

BAB I

PENDAHULUAN

Lebih dari dua puluh tahun yang lalu terjadi suatu perubahan besar dalam bidang pencocokan model secara umum. Ada banyak faktor yang menyebabkan hal itu, diantaranya adalah banyak digunakannya komputer dalam pengolahan statistik. Pencocokan model regresi linier dengan metode kuadrat terkecil digunakan secara luas dalam prosedur statistika.

Beberapa elemen yang menentukan suatu pencocokan persamaan regresi yaitu variabel, pengamatan (*observasi*), dan asumsi model (Chatterjee dan Hadi (1987)). Tulisan ini mempelajari pencocokan model regresi yang ditentukan oleh variabel yang termasuk dalam persamaan, dan bagaimana model terpengaruh oleh penghapusan satu atau beberapa pengamatan tertentu.

Pencocokan yang sangat sensitif terhadap suatu titik data tertentu, variabel tertentu, atau asumsi tertentu akan diperlakukan secara hati-hati dan diterapkan dengan teliti. Biasanya, tidak semua pengamatan mempunyai pengaruh yang sama dalam suatu pencocokan dengan kuadrat terkecil dan dalam kesimpulan-kesimpulan yang dihasilkan dari analisisnya. Pengamatan, secara individu atau kelompok yang terlalu mempengaruhi pencocokan persamaan regresi dibandingkan dengan pengamatan-pengamatan lain disebut *pengamatan berpengaruh* (Chatterjee dan Hadi (1987)).

Dalam kerangka regresi linier, didefinisikan *outlier* yaitu suatu pengamatan yang mempunyai nilai residual besar dibandingkan dengan pengamatan lain dalam himpunan data. Titik *high-leverage* adalah suatu pengamatan yang mempunyai nilai p_{ii} (elemen diagonal ke- i dari matrik Prediksi) besar dibandingkan dengan pengamatan yang lain dalam himpunan data. Outlier, titik *high-leverage* dan pengamatan berpengaruh merupakan tiga konsep yang saling berhubungan (Chatterjee dan Hadi (1987)).

Titik *high-leverage* cenderung untuk mempengaruhi eigenstruktur dari variabel bebas, dimana titik *high-leverage* dapat menimbulkan atau menyembunyikan kolinieritas. Titik yang menyembunyikan atau menampakkan kolinieritas diantara variabel bebas didefinisikan sebagai *titik kolinieritas-berpengaruh* (Chatterjee dan Hadi (1987)).

Kenyataan bahwa suatu pengamatan merupakan outlier jelas tidak baik, namun tidak berarti bahwa pengamatan itu berpengaruh besar dalam estimasi koefisien regresi. Titik outlier menimbulkan dugaan bahwa ia akan berpengaruh besar pada posisi garis regresi disekitar titik tersebut. Salah satu cara mengatasi masalah ini adalah dengan memeriksa apakah penghapusan satu atau beberapa pengamatan kritis mengubah secara berarti persamaan regresinya serta kesimpulan-kesimpulannya (Draper dan Smith (1981)).

Penting bagi seorang analis data untuk dapat mengidentifikasi pengamatan yang berpengaruh dan menaksir

pengaruhnya pada beberapa aspek dari analisisnya. Salah satu metode yang digunakan adalah metode pendekatan penghapusan (*The omission approach*). Metode pendekatan penghapusan menguji bagaimana berbagai kuantitas yang terlibat dalam analisa regresi berubah jika satu atau beberapa pengamatan dihapus (Chatterjee dan Hadi (1987)).

Untuk maksud tersebut, maka dalam penulisan tugas akhir ini akan dibahas metode pendekatan penghapusan untuk mengetahui keberadaan pengamatan berpengaruh dalam himpunan data. Tujuan lainnya untuk menguji sensitifitas faktor-faktor yang menentukan pencocokan model.

Untuk lebih memudahkan dalam pembahasan masalah ini diperlukan pemahaman *aljabar matrik* dan sifat-sifatnya dengan konsep dasar *analisa regresi*. Penulisan tugas akhir ini disusun dengan sistematika pembahasan yang terbagi dalam empat bab sebagai berikut: Bab I berisi pendahuluan. Bab II berisi teori dasar, membahas teori yang mendasari masalah inti. Bab III berisi pembahasan masalah inti, yang membahas tentang pendeteksian multi pengamatan berpengaruh dengan metode pendekatan penghapusan. Bab IV berisi kesimpulan dari pembahasan masalah inti pada bab sebelumnya.