Pemodelan *New Ridge Regression Estimator* pada Tingkat Kemiskinan di Kabupaten/Kota Provinsi Jawa Barat Tahun 2020

Ridho Febriansyah Tambunan*, Suliadi

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

*Rftambunan2@gmail.com, suliadi@gmail.com

Abstract. Linear regression is a statistical method used to predict value dependent variable or response with one or more independent variables. If there is more than one predictor variable, multiple linear regression analysis is used. Ridge regression estimator has been introduced as an alternative to the ordinary least squares estimator (OLS) in the presence of multicollinearity. Ridge regression minimizes the mean square residual by introducing a bias constant and produced biased but stable coefficients estimate. The aim of this research is to apply a method introducing by Al-hassan (2010) to obtaine the bias constant in ridge regression that produces smaller bias than method given by Hoerl & Kennad. We apply this method to model the poverty rate in districts/cities in West Java in 2020. The dependent variable (Y) is the proverty rate and the independet variables are X_1(average length of school), X_2 (unemployment rate), X_3(gross domestic regional product), X_4(human development index), X_5(number of labor force participation rate). The value of the ridge constant using the Al-hassan (2010) method is 1.377633. The ridge regression model for the standardized variables is $y^* = -0.2090Z_1 - 0.0024Z_2 - 0.1141\,Z_3 - 0.2114\,Z_4\ - 0.0809\,Z_5$ with Z_1 , Z_3 & Z_4 that significanly affect the reponse. The regression model based on the original variable is $\hat{y} = 21.56967 - 0.4069871\,x1 - 0.002831064\,x2 - 0.000003772087\,x3 - 0.1257099x4 - 0.0000003769211\ x5$

Keywords: *Multiple regression, ridge regression, new constant of the estimator, poverty level.*

Abstrak. Analisis regresi linier adalah metode statistika yang digunakan untuk membentuk model hubungan antara variabel terikat (*dependent* atau respon) dengan satu atau lebih variabel bebas (*independent* atau prediktor). Apabila variabel prediktor lebih dari satu maka digunakan analisis regresi linier berganda. Ada beberapa asumsi yang harus terpenuhi dalam regresi linier berganda diantaranya asumsi multikolinearitas. Salah satu metode untuk mengatasi masalah multikolinieritas adalah menggunakan metode regresi *ridge*. Regresi *ridge* meminimumkan residual dengan menambahkan tetapan bias (k). Namun metode ini masih memiliki kelemahan yaitu masih terdapat bias. Untuk memperbaiki kelemahan tersebut Al-hassan mengajukan metode baru. Metode ini bertujuan untuk memperkecil nilai bias dari suatu penduga dengan cara memodifikasi nilai *k*. Dalam skripsi ini kami menerapkan metode tersebut untuk memodelkan tingkat kemiskinan di Kabupaten/Kota di Jawa Barat Tahun 2020. Variabel responnya adalah Y (tingkat kemiskinan) dan variabel bebasnya X_1(lama rata-rata sekolah), X_2(tingkat pengangguran terbuka), X_3(produk domestik regional bruto), X_4(indeks pembangunan manusia), X_5(jumlah angkatan kerja). Nilai konstanta *ridge* menggunakan metode Al-hassan (2010) sebesar 1.377633 Sehingga didapatkan model persamaan *ridge* yaitu $: y^* = -0.2090Z_1 - 0.0024Z_2 - 0.1141\,Z_3 - 0.2114\,Z_4\ - 0.0809\,Z_5$ Dengan variabel baku Z_1 , Z_3 dan varibel baku Z_4 yang signifikan terhadap variabel y^*. Dan model berdasarkan variabel aslinya adalah $\hat{y} = 21.56967\ - 0.4069871\,x1 - 0.002831064\,x2 - 0.000003772087\,x3 - 0.1257099x4 - 0.0000003769211\ x5$

Kata Kunci: *Regresi Linier Berganda, Regresi Ridge, Parameter k Baru, Tingkat Kemiskinan.*

Corresponding Author
Email: suliadi@gmail.com

A. Pendahuluan

Dalam pendugaan parameter model regresi, salah satu metode yang paling umum digunakan adalah metode kuadrat terkecil (MKT). Salah satu asumsi yang harus diperhatikan dalam MKT adalah tidak ada multikolinieritas antar variabel bebas. Multikolinearitas dalam model regresi linier diperkenalkan oleh Ragnar Frisch pada tahun 1934 yaitu adanya hubungan linier antar variabel bebas. Jika terjadi masalah multikolinearitas, hasil penduga koefisien regresi tidak stabil dan varians dari koefisien regresi sangat besar (Gujarati, 2007). Untuk mengatasi permasalahan tersebut maka diperlukan metode penduga alternatif, salah satunya adalah regresi ridge.

