

## Penerapan Metode Regresi Ridge untuk Mengatasi Masalah Multikolinearitas pada Kasus Tingkat Pengangguran Terbuka di Pulau Sumatera

Faddila Hayati<sup>1</sup>, Lilis Harianti Hasibuan<sup>✉2</sup>, Syarto Musthofa<sup>3</sup>

<sup>1,2,3</sup> Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Imam Bonjol, Padang, Indonesia

Email corresponding author : [lilisharianti@uinib.ac.id](mailto:lilisharianti@uinib.ac.id)

email : [fadillafadilla303@gmail.com](mailto:fadillafadilla303@gmail.com)<sup>1</sup>, [lilisharianti@uinib.ac.id](mailto:lilisharianti@uinib.ac.id)<sup>2</sup>, [syartom@uinib.ac.id](mailto:syartom@uinib.ac.id)<sup>3</sup>

Received 12 Feb 2026, Accepted 31 Maret 2026, Published 31 Maret 2026

### Abstrak

Tingkat pengangguran terbuka merupakan salah satu indikator penting dalam menilai kondisi ketenagakerjaan dan kesejahteraan ekonomi suatu wilayah. Dalam analisis faktor-faktor yang mempengaruhi tingkat pengangguran terbuka, penggunaan regresi linier berganda sering menghadapi masalah multikolinearitas yang menyebabkan estimasi parameter menjadi tidak stabil dan kurang efisien. Oleh karena itu, penelitian ini bertujuan untuk menerapkan metode regresi ridge dalam mengatasi masalah multikolinearitas pada kasus tingkat pengangguran terbuka di Pulau Sumatera. Metode regresi ridge merupakan pengembangan dari metode kuadrat terkecil (OLS) dengan menambahkan parameter bias melalui konstanta ridge. Dalam penelitian ini digunakan estimator Hoerl, Kennard, dan Baldwin (HKB) untuk menentukan nilai konstanta ridge yang optimal. Hasil penelitian menunjukkan bahwa model regresi ridge mampu mengatasi masalah multikolinearitas yang terjadi pada model regresi linier berganda, yang ditunjukkan dengan nilai Variance Inflation Factor (VIF) seluruh variabel bebas menjadi  $< 10$ . Dengan demikian, regresi ridge menghasilkan model estimasi yang lebih stabil dan efisien dalam menganalisis faktor-faktor yang mempengaruhi tingkat pengangguran terbuka di Pulau Sumatera.

**Kata Kunci:** Regresi Ridge, Multikolinearitas, Tingkat Pengangguran Terbuka, Metode HKB

### Abstract

The open unemployment rate is one of the important indicators in assessing the employment conditions and economic welfare of a region. In analyzing the factors that influence the open unemployment rate, the use of multiple linear regression often encounters multicollinearity problems that cause parameter estimates to become unstable and less efficient. Therefore, this study aims to apply the ridge regression method to overcome the problem of multicollinearity in the case of the open unemployment rate on the island of Sumatra. The ridge regression method is an extension of the ordinary least squares (OLS) method by adding a bias parameter through a ridge constant. In this study, the Hoerl, Kennard, and Baldwin (HKB) estimator was used to determine the optimal ridge constant value. The results show that the ridge regression model is able to overcome the multicollinearity problem that occurs in the multiple linear regression model, as indicated by the Variance Inflation Factor (VIF) value of all independent variables being  $< 10$ . Thus, ridge regression produces a more stable and efficient estimation model in analyzing the factors that influence the open unemployment rate on the island of Sumatra.

**Keywords:** Ridge Regression, Multicollinearity, Open Unemployment Rate, HKB Method

## PENDAHULUAN

Tingkat pengangguran terbuka merupakan salah satu indikator penting dalam menilai kondisi pasar tenaga kerja dan dinamika perekonomian suatu wilayah. Tingginya tingkat pengangguran terbuka tidak hanya mencerminkan kesenjangan antara penawaran dan permintaan tenaga kerja, tetapi juga berpotensi menimbulkan berbagai permasalahan sosial, seperti meningkatnya kemiskinan, ketimpangan ekonomi, dan menurunnya kesejahteraan Masyarakat [1]. Pada tahun 2017-2022 tingkat pengangguran terbuka di pulau Sumatera menunjukkan kondisi yang memprihatinkan. Secara umum, rata-rata tingkat pengangguran terbuka pada provinsi di pulau Sumatera hampir setara dengan tingkat pengangguran nasional [2]. Pada tahun 2022, tingkat pengangguran terbuka di pulau Sumatera meningkat sebesar 5,31. Hal ini terjadi akibat program dari pemerintah dalam rangka pemulihan ekonomi nasional akibat pandemi covid-19, sehingga kondisi perekonomian nasional kembali membaik secara perlahan dan tingkat pengangguran terbuka kembali menurun hingga hampir mendekati pada tingkat yang sama dengan tingkat pengangguran pada tahun-tahun sebelum terdampak pandemi covid-19 [3].

