*Jurnal Matematika Murni dan Terapan Vol.6 No.2 Desember 2012 : 1-12*

# ESTIMASI MODEL LINEAR PARSIAL DENGAN PENDEKATAN KUADRAT TERKECIL DAN SIMULASINYA MENGGUNAKAN PROGRAM S-PLUS

**Nur Salam, Dewi Sri Susanti dan Dewi Anggraini**

Program Studi Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Lambung Mangkurat, Banjarbaru, 70714, Indonesia
Email : nur_salam_2006@yahoo.com

## ABSTRAK

*Model linear parsial (model semiparametrik) merupakan model pendekatan baru dalam regresi diantara 2 model regresi sudah populer yaitu regresi parametrik dan regresi nonparametrik. Model linear parsial merupakan model gabungan yang memuat komponen parametrik dan komponen nonparametrik. Penelitian ini bertujuan untuk mempelajari analisis regresi semiparametrik serta menentukan estimasi parameternya. Model linear parsial mempunyai bentuk :* $Y_i = \boldsymbol{X}_i^T \boldsymbol{\beta} + g(\boldsymbol{T}_i) + \epsilon_i$ *dengan* $\boldsymbol{X}_i$ *dan* $\boldsymbol{T}_i$ *adalah variabel penjelas, g (.) adalah fungsi yang tidak diketahui (fungsi smooth),* $\boldsymbol{\beta}$ *adalah parameter fungsi yang tidak diketahui,* $Y_i$ *variabel respon dan* $\epsilon_i$ *adalah error dengan mean* $(\epsilon_i) = 0$ *dan variansinya* $\sigma_i^2 = E(\epsilon_i^2)$.

*Hasil penelitian menunjukan bahwa estimasi parameter model linear parsial dapat dilakukan dengan menggunakan metode kuadrat terkecil dimana bagian nonparametriknya menggunakan pendekatan kernel dan selanjutnya hasil estimasinya disubtitusi ke model linier parsialnya untuk diestimasi bagian parametriknya menggunakan metode kuadrat terkecil. Hasil estimasi linear parsial yang diperoleh adalah* $\hat{g}_n(t) = \sum_{i=1}^{n} W_{ni} (Y_i - \boldsymbol{X}_i^T + \hat{\boldsymbol{\beta}}_n)$ *dengan* $\hat{\boldsymbol{\beta}}_n = (\tilde{\boldsymbol{x}}^{\boldsymbol{T}} \tilde{\boldsymbol{y}})^{-1} \tilde{\boldsymbol{x}}^{\boldsymbol{T}} \tilde{\boldsymbol{y}}$.

*Berdasarkan hasil simulasi diperoleh output nilai dan grafik yaitu untuk bagian parametrik, tampilan grafik qqnorm dan qqline estimator beta* $(\beta)$ *yaitu* $\beta_0$, $\beta_1$ *dan* $\beta_2$ *dapat terlihat dengan jelas, dimana jika n semakin besar* $(n \to \infty)$ *dan semakin besar replikasi r perulangan, maka titik-titik yang menyebar di sekitar garis lurus makin banyak dan mendekati garis lurus. Hal ini menunjukkan makin besar n dan r, maka beta* $(\beta)$ *semakin mendekati distribusi normal. Hasil simulasi estimator bagian nonparametrik dalam hal ini yang diambil sebagai contoh fungsi kernel normal nilainya mendekati g(T). Jadi dapat diambil kesimpulan singkatnya yaitu jika n semakin besar* $(n \to \infty)$ *maka estimator bagian nonparametriknya semakin mendekati g(T).*

**Kata Kunci** : *Regresi semiparametrik, regresi nonparametrik, fungsi kernel dan metode kuadrat terkecil, simulasi*.

**ABSTRACT**

*Partial linear model (model semiparametric) is a new approach in the regression models between the two regression models are already popular parametric regression and nonparametric regression. Partial linear model is a model that includes both the combination of parametric components and nonparametric components. This study uses literature by studying semiparametric regression analysis, finding and determining the estimated parameters. Partial linear model has the form: :* $Y_i = \boldsymbol{X}_i^T \boldsymbol{\beta} + g(\boldsymbol{T}_i) + \epsilon_i$ *with* $\boldsymbol{X}_i$ *and* $\boldsymbol{T}_i$ *are explanatory variables, g (.) is an unknown function (smooth function),* $\beta$ *is the parameter of unknown function,* $Y_i$ *response variable and* $\epsilon_i$ *is an error with the mean* $(\epsilon_i) = 0$ *and variance* $\sigma_i^2 = E(\epsilon_i^2)$.

