ISBN:1858-1382

J M S K
Jurnal Matematika, Statistika, & Komputasi
http://journal.unhas.ac.id/index.php/jmsk
Vol. 5, No.2, 66-75, Januari 2009

Regresi Semiparametrik dengan Pendekatan *Generalized Estimating Equation* (GEE)

Muhammad Abdy[†]

Abstrak

In generalized linear model, semiparametric model can be denoted as $l(\mu)=\boldsymbol{X'}\beta + \theta(\boldsymbol{T})$, where $\boldsymbol{X}$ and $\boldsymbol{T}$ are covariates, $\boldsymbol{\beta}$ is an unknown parameter vector and $\theta(.)$ is an unknown smooth function. For longitudinal data, estimation of parameter by using Least Squared and Maximum Likelihood is inappropriate because there is tendency correlation within subject. One of the methods to estimate parameter for such data is generalized estimating equation (GEE). In this paper, we will discuss estimation of parametric and nonparametric component in semiparametric regression model for longitudinal data by using GEE and asymptotic distribution of estimator of parametric component. In simulation study, it is obtained that value of the estimator of $\hat{\beta}_1$ and $\hat{\beta}_2$ are tend to same as parameter β_1 and β_2.

Kata Kunci : *Asymptotic, Generalized Estimating Equation, kernel, semiparametric regression.*

1. Pendahuluan

Analisis regresi telah memainkan peranan penting dalam studi tentang hubungan antara variabel respons dan variabel prediktor. Misalkan y adalah variabel respons dan t adalah variabel prediktor untuk n pengamatan, maka secara umum hubungannya dapat dituliskan sebagai berikut:

$$y_i = m(t_i) + \varepsilon_i \ ; \ i = 1, 2, \dots, n. \tag{1}$$

dimana $m(t_i)$ merupakan fungsi regresi yang tidak diketahui dan ε_i adalah sesatan acak yang diasumsikan berdistribusi normal independen dengan mean nol dan varians σ^2.

Tujuan analisis regresi adalah menentukan pendekatan model yang beralasan untuk fungsi $m(t_i)$. Untuk tujuan tersebut dapat digunakan metode regresi parametrik dan metode regresi nonparametrik. Pendekatan metode parametrik digunakan apabila ada informasi sebelumnya tentang bentuk $m(t_i)$ atau memenuhi asumsi bentuk kurva tertentu. Sedangkan pendekatan metode nonparametrik digunakan jika tidak ada informasi tentang bentuk $m(t_i)$ dan tidak tergantung pada asumsi bentuk kurva tertentu, sehingga memberikan fleksibilitas yang lebih besar.

Dalam perkembangan analisis regresi selanjutnya, dikenal regresi semiparametrik yang merupakan penggabungan dari model regresi parametrik dan nonparametrik. Model regresi ini diperkenalkan oleh Engle dkk (1986) dan Wahba (1984). Adapun model regresi semiparametrik dapat dirumuskan sebagai berikut (dalam notasi matriks):

$$\boldsymbol{Y} = \boldsymbol{X}\beta + \theta(\boldsymbol{T}) + \varepsilon \tag{2}$$

dimana β adalah suatu vektor parameter yang tak diketahui dan $\theta(.)$ adalah suatu fungsi mulus (smooth). Hubungan antara $\boldsymbol{Y}$ dan $\boldsymbol{X}$ merupakan model parametrik sedangkan hubungan antara $\boldsymbol{Y}$ dan $\boldsymbol{T}$ merupakan model nonparametrik.

[†] *Jurusan Matematika FMIPA Universitas Negeri Makassar*
Email: muh.abdy@unm.ac.id

Dalam *generalized linear model* (GLM), model dalam (2) di atas dapat ditulis sebagai:

$$l(\mu) = \boldsymbol{X}\boldsymbol{\beta} + \theta(\boldsymbol{T}), \tag{3}$$

dimana μ mean dari variabel prediktor dan $l(.)$ suatu fungsi yang diketahui.

Untuk data non-longitudinal (*cross-sectional*), telah banyak penulis yang mengembangkan metode estimasi model semiparametrik (3). Green (1987) mengusulkan suatu estimasi pada model semiparametrik dengan memaksimumkan fungsi likelihood penalized. Hastie dan Tibshirani (1990) membahas suatu pendekatan estimasi yang didasarkan pada algoritma backfitting. Kemudian Severini dan Staniswalis (1994) mengusulkan suatu estimasi pada model semiparametrik dengan pendekatan quasi-likelihood.

Untuk data longitudinal, model semiparametrik (3) diatas berbentuk:

$$l(\mu_{ij}) = \boldsymbol{X}_{ij}\boldsymbol{\beta} + \theta(\boldsymbol{T}_{ij}) \tag{4}$$

dimana indeks i menyatakan unit pengamatan dan indeks j menyatakan ulangan pengamatan. Untuk mengestimasi model parametrik, yaitu model dalam (4) yang tidak memuat $\theta(\boldsymbol{T}_{ij})$, telah dikembangkan oleh Liang dan Zeger (1986). Mereka mengestimasi parameter $\boldsymbol{\beta}$ dengan menggunakan persamaan :

$$\sum_{i=1}^{n}\left[\frac{\partial g(\boldsymbol{X}_i\boldsymbol{\beta})}{\partial\boldsymbol{\beta}}\right]^T \boldsymbol{V}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{\mu}_i) = \boldsymbol{0} \tag{5}$$

dimana g adalah fungsi invers dari fungsi link dan $\boldsymbol{V}_i$ adalah matriks kovarians $\boldsymbol{Y}_i$. Persamaan ini lebih popular dikenal sebagai *Generalized Estimating Equation* (GEE). Liang dan Zeger (1986) memperlihatkan bahwa estimator GEE $\hat{\boldsymbol{\beta}}$ berdistribusi normal asimptotik dan konsisten. Untuk mengestimasi model nonparametrik, yaitu model dalam (4) yang tidak memuat $\boldsymbol{X}_{ij}\boldsymbol{\beta}$, telah dikembangkan oleh Lin dan Carroll (2000). Mereka menggunakan pendekatan estimator kernel polinomial untuk mengestimasi fungsi mulus $\theta(\boldsymbol{T}_{ij})$. Bentuk persamaan estimasi yang digunakan adalah:

$$\sum_{i=1}^{n}\boldsymbol{G}_i^T(t)\boldsymbol{D}_i\boldsymbol{V}_{2i}^{-1}\boldsymbol{K}_{ih}(t)\{\boldsymbol{Y}_i - \boldsymbol{\mu}_i\} = \boldsymbol{0},$$

dimana $\boldsymbol{D}_i$ adalah suatu matriks diagonal dengan elemen diagonal $g'(\boldsymbol{G}_i^T\boldsymbol{\alpha})$ dan $\boldsymbol{K}_{ih}$ adalah matriks diagonal dengan elemen diagonal $\boldsymbol{K}_h(t)$.