Adapun faktor-faktor yang mempengaruhi tingkat pengangguran terbuka seperti indeks Pembangunan manusia, rata-rata lama sekolah, jumlah penduduk miskin, umur harapan hidup, laju pertumbuhan produk domestik regional bruto, jumlah penduduk, pengeluaran perkapita, kepadatan penduduk, dan tingkat partisipasi angkatan kerja [4]. Dalam kasus sosial analisis regresi linear berganda digunakan untuk melihat adanya hubungan dua atau lebih variabel independen terhadap variabel dependen. Semakin banyak variabel independen yang digunakan dalam model regresi, semakin besar kemungkinan terjadinya multikolinearitas karena adanya korelasi antar variabel bebas [5],[6]. Ukuran yang dapat digunakan untuk mendeteksi adanya multikolinearitas adalah *Variance Inflation Factor* (VIF). VIF digunakan sebagai kriteria dalam mengidentifikasi multikolinearitas pada model regresi linear yang melibatkan lebih dari satu variabel bebas. Apabila nilai VIF melebihi 10, maka hal tersebut menunjukkan adanya multikolinearitas [7],[8],[9],[10].

Apabila multikolinearitas terjadi pada data yang dianalisis, maka terdapat berbagai cara yang dapat diajukan untuk mengurangi atau mengatasinya, seperti menghilangkan variabel bebas yang memiliki korelasi atau hubungan yang sempurna, menambahkan data dan memakai metode analisis lainnya seperti *Regresi Ridge* [11]. Regresi ridge merupakan metode estimasi koefisien regresi yang diperoleh melalui penambahan konstanta bias  $c$  sehingga nilai koefisien menjadi stabil [8].

Regresi ridge digunakan untuk mengatasi masalah multikolinearitas pada suatu data. Keuntungan penggunaan regresi ridge dibandingkan metode lain yaitu regresi ridge mengurangi dampak multikolinearitas dengan menentukan penduga yang bias tetapi mempunyai varians yang lebih kecil dari varians penduga regresi linear berganda [13]. Metode regresi ridge diperoleh dengan cara meminimumkan jumlah

kuadrat terkecil dan dilakukan penambahan konstanta bias agar nilai koefisien berkurang dan mendekati nol. Salah satu pendekatan analitik yang digunakan untuk memilih nilai konstanta bias ( $c$ ) secara iteratif adalah metode Hoerl, Kennard, dan Baldwin [10]. Metode iterasi Hoerl, Kennard dan Baldwin (HKB) adalah cara yang digunakan untuk menentukan nilai konstanta bias ( $c$ ) dalam regresi ridge secara terstruktur. Tahapan awal metode ini yaitu dimulai dengan nilai  $c$  yang dihitung menggunakan rumus HKB berdasarkan parameter estimasi dari metode kuadrat terkecil, kemudian dilakukan iterasi berulang hingga perubahan  $c$  menjadi sangat kecil, sehingga diperoleh  $c$  yang optimal. Metode iterasi Hoerl, Kennard, dan Baldwin adalah cara yang sistematis untuk menentukan nilai konstanta bias  $c$  dalam regresi ridge, agar menghasilkan estimasi koefisien yang stabil.

Berdasarkan latar belakang penulis tertarik melakukan penelitian dengan judul Penerapan metode regresi ridge untuk mengatasi masalah multikolinearitas pada kasus tingkat pengangguran terbuka di pulau Sumatera.

## METODOLOGI

### Jenis Penelitian dan Sumber Data

Penelitian ini menggunakan data kuantitatif yang merupakan data sekunder yang diambil dari *website* BPS pulau sumatera tentang tingkat pengangguran terbuka yang diakses pada tahun 2023. Jumlah data yang digunakan dalam penelitian ini yaitu sebanyak 154 observasi.

### Variabel Penelitian

Data penelitian ini terdiri dari sembilan variabel bebas yaitu:

| Tabel 2.1 Variabel Penelitian |                                     |                        |
|-------------------------------|-------------------------------------|------------------------|
| Variabel                      | Keiteirangan                        | Skala dan Satuan       |
| $Y$                           | Tingkat peingangguran teirbuka      | % (peirsein)           |
| $X_1$                         | Indeiks Peimbangan manusia          | Indeiks (tanpa satuan) |
| $X_2$                         | Rata-rata lama seikolah             | Tahun                  |
| $X_3$                         | Jumlah peinduduk miskin             | Ribu jiwa              |
| $X_4$                         | Umur harapan hidup                  | Tahun                  |
| $X_5$                         | Laju PDRB                           | % (peirsein)           |
| $X_6$                         | Jumlah peinduduk                    | Jiwa                   |
| $X_7$                         | Peingeiluaran peirkapita            | Ribu rupiah            |
| $X_8$                         | Keipadatan peinduduk                | Jiwa/ $km^2$           |
| $X_9$                         | Tingkat partisipasi Angkatan keirja | % (peirsein)           |

### Metode Analisis Data dan Tahapan Penelitian

Metode analisis yang digunakan adalah regresi ridge untuk mengatasi masalah multikolinearitas. Pengolahan data yang digunakan menggunakan bantuan Rstudio.