*The results showed that the partial linear model parameter estimation can be performed using the least squares method in which part of the linear model using nonparametric kernel approach and subsequent estimation results are substituted into the partial linear model to estimate the parametric part of the model by using the linear least squares method. Results obtained partial linear estimation is* $\hat{g}_n(t) = \sum_{i=1}^{n} W_{ni}\,(Y_i - \boldsymbol{X}_i^T + \widehat{\boldsymbol{\beta}}_n)$ *dengan* $\widehat{\boldsymbol{\beta}}_n = (\widetilde{\boldsymbol{x}}^{\boldsymbol{T}}\,\widetilde{\boldsymbol{y}})^{-\boldsymbol{1}}\,\widetilde{\boldsymbol{x}}^{\boldsymbol{T}}\,\widetilde{\boldsymbol{y}}$.

*Based on the simulation results obtained output values and graphs are for the parametric, graphical display and qqline qqnorm estimator beta (β) is* $(\beta)$ *yaitu* $\beta_0$, $\beta_1$ *and* $\beta_2$ *can be seen clearly, where if n is greater (n → ∞) and the greater replication iteration r , then the points are spread around the more straight line and a straight line. This indicates the greater n and r, the beta (β) closer to the normal distribution. Nonparametric estimator simulation results in this section are taken as an example of a normal kernel function values approaching g (T). So it can be concluded briefly that if the larger n (n → ∞), the estimator of the nonparametric part closer to the partial linear model g (T).*

**Keywords**: *semiparametric regression, nonparametric regression, the kernel function and the least squares method, the simulation.*

## 1. PENDAHULUAN

Pendekatan parametrik adalah pendekatan yang dilakukan jika fungsi regresinya diketahui dan bergantung pada parameter, sehingga mengestimasi fungsi regresinya sama dengan mengestimasi parameternya. Sedangkan pendekatan nonparametrik dilakukan jika fungsi regresinya tidak diketahui sehingga mengestimasi fungsi regresinya dilakukan dengan mengestimasi fungsi regresi yang tidak diketahui tersebut dengan pendekatan estimator kernel, spline atau yang lain. Model linear parsial (model semiparametrik) merupakan model pendekatan baru dalam regresi diantara dua model regresi sudah populer yaitu regresi parametrik dan regresi nonparametrik. Model linear parsial merupakan model gabungan yang memuat keduanya yaitu komponen parametrik dan komponen nonparametrik. Estimasi model linear parsial telah dapat dilakukan dengan diberbagai metode yang ada misalnya metode kuadrat terkecil, metode *likelihood*, metode *penalized least square*, *mean square error* (MSE) dan lain-lain. Namun suatu hal yang cukup menarik adalah bagaimana membuat suatu simulasi dari estimasi model linear parsial tersebut. Oleh karena itu pada penelitian ini saya akan mengkaji tentang estimasi model linear parsial dengan pendekatan kuadrat terkecil (*least square*) dan simulasinya menggunakan program S-Plus.

## 2. TINJAUAN PUSTAKA

### 2.1. Estimasi, Estimator dan Estimate.

**Definisi 2.1. (Walpole, 1995)**
*Sebarang nilai yang menjelaskan ciri populasi disebut parameter.*

**Definisi 2.2. (Harinaldi, 2005)**
*Estimasi adalah keseluruhan proses yang menggunakan suatu estimator untuk menghasilkan suatu estimate (hasil estimasi) dari suatu parameter.*

**Definisi 2.3. (Harinaldi, 2005)**
*Estimator adalah setiap statistik (mean sampel, persentase sampel, variansi sampel dan lain-lain) yang digunakan untuk mengestimasi suatu parameter. Jadi, mean sampel ($\bar{x}$) adalah estimator bagi mean populasi ($\mu_x$), persentase sampel (p) adalah estimator bagi persentase papulasi($\pi$) dan variansi sampel ($s^2$) adalah estimator bagi variansi populasi ($\sigma_x^2$).*

**Definisi 2.4. (Harinaldi, 2005)**
*Estimate (hasil estimasi) adalah suatu nilai spesifik atau kuantitas dari suatu statistik seperti nilai mean, persentase sampel atau variansi sampel.*

**Definisi 2.5. (Tirta, 2004)**
*Simulasi adalah tiruan dari proses dunia nyata atau sistem. Simulasi menyangkut pembangkitan proses serta pengamatan dari proses untuk menarik kesimpulan dari sistem yang diwakili. (Banks [1]).*

### 2.2. Metode Kuadrat Terkecil (*Least Square*)

Diberikan suatu model linear sederhana :

$\mathrm{Y} = \mathrm{a} + \mathrm{bX} + \epsilon$ dengan *error* E($\epsilon$) = 0 dan variansi Var ($\epsilon$) = $\sigma^2$.

Apabila populasi dari suatu pasangan nilai ($X_i$,$Y_i$) diketahui, dapat dihitung nilai sebenarnya dari parameter a, b dan Var ($\epsilon$). Dalam prakteknya, tidak diketahui nilai parameter tersebut, akan tetapi dapat diestimasi (diduga) dengan menggunakan data empiris, yakni hasil observasi berdasarkan sampel yang ditarik dari populasi yang tidak terbatas (*infinite population*). Data empiris tersebut sering berupa data deret berkala (*time series data*) : $X_1,X_2,\ldots,X_i,\ldots,X_n$ dan $Y_1,Y_2,\ldots,Y_i,\ldots,Y_n$.

