

## PEMILIHAN MODEL REGRESI LINIER MULTILEVEL TERBAIK (Choice the Best Linear Regression Multilevel Models)

Bertho Tantular<sup>1</sup>, Aunuddin<sup>2</sup>, Hari Wijayanto<sup>2</sup>

<sup>1</sup>Jurusan Statistika FMIPA Universitas Padjadjaran  
<sup>2</sup>Departemen Statistika FMIPA Institut Pertanian Bogor

E-mail : <sup>1</sup>[berthovens@yahoo.co.id](mailto:berthovens@yahoo.co.id)

### **Abstract**

*Linear regression models is used to describe relationship between dependent variable and independent variables. In a survey research, data was used often have hierarchical structure or nested structure. In this research, independent variables can be defined at any level of the hierarchy but dependent variable can only be defined at the lowest level of the hierarchy. Multilevel regression models is one of the methods can be used to analyze this data. Some authors purpose many models can be used to analyze data with hierarchical structures. Deviance as  $-2 \log$  likelihood was defined as the measure goodness of fit. The difference of the deviance for two nested models was a method for comparing that two models.*

**Keywords :** Hierarchy Data, Multilevel Models, Maximum Likelihood, Deviance

### **PENDAHULUAN**

Penelitian sosial seringkali terkonsentrasi pada masalah bagaimana menelusuri hubungan antara individu dengan lingkungannya. Konsep umum suatu individu berkorelasi dengan komunitas sosialnya adalah bahwa suatu individu dipengaruhi lingkungan sosial tempat mereka berada, dan sifat-sifat dari lingkungan sosial tersebut terbentuk oleh individu-individu yang membuat lingkungan tersebut. Secara umum individu dan lingkungan sosial merupakan suatu sistem hierarki atau dapat dikatakan sebagai suatu struktur tersarang (*nested*). Penelitian data yang berstruktur hierarki peubah-peubah dapat didefinisikan pada tingkat individu dan pada tingkat lingkungan. (Hox 2002)

Beberapa peneliti telah membuat beberapa pendekatan untuk menganalisis data berstruktur hierarki. Ringdal (1992) menyebutkan bahwa pada awalnya analisis digunakan tanpa memperhatikan informasi mengenai keanggotaan individu dalam lingkungan meskipun data yang diperoleh berisi informasi tersebut. Hal ini mengakibatkan ketidakpuasan pada hasil analisisnya karena tidak bisa didapatkan simpulan yang lebih khusus untuk masing-masing lingkungan. Selain itu secara teori, mengabaikan informasi ini dapat menimbulkan masalah dalam inferensinya.

Jones & Steenbergen (1997) menyebutkan bahwa masalah yang muncul akibat mengabaikan informasi lingkungan adalah munculnya heteroskedastisitas dalam galat. Pendekatan lain untuk menganalisis data berstruktur hierarki adalah dengan cara membuat model-model yang terpisah

untuk setiap taraf pada tingkat lingkungan. Pendekatan ini menimbulkan masalah yaitu banyak informasi mengenai lingkungan menjadi tidak tercakup. Masalah lain yang muncul dari pendekatan ini adalah bahwa interaksi antar faktor dari tingkat yang berbeda tidak bisa didapatkan.

Pendekatan lain yang lebih baik adalah dengan menggunakan model regresi peubah boneka. Model regresi dengan peubah boneka dapat digunakan untuk mengatasi masalah heterogenitas. Akan tetapi model regresi dengan peubah boneka tetap tidak dapat mengatasi apabila ada hubungan antara peubah pada tingkat yang berbeda.

Model multilevel mulai diperkenalkan oleh Goldstein (1995). Model ini disebutkan dapat mengatasi semua masalah yang muncul dari data dengan struktur hierarki. Dalam model multilevel, struktur hierarki didefinisikan sebagai level. Tingkat yang paling rendah yaitu individu disebut level 1 dan tingkat yang lebih tinggi yaitu lingkungan disebut level 2. Model multilevel selain dapat menentukan keragaman antar lingkungan juga dapat menunjukkan korelasi antar dua individu yang pada model lain diasumsikan tidak ada. Selain itu model multilevel juga dapat mengukur interaksi yang mungkin terjadi antara peubah pada tingkat yang berbeda.

