METODE ESTIMASI *GENERALIZED THREE STAGE RIDGE REGRESSION* UNTUK MENGATASI MULTIKOLINEARITAS PADA MODEL PERSAMAAN SIMULTAN

SKRIPSI

INA PUSPITASARI

H121 15 304

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN

MAKASSAR

NOVEMBER 2019

METODE ESTIMASI *GENERALIZED THREE STAGE RIDGE REGRESSION* UNTUK MENGATASI MULTIKOLINEARITAS PADA MODEL PERSAMAAN SIMULTAN

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuna Alam Universitas Hasanuddin Makassar

INA PUSPITASARI

H121 15 304

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN

MAKASSAR

NOVEMBER 2019

LEMBAR PERNYATAAN KEOTENTIKAN

Saya yang bertanda tangan di bawah ini menyatakan dengan sungguh-sungguh bahwa skripsi yang saya buat dengan judul:

Metode Estimasi *Generalized Three Stage Ridge Regression* Untuk Mengatasi Multikolinearitas Pada Model Persamaan Simultan

Adalah benar hasil karya saya sendiri dan bukan hasil plagiat dan belum pernah dipublikasikan dalam bentuk apapun.

Makassar, 27 Nopember 2019

METERAI TEMPEL

F6AJX109588615

Ina Puspitasari

NIM. H121 15 304

LEMBAR PERSETUJUAN PEMBIMBING

METODE ESTIMASI *GENERALIZED THREE STAGE RIDGE REGRESSION* UNTUK MENGATASI MULTIKOLINEARITAS PADA MODEL PERSAMAAN SIMULTAN

Disetujui oleh:

Pembimbing Utama	Pembimbing Pertama
Dr. La Podje Talangko, M.Si.	**Andi Kresna Jaya, S.Si., M.Si.**
NIP. 19551219 198701 1001	**NIP. 19731228 200003 1001**

Ketua Program Studi

Dr. Nurtiti Sunusi, S.Si., M.Si.

NIP. 19720117 199703 2002

Pada Tanggal: 27 Nopember 2019

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh:

Nama : Ina Puspitasari

NIM : H121 15 304

Program Studi : STATISTIKA

Jusul Skripsi : Metode Estimasi *Generalized Three Stage Ridge Regression* Untuk Mengatasi Multikolinearitass Pada Model Persamaan Simultan

Telah berhasil dipertahankan dihadapan dewan penguji dan diterima sebagai bagian dari persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin .

DEWAN PENGUJI

Tanda Tangan

1. Ketua : Dr. La Podje Talangko, M.Si. (..............................)
2. Sekretaris : Andi Kresna Jaya, S.Si., M.Si. (..............................)
3. Anggota : Drs. Raupong, M.Si. (..............................)
4. Anggota : Dr. Nurtiti Sunusi, S.Si., M.Si. (..............................)

Ditetapkan di : Makassar

KATA PENGANTAR

Segala puji dan syukur dipanjatkan kepada Allah Subhanahu Wa Ta'ala atas rahmat dan kuasaNya, yang telah menyertai penulis selama proses penyelesaian skripsi dengan judul "**Metode Estimasi *Generalized Three Stage Ridge Regression* Untuk Mengatasi Multikolinearitas Pada Model Persamaan Simultan"** sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin Makassar.

Selesainya skripsi ini juga tidak terlepas dari dukungan banyak orang. Izinkan penulis mengucapkan terima kasih serta penghargaan yang setinggi-tingginya kepada orang tua penulis **Settari** dan **Yasse** serta kakak-kakak penulis **Rina, Diana, Enni** dan **Lasu** atas doa dan motivasi serta kasih sayang yang tidak mungkin diperoleh dari orang lain.

Tidak lupa penulis juga mengucapkan terima kasih kepada seluruh pihak yang senantiasa membantu penulis dalam berbagai hal baik berupa materi, tenaga dan dukungan moral selama proses penyelesaian tulisan ini. Penghargaan yang tulus dan terima kasih juga penulis ucapkan kepada:

1. **Ibu Prof. Dwia Aries Tina Palubuhu, MA.**, selaku Rektor Universitas Hasanuddin beserta seluruh jajarannya.
2. **Bapak Dr. Eng. Amiruddin**, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya.
3. **Ibu Dr. Nurititi Sunusi, S.Si., M.Si.**, selaku ketua Departemen Statistika sekaligus Penasehat Akademik dan salah satu Anggota Tim Penguji, segenap Dosen Pengajar dan Staf Departemen Statistika yang telah membekali ilmu dan kemudahan-kemudahan kepada penulis dalam berbagai hal selama menjadi mahasiswa di Departemen Statistika.
4. **Bapak Dr. La Podje Talangko, M.Si.**, selaku Pembimbing Utama sekaligus Ketua Penguji dan **Bapak Andi Kresna Jaya, S.Si.**, selaku Pembimbing Pertama sekaligus Sekretaris Penguji yang telah meluangkan waktu dan pemikirannya untuk memberikan arahan, dorongan dan motivasi kepada penulis mulai dari awal hingga selesainya skripsi ini.

5. **Bapak Drs. Raupong, M.Si.**, selaku salah satu Anggota Tim Penguji yang telah memberikan saran dan kritikan yang membangun dalam menyempurnakan tugas akhir ini.

Ucapan terima kasih juga penulis sampaikan kepada orang-orang yang telah berperan besar serta istimewa yaitu:

1. Sahabat seperjuangan di Statistika terkhusus **Fitri, Noetz, Hesti, Dian, Irma, Mega, Ika** dan seluruh **Statistika 2015** atas kebersaman dan berbagai kenanganya.
2. Keluarga besar **Himatika Fmipa Unhas** yang telah memberikan ilmu yang mungkin tidak bias didapatkan diproses perkuliahan serta berbagai. Penulis bersyukur bias menjadi salah satu bagian dari himpunan ini. **BRAVO Himatika**.
3. **Teman-teman Departemen Matematika dan Statistika** terkhusus keluarga **SIMETRIS 2015** yang memberikan berbagai kenangan dengan rasanya masing-masing selama berproses.
4. **Teman-teman KKN Gel. 99 Kelurahan Bulu Pabbulu Zian, Asri, tami, Virda, Rika** dan lainnya telah menjadi teman sekaligus keluarga selama sebulan lebih. Semoga silaturahmi kita bias tetap berlanjut ke waktu mendatang.

Serta kepada seluruh pihak yang tidak dapat penulis sebutkan satu persatu, terima untuk semuanya. Semoga apa yang tertulis dalam skripsi ini dapat bermanfaat bagi siapa pun. Penulis memohon maaf apabila ada sedikit atau banyak kekurangan dalam tugas akhir ini, karena kesempurnaan hanya milik Allah Subhanahu Wa Ta’ala.