1. Menginput data variabel dependen ( $y_i$ ) dan variabel-variabel independen ( $x_i$ ),
2. Estimasi parameter dengan Metode Kudarat Terkecil ,
3. Mendeteksi multikolinearitas pada tiap variabel bebas dengan menggunakan *variance inflation factor* (VIF). Jika nilai VIF lebih dari 10, maka data mengalami multikolinearitas,
4. Transformasi data melalui pemusatan dan penskalaan,
5. Menentukan estimasi parameter model dengan menggunakan metode regresi ridge menggunakan konstanta estimator Hoerl, Kennard, dan Baldwin (HKB) sebagai berikut:

$$c_{HKB} = \frac{P\sigma^2}{\beta^*\beta^*}$$

6. Membentuk model regresi linear berganda dengan menggunakan parameter estimasi yang sudah diperoleh sebelumnya, berdasarkan konstanta estimasi Hoerl, Kennard dan Baldwin (HKB) pada metode regresi ridge sebagai berikut:

$$\beta_R = (X^T X + cI)^{-1} X^T Y$$

7. Model yang telah diperoleh kemudian diuji signifikansinya dan menginterpretasikan variabel independen yang berpengaruh terhadap variabel dependen,
8. Menginterpretasikan model yang sudah diperoleh dan menarik kesimpulan.

## HASIL DAN PEMBAHASAN

### Analisis Deskriptif Data

Dalam penelitian ini terlebih dahulu disajikan analisis deskriptif terhadap seluruh variabel penelitian. Analisis deskriptif bertujuan untuk memberikan gambaran umum mengenai karakteristik data. Pada penelitian ini terdapat 154 observasi dan data merupakan data sekunder yang diambil dari [www.pulausumatera.bps.id](http://www.pulausumatera.bps.id). Hasil ukuran pemusatan dan ukuran penyebaran data tercantum pada tabel berikut:

Tabel 3.1 Analisis Deskriptif

| Variabel       | N   | Rata-rata | Meidian | Minimum | Maksimum | Standar Deiviasi |
|----------------|-----|-----------|---------|---------|----------|------------------|
| Y              | 154 | 4,5       | 4,42    | 0       | 10,86    | 2,01             |
| X <sub>1</sub> | 154 | 73,6      | 72,94   | 64,68   | 88,32    | 4,20             |
| X <sub>2</sub> | 154 | 9,1       | 8,81    | 6,14    | 13,04    | 1,31             |

|       |     |          |          |       |         |           |
|-------|-----|----------|----------|-------|---------|-----------|
| $X_3$ | 154 | 36,8     | 27,005   | 1,47  | 187,28  | 32,58     |
| $X_4$ | 154 | 69,8     | 69,87    | 63,47 | 75,13   | 2,4       |
| $X_5$ | 154 | 3,0      | 3,125    | -0,71 | 7,32    | 1,11      |
| $X_6$ | 154 | 295760,0 | 205298,5 | 99,8  | 2474166 | 360577,79 |
| $X_7$ | 154 | 11229,9  | 10967    | 6382  | 18990   | 2068,15   |
| $X_8$ | 154 | 633,2    | 116      | 15    | 8859    | 1408,14   |
| $X_9$ | 154 | 70,1     | 69,845   | 57,81 | 86,89   | 5,94      |

Berdasarkan Tabel 3.1 dapat dilihat bahwa rata-rata Tingkat Pengangguran Terbuka ( $Y$ ) adalah 4,5 persen dan Tingkat Pengangguran Terbuka tertinggi adalah 10,86 persen sedangkan yang paling rendah adalah 0 persen. Rata-rata Indeks Pembangunan Manusia ( $X_1$ ) adalah 73,6 dan Indeks Pembangunan Manusia tertinggi 88,32 sedangkan paling rendah adalah 64,68. Rata-rata Lama Sekolah ( $X_2$ ) adalah 9,1 Tahun dan rata-rata Lama Sekolah tertinggi 13,04 Tahun sedangkan paling rendah adalah 6,14. Rata-rata Jumlah Penduduk Miskin ( $X_3$ ) adalah 36,8 ribu jiwa dan Jumlah Penduduk Miskin tertinggi 187,28 ribu jiwa sedangkan paling rendah adalah 1,47 ribu jiwa. Rata-rata Umur Harapan Hidup ( $X_4$ ) adalah 69,8 Tahun dan Umur Harapan Hidup tertinggi 75,13 Tahun sedangkan paling rendah adalah 63,47 Tahun. Rata-rata Laju PDRB ( $X_5$ ) adalah 3 persen dan Laju PDRB tertinggi 7,32 persen sedangkan paling rendah adalah -0,71 persen. Rata-rata Jumlah Penduduk ( $X_6$ ) adalah 295760 Jiwa dan Jumlah Penduduk tertinggi 2474166 jiwa sedangkan paling rendah adalah 99,8 jiwa. Rata-rata Pengeluaran perkapita ( $X_7$ ) adalah 11229,9 Ribu rupiah dan Pengeluaran perkapita tertinggi 18990 Ribu rupiah sedangkan paling rendah adalah 6382 Ribu rupiah. Rata-rata Kepadatan Penduduk ( $X_8$ ) adalah 633,2 Jiwa/ $km^2$  dan Kepadatan Penduduk tertinggi 8859 Jiwa/ $km^2$  sedangkan paling rendah adalah 15 Jiwa/ $km^2$ . Rata-rata TPAK ( $X_9$ ) adalah 70,1 persen dan TPAK tertinggi 86,89 persen sedangkan paling rendah adalah 57,81 persen. Oleh karena itu dapat disimpulkan bahwa tingkat pengangguran terbuka memiliki keragaman yang tinggi

