

MODEL REGRESI RIDGE UNTUK MENGATASI MODEL REGRESI LINIER BERGANDA YANG MENGANDUNG MULTIKOLINIERITAS

(Studi Kasus: Data Pertumbuhan Bayi di Kelurahan Namaelo RT 001, Kota Masohi)

Ridge Regression Model to Overcome Multiple Linear Regression with Multicollinearity

(Case Study: Growth of Infants Data in Sub Namaelo, RT. 001, Masohi City)

T. L. WASILAINÉ¹, M. W. TALAKUA², Y. A. LESNUSSA³

¹Alumni Jurusan Matematika Fakultas MIPA Universitas Pattimura

^{2,3}Staf Jurusan Matematika Fakultas MIPA Universitas Pattimura

Jl. Ir. M. Putuhena, Kampus Unpatti, Poka-Ambon

E-mail: ²ocat_talakua@yahoo.com; ³yopi_a_lesnussa@yahoo.com

ABSTRAK

Model Regresi Linier Berganda merupakan sebuah model yang digunakan untuk menganalisis hubungan antar variabel. Hubungan tersebut dapat diekspresikan dalam bentuk persamaan yang menghubungkan variabel terikat (Y) dengan beberapa variabel bebas (X). Jika adanya hubungan linier yang sempurna atau pasti diantara beberapa atau semua variabel bebas dari model Regresi Berganda disebut Multikolinieritas. Jika korelasi antara dua atau lebih variabel bebas dalam suatu persamaan regresi linier berganda ini terjadi maka taksiran koefisien dari variabel yang bersangkutan tidak lagi tunggal melainkan tidak terhingga banyaknya sehingga tidak mungkin lagi menduganya. Dalam kasus ini peneliti akan melihat hubungan antara variabel-variabelnya. Apabila terdapat hubungan antara variabel-variabel bebasnya. Maka akan diterapkan metode Regresi Ridge untuk menstabilkan nilai koefisien regresi karena adanya Multikolinieritas. Regresi Ridge merupakan metode estimasi koefisien regresi yang diperoleh melalui penambahan konstanta bias c pada diagonal $X^T X$. Sehingga diperoleh persamaan regresi linier yang baru dan tidak mengandung multikolinieritas.

Kata kunci: *Regresi Linier Berganda, Multikolinieritas, Regresi Ridge*

PENDAHULUAN

Analisis regresi adalah suatu analisis yang dilakukan terhadap dua variabel yaitu variabel independen (prediktor) dan variabel dependen (respon) untuk mengetahui apakah ada pengaruh variabel prediktor terhadap variabel respon sehingga variabel respon dapat diduga berdasarkan variabel prediktornya. Berdasarkan jumlah variabel independennya, analisis regresi linier dibagi menjadi dua macam yaitu, analisis regresi linier sederhana dan analisis regresi linier ganda. Pada, analisis regresi linier sederhana, jumlah variabel independen yang digunakan sebagai penduga variabel dependen adalah satu. Sedangkan pada analisis regresi linier ganda, jumlah variabel independen yang digunakan sebagai penduga variabel dependen adalah lebih dari satu.

Saat ini, analisis regresi yang lebih sering digunakan adalah analisis regresi linier ganda. Dapat dilihat dari

berbagai kejadian yang terjadi dalam kehidupan sehari-hari yaitu suatu peristiwa dapat disebabkan oleh berbagai faktor yang mempengaruhinya. Contohnya, tinggi bayi sekarang yang dipengaruhi oleh usia bayi, panjang bayi waktu lahir, berat bayi waktu lahir, dan ukuran dada bayi waktu lahir.

Regresi sendiri memiliki sejarah dan terus dikembangkan. Serangkaian studi yang dikembangkan oleh para statistisi berhasil menunjukkan bahwa dalam banyak kasus, regresi menggunakan teknik kuadrat terkecil (khususnya teknik *Ordinary Least Square/OLS*) sering memberikan hasil yang kurang tepat. Oleh sebab itu, banyak bermunculan teknik “Regresi Modern”, yaitu regresi yang digunakan pada kondisi dimana asumsi-asumsi klasik tidak terpenuhi.

Ada beberapa asumsi klasik salah satunya yaitu tidak terjadi multikolinieritas. Multikolinieritas dapat

menyebabkan analisis regresi menggunakan metode kuadrat terkecil kurang valid.

