

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Regresi linier adalah metode Statistika yang digunakan untuk mengetahui hubungan antara variabel terikat (dependen; respon; Y) dengan satu atau lebih variabel bebas (independen; prediktor; X). Apabila banyaknya variabel bebas hanya ada satu maka disebut regresi linier sederhana, sedangkan apabila terdapat lebih dari satu variabel bebas maka disebut sebagai regresi linier berganda.

Analisis regresi setidak-tidaknya memiliki 3 kegunaan, yaitu untuk tujuan deskripsi dari fenomena data atau kasus yang sedang diteliti, untuk tujuan kontrol serta sebagai prediksi. Regresi mampu mendeskripsikan fenomena data melalui terbentuknya suatu model hubungan yang bersifat numerik. Regresi juga dapat digunakan untuk melakukan pengendalian (kontrol) terhadap suatu kasus atau hal-hal yang sedang diamati melalui penggunaan model regresi yang diperoleh. Selain itu, model regresi juga dapat dimanfaatkan untuk melakukan prediksi variabel terikat.

Metode kuadrat terkecil (*ordinary least square*) merupakan salah satu metode yang sering digunakan untuk mendapatkan nilai-nilai penduga parameter dalam pemodelan regresi. Penggunaan metode kuadrat terkecil memerlukan beberapa asumsi klasik yang harus dipenuhi oleh komponen sisaan atau galat ( $e_i$ ) dalam model yang dihasilkan. Beberapa asumsi itu antara lain bahwa galat harus memenuhi asumsi kenormalan, kehomogenan ragam dan tidak terjadi

autokorelasi. Myers (1990) menyatakan bahwa apabila asumsi itu terpenuhi, maka penduga parameter yang diperoleh bersifat *best linier unbiased estimator* (BLUE)

Dalam berbagai kasus tidak jarang ditemui hal-hal yang menyebabkan tidak terpenuhinya asumsi tersebut. Salah satu penyebabnya adalah adanya pencilan (*outlier*) dalam data amatan. Saat ada asumsi yang tidak terpenuhi, maka penggunaan metode kuadrat terkecil akan memberikan kesimpulan yang bersifat kurang baik atau nilai penduga parameternya bersifat bias sehingga berakibat interpretasi hasil yang diperoleh menjadi tidak valid. Oleh karena itu, saat asumsi klasik tidak terpenuhi maka metode kuadrat terkecil perlu dihindari.

Karena metode kuadrat terkecil tidak dapat digunakan jika asumsinya tidak terpenuhi maka selanjutnya diperlukan alternatif metode penduga parameter lain yang dapat mengatasi adanya pencilan dalam data amatan. Metode Robust dapat menjadi alternatif pilihan untuk menghasilkan model yang lebih baik dari hasil metode kuadrat terkecil berdasarkan kriteria  $R^2$  dan *mean square error* (MSE) bagi masing-masing model.

Regresi *robust* diperkenalkan oleh Andrews (1972) dan merupakan metode regresi yang digunakan ketika distribusi dari galat tidak normal dan atau adanya beberapa pencilan yang berpengaruh pada model. Metode ini merupakan alat penting untuk menganalisa data yang dipengaruhi oleh pencilan sehingga dihasilkan model yang *robust* atau *resistance* terhadap pencilan.

Chen (2002) mengemukakan bahwa regresi *robust* terdiri dari 5 metode penduga, yaitu estimasi *robust M*, estimasi *robust least median of square* (LMS), estimasi *robust least trimmed square* (LTS), estimasi *robust S* dan estimasi *robust MM*.

Dari ke-5 metode diatas, Penulis akan berfokus pada metode estimasi parameter dengan menggunakan metode estimasi regresi *robust least trimmed square* (LTS). Metode LTS menurut Rousseeuw (1984) memiliki kemampuan yang lebih baik dibandingkan dengan metode-metode lainnya karena mampu mengatasi pencilan yang disebabkan baik oleh variabel bebasnya maupun variabel terikatnya. Selain itu, kelebihan dari metode LTS adalah algoritmanya yang lebih mudah dibandingkan metode lainnya karena dalam proses estimasinya LTS hanya akan memangkas sebaran data berdasarkan jumlah pencilan yang teramati sehingga akan menghasilkan fungsi objektif yang mengecil dan konvergen.

## **1.2. Permasalahan**

Berdasarkan latar belakang yang telah dikemukakan sebelumnya, maka permasalahannya adalah bagaimana mendapatkan model regresi terbaik dari data pengamatan yang mengandung pencilan.

## **1.3. Pembatasan Masalah**

Dalam penulisan tugas akhir ini, batasan masalah sangat diperlukan untuk menjamin keabsahan dalam kesimpulan yang diperoleh. Agar tidak terjadi penyimpangan dari tujuan semula dan pemecahan masalah lebih terkonsentrasi, maka pembahasan akan difokuskan pada regresi *robust* dengan metode *least trimmed square* (LTS).

## **1.4. Tujuan Penulisan**

Berdasarkan perumusan masalah yang telah dikemukakan sebelumnya, maka tujuan dari penulisan tugas akhir ini adalah mendapatkan nilai estimasi parameter dari model regresi *least trimmed square* (LTS) sebagai regresi yang *robust* (kuat) terhadap adanya data pencilan

### **1.5. Manfaat Penulisan**

Manfaat penulisan tugas akhir ini adalah untuk dapat mengetahui penggunaan metode *robust least trimmed square* (LTS) dalam menentukan nilai parameter penduga pada model regresi linier sederhana dengan data yang mengandung pencilan.

### **1.6. Sistematika Penulisan**

Sistematika penulisan tugas akhir ini adalah Bab I merupakan Pendahuluan yang membahas mengenai latar belakang, permasalahan, pembatasan masalah, tujuan penulisan dan sistematika penulisan. Bab II merupakan Landasan Teori yang terdiri dari deskripsi dari variabel random, metode maksimum likelihood dan regresi linier, asumsi klasik serta uji kecocokan model. Bab III berisi pembahasan tentang Estimasi *Least Trimmed Square*, regresi *robust*, pencilan, *breakdown point*, koefisien determinasi, *scale estimator*, estimasi parameter dengan menggunakan metode *least trimmed square* serta contoh penerapan metode *Least Trimmed Square* pada kasus data yang mengandung pencilan. Bab IV merupakan kesimpulan dari bab-bab sebelumnya.