

Penerapan Regresi *Complementary Log-Log* dalam Analisis Status Kejadian Tuberkulosis Paru Penduduk Usia Produktif di Provinsi Banten Tahun 2018

DESSY KURNIATI¹, BUDYANRA²

^{1,2)} Politeknik Statistika STIS, Indonesia
e-mail: 111910926@stis.ac.id

ABSTRAK

Regresi *complementary log-log* merupakan salah satu model GLM dengan fungsi hubung (*link function*) yang digunakan adalah *complementary log-log*. Metode ini merupakan salah satu metode yang dinilai tepat untuk digunakan ketika data pada variabel respon termasuk kedalam kategori jarang (*rare events*). Transformasi *complementary log-log* dinilai lebih tepat untuk digunakan karena kebanyakan observasi pada variabel respon akan berada pada kelompok kategori tertentu atau dapat dikatakan bahwa pada variabel respon tidak berdistribusi simetris. Salah satu data yang bersifat *rare events* adalah status kejadian tuberkulosis paru penduduk usia produktif di Provinsi Banten. Dimana berdasarkan data Riskesdas tahun 2018, persentase penduduk usia produktif di Provinsi Banten yang mengalami TB paru hanya sebesar 1,21 persen. Untuk itu, penelitian ini bertujuan untuk menerapkan regresi *complementary log-log* dalam menganalisis determinan status kejadian tuberkulosis paru penduduk usia produktif di Provinsi Banten tahun 2018. Hasil penelitian ini menunjukkan bahwa dengan regresi *complementary log-log* didapatkan hasil bahwa variabel yang signifikan memengaruhi status kejadian tuberkulosis paru yaitu tingkat pendidikan, status bekerja, status gizi, status gangguan mental emosional, status konsumsi buah dan sayur, serta status kelayakan air minum.

Kata Kunci: *complementary log-log, rare events, tuberkulosis paru, usia produktif.*

ABSTRACT

Complementary log-log regression is one of the GLM models where the link function used is the the *complementary log-log*. This method is one of the methods that is considered appropriate to be used when the data on the response variable is included in the rare event category. *Complementary log-log* transformation is considered more appropriate to use because most observations on the response variable will be in certain categories or it can be said that the response variable is not symmetrically distributed. One of the rare events data is the status of pulmonary tuberculosis in the productive age population in Banten Province. Where based on Riskesdas data in 2018, the percentage of the productive age population in Banten Province who experienced pulmonary TB was only 1.21 percent. Therefore, this study aims to apply *complementary log-log* regression in analyzing the determinants of the status of the status of pulmonary tuberculosis in the productive age population in Banten Province in 2018. The results of this study indicate that with *complementary log-log* regression, it is found that the variables that significantly affect the status of the incidence of pulmonary tuberculosis namely education level, work status, nutritional status, emotional mental disorder status, consumption status of fruit and vegetables, and drinking water eligibility status.

Keywords: *complementary log-log, rare events, pulmonary tuberculosis, productive age.*

1. PENDAHULUAN

Generalized Linear Models (GLM) atau Model Linier Tergeneralisasi (MLT) merupakan perluasan dari model linier klasik yang sering digunakan ketika data pada variabel respon “Y” berbentuk kategorik (McCullagh dan Nelder, 1989). Menurut McCullagh dan Nelder (1989), pada model GLM terdapat tiga komponen utama yaitu komponen acak (*random*) yang merupakan komponen pada variabel respon (Y) dan fungsi sebaran peluangnya yang termasuk keluarga eksponensial; komponen sistemik yaitu komponen yang mengacu pada kombinasi linier pada variabel prediktor (X) yang dilambangkan dengan η ; dan fungsi hubung (*link function*) yang menggambarkan hubungan antara pendugaan linier η dan nilai harapan variabel respon (Y). Pada model GLM komponen fungsi hubung diperlukan karena hubungan antara efek prediktor pada variabel respon mungkin tidak linier sehingga dengan mentransformasikan nilai hasil yang diharapkan maka model linier dapat digunakan (Allison, 2012).

Secara umum, terdapat tiga fungsi hubung (*link function*) yang biasa diterapkan saat analisis dengan model respon biner yaitu fungsi logit yang merupakan kebalikan fungsi sebaran kumulatif dari distribusi logistik; fungsi probit yang merupakan kebalikan fungsi sebaran kumulatif dari distribusi normal baku; dan fungsi *complementary log-log* yang merupakan kebalikan fungsi sebaran kumulatif dari distribusi Gompertz (Kismiantini, 2007). Pada umumnya, saat variabel respon bersifat kategori dan hanya memiliki dua kemungkinan (bersifat dikotomis) maka fungsi hubung yang sering digunakan yaitu fungsi hubung logit, sehingga sering juga disebut dengan regresi logistik biner (*binary logistic regression*). Hal ini karena fungsi hubung logit memiliki kelebihan dalam hal kemudahan interpretasi koefisien regresi yaitu interpretasi dapat dilakukan dengan melihat nilai *Odds ratio* (Kismiantini, 2007). Adapun metode estimasi parameter dalam regresi logistik biner yang sering digunakan adalah *maximum likelihood estimation* (MLE). Penggunaan metode *maximum likelihood* memiliki beberapa kelebihan yaitu berdistribusi normal pada sampel besar, konsisten secara asimtotik, konvergen terhadap parameter saat n bertambah, dan efisien secara asimtotik dibandingkan metode estimasi lainnya (Agresti, 2002).