Cara menangani penyimpangan asumsi ini adalah dengan menggunakan analisis Regresi Ridge. Metode Regresi Ridge merupakan penyelesaian yang paling baik, karena mengingat tujuan Regresi Ridge adalah untuk memperkecil variansi estimator koefisien regresi. Berdasarkan latar belakang inilah peneliti memberi judul “Model Regresi Ridge untuk Mengatasi Model Regresi Linier Berganda yang Mengandung Multikolinieritas (Studi Kasus : Data Pertumbuhan Bayi pada Kelurahan Namaelo RT 001, Kota Masohi)”.

Tujuan dari penelitian ini adalah bagaimana menggunakan Regresi Ridge untuk mengatasi masalah Multikolinieritas dalam variabel-variabel bebas sehingga diperoleh persamaan regresi linier berganda yang lebih baik pada Data Pertumbuhan Bayi pada Kelurahan Namaelo RT 001, Kota Masohi.

TINJAUAN PUSTAKA

Regresi Ridge adalah suatu teknik yang dikembangkan untuk menstabilkan nilai koefisien regresi karena adanya Multikolinieritas. Metode Regresi Ridge pertama kali dikemukakan oleh A. E. Hoerl pada tahun 1962. Metode ini ditujukan untuk mengatasi kondisi buruk yang diakibatkan oleh korelasi yang tinggi antara beberapa peubah bebas didalam model regresi, sehingga menyebabkan matriks XX^T nya hampir singular yang pada gilirannya menghasilkan nilai dugaan parameter model regresi yang tidak stabil (Draper dan Smith, 1981).

Regresi Ridge merupakan modifikasi dari Metode Kuadrat Terkecil yang menghasilkan penduga bias dari koefisien regresi (Kutner, et al, 2005). Regresi Ridge mengurangi dampak Multikolinieritas dengan menentukan penduga yang bias tetapi mempunyai variansi yang lebih kecil dari variansi penduga Regresi Linear Berganda.

Istilah Multikolinieritas pertama kali diperkenalkan oleh Ragnar Frisch pada tahun 1934, yang menyatakan Multikolinieritas terjadi jika adanya hubungan linier yang sempurna atau pasti diantara beberapa atau semua peubah bebas dari model Regresi Berganda (Rahardiantoro, 2008). Jadi dalam Regresi Linier Berganda Multikolinieritas merupakan masalah yang sering terjadi dua atau lebih variabel bebas saling berkorelasi satu sama lain. Jika korelasi antara dua atau lebih variabel bebas dalam suatu persamaan regresi linier berganda ini terjadi maka taksiran koefisien dari peubah yang bersangkutan tidak lagi tunggal melainkan tidak terhingga banyaknya sehingga tidak mungkin lagi menduganya, hal ini disebabkan XX^T (Sembiring, 1995). Menurut Montgomery dan Hines (1990) dalam Rahardiantoro (2008), Multikolinieritas dapat menyebabkan koefisien Regresi yang dihasilkan oleh Analisis Regresi berganda menjadi sangat lemah atau tidak dapat menghasilkan hasil analisis yang mewakili sifat atau pengaruh dari peubah bebas yang bersangkutan. Hal ini disebabkan oleh penduga yang dihasilkan oleh metode kuadrat terkecil tidak lagi bersifat BLUE (*Best Linear Unbiased Estimator*).

Multikolinieritas

Multikolinieritas yakni situasi dimana terdapat korelasi atau hubungan linier antara variabel-variabel bebas diantara satu dengan yang lainnya sehingga variabel-variabel bebas tersebut tidak bersifat orthogonal. Variabel-variabel bebas yang bersifat orthogonal adalah variabel bebas yang memiliki nilai korelasi diantara sesamanya sama dengan nol.

Penyebab dan Akibat Multikolinieritas

Hal-hal utama yang sering menyebabkan terjadinya Multikolinieritas pada model regresi antara lain :

- Kesalahan teoritis dalam pembentukan model fungsi regresi yang dipergunakan.
- Terlampau kecilnya jumlah pengamatan yang akan dianalisis dengan model regresi.

Jika terjadi multikolinieritas sempurna maka koefisien regresi dari variabel X tidak dapat ditentukan dan *standard error*nya tak berhingga. Jika terjadi Multikolinieritas kurang sempurna maka akan timbul akibat sebagai berikut:

- Meskipun koefisien regresi dari variabel X dapat ditentukan tetapi nilai *standard error*nya akan cenderung membesar sehingga tingkat kolinieritas antara variabel bebas juga meningkat.
- Nilai *standard error* dari koefisien regresi besar maka interval keyakinan untuk parameter dari populasi juga cenderung melebar.
- Tingginya tingkat kolinieritas, probabilitas untuk menerima hipotesis, padahal hipotesis itu salah menjadi membesar nilainya.
- Bila multikolinieritas tinggi, akan memperoleh nilai R^2 yang tinggi tetapi tidak ada atau sedikit sekali koefisien regresi yang signifikan secara statistik.

