

**ANALISIS REGRESI MULTILEVEL TERHADAP DATA KEPADATAN
PENDUDUK PROVINSI LAMPUNG TAHUN 2016 DENGAN METODE
*RESTRICTED MAXIMUM LIKELIHOOD***

(Skripsi)

**Oleh
ABDUL KODIR**



**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKADAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
2018**

ABSTRACT

MULTILEVEL REGRESSION ANALYSIS OF POPULATION DENSITY OF LAMPUNG PROVINCE IN 2016 WITH MAXIMUM LIKELIHOOD *RESTRICTED MAXIMUM LIKELIHOOD*

By

Abdul Kodir

Hierarchical data is data collected from two or more levels where lower levels are nested in higher levels. Initially, the analysis was done without observing the information based on the higher levels. It leads to dissatisfaction of the analysis result and creating heteroscedasticity in the error. Thus, in order to resolve the problems, multilevel regression model is used. The purpose of this study is to determine the factors affecting population density in Lampung province in 2016 at district and sub-district level and to know the diversity that can be explained at the district and sub-district level using REML method. The results showed that the best multilevel model is a model without Z variables. Factors affecting the population density in Lampung province is sex ratio, population growth, and the original income of the sub-district. The diversity of population density that can be explained at the district level is 36,73% while for sub-district level is 12,17%.

Key words: hierarchical data, population density, multilevel regression

ABSTRAK

ANALISIS REGRESI MULTILEVEL TERHADAP DATA KEPADATAN PENDUDUK PROVINSI LAMPUNG TAHUN 2016 DENGAN METODE *RESTRICTED MAXIMUM LIKELIHOOD*

Oleh

Abdul Kodir

Data hirarki adalah data yang dikumpulkan dari dua atau lebih level dimana level yang lebih rendah bersarang pada level yang lebih tinggi. Awalnya, analisis yang dilakukan tanpa memperhatikan informasi pada level yang lebih tinggi. Hal ini mengakibatkan ketidakpuasan pada hasil analisisnya dan menimbulkan heteroskedastisitas pada galat. Maka untuk mengatasi masalah tersebut, digunakan model regresi multilevel. Tujuan dari penelitian ini adalah untuk mengetahui faktor-faktor yang mempengaruhi kepadatan penduduk di provinsi Lampung tahun 2016 pada level kabupaten dan level kecamatan dan mengetahui keragaman yang dapat dijelaskan di level kabupaten dan di level kecamatan dengan menggunakan metode REML. Dari hasil analisis diperoleh bahwa model multilevel terbaik yaitu model multilevel yang tidak mengikutsertakan variabel Z. Faktor-faktor yang mempengaruhi kepadatan penduduk di provinsi Lampung yaitu rasio jenis kelamin, pertambahan penduduk, dan pendapatan asli kecamatan. Keragaman dari kepadatan penduduk yang dapat dijelaskan pada level kabupaten adalah sebesar 36,73% sedangkan untuk level kecamatan adalah sebesar 12,17%.

Kata kunci: data hirarki, kepadatan penduduk, regresi multilevel

**ANALISIS REGRESI MULTILEVEL TERHADAP DATA KEPADATAN
PENDUDUK PROVINSI LAMPUNG TAHUN 2016 DENGAN METODE
*RESTRICTED MAXIMUM LIKELIHOOD***

Oleh

Abdul Kodir

Skripsi

Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA SAINS

pada

Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Lampung



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2018**

Judul Skripsi

: **ANALISIS REGRESI MULTILEVEL
TERHADAP DATA KEPADATAN
PENDUDUK PROVINSI LAMPUNG TAHUN
2016 DENGAN METODE RESTRICTED
MAXIMUM LIKELIHOOD**

Nama Mahasiswa

: **Abdul Kodir**

Nomor Pokok Mahasiswa : 1417031002

Program Studi

: Matematika

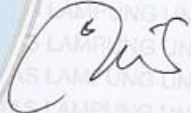
Fakultas

: Matematika dan Ilmu Pengetahuan Alam



1. Komisi Pembimbing


Ir. Netti Herawati, M.Sc., Ph.D.
NIP. 19650125 199003 2 001


Drs. Eri Setiawan, M.Si.
NIP. 19581101 198803 1 002

2. Ketua Jurusan Matematika


Prof. Dra. Wamilliana, M.A., Ph.D.
NIP. 19631108 198902 2 001

MENGESAHKAN

1. Tim Penguji

Ketua : **Ir. Netti Herawati, M.Sc., Ph.D.** 

Sekretaris : **Drs. Eri Setiawan, M.Si** 

Penguji
Bukan Pembimbing : **Prof. Drs. Mustofa Usman, M.A., Ph.D.** 

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Prof. Warsito, S.Si., D.E.A., Ph.D.

NIP. 19710212 199512 1 001

Tanggal Lulus Ujian Skripsi : **13 Maret 2018**

PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini :

Nama : **Abdul Kodir**
Nomor Pokok Mahasiswa : **1417031002**
Jurusan : **Matematika**
Judul Skripsi : **Analisis Regresi Multilevel Terhadap Data
Kepadatan Penduduk Provinsi Lampung
Tahun 2016 dengan Metode *Restricted
Maximum Likelihood***

Dengan ini menyatakan bahwa penelitian ini adalah hasil pekerjaan saya sendiri dan apabila kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 13 Maret 2018

Penulis



Abdul Kodir

RIWAYAT HIDUP

Penulis bernama lengkap Abdul Kodir, anak kelima dari Bapak Darmedi dan Ibu Istinah. Penulis dilahirkan di Merbau Mataram pada tanggal 25 Februari 1996.