Apabila pada sampel merupakan kategori yang jarang (*rare events*) yaitu ketika data pada variabel respon memiliki peluang sangat kecil (biasanya di bawah 5 persen), maka pemodelan menggunakan metode regresi logistik biner dengan fungsi hubung logit akan membuat hasil estimasi parameter melalui metode *maximum likelihood* menjadi bias (King & Zeng, 2001). Kitali et al., (2018) juga telah menjelaskan bahwa fungsi hubung logit dan probit sebaiknya hanya dapat diterapkan jika proporsi kategori sukses (50%) sama dengan proporsi kategori gagal (50%). Oleh karena itu, pada kondisi data *rare events* atau data bersifat asimetris maka diperlukan fungsi hubung alternatif yang dapat diterapkan pada *Generalized Linear Models* (GLM) yaitu dengan *complementary log-log*. Transformasi *complementary log-log* dinilai lebih tepat untuk digunakan karena kebanyakan observasi pada variabel respon akan berada pada kelompok kategori tertentu atau dapat dikatakan bahwa pada variabel respon tidak berdistribusi simetris (Heck et al., 2013). Dengan demikian, model yang asimetris yaitu model *complementary log-log* dinilai lebih tepat dalam menggambarkan hubungan antar variabel (Kitali et al., 2017; Martin dan Wu, 2017 di dalam Kitali et al., (2018)). Pemodelan dengan fungsi hubung *complementary log-log* ini biasa disebut juga dengan regresi *complementary log-log* (da Silva & Utomo, 2020).

Penerapan regresi *complementary log-log* dapat diterapkan dalam beberapa penelitian sosial, salah satunya yaitu tentang analisis status kejadian tuberkulosis paru. Tuberkulosis (TB) merupakan penyakit menular penyebab kesehatan yang buruk dan menjadi salah satu dari 10 penyebab utama kematian di seluruh dunia (WHO, 2020). Bahkan sebelum adanya pandemi COVID-19, penyakit TB masih menempati posisi pertama sebagai penyebab utama kematian akibat agen infeksi tunggal, di atas HIV/AIDS (WHO, 2021). Penyakit ini masih menjadi masalah dalam kesehatan masyarakat secara global, bahkan upaya penanggulangannya menjadi salah satu tujuan pembangunan berkelanjutan yaitu pada tujuan 3.3 (Kemenkes RI, 2021).

Di Indonesia sendiri, salah satu provinsi yang memiliki beban yang cukup tinggi terhadap kasus tuberkulosis (TB), terutama kasus TB Paru adalah Provinsi Banten. Berdasarkan Laporan Nasional Risesdas Tahun 2018, prevalensi TB Paru di Provinsi Banten adalah sebesar 0,76 persen. Nilai ini menempatkan Provinsi Banten pada urutan kedua setelah Provinsi Papua dengan nilai prevalensi TB Paru tertinggi di Indonesia dan menjadi urutan pertama di Pulau Jawa (Risesdas, 2018). Dimana berdasarkan kelompok umur, dari tahun 2015-2020 penemuan pasien baru TB Paru di Provinsi Banten didominasi oleh penduduk usia produktif (15 - 64

tahun). Tingginya kasus TB paru terutama pada usia produktif merupakan masalah mengingat pada usia produktif seseorang akan berada pada tahap untuk bekerja. Apabila orang tersebut menderita TB Paru, maka dapat mengakibatkan individu tersebut kehilangan produktivitasnya bahkan bisa menjadi beban keluarga (Nurjana, 2015). Kondisi ini tentu saja menjadi perhatian mengingat usia produktif sangat potensial untuk digunakan sebagai modal dasar dalam pembangunan (BPS Provinsi Banten, 2020).

Oleh karena itu, penelitian tentang status kejadian tuberkulosis paru pada usia produktif masih sangat diperlukan. Selain itu, nilai prevalensi TB paru yang cukup kecil yaitu di bawah 1 persen mengindikasikan bahwa diperlukan penyesuaian metode analisis yang tepat. Berdasarkan hal tersebut, pada penelitian ini bertujuan untuk menerapkan regresi *complementary log-log* dalam menganalisis determinan status kejadian tuberkulosis paru penduduk usia produktif di Provinsi Banten tahun 2018.

2. METODE PENELITIAN

Metode Pengumpulan Data

Penelitian ini menggunakan data sekunder berupa raw data yang bersumber dari hasil Survei Riset Kesehatan Dasar Nasional (Riskesdas) tahun 2018 yang berasal dari Balitbangkes yang dapat diakses dengan persyaratan dan prosedur tertentu melalui www.litbang.kemkes.go.id. Cakupan wilayah dalam penelitian yaitu Provinsi Banten dengan unit analisis dalam penelitian ini yaitu penduduk usia produktif kelompok usia 15 – 64 tahun. Jumlah responden yang memenuhi kriteria dan dianalisis dalam penelitian ini yaitu sebesar 6.129 jiwa.

Status kejadian tuberkulosis paru penduduk usia produktif merupakan variabel respon dalam penelitian ini. Sedangkan untuk variabel prediktor yang diduga signifikan berpengaruh pada penelitian ini terdiri dari tujuh variabel. Variabel prediktor dalam penelitian ini yaitu tingkat pendidikan, status bekerja, status gizi, status mengalami gangguan mental emosional (GME), status konsumsi buah dan sayur, status merokok, dan status kelayakan sumber air minum. Pada penelitian ini variabel bebas yang digunakan berdasarkan dari beberapa teori yang berkaitan dengan penyakit serta beberapa penelitian yang ada. Adapun beberapa teori yang digunakan sebagai acuan dalam menentukan variabel bebas yaitu teori Kesehatan HL. Blum dan teori Triad Epidemiologi.

