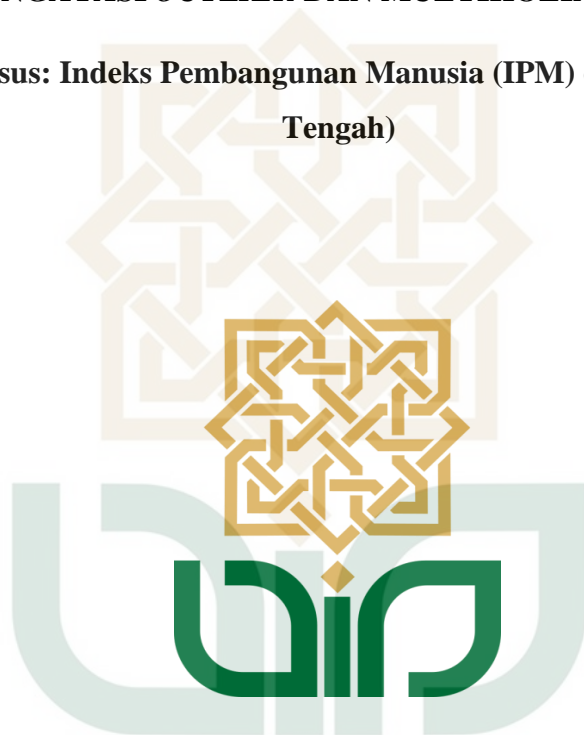


SKRIPSI

**PENERAPAN *GENERALIZED RIDGE REGRESSION (GRR)* PADA
METODE ESTIMASI *ROBUST LEAST MEDIAN SQUARE (LMS)* UNTUK
MENGATASI *OUTLIER* DAN MULTIKOLINIERITAS**

**(Studi Kasus: Indeks Pembangunan Manusia (IPM) di Provinsi Jawa
Tengah)**



**STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA**

**EVI HASTUTI
NIM 20106010031**

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA**

2024



SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi / Tugas Akhir

Lamp :

Kepada

Yth. Dekan Fakultas Sains dan Teknologi

UIN Sunan Kalijaga Yogyakarta

di Yogyakarta

Assalamu 'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Evi Hastuti

NIM : 20106010031

Judul Skripsi : Penerapan Generalized Ridge Regression (GRR) pada Metode Estimasi Robust Least Median Square (LMS) untuk Mengatasi Outlier dan Multikolinieritas

sudah dapat diajukan kembali kepada Program Studi Matematika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Matematika.

Dengan ini kami berharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqasyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu 'alaikum wr. wb.

Yogyakarta, 25 Februari 2024

Pembimbing


Mohammad Farhan Oudratullah

NIP. 19790922 200801 1 011



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS SAINS DAN TEKNOLOGI

Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-447/Un.02/DST/PP.00.9/03/2024

Tugas Akhir dengan judul : Penerapan Generalized Ridge Regression (GRR) pada Metode Estimasi Robust Least Median Square (LMS) untuk Mengatasi Outlier dan Multikolinieritas (Studi Kasus: Indeks Pembangunan Manusia (IPM) di Provinsi Jawa Tengah)

yang dipersiapkan dan disusun oleh:

Nama : EVI HASTUTI
Nomor Induk Mahasiswa : 20106010031
Telah diujikan pada : Kamis, 07 Maret 2024
Nilai ujian Tugas Akhir : A

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Ketua Sidang

Mohammad Farhan Quadratullah, S.Si., M.Si
SIGNED

Valid ID: 65f412ef85e9a



Penguji I

Dr. Muhammad Wakhid Musthofa, S.Si.,
M.Si.
SIGNED

Valid ID: 65f3fe44a118f



Penguji II

Muhamad Rashif Hilmi, S.Si., M.Sc.
SIGNED

Valid ID: 65f3fb0c27870



Yogyakarta, 07 Maret 2024

UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi

Prof. Dr. Dra. Hj. Khurul Wardati, M.Si.
SIGNED

Valid ID: 65f7a8f037416

HALAMAN PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini:

Nama : Evi Hastuti
NIM : 20106010031
Program Studi : Matematika
Fakultas : Sains dan Teknologi

Dengan ini menyatakan bahwa isi skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar sarjana di suatu Perguruan Tinggi dan sesungguhnya skripsi ini merupakan hasil pekerjaan penulis sendiri sepanjang pengetahuan penulis, bukan duplikasi atau saduran dari karya orang lain kecuali bagian tertentu yang penulis ambil sebagai bahan acuan. Apabila terbukti pernyataan ini tidak benar, sepenuhnya menjadi tanggung jawab penulis.

Yogyakarta, 26 Februari 2024



Evi Hastuti

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

MOTTO

Future's gonna be okay.

– Suga



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN PERSEMBAHAN

Alhamdulillah Rabbil Aalamin...

Karya tulis ini saya persembahkan kepada diri saya sendiri, almamater tercinta.

Kepada nenek dan kakek tercinta yang telah membesarkan dan membimbing penulis selama ini, juga teruntuk kedua orang tua yang sangat sabar dan sayang kepada anak-anaknya dalam mendidik kami. Serta untuk adik-adikku tersayang yang selalu memberikan semangat, dukungan, dan keceriaan. Dan juga teman-teman seperjuangan.



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PRAKATA

Assalamualaikum Wr. Wb.

Segala puji dan syukur penulis panjatkan atas kehadiran Allah SWT yang telah memberikan rahmat dan hidayah-Nya berupa keimanan, kesehatan, keberkahan, kesabaran, kelancaran, kekuatan, serta kesempatan. Tidak lupa penulis junjungkan shalawat serta salam kepada Nabi kita Muhammad Saw. Atas berkat rahmat Allah SWT penulis dapat menyelesaikan skripsi yang berjudul “Penerapan *Generallized Ridge Regression* pada Metode Estimasi *Robust Least Median Square* untuk Mengatasi *Outlier* dan Multikolinieritas”.

