

Regresi Logistik Dua Level Respon Biner untuk Pemodelan Data Status Kelulusan Mata Kuliah Statistika Mahasiswa Fakultas X, Y dan Z Universitas Islam Bandung Angkatan 2019

Alma Mariana Sofiah*, Nusar Hajarisman

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

*alma.mariana18@gmail.com, nusarhajarisman@unisba.ac.id

Abstract. Two-level binary logistic regression is a multilevel analysis used to analyze data that has a two-level hierarchical structure with binary response data. Hierarchical data structure is data with observation units nested in higher units. Two-level regression is the simplest multilevel model where the first level is individual data and the second level is group data. In this study, we will discuss the use of two-level binary logistic regression on data on the Passing Status of Statistics Courses of Bandung Islamic University Students Class of 2019. The data used is secondary data obtained from PSITEK Bandung Islamic University, with a total data of 1173 students nested into 3 groups (Faculty). The research variable in this study is the status of passing the statistics course as the dependent variable (Y), with two categories, namely 0 for students who do not pass and 1 for students who pass, as well as gender, regional origin and GPA as predictor variables (X) at level 1 and science clumps as predictor variables (Z) at level 2. The parameter estimator is carried out using the Maximum Likelihood via Quadrature method. Based on the results of the application of 2-level binary logistic regression, the factors that influence the response are the variables of regional origin and GPA (level 1/student) and clumps of knowledge (level 2/Faculty).

Keywords: *Binary Response, Maximum Likelihood Via Quadrature, Course Pass Status.*

Abstrak. Regresi logistik biner dua-level merupakan analisis multilevel yang digunakan untuk menganalisis data yang mempunyai struktur hirarki dua level dengan data respon biner. Struktur data hirarki adalah data dengan unit-unit observasi yang bersarang pada unit yang lebih tinggi. Regresi dua-level merupakan model multilevel yang paling sederhana dimana level kesatu data individu dan level kedua data kelompok. Dalam penelitian ini, akan membahas penggunaan regresi logistik biner dua level pada data Status Kelulusan Mata Kuliah Statistika Mahasiswa Universitas Islam Bandung Angkatan 2019. Data yang digunakan adalah data sekunder yang diperoleh dari PSITEK Universitas Islam Bandung, dengan total data sebanyak 1173 mahasiswa yang tersarang kedalam 3 kelompok (Fakultas). Variabel penelitian dalam penelitian ini adalah status lulus mata kuliah statistika sebagai variabel dependen (Y), dengan dua kategori yaitu 0 untuk mahasiswa yang tidak lulus dan 1 untuk mahasiswa yang lulus, serta jenis kelamin, asal daerah dan IPK sebagai variabel prediktor (X) pada level 1 dan rumpun ilmu sebagai variabel prediktor (Z) pada level 2. Penaksir parameter dilakukan menggunakan metode Maximum Likelihood via Quadrature. Berdasarkan hasil penerapan regresi logistik biner 2-level menyatakan bahwa faktor-faktor yang mempengaruhi respon adalah variabel asal daerah dan IPK (level 1/mahasiswa) dan rumpun ilmu (level 2/Fakultas).

Kata Kunci: *Respon Biner, Maximum Likelihood Via Quadrature, Status Lulus Mata Kuliah.*

A. Pendahuluan

Analisis regresi logistik dipecah menjadi beberapa datarespons termasuk respons biner, multinomial, ordinal, dan Poisson. Analisis regresi logistik yang sangat sederhana adalah analisis regresi logistik dengan respon biner. Regresi logistik biner digunakan untuk menganalisis hubungan antara suatu variabel respon dengan beberapa variabel prediktor, dimana variabel respon berupa data kualitatif dikotomis, dengan nilai 1 menunjukkan adanya suatu kejadian dan nilai 0 menunjukkan tidak adanya sebuah kejadian [3]. Sementara itu, data yang memiliki respon biner tidak hanya digunakan dalam regresi logistik, tetapi juga mulai dikembangkan dalam berbagai metode regresi [11].