Metode Analisis

Penelitian ini menggunakan dua metode analisis, yaitu analisis deskriptif dan analisis inferensia. Analisis deskriptif disajikan dalam bentuk tabel dan peta tematik. Sedangkan, analisis inferensia yang digunakan yaitu regresi *complementary log-log* dengan tujuan untuk mengetahui variabel yang signifikan memengaruhi status kejadian TB paru penduduk usia produktif di Provinsi Banten tahun 2018.

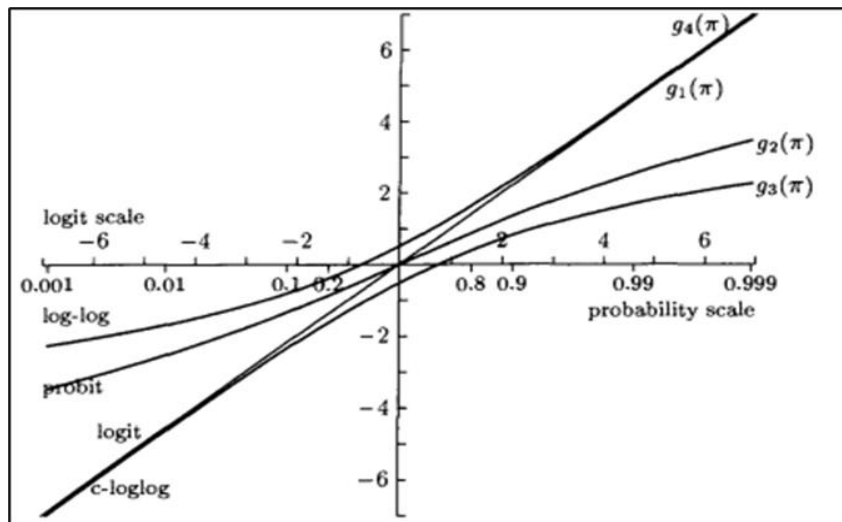
Penggunaan metode regresi *complementary log-log* dinilai lebih tepat dikarenakan kejadian tuberkulosis paru berada pada kasus *rare events*. Pada umumnya, transformasi *complementary log-log* sama dengan transformasi logit dan probit yaitu menggunakan interval 0 sampai 1 dan mengubahnya menjadi persamaan dengan interval $-\infty$ sampai ∞ . Adapun fungsi peluang dari *complementary log-log* sebagai berikut (Allison, 2012):

$$\pi(x) = 1 - \exp[-\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik})] \quad \dots (1)$$

dan model dari *complementary log-log* yaitu:

$$g(x) = \log[-\log(1 - \pi(x))] = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} \quad \dots (2)$$

Meskipun terdapat persamaan antara transformasi logit, probit dan *complementary log-log*. Namun, terdapat perbedaan utama antara transformasi *complementary log-log* dengan fungsi hubung logit dan probit, dimana fungsi hubung logit dan probit bersifat simetris pada sekitaran peluang 0,5, sedangkan transformasi *complementary log-log* bersifat asimetris. Perbedaan ini dapat dilihat pada gambar 1 sebagai berikut.



Gambar 1. Perbandingan Grafis dari Tiga *Link Function*

Pada Gambar 1, dapat dilihat fungsi hubung logit dan probit hampir berhubungan linier pada interval $0,1 \leq \pi \leq 0,9$ (Agresti,2002). Sedangkan pada fungsi hubung *complementary log-log*, kurva yang terbentuk tersebut lebih cepat mendekati 1 dibandingkan mendekati 0. Hal ini menunjukkan keadaan yang asimetris pada interval 0 sampai 1 dan meningkat secara perlahan pada nilai $\pi(x)$ kecil ke pertengahan kemudian meningkat tajam mendekati 1 (Allison,2012). Pada kondisi data bersifat *rare events* atau asimetris, fungsi hubung logit dan probit tidak lagi tepat dan efisien untuk digunakan, karena pada umumnya fungsi hubung logit dan probit memiliki kurva respon $\pi(x)$ yang berbentuk simetris di sekitar $\pi(x) = 0,5$. Oleh karena itu, fungsi hubung *complementary log-log* dapat menjadi pilihan terbaik untuk mendapatkan hasil yang lebih baik (Allison,2012).

Metode *maximum likelihood estimation* (MLE) adalah metode yang sering digunakan untuk mengestimasi parameter ketika data pada variabel respon hanya terdiri dari dua kategori (pada dasarnya bersifat dikotomis) (Deutsch et al., 2010). Berdasarkan hal tersebut, dalam mengestimasi parameter, pada regresi *complementary log-log* juga menggunakan metode *maximum likelihood estimation*.

Jika diasumsikan y_1, y_2, \dots, y_n merupakan nilai observasi dari variabel acak yang bersifat independen yaitu Y_1, Y_2, \dots, Y_n maka Y_i memiliki distribusi binomial dengan indeks m_i dan parameter $\pi(x_i)$. Adapun fungsi kepadatan peluang dari Y_i yang mengikuti distribusi binomial yaitu (McCullagh & Nelder,1989):

$$f(\pi(x_i); y_i) = \binom{m_i}{y_i} [\pi(x_i)]^{y_i} [1 - \pi(x_i)]^{m_i - y_i} \dots (3)$$

$\binom{m_i}{y_i}$ merupakan fungsi konstanta yang tidak terdapat $\pi(x_i)$ sehingga dapat diabaikan. Selain itu, data disusun berdasarkan individual dan bukan berdasarkan kelas *covariate* antara variabel independen sehingga $m_i = 1$ (McCullagh & Nelder,1989). Adapun bentuk fungsi likelihood sebagai berikut:

$$l(\pi(x_i); y_i) = \prod_{i=1}^n f(\pi(x_i); y_i) = \prod_{i=1}^n [\pi(x_i)]^{y_i} [1 - \pi(x_i)]^{1 - y_i} \dots (4)$$

