mengurangi kejadian insiden malaria pada tahun 2030. Peristiwa ini dilihat dari indikator menurunnya angka kesakitan dan angka kematian akibat malaria. *Global Malaria Programme* (GMP) menyatakan bahwa malaria merupakan penyakit yang harus terus menerus dilakukan pengamatan, monitoring dan evaluasi, serta diperlukan kebijakan dan strategi yang tepat. Pada program GMP ditargetkan 80% penduduk terlindungi dan penderita mendapat pengobatan.

Di Indonesia penyakit malaria ditemukan tersebar luas di berbagai daerah dengan derajat infeksi yang bervariasi. Kejadian tersebut disebabkan karena

adanya permasalahan teknis seperti pembangunan yang tidak peduli terhadap kesehatan lingkungan, wilayah terpencil dengan kondisi lingkungan yang tidak baik, sarana transportasi dan komunikasi yang sulit, minimnya akses pelayanan kesehatan dan ekonomi masyarakat yang rendah serta buruknya perilaku masyarakat terhadap kebiasaan hidup sehat. Jumlah Penderita positif penyakit malaria di Indonesia pada tahun 2015 sebanyak 217.025 jiwa. Berdasarkan data dari Kementerian Kesehatan Republik Indonesia bahwa jumlah penderita penyakit malaria sudah mengalami penurunan. Namun, penurunan penderita penyakit malaria masih tetap harus diwaspadai karena adanya disparitas (kesenjangan) penyebaran penyakit malaria antar daerah di Indonesia. Hal tersebut diduga datanya mengandung pencilan (outliers). Oleh sebab itu, perlu dilakukan penelitian lebih lanjut mengenai faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia.

Penelitian sebelumnya yang pernah dilakukan oleh Rahmawati (2011), menerapkan regresi kuantil pada data suhu harian di Kota Sydney dimana suhu kemarin berpengaruh terhadap suhu hari ini. Data yang digunakan tidak simetris, tidak homogen ataupun tidak beraturan. Penelitian ini, diperoleh kesimpulan bahwa regresi kuantil lebih tepat diterapkan pada data yang tidak simetris, karena memberikan lebih banyak informasi tidak hanya pada pusat sebaran, tetapi pada bagian atas dan bawah sebaran. Vandy (2014), menerapkan analisis regresi kuantil untuk menganalisis faktor-faktor yang mempengaruhi IPM di Pulau Jawa, diperoleh hasil bahwa presentase penduduk dengan pendidikan di atas SLTP merupakan faktor yang paling berpengaruh terhadap IPM di Pulau Jawa. Febria (2016), membandingkan metode regresi kuantil median dan transformasi Box-Cox dalam menangani masalah heteroskedastisitas, pada kasus pengaruh pengguna jaminan kesehatan dari Badan Penyelenggara Jaminan Sosial Kesehatan (BPJS Kesehatan) dan jumlah penduduk terhadap penderita Demam Berdarah Dengue (DBD) pada setiap kabupaten di Provinsi Jawa Tengah. Diperoleh hasil bahwa bahwa transformasi Box-Cox lebih baik dari pada regresi kuantil median dalam menangani masalah heteroskedastisitas.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan diatas, maka permasalahan utama yang ingin dibahas dalam penelitian ini adalah membandingkan estimasi parameter regresi kuantil dengan regresi *robust* LTS pada kasus faktor-faktor yang mempengaruhi penyebaran jumlah penderita penyakit malaria di Indonesia.

1.3 Tujuan Penelitian

Dari permasalahan diatas, maka tujuan penelitian ini adalah sebagai berikut :

1. Mendapatkan model dari metode Regresi Kuantil, Regresi *Robust Least Trimmed Square* (LTS) serta Metode *Ordinary Least Square* (OLS) pada kasus faktor - faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia.
2. Membandingkan Model Estimasi Parameter Regresi Kuantil dengan Regresi *Robust Least Trimmed Square* (LTS) serta *Ordinary Least Square* (OLS), pada kasus faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia.
3. Mendapatkan faktor-faktor yang mempengaruhi kasus penyebaran penyakit malaria di Indonesia berdasarkan model terbaik.

1.4 Manfaat Penelitian

Adapun manfaat yang ingin dicapai dalam penelitian ini adalah sebagai berikut :

1. Memberikan sumbangsih keilmuan dalam menerapkan model Regresi Kuantil dan Regresi *Robust Least Trimmed Square* (LTS).
2. Sebagai bahan rujukan kepada Pemerintah dalam hal ini Kementerian Kesehatan, agar perlu memperhatikan faktor-faktor yang berpengaruh secara signifikan terhadap penyebaran penyakit malaria.

1.5 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut :

1. Nilai kuantil yang digunakan dalam penelitian ini adalah $\tau = 0.05, 0.25, 0.50, 0.75, \text{ dan } 0.95$

Nilai kuantil 0,25 menggambarkan model kuartil pertama, nilai kuantil 0.50 menggambarkan nilai tengah pada data, nilai kuantil 0.75 menggambarkan kuartil ketiga serta nilai kuantil 0.95 menggambarkan 95 persen data dan sisanya merupakan data ekstrim.

1. Estimasi parameter pada regresi kuantil dengan menggunakan metode simpleks. Hal ini disebabkan karena metode simpleks lebih sederhana dan prosesnya lebih cepat.
2. Estimasi parameter regresi *robust* dengan menggunakan metode *Least Trimmed Square* (LTS). Hal ini disebabkan karena metode LTS lebih mudah dalam proses algoritmanya, jika dibandingkan dengan metode yang lain.

BAB II

TINJAUAN PUSTAKA

2.1 Analisis Regresi

Analisis regresi pertama kali ditemukan oleh Sir Francis Galton (1822-1911). Analisis ini merupakan salah satu metode yang digunakan untuk menyelidiki pola hubungan antara variabel respon (x) dengan variabel prediktor (y). Secara umum model persamaan regresi dapat dituliskan sebagai berikut :

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

dimana :

y_i : Variabel respon pada pengamatan ke- i

$f(x_i)$: Kurva regresi pada pengamatan ke- i

ε_i : Residual (error) pada pengamatan ke- i

Analisis Regresi selain digunakan sebagai pemodelan, juga dapat digunakan sebagai peramalan (*forecasting*). Analisis regresi dapat didekati melalui tiga pendekatan yaitu regresi parametrik, semiparametrik dan nonparametrik. Dalam pendekatan regresi parametrik terdapat asumsi yang sangat kuat dan kaku yaitu bentuk kurva regresi diketahui misalnya linier, kuadratik, kubik, polynomial berderajat p , dan lain-lain. Apabila pola data mengikuti pola tertentu seperti linier, kuadratik ataupun kubik maka pendekatan regresi parametrik layak untuk digunakan. Sedangkan apabila pola data tidak mengikuti pola tertentu maka dapat digunakan regresi nonparametrik. Sedangkan regresi semiparametrik digunakan apabila terdapat sebagian pola data yang mengikuti pola tertentu dan sebagian lagi tidak mengikuti pola tertentu.

2.2 Pengujian Outliers

Outliers terjadi karena kombinasi unik yang terjadi dan nilai-nilai yang dihasilkan dari observasi tersebut sangat berbeda dari observasi-observasi lainnya. Deteksi *outliers* dapat dilakukan dengan menggunakan *difference in fitted value* (*DFFITs*). (Myers, 1990)

Adapun persamaan untuk *DFFITS* sebagai berikut :

$$DFFITS_i = \frac{\hat{y}_i - \hat{y}_{i,-i}}{\sqrt{S_{-i}^2 h_{ii}}} \text{ untuk } i = 1, 2, \dots, n \quad (2.2)$$

keterangan :

\hat{y}_i : Nilai taksiran y_i dengan proses pemodelan serta menggunakan pengamatan ke- i

$\hat{y}_{i,-i}$: Nilai taksiran y_i dengan proses pemodelan tanpa menggunakan pengamatan ke- i

S_{-i}^2 : *Mean square error* dari model tanpa menggunakan pengamatan ke- i

h_{ii} : Elemen diagonal ke- i dari matrik $H = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$

Pengamatan dikatakan *outliers* jika $|DFFITS_i| > 2\sqrt{\frac{p+1}{n}}$ dengan (p) banyaknya parameter dalam model dan n adalah banyak pengamatan. (Myers, 1990)

2.3 Metode Ordinary Least Square (OLS)

Metode *Ordinary Least Square* (OLS) pada prinsipnya adalah meminimumkan jumlah kuadrat error (residual). Model umum persamaan regresi adalah sebagai berikut :

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.3)$$

dimana :

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{p1} \\ 1 & x_{12} & x_{22} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{pn} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Selanjutnya akan ditunjukkan penaksir untuk $\hat{\beta}$ dengan meminimumkan persamaan berikut :

$$\sum_{i=1}^n \varepsilon_i^2 \quad (2.4)$$

dimana :

$$\varepsilon_i^2 = (y_i - \beta_0 - \beta_1 X_{1i} - \dots - \beta_p X_{pi})^2$$

Langkah berikut dicari turunan parsial terhadap $\beta_0, \beta_1, \dots, \beta_p$ dan menyamakan dengan nol sehingga diperoleh nilai estimasi model regresi linier.

$$\begin{aligned} \frac{\partial s}{\partial \beta_0} &= -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{1i} - \dots - \beta_p x_{pi}) = 0 \\ \frac{\partial s}{\partial \beta_1} &= -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{1i} - \dots - \beta_p x_{pi}) x_{1i} = 0 \\ \frac{\partial s}{\partial \beta_2} &= -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{1i} - \dots - \beta_p x_{pi}) x_{2i} = 0 \\ &\vdots \\ \frac{\partial s}{\partial \beta_p} &= -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{1i} - \dots - \beta_p x_{pi}) x_{pi} = 0 \end{aligned}$$

maka diperoleh persamaan berikut :

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^n x_{1i} + \hat{\beta}_2 \sum_{i=1}^n x_{2i} + \dots + \hat{\beta}_p \sum_{i=1}^n x_{pi} = \sum_{i=1}^n y_i \quad (2.5)$$

$$\hat{\beta}_0 \sum_{i=1}^n x_{1i} + \hat{\beta}_1 \sum_{i=1}^n x_{1i}^2 + \hat{\beta}_2 \sum_{i=1}^n x_{1i} x_{2i} + \dots + \hat{\beta}_p \sum_{i=1}^n x_{1i} x_{pi} = \sum_{i=1}^n x_{1i} y_i \quad (2.6)$$

$$\hat{\beta}_0 \sum_{i=1}^n x_{2i} + \hat{\beta}_1 \sum_{i=1}^n x_{1i} x_{2i} + \hat{\beta}_2 \sum_{i=1}^n x_{2i}^2 + \dots + \hat{\beta}_p \sum_{i=1}^n x_{2i} x_{pi} = \sum_{i=1}^n x_{2i} y_i \quad (2.7)$$

$$\hat{\beta}_0 \sum_{i=1}^n x_{pi} + \hat{\beta}_1 \sum_{i=1}^n x_{1i} x_{pi} + \hat{\beta}_2 \sum_{i=1}^n x_{2i} x_{pi} + \dots + \hat{\beta}_p \sum_{i=1}^n x_{2i}^2 = \sum_{i=1}^n x_{pi} y_i \quad (2.8)$$

Berdasarkan persamaan (2.5), (2.6) dan (2.7) serta (2.8) dapat dibuat dalam persamaan berikut :

$$\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \mathbf{X}'\mathbf{y} \quad (2.9)$$

untuk menyelesaikan persamaan (2.9), maka dikalikan kedua sisinya dengan invers dari $(\mathbf{X}'\mathbf{X})$. Sehingga diperoleh estimator OLS dari $\hat{\boldsymbol{\beta}}$ adalah sebagai berikut :

$$\begin{aligned} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} \\ \mathbf{I}\hat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} \\ \hat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} \end{aligned} \quad (2.10)$$

2.4 Metode *Least Trimmed Square* (LTS)

Metode *Least Trimmed Squares* (LTS) pertama kali diperkenalkan oleh Rousseeuw pada tahun 1984. Metode ini, sebagai alternatif untuk mengatasi kelemahan metode *Ordinary Least Squares* (OLS). Rousseeuw dan Hubert menjelaskan bahwa metode LTS mempunyai prinsip sama dengan metode OLS dalam mengestimasi parameter dengan meminimumkan jumlah kuadrat *residual*. Namun, metode LTS tidak menggunakan seluruh pengamatan dalam perhitungannya, melainkan hanya meminimumkan jumlah kuadrat *residual* dari himpunan bagian data berukuran h pengamatan dengan kuadrat *residual* terkecil, dimana $h < n$. Fungsi tujuan dari metode LTS dapat dituliskan sebagai berikut :

$$\min \sum_{i=1}^h \varepsilon_{(i)}^2 \quad (2.11)$$

Algoritma yang digunakan pada regresi *Robust Least Trimmed Square* (LTS) disebut sebagai *basic resampling algorithm*, seperti pertama kali dikenalkan oleh Rousseeuw dan Leroy (1978) dalam PROGRESS (*Program for Robust Regression*). Selanjutnya diperbaiki kecepatan untuk perhitungan sampel berukuran besar oleh Rousseeuw dan Van Driessen (2006) dalam FAST-LTS. Secara rinci algoritma estimasi LTS diuraikan sebagai berikut :

Langkah 1

Menginput data berpasangan $z_i = (x_{1i}, x_{2i} \cdots x_{pi}, y_i)$ untuk $i = 1, 2, \dots, n$

Langkah 2

Mengambil sebanyak m subset berdasarkan kombinasi p dari n

$$m = C_p^n = \binom{n}{p} = \frac{n!}{p!(n-p)!} \quad (2.12)$$

Beranggotakan sebanyak p pengamatan sehingga diperoleh

$$J_k = \{z_1, z_2, \dots, z_p\}_k \\ = \{(x_i, y_i)_1, (x_i, y_i)_2, \dots, (x_i, y_i)_p\}_k \quad \text{dimana } k = 1, 2, \dots, m$$

Langkah 3

Memperoleh model regresi untuk setiap k dari subset J_k , yang berbentuk

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.13)$$

Langkah 4

Mendefinisikan nilai h yang dipilih, yaitu $h = \lceil (n + p + 1) / 2 \rceil$ dan menghitung nilai *breakdown* melalui persamaan $\varepsilon_n^* = (n - h + 1) / n$

Langkah 5

Mengevaluasi apakah model regresi menggunakan intersep atau tidak. Jika tidak, maka hitung estimasi parameter regresi untuk setiap subset J_k yang terbentuk. nilai estimasi parameter regresi melalui

$$\hat{\beta} = \mathbf{X}^{-1}\mathbf{y} \quad (2.14)$$

Jika YA maka lakukanlah langkah proses adjustment untuk setiap subset J_k yang terbentuk.

- Menghitung estimasi parameter regresi untuk setiap subset J_k sebanyak m kali dinyatakan oleh

$$\hat{\beta} = \mathbf{X}^{-1}\mathbf{y}$$

- Menghitung nilai ε dengan tanpa menggunakan nilai estimasi intersep

$$\varepsilon = \mathbf{y} - \hat{\mathbf{y}} \quad (2.15)$$

- Mengurutkan nilai ε_i , dimana ε_1 berarti nilai ε_i yang paling kecil dan ε_n berarti nilai ε_i yang paling besar.

$$\varepsilon_1 \leq \varepsilon_2 \leq \dots \leq \varepsilon_n$$

- Membentuk kelas interval yang masing-masing berisi sebanyak h pengamatan, sehingga diperoleh sebanyak $(n - h + 1)$ kelas interval.
- Menghitung nilai intersep yang baru $\hat{\beta}'_p$

Langkah 6

- Menghitung Fungsi Tujuan (FT) untuk tiap-tiap estimasi yang diperoleh

$$FT_{LTS} = \sqrt{\frac{1}{h} \sum_{i=1}^h \varepsilon^2(\hat{\beta})_i} \quad (2.16)$$

dengan $\varepsilon(\hat{\beta})_i = \mathbf{y}_i - \hat{\mathbf{y}}_i$

Langkah 7

Menentukan estimasi terbaik dari sejumlah m estimasi yang telah dilakukan dengan nilai fungsi tujuan yang paling minimum

- Menetapkan nilai fungsi tujuan yang terkecil diantara m kali estimasi.
- Menetapkan estimasi parameter regresi $\hat{\beta}$

- Menghitung koefisien determinasi (R^2_{LTS})

2.5 Metode Regresi Quantil

Metode regresi quantil pertama kali diperkenalkan oleh Roger Koenker dan Gilbert Basset pada tahun 1978. Regresi ini merupakan perluasan dari model regresi pada quantil bersyarat dimana distribusi quantil bersyarat dari variabel respon dinyatakan sebagai fungsi dari kovariat yang diamati. Pendekatan ini memungkinkan menduga fungsi quantil dari sebaran bersyarat respon pada berbagai atau setiap nilai quantil yang diinginkan dari peneliti (Chen dan Wei, 2005). Regresi quantil sangat berguna dalam berbagai bidang yaitu keuangan, kesehatan, ekonometrika, lingkungan dan sebagainya. Misalkan diberikan data $\{x_{1i}, x_{2i}, \dots, x_{pi}, y_i\}$ dimana $i = 1, 2, \dots, n$, merupakan himpunan berpasangan variabel random yang terdistribusi secara independen dan tidak identik dengan quantil $\tau \in (0,1)$. Data tersebut memiliki fungsi sebaran peluang bersyarat $F(Y|x_i) = P(Y \leq y|x_i)$ dan fungsi invers $F^{-1}(\tau) = \inf\{y: F(y) \geq \tau\}$ yang merupakan kuantil ke- τ dari variabel respon y (Gob, 2009). Jadi, pada nilai quantil ke- τ dapat didenifisikan sebagai $Q(\tau) = \inf\{y: F(y) \geq \tau\} = F^{-1}(\tau)$ yang merupakan fungsi quantil ke- τ pada fungsi yang bersyarat x . Persamaan umum regresi quantil linier khusus untuk bersyarat $Q_{Y_i}(\tau|x_{1i}, x_{2i}, \dots, x_{pi})$ dari variabel respon y (karena fungsi linier dari regresi $(x_{1i}, x_{2i}, \dots, x_{pi})$) yaitu :

$$y_i = \beta_{\tau 0} + \beta_{\tau 1}x_{1i} + \dots + \beta_{\tau p}x_{pi} + \varepsilon_{\tau i} \quad (2.17)$$

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

apabila model regresi quantil disajikan dalam bentuk matriks, maka persamaan (2.17) dapat ditulis sebagai berikut :

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{p1} \\ 1 & x_{12} & x_{22} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{pn} \end{bmatrix} \begin{bmatrix} \beta_{\tau 0} \\ \beta_{\tau 1} \\ \vdots \\ \beta_{\tau p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{\tau 1} \\ \varepsilon_{\tau 2} \\ \vdots \\ \varepsilon_{\tau n} \end{bmatrix}$$

Selanjutnya persamaan (2.17) dapat ditulis dalam bentuk model linier berikut :

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}(\tau) + \boldsymbol{\varepsilon}(\tau) \quad (2.18)$$

dengan

$\mathbf{y} : [y_1 \ y_2 \ \cdots \ y_n]^T$ merupakan vektor berukuran $n \times 1$

$\mathbf{X} : \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{p1} \\ 1 & x_{12} & x_{22} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{pn} \end{bmatrix}$ merupakan matriks berukuran $n \times (p+1)$

$\boldsymbol{\beta} : [\beta_{\tau 0} \ \beta_{\tau 1} \ \cdots \ \beta_{\tau p}]^T$ merupakan vektor berukuran $\tau p + 1$

$\boldsymbol{\varepsilon} : [\varepsilon_1 \ \varepsilon_2 \ \cdots \ \varepsilon_n]^T$ merupakan vektor berukuran $n \times 1$

