Statistika, Vol. 22 No. 1, 41 – 54
Mei 2022

Pemodelan Kurva Engel Non Makanan dengan Bayesian Quantile Regression di Provinsi Papua

MUHAMMAD FAJAR[1,2], SETIAWAN[1]

[1]Institut Teknologi Sepuluh Nopember, [2]Badan Pusat Statistik,
e-mail: setiawan@statistika.its.ac.id

ABSTRAK

Tujuan paper ini adalah untuk melakukan pemodelan kurva Engel non makanan, dimana pada sebaran data mengindikasikan terjadinya heterokedastisitas. Metode yang digunakan dalam penelitian ini adalah regresi kuantil dengan pendekatan Bayesian. Sumber data yang digunakan adalah data pengeluaran konsumsi rumah tangga dan pengeluaran konsumsi rumah tangga untuk non makanan di Provinsi Papua, yang berasal dari SUSENAS Maret 2018. SUSENAS adalah salah satu survei yang dilaksanakan Badan Pusat Statistik. Rumah tangga sampel SUSENAS Maret 2018 sebanyak 10629 rumah tangga. Hasil dari penelitian ini adalah regresi kuantil Bayesian jauh lebih representatif karena mencakup sebaran data, baik pada sentral data maupun daerah disekitar sentral data, dibandingkan regresi linear biasa, dimana regresi linear biasa hanya mencakup berkisar pada sentral data. Ini berimplikasi bahwa pemodelan kurva Engel dengan regresi kuantil adalah sangat cocok dengan sebaran data. Pengeluaran konsumsi rumah tangga berpengaruh positif dan signifikan pada pengeluran konsumsi rumah tangga untuk non makanan. Proporsi pengeluaran konsumsi rumah tangga untuk non makanan berbeda-beda untuk setiap tingkat pengeluarannya. Semakin tinggi pengeluaran rumah tangga, maka proporsi konsumsi untuk non makanan, semakin besar proporsinya. Dan sebaliknya, rumah tangga dengan level pengeluaran rendah, Proporsi konsumsi untuk makan dan minum lebih besar daripada konsumsi untuk non makanan

Kata Kunci: Engel, Non Makanan, Bayesian, Kuantil, Papua.

ABSTRACT

This paper aims to model a non-food Engel curve, where the distribution of data shows heteroscedasticity. The method used in this research is quantile regression with the Bayesian approach. The data sources used are household consumption expenditure data and household non-food consumption expenditures in Papua Province which are derived from SUSENAS March 2018. SUSENAS is one of the surveys conducted by the Badan Pusat Statistik-Statistics Indonesia. SUSENAS March 2018 sample households were 10629 households. The result of this study is that Bayesian quantile regression is much more representative covering the distribution of data, both in the data center and around it, compared to ordinary linear regression, where ordinary linear regression only has a wide range because of the central data. This implies that the modeling of the Engel curve with quantile regression is very suitable for the distribution of the data. Household consumption expenditure has a positive and significant effect on non-food household consumption expenditure. The proportion of household consumption expenditure on food varies for each level of competition. The higher the household expenditure, the higher the proportion of non-food consumption. And conversely, in households with low levels of expenditure, the proportion of consumption for food is greater than consumption for non-food.

Keywords: Engel, Non-food, Bayesian, Quantile, Papua

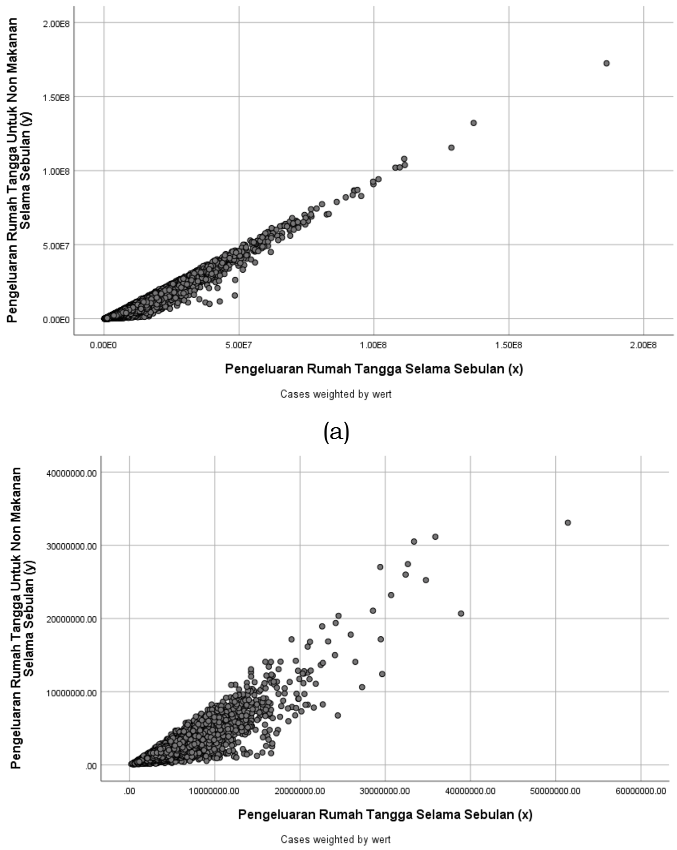
1. PENDAHULUAN

Konsumsi rumah tangga merupakan komponen terbesar dalam PDB Indonesia, dibandingkan komponen lainnya. Berdasarkan data BPS, kontribusi konsumsi rumah tangga terhadap PDB selama periode 2014 – 2020 berkisar 55 sampai 58 persen. Ini berarti bahwa perekonomian masih ditopang oleh magnitudo konsumsi rumah tangga (BPS, 2021). Dalam konsumsi rumah tangga secara garis besar terbagi ke dalam dua kelompok komoditas, konsumsi untuk makanan (termasuk minuman) dan konsumsi untuk non makanan. Berdasarkan hasil susenas selama periode 2018 – 2021 (per bulan Maret) proporsi konsumsi perkapita sebulan untuk makanan dan non makanan terhadap konsumsi total masing-masing berkisar 49.5- 49.1 persen dan 50.5 – 50.9 persen. Ini berarti proporsi konsumsi untuk non makanan lebih besar dibandingkan konsumsi untuk makanan. Menurut Hukum Engel menyatakan bahwa saat pendapatan meningkat, proporsi pendapatan yang dihabiskan untuk konsumsi makanan berkurang (Timmer et al., 1983; Lewbel, 2007; Chai dan Moneta, 2010). Ini berarti mengindikasikan terjadinya peningkatan pendapatan masyarakat dalam situasi perekonomian normal dan semakin meningkatnya proporsi konsumsi non makanan pada masyarakat.

Leser (1963) mengemukakan terdapat lima spesfikasi bentuk kurva Engel. Namun, penulis merujuk kepada salah satu bentuk spesifikasi kurva Engel yang mengakomodir bentuk variabel aslinya tanpa transformasi, yaitu:

$$y = \beta_0 + \beta_1 x + \epsilon. \qquad (1)$$

Dalam paper ini penulis memfokuskan pada pengeluaran konsumsi non makanan karena berkaitan dengan fenomena meningkatnya proporsi non makanan secara nasional, sehingga y adalah pengeluaran konsumsi untuk non makanan rumah tangga, x adalah pengeluaran konsumsi rumah tangga sebagai proxy variabel pendapatan rumah tangga karena data pendapatan rumah tangga sulit diperoleh, kalaupun tersedia terdapat bias cukup besar pada data pendapatan rumah tangga, ϵ adalah eror.



(a)

(b)

Sumber: BPS, diolah.

