

Analisis Regresi Logistik Biner dengan Metode PMLE pada Penyakit Covid-19

Lutfi Anisa, Nur Azizah Komara Rifai*

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

ARTICLE INFO

Article history :

Received : 19/8/2022

Revised : 22/11/2022

Published : 21/12/2022



Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Volume : 2

No. : 2

Halaman : 129-136

Terbitan : **Desember 2022**

ABSTRAK

Estimasi parameter dari model regresi logistik biner biasanya diselesaikan dengan menggunakan metode maximum likelihood estimation (MLE) yang selanjutnya diselesaikan metode iterasi Newton-Raphson. Namun, dalam keadaan tertentu metode MLE tidak bisa digunakan jika terdapat satu atau kombinasi beberapa variabel penjelas dan respon terdapat masalah pemisahan pada data (separation case) yang mengakibatkan penduga menjadi tidak konvergen. Untuk mengatasi masalah tersebut, digunakan pendekatan metode penalized maximum likelihood estimation (PMLE). Hasil analisis terhadap penelitian ini menunjukkan bahwa data pasien Covid-19 RSUD Pringsewu mengandung pemisahan kurang sempurna (quasi-complete separation), sehingga digunakan pendekatan metode PMLE. Penaksiran parameter model terbaik bagi data ini mengandung faktor jenis kelamin, umur, penyakit bawaan (komorbid), dan gejala sesak nafas yang berpengaruh secara signifikan terhadap tingkat kematian akibat Covid-19. Dari hasil analisis penelitian ini jika dibandingkan dengan penelitian terdahulu Sholihin, et al. Dalam jurnal yang berjudul "Implementasi Metode Penalized Maximum Likelihood Estimation Pada Model Regresi Logistik Biner" terbukti bahwa analisis regresi logistik biner jika datanya terdapat masalah separation case dengan menggunakan metode PMLE dapat memberikan hasil yang lebih baik.

Kata Kunci : Covid-19; Regresi Logistik Biner; Penalized Maximum Likelihood

ABSTRACT

Parameter estimation from binary logistic regression models is usually solved using the maximum likelihood estimation (MLE) method which is then completed by the Newton-Raphson iteration method. However, in certain circumstances the MLE method cannot be used if there is one or a combination of several explanatory variables and the response is a separation case problem that causes the estimators to not converge. To overcome this problem, the penalized maximum likelihood estimation (PMLE). The results of the analysis of this study indicate that the Covid-19 patient data at the Pringsewu Hospital contains a quasi-complete separation, so the PMLE method approach is used. The estimation of the best model parameters for this data contains factors of gender, age, congenital disease (comorbid), and symptoms of shortness of breath that significantly affect the mortality rate due to Covid-19. From the results of the analysis of this study when compared with previous studies Sholihin, et al. In the journal entitled "Implementasi Metode Penalized Maximum Likelihood Estimation Pada Model Regresi Logistik Biner" it is proven that binary logistic regression analysis if the data has case separation problems using the PMLE method can give better results.

Keywords : Covid-19; Binary Logistics Regression; Penalized Maximum Likelihood

© 2022 Jurnal Riset Statistika Unisba Press. All rights reserved.

A. Pendahuluan

Analisis Menurut Hosmer dan Lemeshow [1] analisis regresi merupakan salah satu metode yang sangat populer dalam mencari hubungan antara satu atau lebih variabel penjelas (X) dan variabel respon (Y). Analisis regresi merupakan salah satu metode statistika yang digunakan untuk menganalisis hubungan antara satu variabel dependen dengan satu atau lebih variabel independen [2]. Analisis regresi umumnya menggunakan variabel respon kuantitatif. Analisis regresi merupakan pengukur hubungan dua variabel atau lebih yang dinyatakan dengan bentuk hubungan/fungsi [3]. Sedangkan analisis regresi logistik digunakan jika variabel respon berupa kategori. Regresi logistik merupakan suatu metode analisis data yang mendeskripsikan antara variabel respon dengan satu atau lebih variabel prediktor [4]. Regresi logistik biner adalah analisis regresi logistik untuk menganalisis hubungan antara satu variabel respon dalam bentuk biner (dikotomis) terhadap satu atau lebih variabel penjelas. Variabel dikotomi yaitu variabel yang hanya memiliki dua kemungkinan nilai, atau variabel berupa dua kategori yang bernilai 1 untuk menyatakan keberadaan sebuah karakteristik dan bernilai 0 untuk menyatakan ketidakberadaan sebuah karakteristik.