Regresi *ridge* adalah suatu penduga untuk mengatasi masalah multikolinearitas. Penduga *ridge* diperkenalkan oleh Hoerl & Kennad dalam mengatasi multikolinearitas pada suatu data (Supranto, 2005). Regresi *ridge* adalah pengembangan dari MKT (metode kuadrat terkecil) atau OLS (*Ordinary Least Square)* yang telah dilakukan dimodifikasi dengan menambahkan parameter (k) sebagai tetapan bias pada matriks varians-kovarians terhadap variabel bebas

Parameter tetapan bias (k) mempunyai fungsi yang penting dalam regresi *ridge*. Maka dibutuhkan pemilihan yang tepat terhadap parameter tetapan bias. Hasil dari penelitian yang dilakukan oleh Hoerl & Kennard menunjukkan bahwa penduga Hoerl & Kennard memiliki MSE yang lebih baik dibandingkan OLS untuk semua tingkat korelasi dan penduga Hoerl & Kennard memberikan hasil yang cukup baik akan tetapi tetap bersifat bias. Namun metode ini masih memiliki kelemahan yaitu masih terdapat bias. Untuk memperbaiki kelemahan tersebut Al-Hassan (2010) mengajukan metode baru dengan cara menambahkan parameter yang disarankan oleh Alkhamisi & Shukur (2007). Sebelumnya Alkhamisi & Shukur (2007) melakukan modifikasi dengan cara menambahkan ${}^{1}/_{\lambda_{max}}$ terhadap metode yang disampaikan oleh Hocking et al., (1976). Metode ini bertujuan untuk memperkecil nilai bias dari suatu penduga dengan cara memodifikasi nilai *k* Al-Hassan M (2010).

Tujuan yang ingin dicapai pada penelitian ini adalah untuk memodelkan data tingkat kemiskinan di Provinsi Jawa Barat tahun 2020 menggunakan metode *New Ridge Regression Estimator* dalam mengatasi masalah multikolinieritas.

B. Metodologi Penelitian

Peneliti menggunakan metode *New Ridge Regression Estimator*, dimana data yang digunakan adalah data sekunder yang merupakan data tingkat kemiskinan dari 27 kabupaten dan kota di Prov Jawa Barat tahun 2020. Adapun variabel yang diduga mempengaruhi tingkat kemiskinan ialah, lama rata-rata sekolah, tingkat pengangguran terbuka, pendapatan domestik regional bruto, indeks pembangunan manusia, jumlah angkatan kerja. Tahapan analisis dapat dilakukan sebagai berikut:

Regresi Linier

Regresi linier adalah analisis regresi yang menjelaskan bentuk hubungan dari dua atau lebih variabel bebas *X* dengan variabel terikat *Y*. Tujuan analisis regresi linier berganda adalah untuk memprediksi atau perkiraan nilai *Y* atas *X*. Secara model regresi linier berganda Draper & Smith (1998). Metode kuadrat terkecil merupakan metode yang paling terkenal dalam pendugaan parameter regresi (Suliadi, 2015).

Untuk Populasi

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} + \varepsilon_i \quad (1)$$

Untuk sampel

$$Y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + \cdots + b_p x_{pi} + e_i. \quad (2)$$

Normalitas

Untuk mendeteksi apakah galat berdistribusi tertentu atau tidak bisa menggunakan pengujian, yaitu dengan uji Kolmogorov-smirnov. Metode *Kolmogorov – Smirnov* dikemukakan oleh *Andrey Kolmogorov* dan *Nikolai Smirnov* yang merupakan dua matematikawan asal rusia pada tahun 1939. Uji *Kolmogorov – Smirnov* digunakan untuk menguji apakah distribusi sampel yang teramati sesuai dengan distribusi teoritis tertentu atau tidak.

