LAPORAN PENELITIAN

REGRESI KUANTIL MEDIAN UNTUK MENGATASI HETEROSKEDASTISITAS

Diusulkan Oleh:
Dr. Edy Widodo, S.Si., M.Si.
Febria Pradita Prima Andani

PROGRAM STUDI STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS ISLAM INDONESIA
2016

HALAMAN PENGESAHAN

1. Identitas Penelitian

a. Judul Penelitian	: Regresi Kuantil Median untuk Mengatasi Heteroskedastisitas
b. Bidang Ilmu	: Statistika
c. Kategori Penelitian	: Unggulan

2. Ketua Peneliti

a. Nama Lengkap dan Gelar	: Dr. Edy Widodo, S.Si., M.Si.
b. Jenis Kelamin	: Laki-laki
c. Golongan dan Pangkat	: IIIc/Penata
d. NIP/NIK	: 966110103
e. Jabatan Fungsional	: Lektor
f. Fakultas/Jurusan	: MIPA/Statistika

3. Alamat Ketua Peneliti

a. Alamat Kantor	: FMIPA UII, Jalan Kaliurang Km 14,4 Yogyakarta
b. Telp/Fax:	: 0274 895920 Ext 3042/Fax Ext 3020
c. Email	: edywidodo@uii.ac.id
d. Alamat Rumah	: Dusun Mendiro, RT 04 RW 26 Sukoharjo, Ngaglik, Sleman, Yogyakarta
e. Telp/Hp	: 082242482225

4. Jumlah Anggota Peneliti

a. Anggota Peneliti I	: Febria Pradita Prima Andani
5. Lokasi Penelitian	: Laboratorium Statistika FMIPA UII
6. Jangka Waktu Pelaksanaan	: 3 Bulan

Yogyakarta, 15 April 2016

Mengetahui: Ketua Jurusan Statistika	Ketua Peneliti
(Dr. RB. Fajriya Hakim, S.Si., M.Si.) NIP/NIK: 986110101	(Dr. Edy Widodo, S.Si., M.Si.) NIP/NIK: 966110103

Menyetujui, Direktur
DPPM UII

(Prof. Akhmad Fauzi, Ph.D)
NIP/NIK: 956110101

DAFTAR ISI

DAFTAR PUSTAKA

LAMPIRAN

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR LAMPIRAN

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Analisis regresi adalah suatu metode statistik yang memanfaatkan hubungan antara dua variabel atau lebih (Soejoeti, 1986). Tujuan dari analisis regresi adalah mengetahui hubungan antara satu atau lebih variabel prediktor (X) terhadap variabel respon (Y). Analisis regresi dapat digunakan untuk membentuk model dan melakukan estimasi. Menurut Qudratullah (2013), salah satu metode estimasi parameter dalam analisis regresi yang biasa digunakan adalah Metode Kuadrat Terkecil (MKT) atau *Ordinary Least Square* (*OLS*). Prinsip dari metode kuadrat terkecil yaitu dengan meminimumkan jumlah kuadrat *error*. Menurut Uthami (2013), MKT memiliki beberapa asumsi yang harus dipenuhi untuk mendapatkan estimator yang memiliki sifat *Best Linier Unbiased Estimator* (*BLUE*). Asumsi-asumsi yang harus dipenuhi yaitu, *error* berdistribusi normal, homoskedastisitas, tidak terdapat multikolinieritas, dan tidak terdapat autokorelasi. Apabila asumsi tersebut tidak terpenuhi, maka estimator tidak memenuhi sifat *BLUE*.

Pada kenyataannya dalam beberapa kasus sering ditemukan asumsi-asumsi regresi yang tidak terpenuhi, sehingga estimasi dengan menggunakan MKT menjadi kurang tepat. Salah satu asumsi yang tidak terpenuhi adalah asumsi homoskedastisitas, yaitu *error* mempunyai variansi yang konstan. Hal tersebut disebabkan oleh data fluktuatif sehingga terdapat *outlier* pada data tersebut. Menurut Makkulau, dkk (2010) *outlier* adalah pengamatan yang berada jauh dari pengamatan-pengamatan lainnya.

Apabila asumsi homoskedastisitas tidak terpenuhi, berarti *error* tersebut mempunyai variansi yang tidak konstan. Hal ini disebut dengan heteroskedastisitas (Mendenhall, 1996). Menurut Mokosolang, dkk (2015) heteroskedastisitas dapat dideteksi dengan menggunakan beberapa metode statistika, yaitu dengan metode grafik dan uji statistik. Uji statistik yang dapat digunakan dalam pendeteksian heteroskedastisitas antara lain uji Glejser, uji Park, uji Korelasi Rank Spearman, dan uji Goldfeld-Quandt.

Heteroskedastisitas menyebabkan hasil estimasi yang diperoleh mempunyai variansi yang tidak efisien, yang berarti variansi cenderung membesar sehingga tidak lagi memiliki variansi yang kecil. Adanya heteroskedastisitas tersebut dapat mengganggu model yang terbentuk, bahkan dapat menyesatkan kesimpulan yang diambil. Oleh karena itu, penanganan heteroskedastisitas penting untuk dilakukan supaya mendapatkan model yang dapat dipercaya dan tidak menyesatkan kesimpulan (Maziyya, 2015). Terjadinya kesalahan dalam menyimpulkan hasil penelitian akan menimbulkan permasalahan baru, seperti kesalahan dalam pengambilan kebijakan sehingga dapat memperburuk permasalahan atau kondisi yang diteliti. Berdasarkan alasan tersebut maka sangatlah penting untuk menangani heteroskedastisitas.

Regresi kuantil adalah salah satu metode regresi yang digunakan untuk mengatasi permasalahan heteroskedastisitas. Regresi kuantil dilakukan dengan membagi atau memisahkan data menjadi dua bagian atau lebih ketika dicurigai terdapat perbedaan nilai estimator pada kuantil-kuantil tertentu. Regresi kuantil merupakan salah satu metode regresi yang diperkenalkan oleh Roger Koenker dan Gilbert Basset pada tahun 1978. Keuntungan utama dari regresi kuantil adalah sangat berguna ketika digunakan pada data yang distribusinya tidak homogen dan tidak simetris. Bahkan regresi kuantil tidak terpengaruh oleh *outlier*, sehingga tidak mengganggu kestabilan pada data (Furno, 2013).

Berdasarkan uraian di atas, maka penulis ingin mengetahui kemampuan regresi kuantil dalam mengatasi heteroskedastisitas dalam melakukan estimasi parameter. Hal ini sangat penting untuk diteliti, karena regresi kuantil merupakan salah satu metode alternatif dalam melakukan analisis data yang mengandung heteroskedastisitas. Sehingga penanganan pada data yang mengandung heteroskedastisitas benar-benar menggunakan metode yang tepat dan didapatkan hasil analisis yang baik. Maksud dari analisis yang baik, yaitu yang dapat dipercaya dan tidak menyesatkan peneliti dalam mengambil kesimpulan.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan maka permasalahan yang akan dibahas dalam penelitian ini adalah sebagai berikut:

1. Bagaimana kemampuan regresi kuantil median dalam mengatasi data yang mengandung heteroskedastisitas?
2. Bagaimana perbandingan metode regresi yang dihasilkan dari MKT dengan regresi kuantil median?

1.3 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Mengetahui kemampuan metode regresi kuantil median dalam menyelesaikan permasalahan heteroskedastisitas pada data yang mengandung heteroskedastisitas.
2. Mengetahui perbandingan metode regresi yang dihasilkan dengan menggunakan MKT dan regresi kuantil median pada data dengan homoskedastisitas dan data dengan heteroskedastisitas.

1.4 Manfaat Penelitian

Dengan dilakukannya penelitian dengan menggunakan regresi kuantil median pada data yang mengandung heteroskedastisitas, maka didapatkan manfaat sebagai berikut:

1. Memberikan pengetahuan mengenai metode regresi yang lebih baik digunakan ketika terdapat heteroskedastisitas.
2. Sebagai pembelajaran dalam menyelesaikan permasalahan heteroskedastisitas ketika menggunakan analisis regresi.
3. Sebagai referensi mengenai regresi kuantil median.
4. Sebagai acuan untuk melakukan penelitian serupa dengan data yang mengandung heteroskedastisitas dengan metode selain regresi kuantil median.

BAB II

TINJAUAN PUSTAKA

Terkait dengan penelitian yang dilakukan penulis, maka penelitian terdahulu menjadi sangat penting untuk diketahui supaya dapat mengetahui hubungan antara penelitian yang telah dilakukan dengan penelitian yang dilakukan saat ini dan supaya terhindar dari adanya penjiplakan atau duplikasi penelitian. Berikut adalah beberapa penelitian terdahulu yang berkaitan dengan permasalahan heteroskedastisitas pada analisis regresi ataupun yang berkaitan dengan regresi kuantil median.

Kirnasari (2014) dalam penelitiannya menjelaskan bahwa dalam menangani kasus heteroskedastisitas pada data harga saham perusahaan dan kurs nilai tengah *IDR* terhadap *USD*, digunakan dua metode, yaitu regresi kuantil median dan *WLS.* Setelah dilakukan analisis dan perbandingan hasil dengan kedua metode tersebut, didapatkan kesimpulan bahwa metode *WLS* lebih baik dalam menyelesaikan kasus heteroskedastisitas pada data tersebut.

Uthami, dkk (2013) dalam penelitiannya membahas tentang bagaimana nilai estimasi parameter dengan menggunakan regresi kuantil median. Studi kasus dalam penelitian tersebut digunakan data *Passenger Car Milage* yang diperoleh dari buku *Basic Econometrics* dengan pengarang Gujarati (2004). Data tersebut mengandung heteroskedastisitas sehingga untuk melakukan estimasi digunakan regresi kuantil median. Estimasi dilakukan dengan mengestimasi di setiap nilai kuantilnya, kemudian dipilih nilai estimasi pada kuantil median. Hasil penelitian menunjukkan bahwa regresi kuantil median dapat mengatasi heteroskedastisitas.

Penelitian Rahmawati, dkk (2011) membahas tentang bagaimana nilai estimasi parameter dengan menggunakan regresi kuantil. Dalam studi kasus tersebut, plot suhu harian bersifat tidak simetris dan dicurigai terjadi heteroskedastisitas. Jika pendugaan dilakukan dengan menggunakan MKT akan didapatkan estimator parameter yang tidak efisien, maka digunakan regresi kuantil. Hasil dari analisis menunjukkan bahwa nilai estimasi parameter dengan menggunakan regresi kuantil dan MKT memberikan hasil yang berbeda.

Dalam penelitian Fitriah (2009) yang bertujuan untuk mempelajari sifat-sifat hampiran bagi regresi kuantil, diperoleh hasil bahwa sifat hampiran regresi kuantil menyatakan bahwa vektor parameter regresinya dapat meminimumkan nilai harapan dari kuadrat *error* terboboti.

Berdasarkan uraian di atas mengenai regresi kuantil, maka penulis tertarik untuk melakukan penelitian mengenai regresi kuantil median untuk mengatasi heteroskedastisitas dan membandingkan antara regresi menggunakan MKT dengan regresi kuantil median dalam analisis data homoskedastisitas dan heteroskedastisitas.

2.1 Landasan Teori

2.1.1 Analisis Regresi

Analisis regresi adalah suatu metode statistik yang memanfaatkan hubungan antara dua variabel atau lebih (Soejoeti, 1986). Analisis regresi digunakan untuk mengetahui pengaruh variabel prediktor terhadap variabel respon (Sunyoto, 2007). Tujuan analisis regresi, yaitu melakukan prediksi yang dapat dipercaya mengenai nilai variabel respon menggunakan nilai variabel prediktor yang telah diketahui (Qudratullah, 2013).

Di dalam Youlanda (2015), Gujarati pada tahun 2003 menyatakan bahwa terdapat dua jenis regresi yang terkenal, yaitu regresi linier sederhana dan regresi linier berganda. Model regresi linier sederhana adalah sebagai berikut (Montgomery, 1982):

$$Y = \beta_0 + \beta_1 X + \varepsilon \quad (2.1)$$

dengan:

Y : Variabel respon

β_0 : Intersep pada sumbu y, titik potong sumbu y

β_1 : Kemiringan (*slope*) garis regresi

X : Variabel prediktor

ε : Variabel acak

Model regresi linier berganda adalah model regresi yang melibatkan lebih dari satu variabel prediktor. Model regresi linier berganda adalah sebagai berikut (Montgomery, 1982):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon \qquad (2.2)$$

dengan:

Y : Variabel respon yang akan diprediksi

β_0 : Intersep pada sumbu y, titik potong sumbu y

β_j : Parameter; j = 1, 2, . . ., k

X_j : Variabel prediktor; j = 1, 2, . . ., k

ε : Variabel acak

Jika disusun dalam bentuk matriks, maka persamaan (2.2) menjadi:

$$\boldsymbol{Y}_{(nx1)} = \boldsymbol{X}_{nx(j+1)}\boldsymbol{\beta}_{(j+1)x1} + \boldsymbol{\varepsilon}_{nx1} \qquad (2.3)$$

dengan:

$\boldsymbol{Y}$: Vektor amatan yang berukuran (n x 1)

$\boldsymbol{X}$: Matriks berukuran (n × (j+1)) yang diketahui

$\boldsymbol{\beta}$: Vektor parameter yang berukuran ((j+1) × 1)

$\boldsymbol{\varepsilon}$: Vektor *error* yang berukuran (n × 1)

Menurut Sembiring (2003), dalam model tersebut diasumsikan bahwa X_i tidak mempunyai distribusi dan nilainya dapat ditentukan oleh peneliti dengan distribusi ε merupakan *error* acak yang berdistribusi $N(0, \sigma^2)$. Maka Y memiliki distribusi yang sesuai dengan ε.