Skripsi ini disusun sebagai salah satu syarat untuk mendapatkan gelar Sarjana Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga. Selama penulisan skripsi ini, penulis telah mendapatkan bimbingan dan bantuan dari banyak pihak. Oleh karena itu pada kesempatan kali ini penulis ingin menyampaikan ucapan terima kasih kepada :

1. Prof. Dr. Phil Al Makin, MA., selaku Rektorat Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
2. Prof. Dr. Khurul Wardati, M.S., selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
3. Muchammad Abrori, S.Si., M.Kom., selaku Ketua Program Studi Matematika beserta seluruh jajarannya.
4. Arif Munandar , M.Sc., selaku Dosen Pembimbing Akademik yang telah berjasa membimbing penulis selama menjadi mahasiswa program studi Matematika.
5. Mohammad Farhan Qudratullah, S.Si., M.Si., selaku Dosen Pembimbing skripsi yang telah meluangkan waktu ditengah kesibukan beliau, memberikan arahan dan bimbingan kepada penulis dalam proses penulisan skripsi ini.
6. Seluruh Dosen dan Staff Program Studi Matematika Universitas Islam Negeri Sunan Kalijaga Yogyakarta yang telah berbagi ilmu baik dalam bidang akademik maupun non akademik.

7. Dewan penguji yang telah memberikan saran dalam penulisan skripsi ini.
8. Kepada nenek dan kakek penulis, Dasmi dan Darmo Suwito, untuk berdualah skripsi ini penulis persembahkan. Penulis ucapkan terima kasih atas segala cinta dan kasih sayang yang telah diberikan dalam membesarkan dan membimbing penulis selama ini sehingga penulis dapat terus berjuang dalam meraih mimpi dan cita-cita.
9. Kedua orang tua penulis, Ibu Katni dan Bapak Sukamto, untuk berdua juga skripsi ini penulis persembahkan. Terima kasih atas segala kasih sayang yang telah diberikan dalam membimbing dan mendukung penulis selama ini sehingga penulis dapat terus maju untuk meraih mimpi dan cita-cita. Kesuksesan dan segala hal baik yang kedepannya akan penulis dapatkan adalah karena dan untuk kalian berdua.
10. Saudara-saudara penulis, Arif, Linda, dan Rizki, dan Keluarga Besar yang selalu memberikan kekuatan dan doa kepada saya.
11. Teman-teman satu bimbingan tugas akhir (Bimbingan Pak Farhan) yang telah berbagi ilmu, pengalaman dan cerita.
12. Teman yang selalu memberikan semangat dan dukungan di saat suka dan duka selama penyelesaian skripsi ini, Aizah Meilawati.
13. Teman-teman Wonsa yang telah mewarnai kehidupan perkuliahan saya: Cindi, Brenda, Zuva, Sekar, dan Hesty.
14. BTS yang telah menghibur dan memberikan motivasi selama penulisan skripsi ini. Mereka adalah kaca mata penulis untuk melihat indahny berani bermimpi dan keajaiban dalam meraihnya.
15. Dan semua pihak yang tidak dapat penulis sebutkan satu persatu, terima kasih.

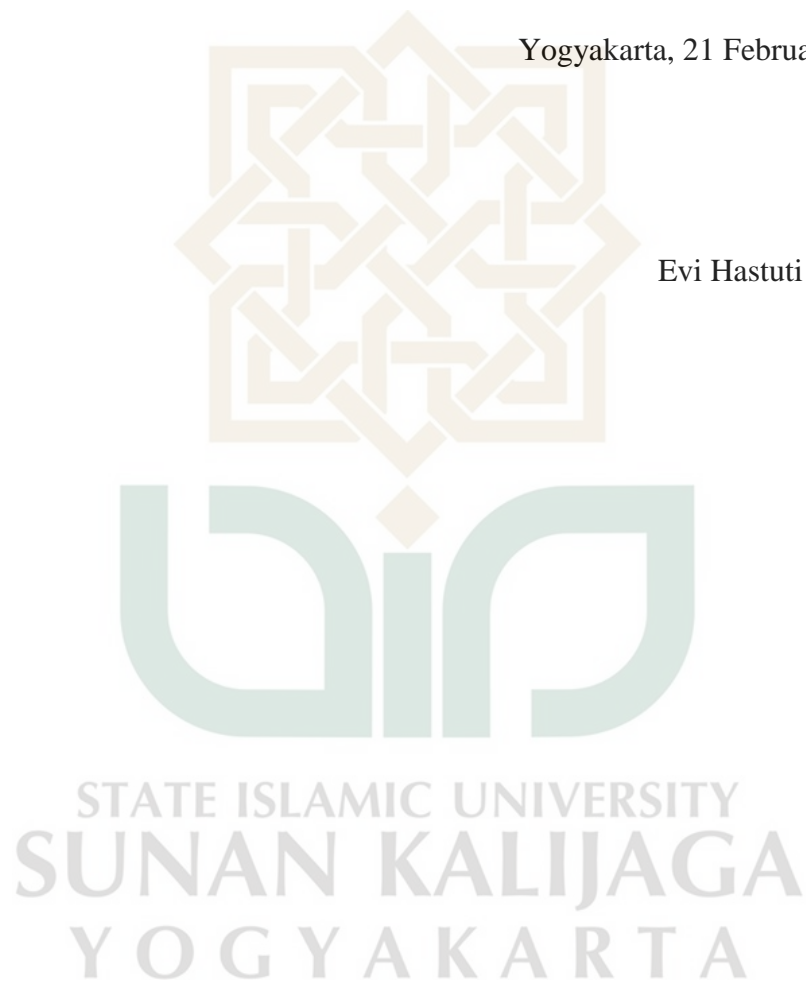
Sebagai manusia biasa Penulis menyadari penyusunan skripsi ini jauh dari kata sempurna karena keterbatasan kemampuan dan ilmu pengetahuan yang dimiliki oleh Penulis. Oleh karena itu atas kesalahan dan kekurangan dalam penulisan skripsi ini, Penulis memohon maaf dan bersedia menerima kritikan yang membangun. Semoga skripsi ini dapat memberikan manfaat khususnya bagi

penulis dan bagi semua yang membutuhkan umumnya. Akhir kata, semoga Allah SWT selalu memberikan rahmat dan berkah-Nya kepada kita semua. Aamiin.

Wassalamualaikum Wr. Wb.