Estimasi parameter dari model regresi logistik biner biasanya menggunakan metode *maximum likelihood estimation* (MLE) yang selanjutnya diselesaikan dengan metode iteratif newton-raphson. *Maximum Likelihood Estimation* merupakan metode yang terdiri atas estimasi dari parameter-parameter yang tidak diketahui dalam perilakunya bahwa probabilitas dalam mengobservasi variabel Y yang telah ditentukan ini dilakukan setinggi mungkin [5]. Namun menurut Albert dan Anderson [6] metode ini tidak bisa digunakan jika satu atau kombinasi beberapa variabel penjelas dan respon terdapat masalah pemisahan pada data (*separation case*), hal ini mengakibatkan penduga menjadi tidak konvergen. Masalah *separation case* perlu diselesaikan karena menimbulkan bias estimasi parameter. Masalah *separation case* umumnya terjadi pada data penelitian medis di rumah sakit, karena setiap gejala penyakit yang dialami semua orang belum tentu sama.

Jika dibandingkan dari Penelitian terdahulu yaitu penelitian Devika, *et al.* [7] diperoleh bahwa metode *penalized maximum likelihood estimation* (PMLE) yang pertama kali diusulkan oleh Firth [8] dapat menjadi solusi untuk mengatasi masalah *separation case* dalam regresi logistik. Menurut Firth [8] PMLE adalah hasil dari perubahan fungsi skor likelihood menjadi fungsi skor *penalized likelihood*. Serta menurut penelitian Sholihin, *et al.* [9] dan Devika, *et al.* [7] bahwa metode PMLE terbukti menjadi solusi ideal untuk menyelesaikan penduga yang tidak konvergen dan masalah *odds ratios* (OR) yang tinggi. Penelitian Ecevit [10] juga menyampaikan hasil kesimpulan bahwa dalam masalah *separation case*, metode PMLE lebih unggul dari metode MLE karena dapat mengurangi nilai bias estimate parameter.

Berdasarkan departemen agama RI di dalam Q.S. Al Jum'ah/62:8 disebutkan:

قُلْ إِنَّ الْمَوْتَ الَّذِي تَفِرُّونَ مِنْهُ فَإِنَّهُ مُلْفِيكُمْ ثُمَّ تُرَدُّونَ إِلَىٰ عَالِمِ الْغَيْبِ وَالشَّهَادَةِ فَيُنبِّئُكُمْ بِمَا كُنْتُمْ تَعْمَلُونَ

Terjemahannya: Katakanlah, “Sesungguhnya kematian yang kamu lari dari padanya, ia pasti menemui kamu, kemudian kamu akan dikembalikan kepada (Allah), yang mengetahui yang gaib dan yang nyata, lalu Dia beritakan kepadamu apa yang telah kamu kerjakan”.

Berdasarkan ayat di atas penulis tertarik untuk membahas tentang kematian khususnya kematian akibat corona virus disease (Covid-19), maka dari itu metode PMLE ini akan diterapkan pada data penyakit Covid-19 karena Indonesia sempat menempati kasus kematian harian tertinggi di dunia selama sepekan penuh.

Pada tahun 2021, Indonesia dilanda gelombang kedua Covid-19 yang bertepatan dengan masuknya libur panjang Hari Raya Idul Fitri dan hadirnya varian baru delta. Gelombang kedua muncul pada bulan Mei yaitu memasuki minggu kesembilan sampai bulan Juli. Indonesia mencatat pada 15 Juli 2021 jumlah kasus tertinggi ada sebanyak 56.757 kasus dan juga mencatat jumlah kematian yang mencapai rekor tertinggi di dunia selama satu pekan berturut-turut, meningkat 34,5% selama periode gelombang kedua (Satuan Tugas (satgas) Covid, 2021), yaitu saat Hari Raya Idul Fitri 2021 atau bertepatan pada tanggal 21 sampai 27 Juli. Di mana pada saat itu Indonesia menempati posisi teratas kasus kematian tertinggi di dunia yang tercatat ada sebanyak 1.383 kasus pada 21 Juli 2021. Berlanjut menempati posisi tertinggi kasus kematian pada 22 Juli tercatat ada sebanyak 1.449 kasus, 23 Juli tercatat ada sebanyak 1.566 kasus, 24 Juli tercatat ada sebanyak 1.415 kasus, 25 Juli tercatat ada sebanyak 1.266 kasus dan pada 26 Juli tercatat ada sebanyak 1.487 kasus. Bahkan sampai 27 Juli, Indonesia masih berada di posisi pertama dunia dengan jumlah kematian sebanyak 2.069 kasus. Sejak

pandemi pertama kali melanda pada Maret 2020, Ini merupakan angka kematian tertinggi di Indonesia hingga saat ini.