Penelitian ini bertujuan menentukan faktor-faktor yang memengaruhi respon pada model regresi linier multilevel dengan memperhitungkan keragaman data yang disebabkan oleh adanya pengaruh acak kelompok. Selain itu juga akan dijelaskan cara menentukan model regresi linier terbaik untuk data berstruktur hierarki atau data

multilevel dan menjelaskan adanya keragaman yang dapat memengaruhi respon pada setiap level dalam model tersebut

## METODOLOGI PENELITIAN

### Data Penelitian

Data yang digunakan adalah data mengenai nilai ujian pertama dan nilai ujian akhir semester mahasiswa pascasarjana Institut Pertanian Bogor mata kuliah Analisis Statistika (STK511) pada tahun ajaran 2008/2009. Data tersebut terdiri atas 5 kelas dimana masing-masing kelas terdiri atas beberapa program studi. Mahasiswa-mahasiswa yang mengambil mata kuliah STK511 berada dalam program studi-program studi tersebut. Setiap mahasiswa dicatat nilai ujian pertama dan ujian akhir semester. Dari data tersebut ingin diketahui apakah nilai ujian pertama mempengaruhi nilai ujian akhir dan bagaimana model yang tepat untuk menggambarkan hubungan itu.

Peubah yang digunakan untuk penelitian ini adalah nilai ujian akhir digunakan sebagai peubah respon dan nilai ujian pertama sebagai peubah penjelasnya. Semua peubah yang digunakan diukur pada level pertama.

### Metode Penelitian

Struktur data yang digunakan adalah mahasiswa merupakan satuan pengamatan pada level 1, program studi level 2 dan kelas adalah level 3. Untuk keperluan analisis penelitian ini ditentukan notasi sebagai berikut :

$i = 1, 2, 3, \dots, n_{jk}$  menyatakan mahasiswa dalam program studi  $j$  kelas  $k$

$j = 1, 2, 3, \dots, m_k$  menyatakan program studi dalam kelas  $k$

$k = 1, 2, 3, 4, 5$  menyatakan kelas.

Notasi tersebut digunakan dalam tahapan analisis sebagai berikut :

1. Membuat pemodelan regresi linier biasa

$$y_{ijk} = \beta_0 + \beta_1 X_{ijk} + \varepsilon_{ijk}$$

Penduga kemungkinan maksimum digunakan untuk model ini (Myers, 1994). Model regresi digunakan untuk melihat pola sebaran data secara gabungan.

2. Menggambarkan struktur tersarang pada level 3 melalui grafik untuk masing-masing kelas dan struktur tersarang pada level 2 juga melalui grafik.
3. Melakukan pengujian terhadap adanya struktur tersarang dari level 3 dan level 2 dengan menggunakan perbedaan *deviance* (Jones dan Steinberg, 1997)

$$diff = D_1 - D_2 \quad (1)$$

dengan  $D_1$  adalah *deviance* pada model komponen ragam 3-level dan  $D_2$  adalah *deviance* pada model komponen ragam 2-level

4. Membangun pemodelan multilevel intersep acak dengan memperhatikan pengaruh kelas dan pengaruh program studi

$$y_{ijk} = \beta_{0jk} + \beta_{1jk} X_{ijk} + \varepsilon_{ijk} \quad (\text{Model level 1})$$

$$\beta_{0jk} = \gamma_{00k} + u_{0jk} \quad (\text{Model untuk level 2})$$

$$\gamma_{00k} = \delta_{000} + v_{0k} \quad (\text{Model untuk level 3})$$

Model ini digunakan untuk menentukan ada tidaknya perbedaan intersep dari setiap kelas/program studi. Penduga kemungkinan maksimum dan galat bakunya

$$\hat{\beta} = (X' V^{-1} X)^{-1} X' V^{-1} y \quad (2)$$

$$cov(\hat{\beta}) = (X' V^{-1} X)^{-1} \quad (3)$$

dengan  $V$  adalah matriks blok diagonal parameter acak. Penduga (2) dan (3) diperoleh melalui proses iterasi. (Hox 2002)