Makassar, 27 November 2019

Ina Puspitasari

PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIK

Sebagai civitas akdemik Universitas Hasanuddin, saya yang bertanda tangan di bawah ini:

Nama : Ina Puspitasari

NIM : H121 15 304

Program Studi : Statistika

Departemen : Statistika

Fakultas : Matematika dan Ilmu Pengetahuan Alam

Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Hasanuddin **Hak Prediktor Royalti Non-Eksklusif (*Non-exclusive Royalty- Free Right*)** atas tugas akhir saya yang berjudul:

"Metode Estimasi *Generalized Three Stage Ridge Regression* Untuk Mengatasi Multikolinearitas Pada Model Persamaan Simultan"

beserta perangkat yang ada (jika diperlukan). Terkait dengan hal di atas, maka pihak universitas berhak menyimpan, mengalih-media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di Makassar pada tanggal 27 November 2019

Yang menyatakan

(Ina Puspitasari)

ABSTRAK

Model persamaan simultan pada penelitian ini merupakan model yang terdiri dari dua persamaan dengan variabel respon satu persamaan menjadi variabel prediktor pada persamaan lain yang menyebabkan adanya korelasi galat antar variabel, Estimasi model persamaan simultan dengan metode kuadrat terkecil (MKT) akan menghasilkan estimasi parameter yang bias dan tidak konsisten, sehingga harus menggunakan metode regresi lain. Salah satu asumsi yang harus terpenuhi dalam analisis regresi yaitu multikolinearitas. Data yang mengandung masalah multikolinearitas dapat diatasi dengan metode regresi ridge. Pada penelitian ini, digunakan data sekunder yaitu jumlah uang beredar, kurs rupiah dalam dollar, indeks harga konsumen, BI *rate*, inflasi dan sertifikat bank Indonesia pada rentang tahun 2013 sampai 2014. Data tersebut memiliki masalah autokorelasi dan multikolinearitas. Metode yang digunakan dalam penelitian ini yaitu gabungan antara *three stage least square* (3SLS) dan *generalized ridge regression* (GRR) yang dinamakan *generalized three stage ridge regression* (G3SRR) untuk mengatasi multikolinearitas pada model persamaan simultan. Dari hasil penerapan metode G3SRR, diperoleh bahwa kurs memiliki pengaruh positif terhadap jumlah uang beredar. Sedangkan jumlah uang beredar memiliki pengaruh negatif terhadap kurs.

Kata kunci : Persamaan simultan, multikolinearitas, *three stage least square*, *generalized ridge regression* dan *generalized three stage ridge regression*.

ABSTRACT

The simultaneous equation model in this study is a model that consists of two equations with the response variable of one equation being a predictor variable in another equation that causes an error correlation between variables. Simultaneous equations model estimation by the method of least squares (OLS) will produce biased parameter estimates and are inconsistent, so it must use another regression method. One assumption that must be fulfilled in a regression analysis is multicollinearity. Data containing multicollinearity problems can be overcome by the ridge regression method. In this study, secondary data are used, namely the money supply, the rupiah exchange rate in dollars, the consumer price index, the BI rate, inflation and Bank Indonesia certificates in the span of 2013 to 2014. These data have autocorrelation and multicollinearity problems. The method used in this study is a combination of three stage least square (3SLS) and generalized ridge regression (GRR) called generalized three stage ridge regression (G3SRR) to overcome multicollinearity in the simultaneous equation model. From the results of applying the G3SRR method, it was found that the exchange rate has a positive influence on the money supply. Whereas the money supply has a negative influence on the exchange rate.

Kata kunci : simultaneous equation, multicollinearity, three stage least square, generalized ridge regression, generalized three stage ridge regression.

DAFTAR ISI

DAFTAR TABEL

DAFTAR LAMPIRAN

BAB I
PENDAHULUAN

1.1 Latar Belakang

Model persamaan simultan adalah suatu model yang mengandung lebih dari satu persamaan yang saling mempengaruhi. Ciri khas dari model persamaan simultan adalah variabel tak bebas dalam suatu persamaan muncul sebagai variabel yang menjelaskan dalam persamaan lain dalam model. Dalam model ini, sejumlah persamaan membentuk suatu sistem persamaan yang menggambarkan ketergantungan di antara berbagai variabel dalam persamaan-persamaan tersebut.

Model suatu persamaan dapat diperoleh dengan melakukan estimasi terhadap parameternya menggunakan analisis regresi. Dalam analisis regresi terdapat beberapa metode yang dapat digunakan diantaranya o*rdinary least square* (OLS) atau biasa juga dikenal sebagai metode kuadrat terkecil sederhana. Namun dalam persamaan simultan, metode OLS menghasilkan estimasi parameter yang bias dan tidak konsisten. Salah satu metode alternatif yang dapat digunakan untuk mengestimasi model persamaan simultan yaitu *three stage least square* (3SLS). Metode ini merupakan pengembangan dari *two stage least square* (2SLS), dimana tahap pertama dan kedua sama dengan 2SLS, selanjutnya tahap ketiga menggunakan *general least square* (GLS). Kelebihan dari metode 3SLS ini yaitu saat menduga parameter pada persamaan tertentu, ia juga mempertimbangkan parameter pada persamaan lainnya.

Untuk menerapkan metode 3SLS, salah satu asumsi yang harus dipenuhi yaitu tidak terjadi multikolinearitas. Multikolinearitas adalah suatu kondisi buruk (*ill condition*) yaitu terjadi korelasi yang kuat di antara variabel-variabel prediktor (X) yang diikutsertakan dalam pembentukan model regresi linear sehingga $\boldsymbol{X'X}$ tidak memenuhi asumsi klasik (Montgomery & Peck, 1992).

Terdapat beberapa metode yang dapat mengatasi masalah multikolinearitas, diantaranya yaitu regresi ridge. Regresi ridge merupakan metode estimasi koefisien regresi yang diperoleh melalui penambahan konstanta bias k pada diagonal $\boldsymbol{X'X}$. Salah satu metode pengembangan regresi ridge yaitu *generalized ridge regression*

(GRR). Nilai k pada regresi ridge sama untuk setiap peubah bebas, sedangkan pada GRR memungkinkan terdapat parameter bias k berbeda untuk setiap peubah bebas.

Retty Catherina Oktaviani (2018) membahas tentang permasalahan persamaan simultan yang mengandung multikolinearitas yang diatasi dengan penggunaan metode gabungan antara *two stage least square* (2SLS) dan regresi ridge diperumum yang disebut *generalized two stage ridge regression* (GTSRR). Titik Purwanti juga pada tahun 2012 membahas tentang permasalahan persamaan simultan yang diatasi dengan penggunaan metode *three stage least square* (3SLS). Berdasarkan pada kedua penelitian tersebut, peneliti akan menjelaskan dan membahas tentang metode estimasi *generalized three stage ridge regression* dalam mengatasi permasalahan data yang mengandung multikolinearitas pada persamaan simultan. Adapun judul tulisan ini yaitu " **Metode Estimasi *Generalized Three Stage Ridge Regression* Untuk Mengatasi Multikolinearitas Pada Model Persamaan Simultan**".

1.2 Rumusan Masalah

Berdasarkan permasalahan yang tertuang dalam latar belakang, terdapat beberapa rumusan masalah yaitu:

1. Bagaimana mengestimasi parameter menggunakan metode *generalized three stage ridge regression*?
2. Bagaimana menerapkan metode *generalized three stage ridge regression* pada data simultan yang mengandung multikolinearitas?