Multikolinieritas

Multikolinieritas pertama kali diperkenalkan oleh *Ragnar Frich* pada tahun 1934, yaitu adanya hubungan linier antara variabel bebas dari model regresi linier berganda. Suatu model dikatakan ada masalah multikolinieritas adalah model yang memiliki nilai faktor *Variance Inflation Factors* (VIF) > 10 mengindikasikan terdapatnya multikolinieritas (Myers, 1990). Rumus VIF adalah:

$$VIF_j = \frac{1}{1 - R_j^2} \tag{3}$$

Regresi *Ridge*

Regresi *ridge* adalah salah satu metode untuk memperbaiki masalah multikolinieritas dengan memodifikasi parameternya untuk mendapatkan penduga bias dari koefisien regresi. Regresi *ridge* ini pertama kali diperkenalkan oleh *Hoerl & Kennard* (1970) dengan cara menambahkan konstanta yang bernilai positif $\boldsymbol{k}$ terhadap elemen diagonal $\boldsymbol{X^tX}$. maka akan diperoleh hasil dari penaksir yang mendekati nilai parameter yang sebenarnya. Untuk menghitung model standar dari regresi linier berganda (Draper & Smith, 1998) menggunakan rumus berikut:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}. \tag{4}$$

Menurut Kutner et al. (2004) penaksiran parameter regresi *ridge* dilakukan dengan cara menstandarisasi variabel bebas dan variabel tak bebas dengan model.

$$\boldsymbol{y_i^*} = \boldsymbol{\beta_1^* z_{i1}^*} + \boldsymbol{\beta_2^* z_{i2}^*} + \cdots + \boldsymbol{\beta_p^* z_{ip}^*} + \boldsymbol{\varepsilon} \tag{5}$$

$$y_i^* = \frac{y_i - \bar{y}}{s_y}, \tag{6}$$

$$Z_j = \frac{x_j - \bar{x}_j}{s_{x_j}}, \tag{7}$$

Dimana

$$S_y = \sqrt{\frac{\sum_i^n (Y_i - \hat{Y})^2}{n-1}}.$$

$$S_{Xj} = \sqrt{\frac{\sum_i^n (X_{ij} - \bar{X}_j)^2}{n-1}}.$$

Dari proses tersebut diperoleh data yang dibakukan yang secara matriks dinotasikan sebagai $\mathbf{Z}$. prinsip dasar regresi *ridge* adalah menambahkan konstanta bias pada diagonal matriks $\mathbf{Z'Z}$. Diperoleh penduga regresi *ridge* yaitu:

$$\boldsymbol{\hat{\beta}_R} = (\mathbf{Z'Z} + \boldsymbol{kI})^{-1}\mathbf{Z'}\boldsymbol{y}^* \tag{8}$$

dimana I adalah matriks identitas berukuran $(\boldsymbol{p\ x\ p})$ dan *k* adalah sebuah bilangan yang positif atau $\boldsymbol{k} \geq \boldsymbol{0}$, umumnya *k* terletak antara interval 0 < *k* < 1. Dalam prakteknya nilai $\boldsymbol{k}$ tidak diketahui. Hoerl dan Kennad (1970) memberikan MSE untuk regresi *ridge* yaitu:

$$\mathbf{MSE}\,(\boldsymbol{\hat{\beta}_R}) = \boldsymbol{\sigma^2} \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} + \boldsymbol{k^2\beta'}(\mathbf{Z'Z} + \boldsymbol{k}\mathbf{I})^{-2}\boldsymbol{\beta}. \tag{9}$$

Misalkan D adalah sebuah matriks ortogonal sehingga $\boldsymbol{D'CD} = \boldsymbol{\Lambda}$, dimana $\boldsymbol{C} = \mathbf{Z'Z}$ dan $\boldsymbol{\Lambda} = \boldsymbol{diag}(\boldsymbol{\lambda_1}, \boldsymbol{\lambda_2}, \cdots, \boldsymbol{\lambda_p})$ berisi nilai eigen dari vektor matriks C, serta $\boldsymbol{X^*} = \boldsymbol{ZD}$ (Al-Hassan M., 2010), maka bentuk kanonik dari model persamaan (4) adalah