2.6 Estimasi Parameter Regresi Quantil

Diberikan himpunan data berpasangan $\{x_{1i}, x_{2i}, x_{3i}, \cdots, x_{pi}, y_i\}$, $i = 1, 2, \cdots, n$, $j = 1, 2, \cdots, p$ merupakan himpunan berpasangan variabel random yang terdistribusi secara independen dan identik dengan quantil $\tau \in (0, 1)$. Data tersebut memiliki fungsi distribusi peluang bersyarat yang didenifikasikan sebagai berikut :

$$F(Y|x_i) = P(Y \leq y|x_i) \quad (2.19)$$

Dan fungsi invers $F^{-1}(\tau) = \inf \{y : F(y) \geq \tau\}$ yang merupakan quantil ke- τ dari variabel respon y . Jadi, persamaan umum regresi quantil linier untuk quantil bersyarat didenifikasikan sebagai berikut :

$$\begin{aligned} Q_y(\tau|x) &= F_y^{-1}(y|x_i) \\ &= \inf \{y : F_y(y|x) \geq \tau\} \end{aligned}$$

Sehingga model fungsi quantil bersyarat yaitu sebagai berikut :

$$Q_y(\tau|x_i) = \mathbf{X}'\boldsymbol{\beta}(\tau), \tau \in (0, 1) \quad (2.20)$$

Persamaan umum regresi kuantil linier dari variabel respon \mathbf{y} dapat ditulis menjadi

$$\mathbf{y}_i = \beta_{\tau 0} + \beta_{\tau 1}x_{1i} + \beta_{\tau 2}x_{2i} + \cdots + \beta_{\tau p}x_{pi} + \varepsilon_{i\tau} \quad (2.21)$$

dimana :

$$\begin{aligned} y_1 &= \beta_{\tau 0} + \beta_{\tau 1}x_{11} + \beta_{\tau 2}x_{12} + \cdots + \beta_{\tau p}x_{p1} + \varepsilon_{1\tau} \\ y_2 &= \beta_{\tau 0} + \beta_{\tau 1}x_{21} + \beta_{\tau 2}x_{22} + \cdots + \beta_{\tau p}x_{p2} + \varepsilon_{2\tau} \\ &\vdots \\ y_n &= \beta_{\tau 0} + \beta_{\tau 1}x_{n1} + \beta_{\tau 2}x_{n2} + \cdots + \beta_{\tau p}x_{pn} + \varepsilon_{n\tau} \end{aligned}$$

Selanjutnya pada persamaan (2.21) dapat dibuat dalam bentuk matriks berikut :

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{p1} \\ 1 & x_{12} & x_{22} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{pn} \end{bmatrix} \begin{bmatrix} \beta_{\tau 0} \\ \beta_{\tau 1} \\ \vdots \\ \beta_{\tau p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1\tau} \\ \varepsilon_{2\tau} \\ \vdots \\ \varepsilon_{n\tau} \end{bmatrix}$$

sehingga diperoleh bentuk model linier sebagai berikut :

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}(\tau) + \boldsymbol{\varepsilon}(\tau) \quad (2.22)$$

Regresi kuantil mempunyai prinsip meminimumkan jumlah kuadrat residual untuk mencari penduga $\hat{\boldsymbol{\beta}}$, sebagaimana metode OLS. Regresi kuantil ke- τ dari F_Y dapat diperoleh dari persamaan (2.23) sebagai berikut :

$$E\rho(Y - \hat{y}) = \int_{-\infty}^{\hat{y}} (1-\tau)(y - \hat{y})f(y)dy + \int_{\hat{y}}^{\infty} \tau(y - \hat{y})f(y)dy \quad (2.23)$$

selanjutnya pada persamaan (2.23) diminimumkan menjadi nol :

$$\frac{\partial}{\partial \hat{y}} E[\rho_{\tau}(Y - \hat{y})] = 0$$

$$\frac{\partial}{\partial \hat{y}} \int_{-\infty}^{\hat{y}} (1-\tau)(y-\hat{y})f(y)dy + \int_{\hat{y}}^{\infty} \tau(y-\hat{y})f(y)dy = 0$$

$$\frac{\partial}{\partial \hat{y}} \left[(1-\tau) \int_{-\infty}^{\hat{y}} (y-\hat{y})f(y)dy + \tau \int_{\hat{y}}^{\infty} (y-\hat{y})f(y)dy \right] = 0$$

$$(1-\tau) \left[(y-\hat{y})f(y) \Big|_{-\infty}^{\hat{y}} + \int_{-\infty}^{\hat{y}} \frac{\partial}{\partial \hat{y}} (y-\hat{y})f(y)dy \right] +$$

$$\tau \left[(y-\hat{y})f(y) \Big|_{\hat{y}}^{\infty} + \int_{\hat{y}}^{\infty} \frac{\partial}{\partial \hat{y}} (y-\hat{y})f(y)dy \right] = 0$$

$$(1-\tau) \left[(y-\hat{y})f(y) \Big|_{y=\hat{y}} + \int_{-\infty}^{\hat{y}} f(y)dy \right] +$$

$$\tau \left[(y-\hat{y})f(y) \Big|_{y=\hat{y}} + \int_{\hat{y}}^{\infty} f(y)dy \right] = 0$$

$$(1-\tau)[0 + F_Y(\hat{y})] + \tau[0 - (1 - F_Y(\hat{y}))] = 0$$

$$(1-\tau)F_Y(\hat{y}) - \tau(1 - F_Y(\hat{y})) = 0$$

$$(1-\tau)F_Y(\hat{y}) - \tau - \tau F_Y(\hat{y}) = 0$$

$$F_Y(\hat{y}) - \tau = 0$$

maka diperoleh

$$F_Y(\hat{y}) = \tau$$

sehingga quantil ke- τ merupakan solusi dari F_Y .

Regresi quantil ke- τ diperoleh dengan meminimumkan jumlah nilai mutlak dari error dengan pembobot (τ) untuk error positif dan pembobot ($1-\tau$) untuk error negatif. sehingga diperoleh solusi persamaan berikut :

$$\hat{\beta}(\tau) = \min_{\beta \in R^{p+1}} \tau \sum_{y \geq x} |y - \mathbf{X}^T \beta| + (1-\tau) \sum_{y < x} |y - \mathbf{X}^T \beta| \quad (2.24)$$

atau dapat ditulis seperti persamaan (2.25) berikut :

$$\hat{\beta}(\tau) = \min_{\beta \in R^{p+1}} \sum_{i=1}^n \rho_{\tau}(u_i) \quad (2.25)$$

dimana :

$$\rho_{\tau}(u_i) = \begin{cases} (\tau-1)u_i & , \text{ dengan } u_i < 0 \\ \tau u_i & , \text{ dengan } u_i \geq 0 \end{cases}$$

dengan

- $\hat{\beta}(\tau)$: Penduga parameter
- τ : Indeks quantil dengan $\tau \in (0,1)$
- $\rho_{\tau}(u_i)$: *Loss function* (fungsi kerugian)
- u_i : Error dari penduga parameter

Dalam estimasi OLS dari suatu model linier terhadap y diperoleh dengan meminimumkan jumlah kuadrat error. Sedangkan dalam estimasi regresi quantil dari suatu model linier terhadap y diperoleh dengan meminimumkan nilai *loss function* yang tidak simetris yaitu meminimumkan nilai harapan $\rho_{\tau}(u)$.

Selanjutnya akan dibuktikan bahwa *loss function* berbentuk asimetris dengan penjelasan berikut :

diberikan

$$\begin{aligned} \rho_{\tau} &= [\tau I(u \geq 0) + (1-\tau)I((u < 0))] |u| \\ &= [\tau - I(u < 0)] u \end{aligned}$$

dengan

$$I(u \geq 0) = \begin{cases} 1, & u \geq 0 \\ 0, & u < 0 \end{cases} \quad \text{dan} \quad |u| = \begin{cases} u, & u \geq 0 \\ -u, & u < 0 \end{cases}$$

dimana

u : Error dari penduga parameter

$I(u)$: Fungsi indikator yang telah didenifisikan

sehingga dapat dibuktikan

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

a. Untuk $u \geq 0$

$$\begin{aligned} \rho_{\tau} &= [\tau I(u \geq 0) + (1 - \tau)I(u < 0)]|u| \\ &= [\tau I(u \geq 0) + (1 - \tau)I(u < 0)]u \\ &= [\tau 1 + (1 - \tau)I(u < 0)]u \\ &= [\tau + I(u < 0) - \tau I(u < 0)]u \\ &= [\tau + (1 - I(u \geq 0)) - \tau(1 - I(u \geq 0))]u \\ &= [\tau + (1 - 1) - \tau(1 - 1)]u \\ &= \tau u \end{aligned}$$

dapat juga diselesaikan dengan cara seperti berikut :

$$\begin{aligned} \rho_{\tau} &= [\tau - I(u < 0)]u \\ &= [\tau - (1 - I(u \geq 0))]u \\ &= [\tau - (1 - 1)]u \\ &= \tau u \end{aligned}$$

b. Untuk $u < 0$

$$\begin{aligned}
\rho_\tau &= [\tau I(u \geq 0) + (1-\tau)I(u < 0)]|u| \\
&= [\tau I(u \geq 0) + (1-\tau)I(u < 0)](-u) \\
&= [\tau 0 + (1-\tau)I(u < 0)](-u) \\
&= [(\tau-1)I(u < 0)]u \\
&= [(\tau-1)(1-I(u < 0))]u \\
&= [(\tau-1)(1-0)]u \\
&= (\tau-1)u
\end{aligned}$$

dapat juga diselesaikan dengan cara seperti berikut :

$$\begin{aligned}
\rho_\tau &= [\tau - I(u < 0)]u \\
&= [\tau - (1 - I(u \geq 0))]u \\
&= [\tau - (1 - 0)]u \\
&= (\tau - 1)u
\end{aligned}$$

sehingga menjadi

$$\rho_\tau = [\tau I(u \geq 0) + (1-\tau)I(u < 0)]|u| = [\tau - I(u < 0)]u, \forall_u$$

dimana

$$\rho_\tau = \begin{cases} \tau u & , \text{ jika } u \geq 0 \\ (\tau-1)u & , \text{ jika } u < 0 \end{cases}$$

Solusi dari persamaan (2.24) tidak dapat diperoleh secara analitik, tetapi secara numerik. Salah satu metode numerik yaitu algoritma simpleks. Metode algoritma simpleks telah dikembangkan oleh Barrodale dan Robert pada tahun 1974. Metode algoritma tersebut memberikan solusi permasalahan program linier yang melibatkan banyak variabel - variabel keputusan dengan bantuan komputasi. (Davino, Furno dan Vistocco, 2014). Adapun langkah-langkah proses optimasi metode simpleks adalah sebagai berikut :

Diberikan bentuk umum regresi kuantil

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}(\tau) + \boldsymbol{\varepsilon}(\tau)$$

Langkah awal digunakan quantil $\tau = 0.05$ untuk mencari nilai estimasi $\hat{\boldsymbol{\beta}}$, maka harus meminimumkan fungsi

$$\min_{\boldsymbol{\beta}} \sum_{i=1}^n |y_i - \mathbf{X}_i\boldsymbol{\beta}|$$

Diberikan $[x]_+$, non negatif bagian dari x , yaitu $s_1 = [y - \mathbf{X}\boldsymbol{\beta}]_+$ dan $s_2 = [\mathbf{X}\boldsymbol{\beta} - y]_+$, maka masalah linier dapat ditulis

$$\min_{\boldsymbol{\beta}} \{1's_1 + 1's_2 \mid \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{s}_1 - \mathbf{s}_2, \{\mathbf{s}_1, \mathbf{s}_2\} \in R_+^n\}$$

Selanjutnya diberikan $\mathbf{B} = [\mathbf{X} \quad -\mathbf{X}\mathbf{I} \quad -\mathbf{I}]$ dan $\boldsymbol{\psi} = \begin{bmatrix} [\boldsymbol{\beta}] \\ [-\boldsymbol{\beta}] \\ [y - \mathbf{X}\boldsymbol{\beta}] \\ [\mathbf{X}\boldsymbol{\beta} - y] \end{bmatrix}$, $\mathbf{d} = \begin{bmatrix} 0_{[p]} \\ 0_{[p]} \\ 1_{[n]} \\ 1_{[n]} \end{bmatrix}$

seperti permasalahan pemrograman linier standar, maka formulasinya dapat ditulis $\min \mathbf{d}'\boldsymbol{\psi}$ dimana untuk $\mathbf{B}\boldsymbol{\psi} = \mathbf{y}$, $\boldsymbol{\psi} \geq 0$. Oleh karena itu parameter dual dapat ditulis $\max_d \mathbf{y}'\mathbf{z}$ dimana untuk $\mathbf{B}'\mathbf{z} \leq \mathbf{d}$.

Mengingat hasil utama program linier yaitu solusi dari masalah minimasi tersebut diselesaikan dengan metode simpleks. Secara sederhana masalah diatas dapat dirumuskan menjadi $\max_z \left\{ \mathbf{y}'\mathbf{z} \mid \mathbf{X}'\mathbf{z} = \mathbf{0}, \mathbf{z} \in [-1, +1]^n \right\}$, untuk $\mathbf{X}'\mathbf{z} = \mathbf{0}$ dapat

ditransformasikan dengan cara dikalikan dengan $\frac{1}{2}$ dan $\frac{1}{2}\mathbf{X}'\mathbf{z} + \frac{1}{2}\mathbf{X}'\mathbf{1} = \frac{1}{2}\mathbf{X}'\mathbf{1}$,

untuk yang ditambah $\frac{1}{2}\mathbf{X}'\mathbf{1}$ dengan kata lain dapat ditulis seperti pada persamaan

berikut :

$$\mathbf{X}' \left(\frac{1}{2}\mathbf{z} + \frac{1}{2}\mathbf{1} \right) = \frac{1}{2}\mathbf{X}'\mathbf{1}. \quad (2.26)$$

Jika $\frac{1}{2}\mathbf{z} + \frac{1}{2}\mathbf{1} = \boldsymbol{\eta}$ dan $\frac{1}{2}\mathbf{X}'\mathbf{1} = \mathbf{b}$, maka persamaan dapat ditulis $\mathbf{X}'\boldsymbol{\eta} = \mathbf{b}$

dan $\max_{\mathbf{J}} \left\{ \mathbf{y}' \mathbf{J} \mid \mathbf{X}' \mathbf{J} = \mathbf{b}, \mathbf{J} \in [0,1]^n \right\}$

Nilai $\frac{1}{2}$ pada persamaan (2.26) di atas, merupakan kunci utama generalisasi untuk kondisi quantil yang lain. Bahkan masalah minimalisasi untuk median bersyarat dapat dipakai untuk quantil bersyarat ke- τ sehingga diperoleh

$$\min_{\boldsymbol{\beta}} \sum_{i=1}^n \rho_{\tau}(\mathbf{y} - \mathbf{X}' \boldsymbol{\beta}(\tau))$$

dan dengan cara yang sama diperoleh $\max_{\mathbf{z}} \left\{ \mathbf{y}' \mathbf{z} \mid \mathbf{X}' \mathbf{z} = (1-\tau) \mathbf{X}' \mathbf{1}, \mathbf{z} \in [0,1]^n \right\}$

dimana $(1-\tau)$ memiliki peran yang sama seperti $\frac{1}{2}$ pada rumus di median.

2.6 Koefisien Korelasi

Menurut Walpole (1995), korelasi merupakan metode statistik yang digunakan untuk mengukur besarnya hubungan linier antara dua variabel atau lebih. Nilai korelasi populasi (ρ) berkisar antara interval $-1 \leq \rho \leq 1$. Jika korelasi bernilai positif, maka hubungan antara dua variabel bersifat searah. Sebaliknya jika korelasi bernilai negatif, maka hubungan antara dua variabel bersifat berlawanan arah. Koefisien korelasi biasanya mengacu pada koefisien korelasi *momen-produk Pearson*, sesuai nama orang yang mengembangkannya yaitu Pearson. Koefisien korelasi Pearson diperoleh dengan membagi kovarians kedua variabel dengan perkalian simpangan bakunya dengan persamaan sebagai berikut :

$$\rho_{x,y} = \frac{\text{cov}(X,Y)}{\sqrt{\text{var}(x)\text{var}(y)}} \quad (2.27)$$

selanjutnya untuk menguji korelasi antar variabel digunakan hipotesis sebagai berikut :

H_0 : Tidak ada hubungan antara kedua variabel

H_1 : Tidak ada hubungan antara kedua variabel

Statistik uji yang digunakan yaitu sebagai berikut :

$$t = \frac{r_{x,y} \sqrt{n-2}}{\sqrt{1-r^2}} \quad (2.28)$$

dimana
$$r_{x,y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left(\sum_{i=1}^n (x_i - \bar{x})^2\right) \left(\sum_{i=1}^n (y_i - \bar{y})^2\right)}}$$

Kriteria penolakan adalah tolak H_0 jika nilai $|t_{hit}| > t_{(\alpha/2; (n-2))}$

2.7 Multikolinieritas

Multikolinieritas merupakan salah satu masalah dalam analisis regresi, apabila beberapa prediktor berkorelasi maka akan sulit membedakan pengaruh dari masing-masing variabel. Dalam pemodelan regresi, korelasi antara variabel-variabel prediktor disebut dengan multikolinearitas. Korelasi antar variabel prediktor akan menyebabkan determinan matriks sama dengan nol sehingga invers matriks $\mathbf{X}^T \mathbf{X}$ tersebut tidak ada dan pendugaan parameter tidak akan diperoleh. Salah satu cara untuk mendeteksi terjadinya multikolinearitas adalah dengan melihat nilai *Variance Inflation Factor* (VIF), yaitu nilai yang menggambarkan kenaikan varians dari dugaan parameter antar variabel prediktor, dan dirumuskan dalam persamaan berikut :

$$VIF_j = \frac{1}{1 - R_j^2} \quad (2.29)$$

dengan R_j^2 adalah koefisien determinasi antara x_j dengan variabel prediktor lainnya. VIF_j yang lebih besar dari 10 menunjukkan adanya kolinearitas antar variabel prediktor.

2.8 Koefisien Determinasi (R^2)

Koefisien determinasi merupakan metode untuk mengukur seberapa jauh kemampuan model dalam menerangkan variasi dari variabel independen. Nilai

koefisien determinasi yaitu antara nol dan satu. Nilai R^2 yang kecil atau mendekati nol, berarti kemampuan variabel independen dalam menjelaskan variasi variabel dependen sangat terbatas. Begitu sebaliknya, jika nilai R^2 mendekati satu berarti variabel independen memberikan hampir semua informasi yang dibutuhkan untuk memprediksi variabel dependen (Ghazali, 2003).

Adapun rumus dari koefisien determinasi (R^2) adalah sebagai berikut :

$$R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (2.30)$$

dimana :

y_i : Observasi respon ke- i

\bar{y} : Rata-rata

\hat{y}_i : Penduga respon ke- i

2.9 Penyakit Malaria

Penyakit malaria merupakan penyakit yang dapat mengancam keselamatan seseorang, apabila tidak diatasi dengan segera. Penyakit ini bahkan bisa mematikan jika tidak ditangani dengan benar. Secara umum penjelasan mengenai penyakit malaria dapat diuraikan sebagai berikut :

2.9.1 Pengertian Penyakit Malaria

Penyakit malaria adalah salah satu penyakit yang penularannya melalui gigitan nyamuk (*anopheles*) betina. Penyakit ini dapat menyerang siapa saja terutama penduduk yang tinggal di daerah dimana tempat tersebut merupakan tempat yang sesuai dengan kebutuhan nyamuk untuk berkembang. (Kemenkes RI, 2015).

2.9.2 Gejala Klinis Malaria dan Masa Inkubasi

Gejala klinis merupakan petunjuk yang penting dalam diagnosa penyakit malaria. Hal ini dipengaruhi oleh jenis *plasmodium* dan jumlah parasit

yang menginfeksi. Waktu mulai terjadinya infeksi sampai timbulnya gejala klinis dikenal sebagai waktu inkubasi, sedangkan waktu antara terjadinya infeksi sampai ditemukannya parasit dalam darah disebut periode prepaten.

