



**PEMODELAN REGRESI QUANTIL DENGAN KERNEL SMOOTHING PADA
FAKTOR-FAKTOR YANG MEMPENGARUHI PENYEBARAN API MALARIA DI
INDONESIA**

*(Quantile Regression Modeling with Kernel Smoothing on Factors Affecting the Spread of
Malaria Fire in Indonesia)*

Muhammad Yahya Matdoan^{a,*}, Mozart Wiston Talakua^{b,*}, & Ronald John Djami^{c,*}

^{ac} Program Studi Statistika FMIPA Universitas Pattimura

^{Jl. Ir. M. Putuhena, Kampus Unpatti-Poka, Ambon, Indonesia}

^bProgram Studi Matematika FMIPA Universitas Pattimura

^{Jl. Ir. M. Putuhena, Kampus Unpatti-Poka, Ambon, Indonesia}

Pos-el: yahya.matdoan@fmipa.unpatti.ac.id

(Diterima: 07 Agustus; Direvisi 09 Agustus; Disetujui: Agustus 2020)

Abstract

Regression analysis method is one of the statistical methods used to describe the relationship between two or more variables, so that a variable can be predicted from another variable. In regression analysis there are two types of approaches, namely parametric and nonparametric approaches. Estimates used to estimate the parameters in the regression analysis using the OLS method. This method is based on the mean distribution, so it is not appropriate to analyze a number of data that are not symmetrical or contain outliers. Therefore, a quantile regression method and kernel smoothing were developed that were not affected by data containing outliers and could also be used as an alternative to solving fluctuating data problems. This study uses quantile regression with kernel smoothing in the case of factors affecting malaria in Indonesia. The results show that the main factors causing the spread of malaria in Indonesia are access to proper sanitation, household factors that behave in a clean and healthy life, and the number of puskesmas and the percentage of medical personnel.

Keywords: *Quantile Regression, Kernel Smoothing, Malaria.*

Abstrak

Metode analisis regresi merupakan salah satu metode statistika yang digunakan untuk mengetahui hubungan antara dua atau lebih variabel, sehingga suatu variabel dapat diprediksikan dari variabel yang lain. Dalam analisis regresi terdapat dua jenis pendekatan yaitu pendekatan parametrik dan nonparametrik. Estimasi yang digunakan untuk menduga parameter pada analisis regresi yaitu dengan menggunakan metode OLS. Metode ini didasarkan pada distribusi mean, sehingga kurang tepat untuk menganalisis sejumlah data yang tidak simetris atau mengandung outlier. Oleh karena itu, dikembangkan metode regresi kuantil dan kernel smoothing yang tidak terpengaruh oleh data yang mengandung pencilan (outlier) dan juga dapat digunakan sebagai alternatif untuk menyelesaikan permasalahan data yang fluktuatif. Penelitian ini menggunakan regresi kuantil dengan kernel smoothing pada kasus faktor-faktor yang mempengaruhi penyakit malaria di Indonesia. Diperoleh hasil bahwa faktor utama penyebab penyebaran malaria di Indonesia yaitu faktor akses sanitasi layak, faktor rumah tangga yang berperilaku hidup bersih dan sehat dan faktor banyaknya jumlah puskesmas serta faktor persentase tenaga medis.

Kata kunci: *Regresi Kuantil, Kernel Smoothing, Malaria.*

PENDAHULUAN

Malaria merupakan suatu penyakit infeksi yang disebabkan oleh parasit *Plasmodium* yang hidup dan berkembang biak dalam sel darah merah manusia, ditularkan oleh nyamuk (*Anopheles*) betina dan dapat menyerang semua orang, baik laki-laki maupun perempuan pada semua golongan umur dari bayi, anak-anak dan orang dewasa. sehingga dapat menyebabkan seseorang sering sakit, daya tahan tubuh rendah dan menyebabkan kematian.