$$\boldsymbol{y^*} = \boldsymbol{X^*\alpha} + \boldsymbol{e}. \tag{10}$$

Al-Hassan (2010) melakukan modifikasi metode Hocking $(\widehat{\boldsymbol{k}}_{\boldsymbol{HSL}})$, dengan menambahkan nilai $\boldsymbol{1}/\boldsymbol{\lambda_{max}}$ seperti yang disarankan oleh Alkhamisi & Shukur (2007), dan diperoleh penduga parameter Ridge sebagai berikut

$$\hat{k}_{NHSL} = \frac{\sigma^2 \lambda_{max} \sum_{i=a}^{p} (\lambda_i \hat{\alpha}_i)^2 + (\sum_{i=1}^{p} \lambda_i \hat{\alpha}_i^i)^2}{\lambda_{max} (\sum_{i=1}^{p} \lambda_i \hat{\alpha}_i^2)^2} \tag{11}$$

$$= \hat{\sigma}^2 \frac{\sum_{i=1}^{p} (\lambda_i \hat{\alpha}_i)^2}{(\sum_{i=1}^{p} \lambda_i \hat{\alpha}_i^2)^2} + \frac{1}{\lambda_{max}} = \hat{k}_{HSL} + \frac{1}{\lambda_{max}}$$

dimana $\hat{\sigma}^2 = \sum e_i^2 \,/ n - p$

karena $1/\lambda_{max} > 0$, maka $\hat{k}_{NHSL}$ lebih besar dari $\hat{k}_{HSL}$.

Pengujian Parameter *Ridge* Secara Parsial Koefisien Regresi

Pengujian ini digunakan untuk membuktikan pengaruh variabel bebas terhadap variabel terikat secara individu (parsial). Uji parsial pada analisis regresi ridge sedikit berbeda dengan uji parsial pada regresi linier berganda karena standard erronya tidak sama. Oleh karena itu akan digunakan pengujian hipotesis untuk melihat signifikansi setiap variabel sebagai berikut (Cule et al., 2011). Statistik uji yang digunakan adalah statistik $\boldsymbol{t_{hitung}}$ dengan rumus:

$$t_{hitung} = \frac{\hat{\beta}_j^R}{se(\hat{\beta}_j^R)}, \tag{12}$$

Dimana:

$Var\,(\hat{\beta}^R) = \sigma^2(\mathbf{Z'Z} + \boldsymbol{k}\,\mathbf{I})^{-1}\mathbf{Z'Z}(\mathbf{Z'Z} + \boldsymbol{k}\,\mathbf{I})^{-\mathbf{1}}$

Untuk $\hat{\sigma}^2 = \frac{(y^* - Z\hat{\beta})'(y^* - Z\hat{\beta})}{n-p}$

$se(\hat{\beta}_j)$ = diagonal ke-j matriks $Var\,(\hat{\beta}^R)$.

Daerah kritis yang digunakan adalah berdasarkan perbandingan nilai t_{hitung} dengan t_{tabel} dimana t_{tabel} didapatkan dari tabel t dengan tingkat signifikansi α, maka jika $t_{hitung} > t_{tabel}$ maka H_0 ditolak. Sedangkan jika $t_{hitung} < t_{tabel}$ maka H_0 diterima. Selain dari daerah kritis tersebut, dapat juga menggunakan nilai p-value < α maka H_0 ditolak.

Setelah nilai signifikansi diperoleh model regresi *ridge* dapat dikembalikan ke dalam bentuk variabel asalnya. Hubungan penduga parameter *ridge* $\hat{\beta}_{R1}, \hat{\beta}_{R2}, \dots, \hat{\beta}_{Rp}$ dengan koefisien parameter regresi linier berganda $\beta_1, \beta_2, \dots, \beta_p$ yang dapat dituliskan sebagai berikut:

$$\beta_j = \left(\frac{S_y}{S_{xj}}\right)\hat{\beta}_R, \tag{13}$$

Menurut rumus untuk mendapatkan β_0 sebagai berikut:

$\beta_0 = \bar{Y} - \beta_1\bar{\mathrm{Z}}_1 - \beta_2\bar{\mathrm{Z}}_2 - \cdots - \beta_P\bar{\mathrm{Z}}_P$.