1.3 Batasan Masalah

Dalam suatu penelitian pada umumnya, distribusi data harus memenuhi asumsi-asumsi regresi klasik yaitu memenuhi normalitas, tidak terjadi autokorelasi, tidak terjadi mutikolinearitas, dan tidak terjadi heterokedastisitas. Namun pada penelitian ini, akan dibahas masalah multikolinearitas yang terjadi pada persamaan simultan dan cara penanganan masalah tersebut menggunakan metode estimasi *generalized three stage ridge regression.*

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah ditetapkan, maka dapat pula ditetapkan tujuan dari penelitian ini yaitu:

1. Untuk mendapatkan estimator *generalized three stage ridge regression*.
2. Untuk mendapatkan estimasi model simultan menggunakan metode *generalized three stage ridge regression*.

1.5 Manfaat Penelitian

Penelitian ini bermanfaat dalam pengembangan pengetahuan khususnya dalam mengatasi masalah multikoliniearitas dalam persamaan simultan dengan menggunakan metode *generalized three stage ridge regression.*

BAB II

TINJAUAN PUSTAKA

2.1 Regresi Linear

Menurut Deny Kurniawan (2008), regresi linear adalah metode statistika yang digunakan untuk membentuk model hubungan antara variabel respon (terikat; dependen; Y) dengan satu atau lebih variabel prediktor (bebas; independen; X). Apabila terdapat hubungan linear antara variabel respon (Y) dengan satu variabel prediktor (X) disebut regresi linear sederhana. Sedangkan hubungan linear antara dua atau lebih variabel prediktor ($X_1, X_2, \dots, X_n$) dengan variabel respon (Y) disebut sebagai regresi linear berganda.

Analisis regresi linear berganda merupakan analisis untuk mendapatkan hubungan secara linear antara dua atau lebih variabel prediktor dengan variabel respon. Tujuannya adalah menduga besarnya koefisien regresi yang akan menunjukkan besarnya pengaruh masing-masing variabel prediktor terhadap variabel respon. Model linearnya dapat dinyatakan sebagai (Myers, 1990):

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip} + \varepsilon_i \tag{2.1}$$

dimana:

Y_i : variabel respon ($i = 1,2,3, \dots, n$)

β_j : koefisien atau parameter regresi ($j = 0,1,2,3, \dots, p$)

X_{ij} : variabel prediktor ke-i pada item ke-j

ε_i : nilai galat regresi yang saling bebas dan berdistribusi normal $\varepsilon_i \sim N(0, \sigma^2)$

Dalam analisis regresi berganda, asumsi yang harus dipenuhi yaitu:

1. Nilai ekspektasi dari vektor residualnya adalah 0

   $E(\varepsilon_i) = 0, i = 1,2,3, \dots, n$

   $$E\begin{bmatrix}\varepsilon_1\\ \varepsilon_2\\ \vdots\\ \varepsilon_n\end{bmatrix} = \begin{bmatrix}E(\varepsilon_1)\\ E(\varepsilon_2)\\ \vdots\\ E(\varepsilon_n)\end{bmatrix} = \begin{bmatrix}0\\ 0\\ \vdots\\ 0\end{bmatrix}$$

2. Variansinya konstan untuk semua residual

   $Var(\varepsilon_i) = \sigma^2, i = 1,2,3, \dots, n$
3. Tidak ada autokorelasi antar residual

   $Cov(\varepsilon_i, \varepsilon_j) = 0, i \neq j$
4. Tidak terdapat hubungan linear antara variabel independen satu dengan yang lainnya atau tidak terjadi multikolinearitas.

2.2 Model Persamaan Simultan

Model persamaan simultan adalah model yang variabel respon pada satu persamaan dapat berperan sebagai variabel prediktor pada persamaan lain. Oleh karena itu, penyebutan variabel prediktor dan variabel respon tidak tepat untuk digunakan pada model persamaan simultan. Sehingga di dalam model persamaan simultan digunakan 2 jenis variabel yaitu variabel endogen dan eksogen. Variabel endogen adalah variabel yang nilainya ditentukan dalam model, karena nilai-nilai ini diperoleh dengan memasukkan nilai variabel lain sebagai akibat adanya hubungan antar variabel. Jumlah variabel endogen sama banyaknya dengan persamaan dalam model. Sedangkan variabel eksogen adalah variabel yang nilainya ditentukan di luar model atau merupakan variabel prediktor asli (Gujarati, 1978).

Ada dua persamaan dalam model simultan, yaitu persamaan struktural dan *reduced form*. Persamaan struktural yaitu persamaan asli yang menggambarkan perilaku hubungan antar variabel dalam persamaan. Sedangkan persamaan *reduced form* merupakan suatu persamaan yang diperoleh dari persamaan-persamaan struktural yang telah dikaitkan. Dengan menyelesaikan persamaan *reduced form*, kita dapat menghitung koefisien-koefisien dalam persamaan struktural. Oleh karena itu, penaksiran terhadap persamaan struktural akan tergantung dari hasil penaksiran pada persamaan *reduced form*. Sebelum menerapkan suatu metode untuk menduga parameter pada setiap persamaan dalam model tersebut, perlu dilakukan identifikasi model terlebih dahulu. Gujarati (2003) mengemukakan bahwa untuk dapat diduga parameternya, suatu model persamaan simultan harus teridentifikasi.

Indentifikasi model ditentukan atas dasar "*order condition*" dan "*rank condition*". Untuk keperluan pengujian identifikasi persamaan perlu dibuat notasi berikut:

m = jumlah variabel endogen dalam model/sistem

m^* = jumlah variabel endogen dalam suatu persamaan

p = jumlah variabel eksogen dalam model/sistem

p^* = jumlah variabel eksogen dalam suatu persamaan

Identifikasi persamaan simultan ada tiga bentuk yaitu:

a. Identifikasi Kurang / tidak teridentifikasi (*under identified)*

Keadaan ini terjadi jika banyaknya variabel eksogen yang tidak tercakup dalam suatu persamaan lebih kecil dari pada banyaknya variabel endogen dalam persamaan tersebut dikurangi satu. Dalam model seperti ini, parameter struktural tidak dapat diestimasi. Notasinya yaitu: $(p - p^*) < (m^* - 1)$. Jika persamaan dalam kondisi tidak teridentifikasi, maka tidak ada teknik yang dapat digunakan untuk menaksir parameternya (Koutsoyiannis, A., 1973).

b. Identifikasi Tepat (*just or exact identified*)

Keadaan ini terjadi jika banyaknya variabel eksogen yang tidak tercakup dalam suatu persamaan sama dengan banyaknya variabel endogen dalam persamaan tersebut dikurangi satu. Dalam model seperti ini, parameter struktural dapat diestimasi. Notasinya yaitu: $(p - p^{*)} = (m^* - 1)$ dengan rank $A = m - 1$. Dimana A adalah matriks koefisien dari variabel yang tidak dimasukkan dalam persamaan tertentu, tetapi terkandung dalam persamaan lain dalam model. Persamaan dalam kondisi teridentifikasi tepat, dapat ditaksir parameternya menggunakan *indirect least square* (ILS) dan *two stage least square* (2SLS) (Ananta, 1987 ; Gurajati, ekonometrika dasar, 1978).

c. Identifikasi Berlebihan (*over identified*)

Keadaan ini terjadi jika banyaknya variabel eksogen yang tidak tercakup dalam suatu persamaan melebihi banyaknya variabel endogen dalam persamaan tersebut dikurangi satu. Dalam model seperti ini, parameter struktural dapat diestimasi. Notasinya yaitu: $(p - p^*) > (m^* - 1)$ dengan rank $A = m - 1$. Pada persamaan dengan kondisi teridentifikasi berlebihan, metode ILS tidak dapat digunakan untuk mengestimasi persamaan tersebut, biasanya ditaksir menggunakan 2SLS atau 3SLS (Koutsoyiannis, A., 1973).