Di Indonesia penyakit malaria ditemukan tersebar luas di berbagai daerah dengan derajat infeksi yang bervariasi. Kejadian tersebut disebabkan karena adanya permasalahan-permasalahan teknis seperti pembangunan yang tidak peduli terhadap kesehatan lingkungan, mobilitas penduduk dari daerah endemis malaria, wilayah terpencil dengan kondisi lingkungan yang tidak baik, akses pelayanan kesehatan kurang dan kondisi ekonomi masyarakat yang rendah serta buruknya perilaku masyarakat terhadap kebiasaan hidup sehat. Berdasarkan data Kementerian Kesehatan RI bahwa pada tahun 2014 terdapat 252.027 penderita malaria, kemudian pada tahun 2015 terdapat 217.025 jiwa, kemudian pada tahun 2016 terdapat 218.461 jiwa, kemudian pada tahun 2017 terdapat 261.617 jiwa dan pada tahun 2018 terdapat sebanyak 222.085 orang.

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 penyebaran penyakit malaria antar daerah di Indonesia. Oleh karena itu, perlu dilakukan penelitian lebih lanjut mengenai faktor-faktor yang mengakibatkan penyebaran penyakit malaria di Indonesia. (Kementerian Kesehatan RI, 2018)

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, sedangkan variabel penaksir disebut sebagai variabel prediktor. Dalam analisis regresi terdapat dua jenis pendekatan dalam menentukan kurva regresi, yaitu pendekatan parametrik dan nonparametrik. Jika bentuk kurva regresi diketahui, maka digunakan pendekatan parametrik. Salah satu metode yang dapat digunakan untuk menduga parameter model regresi adalah metode *Ordinary Least Square* (OLS). Pendekatan nonparametrik merupakan pendekatan regresi yang sesuai untuk pola data yang tidak diketahui bentuknya, atau tidak terdapat informasi masa lalu tentang pola data (I Nyoman Budiantara, 2010).

Estimasi parameter dengan menggunakan metode OLS diperoleh dengan meminimumkan jumlah kuadrat residual dan berdasarkan pada distribusi 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. Regresi kuantil pertama kali diperkenalkan oleh Koenker dan Basset (1978), Metode ini merupakan perluasan dari model regresi pada kuantil bersyarat. Dimana, distribusi kuantil bersyarat dari variabel respon dinyatakan sebagai fungsi dari kovariat yang diamati. Untuk mendapatkan estimator parameter model regresi kuantil diperoleh dengan metode pemrograman linier diantaranya dengan menggunakan metode *simpleks*, *interior-point*, dan *smoothing*.

Selain Regresi kuantil, metode kernel smoothing merupakan metode pemulusan (*smoothing*) dengan menggunakan fungsi kernel. Metode ini dapat digunakan sebagai alternatif untuk menyelesaikan permasalahan data yang fluktuatif. Terdapat beberapa jenis

fungsi kernel, antara lain kernel uniform, kernel triangle, kernel epanechnikov, kernel gaussian, kernel kuartik dan kernel cosinus (Hardle, 1990). Dalam regresi kernel, pemilihan parameter pemulus (*bandwidth*) memiliki peran yang sangat penting untuk melihat kebaikan model. Pemilihan *bandwidth* terbaik dengan menggunakan *Generalized Cross Validation (GCV)* minimum.

LANDASAN TEORI

Analisis Regresi

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

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

dimana

- y_i : Variabel Respon pengamatan ke- i
- $f(x_i)$: Kurva regresi pengamatan ke- i
- ε_i : Residual (*error*) pengamatan ke- i

Ordinary Least Square (OLS)

Metode *Ordinary Least Square (OLS)* pada prinsipnya adalah meminimumkan jumlah kuadrat error (residual). Selanjutnya ditunjukkan Penaksir parameter $\hat{\beta}$ dari metode *Ordinary Least Square (OLS)* adalah sebagai berikut :

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

Regresi Quantil

Regresi quantil merupakan perluasan dari model regresi pada quantil bersyarat dimana distribusi quantil bersyarat dari variabel respon dinyatakan sebagai fungsi dari kovariat yang diamati. Misalkan diberikan data $\{x_{1i}, x_{2i}, \dots, x_{pi}, y_i\}, i = 1, 2, \dots, n$ merupakan himpunan berpasangan variabel random yang terdistribusi secara independen dan identik dengan quantil $\tau \in (0,1)$. Persamaan

umum regresi quantil linier yaitu sebagai berikut :

$$y_i = \beta_{\tau 0} + \beta_{\tau 1}x_{1i} + \beta_{\tau 2}x_{2i} + \dots + \beta_{\tau p}x_{pi} + \varepsilon_{i\tau}$$