Pendeteksian Multikolinieritas

Ada beberapa cara untuk mengetahui ada tidaknya multikolinieritas

Diantaranya adalah :

- Faktor Variansi Inflasi
Adalah merupakan elemen diagonal utama dari invers matriks korelasi. Faktor variansi inflasi yang kecil maka multikolinieritas lebih sederhana. Faktor inflasi yang melebihi 10 maka multikolinieritas dikatakan ada.
- Nilai Determinan
Nilai determinan terletak antara 0 dan 1. Bila nilai determinan satu, kolom matriks X adalah orthogonal dan bila nilainya 0 disanaada sebuah ketergantungan linier yang nyata antara kolom X. nilai yang lebih kecil determinannya aka tingkat multikolinieritasnya lebih besar.
- Kadang-kadang pemeriksaan masing-masing elemen matriks korelasi dapat menolong dalam mendapatkan multikolinieritas. Jika elemen r_{ij} mendekati satu, maka X_i dan X_j mungkin benar-benar ada masalah multikolinieritas.
Karena bila lebih dari dua variabel bebas yang dicakup dalam sebuah multikolinieritas tidak selalu memungkinkan kita untuk mendapatkan keberadaan multikolinieritas.

Regresi Ridge

Regresi Ridge dapat digunakan untuk mengatasi korelasi yang tinggi antara beberapa variabel bebas. Multikolinieritas yang terdapat dalam regresi linier berganda yang mengakibatkan matriks $X^T X$ -nya hampir singular yang pada gilirannya menghasilkan nilai estimasi parameter yang tidak stabil. Regresi Ridge merupakan metode estimasi koefisien regresi yang diperoleh melalui penambahan konstanta bias c pada diagonal $X^T X$. Nilai c untuk koefisien regresi ridge diantara 0 hingga 1.

Dalam bentuknya yang sederhana adalah sebagai berikut :

$$\hat{\beta}(c) = (X^T X + cI)^{-1} X^T Y$$

Umumnya sifat dari penafsiran ridge ini memiliki variansi yang minimum sehingga diperoleh nilai VIF-nya yang merupakan diagonal utama dari matriks

$$(X^T X + cI)^{-1} X^T X (X^T X + cI)^{-1}$$

Pada dasarnya Regresi Ridge merupakan metode kuadrat terkecil. Perbedaannya adalah bahwa pada metode regresi ridge, nilai variabel bebasnya ditransformasikan dahulu melalui prosedur centering dan rescaling.

Ridge Trace

Ridge Trace adalah plot dari estimator regresi ridge secara bersama dengan berbagai kemungkinan tetapan bias c , konstanta c mencerminkan jumlah bias dalam estimator $\hat{\beta}(c)$. Bila $c = 0$ maka estimator $\hat{\beta}(c)$ sama dengan kuadrat terkecil β , tetapi cenderung lebih stabil.

Suatu acuan yang biasa digunakan untuk memilih besarnya c adalah dengan melihat VIF dan melihat pola kecenderungan Ridge Trace. Bila terdapat korelasi yang tinggi antara variabel bebas, maka nilai VIF akan besar. VIF memiliki nilai mendekati 1 jika variabel bebas X tidak saling berkorelasi dengan variabel bebas lainnya.

Pengujian Hipotesis

Pengujian untuk uji keberartian regresi sebagai berikut :

1. Menentukan uji Hipotesis

$H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_k = 0$ (tidak ada hubungan linier antara variabel-variabel bebas dengan variabel terikat)

$H_1: \exists \beta_j \neq 0, j = 1, 2, 3, \dots, k$ (ada hubungan linear antara variabel-variabel bebas dengan variabel terikat)

2. Menentukan tingkat signifikansi (α)

3. Kriteria pengambilan keputusan untuk F_{hitung} .

Terima H_0 jika $F_{hitung} \leq F_{tabel}$ dan Tolak H_0 jika $F_{hitung} > F_{tabel}$.

$$\text{Dengan } F_{hitung} = \frac{JKR/k}{JKR/(n-k-1)}$$

Untuk mengetahui koefisien yang diperoleh berarti atau tidak. Dilakukan pengujian sebagai berikut :

Hipotesis

$H_0: \beta_i = 0$ (koefisien regresi tidak signifikan)

$H_1: \beta_i \neq 0$ (koefisien regresi signifikan)