Gambar 1 Scatter Plot Antara Pengeluaran Konsumsi dan Pengeluaran Untuk Konsumsi Makanan dan Minum Selama 1 bulan pada Rumah Tangga Tahun 2018: (a). Indonesia, (b). Provinsi Papua

Penulis melakukan investigasi terhadap konsumsi rumah tangga di Provinsi Papua, karena scatter plot antara pengeluaran konsumsi dan pengeluaran untuk konsumsi makanan dan minum selama 1 bulan pada rumah tangga di Provinsi Papua jauh lebih melebar ketika level pengeluaran konsumsi semakin meningkat dibandingkan kondisi Nasional (lihat gambar 1). Implikasinya adalah bahwa pemodelan kurva Engel non makanan pada provinsi Papua tidak dapat dilakukan dengan regresi liniar biasa (mean regression) karena sebaran data yang melebar dan indikasi terjadinya heteroscedastisitas.

Oleh karena itu, penulis menggunakan pendekatan regresi kuantil yang diperkenalkan oleh Koenker dan Bassett (1978), untuk spesifikasi dan interpretasi umum untuk pemodelan kedua variabel tersebut dengan pertimbangan bahwa regresi kuantil menghasilkan beberapa garis yang merepresentasikan sebaran data dalam kondisi demikian, sehingga informasi yang dicakup bukan hanya pada sentral data namun kuantil-kuantil disekitar sentral data. Kemudian, inferensi parameter dalam regresi kuantil menggunakan pendekatan Bayesian (Yu dan Moyeed, 2001; Tsionas, 2003; Kozumi dan Kobayashi, 2011). Ini berarti bahwa parameter dalam spesifikasi regresi kuantil bukanlah konstanta semata melainkan variabel random dari distribusi posterior, sehingga mean dan variansi dari distribusi posterior tersebut merupakan nilai estimasi parameter. Dengan demikian, pemodelan kurva Engel non makanan dapat menyerap semua informasi bukan hanya dari data dan berpusat pada sentral data, namun juga menyerap informasi parameter sehingga merefleksikan fenomena riil yang terjadi.

2. METODE PENELITIAN

Sumber Data

Data yang digunakan dalam penelitian ini adalah data total pengeluaran dan pengeluaran non makanan rumah tangga selama sebulan bersumber dari SUSENAS Maret 2018 yang dilaksanakan BPS. Objek yang diteliti adalah sampel rumah tangga di wilayah Provinsi Papua, dengan jumlah rumah tangga sampel sebanyak 10629 observasi. Data pengeluaran rumah tangga sebagai proxy pendapatan rumah tangga dengan pertimbangan yang telah dijelaskan sebelumnya.

Metode Analisis

1. Quantile Bayesian Regression

Regresi kuantil merupakan pengembangan dari mean regression (ordinary regression) yang sering digunakan ketika terdapat hubungan komprehensif antara variabel respon (y) dan beberapa variabel prediktor $(\boldsymbol{x})$. Terdapat model linier berikut:

$$y_i = \boldsymbol{x}_i^T \boldsymbol{\beta}(\tau) + u_i, \quad i = 1,2,\dots,n \tag{1}$$

dengan y_i adalah sebuah variabel respon, $\boldsymbol{x}_i$ adalah vektor berukuran (p+1) x 1 yang terdiri dari p variabel predictor dan konstanta **1**, $\boldsymbol{\beta}(\tau) = \left(\beta_1(\tau)\ \beta_2(\tau) \dots \beta_{p+1}(\tau)\right)^T$ adalah vektor koefisien kuantil ke τ yang berukuran (p+1) x 1, dan $u_i$adalah eror acak. Diberikan kuantil $\tau \in (0,1)$, kemudian regresi kuantil adalah:

$$Q_\tau(y|\boldsymbol{x}) = \boldsymbol{x}_i^T \boldsymbol{\beta}(\tau) \tag{2}$$

dengan $Q_\tau(y|\boldsymbol{x})$ adalah *conditional quantile y* diberikan $\boldsymbol{x}_i$. Untuk mengestimasi $\boldsymbol{\beta}(\tau)$ dengan meminimumkan:

$$\underset{\boldsymbol{\beta}(\tau)}{min} \sum_{i=1}^{n} \rho_\tau(y_i - \boldsymbol{x}_i^T \boldsymbol{\beta}(\tau)) \tag{3}$$

dengan $\rho_\tau(u) = u\tau$ jika $u \geq 0$, dan $\rho_\tau(u) = u(\tau - 1)$ jika $u < 0$, atau dirumuskan dalam bentuk:

$$\rho_\tau(u) = u\left(\tau - I(u < 0)\right) \tag{4}$$

dengan $\rho_\tau(.)$ adalah *loss function*, $I(.)$ adalah fungsi indikator. *Loss function* tidak *differentiable* pada nol, maka tidak mendapatkan solusi turunan eksplisit dalam proses meminimumkan persamaan (3). Oleh karena itu, metode *linear programming* (Koenker dan Basset, 1978) biasa digunakan untuk mengestimasi model regresi kuantil (pendekatan frekuentis).

Yu dan Moyeed (2001) mengusulkan pendekatan *Bayesian* dalam proses meminimumkan persamaan (3), dimana proses tersebut ekivalen dengan memaksimalkan fungsi likelihood yang berdasarkan distribusi Laplace Asimetrik $(\mu = 0{,}1, \tau)$ pada eror yang dihasilkan regresi kuantil:

$$f_\tau(u) = \tau(1-\tau)e^{-\rho_\tau(u)} \quad (5)$$

parameter τ menentukan kemiringan distribusi dan kuantil ke τ distribusi ini adalah nol. Mean dan varians distribusi masing – masing adalah:

$$E(u) = \frac{1-2\tau}{\tau(1-\tau)} \quad \text{dan} \quad \mathrm{Var}(u) = \frac{1-2\tau+\tau^2}{\tau^2(1-\tau)^2}$$

namun persamaan (4) dapat disisipkan scale parameter σ menjadi:

$$f_\tau(u) = \frac{\tau(1-\tau)}{\sigma} e^{-\frac{\rho_\tau(u)}{\sigma}} \quad (6)$$

Dari persamaan (4) dapat dibentuk fungsi likelihood:

$$L(\boldsymbol{y}|\boldsymbol{\beta}(\tau)) = \frac{\tau^n(1-\tau)^n}{\sigma^n} \exp\left(-\frac{\sum_{i=1}^{n} \rho_\tau(y_i - \boldsymbol{x}_i^T\boldsymbol{\beta}(\tau))}{\sigma}\right) \quad (7)$$

Persamaan (7) merupakan fungsi likelihood y_i yang i.i.d mengikuti distribusi Laplace Asimetrik $(\mu = \boldsymbol{x}_i^T\boldsymbol{\beta}(\tau), \sigma, \tau)$, μ adalaha location parameter, σ adalah scale parameter, τ adalah skewness parameter.

Kozumi dan Kobayashi (2011) menunjukkan bahwa distribusi Laplace Asimetrik merupakan representasi gabungan, berikut:

$$u = \theta z + \varphi\sqrt{z}\omega \quad (8)$$

dengan:

$$\theta = \frac{1-2\tau}{\tau(1-\tau)} \qquad \text{dan} \qquad \varphi^2 = \frac{2}{\tau(1-\tau)}$$

z adalah variabel random eksponensial standar (Eksponensial (mean = 1)), ω adalah variabel random normal standar ($N(0,1)$), u adalah variabel random distribusi Laplace Asimetrik.

Masukan persamaan (8) ke persamaan (1) menjadi:

$$y_i = \boldsymbol{x}_i^T\boldsymbol{\beta}(\tau) + \theta z_i + \varphi\sqrt{z_i}\omega_i \quad (9)$$

dengan z_i dan ω_i saling independen, maka distribusi bersyarat y_i diberikan z_i adalah mengikuti $N(\boldsymbol{x}_i^T\boldsymbol{\beta}(\tau) + \theta z_i, \varphi^2 z_i)$. Jadi distribusi $\boldsymbol{y} = (y_1, \dots, y_n)^T$ adalah:

$$f(\boldsymbol{y}|\boldsymbol{\beta}(\tau), \boldsymbol{z}) \propto \left(\prod_{i=1}^{n} z_i^{-\frac{1}{2}}\right) \exp\left(-\sum_{i=1}^{n} \frac{(y_i - \boldsymbol{x}_i^T\boldsymbol{\beta}(\tau) - \theta z_i)^2}{2\varphi^2 z_i}\right)$$

dengan $\boldsymbol{z} = (v_1, \dots, v_n)^T$.