Untuk mengestimasi nilai a dan b digunakan metode kuadrat terkecil yaitu :

model sebenarnya : $\mathrm{Y} = \mathrm{a} + \mathrm{bX} + \epsilon$ dan

model estimasi (pendugaan) : $\mathrm{Y} = \mathrm{A} + \mathrm{BX} + \varepsilon$,

dengan A, B dan $\varepsilon$ estimator (penduga) dari a, b dan $\epsilon$.

Motode kuadrat terkecil adalah suatu metode untuk menghitung nilai A dan B sebagai estimator dari a dan b, sedemikian rupa sehingga jumlah galat memiliki nilai terkecil. Dengan bahasa matematika, dapat dinyatakan sebagai berikut :

$$\mathrm{Y} = \mathrm{A} + \mathrm{BX} + \epsilon_{\mathrm{i}} \quad \mathrm{i} = 1,2,\ldots\mathrm{n}.$$

dengan $\epsilon_{\mathrm{i}} = \mathrm{Y}_{\mathrm{i}} - (\mathrm{A} + \mathrm{BX}_{\mathrm{i}})$ = galat (*error*) sehingga dapat diperoleh :

$\sum \epsilon_{\mathrm{i}}^2 = \sum [\mathrm{Y}_{\mathrm{i}} - (\mathrm{A} + \mathrm{BX}_{\mathrm{i}})]^2$ = jumlah galat kuadrat.

Jadi, metode kuadrat terkecil adalah metode untuk menghitung A dan B sedemikian rupa sehingga $\sum \epsilon_{\mathrm{i}}^2$ = terkecil (minimum).

## 3. METODE PENELITIAN

Penelitian ini merupakan penelitian studi literatur. Metodologi yang digunakan adalah mengumpulkan bahan tulisan dari buku-buku dan jurnal maupun makalah yang membahas konsep model linear parsial, selanjutnya dilakukan analisa dan disusun dalam bentuk laporan penelitian.

## 4. HASIL DAN PEMBAHASAN

### 4.1. Regresi Linear Parsial

Bentuk dari model regresi linear parsial (semiparametrik) didefiniskan :

$$Y_i = \mathbf{X}_i^T \boldsymbol{\beta} + g(\mathbf{T}_i) + \epsilon_i, \quad i = 1,2,\ldots,n. \tag{1}$$

dengan $Y_i$ adalah nilai variabel tak bebas ke-i, (bernilai skalar), $\mathbf{X}_i = (x_{i1},\ldots, x_{ip})^T$ dan $\mathbf{T}_i = (t_{i1},\ldots, t_{ip})^T$ masing-masing merupakan vektor dari variabel bebas, $(\mathbf{X}_i,\mathbf{T}_i)$ adalah titik-titik desain random terdistribusi secara identik dan independen *(fixed design points)*, $\boldsymbol{\beta} = (\beta_1,\ldots, \beta_p)^T$ adalah vektor dari parameter yang tidak diketahui, $g(\mathbf{T}_i)$ adalah suatu fungsi yang tidak diketahui dari $R^n$ ke $R^1$, $\epsilon_i$ adalah *error-error* random (variabel galat) independen dengan mean $(\epsilon_i) = 0$ dan variansi $\sigma_i^2 = E(\epsilon_i^2)$.

### 4.1.1 Estimasi Model Linear Parsial

Dalam model linear parsial pada persamaan (1) g dan $\boldsymbol{\beta}$ merupakan fungsi dan parameter yang akan diestimasi dari data. Estimasi g menggunakan estimator kernel dan estimasi parameter $\boldsymbol{\beta}$ menggunakan metode kuadrat terkecil. Jika $\boldsymbol{\beta}$ merupakan parameter sebenarnya maka persamaan (5.10) dapat dinyatakan sebagai:

$$Y_i - \mathbf{X}_i^T \beta = g(\mathbf{T}_i) + \epsilon_i \tag{2}$$

Jika $Y_i^* = Y_i - \mathbf{X}_i^T\boldsymbol{\beta}$, maka dengan menggunakan teori regresi diperoleh:

$$g(t) = E(Y^* \quad T = t) = \frac{\int y^* \ f\ (t, y^*)\, dy^*}{f(t)} \tag{3}$$

Untuk fungsi densitas bersama $g(t, y^*)$ diestimasi dengan pergandaan kernel:

$$\hat{f}_{h1h2}\ (t, y^*) \quad = \quad n^{-1} \textstyle\sum_{i=I}^{n} K_{hI}\ (t - T_i)\ K_{h2}\ (y^* - Y_i^*)$$

$$\int y^* \hat{f}\, h_1h_2\ (t, y^*) dy^* = n^{-1} \textstyle\sum_{i=I}^{n} K_{hi}\ (t - T_i) \int \frac{y^*}{h_2} K\left(\frac{y^* - Yi^*}{h_2}\right) dy^*$$