Selanjutnya dapat ditulis dalam bentuk model berikut :

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

Estimasi Parameter Regresi Quantil

Diberikan himpunan data berpasangan $\{x_{1i}, x_{2i}, x_{3i}, \dots, x_{pi}, y_i\}, i = 1, 2, \dots, n$, $j = 1, 2, \dots, p$ 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 didenifisikan sebagai berikut :

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

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 didenifisikan sebagai berikut :

$$Q_y(\tau|x) = F_y^{-1}(y|x)$$

Sehi
ngg

a model fungsi quantil bersyarat yaitu sebagai berikut :

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

Persamaan umum regresi quantil linier dari variabel respon y dapat ditulis menjadi

$$y_i = \beta_{\tau 0} + \beta_{\tau 1}x_{1i} + \beta_{\tau 2}x_{2i} + \dots + \beta_{\tau p}x_{pi} + \varepsilon_{i\tau}$$

Selanjutnya dapat dibuat dalam bentuk matriks berikut :

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

Sehingga diperoleh bentuk model linier berikut :

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

Dalam regresi klasik (metode OLS) dengan prinsip meminimumkan jumlah

kuadrat residual untuk mencari penduga $\hat{\beta}$, begitu halnya dengan regresi kuantil.

Kernel Smoothing

Kernel Smoothing adalah metode untuk mendapatkan dan menggunakan bobot yang sesuai $w_{ij(\text{ker})}$ untuk menghasilkan dugaan melalui :

$$\hat{f}(x) = \sum_{j=1}^n w_{ij(\text{ker})} y_j$$

Dari persamaan tersebut sama saja dengan menghitung $\hat{f}(x_i)$ untuk $i = 1, 2, \dots, n$. Setiap titik dari n data diberi bobot $w_{ij(\text{ker})}$, $j = 1, 2, \dots, n$ untuk setiap titik dengan x_i (atau bobot $w_{ij(\text{ker})}$ berbeda untuk suatu titik prediksi x_0). Misal K adalah kernel simetrik yang berpusat di 0, sebuah kernel esensinya adalah suatu pembobot, dalam kasus ini ia merupakan pembobot dalam keanggotaan lingkungan. Kernel yang biasa dipilih adalah kernel sebaran normal standar $K(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$. Maka pembobot bagi y_i diberikan oleh :

$$W_{ij(\text{ker})} = \frac{K\left(\frac{x - x_i}{h}\right)}{\sum_{j=1}^n K\left(\frac{x - x_i}{h}\right)} y_j$$

Sehingga dapat ditulis menjadi

$$\hat{f}(x) = \sum_{j=1}^n \frac{K\left(\frac{x - x_i}{h}\right)}{\sum_{j=1}^n K\left(\frac{x - x_i}{h}\right)} y_j$$

Parameter pemulus h disebut sebagai lebar jendela. Untuk di ketahui bahwa kernel-kernel yang biasa seperti kernel normal, semua titik pengamatan digunakan untuk menghitung pemulusan pada setiap titik, tetapi titik-titik

data yang terletak sangat jauh mendapatkan pembobot yang sangat kecil. Nilai lebar jendela yang besar akan memberikan hasil yang sangat mulus karena pembobotan pada titik-titik data hanya berubah sedikit sepanjang interval pemulusan. Sedangkan

nilai lebar jendela yang kecil akan memberikan hasil yang lebih kasar (*wiggles*) karena titik-titik data yang berdekatan mendapatkan bobot yang jauh lebih besar dibandingkan dengan titik yang jauh (Rizzo, 2008).