Secara matematis, akan lebih mudah jika menggunakan log dari persamaan di atas, sehingga dirumuskan menjadi:

$$L(\pi(x_i); y_i) = \log[l(\pi(x_i); y_i)] = \sum_{i=1}^n \left[y_i \log \left(\frac{\pi(x_i)}{1 - \pi(x_i)} \right) + \log(1 - \pi(x_i)) \right] \dots (5)$$

Bagian sistemis dari model menentukan hubungan antara vektor $\pi(x_i)$ dengan kondisi observasi atau dengan kata lain bagian yang mengacu pada kombinasi linier pada variabel prediktor (X). Pada *Generalized Linear Models* (GLM), bentuk hubungan ini yaitu:

$$g(\pi(x_i)) = \eta_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad \dots (6)$$

Sehingga, log likelihood pada persamaan (5) dapat dinyatakan sebagai fungsi dari parameter yang tidak diketahui yaitu $\beta_1, \beta_2, \dots, \beta_p$. Pada kasus regresi *complementary log-log* kita memiliki fungsi hubung (*link function*) seperti pada persamaan (2). Jika fungsi hubung yang ada di substitusikan ke persamaan (5) maka didapatkan fungsi *log likelihood* $L(\beta)$ sebagai berikut.

$$\begin{aligned} &= \sum_{i=1}^n \left[y_i \log \left(\frac{1 - \exp[-\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)]}{\exp[-\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)]} \right) - \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p) \right] \\ &= \sum_{i=1}^n [y_i \log(\exp[\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)] - 1) - \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)] \end{aligned} \quad \dots (8)$$

Untuk mendapatkan nilai dari β yang memaksimalkan $L(\beta)$, dapat dilakukan turunan $L(\beta)$ terhadap $\beta_j (j = 0, 1, \dots, p)$ dan disama dengankan nol. Dalam prosesnya penurunan fungsi *log likelihood* ini tidak mudah. Sehingga, diperlukan proses iterasi dan telah diprogram pada perangkat lunak yang ada (Kismiantini, 2007). Nilai β yang dihasilkan merupakan estimasi *maximum likelihood* dan dapat dilambangkan dengan $\hat{\beta}$ (Agresti, 2002).

Metode estimasi parameter pada regresi *complementary log-log* dalam penelitian ini diolah menggunakan sintaks pada aplikasi R studio. Sintaks yang digunakan sudah mengakomodir *sampling weight* dari metode pengambilan sampel yang digunakan pada survei Riskesdas. Adapun tahapan pengujian pada regresi *complementary log-log* yaitu sebagai berikut.

1. Pembentukan model awal yang akan digunakan

Model regresi *complementary log-log* dalam penelitian yaitu:

$$g(D) = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 + \beta_4 D_4 + \beta_5 D_5 + \beta_6 D_6 + \beta_7 D_7 \quad \dots (9)$$

Keterangan:

D_1 : Variabel dummy tingkat pendidikan

D_2 : Variabel dummy status bekerja

D_3 : Variabel dummy status gizi

D_4 : Variabel dummy status gangguan mental emosional (GME)

D_5 : Variabel dummy status konsumsi buah dan sayur

D_6 : Variabel dummy status merokok

D_7 : Variabel dummy status kelayakan air minum

2. Melakukan pengujian parameter secara simultan menggunakan statistik uji *likelihood ratio test*. Dimana H_0 menyatakan tidak ada variabel bebas yang memengaruhi status kejadian TB paru. Adapun kriteria uji Tolak H_0 jika $G > \chi^2_{(\alpha; p)}$ atau $p\text{-value} < \alpha$.
3. Melakukan pengujian parameter secara parsial menggunakan statistik uji Wald. Dimana H_0 menyatakan variabel bebas ke-j tidak berpengaruh terhadap status kejadian TB paru. Adapun kriteria uji Tolak H_0 jika $|W| > |Z_{\alpha/2}|$ atau $p\text{-value} < \alpha/2$.
4. Melakukan pengujian kesesuaian model menggunakan statistik uji Hosmer-Lemeshow test. Dimana H_0 menyatakan bahwa model yang digunakan telah sesuai. Adapun kriteria uji Tolak H_0 yaitu ketika $\hat{C}_{hitung} > \chi^2_{(\alpha; (g-2))}$ atau $p\text{-value} < \alpha$.
5. Interpretasi hasil regresi *complementary log-log*. Fungsi peluang untuk model *complementary log-log* dapat diinterpretasikan pada formula berikut:

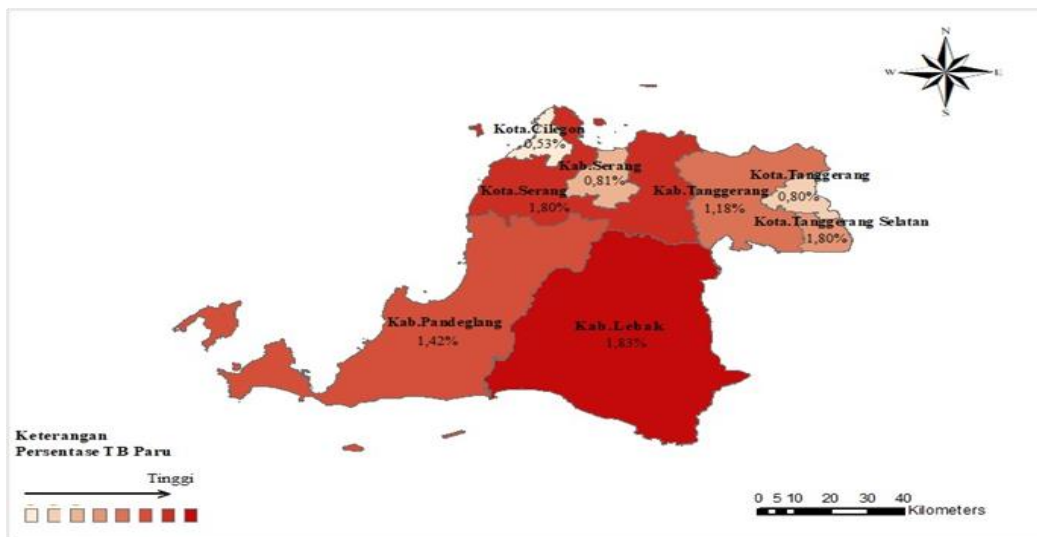
$$\frac{\log[1-\pi(x_j=1)]}{\log[1-\pi(x_j=0)]} \dots (10)$$

