

**PENERAPAN REGRESI *ROBUST* PADA FAKTOR-
FAKTOR YANG MEMENGARUHI KEMISKINAN DI
PROVINSI SULAWESI SELATAN MENGGUNAKAN
PENDUGA *LEAST TRIMMED SQUARE***

SKRIPSI



NAJMAWATI ADE

H051181323

**PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

UNIVERSITAS HASANUDDIN

MAKASSAR

FEBRUARI 2024



**PENERAPAN REGRESI *ROBUST* PADA FAKTOR-
FAKTOR YANG MEMENGARUHI KEMISKINAN DI
PROVINSI SULAWESI SELATAN MENGGUNAKAN
PENDUGA *LEAST TRIMMED SQUARE***

SKRIPSI

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

NAJMAWATI ADE

H051181323

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

MAKASSAR

FEBRUARI 2024

ii



LEMBAR PERNYATAAN KEOTENTIKAN

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

Penerapan Regresi *Robust* pada Faktor-Faktor yang Memengaruhi Kemiskinan di Provinsi Sulawesi Selatan Menggunakan Penduga *Least Trimmed Square*

adalah benar hasil karya saya sendiri, bukan hasil plagiat dan belum pernah dipublikasikan dalam bentuk apapun

Makassar, 1 Februari 2024



Najmawati Ade

NIM H051181323



PENERAPAN REGRESI *ROBUST* PADA FAKTOR-FAKTOR
YANG MEMENGARUHI KEMISKINAN DI PROVINSI
SULAWESI SELATAN MENGGUNAKAN PENDUGA *LEAST
TRIMMED SQUARE*

Disetujui Oleh:

Pembimbing Utama

Drs. Raupong, M.Si.

NIP. 19621015 198810 1 001

Pembimbing Pendamping

Andi Kresna Jwa, S.Si., M.Si.

NIP. 19751228 200003 1 001



NIP. 19770808 200501 2 002

Pada 1 Februari 2024



HALAMAN PENGESAHAN

Skripsi ini diajukan oleh:

Nama : Najmawati Ade
NIM : H051181323
Program Studi : Statistika
Judul Skripsi : Penerapan Regresi *Robust* pada Faktor-Faktor yang Memengaruhi Kemiskinan di Provinsi Sulawesi Selatan Menggunakan Penduga *Least Trimmed Square*

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

DEWAN PENGUJI

1. Ketua : Drs. Raupong, M.Si.

(.....)

2. Sekretaris : Andi Kresna Jaya, S.Si., M.Si.

(.....)

3. Anggota : Dra. Nasrah Sirajang, M.Si.

(.....)

4. Anggota : Siswanto, S.Si., M.Si.

(.....)

Ditetapkan di : Makassar

Tanggal : 1 Februari 2024



KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Segala puji hanya milik Allah *Subhanallahu Wa Ta'ala* atas segala limpahan rahmat dan hidayah-Nya yang telah diberikan kepada penulis sampai saat ini. Shalawat dan salam senantiasa tercurahkan kepada baginda Rasulullah *Shallallahu 'Alaihi Wa sallam*, kepada para keluarga, tabi'in, tabi'ut tabi'in, serta orang-orang sholeh yang haq hingga kadar Allah berlaku atas diri-diri mereka. *Alhamdulillahirobbil'aalamiin*, berkat rahmat dan kemudahan yang diberikan oleh Allah *Subhanallahu Wa Ta'ala*, penulis dapat menyelesaikan skripsi yang berjudul **“Penerapan Regresi Robust pada Faktor-Faktor yang Memengaruhi Kemiskinan di Provinsi Sulawesi Selatan Menggunakan Penduga Least Trimmed Square”** sebagai salah satu syarat memperoleh gelar sarjana pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

Penulis tidak akan sampai pada titik ini, jikalau tanpa dukungan dan bantuan dari pihak yang selalu ada, peduli dan menyayangi penulis. Oleh karena itu, penulis haturkan rasa terima kasih yang setulus-tulusnya serta penghargaan yang setinggi-tingginya untuk orang tua penulis, Ayahanda tercinta **Irwan** dan Ibunda tercinta **Fatmawati Tandi Bali** yang telah membesarkan dan mendidik dengan penuh kesabaran, memberikan cinta dan limpahan kasih sayang, dukungan penuh, pengorbanan, kesabaran hati serta dengan ikhlas telah mengiringi setiap langkah penulis dengan doa dan restunya tanpa henti. Adik-adik tersayang penulis **Atikah Dhiya Ramadhani, Nurfadila Tandean, Ilhamsyah, Rezky Ayudya, Nur Fadia** dan **Nur Fazia**, terima kasih telah menjadi saudara yang sangat baik, selalu ada dan selalu mendoakan serta memberikan dukungan dan semangat kepada penulis. Tak lupa pula terima kasih untuk **Keluarga Besar** atas doa, dukungan, semangat, dan bantuannya selama ini.

Penghargaan yang tulus dan ucapan terima kasih dengan penuh keikhlasan juga penulis ucapkan kepada:

1. **Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.**, 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. Anna Islamiyati, S.Si., M.Si.**, selaku Ketua Departemen Statistika, segenap Dosen Pengajar dan Staf yang telah membekali ilmu dan kemudahan kepada penulis dalam berbagai hal selama menjadi mahasiswa di Departemen Statistika.
4. **Bapak Drs. Raupong, M.Si.**, selaku Pembimbing Utama sekaligus Penasehat Akademik penulis yang dengan penuh kesabaran telah meluangkan waktu dan pemikirannya untuk senantiasa memberikan arahan, dorongan semangat, dan motivasi kepada penulis dari awal perkuliahan, selama penyusunan skripsi hingga selesainya penulisan tugas akhir ini.
5. **Bapak Andi Kresna Jaya, S.Si., M.Si.**, selaku Pembimbing Pertama yang dengan penuh kesabaran telah meluangkan waktu dan pemikirannya di tengah berbagai kesibukan dan prioritasnya untuk senantiasa memberikan arahan, dorongan, dan motivasi kepada penulis mulai dari awal hingga selesainya penulisan tugas akhir ini.
6. **Ibu Dra. Nasrah Sirajang, M.Si.** dan **Bapak Siswanto, S.Si., M.Si.**, selaku Tim Penguji yang telah memberikan saran dan kritikan yang membangun dalam penyempurnaan penyusunan tugas akhir ini serta waktu yang telah diberikan kepada penulis.
7. **Kak Prasalda, Kak Erawati, Kak Titin Hardiyanti, Kak Nurhaeni, Nurwana, Nur Aliya Eva dan Sheldy Arung Lebang** serta **teman-teman Pondok Al-Khaer** yang penulis tidak dapat penulis sebutkan satu-persatu, terima kasih atas segala bantuan baik secara langsung maupun tidak langsung, kebersamaan, suka dukanya, doa, dukungan, dan motivasi yang telah diberikan kepada penulis sejak mahasiswa baru hingga saat ini.
8. Teman-teman seperjuangan di Statistika, terkhusus **Victor Liman, Kevin Ekarinaldo Palungan, Stefani Deborah Tioris, Nehemia Milenium Payung, Intan Permata Sari, Taufiq Akbar, Sri Indriani Amil, Andi Sri Yulianti, Idayah L, Musdalifah** dan seluruh **Statistika 2018** yang tidak dapat penulis sebutkan satu-persatu, terima kasih atas segala bantuan, kerja sama,



kebersamaan, kenangan suka duka serta dukungan kepada penulis selama menjalani pendidikan di Departemen Statistika hingga saat ini

9. Seluruh teman-teman **KKNT Gelombang 106 Posko Tana Toraja 2**, terima kasih atas segala dukungan dan kenangan kebersamaannya selama pelaksanaan KKN.
10. **Keluarga Mahasiswa FMIPA Unhas** terkhusus keluarga **Himatika FMIPA Unhas** dan **Himastat FMIPA Unhas**, terima kasih atas ilmu, pengalaman dan kebersamaan selama penulis kuliah di Universitas Hasanuddin.
11. Kepada semua pihak yang tidak dapat penulis sebutkan satu-persatu, semoga segala dukungan dan partisipasi yang diberikan kepada penulis bernilai ibadah disisi Allah *Subhanallahu Wa Ta'ala*.

Penulis menyadari bahwa masih banyak kekurangan dalam skripsi ini, untuk itu dengan segala kerendahan hati penulis memohon maaf. Akhir kata, semoga tulisan ini memberikan manfaat untuk pembaca.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Makassar, 1 Februari 2024



Najmawati Ade



**PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK
KEPENTINGAN AKADEMIS**

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

Nama : Najmawati Ade
NIM : H051181323
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 Bebas Royalti Non-eksklusif (*Non-exclusive Royalty- Free Right*)** atas tugas akhir saya yang berjudul:

**“Penerapan Regresi *Robust* pada Faktor-Faktor yang Memengaruhi
Kemiskinan di Provinsi Sulawesi Selatan Menggunakan Penduga *Least Trimmed Square*”**

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 memublikasikan 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 1 Februari 2024.

Yang menyatakan



(Najmawati Ade)



ABSTRAK

Metode yang biasa digunakan dalam menduga parameter analisis linier berganda adalah *Ordinary Least Square* (OLS) atau biasa juga disebut Metode Kuadrat Terkecil (MKT). Namun apabila terdapat *outlier* dalam data maka akan mempengaruhi hasil dari metode OLS menjadi kurang efisien. Cara untuk mengatasi masalah *outlier* dalam data adalah dengan menggunakan metode regresi *robust*. Salah satu metode pendugaan parameter pada regresi *robust* adalah penduga *Least Trimmed Square* (LTS). Data yang digunakan dalam penelitian ini adalah data kemiskinan di provinsi Sulawesi Selatan tahun 2021. Tujuan dari penelitian ini adalah memodelkan data kemiskinan serta mendapatkan faktor-faktor yang signifikan memengaruhi persentase penduduk miskin di provinsi Sulawesi Selatan tahun 2021. Penerapan regresi *robust* penduga *Least Trimmed Square* menghasilkan model dengan nilai R^2 sebesar 94.6%. Variabel indeks pembangunan manusia, persentase penduduk miskin yang belum tamat sekolah dasar, angka melek huruf penduduk miskin dan persentase penduduk miskin yang bekerja di sektor informal merupakan faktor yang signifikan memengaruhi persentase penduduk miskin di provinsi Sulawesi Selatan tahun 2021. Pemilihan model terbaik menggunakan kriteria *Cp Mallow*. Model terbaik yang diperoleh adalah model dengan nilai *Cp Mallow* sebesar 5.0 yang menunjukkan kombinasi dari 4 variabel yang signifikan memengaruhi persentase penduduk miskin di provinsi Sulawesi Selatan tahun 2021.

Kata Kunci: kemiskinan, OLS, *outlier*, penduga *Least Trimmed Square*, regresi *robust*.



ABSTRACT

The method commonly used in estimating multiple linear analysis parameters is *Ordinary Least Square* (OLS) or also called the *Least Squares Method* (MKT). However, if there are *outliers* in the data, it will affect the results of the OLS method to be less efficient. A way to overcome the problem of *outliers* in the data is to use *robust* regression methods. One of the parameter estimation methods in *robust* regression is the *Least Trimmed Square* (LTS) estimator. The data used in this study is poverty data in South Sulawesi province in 2021. The purpose of this study is to model poverty data and obtain factors that significantly affect the percentage of poor people in South Sulawesi province in 2021. The application of *robust* regression of the *Least Trimmed Square* estimator produces a model with an R^2 value of 94.6%. The human development index variable, the percentage of poor people who have not graduated from primary school, the literacy rate of the poor and the percentage of poor people who work in the informal sector are factors that significantly affect the percentage of poor people in South Sulawesi province in 2021. The best model selection uses the *Cp Mallow* criterion. The best model obtained is the model with a *Cp Mallow* value of 5.0 which shows a combination of 4 variables that significantly affect the percentage of poor people in South Sulawesi province in 2021.