1. Menentukan tingkat signifikansi (α)

2. Uji Statistik yang digunakan, dalam hal ini uji statistik yang digunakan adalah uji-t

$$t = \frac{b_i}{s_{b_i}}$$

b_i = Koefisien regresi variabel $-i$

s_{b_i} = Standar error variabel $-i$

3. Kriteria Keputusan

H_0 diterima jika

$$|t_{hitung}| \leq t_{(\frac{\alpha}{2}, n-k-1)}$$

H_0 ditolak jika

$$|t_{hitung}| > t_{(\frac{\alpha}{2}, n-k-1)}$$

HASIL DAN PEMBAHASAN

Pada bagian pembahasan dibahas cara mendeteksi multikolinieritas pada data, penerapan regresi ridge dalam menyelesaikan multikolinieritas pada data, dan pengujian keberartian regresi pada data pertumbuhan bayi.

Data Pertumbuhan Bayi

Bayi merupakan anak-anak kecil yang berusia 0-1 tahun. Pertumbuhan bayi pun menjadi pusat perhatian para orang tua. Pertumbuhan seorang bayi dipengaruhi oleh berbagai macam aspek. Tinggi seorang bayi pun menjadi perhatian yang sangat penting bagi orang tua. Tinggi bayi juga dipengaruhi oleh beberapa faktor-faktor misalnya berat bayi waktu lahir, usia bayi dan lain-lain. Dan yang akan peneliti teliti yaitu Tinggi bayi sekarang (Y) yang dipengaruhi oleh Usia Bayi (X1), Tinggi bayi waktu lahir (X2), Berat bayi waktu lahir (X3), dan Ukuran dada bayi waktu lahir. Untuk mendapatkan Data Pertumbuhan Bayi dengan faktor-faktor yang mempengaruhi seperti yang telah dipaparkan peneliti. Peneliti melakukan Penelitian pada Posyandu Binaya. Posyandu Binaya merupakan salah satu Posyandu yang bertugas untuk melihat perkembangan Bayi pada kota Masohi terkhusus pada Kelurahan Namaelo RT 001. Data yang peneliti peroleh dapat terlihat pada Tabel 1.

Tabel 1. Data Pertumbuhan Bayi

No	Y	X1	X2	X3	X4
1	62.0	216.0	51.0	2.7	19.0
2	50.0	104.0	50.0	3.1	16.0
3	61.0	106.0	51.0	2.4	21.0
4	60.0	147.0	52.0	2.0	26.0
5	60.0	71.0	52.0	2.3	23.0
6	63.0	144.0	49.0	2.2	22.0
7	58.0	73.0	50.0	2.3	22.0
8	63.0	235.0	50.0	2.1	24.0
9	55.0	23.0	51.0	3.2	16.0
10	56.0	92.0	50.0	2.9	17.0
11	56.0	53.0	51.0	3.0	17.0
12	67.0	85.0	52.0	2.6	20.0
13	59.0	100.0	48.0	2.4	20.0
14	55.0	80.0	48.0	3.0	16.0
15	55.0	80.0	48.0	3.0	16.0
16	64.0	133.0	51.0	2.4	21.0
17	60.0	140.0	52.0	2.6	20.0
18	62.0	141.0	51.0	2.3	22.0
19	68.0	233.0	50.0	2.3	22.0
20	60.0	167.0	48.0	2.3	21.0

Keterangan :

Y =Tinggi bayi sekarang (cm)

X₁ =Usia Bayi (hari)

X₂ =Tinggi bayi waktu lahir (cm)

X_3 =Berat bayi waktu lahir (kg)
 X_4 =Ukuran dada bayi waktu lahir (cm)

Regresi Linier Berganda

Akan dibentuk suatu model regresi linier berganda berdasarkan Tabel 1. Analisa regresi dengan menggunakan SPSS terhadap data menghasilkan nilai estimator parameter (Tabel 2)

Tabel 2. Coefficients^a

Model	Unstandardized Coefficients	
	B	Std. Error
1 (Constant)	73.715	33.961
X1	.033	.012
X2	1.839	.669
X3	-22.595	10.368
X4	-2.602	1.330

a. Dependent Variable: Y

Berdasarkan Tabel 2 diperoleh persamaan regresi linier berganda yaitu

$$Y = 73,715 + 0,033X_1 + 1,839X_2 - 22,595X_3 - 2,602X_4$$

Pada persamaan diatas terlihat bahwa koefisien regresi dari setiap variabel bebas memiliki nilai yang positif dan negatif. Sehingga dapat diartikan bahwa variabel bebas X_1, X_2 yang memiliki koefisien regresi yang positif menunjukkan bahwa antara X_1, X_2 dengan variabel terikat Y berjalan satu arah, dimana setiap penurunan atau peningkatan variabel bebas X_1, X_2 akan diikuti dengan penurunan atau peningkatan variabel terikat Y. Sedangkan variabel bebas X_3, X_4 yang memiliki koefisien regresi negatif menunjukkan bahwa antara X_3, X_4 berjalan dua arah dimana setiap peningkatan pada variabel bebas X_3, X_4 akan diikuti dengan penurunan variabel terikatnya dan setiap penurunan pada variabel bebas X_3, X_4 akan diikuti dengan peningkatan variabel terikat Y.