C. Hasil Penelitian dan Pembahasan

Peneliti menggunakan metode *New Ridge Regression Estimator*, dimana data yang digunakan adalah data sekunder yang merupakan data tingkat kemiskinan dari 27 kabupaten dan kota di Provinsi Jawa Barat tahun 2020. Dimana terlebih dahulu kita harus menghitung model dari regresi linier berganda menggunakan persamaan (2), setelah itu melihat apakah data mengandung multikolinieritas mengggunakan persamaan (3), jika terdapat multikolinieritas pada data maka selanjutnya dilakukan pendugaan menggunakan metode *ridge*. Sebelum melakukan pendugaan regresi *ridge*, tahap awal ialah dilakukannya standarisasi pada variabel X dan Y menggunakan persamaan (6) dan (7), kemudian dilakukan perhitungan paramater *ridge* menggunakan rumus persamaan (11), setelah diketahui nilai parameter *ridge* nya maka bisa dilakukan pengujian parameter *ridge* menggunakan persamaan (8). Selanjutnya dilakukan pengujian parsial koefisien regresi *ridge* menggunakan persamaan (12), dan dilakukan proses mengembalikan ke dalam bentuk variabel asalnya menggunakan persamaan (13).

Model Regresi Linier Berganda

Analisis regresi linier berganda menggunakan Metode Kuadrat Terkecil (MKT) untuk mengestimasi koefisien regresi dan melihat pengaruh antara variabel prediktor dan variabel respon. Hasil analisis MKT diperoleh sebagai berikut:

$$\hat{Y} = 27.211 - 0.975\,X_1 + 0.209\,X_2 - 0.000006063\,X_3 - 0.157\,X_4 - 0.00000094311X_5$$

Tabel 1 Hasil ANOVA Pengujian Parameter

Sumber Seragam	Jumlah Kuadrat	Derajat Bebas (db)	Kuadrat Tengah	F Hitung
Regresi	132.598	5	26.250	7.690
Sisaan	72.424	21	3.449	
Total	205.022	26		

Dengan kriteria uji Tolak H_0 jika $F_{hitung} > F_{tabel}$ Dari hasil perhitungan yang telah dilakukan diatas maka dapat disimpulkan H_0 ditolak karena $F_{hitung} > F_{tabel}$ atau 7.690 > 2.5717. Artinya, minimal ada satu koefisien regresi yang tidak sama dengan nol, atau minimal ada satu variabel prediktor yang mempengaruhi variabel respon secara signifikan.

Normalitas

Pada uji normalitas, residual (eror/galat) harus berdistribusi normal. Dengan nilai Kolmogorov-smirnov 0.126 dan nilai P-value sebesar 0.200 maka dapat disimpulkan H_0 diterima karena P-value > α , 0.200 > 0.05 maka galat mengikuti distribusi normal.

Heteroskedastisitas

Uji asumsi ini dilakukan untuk mengetahui dalam model regresi terjadi ketidaksamaan variansi dari residual antara satu pengamatan ke pengamatan lain. Jika variansi residual antara satu observasi ke observasi lain berbeda maka disebut heteroskedastisitas. Metode Breush-Pagan mengasumsikan bahwa galat harus berdistribusi normal. Sebelumnya telah di uji bahwa galat berdistribusi normal. Misalkan ragam merupakan fungsi dari $\hat{Y}$ hasil analisis MKT maka metode breush-pagan dari hasil analisis regresi didapat nilai JKS = 72.424, dan nilai dari regresi dengan e_i^2 sebagai variabel respon dan $\hat{Y}$ sebagai variabel bebas . diperoleh JKR* = 2.316. H_0 diterima sebab 0.1069 < 3.8415 Artinya ragam galat homogen pada taraf

Uji Multikolinieritas

Uji asumsi multikolinieritas pada analisis regresi linier berganda yaitu menggunakan nilai VIF. Apabila nilai VIF > 10 maka mengindikasikan ada masalah multikolinieritas. Nilai VIF untuk masih masing variabel sebagai berikut:

Tabel 2 Uji Multikolinieritas

Variabel	VIF	Keterangan
X_1	11.124	Multikolinieritas
X_2	1.338	Tidak Multikolinieritas
X_3	2.978	Tidak Multikolinieritas
X_4	11.635	Multikolinieritas
X_5	2.488	Tidak Multikolinieritas

Berdasarkan tabel diatas dapat dilihat bahwa X_1 dan X_2 memiliki nilai VIF > 10 sehingga dapat disimpulkan bahwa terjadi multikolinieritas pada variabel-variabel prediktor sehingga perlu diatasi menggunakan metode-metode tertentu, salah satunya regresi *Ridge*.

