PENDETEKSIAN HETEROSKEDASTISITAS DENGAN PENGUJIAN KORELASI RANK SPEARMAN DAN TINDAKAN PERBAIKANNYA

Skripsi
disajikan sebagai salah satu syarat
untuk memperoleh gelar Sarjana Sains
Program Studi Matematika

Oleh
Layyinatus Syifa
4150404021

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI SEMARANG
2009

SURAT PERNYATAAN KEASLIAN TULISAN

Dengan ini saya menyatakan bahwa isi skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya tidak terdapat karya yang diterbitkan oleh orang lain, kecuali yang secara tertulis dirujuk dalam skripsi ini dan disebutkan dalam daftar pustaka.

Semarang, 27 Maret 2009

Layyinatus Syifa
NIM 4150404021

ABSTRAK

Syifa, Layyinatus. 2009. *Pendeteksian Heteroskedastisitas Dengan Pengujian Korelasi Rank Spearman dan Tindakan Perbaikannya*. Skripsi Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang. Prof. Dr. Y.L. Sukestiyarno. Drs. Arief Agoestanto, M.Si.

Kata kunci : Heteroskedastisitas, Korelasi rank Spearman, Transformasi.

Ordinary Least Square (OLS) merupakan suatu metode analisis regresi yang digunakan untuk memperoleh koefisien regresi yang meminimumkan jumlah kuadrat residu. Model regresi yang diperoleh dengan OLS merupakan estimator yang baik bila model regresi memenuhi asumsi model regresi linear klasik, salah satunya adalah homoskedastisitas. Penyimpangan asumsi homoskedastisitas atau yang disebut heteroskedastisitas, tidak merusak sifat kebiasan dan konsistensi dari penaksir OLS. Tetapi penaksir ini tidak lagi mempunyai varians minimum atau efisien sehingga membuat prosedur pengujian hipotesis yang biasa nilainya diragukan.

Permasalahan pada skripsi ini adalah bagaimana mendeteksi heteroskedastisitas dengan pengujian korelasi rank Spearman dan bagaimana tindakan perbaikannya jika terjadi heteroskedastisitas. Tujuan dari skripsi ini adalah mengetahui cara mendeteksi heteroskedastisitas dengan pengujian korelasi rank Spearman dan mengetahui bagaimana tindakan perbaikannya jika terjadi heteroskedastisitas.

Metode yang digunakan dalam penelitian ini adalah studi pustaka. Langkah pertama yang dilakukan dalam penelitian ini adalah menemukan masalah. Kemudian merumuskan masalah, selanjutnya dengan menggunakan pendekatan teoritik maka dapat ditemukan jawaban permasalahan sehingga tercapai tujuan penulisan skripsi.

Berdasarkan pembahasan hasil penelitian, diperoleh simpulan bahwa cara mendeteksi heteroskedastisitas dengan pengujian korelasi rank Spearman yaitu dengan mencari nilai koefisien korelasi rank Spearman (r_s) antara setiap variabel bebas dengan $|e|$, kemudian melakukan statistik uji dengan pengujian $t = \dfrac{r_s\sqrt{N-2}}{\sqrt{1-r_s^2}}$ dengan kriteria uji terjadi heteroskedastisitas jika nilai t hitung lebih dari nilai t kritis. Tindakan perbaikan untuk menghilangkan heteroskedastisitas dapat dilakukan dengan dua cara yaitu dengan mencari model regresi baru melalui prosedur metode kuadrat terkecil tertimbang jika σ_i^2

diketahui atau dengan melakukan transformasi jika σ_i^2 tidak diketahui. Setelah diperoleh model regresi yang baru harus diperiksa kembali apakah masih terjadi heteroskedastisitas atau tidak.

Dari hasil penelitian diperoleh saran bahwa jika pada suatu model regresi terjadi penyimpangan asumsi heteroskedastisitas, maka harus dilakukan tindakan perbaikan untuk menghilangkan heteroskedastisitas tersebut. Setelah dilakukan tindakan perbaikan harus dideteksi kembali apakah masih terjadi heteroskedastisitas atau tidak.

PENGESAHAN

Skripsi ini telah dipertahankan di hadapan sidang Panitia Ujian Skripsi FMIPA Unnes pada tanggal 16 Maret 2009.

Panitia:

Ketua	Sekretaris
Drs. Kasmadi Imam S, M.S.	Drs. Edy Soedjoko, M.Pd.
NIP 130781011	NIP 131693657
	Penguji
	Dr. Scolastika Mariani, M.Si.
	NIP 131931636
Penguji/Pembimbing I	Penguji/Pembimbing II
Prof. Dr. Y.L. Sukestiyarno	Drs. Arief Agoestanto, M.Si
NIP 131404322	NIP 132046855

MOTTO DAN PERSEMBAHAN

Motto :

Jika kita bangun dengan mengingat Tuhan dan tidur setelah memastikan bahwa seharian ini tak ada sedetikpun yang terlewatkan tanpa mengingat Tuhan, maka apapun bentuk kegagalan, kesuksesan, kesedihan ataupun kebahagiaan, akan terlihat sama nantinya, "a better human".

Persembahan

Skripsi ini ku persembahkan untuk

Ibu, ibu, ibu dan bapak,,,

Lovely Rae, Maz K_c0en dan Ataka

Terimakasih untuk setiap ketulusan kasih sayang, restu dan do'a.

KATA PENGANTAR

Puji syukur dipanjatkan ke hadirat Allah SWT atas limpahan rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan penulisan skripsi yang berjudul **"Pendeteksian Heteroskedastisitas Dengan Pengujian Korelasi Rank Spearman dan Tindakan Perbaikannya".**

Penulis menyadari bahwa dalam penyusunan skripsi ini, penulis banyak memperoleh bantuan dari berbagai pihak. Oleh karena itu, dalam kesempatan ini penulis menyampaikan terima kasih dan pengahargaan yang sebesar-besarnya kepada:

1. Prof. Dr. H. Sudijono Sastroatmodjo, M.Si., Rektor Universitas Negeri Semarang;
2. Drs. Kasmadi Imam S., M.S., Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang;
3. Drs. Edy Soedjoko, M.Pd., Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang;
4. Prof. Dr. Y.L. Sukestiyarno, Pembimbing Utama yang selalu memberikan bimbingan dan motivasi kepada penulis;
5. Drs. Arief Agoestanto, M.Si, Pembimbing Pendamping yang telah memberikan arahan dan motivasi kepada penulis;
6. Keluarga besar Universitas Negeri Semarang, FMIPA, Seluruh Dosen dan rekan-rekan di Jurusan Matematika yang selalu memberikan ilmu, bimbingan dan motivasinya.

7. Bapak, Ibu, kakak dan adik, semua keluarga besar yang tak henti-hentinya memberikan do’a, semangat, dan dukungan moral serta spiritual.
8. Sahabat-sahabatku di kost Griya Bunda, terima kasih untuk setiap dukungan, kebersamaan dan persahabatan.
9. My class mate, menjenk, blandonk, kucing dan erwin.
10. Teman-teman seperjuangan, semua mahasiswa matematika’04.
11. Semua pihak yang turut membantu dalam penyusunan skripsi ini, yang tidak dapat penulis sebutkan satu persatu.

Penulis mengucapkan terima kasih kepada semua pihak yang telah berkenan membaca skripsi ini. Semoga skripsi ini dapat bermanfaat bagi mahasiswa matematika khususnya dan bagi semua pihak pada umumnya.

Semarang, 27 Maret 2009

Penulis

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR LAMPIRAN

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Dalam kehidupan sehari-hari sering dijumpai adanya hubungan antara satu variabel yang disebut variabel tak bebas dengan satu atau lebih variabel yang disebut variabel bebas. Misalnya perubahan harga suatu barang akan bepengaruh terhadap jumlah yang diminta bagi barang tersebut, tekanan semacam gas bergantung pada temperatur, hasil produksi padi tergantung pada jumlah pupuk yang digunakan, banyak hujan, cuaca dan sebagainya.

Hubungan tersebut dapat dinyatakan dalam bentuk persamaan matematik yang menyatakan hubungan fungsional antara variabel-variabel yang disebut model regresi. Analisis regresi berkenaan dengan penaksiran dan/atau peramalan nilai rata-rata hitung atau nilai rata-rata (populasi) variabel tak bebas atas dasar nilai variabel bebas yang tetap (*fixed)* atau diketahui.

Model regresi linear yang paling sederhana berupa garis lurus antara variabel tak bebas dengan satu variabel bebas. Apabila model tersebut dikembangkan dengan beberapa variabel bebas, maka model regresi tersebut dikenal dengan model regresi linear berganda. Secara umum, model regresi dapat dituliskan sebagai $\boldsymbol{Y=X\beta+\varepsilon}$. Dimana $\boldsymbol{Y}$ menyatakan variabel tak bebas, $\boldsymbol{X}$ menyatakan variabel bebas, $\boldsymbol{\beta}$ merupakan parameter dan $\boldsymbol{\varepsilon}$ merupakan faktor galat (*error*). Dari model regresi ini dapat diestimasi besarnya pengaruh secara kuantitatif dari masing-masing variabel bebas terhadap variabel tak bebas. Selain itu juga dapat diestimasi nilai variabel tak bebas bila nilai variabel bebas telah diketahui.

Ada beberapa metode penyusunan model regresi, tetapi sejauh yang menyangkut analisis regresi, metode yang paling luas digunakan adalah metode kuadrat terkecil biasa (*method of ordinary least square, OLS*). Estimasi parameter model regresi yang diperoleh dengan OLS merupakan estimator yang baik bila model regresi memenuhi asumsi-asumsi tertentu yang sering disebut dengan asumsi model regresi linear klasik.

Satu asumsi kritis dari model regresi linear klasik adalah bahwa gangguan u_i (dalam penulisan skripsi ini yang dimaksud dengan gangguan u_i adalah *error* ε_i) diasumsikan semuanya mempunyai varians yang sama. Jika asumsi ini tidak dipenuhi, salah satunya kita mempunyai kasus heteroskedastisitas. Heteroskedastisitas tidak merusak sifat kebiasan dan konsistensi dari penaksir OLS. Tetapi penaksir ini tidak lagi mempunyai varians minimum atau efisien. Dengan perkataan lain, mereka tidak lagi BLUE (*Best Linear Unbiased Estimation*) (Gujarati, 1978:194). Maksudnya, dengan meningkatnya ukuran sampel sampai tak terhingga, penaksir tadi mengarah pada nilai sebenarnya (sifat konsistensi) tetapi variansnya tidak lagi minimum bahkan jika besarnya sampel meningkat secara tak terbatas. Ketiadaan efisiensi ini membuat prosedur pengujian hipotesis yang biasa nilainya diragukan.

Mengingat pentingnya model regresi pada suatu data maka perlu dideteksi apakah terjadi heterokedastisitas dan bagaimanakah cara mendeteksi heteroskedastisitas tersebut. Salah satunya adalah dengan pengujian korelasi rank spearman. Korelasi rank spearman merupakan salah satu ukuran korelasi berdasarkan rank yaitu ukuran korelasi yang diperoleh dengan mengganti observasi-observasi dengan ranking-ranking mereka. Jika ternyata pada suatu data terjadi heteroskedastisitas, maka diperlukan suatu usaha untuk memperbaiki heteroskedastisitas tersebut.

Berdasarkan latar belakang tersebut, maka penulis tertarik untuk mengadakan penelitian dengan judul " Pendeteksian Heteroskedastisitas dengan Pengujian Korelasi Rank Spearman dan Tindakan Perbaikannya".

1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, permasalahan yang akan diteliti adalah sebagai berikut :

(1) Bagaimana mendeteksi heteroskedastisitas dengan pengujian korelasi rank Spearman?

(2) Bagaimana tindakan perbaikannya jika terjadi heteroskedastisitas?

1.3 Tujuan Penelitian

Berdasarkan pada permasalahan yang ada, maka tujuan yang ingin dicapai dalam penulisan skripsi ini adalah sebagai berikut :

(1) Mendeteksi heteroskedastisitas dengan pengujian korelasi rank Spearman.

(2) Melakukan tindakan perbaikan jika terjadi heteroskedastisitas.

1.4 Manfaat Penelitian

Dalam penulisan skripsi ini diharapkan dapat memberikan manfaat antara lain sebagai berikut :

(1) Secara teoritis dapat memberikan sumbangan yang nyata terhadap perkembangan ilmu statistika terutama tentang heteroskedastisitas pada suatu model regresi linear.

(2) Secara praktis dapat memberikan jalan keluar bagi data yang tidak memenuhi asumsi homokedastisitas untuk dapat memperoleh estimasi parameter model regresi yang tebaik sehingga model regresi layak digunakan.

1.5 Sistematika Penulisan Skripsi

Secara garis besar skripsi ini terdiri dari tiga bagian yaitu bagian awal, bagian inti dan bagian akhir.

Bagian awal terdiri dari halaman sampul, halaman judul, abstrak, pengesahan, motto dan persembahan, daftar isi dan daftar lampiran.

Bagian inti terdiri dari lima bab yaitu:

BAB 1 : PENDAHULUAN

Bab ini berisi latar belakang masalah, perumusan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan skripsi.

BAB 2 : LANDASAN TEORI

Landasan teori berisi mengenai teori-teori yang mendukung dan berkaitan dengan permasalahan skripsi sehingga dapat dijadikan sebagai teori dan analisis data. Pada bagian ini terdiri dari matriks, model regresi linear, metode kuadrat terkecil, asumsi model regresi linear klasik, koefisien determinasi, heteroskedastisitas, rank, koefisien korelasi berdasarkan rank, pengujian heteroskedastisitas, metode kuadrat terkecil tertimbang dan kerangka berfikir.

BAB 3 : METODE PENELITIAN

Metode penelitian berisi tentang proses atau langkah penelitian. Bab ini meliputi pemilihan masalah, perumusan masalah, studi pustaka, pemecahan masalah dan penarikan kesimpulan.

BAB 4 : HASIL PENELITIAN DAN PEMBAHASAN

Pada bab ini berisi pembahasan dari permasalahan yang disajikan yang terbagi tiga sub bagian, yaitu pendeteksian heteroskedastisitas, tindakan perbaikan dan contoh kasus.

BAB 5 : PENUTUP

Bab ini berisi tentang simpulan dan saran yang diperoleh dari penelitian.

Bagian akhir skripsi berisi daftar pustaka dan lampiran-lampiran.

BAB 2

LANDASAN TEORI

2.1 Matriks

Asumsi homoskedastisitas adalah salah satu asumsi yang harus dipenuhi agar suatu model regresi layak digunakan. Dalam menentukan model regresi tersebut, notasi matriks banyak digunakan dalam menaksir nilai parameter. Oleh karena itu perlu adanya definisi matriks dan operasi dalam matriks.

Definisi 2.1

Sebuah matriks adalah susunan bilangan-bilangan yang membentuk segiempat siku-siku. Bilangan-bilangan dalam susunan tersebut dinamakan *entry* dalam matriks. (Anton, 1992:22)

Contoh : $\boldsymbol{A_{m\times n}} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$

Bilangan $a_{11}, a_{22}, a_{33}, \ldots, a_{mn}$ disebut elemen atau unsur dari matriks $\boldsymbol{A}$. Indeks pertama dari elemen menunjukkan baris dan indeks kedua menunjukkan kolom dimana elemen itu berada. Ukuran (*ordo*) sebuah matriks ditentukan oleh banyaknya baris dan kolom, karena matriks $\boldsymbol{A}$ tersebut mempunyai *m* baris dan *n* kolom, maka matriks *A* tersebut berukuran $m \times n$.

Sebuah matriks dengan n baris dan n kolom dinamakan matriks bujur sangkar n (*square matrix of order n*) dan unsur-unsur $a_{11}, a_{22}, a_{33}, \ldots, a_{nn}$ dikatakan berada pada diagonal utama dari $\boldsymbol{A}_{m\times n}$.

Definisi 2.2

Jika $\boldsymbol{A}$ adalah suatu matriks dan c adalah suatu skalar, maka hasil kali (*product*) $c\boldsymbol{A}$ adalah matriks yang diperoleh dengan mengalikan masing-masing entri dari $\boldsymbol{A}$ dengan c (Anton, 1992:24).