Jika $\frac{y^* - Y_i^*}{h_2} = S$ maka diperoleh $y^* = Sh_2 + y_1^*$ dan $dy^* = dS$, selanjutnya di peroleh :

$$= n^{-1} \textstyle\sum_{i=I}^{n} K_{hi}\ (t - T_i) \int (Sh_2 + Y_i^*)\ K\ (s)\ ds$$

$$= n^{-1} \textstyle\sum_{i=I}^{n} K_{hi}\ (t - T_i)\ Y_i^*$$

sehingga diperoleh:

$$\hat{g}_n\ (t) = \frac{\sum_{i=1}^{n} K_h\ (t - T_i)\, Y_i^*}{\sum_{i=1}^{n} K_h\ (t - T_j)} = \textstyle\sum_{i=1}^{n} W_n\ (t)\, \left(Y_i - X_i^T\ \boldsymbol{\beta}\right) \tag{4}$$

dengan :

$$W_{ni}(t) = \frac{K_h(t - T_i)}{\sum_{i=1}^{n} K_h(t - T_i)} = \frac{K\left(\frac{t - T_i}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{t - T_i}{h}\right)}$$

Selanjutnya $\hat{g}_n\ (t)$ menggantikan $g(t)$ persamaan (5.10), sehingga didapat model estimasi sebagai berikut : $Y_i = \mathbf{X}_i^T \boldsymbol{\beta} + \hat{g}(\mathbf{T}_i) + \epsilon_i$ (5)

Berdasarkan model di atas, dengan menggunakan metode kuadrat terkecil diperoleh estimasi $\boldsymbol{\beta}$ dengan meminimumkan jumlah kuadrat *error*nya maka diperoleh keadaan sebagai berikut :

$$
\begin{aligned}
\sum \epsilon_i^2 &= \sum_{i=1}^{n}\left[\mathrm{Y_i} - \mathbf{X}_i^T\beta + \hat{g}_n(\mathbf{T}_i)\right]^2 \\
&= \sum_{i=1}^{n}\left[\mathrm{Y_i} - \mathbf{X}_i^T\boldsymbol{\beta} - \sum_{i=1}^{n} W_{nj}\,(\mathrm{T_i})(\mathrm{Y_i} - \mathbf{X}_i^T\boldsymbol{\beta}\right]^2 \\
&= \sum_{i=1}^{n}\left[\left(\mathrm{Y_i} - \sum_{i=1}^{n} W_{nj}\,(\mathrm{T_i})\mathrm{Y_i}\right) - \left(\mathbf{X}_i^T - \sum_{i=1}^{n} W_{nj}\,(\mathrm{T_i})\mathbf{X}_i^T\right)\boldsymbol{\beta}\right]^2 \\
&= \sum_{i=1}^{n}\left[\tilde{\mathrm{Y}}_i - \tilde{\mathbf{X}}_i\boldsymbol{\beta}\right]^2 \\
&= \left(\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}\boldsymbol{\beta}\right)^T(\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}\boldsymbol{\beta}) \\
&= \tilde{\mathbf{y}}^{\mathbf{T}}\tilde{\mathbf{y}} + \boldsymbol{\beta}^{\mathbf{T}}\tilde{\mathbf{x}}^{\mathbf{T}}\tilde{\mathbf{y}} - \tilde{\mathbf{y}}^{\mathbf{T}}\tilde{\mathbf{x}}\boldsymbol{\beta} + \boldsymbol{\beta}^{\mathbf{T}}\tilde{\mathbf{x}}^{\mathbf{T}}\tilde{\mathbf{x}}\boldsymbol{\beta}
\end{aligned}
$$

Karena $\boldsymbol{\beta}^{\mathbf{T}}\tilde{\mathbf{x}}^{\mathbf{T}}\tilde{\mathbf{y}} = (\boldsymbol{\beta}^{\mathbf{T}}\tilde{\mathbf{x}}^{\mathbf{T}}\tilde{\mathbf{y}})^{\mathbf{T}} = \tilde{\mathbf{y}}^{\mathbf{T}}\tilde{\mathbf{x}}\boldsymbol{\beta}$ maka diperoleh persamaan menjadi :

$$
\begin{aligned}
\in^{\mathbf{T}}\in &= \tilde{\mathbf{y}}^{\mathbf{T}}\tilde{\mathbf{y}} + 2\boldsymbol{\beta}^{\mathbf{T}}\tilde{\mathbf{x}}^{\mathbf{T}}\tilde{\mathbf{y}} + \boldsymbol{\beta}^{\mathbf{T}}\tilde{\mathbf{x}}^{\mathbf{T}}\tilde{\mathbf{x}}\boldsymbol{\beta} \\
\frac{\partial \in^{\mathbf{T}}\in}{\partial\beta} &= 0
\end{aligned}
$$

maka diperoleh :

$$
\begin{aligned}
-\mathbf{2}\tilde{\mathbf{x}}^{\mathbf{T}}\tilde{\mathbf{y}} + \mathbf{2}\tilde{\mathbf{x}}^{\mathbf{T}}\tilde{\mathbf{x}}\hat{\boldsymbol{\beta}}_n &= 0 \\
\tilde{\mathbf{x}}^{\mathbf{T}}\tilde{\mathbf{y}} &= \tilde{\mathbf{x}}^{\mathbf{T}}\tilde{\mathbf{x}}\hat{\boldsymbol{\beta}}_n \qquad (6)
\end{aligned}
$$

