

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis regresi adalah salah satu diantara analisis yang digunakan dalam menganalisa hubungan sebab akibat variabel independen (terikat) dan variabel dependen (bebas) adalah analisis regresi. Berdasarkan jumlah variabel dependennya, analisis regresi terbagi atas dua macam yaitu : analisis regresi sederhana dan analisis regresi berganda yang dapat diperoleh dengan melakukan estimasi terhadap parameter-parameternya menggunakan metode kuadrat terkecil (ordinary least square). Dalam ilmu statistik sebuah model dikatakan cocok atau apabila garis regresi harus memenuhi sebuah asumsi klasik antara data tidak mengalami heteroskedastisitas, multikolinieritas dan autokorelasi.

Permasalahan yang sering terjadi pada multikolinieritas adalah jika uji T tidak signifikan walaupun variabel dependennya diregresikan secara terpisah dengan variabel dependennya (simple regression) uji menunjukkan hasil yang signifikan. Hal tersebut menandakan pada regresi bertentangan maka dapat diartikan pada masing-masing regresi linier berganda untuk masing-masing peubah seringkali tidak sesuai (Montgomery, 1990). Selain itu korelasi yang cukup tinggi antar variabel prediktor (Astuti, 2014). Multikolinieritas mengakibatkan determinan $X^t X$ mendekati nol sehingga menyebabkan matriks hampir singular yang berakibat nilai penduga parameter bisa mempunyai tanda

yang salah atau lebih besar dari perkiraan. Oleh karena itu masalah multikolinieritas harus dihindari.

Multikolinieritas dalam model regresi linier dapat dideteksi dengan beberapa cara, diantaranya dengan menghitung nilai *Variance Inflation Factor* (VIF) dan nilai *Tolerance* (TOL). Jika terdapat pelanggaran asumsi multikolinieritas, terdapat prosedur yang dapat digunakan untuk mengatasinya, seperti menabahkan data, menghilangkan satu atau beberapa variabel prediktor yang memiliki korelasi tinggi dari model regresi dan menggunakan metode analisis yang lain seperti Regresi ridge (Ghozali, 2013). Regresi Ridge diajukan sebagai suatu cara untuk mengatasi multikolinieritas. Keuntungan Regresi Ridge dibanding dengan metode lain yaitu regresi Ridge berdampak mengurangi multikolinieritas dengan menentukan penduga yang bias tetapi mempunyai varians yang lebih kecil dari varians penduga regresi linier berganda (Pratiei, 2016). Metode regresi Ridge diperoleh dengan cara yang sama seperti metode kuadrat terkecil, yaitu dengan meminimumkan jumlah kuadrat sisaan. Regresi Ridge menambahkan kendala (tetapan bias) pada kuadrat terkecil sehingga koefisien ridge berkurang dan mendekati nol (Hastie, Robert & Jerome, 2008). Penggunaan regression Ridge adalah komponen yang utama dapat dihilangkan korelasi secara bersih ($korelasi=0$) tanpa menghilangkan variabel bebas sehingga masalah multikolinieritas dapat teratasi. Selain itu metode komponen utama dapat digunakan untuk semua jenis data baik musiman atau non musiman dengan cara memodifikasi metode kuadrat terkecil untuk mendapatkan pengurangan varian dengan cara menambahkan suatu ketetapan k dalam menstabilkan koefisien (Mardikya dan cetin, 2008).

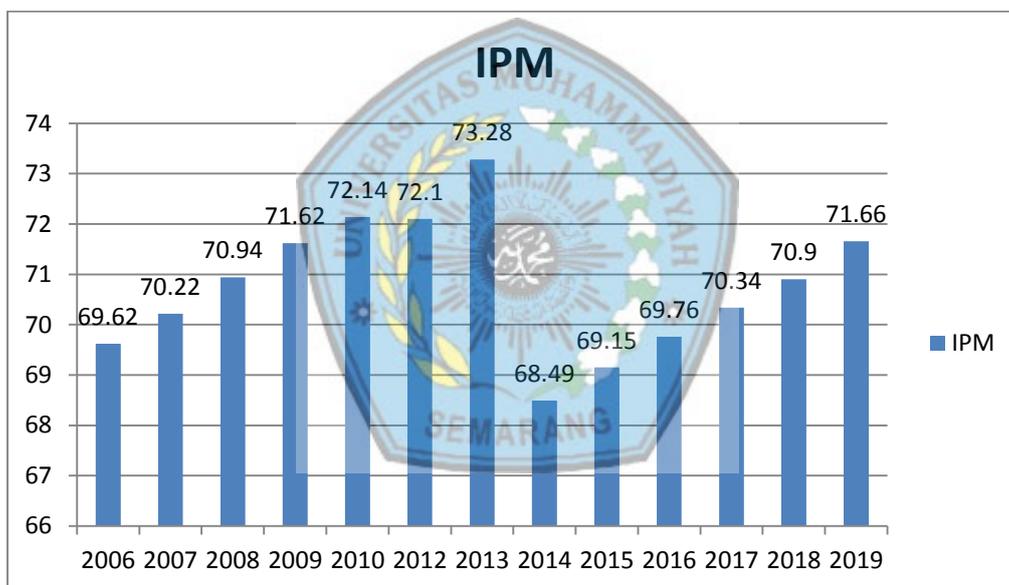
Metode lain yang digunakan dalam mengatasi multikolinieritas adalah Principal Component Regression merupakan salah satu teknik dalam mengatasi multikolinieritas

