

ANALISIS REGRESI MULTILEVEL TERHADAP DATA KEPADATAN PENDUDUK PROVINSI LAMPUNG TAHUN 2021 DENGAN METODE *MAXIMUM LIKELIHOOD*

Vinny Adhania Piliang^{1*}, Mustofa Usman², Notiragayu³, Dian Kurniasari⁴

^{1,2,3,4} Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung, Indonesia

*Email korespondensi: vinnyadhp14@gmail.com

Dikirim: dd-mm-yyyy, Diterima: dd-mm-yyyy, Diterbitkan: dd-mm-yyyy

Abstrak

Analisis regresi multilevel digunakan untuk menganalisis struktur data hirarki dan mengukur interaksi yang mungkin terjadi antar variabel pada level yang berbeda. Penelitian ini bertujuan untuk menentukan model regresi multilevel terhadap faktor-faktor yang mempengaruhi kepadatan penduduk di Provinsi Lampung tahun 2021 berdasarkan level kecamatan dan kabupaten serta menguraikan keragaman yang dapat dijelaskan oleh level 1 atau kecamatan dan level 2 kabupaten terhadap kepadatan penduduk di Provinsi Lampung tahun 2021. Metode yang digunakan pada penelitian ini yaitu *Maximum Likelihood* untuk menaksir parameter populasi yang tidak diketahui. Data yang digunakan adalah data kepadatan penduduk provinsi lampung tahun 2021. Hasil penelitian menunjukkan bahwa faktor-faktor yang mempengaruhi kepadatan penduduk di Provinsi Lampung tahun 2021 adalah luas daerah dan rasio jenis kelamin. Model regresi multilevel yang diperoleh yaitu kepadatan penduduk = $141220 - 0,6513 (\text{luas daerah}) - 6,2751 (\text{rasio jenis kelamin}) X_2 + \varepsilon_{ij} + u_{0j}$. Keragaman yang dapat dijelaskan oleh level 1 adalah sebesar 41% dan level 2 sebesar 29%.

Kata kunci : analisis regresi multilevel, *maximum likelihood*, kepadatan penduduk

Abstract

Multilevel regression analysis is used to analyze the hierarchical data structure and measure the possible interactions between variables at different levels. This study aims to determine a multilevel regression model for the factors that affect population density in Lampung Province in 2021 based on the sub-district and district levels and describe the diversity that can be explained by level 1 or sub-district and district level 2 on population density in Lampung Province in 2021. The method used in this study is Maximum Likelihood to estimate unknown population parameters. The data used is data on population density of Lampung Province in 2021. The results show that the factors that affect population density in Lampung Province in 2021 are area size and sex ratio. The multilevel regression model obtained is population density = $141220 - 0.6513 (\text{area}) - 6.2751 (\text{gender ratio}) X_2 + \varepsilon_{ij} + u_{0j}$. The diversity that can be explained by level 1 is 41% and level 2 is 29%.

Keywords: *multilevel regression analysis, maximum likelihood, population density*

1. Pendahuluan

Kepadatan penduduk merupakan salah satu permasalahan yang sampai saat ini belum dapat diatasi oleh Indonesia. Hal ini disebabkan terjadinya perubahan jumlah penduduk setiap tahunnya. Berdasarkan hasil sensus penduduk oleh Badan Pusat Statistik (BPS) pada tahun 2020, penduduk di Provinsi Lampung berjumlah 8.521.201 jiwa dengan kepadatan 5.332 km². Kepadatan Penduduk dapat dilihat berdasarkan beberapa kecamatan yang berada dalam suatu kabupaten. Struktur hirarki data kabupaten dan provinsi terdiri dari kecamatan yang dikelompokkan dalam kabupaten, dan kabupaten dikelompokkan dalam provinsi. Data hierarkis ialah data yang muncul akibat individu-individu dikumpulkan ke dalam kelompok-kelompok, dan individu-individu pada kelompok yang sama mempunyai persamaan karakteristik. Struktur hierarkis menunjukkan data

yang dianalisis bersumber dari berbagai level, dan level yang lebih rendah dinamakan level mikro, sedangkan level yang lebih tinggi secara hirarkis dinamakan level makro [1].