Dimana nilai $\exp(\hat{\beta}_j)$ memiliki arti besaran risiko dari variabel bebas yang memengaruhi status kejadian TB paru penduduk usia produktif di Provinsi Banten Tahun 2018.

3. HASIL DAN PEMBAHASAN

Gambaran Umum Kejadian Tuberkulosis (TB) Paru Penduduk Usia Produktif di Provinsi Banten tahun 2018

Berdasarkan hasil dari pengolahan data, didapatkan sebesar 1,21 persen penduduk usia produktif yang mengalami kejadian TB paru di provinsi tahun 2018. Kondisi ini mengindikasikan bahwa diperkirakan 1 sampai 2 dari 100 penduduk usia produktif di Provinsi Banten mengalami kejadian TB paru. Kemudian, persebaran persentase penduduk usia produktif yang mengalami kejadian TB paru dapat juga dilihat berdasarkan kabupaten/kota seperti pada gambar 2 berikut.



Gambar 2. Persentase Penduduk Usia Produktif (15-64 Tahun) yang Mengalami Kejadian TB Paru Menurut Kabupaten/Kota di Provinsi Banten Tahun 2018

Berdasarkan gambar 2 dapat dilihat bahwa persebaran persentase status kejadian TB paru pada penduduk usia produktif di Provinsi Banten tahun 2018 berbeda-beda menurut kabupaten/kota yang ada. Kabupaten Lebak merupakan kabupaten/kota dengan persentase status kejadian TB paru pada penduduk usia produktif tertinggi di Provinsi Banten yaitu sebesar 1,83 persen. Sedangkan, Kota Cilegon merupakan kabupaten/kota dengan persentase status kejadian TB paru pada penduduk usia produktif terendah di Provinsi Banten yaitu hanya sebesar 0,53 persen.

Jika status kejadian TB paru dikaitkan dengan variabel prediktor yaitu tingkat pendidikan, status bekerja, status gizi, status mengalami gangguan mental emosional (GME), status konsumsi buah dan sayur, status merokok, dan status kelayakan sumber air minum maka persentase kejadian TB paru pada penduduk usia produktif dapat dilihat pada tabel 1 sebagai berikut:

Tabel 1. Persentase Kejadian TB Paru Penduduk Usia Produktif Berdasarkan Variabel Prediktor

Variabel	Kategori	Status Kejadian TB paru	
		Mengalami	Tidak Mengalami
Tingkat pendidikan	\geq SMP	0,76%	99,24%
	< SMP	1,73%	98,27%
Status bekerja	Bekerja	1,04%	98,96%
	Tidak bekerja	1,68%	93,32%
Status gizi	Gizi baik/berlebih	0,92%	99,08%
	Gizi buruk	4,09%	95,91%
Status GME	Tidak mengalami GME	1,09%	98,91%
	Mengalami GME	2,07%	97,93%
Status konsumsi buah & sayur	Cukup	1,01%	98,99%
	Kurang	1,23%	98,77%
Status merokok	Tidak merokok	1,30%	98,70%
	Merokok	1,15%	98,85%
Status kelayakan sumber air minum	Layak	1,16%	98,84%
	Tidak layak	1,92%	98,08%

Berdasarkan Tabel 1 dapat diketahui bahwa pada saat penduduk usia produktif memiliki pendidikan yang ditamatkan kurang dari SMP, tidak bekerja, gizi kurang, mengalami gangguan mental emosional (GME), porsi konsumsi buah dan sayur harian kurang dan sumber air minumannya tidak layak menunjukkan persentase mengalami TB paru lebih tinggi dibandingkan pada kondisi penduduk usia produktif memiliki pendidikan yang ditamatkan minimal SMP, bekerja, memiliki gizi baik/berlebih, tidak mengalami gangguan mental emosional (GME), porsi konsumsi buah dan sayur harian cukup, dan sumber air minumannya layak. Sedangkan, pada saat penduduk usia produktif berstatus merokok memiliki persentase mengalami TB lebih rendah dibandingkan pada kondisi penduduk tidak merokok. Meskipun terlihat bahwa persentase yang tidak merokok cenderung lebih tinggi, akan tetapi nilai tersebut tidak berbeda jauh jika dibandingkan dengan persentase penduduk yang merokok, dimana hal ini juga dapat dilihat berdasarkan hasil regresi yang dilakukan bahwa variabel status merokok tidak signifikan memengaruhi status kejadian tuberkulosis paru.

Variabel-variabel yang Memengaruhi Status Kejadian Tuberkulosis (TB) Paru Penduduk Usia Produktif di Provinsi Banten tahun 2018.