Regresi *Ridge*

Sebelum melakukan pemodelan regresi *Ridge*. perlu dilakukan transformasi data terlebih dahulu yaitu dengan dilakukan standarisasi atau pembakuan data untuk meminimumkan kesalahan dalam pembulatan data dan juga transformasi ini menghilangkan β_0 yang akan membuat perhitungan regresi menjadi lebih mudah dan sederhana.

Langkah pertama yang dilakukan untuk mencari nilai k_{NHSL} adalah menentukan nilai matriks Λ=Diag(63.69515, 41.9983, 17.2498, 5.907652, 1.149099). Setelah mendapat nilai Λ, kemudian diperoleh nilai $\hat{\alpha}$ sebagai berikut:

Setelah mendapat nilai Λ, kemudian diperoleh nilai $\hat{\alpha}$ sebagai berikut:

$$\hat{\alpha} = \begin{bmatrix} 0.4530641 \\ -0.1597883 \\ 0.3852520 \\ -0.1098683 \\ -0.1797787 \end{bmatrix}$$

Langkah selanjutnya untuk menentukan parameter k_{NHSL} diperlukan nilai $\hat{\sigma}^2$, maka diperoleh nilai $\hat{\sigma}^2$ sebagai berikut:

Nilai total dari e_i^2 = 9.18452 dengan nilai n = 27 dan p = 5 maka,

$\hat{\sigma}^2 = \frac{9.18452}{22} = 0.4174782$

Kemudian dilakukan perhitungan nilai k_{NHSL} menggunakan persamaan (11) sebagai berikut:

$$k_{NHSL} = 0.4174782\ \frac{922.4455}{282.7604} + 0.01569978$$

$$k_{NHSL} = 1.377633$$

maka didapat nilai k_{NHSL} sebesar 1.377633

Menghitung Parameter *Ridge*

Langkah pertama dalam menghitung penduga parameter adalah *New Ridge Regression Estimator* adalah menentukan nilai parameter *ridge* sehingga diperoleh nilai k_{NHSL} = 1.377633. Berdasarkan nilai parameter k_{NHSL} pada regresi *ridge* menyebabkan koefisien regresi yang dihasilkan semakin menyusut. Koefisien analisis disajikan sebagai berikut

Berdasarkan perhitungan menggunakan *Software R Studio* koefisien regresi *Ridge* dapat dituliskan kedalam persamaan model regresi *Ridge* sebagai berikut:

$$y^* = -0.2090Z_1 - 0.0024Z_2 - 0.1141\,Z_3 - 0.2114\,Z_4 - 0.0809\,Z_5$$

Pengujian Parsial Koefisien Regresi *Ridge*

Dari persamaan model regresi *ridge* selanjutnya dilakukan pengujian signifikan parameter koefisien regresi secara parsial dengan tingkat signifikansi sebesar 5%. Dengan statistik uji t_{tabel} maka $t_{(27,0,05/2)}$ = 2.0518 dan kriteria uji Tolak H_0 jika $|t_{hitung}| > t_{tabel}$ atau P-value < α.

Variabel Z_1 (lama rata-rata sekolah) signifikan
4.7570 > 2.0518
Variabel Z_3 (produk domestik regional bruto) signifikan
2.2914 > 2.0518
Variabel Z_4 (indeks pembangunan manusia) signifikan
4.8891 > 2.0518

dan ada dua variabel yaitu Z_2 (tingkat pengangguran terbuka) dan Z_5 (jumlah angkatan kerja) yang tidak signifikan karena $t_{hitung} < t_{tabel}$.

Setelah dilakukan uji parsial *ridge*, maka selanjunya akan dikembalikan kedalam variabel asli. Berikut hasil penaksiran parameter regresi linier berganda dengan analisis regresi *ridge.*

Sehingga persamaan model regresi linier berganda dengan analisis regresi *ridge* menggunakan metode *New Estimator New Ridge Regression* sebagai berikut:

$$\hat{y} = 21.56967 - 0.4069871\,x1 - 0.002831064\,x2 - 0.000003772087\,x3 - 0.1257099x4 - 0.0000003769211\ x5$$

Karena hanya x_1, x_3 dan x_4 yang signifikan, maka hanya variabel tersebut yang akan di intepretasikan.