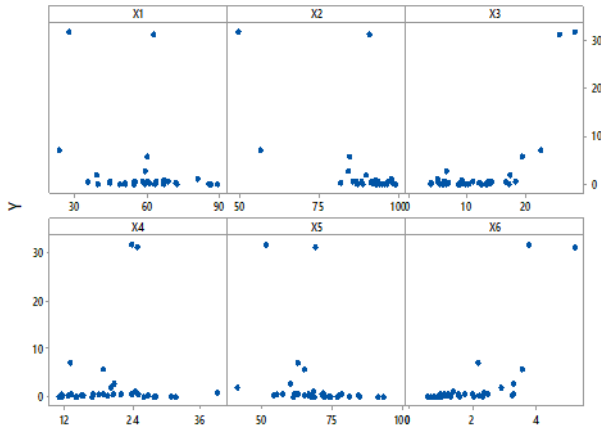
METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Kementerian Kesehatan Republik Indonesia dan Badan Pusat Statistika (BPS) tahun 2017. Data yang diambil mengenai jumlah penderita penyakit malaria di Indonesia beserta faktor-faktor yang diduga mempengaruhi. Variabel respon yang digunakan yaitu jumlah penderita malaria (Y) dan variabel prediktor yang digunakan yaitu persentase akses sanitasi layak ($X1$), persentase akses air bersih ($X2$), persentase rumah tangga yang berperilaku hidup bersih dan sehat ($X3$), persentase jumlah puskesmas ($X4$), dan persentase penduduk miskin ($X5$) serta persentase jumlah tenaga medis ($X6$).

PEMBAHASAN

Statistik Deskriptif

Studi kasus pada penelitian ini adalah faktor-faktor penyebab penyebaran jumlah penderita malaria di Indonesia, dimana akan diteliti faktor penyebab penyebaran penyakit malaria (X) terhadap banyaknya penderita penyakit malaria (Y). Scatter plot antara faktor-faktor penyebab penyakit malaria dengan jumlah penderita penyakit malaria dapat dilihat pada Gambar 1. berikut :



Gambar 1. Scatter Plot antara Variabel X dan variabel Y

Berdasarkan Gambar 1, dapat dilihat bahwa data yang digunakan dalam penelitian ini memiliki pola menyebar dan mengandung outlier. Apabila analisis regresi digunakan dengan menggunakan metode OLS akan mengakibatkan varians dari penduga parameter besar, sehingga tidak efisien. Oleh karena itu, digunakan regresi kuantil untuk menentukan parameter dari faktor-faktor yang mempengaruhi penyebaran jumlah penderita penyakit malaria di Indonesia.

Estimasi Parameter Regresi Linier Berganda

Estimasi parameter regresi linier berganda dengan menggunakan metode *Ordinary Least Square* (OLS).

Tabel 1. Estimasi Parameter Regresi Linier Berganda

Variabel	Estimation	P_Value
Constant	-12480	0,629
X ₁	-465,7	0,042
X ₂	78,05	0,778
X ₃	279,1	0,251
X ₄	1,845	0,958
X ₅	340,5	0,945
X ₆	2163	0,004

Berdasarkan Tabel 1, Dapat dilihat bahwa hanya ada dua variabel independen yang berpengaruh signifikan terhadap banyaknya jumlah penderita malaria di

Indonesia yaitu variabel persentase sanitasi layak (X₁) dan variabel persentase banyaknya tenaga medis (X₆). Hasil Estimasi tersebut dapat dimodelkan sebagai berikut :

$$\hat{Y} = -12480 - 465,7X_1 + 78,05X_2 + 279,1X_3 + 1,845X_4 + 340,5X_5 + 2163X_6$$

Estimasi Parameter Regresi Kuantil

Berdasarkan statistika deskriptif menunjukkan bahwa terdapat data dalam penelitian ini memiliki pola menyebar dan tidak simetris. Oleh karena itu, digunakan analisis regresi kuantil. Langkah yang dilakukan dalam analisis regresi kuantil adalah melakukan estimasi pada nilai kuantil ($\tau = 0.05, 0.25, 0.5, 0.75$ dan 0.95).

Tabel 2. Estimasi Parameter Regresi Kuantil

Nilai kuantil (τ) = 0,05		
Variabel	Estimasi	P_Value
Constant	1571,654	0,273
X ₁	-25,001	0,047
X ₂	17,011	0,270
X ₃	13,383	0,315
X ₄	1,864	0,339
X ₅	-732,486	0,012
X ₆	-15,314	0,688

Nilai kuantil (τ) = 0,25		
Variabel	Estimasi	P_Value
Constant	151,576	0,948
X ₁	-49,346	0,019
X ₂	1,895	0,939
X ₃	45,528	0,044
X ₄	-2,370	0,457
X ₅	-405,017	0,372
X ₆	238,733	0,000

