

Model Regresi Linear Multilevel untuk Nilai UN SMP Tahun 2010 di Propinsi Sumatera Barat

Gurfany Falensia¹, Dodi Vionanda², Riry Sriningsih³

¹ Student of Mathematics Department State University of Padang, Indonesia

^{2,3} Lecturers of Mathematics Department State University of Padang, Indonesia

¹gurfanyfalensia@gmail.com

Abstract --- In many social research, data often have hierarchical structure or nested structure. Generally, regression classical model is not suitable to analyze hierarchical data, because the observations is not fully independent. The suitable model is regression multilevel model. National Exam Data on 2010 for Junior high School in West Sumatera is one of hierarchical structure. Hierarchical structure obtain from two levels, first level is school, second level is region in West Sumatera. This research is to see the influence status of school and school quantity in every region in West Sumatera on National Exam for Junior High School. Using regression multilevel model we can obtain the best model $Y_{ij} = 26,794034 + 1,273654X_{ij} + 0,036334Z_j$, which show that average of National Exam which considered by status dan quantity of school in the area is 26,794034 and expected value of state school is 26,794034 higher compared with private school and every increased of one school's quantity in the area, grade of National Exam will increase 0,036334.

Keywords: Hierarchical Data, Multilevel Models, Maximum Likelihood, Deviance.

Abstrak --- Pada beberapa penelitian social sering ditemukan data yang mempunyai struktur hirarki atau struktur tersarang. Pada umumnya, model regresi klasik tidak cocok untuk analisis data hirarki, karena observasi tidak saling bebas. Model yang pantas adalah model regresi multilevel. Data ujian nasional (UN) untuk SMP di Sumatera Barat adalah salah satu dari struktur hirarki. Struktur hirarki diperoleh dari dua tingkatan, tingkatan pertama adalah sekolah, tingkatan kedua adalah wilayah di Sumatera Barat. Penelitian ini untuk melihat pengaruh status dari sekolah jumlah sekolah di setiap wilayah di Sumatera Barat pada Ujian Nasional SMP. Dengan menggunakan model regresi multilevel kita dapat memperoleh model terbaik yaitu: $Y_{ij} = 26,794034 + 1,273654X_{ij} + 0,036334Z_j$, dimana menunjukkan bahwa rata-rata dari Ujian Nasional dengan mempertimbangkan status dan jumlah dari sekolah dalam area 26,794034 dan nilai harapan dari status sekolah adalah 26,794034 tertinggi dan setiap peningkatan dari satu jumlah sekolah dalam wilayah akan meningkatkan nilai Ujian Nasional sebesar 0,036334.

Kata Kunci – Data hirarki, model multilevel, maksimum likelihood,

PENDAHULUAN

Aktifitas penelitian tidak akan terlepas dari data yang merupakan bahan informasi untuk memberikan gambaran spesifik mengenai objek penelitian. Di berbagai disiplin ilmu, sering dijumpai data populasi yang berstruktur hirarki (Goldstein, 2011: 1). Data berstruktur hirarki merupakan data yang timbul karena unit-unit terkumpul dalam kelompok-kelompok sosialnya. Keberadaan struktur hirarki dalam data tidak bisa diabaikan. Hal ini disebabkan pada pemodelan data hirarki parameter pada unit di level bawah dikendalikan oleh parameter yang berada di level kelompok (Andrew Gelman & Jennifer Hill, 2011: 1).

Secara umum, model regresi linear klasik kurang tepat digunakan untuk menganalisis data hirarki, karena observasi-observasi pada data hirarki tidak sepenuhnya independen. Oleh karena itu, untuk data hirarki model regresi yang tepat digunakan adalah model regresi multilevel. Model regresi multilevel merupakan perluasan dari regresi yang datanya disusun dalam kelompok dan koefisiennya dapat diubah oleh kelompok.

Salah satu data tentang sekolah yang memiliki struktur hirarki adalah data nilai UN SMP yang diperoleh dari Badan Standar Nasional Pendidikan (BSNP). Informasi yang dapat diperoleh dari data diantaranya rata-rata total UN masing-masing SMP seluruh propinsi di Indonesia. Di samping itu, diperoleh informasi tentang rata-rata nilai

UN SMP per bidang studi, status sekolah, jumlah peserta UN, persentase kelulusan masing-masing SMP, dan jumlah sekolah masing-masing wilayah.

