

Model Regresi Dummy untuk Indeks Prestasi Kumulatif Mahasiswa

Dita Monita

Universitas Palangka Raya, ditamonita@mipa.upr.ac.id

ABSTRAK, Regresi yang melibatkan variabel dummy (bernilai 0 atau 1) dapat digunakan untuk menentukan pengaruh dari variabel dummy terhadap suatu variabel kuantitatif. Penelitian ini bertujuan untuk menentukan model regresi yang menyatakan hubungan antara IPK mahasiswa dengan nilai UN (variabel kuantitatif) dan empat variabel dummy yaitu status SMA/SMK (SMA), tingkat akreditasi SMA/SMK (AK), asal kabupaten dari SMA/SMK (KAB), dan jalur masuk perguruan tinggi (PT). Populasinya adalah semua mahasiswa pendidikan matematika angkatan (yang masuk) di tahun akademik 2019/2020 dari salah satu universitas di Palangka Raya, Kalimantan Tengah. Penelitian ini dilaksanakan pada populasi. Banyak anggota populasi adalah 59 mahasiswa. Peneliti mengumpulkan data menggunakan survey online (google form). Analisis data dilakukan secara deskriptif (tabel dan grafik), dan inferensia (analisis regresi dummy). Hasilnya menunjukkan bahwa UN, SMA, dan AK secara signifikan berpengaruh terhadap IPK, sedangkan dua variabel dummy lainnya tidak signifikan. Model regresi terbaiknya adalah $IPK = 2,237 + 0,012 \cdot UN + 0,323 \cdot SMA + 0,2 \cdot AK$.

Kata Kunci: akreditasi, indeks prestasi kumulatif, nilai ujian nasional, regresi dummy, status sekolah

1. PENDAHULUAN

Secara teoritis, Indeks Prestasi Kumulatif (IPK) yang merepresentasikan hasil belajar mahasiswa dipengaruhi kemampuan awal dari mahasiswa [1]. Pada mahasiswa tingkat 2, kemampuan awalnya adalah nilai Ujian Nasional (UN) dan hasil belajar saat mahasiswa tersebut berada di jenjang SMA/SMK. Hasil belajar siswa itu sendiri dipengaruhi oleh kondisi sekolah dan lingkungan sosial [2]. Kualitas dari kedua faktor tersebut terakut dengan status (negeri atau swasta) dan tingkat akreditasi dari sekolah. Pada konteks pendidikan di Kalimantan Tengah, kualitas dari kedua faktor tersebut juga dipengaruhi oleh asal kabupaten dari lokasi sekolah. Sekolah yang berlokasi di kabupaten lama memiliki kondisi sekolah dan lingkungan sosial yang lebih baik dari sekolah pada kabupaten pemekaran secara historis. Selain itu,

IPK juga dipengaruhi oleh jalur masuk Perguruan Tinggi (PT) [3].

Penelitian ini bertujuan untuk menentukan model regresi yang menyatakan hubungan antara IPK mahasiswa Pendidikan Matematika dari salah satu Universitas di Palangka Raya, Kalimantan Tengah angkatan 2019 dengan nilai UN, status sekolah SMA/SMK, tingkat akreditasi SMA/SMK, asal kabupaten dari SMA/SMK, dan jalur masuk PT dengan menggunakan regresi dummy.

2. TINJAUAN PUSTAKA

Analisis Regresi

Analisis regresi adalah suatu teknik Statistika yang ditujukan untuk menyelidiki dan memodelkan hubungan antara variabel terikat dan setidaknya satu variabel bebas [4]. Kegunaan analisis ini adalah mendeskripsikan data, menduga parameter, interpolasi, dan kontrol [5].

Regresi secara umum dibedakan atas regresi linear dan regresi nonlinear. Regresi linear apabila hubungan antara variabel terikat dan variabel-variabel bebasnya bersifat linear, Pada regresi nonlinear, hubungan tersebut bersifat nonlinear. Contoh model regresi linear dengan variabel terikat y dan dua variabel bebas, x_1 dan x_2 , adalah

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon \quad (2.1)$$

dengan β_0 , β_1 dan β_2 adalah koefisien-koefisien regresi dan ε adalah residual. Persamaan tersebut disebut dengan model regresi linear berganda (*multiple regression model*) dengan dua variabel bebas. Contoh model regresi nonlinear dengan variabel-variabel di atas adalah

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \varepsilon \quad (2.2)$$

Persamaan yang terakhir ini disebut dengan model regresi polinomial.