Dalam tulisan ini akan dikaji bentuk estimator parameter $\boldsymbol{\beta}$ dan fungsi mulus $\theta(\boldsymbol{T}_{ij})$ dalam model semiparametrik (4) secara bersama-sama dengan menggunakan modifikasi metode estimasi yang diusulkan oleh Liang dan Zeger (1986) dan Lin dan Carroll (2000).

Perbedaan kemakmuran dan standar kehidupan suatu daerah atau negara dengan daerah atau negara lain sering menimbulkan pertanyaan yang sulit dijelaskan. Mengapa suatu daerah atau negara yang satu lebih berkembang dari daerah lainnya. Faktor apa saja yang mempengaruhi perkembangan suatu daerah. Faktor apa saja yang mempengaruhi kemakmuran suatu daerah atau negara. Berapa lama waktu yang diperlukan untuk menuju kemakmuran.

Memang tidak dapat dipungkiri bahwa dunia sekarang lebih makmur dari dunia masa lalu. Fenomena ini, tentunya sangat menarik untuk dikaji. Banyak model telah dibuat dalam upaya menerangkan fenomena ini. Ada model yang dapat bertahan, demikian pula banyak model matematika yang hilang dengan sendirinya yang disebabkan dengan ketidaksuaian dengan fenomena yang ada. Diantara segelintir model yang dapat bertahan adalah model Solow yang bertumpu pada empat pilar pertumbuhan yakni hasil produksi (Y), modal/kapital (K), tenaga kerja/labour (L) dan pengetahuan atau efektifitas tenaga kerja (A) yang semuanya merupakan fungsi yang kontinu terhadap waktu.

2. *Generalized Estimating Equation* (GEE)

Generalized Estimating Equation (GEE) merupakan salah satu metode untuk mengestimasi parameter regresi. Metode ini diperkenalkan pertama kali oleh Liang dan Zeger (1986), dan banyak dipakai dalam konteks analisis data longitudinal. Kerangka dasar dari methodologi GEE adalah sebagai berikut:

Misalkan $Y_i = (Y_{i1}, ..., Y_{im_i})^T$ adalah vektor m_i x 1 yang menyatakan variabel respons, $X_i = (X_{i1}, ..., X_{im_i})^T$ adalah matriks m_i x p yang merupakan nilai dari variabel prediktor untuk subjek ke-i (i = 1, …, n ; j = 1, …, m_i) dan μ_{ij} menyatakan ekspektasi bersyarat dari respons Y_{ij}, diberikan matriks X_i, yaitu: $\mu_{ij} = E(Y_{ij} \mid X_i)$, serta diasumsikan hubungan $l(\mu_{ij}) = X_{ij}\beta$, dengan l adalah fungsi link. Selanjutnya, misalkan $\text{var}(Y_{ij}) = \phi V(\mu_{ij})$, dimana V adalah fungsi dari μ_{ij} yang diketahui dan ϕ adalah suatu parameter dispersi. Jika varians Y_i pada quasi-likelihood diganti dengan matriks kovarians Y_i, yaitu matriks yang berbentuk $V_i = A_i^{1/2} R_i A_i^{1/2}$, dimana A_i = diag $(\phi V(\mu_{ij}(\beta))$ dan elemen ke-(j, k) dari R_i adalah $\text{Corr}[Y_{ij}, Y_{ik}]$, maka persamaan estimasi quasi-likelihood akan menjadi suatu persamaan yang biasa diistilahkan dengan Generalized Estimating Equation, yang berbentuk :

$$U(\beta) = \sum_{i=1}^{n} \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}} \right)^t \boldsymbol{V}_i^{-1} (\boldsymbol{Y}_i - \boldsymbol{\mu}_i) = 0, \qquad (6)$$

dimana $\mu_i = (\mu_{i1}, ..., \mu_{in_i})^t$. Estimator untuk β, dinyatakan dengan $\hat{\boldsymbol{\beta}}_{GEE}$, diperoleh dengan menyelesaikan persamaan estimasi: $U(\beta) = \sum_{i=1}^{n} \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}} \right)^t \boldsymbol{V}_i^{-1} (\boldsymbol{Y}_i - \boldsymbol{\mu}_i) = 0$.

Liang dan Zeger (1986) memperlihatkan bahwa estimator $\hat{\boldsymbol{\beta}}_{GEE}$ berdistribusi normal asimptotik, yaitu: $\sqrt{n}\left(\hat{\boldsymbol{\beta}}_{GEE} - \boldsymbol{\beta}\right) \rightarrow N_p(0, V_{GEE})$ untuk $n \rightarrow \infty$,

dimana $V_{GEE} = \lim_{N \to \infty} (A \times B \times C)$,

dengan

$$A = \left\{ \sum_{i=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}} \right)^t \boldsymbol{V}_i^{-1} \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}} \right) \right\}^{-1},$$

$$B = \left\{ \sum_{i=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}} \right)^t \boldsymbol{V}_i^{-1} \text{cov}[\boldsymbol{Y}_i] \boldsymbol{V}_i^{-1} \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}} \right) \right\},$$

$$C = \left\{ \sum_{i=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}} \right)^t \boldsymbol{V}_i^{-1} \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}} \right) \right\}^{-1}.$$