Penelitian dalam berbagai bidang sering dilakukan dengan data yang dimana unit-unit yang diteliti diambil dari kelompok-kelompok yang membawahi unit-unit yang diteliti, atau sering disebut dengan data yang berstruktur secara hierarkis. Dalam data dengan struktur hirarki memang terdapat pengaruh kelompok yang membawahi unit-unit yang diteliti [7]. Jika ingin menganalisis variasi antara kelompok-kelompok yang membawahi unit-unit yang diteliti pada data hirarkis, maka penggunaan model regresi biasa tidak tepat karena analisis regresi biasa tidak memberikan informasi tersebut. Oleh karena itu, model hierarkis atau multilevel merupakan analisis yang tepat untuk mengatasi kekurangan pada model regresi biasa saat menganalisis data yang berstruktur hierarkis. Model multilevel adalah teknik statistik yang berevolusi dari regresi sederhana. Perkembangan ini didasarkan pada fakta bahwa dalam kasus penelitian seringkali terdapat perbedaan latar belakang responden yang sedang diteliti, maka data tersebut memiliki struktur yang bertingkat (*hierarchy*) dan berkelompok (*cluster*). Bila variabel respon yang digunakan adalah biner, maka metode analisis yang dapat dilakukan adalah regresi logistik biner multilevel.

Hasil belajar adalah keterampilan yang dimiliki siswa setelah memperoleh pengalaman belajar [10]. Ada beberapa faktor yang dapat mempengaruhi hasil belajar. Hasil penelitian menemukan bahwa perempuan memiliki kecerdasan yang lebih tinggi dibandingkan laki-laki, terutama dalam studi di perguruan tinggi [2]. Hasil penelitian juga menunjukkan bahwa perbedaan latar belakang sekolah dan tempat tinggal berdampak pada pembelajaran atau hasil belajar [8]. Didalam penelitian juga dijelaskan bahwa ada hubungan signifikan antara hasil belajar dengan IPK mahasiswa [6]. Keberhasilan mahasiswa dalam belajar tercermin dari penerapan model dan media pembelajaran yang sesuai dengan karakteristik siswa [4]. Hal tersebut dapat mempengaruhi kreativitas, kinerja, dan produktivitas pada mahasiswa. Berkat penyesuaian dengan karakteristik yang berbeda-beda berdasarkan rumpun ilmu dan kebutuhan setiap kelompok, mahasiswa dapat berhasil melaksanakan pembelajaran.

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah “Bagaimana penerapan regresi logistik dua level respon biner untuk pemodelan data status kelulusan mata kuliah statistika mahasiswa Fakultas X, Y dan Z Universitas Islam Bandung angkatan 2019”, selanjutnya tujuan dalam penelitian ini diuraikan dalam pokok-pokok sbb.

1. Mengetahui faktor – faktor yang berpengaruh terhadap status kelulusan pada mata kuliah yang berhubungan dengan statistika mahasiswa angkatan 2019 Universitas Islam Bandung.
2. Mengetahui faktor-faktori yang mempengaruhi status kelulusan pada mata kuliah yang berhubungan dengan statistika mahasiswa angkatan 2019 Universitas Islam Bandung yang disebabkan oleh adanya keragaman didalam mahasiswa (Level 1).
3. Mengetahui status kelulusan pada mata kuliah yang berhubungan dengan statistika mahasiswa angkatan 2019 Universitas Islam Bandung yang disebabkan oleh adanya keragaman didalam fakultas (Level 2).

B. Metodologi Penelitian

Bahan yang digunakan untuk mengaplikasikan metode yang akan dibahas dalam penelitian ini adalah data sekunder yang diperoleh dari Pengembangan Sistem Informasi dan Teknologi Universitas Islam Bandung. Data yang digunakan adalah data hasil penilaian mata kuliah yang berhubungan dengan statistika mahasiswa Fakultas X, Y dan Z Universitas Islam Bandung

angkatan 2019 sebanyak 1173 mahasiswa. Variabel respon berupa status lulus mata kuliah yang berbentuk biner yang dinyatakan dalam bentuk “1: Lulus” dan “0: Tidak lulus” untuk dua buah kategori. Sedangkan variabel prediktor pada setiap level adalah:

Level 1 (mahasiswa):

1. Jenis kelamin (0: Laki-laki dan 1: Perempuan)
2. Asal daerah (0: Luar pulau Jawa dan 1: Luar Pulau Jawa)
3. IPK (skala 0-4)

Level 2 (fakultas)

1. Rumpun ilmu (0: Soshum dan 1: Saintek)

Metode Analisis Data

Analisis Regresi Logistik Biner

Analisis regresi logistik merupakan suatu metode terbaik untuk analisis regresi dengan variabel respon dikotomis atau biner [5]. Model regresi logistik yang paling sederhana adalah model regresi logistik untuk satu variabel prediktor X dengan variabel respon biner Y . Nilai variabel $Y = 1$ menunjukkan adanya suatu karakteristik dan $Y = 0$ menunjukkan tidak adanya suatu karakteristik.

