

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis regresi adalah sebuah alat statistika untuk mengetahui bentuk hubungan antara peubah respon sebagai fungsi dari peubah prediktor. Model yang menggunakan lebih dari dua peubah prediktor dan memiliki parameter linier disebut model regresi linier berganda. Istilah regresi diperkenalkan oleh Francis Galton dalam artikel *Family Likeness in Stature*. Galton menunjukkan bahwa tinggi badan anak laki-laki dari ayah yang lebih tinggi setelah beberapa generasi cenderung mundur (*regressed*) mendekati nilai tengah populasi.

Metode pendugaan parameter regresi yang paling sering digunakan adalah Metode Kuadrat Terkecil (MKT). MKT dikemukakan oleh Carl Friedrich Gauss, seorang ahli matematika asal Jerman. Gauss mengatakan, apabila asumsi analisis regresi terpenuhi, penduga parameter regresi akan memiliki sifat BLUE (*Best Linear Unbiased Estimator*). Asumsi dalam analisis regresi adalah kenormalan sisaan, tidak terdapat ketergantungan antar sisaan (non autokorelasi), kehomogenan ragam sisaan, dan non multikolinieritas.

Pada analisis regresi linier berganda, pelanggaran terhadap asumsi non multikolinieritas sering terjadi. Multikolinieritas terjadi ketika terdapat hubungan linier antar peubah prediktor, menyebabkan kondisi buruk (*ill conditioned*) di mana matriks $X'X$ hampir singular, yang mengakibatkan penduga parameter regresi menggunakan MKT tidak bersifat BLUE. Oleh karena itu, multikolinieritas harus diatasi dengan metode regresi gulud (*ridge regression*).

Regresi *ridge* pertama kali dikemukakan oleh A.E. Hoerl pada tahun 1970 dalam makalah *Ridge Regression: Biased Estimation for Nonorthogonal Problem* (Draper dan Smith, 1992). Pendugaan parameter regresi *ridge* dapat dilakukan menggunakan MKT yang telah dimodifikasi dengan menambahkan konstanta bias (c), yang menghasilkan penduga parameter regresi *ridge* bersifat bias namun memiliki ragam minimum.

Terdapat beberapa metode dalam regresi *ridge*, seperti *Ordinary Ridge Regression* (ORR), *Generalized Ridge Regression* (GRR), dan *Directed Ridge Regression* (DRR). El-Dereny dan Rashwan (2011) dalam *Solving Multicollinearity Problem Using Ridge Regression Model* mengatakan bahwa ketiga metode regresi

ridge tersebut lebih baik dibandingkan MKT ketika data mengandung multikolinieritas. Selain itu, penelitian juga dilakukan oleh Utami, *et.al.* (2013) berjudul Penerapan Metode Generalized Ridge Regression dalam Mengatasi Masalah Multikolinieritas. Secara umum, konsep ORR adalah memberikan konstanta bias (c) yang sama untuk setiap peubah prediktor, sedangkan GRR merupakan pengembangan dari ORR, di mana GRR menggunakan konstanta bias (c) yang berbeda untuk setiap peubah prediktor ($c = c_1, c_2, \dots, c_k$), $k =$ banyaknya peubah prediktor.

Pada penelitian ini akan dibandingkan penduga parameter regresi hasil simulasi menggunakan *Ordinary Ridge Regression* dan *Generalized Ridge Regression* pada tingkat koefisien korelasi berbeda.

1.2 Rumusan Masalah

1. Bagaimana proses pembangkitan data untuk simulasi *Ordinary Ridge Regression* dan *Generalized Ridge Regression* pada tingkat koefisien korelasi berbeda?
2. Bagaimana proses pendugaan parameter regresi menggunakan *Ordinary Ridge Regression* dan *Generalized Ridge Regression* pada tingkat koefisien korelasi berbeda?
3. Bagaimana perbandingan hasil pendugaan parameter regresi dua metode?

1.3 Tujuan Penelitian

1. Membangkitkan data untuk simulasi *Ordinary Ridge Regression* dan *Generalized Ridge Regression* pada tingkat koefisien korelasi berbeda.
2. Menduga parameter regresi menggunakan *Ordinary Ridge Regression* dan *Generalized Ridge Regression* pada tingkat koefisien korelasi berbeda.
3. Membandingkan hasil pendugaan parameter regresi dua metode.

1.4 Manfaat Penelitian

Peneliti dapat menduga parameter regresi menggunakan *Ordinary Ridge Regression* dan *Generalized Ridge Regression* kemudian menjadikan metode ini sebagai alternatif pendugaan parameter regresi jika data mengandung multikolinieritas.

1.5 Batasan Masalah

Penelitian ini dibatasi pada:

1. Pendugaan parameter regresi menggunakan *Ordinary Ridge Regression* dan *Generalized Ridge Regression*.
2. Koefisien korelasi dalam pembangkitan data bernilai positif.
3. Penentuan konstanta bias menggunakan metode usulan Hoerl dan Kennard (1970) yaitu $\mathbf{C} = \text{diag}\left(\frac{\hat{\sigma}^2}{\hat{\alpha}_j^2}\right), j = 1, 2, \dots, k$.