Secara umum, model regresi linear berganda dengan k variabel bebas adalah

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (2.3)$$

Jika banyak data adalah n , maka model regersinya adalah

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_k x_{1k} + \varepsilon_1 \\ y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_k x_{2k} + \varepsilon_2 \\ y_3 &= \beta_0 + \beta_1 x_{31} + \beta_2 x_{32} + \dots + \beta_k x_{3k} + \varepsilon_3 \\ &\vdots \\ y_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_k x_{nk} + \varepsilon_n \end{aligned}$$

Sistem persamaan tersebut dapat ditulis

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i \tag{2.4}$$

dengan $i = 1, 2, 3, \dots, n$.

Persamaan (4) dapat ditulis dalam bentuk matriks

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ 1 & x_{31} & x_{32} & \dots & x_{3k} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{2.5}$$

atau

$$y = X\beta + \varepsilon \tag{2.6}$$

Parameter persamaan regresi β diduga menggunakan metode kuadrat terkecil (*least-squares method*). Hasilnya adalah persamaan normal kuadrat terkecil (*least-squares normal equation*) berikut.

$$X'X\hat{\beta} = X'y \tag{2.7}$$

dengan $\hat{\beta}$ adalah dugaan dari β . Selesaian persamaan (7) diperoleh dengan mengalikan sebelah kiri di kedua ruas dengan $(X'X)^{-1}$ diperoleh persamaan:

$$\hat{\beta} = (X'X)^{-1} X'y \tag{2.8}$$

Variabel bebas X dalam suatu model regresi dapat berupa variabel kuantitatif atau variabel kualitatif (kategori) [4]. Variabel kuantitatif apabila skala datanya adalah rasio atau interval. Skala rasio melibatkan data yang dapat diurutkan dimana perbedaan antardata bermakna, dan memiliki nol alamiah, contohnya adalah tinggi atau berat badan. Skala interval melibatkan data yang dapat diurutkan dimana perbedaan antardata bermakna, tetapi tidak mempunyai nol alamiah, contohnya adalah nilai UN (ujian nasional) [6].

Variabel kualitatif apabila skala datanya adalah ordinal atau nominal. Skala ordinal melibatkan data dalam bentuk kategori yang dapat diurutkan, tetapi perbedaan antar data

tanpa makna, contohnya adalah akreditasi SMA/SMK. Skala nominal dicirikan oleh data yang berupa nama atau kategori, tetapi data tersebut tidak dapat diurutkan, contohnya adalah status SMA/SMK, asal kabupaten dari SMA/SMK, atau jalur masuk PT (perguruan tinggi). Variabel kualitatif tersebut disebut juga dengan variabel indikator atau dummy [4].

Regresi Dummy

Regresi yang melibatkan variabel dummy disebut regresi dummy. Regresi digunakan untuk melihat hubungan antara variabel terikat (variabel kuantitatif) dengan variabel dummy tersebut. Sebagai contoh, variabel terikat adalah indeks prestasi kumulatif dari mahasiswa (*IPK*). Variabel bebasnya adalah nilai UN SMA/SMK dari mahasiswa tersebut disimbolkan *UN* (variabel kuantitatif), dan akreditasi SMA/SMK disimbolkan *AK* (variabel dummy). Model regresi untuk variabel-variabel tersebut adalah

$$IPK = \beta_0 + \beta_1 UN + \beta_2 AK + \varepsilon \tag{2.9}$$

Jika hubungan antara *IPK* dan *AK* yang dinyatakan dengan koefisien regresi β_2 tidak signifikan dengan tingkat kesalahan jenis I sebesar α , maka model regresi terbaik yang menyatakan hubungan tersebut dapat ditulis

$$IPK = \beta_0 + \beta_1 UN + \varepsilon \tag{2.10}$$

Sebaliknya, apabila hubungan tersebut signifikan, maka model regresi untuk SMA/SMK dengan akreditasi C atau belum akreditasi ($AK = 0$) dinyatakan oleh persamaan (11), sedangkan untuk akreditasi A atau B ($AK = 1$) dinyatakan oleh persamaan (12).

$$IPK = \beta_0 + \beta_1 UN + \beta_2(0) + \varepsilon = \beta_0 + \beta_1 UN + \varepsilon \tag{2.11}$$

$$IPK = \beta_0 + \beta_1 UN + \beta_2(1) + \varepsilon = (\beta_0 + \beta_2) + \beta_1 UN + \varepsilon \tag{2.12}$$

Kedua model regresi tersebut memiliki kemiringan (*slope*) yang sama yaitu β_1 , tetapi interceptnya berbeda yaitu β_0 pada saat $AK = 0$, dan $\beta_0 + \beta_2$ pada saat $AK = 1$. Jika $\beta_2 > 0$, maka garis regresi dari $AK = 1$ di atas garis regresi $AK = 0$ dengan kemiringan yang sama. Perbedaan ketinggian tersebut menunjukkan *IPK* mahasiswa yang berasal dari sekolah dengan akreditasi A atau B lebih dari akreditas C atau belum akreditasi. Dengan kata lain, *IPK*

mahasiswa dipengaruhi oleh akreditasi SMA/SMK.