Langkah-langkah identifikasi persamaan, yaitu:

1. Syarat order

Syarat order merupakan syarat yang diperlukan tetapi belum cukup untuk memastikan kondisi identifikasi. Syarat order menyatakan bahwa syarat identifikasi dari suatu persamaan struktural adalah jumlah variabel eksogen yang tidak dimasukkan dalam suatu persamaan, sekurang-kurangnya harus sama dengan jumlah variabel endogen yang terdapat dalam suatu persamaan dikurangi satu. Dalam bentuk notasi yaitu : $(p - p^*) \geq (m^* - 1)$.

Apabila suatu persamaan dalam model tidak memenuhi syarat order, maka tidak diperlukan menyelidiki syarat rank, sebaliknya bila syarat order dipenuhi, maka dilanjutkan menyelidiki syarat rank.

2. Syarat rank

Syarat rank menyatakan bahwa dalam suatu sistem terdiri dari m persamaan. Suatu persamaan dikatakan teridentifikasi jika sekurang-kurangnya memiliki satu determinan yang tidak nol berdimensi $(m - 1)$, dari suatu matriks koefisien-koefisien variabel yang tidak terdapat dalam suatu persamaan tertentu, tetapi muncul dalam persamaan lain.

Model umum persamaan simultan adalah sebagai berikut:

$$Y_{1t} = \beta_{12}Y_{2t} + \beta_{13}Y_{3t} + \cdots + \beta_{1m}Y_{mt} + \gamma_{11}X_{1t} + \cdots + \gamma_{1p}X_{pt} + \mu_{1t}$$

$$Y_{2t} = \beta_{21}Y_{1t} + \beta_{23}Y_{3t} + \cdots + \beta_{2m}Y_{mt} + \gamma_{21}X_{1t} + \cdots + \gamma_{2p}X_{pt} + \mu_{2t}$$

...

$$Y_{mt} = \beta_{m1}Y_{1t} + \beta_{m2}Y_{2t} + \cdots + \beta_{m(m-1)}Y_{(m-1)t} + \gamma_{m1}X_{1t} + \cdots + \gamma_{mp}X_{pt} + \mu_{mt} \quad (2.2)$$

dengan:

$Y_{1t}, Y_{2t}, \dots, Y_{mt}$	: variabel endogen
$X_{1t}, X_{2t}, \dots X_{pt}$	: variabel eksogen
$\mu_{1t}, \mu_{2t}, \dots, \mu_{mt}$	: nilai error persamaan simultan
$t = 1,2, \dots n$	: banyak observasi
β	: koefisien variabel endogen
γ	: koefisien variabel eksogen

Model pada persamaan (2.2) merupakan model lengkap yang secara umum dapat diselesaikan dengan model *reduced form.* Model *reduced form* ditulis sebagai berikut:

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \\ \vdots \\ Y_{mt} \end{bmatrix} = \begin{bmatrix} \pi_{11} & \pi_{12} & \cdots & \pi_{1p} \\ \pi_{21} & \pi_{22} & \cdots & \pi_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{m1} & \pi_{m2} & \cdots & \pi_{mp} \end{bmatrix} \begin{bmatrix} X_{1t} \\ X_{2t} \\ \vdots \\ X_{pt} \end{bmatrix} + \begin{bmatrix} \nu_{1t} \\ \nu_{2t} \\ \vdots \\ \nu_{mt} \end{bmatrix}$$

dengan :

$Y_{1t}, Y_{2t}, \dots, Y_{mt}$: variabel endogen

$\pi_{11}, \pi_{12}, \dots, \pi_{mp}$: koefisien regresi *reduced form*

$X_{1t}, X_{2t}, \dots X_{pt}$: variabel eksogen

$\nu_{1t,}\nu_{2t}, \dots, \nu_{mt}$: residual *reduced form* pada waktu t

Untuk estimasi parameter persamaan simultan, dapat dipakai beberapa metode, diantaranya metode OLS (*ordinary least square*), ILS (*indirect least square*), 2SLS (*two stage least square*), dan 3SLS (*three stage least square*). OLS yaitu metode yang meminimumkan jumlah kuadrat error, sehingga diperoleh estimator dengan variansi yang terkecil dan dapat digunakan untuk persamaan yang memenuhi asumsi diterapkannya OLS. ILS yaitu metode kuadrat terkecil tak langsung yang parameter persamaan strukturalnya ditaksir secara tak langsung dari persamaan *reduced form* dengan taksirannya yang bias tapi tetap konsisten. 2SLS yaitu metode kuadrat terkecil dua tahap dan digunakan untuk persamaan yang mengandung korelasi di antara variabel gangguan dengan variabel tak bebas (endogen). 3SLS yaitu metode metode kuadrat terkecil tiga tahap. Untuk tahap satu dan dua, sama dengan 2SLS dan tahap ketiga menggunakan metode GLS *modified* yaitu *seemingly unrelated* regressions (SUR).

2.3 Metode Kuadrat Terkecil

Metode kuadrat terkecil atau *ordinary least square* (OLS) merupakan salah satu metode untuk menghitung taksiran koefisien regresi. Metode OLS pada prinsipnya bertujuan untuk meminimumkan jumlah kuadrat galat (Montgomery & Peck, 1992). Misalkan model yang akan ditaksir adalah parameter dari persamaan

(2.1), jika dimodelkan dalam bentuk matriks, dapat ditulis dengan menggunakan persamaan sebagai berikut (Sembiring, 1995):

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{2.3}$$

$$\boldsymbol{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}; \ \boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}; \ \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}; \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

dengan :

$\boldsymbol{Y}$: vektor matriks respon berukuran $n \times 1$

$\boldsymbol{X}$: matriks variabel prediktor berukuran $n \times (p+1)$

$\boldsymbol{\beta}$: vektor parameter regresi yang akan diduga, berukuran $(p+1) \times 1$

$\boldsymbol{\varepsilon}$: vektor galat yang berukuran $n \times 1$

n : jumlah observasi

p : jumlah variabel

Berdasarkan persamaan (2.3) diperoleh

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

Jumlah kuadrat galat dapat dituliskan sebagai berikut:

$$\boldsymbol{G} = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}$$

$$\boldsymbol{G} = (\boldsymbol{Y} - \boldsymbol{X\beta})'(\boldsymbol{Y} - \boldsymbol{X\beta})$$