Keywords: *poverty, OLS, outliers, Least Trimmed Square estimator, robust regression.*



DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL	ii
LEMBAR PERNYATAAN KEOTENTIKAN	iii
HALAMAN PERSETUJUAN PEMBIMBING	iv
HALAMAN PENGESAHAN	v
KATA PENGANTAR	vi
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI	ix
ABSTRAK	x
ABSTRACT	xi
DAFTAR ISI	xii
DAFTAR TABEL	xiv
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
BAB II TINJAUAN PUSTAKA	4
2.1 Analisis Regresi	4
2.2 Uji Asumsi	5
2.3 Pencilan.....	6
2.4 Metode Kuadrat Terkecil	7
2.5 Regresi <i>Robust</i>	7
2.6 Penduga <i>Least Trimmed Square</i>	8
2.7 Koefisien Determinasi	9
2.8 Pengujian Signifikansi Parameter	10
2.9 Kemiskinan	11
2.9.1 Pendidikan	12
2.9.2 Sektor Informal.....	13



2.9.3	Pengeluaran Perkapita untuk Makanan	14
2.9.4	Rumah Tangga yang Menggunakan Air Layak.....	14
2.9.5	Rumah Tangga yang Menggunakan Jamban Sendiri/Bersama	15
2.9.6	Rumah Tangga Miskin Penerima Rastra atau BPNT	15
2.10	Pemilihan Model Terbaik	15
BAB III METODOLOGI PENELITIAN		16
3.1	Sumber Data.....	16
3.2	Variabel Penelitian.....	16
3.3	Tahapan Analisis Data	16
BAB IV HASIL DAN PEMBAHASAN.....		18
4.1	Analisis Deskriptif	18
4.2	Uji Normalitas.....	19
4.3	Uji Asumsi Multikolinearitas.....	19
4.4	Uji Heteroskedastisitas.....	20
4.5	Pemodelan Regresi Linier Berganda dengan Metode <i>Ordinary Least Square</i>	21
4.6	Pendeteksian <i>outlier</i> dengan Metode <i>DfFITS</i>	22
4.7	Penerapan Metode Regresi <i>Robust</i> dengan Penduga <i>Least Trimmed Square</i>	23
4.8	Koefisien Determinasi (R^2)	25
4.9	Uji Signifikansi Parameter	25
4.10	Pemilihan Model Terbaik	27
BAB V KESIMPULAN DAN SARAN		28
5.1	Kesimpulan	28
5.2	Saran	28
DAFTAR PUSTAKA		29
LAMPIRAN.....		33



DAFTAR TABEL

Tabel 3.1. Variabel Penelitian	16
Tabel 4.1. Analisis Deskriptif.....	18
Tabel 4.2. Uji Normalitas	19
Tabel 4.3. Nilai VIF Variabel Prediktor.....	20
Tabel 4.4. Hasil Uji Heteroskedastisitas.....	20
Tabel 4.5. Hasil Identifikasi <i>Outlier</i>	23
Tabel 4.6. Nilai Estimasi Regresi <i>Robust LTS</i>	24
Tabel 4.7. Nilai R^2 dan $R^2_{adjusted}$	25
Tabel 4.8. Hasil Pengujian Simultan	25
Tabel 4.9. Nilai $ t_{hitung} $ Model Regresi <i>Robust LTS</i>	26



DAFTAR GAMBAR

Gambar 4.1. *Scatter* Plot Nilai Galat Data 22



DAFTAR LAMPIRAN

Lampiran 1. Data Persentase Kemiskinan di Provinsi Sulawesi Selatan Tahun 2021 34

Lampiran 2. Hasil Uji DfFFITS Data Persentase Kemiskinan di Provinsi Sulawesi Selatan Tahun 2021..... 36

Lampiran 3. Uji Normalitas, Multikolinearitas dan Heteroskedastisitas 37

Lampiran 4. Regresi *Robust* Penduga LTS 38

Lampiran 5. Pemilihan Model Terbaik Berdasarkan Nilai *Cp Mallow* 39



BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemiskinan merupakan salah satu permasalahan umum yang terjadi di beberapa negara di dunia, terutama di negara berkembang. Negara dalam kategori ini memiliki tugas untuk menumpas kemiskinan dan memakmurkan masyarakatnya (Pratama, 2014). Provinsi Sulawesi Selatan adalah salah satu daerah di Indonesia yang masih menghadapi permasalahan kemiskinan. Meski menjadi salah satu provinsi yang mempunyai tingkat pertumbuhan ekonomi yang cukup baik, angka kemiskinan di Provinsi Sulawesi Selatan masih terbilang cukup tinggi. Berdasarkan hasil survei sosial ekonomi nasional Maret 2021, jumlah penduduk miskin di Sulawesi Selatan pada bulan Maret 2021 sekitar 8,78 persen dari total penduduk. Jumlah penduduk miskin di Provinsi Sulawesi Selatan terus mengalami fluktuasi setiap tahunnya (BPS, 2021). Tingginya jumlah kemiskinan di Provinsi Sulawesi Selatan tidak terlepas dari berbagai faktor yang mempengaruhi. Salah satu metode yang dapat digunakan untuk menganalisis faktor-faktor yang mempengaruhi kemiskinan adalah menggunakan analisis regresi.