#### Estimasi Parameter Menggunakan MKT

Tahapan selanjutnya melakukan estimasi parameter regresi linear berganda menggunakan metode kuadrat terkecil. Estimasi ini digunakan untuk membentuk model dasar yang menggambarkan hubungan antara variabel independen dan variabel dependen. Hasil analisisnya dapat diperoleh sebagai berikut:

Tabel 3.2 Hasil Estimasi Parameter Metode Kuadrat Terkecil

| Variabel  | Estimasi Parameter | Standar Elrror |
|-----------|--------------------|----------------|
| Konstanta | 3,6476             | 5,187          |
| $X_1$     | 0,151              | 0,121          |
| $X_2$     | 0,121              | 0,295          |
| $X_3$     | 0,002              | 0,00471        |
| $X_4$     | -0,030             | 0,065          |
| $X_5$     | 0,021              | 0,119          |
| $X_6$     | $9,618e^{-7}$      | $4,298e^{-7}$  |
| $X_7$     | $-9,156e^{-5}$     | $1,124e^{-5}$  |
| $X_8$     | $1,8001e^{-4}$     | $1,096e^{-4}$  |
| $X_9$     | -0,144             | 0,021          |

Sehingga diperoleh persamaan regresi linear berganda dengan analisis MKT sebagai berikut:

$$Y_{MKT} = 3,6476 + 0,1516X_1 + 0,1210X_2 + 0,0026X_3 - 0,0305X_4 + 0,0215X_5 + 0,00000096X_6 - 0,00009156X_7 + 0,00018001X_8 - 0,14401X_9 \quad (1)$$

### Matriks Korelasi

Berdasarkan model regresi linear berganda yang telah diestimasi, langkah selanjutnya adalah menganalisis hubungan antar variabel independen melalui matriks korelasi. Matriks korelasi digunakan untuk mengidentifikasi adanya hubungan linear yang kuat antar variabel bebas. Langkah awal yang dilakukan adalah melakukan korelasi data terlebih dahulu yang didapat dari transformasi data.

Sehingga diperoleh matriks korelasi sebagai berikut:

$$X^T X = \begin{bmatrix} 1 & 0,917657 & -0,00245 & 0,605638 & 0,29933 & 0,219102 & 0,762976 & 0,587779 & -0,25299 \\ 0,917657 & 1 & -0,06598 & 0,482162 & 0,265302 & 0,19419 & 0,564179 & 0,581972 & -0,17287 \\ -0,00245 & -0,065982 & 1 & 0,098944 & 0,159096 & 0,561627 & 0,081374 & 0,214032 & -0,16217 \\ 0,605638 & 0,482162 & 0,098944 & 1 & 0,321993 & 0,219728 & 0,483651 & 0,388853 & -0,10226 \\ 0,29933 & 0,265302 & 0,159096 & 0,321993 & 1 & 0,240453 & 0,361868 & 0,245289 & 0,088872 \\ 0,219102 & 0,19419 & 0,561627 & 0,219728 & 0,240453 & 1 & 0,307775 & 0,252472 & -0,16427 \\ 0,762976 & 0,564179 & 0,081374 & 0,483651 & 0,361868 & 0,307775 & 1 & 0,458943 & -0,22277 \\ 0,587779 & 0,581972 & 0,214032 & 0,388853 & 0,245289 & 0,252472 & 0,458943 & 1 & -0,16636 \\ -0,252987 & -0,17287 & -0,162174 & -0,102264 & 0,088872 & -0,164275 & -0,22277 & -0,166362 & 1 \end{bmatrix}$$

Berdasarkan matriks korelasi yang diperoleh, terlihat bahwa sebagian variabel independen memiliki nilai korelasi yang cukup tinggi, bahkan mendekati satu, yang menunjukkan adanya hubungan linear yang kuat antar variabel bebas. Dari kondisi ini dapat diinterpretasikan terjadinya multikolinearitas dalam model regresi, dimana

beberapa variabel independen saling menurun secara signifikan sehingga menyebabkan ketidakstabilan dalam pendugaan koefisien regresi apabila menggunakan metode kuadrat terkecil. Selain itu, meskipun terdapat variabel yang memiliki korelasi cukup kuat dengan variabel dependen, pengujian VIF serta penerapan regresi ridge diperlukan untuk memperoleh estimasi parameter yang lebih stabil dan dapat diinterpretasikan dengan baik.

### Uji Multikolinearitas

Uji multikolinearitas digunakan untuk melihat apakah ada korelasi antar variabel independen pada model regresi. Model dianggap baik jika terdapat korelasi antar variabel independen. *Variance Inflation Factor* (VIF) digunakan untuk mendeteksi apakah ada atau tidak multikolinearitas pada model regresi. Kriteria pengujian VIF yaitu apabila nilai  $VIF < 10$  maka tidak terjadi multikolienaritas pada data.