Analisis regresi memiliki parameter-parameter yang perlu diestimasi karena nilainya belum diketahui. Metode estimasi parameter yang sering digunakan adalah Metode Kuadrat Terkecil (MKT). MKT akan menemukan nilai-nilai estimasi dengan cara meminimumkan jumlah kuadrat *error* (Draper dan Smith, 1992). Jumlah kuadrat *error* (J) dari persamaan (2.2) adalah sebagai berikut (Montgomery, 1982):

$$J = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \cdots - \beta_k x_{ik})^2 \qquad (2.4)$$

Pada persamaan (2.3) jumlah *error* diminimumkan untuk mendapatkan nilai-nilai bagi β dengan cara melakukan penurunan persamaan (2.4) secara parsial terhadap β_j dan disamadengankan nol, sehingga diperoleh:

$$\begin{aligned}
\frac{\partial J}{\partial \beta_0} &= -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \cdots - \beta_k x_{ik}) &= 0 \\
\frac{\partial J}{\partial \beta_1} &= -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \cdots - \beta_k x_{ik})\, x_{i1} &= 0 \\
\frac{\partial J}{\partial \beta_2} &= -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \cdots - \beta_k x_{ik})\, x_{i2} &= 0 \\
\vdots \qquad & \qquad\qquad\qquad\qquad \vdots & \vdots \\
\frac{\partial J}{\partial \beta_k} &= -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \cdots - \beta_k x_{ik})\, x_{ik} &= 0
\end{aligned} \qquad (2.5)$$

Kemudian persamaan (2.5) dijabarkan dan menghasilkan persamaan sebagai berikut:

$$\begin{aligned}
n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{i2} + \cdots + \hat{\beta}_\mathrm{k} \sum_{i=1}^{n} x_{ik} &= \sum_{i=1}^{n} y_i \\
\hat{\beta}_0 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_1 \sum_{i=1}^{n} {x_{i1}}^2 + \hat{\beta}_2 \sum_{i=1}^{n} x_{i1}x_{i2} + \cdots + \hat{\beta}_\mathrm{k} \sum_{i=1}^{n} x_{i1}x_{ik} &= \sum_{i=1}^{n} x_{i1}y_i \\
\hat{\beta}_0 \sum_{i=1}^{n} x_{i2} + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1}x_{i2} + \hat{\beta}_2 \sum_{i=1}^{n} {x_{i2}}^2 + \cdots + \hat{\beta}_\mathrm{k} \sum_{i=1}^{n} x_{i2}x_{ik} &= \sum_{i=1}^{n} x_{i2}y_i \\
\vdots \qquad\qquad\qquad\qquad \vdots \qquad\qquad & \\
\hat{\beta}_0 \sum_{i=1}^{n} x_{ik} + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1}x_{ik} + \hat{\beta}_2 \sum_{i=1}^{n} x_{i2}x_{ik} + \cdots + \hat{\beta}_\mathrm{k} \sum_{i=1}^{n} {x_{ik}}^2 &= \sum_{i=1}^{n} x_{ik}y_i
\end{aligned}$$

(2.6)

Jika disusun dalam bentuk persamaan matriks, maka persamaan (2.6) menjadi:

$$\boldsymbol{X'X\hat{\beta}} = \boldsymbol{X'Y} \qquad (2.7)$$

dengan,

$$\boldsymbol{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \qquad \boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \quad \widehat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_{\mathrm{k}} \end{bmatrix}$$

$$\boldsymbol{X'X} = \begin{bmatrix} n & \sum_{i=1}^{n} x_{i1} & \cdots & \sum_{i=1}^{n} x_{ik} \\ \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i1}^2 & \cdots & \sum_{i=1}^{n} x_{i1}x_{ik} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} x_{ik} & \sum_{i=1}^{n} x_{i1}x_{ik} & \cdots & \sum_{i=1}^{n} x_{ik}^2 \end{bmatrix}$$

$$\boldsymbol{X'Y} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_{21} & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{1k} & x_{2k} & x_{3k} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i1}y_i \\ \vdots \\ \sum_{i=1}^{n} x_{ik}y_i \end{bmatrix}$$

Nilai estimasi β dicari dengan menggunakan persamaan (2.7) yang kedua ruasnya dikalikan dengan invers dari $(\boldsymbol{X'X})$, sehingg diperoleh:

$$(\boldsymbol{X'X})^{-1}\boldsymbol{X'X}\widehat{\boldsymbol{\beta}} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'Y}$$

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'Y} \qquad (2.8)$$

Perhitungan nilai $\hat{\beta}$ dapat diterapkan dalam regresi sederhana maupun regresi berganda. Berikut adalah contoh penerapan MKT.

<u>Contoh 1</u> (Montgomery, 1982):

Seorang yang membotoli minuman ringan menganalisis rute layanan mesin penjual otomatis di sistem distribusinya. Ia tertarik dalam memprediksi jumlah waktu yang diperlukan oleh pengemudi dengan layanan mesin penjual otomatis di *outlet*. Kegiatan layanan ini termasuk mengisi mesin dengan produk minuman dan pemeliharaan ringan atau pembenahan. Insinyur industri yang bertanggung jawab untuk penelitian telah menyarankan bahwa dua variabel yang paling penting yang mempengaruhi waktu pengiriman adalah jumlah kasus produk tersedia dan jarak mengantar dengan rute pengemudi. Insinyur telah mengumpulkan 25 pengamatan pada waktu pengiriman yang ditunjukkan pada tabel berikut:

Tabel 2.1 Data Waktu Pengiriman

No	Waktu Pengiriman (menit)	Jumlah Kasus	Jarak (*feet*)	No	Waktu Pengiriman (menit)	Jumlah Kasus	Jarak (*feet*)
1	16.68	7	560	14	19.75	6	462
2	11.50	3	220	15	24.00	9	448
3	12.03	3	340	16	29.00	10	776
4	14.88	4	80	17	15.35	6	200
5	13.75	6	150	18	19.00	7	132
6	18.11	7	330	19	9.50	3	36
7	8.00	2	110	20	35.10	17	770
8	17.83	7	210	21	17.90	10	140
9	79.24	30	1460	22	52.32	26	810
10	21.50	5	605	23	18.75	9	450
11	40.33	16	688	24	19.83	8	635
12	21.00	10	215	25	10.75	4	150
13	13.50	4	255				

Sumber: (Montgomery, 1982)

Nilai β_0, β_1, dan β_2 dihitung dengan menggunakan matriks. Matriks $\boldsymbol{X}$ dan vektor $\boldsymbol{Y}$ untuk contoh kasus di atas adalah sebagai berikut:

$$\boldsymbol{X} = \begin{pmatrix} 1 & 7 & 560 \\ 1 & 3 & 220 \\ 1 & 3 & 340 \\ \vdots & \vdots & \vdots \\ 1 & 4 & 150 \end{pmatrix} \qquad \boldsymbol{Y} = \begin{pmatrix} 16.68 \\ 11.50 \\ 12.03 \\ \vdots \\ 10.75 \end{pmatrix}$$

Selanjutnya dihitung matriks $\boldsymbol{X'X}$ dan $\boldsymbol{X'Y}$ sebagai berikut:

$$\boldsymbol{X'X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 7 & 3 & \cdots & 4 \\ 560 & 220 & \cdots & 150 \end{bmatrix} \begin{bmatrix} 1 & 7 & 560 \\ 1 & 3 & 220 \\ \vdots & \vdots & \vdots \\ 1 & 4 & 150 \end{bmatrix}$$

$$= \begin{bmatrix} 25 & 219 & 10232 \\ 219 & 3055 & 133899 \\ 10232 & 133899 & 6725688 \end{bmatrix}$$

$$\mathbf{X'Y} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 7 & 3 & \cdots & 4 \\ 560 & 220 & \cdots & 150 \end{bmatrix} \begin{bmatrix} 16.68 \\ 11.50 \\ \vdots \\ 10.75 \end{bmatrix} = \begin{bmatrix} 559.60 \\ 7375.44 \\ 337072.00 \end{bmatrix}$$

Estimator dari $\boldsymbol{\beta}$ adalah $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X'X})^{-\mathbf{1}}\boldsymbol{X'Y}$ atau:

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 25 & 219 & 10232 \\ 219 & 3055 & 133899 \\ 10232 & 133899 & 6725688 \end{bmatrix}^{-1} \begin{bmatrix} 559.60 \\ 7375.44 \\ 337072.00 \end{bmatrix}$$

$$= \begin{bmatrix} 0.11321518 & -0.00444859 & -0.00008367 \\ -0.00444859 & 0.00274378 & -0.00004786 \\ -0.00008367 & -0.00004786 & 0.00000123 \end{bmatrix} \begin{bmatrix} 559.60 \\ 7375.44 \\ 337072.00 \end{bmatrix}$$

$$= \begin{bmatrix} 2.34123115 \\ 1.61590712 \\ 0.01438483 \end{bmatrix}$$

Berdasarkan perhitungan tersebut, maka model yang terbentuk adalah:

$$\hat{Y} = 2.34123115 + 1.61590712X_1 + 0.01438483X_2$$

Sifat-sifat estimasi kuadrat terkecil yang sering disebut *BLUE* (*Best Linear Unbias Estimator*), (Gujarati, 2004):

a. Linier

Estimator bersifat linier, yaitu fungsi linier dari variabel acak, seperti halnya variabel dependen *Y* dalam model regresi. Dalam persamaan estimator $\hat{\beta}$ merupakan estimator linier karena merupakan fungsi linier dari *Y*, yaitu:

$$\hat{\beta} = \frac{\sum_{i=1}^{n} Y_i(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \sum_{i=1}^{n} l_i Y_i \tag{2.9}$$

dengan

$l_i = \frac{(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$, untuk i = 1, 2, ..., n

b. Tak Bias

Tak bias yaitu nilai ekspektasi β, $E(\hat{\beta})$**,** adalah nilai β yang sebenarnya. Hal tersebut dibuktikan dengan subtitusi persamaan (2.1) ke dalam persamaan (2.9) sebagai berikut:

$$\hat{\beta} = \sum_{i=1}^{n} l_i(\beta_0 + \beta_1 X_i + \varepsilon_i)$$
$$= \beta_i + \sum_{i=1}^{n} li\ \varepsilon_i \qquad (2.10)$$

Nilai eskpektasi $\hat{\beta}$ adalah sebagai berikut:

$$E(\hat{\beta}) = \beta_i + \sum_{i=1}^{n} li\ \varepsilon_i$$
$$E(\hat{\beta}) = \beta + \sum_{i=1}^{n} li\ (0)$$
$$= \beta \qquad (2.11)$$

Persamaan (2.11) menunjukkan bahwa estimasi kuadrat terkecil bersifat tak bias.

c. Variansi Minimum

Estimasi yang efisien yaitu suatu estimasi tak bias dengan variansi terkecil. Dengan digunakannya definisi variansi, dapat ditunjukkan bahwa estimasi kuadrat terkecil menghasilkan variansi yang minimum.

$$var(\hat{\beta}) = E\left[\hat{\beta} - E(\hat{\beta})\right]^2$$
$$= E\left(\hat{\beta} - \beta\right)^2$$
$$= E(\beta_i + \sum_{i=1}^{n} li\ \varepsilon_i\ - \beta)^2$$
$$= E(\sum_{i=1}^{n} li\ \varepsilon_i\ )^2$$
$$= \sum_{i=1}^{n} {l_i}^2\ E[{\varepsilon_i}^2] \qquad (2.12)$$

Karena $E[{\varepsilon_i}^2] = \sigma^2$ untuk setiap *i* dan $E\left[\varepsilon_i \varepsilon_j\right] = 0, i \neq j$ maka,

$$var(\hat{\beta}) = \sigma^2 \sum_{i=1}^{n} {l_i}^2$$
$$= \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \text{ (menggunakan definisi } {l_i}^2\text{)} \qquad (2.13)$$

Misalkan suatu estimator linier (β) sebagai berikut:

$$\beta^* = \sum_{i=1}^{n} v_i Y_i \qquad (2.14)$$

dengan v_i tidak perlu sama dengan k_i, maka:

$$
\begin{aligned}
E(\beta^*) &= \sum_{i=1}^{n} v_i E(Y_i) \\
&= \sum_{i=1}^{n} v_i (\beta_0 + \beta_1 X_i) \\
&= \beta_0 \sum_{i=1}^{n} v_i + \beta_1 \sum_{i=1}^{n} v_i X_i
\end{aligned} \tag{2.15}
$$

Supaya β^* tak bias, maka $\sum_{i=1}^{n} v_i = 0$ dan $\sum_{i=1}^{n} v_i X_i = 1$, hasilnya adalah sebagai berikut:

$$
\begin{aligned}
var(\beta^*) &= var\ \sum_{i=1}^{n} v_i Y_i \\
&= \sum_{i=1}^{n} {v_i}^2 var\ Y_i \\
&= \sigma^2 \sum_{i=1}^{n} {v_i}^2\ = \sigma^2 \sum_{i=1}^{n} \left(v_i - \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} + \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \right)^2 \\
&= \sigma^2 \sum_{i=1}^{n} \left(v_i - \frac{(x_i - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} + \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \right) + \sigma^2 \left(\frac{1}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \right) \\
&= \sigma^2 + var(\hat{\beta})
\end{aligned} \tag{2.16}
$$

Hasil akhir dari persamaan (2.16) menunjukkan bahwa $var(\hat{\beta}) \leq var(\hat{\beta}^*)$, sehingga estimator tak bias memiliki variansi minimum.

Analisis regresi memiliki asumsi-asumsi yang harus dipenuhi. Asumsi tersebut diperlukan supaya dapat menilai kebaikan suatu persamaan regresi (Sembiring, 2003). Hasil estimasi regresi yang didapatkan dari MKT merupakan hasil estimasi dengan sifat *BLUE* apabila asumsi-asumsinya terpenuhi. Asumsi-asumsi tersebut disebut dengan asumsi klasik, yaitu (Algifari, 1997):

a. Normalitas

Analisis regresi linier mengasumsikan bahwa *error* berdistribusi normal dengan rata-rata 0 dan variansi σ^2 atau dapat ditulis $\varepsilon \sim N(0, \sigma^2)$ (Gujarati, 2004). Salah satu metode yang digunakan untuk uji normalitas adalah Shapiro-Wilk, yang pada dasarnya adalah mengkuadratkan korelasi antara urutan statistik yang diamati dan urutan statistik yang diharapkan. Normalitas akan ditolak jika nilai *W* terlalu kecil (Weisberg, 2005). Uji Shapiro-Wilk digunakan untuk banyaknya data yang sedikit, yaitu kurang atau sama dengan 50. Sedangkan untuk banyaknya data yang lebih dari 50, digunakan uji Kolmogorov-Smirnov (Dahlan, 2008).

b. Homoskedastisitas

Homoskedastisitas adalah kondisi variansi dari variabel prediktor mempunyai nilai yang tidak bertambah dan tidak berkurang (konstan) untuk setiap nilai variabel prediktor tersebut. Adanya heteroskedastisitas menyebabkan estimasi regresi menjadi tidak efisien, baik dalam sampel kecil maupun sampel besar (Algifari, 1997). Menurut Gujarati (2004) pemeriksaan asumsi ini dapat dilakukan dengan berbagai uji, salah satunya pengujian GoldFeld-Quandt.

c. Non-multikolinearitas

Asumsi non-multikolinearitas diterapkan untuk analisis regresi yang memiliki dua atau lebih variabel prediktor. Non-multikoliniertas berarti antara variabel prediktor tersebut tidak memiliki hubungan atau korelasi. Apabila terjadi multikolinieritas antar variabel prediktor tersebut, maka model regresi tidak valid dalam melakukan pendugaan nilai variabel prediktor (Sunyoto, 2007).

Multikolinieritas dapat diperiksa dengan melakukan pengujian terhadap nilai *Variance Inflation Factor* (*VIF*). Apabila nilai *VIF* lebih dari 10, maka asumsi tidak terpenuhi, berarti terdapat multikolinearitas pada model regresi tersebut (Ghozali, 2005).

Nilai *VIF* dihitung dengan menggunakan persamaan berikut:

$$VIF_j = \frac{1}{1-R_j^2} \qquad (2.17)$$

dengan

j : Jumlah variabel prediktor; $j = 1, 2, \dots, k$

R_j^2 : Koefisien determinasi dari variabel prediktor X_j dengan variabel prediktor lain

Adanya multikolinieritas dalam regresi dapat diatasi dengan menggunakan beberapa cara, yaitu sebagai berikut (Sunyoto, 2007):

i. Menghilangkan salah satu atau lebih variabel prediktor yang memiliki korelasi tinggi.

ii. Variabel prediktor yang memiliki korelasi tinggi dapat tetap digunakan, namun hanya untuk membantu memprediksi dan tidak diinterpretasikan.
iii. Hubungan linier antar variabel prediktor dikurangi dengan menggunakan logaritma natural (*ln*).
iv. Menggunakan metode lain dalam melakukan analisis, seperti regresi *ridge*.

d. Non-Autokorelasi

Non-autokorelasi yaitu tidak terdapat korelasi antar pengamatan dalam satu variabel prediktor. Apabila asumsi non-autokorelasi tidak terpenuhi, maka variansi sampel tidak dapat menggambarkan variansi populasi (representatif), sehingga model regresi tidak dapat digunakan untuk menduga nilai variabel dependen. Asumsi non-autokorelasi dapat diperiksa dengan menggunakan uji Durbin-Watson. Statistik uji Durbin-Watson mempunyai (*d*) dengan ketentuan sebagai berikut (Gujarati, 2004):

1. Jika $d < d_L$ atau $d > 4 - d_L$, maka H_0 ditolak, berarti terdapat autokorelasi.
2. Jika $d_U < d < 4 - d_U$, maka H_0 gagal tolak, berarti tidak terdapat autokorelasi.
3. Jika $d_L < d < d_U$ atau $4 - d_U < d < 4 - d_L$, maka maka tidak dapat diputuskan apakah H_0 gagal tolak atau ditolak, berarti tidak dapat disimpulkan ada atau tidak ada autokorelasi.