Persamaan (6) di atas merupakan persamaan normal. Dengan asumsi $\tilde{\mathbf{x}}$ mempunyai rank penuh (*fullrank*), maka persamaan (6) mempunyai persamaan tunggal :

$$\hat{\boldsymbol{\beta}}_n = (\tilde{\mathbf{x}}^{\mathbf{T}}\tilde{\mathbf{y}})^{-\mathbf{1}}\tilde{\mathbf{x}}^{\mathbf{T}}\tilde{\mathbf{y}}$$

Karena $\frac{\partial^2 \ \in^{\mathbf{T}}\in}{\partial\beta\,\partial\beta^T} = 2\ \tilde{\mathbf{x}}^{\mathbf{T}}\ \tilde{\mathbf{x}}$ merupakan matriks definit positif maka $\hat{\boldsymbol{\beta}}_n$ merupakan estimator kuadrat terkecil untuk $\boldsymbol{\beta}$ dengan :

$$
\begin{aligned}
\tilde{\mathbf{X}}_i &= \left[\mathbf{X}_i^T - \sum_{i=1}^{n} W_{ni}\,(\mathrm{T_i})\mathbf{X}_i^T\right] \\
\tilde{\mathbf{Y}}_i &= \left[\mathrm{Y_i} - \sum_{i=1}^{n} W_{nj}\,(\mathrm{T_i})\mathrm{Y_i}\right] \\
\tilde{\mathbf{Y}}^T &= \left(\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_n\right) \\
\tilde{\mathbf{X}}^T &= \left(\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n\right)
\end{aligned}
$$

Setelah mendapatkan $\hat{\boldsymbol{\beta}}_n$ maka $\boldsymbol{\beta}$ pada persamaan (5.13) diganti dengan $\hat{\boldsymbol{\beta}}_n$ untuk menghitung $\hat{g}_n(t)$ dan di peroleh $\hat{g}_n\,(t) = \sum_{i=1}^{n} W_{ni}\left(\mathrm{Y_i} - \mathbf{X}_i^T\,\hat{\boldsymbol{\beta}}_n\right)$.

Berdasarkan uraian di atas maka diperoleh estimasi model linear persial dengan menggunakan metode kuadrat terkecil akan berbentuk :

$\hat{g}_n\,(t) = \sum_{i=1}^{n} W_{ni}\left(\mathrm{Y_i} - \mathbf{X}_i^T\,\hat{\boldsymbol{\beta}}_n\right)$ dengan $\hat{\boldsymbol{\beta}}_n = (\tilde{\mathbf{x}}^{\mathbf{T}}\tilde{\mathbf{y}})^{-\mathbf{1}}\tilde{\mathbf{x}}^{\mathbf{T}}\tilde{\mathbf{y}}$ .

### 4.2. Simulasi Estimasi Model Linear Parsial Menggunakan Program S-Plus.

#### 4.2.1. Permasalahan dan Pembahasan Simulasi

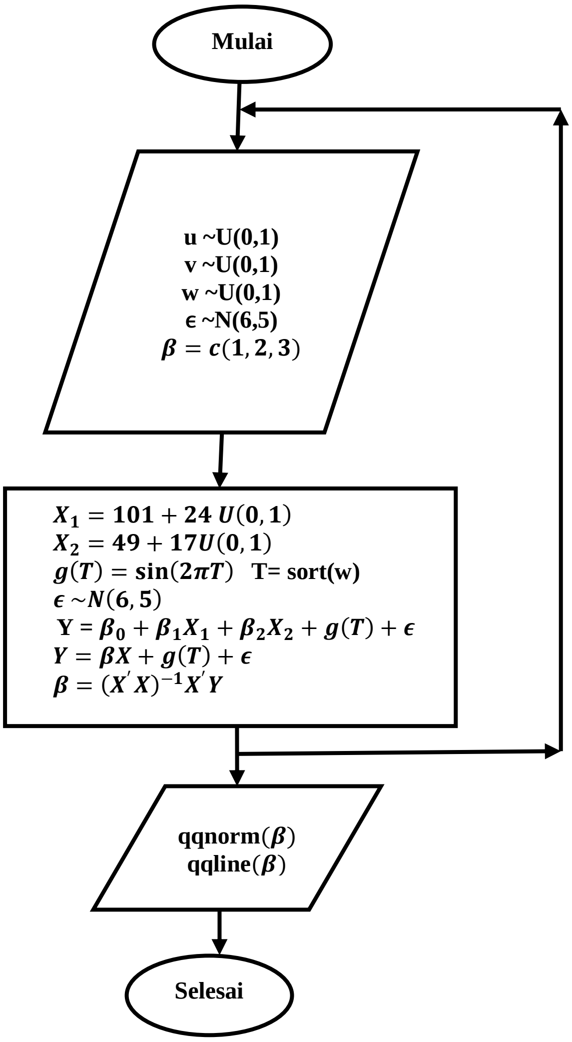
Simulasi estimasi model linear parsial dilakukan dengan mengambil dengan mengambil salah satu contoh permasalahan serta membahasnya yaitu : simulasikan estimasi model linear parsial $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + g(T) + \epsilon$ dengan $X_1 = 101 + 24\,U(0{,}1)$, $X_2 = 49 + 17U(0{,}1)$, $g(T) = \sin(2\pi T)$, $\epsilon \sim N(6{,}5)$.