[9] Bentuk umum regresi logistik dengan melibatkan dua atau lebih variabel prediktor (X_i) dapat ditulis sebagai berikut:

$$E(Y|X) = \frac{e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)}}{1 + e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)}} \quad (1)$$

Fungsi $\pi(X_i)$ adalah fungsi nonlinier sehingga perlu menggunakan fungsi hubung logit untuk memperoleh fungsi yang linier agar dapat dilihat hubungan antara variabel respon (Y_i) dengan variabel prediktornya (X_i). Bentuk logiti dari $\pi(X_i)$ dinyatakan sebagai $g(X)$, yaitu sebagai berikut:

$$g(X) = \ln \left[\frac{\pi(X_i)}{1 - \pi(X_i)} \right] \quad (2)$$

Sehingga didapat model regresi logistik seperti pada persamaan (3)

$$\ln \left[\frac{\pi(X_i)}{1 - \pi(X_i)} \right] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad (3)$$

dengan:

$X_1, X_2, \dots, X_p$: variabel prediktor

$\beta_1, \beta_2, \dots, \beta_p$: koefisien dari variabel prediktor (parameter dari model)

Analisis Regresi Multilevel

Penelitian sosial sering berfokus pada studi tentang hubungan antara individu dan lingkungan. Pengertian umum bahwa individu berhubungan dengan komunitas sosialnya adalah bahwa individu tersebut dipengaruhi oleh lingkungan sosial tempatnya tinggal dan bahwa individu yang membentuk lingkungan tersebut merupakan ciri-ciri dari lingkungan sosial tersebut. Secara umum, individu dan lingkungan sosial merupakan sistem hirarki atau dapat digambarkan sebagai struktur bersarang. Data penelitian yang berstruktur hirarki, variabel-variabelnya dapat diartikan pada level individu dan pada level lingkungan [7].

Regresi Dua-Level untuk Respon Biner dengan *Random Intercept*

Model regresi multilevel bisa diklasifikasikan menjadi dua bentuk, yakni model multilevel dengan intersep acak dan model multilevel dengan random slope [7]. Model multilevel dengan intersep acak merupakan model dimana intersep dimodelkan sebagai efek acak dari variabel pada level 2, tetapi memiliki kemiringan atau slope yang sama sehingga pengaruh setiap variabel penjelas terhadap variabel respon sama untuk masing-masing kelompok.

Pada model regresi logistik dua level, misalkan j menyatakan unit level-2 (*cluster*) dan i menyatakan unit level-1 (pengamatan tersarang). Asumsi bahwa unit level-2 adalah $j = 1, \dots$,

m dan unit level-1 adalah $i=1, \dots, n_j$ tersarang dalam tiap unit level-2 j . Jumlah total observasi level-1 dalam unit-unit level-2 adalah:

$$n = \sum_{j=1}^m n_j$$

Misalkan terdapat P variabel prediktor pada level 1 dan terdapat Q variabel prediktor pada level 2, maka bentuk representasi model multilevel biner menggunakan intersep acak dapat dilihat sebagai berikut:

Model level-1

$$\ln \left[\frac{\pi_{ij}}{(1 - \pi_{ij})} \right] = \beta_{0j} + \sum_{p=1}^P \beta_{pij} X_{pij} \quad (4)$$

dan model level-2

$$\beta_{0j} = \beta_{00} + \sum_{q=1}^Q \beta_{0q} Z_{qj} + u_{0j} \quad (5)$$

Apabila model level-1 dan level-2 digabungkan maka model intersep acak 2-level dengan respon biner diperoleh:

$$\ln \left[\frac{\pi_{ij}}{(1 - \pi_{ij})} \right] = \beta_{00} + \sum_{q=1}^Q \beta_{0q} Z_{qj} + \sum_{p=1}^P \beta_{pij} X_{pij} + u_{0j} \quad (6)$$

dimana:

β_{00} = *intercept*, merupakan rata-rata secara keseluruhan

β_{p0} = efek tetap untuk variabel penjelas ke- p (*fixed slope*) dimana $p=1, 2, \dots, P$

X_{pij} = variabel penjelas ke- p pada level 1 untuk individu ke- i pada level 1 dalam kelompok ke- j pada level 2 dimana $p=1, 2, \dots, P$,

Z_{qj} = variabel penjelas ke- q pada level 2 untuk kelompok ke- j dimana $q=1, 2, \dots, Q$,

u_{0j} = efek acak kelompok ke- j pada level 2.