Menurut laporan Satgas Covid-19 pada Juli 2021, Lampung menjadi salah satu dari lima Provinsi tertinggi dengan tingkat kematian Covid-19 sebanyak 1.347 kasus (4,24%). Salah satu kabupaten di Lampung yaitu Pringsewu mengalami kenaikan kasus tertinggi sebanyak 2 kali lipat dari bulan sebelumnya yaitu 163 kasus menjadi 377 kasus, dan mencapai angka kematian penduduk tertinggi 23,74 kasus per 100.000 penduduk. Kematian penduduk akibat Covid-19 ini disebabkan oleh beberapa faktor, seperti faktor umur, jenis kelamin, gejala sesak nafas dan adanya penyakit bawaan (komorbid) dari pasien (WHO, 2020). Menurut laporan Satgas Covid-19 85,31% kematian berasal dari pasien dengan umur 46-59 tahun dan 53,50% kematian berjenis kelamin laki-laki. Oleh karena itu Firth (4) mengusulkan bahwa bias orde pertama pertama $O(N^{-1})$ dapat dihilangkan dengan memodifikasi skor likelihood (\mathbf{u}) yaitu:

$$\mathbf{b}(\boldsymbol{\beta}) = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\boldsymbol{\xi} \tag{1}$$

Dari modifikasi tersebut didapatkan fungsi *penalized likelihood* yaitu:

$$L^*(\boldsymbol{\beta}) = L(\boldsymbol{\beta})|I(\boldsymbol{\beta})|^{\frac{1}{2}} \tag{2}$$

Dengan *penalized log-likelihood* yaitu;

$$l^*(\boldsymbol{\beta}) = l(\boldsymbol{\beta}) + \frac{1}{2} \ln |I(\boldsymbol{\beta})| \tag{3}$$

Berdasarkan latar belakang yang telah diuraikan, maka pada penelitian ini akan dilakukan analisis regresi logistik biner dengan metode *penalized maximum likelihood estimation* (PMLE) pada penyakit Covid-19 di RSUD Pringsewu. Selanjutnya, tujuan dalam penelitian ini diuraikan dalam pokok-pokok sbb. (1) Untuk mengetahui penaksiran parameter model regresi logistik menggunakan metode *penalized maximum likelihood estimation* pada kasus Covid-19 di RSUD Kabupaten Pringsewu pada bulan Mei sampai November tahun 2021; (2) Untuk mengetahui faktor-faktor apa saja yang berpengaruh secara signifikan terhadap kematian pasien Covid-19 di RSUD Kabupaten Pringsewu pada bulan Mei sampai November tahun 2021.

B. Metode Penelitian

Data Penelitian

Dalam penelitian ini data yang digunakan yaitu data sekunder, yang didapatkan dari bagian rekam medis RSUD Kabupaten Pringsewu yang beralamat di Jl. Jend. Ahmad Yani, Pajar Agung, Kecamatan Pringsewu, Kabupaten Pringsewu, Lampung 35376. Data penelitian ini terdiri dari 1017 pasien Covid-19.

Identifikasi Variabel

Variabel yang digunakan dalam penelitian ini adalah satu buah variabel respon dan empat buah variabel penjelas. Variabel respon berupa kondisi terakhir pasien saat di rumah sakit, empat variabel penjelas yang digunakan yaitu jenis kelamin, umur, penyakit bawaan atau *komorbid*, dan gejala sesak nafas. Variabel-variabel tersebut dinyatakan sebagai berikut:

Status akhir pasien (Y) : 0 (pasien tidak meninggal), 1 (pasien meninggal); Jenis Kelamin (X_1) : 0 (perempuan), 1 (laki-laki); Umur (X_2) dalam tahun; Penyakit bawaan (X_3) : 0 (pasien tidak ada komorbid), 1 (pasien ada komorbid); Gejala sesak nafas (X_4) : 0 (pasien tidak ada gejala sesak nafas), 1 (pasien ada gejala sesak nafas).