Penulis menyelesaikan pendidikan di Sekolah Dasar Negeri 1 Triharjo Kabupaten Lampung Selatan pada tahun 2008, Sekolah Menengah Pertama Negeri 1 Tanjung Bintang Kabupaten Lampung Selatan pada tahun 2011, dan Sekolah Menengah Atas Negeri 1 Tanjung Bintang Kabupaten Lampung Selatan pada tahun 2014. Pada tahun 2014 penulis diterima sebagai mahasiswa S1 di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN).

Sebagai bentuk aplikasi bidang ilmu di dunia kerja, penulis melaksanakan Praktik Kerja Lapangan di Kantor Badan Kependudukan dan Keluarga Berencana Nasional Provinsi Lampung. Sebagai bentuk aplikasi bidang ilmu kepada masyarakat, penulis melaksanakan Kuliah Kerja Nyata di Desa Antar Brak, Kecamatan Limau, Tanggamus.

KATA INSPIRASI

“Dan barangsiapa yang bertakwa kepada Allah, niscaya Allah menjadikan baginya kemudahan dalam urusannya.”

(Q.S. At-Talaq: 4)

“Dan Allah mengeluarkan kamu dari perut ibumu dalam keadaan tidak mengetahui sesuatu apapun, dan Dia memberimu pendengaran, penglihatan, dan hati agar kamu bersyukur.”

(Q.S. An-Nahl: 78)

“Saya memohon kebijakan, dan Allah memberi saya persoalan untuk diselesaikan. Saya memohon kemakmuran, dan Allah memberi saya otak dan tenaga untuk bekerja. Saya memohon keteguhan hati, dan Allah memberi saya bahaya untuk diatasi. Saya memohon cinta, dan Allah memberi saya orang-orang bermasalah untuk ditolong. Saya memohon kemurahan hati, dan Allah memberi kesempatan-kesempatan. Saya tidak memperoleh yang saya inginkan, saya mendapatkan segala yang saya butuhkan.”

(H.R. Ahmad, Ibnu Hibban, dan Al Baghawiy)

“Ada Allah.”

(Abdul Kodir)

PERSEMBAHAN

Alhamdulillah, puji syukur kehadirat Allah SWT yang Maha Pengasih lagi Maha Penyayang. Dengan segala kerendahan hati penulis persembahkan skripsi ini kepada:

Kedua orangtuaku yang selalu berkorban, membimbing, selalu memberikan semangat, rela menjadi pendengar yang baik dan mendoakan setiap waktu untuk keberhasilan penulis.

Untuk kakak-kakak dan adikku tersayang yang selalu memberikan dukungan dan semangat serta do'a yang tak pernah henti untukku.

Untuk sahabat-sahabat terbaikku, terimakasih untuk semua kebahagiaan dan keceriaan yang telah kalian berikan. Terimakasih juga telah selalu ada untuk penulis.

SANWACANA

Puji dan syukur penulis panjatkan kehadirat Allah SWT, yang telah memberikan rahmad dan hidayah-Nya kepada penulis, sehingga penulis dapat menyelesaikan skripsi yang berjudul “ **Analisis Regresi Multilevel Terhadap Data Kepadatan Penduduk Provinsi Lampung Tahun 2016 dengan Metode *Restricted Maximum Likelihood*** ”.

Penulis menyadari bahwa skripsi ini tidak akan terwujud tanpa adanya, bantuan bimbingan, dan doa dari berbagai pihak sehingga skripsi ini dapat terselesaikan.

Dengan segala kerendahan hati penulis mengucapkan terima kasih kepada:

1. Ibu Ir. Netti Herawati, M.Sc., Ph.D., selaku dosen pembimbing utama yang memberikan motivasi, bimbingan, pengarahan, dan saran kepada penulis dalam menyelesaikan skripsi.
2. Bapak Drs. Eri Setiawan, M.Si., selaku pembimbing kedua yang memberikan saran, solusi serta pembelajaran yang sangat bermanfaat bagi penulis.
3. Bapak Prof. Drs. Mustofa Usman, M.A., Ph.D., selaku pembahas dan penguji skripsi yang telah memberikan evaluasi dan saran bagi perbaikan skripsi penulis.
4. Ibu Ir. Netti Herawati, M.Sc., Ph.D., selaku dosen pembimbing akademik.

5. Ibu Prof. Dra. Wamiliana, M.A., Ph.D., selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Bapak Prof. Warsito, S.Si., D.E.A., Ph.D., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam.
7. Seluruh dosen, staf dan karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
8. Ayah, ibu, dan seluruh keluarga besar penulis yang telah membantu dan memberi dukungan kepada penulis.
9. Teman-teman angkatan Jurusan Matematika 2014 dan Keluarga Besar Matematika Unila.
10. Seluruh pihak yang telah membantu penulis dalam menyelesaikan kuliah.

Penulis menyadari bahwa dalam penulisan skripsi ini masih jauh dari kesempurnaan, oleh karena itu kritik dan saran sangat penulis harapkan. Semoga skripsi ini dapat bermanfaat bagi penulis dan bagi para pembaca.