Estimasi Parameter

Metode penaksir parameter yang digunakan pada penelitian ini adalah *Maximum Likelihood Estimation* (MLE). Dalam estimasi parameter model regresi logistik multilevel biner sulit jika diselesaikan secara analitik, maka digunakan pendekatan metode *Gauss Hermite Quadrature*. *Gauss-Hermite* menggantikan integrasi dengan jumlah bobot pada fungsi yang dihitung pada serangkaian titik tertentu. Secara umum dapat dirumuskan seperti persamaan

$$\frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (g(v)) \exp(-v^2) dv \approx \sum_{k=1}^K w_k g(v_k) \quad (7)$$

dimana:

w_k = Bobot *quadrature* dan v_k = Titik *quadrature*

Pengujian Signifikansi *Random Effect*

Metode estimasi kemungkinan maksimum dapat menghasilkan suatu statistik yang disebut deviance yang mampu menunjukkan seberapa cocok model dengan data [7]. Pada model berhierarki, tes *deviance* dapat digunakan dengan tujuan untuk mengetahui apakah model dengan efek acak lebih baik daripada model tanpa efek acak. Pengujian dilakukan dengan hipotesis:

$H_0: \sigma_{u0}^2 = 0$ (efek acak tidak signifikan)

$H_1: \sigma_{u0}^2 \neq 0$ (efek acak signifikan)

Statistik uji:

$$LR = -2\ln \left[\frac{L(\text{Likelihood model logistik tanpa efek acak})}{L(\text{Likelihood model logistik dengan efek acak})} \right] \quad (8)$$

Tolak H_0 jika $LR > \chi^2_{(a,v)}$ atau $p\text{-value} < \alpha$, dimana v adalah selisih jumlah parameter dari kedua model. H_0 ditolak maka dapat disimpulkan bahwa *random effect* signifikan.

Pengujian Signifikansi Parameter secara Simultan

G-test merupakan pengujian signifikansi keseluruhan variabel prediktor di dalam model secara bersama-sama. Pengujian dilakukan dengan hipotesis:

$H_0: \beta_{10} = \beta_{20} = \dots = \beta_{p0} = \beta_{01} = \dots = \beta_{0q} = 0$; (tidak ada pengaruh variabel prediktor terhadap variabel respon)

H_0 : minimal ada satu $\beta \neq 0$; (minimal ada satu variabel prediktor yang berpengaruh terhadap variabel respon)

Statistik uji:

$$G^2 = -2\ln \left[\frac{L(\text{Likelihood tanpa variabel prediktor sedangkan})}{L(\text{Likelihood dengan variabel prediktor})} \right] \quad (9)$$

Tolak H_0 jika $G^2 > \chi^2_{(a,r)}$ atau $p\text{-value} < \alpha$, dimana derajat bebas r adalah jumlah parameter di level 1 dan level 2. H_0 ditolak maka dapat disimpulkan bahwa paling tidak terdapat satu variabel prediktor yang memengaruhi variabel respons pada tingkat signifikansi α .

Pengujian Signifikansi Parameter Secara Parsial

Uji parsial digunakan untuk menentukan variabel prediktor mana yang memiliki pengaruh terhadap variabel respon. Pengujian dilakukan dengan hipotesis:

$H_0: \beta_{pq} = 0$; (tidak ada pengaruh variabel prediktor ke- p pada level 1 atau variabel penjelas ke- q pada level 2 terhadap respon)

$H_0: \beta_{pq} \neq 0$; (ada pengaruh variabel prediktor ke- p pada level 1 atau variabel penjelas ke- q pada level 2 terhadap respon)

Statistik uji:

$$W = \left[\frac{\hat{\beta}_{pq}}{SE(\hat{\beta}_{pq})} \right] \quad (10)$$

Tolak H_0 jika $|W| > Z_{1-\frac{\alpha}{2}}$ atau $-value < \alpha$. H_0 ditolak, dapat disimpulkan bahwa variabel prediktor signifikan berpengaruh terhadap variabel respons pada tingkat signifikansi α .