**Tabel 3.2 Nilai VIF untuk Variabel Bebas**

| Variabel Bebas                                | VIF    |
|-----------------------------------------------|--------|
| Indeiks Peimbangan manusi ( $X_1$ )           | 19,039 |
| Rata-rata lama seikolah ( $X_2$ )             | 11,040 |
| Jumlah peinduduk miskin ( $X_3$ )             | 1,719  |
| Umur harapan hidup ( $X_4$ )                  | 1,819  |
| Laju PDRB ( $X_5$ )                           | 1,297  |
| Jumlah peinduduk ( $X_6$ )                    | 1,747  |
| Peingeiluaran peirkapita ( $X_7$ )            | 3,933  |
| Keipadatan peinduduk ( $X_8$ )                | 1,732  |
| Tingkat partisipasi Angkatan keirja ( $X_9$ ) | 1,188  |

Pada Tabel 3.3 dapat dilihat bahwa terdapat 2 data dengan nilai  $VIF > 10$  yaitu  $X_1, X_2$ , sehingga bisa disimpulkan bahwa terdapat multikolinearitas pada data indikator tingkat pengangguran terbuka di pulau Sumatera tahun 2023. Karena asumsi multikolinearitas terlanggar maka dapat diselesaikan dengan metode regresi ridge.

### Pemusatan dan Penskalaan

Setelah teridentifikasi adanya multikolinearitas pada model regresi linear berganda, tahap selanjutnya adalah melakukan pemusatan dan penskalaan data

(standarisasi variabel). Pemusatan dan penskalaan data yang bertujuan untuk menyamakan skala pengukuran seluruh variabel independen sehingga tidak ada variabel yang mendominasi proses estimasi akibat perbedaan satuan atau rentang nilai variabel. Berdasarkan persamaan 2.7 dan 2.8 diperoleh hasil standarisasi sebagai berikut:

**Tabel 3.4 Hasil Standarisasi Data**

| $Y$   | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$  | $X_6$  | $X_7$  | $X_8$ | $X_9$ |
|-------|-------|-------|-------|-------|--------|--------|--------|-------|-------|
| -2,23 | 48,2  | 0,85  | 12,6  | 36,05 | -0,17  | 90405  | 6885   | 14,5  | 68,2  |
| 2,51  | 54,7  | 1,67  | 33,9  | 42,47 | -0,008 | 523401 | 9947   | 90,5  | 53,2  |
| 2,75  | 54,4  | 0,99  | 26,2  | 40,53 | 0,36   | 401370 | 10711  | 107,5 | 64,9  |
| 2,47  | 54,7  | 1,66  | 13,8  | 37,97 | 0,24   | 243307 | 110112 | 77,5  | 56,7  |
| 3,11  | 58,05 | 2,11  | 25,8  | 41,79 | 0,50   | 378796 | 11012  | 286,5 | 57,6  |
| 4,45  | 57,02 | 1,50  | 32,2  | 40,65 | 0,98   | 439170 | 10899  | 338,5 | 53,3  |
| 2,72  | 56,9  | 2,31  | 25,8  | 44,18 | 0,43   | 547577 | 11426  | 305,5 | 57,3  |
| 1,71  | 54,5  | 1,21  | 18,8  | 41,25 | 0,63   | 391053 | 10396  | 110,5 | 60,8  |
| 2,85  | 52,3  | 1,22  | 10,7  | 39,24 | 0,09   | 311743 | 10279  | 78,5  | 61,2  |

Berdasarkan Tabel 3.4 hasil standarisasi data, dapat diketahui bahwa variabel dependen dan seluruh variabel independen memiliki rentang nilai yang sangat berbeda sebelum dilakukan pemusatan dan penskalaan, khususnya pada variabel  $X_6$ ,  $X_7$ , dan  $X_8$  yang memiliki nilai jauh lebih besar dibandingkan variabel lainnya. Kondisi ini menunjukkan adanya perbedaan skala pengukuran yang signifikan antar variabel, sehingga berpotensi menyebabkan dominasi variabel tertentu dalam proses estimasi regresi. Oleh karena itu, pemusatan dan penskalaan data dilakukan untuk menyamakan skala seluruh variabel dengan cara menghilangkan pengaruh satuan pengukuran dan jarak nilai yang berbeda. Dengan dilakukan standarisasi, setiap variabel berada pada skala yang sebanding, sehingga estimasi parameter regresi dapat dilakukan secara lebih stabil, serta pengaruh multikolinearitas dapat dikurangi.