Adapun gejala-gejala malaria yaitu sebagai berikut :

1. Gejala Malaria Ringan (Malaria Tanpa Komplikasi)

Gejala ini, meskipun dikatakan gejala ringan, namun untuk penderita cukup merasakan siksaan. Gejalanya yaitu demam dan menggigil disertai sakit kepala, mual, muntah, diare, nyeri otot atau pegal-pegal. Adapun gejala malaria ini ada tiga stadium berurutan yaitu sebagai berikut :

a. Stadium Dingin

Stadium ini dimulai dengan merasakan dingin pada badan kurang lebih 15 menit hingga 1 jam. Dimulai dengan menggigil dan perasaan sangat dingin, denyut nadi cepat lemah, kulit kering dan terkadang disertai muntah.

b. Stadium Demam

Setelah stadium dingin, pada stadium ini penderita mengalami serangan demam. wajah penderita menjadi merah, kulitnya kering dan merasakan sangat panas seperti terbakar, sakit kepala bertambah keras, dan sering disertai dengan rasa mual atau muntah-muntah. Stadium ini berlangsung selama 2 hingga 4 jam.

c. Stadium Berkeringat

Pada stadium ini penderita berkeringat sangat banyak. Stadium ini berlangsung lebih dari 2 hingga 4 jam. Suhu tubuh penderita kembali turun, kadang-kadang sampai dibawah normal. Setelah itu, biasanya penderita beristirahat hingga tertidur. Setelah bangun tidur penderita merasa lemah tetapi tidak ada gejala lain sehingga dapat kembali melakukan kegiatan sehari-hari.

2. Gejala Malaria Berat (Malaria Dengan Komplikasi)

Penderita dikatakan menderita malaria berat bila di dalam darahnya ditemukan parasit malaria melalui pemeriksaan laboratorium sediaan darah tepi atau *rapid diagnostic test* (rdt) dan disertai satu atau beberapa gejala/komplikasi yaitu sebagai berikut :

- a. Gangguan kesadaran dalam berbagai derajat (mulai dari koma sampai penurunan kesadaran lebih ringan dengan manifestasi seperti : menggigau, bicara salah, tidur terus dan tingkah laku berubah).
- b. Panas sangat tinggi
- c. Mata atau tubuh kuning
- d. Tanda-tanda dehidrasi (mata cekung, bibir kering dan sesak napas)

2.9.3 Faktor-faktor Penyebab Penyakit Malaria

Menurut Kementerian Kesehatan Republik Indonesia (2016), meningkatnya jumlah penderita malaria sangat berkaitan erat dengan beberapa hal yaitu sebagai berikut :

1. Adanya perubahan lingkungan yang berakibat meluasnya tempat perindukkan nyamuk penular malaria.
2. Krisis ekonomi yang berkepanjangan, memberikan dampak pada daerah tertentu dengan adanya masyarakat yang mengalami gizi buruk sehingga lebih rentan untuk terserang malaria.
3. Tidak efektifnya pengobatan.
4. Kurangnya perhatian dan kepedulian masyarakat terhadap upaya penanggulangan malaria terpadu.

Secara umum menurut Teori Hendrik L. Blum (1974), mengatakan bahwa ada empat faktor yang mempengaruhi derajat kesehatan manusia yaitu faktor lingkungan, faktor perilaku, faktor pelayanan kesehatan, dan faktor genetik atau keturunan. Sementara itu menurut Green, perilaku kesehatan seseorang dipengaruhi oleh tiga faktor, yaitu: faktor pemudah (*predisposing factor*) yang mencakup pengetahuan, sikap, keyakinan dan sebagainya. Kemudian faktor pendukung (*enabling factor*) mencakup lingkungan fisik, tersedia atau tidak tersedianya

sarana-sarana kesehatan dan faktor pendorong (*reinforcing factor*) mencakup perilaku dari petugas kesehatan.

2.9.4 Penelitian Sebelumnya

Penelitian sebelumnya pernah dilakukan oleh Indriani dan Yuyun (2010), menganalisis faktor-faktor yang mempengaruhi angka kejadian penyakit malaria di Kabupaten Sorong Selatan. Analisis yang dilakukan dengan menggunakan metode klasifikasi pohon. Diperoleh hasil bahwa faktor utama penyebab penyebaran penyakit malaria di Kabupaten Sorong Selatan yaitu faktor kondisi kemiskinan yang terjadi di lingkungan masyarakat. Selanjutnya Susilowati, P, A. (2013), menganalisis prevelensi penyebaran penyakit malaria di Provinsi Maluku, Maluku Utara, Papua dan Papua Barat dengan faktor-faktor yang mempengaruhinya. Analisis ini dilakukan dengan menggunakan metode *Ordinary Least Square* (OLS) dan regresi *robust*. Dalam penelitian ini, diperoleh kesimpulan yang sama pada kedua metode tersebut yaitu variabel yang paling berpengaruh terhadap penyebaran penyakit malaria yaitu presentase rumah tangga yang tidak memiliki saluran pembuangan air limbah (sanitasi layak) dan presentase rumah tangga mengakses air bersih. Penelitian selanjutnya dilakukan oleh Datukramat, Mayulu dan Masi (2013), menganalisis tentang hubungan sanitasi lingkungan dengan kejadian malaria pada murid sekolah dasar di Kabupaten Bolaang Mongondow Utara. Dalam penelitian ini diperoleh kesimpulan bahwa terdapat hubungan yang signifikan antara sanitasi lingkungan dengan kejadian malaria. Penelitian Selanjutnya dilakukan oleh Sari, Ambarita dan Sitorus (2013), menganalisis tentang hubungan antara akses pelayanan kesehatan dengan kejadian malaria di Provinsi Bengkulu. Dalam penelitian ini, diperoleh kesimpulan bahwa terdapat hubungan yang signifikan antara akses pelayanan kesehatan dengan kejadian malaria di Provinsi Bengkulu. Selanjutnya penelitian dilakukan oleh Miron-toneng, Ismanto dan Malara (2014), penelitian ini dilakukan untuk menganalisis faktor-faktor yang berhubungan dengan kejadian malaria pada anak di wilayah kerja PKM Tona Kecamatan Tahuna Kabupaten Sangihe. Metode yang digunakan yaitu uji Statistik *chi-square*. Diperoleh hasil bahwa faktor pelayanan kesehatan dan faktor lingkungan mempengaruhi penyebaran penyakit malaria. Selanjutnya Jullen,

Nyorong dan Ibnu (2014), menganalisis perilaku pencarian pengobatan masyarakat terhadap penyakit malaria di Kelurahan Remu Utara, Distrik Sorong, Kota Sorong Papua Barat. Diperoleh hasil bahwa, keberadaan Puskesmas dan pelayanan kesehatan kepada masyarakat merupakan faktor utama penyebaran penyakit malaria. Penelitian selanjutnya dilakukan oleh Tangkilisan, Sorisi dan Tuda (2015), menganalisis peran sarana pelayanan kesehatan terhadap kejadian malaria di Kecamatan Silia Raya Kabupaten Minahasa Tenggara. Penelitian ini memperoleh hasil bahwa pelayanan kesehatan dari pihak medis dan keberadaan puskesmas sangat berpengaruh terhadap penyebaran penyakit malaria.

Berdasarkan penelitian - penelitian yang telah dilakukan sebelumnya, mengenai faktor-faktor yang mempengaruhi penyebaran penyakit malaria. Belum ada peneliti yang menggunakan metode regresi kuantil dan regresi *robust* LTS pada kasus tersebut. Berdasarkan data faktor-faktor yang mempengaruhi penyakit malaria, diduga data tersebut mengandung pencilan (outliers). Oleh karena itu, pada penelitian ini dilakukan dengan menggunakan metode regresi kuantil dan regresi *robust* LTS pada kasus faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia.

(Halaman Ini Sengaja Dikosongkan)

BAB 3

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Kementerian Kesehatan Republik Indonesia dan Badan Pusat Statistika (BPS) pada tahun 2015. Data yang diambil mengenai *Annual Parasite Incidence* (API) Malaria di Indonesia beserta faktor-faktor yang diduga mempengaruhi. Adapun data penelitian dapat dilihat pada lampiran 1.

3.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri dari variabel dependen dan variabel independen. Variabel dependen pada penelitian ini adalah jumlah kasus penyebaran penyakit malaria di Indonesia pada tahun 2015. Variabel independen pada penelitian ini yaitu faktor-faktor yang diduga mempengaruhi jumlah kasus penyebaran penyakit malaria. Adapun variabel penelitian dapat dilihat pada Tabel 3.1.

Tabel 3.1 Variabel Penelitian

| Kode | Keterangan |
|-------|--|
| Y | <i>Annual Parasite Incidence</i> (API) Malaria |
| X_1 | Presentase Akses Sanitasi Layak |
| X_2 | Presentase Rumah Layak Huni |
| X_3 | Presentase Penduduk Miskin |
| X_4 | Rasio Tenaga Dokter |
| X_5 | Presentase Akses Air Bersih |
| X_6 | Rasio Puskesmas |

Berikut definisi variabel-variabel penelitian yang terdapat pada Tabel 3.1

1. *Annual Parasite Incidence* (API) Malaria (Y)

Annual Parasite Incidence adalah angka kesakitan malaria (berdasarkan hasil pemeriksaan laboratorium) per 1000 penduduk dalam 1 tahun.

$$API = \frac{\text{Jumlah Penderita Malaria}}{\text{Jumlah Penduduk}} \times 1000$$

2. Presentase Akses Sanitasi Layak (X_1)

Presentase akses sanitasi layak adalah perbandingan antara penduduk atau rumah tangga yang memiliki akses terhadap fasilitas sanitasi yang layak dengan seluruh penduduk atau rumah tangga, yang dinyatakan dalam persentase. Indikator ini digunakan untuk menggambarkan tingkat kesejahteraan rakyat dari aspek kesehatan. rumus yang digunakan adalah sebagai berikut :

$$\text{Sanitasi Layak} = \frac{\text{Jumlah rumah tangga dengan sanitasi layak}}{\text{Jumlah seluruh rumah tangga}} \times 100$$

3. Presentase Rumah Layak Huni (X_2)

Presentase rumah layak huni adalah hasil bagi antara banyaknya jumlah rumah layak huni di suatu wilayah dibagi dengan jumlah rumah di suatu wilayah dikali 100%. Rumus umum rumah presentase rumah layak huni menurut kementerian perumahan rakyat RI adalah sebagai berikut :

$$RLH = \frac{\text{Jumlah rumah layak huni di suatu wilayah}}{\text{Jumlah rumah di suatu wilayah}} \times 100$$

4. Presentase Penduduk Miskin (X_3)

Persentase Penduduk Miskin adalah persentase penduduk yang memiliki rata-rata pengeluaran perkapita perbulan dibawah garis kemiskinan. Rumus umum persentase kemiskinan yaitu sebagai berikut :

$$\%PM_p = \frac{PM_p}{P_p} \times 100$$

Keterangan :

$\%PM_p$: Presentase Penduduk Miskin di Provinsi p .

PM_p : Penduduk Miskin di Provinsi p .

P_p : Jumlah Penduduk di Provinsi p .

5. Rasio Dokter Umum (X_4)

Rasio Tenaga Medis adalah hasil bagi antara banyaknya dokter umum dengan jumlah penduduk di suatu daerah. Menurut standar WHO, Satu dokter umum melayani 2.500 penduduk. Berikut merupakan perhitungan rasio dokter umum.

$$\text{Rasio Tenaga Dokter} = \frac{\text{Banyaknya Dokter Umum}}{\text{Jumlah Penduduk}}$$

6. Presentase Akses Air Bersih (X_5)

Akses Air Bersih adalah air leding eceran / meteran, air hujan, dan pompa / sumur terlindung/mata air terlindung dengan jarak ke tempat penampungan kotoran / tinja ≥ 10 m. Untuk mendapatkan presentase rumah tangga dengan akses air bersih yaitu dengan membagi rumah tangga yang memiliki akses air bersih dengan seluruh rumah tangga dikali 100%.

7. Rasio Puskesmas (X_6)

Rasio Puskesmas adalah hasil bagi antara banyaknya jumlah puskesmas dengan jumlah penduduk di suatu wilayah. Menurut standar Dinas Kesehatan, satu puskesmas melayani 30.000 penduduk. Berikut merupakan perhitungan rasio puskesmas :

$$\text{Rasio Puskesmas} = \frac{\text{Banyaknya Jumlah Puskesmas}}{\text{Jumlah Penduduk}}$$

3.3 Struktur Data

Dalam penelitian ini struktur data yang digunakan dapat dilihat pada tabel 3.2 berikut :

Tabel 3.2 Struktur Data Penelitian

| Subyek | Y | X ₁ | X ₂ | X ₃ | X ₄ | X ₅ | X ₆ |
|-----------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| 1 | y _{1,1} | x _{1,1} | x _{2,1} | x _{3,1} | x _{4,1} | x _{5,1} | x _{6,1} |
| 2 | y _{1,2} | x _{1,2} | x _{2,2} | x _{3,2} | x _{4,2} | x _{5,2} | x _{6,2} |
| 3 | y _{1,3} | x _{1,3} | x _{2,3} | x _{3,3} | x _{4,3} | x _{5,3} | x _{6,3} |
| 4 | y _{1,4} | x _{1,4} | x _{2,4} | x _{3,4} | x _{4,4} | x _{5,4} | x _{6,4} |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 34 | y _{1,34} | x _{1,34} | x _{2,34} | x _{3,34} | x _{4,34} | x _{5,34} | x _{6,34} |

3.4 Langkah-Langkah Penelitian

Secara umum langkah-langkah yang akan dilakukan dalam penelitian ini yaitu sebagai berikut :

- A. Langkah-langkah estimasi parameter regresi kuantil dengan regresi *robust least trimmed square* (LTS).
 1. Prosedur estimasi parameter regresi kuantil
 - a. Mengestimasi parameter model regresi kuantil, yaitu dengan mencari nilai estimasi $\hat{\beta}$ dengan meminimumkan fungsi

$$\min_{\beta} \sum_{i=1}^n |y_i - \mathbf{x}'_i \beta|$$

b. Mengoptimasi simpleks dengan proses sebagai berikut :

- Menghitung nilai $\mathbf{s}_1 = [\mathbf{y} - \mathbf{X}\boldsymbol{\beta}]_+$ dan nilai $\mathbf{s}_2 = [\mathbf{X}\boldsymbol{\beta} - \mathbf{y}]_+$ sehingga diperoleh permasalahan linier. Selanjutnya
- Menghitung nilai $\min_{\boldsymbol{\psi}} \mathbf{d}'\boldsymbol{\psi}$ dan memperoleh nilai parameter dual dari $\max_{\mathbf{d}} \mathbf{y}'\mathbf{z}$ dimana $\mathbf{B}'\mathbf{z} \leq \mathbf{d}$
- Menentukan permasalahan minimalisasi yaitu dengan formulasi $\max_{\mathbf{z}} \left\{ \mathbf{y}'\mathbf{z} \mid \mathbf{X}'\mathbf{z} = \mathbf{0}, \mathbf{z} \in [-1, +1]^n \right\}$
- Untuk $\mathbf{X}'\mathbf{z} = \mathbf{0}$ dapat ditransformasikan menjadi $\mathbf{X}' \left(\frac{1}{2}\mathbf{z} + \frac{1}{2}\mathbf{1} \right)$
- Jika $\frac{1}{2}\mathbf{z} + \frac{1}{2}\mathbf{1} = \boldsymbol{\eta}$ dan $\frac{1}{2}\mathbf{X}'\mathbf{1} = \mathbf{b}$ maka diperoleh $\max \left\{ \mathbf{y}'\boldsymbol{\eta} \mid \mathbf{X}'\boldsymbol{\eta} = \mathbf{b}, \boldsymbol{\eta} \in [0, 1]^n \right\}$

c. Mendapatkan estimator $\hat{\boldsymbol{\beta}}$ yaitu sebagai berikut :

$$\hat{\boldsymbol{\beta}}(\tau) = \min_{\boldsymbol{\beta} \in \mathbb{R}^{p+1}} \tau \sum_{y \geq x} |\mathbf{y} - \mathbf{X}'\boldsymbol{\beta}| + (1-\tau) \sum_{y < x} |\mathbf{y} - \mathbf{X}'\boldsymbol{\beta}|$$

2. Prosedur Estimasi LTS dapat diuraikan sebagai berikut :

Mendapatkan penduga koefisien regresi dari model melalui pendekatan *Least Trimmed Squares* (LTS) dengan menjabarkan algoritma untuk mengestimasi parameter regresi dengan estimator LTS.

1. Menginput data berpasangan $z_i = (x_i, y_i)$ untuk $i = 1, 2, \dots, n$

2. Menentukan J_k dengan

$$J_k = \left\{ z_{i_1}, z_{i_2}, \dots, z_{i_p} \right\}_k$$

$$= \left\{ (x_i, w_i)_1, (x_i, w_i)_2, \dots, (x_i, w_i)_p \right\}_k$$

3. Memperoleh model regresi untuk setiap k dari subset J_k

4. Mendenifikasikan nilai h yang dipilih

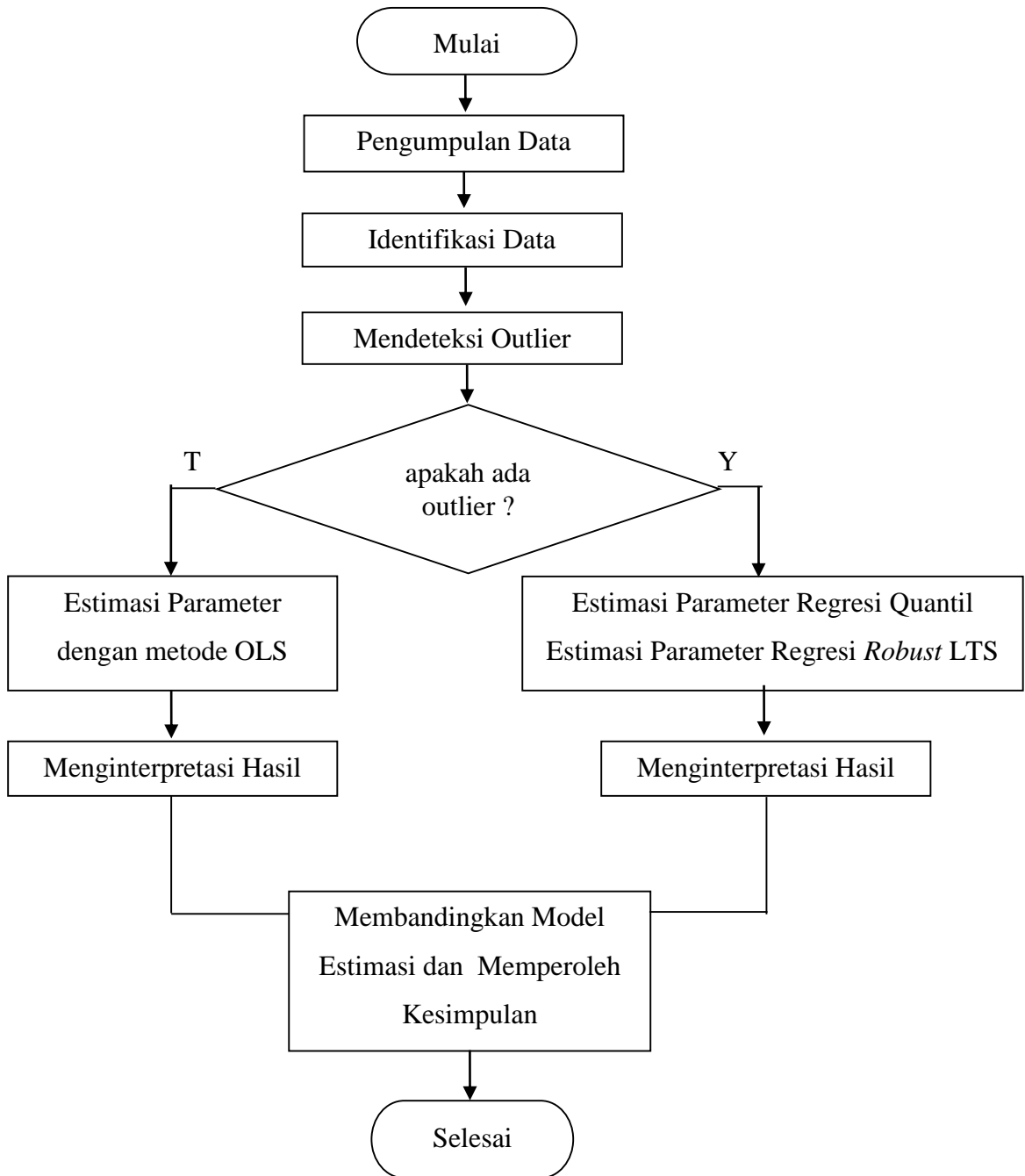
5. Menghitung fungsi tujuan (FT) untuk tiap-tiap estimasi yang diperoleh

6. Menentukan estimasi terbaik dari sejumlah m estimasi.

B. Langkah-langkah untuk menerapkan model regresi quantil dengan regresi *robust* LTS terhadap faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia.