Bandar Lampung, 13 Maret 2018

Penulis

Abdul Kodir

DAFTAR ISI

Halaman

DAFTAR TABEL	v
DAFTAR GAMBAR	vi
I. PENDAHULUAN	1
1.1 Latar Belakang dan Masalah	1
1.2 Tujuan Penelitian	2
1.3 Manfaat Penelitian	2
II. TINJAUAN PUSTAKA	3
2.1 Analisis Regresi Linier	3
2.2 Struktur Data Model Multilevel	4
2.3 Model Regresi Multilevel	5
2.4 Uji Asumsi	7
2.4.1 Uji Normalitas	7
2.4.2 Uji Multikolinieritas	8
2.5 Penduga Parameter	8
2.6 Pengujian Hipotesis	12
2.7 Pemilihan Model Terbaik	13
2.8 Koefisien Korelasi Interklas	14
2.9 Keragaman Model	14
III. METODOLOGI PENELITIAN	16
3.1 Waktu dan Tempat Penelitian	16
3.2 Data Penelitian.....	16
3.3 Metode Penelitian	17

IV. HASIL DAN PEMBAHASAN	19
4.1 Deskriptif Data	19
4.2 Pengujian Asumsi	21
4.2.1 Uji Normalitas	21
4.2.2 Uji Multikolinieritas	23
4.3 Regresi Linier Berganda	25
4.4 Analisis Regresi Multilevel	27
4.4.1 Analisis Regresi Multilevel Tanpa Variabel Z	27
4.4.2 Analisis Regresi Multilevel dengan Variabel Z	28
4.5 Pemilihan Model Terbaik	31
4.6 Koefisien Korelasi Interklas	33
4.7 Koefisien Determinasi	34
V. KESIMPULAN	36
DAFTAR PUSTAKA	38
LAMPIRAN	

DAFTAR TABEL

Tabel	Halaman
1. Struktur Data Multilevel dengan 2 Level.....	5
2. Statistik Deskriptif	19
3. Uji Normalitas dengan Kolmogorov-Smirnov.....	22
4. Uji Normalitas dengan Kolmogorov-Smirnov.....	23
5. Nilai VIF pada Variabel Bebas Level 1	23
6. Nilai VIF pada Variabel Bebas Level 2	24
7. Hasil Dugaan Parameter Analisis Regresi Berganda.....	25
8. Hasil Dugaan Parameter Analisis Regresi Berganda Tanpa Variabel X_1	26
9. Hasil Dugaan Parameter Analisis Regresi Multilevel Tanpa Variabel Z	27
10. Hasil Dugaan Parameter Analisis Regresi Multilevel dengan Variabel Z.....	28
11. Hasil Dugaan Parameter Analisis Regresi Multilevel dengan Variabel Z secara <i>Backwise</i>	30
12. Nilai Deviasi Setiap Model	31
13. Nilai Dugaan Parameter Ragam Tanpa Variabel Bebas	34
14. Hasil Pendugaan Ragam Model.....	35

DAFTAR GAMBAR

Gambar	Halaman
1. <i>Normal Probability Plot</i> Variabel Kepadatan Penduduk.....	21
2. <i>Normal Probability Plot</i> Variabel Kepadatan Penduduk Hasil Transformasi ..	22

I. PENDAHULUAN

1.1. Latar Belakang dan Masalah

Data yang terstruktur hirarki merupakan data yang timbul karena individu-individu terkumpul dalam kelompok-kelompoknya, dimana individu-individu dalam kelompok yang sama memiliki karakteristik yang cenderung sama. Struktur hirarki mengindikasikan bahwa data yang dianalisis berasal dari beberapa level, dimana level yang lebih rendah tersarang pada level yang lebih tinggi. Analisis dari data seperti ini disebut sebagai analisis data multilevel. Dalam beberapa sumber penelitian model regresi multilevel dikenal dengan *Hierarchical Linear Model* (Harlan, 2016).

Metode yang digunakan untuk menduga parameter model regresi multilevel cukup bervariasi, diantaranya metode *Maximum Likelihood* dan *Restricted Maximum Likelihood*. Pendugaan parameter pada analisis regresi biasanya menggunakan metode *Maximum Likelihood*. Namun metode *Maximum Likelihood* menghasilkan penduga ragam yang bias, sehingga untuk mendapatkan penduga ragam yang tak bias dilakukan pendugaan dengan metode *Restricted Maximum Likelihood* (West, dkk., 2007).

Oleh karena itu, pada penelitian ini akan dibahas mengenai faktor-faktor yang mempengaruhi kepadatan penduduk di provinsi Lampung tahun 2016 pada level kecamatan dan level kabupaten menggunakan analisis regresi multilevel dengan metode *Restricted Maximum Likelihood*.

1.2. Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Menentukan model regresi multilevel terhadap faktor-faktor yang mempengaruhi kepadatan penduduk di provinsi Lampung tahun 2016.
2. Menguraikan keragaman yang dapat dijelaskan oleh level kecamatan dan level kabupaten terhadap kepadatan penduduk di provinsi Lampung tahun 2016.

1.3. Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut:

1. Dapat mengetahui model regresi multilevel terhadap faktor-faktor yang mempengaruhi kepadatan penduduk di provinsi Lampung tahun 2016.
2. Dapat mengetahui keragaman yang dapat dijelaskan oleh level kecamatan dan level kabupaten terhadap kepadatan penduduk di provinsi Lampung tahun 2016.