**Perhitungan Nilai  $c$  menggunakan iterasi Hoerl, Kennard, dan Balwin (HKB)**

Setelah data distandarisasi, tahapan selanjutnya dalam analisis regresi adalah penentuan nilai bias parameter  $c$  menggunakan metode Hoerl, Kennard, dan Balwin. Tahap awal dalam iterasi dimulai dengan penentuan nilai tetapan bias  $c$  awal yang diperoleh dari estimasi parameter menggunakan metode kuadrat terkecil. Nilai tetapan bias awal tersebut selanjutnya digunakan sebagai dasar dalam proses estimasi parameter pada regresi ridge, yang dijelaskan dalam rumus iterasi Hoerl, Kennard, dan Balwin (HKB) dengan estimator kuadrat terkecil:

$$c_{HKB} = \frac{P\sigma^2}{\beta^*\beta^*}$$

Nilai  $c_{HKB \text{ optimal}} = 19.81895$

Sehingga diperoleh nilai tetapan  $c$  bias sebagai berikut:

**Tabel 3.5 Nilai  $\beta(c)$  dengan berbagai nilai  $c$**

| $c$   | $X_1$ | $X_2$ | $X_3$ | $X_4$  | $X_5$  | $X_6$ | $X_7$  | $X_8$ | $X_9$   |
|-------|-------|-------|-------|--------|--------|-------|--------|-------|---------|
| 0     | 6,36  | 0,159 | 0,085 | -0,073 | 0,024  | 0,346 | -0,189 | 0,253 | -0,085  |
| 0,68  | 0,593 | 0,188 | 0,88  | -0,066 | 0,021  | 0,341 | -0,171 | 0,251 | -0,854  |
| 1,36  | 0,559 | 0,210 | 0,091 | -0,061 | 0,019  | 0,337 | -0,157 | 0,250 | -0,0851 |
| 2,05  | 0,532 | 0,228 | 0,093 | -0,057 | 0,017  | 0,333 | -0,144 | 0,249 | -0,849  |
| 2,73  | 0,509 | 0,241 | 0,095 | -0,053 | 0,016  | 0,330 | -0,134 | 0,248 | -0,843  |
| 10,25 | 0,393 | 0,301 | 0,107 | -0,029 | 0,005  | 0,307 | -0,068 | 0,242 | -0,810  |
| 10,93 | 0,388 | 0,303 | 0,108 | -0,028 | 0,004  | 0,305 | -0,064 | 0,241 | -0,807  |
| 11,61 | 0,383 | 0,305 | 0,109 | -0,026 | 0,004  | 0,304 | -0,06  | 0,241 | -0,804  |
| ⋮     | ⋮     | ⋮     | ⋮     | ⋮      | ⋮      | ⋮     | ⋮      | ⋮     | ⋮       |
| 19,81 | 0,344 | 0,311 | 0,115 | -0,013 | -0,001 | 0,288 | -0,025 | 0,237 | -0,767  |

Berdasarkan Tabel 3.5 yang menyajikan nilai koefisien regresi  $\beta(c)$  dengan tetapan bias yang diperoleh melalui iterasi Hoerl, Kennard, dan Baldwin (HKB), terlihat bahwa peningkatan nilai parameter ridge  $c$  menyebabkan terjadinya penyusutan koefisien regresi secara bertahap dan stabil. Secara umum, sebagian besar koefisien regresi menunjukkan kecenderungan menurun menuju nol seiring dengan bertambahnya nilai  $c$ , tanpa mengalami perubahan arah pengaruh. Hal ini menunjukkan bahwa penerapan regresi ridge dengan penentuan parameter menggunakan metode HKB mampu mengurangi dampak multikolinearitas serta menghasilkan estimator yang lebih stabil dibandingkan metode kuadrat terkecil. Meskipun demikian, pada beberapa variabel tertentu, seperti  $X_2$  dan  $X_3$ , terjadi peningkatan nilai koefisien pada tahap awal penambahan nilai  $c$ , yang disebabkan oleh adanya redistribusi pengaruh antar variabel bebas yang saling menghambat akibat adanya penalti ridge. Pola perubahan koefisien yang berlangsung secara halus dan konsisten hingga mendekati nilai konvergen  $c$  menunjukkan bahwa nilai tetapan bias yang diperoleh telah optimal dan layak digunakan dalam pembentukan model regresi ridge akhir.

Langkah selanjutnya adalah mensubstitusi nilai konstanta bias  $c$  yang diperoleh melalui proses iterasi metode Hoerl, Kennard, dan Baldwin (HKB) kedalam estimasi parameter pada model regresi ridge:

$$\beta_R = (X^T X + cI)^{-1} X^T Y$$

Sehingga diperoleh estimasi parameternya sebagai berikut:

**Tabel 3.6 Estimasi Parameter Regresi Ridge dengan iterasi HKB**

| Variabel | Nilai Parameter | VIF      |
|----------|-----------------|----------|
| $X_1$    | 0,08838887      | 0,006513 |
| $X_2$    | 0,2337026       | 0,006980 |
| $X_3$    | 0,003402276     | 0,006530 |
| $X_4$    | -0,0009631120   | 0,006746 |
| $X_5$    | 0,002227014     | 0,005935 |
| $X_6$    | 0,0000008322288 | 0,006650 |
| $X_7$    | -0,00002448843  | 0,007517 |
| $X_8$    | 0,0001707633    | 0,007039 |
| $X_9$    | -0,1337326      | 0,005548 |