Nilai *d* dihitung dengan menggunakan rumus berikut:

$$d = \frac{\sum_{i=2}^{n}(e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2} \tag{2.18}$$

Nilai d_L dan d_U ditentukan berdasarkan tabel Durbin-Watson.

Misalkan X adalah variabel random dan nilai harapan dari X adalah $E(X) = \mu$ maka ditribusi dari nilai X disekitar rata-ratanya disebut dengan variansi. Variansi dapat didefinisikan sebagai berikut:

$$Var(X) = \sigma_x^2 = E(X - \mu)^2 \quad (2.19)$$

Menggunakan variansi bisa didapatkan standar deviasi dari X, yaitu akar dari variansi X. Variansi dan standar deviasi menunjukkan seberaba dekat data X dengan rata-ratanya.

Salah satu asumsi dalam MKT yang harus terpenuhi adalah variansi *error* yang konstan. Namun pada kenyataannya, seringkali ditemukan variansi *error* yang tidak konstan atau sering disebut dengan heteroskedastisitas. Jika tetap menggunakan MKT pada model regresi linier yang mengandung heteroskedastisitas dan memenuhi asumsi yang lain, maka estimator dari β_1 adalah sebagai berikut:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} \quad (2.20)$$

dengan :

$x_i = X_i - \bar{X}$

$y_i = Y_i - \bar{Y}$

Dalam persamaan tersebut estimator $\hat{\beta}_1$ bersifat linier dan tidak bias. Apabila varian *error* homoskedastisitas, maka persamaannya adalah sebagai berikut:

$$var(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{n} x_i^2} \quad (2.21)$$

Namun jika variansi *error* heteroskedastisitas, maka persamaannya adalah sebagai berikut:

$$var(\hat{\beta}_1) = \frac{\sum_{i=1}^{n} x_i \sigma_i^2}{(\sum_{i=1}^{n} x_i^2)^2} \quad (2.22)$$

Jadi, jika tetap digunakan MKT estimator $\hat{\beta}_1$ tidak lagi mempunyai variansi yang minimum (Widarjono, 2005).

Apabila MKT diterapkan pada data yang mengandung heteroskedastisitas, maka akan berakibat pada hal-hal berikut (Qudratullah, 2013):

a. Penaksiran tetap *unbiased*, karena asumsi heteroskedastisitas tidak digunakan dalam membuktikan sifat *unbiased* parameter. Jadi, hal ini bisa dimaklumi.
b. Variansi estimator akan salah, karena variansi estimator tersebut akan berubah-ubah nilainya atau tidak tetap.
c. Estimator-estimator menjadi tidak efisien. Meskipun menggunakan MKT, estimator yang dihasilkan bersifat *unbiased* tetapi memiliki variansi yang lebih besar dari estimator lain yang bersifat bias.
d. Prediksi terhadap koefisien populasi tidak akan tepat atau keliru.

Pendeteksian heteroskedastisitas dapat dilakukan dengan grafik atau dengan menggunakan uji hipotesis supaya lebih akurat. Salah satu metode yang digunakan untuk melakukan pendeteksian heteroskedastisitas Metode GoldFeld-Quandt yang dikembangkan oleh GoldFeld-Quandt. Uji tersebut lebih baik digunakan pada sampel yang besar. Langkah-langkah pengujiannya adalah sebagai berikut (Widarjono, 2005):

1. Pengamatan – pengamatan disusun berdasarkan nilai *X*.
2. Menghilangkan beberapa pengamatan yang ada di tengah, kemudian jumlah pengamatan sisanya dibagi menjadi dua kelompok dengan satu kelompok tersusun dari nilai *X* yang kecil, sedangkan kelompok lainnya terdiri dari nilai *X* yang besar.
3. Melakukan analisis regresi secara terpisah pada kedua kelompok tersebut.
4. Setelah didapatkan nilai Jumlah Kuadrat *Error* (JKE) pada masing-masing kelompok, hitung perbandingan :

$$F_{\text{hitung}} = \frac{\text{JKE}_2/\text{db}_2}{\text{JKE}_1/\text{db}_1} \tag{2.23}$$

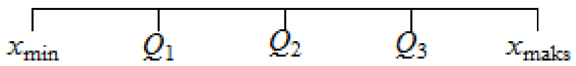
dengan :

JKE_1 : Jumlah Kuadrat *Error* kelompok *X* kecil
JKE_2 : Jumlah Kuadrat *Error* kelompok *X* besar
db_i : Derajat kebebasan kelompok ke-*i*

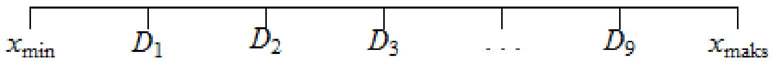
5. Membandingkan nilai F_{hitung} dengan $F_{tabel} = (\alpha;\ v_1;\ v_2)$. Data mengandung unsur heteroskedastisitas apabila F_{hitung} lebih besar dari F_{tabel}.

2.1.2 Kuantil

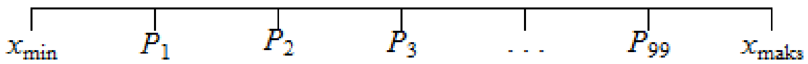
Kuantil adalah nilai-nilai yang membagi sederet data yang telah diurutkan menjadi bagian-bagian yang sama. Kuantil yang membagi data terurut menjadi dua bagian disebut median, menjadi empat bagian disebut kuartil (Q_1, Q_2, Q_3), menjadi sepuluh bagian disebut desil (D_1, D_2, ... D_9), dan menjadi seratus bagian disebut persentil (P_1, P_2, ... P_{99}) (Harinaldi, 2005).



Gambar 2.1 Kuartil



Gambar 2.2 Desil



Gambar 2.3 Persentil

dengan:

x_{min} : Data terkecil

x_{maks} : Data terbesar

Q_i : Kuartil ke-*i*

D_i : Desil ke-*i*

P_i : Persentil ke-*i*

2.1.3 Median

Median atau nilai tengah adalah tengah dari sekumpulan data yang telah diurutkan sebelumnya (Diaz, 2007). Bila X_1, X_2, ..., X_n menyatakan sampel ukuran acak n, diurutkan dari kecil ke besar, maka nilai median dapat ditentukan dengan (Walpole dan Myers, 1995):

$$\tilde{X} = \begin{cases} X_{\left(\frac{n+1}{2}\right)} & ; \text{bila } n \text{ ganjil} \\ \frac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2} & ; \text{bila } n \text{ genap} \end{cases} \quad (2.24)$$

2.1.4 *Loss Function* Asimetrik

Jika L_p merupakan *loss function* asimetrik ke-*p*, maka (Koenker, 2005):

$$\begin{aligned} L_p &= [pI(u \geq 0) + (I - p)I(u < 0)]|u| \\ &= [p - I(u < 0)]u \end{aligned} \quad (2.25)$$

Sehingga diperoleh $L_p = \begin{cases} pu & , \text{jika } u \geq 0 \\ (p-1)u & , \text{jika } u < 0 \end{cases}$

dengan:

u : *Error* dari estimasi

$I(u)$: Fungsi indikator yang didefinisikan

$$I(u) = \begin{cases} 1 & , \text{jika } u \geq 0 \\ 0 & , \text{jika } u < 0 \end{cases}$$

2.1.5 Fungsi Kepadatan Bersyarat

Fungsi kepadatan bersyarat dari *Y* dengan syarat *X*=*x*, didefinisikan dengan (Walpole dan Myers, 1995):

$$f_Y(y|X = x) = \frac{f_{X,Y}(x,y)}{f_X(x)} \quad (2.26)$$

dengan :

$f_{X,Y}(x,y)$: Fungsi kepadatan bersama dari *X* dan *Y*

$f_X(x)$: Fungsi kepadatan marginal untuk *X*

Nilai harapan bersyarat dari *Y* dengan syarat *X*=*x*, didefinisikan dengan:

$$E(Y|X = x) = \int_{-\infty}^{\infty} y f_{Y|X}(y,x) dy \quad (2.27)$$

2.1.6 Regresi Kuantil Median

Tahun 1978 Roger Koenker dan Gilbert Basset dalam (Furno, 2013) memperkenalkan metode regresi kuantil. Regresi kuantil digunakan dengan membagi atau memisahkan data menjadi dua bagian atau lebih dimana dicurigai terdapat perbedaan nilai estimator pada kuantil-kuantil tertentu. Regresi kuantil dapat dituliskan dengan persamaan berikut (Wardani, 2014):

$$Y_{i,\tau} = \beta_{0,\tau} + \beta_{1,\tau}X_{i1} + \cdots + \beta_{p,\tau}X_{ip} + \varepsilon_{i,\tau} \tag{2.28}$$

dengan:

$Y_{i,\tau}$: Nilai pengamatan ke-i pada kuantil ke-τ

X_{ip} : Nilai pengamatan ke-i variabel prediktor ke-p

β_{τ} : Estimator parameter pada kuantil ke-τ

$\varepsilon_{i,\tau}$: *Error* ke-i dan kuantil ke-τ

i = 1, 2, ..., n; j = 1, 2, ..., p; dan $0 < \tau < 1$

Regresi kuantil dilakukan dengan meminimumkan jumlah nilai mutlak dari *error* yang merupakan minimum penjumlahan *error* positif dan *error* negatif. Hal tersebut memberikan perbedaan bobot τ untuk *error* positif dan pembobot $1 - \tau$ untuk *error* negatif (Wardani, 2014). Pada regresi kuantil median dapat mendefinisikan median sebagai solusi untuk meminimumkan jumlah nilai mutlak dari *error* (Uthami, 2013).

Keuntungan utama dari regresi kuantil dibandingkan dengan MKT adalah fleksibilitas ketika memodelkan data dengan sebaran bersyarat yang heterogen. Regresi kuantil mampu untuk mengukur efek variabel prediktor tidak hanya di pusat sebaran data tetapi juga pada bagian atas atau bawah ekor sebaran (Djuraidah dan Wigena, 2011). Menurut Lee (2011) dalam Diaz (1819) penggunaan regresi kuantil median untuk estimasi pada data dengan heteroskedastisitas lebih efisien.

2.1.7 Pengujian Signifikansi Parameter

Pengujian signifikansi parameter dibagi menjadi dua, yaitu uji *overall* (serentak) dan uji *partial* (individu) (Montgomery, 1982):

a. Uji *Overall*

Uji *overall* atau uji serentak atau uji F merupakan pengujian untuk mengetahui ada atau tidaknya pengaruh secara bersama-sama variabel prediktor terhadap variabel respon. Hipotesis dalam pengujian ini adalah sebagai berikut:

$H_0 : \beta_1 = \beta_2 = ... = \beta_k = 0$

H_1 : Minimal ada satu $\beta_j \neq 0$ untuk $j = 0, 1,..., k$

Tabel untuk perhitungan adalah sebagai berikut:

Tabel 2.2 Tabel Anova

Sumber Variansi	Derajat Bebas	Jumlah Kuadrat	Kuadrat Tengah	F_{hitung}
Regresi	1	$JKR = \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2$	$KTR = \frac{JKR}{1}$	$F_{\text{hitung}} = \frac{KTR}{KTE}$
Error	$n-2$	$JKE = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$	$KTE = \frac{JKE}{n-2}$	
Total	$n-1$	$JKT = \sum_{i=1}^{n}(Y_i - \bar{Y})^2$		

Sumber: Pangesti dan Soejoeti, 1987

Dasar pengambilan keputusannya yaitu apabila $F_{\text{hitung}} > F_{\text{tabel}}$ maka H_0 ditolak, artinya minimal ada satu β_j yang tidak sama dengan nol. Selain menggunakan F_{hitung}, pengambilan keputusan dapat dilakukan dengan menggunakan *P-value*. Dalam hal ini H_0 ditolak jika *P-value* $< \alpha$.

b. Uji *Partial*

Uji *partial* atau uji t adalah pengujian secara sendiri-sendiri pada parameter. Tujuan dari uji *partial* ini adalah untuk mengetahui adanya pengaruh antara variabel prediktor ke-j dengan j=1, 2, ..., k dengan variabel respon. Hipotesis yang digunakan dalam pengujian ini adalah sebagai berikut:

$H_0 : \beta_j = 0$

$H_1 : \beta_j \neq 0$ untuk j=0, 1, ...,k

Statistik uji yang digunakan adalah sebagai berikut:

$$t_{\text{hitung}} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \tag{2.29}$$

dengan $SE(\hat{\beta}_j)$ adalah nilai *standar error* dari $\hat{\beta}_j$.

Dasar pengambilan keputusannya yaitu apabila $|t_{\text{hitung}}| > t_{(1-(\alpha/2),\, n-k-1)}$ dengan k adalah parameter maka tolak H_0, berarti ada pengaruh x_i terhadap model. Selain menggunakan t_{hitung}, pengambilan keputusan dilakukan dengan menggunakan *P-value*. Dalam hal ini H_0 ditolak jika *P-value* $< \alpha$.

2.1.8 Standar *Error* Estimasi

Standar *error* estimasi adalah deviasi standar yang memberikan ukuran penyebaran nilai-nilai yang teramati di sekitar garis regresi. Standar *error* estimasi dirumuskan sebagai berikut (Harinaldi, 2005):

$$S_{y,x} = \sqrt{\frac{\sum_{i=1}^{n}(y-\hat{y})^2}{n-2}} = \sqrt{\frac{\sum_{i=1}^{n}(y)^2 - \beta_0(\sum_{i=1}^{n} y) - \beta_1(\sum_{i=1}^{n} xy)}{n-2}} \tag{2.30}$$

dengan:

y : Variabel respon

$\hat{y}$: Estimasi variabel respon

n : Jumlah data

β_0 : *Slope*

β_1 : Intersep

2.1.9 Kuadrat Tengah *Error* (KTE)

Dalam memilih metode yang baik dalam melakukan estimasi didasarkan pada KTE. KTE adalah salah satu ukuran statistik yang digunakan untuk membandingkan metode-metode estimasi yang digunakan, yaitu untuk menentukan estimasi yang paling mendekati data aktual. Perhitungan KTE dilakukan dengan mengkuadratkan masing-masing nilai *error* kemudian menjumlahkannya, setelah itu dibagi dengan jumlah pengamatan. Persamaan KTE adalah sebagai berikut (Makridakis, 1999):

$$\text{KTE} = \sum_{i=1}^{n} e_i^2 \tag{2.31}$$

dengan

$$e_i^2 = (y_i - \hat{y}_i)^2 \tag{2.32}$$

BAB III
METODOLOGI PENELITIAN

3.1 Data

Data yang digunakan dalam penelitian ini ada dua macam, yaitu data yang homoskedastisitas dan data yang heteroskedastisitas. Data yang homoskedastisitas adalah data balita penderita gizi buruk (Y) dalam rasio dan bayi yang lahir dengan berat badan rendah (X) dalam rasio pada setiap kabupaten/kota di Provinsi Jawa Timur tahun 2012. Data tersebut diperoleh dari Dinas Kesehatan Provinsi Jawa Timur.