II. TINJAUAN PUSTAKA

2.1. Analisis Regresi Linier

Analisis regresi merupakan suatu model yang digunakan untuk menganalisis hubungan antara variabel terikat (Y) dengan variabel bebas (X). Menurut Gujarati (2006) ada tiga tujuan dari model regresi yaitu yang pertama menjelaskan pola hubungan sebab akibat yang terjadi antar variabel terikat dengan variabel bebas, tujuan kedua untuk mengetahui kontribusi tiap variabel bebas untuk menjelaskan variabel terikat dan yang ketiga adalah memprediksi nilai variabel terikat untuk beberapa nilai variabel bebas tertentu.

Menurut Draper dan Smith (1992), persamaan model regresi dapat ditunjukkan sebagai berikut

$$Y = a + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon_i \quad (2.1)$$

dengan

Y = variabel terikat

X = variabel bebas

a = intersep atau konstanta regresi

β = slope atau koefisien regresi

ε = galat yang diasumsikan IIDN.

i = indeks ($i = 1, 2, \dots, n$)

2.2. Struktur Data Model Multilevel

Data berstruktur hirarki yaitu data yang terdiri dari unit-unit yang diobservasi bersarang atau terkelompokkan dalam unit level yang lebih tinggi. Data hirarki disebut juga data multilevel (Dewi, 2008).

Pada data multilevel, variabel terikat diukur pada level 2 atau di tingkat terendah saja, sedangkan variabel bebas diukur pada level 2 dan level yang lebih tinggi (Goldstein, 2011). Jika terdapat data multilevel dengan 2 level sebagai berikut.

- a. Terdapat m buah kelompok, dimana banyaknya pengamatan pada masing- masing kelompok sebesar n_1, n_2 dan seterusnya sampai dengan n_m .
- b. Variabel terikat diamati hanya pada level 2
- c. Variabel bebas dengan 2 level yaitu
 - Level 1 : Variabel bebas Z_1 sampai Z_q
 - Level 2 : Variabel terikat X_1 sampai X_p

Gambaran dari struktur data model regresi 2 level dapat dilihat pada Tabel 1 berikut.

Tabel 1. Struktur Data Multilevel dengan 2 Level.

Kelompok	Pengamatan	Variabel Terikat (Y)	Variabel bebas Level 1			Variabel Bebas Level 2		
			Z ₁	...	Z _q	X ₁	...	X _p
1	1	y ₁₁	Z ₁₁		Z _{q1}	x ₁₁₁		x _{p11}
	2	y ₂₁				x ₁₂₁		x _{p21}
	:	:				:		:
	n ₁	y _{n11}				x _{1n11}		x _{pn11}
2	1	y ₁₂	Z ₁₂		Z _{q2}	x ₁₁₂		x _{p12}
	2	y ₂₂				x ₁₂₂		x _{p22}
	:	:				:		:
	n ₂	y _{n22}				x _{1n22}		x _{pn22}
:	:	:	:	...	:	:	...	:
M	1	y _{1m}	Z _{1m}		Z _{qm}	x _{11m}		x _{p1m}
	2	y _{2m}				x _{12m}		x _{p2m}
	:	:				:		:
	n _m	y _{nmm}				x _{1nmm}		x _{pnmm}

2.3. Model Regresi Multilevel

Model regresi multilevel adalah model linier dengan variabel terikat yang nilainya tidak hanya dipengaruhi oleh variabel dengan efek tetap saja namun juga dipengaruhi oleh variabel dengan efek acak. Istilah model regresi multilevel menangkap dua hal dalam mendefinisikan model. Hal pertama, data yang sesuai untuk model berstruktur hirarki, dengan unit level 3 tersarang dalam unit level 2, unit level 2 tersarang di dalam unit level 1, dan seterusnya. Sedangkan hal kedua, parameter model tersebut terlihat seperti memiliki struktur hirarki (Raudenbush dan Bryk, 2002).

Bentuk paling dasar dari analisis regresi multilevel adalah analisis regresi 2 level.

Model regresi 2 level adalah model dua level dimana level 1 adalah data kelompok dan level 2 adalah data individu (West, dkk., 2007). Representasi persamaan regresi 2 level adalah sebagai berikut:

- Untuk model level 1

$$\beta_{pj} = \gamma_{p0} + \sum_{q=1}^s \gamma_{pq} Z_{qj} + u_{pj} \quad (2.2)$$

dengan:

β_{pj} = variabel terikat untuk kelompok ke- j level 1

γ_{p0} = intersep level 1

γ_{pq} = slope ke- p untuk variabel bebas ke- q level 1

Z_{qj} = variabel bebas ke- q untuk kelompok ke- j level 1

u_{oj} = galat level 1, diasumsikan berdistribusi $N(0, \sigma_u^2)$

j = indeks pada level 1 ($j = 1, 2, \dots, m$)

q = banyak variabel bebas level 1 ($q = 1, 2, \dots, s$)

p = banyak variabel bebas level 2 ($p = 1, 2, \dots, r$)

- Untuk model level 2:

$$Y_{ij} = \beta_{0j} + \sum_{p=1}^r \beta_{pj} X_{pij} + \varepsilon_{ij} \quad (2.3)$$

dengan:

Y_{ij} = variabel terikat untuk individu ke- i dalam kelompok ke- j

β_{0j} = intersep

β_{pj} = slope acak untuk variabel bebas ke- q pada kelompok ke- j

Z_{qj} = variabel bebas ke- q untuk kelompok ke- j

X_{pij} = variabel bebas ke- p untuk individu ke- i dalam kelompok ke- j

ε_{ij} = galat level 2, diasumsikan berdistribusi $N(0, \sigma_{\varepsilon}^2)$

j = indeks pada level 1 ($j = 1, 2, \dots, m$)

i = indeks pada level 2 ($i = 1, 2, \dots, n_j$)

q = banyak variabel bebas level 1 ($q = 1, 2, \dots, s$)

p = banyak variabel bebas level 2 ($p = 1, 2, \dots, r$)

2.4. Uji Asumsi

2.4.1. Uji Normalitas

Uji normalitas adalah uji untuk melihat apakah data berdistribusi normal atau tidak. Uji statistik yang sering digunakan untuk menguji normalitas adalah uji Kolmogorov-Smirnov. Uji Kolmogorov-Smirnov bekerja dengan cara membandingkan dua distribusi atau sebaran data, yaitu distribusi yang dihipotesiskan dan distribusi yang teramati. Apabila distribusi yang teramati mirip dengan distribusi yang dihipotesiskan, maka dapat disimpulkan bahwa data yang diamati memiliki distribusi atau sebaran normal (Kurniawan, 2008). Selain uji Kolmogorov-Smirnov, uji normalitas juga dapat dilakukan dengan melihat *Normal Probability Plot*.

2.4.2. Uji Multikolinieritas

Uji multikolinieritas adalah uji untuk melihat ada atau tidaknya korelasi yang tinggi antara variabel-variabel bebas dalam suatu model regresi linier berganda.

Uji statistik yang sering digunakan untuk menguji gangguan multikolinieritas adalah *Variance Inflation Factor* (VIF). VIF dapat menginterpretasikan akibat dari korelasi antar variabel bebas ke-*i* pada ragam penduga koefisien regresi.

Perhitungan VIF sebagai berikut:

$$VIF = \frac{1}{1-R_i^2}; \text{ dimana } i = 1, 2, \dots, p - 1. \quad (2.4)$$

Pada uji statistik VIF, apabila nilai VIF lebih besar dari 10 mengindikasikan adanya multikolinieritas yang serius (Zulmi, 2011).

2.5. Penduga Parameter

Metode pendugaan parameter yang dapat digunakan pada pemodelan regresi multilevel adalah metode *Maximum Likelihood* atau *Restricted Maximum Likelihood* (Goldstein, 1995). Metode *Maximum Likelihood* mengestimasi parameter dengan cara memaksimumkan fungsi *likelihood* (Hox, 2010). Berikut ini adalah pendugaan parameter dengan metode *Maximum Likelihood*.

Diberikan persamaan model regresi multilevel sebagai berikut :

$$Y = X\beta + Zu + \varepsilon \quad (2.5)$$

dimana $E(\boldsymbol{\varepsilon}) = \mathbf{0}$, $cov(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}_n = \boldsymbol{\Sigma}_t$ dan $E(\mathbf{u}_i) = \mathbf{0}$, $cov(\mathbf{u}_i) = \sigma^2 \mathbf{I}_{r_i} = \mathbf{D}$

atau $\begin{bmatrix} \mathbf{u}_i \\ \boldsymbol{\varepsilon} \end{bmatrix} \sim N_{mq+n} \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{G} & \mathbf{0}_{mq+n} \\ \mathbf{0}_{mq+n} & \mathbf{R} \end{bmatrix} \right)$.

$$\mathbf{G} = \begin{bmatrix} \mathbf{D} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{D} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{D} \end{bmatrix} \text{ dan } \mathbf{R} = \begin{bmatrix} \boldsymbol{\Sigma}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \boldsymbol{\Sigma}_n \end{bmatrix}$$

dengan:

\mathbf{Y} = vektor variabel terikat ($n \times 1$)

\mathbf{X} = matriks variabel bebas dengan efek tetap ($n \times p$)

$\boldsymbol{\beta}$ = vektor parameter efek tetap ($p \times 1$)

\mathbf{Z} = matriks variabel bebas dengan efek acak ($n \times q$)

\mathbf{u} = vektor parameter efek acak ($q \times 1$)

$\boldsymbol{\varepsilon}$ = vektor galat ($n \times 1$)

\mathbf{D} = matriks kovarian dari efek acak \mathbf{u}

\mathbf{R} = matriks kovarian dari vektor galat $\boldsymbol{\varepsilon}$

Misalkan

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}$$

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}^* \tag{2.6}$$

dengan:

$$\boldsymbol{\varepsilon}^* = \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}$$

$$= [\mathbf{Z} \quad \mathbf{I}_{n \times n}] \begin{bmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{bmatrix}$$

$$= \mathbf{A} \begin{bmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{bmatrix}$$

$$\boldsymbol{\varepsilon}^* \sim N_n(\mathbf{0}, \mathbf{V})$$

dimana

$$\begin{aligned}
 V &= A \begin{bmatrix} G & \mathbf{0} \\ \mathbf{0} & R \end{bmatrix} A^t \\
 &= [Z \quad I_{n \times n}] \begin{bmatrix} G & \mathbf{0} \\ \mathbf{0} & R \end{bmatrix} \begin{bmatrix} Z \\ I_{n \times n} \end{bmatrix} \\
 &= [ZG \quad R] \begin{bmatrix} Z \\ I_{n \times n} \end{bmatrix} \\
 &= ZGZ^t + R
 \end{aligned} \tag{2.7}$$

Vektor parameter kovarian θ diduga oleh $\hat{\theta}_{ML}$ dengan memaksimumkan fungsi *likelihood*.

$$Y = X\beta + Zu + \varepsilon \tag{2.8}$$

dimana $\begin{bmatrix} \mathbf{u}_i \\ \varepsilon \end{bmatrix} \sim N_{mq+n} \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} G & \mathbf{0}_{mq+n} \\ \mathbf{0}_{mq+n} & R \end{bmatrix} \right)$.

$$Y = X\beta + \varepsilon^*$$

dengan $\varepsilon^* \sim N_n(\mathbf{0}, V)$ dengan $V = ZG(\theta)Z^t + R(\theta)$.

