

Received: 25 November 2023

Revised: 27 December 2023

Accepted: 30 December 2023

Published: 31 December 2023

Log Linier Binomial Negatif dalam Memodelkan Data Siswa Putus Sekolah Tingkat SMA/SMK

Vera Maya Santi^{1, a)}, Fanya Izmi Hawa^{1, b)}, Bagus Sumargo^{1, c)}

¹*Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta, Jl. Rawamangun muka, Kota Jakarta Timur, DKI Jakarta, 13220..*

Email: ^{a)} ymsanti@unj.ac.id, ^{b)} fnyaizmi@gmail.com, ^{c)} bagussumargo@unj.ac.id

Abstract

Log linear negative binomial model is an extension of the linear regression model that can be used to analyze count data in the form of a contingency table when dealing with overdispersion where the variance value is greater than the mean value. Research in the field of education often produces contingency table data, one of which is data on students dropping out of school. The 2018 BPS survey showed that the number of students dropping out of school at the secondary school level in West Java province had the highest number of dropouts. Unfortunately, quantitative research on dropout students is still rarely done. Log linear negative binomial model is applied to determine the factors that influence the number of SMA/SMK dropout students in West Java province in 2021. The results show that gender, type of middle school, school status, interaction of gender and school status, and the interaction of middle school type and school status significantly affects the number of students dropping out of school. Furthermore, SMA/SMK dropouts in the province of West Java are dominated by male students, vocational students, and public-school students. Females are 1% more risky to drop out than males in public schools, while males are 1% more risky to drop out than females in private schools. SMA students are 59% more risky to drop out than SMK students in public schools, while SMK students are 71% more risky to drop out than SMA students in private school.

Keywords: contingency table, dropouts, log linear negative binomial, overdispersion.

Abstrak

Model log linier binomial negatif adalah perluasan model regresi linier yang dapat digunakan untuk menganalisis data cacahan berbentuk tabel kontingensi saat berhadapan dengan overdispersi dimana nilai varians lebih besar dari nilai rata-rata. Penelitian di bidang pendidikan seringkali menghasilkan data tabel kontingensi, salah satunya data siswa putus sekolah. Survei BPS tahun 2018 menunjukkan bahwa jumlah siswa putus sekolah pada tingkat sekolah menengah di provinsi Jawa Barat memiliki angka siswa putus sekolah tertinggi. Sayangnya, penelitian mengenai siswa putus sekolah secara kuantitatif masih jarang dilakukan. Model log linier binomial negatif diterapkan untuk mengetahui faktor-faktor yang memengaruhi jumlah siswa putus sekolah tingkat SMA/SMK di provinsi Jawa Barat tahun 2021. Hasil penelitian menyatakan

bahwa jenis kelamin, jenis sekolah menengah, status sekolah, interaksi jenis kelamin dan status sekolah, dan interaksi jenis sekolah menengah dan status sekolah secara signifikan memengaruhi jumlah siswa putus sekolah. Lebih lanjut, siswa putus sekolah tingkat SMA/SMK di provinsi Jawa Barat didominasi oleh siswa laki-laki, siswa SMK, dan siswa sekolah negeri. Perempuan 1% lebih berisiko putus sekolah dibandingkan laki-laki di sekolah negeri, sedangkan laki-laki 1% lebih berisiko putus sekolah dibandingkan perempuan di sekolah swasta. Siswa SMA 59% lebih berisiko putus sekolah dibandingkan siswa SMK di sekolah negeri, sedangkan siswa SMK 71% lebih berisiko putus sekolah dibandingkan siswa SMA di sekolah swasta.

Keywords: tabel kontingensi, putus sekolah, log linier binomial negatif, overdispersi.

PENDAHULUAN

Analisis regresi merupakan salah satu metode untuk menganalisis hubungan antar variabel penjelas dan variabel respon pada data kontinu. Namun, tidak menutup kemungkinan bahwa data yang dianalisis dapat berupa data diskrit atau data cacahan, salah satunya menggunakan regresi poisson (David & Djamaris, 2018; Kurniawan, 2016).