Sedangkan data yang heteroskedastisitas adalah data pengeluaran konsumsi (Y) dengan satuan dolar dan pendapatan (X) dengan satuan dolar. Data tersebut diperoleh dari buku "*Basic Econometrics*" (Gujarati, 2004).

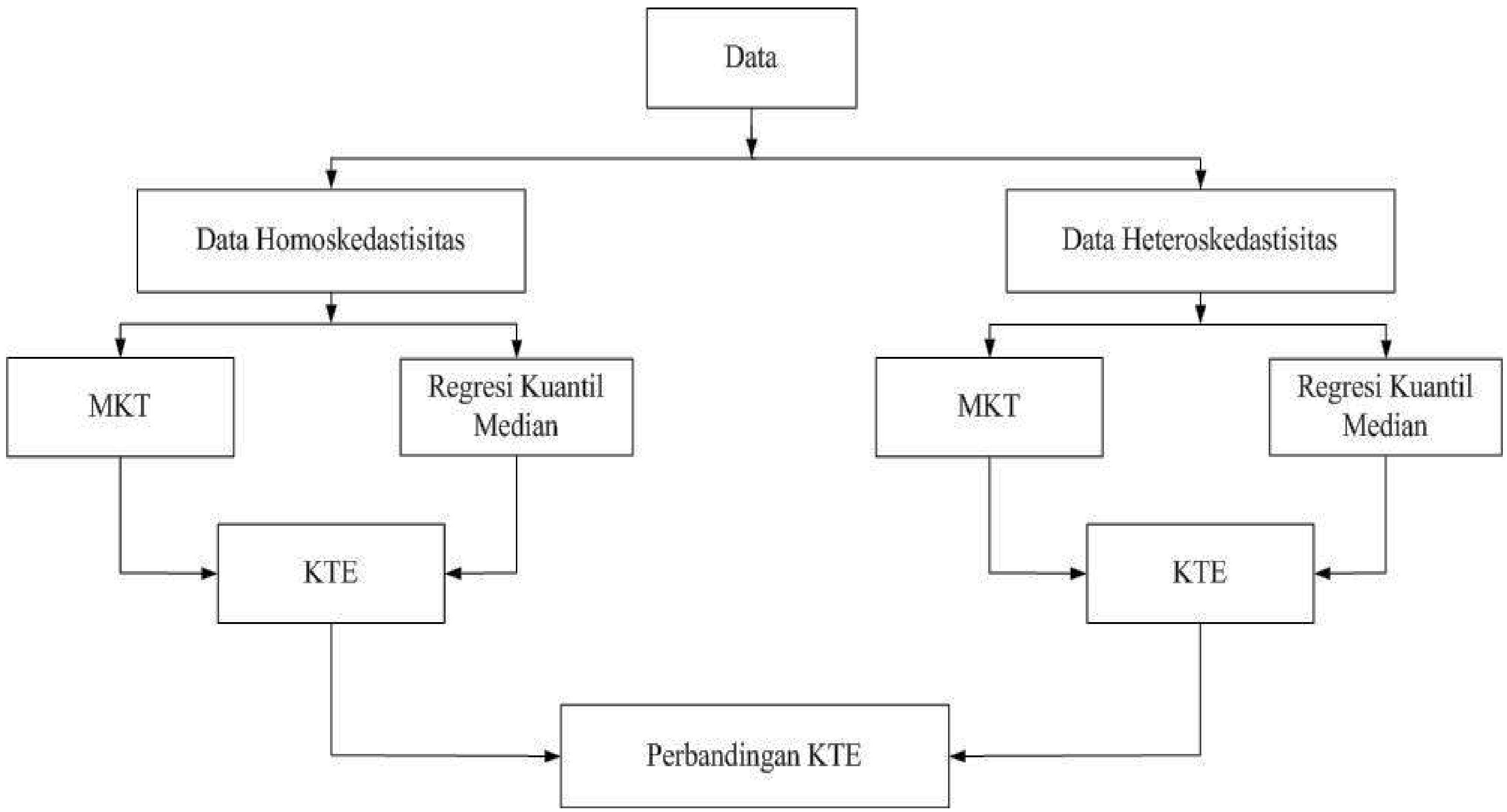
3.2 Tahapan Analisis Data

Pengujian heteroskedastisitas pada data dilakukan dengan menggunakan aplikasi Microsoft Excel, sedangkan dalam melakukan estimasi dengan MKT dan regresi kuantil median dilakukan dengan menggunakan *software R* 3.2.3. Penelitian ini dilakukan dengan tahapan-tahapan sebagai berikut:

1. Tahapan estimasi data homoskedastisitas
   a. Memasukkan data balita penderita gizi buruk (Y) dan bayi yang lahir dengan berat badan rendah (X) pada setiap kabupaten di Provinsi Jawa Timur tahun 2012 ke *software R* 3.2.3.
   b. Melakukan estimasi regresi dengan MKT.
   c. Mendeteksi heteroskedastisitas pada residual data dengan menggunakan uji GoldFeld-Quandt pada *Microsoft Excel.*
   d. Melakukan estimasi dengan regresi kuantil pada $0 < \tau < 1$.
   e. Mendeteksi heteroskedastisitas pada residual data dengan menggunakan uji GoldFeld-Quandt pada *Microsoft Excel.*

f. Menghitung dan membandingkan KTE.

2. Tahapan estimasi data heteroskedastisitas
   a. Memasukkan data pengeluaran konsumsi (Y) dan pendapatan (X) ke *software R* 3.2.3.
   b. Melakukan estimasi regresi dengan MKT.
   c. Mendeteksi heteroskedastisitas pada residual data dengan menggunakan uji GoldFeld-Quandt pada *Microsoft Excel.*
   d. Melakukan estimasi dengan regresi kuantil pada $0 < \tau < 1$.
   e. Mendeteksi heteroskedastisitas pada residual data dengan menggunakan uji GoldFeld-Quandt pada *Microsoft Excel.*
   f. Menghitung dan membandingkan KTE



Gambar 3.1 Diagram Alir Tahapan Analisis Data

BAB IV
PEMBAHASAN

MKT adalah salah satu cara untuk melakukan estimasi parameter dalam model regresi. Estimasi parameter dalam MKT dilakukan dengan meminimumkan jumlah kuadrat *error*. MKT memiliki beberapa asumsi yang harus terpenuhi supaya mendapatkan estimator yang bersifat *BLUE*. Salah satu asumsi tersebut adalah homoskedastisitas, yaitu variansi *error* harus memiliki nilai yang konstan. Namun pada kenyataannya tidak semua *error* memiliki nilai yang konstan sehingga asumsi homoskedastisitas tidak terpenuhi. Tidak terpenuhinya asumsi tersebut menyebabkan variansi yang diperoleh menjadi tidak efisien dan dapat mengganggu model yang akan diestimasi. Maka dari itu diperlukan penanganan untuk mengatasi masalah homoskedastisitas. Karena ketika asumsi homoskedastisitas tidak terpenuhi estimasi dengan MKT menjadi tidak tepat. Metode lain yang dapat digunakan untuk melakukan estimasi ketika terdapat masalah heteroskedastisitas adalah metode regresi kuanti median.

4.1 Regresi Kuantil

Regresi kuantil adalah salah satu metode regresi yang sering digunakan dalam permasalahan ekonometrika. Regresi kuantil dilakukan dengan membagi atau mengelompokkan data yang telah diurutkan sehingga lebih mudah menentukan letaknya dan dapat mendefinisikan kuantil dengan alternatif sederhana sebagai masalah optimasi.

Regresi kuantil median dalam model dengan sampel acak $\{y_1, y_2, \dots, y_n\}$ didefinisikan dengan menggunakan:

$$min_{\mu \epsilon R} \sum_{i=1}^{n} (y_i - \mu)^2 \qquad (4.1)$$

Jika $\mu = \boldsymbol{X}'\boldsymbol{\beta}$ maka persamaan (5.1) menjadi:

$$\hat{\beta} = \arg min_{\beta \epsilon R^p} \sum_{i=1}^{n} (\boldsymbol{y_i} - \boldsymbol{X_i'\beta})^2 \quad (4.2)$$

dengan:

$\boldsymbol{X_i}$: Variabel prediktor ke-i

$\boldsymbol{\beta}$: Parameter

$\boldsymbol{y_i}$: Variabel respon ke-i

Menurut Koenker dan Basset (1978) dalam penelitiannya yang membahas masalah regresi tersebut berkembang menjadi median sampel dinyatakan dalam persamaan:

$$min_{\beta \epsilon R} \sum_{i=1}^{n} |\boldsymbol{y_i} - \boldsymbol{X_i'\beta}| \quad (4.3)$$

Kemudian secara umum dispesifikasikan dalam regresi kuantil bersyarat ke-τ dengan mempertimbangkan estimator bagi $\beta(\tau)$ yaitu $(\hat{\beta}(\tau))$ sehingga diperoleh ide bahwa masalah tersebut dapat dinyatakan dengan menggunakan persamaan:

$$min_{\beta \epsilon R} \sum_{i=1}^{n} \rho_\tau (y_i - Q_\tau(Y|X)) \quad (4.4)$$

dengan:

τ : Indeks kuantil $\in (0,1)$

$\rho_\tau(.)$: *Loss function* yang asimetrik

$Q_\tau(Y|X) = \boldsymbol{X'\beta(\tau)}$: Fungsi kuantil ke-τ dari *Y* dengan syarat *X*

Jika *Y* merupakan sebaran variabel acak kontinu dan *x* adalah salah satu vektor regresor *X*, fungsi kuantil bersyarat dalam fungsi kuantil ke-τ didefinisikan dengan:

$$Q_\tau(Y|X) = \inf\{y: F_y(y|X) \geq \tau\} \quad (4.5)$$

dengan:

$F_y(y|X)$: Fungsi sebaran dari *Y* dengan syarat *X* dan fungsi kepadatan bersyaratnya $F_y(y|X)$.

Regresi kuantil dapat terpenuhi dengan mengganti suatu model linier $q(X)$ pada persamaan (5.5) sehingga diperoleh masalah minimisasi berikut:

$$\beta(\tau) = arg\ min_{\beta \in R^d} E[\boldsymbol{\rho_\tau}(\boldsymbol{Y} - \boldsymbol{X_i'\beta})] \tag{4.6}$$

Jika dalam MKT pendekatan dilakukan dengan meminimumkan nilai harapan kuadrat *error*, maka dalam regresi kuantil pendekatan dilakukan dengan meminimumkan nilai harapan *loss function* yang asimetrik, yaitu dengan meminimumkan nilai harapan $\rho_\tau(y)$. Jika *loss function* didefinisikan sebagai $\rho_\tau(y) = |y(\tau - I_{(y<0)})|$ dengan I adalah fungsi indikator. Maka nilai kuantil ke-τ dengan meminimumkan fungsi terhadap *u* :

$$\begin{aligned} min_u E(\rho_\tau(Y-u)) &= min_u(\tau - 1)\int_{-\infty}^{u}(y-u)dF(y) + \tau\int_{u}^{\infty}(y-u)dF(y) \\ 0 &= (1-\tau)\int_{-\infty}^{q_\tau} dF(y) - \tau\int_{q_\tau}^{\infty} dF(y) \\ 0 &= (1-\tau)F(q_\tau) - \tau[1 - F(q_\tau)] \\ 0 &= F(q_\tau) - \tau \\ F(q_\tau) &= \tau \end{aligned} \tag{4.7}$$

Sehingga kuantil ke-τ merupakan solusi dari *F*.

Jika *X* merupakan fungsi dari *Y* yang telah diketahui dan mempunyai peluang *F*(*y*), maka kuantil ke-τ dari fungsi tersebut dapat ditulis sebagai berikut:

$$\begin{aligned} q_\tau &= Q(\tau) \\ &= F^{-1}(\tau) \end{aligned} \tag{4.8}$$

$Q(\tau)$ merupakan fungsi dari *Y* yang dapat diselesaikan dengan menggunakan persamaan sebagai berikut:

$$min(\tau - 1)\int_{-\infty}^{q_\tau}(y - q_\tau)dF(y) + \tau\int_{q_\tau}^{\infty}(y - q_\tau)dF(y) \tag{4.9}$$

Q(0,5) adalah median *X* yang menunjukkan titik simetri dari *F*. Dalam notasi matriks jika $Q(\tau)$ adalah fungsi linier $\boldsymbol{X'\beta}$ yang dinotasikan dengan $\xi(\tau)$, maka persamaan (4.9) menjadi:

$$min(\tau - 1)\int_{-\infty}^{\boldsymbol{X'\beta}}(\boldsymbol{y} - \boldsymbol{X'\beta})dF(y) + \tau\int_{\boldsymbol{X'\beta}}^{\infty}(\boldsymbol{y} - \boldsymbol{X'\beta})dF(y) \qquad (4.10)$$

Solusi dari persamaan (4.10) dinotasikan dengan β dan kuantil X ke-τ adalah $\boldsymbol{Q(\tau) = X'\beta}$

Misal diberikan data (y_i, x_i) untuk i = 1, 2, ..., n, maka model linier dari persamaan regresi kuantil dapat dituliskan dengan:

$$\boldsymbol{y_i = X_i'\beta + e_i}$$

dengan $Q_\tau(y_i|x_i) = \boldsymbol{X_i'\beta}$ merupakan kuantil ke-τ dari y dengan suatu nilai x_i tertentu. Estimator β dari regresi kuantil ke-τ diperoleh dengan meminimumkan jumlah nilai mutlak dari *error* dengan pembobot τ untuk *error* positif dan pembobot $1 - \tau$ untuk *error* negatif, yaitu sebagai berikut:

$$\hat{\beta}(\tau) = min_\beta\{\tau\Sigma_{i;y_i \geq x_i}|\boldsymbol{y_i - X_i'\beta}| + (1-\tau)\Sigma_{i;y_i < x_i}|\boldsymbol{y_i - X_i'\beta}|\} \qquad (4.11)$$

$$min_{\xi \in R}\sum_{i=1}^{n}\boldsymbol{\rho_\tau}(\boldsymbol{y_i - \xi}) \qquad (4.12)$$

dengan $\rho_\tau(u) = u(\tau - 1(u < 0))$

Estimasi koefisien variabel prediktor dapat diperoleh dengan memecahkan fungsi sebagai berikut:

$$\hat{\beta}(\tau) = arg\ min_{\xi \in R}\sum_{i=1}^{n}\boldsymbol{\rho_\tau}(\boldsymbol{y_i - X_i'\beta_\tau}) \qquad (4.13)$$

4.2 Sifat Pendekatan Regresi Kuantil

Hasil teori utama dari prinsip regresi kuantil adalah suatu vektor parameter regresi kuantil yang dapat meminimumkan jumlah *error* kuadrat terboboti. Jika indeks kuantil $\tau \in (0,1)$, maka *error* dari regresi kuantil adalah sebagai berikut (Fitriah, 2009):

$$\Delta_{\boldsymbol{\tau}}(\boldsymbol{X}, \boldsymbol{\beta}) = \boldsymbol{X'\beta - Q_\tau(Y|X)} \qquad (4.14)$$

Misalkan ϵ_τ adalah *error* dari regresi kuantil yaitu penyimpangan peubah respon dari kuantil bersyarat, ditulis sebagai berikut:

$$\epsilon_\tau = Y - Q_\tau(Y|X) \tag{4.15}$$

dengan fungsi kepadatan bersyarat $f_{\epsilon_\tau}(e|X)$ pada $\epsilon_\tau = e$.