Pengolahan data dalam penelitian ini menggunakan *software* R, dan *package* yang digunakan adalah *Logistf*. *Package* *Logistf* adalah salah satu *package* statistika yang secara khusus menangani masalah data yang mengandung pemisahan (*separation case*) dengan menggunakan metode *penalized maximum likelihood estimation* (PMLE).

Langkah Penelitian

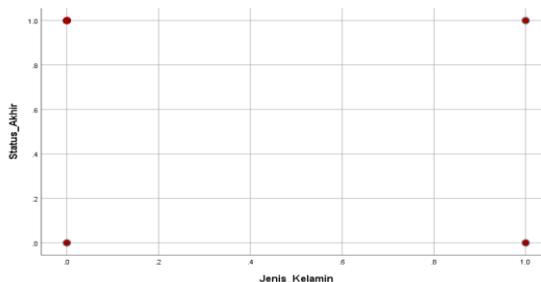
Langkah-langkah yang dilakukan untuk menyelesaikan penelitian ini adalah sebagai berikut: (1) Menginput data penelitian ke dalam *software* R; (2) Lakukan eksplorasi data dengan membuat 4 buah *scatter plot*, yaitu antara variabel respon dengan masing-masing variabel penjelas; (3) Menentukan penaksiran parameter model regresi logistik menggunakan metode PMLE; (4) Melakukan uji signifikansi parameter ; (5) Melakukan uji signifikansi parameter secara serentak (simultan) dengan Uji G; (6) Melakukan uji signifikansi parameter

secara individu (parsial) dengan Uji *Wald*; (7) Melakukan uji kecocokan model untuk menguji apakah sudah layak model yang dihasilkan berdasarkan regresi logistik biner; (8) Interpretasi *odds ratio* yang didapat. *Odds ratio* yaitu ukuran rata-rata besarnya kecenderungan variabel respon; (9) Membuat kesimpulan dan saran dari hasil analisis dan pembahasan.

C. Hasil dan Pembahasan

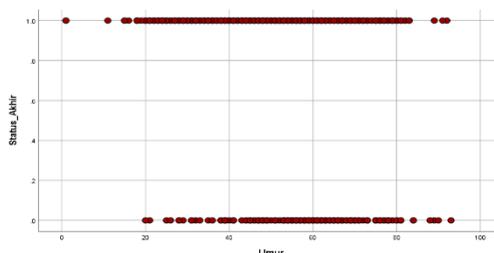
Eksplorasi Data Untuk Variabel Penjelas

Sebelum menaksir parameter model regresi logistik biner, maka dilakukan eksplorasi data terlebih dahulu menggunakan *scatter plot* untuk memeriksa ada atau tidaknya masalah pemisahan (*separation case*) pada data penelitian, dapat dilihat pada gambar di bawah ini.



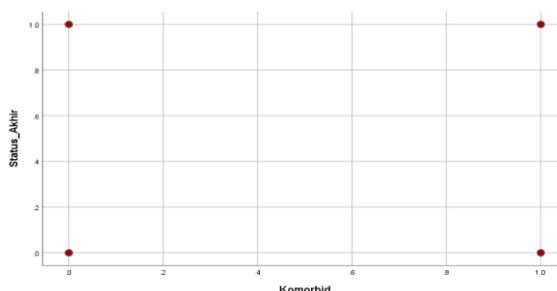
Gambar 1. *Scatter plot* antara Variabel Status Akhir dengan Variabel Jenis Kelamin

Dari gambar *scatter plot* di atas antara status akhir dengan jenis kelamin pasien dapat dilihat bahwa tidak terjadi *separation case* karena ketika jenis kelaminnya adalah 0 maka nilai status akhirnya ada di keduanya (ada di $Y = 0$ dan $Y = 10$), begitupun ketika jenis kelaminnya adalah 1 maka nilai status akhirnya juga ada di keduanya (ada di $Y = 0$ dan $Y = 10$).



Gambar 2. *Scatter plot* antara Variabel Status Akhir dengan Variabel Umur

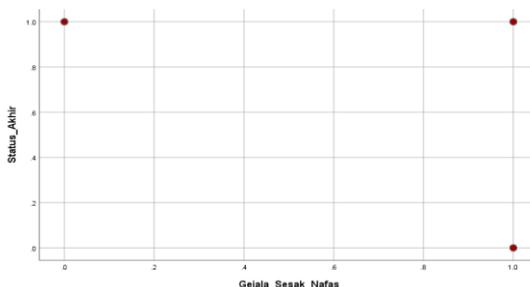
Dari *scatter plot* antara status akhir pasien dengan umur pasien di atas dapat dilihat bahwa tidak terjadi *separation case* karena untuk nilai setiap umur ada di kedua nilai status akhir (ada di $Y = 0$ dan $Y = 10$).