Nilai kuantil (τ) = 0,5		
Variabel	Estimasi	P_Value
Constant	1085,696	0,749
X ₁	-82,434	0,008
X ₂	58,664	0,116
X ₃	31,036	0,330
X ₄	-2,123	0,646
X ₅	-1334,143	0,050
X ₆	310,331	0,002

Nilai quantil (τ) = 0,75		
Variabel	Estimasi	P_Value
Constant	-5490,693	0,659
X ₁	-169,669	0,119
X ₂	79,274	0,555
X ₃	110,745	0,344
X ₄	-7,454	0,661
X ₅	-266,573	0,911
X ₆	1088,226	0,003

Nilai quantil (τ) = 0,95		
Variabel	Estimasi	P_Value
Constant	-51107,51	0,001
X ₁	-383,817	0,004
X ₂	99,883	0,522
X ₃	726,741	0,000
X ₄	-38,537	0,059
X ₅	1651,625	0,556
X ₆	5217,856	0,000

Setelah mendapatkan estimasi model dengan beberapa nilai quantil, maka langkah selanjutnya adalah memilih model terbaik dengan melihat nilai *Koefisien Determinasi* (R^2) yang paling tinggi. *Koefisien determinasi* (R^2) merupakan besaran yang digunakan untuk mengukur kelayakan model regresi dan menunjukkan besar kontribusi X terhadap perubahan Y. Semakin tinggi nilai R^2 (mendekati 1) maka semakin baik model regresi yang terbentuk.

Tabel 3. Koefisien Determinasi (R^2)

Nilai Quantil	Koefisien Determinasi (R^2)
$Q_{(0,05)}$	0,072
$Q_{(0,25)}$	0,064
$Q_{(0,50)}$	0,101
$Q_{(0,75)}$	0,190
$Q_{(0,95)}$	0,740
OLS	0,520

Berdasarkan Tabel 3, dapat dilihat bahwa pada nilai quantil $Q_{(0,95)}$, memiliki

nilai R^2 yaitu sebesar (0,740) lebih besar diantara beberapa nilai quantil dan lebih besar daripada OLS yaitu dengan nilai R^2 sebesar (0,523). Hal ini menunjukkan bahwa model regresi pada quantil $Q_{(0,95)}$ merupakan model terbaik.

$$Q_{0,95}(y|x) = -51107,51 - 383,817X_1 + 99,883X_2 + 726,741X_3 - 38,537X_4 + 1651,625X_5 + 5217,856X_6$$

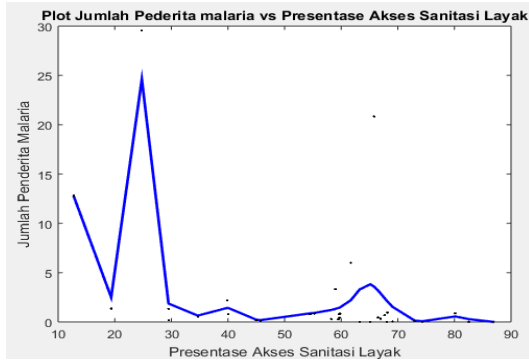
Berdasarkan model regresi quantil pada $\tau=0,95$, Pada koefisien persentase akses sanitasi layak (X₁) berkurang sebesar (383,817), maka akan meningkatkan jumlah penderita malaria dengan asumsi variabel bebas yang lainnya tetap. apabila persentase akses air bersih (X₂) meningkat sebesar (99,881), maka jumlah penderita malaria akan berkurang dengan asumsi variabel bebas yang lainnya tetap. Koefisien persentase rumah tangga yang berperilaku hidup bersih dan sehat (X₃) meningkat sebesar (726,741), maka akan mengurangi jumlah penderita malaria dengan asumsi variabel bebas yang lainnya tetap. Koefisien banyaknya fasilitas jumlah puskesmas (X₄) berkurang sebesar (38,537), maka akan meningkatkan jumlah penderita malaria dengan asumsi variabel bebas yang lainnya tetap. Koefisien persentase penduduk miskin (X₅) meningkat sebesar (1651,625), Maka akan meningkatkan jumlah penderita malaria dengan asumsi variabel bebas yang lainnya tetap. Koefisien persentase tenaga medis (X₆) meningkat sebesar (5217,856), maka akan mengurangi jumlah penderita malaria dengan asumsi variabel bebas yang lainnya tetap. Selanjutnya dengan menggunakan $\alpha = 5\%$ pada quantil $\tau=0,95$, maka diperoleh hasil bahwa indikator yang berpengaruh signifikan terhadap penyebaran jumlah penderita malaria di Indonesia yaitu persentase akses sanitasi layak (X₁), rumah tangga yang berperilaku hidup bersih dan sehat (X₃) dan banyaknya jumlah puskesmas (X₄) serta banyaknya tenaga medis (X₆).