Yogyakarta, 21 Februari 2024

Evi Hastuti



DAFTAR ISI

SKRIPSI.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iv
HALAMAN PERNYATAAN KEASLIAN.....	v
MOTTO.....	vi
HALAMAN PERSEMBAHAN.....	vii
PRAKATA.....	viii
DAFTAR ISI.....	xi
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR.....	xv
DAFTAR LAMPIRAN.....	xvi
DAFTAR SIMBOL.....	xvii
INTISARI.....	xix
ABSTRACT.....	xx
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Batasan Masalah.....	4
1.3 Rumusan Masalah.....	4
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian.....	5
1.6 Tinjauan Pustaka.....	6
1.7 Sistematika Penulisan.....	9
BAB II LANDASAN TEORI.....	11
2.1 Variabel Random.....	11
2.2 Matriks dan Operasi Matriks.....	11
2.2.1 Definisi Matriks.....	11
2.2.2 Penjumlahan Matriks.....	14
2.2.3 Perkalian Matriks.....	14

2.2.4 Transpose Matriks.....	14
2.2.5 Matriks Invers	15
2.2.6 Determinan.....	15
2.3 Analisis Regresi Linier	16
2.3.1 Model Analisis Regresi Linier	18
2.3.2 Estimasi Parameter Regresi Linier	19
2.4 Ordinary Least Square (OLS).....	22
2.4.1 Uji Parameter OLS.....	28
2.5 Uji Asumsi Klasik Analisis Regresi	29
2.5.1 Uji Multikolinearitas.....	30
2.5.2 Uji Heteroskedastisitas	31
2.5.3 Uji Autokorelasi.....	32
2.5.4 Uji Normalitas.....	34
2.7 Outlier.....	34
2.7.1 Pengaruh Outlier	36
2.7.2 Pendeteksi Outlier.....	36
2.8 Multikolinieritas	41
BAB III METODE PENELITIAN.....	46
3.1 Jenis Data dan Penelitian.....	46
3.2 Sumber Data	47
3.3 Variabel Penelitian	47
3.4 Prosedur Analisis.....	48
3.5 Interpretasi Kesimpulan.....	50
3.6 Flowchart.....	50
BAB IV PEMBAHASAN.....	51
4.1 Regresi <i>Robust</i>	51
4.1.1 Estimasi LMS (Least Median Square).....	54
4.1.2 Prosedur Estimasi <i>Least Median Square</i> (LMS)	56
2.9 Regresi <i>Ridge</i>	58
2.9.1 Pembakuan Data	61
2.9.2 Generalized Ridge Regression (GRR).....	63

2.10 <i>Generalized Ridge Regression - Least Median Square (GRR-LMS)</i>	65
2.11 Pengujian Parameter GRR-LMS	66
2.12 Uji Kesesuaian Model	67
2.12.1 Koefisien Determinasi (R^2)	68
2.12.2 Mean Square Error (MSE)	68
2.12.3 Akaike Information Criterion (AIC)	69
BAB V STUDI KASUS	70
5.1 Statistika Deskriptif	70
5.2 Regresi Linier Berganda	73
5.2.1 Ordinary Least Square (OLS)	73
5.2.2. Asumsi Regresi Klasik	75
5.3 Pendeteksian <i>Outlier</i> dan Pengamatan Berpengaruh	78
5.3.1 Studentized Deleted Residual	78
5.3.2 Leverage Value	80
5.3.3 DFFITS	82
5.4 Estimasi Parameter Regresi Robust dengan Metode Least Median Square (LMS)	83
5.5 Pendugaan Least Median Square-Generalized Ridge Regression (LMS-GRR)	84
5.6 Pengujian Parameter <i>LMS-GRR</i>	86
5.7 Pemilihan Model Terbaik	87
BAB VI PENUTUP	89
5.1 Kesimpulan	89
5.2 Saran	90
DAFTAR PUSTAKA	91
LAMPIRAN	93
DAFTAR RIWAYAT HIDUP	109

DAFTAR TABEL

Tabel 1. 1 Perbedaan Penelitian Sebelumnya dengan Penulis.....	7
Tabel 5. 1 Statistika Deskriptif.....	70
Tabel 5. 2. Pendugaan Model Regresi dengan Metode OLS	73
Tabel 5. 3 Hasil Uji Non Multikolinieritas	76
Tabel 5. 4. Penduga model regresi <i>robust</i> dengan LMS	84
Tabel 5. 5. Penduga model regresi dengan <i>LMS-GRR</i>	84
Tabel 5. 6. Perhitungan uji parsial	86
Tabel 5. 7. Perbandingan metode OLS dengan LMS-GRR.....	87



DAFTAR GAMBAR

Gambar 1. <i>Scatter Plot</i>	37
Gambar 2. <i>Box Plot</i>	38
Gambar 5. 1. Grafik <i>Line Plot</i> Variabel.....	72
Gambar 5. 2 Grafik <i>Normal Probability Plot of Residuals</i>	76
Gambar 5. 3. Grafik <i>studentized deleted residuals plot</i>	79
Gambar 5. 4. Grafik <i>leverage value</i> variabel independen (<i>X</i>).....	81
Gambar 5. 5. Grafik <i>DFITS</i>	83



DAFTAR LAMPIRAN

Lampiran 1. Data faktor-faktor yang mempengaruhi IPM di Kabupaten dan Kota Provinsi Jawa Tengah pada Tahun 2022.....	93
Lampiran 2. Syntax R	94
Lampiran 3. Statistika Deskriptif	99
Lampiran 4. Line plot variabel dependen dan independen <i>Line plot</i> variabel dependen	100
Lampiran 5. Metode Ordinary Least Square (OLS).....	102
Lampiran 6. Asumsi Regresi Klasik	103
Lampiran 7. Outlier dan Pengamatan Berpengaruh.....	104
Lampiran 8. Pendugaan Least Median of Square (LMS).	107
Lampiran 9. Pendugaan Generalized Ridge Regression-Least Median of Square (LMS-GRR)	107
Lampiran 10. Pengujian Parameter LMS-GRR	107
Lampiran 11. Uji Model Terbaik	107