Gambar 3. *Scatter plot* antara Variabel Status Akhir dengan Variabel Komorbid

Dari gambar *scatter plot* di atas antara status akhir dengan komorbid pasien dapat dilihat bahwa tidak terjadi *separation case* karena ketika komorbidnya adalah 0 maka nilai status akhirnya ada di keduanya (ada

di $Y = 0$ dan $Y = 1$), begitupun ketika komorbidnya adalah 1 maka nilai status akhirnya juga ada di keduanya (ada di $Y = 0$ dan $Y = 1$).



Gambar 4. Scatter plot antara Variabel Status Akhir dengan Variabel gejala sesak nafas

Dari gambar *scatter plot* di atas antara status akhir dengan gejala sesak nafas pasien dapat dilihat bahwa telah terjadi *separation case* karena ketika gejala sesak nafasnya adalah 0 sudah dipastikan nilai status akhirnya hanya ada di $Y = 1$, hal ini yang mengindikasikan adanya masalah *separation case* pada data.

Penaksiran Parameter Model Regresi Logistik Menggunakan Metode *Penalized Maximum Likelihood Estimation* (PMLE)

Model umum regresi logistik biner adalah:

$$\pi(x_i) = \frac{\exp(\beta_0 + \sum_{j=1}^p \beta_j x_{ij})}{1 + \exp(\beta_0 + \sum_{j=1}^p \beta_j x_{ij})} \tag{4}$$

Penaksiran parameter model regresi logistik biner menggunakan metode *penalized maximum likelihood estimation* (PMLE). Berikut merupakan hasil penaksiran parameter regresi logistik biner menggunakan *software R* disajikan pada Tabel 1.

Tabel 1. Nilai Taksiran Parameter Model Regresi Logistik Biner

Peubah Penjelas (X_j)	$\hat{\beta}_j$	$SE(\hat{\beta}_j)$	P-Value
(Intercept)	8.2514	1.4544	0.0000
Jenis Kelamin	0.3430	0.1690	0.0419
Umur	-0.0364	0.0061	0.0000
Komorbid	0.3608	0.1788	0.0418
Gejala Sesak Nafas	-5.9807	1.4130	0.0000

Berdasarkan hasil penaksiran pada Tabel 1 di atas, diperoleh persamaan untuk taksiran peluang pasien Covid-19 yang meninggal yaitu:

$$\hat{\pi}(x) = \frac{e^{8.2514+0.3430x_1-0.0364x_2+0.3608x_3-5.9807x_4}}{1 + e^{8.2514+0.3430x_1-0.0364x_2+0.3608x_3-5.9807x_4}} \tag{5}$$

Maka model taksiran logitnya yaitu:

$$\text{logit}[\hat{\pi}(x)] = 8.2514 + 0.3430x_1 - 0.0364x_2 + 0.3608x_3 - 5.9807x_4 \tag{6}$$

Berdasarkan model taksiran logit di atas maka dapat diinterpretasikan sebagai berikut, karena $\hat{\beta}_1 = 0.3430 > 0$ maka pasien dengan jenis kelamin laki-laki (X_1) memiliki efek positif, atau dengan kata lain peluang pasien Covid-19 yang meninggal akan meningkat dengan meningkatnya pasien yang berjenis kelamin laki-laki ketika variabel lain dalam keadaan konstan. Begitu pula dengan $\hat{\beta}_2 = -0.0364 < 0$ maka umur pasien (X_2) memiliki efek negatif, atau dengan kata lain peluang pasien Covid-19 yang meninggal akan meningkat dengan menurunnya umur pasien ketika variabel lain dalam keadaan konstan. Untuk $\hat{\beta}_3 = 0.3608 > 0$ maka pasien dengan adanya Komorbid (X_3) memiliki efek positif, atau dengan kata lain peluang

pasien Covid-19 yang meninggal akan meningkat dengan meningkatnya pasien yang memiliki gejala komorbid ketika variabel lain dalam keadaan konstan ketika variabel lain dalam keadaan konstan. Dan begitu pula dengan $\hat{\beta}_4 = -5.9807 < 0$ maka pasien dengan adanya gejala sesak nafas (X_4) memiliki efek negatif, atau dengan kata lain peluang pasien Covid-19 yang meninggal akan meningkat dengan menurunnya pasien yang terdapat gejala sesak nafas ketika variabel lain dalam keadaan konsta

