

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Kehidupan manusia di dunia ini sangat berkaitan erat dengan matematika, hal paling sederhana seperti menghitung jari, hingga hal sulit seperti menghitung nilai investasi di suatu kabupaten atau kota di sebuah provinsi. Karena keterlibatannya yang mencakup banyak bidang dalam kehidupan, maka diperlukan berbagai macam metode yang digunakan untuk memaksimalkan penggunaan matematika dalam kehidupan sehari-hari. Dan salah satu metode yang paling sering digunakan adalah ilmu statistika.

Statistika merupakan salah satu bagian dari matematika yang sangat bermanfaat dan dapat digunakan dalam segala aspek kehidupan manusia. Dengan menggunakan statistika, data yang jumlahnya sangat banyak dapat diolah dengan cepat dibandingkan dengan mengolahnya secara manual. Pada prinsipnya, statistika bekerja pada proses pengumpulan data, analisis data, dan interpretasi data yang dapat disampaikan kepada masyarakat luas dengan bahasa yang lebih mudah dimengerti. Statistika tentunya mempunyai banyak metode yang digunakan untuk menganalisis data yang penggunaannya disesuaikan dengan kebutuhan peneliti, dan salah satu metode dari statistika yang cukup sering digunakan untuk menganalisis data adalah analisis regresi.

Analisis regresi berkaitan dengan hubungan antar dua variabel atau lebih, yang bertujuan untuk mengetahui variabel mana yang lebih dominan diantara variabel lainnya sehingga dengan hasil analisis tersebut dapat dibuatkan suatu kebijakan atau tindakan berdasarkan hasil analisis tersebut sehingga kesalahan yang digunakan harus dapat diminimalisir.

Karena penggunaannya yang sangat penting, maka analisis regresi terus mengalami perkembangan dari waktu ke waktu untuk memaksimalkan penggunaannya dan meminimalisir kesalahan dalam penggunaannya. Dimulai dari

analisis regresi sederhana yang terdiri dari satu variabel bebas dan satu variabel tak bebas, selanjutnya berkembang menjadi analisis regresi berganda yang terdiri dari dua variabel bebas atau lebih. Penggunaan analisis regresi perlu memperhatikan beberapa asumsi regresi yaitu normalitas, linearitas, homogenitas, dan multikolinieritas. Analisis regresi linier sederhana dan berganda tidak dapat dilanjutkan apabila salah satu uji asumsi tidak terpenuhi. Alternatif pendekatan analisis lain dengan menggunakan regresi kuantil atau regresi kuantil bayesian ketika terjadi heteroskedastisitas pada data.

Regresi kuantil bayes (*bayesian quantile regression*) pertama kali diperkenalkan oleh Yu dan Moyeed, di mana mereka menggunakan pendekatan Bayes untuk mengestimasi setiap parameter yang digunakan. Cara yang digunakan adalah dengan mengestimasi nilai distribusi posterior berdasarkan fungsi *likelihood* yang dibentuk oleh *asymmetric laplace distribution* (ALD) yang setiap parameternya akan diestimasi dengan menggunakan *Markov Chain Monte Carlo* (MCMC) dengan salah satu algoritmanya yaitu *gibbs sampling*.

Beberapa penelitian sebelumnya yang berkaitan dengan analisis regresi kuantil bayes dilakukan oleh beberapa orang dengan jenis publikasi yang berbeda diantaranya adalah Annisa Hanif [1] membahas mengenai faktor apa saja yang mempengaruhi harga emas, Ridha Anullah [2] dalam skripsinya yang membahas mengenai pengaruh dari usia dan paritas terhadap nilai glukosa darah ibu hamil, dan kelompok peneliti dari Malaysia yaitu Ros Idayuwati Alaudin, Noriszura Ismail, dan Zaidi Isa dalam jurnal [3] yang membahas mengenai pengaruh variabel terkait pekerjaan dan non-pekerjaan terhadap pengeluaran perkapita untuk pensiunan di Malaysia, sekaligus menjadi acuan utama dalam penyusunan skripsi ini.

Pada skripsi ini penulis akan menjabarkan secara detil apa itu regresi kuantil bayes secara bertahap dimulai dari regresi linier sederhana, selanjutnya penjelasan kuantil, selanjutnya bagaimana pembentukan model regresi kuantil, penjelasan metode Bayes hingga akhirnya membentuk metode regresi kuantil bayes.. Setelah itu akan ditunjukkan juga bagaimana cara kerjanya dengan menggunakan *software R*, yang di implementasikan untuk mengetahui seberapa besar pengaruh dari pajak,

realisasi PMDN, laju pertumbuhan PDRB, tingkat pengangguran terbuka, jumlah penduduk dan jumlah UMKM yang ada di Jawa Barat terhadap pendapatan asli daerah (PAD) setiap Kota/Kabupaten di Jawa Barat di tahun 2019, yang dibagi menjadi tiga studi kasus.

## 1.2 Rumusan Masalah

Rumusan masalah pada pembahasan Skripsi ini dapat dijabarkan sebagai berikut:

1. Bagaimana cara membangun model regresi kuantil bayes?
2. Bagaimana langkah-langkah estimasi parameter model regresi kuantil bayes dengan menggunakan algoritma *Gibbs sampling*?
3. Bagaimana penerapan model regresi kuantil bayes dengan menggunakan algoritma *gibbs sampling*?