Dalam penelitian ini, sekolah-sekolah yang berada di Sumatera Barat dijadikan sebagai objek penelitian, dimana SMP sebagai unit terkecil yang berada pada level-1 dan wilayah Kota/Kabupaten sebagai unit level-2 yang membawahi unit level-1. Status sekolah yaitu negeri dan swasta merupakan informasi pada level-1 yang berperan sebagai peubah bebas level-1 (X) dan jumlah SMP di wilayah sebagai informasi level-2 berperan sebagai peubah bebas level-2 (Z). Sementara peubah respons (Y) berada pada level-1, yaitu rata-rata total UN masing-masing SMP.

Oleh karena itu, akan dilihat bagaimana nilai UN yang dicapai SMP tahun 2010 dengan mempertimbangkan status sekolah dan jumlah SMP di wilayah menggunakan model regresi multilevel.

Secara umum, Hox (2002: 11) menjelaskan bahwa model regresi multilevel mempunyai struktur hirarki dimana terdapat sebuah peubah respons (*dependent variable*) yang diukur pada level paling bawah atau level-1 dan satu atau lebih peubah penjelas (*explanatory variabel*) yang diukur pada setiap level. Bentuk umum regresi linier 2-level adalah:

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

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j} \quad (2)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + u_{1j} \quad (3)$$

Dengan menstutbutisikan persamaan level-2 ke persamaan level-1 diperoleh :

$$Y_{ij} = (\gamma_{00} + \gamma_{01}Z_j + \gamma_{10}X_{ij} + \gamma_{11}X_{ij}Z_j)(u_{0j} + u_{1j}X_{ij} + \epsilon_{ij}) \quad (4)$$

Model di atas dikenal juga dengan Model Koefisien Acak.

Pada regresi multilevel terdapat beberapa model yang akan dibandingkan diantaranya Model Null, Model Intersep Acak dan Model Koefisien Acak. Model Null digunakan sebagai langkah awal dalam analisis multilevel yang dapat memberikan keberagaman respons pada masing-masing level dalam data hirarki. Model Null dipresentasikan sebagai berikut:

$$\text{Model level-1 } Y_{ij} = \beta_{0j} + \epsilon_{ij} \quad (5)$$

$$\text{Model level-2 } \beta_{0j} = \gamma_{00} + u_{0j} \quad (6)$$

Dengan mensubstitusikan persamaan (6) ke dalam persamaan (5) diperoleh model tunggalnya:

$$Y_{ij} = \gamma_{00} + u_{0j} + \epsilon_{ij} \quad (7)$$

(Hox,2002: 15)

Untuk melihat hubungan antara 2 unit level-1 dalam unit level-2 yang sama, terdapat metode yang dikenal dengan *Intraclass Correlation* (ICC). Dalam model ini dapat diperoleh besarnya ICC, yaitu:

$$\rho = \frac{\sigma_{u_0}^2}{(\sigma_{\epsilon}^2 + \sigma_{u_0}^2)} \quad (8)$$

Model Intersep Acak dipresentasikan sebagai berikut:

Model level-1:

$$Y_{ij} = \beta_{0j} + \sum_{p=1}^P \beta_{pj} x_{pij} + \epsilon_{ij} \quad (9)$$

Model level-2:

$$\beta_{0j} = \gamma_{00} + \sum_{q=1}^Q \gamma_{0q} Z_{qj} + u_{0j} \quad (10)$$

$$\beta_{pj} = \gamma_{p0} \quad (11)$$

Persamaan (10) dan (11) disubstitusikan ke dalam model (9) sehingga diperoleh model persamaan tunggal:

$$Y_{ij} = \gamma_{00} + \sum_{p=1}^P \gamma_{p0} x_{pij} + \sum_{q=1}^Q \gamma_{0q} Z_{qj} + u_{0j} + \epsilon_{ij} \quad (12)$$