Selanjutnya akan dilakukan pengujian signifikansi parameter model regresi logistik biner dengan metode PMLE secara serentak menggunakan uji G . Berikut merupakan hasil perhitungan statistik uji G , nilai kritis, nilai signifikansi, dan keputusan uji:

Tabel 2. Hasil Pengujian Parameter Model Regresi Logistik Secara Serentak

Uji G	$X^2_{(4;0.05)}$	P -value	Keputusan
285.991	9.4877	0.000	H_0 ditolak

Berdasarkan Tabel 2 di atas, didapatkan keputusan bahwa H_0 ditolak pada taraf nyata $\alpha = 5\%$, karena nilai statistik uji $G = 285.991 > X^2_{(4;0.05)} = 9.4877$ atau dapat juga dilihat dari nilai P -value $= 0.000 < \alpha = 0.05$. Sehingga dapat disimpulkan bahwa jenis kelamin, umur, penyakit bawaan dan gejala sesak nafas berpengaruh secara serentak terhadap kondisi akhir pasien Covid-19 di RSUD Kabupaten Pringsewu pada bulan Mei sampai November tahun 2021.

Selanjutnya dilakukan pengujian signifikansi parameter model regresi logistik biner dengan metode PMLE secara parsial menggunakan statistik uji $Wald$. Berikut merupakan hasil perhitungan statistik uji $Wald$, nilai kritis, nilai signifikansi, dan keputusan uji:

Tabel 3. Hasil Pengujian Parameter Model Regresi Logistik Secara Parsial

Peubah Penjelas (X_j)	$\hat{\beta}_j$	$SE(\hat{\beta}_j)$	$ W_j $	$Z_{0.05/2}$	P -Value	Keputusan
X_1	0.3430	0.1690	2.0296	1.96	0.0419	H_0 ditolak
X_2	-0.0364	0.0061	5.9672	1.96	0.0000	H_0 ditolak
X_3	0.3608	0.1788	2.0179	1.96	0.0418	H_0 ditolak
X_4	-5.9807	1.4130	4.2326	1.96	0.0000	H_0 ditolak

Didapatkan keputusan bahwa H_0 ditolak pada taraf nyata $\alpha = 5\%$, karena nilai mutlak statistik $Wald$ dari masing-masing variabel penjelas memiliki nilai lebih besar dari nilai $Z_{tabel} = 1.96$ atau dapat juga dilihat dari masing-masing nilai P -value lebih kecil dari $\alpha = 0.05$. sehingga dapat disimpulkan bahwa jenis kelamin, umur, penyakit bawaan dan gejala sesak nafas masing-masing berpengaruh secara parsial terhadap kondisi akhir pasien Covid-19 di RSUD Kabupaten Pringsewu pada bulan Mei sampai November tahun 2021.

Setelah dilakukan uji signifikansi parameter model, selanjutnya melakukan uji kecocokan model (*goodness of fit*) menggunakan metode *Hosmer Lemeshow*. Berikut merupakan hasil keputusan uji:

Tabel 4. Hasil Pengujian Parameter Model Regresi Logistik Secara Parsial

Uji Kecocokan Model			Keputusan
X -squared	df	P -value	
5.5596	8	0.6964	H_0 diterima

Berdasarkan nilai statistik uji \hat{C} pada Tabel 7 di atas, didapatkan keputusan bahwa H_0 diterima pada taraf nyata $\alpha = 0.05$, karena nilai statistik uji $\hat{C} = 5.5596 < X^2_{(0.05;df)} = 15.5073$ atau bias juga dilihat dari nilai P -value $= 0.6964 > \alpha = 0.05$. sehingga dapat disimpulkan bahwa Model regresi logistik dengan metode *panalized maximum likelihood estimation* (PMLE) cocok dengan data pasien Covid-19 di RSUD Kabupaten Pringsewu pada bulan Mei sampai November tahun 2021, atau dapat dikatakan bahwa model regresi logistik sudah layak untuk digunakan.