- a. Diberikan data jumlah kasus penyebaran penyakit malaria di Indonesia
- b. Tahap Identifikasi Data
Setelah melakukan analisis deskriptif, tahap selanjutnya yaitu tahap identifikasi data riil dengan membuat *scatter plot* antara masing-masing variabel prediktor dengan variabel respon untuk mendeteksi pola hubungannya.
- c. Mendeteksi Data Outlier.
- d. Mendapatkan model regresi klasik (OLS) pada faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia.
- e. Mendapatkan model regresi *robust* LTS pada faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia.
- f. Mendapatkan model regresi quantil pada faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia.
- g. Mendapatkan model terbaik
- h. Menentukan faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia.
- i. Memperoleh Kesimpulan

Langkah-langkah dalam penelitian ini dapat digambarkan dalam diagram alir sebagai berikut :



Gambar 3.1 Diagram alir pemodelan faktor-faktor yang mempengaruhi

(Halaman ini Sengaja Dikosongkan)

BAB IV

HASIL DAN PEMBAHASAN

Pada Bab 4 ini akan membahas tentang perbandingan metode Regresi Quantil dengan metode Regresi *Robust Least Trimmed Square* (LTS) dan metode *Ordinary Least Square* (OLS) serta mendapatkan faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia.

4.1 Pemodelan Penyebaran Penyakit Malaria Di Indonesia

4.1.1 Statistika Deskriptif

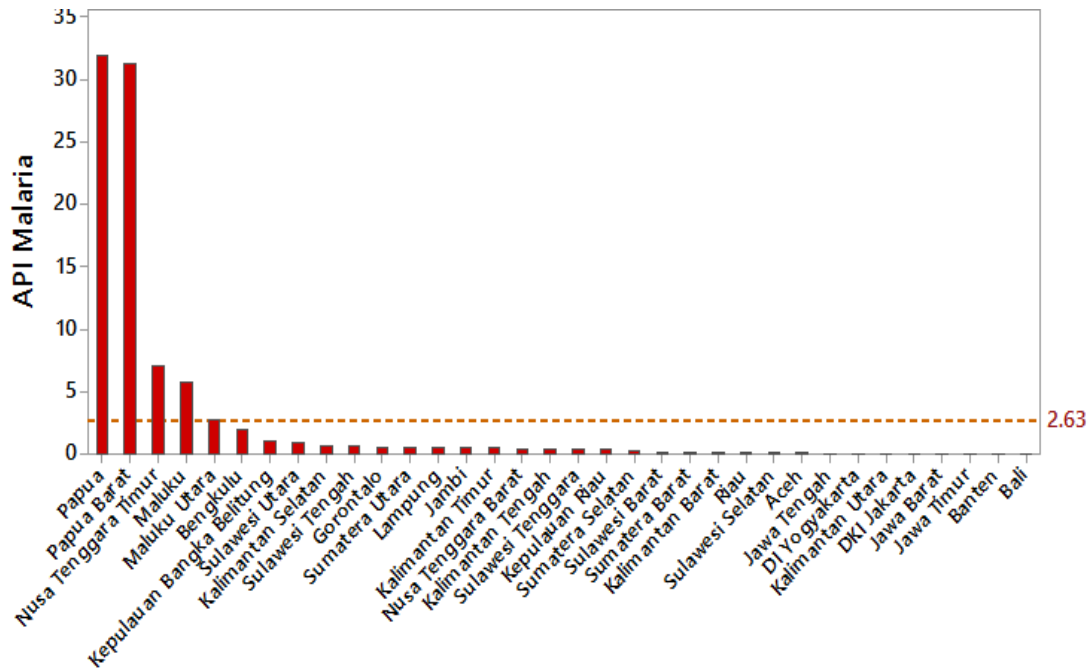
Statistika deskriptif dapat digunakan untuk mengetahui ukuran pemusatan data (*mean*), minimum dan maksimum, standar deviasi serta variansi dari suatu data. Hasil statistika deskriptif pada kasus faktor-faktor yang mempengaruhi penyebaran penyakit malaria, disajikan pada Tabel 4.1.

Tabel 4.1 Statistik Deskriptif Variabel

| Deskriptif | Y | X ₁ | X ₂ | X ₃ | X ₄ | X ₅ | X ₆ |
|-------------|-------|----------------|----------------|----------------|----------------|----------------|----------------|
| Mean | 2,63 | 59,08 | 89,83 | 11,70 | 20,17 | 68,62 | 1,86 |
| Minimum | 0,00 | 23,90 | 49,23 | 3,61 | 10,95 | 41,08 | 0,58 |
| Maksimum | 31,93 | 89,28 | 99,21 | 28,40 | 39,18 | 93,40 | 5,20 |
| Range | 31,93 | 65,38 | 49,98 | 24,79 | 28,23 | 52,32 | 4,62 |
| Std Deviasi | 7,51 | 15,37 | 10,41 | 6,19 | 6,81 | 11,04 | 1,07 |
| Variansi | 56,47 | 236,17 | 108,33 | 38,26 | 46,40 | 121,79 | 1,15 |

Berdasarkan data dari Kementerian Kesehatan Republik Indonesia, menunjukkan bahwa *annual parasite incidence* (API) hingga tahun 2015 terus mengalami penurunan, rata-rata Nasional API malaria (Y) sebanyak 2,63 persen. Berdasarkan Tabel 4.1,

penderita penyakit malaria tertinggi berada di Provinsi Papua yaitu sebanyak 31,93 persen. Sedangkan Provinsi Bali merupakan propinsi yang memiliki presentase penderita malaria terendah yaitu sebanyak 0 persen. Selanjutnya pada Gambar 4.1 dapat dilihat sebaran API malaria di Indonesia.

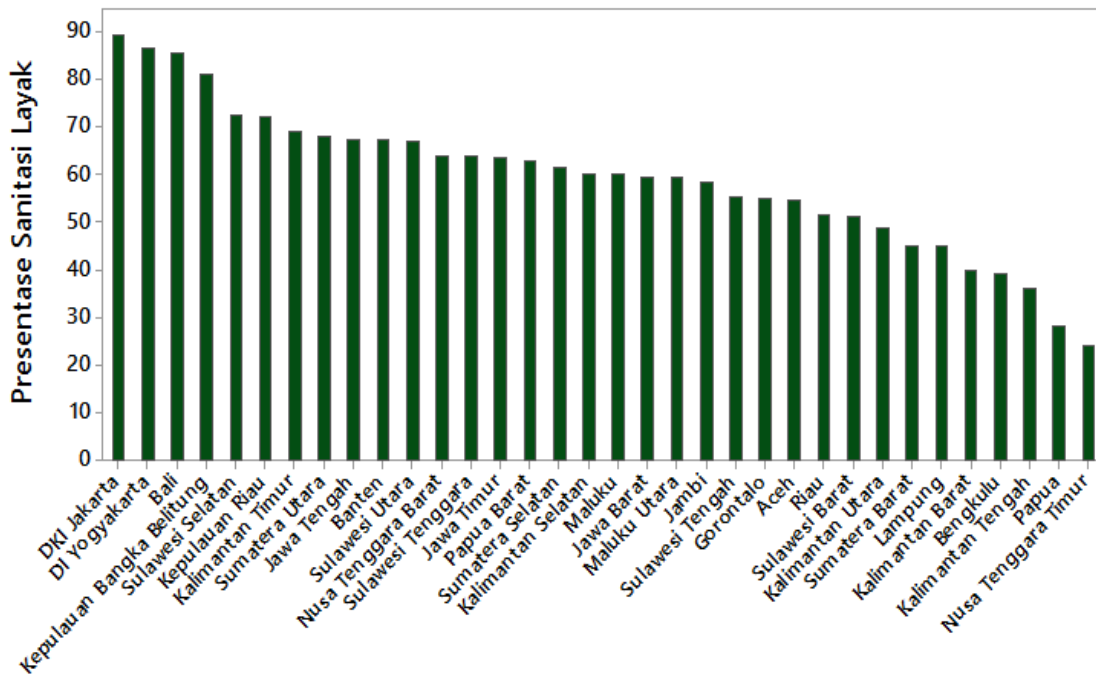


Gambar 4.1 Grafik *Annual Parasite Incidence* (API) malaria pada tahun 2015

Berdasarkan Gambar 4.1, dapat diketahui bahwa pada tahun 2015 masih terdapat beberapa propinsi yang memiliki tingkat API malaria jauh atas rata-rata Nasional. Beberapa Propinsi dengan tingkat API malaria diatas rata-rata nasional yaitu Propinsi Papua sebanyak 31,93 persen, Propinsi Papua Barat sebanyak 31,29 persen, Propinsi Nusa Tenggara Timur (NTT) sebanyak 7,04 persen, Propinsi Maluku sebanyak 5,81 persen, dan Propinsi Maluku Utara sebanyak 2,77 persen. Hal ini menunjukkan bahwa daerah dengan penderita malaria terbanyak berada di wilayah Indonesia bagian Timur.

Salah satu faktor yang mempengaruhi penyebaran penyakit malaria adalah faktor Akses Sanitasi Layak. Rata-rata rumah tangga di Indonesia yang memiliki akses sanitasi layak yaitu sebanyak 59,08 persen rumah tangga. Provinsi DKI Jakarta merupakan

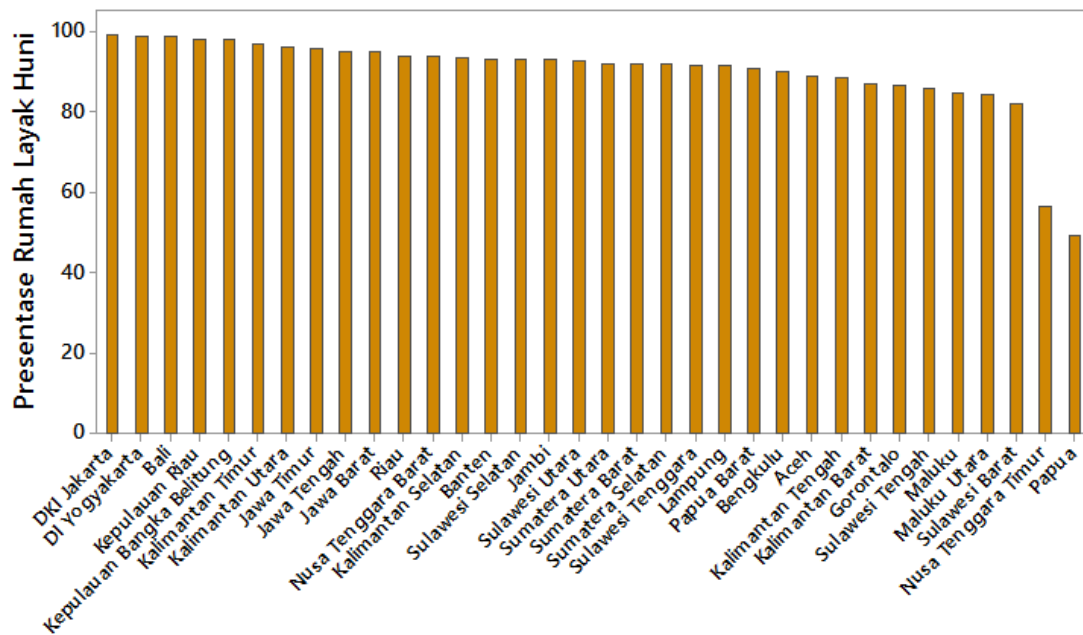
propinsi yang memiliki akses sanitasi layak tertinggi di Indonesia dengan presentase sebanyak 89,28 persen rumah tangga. Selain itu, Propinsi DKI Jakarta juga merupakan salah satu propinsi yang memiliki presentase penderita malaria terendah yaitu sebanyak 0 persen penderita malaria. Sedangkan, Provinsi Nusa Tenggara Timur (NTT) merupakan Propinsi dengan kepemilikan akses akses sanitasi layak terendah yaitu sebanyak 23,90 persen rumah tangga. Selain itu, Propinsi Nusa Tenggara Timur (NTT) juga merupakan salah satu Propinsi yang memiliki presentase *annual parasite incidence* (API) malaria di atas rata-rata nasional yaitu sebanyak 7,04 persen penderita malaria. Hal ini menunjukkan bahwa akses sanitasi layak memiliki hubungan terhadap peningkatan jumlah penderita penyakit malaria. Selanjutnya pada Gambar 4.2, dapat dilihat presentase akses sanitasi layak di Indonesia.



Gambar 4.2 Grafik presentase akses sanitasi layak di Indonesia pada tahun 2015

Salah satu faktor pemicu meningkatnya penyebaran penyakit malaria di Indonesia adalah faktor kepemilikan rumah layak huni. Rata-rata presentase rumah

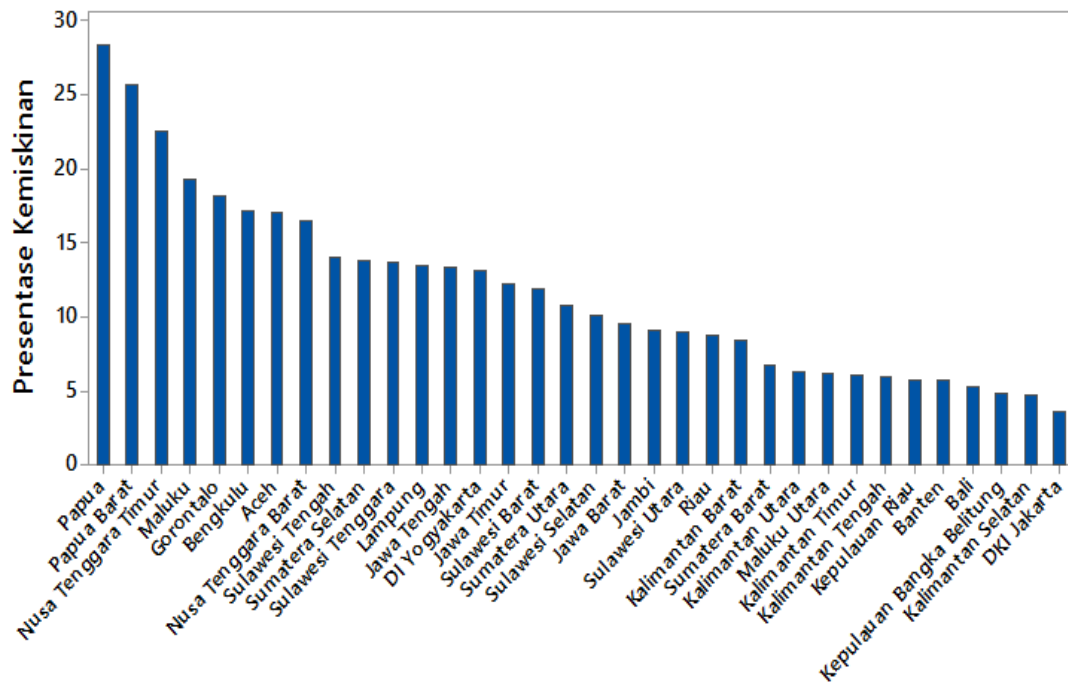
tangga yang memiliki rumah layak huni di Indonesia yaitu sebanyak 89,83 persen rumah tangga. Propinsi DKI Jakarta merupakan propinsi yang memiliki presentase rumah layak huni tertinggi yaitu sebanyak 99,21 persen. Sedangkan propinsi yang memiliki presentase rumah layak huni terendah yaitu Propinsi Papua, dengan presentase sebanyak 49,23 persen dan sekaligus merupakan propinsi yang memiliki presentase penderita malaria tertinggi yaitu sebanyak 31,93 persen penderita malaria. Hal ini menunjukkan bahwa kepemilikan rumah layak huni, memiliki hubungan atau dampak terhadap peningkatan jumlah penderita penyakit malaria. Selanjutnya pada Gambar 4.3, dapat dilihat sebaran presentase rumah layak huni di Indonesia.



Gambar 4.3 Grafik presentase rumah layak huni pada tahun 2015

Selanjutnya salah satu pemicu bertambahnya presentase penderita penyakit malaria di suatu daerah, disebabkan karena masalah kemiskinan. Rata-rata presentase penduduk miskin di Indonesia pada tahun 2015 sebanyak 11,70 persen penduduk miskin. Propinsi yang memiliki penduduk miskin terendah yaitu Propinsi DKI Jakarta dengan

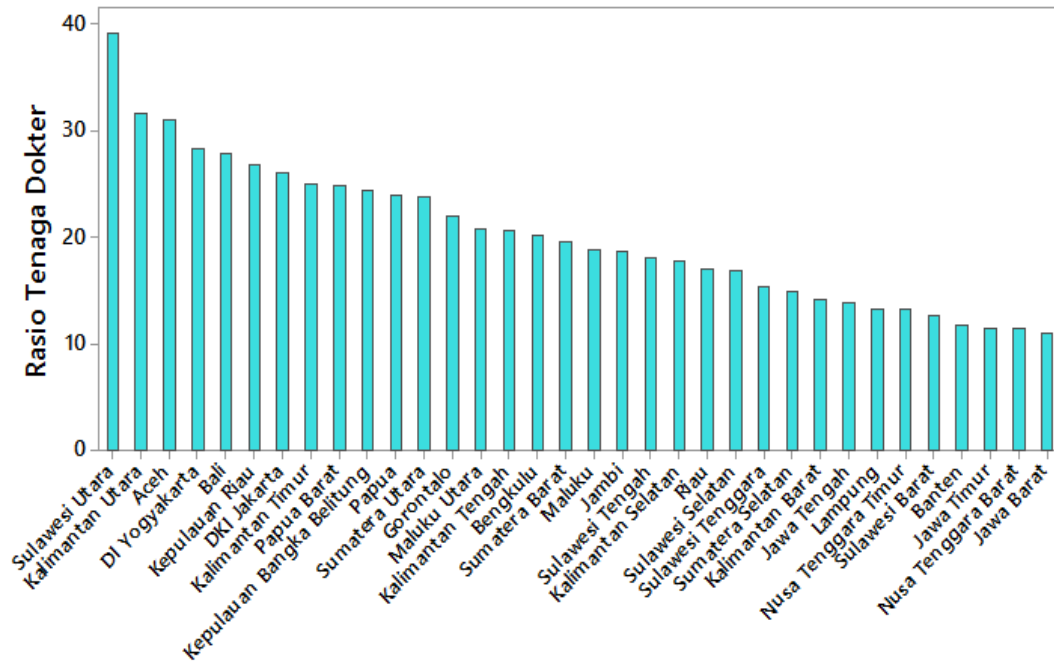
presentase sebanyak 3,61 persen penduduk miskin, disusul oleh Propinsi Kalimantan Selatan sebanyak 4,72 persen penduduk miskin. Sementara itu, presentase penduduk miskin tertinggi berada di Provinsi Papua yaitu sebanyak 28,4 persen penduduk miskin, disusul oleh Propinsi Papua Barat sebanyak 25,73 persen penduduk miskin dan Propinsi Nusa Tenggara Timur (NTT) sebanyak 22,58 persen penduduk miskin. Selanjutnya pada Gambar 4.4, dapat dilihat sebaran presentase penduduk miskin di Indonesia.