Diberikan fungsi *likelihood* sebagai berikut :

$$\ln L(\beta, \sigma^2) = -\frac{1}{2} \{ \ln |V(\theta)| + (Y - X\beta)^t V(\theta)^{-1} (Y - X\beta) \} \tag{2.9}$$

Dengan memaksimumkan fungsi log *likelihood* untuk θ terhadap β maka diperoleh:

$$\hat{\beta}(\theta) = (X^t V(\theta)^{-1} X)^{-1} X^t V(\theta)^{-1} Y \tag{2.10}$$

$$Y \sim N_n(X\beta, V) \text{ dan } \mathbf{u} \sim N_{mq}(\mathbf{0}, G).$$

$$\begin{aligned}
 Cov(Y, \mathbf{u}) &= cov(X\beta + Zu + \varepsilon, \mathbf{u}) \\
 &= cov(X\beta, \mathbf{u}) + cov(Zu, \mathbf{u}) + cov(\varepsilon, \mathbf{u})
 \end{aligned}$$

$$\begin{aligned}
&= \mathbf{0} + \mathbf{ZG} + \mathbf{0} \\
&= \mathbf{ZG}
\end{aligned} \tag{2.11}$$

$$\begin{aligned}
E(\mathbf{u}|\mathbf{Y}) &= E(\mathbf{u}) + cov(\mathbf{u}, \mathbf{Y})[cov(\mathbf{Y})]^{-1}[\mathbf{Y} - E(\mathbf{Y})] \\
&= \mathbf{0} + \mathbf{GZ}^t\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \\
&= \mathbf{GZ}^t\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})
\end{aligned} \tag{2.12}$$

$\mathbf{GZ}^t\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$ adalah *Best Linier Unbiased Predictor* (BLUP) dari \mathbf{u} . Menurut West, dkk. (2007) efek tetap $\boldsymbol{\beta}$ dan efek acak \mathbf{u} diduga oleh

$$\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}) = (\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{V}^{-1}\mathbf{Y}; \mathbf{u} = \mathbf{GZ}^t\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$$

dimana $\hat{\mathbf{V}} = \mathbf{V}(\hat{\boldsymbol{\theta}}_{ML})$.

Penduga yang didapatkan dengan metode *Maximum Likelihood* merupakan penduga yang bias sehingga untuk mendapatkan penduga yang tak bias perlu dilakukan suatu modifikasi dari prosedur pendugaannya. Prosedur pendugaan *Maximum Likelihood* akan menghasilkan penduga tak bias apabila dilakukan modifikasi dalam fungsi *profile log-likelihood* yang disebut fungsi *Restricted log-likelihood* berikut ini

$$\ln L(\boldsymbol{\beta}, \boldsymbol{\sigma}^2) = -\frac{1}{2} \{\ln|\mathbf{V}(\boldsymbol{\theta})| + \ln|\mathbf{X}^t\mathbf{V}(\boldsymbol{\theta})^{-1}\mathbf{X}|\} \tag{2.13}$$

Prosedur ini disebut sebagai *Restricted Maximum Likelihood* atau *Residual*

Maximum Likelihood (West, dkk., 2007).

2.6. Pengujian Hipotesis

Penduga parameter yang didapatkan dari metode *Restricted Maximum Likelihood* selanjutnya digunakan untuk menguji signifikansi parameter pada model regresi multilevel secara individual.

Hipotesis yang digunakan adalah sebagai berikut :

a. Parameter level 1

$$H_0: \gamma_{qj} = 0$$

$$H_1: \gamma_{qj} \neq 0$$

dengan indeks $q = 1, 2, \dots, s$ dan s menyatakan banyak variabel bebas level 1.

Statistik uji yang digunakan adalah statistik Wald sebagai berikut:

$$t = \frac{\hat{\gamma}_{qj}}{\sqrt{v(\hat{\gamma}_{qj})}} \quad (2.14)$$

b. Parameter level 2

$$H_0: \beta_{pj} = 0$$

$$H_1: \beta_{pj} \neq 0$$

dengan indeks $p = 1, 2, \dots, r$ dan r menyatakan banyak variabel bebas level 2.

Statistik uji yang digunakan adalah statistik Wald sebagai berikut:

$$t = \frac{\hat{\beta}_{pj}}{\sqrt{v(\hat{\beta}_{pj})}} \quad (2.15)$$

Dalam hal ini t mengikuti sebaran t student dengan derajat bebas untuk penduga parameter level 2 adalah $i - r - 1$ sedangkan derajat bebas untuk penduga parameter level 1 adalah $j - s - 1$ (Jones & Steenbergen, 2002).

