

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Dalam aplikasi sosial, kesehatan, pendidikan, dan lainnya, seringkali dijumpai data yang memiliki struktur hirarki, *tercluster*, atau bersarang (*nested*). Struktur hirarki artinya unit-unit pada level yang lebih rendah, tersarang atau *tercluster* dalam unit-unit pada level yang lebih tinggi. Data hirarki disebut juga data *multilevel*.

Sebagai contoh, pada suatu penelitian mengenai pendidikan, data murid-murid yang diteliti tersarang pada sekolah-sekolah. Data populasi yang demikian mempunyai struktur hirarki dengan dua tingkatan, maka dapat disebut sebagai data dua-level. Dalam data dua-level, dikenal istilah level-1 dan level-2. Level-2 biasanya menunjukkan kelompok yang membawahi unit-unit di level-1. Pada masalah di atas, murid merupakan unit level-1 dan sekolah merupakan unit level-2.

Pada data *multilevel*, observasi-observasi dalam grup (unit level-2) yang sama cenderung berkorelasi atau mempunyai karakteristik yang similar dibandingkan dengan observasi-observasi dari grup yang berbeda. Pada contoh data dua-level di atas, dimana murid bersarang dalam sekolah, murid-murid yang bersekolah di sekolah yang sama akan cenderung mempunyai

karakteristik yang hampir sama dibandingkan dengan murid-murid yang bersekolah di sekolah yang berbeda.

Salah satu karakteristik dari data *multilevel* adalah dimungkinkan banyaknya individu dalam setiap *cluster* berbeda-beda sehingga data yang dihasilkan tidak seimbang (*unbalanced*).

Jika masalah multilevel dianalisis menggunakan metode analisis standar, seperti model regresi berganda, ANOVA dan lainnya, maka analisis dapat dilakukan dengan mengumpulkan semua variabel pada level rendah (*disaggregated*) atau pada level tinggi (*aggregated*). Hal ini akan menimbulkan beberapa masalah, diantaranya:

- Jika analisis dilakukan pada level tinggi (*aggregated*), maka informasi di level rendah akan hilang. Akibatnya, power dari pengujian statistik pada level ini juga dapat berkurang karena banyaknya informasi hilang dari level rendah.
- Jika analisis dilakukan pada level rendah (*disaggregated*), maka pengelompokan data diabaikan. Masalah yang akan timbul adalah multikolinieritas sehingga model yang dihasilkan menjadi kurang baik.

Untuk menganalisis data yang mempunyai struktur hirarki maka digunakan model regresi *multilevel* (model *multilevel*). Model regresi *multilevel* merupakan generalisasi model regresi berganda yang dikembangkan untuk menganalisis data ter*cluster* atau bersarang. Perbedaan antara model *multilevel* dengan model regresi berganda adalah, pada model

regresi berganda hanya terdapat satu *error*, sedangkan pada model *multilevel* terdapat lebih dari satu *error*.

Tujuan utama dari model ini adalah untuk mengetahui bagaimana variabel respon dapat dijelaskan oleh variabel penjelas dari setiap level dan untuk mengetahui variabilitas dalam setiap level.

Pada skripsi ini akan dibahas mengenai pengembangan model regresi *multilevel* untuk variabel respon berskala ordinal. Untuk struktur data dua level, model ini biasa dikenal dengan model regresi ordinal dua level. Model ini berguna untuk menganalisis data respon ordinal *tercluster* dan longitudinal. Dalam hal ini variabel respon ordinal yang diketahui, dibentuk dari suatu variabel laten kontinu yang tak diketahui nilainya. Nilai batas kategorik (*threshold*) pada variabel laten perlu diestimasi bersama-sama dengan koefisien regresi ordinal dua level.

Data longitudinal adalah pengamatan berulang yang diambil untuk beberapa individu (subjek) pada beberapa periode waktu tertentu. Pengamatan berulang ini cenderung akan menghasilkan pengamatan-pengamatan yang berkorelasi satu sama lain. Data longitudinal juga dapat dipandang sebagai data *multilevel*. Dalam hal ini, perulangan observasi bersarang dalam individu, individu bersarang dalam grup, dan seterusnya.

Dalam skripsi ini akan dijelaskan lebih lanjut mengenai estimasi parameter pada model regresi ordinal dua level.

1.2 PERUMUSAN MASALAH

Bagaimana mengestimasi parameter-parameter dalam model regresi dua level untuk variabel respon berskala ordinal ?

1.3 TUJUAN PENULISAN

Mempelajari metode *maximum marginal likelihood* (MMLE) untuk mengestimasi parameter-parameter dalam model regresi ordinal dua level.

1.4 PEMBATASAN MASALAH

Struktur data hirarki yang dibahas pada skripsi ini dibatasi hanya sampai struktur data dua level.

1.5 SISTEMATIKA PENULISAN

Penulisan skripsi ini dibagi menjadi lima bab, yaitu :

Bab I : Pendahuluan

Berisi latar belakang masalah, perumusan masalah, tujuan penulisan, pembatasan masalah dan sistematika penulisan.

Bab II : Landasan Teori

Membahas teori-teori dasar yang melandasi materi skripsi ini yaitu, model regresi dua level, metode estimasi *maximum*

likelihood, matriks partisi, *kroncker product*, dekomposisi *Cholesky*, dan *Gauss-Hermite quadrature*.

Bab III : Estimasi Parameter pada Model Regresi Ordinal Dua Level

Bab ini membahas mengenai mencari taksiran dari parameter-parameter pada model regresi ordinal dua level dengan metode estimasi *maximum marginal likelihood* (MMLE).

Bab IV : Contoh Aplikasi Model Regresi Ordinal Dua Level

Membahas penerapan dari model regresi dua level untuk data respon ordinal *tercluster* dan data respon ordinal longitudinal .

Bab V : Penutup

Berisi kesimpulan dan saran.