Setelah membentuk model awal yang akan digunakan seperti pada persamaan (3), maka dilakukan pengujian parameter secara simultan. Berdasarkan hasil dari pengolahan data, diperoleh nilai statistik uji *likelihood ratio test* (G) sebesar 402,783. Nilai statistik uji G tersebut jauh lebih besar dari nilai *Chi-square* tabel $\chi^2_{(7,0,05)}$ yaitu 14,0671 sehingga keputusannya adalah tolak H_0 yang artinya minimal ada satu variabel bebas yang memengaruhi status kejadian TB paru. Selanjutnya, melakukan pengujian parameter secara parsial. Berdasarkan hasil pengolahan data, diperoleh nilai koefisien regresi dari semua variabel prediktor beserta nilai statistik uji *Wald* sebagaimana tercantum pada tabel 2.

Tabel 2. Hasil Uji Signifikansi Estimasi Parameter Secara Parsial

Kode Variabel	Nama Variabel	Kategori	$\hat{\beta}$	SE ($\hat{\beta}$)	Wald	p-value
Const	Konstanta	-	-5,258	0,022	-238,077	0,000
D_1	Tingkat Pendidikan	\geq SMP	-	-	-	-
		$<$ SMP	0,548	0,011	50,574	0,000
D_2	Status Bekerja	Bekerja	-	-	-	-
		Tidak Bekerja	0,619	0,011	54,124	0,000
D_3	Status Gizi	Gizi Baik/Berlebih	-	-	-	-
		Gizi Buruk	1,343	0,011	119,868	0,000
D_4	Status GME	Tidak Mengalami GME	-	-	-	-
		Mengalami GME	0,181	0,014	13,347	0,000
D_5	Status Konsumsi Buah dan Sayur	Cukup	-	-	-	-
		Kurang	0,130	0,020	6,549	0,000
D_6	Status Merokok	Tidak Merokok	-	-	-	-
		Merokok	-0,016	0,011	-1,375	0,169
D_7	Status Kelayakan Air Minum	Layak	-	-	-	-
		Tidak layak	0,048	0,020	2,388	0,017

Jika nilai mutlak statistik uji *Wald* dibandingkan dengan nilai Z tabel $|Z_{(\frac{\alpha}{2})}|$ yaitu 1,96 maka keputusan untuk variabel tingkat pendidikan, status bekerja, status gizi, status gangguan mental emosional (GME), status konsumsi buah dan sayur, dan status kelayakan air minum yaitu tolak H_0 karena $|W_i| > 1,96$. Sedangkan keputusan untuk variabel merokok yaitu gagal tolak H_0 karena $1,375 < 1,96$. Kesimpulan dari pengujian secara parsial yaitu dari tujuh variabel bebas, hanya terdapat enam variabel yang signifikan memengaruhi status kejadian TB paru. Variabel bebas yang signifikan memengaruhi status kejadian TB paru yaitu tingkat pendidikan, status bekerja, status gizi, status gangguan mental emosional (GME), status konsumsi buah dan sayur, serta status kelayakan air minum. Sedangkan, variabel status merokok tidak berpengaruh signifikan terhadap status kejadian TB paru.

Sebelum model digunakan harus diuji apakah model yang digunakan sudah sesuai. Hasil dari pengolahan data didapatkan nilai statistik uji *Hosmer-Lemeshow test* (\hat{C}) yaitu 7,0838. Karena nilai \hat{C} lebih kecil dibandingkan nilai *Chi-square* tabel $\chi^2_{(8;0,05)}$ yaitu 15,507 maka keputusan yang diambil adalah gagal tolak H_0 , sehingga dapat disimpulkan bahwa model telah sesuai atau dengan kata lain tidak terdapat perbedaan yang signifikan antara hasil prediksi model dengan hasil pengamatan.

Berdasarkan pengujian pendugaan parameter serta telah disimpulkan bahwa model yang digunakan telah sesuai, maka bentuk persamaan regresi *complementary log-log* yang diperoleh sebagai berikut:

$$\hat{g}(D) = -5,258 + 0,548D_1^* + 0,619D_2^* + 1,343D_3^* + 0,181D_4^* + 0,130D_5^* - 0,016D_6 + 0,048D_7^* \dots \text{ (11)}$$

Keterangan: *) signifikan pada alfa 0,05

Nilai Risiko Variabel yang Memengaruhi Status Kejadian Tuberkulosis Paru Penduduk Usia Produktif di Provinsi Banten Tahun 2018

Berdasarkan persamaan yang telah didapatkan pada persamaan (11), maka dapat diketahui nilai risiko dari variabel bebas yang signifikan memengaruhi status kejadian TB paru yang ditunjukkan dengan $\exp(\hat{\beta})$ pada tabel 3 sebagai berikut.

Tabel 3. Nilai Risiko Hasil Regresi *Complementary log-log*.

Kode Variabel	Nama Variabel	Kategori	$\exp(\hat{\beta})$
D_1	Tingkat Pendidikan	\geq SMP	-
		$<$ SMP	1,7299
D_2	Status Bekerja	Bekerja	-
		Tidak Bekerja	1,8564
D_3	Status Gizi	Gizi Baik/Berlebih	-
		Gizi Buruk	3,8319
D_4	Status Gangguan Mental Emosional (GME)	Tidak Mengalami GME	-
		Mengalami GME	1,1981
D_5	Status Konsumsi Buah dan Sayur	Cukup	-
		Kurang	1,1391
D_7	Status Kelayakan Air Minum	Layak	-
		Tidak Layak	1,0492

Tingkat Pendidikan

Nilai $\exp(\hat{\beta})$ untuk variabel tingkat pendidikan sebesar 1,7299. Artinya, penduduk dengan pendidikan tertinggi yang ditamatkan di bawah SMP memiliki risiko lebih besar yaitu 1,7299 untuk mengalami kejadian TB paru dibandingkan penduduk dengan pendidikan tertinggi yang ditamatkan minimal SMP dengan asumsi variabel bebas yang lain dianggap konstan. Hal ini dapat dilihat juga dari nilai koefisien $\hat{\beta}$ variabel tingkat pendidikan yang bertanda positif. Hasil ini juga sejalan dengan penelitian oleh Pangaribuan et al. (2020) yang menemukan bahwa responden yang berpendidikan di bawah SMP memiliki risiko lebih besar hampir 1,5 kali untuk mengalami TB paru dibandingkan yang berpendidikan minimal SMP. Hal ini karena seseorang yang berpendidikan tinggi akan cenderung memiliki wawasan yang lebih luas sehingga akan mendorong untuk dapat berperilaku hidup sehat.