2.7. Pemilihan Model Terbaik

Dalam regresi multilevel, pemilihan model terbaik dilakukan dengan melihat nilai deviasi. Deviasi merupakan suatu ukuran yang dapat digunakan untuk menentukan kesesuaian suatu model. Menurut Tantular (2009), untuk memilih model regresi multilevel yang terbaik dapat menggunakan uji rasio *likelihood* atau disebut juga sebaran deviasi yaitu ukuran untuk menentukan cocok tidaknya suatu model. Perhitungan untuk pengujian ini adalah selisih nilai deviasi antara dua model atau *diff* yaitu sebagai berikut:

$$Diff = -2 \log_e \left(\frac{\lambda_0}{\lambda_1} \right) \quad (2.16)$$

dengan

λ_0 = nilai deviasi untuk *null model*

λ_1 = nilai deviasi untuk *full model*

Menurut Hox (2010), semakin kecil nilai deviasi pada model tersebut maka model tersebut dikatakan semakin cocok dan apabila nilai $diff > \chi^2_{(a,db)}$ maka tolak H_0 , dimana db merupakan selisih jumlah parameter dari kedua model. Sehingga dapat disimpulkan bahwa efek acak signifikan, artinya terdapat keragaman atau variasi variable dependen yang signifikan antarkelompok.

2.8. Koefisien Korelasi Interklas

Jika data yang dimiliki adalah data multilevel, maka regresi multilevel dapat digunakan untuk memberikan nilai dugaan bagi korelasi intraklas. Model yang digunakan untuk tujuan ini adalah model yang tidak memiliki variabel bebas dalam setiap levelnya, yang dikenal sebagai *null model*.

Dengan menggunakan *null model* korelasi intraklas ρ dapat diformulasikan sebagai berikut:

$$\rho = \frac{\sigma_{u_0}^2}{\sigma_{u_0}^2 + \sigma_{e_0}^2} \quad (2.17)$$

dengan $\sigma_{u_0}^2$ adalah ragam dari galat pada level tertinggi u_{0j} dan $\sigma_{e_0}^2$ adalah ragam dari galat pada level terendah. Korelasi intraklas (ρ) menunjukkan proporsi keragaman yang dijelaskan oleh struktur kelompok dalam populasi, yang dapat juga diinterpretasikan sebagai korelasi antara dua individu yang dipilih secara acak yang berada dalam kelompok yang sama (Hox, 2002).

2.9. Keragaman Model

Keragaman variable terikat yang dapat dijelaskan oleh variable bebas dalam model disebut koefisien determinasi. Koefisien determinasi digunakan untuk mengukur seberapa besar keragaman variabel terikat yang dapat dijelaskan oleh model yang ditetapkan. Pada model multilevel akan didapatkan koefisien determinasi lebih dari satu. Menurut Hox (2002) koefisien determinasi akan didapat pada masing-masing level.

a. Level 1 :

$$R_2^2 = 1 - \frac{\hat{\sigma}_{u0q}^2}{\hat{\sigma}_{u0}^2} \quad (2.18)$$

dengan:

$\hat{\sigma}_{u0q}^2$ = penduga ragam dari galat pada level 1 dengan q variabel bebas

$\hat{\sigma}_{u0}^2$ = penduga ragam dari galat pada level 1 tanpa variabel bebas.

b. Level 2 :

$$R_1^2 = 1 - \frac{\hat{\sigma}_{ep}^2}{\hat{\sigma}_{e0}^2} \quad (2.19)$$

dengan:

$\hat{\sigma}_{ep}^2$ = penduga ragam dari galat pada level 2 dengan p variabel bebas

$\hat{\sigma}_{e0}^2$ = penduga ragam dari galat pada level 2 tanpa variabel bebas.

III. METODOLOGI PENELITIAN

3.1. Waktu dan Tempat Penelitian

Penelitian ini dilakukan di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung pada semester ganjil tahun ajaran 2017/2018.

3.2. Data Penelitian

Data yang akan digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik Provinsi Lampung tahun 2016. Unit analisis pada penelitian level 1 adalah 13 Kabupaten di Provinsi Lampung. Sedangkan unit analisis pada penelitian level 2 adalah 188 kecamatan yang ada di Provinsi Lampung. Variabel yang diteliti adalah kepadatan penduduk, rasio jenis kelamin, pertumbuhan penduduk, pendapatan asli kecamatan, indeks pembangunan manusia, dan laju pertumbuhan ekonomi.

Pada penelitian ini terdiri dari variabel bebas dan variabel terikat, karena data yang digunakan merupakan data multilevel, maka pada variabel bebas terdapat data level 1 (kabupaten) dan data level 2 (kecamatan). Berikut ini penjelasan variabel yang digunakan:

1. Variabel terikat : Kepadatan penduduk (Y)
2. Variabel bebas :
 - a. Level 1 (Kabupaten) : Indeks pembangunan manusia (Z_1) dan laju pertumbuhan ekonomi (Z_2).
 - b. Level 2 (Kecamatan) : Rasio jenis kelamin (X_1), pertumbuhan penduduk (X_2) dan pendapatan asli kecamatan (X_3).

3.3. Metode Penelitian

Penelitian ini dilaksanakan dengan studi pustaka yaitu dengan pengkajian secara teoritis dan praktik komputasi. Langkah-langkah penelitian yang akan dilakukan dalam penelitian ini sesuai dengan tujuan yang ingin dicapai yaitu.