Interpretasi Parameter

$$\frac{(\pi_{ij}|x_{ij}, u_{0j})}{1 - (\pi_{ij}|x_{ij}, u_{0j})} = \exp[\beta_{00} + \beta_{10}x_{ij} + u_{0j}] \quad (11)$$

[1] Untuk interpretasi intersep:

Ketika $x_{ij} = 0$, *odds ratio* didalam kelompok j sama dengan $\exp[\beta_{00} + u_{0j}]$

Ketika $x_{ij} = 0$, *odds ratio* didalam sebuah kelompok sama dengan $\exp[\beta_{00}]$

Untuk interpretasi slope:

Odds Ratio didalam sebuah kelompok untuk perubahan satu unit dalam x_{ij} adalah sebagai berikut:

$$\frac{\exp(\beta_{00}) \exp(\beta_{10}(x_{ij} + 1)) \exp(u_{0j})}{\exp(\beta_{00}) \exp(\beta_{10}x_{ij}) \exp(u_{0j})} = \exp(\beta_{10}) \quad (12)$$

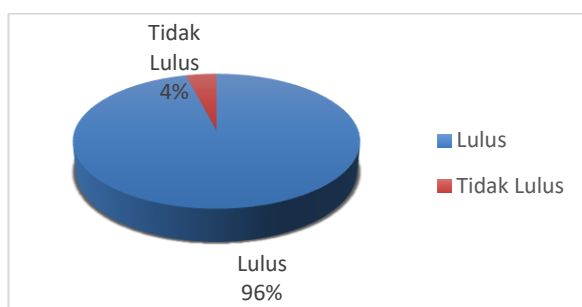
Intraclass Correlation Coefficient (ICC)

Jika data yang dimiliki adalah data dengan struktur hirarkhi sederhana, maka regresi multilevel dapat digunakan untuk memperkirakan nilai estimasi bagi korelasi intraklas [7]. Model yang digunakan untuk tujuan ini adalah model yang hanya mempunyai intersep.

$$ICC = \rho = \frac{\sigma_{u_{ij}}^2}{\sigma_{u_{ij}}^2 + \sigma_e^2}$$

C. Hasil Penelitian dan Pembahasan

Deskripsi Data



Gambar 1. Persentase Status Kelulusan Mahasiswa Fakultas X, Y dan Z

Berdasarkan gambar 1 diatas menunjukkan bahwa mayoritas mahasiswa angkatan 2019 di Fakultas X, Y dan Z lulus pada mata kuliah yang berhubungan dengan statistika dibanding tidak lulus.

Estimasi Parameter

Dengan bantuan software R Studio hasil estimasi parameter dalam model regresi logistik biner multilevel dengan *random intercept* dapat dilihat pada Tabel 1 dan untuk nilai *Random Effect* untuk setiap Fakultas dapat dilihat pada Tabel 2.

Tabel 1. Estimasi Parameter untuk Model Regresi Logistik Biner Multilevel dengan *random intercept*.

Variabel	Koefisien	Standard Error	Nilai Z	p-value	Keputusan
Konstanta ($\hat{\beta}_{00}$)	-6,8135	1,0863	-6,272	0,0000	Tolak H_0
Jenis Kelamin ($\hat{\beta}_{10}$)	0,4495	0,5080	0,885	0,3762	Terima H_0
Asal Daerah ($\hat{\beta}_{20}$)	1,3519	0,5829	2,319	0,0204	Tolak H_0
IPK ($\hat{\beta}_{30}$)	3,0676	0,3269	9,383	0,000	Tolak H_0
Rumpun Ilmu ($\hat{\beta}_{01}$)	1,0509	0,4591	2,289	0,0221	Tolak H_0

Sumber: Data Penelitian yang Sudah Diolah, 2023.

Tabel 2. *Random Effect* untuk Fakultas

Fakultas	Nilai
u_{01}	0,3167
u_{02}	-0,2131
u_{03}	-0,1036

Sumber: Data Penelitian yang Sudah Diolah, 2023.

