



## Estimasi Parameter Regresi Linear Menggunakan Regresi Kuantil

**Baiq Devi Rachmawati<sup>a\*</sup>, Qurratul Aini<sup>b</sup>**

<sup>a</sup>Program Studi Matematika, FMIPA, Universitas Mataram, Jl. Majapahit No. 62, Mataram, 83125, Indonesia. Email: [devirch96@gmail.com](mailto:devirch96@gmail.com)

<sup>b</sup>Program Studi Matematika, FMIPA, Universitas Mataram, Jl. Majapahit No. 62, Mataram, 83125, Indonesia. Email: [qurratulaini.aini@gmail.com](mailto:qurratulaini.aini@gmail.com)<sup>b</sup>

### ABSTRACT

Regression analysis is a statistical analysis method for estimating the relationship between dependent variables (Y) and one or more independent variables (X). As the purpose of this study is to theoretically examine the quantile regression method in estimating linear regression parameters. In regression analysis usually the method used to estimate parameters is the least square method with assumptions that must be met that normal assumption, homoskedasticity, no autocorrelation and non multicollinearity. Basically the least square method is sensitive to the assumptions of deviations in the data, so that the estimations results will be less good if the assumptions are not fulfilled. Therefore, to overcome the limitations of the least square method developed a quantile regression method for estimating linear regression parameters. Based on the result of research that has been done shows that the estimation of linear regression parameters using the quantile regression method is obtained by minimizing the absolute number of errors through the simplex algorithm.

Keywords: Least square method, Quantile regression method, Simplex algorithm.

### 1. Pendahuluan

Walpole et al. (2011) menyatakan bahwa analisis regresi merupakan suatu metode statistika yang mempelajari tentang pola hubungan secara sistematis antara variabel dependen (Y) dengan satu atau lebih variabel independen (X). Dalam analisis regresi biasanya metode yang digunakan untuk melakukan estimasi parameter adalah Metode Kuadrat Terkecil.

Estimasi parameter metode kuadrat terkecil pada prinsipnya yaitu meminimumkan jumlah kuadrat *error* nilai observasi terhadap rata-ratanya dengan cara menurunkan jumlah kuadrat *error* dengan masing-masing parameter regresi kemudian disamakan dengan nol. Terdapat beberapa asumsi model yang harus dipenuhi pada metode kuadrat terkecil agar estimator yang dihasilkan memiliki sifat BLUE (*Best Linear Unbiased Estimator*) adalah asumsi kenormalan, asumsi

homoskedastisitas, asumsi tidak terjadi multikolinieritas dan tidak ada autokorelasi (Gujarati, 2004).

Menurut Youlanda (2015), pada dasarnya metode kuadrat terkecil peka terhadap adanya penyimpangan asumsi pada data, sehingga hasil estimasinya akan kurang baik bila asumsi-asumsi tersebut tidak terpenuhi. Oleh karena itu, untuk mengatasi keterbatasan dari metode kuadrat terkecil berkembanglah metode regresi kuantil untuk estimasi parameter regresi linear.

Menurut Uthami.dkk (2013), regresi kuantil pertama kali diperkenalkan oleh Roger Koenker dan Gilbert Basset 1978, regresi ini bertujuan untuk memperluas ide-ide dalam estimasi model fungsi kuantil bersyarat, dimana distribusi kuantil bersyarat dari variabel dependen dinyatakan sebagai fungsi dari kovariat yang diamati. Analisis ini sangat berguna untuk sebaran bersyarat yang tidak simetris, padat di ekor sebarannya, atau sebarannya terpotong.

\* Corresponding author. Alamat e-mail: [devirch96@gmail.com](mailto:devirch96@gmail.com)

Berdasarkan uraian di atas, maka akan dilakukan penelitian tentang estimasi parameter regresi linear menggunakan regresi kuantil. Dalam hal ini, penelitian ini mengkaji secara teori regresi kuantil dalam mengatasi pelanggaran asumsi pada metode kuadrat terkecil.