Gambar 4.4 Grafik presentase penduduk miskin di Indonesia pada tahun 2015

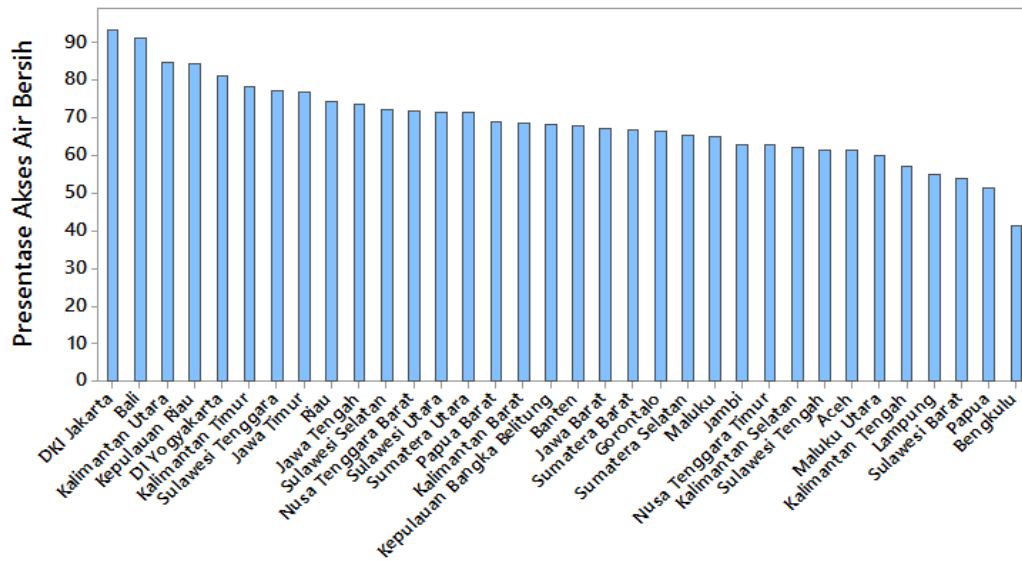
Tenaga dokter memiliki peranan penting dalam kesehatan masyarakat di suatu daerah. Minimnya tenaga dokter akan mempersulit masyarakat untuk memeriksa kesehatan serta berkonsultasi mengenai masalah kesehatan dan lain-lain, begitu sebaliknya. Rata-rata rasio tenaga dokter di Indonesia yaitu sebanyak 20,17 tenaga dokter. Propinsi Sulawesi Utara merupakan propinsi yang memiliki rasio tenaga dokter tertinggi yaitu sebanyak 39,18 tenaga dokter, disusul oleh Propinsi Kalimantan Utara yaitu sebanyak 31,62 tenaga dokter. Sedangkan Propinsi Jawa Barat merupakan propinsi yang memiliki rasio tenaga dokter terendah yaitu sebanyak 10,95 tenaga dokter, disusul

oleh Propinsi Nusa Tenggara Barat (NTB) dengan rasio tenaga dokter sebanyak 11,44 tenaga dokter. Selanjutnya pada Gambar 4.5 dapat dilihat rasio tenaga dokter di Indonesia.



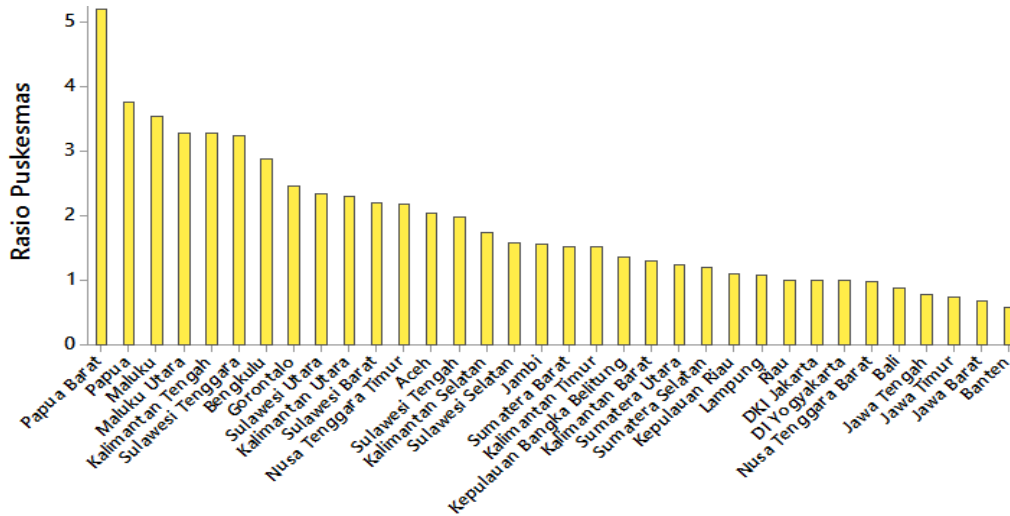
Gambar 4.5 Grafik rasio tenaga dokter di Indonesia pada tahun 2015

Selanjutnya rata-rata presentase rumah tangga di Indonesia yang memiliki akses air bersih yaitu sebanyak 68,62 persen rumah tangga. Propinsi yang memiliki presentase akses air bersih tertinggi yaitu Provinsi DKI Jakarta dengan presentase sebanyak 93,4 persen rumah tangga. Sedangkan, Provinsi Bengkulu merupakan propinsi yang memiliki presentase rumah tangga dengan akses air bersih terendah yaitu sebanyak 41,08 persen rumah tangga. Selanjutnya pada Gambar 4.6 dapat dilihat presentase rumah tangga yang memiliki akses air bersih di Indonesia.



Gambar 4.6 Grafik presentase akses air bersih di Indonesia pada tahun 2015

Rata-rata rasio puskesmas di Indonesia yaitu sebanyak 1,86 puskesmas. Keberadaan puskesmas terbanyak berada di Provinsi Papua Barat dengan rasio sebanyak 5,2 puskesmas. Sementara itu, Propinsi Banten merupakan propinsi yang memiliki puskesmas paling sedikit dengan rasio sebanyak 0,58 puskesmas. Selanjutnya pada Gambar 4.7, dapat dilihat rasio keberadaan puskesmas di Indonesia.

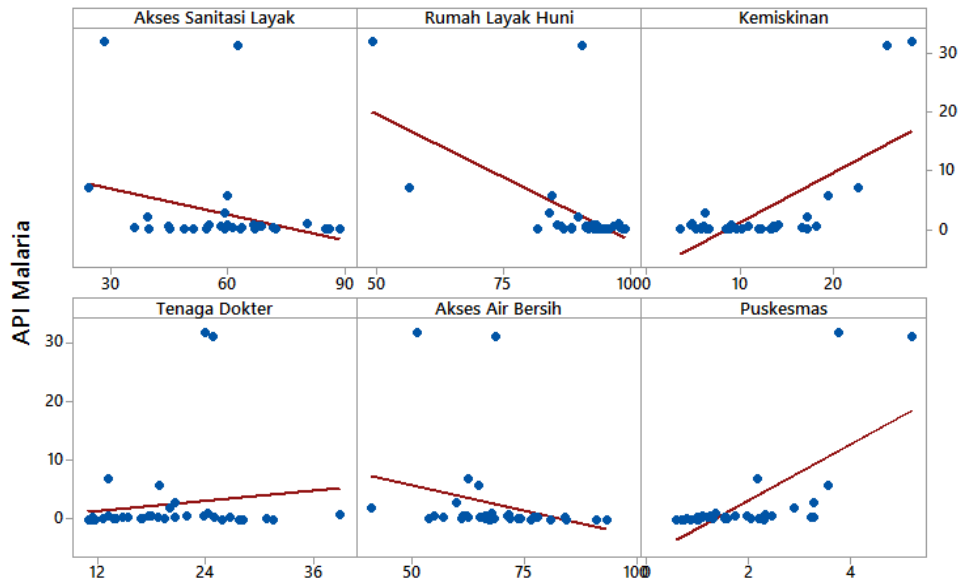


Gambar 4.7 Grafik rasio puskesmas di Indonesia pada tahun 2015

Selanjutnya pada nilai standar deviasi dimaknai bahwa semakin tinggi nilai standar deviasi, maka semakin besar penyimpangan data dari rata-rata hitungannya, sehingga data dikatakan memiliki variabilitas yang tinggi. Pada penelitian ini, variabel yang memiliki standar deviasi paling tinggi adalah variabel presentase akses sanitasi layak (X_1) yaitu sebanyak 15,37 persen. Artinya, pada variabel presentase akses sanitasi layak adalah heterogen. Sebaliknya, semakin rendah nilai standar deviasi, maka semakin rendah pula penyimpangan data dari rata-rata hitungannya, sehingga dapat dikatakan memiliki variabilitas yang rendah. Seperti halnya standar deviasi, nilai varians menunjukkan tingkat keragaman pada data masing-masing variabel. Keragaman data terbesar berada pada variabel presentase akses sanitasi layak (X_1) yaitu sebanyak 236,17 persen. Nilai varians yang besar terjadi apabila data memiliki nilai range yang luas. Seperti pada variabel presentase akses sanitasi layak memiliki nilai jarak paling bawah yaitu 23,90 dan jarak paling atas yaitu 89,28.

4.1.2 Identifikasi Pola Hubungan dan Pengujian Outliers

Studi kasus pada penelitian ini adalah faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia, dimana akan diteliti faktor penyebab penyebaran penyakit malaria (X) terhadap banyaknya presentase penderita penyakit malaria di Indonesia (Y). Selanjutnya pada Gambar 4.8 dapat dilihat pola hubungan antara presentase penderita penyakit malaria dengan faktor-faktor penyebab penyakit malaria.



Gambar 4.8 Scatterplot antara variabel respon dengan variabel prediktor

Pada Gambar 4.8 di atas, menunjukkan bahwa pola hubungan antara variabel *annual parasite incidence* (API) malaria (Y) dengan faktor-faktor penyebab penyakit malaria (X) memiliki pola yang tidak jelas dan sebaran data menyebar. Selanjutnya pada Tabel 4.2 akan ditunjukkan nilai koefisien korelasi antara variabel *annual parasite incidence* (API) malaria (Y) dengan presentase akses sanitasi layak (X_1), presentase rumah layak huni (X_2), presentase penduduk miskin (X_3), rasio banyaknya tenaga dokter (X_4), presentase akses air bersih (X_5) dan rasio banyaknya puskesmas (X_6). Koefisien korelasi dilakukan untuk mengukur tingkat keeratan hubungan linier antar dua variabel.

Pemeriksaan hubungan antar variabel secara statistik dilakukan dengan menggunakan *Korelasi Pearson*. Pengujian hipotesisnya adalah sebagai berikut :

H_0 : Tidak ada hubungan antar kedua variabel

H_1 : Terdapat hubungan antar kedua variabel

Selanjutnya pada Tabel 4.2, menunjukkan nilai koefisien korelasi antar variabel respon dengan variabel prediktor.

Tabel 4.2 Koefisien korelasi antar variabel

| Variabel | Y | X1 | X2 | X3 | X4 | X5 |
|----------|----------------------|----------------------|----------------------|----------------------|--------------------|----------------------|
| X1 | [-0,298] (0,087) | | | | | |
| X2 | [-0,594] (0,000)* | [0,717] (0,000)* | | | | |
| X3 | [0,700] (0,000)* | [-0,446] (0,008)* | [-0,667] (0,000)* | | | |
| X4 | [0,134] (0,449) | [0,289] (0,097) | [0,149] (0,400) | [-0,113] (0,524) | | |
| X5 | [-0,255] (0,145) | [0,702] (0,000)* | [0,540] (0,001)* | [-0,437] (0,010)* | [0,317] (0,067) | |
| X6 | [0,683] (0,000)* | [-0,395] (0,021)* | [-0,471] (0,005)* | [0,548] (0,001)* | [0,247] (0,158) | [-0,409] (0,016)* |

Ket : [] Korelasi Pearson,
() *P-value*, * sig pada $\alpha = 5\%$

Berdasarkan Tabel 4.2 diatas, menunjukkan bahwa terdapat tiga variabel yang berkorelasi positif terhadap *annual parasite incidence* (API) malaria (Y) yaitu presentase rumah layak huni (X₂), variabel presentase penduduk miskin (X₃) dan variabel rasio puskesmas (X₆). Hal ini menunjukkan bahwa jika terjadi peningkatan pada variabel presentase rumah layak huni (X₂), variabel presentase penduduk miskin (X₃) dan variabel rasio puskesmas (X₆) akan mengakibatkan semakin tingginya *annual parasite incidence* (API) malaria (Y) begitu sebaliknya. Sementara itu, variabel presentase akses sanitasi layak (X₁), variabel rasio tenaga dokter (X₄) dan variabel presentase akses air bersih (X₅) memiliki korelasi negatif artinya jika terjadi peningkatan pada variabel tersebut maka

akan mengakibatkan penurunan pada *annual parasite incidence* (API) malaria (Y). Selanjutnya untuk memastikan ada atau tidaknya kasus multikolinieritas, dapat dilihat dari nilai *Variance Inflation Factors* (VIF) dari setiap variabel prediktor. Berikut pada Tabel 4.3 menunjukkan nilai *Variance Inflation Factors* (VIF) dari masing-masing variabel prediktor yang digunakan dalam penelitian ini.

Tabel 4.3 Nilai *Variance Inflation Factor* (VIF)

| Variabel Prediktor | X ₁ | X ₂ | X ₃ | X ₄ | X ₅ | X ₆ |
|--------------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Nilai VIF | 3,05 | 3,01 | 2,20 | 1,48 | 2,24 | 1,99 |

Berdasarkan Tabel 4.3 diatas, menunjukkan bahwa nilai VIF untuk semua variabel prediktor adalah lebih kecil dari 10 yang artinya tidak terjadi multikolinieritas diantara variabel prediktor.

Selanjutnya perlu diuji data outliers dengan menggunakan metode *DFFITs*. Pengamatan dinyatakan *outliers* jika $|DFFITs| > 2\sqrt{(p+1)/n} = 0,907$

Tabel 4.4 Pengujian Data Pencilan (Outliers)

| Propinsi | <i>DFFITs</i> |
|---------------------|---------------|
| Nusa Tenggara Timur | 1,69655 |
| Sulawesi Tenggara | 1,28412 |
| Papua Barat | 6,41416 |
| Papua | 4,55268 |

Berdasarkan Tabel 4.4 diatas, dapat dilihat bahwa terdapat empat Propinsi yang dinyatakan sebagai outliers yaitu Propinsi Nusa Tenggara Timur (NTT), Propinsi

Sulawesi Tenggara, Propinsi Papua Barat dan Propinsi Papua. Apabila analisis regresi dilakukan dengan menggunakan metode OLS, akan mengakibatkan varians dari penduga parameter besar, sehingga tidak efisien. Oleh karena itu, digunakan regresi kuantil dan regresi *robust* LTS untuk memperoleh hasil estimasi parameter dari faktor-faktor yang mempengaruhi penyebaran presentase penderita penyakit malaria di Indonesia.

4.1.3 Estimasi Parameter dengan menggunakan *Ordinary Least Square* (OLS)

Pada kasus ini akan diestimasi dengan menggunakan metode *Ordinary Least Square* (OLS). Nilai dugaan bagi parameter pada metode OLS diperoleh dengan meminimumkan jumlah kuadrat residual. Selain itu, metode OLS didasarkan pada distribusi mean sehingga tidak dapat merepresentasikan keseluruhan data. Pengolahan data analisis dilakukan dengan menggunakan program R. Syntax metode OLS dapat dilihat pada lampiran 3. Berikut pada Tabel 4.5 dapat dilihat hasil estimasi dengan menggunakan metode *Ordinary Least Square* (OLS).

Tabel 4.5 Estimasi Parameter menggunakan metode *Ordinary Least Square* (OLS)

| Parameter | Koefisien | Std Error | <i>t-value</i> | <i>p-value</i> | <i>Keputusan</i> |
|-----------|-----------|-----------|----------------|----------------|-------------------|
| β_0 | 2,211 | 12,916 | 0,171 | 0,865 | Gagal Tolak H_0 |
| β_1 | 0,095 | 0,092 | 1,025 | 0,314 | Gagal Tolak H_0 |
| β_2 | -0,255 | 0,136 | -1,878 | 0,071 | Gagal Tolak H_0 |
| β_3 | 0,451 | 0,195 | 2,307 | 0,028 | Tolak H_0 |
| β_4 | 0,024 | 0,145 | 0,165 | 0,870 | Gagal Tolak H_0 |
| β_5 | 0,091 | 0,110 | 0,824 | 0,417 | Gagal Tolak H_0 |
| β_6 | 3,073 | 1,072 | 2,865 | 0,008 | Tolak H_0 |

Ket : Signifikan pada $\alpha = 5\%$, $R^2 = 0,681$

Selanjutnya dari hasil estimasi dengan metode OLS dapat dimodelkan sebagai berikut :

$$\hat{Y} = 2,211 + 0,095X_1 - 0,255X_2 + 0,451X_3 + 0,024X_4 + 0,091X_5 + 3,073X_6$$

Berdasarkan Tabel 4.5 diatas, dapat diketahui bahwa dengan menggunakan taraf signifikan ($\alpha = 5\%$), maka terdapat dua variabel yang berpengaruh secara signifikan terhadap presentase *annual parasite incidence* (API) malaria di Indonesia yaitu variabel presentase penduduk miskin (X_3) dan variabel rasio banyaknya puskesmas (X_6). Metode OLS menghasilkan *koefisien determinasi* (R^2) sebesar 0,681. Hal ini menunjukkan bahwa 68% penyebaran penyakit malaria di Indonesia mampu dijelaskan oleh model, sedangkan sisanya sebesar 32 % dijelaskan oleh peubah lain di luar model.

4.1.4 Estimasi Parameter Regresi *Robust LTS*

Metode regresi *robust LTS* dapat digunakan untuk mengatasi pengamatan yang mengandung pencilan (*outliers*). Pengolahan data dilakukan dengan menggunakan program SAS. Adapun syntax program SAS dapat dilihat pada lampiran 3. Berikut pada Tabel 4.6 dapat dilihat hasil estimasi dengan menggunakan metode Regresi *Robust Least Trimmed Square* (LTS).

Tabel 4.6 Estimasi Parameter Regresi *Robust LTS*

| Parameter | Estimasi | Std Error | <i>t-value</i> | <i>p-value</i> | <i>Keputusan</i> |
|-----------|----------|-----------|----------------|----------------|-------------------|
| β_0 | -1,324 | -1,324 | -1,24 | 0,231 | Gagal Tolak H_0 |
| β_1 | -0,001 | 0,003 | -0,29 | 0,774 | Gagal Tolak H_0 |
| β_2 | 0,018 | 0,013 | 1,41 | 0,177 | Gagal Tolak H_0 |
| β_3 | 0,035 | 0,008 | 4,04 | 0,000 | Tolak H_0 |
| β_4 | 0,029 | 0,005 | 5,43 | 0,000 | Tolak H_0 |
| β_5 | -0,014 | 0,005 | -2,53 | 0,022 | Tolak H_0 |
| β_6 | 0,069 | 0,052 | 1,32 | 0,204 | Gagal Tolak H_0 |

Ket : Signifikan pada $\alpha = 5\%$, $R^2 = 0,732$

Selanjutnya dari hasil estimasi dengan metode LTS dapat dimodelkan sebagai berikut :

$$\hat{Y} = -1,324 - 0,001X_1 + 0,018X_2 + 0,035X_3 + 0,029X_4 - 0,014X_5 + 0,069X_6$$

Berdasarkan tabel 4.6 diatas, dapat diketahui bahwa dengan menggunakan taraf signifikan ($\alpha = 5\%$), maka terdapat tiga variabel yang berpengaruh signifikan terhadap peningkatan presentase *annual parasite incidence* (API) malaria di Indonesia yaitu variabel presentase kemiskinan (X_3), variabel rasio tenaga dokter (X_4), dan variabel presentase akses air bersih (X_5). Nilai dari *koefisien determinasi* (R^2) sebesar 0,732. Hal ini menunjukkan bahwa 73% penyebaran penyakit malaria di Indonesia mampu dijelaskan oleh model, sedangkan sisanya sebesar 27% dijelaskan oleh peubah lain di luar model.