## 1.3 Batasan Masalah

Dalam pembahasan Skripsi ini terdapat beberapa keterbatasan, diantaranya sebagai berikut:

1. Algoritma MCMC yang digunakan adalah *Gibbs Sampling*.
2. Data yang digunakan adalah data nilai pendapatan asli daerah (PAD), pajak, realisasi PMDN, laju pertumbuhan PDRB, tingkat pengangguran terbuka, jumlah penduduk, dan jumlah UMKM yang diambil secara sekunder.
3. Sampel yang digunakan adalah seluruh Kabupaten/Kota di Jawa Barat pada tahun 2019.
4. Data diperoleh secara *cross section*.
5. Uji asumsi yang dilakukan hanya uji heteroskedastisitas.
6. Prior yang digunakan  $\beta(\theta)$  dan  $\sigma(\theta)$ .
7. Prior  $\beta(\theta)$  berdistribusi secara normal, prior  $\sigma(\theta)$  berdistribusi secara invers gamma.
8. *Software* pengolah data yang digunakan adalah R Studio.

#### 1.4 Tujuan dan Manfaat

Berdasarkan rumusan masalah yang telah dijabarkan sebelumnya, maka tujuan yang ingin diperoleh di penelitian ini adalah:

1. Mengetahui bagaimana proses untuk membangun model regresi kuantil bayes.
2. Memahami bagaimana langkah-langkah dalam mengestimasi parameter dari model regresi kuantil bayes dengan algoritma *gibbs sampling* pada studi kasus yang digunakan.
3. Mengetahui cara penerapan model regresi kuantil bayes dengan menggunakan algoritma *gibbs sampling*.

Manfaat yang didapat dari penelitian ini adalah untuk mengetahui cara mengatasi segala jenis data yang terkena heteroskedastisitas yang disebabkan oleh adanya perubahan nilai varians dari setiap sampel yang digunakan, dan dapat menjadi referensi tambahan bagi pembaca yang berminat dalam statistika khususnya pada analisis regresi, karena ruang lingkup pembahasannya akan semakin luas dengan memaksimalkan potensi dari regresi kuantil bayes ini.

#### 1.5 Metodologi Penelitian

1. Studi Literatur

Studi literatur merupakan tahapan di mana penulis mencari referensi yang dibutuhkan dalam penyusunan skripsi ini dan mendukung topik yang dipahas pada skripsi ini yaitu mengenai analisis regresi, regresi kuantil, regresi kuantil bayes, dan lainnya. Referensi yang diperoleh dari studi literatur dapat berupa buku – buku, jurnal nasional dan internasional.

2. Simulasi

Pada tahapan simulasi, penulis melakukan simulasi tak langsung menggunakan simulasi Markov chain monte carlo (MCMC) dengan algoritma *gibbs sampling* yang digunakan untuk mengestimasi nilai dari setiap parameter.

### 3. Analisis

Pada tahap analisis, penulis menganalisis dan menginterpretasikan seluruh hasil estimasi parameter yang ada setelah meselanjutnyai proses simulasi di setiap kuantil yang digunakan.

### 4. Kesimpulan

Tahapan ini adalah tahap terakhir dari proses analisis model regresi kuantil bayes, karena pada tahap ini penulis akan menyimpulkan dan menginterpretasikan hasil yang diperoleh dari analisis yang telah dilakukan.

## 1.6 Sistematika Penulisan

Skripsi ini menggunakan sistematika penulisan yang terdiri dari lima bab yang tiap babnya terdiri dari beberapa subbab. Lima bab yang menyusun penulisan skripsi ini adalah:

### BAB I PENDAHULUAN

Pada bab ini meliputi latar belakang masalah, rumusan masalah, tujuan penelitian, ruang lingkup penelitian, metedologi penelitian, dan sistematika penulisan.

### BAB II LANDASAN TEORI

Bab ini akan menjelaskan hal – hal yang menjadi acuan dalam penulisan skripsi ini. Hal – hal tersebut meliputi analisis regresi, statistika urutan, kuantil, metode kuadrat terkecil (MKT), pemeriksaan asumsi regresi, fungsi *likelihood*, analisis bayes, *Markov chain monte carlo* (MCMC), dan *R Studio*.

### BAB III ANALISIS MODEL REGRESI KUANTIL BAYES MENGGUNAKAN ALGORITMA GIBBS SAMPLING

Bab ini akan menjelaskan masalah utama pada skripsi ini yaitu definisi dari model regresi kuantil, bagaimana pembentukan model regresi kuantil, definisi dari *asymmetric laplace distribution* (ALD), metode estimasi bayes, model regresi kuantil bayes, pembentukan parameter regresi kuantil bayes, pembentukan

distribusi prior dan posterior, dan yang terakhir adalah bagaimana cara estimasi parameter regresi kuantil dengan menggunakan algoritma *gibbs sampling*.

#### BAB IV STUDI KASUS

Pada bab ini akan dijelaskan seluruh proses bagaimana cara menerapkan regresi kuantil bayes dengan studi kasus untuk mengetahui pengaruh pajak, realisasi PMDN, laju pertumbuhan PDRB, tingkat pengangguran terbuka, jumlah penduduk dan jumlah UMKM yang ada di Jawa Barat terhadap pendapatan asli daerah (PAD) setiap Kota/Kabupaten di Jawa Barat di tahun 2019, yang terbagi kedalam tiga studi kasus.

#### BAB V PENUTUP

Bab ini akan menyimpulkan keseluruhan materi yang ada pada skripsi ini khususnya hasil penelitian yang dilakukan dan terdapat juga saran yang dapat digunakan untuk mengembangkan tulisan dan analisa dari masalah yang dikaji dalam skripsi ini.