2. Algoritma Gibbs pada Estimasi Quantile Bayesian Regression

Dari persamaan (9) dapat ditulis ulang menjadi:

$$y_i = \boldsymbol{x}_i^T\boldsymbol{\beta}(\tau) + \sigma\theta z_i + \sigma\varphi\sqrt{z_i}\omega_i \quad (10)$$

Pada persamaan (10) sulit dibuat algoritma *Gibbs Sampling* yang menghadirkan keberadaan scale parameter dalam rata-rata bersyarat y_i. Oleh karena itu, persamaan (10) dilakukan reparameterisasi menjadi (Kozumi dan Kobayashi, 2011):

$$y_i = \boldsymbol{x}_i^T \boldsymbol{\beta}(\tau) + \theta v_i + \varphi\sqrt{\sigma v_i}\omega_i \tag{11}$$

dengan $v_i = \sigma z_i$. Untuk melengkapi spesifikasi model, asumsikan bahwa $\boldsymbol{\beta}(\tau) \sim N(\boldsymbol{\beta}_0(\tau), \boldsymbol{B}_0)$ dan $\sigma \sim IG({}^{n_0}/_2, {}^{s_0}/_2)$, $IG(a,b)$ menyatakan distribusi inverse-Gamma dengan parameter a dan b. Sekarang dibutuhkan sampel $\boldsymbol{\beta}(\tau), \boldsymbol{v} = (v_1, \dots, v_n)^T$ dan σ dari masing-masing distribusi bersyarat mereka. Distribusi bersyarat $\boldsymbol{\beta}(\tau)$ adalah:

$$\boldsymbol{\beta}(\tau)|\boldsymbol{y}, \boldsymbol{v}, \sigma \sim N(\tilde{\boldsymbol{\beta}}(\tau), \tilde{\boldsymbol{B}}) \tag{12}$$

dengan:

$$\tilde{\boldsymbol{B}}^{-1} = \sum_{i=1}^{n}\left(\frac{\boldsymbol{x}_i \boldsymbol{x}_i^T}{\varphi^2 \sigma v_i}\right) + \boldsymbol{B}_0^{-1} \text{ dan } \tilde{\boldsymbol{\beta}}(\tau) = \tilde{\boldsymbol{B}}\left(\sum_{i=1}^{n}\left(\frac{\boldsymbol{x}_i(y_i - \theta v_i)}{\varphi^2 \sigma v_i} + \boldsymbol{B}_0^{-1}\boldsymbol{\beta}(\tau)_0\right)\right)$$

Kemudian:

$$v_i|\boldsymbol{y}, \boldsymbol{\beta}(\tau), \sigma \sim GIG\left(\frac{1}{2}, \tilde{\delta}_i, \tilde{\gamma}_i\right) \tag{13}$$

$GIG(d,q,r)$ adalah fungsi *generalized inverse* Gaussian yang dirumuskan:

$$f(w|d,q,r) = \frac{(r/q)^d}{2K_d(qr)} w^{d-1} \exp\left(-\frac{1}{2}(q^2 w^{-1} + r^2 w\right), \quad w > 0, -\infty < d < \infty, \qquad q, r \geq 0$$

dengan $K_d(.)$ adalah modifikasi fungsi Bessel jenis ketiga.

$$\tilde{\delta}_i^2 = \frac{(y_i - \boldsymbol{x}_i^T \boldsymbol{\beta}(\tau))^2}{2\varphi^2 \sigma} \quad \text{dan} \quad \tilde{\gamma}_i^2 = \frac{2}{\sigma} + \frac{\theta^2}{2\varphi^2 \sigma}$$

dengan catatan bahwa $v_i \sim$ Eksponensial $(Mean = \sigma)$ sehingga distribusi bersyarat σ proporsional terhadap:

$$\left(\frac{1}{\sigma}\right)^{\left(\frac{n_0}{2}\right) + \left(\frac{3}{2}\right)n + 1} \exp\left(-\frac{1}{\sigma}\left(\frac{s_0}{2} + \sum_{i=1}^{n} v_i + \sum_{i=1}^{n} \frac{\left(y_i - \boldsymbol{x}_i^T \boldsymbol{\beta}(\tau) - \theta v_i\right)^2}{2\varphi^2 v_i}\right)\right)$$

Jadi:

$$\sigma|\boldsymbol{y}, \boldsymbol{\beta}(\tau), \boldsymbol{v} \sim IG\left(\frac{n_0 + 3n}{2}, \frac{\tilde{s}}{2}\right) \tag{14}$$

dengan:

$$\tilde{s} = s_0 + 2\sum_{i=1}^{n} v_i + \sum_{i=1}^{n} \frac{\left(y_i - \boldsymbol{x}_i^T \boldsymbol{\beta}(\tau) - \theta v_i\right)^2}{\varphi^2 v_i}$$

sehingga dapat dikonstruksi distribusi posterior gabungan:

$$\pi(\boldsymbol{\beta}(\tau), \boldsymbol{v}, \sigma|\boldsymbol{y}) \quad \propto \quad p(\boldsymbol{\beta}(\tau)|\boldsymbol{y}, \boldsymbol{v}, \sigma)\left(\prod_{i=1}^{n} p(v_i|\boldsymbol{y}, \boldsymbol{\beta}(\tau), \sigma)\right) p(\sigma|\boldsymbol{y}, \boldsymbol{\beta}(\tau), \boldsymbol{v})$$

Algoritma Gibbs dapat diimplementasikan (Tsionas, 2003; Kozumi dan Kobayashi, 2011; Choi dan J.P. Hobert, 2013):

(1) Tentukan initial value $\Theta_{\tau,(0)} = (\boldsymbol{\beta}(\tau)_{(0)}, \boldsymbol{v}_{(0)}, \quad \sigma_{(0)})$, dengan $\boldsymbol{\beta}(\tau)_{(0)} = \{\beta_1(\tau)_{(0)}, \dots, \beta_p(\tau)_{(0)}\}$,

$\boldsymbol{v}_{(0)} = \{v_{1,(0)}, \dots, v_{n,(0)}\}$, $\boldsymbol{B}_0$, n_0, dan s_0. Set m = 0.

(2) Bangkitkan $\Theta_{\tau,(m)}$ berdasarkan:

- Bangkitkan $\boldsymbol{\beta}(\tau)_{(m+1)}$ dari distribusi $\boldsymbol{\beta}(\tau)|\boldsymbol{y},\boldsymbol{v},\sigma$ pada persamaan (12).
- Bangkitkan $\boldsymbol{v}_{(m+1)}$ dari distribusi $\boldsymbol{v}|\boldsymbol{y},\boldsymbol{\beta}(\tau),\sigma$ pada persamaan (13), untuk $i = 1, \dots, n$.
- Bangkitkan σ dari distribusi $\sigma|\boldsymbol{y},\boldsymbol{\beta}(\tau),\boldsymbol{v}$ pada persaman (14).

(3) Lakukan $m = m + 1$ dan ulangi ke langkah (1).