Pembahasan simulasi estimasi dari masalah di atas adalah sebagai berikut :

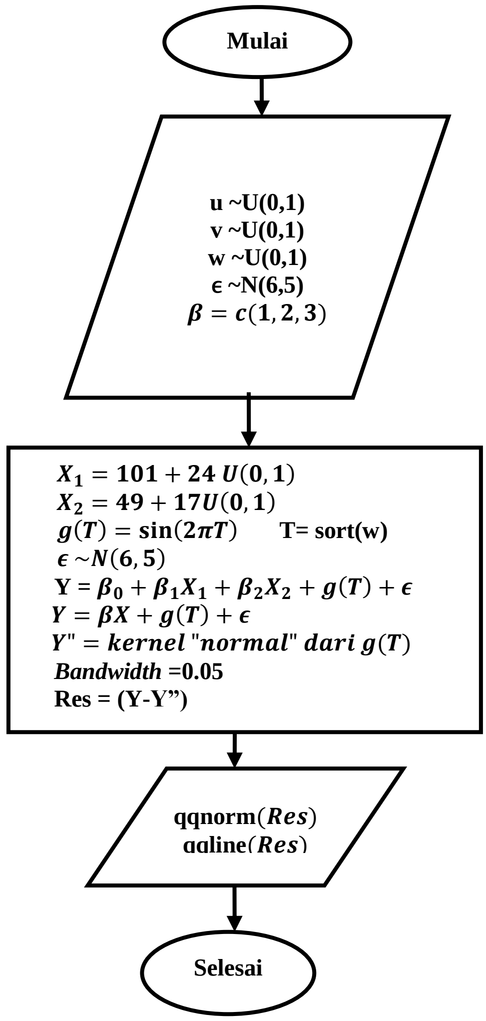
### 4.2.1.1. Algoritma Program

Adapun algoritma program simulasi estimasi model linear parsial menggunakan program S-Plus, yaitu :

a. Estimasi bagian parametrik yaitu estimator beta ($\beta$) dan b). Estimasi bagian nonparametrik yaitu estimator g(T) adalah sebagai berikut :




a




b

### 4.2.1.2. Membuat Fungsi Estimasi Parameter Pada Bagian Parametrik Model Regresi Semiparametrik

Adapun estimasi pada bagian parametrik dari model regresi semiparametrik yaitu :

1. Estimasi Parameter Beta 0 ($\beta 0$)

Bentuk fungsi estimasi parameter beta0 ($\beta 0$) adalah sebagai berikut :

```
> beta0<-function(n,r)
+ {
+ }
> fix(beta0)
   function(n,r)
{
```

```
        x<-matrix(0,n,3)
        x[,1]<-rep(1,n)
        u<-runif(n)
        x[,2]<-100*x[,1]+25*u
        v<-runif(n)
        x[,3]<-50*x[,1]+16*v
        b<-c(1,2,3)
        beta0<-rep(0,r)
        for(i in 1:r) {
            e<-rnorm(n,5,2)
            w<-runif(n)
            h<-sort(w)
            gt<-sin(2*pi*h)
            y<-x%*%b+gt+e
            xt<-t(x)
            xtx<-solve(xt%*%x)
            xty<-xt%*%y
            beta<-xtx%*%xty
            beta0[i]<-beta[1]
         }
        qqnorm(beta0)
        qqline(beta0)
}
```

```
> win.graph()
> par(mfrow=c(4,3))
> beta0(100,60)
 intercept    slope
  6.133159 3.922958
> beta0(100,90)
 intercept    slope
  5.919716 5.117378
> beta(100,140)
> beta0(100,140)
 intercept    slope
  6.648221 4.304291
```

```
> beta0(1000,60)
 intercept    slope
  6.209254 1.310699
> beta0(1000,90)
 intercept     slope
  5.308048 0.9806193
> beta0(1000,140)
 intercept    slope
  4.778199 1.335617
```

```
> beta0(500,60)
 intercept    slope
   6.42679 1.538664
> beta0(500,90)
 intercept    slope
  5.512805 1.536441
> beta0(500,140)
 intercept    slope
  4.573788 1.433025
```

```
> beta0(1700,60)
 intercept     slope
  5.640636 0.9140424
> beta0(1700,90)
 intercept     slope
  5.906575 0.9756339
> beta0(1700,140)
 intercept    slope
  5.725705 1.007656
```