Analisis yang dapat digunakan untuk mengatasi permasalahan yang ada pada data hierarkis adalah analisis regresi multilevel. Analisis regresi multilevel merupakan analisis yang digunakan untuk mengetahui hubungan antara peubah respon dengan peubah prediktor pada level yang berbeda. Peubah prediktor pada regresi multilevel diukur pada dua level, yaitu pada level rendah dan level tinggi [2].

Salah satu metode yang dapat digunakan untuk analisis regresi multilevel adalah *maximum likelihood*. Metode tersebut dapat menghasilkan parameter dugaan yang lebih konsisten dan efisien daripada metode lain [2]. Pendugaan parameter dengan metode *maximum likelihood* dilakukan dengan memaksimalkan fungsi *likelihood*. Pendugaan menggunakan metode *maximum likelihood* menghasilkan parameter-parameter dugaan dan kemudian parameter dugaan tersebut harus diuji signifikansinya.

Beberapa penelitian sebelumnya yang menerapkan analisis regresi multilevel. Salah satunya adalah penelitian yang dilakukan oleh Tantular, Aunuddin, dan Wijayanto (2009), mengenai pemilihan model regresi linier multilevel terbaik. Peneliti terdahulu menganalisis regresi multilevel yaitu dengan metode pendugaan *maximum likelihood*. Hasilnya yaitu level kelas tidak memiliki pengaruh terhadap respon. Selanjutnya Islamiyah dan Subekti (2017), melakukan penelitian mengenai analisis regresi multilevel dan menghasilkan faktor-faktor yang dapat mempengaruhi kepadatan penduduk, seperti jumlah penduduk pindah, penduduk mati, dan luas wilayah.

Oleh karena itu, penulis akan mencoba menerapkan analisis regresi multilevel pada data kepadatan penduduk Provinsi Lampung tahun 2021 dengan metode *maximum likelihood* guna mengetahui faktor-faktor yang mempengaruhi kepadatan penduduk di Provinsi Lampung dan mengetahui keragaman pada level kecamatan maupun kabupaten.

2. Metodologi Penelitian

2.1 Data Hirarki

Data berstruktur hirarki merupakan data yang terdiri dari unit-unit yang diobservasi bersarang atau terkelompokkan dalam unit level yang lebih tinggi. Data hirarki disebut juga data multilevel [3]. Jika analisis dilakukan pada level 1, struktur hirarki atau pengelompokkan data diabaikan, artinya model regresi dibentuk dari seluruh data pengamatan level 1. Keragaman antar unit-unit tidak dapat diketahui secara langsung, tetapi masih bisa diukur dengan membuat model regresi untuk tiap unit level 2 [4].

2.2 Analisis Regresi

Analisis regresi ialah metode statistik untuk membangun model untuk menentukan hubungan kausal antara dua variabel atau lebih [5]. Analisis regresi digunakan untuk menganalisis hubungan antara variabel terikat (Y) dengan variabel bebas (X). Variabel bebas adalah variabel yang mempengaruhi atau menjadi sebab perubahan atau timbulnya variabel tak bebas, sedangkan variabel terikat adalah variabel yang dipengaruhi atau menjadi akibat karena adanya variabel bebas.

2.3 Regresi Multilevel

Regresi multilevel merupakan salah satu metode statistik yang digunakan untuk menganalisis struktur data hirarki [4]. Model multilevel selain dapat menentukan keragaman antar kelompok juga dapat mengukur interaksi yang mungkin terjadi antar variabel pada level yang berbeda. Pada level terendah atau level 1, terdapat peubah respon Y_{ij} dan peubah penjelas X_{ij} serta pada level yang lebih tinggi yaitu level 2, peubah penjelasnya adalah Z_j . Persamaan umum regresi dua level sebagai berikut:

1. Persamaan pada level 1:

$$Y_{ij} = \beta_{0j} + \sum_i^j \beta_{ij} X_{ij} + \varepsilon_{ij} \quad (1)$$

2. Persamaan pada level 2:

$$\beta_{0j} = \gamma_{00} + \sum_i \gamma_{0i} Z_{ij} + u_{0j} \quad (2)$$

Jika persamaan 1 disubstitusikan ke persamaan 2 maka akan membentuk model regresi multilevel dengan intersep acak yang ditunjukkan pada persamaan berikut:

$$Y_{ij} = \gamma_{00} + \gamma_{01}Z_j + \beta_{1j}X_{ij} + u_{0j} + \varepsilon_{ij} \quad (3)$$

Asumsi pada regresi multilevel adalah sebagai berikut:

- Rata-rata sama dengan nol, $E(u_{0j}) = E(\varepsilon_{ij}) = 0$.
- Galat terdiri dari atas dua komponen, yaitu u_{0j} dan ε_{ij} . Ragam galat pada level 1 adalah $Var(\varepsilon_{ij}) = \sigma_e^2$, sedangkan ragam galat pada level 2 yaitu $Var(u_{0j}) = \sigma_{u0}^2$ [2].