3. Estimasi Model Semiparametrik

Diberikan model regresi semiparametrik:

$$\boldsymbol{Y}_{ij} = \boldsymbol{X}_{ij}\boldsymbol{\beta} + \boldsymbol{\theta}(\boldsymbol{T}_{ij}) + \boldsymbol{\varepsilon}_{ij}, \quad i=1,..., n\,;\, j=1,...,m \qquad (7)$$

dimana $\boldsymbol{Y}_{ij}$ adalah variable respons kelompok ke-i untuk pengamatan ke-j, $(\boldsymbol{X}_{ij}, \boldsymbol{T}_{ij})$ adalah kovariat dan $\boldsymbol{\theta}(\boldsymbol{T}_{ij})$ adalah komponen nonparametrik yang merupakan fungsi mulus yang tidak diketahui. Dalam model GLM, bentuk (7) di atas dapat dinyatakan dalam bentuk:

$$l(\mu_{ij}) = \boldsymbol{X}_{ij}\boldsymbol{\beta} + \boldsymbol{\theta}(\boldsymbol{T}_{ij}) \qquad (8)$$

dimana μ_{ij} = E[$\boldsymbol{Y}_{ij}$| $\boldsymbol{X}_{ij}$] dan l(.) adalah suatu fungsi link. Fungsi mulus $\theta(t)$ akan diestimasi dengan persamaan estimasi kernel dan $\boldsymbol{\beta}$ akan diestimasi dengan persamaan estimasi profil (suatu persamaan yang diperoleh dengan cara mencari terlebih dahulu estimasi $\theta(t)$ kemudian mengkonstruksi suatu persamaan estimasi dengan mengetahui $\theta(t)$ untuk mencari estimasi $\boldsymbol{\beta}$). Pendekatan persamaan estimasi kernel untuk memperoleh bentuk estimasi $\theta(t)$ akan diuraikan sebagai berikut:

Misalkan bahwa $\theta(.)$ adalah fungsi polinom orde p, yaitu: $\theta(.) = \boldsymbol{G}_{\boldsymbol{p}}^{\boldsymbol{T}}(.)\alpha$, dimana $\boldsymbol{G}_{\boldsymbol{p}}(z)$ =(1, z, …, z^p)$^{\mathrm{T}}$ dan $\boldsymbol{\alpha} = (\alpha_0,...,\alpha_p)^{\mathrm{T}}$. Misalkan $\boldsymbol{Y}_i$=(Y_{i1},…, Y_{im})$^{\mathrm{T}}$ dan $\boldsymbol{G}_{ip}$=[$\boldsymbol{G}_p$(T_{i1}),…,$\boldsymbol{G}_p$(T_{im})]$^{\mathrm{T}}$. Koefisien regresi $\boldsymbol{\alpha}$ dapat diestimasi dengan menggunakan GEE, yaitu menyelesaikan persamaan estimasi:

$$\sum_{i=1}^{n} \boldsymbol{G}_{ip}^{T}\boldsymbol{D}_i\boldsymbol{V}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{\mu}_i) = \boldsymbol{0} \qquad (9)$$

dimana $\boldsymbol{\mu}_i = E(\boldsymbol{Y}_i)$ yang mempunyai komponen ke-j $\mu_{ij} = g[\boldsymbol{G}_p^T(\boldsymbol{T}_{ij})\boldsymbol{\alpha}]$, $\boldsymbol{D}_{\boldsymbol{i}} = diag\{g'(\boldsymbol{G}_p^T(\boldsymbol{T}_{ij})\boldsymbol{\alpha})\}$, g(.) merupakan fungsi invers dari fungsi link dari l(.), $\boldsymbol{V}_i$ = $\boldsymbol{A}_i^{1/2}\boldsymbol{R}_i\boldsymbol{A}_i^{1/2}$ dimana $\boldsymbol{A}_i$ = diag{$\phi \boldsymbol{V}(\mu_{ij})$} dan $\boldsymbol{R}_i$ adalah matriks korelasi yang invertibel. Selanjutnya, apabila θ(.) suatu fungsi nonparametrik, maka estimator $\hat{\theta}(.)$ dapat diperoleh dengan memfitkan polinom $\alpha_0 + \alpha_1(.-t) + ... + \alpha_p(.-t)^p$ terhadap (t_i, $\boldsymbol{Y}_i$) dengan menggunakan GEE terboboti dengan pembobot kernel $K_h(T_{ij} - t)$. Nilai dari $\hat{\theta}(.)$ adalah nilai dari $\hat{\alpha}_0$, dimana $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_0,...,\hat{\alpha}_p)^{\mathrm{T}}$ solusi dari persamaan estimasi:

$$\sum_{i=1}^{n} \boldsymbol{G}_{ip}^{T}(t)\boldsymbol{D}_i\boldsymbol{V}_i^{-1}\boldsymbol{K}_{ih}(t)\{\boldsymbol{Y}_i - g(\boldsymbol{G}_p^T(t)\boldsymbol{\alpha})\} = \boldsymbol{0} \qquad (10)$$

dimana $\boldsymbol{K}_{ih}(t)$ = diag{$K_h(T_{ij} - t)$}. Untuk mengestimasi $\theta(t)$ pada model (8) akan digunakan orde p=1 (estimator kernel linier lokal), sehingga $\boldsymbol{G}_{ip}(t)$ adalah suatu matriks yang berukuran m_i x 2 dengan baris ke-j {1, (T_{ij} - t)/h}. Untuk mendapatkan nilai estimasi $\theta(t)$ pada model (8), maka terlebih dahulu diasumsikan bahwa $\hat{\boldsymbol{\beta}}$ diketahui (nilai initial), kemudian $\theta(t)$ diestimasi dengan menggunakan (10), yaitu menyelesaikan persamaan estimasi:

$$\sum_{i=1}^{n} \boldsymbol{G}_i^{T}(t)\boldsymbol{D}_i\boldsymbol{V}_{2i}^{-1}\boldsymbol{K}_{ih}(t)\{\boldsymbol{Y}_i - g(\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1(\boldsymbol{T}_i - \boldsymbol{t})/\boldsymbol{h}))\} = 0 \qquad (11)$$