Contoh : $\boldsymbol{A} = \begin{bmatrix} 2 & -1 & 5 \\ 4 & 3 & 9 \\ 6 & -4 & 5 \end{bmatrix}$ dan $c = 2$

Maka $c\boldsymbol{A} = \begin{bmatrix} 2\times 2 & 2\times(-1) & 2\times 5 \\ 2\times 4 & 2\times 3 & 2\times 9 \\ 2\times 6 & 2\times(-4) & 2\times 5 \end{bmatrix} = \begin{bmatrix} 4 & -2 & 10 \\ 8 & 6 & 18 \\ 12 & -8 & 10 \end{bmatrix}$

Definisi 2.3

Jika $\boldsymbol{A}$ adalah matriks $m \times r$ dan $\boldsymbol{B}$ adalah matriks $r \times n$, maka hasil kali $\boldsymbol{AB}$ adalah matriks $m \times n$ yang entri-entrinya ditentukan sebagai berikut. Untuk mencari entri dalam baris i dan kolom j dari matriks $\boldsymbol{AB}$, pilih baris i dari matriks $\boldsymbol{A}$ dan kolom j dari matriks $\boldsymbol{B}$. Kalikanlah entri-entri yang bersesuaian dari baris dan kolom tersebut bersama-sama dan kemudian tambahkan hasil kali yang dihasilkan (Anton, 1992:25).

Contoh : $\boldsymbol{A} = \begin{bmatrix} 2 & -3 & 1 \\ 6 & 7 & 4 \end{bmatrix}$ dan $\boldsymbol{B} = \begin{bmatrix} 5 & -1 \\ 2 & 0 \\ 6 & 3 \end{bmatrix}$, maka :

$$AB = \begin{bmatrix} 2 & -3 & 1 \\ 6 & 7 & 4 \end{bmatrix} \begin{bmatrix} 5 & -1 \\ 2 & 0 \\ 6 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} (2\times5)+(-3\times2)+(1\times6) & (2\times-1)+(-3\times0)+(1\times3) \\ (6\times5)+(7\times2)+(4\times6) & (6\times-1)+(7\times0)+(4\times3) \end{bmatrix}$$

$$= \begin{bmatrix} 10 & 1 \\ 68 & 6 \end{bmatrix}$$

Definisi 2.4

Jika $\boldsymbol{A}$ adalah sebarang matriks berukuran $m \times n$ maka transpose $\boldsymbol{A}$ dinyatakan oleh $\boldsymbol{A^T}$ dan didefinisikan sebagai matriks berukuran $n \times m$ dengan kolom pertamanya adalah baris pertama dari $\boldsymbol{A}$ dan kolom kedua adalah baris kedua dari $\boldsymbol{A}$ dan seterusnya (Anton, 1992:27). Sifat dari matriks transpose: $(\boldsymbol{AB})^{\boldsymbol{T}}=\boldsymbol{B^T A^T}$, dengan $\boldsymbol{A}$ dan $\boldsymbol{B}$ suatu matriks.

Contoh : $\boldsymbol{A} = \begin{bmatrix} 1 & 6 \\ 2 & -3 \\ 5 & 4 \end{bmatrix}$ maka $\boldsymbol{A^T} = \begin{bmatrix} 1 & 2 & 5 \\ 6 & -3 & 4 \end{bmatrix}$

Definisi 2.5

Matriks bujur sangkar dengan diagonal utama memuat unsur-unsur $a_{11}, a_{22}, a_{33}, \ldots, a_{nn}$ serta semua unsur di atas dan di bawah diagonal utamanya nol $(a_{jk} = 0, j \neq k)$ disebut matriks diagonal. (Anton, 1992:29)

Matriks diagonal yang semua unsur diagonal utamanya adalah 1 disebut matriks satuan atau matriks identitas.

Definisi 2.6

Jika $\boldsymbol{A}$ adalah matriks bujur sangkar dan ditemukan matriks $\boldsymbol{B}$ sehingga $\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{I}$, dengan $\boldsymbol{I}$ matriks identitas maka $\boldsymbol{A}$ dapat dikatakan mempunyai invers dan $\boldsymbol{B}$ dikatakan invers dari $\boldsymbol{A}$ atau dapat dinyatakan $\boldsymbol{B}=\boldsymbol{A}^{-1}$. (Anton, 1992:34)

Definisi 2.7

Jika $\boldsymbol{A}$ adalah matriks $n \times n$ mempunyai invers maka untuk setiap matriks $\boldsymbol{B}$ yang berukuran $n \times 1$, sistem persamaan $\boldsymbol{AX}=\boldsymbol{B}$ mempunyai tepat satu penyelesaian yaitu $\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{B}$. (Anton, 1992:52)

2.2 Model Regresi Linear

Jika kita mempunyai data yang terdiri dari atas dua atau lebih variabel, adalah sewajarnya untuk mempelajari cara bagaimana variabel-variabel itu berhubungan. Hubungan yang didapat pada umumnya dinyatakan dalam bentuk persamaan matematik yang menyatakan hubungan fungsional antara variabel-variabel. Studi yang menyangkut masalah ini dikenal dengan analisis regresi. (Sudjana, 2001:310)

Hubungan fungsional yang dituliskan dalam bentuk persamaan matematik yang bergantung pada parameter-parameter disebut persamaan regresi. Dalam analisis regresi dibedakan dua jenis variabel yaitu variabel bebas (*independent*) dan variabel tak bebas (*dependent*). Penentuan variabel mana yang bebas dan mana yang tak bebas dalam beberapa hal tidak mudah dapat dilaksanakan. Studi yang cermat, diskusi yang seksama, berbagai pertimbangan, kewajaran masalah yang dihadapi dan pengalaman akan membantu memudahkan penentuan. Untuk keperluan analisis, variabel bebas akan dinyatakan dengan $X_1, X_2, \cdots, X_k, (k \geq 1)$ sedangkan variabel tak bebas dinyatakan dengan Y.

Bentuk umum model regresi linear yang mengandung *k* variabel bebas dapat dituliskan sebagai (Sembiring, 1995:134) :

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$$

Persamaan diatas adalah bentuk ringkas untuk sekumpulan *n* persamaan simultan berikut:

$$\begin{aligned} Y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_k x_{1k} + \varepsilon_1 \\ Y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \cdots + \beta_k x_{2k} + \varepsilon_2 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ Y_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_k x_{nk} + \varepsilon_n \end{aligned}$$

Dengan lambang matriks dapat ditulis menjadi :

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_k \end{bmatrix}$$

Sehingga bentuk umum model regresi linear dalam bentuk matriks dituliskan sebagai berikut:

$$\boldsymbol{Y=X\beta+\varepsilon}$$

dengan

$\boldsymbol{Y}$ = vektor variabel tak bebas ukuran $n \times 1$

$\boldsymbol{X}$ = matriks variabel bebas ukuran $n \times (k+1)$

$\boldsymbol{\beta}$ = vektor parameter ukuran $(k+1) \times 1$

$\boldsymbol{\varepsilon}$ = vektor *error* ukuran $n \times 1$

2.3 Metode Kuadrat Terkecil (Ordinary Least Square / OLS)

Menurut Gujarati (1978:34), metode kuadrat terkecil adalah suatu metode analisis regresi untuk memperoleh koefisien regresi yang memenuhi asumsi-asumsi sebagai berikut:

(1) Nilai rata-rata dari *error* random sama dengan nol, yaitu $E(\varepsilon_i) = 0$ untuk setiap $i = 1,2,\cdots,n$.

(2) Tidak ada korelasi antara *error* yang satu dengan *error* yang lain, berarti $Kov(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i, \varepsilon_j) = 0, i \neq j$.

(3) $Var(\varepsilon_i) = E(\varepsilon_i^2) = \sigma^2$ artinya setiap *error* mempunyai variansi yang sama.

(4) Variabel *error* berdistribusi normal.

Misalkan regresi garis populasi adalah $Y_i = \alpha + \beta X_i + \varepsilon_i$. Menurut Sumodiningrat (2002:106), yang dimaksud penaksiran α dan β dengan metode kuadrat terkecil adalah menemukan nilai-nilai taksiran $\hat{\alpha}$ dan $\hat{\beta}$ yang meminimumkan jumlah kuadrat residu : $\sum e_i^2$.

Dari garis regresi sampel $Y = \hat{\alpha} + \hat{\beta} X_i + e_i$ diperoleh :

$$e_i = Y_i - \left(\hat{\alpha} + \hat{\beta} X_i\right)$$

dan

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left(Y_i - \hat{\alpha} - \hat{\beta} X_i\right)^2 .$$

Nilai-nilai $\hat{\alpha}$ dan $\hat{\beta}$ yang meminimumkan jumlah kuadrat, diperoleh dengan menurunkan secara parsial fungsi kuadrat residu, $\sum e_i^2$, dan menyamakan turunan ini dengan nol.

$$\frac{\partial \sum e_i^2}{\partial \hat{\alpha}} = -2\sum \left(Y_i - \hat{\alpha} - \hat{\beta} X_i\right) = 0$$

$$\frac{\partial \sum e_i^2}{\partial \hat{\beta}} = -2\sum X_i \left(Y_i - \hat{\alpha} - \hat{\beta} X_i\right) = 0$$

atau

(2.1) $$\sum Y_i - n\hat{\alpha} - \beta \sum X_i = 0 \Leftrightarrow \sum Y_i = n\hat{\alpha} + \hat{\beta}\sum X_i$$

(2.2) $$\sum X_i Y_i - \hat{\alpha}\sum X_i - \hat{\beta}\sum X_i^2 = 0 \Leftrightarrow \sum X_i Y_i = \hat{\alpha}\sum X_i + \hat{\beta}\sum X_i^2$$

Bila kita nyatakan $\overline{X} = \frac{\sum X_i}{n}$ dan $\overline{Y} = \frac{\sum Y_i}{n}$ maka persamaan (2.1) memberikan

(2.3) $$\hat{\alpha} = \frac{\sum Y_i}{n} - \frac{\hat{\beta}\sum X_i}{n} = \overline{Y} - \hat{\beta}\overline{X}$$

Dengan menyubstitusikan nilai $\hat{\alpha}$ ke dalam persamaan (2.2), diperoleh :

$$\sum X_i Y_i = \left(\overline{Y} - \hat{\beta}\overline{X}\right)\sum X_i + \hat{\beta}\sum X_i^2$$

$$\Leftrightarrow \sum X_i Y_i = \left(\frac{\sum Y_i}{n} - \hat{\beta}\frac{\sum X_i}{n}\right)\sum X_i + \hat{\beta}\sum X_i^2$$

$$\Leftrightarrow \sum X_i Y_i - \frac{\sum Y_i \sum X_i}{n} - \hat{\beta}\left(\sum X_i^2 - \frac{\left(\sum X_i\right)^2}{n}\right) = 0$$

(2.4) $$\Leftrightarrow \hat{\beta} = \frac{\sum X_i Y_i - \frac{\sum Y_i \sum X_i}{n}}{\sum X_i^2 - \frac{\left(\sum X_i\right)^2}{n}}$$

Persamaan (2.4) dapat disederhanakan menjadi :

$$\hat{\beta} = \frac{\sum X_i Y_i - n\overline{YX}}{\sum X_i^2 - n\bar{X}^2}$$

$$= \frac{\sum X_i Y_i - n\overline{YX} - n\overline{YX} + n\overline{YX}}{\sum X_i^2 - 2n\bar{X}^2 + n\bar{X}^2}$$

$$= \frac{\sum X_i Y_i - \bar{X}\sum Y_i - \bar{Y}\sum X_i + n\overline{YX}}{\sum X_i^2 - 2n\left(\frac{\sum X_i}{n}\right)^2 + n\bar{X}^2}$$

$$= \frac{\sum X_i Y_i - \sum \bar{X} Y_i - \sum \bar{Y} X_i + \sum \overline{YX}}{\sum X_i^2 - 2\bar{X}\sum X_i + n\bar{X}^2}$$

(2.5)
$$\hat{\beta} = \frac{\sum \left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sum \left(X_i - \bar{X}\right)^2}$$

2.3.1 Metode Kuadrat Terkecil Dengan Matriks

Telah diketahui bahwa bentuk umum model regresi linear dalam bentuk matriks adalah **Y=Xβ+ε**, maka persamaan hasil estimasinya dapat ditulis $\mathbf{Y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e}$.

Diperoleh : $\mathbf{e} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$.

Sekarang dengan meminimumkan jumlah kuadrat residu, diperoleh :

$$JKS = \sum_{i=1}^{k} e_i^2 = e'e = \left(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right)'\left(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right)$$

$$= \mathbf{Y'Y} - \hat{\boldsymbol{\beta}}'\mathbf{X'Y} - \mathbf{Y'X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}'\mathbf{X'}\hat{\boldsymbol{\beta}}\mathbf{X}$$

$$= \mathbf{Y'Y} - 2\hat{\boldsymbol{\beta}}'\mathbf{X'Y} + \hat{\boldsymbol{\beta}}'\mathbf{X'X}\hat{\boldsymbol{\beta}}.$$

Perhatikan bahwa $\hat{\boldsymbol{\beta}}'\mathbf{X'Y} = \mathbf{Y'X}\hat{\boldsymbol{\beta}}$ karena keduanya skalar. Ada dua cara dalam menyelesaikan masalah ini. Pertama ialah dengan menurunkan *JKS* terhadap $\hat{\boldsymbol{\beta}}$ dan menyamakan turunannya dengan nol diperoleh :

$$\frac{\partial JKS}{\partial \hat{\beta}} = -2\mathbf{X'Y} + 2\mathbf{X'X}\hat{\boldsymbol{\beta}} = 0$$

$$\Leftrightarrow 2\mathbf{X'X}\hat{\boldsymbol{\beta}} = 2\mathbf{X'Y}$$

$$\Leftrightarrow \mathbf{X'X}\hat{\boldsymbol{\beta}} = \mathbf{X'Y}.$$

Jika **X'X** tidak singulir maka ada balikannya (inversnya). Jadi $\mathbf{X'X}\hat{\boldsymbol{\beta}} = \mathbf{X'Y}$ mempunyai penyelesaian tunggal, maka diperoleh nilai estimasi $\boldsymbol{\beta}$:

(2.6) $$\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'Y}$$

Sedangkan cara yang kedua lebih menguntungkan dari segi komputasi, cara ini disebut pemfaktoran **QR**. Suatu matriks sebarang, misal **A**, berukuran $n \times n$ selalu dapat diuraikan menjadi **A**=**QR**.

Dengan **Q'Q**=**I** dan **R** matriks segitiga yang semua unsurnya dibawah diagonal utama nol.

Penguraian **QR** sangat memudahkan perhitungan dalam metode kuadrat terkecil dan dengan ketelitian yang sangat tinggi. Lihat kembali peminimuman persamaan:

$$JKS = \mathbf{Y'Y} - 2\mathbf{Y'X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}'\mathbf{X'}\hat{\boldsymbol{\beta}}$$

Dalam hal ini **X** berukuran $n \times p$, dengan $p=k+1$. Misalkan **X**=**QR**, dengan $\mathbf{Q'Q} = \mathbf{I}_{\mathbf{p \times p}}$ dan **R** matriks segitiga berukuran $p \times p$. Ganti **X** dengan **QR** pada persamaan diatas kemudian tambahkan dan kurangi dengan **Y'QQ'Y**, maka kita peroleh:

$$\begin{aligned} JKS &= \mathbf{Y'Y} - 2\mathbf{Y'QR}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}'\mathbf{R'Q'QR}\hat{\boldsymbol{\beta}} \\ &= (\mathbf{Y'Y} - \mathbf{Y'QQ'Y}) + \left(\mathbf{Y'QQ'Y} - 2\mathbf{Y'QR}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}\mathbf{R'R}\hat{\boldsymbol{\beta}}\right) \\ &= \mathbf{Y'}(\mathbf{I} - \mathbf{QQ'})\mathbf{Y} + \left(\mathbf{Q'Y} - \mathbf{R}\hat{\boldsymbol{\beta}}\right)'\left(\mathbf{Q'Y} - \mathbf{R}\hat{\boldsymbol{\beta}}\right). \end{aligned}$$

Kedua suku di ruas kanan berbentuk kuadrat, suku pertama tidak mengandung parameter $\hat{\boldsymbol{\beta}}$, jadi tidak terpengaruh proses peminimuman. Jadi *JKS* akan minimum, yaitu nol, jadi jika

$$\mathbf{Q'Y} - \mathbf{R}\hat{\boldsymbol{\beta}} = 0,$$

atau

$$\mathbf{Q'Y} = \mathbf{R}\boldsymbol{\beta}.$$

Jadi diperoleh nilai estimasi $\boldsymbol{\beta}$:

$$\hat{\boldsymbol{\beta}} = \mathbf{R}^{-1}\mathbf{Q'Y}.$$

Nyatakan **R** dan **Q** kembali dalam **X** maka diperoleh:

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= \mathbf{R}^{-1}\mathbf{Q'Y} \\ &= \mathbf{R}^{-1}\left(\mathbf{XR}^{-1}\right)'\mathbf{Y}\end{aligned}$$

Karena dari **X=QR** diperoleh $\mathbf{Q} = \mathbf{XR}^{-1}$, maka

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= \mathbf{R}^{-1}\left(\mathbf{XR}^{-1}\right)'\mathbf{Y} \\ &= \mathbf{R}^{-1}\left(\mathbf{R}^{-1}\right)'\mathbf{X'Y} \\ &= \left(\mathbf{R'R}\right)^{-1}\mathbf{X'Y}.\end{aligned}$$

Karena **Q'Q=I** maka

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{R'Q'QR}\right)^{-1}\mathbf{X'Y}.$$

Karena **QR=X** dan **R'Q'=X'** maka

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X'X}\right)^{-1}\mathbf{X'Y}.$$

2.3.2 Sifat-Sifat Penaksir Kuadrat Terkecil

Menurut Sumodiningrat (2002:109), terdapat tiga sifat penaksir kuadrat terkecil, yaitu :

(1) Linear (*Linearity*)

Pandang kembali persamaan (2.5) :

$$\hat{\beta} = \frac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{\sum (X_i - \overline{X})^2}$$

$$\hat{\beta} = \frac{\sum Y_i (X_i - \overline{X}) - \overline{Y} \sum (X_i - \overline{X})}{\sum (X_i - \overline{X})^2}$$

Karena $\overline{Y} \sum (X_i - \overline{X}) = 0$, maka $\hat{\beta} = \frac{\sum Y_i (X_i - \overline{X})}{\sum (X_i - \overline{X})^2}$.