## 2. Analisis Regresi

Analisis Regresi (*regression analysis*) merupakan suatu teknik untuk membangun persamaan garis lurus yang selanjutnya digunakan untuk membuat suatu perkiraan (Algifari, 2000). Terdapat dua jenis persamaan untuk menyatakan hubungan antara variabel independen X dengan variabel dependen Y pada analisis regresi yaitu analisis regresi linear sederhana dan analisis regresi linear berganda.

Walpole et al. (2011) mengatakan bahwa untuk menyatakan hubungan Y dan X menggunakan regresi linear berganda dengan  $k$  variabel independen dapat dinyatakan dalam persamaan berikut:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i \quad (1)$$

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

Keterangan:

- $Y_i$  : variabel dependen ke- $i$   
 $X_{ki}$  : variabel independen dengan  $k = 1, 2, \dots, j$   
 $\beta_0, \beta_1, \dots, \beta_k$  : parameter yang tidak diketahui nilainya dan akan diestimasi  
 $\varepsilon_i$  : variabel galat/kesalahan regresi  
 $n$  : banyaknya data observasi

Dengan menggunakan lambang matriks, persamaan (1) di atas dapat ditulis sebagai

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2)$$

dimana

$$\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1k} \\ 1 & X_{21} & X_{22} & \dots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{nk} \end{bmatrix},$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

## 3. Metode Kuadrat Terkecil

Metode kuadrat terkecil adalah metode estimasi parameter yang sering digunakan dalam analisis regresi dengan tujuan meminimumkan jumlah kuadrat *error* (*sum square error*) (Andani dan Widodo, 2016). Menurut Youlanda (2015), untuk melakukan analisis regresi yang benar berdasarkan metode kuadrat terkecil maka harus memenuhi asumsi-asumsi yang telah ditentukan. Asumsi

yang diperlukan dalam metode kuadrat terkecil, yaitu asumsi kenormalan, asumsi homoskedastisitas, asumsi tidak ada autokorelasi dan asumsi non-multikolinieritas.

Menurut Gujarati (2004), estimasi  $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$  dengan menggunakan metode kuadrat terkecil dapat diperoleh dengan meminimumkan jumlah kuadrat *error*. Dalam notasi matriks, meminimumkan jumlah kuadrat *error* dapat dicapai dengan meminimumkan  $\sum_{i=1}^n \varepsilon_i^2$ .

$$\begin{aligned} \sum_{i=1}^n \varepsilon_i^2 &= \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} \\ &= (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})^T (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \\ &= (\mathbf{Y}^T - \hat{\boldsymbol{\beta}}^T \mathbf{X}^T) (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \\ &= \mathbf{Y}^T \mathbf{Y} - \mathbf{Y}^T \mathbf{X}\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{Y} + \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{X}\hat{\boldsymbol{\beta}} \\ &= \mathbf{Y}^T \mathbf{Y} - \mathbf{Y}^T \mathbf{X}\hat{\boldsymbol{\beta}} - (\mathbf{Y}^T \mathbf{X}\hat{\boldsymbol{\beta}})^T + \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{X}\hat{\boldsymbol{\beta}} \\ &= \mathbf{Y}^T \mathbf{Y} - \mathbf{Y}^T \mathbf{X}\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{Y} + \hat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{X}\hat{\boldsymbol{\beta}} \end{aligned} \quad (3)$$

Selanjutnya, untuk menentukan estimator regresi linear dapat dihitung dari turunan pertama persamaan (3) secara parsial terhadap  $\hat{\boldsymbol{\beta}}^T$  dan disamakan hasilnya dengan nol. Hasil yang diperoleh sebagai berikut:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{Y}) \quad (4)$$

$$\begin{aligned} \Rightarrow \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} &= \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{12} & \dots & x_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{k1} & x_{k2} & \dots & x_{kn} \end{pmatrix}^{-1} \begin{bmatrix} 1 & x_{11} & \dots & x_{k1} \\ 1 & x_{12} & \dots & x_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \dots & x_{kn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \\ \Rightarrow \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} &= \begin{pmatrix} N & \sum x_1 & \dots & \sum x_k \\ \sum x_1 & \sum x_1^2 & \dots & \sum x_1 x_2 \\ \vdots & \vdots & \ddots & \vdots \\ \sum x_k & \sum x_1 x_k & \dots & \sum x_k^2 \end{pmatrix}^{-1} \begin{bmatrix} \sum y \\ \sum x_1 y \\ \vdots \\ \sum x_k y \end{bmatrix} \\ \Rightarrow \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} &= \frac{1}{\det(\mathbf{X}^T \mathbf{X})} \text{adj}(\mathbf{X}^T \mathbf{X}) (\mathbf{X}^T \mathbf{Y}) \end{aligned} \quad (5)$$