Berdasarkan output yang tersaji pada Tabel 1 diperoleh nilai-nilai estimasi parameter untuk model regresi logistik biner multilevel. Dengan demikian model untuk regresi logistik biner multilevel dengan random intercept adalah sebagai berikut:

$$\log\left(\frac{\widehat{\pi_{ij}}}{1 - \pi_{ij}}\right) = -6,8135 + 0,4495 \text{ jenis kelamin}_{ij} + 1,3519 \text{ asal daerah}_{ij} \\ + 3,0676 \text{IPK}_{ij} + 1,0509 \text{rumpun_ilmu}_j + u_{0j}$$

dimana, berdasarkan Tabel 2 model regresi logistik biner multilevel dengan *random intercept* untuk setiap kelompok adalah sebagai berikut:

Kelompok (fakultas) 1:

$$\log\left(\frac{\widehat{\pi_{ij}}}{1 - \pi_{ij}}\right) = (-6,4968) + 0,4495 \text{ jenis kelamin}_{ij} + 1,3519 \text{ asal daerah}_{ij} \\ + 3,0676 \text{IPK}_{ij} + 1,0509 \text{rumpun_ilmu}_j$$

Kelompok (fakultas) 2:

$$\log\left(\frac{\widehat{\pi_{ij}}}{1 - \pi_{ij}}\right) = (-7,0266) + 0,4495 \text{ jenis kelamin}_{ij} + 1,3519 \text{ asal daerah}_{ij} \\ + 3,0676 \text{IPK}_{ij} + 1,0509 \text{rumpun_ilmu}_j$$

Kelompok (fakultas) 3:

$$\log\left(\frac{\widehat{\pi_{ij}}}{1 - \pi_{ij}}\right) = (-6,9171) + 0,4495 \text{ jenis kelamin}_{ij} + 1,3519 \text{ asal daerah}_{ij} \\ + 3,0676 \text{IPK}_{ij} + 1,0509 \text{rumpun_ilmu}_j$$

Hasil Pengujian Signifikansi *Random Effect*

Tabel 3. Uji Signifikansi *Random Effect*

Model	df	AIC	logLik	LR	p-value
Model logistik tanpa efek acak	1	396,5	-197,2527	4,3381	0,0373
Model logistik dengan efek acak	2	394,2	-195,0836		

Sumber: Data Penelitian yang Sudah Diolah, 2023.

$$LR = -22 \ln \left[\frac{L(\text{Likelihood model logistik tanpa efek acak})}{L(\text{Likelihood model logistik dengan efek acak})} \right] \\ = -2 \times [-197,2527 - (-195,0836)] \\ = 4,3381$$

Berdasarkan hasil dari pengolahan data R Studio menggunakan persamaan (8) yang tersaji pada Tabel 3, diperoleh nilai p-value 0,0373 dan nilai statistik uji LR sebesar 4,3381. Nilai p-value yang diperoleh lebih kecil dari $\alpha=5\%$ dan nilai statistik uji LR lebih besar dari $\chi^2_{0,05;1} = 3,84$ sehingga dapat disimpulkan bahwa H_0 ditolak pada taraf nyata 5%. Artinya dengan tingkat kepercayaan 95% terdapat *random effect* yang signifikan.

Hasil Pengujian Parameter Secara Simultan

Tabel 4. Uji Simultan

Model	df	logLik	Deviance	Chisq	p-value
Regresi logistik multilevel tanpa variabel prediktor	2	-195,084	390,17	209,75	0,000
Regresi logistik multilevel dengan variabel prediktor	6	-90,211	180,42		

Sumber: Data Penelitian yang Sudah Diolah, 2023.

$$G^2 = -2 \ln \left[\frac{L(\text{Likelihood tanpa variabel prediktor sedangkan})}{L(\text{Likelihood dengan variabel prediktor})} \right] \\ = -2 \times [-195,084 - (-90,211)] = 209,75$$

Berdasarkan hasil dari pengolahan data R Studio menggunakan persamaan (9) yang tersaji pada Tabel 4, diperoleh nilai p-value 0,000 dan nilai statistik uji G^2 sebesar 209,75. Nilai p-value yang diperoleh lebih kecil dari $\alpha=5\%$ dan nilai statistik uji G^2 lebih besar dari $\chi^2_{0,05;4} = 9,49$ sehingga dapat disimpulkan bahwa H_0 ditolak pada taraf nyata 5%. Artinya dengan tingkat kepercayaan 95% minimal ada salah satu dari variabel jenis kelamin (X1), asal daerah (X2), IPK (X3) dan rumpun ilmu (Z) yang berpengaruh terhadap status lulus mata kuliah statistika (Y).