Misalkan $x_i = X_i - \overline{X}$, diperoleh : $\hat{\beta} = \frac{\sum Y_i x_i}{\sum x_i^2}$

$$\Leftrightarrow \hat{\beta} = \sum k_i Y_i \qquad (2.7)$$

dimana $k_i = \frac{x_i}{\sum x_i^2}$. Persamaan (2.7) menunjukkan bahwa $\hat{\beta}$ adalah penaksir linear karena merupakan fungsi linear dari *Y*.

Begitu juga dengan (2.3) yang memberikan :

$$\hat{\alpha} = \overline{Y} - \hat{\beta}\overline{X}$$

$$= \frac{\sum Y_i}{n} - \overline{X} \sum k_i Y_i$$

$$= \sum \left(\frac{1}{n} - \overline{X} k_i \right) Y_i \qquad (2.8)$$

Persamaan (2.8) menunjukkan bahwa $\hat{\alpha}$ juga merupakan fungsi linear dari *Y*.

(2) Tak Bias (*Unbiasedness*)

Dari (2.7) menunjukkan bahwa : $\hat{\beta} = \sum k_i Y_i$ dan telah didefinisikan sebelumnya bahwa : $Y_i = \alpha + \beta X_i + \varepsilon_i$, maka :

$$\hat{\beta} = \sum k_i(\alpha + \beta X_i + \varepsilon_i)$$

(2.9) $$= \alpha \sum k_i + \beta \sum k_i X_i + \sum k_i \varepsilon_i$$

Karena $k_i = \dfrac{x_i}{\sum x_i^2}$ dan $\sum k_i = \dfrac{\sum x_i}{\sum x_i^2} = 0$, maka :

$$\sum k_i X_i = \sum k_i(x_i + \overline{X})$$

$$= \sum k_i x_i + \sum k_i \overline{X}$$

$$= \frac{\sum x_i^2}{\sum x_i^2} + \overline{X} \sum k_i$$

$$= \frac{\sum x_i^2}{\sum x_i^2} \qquad \text{(karena } \textstyle\sum k_i = 0\text{)}$$

$$\sum k_i X_i = 1$$

Substitusikan nilai-nilai tersebut ke dalam (2.9) sehingga diperoleh :

(2.10) $$\hat{\beta} = \beta + \sum k_i \varepsilon_i$$

$$E\left[\hat{\beta}\right] = E[\beta] + \sum k_i(E[\varepsilon_i])$$

Dengan mengingat asumsi ε_i, maka sebagai akibatnya :

$$E\left[\hat{\beta}\right] = \beta \text{ .}$$

Persamaan (2.8) memberikan :

$$\hat{\alpha} = \sum \left(\frac{1}{n} - \overline{X} k_i \right) Y_i$$

$$= \sum \left(\frac{1}{\mathrm{n}} - \overline{\mathrm{X}}\mathrm{k}_{\mathrm{i}} \right) \left(\alpha + \beta \mathrm{X}_{\mathrm{i}} + \varepsilon_{\mathrm{i}} \right)$$

$$= \alpha + \beta \frac{1}{n} \sum X_i + \frac{1}{n} \sum \varepsilon_i - \alpha \overline{X} \sum k_i - \beta \overline{X} \sum k_i X_i - \overline{X} \sum k_i \varepsilon_i$$

$$= \alpha + \beta \overline{X} + \frac{1}{n} \sum \varepsilon_i - \beta \overline{X} - \overline{X} \sum k_i \varepsilon_i$$

(2.10) $$\hat{\alpha} = \alpha + \frac{1}{n} \sum \varepsilon_i - \overline{X} \sum k_i \varepsilon_i$$

Sehingga : $E[\hat{\alpha}] = \alpha + \frac{1}{n} \sum E[\varepsilon_i] - \overline{X} \sum k_i E[\varepsilon_i]$

$$E[\hat{\alpha}] = \alpha$$

(3) Varians Minimum dari $\hat{\alpha}$ dan $\hat{\beta}$

Sekarang akan dibuktikan $\hat{\alpha}$ dan $\hat{\beta}$ memiliki varians sampel terkecil dibandingkan dengan penaksir-penaksir linear tak bias lainnya. Pertama, akan dicari varians $\hat{\beta}$ kemudian akan dibuktikan bahwa variansnya minimum.

$$Var(\hat{\beta}) = E\left[(\hat{\beta} - \beta)^2 \right]$$

$$= E\left[\left(\sum k_i \varepsilon_i \right)^2 \right] \qquad \text{(dari persamaan 2.10)}$$

$$= E\left[k_1^2 \varepsilon_1^2 + k_2^2 \varepsilon_2^2 + \cdots + k_n^2 \varepsilon_n^2 + 2k_1 k_2 \varepsilon_1 \varepsilon_2 + \cdots + 2k_{n-1} k_n \varepsilon_{n-1} \varepsilon_n \right]$$

$$= E\left[k_1^2 \varepsilon_1^2 + k_2^2 \varepsilon_2^2 + \cdots + k_n^2 \varepsilon_n^2 \right] + E\left[2k_1 k_2 \varepsilon_1 \varepsilon_2 + \cdots + 2k_{n-1} k_n \varepsilon_{n-1} \varepsilon_n \right]$$

$$= E\left[\sum \left(k_i^2 \varepsilon_i^2 \right) \right] + 2E\left[\sum k_i k_j \varepsilon_i \varepsilon_j \right] i \neq j$$

$$= \sum k_i^2 E\left[\varepsilon_i^2 \right] + 2 \sum k_i k_j E\left[\varepsilon_i \varepsilon_j \right]$$

$$= \sigma^2 \sum k_i^2 \qquad \left(\text{karena } E\left[\varepsilon_i \varepsilon_j \right] = 0 \right)$$

Karena $k_i = \frac{x_i}{\sum x_i^2}$ dan $\sum k_i^2 = \frac{\sum x_i^2}{\left(\sum x_i^2 \right)^2} = \frac{1}{\sum x_i^2}$, maka :

(2.12) $$Var(\hat{\beta}) = \sigma^2 \sum k_i^2 = \frac{\sigma^2}{\sum x_i^2}$$

Sedangkan $Var(\hat{\alpha}) = E\left[(\hat{\alpha} - \alpha)^2\right]$

$$= E\left[\sum\left(\frac{1}{n} - \overline{X}k_i\right)\varepsilon_i\right]^2 \qquad \text{(dari persamaan (2.11)}$$

$$= \sigma^2 \sum\left(\frac{1}{n} - \overline{X}k_i\right)^2$$

$$= \sigma^2 \sum\left(\frac{1}{n^2} - \frac{2}{n}\overline{X}k_i + \overline{X}^2 k_i^2\right)$$

$$= \sigma^2\left(\frac{1}{n} - \frac{2\overline{X}}{n}\sum k_i + \overline{X}^2 \sum k_i^2\right)$$

$$= \sigma^2\left(\frac{1}{n} + \frac{\overline{X}^2}{\sum x_i^2}\right) \qquad \left(\text{karena} \sum k_i = 0 \text{ dan } \sum k_i^2 = \frac{1}{\sum x_i^2}\right)$$

$$= \sigma^2\left(\frac{\sum x_i^2 + n\overline{X}^2}{n\sum x_i^2}\right)$$

(2.13) $$Var(\hat{\alpha}) = \sigma^2\left(\frac{\sum X_i^2}{n\sum x_i^2}\right)$$

Untuk membuktikan bahwa $\hat{\beta}$ memiliki varians minimum, maka perlu dibandingkan varians $\hat{\beta}$ dengan beberapa penaksir β (katakanlah β^*) yang tidak bias.

Misalkan $\beta^* = \sum w_i Y_i$, dimana konstanta $w_i \neq k_i$ tetapi $w_i = k_i + c_i$, sehingga :

$$\beta^* = \sum w_i(\alpha + \beta X_i + \varepsilon_i)$$

(2.14) $$= \alpha\sum w_i + \beta\sum w_i X_i + \sum w_i \varepsilon_i$$

dan $E[\beta^*] = \alpha\sum w_i + \beta\sum w_i X_i$ $\qquad$ $(\text{karena } E[\varepsilon_i] = 0)$

Karena β^* diasumsikan penaksir yang tak bias, berarti pada persamaan diatas $\sum w_i = 0$ dan $\sum w_i X_i = 1$.

Tetapi $\sum w_i = \sum (k_i + c_i) = \sum k_i + \sum c_i$

karena $\sum c_i = 0$ dan $\sum k_i = \sum w_i = 0$

Maka $\sum w_i X_i = \sum (k_i + c_i) X_i = \sum k_i X_i + \sum c_i X_i$

karena $\sum c_i X_i = 0$, $\sum w_i X_i = 1$ dan $\sum k_i X_i = \sum k_i x_i = 1$

Juga $\sum c_i x_i = \sum c_i X_i + \bar{X} \sum c_i = 0$

Telah ditunjukkan jika β^* merupakan penaksir yang tak bias, hasil-hasil berikut berlaku :

(2.15) $\sum w_i = 0$, $\sum w_i X_i = 1$, $\sum c_i = 0$, $\sum c_i X_i = \sum c_i x_i = 0$

Varians dari β^* yang diasumsikan menjadi :

$$Var(\beta^*) = E\left[(\beta^* - \beta)^2\right]$$

$$= E\left[\left(\sum w_i \varepsilon_i\right)^2\right] \quad \text{(dari 2.14)}$$

$$= \sigma^2 \sum w_i^2 \quad \text{(dengan prosedur yang sama ketika mendapatkan Var}(\hat{\beta})\text{)}$$

Akan tetapi $\sum w_i^2 = \sum (k_i + c_i)^2 = \sum k_i^2 + 2\sum k_i c_i + \sum c_i^2$.

Sedangkan $\sum k_i c_i = \sum c_i \dfrac{x_i}{\sum x_i^2} = \dfrac{\sum c_i x_i}{\sum x_i^2} = 0$ (dengan 2.15)

$$\sum w_i^2 = \sum k_i^2 + \sum c_i^2$$

Sehingga : $Var(\beta^*) = \sigma^2 \left(\sum k_i^2 + \sum c_i^2\right) = \sigma^2 \sum k_i^2 + \sigma^2 \sum c_i^2$

$$Var(\beta^*) = Var(\beta) + \sigma^2 \sum c_i^2$$

Karena $\sum c_i^2$ selalu positif, maka $Var(\beta^*) > Var(\hat{\beta})$. Kecuali jika $\sum c_i^2 = 0$ (tapi tidak mungkin), maka $Var(\beta^*)$ akan sama dengan $Var(\hat{\beta})$. Hal ini menunjukkkan $\hat{\beta}$ memiliki sifat varians minimum.

Dengan cara yang sama dapat ditunjukkan bahwa konstanta intercept $(\hat{\alpha})$ metode kuadrat terkecil juga memiliki varians minimum. Misalkan α^* adalah sebuah penaksir baru yang diasumsikan fungsi linear dari Y_i dengan bobot $w_i = k_i + c_i$ seperti sebelumnya.

Kuadrat terkecil $\hat{\alpha}$ adalah : $\hat{\alpha} = \sum \left(\frac{1}{n} - \overline{X}k_i \right) Y_i$ (dari 2.8)

Dengan analogi, $\alpha^* = \sum \left(\frac{1}{n} - \overline{X}k_i \right) Y_i = f(Y)$

Karena diinginkan agar α^* menjadi penaksir tidak bias bagi α, maka $E[\alpha^*] = \alpha$. Substitusikan α^* ke dalam persamaan $Y_i = \alpha + \beta X_i + \varepsilon_i$ tetapi dalam notasi α^*, dan temukan nilai harapan dari α^*.

$$\alpha^* = \alpha\left(1 - X\sum w_i\right) + \beta\left(\overline{X} - \overline{X}\sum w_i X_i\right) + \sum \left(\frac{1}{n} - \overline{X}w_i \right) \varepsilon_i$$

$$E[\alpha^*] = \alpha\left(1 - \overline{X}\sum E[w_i]\right) + \beta\left(\overline{X} - \overline{X}\sum E[w_i X_i]\right) + \sum E\left[\frac{1}{n} - \overline{X}w_i \right] \varepsilon_i$$

Sekarang, $E[\alpha^*] = \alpha$ jika dan hanya jika $E[w_i] = 0$, $E[w_i X_i] = \frac{1}{n}$ dan $E[w_i \varepsilon_i] = 0$, yaitu jika $\sum w_i = 0$, $\sum w_i X_i = 1$ dan $\sum w_i \varepsilon_i = 0$. Syarat ini mengandung arti bahwa $\sum c_i = 0$ dan $\sum c_i X_i = 0$. Maka varians dari penaksir ini adalah :

$$Var(\alpha^*) = E[\alpha^* - \alpha]^2 = \sigma^2 \sum \left(\frac{1}{n} - \bar{X} w_i \right)^2$$

$$= \sigma^2 \sum \left(\frac{1}{n^2} + \bar{X}^2 w_i^2 - 2\frac{1}{n} \bar{X} w_i \right)$$

$$= \sigma^2 \left(\frac{n}{n^2} + \bar{X}^2 \sum w_i^2 - 2\bar{X} \frac{1}{n} \sum w_i \right)$$

$$= \sigma^2 \left(\frac{1}{n} + \bar{X}^2 \sum w_i^2 - \frac{2}{n} \bar{X} \sum w_i \right)$$

Karena $\sum w_i = 0$ dan $\sum w_i^2 = \sum k_i^2 + \sum c_i^2$, maka :

$$Var(\alpha^*) = \sigma^2 \left(\frac{1}{n} + \bar{X}^2 \left(\sum k_i^2 + \sum c_i^2 \right) \right)$$

$$Var(\alpha^*) = \sigma^2 \left(\frac{1}{n} + \frac{\bar{X}^2}{\sum x_i^2} \right) + \left(\sigma^2 \bar{X}^2 \sum c_i^2 \right)$$

Komponen pertama di sebelah kanan tanda sama dengan adalah varians dari $\hat{\alpha}$, sehingga :

$$Var(\alpha^*) = Var(\hat{\alpha}) + \sigma^2 \bar{X}^2 \sum c_i^2$$

Akan tetapi $\sum c_i^2 > 0$ karena tidak semua c_i adalah nol. Oleh karena itu, $Var(\alpha^*) > Var(\hat{\alpha})$. Sehingga terbukti bahwa penaksir-penaksir kuadrat terkecil model regresi linear merupakan penaksir-penaksir yang memenuhi syarat-syarat BLU (*Best, Linear, Unbiased*).

2.4 Asumsi Model Regresi Linear Klasik

Model regresi yang diperoleh dari OLS merupakan model regresi yang menghasilkan estimator linear tidak bias yang terbaik (Best Linear Unbias Estimator). Kondisi ini akan terjadi jika dipenuhi beberapa asumsi yang disebut dengan asumsi klasik, sebagai berikut (Algifari, 2000:83):

(1) Nonmultikolinearitas. Artinya antara variabel *independent* yang satu dengan *independent* yang lain dalam model regresi tidak saling berhubungan secara sempurna atau mendekati sempurna.

(2) Homoskedastisitas. Artinya varians semua variabel adalah konstan (sama).

(3) Nonautokorelasi. Artinya, tidak terdapat pengaruh dari variabel dalam model melalui tenggang waktu. Misalnya nilai suatu variabel saat ini akan berpengaruh terhadap nilai variabel lain pada masa yang akan datang.

(4) Nilai rata-rata kesalahan (*error*) populasi pada model stokastiknya sama dengan nol.

(5) Variabel *independent* adalah nonstokastik (nilainya konstan pada setiap kali percobaan yang dilakukan secara berulang).

(6) Distribusi kesalahan (*error*) adalah normal.

