

BAB I

PENDAHULUAN

1.1 Latar Belakang

Secara alamiah semua objek akan berubah seiring dengan berjalannya waktu. Untuk melihat dan meneliti fenomena ini, banyak peneliti mengumpulkan data dengan cara melakukan observasi atau mengukur objek (unit analisis) yang menjadi perhatian lebih dari satu kali pada waktu yang berbeda. Data yang diperoleh dari penelitian semacam ini disebut sebagai data longitudinal [1].

Data longitudinal adalah data hasil pengukuran berulang untuk satu atau beberapa variabel pada setiap anggota sejumlah subjek atau individu yang sama, yang diamati pada sejumlah titik waktu berbeda [2]. Data longitudinal merupakan salah satu jenis data yang menggabungkan antara pengamatan *cross section* dan *time series*. Pengamatan *cross section* yaitu suatu pengamatan yang dilakukan kepada objek yang banyak pada waktu yang sama, sedangkan pengamatan *time series* berlaku sebaliknya, yaitu suatu pengamatan yang dilakukan pada satu objek dari waktu ke waktu. Informasi yang diambil dari tiap individu, atau variabel dalam penelitian longitudinal biasanya lebih dari satu variabel, yang biasa disebut sebagai variabel respon (variabel dependen) dan variabel penjelas (variabel independen). Salah satu tujuan menggunakan analisis data longitudinal adalah untuk meneliti apakah ada pengaruh antara variabel penjelas terhadap variabel respon, termasuk meneliti pengaruh variabel penjelas terhadap besarnya perubahan terhadap variabel respon. Sehingga pada dasarnya analisis data longitudinal adalah analisis regresi pada data longitudinal [3].

Analisis regresi data longitudinal memiliki beberapa model, yaitu model koefisien konstan (*Common Effect Model*), model efek tetap (*Fixed Effect Model*), model efek acak (*Random Effect Model*) dan model linier efek campuran (*Linear Mixed Effect Model*). Model linier efek campuran merupakan penggabungan dari model efek tetap dan model efek acak, dimana berdasarkan struktur modelnya variabel yang satu berlaku sebagai efek tetap sedangkan variabel lainnya sebagai efek acak.

Dalam menganalisis model linier efek campuran, diperlukan uji kecocokan model untuk mengevaluasi cocok tidaknya model dengan data, nilai observasi yang diperoleh sama atau mendekati dengan yang diharapkan model. Evaluasi terhadap tingkat kecocokan data dengan model memiliki beberapa tahapan, yaitu kecocokan keseluruhan model (*overall model fit*), kecocokan model pengukuran (*measurement model fit*), dan kecocokan model struktural (*structural model fit*) [4]. Kecocokan keseluruhan model (*overall model fit*) terdiri dari *chi-square* (χ^2), *Non-Centrality Parameter* (NCP), *Goodness of Fit Index* (GFI), *Adjust Goodness of Fit Index* (AGFI), *Root Mean Square Error of Approximation* (RMSEA) dan *Parsimonious Normed Fit Index* (PNFI). Kecocokan model pengukuran (*measurement model fit*) terdiri dari nilai *Construct Reability* dan *Variance Extracted*. Sedangkan kecocokan model struktural (*structural model fit*) terdiri dari uji signifikan parameter dan koefisien determinasi (R^2). Berdasarkan struktur model linier efek campuran, hasil estimasi dari model, dan residual antara nilai observasi dan nilai hasil estimasi, maka pada skripsi ini digunakan nilai R^2 untuk meminimumkan residual sebagai uji kecocokan pada model linier efek campuran data longitudinal.

Koefisien determinasi (R^2) merupakan suatu ukuran yang penting dalam regresi, karena ukuran ini dapat menginformasikan baik atau tidaknya model regresi yang terestimasi. Atau dengan kata lain, angka tersebut dapat mengukur seberapa dekat garis regresi yang terestimasi dengan data sesungguhnya [5]. Dalam penelitian sebelumnya oleh McFadden [6] menjelaskan rasio *log-likelihood* R^2 untuk model regresi logistik, dan penelitian oleh Cox dan Snell [7] menjelaskan generalisasi koefisien determinasi untuk model linier umum. Pada saat ini belum ada penelitian yang membahas tentang uji kecocokan R^2 untuk model linier efek campuran data longitudinal.