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

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

Karena $\boldsymbol{Y}'\boldsymbol{X\beta}$ merupakan skalar, maka $\boldsymbol{Y}'\boldsymbol{X\beta} = (\boldsymbol{Y}'\boldsymbol{X\beta})' = \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{Y}$ sehingga diperoleh

$$\boldsymbol{G} = \boldsymbol{Y}'\boldsymbol{Y} - 2\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{Y} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X\beta}$$

Kemudian $\boldsymbol{G}$ diturunkan secara parsial terhadap $\boldsymbol{\beta}$ dan disamakan dengan nol,

$$\left.\frac{\partial G}{\partial \boldsymbol{\beta}}\right|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}} = \mathbf{0}$$

$$-2\boldsymbol{X}'\boldsymbol{Y} + 2\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \mathbf{0}$$

$$\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{X}'\boldsymbol{Y}$$

sehingga diperoleh rumus penaksir parameter kuadrat terkecil, yaitu:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{OLS}} = (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{Y} \tag{2.4}$$

2.4 Metode Kuadrat Terkecil Diperumum

Metode kuadrat terkecil diperumum atau g*eneralized least square* (GLS) merupakan varian lain dari metode *least square*. Metode ini digunakan ketika asumsi bebas autokorelasi dan atau heterokedastisitas yang diisyaratkan oleh metode OLS tidak terpenuhi. Metode GLS dapat digunakan untuk mengestimasi parameter ketika model regresi linear berbentuk:

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

dengan $\boldsymbol{\varepsilon} \sim N\ (0, \boldsymbol{\sigma}^{\mathbf{2}}\mathbf{V})$, dimana $\boldsymbol{V}$ merupakan matriks berukuran $n \times n$. Matriks kovariansi galat berbentuk $\boldsymbol{\sigma}^{\mathbf{2}}\mathbf{V}$ dengan $\boldsymbol{V}$ merupakan matriks nonsingular dan definit positif sehingga terdapat matriks $\boldsymbol{K}$ simetris nonsingular berukuran $n \times n$ dengan $\boldsymbol{K}'\boldsymbol{K} = \boldsymbol{KK} = \boldsymbol{V}$ dan $\boldsymbol{K}$ disebut akar kuadrat dari $\boldsymbol{V}$.

Didefinisikan variabel-variabel baru sebagai berikut:

$$\boldsymbol{Y}^* = \boldsymbol{K}^{-1}\boldsymbol{Y}\,, \qquad \boldsymbol{X}^* = \boldsymbol{K}^{-1}\boldsymbol{X}\,, \qquad \boldsymbol{\varepsilon}^* = \boldsymbol{K}^{-1}\boldsymbol{\varepsilon}$$

Sehingga model regresi $\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$ menjadi $\boldsymbol{K}^{-1}\boldsymbol{Y} = \boldsymbol{K}^{-1}\boldsymbol{X\beta} + \boldsymbol{K}^{-1}\boldsymbol{\varepsilon}$ atau

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

Dengan demikian dapat dilakukan langkah-langkah pada metode OLS untuk mencari parameter model regresi metode GLS. Akan dicari estimator dari parameter yang meminimumkan bentuk kuadrat galat berikut:

$$\boldsymbol{\varepsilon}^{*\prime}\boldsymbol{\varepsilon}^* = (\boldsymbol{Y}^* - \boldsymbol{X}^*\boldsymbol{\beta})'(\boldsymbol{Y}^* - \boldsymbol{X}^*\boldsymbol{\beta})$$

Analogi dengan OLS, estimasi parameter-parameter pada $\boldsymbol{\beta}$ untuk model transformasi statistik linear umum yang disebut sebagai *generalized least square estimator* (GLSE), yaitu sebagai berikut:

$$\widehat{\boldsymbol{\beta}}_{gls} = (\boldsymbol{X}^{*'}\boldsymbol{X}^{*})^{-1}\boldsymbol{X}^{*}\boldsymbol{Y}^{*} = (\boldsymbol{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\boldsymbol{X}'\mathbf{V}^{-1}\boldsymbol{Y} \tag{2.5}$$

merupakan estimator yang *best linier unbiased estimator* (BLUE) (Aziz, 2010).

2.5 Korelasi Kesebayaan

Korelasi kesebayaan (*contemporaneous correlation*) merupakan ukuran hubungan antara galat dari m persamaan yang berbeda pada waktu yang sama (Dofour, 2003). Korelasi ini dapat diuji menggunakan statistik uji *lagrange multiplier* dengan

$H_0: s_{ij} = 0$ (semua kovarian bernilai nol/tidak terdapat korelasi kesebayaan)

H_1: minimal ada salah satu $s_{ij} \neq 0$ (terdapat korelasi kesebayaan)

Statistik uji yang digunakan yaitu sebagai berikut:

$$\lambda = n\sum_{i=2}^{M}\sum_{j=1}^{i-2} r_{ij}^2 \tag{2.6}$$

dengan $r_{ij}^2 = \frac{\sigma_{ij}^2}{\sigma_{ii}\sigma_{jj}}$ merupakan korelasi galat antara persamaan ke-i dan ke-j, σ_{ij}^2 merupakan kovariansi persamaan ke-i dengan persamaan ke-j, σ_{ii} merupakan variansi persamaan ke-i dan σ_{jj} merupakan variansi persamaan ke-j. λ mempunyai sebaran χ^2 dengan derajat bebas $\frac{m(m-1)}{2}$ dan taraf signifikan α. Keputusan H_0 akan ditolak jika λ lebih besar dari nilai χ^2.

2.6 *The Seemingly Unrelated Regressions*

Seemingly unrelated regressions (SUR) merupakan metode yang dikemukakan pertama kali oleh Arnold Zellner pada tahun 1962 yang akan menghasilkan model SUR. Model ini menunjukkan adanya galat pada persamaan struktural, dimana galat tersebut mengalami autokorelasi atau saling berkorelasi antara satu galat dengan galat lainnya.

Model SUR yang terdiri dari m persamaan dapat dituliskan dalam bentuk:

$$Y_{1t} = \beta_{10} + \beta_{11}X_{11,t} + \cdots + \beta_{1p_1}X_{1p_1,t} + \varepsilon_{1t}$$

$$Y_{2t} = \beta_{20} + \beta_{21}X_{21,t} + \cdots + \beta_{2p_2}X_{2p_2,t} + \varepsilon_{2t}$$

$$\cdots$$

$$Y_{mt} = \beta_{m0} + \beta_{m1}X_{m1,t} + \cdots + \beta_{mp_1}X_{mp_m,t} + \varepsilon_{mt} \tag{2.7}$$

dengan $p_i(i = 1,2,3, \ldots . m)$ merupakan jumlah variabel prediktor persamaan ke-i. Persamaan (2.7) dapat juga ditulis dalam bentuk berikut:

$$\boldsymbol{Y_i} = \boldsymbol{X_i\beta_i} + \boldsymbol{\varepsilon_i}\ , i = 1,2, \ldots , m$$

dimana $\boldsymbol{Y_i}(m \times 1), \boldsymbol{X_i}\left(m \times (m + \sum_{i=1}^{m} p_i)\right), \boldsymbol{\beta_i}\left((m + \sum_{i=1}^{m} p_i) \times 1\right), \boldsymbol{\varepsilon_i}(m \times 1)$ (m merupakan jumlah persamaan). Jika $\boldsymbol{\varepsilon_{it}}$ dimana $t = 1,2, \ldots n$ (n merupakan jumlah observasi) adalah elemen dari $\boldsymbol{\varepsilon_i}$, dengan asumsi bahwa $(\varepsilon_{1t}, \varepsilon_{2t}, \ldots , \varepsilon_{mt})$ berdistribusi bebas identik, dimana