Model regresi poisson adalah salah satu kasus khusus pada *Generalized Linear Model* (GLM) yang umumnya digunakan untuk memodelkan hubungan antar variabel respons cacahan dan beberapa variabel penjelas berupa data diskrit, kontinu, kategorik, atau campuran (Handayani *et al.*, 2021; Ruliana *et al.*, 2016). Dalam kasus dimana data variabel penjelas adalah kategorik, pada umumnya data akan dimodelkan menjadi jumlah sel dari tabel klasifikasi silang atau tabel kontingensi. Penggunaan regresi poisson pada tabel kontingensi sangatlah umum hingga ia memiliki nama alternatif yaitu model log linier yang menggunakan *link function* log natural (Bilder & Loughin, 2015; Julianda *et al.*, 2019).

Model log linier memiliki fungsi utama memodelkan jumlah sel dalam tabel kontingensi data variabel respons kategorik dengan asumsi data berdistribusi poisson (Agresti, 2019; Sihotang & Zuhri, 2020). Asumsi distribusi poisson dalam model log linier memiliki makna nilai varians dan nilai rata-rata adalah sama, kondisi ini disebut juga dengan equidispersi. Dalam kehidupan nyata, varians dari respons cacahan yang diamati dari data sampel bisa jauh melebihi rata-rata sampel. Fenomena ini biasa dikenal dengan overdispersi.

Metode alternatif yang umumnya diterapkan untuk pemodelan data jumlah sel yang memiliki fenomena overdispersi pada model log linier poisson adalah dengan menggunakan distribusi binomial negatif daripada distribusi poisson standar (Svetliza & Paula, 2001). Penelitian oleh Patummasut *et al.* (2019) serta Svetliza dan Paula (2001) memaparkan bahwa pilihan mengenai bagaimana memodelkan overdispersi memiliki dampak yang jelas pada hasil analisis data, dimana kedua penelitian tersebut menunjukkan bahwa model log linier binomial negatif dinilai lebih cocok untuk data sebagai alternatif overdispersi dalam model log linier poisson.

Penelitian di bidang pendidikan seringkali menghasilkan data tabel kontingensi, salah satunya yaitu data siswa putus sekolah. Survei BPS tahun 2018 menunjukkan bahwa siswa pendidikan menengah konsisten mendominasi angka putus sekolah selama tiga tahun ajaran dengan provinsi Jawa Barat secara tetap menduduki posisi pertama sebagai provinsi dengan angka siswa putus paling tinggi di Indonesia. Namun, penelitian mengenai siswa putus sekolah secara kuantitatif masih jarang dilakukan. Dalam penelitian ini, model log linier binomial negatif akan diterapkan untuk menangani overdispersi pada data jumlah siswa putus sekolah tingkat SMA/SMK di provinsi Jawa Barat tahun 2021 guna mengetahui faktor-faktor yang memengaruhinya.

METODOLOGI

Bahan dan Data

Data yang akan dianalisis dalam penelitian ini merupakan data sekunder diperoleh dari Statistik Persekolahan SMA tahun ajaran 2020/2021 dan Statistik Persekolahan SMA tahun ajaran 2020/2021

yang dipublikasikan oleh Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia ditampilkan pada TABEL 1.

TABEL 1. Data Jumlah Siswa Putus Sekolah

Jenis Kelamin	Jenis Sekolah Menengah	Status Sekolah		Jumlah
		Negeri	Swasta	
Laki-laki	SMA	408	230	638
	SMK	481	624	1.105
Perempuan	SMA	314	145	459
	SMK	208	346	554
Total		1.411	1.345	2.756

Populasi penelitian ini adalah seluruh siswa SMA dan SMK di daerah provinsi Jawa Barat yang putus sekolah pada tahun ajaran 2020/2021. Variabel penelitian yang digunakan terdiri dari tiga variabel penjelas (X) seperti yang ditampilkan pada TABEL 2.