2.4 Uji Asumsi

Uji asumsi klasik merupakan sebutan dari beberapa asumsi yang harus terpenuhi dalam analisis regresi. Kesalahan baik kecil maupun besar dapat terjadi dalam melakukan pendugaan parameter. Ketika beberapa asumsi terpenuhi, maka kesalahan pendugaan merupakan kesalahan terkecil [6].

2.4.1 Uji Normalitas

Uji normalitas adalah uji untuk melihat apakah galat berdistribusi normal atau tidak. Uji statistik yang sering digunakan untuk menguji kenormalan suatu galat adalah uji Kolmogorov-Smirnov. Perumusan hipotesis uji Kolmogorov-Smirnov secara matematis sebagai berikut:

H_0 : galat berdistribusi normal

H_1 : galat tidak berdistribusi normal

2.4.2 Uji Multikolinearitas

Uji multikolinearitas adalah uji yang digunakan untuk melihat ada atau tidaknya korelasi yang tinggi antara variabel-variabel independen dalam suatu model regresi. Pendeteksian multikolinearitas antara variabel-variabel independen dapat dilakukan melalui nilai *Variance Inflation Factor* (VIF) yang didefinisikan dalam persamaan berikut:

$$VIF = \frac{1}{1-R_j^2}; \quad j=1,2,\dots,m \quad (4)$$

2.5 Metode Pendugaan Parameter

Metode pendugaan yang dapat digunakan dalam regresi multilevel adalah metode *maximum likelihood* [2]. Metode *Maximum Likelihood* memiliki kelebihan dibandingkan dengan REML yaitu relatif lebih mudah dari segi komputasi dan metode ini sering digunakan untuk mengestimasi efek tetap, sedangkan untuk mengestimasi efek acak lebih baik menggunakan REML [4].

Penduga *maksimum likelihood*nya sebagai berikut:

$$\begin{aligned} \hat{\beta} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \\ \sigma^2 &= \frac{1}{n}\mathbf{Y}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'] \\ \sigma^2 &= \frac{1}{n}\mathbf{Y}'[\mathbf{I} - \mathbf{X}\mathbf{X}^{-}]\mathbf{Y} \end{aligned} \quad (5)$$

2.6 Pengujian Hipotesis

Pendugaan parameter menggunakan metode *maximum likelihood* menghasilkan parameter-parameter dugaan dan kemudian parameter dugaan tersebut harus diuji signifikansinya. Jones dan Steenbergen (2002), menyatakan statistik uji Wald yang digunakan mengikuti sebaran *t-Student* yang didefinisikan berikut:

$$\text{Level 1: } t = \frac{\hat{\beta}_{ij}}{SE(\beta_{ij})} \quad (6)$$

$$\text{Level 2: } t = \frac{\hat{\gamma}_{ij}}{SE(\gamma_{ij})} \quad (7)$$

2.7 Pemilihan Model Terbaik

Null model merupakan model rata-rata tanpa syarat yang memiliki peran penting dalam analisis regresi multilevel. Perhitungan untuk pengujian ini adalah selisih nilai deviasi antara dua model atau *diff* yaitu sebagai berikut:

$$Diff = -2 \log e \left(\frac{\lambda_0}{\lambda_1} \right) \quad (8)$$

2.8 Koefisien Korelasi Intraklas

Korelasi adalah ukuran kedekatan hubungan antara dua variabel [7]. Model yang digunakan untuk merepresentasikan koefisien korelasi antar kelas adalah model tanpa variabel bebas pada setiap levelnya (null model). Nilai Koefisien Korelasi Antar Kelas atau *Interclass Correlation Coefficient* (ICC) dapat didefinisikan dalam persamaan berikut:

$$\rho = \frac{\sigma^2_{e_{ij}}}{\sigma^2_{e_{ij}} + \sigma^2_{u_{0j}}} \quad (9)$$