DAFTAR SIMBOL

Y_i	Variabel dependen/terikat
β_0	Konstanta regresi
β_1	Kemiringan garis regresi
X_1	Variabel independen/bebas
ε_i	<i>Error</i>
β_i	Parameter/koeffisien regresi ke- i
β_k	Koeffisien regresi pada variabel X_k
X_k	Variabel independen/bebas
β_R	Vektor parameter regresi <i>ridge</i>
\bar{Y}	Rata-rata Y
\bar{X}	Rata-rata X
X^T	X transpose
I	Matriks identitas berukuran $p \times p$
λ	Nilai <i>eigen</i>
k	Vektor <i>eigen</i>
$\hat{\beta}^*$	Estimasi parameter dari metode kuadrat terkecil
$\hat{\beta}_R$	Estimasi koefisien regresi <i>ridge</i>
α	Penaksir tertentu yang belum diketahui
Y	Vektor pengamatan variabel dependen
n	Banyak data
\hat{Y}	Prediksi model regresi
JKT	Jumlah Kuadrat Total
JKR	Jumlah Kuadrat Regresi
JKE	Jumlah Kuadrat <i>Error</i>
MSE	<i>Means Square Error</i>
$E(X)$	Nilai ekspektasi X
Cov atau C	Kovariansi
Var atau σ^2	Variansi

<i>VIF</i>	<i>Variance Inflating Factors</i>
<i>TOL</i>	Toleransi
<i>R</i>	Korelasi
z_i	Komponen utama
p	parameter
c	Koefisien regresi <i>ridge</i>



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

INTISARI

PENERAPAN *GENERALIZED RIDGE REGRESSION* PADA METODE ESTIMASI *ROBUST LEAST MEDIAN SQUARE* UNTUK MENGATASI *OUTLIER* DAN MULTIKOLINIERITAS

Oleh

EVI HASTUTI

20106010031

Analisis regresi linier berganda merupakan analisis regresi yang digunakan untuk menganalisis data tentang hubungan linier dari suatu variabel dependen terhadap beberapa variabel independen. Metode *ridge-robust* adalah salah satu penyelesaian yang digunakan untuk menangani permasalahan multikolinieritas dan *outlier* dalam data. Penggabungan kedua metode regresi tersebut menjadi pilihan yang efektif dibandingkan menerapkannya secara terpisah. Seperti pada data Indeks Pembangunan Manusia (IPM), usia harapan hidup saat lahir, harapan lama sekolah, rata-rata lama sekolah, pengeluaran perkapita disesuaikan, dan banyaknya puskesmas pada 35 kabupaten/kota di Provinsi Jawa Tengah pada tahun 2022 yang melanggar asumsi non multikolinieritas dan terdapat indikasi adanya *outlier* pada data. Berdasarkan permasalahan yang ada, penggunaan metode *Generalized Ridge Regression (GRR)* dengan estimasi *Least Median Square (LMS)* adalah pilihan yang tepat untuk menduga parameter regresi. Pengujian untuk menentukan model terbaik dilakukan dengan menggunakan *standard error*, R^2 (koefisien determinasi), dan *AIC (Akaike Determination Criterion)*, didapat kesimpulan bahwa metode regresi *ridge-robust* menghasilkan model yang lebih baik dibandingkan dengan *Ordinary Least Square (OLS)*.

Kata kunci: Regresi Ridge, Regresi Robust, *Generalized Ridge Regression*, *Least Median Square*, *Ordinary Least Square*

ABSTRACT

APPLICATION OF GENERALIZED RIDGE REGRESSION IN ROBUST LEAST MEDIAN SQUARE ESTIMATION METHOD TO ADDRESS OUTLIERS AND MULTICOLLINEARITY

By

EVI HASTUTI

20106010031

Multiple linear regression analysis is a statistical method used to analyze the linear relationship between a dependent variable and multiple independent variables. The ridge-robust method is one of the solutions used to handle multicollinearity and outliers in data. Combining these two regression methods is an effective choice compared to applying them separately. For instance, in the data on the Human Development Index (HDI), life expectancy at birth, expected years of schooling, mean years of schooling, adjusted per capita expenditure, and the number of health centers in 35 districts/cities in Central Java Province in 2022, which violate the assumption of non-multicollinearity and indicate the presence of outliers in the data. Based on the existing problems, the use of Generalized Ridge Regression (GRR) with Least Median Square (LMS) estimation is the right choice for estimating regression parameters. Testing to determine the best model is done using standard error, R^2 (coefficient of determination), and AIC (Akaike Information Criterion), concluding that the ridge-robust regression method produces a better model compared to Ordinary Least Squares (OLS).

Keywords: Ridge Regression, Robust Regression, Generalized Ridge Regression, Least Median Square, Ordinary Least Square

BAB I

PENDAHULUAN

1. 1 Latar Belakang Masalah

Menurut Gaspersz (1989), statistika merupakan ilmu pengetahuan yang berkaitan dengan bagaimana pengumpulan data, pengolahan serta analisis data, pengambilan kesimpulan hingga keputusan yang didasarkan pada fakta. Statistika memiliki peranan penting dalam penyelesaian permasalahan yang ada dalam kehidupan manusia, seperti dalam bidang ekonomi, kesehatan, kependudukan, dan lain sebagainya. Permasalahan-permasalahan yang ada dalam bidang-bidang tersebut menjadikan munculnya berbagai macam solusi yang berupa suatu hasil analisis regresi, di mana hasil analisis regresi dapat dijadikan sebagai pengambilan keputusan dari permasalahan yang ada.

Istilah regresi pertama kali diperkenalkan oleh Sir Francis Galton (1822-1911) yang terdapat dalam *Journal of the Anthropological Institute, Volume 12*, halaman 246-263 tahun 1885. Pada jurnal tersebut dijelaskan bahwa keturunan dari biji cenderung tidak memiliki kesamaan dengan induknya dalam hal ukuran atau besarnya, akan tetapi lebih mendekati rata-rata. Maksudnya yaitu ukuran keturunan dari biji cenderung lebih kecil daripada induknya apabila induknya lebih besar dan sebaliknya, apabila induknya lebih kecil maka keturunannya akan lebih besar dibandingkan dengan induknya (Draper dan Smith, 1992).