4.1.5 Estimasi Parameter Regresi Quantil

Berdasarkan identifikasi pola hubungan dan pengujian outliers menunjukkan bahwa terdapat pencilan (outliers) pada data. Oleh karena itu, digunakan analisis regresi quantil untuk mengestimasi faktor-faktor yang mempengaruhi peningkatan morbiditas atau *annual parasite incidence* (API) malaria. Langkah pertama dalam analisis regresi quantil adalah memilih nilai quantil yang ingin diestimasi yaitu quantil ($\tau = 0,5 ; 0,25 ; 0,50 ; 0,75$ dan $0,95$), selanjutnya melakukan estimasi regresi quantil untuk masing-masing koefisien regresi. Pengolahan data analisis dilakukan dengan menggunakan program R dapat dilihat pada lampiran 4. Berikut pada Tabel 4.7 dapat dilihat hasil estimasi dengan menggunakan metode regresi quantil.

Tabel 4.7 Estimasi Regresi Quantil Pada Quantil 5%

| Parameter | Koefisien | Std Error | <i>t-value</i> | <i>p-value</i> | <i>Keputusan</i> |
|-----------|-----------|-----------|----------------|----------------|-------------------|
| β_0 | 0,048 | 0,503 | 0,095 | 0,924 | Gagal Tolak H_0 |
| β_1 | 0,003 | 0,003 | 0,973 | 0,338 | Gagal Tolak H_0 |
| β_2 | -0,008 | 0,005 | -1,682 | 0,104 | Gagal Tolak H_0 |
| β_3 | 0,011 | 0,007 | 1,459 | 0,155 | Gagal Tolak H_0 |
| β_4 | -0,000 | 0,005 | -1,073 | 0,292 | Gagal Tolak H_0 |
| β_5 | 0,005 | 0,004 | 1,213 | 0,235 | Gagal Tolak H_0 |
| β_6 | 0,151 | 0,041 | 3,625 | 0,001 | Tolak H_0 |

Ket : Signifikan pada taraf $\alpha = 5\%$, $R^2 = 0,039$

Selanjutnya dari hasil estimasi regresi quantil pada quantil 5% dapat dimodelkan sebagai berikut :

$$Q_{0,05}(y|x) = 0,048 + 0,003X_1 - 0,008X_2 + 0,011X_3 - 0,006X_4 + 0,005X_5 + 0,151X_6$$

Berdasarkan Tabel 4.7, dapat diketahui bahwa dengan menggunakan taraf signifikan ($\alpha = 5\%$) pada quantil ($\tau = 0,05$), maka dapat disimpulkan bahwa variabel yang berpengaruh secara signifikan terhadap peningkatan *annual parasite incidence* (API) malaria di Indonesia adalah variabel rasio keberadaan puskesmas (X_6). Pada model regresi quantil ($\tau = 0,05$), diperoleh *koefisien determinasi* (R^2) sebesar 0,039. Hal ini menunjukkan bahwa 3% kasus penyebaran penyakit malaria di Indonesia mampu dijelaskan oleh model, sedangkan sisanya sebesar 97% dijelaskan oleh peubah lain di luar model.

Tabel 4.8 Estimasi Regresi Quantil Pada Quantil 25%

| Parameter | Koefisien | Std Error | <i>t-value</i> | <i>p-value</i> | <i>Keputusan</i> |
|-----------|-----------|-----------|----------------|----------------|-------------------|
| β_0 | 6,139 | 1,810 | 3,391 | 0,002 | Tolak H_0 |
| β_1 | 0,009 | 0,013 | 0,709 | 0,483 | Gagal Tolak H_0 |
| β_2 | -0,075 | 0,019 | -3,945 | 0,000 | Tolak H_0 |
| β_3 | 0,001 | 0,027 | 0,068 | 0,945 | Gagal Tolak H_0 |
| β_4 | 0,013 | 0,020 | 0,637 | 0,529 | Gagal Tolak H_0 |
| β_5 | 0,000 | 0,015 | 0,031 | 0,975 | Gagal Tolak H_0 |
| β_6 | 0,091 | 0,150 | 0,606 | 0,549 | Gagal Tolak H_0 |

Ket : Signifikan pada $\alpha = 5\%$, $R^2 = 0,074$

Selanjutnya dari hasil estimasi regresi quantil pada quantil 25% dapat dimodelkan sebagai berikut :

$$Q_{0,25}(y|x) = 6,139 + 0,009X_1 - 0,075X_2 + 0,001X_3 + 0,013X_4 + 0,000X_5 + 0,091X_6$$

Berdasarkan Tabel 4.8, dapat diketahui bahwa dengan menggunakan taraf signifikan ($\alpha = 5\%$) pada quantil ($\tau = 0,25$), maka dapat disimpulkan bahwa variabel yang berpengaruh secara signifikan terhadap peningkatan *annual parasite incidence* (API) malaria di Indonesia yaitu variabel presentase rumah layak huni (X_2). Pada model regresi quantil ($\tau = 0,25$), diperoleh *koefisien determinasi* (R^2) sebesar 0,074. Hal ini menunjukkan bahwa 7% kasus penyebaran penyakit malaria di Indonesia mampu dijelaskan oleh model, sedangkan sisanya sebesar 93 % dijelaskan oleh peubah lain di luar model.

Tabel 4.9 Estimasi Regresi Quantil Pada Quantil 50%.

| Parameter | Koefisien | Std Error | <i>t-value</i> | <i>p-value</i> | <i>Keputusan</i> |
|-----------|-----------|-----------|----------------|----------------|-------------------|
| β_0 | 13,704 | 2,627 | 5,216 | 0,000 | Tolak H_0 |
| β_1 | 0,027 | 0,018 | 1,472 | 0,152 | Gagal Tolak H_0 |
| β_2 | -0,186 | 0,027 | -6,749 | 0,000 | Tolak H_0 |
| β_3 | 0,034 | 0,039 | 0,862 | 0,396 | Gagal Tolak H_0 |
| β_4 | -0,014 | 0,029 | -0,493 | 0,625 | Gagal Tolak H_0 |
| β_5 | 0,021 | 0,022 | 0,957 | 0,346 | Gagal Tolak H_0 |
| β_6 | 0,588 | 0,218 | 2,698 | 0,011 | Tolak H_0 |

Ket : Signifikan pada taraf $\alpha = 5\%$, $R^2 = 0,184$

Selanjutnya dari hasil estimasi regresi quantil pada quantil 50% dapat dimodelkan sebagai berikut :

$$Q_{0,50}(y|x) = 13,704 + 0,027X_1 - 0,186X_2 + 0,034X_3 - 0,0146X_4 + 0,021X_5 + 0,588X_6$$

Berdasarkan Tabel 4.9, dapat diketahui bahwa dengan menggunakan taraf signifikan ($\alpha = 5\%$) pada quantil ($\tau = 0,50$), maka dapat disimpulkan bahwa variabel yang berpengaruh secara signifikan terhadap peningkatan *annual parasite incidence* (API) malaria di Indonesia yaitu variabel presentase rumah layak huni (X_2) dan variabel rasio banyaknya puskesmas (X_6). Pada model regresi quantil ($\tau = 0,50$), diperoleh koefisien determinasi (R^2) sebesar 0,184. Hal ini menunjukkan bahwa 18% kasus penyebaran penyakit malaria di Indonesia mampu dijelaskan oleh model, sedangkan sisanya sebesar 82% dijelaskan oleh peubah lain di luar model.

Tabel 4.10 Estimasi Regresi Quantil Pada Quantil 75%.

| Parameter | Koefisien | Std Error | <i>t-value</i> | <i>p-value</i> | <i>Keputusan</i> |
|-----------|-----------|-----------|----------------|----------------|-------------------|
| β_0 | 44,537 | 9,161 | 4,861 | 0,000 | Tolak H_0 |
| β_1 | 0,116 | 0,065 | 1,767 | 0,088 | Gagal Tolak H_0 |
| β_2 | -0,616 | 0,096 | -6,387 | 0,000 | Tolak H_0 |
| β_3 | 0,202 | 0,138 | 1,456 | 0,156 | Gagal Tolak H_0 |
| β_4 | 0,043 | 0,103 | 0,424 | 0,674 | Gagal Tolak H_0 |
| β_5 | 0,036 | 0,078 | 0,469 | 0,642 | Gagal Tolak H_0 |
| β_6 | 1,545 | 0,760 | 2,031 | 0,052 | Tolak H_0 |

Ket : Signifikan pada taraf $\alpha = 5\%$, $R^2 = 0,355$

Selanjutnya dari hasil estimasi regresi quantil pada quantil 75% dapat dimodelkan sebagai berikut :

$$Q_{0,75}(y|x) = 44,537 + 0,116X_1 - 0,616X_2 + 0,202X_3 + 0,043X_4 + 0,036X_5 + 1,545X_6$$

Berdasarkan Tabel 4.10, dapat diketahui bahwa dengan menggunakan taraf signifikan ($\alpha = 5\%$) pada quantil ($\tau = 0,75$), maka dapat disimpulkan bahwa variabel yang berpengaruh secara signifikan terhadap peningkatan *annual parasite incidence* (API) malaria di Indonesia yaitu variabel presentase rumah layak huni (X_2) dan variabel rasio puskesmas (X_6). Pada model regresi quantil ($\tau = 0,75$), diperoleh *koefisien determinasi* (R^2) sebesar 0,355. Hal ini menunjukkan bahwa 35% kasus penyebaran penyakit malaria di Indonesia mampu dijelaskan oleh model, sedangkan sisanya sebesar 65% dijelaskan oleh peubah lain di luar model.

Tabel 4.11 Estimasi Regresi Quantil Pada Quantil 95%.

| Parameter | Koefisien | Std Error | <i>t-value</i> | <i>p-value</i> | <i>Keputusan</i> |
|-----------|-----------|-----------|----------------|----------------|------------------|
| β_0 | 11,934 | 5,266 | 2,660 | 0,031 | Tolak H_0 |
| β_1 | 0,170 | 0,037 | 4,502 | 0,000 | Tolak H_0 |
| β_2 | -0,394 | 0,055 | -7,108 | 0,000 | Tolak H_0 |
| β_3 | 0,407 | 0,079 | 5,102 | 0,000 | Tolak H_0 |
| β_4 | -0,275 | 0,059 | -4,637 | 0,000 | Tolak H_0 |
| β_5 | 0,161 | 0,045 | 3,579 | 0,001 | Tolak H_0 |
| β_6 | 5,704 | 0,437 | 13,041 | 0,000 | Tolak H_0 |

Ket : Signifikan pada $\alpha = 5\%$, $R^2 = 0,832$

Selanjutnya dari hasil estimasi regresi quantil pada quantil 95% dapat dimodelkan sebagai berikut :

$$Q_{0,95}(y|x) = 11,934 + 0,170X_1 - 0,394X_2 + 0,407X_3 - 0,275X_4 + 0,161X_5 + 5,704X_6$$

Berdasarkan Tabel 4.11, dapat diketahui bahwa dengan menggunakan taraf signifikan ($\alpha = 5\%$) pada quantil ($\tau = 0,95$), maka dapat disimpulkan bahwa variabel yang berpengaruh secara signifikan terhadap peningkatan *annual parasite incidence* (API) malaria di Indonesia yaitu variabel presentase akses sanitasi layak (X_1), variabel presentase rumah layak huni (X_2), variabel presentase kemiskinan (X_3), variabel rasio banyaknya tenaga dokter (X_4) dan variabel presentase akses air bersih (X_5) serta variabel rasio banyaknya puskesmas (X_6). Pada quantil ($\tau = 0,95$), diperoleh *koefisien determinasi* atau R^2 sebanyak 0,832. Hal ini menunjukkan bahwa 83% kasus penyebaran penyakit malaria di Indonesia mampu dijelaskan oleh model, sedangkan sisanya sebesar 17% dijelaskan oleh peubah lain di luar model. Selanjutnya untuk melihat pola data hasil estimasi regresi quantil, dapat dilihat pada lampiran 9.

4.2 Pemilihan Model Terbaik

Setelah mendapatkan model dari metode *Ordinary Least Square* (OLS), metode regresi *robust* LTS dan metode regresi kuantil, maka langkah selanjutnya adalah menentukan model terbaik dengan melihat nilai *koefisien determinasi* (R^2) yang paling tinggi. Berikut pada Tabel 4.12 dapat lihat hasil *koefisien determinasi* (R^2) dari metode Regresi Kuantil dan Regresi *Robust Least Trimmed Square* (LTS) serta *Ordinary Least Square* (OLS).

Tabel 4.12 *Koefisien Determinasi* (R^2)

| Model | <i>Koefisien Determinasi</i> (R^2) |
|-------------------|--|
| $Q_{(0,05)}$ | 0,039 |
| $Q_{(0,25)}$ | 0,074 |
| $Q_{(0,50)}$ | 0,184 |
| $Q_{(0,75)}$ | 0,355 |
| $Q_{(0,95)}$ | 0,832* |
| OLS | 0,681 |
| <i>Robust</i> LTS | 0,732 |

Ket : *) *Koefisien Determinasi* Tertinggi

Koefisien determinasi (R^2) merupakan besaran yang digunakan untuk mengukur kelayakan model regresi dan menunjukkan besar kontribusi variabel x terhadap perubahan variabel y . Semakin tinggi nilai R^2 mendekati 1 semakin baik model regresi yang terbentuk. Berdasarkan Tabel 4.13 dapat dilihat bahwa pada model kuantil $Q_{(0,95)}$, memiliki nilai *koefisien determinasi* (R^2) sebanyak 0,832 lebih besar dibandingkan dengan model kuantil yang lain dan lebih besar pula jika dibandingkan dengan model regresi *robust* LTS dengan R^2 sebanyak 0,732 serta model *Ordinary Least Square* (OLS)

dengan R^2 sebanyak 0,681. Hal ini menunjukkan bahwa model regresi kuantil $Q_{(0,95)}$ merupakan model terbaik.

4.3 Faktor-faktor yang Berpengaruh Terhadap Penyebaran Penyakit Malaria di Indonesia

Berdasarkan model terbaik yaitu model estimasi regresi kuantil $Q_{(0,95)}$, menunjukkan bahwa variabel presentase akses sanitasi layak (X_1) berpengaruh positif dan signifikan terhadap *annual parasite incidence* (API) malaria. Pada kuantil $Q_{(0,95)}$ menunjukkan bahwa koefisien variabel presentase akses sanitasi layak (X_1) sebanyak 0,17048. Hal ini berarti bahwa daerah dengan akses sanitasi layak tinggi cenderung memiliki tingkat kerentanan terhadap penyakit malaria. Hal ini tidak sesuai dengan hipotesis yang diharapkan yaitu akses sanitasi layak dapat menurunkan morbiditas penyakit malaria. Ketidaksesuaian hasil ini, diduga disebabkan karena kurang bervariasinya data morbiditas atau API penderita penyakit malaria.

Selanjutnya pada kuantil $Q_{(0,95)}$ menunjukkan bahwa variabel presentase rumah layak huni (X_2), berpengaruh negatif dan signifikan terhadap API malaria. Koefisien presentase rumah layak huni (X_2) sebanyak minus 0,39447. Artinya jika presentase rumah layak huni di suatu daerah naik sebanyak 1 persen, maka akan menurunkan tingkat *annual parasite incidence* (API) malaria sebanyak 0,39 per 1000 penduduk. Hal ini sesuai dengan hipotesis yang diharapkan bahwa kondisi rumah layak huni dapat mengurangi *annual parasite incidence* (API) malaria. Kondisi rumah dengan dinding yang terbuat dari kayu memungkinkan lebih banyak lubang untuk masuknya nyamuk. Demikian pula lantai rumah dari tanah menyebabkan kondisi rumah menjadi lembab sehingga menjadi tempat yang ideal bagi nyamuk.

Selanjutnya pada kuantil $Q_{(0,95)}$ menunjukkan bahwa variabel presentase kemiskinan (X_3) berpengaruh positif dan signifikan terhadap *annual parasite incidence* (API) malaria. Koefisien presentase kemiskinan (X_3) sebanyak 0,40715. Artinya jika presentase kemiskinan di suatu daerah naik sebanyak 1 persen, maka akan meningkatkan

tingkat *annual parasite incidence* (API) malaria sebanyak 0,40 per 1000 penduduk. Hal ini menunjukkan bahwa ada kecenderungan daerah dengan presentase kemiskinan relatif tinggi akan diikuti tingkat *annual parasite incidence* (API) malaria tinggi pula. Hal ini sesuai dengan hipotesis yang diharapkan bahwa tingkat kemiskinan dapat meningkatkan morbiditas malaria. Hubungan kemiskinan dengan tingkat *annual parasite incidence* (API) malaria dapat dilihat dari turunan variabel kemiskinan, misalnya penggunaan kelambu dan pekerjaan. Pada banyak penelitian disebutkan bahwa penduduk yang menggunakan kelambu resiko terkena malaria lebih rendah dibanding yang tidak memakai. Demikian pula ada beberapa pekerjaan beresiko seperti nelayan dan pertanian yang lebih rentan terkena malaria dibanding jenis pekerjaan lainnya.

Selanjutnya pada kuantil $Q_{(0,95)}$ menunjukkan bahwa variabel rasio tenaga dokter (X_4) berpengaruh negatif dan signifikan terhadap *annual parasite incidence* (API) malaria. Koefisien variabel rasio tenaga dokter (X_4) sebanyak minus 0,27588. Artinya jika rasio tenaga dokter di suatu daerah naik sebanyak 1, maka akan menurunkan tingkat *annual parasite incidence* (API) malaria sebanyak 0,27 per 1000 penduduk. Artinya ada kecenderungan daerah dengan rasio tenaga dokter relatif tinggi akan diikuti tingkat *annual parasite incidence* (API) malaria yang relatif lebih rendah. Hal ini sesuai dengan hipotesis yang diharapkan bahwa tenaga kesehatan dapat menurunkan tingkat *annual parasite incidence* (API) malaria. Hubungan tenaga kesehatan terhadap kerentanan penyakit malaria dapat dijelaskan melalui kegiatan pelayanan kesehatan. Tenaga kesehatan dapat melakukan penyuluhan, penyemprotan, dan pengobatan sehingga dapat mempengaruhi *annual parasite incidence* (API) malaria di suatu daerah.

Selanjutnya pada kuantil $Q_{(0,95)}$ menunjukkan bahwa variabel presentase akses air bersih (X_5) berpengaruh positif dan signifikan terhadap *annual parasite incidence* (API) malaria. Koefisien variabel presentase akses air bersih (X_5) sebanyak 0,16171. Artinya daerah dengan akses air bersih tinggi cenderung memiliki tingkat kerentanan terhadap penyakit malaria. Hal ini tidak sesuai dengan hipotesis yang diharapkan yaitu presentase akses air bersih dapat menurunkan tingkat *annual parasite incidence* (API) malaria. Ketidaksesuaian hasil ini diduga disebabkan karena kurang bervariasinya data *annual*

parasite incidence (API) malaria. Hal lainnya diduga karena akses air bersih pada daerah yang rendah tingkat *annual parasite incidence* (API) malariannya adalah air leding atau air mengalir lainnya. Nyamuk penyebab penyakit malaria biasanya berada pada air tergenang seperti seperti air sumur yang tidak lindungi dan memiliki presentase penggunaannya rendah di Indonesia. Sehingga, dampak dari pengaruh air bersih terhadap tingkat *annual parasite incidence* (API) malaria menjadi tidak bermakna.