TABEL 2. Variabel Penelitian

Notasi	Nama Variabel	Tipe	Kategori
X_1	Jenis Kelamin	Nominal	1 : Laki-laki 2 : Perempuan
X_2	Jenis Sekolah Menengah	Nominal	1 : SMA 2 : SMK
X_3	Status Sekolah	Nominal	1 : Negeri 2 : Swasta

Metode Penelitian

Prosedur analisis data dalam penelitian ini adalah sebagai berikut:

1. Menginput data ke *software* R Studio untuk melakukan analisis pada data.
2. Mengolah statistika deskriptif data untuk mengenal karakteristik siswa putus sekolah tingkat SMA/SMK di Jawa Barat guna mempermudah proses analisis data yang akan dilakukan.
3. Menguji asumsi syarat penggunaan metode analisis log linier, yaitu:
 - 1) Observasi independen satu sama lain, setiap pengamatan berasal dari subjek yang berbeda.
 - 2) Memiliki distribusi observasi yang identik. data diambil dari populasi yang homogen dengan cara perolehan yang sama.
 - 3) Jumlah observasi cukup banyak. Model log linier membutuhkan sampel besar dan sel kosong tidak diperbolehkan.
 - 4) Frekuensi harapan tidak boleh terlalu kecil. Frekuensi harapan dengan nilai < 5 tidak boleh melebihi 20% sel, dan tidak diperbolehkan sel dengan frekuensi harapan < 1 .
4. Melakukan pendugaan parameter model log linier poisson. Model log linier yang dispesifikasikan dalam penelitian ini menggunakan persamaan sebagai berikut:

$$\log(\mu_i) = \log(t_i) + \mathbf{x}_i^T \boldsymbol{\beta}, \tag{1}$$

dimana $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})^T$ berisi nilai dari p variabel penjelas, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ dan $\log(t_i)$ adalah *offset* (Hilbe, 2011; Patummasut *et al.*, 2019) pendugaan parameter model log linier metode *maximum likelihood* pada model log linier.

5. Melakukan uji overdispersi pada model log linier dengan mengikuti Crawley (2013) serta Rashwan dan Kamel (2011), dapat diamati melalui perhitungan rasio antara nilai *deviance* dengan derajat bebas sebagai berikut:

$$\text{rasio dispersi} = \frac{\text{Deviance}}{\text{db}} \tag{2}$$

Dengan kriteria jika rasio dispersi > 1 artinya model log linier mengalami overdispersi dan tidak cocok untuk data yang diamati.

6. Melakukan pendugaan parameter model log linier binomial negatif menggunakan persamaan (2) dengan metode *maximum likelihood* pada model log linier.
7. Pemodelan akan dilanjutkan dengan mencoba seluruh model yang mungkin untuk analisis log linier tiga variabel kategorik.
8. Melakukan uji evaluasi pada model-model yang terbentuk pada langkah sebelumnya menggunakan uji *goodness-of-fit*. Hipotesis yang berlaku dalam kasus tiga variabel kategorik dimana n_{ijk} menjadi jumlah unit yang memiliki variabel X_1 terdiri dari i kelompok, variabel X_2 terdiri dari j kelompok, dan variabel X_3 terdiri dari k kelompok, yaitu:

H_0 : Model sesuai (n_{ijk} mendekati μ_{ijk})

H_1 : Model tidak sesuai (n_{ijk} tidak mendekati μ_{ijk})

dengan rumus statistik uji *Pearson Chi Square* (χ^2) atau *Likelihood Ratio Test* (G^2) menurut Christensen (1997) adalah sebagai berikut :

$$\chi^2 = \sum_i \sum_j \sum_k \frac{(n_{ijk} - \mu_{ijk})^2}{\mu_{ijk}} \tag{3}$$

$$G^2 = 2 \sum_i \sum_j \sum_k n_{ijk} \log \left(\frac{n_{ijk}}{\mu_{ijk}} \right) \tag{4}$$

dimana

n_{ijk} = jumlah amatan sampel pada baris ke- i , layer ke- j dan kolom ke- k .

μ_{ijk} = nilai harapan pada baris ke- i , layer ke- j dan kolom ke- k .

Dengan kriteria uji terima H_0 ketika χ^2_{hitung} atau $G^2_{hitung} \leq \chi^2_{(df; \alpha)}$ dan nilai *p-value* < α .

9. Memilih model log linier binomial negatif terbaik untuk data dengan melihat nilai AIC terkecil dengan rumus sebagai berikut.

$$AIC = -2l + 2p \tag{5}$$

dimana l adalah nilai fungsi *log-likelihood*, p merupakan jumlah parameter.

10. Menginterpretasi kecocokan model log linier binomial negatif pada data menggunakan nilai risiko relatif (Tenny & Hoffman, 2022).