Hasil Pengujian Parameter Secara Parsial

Berdasarkan hasil dari pengolahan data R Studio menggunakan persamaan (10) yang tersaji pada tabel 1 di atas, untuk parameter-parameter pada level 1 variabel asal daerah (X2) dan IPK (X3) berpengaruh secara signifikan terhadap status kelulusan pada mata kuliah statistika (Y) pada taraf nyata 5%, karena nilai p-value dari variabel asal daerah (X2) dan IPK (X3) secara berturut-turut sebesar 0,0204 dan 0,000 yang lebih kecil $\alpha=5\%$ dan variabel asal daerah (X2) dan IPK (X3) menghasilkan nilai statistik uji $|W|$ secara berturut-turut sebesar 2,319 dan 9,383 yang lebih besar $Z_{0,475} = 1,96$. Dan sisanya yaitu variabel jenis kelamin (X1) secara individual tidak berpengaruh signifikan terhadap status kelulusan pada mata kuliah statistika (Y) pada taraf nyata 5%.

Sedangkan untuk parameter pada level 2 menunjukkan bahwa secara individual rumpun ilmu (Z) berpengaruh secara signifikan terhadap status kelulusan pada mata kuliah statistika (Y) pada taraf nyata 5%, karena menghasilkan nilai p-value sebesar 0,0221 yang lebih kecil $\alpha=5\%$ dan menghasilkan nilai statistik uji $|W|$ sebesar 2,289 yang lebih besar dari $Z_{0,475} = 1,96$.

Interpretasi Parameter

Interpretasi perbedaan intersep antar kelompok (fakultas):

Kelompok 1 dan 2:

Variabel rumpun ilmu membantu menjelaskan perbedaan intersep antar kelompok. Didalam kelompok 1, untuk mahasiswa dengan jenis kelamin, asal daerah dan IPK yang sama kemungkinan seorang mahasiswa lulus adalah

$$\begin{aligned} & \frac{\exp[-6,8135 + 1,0509 \text{ rumpun_ilmu}_1 + 0,3167]}{\exp[-6,8135 + 1,0509 \text{ rumpun_ilmu}_2 - 0,2131]} \\ &= \exp[1,0509 (1 - 0) + (0,3167 - (-0,2131))] \\ &= \exp[1,0509 + 0,53] \\ &= 4,859 \text{ kali dibanding mahasiswa yang berasal dari kelompok 2.} \end{aligned}$$

Kelompok 1 dan 3:

Didalam kelompok 1, untuk mahasiswa dengan jenis kelamin, asal daerah dan IPK yang sama kemungkinan seorang mahasiswa lulus adalah

$$\begin{aligned} & \frac{\exp[-6,8135 + 1,0509 \text{ rumpun_ilmu}_1 + 0,3167]}{\exp[-6,8135 + 1,0509 \text{ rumpun_ilmu}_3 - 0,1036]} \\ &= \exp[1,0509 (1 - 0) + (0,3167 - (-0,1036))] \\ &= \exp[1,0509 + 0,42] \\ &= 4,358 \text{ kali dibanding mahasiswa yang berasal dari kelompok 3.} \end{aligned}$$

Jadi, perbedaan antar kelompok (fakultas) bisa sebagian dijelaskan oleh rumpun ilmu, sehingga mahasiswa yang berasal dari rumpun ilmu Saintek memiliki kecenderungan lulus lebih tinggi dibanding mahasiswa yang berasal dari Soshum.

Interpretasi variabel asal daerah (X2) dan IPK (X3) yang signifikan pada taraf nyata 5%. Pegang variabel prediktor lain konstan. Didalam sebuah kelompok, mahasiswa yang berasal dari pulau Jawa memiliki kecenderungan lulus pada mata kuliah statistika sebesar $\exp(1,3519) = 3,8650$ kali dibanding dengan mahasiswa yang berasal dari luar pulau Jawa. Didalam sebuah kelompok, setiap peningkatan 1 satuan IPK akan meningkatkan kecenderungan mahasiswa lulus pada mata kuliah statistika sebesar $\exp(3,0676) = 21,4902$ kalinya kecenderungan lulus dari mahasiswa yang berIPK satu satuan lebih rendah.

Intraclass Correlation Coefficient (ICC)

Nilai *interclass correlation* (ICC) digunakan untuk menggambarkan besarnya variasi status kelulusan pada mata kuliah statistika antar Fakultas di Fakultas x, Y dan Z. Hasil dari pengolahan data R menggunakan persamaan (13) diperoleh nilai ICC sebesar 0,0638 artinya sebesar 6,38% keragaman status kelulusan mata kuliah statistika di Fakultas X, Y dan Z dijelaskan oleh adanya perbedaan karakteristik antar fakultas, sedangkan sisanya yaitu sebesar 93,62% dijelaskan oleh keragaman antar mahasiswa.