Pendeteksian Multikolinieritas Nilai VIF dan Tolerance

Mendeteksi multikolinieritas dengan menggunakan nilai VIF dan nilai tolerance dapat disimpulkan terdapat multikolinieritas apabila nilai VIF dari variabel-variabel bebasnya lebih dari 10 dan nilai tolerance dari variabel-variabel bebasnya mendekati 0.

Tabel 3. Coefficients^a

Model	Collinearity Statistics	
	Tolerance	VIF
1 (Constant)		
X1	.664	1.507
X2	.346	2.894
X3	.022	45.831
X4	.020	49.386

a. Dependent Variable: Y

Dari Tabel 3 terlihat bahwa nilai VIF dari $X_3 = 45,831$ dan $X_4 = 49,386$. karena nilai VIF dari X_3 dan X_4 lebih dari 10 dan juga nilai tolerance dari $X_3 = 0,022$ dan $X_4 = 0,020$ karena nilai tolerance mendekati 0 maka dapat disimpulkan bahwa adanya multikolinieritas antara variabel bebasnya.

Determinan Korelasi

Mendeteksi multikolinieritas dengan menggunakan determinan korelasi dapat disimpulkan terdapat multikolinieritas apabila nilai determinan korelasinya mendekati 0. Untuk memperoleh koefisien korelasi antara variabel-variabel bebasnya peneliti menggunakan SPSS 16.0. Hasilnya dapat terlihat pada matriks korelasi berikut

$$C = \begin{bmatrix} 1 & -0,334 & -0,294 & 0,237 \\ -0,334 & 1 & 0,062 & 0,186 \\ -0,294 & 0,062 & 1 & -0,956 \\ 0,237 & 0,186 & -0,956 & 1 \end{bmatrix}$$

Dari matriks korelasi tersebut akan diperoleh determinannya yaitu $|C| = 0,019674$

Karena nilai determinan dari matriks korelasinya mendekati 0 maka dapat disimpulkan bahwa terdapat korelasi antara variabel-variabel bebasnya.

Pemodelan Regresi Ridge

Software NCSS merupakan salah satu program matematika yang dipakai untuk menyelesaikan beberapa masalah yang terjadi dalam bidang matematika terkhususnya dalam bidang statistika. NCSS diciptakan pada tahun 1981.NCSS LLC merupakan perusahaan yang menciptakan NCSS. NCSS LLC khusus menyediakan software analisis statistika untuk penelitian, bisnis, dan lembaga akademik.

NCSS (Number Cruncher Statistical System) memasukkan lebih dari 230 dokumen statistik dan menggambarkan caranya. NCSS dapat menyelesaikan berbagai macam masalah Statistika dan salah satunya yaitu masalah multikolinieritas dengan menggunakan Metode Regresi Ridge.

Untuk menghilangkan kondisi buruk yang tidak menguntungkan yang diakibatkan oleh adanya Multikolinieritas dalam data yang dianalisis dan juga untuk memudahkan kerja dalam proses pendeteksian maupun dalam penanganan multikolinieritas maka dilakukan proses pemusatan dan penskalaan terhadap data atau variabel.

Tabel 4. One-Sample Kolmogorov-Smirnov Test

		Unstandardized Residual
N		20
Normal Parameters ^a	Mean	.0000000
	Std. Dev	2.14694314
Most Extreme Differences	Absolute	.175
	Positive	.175
	Negative	-.119
Kolmogorov-Smirnov Z		.782
Asymp. Sig. (2-tailed)		.574

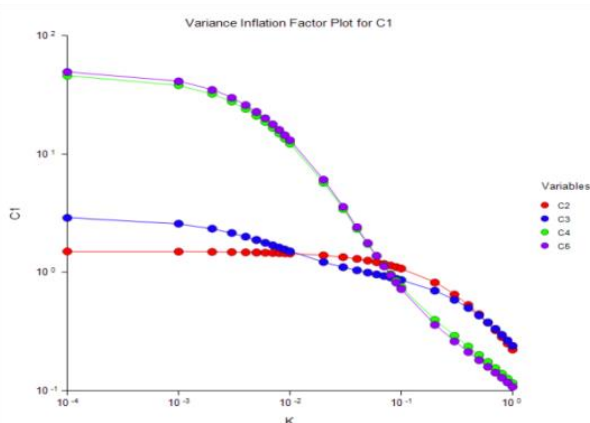
Sebelumnya akan dilakukan pengujian terhadap data yang ada. Apakah data berdistribusi normal. Dalam penelitian ini, peneliti menggunakan pengujian Kolmogorov-Smirnov.