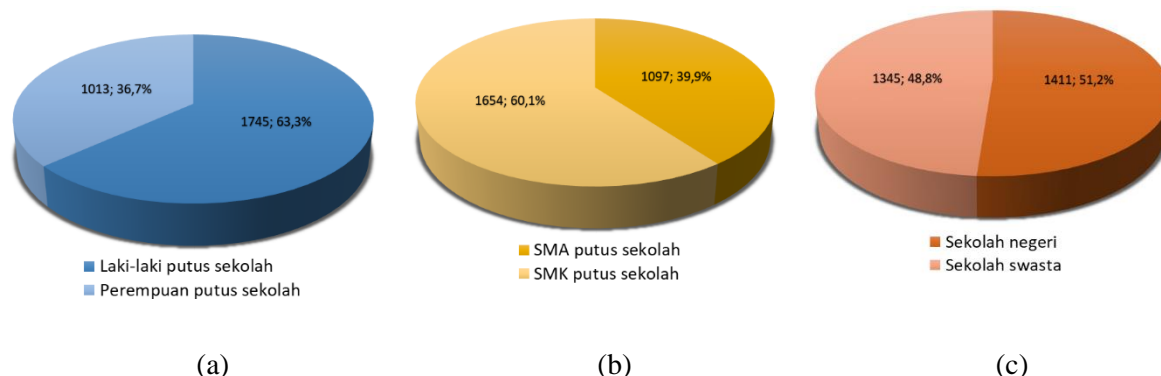
$$\text{Relative risk} = \frac{\pi_1}{\pi_2} \tag{6}$$

dimana π_1 adalah peluang berhasil dengan untuk subjek di baris 1 dan dengan π_2 untuk subjek di baris 2.

HASIL DAN PEMBAHASAN

Statistik Deskriptif Data

Statistik deskriptif data pada penelitian ini bertujuan untuk melihat karakteristik responden dari perbandingan hubungan variabel penjelas menggunakan diagram pie seperti yang disajikan pada gambar di bawah ini.



GAMBAR 1. Persentase siswa putus sekolah pada tahun 2021 di provinsi Jawa Barat berdasarkan kategori : (a) jenis kelamin, (b) jenis sekolah menengah, (c) status sekolah

Terlihat dari GAMBAR 1 (a) persentase siswa putus sekolah berjenis kelamin laki-laki lebih besar dibandingkan dengan siswa putus sekolah berjenis kelamin perempuan. Berdasarkan GAMBAR 1 (b) persentase siswa putus sekolah di SMK lebih besar dibandingkan dengan siswa putus sekolah di SMA. Menurut GAMBAR 1 (c) persentase siswa putus sekolah di sekolah negeri lebih besar dibandingkan dengan siswa putus sekolah di sekolah swasta.

Uji Asumsi

Frekuensi sel harapan pada data yang digunakan pada penelitian ini dihitung untuk melihat apakah data layak untuk dimodelkan menggunakan analisis log linier.

TABEL 3. Estimasi Frekuensi Sel Harapan

Jenis Kelamin	Jenis Sekolah Menengah	Status Sekolah	
		Negeri	Swasta
Laki-laki	SMA	325,406	312,594
	SMK	563,594	541,406
Perempuan	SMA	236,523	222,477
	SMK	285,477	268,523

TABEL 3 menunjukkan bahwa nilai frekuensi sel harapan yang kurang dari 5, tidak lebih besar dari 20%, dan tidak ada sel yang memiliki frekuensi harapan < 1. Maka data disimpulkan memenuhi asumsi untuk dilakukan pemodelan log linier secara lanjut.

Pemodelan Log linier

Jika asumsi-asumsi pada data sudah terpenuhi, maka dilakukan pemodelan log linier poisson pada data penelitian untuk mendapatkan nilai estimasi parameter dari model melalui metode *maximum likelihood* dengan keluaran sebagai berikut.