Analisis regresi merupakan suatu metode yang digunakan untuk menggambarkan hubungan antara variabel prediktor dengan variabel respon. Metode umum yang sering digunakan dalam mengestimasi koefisien regresi adalah Metode *Ordinary Least Square* (OLS) atau biasa juga disebut Metode Kuadrat Terkecil (Semar dkk., 2020). Model terbaik yang dihasilkan pada metode OLS harus memenuhi beberapa asumsi klasik seperti tidak terdapat multikolinearitas dan heteroskedastisitas. Selain itu, adanya *outlier* (pencilan) pada data akan mempengaruhi hasil dari metode OLS menjadi kurang efisien (Setyowati dkk., 2021). Menurut Soemartini (2007) *outlier* merupakan pengamatan yang jauh dari kelompok data yang mungkin berpengaruh besar terhadap koefisien regresi. Faktor yang dapat menyebabkan adanya *outlier* adalah kekeliruan pada sistem, kesalahan input data ataupun karena terjadinya peristiwa luar biasa seperti bencana atau krisis. *Outlier* merupakan pengamatan yang tidak mengikuti pola sebagian besar data (Chatzinakos & Zioutas). Cara untuk mengatasi



adanya data outlier adalah dengan menggunakan metode yang tepat. Salah satu metode yang bisa digunakan yaitu metode regresi *robust*.

Regresi *robust* adalah metode regresi yang digunakan ketika terdapat masalah *outlier* pada data. Metode ini merupakan alat penting untuk menganalisis data yang dipengaruhi *outlier* sehingga dihasilkan model *robust* terhadap *outlier* (Ramadhani, 2023). Regresi Robust terdiri dari 5 metode yaitu estimasi M (*M Estimation*), estimasi LMS (*Least Median of Square*), estimasi LTS (*Least Trimmed Square*), estimasi S (*Scale Estimation*) dan estimasi MM (*Method of Moment*). Regresi *robust* yang akan digunakan adalah regresi *robust* dengan estimasi *Least Trimmed Square*. Menurut Rousseeuw dan Driessen (2006), metode LTS lebih efisien dan memiliki fungsi objektif yang lebih *smooth*. Metode LTS juga merupakan salah satu metode regresi *robust* dengan nilai *breakdown point* yang tinggi dibandingkan metode lainnya dalam mengatasi masalah *outlier* (Perihatini, 2018). Regresi *robust* estimasi LTS memiliki standar galat terkecil dan nilai R-Square yang tinggi (Akolo & Nadjamuddin, 2022).

Penelitian sebelumnya, telah dilakukan oleh Shodiqin dkk (2018) mendapatkan hasil bahwa penduga *Least Trimmed Square* lebih baik daripada penduga MM yang didasarkan pada nilai koefisien determinasi. Suci Wulandari (2020), tentang perbandingan model Regresi *Robust* dengan estimasi *Least Trimmed Square*, *Maximum Likelihood Type*, dan *Scale* pada data *outlier* pada faktor-faktor yang mempengaruhi TPT di provinsi Jawa Timur tahun 2018. Hasil dari penelitian ini menunjukkan bahwa metode estimasi *robust* yang paling baik digunakan adalah estimasi LTS karena estimasi ini memiliki nilai koefisien determinasi terbesar dan MSE terkecil dibandingkan metode estimasi lainnya. Berdasarkan uraian dan penelitian yang telah dilakukan sebelumnya, maka penulis tertarik untuk mengangkat judul tentang Penerapan Regresi *Robust* Pada Faktor-Faktor yang Memengaruhi Kemiskinan di Provinsi Sulawesi Selatan dengan Menggunakan Penduga *Least Trimmed Square*.

1.2 Rumusan Masalah

Adapun rumusan masalah yang dapat diambil berdasarkan latar belakang,



1. Bagaimana pemodelan regresi *Robust* menggunakan metode *Least Trimmed Square (LTS)* pada data kemiskinan di provinsi Sulawesi Selatan tahun 2021?
2. Faktor-faktor apa saja yang signifikan memengaruhi persentase penduduk miskin di provinsi Sulawesi Selatan tahun 2021?

1.3 Batasan Masalah

Penelitian ini dibatasi pada 9 variabel prediktor yang diduga memengaruhi kemiskinan di provinsi Sulawesi Selatan tahun 2021 dan menggunakan metode *Difference in Fit Standardized (DFFITS)* untuk mendeteksi *outliers* dalam data.

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini, yaitu:

1. Mendapatkan model regresi *Robust* menggunakan metode *Least Trimmed Square* pada data kemiskinan di provinsi Sulawesi Selatan tahun 2021.
2. Memperoleh faktor-faktor yang signifikan memengaruhi persentase penduduk miskin di provinsi Sulawesi Selatan tahun 2021.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan pengetahuan dan pemahaman kepada para pembaca terutama bagi penulis mengenai penerapan metode Regresi *Robust* dalam mengatasi masalah *outlier* pada faktor-faktor yang memengaruhi kemiskinan di provinsi Sulawesi Selatan tahun 2021 dengan menggunakan penduga *Least Trimmed Square (LTS)*.



BAB II

TINJAUAN PUSTAKA

2.1 Analisis Regresi

Analisis regresi merupakan analisis statistik yang bertujuan untuk memodelkan hubungan antara variabel respon dengan variabel prediktor (Purhadi & Yasin, 2012). Analisis regresi juga dapat diartikan sebagai salah satu metode statistika yang digunakan untuk memodelkan satu atau lebih variabel. Jenis variabel yang digunakan dalam analisis regresi terdiri dari:

- a. Variabel terikat yaitu variabel yang dipengaruhi oleh variabel lainnya yang biasa juga disebut sebagai variabel *dependent*.
- b. Variabel bebas yaitu variabel yang tidak dipengaruhi oleh variabel lainnya yang biasa juga disebut variabel *independent*.