Status Bekerja

Nilai $\exp(\hat{\beta})$ untuk variabel status bekerja sebesar 1,8564. Artinya, penduduk yang tidak bekerja memiliki risiko lebih besar yaitu 1,8564 untuk mengalami kejadian TB paru dibandingkan penduduk yang bekerja dengan asumsi variabel bebas yang lain dianggap konstan. Hal ini dapat dilihat juga dari nilai koefisien $\hat{\beta}$ variabel status bekerja yang bertanda positif. Hasil ini juga sejalan dengan penelitian oleh Simbolon et al. (2019) yang menyatakan bahwa seseorang yang tidak bekerja memiliki resiko 3,4 kali lebih tinggi untuk terkena TB paru dibandingkan orang yang bekerja. Hal ini karena orang yang bekerja akan memiliki pendapatan untuk memenuhi kebutuhannya serta akan lebih jarang dirumah sehingga dapat menghindari penularan apabila tinggal serumah dengan penderita TB.

Status Gizi

Nilai $\exp(\hat{\beta})$ untuk variabel status gizi sebesar 3,8319. Artinya, penduduk dengan kondisi gizi buruk memiliki risiko lebih besar yaitu 3,8319 untuk mengalami kejadian TB paru dibandingkan

penduduk dengan kondisi gizi baik/berlebih dengan asumsi variabel bebas yang lain dianggap konstan. Hal ini dapat dilihat juga dari nilai koefisien $\hat{\beta}$ variabel status gizi yang bertanda positif. Ellyse S, (2021) dalam penelitiannya juga menemukan bahwa orang yang mengalami malnutrisi ($< 18,5 \text{ kg/m}^2$) beresiko lebih besar 6,245 kali untuk mengalami kejadian TB paru dibandingkan yang tidak mengalami malnutrisi ($\text{IMT} \geq 18,5$). Hal ini karena kekurangan gizi dapat membuat daya tahan tubuh menjadi lemah.

Status Gangguan Mental Emosional (GME)

Nilai $\exp(\hat{\beta})$ untuk variabel status gangguan mental emosional sebesar 1,1981. Artinya, penduduk yang mengalami GME memiliki risiko lebih besar yaitu 1,1981 untuk mengalami kejadian TB paru dibandingkan penduduk yang tidak mengalami GME dengan asumsi variabel bebas yang lain dianggap konstan. Hal ini dapat dilihat juga dari nilai koefisien $\hat{\beta}$ variabel status gangguan mental emosional yang bertanda positif. Hal ini sejalan dengan pernyataan bahwa gangguan mental emosional akan dapat menyebabkan seseorang mengalami masalah dalam kesehatan (Nurjanah, 2020). Hal ini karena gangguan mental emosional dapat mengganggu hubungan individu dengan lingkungan yang menyebabkan penurunan kekebalan tubuh.

Status Konsumsi Buah & Sayur

Nilai $\exp(\hat{\beta})$ untuk variabel status konsumsi buah dan sayur sebesar 1,1391. Artinya, penduduk yang porsi konsumsi buah dan/atau sayur perharinya kurang memiliki risiko lebih besar yaitu 1,1391 untuk mengalami kejadian TB paru dibandingkan penduduk yang porsi konsumsi buah dan/atau sayur perharinya cukup dengan asumsi variabel bebas yang lain dianggap konstan. Hal ini dapat dilihat juga dari nilai koefisien $\hat{\beta}$ variabel status konsumsi buah dan sayur yang bertanda positif. Hal ini juga sejalan dengan penelitian yang dilakukan oleh Setyawati & Putri (2008) di dalam penelitiannya menemukan bahwa responden yang tidak mengkonsumsi buah dan sayur dalam keadaan cukup memiliki resiko 1,2 kali lebih besar untuk mengalami TB paru dibandingkan responden yang konsumsi buah dan sayur dalam keadaan cukup. Hal ini karena didalam buah dan sayur terdapat senyawa antioksidan yang dapat membantu menjaga daya tahan tubuh.

Status Kelayakan Air Minum

Nilai $\exp(\hat{\beta})$ untuk variabel status kelayakan air minum sebesar 1,0492. Artinya, penduduk yang tinggal pada rumah tangga dengan kondisi air minum tidak layak memiliki risiko lebih besar yaitu 1,0492 untuk mengalami kejadian TB paru dibandingkan penduduk yang tinggal pada rumah tangga dengan kondisi air minum layak dan dengan asumsi variabel bebas yang lain dianggap konstan. Hal ini dapat dilihat juga dari nilai koefisien $\hat{\beta}$ variabel status kelayakan air minum yang bertanda positif. Hal ini sejalan dengan penelitian Girsang et al. (2011) yang menjelaskan bahwa salah satu faktor penyebab penyakit tuberkulosis yaitu konsumsi air minum yang tidak memenuhi syarat. Dimana dia menjelaskan bahwa anggota rumah tangga yang apabila terus menerus mengkonsumsi air minum yang tidak memenuhi syarat memiliki kemungkinan 1,8 kali untuk terpapar penyakit menular. Hal ini karena konsumsi air minum yang layak sangat dibutuhkan tubuh dan dapat menjaga daya tahan tubuh.