TABEL 4. Estimasi Parameter Log linier Poisson

Parameter	Estimasi	Std. Error	Z-value	Sig.	Residual deviance	Db
Intersep	-6,108	0,036	-165,63	<2e ⁻¹⁶	49,904	4
StatusSwasta	-0,431	0,040	-10,60	<2e ⁻¹⁶		
JenisSMK	0,042	0,041	1,01	0,312		
GenderPerempuan	-0,462	0,039	-11,60	<2e ⁻¹⁶		

Terlihat bahwa model di atas merupakan model yang baik karena hampir seluruh faktor bernilai signifikan pada data, namun dapat dilihat nilai rasio dispersi > 1 pada model log linier distribusi poisson yaitu sebesar 12,708 yang menandakan adanya overdispersi. Oleh karena itu, model log linier binomial negatif akan digunakan untuk menghitung estimasi nilai harapan melalui metode *maximum likelihood*.

Uji Goodness-of-it

Setelah mendapat estimasi nilai harapan model log linier binomial negatif maka dilanjutkan dengan melakukan uji *goodness-of-it* untuk mengetahui apakah model tersebut dapat menggambarkan data dengan baik atau tidak menggunakan uji pearson chi square (χ^2).

TABEL 5. Uji Goodness-of-it

Tahap	Model	Pearson Chi Square (χ^2_{hitung})	Db	$\chi^2_{(df; \alpha)}$	AIC
1	$\log(\mu_{ijk}) = \log(t_i) + \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$	56,994	4	9,487	92,225
2	$\log(\mu_{ijk}) = \log(t_i) + \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1 x_2$	51,367	3	7,814	93,816
3	$\log(\mu_{ijk}) = \log(t_i) + \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_5 x_1 x_3$	45,859	3	7,814	93,419
4	$\log(\mu_{ijk}) = \log(t_i) + \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_6 x_2 x_3$	11,222	3	7,814	82,816
5	$\log(\mu_{ijk}) = \log(t_i) + \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1 x_2 + \beta_5 x_1 x_3$	42,338	2	5,991	94,816
6	$\log(\mu_{ijk}) = \log(t_i) + \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1 x_2 + \beta_6 x_2 x_3$	6,278	2	5,991	80,991
7	$\log(\mu_{ijk}) = \log(t_i) + \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_5 x_1 x_3 + \beta_6 x_2 x_3$	3,836	2	5,991	78,531
8	$\log(\mu_{ijk}) = \log(t_i) + \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1 x_2 + \beta_5 x_1 x_3 + \beta_6 x_2 x_3$	1,503	1	3,841	78,186
9	$\log(\mu_{ijk}) = \log(t_i) + \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1 x_2 + \beta_5 x_1 x_3 + \beta_6 x_2 x_3 + \beta_7 x_1 x_2 x_3$	0	0	0	104,749

Berdasarkan TABEL 5 didapatkan model log linier binomial negatif terbaik yaitu model dimana nilai $\chi^2_{hitung} \leq \chi^2_{(df; \alpha)}$ dengan nilai AIC terkecil adalah model yang memuat tiga interaksi dua faktor dengan persamaan sebagai berikut:

$$\log(\mu_{ijk}) = \log(t_i) + \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1 x_2 + \beta_5 x_1 x_3 + \beta_6 x_2 x_3 \tag{6}$$

dimana

μ_{ijk} = frekuensi nilai harapan dalam sel-*ijk*

$\log(t_{ijk})$ = *offset* sel-*ijk*

β_0 = efek "keseluruhan", atau *mean* besar dari logaritma dari jumlah yang diharapkan

$\beta_1, \beta_2, \beta_3$ = efek "utama" dari variabel $X_1, X_2,$ dan $X_3,$ atau penyimpangan dari *mean* utama

$\beta_4, \beta_5, \beta_6$ = interaksi/asosiasi antara dua variabel selagi mengendalikan yang ketiga

x_1 = variabel jenis kelamin (X_1)

x_2 = variabel jenis sekolah menengah (X_2)

x_3 = variabel status sekolah (X_3)

Dengan $\log(t_{ijk})$ sebagai *offset*, dimana *offset* yang digunakan adalah jumlah siswa sekolah tingkat SMA/SMK di Provinsi Jawa Barat. Estimasi parameter untuk model terbaik log linier binomial negatif dapat dilihat seperti pada TABEL 6.