Secara umum, analisis regresi terbagi menjadi dua yaitu analisis regresi linier dan analisis regresi nonlinier. Analisis regresi linier sederhana merupakan analisis regresi yang digunakan untuk mengetahui hubungan antara satu variabel prediktor dengan variabel respon (Irwan dkk., 2012). Persamaan model regresi linier sederhana adalah sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X + \varepsilon_i; \quad i = 1, 2, \dots, n \quad (2.1)$$

Selain model analisis regresi linier sederhana terdapat pula analisis regresi linier berganda yang mana analisis ini digunakan untuk mengetahui hubungan antara beberapa variabel prediktor terhadap variabel respon (Roozbeh dkk., 2022). Persamaan umum regresi linier berganda dapat dituliskan sebagai berikut

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i; \quad i = 1, 2, \dots, n; n \geq k + 1$$

atau

$$Y_i = \beta_0 + \sum_{j=1}^k \beta_j X_{ij} + \varepsilon_i \quad (2.2)$$

dengan:

Y_i : nilai variabel respon pada pengamatan ke-i

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

β_k : parameter regresi

ε_i : galat (error)



2.2 Uji Asumsi

1. Uji Normalitas *Shapiro Wilks*

Uji normalitas digunakan untuk mengetahui data berasal dari populasi yang memiliki sebaran atau distribusi normal. Salah satu metode yang digunakan untuk uji normalitas yaitu uji *Shapiro Wilk*. Uji normalitas digunakan untuk mengetahui data berasal dari populasi yang memiliki sebaran atau distribusi normal.

Hipotesis uji normalitas *Shapiro Wilk* yaitu:

H_0 : galat berdistribusi normal

H_1 : galat tidak berdistribusi normal

Statistik uji *Shapiro-Wilk* dirumuskan sebagai berikut:

$$W = \frac{1}{S^2} \left[\sum_{i=1}^n a_i (e_{n-i+1} - e_i) \right]^2$$

Keterangan:

S^2 = Varians galat

a_i = Koefisien *Shapiro-Wilk test*

e_{n-i+1} = galat ke $n - i + 1$

e_i = galat ke- i

\bar{e} = rata-rata galat

Kriteria pengambilan keputusan yaitu jika nilai W hitung lebih besar dari nilai kritis pada tabel $W_{(n,\alpha)}$ atau $W > W_{(n,\alpha)}$ maka asumsi normalitas terpenuhi (Syamsudin dan Wachidah, 2020).

2. Uji Multikolinearitas

Uji multikolinearitas adalah uji yang digunakan untuk mengetahui adanya hubungan antara variabel-variabel prediktor dalam model regresi. Multikolinearitas dapat diketahui melalui nilai *VIF* (*Variance Inflation Factor*). Multikolinearitas terjadi apabila nilai $VIF > 10$ pada masing-masing variabel prediktor (Yaziz dkk., 2019). Nilai *VIF* didapatkan melalui perhitungan sebagai berikut:

$$VIF = \frac{1}{1 - R_j^2} \quad (2.3)$$

h koefisien determinasi antara variabel prediktor X_{ij} dengan variabel lainnya.



3. Uji Heteroskedastisitas

Heteroskedastisitas berarti bahwa varians pada galat tidak sama atau tidak identik. Model regresi yang baik adalah jika tidak terjadi heteroskedastisitas atau terjadi homokedastisitas. Uji heteroskedastisitas dapat dilakukan dengan menggunakan metode *Glejser* (Nurahman dkk., 2016). Hipotesis uji heteroskedastisitas metode *Glejser* yaitu:

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 \text{ (tidak terjadi heteroskedastisitas)}$$

$$H_1: \text{ada } \sigma_i^2 \neq \sigma_{i^*}^2; i = 1, 2, \dots, n \text{ (terjadi heteroskedastisitas)}$$

Statistik uji yang digunakan adalah sebagai berikut:

$$F_{hitung} = \frac{\frac{\sum_{i=1}^n (|\varepsilon_i| - |\bar{\varepsilon}|^2)}{p - 1}}{\frac{\sum_{i=1}^n (|\varepsilon_i| - |\hat{\varepsilon}_i|^2)}{n - p}}$$

Kriteria pengambilan keputusan yaitu H_0 ditolak jika $F_{hitung} > F_{\alpha, (p-1), (n-p)}$ atau $p - value < \alpha$ berarti terjadi heteroskedastisitas (Alwi dkk., 2021).

2.3 Pencilan

Pencilan (*outlier*) merupakan pengamatan yang jauh dari kisaran nilai yang diharapkan dalam suatu penelitian atau dapat artikan sebagai titik yang tidak mengikuti pola linier utama data (Roozbeh dkk., 2018). Data yang mengandung *outlier* jika dianalisis menggunakan analisis regresi maka hasilnya kurang tepat. Salah satu cara mendeteksi adanya *outlier* yaitu dengan menggunakan *Difference in Fit Standardized* (DFFITS) dengan persamaan sebagai berikut (Kurniawan dkk., 2023):

$$DFFITS_i = t_i \sqrt{\frac{h_{ii}}{1 - h_{ii}}} \tag{2.4}$$

dengan:

t_i : *studentized deleted* galat untuk kasus ke-i

h_{ii} : nilai *leverage* untuk kasus ke-i

Data dinyatakan *outlier* apabila nilai:

$$|DFFITS| > 2 \sqrt{\frac{p}{n}}$$



dengan p adalah banyaknya parameter dan n banyaknya data pengamatan (Rahmiatun dkk., 2022).

2.4 Metode Kuadrat Terkecil

Metode *Ordinary Least Square* (OLS) adalah metode yang biasa digunakan dalam analisis regresi (Subzar dkk., 2020). Prinsip umum dari metode OLS (metode kuadrat terkecil) adalah meminimumkan jumlah kuadrat galat. Persamaan (2.2) dapat dituliskan dalam bentuk matriks seperti pada persamaan sebagai berikut:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.5)$$

dengan:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_3 \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \text{ dan } \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Keterangan:

\mathbf{y} = vektor variabel respon ($n \times 1$)

\mathbf{X} = matriks dari variabel prediktor ($n \times (p + 1)$)

$\hat{\boldsymbol{\beta}}$ = vektor parameter model regresi ($(p + 1) \times 1$)

$\boldsymbol{\varepsilon}$ = vektor error ($n \times 1$) (Permai & Tanty, 2018).