dengan cara mereduksi variabel-variabel yang ada menjadi variabel baru yang saling bebas dan merupakan kombinasi linier dari variabel asal (maitra & yan, 2008). Pricipal Component Regression adalah metode yang menggabungkan antara analisi regresi dengan etode pricipal component analisis (PCA). PCA bertujuan untuk menyederhanakan variabel yang diamati dengan cara mereduksi dimensinya tanpa kehilangan banyak informasi dari variabel asal. Prinsip utama PCA adalah adanya nilai korelasi antar variabel sehingga memungkinkan bahwa variabel-variabel itu dapat direduksi. Hal ini dilakukan dengan cara menghilangkan korelasi antar variabel prediktor baru yang tidak saling berkorelasi sama sekali atau yang bisa disebut dengan pricipal component (Johnson & Wichern, 2007). Dari sebuah komponen utama, dipilih buah komponen saja maka sudah mampu menerangkan keragaman data yang cukup tinggi (Johnson & Wichern, 1996, hal 336). Komponen utama yang tadi sudah dipilih buah dapat menggantikan buah variabel asalnya. Kelebihan Pricipal Component Regression adalah dapat menghilangkan korelasi, dapat digunakan untuk segala kondisi data dan dapat digunakan tanpa mengurangi variabel asal (Soemartini, 2008).

Penelitian terdahulu dilakukan oleh Hasriani (2014) perbandingan regresi ridge (regresi gulud) dan pricipal component analysis (PCA) dalam mengatasi multikolinieritas dengan hasil metode Regresi ridge (Regresi Gulud) lebih baik mengatasi multikolinieritas dibandingkan metode Pricipal Component Analysis, T.L. Wasilaine, M.W. Talakua dan Y.A Lessnussa (2014) tentang regresi ridge dalam mengatasi multikolinieritas pada data pertumbuhan bayi di Kelurahan Namaeleo RT 001, Kota Masohi dengan hasil bayi yang diteliti berpengaruh signifikan, Eko Supriyadi (2016) perbandingan Partial Least Square(PLS) dan Pricipal Component Regression (PCR) dalam mengatasi multikolinieritas pada faktor-faktor yang mempengaruhi kemiskinan di Jawa Tengah dengan hasil metode partial least square lebih baik mengatasi masalah multikolineirtas

dibanding metode Pricipal Component Regression dan Nur Hasana (2017) Metode Regresi Ridge dengan Iterasi HKB dalam mengatasi Multikolinieritas dengan hasil regresi ridge dengan iterasi HKB memberikan hasil yang lebih baik, sehingga metode regresi ridge dengan iterasi HKB dapat dijadikan alternatif dalam penggunaan untuk mengatasi multikolinieritas.

Berdasarkan hal tersebut, peneliti tertarik untuk menbandingkan metode Pricipal Component Regression dan Regression Ridege dalam mengatasi multikolinieritas. Kedua metode ini akan diterapkan pada kasus Indeks Pembangunan Manusia (IPM) di Provinsi Sulawesi Selatan.



Gambar 1.1 Grafik IPM

Berdasarkan data yang diperoleh dari Badan Pusat Statistik menunjukkan bahwa Indeks Pembangunan Manusia (Y) mengalami fluktuasi pada 2006 – 2019 dengan rata-rata peningkatan tiap tahunnya sebesar 4.79% dan angka tertinggi sebesar 73.28% yang artinya pemerintah membenahi namun pada tahun 2013 menurun drastis bahkan terendah selama 13 tahun terakhir dengan sebesar 68.49% pada tahun 2014 yang

disebabkan nilai digunakannya metode baru pada IPM yang diharapkan lebih mempresentasikan kondisi saat ini ditambah dengan kurangnya infrastruktur penunjang pendidikan. Dengan variabel yang mempengaruhi Indeks Pembangunan manusia (IPM) adalah harapan lama sekolah (HLS), rata-rata lama sekolah (RLS), angka harapan hidup saat lahir (AHH), angka partisipasi sekolah (APS), produk domestik regional bruto (PDRB), angka melek huruf (AMH), tingkat pengangguran terbuka (TPT), tingkat angka partisipasi angkatan kerja (TPAK), pengeluaran perkapita non makanan (PPK) dan status kepemilikan rumah milik sendiri (SKRMS).