Teorema berikut menunjukkan bahwa regresi kuantil merupakan pendekatan kuadrat terkecil terboboti dengan seluruh peubah prediktor.

Teorema 1

Asumsi yang digunakan pada teorema ini adalah:

1. Terdapat fungsi kepadatan bersyarat $f_Y(y|X)$
2. $E[Y], E[Q_\tau(Y|X)],$ dan $E\|X\|$ terbatas
3. $\beta(\tau)$ solusi unik bagi persamaan (5.11)

Dari asumsi-asumsi tersebut diperoleh persamaan:

$$\beta(\tau) = argmin_{\beta\epsilon R^d} E[w_\tau(X,\beta).\Delta_\tau^2(X,\beta)] \tag{4.16}$$

dengan:

$\boldsymbol{\beta(\tau)}$: Vektor parameter regresi kuantil

$\Delta_{\boldsymbol{\tau}}^{\mathbf{2}}(\boldsymbol{X},\boldsymbol{\beta})$: Kuadrat *error* dari regresi kuantil

$\boldsymbol{w_\tau}(\boldsymbol{X},\boldsymbol{\beta})$: Fungsi pembobot (*importance weights*)

$= \int_0^1 (1-u).f_{\epsilon_\tau}(u\Delta_\tau(\boldsymbol{X},\boldsymbol{\beta})|\boldsymbol{X})d$

$=\int_0^1 (1-u).\boldsymbol{\Psi} du \geq 0,$

dengan:

$$\boldsymbol{\Psi} = f_Y(u.\boldsymbol{X_i'\beta} + (1-u).Q_\tau(Y|X)|X)$$

Teorema 1 menyatakan bahwa vektor parameter regresi kuantil dapat meminimumkan nilai harapan dari kuadrat tengah terboboti, yaitu kuadrat dari selisih antara fungsi kuantil bersyarat yang sebenarnya dan pendekatan garis linier dengan fungsi pembobot $\boldsymbol{w_\tau}(\boldsymbol{X},\boldsymbol{\beta})$. Fungsi pembobot $\boldsymbol{w_\tau}(\boldsymbol{X},\boldsymbol{\beta})$ adalah fungsi yang menentukan regresi kuantil yang dapat meminimumkan nilai variabel respon dari variabel prediktor yang diberikan. Fungsi pembobot tersebut ditentukan dengan menggunakan (Fitriah, 2009):

$$\boldsymbol{w_\tau}(\boldsymbol{X},\boldsymbol{\beta}) = \frac{1}{2}.\boldsymbol{f_Y}(\boldsymbol{Q_\tau}(\boldsymbol{Y}|\boldsymbol{X})|\boldsymbol{X}) + \boldsymbol{\varrho_\tau}(\boldsymbol{X})$$

dengan:

$$|\boldsymbol{\varrho_\tau}(\boldsymbol{X})| = \left|\int_0^1 (1-u).\, f_{\epsilon_\tau}(\boldsymbol{u}.\, \Delta_{\boldsymbol{\tau}}(\boldsymbol{X},\boldsymbol{\beta})|\boldsymbol{X}) - f_{\epsilon_\tau}(\boldsymbol{0}|\boldsymbol{X})\right| <$$

$$|\Delta_{\boldsymbol{\tau}}(\boldsymbol{X},\boldsymbol{\beta})|\ .\, \bar{\boldsymbol{f}}'(\boldsymbol{X}).\int_0^1 (1-u).\, udu$$

$$= \frac{1}{6}.\, |\Delta_{\boldsymbol{\tau}}(\boldsymbol{X},\boldsymbol{\beta})|\ .\, \bar{\boldsymbol{f}}'(\boldsymbol{X}) \tag{4.17}$$

dengan:

$\boldsymbol{\varrho_\tau}(\boldsymbol{X})$	: *Error*
$\boldsymbol{w_\tau}(\boldsymbol{X},\boldsymbol{\beta})$	: Fungsi pembobot (*importance weights)*
$\boldsymbol{f_Y}(\boldsymbol{y}|\boldsymbol{X})$	: Fungsi kepadatan bersyarat yang diasumsikan mempunyai turunan pertama pada *y* yang terbatas pada nilai mutlaknya
$\bar{\boldsymbol{f}}'(\boldsymbol{X})$	: Turunan pertama pada *y* dari fungsi kepadatan $f_Y(y|X)$

Fungsi kepadatan yang terboboti di dalam beberapa kasus merupakan faktor penentu bagi fungsi pembobot. Dengan $f_Y(\boldsymbol{Q_\tau}(\boldsymbol{Y}|\boldsymbol{X})|\boldsymbol{X})$ adalah nilai tetap dari *X* pada model yang berbanding terbalik dengan simpangan baku bersyaratnya.

4.3 Studi Kasus

4.3.1 Data dengan Homoskedastisitas

a. Estimasi dengan Metode Kuadrat Terkecil pada Data dengan Homoskedastisitas

Langkah pertama yang dilakukan adalah melakukan estimasi dengan MKT pada data balita penderita gizi buruk (%) dan balita lahir dengan berat badan rendah (%). Estimasi dilakukan dengan menggunakan *software R* 3.2.3. Hasil estimasi dengan MKT adalah sebagai berikut:

Tabel 4.1 Hasil Estimasi dengan MKT pada Data dengan Homoskedastisitas

Variabel	Nilai Estimasi	*Standard Error*	*P-value*
(intersep)	0.8983	0.3171	0.0075
X	1.2830	0.2469	8.23E-06

Berdasarkan tabel 4.1 didapatkan model sebagai berikut:

$$Y = 0.8983 + 1.2830X$$

dengan :

Y : Balita penderita gizi buruk (%)

X : Bayi yang lahir dengan berat badan rendah (%)

Model tersebut dapat dijelaskan sebagai berikut:

a. Nilai konstanta sebesar 0.8983, artinya jika bayi lahir dengan berat badan rendah tidak mengalami penambahan ataupun pengurangan maka balita penderita gizi buruk adalah sebesar 0.8983%.

b. Koefisien regresi variabel bayi lahir dengan berat badan rendah sebesar 1.2830, artinya jika bayi lahir dengan berat badan rendah mengalami kenaikan 1 dolar maka banyaknya balita penderita gizi buruk akan mengalami kenaikan sebesar 1.2830%.

Tabel perhitungan analisis variansinya adalah sebagai berikut:

Tabel 4.2 Tabel Analisis Variansi

Sumber Variansi	Derajat Bebas	Jumlah Kuadrat	Kuadrat Tengah	F_{hitung}
Regresi	1	23.6805	23.6805	27.0012
Error	$n-2$	31.5726	0.8770	
Total	$n-1$	55.2539		

Uji hipotesis untuk uji *overall* adalah sebagai berikut :

- Hipotesis:

  H_0 : $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ (Model tidak sesuai)

  H_1 : Minimal ada satu $\beta_i \neq 0$, $i = 1, 2, 3, \ldots, k$ (Model sesuai)

- Tingkat signifikansi:

  $\alpha = 0.05$

- Daerah kritis:

  H_0 ditolak jika $F_{\text{hitung}} > F_{\text{tabel}}$

- Statistik Uji

  $F_{\text{hitung}} = 27.0012 > F_{\text{tabel}} = 4.1132$

- Keputusan

  $F_{\text{hitung}} = 27.0012 > F_{\text{tabel}} = 4.1132$ maka tolak H_0.

- Kesimpulan

  Dengan menggunakan tingkat signifikansi sebesar $\alpha = 0.05$ maka dapat diambil kesimpulan bahwa minimal ada satu $\beta_i \neq 0$ atau dengan kata lain model sesuai.

Untuk melihat kesalahan pengukuran dari persamaan model dalam melakukan estimasi variabel respon digunakan nilai KTE. Nilai KTE yang didapatkan adalah sebesar 0. 8770.

b. Uji GoldFeld-Quandt pada Data dengan Homoskedastisitas

Selanjutnya dilakukan pengujian asumsi heteroskedastisitas. Untuk uji heteroskedastisitas digunakan uji GoldFeld-Quandt. Pengujian dilakukan dengan mengurutkan data sesuai dengan nilai presentase bayi lahir dengan berat badan rendah (*X*), dimulai dari yang paling kecil hingga paling besar. Kemudian menghilangkan nilai presentase bayi lahir dengan berat badan rendah yang di tengah (c) sebanyak 4 data (c=4 jika n=30) (Gujarati, 2004). Setelah itu data dibagi menjadi dua kelompok. Kelompok pertama adalah data dengan nilai presentase bayi lahir dengan berat badan rendah (*X*) yang kecil, sedangkan kelompok kedua adalah data dengan nilai presentase bayi lahir dengan berat badan rendah (*X*) yang besar. Selanjutnya melakukan regresi pada setiap kelompok secara terpisah. Hasil regresi untuk kelompok data dengan nilai pendapatan (*X*) kecil adalah sebagai berikut:

Tabel 4.3 Hasil Regresi Kelompok Data Nilai Pendapatan (*X*) Kecil

	Db	*SS*
Regresi	1	1.8499
Error	15	7.1701
Total	16	9.0200

Hasil regresi untuk kelompok data dengan nilai pendapatan (*X*) besar adalah sebagai berikut:

Tabel 4.4 Hasil Regresi Kelompok Data Nilai Pendapatan (X) Besar

	Db	*SS*
Regresi	1	7.8799
Error	15	14.6188
Total	16	22.4988

Hasil regresi pada tabel 4.3 dan tabel 4.4 di atas kemudian digunakan untuk menghitung F_{hitung} sebagai berikut:

$$F_{\text{hitung}} = \frac{7.1700/15}{14.6189/15} = 2.0389$$

Uji hipotesisnya adalah sebagai berikut:

- Hipotesis:

  H_0 : Hasil regresi tidak mengandung heteroskedastisitas

  H_1 : Hasil regresi mengandung heteroskedastisitas
- Tingkat signifikansi:

  $\alpha = 0.05$
- Daerah kritis:

  H_0 ditolak jika $F_{\text{hitung}} > F_{\text{tabel}}$
- Statistik uji:

  $F_{\text{hitung}} = \frac{7.1700/15}{14.6189/15} = 2.0389$

  $F_{\text{tabel}} = F_{(0.05;15;15)} = 2.4034$
- Keputusan:

  $F_{\text{hitung}} = 2.0389 < F_{\text{tabel}} = 2.4034$ maka gagal tolak H_0.
- Kesimpulan:

  Dengan menggunakan tingkat signifikansi sebesar $\alpha = 0.05$ maka dapat diambil kesimpulan bahwa hasil regresi tidak mengandung heteroskedastisitas.

c. Estimasi dengan Regresi Kuantil Median pada Data dengan Homoskedastisitas

Selanjutnya adalah analisis menggunakan regresi kuantil median. Hal tersebut dilakukan untuk mengetahui bagaimana jika data yang homoskedastisitas dianalisis menggunakan regresi kuantil median. Didapatkan hasil analisis sebagai berikut:

Tabel 4.5 Hasil Estimasi Koefisien β_0 dengan Regresi Kuantil Median pada Data dengan Homoskedastisitas

Variabel	Kuantil	Koefisien	*Standard Error*	Signifikansi
intersep	**0.15**	0.2546	0.3965	0.5248
	0.20	0.3277	0.4354	0.4566
	0.25	0.1954	0.4291	0.6516
	0.30	0.4143	0.3116	0.1921
	0.35	0.4579	0.3018	0.1379
	0.40	0.3839	0.3443	0.2721
	0.45	0.4350	0.3945	0.2774
	0.50	0.6969	0.4348	0.1177
	0.55	0.7369	0.4035	0.0761
	0.60	0.7069	0.4012	0.0866
	0.65	0.9281	0.4206	0.0228
	0.70	0.9623	0.6036	0.1196
	0.75	1.3782	0.5963	0.0266
	0.80	1.7853	0.8654	0.0464
	0.85	1.8066	0.8328	0.0367

Jika H_0 adalah variabel berpengaruh secara signifikan dan H_1 adalah variabel tidak berpengaruh secara signifikan, maka berdasarkan tabel 4.5 diketahui bahwa tidak semua koefisien signifikan, karena nilai signifikansinya lebih besar dari nilai $\alpha = 0.05$. Variabel yang berpengaruh secara signifikan adalah pada kuantil 0.65, 0.75, 0.80, dan 0.85. Kemudian nilai koefisien yang dipilih adalah nilai koefisien pada kuantil 0.5, yaitu 0.6969.

Tabel 4.6 Hasil Estimasi Koefisien β_1 dengan Regresi Kuantil Median pada Data dengan Homoskedastisitas

Variabel	Kuantil	Koefisien	*Standard Error*	Signifikansi
X	**0.15**	1.0309	0.5282	0.0588
	0.20	0.9901	0.3567	0.0087
	0.25	1.3846	0.3691	0.0006
	0.30	1.4286	0.3325	0.0001
	0.35	1.4035	0.3222	0.0001
	0.40	1.5267	0.3431	0.0001
	0.45	1.5000	0.3598	0.0002
	0.50	1.3846	0.3947	0.0012
	0.55	1.4130	0.2989	0.0000
	0.60	1.4685	0.2190	0.0000
	0.65	1.4035	0.2318	0.0000
	0.70	1.3934	0.3570	0.0004
	0.75	1.2727	0.3206	0.0003
	0.80	1.2196	0.6467	0.0674
	0.85	1.2088	0.6497	0.0710

Jika H_0 adalah variabel berpengaruh secara signifikan dan H_1 adalah variabel tidak berpengaruh secara signifikan maka berdasarkan tabel 4.6 diketahui bahwa ada beberapa koefisien yang tidak signifikan, karena nilai signifikansinya lebih dari nilai $\alpha = 0.05$. Nilai koefisien yang tidak signifikan yaitu pada kuantil 0.15, 0.8, dan 0.85. Kemudian nilai koefisien yang dipilih adalah nilai koefisien pada kuantil 0.5, yaitu 1.3846. Berdasarkan tabel tersebut didapatkan model sebagai berikut:

$$Y_{0.5} = 0.6969 + 1.3846X$$

dengan :

Y : Balita penderita gizi buruk (%)

X : Bayi yang lahir dengan berat badan rendah (%)

Model tersebut dapat dijelaskan sebagai berikut:

a. Konstanta sebesar 0.6969, artinya jika bayi yang lahir dengan berat badan rendah (*X*) adalah nol, maka balita penderita gizi buruk (*Y*) memiliki nilai sebesar 0.6969.

b. Koefisien regresi variabel bayi yang lahir dengan berat badan rendah (*X*) sebesar 1.384, artinya jika bayi yang lahir dengan berat badan rendah (*X*) mengalami kenaikan sebesar 1%, maka balita penderita gizi buruk (*Y*) akan mengalami peningkatan sebesar 1.384. Koefisien bernilai positif berarti bahwa bayi yang lahir dengan berat badan rendah dengan balita penderita gizi buruk memiliki hubungan yang positif. Semakin banyak bayi yang lahir dengan berat badan rendah (*X*) maka balita penderita gizi buruk (*Y*) semakin meningkat.