Analisis regresi adalah analisis statistik yang digunakan untuk menentukan model hubungan sebab akibat antara variabel dependen (terikat) dengan variabel independen (bebas). Variabel dependen (Y) merupakan variabel yang dipengaruhi oleh variabel variabel independen, sebaliknya variabel independen (X) merupakan variabel yang nilainya dapat ditentukan terlebih dahulu dan tidak tergantung oleh variabel lainnya. Jika variabel independen (X) hanya satu, maka analisis tersebut dapat dikatakan sebagai analisis regresi linear sederhana. Akan tetapi apabila variabel independennya terdiri lebih dari satu variabel maka dikatakan sebagai analisis regresi linear berganda (Supandi, 2020).

Analisis regresi linear berganda adalah perluasan dari analisis regresi linear sederhana. Di mana dalam regresi linear sederhana terdapat analisis hubungan dua variabel, yaitu sebuah variabel independen (X) dengan sebuah variabel dependen yang dapat ditulis dalam persamaan linear $Y' = a + bX$, dengan tujuan membuat penduga tentang besarnya nilai variabel dependen (Y) yang didasarkan pada nilai variabel independen (X).

Pada umumnya analisis regresi digunakan dalam menentukan model hubungan sebab akibat antara variabel independen ($X_1, X_2, X_3, \dots, X_k$) dengan variabel dependen (Y). Pembentukan model regresi diawali dengan dugaan awal jika suatu variabel memiliki hubungan terhadap satu atau beberapa variabel lainnya. Dalam penentuan model regresi dilakukan penaksiran parameter agar diperoleh suatu informasi dari data. Metode analisis regresi yang sering digunakan untuk mendapatkan penduga regresi adalah Metode Kuadrat Terkecil atau *Ordinary Least Square* (OLS). Metode ini merupakan salah satu metode estimasi yang sering dijumpai dalam melakukan analisis regresi. Metode ini ditemukan oleh seorang matematikawan Jerman bernama Carl Friedrich Gauss.

OLS merupakan metode yang digunakan untuk mengestimasi suatu garis regresi dengan cara mencari nilai prediksi dengan nilai kenyataannya. Metode ini memerlukan beberapa asumsi klasik yang harus dipenuhi, seperti uji normalitas, multikolinieritas, heteroskedastisitas, dan autokorelasi. Asumsi-asumsi yang ada harus dipenuhi supaya didapatkan dugaan parameter yang tidak bias.

Akan tetapi, tidak jarang ketika melakukan uji asumsi klasik sering terjadi adanya pelanggaran asumsi sehingga model regresi yang dihasilkan menjadi tidak tepat dan hasil estimasi menjadi kurang akurat. Terjadinya pelanggaran asumsi dapat disebabkan beberapa faktor, beberapa diantaranya adalah asumsi non multikolinieritas dan data observasi mengandung *outlier* atau pencilan.

Pengingkaran terhadap asumsi non multikolinieritas dapat mengakibatkan model regresi yang diperoleh menjadi bias, akibatnya dugaan koefisien parameter menjadi tidak akurat. Salah satu penyelesaian yang dapat dilakukan untuk mengatasi masalah non multikolinieritas yaitu dengan menggunakan regresi *ridge*. Regresi *ridge* diperkenalkan oleh Hoerl dan Kennard pada 1970 untuk menangani

multikolinieritas. Masih di tahun yang sama Hoerl dan Kennard memperkenalkan *Generalized Ridge Regression (GRR)* yang merupakan pengembangan dari *Ordinary Ridge Regression (ORR)* atau regresi *ridge* standar.

Permasalahan selanjutnya yaitu data observasi mengandung *outlier* (pencilan). Adanya *outlier* dapat mempengaruhi dalam pendugaan koefisien regresi. Untuk mengatasi permasalahan tersebut diperlukan metode estimasi yang akan tetap berjalan dengan baik saat mengestimasi parameter-parameter regresi, yaitu dengan regresi *robust*.

Regresi *robust* merupakan metode regresi yang digunakan ketika data pengamatan terdapat *outlier* atau distribusi dari sisaan tidak normal yang memberikan pengaruh pada model regresi yang dihasilkan. Keberadaan *outlier* dalam suatu analisis regresi dapat menimbulkan beberapa masalah yang pada akhirnya akan mempengaruhi keakuratan hasil analisis yang dibuat oleh peneliti, sehingga asumsi yang dihasilkan menjadi tidak tepat atau estimasi yang dihasilkan menjadi bias.

Akan tetapi menghilangkan keberadaan *outlier* dalam melakukan analisis bukanlah pilihan yang tepat untuk dilakukan. Terkadang adanya *outlier* dalam suatu data penelitian justru memberikan informasi penting yang tidak diberikan oleh data lainnya, seperti adanya keadaan yang mungkin saja sangat penting dan perlu dilakukan penyelidikan lebih lanjut. Salah satu metode yang dapat digunakan untuk mendapatkan nilai penduga parameter dalam pemodelan regresi adalah dengan menggunakan metode *Least Median Square (LMS)*. LMS merupakan metode yang memiliki nilai *breakdown* yang tinggi dan merupakan metode penduga parameter regresi *robust* dengan cara meminimumkan median kuadrat galat.

Permasalahan-permasalahan yang disebutkan di atas akan menjadi efektif apabila pelanggaran asumsi keduanya diatasi dengan cara menggabungkan metode regresi *robust* dan regresi *ridge*. Hal pertama yang harus dilakukan ketika melakukan pemodelan regresi dengan metode LMS yaitu dengan melakukan uji asumsi dan identifikasi adanya *outlier* pada data observasi. Setelah itu, dilakukan analisis regresi *robust* dengan estimasi LMS dan GRR. Pada penelitian ini akan

dilakukan analisis regresi *robust* dengan diterapkan pada *Generalized Ridge Regression (GRR)*.

Oleh karena itu, peneliti tertarik untuk melakukan penelitian. Pada penelitian ini akan menggunakan data Indeks Pembangunan Manusia (IPM). Hasil dari penelitian akan memberikan informasi tentang analisis regresi *robust* dan regresi *ridge* yang tentunya akan sangat bermanfaat bagi pembaca.

1.2 Batasan Masalah

Guna memperjelas pembahasan dalam skripsi ini, maka penulis perlu memberikan batasan-batasan yaitu sebagai berikut:

1. Berfokus pada penentuan estimasi parameter regresi yang mengandung *outlier* dan adanya pelanggaran asumsi non multikolinieritas, dengan kata lain data yang digunakan merupakan data yang mengandung *outlier* dan melanggar uji asumsi klasik non multikolinieritas.
2. Metode yang digunakan dalam penelitian adalah *Least Median Square (LMS)*, *Generalized Ridge Regression (GRR)*, dan *Ordinary Least Square (OLS)*.