3. METODOLOGI

Penelitian ini adalah penelitian statistika terapan. Populasi dalam penelitian ini adalah semua mahasiswa pendidikan matematika angkatan (yang masuk) di tahun akademik 2019/2020 dari salah satu universitas di Palangka Raya, Kalimantan Tengah. Penelitian ini dilaksanakan pada populasi. Banyak anggota populasi adalah 59 mahasiswa yang terbagi dalam dua kelas yaitu A dan B. Mahasiswa dengan nomor induk ganjil di kelas A, dan yang genap di kelas B.

Data yang dikumpulkan adalah data primer yang terdiri dari IPK, nilai UN Matematika jenjang SMA, akreditasi SMA, status sekolah (negeri/swasta), kabupaten dimana siswa menempuh jenjang SMA, dan jalur masuk PT. Data tersebut diperoleh melalui survey online dengan google form. Link google form: <https://forms.gle/i8ETn2hs69Lb5ZwL9>.

Analisis data dilakukan dalam empat tahap. Pertama, peneliti mengeksplorasi data menggunakan tabel dan grafik. Kedua, peneliti menguji empat asumsi analisis regresi. Asumsi kelinearan diperiksa menggunakan uji *lack-of-fit*. Asumsi kenormalan residual diperiksa menggunakan uji Kolmogorov-Smirnov. Asumsi kebebasan residual diperiksa menggunakan plot antara no urut residual dengan residual itu sendiri. Asumsi homoskedastisitas diperiksa menggunakan plot antara dugaan (*fit*) terhadap residual. Keempat, peneliti melakukan analisis regresi dummy, dan membandingkan model-model regresi untuk menentukan model terbaik untuk variabel terikat IPK [4]. Semua analisis tersebut dilakukan menggunakan Minitab 18.

Variabel terikat dalam penelitian ini berdasarkan tujuannya adalah Indeks Prestasi Kumulatif untuk 3 semester disimbolkan dengan IPK. Variabel bebasnya adalah nilai UN SMA/SMK pada mata pelajaran Matematika, tingkat akreditasi, status sekolah (negeri/swasta), asal kabupaten dari lokasi SMA/SMK, dan jalur masuk PT yang disimbolkan secara berturut-turut dengan UN, AK, SMA, KAB, dan PT. UN adalah variabel bebas kuantitatif, sedangkan

variabel bebas lainnya adalah variabel kualitatif atau dummy.

Nilai untuk setiap variabel bebas dummy adalah 0 atau 1. Nilai AK = 0 jika akreditasi SMA/SMK adalah C atau belum terakreditasi, dan AK = 1 jika akreditasinya A atau B. Nilai SMA = 0 jika SMA/SMK adalah sekolah swasta, dan SMA = 1 jika negeri. Nilai KAB = 0 jika asal SMA/SMK dari kabupaten pemekaran di Kalimantan Tengah yaitu Pulang Pisau, Katingan, Sukamara, Lamandau, Seruyan, Barito Timur, Murung Raya, dan Gunung Mas, dan KAB = 1 jika asal SMA/SMK adalah kabupaten/kota lama di Kalimantan Tengah, atau salah satu kabupaten di luar Kalimantan Tengah. Kabupaten lama tersebut adalah Palangka Raya, Kapuas, Kotawaringin Timur, Kotawaringin Barat, Barito Utara dan Barito Selatan. Nilai PT = 0 jika jalur masuk mahasiswa di perguruan tinggi adalah SMMPTN Barat atau Jalur Mandiri, dan PT = 1 jika jalur masuknya adalah SNMPTN atau SBMPTN.

Model regresi dummy dari variabel-variabel tersebut adalah

$$IPK = \beta_0 + \beta_1 UN + \beta_2 SMA + \beta_3 AK + \beta_4 PT + \beta_5 KAB + \varepsilon \quad (2.11)$$

Parameter-parameter $\beta_0, \beta_1, \dots, \beta_5$ diduga menggunakan persamaan (8). Hipotesisnya adalah

$$H_0: \beta_1 = \dots = \beta_5 = 0$$

$$H_1: \text{setidaknya ada satu } \beta_j \neq 0 \text{ dengan } j = 1, 2, \dots, 5$$

Statistik hitung untuk hipotesis tersebut adalah [5]:

$$F_{\text{hitung}} = \frac{KTR}{KTG} = \frac{JKR/k}{JKG/(n-k-1)} \quad (2.12)$$

dengan

KTR = kuadrat tengah regresi

KTG = kuadrat tengah galat

JKR = Jumlah kuadrat regresi = $\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$