Selanjutnya pada kuantil $Q_{(0,95)}$ menunjukkan bahwa variabel rasio banyaknya puskesmas (X_6) berpengaruh positif dan signifikan terhadap *annual parasite incidence* (API) malaria. Koefisien rasio banyaknya puskesmas (X_6) sebanyak 5,7049. Artinya bahwa ada kecenderungan daerah dengan rasio banyaknya puskesmas relatif tinggi, akan diikuti dengan tingkat *annual parasite incidence* (API) malaria yang relatif tinggi pula. Hal ini tidak sesuai dengan hipotesis yang diharapkan bahwa banyaknya puskesmas dapat menekan tingkat *annual parasite incidence* (API) malaria. Ketidaksesuaian hubungan ini, diduga disebabkan karena jumlah puskesmas yang banyak tidak menjamin akses ke sarana kesehatan yang lebih baik bagi masyarakat. Hal ini, dimungkinkan pada kawasan Timur Indonesia, sarana puskesmas yang hanya terletak di Ibu Kota Kabupaten dan beberapa Ibu Kota Kecamatan sehingga persebarannya tidak merata. Selain itu, jumlah puskesmas yang banyak namun tidak diikuti dengan pelayanan yang maksimal akan menyebabkan ketimpangan dalam akses sarana kesehatan bagi masyarakat.

(Halaman Ini Sengaja Dikosongkan)

BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil dan pembahasan, maka diperoleh kesimpulan dari penelitian ini adalah sebagai berikut :

1. a. Model Regresi Quantil pada faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia dapat dinyatakan dengan model regresi quantil ($\tau = 0,05 ; 0,25 ; 0,50 ; 0,75 ; 0,95$) secara berturut-turut adalah sebagai berikut :

$$Q_{0,05}(y|x) = 0,048 + 0,003X_1 - 0,008X_2 + 0,011X_3 - 0,006X_4 + 0,005X_5 + 0,151X_6$$

$$Q_{0,25}(y|x) = 6,139 + 0,009X_1 - 0,075X_2 + 0,001X_3 + 0,013X_4 + 0,000X_5 + 0,091X_6$$

$$Q_{0,50}(y|x) = 13,704 + 0,027X_1 - 0,186X_2 + 0,034X_3 - 0,014X_4 + 0,021X_5 + 0,588X_6$$

$$Q_{0,75}(y|x) = 44,537 + 0,116X_1 - 0,616X_2 + 0,202X_3 + 0,043X_4 + 0,036X_5 + 1,545X_6$$

$$Q_{0,95}(y|x) = 11,934 + 0,170X_1 - 0,394X_2 + 0,407X_3 - 0,275X_4 + 0,161X_5 + 5,704X_6$$

- b. Model Regresi *Robust Least Trimmed Square* (LTS) pada faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia adalah sebagai berikut :

$$\hat{Y} = -1,324 - 0,001X_1 + 0,018X_2 + 0,035X_3 + 0,029X_4 - 0,014X_5 + 0,069X_6$$

- c. Model metode *Ordinary Least Square* (OLS) pada faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia adalah sebagai berikut :

$$\hat{Y} = 2,211 + 0,0951X_1 - 0,255X_2 + 0,451X_3 + 0,024X_4 + 0,091X_5 + 3,073X_6$$

2. Hasil perbandingan metode estimasi yang diperoleh yaitu metode Regresi Quantil lebih baik dibandingkan dengan metode Regresi *Robust Least Trimmed Square* (LTS) dan metode *Ordinary Least Square* (OLS) pada kasus faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia. Hal ini, ditunjukkan dengan nilai R^2 dari model Regresi Quantil sebesar 0,832, model Regresi *Robust Least Trimmed Square* (LTS) sebesar 0,732 dan model *Ordinary Least Square* (OLS) sebesar 0,681.
3. Faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia dengan menggunakan model terbaik, pada model regresi kuantil dengan $\tau = 0,95$ adalah faktor rumah layak huni, faktor penduduk miskin dan faktor tenaga dokter. Oleh karena itu, Pemerintah beserta seluruh elemen masyarakat perlu mengevaluasi dan melakukan proses pencegahan agar terhindar dari penyebaran penyakit malaria.

5.2 Saran

Berdasarkan analisis data dan kesimpulan yang telah diperoleh dalam penelitian ini, maka dapat diberikan saran sebagai berikut :

1. Penelitian ini menggunakan metode simpleks untuk data dengan jumlah data kecil, pada penelitian selanjutnya bisa dilakukan untuk jumlah data besar dan memiliki variabel yang banyak.
2. Penelitian ini membandingkan metode Regresi Quantil dengan Regresi *Robust Least Trimmed Square* (LTS). Penelitian selanjutnya bisa menggunakan perbandingan antara metode regresi kuantil dengan penaksir regresi *robust* lainnya, seperti *Least Median of Squares* (LMS), penaksir M, penaksir S dan penaksir MM dan lain sebagainya.

DAFTAR PUSTAKA

- Aulia, Triastuti dan Ichsan. (2016). *Hubungan Perilaku Hidup Bersih dan Sehat Dengan Kejadian Malaria Di Wilayah Kerja Puskesmas Kota Sorong, Propinsi Papua Barat Tahun 2015*. Fakultas Kedokteran, Universitas Muhammadiyah Surakarta.
- Badan Pusat Statistika. (2014). *Hasil Survei Sosial Ekonomi Nasional Tahun 2014*. Jakarta
- Budiantara, I.N. (2011), *Penelitian Bidang Regresi Spline Menuju Terwujudnya Penelitian Statistika yang Mandiri dan Berkarakter*. Seminar Nasional FMIPA Universitas Pendidikan Ganesha.
- Chen, C. dan Wei. (2005). Computational Issues for Quantile Regression. *The Indian Journal of Statistics*. Vol. 67, hal. 399-417.
- Davino, Furno dan Vistocco, (2014). *Quantile regression (Theory and application)*. Standford Weisberg.
- Datukramat, Mayulu dan Masi. (2013). *Menganalisis Hubungan Anantara Sanitasi Lingkungan Dengan Kejadian Malaria Pada Murid Sekolah Dasar Di Kabupaten Bolaang Mongondow Utara*. journal keperawatan (e-Kp) Fakultas Kedokteran Universitas Sam Ratulangi.
- Febria dan Widodo, E. (2016). *Perbandingan Kemampuan Regresi Quantil Median dan Transformasi Box-Cox Dalam Menangani Heteroskedastisitas*. Jurnal Prosiding, Program Studi Statistika FMIPA Universitas Islam Indonesia.
- Ghozali, Imam. 2013. *Aplikasi Analisis Multivariat dengan Program SPSS*. Edisi ketujuh. Semarang
- Gob, S. C. dan Knight, K. (2009). “*Nonstandard Quantile-Regression Inference*”. *Econometric Theory*. Vol. 25, hal. 1415-1432

- Gujarati, D. N. (2004). *Basic Econometrics, 4th Edition*. New York: McGrahill. Co.
- Hardle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press. New York.
- Hastie, T. J. dan Tibshirani, R. J. (1990). *Generalized Additive Models*. Chapman and Hall. New York. London.
- Indriani dan Yuyun. (2010). *Analisis faktor-faktor yang mempengaruhi angka kejadian penyakit tropis malaria dengan metode klasifikasi pohon (Studi kasus Kabupaten Sorong Selatan, Provinsi Papua Barat)*. Jurnal, Institut Teknologi Sepuluh Nopember
- Kementrian Kesehatan Republik Indonesia. (2014). *Profil kesehatan Indonesia*, Pusat data dan Informasi Kementrian Kesehatan. Jakarta.
- Jullen, Nyorong dan Ibnu, I. F. (2014). Perilaku pencarian pengobatan masyarakat terhadap penyakit malaria di Kelurahan Remu Utara, Distrik Sorong Kota Sorong Papua Barat. Bagian PKIP, Fakultas Kesehatan Masyarakat Universitas Hasanuddin.
- Koenker, R. dan Machado, J. A. F. (1999). *Goodness of fit and Related Inference Process for Quantile Regression*. Journal of the American statistical Association. Vol. 94, no 448, hal 1296-1310.
- Koenker, Roger. (2005). "Quantile Regression". *Ekonometric society monographs*. Cambridge University, New York.
- Lin W, Zongwu, Li. (2015). *Optimal smoothing in nonparametric conditional quantile derivative function estimation*. Journal of Econometrics ScienceDirect vol. 188, hal. 502-513.
- Mirontoneng, A.R, Ismanto, A.Y dan Malara, R. (2014). *Analisis faktor-faktor yang berhubungan kejadian malaria pada anak di wilayah kerja PKM*

Tona kecamatan tahuna kabupaten Sangihe. Program Studi Ilmu Keperawatan Fakultas kedokteran Universitas Sam Ratulangi Manado.

Myers, R. H. (1990). *Classical and Modern Regression with Application*. Boston: PWS.

Rahmawati, R, Widiarti dan Novianti, P. (2011). Regresi Quantil (Studi Kasus pada Data Suhu Harian). *Peran dan Implementasi Statistika dalam Analisis Finansial dan Pengambilan Keputusan Bisnis: Prosiding Seminar Nasional Statistika*. Diselenggarakan oleh Program Studi Statistika, Undip, 21 mei 2011. Semarang.

Rita R, Widiarti, dan Novianti P. (2011). *Analisis Regresi Quantil (Studi Kasus Pada Data Suhu Harian)*. Sekolah Pascasarjana Departemen Statistika Institut Pertanian Bogor.

Ronald E. Walpole. (1993). *Pengantar Statistika*. Edisi ke-3. Gramdia. Jakarta

Sari, Ambarita dan Sitorus. (2013). *Akses Pelayanan Kesehatan Dan Kejadian Malaria di Provinsi Bengkulu*. Jurnal Media Litbangkes Vol 23 No. 4, hal 158-164.

Susilowati, P, A. (2013). *Analisis regresi pada prevalensi malaria di Provinsi Maluku Utara, Malauku, Papua Barat dan Papua dengan faktor yang mempengaruhinya*. Skripsi, Jurusan Statistika Institut Teknologi Sepuluh Nopember.

Tangkilisan, M. Sorisi, A dan Tuda, J. (2015). *Peran sarana pelayanan kesehatan terhadap kejadian malaria di Kecamatan Silian Raya Kabupaten Minahasa Tenggara*. Jurnal *e-Biomedik (eBm)*, Volume 3, Nomor 1. Bagian Fisika kedokteran Universitas Sam Ratulangi manado.

Vendy dan Ismaini, Z. (2014). *Analisis IPM di Pulau Jawa Menggunakan Analisis Regresi Quantil*. Tesis. Jurausan Statistika, Fakultas Matematika

dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember, Surabaya.

Werner, Mark. (2003). *Identification of Multivariate Outliers in Large Data Sets*, University of Colorado at Denver.

Willems, G. and Aelst, S.V. (2005). *Fast and Robust Bootstrap for LTS* Computational Statistics and Data Analysis, No. 48, hal. 703-715.

LAMPIRAN

Lampiran 1.

Data penderita penyakit malaria dan faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia pada tahun 2015.

| Propinsi | Y | X1 | X2 | X3 | X4 | X5 | X6 |
|---------------------------|------|-------|-------|-------|-------|-------|------|
| Aceh | 0,08 | 54,68 | 88,60 | 17,11 | 30,99 | 61,23 | 2,03 |
| Sumatera Utara | 0,49 | 67,89 | 91,94 | 10,79 | 23,85 | 71,41 | 1,23 |
| Sumatera Barat | 0,14 | 45,02 | 91,90 | 6,71 | 19,53 | 66,58 | 1,52 |
| Riau | 0,10 | 51,30 | 93,89 | 8,82 | 16,96 | 74,24 | 1,00 |
| Jambi | 0,47 | 58,21 | 92,91 | 9,12 | 18,73 | 62,75 | 1,55 |
| Sumatera Selatan | 0,31 | 61,30 | 91,83 | 13,77 | 14,89 | 65,16 | 1,20 |
| Bengkulu | 2,03 | 39,22 | 89,80 | 17,16 | 20,11 | 41,08 | 2,88 |
| Lampung | 0,49 | 44,83 | 91,29 | 13,53 | 13,23 | 55,06 | 1,08 |
| Kepulauan Bangka Belitung | 1,08 | 80,80 | 97,77 | 4,83 | 24,40 | 68,03 | 1,35 |
| Kepulauan Riau | 0,35 | 71,97 | 97,97 | 5,78 | 26,76 | 84,12 | 1,09 |
| DKI Jakarta | 0,00 | 89,28 | 99,21 | 3,61 | 25,99 | 93,40 | 1,00 |
| Jawa Barat | 0,00 | 59,43 | 94,75 | 9,57 | 10,95 | 67,20 | 0,67 |
| Jawa Tengah | 0,06 | 67,20 | 94,97 | 13,32 | 13,86 | 73,63 | 0,78 |
| DI Yogyakarta | 0,03 | 86,31 | 98,72 | 13,16 | 28,24 | 80,99 | 0,99 |
| Jawa Timur | 0,00 | 63,48 | 95,51 | 12,28 | 11,47 | 76,64 | 0,74 |
| Banten | 0,00 | 67,04 | 93,12 | 5,75 | 11,69 | 67,68 | 0,58 |
| Bali | 0,00 | 85,46 | 98,66 | 5,25 | 27,79 | 91,27 | 0,87 |
| Nusa Tenggara Barat | 0,42 | 63,72 | 93,69 | 16,54 | 11,44 | 71,70 | 0,98 |
| Nusa Tenggara Timur | 7,04 | 23,90 | 56,43 | 22,58 | 13,02 | 62,72 | 2,17 |
| Kalimantan Barat | 0,13 | 39,78 | 86,97 | 8,44 | 14,11 | 68,39 | 1,30 |
| Kalimantan Tengah | 0,42 | 35,88 | 88,46 | 5,91 | 20,64 | 57,01 | 3,27 |
| Kalimantan Selatan | 0,68 | 60,13 | 93,36 | 4,72 | 17,07 | 62,23 | 1,73 |
| Kalimantan Timur | 0,46 | 68,83 | 96,76 | 6,10 | 25,04 | 78,13 | 1,52 |
| Kalimantan Utara | 0,03 | 48,80 | 96,13 | 6,32 | 31,62 | 84,59 | 2,29 |
| Sulawesi Utara | 0,88 | 66,79 | 92,77 | 8,98 | 39,18 | 71,53 | 2,33 |

| | | | | | | | |
|-------------------|-------|-------|-------|-------|-------|-------|------|
| Sulawesi Tengah | 0,68 | 55,37 | 85,61 | 14,07 | 18,08 | 61,49 | 1,97 |
| Sulawesi Selatan | 0,10 | 72,36 | 92,97 | 10,12 | 16,91 | 72,07 | 1,58 |
| Sulawesi Tenggara | 0,41 | 63,62 | 91,36 | 3,74 | 15,36 | 77,19 | 3,23 |
| Gorontalo | 0,57 | 54,96 | 86,53 | 18,16 | 21,97 | 66,47 | 2,46 |
| Sulawesi Barat | 0,17 | 51,21 | 81,77 | 11,90 | 12,63 | 53,89 | 2,20 |
| Maluku | 5,81 | 60,02 | 84,55 | 19,36 | 18,86 | 64,96 | 3,54 |
| Maluku Utara | 2,77 | 59,17 | 84,13 | 6,22 | 20,73 | 60,07 | 3,28 |
| Papua Barat | 31,29 | 62,81 | 90,65 | 25,73 | 24,90 | 68,85 | 5,20 |
| Papua | 31,93 | 28,04 | 49,23 | 28,4 | 23,97 | 51,27 | 3,75 |

Lampiran 2. Hasil Pengujian Data Pencilan (*Outlier*)

| No | Propinsi | <i>DFITS</i> |
|-----------|---------------------------|---------------------|
| 1. | Aceh | -0,72810 |
| 2. | Sumatera Utara | -0,02532 |
| 3. | Sumatera Barat | 0,24903 |
| 4. | Riau | 0,23589 |
| 5. | Jambi | 0,07665 |
| 6. | Sumatera Selatan | -0,03046 |
| 7. | Bengkulu | -0,34254 |
| 8. | Lampung | 0,28016 |
| 9. | Kepulauan Bangka Belitung | 0,36689 |
| 10. | Kepulauan Riau | 0,16117 |
| 11. | DKI Jakarta | 0,10651 |
| 12. | Jawa Barat | 0,28231 |
| 13. | Jawa Tengah | 0,02133 |
| 14. | DI Yogyakarta | -0,28273 |
| 15. | Jawa Timur | 0,10161 |
| 16. | Banten | 0,42992 |
| 17. | Bali | 0,09336 |
| 18. | Nusa Tenggara Barat | -0,12657 |
| 19. | Nusa Tenggara Timur | -1,69655 |
| 20. | Kalimantan Barat | 0,20602 |
| 21. | Kalimantan Tengah | -0,16756 |
| 22. | Kalimantan Selatan | 0,26433 |
| 23. | Kalimantan Timur | 0,07614 |
| 24. | Kalimantan Utara | -0,14243 |
| 25. | Sulawesi Utara | -0,43172 |
| 26. | Sulawesi Tengah | -0,16607 |
| 27. | Sulawesi Selatan | -0,10936 |
| 28. | Sulawesi Tenggara | -1,28412 |

| | | |
|-----|----------------|----------------|
| 29. | Gorontalo | -0,46628 |
| 30. | Sulawesi Barat | -0,33163 |
| 31. | Maluku | -0,67649 |
| 32. | Maluku Utara | -0,42709 |
| 33. | Papua Barat | 6,41416 |
| 34. | Papua | 4,55268 |

Lampiran 3. Ouput Koefisien Korelasi dan Nilai Variance Inflation Factor (VIF)

| Correlation: Y, X1, X2, X3, X4, X5, X6 | | | | | | |
|--|-----------------|-----------------|-----------------|-----------------|----------------|-----------------|
| | Y | X1 | X2 | X3 | X4 | X5 |
| X1 | -0.298 0.087 | | | | | |
| X2 | -0.594 0.000 | 0.717 0.000 | | | | |
| X3 | 0.700 0.000 | -0.446 0.008 | -0.667 0.000 | | | |
| X4 | 0.134 0.449 | 0.289 0.097 | 0.149 0.400 | -0.113 0.524 | | |
| X5 | -0.255 0.145 | 0.702 0.000 | 0.540 0.001 | -0.437 0.010 | 0.317 0.067 | |
| X6 | 0.683 0.000 | -0.395 0.021 | -0.471 0.005 | 0.548 0.001 | 0.247 0.158 | -0.409 0.016 |

=====

Nilai Variance Inflation Factor (VIF)

Coefficients

| Term | Coef | SE Coef | T-Value | P-Value | VIF |
|----------|--------|---------|---------|---------|------|
| Constant | 2.2 | 12.9 | 0.17 | 0.865 | |
| X1 | 0.0951 | 0.0929 | 1.02 | 0.315 | 3.05 |
| X2 | -0.256 | 0.136 | -1.88 | 0.071 | 3.01 |
| X3 | 0.451 | 0.196 | 2.31 | 0.029 | 2.20 |
| X4 | 0.024 | 0.146 | 0.16 | 0.870 | 1.48 |
| X5 | 0.091 | 0.111 | 0.82 | 0.417 | 2.24 |
| X6 | 3.07 | 1.07 | 2.86 | 0.008 | 1.99 |