1. Menginput data dari Badan Pusat Statistik tahun 2016 ke *Microsoft Excel* dan mengelompokkannya berdasarkan level kecamatan dan level kabupaten.
2. Melakukan analisis deskriptif pada masing-masing variabel pada level kecamatan dan level kabupaten.
3. Melakukan uji asumsi pada data.
 - a. Uji normalitas menggunakan uji Kolmogorov-Smirnov dan *Normal Probability Plot*.
 - b. Uji multikolinieritas dengan melihat nilai *Variance Inflation Factor* (VIF).

4. Memodelkan menggunakan analisis regresi multilevel untuk mengetahui faktor-faktor yang mempengaruhi kepadatan penduduk di provinsi Lampung tahun 2016 pada level kecamatan dan level kabupaten dengan metode *Restricted Maximum Likelihood*.
5. Membandingkan model-model regresi multilevel menggunakan nilai deviasi dan *diff* untuk menentukan model terbaik.
6. Menghitung nilai koefisien korelasi interklas.
7. Menguraikan keragaman yang dapat dijelaskan oleh level kecamatan dan level kabupaten terhadap kepadatan penduduk di provinsi Lampung tahun 2016.

V. KESIMPULAN

Berdasarkan hasil dan pembahasan pada bab sebelumnya, diperoleh kesimpulan sebagai berikut:

1. Faktor-faktor yang mempengaruhi kepadatan penduduk di Provinsi Lampung tahun 2016 adalah rasio jenis kelamin, pertumbuhan penduduk dan pendapatan asli kecamatan.
2. Berdasarkan hasil analisis regresi multilevel diperoleh model regresi yaitu
kepadatan penduduk = $11,235786 - 4,317228$ rasio jenis kelamin +
 $0,047744$ penambahan penduduk – $0,086954$
pendapatan asli kecamatan.

Artinya, jika variabel rasio jenis kelamin, pertumbuhan penduduk dan laju pertumbuhan ekonomi bernilai nol, maka kepadatan penduduk bernilai 11,23578. Untuk setiap penambahan rasio jenis kelamin sebesar satu satuan akan menyebabkan menurunnya kepadatan penduduk sebesar -4,317228, variabel penambahan penduduk akan menyebabkan meningkatnya kepadatan penduduk sebesar 1,247063 dan untuk setiap penambahan pendapatan asli kecamatan sebesar satu satuan akan menyebabkan menurunnya kepadatan penduduk sebesar -0,086954.

3. Nilai korelasi interklas sebesar 41,53% menunjukkan proporsi ragam pada level kabupaten sebesar 41,53%. Artinya tanpa mempertimbangkan faktor-faktor yang mempengaruhi variabel kepadatan penduduk, 41,53% proporsi keragaman variabel kepadatan penduduk dapat dijelaskan oleh variasi antar kabupaten. Selain itu dapat dijelaskan pula bahwa korelasi dua kecamatan dalam satu kabupaten yaitu 0,4153.
4. Keragaman yang dapat dijelaskan antarkabupaten adalah sebesar 36,73% artinya besarnya pengaruh perbedaan kabupaten adalah sebesar 36,73%. Sedangkan keragaman yang dapat dijelaskan antarkecamatan adalah sebesar 12,17%, artinya besarnya pengaruh variabel rasio jenis kelamin, penambahan penduduk dan pendapatan asli kecamatan terhadap kepadatan penduduk adalah sebesar 12,17%.

DAFTAR PUSTAKA

- Badan Pusat Statistik. 2017. *Provinsi Lampung dalam Angka 2017*. CV. Jaya Wijaya, Bandar Lampung.
- Dewi, A.L. 2008. Estimasi Parameter Logistik Multilevel. Skripsi. Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia, Depok.
- Draper, N., dan Smith. 1992. *Analisis Regresi Terapan*. Ed. Ke-2. PT. Gramedia Pustaka Utama, Jakarta.
- Goldstein. 1995. *Multilevel Statistical Models*. Ed. Ke-2. E-book of Arnold, London.
- Goldstein. 2011. *Multilevel Statistical Models*. Ed. Ke-4. John Wiley & Sons, Ltd., London.
- Gujarati, D. 2006. *Dasar-Dasar Ekonometrika*. Erlangga, Jakarta.
- Harlan, J. 2016. *Analisis Multilevel*. Gunadarma, Jakarta.
- Hox, J.J. 2002. *Multilevel Analysis: Techniques and Applications*. Routledge. Lawrence Erlbaum Associates, Inc., New Jersey.
- Hox, J.J. 2010. *Multilevel Analysis: Techniques and Applications*. Routledge, Great Britain.
- Jones, B.S., dan Steenbergen, M.R. 1997. Modelling Multilevel Data Structures. *American Journal of Political Science*. **46**(1): 218-237.
- Kurniawan, D. 2008. *Regresi Linier*. R Development Core Team, Australia.

- Raudenbush, S.W., dan Bryk, A.S. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Sage, Thousand Oaks.
- Tantular, B. 2009. Penerapan Model Regresi Multilevel Pada Data Pendidikan dan Data Nilai Ujian. Tesis. Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Pertanian Bogor, Bogor.
- West, B.T., Welch, K.B. and Galecki, A.T. 2007. *Linear Mixed Models: A Practical Guide Using Statistical Software*. Chapman and Hall, Boca Raton.
- Zulmi, R. Pengaruh Luas Lahan, Tenaga Kerja, Penggunaan Benih dan Pupuk Terhadap Produksi Padi di Jawa Tengah Tahun 1994-2008. Skripsi. Jurusan Ilmu Ekonomi dan Studi Pembangunan Fakultas Ekonomi Universitas Diponegoro, Semarang.