5. Membangun pemodelan multilevel koefisien acak

$$y_{ijk} = \beta_{0jk} + \beta_{1jk} X_{ijk} + \varepsilon_{ijk} \quad (\text{Model level 1})$$

$$\beta_{0jk} = \gamma_{00k} + u_{0jk} \quad (\text{Model untuk level 2})$$

$$\beta_{1jk} = \gamma_{10k} + u_{1jk}$$

$$\gamma_{00k} = \delta_{000} + v_{0k} \quad (\text{Model untuk level 3})$$

$$\gamma_{10k} = \delta_{100} + v_{1k}$$

Model ini digunakan untuk menentukan ada tidaknya perbedaan kemiringan (*slope*) dari setiap kelas/program studi. Penduganya adalah Persamaan (2).

6. Membandingkan model-model tersebut dengan menggunakan pengujian perbedaan *deviance* pada Persamaan (1) untuk diperoleh model terbaiknya.
7. Menentukan keragaman yang dapat dijelaskan pada setiap level menggunakan persamaan berikut (Bliese 2006, Kramer 2005) yaitu :

$$R_1^2 = 1 - \frac{\hat{\sigma}_{ep}^2}{\hat{\sigma}_{e0}^2} \quad (4)$$

dengan  $\hat{\sigma}_{ep}^2$  penduga ragam residu pada level 1 dengan  $p$  peubah penjelas dan  $\hat{\sigma}_{e0}^2$  penduga ragam residu pada level 1 tanpa peubah penjelas,

$$R_2^2 = 1 - \frac{\hat{\sigma}_{u0p}^2}{\hat{\sigma}_{u0}^2} \quad (5)$$

dengan  $\hat{\sigma}_{u0p}^2$  penduga ragam residu pada level 2 dengan  $p$  peubah penjelas dan  $\hat{\sigma}_{u0}^2$  penduga ragam residu pada level 2 tanpa peubah penjelas.

## HASIL DAN PEMBAHASAN

### Deskripsi Data Nilai Ujian

Data nilai ujian STK511 terdiri atas 447 mahasiswa yang tersarang didalam 36 program studi. Program studi-program studi tersebut tersarang didalam 6 kelas. Dari setiap mahasiswa tersebut diukur nilai ujian pertama dan nilai ujian

akhir. Kemudian ingin diketahui apakah ada pengaruh ujian pertama terhadap ujian akhir. Perbedaan karakter mahasiswa antar program studi tentu akan ada pengaruhnya terhadap pola belajar mahasiswa dalam program studi tersebut.

Perbedaan dosen pengajar mata kuliah STK511 pada setiap kelas diduga akan memberikan pengaruh terhadap hasil penilaian nilai ujian meskipun telah ditetapkan standar dalam penilaian. Oleh sebab itu keragaman antar program studi yang ada di dalam kelas dan keragaman antar kelas harus diperhitungkan dalam model yang akan dibangun.

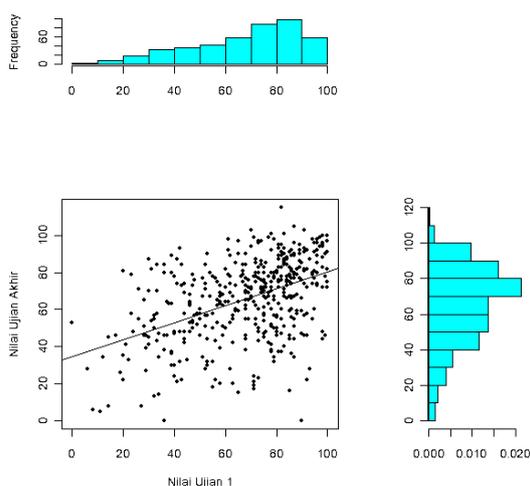
### Pemodelan Data Nilai Ujian

Model regresi dalam hal ini digunakan sebagai model dasar. Model ini menjadi pembandingan bagi model-model lain. Berikut adalah hasil pendugaan parameter dengan metode kuadrat terkecil dan plot grafik analisis regresi yang diperoleh menggunakan *software* R adalah sebagai berikut