dimana $\boldsymbol{K}_{ih}(t)$, $\boldsymbol{D}_i$, $\boldsymbol{V}_{2i}$ sama seperti pada (10) tetapi dievaluasi pada $\mu_{ij}(\boldsymbol{X}_{ij}, t) = g(\boldsymbol{X}_{ij}\boldsymbol{\beta} + \boldsymbol{\alpha_0} + \boldsymbol{\alpha_1}(\boldsymbol{T}_{ij} - \boldsymbol{t})/\boldsymbol{h})$. Setelah mengestimasi $\boldsymbol{\alpha}$ pada t, yaitu $\hat{\boldsymbol{\alpha}}(\boldsymbol{\beta})$, maka estimator $\theta(t)$ adalah $\hat{\theta}(\mathrm{t};\boldsymbol{\beta}) = \hat{\alpha}_0(\boldsymbol{\beta})$. Selanjutnya, parameter $\boldsymbol{\beta}$ pada model (8) diestimasi dengan menggunakan GEE parametrik (5). Proses pengestimasiannya dilakukan dengan menyelesaikan persamaan estimasi:

$$\sum_{i=1}^{n}\left[\frac{\partial g(\boldsymbol{X}_i\boldsymbol{\beta}+\hat{\boldsymbol{\theta}}(\boldsymbol{T}_i;\boldsymbol{\beta}))}{\partial\boldsymbol{\beta}}\right]^{\boldsymbol{T}}\boldsymbol{V}_{1i}^{-1}\{\boldsymbol{Y}_i-g(\boldsymbol{X}_i\boldsymbol{\beta}+\hat{\boldsymbol{\theta}}(\boldsymbol{T}_i;\boldsymbol{\beta}))\}=\mathbf{0} \qquad (12)$$

dimana

$\hat{\boldsymbol{\theta}}(\boldsymbol{T}_i;\boldsymbol{\beta}) = [\,\hat{\theta}(T_{i1};\boldsymbol{\beta}),\ldots,\hat{\theta}(T_{im_i};\boldsymbol{\beta})\,]^{\mathrm{T}}$, $\boldsymbol{V}_{1i}=\boldsymbol{S}_i^{1/2}\boldsymbol{R}_{1i}\boldsymbol{S}_i^{1/2}$,

$\boldsymbol{S}_i^{1/2}=\mathrm{diag}\{\phi\boldsymbol{V}(g(\boldsymbol{X}_{ij}^T\boldsymbol{\beta}+\hat{\theta}(\boldsymbol{T}_{ij};\boldsymbol{\beta}))\}$.

Untuk mendapatkan $\hat{\theta}(t)$ dan $\hat{\boldsymbol{\beta}}$ dilakukan dengan menyelesaikan (11) dan (12) secara iteratif dengan menggunakan algoritma scoring Fisher. Nilai $\hat{\theta}(t)$ diperoleh dengan mencari nilai dari:

$$\hat{\boldsymbol{\alpha}}^{(v)}=\hat{\boldsymbol{\alpha}}^{(v-1)}-\left[\sum_{i=1}^{n}[\boldsymbol{G}_i^T(t)\boldsymbol{W}_i(t)\boldsymbol{G}_i(t)]\right]^{-1}\sum_{\mathrm{i=1}}^{\mathrm{n}}\mathbf{G}_{\mathrm{i}}^{\mathrm{T}}(\mathrm{t})\mathbf{D}_{\mathrm{i}}\mathbf{V}_{2\mathrm{i}}^{-1}\mathbf{K}_{\mathrm{ih}}(\mathrm{t})\{\mathbf{Y}_{\mathrm{i}}-\boldsymbol{\mu}_{\mathrm{i}}\} \qquad (13)$$

dimana (v) adalah aproksimasi ke-v. Selanjutnya, nilai $\boldsymbol{\beta}$ dapat diperbaharui dengan menyelesaikan:

$$\hat{\boldsymbol{\beta}}^{(v)}=\hat{\boldsymbol{\beta}}^{(v-1)}-\left[\sum_{i=1}^{n}\tilde{\boldsymbol{X}}_i^T\boldsymbol{D}_i[\mathbf{V}_{1i}^{-1}]^T\boldsymbol{D}_i\tilde{\boldsymbol{X}}_i\right]^{-1}\left[\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}}\tilde{\boldsymbol{X}}_i^T\boldsymbol{D}_i\mathbf{V}_{1i}^{-1}\{\boldsymbol{Y}_i-\boldsymbol{\mu}(\boldsymbol{X}_i\beta+\hat{\theta}(\boldsymbol{T}_i;\beta^*))\}\right] \qquad (14)$$

Persamaan (13) dan (14) diiterasikan secara simultan sampai konvergen.

4. Distribusi Asimptotik Estimator Komponen Parametrik

Pada bagian ini akan dibahas distribusi asimptotik dari estimator $\hat{\boldsymbol{\beta}}$ dengan membuktikan teorema berikut:

Teorema 1.

$$\sqrt{n}\{\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\}\rightarrow \boldsymbol{N}(\boldsymbol{0},\boldsymbol{V}_\beta)$$

dimana

$\boldsymbol{V}_\beta=\boldsymbol{D}_n^{-1}(\boldsymbol{Z}_{1i}^T-\boldsymbol{Z}_{2i}^T)\boldsymbol{\Sigma}_i(\boldsymbol{Z}_{1i}^T-\boldsymbol{Z}_{2i}^T)^T\boldsymbol{D}_n^{-1}$,

dengan $\boldsymbol{\Sigma}_i=\mathrm{cov}(\boldsymbol{Y}_i\,|\,\boldsymbol{X}_i,\boldsymbol{T}_i)$.

Sebelum membuktikan teorema tersebut di atas, akan diberikan beberapa lemma yang mendukung pembuktian tersebut.

Lemma 1.

$$\sum_{i=1}^{n}\boldsymbol{D}_i\boldsymbol{V}_{2i}^{-1}\boldsymbol{K}_{ih}(t)\{\boldsymbol{Y}_i-g(\boldsymbol{X}_i\boldsymbol{\beta}+\hat{\boldsymbol{\theta}}(t;\boldsymbol{\beta}))\}=\mathbf{0} \qquad (15)$$

dapat dinyatakan sebagai:

$$\hat{\theta}(t;\beta)-\theta(t)=(\boldsymbol{N}_n)^{-1}\boldsymbol{M}_n+\mathrm{o}_p(1) \qquad (16)$$

dimana:

$$\mathbf{N}_{\mathrm{n}} = \sum_{i=1}^{n} \boldsymbol{D}_i \boldsymbol{V}_{2i}^{-1} \boldsymbol{K}_{ih}(t) \boldsymbol{D}_i \;\text{ dan }\; \mathbf{M}_{\mathrm{n}} = \sum_{i=1}^{n} \boldsymbol{D}_i \boldsymbol{V}_{2i}^{-1} \boldsymbol{K}_{ih}(t) \{\boldsymbol{Y}_i - g(\boldsymbol{X}_i \boldsymbol{\beta} + \hat{\boldsymbol{\theta}}(t;\boldsymbol{\beta})\}$$

