

BAB I

PENDAHULUAN

1.1 Latar Belakang

Regresi linear (*linear regression*) adalah metode yang digunakan untuk memperoleh model hubungan antara satu variabel terikat (dependen) dengan satu atau lebih variabel bebas (independen). Harlan (2018:5) menyebutkan bahwa Variabel dependen pada regresi linear disebut juga sebagai respons, sedangkan variabel independen dikenal pula sebagai prediktor atau regresor. Jika hanya terdapat satu variabel independen dalam model maka teknik ini disebut sebagai regresi linear sederhana (*simple linear regression*), sedangkan jika terdapat beberapa variabel independen, teknik ini disebut regresi linear ganda (*multiple linear regression*). Analisis regresi memiliki 3 kegunaan, yaitu untuk tujuan deskripsi fenomena data atau kasus yang sedang diteliti, untuk tujuan kontrol, serta untuk tujuan prediksi.

Bawono dan Shina (2018:20) mengatakan model regresi linier dapat disebut valid apabila model tersebut memenuhi asumsi-asumsi klasik yaitu tidak terjadinya autokorelasi, heteroskedastisitas, multikolinearitas dan normalitas. Autokorelasi adalah keadaan dimana nilai korelasi dalam persamaan regresi bernilai nol, heteroskedastisitas terjadi ketika terdapat ketaksamaan varian dari residual untuk semua pengamatan pada model regresi, multikolinearitas terjadi ketika ada korelasi diantara variabel bebas sedangkan normalitas terjadi ketika variabel pengganggu atau residual dalam persamaan regresi memiliki distribusi normal. Jika salah satu dari asumsi ini tidak dipenuhi maka persamaan regresi yang dihasilkan dengan metode kuadrat terkecil akan bersifat tidak valid, sedangkan jika seluruh asumsi dipenuhi maka model tersebut dikatakan BLUE (*Best Linier Unbiased Estimation*). Asumsi yang akan dibahas dalam penelitian ini adalah asumsi multikolinearitas.

Multikolinearitas adalah situasi dimana terdapat hubungan linier sempurna atau pasti antara beberapa atau semua variabel bebas dari model regresi berganda. Dalam arti luas berarti terdapat korelasi yang tinggi antara variabel bebas (Bawono dan Shina, 2018:46). Jika terjadi korelasi yang tinggi antar variabel bebas, pengujian hipotesis parameter berdasarkan metode kuadrat terkecil atau *Ordinary*

Least Square (OLS) akan memberikan hasil yang tidak valid (galat yang dihasilkan akan menjadi besar, variansi dan kovariansi parameter tidak berhingga). Nilai *standar error* setiap koefisien regresi menjadi besar sehingga hasil estimasi menggunakan OLS menjadi tidak efisien (variansi tidak minimum). Selang kepercayaan juga akan melebar. Melebarnya selang kepercayaan diakibatkan oleh membesarnya *standar error*. Melebarnya selang kepercayaan berarti hasil estimasi menjadi tidak dapat dipercaya. Asumsi multikolinearitas harus terpenuhi agar menghasilkan model regresi linier berganda yang baik.

Ada beberapa cara yang dapat digunakan untuk mengatasi masalah multikolinearitas diantaranya yaitu regresi komponen utama, regresi *ridge*, dan regresi PLS (*Partial Least Square*). Metode regresi komponen utama bertujuan untuk menyederhanakan variabel yang diamati dengan cara mereduksi dimensinya, metode regresi *ridge* bertujuan untuk memperkecil variansi estimator koefisien regresi, sedangkan metode regresi PLS bertujuan untuk perancangan model. Metode analisis regresi komponen utama telah diteliti oleh Silalahi (2011), Markus, et al (2012) dan Sudrajat (2016). Data yang mengandung multikolinearitas pada penelitian-penelitian tersebut dapat diatasi dengan menggunakan Metode Regresi Komponen Utama sehingga diperoleh persamaan regresi yang sudah tepat untuk analisis data. Namun metode yang akan digunakan dalam penelitian ini adalah Metode Regresi *Ridge*.