Tabel 1. Hasil Pendugaan Parameter Model Regresi untuk Data Nilai Ujian

	Dugaan	Galat Baku	t	Nilai-p
Intersep	34.6047	3.03589	11.09	0.000
Ujian Pertama	0.45643	0.04244	10.75	0.000

Berdasarkan Tabel 1 jelas terlihat bahwa nilai ujian pertama berpengaruh sangat nyata terhadap nilai ujian akhir pada taraf 5%. Nilai koefisien determinasi ( $R^2$ ) adalah 0.3084 yang artinya Nilai Ujian Akhir hanya sebesar 20.59% dapat dijelaskan oleh peubah Nilai Ujian Pertama, sisanya dijelaskan oleh peubah lain.

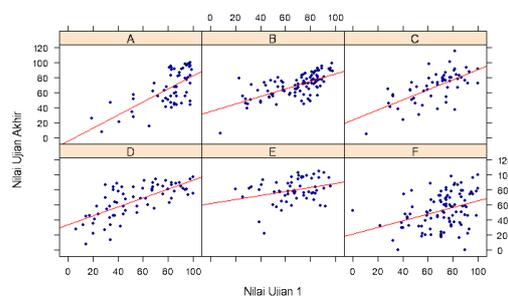


Gambar 1. Plot Analisis Regresi Data Nilai Ujian

Pada Gambar 1 tampak data menyebar secara merata mengikuti pola garis regresinya. Dari gambaran tersebut terlihat bahwa data meskipun

mengikuti pola garis regresinya tetapi mempunyai keragaman yang cukup besar.

**Penjelasan Struktur Tersarang Model Multilevel**  
Untuk Level 3 data nilai ujian dapat dideskripsikan dengan cara membuat plot grafik analisis regresi dari masing-masing kelas. Dari plot ini dapat dilihat keragaman pola regresi antar kelas.



Gambar 2. Plot Regresi untuk Masing-masing Kelas

Dari Gambar 2 terlihat pola dari setiap kelas memiliki kemiripan, kecuali untuk Kelas A. Pola sebaran data pada plot untuk Kelas A terlihat cenderung berkumpul di bagian kanan atas yang berarti mayoritas mahasiswa Kelas A memiliki nilai ujian yang tinggi. Dari kemiripan pola ini menimbulkan dugaan bahwa kelas tidak mempunyai pengaruh yang nyata terhadap Nilai Ujian Akhir. Hal ini akan dibuktikan dengan menggunakan prosedur berikut ini.

Analisis data menggunakan model multilevel diawali dengan membuat model tanpa melibatkan peubah penjelas. Model ini disebut sebagai model komponen ragam (*Variance Component Model*). Model ini digunakan untuk melihat ada tidaknya pengaruh dari struktur multilevel (tersarang) dari data tersebut, selain itu dapat ditentukan pula besarnya proporsi keragaman dalam level yang lebih tinggi (West 2007).

Tabel 2 memperlihatkan nilai komponen ragam untuk masing-masing level pada data nilai ujian. Pada Tabel 2 terlihat bahwa nilai ragam pada level 3 cukup kecil dibandingkan dengan nilai ragam level 2 dan level 1. Hal ini mengindikasikan bahwa tidak adanya keragaman dalam kelas atau dapat juga dikatakan antara kelas cenderung homogen.

Tabel 2. Nilai Dugaan Parameter Acak Model Tiga-Level tanpa Peubah Penjelas

Parameter	Dugaan
$\sigma^2_{v0}$ (level 3)	36.49518
$\sigma^2_{u0}$ (level 2)	95.35708
$\sigma^2_{e0}$ (level 1)	327.48897

Berdasarkan nilai dugaan parameter tersebut dapat diperoleh nilai koefisien korelasi *intra class* sebagai berikut :

$$\rho_{ps} = \frac{36.49518 + 95.35708}{36.49518 + 95.35708 + 327.48897} = 0.287$$

$$\rho_{kelas} = \frac{36.49518}{36.49518 + 95.35708 + 327.48897} = 0.0795$$

Nilai ini mengandung arti bahwa proporsi ragam pada level program studi sebesar 28,7 % dan proporsi ragam pada level kelas sebesar 7,95 %. Selain itu dapat dijelaskan pula bahwa korelasi antara dua mahasiswa dalam satu program studi sebesar 0.287. Sedangkan dalam kelas sebesar 0.0795. Memperhatikan nilai ini bahwa korelasi dalam kelas cukup kecil tetapi perlu diuji keberadaan keragaman dalam kelas.