Peneliti menggunakan SPSS 16.0 untuk mendapatkan hasilnya. Hasil yang diperoleh dapat terlihat pada Tabel 4

Dari Tabel 4 dapat terlihat bahwa nilai Kolmogorov-Smirnov Z sebesar 0,782 dan Asymp Sig sebesar 0,574 lebih besar dari Nilai signifikan yang dipilih yaitu 0,05 maka dapat disimpulkan bahwa data berdistribusi normal. Karena data yang diperoleh berdistribusi normal maka tidak perlu dilakukan pentransformasian terhadap data.

Tabel 5. Nilai VIF $\hat{\beta}(c)$ dengan berbagai nilai c

Nilai c	VIF $\hat{\beta}_1(c)$	VIF $\hat{\beta}_2(c)$	VIF $\hat{\beta}_3(c)$	VIF $\hat{\beta}_4(c)$
0,000	1,5067	2,894	45,831	49,386
0,001	1,4995	2,584	38,213	41,156
0,002	1,4927	2,345	32,363	34,836
0,003	1,4862	2,157	27,774	29,879
0,004	1,4800	2,006	24,108	25,918
0,005	1,4740	1,884	21,133	22,704
0,006	1,4682	1,782	18,685	20,059
0,007	1,4625	1,697	16,647	17,858
0,008	1,4569	1,626	14,932	16,006
0,009	1,4514	1,564	13,475	14,432
0,01	1,4460	1,511	12,227	13,085
0,02	1,3949	1,225	5,737	6,077
0,03	1,3474	1,109	3,424	3,582
0,04	1,3028	1,043	2,339	2,415
0,05	1,2606	0,998	1,743	1,774
0,06	1,2206	0,963	1,378	1,3836
0,07	1,1826	0,934	1,138	1,1275
0,08	1,1464	0,908	0,970	0,9498
0,09	1,1121	0,885	0,848	0,8209
0,1	1,0794	0,864	0,756	0,7240
0,2	0,8223	0,702	0,398	0,3607
0,3	0,6515	0,588	0,292	0,2618
0,4	0,5315	0,501	0,237	0,2127
0,5	0,4436	0,4334	0,2012	0,1817
0,6	0,3771	0,3789	0,1755	0,1595
0,7	0,3253	0,3344	0,1557	0,1425
0,8	0,2841	0,2974	0,1400	0,1289
0,9	0,2507	0,2664	0,1271	0,1177
1,0	0,2232	0,2401	0,1162	0,1082



Gambar 1. VIF plot

Dalam proses pengestimasi regresi ridge, pemilihan tetapan bias c merupakan hal yang paling penting dalam penelitian ini, penentuan tetapan bias c ditempuh melalui pendekatan nilai VIF dan gambar Ridge Trace. Nilai dari VIF $\hat{\beta}(c)$ dengan berbagai kemungkinan tetapan bias c dapat dilihat pada Tabel 5.

Tabel 5 menunjukkan bahwa $0,000 \leq c \leq 1,000$ dengan berbagai nilai VIF-nya. Dari Tabel 5 juga dapat terlihat bahwa semakin besar nilai c , semakin kecil nilai VIF koefisien estimator $\hat{\beta}(c)$. Nilai VIF dapat diplot pada Gambar 1.

Berdasarkan Gambar 1 terlihat bahwa setiap nilai VIF semakin mengalami penurunan. VIF dari setiap koefisien $\hat{\beta}(c)$ mengalami penurunan menuju 0. Pada saat $c = 0,2$ nilai VIF mendekati 1 dari 0. Nilai VIF $\hat{\beta}_1(c) = 0,8223$, VIF $\hat{\beta}_2(c) = 0,7015$, VIF $\hat{\beta}_3(c) = 0,3977$, VIF $\hat{\beta}_4(c) = 0,3607$.