Tabel 3.6 dapat disimpulkan bahwa tidak terjadi multikolinearitas pada analisis data, dikarenakan nilai *Varians Inflation Factors (VIF)* <10, sehingga didapatkan persamaan regresinya:

$$Y = 5,686689 + 0,08838887X_1 + 0,2337026X_2 + 0,003402276X_3 - 0,0009631120X_4 + 0,002227014X_5 + 0,0000008322288X_6 - 0,00002448843X_7 + 0,0001707633X_8 - 0,1337326X_9 \quad (4.2)$$

### Pengujian Signifikan Parameter Model

#### Uji F

Uji F bertujuan untuk mengetahui apakah seluruh variabel independen secara simultan berpengaruh terhadap variabel dependen dalam suatu model regresi. Uji F dinyatakan sebagai berikut:

$H_0$ : Tidak ada pengaruh variabel bebas secara bersama-sama terhadap variabel terikat

$H_1$ : Ada pengaruh variabel bebas secara bersama-sama terhadap variabel terikat

Kriteria pengambilan keputusan dalam uji ini dilakukan berdasarkan nilai signifikansinya, yaitu:

$H_0$  diterima jika nilai  $F_{tabel} \geq 0,05$

$H_1$  ditolak jika nilai  $F_{tabel} < 0,05$

Tabel 3.7 Hasil Uji F

| $F_{hitung}$ | $F_{tabel}$ | Keterangan |
|--------------|-------------|------------|
| 16,813       | 1,73        | Signifikan |

Berdasarkan tabel 3.7 diperoleh  $F_{hitung} = 16,813$  dan diperoleh nilai  $F_{tabel} = 1,73$ , karena  $F_{hitung} \geq F_{tabel}$  maka tolak  $H_0$ , sehingga dapat disimpulkan bahwa variabel independen memiliki pengaruh secara Bersama-sama terhadap variabel dependen.

#### Uji t

Uji t adalah pengujian statistik yang digunakan untuk mengetahui pengaruh masing-masing variabel independen secara parsial terhadap variabel dependen dalam suatu model regresi. Uji t digunakan untuk mengetahui variabel-variabel yang berpengaruh secara signifikan terhadap model. Hipotesis yang digunakan dalam uji ini yaitu:

$H_0$ : Tidak ada pengaruh variabel bebas secara parsial terhadap variabel terikat

$H_1$ : Ada pengaruh variabel bebas secara parsial terhadap variabel terikat

Adapun kriteria pengambilan keputusan pada uji t yaitu:

$H_0$  diterima jika nilai  $t_{tabel} \geq 0,05$

$H_1$  ditolak jika nilai  $t_{tabel} < 0,05$

$t_{tabel} = 1,655$

**Tabel 3.8 Hasil Uji t**

| Peinduga parameteir | $t_{hitung}$ | Keiteirangan     |
|---------------------|--------------|------------------|
| $\beta_1$           | 7,459        | Signifikan       |
| $\beta_2$           | 6,730        | Signifikan       |
| $\beta_3$           | 2,774        | Signifikan       |
| $\beta_4$           | 3,680        | Signifikan       |
| $\beta_5$           | 1,529        | Tidak Signifikan |
| $\beta_6$           | 4,577        | Signifikan       |
| $\beta_7$           | 5,181        | Signifikan       |
| $\beta_8$           | 5,814        | Signifikan       |
| $\beta_9$           | -8,098       | Tidak Signifikan |

Berdasarkan Tabel 3.8 diperoleh 2 variabel tidak signifikan yaitu laju pertumbuhan domestik bruto ( $X_5$ ) dan tingkat partisipasi angkatan kerja ( $X_9$ ). Hal ini dapat dilihat berdasarkan perbandingan antara nilai  $t_{tabel}$  dan  $t_{hitung}$ . Apabila  $|t_{hitung}| \geq t_{tabel}$  maka terima  $H_0$ , sehingga variabel independen tersebut tidak signifikan terhadap variabel dependen dan sebaliknya apabila  $|t_{hitung}| > t_{tabel}$  maka tolak  $H_0$ , sehingga variabel independen signifikan terhadap variabel dependen.

## SIMPULAN

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan, diperoleh Kesimpulan sebagai berikut:

1. Metode Regresi ridge menggunakan konstanta estimator Hoerl, Kennard dan Baldwin terbukti dapat mengatasi masalah multikolinearitas. hal ini dilihat dari nilai VIF regresi ridge yang berkurang setelah di transformasi dengan persamaan regresi ridge.
2. Berdasarkan hasil estimasi parameter regresi ridge menggunakan iterasi Hoerl Kennard dan Baldwin diperoleh  $C_{Hkb}$  yang konvergen sebesar 19,81. Konvergensi nilai  $C_{Hkb}$  ini menunjukkan bahwa proses iterasi telah mencapai kondisi stabil, sehingga nilai parameter yang dihasilkan tidak lagi mengalami perubahan signifikan pada iterasi selanjutnya.
3. Diperoleh model persamaan regresi linear berganda dengan analisis *Ridge Regression* menggunakan metode Hoerl, Kennard dan Baldwin (HKB) sebagai:

$$Y = 5,686689 + 0,08838887X_1 + 0,2337026X_2 + 0,003402276X_3 - 0,0009631120X_4 + 0,002227014X_5 + 0,0000008322288X_6 - 0,00002448843X_7 + 0,0001707633X_8 - 0,1337326X_9$$

4. Berdasarkan model Regresi Ridge yang diestimasi menggunakan metode Hoerl, Kennard dan Baldwin (HKB) maka dapat diidentifikasi laju PDRB, dan tingkat partisipasi angkatan kerja yang tidak berpengaruh terhadap tingkat pengangguran terbuka di pulau Sumatera pada tahun 2023.