Bukti:
Dengan ekspansi deret Taylor terhadap (15) diperoleh:

$$\sum_{i=1}^{n} \boldsymbol{D}_i \boldsymbol{V}_{2i}^{-1} \boldsymbol{K}_{ih}(t) \{\boldsymbol{Y}_i - g(\boldsymbol{X}_j \boldsymbol{\beta} + \hat{\boldsymbol{\theta}}(t;\boldsymbol{\beta})\} = \sum_{i=1}^{n} \boldsymbol{D}_i \boldsymbol{V}_{2i}^{-1} \boldsymbol{K}_{ih}(t) \{\boldsymbol{Y}_i - g(\boldsymbol{X}_i \boldsymbol{\beta} + \hat{\boldsymbol{\theta}}(t;\boldsymbol{\beta})\} \Big|_{\hat{\boldsymbol{\theta}}(t;\boldsymbol{\beta}) = \boldsymbol{\theta}(t)}$$

$$+ \left(\frac{\partial}{\partial \hat{\boldsymbol{\theta}}(t)} \sum_{i=1}^{n} \boldsymbol{D}_i \boldsymbol{V}_{2i}^{-1} \boldsymbol{K}_{ih}(t) \{\boldsymbol{Y}_i - g(\boldsymbol{X}_i \boldsymbol{\beta} + \hat{\boldsymbol{\theta}}(t;\boldsymbol{\beta})\} \Big|_{\hat{\boldsymbol{\theta}}(t;\boldsymbol{\beta}) = \boldsymbol{\theta}(t)} \right) \left(\hat{\boldsymbol{\theta}}(t;\boldsymbol{\beta}) - \boldsymbol{\theta}(t) \right) + \mathrm{o}_p(1)$$

$$= \boldsymbol{M}_n - \left(\sum_{i=1}^{n} \boldsymbol{D}_i \boldsymbol{V}_{2i}^{-1} \boldsymbol{K}_{ih}(t) \boldsymbol{D}_i \right) \left(\hat{\boldsymbol{\theta}}(t;\boldsymbol{\beta}) - \boldsymbol{\theta}(t) \right) + \mathrm{o}_p(1) = \mathbf{0}$$

Jadi

$$\hat{\boldsymbol{\theta}}(t;\boldsymbol{\beta}) - \boldsymbol{\theta}(t) = (\boldsymbol{N}_n)^{-1} \boldsymbol{M}_n + \mathrm{o}_p(1) \quad \blacksquare$$

<u>Lemma 2.</u>

$$\sum_{i=1}^{n} \left[\frac{\partial g(\boldsymbol{X}_i \boldsymbol{\beta} + \hat{\boldsymbol{\theta}}(\boldsymbol{T}_i;\boldsymbol{\beta}))}{\partial \boldsymbol{\beta}} \right]^T \boldsymbol{V}_{1i}^{-1} \{\boldsymbol{Y}_i - g(\boldsymbol{X}_i \boldsymbol{\beta} + \hat{\boldsymbol{\theta}}(\boldsymbol{T}_i;\boldsymbol{\beta}))\} = \mathbf{0} \tag{17}$$

dapat dinyatakan sebagai:

$$\sqrt{n}\{\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\} = \boldsymbol{D}_n^{-1}\{\sqrt{n}\boldsymbol{C}_n\} + o_p(n^{-1/2}),$$

dimana

$$\boldsymbol{D}_n = \frac{1}{n} \sum_{i=1}^{n} \tilde{\boldsymbol{X}}_i \boldsymbol{D}_i \boldsymbol{V}_{1i}^{-1} \boldsymbol{D}_i \tilde{\boldsymbol{X}}_i = E[\tilde{\boldsymbol{X}} \boldsymbol{D} \boldsymbol{V}_1^{-1} \boldsymbol{D} \tilde{\boldsymbol{X}}]$$

$$\boldsymbol{C}_n = \frac{1}{n} \sum_{i=1}^{n} \tilde{\boldsymbol{X}}_i \boldsymbol{D}_i \boldsymbol{V}_{1i}^{-1} \{\boldsymbol{Y}_i - g(\boldsymbol{X}_i \boldsymbol{\beta} + \hat{\boldsymbol{\theta}}(\boldsymbol{T}_i;\boldsymbol{\beta}))\} \text{ dan}$$

$$\tilde{\boldsymbol{X}}_i = \boldsymbol{X}_i + \left(\frac{\partial \hat{\boldsymbol{\theta}}(\boldsymbol{T}_i;\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \right)$$

Bukti:
Persamaan (17) dapat ditulis sebagai:

$$\sum_{i=1}^{n} \tilde{\boldsymbol{X}}_i^T \boldsymbol{D}_i \boldsymbol{V}_{1i}^{-1} \{\boldsymbol{Y}_i - g(\boldsymbol{X}_i \boldsymbol{\beta} + \hat{\boldsymbol{\theta}}(\boldsymbol{T}_i;\boldsymbol{\beta}))\} = \mathbf{0} \tag{18}$$

Dengan ekspansi deret Taylor pada (18), maka didapat:

$$\sum_{\mathrm{i}=1}^{\mathrm{n}} \tilde{\mathbf{X}}_{\mathrm{i}}^{\mathrm{T}} \mathbf{D}_{\mathrm{i}} \mathbf{V}_{1\mathrm{i}}^{-1} \{\mathbf{Y}_{\mathrm{i}} - \mathrm{g}(\mathbf{X}_{\mathrm{i}} \boldsymbol{\beta} + \hat{\boldsymbol{\theta}}(\mathbf{T}_{\mathrm{i}};\boldsymbol{\beta}))\} = \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \tilde{\mathbf{X}}_{\mathrm{i}}^{\mathrm{T}} \mathbf{D}_{\mathrm{i}} \mathbf{V}_{1\mathrm{i}}^{-1} \{\mathbf{Y}_{\mathrm{i}} - \mathrm{g}(\mathbf{X}_{\mathrm{i}} \boldsymbol{\beta} + \hat{\boldsymbol{\theta}}(\mathbf{T}_{\mathrm{i}};\boldsymbol{\beta}))\}_{\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}}$$