Koefisien regresi rata-rata lama sekolah (x1) sebesar -0.4069871 yang berarti bahwa setiap peningkatan satu tahun rata-rata lama sekolah, maka akan menurunkan tingkat kemiskinan sebesar 40.70%. Koefisien regresi produk domestik regional bruto (x3) sebesar -0.000003772087 yang berarti bahwa setiap peningkatan satu juta rupiah produk domestik regional bruto, maka akan menurunkan tingkat kemiskinan sebesar 0.000003772087 atau produk domestik regional bruto meningkat 1 triliyun maka tingkat kemiskinan akan menurun sebesar 3.78%. Koefisien regresi indeks pembangunan manusia (x4) sebesar -0.1257099 yang berarti bahwa setiap peningkatan satu satuan indeks pembangunan manusia , maka akan menurunkan tingkat kemiskinan sebesar 0.1257099%

D. Kesimpulan

Berdasarkan pembahasan diatas dalam penelitian ini, peneliti menyimpulkan terdapat masalah multikolinieritas yaitu pada variabel x1 (lama rata-rata sekolah) dan x4 (indeks pembangunan manusia). Nilai konstanta bias dengan menggunakan metode Al-hassan diperoleh sebesar 1.377633. Model regresi melalui *New Ridge Estimator* diperoleh sebagai berikut : $\hat{y} = 21.56967 - 0.4069871\,x1 - 0.002831064\,x2 - 0.000003772087\,x3 - 0.1257099x4 - 0.0000003769211\ x5$. Pada penelitian ini hanya x_1(lama rata-rata sekolah), x_3 (produk domestik regional bruto) dan x_4 (indeks pembangunan manusia) yang signifikan. Dimana lama rata-rata sekolah (x_1) berpengaruh negatif terhadap tingkat kemiskinan, produk domestik regional bruto (x_3) berpengaruh negatif terhadap tingkat kemiskinan, indeks pembangunan manusia (x_4) berpengaruh negatif terhadap tingkat kemiskinan.x

Acknowledge

Terima kasih saya ucapkan kepada Allah Swt yang telah memberikan rahmat dan hidayah kepada saya sebagai. Terima kasih juga saya ucapkan kepada kedua orang tua yang membimbing saya hingga bisa sampai titik ini. Terima kasih juga saya ucapkan kepada bapak Suliadi, M.si, P.hd, yang sudah membantu dalam proses mengerjakan artikel ini.

Daftar Pustaka

[1] Al-Hassan M., Y. (2010). Performance of a *New Ridge Regression Estimator*. *Journal of the Association of Arab Universities for Basic and Applied Sciences*, *9*(1), 23–26. https://doi.org/10.1016/j.jaubas.2010.12.006

[2] Draper, R. N., & Smith, H. (1998). *Applied Regression Analysis* (3rd ed.). Canada.

[3] Gujarati, S. (2007). *Ekonometrika Dasar* (1st ed.). Erlangga.

[4] Hocking, R., Speed, F, M., & Lynn, M, J. (1976). A class of biased estimators in linear regression. *Technometrics*, *18*(4), 425–437. https://doi.org/https://doi.org/10.2307/1268658

[5] Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2004). *Applied Linear Statistical Models* (5th ed.). McGraw-Hill Irwin.

[6] Myers, R. (1990). *Classical and Modern Regression with Applications*. Duxburry.

[7] Suliadi. (2015). Analisis Regresi. In *Diktat Kuliah Program Studi Statistika*. Universitas Islam Bandung.

[8] Supranto, J. (2005). *Ekonometri* (1st ed.). Bogor Ghalia Indonesia.

[9] Khoeriyah, Risti Yulianti. (2021). *Regresi Terboboti Geografis Semiparametrik (RTG-S) untuk Pemodelan Indeks Pembangunan Kesehatan Masyarakat Kabupaten/Kota di Sumatera Utara*, Jurnal Riset Statistika, 1(1), 43-50.