Sedangkan nilai koefisien estimator parameter $\hat{\beta}(c)$ dengan berbagai kemungkinan tetapan bias c dapat dilihat pada Tabel 6 berikut :

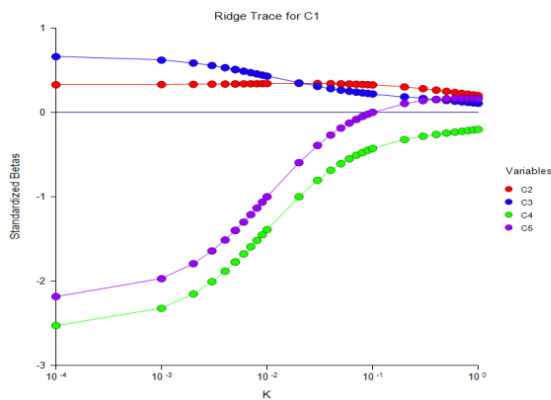
Tabel 6. Nilai $\hat{\beta}(c)$ dengan berbagai nilai c

Nilai (c)	$\hat{\beta}_1(c)$	$\hat{\beta}_2(c)$	$\hat{\beta}_3(c)$	$\hat{\beta}_4(c)$
0.0000	0.5069	0.6914	-2.181	-2.0323
0.0010	0.5090	0.6542	-1.998	-1.8418
0.0020	0.5107	0.6229	-1.845	-1.6820
0.0030	0.5120	0.5963	-1.714	-1.5460
0.0040	0.5130	0.5733	-1.602	-1.4290
0.0050	0.5137	0.5533	-1.505	-1.3271
0.0060	0.5143	0.5356	-1.419	-1.2376
0.0070	0.5148	0.5199	-1.343	-1.1584
0.0080	0.5151	0.5059	-1.275	-1.0878
0.0090	0.5153	0.4934	-1.215	-1.0245
0.0100	0.5154	0.4820	-1.160	-0.9674
0.0200	0.5136	0.4080	-0.813	-0.6023
0.0300	0.5093	0.3690	-0.638	-0.4173
0.0400	0.5042	0.3442	-0.534	-0.3053
0.0500	0.4987	0.3266	-0.464	-0.2300
0.0600	0.4931	0.3132	-0.415	-0.1758
0.0700	0.4875	0.3024	-0.379	-0.1349
0.0800	0.4820	0.2934	-0.350	-0.1029
0.0900	0.4766	0.2857	-0.328	-0.0771
0.1000	0.4713	0.2790	-0.310	-0.0558
0.2000	0.4246	0.2360	-0.227	0.0477
0.3000	0.3878	0.2104	-0.201	0.0854
0.4000	0.3581	0.1917	-0.187	0.1041
0.5000	0.3335	0.1770	-0.179	0.1145
0.6000	0.3127	0.1649	-0.172	0.1206
0.7000	0.2948	0.1546	-0.167	0.1240
0.8000	0.2793	0.1458	-0.163	0.1257
0.9000	0.2655	0.1381	-0.158	0.1264
1.0000	0.2533	0.1313	-0.154	0.1263

Tabel 6 menunjukkan berbagai koefisien estimator parameter $\hat{\beta}(c)$ dengan berbagai nilai c . Berdasarkan Tabel 6 dapat terlihat bahwa semakin besar nilai c , koefisien regresi ridge dari variabel bebas X_1, X_2 semakin kecil sedangkan koefisien regresi ridge dari variabel bebas X_3, X_4 . Dari Tabel 6 dapat membuat sebuah Ridge Trace seperti pada Gambar 2

Dari Gambar 2 terlihat bahwa koefisien β lebih stabil pada saat $c = 0,2$. Dengan demikian persamaan Regresi Ridge yang dihasilkan dari $c = 0,2$ yaitu

$$Y^* = 0,4246 X_1^* + 0,2360X_2^* - 0,2274X_3^* + 0,0477 X_4^*$$



Gambar 2. Ridge Trace

Uji Keberartian Regresi

Persamaan regresi yang diperoleh yaitu:

$$Y^* = 0,4246 X_1^* + 0,2360X_2^* - 0,2274X_3^* + 0,0477 X_4^*$$

Akan diuji keberartian dari model tersebut

1. Hipotesis
 $H_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$
 (tidak ada hubungan antara variabel-variabel bebas dengan variabel terikat)
 $H_0: \exists \beta_j \neq 0, j = 1,2,3,4$
 (ada hubungan linier antara variabel-variabel bebas dengan variabel terikat)
2. Taraf signifikansi $\alpha = 0,05$
3. Nilai F_{hitung} dapat dilihat pada Tabel 7

Tabel 7. Anava Ridge

S. Varians	JK	DK	KT	F_{hitung}	F_{tabel}
Regresi	0,687	4	0,1718	8,23	3,056
Sis	0,313	15	0,0209		
Total	1	19			

4. Dari Tabel 7 terlihat bahwa $F_{hitung} = 8,23 > F_{tabel} = 3,06$ maka tolak H_0 sehingga ada hubungan antara variabel-variabel bebas dengan variabel terikat.