$$+ \left(\frac{\partial}{\partial \boldsymbol{\beta}} \frac{1}{n} \sum_{i=1}^{n} \tilde{\boldsymbol{X}}_i^T \boldsymbol{D}_i \boldsymbol{V}_{1i}^{-1} \{\boldsymbol{Y}_i - g(\boldsymbol{X}_i \boldsymbol{\beta} + \hat{\boldsymbol{\theta}}(\boldsymbol{T}_i;\boldsymbol{\beta}))\} \Big|_{\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}} \right) (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) + \mathrm{o}_p(1/n) = \mathbf{0}$$

$$= \boldsymbol{C}_n - \left(\tfrac{1}{n}\sum_{i=1}^{n} \tilde{\boldsymbol{X}}_i^T \boldsymbol{D}_i \boldsymbol{V}_{1i}^{-1} \boldsymbol{D}_i \tilde{\boldsymbol{X}}_i \right)(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) + o_p(1/n) = \mathbf{0}$$

sehingga

$$\sqrt{n}\{\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\} = \boldsymbol{D}_n^{-1}\{\sqrt{n}\boldsymbol{C}_n\} + o_p(n^{-1/2}) \quad \blacksquare$$

Lemma 3.

$$\boldsymbol{C}_n = \tfrac{1}{n}\sum_{i=1}^{n} \tilde{\boldsymbol{X}}_i^T \boldsymbol{D}_i \boldsymbol{V}_{1i}^{-1}\{\boldsymbol{Y}_i - g(\boldsymbol{X}_i\boldsymbol{\beta} + \hat{\boldsymbol{\theta}}(\boldsymbol{T}_i;\boldsymbol{\beta}))\}$$

dapat dinyatakan sebagai: $\boldsymbol{C}_n = \boldsymbol{C}_{1n} - \boldsymbol{C}_{2n} + o_p(1)$,
dimana:

$$\boldsymbol{C}_{1n} = \tfrac{1}{n}\sum_{i=1}^{n} \tilde{\boldsymbol{X}}_i^T \boldsymbol{D}_i \boldsymbol{V}_{1i}^{-1}\{\boldsymbol{Y}_i - \boldsymbol{\mu}_i) \; dan \;\; \boldsymbol{C}_{2n} = \tfrac{1}{n}\sum_{i=1}^{n} \tilde{\boldsymbol{X}}_i^T \boldsymbol{D}_i \boldsymbol{V}_{1i}^{-1}\boldsymbol{D}_i\{\hat{\boldsymbol{\theta}}(\boldsymbol{T}_i;\boldsymbol{\beta}) - \boldsymbol{\theta}(\boldsymbol{T}_i)\}$$

Bukti:
Dengan ekspansi deret Taylor pada $\boldsymbol{C}_n$, didapatkan:

$$\begin{aligned}
\boldsymbol{C}_n &= \tfrac{1}{n}\sum_{i=1}^{n} \tilde{\boldsymbol{X}}_i^T \boldsymbol{D}_i \boldsymbol{V}_{1i}^{-1}\{\boldsymbol{Y}_i - \boldsymbol{\mu}_i\}\Big|_{\hat{\boldsymbol{\theta}}(t;\boldsymbol{\beta})=\boldsymbol{\theta}(t)} \\
&\quad + \left(\tfrac{1}{n}\frac{\partial}{\partial\hat{\boldsymbol{\theta}}(t)} \sum_{i=1}^{n} \tilde{\boldsymbol{X}}_i^T \boldsymbol{D}_i \boldsymbol{V}_{1i}^{-1}\{\boldsymbol{Y}_i - \boldsymbol{\mu}_i\}\Big|_{\hat{\boldsymbol{\theta}}(t;\boldsymbol{\beta})=\boldsymbol{\theta}(t)} \right)(\hat{\boldsymbol{\theta}}(\boldsymbol{T}_i;\boldsymbol{\beta}) - \boldsymbol{\theta}(\boldsymbol{T}_i)) + o_p(1) \\
&= \boldsymbol{C}_{1n} - \left(\tfrac{1}{n}\sum_{i=1}^{n} \tilde{\boldsymbol{X}}_i^T \boldsymbol{D}_i \boldsymbol{V}_{1i}^{-1} \boldsymbol{D}_i \right)(\hat{\boldsymbol{\theta}}(\boldsymbol{T}_i;\boldsymbol{\beta}) - \boldsymbol{\theta}(\boldsymbol{T}_i)) + o_p(1) \\
&= \boldsymbol{C}_{1n} - \boldsymbol{C}_{2n} + o_p(1) \quad \blacksquare
\end{aligned}$$

Selanjutnya akan dibuktikan **Teorema 1** yang ekivalen dengan mencari distribusi asimptotik $\boldsymbol{D}_n^{-1}\{\sqrt{n}\boldsymbol{C}_n\}$.

Diketahui: $C_{1n} = \tfrac{1}{n}\sum_{i=1}^{n} \tilde{\boldsymbol{X}}_i^T \boldsymbol{D}_i \boldsymbol{V}_{1i}^{-1}\{\boldsymbol{Y}_i - \boldsymbol{\mu}_i) = \tfrac{1}{n}\sum_{i=1}^{n} \boldsymbol{Z}_{1i}^T(\boldsymbol{Y}_i - \boldsymbol{\mu}_i)$