Lampiran 4. Syntax Metode *Ordinary Least Square* (OLS)

```
k1=lm(Y~X1+X2+X3+X4+X5+X6, data=data)

summary(k1)

Ouput OLS

Coefficients:

              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.21114    12.91636   0.171  0.86535
X1            0.09513     0.09285   1.025  0.31467
X2           -0.25561     0.13609  -1.878  0.07118 .
X3            0.45146     0.19567   2.307  0.02894 *
X4            0.02404     0.14589   0.165  0.87032
X5            0.09131     0.11080   0.824  0.41714
X6            3.07310     1.07277   2.865  0.00798 **

Residual standard error: 4.692 on 27 degrees of freedom

Multiple R-squared:  0.681,    Adjusted R-squared:
 0.6101
```

Lampiran 5. Syntax Metode Regresi *Robust Least Trimmed Square (LTS)* dengan Menggunakan Program SAS

```

title 'data';

proc iml;

print 'data';

data1=

{0.08    54.68    88.6    17.11    30.99    61.23    2.03,
0.49    67.89    91.94    10.79    23.85    71.41    1.23,
0.14    45.02    91.9    6.71    19.53    66.58    1.52,
0.1     51.3     93.89    8.82    16.96    74.24    1,00
0.47    58.21    92.91    9.12    18.73    62.75    1.55,
0.31    61.3     91.83    13.77   14.89    65.16    1.2,
2.03    39.22    89.8     17.16   20.11    41.08    2.88,
0.49    44.83    91.29    13.53   13.23    55.06    1.08,
1.08    80.8     97.77    4.83    24.4     68.03    1.35,
0.35    71.97    97.97    5.78    26.76    84.12    1.09,
0       89.28    99.21    3.61    25.99    93.4     1,00
0       59.43    94.75    9.57    10.95    67.2     0.67,
0.06    67.2     94.97    13.32   13.86    73.63    0.78,
0.03    86.31    98.72    13.16   28.24    80.99    0.99,
0       63.48    95.51    12.28   11.47    76.64    0.74,
0       67.04    93.12    5.75    11.69    67.68    0.58,
0       85.46    98.66    5.25    27.79    91.27    0.87,
0.42    63.72    93.69    16.54   11.44    71.7     0.98,

```

```
0.68      60.13      93.36      4.72      17.7      62.23      1.73,  
0.46      68.83      96.76      6.1       25.04     78.13     1.52,  
0.03      48.8       96.13     6.32     31.62     84.59     2.29,  
0.88      66.79     92.77     8.98     39.18     71.53     2.33,  
0.68      55.37     85.61     14.07    18.08     61.49     1.97,  
0.1       72.36     92.97     10.12    16.91     72.07     1.58,  
0.41      63.62     91.36     13.74    15.36     77.19     3.23,  
0.57      54.96     86.53     18.16    21.97     66.47     2.46,  
0.17      51.21     81.77     11.9     12.63     53.89     2.2,  
5.81      60.02     84.55     19.36    18.86     64.96     3.54,  
2.77      59.17     84.13     6.22     20.73     60.07     3.28,  
31.29     62.81     90.65     25.73    24.9      68.85     5.2,  
31.93     28.04     49.23     28.4     23.97     51.27     3.75};
```

```
a=data1[,2:7]; b=data1[,1];
```

```
optn=j(9,1,.);
```

```
optn[2]=2;
```

```
optn[3]=3;
```

```
optn[5]=0;
```

```
optn[8]=3;
```

Lampiran 6. Ouput Regresi *Robust Least Trimmed Square (LTS)*

| | | | | |
|--|------------|------------|---------|--------|
| Least Trimmed Squares (LTS) Method | | | | |
| Least Trimmed Squares (LTS) Method | | | | |
| Minimizing Sum of 21 Smallest Squared Residuals. | | | | |
| Highest Possible Breakdown Value = 41.18 % | | | | |
| Random Selection of 500 Subsets | | | | |
| All 500 Subsets were Nonsingular | | | | |
| LTS Objective Function = 0.0892078998 | | | | |
| Preliminary LTS Scale = 0.1862241111 | | | | |
| Robust R Squared = 0.7323199365 | | | | |
| Final LTS Scale = 0.1667540456 | | | | |
| RLS Parameter Estimates Based on LTS | | | | |
| Variable | Estimate | Std Err | t Value | Pr > t |
| VAR1 | -0.0011289 | 0.00387548 | -0.29 | 0.7746 |
| VAR2 | 0.01838753 | 0.01302295 | 1.41 | 0.1771 |
| VAR3 | 0.03557334 | 0.00880176 | 4.04 | 0.0009 |
| VAR4 | 0.02985777 | 0.00549978 | 5.43 | <.0001 |
| VAR5 | -0.0146698 | 0.00579642 | -2.53 | 0.0222 |
| VAR6 | 0.06953609 | 0.05258493 | 1.32 | 0.2046 |
| Intercep | -1.3249573 | 1.06570376 | -1.24 | 0.2317 |

Lampiran 7. Syntax Metode Regresi Kuantil dengan Menggunakan Program R

```
library(quantreg)

data=read.csv("F:/DATA.csv", sep=", ")

k1=rq(Y~X1+X2+X3+X4+X5+X6, data=data,
      tau=c(0.05,0.25,0.5,0.75,0.95))

summary(k1, se="iid")

=====

mancari R-square

fit0 <- rq(Y~1, tau=0.05, data=data)
fit1 <- rq(Y~X1+X2+X3+X4+X5+X6, tau=0.05, data=data)
rho <- function(u, tau=0.05) u*(tau- (u < 0))
R1 <- 1-fit1$rho/fit0$rho
R1

fit0 <- rq(Y~1, tau=0.25, data=data)
fit1 <- rq(Y~X1+X2+X3+X4+X5+X6, tau=0.25, data=data)
rho <- function(u, tau=0.25) u*(tau- (u < 0))
R1 <- 1-fit1$rho/fit0$rho
R1

fit0 <- rq(Y~1, tau=0.5, data=data)
fit1 <- rq(Y~X1+X2+X3+X4+X5+X6, tau=0.5, data=data)
rho <- function(u, tau=0.5) u*(tau- (u < 0))
R1 <- 1-fit1$rho/fit0$rho
```

```

fit0 <- rq(Y~1,tau=0.75,data=data)

fit1 <- rq(Y~X1+X2+X3+X4+X5+X6,tau=0.75,data=data)

rho <- function(u,tau=0.75)u*(tau-(u < 0))

R1 <- 1-fit1$rho/fit0$rho

R1

fit0 <- rq(Y~1,tau=0.95,data=data)

fit1 <- rq(Y~X1+X2+X3+X4+X5+X6,tau=0.95,data=data)

rho <- function(u,tau=0.95)u*(tau-(u < 0))

R1 <- 1-fit1$rho/fit0$rho

R1

=====

plot data

attach(data)

plot(X1,Y, type="n", xlab="Akses Sanitasi Layak",
      ylab="Annual Parasite Incidence (API) Malaria")

points(X1, Y, cex=0.70, col="blue")

abline(rq(Y~X1, tau=0.05), col="red")

abline(rq(Y~X1, tau=0.25), col="orange")

abline(rq(Y~X1, tau=0.5), col="black")

abline(rq(Y~X1, tau=0.75), col="green")

abline(rq(Y~X1, tau=0.95), col="violet")

legend(75, 32, c("Q5","Q25","Q50","Q75","Q95"),
      lty=c(1,1),lwd=c(2.2),col=c("blue","red","orange",""
      green","violet"))

```

```

attach(data)

plot(X2,Y, type="n", xlab="Rumah Layak Huni", ylab="Annual
      Parasite Incidence (API) Malaria")

points(X2, Y, cex=0.70, col="blue")

abline(rq(Y~X2, tau=0.05), col="red")

abline(rq(Y~X2, tau=0.25), col="orange")

abline(rq(Y~X2, tau=0.5), col="black")

abline(rq(Y~X2, tau=0.75), col="green")

abline(rq(Y~X2, tau=0.95), col="violet")

legend(88, 32, c("Q5","Q25","Q50","Q75","Q95"),
      lty=c(1,1),lwd=c(2.2),col=c("blue","red","orange","
      green","violet"))

attach(data)

plot(X3,Y, type="n", xlab="Penduduk Miskin", ylab="Annual
      Parasite Incidence (API) Malaria")

points(X3, Y, cex=0.70, col="blue")

abline(rq(Y~X3, tau=0.05), col="red")

abline(rq(Y~X3, tau=0.25), col="orange")

abline(rq(Y~X3, tau=0.5), col="black")

abline(rq(Y~X3, tau=0.75), col="green")

abline(rq(Y~X3, tau=0.95), col="violet")

legend(3, 32, c("Q5","Q25","Q50","Q75","Q95"),
      lty=c(1,1),lwd=c(2.2),col=c("blue","red","orange","
      green","violet"))

```

```

attach(data)

plot(X4,Y, type="n", xlab="Tenaga Dokter", ylab="Annual
      Parasite Incidence (API) Malaria")

points(X4, Y, cex=0.70, col="blue")

abline(rq(Y~X4, tau=0.05), col="red")

abline(rq(Y~X4, tau=0.25), col="orange")

abline(rq(Y~X4, tau=0.5), col="black")

abline(rq(Y~X4, tau=0.75), col="green")

abline(rq(Y~X4, tau=0.95), col="violet")

legend(33, 32, c("Q5","Q25","Q50","Q75","Q95"),
      lty=c(1,1),lwd=c(2.2),col=c("blue","red","orange","
      green","violet"))

attach(data)

plot(X5,Y, type="n", xlab="Akses Air Bersih",
      ylab="Annual Parasite Incidence (API) Malaria")

points(X5, Y, cex=0.70, col="blue")

abline(rq(Y~X5, tau=0.05), col="red")

abline(rq(Y~X5, tau=0.25), col="orange")

abline(rq(Y~X5, tau=0.5), col="black")

abline(rq(Y~X5, tau=0.75), col="green")

abline(rq(Y~X5, tau=0.95), col="violet")

legend(82, 32, c("Q5","Q25","Q50","Q75","Q95"),
      lty=c(1,1),lwd=c(2.2),col=c("blue","red","orange","
      green","violet"))

```

```
attach(data)

plot(X5,Y, type="n", xlab="Puskesmas", ylab="Annual
      Parasite Incidence (API) Malaria")

points(X5, Y, cex=0.70, col="blue")

abline(rq(Y~X5, tau=0.05), col="red")

abline(rq(Y~X5, tau=0.25), col="orange")

abline(rq(Y~X5, tau=0.5), col="black")

abline(rq(Y~X5, tau=0.75), col="green")

abline(rq(Y~X5, tau=0.95), col="violet")

legend(82, 32, c("Q5","Q25","Q50","Q75","Q95"),
      lty=c(1,1),lwd=c(2.2),col=c("blue","red","orange"
      ,"green","violet"))
```

Lampiran 8. Ouput Metode Regresi Kuantil

```
tau: [1] 0.05
```

Coefficients:

| | Value | Std.Error | t value | Pr(> t) |
|-------------|----------|-----------|----------|----------|
| (Intercept) | 0.04803 | 0.50301 | 0.09549 | 0.92463 |
| X1 | 0.00352 | 0.00362 | 0.97360 | 0.33890 |
| X2 | -0.00892 | 0.00530 | -1.68224 | 0.10405 |
| X3 | 0.01112 | 0.00762 | 1.45964 | 0.15593 |
| X4 | -0.00610 | 0.00568 | -1.07331 | 0.29263 |
| X5 | 0.00524 | 0.00432 | 1.21341 | 0.23548 |
| X6 | 0.15145 | 0.04178 | 3.62513 | 0.00118 |

Koefisien Determinasi : 0.03948024

```
tau: [1] 0.25
```

Coefficients:

| | Value | Std.Error | t value | Pr(> t) |
|-------------|----------|-----------|----------|----------|
| (Intercept) | 6.13911 | 1.81009 | 3.39162 | 0.00216 |
| X1 | 0.00924 | 0.01301 | 0.70978 | 0.48392 |
| X2 | -0.07525 | 0.01907 | -3.94567 | 0.00051 |
| X3 | 0.00188 | 0.02742 | 0.06841 | 0.94596 |
| X4 | 0.01303 | 0.02044 | 0.63739 | 0.52924 |
| X5 | 0.00048 | 0.01553 | 0.03122 | 0.97532 |
| X6 | 0.09122 | 0.15034 | 0.60676 | 0.54908 |

tau: [1] 0.5

Coefficients:

| | Value | Std.Error | t value | Pr(> t) |
|-------------|----------|-----------|----------|----------|
| (Intercept) | 13.70497 | 2.62703 | 5.21690 | 0.00002 |
| X1 | 0.02780 | 0.01888 | 1.47207 | 0.15256 |
| X2 | -0.18683 | 0.02768 | -6.74985 | 0.00000 |
| X3 | 0.03432 | 0.03980 | 0.86241 | 0.39606 |
| X4 | -0.01465 | 0.02967 | -0.49389 | 0.62538 |
| X5 | 0.02159 | 0.02254 | 0.95786 | 0.34663 |
| X6 | 0.58883 | 0.21819 | 2.69873 | 0.01186 |

Koefisien Determinasi : 0.1841829

tau: [1] 0.75

Coefficients:

| | Value | Std.Error | t value | Pr(> t) |
|-------------|----------|-----------|----------|----------|
| (Intercept) | 44.53768 | 9.16107 | 4.86162 | 0.00004 |
| X1 | 0.11639 | 0.06586 | 1.76728 | 0.08848 |
| X2 | -0.61652 | 0.09652 | -6.38738 | 0.00000 |
| X3 | 0.20206 | 0.13878 | 1.45600 | 0.15692 |
| X4 | 0.04393 | 0.10347 | 0.42459 | 0.67450 |
| X5 | 0.03691 | 0.07859 | 0.46966 | 0.64237 |
| X6 | 1.54560 | 0.76087 | 2.03135 | 0.05217 |

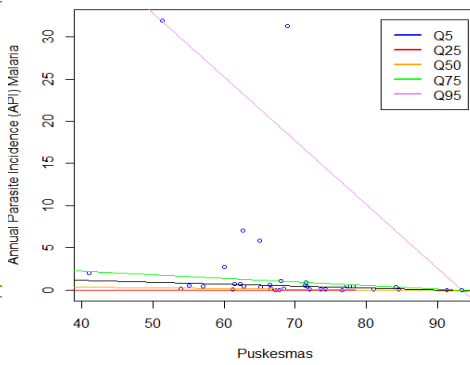
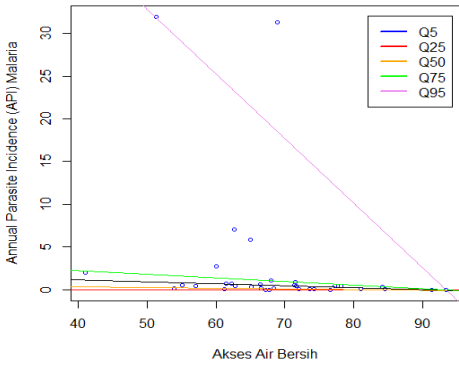
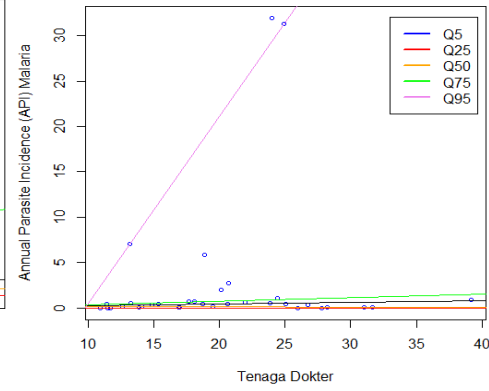
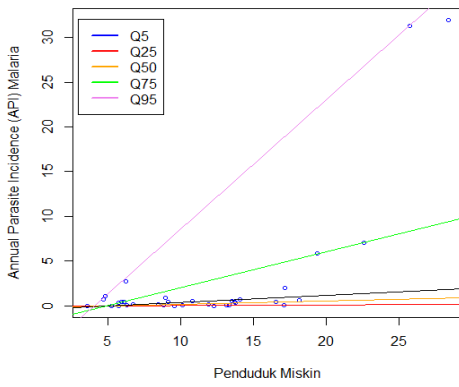
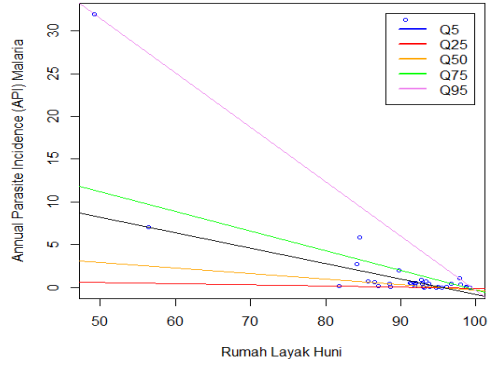
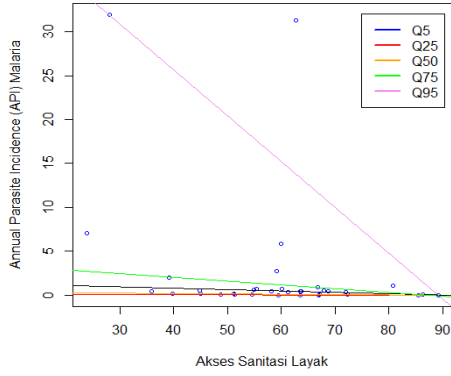
tau: [1] 0.95

Coefficients:

| | Value | Std.Error | t value | Pr(> t) |
|-------------|----------|-----------|----------|----------|
| (Intercept) | 11.93489 | 5.26685 | 2.26604 | 0.03168 |
| X1 | 0.17048 | 0.03786 | 4.50278 | 0.00012 |
| X2 | -0.39447 | 0.05549 | -7.10869 | 0.00000 |
| X3 | 0.40715 | 0.07979 | 5.10298 | 0.00002 |
| X4 | -0.27588 | 0.05949 | -4.63757 | 0.00008 |
| X5 | 0.16171 | 0.04518 | 3.57904 | 0.00133 |
| X6 | 5.70499 | 0.43744 | 13.04180 | 0.00000 |

Koefisien Determinasi : 0.8323031

Lampiran 9. Plot Regresi Kuantil



BIOGRAFI PENULIS



Penulis dilahirkan di Tual Provinsi Maluku, pada 10 juli 1993, merupakan anak ke-empat dari pasangan. Husin Matdoan dan Maryam Dfinubun. Pendidikan formal yang pernah ditempuh di SD Negeri Jati Mekar 7 Jakarta (2000-2002), kemudian pindah ke SD Inpres Namasula (2002-2004), MTS Negeri Masohi (2004-2007) dan SMK Negeri Masohi (2007-2010). Penulis kemudian melanjutkan pendidikan tinggi di Jurusan Matematika Universitas Pattimura (2010-2014).

Pada tahun 2015, penulis mendapat kesempatan untuk melanjutkan studi S2 di Jurusan Statistika Institut Teknologi Sepuluh Nopember. Karena kesempatan dan dukungan keluarga serta keinginan penulis untuk mengembangkan ilmu statistikalah yang mendorong penulis untuk melanjutkan studi S2 di Institut Teknologi Sepuluh Nopember. Penulis memiliki motto hidup, sebaik-baik manusia adalah yang paling bermanfaat kepada manusia lain dan sampaikanlah walaupun satu ayat. Hal ini menjadi penggerak penulis untuk selalu berbuat yang terbaik di lingkungan keluarga, teman maupun masyarakat. Mudah-mudahan penulis dapat mengamalkan ilmu yang telah diperoleh serta dapat mengaplikasikannya kelak dalam kehidupan sehari-hari, Semoga sekaligus menjadi amal bakti yang akan mendapatkan balasan terbaik dan dicatat serta dilipatgandakan pahalanya oleh Allah SWT di kemudian hari kelak. Amin.

Email. keepyahya@gmail.com