- $E(\varepsilon_{it}) = 0$
- $E(\varepsilon_{it}, \varepsilon_{js}) = \begin{matrix} \sigma_{ij}\ , jika\ t = s \\ 0 \quad , jika\ t \neq s \end{matrix}$
- $i = j = 1,2, \ldots , m$
- $t = s = 1,2, \ldots , n$

maka persamaan dapat ditulis dalam bentuk matriks:

$$\begin{bmatrix} \boldsymbol{Y_1} \\ \boldsymbol{Y_2} \\ \vdots \\ \boldsymbol{Y_m} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X_1} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{X_2} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{X_m} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{\beta_1} \\ \boldsymbol{\beta_2} \\ \vdots \\ \boldsymbol{\beta_m} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon_1} \\ \boldsymbol{\varepsilon_2} \\ \vdots \\ \boldsymbol{\varepsilon_m} \end{bmatrix}$$

sehingga

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

dan

- $E(\boldsymbol{\varepsilon}) = \boldsymbol{0}$
- $E(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') = \boldsymbol{V} = \boldsymbol{\Sigma} \otimes \boldsymbol{I_n}$ (berukuran $mn \times mn$)

dimana

$$\boldsymbol{V}(\boldsymbol{\varepsilon}) = \begin{bmatrix} \sigma_{11}\boldsymbol{I_n} & \sigma_{12}\boldsymbol{I_n} & \cdots & \sigma_{1m}\boldsymbol{I_n} \\ \sigma_{21}\boldsymbol{I_n} & \sigma_{22}\boldsymbol{I_n} & \cdots & \sigma_{2m}\boldsymbol{I_n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1}\boldsymbol{I_n} & \sigma_{2m}\boldsymbol{I_n} & \cdots & \sigma_{mm}\boldsymbol{I_n} \end{bmatrix}$$

Matriks variansi kovariansi non-skalar ini merupakan gabungan dari matriks

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1} & \sigma_{m2} & \cdots & \sigma_{mm} \end{bmatrix}$$

dan matriks identitas $(n \times n)\boldsymbol{I_n}$**.** Maka penaksir GLS dari $\boldsymbol{\beta}$ yaitu

$$\widehat{\boldsymbol{\beta}}_{gls} = (\boldsymbol{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\boldsymbol{X}'\mathbf{V}^{-1}\boldsymbol{Y} \text{ (dimana } \boldsymbol{V} = (\boldsymbol{\Sigma_{mm}} \otimes \boldsymbol{I_n})$$

$$= (\boldsymbol{X}'(\Sigma \otimes \boldsymbol{I_n})^{-1}\mathbf{X})^{-1}\boldsymbol{X}'(\Sigma \otimes \boldsymbol{I_n})^{-1}\boldsymbol{Y}$$

$$= (\boldsymbol{X}'(\Sigma^{-1} \otimes \boldsymbol{I_n})\mathbf{X})^{-1}\boldsymbol{X}'(\Sigma^{-1} \otimes \boldsymbol{I_n})\boldsymbol{Y}$$

$$= \begin{bmatrix} \sigma_{11}(\boldsymbol{X_1'X_1}) & \sigma_{12}(\boldsymbol{X_1'X_2}) & \cdots & \sigma_{1m}(\boldsymbol{X_1'X_m}) \\ \sigma_{21}(\boldsymbol{X_2'X_1}) & \sigma_{22}(\boldsymbol{X_2'X_2}) & \cdots & \sigma_{2m}(\boldsymbol{X_2'X_m}) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1}(\boldsymbol{X_m'X_1}) & \sigma_{m2}(\boldsymbol{X_2'X_m}) & \cdots & \sigma_{mm}(\boldsymbol{X_m'X_m}) \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{X_1'}\sum_{\boldsymbol{j}}\sigma_{1j}Y_j \\ \boldsymbol{X_2'}\sum_{\boldsymbol{j}}\sigma_{2j}Y_j \\ \vdots \\ \boldsymbol{X_m'}\sum_{\boldsymbol{j}}\sigma_{mj}Y_j \end{bmatrix} \quad (2.8)$$

dimana σ_{ij} didefinisikan sebagai unsur pada baris ke-i dan kolom ke-j dari Σ^{-1}**.** Persamaan (2.8) adalah model penaksir GLS yang diperluas yang merupakan model penaksir SUR. Untuk efisiensi dalam penulisan, Greene (2000) menyatakan model penaksir SUR yang merupakan pengembangan dari metode GLS dapat ditulis dalam notasi:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{SUR}} = (\boldsymbol{X}'(\Sigma \otimes \boldsymbol{I_n})^{-1}\mathbf{X})^{-1}\boldsymbol{X}'(\Sigma \otimes \boldsymbol{I_n})^{-1}\boldsymbol{Y} \quad (2.9)$$

2.7 Metode Kuadrat Terkecil Tiga Tahap

Metode kuadrat terkecil tiga tahap atau *three stage least square* (3SLS) adalah suatu metode yang diaplikasikan untuk semua persamaan yang terdapat pada parameter secara simultan. Metode ini dikembangkan oleh Theil dan Zellner tahun 1962, sebagai lanjutan dari *two stage least square* (2SLS). Metode 3SLS dapat digunakan untuk menaksir persamaan simultan dengan identifikai *exactly identified*, dan *over identified*.

Berikut adalah tahap-tahap penaksiran persamaan simultan dengan menggunakan metode 3SLS:

Tahap I.

Ide dasar yang dilakukan pada tahap ini adalah untuk menghilangkan korelasi yang terjadi antara variabel endogen yang menjadi variabel prediktor pada persamaan lain yang menyebabkan OLS tidak dapat digunakan untuk menaksir parameter-parameter dari persamaan yang diberikan. Penghilangan korelasi tersebut dilakukan dengan mengubah persamaan simultan menjadi bentuk tereduksi, dimana suatu variabel endogen hanya dijelaskan oleh variabel eksogen. Dengan demikian metode OLS dapat digunakan untuk menaksir parameter struktural dari bentuk yang tereduksi.

Tahap II.

Pada tahap ini, kembali dilakukan OLS pada persamaan struktural yang telah diperoleh dari hasil taksiran pada tahap II dan diperoleh matriks variansi kovariansi galatnya.

Tahap III.

Pada tahap III ini, dengan memanfaatkan matriks variansi kovariansi galat yang diperoleh pada tahap II, persamaan struktural ditaksir kembali menggunakan metode SUR dengan mengganti variabel endogen sebelah kanan persamaan dengan nilai taksirannya yang diperoleh pada tahap I.