JKG = Jumlah kuadrat galat = $\sum_{i=1}^n (y_i - \hat{y}_i)^2$

k = banyak variabel bebas

n = banyak data

\hat{y}_i adalah entri baris ke- i dari $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} + \boldsymbol{\varepsilon}$

Hasil $JKR + JKG$ dinotasikan dengan JKT (jumlah kuadrat total). Kriteria ujinya adalah tolak H_0 jika $F_{\text{hitung}} > F_{\text{tabel}}(k, n - k - 1)$. Jika hasil uji adalah tolak H_0 , maka peneliti menyelidiki lebih lanjut mana dari β_j untuk suatu

$j = 1, 2, \dots, 5$ yang tidak nol. Hipotesisnya adalah

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0$$

Statistik hitungnya adalah [4]

$$t_{hitung} = \frac{\hat{\beta}_j}{\sqrt{KTG \cdot C_{jj}}} \quad (2.13)$$

dengan C_{jj} adalah unsur diagonal dari matriks $(X'X)^{-1}$. Kriteria ujinya adalah tolak H_0 jika $t_{hitung} > t_{tabel}(\alpha/2, n - 6)$ dengan α adalah kesalahan pengambilan kesimpulan jenis I [4].

Salah satu indikator kebaikan dari suatu model regresi adalah koefisien determinasi yang dinotasikan dengan R^2 . Koefisien tersebut menyatakan seberapa besar model regresi dapat menjelaskan keragaman dari variabel terikat [6]. Rumus hitungnya

$$R^2 = \frac{JKR}{JKT} \quad (2.14)$$

Bila ada satu parameter β_j yang tidak signifikan, peneliti mempertimbangkan untuk mengeliminasi variabel bebas yang berkaitan dengan β_j . Statistik yang digunakan untuk membandingkan model regresi sebelum dan sesudah pengeliminasian tersebut adalah koefisien determinasi terkoreksi yang disimbolkan dengan R^2_{Adj} . Rumus hitungnya adalah

$$R^2_{Adj} = 1 - \frac{KTR}{JKT/(n-1)} \quad (2.15)$$

Setiap uji dalam Statistika memiliki asumsi, begitu pula analisis regresi termasuk regresi dummy. Asumsi pertama adalah asumsi kelinearan hubungan antara peubah terikat dan variabel-variabel bebas. Asumsi kedua adalah residual menyebar normal. Asumsi ketiga adalah kebebasan (independensi) dari residual. Asumsi keempat adalah homoskedastisitas (kesamaan varians). Keempat asumsi tersebut diperiksa peneliti menggunakan Minitab 18.

4. PEMBAHASAN

Hasil Penelitian

Rangkuman IPK untuk setiap peubah dummy menunjukkan bahwa IPK mahasiswa yang memiliki nilai 1 pada variabel dummy status SMA/SMK lebih dari yang memiliki nilai 0. Hal yang sama untuk variabel dummy

akreditasi, jalur masuk dan kabupaten (Tabel 4.1). Kondisi tersebut sesuai dengan pengkodean 1 dan 0 dalam regresi [5]. Begitu pula, nilai maksimum dari IPK terjadi pada saat nilai = 1 untuk setiap variabel dummy. Walaupun demikian, ada perbedaan dalam nilai minimum (IPK = 1,79). Pada variabel status dan akreditasi SMA/SMK, nilai minimum tersebut terjadi pada saat nilai variabel dummy = 0. Pada variabel jalur masuk dan kabupaten, nilai minimum tersebut terjadi pada saat nilai variabel dummy = 1.

Hasil analisis regresi dummy menunjukkan bahwa variabel kuantitatif nilai UN berpengaruh secara signifikan ($p\text{-value} = 0 < 0,1$) terhadap IPK dengan tingkat kepercayaan 90%. Begitu pula dengan variabel dummy status SMA/SMK ($p\text{-value} = 0,01 < 0,1$), dan Akreditasi ($p\text{-value} = 0,04 < 0,1$). Akan tetapi, dua peubah dummy lainnya yaitu Jalur Masuk ($p\text{-value} = 0,526 \geq 0,1$), dan Kabupaten tidak signifikan ($p\text{-value} = 0,382 \geq 0,1$) dengan tingkat kepercayaan 90%. Koefisien determinasi untuk model regresi tersebut, $R^2 = 44,95\%$ artinya model tersebut dapat menjelaskan keragaman dari variabel terikat IPK sebesar 44,95%.