**Estimator Regresi Kernel
Memodelkan Jumlah Penderita Malaria dengan Persentase Akses Sanitasi Layak**

Berdasarkan metode GCV diperoleh bandwidth optimal sebesar 2.1300 dengan nilai GCV minimum 32.4701, Sehingga estimasi model adalah sebagai berikut :

$$\hat{f}(x) = \frac{\sum_{i=1}^{34} K\left(\frac{X_{i1} - X_{j1}}{2.1300}\right) Y_j}{\sum_{i=1}^{34} K\left(\frac{X_{i1} - X_{j1}}{2.1300}\right)}$$

kurva estimasi dapat dilihat pada Gambar 2 berikut :



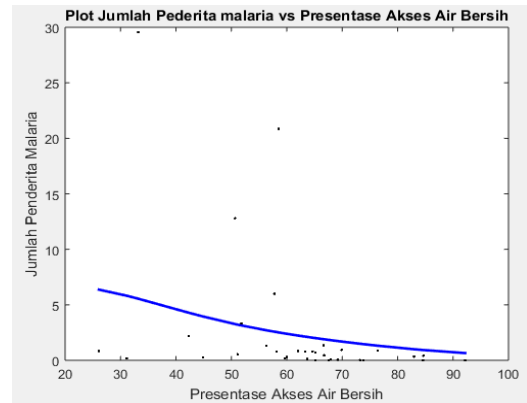
Gambar 2. Estimasi Kurva antara Jumlah Penderita Malaria dengan Persentase Akses Sanitasi Layak

Memodelkan Jumlah Penderita Malaria dengan Persentase Akses Air Bersih

Berdasarkan metode GCV diperoleh bandwidth optimal sebesar 15.3600 dengan nilai GCV minimum 39.3309, Sehingga estimasi model adalah sebagai berikut

$$\hat{f}(x) = \frac{\sum_{i=1}^{34} K\left(\frac{X_{i1} - X_{j1}}{15.3600}\right) Y_j}{\sum_{i=1}^{34} K\left(\frac{X_{i1} - X_{j1}}{15.3600}\right)}$$

Kurva estimasi dapat dilihat pada Gambar 3 berikut :



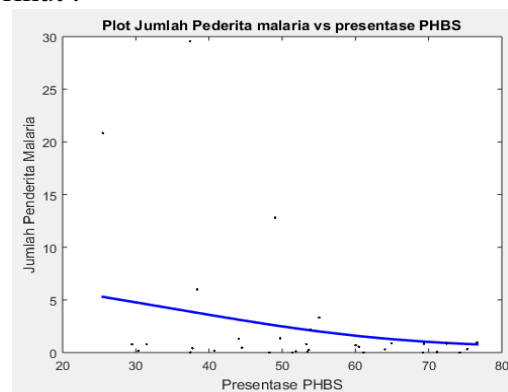
Gambar 3. Estimasi Kurva antara Jumlah Penderita Malaria dengan Persentase Akses Air Bersih

Memodelkan Jumlah Penderita Malaria dengan Persentase Rumah Tangga Yang Berperilaku Hidup Bersih dan Sehat (PHBS)

Berdasarkan metode GCV diperoleh bandwidth optimal sebesar 14.1800 dengan nilai GCV minimum 39.1658. Sehingga estimasi model adalah sebagai berikut :

$$\hat{f}(x) = \frac{\sum_{i=1}^{34} K\left(\frac{X_{i1} - X_{j1}}{14.1800}\right) Y_j}{\sum_{i=1}^{34} K\left(\frac{X_{i1} - X_{j1}}{14.1800}\right)}$$