Tabel perhitungan analisis variansinya adalah sebagai berikut:

Tabel 4.7 Tabel Analisis Variansi

Sumber Variansi	Derajat Bebas	Jumlah Kuadrat	Kuadrat Tengah	F_{hitung}
Regresi	1	27.8669	27.8669	31.3427
Error	$n-2$	32.0077	0.8891	
Total	$n-1$	59.8745		

Uji hipotesis untuk uji *overall* adalah sebagai berikut :

- Hipotesis:

  H_0 : $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ (Model tidak sesuai)

  H_1 : Minimal ada satu $\beta_i \neq 0$, $i = 1, 2, 3, \ldots, k$ (Model sesuai)
- Tingkat signifikansi:

  $\alpha = 0.05$
- Daerah kritis:

  H_0 ditolak jika $F_{\text{hitung}} > F_{\text{tabel}}$
- Statistik Uji

  $F_{\text{hitung}} = 31.3427 > F_{\text{tabel}} = 4.1132$
- Keputusan

  $F_{\text{hitung}} = 31.3427 > F_{\text{tabel}} = 4.1132$ maka tolak H_0.

- Kesimpulan

  Dengan menggunakan tingkat signifikansi sebesar $\alpha = 0.05$ maka dapat diambil kesimpulan bahwa minimal ada satu $\beta_i \neq 0$ atau dengan kata lain model sesuai.

Untuk melihat kesalahan pengukuran dari persamaan model dalam melakukan estimasi variabel respon digunakan nilai JKE. Nilai JKE yang didapatkan adalah sebesar 0.8891. Selanjutnya dilakukan uji GoldFeld-Quandt kembali untuk mendeteksi heteroskedastisitas. Pengujian dilakukan dengan mengurutkan data sesuai dengan nilai pendapatan (*X*), dimulai dari yang paling kecil hingga paling besar. Kemudian menghilangkan nilai pendapatan yang di tengah (*c*) sebanyak 4 data (c=4 jika n=30) (Gujarati, 2004). Setelah itu data dibagi menjadi dua kelompok. Kelompok pertama adalah data dengan presentase bayi yang lahir dengan berat badan rendah (*X*) yang kecil, sedangkan kelompok kedua adalah data dengan presentase bayi yang lahir dengan berat badan rendah (*X*) yang besar. Selanjutnya melakukan regresi pada setiap kelompok secara terpisah. Variabel *Y* yang digunakan adalah $\hat{Y}$. Hasil regresi untuk kelompok data presentase bayi yang lahir dengan berat badan rendah (*X*) kecil adalah sebagai berikut:

Tabel 4.8 Hasil Regresi Kelompok Data Presentase Bayi Yang Lahir dengan Berat Badan Rendah (*X*) Kecil

	db	*SS*
Regresi	1	0.6774
Error	15	9.5741
Total	16	10.2515

Hasil regresi untuk kelompok data dengan presentase bayi yang lahir dengan berat badan rendah (*X*) besar adalah sebagai berikut:

Tabel 4.9 Hasil Regresi Kelompok Data Presentase Bayi Yang Lahir dengan Berat Badan Rendah (*X*) Besar

	db	*SS*
Regresi	1	0.01186
Error	15	12.9709
Total	16	12.9827

Hasil regresi di atas kemudian digunakan untuk menghitung F_{hitung} sebagai berikut:

$$F_{\text{hitung}} = \frac{12.9709/15}{9.5741/15} = 1.3548$$

Uji hipotesisnya adalah sebagai berikut:

- Hipotesis:

  H_0 : Hasil regresi tidak mengandung heteroskedastisitas

  H_1 : Hasil regresi mengandung heteroskedastisitas
- Tingkat signifikansi:

  $\alpha = 0.05$
- Daerah kritis:

  H_0 ditolak jika $F_{\text{hitung}} > F_{\text{tabel}}$
- Statistik uji:

  $F_{\text{hitung}} = \frac{12.9709/15}{9.5741/15} = 1.3548$

  $F_{\text{tabel}} = F_{(0.05;15;15)} = 2.4034$
- Keputusan:

  $F_{\text{hitung}} = 1.3548 < F_{\text{tabel}} = 2.4034$ maka gagal tolak H_0.
- Kesimpulan:

  Dengan menggunakan tingkat signifikansi sebesar $\alpha = 0.05$ maka dapat diambil kesimpulan bahwa hasil regresi tidak mengandung heteroskedastisitas.

4.3.2 Data dengan Heteroskedastisitas

a. Estimasi dengan Metode Kuadrat Terkecil pada Data dengan Heteroskedastisitas

Sama seperti pada data dengan homoskedastisitas, langkah pertama yang dilakukan adalah melakukan estimasi dengan MKT. Estimasi dilakukan dengan menggunakan *software R* 3.2.3. Hasil estimasi tersebut adalah sebagai berikut:

Tabel 4.10 Hasil Estimasi dengan MKT pada Data dengan Homoskedastisitas

Variabel	Nilai Estimasi	*Standard Error*	*P-value*
(intersep)	9.2903	5.2314	0.0866
X	0.6378	0.0286	2.33E-19

Tabel 4.10 di atas adalah hasil estimasi data pengeluaran konsumsi (dolar) dengan pendapatan (dolar) menggunakan MKT. Berdasarkan tabel 4.10 didapatkan model sebagai berikut:

$$Y = 9.2903 + 0.6378X$$

dengan :

Y : Pengeluaran konsumsi (dolar)

X : Pendapatan (dolar)

Model tersebut dapat dijelaskan sebagai berikut:

a. Nilai konstanta sebesar 9.2903, artinya jika pendapatan tidak mengalami penambahan ataupun pengurangan maka pengeluaran konsumsi sebesar 9.2903 dolar.

b. Koefisien regresi variabel pendapatan sebesar 0.6378, artinya jika pendapatan mengalami kenaikan sebesar 1 dolar maka besarnya pengeluaran konsumsi akan mengalami kenaikan sebesar 0.6378 dolar. Koefisien bernilai positif artinya terjadi hubungan positif antara pendapatan dan pengeluaran konsumsi. Semakin besar pendapatan maka pengeluaran konsumsi juga akan semakin besar.

Tabel perhitungan analisis variansinya adalah sebagai berikut:

Tabel 4.11 Tabel Analisis Variansi

Sumber Variansi	Derajat Bebas	Jumlah Kuadrat	Kuadrat Tengah	F_{hitung}
Regresi	1	41886.1148	41886.1148	496.7112
Error	n - 2	2361.1533	84.3269	
Total	n - 1	44247.8667		

Uji hipotesis untuk uji *overall* adalah sebagai berikut :

- Hipotesis:

  H_0 : $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ (Model tidak sesuai)

  H_1 : Minimal ada satu $\beta_i \neq 0$, $i = 1, 2, 3, \ldots, k$ (Model sesuai)
- Tingkat signifikansi:

  $\alpha = 0.05$
- Daerah kritis:

  H_0 ditolak jika $F_{hitung} > F_{tabel}$
- Statistik Uji

  $F_{hitung} = 496.7112 > F_{tabel} = 4.1960$
- Keputusan

  $F_{hitung} = 496.7112 > F_{tabel} = 4.1960$ maka tolak H_0.
- Kesimpulan

  Dengan menggunakan tingkat signifikansi sebesar $\alpha = 0.05$ maka dapat diambil kesimpulan bahwa minimal ada satu $\beta_i \neq 0$ atau dengan kata lain model sesuai.

Untuk melihat kesalahan pengukuran dari persamaan model dalam melakukan estimasi variabel respon digunakan nilai KTE. Nilai KTE yang didapatkan adalah sebesar 84.3269.

b. Uji GoldFeld-Quandt pada Data dengan Heteroskedastisitas

Selanjutnya dilakukan uji heteroskedastisitas pada *error*. Untuk uji heteroskedastisitas digunakan uji GoldFeld-Quandt. Pengujian dilakukan dengan mengurutkan data sesuai dengan nilai pendapatan (X), dimulai dari yang paling kecil hingga paling besar. Kemudian menghilangkan nilai pendapatan yang di tengah (c) sebanyak 4 data (c=4 jika n=30) (Gujarati, 2004). Setelah itu data dibagi menjadi dua kelompok. Kelompok pertama adalah data dengan nilai pendapatan (X) yang kecil, sedangkan kelompok kedua adalah data dengan nilai pendapatan (X) yang besar. Selanjutnya melakukan regresi pada setiap kelompok secara terpisah. Hasil regresi untuk kelompok data dengan nilai pendapatan (X) kecil adalah sebagai berikut:

Tabel 4.12 Hasil Regresi Kelompok Data Nilai Pendapatan (*X*) Kecil

	Db	*SS*
Regresi	1	3010.065
Error	11	377.1663
Total	12	3387.231

Hasil regresi untuk kelompok data dengan nilai pendapatan (*X*) besar adalah sebagai berikut:

Tabel 4.13 Hasil Regresi Kelompok Data Nilai Pendapatan (*X*) Besar

	db	*SS*
Regresi	1	5088.893
Error	11	1536.8
Total	12	6625.692

Hasil regresi di atas kemudian digunakan untuk menghitung F_{hitung} sebagai berikut:

$$F_{\text{hitung}} = \frac{377.1663/11}{1536.8/11} = 4.0746$$

Uji hipotesisnya adalah sebagai berikut:

- Hipotesis:

  H_0 : Hasil regresi tidak mengandung heteroskedastisitas

  H_1 : Hasil regresi mengandung heteroskedastisitas
- Tingkat signifikansi:

  $\alpha = 0.05$
- Daerah kritis:

  H_0 ditolak jika $F_{\text{hitung}} > F_{\text{tabel}}$
- Statistik uji:

  $F_{\text{hitung}} = \frac{377.1663/11}{1536.8/11} = 4.0746$

  $F_{\text{tabel}} = F_{(0.05;11;11)} = 2.8179$
- Keputusan:

  $F_{\text{hitung}} = 4.0746 > F_{\text{tabel}} = 2.8179$ maka tolak H_0.

- Kesimpulan:

  Dengan menggunakan tingkat signifikansi sebesar $\alpha = 0.05$ maka dapat diambil kesimpulan bahwa hasil regresi mengandung heteroskedastisitas.

c. Estimasi dengan Regresi Kuantil Median pada Data dengan Heteroskedastisitas

Berdasarkan uji GoldFeld-Quandt yang telah dilakukan, diketahui bahwa terdapat masalah heteroskedastisitas. Untuk mengatasi masalah heteroskedastisitas tersebut dilakukan estimasi dengan regresi kuantil median. Berikut adalah hasil estimasi regresi kuantil median dengan menggunakan *software R* 3.2.3:

Tabel 4.14 Hasil Estimasi Koefisien β_0 dengan Regresi Kuantil Median pada Data dengan Heteroskedastisitas

Variabel	Kuantil	Koefisien	*Standard Error*	Signifikansi
intersep	**0.15**	10.0000	4.2197	0.0249
	0.20	9.2857	5.5748	0.1069
	0.25	7.0000	6.6343	0.3004
	0.30	2.0400	6.3477	0.7503
	0.35	5.7500	6.3035	0.3695
	0.40	4.7000	7.2738	0.5234
	0.45	9.6000	7.0907	0.1866
	0.50	8.0000	6.8686	0.2540
	0.55	6.6667	7.0614	0.3532
	0.60	6.6667	5.7974	0.2590
	0.65	7.8621	5.6856	0.1777
	0.70	16.0000	5.8512	0.0107
	0.75	11.7200	6.0204	0.0617
	0.80	14.4054	6.1137	0.0257
	0.85	15.1053	3.5789	0.0002

Jika H_0 adalah variabel berpengaruh secara signifikan dan H_1 adalah variabel tidak berpengaruh secara signifikan maka berdasarkan tabel 4.14 diketahui bahwa tidak semua koefisien signifikan, karena nilai signifikansinya lebih besar dari nilai $\alpha = 0.05$. Variabel yang berpengaruh secara signifikan adalah pada kuantil 0.15, 0.70, 0.80, dan 0.85. Kemudian nilai koefisien yang dipilih adalah nilai koefisien pada kuantil 0.5, yaitu 8.0000.

Tabel 4.15 Hasil Estimasi Koefisien β_1 dengan Regresi Kuantil Median pada Data dengan Heteroskedastisitas

Variabel	Kuantil	Koefisien	*Standard Error*	Signifikansi
X	**0.15**	0.5625	0.0395	0.0000
	0.20	0.5800	0.0447	0.0000
	0.25	0.6000	0.0468	0.0000
	0.30	0.0664	0.0455	0.0000
	0.35	0.6500	0.0456	0.0000
	0.40	0.6600	0.0464	0.0000
	0.45	0.6400	0.0484	0.0000
	0.50	0.6546	0.0443	0.0000
	0.55	0.6667	0.0435	0.0000
	0.60	0.6667	0.0290	0.0000
	0.65	0.6621	0.0323	0.0000
	0.70	0.6308	0.0399	0.0000
	0.75	0.6640	0.0399	0.0000
	0.80	0.6541	0.0398	0.0000
	0.85	0.65263	0.0313	0.0000

Jika H_0 adalah variabel berpengaruh secara signifikan dan H_1 adalah variabel tidak berpengaruh secara signifikan maka berdasarkan tabel 4.15 diketahui bahwa semua koefisien signifikan, karena nilai signifikansinya kurang dari nilai $\alpha = 0.05$. Kemudian nilai koefisien yang dipilih adalah nilai koefisien pada kuantil 0.5, yaitu 0.6546. Berdasarkan tabel 4.12 didapatkan model sebagai berikut:

$$Y_{0.5} = 8.0000 + 0.6546X$$

dengan :

Y : Pengeluaran konsumsi (dolar)

X : Pendapatan (dolar)

a. Nilai konstanta sebesar 8.0000, artinya jika pendapatan tidak mengalami penambahan ataupun pengurangan maka pengeluaran konsumsi sebesar 8.0000 dolar.

b. Koefisien regresi variabel pendapatan sebesar 0.6546, artinya jika pendapatan mengalami kenaikan sebesar 1 dolar maka besarnya pengeluaran konsumsi akan mengalami kenaikan sebesar 0.6546 dolar. Koefisien bernilai positif artinya terjadi hubungan positif antara pendapatan dan pengeluaran konsumsi. Semakin besar pendapatan maka pengeluaran konsumsi juga akan semakin besar.

Tabel perhitungan analisis variansinya adalah sebagai berikut:

Tabel 4.16 Tabel Analisis Variansi

Sumber Variansi	Derajat Bebas	Jumlah Kuadrat	Kuadrat Tengah	F_{hitung}
Regresi	1	44203.4342	44203.4342	501.2632
Error	n - 2	2469.1545	88.1841	
Total	n - 1	44247.8667		

Uji hipotesis untuk uji *overall* adalah sebagai berikut :

- Hipotesis:

  H_0 : $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ (Model tidak sesuai)

  H_1 : Minimal ada satu $\beta_i \neq 0$, $i = 1, 2, 3, \ldots, k$ (Model sesuai)
- Tingkat signifikansi:

  $\alpha = 0.05$
- Daerah kritis:

  H_0 ditolak jika $F_{\text{hitung}} > F_{\text{tabel}}$
- Statistik Uji

  $F_{\text{hitung}} = 501.2632 > F_{\text{tabel}} = 4.1960$
- Keputusan

  $F_{\text{hitung}} = 501.2632 > F_{\text{tabel}} = 4.1960$ maka tolak H_0.