Lebih lanjut, koefisien determinasi terkoreksinya, $R^2_{Adj} = 39,76\%$. Hasil pendugaan terhadap parameter model regresi adalah $\hat{\beta}_0 = 2,259$, $\hat{\beta}_1 = 0,014$, $\hat{\beta}_2 = 0,321$, $\hat{\beta}_3 = 0,225$, $\hat{\beta}_4 = -0,056$, dan $\hat{\beta}_5 = -0,075$ sehingga persamaan regresinya adalah

$$IPK = 2,259 + 0,014 \cdot UN + 0,321 \cdot SMA + 0,225 \cdot AK - 0,056 \cdot PT - 0,075 \cdot KAB \quad (2.16)$$

Peneliti mempertimbangkan untuk mengeliminasi variabel dummy yang tidak signifikan dari model regresi pertama. Hasil analisis regresi kedua dengan hanya 2 peubah dummy yaitu Status SMA/SMK dan Akreditasi menunjukkan bahwa $R^2_{Adj} = 40,48\%$. Dengan demikian, koefisien determinasi dengan hanya 2 peubah dummy lebih dari 4 peubah dummy, tetapi variabel bebas yang dilibatkan dalam model regresi lebih sedikit sehingga peneliti memilih menggunakan model regresi yang melibatkan 1 peubah kontinu (UN) dan 2 peubah dummy (Status SMA/SMK dan Akreditasi). Hasil analisis ragam untuk ketiga peubah bebas berpengaruh terhadap IPK dengan tingkat

kepercayaan 90% dimana $p\text{-value} = 0 < 0,1$
(Tabel 4.3).

Tabel 4.1. Rangkuman IPK untuk Setiap Variabel Dummy

Variabel Dummy	Nilai	N	Mean	StDev	Min	Q1	Median	Q3	Maks
Status SMA/SMK (SMA)	0 (Swasta)	6	2,85	0,58	1,79	2,52	2,95	3,25	3,46
	1 (Negeri)	53	3,30	0,26	2,64	3,12	3,31	3,52	3,72
Akreditasi (AK)	0 (C/Belum Akreditasi)	8	2,93	0,53	1,79	2,86	2,94	3,32	3,54
	1 (A/B)	51	3,31	0,26	2,64	3,12	3,31	3,52	3,72
Jalur Masuk (PT)	0 (SMMPTN Barat/Mandiri)	13	3,13	0,28	2,76	2,90	3,05	3,42	3,57
	1 (SNMPTN/SBMPTN)	46	3,29	0,33	1,79	3,15	3,31	3,52	3,72
Kabupaten (KAB)	0 (Pemekaran)	13	3,22	0,26	2,64	3,10	3,24	3,41	3,57
	1 (Lama)	46	3,27	0,35	1,79	3,05	3,31	3,52	3,72

Tabel 4.2. Analisis Ragam dari Regresi IPK terhadap UN, Status SMA/SMK, Akreditasi, Jalur Masuk dan Kabupaten

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	5	2,77937	0,55587	8,66	0,000
UN	1	1,23418	1,23418	19,22	0,000
Status SMA/SMK (SMA)	1	0,46410	0,46410	7,23	0,010
Akreditasi (AK)	1	0,28532	0,28532	4,44	0,040
Jalur masuk (PT)	1	0,02613	0,02613	0,41	0,526
Kabupaten (KAB)	1	0,04987	0,04987	0,78	0,382
Error	53	3,40329	0,06421		
Lack-of-Fit	32	2,22641	0,06958	1,24	0,306
Pure Error	21	1,17688	0,05604		
Total	58	6,18265			

Tabel 4.3. Analisis Ragam dari Regresi IPK terhadap UN, Status SMA/SMK, dan Akreditasi

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	3	2,6932	0,89772	14,15	0,000
UN	1	1,1694	1,16940	18,43	0,000
Status SMA/SMK (SMA)	1	0,4898	0,48975	7,72	0,007
Akreditasi (AK)	1	0,2362	0,23615	3,72	0,059
Error	55	3,4895	0,06345		
Lack-of-Fit	23	1,8426	0,08011	1,56	0,122
Pure Error	32	1,6469	0,05147		
Total	58	6,1827			

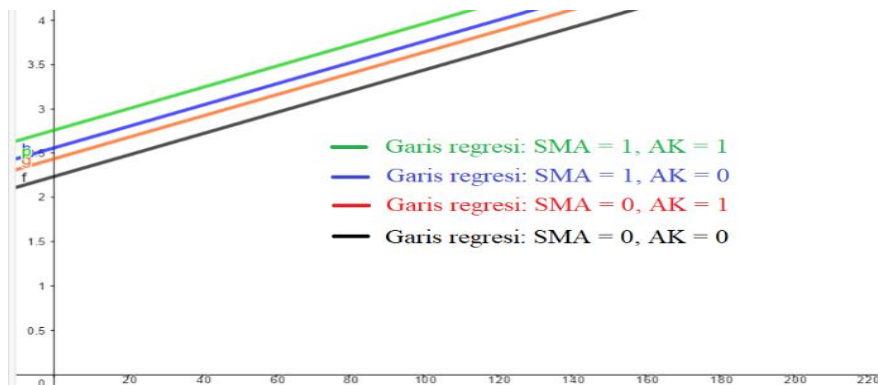
Hasil pendugaan terhadap parameter model regresi adalah $\widehat{\beta}_0 = 2,237$, $\widehat{\beta}_1 = 0,012$, $\widehat{\beta}_2 = 0,323$, dan $\widehat{\beta}_3 = 0,2$ sehingga persamaan regresinya adalah