Menentukan penduga \mathbf{y} yaitu $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ diperlukan nilai penduga untuk parameter $\boldsymbol{\beta}$ yaitu $\hat{\boldsymbol{\beta}}$ dengan parameter $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$ tidak diketahui dan perlu ditentukan nilai dugaannya dengan meminimumkan jumlah kuadrat galat. Nilai pendugaan parameter model regresi dapat dinyatakan sebagai berikut:

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

2.5 Regresi Robust

Regresi *robust* adalah analisis regresi yang digunakan untuk mendeteksi pencilan dan memberikan hasil yang resisten terhadap adanya pencilan (Chen, 2002). Metode ini diperkenalkan oleh Andrews (1972). Metode regresi *robust* merupakan suatu metode yang digunakan untuk mengatasi masalah pencilan

(Law dan Leroy, 1987). Regresi *robust* adalah metode yang digunakan untuk analisis data yang mengandung *outlier*. Penerapan metode ini dalam analisis data yang mengandung pencilan dapat memberikan hasil model yang



robust atau *resistance* terhadap *outlier* (Shodiqin dkk., 2018). Regresi *robust* merupakan metode regresi yang digunakan apabila terdapat *outlier* yang mempengaruhi model regresi (Begashaw & Yohannes, 2020).

2.6 Penduga *Least Trimmed Square*

Least Trimmed Square (LTS) diperkenalkan pertama kali oleh Rousseeuw pada tahun 1984. Penduga ini merupakan metode regresi robust yang memiliki nilai *breakdown point* tinggi (Chen, 2002). Menurut Rousseeuw dan Hubert prinsip dari penduga LTS ini sama dengan metode OLS dalam mengestimasi parameter dengan meminimumkan jumlah kuadrat galat. Penduga *Least Trimmed Square* (LTS) didefinisikan oleh Rousseeuw dan Leroy (1987) sebagai berikut:

$$\hat{\beta}_{LTS} = \min \sum_{i=1}^h \varepsilon_i^2 \quad (2.7)$$

fungsi tujuan dari metode LTS ini di tuliskan sebagai berikut:

$$\min \sum_{i=1}^h \varepsilon_i^2 = \min \sum_{i=1}^h (Y_{(i)} - \hat{Y}_{(i)})^2 \quad (2.8)$$

dengan meminimumkan jumlah kuadrat galat sebanyak h , maka:

$$h = \frac{n}{2} + \frac{k+1}{2} = \frac{n+k+1}{2}, h \leq n$$

dengan:

ε_i^2 : kuadrat galat yang telah diurutkan dari terkecil ke terbesar $\varepsilon_1^2 \leq \varepsilon_2^2 \leq \dots \leq \varepsilon_n^2$;

k : banyaknya variabel prediktor

n : banyaknya pengamatan

Jumlah h menunjukkan sejumlah *subset* data dengan kuadrat galat terkecil (Willems & Aelst, 2005). Sama halnya dengan penduga lain pada regresi *robust*, pada metode LTS diberikan pembobot (w_{ii}) pada data sehingga data *outlier* tidak mempengaruhi model parameter hasil estimasi. Pembobot (w_{ii}) dapat dituliskan dalam bentuk matriks, yang dinotasikan sebagai \mathbf{W} ,

$$\mathbf{W} = \begin{pmatrix} w_{11} & 0 & \dots & 0 \\ 0 & w_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_{nn} \end{pmatrix}$$



Pembentukan estimasi parameter dilakukan hingga *Final Weighted Least Square* dengan fungsi pembobot seperti pada persamaan (2.9).

$$w_{ii} = \begin{cases} 1, & \frac{|e_i|}{S_{LTS}} \leq 3 \\ 0, & \text{lainnya} \end{cases} \quad (2.9)$$

dengan,

$$S_{LTS} = d_{h,n} \sqrt{\frac{1}{h} \sum_{i=1}^h e_{(i)}^2}$$

$$d_{h,n} = \frac{1}{\sqrt{1 - \frac{2n}{hc_{h,n}} \Phi\left(\frac{1}{c_{h,n}}\right)}}$$

$$c_{h,n} = \frac{1}{\Phi^{-1}\left(\frac{h+n}{2n}\right)}$$

dengan Φ merupakan fungsi kepadatan peluang dari distribusi normal standar, Φ merupakan fungsi distribusi kumulatif normal standar dan p merupakan banyaknya parameter pada model (Savitri, 2015). Penduga parameter regresi *robust LTS* adalah sebagai berikut (Maharani dkk., 2014):

$$\beta_{LTS} = (X'WX)^{-1}X'Wy \quad (2.10)$$

2.7 Koefisien Determinasi

Koefisien determinasi (R^2) digunakan untuk mengukur seberapa besar pengaruh variabel prediktor terhadap variabel respon. Semakin tinggi nilai R^2 maka semakin besar proporsi dari variabel respon yang dapat dijelaskan oleh variabel prediktor. Model yang paling baik adalah model yang mempunyai nilai koefisien determinasi yang besar. Nilai R^2 berkisar antara $0 \leq R^2 \leq 1$ (Ndruru dkk., 2014).