Baris ke-r dari persamaan (16) dapat dinyatakan sebagai:

$$Q\left(\sum_{i=1}^{n}\left(\sum_{k=1}^{m}\left(\sum_{j=1}^{m} d_{jj}^{i} K_h^i(T_{jj} - t) v_{jk}^{2i} \right)(Y_{ik} - \mu_{ik}) \right) \right) \tag{19}$$

dimana Q adalah baris ke-r dari $(\boldsymbol{N}_n)^{-1}$. Selanjutnya, baris ke-r dari $\boldsymbol{C}_{2n}$ dapat juga dinyatakan sebagai :

$$\tfrac{1}{n}\left(\sum_{i=1}^{n}(\sum_{k=1}^{m}[\sum_{j=1}^{m} \tilde{X}_{ij}(d_{jj}^{i})^2 v_{jk}^{1i}] \right)(\hat{\theta}(T_{ik};\beta) - \theta(T_{ik})) \tag{20}$$

dimana d_{jj}^{i} adalah elemen ke-jj dari matriks $\boldsymbol{D}_i$, $K_h^i(T_{jj} - t)$ adalah elemen ke-jj dari matriks $\boldsymbol{K}_{ih}(t)$ dan v_{jk}^{2i} adalah elemen ke-jk dari matriks $\boldsymbol{V}_{2i}^{-1}$. Dengan menggunakan **Lemma 1**, **Lemma 2,** dan **Lemma 3,** maka diperoleh:

$$\boldsymbol{C}_{2n} = \tfrac{1}{n}\sum_{i=1}^{n} \boldsymbol{Z}_{2i}\left(\boldsymbol{Y}_i - \boldsymbol{\mu}_i\right), \text{ dimana } \boldsymbol{Z}_{2i} = [\boldsymbol{Z}_{2i1}, \ldots, \boldsymbol{Z}_{2im}]^{\mathrm{T}}$$

Kemudian dengan menggunakan **Lemma 2**, diperoleh:

$$\boldsymbol{C}_n = \tfrac{1}{n}\sum_{i=1}^{n} \boldsymbol{Z}_{1i}^T(\boldsymbol{Y}_i - \boldsymbol{\mu}_i) - \tfrac{1}{n}\sum_{i=1}^{n} \boldsymbol{Z}_{2i}^T(\boldsymbol{Y}_i - \boldsymbol{\mu}_i) = \tfrac{1}{n}\sum_{i=1}^{n} (\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)(\boldsymbol{Y}_i - \boldsymbol{\mu}_i)$$

Jadi $\boldsymbol{D}_n^{-1}\{\sqrt{n}\boldsymbol{C}_n\} = \boldsymbol{D}_n^{-1}\{\sqrt{n}\,\tfrac{1}{n}\sum_{i=1}^{n}(\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)(\boldsymbol{Y}_i - \boldsymbol{\mu}_i)\} = \boldsymbol{D}_n^{-1}\{\tfrac{1}{\sqrt{n}}\sum_{i=1}^{n}(\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)(\boldsymbol{Y}_i - \boldsymbol{\mu}_i)\}$

Selanjutnya akan dicari distribusi asimptotik dari $\sqrt{n}\boldsymbol{C}_n$, sebagai berikut:

$$E\left((\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)(\boldsymbol{Y}_i - \boldsymbol{\mu}_i)\right) = \boldsymbol{0}$$

$$\begin{aligned}\mathrm{Var}\left((\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)(\boldsymbol{Y}_i - \boldsymbol{\mu}_i)\right) &= E\left((\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)(\boldsymbol{Y}_i - \boldsymbol{\mu}_i) - E[(\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)(\boldsymbol{Y}_i - \boldsymbol{\mu}_i)]\right)^2 \\ &= (\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)E\left((\boldsymbol{Y}_i - \boldsymbol{\mu}_i)^2\right)(\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)^T \\ &= (\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)\boldsymbol{\Sigma}_i(\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)^T\end{aligned}$$

Dengan menggunakan Teorema Limit Pusat diperoleh:

$$\frac{\sqrt{n}(\boldsymbol{C}_n - \boldsymbol{0})}{\sqrt{Var\left((\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)(\boldsymbol{Y}_i - \boldsymbol{\mu}_i)\right)}} = \frac{\sqrt{n}\boldsymbol{C}_n}{\sqrt{Var\left((\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)(\boldsymbol{Y}_i - \boldsymbol{\mu}_i)\right)}} \xrightarrow{d} \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{I})$$

jadi $\sqrt{n}\boldsymbol{C}_n \xrightarrow{d} \boldsymbol{N}\left(\boldsymbol{0}, Var(\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)(\boldsymbol{Y}_i - \boldsymbol{\mu}_i)\right)$.

Sehingga didapatkan distribusi asimptotik dari $\boldsymbol{D}_n^{-1}\{\sqrt{n}\boldsymbol{C}_n\}$, yaitu:

$$\boldsymbol{N}\left(\boldsymbol{0}, \boldsymbol{D}_n^{-1}(\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)\boldsymbol{\Sigma}_i(\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)^T\boldsymbol{D}_n^{-1}\right).$$

Akibatnya, untuk $n \to \infty$ maka:

$$\sqrt{n}\{\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\} \xrightarrow{d} \boldsymbol{N}\left(\boldsymbol{0}, \boldsymbol{D}_n^{-1}(\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)\boldsymbol{\Sigma}_i(\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)^T\boldsymbol{D}_n^{-1}\right). \quad \blacksquare$$

4. Studi Simulasi

Studi simulasi dilakukan untuk memperlihatkan bahwa estimator yang diturunkan dari teori adalah valid. Model yang dipilih adalah: $Y_{ij} = \beta_1 X_{1ij} + \beta_2 X_{2ij} + \theta(T_{ij}) + \varepsilon_{ij}$ dimana $\theta(T_{ij})$ dianggap bagian nonparametrik. X_{1ij}, T_{ij} dan ε_{ij} dibangkitkan masing-masing dengan $X_{1ij} \sim U(-2, 2)$ dan $T_{ij} \sim U(-2, 2)$ dan $\varepsilon_{ij} \sim N(0, 1)$. Bagian nonparametrik $\theta(T_{ij})$ dipilih tiga bentuk fungsi, yaitu: trigonometri, eksponensial dan polynomial. β_1 dan β_2 masing-masing diberikan nilai tertentu (β_1=1, β_2=3) dan tiap subjek dipilih 3 pengamatan (m=3). Simulasi ini dilakukan dengan menggunakan software MATLAB. Hasil simulasi diberikan berikut.