IPM didefinisikan sebagai proses perluasan pilihan bagi penduduk (*enlarging people's choice*). IPM merupakan indikator penting dalam mengukur keberhasilan dalam upaya membangun kualitas hidup manusia (masyarakat atau penduduk). IPM menjelaskan bagaimana penduduk dapat mengakses hasil pembangunan dalam memperoleh pendapatan, kesehatan, pendidikan dan sebagainya. IPM diperkenalkan oleh United Nations Development Programme (UNDP) pada tahun 1990 dan metode perhitungan direvisi pada tahun 2010. BPS mengadopsi perubahan metodeologi perhitungan IPM yang baru pada tahun 2014 dan melakukan backcastin sejak 2010. IPM dibentuk oleh tiga dimensi dasar, yaitu umur panjang dan hidup sehat (*a long and healthy life*), pengetahuan (*knowledge*) dan standar hidup layak (*decent standar of living*). Umur panjang dan hidup sehat digambarkan oleh AHH yaitu jumlah tahun yang diharapkan dapat dicapai oleh bayi yang baru lahir untuk hidup dengan asumsi bahwa pola angka kematian menurut umur pada saat kelahiran sama panjang usia bayi. Pengetahuan diukur melalui indikator RLS dan HLS. RLS adalah rata-rata lamanya (tahun) penduduk usia 25 tahun keatas dalam menjalani pendidikan formal. HLS didefinisikan sebagai lamanya (tahun) sekolah formal yang diharapkan akan dirasakan oleh anak-anak pada umur tertentu dimasa mendatang. Standar hidup layak didefinisikan oleh pengeluaran perkapita disesuaikan dengan dari

nilai pengeluaran perkapita dan paritas daya beli. IPM dihitung berdasarkan rata-rata geometri indeks kesehatan, indeks pengetahuan dan indeks pengeluaran. Perhitungan ketiga indeks ini dilakukan dengan cara standarisasi dengan nilai minimum dan maksimum masing-masing indeks. IPM merupakan indikator yang digunakan untuk melihat perkembangan pembangunan dalam jangka panjang. Untuk melihat kemajuan pembangunan manusia, terdapat dua aspek yang perlu diperhatikan yaitu kecepatan dan status pencapaian.

Dengan latar belakang yang membahas masalah multikolinieritas pada IPM di Provinsi Sulawesi Selatan maka penelitian akan diselesaikan dengan menggunakan perbandingan metode Principal Component Regression dan Regression Ridge.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, permasalahan yang akan dibahas dalam penelitian ini adalah sebagai berikut :

1. Bagaimana penerapan model regresi pada metode Principal Component Regression dalam mengatasi multikolinieritas pada IPM di Provinsi Sulawesi Selatan ?
2. Bagaimana penerapan model regresi pada metode Regression Ridge dalam mengatasi multikolinieritas pada IPM di Provinsi Sulawesi Selatan ?
3. Bagaimana perbandingan metode Principal Component Regression dengan Regression Ridge dalam mengatasi multikolinieritas pada IPM di Provinsi Sulawesi Selatan ?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah di atas tujuan yang ingin dicapai dari penelitian ini adalah sebagai berikut :

1. Mengetahui penerapan metode Pricipal Component Regression dalam mengatasi Multikolinieritas pada IPM di Provinsi Sulawesi Selatan.
2. Mengetahui penerapan metode Regression Ridge dalam mengatasi Multikolinieritas pada IPM di Provinsi Sulawesi Selatan.
3. Mengetahui perbandingan metode Pricipal Component Regression dengan Regression Ridge dalam mengatasi multikolinieritas pada IPM di Provinsi Sulawesi Selatan.

1.4 Manfaat Penelitian

1. Manfaat Teoritis

Bagi pihak pembaca penelitian ini dapat menambah wawasan mengenai metode Pricipal Component Regression dan Regression Ridge dalam mengatasi multikolinieritas.

2. Manfaat Praktis

Bagi peneliti mampu menerapkan metode yang sesuai materi yang dipelajari serta peneliti menambah wawasan mengenai metode Pricipal Cmponent Regression dan Regression Ridge dalam mengatasi multikolinieritas.

1.5 Batasan Masalah

Batasan masalah adalah sebagai berikut :

1. Meneliti tentang masalah multikolinieritas
2. Identifikasi masalah dengan hanya mengatasi multikolinieritas .
3. Metode yang digunakan adalah *Principal Component Regression (PCR)* dan metode *regeression ridge (RR)*