2.8 Uji Asumsi Klasik

2.8.1 Uji Normalitas

Uji normalitas data bertujuan untuk menguji apakah dalam model regresi, variabel pengganggu atau galat memiliki distribusi normal atau tidak (Ghozali, 2006). Jika residual yang kita dapatkan berdistribusi normal maka uji signifikansi pengaruh variabel prediktor terhadap variabel respon akan valid. Uji normalitas dilakukan pada nilai residualnya, bukan pada masing-masing variabelnya. Dalam penelitian ini, pengujian normalitas menggunakan uji Shapiro-Wilk dengan pengujian hipotesis sebagai berikut:

1. Hipotesis:

   H_o: Sampel berdistribusi normal

H_1: Sampel tidak berdistribusi normal

2. Taraf signifikansi: α
3. Statistik uji:

   $T_3 = \frac{1}{D}\left[\sum_{i=1}^{k} a_i(X_{n-i+1} - X_i)\right]^2$ $\qquad$ $D = \sum_{i=1}^{n}(X_i - \bar{X})^2$

   Keterangan:

   a_i : koefisien Shapiro-Wilk $\qquad$ X_i : data ke-i

   X_{n-i+1} : data ke $n - i + 1$ $\qquad$ $\bar{X}$: rata-rata data
4. Kriteria keputusan:

   H_0 diterima jika $T_{3\ hitung} < T_{3\ tabel}$

2.8.2 Uji Heterokedastisitas

Tujuan uji heterokedastisitas adalah untuk menguji apakah dalam model regresi terjadi ketidaksamaan variansi dari residual satu pengamatan ke pengamatan yang lain. Jika variansi dari residual satu pengamatan ke pengamatan yang lain tetap, maka disebut homokedastisitas dan jika berbeda disebut heterokedastisitas. Menurut Ghozali (2011), model regresi yang baik adalah homokedastisitas atau tidak terjadi heterokedastisitas. Untuk mengetahui ada atau tidaknya masalah heterokedastisitas, dalam penelitian ini digunakan metode uji Park yaitu meregresikan galat dengan variabel prediktornya dengan model sebagai berikut:

$$\ln(\varepsilon_i^2) = a_0 + a_1 \ln X_{1i} + a_2 \ln X_{2i} + \cdots + a_p \ln X_{pi} + u_i.$$

dengan:

$\ln(\varepsilon_i^2)$: variabel respon

$\ln X_{ji}$: variabel prediktor ($j = 1,2,3, \ldots, n$)

a_i : parameter regresi

u_i : galat ke-i

Langkah-langkah uji hipotesis yang digunakan dalam uji Park yaitu:

1. Hipotesis:

   H_0: sampel tidak mengandung heterokedastisitas

   H_1: sampel mengandung heterokedastisitas
2. Taraf signifikansi: α
3. Statistik uji:

$t_{hitung} = \frac{\hat{\alpha}_i}{Sd(\hat{\alpha}_i)}$

Keterangan:

$\hat{\alpha}_i$: estimator parameter model regresi

$Sd(\hat{\alpha}_i)$: standar deviasi variabel dari $\hat{\alpha}_i$

4. Kriteria keputusan:

   H_0 diterima jika $t_{hitung} \leq t_{tabel}$ atau $t_{hitung} \geq -t_{tabel}$

2.8.3 Uji Multikolinearitas

Istilah multikolinearitas pertama kali ditemukan oleh Ragnar Friech yang berarti adanya hubungan linear yang sempurna atau pasti di antara beberapa atau semua variabel prediktor dari model regresi berganda (Gurajati, 2004). Uji multikolinearitas bertujuan untuk menguji apakah terdapat korelasi antar variabel prediktor pada model regresi. Model regresi yang baik adalah tidak terjadi korelasi di antara variabel prediktor. *Variance inflation factor (VIF)* adalah salah satu cara mendeteksi adanya multikolinearitas. Data dikatakan mengandung masalah multikolinearitas apabila $VIF \geq 10$ (Myers, 1990). *Variance inflation factor* dinyatakan dengan rumus:

$$(VIF)_j = \frac{1}{1-R_j^2}, j = 1,2, \dots, p$$

dimana R_j^2 adalah koefisien determinasi yang dihasilkan bila X_j dijadikan sebagai variabel respon dan diregresikan dengan variabel prediktor lainnya. R_j^2 dapat dihitung dengan rumus berikut:

$$R_j^2 = \frac{\sum_{i=1}^{p} \beta_i \sum_{k=1}^{n} X_{jk} X_{ik}}{\sum_{k=1}^{n} X_{jk}^2}, i \neq j$$

1. Hipotesis :

   $H_0 : VIF < 10$ artinya tidak terdapat multikolinearitas

   $H_1 : VIF \geq 10$ artinya terdapat multikolinearitas

2. Taraf nyata: α
3. Statistik uji:

$$(VIF)_j = \frac{1}{1-R_j^2}$$

4. Kriteria keputusan:

   Jika $F < 10$: H_0 tidak ditolak (tidak terdapat multikolinearitas)

   Jika $VIF \geq 10$: H_0 ditolak (terdapat multikolinearitas)

2.8.4 Uji Autokorelasi

Uji autokorelasi dilakukan untuk melihat adanya korelasi di antara galat variabel-variabelnya. Akibat dari adanya autokorelasi yaitu estimasi masih linear dan tidak bias, namun estimator-estimator tidak efisien. Untuk mengetahui adanya autokorelasi dapat dilakukan uji Durbin Watson (DW).

1. Hipotesis :

   $H_0 : \rho = 0$ artinya tidak ada autokorelasi

   $H_1 : \rho \neq 0$ artinya ada autokorelasi
2. Taraf nyata: α
3. Statistik uji:

   $$d = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} {e_t}^2}$$

   Keterangan:

   e_t : galat waktu t

   e_{t-1} : galat waktu $t-1$
4. Kriteria keputusan:

   Jika $0 < d \leq d_L$, H_0 ditolak (ada autokorelasi positif)

   Jika $d_L < d \leq d_U$, tidak adak keputusan

   Jika $d_U < d \leq 4 - d_U$, H_0 diterima (tidak ada autokorelasi)

   Jika $4 - d_U < d \leq 4 - d_L$, tidak ada keputusan

   Jika $4 - d_L < d \leq 4$, H_0 ditolak (ada autokorelasi negatif)

2.9 Uji Simultan

Masalah simultanitas di dalam persamaan regresi muncul karena variabel endogen berhubungan dengan faktor galat. Salah satu metode uji simultan dikemukakan oleh Hausman (Widarjono, 2013). Uji Hausman dilakukan dengan melakukan regresi antara variabel endogen dengan faktor galat *reduced form*.

1. Hipotesis:

   H_0 : terjadi simultanitas pada data

   H_1 : tidak terjadi simultanitas pada data
2. Taraf signifikansi: α
3. Statistik uji:

   $$r = \frac{\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{\sqrt{\left(\sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}\right)\left(\sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}\right)}} \qquad t_{hitung} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

   Keterangan:

   X_i : nilai variabel X yang ke-i

   Y_i : nilai variabel Y yang ke-i

   r : korelasi antara variabel X dan Y

   n : jumlah observasi
4. Kriteria keputusan:

   Jika $t_{hitung} > t_{tabel}$ maka terjadi simultanitas pada data.