TABEL 6. Estimasi Parameter Log linier Binomial Negatif

Parameter	Estimasi	Std. Error	z-value	Sig.	Residual deviance	Db
Intersep	-6,139	0,046	-130,859	< 2e-16 ***		
GenderPerempuan	-0,615	0,067	-9,152	< 2e-16 ***		
JenisSMK	0,196	0,061	3,189	0,00143 **		
StatusSwasta	-0,187	0,072	-2,607	0,00914 **	1,507	1
GenderPerempuan × JenisSMK	-0,127	0,083	1,532	0,12552		
GenderPerempuan × StatusSwasta	0,179	0,082	2,191	0,02846 *		
JenisSMK × StatusSwasta	-0,491	0,081	-6,060	1,36e-09 ***		

Diketahui bahwa pada model log linier binomial negatif terbaik, seluruh faktor signifikan karena nilai Sig. < $\alpha = 0,05$ selain untuk interaksi faktor jenis kelamin dan jenis sekolah menengah (X_1X_2). Nilai rasio dispersi yang dihasilkan sebesar 1,507. Rasio dispersi yang dihasilkan mendekati satu atau lebih kecil dibandingkan model log linier poisson yang sebesar 12,708. Hal ini menunjukkan bahwa model log linier binomial negatif lebih cocok dengan data yang dianalisis pada penelitian ini, yang artinya juga mampu meminimumkan masalah overdispersi yang terjadi pada data. Berdasarkan TABEL 6 dapat juga diperoleh koefisien parameter model guna menginterpretasi lebih dalam mengenai pengaruh masing-masing faktor pada data yaitu:

- 1) Koefisien parameter Jenis Kelamin (X_1) sebesar -0,615 menunjukkan bahwa tanpa memperhatikan jenis sekolah menengah dan status sekolah, peluang siswa putus sekolah berjenis kelamin perempuan adalah $\exp(-0,615) = 0,540$ kali dibandingkan laki-laki. Atau peluang siswa putus sekolah berjenis kelamin laki-laki $\frac{1}{0,540} = 1,851$ kali dibandingkan perempuan.
- 2) Koefisien parameter Jenis Sekolah Menengah (X_2) sebesar 0,196 menunjukkan bahwa tanpa memperhatikan jenis kelamin dan status sekolah, peluang siswa putus sekolah di SMK adalah $\exp(0,196) = 1,216$ kali dibandingkan SMA. Atau peluang siswa putus sekolah di SMA $\frac{1}{1,216} = 0,822$ kali dibandingkan SMK.
- 3) Koefisien parameter Status Sekolah (X_3) sebesar -0,187 menunjukkan bahwa tanpa memperhatikan jenis kelamin dan jenis sekolah menengah, peluang siswa putus sekolah di sekolah swasta adalah $\exp(-0,187) = 0,828$ kali dibandingkan sekolah negeri. Atau peluang siswa putus sekolah di sekolah negeri $\frac{1}{0,828} = 1,207$ kali dibandingkan sekolah swasta.
- 4) Koefisien parameter interaksi jenis kelamin dan status sekolah (X_1X_3) sebesar 0,179 menunjukkan bahwa tanpa memperhatikan jenis sekolah menengah, *odds* siswa putus sekolah di sekolah swasta (dibandingkan sekolah negeri) jika dia berjenis kelamin perempuan adalah $\exp(0,179) = 1,197$ kali dibandingkan *odds* yang sama jika dia laki-laki.
- 5) Koefisien parameter interaksi jenis sekolah menengah dan status sekolah (X_2X_3) sebesar -0,491 menunjukkan bahwa tanpa memperhatikan jenis kelamin, *odds* siswa putus sekolah di sekolah swasta (dibandingkan sekolah negeri) jika dia dari SMK adalah $\exp(-0,491) = 0,611$ kali dibandingkan *odds* yang sama jika dia dari SMA.

Risiko Relatif

Berdasarkan model log linier binomial negatif terbaik yang terpilih, terdapat dua interaksi dua faktor yang dinyatakan signifikan pada data, yaitu interaksi faktor jenis kelamin dan status sekolah (X_1X_3) serta interaksi faktor jenis sekolah menengah dan status sekolah (X_2X_3). Namun, cukup sulit mengartikan interaksi faktor melalui koefisien parameternya saja. Melalui nilai risiko relatif akan dilihat lebih jelas hubungan antara faktor tersebut terhadap data penelitian.