2. Estimasi Parameter Beta1 ($\beta 1$)

Bentuk fungsi estimasi parameter beta1 ($\beta 1$) adalah sebagai berikut :

```
> beta1<-beta0
> fix(beta1)

function(n, r)
{
        x <- matrix(0, n, 3)
        x[, 1] <- rep(1, n)
        u <- runif(n)
        x[, 2] <- 100 * x[, 1] + 25 * u
        v <- runif(n)
        x[, 3] <- 50 * x[, 1] + 16 * v
        b <- c(1, 2, 3)
        beta1 <- rep(0, r)
        for(i in 1:r) {
                e <- rnorm(n, 5, 2)
                w <- runif(n)
                h <- sort(w)
                gt <- sin(2 * pi * h)
                y <- x %*% b + gt + e
                xt <- t(x)
                xtx <- solve(xt %*% x)
                xty <- xt %*% y
                beta <- xtx %*% xty
                beta1[i] <- beta[2]
        qqnorm(beta1)
        qqline(beta1)
```

```
> win.graph()
> par(mfrow=c(4,3))
> beta1(100,60)
 intercept      slope
  2.001763 0.03243128
> beta1(100,90)
 intercept      slope
  2.020575 0.02947815
> beta1(100,140)
 intercept      slope
  2.012777 0.02746861
> beta1(500,60)
 intercept      slope
   1.99559 0.01198506
> beta1(500,90)
 intercept      slope
  1.990491 0.01210625
> beta1(500,140)
 intercept      slope
  1.998861 0.01252757
  1.995953 0.008512038
```

```
> beta1(1000,60)
 intercept      slope
  1.995953 0.008512038
> beta1(1000,90)
 intercept      slope
  1.998538 0.009595129
> beta1(1000,140)
 intercept      slope
  1.992354 0.007681676
> beta1(1700,60)
 intercept      slope
  1.998964 0.006564065
> beta1(1700,90)
 intercept      slope
  1.999104 0.006924665
> beta1(1700,140)
 intercept      slope
  2.000598 0.007215749
```

3. Estimasi Parameter Beta2 ($\beta 2$)

Bentuk fungsi estimasi parameter beta2 ($\beta 2$) adalah sebagai berikut :

```
> beta2<-beta0
> fix(beta2)
>beta2
function(n, r)
{
	x <- matrix(0, n, 3)
	x[, 1] <- rep(1, n)
	u <- runif(n)
	x[, 2] <- 100 * x[, 1] + 25 * u
	v <- runif(n)
	x[, 3] <- 50 * x[, 1] + 16 * v
	b <- c(1, 2, 3)
	beta2 <- rep(0, r)
	for(i in 1:r) {
		e <- rnorm(n, 5, 2)
		w <- runif(n)
		h <- sort(w)
		gt <- sin(2 * pi * h)
		y <- x %*% b + gt + e
		xt <- t(x)
		xtx <- solve(xt %*% x)
		xty <- xt %*% y
		beta <- xtx %*% xty
		beta2[i] <- beta[3]
	}
	qqnorm(beta2)
	qqline(beta2)
}
```

```
> win.graph()
> par(mfrow=c(4,3))
> beta2(100,60)
 intercept      slope
  3.019461 0.04276547
> beta2(100,90)
 intercept      slope
  3.004528 0.04574629
> beta2(100,140)
 intercept      slope
  2.988473 0.04374792
> beta2(500,60)
 intercept    slope
  3.005658 0.019245
> beta2(500,90)
 intercept      slope
  2.992396 0.02100095
> beta2(500,140)
 intercept      slope
  3.005576 0.01996945
> beta2(1000,60)
 intercept      slope
  2.998361 0.01243464
> beta2(1000,90)
 intercept      slope
  2.999569 0.01546367
> beta2(1000,140)
 intercept      slope
   3.00596 0.01553633
> beta2(1700,60)
 intercept      slope
  3.000804 0.01023663
> beta2(1700,90)
 intercept       slope
  2.999323 0.008055285
> beta2(1700,140)
 intercept      slope
   2.99026 0.01136875
```

**4.2.1.3. Membuat Fungsi Estimasi Parameter pada Bagian Nonparametrik yaitu g(T) Model Regresi Semiparametrik**

Adapun estimasi pada bagian nonparametrik yaitu g(T) model regresi semiparametrik yaitu :

```
> gete<-function(n)
+ {
+ }
> fix(gete)
function(n)
{
        x<-matrix(0,n,3)
        x[,1]<-rep(1,n)
        u<-runif(n)
        x[,2]<-101*x[,1]+24*u
        v<-runif(n)
        x[,3]<-49*x[,1]+17*v
        b<-c(1,2,3)
        e<-rnorm(n,5,2)
        w<-runif(n)
        h<-sort(w)
        gt<-sin(2*pi*h)
        y<-x%*%b+gt+e
        g<-ksmooth(h,y, ker = "normal", bandwidth = 0.05)
        gy<-g$y
        res<-(y-gy)
        qqnorm(res)
        qqline(res)
```