Rumus yang digunakan untuk menghitung nilai R^2 adalah sebagai berikut:

$$R^2 = 1 - \frac{JKG}{JKT} \quad (2.11)$$



an:
Jumlah Kuadrat Galat
Jumlah Kuadrat Total

2.8 Pengujian Signifikansi Parameter

Ada dua tahap pengujian parameter yaitu pengujian secara serentak dan pengujian secara parsial satu-satu. Pengujian secara serentak menggunakan uji F dan pengujian secara parsial satu-satu menggunakan uji t.

a) Uji F

Tujuan digunakannya uji serentak adalah untuk mengetahui apakah parameter model regresi yang digunakan sudah signifikan atau belum. Parameter yang digunakan dalam pengujian serentak ini adalah parameter yang terdapat dalam model. Hipotesis untuk uji serentak yaitu:

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

$$H_1: \text{Minimal ada } \beta_j \neq 0 \text{ untuk } j = 1, 2, \dots, k$$

$$F_{hitung} = \frac{KTR}{KTG} \quad (2.12)$$

Kriteria pengambilan keputusannya yaitu:

Tolak H_0 jika $F_{hitung} > F_{tabel}$ atau $p\text{-value} < \alpha(0.05)$ (Atamia dkk., 2021).

b) Uji t

Uji parsial satu-satu atau biasa juga disebut uji individu digunakan untuk menguji pengaruh masing-masing variabel prediktor secara sendiri-sendiri terhadap variabel respon. Uji ini dapat dilakukan dengan membandingkan t hitung dengan t tabel atau dengan melihat kolom signifikan pada masing-masing t hitung. Hipotesis untuk uji parsial satu-satu yaitu:

$$H_0: \beta_j = 0; \quad j = 1, 2, \dots, k$$

$$H_1: \beta_j \neq 0$$

Statistik uji yang digunakan:

$$t_{hitung} = \frac{\hat{\beta}_j}{Se(\hat{\beta}_j)} \quad (2.13)$$

$$Se(\hat{\beta}_j) = \sqrt{Var(\hat{\beta}_j)} \quad (2.14)$$

$Se(\hat{\beta}_j)$ menunjukkan nilai standar galat dari β_j

pengambilan keputusannya yaitu:

Tolak H_0 jika $t_{hitung} > t_{tabel}$ atau $p\text{-value} < \alpha(0.05)$ (Atamia dkk., 2021).



2.9 Kemiskinan

Kemiskinan dapat diartikan sebagai ketidakmampuan memenuhi standar minimum kebutuhan dasar berupa kebutuhan makanan maupun non makanan (Syairozi, 2020). Badan Pusat Statistik (BPS) menggunakan konsep kemampuan memenuhi kebutuhan dasar (*basic needs approach*) dalam mengukur kemiskinan. Dengan pendekatan ini, kemiskinan dipandang sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran. Presentase penduduk miskin dapat diperoleh dengan menggunakan metode *Head Count Index (HCI-PO)* yang menunjukkan presentase penduduk yang berada dibawah Garis Kemiskinan (GK). Metode ini secara sederhana mengukur proporsi yang dikategorikan miskin yang dinyatakan sebagai berikut:

$$P_{\alpha} = \frac{1}{n} \sum_{i=1}^q \left[\frac{Z - y_i}{Z} \right]^{\alpha}$$

dengan

P_{α} : persentase penduduk miskin

α : 0, 1, 2

Z : garis kemiskinan

y_i : rata-rata pengeluaran per kapita per bulan yang berada di bawah garis kemiskinan

q : banyaknya penduduk yang berada di bawah garis kemiskinan

n : jumlah penduduk

Penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran perkapita per bulan dibawah garis kemiskinan. Garis Kemiskinan (GK) yaitu penjumlahan dari garis kemiskinan makanan dan garis kemiskinan non makanan. Garis Kemiskinan Makanan (GKM) merupakan jumlah nilai pengeluaran dari 52 komoditi dasar makanan yang riil dikonsumsi penduduk yang setara dengan 2100 kilokalori perkapita perhari. Garis Kemiskinan Non Makanan (GKNM) adalah penjumlahan nilai kebutuhan minimum dari komoditi-komoditi non makanan

yang meliputi perumahan, sandang, pendidikan dan kesehatan (BPS, 2020).

Faktor-faktor yang berpengaruh terhadap presentase kemiskinan antara lain:

Perkembangan manusia, angka melek huruf, tingkat pengangguran terbuka,



pengeluaran perkapita penduduk miskin untuk makanan, pertumbuhan ekonomi dan rata-rata lama sekolah. Hasil penelitian Ardiansyah (2019) menyatakan faktor-faktor yang signifikan terhadap kemiskinan meliputi tingkat pengangguran terbuka, pertumbuhan ekonomi dan rata-rata lama sekolah. angka melek huruf berpengaruh terhadap presentase kemiskinan (Astuti dkk., 2016). Hasil penelitian Padambo dkk (2021) menyatakan bahwa Indeks Pembangunan Manusia mempunyai pengaruh yang signifikan terhadap kemiskinan.

2.9.1 Pendidikan

Tahap dalam mengembangkan segala aspek kepribadian manusia, yang meliputi pengetahuannya, nilai serta sikapnya dan keterampilannya. Pada dasarnya, Pendidikan formal terdiri atas tiga tahap, yaitu Pendidikan dasar, Pendidikan menengah dan Pendidikan tinggi. Pendidikan dasar adalah jenjang Pendidikan yang berbentuk Sekolah Dasar (SD) ataupun Madrasah Ibtidaiyah (MI) yang mendasari jenjang pendidikan menengah. Sehingga dapat dinyatakan bahwa pendidikan di sekolah dasar merupakan salah satu hal penting yang menjadi dasar dalam mewujudkan tujuan pendidikan nasional. Presentase penduduk miskin usia 15 tahun ke atas yang belum tamat SD di nyatakan sebagai berikut:

$$a = \frac{b}{c} \times 100\%$$

dengan,

a: Persentase penduduk miskin usia 15 tahun ke atas yang belum tamat SD

b: Jumlah penduduk miskin usia 15 tahun ke atas belum tamat SD

c: Jumlah penduduk miskin usia 15 tahun ke atas

Salah satu indikator yang digunakan untuk mengetahui perkembangan pendidikan penduduk adalah Angka Melek Huruf (AMH). Menurut Badan Pusat Statistik (2021), AMH adalah proporsi penduduk miskin yang dapat membaca dan menulis kalimat sederhana dalam aksara tertentu, yaitu huruf latin, huruf arab, ataupun huruf lainnya.

$$AMH_{15+}^t = \frac{MH_{15+}^t}{P_{15+}^t} \times 100\%$$

: Angka Melek Huruf penduduk usia 15 tahun ke atas



MH_{15+}^t : penduduk usia 15 tahun ke atas yang mempunyai kemampuan membaca dan menulis huruf latin dan huruf lainnya.