Menurut nilai risiko relatif dari interaksi faktor jenis kelamin dan status sekolah bahwa siswa berjenis kelamin perempuan 1% lebih berisiko putus sekolah dibandingkan laki-laki di sekolah negeri, sedangkan laki-laki 1% lebih berisiko putus sekolah dibandingkan perempuan di sekolah swasta.

Nilai risiko relatif dari interaksi faktor jenis sekolah menengah dan status sekolah memperlihatkan bahwa siswa SMA 59% lebih berisiko putus sekolah dibandingkan siswa SMK di sekolah negeri, sedangkan siswa SMK 71% lebih berisiko putus sekolah dibandingkan siswa SMA di sekolah swasta.

KESIMPULAN

Berdasarkan hasil analisis yang sudah dibahas sebelumnya, model log linier binomial negatif dinyatakan layak menjadi salah satu metode alternatif dari analisis kasus overdispersi pada model log linier poisson. Analisis log linier binomial negatif menghasilkan model terbaik yaitu jenis model tiga interaksi dua faktor dimana faktor jenis kelamin, jenis sekolah menengah, status sekolah, interaksi jenis kelamin dan status sekolah, serta interaksi jenis sekolah menengah dan status sekolah dinyatakan memengaruhi jumlah siswa putus sekolah tingkat SMA/SMK di provinsi Jawa Barat tahun 2021.

REFERENSI

- Agresti, A. (2019). *Wiley Series in Probability and Statistics*. <http://www.wiley.com/go/wsp>
- Bilder, C. R., & Loughin, T. M. (2015). *Analysis of Categorical Data with R* (F. Dominici, J. J. Julian J. Faraway, M. Martin Tanner, & J. Jim Zidek, Eds.). CRC Press Taylor & Francis Group.
- Christensen, R. (1997). *Log-Linear Models and Logistic Regression* (2nd ed.). Springer-Verlag New York, Inc.
- Crawley, M. J. (2013). *The R book* (2nd ed.).
- David, W., & Djamaris, Aurino. R. A. (2018). *Metode Statistik Untuk Ilmu dan Teknologi Pangan*. Penerbitan Universitas Bakrie.
- Handayani, D., Artari, A. F., Safitri, W., Rahayu, W., & Santi, V. M. (2021). Count Regression Models for Analyzing Crime Rates in The East Java Province. *Journal of Physics: Conference Series*, 2123(1), 012028. <https://doi.org/10.1088/1742-6596/2123/1/012028>
- Hilbe, J. M. (2011). *Negative Binomial Regression* (2nd ed.). Cambridge University Press.
- Julianda, R., Herrhyanto, N., & Avip, B. (2019). Penerapan Data Count Dengan Menggunakan Regresi Hurdle Poisson. *Jurnal Eureka Matika*, 7(1), 11–23.
- Kurniawan, C. (2016). Analisis Data Hubungan Antar Variabel Sebagai Metode Alternatif Penentuan Hubungan Kausalitas. *Sinteks : Jurnal Teknik*, 5(2).
- Patummasut, M., Phewchean, N., & Sirirattanapa, J. (2019). *Modelling Motorcycle-Related Head Injury Trends for Thailand Following the 100% Motorcycle Helmet Use Campaign Using Log-linear Model* (Vol. 17, Issue 1). <http://statassoc.or.th>
- Rashwan, N. A., & Kamel, M. M. (2011). Using Generalized Poisson Log Linear Regression Models in Analyzing Two-Way Contingency Tables. *Applied Mathematical Sciences*, 5(5), 213–222.
- Ruliana, Hendikawati, P., & Agoestanto, A. (2016). Pemodelan Generalized Poisson Regression (GPR) Untuk Mengatasi Pelanggaran Equidispersi Pada Regresi Poisson Kasus Campak di Kota Semarang Tahun 2013. *UNNES Journal of Mathematics*, 5(1), 40–46.
- Sihotang, S. F., & Zuhri. (2020). Analisis Model Log Linier Tiga Dimensi Untuk Data Kualitatif Dengan Metode Forward. *Journal of Mathematics Education and Science*, 6(1), 62–69.
- Svetliza, C. F., & Paula, G. A. (2001). On diagnostics in log-linear negative binomial models. *Journal of Statistical Computation and Simulation*, 71(3), 231–244. <https://doi.org/10.1080/00949650108812145>
- Tenny, S., & Hoffman, M. R. (2022). *Relative Risk*.