Dalam penelitian ini initial value berdasarkan default pada package bayesQR dan kuantil yang digunakan dalam penelitian ini adalah kuantil 0.2, 0.4, 0.6, dan 0.8 merujuk pada Tilak (2000) dan penulis menambahkan kuantil ke 0.05 dan 0.96 menyesuaikan kondisi data. Dalam hasil estimasi dalam proses package bayesQR (Benoit dan Van de Poel, 2017) hanya disajikan mean dan varians posterior $\boldsymbol{\beta}(\tau)$ dan σ yang merupakan fokus penelitian ini serta tidak disajikan mean dan varians posterior $\boldsymbol{v}$ karena hanya sebagai variabel artifisial. Iterasi yang diterapkan pada proses algoritma *Gibbs Sampling* sebanyak 1 juta iterasi untuk setiap kuantilnya. Proses satu kali iterasi melibatkan 10629 rumah tangga sampel untuk proses kalkulasi sehingga memerlukan waktu yang lama. Dari proses algoritma *Gibbs Sampling* menghasilkan data $\boldsymbol{\beta}(\tau)$ dan σ masing – masing sebanyak 1 juta unit, kemudian 2500 unit pertama dibuang sehingga 997500 unit sisanya digunakan untuk menghitung mean dan variansi dari distribusi posterior bersyarat $\boldsymbol{\beta}(\tau)$ dan σ.

3. *Credible Interval*

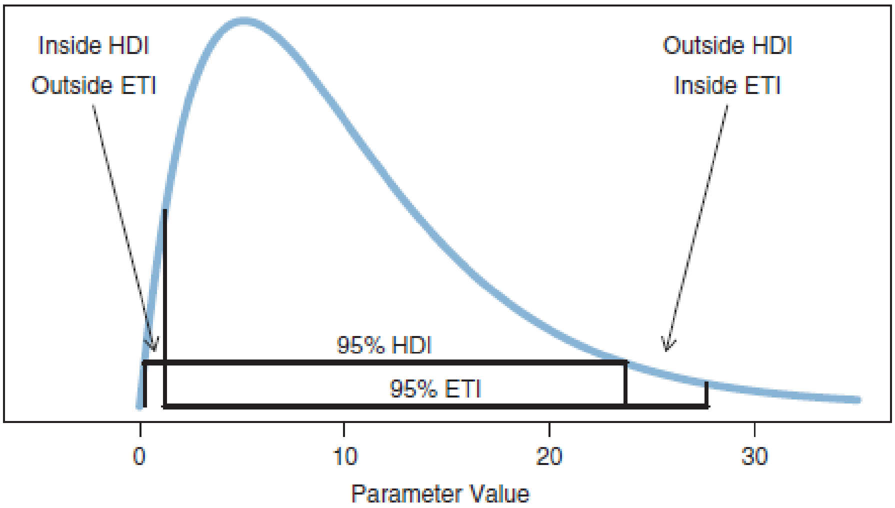
Dalam inferensi *Bayesian*, estimasi interval dikenal dengan credible interval (penulis lain ada yang menyebutnya HDI, (*the high density interval*) yang diperoleh distribusi posterior. Cara membuat *credible interval* adalah dengan menentukan distribusi posterior marginal untuk suatu parameter dan dicari pada himpunan:

$$A_\gamma = \{\psi \in \Psi : \pi(\boldsymbol{\beta}(\tau), \boldsymbol{v}, \sigma|\boldsymbol{y}) > \gamma\}$$

yang memiliki probabilitas:

$$\iiint_{A_\gamma} \pi(\boldsymbol{\beta}(\tau), \boldsymbol{v}, \sigma|\boldsymbol{y})\, d\boldsymbol{\beta}(\tau)\, d\boldsymbol{v}\, d\sigma = 1 - \alpha$$

A_γ adalah $100(1-\alpha)\%$ daerah Bayesian credible. Menurut Kruschke (2010) dalam inferensi *Bayesian* yang seanalogi dengan NHST (*Null Hypothesis Significance Testing*) adalah HDI (*High Density Interval*). Ada tiga keuntungan HDI atau credible interval, yaitu: (1). HDI secara eksplisit di sekitar $\pi(\boldsymbol{\beta}(\tau), \boldsymbol{v}, \sigma|\boldsymbol{y})$, yang merupakan tujuan parameter yang diestimasi, (2). HDI tidak tergantung pada *sampling statistic*, dan (3) HDI responsif terhadap prior yang digunakan. Interpretasi 95%HDI adalah bahwa HDI tersebut terdiri dari nilai – nilai $\boldsymbol{\beta}(\tau), \boldsymbol{v}, \text{dan } \sigma$ yang memiliki tingkat kredibilitas, yang mana nilai probabilitas total semua nilai tersebut adalah 95%. Kruschke (2010) secara implisit menunjukkan bahwa HDI (*credible interval*) dapat digunakan untuk signifikansi pengujian parameter. Pemeriksaan signifikansi hasil estimasi dapat dilakukan melalui credible interval. Jika credible interval mengandung nilai 0, maka estimasi tidak signifikan. Dan sebaliknya, jika credible interval tidak mengandung nilai 0, maka estimasi signifikan. Secara visual, HDI dan interval estimasi pendekatan frekuentis (ETI, *equal tailed interval*) pada gambar 2.



Sumber: Kruschke (2010), p.342.

Gambar 2 Sebuah Distribusi Menceng (Skewed Distribution) yang Memiliki 95% HDI yang Berbeda Dibandingkan 95% ETI

3. HASIL DAN PEMBAHASAN

Penggunaan algoritma *Gibbs Sampling*, dimana distribusi bersyarat untuk suatu parameter diketahui menyebabkan nilai kandidat untuk estimasi $\beta_1(\tau)$ dan $\beta_2(\tau)$ yang dibangkitkan distribusi tersebut pada kuantil 0.05, 0.2, 0.4, 0.6, 0.8, dan 0.96 tidak ada yang tertolak sehingga berimplikasi pada *trace plot* kandidat nilai tersebut tidak terjadi *jump process* atau dengan kata lain pergerakan kandidat nilai untuk estimasi $\beta_1(\tau)$ dan $\beta_2(\tau)$ stasioner dan histogram pada estimatornya smooth, seperti yang disajikan secara visual pada lampiran 1 dan 2.

Berikut hasil estimasi *Bayesian* untuk regresi kuantil:

- Pada Kuantil 0.05 $(\tau = 0.05)$ pada Pengeluaran Rumah Tangga Untuk Makanan

Tabel 1 Estimasi *Bayesian* Regresi Kuantil 0.05

Parameter	Estimasi	*Lower of Credible Interval: 2.5%*	*Upper Credible of Credible Interval: 97.5%*
$\beta_1(0.05)$	1.75×10^4	1.75×10^4	1.75×10^4
$\beta_2(0.05)$	0.194	0.194	0.194
σ	0.00649	0.00639	0.0066

Sumber: Pengolahan penulis.

Credible interval pada kuantil 0.05 yang disajikan tabel 1, tidak mengandung nilai nol. Maka, intersep dan koefisien pada regresi kuantil 0.05 adalah signifikan pada level signifikansi 5%. Ini berarti artinya variabel pengeluaran rumah tangga signifikan mempengaruhi pengeluaran rumah tangga untuk non makanan. Pada regresi kuantil 0.05 diperoleh koefisien 0.194, yang berarti setiap kenaikan sebesar Rp 1 pada pengeluaran rumah tangga pada kuantil 0.05, maka pengeluaran konsumsi non makanan meningkat sebesar Rp 0.194. Jadi setiap terjadi peningkatan sebesar 1 unit satuan moneter, maka sebanyak 19.4 persen untuk konsumsi non makanan, dan sisanya 80.6 persen untuk konsumsi makan dan minum. Garis regresi kuantil 0.01 pendekatan *Bayesian* berada paling bawah dibandingkan garis regresi kuantil *Bayesian* lainnya.