## 4. Regresi Kuantil

Menurut Uthami.dkk (2013), regresi kuantil pertama kali diperkenalkan oleh Roger Koenker dan Gilbert Basset tahun 1978. Kuantil dapat dioperasikan dengan penyusunan atau pengurutan sampel pengamatan sehingga lebih mudah menentukan letaknya dan dapat mendefinisikan kuantil melalui alternatif yang sederhana sebagai masalah optimisasi.

Menurut Hapsery (2017), persamaan umum regresi kuantil linear khusus untuk kuantil bersyarat  $Q(\tau | X_{1i}, X_{2i}, \dots, X_{pi})$  dari variabel dependen  $Y_i$  yaitu:

$$Y_i = \beta_0(\tau) + \beta_1(\tau)X_{1i} + \dots + \beta_p(\tau)X_{pi} + \varepsilon_i(\tau) \quad (6)$$

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

dimana  $\beta_0, \beta_1, \dots, \beta_p$  merupakan parameter yang tidak diketahui pada kuantil ke- $\tau$  dan  $\varepsilon_i$  merupakan *error* dari model regresi pada sampel sebanyak  $n$  dan pada kuantil ke- $\tau$ . Apabila model regresi kuantil disajikan dalam bentuk matriks, dapat ditulis sebagai berikut:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{p1} \\ 1 & X_{12} & X_{22} & \cdots & X_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{pn} \end{bmatrix} \begin{bmatrix} \beta_0(\tau) \\ \beta_1(\tau) \\ \vdots \\ \beta_p(\tau) \end{bmatrix} + \begin{bmatrix} \varepsilon_1(\tau) \\ \varepsilon_2(\tau) \\ \vdots \\ \varepsilon_n(\tau) \end{bmatrix} \quad (7)$$

Selanjutnya persamaan (7), dapat ditulis dalam bentuk model linear berikut:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}(\boldsymbol{\tau}) + \boldsymbol{\varepsilon} \quad (8)$$

Keterangan:

$\mathbf{y}$  : vektor kolom berukuran  $n \times 1$  dari variabel dependen ( $Y$ )

$\mathbf{X}$  : matriks berukuran dengan  $n$  observasi pada  $p$  variabel  $X_j$  dimana  $j = 1, 2, \dots, p$

$\boldsymbol{\beta}$  : vektor kolom berukuran  $(p + 1) \times 1$  parameter  $\beta_j$  dimana  $j = 1, 2, \dots, p$

$\boldsymbol{\varepsilon}$  : vektor kolom berukuran  $n \times 1$  *error*

Jika fungsi bersyarat dari kuantil ke- $\tau$  dengan variabel independen  $X$  tertentu, maka fungsi bersyarat tersebut didefinisikan sebagai berikut:

$$\begin{aligned} Q(\tau | X_{1i}, X_{2i}, \dots, X_{pi}) &= Q_{Y_i}(\tau | \mathbf{X}) \\ &= \mathbf{X}^T \boldsymbol{\beta}(\boldsymbol{\tau}) \end{aligned} \quad (9)$$

Akibatnya solusi optimasi pada regresi kuantil adalah sebagai berikut:

$$\begin{aligned} &\min_{\beta \in R^p} \sum_{i=1}^n \rho(\tau)(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}(\boldsymbol{\tau})), \quad i = 1, 2, \dots, n \quad \tau \in (0, 1) \\ \text{atau} \\ \hat{\beta}(\tau) &= \min_{\beta \in R^p} \sum_{i=1}^n \rho(\tau)(\varepsilon_i) \end{aligned} \quad (10)$$

## 5. Estimasi Parameter Regresi Linear menggunakan Regresi Kuantil

Pada bagian ini, akan dijabarkan secara teori mengenai estimasi parameter regresi linear menggunakan regresi kuantil. Dalam proses analisis tersebut, terdapat beberapa hal yang dilakukan yaitu menunjukkan median sebagai solusi dari masalah minimasi, menunjukkan fungsi kuantil bersyarat ke- $\tau$  solusi dari masalah minimasi, menunjukkan *loss function* asimetris dan menentukan

estimator  $\hat{\beta}(\tau)$  menggunakan algoritma simpleks. Berikut ini adalah hasil analisis yang diperoleh.