2.5 Koefisien Determinasi

Untuk mengetahui kecocokan data terhadap model regresinya dapat dilihat dari besar nilai koefisien determinasi, yang dinotasikan R^2. Nilai R^2 mendekati satu menunjukkan model regresinya semakin baik. Menurut Gujarati (1978:45), besaran R^2 yang dikenal sebagai koefisien determinasi (sampel) dan merupakan besaran yang paling lazim digunakan untuk mengukur kebaikan-suai (*goodness of fit*) garis regresi. Secara verbal, R^2 mengukur proporsi (bagian) atau prosentase total variansi dalam *Y* yang dijelaskan oleh model regresi.

Menurut Gujarati (1978:45), koefisien determinasi dihasilkan dari pembagian jumlah kuadrat regresi dengan jumlah kuadrat total, yang dapat ditulis sebagai berikut:

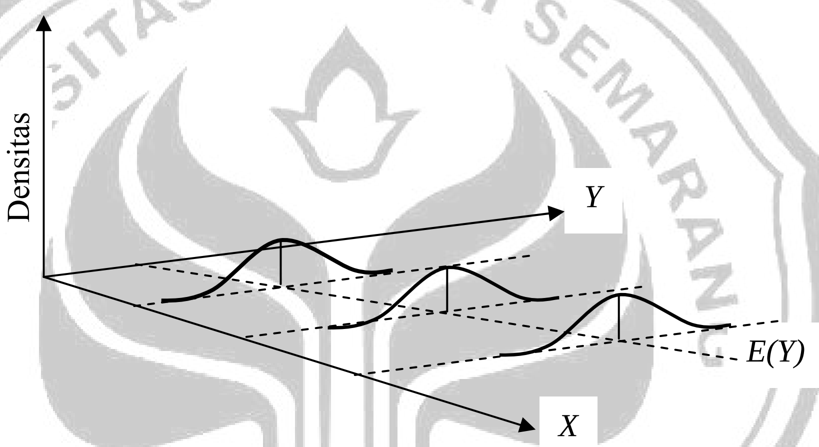
(2.16) $$R^2 = \frac{JKR}{JKT} = \frac{\sum_{i=1}^{n}\left(\hat{Y}_i - \bar{Y}\right)^2}{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2}$$

dengan *JKR* = Jumlah Kuadrat Regresi dan *JKT* = Jumlah Kuadrat Total.

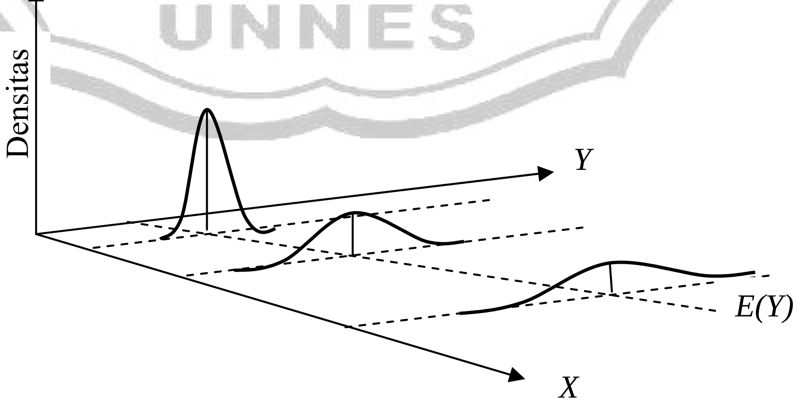
2.6 Heteroskedastisitas

Salah satu asumsi model klasik adalah asumsi homoskedastisitas yaitu asumsi yang menyatakan bahwa varian setiap ε_i di sekitar rerata nolnya tidak tergantung pada nilai variabel bebas. Varian setiap ε_i masih tetap sama baik untuk nilai-nilai *X* (variabel bebas) yang kecil maupun besar (Sumodiningrat, 2002:261). Dengan kata lain, varian tiap unsur *error* merupakan suatu bilangan konstan atau $\text{var}(\varepsilon_i) = E(\varepsilon_i^2) = \sigma^2$, $i = 1,2,3,\cdots,n$. Jika terjadi pelanggaran terhadap asumsi ini, maka model regresinya dikatakan berada dalam keadaan heterokedastisitas, yaitu varian tiap unsur *error* tidak konstan atau $\text{var}(\varepsilon_i) = E(\varepsilon_i^2) = \sigma_i^2$, $i = 1,2,3,\cdots,n$.

Untuk memperjelas perbedaan antara homoskedastisitas dan heteroskedastisitas, lihat gambar 2.1 dan gambar 2.2. Jika varians dari ε sama pada setiap titik atau untuk seluruh nilai *X*, maka pola yang tertentu akan terbentuk bila sebaran *Y* diplot dengan sebaran *X*. Bila digambarkan dalam tiga dimensi, polanya akan mendekati pola pada gambar 2.1. Sebaliknya, pada gambar 2.2, menunjukkan varian kondisional dari Y_i (yaitu ε_i) meningkat bersama dengan meningkatnya *X*. Situasi ini secara tepat dikenal sebagai heteroskedastisitas .



Gambar 2.1 Ilustrasi Homoskedastisitas



Gambar 2.2 Ilustrasi Heteroskedastisitas

Masalah heteroskedastisitas umum terjadi pada data *cross-section*, yaitu data yang diambil pada satu waktu saja tetapi dengan responden yang besar, misalnya jika kita melakukan survey. Sebagai contohnya, kita akan menganalisis data *cross-section* penjualan perusahaan-perusahaan dalam suatu industri. Variabel gangguan (*error terms*) akan sangat terkait dengan besar kecilnya perusahaan. Perusahaan yang besar akan mempunyai varian variabel gangguan yang besar, sebaliknya perusahaan yang kecil karena penjualan kemungkinan akan mempunyai varian variabel gangguan yang lebih kecil. Hal ini dikarenakan perusahaan yang lebih besar lebih fluktuatif daripada penjualan perusahaan kecil. Contoh yang lainnya adalah hubungan antara pendapatan dan pengeluaran konsumsi rumah tangga. Pengeluaran konsumsi rumah tangga kelompok kaya akan lebih fluktuatif daripada pengeluaran rumah tangga kelompok miskin (Widarjono, 2007: 126).

Konsekuensi adanya heteroskedastisitas dalam model regresi adalah penaksir (*estimator*) yang diperoleh tidak efisien, baik dalam sampel yang besar walaupun penaksir yang diperoleh menggambarkan populasinya (tidak bias) dan bertambahnya sampel yang digunakan akan mendekati nilai sebenarnya (konstan). Ini disebabkan oleh variansnya yang tidak minimum (tidak efisien) (Algifari, 2000:85). Jika varian tidak minimum maka menyebabkan perhitungan *standard error* metode OLS tidak lagi bisa dipercaya kebenarannya dalam pengujian hipotesis.

Dalam banyak kasus, kita tidak dapat mengetahui varian variabel gangguan. Katakanlah suatu model asli adalah : $Y_i = \alpha + \beta X_i + \varepsilon_i$, dimana ε_i memenuhi asumsi kecuali bahwa ε_i adalah heteroskedastik, yaitu :

$$\varepsilon_i = N\left(0, \sigma_i^2\right)$$
$$E\left[\varepsilon_i^2\right] = \sigma_i^2 = f\left(X_i\right)$$

Disini masalahnya, bagaimanakah bentuk dari $f\left(X_i\right)$. Secara empiris tidak mungkin mendapatkan bentuk pasti dari hubungan ini, maka yang dapat dilakukan hanyalah mengasumsikan kemungkinan tipe dari struktur heteroskedastik (Sumodiningrat, 2002:272).

(1) Asumsi 1

Diasumsikan bahwa pola varian variabel gangguan adalah proporsional dengan X_i^2, yaitu $Var\left(\varepsilon_i | X_i\right) = E\left[\varepsilon_i^2\right] = \sigma^2 X_i^2$.

(2) Asumsi 2

Diasumsikan bahwa pola varian variabel gangguan adalah proporsional dengan X_i, yaitu $Var\left(\varepsilon_i | X_i\right) = E\left[\varepsilon_i^2\right] = \sigma^2 X_i$.

(3) Asumsi 3

Diasumsikan bahwa pola varian variabel gangguan adalah proporsional terhadap rerata hitung kuadrat dari variabel terikat $\left(E\left[Y_i\right]\right)^2$, yaitu $Var\left(\varepsilon_i | X_i\right) = E\left[\varepsilon_i^2\right] = \sigma^2 \left(E\left[Y_i\right]\right)^2$.

Dalam memperbaiki heteroskedastisitas, beberapa asumsi tersebut akan digunakan ketika variabel gangguan tidak diketahui secara pasti.

2.7 Rank

Pandang peubah acak X_1, X_2, …,X_n yang masing-masing mempunyai nilai pengamatan $x_1, x_2, \ldots, x_n$ .nilai-nilai pengamatan ini diurutkan dari yang terkecil hingga yang terbesar kemudian diberi nomor 1, nomor 2, dan seterusnya yang terbesar diberi nilai *n*. Nomor-nomor urut tersebut adalah rank, yaitu bilamana yang diberikan pada setiap pengamatan sesuai dengan urutan besarnya peubah acaknya. Susunan keseluruhan rank disebut ranking, dimana setiap anggotanya memiliki nilai rank (Widiastuti, 2005:19). Misal diambil data sebagai berikut 8, 2, 4, 7, 3 kemudian data tersebut diurutkan dari yang terkecil ke yang terbesar sehingga diperoleh 2, 3, 4, 7, 8. Selanjutnya data yang telah diurutkan diberi rank masing-masing dari 1-5 dimana rank 1 = 2, rank 2 = 3, rank 3 = 4, rank 4 = 7, dan rank 5 = 8.

Misalkan dalam observasi terdapat data yang nilainya sama, maka observasi-observasi yang berangka sama diberi ranking rata-rata dari posisi yang seharusnya. Misal dipunyai data 2, 5, 7, 9, 5, kemudian kita urutkan datanya sehingga menjadi 2, 5, 5, 7, 9. Selanjutnya dari data yang telah diurutkan tersebut diberikan rank sehingga diperoleh rank 1 = 2, rank ke-2 dan rank ke-3 nilainya sama yaitu 5, sehingga untuk untuk rank 3 dan rank 4 kita jumlahkan ranknya kemudian kita bagi 2, sehingga diperoleh rank rata-rata yaitu $\frac{2+3}{2} = 2{,}5$ sebanyak dua kali. Jadi rank 2,5 = 5, rank 2,5 = 5, rank 4 = 7 dan rank 5 = 9.

2.8 Koefisien Korelasi Berdasarkan Rank

Koefisien korelasi merupakan salah satu ukuran untuk mengetahui keeratan hubungan (korelasi) linear antara 2 variabel. Misal dipunyai sampel random bivariat berukuran n yaitu (X_1, Y_1), (X_2, Y_2), … , (X_n, Y_n), maka ukuran korelasi antara X dan Y harus memenuhi syarat sebagai berikut :

(1) Nilai koefisien korelasi hanya antara -1 sampai dengan 1.

(2) Jika harga X makin besar berpasangan dengan Y yang juga semakin besar (dan sebaliknya), maka korelasi dikatakan positif atau korelasi ukuran positif. Hal seperti ini disebut hubungan langsung antara X dan Y.

(3) Jika harga X makin besar berpasangan dengan Y yang semakin kecil (dan sebaliknya), maka korelasi dikatakan negatif atau korelasi ukuran negatif. Hal seperti ini disebut hubungan invers antara X dan Y.

(4) Jika harga X tampak berpasangan secara acak dengan Y maka korelasi mendekati nol. Hal ini terjadi bila X dan Y independent, sehingga dapat dikatakan bahwa X dan Y tidak berhubungan.

Ukuran korelasi yang paling umum digunakan adalah koefisien korelasi Pearson's product moment (r), didefinisikan sebagai (Conover, 1971:251):

$$r = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\left[\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2\right]^{\frac{1}{2}}} \tag{2.17}$$

Disini r merupakan variabel acak, jika demikian maka r mempunyai fungsi distribusi. Fungsi distribusi r sangat bergantung pada distribusi (X,Y), maka r tidak dapat digunakan dalam statistik nonparametrik.

Kruskal telah menemukan cara untuk mencari ukuran korelasi yang fungsi distribusinya tidak tergantung kepada fungsi distribusi (*X,Y*), sehingga korelasi ini masuk dalam kawasan statistik nonparametrik. Ukuran korelasi yang dicari mendasarkan pada rank pengamatan atau dengan kata lain ukuran korelasi diperoleh dengan mengganti observasi-observasi dengan ranking-ranking mereka.

2.9 Pengujian Heteroskedastisitas

Berbagai pengujian disarankan untuk menyelidiki masalah heteroskedastisitas. Beberapa diantaranya adalah :

2.9.1 Pengujian Goldfeld-Quandt

Pengujian ini didasarkan atas dua asumsi dasar, yaitu : jumlah pengamatan sekurang-kurangnya dua kali jumlah variabel bebas dalam model dan ε_i adalah nonautokorelasi dan berdistribusi normal (Sumodiningrat, 2002:269). Uji Goldfeld-Quandt ini hanya untuk sample-sampel besar, dan meliputi langkah-langkah berikut:

(1) Langkah pertama

Susunlah pengamatan-pengamatan menurut besaran variabel bebas.

(2) Langkah kedua

Hilangkan sejumlah tertentu pengamatan yang ditengah-tengah (katakanlah *c*) dari analisis. Jumlah pengamatan sisanya, yaitu (*n*-*c*) pengamatan, kemudian dibagi menjadi dua bagian yang sama. Dengan demikian, masing-masing bagian terdiri dari $\frac{1}{2}(n-c)$ jumlah pengamatan. Satu bagian terdiri dari nilai-

nilai *X* yang kecil, sedangkan bagian lainnya mencakup nilai-nilai *X* yang besar.

(3) Langkah ketiga

Taksirlah regresi secara terpisah dengan prosedur OLS untuk setiap bagian, dan dapatkan jumlah residu kuadrat setiap bagian. Katakanlah $\sum e_1^2$ menunjukkan jumlah residu kuadrat dari sampel yang mengandung nilai-nilai *X* kecil, dan $\sum e_2^2$ dari sample yang mengandung nilai-nilai *X* besar.

(4) Langkah keempat

Hitunglah $F = \frac{\left(\sum e_2^2\right)}{\left(\sum e_1^2\right)}$; yang akan mempunyai distribusi F dengan derajat bebas $\left[\frac{1}{2}(n-c)-k\right]$ baik untuk pembilang maupun untuk penyebut dari ratio itu (*n*=jumlah pengamatan, *c*=jumlah pengamatan ditengah-tengah yang dihilangkan, dan *k*=jumlah parameter yang ditaksir).

Jika ε adalah homokedastik, maka kedua varian yaitu $\sum e_1^2$ dan $\sum e_2^2$ seharusnya sama, karena itu F akan cenderung sama dengan satu. F besar menunjukkan adanya heteroskedastisitas.

2.9.2 Pengujian Park

Prof. R.E. Park menyarankan sutu bentuk fungsi spesifik diantara σ_i^2 dan variabel bebas untuk menyelidiki adanya heteroskedastisitas :

$$\sigma_i^2 = \sigma^2 X_i^{\beta} e^{v_i}$$

$$\ln \sigma_i^2 = \ln \sigma^2 + \beta \ln X_i + v_i$$

Karena σ_i^2 biasanya tidak diketahui, Park menyarankan untuk menggunakan e_i^2 sebagai pendekatan dan melakukan regresi berikut (Gujarati, 1978:86):

$$\begin{aligned}\ln e_i^2 &= \ln \sigma^2 + \beta \ln X_i + v_i \\ &= \alpha + \beta \ln X_i + v_i\end{aligned}$$

Menurut Park, jika β ternyata signifikan secara statistik, maka berarti terdapat heteroskedastisitas di dalam data. Apabila ternyata tidak signifikan, maka kita dapat menerima asumsi homoskedastisitas.

Uji Park ini merupakan prosedur dua langkah :

Langkah 1 : lakukan regresi OLS tanpa memperdulikan heteroskedastisitas.

Langkah 2 : lakukan regresi log-linear antara e_i^2 dan X_i dan ujilah apakah β signifikan atau tidak.

Meskipun secara empiris menarik, uji Park mempunyai beberapa masalah. Goldfeld dan Quandt telah mengemukakan pendapat bahwa unsur kesalahan v_i mungkin tidak memenuhi asumsi OLS dan mungkin dengan sendirinya menjadi heteroskedastik.