Dalam pendugaan parameter model multilevel, berbagai prosedur pendugaan telah diusulkan oleh beberapa ahli statistika. Salah satu metode yang digunakan adalah Metode Kemungkinan Maksimum (*Maximum Likelihood*). Secara umum, fungsi kemungkinan maksimum untuk model linear campuran adalah sebagai berikut:

$$L(\beta, \theta) = \prod_{j=1}^J 2\pi^{-\frac{n_j}{2}} |V_j|^{-\frac{1}{2}} \exp -0,5 \sum (y_j - X_j \beta)' V_j^{-1} (y_j - X_j \beta) \quad (13)$$

dengan θ adalah vektor parameter acak dari elemen matriks V_j . Fungsi log-likelihood dari persamaan (10) adalah:

$$l(\beta, \theta) = \ln(\beta, \theta) = 0,5n. \ln 2\pi - 0,5 \ln |V_j| - 0,5 \sum_{j=1}^J (y_j - X_j \beta_j)' V_j^{-1} (y_j - X_j \beta_j) \quad (14)$$

dalam fungsi ini, V_j merupakan fungsi dari θ . Diperoleh penduga koefisien regresi sebagai berikut:

$$\hat{\beta} = (\sum_{j=1}^J (X_j' V_j^{-1} X_j))^{-1} \sum_{j=1}^J X_j' V_j^{-1} y_j \quad (15)$$

dan galat baku adalah akar diagonal utama matriks:

$$\text{var}(\hat{\beta}) = (\sum_{j=1}^J (X_j' V_j^{-1} X_j))^{-1} \quad (16)$$

Untuk mendapatkan nilai V_j dibentuk fungsi *profile log-likelihood* ($l_{ML}(\beta, \theta)$). Dengan mensubstitusikan nilai β dengan penduga persamaan (13) untuk nilai V_j yang telah ditentukan, diperoleh:

$$l(\beta, \theta) = \ln(\beta, \theta) = -0,5n. \ln 2\pi - 0,5 \ln |V_j| - 0,5 \sum_{j=1}^J (y_j - X_j \hat{\beta}_j)' V_j^{-1} (y_j - X_j \hat{\beta}_j) \quad (17)$$

sehingga diperoleh penduga V_j sebagai berikut:

$$\hat{V}_j = Z_j \hat{D} Z_j' + R_j' \quad (18)$$

dengan mensubstitusikan persamaan (18) ke persamaan (16) diperoleh:

$$\hat{\beta} = (\sum_{j=1}^J (X_j' (Z_j \hat{D} Z_j')^{-1} X_j))^{-1} \sum_{j=1}^J X_j' (Z_j \hat{D} Z_j' + R_j')^{-1} y_j \quad (19)$$

dengan galat baku diperoleh dari akar diagonal utama matriks:

$$\text{var}(\hat{\beta}) = (\sum_{j=1}^J (X_j' (Z_j \hat{D} Z_j' + R_j')^{-1} X_j))^{-1} \quad (20)$$

Prosedur pendugaan ini harus melalui proses iterasi sampai mendapatkan penduga yang konvergen.

Agar diperoleh penduga yang tak bias, dilakukan suatu modifikasi dari prosedur pendugaannya. Fungsi yang dimodifikasi ini dinamakan *Restricted log-likelihood* sebagai berikut:

$$l_{REML} = -0,5(n-p)(\ln 2\pi - 0,5 \sum_{j=1}^J \ln |V_j|) \\ [-0,5 \sum_{j=1}^J (y_j - X_j \hat{\beta}_j)' V_j^{-1} (y_j - X_j \hat{\beta}_j) - \\ 0,5 \sum_{j=1}^J \ln |X_j' V_j^{-1} X_j|] \quad (21)$$

dengan p adalah banyak parameter tetap dalam model. Dengan menggunakan persamaan (21) dapat diperoleh nilai dugaan V_j yang selanjutnya digunakan untuk menduga koefisien regresi dan galat bakunya menggunakan persamaan (19) dan (20). Prosedur ini disebut *Restricted Maximum Likelihood* atau *Residual Maximum Likelihood* (REML).