### 5.1 Menunjukkan Median Solusi dari Masalah Minimasi

Dalam menunjukkan median sebagai solusi dari masalah minimasi dilakukan dengan meminimumkan fungsi berikut ini terhadap  $\hat{y}$  dengan mengasumsikan bahwa bobot yang diberikan adalah sama sehingga *loss function* berbentuk simetris.

$$\arg \min_{\beta \in R^p} E[|y - XB(\tau)|] \quad (11)$$

Misalkan  $XB(\tau) = \hat{y}$  maka diperoleh:

$$\begin{aligned} &\frac{\partial}{\partial \hat{y}} E[|y - \hat{y}|] = 0 \\ \Rightarrow &\frac{\partial}{\partial \hat{y}} \int_{-\infty}^{\infty} |y - \hat{y}| f(y) dy = 0 \\ \Rightarrow &\frac{\partial}{\partial \hat{y}} \left[ \int_{-\infty}^{\hat{y}} |y - \hat{y}| f(y) dy + \int_{\hat{y}}^{\infty} |y - \hat{y}| f(y) dy \right] = 0 \\ \Rightarrow &\lim_{\alpha \rightarrow -\infty} \left[ \frac{\partial}{\partial \hat{y}} \int_{\alpha}^{\hat{y}} |y - \hat{y}| f(y) dy \right] + \lim_{\omega \rightarrow \infty} \left[ \frac{\partial}{\partial \hat{y}} \int_{\hat{y}}^{\omega} |y - \hat{y}| f(y) dy \right] = 0 \\ \Rightarrow &\lim_{\alpha \rightarrow -\infty} \left[ \frac{\partial}{\partial \hat{y}} \int_{\alpha}^{\hat{y}} (\hat{y} - y) f(y) dy \right] + \lim_{\omega \rightarrow \infty} \left[ \frac{\partial}{\partial \hat{y}} \int_{\hat{y}}^{\omega} (y - \hat{y}) f(y) dy \right] = 0 \\ \Rightarrow &\left[ 0 + \lim_{\alpha \rightarrow -\infty} \int_{\alpha}^{\hat{y}} f(y) dy \right] + \left[ 0 + \lim_{\omega \rightarrow \infty} \int_{\hat{y}}^{\omega} -f(y) dy \right] = 0 \\ \Rightarrow &[F(\hat{y})] - [1 - F(\hat{y})] = 0 \\ \Rightarrow &F(\hat{y}) = 1/2 \end{aligned}$$

### 5.2 Menunjukkan Fungsi Kuantil Bersyarat ke- $\tau$ Solusi dari Masalah Minimasi

Sama halnya dengan menunjukkan median pada bagian (5.1), dalam menunjukkan fungsi kuantil bersyarat ke- $\tau$  dari  $F_Y$  dilakukan dengan meminimumkan fungsi terhadap  $\hat{y}$  dengan menambahkan bobot  $\tau$  dan  $1 - \tau$  yang tidak simetris:

$$\begin{aligned} &\frac{\partial}{\partial \hat{y}} E[\rho(\tau)|y - \hat{y}|] = 0 \\ \Rightarrow &\frac{\partial}{\partial \hat{y}} \int_{-\infty}^{\infty} \rho(\tau)|y - \hat{y}| f(y) dy = 0 \\ \Rightarrow &\frac{\partial}{\partial \hat{y}} \left[ \int_{-\infty}^{\hat{y}} (1 - \tau)|y - \hat{y}| f(y) dy + \int_{\hat{y}}^{\infty} \tau |y - \hat{y}| f(y) dy \right] = 0 \\ \Rightarrow &\lim_{\alpha \rightarrow -\infty} \frac{\partial}{\partial \hat{y}} \left[ (1 - \tau) \int_{\alpha}^{\hat{y}} |y - \hat{y}| f(y) dy \right] + \lim_{\omega \rightarrow \infty} \frac{\partial}{\partial \hat{y}} \left[ \tau \int_{\hat{y}}^{\omega} |y - \hat{y}| f(y) dy \right] = 0 \end{aligned}$$