2.9.3 Pengujian Glejser

Pengujian Glejser serupa dengan pengujian Park. Perbedaannya hanyalah Glejser menyarankan tujuh bentuk fungsi sebagai ganti dari hanya satu bentuk fungsi yang disarankan oleh Park. Bentuk-bentuk fungsi yang disarankan Glejser untuk menyelidiki adanya heteroskedastisitas adalah sebagai berikut (Sumodiningrat, 2002:271) :

$$|e_i| = \beta X_i + v_i$$

$$|e_i| = \beta \sqrt{X_i} + v_i$$

$$|e_i| = \frac{\beta}{X_i} + v_i$$

$$|e_i| = \frac{\beta}{\sqrt{X_i}} + v_i$$

$$|e_i| = \alpha + \beta X_i + v_i$$

$$|e_i| = \sqrt{(\alpha + \beta X_i) + v_i}$$

$$|e_i| = \sqrt{(\alpha + \beta X_i^2) + v_i}$$

dimana v_i adalah faktor kesalahan. Jika β pada regresi-regresi tersebut diatas adalah signifikan, maka bararti ada heteroskedastisitas di dalam data.

Sekali lagi sebagai persoalan praktis atau empiris, orang bisa menggunakan pendekatan Glejser. Tetapi Goldfeld dan Quandt menunjukkkan bahwa unsur kesalahan v_i mempunyai beberapa masalah dalam hal nilai yang diharapkan (*expected value*) tidak sama dengan nol, berkorelasi secara parsial dan ironiknya bersifat heteroskedastisitas.

2.9.4 Pengujian Korelasi Rank Spearman

Korelasi rank Spearman didefinisikan sebagai (Gujarati, 1978:188) :

(2.18) $$r_s = 1 - \frac{6\sum_{i=1}^{N} d_i^2}{N^3 - N}$$

dimana d_i = perbedaan dalam rank yang ditepatkan untuk dua karakteristik yang berbeda dari individual atau fenomena ke-*i* dan *N*= banyaknya individual atau fenomena yang di rank. Koefisien rank korelasi tadi dapat digunakan untuk mendeteksi heteroskedastisitas dengan langkah-langkah sebagai berikut :

(1) Langkah pertama

Estimasi *Y* (variabel tak bebas) terhadap *X* (variabel bebas) untuk mendapatkan residu-residu (*e*) yang merupakan taksiran bagi faktor-faktor galat (ε).

(2) Langkah kedua

Dengan mengabaikan tanda dari *e*, yaitu dengan mengambil nilai mutlaknya $|e|$, ranking harga mutlak $|e|$ dan *X* sesuai dengan urutan yang meningkat atau menurun dan menghitung koefisien korelasi rank Spearman yang telah diberikan sebelumnya tadi.

(3) Langkah ketiga

Dengan mengasumsikan bahwa koefisien rank korelasi populasi ρ_s adalah nol dan $N > 8$, tingkat signifikan dari r_s yang disampel dapat diuji dengan pengujian *t* sebagai berikut :

(2.19) $$t = \frac{r_s\sqrt{N-2}}{\sqrt{1-r_s^2}}$$

dengan derajat kebebasan = *N-2* . jika nilai *t* yang dihitung melebihi nilai kritis, kita bisa menerima hipotesis adanya heteroskedastisitas. Jika model regresi meliputi lebih dari satu variabel *X*, r_s dapat dihitung antara $|e|$ dan tiap-tiap variabel *X* secara terpisah dan dapat diuji dengan pengujian *t* yang diberikan di atas.

2.10 Metode kuadrat Terkecil Tertimbang

Metode kuadrat terkecil tertimbang pada prinsipnya sama dengan metode kuadrat terkecil, bedanya pada metode kuadrat terkecil tertimbang terdapat penambahan variabel baru, yaitu variabel *w* yang menunjukkan bobot atau timbangan.

Metode kuadrat terkecil meminimumkan :

$$\sum \varepsilon_i^2 = \sum_{i=1}^{k} \left(Y_i - \hat{Y}_i\right)^2$$

untuk mendapatkan taksiran, sedangkan metode kuadrat terkecil tertimbang meminimumkan jumlah kuadrat *error* tertimbang (Gujarati, 1978:200) :

$$\sum w_i \varepsilon_i^2 = \sum_{i=1}^{k} w_i \left(Y_i - \hat{Y}_i\right)^2$$

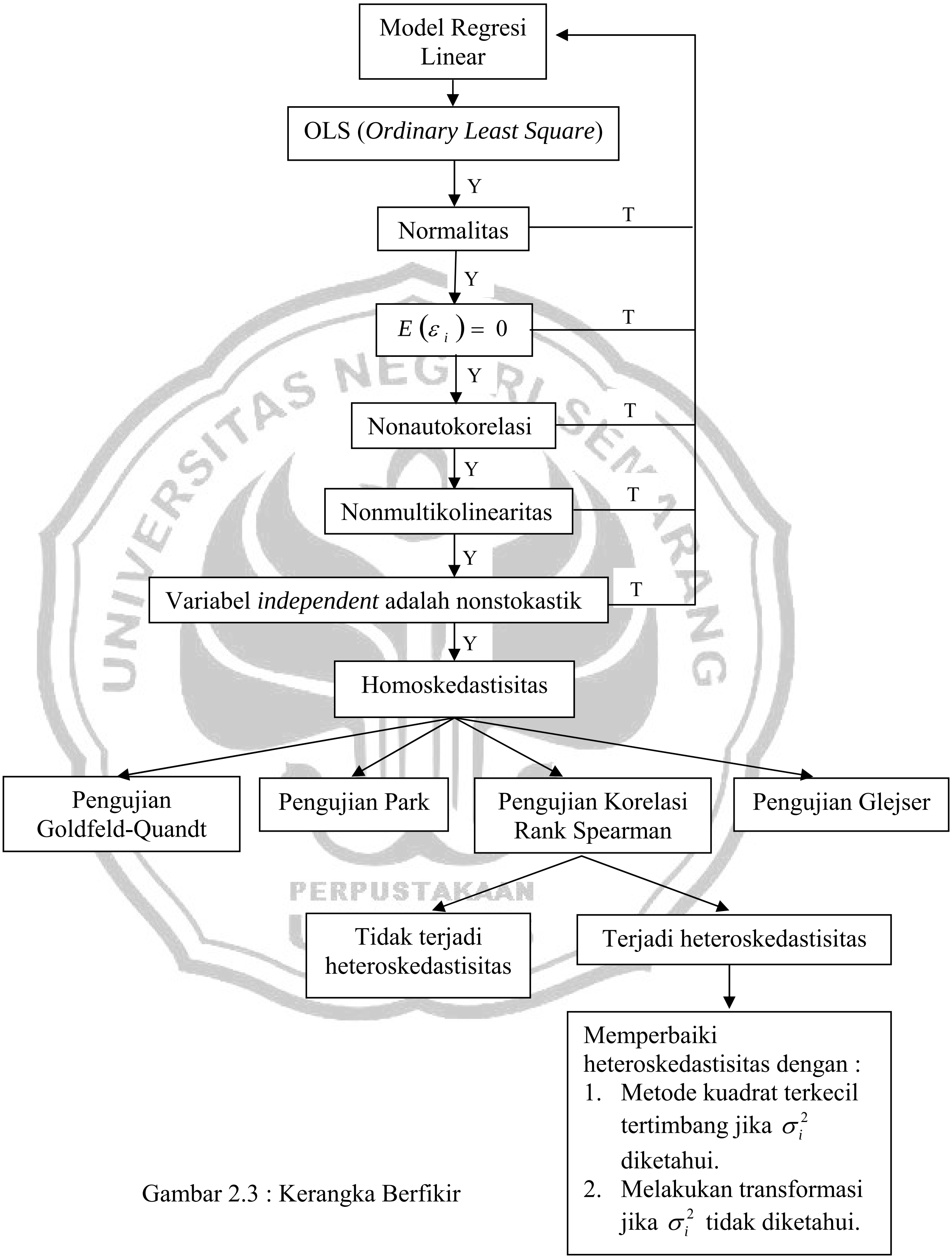
$$= \sum_{i=1}^{k} w_i \left[Y_i - \left(b_0 + b_1 x_1 + \cdots + b_k x_k\right)\right]^2 \qquad (2.20)$$

dimana $b_0, b_1, \cdots, b_k$ adalah penaksir kuadrat terkecil tertimbang dan dimana timbangannya w_i dengan $w_i = \dfrac{1}{\sigma_i^2}$.

2.11 Kerangka Berfikir

Dalam penyusunan model regresi linear, metode kuadrat terkecil biasa (*method of ordinary least square, OLS*) merupakan metode yang paling luas digunakan. Estimasi parameter model regresi yang diperoleh dengan OLS merupakan estimator yang baik bila model regresi memenuhi asumsi model regresi linear klasik, salah satunya adalah asumsi homoskedastisitas. Jika terjadi penyimpangan asumsi ini, atau yang disebut dengan heteroskedastisitas, maka penaksir yang diperoleh tidak lagi mempunyai varian yang minimum sehingga prosedur pengujian hipotesis yang biasa nilainya diragukan.

Dapat dilihat pada gambar 2.3 bahwa pengujian Goldfeld-Quandt, pengujian Park, pengujian Geljser dan pengujian korelasi rank Spearman merupakan cara untuk mendeteksi heteroskedastisitas. Dari uraian diketahui bahwa pengujian Goldfeld-Quandt hanya dapat digunakan untuk sampel-sampel besar, sedangkan untuk pengujian Park dan Glejser, Goldfeld dan Quandt telah mengemukakan pendapat tentang beberapa masalah berkaitan dengan unsur kesalahan v_i. Oleh karena itu, pada penelitian ini pengujian yang digunakan untuk mendeteksi heteroskedastisitas adalah pengujian korelasi rank Spearman. Selain dapat digunakan pada sampel besar maupun kecil, langkah-langkahnya yang relatif sederhana memungkinkan setiap orang untuk menggunakannya sekalipun bagi mereka yang tidak mempunyai keahlian khusus di bidang matematika. Dalam pengujian ini diambil hipotesis awal dengan H_0 = terjadi heteroskedastisitas. Kita bisa menerima hipotesis adanya heteroskedastisitas jika nilai *t* hitung melebihi nilai *t* kritis. Jika ternyata terjadi heteroskedastisitas, maka ada dua pendekatan yang dapat dilakukan untuk memperbaiki heteroskedastisitas tersebut, yaitu dengan metode kuadrat terkecil tertimbang jika σ_i^2 diketahui dan melakukan transformasi jika σ_i^2 tidak diketahui.



Gambar 2.3 : Kerangka Berfikir

BAB 3

METODE PENELITIAN

Peranan metode dalam suatu penelitian sangatlah penting sehingga dengan metode penelitian dapat mencapai tujuan penelitian yang telah ditetapkan. Melalui metode penelitian, masalah yang dihadapi dapat diatasi dan dipecahkan dari perolehan data atau informasi yang telah dikumpulkan.

Langkah-langkah yang dilakukan pada penelitian ini meliputi beberapa hal yaitu sebagai berikut:

3.1 Pemilihan Masalah

Dalam tahap ini dilakukan pencarian sumber pustaka dan memilih bagian dalam sumber pustaka tersebut yang dapat dijadikan permasalahan yang kemudian dijadikan bahan dasar untuk melakukan penelitian lebih lanjut.

3.2 Perumusan Masalah

Perumusan masalah diperlukan untuk membatasi permasalahan sehingga diperoleh kajian yang jelas. Sehingga akan lebih mudah untuk menentukan langkah dalam memecahkan permasalah tersebut.

3.3 Studi Pustaka

Setelah diperoleh permasalan untuk diteliti, penulis melakukan studi pustaka. Studi pustaka adalah penelaahan sumber pustaka yang relevan, digunakan untuk mengumpulkan data informasi yang diperlukan dalam penelitian. Studi pustaka diawali dengan mengumpulkan sumber pustaka yang berupa buku atau literatur, jurnal dan sebagainya. Setelah pustaka terkumpul dilanjutkan dengan pemahaman isi sumber pustaka tersebut yang pada akhirnya sumber pustaka ini dijadikan landasan untuk menganalisis permasalahan.

3.4 Pemecahan Masalah

Setelah permasalahan dirumuskan dan sumber pustaka terkumpul, langkah selanjutnya adalah pemecahan masalah melalui pengkajian secara teoritis yang selanjutnya disusun secara rinci dalam bentuk pembahasan.

Dalam pembahasan masalah dilakukan beberapa langkah pokok yaitu sebagai berikut:

(1) Menjelaskan tentang koefisien korelasi rank Spearman (r_s).

(2) Menjelaskan bagaimana cara mendeteksi heteroskedastisitas dengan pengujian korelasi rank Spearman.

(3) Melakukan tindakan perbaikan jika terjadi heteroskedastisitas, yaitu dengan cara :

(a) Melakukan prosedur metode kuadrat terkecil tertimbang jika σ_i^2 diketahui.

(b) Melakukan transformasi pada model regresi jika σ_i^2 tidak diketahui.

(4) Memberikan contoh kasus dalam mendeteksi heteroskedastisitas dan tindakan perbaikannya. Pada contoh kasus ini digunakan software SPSS 15 untuk mencari model regresi awal.

3.5 Penarikan Kesimpulan

Langkah terakhir dalam kegiatan penelitian ini adalah menarik kesimpulan dari keseluruhan permasalahan yang telah dirumuskan dengan berdasarkan pada landasan teori dan hasil pemecahan.

BAB 4

HASIL PENELITIAN DAN PEMBAHASAN

4.1 Hasil Penelitian

Sebuah model regresi dengan heteroskedastisitas mengandung konsekuensi serius pada estimator metode kuadrat terkecil / *ordinary least square* (*OLS*) karena tidak lagi BLUE (*Best Linear Unbiased Estimator*). Oleh karena itu sangat penting bagi kita untuk mengetahui apakah suatu model regresi mengandung unsur heteroskedastisitas atau tidak. Pada bab ini akan dibahas mengenai bagaimana mendeteksi heteroskedastisitas dengan pengujian korelasi rank Spearman dan tindakan perbaikannya jika terjadi heteroskedastisitas. Selanjutnya untuk mempermudah pemahaman mengenai pembahasan diberikan contoh kasus.

4.1.1 Pendeteksian Heteroskedastisitas dengan Pengujian Korelasi Rank Spearman

Sebelum penjelasan lebih lanjut tentang langkah-langkah penggunaan korelasi rank Spearman dalam mendeteksi heteroskedastisitas, akan dijelaskan terlebih dahulu bagaimana diperoleh persamaan korelasi rank Spearman dari persamaan umum suatu koefisien korelasi.

4.1.1.1 Korelasi Rank Spearman

Korelasi rank Spearman didasarkan atas pemikiran bahwa, misalkan N individu diranking menurut dua variabel. Misalnya: kita ingin mengatur sekelompok siswa dalam urutan berdasarkan skor-skor mereka pada tes masuk perguruan tinggi, dan juga dalam urutan berdasarkan indeks prestasi mereka pada akhir tahun pertama. Jika ranking pada skor tes masuk itu dinyatakan sebagai $X_1, X_2, \ldots, X_N$ dan ranking indeks prestasi dinyatakan sebagai $Y_1, Y_2, \ldots, Y_N$, maka kita dapat menggunakan suatu ukuran korelasi rank untuk menetapkan hubungan antara *X* dan *Y*.

Dapat kita lihat bahwa korelasi antara rank skor tes masuk perguruan tinggi dan indeks prestasi akan sempurna jika dan hanya jika $X_i = Y_i$ untuk semua *i*. Oleh sebab itu kita menggunakan selisih-selisih $d_i = X_i - Y_i$ sebagai petunjuk perbedaan antara kedua himpunan ranking itu. Ukuran besar berbagai d_i ini berkenaan mengenai seberapa erat hubungan antara skor ujian masuk dengan indeks prestasi. Jika hubungan antara kedua himpunan rank itu sempurna, setiap d_i akan sama dengan nol. Penjabaran rumus untuk menghitung r_s cukup sederhana, yaitu sebagai berikut :

Jika $x = X - \overline{X}$ dimana $\overline{X}$ *mean* skor pada variabel *X*, dan jika $y = Y - \overline{Y}$, maka rumus umum suatu koefisien korelasi pada persamaan (2.17) dapat kita tuliskan kembali sebagai berikut :

(4.1) $$r = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}}$$

Sekarang bila *X* dan *Y* adalah harga-harga ranking, maka jumlah *N* bilangan bulat $1, 2, \ldots, N$ adalah :

$$\sum X = \frac{N(N+1)}{2}$$

maka jumlah kuadrat bilangan-bilangan itu $1^2, 2^2, \dots, N^2$ dapat ditunjukkan sebagai:

$$\sum X^2 = \frac{N(N+1)(2N+1)}{6}.$$

Oleh sebab itu $\sum x^2 = \sum \left(X - \overline{X}\right)^2 = \sum X^2 - \frac{\left(\sum X\right)^2}{N}$

$$= \frac{N(N+1)(2N+1)}{6} - \frac{\left(\frac{N(N+1)}{2}\right)^2}{N}$$

$$= \frac{N(N+1)(2N+1)}{6} - \frac{N(N+1)^2}{4} = \frac{N^3 - N}{12}$$

demikian pula $\sum y^2 = \frac{N^3 - N}{12}$.