➢ Pada Kuantil 0.2 $(\tau = 0.2)$ pada Pengeluaran Rumah Tangga Untuk Makanan

Tabel 2 Estimasi *Bayesian* Regresi Kuantil 0.2

Parameter	Estimasi	*Lower of Credible Interval: 2.5%*	*Upper Credible of Credible Interval: 97.5%*
$\beta_1(0.2)$	-1.12×10^5	-1.12×10^5	-1.12×10^5
$\beta_2(0.2)$	0.323	0.323	0.323
σ	0.00649	0.00639	0.0066

Sumber: Pengolahan penulis.

Credible interval pada kuantil 0.2 yang disajikan tabel 2, tidak mengandung nilai nol. Maka, intersep dan koefisien pada regresi kuantil 0.2 adalah signifikan pada level signifikansi 5%. Ini berarti artinya variabel pengeluaran rumah tangga signifikan mempengaruhi pengeluaran rumah tangga untuk non makanan. Pada regresi kuantil 0.2 diperoleh koefisien 0.323, yang berarti setiap kenaikan sebesar Rp 1 pada pengeluaran rumah tangga pada kuantil 0.2, maka pengeluaran konsumsi non makanan meningkat sebesar Rp 0.323. Jadi setiap terjadi peningkatan sebesar 1 unit satuan moneter, maka sebanyak 32.3 persen untuk konsumsi non makanan, dan sisanya 67.7 persen untuk konsumsi makan dan minum.

➢ Pada kuantil 0.4 $(\tau = 0.4)$ pada Pengeluaran Rumah Tangga Untuk Makanan

Tabel 3 Estimasi *Bayesian* Regresi Kuantil 0.2

Parameter	Estimasi	*Lower of Credible Interval: 2.5%*	*Upper Credible of Credible Interval: 97.5%*
$\beta_1(0.4)$	-1.53×10^5	-1.53×10^5	-1.53×10^5
$\beta_2(0.4)$	0.408	0.408	0.408
σ	0.00649	0.00639	0.0066

Sumber: Pengolahan penulis.

Credible interval pada kuantil 0.4 yang disajikan tabel 3, tidak mengandung nilai nol. Maka, intersep dan koefisien pada regresi kuantil 0.4 adalah signifikan pada level signifikansi 5%. Ini berarti artinya variabel pengeluaran rumah tangga signifikan mempengaruhi pengeluaran rumah tangga untuk non makanan. Pada regresi kuantil 0.4 diperoleh koefisien 0.408, yang berarti setiap kenaikan sebesar Rp 1 pada pengeluaran rumah tangga pada kuantil 0.4, maka pengeluaran konsumsi non makanan meningkat sebesar Rp 0.408. Jadi setiap terjadi peningkatan sebesar 1 unit satuan moneter, maka sebanyak 40.8 persen untuk konsumsi non makanan, dan sisanya 59.2 persen untuk konsumsi makan dan minum.

➢ Pada kuantil 0.6 $(\tau = 0.6)$ pada Pengeluaran Rumah Tangga Untuk Makanan

Tabel 4 Estimasi *Bayesian* Regresi Kuantil 0.6

Parameter	Estimasi	*Lower of Credible Interval: 2.5%*	*Upper Credible of Credible Interval: 97.5%*
$\beta_1(0.6)$	-1.96×10^5	-1.96×10^5	-1.53×10^5
$\beta_2(0.6)$	0.494	0.494	0.494
σ	0.00649	0.00639	0.0066

Sumber: Pengolahan penulis.

Credible interval pada kuantil 0.4 yang disajikan tabel 4, tidak mengandung nilai nol. Maka, intersep dan koefisien pada regresi kuantil 0.6 adalah signifikan pada level signifikansi 5%. Ini berarti artinya variabel pengeluaran rumah tangga signifikan mempengaruhi pengeluaran rumah tangga untuk non makanan. Pada regresi kuantil 0.6 diperoleh koefisien 0.408, yang berarti setiap kenaikan sebesar Rp 1 pada pengeluaran rumah tangga pada kuantil 0.6, maka pengeluaran konsumsi non makanan meningkat sebesar Rp 0.494. Jadi setiap terjadi peningkatan sebesar 1 unit satuan moneter, maka sebanyak 49.4 persen untuk konsumsi non makanan, dan sisanya 50.6 persen untuk konsumsi makan dan minum.

- Pada kuantil ke 0.8 $(\tau = 0.8)$ pada Pengeluaran Rumah Tangga Untuk Makanan

Tabel 5 Estimasi *Bayesian* Regresi Kuantil 0.8

Parameter	Estimasi	*Lower of Credible Interval: 2.5%*	*Upper Credible of Credible Interval: 97.5%*
$\beta_1(0.8)$	-2.53×10^5	-2.53×10^5	-2.53×10^5
$\beta_2(0.8)$	0.609	0.609	0.609
σ	0.00649	0.00639	0.0066

Sumber: Pengolahan penulis.

Credible interval pada kuantil 0.8 yang disajikan tabel 5, tidak mengandung nilai nol. Maka, intersep dan koefisien pada regresi kuantil 0.8 adalah signifikan pada level signifikansi 5%. Ini berarti artinya variabel pengeluaran rumah tangga signifikan mempengaruhi pengeluaran rumah tangga untuk non makanan. Pada regresi kuantil 0.8 diperoleh koefisien 0.609, yang berarti setiap kenaikan sebesar Rp 1 pada pengeluaran rumah tangga pada kuantil 0.8, maka pengeluaran konsumsi non makanan meningkat sebesar Rp 0.609. Jadi setiap terjadi peningkatan sebesar 1 unit satuan moneter, maka sebanyak 60.9 persen untuk konsumsi non makanan, dan sisanya 39.1 persen untuk konsumsi makan dan minum.

- Pada kuantil ke 0.96 $(\tau = 0.96)$ pada Pengeluaran Rumah Tangga Untuk Makanan

Credible interval pada kuantil 0.96 yang disajikan tabel 6, tidak mengandung nilai nol. Maka, intersep dan koefisien pada regresi kuantil 0.96 adalah signifikan pada level signifikansi 5%. Ini berarti artinya variabel pengeluaran rumah tangga signifikan mempengaruhi pengeluaran rumah tangga untuk non makanan. Pada regresi kuantil 0.96 diperoleh koefisien 0.763, yang berarti setiap kenaikan sebesar Rp 1 pada pengeluaran rumah tangga pada kuantil 0.96, maka pengeluaran konsumsi non makanan meningkat sebesar Rp 0.763. Jadi setiap terjadi peningkatan sebesar 1 unit satuan moneter, maka sebanyak 76.3 persen untuk konsumsi non makanan, dan sisanya 23.7 persen untuk konsumsi makan dan minum.

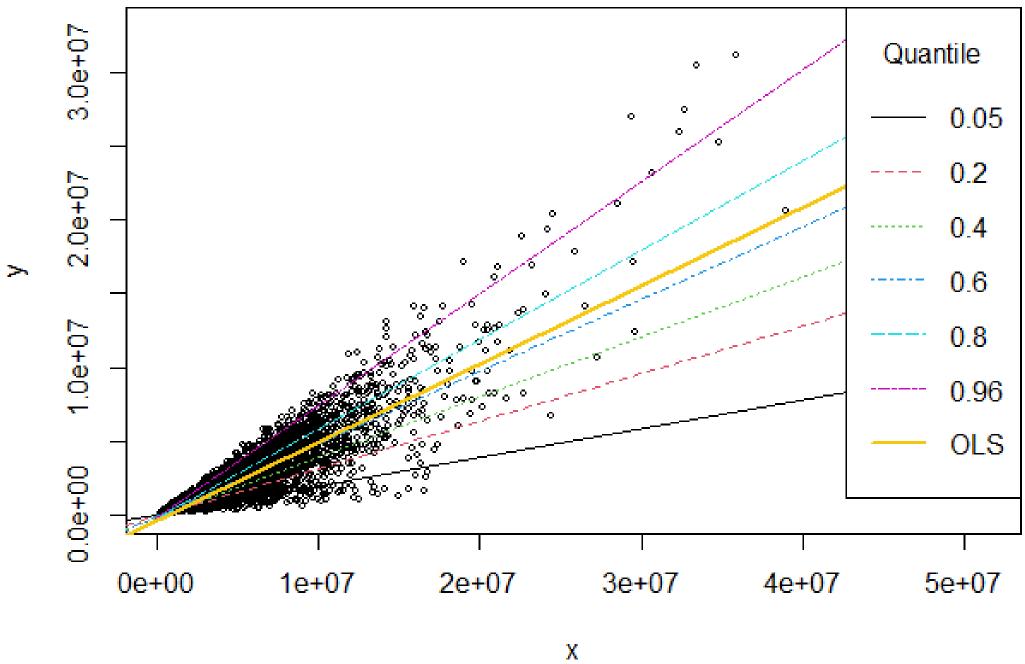
Tabel 6 Estimasi *Bayesian* Regresi Kuantil 0.96