## UCAPAN TERIMA KASIH

Terima kasih kepada Program Studi Matematika, Fakultas Sains dan Teknologi, UIN Imam Bonjol Padang yang telah menyediakan banyak fasilitas dan dukungan kepada penulis.

## DAFTAR PUSTAKA

- [1] R. E. Christine, "DETERMINAN TINGKAT PENGANGGURAN TERBUKA DI KABUPATEN/KOTA PROVINSI SUMATERA UTARA PERIODE 2020-2024 Ruth," *E-Jurnal EP Unud*, vol. 32, no. 3, pp. 167-186, 2021.
- [2] R. E. N. Rongcai, W. U. Guoxiong, and C. A. I. Ming, "INDIKATOR KINERJA UTAMA DINAS PENDIDIKAN DAN KEBUDAYAAN KOTA MOJOKERTO".
- [3] D. Alrakhman, D. Susetyo, Taufiq, and Azwardi, "Effect of Unemployment Rate, Inequality and Investment Against Economic Growth on The Island of Sumatra," *Proc. 7th Sriwij. Econ. Accounting, Bus. Conf. (SEABC 2021)*, vol. 647, no. Seabc 2021, pp. 48-56, 2022, doi: 10.2991/aebmr.k.220304.007.
- [4] A. Andini, E. Sunandi, P. Novianti, I. Sriliana, and W. Agwil, "Perbandingan Metode Regresi Ridge dan Jackknife Ridge Regression pada Data Tingkat Pengangguran Terbuka," *J. Math. Its Appl.*, vol. 22, no. 1, pp. 79-86, 2025, doi: 10.12962/limits.v22i1.3374.
- [5] M. A. Yusuf, Herman, T. H. A. Abraham, and H. Rukmana, "Analisis Regresi Linier Sederhana dan Berganda Beserta Penerapannya," *J. Educ.*, vol. 06, no. 02, pp. 13331-13344, 2024.
- [6] L. H. Hasibuan and S. Musthofa, "Penerapan Metode Regresi Linear Sederhana Untuk Prediksi Harga Beras di Kota Padang," *J. Sci. Technol.*, vol. 2, no. 1, pp. 85-95, 2022.
- [7] M. Sriningsih, D. Hatidja, and J. D. Prang, "Multicollinierity Handling Using Principal Components Refression on Imported Rice Case in North Sulawesi Province," *J. Stat. Dan Mat.*, vol. 18, no. 01, pp. 18-24, 2018.
- [8] L. H. Hasibuan, F. Yanuar, and D. Devianto, "Hypertension Status Model with Bayesian Adaptive LASSO Binary Quantile Regression Using Asymmetric Laplace Distribution," *IAENG Int. J. Comput. Sci.*, vol. 52, no. 8, 2025.
- [9] L. H. Hasibuan, F. Yanuar, D. Devianto, and M. Maiyastri, "Modeling Classification Of Stunting Toddler Height Using Bayesian Binary Quantile Regression With Penalized Lasso," *Mathline J. Mat. dan Pendidik. Mat.*, vol. 10, no. 2, pp. 373-386, 2025.
- [10] L. H. Hasibuan, F. Yanuar, V. P. Harahap, and L. Qalbi, "Performance Quantile Regression and Bayesian Quantile Regression in Dealing with Non-normal Errors (Case Study on Simulated Data)," *Numer. J. Mat. dan Pendidik. Mat.*, vol. 8, no. 2, pp. 46-54, 2024.
- [11] Rosyadi, "PENERAPAN METODE REGRESI RIDGE UNTUK MENGATASI MASALAH MULTIKOLINEARITAS PADA KASUS INDEKS PEMBANGUNAN MANUSIA DI PROVINSI JAWA TENGAH," vol. 3, no. 32, pp. 1-44, 2018.
- [12] S. N. Sugeng, B. Pratikno, D. Matematika, and U. J. Soedirman, "Regresi ridge dan simulasinya pada data kemiskinan," no. 2020, pp. 956-962, 2023.
- [13] N. B. Pratiwi, "Perbandingan Regresi Komponen Utama Dengan Regresi Ridge Untuk Mengatasi Masalah Multikolinieritas," pp. 1-74, 2016.

- [14] N. A. Solekakh, D. Ispriyanti, and S. Sudarno, "Estimasi Parameter Regresi Ridge Menggunakan Iterasi Hoerl, Kennard, dan Baldwin (HKB) Untuk Penanganan Multikolinearitas," *J. Gaussian*, vol. 4, no. 4, pp. 1109–1116, 2015.