Selanjutnya dibangun model 2 level tanpa peubah penjelas. Model ini digunakan untuk melihat apakah dengan dikeluarkannya struktur kelas dari model tidak memengaruhi keragaman pada level program studi dan tidak memengaruhi koefisien korelasi *intra-class* untuk level 2. Tabel berikut dapat menjelaskan hal tersebut

Tabel 3. Nilai Dugaan Parameter Acak Model Dua-Level tanpa Peubah Penjelas

Parameter	Dugaan
$\sigma_{u0}^2$ (level 2)	128.5939
$\sigma_{e0}^2$ (level 1)	326.6247

Dari tabel 3 jelas sekali terlihat bahwa nilai ragam pada program studi menjadi lebih besar setelah struktur kelas telah dikeluarkan dalam model bila dibandingkan nilai ragam pada Tabel 2. Hasil perbandingan *deviance* dapat dilihat pada Tabel 4. Berdasarkan Tabel 4 terlihat bahwa perbedaan model 3-level dengan model 2-level tidak nyata pada taraf 5% artinya bahwa tidak terdapat perbedaan yang nyata antara model 2-level dengan model 3-level. Hal ini menunjukkan bahwa keragaman dalam kelas tidak berpengaruh terhadap nilai ujian akhir.

Tabel 4. Nilai *Deviance* untuk Model 3-level dan Model 2-level Data Nilai Ujian

Model	Deviance	Parameter	Diff	db	Nilai-p
3-level	3908.37	4			
2-level	3909.60	3	1.23	1	0.27

Keragaman Level 2 dapat dideskripsikan secara visual menggunakan pola gambar plot regresi dari setiap program studi. Plot regresi ditampilkan untuk masing-masing program studi di dalam kelas. Pada Gambar 3 terlihat terlihat bahwa plot regresi Kelas

A ada tiga program studi yaitu IPN, SVK dan THH memiliki pola yang hampir sama, sementara program studi KVT dan MEJ memiliki pola yang sama sekali berbeda dengan yang lainnya. Keragaman pada pola plot regresi dari 8 program studi di dalam kelas B. Empat program studi memiliki pola yang hampir sama yaitu PPN, SPD, MAT dan IPH. Tiga program studi lainnya memiliki pola yang mirip yaitu KMP, MAN dan GMS. Sementara program studi APL memiliki pola yang sama sekali berbeda dengan yang lainnya. Untuk Kelas C plot regresi Pola plot regresi Enam program studi memiliki pola yang hampir sama yaitu SPT, IPH, TPT, ITP, IFO dan BRP. Sementara program studi MKM dan INP memiliki pola yang sama sekali berbeda dengan yang lainnya. Hal yang sama terlihat pada Kelas D untuk program studi TIP dan STK mempunyai pola yang hampir sama, sementara program studi TPP dan ATT memiliki pola yang berbeda dengan yang lainnya. Sementara untuk Kelas E perbedaan plot regresi terlihat pada program studi BIK dan BIO seperti terlihat pada Gambar 3 dan program studi FIT pada kelas F berbeda dengan yang lainnya. Beberapa perbedaan pola garis regresi juga disebabkan oleh sedikitnya jumlah data dalam suatu program studi seperti pada program studi IFO, IPH, MKM dan PS5 yang datanya kurang dari 5 buah.

Secara umum dari Gambar 3 plot regresi untuk masing-masing program studi terlihat adanya perbedaan pola. Hal ini mengindikasikan adanya pengaruh program studi terhadap nilai ujian akhir STK511.

### Model Intersep Acak

Dalam model ini ingin diketahui adanya pengaruh acak dari program studi. Hasil dugaan parameter untuk model ini dapat dilihat pada tabel berikut.