Parameter	Estimasi	*Lower of Credible Interval: 2.5%*	*Upper Credible of Credible Interval: 97.5%*
$\beta_1(0.96)$	-2.69×10^5	-2.69×10^5	-2.69×10^5
$\beta_2(0.96)$	0.763	0.763	0.763
σ	0.00649	0.00639	0.0066

Sumber: Pengolahan penulis.

Dari hasil estimasi regresi kuantil *Bayesian* yang tersaji pada tabel 1 – 6, menunjukkan bahwa proporsi pengeluaran konsumsi rumah tangga untuk non makanan berbeda-beda untuk setiap tingkat pengeluarannya. Semakin tinggi pengeluaran rumah tangga, maka proporsi konsumsi untuk non makanan, semakin besar proporsinya. Dan sebaliknya, rumah tangga dengan level pengeluaran rendah, Proporsi konsumsi untuk makan dan minum lebih besar daripada konsumsi untuk non makanan. Kemudian nilai estimasi σ yang sama pada setiap kuantil, karena σ tidak terikat pada kuantil yang didasarkan pada persamaan (10).

Secara visual, garis regresi kuantil *Bayesian* untuk kuantil 0.05, 0.2, 0.4, 0.6, 0.8 dan 0.96 disajikan gambar 3 berikut:



Sumber: Pengolahan penulis.

Gambar 3 Garis Regresi Kuantil *Bayesian* dan Regresi *Ordinary Least Square* (OLS)

Terlihat pada gambar 3, regresi kuantil *Bayesian* jauh lebih representatif mereflesikan sebaran data dibandingkan regresi linear biasa yang hanya sebaran data disekitar rata-rata. Sehingga, spesifikasi pemodelan antara pengeluaran konsumsi rumah tangga terhadap pengeluaran konsumsi rumah tangga untuk non makanan yang secara visual disajikan dalam kurva Engel, dimana menyiratkan pemodelan kurva Engel dengan regresi kuantil adalah tepat. Kemudian pendekatan *Bayesian* dalam estimasi regresi kuantil menghasilkan nilai estimasi yang memiliki probabilitas yang besar mendekati nilai parameter karena adanya informasi prior dan proses MCMC yang membangkitkan nilai – nilai kandidat untuk estimasi parameter dari distribusi posterior bersyarat parameternya.

4. KESIMPULAN

Berdasarkan pembahasan dapat ditarik kesimpulan, yaitu: (1). Regresi kuantil Bayesian jauh lebih representatif karena mencakup sebaran data, baik pada sentral data maupun daerah disekitar sentral data, dibandingkan regresi linier sederhana yang hanya mencakup berkisar apda senral data. Ini berimplikasi bahwa pemodelan kurva Engel dengan regresi kuantil adalah tepat, dan (2). Pengeluaran konsumsi rumah tangga berpengaruh positif dan signifikan pada pengeluran konsumsi rumah tangga untuk non makanan. Proporsi pengeluaran konsumsi rumah tangga untuk non makanan berbeda-beda untuk setiap tingkat pengeluarannya. Semakin tinggi pengeluaran rumah tangga, maka proporsi konsumsi untuk non makanan, semakin besar proporsinya. Dan sebaliknya, rumah tangga dengan level pengeluaran rendah, Proporsi konsumsi untuk makan dan minum lebih besar daripada konsumsi untuk non makanan.

Untuk penelitian selanjutnya, penambahan variabel prediktor lainnya sangat memungkinkan dalam spesifikasi kurva Engel dan metode estimasi parameter regresi kuantil kurva Engel dengan menggunakan neural network dapat diaplikasikan.

DAFTAR PUSTAKA

Benoit D.F. & Van den Poel, D. (2017). BayesQR: A Bayesian approach to quantile regression. Journal of Statistical Software, 76(7), 1-32, 2017.

Chai, A. & A. Moneta, A. (2010). Retrospectives: Engel curves. Journal of Economic Perspectives, 24(1), 225–240.

Choi, H.M. & and J.P. Hobert, J.P. (2013). Analysis of MCMC algorithms for Bayesian linear regression with Laplace errors. Journal of Multivariate Analysis, 117, 32–40.

Koenker, R & Bassett, G. (1978). Regression quantiles. Econometrica, 46(1), 33–50.

Kozumi, H. & Kobayashi, G. (2011). Gibbs Sampling methods for Bayesian quantile regression. Journal of Statistical Computation and Simulation, 81(11), 1565–1578.

Kruschke, J.K. (2010). Doing Bayesian Data Analysis: A Tutorial with R and BUGS. Academic Press/Elsevier.

Leser, C.E.V. (1963). Form of Engel functions. Econometrica, 31(4), 694-703.

Lewbel, A. (2007). Engel Curves. The New Palgrave Dictionary of Economics. https://www2.bc.edu/arthur-lewbel/palengel.pdf

Tilak, T.B.G. (2000). Education poverty in India. NIEPA Occasional Paper 29.

Timmer, C.P., Falcon, W.P. dan Pearson, S. R. (1983). Food policy Analysis. Baltimore: Johns Hopkins University Press.

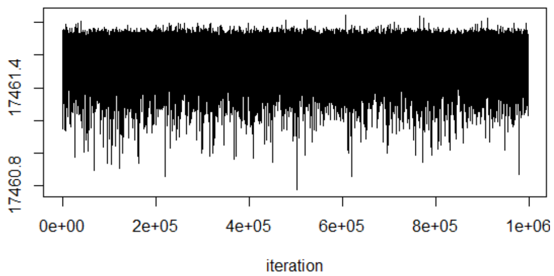
Tsionas, E.G. (2003). Bayesian quantile inference. Journal of Statistical Computation and Simulation,73(9), 659–674.

Yu, K. & R.A. Moyeed, R.A. (2001). Bayesian quantile regression. Statistics & Probability Letters, 54, 437–447.