2.9 Keragaman Model

Keragaman variabel dependen yang dapat dijelaskan oleh variabel bebas dalam model disebut koefisien determinasi. Terdapat dua jenis koefisien determinasi dalam model regresi multilevel, yaitu koefisien determinasi pada level 1 dan level 2. Koefisien determinasi level 1 bertujuan untuk menilai rasio ragam galat terhadap ragam total yang ditunjukkan dalam persamaan berikut:

$$R_1^2 = 1 - \frac{\sigma^2_{e_0}}{\sigma^2_{e_{ij}}} \quad (10)$$

Koefisien determinasi pada level 2 ditunjukkan dalam persamaan berikut:

$$R_2^2 = 1 - \frac{\sigma^2_{u_0}}{\sigma^2_{u_{0j}}} \quad (11)$$

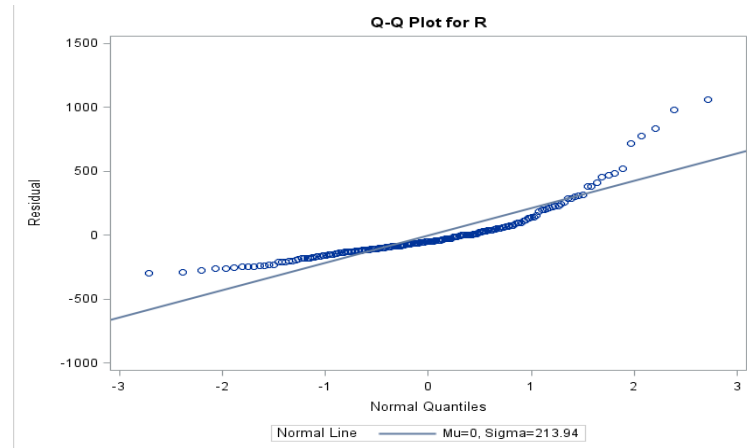
3. Hasil dan Pembahasan

Langkah pertama yang perlu dilakukan yaitu deskripsi data. Pada tahap ini akan dilakukan untuk melihat variasi pada data. Hasil deskripsi data sebagai berikut:

Tabel 1. Deskripsi Data

Variabel	Rata-rata	Standar Deviasi	Minimum	Maximum
Kepadatan Penduduk	300.65	253.14	2.95	1534.55
Luas Daerah	170.77	140.73	25.68	943.7
Rasio Jenis Kelamin	106.25	2.77	99.8	116.6
Indeks Pembangunan Manusia	67.84	1.73	64.04	70.45
Angka Harapan Hidup	69.12	1.6	63.66	70.78
Jumlah Bayi Lahir	11381.15	5740.07	3181	22266
Tingkat Pengangguran Terbuka	3.63	1.00	1.68	5.09

Setelah dilakukan deskripsi data, tahap selanjutnya yaitu melakukan uji normalitas. Uji normalitas dilakukan untuk mengetahui apakah residual berdistribusi normal atau tidak. Uji normalitas dilakukan dengan melihat *normal probability plot* dan uji Kolmogorov-Smirnov Hasil dari uji normalitas sebagai berikut:



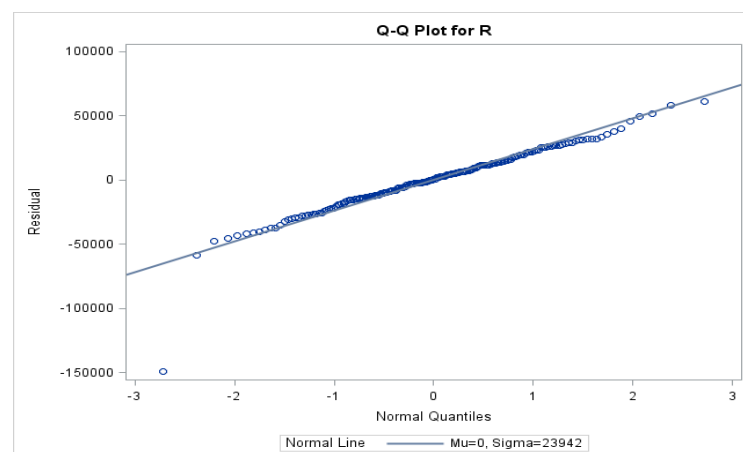
Gambar 1. *Normal Probability Plot*.