**}**

```
> win.graph()
> par(mfrow=c(3,2))
>gete(60)
 intercept    slope
  1.495798 20.13909
> gete(90)
  intercept    slope
 -0.7991943 15.24517
> gete(140)
  intercept    slope
 -0.4320971 17.89737
> gete(500)
 intercept    slope
 0.3542298 21.20898
> gete(1700)
   intercept    slope
 -0.04457853 21.65917
> gete(2000)
 intercept    slope
 0.4421019 22.30143
```

## 5. KESIMPULAN

Berdasarkan hasil dan pembahasan maka diperoleh hasil estimasi dan hasil simulasi *output* baik berupa nilai maupun grafik dengan program S-Plus, dapat ditarik beberapa kesimpulan yaitu :

1. Estimasi model linear persial dengan menggunakan metode kuadrat terkecil akan berbentuk :
   $\hat{g}_n\,(t) = \sum_{i=1}^{n} W_{ni}\left(Y_i - \mathbf{X}_i^T\,\hat{\boldsymbol{\beta}}_n\right)$ dengan $\hat{\boldsymbol{\beta}}_n = (\tilde{\mathbf{x}}^{\mathbf{T}}\,\tilde{\mathbf{y}})^{-\mathbf{1}}\,\tilde{\mathbf{x}}^{\mathbf{T}}\,\tilde{\mathbf{y}}$ .

2. Bagian Parametrik
   1. Nilai $\epsilon \sim N(5{,}4)$ hanya mempengaruhi estimator beta 0 ($\beta_0$), dalam program simulasi ini dimisalkan $\beta_0$ = 1, sedangkan hasil estimasi $\beta_0$ berada dalam interval 5.5 < $\beta_0$ < 7.3. Sedangkan untuk estimator beta 1 ($\beta_1$) dan beta 2 ($\beta_2$), nilai yang diperoleh hampir sama dengan nilai yang dimisalkan di dalam list program S-Plus yaitu $\beta_1$ = 1 dan $\beta_2$ = 3.
   2. Berdasarkan tampilan grafik qqnorm dan qqline estimator beta ($\beta$) yaitu $\beta_0$, $\beta_1$ dan $\beta_2$ dapat terlihat dengan jelas, dimana jika n semakin besar ($n \to \infty$) dan semakin besar replikasi r perulangan, maka titik-titik yang menyebar disekitar garis lurus makin banyak dan mendekati garis lurus. Hal ini menunjukkan makin besar n dan r, maka beta ($\beta$) semakin mendekati distribusi normal.
3. Bagian Nonparametrik
   Berdasarkan hasil simulasi berupa nilai maupun grafik program S-Plus terlihat bahwa jika n diambil semakin membesar maka titik-titik semakin banyak menyebar disekitar garis lurus dan jika n di ambil semakin besar lagi ($n \to \infty$) maka titik-titik akan membentuk lambang integral yaitu $\int$ . Artinya titik-titik yang berada di atas dan di bawah garis, baik jarak maupun jumlahnya relatif sama (sifat simetris), akibatnya nilai residunya akan cenderung mendekati nol (0). Hal ini, berarti estimator bagian nonparametrik dalam hal ini yang diambil sebagai contoh fungsi kernel normal nilainya mendekati g(T). Jadi dapat diambil kesimpulan singkatnya yaitu jika n semakin besar ($n \to \infty$) maka estimator bagian nonparametriknya semakin mendekati g(T).

## 6. DAFTAR PUSTAKA

[1] Anton, H. 1987. *Aljabar Linear Elementer,* Edisi ke 5, Terjemahan Pantur Silaban Erlangga, Jakarta.
[2] Draper, R. N. and Smith, H. 1998. *Applied Regression Analysis*, Johan Wiley & Sons,INC.
[3] Hardle, W. 1990. *Applied Nonparametric Regression*, The Syndicate of the Universitas of Cambridge.
[4] Hardle, W, Liang. H and Gao. J. 2000. *Partially Linear Model*, Physica-Verlag, Heidelberg.
[5] Harinaldi. 2005. *Prinsip-prinsip Statistika Untuk Teknik dan Sains*, Penerbit Airlangga. Jakarta.
[6] Montgomery, C.D and Peck, E.A. 1982. *Introduction to Linear Regression Analysis*, John Wiley & Sons.
[7] Myers, H.R and Milton, S.J. 1991. *A First Course in the Theory of Linear Statistical Models*, PWS-KENT Publising Company, Boston.
[8] Tirta, I. M, 2003. *Pengantar Metode Simulasi*, Diterbitkan oleh Unit Penerbit FMIPA Universitas Jember.
[9] Wang, Q, Linton. O and Hardle, W. 2003. *Semiparametric Regression Analysis with Missing at Random.* Journal of American Statistical Association, June 2004, Vol. 99, no. 466,