$$IPK = 2,237 + 0,012 \cdot UN + 0,323 \cdot SMA + 0,2 \cdot AK \quad (2.17)$$

Lebih lanjut, model regresi untuk setiap pasang kemungkinan nilai dari peubah dummy SMA dan AK memiliki kemiringan yang sama sebesar 0,012 (Tabel 4.4). Garis yang paling tinggi yang ditunjukkan oleh intercept terbesar pada saat nilai variabel dummy SMA = 1 dan AK = 1 (Gambar 4.1).

Tabel 4.4. Model Regresi untuk Setiap Kemungkinan Nilai Peubah Dummy Status SMA/SMK (SMA) dan Akreditasi (AK)

SMA	AK	Regression Equation
0	0	$IPK = 2,237 + 0,012 UN + 0,323 (0) + 0,2 (0) = 2,237 + 0,012 UN$
0	1	$IPK = 2,237 + 0,012 UN + 0,323 (0) + 0,2 (1) = 2,437 + 0,012 UN$
1	0	$IPK = 2,237 + 0,012 UN + 0,323 (1) + 0,2 (0) = 2,560 + 0,012 UN$
1	1	$IPK = 2,237 + 0,012 UN + 0,323 (1) + 0,2 (1) = 2,760 + 0,012 UN$

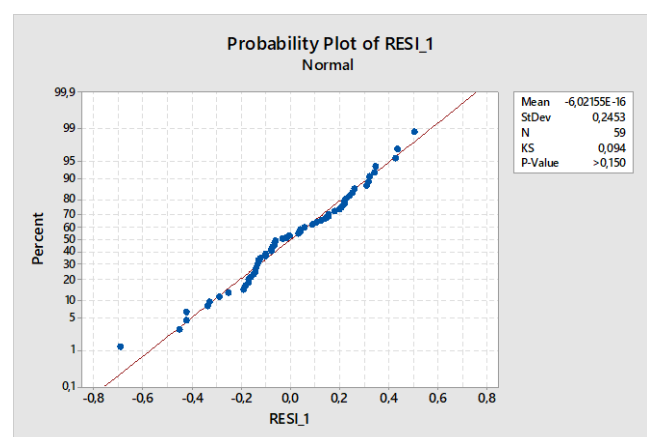


Gambar 4.1. Garis Regresi untuk Semua Kemungkinan Nilai Variabel Dummy

Ada dua data yang diidentifikasi pengamatan yang tidak biasa (*unusual observations*) karena memiliki residual yang besar yaitu mahasiswa no 16 dan 29. Peneliti mempertimbangkan untuk mengeliminasi kedua data tersebut. Hasilnya adalah model regresi dengan $R^2 = 32,99\%$ kurang dari koefisien determinasi sebelum pengeliminasian ($R^2 = 43,56\%$). Jadi, peneliti memilih menggunakan model regresi yang menggunakan seluruh data populasi (sebelum pengeliminasian).

Peneliti memeriksa asumsi analisis regresi untuk menentukan kevalidan dari model regresi yang digunakan. Asumsi pertama diperiksa menggunakan uji *lack-of-fit*.

Hasilnya adalah $p\text{-value} = 0,122 > 0,1$ yang berarti asumsi kelinearan terpenuhi (Tabel 4.3),