D. Kesimpulan

Berdasarkan pembahasan dalam penelitian ini, peneliti menyimpulkan beberapa hasil penelitian sebagai berikut:

1. Berdasarkan hasil penerapan regresi logistik biner multilevel menyatakan bahwa faktor-faktor yang mempengaruhi status kelulusan mata kuliah yang berhubungan dengan statistika pada mahasiswa Universitas Islam Bandung Fakultas X, Y dan Z angkatan 2019 adalah variabel asal daerah (X2), IPK (X3) dan rumpun ilmu (Z).
2. Variabel pada level 1 yang mempengaruhi status kelulusan mata kuliah yang berhubungan dengan statistika pada mahasiswa Universitas Islam Bandung Fakultas X, Y dan Z angkatan 2019 adalah variabel asal daerah (X2) dan IPK (X3). Dimana mahasiswa yang berasal dari pulau Jawa cenderung untuk lulus lebih tinggi pada mata kuliah yang berhubungan dengan statistika dibanding mahasiswa yang berasal dari luar pulau Jawa. Serta, semakin tingginya IPK yang diperoleh mahasiswa maka kecenderungan mahasiswa untuk lulus pada mata kuliah yang berhubungan dengan statistika juga akan semakin tinggi.
3. Variabel pada level 2 yaitu rumpun (Z) secara signifikan mempengaruhi status kelulusan mata kuliah yang berhubungan dengan statistika pada mahasiswa Universitas Islam Bandung Fakultas X, Y dan Z angkatan 2019. Dimana mahasiswa yang berasal dari rumpun ilmu Saintek cenderung untuk lulus lebih tinggi pada mata kuliah yang berhubungan dengan statistika dibanding mahasiswa yang berasal dari rumpun ilmu Soshum.

Acknowledge

Alhamdulillah, all praise and gratitude to Allah SWT who has given health and strength to the author so that author can complete this research. Thank you to my beloved parents, mother and father, younger siblings and family who always pray for and provide both moral and material support to the author. Thank you to Dr. Nusar Hajarisman, M.Si. who has provided guidance to the author until this research is completed. Lecturers of the Statistics Study Program at the Bandung Islamic University who have provided a lot of knowledge. My friends and all those who have accompanied the author until this research is completed.

Daftar Pustaka

- [1] Anderson, C., J. (2020). *Multilevel Logistic Regression*. Urbana: University of Illinois.
- [2] Anwar, S., dkk. (2019). Laki-laki atau Perempuan, Siapa yang Lebih Cerdas dalam Proses Belajar? Sebuah Bukti dari Pendekatan Analisis Survival, 18(2), 281-296.
- [3] Agresti, Alan. (2007). *An Introduction to Categorical Data Analysis Second Edition*. New Jersey: John Wiley & Sons.
- [4] Dewi, W. A. (2020). Dampak Covid-19 terhadap Implementasi Pembelajaran Daring di Sekolah Dasar. Edukatif: Jurnal Ilmu Pendidikan, 55-61.
- [5] Hajarisman, N. (2009). *Analisis Data Kategorik*. Bandung: Universitas Islam Bandung.
- [6] Hasanah, H. dkk. (2018). Analisis Hasil Belajar Dengan Nilai Indeks Prestasi Kumulatif (IPK) Mahasiswa STAI At-taqwa Bondowoso, 5(2), 29-39.
- [7] Hox, J. (2002). *Multilevel analysis techniques and applications*. Lawrence Erlbaum Associates Publishers.
- [8] Indriyani, R. (2018). Pengaruh Asal Sekolah dan Tempat Tinggal terhadap Prestasi Belajar

- Mahasiswa, 34-39.
- [9] Poedjiati, A. S, Susanti L., Dwiatmono. (2008). Perbandingan Regresi Logistik dan Model Multilevel pada Data Ordinal dan Biner untuk Mengetahui Variabel yang Mempengaruhi Hasil Belajar Mengajar Kalkulus I di ITS. Magister Statistik. Institut Teknologi Sepuluh November: Surabaya.
 - [10] Sudjana, Nana. (2010). *Dasar-Dasar Proses Belajar*, Bandung: Sinar Baru.
 - [11] W. Fadila and M. Herlina, “Penerapan Metode Generalized Structure Component Analysis pada Pengguna Dompot Digital Menggunakan Model UTAUT 2,” *Jurnal Riset Statistika*, pp. 27–34, Jul. 2023, doi: 10.29313/jrs.v3i1.1772.