- Kesimpulan

  Dengan menggunakan tingkat signifikansi sebesar $\alpha = 0.05$ maka dapat diambil kesimpulan bahwa minimal ada satu $\beta_i \neq 0$ atau dengan kata lain model sesuai.

Untuk melihat kesalahan pengukuran dari persamaan model dalam melakukan estimasi variabel respon digunakan nilai KTE. Nilai KTE yang didapatkan adalah sebesar 88.1841. Selanjutnya dilakukan uji GoldFeld-Quandt kembali untuk mendeteksi heteroskedastisitas. Pengujian dilakukan dengan mengurutkan data sesuai dengan nilai pendapatan (X), dimulai dari yang paling kecil hingga paling besar. Kemudian menghilangkan nilai pendapatan yang di tengah (c) sebanyak 4 data (c=4 jika n=30) (Gujarati, 2004). Setelah itu data dibagi menjadi dua kelompok. Kelompok pertama adalah data dengan nilai pendapatan (X) yang kecil, sedangkan kelompok kedua adalah data dengan nilai pendapatan (X) yang besar. Selanjutnya melakukan regresi pada setiap kelompok secara terpisah. Variabel Y yang digunakan adalah $\hat{Y}$. Hasil regresi untuk kelompok data dengan nilai pendapatan (X) kecil adalah sebagai berikut:

Tabel 4.17 Hasil Regresi Kelompok Data Nilai Pendapatan (X) Kecil

	Db	*SS*
Regresi	1	1575.822
Error	11	1753.303
Total	12	3329.124

Hasil regresi untuk kelompok data dengan nilai pendapatan (X) besar adalah sebagai berikut:

Tabel 4.18 Hasil Regresi Kelompok Data Nilai Pendapatan (X) Besar

	Db	*SS*
Regresi	1	604.8775
Error	11	3465.884
Total	12	4070.761

Hasil regresi di atas kemudian digunakan untuk menghitung F_{hitung} sebagai berikut:

$$F_{\text{hitung}} = \frac{3465.884/11}{1753.303/11} = 1.9768$$

Uji hipotesisnya adalah sebagai berikut:

- Hipotesis:

  H_0 : Hasil regresi tidak mengandung heteroskedastisitas

  H_1 : Hasil regresi mengandung heteroskedastisitas

- Tingkat signifikansi:

  $\alpha = 0.05$

- Daerah kritis:

  H_0 ditolak jika $F_{\text{hitung}} > F_{\text{tabel}}$

- Statistik uji:

  $F_{\text{hitung}} = \frac{3465.884/11}{1753.303/11} = 1.9768$

  $F_{\text{tabel}} = F_{(0.05;11;11)} = 2.8179$

- Keputusan:

  $F_{\text{hitung}} = 4.0746 < F_{\text{tabel}} = 2.8179$ maka gagal tolak H_0.

- Kesimpulan:

  Dengan menggunakan tingkat signifikansi sebesar $\alpha = 0.05$ maka dapat diambil kesimpulan bahwa hasil regresi tidak mengandung heteroskedastisitas.

Jadi, terbukti bahwa dengan menggunakan regresi kuantil median dapat menghilangkan heteroskedastisitas pada data pengeluaran konsumsi (*Y*) dan pendapatan (*X*).

4.3.2 Perbandingan KTE

Menentukan metode terbaik untuk analisis data dengan homoskedastisitas dan homoskedastisitas tersebut dengan menggunakan KTE. Dari kedua metode yang digunakan, yaitu MKT dan regresi kuantil median yang digunakan untuk analisis pada data dengan homoskedastisitas dan heteroskedastisitas diperoleh nilai *MSE* secara keseluruhan sebagai berikut:

Tabel 4.19 Perbandingan KTE antara Metode MKT dan Regresi Kuantil Median pada Data dengan Homoskedastisitas dan Heteroskedastisitas

Data	Metode	
	MKT	Regresi Kuantil Median
Homoskedastisitas	0.8770	0.8891
Heteroskedastisitas	84.3269	88.1841

Pada tabel 4.19 diketahui bahwa pada data dengan homoskedastisitas KTE dengan metode regresi kuantil median lebih besar dari KTE MKT. Hal tersebut menunjukkan bahwa MKT lebih cocok digunakan untuk analisis pada data dengan homoskedastisitas. Sedangkan pada data dengan heteroskedastisitas menunjukkan bahwa KTE regresi kuantil median lebih besar dari KTE MKT.

BAB V
PENUTUP

5.1 Kesimpulan

Berdasarkan pada hasil pembahasan, dapat diambil kesimpulan sebagai berikut:

1. Kemampuan metode regresi kuantil median dalam menyelesaikan permasalahan heteroskedastisitas dilihat pada hasil uji heteroskedastisitas pada *error* data setelah dilakukan analisis regresi kuantil. Setelah dilakukan analisis regresi kuantil median, data yang semula mengandung heteroskedastisitas menjadi tidak mengandung heteroskedastisitas atau dengan kata lain memenuhi homoskedastisitas. Hal ini menunjukkan bahwa regresi kuantil median mampu untuk menangani heteroskedastisitas pada data.
2. Analisis pada data homoskedastisitas lebih tepat dengan menggunakan MKT, sedangkan pada data heteroskedastisitas lebih tepat dengan menggunakan regresi kuantil median karena dengan menggunakan regresi kuantil median dapat menghilangkan heteroskedastisitas. Namun pada penelitian ini nilai KTE yang didapatkan regresi kuantil median lebih besar dari MKT.

5.2 Saran

Berdasarkan kesimpulan yang diperoleh dari hasil pembahasan, diketahui bahwa regresi kuantil median mampu mengatasi heteroskedastisitas pada data dan regresi kuantil lebih baik daripada MKT untuk analisis pada data yang mengandung heteroskedastisitas, sehingga dapat digunakan sebagai referensi ketika ditemukan data yang mengandung heteroskedastisitas.

Pada penelitian selanjutnya diharapkan peneliti mampu mengetahui bagaimana tahapan estimasi pada regresi kuantil.

DAFTAR PUSTAKA

Algifari. 1997. *Analisis Regresi Teori, Kasus dan Solusi.* Yogyakarta: BPFE.

Dinas Kesehatan Provinsi Jawa Timur. http://dinkesjatim.go.id diakses pada tanggal 12 Februari 2016 pukul 08.40.

Buhai, S. 2005. *Quantile Regression: Overview and Selected Applications*. Journal Ad Astra. Vol. 4:1-17.

Dahlan, M. S. 2008. *Statistik untuk Kedokteran dan Kesehatan*. Yogyakarta: Salemba Medika.

Diaz, G. 2007. *Encyclopedia of Statistics.* Delhi: Global Media.

Djuraidah, A. dan Wigena, A. H. 2011. *Regresi Kuantil untuk Eksplorasi Pola Curah Hujan di Kabupaten Indramayu.* Jurnal Ilmu Dasar. Vol. 12, No. 1: 50-56.

Fitriah, R. 2009. *Analisis Regresi Menggunakan Metode Kuantil.* Skripsi Departemen Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam. Bogor: Institut Pertanian Bogor.

Furno, M. dan Vistocco, D. 2013. *Qu Test for Structural Breaks in Quantile Regressions*. International Journal of Statistics and Probability. Vol. 2, No. 4: 42-55.

Gujarati, D. N. 2004. *Basic Econometrics Fourth Edition.* New York: McGraw-Hill.

Harinaldi. 2005. *Prinsip-Prinsip Statistik untuk Teknik dan Sains.* Jakarta: Erlangga.

Kirnasari, T. P. 2014. *Perbandingan Metode Weighted Least Square dan Regresi Kuantil Median dalam Menyelesaikan Kasus Heteroskedastisitas pada Analisis Regresi.* Skripsi. Malang: Fakultas MIPA Universitas Malang.

Koenker, R. 2005. *Quantile Regression*. New York: Cambridge University Press.

Makkulau, Linuwih, S., Purhadi, dan Mashuri, M. 2010. *Pendeteksian Outlier dan Penentuan Faktor-Faktor yang Mempengaruhi Produksi Gula dan Tetes Tebu dengan Metode Likelihood Displacement Statistic-Lagrange.* Jurnal Teknik Industri, Vol. 12, No.2: 95-100.

Makridakis, S., Wheelwright, S. C., dan McGee, V. E. 1999. *Metode dan Aplikasi Peramalan*, Terj. dari *Forecasting: Methods and Applications*, oleh Hari Suminto. Jakarta: Binarupa Aksara.

Maziyya, P. A., Sukarsa, I. K. G., dan Asih, N. M. 2015. *Mengatasi Heteroskedastisitas pada Regresi dengan Menggunakan Weighted Least Square.* E-Jurnal Matematika, Vol. 4, No. 1: 20-25.

Mendenhall, W., dan Sincich, T. 1996. *A Second Course in Statistics: Regression Analysis, Seventh Edition.* United States of America: Pearson Education, Inc.

Mokosolang, C. A., Prang, J. D., dan Mananohas, M. L. 2015. *Analisis Heteroskedastisitas pada Data Cross Section dengan White Heteroscedasticity Test dan Weighted Least Squares.* JdC, Vol. 4, No. 2: 172-179.

Montgomery, D. C. dan Peck, E. A. 1982. *Introduction to Linear Regression Analysis.* United States of America: John Wiley & Sons.

Qudratullah, M. F. 2013. *Analisis Regresi Terapan: Teori, Contoh Kasus, dan Aplikasi dengan SPSS.* Yogyakarta: Penerbit ANDI.

Rahmawati, R., Widiarti, dan Novianti, P. 2011. *Regresi Kuantil (Studi Kasus pada Data Suhu Harian).* Prosiding Seminar Nasional Statistika Universitas Diponegoro, A-23.

Soejoeti, Zanzawi. 1986. *Buku Materi Pokok Metode Statistika II.* Jakarta: Karunika Jakarta.

Sunyoto, D. 2007. *Analisis Regresi dan Korelasi Bivariat: Ringkasan dan Kasus.* Yogyakarta: Amara Books.

Uthami, I. A. P., Sukarsa, I. K. G., dan Kencana, I. P. E. N. 2013. *Regresi Kuantil Median untuk Mengatasi Heteroskedastisitas pada Analisis Regresi.* E-Jurnal Matematika, Vol. 2, No. 1:6-13.

Wahyudi, V. E. dan Zain, I. 2014. *Analisis IPM di Pulau Jawa Menggunakan Analisis Regresi Kuantil.* Jurnal Statistika, Vol. 2, No. 1: 64-69.

Walpole, R. E. dan Myers, R. H. 1995. *Ilmu Peluang dan statistika untuk Insinyur dan Ilmuwan*, Terj. *Probability and Statistics for Engineers and Scientists*, oleh RK Sembiring. Bandung: Penerbit ITB.

Widarjono, A. 2005. *Ekonometrika: Teori dan Aplikasi untuk Ekonomi dan Bisnis.*Yogyakarta: Penerbit Ekonisia.

Youlanda, S. R. 2015. *Perbandingan Metode Regresi Kuantil Median dengan Metode Weighted Least Square (WLS) Untuk Menyelesaikan Heteroskedastisitas pada Analisis Regresi*. Skripsi. Jember: Fakultas MIPA Universitas Jember.

LAMPIRAN

Lampiran 1. Data dengan Homoskedastisitas

Tabel 1. Data Balita Penderita Gizi Buruk (*Y*) dan Bayi Lahir dengan Berat Badan Rendah (*X*) pada Kabupaten/Kota di Provinsi Jawa Timur Tahun 2012

Kabupaten/Kota	Balita Penderita Gizi Buruk (%)	BGM (%)
Kabupaten Pacitan	1.1	0.82
Kabupaten Ponorogo	1.5	0.54
Kabupaten Trenggalek	0.8	0.55
Kabupaten Tulungagung	1.4	0.41
Kabupaten Blitar	1.5	0.91
Kabupaten Kediri	2.1	1.18
Kabupaten Malang	1.6	0.81
Kabupaten Lumajang	3.1	0.68
Kabupaten Jember	3.4	0.2
Kabupaten Banyuwangi	1.4	0.87
Kabupaten Bondowoso	2.9	1.74
Kabupaten Situbondo	4.7	1.74
Kabupaten Probolinggo	5.7	3.4
Kabupaten Pasuruan	2.1	2.00
Kabupaten Sidoarjo	2.5	1.24
Kabupaten Mojokerto	2.3	1.52
Kabupaten Jombang	1.5	0.58
Kabupaten Nganjuk	2.4	1.31
Kabupaten Madiun	1.5	1.35
Kabupaten Magetan	1.7	0.55
Kabupaten Ngawi	2.1	1.79
Kabupaten Bojonegoro	0.6	1.13
Kabupaten Tuban	2.3	0.96
Kabupaten Lamongan	1.4	0.65
Kabupaten Gresik	2.4	1.23
Kabupaten Bangkalan	2.2	0.34
Kabupaten Sampang	4.2	1.98
Kabupaten Pamekasan	3.3	1.91
Kabupaten Sumenep	4.9	1.34
Kota Kediri	2.6	0.96
Kota Blitar	0.3	0.63
Kota Malang	1.9	0.50
Kota Probolinggo	3.1	1.07
Kota Pasuruan	4.1	1.60
Kota Mojokerto	3.3	1.51
Kota Madiun	1.3	0.60
Kota Surabaya	2.8	1.46
Kota Batu	1.1	0.78

Keterangan: BGM : Bawah Garis Merah

Sumber: Dinas Kesehatan Provinsi Jawa Timur, 2012

Lampiran 2. Data dengan Heteroskedastisitas

Tabel 2. Data Pengeluaran Konsumsi dan Pendapatan

Pengeluaran Konsumsi (dolar)	**Pendapatan (dolar)**
55	80
65	100
70	85
80	110
79	120
84	115
98	130
95	140
90	125
75	90
74	105
110	160
113	150
125	165
108	145
115	180
140	225
120	200
145	240
130	185
152	220
144	210
175	245
180	260
135	190
140	205
178	265
191	270
137	230
189	250

Sumber: Gujarati, 2004

Lampiran 3. Perintah pada *Software R* 3.2.3 untuk Analisis Regresi dengan MKT dan Regresi Kuantil

```
data=read.delim("clipboard")
data
Y=data$Y
Y
ccy=matrix(Y)
ccy
X=data$X
X
ccx=matrix(X)
ccx

#Estimasi MKT
model <- lm(Y ~ X)
summary(model)

#Regresi Kuantil
library(quantreg)
regresiquantile <- rq(Y ~ X, tau=0.15,data=data,na.action=na.omit,
                      method="br", model = TRUE, contrasts=NULL)
summary(regresiquantile, se="nid")
```

Lampiran 4. Hasil Analisis dengan MKT pada Data dengan Homoskedastisitas Menggunakan *Software R* 3.2.3

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.8983     0.3171   2.833  0.00751 **
X             1.2830     0.2469   5.196 8.23e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9365 on 36 degrees of freedom
Multiple R-squared:  0.4286,    Adjusted R-squared:  0.4127
F-statistic:    27 on 1 and 36 DF,  p-value: 8.235e-06
```

Lampiran 5. Hasil Analisis dengan Regresi Kuantil pada Data dengan Homoskedastisitas Menggunakan *Software R* 3.2.3

```
tau: [1] 0.15

Coefficients:
            Value   Std. Error t value Pr(>|t|)
(Intercept) 0.25464 0.39653    0.64217 0.52483
X           1.03093 0.52819    1.95181 0.05877
```