Untuk mengetahui koefisien yang diperoleh berarti atau tidak dilakukan pengujian sebagai berikut :

1. Hipotesis
 $H_0: \beta_i = 0$ (koefisien regresi tidak signifikan)
 $H_1: \beta_i \neq 0$ (koefisien regresi signifikan)
2. Tingkat signifikansi $\alpha = 0,05$
3. Uji-t

$$t = \frac{b_i}{s_{b_i}}$$
 $b_i =$ Koefisien regresi variabel $-i$
 $s_{b_i} =$ Standar error variabel $-i$

Dari Tabel 8 dapat disimpulkan bahwa variabel X_1 berpengaruh secara signifikan terhadap variabel terikat Y , sedangkan variabel X_2, X_3, X_4 tidak berpengaruh secara

signifikan terhadap variabel Y . Proses Pengembalian X^* ke X bentuk semula dengan $\bar{Y} = 60,11, \bar{X}_1 = 121,15, \bar{X}_2 = 50,25, \bar{X}_3 = 2,55, \bar{X}_4 = 20,05 S_Y = 3,751347,$

$$S_{X_1} = 58,16336, S_{X_2} = 1,409554, S_{X_3} = 0,361976, S_{X_4} = 5,448563$$

Sehingga persamaan regresinya menjadi :

$$\hat{Y} = 30,12984 + 0,0274X_1 + 0,628X_2 - 2,3551X_3 + 0,061X_4$$

Tabel 8. Uji Keberartian Koefisien

	b_i	s_{b_i}	$ t_{hitung} $	t_{tabel}	Kesimpulan
1	0,425	0,011	40,46	2,13	signifikan
2	0,2360	0,399	0,59		Tidak signifikan
3	-0,227	1,173	0,19		Tidak signifikan
4	0,048	0,138	0,346		Tidak signifikan

Dari persamaan regresi diatas dapat diartikan bahwa variabel bebas X_1, X_2, X_4 yang memiliki koefisien regresi yang positif menunjukkan bahwa antara X_1, X_2, X_4 dengan variabel terikat Y berjalan satu arah, dimana setiap penurunan atau peningkatan variabel bebas X_1, X_2, X_4 akan diikuti dengan penurunan atau peningkatan variabel terikat Y . Sedangkan variabel bebas X_3 yang memiliki koefisien regresi negatif menunjukkan bahwa antara X_3 berjalan dua arah dimana setiap peningkatan pada variabel bebas X_3 akan diikuti dengan penurunan variabel terikatnya dan setiap penurunan pada variabel bebas X_3 akan diikuti dengan peningkatan variabel terikat Y .

KESIMPULAN

Dari pembahasan diatas dapat disimpulkan bahwa :

1. Persamaan regresi yang diperoleh dengan menggunakan kuadrat terkecil mengandung multikolinieritas. Sehingga diterapkan Metode Regresi Ridge untuk memperoleh persamaan regresi yang baru dan tidak mengandung multikolinieritas. dan persamaan regresi yang baru yaitu

$$\hat{Y} = 30,12984 + 0,0274X_1 + 0,628X_2 - 2,3551X_3 + 0,061X_4$$
2. Berdasarkan data bayi yang peneliti teliti, usia bayi sangat berpengaruh secara signifikan terhadap tinggi bayi sekarang, sedangkan tinggi bayi waktu lahir, berat bayi waktu lahir dan ukuran dada bayi waktu lahir tidak berpengaruh secara signifikan terhadap tinggi bayi sekarang.

DAFTAR PUSTAKA

[1] Agriska Prenadita Putri. 2011. *Penggunaan Metode Ridge Trace dan Variance Inflation Factor*.

- [2] Draper N dan Smith H. 1981. *Analisis Regresi Terapan Edisi 2 (Terjemahan Cmbang-Sumantri)*. Jakarta: PT. Gramedia Pustaka Utama
- [3] Kariyam. 2000. *Studi Penanganan Kasus Multikolinieritas dengan Pendekatan Analisis Regresi Faktor*. ISSN: 1410-2315
- [4] Pradipta, Nanan. 2009. *Model Regresi Ridge untuk Mengatasi Model Regresi Linier Berganda yang Mengandung Multikolinieritas*. USU Repository
- [5] Walpole, RE dan Myers, R H. 1995. *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan*. Penerbit ITB