Untuk mengukur baik atau tidaknya suatu penduga, dapat dihitung galat baku (Se). Menurut Goldstein (2011: 40), penduga untuk parameter tetap mempunyai ragam penarikan sampel:

$$\text{cov}(\hat{\beta}) = (X' V^{-1} X)^{-1} \quad (22)$$

dengan galat baku penduga parameter β adalah akar positif diagonal utama matriks persamaan (19). Sedangkan ragam penarikan sampel untuk penduga parameter acak θ adalah:

$$\text{cov}(\hat{\theta}) = (Z^* V^{*-1} Z^*)^{-1} \quad (23)$$

Galat baku koefisien acak diperoleh dari akar positif diagonal utama matriks tersebut.

Pengujian hipotesis parameter dalam model multilevel dilakukan secara individual. Adapun uji yang digunakan adalah uji Wald. Perumusan hipotesis disusun pada setiap level, dimana:

Level-1: $H_0: \beta_{kj} = 0$ dan $H_1: \beta_{kj} \neq 0$ dengan $k=1,2,3,\dots,q$ menyatakan banyak parameter tetap level-1

Level-2: $H_0: \gamma_{lj} = 0$ dan $H_1: \gamma_{lj} \neq 0$ dengan $l=1,2,3,\dots,r$ menyatakan banyak parameter tetap level-2.

$H_0: \beta_{kj} = 0$ maupun $H_0: \gamma_{lj} = 0$ berarti parameter tidak signifikan, dan sebaliknya.

Kecocokan model dalam model multilevel dapat diukur dengan menggunakan deviansi. Secara umum deviansi dapat didefinisikan:

$$D = -2 \log \left(\frac{\lambda_0}{\lambda_1} \right) \quad (24)$$

(Goldstein, 2011: 41)

dengan λ_0 adalah fungsi kemungkinan di bawah hipotesis nol pada saat mencapai konvergen λ_1 adalah fungsi kemungkinan di bawah hipotesis alternatif pada saat mencapai konvergen.

Hox (2002: 16) menyatakan bahwa semakin kecil nilai deviansi, model tersebut semakin baik. Akan tetapi tidak ada ketentuan yang pasti berapa besar ukuran untuk nilai deviansi ini. Sehingga untuk mengetahui model cocok

atau tidaknya harus dibandingkan dengan model lain. Prosedur membandingkan dua model tersebut dengan menggunakan suatu pengukuran perbedaan deviansi yaitu:

$$\text{diff} = D_1 - D_2 \quad (25)$$

diff mengikuti sebaran chi-kuadrat dengan derajat bebas $k = p_2 - p_1$ dimana p_1 adalah banyak parameter pada model M_1 dan p_2 adalah banyak parameter pada model M_2 (Goldstein, 2007: 41). Jika nilai diff lebih besar dari nilai $\chi_{\alpha}^2(k)$, maka model yang memiliki parameter yang lebih banyak yang lebih cocok untuk data tersebut.

Dalam analisis regresi, keragaman dapat dijelaskan melalui koefisien determinasi. Koefisien determinasi pertama didefinisikan pada level-1 dengan tujuan untuk menilai rasio ragam galat terhadap ragam total dengan rumus:

$$R_1^2 = 1 - \frac{\hat{\sigma}_{ep}^2}{\hat{\sigma}_{e0}^2} \quad (26)$$

dimana:

$\hat{\sigma}_{ep}^2$: penduga ragam residu pada level-1 dengan p peubah penjas

$\hat{\sigma}_{e0}^2$: penduga ragam residu pada level-1 tanpa peubah penjas.

$R_1^2 = 1$ artinya peubah penjas level-1 secara sempurna mampu menjelaskan respons sedangkan $R_1^2 = 0$ artinya peubah penjas tidak dapat menjelaskan apapun terhadap respons.

Koefisien determinasi juga dapat dihitung pada setiap level-2 dalam model.

Dirumuskan:

$$R_2^2 = 1 - \frac{\hat{\sigma}_{u0p}^2}{\hat{\sigma}_{u0}^2} \quad (27)$$

dimana :

$\hat{\sigma}_{u0p}^2$: penduga ragam residu pada level-2 dengan p peubah penjas

$\hat{\sigma}_{u0}^2$: penduga ragam residu pada level-2 tanpa peubah penjas

METODE

Penelitian ini merupakan penelitian terapan. Pada penelitian ini, data yang diperoleh dideskripsikan setelah dilakukan penerapan dari analisis regresi multilevel pada data tersebut. Data yang digunakan dalam penelitian ini adalah data sekunder yang diambil dari Badan Standar Nasional Pendidikan (BSNP). Data yang diperoleh berupa data tahunan yaitu data pada tahun 2010. Data yang digunakan adalah data rata-rata total UN SMP baik negeri maupun swasta. Terdapat 19 Kota/Kabupaten yang terdiri dari 524 sekolah negeri dan 87 sekolah swasta.