$$\begin{aligned} \Rightarrow & (1 - \tau) \left[ ((\hat{y} - y)f(y))|_{y=\hat{y}} + \lim_{\alpha \rightarrow -\infty} \int_{\alpha}^{\hat{y}} \frac{\partial}{\partial \hat{y}} (\hat{y} - y)f(y)dy \right] + \tau \left[ ((y - \hat{y})f(y))|_{y=\hat{y}} + \lim_{\omega \rightarrow \infty} \int_{\hat{y}}^{\omega} \frac{\partial}{\partial \hat{y}} (y - \hat{y})f(y)dy \right] = 0 \\ \Rightarrow & (1 - \tau) \left[ 0 + \lim_{\alpha \rightarrow -\infty} \int_{\alpha}^{\hat{y}} f(y)dy \right] + \tau \left[ 0 + \lim_{\omega \rightarrow \infty} \int_{\hat{y}}^{\omega} -f(y)dy \right] = 0 \\ \Rightarrow & [(1 - \tau)F(y)] - \tau[1 - F(y)] = 0 \\ \Rightarrow & F(y) = \tau \end{aligned}$$

### 5.3 Menunjukkan Loss Function Asimetris

Dalam melakukan estimasi menggunakan metode regresi kuantil, estimator didapatkan dengan meminimumkan nilai *loss function* atau perkalian antara bobot dan *error* pada regresi kuantil berbentuk asimetris untuk  $\tau \in (0,1)$  dengan penjelasan sebagai berikut:

Dimana *loss function* didefinisikan sebagai:

$$\rho(\tau)(\varepsilon_i) = \begin{cases} \tau \varepsilon_i & ; \varepsilon_i \geq 0 \\ (1 - \tau) \varepsilon_i & ; \varepsilon_i < 0 \end{cases}$$

dengan

$$\begin{aligned} \rho(\tau) &= [\tau \mathbf{I}_{(\varepsilon \geq 0)} + (1 - \tau) \mathbf{I}_{(\varepsilon < 0)}] |\varepsilon| \\ &= [(\tau - \mathbf{I}_{(\varepsilon < 0)}) \varepsilon] \\ \mathbf{I}_{(\varepsilon \geq 0)} &= \begin{cases} 1 & ; \varepsilon \geq 0 \\ 0 & ; \varepsilon < 0 \end{cases} \end{aligned}$$

Sehingga dapat dibuktikan:

1. Untuk  $\varepsilon \geq 0$

$$\begin{aligned} \rho(\tau) &= [\tau \mathbf{I}_{(\varepsilon \geq 0)} + (1 - \tau) \mathbf{I}_{(\varepsilon < 0)}] |\varepsilon| \\ &= [\tau \mathbf{I}_{(\varepsilon \geq 0)} + (1 - \tau) \mathbf{I}_{(\varepsilon < 0)}] \varepsilon \\ &= [\tau(1) + (1 - \tau) \mathbf{I}_{(\varepsilon < 0)}] \varepsilon \\ &= [\tau + \mathbf{I}_{(\varepsilon < 0)} - \tau \mathbf{I}_{(\varepsilon < 0)}] \varepsilon \\ &= [\tau + (1 - \mathbf{I}_{(\varepsilon \geq 0)}) - \tau(1 - \mathbf{I}_{(\varepsilon \geq 0)})] \varepsilon \\ &= [\tau + (1 - 1) - \tau(1 - 1)] \varepsilon \\ &= \tau \varepsilon \end{aligned}$$