2.10 Metode Pemusatan dan Penskalaan

Pemusatan dan penskalaan data merupakan bagian dari membakukan (*standardized)* variabel. Pemusatan merupakan perbedaan antara masing-masing pengamatan dengan rata-rata dari semua pengamatan setiap variabel, sedangkan penskalaan meliputi gambaran pada kesatuan (unit) standar deviasi dari pengamatan untuk setiap variabel (Kutner, 2004). Berikut ini merupakan pembakuan variabel dependen Y dan variabel independen X:

$$Y_i^* = \frac{Y_i - \bar{Y}}{S_y}, S_y = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1}} \qquad (2.10)$$

$$X_{ij}^* = \frac{X_{ij} - \bar{X}_j}{S_{xj}}, S_{xj} = \sqrt{\frac{\sum_{i=1}^{n}\left(X_{ij} - \bar{X}_j\right)^2}{n-1}}, j = 1,2, \dots, p \qquad (2.11)$$

dimana:

$\bar{Y}$: rata-rata dari Y

$\bar{X}_j$: rata-rata dari X_j

S_{x_j} : standar deviasi dari X_j

S_y : standar deviasi dari Y

Model pemusatan dan penskalaan diperoleh dari model regresi linear berganda persamaan (2.1) yang dibentuk menjadi :

$$Y_i = \beta_0 + \beta_1(X_{i1} - \bar{X}_1) + \beta_1\bar{X}_1 + \beta_2(X_{i2} - \bar{X}_2) + \beta_2X_{i2} + \cdots + \beta_pX_{ip} + \beta_p(X_{ip} - \bar{X}_p) + \varepsilon_i$$

$$= (\beta_0 + \beta_1\bar{X}_1 + \beta_2X_{i2} + \cdots + \beta_pX_{ip}) + \beta_1(X_{i1} - \bar{X}_1) + \beta_2(X_{i2} - \bar{X}_2) + \ldots + \beta_p(X_{ip} - \bar{X}_p) + \varepsilon_i$$

Jika Y_i diberlakukan prosedur pemusatan menggunakan

$$\bar{Y} = \beta_0 + \beta_1\bar{X}_1 + \beta_2\bar{X}_2 + \cdots + \beta_p\bar{X}_p$$

maka diperoleh

$$Y_i - \bar{Y} = \beta_1(X_{i1} - \bar{X}_1) + \beta_2(X_{i2} - \bar{X}_2) + \cdots + \beta_p(X_{ip} - \bar{X}_p) + \varepsilon_i \qquad (2.12)$$

Kemudian pada persamaan (2.12) diberlakukan prosedur penskalaan sesuai persamaan (2.10) dan (2.11), maka diperoleh model regresi berikut:

$$Y_i^* = \beta_1^*X_{1i}^* + \beta_2^*X_{2i}^* + \cdots + \beta_p^*X_{pi}^* + \varepsilon_i^*. \qquad (2.13)$$

Hubungan model regresi terstandarisasi dengan model regresi awal yaitu :

$$\beta_j = \left(\frac{s_y}{s_{xj}}\right)\beta_j^*, j = 1,2, \ldots, p \qquad (2.14)$$

$$\beta_0 = \bar{Y} - \beta_1\bar{X}_1 - \beta_2\bar{X}_2 - \cdots - \beta_p\bar{X}_p \qquad (4.15)$$

Model pada persamaan (2.13) disebut sebagai model regresi yang baku (*standardized regression model)*(Kutner, 2005).

2.11 Metode Regresi Ridge

Regresi ridge diperkenalkan pertama kali oleh Hoerl pada tahun 1962 untuk mengendalikan ketidakstabilan penduga kuadrat terkecil (Hoerl & Kennard, 1970). Regresi ridge bertujuan untuk mengatasi multikolinearitas yang terdapat dalam regresi linear berganda yang mengakibatkan nilai estimasi parameter yang tidak stabil. Regresi ridge merupakan metode estimasi koefisien regresi yang diperoleh melalui penambahan konstanta bias k pada diagonal $\boldsymbol{X'X}$. Nilai estimasi parameter untuk regresi ridge dihitung dengan rumus (Drapper & Smith, 1992):

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{R}} = (\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I})^{-1}\boldsymbol{X}'\boldsymbol{Y}$$

Nilai k pada regresi ridge sama untuk setiap peubah bebas, sedangkan *generalized ridge regression* (GRR) merupakan pengembangan dari prosedur regresi *ridge*, memungkinkan terdapat parameter bias k berbeda untuk setiap peubah bebas. Maka terdapat **K** dimana **K** adalah matriks diagonal $(k_1, k_2, k_3, \dots, k_p)$, sehingga estimasi GRR yaitu:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{GRR}} = (\boldsymbol{X}^{*\prime}\boldsymbol{X}^{*} + \boldsymbol{K})^{-1}\boldsymbol{X}^{*\prime}\boldsymbol{Y}^{*}$$

Menurut Hoerl & Kennard (1970), nilai k ditentukan dengan rumus:

$$k_j = \frac{\sigma^2}{\hat{a}_j^2} \qquad (2.16)$$

dengan $\hat{a}_j^2$: estimator OLS pada data terstandarisasi dan

$$\sigma^2 = \frac{sum\ square\ error}{n-p-1} = \frac{\sum_{i=1}^{n} \varepsilon_i^2}{n-p-1}$$

2.12 Uji Hipotesis F

Sugiyanto (1995) menjelaskan bahwa uji F dilakukan untuk mengetahui apakah variabel prediktor berpengaruh secara bersama-sama terhadap variabel respon. Sebelum pengujian hipotesis, terlebih dahulu akan dicari F hitung yaitu:

$$F_{hitung} = \frac{MSR}{MSE} = \frac{SSR/p}{SSE/(n-p-1)} = \frac{\left(1-\sum_{i=1}^{n} \varepsilon_i^2\right)/p}{\sum_{i=1}^{n} \varepsilon_i^2/(n-p-1)}$$

Langkah-langkah pengujian hipotesis yaitu sebagai berikut:

1. Hipotesis:
   $H_0 : \beta_1 = \beta_2 = \cdots = \beta_p = 0$ (artinya variabel prediktor secara bersama-sama tidak mempunyai pengaruh signifikan terhadap variabel respon.
   $H_1 : \beta_j \neq 0 , j = 1,2, \dots, p$ (artinya variabel prediktor secara bersama-sama mempunyai pengaruh yang signifikan terhadap variabel respon.
2. Tingkat signifikansi : α
3. Statistik uji :

$F_{hitung} = \frac{MSR}{MSE}$

4. Daerah kritis :

   H_0 ditolak jika$F_{hitung} > F_{(\alpha;p,n-p-1)}$

5. Kesimpulan :

   Jika H_0 ditolak berarti variabel prediktor secara bersama-sama mempunyai pengaruh yang signifikan terhadap variabel respon.