Berdasarkan Gambar 1, dapat dilihat bahwa residual tidak tersebar merata di sekitar garis normal. Sehingga dapat diasumsikan residual tidak berdistribusi normal. Selanjutnya dilakukan uji Kolmogorov-Smirnov dengan hasil sebagai berikut:

Tabel 2. Uji Kolmogorov-Smirnov

<i>Tests for Normality</i>				
<i>Test</i>	<i>Statistic</i>		<i>p Value</i>	
Kolmogorov-Smirnov	D	0.162787	Pr > D	<0.0100

Pada Tabel 2 dapat dilihat bahwa *p value* adalah <0,01. Karena *p value* kurang dari 0,05 maka dapat disimpulkan bahwa residual tidak berdistribusi normal. Oleh karena itu, akan dilakukan transformasi dengan menggunakan transformasi logaritma. Adapun hasil yang diperoleh adalah sebagai berikut:



Gambar 2. *Normal Probability Plot Setelah Transformasi*.

Berdasarkan Gambar 2 terlihat bahwa residual tersebar merata di sekitar garis normal sehingga dapat disimpulkan residual berdistribusi normal. Selanjutnya uji kolmogorov-smirnov data transformasi dengan hasil sebagai berikut:

Tabel 3. Uji Kolmogorov-Smirnov Data Transformasi

<i>Tests for Normality</i>				
<i>Test</i>	<i>Statistic</i>		<i>p Value</i>	
Kolmogorov-Smirnov	D	0.0604	Pr > D	0.0905

Pada Tabel 3 dapat dilihat bahwa *p value* sebesar 0,0905. Karena *p value* lebih dari 0,05 maka dapat disimpulkan bahwa residual berdistribusi normal. Selanjutnya melakukan uji multikolinieritas. Uji multikolinearitas dilakukan untuk mengetahui korelasi antara variabel independen dengan cara melihat nilai *Variance Inflation Factor* (VIF). Hasil yang diperoleh sebagai berikut:

a. Nilai VIF Level

Tabel 4. Nilai VIF Level 1

Variabel Independen	VIF
Luas Daerah (X1)	1.09082
Rasio Jenis Kelamin (X2)	1.56967

Adapun hipotesis uji multikolinearitas adalah sebagai berikut:

H_0 : Terdapat multikolinearitas antara variabel independen

H_1 : Tidak terdapat multikolinearitas antara variabel independen

Taraf signifikan : $\alpha = 5\%$

Statistik uji : $VIF = \frac{1}{1-R_j^2}$

Daerah kritis : Tolak H_0 jika $VIF < 10$

Keputusan : Tolak H_0 karena nilai VIF dari masing-masing variabel independen kurang dari 10

Kesimpulan, tidak terdapat multikolinearitas antara variabel pada level 1.

b. Nilai level 2

Tabel 5. Nilai VIF Level 2

Variabel Independen	VIF
Indeks Pembangunan Manusia (Z1)	3.05432
Angka Harapan Hidup (Z2)	2.58516
Jumlah Bayi Lahir (Z3)	2.58207
Tingkat Pengangguran Terbuka (Z4)	1.21527

Adapun hipotesis uji multikolinearitas adalah sebagai berikut:

H_0 : Terdapat multikolinearitas antara variabel independen

H_1 : Tidak terdapat multikolinearitas antara variabel independen

Taraf signifikan : $\alpha = 5\%$

Statistik uji : $VIF = \frac{1}{1-R_j^2}$

Daerah kritis : Tolak H_0 jika $VIF < 10$

Keputusan : Tolak H_0 karena nilai VIF dari masing-masing variabel independen kurang dari 10

Kesimpulan, tidak terdapat multikolinearitas antara variabel pada level 2. Berdasarkan hasil tersebut maka dapat disimpulkan bahwa nilai VIF kurang dari 10 sehingga tidak terdapat multikolinieritas antara variabel pada level 2. Selanjutnya melakukan analisis regresi multilevel yang digunakan pada penelitian ini yaitu model regresi multilevel pada level 1 yaitu tanpa mengikutsertakan variabel Z. Setelah itu, pemodelan level 2 akan dilakukan dengan mengikutsertakan variabel Z. Pada level 1 pemodelannya sebagai berikut:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{1j} + \dots + \beta_{ij}X_{ij} + \varepsilon_{ij}$$