2. Untuk  $\varepsilon < 0$

$$\begin{aligned} \rho(\tau) &= [\tau \mathbf{I}_{(\varepsilon \geq 0)} + (1 - \tau) \mathbf{I}_{(\varepsilon < 0)}] |\varepsilon| \\ &= [\tau \mathbf{I}_{(\varepsilon \geq 0)} + (1 - \tau) \mathbf{I}_{(\varepsilon < 0)}] (-\varepsilon) \\ &= [\tau(0) + (1 - \tau) \mathbf{I}_{(\varepsilon < 0)}] (-\varepsilon) \quad (-1) \\ &= [(\tau - 1) \mathbf{I}_{(\varepsilon < 0)}] (\varepsilon) \\ &= [(\tau - 1)(1 - \mathbf{I}_{(\varepsilon \geq 0)})] (\varepsilon) \\ &= [(\tau - 1)(1 - 0)] (\varepsilon) \\ &= (\tau - 1) \varepsilon \end{aligned}$$

Sehingga menjadi:

$$\begin{aligned} \rho(\tau) &= [\tau \mathbf{I}_{(\varepsilon \geq 0)} + (1 - \tau) \mathbf{I}_{(\varepsilon < 0)}] |\varepsilon| \\ &= [(1 - \tau) \mathbf{I}_{(\varepsilon < 0)}] \varepsilon, \quad \forall \varepsilon \end{aligned}$$

### 5.4 Menentukan Estimator $\hat{\beta}(\tau)$ menggunakan Algoritma Simpleks

Berikut adalah langkah yang dilakukan dalam

menentukan estimator persamaan regresi kuantil menggunakan algoritma simpleks pada kasus dua variabel dan  $p$ -variabel:

#### 1. Masalah dua variabel

Dalam kasus dua variabel model yang dimiliki sebagai berikut:

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (12)$$

Untuk mendapatkan estimator terbaik dari  $\beta_0$  dan  $\beta_1$  diperoleh dengan meminimumkan jumlah *absolute error* ( $\sum_{i=1}^n |\varepsilon_i|$ ).

dimana,

$$\varepsilon_i = \beta_0 + \beta_1 x_i - y_i$$

sehingga,

$$\min_{\beta_0, \beta_1} \sum_{i=1}^n |\beta_0 + \beta_1 x_i - y_i| \quad (13)$$

Fungsi di atas merupakan fungsi yang tidak linear, namun dapat diubah menjadi linear dengan bantuan program linear.

Fungsi tujuan : meminimumkan  $\sum_{i=1}^n \varepsilon_i$

Fungsi kendala :

$$\begin{aligned} \varepsilon_i &\geq \beta_0 + \beta_1 x_i - y_i & , \quad i = 1, \dots, n \\ \varepsilon_i &\geq -(\beta_0 + \beta_1 x_i - y_i) & , \quad i = 1, \dots, n \end{aligned}$$

Setiap  $\varepsilon_i$  merupakan variabel tambahan yang menunjukkan *error* ke- $i$  dengan kendala-kendala yang memenuhi:

$$\begin{aligned} \varepsilon_i &\geq \max\{\beta_0 + \beta_1 x_i - y_i, \\ &-(\beta_0 + \beta_1 x_i - y_i)\} = |\beta_0 + \beta_1 x_i - y_i| \end{aligned}$$

#### 2. Masalah $p$ -variabel

Dalam kasus  $p$ -variabel, model linear untuk regresi kuantil sebagai berikut:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta}(\tau) + \boldsymbol{\varepsilon} \quad (14)$$

Notasi sederhana  $\beta$  akan digunakan untuk kondisi median ( $\tau = 0.5$ ), dimana  $\hat{\beta}$  akan diperoleh sebagai solusi dari masalah minimasi:

$$\min_{\beta} \sum_{i=1}^n |\mathbf{y}_i - \mathbf{x}_i^T \boldsymbol{\beta}| \quad (15)$$

Persamaan (15) merupakan *error* ke- $i$  ( $\varepsilon_i$ ) yang akan diminimumkan terhadap  $\beta$ . Dimana  $\varepsilon_i$  selalu tak negatif. Karena persamaan tersebut tidak dapat diselesaikan secara analitik maka dapat diselesaikan melalui program linear dengan langkah sebagai berikut:

Diberikan  $[X]_+$ , bagian tak negatif dari  $x$  yaitu,

$$\begin{aligned} [X]_+ &= x && \text{jika } x_i \geq 0, \forall i = 1, \dots, n \\ [X]_+ &= -x && \text{jika } x_i \leq 0, \forall i = 1, \dots, n \end{aligned}$$

Sehingga, ambil bagian positif dari *error* yaitu  $|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}|$  dan bagian positif dari  $-|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}|$ .