Diketahui bahwa $d = x - y$.

Maka $d^2 = (x-y)^2 = x^2 - 2xy + y^2$ dan $\sum d^2 = \sum x^2 + \sum y^2 - 2\sum xy$

Tetapi persamaan (4.1) menyatakan bahwa :

$$r = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}} = r_s$$

jika observasi-observasi itu diranking. Oleh sebab itu,

$$\sum d^2 = \sum x^2 + \sum y^2 - 2r_s\sqrt{\sum x^2 \sum y^2}$$

dan dengan demikian

(4.2) $$r_s = \frac{\sum x^2 + \sum y^2 - \sum d^2}{2\sqrt{\sum x^2 \sum y^2}}$$

Dengan X dan Y dalam rank, kita dapat menyubstitusikan nilai $\sum x^2 = \frac{N^3 - N}{12} = \sum y^2$ ke dalam persamaan (4.2) dan diperoleh :

$$r_s = \frac{\frac{N^3 - N}{12} + \frac{N^3 - N}{12} - \sum d^2}{2\sqrt{\left(\frac{N^3 - N}{12}\right)\left(\frac{N^3 - N}{12}\right)}}$$

$$= \frac{2\left(\frac{N^3 - N}{12}\right) - \sum d^2}{2\left(\frac{N^3 - N}{12}\right)}$$

$$= 1 - \frac{\sum d^2}{\frac{N^3 - N}{6}}$$

$$r_s = 1 - \frac{6\sum d^2}{N^3 - N}$$

Jadi diperoleh persamaan korelasi rank Spearman

$$r_s = 1 - \frac{6\sum_{i=1}^{N} d_i^2}{N^3 - N} \tag{4.3}$$

Pada pendeteksian heteroskedastisitas, variabel yang digunakan untuk menghitung koefisien korelasi rank Spearmannya adalah variabel bebas dan nilai mutlak $|e|$.

4.1.1.2 Langkah-Langkah Pendeteksian

Langkah-langkah penggunaan korelasi rank Spearman dalam mendeteksi heteroskedastisitas adalah sebagai berikut :

(1) Estimasi Y (variabel tak bebas) terhadap X (variabel bebas) untuk mendapatkan residu-residu (e) yang merupakan taksiran bagi faktor-faktor galat (ε).

(2) Cari nilai absolut residu, $|e|$, kemudian diranking dari nilai yang paling besar atau dari nilai yang paling kecil. Lakukan hal yang sama untuk variabel bebas (X) dan selanjutnya menghitung koefisien korelasi rank Spearman (r_s).

(3) Mengambil hipotesis :

H_0 = tidak terjadi heteroskedastisitas

H_1 = terjadi heteroskedastisitas

(4) Mencari nilai statistik t hitung dengan pengujian t sebagai berikut :

$t = \frac{r_s\sqrt{N-2}}{\sqrt{1-r_s^2}}$ dengan derajat kebebasan $db = N\text{-}2$.

(5) Kriteria uji :

Menolak H_0 jika nilai t hitung lebih dari nilai kritis.

Jika model regresi meliputi lebih dari satu variabel X, r_s dapat dihitung antara $|e|$ dan tiap-tiap variabel X secara terpisah dan dapat diuji dengan pengujian t yang diberikan di atas.

4.1.2 Tindakan Perbaikan

Diketahui bahwa heteroskedastisitas tidak merusak sifat kebiasan dan konsistensi dari penaksir OLS, tetapi penaksir tadi tidak lagi efisien yang membuat prosedur pengujian hipotesis yang biasa nilainya diragukan. Oleh karena itu diperlukan suatu tindakan perbaikan pada model regresi untuk menghilangkan masalah heteroskedastisitas pada model regresi tersebut. Tindakan perbaikan ini tergantung dari pengetahuan kita tentang varian dari variabel gangguan. Ada dua pendekatan untuk melakukan tindakan perbaikan, yaitu jika σ_i^2 diketahui dan jika σ_i^2 tidak diketahui.

4.1.2.1 Jika σ_i^2 Diketahui

Jika σ_i^2 diketahui atau dapat ditaksir, metode yang paling berkaitan dengan heteroskedastisitas adalah dengan menggunakan metode kuadrat terkecil tertimbang (*weighted least square*). Pada metode kuadrat terkecil meminimumkan $\sum \varepsilon_i^2$ untuk mendapatkan taksiran, sedangkan pada metode kuadrat terkecil tertimbang meminimumkan jumlah kuadrat residual tertimbang :

$$\sum w_i \varepsilon_i^2 = \sum_{i=1}^{k} w_i \left(Y_i - \hat{Y}_i\right)^2$$

$$= \sum_{i=1}^{k} w_i \left[Y_i - \left(b_0 + b_1 x_1 + \cdots + b_k x_k\right)\right]^2$$

dimana $b_0, b_1, \cdots, b_k$ adalah penaksir kuadrat terkecil tertimbang dan dimana timbangannya w_i dengan $w_i = \dfrac{1}{\sigma_i^2}$.

Untuk menggambarkan metode ini, perhatikan model regresi sampel dua variabel berikut :

$$Y_i = \alpha + \beta X_i + e_i$$

Dengan meminimumkan jumlah kuadrat residual tertimbang, diperoleh :

(4.4) $$\sum w_i e_i^2 = \sum w_i \left(Y_i - \alpha^* - \beta^* X_i\right)^2$$

Dengan mendeferensialkan (4.4) terhadap α^* dan β^*, diperoleh :

$$\frac{\partial \sum w_i e_i^2}{\partial \alpha^*} = -2\sum w_i \left(Y_i - \alpha^* - \beta^* X_i\right)$$

$$\frac{\partial \sum w_i e_i^2}{\partial \beta^*} = 2\sum w_i \left(Y_i - \alpha^* - \beta^* X_i\right)\left(-X_i\right)$$

Dengan menetapkan persamaan-persamaan diatas dengan nol, maka diperoleh persamaan normal berikut :

$$-2\sum w_i \left(Y_i - \alpha^* - \beta^* X_i\right) = 0$$

$$\Leftrightarrow \sum w_i Y_i - \alpha^* \sum w_i - \beta^* \sum w_i X_i = 0$$

(4.5) $$\Leftrightarrow \sum w_i Y_i = \alpha^* \sum w_i + \beta^* \sum X_i$$

$$2\sum w_i \left(Y_i - \alpha^* - \beta^* X_i\right)\left(-X_i\right) = 0$$

$$\Leftrightarrow -\sum w_i X_i Y_i + \alpha^* \sum w_i X_i + \beta^* \sum w_i X_i^2 = 0$$

(4.6) $$\Leftrightarrow \sum w_i X_i Y_i = \alpha^* \sum w_i X_i + \beta^* \sum w_i X_i^2$$

Nyatakan $\overline{X}^* = \frac{\sum w_i X_i}{\sum w_i}$ dan $\overline{Y}^* = \frac{\sum w_i Y_i}{\sum w_i}$, maka persamaan (4.5) memberikan :

$$\alpha^* = \frac{\sum w_i Y_i}{\sum w_i} - \frac{\beta^* \sum w_i X_i}{\sum w_i}$$

$$\Leftrightarrow \alpha^* = \overline{Y}^* - \beta^* \overline{X}^*$$

Dengan menyubstitusikan nilai α^* kedalam persamaan (4.6), diperoleh :

$$\sum w_i X_i Y_i = \left(\overline{Y}^* - \beta^* \overline{X}^*\right) \sum w_i X_i + \beta^* \sum w_i X_i^2$$

$$\Leftrightarrow \sum w_i X_i Y_i = \left(\frac{\sum w_i Y_i}{\sum w_i} - \beta^* \frac{\sum w_i X_i}{\sum w_i}\right) \sum w_i X_i + \beta^* \sum w_i X_i^2$$

$$\Leftrightarrow \sum w_i X_i Y_i - \frac{\sum w_i Y_i \sum w_i X_i}{\sum w_i} - \beta^* \left(\sum w_i X_i^2 - \frac{\left(\sum w_i X_i\right)^2}{\sum w_i} \right) = 0$$

(4.7)
$$\Leftrightarrow \beta^* = \frac{\sum w_i X_i Y_i - \dfrac{\sum w_i Y_i \sum w_i X_i}{\sum w_i}}{\sum w_i X_i^2 - \dfrac{\left(\sum w_i X_i\right)^2}{\sum w_i}}$$

Persamaan (4.7) dapat disederhanakan menjadi :

$$\beta^* = \frac{\sum w_i X_i Y_i - \left(\sum w_i\right)\bar{Y}^* \bar{X}^*}{\sum w_i X_i^2 - \left(\sum w_i\right)\bar{X}^{*2}}$$

$$= \frac{\sum w_i X_i Y_i - \left(\sum w_i\right)\bar{Y}^* \bar{X}^* - \left(\sum w_i\right)\bar{Y}^* \bar{X}^* + \left(\sum w_i\right)\bar{Y}^* \bar{X}^*}{\sum w_i X_i^2 - 2\left(\sum w_i\right)\bar{X}^{*2} + \left(\sum w_i\right)\bar{X}^{*2}}$$

$$= \frac{\sum w_i X_i Y_i - \bar{X}^*\left(\sum w_i Y_i\right) - \bar{Y}^*\left(\sum w_i X_i\right) + \left(\sum w_i\right)\bar{Y}^* \bar{X}^*}{\sum w_i X_i^2 - 2\left(\sum w_i\right)\left(\dfrac{\sum w_i X_i}{\sum w_i}\right)^2 + \left(\sum w_i\right)\bar{X}^{*2}}$$

$$= \frac{\sum w_i X_i Y_i - \sum w_i \bar{X}^* Y_i - \sum w_i \bar{Y}^* X_i + \sum w_i \bar{Y}^* \bar{X}^*}{\sum w_i X_i^2 - 2\bar{X}^* \sum w_i X_i + \left(\sum w_i\right)\bar{X}^{*2}}$$

(4.8)
$$\beta^* = \frac{\sum w_i \left(X_i - \bar{X}^*\right)\left(Y_i - \bar{Y}^*\right)}{\sum w_i \left(X_i - \bar{X}^*\right)^2}$$

4.1.2.2 Jika σ_i^2 Tidak Diketahui

Pada kenyataanya sulit kita mengetahui besarnya varian variabel gangguan. Jika hal itu terjadi, maka tindakan perbaikan yang dapat dilakukan adalah dengan melakukan transformasi pada masing-masing asumsi kemungkinan tipe dari struktur heteroskedastik.

(1) Asumsi 1

Heteroskedastisitas berbentuk : $Var(\varepsilon_i | X_i) = E[\varepsilon_i^2] = \sigma^2 X_i^2$.

Dalam kasus ini diasumsikan bahwa pola varians variabel gangguan adalah proporsional dengan X_i^2 , maka transformasi yang dibutuhkan adalah:

$$\frac{Y_i}{X_i} = \alpha \frac{1}{X_i} + \beta + \frac{\varepsilon_i}{X_i}$$

$$= \alpha \frac{1}{X_i} + \beta + v_i$$

dimana $v_i = \frac{\varepsilon_i}{X_i}$ adalah faktor gangguan baru yang telah ditransformasi. Untuk menyelidiki apakah faktor-faktor gangguan v_i homoskedastik atau tidak, maka harus diperoleh varians dari $\left(\frac{\varepsilon_i}{X_i}\right)$.

$$Var\left(\frac{\varepsilon_i}{X_i}\right) = E\left[\frac{\varepsilon_i}{X_i}\right]^2 = \frac{1}{X_i^2} E[\varepsilon_i^2].$$

Karena telah diasumsikan bahwa $E[\varepsilon_i^2] = \sigma^2 X_i^2$, maka :

$$Var\left(\frac{\varepsilon_i}{X_i}\right) = \frac{1}{X_i^2} \sigma^2 X_i^2 = \sigma^2$$

Terbukti bahwa faktor gangguan yang baru di dalam model memiliki sebuah varians konstan tertentu. Oleh karena itu, OLS dapat diterapkan pada model transformasi, yaitu sebagai berikut :

$$\frac{Y_i}{X_i} = \alpha \frac{1}{X_i} + \beta + \frac{\varepsilon_i}{X_i}$$

Dalam transformasi ini, posisi dari koefisien-koefisien telah berubah. Parameter variabel $\left(\frac{1}{X_i}\right)$ dalam model transformasi merupakan intercept pada model aslinya, sedangkan faktor konstanta dalam model transformasi merupakan parameter dari variabel bebas *X* pada model aslinya. Oleh karena itu, untuk memperoleh kembali model aslinya harus mengalikan regresi dengan X_i.

(2) Asumsi 2

Heteroskedastisitas berbentuk : $Var(\varepsilon_i | X_i) = E[\varepsilon_i^2] = \sigma^2 X_i$.

Dalam kasus ini diasumsikan bahwa pola varians variabel gangguan adalah proporsional dengan X_i, maka transformasi yang dibutuhkan seharusnya adalah :

$$\frac{Y_i}{\sqrt{X_i}} = \frac{\alpha}{\sqrt{X_i}} + \frac{\beta X_i}{\sqrt{X_i}} + \frac{\varepsilon_i}{\sqrt{X_i}}$$

atau :

$$\frac{Y_i}{\sqrt{X_i}} = \alpha \frac{1}{\sqrt{X_i}} + \beta \sqrt{X_i} + v_i$$

dimana $v_i = \frac{\varepsilon_i}{\sqrt{X_i}}$ adalah faktor gangguan baru yang telah ditransformasi dan $X_i > 0$. Untuk menyelidiki apakah faktor-faktor gangguan dalam bentuk transformasi homoskedastik atau tidak, maka harus diperoleh varians dari $\left(\frac{\varepsilon_i}{\sqrt{X_i}}\right)$.

$$Var\left(\frac{\varepsilon_i}{\sqrt{X_i}}\right) = E\left[\frac{\varepsilon_i}{\sqrt{X_i}}\right]^2 = \frac{1}{X_i}E\left[\varepsilon_i^2\right]$$

Karena telah diasumsikan bahwa $E\left[\varepsilon_i^2\right] = \sigma^2 X_i$, maka :

$$Var\left(\frac{\varepsilon_i}{X_i}\right) = \frac{1}{X_i}\sigma^2 X_i = \sigma^2 .$$

Jadi faktor gangguan dalam model transformasi ini adalah homoskedastik. Dengan kata lain, OLS dapat diterapkan pada model transformasi :

$$\begin{aligned} \frac{Y_i}{\sqrt{X_i}} &= \frac{\alpha}{\sqrt{X_i}} + \beta\sqrt{X_i} + \frac{\varepsilon_i}{\sqrt{X_i}} \\ &= \alpha\frac{1}{\sqrt{X_i}} + \beta\sqrt{X_i} + v_i \end{aligned}.$$

Tidak ada faktor intercept dalam model transformasi ini. Oleh karena itu, harus digunakan model regresi yang melalui titik nol dalam menaksir α dan β. Untuk memperoleh kembali model aslinya harus mengalikan taksiran regresi itu dengan $\sqrt{X_i}$.

(3) Asumsi 3

Heteroskedastisitas berbentuk : $Var\left(\varepsilon_i | X_i\right) = E\left[\varepsilon_i^2\right] = \sigma^2\left(E\left[Y_i\right]\right)^2$.

Dalam kasus ini diiasumsikan bahwa pola varians variabel gangguan adalah proporsional terhadap rerata hitung kuadrat dari variabel terikat $\left(E\left[Y_i\right]\right)^2$, dimana $E\left[Y_i\right] = \alpha + \beta X_i$, maka transformasi yang dibutuhkan adalah :

$$\frac{Y_i}{\alpha + \beta X_i} = \frac{\alpha}{\alpha + \beta X_i} + \frac{\beta X_i}{\alpha + \beta X_i} + \frac{\varepsilon_i}{\alpha + \beta X_i} \tag{4.9}$$

$$= \frac{\alpha}{\alpha + \beta X_i} + \frac{\beta X_i}{\alpha + \beta X_i} + v_i$$

dimana $v_i = \dfrac{\varepsilon_i}{\alpha + \beta X_i}$ adalah faktor gangguan baru yang telah ditransformasi dan $X_i > 0$. Untuk menyelidiki apakah faktor-faktor gangguan dalam bentuk transformasi homoskedastik atau tidak, maka harus diperoleh varians dari $\dfrac{\varepsilon_i}{\alpha + \beta X_i}$.

$$Var\left(\frac{\varepsilon_i}{\alpha + \beta X_i}\right) = E\left[\frac{\varepsilon_i}{\alpha + \beta X_i}\right]^2$$

$$= \left(\frac{1}{\alpha + \beta X_i}\right)^2 E\left[\varepsilon_i^2\right]$$

$$= \left(\frac{1}{\alpha + \beta X_i}\right)^2 \sigma^2 \left(\alpha + \beta X_i\right)^2$$

$$= \sigma^2$$

Faktor gangguan yang baru adalah homoskedastik. Namun pada model transformasi yang digambarkan pada (4.9) diatas tidak operasional dalam kasus ini. Hal ini disebabkan nilai-nilai α dan β tidak diketahui. Tetapi karena regresi $\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i$ bisa diperoleh, maka transformasi dapat dilakukan melalui dua langkah berikut :

Pertama, lakukan regresi OLS biasa tanpa memperhatikan heteroskedastisitas yang terkandung dalam data dan mendapatkan $\hat{Y}_i$. Dengan menggunakan taksiran $\hat{Y}_i$, kita mentransformasikan model sebagai berikut :

(4.10) $$\frac{Y_i}{\hat{Y}_i} = \alpha\left(\frac{1}{\hat{Y}_i}\right) + \beta\left(\frac{X_i}{\hat{Y}_i}\right) + \frac{\varepsilon_i}{\hat{Y}_i}.$$

Kedua, lakukan regresi (4.10) untuk mendapatkan α dan β. Secara umum, jika heteroskedastisitas berbentuk $E\left[\varepsilon_i^2\right] = \sigma_i^2 = \sigma^2 f\left(X_i\right)$, maka versi transformasi dari model bisa diperoleh dengan membagi keseluruhan komponen model aslinya dengan $\sqrt{f\left(X_i\right)}$.