```
tau: [1] 0.2

Coefficients:
            Value   Std. Error t value Pr(>|t|)
(Intercept) 0.32772 0.43542    0.75265 0.45655
X           0.99010 0.35666    2.77605 0.00868
```

```
tau: [1] 0.25

Coefficients:
            Value   Std. Error t value Pr(>|t|)
(Intercept) 0.19538 0.42905    0.45539 0.65156
X           1.38462 0.36910    3.75130 0.00062
```

```
tau: [1] 0.3

Coefficients:
            Value   Std. Error t value Pr(>|t|)
(Intercept) 0.41429 0.31164    1.32936 0.19209
X           1.42857 0.33250    4.29641 0.00013
```

```
tau: [1] 0.35

Coefficients:
            Value   Std. Error t value Pr(>|t|)
(Intercept) 0.45789 0.30181    1.51715 0.13796
X           1.40351 0.32219    4.35621 0.00011
```

```
tau: [1] 0.4

Coefficients:
            Value   Std. Error t value Pr(>|t|)
(Intercept) 0.38397 0.34426    1.11536 0.27209
X           1.52672 0.34311    4.44965 0.00008
```

```
tau: [1] 0.45

Coefficients:
            Value   Std. Error t value Pr(>|t|)
(Intercept) 0.43500 0.39445    1.10281 0.27743
X           1.50000 0.35977    4.16929 0.00018
```

```
tau: [1] 0.5

Coefficients:
            Value   Std. Error t value Pr(>|t|)
(Intercept) 0.69692 0.43477    1.60296 0.11768
X           1.38462 0.39478    3.50727 0.00123
```

```
tau: [1] 0.55

Coefficients:
            Value   Std. Error t value Pr(>|t|)
(Intercept) 0.73696 0.40345    1.82666 0.07605
X           1.41304 0.29887    4.72799 0.00003

tau: [1] 0.6

Coefficients:
            Value   Std. Error t value Pr(>|t|)
(Intercept) 0.70699 0.40124    1.76201 0.08656
X           1.46853 0.21901    6.70536 0.00000

tau: [1] 0.65

Coefficients:
            Value   Std. Error t value Pr(>|t|)
(Intercept) 0.92807 0.42064    2.20632 0.03383
X           1.40351 0.23179    6.05520 0.00000

tau: [1] 0.7

Coefficients:
            Value   Std. Error t value Pr(>|t|)
(Intercept) 0.96230 0.60362    1.59420 0.11963
X           1.39344 0.35702    3.90296 0.00040

tau: [1] 0.75

Coefficients:
            Value   Std. Error t value Pr(>|t|)
(Intercept) 1.37818 0.59625    2.31143 0.02664
X           1.27273 0.32060    3.96980 0.00033

tau: [1] 0.8

Coefficients:
            Value   Std. Error t value Pr(>|t|)
(Intercept) 1.78537 0.86540    2.06305 0.04638
X           1.21951 0.64666    1.88587 0.06740

tau: [1] 0.85

Coefficients:
            Value   Std. Error t value Pr(>|t|)
(Intercept) 1.80659 0.83279    2.16933 0.03674
X           1.20879 0.64971    1.86051 0.07100
```

Lampiran 6. Hasil Analisis dengan MKT pada Data dengan Heteroskedastisitas Menggunakan *Software R* 3.2.3

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.29031    5.23139   1.776   0.0866 .
X            0.63778    0.02862  22.287   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.:

Residual standard error: 9.183 on 28 degrees of freedom
Multiple R-squared:  0.9466,    Adjusted R-squared:  0.:
F-statistic: 496.7 on 1 and 28 DF,  p-value: < 2.2e-16
```

Lampiran 7. Hasil Analisis dengan Regresi Kuantil pada Data dengan Heteroskedastisitas Menggunakan *Software R* 3.2.3

```
tau: [1] 0.15

Coefficients:
            Value    Std. Error t value  Pr(>|t|)
(Intercept) 10.00000  4.21970    2.36984  0.02492
X            0.56250  0.03949   14.24568  0.00000

tau: [1] 0.2

Coefficients:
            Value    Std. Error t value  Pr(>|t|)
(Intercept)  9.28571  5.57479    1.66566  0.10693
X            0.58095  0.04469   12.99836  0.00000

tau: [1] 0.25

Coefficients:
            Value    Std. Error t value  Pr(>|t|)
(Intercept)  7.00000  6.63431    1.05512  0.30039
X            0.60000  0.04684   12.80885  0.00000

tau: [1] 0.3

Coefficients:
            Value    Std. Error t value  Pr(>|t|)
(Intercept)  2.04000  6.34766    0.32138  0.75031
X            0.66400  0.04551   14.59042  0.00000
```

```
tau: [1] 0.35

Coefficients:
             Value    Std. Error t value  Pr(>|t|)
(Intercept)  5.75000  6.30354     0.91219  0.36946
X            0.65000  0.04560    14.25310  0.00000

tau: [1] 0.4

Coefficients:
             Value    Std. Error t value  Pr(>|t|)
(Intercept)  4.70000  7.27382     0.64615  0.52344
X            0.66000  0.04636    14.23504  0.00000

tau: [1] 0.45

Coefficients:
             Value    Std. Error t value  Pr(>|t|)
(Intercept)  9.60000  7.09066     1.35389  0.18660
X            0.64000  0.04836    13.23305  0.00000

tau: [1] 0.5

Coefficients:
             Value    Std. Error t value  Pr(>|t|)
(Intercept)  8.00000  6.86864     1.16471  0.25396
X            0.65455  0.04434    14.76099  0.00000

tau: [1] 0.55

Coefficients:
             Value    Std. Error t value  Pr(>|t|)
(Intercept)  6.66667  7.06139     0.94410  0.35319
X            0.66667  0.04351    15.32054  0.00000

tau: [1] 0.6

Coefficients:
             Value    Std. Error t value  Pr(>|t|)
(Intercept)  6.66667  5.79737     1.14995  0.25989
X            0.66667  0.02998    22.23478  0.00000

tau: [1] 0.65

Coefficients:
             Value    Std. Error t value  Pr(>|t|)
(Intercept)  7.86207  5.68555     1.38282  0.17765
X            0.66207  0.03230    20.49876  0.00000
```

```
tau: [1] 0.7

Coefficients:
            Value    Std. Error t value  Pr(>|t|)
(Intercept) 16.00000  5.85117    2.73450  0.01071
X            0.63077  0.03987   15.81901  0.00000

tau: [1] 0.75

Coefficients:
            Value    Std. Error t value  Pr(>|t|)
(Intercept) 11.72000  6.02041    1.94671  0.06167
X            0.66400  0.03989   16.64437  0.00000

tau: [1] 0.8

Coefficients:
            Value    Std. Error t value  Pr(>|t|)
(Intercept) 14.40541  6.11369    2.35626  0.02569
X            0.65405  0.03980   16.43318  0.00000

tau: [1] 0.85

Coefficients:
            Value    Std. Error t value  Pr(>|t|)
(Intercept) 15.10526  3.57885    4.22071  0.00023
X            0.65263  0.03128   20.86649  0.00000
```

Lampiran 8. Pembuktian Rumus

Bukti (Fitriah, 2009):

Diketahui persamaan (5.6), persamaan tersebut dapat dikurangi tanpa optimisasi karena tidak tergantung pada β dan terbatas pada asumsi (2), maka

$$\beta(\tau) = \arg min_{\beta \in R^d} \left\{E[\rho_\tau(Y - X'\beta)] - E\left[\rho_\tau\left(Y - Q_\tau(Y|X)\right)\right]\right\}$$

Dengan menggunakan *error* regresi kuantil yang didefinisikan pada persamaan (5.13) dan (5.14) kemudian dituliskan

$$E[\rho_\tau(Y - X'\beta)] - E\left(\rho_\tau\left(Y - Q_\tau(Y|X)\right)\right]$$
$$= E[\rho_\tau(Y - Q_\tau(Y|X) + Q_\tau(Y|X) - X'\beta)] - E[\rho_\tau(\epsilon_\tau)]$$
$$= E\left[\rho_\tau\left(\epsilon_\tau - \left(X'\beta - Q_\tau(Y|X)\right)\right)\right] - E[\rho_\tau(\epsilon_\tau)]$$
$$= E[\rho_\tau(\epsilon_\tau - \Delta_\tau(X,\beta))] - E[\rho_\tau(\epsilon_\tau)] \qquad \text{; dengan } \rho_\tau(\epsilon_\tau) = (\tau - 1\{\epsilon_\tau < 0\})\epsilon_\tau$$
$$= E\left[\left(\tau - 1\left\{\left(\epsilon_\tau - \Delta_\tau(X,\beta)\right) < 0\right\}\right)\left(\epsilon_\tau - \Delta_\tau(X,\beta)\right)\right] - E[(\tau - 1\{\epsilon_\tau < 0\})(\epsilon_\tau)]$$
$$= E\left[(\tau - 1\{\epsilon_\tau < \Delta_\tau(X,\beta)\})\left(\epsilon_\tau - \Delta_\tau(X,\beta)\right)\right] - E[(\tau - 1\{\epsilon_\tau < 0\})(\epsilon_\tau)]$$
$$= E[\tau\epsilon_\tau - \tau\Delta_\tau(X,\beta) - 1\{\epsilon_\tau < \Delta_\tau(X,\beta)\}\epsilon_\tau + 1\{\epsilon_\tau < \Delta_\tau(X,\beta)\}\Delta_\tau(X,\beta)] - E[\tau\epsilon_\tau - 1\{\epsilon_\tau < 0\}\epsilon_\tau]$$
$$= E(\tau\epsilon_\tau) + E[(1\{\epsilon_\tau < \Delta_\tau(X,\beta)\} - \tau)\Delta_\tau(X,\beta)] - E[(1\{\epsilon_\tau < \Delta_\tau(X,\beta)\})\epsilon_\tau] - E(\tau\epsilon_\tau) + E[1\{\epsilon_\tau < 0\}\epsilon_\tau]$$
$$= E[(1\{\epsilon_\tau < \Delta_\tau(X,\beta)\} - \tau)\Delta_\tau(X,\beta)] - E[(1\{\epsilon_\tau < \Delta_\tau(X,\beta)\} - 1\{\epsilon_\tau < 0\})\epsilon_\tau] \quad (14)$$

Misalkan $\mathcal{A}(X,\beta) = E[(1\{\epsilon_\tau < \Delta_\tau(X,\beta)\} - \tau)\Delta_\tau(X,\beta)|X]$

$\mathcal{B}(X,\beta) = E[(1\{\epsilon_\tau < \Delta_\tau(X,\beta)\} - 1\{\epsilon_\tau < 0\})\epsilon_\tau|X]$

Maka

$$\beta(\tau) = \arg min_{\beta \in \mathbb{R}^d}\{E[\mathcal{A}(X,\beta)] - E[\mathcal{B}(X,\beta)]\}$$

Sehingga perlu membuktikan:

$$\mathcal{A}(X,\beta) = E[(1\{\epsilon_\tau < \Delta_\tau(X,\beta)\} - \tau)\Delta_\tau(X,\beta)|X]$$
$$= E[E[(1\{\epsilon_\tau < \Delta_\tau(X,\beta)\} - \tau)|X]\Delta_\tau(X,\beta)]$$
$$= E\left[\left[F_{\epsilon_\tau}(\Delta_\tau(X,\beta)|X) - F_{\epsilon_\tau}(0|X)\right]\Delta_\tau(X,\beta)\right] \quad (15)$$
$$= E\left[\left(\int_0^1 f_{\epsilon_\tau}(u\Delta_\tau(X,\beta)|X)\,\Delta_\tau(X,\beta)du\right)\Delta_\tau(X,\beta)\right]$$

$$= E\left[\left(\int_0^1 f_{\epsilon_\tau}(u\Delta_\tau(X,\beta)|X)\,du\right).\Delta_\tau^2(X,\beta)\right] \tag{16}$$

Jika $u_\tau = \frac{\epsilon_\tau}{\Delta_\tau(X,\beta)}$, untuk nilai $\Delta_\tau(X,\beta) > 0$

$$\mathcal{B}(X,\beta) = E[(1\{\epsilon_\tau < \Delta_\tau(X,\beta)\} - 1\{\epsilon_\tau < 0\})\epsilon_\tau|X]$$

$$= E[1\{u_\tau \in [0,1]\}.u_\tau.\Delta_\tau(X,\beta)|X]$$

$$= E\big[[1\{u_\tau \in [0,1]\}.u_\tau|X]\Delta_\tau(X,\beta)\big]$$

$$= E\left[(\int_0^1 u f_{u_\tau}(u|X)du).\Delta_\tau(X,\beta)\right]$$

$$= E\left[\left(\int_0^1 u f_{\epsilon_\tau}(u\Delta_\tau(X,\beta)|X).\Delta_\tau(X,\beta)du\right).\Delta_\tau(X,\beta)\right]$$

$$= E\left[\left(\int_0^1 u f_{\epsilon_\tau}(u\Delta_\tau(X,\beta)|X)\,du\right).\Delta_\tau^2(X,\beta)\right]$$

$$= E\left[\left(\int_0^1 u f_{\epsilon_\tau}(u\Delta_\tau(X,\beta)|X)\right).\Delta_\tau^2(X,\beta)\right] \tag{17}$$

Dari persamaan (16) dan (17) maka

$E[\rho_\tau(Y - X'\beta)] - E\big(\rho_\tau\big(Y - Q_\tau(Y|X)\big)\big]$

$= \mathcal{A}(X,\beta) - \mathcal{B}(X,\beta)$

$= E\left[\left(\int_0^1 f_{\epsilon_\tau}(u\Delta_\tau(X,\beta)|X)\,du\right).\Delta_\tau^2(X,\beta)\right] - E\left[\left(\int_0^1 u f_{\epsilon_\tau}(u\Delta_\tau(X,\beta)|X)\right).\Delta_\tau^2(X,\beta)\right]$

$= \left[\left(\int_0^1 (1-u) f_{\epsilon_\tau}(u\Delta_\tau(X,\beta)|X)\,du\right).\Delta_\tau^2(X,\beta)\right]$

$= E[w_\tau(X,\beta).\Delta_\tau^2(X,\beta)]$

Terbukti.

Lampiran 9. Sertifikat Seminar Makalah Tugas Akhir dalam Konferensi Nasional Penelitian Matematika dan Pembelajarannya

PROGRAM STUDI PENDIDIKAN MATEMATIKA
FAKULTAS KEGURUAN DAN ILMU PENDIDIKAN
UNIVERSITAS MUHAMMADIYAH SURAKARTA

SERTIFIKAT

diberikan kepada

Febria Pradita Prima Andani

sebagai

PEMAKALAH

Judul

PERBANDINGAN REGRESI KUANTIL DAN TRANSFORMASI BOX-COX DALAM MENGATASI MASALAH HETEROSKEDASTISITAS

pada Konferensi Nasional Penelitian Matematika dan Pembelajarannya I dengan Tema:
"Tren dan Isu-isu Baru dalam Penelitian Matematika dan Pembelajarannya"
yang diselenggarakan oleh Program Studi Pendidikan Matematika FKIP UMS
di Auditorium Mohamad Djazman UMS, 12 Maret 2016

Rektor UMS,

Prof. Dr. Bambang Setiaji

Surakarta, 12 Maret 2016
Ketua Panitia,

Rita P. Khotimah, M.Sc