Dalam penelitian ini, peubah respon Y merupakan rata-rata total UN SMP pada masing-masing SMP di Propinsi Sumatera Barat yang berupa data kontinu. Peubah bebas level-1 adalah status sekolah (X) yang berupa data nominal, yaitu negeri (1) dan swasta (0).

Sedangkan peubah bebas level-2 adalah jumlah SMP masing-masing wilayah Kota/Kabupaten yang berupa data diskrit.

Analisis data dalam penelitian ini dilakukan dengan bantuan software SPSS, dengan tahapan analisis sebagai berikut:

1. Eksplorasi perilaku hirarki yang terdapat dalam data
2. Pembentukan model, dimana terdapat 3 jenis model:
 - a. Model null yang tidak melibatkan peubah bebas pada level-1 maupun level-2 berdasarkan persamaan (7). Dengan menggunakan informasi pada model ini, dapat dihitung nilai koefisien korelasi intraclass menggunakan persamaan (8). Model ini juga digunakan untuk melihat adanya struktur hirarki dalam data.
 - b. Model intersep acak, yaitu model dengan nilai intersep merupakan koefisien acak, tetapi koefisien kemiringan tetap berdasarkan persamaan (12)
 - c. Model koefisien acak yang melibatkan peubah bebas level-1 dan peubah bebas level-2 dan memperhatikan adanya pengaruh peubah pada level-2 berdasarkan persamaan (4)
3. Pendugaan parameter tetap dan parameter acak menggunakan software SPSS 17.0
4. Model-model tersebut dibandingkan antara yang satu dengan yang lain dengan menggunakan deviansi untuk memperoleh model terbaik berdasarkan persamaan (24) dan (25)
5. Dari model terpilih dihitung keragaman yang dapat dijelaskan pada masing-masing level menggunakan persamaan (26) dan (27)
6. Bila model yang terpilih mengandung koefisien kemiringan acak, langkah selanjutnya adalah menghitung keragaman yang ada pada koefisien kemiringan tersebut
7. Interpretasi dan kesimpulan.

HASIL DAN PEMBAHASAN

1. Eksplorasi Perilaku Hirarki Data

Struktur hirarki dalam data dapat dilihat pada plot data masing-masing wilayah terhadap rata-rata total UN sekolah yang disajikan dalam Gambar.1.

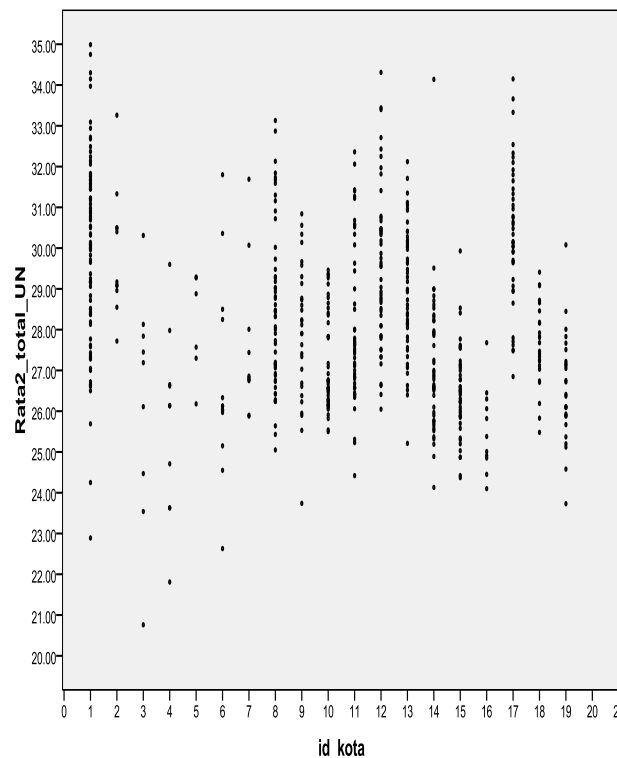
Berdasarkan plot data pada gambar 1, dapat diperhatikan bahwa terdapat struktur hirarki dalam data yang menggambarkan pengaruh wilayah terhadap unit-unit yang berada di bawahnya.