4.2 Pembahasan

Heteroskedastisitas pada suatu data dapat menimbulkan konsekuensi serius pada estimator OLS karena estimator yang diperoleh dengan metode OLS tidak lagi mempunyai varians yang minimum sehingga menyebabkan perhitungan *standard error* metode OLS tidak lagi bisa dipercaya kebenarannya dalam pengujian hipotesis. Oleh karena itu sangat penting untuk mengetahui ada tidaknya heteroskedastisitas pada suatu data dan pendeteksian ini dapat dilakukan dengan pengujian korelasi rank Spearman.

Heteroskedastisitas sering kali terjadi pada data *cross-section* seperti dalam contoh kasus ini berikut ini, yaitu data jumlah tenaga kerja dan output yang dihasilkan industri ISIC 3 digit tahun 1993 (Widarjono, 2007:151). Data ini merupakan data *cross-section* karena data ini terdiri dari 30 jenis industri dan masing-masing jenis industri tentu mempunyai skala yang berbeda-beda sehingga tingkat penyerapan tenaga kerja juga berbeda-beda. Datanya sebagai berikut :

	Klasifikasi industri besar dan sedang ISIC 3 digit	Jumlah tenaga kerja (orang)	Jumlah output yang dihasilkan (rupiah)
311	Makanan	369248	18740153851
312	Makanan	143493	3488229886
313	Minuman	21127	829310486
314	Pengolahan tembakau dan	184304	8215641238

	bumbu rokok		
321	Tekstil	580519	14182392561
322	Pakaian jadi, kecuali alas kaki	350039	6300671710
323	Kulit dan barang dari kulit	23296	510143485
324	Alas kaki	230901	4415114453
331	Kayu, bambu, rotan, rumput	378093	11626808024
332	Perabotan dan perlengkapan rumah tangga	123428	1420820377
341	Kertas, barang dari kertas	74060	3663693185
342	Percetakan dan penerbitan	48757	1287837008
351	Bahan kimia industri	60112	5637636226
352	Kimia lain	100194	5433798845
354	Barang-barang hasil kilang minyak bumi dan batu bara	772	49824840
355	Karet dan barang-barang dari karet	121505	3557350068
356	Barang plastik	119574	2779130487
361	Porselin	38691	803480230
362	Gelas dan barang-barang dari gelas	20121	878305620
363	Semen, kapur dan barang-barang dari semen dan kapur	42102	2292618803
364	Pengolahan tanah liat	29226	123058917
369	Barang galian lain bukan logam	18222	419136783
371	Logam dasar besi dan baja	31465	6163470615
372	Logam dasar bukan besi	12047	1241100474
381	Barang dari logam kecuali mesin dan peralatannya	117748	3821375468
382	Mesin dan perlengkapannya kecuali mesin listrik	36158	1522240396
383	Mesin, peralatan dan	107172	5689820639

perlengkapan listrik serta bahan keperluan listrik		
384 Alat angkutan	100226	6495988172
385 Peralatan profesional, ilmu pengetahuan, pengukuran dan pengatur	6278	199302084
390 Pengolahan lainnya	70500	1043723314

Dari data diatas, akan diselidiki apakah terjadi heteroskedastisitas atau tidak. Jika terjadi heteroskedastisitas, maka akan dilakukan tindakan perbaikan.

Penyelesaian :

(1) Pendeteksian heteroskedastisitas :

Dengan menggunakan program SPSS 15.0 diperoleh model regresinya sebagai berikut (lihat lampiran 2 pada tabel *Coefficients*[a]) :

$$\hat{Y} = 12258{,}693 + 0{,}000026X \; .$$

Mencari nilai koefisien korelasi rank Spearman (nilai d^2 dapat dilihat pada lampiran 3) :

$$r_s = 1 - \frac{6\sum_{i=1}^{N} d_i^2}{N^3 - N}$$

$$= 1 - \frac{6 \times 822}{30^3 - 30}$$

$$= 0{,}817$$

Maka diperoleh nilai t_{hitung} :

$$t = \frac{r_s\sqrt{N-2}}{\sqrt{1-r_s^2}}$$

$$= \frac{0{,}817\left(\sqrt{30-2}\right)}{\sqrt{1-(0{,}817)^2}}$$

$$= 7{,}501$$

Dari distribusi t (lampiran 7) dengan $\alpha = 5\%$, diperoleh nilai $t_{kritis} = t_{(1-\alpha,df)} = t_{(0,95;28)} = 1{,}70$. Karena nilai t_{hitung} = 7,501 > t_{kritis} = 1,70 maka H_0 ditolak. Artinya terjadi heteroskedastisitas.

(2) Tindakan perbaikan :

Dari model regresi awal diperoleh persamaan sebagai berikut : $\hat{Y} = 12258{,}693 + 0{,}000026X$. Jika diasumsikan heteroskedastisitas berbentuk $Var(\varepsilon_i|X_i) = E[\varepsilon_i^2] = \sigma^2 X_i^2$, maka transformasi yang sesuai adalah :

$$\frac{Y_i}{X_i} = \alpha\frac{1}{X_i} + \beta + \frac{\varepsilon_i}{X_i}.$$

Variabel-variabel yang telah ditransformasikan $\left(\frac{Y}{X} \text{ dan } \frac{1}{X}\right)$ disajikan pada lampiran 4. Dari variabel-variabel yang telah ditransformasikan tersebut diperoleh model regresi baru sebagai berikut (lihat lampiran 2 pada tabel *Coefficients*[a)] : $\frac{Y}{X} =$ 7,68E-10 $\frac{1}{X}$ + 2,17E-05.

Untuk mendapatkan model aslinya, maka taksiran model regresi diatas harus dikalikan dengan *X*, sehingga diperoleh : $Y^* =$ 7,68E-10 + 2,17E-05 *X*.

Selanjutnya akan diperiksa apakah model regresi setelah dilakukan transformasi masih terdapat heteroskedastisitas atau tidak. Dari model tersebut, diperoleh nilai koefisien korelasi Spearman (nilai d^2 dapat dilihat pada lampiran 6) :

$$r_s = 1 - \frac{6\sum_{i=1}^{N} d_i^2}{N^3 - N}$$

$$= 1 - \frac{6 \times 4788}{30^3 - 30}$$

$$= -0,0652$$

Maka diperoleh nilai t_{hitung} :

$$t = \frac{r_s\sqrt{N-2}}{\sqrt{1-r_s^2}}$$

$$= \frac{(-0,0652)\left(\sqrt{30-2}\right)}{\sqrt{1-\left(-0,0652\right)^2}}$$

$$= -0,346$$

Dari distribusi t (lampiran 7) dengan $\alpha = 5\%$, diperoleh nilai $t_{kritis} = t_{(1-\alpha,df)} = t_{(0,95;28)} = 1,70$. Karena nilai t_{hitung} = -0,346 < t_{kritis} = 1,70 maka H_0 diterima. Artinya tidak terjadi heteroskedastisitas. Sehingga model regresi $Y^* =$ 7,68E-10 + 2,17E-05 X layak digunakan.

BAB 5

PENUTUP

5.1 Simpulan

Berdasarkan hasil pembahasan pada penilitian ini, simpulan yang dapat diambil adalah:

(1) Pendeteksian heteroskedastisitas dengan pengujian korelasi rank Spearman yaitu dengan mencari nilai koefisien korelasi rank Spearman (r_s) untuk setiap variabel bebas dengan $|e|$ kemudian melakukan statistik uji dengan pengujian $t = \frac{r_s\sqrt{N-2}}{\sqrt{1-r_s^2}}$ dengan kriteria uji terjadi heteroskedastisitas jika nilai t hitung lebih dari nilai t kritis. Kasus heteroskedastisitas ini sering terjadi pada data *cross-section*.

(2) Tindakan perbaikan untuk menghilangkan heteroskedastisitas dapat dilakukan dengan dua cara yaitu dengan mencari model regresi baru melalui prosedur metode kuadrat terkecil tertimbang jika σ_i^2 diketahui atau dengan melakukan transformasi jika σ_i^2 tidak diketahui. Setelah diperoleh model regresi yang baru harus diperiksa kembali apakah masih terjadi heteroskedastisitas atau tidak.

5.2 Saran

(1) Jika pada suatu model regresi terjadi penyimpangan asumsi heteroskedastisitas, maka harus dilakukan tindakan perbaikan untuk menghilangkan heteroskedastisitas tersebut. Setelah dilakukan tindakan perbaikan harus dideteksi kembali apakah masih terjadi heteroskedastisitas atau tidak.

(2) Pada penelitian ini, pengujian yang digunakan untuk mendeteksi heteroskedastisitas adalah pengujian korelasi dengan rank Spearman, pada penelitian selanjutnya diharapkan menggunakan jenis pengujian yang berbeda.

DAFTAR PUSTAKA

Algifari. 2000. *Analisis Regresi Teori, Kasus dan Solusi*. Yogyakarta : BPFE.

Anton, H. 1992. *Aljabar Linear Elementer*. Terjemahan oleh Pantur Silaban. Jakarta : Erlangga.

Conover, W. J. 1971. *Practical Nonparametric Statistic*. Canada : John Wiley & Sons.

Gujarati, D. 1978. *Ekonometrika Dasar*. Terjemahan oleh Sumarno Zain. Jakarta : Erlangga.

Praptono. 1986. *Metode Statistika Nonparametrik*. Jakarta : Universitas Terbuka.

Sembiring, R. K. 1995. *Analisis Regresi Edisi Kedua*. Bandung : ITB.

Siegel, S. 1994. *Statistik Nonparametrik Untuk Ilmu-Ilmu Sosial*. Jakarta : PT. Gramedia Pustaka Utama.

Sudjana. 2001. *Metoda Statistika*. Bandung : Tarsito.

Sumodiningrat, G. 2002. *Ekonometrika Pengantar*. Yogyakarta : BPFE Yogyakarta.

Widarjono, A. 2007. *Ekonometrika : Teori dan Aplikasi Untuk Ekonomi dan Bisnis.* Yogyakarta : EKONESIA fakultas Ekonomi UI.

Widiastutik. 2005. *Uji Hipotesis Berdasarkan Rank Spearman dan Simulasinya dengan Program SPSS.* Semarang : UNNES.

Lampiran 1

Data Jumlah Tenaga Kerja dan Output Yang Dihasilkan Industri Isic 3 Digit
Tahun 1993

Klasifikasi industri besar dan sedang ISIC 3 digit	Jumlah tenaga kerja (orang)	Jumlah output yang dihasilkan (rupiah)
311 Makanan	369248	18740153851
312 Makanan	143493	3488229886
313 Minuman	21127	829310486
314 Pengolahan tembakau dan	184304	8215641238
bumbu rokok	580519	14182392561
321 Tekstil	350039	6300671710
322 Pakaian jadi, kecuali alas kaki	23296	510143485
323 Kulit dan barang dari kulit	230901	4415114453
324 Alas kaki	378093	11626808024
331 Kayu, bambu, rotan, rumput	123428	1420820377
332 Perabotan dan perlengkapan	74060	3663693185
rumah tangga	48757	1287837008
341 Kertas, barang dari kertas	60112	5637636226
342 Percetakan dan penerbitan	100194	5433798845
351 Bahan kimia industri	772	49824840
352 Kimia lain		
354 Barang-barang hasil kilang minyak	121505	3557350068
bumi dan batu bara	119574	2779130487
355 Karet dan barang-barang dari	38691	803480230
karet	20121	878305620
356 Barang plastik	42102	2292618803
361 Porselin		
362 Gelas dan barang-barang dari	29226	123058917
gelas	18222	419136783
363 Semen, kapur dan barang-barang	31465	6163470615
dari semen dan kapur	12047	1241100474
364 Pengolahan tanah liat	117748	3821375468
369 Barang galian lain bukan logam		
371 Logam dasar besi dan baja	36158	1522240396
372 Logam dasar bukan besi		
381 Barang dari logam kecuali mesin	107172	5689820639
dan peralatannya		
382 Mesin dan perlengkapannya	100226	6495988172
kecuali mesin listrik	6278	199302084
383 Mesin, peralatan dan		
perlengkapan listrik serta bahan	70500	1043723314

keperluan listrik 384 Alat angkutan 385 Peralatan profesional, ilmu pengetahuan, pengukuran dan pengatur 390 Pengolahan lainnya		

Sumber : BPS, Statistik Industri Besar dan Sedang tahun 1993

Lampiran 2

Hasil Output SPSS Data Jumlah Tenaga Kerja dan Output Yang Dihasilkan Industri Isic 3 Digit Tahun 1993

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	X[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: Y

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,839[a]	,704	,694	75436,360	1,309

a. Predictors: (Constant), X

b. Dependent Variable: Y

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3,8E+011	1	3,796E+011	66,700	,000[a]
	Residual	1,6E+011	28	5690644369		
	Total	5,4E+011	29			

a. Predictors: (Constant), X

b. Dependent Variable: Y

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	12258,693	18957,238		,647	,523		
	X	2,60E-005	,000	,839	8,167	,000	1,000	1,000

a. Dependent Variable: Y

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions	
				(Constant)	X
1	1	1,687	1,000	,16	,16
	2	,313	2,322	,84	,84

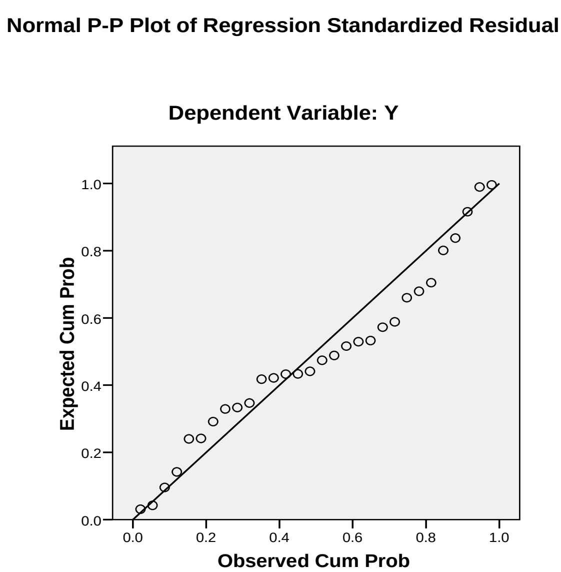
a. Dependent Variable: Y

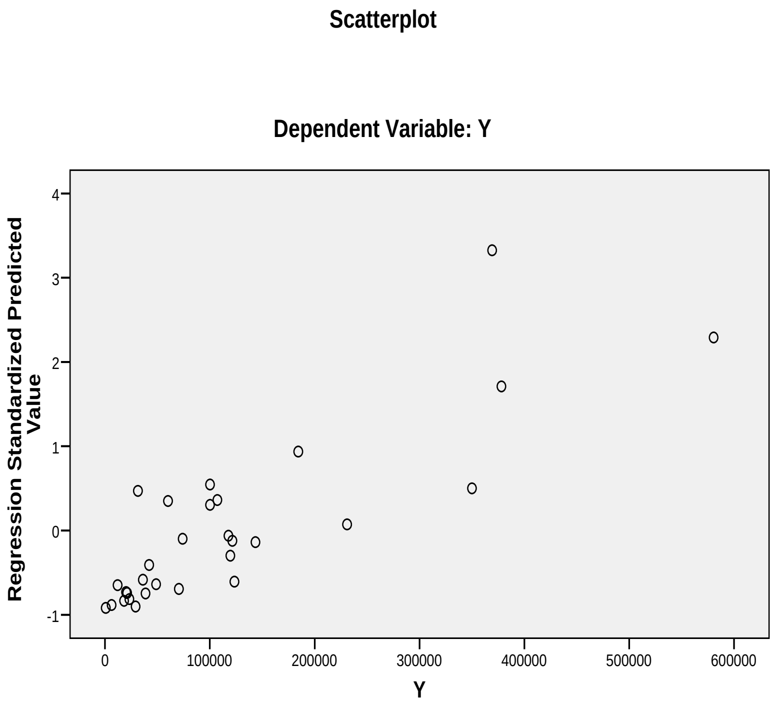
Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	13553,32	499194,59	118645,93	114404,940	30
Residual	-140943	199751,3	,000	74124,323	30
Std. Predicted Value	-,919	3,326	,000	1,000	30
Std. Residual	-1,868	2,648	,000	,983	30

a. Dependent Variable: Y

Charts


Normal P-P Plot of Regression Standardized Residual
Dependent Variable: Y
Expected Cum Prob
Observed Cum Prob
0.0
0.2
0.4
0.6
0.8
1.0