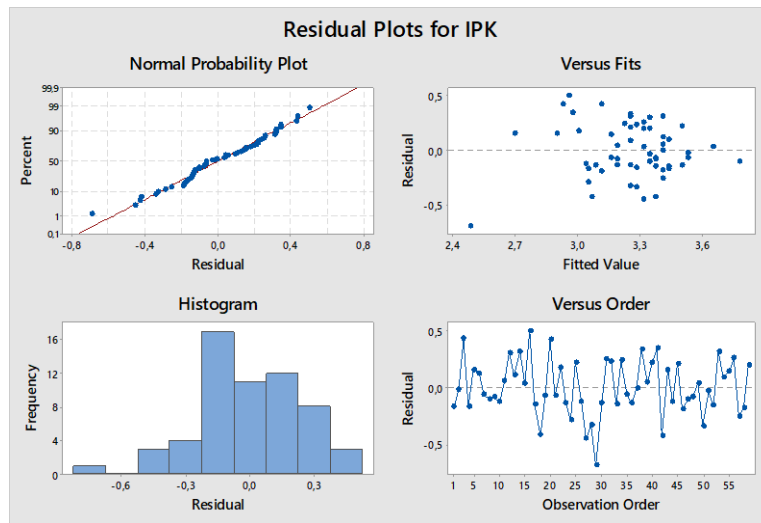


Gambar 4.2. Uji Kenormalan Kolmogorov-Smirnov terhadap Residual

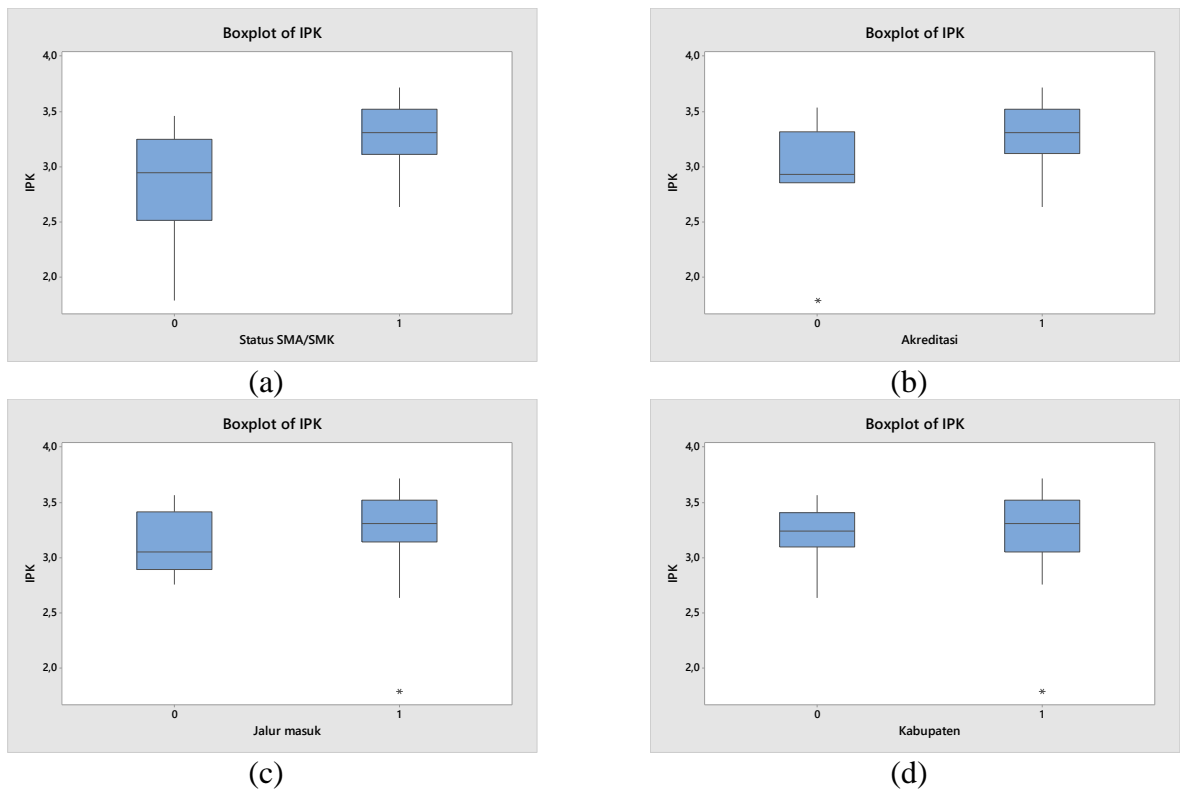
Asumsi kedua diperiksa menggunakan uji Kolmogorov Smirnov terhadap residual. Hasilnya diperoleh $p\text{-value} > 0,15 \geq 0,1$ (Gambar

4.2) yang berarti asumsi kenormalan dari residual terpenuhi. Asumsi ketiga (kebebasan residual) terpenuhi dari keacakan titik-titik pada plot antara Residual vs Observation Order. Terakhir,

asumsi keempat (homoskedastisitas) terpenuhi dari keacakan titik-titik pada plot antara Residual vs Fitted Value (Gambar 4.3). Dengan demikian, keempat asumsi dari analisis regresi terpenuhi.



Gambar 4.3. Plot antara Residual terhadap Dugaan dari Urutan Residual



Gambar 4.4. Boxplot IPK untuk Setiap Nilai dari Keempat Peubah Dummy

Hasil penelitian menunjukkan bahwa variabel kuantitatif UN, variabel dummy status SMA (Negeri/Swasta), dan variabel dummy AK (akreditasi sekolah) berpengaruh signifikan terhadap IPK mahasiswa. Signifikansi dari ketiga

Pembahasan

variabel tersebut sejalan dengan hasil penelitian Amalita dan Kurniawati [7]. Akan tetapi, variabel dummy PT (jalur masuk perguruan tinggi) dan KAB (kabupaten lama/baru) tidak berpengaruh signifikan. Pengaruh dari variabel SMA dan AK juga dapat dilihat dari posisi dari boxplot dengan nilai = 1 lebih tinggi dari yang nilai = 0 (Gambar 4.4 (a) dan 4.4 (b)).