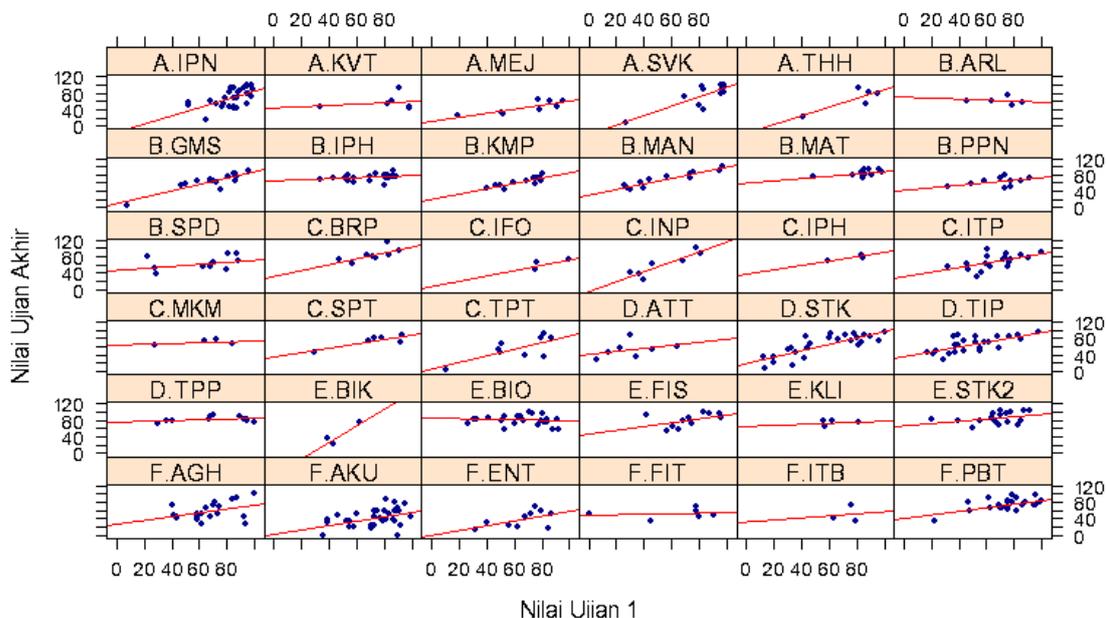
Tabel 5. Hasil pendugaan parameter model intersep acak untuk Data Nilai Ujian

	Dugaan	Galat Baku	db	t	Nilai-p
Intersep	31.080	3.228	410	9.629	0
ujian pertama	0.5058	0.037	410	13.66	0

Dilihat dari tabel diatas jelas terlihat bahwa nilai ujian pertama berpengaruh nyata terhadap nilai ujian akhir pada taraf 5%.

### Model Koefisien Acak

Dalam model koefisien acak peubah Nilai ujian pertama dipengaruhi oleh keragaman dalam program studi. Hasil dugaan parameter untuk model koefisien acak dapat dilihat pada tabel 6. Dilihat dari tabel 6 jelas terlihat bahwa nilai ujian pertama berpengaruh sangat nyata terhadap nilai ujian akhir pada taraf 5%.



Gambar 3 Plot Regresi Masing-masing Program Studi dalam Kelas

Tabel 6. Hasil Pendugaan Parameter Model Koefisien Acak untuk Data Nilai Ujian

	Dugaan	Galat Baku	db	t	Nilai-p
Intersep	31.2694	4.7408	410	0.06	0
Ujian Pertama	0.5012	0.0571	410	8.78	0

**Pemilihan Model Terbaik**

Seperti pada contoh data yang pertama, untuk membandingkan model-model untuk data ini digunakan nilai *deviance*. Tabel berikut menjelaskan perbedaan *deviance* untuk ketiga model.