Lampiran 3

Tabel Perhitungan Koefisien Korelasi Rank Spearman

No	Y	X	e	$\lvert e \rvert$	rank $\lvert e \rvert$	rank X	d	d^2
1	369248	18740153851	-130254,6931	130254,6931	27	30	3	9
2	143493	3488229886	40540,32996	40540,32996	18	16	-2	4
3	21127	829310486	-12693,76564	12693,76564	7	7	0	0
4	184304	8215641238	-41561,36519	41561,36519	19	27	8	64
5	580519	14182392561	199518,1004	199518,1004	30	29	-1	1
6	350039	6300671710	173962,8425	173962,8425	29	25	-4	16
7	23296	510143485	-2226,42361	2226,42361	1	5	4	16
8	230901	4415114453	103849,3312	103849,3312	26	20	-6	36
9	378093	11626808024	63537,29838	63537,29838	22	28	6	36
10	123428	1420820377	74227,9772	74227,9772	23	12	-11	121
11	74060	3663693185	-33454,71581	33454,71581	16	18	2	4

12	48757	1287837008	3014,544792	3014,544792	2	11	9	81
13	60112	5637636226	-98725,23488	98725,23488	25	22	-3	9
14	100194	5433798845	-53343,46297	53343,46297	21	21	0	0
15	772	49824840	-12782,13884	12782,13884	8	1	-7	49
16	121505	3557350068	16755,20523	16755,20523	12	17	5	25
17	119574	2779130487	35057,91434	35057,91434	17	15	-2	4
18	38691	803480230	5541,82102	5541,82102	4	6	2	4
19	20121	878305620	-14973,63912	14973,63912	10	8	-2	4
20	42102	2292618803	-29764,78188	29764,78188	13	14	1	1
21	29226	123058917	13767,77516	13767,77516	9	2	-7	49
22	18222	419136783	-4934,249358	4934,249358	3	4	1	1
23	31465	6163470615	-141043,929	141043,929	28	24	-4	16
24	12047	1241100474	-32480,30532	32480,30532	15	10	-5	25
25	117748	3821375468	6133,544832	6133,544832	5	19	14	196
26	36158	1522240396	-15678,9433	15678,9433	11	13	2	4
27	107172	5689820639	-53022,02961	53022,02961	20	23	3	9
28	100226	6495988172	-80928,38547	80928,38547	24	26	2	4
29	6278	199302084	-11162,54718	11162,54718	6	3	-3	9
30	70500	1043723314	31104,50084	31104,50084	14	9	-5	25
Jumlah								822

Lampiran 4

Tabel Perhitungan Transformasi Variabel

No	Y	X	$\frac{Y}{X}$	$\frac{1}{X}$
1	369248	18740153851	1,97036E-05	5,33614E-11
2	143493	3488229886	4,11363E-05	2,86678E-10
3	21127	829310486	2,54754E-05	1,20582E-09
4	184304	8215641238	2,24333E-05	1,21719E-10
5	580519	14182392561	4,09324E-05	7,051E-11
6	350039	6300671710	5,55558E-05	1,58713E-10
7	23296	510143485	4,56656E-05	1,96023E-09
8	230901	4415114453	5,22979E-05	2,26495E-10
9	378093	11626808024	3,25191E-05	8,60081E-11
10	123428	1420820377	8,68709E-05	7,03819E-10
11	74060	3663693185	2,02146E-05	2,72949E-10

12	48757	1287837008	3,78596E-05	7,76496E-10
13	60112	5637636226	1,06626E-05	1,77379E-10
14	100194	5433798845	1,8439E-05	1,84033E-10
15	772	49824840	1,54943E-05	2,00703E-08
16	121505	3557350068	3,4156E-05	2,81108E-10
17	119574	2779130487	4,30257E-05	3,59825E-10
18	38691	803480230	4,81543E-05	1,24459E-09
19	20121	878305620	2,29089E-05	1,13856E-09
20	42102	2292618803	1,83642E-05	4,36182E-10
21	29226	123058917	0,000237496	8,12619E-09
22	18222	419136783	4,34751E-05	2,38586E-09
23	31465	6163470615	5,10508E-06	1,62246E-10
24	12047	1241100474	9,70671E-06	8,05737E-10
25	117748	3821375468	3,0813E-05	2,61686E-10
26	36158	1522240396	2,37531E-05	6,56926E-10
27	107172	5689820639	1,88357E-05	1,75752E-10
28	100226	6495988172	1,54289E-05	1,53941E-10
29	6278	199302084	3,14999E-05	5,01751E-09
30	70500	1043723314	6,75466E-05	9,58108E-10

Lampiran 5

Hasil Output Setelah Transformasi

Regression

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	X[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: Y

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,233[a]	,054	,021	**********	2,141

a. Predictors: (Constant), X

b. Dependent Variable: Y

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,000	1	,000	1,609	,215[a]
	Residual	,000	28	,000		
	Total	,000	29			

a. Predictors: (Constant), X

b. Dependent Variable: Y

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	7,68E-010	,000		,794	,434		
	X	2,17E-005	,000	,233	1,268	,215	1,000	1,000

a. Dependent Variable: Y

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions	
				(Constant)	X
1	1	1,692	1,000	,15	,15
	2	,308	2,343	,85	,85

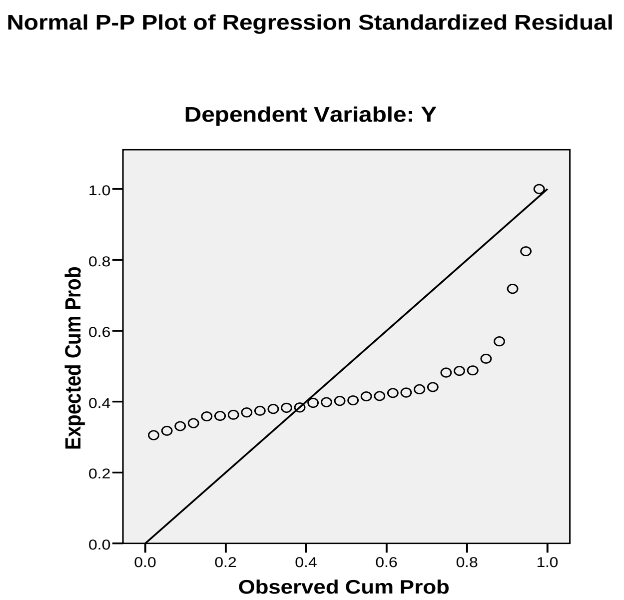
a. Dependent Variable: Y

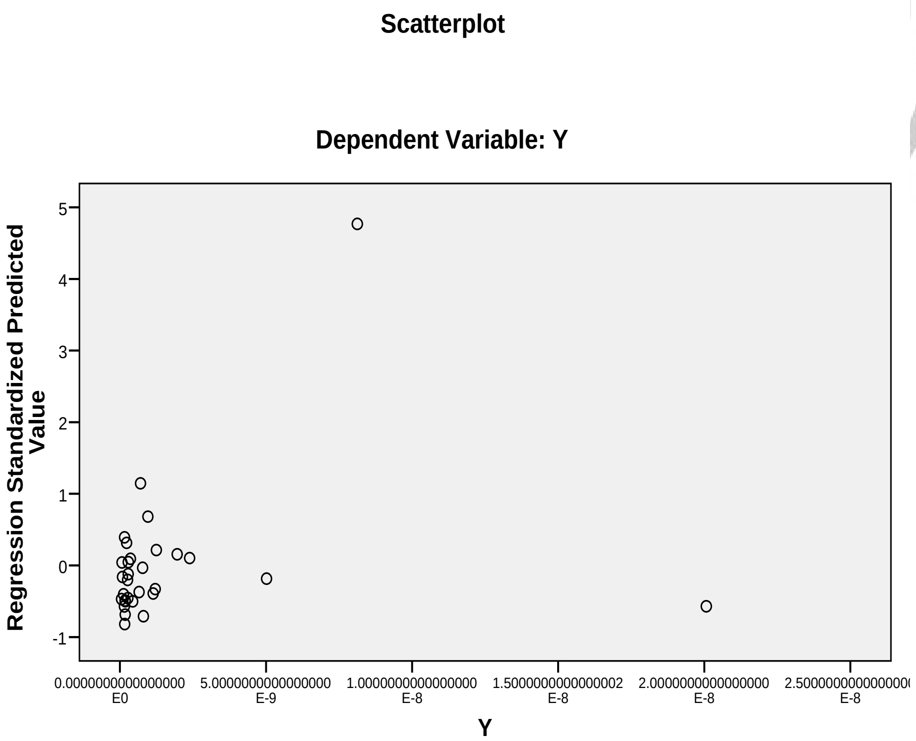
Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	********	********	********	***********	30
Residual	********	********	********	***********	30
Std. Predicted Value	-,819	4,768	,000	1,000	30
Std. Residual	-,509	4,958	,000	,983	30

a. Dependent Variable: Y

Charts


Normal P-P Plot of Regression Standardized Residual
Dependent Variable: Y
Expected Cum Prob
Observed Cum Prob
0.0
0.2
0.4
0.6
0.8
1.0


Scatterplot
Dependent Variable: Y
Regression Standardized Predicted Value
Y
-1
0
1
2
3
4
5
0.0000000000000000 E0
5.0000000000000000 E-9
1.0000000000000000 E-8
1.5000000000000002 E-8
2.0000000000000000 E-8
2.5000000000000000 E-8

Lampiran 6

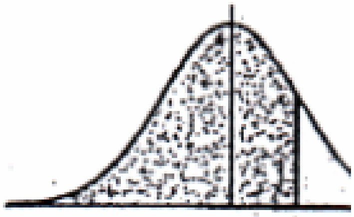
Tabel Perhitungan Koefisien Korelasi Rank Spearman Setelah Transformasi

No	Y	X	*e*	$\|e\|$	rank $\|e\|$	rank *X*	*d*	d^2
1	369248	18740153851	-37413,33857	37413,33857	15	30	15	225
2	143493	3488229886	67798,41147	67798,41147	19	16	-3	9
3	21127	829310486	3130,962454	3130,962454	29	7	-22	484
4	184304	8215641238	6024,585135	6024,585135	26	27	1	1
5	580519	14182392561	272761,0814	272761,0814	3	29	26	676
6	350039	6300671710	213314,4239	213314,4239	11	25	14	196
7	23296	510143485	12225,88638	12225,88638	4	5	1	1
8	230901	4415114453	135093,0164	135093,0164	20	20	0	0
9	378093	11626808024	125791,2659	125791,2659	22	28	6	36
10	123428	1420820377	92596,19782	92596,19782	7	12	5	25
11	74060	3663693185	-5442,142115	5442,142115	24	18	-6	36
12	48757	1287837008	20810,93693	20810,93693	27	11	-16	256
13	60112	5637636226	-62224,7061	62224,7061	14	22	8	64

14	100194	5433798845	-17719,43494	17719,43494	12	21	9	81
15	772	49824840	-309,199028	309,199028	18	1	-17	289
16	121505	3557350068	44310,50352	44310,50352	21	17	-4	16
17	119574	2779130487	59266,86843	59266,86843	25	15	-10	100
18	38691	803480230	21255,47901	21255,47901	1	6	5	25
19	20121	878305620	1061,768046	1061,768046	28	8	-20	400
20	42102	2292618803	-7647,828025	7647,828025	16	14	-2	4
21	29226	123058917	26555,6215	26555,6215	30	2	-28	784
22	18222	419136783	9126,731809	9126,731809	17	4	-13	169
23	31465	6163470615	-102282,3123	102282,3123	13	24	11	121
24	12047	1241100474	-14884,88029	14884,88029	2	10	8	64
25	117748	3821375468	34824,15234	34824,15234	10	19	9	81
26	36158	1522240396	3125,383407	3125,383407	23	13	-10	100
27	107172	5689820639	-16297,10787	16297,10787	9	23	14	196
28	100226	6495988172	-40736,94333	40736,94333	8	26	18	324
29	6278	199302084	1953,144777	1953,144777	6	3	-3	9
30	70500	1043723314	47851,20409	47851,20409	5	9	4	16
jumlah								4788

Tabel Distribusi t

Nilai Persentil
Untuk Distribusi t
V = dk
(Bilangan Dalam Badan Daftar
Menyatakan t_p)



t_p

V	$t_{0,995}$	$t_{0,99}$	$t_{0,975}$	$t_{0,95}$	$t_{0,90}$	$t_{0,80}$	$t_{0,75}$	$t_{0,70}$	$t_{0,60}$
1	63,66	31,82	12,71	6,31	3,08	1,376	1,000	0,727	0,325
2	9,92	6,96	4,30	2,92	1,89	1,061	0,816	0,617	0,289
3	5,84	4,54	3,18	2,35	1,64	0,978	0,765	0,584	0,277
4	4,60	3,75	2,78	2,13	1,53	0,941	0,741	0,569	0,271
5	4,03	3,36	2,57	2,02	1,48	0,920	0,727	0,559	0,267
6	3,71	3,14	2,45	1,94	1,44	0,906	0,718	0,553	0,265
7	3,50	3,00	2,36	1,90	1,42	0,896	0,711	0,549	0,263
8	3,36	2,90	2,31	1,86	1,40	0,889	0,706	0,546	0,262
9	3,25	2,82	2,26	1,83	1,38	0,883	0,703	0,543	0,261
10	3,17	2,76	2,23	1,81	1,37	0,879	0,700	0,542	0,260
11	3,11	2,72	2,20	1,80	1,36	0,876	0,697	0,540	0,260
12	3,06	2,68	2,18	1,78	1,36	0,873	0,695	0,539	0,259
13	3,01	2,65	2,16	1,77	1,35	0,870	0,694	0,538	0,259
14	2,98	2,62	2,14	1,76	1,34	0,868	0,692	0,537	0,258
15	2,95	2,60	2,13	1,75	1,34	0,866	0,691	0,536	0,258
16	2,92	2,58	2,12	1,75	1,34	0,865	0,690	0,535	0,258
17	2,90	2,57	2,11	1,74	1,33	0,863	0,689	0,534	0,257
18	2,88	2,55	2,10	1,73	1,33	0,862	0,688	0,534	0,257
19	2,86	2,54	2,09	1,73	1,33	0,861	0,688	0,533	0,257
20	2,84	2,53	2,09	1,72	1,32	0,860	0,687	0,533	0,257
21	2,83	2,52	2,08	1,72	1,32	0,859	0,686	0,532	0,257
22	2,82	2,51	2,07	1,72	1,32	0,858	0,686	0,532	0,256
23	2,81	2,50	2,07	1,71	1,32	0,858	0,685	0,532	0,256
24	2,80	2,49	2,06	1,71	1,32	0,857	0,685	0,531	0,256
25	2,79	2,48	2,06	1,71	1,32	0,856	0,684	0,531	0,256
26	2,78	2,48	2,06	1,71	1,32	0,856	0,684	0,531	0,256
27	2,77	2,47	2,05	1,70	1,31	0,855	0,684	0,531	0,256
28	2,76	2,47	2,05	1,70	1,31	0,855	0,683	0,530	0,256
29	2,76	2,46	2,04	1,70	1,31	0,854	0,683	0,530	0,256
30	2,75	2,46	2,04	1,70	1,31	0,854	0,683	0,530	0,256
40	2,70	2,42	2,02	1,68	1,30	0,851	0,681	0,529	0,255
60	2,66	2,39	2,00	1,67	1,30	0,848	0,679	0,527	0,254
120	2,62	2,36	1,98	1,66	1,29	0,845	0,677	0,526	0,254
∞	2,58	2,33	1,96	1,645	1,28	0,842	0,674	0,524	0,253

Sumber : Statistical Tables for Biological, Agricultural and Medical Research, Fisher, R.A. dan Yates, F. Table III, Oliver & Boyd Ltd, Edinburgh.