Kurva estimasi dapat dilihat pada Gambar 4 berikut :

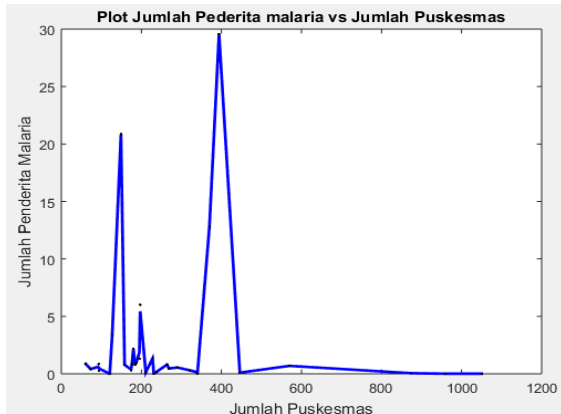


Gambar 4. Estimasi Kurva antara Jumlah Penderita Malaria dengan Persentase Rumah Tangga yang Berperilaku Hidup Bersih dan Sehat
Memodelkan Jumlah Penderita Malaria Dengan Jumlah Puskesmas

Berdasarkan metode GCV diperoleh bandwidth optimal sebesar 1 dengan nilai GCV minimum 5.7081, Sehingga estimasi model tersebut adalah

$$\hat{f}(x) = \frac{\sum_{i=1}^{34} K\left(\frac{X_{i1} - X_{j1}}{1}\right)}{\sum_{i=1}^{34} K\left(\frac{X_{i1} - X_{j1}}{1}\right)} Y_j$$

Kurva estimasi dapat dilihat pada Gambar 5 berikut :



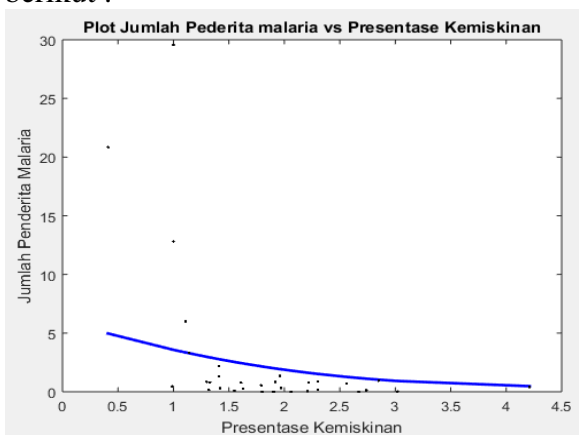
Gambar 5. Estimasi Kurva antara Jumlah Penderita Malaria dengan Jumlah Puskesmas

Memodelkan Jumlah Penderita Malaria dengan Persentase Kemiskinan

Berdasarkan metode GCV diperoleh bandwidth optimal sebesar 1 dengan nilai GCV minimum 36.8120, Sehingga estimasi model adalah sebagai berikut :

$$\hat{f}(x) = \frac{\sum_{i=1}^{34} K\left(\frac{X_{i1} - X_{j1}}{1}\right)}{\sum_{i=1}^{34} K\left(\frac{X_{i1} - X_{j1}}{1}\right)} Y_j$$

Kurva estimasi dapat dilihat pada Gambar 6 berikut :



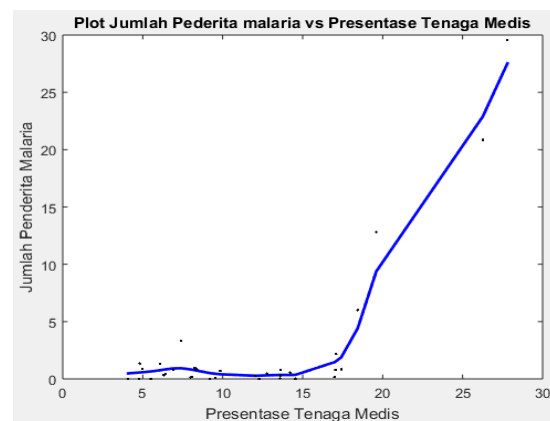
Gambar 6. Estimasi Kurva antara Jumlah Penderita Malaria dengan Persentase Penduduk Miskin

