

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Regresi Logistik digunakan untuk mengetahui bentuk hubungan antar satu peubah respon dengan satu atau lebih peubah prediktor bersifat kategorik maupun kontinyu. Banyaknya kategori peubah respon dapat bersifat dikotom dan politom. Kadang suatu penelitian menggunakan data dengan dua atau lebih kategori berskala ordinal, sehingga digunakan regresi logistik ordinal. Fahrmeir dan Tutz (1994) menjelaskan bahwa apabila terdapat peubah respon (Y) dengan  $k \geq 2$  kategori berskala ordinal, digunakan model regresi logistik ordinal. Individu yang diklasifikasikan ke dalam dua atau lebih kategori didasarkan pada sebaran multinomial ditemukan pada data longitudinal.

Data longitudinal diperoleh dari hasil pengamatan pada beberapa unit *cross-sectional* selama beberapa periode waktu tertentu. Data *cross-sectional* merupakan hasil pengamatan yang dilakukan pada individu berbeda pada waktu tertentu. Data longitudinal dihasilkan dari penelitian The National Youth Survey tentang penggunaan Marijuana selama 5 tahun sejak 1976 – 1980 oleh 237 responden yang telah berumur 13 tahun pada 1976. Peubah respon yakni frekuensi penggunaan Mariyuana dalam sebulan dengan 3 kategori (tidak pernah, kurang dari sekali dan lebih dari sekali sebulan). Unit *cross-sectional* adalah setiap responden, di mana waktu pengamatan berbeda satu dengan yang lain. Setiap unit *cross-sectional* diamati dalam kurun waktu lima tahun, pengamatan inilah yang disebut pengamatan berulang yang menyebabkan autokorelasi antar waktu pengamatan yang harus diatasi agar model data longitudinal menghasilkan kesimpulan yang tepat.

Liang dan Zeger (1986) pertama kali memperkenalkan *Generalized Estimating Equation* (GEE) sebagai metode pengembangan GLM yang digunakan untuk menduga parameter berdasarkan data yang mengandung autokorelasi. Permasalahan autokorelasi dalam GEE digambarkan pada matriks korelasi di setiap unit *cross-sectional*. Diperlukan pemilihan struktur korelasi yang tepat dan dapat menggambarkan korelasi agar hasil yang didapat tidak bias. Pemilihan struktur korelasi pada GEE menggunakan *Quasi-likelihood under the independence Information Criterion* (QIC). Pada penelitian ini akan diterapkan GEE dan pemilihan struktur korelasi terbaik berdasarkan QIC pada data longitudinal di mana peubah respon bersifat ordinal.

## 1.2 Rumusan Masalah

Permasalahan dalam penelitian ini adalah bagaimana menentukan model *Generalized Estimating Equation* dan pemilihan struktur korelasi terbaik pada data longitudinal di mana peubah respon bersifat ordinal.

## 1.3 Batasan Masalah

Permasalahan yang dibahas dalam penelitian ini dibatasi pada:

1. Data yang digunakan adalah data longitudinal seimbang dan lengkap.
2. Peubah respon bersifat ordinal.

## 1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk membentuk model *Generalized Estimating Equation* dan memilih struktur korelasi terbaik pada data longitudinal di mana peubah respon bersifat ordinal

## 1.5 Manfaat Penelitian

Manfaat penelitian ini adalah penggunaan GEE dalam mengatasi autokorelasi pada data longitudinal dan pemilihan struktur korelasi yang tepat berdasarkan QIC.