Tabel 7. Nilai *Deviance* untuk Setiap Model Nilai Ujian untuk Data Nilai Ujian

Model	Deviance	Parameter	Diff	db	Nilai-p
Regresi	3914.98	3			
Intersep acak	3760.34	4	154.64	1	0.00000
Koefisien acak	3743.15	6	17.19	2	0.00002

Dari Tabel 7 terlihat bahwa Model koefisien acak mempunyai nilai *deviance* yang paling kecil. Hal ini menyiratkan bahwa model koefisien acak merupakan model yang lebih baik dari model lainnya. Hasil pengujian memperkuat simpulan ini, berdasarkan pengujian Model Regresi jauh lebih buruk dibandingkan dengan model intersep acak, sedangkan model koefisien acak yang lebih baik

dibandingkan dengan model intersep acak. Sehingga dapat disimpulkan bahwa model koefisien acak merupakan model yang terbaik untuk permasalahan ini.

Hasil pengujian pada Tabel 7 juga memberikan arti bahwa keragaman yang ada pada kemiringan (*slope*) berarti sehingga dapat disimpulkan bahwa program studi memberikan pengaruh nyata terhadap nilai ujian akhir.

Persamaan yang dapat dibentuk untuk menggambarkan permasalahan ini adalah sebagai berikut

$$y_{ijk} = \beta_{0jk} + \beta_{1jk}X_{ijk} + \varepsilon_{ijk} \quad (\text{Level 1})$$

$$\beta_{0jk} = 31.2694 + u_{0jk} \quad (\text{Level 2})$$

$$\beta_{1jk} = 0.5012 + u_{1jk}$$

atau dalam bentuk model campuran (West, 2006) adalah substitusi Level 2 ke Level 1 yaitu

$$y_{ijk} = 31.2694 + 0.5012X_{ijk} + \{u_{1jk}X_{ijk} + u_{0jk} + \varepsilon_{ijk}\}$$

Dari model tersebut dapat dijelaskan bahwa apabila nilai ujian pertama meningkat satu satuan maka nilai ujian akhir akan meningkat sebesar 0.5012 satuan, tetapi peningkatan ini juga dipengaruhi oleh keragaman antar program studi.

Tabel 8 menjelaskan nilai dugaan komponen ragam galat untuk model koefisien acak. Dari Tabel 8 terlihat bahwa pengaruh acak pada koefisien nilai ujian pertama memperbesar ragam pada intersep yang berarti bahwa nilai ujian pertama

menyebabkan keragaman antar program studi menjadi semakin jelas.

Tabel 8. Nilai Dugaan Komponen Ragam Model Koefisien Acak untuk Data Nilai Ujian

Komponen	Ragam
Intersep Level 2 ( $u_0$ )	492.590
Nilai ujian pertama Level 2 ( $u_1$ )	0.05689
Residual ( $e_0$ )	204.4506

### Keragaman yang Dapat Dijelaskan Pada Setiap Level

Keragaman pada setiap level yang dapat dijelaskan oleh model dapat diperoleh dari perbandingan model dengan melibatkan peubah penjelas dan model tanpa melibatkan peubah penjelas. (Kramer, 2005)

Untuk memperoleh nilai keragaman yang dapat dijelaskan pada setiap level untuk data ini digunakan model tanpa peubah penjelas level 1 sebagai dasar dan model intersep acak. Dalam data ini tidak digunakan model koefisien acak yang merupakan model terbaik pada bagian sebelumnya karena model koefisien acak berfungsi sebagai model dasar bagi keragaman yang ada pada *slope* pada setiap level dan tidak dapat digunakan untuk menjelaskan keragaman yang ada pada intersep.

Tabel 9. Nilai Dugaan Komponen Ragam Model Intersep Acak untuk Data Nilai Ujian

Komponen	Ragam
Intersep Level 2 ( $u_0$ )	122.9820
Residual ( $e_0$ )	226.8315

Nilai ragam ini akan lebih berarti apabila dapat dicari nilai keragaman yang dapat dijelaskan pada setiap level. Untuk level 1 dan level 2 dapat digunakan Persamaan 4 dan Persamaan 5. Nilai keragaman ini dapat diperoleh menggunakan dugaan pada Tabel 3 dan Tabel 9 yaitu sebagai berikut:

$$R_1^2 = 1 - \frac{226.8315}{326.6247} = 0.3055$$

$$R_2^2 = 1 - \frac{122.9820}{128.5934} = 0.0434$$

Dari hasil tersebut terlihat bahwa keragaman yang dapat dijelaskan oleh model pada level 1 sebesar 30,55 % dan keragaman yang dapat dijelaskan oleh model pada level 2 sebesar 4,34 %. Hal ini mengandung arti bahwa keragaman program studi memberikan kontribusi terhadap nilai ujian akhir sebesar 3,43% dan sebesar 30,55% kontribusi diberikan oleh nilai ujian pertama. Keragaman antar program studi secara umum dapat disebabkan oleh

adanya perbedaan pada program studi pendidikan S1 yang ditempuh mahasiswa.

Selain itu, perbedaan universitas tempat mahasiswa menempuh program studi S1 juga berperan dalam keragaman ini karena universitas yang berbeda tentu mempunyai tingkat kompetensi yang berbeda-beda pula terhadap mata kuliah Statistika yang pernah diperolehnya.

### SIMPULAN

Model multilevel dapat digunakan untuk data dengan struktur hierarki yaitu data dengan struktur tersarang. Model multilevel dapat mengatasi masalah-masalah yang muncul dari data berstruktur hierarki yaitu masalah heterogenitas, masalah hubungan antara peubah pada level yang berbeda dan masalah terdapatnya komponen acak dalam model. Selain daripada itu model multilevel juga dapat mengukur keragaman dari setiap level.

Struktur data nilai ujian STK511 memiliki 3 level hierarki, mahasiswa tersarang dalam program studi dan program studi tersarang dalam kelas. Dapat disimpulkan bahwa struktur kelas tidak memiliki pengaruh terhadap respon. Hal ini dijelaskan oleh hasil pengujian, sehingga model yang digunakan hingga level kedua. Hasil ini menunjukkan bahwa perbedaan dosen pengajar mata kuliah STK511 tidak memengaruhi nilai ujian.

Model koefisien acak merupakan model yang paling cocok untuk menggambarkan pengaruh nilai ujian pertama terhadap nilai ujian akhir dengan memperhitungkan adanya keragaman antar program studi. Hasil yang diperoleh adalah program studi yang berbeda akan memberikan pengaruh yang berbeda dalam peningkatan nilai ujian akhir selain pengaruh yang diberikan oleh peningkatan nilai ujian pertama.

Keragaman nilai ujian akhir dapat dijelaskan oleh keragaman nilai ujian pertama dan keragaman antar program studi. Hal ini menunjukkan adanya pengaruh program studi terhadap nilai ujian akhir. Keragaman antar program studi secara umum dapat disebabkan oleh adanya perbedaan pada program studi pendidikan S1 yang ditempuh mahasiswa. Selain itu, perbedaan universitas tempat mahasiswa menempuh program studi S1 juga berperan dalam keragaman ini karena universitas yang berbeda tentu mempunyai tingkat kompetensi yang berbeda-beda pula terhadap mata kuliah Statistika yang pernah diperolehnya.

### DAFTAR PUSTAKA

- Bliese P. 2006. *Multilevel Models in R* (2.2). R Development Core Team.  
 Goldstein H. 1995. *Multilevel Statistical Models 2<sup>nd</sup> Ed*. London. Arnold London.  
 Hox JJ. 2002. *Multilevel Analysis: Techniques and Applications*. New Jersey. Lawrence Erlbaum Associates Publishers.

- Jones BS & Steenbergen MR. 1997. *Modelling Multilevel Data Structures*. Paper prepared in 14<sup>th</sup> annual meeting of the political methodology society. Columbus. OH.
- Kramer M. 2005. *R<sup>2</sup> Statistics for Mixed Models*. Published Paper in Biometrical Consulting Service, ARS (Beltsville, MD), USDA.
- Myers RH. 1990. *Classical and Modern Regression with Applications*. Boston : PWS-KENT Publishing Company.
- Ringdal K. 1992. Methods for Multilevel Analysis. *Acta Sociologica* 35:235-243.
- West BT, Welch KB, Galecki AT. 2006. *Linear Mixed Models: A Practical Guide Using Statistical Software*. Boca Raton : Chapman & Hall.