Memodelkan Jumlah Penderita Malaria dengan Persentase Tenaga Medis

Berdasarkan metode GCV diperoleh bandwidth optimal sebesar 1 dengan nilai GCV minimum 1.8900, Sehingga estimasi model adalah sebagai berikut :

$$\hat{f}(x) = \frac{\sum_{i=1}^{34} K\left(\frac{X_{i1} - X_{j1}}{1}\right)}{\sum_{i=1}^{34} K\left(\frac{X_{i1} - X_{j1}}{1}\right)} Y_j$$

Kurva estimasi dapat dilihat pada Gambar 7 berikut :



Gambar 7. Estimasi Kurva antara Jumlah Penderita Malaria dengan Persentase Banyaknya Tenaga Medis

KESIMPULAN

Berdasarkan hasil dan pembahasan dapat disimpulkan bahwa dengan menggunakan regresi kuantil, model terbaik yang diperoleh yaitu pada nilai kuantil $\tau=0,95$ dengan faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia adalah Akses Sanitasi layak, dengan GCV minimum sebesar (32.4701), faktor rumah tangga yang berperilaku hidup bersih dan sehat dengan GCV minimum sebesar (39.1658), dan faktor banyaknya jumlah puskesmas dengan GCV minimum sebesar (5.7081) serta faktor banyaknya tenaga medis dengan GCV minimum sebesar (1.8900).

DAFTAR PUSTAKA

- Budiantara, I. N. (2011, December). Penelitian Bidang Regresi Spline Menuju Terwujudnya Penelitian Statistika yang Mandiri dan Berkarakter. In *Prosiding Seminar Nasional MIPA*.
- Statistik, B. P. (2017). Badan pusat statistik. *Badan Pusat Statistik*.
- Goh, S. C., & Knight, K. (2009). Nonstandard quantile-regression inference. *Econometric Theory*, 1415-1432.
- Härdle, W. (1990). *Applied nonparametric regression* (No. 19). Cambridge university press.
- Hubert, M., Rousseeuw, P. J., & Van Aelst, S. (2008). High-breakdown robust multivariate methods. *Statistical science*, 92-119.
- Kementerian Kesehatan RI. (2017). *Profil kesehatan Indonesia*. Jakarta: Pusat data dan Informasi Kementerian Kesehatan.
- Koenker, R. (2005). *Quantile Regression*. . New York: Econometric society monographs Cambridge University.
- Lin W, Z. L. (2015). Optimal Smoothing In Nonparametric Conditional Quantile Derivative Function Estimation. *Journal of Econometrics ScienceDirect*, Vol 8. Hal. 502-513.
- Matdoan, M. Y. (2020). Estimasi Parameter Regresi Quantil Dengan Regresi Robust Least Trimmed Square (LTS) (Studi Kasus: Faktor-Faktor yang Mempengaruhi Penyebaran Penyakit Malaria di Indonesia). *Jurnal Riset dan Aplikasi Matematika (JRAM)*, 4(1), 50-62.
- Matdoan, M. Y., Persulesy, E. R., & Lembang, F. K. (2016). Analisis Preferensi Pelanggan PT. PLN (Persero) dalam Menentukan Atribut Rekening Listrik Prabayar di Kota Ambon dengan Menggunakan Metode Analisis Konjoin. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 10(1), 37-46.
- Matdoan, M. Y., & Balami, A. M. (2019). Estimasi Parameter Regresi Kuantil Dengan Fungsi Spline Truncated Pada Kasus Demam Berdarah Dengue di Kota Surabaya. *Jurnal MSA (Matematika dan Statistika serta Aplikasinya)*, 7(1), 44-53.
- Meentemeyer, R. K., Anacker, B. L., Mark, W., & Rizzo, D. M. (2008). Early detection of emerging forest disease using dispersal estimation and ecological niche modeling. *Ecological Applications*, 18(2), 377-390
- Shodiqin, A., Aini, A. N., & Rubowo, M. R. (2018). Perbandingan Metode Regresi Robust yakni Metode Least Trimmed Squares (LTS) dengan metode Estimator-MM (Estimasi-MM) (Studi Kasus Data Ujian Tulis Masuk Terhadap Hasil IPK Mahasiswa UPGRIS).