1.3 Rumusan Masalah

Berdasarkan pada latar belakang yang telah dijelaskan sebelumnya maka diperoleh rumusan masalah dari penelitian ini yaitu:

1. Bagaimana langkah-langkah penerapan *Generalized Ridge Regression (GRR)* pada *Least Median Square (LMS)* untuk mengatasi data observasi yang mengandung adanya *outlier* dan multikolinieritas?
2. Bagaimana persamaan model *Ordinary Least Square (OLS)* dan *Generalized Ridge Regression (GRR)*- *Least Median Square (LMS)* dalam menganalisis Indeks Pembangunan Manusia (IPM) di Provinsi Jawa Tengah tahun 2022?
3. Bagaimana perbandingan model antara estimasi *Generalized Ridge Regression (GRR)*-*Least Median Square (LMS)* dengan metode *Ordinary Least Square (OLS)* dalam menganalisis Indeks Pembangunan Manusia (IPM) di Provinsi Jawa Tengah tahun 2022?

1.4 Tujuan Penelitian

Berdasarkan pada latar belakang dan rumusan masalah yang telah dijelaskan sebelumnya maka penelitian ini dilakukan dengan tujuan untuk:

1. Mengetahui langkah-langkah penerapan *Generalized Ridge Regression* (GRR) pada estimasi *Least Median Square* untuk mengatasi data yang observasi yang mengandung *outlier* dan multikolinieritas.
2. Mengetahui persamaan model *Ordinary Least Square (OLS)* dan *Generalized Ridge Regression (GRR)-Least Median Square (LMS)* dalam menganalisis Indeks Pembangunan Manusia (IPM) di Provinsi Jawa Tengah tahun 2022.
3. Mengetahui perbandingan model antara estimasi *Generalized Ridge Regression (GRR)-Least Median Square (LMS)* dengan metode *Ordinary Least Square (OLS)* dalam menganalisis Indeks Pembangunan Manusia (IPM) di Provinsi Jawa Tengah tahun 2022

1.5 Manfaat Penelitian

Hasil dari penelitian ini diharapkan dapat memberikan manfaat untuk pihak-pihak berikut ini, antara lain:

1. Penulis
Penelitian ini memberikan sejumlah manfaat yang dapat memengaruhi pengembangan diri dan karir penulis di masa depan. Dengan penelitian ini memungkinkan penulis untuk mengembangkan ketrampilan penelitian yang lebih baik, termasuk dalam perumusan masalah penelitian, pemilihan metodologi yang tepat, pengumpulan data, analisis statistik, dan interpretasi temuan yang mana akan berguna dalam penelitian masa depan. Selain itu dengan penelitian ini juga menjadi bentuk partisipasi dalam pengembangan ilmu matematika khususnya dalam bidang statistika tentang pemodelan regresi *robust*. Serta menjadi bagian penting portofolio akademik penulis yaitu sebagai salah satu syarat kelulusan dalam mendapatkan gelar Strata 1.
2. Pembaca
Adapun manfaat yang diperoleh pembaca yaitu dapat memberikan pemahaman yang mendalam mengenai pemodelan regresi *robust* pada

penerapan regresi *ridge*. Selain itu penelitian ini diharapkan dapat menjadi dasar bagi pengambilan keputusan yang berguna untuk merancang kebijakan yang lebih efektif dan relevan.

1.6 Tinjauan Pustaka

Pada bagian ini akan menguraikan beberapa penelitian-penelitian yang telah dilakukan sebelumnya untuk dimanfaatkan sebagai materi pendukung penelitian yang dilakukan penulis. Melalui pemeriksaan mendalam terhadap literature, penulis memberikan gambaran menyeluruh mengenai pemahaman yang telah diperoleh dalam kerangka konseptual ini. Selain itu, bab ini berperan sebagai penjelasan mengenai hubungan antara penelitian terdahulu dengan pemilihan metode dalam penelitian. Adapun tinjauan pustaka dalam penulisan skripsi ini adalah sebagai berikut:

1. Penelitian berjudul: “Mengatasi Dampak Multikolinieritas dalam Analisis Regresi Berganda dengan Metode Regresi *Ridge* dan *Principal Component Analysis (PCA)*” oleh Sri Wahyuningsih Mahasiswi Program Studi Mathematical UIN Sunan Kalijaga Fakultas Sains dan Teknologi tahun 2018. Penelitian ini membandingkan hasil analisis dengan metode regresi *ridge* dan *principal component analysis (PCA)* dalam mengatasi multikolinieritas.
2. “Perbandingan Estimasi *Least Median Square* dengan Estimasi *Least Trimmed Square* pada Regresi *Robust*” oleh Noor Hidayati Tri Puji Lestari Mahasiswi Program Studi Matematika UIN Sunan Kalijaga Fakultas Sains dan Teknologi tahun 2012. Penelitian ini membandingkan tingkat kerobustan dua metode estimasi antara *Least Median Square (LMS)* dengan estimasi *Least Trimmed Square (LTS)* pada data laporan keuangan Bank Syariah Mandiri (BSM).
3. Penelitian berjudul: “Penerapan Metode *Generalized Ridge Regression* dalam Mengatasi Masalah Multikolinieritas” oleh Ni Ketut Tri Utami, I Komang Gde Sukarsa, I Putu Eka Nila Kencana Mahasiswa Program Studi Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Udayana tahun 2013. Penelitian ini membahas tentang penerapan *Generalized Ridge Regression* dalam mengatasi permasalahan multikolinieritas terhadap data kebutuhan tenaga kerja pada 17 Rumah Sakit Angkatan Laut U.S.

