

BAB I

PENDAHULUAN

1.1 Latar Belakang

Studi *longitudinal* merupakan salah satu dari banyak metode penelitian yang sering digunakan dalam statistika. Studi ini digunakan untuk memahami bagaimana dan mengapa perubahan terjadi dalam suatu subjek baik individu atau kelompok dari waktu ke waktu. Penelitian jenis ini biasanya lebih mahal karena melibatkan lebih banyak pihak dibandingkan penelitian *cross-sectional*, tetapi dapat diperoleh hasil yang lebih akurat dalam mengungkapkan informasi tentang dinamika perubahan [2]. Selain itu, tergantung pada bagaimana teori dan metode penelitian dioperasionalkan, studi *longitudinal* dapat menghasilkan data yang lebih komprehensif.

Data *longitudinal* adalah data yang diperoleh dari hasil observasi atau pengamatan yang dilakukan secara berulang atau berkala pada setiap unit *cross-sectional* dalam kurun waktu tertentu [8]. Data *cross-sectional* yang berulang dapat digunakan untuk mempelajari perubahan dinamis melalui data *longitudinal*. Pengamatan berulang pada masing-masing unit *cross-sectional* tersebut akan menyebabkan adanya autokorelasi atau hubungan antar perulangan yang dilakukan pada setiap unit *cross-sectional*. Saat memodelkan data *longitudinal*, autokorelasi ini perlu diperhitungkan agar dapat menarik kesimpulan yang tepat [12].

Pada tahun 1986 Liang dan Zeger memperkenalkan metode *Generalized Estimating Equation* (GEE) untuk menyelesaikan data yang berkorelasi dan tidak berdistribusi normal khususnya pada data *longitudinal* [19]. Metode GEE merupakan perluasan dari metode *Generalized Linear Model* (GLM) yang dimana metode ini memodelkan fungsi yang diperoleh dari ekspektasi marjinal *variabel prediktor* sebagai penjelas yang dapat digunakan untuk memperkirakan parameter model. Karena asumsi yang digunakan dalam metode GLM adalah responnya saling bebas atau saling *independent*, sehingga kurang cocok digunakan untuk menyelesaikan permasalahan autokorelasi pada data *longitudinal*. Oleh karena itu, metode GEE banyak digunakan untuk menganalisis data *longitudinal* karena dalam

pendekatannya hanya membutuhkan dua momen pertama dari *variabel respon* dan penduga parameter yang dihasilkan.

Struktur korelasi dalam bentuk matriks yang menjelaskan proses pengulangan pengukuran unit *cross-sectional* digunakan untuk mengungkapkan masalah autokorelasi didalam GEE. Matriks korelasi dalam GEE mempunyai beberapa struktur, diantaranya *Independent*, *Exchangeable*, *Autoregressive* atau AR (1) dan *Unstructured* [18]. Hasil analisis yang akan diperoleh dari beberapa struktur korelasi tersebut akan tetap konsisten. Tetapi ketika sembarang dalam menentukan struktur korelasinya akan berpengaruh terhadap efisiensi penduga parameter yang diperoleh. Oleh karena itu, sangat penting memilih struktur korelasi yang tepat dalam metode GEE.

Pengembangan metode statistik untuk menentukan efisiensi dari penduga parameter GEE diperlukan dalam mengatasi masalah tersebut. Oleh karena itu, bisa menggunakan metode *Mixture* GEE atau GEE Campuran. GEE Campuran ini merupakan pengembangan metode untuk menambah efisiensi dari GEE yang diperkenalkan oleh Zhao dan Prentice pada tahun 1990 [6]. Pada GEE Campuran, populasi dari data yang diamati diasumsikan terdiri dari campuran (*mixture*) dari beberapa komponen atau kelompok yang berbeda. Setiap komponen dalam campuran ini mewakili distribusi probabilitas tertentu, dan data yang diamati dianggap berasal dari salah satu komponen ini, meskipun komponen yang memberikan kontribusi sebenarnya tidak diketahui.

Pada penelitian yang dilakukan oleh Zhao dan Prentice dihasilkan nilai AIC dan kesalahan standar pada metode GEE Campuran secara signifikan lebih rendah daripada GEE biasa, penelitiannya menunjukkan bahwa model GEE Campuran lebih unggul daripada GEE biasa. Hal ini karena GEE biasa tidak memodelkan struktur korelasi, sedangkan GEE Campuran melakukannya. Beberapa penelitian yang mengkaji mengenai GEE dan GEE Campuran ini adalah jurnal yang ditulis oleh Ming Wang membahas mengenai metode GEE dan menerapkannya pada data simulasi [10], jurnal yang disusun oleh Lili Xiu dkk dengan pembahasan mengenai metode GEE Campuran kemudian menerapkannya pada data simulasi [6], serta

jurnal yang ditulis oleh Niansheng Tang dan Wenjun Wang yang membahas tentang metode GEE Campuran dengan memperkirakan nilai kovariat yang hilang [7].