Akibatnya,

$$\begin{aligned} \mathbf{s}_1 &= [\mathbf{y} - \mathbf{X}\boldsymbol{\beta}]_+ \\ \mathbf{s}_2 &= [\mathbf{X}\boldsymbol{\beta} - \mathbf{y}]_+ \end{aligned}$$

Masalah tersebut dapat ditulis sebagai masalah linear yang diformulasikan sebagai berikut:

$$\min_{\boldsymbol{\beta} \in \mathbb{R}^p} \{ \mathbf{1}^T \mathbf{s}_1 + \mathbf{1}^T \mathbf{s}_2 \mid \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{s}_1 - \mathbf{s}_2, \{\mathbf{s}_1, \mathbf{s}_2\} \in \mathbb{R}_+^n \} \quad (16)$$

Persamaan (16) tersebut belum jelas apa yang menjadi fungsi tujuan dan kendala-kendalanya, sehingga dibentuklah seperti berikut ini:

Diberikan

$$\mathbf{B} = [\mathbf{X} \quad -\mathbf{X} \quad \mathbf{I} \quad -\mathbf{I}]$$

dan

$$\boldsymbol{\psi} = \begin{bmatrix} [\boldsymbol{\beta}]_+ \\ [-\boldsymbol{\beta}]_+ \\ [\mathbf{y} - \mathbf{X}\boldsymbol{\beta}]_+ \\ [\mathbf{X}\boldsymbol{\beta} - \mathbf{y}]_+ \end{bmatrix}$$

$$\mathbf{d} = \begin{bmatrix} \mathbf{0}_{[p]} \\ \mathbf{0}_{[p]} \\ \mathbf{1}_{[n]} \\ \mathbf{1}_{[n]} \end{bmatrix}$$

Sebagai reformulasi dari permasalahan program linear standar, maka formulasinya dapat ditulis sebagai berikut:

Bentuk primal:

$$\text{Fungsi tujuan: } \min_{\boldsymbol{\psi}} \mathbf{d}^T \boldsymbol{\psi} \quad \text{dengan kendala } \mathbf{B}\boldsymbol{\psi} = \mathbf{y}, \boldsymbol{\tau} \geq \mathbf{0}$$

Bentuk dual:

$$\text{Fungsi tujuan: } \max_{\mathbf{z}} \mathbf{y}^T \mathbf{z} \quad \text{dengan kendala } \mathbf{B}^T \mathbf{z} \leq \mathbf{d}$$

Secara sederhana masalah di atas dapat dirumuskan sebagai berikut:

$$\max_{\mathbf{z}} \{ \mathbf{y}^T \mathbf{z} \mid \mathbf{X}^T \mathbf{z} = \mathbf{0}, \mathbf{z} \in [-1, +1]^n \} \quad (17)$$

Faktanya, untuk  $\mathbf{X}^T \mathbf{z} = \mathbf{0}$  yang ditransformasikan dengan cara dikalikan dengan  $\frac{1}{2}$  dan menambah  $\frac{1}{2} \mathbf{X}^T \mathbf{1}$ .