Metode pendugaan parameter menggunakan *maximum likelihood* hasil perhitungan disajikan pada Tabel 6 sebagai berikut:

Tabel 6. Hasil Dugaan Parameter Analisis Regresi Multilevel Tanpa Variabel Z

Parameter	Dugaan	Galat Baku	t value	P value
Intersep	2026682	387957	5.22	0.0002
Luas Daerah (X1)	-0.6237	0.06560	-9.51	<.0001
Rasio Jenis Kelamin (X2)	-8.1913	1.9321	-4.24	<.0001

Berdasarkan Tabel 6 dapat diketahui bahwa hasil dugaan parameter untuk masing-masing variabel independen dengan adanya pengaruh level 1 yaitu kecamatan, variabel luas daerah (X1), dan rasio jenis kelamin (X2) berpengaruh terhadap kepadatan penduduk. Sehingga, model regresi multilevel tanpa variabel Z yaitu sebagai berikut:

$$Y_{ij} = 2026682 - 0,6237X_1 - 8,1913X_2 + \varepsilon_{ij}$$

Selanjutnya Pemodelan analisis regresi multilevel dengan mengikutsertakan variabel independen pada level kabupaten adalah sebagai berikut:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_{1j} + \gamma_{02}Z_{2j} + \dots + \gamma_{0i}Z_{ij} + u_{0j}$$

Hasil dugaan parameter dengan variabel Z dapat dilihat pada Tabel 7 sebagai berikut:

Tabel 7. Hasil dugaan parameter dengan variabel Z

Parameter	Dugaan	Galat Baku	t value	P value
Intersep	141220	913118	0.15	0.8805
Luas Daerah (X1)	-0.6513	0.06432	-10.13	<.0001
Rasio Jenis Kelamin (X2)	-6.2751	1.9570	-3.21	0.0016
Indeks Pembangunan Manusia (Z1)	0.8724	4.3682	0.20	0.8419
Angka Harapan Hidup (Z2)	6.7560	4.3316	1.56	0.1207
Jumlah Bayi Lahir (Z3)	0.2320	0.1960	1.18	0.2382
Tingkat Pengangguran Terbuka (Z4)	17183	26182	0.66	0.5125

Berdasarkan Tabel 7 dapat dilakukan pengujian dugaan parameter secara parsial untuk variabel independen level 2 indeks pembangunan manusia (Z1), angka harapan hidup (Z2), jumlah bayi lahir (Z3), dan tingkat pengangguran terbuka (Z4) yang tidak signifikan sebagai berikut:

H_0 : $\gamma_{ij} = 0$ (variabel independen level 2 tidak berpengaruh terhadap variabel kepadatan penduduk)

H_1 : $\gamma_{ij} \neq 0$ (variabel independen level 2 berpengaruh terhadap variabel kepadatan penduduk)

Tingkat signifikansi : $\alpha = 5\%$

Statistik uji : $t = \frac{\hat{\gamma}_{ij}}{SE(\gamma_{ij})}$

Daerah kritis : Tolak H_0 jika t hitung $> t$ Tabel dan p value $< 0,05$

Keputusan : Tidak tolak H_0 karena nilai t hitung $Z1 < t$ tabel ($0,20 < 1,973$), nilai t hitung $Z2 < t$ tabel ($1,56 < 1,973$), nilai t hitung $Z3 < t$ tabel ($1,18 < 1,973$), dan nilai t hitung $Z4 < t$ tabel ($0,66 < 1,973$) dan p value $> 0,05$

Kesimpulan, dapat diketahui bahwa variabel indeks pembangunan manusia (Z1) dan angka harapan hidup (Z2), jumlah bayi lahir (Z3), dan tingkat pengangguran terbuka (Z4) tidak berpengaruh terhadap variabel kepadatan penduduk. Setelah dilakukan pengujian dugaan parameter dengan metode *maximum likelihood* untuk masing-masing variabel independen pada level 1 dan 2 maka variabel luas daerah (X1) dan rasio jenis kelamin (X2) berpengaruh terhadap kepadatan penduduk. Sehingga model regresi multilevel untuk kepadatan penduduk adalah sebagai berikut:

$$\gamma_{ij} = 141220 - 0,6513 X1 - 6,2751 X2 + \varepsilon_{ij} + u_{0j}$$

Hasil analisis yang telah diperoleh, dapat diketahui nilai deviasi dari masing-masing model yang dihasilkan. Pemilihan model terbaik dilakukan dengan cara membandingkan nilai deviasi setiap model dengan hasil sebagai berikut:

Tabel 8. Nilai Deviasi Setiap Model

No	Model	Deviasi	Diff	$\chi^2_{(a,db)}$
1.	<i>Null model</i>	4449.3		
2.	Model multilevel tanpa variabel Z	4344.0	8793.3	3.841
3.	Model multilevel dengan variabel Z	4333.3	8781.6	11.07

H_0 : Model tidak sesuai

H_1 : Model sesuai

Taraf signifikansi : $\alpha = 5\%$

Statistik uji :

Diff : $-2 \log e \left(\frac{\lambda_0}{\lambda_1} \right)$

Daerah kritis : Tolak H_0 jika nilai deviasi paling kecil dan diff $> \chi^2_{(a,db)}$

Keputusan : tolak H_0 karena nilai deviasi model multilevel dengan variabel Z paling kecil diantara nilai deviasi model lain yaitu 4333,3 dan nilai diff $> \chi^2_{(a,db)}$ ($8781,6 > 11,07$).

Kesimpulan : dapat disimpulkan bahwa model multilevel dengan variabel Z lebih sesuai dibandingkan model multilevel tanpa variabel Z

Hasil model regresi multilevel adalah sebagai berikut:

(1) Model level 1

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{1j} + \beta_{2j}X_{2j} + \varepsilon_{ij}$$

(2) Model level 2

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

Sehingga diperoleh model regresi multilevel sebagai berikut:

Nilai koefisien regresi pada variabel-variabel independennya menggambarkan apabila diperkirakan variabel independennya naik sebesar satu unit dan nilai variabel independen lainnya diperkirakan konstan atau sama dengan nol, maka nilai variabel dependen diperkirakan bias naik atau bias turun dengan tanda koefisien variabel independennya. Konstanta regresi signifikan terhadap model, hal ini menunjukkan bahwa apabila variabel independennya konstan atau sama dengan nol, maka nilai variabel dependen diperkirakan adalah 141220. Artinya jika variabel luas daerah dan rasio jenis kelamin bernilai nol maka kepadatan penduduk bernilai 141220.

Koefisien regresi untuk variabel independen level 1 yaitu luas daerah bernilai negatif, hal ini menunjukkan adanya hubungan yang tidak searah antara kepadatan penduduk dengan luas daerah sebesar -0,6513 yang mengandung arti untuk setiap pertambahan luas daerah sebesar satu satuan akan menyebabkan menurunnya kepadatan penduduk sebesar 0,6513. Koefisien regresi untuk variabel independen rasio jenis kelamin bernilai negatif, hal ini menunjukkan adanya hubungan yang tidak searah antara kepadatan penduduk dengan rasio jenis kelamin sebesar -6,2751 mengandung arti untuk setiap pertambahan rasio jenis kelamin sebesar satu satuan akan menyebabkan menurunnya kepadatan penduduk sebesar 6,2751.

Selanjutnya mencari nilai koefisien intraklas. Nilai koefisien korelasi intraklas diperoleh dari penduga ragam pada setiap level untuk model multilevel dengan melibatkan peubah penjelas. Nilai korelasi intraklas berfungsi sebagai variabel kehomogenan antar unit disetiap level yang sama. Semakin tinggi nilai korelasi intraklas maka semakin homogen unit-unit tersebut.

Tabel 9. Nilai Dugaan Parameter Acak *Full Model*

Parameter	Dugaan
$\sigma^2_{e_{ij}}$ (Level 1)	4.8585
$\sigma^2_{u_{0j}}$ (Level 2)	4.8549

Berdasarkan nilai dugaan parameter tersebut dapat diperoleh koefisien korelasi intraklas menggunakan persamaan 9 sebagai berikut:

$$\rho = \frac{4.8585}{4.8585 + 4.8549} = 0,50$$

Berdasarkan perhitungan, diperoleh nilai korelasi antara dua kecamatan dalam satu kabupaten sebesar 0,50. Nilai korelasi ini melebihi batas ketentuan yaitu 0,05 sehingga dapat dikatakan bahwa data kepadatan penduduk tahun 2021 memiliki struktur hirarki. Selain itu, nilai korelasi intraklas juga menjelaskan proporsi ragam pada level kabupaten sebesar 50%. Selanjutnya mencari nilai koefisien determinasi. Koefisien determinasi harus diperoleh menggunakan penduga ragam galat level 1 dan level 2 tanpa variabel independen atau yang disebut juga *null model*. Hasil dugaan parameter acak untuk model sebagai berikut:

Tabel 10. Nilai Dugaan Parameter Acak *Full Model* dan *Null Model*

Parameter		Nilai Duga
Dengan Peubah Prediktor	$\sigma^2_{e_0}$	4.8585
	$\sigma^2_{u_0}$	4.8549
Tanpa Peubah Prediktor	$\sigma^2_{e_{ij}}$	8.2416
	$\sigma^2_{u_{0j}}$	6.8848

Berdasarkan tabel tersebut maka diperoleh koefisien determinasi level 1 sebagai berikut:

$$R_1^2 = 1 - \frac{4,8585}{8,2416} = 0,41$$

Koefisien determinasi pada level 1 menjelaskan kontribusi luas daerah dan rasio jenis kelamin di setiap kecamatan terhadap model sebesar 41%. Perhitungan koefisien determinasi level 2 menggunakan persamaan 2.21 berikut:

$$R_2^2 = 1 - \frac{4,8549}{6,8848} = 0,29$$

Koefisien determinasi pada level 2 menjelaskan kontribusi yang diberikan sebesar 29%. Total keragaman model kepadatan penduduk Provinsi Lampung tahun 2021 sebesar 70% yang

merupakan kontribusi dari faktor-faktor yang diteliti (luas daerah dan rasio jenis kelamin) sedangkan 30% merupakan kontribusi dari faktor lain.

4. Kesimpulan

Faktor-faktor yang mempengaruhi kepadatan penduduk di Provinsi Lampung tahun 2021 adalah luas daerah dan rasio jenis kelamin. Berdasarkan hasil analisis regresi multilevel diperoleh model yaitu $\text{Kepadatan penduduk} = 141220 - 0,6513 (\text{luas daerah}) - 6,2751 (\text{rasio jenis kelamin}) X2 + \varepsilon_{ij} + u_{0j}$. Artinya, jika variabel independen bernilai nol maka kepadatan penduduk bernilai 141220. Untuk setiap pertambahan satu satuan pada variabel luas daerah akan menyebabkan menurunnya kepadatan penduduk sebesar 0,6513. Variabel rasio jenis kelamin akan menyebabkan menurunnya kepadatan penduduk sebesar 6.2751. Keragaman yang dapat dijelaskan oleh level 1 adalah sebesar 41%, artinya besarnya pengaruh variabel independen pada level kecamatan yaitu luas daerah dan rasio jenis kelamin terhadap kepadatan penduduk adalah sebesar 41%. Sedangkan keragaman pada level 2 memberikan pengaruh sebesar 29%. Total keragaman model kepadatan penduduk Provinsi Lampung tahun 2021 sebesar 70% yang merupakan kontribusi dari faktor-faktor yang diteliti (luas daerah dan rasio jenis kelamin) sedangkan 30% merupakan kontribusi dari faktor lain.

Daftar Pustaka:

- [1] J. Harlan, *Analisis Multilevel*. Depok: Gunadarma, 2016.
- [2] J.J. Hox, *Multilevel Analysis: Techniques and Applications*, 2nd ed. New York: Routledge, 2010.
- [3] B. Tantular, Annuddin, and H. Wijayanto, "Pemilihan Model Regresi Linier Multilevel Terbaik," *Forum Stat. dan Komputasi*, vol. 14, no. 2, pp. 1–7, 2009.
- [4] J.J.Hox, *Applied Multilevel Analysis*. Amsterdam: TT-Publikaties, 1995.
- [5] R.K. Sembiring, *Analisis Regresi*. Bandung: ITB, 1995.
- [6] D. Kurniawaaan, *Regresi Linier*. Australia: R Development Core Team, 2008.
- [7] S. Kristiana, L. S. Nasution, Naibaho, and M.M.P, "Faktor Kontekstual dan Individual Terhadap Jumlah Anak Lahir Hidup: Sebuah Analisis Multilevel," *J. Kependud. Indones.*, vol. 15, no. 1, pp. 33–38, 2020.