Berdasarkan uraian latar belakang di atas, maka penulis ingin mengkaji lebih dalam mengenai GEE Campuran pada data *longitudinal* yang diberi judul ”**Analisis Generalized Estimating Equation (GEE) Campuran Untuk Data Longitudinal**”. Metode GEE Campuran kemudian akan diaplikasikan pada data Indeks Pembangunan Manusia (IPM) pada 27 Kabupaten dan Kota di Provinsi Jawa Barat.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, maka penulis merumuskan masalah pada penelitian ini yaitu:

1. Bagaimana langkah-langkah analisis metode GEE Campuran untuk data *longitudinal* dengan kombinasi struktur korelasi *Exchangeable*, AR (1) dan *Unstructured* ?
2. Bagaimana penerapan metode GEE Campuran untuk data *longitudinal* Indeks Pembangunan Manusia (IPM) pada 27 Kabupaten dan Kota di Provinsi Jawa Barat ?

1.3 Batasan Masalah

Adapun batasan masalah pada penulisan penelitian ini untuk menjawab rumusan masalah di atas yaitu, antara lain:

1. Jenis data yang digunakan adalah data *longitudinal* seimbang (*balanced data*).
2. Struktur korelasi yang digunakan dalam metode GEE Campuran adalah kombinasi dari struktur korelasi *Exchangeable*, AR (1) dan *Unstructured*.
3. Data yang digunakan dalam penerapan metode GEE Campuran adalah data Indeks Pembangunan Manusia (IPM) pada 27 Kabupaten dan Kota di Provinsi Jawa Barat dengan rentang 3 tahun yaitu dari tahun 2021 sampai tahun 2023.

1.4 Tujuan Penelitian

Berdasarkan latar belakang dan rumusan masalah yang sudah disampaikan penulis sebelumnya, terdapat tujuan dalam penulisan penelitian ini, yaitu:

1. Mengetahui bagaimana langkah-langkah analisis metode GEE Campuran pada data *longitudinal* dengan kombinasi struktur korelasi *Exchangeable*, AR (1) dan *Unstructured*.
2. Mengetahui penerapan metode GEE Campuran untuk data *longitudinal* Indeks Pembangunan Manusia (IPM) pada 27 Kabupaten dan Kota di Provinsi Jawa Barat.

1.5 Metode Penelitian

Pada penelitian ini, penulis melakukan pengambilan data dari *website* Badan Pusat Statistik (BPS) provinsi Jawa Barat. Data yang digunakan yaitu data Indeks Pembangunan Manusia (IPM) pada 27 Kabupaten dan Kota di Provinsi Jawa Barat periode 2021-2023. Penulis mengumpulkan data, teori, dan informasi mengenai data *longitudinal*, metode GEE Campuran, dan struktur korelasi pada GEE yang didapatkan dari berbagai sumber pustaka berupa buku, jurnal, dan lain sebagainya.

Data yang telah diperoleh selanjutnya akan dianalisis menggunakan metode GEE Campuran. Uji kecocokan parameter menggunakan uji *Wald*. Pada proses analisisnya digunakan bantuan aplikasi R Studio versi 4.1.2.

1.6 Sistematika Penelitian

Berdasarkan sistematika penulisannya, skripsi ini terdiri dari atas lima bab serta daftar pustaka dan lampiran. Dalam setiap bab terdapat beberapa subbab yang dirincikan sebagai berikut:

BAB I : PENDAHULUAN

Bab ini berisi tentang pendahuluan diantaranya adalah latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, metode penelitian serta sistematika penulisan dari masalah yang dikaji.

BAB II : LANDASAN TEORI

Bab ini berisi penjelasan mengenai teori-teori yang berkaitan dengan masalah yang dikaji. Secara garis besar, bab ini mencakup semua yang berkaitan dengan penelitian serta data *longitudinal*.

BAB III : ANALISIS *GENERALIZED ESTIMATING EQUATION* (GEE) CAMPURAN UNTUK DATA *LONGITUDINAL*

Bab ini berisi pembahasan utama dari studi yang dikaji, meliputi pembahasan mengenai *Generalized Estimating Equations* (GEE) dan GEE Campuran dan bagaimana penerapannya pada data *longitudinal*.

BAB IV : STUDI KASUS DAN ANALISIS DATA

Bab ini merupakan intisari dari bab-bab sebelumnya yang terdiri dari penerapan metode GEE Campuran pada data *longitudinal* Indeks Pembangunan Manusia (IPM) 27 kabupaten dan Kota di Provinsi Jawa Barat.

BAB V : PENUTUP

Bab ini berisi penjelasan mengenai beberapa hal yang menjadi kesimpulan atas penelitian yang telah dilakukan serta beberapa saran pengembangan tulisan ini.