2. Pembentukan Model Regresi Multilevel dan Pendugaan Parameter

Pada model null dapat dihitung nilai *Intraclass Correlation* (ICC) untuk melihat adanya struktur hirarki pada data. Berdasarkan persamaan (8) dapat dihitung:

$$\rho = \frac{\sigma_{u0}^2}{(\sigma_{\epsilon}^2 + \sigma_{u0}^2)} = \frac{1,930913}{(1,930913 + 3,234387)} = 0,3738$$

Nilai ini mengandung arti bahwa proporsi ragam pada level wilayah sebesar 37,38%. Selain itu dapat dijelaskan bahwa korelasi antara dua sekolah dalam satu wilayah sebesar 0,3738.



Gambar.1 Plot data nilai UN SMP yang berada di masing-masing wilayah

3. Pemilihan Model Terbaik

Berdasarkan tabel II, nilai *diff* terkecil diperoleh pada Model Koefisien Acak, namun pada model ini terdapat beberapa parameter yang tidak signifikan. Oleh karena itu, dipilih Model Intersep Acak sebagai model terbaik. Bentuk model:

$$Y_{ij} = \beta_{0j} + 1,273654X_{ij} + \epsilon_{ij} \text{ (Model Level-1)}$$

$$\beta_{0j} = 26,794034 + 0,036334Z_j + u_{0j}$$

$$\beta_{1j} = 1,273654 \text{ (Model Level-2)}$$

atau dalam bentuk model campuran:

$$Y_{ij} = 26,794034 + 1,273654X_{ij} + 0,036334Z_j + u_{0j} + \epsilon_{ij}$$

Model ini menunjukkan bahwa rata-rata total UN sekolah dengan mempertimbangkan peranan status sekolah dan banyaknya SMP di wilayah sebesar $\gamma_{00} = 26,794034$. Nilai ini lebih rendah dibandingkan Model Null dengan $\gamma_{00} = 27,822081$ yang tidak memperhatikan adanya pengaruh status sekolah dan banyaknya SMP di wilayah. $\gamma_{10} = 1,273654$ menunjukkan sekolah yang berstatus negeri nilai UN yang diharapkan lebih tinggi 1,273654 dibandingkan sekolah berstatus swasta. $\gamma_{01} = 0,036334$ menunjukkan setiap kenaikan satu unit sekolah di suatu wilayah akan meningkatkan rata-rata total UN sebesar 0,036334.

TABEL I
HASIL PENDUGAAN PARAMETER TETAP DAN PARAMETER ACAK MODEL
NULL, MODEL INTERSEP ACAK DAN MODEL KOEFISIEN ACAK

Model Null (<i>Intersept Only Model</i>)				
	Penduga	Standar Error (Se)	t	P_value
<i>Fixed Parameter</i>				
γ_{00}	27,822081	0,332318	83,72	0,000
<i>Random Parameter</i>				
σ_{u0}^2	1,930913	0,699030	2,762	0,006
σ_{ϵ}^2	3,234387	0,187990	17,20	0,000
Deviansi	2503,362			
Model Intersep Acak (<i>Intersept Only Model</i>)				
<i>Fixed parameter</i>				
γ_{00}	26,79403	0,569558	47,004	0,000
γ_{10}	1,273654	0,234061	5,442	0,000
γ_{01}	0,036334	0,014481	2,509	0,023
<i>Random parameter</i>				
σ_{u0}^2	1,519556	0,578178	2,628	0,009
σ_{ϵ}^2	3,081207	0,179262	17,188	0,000
Deviansi	2476,325			
Model Koefisien Acak (<i>Random Coefficient Model</i>)				
<i>Fixed parameter</i>				
γ_{00}	26,58032	0,596791	44,539	0,000
γ_{10}	0,386607	0,823813	0,469	0,647
γ_{01}	0,042099	0,015175	2,774	0,013
γ_{11}	0,009054	0,019517	0,464	0,653
<i>Random parameter</i>				
σ_{u0}^2	2,441976	1,327593	1,839	0,066
σ_{u1}^2	1,644162	0,628637	2,615	0,009
σ_{u01}	1,133509	0,688391	1,647	0,100
σ_{ϵ}^2	2,907953	0,171493	16,597	0,000
Deviansi	2461,757			