4. Penelitian berjudul: “Penerapan Estimasi Parameter Regresi *Ridge* Menggunakan Metode Algoritma *Schall* dan *Akaike Information Criterion* (AIC).” Oleh Linaksanan Program Studi Matematika Fakultas Sains dan Teknologi UIN Sunan Kalijaga. Penelitian ini membahas tentang cara menentukan model terbaik regresi *ridge* dengan menggunakan metode Algoritma *Schall* dan *Akaike Information Criterion* (AIC).
5. Penelitian berjudul: “Pemilihan Model Terbaik Regresi *Ridge* Tingkat Laju Inflasi Indonesia Tahun 2019 Menggunakan Metode Algoritma *Schall* dan *Akaike’s Information Criterion* (AIC)” oleh Royhana Devi Program Studi Matematika Fakultas Sains dan Teknologi UIN Sunan Kalijaga. Penelitian ini membandingkan efektifitas pemilihan model terbaik pada regresi *ridge*, metode Algoritma *Schall* dan *Akaike’s Information Criterion* (AIC).

Beberapa penelitian sebelumnya menjadi inspirasi untuk penulis melakukan penelitian. Perbedaan penelitian ini dengan penelitian sebelumnya terletak pada penggunaan metode estimasi dan data observasi yang digunakan pada model estimasi regresi *robust* terhadap suatu permasalahan. Pada penelitian-penelitian sebelumnya fokus pembahasan hanya membandingkan metode estimasi *robust* atau menggunakan estimasi regresi *robust* lainnya yang berbeda. Rangkuman perbedaan penelitian terdahulu dengan penulis disajikan pada **Tabel 1.1** di bawah:

Tabel 1. 1 Perbedaan Penelitian Sebelumnya dengan Penulis

No.	Nama	Permasalahan	Metode	Objek
1	Sri Wahyuning-sih	Mengatasi Dampak Multikolinieritas	Regresi <i>Ridge</i> dan <i>Principal Component</i>	Hasil penjualan toko roti

No.	Nama	Permasalahan	Metode	Objek
	(UIN, 2018)		<i>Analysis (PCA)</i>	Ana Bakery tahun 2017
2	Noor Hidayati Tri Puji Lestari (UIN, 2012)	Mengatasi indikasi adanya <i>outlier</i>	<i>Least Median Square (LMS)</i> dan <i>Least Trimmed Square (LTS)</i>	Data laporan keuangan Bank Syariah Mandiri (BSM).
3	Ni Ketut Tri Utami, I Komang Gde Sukarsa, I Putu Eka Nila Kencana (Universitas Udayana, 2013)	Mengatasi Masalah Multikolinieritas	<i>Generalized Ridge Regression (GRR)</i>	Data kebutuhan tenaga kerja pada 17 Rumah Sakit Angkatan Laut U.S.
4	Linaksanan (UIN, 2021)	Mengatasi permasalahan multikolinieritas	Algoritma <i>Schall</i> dan <i>Akaike Information Criterion (AIC)</i>	Data Tingkat Pengangguran Terbuka (TPT) Provinsi Jawa Tengah.
5	Royahana Devi (2021)	Mengatasi permasalahan multikolinieritas	Algoritma <i>Schall</i> dan <i>Akaike's</i>	Tingkat laju inflasi tahun 2019

No.	Nama	Permasalahan	Metode	Objek
			<i>Information Criterion (AIC)</i>	
6	Evi Hastuti (UIN, 2024)	Mengatasi permasalahan <i>outlier</i> dan multikolinieritas	<i>Generalized Ridge Regression (GRR)</i> dan <i>Least Median Square (LMS)</i>	Indeks Pembangunan Manusia (IPM) Provinsi Jawa Tengah tahun 2022

Pada penelitian penulis memiliki beberapa perbedaan dan kesamaan antara penelitian-penelitian sebelumnya, yaitu metode yang digunakan terdapat kesamaan yang mana sama-sama menerapkan pada regresi *ridge* dan menggunakan metode estimasi regresi *robust*. Akan tetapi berbeda dalam pemilihan obyek penelitian yang digunakan.

1.7 Sistematika Penulisan

Penulisan skripsi ini dibagi dalam enam bab supaya memberikan gambaran secara menyeluruh serta mempermudah dalam memahami penelitian ini. Adapun susunan pembagian masing-masing bab dalam penelitian ini yaitu sebagai berikut:

Bab I : Pendahuluan

Bab ini berisi pembahasan tentang pendahuluan dari tema yang diangkat dalam tugas akhir yang meliputi latar belakang masalah, rumusan masalah, tujuan penelitian, tinjauan pustaka, dan sistematika penulisan.

Bab II : Landasan Teori

Bab ini membahas tentang landasan teori yang digunakan sebagai dasar dalam penelitian, meliputi variabel random, matriks dan operasi matriks, regresi linier, *outlier*, uji asumsi klasik, metode kuadrat terkecil, regresi *robust* dan *ridge*, serta uji kesesuaian model.

Bab III : Metode Penelitian

Bab ini membahas mengenai metodologi penelitian yang digunakan dalam penelitian penulis, jenis penelitian, sumber dan data penelitian, populasi dan sampel, variabel penelitian, alat penelitian, metode analisis data, tahap penelitian dan penyimpulan data.

Bab IV : Pembahasan

Bab ini membahas tentang regresi *robust* estimasi *Least Median Square (LMS)* dan penerapannya pada regresi *ridge Generalized Ridge Regression* dalam mengatasi permasalahan multikolinieritas dan *outlier*.

Bab V : Studi Kasus

Bab ini membahas hasil dari analisis regresi *robust* dari analisis data dan pembahasan penelitian pada Indeks Pembangunan Manusia di Provinsi Papua menggunakan bantuan *software* R dalam data yang mengandung *outlier* dan non multikolinieritas.

Bab VI : Penutup

Bab ini membahas tentang kesimpulan yang didapatkan dari pembahasan sebelumnya dan juga berisi saran-saran terkait dengan penulisan tugas akhir sebagai kekurangan atau kelebihan dari hasil analisis yang dilakukan serta saran atas hasil penelitian yang berkaitan dengan penelitian selanjutnya.

Daftar Pustaka

Bagian ini berisi sumber informasi yang digunakan penulis dalam penyusunan tugas akhir, yang berupa: jurnal, buku, artikel maupun sumber-sumber penunjang yang terdapat di internet.