Pada variabel PT, mahasiswa yang masuk di jalur pertama (SNMPTN) dan kedua (SBMPTN) (nilai PT = 1) diharapkan memiliki IPK yang lebih dari yang masuk di jalur ketiga SMMPTN wilayah barat/Mandiri (nilai PT = 0). Akan tetapi, rata-rata IPK yang relatif sama di PT = 1 sebesar 3,29; dan PT = 0 sebesar 3,13 (Tabel 4.1), dan boxplot PT = 1 dan PT = 0 yang tingginya relatif sama (Gambar 4.4 (c)) menunjukkan variabel PT tidak signifikan terhadap IPK. Lebih lanjut, adanya outlier (mahasiswa ke-29) yang nilainya jauh lebih rendah dari kumpulannya pada boxplot IPK di PT = 1 menyebabkan kondisi tersebut, dan mendorong koefisien regresi untuk variabel ini bernilai negatif ($\widehat{\beta}_4 = -0,056$).

Begitu pula pada variabel KAB, mahasiswa dari kabupaten lama (nilai KAB = 1) diharapkan memiliki fasilitas belajar yang lebih baik dari kabupaten baru/pemekaran (nilai KAB = 0). Akan tetapi, rata-rata IPK untuk KAB = 1 sebesar 3,27 relatif sama dengan KAB = 0 dengan rata-rata IPK = 3,22. Adanya outlier pada boxplot IPK di KAB = 1 juga mendorong koefisien regresi untuk variabel ini bernilai negatif ($\widehat{\beta}_5 = -0,075$).

5. KESIMPULAN

Hasil penelitian ini menunjukkan bahwa variabel kuantitatif UN (nilai UN SMA) berpengaruh signifikan terhadap IPK. Pada variabel dummy, status SMA (negeri/swasta) dan AK (akreditasi sekolah) berpengaruh signifikan, sedangkan PT (jalur masuk PT) dan KAB (kabupaten lama/pemekaran) tidak berpengaruh signifikan terhadap IPK. Regresi terbaik untuk IPK diperoleh dengan mengeliminasi variabel-variabel yang tidak signifikan yaitu PT dan KAB. Model regresinya adalah:

$$IPK = 2,237 + 0,012.UN + 0,323.SMA + 0,2.AK.$$

Pengeliminasian kedua variabel tersebut meningkatkan R_{Adj}^2 dari 39,76% (sebelum pengeliminasian) menjadi 40,48%. (sesudah pengeliminasian).

Hasil penelitian ini diharapkan dapat dimanfaatkan pihak kampus sebagai dasar perekrutan mahasiswa baru dan pihak pemegang kebijakan pada pemerintah provinsi maupun kabupaten/kota dalam rangka meningkatkan kualitas sumber daya manusia di daerahnya masing-masing.

6. DAFTAR PUSTAKA

- [1] Lestari, W, 2017, “Pengaruh Kemampuan Awal Matematika dan Motivasi Belajar terhadap Hasil Belajar Matematika”, *Jurnal Analisa*, 3 (1), 76-84.
- [2] Sawawa, D., Solehudin, A., & Sabri, 2018, “Pengaruh Faktor Internal dan Eksternal Siswa Terhadap Hasil Belajar pada Mata Pelajaran Mekanika Teknik dan Elemen Mesin”, *Journal of Mechanical Engineering Education*, 5 (1), 21-26.
- [3] Amalita, N., & Kurniati, Y, 2013, “Model Regresi Dummy dalam Memprediksi Performansi Akademik Mahasiswa Jurusan Matematika FMIPA UNP”, *Semirata FMIPA BKS Barat* (pp. 387-391), Lampung: Unila.
- [4] Montgomery, D. C., Peck, E. A., & Vining, G. G, 2012, “*Introduction to Linear Regression Analysis (5th Edition ed.)*”, Hoboken, New Jersey: John Wiley & Sons, Inc.
- [5] Chatterjee, S., & Simonoff, J. S, 2013, “*Handbook of Regression Analysis*”, Hoboken, New Jersey: John Wiley & Sons, Inc.
- [6] Mairing, J. P. (2016). *Statistika Pendidikan: Konsep dan Aplikasi Menggunakan Minitab dan Excel*. Yogyakarta: Andi Offset.
- [7] Amalita, N., & Kurniawati, Y, 2013, “Model Regresi Dummy Indeks Prestasi Akademik Mahasiswa Angkatan 2010 Jurusan Matematika FMIPA UNP”, *Eksakta*, 2, 17-23.