Indahwati, et al (2014) telah melakukan penelitian tentang metode PLS untuk mengatasi multikolinearitas pada model regresi linear berganda serta membandingkan tingkat efisiensi dari metode OLS dan PLS. Data yang digunakan adalah data yang dibangkitkan menggunakan program R dengan ukuran pengamatan 30, 100 dan 200 pengamatan. Variabel yang digunakan terdiri dari tiga variabel bebas x dan satu variabel terikat y . Hasil penelitian tersebut menyimpulkan bahwa penduga parameter regresi yang dihasilkan oleh metode PLS menjadi bias dan tidak efisien ketika digunakan untuk mengestimasi parameter regresi ketika terdapat korelasi negatif, tidak ada korelasi dan korelasi kecil antar variabel bebasnya. Hal tersebut terlihat dari nilai bias dan MSE yang dihasilkan metode PLS lebih besar dibandingkan dengan metode OLS.

Penelitian lainnya juga dilakukan oleh Rosyadi (2018) yaitu variansi estimator koefisien regresi dalam masalah multikolinearitas dapat diperkecil dengan menggunakan

regresi *Ridge* pada data Indeks Pembangunan Manusia (IPM) di Provinsi Jawa Tengah. Suriyanto (2016) telah melakukan penelitian mengenai perbandingan metode analisis regresi komponen utama dan Regresi *Ridge* dengan membandingkan nilai MSE dari kedua metode tersebut. Diperoleh hasil bahwa Regresi *Ridge* memiliki nilai MSE jauh lebih kecil dan Regresi *Ridge* lebih efektif dibanding dengan metode analisis regresi komponen utama dalam mengatasi multikolinearitas.

Berdasarkan uraian latar belakang di atas, penulis tertarik ingin melakukan penelitian mengenai “Model Regresi *Ridge* Untuk Mengatasi Masalah Multikolinearitas Dalam Mengestimasi Parameter Regresi Linier Berganda”.

1.2 Rumusan Masalah

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

1. Bagaimana penggunaan metode regresi *ridge* untuk mengatasi masalah multikolinearitas yang terjadi pada regresi linier berganda pada data Produk Domestik Regional Bruto (PDRB) Provinsi Sumatera Utara?
2. Apakah dengan menggunakan metode regresi *ridge* dapat memberikan model yang baik ditunjukkan dengan *mean squared error* yang lebih efisien?

1.3 Batasan Masalah

Agar penelitian ini tidak meluas dan lebih terarah, maka penulis memberikan batasan masalah yaitu:

1. Masalah yang ingin diatasi yaitu masalah multikolinearitas dan diasumsikan bahwa data telah memenuhi asumsi klasik yang lain.
2. Metode yang digunakan dalam mengatasi masalah multikolinearitas adalah metode regresi *ridge*.
3. Data yang digunakan yaitu data Produk Domestik Regional Bruto (PDRB) Provinsi Sumatera Utara tahun 2011-2020 yang diperoleh melalui *website* Badan Pusat Statistik (BPS).

1.4 Tujuan Penelitian

Sesuai dengan latar belakang masalah dan perumusan masalah yang telah diuraikan di atas, maka penelitian ini mempunyai tujuan sebagai berikut:

1. Mengetahui penggunaan metode regresi *ridge* untuk mengatasi masalah multikolinearitas yang terjadi pada regresi linier berganda pada data Produk Domestik Regional Bruto (PDRB) Provinsi Sumatera Utara.
2. Menghasilkan model regresi yang baik ditunjukkan dengan *mean squared error* yang lebih efisien.

1.5 Manfaat Penelitian

Adapun manfaat yang diharapkan pada penelitian ini yaitu sebagai berikut:

1. Bagi penulis, penelitian ini berguna untuk menambah wawasan mengenai solusi masalah multikolinearitas dengan metode regresi *ridge*.
2. Bagi pembaca, sebagai tambahan ilmu dan wawasan tentang regresi *ridge* sebagai salah satu metode estimasi alternatif untuk mengatasi masalah multikolinearitas yang terjadi pada regresi linear berganda.
3. Bagi akademik, dapat digunakan sebagai tambahan informasi dan referensi bacaan mengenai metode regresi *ridge* bagi yang hendak melakukan penelitian serupa.