BAB VI PENUTUP

5.1 Kesimpulan

Berdasarkan hasil dan pembahasan yang telah dilakukan tentang regresi *robust* estimasi *Least Median Square (LMS)* dan regresi *ridge Generalized Ridge Regression (GRR)* dengan studi kasus Indeks Pembangunan Manusia (IPM) Provinsi Jawa Tengah pada tahun 2022 maka, dapat diambil kesimpulan sebagai berikut:

a. Langkah-langkah sistematis penerapan *Generalized Ridge Regression (GRR)* pada estimasi *Least Median* untuk mengatasi data observasi yang mengandung adanya *outlier* dan multikolinieritas yaitu:

1. Uji asumsi klasik
2. Deteksi *outlier* dan pengamatan berpengaruh
3. Estimasi LMS dan penerapan GRR
4. Uji signifikansi parameter dan kesesuaian model
5. Pemilihan model terbaik.

b. Persamaan model *Ordinary Least Square (OLS)* dan *Generalized Ridge Regression (GRR)- Least Median Square (LMS)* dalam menganalisis Indeks Pembangunan Manusia (IPM) di Provinsi Jawa Tengah tahun 2022 yaitu:

Metode OLS menghasilkan model sebagai berikut:

$$Y = 6,5499 + 0,4626X_1 + 0,8971X_2 + 1,3414X_3 + 0,8329X_4$$

Dengan nilai $R^2 = 0,9987$ dan nilai $AIC = -20,8375$.

Sedangkan, untuk penerapan GRR pada LMS menghasilkan model sebagai berikut:

$$Y = -59,557 + 1,512X_1 + 1,045X_2 + 0,763X_3 + 4,022e - 06X_4 - 0,016X_5$$

Dengan nilai $R^2 = 0,9994$ dan nilai $AIC = -44,9374$.

c. Perbandingan model antara estimasi *Generalized Ridge Regression (GRR)-Least Median Square (LMS)* dengan metode *Ordinary Least Square*

(*OLS*) dalam menganalisis Indeks Pembangunan Manusia (IPM) di Provinsi Jawa Tengah tahun 2022 menghasilkan model di mana, metode *Generalized Ridge Regression* dengan estimasi *Least Median Square* merupakan pilihan yang tepat apabila data penelitian terdapat indikasi adanya *outlier* dan terjadi pelanggaran asumsi non multikolinieritas. Berdasarkan indikator pemilihan model terbaik, diambil kesimpulan bahwa model yang didapat dari metode LMS-GRR menghasilkan model yang lebih baik dibandingkan dengan model yang didapat dari metode OLS. Di mana untuk GRR-LMS memiliki

5.2 Saran

Saran yang dapat diberikan penulis untuk penelitian selanjutnya yaitu dapat menerapkan regresi *robust* terhadap metode lainnya seperti estimasi LTS atau dapat diterapkan pada data di bidang yang lainnya.



DAFTAR PUSTAKA

- Adiningsih, S. (2001). *Statistik*. Yogyakarta: BPFE-Yogyakarta.
- Akbar, dan Maftukhah. 2007. *Optimasi Kekuatan Tourqe pada Lampu TL*. Jurnal ilmiah Sains dan Teknologi. 6(1): 218-229
- Chen, Colin. 2002. *Robust Regression and Outlier Detection with the ROBUSTREG Procedur*. Paper 256-27
- Derquenne, C. (1993). *Outlier Detection Before Running Statistical Methods. Theory of Probability and its Applications*, 37(2): 323-4. doi:10.1137/1137066.
- Draper, N. R., & Smith, H. 1992. *Analisis Regresi Terapan*. Jakarta: Gramedia Pustaka Utama
- Eko, H. (2024). *Statistika Dasar*. DI. Yogyakarta: Sinar Pustaka.
- Firdaus, Muhammad. 2004. *Ekonometrika Suatu Pendekatan Aplikasi*. Jakarta: PT Bumi Aksara
- Gaspersz, V. (1989). *Statistika*. Bandung: Armico.
- Ghozali, I. (2016) *Aplikasi Analisis Multivarite dengan Program IBM SPSS 23*. Edisi 8. Semarang: Badan Penerbit Universitas Diponegoro.
- Gujarati, D. N. 2003. *Basic Econometric Fourth Edition*. New York: Mc Graw-Hill Companies
- Huber, P.J. 1981. *Robust Statistic*. Canada: John Wiley & Sons Inc.
- Hoerl, A.E. and R.W. Kennard. 1970. "Ridge Regression: Biased Estimation for Nonorthogonal Problems". *Technometrics*, 12 (1970a), 55- 67. (Republished in *Technometrics*. 42 (2000), 80-86).
- Hoerl, A.E. and R.W. Kennard. 1970. "Ridge Regression: Applications to Nonorthogonal Problems". *Technometrics*, Vol. 12, No. 1. (Feb., 1970b), pp. 69-82. <http://statgen.ucr.edu/download/course/STAT288/hoerl70b.pdf>.

Diakses tanggal 14 Juni 2012.

- Imrona, Mahmud. 2009. *Aljabar Linier Dasar*. Jakarta: Penerbit Erlangga.
- J. Supranto, *Statistika, Teori dan Aplikasi*, edisi-5, (Jakarta: Erlangga, 1989) hal. 276
- Kerns, G. J. 2010. *Introduction to Probability Statistics Using R First Edition*. ISBN: 978-0-557-24979-4.
- Kutner, M. H., C. J. Nachtsheim, J. Neter, dan W. Li. 2005. *Applied Linear Statistical Models: Fifth Edition*. New York: McGrawHill
- Maronna, R.A., Martin, R.D., Yohai, V.J., dan Barrera, M.S. 2019. *Robust Statistics*. John Wiley and Sons Ltd, Hoboken
- Mortgomery, D.C., Peck, E. A., dan Vining, G. G. (1982). *Introducing to Linear Regression Analysis*. New York: John Whilley and Sons Inc.
- Rousseeuw, P. J. (1987). *Robust Regression and Outlier Detection*. New York: Wiley and Sons.
- Sugiyarto. 2021. *Pengantar Statistika Matematika 1*. 1st ed. DI Yogyakarta: Magnum Pustaka Utama.
- Supandi, Epha Diana. 2020. *Statistika dan Terapannya*. Bandung: PT Refika Aditama.
- Susanti, Yuliana. 2013. *Optimasi Model Regresi Robust Untuk Memprediksi Produksi Kedelai Di Indonesia*. Prosiding ISBN : 978-979-16353-9-4. FMIPA