Tabel 1. Hasil estimasi parameter model

$Y_{ij} = \beta_1 X_{1ij} + \beta_2 X_{2ij} + \sin(3T_{ij}^2)/T_{ij} + \varepsilon_{ij}$, $\beta_1 = 1$, $\beta_2 = 3$

N	**mean** $\hat{\beta}_1$	**mean** $\hat{\beta}_2$
50	0.9926	3.0246
100	0.9832	2.9879
200	1.0048	2.9924
400	1.0014	3.0128

Tabel 2. Hasil estimasi parameter model

$Y_{ij} = \beta_1 X_{1ij} + \beta_2 X_{2i} + \exp(T_{ij}) + \varepsilon_{ij}$, $\beta_1 = 1$, $\beta_2 = 3$

N	**mean** $\hat{\beta}_1$	**mean** $\hat{\beta}_2$
50	1.0042	3.0343
100	0.9959	3.0501
200	0.9957	3.0558
400	1.0069	3.0526

Tabel 3. Hasil estimasi parameter model

$Y_{ij} = \beta_1 X_{1ij} + \beta_2 X_{2i} + T_{ij}^2 + \varepsilon_{ij}$, $\beta_1 = 1$, $\beta_2 = 3$

N	mean $\hat{\beta}_1$	mean $\hat{\beta}_2$
50	1.0027	2.9238
100	1.0238	3.0168
200	0.9973	3.0611
400	1.0032	2.8915

Dari Tabel 1 sampai Tabel 3, terlihat bahwa nilai estimasi $\hat{\beta}_1$ dan $\hat{\beta}_2$ pada berbagai nilai *n* dan bentuk fungsi regresi yang berbeda adalah cenderung hampir sama dengan nilai parameter sebenarnya (β_1 dan β_2). Ini memperlihatkan bahwa estimator $\hat{\boldsymbol{\beta}}$ yang diperoleh dari teori adalah valid.

5. Kesimpulan

Ada beberapa kesimpulan yang didapatkan dari kajian ini, yaitu:

1. Jika diberikan model: $\boldsymbol{Y}_{ij} = \boldsymbol{X}_{ij}\boldsymbol{\beta} + \boldsymbol{\theta}(\boldsymbol{T}_{ij}) + \boldsymbol{\varepsilon}_{ij}$, maka parameter komponen parametrik dan bentuk nonparametriknya dapat diestimasi dengan metode pendekatan GEE dengan menyelesaikan persamaan berikut:

$$\hat{\boldsymbol{\alpha}}^{(v)} = \hat{\boldsymbol{\alpha}}^{(v-1)} - \left[\sum_{i=1}^{n}[\boldsymbol{G}_i^T(t)\boldsymbol{W}_i(t)\boldsymbol{G}_i(t)]\right]^{-1} \sum_{\mathrm{i=1}}^{\mathrm{n}} \mathbf{G}_{\mathrm{i}}^{\mathrm{T}}(\mathrm{t})\mathbf{D}_{\mathrm{i}}\mathbf{V}_{2\mathrm{i}}^{-1}\mathbf{K}_{\mathrm{ih}}(\mathrm{t})\{\mathbf{Y}_{\mathrm{i}} - \boldsymbol{\mu}_{\mathrm{i}}\}$$

$$\hat{\boldsymbol{\beta}}^{(v)} = \hat{\boldsymbol{\beta}}^{(v-1)} - \left[\sum_{i=1}^{n}\tilde{\boldsymbol{X}}_i^T\boldsymbol{D}_i[\mathbf{V}_{1i}^{-1}]^T\boldsymbol{D}_i\tilde{\boldsymbol{X}}_i\right]^{-1}$$

   Kedua persamaan di atas diselesaikan secara simultan dan iteratif sampai konvergen.

2. Estimator $\hat{\boldsymbol{\beta}}$ berdistribusi normal asymptotik, yaitu:

$$\sqrt{n}\{\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\} \to \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{V}_\beta) \text{ untuk } n \to \infty,$$

dimana $\boldsymbol{V}_\beta = \boldsymbol{D}_n^{-1}(\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)\boldsymbol{\Sigma}_i(\boldsymbol{Z}_{1i}^T - \boldsymbol{Z}_{2i}^T)^T \boldsymbol{D}_n^{-1}$

3. Dari studi simulasi terlihat bahwa model-model regresi yang diberikan, nilai $\hat{\beta}_1$ dan $\hat{\beta}_2$ yang diperoleh dengan menggunakan estimasi yang didapatkan dari teori adalah cenderung mendekati nilai β_1 dan β_2 yang sebenarnya. Ini menunjukkan bahwa estimator yang didapatkan adalah cukup layak digunakan untuk mengestimasi.

Daftar Pustaka

1. Abdy, M., 2007, *Suatu Tinjauan Tentang Generalized Estimating Equation, Jurnal Matematika Statistika dan Komputasi*, Jurusan Matematika UNHAS, 4, 21 – 27.

2. Engle, R., Granger, C. W. J., Rice, J., and Weiss, A., 1986, *Semiparametric Estimation of the Relation Between Weather and Electricity Sales*, *Journal of the American Statistical Association, 81*, 310 – 320.

3. Green, P. J., 1986, *Penalized Likelihood for General Semiparametric Regression Models*, *International Statistical Review, 55*, 245 – 259.

4. Hastie, T. J., and Tibshirani, R. J., 1990, *Generalized Additive Models*, Chapman & Hall, London.

5. Liang, K. Y., and Zeger, S. L., 1986, *Longitudinal Data Analysis Using Generalized Linear Model*, *Biometrika, 73*, 13 – 22.

6. Lin, X., and Carroll, R. J., 2000, *Nonparametric Function Estimation for Clustered Data When the Predictor is Measured Without/With Error*, *Journal of the American Statistical Association, 95*, 520 – 534.

7. ______, 2001, *Semiparametric Regression for Clustered Data*, *Biometrika*, 88, 1179-1185.

8. McCullagh, P., 1983, *Quasi-likelihood Functions*, *The Annals of Statistics*, 11, 59-67.

9. McCullagh, P., and Nelder, J.A., 1989, *Generalized Linear Model* (2nd ed.), Chapman & Hall. London.

10. Severini, T. A., and Staniswalis, J. G., 1994, *Quasi-likelihood Estimation in Model Semiparametric*, *Journal of the American Statistical Association, 89*, 501- 511.

11. Wahba, G., 198), *Partial Spline Models for the Semiparametric Estimation of several Variable*, *Statistics Analysis of Time Series*, *Proceeding of the Japan US, Joint Seminar, Tokyo*, pp. 319 -329.

12. Wedderburn, R.,W. M., 1974, *Quasi-likelihood Functions*, *Generalized Linear Model*, *and the Gauss-Newton Method*, *Biometrika*, 61, 439 – 447.