4. SIMPULAN DAN SARAN

Dengan menggunakan metode *complementary log-log* didapatkan kesimpulan bahwa variabel prediktor yang signifikan memengaruhi status kejadian tuberkulosis paru pada penduduk usia produktif di Provinsi Banten yaitu tingkat pendidikan, status bekerja, status gizi, status gangguan mental emosional, status konsumsi buah dan sayur, serta status kelayakan air minum. Dimana Seseorang yang berada pada kelompok usia produktif memiliki ukuran risiko lebih besar untuk mengalami kejadian TB paru apabila memiliki karakteristik pendidikan tertinggi yang ditamatkan di bawah SMP, tidak bekerja, keadaan gizi buruk, mengalami gangguan mental emosional, porsi konsumsi buah dan/atau sayuran kurang, dan sumber air minum yang dikonsumsi tidak layak.

Saran dalam penelitian ini yaitu bagi penelitian selanjutnya dapat melakukan perbandingan model dan memilih model terbaik dari beberapa pemodelan yang sering digunakan untuk data kategori jarang (*rare events*).

DAFTAR PUSTAKA

- Agresti, A. (2002). *Categorical Data Analysis* (2nd ed). New York: John Wiley & Sons.
- Allison, P.D. (2012). *Logistic Regression Using the SAS System : Theory & Application*. Cary, NC: SAS Institute Inc.
- BPS. (2020). *Indikator kesejahteraan Rakyat Provinsi Banten*. Serang : BPS Provinsi Banten
- Balitbangkes. (2018). *Laporan Nasional Riskesdas 2018*. Jakarta: Kementerian Kesehatan RI
- Da Silva, M. M., & Utomo, A. P. (2020). Penerapan Regresi *Complementary log-log* Dalam Analisis Kemiskinan Anak Di Daerah Perkotaan Di Provinsi Bengkulu Tahun 2018. *Seminar Nasional Official Statistics*, 2020(1), 825–834.
- Ellyse S, B. (2021). Pengaruh Kepadatan Hunian, Malnutrisi, Umur dan Jenis Kelamin terhadap Kejadian Tuberkulosis Paru (Studi Observasional di Wilayah Kerja Puskesmas Batulicin I Kecamatan Karang Bintang Kabupaten Tanah Bumbu Provinsi Kalimantan Selatan Tahun 2019). *Universitas Islam Kalimantan MAB*.
- Girsang, M., Tobing, K., & Rafrizal, R. (2011). Faktor Penyebab Kejadian Tuberculosis Serta Hubungannya dengan Lingkungan Tempat Tinggal di Provinsi Jawa Tengah (Analisis Lanjut Riskesdas 2007). *Indonesian Bulletin of Health Research*, 39(1), 34–41.
- Hosmer, D.W., Lemeshow, S., & Sturdivant, R. X. (2000). *Applied Logistic Regression Second Edition*. New Jersey: John Wiley & Sons Inc
- Kemendes RI. (2021). *Profil Kesehatan Indonesia 2020*. Jakarta: Kementerian Kesehatan RI
- King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political Analysis*, 9(2), 137–163.
- Kitali, A. E., Alluri, P., Sando, T., Haule, H., Kidando, E., & Lentz, R. (2018). Likelihood estimation of secondary crashes using Bayesian *complementary log-log* model. *Accident Analysis & Prevention*, 119, 58–67.
- Kismiantini, K. (2007). Analisis Peubah Respon Biner. *Pythagoras: Jurnal Pendidikan Matematika*, 3(1).
- Localio, A. R., Margolis, D. J., & Berlin, J. A. (2007). Relative risks and confidence intervals were easily computed indirectly from multivariable logistic regression. *Journal of Clinical Epidemiology*, 60(9), 874–882.
- Nurjana, M. A. (2015). Faktor risiko terjadinya Tuberculosis paru usia produktif (15-49 tahun) di Indonesia. *Media Penelitian Dan Pengembangan Kesehatan*, 25(3), 20736.
- Nurjanah, S. (2020). Gangguan mental emosional pada klien pandemi covid 19 di rumah karantina. *Journal Ilmu Keperawatan Jiwa*, 3(3), 329–334.
- McCullagh, P., & J.A. Nelder (1989). *Generalized Linear Models*. Second Edition. New York: Chapman & Hall.
- Pangaribuan, L., Kristina, K., Perwitasari, D., Tejayanti, T., & Lolong, D. B. (2020). Faktor-Faktor yang memengaruhi kejadian tuberkulosis pada umur 15 tahun ke atas di Indonesia. *Buletin Penelitian Sistem Kesehatan*, 23(1), 10–17.
- Setyawati, B., & Putri, D. S. K. (2008). Hubungan Pola Konsumsi Bahan Makanan Sumber Antioksidan Dengan Risiko Tuberculosis (TB) Paru. Diakses pada Tanggal 15 Maret 2022 melalui <http://repository.litbang.kemkes.go.id/2247/>
- Simbolon, D. R., Mutiara, E., & Lubis, R. (2019). Analisis spasial dan faktor risiko tuberkulosis paru di Kecamatan Sidikalang, Kabupaten Dairi-Sumatera Utara tahun 2018. *Berita Kedokteran Masyarakat*, 35(2), 65–71.
- Susanti, D. S., Sukmawaty, Y., & SALAM, N. U. R. (2019). Analisis Regresi dan Korelasi. *IRDH*.
- World health Organization. (2020). *Global Tuberculosis Report 2020*. Geneva: World Health Organization
- World health Organization. (2021). *Global Tuberculosis Report 2021*. Geneva: World Health Organization