Selanjutnya karena model taksiran logit yang didapat sudah cocok atau layak digunakan, maka interpretasi hasil model dapat dilakukan menggunakan nilai *Odds Ratio (OR)*. Berikut hasil nilai *odds ratio* untuk penduga parameter model regresi logistik biner:

Tabel 5. Hasil Pengujian Parameter Model Regresi Logistik Secara Parsial

Peubah Penjelas	Penduga Parameter	Odds Ratio
X ₁	0.3430	1.4091
X ₂	-0.0364	0.9642
X ₃	0.3608	1.4345
X ₄	-5.9807	0.0025

Berdasarkan Tabel 5 dapat diinterpretasikan sebagai berikut:

(a) Jenis Kelamin Pasien Covid-19 ; Setiap kenaikan nilai X₁ (jenis kelamin) sebesar satu satuan akan menyebabkan perubahan odds atau peluang pasien Covid-19 yang meninggal secara multiplikatif sebesar $e^{\beta_j} = e^{0.3430} = 1.4091$ kali, dengan syarat pengaruh variabel lain konstan. Dengan kata lain, peluang pasien yang meninggal akibat Covid-19 di RSUD Kabupaten Pringsewu bagi yang berjenis kelamin laki-laki adalah 1.4091 kali dibandingkan pasien Covid-19 yang berjenis kelamin perempuan; (b) Umur Pasien Covid-19 Setiap kenaikan nilai X₂ (umur) sebesar satu satuan akan menyebabkan perubahan *odds* atau peluang pasien Covid-19 yang meninggal secara multiplikatif sebesar $e^{\beta_j} = e^{-0.0364} = 0.9642$ kali. Dengan kata lain, peluang pasien yang meninggal akibat Covid-19 di RSUD Kabupaten Pringsewu bagi yang berumur 1 tahun lebih tua adalah 0.9642 kali dibandingkan pasien Covid-19 lainnya. Karena umur pasien (X₂) memiliki efek negatif ($\hat{\beta}_2 = -0.0364$), atau dengan kata lain peluang pasien Covid-19 yang meninggal akan meningkat dengan menurunnya umur pasien ketika variabel lain dalam keadaan konstan.

(c) Penyakit bawaan (Komorbid) Pasien Covid-19 ; Setiap kenaikan nilai X₃ (penyakit bawaan) sebesar satu satuan akan menyebabkan perubahan *odds* atau peluang pasien Covid-19 yang meninggal secara multiplikatif sebesar $e^{\beta_j} = e^{0.3608} = 1.4345$ kali, dengan syarat pengaruh variabel lain konstan. Dengan kata lain, peluang pasien yang meninggal akibat Covid-19 di RSUD Kabupaten Pringsewu bagi pasien yang memiliki penyakit bawaan atau komorbid adalah 1.4345 kali dibandingkan pasien Covid-19 yang tidak memiliki penyakit bawaan atau komorbid; (d) Gejala sesak nafas Pasien Covid-19 : Setiap kenaikan nilai X₄ (gejala sesak nafas) sebesar satu satuan akan menyebabkan perubahan *odds* atau peluang pasien Covid-19 yang meninggal secara multiplikatif sebesar $e^{\beta_j} = e^{-5.9807} = 0.0025$ kali, dengan syarat pengaruh variabel lain konstan. Dengan kata lain, peluang pasien yang meninggal akibat Covid-19 di RSUD Kabupaten Pringsewu bagi pasien yang memiliki gejala sesak nafas adalah 0.0025 kali dibandingkan pasien Covid-19 yang tidak memiliki gejala sesak nafas. Karena pasien dengan adanya gejala sesak nafas (X₄) memiliki efek negatif ($\hat{\beta}_4 = -5.980$), atau dengan kata lain peluang pasien Covid-19 yang meninggal akan meningkat dengan menurunnya pasien yang terdapat gejala sesak nafas ketika variabel lain dalam keadaan konstan.

Untuk variabel umur pasien dan gejala sesak nafas dalam hasil interpretasi nilai peluang tidak sesuai dengan yang diharapkan, dimana pasien yang berumur tua maupun pasien sesak nafas memiliki peluang sukses lebih kecil kemungkinan terjadinya dibandingkan peluang gagal. Hal tersebut terjadi karena nilai penduga parameter model yang didapat untuk variabel umur dan gejala sesak nafas bernilai negatif maka mengakibatkan nilai *odds ratio* yang diperoleh bernilai < 1 dan hal tersebut bisa terjadi juga ketika terdapat rentang yang cukup jauh antara variabel respon dengan variabel penjelas (Dougherty, 2002). Sehingga untuk mengurangi hal tersebut untuk variabel umur pasien datanya bisa diubah dalam bentuk rentang yang sama seperti diubah dalam bentuk data kategori untuk lanjut usia dan bukan lanjut usia. Sedangkan untuk variabel gejala sesak nafas yang membuat nilai *odds ratio* menjadi kecil yaitu dalam datanya terdapat nilai 0 yang dapat dilihat pada Gambar 4. memperlihatkan data bernilai 0 yaitu ketika tidak ada gejala sesak nafas (X = 0) untuk nilai status akhirnya hanya ada di Y = 1. Atau dapat dilihat pada tabel kontingensi berikut dimana dalam sel terdapat nilai 0 untuk variabel gejala sesak nafas:

Tabel 6. Tabel Kontingensi antara Variabel Respon dengan Variabel Penjelas

Variabel Penjelas	Kategori	Status Akhir	
		Tidak Meninggal	Meninggal
Jenis Kelamin	0 = Perempuan	121	365
	1 = Laki-laki	114	417
Komorbid	0 = Tidak Ada	161	490
	1 = Ada	74	292
Gejala Sesak Nafas	0 = Tidak Ada	0	365
	1 = Ada	235	417

D. Kesimpulan

Berdasarkan rumusan masalah dan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa berdasarkan Penaksiran parameter model taksiran logitnya menggunakan metode PMLE pada kasus Covid-19 di RSUD Kabupaten Pringsewu pada bulan Mei sampai November tahun 2021 sebagai berikut:

$$\text{logit}[\hat{\pi}(x)] = 8.2514 + 0.3430x_1 - 0.0364x_2 + 0.3608x_3 - 5.9807x_4$$

Faktor yang mempengaruhi secara signifikan terhadap kematian pasien Covid-19 di RSUD Kabupaten Pringsewu pada bulan Mei sampai November tahun 2021 menggunakan regresi logistik biner adalah jenis kelamin, umur, penyakit bawaan (*komorbid*), dan gejala sesak nafas. Sehingga pasien Covid-19 berpeluang besar meninggal apabila berjenis kelamin laki-laki, memiliki umur yang lebih muda, memiliki penyakit bawaan, dan tidak memiliki gejala sesak nafas.

Daftar Pustaka

- [1] D. W. Hosmer Jr., S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression, Third Edition*. John Wiley & Sons, Inc., 2013.
- [2] Muhammad Bangkit Riksa Utama and N. Hajarisman, "Metode Pemilihan Variabel pada Model Regresi Poisson Menggunakan Metode Nordberg," *J. Ris. Stat.*, vol. 1, no. 1, pp. 35–42, Jul. 2021, doi: 10.29313/jrs.v1i1.24.
- [3] R. Kurniawan, *Analisis Regresi*. 2016.
- [4] F. A. Novianti and S. W. Purnami, "Analisis Diagnosis Pasien Kanker Payudara Menggunakan Regresi Logistik dan Support Vector Machine (SVM) Berdasarkan Hasil Mamografi," *J. SAINS dan Seni ITS*, vol. 1, no. 1, pp. D147–D152, 2012.
- [5] L. Muftadiah, "Estimasi parameter model regresi spasial error dengan metode maximum likelihood estimation," 2011, [Online]. Available: <http://etheses.uin-malang.ac.id/id/eprint/6860>
- [6] A. A. and A. J.A., "On the Existence of Maximum Likelihood Estimates in Logistic Regression Models," *Biometrika*, vol. 71, pp. 1–10, 1984, doi: <https://doi.org/10.1093/biomet/71.1.1>.
- [7] S. Devika, L. Jeyaseelan, and G. Sebastian, "Analysis of sparse data in logistic regression in medical research: A newer approach," *J. Postgrad. Med.*, vol. 62, no. 1, pp. 26–31, 2016, doi: 10.4103/0022-3859.173193.
- [8] F. D., "Bias Reduction of Maximum Likelihood Estimates," *Biometrics*, vol. 80, no. 1, pp. 27–38, 1993.
- [9] M. Sholihin *et al.*, "Implementasi Metode Penalized Maximum Likelihood Estimation Pada Model Regresi Logistik Biner 1 Pendahuluan," no. November, pp. 215–222, 2014.
- [10] E. Eyduran, "USAGE OF PENALIZED MAXIMUM LIKELIHOOD ESTIMATION METHOD IN MEDICAL RESEARCH: AN ALTERNATIVE TO MAXIMUM LIKELIHOOD ESTIMATION METHOD," *J. Res. Med. Sci.*, vol. 13, pp. 325–330, 2008.