4. Keragaman Masing-masing Level

Untuk memperoleh keragaman yang dapat dijelaskan pada setiap level pada Model Intersep Acak, digunakan Model Null sebagai dasar. Dengan menggunakan persamaan (26) dan persamaan (27) serta informasi pada Tabel 1 dan Tabel 2 dapat diperoleh:

$$R_1^2 = 1 - \frac{3,081207}{3,234387} = 0,0474$$

$$R_2^2 = 1 - \frac{1,519556}{1,930913} = 0.2130$$

Hasil tersebut dapat diinterpretasikan bahwa keragaman yang dijelaskan oleh model pada Level 1 sebesar 4,74% dan keragaman yang dapat dijelaskan oleh model pada Level 2 sebesar 21,30%. Hal ini mengandung arti bahwa keragaman status sekolah memberikan pengaruh terhadap rata-rata total UN sekolah sebesar 4,74% dan sebesar 21,30% pengaruh diberikan oleh keragaman banyak SMP di wilayah. Hal ini menunjukkan bahwa meskipun keragaman lebih besar disebabkan oleh

struktur wilayah, keragaman antar sekolah juga mempunyai pengaruh terhadap keragaman data.

TABEL II
NILAI DEVIASI MASING-MASING MODEL

Model	Deviansi	Parameter	diff	db	Nilai p
Model Null	2503,362	3			
Model Intersep Acak	2476,325	5	27,03	2	5,991
Model Koefisien Acak	2461,757	8	14.56	3	7,815

SIMPULAN

Model regresi multilevel nilai UN SMP tahun 2010 di Propinsi Sumatera Barat adalah:

$$Y_{ij} = \beta_{0j} + 1,273654X_{ij} + \epsilon_{ij} \quad (\text{Model Level-1})$$

$$\beta_{0j} = 26,794034 + 0,036334Z_j + u_{0j} \quad (\text{Model Level-2})$$

$$\beta_{1j} = 1,273654$$

Model campuran:

$$Y_{ij} = 26,794034 + 1,273654X_{ij} + 0,036334Z_j$$

Model ini menjelaskan rata-rata total UN SMP dengan mempertimbangkan peranan status sekolah dan jumlah sekolah di wilayah sebesar 26,794034. Di samping itu, sekolah berstatus negeri nilai UN yang diharapkan 1,2736 lebih tinggi dibanding sekolah berstatus swasta, dan setiap peningkatan satu unit sekolah akan meningkatkan rata-rata total UN sebesar 0,036334.

Total keragaman sebesar 26,04% belum cukup menjelaskan rata-rata total UN SMP. Bagi peneliti yang ingin melakukan penelitian serupa diharapkan dapat menambahkan peubah bebas yang mungkin baik di level sekolah maupun di level wilayah.

Peningkatan mutu pendidikan daerah dapat dilakukan melalui evaluasi dan pengawasan terhadap pertumbuhan jumlah sekolah dari pemerintah daerah setempat. Di samping itu, pemerintah daerah agar membangun lebih banyak lagi sekolah negeri di banding sekolah swasta.

Data hirarki sangat banyak dijumpai di berbagai bidang kehidupan, diantaranya sosial, pendidikan, kesehatan, kependudukan, perusahaan, pemerintahan, dan lain sebagainya. Analisis tidak hanya terbatas pada 2 level, sangat dimungkinkan penggunaan untuk level yang lebih banyak lagi.

REFERENSI

- [1] Gelman, Andrew & Jennifer Hill, *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press, 2007.
- [2] Goldstein, *Multilevel Statistical Models 4th Edition*. United Kingdom: University of Bristol, 2011.
- [3] Hox, Joop, *Multilevel Analysis Techniques and Application*. New Jersey: Lawrence Erlbaum Associates, 2002.
- [4] West B., et al, *A Practical Guide Using Statistical Software*. Taylor and Francis Group: Chapman & Hall /CRC, 2007.