Sehingga

$$\begin{aligned} \frac{1}{2} \mathbf{X}^T \mathbf{z} &= \mathbf{0} \\ \Leftrightarrow \frac{1}{2} \mathbf{X}^T \mathbf{z} + \frac{1}{2} \mathbf{X}^T \mathbf{1} &= \frac{1}{2} \mathbf{X}^T \mathbf{1} \end{aligned}$$

Hasil dari transformasi dapat ditulis seperti pada persamaan berikut ini:

$$\mathbf{X}^T \left( \frac{1}{2} \mathbf{z} + \frac{1}{2} \mathbf{1} \right) = \frac{1}{2} \mathbf{X}^T \mathbf{1} \mathbf{X}^T \quad (18)$$

Jika  $\frac{1}{2} \mathbf{z} + \frac{1}{2} \mathbf{1} = \mathbf{J}$  dan  $\frac{1}{2} \mathbf{X}^T \mathbf{1} = \mathbf{b}$ , maka persamaan tersebut dapat ditulis  $\mathbf{X}^T \mathbf{J} = \mathbf{b}$ . Akibatnya diperoleh masalah dual sebagai berikut:

$$\max_{\mathbf{J}} \{ \mathbf{y}^T \mathbf{J} \mid \mathbf{X}^T \mathbf{J} = \mathbf{b}, \mathbf{J} \in [0, 1]^n \} \quad (19)$$

Nilai  $\frac{1}{2}$  pada persamaan (18) merupakan kunci utama generalisasi untuk kondisi median yang lain. Akibatnya masalah minimasi untuk median bersyarat dapat digunakan untuk kuantil bersyarat ke- $\tau$  sehingga diperoleh hasil sebagai berikut:

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

Dengan cara yang sama diperoleh formulasi dari persamaan dual

$$\max_{\mathbf{z}} \{ \mathbf{y}^T \mathbf{z} \mid \mathbf{X}^T \mathbf{z} = (1 - \tau) \mathbf{X}^T \mathbf{1}, \mathbf{z} \in [0, 1]^n \} \quad (20)$$

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

## 7. Kesimpulan

Berdasarkan hasil analisis yang telah dilakukan, maka dapat disimpulkan bahwa estimasi parameter regresi linear menggunakan metode regresi kuantil diperoleh dengan mendefinisikan median sebagai solusi untuk masalah meminimumkan jumlah *absolute error* ( $|\sum_{i=1}^n e_i|$ ) dengan meminimumkan nilai *Loss function* atau perkalian antara bobot dan *error* pada regresi kuantil berbentuk asimetris untuk  $\tau \in (0, 1)$  seperti persamaan berikut ini:

$$\hat{\boldsymbol{\beta}}(\tau) = \min_{\boldsymbol{\beta} \in \mathbb{R}^p} \sum_{i=1}^n \rho(\tau)(\varepsilon_i)$$

Hasil dari masalah minimasi tersebut diperoleh menggunakan algoritma simpleks.

## DAFTAR PUSTAKA

- Algifari, 2000, *Analisis Regresi: Teori, Kasus dan Solusi*, Edisi kedua, BPFE Yogyakarta, Yogyakarta.  
Andani, F.P.P. dan Widodo, E., 2016, *Perbandingan Kemampuan Regresi Kuantil Median dan*

- Transformasi Box-Cox dalam Mengatasi Heteroskedastisitas*, Konferensi Nasional Penelitian Matematika dan Pembelajarannya (KNPMP I), Universitas Muhammadiyah Surakarta, 12 Maret 2016.
- Gujarati, D.N., 2004, *Basic Econometric*, Mc Graw Hill, New York.
- Hapsery, A., 2017, *Regresi Kuantil Berbasis Model Rekursif dan Estimasi Sparsity untuk Analisis Publikasi Doses ITS di Scopis*, Tesis, Jurusan Statistika Fakultas MIPA Institut Teknologi Sepuluh November Surabaya.
- Uthami, I.A.P., Sukarsa, I.K.G., dan Kencana, I.P.E., 2013, *Regresi Kuantil Median untuk Mengatasi Heteroskedastisitas pada Analisis Regresi*, E-Jurnal Matematika Vol.2, No.1, Januari 2013, 6-13.
- Walpole, R.E., Raymond, H.M., Shaton, L.M., dan Keying, Y., 2011, *Probability and Statistics for Engineers and Scientist*, Ninth Edition, Lifland et al. Bookmakers, New York.
- Youlanda, S.R., 2015, *Perbandingan Metode Regresi Kuantil Median dengan Metode Weighted Least Square (WLS) untuk Menyelesaikan Heteroskedastisitas pada Analisis regresi*, Skripsi, Jurusan Matematika Fakultas MIPA Universitas Jember.