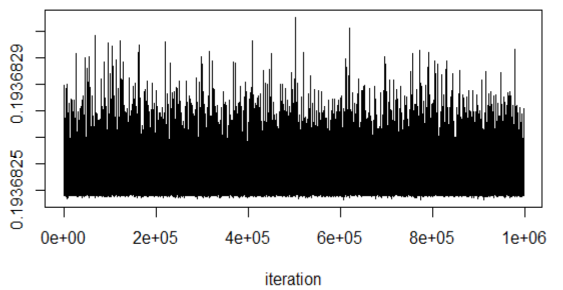
LAMPIRAN

1. Trace Plot

Kuantil 0.05

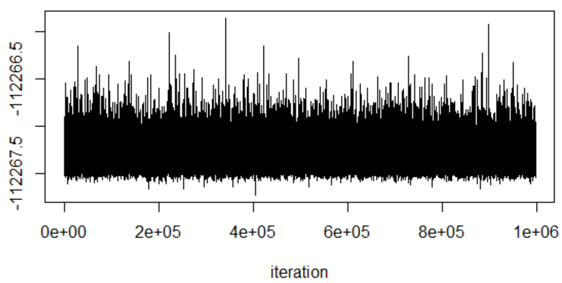


Intersep

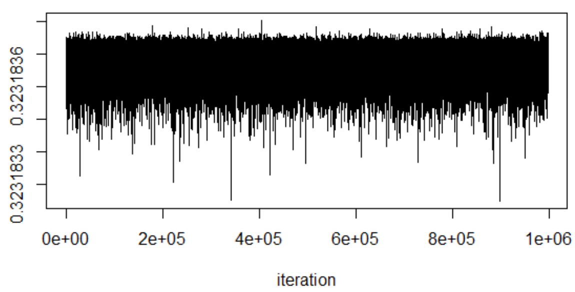


Koefisien

Kuantil 0.2

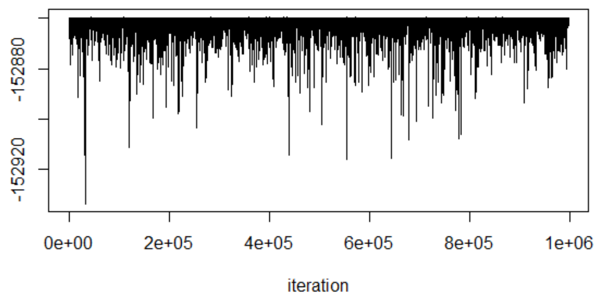


Intersep

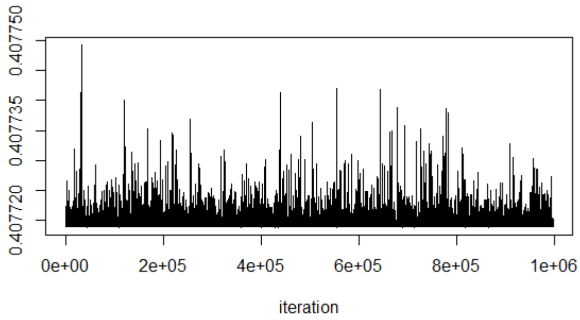


Koefisien

Kuantil 0.4

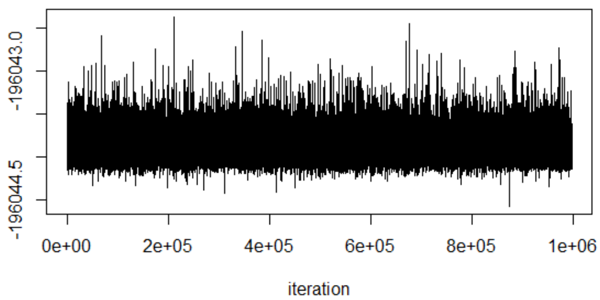


Intersep

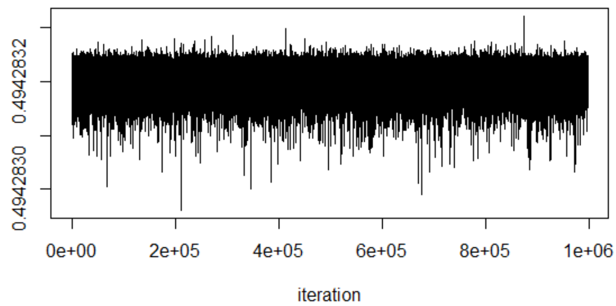


Koefisien

Kuantil 0.6

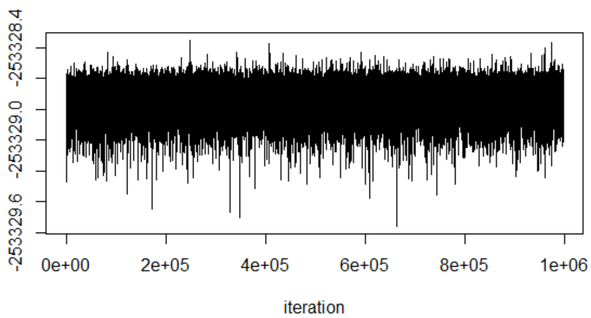


Intersep

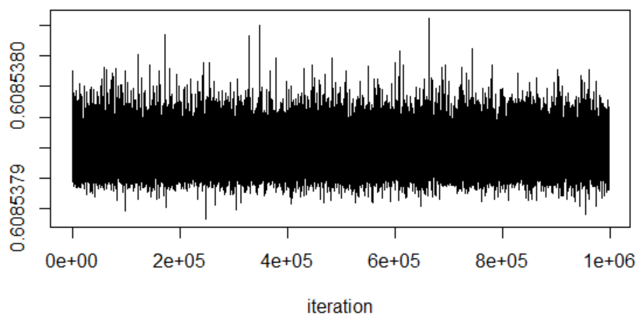


Koefisien

Kuantil 0.8

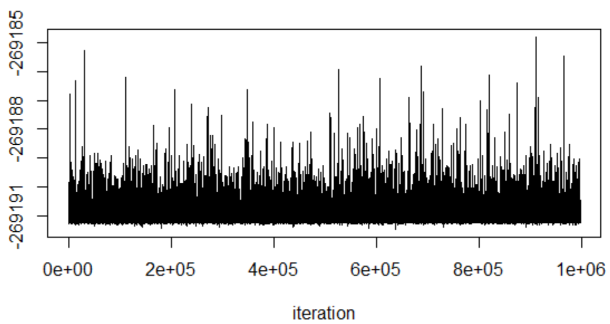


Intersep

