

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Analisis regresi merupakan suatu alat dalam statistik yang mengalami perkembangan pesat dan banyak digunakan dalam berbagai bidang kehidupan. Analisis tersebut mempunyai tujuan untuk mengestimasi hubungan antara variabel *dependent* (respon) dengan variabel *independent* (bebas).

Ada beberapa metode yang digunakan untuk mengestimasi parameter-parameter dalam persamaan regresi, salah satunya yang paling sering digunakan adalah metode *Ordinary Least Square* (OLS). Secara matematik, penentuan parameter regresi ini dengan cara meminimumkan jumlah kuadrat dari residualnya [24]. Penggunaan metode OLS dilandasi pada beberapa asumsi yaitu kenormalan sisaan, non-multikolinieritas, kehomogenan ragam sisaan dan non-autokorelasi. Semua asumsi harus terpenuhi supaya didapatkan penduga parameter yang bersifat BLUE (*Best Linear Unbiased Estimator*). Namun pada metode OLS sering ditemukan penyimpangan asumsi pada data, jika data tidak memenuhi salah satu asumsi, maka penduga OLS tidak lagi baik digunakan.

Salah satu asumsi yang seringkali tidak terpenuhi ketika data mengandung pencilan (*outlier*) adalah asumsi normalitas. Uji normalitas dilakukan un-

tuk melihat apakah nilai residual terdistribusi normal atau tidak. Model regresi yang baik memiliki nilai residual yang terdistribusi normal. Jadi uji normalitas bukan dilakukan pada masing-masing variabel tetapi pada nilai residualnya. Sering terjadi kesalahan yang jamak yaitu bahwa uji normalitas dilakukan pada masing-masing variabel. Hal ini tidak dilarang tetapi model regresi memerlukan normalitas pada nilai residualnya bukan pada masing-masing variabel penelitian. Untuk mengatasi hal tersebut, peneliti melakukan transformasi terhadap data agar asumsi terpenuhi. Namun seringkali asumsi tersebut tidak terpenuhi meskipun telah dilakukan transformasi yang akhirnya mengakibatkan dugaan berbias. Kemudian berkembanglah metode Regresi Median (Median Regression) sebagai pengganti pendekatan rata-rata pada OLS menjadi median. Namun permasalahannya adalah metode Regresi Median juga dianggap kurang tepat karena metode ini hanya dapat melihat pada dua kelompok data, padahal ada kemungkinan data bisa terbagi lebih dari dua kelompok.

Selanjutnya dikembangkan metode Regresi Kuantil yang pada umumnya dipakai dalam kasus ekonometrika. Regresi kuantil pertama kali diperkenalkan oleh Koenker dan Bassett[11]. Metode ini tidak membutuhkan asumsi error dalam model dan estimatornya bersifat tegar terhadap variabel respon yang mengandung pencilan. Pendekatan pada regresi kuantil adalah memisahkan atau membagi data menjadi kuantil-kuantil, dengan menduga fungsi kuantil bersyarat pada suatu sebaran data tersebut dan meminimumkan sisaan mutlak berbobot yang tidak simetris. Pada metode regresi kuantil biasanya memerlukan ukuran data besar. Pengambilan data sampel berukuran besar akan

membutuhkan waktu yang lama dan tenaga yang banyak. Oleh karena itu digunakan metode Bayes yang tidak membutuhkan data besar.

Bayes memperkenalkan suatu metode untuk mengestimasi parameter yang akan diestimasi dengan memanfaatkan informasi awal yang disebut distribusi awal (prior). Metode ini dikenal dengan metode Bayesian. Distribusi prior ini dapat berasal dari data penelitian sebelumnya atau berdasarkan intuisi seorang peneliti. Informasi prior dari sebaran parameter tersebut kemudian digabungkan dengan informasi dari data yang didapat dari pengambilan sampel atau yang disebut juga dengan fungsi *likelihood* sehingga didapat distribusi posterior dari parameter. Rataan dan varian dari distribusi posterior ini atau posterior mean dan posterior varian yang akan menjadi penduga bagi parameter regresi dengan metode Bayesian [3].

Kadang kala untuk memperoleh marginal posterior sulit untuk ditentukan secara analitik. Dalam kasus ini, mengintegalkan parameter dari distribusi posterior bersama atau bahkan menentukan kenormalan dari distribusi posterior secara umum adalah hal yang tidak mungkin. Terlebih lagi menghitung fungsi distribusi posterior dari parameter itu sulit dilakukan. Metode Bayesian mengatasi hal ini dengan bantuan algoritma MCMC (Markov Chain Monte Carlo). MCMC (Markov Chain Monte Carlo) dapat dengan mudah digunakan untuk mendapatkan distribusi posterior bahkan dalam situasi yang kompleks.

Pada penelitian ini akan dilakukan estimasi parameter model dengan menggabungkan metode Regresi Kuantil dan metode Regresi Bayesian yang

disebut metode Regresi Kuantil Bayesian. Dari uraian diatas peneliti akan membandingkan hasil estimasi dari metode Regresi Kuantil dan metode Regresi Kuantil Bayesian dengan menggunakan data bangkitan dan data empirik.

1.2 Perumusan Masalah

Permasalahan yang akan dibahas pada tesis ini adalah :

1. Bagaimana mengestimasi parameter model dengan metode Regresi Kuantil pada *error* yang tak normal untuk data bangkitan dan data empirik ?
2. Bagaimana mengestimasi parameter model dengan metode Regresi Kuantil Bayesian pada *error* yang tak normal untuk data data bangkitan dan data empirik ?
3. Bagaimana perbandingan model hasil penerapan metode Regresi Kuantil dan metode Regresi Kuantil Bayesian dalam mengestimasi parameter pada *error* yang tak normal untuk data bangkitan dan data empirik ?

1.3 Tujuan Penelitian

Adapun tujuan penelitian pada tesis ini adalah :

1. Menerapkan metode Regresi Kuantil dalam mengestimasi parameter pada *error* yang tak normal untuk data bangkitan dan data empirik.
2. Menerapkan metode Regresi Kuantil Bayesian dalam mengestimasi parameter pada *error* yang tak normal untuk data bangkitan dan data empirik.

3. Membandingkan metode hasil estimasi parameter antara metode Regresi Kuantil dan metode Regresi Kuantil Bayesian dalam mengestimasi parameter pada *error* yang tak normal untuk data bangkitan dan data empirik.

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat menjadi referensi bagi peneliti yang ingin meneliti lebih lanjut mengenai Regresi Kuantil Bayesian dan juga dapat menjadi referensi untuk melakukan inovasi-inovasi lainnya yang berhubungan dengan Regresi Kuantil Bayesian.