Sehingga pada penelitian skripsi ini penulis memberikan judul “*MENENTUKAN UJI KECOCOKAN PADA MODEL LINIER EFEK CAMPURAN DATA LONGITUDINAL*”.

1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang, maka yang menjadi pokok permasalahan pada penelitian ini adalah :

1. Bagaimana menentukan bentuk model linier efek campuran data longitudinal ?
2. Bagaimana menentukan estimasi dan uji kecocokan pada model linier efek campuran data longitudinal ?
3. Bagaimana penerapan model linier efek campuran data longitudinal dan uji kecocokan modelnya ?

1.3 Batasan Masalah

Penulis membatasi beberapa masalah yaitu :

1. Metode estimasi yang digunakan adalah *Maximum Likelihood Estimation* (MLE) dan *Restricted Maximum Likelihood* (REML).
2. Uji kecocokan yang digunakan adalah R^2 .

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut :

1. Untuk menentukan bentuk model linier efek campuran data longitudinal.
2. Untuk menentukan estimasi dan uji kecocokan pada model linier efek campuran data longitudinal.
3. Menerapkan model linier efek campuran data longitudinal dan uji kecocokan modelnya pada kasus tingkat kemiskinan di Pulau Jawa.

1.5 Metode Penelitian

Metode penelitian pada skripsi ini memiliki beberapa tahap. Pertama, tahap studi literatur merupakan tahap mengetahui dan memahami materi yang terkait dengan uji kecocokan R^2 pada model linier efek campuran data longitudinal dari berbagai sumber, diantaranya jurnal, skripsi, buku, dan lain-lain. Kedua, tahap penelitian yaitu menganalisis dan melakukan penelitian terhadap R^2 sebagai uji kecocokan pada model linier efek campuran data longitudinal dengan mengaplikasikannya pada data sekunder dan dihitung menggunakan bantuan *software* R versi 3.6.3 dan paket *nlme*.

Adapun beberapa fungsi pada paket `nlme`, yang akan digunakan `lme` untuk mencari nilai estimasi model linier efek campuran data longitudinal yaitu [8]:

```
lme(fixed, data, random, correlation, weights, subset, method, na.action, control, contrast=NULL, keep.data=TRUE).
```

Beberapa penjelasan tentang komponen dalam *script* `lme` :

`object` = objek dari kelas `lme`, mewakili model linier efek campuran data longitudinal.

`fixed` = rumus dua sisi linier memberikan efek tetap bagian dari model.

`data` = kerangka data opsional yang berisi variabel bernama. Secara *default* variabel yang diambil dari lingkungan dari mana `lme` disebut.

`random` = formula yang menggambarkan efek acak.

`correlation` = struktur korelasi antar kelompok.

`weights` = untuk menentukan adanya heteroskedastisitas dari kelompok.

`subset` = pilihan yang menunjukkan subset dari baris pada data.

`method` = untuk mengetahui metode maksimum *likelihood*.

`na.action` = fungsi yang menunjukkan apabila ada data yang kosong.

`control` = nilai kontrol ketika menjalankan algoritma estimasi.

`contrast` = daftar opsional.

1.6 Sistematika Penulisan

Sistematika penulisan skripsi ini terdiri dari lima bab dan daftar pustaka.

BAB I PENDAHULUAN

Bab pendahuluan berisi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, metode penelitian dan sistematika penulisan.

BAB II LANDASAN TEORI

Bab landasan teori berisi teori-teori yang melandasi pembahasan dalam skripsi ini, yang terdiri atas data longitudinal, model efek tetap, model efek acak, metode estimasi, estimasi maksimum *likelihood*, REML dan uji kecocokan model.

**BAB III MENENTUKAN UJI KECOCOKAN PADA MODEL LINIER
EFEK CAMPURAN DATA LONGITUDINAL**

Bab ini berisi pembahasan mengenai model linier efek campuran data longitudinal, estimasi parameter menggunakan metode maksimum *likelihood* dan REML, uji kecocokan menggunakan koefisien determinasi (R^2).

BAB IV STUDI KASUS DAN ANALISA

Bab ini berisi pembahasan utama dari penelitian skripsi mengenai analisis deskriptif, mencari nilai estimasi dari model tingkat kemiskinan di Pulau Jawa, kemudian menentukan uji kecocokan modelnya menggunakan nilai R^2 .

BAB V PENUTUP

Bab ini berisi kesimpulan yang diperoleh dari hasil analisis deskriptif serta berisi saran untuk penelitian kedepannya.