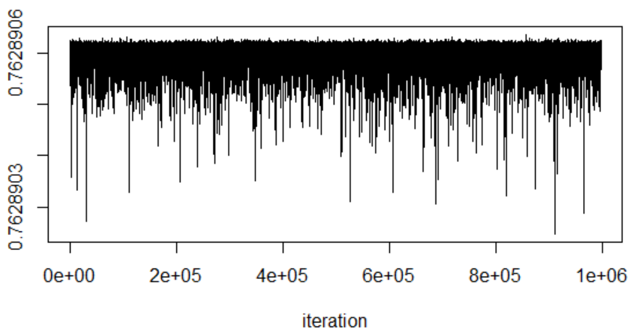


Koefisien

Kuantil 0.96



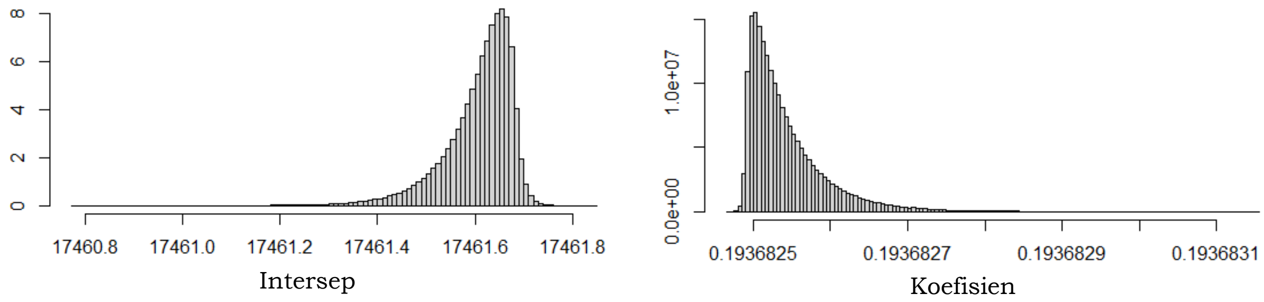
Intersep



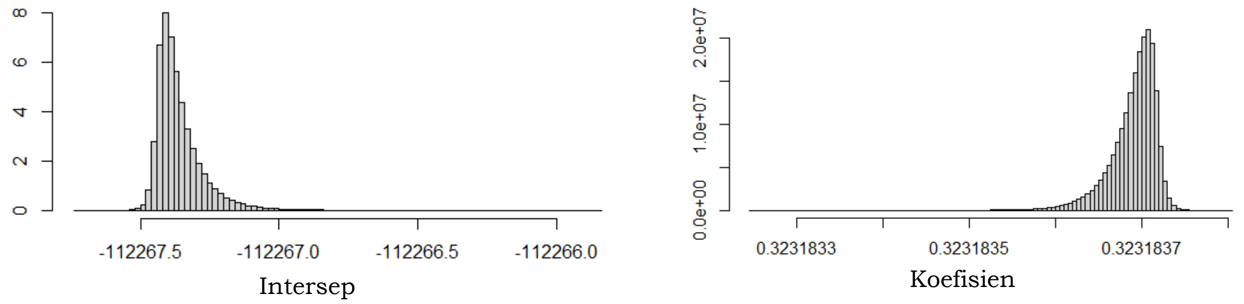
Koefisien

2. Histogram

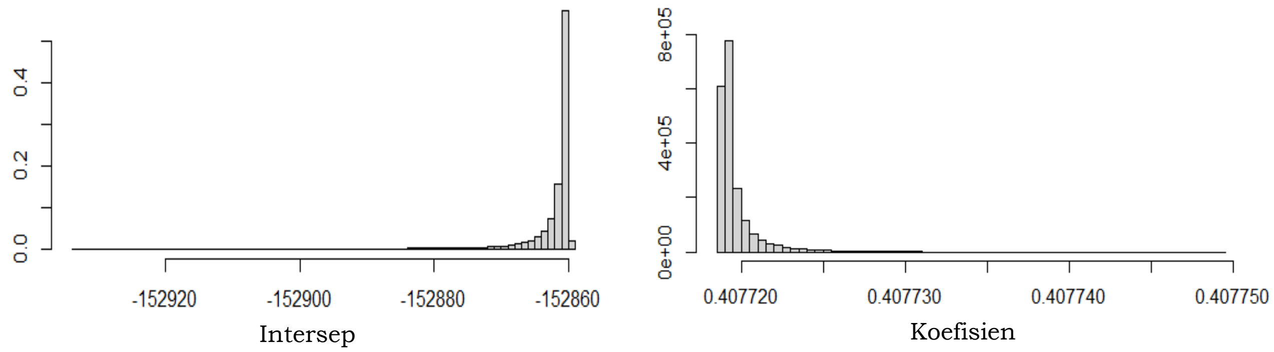
Kuantil 0.05



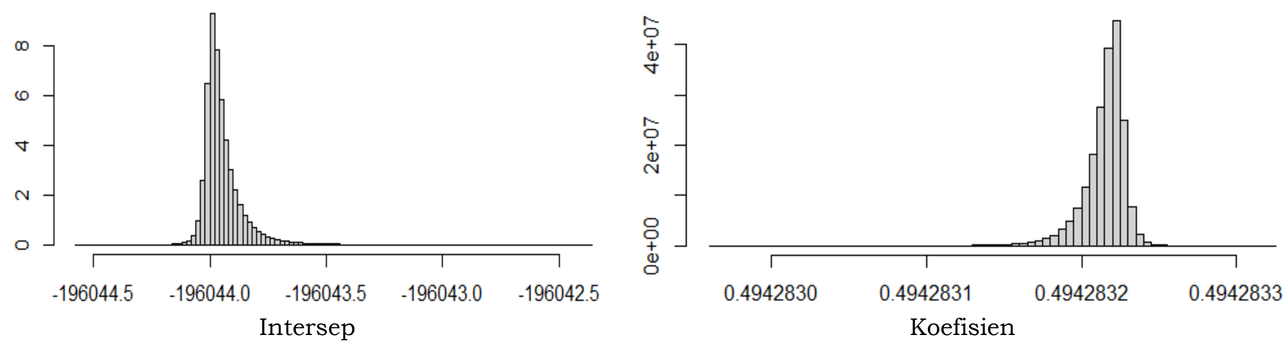
Kuantil 0.2

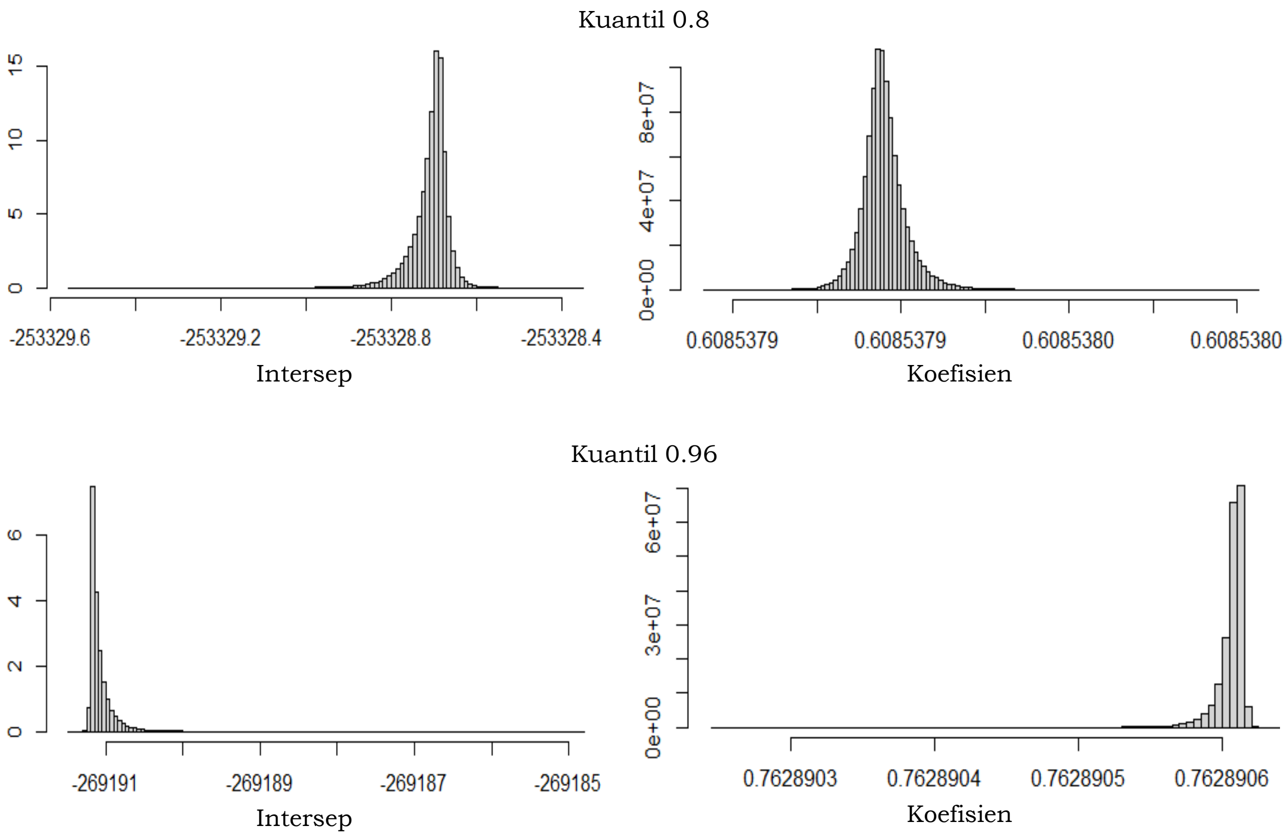


Kuantil 0.4



Kuantil 0.6




Kuantil 0.8
0
5
10
15
-253329.6
-253329.2
-253328.8
-253328.4
Intersep
0e+00
4e+07
8e+07
0.6085379
0.6085379
0.6085380
0.6085380
Koefisien
Kuantil 0.96
0
2
4
6
-269191
-269189
-269187
-269185
Intersep
0e+00
3e+07
6e+07
0.7628903
0.7628904
0.7628905
0.7628906
Koefisien