P_{15+}^t : penduduk usia 15 tahun ke atas

Ardiansyah (2019) menyatakan bahwa salah satu faktor yang signifikan terhadap kemiskinan adalah Tingkat Pengangguran Terbuka (TPT). Menurut Badan Pusat Statistik, TPT adalah persentase sejumlah pengangguran terhadap jumlah angkatan kerja. Angkatan kerja adalah penduduk usia kerja (15 tahun ke atas) yang bekerja atau punya pekerjaan namun sementara tidak bekerja dan pengangguran.

$$TPT = \frac{PP}{PAK} \times 100\%$$

dengan,

TPT : Tingkat Pengangguran Terbuka

PP : jumlah pengangguran

PAK : jumlah Angkatan kerja

Indeks Pembangunan Manusia (IPM) merupakan indikator yang menjelaskan penduduk dapat mengakses hasil pembangunan dalam memperoleh pendapatan, kesehatan, pendidikan dan sebagainya.

$$IPM = \sqrt[3]{I_{kesehatan} \times I_{pendidikan} \times I_{pengeluaran}} \times 100\%$$

dengan,

IPM : Indeks Pembangunan Manusia

$I_{kesehatan}$: indeks untuk dimensi kesehatan

$I_{pendidikan}$: indeks untuk dimensi pendidikan

$I_{pengeluaran}$: indeks untuk dimensi pengeluaran

2.9.2 Sektor Informal

BPS mendefinisikan bahwa pekerja sektor informal merupakan suatu pekerjaan utama seseorang yang mencakup berusaha sendiri, berusaha dibantu buruh tidak tetap, buruh/karyawan, pekerja bebas di pertanian, pekerja bebas di nonpertanian dan pekerja keluarga/tidak dibayar. Persentase penduduk miskin usia

ke atas yang bekerja di sektor informal dinyatakan sebagai berikut:

$$d = \frac{e}{c} \times 100\%$$



dengan,

d: Presentase penduduk miskin usia 15 tahun ke atas yang bekerja di sektor informal

e: Jumlah penduduk miskin usia 15 tahun ke atas yang berkerja di sektor informal

2.9.3 Pengeluaran Perkapita untuk Makanan

Definisi pengeluaran per kapita menurut BPS adalah pengeluaran rumah tangga dibagi banyaknya anggota rumah tangga dan untuk persentase pengeluaran per kapita untuk makanan didefinisikan sebagai pengeluaran per kapita untuk makanan dibagi dengan banyaknya anggota rumah tangga.

$$f = \frac{g}{h} \times 100\%$$

dengan,

f: Persentase pengeluaran per kapita untuk makanan

g: Pengeluaran per kapita untuk makanan

h: Banyaknya anggota rumah tangga

2.9.4 Rumah Tangga yang Menggunakan Air Layak

BPS mendefinisikan bahwa rumah tangga pengguna air layak merupakan rumah tangga miskin yang menggunakan sumber utama air minum terlindungi yang meliputi leding, sumur bor atau sumur pompa, sumur terlindung, mata air terlindung dan air hujan. Rumah tangga yang air minum utamanya juga menggunakan air kemasan bermerk atau air isi ulang dan sumber utama air mandi/cuci dan lain-lain yang digunakan adalah leding, sumur bor atau sumur pompa, sumur terlindung, mata air terlindung dan air hujan juga termasuk definisi dari rumah tangga pengguna air layak. Persentase rumah tangga miskin yang menggunakan air layak dinyatakan sebagai berikut:

$$i = \frac{j}{l} \times 100\%$$

dengan,

i: Persentase rumah tangga miskin yang menggunakan air layak

j: rumah tangga miskin yang menggunakan air layak

l: rumah tangga miskin



2.9.5 Rumah Tangga yang Menggunakan Jamban Sendiri/Bersama

Definsi rumah tangga pengguna jamban sendiri/bersama menurut BPS adalah rumah tangga yang menggunakan fasilitas tempat pembuangan air besar yang digunakan oleh rumah tangga sendiri atau bersama dengan rumah tangga tertentu. Persentase rumah tangga miskin yang menggunakan jamban sendiri/Bersama dinyatakan sebagai berikut:

$$m = \frac{o}{l} \times 100\%$$

dengan,

m : Persentase rumah tangga miskin yang menggunakan jamban sendiri/Bersama

o : Jumlah rumah tangga yang menggunakan jamban sendiri/Bersama

2.9.6 Rumah Tangga Miskin Penerima Rastra atau BPNT

Persentase rumah tangga miskin yang menerima bantuan beras sejahtera/BPNT dinyatakan sebagai berikut:

$$q = \frac{r}{l} \times 100\%$$

dengan,

q : Persentase rumah tangga miskin penerima rastra/BPNT

r : Jumlah rumah tangga miskin penerima rastra/BPNT

2.10 Pemilihan Model Terbaik

Model terbaik adalah model yang seluruh koefisien regresinya signifikan dan mempunyai kriteria terbaik optimum (Pramessti & Suhardiyah, 2013). Salah satu kriteria yang bisa digunakan untuk memilih model terbaik adalah menggunakan statistik *Cp Mallow*. Statistik *Cp Mallow* adalah teknik untuk pemilihan model dalam regresi yang didefinisikan sebagai kriteria yang diterima untuk menilai kecocokan saat model yang berbeda parameter sedang dibandingkan (Sartika dkk., 2020). Rumus yang digunakan untuk menghitung nilai *Cp Mallow* adalah:

$$C_p = \frac{JKG_k}{\hat{\sigma}^2} + 2(k + 1) - n \quad (2.15)$$



Jumlah Kuadrat Galat untuk variabel prediktor yang signifikan
Mean Square Error untuk semua variabel prediktor (Alcantara dkk., 2022).